مجلة العلوم الهندسية والتقنية

مجلة علمية نصف سنوية محكمة

تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية- جامعة ذمار

المجلد: 3 ، **العدد** : 1

مجلة العلوم
الهندسية والتقنية

مجلة علمية محكمة تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية - جامعة ذمار

مجلة العلوم الهندسية والتقنية JOEATS - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية بجامعة ذمار، وتنشر مقالات علمية في الهندسة والحاسبات ونظم المعلومات، والهندسة المعمارية والمدنية والكهربائية والميكانيكية والاتصالات والتخطيط العمراني والتصميم والميكاترونكس والكمبيوتروتكنولوجيا المعلومات والبيئة والطاقة المتجددة وغيرها من المجالات الهندسية.

الهيئة الاستشارية

سكرتير التحرير

مجلة العلوم الهندسية والتقنية

مجلة علمية نصف سنوية محكمة
تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية
جامعة ذمار، ذمار، الجمهورية اليمنية.

المجلد (3)

العدد (1)

يونيو 2024

ISSN: 2958-809X
EISSN: 2958-8103

شروط النشر

أولا: خطوات النشر في المجلة

تسير عملية النشر في المجلة وفق الخطوات الآتية:

1. أن تتسم بالأصالة والمنهجية العلمية السليمة.
2. أن لا تكون قد سبق نشرها أو تقديمها للنشر إلى جهة أخرى، ويقدم الباحث إقرارا خطيًّا بذلك.
3. تكتب البحوث بلغة سليمة، وتراعى فيها قواعد الضبط ودقة الأشكال -إن وجدت- بصيغة (Word).
4. تسلم نسخة من البحوث المطلوب نشرها إلى مكتب مجلة العلوم الهندسية والتقنية - جامعة ذمار. أو ترسل إلى رئيس تحرير المجلة على الإيميل: joeats-tu@tu.edu.ye
2. بعد استلام البحث سيتم تأكيد ذلك بإرسال الإيميل للكاتب الأول.
3. تتم المراجعة الأولية لجودة البحث لمعرفة مدى تطابقه مع متطلبات النشر والجودة في المجلة.
4. عند قبول البحث أوليا للنشر في المجلة يتم إرسال طلب سداد رسوم النشر بواسطة الإيميل.
4. بعد استلام رسوم النشر غير القابلة للاسترداد (نقدًا- أو تحويل بنكي) تبدأ مرحلة تقييم البحث وتحكيمه.
5. يرسل البحث للتقييم والتحكيم من قبل (اثنين محكمين).
6. بحسب قرار المحكمين يتخذ رئيس التحرير القرار، ويتم إرساله إلى الكاتب - الكُتاب. والقرار المتخذ له ثلاثة حالات هي:

- قبول النشر بدون تعديل - أو قبول النشر مع تعديلات ثانوية - أساسية (يطلب تعديلها) أو يرفض نشر البحث.

- بعد قيام الكاتب - أو الكُتاب بإجراء التعديلات المطلوبة وإرسالها إلى المحكمين والموافقة عليها من قبل المحكمين وقبول نشر البحث، يرسل إلى الكاتب قرار قبول النشر ومرفق معه استمارة حقوق النشر، ليقوم الكاتب بالتوقيع عليها وإعادة إرسالها بالإيميل إلى رئيس التحرير خلال 6 أيام.

ثانيا: رسوم النشر

- يطالب الكاتب والكُتاب بتسديد المبالغ المستحقة غير القابلة للاسترداد (ينظر الجدول أدناه) باعتبارها رسوم نشر البحوث ولا يتم ذلك إلا بعد إبلاغ رئيس التحرير للكاتب أو الكُتاب أن البحث مناسب ويلتزم بمتطلبات المجلة، ولا يتم استكمال إجراءات تحكيم البحث وإرساله إلى المحكمين إلا بعد استلام رسوم النشر (نقدًا -تحويلا بنكيا).

صفة الباحث الأول	الرسوم
أعضاء هيئة التدريس/ جامعة ذمار	20000 ريال يمني.
أعضاء هيئة التدريس/ الجامعات والهيئات والمنظمات اليمنية	30000 ريال يمني.
الباحثين الأجانب من الدول والمنظمات الخارجية	150دولار أمريكي.

ثالثا: دليل تقديم الأبحاث

- تنشر المجلة البحوث العلمية في مجالات الهندسة المعمارية، المدنية، الكهربائية، الميكانيكية، الاتصالات، التخطيط والتصميم العمراني، الميكاترونكس، الحاسوب وتقنية المعلومات، البيئة والطاقة المتجددة وغيرها من المجالات الهندسية بإحدى اللغتين (العربية أو الإنجليزية)، وتسلم البحوث إلكترونيًّا بصيغة وورد (MS WORD)، ويتطلب توفير المتطلبات الضرورية للنشر على النحو الآتي:
- **مكونات البحث**
- يجب أن يحتوي البحث على الآتي:

- الصفحة الأولى (عنوان البحث - اسم الكاتب- الكُتاب، وعناوينهم، وملخص البحث (أقل من 200 كلمة) تحتوي على هدف البحث، منهجيته، أبرز النتائج التي توصل إليها، خمس كلمات مفتاحيه.
- تحتوي الصفحة الثانية على ترجمة لما جاء في الصفحة الأولى.
- ويبدأ من الصفحة الثالثة: (محتويات البحث التي تشمل: المقدمة والمشكلة البحثية وأهداف البحث - وخلفية البحث والمنهجية المتبعة، والتحليل، والنتائج، والتوصيات إن وجدت، والمراجع شاملة الجداول والاشكال والمخططات والصور التوضيحية)، ويجب ألا يزيد البحث بكامل محتوياته عن (8000) كلمة.

رابعًا: مراحل التقييم والتحكيم

كل بحث يتم تقييمه أوليًا بواسطة هيئة التحرير، ومن ثم ترسل الأبحاث المطابقة للمعايير إلى (اثنين محكمين مستقلين) على الأقل، وبناء على توصية المحكمين، وبالتشاور مع هيئة التحرير، يقرر رئيس التحرير إحدى القرارات الثلاثة الآتية:

قبول النشر بدون تعديل، أو قبول النشر مع تعديلات ثانوية - أساسية (يطلب تعديلها)، أو يرفض نشر البحث.

- **ترسل الابحاث** بواسطة الإيميل الآتي: joeats-tu@tu.edu.ye

بعد استلام البحث سيتم إبلاغ الباحث باستلام بحثه عبر الإيميل، وبعد سداده رسوم النشر ستبدأ مرحلة تحكيم البحث بوساطة محكمين مستقلين من جنسيات مختلفة.

مجلــة العلــوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

Using Machine Learning for Arabic Sentiment Analysis in Higher Education: Investigating the Impact of Utilizing the ChatGPT and Bard Google

Salah AL-Hagree[1,2], Ghaleb Al-Gaphari[1], Fuaad Hasan Abdulrazzak[3], Maher Al-Sanabani[3], Ahmed Al-Shalabi[1]

[1] Computer Science Department, Faculty of Computer and Information Technology, Sana'a University, Yemen
[2] Computer Science Department, Faculty of Sciences, Ibb University, Yemen.
[3] Computer Science Department, Faculty of Computer and Information Technology, Thamar University, Yemen.
s.alhagree@gmail.com, s.alhagree@su.edu.ye, drghalebh@su.edu.ye, fuaad.abdulrazzak@tu.edu.ye, M.sanabani@gmail.com , a.alshalabi@su.edu.ye

Abstract.

Gaining an understanding of the application's quality and meeting, the user's needs are crucial in the development of applications. To enhance the quality of applications, it is important to comprehend the requirements of the users. One effective approach for achieving this is through the utilization of application review-based sentiment analysis (SA). In this study, the objective was to assess students' opinions regarding mobile applications of universities in order to update and maintain them accordingly. Mobile applications of universities have become an integral part of students' lives, thus making it imperative to analyze user comments on these apps for SA purposes, where student input is crucial for assessing the effectiveness of educational institutions. This paper presents a machine learning (ML) based approach to sentiment analysis on students' evaluation of higher education institutions. The study analyzes a corpus containing approximately 275 student reviews written in Arabic, It also evaluates the performance of three ML techniques, including K-Nearest Neighbors (K-NN), Decision Tree (DT), and Random Forest (RF) using an accuracy, precision, recall, and f-score measures. In addition, the study compares one method of labeling the data for ASA, including manual labeling by humans, labeling by Bard Google and labeling by ChatGPT. Experimental results show that the K-NN technique performed the best, achieving an accuracy of 74.91% by ChatGPT models for Arabic sentiment analysis (ASA). Moreover, utilizing proposed active labeling method with Bard Google achieved higher accuracy compared to other labeling methods. The study proposed study suggests that the K-NN technique with ChatGPT models and proposed active labeling method are effective approaches for ASA by ChatGPT. It is indicated by the empirical results that promising results are yielded on the evaluation of students' opinions of higher educational institutions by an ML based approach.

Keywords: *Arabic Sentiment Analysis, Higher Educational, Bard Google, ChatGPT, Machine learning.*

I. Introduction

Over the past few years, there has been a growing fascination with sentiment analysis (SA) within the field of text mining. It has gained popularity as a prominent research area in higher education, specifically in the realm of opinion mining. This field focuses on analyzing and comprehending students' opinions regarding their educational institutions, aiming to enhance the decision-making process by improving its overall quality. This study aimed to investigate the utilization of SA in higher education by conducting a comprehensive analysis. It sought to identify and categorize the commonly employed and effective SA techniques and methods within the higher education domain. SA involves employing natural language processing techniques, text analysis, and statistical methods to analyze subjective information, such as opinions, attitudes, impressions, and emotions. Its purpose is to extract and categorize students' opinions and emotions, which are crucial in various aspects [1]. This approach finds wide application in processing, searching, and extracting factual information from diverse platforms such as blogs, Google Play , Tumblr, Instagram, Twitter, Facebook, and more [2]. As students form a significant part of universities, their perspectives and opinions play a vital role in enhancing teaching and addressing institutional challenges and concerns. Therefore, evaluating student opinions in digital educational resources and learning environments becomes essential for assessing the institution, teachers, and teaching effectiveness. Smartphone applications have become increasingly integral to our daily lives, experiencing a surge in usage. The Google Play Store is a widely recognized platform that provides access to a diverse range of Android applications. Among these applications, those related to higher education play a crucial role in enhancing the delivery of services to students with greater efficiency and effectiveness [3]. In SA tasks, text records are categorized into three classes: positive, negative, or neutral, representing different levels of text polarity [4]. There are two primary approaches to conducting SA: lexicon-based sentiment analysis (LBSA) and ML Arabic sentiment analysis (MLSA). LBSA utilizes a vocabulary dictionary to calculate the polarity of each text record, while MLSA relies on ML models to predict the polarity of text records. Although MLSA is more efficient, it requires human-annotated data for training on polarity prior to the prediction process [4]. SA encounters a challenge when it comes to capturing the experiences of university students in higher education. Presently, numerous companies have emerged with the advent of artificial intelligence, offering a range of tools designed to assist students in their higher education journey. Among the most renowned tools in this domain are those powered by artificial intelligence, such as

Lectomate[1], Qonqur[2], Quetab [3], Dunno [4], Conker [5], Tutor AI [6], Resoomer[7], Lectomate [8], Gistvid [9] [10]. The Arabic language holds significant popularity and widespread usage, making it crucial to employ SA tools for Arabic adoption. However, the complexity of the Arabic language, including its morphology, structure, and variations, poses challenges. Further efforts are needed to improve Arabic language SA [1]. In this study, a specific Arabic dataset was manually collected and intentionally annotated SA tasks. The K-Nearest Neighbors (K-NN), Decision Tree (DT), and Random Forest (RF) algorithms were utilized to conduct ASA. The study focused on the challenges faced in SA within higher education institutions, with a specific focus on mobile applications in Yemeni universities. The models were evaluated using accuracy metrics for ML evaluation, and a comparison among the three models revealed the superiority of the K-NN algorithm with ChatGPT.

The structure of this paper is as follows. Section 2 provides an overview of recent studies on SA that utilize app feedback data from Google Play, with particular emphasis on the Arabic language. Section 3 presents the proposed models and methods, including a brief description of the dataset and the preprocessing techniques employed. The findings and discussions of the experiments are presented in Section 4. Lastly, Section 5 presents the conclusions of the study.

II. LITERATURE REVIEW

To identify the research gap in SA, we conducted a thorough review of important and relevant studies. SA is the process of extracting patterns from textual data, which can include categorizing and interpreting sentiment into negative, positive, or neutral comments using techniques such as ML. With the increasing availability of user information on the web, including social networks and other platforms, SA has become an important tool for understanding user opinions and behaviors. Many studies have focused on developing solutions for SA based on ChatGPT and Bard Google. Additionally, reviews on SA are often based on various

[1] https://topai.tools/t/bloombot-ai
[2] https://topai.tools/t/qonqur /
[3] https://topai.tools/t/quetab
[4] https://topai.tools/t/Dunno
[5] https://topai.tools/t/conker
[6] https://topai.tools/t/tutor-ai
[7] https://topai.tools/t/resoomer-com
[8] https://topai.tools/t/lectomate
[9] https://topai.tools/t/Gistvid

platforms, such as Google Play. By exploring these studies and platforms, we aim to contribute to the field of ASA and provide a more effective and accurate approach to analyzing sentiment in Arabic text. In [1] the previous study explored the challenges of SA in Arabic language and the lack of research on ASA compared to English and other Latin languages. The study proposed a new approach to analyzing sentiment in Arabic script using the comments dataset of users of some mobile applications reviews available on the Google Play Store. The approach involved improving algorithms such as the Levenshtein distance (LD) algorithm for data preprocessing and combining it with the K-NN algorithm. The study conducted experiments to investigate the impact of utilizing the K-NN and LD algorithms for ASA on mobile applications reviews effectively. The results showed that the K-NN with LD algorithm achieved the highest accuracy, recall, precision, and F-score evaluation measures. The study demonstrated the potential of the proposed approach to improve the accuracy and effectiveness of SA in Arabic text. In [2], the previous study aimed to understand customer opinions towards mobile banking services' applications and to improve and maintain these applications. The study used application review-based SA to analyze user comments collected from banking mobile apps on Google Play Store. The dataset was labeled manually into three classes: positive, negative, and neutral. ML techniques, including NB, KNN, DT, and SVM models, were utilized for ASA. The NB model outperformed the other algorithms, achieving the highest accuracy, recall, precision, and F-score measures. The study demonstrated the potential of using SA to understand user requirements and improve application quality in mobile banking services. In[4], the study introduces a machine learning (ML) based method for SA, focusing on students' evaluations of higher educational institutions. The researchers analyze a dataset consisting of approximately 700 student reviews written in Turkish. They employ conventional text representation schemes and ML classifiers in their analysis. In the experimental analysis, three conventional text representation schemes (term-presence, term-frequency, and TF-IDF) and three N-gram models (1-gram, 2-gram, and 3-gram) are considered alongside four classifiers (support vector machines, Naïve Bayes, logistic regression, and RF algorithm). Additionally, the predictive performance of four ensemble learners (AdaBoost, Bagging, Random Subspace, and voting algorithm) is evaluated. The empirical findings demonstrate that the ML based approach shows promising results in assessing students' evaluations of higher educational institutions. In [5], the study aimed to systematically review the recent advancements in the application of SA in higher education. The primary objectives were to categorize the commonly used and successful SA techniques and methods within the higher education domain. The researchers conducted a systematic mapping review of 840 articles, ultimately selecting 22 relevant studies based on predetermined criteria. The

findings indicated that the selected studies primarily focused on six domains in applying SA within higher education, with a particular emphasis on evaluating teaching quality. The study also revealed that the utilization of specific SA techniques could prove to be a valuable tool for institutions in addressing specific learning challenges, improving the quality of higher education institutions, and evaluating the teaching process and teachers' performance. In [6], to analyze healthcare researchers' emotions towards ChatGPT, the researchers utilized the pre-trained BERT (Bidirectional Encoder Representations from Transformers) model and conducted SA and topic modeling on social media posts. In [7], the researchers conducted an initial assessment of ChatGPT's ability to comprehend opinions, sentiments, and emotions within the text. This evaluation included four settings: standard evaluation, open-domain evaluation, polarity shift evaluation, and sentiment inference evaluation. To carry out this evaluation, they utilized 18 benchmark datasets and 5 SA tasks, comparing ChatGPT, performance with fine-tuned BERT and other state-of-the-art models in end-task scenarios. Additionally, they conducted human evaluation and presented qualitative case studies to gain further insight into ChatGPT's SA capabilities. In [8], the researchers utilized BERT and ChatGPT to perform SA on Lyme disease. Their study provides a practical guide to conducting SA in the domain of tick-borne diseases using Natural Language Processing (NLP) techniques. The researchers aimed to demonstrate how the occurrence of bias in the discourse neighboring chronic manifestations of the disease can be evaluated. They used a dataset of 5643 abstracts from academic journals on the topic of chronic Lyme disease to show the steps involved in conducting SA using pre-trained language models with Python. The researchers validated their preliminary results using interpretable ML tools and a novel methodology that employs emerging state-of-the-art large language models like ChatGPT. In [9], the researchers aimed to investigate whether ChatGPT could effectively replace human-generated label annotations in social computing tasks, potentially reducing the cost and complexity of such research. To test this, they used ChatGPT to re-label five influential datasets covering SA, stance detection (twice), bot detection, and hate speech. Their findings suggest that ChatGPT does have the potential to handle these annotation tasks, although there are still some challenges to overcome. Overall, ChatGPT achieved an average accuracy of 0.609, with the highest performance observed in the SA dataset, correctly annotating 64.9% of tweets. However, the researchers noted significant variation in performance across different labels. This study has the potential to inspire new avenues of analysis and serve as a foundation for future research exploring the use of ChatGPT for human annotation tasks. However, the present study showed us a main motive to investigate the impact of ChatGPT on the analysis of sentiment of the Arabic language.

III. Methodology

This study aimed to investigate the impact of ChatGPT and Bard Google on ASA. It included 275 comments. The proposed method for labeling ASA using ChatGPT and Bard Google involves two different approaches, which are outlined below. The approach consists of five main phases, as illustrated in Fig. 1. In this paper, there are several stages. In the initial stage, the researchers collected "student comments" and review data for several mobile applications in Yemeni universities taken from the Google Play Store for the period from August 15, 2023 to September 10, 2023. The data was extracted and collected manually for six different Yemeni universities applications, which are available on Google Play. The list includes (Sana'a University, Thamar University, University of Science and Technology, Yemeni Jordanian University, Queen Arwa University, and *Unified Electronic Coordination Portal for Yemeni Universities*.

A. Data collection

This is the first phase of the approach. It included only aggregated Arabic reviews. The total number of reviews was 300 extracts from Google Play, 275 were considered. Table 1 shows statistics for reviews based on their apps.

B. Labelingmanually by humans, Labeling by ChatGPT, and Labeling by Bard Google.

The second phase was divided into three approaches as follows:

- **Labeling manually by humans**

This approach involves manually labeling Arabic text for SA by human annotators. The results show that the majority of the selected comments are classified as positive, with 150 positive records. The number of negative records is 102, while the number of neutral records is 23. Therefore, this approach involves labeling Arabic text for SA manually by human annotators. The annotators read and analyze the text to determine the sentiment expressed and label it as positive, negative, or neutral. This approach serves as a benchmark for evaluating the accuracy of the other labeling approaches.

- **Labeling by ChatGPT**

A novel approach to labeling for ASA utilizing ChatGPT. In this paper, we propose one method for investigating labeling of ASA using ChatGPT as follows:

This approach involves using ChatGPT to label Arabic text for SA. The results show that the majority of the selected comments are classified as positive, with 156 positive records. The number of negative records is 99, while the number of neutral records is 20. This approach involves using ChatGPT, a state-of-the-art natural language processing model, to label Arabic text for SA. ChatGPT is used to classify text into positive, negative,

or neutral sentiment categories. Overall, the proposed approach aims to investigate the effectiveness of labeling approaches by ChatGPT for ASA. The use of ChatGPT, a highly advanced natural language processing model, provides an opportunity to improve the accuracy and efficiency of labeling for ASA.

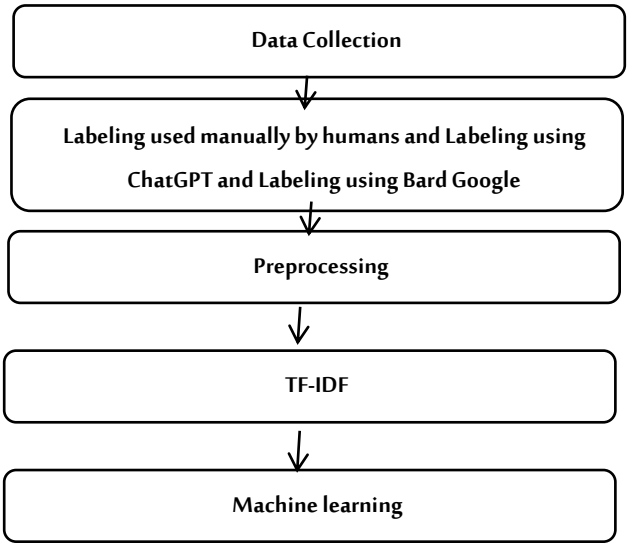


Fig. 1. Data collection and preprocessing steps.

TABLE 1. STATISTICS OF THE ARB-APPS COMMENTS DATASET.

No	Applications Name	Arabic Name	Comments Number
1	Sana'a University [10]	جامعة صنعاء	72
2	Thamar University [11]	جامعة ذمار	127
3	UST-DEV-TEAM [12]	جامعة العلوم والتكنولوجيا	18
4	ALSHIBANI Group [13]	الجامعة اليمنية الأردنية	3
5	QAU Dev. [14]	جامعة الملكة أروى	42
6	YCIT-HE. [15]	بوابة التنسيق الأهلي	13
Comments Total			**275**

10 https://play.google.com/store/apps/details?id=libosft.ye.com.sanaunif2

11 https://play.google.com/store/apps/details?id=thamar.univ.studentgate

12 https://play.google.com/store/apps/details?id=com.saqib.uststudentapp

13 https://play.google.com/store/apps/details?id=com.it.group

14 https://play.google.com/store/apps/details?id=edu.qau.queenarwauniversity.yemen

15 https://play.google.com/store/apps/details?id=org.ycit_he.p.nasseq

- **Labeling by Bard Google**

This is a novel approach to labeling for ASA utilizing Bard Google. In this paper, we propose one methods for investigating labeling of ASA using on Bard Google as follows: This approach involves using the Bard Google to label Arabic text for SA. The results show that the majority of the selected comments are classified as positive, with 135 positive records. The number of negative records is 89, while the number of neutral records is 51. This approach involves using Bard Google, a state-of-the-art natural language processing model, to label Arabic text for SA. Bard Google is used to classify text into positive, negative, or neutral sentiment categories. Overall, the proposed approach aims to investigate the effectiveness of labeling approaches by Bard Google for ASA. The use of Bard Google, a highly advanced natural language processing model, provides an opportunity to improve the accuracy and efficiency of labeling for ASA.

C. Pre-processing

In the third phase, the dataset is preprocessed by removing stop words, stemming, and applying other text normalization techniques. Data pre-processing is a crucial step in enhancing and extracting meaningful insights from data. This step helps remove inconsistencies and errors that may be present in the data, which can affect the accuracy of the analysis. There are several techniques involved in data pre-processing, which are summarized as follows [1]: Table 2. shows example of review data and Table 3 shows dataset statistics that have been preprocessed.

- **Removal**: This involves removing any irrelevant or redundant data that does not contribute to the analysis.
- **Folding of Case**: This step involves converting all the text data to a standard case; typically lowercase, to reduce the number of unique words in the dataset.
- **Tokenization**: This step involves breaking down the text data into individual words, known as tokens, to facilitate analysis.
- **Stop Word Filtering**: This technique involves removing commonly used words, known as stop words, from the dataset, as they do not add any significant value to the analysis.
- **Rooting or Stemming**: This step involves reducing the words in the dataset to their root form, which helps to reduce the number of unique words and improve the accuracy of the analysis.

TABLE 2. EXAMPLE OF REVIEW DATA.

ID	Original Comments	Sentiment Polarity
12	تطبيق في قمه الروعه والفكره ممتازه جدا مماخلقت لدينا حجات كثيره جدا .....واهمها ارتباط الطالب بالجامعه وووو.	Positive
32	لا أرى فيه أي فائده للطالب الجامعي في جامعة صنعاء	Negative
250	الف شكر لجامعة ذمار على هذا الانجاز وكذلك ملتقى الطالب الجامعي	Positive
264	قوه القوه هذا التنسيق عن بعد يتيح للطلاب التسجيل من اي مكان وفي لي وقت	Positive
186	ارجوا منكم اتاحة تعديل البيانات الشخصية!!	Neutral
98	البرنامج مايفتح تظهر شعار الجامعه فقط ولايفتح شيئ	Negative

TABLE 3. DATASET STATISTICS WHICH HAVE BEEN PREPROCESSED.

Unique word	2326
Not Letter	19
Punctuation	32
Stopword	473
Stemming word	1586
Not Stemming	216

D. Features Extraction:

TF-IDF algorithm is the fourth phase of the proposed approach. This algorithm is used to determine the importance of words in the dataset. It calculates the TF-IDF value for each word based on its frequency in the documents within the dataset. The TF-IDF value is increased for words that appear in fewer documents and decreased for words that appear in more documents. This technique is used to determine the significance of words in representing the text dataset and to reduce the impact of common words that do not provide any distinctive meaning [1].

E. Machine learning (ML)

In the fifth phase, and the last ML is a subfield of artificial intelligence that involves designing algorithms that enable computer systems to learn from data and improve their performance over time without being explicitly programmed. In SA and ML, algorithms are used to identify the sentiment expressed in a given text,

such as whether it is negative, positive, or neutral. K-NN, DT, and RF, are all ML algorithms that can be used for SA. Each algorithm has its strengths and weaknesses and may perform better or worse depending on the characteristics of the data being analyzed.

- **K-Nearest Neighbors (K-NN)**

K-NN is a non-parametric algorithm that classifies a given text based on the sentiment of its nearest neighbors, which are determined based on a similarity metric such as Euclidean distance. K-NN does not make any assumptions about the underlying distribution of the data and can be effective in SA when the data is high dimensional and complex [1].

- **Decision Tree (DT)**

DT is a popular machine-learning algorithm used for both classification and regression tasks. It models decisions and their possible consequences in a tree-like structure. In a DT, the dataset is split based on different attributes/features at each node of the tree. The algorithm selects the best attribute to split the data based on certain criteria (e.g., information gain or Gini impurity) to maximize the homogeneity or purity of the resulting subsets. The tree continues to grow by recursively splitting the data until a stopping condition is met. This can be a predefined depth limit, a minimum number of instances per leaf, or when there are no more attributes to split. Each leaf node represents a class or a predicted value for regression tasks. During the prediction phase, a new instance is traversed down the tree by following the decision paths based on the attribute values. The final prediction is determined by the class or value associated with the leaf node reached. DT have several advantages. They are easy to understand and interpret, as the resulting tree structure can be visualized. DT can handle both categorical and numerical data and are robust against noise and missing values. They can also capture non-linear relationships between features. However, DT are prone to overfitting, where they memorize the training data excessively. To overcome this, techniques like pruning or using ensemble methods like RF can be employed [3] (Al-Ghobesi, 2025).

- **Random Forest (RF)**

Random Forest (RF) is a popular ML technique that is used for both classification and regression tasks. It is an ensemble learning method that combines multiple DT to make predictions. In a RF, a collection of DT is created, where each tree is trained on a different subset of the data. To build each DT, a random subset of features is selected for each split, hence the term "random" in Random Forest. This random feature selection helps to introduce diversity among the trees and reduce the risk of overfitting. During the prediction phase, each DT in the RF independently makes predictions, and the final prediction is determined by majority voting

(in classification) or averaging (in regression) of the individual tree predictions. RF have several advantages. They can handle high-dimensional datasets with a large number of features, and they are robust against overfitting. They are also capable of capturing complex relationships between variables and handling missing data. Additionally, RF provide estimates of feature importance, allowing insights into the relative importance of different features in the prediction process (Omer, 2024, & Alasmari, 2023, Mleiki, 2025).

IV. Experiments and Results

In this study, we conducted SA of Arabic text using three ML techniques: K-NN, DT, and RF. We evaluated the performance of these techniques accuracy as the metric, and applied them to three different methods of labeling the data for ASA. These methods included manual labeling by humans, labeling by ChatGPT, and labeling by Bard Google. The goal of the experiment was to compare the performance of these techniques and methods and determine the most effective approach for SA of Arabic text. The accuracy scale was used to measure the performance of each method and technique, with higher accuracy indicating better performance. In the following section, we present the results of the experiment and compare the accuracy, precision, recall, and f-score measures of each technique and method, providing insights into the strengths and weaknesses of each approach for SA of Arabic text. The baseline models used in our experiments are in Section B. All the scripts constructed for the experiment are in RapidMiner. RapidMiner was used to create baseline models.

A. Evaluation criterion

To evaluate the effectiveness of a proposed approach, which involves using manual labeling by humans, labeling by ChatGPT, and labeling by Bard Google, we utilized one evaluation metrics in this paper. Four evaluation metrics have been utilized to evaluate ML models in this study. They are accuracy, precision, recall, and f-score measures. Before evaluating the effectiveness of the prediction model, the dataset was partitioned into testing and training sets. As shown in Eq (1), (2), (3), and (4).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+TN}} \quad (2)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

$$\text{F} - \text{score} = \frac{2*\text{Recall}*\text{Precision}}{\text{Recall+Precision}} \quad (4)$$

B. Baseline models

In the present study, we implemented four ML models as a baseline in our experiments. They are commonly used in the classification of approaches and constructed based on the training data. Examples include the K-NN, DT, and RF techniques.

C. Experimental results

Table 4 presents the results of different ML algorithms for SA, with each algorithm being labeled using different methods: manual labeling by humans, labeling by ChatGPT, and labeling by Bard Google، For the K-NN algorithm. Manual labeling by humans achieved an accuracy of 73.82%, while ChatGPT labeling achieved a slightly higher accuracy of 74.91%. Bard Google labeling had the lowest accuracy at 66.55%. In terms of recall, manual labeling had the highest value at 58.51%, followed by ChatGPT labeling at 56.82%. Bard Google labeling had a recall of 54.66%. For precision, both manual labeling and ChatGPT labeling had similar high values of 87.46% and 87.47%, respectively. Bard Google labeling had a precision of 81.15%. The F-score, which combines precision and recall, was the highest for manual labeling at 70.11%, followed by ChatGPT labeling at 68.88%. Bard Google labeling had an F-score of 65.32%.

TABLE 4. THE RESULTS ASA THROUGH MANUAL LABELING BY HUMANS, LABELING BY CHATGPT AND LABELING BY BARD GOOGLE USING K-NN TECHNIQUE.

Approaches	Accuracy	Recall	Precision	F-score
Labeling using manually by humans	73.82%	58.51%	87.46%	70.11%
Labeling By ChatGPT	**74.91%**	56.82%	87.47%	68.88%
Labeling By Bard Google	66.55%	54.66%	81.15%	65.32%

Table 5 presents the results of different ML algorithms for SA, with each algorithm being labeled using different methods: manual labeling by humans, labeling by ChatGPT, and labeling by Bard Google، For the DT algorithm. Manual labeling by humans achieved an accuracy of 57.45%, while ChatGPT labeling resulted in a slightly higher accuracy of 58.18%. Bard Google labeling had the lowest accuracy, with 53.45%. In terms of recall, manual labeling had the highest value at 44.93%, followed by Bard Google labeling at 41.18%. ChatGPT labeling had the lowest recall at 40.00%. For precision, ChatGPT labeling achieved the highest value of 52.52%, followed by Bard Google labeling at 50.44%. Manual labeling had a precision of 52.06%. The F-score, which combines precision and recall, was highest for manual labeling at 48.23%. Bard Google labeling and ChatGPT labeling had F-scores of 45.34% and 45.41%, respectively.

TABLE 5. THE RESULTS ASA THROUGH MANUAL LABELING BY HUMANS, LABELING BY CHATGPT AND LABELING BY BARD GOOGLE USING DT TECHNIQUE.

Approaches	Accuracy	Recall	Precision	F-score
Labeling using manually by humans	57.45%	44.93%	52.06%	48.23%
Labeling By ChatGPT	**58.18%**	40.00%	52.52%	45.41%
Labeling By Bard Google	53.45%	41.18%	50.44%	45.34%

Table 6 presents the results of different ML algorithms for SA, with each algorithm being labeled using different methods: manual labeling by humans, labeling by ChatGPT, and labeling by Bard Google. For the DT algorithm. Manual labeling by humans achieved an accuracy of 55.64%, while ChatGPT labeling resulted in a slightly higher accuracy of 57.09%. Bard Google labeling had the lowest accuracy, with 51.64%. In terms of recall, Bard Google labeling had the highest value at 35.96%, followed by manual labeling at 34.31%. ChatGPT labeling had the lowest recall at 33.67%. For precision, ChatGPT labeling achieved the highest value of 52.31%, followed by Bard Google labeling at 50.12%. Manual labeling had a precision of 51.72%. The F-score, which combines precision and recall, was highest for Bard Google labeling at 41.87%. Manual labeling and ChatGPT labeling had F-scores of 41.25% and 40.96%, respectively.

TABLE 6. THE RESULTS ASA THROUGH MANUAL LABELING BY HUMANS, LABELING BY CHATGPT AND LABELING BY BARD GOOGLE USING RF TECHNIQUE.

Approaches	Accuracy	Recall	Precision	F-score
Labeling using manually by humans	55.64%	34.31%	51.72%	41.25%
Labeling By ChatGPT	**57.09%**	33.67%	52.31%	40.96%
Labeling By Bard Google	51.64%	35.96%	50.12%	41.87%

.Fig. 2. This graph compares the accuracy of different labeling methods across various algorithms.

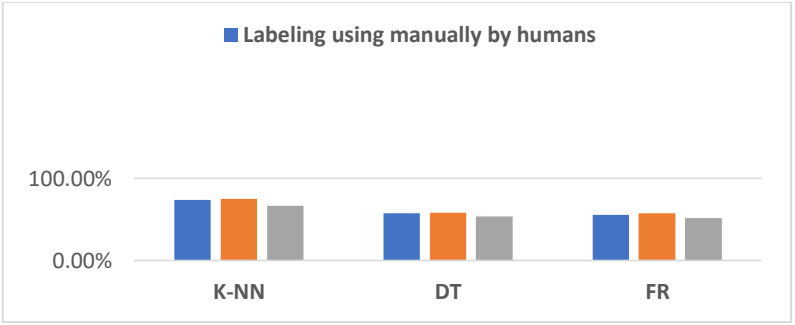


Fig. 2. A graphical representation of the Accuracy metrics for manual labeling by humans, labeling by ChatGPT and labeling by Bard Google.

V. conclusion

This study is dedicated to analyze the Arabic sentiments of manually collected Arabic dataset. It related to mobile applications for Yemeni universities. The SA tasks were based on the reviews collected from Google Play Store. The machine learning models used were K-NN, DT, and RF. Thus, the advantages and disadvantages of a particular mobile application can be seen based on its users' reviews. Several mobile applications in Yemeni universities providers can concentrate on correcting defects and better adjustment to the demands of their users. These methods included manual labeling by humans, labeling by ChatGPT and labeling by Bard Google. Based on the experimental findings, the K-NN technique demonstrated superior performance in Arabic sentiment analysis (ASA) using ChatGPT models, achieving an accuracy of 74.91%. Furthermore, the utilization of the proposed active labeling method with ChatGPT resulted in higher accuracy compared to other labeling methods. The study proposes that combining the K-NN technique with ChatGPT models and employing the suggested active labeling method are effective approaches for ASA using ChatGPT. The empirical results highlight the promising outcomes of the machine learning-based approach in evaluating students' opinions regarding higher education institutions. For future work, expanding the quantity of data (big data) and incorporating Arabic and English reviews are recommended.

References

[1] Alasmari, J. S. . (2023). The Dynamics of Verbal and Non-Verbal Linguistic Communication in The Saudi Sports Community. *Arts for Linguistic & Literary Studies*, *5*(4), 539–569. https://doi.org/10.53286/arts.v5i4.1676

[2] Al-Ghobesi, A. A. H. (2025). Risks of Relying on Artificial Intelligence in Learning Arabic Language Sciences Through the Meta Application. *Arts for Linguistic & Literary Studies*, *7*(1), 396–419. https://doi.org/10.53286/arts.v7i1.2420

[3] Al-Hagree, S., & Al-Gaphari, G. (2022). "Arabic Sentiment Analysis Based Machine Learning for Measuring User Satisfaction with Banking Services' Mobile Applications: Comparative Study". In 2022 2nd International Conference on Emerging Smart Technologies and Applications (eSmarTA) (pp. 1-4). IEE .

[4] Al-Hagree, S., & Al-Gaphari, G. (2022). Arabic Sentiment Analysis on Mobile Applications Using Levenshtein Distance Algorithm and Naive Bayes". In 2022 2nd International Conference on Emerging Smart Technologies and Applications (eSmarTA) (pp. 1-6). IEEE.

[5] Al-Shalabi, A. A., Al-Gaphari, G. (2023). Salah, A. H., & Alqasemi, F. Investigating the Impact of Utilizing the K-Nearest Neighbor and Levenshtein Distance Algorithms for Arabic Sentiment Analysis on Mobile Applications. JAST, 1(2).

[6] Baragash, R., & Aldowah, H. (2021). Sentiment analysis in higher education: a systematic mapping review. In Journal of Physics: Conference Series (Vol. 1860, No. 1, p. 012002). IOP Publishing.

[7] Mleiki, A. K. (2025). Exploring Saudi EFL Learners' Perspectives on Digital Writing Tools for Mitigating Emotional Challenges in Foreign Language Writing. *Arts for Linguistic & Literary Studies*, *7*(1), 577–590. https://doi.org/10.53286/arts.v7i1.2373

[8] Motlagh, N. Y. (2023). Khajavi, M., Sharifi, A., & Ahmadi, M. The Impact of Artificial Intelligence on the Evolution of Digital Education: A Comparative Study of OpenAI Text Generation Tools including ChatGPT, Bing Chat, Bard, and Ernie. arXiv preprint arXiv:2309.02029.

[9] Omer, N. I. M. (2024). Maintaining Meaningful Human Interaction in AI-Enhanced Language Learning Environments: A Systematic Review. *Arts for Linguistic & Literary Studies*, *6*(3), 533–552. https://doi.org/10.53286/arts.v6i3.2083

[10] Praveen, S. V., & Vajrobol, V. (2023). Understanding the perceptions of healthcare researchers regarding ChatGPT: a study based on bidirectional encoder representation from transformers (BERT) sentiment analysis and topic modeling. Annals of Biomedical Engineering, 1-3.

[11] Susnjak, T. (2023). Applying bert and chatgpt for sentiment analysis of lyme disease in scientific literature. arXiv preprint arXiv:2302.06474.

[12] Toçoğlu, M. A., & Onan, A. (2021). Sentiment analysis on students' evaluation of higher educational institutions. In Intelligent and Fuzzy Techniques: Smart and Innovative Solutions: Proceedings of the INFUS 2020 Conference, Istanbul, Turkey, July 21-23, 2020 (pp. 1693-1700). Springer International Publishing.

[13] Wang, Z., Xie, Q. (2023). Ding, Z., Feng, Y., & Xia, R. Is ChatGPT a good sentiment analyzer? A preliminary study. arXiv preprint arXiv:2304.04339.

[14] Zhu, Y., Zhang, P., Haq, E. U., Hui, P., & Tyson, G. (2023). Can chatgpt reproduce human-generated labels? a study of social computing tasks. arXiv preprint arXiv:2304.10145.

مجلــة العلوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

A Proposed Vision for Developing the Administrative System in Higher Education in the light of Artificial Intelligence Applications

Scientific Research

Taghreed Mahfooth Sultan Al-Zubairi

Abstract

This research aimed to plant a proposed conception for developing the administrative system in light of artificial intelligence applications in higher education in Yemen. The researcher explained the problem, objectives, and importance of the research, and referred to some relevant Arabic and foreign studies. The research tool also consisted of a questionnaire that was applied electronically, with a sample consisted of 137 individuals from higher education institutions in Yemen. The results demonstrated a high relative importance of using artificial intelligence in the administrative system. The research also mentioned the foundations of the proposed conception, the most important of which were: predicting the possibilities and imagining the desired future, taking into account that this future is variable due to artificial intelligence applications. Realism is achieved by monitoring the current reality and basing the goals that it seeks to achieve on available resources. Specific and prioritized goals are set based on statistics, data, and accurate information obtained through analysis of the internal and external environment. Flexibility and continuity are crucial, ensuring that the conception is an interconnected series of overlapping processes. Continuous follow-up and evaluation are necessary to track the success of the conception and assess the current situation, identifying its strengths and weaknesses. Several recommendations have been made, including: the development of higher education in light of artificial intelligence applications focuses on improving the technical infrastructure to enhance administrative performance efforts and to improve funding, spending, and work opportunities. The researcher also suggested providing highly qualified specialists to support the technical aspects by fixing network damage.

key words: *institutional trends, proposed vision, administrative system, higher education, artificial intelligence applications.*

1. Introduction

The world is currently experiencing a prosperous era of science and technology, and achieving distinctive levels of scientific and technical progress in a significant way that includes various aspects of life. Technology has become one of the necessities of the time, the basic driver for any human progress, and one of the most important factors in the development of human society. Today, the world is confronted with numerous challenges that have triggered a profound transformation in humanity. The outcome of this transformation will largely depend on how we address risks and seize opportunities. This has made the world as an electronic village. (Ali, 2020, Al-Ghobesi, 2025)

Scientific developments and their advanced and renewable applications have created an administrative, educational, and technical reality different from what it was in previous decades. Artificial intelligence is considered one of the most important outcomes of the Fourth Industrial Revolution due to its multiple uses. It is expected to open the door to limitless innovations, and this will lead to more industrial revolutions. This radical change in human life positions artificial intelligence as the driving force behind progress, growth, and prosperity during the next few years. The resulting innovations will establish a new world that may now have seemed unimaginable until now. Many smart technologies based on artificial intelligence have emerged, surpassing limits in production capabilities and effectiveness of use. Human minds have begun to work hard and study them to develop them in service of educational institutions, benefiting them and improving the quality of their performance (Deesing, 2017, Omer, 2024).

The trends of most institutions have shifted from traditional administrations to modern administrations characterized by the use of technology and employing and benefiting from artificial intelligence techniques, which in turn work to increase the efficiency of the performance of public institutions and private universities (Luo, 2018, Alasmari, 2023).

Hence, higher education institutions in Yemen must try to develop their administrative and technical systems, and we must interact with global changes and the modern requirements, conditions, and necessities of the new reality and possess the flexibility necessary to advance towards the future because they are obligated to adapt to the reality imposed by the era as, today, globalization and transformations are using modern technologies to be the best source of information.

2. The Problem of the Research

In the present, the most universities develop their countries which have turned to improving and developing their management systems. In this era, there is no other alternative options for these universities

but to turn to electronic management by employing artificial intelligence in their management (Al-Khataybah, 2015). The Fourth Industrial Revolution also brought a set of challenges, the most important of which are outlined by (Al-Arabi and Al-Qashlan, 2009-509, Mleiki, 2025):

I. **Electronic Challenges**: They are represented by the dangers of the Internet, which requires the state to work to protect cyberspace.

II. **Artificial Intelligence Challenges**: They are represented in the software system that will enable machines to think and decide their actions without human intervention.

III. **Geopolitical Challenges**: They are represented by conflicts and divisions that will harm the global development process.

IV. **Economic Challenges:** They are represented by the disparity in the distribution of wealth and income levels between social groups and occupy second place in the list of risks mentioned in the report (2018). This requires the world to make highly efficient technologies available to all countries.

In light of what has been mentioned above, it is clear that there is a large gap between scientific development and technological development. Given the disability of higher education institutions in Yemen to respond to these developments, the administrative system in higher education lacks clear-cut strategies that are characterized by a bureaucratic nature, which makes it difficult for higher education institutions to face challenges. The Fourth Industrial Revolution began with traditional methods of management.

In light of what has been mentioned above, the research problem is determined by answering the following questions:

1- What is the proposed scenario for the development of the administrative system in higher education in light of artificial intelligence applications in Yemen?

2- What are the obstacles to the proposed conception for developing the administrative system in higher education in light of artificial intelligence applications?

3- Is there an influential relationship between artificial intelligence applications and the development of the administrative system in higher education in Yemen?

4- What are the most important recommendations and suggestions for developing the administrative system in higher education in light of artificial intelligence applications in Yemen?

3. Objectives:

1- To review the proposed conception for the development of the administrative system in higher education in light of artificial intelligence applications in Yemen.

2- To address the obstacles hindering the implementation of the proposed conception for developing the higher education administrative system in light of artificial intelligence applications in Yemen.

3- To explain the relationship between artificial intelligence applications and the development of the administrative system in higher education in Yemen.

4- To provide recommendations and suggestions to develop the administrative system in higher education in light of artificial intelligence applications in Yemen.

4. The Importance of Research:

1- The theoretical significance:

This research addresses a modern and universal topic in the Arab, regional and global countries. The research aims to contribute to the development of a proposed conception for developing the educational administrative system in Yemen in light of the applications of artificial intelligence. This aspect of the research emphasizes the significance of artificial intelligence.

2- The applied importance:

The applied importance highlights the practical significance of this research, as it aims to draw the attention of professionals in the field of higher education to the importance of employing artificial intelligence in the administrative systems of both public and private universities in Yemen. It is also expected that this research will benefit all decision-makers in higher education by providing them with a proposed conception to develop the administrative system in light of artificial intelligence applications. Furthermore, it can serve as a valuable resource for planners providing them with the necessary information to effectively plan for the future. Additionally, it may also benefit those in charge of higher- education institutions in order to advance education in light of artificial intelligence.

5. **Theoretical framework**:

The proposed conception for developing the administrative system in higher education in light of artificial intelligence applications in Yemen provides better educational opportunities for administrators. It is used to

create interactive learning environments under the supervision of efficient administrative system professionals, thereby improving their time management and efficiency. The state's efforts to develop the administrative system in higher education, to keep pace with the changes in the light of artificial intelligence applications, involve the following strategies (Abu Al-Nour, 2023, Ahmed, 2025):

1. 1-Setting clear and relevant goals that meet the needs of education and its divisions.
2. Prioritizing the achievement of objectives with material capabilities.
3. Employing a technical approach to analysis the future prospects.
4. Recognizing that the future is constantly changing due to technical development and artificial intelligence applications.
5. Emphasizing realism by monitoring the current reality and building upon it.
6. Possibilities: The conception of different activities in the educational process.
7. Flexibility: The perception of the greatest degree of freedom to face unrealistic changes.
8. Implementing continuous follow-up and evaluation to ensure ongoing improvement.

The requirements for developing the higher education system:

1. Developing administrative systems based on the Internet and the applications of artificial intelligence.
2. Creating functions related to the management of higher education institutions based on artificial intelligence applications.
3. Improving administrative performance using artificial intelligence applications.
4. Holding workshops for administrative leaders to adopt the culture of artificial intelligence with education.

Mechanisms to activate the requirements for the employment of artificial intelligence applications

1- To improve the administrative systems by relying on the Internet and artificial intelligence.

2- Management of higher education institutions according to artificial intelligence applications.

3- Building an administrative system to regulate the management of digital content by the applications of artificial intelligence.

4- Training leaders to spread the culture of artificial intelligence applications in the development of education.

Obstacles of applying the proposed conception:

1- The weakness of the national economy that limits the applications of artificial intelligence in the higher education systems in Yemen.

2- The widening gap between contemporary real needs and the traditional system of higher education policies.

3- Some officials make decisions alone and do not allow the other opinion or criticism.

4- There are not the legislations obligating which is the supporting existence, such as the ministries, organizations and institutions to develop the structure system in higher education.

Ways to overcome the obstacles of applying the proposed conception:

1- To activate more projects, seminars and conferences on artificial intelligence applications in education.

2- Opening channels of communication between local education institutions and external education to gain experiences.

3- Enact legislation obligating the state ministries to support the development of the administrative system of higher education in light of artificial intelligence applications.

4- Motivating leaders to look up and practice the applications of artificial intelligence in higher education.

6. **Related works**:

1- Alia Al-Hwaiti (2022): It aimed to examine the degree of acceptance of faculty members in Jordanian universities to use the applications of artificial intelligence in light of the unified theory of acceptance and use of technology, UTAUT, and the study sample consisted of 250 faculty members, and the study tools consisted of a questionnaire that was applied electronically.

2- Majed Al Habib (2022): Employing artificial intelligence applications in training faculty members in Saudi universities, from the point of view of education experts, and the obstacles that limit their employment. The results showed that the sample members agreed to a medium degree on the fact that these applications related to artificial intelligence were employed in the training of faculty members, and they agreed that there were obstacles that limit their application in the training of the faculty members, and then presented a proposed conception to overcome them.

3- The study (Zhao, Chen, Liu, Zhang & Copland, 2019) in China: It aimed to reveal the impact of the use of online AI-based teaching systems. In order to analyze the studies that used the Internet AI- based teaching systems. The results indicated that the use of Internet AI- based teaching systems positively affected the degree of academic achievement of students.

4- Hudasi & Ady study (2020), conducted in Hungary, aimed at discovering data management methods in smart city systems by using artificial intelligence. To achieve the goal of the study, a specific methodology based on data generation was used, and the result of the study indicated that smart cities adopt artificial intelligence systems that are characterized by their high sensitivity.

7. **Research tools**:

To achieve the objectives of the research and to collect field data from the vocabulary of the random sample and to answer its questions, the data was processed through the use of many statistical methods and procedures by the Social Program Statistical Sciences Analysis (SPSS). The research included higher education institutions in the Republic of Yemen, and the research sample was a random sample, and it targeted the 137 members of the administrations. This indicates that the percentage of the 11 to 15-year-old category is suitable for enhancing confidence. The validity of the answers obtained from the respondents.

Table (2) shows the mean and standard deviation

	PHRASES	NOT TOTALLY AGREE	NOT AGREE		AGREE	TOTALLY AGREE	THE MEAN	STANDARD DEVIATION
		Number	number	number	number	number		
		Percentage	percentage	percentage	percentage	percentage		
1	Adding a new subject to increase sensory comprehension and technical skills in the administrative system	9	17	34	38	39	3.5912	1.2099
		6.6	12.4	24.8	27.7	28.5		
2	Artificial intelligence application drawing contributes to the development of the administrative system	6	13	31	56	31	3.6788	1.0637
		4.4	9.5	22.6	40.9	22.6		
3	Providing the modern proposed visualization training programs to expand the use of artificial intelligence applications for system development	6	11	34	58	28	3.6493	1.0848
		4.4	8.0	24.8	42.3	20.4		
4	To show the proposed conception of presenting creative ideas and solutions in the administrative system at higher education	10	14	30	46	37	3.6277	1.1944
		7.3	10.2	21.9	33.6	27.0		

	PHRASES	NOT TOTALLY AGREE	NOT AGREE		AGREE	TOTALLY AGREE	THE MEAN	STANDARD DEVIATION
		Number	number	number	number	number		
		Percentage	percentage	percentage	percentage	percentage		
5	The proposed conception is based on artificial intelligence systems that are characterized by their precise sensitivity to rapid development of the administrative system	9	11	36	52	29	3.5912	1.1085
		6.6	8.0	26.3	38.0	21.2		
6	The proposed conception of management leadership develops its expertise by working efficiently and effectively.	15	25	31	45	22	3.5474	3.6941
		10.9	17.5	22.6	32.8	16.1		
7	The proposed conception is used to improve the outputs of higher education	8	8	36	56	29	3.6569	1.0602
		5.8	5.8	26.3	40.9	21.2		
8	The holding of training workshops helps the administrative leadership to develop the ideas of the administrative system	6	11	35	58	27	3.6642	1.0329
		4.4	8.0	25.5	42.3	19.7		
	The result	76	125	296	464	274	3.6355	1.3978
		6.17	9.96	24	37.62	22.23		

Source/ prepared by the researcher according to the results of the statistical analysis using SPSS.

8. The results

1- The results of the analysis are noted in the table, as the arithmetic mean of the answers of the sample members reached (3.6355), which is greater than the hypothetical average of (3), and this indicates that the total score is for answers to the sample members on aggregate.

2- According to the results shown in the above table, we note that there are statistically significant differences and a statistically significant relationship between the applications of artificial intelligence and the development of the administrative system of higher education in Yemen.

3- Highlighting the proposed conception of artificial intelligence applications on the development of the administrative system in higher education, and thus the main question was answered.

4- The results of this research agreed with some studies as (Zhao, Chen, Liu, Zhang & Copland, 2019) study to reveal an impact. The use of online AI-based teaching systems, and the results indicated that the use of Internet AI- based teaching systems positively affected the degree of academic achievement of students.

9. **Recommendations and Suggestions:**

Recommendations

1- The practice of the fourth and fifth generation techniques of artificial intelligence and its applications can be used in developing the administrative system in higher education to produce correct decisions that support keeping pace with contemporary and modern education in the era of contemporary technology.

2- To highlight that artificial intelligence application is a major form of scientific and technological progress, so it is necessary to take the advantages and avoid its disadvantages.

3- To activate the artificial intelligence applications in education, which aim to enhance human capabilities through cooperation between humans and machines in education, learning and work.

4- The development of higher education in light of artificial intelligence applications focuses on improving the technical structure to enhance management performance and improve funding and spending opportunities.

Suggestions:

1- Training administrative leaders and designing an electron-learning environment that is motivating through artificial intelligence applications

2- Spreading the culture of artificial intelligence among management personnel, developing awareness of its importance in the administrative system, and encouraging researchers and thinkers to do research.

3- Organizing continuous meetings with external countries using electronic platforms. And adapting to contemporary updates.

4- Providing highly qualified professionals with technical support to address network malfunctions before the application of artificial intelligence.

10. Conclusion

In the present, more people are talking about artificial intelligence, so this topic is anxiety-provoking, and its information is renewed in the world every second time, but it is not beneficial unless the foundations and proposed plans are developed that transform the dream into a reality in which it is applied to the development of the high-knowledge in light of the applications of artificial intelligence. So, everyone should

be careful and be careful not to suddenly become in another world, modern word. We must Provide highly qualified professionals with technical support to address their subjects by the application of artificial intelligence to make subjects more attractive and enjoyable exchange in traditional subjects, which is the stagnation in the institutes after that has it in the country

11. **References**:

Arabic References

1. Abdalla Ali Mohammad Ebrahaim (2023). *"Awareness level of science education practices in Al-Azhar and general education applications in secondary school teaching*": (analytical study) Al-Azhar University, Dakahlia Governorate, Egypt.
2. Abu Al-Nour Misbah Abu Al-Nour, (2023)" Development of Industrial Technical Education in Egypt in the Light of Artificial Intelligence Applications ": Suggested Visionary"" Beni Suef University, Journal of the Faculty of Education, January issue, Part One 2023 Badawi.
3. Al-Araby, Sherif Abdel-Moati& Ahmed Hassan (2009): "Developing the performance of higher education institutions in the light of organizational learning and total quality management", Al-Arabiya Journal of Quality Assurance University Education, No. (3) -2009
4. Al-Ghobesi, A. A. H. (2025). Risks of Relying on Artificial Intelligence in Learning Arabic Language Sciences Through the Meta Application. *Arts for Linguistic & Literary Studies*, *7*(1), 396–419. https://doi.org/10.53286/arts.v7i1.2420
5. Ali, Shaima Ali Abbas (2020): Activating the Principles of Governance at the Egyptian University to meet the challenges of the Fourth Industrial Revolution." Educational Journal, Seventy-sixth Issue, August 2020.
6. Alia Al-Hwaiti (2022). "The degree to accept the faculty members in Jordanian universities using of artificial intelligence applications in the light of the unified theory of acceptance and use UTAUT technology". Unpublished master's thesis. Middle East University, College of Educational Sciences, Jordan.
7. Majad Abn Abdallh Al habib 2022 " Employing the applications of artificial intelligence in training faculty members in Saudi Arabia, from the point of view of education experts: a proposed conception," journal of the Islamic University of Educational and Social Sciences. The Islamic university, Al maddena Al meenoura , Saudi 9 march, 317-276.
8. Mohamed Mohamed Abdel Hadi (2022)" Artificial Intelligence Applications in Education: Challenges and Future Prospects", Tenth Magazine, Issue Two, Issue Series (20), December 2022 Port Said - Egypt Magazine website via the Egyptian Knowledge Bank http://eaec.journals.ekb.eg Al-Khattabiya.
9. Nour Abdul-Hakim (2015)." *The impact of the application of electronic management on job performance*": Field study on female deanship employees at King Abdulaziz University in Jeddah. Al-Azhar University, Scientific Journal of the Faculties of Commerce, p. (10), 102-83.
10. The shipment Abdel Moneim El-Desouky Hassan, (2021) "A proposed conception to develop the performance of higher education institutions in Egypt in the light of artificial intelligence", Journal of the Faculty of Education,Port Said University, Issue 36 - October.

English References

1. Ahmed, M. R. A. (2025). Accreditation and Quality Assurance: Exploring Impact and Assessing Institutional Change in the US and Saudi Arabian Higher Education Institutions. *Arts for Linguistic & Literary Studies*, *7*(1), 626–639. https://doi.org/10.53286/arts.v7i1.2419
2. Alasmari, J. S. . (2023). The Dynamics of Verbal and Non-Verbal Linguistic Communication in The Saudi Sports Community. *Arts for Linguistic & Literary Studies*, *5*(4), 539–569. https://doi.org/10.53286/arts.v5i4.1676
3. Deesing ,J.(2017). The Role of Artificial Intelligence in *the Aviation Industry*, retrieved 2018-8-11 From www.lifewire.com,Edited.
4. Hudasi,L., & Ady,L.(2020). Artificial intelligence usage opportunities in smart city data management. Interdisciplinary Description of Complex Systems: INDECS,18(3),382-388.
5. Luo,D.(2018), "Guide teaching system based on artificial intelligence",International Journal of Emerging ,
6. Mleiki, A. K. (2025). Exploring Saudi EFL Learners' Perspectives on Digital Writing Tools for Mitigating Emotional Challenges in Foreign Language Writing. *Arts for Linguistic & Literary Studies*, *7*(1), 577–590. https://doi.org/10.53286/arts.v7i1.2373
7. Omer, N. I. M. (2024). Maintaining Meaningful Human Interaction in AI-Enhanced Language Learning Environments: A Systematic Review. *Arts for Linguistic & Literary Studies*, *6*(3), 533–552. https://doi.org/10.53286/arts.v6i3.2083
8. Zhao,L .,chen, L., Liu,Q., Zhang,M. &Copland,H .(2019),"Artificial intelligence-based platform for online teaching management systems", Journal of Intelligent & Fuzzy Systems, 37(1),45-51.

مجلـــة العلـوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

Review on AI in Higher Education Institutes in Yemen

Abdullah Rashed[1] **Fua'ad Abdulrazzaq**[2] **Ghadah Adel**[3]

[1,2]YCIT-HE, [3]Lebanese International University, Sana'a, Yemen

Thursday, March 6, 2025

Abstract

Nowadays, artificial intelligence is increasingly used in different sectors of industry. Its positive influence changes the operation process and forces these industries to implement artificial intelligence. Education is one of these industries that implement artificial intelligence. Despite this positive impact, in Yemen, up to our knowledge, AI has not been experienced in any education system. This paper aims to overview the AI in higher education institutions with special focus on Yemeni situation. A questionnaire was distributed among public and private Yemeni universities to measure the degree of AI implementation in their environment. Results show that only 30% of the responded public universities have AI programs and only 10% from private universities. These percentages are not promising and still in the traditional area. So, we recommend increasing the awareness of the importance of AI and its role on enhancing the quality of education in Yemen. Furthermore, we hope YCIT-HE and Higher Education Institutes will work on education technology.

Keywords: AI, education quality, Yemen AI, Yemen Education

Introduction

Nowadays, artificial intelligence (AI) is increasingly used in different sectors of industry (Mishra, 2019). Education is one of these industries that implement AI (Zhang, 2023). Traditional learning methods have become obsolete and rejected by various educational institutions (Krstić, 2022). Besides, AI systems have significantly changed the operation process (Albawwat, 2021) specially the education process (Krstić, 2022, Ahmed, 2025). On the other hand, globalization demands more talent cultivation in education (Zhang, 2023, Al-Ghobesi, 2025) and AI. For this reason, using AI in education has become increasingly apparent (Huang, 2021, Omer, 2024).

Implementing AI in education increases the urgency of embedding innovative technologies and new teaching and learning methods (Krstić, 2022, Alasmari, 2023, Mleiki, 2025).

Innovative AI technologies have an impact on the methods of teaching and learning (Huang, 2021). For instance, educators and policymakers would take better decisions and harness the benefits of AI technology and maximize its impact on developing effective communication skills among English language learners (Rusmiyanto, 2023). As result, AI in education changes the way teachers teach and the way students learn (Krstić, 2022). It will achieve both business and students' objectives through the integration of smart classroom teaching and AI technology (Zhang, 2023). Furthermore, the application of AI technology assists teachers in guiding students to exercise their lessons with application skills (Mishra, 2019). AI also provides personalized, adaptable experiences, interactive engagement (Zhang, 2023), real time teaching and feedback (Krstić, 2022) and that leads to student satisfaction and business growth for education institutions (Mishra, 2019).

Using AI in the education process will increase the competitive environment in private education institutions (Mishra, 2019). It is a viable aid in the field of education, including language acquisition, as technology advances (Rusmiyanto, 2023). It also improves practical abilities like ability to respond to a range of learning styles shortly (Krstić, 2022). For students, it will improve their abilities such as writing and offer a trustworthy simulation dialogue platform like spoken languages (Ghafar, 2023). AI systems are rapidly changing educational institutions at all levels of education, to help people learn effectively and meet their learning goals (Krstić, 2022).

However, AI creates challenges for qualifications on online education (Li and Su, 2020). One of the reasons is that education institutes today are not as flexible as those which will be supported shortly using AI (Krstić, 2022).

In short, AI has been introduced into the field of education, where their use has enormous potential to enhance the teaching and learning processes (Salas-Pilco, 2022).

In Yemen, to the best of the researchers' knowledge, AI has not been experienced in any education system. Moreover, it is seldom to find any scientific research which overviews this obstacle. Hereby, we overview AI applications and the possibilities of improving the quality of education. Thus, this study investigates the idea of the national experience of AI in education and recommending national policy.

This paper is organized as follows: introduction is presented in section 1, followed by literature review in section 2; discussion comes in section 3; and finally, conclusion is provided in section 4.

Literature review

Diverse studies handled Artificial Intelligence (AI) in education issues from different perspectives. Some discussed it as a review while others highlighted its effect and ethical matters. These related studies can be summed up in the following section.

Chheda et. al (2023) and Wang et. al (2023) both conducted a comprehensive examination of the existing literature regarding the utilization of AI in the field of education. This examination employed a combination of research methods, such as bibliometric analysis and content analysis. Chheda et. al (2023) study not only provided a wide-ranging overview of the research field, including major categories of applications and theoretical perspectives, but also delves deeply into the impact of AI on education. It identified four primary categories of AI applications: adaptive learning and personalized tutoring, profiling and prediction, intelligent assessment and management, and natural language processing. The study also emphasized the importance of integrating theories to guide research in this domain. Moreover, it underscored the necessity of considering a variety of outcomes resulting from AI applications, extending beyond mere teaching effectiveness and student learning performance. The outcomes of this study bore significant consequences for researchers and practitioners in the field of education, as they offered valuable insights into the current state of AI research in education and potential areas for further investigation. The study proposed future research directions, including the in-depth analysis of specific AI applications in education and exploration of the societal, organizational, and individual impacts of adopting AI. While Wang et. al (2023) study delved deeply into the impact of AI on education. The outcomes of this study bore significant consequences for researchers and practitioners in the field of education, as they offer valuable insights into the current state of AI research in education and potential areas for further investigation. However, both studies suggested the inclusion of a broader range of theories and frameworks to enhance

the understanding of the role of AI in education. In the same stream and on the other hand, Salas-Pilco et al. (2022) carried out a systematic review of the literature on AI and Learning Analytics (LA) in teacher education. Their findings indicate that machine learning algorithms were employed in most of the studies. Furthermore, the implications would be valuable for teachers and educational authorities, informing their decisions regarding the effective use of AI and LA technologies to support teacher education.

The study of Saputra et al. (2023) undertook a partly systematic review of the existing literature to examine the potential, difficulties, risks, and hindrances associated with the integration of AI in the field of education. It discerned the capabilities of AI in the provision of educational materials, assessment, administrative systems, and policymaking. The challenges encompass aspects of pedagogy, educational frameworks, and literacy. The threats affect the security of personal data, character development, and ethical concerns in education. The obstacles involve significant financial investments, inadequate teacher training, and adjustments in the structure of the curriculum. The study underscored the advantages of AI in personalized education and advocates for substantial investment in AI in education to enhance the quality of education. The research methodology employed a metanarrative approach, which entailed an examination of secondary data from Google Scholar. The opportunities, challenges, threats, and obstacles in the field of AI education were identified. The study underscored the significance of addressing ethical considerations, biases, privacy concerns, and curriculum modifications associated with AI. The recommendations put forth included the formulation of comprehensive policies, the adoption of ethical pedagogy, the preparedness of educators, collaboration among stakeholders, and the prioritization of AI literacy in the design of the curriculum to achieve an optimal AI revolution in education.

The research paper of Kassymova et al. (2023) explored the ethical issues surrounding the digitalization and utilization of AI in the field of education, with a focus on the importance of valuing human beings as much as the machines they created. It emphasized how the creators' perspectives on life can influence the development of AI systems, potentially leading to competition, manipulation, and a decline in human values. Additionally, the study underscored the global implications of digital education, including unforeseen distortions, challenges, and risks, while expressing concerns about its impact on emotional intelligence and the potential alienation from human interaction. The authors argued for a comprehensive understanding of the purpose and consequences of digital education, with an emphasis on incorporating cultural and ethical considerations into the development and implementation of AI systems in the educational context. Although the study did not explicitly state its limitations or provide specific recommendations, it suggested potential areas for improvement. These areas included the need for further empirical research to explore the ethical

implications of digitalization and AI in education, given that the paper primarily relies on theoretical analysis and literature review. Policymakers and educators are encouraged to consider cultural and ethical dimensions when developing and implementing AI systems in education. Collaboration among experts in education, technology, and ethics was also considered crucial to ensure responsible and ethical use of AI in education. Continuous monitoring and evaluation of the impact of digital education and AI systems on students' well-being, social interactions, and learning outcomes were proposed as ongoing research priorities.

Alternatively, some studies were aiming to improve education, such as the study of Trifonov et al. (2020) which introduced AI system that they provided dynamic principles and personalization in the curriculum. They studied the effect of AI applications to improve education, especially Cyber-security.

Another study by Chiu (2021) aimed to develop a curriculum model for AI with four aspects: produce, process, and praxis. His thematic analysis of collected data included individual interviews, teaching documents, and meeting minutes from teachers. His findings revealed six key components: AI knowledge, AI processes, and the impact of AI, student relevance, teacher-student communication, and flexibility.

Smolin et al. (2012) proposed an AI based framework to manage the quality of the syllabus. They applied AI methods to evaluate a syllabus based on such characteristics as validity, usability, and efficiency automatically. In addition, they provided user trials to show the advantages of the developed approach against the traditional human-based process of syllabi verification and evaluation.

Khan et al. (2023) delved into the transformative potential of AI in the field of education, with a particular focus on its ability to personalize learning paths and provide targeted interventions. It also underscored the significance of ethical considerations in the implementation of AI technology. However, the study suggested the necessity for additional research to optimize the applications of AI in education for future endeavors. The study thoroughly examined the role of AI in customizing the educational experience by offering tailored learning materials, facilitating online interactions, and providing flexible learning pathways. Additionally, it highlighted the ethical concerns that arise from the use of AI. The study recommended the continuous assessment of these ethical issues and suggested conducting more comprehensive research to fully maximize the potential of AI in the field of education. The paper placed significant importance on AI's capacity to personalize learning, identified gaps in knowledge, and offered targeted interventions. It emphasized the positive impact of AI on students' autonomy and metacognitive development. However, the paper acknowledged that ethical concerns, data privacy, and potential bias in AI

algorithms must be carefully addressed for successful integration. The collaboration between AI experts, academics, and practitioners was deemed crucial in this endeavor. The study concluded by emphasizing the transformative potential of AI in education and the necessity for further research, including the implementation of surveys, interviews, and longitudinal studies.

The study by Mara et al. (2023) highlighted the significance of AI in the field of education, particularly in higher education, by placing emphasis on its role in enhancing academic and professional performance. It highlighted the ability of AI to automate repetitive tasks, thereby allowing educators to allocate more time to meaningful teaching experiences. Furthermore, the study emphasized the potential for personalization in the learning process, which positively impacted teaching effectiveness, student engagement, and overall performance. While the study's practical implications highlight the advantages of AI in education, it cascaded short in terms of addressing potential limitations and lacks specific recommendations for future research in this domain.

Awad et al. (2022) investigated the interdisciplinary nature of AI, with a particular focus on its applications in the field of education. They focused primarily on around discussing theoretical concepts and practical applications rather than presenting concrete outcomes or sources of data It emphasized the significant role that AI plays in assisting students in making well-informed career decisions through the utilization of predictive modeling techniques. They found that advancements in AI, specifically in the areas of machine learning and deep learning, were highlighted as catalysts for enhancing productivity and minimizing errors in educational environments. However, it lacked specific empirical evidence or research findings to substantiate its assertions.

In Brazil, Reis et al. (2006) suggested to use an AI approach to develop the computer environment. They found that using this approach would increase the quality of the teaching system.

Cui et al. (2023) studied understanding the problems and challenges faced by innovative online education in the context of the new coronavirus epidemic and look forward to the future on this basis.

Li and Su (2020) designed an evaluation method for online teaching quality of basic education in the context of AI. Their results provided a good reference for the application of online teaching and AI in basic education.

In contrast, other studies were aiming to help students become familiar with the English context and cultural background; such as Zhang et al. (2023) who proposed a method and application assessment for designing a smart classroom for English language and literature based on AI technology. They found that

using smart classroom approach has many potential benefits for English language and literature education. Furthermore, educators can develop a more effective and inclusive approach to language learning that leverages the power of technology while respecting the cultural diversity and individual needs of their students

Ghafar, et al. (2023) found an indication that AI provides a positive learning environment for learning English. The study aimed to understand the function of AI in ELT and examine AI technologies in ELT.

Al-Maliki (2023) explored the interdisciplinary nature of AI with a specific focus on its applications within the field of education. It highlighted the importance of AI in assisting students in making well-informed decisions about their future careers using predictive modeling. The advancements in AI, particularly in the areas of machine learning and deep learning, were emphasized as catalysts for increased efficiency and a reduction in errors within educational environments. Furthermore, the paper examines the potential for AI to provide personalized learning experiences and its transformative impact on the technology industry. However, it acknowledged the ethical concerns and potential job displacement that need to be addressed. The paper lacked specific empirical evidence or research findings to support its claims.

Rusmiyanto et al. (2023) highlighted the transformative role of AI in English language education and its potential to address the diverse needs of language learners. They found that AI had the potential to significantly enhance English language learners' communication skills by providing personalized and interactive learning experiences.

Li (2020) studied the usage of artificial English-learning as mobile application to improve IELTS performance. The results showed that the application was affordable to be used as an online platform for foreign English learning.

Huang et al. (2021) studied the effect of AI application in the field of education especially in learning, teaching, and virtual classroom evaluation. The study found quality improvement in teaching and learning methods among teachers and students who used AI technology, as it made students' learning styles more diversified and personalized.

On the Middle East, some studies have mentioned the use of AI in education. One of these studies is the distribution of AI in the Midde East, which was presented by Jain (2018). The study furnished a comprehensive overview encompassing AI's definition, its pragmatic applications, its economic implications, and the imperative nature of digital transformation across diverse sectors in the Middle East. Furthermore, it underscored the potential economic ramifications of AI, particularly within the confines of Saudi Arabia,

where it is anticipated to make a substantial contribution to the Gross Domestic Product by the year 2030. It emphasized the significance of digital transformation in the realm of healthcare, with the objective of augmenting the proportion of Saudi citizens who possess a unified digital health record. Moreover, the study concentrated on AI's role in enhancing labor productivity through automation and offered an estimation of the magnitude of AI's influence on various industries in the Middle East until 2030.

Another study was implemented on Kuwait by Al-Husseini (2023) to examine the importance of AI in the progression of primary stage education in Kuwait, aligning with the State of Kuwait 2035 Vision. It scrutinized the obstacles encountered in implementing AI in education, especially from the viewpoint of science educators in Kuwait. The research, carried out on 50 male and female science educators in the Hawalli educational district, reveals a minimal level of consciousness among these educators concerning AI's role in science education, a significant decrease in awareness of how to utilize AI applications in this context, and a general lack of comprehension regarding the significance of AI in science education. These findings emphasize the necessity for augmented awareness, training, and support for science educators to effectively incorporate AI into their classrooms. The constraints of the study encompassed its concentration on a specific sample of 50 science educators in one educational district, which might not be entirely representative of all science educators in Kuwait. Moreover, the descriptive approach employed in the research might restrict the comprehensiveness of the analysis of the obstacles and the importance of AI in science education.

Yemen, on the other hand, was highlighted in a study of Mutair (2022) which delved into the incorporation of AI within the educational framework of Yemeni universities. It shed light on the pivotal role that AI plays in augmenting the educational process, delineated the impediments encountered in the implementation of AI, and underscores the significance of adapting educational practices to the technological progressions witnessed in the 21st century.

AI is expounded upon as a revolutionary technology that possesses the capacity to enhance pedagogical and didactic approaches employed within Yemeni universities. It presents opportunities for interactive and tailor-made learning experiences, multimedia-enriched content, and adaptive learning trajectories. Nonetheless, the paper also delineates the barriers impeding the complete realization of AI in education, which encompass fiscal constraints, insufficient information technology resources, and resistance to change. The research done by Mutair (2022) underscored the imperative for Yemeni universities to modernize their pedagogical methodologies and curricula to align with the intellectual and technological revolution. It called

for a comprehensive approach to effectively harness AI, ameliorate the caliber of education, and equip students with the necessary tools to confront the challenges of the contemporary era.

Overall, the study accentuated the potential advantages of AI in the realm of Yemeni education, while simultaneously acknowledging and addressing the challenges that must be confronted for successful implementation. It emphasized the utmost importance of embracing technological advancements to furnish education of the highest quality and adapt to the ever-evolving educational landscape.

Discussion

The researchers distributed the questionnaire via two WhatsApp groups with periodic reminders. The first group targeted rectors of the universities while the second group for the university's representatives in Ministry of Higher Education and Scientific Research.

The questionnaire was minted to measure whether these universities have an AI program and the resources (program description, tools, labs, and qualified instructors) for implementing this program.

Table 1: shows that twelve participants responses (40% of the targeted community) were received. Three public universities (25% of the received respondents and 11% of the targeted community) and nine private universities with 30%. No AI system or technique were mentioned when the sample was asked about teaching techniques systems and tools. Two universities of the three reported to have AI lab as in table 2.

Only three universities (30%) of the received respondents have AI programs. One of them has two Ph.D. holders while the other has one. However, the third one claimed to have 5 Ph.D. holders. None of the staff were from Western Europe or America.

Figure 1 shows that most of the staff graduated from an Indian university with 75% of the received respondents. Nobody graduated from a western country. Yemen came in second position.

The results show that most of the respondents have no AI programs even though some of these universities have qualified instructors who are majoring in AI. However, instructors in AI are still few which is not enough to build an AI. Also, tools and Yemen infrastructure do not support the implementation of AI. The program description for AI is still limited and does not include the major courses that should be involved in such a program until now. Most of these descriptions contain only three or four AI courses and the rest are related to either IT or networking programs. This indicates the weakness of resources in Yemeni educational institutes. 7% received public universities reported that they have AI Programs.

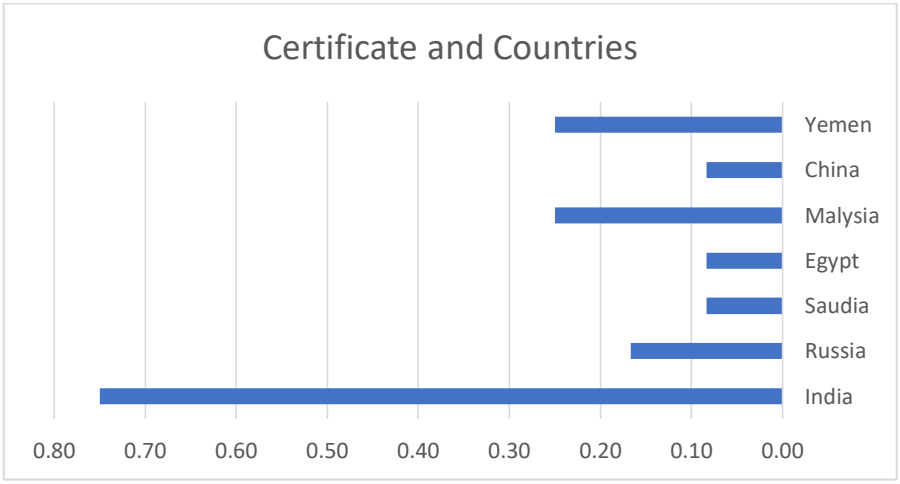


Figure 1: Certificates Origins

No	University Name	Year	Have AI Program	Staff	Degree	Graduation Country
1	Hodeidah	1996	No	3	Ph.D.	1 India + 2 Russia
2	Yemen Gulf	2014	No			
3	IBB	1996	No	4	3 Ph.Ds. +MSc.	1 Saudi 2 India 1 Egypt
4	Razi	2009	Yes	2	Ph.D.	
5	Andalus	1994	No			
6	Wehdah	2014	No			
7	Bayda	2008	No			
8	Shafaee	2012	No			
9	Nokhba	2020	No			
10	Watania	1994	Yes	1	Ph.D.	
11	Lebanese	2006	No	1	Ph.D.	1 China
12	University of Science and Technology	1994	Yes	5	Ph.D.	3 Yemen 2Malaysia

Table 1: Received Participants

N	University	Owns AI Systems	Owns AI Tools
1	Watania	No	No
2	Razi	No	Simulation lab
3	Science and Technology	No	Robotic Lab

Table 2: AI Systems and tools Architecture

Conclusion

This study presents a literature review to investigate the function of AI in the development of communication skills in education. Thus, it aims to look at the existing research and literature on the use of AI-based technologies in education. Furthermore, it opens with an overview about AI and its uses in education. It then investigates the various methods in which AI might help education systems to achieve goals.

There is a gap between YCIT-HE and the higher education institutions as the first one has the willingness to tackle future issues, while the latter did not respond to questionnaires and activities.

The study found that Yemen is still in infant stage considering AI in higher education institutions. The process will need huge efforts to leverage the level of education and AI. Besides, it needs experts and methodology to streamline them towards the objectives. It is necessary to bridge the gap before establishing AI national strategy or to be considered at the head of the table.

The study recommends universities to follow up the international standards for building an AI structure. This means developing AI programs that are fully specific to AI, not as a part of other programs. Universities should follow the standards for implementing AI infrastructure as it needs special equipment, labs, software, and tools. The monitoring and the approval process should be done under the supervision of YCIT-HE to ensure the quality of AI programs in Yemen. Finally, we should shed the light on Education Technology via workshops, seminars and projects.

References

1. Ahmed, M. R. A. (2025). Accreditation and Quality Assurance: Exploring Impact and Assessing Institutional Change in the US and Saudi Arabian Higher Education Institutions. *Arts for Linguistic & Literary Studies*, *7*(1), 626–639. https://doi.org/10.53286/arts.v7i1.2419
2. Alasmari, J. S. . (2023). The Dynamics of Verbal and Non-Verbal Linguistic Communication in The Saudi Sports Community. *Arts for Linguistic & Literary Studies*, *5*(4), 539–569. https://doi.org/10.53286/arts.v5i4.1676
3. Albawwat, I., Frijat, Y. (2021). An Analysis of Auditors' Perceptions Towards Artificial Intelligence and Its Contribution to Audit Quality, *Accounting*, 7(4), 755-762.
4. Al-Ghobesi, A. A. H. (2025). Risks of Relying on Artificial Intelligence in Learning Arabic Language Sciences Through the Meta Application. *Arts for Linguistic & Literary Studies*, *7*(1), 396–419. https://doi.org/10.53286/arts.v7i1.2420
5. Al-Husseini B. (2023). The role of artificial intelligence in teaching science to primary school students to achieve the vision of the State of Kuwait 2035. *Educational Journal of the College of Education in Sohag*, 108 (108), 153-176.
6. Al-Maliki W. (2023). The role of artificial intelligence applications in enhancing educational strategies in higher education (literature review). *Journal of Educational and Psychological Sciences*, 7(5), 93-107

7. Awad O., Mohamed Y., Shaheen, R. (2022). Applications of Artificial Intelligence in Education. Al-Azkiyaa: *International Journal Languages and Education*, Vol. 1., No. 1.
8. Chheda K., Chandrashekara A., Shankar P., Lima-Narváez M., Otero-Potosi S., Fuertes-Narváez E. (2023). The Role of Artificial Intelligence in Modern Education System. *Journal of Namibian Studies: History Politics Culture*, 35, 952-966.
9. Chiu T. (2021). A holistic approach to the design of artificial intelligence (AI) education for K-12 schools. *TechTrends*, 65(5), 796-807.
10. Cui Y., Ma Z., Wang L., Yang A., Liu Q., Kong S., and Wang H., 2023. A survey on big data-enabled innovative online education systems during the COVID-19 pandemic. *Journal of Innovation & Knowledge*, 8(1), p.100295.
11. Ghafar Z., Salh H., Abdulrahim M., Farxha S., Arf S., Rahim R. The Role of Artificial Intelligence Technology on English Language Learning: A Literature Review. *Canadian Journal of Language and Literature Studies*. 2023 Mar 9;3(2):17-31.
12. Huang J, Saleh S, Liu Y. A review on artificial intelligence in education. *Academic Journal of Interdisciplinary Studies*. 2021 May;10(206).
13. Jain, S. (2018). US $320 billion by 2030? The potential impact of AI in the Middle East, PwC.
14. Kassymova, G. K., Malinichev, D. M., Lavrinenko, S. V., Panichkina, M. V., Koptyaeva, S. V., & Arpentieva, M. R. (2023). Ethical Problems of Digitalization and Artificial Intelligence in Education: A Global Perspective. *Journal of Pharmaceutical Negative Results*, 2150-2161.
15. Khan D., Kumar A., Anil C., Solar C., Eswarareddy, S., Surbakti, H. (2023). Artificial Intelligence in Education: Enhancing Learning Experiences and Personalization. *Journal of Informatics Education and Research*, 3(2).
16. Li M. and Su Y., (2020). Evaluation of online teaching quality of basic education based on artificial intelligence. *International Journal of Emerging Technologies in Learning* (iJET), 15(16), pp. 147-161
17. Li, R., (2020), November. Using artificial intelligence in learning English as a foreign language: an examination of IELTS LIULISHUO as an online platform. In *Conference Proceedings. Innovation in Language Learning* 2020.
18. Mara, Cabanilla, Guerra., Galo, Ernestro, Cabanilla, Guerra. (2023). The Role of Artificial Intelligence in the Development of Teaching Effectiveness: A Tool for Personalization of Learning in Higher Education. doi: 10.22492/issn.2435-1202.2023.14
19. Mleiki, A. K. (2025). Exploring Saudi EFL Learners' Perspectives on Digital Writing Tools for Mitigating Emotional Challenges in Foreign Language Writing. *Arts for Linguistic & Literary Studies*, *7*(1), 577–590. https://doi.org/10.53286/arts.v7i1.2373
20. Mutair, A. Y. A. (2022). The Role of manufacturing cleverness on the developing education process at Arabic Language Departments in Yemeni Universities. *Al-Isbaah Journal*, Issue 7, July 2, 2022.
21. Omer, N. I. M. (2024). Maintaining Meaningful Human Interaction in AI-Enhanced Language Learning Environments: A Systematic Review. *Arts for Linguistic & Literary Studies*, *6*(3), 533–552. https://doi.org/10.53286/arts.v6i3.2083
22. Reis, M. M., Paladini, E. P., Khator, S., & Sommer, W. A. (2006). Artificial Intelligence Approach to Support Statistical Quality Control Teaching. Computers & Education, 47(4), 448-464.
23. Rusmiyanto R, Huriati N, Fitriani N, Tyas NK, Rofi'i A, Sari MN. (2023) The Role of Artificial Intelligence (AI) In Developing English Language Learner's Communication Skills. *Journal on Education*. ;6(1):750-7.
24. Salas-Pilco SZ, Xiao K, Hu X. (2022). Artificial intelligence and learning analytics in teacher education: a systematic review. *Educ. Sci.* 12 (8), 569.

25. Saputra, I., Astuti, M., Sayuti, M., & Kusumastuti, D. (2023). Integration of Artificial Intelligence in Education: Opportunities, Challenges, Threats and Obstacles. A Literature Review. *Indonesian Journal of Computer Science*, 12(4).
26. Smolin D., Butakov S., (2012). Applying Artificial Intelligence to The Educational Data: An Example of Syllabus Quality Analysis, LAK '12: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge April* 2012 Pages 164–169. https://Doi.Org/10.1145/2330601.2330644
27. Trifonov, R., Nakov, O., Manolov, S., Tsochev, G., & Pavlova, G. (2020, October). Possibilities for Improving the Quality of Cyber Security Education Through Application of Artificial Intelligence Methods. In *2020 International Conference Automatics and Informatics (ICAI)* (Pp. 1-4). IEEE.
28. Wang, S., Tran, T., Wang, J., Wang, F., Zhu, Z. and Du, Z., (2023). Artificial Intelligence in Education: A Systematic Literature Review.
29. Zhang X, Sun J, Deng Y. (2023). Design and Application of Intelligent Classroom for English Language and Literature Based on Artificial Intelligence *Technology. Applied Artificial Intelligence*. ;37(1):2216051.

مجلــة العلوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

Exploring the Potential of Yemeni Universities for Rankings: An Analytical Study of Prominent International University Ranking Systems

Fuaad Hasan Abdulrazzak[3], Salah AL-Hagree[1,2,4], Maher Al-Sanabani [3], Ahmed Al-Shalabi[1], Ayedh Abdulaziz Mohsen[2], Nashwan Ahmed Al-Majmar[2,4],Ghaleb Al-Gaphari[1], Mossa Ghurab[1],Mohammed Gamal[2]

[1] Computer Science Department, Faculty of Computer and Information Technology, Sana'a University, Yemen.

[2] Computer Science Department, Faculty of Sciences, Ibb University, Yemen.

[3] Computer Science Department, Faculty of Computer and Information Technology, Thamar University, Yemen.

[4] Computer Science, Aljazeera University, Yemen.

fuaad.abdulrazzak@tu.edu.ye, s.alhagree@gmail.com, s.alhagree@su.edu.ye, M.sanabani@gmail.com, a.alshalabi@su.edu.ye, ayedh992001@hotmail.com, almojammer2015@gmail.com, drghalebh@su.edu.ye, musa@su.edu.ye, m.alhoomaidi@gmail.com

Abstract:

Specifically targeting Yemeni universities, this research focuses on their potential to enter and enhance their positions in the international rankings such as Academic Ranking of World Universities (ARWU), Times Higher Education World University Ranking (THE), Quacquarelli Symonds World University Ranking (QS), and Webometrics Ranking and SCImago Institutions Rankings (SIR). Its objective is to encourage Yemeni universities to prioritize the criteria outlined in these rankings and view them as a means to access other international rankings. The research tackles the challenge of comprehending the criteria and application process required for Yemeni universities to participate in these rankings, as well as the rankings they have achieved. The study highlights the importance of this research by providing Yemeni university administrators with valuable insights regarding the criteria used for international rankings and the rankings obtained by Yemeni universities. This understanding can facilitate efforts to enhance the quality of higher education within these institutions. The findings of the study reveal that two Yemeni universities are listed in the QS. The University of Science and Technology, Yemen, ranks 151-170, while Thamar University in Thamar, Yemen, ranks 171-200. The findings of the study reveal that Ibb University secures the highest position among Yemeni universities in the SIR for 2023, with an overall score of 15.4 and a world rank of 3130. Sana'a University follows closely behind, ranking second among Yemeni universities according to these criteria, with a total score of 14.9 and a world rank of 3190.

Keywords: *QS, Higher Education, University Rankings, Yemeni Universities.*

Introduction

In recent years, global university rankings have become increasingly influential in assessing the quality and reputation of higher education institutions worldwide. These rankings serve as essential benchmarks for universities, students, researchers, and policymakers in evaluating academic performance, research productivity, and international standing (Ahmed, 2025)[14]. Recognizing the significance of these rankings, this study aims to investigate the rankings of Yemeni universities and shed light on their current positions and potential for improvement. Yemen, a country located in the Arabian Peninsula, has a diverse higher education sector comprising several universities that play a vital role in shaping the nation's intellectual and socio-economic development. However, Yemeni universities have faced numerous challenges, including limited resources, political instability, and regional conflicts, which have hindered their progress and international recognition. Understanding the rankings of Yemeni universities is crucial for several reasons. Firstly, it provides valuable insights into the current state of higher education in Yemen and allows for a comparative analysis with universities in other countries. This comparison helps identify areas of strength and weakness, enabling universities to focus on enhancing their academic programs, research output, and overall institutional performance[16]. Moreover, investigating the rankings of Yemeni universities can contribute to the ongoing efforts of university administrators, policymakers, and stakeholders to improve the quality of higher education in the country. By comprehending the criteria and methodologies used in these rankings, universities can align their strategies and priorities accordingly, aiming to meet the international standards set by renowned ranking systems. The purpose of this study is not only to examine the rankings of Yemeni universities but also to evaluate their potential for enhancing their positions in these rankings. By analyzing five prominent international rankings, namely Academic Ranking of World Universities (ARWU), Times Higher Education World University Ranking (THE), Quacquarelli Symonds World University Ranking(QS), Webometrics Ranking, and SCImago Institutions Rankings (SIR), this research aims to identify the strengths and weaknesses of Yemeni universities within the global higher education landscape. Furthermore, this study aims to encourage Yemeni universities to prioritize the criteria outlined in these rankings and view them as valuable tools for accessing other international rankings. By adopting a proactive approach and aligning their practices with the requirements of these rankings, Yemeni universities can enhance their visibility, attract international collaborations, and improve their overall academic reputation. To accomplish these objectives, this research will delve into the specific criteria and application processes required for Yemeni universities to

[16] https://www.elsevier.com/academic-and-government/university-rankings-guide

participate in these rankings. It will also examine the rankings already achieved by Yemeni universities, providing a comprehensive assessment of their current standing. By gaining a deeper understanding of the rankings of Yemeni universities, this study seeks to contribute to the larger discourse on improving the quality of higher education in Yemen. The findings of this research can guide university administrators, policymakers, and other stakeholders in formulating strategies to enhance the academic performance and international competitiveness of Yemeni universities.

The research problem is that the development of higher education has led to the emergence of international university rankings, which serve as indicators of the quality of education. Yemeni universities strive to enter these rankings and improve their positions. The research aims to answer the following questions:

- What are the criteria for entering international rankings, such as QS, ARWU, WEBOMETRICS, THE, and SCIMAGOI (SIR), and how are they applied by universities in Yemen?
- What are the ranks achieved by Yemeni universities in international rankings such as QS, ARWU, WEBOMETRICS, THE, and SCIMAGOI (SIR)? The goal is to understand the importance and impact of these rankings on improving the quality of higher education in Yemen.

The importance of this paper lies in its ability to provide university administrators in Yemen with an understanding of the criteria for international rankings such as QS, ARWU, WEBOMETRICS, THE, and SCIMAGOI (SIR), as well as the ranks achieved by Yemeni universities. This will allow them to recognize the importance and impact of these rankings on improving the quality of higher education at their universities. The paper aims to encourage Yemeni universities to pay attention to these international ranking criteria and consider them a gateway to entering other international rankings.

The paper aims to achieve the following objectives:

- To identify the criteria for entering international rankings, such as QS, ARWU, WEBOMETRICS, THE, and SCIMAGOI (SIR), and understand how they are applied by universities.
- To examine the ranks achieved by universities in Yemen in international rankings such as QS, ARWU, WEBOMETRICS, THE, and SCIMAGOI (SIR)
- To compare these rankings based on their indicators and weights.
- To provide recommendations for universities in Yemen based on the findings of the study.

The remainder of this paper is organized as follows: Section II provides an overview of previous literature on the topic; Section III explains the methodology; Section IV presents the investigation of the rankings of Yemeni universities and the study findings; and the final section concludes the paper.

I. LITERATURE REVIEW

In this section, let's have a look at the analytical study of **QS, ARWU, WEBOMETRICS, THE,** and **SIR**. In [1], the data used in the study is **SciVal data** from the **Scopus database**. This data was used to analyze the publication performance and the role of mentors at four Hungarian universities that are ranked in the QS rankings by subject. The methods used in the study include analyzing the publication performance and the role of mentors at four Hungarian universities that are ranked in the QS rankings by subject. The study also proposes a formula to calculate the vulnerability of the universities in the event of losing their mentors. The study aims to investigate the impact of mentors on the scientific competitiveness and ranking position of four Hungarian universities. The study identifies the topic clusters and topics that contribute to the publication output and citation impact of each university and assesses the risk level of losing mentors based on their age and contribution. The study also suggests some strategic implications for the universities to maintain and improve their ranking performance. In [2], the study uses a statistical analysis of a survey applied to 189 officials of various positions from higher education institutions, of which 41.2% correspond to Venezuela and 58.8% to Chile, who make up a total of 33 institutions between both countries. The survey measures the perception of nine dimensions and four context variables that use university rankings and how they influence institutional performance. The data used in the study is obtained from the survey responses of the officials, who rated the importance of each dimension and context variable on a Likert scale from 1 to 5. The data was analyzed using SPSS software and non-parametric tests such as U Mann-Whitney and Kruskal-Wallis. The study aims to comparatively analyze the perceptions that experts from universities in Chile and Venezuela have regarding university rankings and their impact on institutional performance. The study identifies the dimensions and context variables that are most relevant for the evaluation and positioning of the universities and examines the differences by country, type of institution, position, seniority, age, and gender of the respondents. The study also suggests some strategic implications for the universities to improve their ranking performance. In [3], the data used in the study are the 2014 QS World University Rankings by Mathematics subject, obtained from the QS official website. The data consist of the overall score and the four variables for each university. The study uses a hierarchical cluster analysis with Euclidean distance and average linkage to group the top 200 universities in mathematics based on four variables: academic reputation, employer reputation, citation per paper, and H-index citations. The study also uses the cophenetic correlation

coefficient to compare different clustering methods and evaluate the agreement between cluster analysis and QS rankings. The study aims to rank the universities using cluster analysis and to provide a different perspective from the existing QS ranking technique. The study identifies the natural clusters of the universities based on their similarities in mathematics and suggests that the distance between universities can be used as an alternative measure to rank them. The study also verifies the strong correlation between the overall score and the distance from the top-ranked university. In [4], the study used data from various sources, such as the official websites of universities, national agencies, metric databases, search engines, and surveys. The study covered 154 universities in Nigeria, including federal, state, and private universities. The study used a web crawling algorithm to extract data from university websites and other sources, and then aggregated nine existing ranking systems to form a unified system. The study also assigned percentage weights to 17 criteria based on their relevance and importance for ranking universities in developing nations. The study aimed to develop a purpose-built web ranking system for tertiary institutions in developing nations, using Nigeria as a case study. The study compared the results of the proposed ranking system with other existing ranking systems and found that the proposed system gave a better and more reliable ranking result for universities in developing nations. The study also suggested some recommendations for improving the quality and visibility of universities in developing nations. In [5], Bibliometric data from Thomson ISI, ARC, ERIH, Melbourne Institute and other sources; survey data from academic peers, funding agencies, subject associations and research centers; outlet ranking lists from various disciplines. Review of existing ranking systems and their limitations for HASS disciplines; analysis of indicators such as citation analysis and tiered outlets; discussion of the Australian Research Quality Framework (RQF) and its implications for HASS. Study description: The study focuses on the exclusion of HASS from university and discipline cluster ranking systems and examines some ways to address it. It argues that HASS disciplines need to develop and agree on suitable metrics that reflect their diversity and quality of research outputs. It draws on the Australian experience of introducing a new RQF that incorporates tiered outlets for all disciplines. In [6], Critical policy analysis and affect theories applied to publicly available texts from two commercial rankers (THE and QS). The study explored how THE and QS utilized emoscapes (affective landscapes) in global higher education policy. It demonstrated how rankers capitalized on stakeholders' emotions, desires, moods, and attitudes through affective infrastructures. The study emphasized the role of emoscapes in framing policy issues and promoting policy solutions. In [7], the study used an ordinary least squares model with ResearchGate Score as the dependent variable. It considered four explanatory variables: postgraduate programs, teacher's profiles in Google Scholar, institutional YouTube channel subscribers, and GDP per capita of the university's origin country. The sample consisted of the first

100 Latin American universities of the Webometrics Ranking. The study aimed to identify the variables influencing ResearchGate Score, which measures the scientific reputation and activity of researchers and institutions. The significant variables were found to be the number of postgraduate programs, teacher's profiles in Google Scholar, and the GDP per capita of the university's origin country. However, the number of subscribers to the institutional YouTube channel had a negative impact on the ResearchGate Score. In [8], the study employed a Gaussian Mixture Model (GMM) to cluster universities using the 2022 QS World University Ranking dataset. It utilized four ranking indicators: academic reputation score, faculty-student score, citation per faculty score, and international student score. om the official website of the 2022 QS World University Ranking. The sample included 1300 universities from various regions and countries. The study aimed to provide a fresh perspective on university rankings beyond the overall score. It identified four distinct clusters of universities with varied characteristics and performance in the ranking indicators. The study also discussed the implications of these clustering results for university managers and policymakers. In [9], the study critically analyzed three global university rankings (ARWU, QS, and THE) and their parameters, methods, data sources, and inconsistencies. It employed correlation analysis, case studies, and examples to illustrate the weaknesses and potential abuses of these rankings Data used from the official websites of the three ranking organizations and other sources such as SciVal, Wikipedia and Wikiwand. The sample consisted of the universities that appeared in all three rankings for the year 2017. The study examined the problems and challenges of university ranking systems and their impact on university policies and practices. It revealed that rankings prioritize quantity over quality, tolerate academic misconduct, exhibit bias, and lack transparency. The study suggested measures to improve the rankings and emphasized the importance of cautious usage. In [10], the study compared eight global university rankings (WR, THE, QS, ARWU, HEEACT, Leiden, URAP, and SIR) for the Mediterranean and Black Sea regions using three analytical procedures. These procedures included constructing matrices of university and country entries in the rankings, calculating proximity indicators, and performing correlation analysis using data from the official websites of the eight ranking organizations for the year 2011. The sample consisted of 29 countries and their universities in the region. The study aimed to develop a quantitative methodology for comparing global university rankings using spatial analysis. The findings indicated that Italy, Spain, and France had the highest number of university entries in the rankings, followed by Israel, Greece, Turkey, and Russia. Additionally, the study revealed a high correlation among the rankings. In [11], the study used data related to higher education systems from various sources to perform the ranking and analysis. The study utilizes a ranking methodology to evaluate and rank 94 higher education systems. It also employs Pearson coefficients of skewness and kurtosis calculation to

analyze global inequalities in higher education. The study analyzed global higher education systems' competitive positioning, identifying leaders, followers, and underperformers through cluster analysis. It discussed inequalities in enrollment rates, research quality, and university-industry collaboration. Enhancing quality assurance systems in developing countries was emphasized for improved competitiveness. The authors anticipated increased university specialization and their role in national innovation systems. In [12], the study utilized the numerical values of criteria from the QS World University Rankings and Times Higher Education rankings. The analysis focused on the Top 50 universities according to the QS ranking. The study utilized clustering (k-means) and classification processing (decision trees) as methods of analysis. The study identified factors determining the leadership of top universities in international rankings. It analyzed QS and Times Higher Education criteria, categorized universities into clusters, and determined influential criteria combinations. The findings serve as guidelines for improving rankings.

II. Methodology

Universities and educational institutions play a crucial role in higher education, and assessing the quality and ranking of these institutions is of paramount importance to students, parents, and the academic community. Various methodologies and indicators are available for evaluating and ranking universities, and among these well-known methodologies are **QS, ARWU, WEBOMETRICS, THE,** and **SIR**. This research aims to conduct an analytical study of these different methodologies with the goal of understanding, analyzing, and comparing them. The research focuses on analyzing the foundations and criteria used in these methodologies, in addition to evaluating the impact of each methodology on university rankings and positioning. In this section, we list the best university rankings in the world as follows:

1- Quacquarelli Symonds (QS)

Quacquarelli Symonds (QS)[17] is a global higher education ranking system that evaluates universities based on a range of criteria. These criteria include academic reputation, employer reputation, faculty/student ratio, citations per faculty, international faculty ratio, and international student ratio. The QS ranking system is widely recognized and respected by universities and employers around the world. It provides valuable information about the quality of education and research at universities and can help students make informed decisions about where to study. The table 1 shows the approved indicators in QS.

[17] https://www.topuniversities.com/qs-world-university-rankings

Table 1. Rating indicators adopted in QS.

No	Criteria	Indicators	Weight [15]
1	Academic Reputation	This measures a university's reputation among academics worldwide.	40%
2	Faculty/Student Ratio	This measures the number of academic staff relative to the number of students, and is used as an indicator of teaching quality.	20%
3	Citations per Faculty	This measures the number of citations received by a university's research papers, normalized by the number of academic staff at the university.	20%
4	Employer Reputation	This measures a university's reputation among employers worldwide.	10%
5	International Faculty Ratio	This measures the proportion of international academic staff at a university.	5%
6	International Student Ratio	This measures the proportion of international students at a university.	5%

2- Times Higher Education (THE)

THE[18] is another global university ranking system that evaluates universities based on a range of criteria. These criteria include teaching, research, citations, international outlook, and industry income. THE ranking system is widely recognized and respected by universities and employers around the world. It provides valuable information about the quality of education and research at universities and can help students make informed decisions about where to study. It evaluates universities based on five criteria. Table 2 shows the approved indicators in THE.

Table 2. Rating indicators adopted in THE.

No	Criteria	Indicators	Weight [15]
1	Teaching	This measures the learning environment and the quality of teaching at a university.	30%
2	Research	This measures the volume, income, and reputation of a university's research.	30%
3	Citations	This measures the influence of a university's research by counting the number of times its research papers are cited by other researchers.	30%
4	International Outlook	This measures the proportion of international students and staff at a university, as well as the university's international collaborations.	7.5%
5	Industry Income	This measures a university's ability to transfer knowledge to industry and attract research funding from businesses. Is there anything else you would like to	2.5%

[18] https://www.timeshighereducation.com/world-university-rankings

3- SCImago Institutions Rankings (SIR)

SIR[19] is a classification of academic and research-related institutions ranked by a composite indicator that combines three different sets of indicators based on research performance, innovation outputs and societal impact measured by their web visibility. The SIR focuses on research institutions, universities, hospitals, government agencies, and private non-profit organizations. The ranking is based on a number of indicators, including the number of publications, citations, international collaboration, normalized impact and others. Table 3 shows the approved indicators in SIR.

Table 3 Rating indicators adopted in SIR.

No	Criteria	Indicators
1	Research Performance	This criterion evaluates the research output of the institution and includes elements such as the number of published articles, citation counts, participation in peer-reviewed scientific journals, and research collaboration with other entities.
2	Innovation Outputs	These indicators consider non-academic activities that promote innovation and technological development. They may include factors such as patent rates, scientific investigations, collaboration with industry, and technology transfer.
3	Societal Impact Measured by Web Visibility	The social impact of an institution is assessed through factors such as the dissemination of scientific publications on the web, social interaction, and the digital impact achieved by the institution.

4- Webometrics

Webometrics is the largest academic ranking of higher education institutions. It is published by the Cybermetrics Lab, a research group of the Spanish National Research Council (CSIC), the largest public research body in Spain. The ranking aims to promote academic web presence and support open access initiatives. It provides web indicators for more than 20,000 universities worldwide, with the top 16,000 being published. The ranking is based on a composite indicator that takes into account both the volume of the web content (size, visibility, rich files) and the impact and quality of these contents (scholarship, excellence) measured by their visibility and impact [16]. The ranking is updated every January and July. Table 4 shows the approved indicators in webometrics.

[19] https://www.shanghairanking.com/

Table 4. Rating indicators adopted in Webometrics.

No	Criteria	Indicators	Weight [15]
1	Size	These measures the number of pages automatically linked in a single site.	20%
2	Rich Files	This measures the number of rich files (documents and textual information) that are present in the search engine and belong to the university's site.	15%
3	Scholar	This measures the scientific material in Google Scholar, including peer-reviewed research, reports, dissertations, abstracts in various scientific topics, as well as images, films, maps and others published electronically under the domain of the university's site.	15%
4	Visibility	This measures the visibility of a university's site on the internet through search engines and its appearance in these engines	50%

5- Acdemic Ranking of World Universities (ARWU)

ARWU, also known as the **Shanghai** Ranking, is a ranking of universities published by the Shanghai Ranking Consultancy. The ranking uses six objective indicators to rank world universities. Table 5 shows the approved indicators in ARWU.

. Table 5. Rating indicators adopted in ARWU.

No	Criteria	Indicators	Weight [15]
1	Number of alumni and staff winning Nobel Prizes and Fields Medals	This measures the number of alumni and staff from a university who have won Nobel Prizes or Fields Medals.	10%
2	Number of highly cited researchers selected by Clarivate	This measures the number of researchers from a university who are among the most highly cited in their fields.	20%
3	Number of articles published in journals of Nature and Science	This measures the number of articles published by a university's researchers in the journals Nature and Science.	20%
4	Number of articles indexed in Science Citation Index Expanded™ and Social Sciences Citation Index™ in the Web of Science™	This measures the number of articles published by a university's researchers that are indexed in these two databases.	20%
5	Per capita performance of a university	This measures the overall performance of a university, normalized by its size.	10%

6- AD Scientific

They based ranking system on the number of meritorious scientists. Four criteria are used to rank the countries. The first one is the number of scientists in the top 3 % list. The second criterion is the number of scientists in the top 10%, Top 20%, Top 40%, top 60%, and top 80% list. The third one is the number of scientists listed in the AD Scientific Index. In the case of equalities after applying all these three criteria, the world rank of the meritorious scientist of that country is used. Table 6 shows the approved indicators in AD Scientific Index[20].

Table 6. Rating indicators adopted in AD Scientific Index.

No	Criteria	Indicators	Weight
1	Number of scientists in the top 3% list:	This criterion focuses on the number of scientists from each university who fall within the top 3% in terms of their scientific achievements. These achievements could include publications, citations, research impact, or other relevant factors. The higher the number of scientists in this elite category, the higher the ranking for the university.	3%
2	Number of scientists in various percentile lists:	AD Scientific considers the number of scientists from each university who are ranked within different percentile categories, including the top 10%, top 20%, top 40%, top 60%, and top 80%. This criterion provides a broader assessment of the distribution of meritorious scientists across different performance levels.	top 10%, top 20%, top 40%, top 60%, and top 80%
3	Number of scientists listed in the AD Scientific Index:	AD Scientific likely maintains its own index or database of scientists, which includes individuals who have demonstrated significant contributions in their respective fields. This criterion assesses the number of scientists from each university who are listed in the AD Scientific Index.	
4	World rank of meritorious scientists:	In the event of ties or equalities after applying the previous three criteria, AD Scientific uses the world rank of the meritorious scientists from each university to break the tie. This means that if two or more universities have an equal number of highly accomplished scientists, the overall world ranking of those scientists will be taken into account to determine the ranking of the respective universities.	20%

[20] https://www.adscientificindex.com/country-ranking/?country_code=ye

7- EduRank

EduRank is an independent metric-based ranking of 14,131 universities from 183 countries. They utilize the world's largest scholarly papers database with 83,166,817 scientific publications and 1,801,313,576 citations to rank universities across 246 research topics. In the overall rankings, they add non-academic prominence and alumni popularity indicators. Always check official university websites for the latest enrollment information. Table 7 shows the approved indicators in EduRank[21].

Table 7. Rating indicators adopted in EduRank.

No	Criteria	Indicators	Weight
1	Research performance.	We use the OpenAlex database as a proxy to retrieve scientific publications and links between them (citations). Rather than just summarizing them, we build a graph with publications as nodes and citations as edges to calculate the weight of each publication. Then we adjust that weight for the publication date and share of university representatives in the list of authors.	45%
2	Non-academic prominence.	They use the same approach that Google and other modern search engines use to calculate the reputation of individual web pages - backlinks to a university from other sites. We use the data from Ahrefs as a source with the largest available index of pages and links.	45%
3	Alumni score.	The indicator reflects the combined number of page views that a university's graduates and other affiliated individuals have on all 43 language versions of Wikipedia.	10%

8- uniRank

uniRank's university ranking utilizes a simple and transparent methodology to evaluate universities worldwide. However, it's important to note that uniRank's ranking is not a comprehensive alternative to more complex and comprehensive rankings such as QS or Times Higher Education. uniRank's ranking is based on the following indicators: The current uniRank University Ranking™ is based upon an algorithm including four unbiased and independent web metrics extracted from three different web intelligence sources:. [25] [24] [2322]

Table 8 shows the approved indicators in uniRank.

[21] https://edurank.org/geo/ye/

[22] https://majestic.com/help/glossary#RefDomain

[23] https://support.similarweb.com/hc/en-us/articles/213452305-Rank

[24] https://moz.com/learn/seo/domain-authority

[25] https://www.4icu.org/ye/universities/

Table 8. Rating indicators adopted in uniRank.

No	Criteria	Indicators
1	Global Ranking:	Universities are ranked based on their overall global position. This is determined using a wide range of available online sources, including academic databases and official university websites.
2	Institutional Presence:	The online presence of the university and its institutional connections with other academic and research entities are taken into consideration.
3	Accreditation:	The official accreditation of the university by recognized accrediting bodies is considered. This includes accreditation from government bodies, professional bodies, and reputable academic organizations.

uniRank aims to provide a non-academic League Table of the top Yemeni Universities and Colleges based on valid, unbiased and non-influenceable web metrics provided by independent web intelligence sources, rather than data submitted by the Universities themselves.

9- CWUR

The Center for World University Rankings (**CWUR**) is responsible for publishing a unique global university ranking that evaluates the quality of education, alumni employment, faculty excellence, and research performance. What sets CWUR apart is that it doesn't rely on surveys or data submissions from universities. CWUR utilizes seven objective and reliable indicators categorized into four areas to determine the rankings of universities worldwide. The Center for World University Rankings grades universities on four factors without relying on surveys and university data submissions: education (25%), employability (25%), faculty (10%), and research (40%). This year, 62 million outcome-based data points were analyzed for the rankings. Table 9 shows the approved indicators in CWUR.

Table 9. Rating indicators adopted in CWUR.

No	Criteria	Indicators	Weight
1	Education:	This criterion evaluates the academic achievements of a university's alumni. It considers the number of alumni who have earned prestigious academic distinctions in relation to the size of the university, accounting for 25% of the ranking.	25%
2	Employability	The employability category assesses the professional success of a university's alumni. It examines the number of alumni who have attained top positions in major companies relative to the university's size, contributing to 25% of the ranking.	25%
3	Faculty:	This indicator focuses on the quality of the university's faculty members. It takes into	

No	Criteria	Indicators	Weight
		account the number of faculties who have received prestigious academic distinctions, contributing to 10% of the ranking.	10%
4	Research	**Research Output:** This aspect measures the total number of research papers produced by the university, accounting for 10% of the ranking. • **High-Quality Publications:** It considers the number of research papers published in top-tier journals, indicating the impact and quality of the university's research. This category contributes to 10% of the ranking. • **Influence**: The influence category evaluates the number of research papers published in highly influential journals, signifying the reach and significance of the university's research. It accounts for 10% of the ranking. • **Citations**: This criterion measures the number of highly-cited research papers affiliated with the university, reflecting the impact and recognition of its research. It contributes to 10% of the ranking.	40%

By considering these indicators, CWUR provides a comprehensive assessment of universities worldwide, allowing individuals to gauge the quality of education, alumni success, faculty excellence, and research performance across different institutions[26].

10- Leiden Ranking

The **Leiden**[27] Ranking is a ranking that focuses on the research performance of universities. It is developed by the Leiden Centre for Science and Technology Studies at Leiden University in the Netherlands. Here is an overview of the methodologies used in the Leiden Ranking:

The CWTS Leiden Ranking 2023 offers a sophisticated set of bibliometric indicators that provide statistics at the level of universities on scientific impact, collaboration, open access publishing, and gender diversity. The indicators available in the Leiden Ranking are discussed in detail below.

The Leiden Ranking aims to provide a comprehensive and objective view of the research performance of universities worldwide. While it primarily focuses on scientific research, it offers a greater understanding of international collaboration and the scientific impact of universities. It's important to note that there may be updates and changes to the ranking methodology over time. Interested individuals can refer to the official Leiden Ranking website for up-to-date information on the methodologies used[28]. Table 10 shows the approved indicators in Leiden.

[26] https://cwur.org/2023.php

[27] https://www.leidenranking.com/

[28] https://www.leidenranking.com/ranking/2023/list

Table 10. Rating indicators adopted in Leiden.

No	Criteria	Indicators
1	Core Indicators:	The Leiden Ranking uses a set of core indicators to measure the research performance of universities, including: • Number of research publications in scientific journals. • Number of citations received by these publications from other researchers. • Number of collaborative research between the university and other researchers or institutions.
2	Global Performance Analysis:	University performance is analyzed based on global data available in scientific databases such as Scopus. This data is used to analyze the research and collaborative performance of universities at the global level.
3	Achieved Performance:	Universities are ranked based on their actual and achieved research performance, rather than relying on factors such as university reputation or financial resources.

11- NTU Ranking

NTU Ranking (Nanyang Technological University Ranking) is a global university ranking that relies on several criteria to evaluate the quality of scientific research and academic performance. The ranking focuses particularly on scientific, technological, and engineering fields. Table 11 shows the approved indicators in NTU. Here are some key criteria used in the NTU Ranking:

Table 11. Rating indicators adopted in NTU.

No	Criteria	Indicators
1	Research Productivity:	Measures the number of research papers published in prestigious scientific journals and the level of citations received by that research.
2	Research Performance:	Measures various indicators related to published research and scientific impact, such as citation counts and the H-index.
3	Research Quality:	Relies on the ranking of journals where research papers are published and their inclusion in prestigious lists like the Journal Impact Factor.
4	International Collaboration:	Measures the extent of research collaboration between universities and researchers from different countries, such as the number of co-authors from different countries in research papers.
5	Innovation and Societal Impact:	Measures the level of innovation and the economic and social impact achieved by the research, such as the number of patents and the societal relevance of the research.

The NTU Ranking covers multiple fields, including medical sciences, engineering, technology, and social sciences. It aims to provide a comprehensive snapshot of universities' performance in scientific research and academic excellence[29].

12- Yemen research

The initial release of the **Yemen research** university ranking for the best universities in Yemen is based on a wide range of data sources, including the Information Technology Center under the Ministry of Higher Education. Bibliometric data was collected to estimate citation-based metrics as of August 1, 2023. The ranking relies on a set of criteria, with the most important criterion being the h-index for all affiliated researchers at a particular university. The h-index takes into account the papers and citation data across all disciplines. Their goal is to encourage scientists, business professionals, and administrative bodies throughout Yemen to explore the trends of top experts and provide an opportunity for the entire research community to identify leading experts in various research fields within Yemen or even within research institutions [13]. Table 12 shows the approved indicators in Yemen research[30].

Table 12. Rating indicators adopted in Yemen research.

On	Attribute(Ranking Indicator)	Description
1	H INDEX	It is known that the H index measures research productivity at the level of the researcher (it is calculated by determining the number of H publications that have been cited at least H times).
2	H INDEX Last 5 Year	
3	CITATION	Measure from research impact 20 for the metric of QS. The indicator is calculated by the citation count for six years for papers published over five years.
4	CITATION Last 5 Year	
5	i10 INDEX	The I10 index measures the number of academic publications produced by the researcher and cited with a minimum of ten citations.
6	i10 INDEX Last 5 Year	
7	Student Ratio	Student and faculty ratios measure teaching commitment. 20% for the metric of QS; international faculty and student ratio: 10% for the metric of QS. The indicator calculates the ratio of new students to overall students.

III. Investigating the Rankings of Yemeni universities

Aligning practices with the requirements of international rankings can be a strategic approach for Yemeni universities to improve their positions and enhance their academic reputation. Here are some strategies that

[29] http://nturanking.csti.tw/ranking/ByContinentCountry?

[30] https://ieeexplore.ieee.org/document/10293385

Yemeni universities can consider: Research Productivity, Faculty Quality and Development, International Collaboration, Student Success and Employability, Infrastructure and Facilities, Quality Assurance and Accreditation, Visibility and Outreach, and Continuous Improvement. It is important to note that rankings should not be the sole focus of Yemeni universities. The ultimate goal should be to provide quality education, advance knowledge, and contribute to societal development. Aligning practices with ranking requirements should be seen as a means to enhance overall academic excellence and reputation, rather than an end in itself.

1- Quacquarelli Symonds (QS)

According to the specified text you provided, it appears that there are two Yemeni universities listed in the QS World University Rankings. The **University of Science and Technology**, Yemen, ranks 151–170, while **Thamar University** in Thamar, Yemen, ranks 171-200. This information is taken from the current web page, which has changed to a page dedicated to university rankings.

2- SCImago Institutions Rankings (SIR)

According to the current web page context, the ranking of Yemeni universities for the year 2023 according to the SCImago Institutions Rankings (SIR) criteria is as follows:

- **Ibb University** is the first ranked university in Yemen, with a total score of **15.4** and a world rank of **3130**. It has the third highest score in research performance among the Yemeni universities.
- **Sana'a University** is the second ranked university in Yemen, with a total score of **14.9** and a world rank of **3190**. It has the highest scores in research performance and societal impact among the Yemeni universities.
- **Taiz University** is the third ranked university in Yemen, with a total score of **14.2** and a world rank of **3251**. It has the second highest score in research performance among the Yemeni universities.
- **Hadhramout University** is the fourth ranked university in Yemen, with a total score of **13.8** and a world rank of **3296**. It has the second highest score in innovation outputs among the Yemeni universities.
- **University of Aden** is the fifth ranked university in Yemen, with a total score of **13.5** and a world rank of **3327**. It has the highest score in innovation outputs among the Yemeni universities.
- **Hodeidah University** is the sixth ranked university in Yemen, with a total score of **12.9** and a world rank of **3404**. It has the fourth highest score in research performance among the Yemeni universities.

- **Thamar University** is the seventh ranked university in Yemen, with a total score of **12.3** and a world rank of **3479**. It has the third highest score in innovation outputs among the Yemeni universities.
- **University of Science and Technology (Yemen)** is the eighth ranked university in Yemen, with a total score of **11.6** and a world rank of **3568**. It has the fifth highest score in research performance among the Yemeni universities.

Table 13. Sample rankings list of Yemeni universities using SIR

No	University name	University in Arabic	total score	Rankings 2023
1	**Ibb University**	**جـامعـة إب**	**15.4**	**1**
2	**Sana'a University**	**جـامعـة صنعـاء**	**14.9**	**2**
3	**Taiz University**	**جـامعـة تـعـز**	**14.2**	3
4	**Hadhramout University**	**جـامعـة حضرمـوت**	**13.8**	4
5	**University of Aden**	**جـامعـة عـدن**	**13.5**	5
6	**Hodeidah University**	**جـامعـة الحديدة**	**12.9**	6
7	**Thamar University**	**جـامعـة ذمـار**	**12.3**	7
8	**University of Science and Technology**	**جـامعـة العلوم والتكنولوجيا**	**11.6**	8

3- Webometrics

According to the current web page context, the top Ten Yemeni Universities according to the webometrics criteria are:

- **Queen Arwa University** is ranked **9788** in the world and has the best scores in size, visibility, and rich files among the Yemeni universities.
- **University of Science and Technology Yemen Aden** is ranked **3701** in the world and has the best score in scholar among the Yemeni universities.
- **Future University Yemen** is ranked **13264** in the world and has the second best score in visibility among the Yemeni universities.
- **Al Razi University Yemen** is ranked **10944** in the world and has the third best score in visibility among the Yemeni universities.
- **Al Nasser University** is ranked **15275** in the world and has the fourth best score in visibility among the Yemeni universities.
- **University of Aden** is ranked **5084** in the world and has the second best score in scholar among the Yemeni universities.

- **Sana'a University** is ranked **3897** in the world and has the third best score in scholar among the Yemeni universities.
- **Al Ahgaff University** is ranked **17568** in the world and has a good score in visibility among the Yemeni universities.
- **Saba University** is ranked **9000** in the world and has a good score in visibility among the Yemeni universities.
- **University of Science and Technology Yemen Sana'a** is ranked **15914** in the world and has a good score in visibility among the Yemeni universities.

Table 13. Sample rankings list of Yemeni universities using webometrics.

No	University name	University in Arabic	total score	Rankings 2023
1	**Queen Arwa University**	**جـامعـة الملكة أروى**	**15.4**	**1**
2	**University of Science and Technology**	**جـامعـة العلوم والتكنولوجيا**	**14.9**	**2**
3	**Future University**	**جـامعـة المستقبل**	**14.2**	3
4	**Al Razi University**	**جـامعـة الرازي**	**13.8**	4
5	**Al Nasser University**	**جـامعـة الناصر**	**13.5**	5
6	**University of Aden**	**جـامعـة عدن**	**12.9**	6
7	**Sana'a University**	**جـامعـة صنعاء**	**12.3**	7
8	**Al Ahgaff University**	**جـامعـة حضرموت**	**13.8**	4
9	**University**	**جـامعـة الحديدة**	**12.9**	6
10	**University**	**جـامعـة تـعـز**	**14.2**	3

4- Times Higher Education (THE)

Yemeni universities are not available in this Ranking.

5- Academic Ranking of World Universities (ARWU)

Yemeni universities are not available in this Ranking.

6- AD Scientific

They based ranking system on the number of meritorious scientists. Four criteria are used to rank the countries.

Table 14. Sample rankings list of Yemeni universities using AD Scientific

No	University name	University in Arabic	Rankings 2023
1	**Sana'a University**	**جـامعـة صنعاء**	**1**
2	**Taiz University**	**جـامعـة تعز**	**2**
3	**Thamar University**	**جـامعـة ذمار**	3
4	**Hadhramout University**	**جـامعـة حضـرمـوت**	4
5	**Ibb University**	**جـامعـة إب**	5
6	**University of Aden**	**جـامعـة عدن**	6
7	**Hodeidah University**	**جـامعـة الحديدة**	7
8	**Colleges Sana'a**	**كلية المجتمع صنعاء**	8
9	**Queen Arwa University**	**جـامعـة الملكة أروى**	9
10	**Azal University of Human Development**	**جـامعـة ازال**	10

7- EduRank

Here are the top 10 universities in Yemen ranked by EduRank based on the adopted indicators. According to the analysis, Sana'a University takes the top spot in the ranking, followed by the University of Aden and the University of Science and Technology - Yemen. These universities have a long history of establishment and enjoy widespread recognition in Yemen and globally[31].

Table 15. Sample rankings list of Yemeni universities using EduRank

No	University name	University in Arabic	Rankings 2023
1	**Sana'a University**	**جـامعـة صنعاء**	**1**
2	**University of Aden**	**جـامعـة عدن**	**2**
3	**University of Science and Technology Yemen**	**جـامعـة العلوم والتكنولوجيا**	3
4	**Taiz University**	**جـامعـة تـعـز**	4
5	**Thamar University**	**جـامعـة ذمار**	5
6	**Hodeidah University**	**جـامعـة الحديدة**	6
7	**Hadhramout University**	**جـامعـة حضـرمـوت**	7
8	**Ibb University**	**جـامعـة إب**	8
9	**Queen Arwa University**	**جـامعـة الملكة أروى**	9
10	**Al Razi University**	**جـامعـة الرازي**	10

[31] https://edurank.org/geo/ye/

8- uniRank

uniRank University Ranking lists the top 10 Yemeni higher education institutions that meet the uniRank selection criteria. uniRank aims to provide a non-academic League Table of the top Yemeni universities and colleges based on valid, unbiased and non-influenceable web metrics provided by independent web intelligence sources, rather than data submitted by the universities themselves.

Table 16 Sample rankings list of Yemeni universities using uniRank

No	University name	University in Arabic	Rankings 2023
1	University of Science and Technology Yemen	جـامعـة العلوم والتكنولوجيا	1
2	Queen Arwa University	جـامعـة الملكة أروى	2
3	Sana'a University	جـامعـة صنعاء	3
4	Lebanese International University Yemen	الجامعة اللبنانية الدولية	4
5	Al Razi University	جـامعـة الرازي	5
6	Al-Ahgaff University	جـامعـة الاحقاف	6
7	Emirates International University	جـامعـة الإماراتية الدولية	7
8	Hadhramout University	جـامعـة حضرموت	8
9	Saba University	جـامعـة سباء	9
10	University of Aden	جـامعـة عدن	10

9- CWUR

Yemeni universities are not available in this Ranking.

10- Leiden Ranking

Yemeni universities are not available in this Ranking.

11- NTU Ranking

Yemeni universities are not available in this Ranking.

12- Yemen research

The researchers extracted study data from the Information Technology Center at the Ministry of Higher Education and Scientific Research in Yemen, as well as the "AD SCIENTIFIC INDEX" world ranking dataset, to identify the most important indicators for ranking Yemeni universities. The research identified seven indicators for 78 Yemeni universities, both governmental and private, and used cluster analysis to reveal the best performers. The results showed that Sana'a University ranked first, followed by the University of Aden. The study recommends that Yemeni universities align their goals with the most important indicators based on

international strategies and support scientific research through the support of scholars. Table 17 provides information about the clusters to which each university belongs, as well as their respective rankings for the year 2023. Universities like **Sana'a University** and **Aden University**, belonging to Cluster 0, occupy the top positions. Similarly, universities in Clusters 1, 2, and 3 also have notable rankings within the Yemeni context.

Table 17. Sample rankings list of Yemeni universities using Yemen research.

No	University name	University in Arabic	Cluster	Rankings 2023
1	Sana'a University	جامعة صنعاء	0	1
2	Aden University	جامعة عدن	0	2
3	Ibb University	جامعة إب	1	3
4	Hodeidah University	جامعة الحديدة	1	4
5	Hadhramout University	جامعة حضرموت	1	5
6	Sana'a Community College	كلية المجتمع صنعاء	1	6
7	University of Science and Technology	جامعة العلوم والتكنولوجيا	1	7
8	Taiz University	جامعة تعز	2	8
9	Thamar University	جامعة ذمار	2	9
10	Al-Razi University	جامعة الرازي	3	10

IV. conclusion

The study focuses on Yemeni universities and their potential to enter and improve their positions in international rankings. The research examines five prominent international rankings: **QS, ARWU, WEBOMETRICS, THE,** and **SIR**. The objective of the study is to encourage Yemeni universities to prioritize the criteria outlined in these rankings and consider them as a pathway to access other international rankings. The research addresses the challenge of understanding the criteria and application process required for Yemeni universities to participate in these rankings, as well as the rankings they have achieved. The study emphasizes the importance of providing Yemeni university administrators with valuable insights into the criteria used for international rankings and the rankings obtained by Yemeni universities. The findings of the study reveal that two Yemeni universities are listed in the **QS**. The **University of Science and Technology**, Yemen, ranks 151-170, while **Thamar University** in Thamar, Yemen, ranks 171-200.

Additionally, the findings of the study reveal the ranking positions of two Yemeni universities in the **SCImago** Institutions Rankings (**SIR**) for 2023:

- **Ibb University** secures the highest position among Yemeni universities with an overall score of 15.4 and a world rank of 3130.

- **Sana'a University** closely follows , ranking second among Yemeni universities with a total score of 14.9 and a world rank of 3190.

The study highlights the importance of international rankings for Yemeni universities and their potential to improve their positions. **Ibb University** and **Sana'a University** have achieved notable rankings in the SCImago Institutions Rankings for 2023 among Yemeni universities. Yemeni university administrators can use the insights provided by the study to understand the criteria used for rankings and work towards enhancing the quality of higher education in their institutions. The findings of the study offer Yemeni university leadership a valuable benchmark to evaluate their institutions' global performance. These rankings shed light on the current positions of the universities and indicate areas where improvements can be made.

University administrators can carefully examine the ranking criteria and identify specific areas that require enhancement to elevate the quality of higher education within their institutions. By gaining an understanding of the methodologies and indicators utilized in these rankings, leadership can align their strategic goals and initiatives to address the identified areas of improvement. This may involve measures such as boosting research output, enhancing faculty quality, improving student satisfaction, fostering international collaborations, or focusing on other factors that contribute to higher rankings. The study's findings provide valuable insights into the strengths and weaknesses of Yemeni universities compared to their global counterparts, empowering leadership to make informed decisions and implement strategies aimed at enhancing their institutions' overall standing.

Regarding potential future work, some suggestions are:

- Conducting a comparative analysis of the ranking criteria and methodologies to identify common areas for improvement across different rankings.

- Investigating the specific strengths and weaknesses of Ibb University and Sanaa University in relation to the criteria used in the SCImago Institutions Rankings and developing strategies to further enhance their performance.

- Exploring the impact of improved rankings on the reputation, funding opportunities, and international collaborations of Yemeni universities.

- Conducting surveys or interviews with Yemeni university stakeholders to gather their perspectives on the rankings and their potential implications for the higher education sector in Yemen.

These suggestions can help expand on the study's findings and contribute to the ongoing efforts to improve the rankings and overall quality of Yemeni universities. The original approach in ranking Yemeni institutions is crucial as it assists scholars and students in finding educational institutions that align with their specific academic needs.

References

1. Ahmed, M. R. A. (2025). Accreditation and Quality Assurance: Exploring Impact and Assessing Institutional Change in the US and Saudi Arabian Higher Education Institutions. *Arts for Linguistic & Literary Studies, 7*(1), 626–639. https://doi.org/10.53286/arts.v7i1.2419
2. Alshammari, A., Alshammari, M., & Alshammari, T. (2022). University Rankings and Goals: A Cluster Analysis. In J. Kacprzyk & L. C. Jain (Eds.), Intelligent Systems Reference Library (pp. 1-15). Springer Nature Switzerland AG.
3. Antoniuk, L., Kalenyuk, I., Tsyrkun, O., & Sandul, M. (2019). Rankings in the higher education competitiveness management system. Problems and Perspectives in Management, 17(4), 325-339. doi:10.21511/ppm.17(4).2019.27
4. Chowdhury, A. R., & Rahman, Z. (2021). Global ranking framework & indicators of higher educational institutions: A comparative study. Library Philosophy and Practice (e-journal), 5268.
5. Cunningham, S. D. (2008). University and discipline cluster ranking systems and the humanities, arts and social sciences21. Higher Education in Europe, 33(2), 245-258.
6. Das, S. S., Balasubramanian, P., & Chowdhury, A. R. (2019). Webometrics ranking (WR) of world universities and national institutional ranking framework (NIRF): A comparative study. Journal of Information and Knowledge, 154-158.
7. Gabriel, A., & Aribisala, B. (2020). Automatic ranking of tertiary institutions in developing nations (a case study of Nigeria universities).
8. Ganga-Contreras, F., Suarez-Amaya, W., Valderrama-Hidalgo, C., & Salazar-Botello, C. (2022). University rankings and influence on performance: Comparing perspectives in Chile and Venezuela. RETOS. Revista de Ciencias de la A ministración y Economía, 12(24), 200-210.
9. Gopal, K., & Shitan, M. (2015, May). Cluster analysis of top 200 universities in Mathematics. In 2015 International Symposium on Mathematical Sciences and Computing Research (iSMSC) (pp. 408-413). IEEE.
10. Henao-Rodríguez, C., Lis-Gutiérrez, J. P., Gaitán-Angulo, M., Vásquez, C., Torres, M., & Viloria, A. (2019). Determinants of ResearchGate (RG) Score for the Top100 of Latin American Universities at Webometrics1. In Y. Tan & Y. Shi (Eds.), Data Mining and Big Data (pp. 320-327). Springer Nature Singapore Pte Ltd.
11. Moskovkin, V.M., Pupynina, E.V., Zaitseva, N.P., & Lesovik, R.V. (2013). Methodology for comparative analysis of university rankings, with the Mediterranean and Black Sea region countries taken as an example. Middle-East Journal of Scientific Research, 18(11), 1656-1665.
12. Mussard, M., & James, A. P. (2018). Engineering the global university rankings: Gold standards, limitations and implications1. IEEE Access, 6, 6765-6775.
13. Polyakov, M., Bilozubenko, V., Korneyev, M., & Nebaba, N. (2020). Analysis of key university leadership factors based on their international rankings (QS World University Rankings and Times Higher Education). Problems and Perspectives in Management, 18(4), 142-152. doi:10.21511/ppm.18(4).2020.13.

14. Rafique, T., Awan, M. U., Shafiq, M., & Mahmood, K. (2023). Exploring the role of ranking systems towards university performance improvement: A focus group-based study. Heliyon, 9(10).
15. Salah Al-Hagree; Ayedh Abdulaziz Mohsen et al, Universities the Best Performers: A Cluster Analysis of Yemeni Universities Rankings, 2023 3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA).
16. Sasvári, P., & Urbanovics, A. (2022). Subject Area Risk Assessment of Four Hungarian Universities with a View to the QS University Rankings by Subject. Journal of Data and Information Science, 7(4), 61-80.
17. Shahjahan, R. A., Sonneveldt, E. L., Estera, A. L., & Bae, S. (2020). Emoscapes and commercial university rankers: the role of affect in global higher education policy. Critical Studies in Education. https://doi.org/10.1080/17508487.2020.1748078

مجلــة العلــوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

A Literature Review on Arabic Automatic Question Generation

Abdulkhaleq Amin Abdullah[1,2], Khaled A. Al-Soufi[2]

[1]Information Technology, Faculty of Engineering and Information Technology
[2]Qalam University, Yemen
a7q2014@gmail.com, kalsoufi@gmail.com

Abstract:

This comprehensive literature review is dedicated to the field of Arabic Automatic Question Generation (AQG), which focuses on the development of computational models and algorithms for the automatic generation of questions. The review systematically covers key concepts in AQG, including question types and evaluation metrics. Additionally, it delves into the specific challenges associated with applying AQG techniques to the Arabic language, considering factors like complex morphology and dialectal variations. The review introduces a taxonomy of Arabic AQG approaches, classifying them into rule-based, template-based, and machine learning-based methods. It examines the pivotal role of datasets, resources, and evaluation methodologies in the training and assessment of AQG systems. Advancements in Arabic AQG are highlighted, and the review identifies emerging research directions, such as domain-specific question generation and integration into educational platforms.

In conclusion, the review provides valuable insights for researchers, developers, and educators interested in Arabic AQG. It addresses current advancements, challenges, and outlines potential future research directions, including the scarcity of labeled data and the necessity for domain-specific approaches. Overall, this review serves as a comprehensive resource in the realm of Arabic AQG.

Keywords: Question generation, Arabic language, Artificial intelligence, Contexts.

1. Introduction

Arabic Automatic Question Generation (AAQG) is an emerging field of research that focuses on developing systems and techniques for automatically generating questions in the Arabic language. This field is of great importance in various educational applications as it can be used to create language learning materials, assess comprehension skills, and enhance the overall learning experience for Arabic learners.

Automatic Question Generation is of great significance for the Arabic language for several reasons. Firstly, the Arabic language is one of the most widely spoken languages globally, with millions of Arabic speakers around the world. Secondly, the Arabic language has a rich literary and intellectual heritage, making it an important language for educational purposes. Furthermore, the increasing availability of Arabic educational content on the internet necessitates the development of efficient and effective methods for generating relevant and accurate questions to enhance learning outcomes for Arabic-speaking students [1].

Question answering (QA) can benefit from QG, where the training dataset of QA can be enriched using QG to improve the learning and performance of QA algorithms [2]. Additionally, in the future, QG can be greatly helpful to the education industry by generating teaching materials with less effort and motivating students' intention to read.

Arabic question-generation systems have been limited and are mainly focused on factoid questions and short-answer generation [3]. Current approaches rely heavily on rule-based methods or manual construction of question styles using specific text sources, which lack scalability and generalizability [4]. The complexity of the Arabic language presents challenges in morphological analysis, syntactic structures, and semantic nuances (Ahmed, 2025, Alharbi, 2022)[5][6]. Addressing these challenges requires the creation of large and diverse Arabic question-generation datasets, as well as the exploration of advanced Natural Language Processing techniques specifically tailored for the Arabic language [5]. Deep learning models have demonstrated their effectiveness in overcoming these obstacles by leveraging extensive Arabic datasets and educational resources to comprehend the intricacies of the language (Al-Ghobesi, 2025)[7].

To develop an effective Automatic Question Generation system for the Arabic language, it is imperative to adopt strategies that leverage recent advancements in deep learning techniques, particularly those that have been successful in the development of AQG systems for other languages [8]. Recent research has demonstrated the efficacy of Transformer-based models and recurrent neural networks in generating questions from textual sources [5][9][10]. These models have exhibited a high degree of accuracy when applied to languages with complex syntactic and semantic structures, making them viable candidates for developing Arabic question-generation systems [11].

The field of educational question generation has made significant advancements, particularly with the integration of deep learning techniques [12]. This attempts to create questions from a text paragraph, where certain sub-spans of the passage in question will answer the questions produced. Artificial intelligence (AI) techniques are now extensively used to minimize the manual effort required by teachers, and research progress has been made in Automatic Question Generation (AQG) from textual data based on syntax and semantics [13].

The integration of question-generation technology in education has the potential to greatly enhance the learning experience by automatically generating thought-provoking questions [14]. This advancement relies on deep learning techniques, which have shown promise in capturing complex patterns and semantic connections within Arabic QG systems used in educational contexts.

The review has made significant contributions, which are listed below:

-The current state of research on generating Arabic questions.

-What existing methods for creating Arabic questions could be adjusted to fulfill the specific demands of education.

-Strategies for future research to address deficiencies and fill gaps in educational Arabic question generation.

This paper is organized as follows: We first formally represent the problem of Arabic automatic question generation and discuss the various question categories, providing a technical overview of such a system. In section 2, an overview of previous works on Arabic question generation is provided. Section 3 presents the methodologies and techniques employed in Arabic question generation using deep learning, including Transformer-based models and recurrent neural networks. In section 4, we discuss the performance evaluation metrics used to assess the quality of generated questions and compare different approaches. Section 5 presented the challenges and limitations faced in Arabic question generation, such as the lack of large-scale annotated datasets and the complexity of Arabic language structures. Future works are discussed in section 6. Finally, the conclusion is presented in section 7.

1.1. ARABIC LANGUAGE Challenges

Arabic is a Semitic language known for its rich morphology, complex grammatical structures, and varied dialects. It is spoken by hundreds of millions of people worldwide and is the official language of over 20 countries [15]. Although Arabic is popular, work on the Arabic language is still limited, especially in QG [16]. To understand the challenges of the Arabic language, the basics of Arabic should be understood.

Given the complexity of the Arabic language, it presents unique challenges in the field of Natural Language

Processing. One of the notable challenges is the shortage of large datasets and limited tools for tasks such as sentiment analysis and question generation [17]. Although deep learning techniques have shown promising results in addressing these challenges for other languages, there is a lack of research specifically focusing on deep learning methods for the Arabic language. The intricate morphology and structural complexity of Arabic create obstacles that necessitate the adaptation and customization of existing deep learning models to meet the specific requirements of Arabic NLP [18].

1.2. Classification of question generation:

The question generation problem involves finding a model that approximates the generated question. The dataset is pre-processed to make the data available in the desired format, and an appropriate strategy is employed based on the question type for generating questions [19]. Depending on the type of system being used, either text or image datasets can be selected [20]. Categorization aids in outlining specific use cases, and the following is a classification of questions based on existing research in question generation.

1.2.1. Factual questions:

Factual questions aim to elicit specific information or facts. These questions typically start with words like "من who," "ماذا/ما what," "أين where," and "متى when." Examples of factual questions in an educational context could be: "ماهي عاصمة اليمن؟ What is the capital of Yemen?"

1.2.2. Inferential questions:

Inferential questions require the reader to analyze information and draw conclusions based on the given context. These questions often start with words like "لِمَاذا why," "كيف how," or phrases like "بماذا تستنتج؟ What can you infer from this?

1.2.3. Multiple sentences spanning questions:

Some questions may necessitate answers consisting of multiple sentences from a paragraph, as the relevant information is distributed across several sentences. These types of questions are typically W4H (What/Where/When/Why/How) and can be addressed using methods similar to those employed for answering factual questions.

1.2.4. Yes/no type questions:

Yes/no type questions aim to elicit a binary response of "yes" or "no." Examples of yes/no type questions in an educational context could be: "هل الشمس تدور حول الأرض ؟ Does the sun revolve around the Earth?".

2. Literature Review

The recent work on Arabic QA has focused on leveraging the advancements made in English QA, particularly the utilization of deep learning models such as BERT and recurrent neural networks [21]. These

models have shown promise in improving the performance of Arabic QG systems by capturing contextual information and semantic relationships [22]. These models have shown promising results in generating questions from Arabic healthcare texts based on word embeddings. However, there is still a lack of research on Arabic QG specifically for educational purposes.

In addition to applying advancements from English QA to Arabic QA, recent studies have also demonstrated the effectiveness of deep learning in the domain of Arabic text classification. Techniques such as word embeddings and deep learning models have shown promising classification accuracy, highlighting their potential for use in educational applications. Furthermore, the integration of deep learning in Arabic QG can significantly enhance the automatic generation of thought-provoking questions by capturing complex patterns and semantic connections within the educational context [23].

The literature on Arabic Question Generation using Deep Learning in educational contexts reveals a comprehensive landscape of research endeavors focused on advancing the integration of DL techniques for QG in Arabic language. Numerous studies have explored the utilization of recurrent neural networks, convolutional neural networks, and transformer models in the domain of Arabic QG, emphasizing the potential of these DL architectures in addressing the complexities inherent to Arabic language.

A study proposes a method to analyze the question and retrieve the passage answer in Arabic language, along with experimenting with the generation of a logical representation from the declarative form of each question [11]. This includes extensive evaluation of various Arabic language tasks such as Sentiment Analysis, Named Entity Recognition, and QA using different datasets. Additionally, it discusses advancements in word vector representations for Arabic compared to English models like Word2Vec and highlights initiatives such as AraVec which provides powerful word embedding models developed specifically for Arabic NLP through pre-trained word representations sourced from different domains like Wikipedia, World Wide Web pages, and tweets. Furthermore, it describes real-time embedding schemes that aim at improving performance in various NLP tasks without relying on tabulated pre-embedding or pre-trained transfer learning models [24].

In 2015, [26] introduced rule-based methods that utilized rule-based knowledge sources to assess the comprehension of crucial domain rules. Our interpretation aligns rule-based approaches with semantic-based methods due to their requirement for a deeper understanding beyond mere syntax. Regarding the fifth category, schemas are similar to templates but more abstract in nature. They group templates that represent variations of the same problem. However, we find the differentiation between template and schemas to be unclear.

In 2010, [27] developed a method for creating definitional queries, which are inquiries that seek the meaning of a chosen term referred to as an "up-key". The identification of the "up-key" can be achieved through statistical approaches such as term frequency. Once the named entity of the "up-key" is identified, a question word can be selected. This "up-key" and question word are then inserted into the template "<Question word> is <up-key>?" to generate a definitional query.

In 2010, [28] proposed an approach to automatically generate questions to support students in writing literature reviews. The approach generates questions e.g. What is the research question framed by X? The approach automatically extracts and classifies citations from a student's review. The approach uses templates to generate questions based on the extracted information.

In 2018, [8] presented a Corpus-Based Approach that utilizes a large collection of texts, known as a corpus, to automatically generate questions. This approach involves analyzing the corpus to identify patterns and generate questions based on the content found in the texts. The generated questions are based on patterns and structures observed in the corpus. The corpus-based approach for automatic question generation relies on analyzing a large collection of texts to identify patterns and generate questions. These questions are generated based on patterns and structures observed in the analyzed corpus, allowing for a more data-driven approach to question generation.

In 2014, Google introduced a text-generation task defined as a sequence-by-sequence task. This approach facilitates end-to-end learning for tasks that involve input and output sequences of tokens. Encoder-decoder models are frequently applied to these types of problems and typically include an encoder to process the input sequence through layers of recurrent neural networks, along with a decoder designed to generate the output sequence.

The research by [12] focused on creating queries from a text paragraph for machine-reading comprehension using an attention-driven mechanism based on a sequence-to-sequence model and an RNN-based encoder-decoder system with two distinct encoders. These approaches demonstrate the use of templates and corpus-based analysis, as well as the application of sequence-to-sequence models with attention mechanisms for automatic question generation from text.

In 2019, [35] employed transformers to generate questions based on passages from the SQuAD dataset. They used word error rate as a metric to compare the generated questions with the target questions. The authors observed the syntactic correctness of the generated questions and their relevance to the passage. Furthermore, they found that WER was low for shorter questions but increased for longer ones.

In 2022, [5] improved the transformer model for generating questions on the mMARCO dataset and

optimized it using the AdamW optimizer model. They proposed an end-to-end Arabic automatic question generation model based on the Transformer architecture to generate N interrogative questions for educational content from a single unlimited-length document. This constitutes a transfer-learning process. The authors also incorporated TextRank and Sentence Extraction techniques in their model to improve the quality of the generated questions.

3. Categories of Arabic question generation Approaches

Based on the sources reviewed, the categories of Arabic question generation approaches can be classified into the following:

3.1. Rule-Based Approach

A method based on predefined rules is employed to generate questions from declarative sentences [25]. This method simplifies the sentence, applies a transformation technique for question generation, ranks the generated questions using logistic regression to assess their quality, and then labels them for approval. For example, the rule-based approach may involve transforming a declarative sentence like " أحمد ذهب إلى السوق" into a question like "إلى أين ذهب أحمد؟".

3.2. Template-Based Approach

A method based on templates entails the creation of pre-established question templates, which are then populated with pertinent information derived from the input text. These templates act as a framework for generating questions and ensure that the resulting questions adhere to proper grammar while remaining relevant to the content at hand (Ushio et al., 2023).

3.3. Corpus-Based Approach

The Corpus-Based Approach utilizes a framework for automatic question generation that involves analyzing a large corpus of text. This approach extracts patterns and linguistic features from the corpus to generate questions [51].

3.4. Sequence-to-Sequence Approach

One method for generating questions is the sequence-to-sequence approach, where a neural network model is trained to generate questions by treating it as a sequence-generation task. The model takes an input sequence (such as a sentence or passage) and produces an output sequence (the question). This output sequence is trained to match human-generated questions, allowing the model to learn how to generate questions based on the given input [51].

3.5. Transformer-based Approach

A transformer-based approach [29] involves utilizing the Transformer architecture, a popular deep-learning model, for automatic question generation. This approach relies on the use of transformer models, such as the Transformer architecture, to automatically generate questions by analyzing the syntactic and semantic structures of sentences [5].

These different categories of automatic question generation approaches provide a range of methods for generating questions from text, including rule-based approaches, template-based approaches, corpus-based approaches, sequence-to-sequence approaches, and transformer-based approaches.

Furthermore, the use of multilingual and cross-lingual models allows for the transfer of question-generation techniques and models from other languages to Arabic [49]. Overall, the existing research on Arabic question generation using deep learning techniques demonstrates various approaches and models, including template-based methods, sequence-to-sequence models with attention mechanisms, and transformer-based approaches different QG approaches used for Arabic question generation.

Table 1 provides a comparison summary of different approaches for Arabic question generation, including rule-based approaches, template-based approaches, corpus-based approaches, sequence-to-sequence approaches, and transformer-based approaches.

Table 2 presents an analysis of papers that discussed Arabic QG from multiple perspectives.

Table 1 Summary of Arabic QG Approaches

Approach	Description	Advantages	Limitations
Rule-based approaches	Arabic Question generation methods have traditionally relied on pre-established rules and syntactic transformations.	These methods are simple to execute and can produce grammatically accurate questions.	These methodologies heavily depend on manually created rules and templates, which might not encompass all linguistic variations in Arabic.
Template-based approaches	These methods utilize predetermined templates to produce questions by substituting specific placeholders with pertinent details from the given text.	One approach that is relatively simple to implement and can generate questions with correct syntax is the use of templates.	The ability of these methods to handle complex or diverse sentence structures may be limited, and the quality of the produced questions relies heavily on the quality of the templates used.
Corpus-based approaches	These methods leverage a vast collection of pre-existing questions and their corresponding answers to generate novel questions.	Corpus-driven methods can capture a wide range of question patterns and enhance the diversity of generated questions.	Obtaining a sufficient and diverse question-answer corpus for Arabic may pose challenges for these approaches.

Approach	Description	Advantages	Limitations
Sequence-to-sequence approaches	These methods utilize neural networks to generate questions from the input text by considering the process of generation as a prediction task based on sequences.	Neural sequence-to-sequence methods have been found to produce a wider range of questions that are more relevant in context, compared to traditional rule-based or template-based approaches.	These methods might necessitate a substantial amount of training data and could be computationally demanding.
Transformer-based approaches	These methods make use of transformer models, like the Transformer architecture, to generate questions by capturing the connections between input and output sequences.	Recent advancements in natural language processing tasks have demonstrated the effectiveness of Transformer-based approaches, particularly for question generation.	Optimal performance of transformer-based methods may necessitate a substantial amount of computational resources and training data.

Table 2 Presents an analysis of papers that discussed Arabic QG:

Ref.	*Approach*	*Algorithm*	*Scientific additions*
[5]	Transformer-based approaches	The **TextRank** algorithm operates by extracting key sentences from paragraphs and organizing them into a list.	The system creates questions without answers and functions by segmenting the document into paragraphs while identifying key sentences.
[1]	Template-based approaches	It seamlessly incorporates models for semantic role labeling and question generation.	A novel method for generating questions from Arabic text involves the fusion of semantic role labeling (SRL) for capturing linguistic relationships and the flexibility inherent in question models.
[22]	Transformer-based approaches	ARAGPT2 is a stacked transformer-decoder model that has undergone training using the causal language modeling objective.	ARAGPT2-MEGA, with its impressive 1.46 billion parameters, stands as the most extensive Arabic language model currently accessible. This model, AraGPT2-mega, demonstrates remarkable proficiency in generating news articles that pose a challenge in distinguishing them from those crafted by human authors.

Ref.	*Approach*	*Algorithm*	*Scientific additions*
[30]	Transformer-based approaches	Fine-tuning AraT5 model on ARGENQG dataset.	The introduction involves three potent variants of the T5 model specifically designed for Arabic language generation. Furthermore, a groundbreaking proposition is made for a novel benchmark in Arabic natural language generation, denoted as ARGEN, comprising seven distinct tasks.
[9]	Transformer-based approaches	Fine-tuning GPT-2 architecture	The training of GPT-2 for the purpose of generating Arabic poems has resulted in superior performance, surpassing existing models in the realm of Arabic poetry generation.
[31]	Transformer-based approaches	The authors put forward a multilingual data-driven approach for the generation of reading comprehension questions, employing dependency trees in the process.	The assumed position for the question word in Arabic is the first processed word, essentially representing the end of the sentence in Arabic sentence structure.
[34]	Visual Arabic Question Answering (VAQA)	The envisioned system comprises five key modules: visual feature extraction, question preprocessing, textual feature extraction, feature fusion, and answer prediction.	The VQA task is conceptualized as a binary classification problem.

4. Datasets

The development and assessment of natural language processing models rely on the availability and quality of Arabic question generation datasets. However, obtaining such datasets for Arabic language poses challenges due to factors such as limited resources, the absence of standardized datasets, and the intricate linguistic nature of Arabic. This scarcity poses a notable challenge for researchers and developers aiming to enhance Arabic natural language processing models. Addressing this issue by creating extensive Arabic question-generation datasets is essential [1].

To address the scarcity of Arabic question generation datasets, recent translation techniques have been applied to crowd-sourced annotated datasets. This approach has produced reasonable results on training data for different languages, enabling researchers to overcome the lack of datasets by translating existing ones into Arabic [1]. For instance, the SQuAD dataset has been translated into Arabic, creating a valuable resource for Arabic question generation [38].

Although the availability of specific Arabic question generation datasets may still be limited compared to English, several datasets have been developed for this purpose [39]. These datasets, such as the translated SQuAD dataset and the Arabic Question Answering Web Corpus, provide valuable resources for training and evaluating Arabic question generation models, enabling the development of more accurate and robust systems for generating questions in Arabic.

THE DETAILS OF THE DATASETS USED FOR QUESTION GENERATION ARE SUMMARIZED IN TABLE 3.

Table 3: *Summary of Arabic QG datasets*

DATASET NAME	**Source**	**Content**
ARABIC SQUAD: Arabic Stanford Question Answering Dataset [38].	Translated from English	Translated into Arabic, providing a valuable resource for Arabic question generation.
ARCD: Arabic Reading Comprehension Dataset [38].	Custom	Crowdsourced Arabic questions based on the CNN and Daily Mail datasets in English.
MLQA: Multilingual Question Answering [40].	Multilingual	Parallel sentences in 7 languages, including Arabic, for machine translation evaluation.
XQUAD: Cross-Lingual Question Answering Dataset [41].	Multilingual	Multilingual dataset for question-answering tasks, including Arabic.
TYDI QA: Typologically Diverse Question Answering [41].	Multilingual	The data set is a question-answering benchmark based on Wikipedia articles for 11 typologically diverse languages. Eight of these languages have available UD treebanks and trained dependency parsers in the Stanza package
LAREQA: Low-Resource Cross-Language Question Answering [42].	Custom	A dataset for cross-lingual question answering in low-resource languages, including Arabic.
DAWQAS: Dataset For Web-Based Question Answering In Arabic Script [43].	Custom	A dataset for Arabic web-based question answering.
EXAMS [44].	Custom	A dataset for evaluating extractive question-answering models, including Arabic passages.
MS MARCO [5].	Custom	The MMARCO dataset is a machine-translated multilingual version of the MS MARCO passage ranking dataset covering 13 typologically diverse languages [45].
QUIZITO'S Platform [1].	Custom	The collection comprises 40,435 questions gathered from kids' books and summaries manually curated by Quizito.
Image-Question-Answer (IQA) Triplet [43].	VAQA	The VAQA dataset contains 5000 images. The first Visual Arabic Question Answering (VAQA) dataset is generated. first dataset and system for VQA in Arabic.
ARGENQG [30]	Custom	The researcher built ARGENQG by extracting 96K (passage, answer, and question) triplets from the ARCD dataset, and three multi-lingual QA datasets: MLQA, XQuAD, and TyDi QA.

4. Evaluation Metrics

To evaluate Arabic question generation, several metrics can be used to assess the quality and effectiveness of the generated questions. Some commonly used metrics for automatic evaluation include BLEU-4, METEOR, and ROUGE-L. These metrics compare the generated questions with reference questions to measure the similarity and overlap in terms of n-grams and recall scores. Apart from these automatic metrics, human evaluation can also be conducted to gather subjective feedback on the quality and comprehensibility of the generated questions [3]. The commonly used similarity metrics for evaluating Arabic question generation are as follows:

4.1. BLEU Score:

This metric measures the similarity between the generated questions and reference questions based on n-gram overlap. BLEU-4 is a commonly used automatic evaluation metric in question generation tasks. It measures the n-gram similarity between the generated questions and the reference questions. The BLEU-4 metric calculates the precision of n-grams up to four words in length, comparing the generated questions with the reference questions [46].

4.2. ROUGE Score:

This metric evaluates the quality of the generated questions by comparing them with reference questions based on n-gram overlap and word order. ROUGE-L is a metric specifically designed for evaluating the quality of summaries or generated questions. It calculates the recall score by comparing the generated questions with the reference questions based on the longest common subsequence between them. This metric is particularly useful in assessing the quality of generated questions when they may have different complexities than the input questions. ROUGE-L is a metric that measures the recall-oriented evaluation of generated questions by comparing them with reference questions [47].

4.3. METEOR Score:

This metric takes into account various linguistic aspects such as synonyms, stemming, and word order to measure the quality of the generated questions. METEOR is another automatic evaluation metric commonly used in question generation. METEOR is a metric that combines precision and recall to measure the quality of generated questions compared to the reference questions. It takes into account not only the overlap in n-grams but also considers synonyms and paraphrases to evaluate the semantic similarity between the generated questions and the reference questions [48].

4.4. CIDEr

CIDEr, also known as Consensus-based Image Description Evaluation, is an automated metric that was

developed to assess the quality of image descriptions. This metric compares a sentence generated by a model with a set of human-written sentences considered as ground truth. It measures how similar the model-generated sentence is to the consensus among the ground truth sentences for that particular image. The concept of consensus in CIDEr refers to how often most of the sentences used to describe an image are similar. Additionally, CIDEr claims that its metric inherently captures aspects such as grammar, importance, accuracy, and saliency [50].

4.5. Human Evaluation:

In human evaluation studies, human evaluators assess the quality of the generated questions based on criteria such as relevance, syntactic correctness, ambiguity, and difficulty. They evaluate if the generated questions are relevant to the given information, if they are grammatically correct if there is any ambiguity in the wording, and if the difficulty level of the questions aligns with the intended target audience. Additionally, the evaluation criteria used in human evaluation studies provide guidelines for the evaluators and help assess the performance of each template. The human evaluators rate the difficulty level of the questions to ensure that there is a syntactic divergence between the input sentence and the generated question [8].

In summary, the evaluation of Arabic question generation can be done using a combination of automatic metrics and human evaluation studies. These evaluation methods provide a comprehensive analysis of the generated questions, taking into account both quantitative metrics for similarity and quality, as well as qualitative assessments by human evaluators, as well as human evaluation studies to assess the relevance, syntactic correctness, ambiguity, and difficulty of the generated questions.

In Table 4, we list the different evaluation metrics used for question generation.

Table 4 Summery of Evaluation metrics and results

Ref.	*BLEU-4*	*ROUGE*	METEOR	CIDEr	f-measure
[5]	19.12	23.00	51.99	-	-
[1]	-	-	-	-	86%
[22]	-	-	-	-	98.7%
[30]	16.99	-	-	-	-
[9]	0.187	-	-	-	-
[31]	2.90	13.12	24.69	22.70	-
[34]	-	-	-	-	84.94%

5.Arabic Question Generation Challenges:

When it comes to Arabic question generation, several unique challenges need to be considered. These challenges include the complexity of Arabic language, which has different grammatical rules and structures

compared to other languages. Additionally, Arabic Question Generation faces several unique challenges that need to be addressed for effective implementation [1]:

5.1. Lack of Arabic-specific datasets:

One of the major challenges in Arabic question generation is the scarcity of high-quality datasets specific to the Arabic language. The availability of comprehensive datasets is crucial for training and evaluating question generation models effectively [22].

5.2. Limited linguistic resources:

Arabic question generation requires robust linguistic resources, including morphological analyzers, part-of-speech taggers, and syntactic parsers. These resources are essential for accurately understanding the semantics and syntax of Arabic text and generating grammatically correct and meaningful questions [1].

5.3. Inconsistencies in discretization:

Arabic words can have different meanings depending on the placement of diacritical marks. This creates challenges in accurately representing the intended meaning of words and generating questions that capture the desired information [36].

5.4. Complex sentence structures:

Arabic sentences often follow different syntactic structures compared to other languages. These complex sentence structures pose a challenge for question-generation models, as they need to understand and parse the sentences correctly to generate relevant and accurate questions [3].

5.5. Lack of domain-specific knowledge:

Arabic question generation for educational purposes requires a deep understanding of the subject matter and the ability to generate contextually relevant questions that align with the educational content [4].

6. Future work

To address these challenges and improve Arabic question generation, future research can focus on the following areas:

6.1. Development of Arabic-specific datasets:

Developing high-quality datasets dedicated for generating Arabic questions is crucial for the effective training and evaluation of models. These datasets should cover a wide range of diverse topics and domains, including educational content, to ensure the generation of contextually relevant questions [38].

6.2. Integration of advanced linguistic resources:

Research efforts can focus on improving the availability and quality of linguistic resources for Arabic

question generation. This can include developing more accurate morphological analyzers, part-of-speech taggers, and syntactic parsers specifically tailored for Arabic language processing [3].

6.3. Application of transfer learning:

Investigating transfer learning techniques provides an avenue to utilize pre-existing question-generation models trained in different languages and modify them for application in Arabic. This strategy has the potential to address the constraints posed by the scarcity of Arabic linguistic resources, thereby enhancing the effectiveness of models dedicated to generating questions in Arabic [4].

6.4. Designing hybrid models:

The development of hybrid models, integrating rule-based methodologies with deep learning techniques, offers a promising solution to tackle the complexities associated with Arabic question generation. These hybrid models can harness the advantages of rule-based approaches, adept at managing intricate sentence structures and incorporating domain-specific knowledge, while concurrently leveraging the capabilities of deep learning models to generate questions that are diverse and contextually relevant [4].

In conclusion, future research in Arabic question generation should focus on the development of Arabic-specific datasets, integration of advanced linguistic resources, application of transfer learning, and design of hybrid models to improve the accuracy and contextual relevance of generated questions.

7. Conclusion

This literature review discussed the current state of research on Arabic question generation using deep learning techniques for educational purposes. It highlighted the challenges faced in Arabic question generation, such as the lack of Arabic-specific datasets and linguistic resources, and suggests potential solutions, including the development of Arabic-specific datasets, integration of advanced linguistic resources, application of transfer learning, fine-tuning pre-trained multilingual models and the design of hybrid models. Overall, the literature review provided insights into the current state of research on Arabic question generation using deep learning techniques for educational purposes. Future research in Arabic question generation should focus on the development of Arabic-specific datasets, integration of advanced linguistic resources, application of transfer learning, and the design of hybrid models to improve the accuracy and contextual relevance of generated questions.

References

[1] Bousmaha, K. Z., Chergui, N. H., Mbarek, M. S. A., & Belguith, L. H. (2020). AQG: Arabic Question Generator. Rev. d'Intelligence Artif., 34(6), 721-729.

[2] Rahmani, A. M., Yousefpoor, E., Yousefpoor, M. S., Mehmood, Z., Haider, A., Hosseinzadeh, M., & Ali Naqvi, R. (2021). Machine learning (ML) in medicine: Review, applications, and challenges. Mathematics, 9(22), 2970.

[3] Biltawi, M. M., Tedmori, S., & Awajan, A. (2021). Arabic question answering systems: gap analysis. IEEE Access, 9, 63876-63904.

[4] Alwaneen, T H., Azmi, A M., Aboalsamh, H., Wang, Z., & Hussain, A. (2021, July 12). Arabic question answering system: a survey. https://doi.org/10.1007/s10462-021-10031-1.

[5] Alhashedi, S., Suaib, N. M., & Bakri, A. (2022). Arabic Automatic Question Generation Using Transformer Model (No. 8588). EasyChair.

[6] Al-Janabi, A., Al-Zubaidi, E. A., Al-Sagheer, R. H. A., & Hussein, R. (2020). Encapsulation of semantic description with syntactic components for the Arabic language. Indonesian Journal of Electrical Engineering and Computer Science, 22(2), 961-967.

[7] Fuad, A., & Al-Yahya, M. (2022). Recent developments in Arabic conversational AI: a literature review. IEEE Access, 10, 23842-23859.

[8] Keklik, O. (2018). Automatic question generation using natural language processing techniques (Doctoral dissertation, Izmir Institute of Technology (Turkey)).

[9] Beheitt, M. E. G., & Hmida, M. B. H. (2022). Automatic Arabic Poem Generation with GPT-2. In ICAART (2) (pp. 366-374).

[10] Ushio, A., Alva-Manchego, F., & Camacho-Collados, J. (2023). A Practical Toolkit for Multilingual Question and Answer Generation. arXiv preprint arXiv:2305.17416.

[11] Alsubhi, K., Jamal, A., & Alhothali, A. (2022). Deep learning-based approach for Arabic open domain question answering. PeerJ Computer Science, 8, e952.

[12] Jayarajan, A K. (2020, March 30). Automatic Question Generation u sing Sequence to Sequence RNN Model. https://scite.ai/reports/10.35940/ijitee.e2675.039520.

[13] Das, B., Majumder, M., Phadikar, S., & Sekh, A. A. (2021). Automatic question generation and answer assessment: a survey. Research and Practice in Technology Enhanced Learning, 16(1), 1-15.

[14] Khan, S., & Emara, S. A. (2018). Effect of Technology Use in Education. International Journal of Pedagogical Innovation, 6(2).

[15] Alotaibi, F., Abdullah, M T., Abdullah, R., Rahmat, R W O K., & Murad, M A A. (2019, November 22). A Method for Arabic Handwritten Diacritics Characters. https://scite.ai/reports/10.35940/ijeat.f1034.0986s319

[16] Ali, M. N., Tan, G., & Hussain, A. (2018). Bidirectional recurrent neural network approach for Arabic named entity recognition. Future Internet, 10(12), 123.

[17] Abdelali, A., Mubarak, H., Chowdhury, S. A., Hasanain, M., Mousi, B., Boughorbel, S., ... & Alam, F. (2023). Benchmarking Arabic AI with Large Language Models. arXiv preprint arXiv:2305.14982.

[18] Farghaly, A., & Shaalan, K. (2009). Arabic natural language processing: Challenges and solutions. ACM Transactions on Asian Language Information Processing (TALIP), 8(4), 1-22.

[19] Pranav, D S., & R, B P V. (2023, May 28). Histobot: Question Generation System Using Deep Learning Techniques. https://doi.org/10.47392/irjash.2023.s071

[20] Krishna, R., Bernstein, M S., & Li, F. (2019, March 26). Information Maximizing Visual Question Generation. http://arxiv.org/abs/1903.11207

[21] Abuali, B., & Kurdy, M. B. (2022). Full Diacritization of the Arabic Text to Improve Screen Readers for the Visually Impaired. Advances in Human-Computer Interaction.

[22] Antoun, W., Baly, F., & Hajj, H. (2020). AraGPT2: Pre-trained transformer for Arabic language generation. arXiv preprint arXiv:2012.15520.

[23] Alami, H., El Mahdaouy, A., Benlahbib, A., En-Nahnahi, N., Berrada, I., & Ouatik, S. E. A. (2023). DAQAS: Deep Arabic Question Answering System based on duplicate question detection and machine reading comprehension. Journal of King Saud University-Computer and Information Sciences, 35(8), 101709.

[24] Alkhurayyif, Y., & Sait, A. R. W. (2023). Developing an Open Domain Arabic Question Answering System Using a Deep Learning Technique. IEEE Access.

[25] Wyse, B., & Piwek, P. (2009). Generating questions from openlearn study units.

[26] Alsubait, T., Parsia, B., & Sattler, U. (2012). Automatic generation of analogy questions for student assessment: an Ontology-based approach. Research in Learning Technology, 20.

[27] Kalady, S., Elikkottil, A., & Das, R. (2010, June). Natural language question generation using syntax and keywords. In Proceedings of QG2010: The Third Workshop on Question Generation (Vol. 11, pp. 1-10). questiongeneration. org.

[28] Liu, M., Calvo, R. A., & Rus, V. (2010). Automatic question generation for literature review writing support. In Intelligent Tutoring Systems: 10th International Conference, ITS 2010, Pittsburgh, PA, USA, June 14-18, 2010, Proceedings, Part I 10 (pp. 45-54). Springer Berlin Heidelberg.

[29] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

[30] Nagoudi, E. M. B., Elmadany, A., & Abdul-Mageed, M. (2021). AraT5: Text-to-text transformers for Arabic language generation. arXiv preprint arXiv:2109.12068.

[31] Kalpakchi, D., & Boye, J. (2023). Quinductor: a multilingual data-driven method for generating reading-comprehension questions using universal dependencies. Natural Language Engineering, 1-39.

[32] El-Khair, I. A. (2016). 1.5 billion words Arabic corpus. arXiv preprint arXiv:1611.04033.

[33] Zeroual, I., Goldhahn, D., Eckart, T., & Lakhouaja, A. (2019, August). OSIAN: Open source international Arabic news corpus-preparation and integration into the CLARIN-infrastructure. In Proceedings of the fourth arabic natural language processing workshop (pp. 175-182).

[34] kamel, S. M., Hassan, S. I., & Elrefaei, L. (2023). VAQA: Visual Arabic Question Answering. Arabian Journal for Science and Engineering, 1-21.

[35] Kriangchaivech, K., & Wangperawong, A. (2019). Question generation by transformers. arXiv preprint arXiv:1909.05017.

[36] Shaalan, K., Siddiqui, S., Alkhatib, M., & Abdel Monem, A. (2019). Challenges in Arabic natural language processing. In Computational linguistics, speech and image processing for Arabic language (pp. 59-83).

[37] Suárez, P. J. O., Romary, L., & Sagot, B. (2020). A monolingual approach to contextualized word embeddings for mid-resource languages. arXiv preprint arXiv:2006.06202.

[38] Mozannar, H., Hajal, K. E., Maamary, E., & Hajj, H. (2019). Neural Arabic question answering. arXiv preprint arXiv:1906.05394.

[39] Al-Bataineh, H., Farhan, W., Mustafa, A., Seelawi, H., & Al-Natsheh, H. T. (2019, November). Deep contextualized pairwise semantic similarity for Arabic language questions. In 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI) (pp. 1586-1591). IEEE.

[40] Lewis, P., Oğuz, B., Rinott, R., Riedel, S., & Schwenk, H. (2019). MLQA: Evaluating cross-lingual extractive question answering. arXiv preprint arXiv:1910.07475.

[41] Artetxe, M., Ruder, S., & Yogatama, D. (2019). On the cross-lingual transferability of monolingual representations. arXiv preprint arXiv:1910.11856.

[42] Roy, U., Constant, N., Al-Rfou, R., Barua, A., Phillips, A., & Yang, Y. (2020). LAReQA: Language-agnostic answer retrieval from a multilingual pool. arXiv preprint arXiv:2004.05484.

[43] Ismail, W. S., & Homsi, M. N. (2018). Dawqas: A dataset for Arabic why question answering system. Procedia computer science, 142, 123-131.

[44] Hardalov, M., Mihaylov, T., Zlatkova, D., Dinkov, Y., Koychev, I., & Nakov, P. (2020). EXAMS: A multi-subject high school examinations dataset for cross-lingual and multilingual question answering. arXiv preprint arXiv:2011.03080.

[45] Bonifacio, L., Jeronymo, V., Abonizio, H. Q., Campiotti, I., Fadaee, M., Lotufo, R., & Nogueira, R. (2021). mmarco: A multilingual version of the ms marco passage ranking dataset. arXiv preprint arXiv:2108.13897.

[46] Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). Bleu: a method for automatic evaluation of machine translation. In Proceedings of the 40th annual meeting of the Association for Computational Linguistics (pp. 311-318).

[47] Lin, C. Y. (2004, July). Rouge: A package for automatic evaluation of summaries. In Text summarization branches out (pp. 74-81).

[48] Banerjee, S., & Lavie, A. (2005, June). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization (pp. 65-72).

[49] Riabi, A., Scialom, T., Keraron, R., Sagot, B., Seddah, D., & Staiano, J. (2021, January 1). Synthetic Data Augmentation for Zero-Shot Cross-Lingual Question Answering.https://scite.ai/reports/10.18653/v1/2021.emnlp-main.562

[50] Vedantam, R., Zitnick, C.L., Parikh, D. (2015). CIDEr: Consensusbased image description evaluation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 07–12-June, pp. 4566–4575. https://doi.org/10.1109/CVPR.2015.7299087

[51] Mulla, N., & Gharpure, P. (2023). Automatic question generation: a review of methodologies, datasets, evaluation metrics, and applications. *Progress in Artificial Intelligence*, *12*(1), 1-32.

[52] Ahmed, M. R. A. (2025). Accreditation and Quality Assurance: Exploring Impact and Assessing Institutional Change in the US and Saudi Arabian Higher Education Institutions. *Arts for Linguistic & Literary Studies*, *7*(1), 626–639. https://doi.org/10.53286/arts.v7i1.2419

[53] Al-Ghobesi, A. A. H. (2025). Risks of Relying on Artificial Intelligence in Learning Arabic Language Sciences Through the Meta Application. *Arts for Linguistic & Literary Studies*, *7*(1), 396–419. https://doi.org/10.53286/arts.v7i1.2420.

[54] Alharbi, K. N. . (2022). Female Saudi ESL Learners' Attitudes Toward Communication in Mixed Gender Classes in the USA. *Arts for Linguistic & Literary Studies*, *1*(13), 7–30. https://doi.org/10.53286/arts.v1i13.838.

مجلــة العلــوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

Approach to Plagiarism Detection in Programming Assignments

Malek Algabri[1,2], Firdaus Alhrazi[1]

[1]College of Computer Science and Technology, Sana'a University

2College of Engineering and Information Technology, Emirates International University, Sana'a, Yemen

malekye@su.edu.ye, dr.malekye@eiu-edu.net , firdaus.mansoor2024@gmail.com

Abstract:

People tend to shortcut ways that save them time and effort to do the tasks required by them, either by taking tasks ready-made online, or stealing someone's work as their own. Since everything now is connected to the Internet, there is a very high potential of duplicating or stealing someone else's work, which is known as plagiarism. With the advancement of technology, it has become quite simple to do all tasks through the Internet. Plagiarism is the copying of other people's ideas and actions; it is considered a crime. Plagiarism occurs due to laziness, fear of failure, and the desire to perform the required tasks without fatigue or effort. In this paper, a methodology for detecting plagiarism in programming tasks, in particular in the visual programming category using deep learning and machine learning algorithms is proposed. Also, a solution has been proposed to detect plagiarism in the source code and interfaces that pertainto programming assignments.

Keywords: Index Terms Plagiarism Detection, Programming Assignments, Ma chine Learning, CNN, MILEPOST GCC.

I. INTRODUCTION

Plagiarism detection is useful in the era of technological development. Plagiarism detection helps to maintain integrity in the academic field, as it is easy to copy or steal the worksof others via the Internet. Most people tend to steal or copy someone's work to save effort and time in getting the costsor work required of them done. If people continue to steal the works of others, the original work will not be valuable at all. Plagiarism in programming tasks is a big problem, where students copy codes from their classmates, or find codes onlineand utilize them without indicating to original code [1].

There are many plagiarism detection tools, and these tools have greatly reduced manual checking. Most plagiarism de- tection tools use syntactic-based methods and a text-based ap- proach [2]. These tools tend to fail when the code obfuscation process is done. In programming tasks, code obfuscation can be done by changing variable names, rearranging sentences, replacing loops and control sentences, converting variables, and adding comments.

It is very difficult to identify copied work in programming languages because there are different types of plagiarism, which are difficult to find with plagiarism detection tools.Tools like JPlag [3] and MOSS [4] remove comments before testing for plagiarism, which can lessen their importance for plagiarism detection. There are different types of plagiarism

[5] such as direct plagiarism, smart plagiarism, self-plagiarism, and unintentional plagiarism. Direct plagiarism is copying without any change or quotations. Smart plagiarism is when someone plagiarizes an original work skillfully, they rephrase sentences in their way. Self-impersonation is when someone presents the same as their previous work in other situations andtasks. Unintentional plagiarism where this type occurs when someone inadvertently reformulates a topic [5].

In this paper, the focus will be on plagiarism detection of source code and interfaces in visual programming class. The goal is to solve the problem of plagiarism in students' softwaretasks.

II. RELATED WORK

Most universities use online code evaluation platforms [6], making plagiarism an easy task in programming tasks. Imper- sonation in general in the field of education is considered a betrayal of academic honesty and a moral crime. There aretwo ways to detect plagiarism [5]:

- Document plagiarism.

- Source code plagiarized.

A. *Existing Tools*

- Tools for detecting plagiarism online:

  An online plagiarism detector is a tool that takes text or documents as input and checks or compares them with several articles and papers published that have been published online. It produces a report at the end for reference that shows the proportion of information that has been plagiarized [5].

- Independent tools for detecting plagiarism:

  Unlike online plagiarism services, plagiarism detection apps may be installed and used on individual PCs andare used to check for plagiarism in specified input textsor documents. To find text matches, it does a huge amountof internet searches using articles. For these resources to work, the device must be connected to the Internet [5].

B. *Plagiarism Detection for Source code and Texts*

Based on a survey [7] of computer undergraduate students, 54% admitted plagiarizing and over 87% reported that their assignment was plagiarized. According to article [8] students cheat for a variety of reasons, including the ease with which

they may access or exchange material online, the low like- lihood of being discovered, a lack of time, a fear of being judged, and so on. As a result, when students complete the course, they are deficient in important knowledge, and their entitlement to a fair and equal evaluation of their expertise is denied. Due to the large number of students in each course, educators rely on assignments to determine which students need further help and whether the course can be made better [6].

C. *Detection of Source Code Plagiarism*

Source code plagiarism occurs when someone copies or reproduces the same source code without properly citing the original source code author [5]. There are two methods available:

- Syntaxial Change

  The shift is the simplest form in syntaxial. They may be completed by making a few straightforward adjustments to the main code. In this situation, knowing how to code is not necessary [5].

- Structural Change

Changes to calls made within the process body and vice versa, changes to iterations, conditional statements, and statement order, as well as the conversion of procedures to works and the inclusion of statements that will not modify the code's output are all instances of structural changes [5].

D. *Process Model for Plagiarism Detection*

An enhancement to the process model used to identify plagiarism. Five phases make up the comparable model seenin [1]:

- By tokenizing the source code, detection becomes re- sistant to changes in the source code's structure, likelanguage translations and variable renaming.
- Pre-processing improves the detection of source code changes such as removing comment blocks, splitting or merging variable declarations, rearrangement of state-ments, etc.
- deleting and creating a tokenized version of source code that is allowed.
- Utilizing similarity measurement to find similarities.
- Finally, calculating the amount of similarity between two source codes.

E. *Detecting Source Code Plagiarism Bytecode Approach*

A technique for detecting source code plagiarism that uses low-level instructions instead of source code tokens may catch the majority of plagiarism assaults from beginner pro- gramming courses at any level. Additionally, a number of techniques, including instruction reinterpretation, instruction generalization, method linearization, and method-based com- parison are incorporated to increase its efficacy. The choice ofJava as the targeted programming language with Bytecode as its low-level instruction was made, due to low-level instructionbeing a language-dependent characteristic based on analysis

[9]. Most source code plagiarism detection tools are created with language independence in mind. This kind of approach generalizes all computer languages and treats them as raw text,assuming that the most obvious plagiarism indications are not language-dependent [10]. Suggest a controlled experiment to gather plagiarism assaults because more accurate plagiarism detection may be built if the most likely plagiarism attacks are identified and reported. Based on 378 source codes created by respondents who stole work, we investigate potential plagia- rism assaults in a controlled experiment. Respondents are only allowed to be skilled programmers from different courses in order to gather more diverse plagiarism attacks. Additionally, the majority of them are trained to spot plagiarism in students'homework as lecturer assistants.

F. Source Code Plagiarism Detection using Machine Intelli- gence Approach

There are several methods available to assist preserve the requisite integrity and identify plagiarism. where it deals with plagiarism in a particular category of C programming tasks. The benefits and drawbacks of various deep learning and Machine Learning techniques are thoroughly examined. To identify plagiarism in source code, algorithms like SVM, KNN, RNNs, D-Trees, and attention-based transformer net- works are put to the test. During the course of this investigation, a large dataset made up of code pairings was created. The results collected demonstrate that the state-of-the-art text-based plagiarism detectors used today are not as accurate at identifying plagiarism as machine learning and deep learning approaches [2].

The authors in [11] use feature-based neural networks to try to detect plagiarism. They analyze the usefulness of the various characteristics for identifying plagiarism and use a neural network approach to detect plagiarism. The similarity provided by MOSS is employed as one of the characteristics of their method and is also demonstrated to have the highest significance.

The authors in [12] go through various Machine Learning approaches to plagiarism detection. The outcomes of several experiments are used to illustrate them. Result of testing several machine learning techniques on feature pairings, conclusions are drawn.

III. METHODOLOGY

To detect plagiarism in source code and interfaces in visual programming, a methodology consisting of steps is proposed, which are as follows:

A. Dataset Preparing

There is the study [2] was done before the dataset is created. conducted to create a dataset containing pairs of plagiarized and non-plagiarized programs, and different methods were used to obfuscate the source code, and then collect programs from various tutorials and online platforms. In this section, the dataset will be used in the study [2], and data augmentation will be introduced.

Here, different techniques for code obfuscation were utilized to provide a sufficient dataset that could accommodate a range of potential instances of source code and interface theft. Programs were gathered from schoolwork assignments, various tutorials, and from internet coding platforms to include a range of source codes and interfaces. Out of them, 519 pairs included plagiarism, whereas the remaining 722 pairs did not. Pair of copied source codes and interfaces that had been obfuscated using well-known methods. In

order to make thedataset impenetrable against a range of problem sets, it wasmade sure that a variety of strategies were incorporated where

the different code obfuscation methods are as follows:

- Change in Data Types.
- Change in Variable Name.
- Changing from For Loop to While Loop.
- Rearranging the Statements.
- Coding Block Reordering.
- Expression Reordering.
- If Else converted statements.
- Change colors in interfaces.
- Change the order of tools in interfaces.

To detect plagiarism in visual programming, we must take two approaches. First, plagiarism detection in source code. Second, plagiarism detection in the interfaces. Below are more details.

B. *Textual Feature Extraction for Source Code*

In this approach, plagiarism of source code will be detected by extracting features from source code using machine learn- ing. The following is the plagiarism detection mechanism.

To extract features from a source code MILEPOST GCC with O3 was utilized [13]. A machine learning compiler called MILEPOST GCC is used for testing and research. The scale of each characteristic ranged from zero (the least value) to one(the greatest value).

In [2] Support Vector Machine (SVM), K Nearest Neighbors (KNN), and Decision Trees were the classification techniques employed on the condensed dataset. The next sections show how these algorithms are implemented and the accuracy that may be attained using them. The weights assigned based on the distance between the 7 nearest neighbors were taken into account when the K Nearest Neighbors algorithm was developed. The Gini Index served as the foundation for the Decision Trees Classifier method that was used. The maximum depth was 4, and the minimum number of leaves for the sample was five. Support Vector Machine (SVM) was the most recent algorithm used. The Grid Search Technique was used to get the SVM model's ideal parameters[14].

C. *Image Feature Extraction for Interface*

In this approach, plagiarism of interfaces will be detected by extracting features from interfaces as images using a Convolutional Neural Network (CNN). The following is the plagiarism detection mechanism.

When using traditional feed-forward neural networks to handle picture classification issues, image pixels might be

utilized directly as input. Compared to conventional fully connected neural networks, CNN networks are significantly quicker and more reliable [15]. The CNN model has excelled in a variety of computer vision tasks, including image process-ing. Additionally, we may utilize the CNN model to extractthe features. pull features from various network levels for use in other tasks. In general, global feature extraction based on CNN consists of feature extraction, pretraining the CNN model, and fine-tuning the CNN model [16].

The feature extraction pooling layer and convolution layer are two different types of network layers in the CNN model. To extract various picture information, the convolution layer convolves input data using several convolution kernels. The pooling layer samples the data coming in. The activation function then nonlinearly abstracts input information. After entering the CNN model, the original picture moves ahead through pooling, multi-layer convolution, and non-linear addressing. Additionally, the picture data is continuously abstracted. The final output, characteristics increasingly abstract from low-level semantic data to local specifics [16]. As a result, the CNN model has more low-level data and higher-level semantic information the closer it is to the input convolution layer and full connection layer, respectively.

In this step, the focus will be on the detection of plagiarism in the software interfaces, as students may perform plagiarism operations for the software interfaces, the solution is to use CNN so as to detect plagiarism between the interfaces. Through CNN the features of the interfaces will be extracted, and then the features extracted from the interfaces to be detected for plagiarism will be compared with the dataset createdin advance in the first step. The result of this comparison will be calculated as a percentage of the amount of differencebetween the interfaces to be examined and the previously prepared dataset.

IV. RESULT AND DISCUSSION

Using machine learning and deep learning, an algorithm was created to identify interface and source code plagiarism. By contrasting them and determining the proportion by utilizing the Machine Learning algorithm and CNN, we can determine the amount of plagiarism that has been done in visual pro-gramming.

Figure. 1 shows the performance of our method when compared to MOSS and JPlag. The blue bars represent the performance of our method, and the red bars represent MOSS and yellow JPlag.

We can find the amount of plagiarism performed on a pro- gram by comparing the source code and interfaces in the pro- gram and calculating the amount of percentage of the amount of difference between the program and the programs stored in the dataset prepared for this task regarding the source code,the percentage of difference is calculated by Machine Learningand the percentage of the difference between interfaces by CNN's deep learning algorithms. Plagiarism is verified by taking outputs and calculating the average percentages to find the final plagiarism percentage. After training on source code data [2], the results were tested on a test suite including 496 programs, which was processed into 248 pairs. The top levelof accuracy was obtained using the SVM model. The details obtained for each model are given in Table. I.

Model	Accuracy
SVM	97.42%
D-Trees	96.35%
KNN	96.95%
Proposed Work	98.95%

TABLE I

ACCURACY OF CLASSIFIER MODELS

A simple Convolutional Neural Network has also been pro- posed to detect plagiarism in the interface by classifying images. This was done by taking the features of the interfaces andcomparing them with the data in the dataset; then calculating all the proportions of differences between the data stored in the dataset and the interface that is required to be examined, and after that calculating the average of differences between those interfaces. In addition, in this methodology, interfaces are treated as images. This simple Convolutional Neural Network imposes a lower computational cost as well as better results, which is what distinguishes the methodology in this paper compared to plagiarism detection tools.

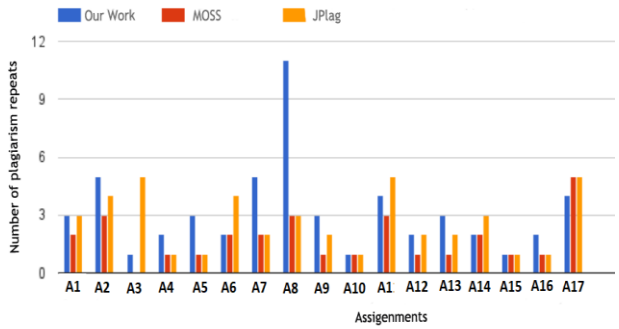


Figure I

Performance of our method when compared to MOSS and JPlag

V. CONCLUSION AND FUTURE WORK

Academic dishonesty is defined as plagiarism. Even though stealing someone else's work is legal, it may still violate their copyright. It is a serious, unethical infraction in academia. The law does not penalize plagiarism. Plagiarism is defined as using someone else's words or ideas without their permission or referencing the actual author of the material in order to pass it off as your own. This study was done on the application of deep learning and machine learning algorithms for plagiarism detection shows significant promise. Compared to earlier techniques, these new algorithms are far more potent. There is an opportunity to strengthen the algorithms, though. Future promises may include creating a tool that uses machine learning methods to show precisely where code noise occurs. Utilizing the Bag of Words algorithm to identify plagiarism is a further research field. There may be further techniques to obtain dynamic features from the program that can spur more advancement in this particular field and lengthen the useful life of plagiarism detection.

REFERENCES

1. Kermek, D. and M.J.I.i.E. Novak, *Process model improvement for source code plagiarism detection in student programming assignments*. 2016. **15**(1): p. 103-126.
2. Eppa, A. and A. Murali. *Source Code Plagiarism Detection: A Machine Intelligence Approach*. in *2022 IEEE Fourth International Conference on Advances in Electronics, Computers and Communications (ICAECC)*. 2022. IEEE.
3. Prechelt, L., G. Malpohl, and M.J.J.U.C.S. Philippsen, *Finding plagiarisms among a set of programs with JPlag*. 2002. **8**(11): p. 1016-.
4. Schleimer, S., D.S. Wilkerson, and A. Aiken. *Winnowing: local algorithms for document fingerprinting*. in *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*. 2003.
5. Yasmeen, S., et al. *Plagiarism Detection for Source Codes and Texts*. in *Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021*. 2022. Springer.
6. Alexandra-Cristina, C. and A.-C. Olteanu. *Material survey on source code plagiarism detection in programming courses*. in *2022 International Conference on Advanced Learning Technologies (ICALT)*. 2022. IEEE.
7. Sraka, D. and B. Kaucic. *Source code plagiarism*. in *Proceedings of the ITI 2009 31st international conference on information technology interfaces*. 2009. IEEE.
8. Albluwi, I.J.A.T.o.C.E., *Plagiarism in programming assessments: a systematic review*. 2019. 20(1): p. 1-28.
9. Karnalim, O. *Detecting source code plagiarism on introductory programming course assignments using a bytecode approach*. in *2016 International Conference on Information & Communication Technology and Systems (ICTS)*. 2016. IEEE.
10. Brixtel, R., et al. *Language-independent clone detection applied to plagiarism detection*. in *2010 10th IEEE Working Conference on Source Code Analysis and Manipulation*. 2010. IEEE.
11. Engels, S., V. Lakshmanan, and M. Craig. *Plagiarism detection using feature-based neural networks*. in *Proceedings of the 38th SIGCSE technical symposium on Computer science education*. 2007.
12. Katta, J.Y.B., *Machine learning for source-code plagiarism detection*. 2018, International Institute of Information

Technology Hyderabad, University of ....

13. Fursin, G., et al. *MILEPOST GCC: machine learning based research compiler*. in *GCC summit*. 2008.
14. Platt, J.J.A.i.l.m.c., *Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods*. 1999. **10**(3): p. 61-74.
15. Li, Q., et al. *Medical image classification with convolutional neural network*. in *2014 13th international conference on control automation robotics & vision (ICARCV)*. 2014. IEEE.
16. Peng, X., et al., *Research on image feature extraction and retrieval algorithms based on convolutional neural network*. 2020. **69**: p. 102705.

مجلــة العلــوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

Synergizing Software Engineering and AI: A Sustainable Intelligent Model for Automating Higher Education Quality Assurance Auditing

Ibrahim Ahmed Al-Baltah

Department of Information Technology, Faculty of Science and Engineering, Al-Hikma University, Sana'a, Yemen.
albalta2020@gmail.com

RECEIVED:00/00/2024

REVISED : 00/00/2024

ACCEPTED : 00/00/2024

PUBLISHED :00/00/2023

Abstract:

Higher education quality assurance is a very important task for improving the quality of education, which needs regular auditing. This auditing process requires a lot of manual work to evaluate the higher education institutions against quality assurance standards, guides, and forms. Therefore, this paper aims to propose a new AI-based model to automate the auditing process and to ensure as accurate results as possible when needed. This model utilizes artificial intelligence and software engineering. The software architecture of the proposed model consists of three layers that systematically interact with each other to achieve the model goal.

Keywords: Artificial intelligence, software architecture design, quality assurance, auditing, natural language processing, deep learning.

1. Introduction

Artificial Intelligence (AI) is a concept that started a long time ago and gradually became a very important and yet very popular and active field [1]. AI technology is no longer a research subject, rather it is an integral component model for most sectors such as finance, medicine, manufacturing, education, and others [2]. AI has a great power of decision making that helps business processes to take better action to ensure obtaining and achieving optimal competitive advantages. Higher education quality assurance is not an exception; it utilizes many technologies to improve the quality of education at different levels. The utilization of AI in education will significantly impact education from kindergarten education level to the highest postgraduate level (Alasmari, 2023) [3].

According to (Omer, 2024)[4], the process of the Council for Accreditation and Quality Assurance (CAQA) evolves some processes to control and develop the required higher education standards, guides, and forms to improve the output quality of the education service.

However, researchers have tried to investigate the role and impact of AI on higher education quality assurance (Al-Ghobesi, 2025, Ahmed, 2025)[5]. Meanwhile, others are utilizing software engineering during the development of software that is used in higher education quality assurance. Therefore, this paper aims to fill the existing gap through the consideration of these three areas and reveal the possible advantages of their utilization. Moreover, it aims to propose a new AI-based model for automating the process of quality assurance auditing that CAQA and higher education institutions could benefit from. This is because using AI and software engineering could yield a good combination that can help in improving the quality of services.

Current software systems combine several components and technologies from different fields. For example, automatic student assessment is made up of different components including IoT devices, AI algorithms, and other components. From a technical point of view, those components are not separate, hence, they are working together in a harmonized way, which could be achieved through software engineering. Software engineering plays a crucial role in developing and delivering high quality large systems [6], where auditing quality assurance is one of them.

Nonetheless, the authors of [7] presented a detailed case study on sustainable quality assurance at Prince Sultan University. They proposed a data warehouse for accreditation from a theoretical perspective. Yet, they ignored the auditing process. On the other hand, [4] proposed an internal quality assurance system based on AI that aimed at increasing the chance of accreditation; consequently, improving the quality of graduations.

However, the main purpose of this paper is to propose a new sustainable AI-based model for automating the process of auditing higher education quality assurance called AI-HEQA. This is due to the fact that CAQA and the higher education institutions are still conducting the auditing task manually, and thus needs a lot of effort. Therefore, AI-HEQA aims to produce instant and accurate auditing results with less effort.

The remainder of this paper is organized as follows: section 2 presents the proposed model. Section three presents the system architecture design of AI-HEQA from a software engineering point of view. Section four concludes the paper.

2. The proposed AI-based model

According to [8], automating the auditing process through the utilization of AI technologies can sharply reduce the required manual work and yield effective results with less cost. As mentioned earlier, the main goal of this paper is to propose a sustainable model that aims at automating the process of auditing higher education quality assurance. This main goal can be divided into two sub-goals. This first one is to help CAQA easily and instantly investigate to what extent higher education institutions are adhering to the imposed standards, guides, and forms by CAQA through visual representations and detailed reports. Meanwhile, CAQA can make a realistic decision based on the data that has been collected from the higher education institutions and processed by AI-HEQA. The second sub-objective is to help higher education institutions to monitor themselves and instantly identify the standards, guides, and forms that are not fully applied to improve their quality.

This model AI-HEQA consists of the following main steps:

a. **Defining Objectives and Criteria:** The main objective of the proposed model is derived from the objective of quality assurance auditing. In fact, the main objective of the auditing process is not only to identify and inform the higher education institutions of their deviations from the stated standards; it is rather meant for monitoring the progress of education quality by highlighting the main shortcomings that should be efficiently addressed to improve the education quality. This improvement will certainly improve the quality of the institution's outputs including students' knowledge and skills, which have direct impacts on our society and environments. On the other hand, it aims at helping quality assurance and the institutions to take the right actions when needed based on the information generated by AI-HEQA.

b. **Data Collection:** Data is the vital input for any AI model which are used for training and testing the model. This data should be collected from two sources where CAQA is the first source, in which standards, guides, forms, and other data relevant to quality assurance are collected. On the other

hand, the second data source is the higher education institutions, which have to provide the curriculums, institutional reports, programs' infrastructures, staff, and other related data. It is important to mention that, data collection should be carried out carefully and by specialists to ensure the quality of the collected data, which will certainly affect the results of the proposed model.

c. **AI Model:** From among different AI models, natural language processing (NLP) and deep learning (DL) are chosen. The reason for choosing NLP is that NLP has the power to deal with textual data [9], which represents most of the quality assurance data. Meanwhile, DL is used to develop predictive models to help the auditors efficiently and effectively identify the possible deviation from the standards, guides, and forms that are being stated by the CAQA. Therefore, the auditing process should be formulated and encoded into the AI model.

d. **Model Training:** After the construction of the proposed model, it is mandatory to sufficiently train the model until it obtains acceptable results. AI-HEQA should have sufficient learning to recognize patterns and relationships between features of the different forms of data. The training data should be prepared carefully and should cover all types of data to help the model generate as accurate results as possible.

e. **Model Validation and Evaluation**: Before deploying the model to the execution environment, it has to pass the validation and evaluation process. This is done by using testing data, which should be different from the training data to judge its ability to process unseen data to be generalized. For evaluating the performance of the proposed model, some relevant metrics shall be used such as accuracy, precision, recall, and F2 score [10].

f. **Interpretability and Explainability:** To make the auditing process using the AI-HEQA model easy and smooth, the produced results should be clear, easy to understand, and yet interpretable. Besides, the results should also be explainable in such a way that stakeholders can understand them correctly. This is a very important step to ensure a unified understanding and explanation, and thus, different stakeholders arrive at common decisions and conclusions. Without contradiction of interpretation and explanation of the model's results, it could help quality assurance to make suitable decisions with confidence.

g. **Feedback and continuous improvement:** To increase the trust of the stakeholders, the proposed model must be regularly improved continuously. This improvement is based on the feedback given by the CAQA experts and the higher education institutions. Despite the power of AI technology, the produced results by AI-HEQA should be further investigated by experts continuously to identify

possible improvements. Continuous feedback is very important to ensure the **sustainability of** AI-HEQA.

3. System architecture design of AI-HEQA from the software engineering point of view

According to the European Middle Market Report 2H 2015 [11], and U.S. Middle Market Report Q4 2015 [12], the rate of building unsuccessful software products is 75%, which wastes 322 billion USD of companies capital. Therefore, it is very important to adopt the best practices of software engineering to increase the rate of success of the proposed model. From a software engineering perspective, the suitable system architecture design of AI-HEQA consists of different layers as illustrated in Figure 1.

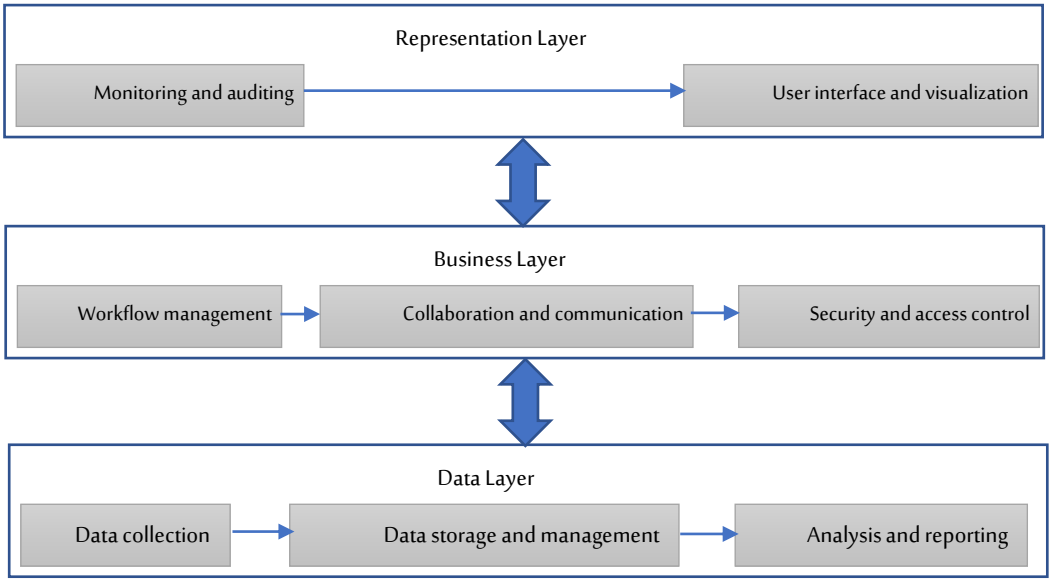


Figure 1: system architecture design of AI-HEQA

As can be shown in Figure 1, the proposed model consists of three main layers including representation, business, and data layer. It is worth pointing out that, the AI components are not isolated in a separate layer; instead, they are integrated into each layer to support different aspects of the auditing process.

a. **Data Layer:** it is the fundamental layer for the proposed model, which is responsible for collecting data from different sources, and storing and managing the collected data related to the auditing process. It has the capabilities to perform data analytics and generate the required reports.

   Data should be collected from CAQA and higher education institutions. During this process, the data collection component should employ unified data formats and ensure data integrity.

Furthermore, the output of this component is clean data without contradiction, which is valid for storage to be used for the auditing process.

The data storage and management component are responsible for storing the data that were generated from the data collection component, organizing this data for efficient use, and retrieving the stored data. A centralized repository is recommended at CAQA where other parties can have the right access to the right data only.

The last component in this layer is analysis and reporting. This component has the capability to perform intelligent data analytics to extract meaningful insights from the collected data. These insights are crucial for making the right auditing decisions. Moreover, this component is capable of generating comprehensive reports that provide quality assurance with very important information that presents to what extent the higher education institutions have adhered to the quality assurance standards, forms, and guides. Additionally, it identifies and highlights some points for quality improvement.

b. **Business layer:** it is a middle layer, in which the presentation layer and data layer communicate through it. It contains the core business logic of the whole model, including workflow management, collaboration and communication, and security and access control.

The first component of this layer is workflow management. It is the orchestrator for the different components that are involved in the auditing process. The workflow is task-based, in which the automated auditing process is divided into tasks. This could include automated data collection, data storage and management, analysis, and report generation, feedback collection.

The second component is collaboration and communication which consumes the output of the first component to perform its tasks. This component aims to facilitate the process of communication and collaboration among different parties namely: auditors, higher education institutions, and other stakeholders. It should provide sufficient communication and collaboration tools including, email, discussion forums, and chat. It is worth mentioning that, the model could be integrated with Chabot as a technique for communication and feedback collection.

The last component of this layer is the security and access control. This is a very critical component, especially with the daily increase of cyber-attacks. For that reason, it is recommended to apply software security best practices. The ultimate goal of this component is to protect the

collected data from unauthorized access. Therefore, this component is capable of data encryption, user authentication, and access control mechanisms.

c. **Representation layer:** This layer aims to provide users with interactive user interfaces to easily and smoothly interact with the model. Thus, auditors and stakeholders are able to navigate and access all of the model's functionalities. In addition, it provides a dashboard to display clear and concise charts and graphics that are easily understandable and interpretable by different users. Therefore, auditors can make initial decisions from the data presented on the dashboard. The final decisions can be made after generating detailed reports.

4. Conclusion

To evaluate to what extent the higher education institutions follow quality assurance is a tedious task. As a result, this paper proposed a new sustainable model with the aim of automatizing the process of quality assurance auditing called AI-HEQA. The model utilizes two different fields: AI and software engineering. This model consists of three layers namely the data layer, the business layer, and the presentation layer. Every layer also consists of some components, which are systematically integrated.

In future work, we are planning to develop the proposed model based on the provided system architecture and conduct experimental evaluations to ensure the applicability of the model. Consequently, the model can be deployed in the real environment for real use.

5. References

1. Xu, Y., et al., *Artificial intelligence: A powerful paradigm for scientific research.* 2021. **2**(4).
2. Dwivedi, Y.K., et al., *Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy.* 2021. **57**: p. 101994.
3. Ocaña-Fernández, Y., L.A. Valenzuela-Fernández, and L.L.J.J.o.E.P.-P.y.R. Garro-Aburto, *Artificial Intelligence and Its Implications in Higher Education.* 2019. **7**(2): p. 553-568.
4. Elviwani, M.Z., A. Dilham, and R.J.I.J.S.R.C.S.E.I.T. Buaton, *Higher Education Quality Assurance System Based Artificial Intelligence.* 2020. **3307**: p. 274-279.
5. Zawacki-Richter, O., et al., *Systematic review of research on artificial intelligence applications in higher education–where are the educators?* 2019. **16**(1): p. 1-27.
6. Van Vliet, H., H. Van Vliet, and J. Van Vliet, *Software engineering: principles and practice*. Vol. 13. 2008: John Wiley & Sons Hoboken, NJ.
7. Javed, Y. and M. Alenezi, *A Case Study on Sustainable Quality Assurance in Higher Education.* 2023. **15**(10): p. 8136.
8. Noordin, N.A., K. Hussainey, and A.F. Hayek, *The Use of Artificial Intelligence and Audit Quality: An Analysis from the Perspectives of External Auditors in the UAE.* 2022. **15**(8): p. 339.
9. Yogish, D., T. Manjunath, and R.S. Hegadi. *Review on natural language processing trends and techniques using NLTK.* in

Recent Trends in Image Processing and Pattern Recognition: Second International Conference, RTIP2R 2018, Solapur, India, December 21–22, 2018, Revised Selected Papers, Part III 2. 2019. Springer.

10. Abd El-Haleem, A.M., et al., *A Generic AI-Based Technique for Assessing Student Performance in Conducting Online Virtual and Remote Controlled Laboratories*. 2022. **10**: p. 128046-128065.
11. I, P.D., *European Middle Market Report 2H 2015*. 20215.
12. I, P.D., *U.S. Middle market report Q4 2015. Technical report*. 2015.
13. Alasmari, J. S. . (2023). The Dynamics of Verbal and Non-Verbal Linguistic Communication in The Saudi Sports Community. *Arts for Linguistic & Literary Studies*, *5*(4), 539–569. https://doi.org/10.53286/arts.v5i4.1676
14. Ahmed, M. R. A. (2025). Accreditation and Quality Assurance: Exploring Impact and Assessing Institutional Change in the US and Saudi Arabian Higher Education Institutions. *Arts for Linguistic & Literary Studies*, *7*(1), 626–639. https://doi.org/10.53286/arts.v7i1.2419
15. Al-Ghobesi, A. A. H. (2025). Risks of Relying on Artificial Intelligence in Learning Arabic Language Sciences Through the Meta Application. *Arts for Linguistic & Literary Studies*, *7*(1), 396–419. https://doi.org/10.53286/arts.v7i1.2420
16. Omer, N. I. M. (2024). Maintaining Meaningful Human Interaction in AI-Enhanced Language Learning Environments: A Systematic Review. *Arts for Linguistic & Literary Studies*, *6*(3), 533–552. https://doi.org/10.53286/arts.v6i3.2083

مجلــة العلــوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

AI Powered Chatbots-Driven Assistant in Educational and Commercial Foundations

Ammar Zahary

Sana'a University, University of Science and Technology, Sana'a, Yemen
a.zahary@ust.edu.ye

Nesmah Al-khulaidi

Yemen Academy for Graduate Studies, Sana'a, Yemen
nesmaalkhulaidi88@ieee.org

ABSTRACT

AI-driven chatbots have become adaptable tools with enormous promise in both educational and professional contexts. This paper examines the use of chatbots with AI as motivated helpers in institutional and commercial settings. The article investigates the advantages and difficulties of using chatbots in various fields. Chatbots can offer tailored learning experiences, help in administrative duties, and give students immediate support in the educational sector. Additionally, chatbots may be used in business settings to improve customer service, automate sales procedures, and offer effective product suggestions. The study also explores the underlying technologies, including natural language processing, machine learning, and knowledge bases that support these chatbots. Additionally, privacy issues and ethical issues relating to chatbots powered by AI are highlighted. This article emphasizes the transformational potential of AI-powered chatbots in commercial and educational contexts, stressing their capacity to increase productivity, enhance user experiences, and promote creativity.

1. INTRODUCTION

A number of sectors have been transformed by artificial intelligence (AI), and chatbots that use AI are one area where this promise is progressively being fulfilled. An adaptable tool that may support and enhance numerous activities in educational and commercial foundations, chatbots are driven by AI algorithms. This article explores the use of chatbots that are powered by AI as driving assistants in different fields, outlining their advantages, difficulties, and underpinning technology. Chatbot integration in educational settings has the potential to change how students learn. Personalized learning has become a priority, and chatbots can play a crucial role in providing tailored educational support to students. From answering queries and providing instant feedback to offering educational resources and guidance, chatbots can enhance students' engagement and facilitate self-paced learning. Additionally, chatbots can assist in administrative tasks, such as course registration, scheduling, and grading, reducing the burden on educators and administrators (Omer, 2024, Al-Ghobesi, 2025) [1],[4],[6].

AI-powered chatbots have the potential to transform consumer interactions and expedite company procedures in commercial settings. Chatbots can improve customer service by delivering round-the-clock assistance, responding to frequent questions, and making tailored recommendations. Additionally, chatbots have the ability to automate sales procedures, allowing easy transactions and raising conversion rates [12]. Chatbots may evaluate client data, forecast preferences, and give tailored marketing tactics by utilizing AI algorithms, which improve customer happiness and boost revenues. The underlying technologies that enable AI-powered chatbots include natural language processing (NLP), machine learning (ML), and knowledge bases. NLP allows chatbots to understand and interpret human language, enabling them to engage in meaningful conversations with users [7]. ML algorithms empower chatbots to continuously learn from user interactions, improving their responses and problem-solving capabilities over time. Knowledge bases store extensive information about educational content or product catalogs, serving as valuable resources for chatbots to provide accurate and relevant information to users. While the potential benefits of AI-powered chatbots in educational and commercial foundations are promising, there are also challenges to consider. Ethical considerations, privacy concerns, and the need for transparent decision-making are crucial aspects that must be addressed when implementing chatbot systems. Ensuring data security, maintaining user privacy, and addressing biases in AI algorithms are essential to foster trust and responsible use of AI-powered chatbots [7],[16].

In conclusion, this study seeks to offer a thorough comprehension of the function of chatbots powered by AI as motivated assistants in academic and commercial foundations. This paper adds to the body of

information on the revolutionary potential of AI-powered chatbots by examining their advantages, difficulties, and underlying technology. Understanding chatbots' capabilities and constraints may help educational and business institutions make the most of these intelligent tools, boosting productivity, enhancing user experiences, and encouraging innovation in their respective fields.

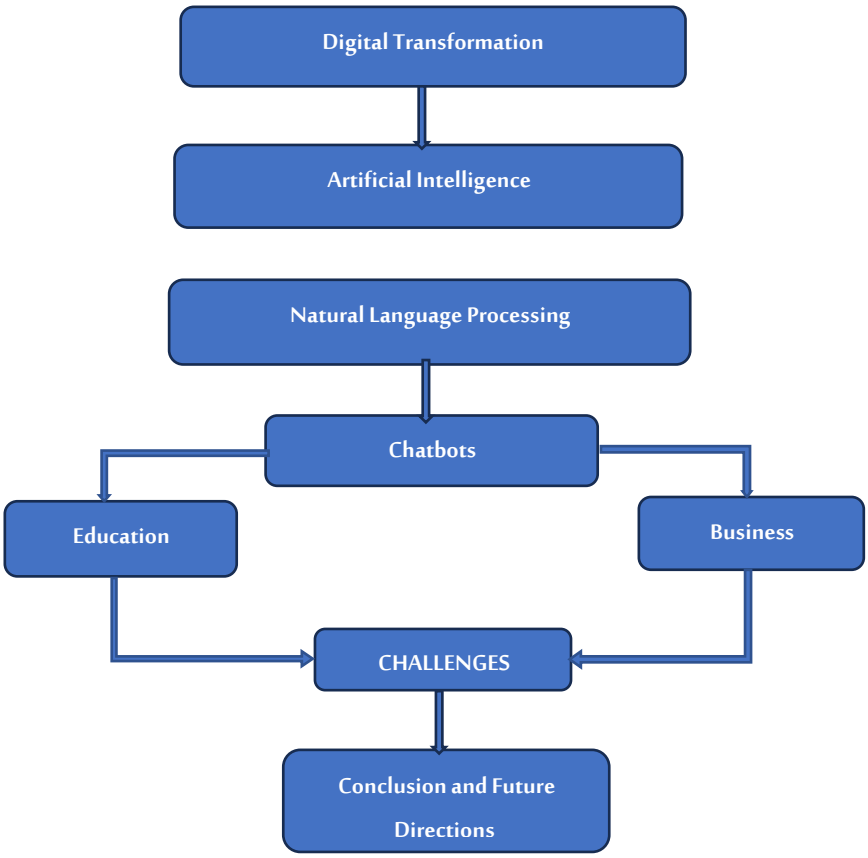


Fig.1 Research Scope

2. CHATBOTS IN DIGITAL TRANSFORMATION

Rule-based chatbots and machine learning chatbots are the two primary operations carried out by chatbot systems [18]. Rule-based chatbots function in accordance with a predetermined set of rules that specify how they must react to user input. The chatbot often uses certain words or phrases to set these rules, and the replies that follow are planned. Although rule-based chatbots can successfully respond to

straightforward questions, they are less flexible to complicated or unexpected user interactions because their replies are constrained by the preset rules [8].

Initiatives for the digital transformation now cannot be completed without the use of AI, ML, and chatbots. The development of intelligent chatbot systems is being driven by AI, which can imitate human intelligence, and ML, which allows systems to learn from data and improve performance. By automating consumer contacts, improving customer experiences, and optimizing corporate processes, chatbots—conversational agents—play a significant part in the digital transformation of businesses. These intelligent assistants are able to provide recommendations, respond to questions, and offer individualized assistance in natural language [10],[12]. Chatbots can analyze enormous volumes of data, spot patterns and trends, and offer insightful data for data-driven decision-making by utilizing AI and ML algorithms. The fusion of AI, ML, and chatbots in the context of digital transformation enables businesses to provide effective customer service, streamline processes, and open up fresh doors for development and innovation [16].

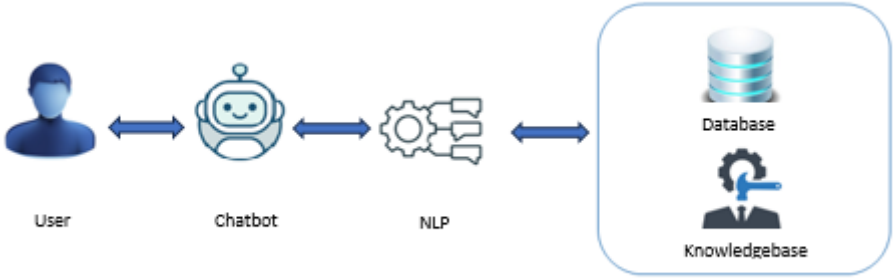


Fig.2 AI Powered Chatbot

3. NATURAL LANGUAGE PROCESSING IN CHATBOTS

In order for chatbots to properly grasp and respond to human language, NLP is essential to their growth. NLP includes a number of methods and algorithms that chatbots may utilize to understand and process user input in natural language. Language comprehension, sentiment analysis, entity recognition, and intent categorization are just a few of the duties involved. Chatbots may properly understand user questions, gather pertinent data, and produce suitable replies by utilizing NLP [5],[10]. Chatbots can manage complications like synonyms, context, and linguistic variances thanks to NLP algorithms, resulting in more conversational and user-friendly interactions.

Microsoft Azure's [15]: Azure Bot Service offers a complete platform for creating and deploying chatbots. To improve chatbot functionality, it makes use of Azure's robust cloud architecture and combines a number of AI services, including NLP capabilities. Developers may benefit from pre-built NLP models, train their own models, and easily incorporate them into their chatbot apps by employing Azure Bot Service. The service provides tools like language identification, sentiment analysis, and entity recognition that let programmers build chatbots that can comprehend user input and react properly in a natural and context-aware way. Additionally, the scalability, security, and analytics features of Azure Bot Service enable businesses to deploy and operate chatbots at scale while gathering knowledge about user interactions for ongoing development.

4. CHATBOTS IN EDUCATION

4.1. ENHANCE SKILLS AND KNOWLEDGE

The design and implementation of an educational chatbot to improve learners' knowledge and abilities centers around four important steps: knowledge linkage, emotional expression, dialogue management, and user modeling [2],[5]. In the beginning, knowledge connection entails linking the chatbot with pertinent educational resources and content to provide it with access to correct and current information. In order to improve learners' comprehension and learning process, the chatbot is able to offer them insightful information, justifications, and examples. The second component of emotional expressiveness is based on using sympathetic and encouraging words in the chatbot's answers. The chatbot may generate a more engaging and individualized engagement by recognizing and addressing the emotions of the learners, promoting a good learning environment. Thirdly, creating a productive discussion flow is a component of dialogue management [7]. The chatbot should be able to handle user requests, offer pertinent replies, and lead students along a logical learning path. Last but not least, user modeling entails gathering and assessing user information to customize the chatbot's interactions. This entails monitoring user progress, finding areas for development, and customizing advice and criticism to suit each learner's need. This approach enables educational chatbots to successfully help skill learners by giving them useful information, emotional support, meaningful conversation, and individualized learning experiences.

4.2. COST EFFICIENCY AND AUTOMATED TASKS

By automating many administrative and support activities, chatbots significantly reduce the cost of education. Chatbots relieve the stress on educators and administrative staff by handling repeated and time-consuming enquiries, enabling them to concentrate on more important tasks [13]. Chatbots eliminate the need for human interaction and streamlined administrative procedures by giving quick and accurate answers to frequent student questions such as course registration, scheduling, or gaining access to learning

materials. Additionally, chatbots can provide tailored learning experiences, giving students at a scale specific advice and support. Educational institutions may optimize resource allocation, increase operational efficiency, and provide improved services to students by utilizing chatbots' cost-effective technology, ultimately resulting in a more cost-effective education environment [14].

5. CHATBOTS IN BUSINESS

5.1.MARKETING

Chatbots have become effective marketing tools, altering consumer relationships and increasing engagement. Chatbots serve as virtual assistants in marketing, allowing companies to automate customer support, respond to questions, and provide tailored recommendations. They may be included into a variety of channels, such as websites, social media, and chat programs, enabling businesses to provide their client's real-time support. Marketers may improve customer experiences by using chatbots to give timely and pertinent information, assist consumers with the purchase process, and cater to their unique requirements and preferences. With the use of chatbots, companies may also gather crucial customer information that allows marketers to better understand consumer behavior and preferences and adjust their marketing strategies. Chatbots increase productivity and scalability by managing many customer conversations at once, allowing businesses to reach and connect with a broader audience successfully [1],[4],[6].

5.2. CUSTOMER SATISFACTION

Customer satisfaction has grown as a result of the use of chatbots in marketing. Customers no longer need to wait for assistance since chatbots respond to their queries right away. This immediate accessibility and reactivity enhance the client experience. The relevance and value of the customer's connection with the business are further increased by chatbots' ability to provide customized recommendations and help based on preferences and prior interactions [8],[11]. Chatbots assist clients in making educated selections and navigating through products and services with ease by providing quick and accurate information. Chatbots may also recall past interactions and preferences from customers, giving interactions with them a feeling of continuity and customization. Overall, chatbots' accessibility, response, personalization, and efficiency result in greater customer satisfaction levels, fostering stronger customer relationships and loyalty to the brand [12].

LPP [1], a well-known Polish clothes shop, stands out as a business that has successfully leveraged the potential of AI in marketing. It is an example of consumer pleasure. The use of AI-powered technologies, especially a smart chatbot, has allowed LPP to completely revamp its customer contact approach. The chatbot answers client questions about items fluently, thanks to its sophisticated natural language processing skills.

6. CHATBOTS CHALLENGES

6.1. Challenges Related to Education

7.1.1. Lack of Emotional Intelligence:

Chatbots may have trouble identifying and responding to students' emotional states [5]. For chatbots, understanding and treating emotions like irritation, perplexity, or worry can be difficult since they often rely on pre-programmed replies and may lack the empathy that human educators have.

7.1.2. Privacy and Data Security:

Chatbots used in education must handle sensitive student data, such as academic records and personal information [1]. To win the trust of students and adhere to data protection requirements, it is crucial to maintain privacy standards and provide reliable data protection procedures.

7.1.3. Ethical Considerations:

Ethical issues should be taken into account while creating and programming chatbots [1],[6]. When deploying educational chatbots, it is crucial to take into account issues like ensuring justice, openness, and accountability in decision-making processes, eliminating biases, and preventing discriminatory consequences.

6.2. Challenges Related to Business

7.2.1 Integration with Existing Systems:

Complex and varied IT infrastructures and systems are frequently used by businesses. It can be difficult and requires extensive technical skill to integrate chatbots with current systems, such as customer relationship management (CRM) software, enterprise resource planning (ERP) systems, or knowledge bases [4],[6].

7.2.2 Natural Language Understanding:

It can be difficult for chatbots to effectively understand and comprehend user questions [7]. Businesses deal with a variety of specialized terminology, jargon, and contextual cues; thus, it is essential that chatbots have cutting-edge natural language processing abilities to deliver precise and pertinent replies.

7.2.3. Constant Improvement and Maintenance:

For chatbots to continue functioning and being successful, regular monitoring, optimization, and maintenance are necessary. It takes time and dedication to update information often, improve conversational flows [14], and respond to user comments [17].

7. FRAMEWORK EXAMPLES

The AI readiness framework is a useful resource for businesses thinking about implementing AI technology [4]. It enables companies to evaluate their present level of preparedness, pinpoint problem areas, and develop a thorough plan for successfully deploying AI. The framework takes into account factors including corporate culture, talent and skills, technology infrastructure, data preparedness, and ethical issues. Organizations may assess each factor to have a comprehensive understanding of their AI preparedness and to spot any gaps or impediments that need to be filled. Another theoretical framework for learning analytics in STEM education, Applications for virtual reality (VR) must take into account the specific characteristics and difficulties presented by VR. Students may connect with challenging ideas in a virtual setting thanks to the immersive and engaging learning experiences offered by VR. For learning analytics, recording and evaluating students' interactions in VR, such as gaze, gestures, and movements, becomes essential. To effectively extract useful learning patterns from the massive amounts of data generated by VR applications, effective data collection, management, and analysis strategies are required [13]. STEM students may benefit from richer learning experiences by creating an effective learning analytics framework for VR that can improve instructional design, adaptive learning, and evaluation methods.

8. CONCLUSIONS, RECOMMENDATIONS, AND FUTURE DIRECTIONS

The possibility of chatbots with AI capacity as driven assistants in business and educational foundations has been examined in this article. The research has highlighted a number of advantages of chatbots in various settings, including improved customer services, individualized learning experiences, and higher operational effectiveness. In addition to helping in everyday activities and supporting students in their academic endeavors [11],[16], chatbots have the capacity to deliver timely and accurate information. Additionally, the use of AI technology enables chatbots to constantly learn and advance, adjusting to changing demands and preferences. To improve the efficiency of chatbots in these fields, however, a number of issues such natural language understanding, privacy issues, and user acceptability must be resolved [18].

There are a number of interesting future areas that entail additional investigation. First off, more study is required to improve chatbots' ability to grasp complicated questions, context, and emotions through natural language processing. The processing of users' personal information should be handled with confidence by taking steps to maintain strong data privacy and security protections. Additionally, the creation of more advanced chatbots with emotional intelligence and the ability to offer users sympathetic assistance will considerably increase their efficacy in academic and professional contexts and commercial settings.

9. REFERENCES

[1] E. Gołąb-Andrzejak, "AI-Powered Digital Transformation: Tools, Benefits and Challenges for Marketers – Case Study of LPP," Procedia Computer Science, vol. 219, pp. 397–404, 2023, doi: 10.1016/j.procs.2023.01.305.

[2] C. Kooli, "Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions," Sustainability, vol. 15, no. 7, p. 5614, Mar. 2023, doi: 10.3390/su15075614.

[3] B. Nyagadza, "Sustainable digital transformation for ambidextrous digital firms: systematic literature review, meta-analysis and agenda for future research directions," Sustainable Technology and Entrepreneurship, vol. 1, no. 3, p. 100020, Sep. 2022, doi: 10.1016/j.stae.2022.100020.

[4] J. Holmström, "From AI to digital transformation: The AI readiness framework," Business Horizons, vol. 65, no. 3, pp. 329–339, May 2022, doi: 10.1016/j.bushor.2021.03.006.

[5] C. Zhang, G. Li, H. Hashimoto, and Z. Zhang, "Digital Transformation (DX) for Skill Learners: The Design Methodology and Implementation of Educational Chatbot using Knowledge Connection and Emotional Expression," 2022 IEEE Global Engineering Education Conference (EDUCON), Mar. 2022, doi: 10.1109/educon52537.2022.9766384.

[6] M. Matarazzo, L. Penco, G. Profumo, and R. Quaglia, "Digital transformation and customer value creation in Made in Italy SMEs: A dynamic capabilities perspective," Journal of Business Research, vol. 123, pp. 642–656, Feb. 2021, doi: 10.1016/j.jbusres.2020.10.033.

[7] M.-H. Chao, A. J. C. Trappey, and C.-T. Wu, "Emerging Technologies of Natural Language-Enabled Chatbots: A Review and Trend Forecast Using Intelligent Ontology Extraction and Patent Analytics," Complexity, vol. 2021, pp. 1–26, May 2021, doi: 10.1155/2021/5511866.

[8] F. Brachten, T. Kissmer, and S. Stieglitz, "The acceptance of chatbots in an enterprise context – A survey study," International Journal of Information Management, vol. 60, p. 102375, Oct. 2021, doi: 10.1016/j.ijinfomgt.2021.102375

[9] A. Miklosik, N. Evans, and A. M. A. Qureshi, "The Use of Chatbots in Digital Business Transformation: A Systematic Literature Review," IEEE Access, vol. 9, pp. 106530–106539, 2021, doi: 10.1109/access.2021.3100885

[10] S. Magistretti, C. T. A. Pham, and C. Dell'Era, "Enlightening the dynamic capabilities of design thinking in fostering digital transformation," Industrial Marketing Management, vol. 97, pp. 59–70, Aug. 2021, doi: 10.1016/j.indmarman.2021.06.014

[11] K. Stalmachova, R. Chinoracky, and M. Strenitzerova, "Changes in Business Models Caused by Digital Transformation and the COVID-19 Pandemic and Possibilities of Their Measurement—Case Study," Sustainability, vol. 14, no. 1, p. 127, Dec. 2021, doi: 10.3390/su14010127

[12] S. Zel and E. Kongar, "Transforming Digital Employee Experience with Artificial Intelligence," 2020 IEEE / ITU International Conference on Artificial Intelligence for Good (AI4G), Sep. 2020, doi: 10.1109/ai4g50087.2020.9311088.

[13] A. Christopoulos, N. Pellas, and M.-J. Laakso, "A Learning Analytics Theoretical Framework for STEM Education Virtual Reality Applications," Education Sciences, vol. 10, no. 11, p. 317, Nov. 2020, doi: 10.3390/educsci10110317

[14] N. R. Carvalho and L. S. Barbosa, "Deep Learning Powered Question-Answering Framework for Organizations Digital Transformation," Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance, Apr. 2019, doi: 10.1145/3326365.3326375.

[15] C. Waghmare, "Introducing Azure Bot Service," 2019, doi: 10.1007/978-1-4842-4888-1.

[16] V. Damnjanovic, "Entry Market Strategy for Weaver Chatbot using the Digital B2B Model," 2019 International Conference on Artificial Intelligence: Applications and Innovations (IC-AIAI), Sep. 2019, doi: 10.1109/Ic-aiai48757.2019.00014.

[17] D. Kaczorowska-Spychalska, "How chatbots influence marketing," Management, vol. 23, no. 1, pp. 251–270, Jun. 2019, doi: 10.2478/manment-2019-0015.

[18] A. López, J. Sànchez-Ferreres, J. Carmona, and L. Padró, "From Process Models to Chatbots," Lecture Notes in Computer Science, pp. 383–398, 2019, doi: 10.1007/978-3-030-21290-2_24.

[19] Al-Ghobesi, A. A. H. (2025). Risks of Relying on Artificial Intelligence in Learning Arabic Language Sciences Through the Meta Application. *Arts for Linguistic & Literary Studies*, *7*(1), 396–419. https://doi.org/10.53286/arts.v7i1.2420

[20] Omer, N. I. M. (2024). Maintaining Meaningful Human Interaction in AI-Enhanced Language Learning Environments: A Systematic Review. *Arts for Linguistic & Literary Studies*, *6*(3), 533–552. https://doi.org/10.53286/arts.v6i3.2083

مجلــة العلــوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

Improving Learning and Teaching Methods Using Chatgpt in Higher Education: A Case Study of IBB University

Muneer Alsurori[1], Ayedh Abdulaziz Mohsen[1,2], Hezam Gawbah[1], Soliman Albaraky[1], Shaima Abdulkreem[1]

[1]Department of Computer Sciences & Information Technology IBB University, Yemen.

[2]Department of IT, Alqalam University, Yemen.

msurory@yahoo.com, ayedh992001@hotmail.com, hezammaster2022@gmail.com, Solimanalbarmaky2018@gmail.com, Shaimaalabasi@gmail.com

Abstract:

ChatGPT's impressive natural language generation capabilities represent a major advance in practical applications of AI-generated information. The potential use cases for ChatGPT are wide-ranging, including assessment setup, translation tasks, and custom source code creation. It can also handle more complex aspects of scientific writing, such as summarizing the literature and paraphrasing text. The difficulties and challenges of using ChatGPT in learning and teaching methods vary from one society to another and from one country to another, and there are many factors. Therefore, this paper aims to investigate the possibility of improving learning and teaching methods using ChatGPT in higher education. It presents the main challenges and notable opportunities that arise from the introduction of ChatGPT in the context of higher education. In line with the aim, the question guiding the study is: "How can learning and teaching methods be improved using ChatGPT in higher education?" To answer this question, a qualitative exploratory case study was conducted at the Department of Information Technology, Faculty of Science, IBB University. An electronic questionnaire containing closed questions was sent to all students. 93 responses were obtained via email, which constituted the sample of participants in this study. The primary contribution of this study is to provide a proposal on how to effectively integrate ChatGPT into higher education and improve learning and teaching methods in key areas.

Keywords: ChatGPT, Higher Education, Learning Methods, Teaching Methods, Artificial Intelligence, Natural Language Generation.

1. Introduction

Since the release of ChatGPT in November 2022, the field of higher education has quickly recognized the potential of generative AI in enhancing various traditional tasks (Omer, 2024, & Al-Ghobesi, 2025, Ahmed, 2025) [2]. The remarkable natural language capabilities of ChatGPT have revolutionized the practical applications of AI-generated information. Its use in the realm of higher education are extensive, ranging from assessment preparation and translation tasks to the generation of customized source code [3]. Furthermore, it can address more complex aspects of scientific writing, such as literature summarization and text paraphrasing [4]. Given the impressive natural language generation capabilities of ChatGPT, it represents a significant advancement in the practical applications of AI-generated information [6]. However, the problem with learning and teaching methods is the challenges and difficulties associated with using ChatGPT which may vary across societies and countries, and are affected by different factors.

This paper aims to foster a discussion on the possibility of improving learning and teaching methods in higher education through the integration of ChatGPT. To address this objective, a qualitative exploratory case study was conducted in the Department of Information Technology at IBB University. A questionnaire containing closed questions was distributed to all students, and a total of 94 responses were collected via email, forming the sample for this study. By identifying the main challenges and notable opportunities arising from the introduction of ChatGPT in the higher education context, this study seeks to provide valuable insights for effectively integrating ChatGPT into higher education and enhancing learning and teaching methods in key areas.

The primary contribution of this research is to provide suggestions on the effective incorporation of ChatGPT into higher education to enhance learning and teaching methods in key areas.

2. Review of the Related Literature

In [1], the much-touted artificial intelligence (AI) is yet to attain a level of intelligence commensurate with these claims. GPT-4 and its predecessor demonstrate superior performance, while Bing Chat and Bard resemble underperforming students with consistently failing grades. This article presents four distinct categories of recommendations tailored to key stakeholders in higher education. In [2], the aim was to

determine how the use of ChatGPT on the digitalized teaching system among university students in Peru. Descriptive statistics and linear regression analysis were utilized to analyze the data collected randomly from 216 students' responses on the Twitter website regarding their experiences with ChatGPT. The findings demonstrate that ChatGPT has a significant influence on the digitalized learning process, as a substantial number of students express a preference for utilizing ChatGPT to fulfill various academic tasks. In [3], drawing upon the theoretical framework of creative writing, this scholarly article endeavors to examine the ChatGPT system and its influence on the dearth of creativity observed in students' writing abilities. Employing qualitative methodologies, the study employs library research as the primary data collection approach, analyzing scientific journals and other pertinent articles to facilitate an in-depth discussion on the topic at hand. In [4], examined ChatGPT's potential to take the position of instructors in the classroom teaching process in great detail. The nature of research is qualitative. Careful documentation of pertinent information and attentive listening were key components of the data collection process. The material was then submitted to analytical processes such data reduction, data presentation, and conclusion drafting. The study came to the conclusion that technology can only be used as a tool when it comes to employing ChatGPT in teaching. In [5], the central questions and problem statements pertaining to the utilization of ChatGPT for self-determined learning within higher education were examined. The paper intended to synthesize and critically evaluate the existing literature on the potential of ChatGPT to support self-directed and self-determined learning and highlight the main challenges and concerns associated with its use. The findings suggest that the progression of ChatGPT-based interventions for self-determined learning in higher education necessitates an intricate and interdisciplinary approach that takes into account the perspectives of educators, researchers, learners, and other relevant stakeholders. In [6], content analysis was used to examine 100 news articles about how ChatGPT is disrupting higher education, concentrating specifically on Australia, New Zealand, the United States, and the United Kingdom. The analysis aimed to explore various key themes including university responses, academic integrity concerns, the limitations and weaknesses of AI tool outputs, and opportunities for student learning. The data reveals mixed public discussion and university responses, mostly centered on issues related to academic integrity and creative possibilities for assessment design.

Table. (1) Review of the Related Literature.

Re.	Year	Method	Objectives	Data	Results
[1]	2023	Descriptive comparative	describe the comparative method and present a method of comparison of a few chosen chatbots on a multidisciplinary test that is pertinent to higher education.	Some of the most promising chatbots in the English and Chinese-language spaces	There are currently no A-students and no B-students in this bot cohort, despite all publicized and sensationalist claims to the contrary. GPT-4 and its predecessor did best, whilst Bing Chat and Bard were akin to at-risk students with F-grade averages.
[2]	2023	Descriptive statistics and linear regression analysis	To determine the influence of ChatGPT on the digitalized teaching system among Peruvian University students.	216 students' responses on the Twitter website on the various experiences they have of ChatGPT	The study concludes that ChatGPT significantly affects the digitalized learning process, as many students prefer to use ChatGPT to handle tasks.
[3]	2023	Qualitative	To discuss the ChatGPT system and its influence on pupils' lack of writing inventiveness.	Scientific journals and other articles relevant	ChatGPT, with its ability to provide answers according to the keywords entered by the user, can positively influence the world of teaching and learning.
[4]	2023	Qualitative	To explore the role of technology in the classroom learning process, particularly in terms of altering the teacher's position as the primary instructor.	Literature review	In the context of utilizing ChatGPT for teaching all purposes, it should be noted that technology can only be a tool and cannot replace the role of the teacher entirely. Therefore, it is necessary to integrate technology in learning in an appropriate and effective way and develop the competence of teachers in managing learning with technology.
[5]	2023	Analysis	To e examine the major research questions and problem statements concerning the usage of Chat GPT for self-determined learning in higher education.	Literature review	Advancement of ChatGPT-based interventions for self-determined learning in higher education necessitates a nuanced and multidisciplinary approach that considers the viewpoints of educators, researchers, learners, and other interested stakeholders.
[6]	2023	Analysis	To conduct a sentiment analysis about how ChatGPT is disrupting higher education.	News articles (N=100) about how ChatGPT	Relatively balanced in the number of times positive (n=912) and negative (n=1034) language was coded.
[7]	2023	Exploration	To explore the perspectives of students and educators on the implications of ChatGPT and AI integration in the	Responses of seven scholars and 14 PhD students from four countries – Turkey, Sweden, Canada and	Based on these findings, suggestions for future research include further exploration of the ethical implications of AI for teaching, the development of strategies to manage privacy concerns, and the investigation of how

Re.	Year	Method	Objectives	Data	Results
			context of universities	Australia	teaching al institutions can best prepare for the integration of AI technologies
[8]	2023	Qualitative	To explore the perceptions of educators and students on the use of ChatGPT in teaching during the digital era	Ten educators and 15 students from different academic institutions in Krabi, Thailand	The findings showed some concerns regarding the use of ChatGPT in teaching. Participants were worried about the accuracy of information provided by the chatbot and the potential loss of personal interaction with teachers. The need for privacy and data security was also raised as a significant concern
[9]	2023	-	To address the need for discussion of potential approaches for integrating ChatGPT into higher education	five out of 55 preprints and discussions and talks with other lecturers and researchers and took into account the authors' test results from using ChatGPT	A proposal for how to integrate ChatGPT into higher education in four main areas
[10]	2023	-	To examine the potential benefits and challenges of using the generative AI model, ChatGPT, in higher education, in the backdrop of the constructivist theory of learning	-	Integrating ChatGPT within higher education necessitates the establishment of a delicate equilibrium that addresses both the prevention of academic misconduct and the promotion of academic freedom and innovation, all while prioritizing the cultivation of essential graduate skills. By striking this balance, ChatGPT can emerge as a valuable instrument that enriches, rather than impedes, students' teaching al journeys

3. Method

This section provides a concise overview of the research methodology used in the study. A qualitative and descriptive statistics methodology was employed to identify the opportunities and challenges of implementing ChatGPT in higher education in Yemen. A qualitative exploratory case study was conducted in the Information Technology Department at IBB University. An electronic questionnaire with closed questions was distributed to all students, resulting in 94 responses as the participant sample. Data were analyzed using qualitative and descriptive statistical techniques. Measures were taken to ensure the study credibility included careful questionnaire design, participant confidentiality, and validation through multiple perspectives.

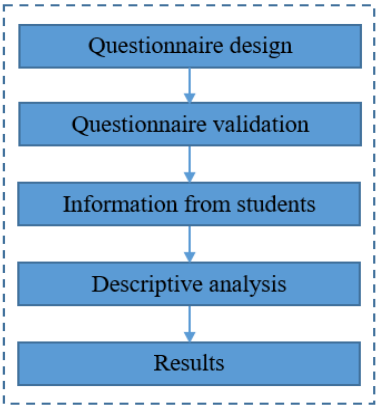


FIGURE 1: Method.

3.1. Questionnaire design

the questionnaire was designed and then divided into two axes: opportunities and threats, as in Figure 1.

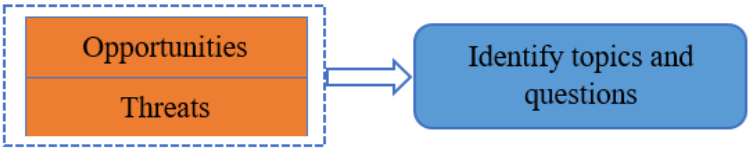


FIGURE 2: The stage of questionnaire design.

3.2. Questionnaire validation

The purpose of this verification is that the opportunities and challenges in the questionnaire can be used to reach a general conclusion regarding the variables that improve learning and teaching methods using ChatGPT in higher education, as in Figure 2.

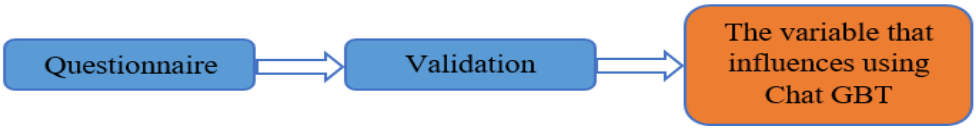


FIGURE 3: The stage of questionnaire validation.

3.3. Information from students

One of the crucial steps in exploring opportunities and threats in learning and teaching methods using ChatGPT in higher education is to collect information directly from academics. One way to collect data

is to use a questionnaire, as shown in Figure 3. The questionnaire focuses on two main axes: opportunities and threats. 94 students participated in the survey, providing valuable insights into improving learning and teaching methods using ChatGPT in higher education.

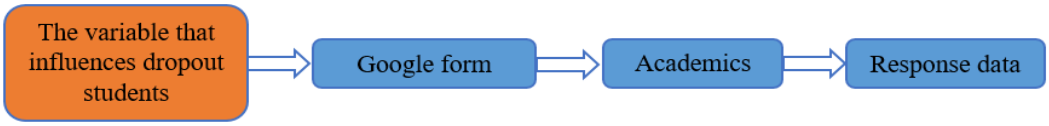


FIGURE 4: The stage of seeking information from students.

3.4. Descriptive analysis

Descriptive analysis is shown in Figure 4. In this step, we apply descriptive analysis after preparing the data collected from the students using SPSS to obtain the results.

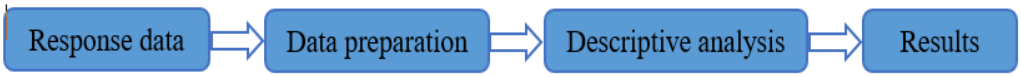


FIGURE 5: The stage of descriptive analysis.

4. Result

The research question guiding this study was "How can learning and teaching methods be improved using ChatGPT in higher education?" A study design was used to explore the opportunities offered by ChatGPT and the challenges that arise from its use with the aim of improving learning and teaching methods in higher education. Data were collected from 93 participants from the Department of Information Technology, University IBB using a questionnaire that included closed questions. After collecting the questionnaires, they were reviewed to ensure their validity and suitability for statistical analysis. They were computer tabulated and formatted using the Statistical Package for the Social Sciences (SPSS) program.

In the analysis, the researcher relied on a set of statistical methods for the aforementioned questionnaire topics, including:

- Descriptive analysis methods represented by frequencies and percentages of the main variables and factors for identifying the general demographic characteristics of the research sample. Arithmetic averages (Means) and standard deviations (Std. Deviation) were also used in order to answer questions related to graduated scales (Scales) to know the students' answers to the questionnaire statements to determine the main reasons for dropout.

After tabulating the data, copying it to the computer, and processing it statistically, the results of the statistical analysis showed the reliability of the questionnaire questions (honesty). A three-point Likert scale was used in its design. It has also been proven that this measure is stable and does not contradict itself, meaning that the measure gives the same results with a probability equal to the value of the coefficient if it is re-applied to the same sample. The total Cronbach's alpha value for the first axis (opportunities) is equal to 75%, while for the second axis (challenges), the Cronbach's alpha value is equal to 64%.

4.1. Opportunities

Table. (2) the statistical analysis of the first axis, "Opportunities".

No.	Opportunities	Agree	Disagree	Neutral	Mean	Std. De	Ranking
1	A free resource available to everyone around the clock	75.3%	16.1%	8%	1.3333	.63131	7
2	Easy to use and handle	92.5%	2.2%	5.4%	1.1290	.47149	1
3	It contributes to overcoming the disparity in capabilities and resources between teaching al institutions	55.9%	11.8%	32.3%	1.7634	.91364	13
4	It has an extensive database of learning and teaching resources	89.2%	3.2%	7.5%	1.1828	.55062	2
5	It is used to determine your level for the course or scientific field	33.3%	38.7	28.0	1.9462	.78527	16
6	It provides the learner with clear information and increases his motivation to learn	74.2%	10.8%	15.1%	1.4086	.74069	8
7	It facilitates and understands the application of theories, rules and laws	75.3%	7.5%	17.2%	1.4194	.77069	6
8	It helps teachers prepare for scientific content to save time and effort	79.6%	7.5%	12.9%	1.3333	.69678	4
9	It trains the learner to employ information and practice skills, making learning lasting	62.4%	18.3%	19.4%	1.5699	.79943	12
10	It has an important and effective role in solving the problems of guidance and counseling for learners	68.8%	8.6%	22.6%	1.5376	.84131	9
11	It helps in correcting grammatical and spelling errors and finding linguistic synonyms	67.7%	10.8%	21.5%	1.5376	.82829	10
12	Helps manage classroom logistics and announcements including scheduling, alerts, wording announcements, and even setting appointments	36.6%	30.1%	33.3%	1.9677	.83992	15

No.	Opportunities	Agree	Disagree	Neutral	Mean	Std. De	Ranking
13	It is used to suggest effective strategies to improve communication between the teacher and the student	55.9%	22.6%	21.5%	1.6559	.81420	14
14	Used to help you organize information	77.4%	11.8%	10.8%	1.3333	.66485	5
15	Used to search and collect relevant content and links for your lessons	83.9%	7.5%	8.6%	1.2473	.60174	3
16	It helps in developing teaching al courses in line with modern knowledge and advanced technological mechanisms	64.5%	12.9%	22.6%	1.5806	.83825	11

From Table (2), the results of the statistical analysis of the opportunities indicate that the opportunity ranked 1 improving learning and teaching methods using ChatGPT in higher education by 92.5%, with an arithmetic mean (1.1290) and a standard deviation (.47149). The opportunity ranked 2 improving learning and teaching methods using ChatGPT in higher education by 89.2%, with a mean (1.1828) and standard deviation (.55062). The opportunity ranked 3 improving learning and teaching methods using ChatGPT in higher education by 83.9%, with a mean (1.2473) and standard deviation (.60174). The rest of the opportunities are as shown in the table according to the ranking column.

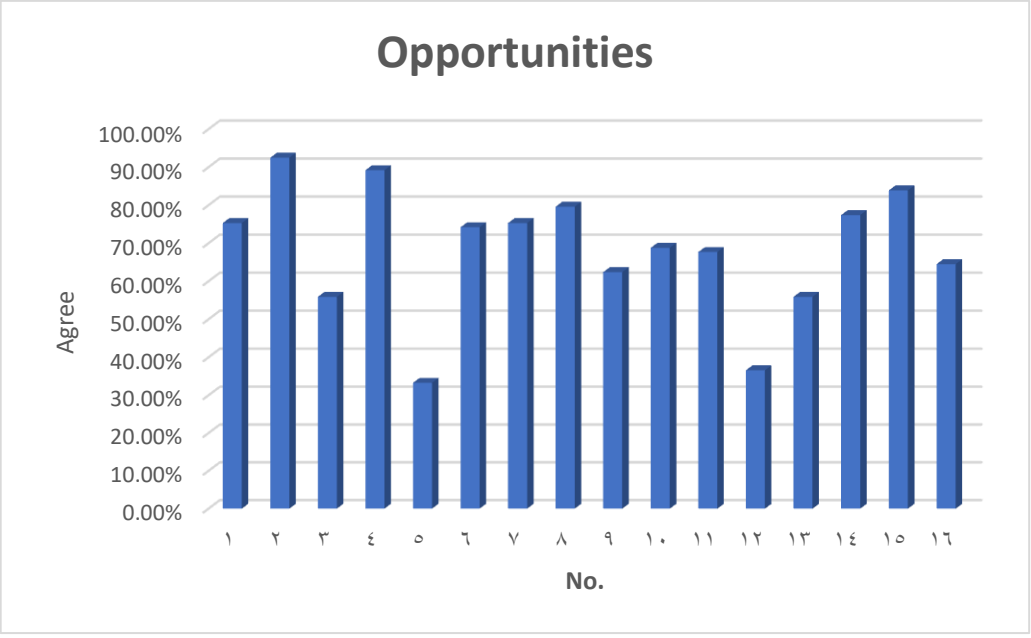


FIGURE 6: the agree rate of the first axis, "Opportunities".

4.2. Threats

Table. (3) the statistical analysis of the second axis, "Threats".

No.	Threats	Agree	Disagree	Neutral	Mean	Std. De	Ranking
1	It threatens the originality of ideas and scientific rights and increases the possibility of plagiarism	65.6%	20.4%	14.0%	1.4839	.73148	4
2	Concerns about breaching user privacy	58.1%	14.0%	28.0%	1.6989	.88201	5
3	Difficulty in assessing students' knowledge using the human mind if students submit assignments created by using ChatGPT	73.1%	11.8%	15.1%	1.4194	.74195	1
4	There is data misuse when used to generate a large amount of data from unknown sources	65.6%	17.2%	17.2%	1.5161	.77478	3
5	Often he did not succeed in providing an accurate answer, but rather provided false and illogical information	65.6%	15.1%	19.4%	1.5376	.80162	2
6	Answering religious questions and fatwas	25.8%	44.1%	30.1%	2.0430	.75056	6

As shown in Table (3), the results of the statistical analysis of the threats indicate that the threat ranked 1 improving learning and teaching methods using ChatGPT in higher education by 73.1%, with an arithmetic mean (1.4194) and a standard deviation (.74195). The threat ranked 2 improving learning and teaching methods using ChatGPT in higher education by 65.6%, with a mean (1.5376) and standard deviation (.80162). The threat ranked 3 improving learning and teaching methods using ChatGPT in higher education by 65.6%, with a mean (1.5376) and standard deviation (.80162). The rest of the threats are as shown in the table according to the ranking column.

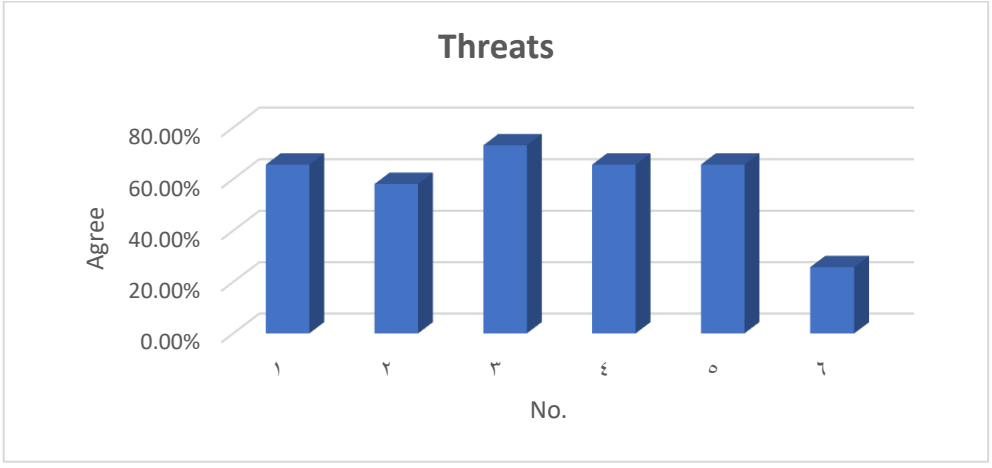


FIGURE 7: the agree rate of the second axis, "Threats".

5. Recommendations

In accordance with the findings of this study, and in the context using of ChatGPT, the study offers some significant recommendations for improving learning and teaching methods using ChatGPT in higher education as follows:

- Banning or prohibiting

Banning or prohibiting ChatGPT will not help. Such action may frustrate instructors, learners, and regulators. It is better to embrace it and set clear ground rules for its use in higher education teaching.

- Educators

For educators, it is going to be as transformational as Google was in 1998, and requires a serious conversation about the benefits, challenges and implications for schools and learners. The future will be changed by it indelibly. Educators have to start engaging with it in a meaningful way through peer support and mentoring for faculty members to increase skill level and share good practices for teaching and ways of using ChatGPT in research

- Oversight, transparency, and assessment

ChatGPT still constitutes a real threat to conventional assessment models. We need more oversight and transparency to discern human-created content from AI-created content. It is clear that future models of assessment will need to prioritize independent and critical thinking, deductive reasoning, creative thinking, and questioning and validating data inputs.

- Academic integrity

Examine and revise academic integrity/honesty rules in connection to ChatGPT and other AI tools.

- Teaching in higher education

Universities should integrate artificial intelligence into their curricula to enhance learning quality and prepare students for the ever-changing technological world. This includes incorporating AI into existing curricula and developing new courses focusing on technology, enabling students to acquire transferable skills to effectively utilize and manage AI tools.

- Self-learning

Universities must teach students the skills needed to tackle the challenges posed by ChatGPT/AI, sometimes the information he had learned could be wrong, or it may not be accurate or complete.

- Prompts (How to question AI?)

To produce a more relevant result, provide ChatGPT with prompts about how you want it to respond. You also need to keep in mind that the answers provided by generative AI need to be reviewed and checked, as many tools are still being trialed or have fewer features than the paid versions. teachers, researchers, and students should be trained to improve the queries they pose to ChatGPT.

- Build capacity to understand and manage ChatGPT

For increasing ability to comprehend and handle ChatGPT, adapting to higher education in the ChatGPT era necessitates that HEIs consider their role in developing ability to comprehend and manage ChatGPT and AI through new programmes, workshops, seminars and courses that focus on ChatGPT.

1. Conclusion

While incorporating ChatGPT technology into learning and teaching can have immense benefits, it is crucial to remember that technology should only serve as a tool and not to take the place of a teacher's role entirely. As such, it is essential for both tertiary educators and students to handle this technology with caution and use it ethically, reliably, and effectively for academic purposes. Higher education institutions bear the responsibility of teaching students about the responsible and ethical use of ChatGPT and other generative AI tools to ensure its accurate implementation. In addition, educators can come up with innovative assessment methods that ChatGPT cannot easily replicate, such as evaluating the learning process rather than just the end result.

On a more positive note, we are optimistic about incorporating ChatGPT or similar AI tools in the future to automate certain tasks and processes such as study guides, summaries, concept maps, and practice tests. This will allow for a shift in focus towards the human aspect of teaching, moving away from reliance on memorization and instead emphasizing the acquisition of enduring skills necessary for the future workforce. As we have already witnessed, by utilizing ChatGPT in the right way, we can unlock its full potential and enhance the educational process.

References

[1] Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, *6*(1).

[2] Castillo, A. G. R., Silva, G. J. S., Arocutipa, J. P. F., Berrios, H. Q., Rodriguez, M. A. M., Reyes, G. Y., ... & Arias-Gonzáles, J. L. (2023). Effect of Chat GPT on the digitized learning process of university students. *Journal of Namibian Studies: History Politics Culture*, *33*, 1-15.

[3] Shidiq, M. (2023, May). The use of artificial intelligence-based chat-gpt and its challenges for the world of teaching; from the viewpoint of the development of creative writing skills. In *Proceeding of International Conference on Teaching, Society and Humanity* (Vol. 1, No. 1, pp. 353-357).

[4] Ausat, A. M. A., Massang, B., Efendi, M., Nofirman, N., & Riady, Y. (2023). Can ChatGPT replace the role of the teacher in the classroom: A fundamental analysis. *Journal on Teaching*, *5*(4), 16100-16106.

[5] Baskara, F. R. (2023, May). The Promises and Pitfalls of Using Chat GPT for Self-Determined Learning in Higher education: An Argumentative Review. In *Prosiding Seminar Nasional Fakultas Tarbiyah dan Ilmu Keguruan IAIM Sinjai* (Vol. 2, pp. 95-101).

[6] Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning.

[7] Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*, *6*(1).

[8] Limna, P., Kraiwanit, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and teaching*, *6*(1).

[9] Neumann, M., Rauschenberger, M., & Schön, E. M. (2023). "We Need to Talk about ChatGPT": The Future of AI and Higher education.

[10] Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., ... & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, *6*(1).

[11] Ahmed, M. R. A. (2025). Accreditation and Quality Assurance: Exploring Impact and Assessing Institutional Change in the US and Saudi Arabian Higher Education Institutions. *Arts for Linguistic & Literary Studies*, *7*(1), 626–639. https://doi.org/10.53286/arts.v7i1.2419

[12] Al-Ghobesi, A. A. H. (2025). Risks of Relying on Artificial Intelligence in Learning Arabic Language Sciences Through the Meta Application. *Arts for Linguistic & Literary Studies*, *7*(1), 396–419. https://doi.org/10.53286/arts.v7i1.2420

[13] Omer, N. I. M. (2024). Maintaining Meaningful Human Interaction in AI-Enhanced Language Learning Environments: A Systematic Review . *Arts for Linguistic & Literary Studies*, *6*(3), 533–552. https://doi.org/10.53286/arts.v6i3.2083

مجلــة العلوم
الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

The Use of Artificial Intelligence in Records Management at Higher Education Institutions: An Analytical Theoretical Study

Chemaa souheir
souheir.chemaa@univ-constantine2.dz
Kamal Boukerzaza
kamel.boukerzaza@univ-constantine2.dz
University of Constantine 2 Abdelhamid Mehri. Algeria

Abstract:

This analytical theoretical study explores the transformative impact of artificial intelligence (AI) on records management within higher education institutions, focusing on document handling and administrative operations. This scientific paper aims to elucidate the potential benefits derived from AI applications in modernizing document management processes and controlling the document life cycle within educational institutions. Employing an analytical approach and drawing on various scientific research, the study emphasizes the importance of harnessing AI technology in this context. It reveals significant advantages, such as streamlined workflows and enhanced student experiences, underscoring AI's pivotal role in reshaping the records management landscape in higher education.

Key words: artificial intelligence (AI), records management, higher education institutions

Introduction

The incorporation of cutting-edge technologies becomes essential in the constantly changing world of information management. Artificial Intelligence (AI) stands out among these technologies as a driver for transformation with enormous potential. In higher education institutions, the nexus between AI and records management is of special relevance. This scholarly investigation aims to clarify the theoretical foundations and analytical aspects of using AI to records management in the context of higher education.

There are persistent inherent challenges in traditional approaches to records management within academic institutions. The rapid increase in digital information combined with the complex nature of institutional records frequently leads to inefficiencies in retrieving, categorizing and preserving such records. Acknowledging these difficulties is important for understanding how artificial intelligence applications could impact higher education records management. the ability of AI to adapt and evolve through continuous learning aligns seamlessly with the dynamic nature of information within academic settings. However, as we navigate this intersection, questions arise regarding the ethical implications, the need for transparent governance, and the balance between automation and human oversight in records management processes. These considerations underscore the complexity of the theoretical landscape we aim to explore, providing a foundation for a nuanced understanding of AI's impact on higher education records management.

The Critical Importance:

This study is critical because it endeavors to illuminate the theoretical underpinnings and analytical dimensions of integrating AI in records management within higher education institutions. By exploring the potential benefits of AI applications. The results hold the prospect of advancing both the field's theoretical understanding and offering useful suggestions for higher education records management decision-makers.

The Objective:

This study intends to bridge the gap between traditional practices and the revolutionary potential of AI by providing a balanced perspective on the importance of using such technology in records management for the benefit of post-secondary education environments. It also seeks to offer valuable insights for enhancing the efficiency, accuracy, and strategic utilization of records in academic settings.

Problem Statement:

The strategic incorporation of AI in records management processes has the potential to revolutionize how information is handled, assessed, and leveraged within academic institutions. It offers unprecedented efficiency in sorting, categorizing, and safeguarding records, thereby influencing the overall operational dynamics of higher education organizations. However, as we delve into this promising realm, a critical

problem emerges: How beneficial can the implementation of AI be in enhancing records management practices for higher education institutions?

1. literature review:

Study 1: Mashilo Thomas, M. (2021). *Utilizing artificial intelligence technology for the management of records at the Council for Scientific and Industrial Research in South Africa.* [PHD Thesis].

Study 2: Fayaz Ahmad, S., Alam, M., Rahmat, M., Mubarik, M., & Hyder, S. (2022). Academic and Administrative Role of Artificial Intelligence in Education. *Sustainability.* [Article].

	Purpose	Methodology	Results/Features	Disadvantages
Study 1	Investigating the use of AIT for records management at the CSIR in South Africa	Convergent mixed methods with parallel sampling of participants utilized multiple data collection techniques such as: interviews, questionnaire, and document analysis	the key results: -Review found ineffective records management at CSIR. -AI would enhance digital records management. -Cloud storage and robotic machines recommended. The key features: -Reviewed current records management system at CSIR using system analysis. -Integrated quantitative and qualitative findings to reveal ineffective records management. -Recommended AI technology framework to improve CSIR records management.	- Small sample size of only 8 participants limits generalizability of the results. A larger and more diverse sample could yield more robust findings. - Does not provide granular technical guidance on designing the AI systems proposed. More implementation details are needed.
Study 2	Explores academic and administrative applications of AI that can be in education.	Literature review and conceptual analysis.	The key results: - AI reduces administrative workload for teachers. - AI has significant potential in education. The key features: - Analyzes how AI can assist with both academic activities and administrative tasks -Assesses the potential benefits of AI in enhancing student learning, reducing teacher workload, and improving educational efficiency.	-It has not been strengthened by some examples of implementation. -Lacks concrete guidance on challenges of integrating AI in records management.

In comparing the two previous studies with our current study, we find that the latter has addressed the subject of artificial intelligence in higher education from a different angle. It has focused on the field of administration, specifically on the adoption of artificial intelligence in records management at higher

education institutions, which will positively impact the educational process and the outputs of higher education institutions, as well as their reputation, if this modern technology is intelligently employed in the area of records management. Subsequently, this could enable advanced control over information flows.

2. Theoretical concepts:

2.1. Artificial Intelligence (AI): Odlis dictionary defines AI as**:** "Mechanical and electronic devices and applications designed to closely mimic the human ability to learn, reason, and make decisions. AI is used in voice recognition technology, expert systems, natural language and foreign language processing, and robotics."(Reitz, 2014, Omer, 2024, Al-Ghobesi, 2025, Ahmed, 2025)

AI is defined as a branch of computer science that focuses on the creation of intelligent machines that function similarly to HI.(Mashilo Thomas, 2021)

In line with the above definitions, we understand that: AI encompasses mechanical and electronic devices and applications intricately designed to replicate human cognitive functions, including learning, reasoning, and decision-making. As a pivotal branch of computer science, AI is fundamentally concerned with developing intelligent machines that emulate the functionalities of Human Intelligence (HI). This involves the creation and implementation of algorithms and systems capable of cognitive processes, contributing to the advancement of technology in various fields.

2.2. Records Management: is the branch of management that is in charge of the competent and systematic control of the creation, reception, preservation, use, and disposition of records, as well as the procedures for understanding and preserving evidence of and information about commercial functions and transactions in the records process. Records management is based on notions such as the records life cycle and the records continuum.(Mashilo Thomas, 2022)

According to this definition, we can say that records management constitutes a discipline within management practices that assumes responsibility for the proficient and methodical oversight of the entire life cycle of records. This encompasses the systematic control of records from their creation and reception through to preservation, utilization, and eventual disposition.

3. The use of AI in Administration Records Management:

AI has significant promise for use in the administration management of records. The evolution of AI-assisted text production, represented by for example GPT-3, allows the algorithm to emulate domain-

specific writing styles, assisting in the synthesis of legal writings. Second, machine-learning systems that use semantic algorithms look for rules in legal texts autonomously, presenting and applying them in an intelligible manner. These algorithms can forecast choices for administrative procedures by using metadata and fact-of-the-case assessments.(Parycek et al., 2023, Mleiki, 2025)

AI can be utilized to manage records for the administration in a variety of ways, including:

- Automated classification and categorization: AI can assess and categorize records based on their content, information, and context. This aids in more efficiently arranging and categorizing records, making them easier to obtain and manage.
- Intelligent search and retrieval: Search algorithms enabled by AI can interpret natural language queries and return more accurate and relevant search results. This allows users to find the records they need fast, saving time and effort.
- Data extraction and indexing: AI can automatically extract relevant information from records and generate indexes, making it easier to find specific data within a record or across numerous records.
- Automated retention and disposal: AI can help determine the proper retention periods for various types of records depending on legal and regulatory requirements. It can also identify records that are eligible for destruction, ensuring that retention requirements are followed.
- Data privacy and security: AI can aid in the identification and protection of sensitive information inside records, such as personally identifiable information (PII) or confidential data. It can detect and highlight potential privacy or security breaches, hence improving data protection.
- Predictive analytics: AI can evaluate patterns and trends within records to generate insights and predictions through predictive analytics. Making educated decisions, detecting potential dangers, and streamlining records management processes can all benefit from this.(Rolan et al., 2019)

Comprehensively, by automating, AI can dramatically increase the efficiency, accuracy, and efficacy of records administration. So, there are various uses of AI in records administration. To begin, AI, as represented by GPT-3, mimics writing styles, assisting in the synthesis of legal papers. Furthermore, machine-learning systems that use semantic algorithms read legal documents autonomously, presenting regulations in a comprehensible manner and forecasting administrative process alternatives. AI also has practical applications in records management, including automated classification, intelligent search, data extraction, automated retention, privacy and security enhancement, and predictive analytics. These features improve efficiency, accessibility, compliance, and decision-making in records management processes.

4. **The Importance of AI in Records Managements at Higher Educational Institutions:**

Document management is one area where artificial intelligence has the potential to significantly improve education. Administrators and teachers are increasingly confronted with a deluge of data, documentation, and student records. All of this information must be organized and secured. Student information privacy is mandated by law; therefore, document management solutions must be fully compliant with all relevant legislation.

These documents can be managed much more efficiently with the help of artificial intelligence. AI-powered document management solutions can handle massive volumes of content thanks to the power of machine learning – where the computer learns and improves as it works. This technology also allows for improved reporting. AI systems can find trends and patterns in data by watching it.(Kelley Connect, 2020, Alasmari, 2023)

The growth of technology and as information is scattered across different systems, the proliferation of technology and data on college campuses confronts obstacles. AI tackles this issue by making campuses more linked, allowing institutions to access data from many platforms. AI applications aid in refining course offerings and increasing student retention. As data from numerous campus systems is combined, AI-driven "hyper-learning" opportunities emerge, facilitating informed decision-making. Overall, AI improves the effectiveness of student success tools and adds to better higher education decision-making. (Klutka et al., n.d.)

Higher education institutions can benefit from AI in administrative record management in the following ways:

- Improved Efficiency: AI can automate time-consuming administrative operations like data entry, document management, and record retrieval. Administrative staff saves time, allowing them to focus on more strategic and value-added initiatives.
- Intelligent Document Classification: AI algorithms can classify and categorize many sorts of documents, making record organization and retrieval easier. This improves record management accuracy and speed.
- Improved Accuracy: AI algorithms can reduce errors caused by manual data entry and record maintenance. This enhances the correctness and dependability of administrative records, lowering the likelihood of data discrepancies or errors.

- Streamlined Processes: By automating regular operations like record retention and disposal, document classification, and search and retrieval, AI can optimize administrative workflows. This improves operating efficiency while decreasing administrative load.
- Data-Driven Decision Making: AI-powered analytics can mine vast amounts of administrative data for important insights and patterns. This enables higher education institutions to make data-driven decisions including finding patterns in student performance, forecasting enrollment trends, and optimizing resource allocation.
- Streamlined Workflows: AI can automate common administrative operations. This allows administrative personnel to concentrate on more difficult and value-added operations.
- Enhanced Student Experience: AI applications in administrative record management can enhance the student experience by delivering individualized services and assistance. AI-powered chatbots, for example, can assist students with questions, admissions processes, and course selection, increasing their engagement.(Fayaz Ahmad et al., 2022)

Accordingly, we can conclude that the integration of AI in document management is pivotal for higher education institutions. It not only handles data volume, but it also improves efficiency and decision-making. AI's significance in connecting campuses, streamlining procedures, and providing individualized student experiences highlights the technology's critical role in modern educational record management.

5. Initiatives for AI in Higher Education Records Management:

Higher education institutions and technology providers have recognized the vast potential of AI for transforming records management and have spearheaded various initiatives to promote adoption. Some notable examples include:

The Digital Records Project by MIT Libraries: This project explores AI for automating the appraisal, arrangement, description and delivery of archival collections. Computer vision and NLP are applied to extract information from records and automate processing. The goal is to expand access to archival materials at lower costs. (MIT Libraries, n.d.)

Jisc Intelligent Campus Program: Bolton College implemented the Ada Chabot integrated with internal systems to offer tailored responses on academic issues to students and staff; inspired by this, Jisc's national AI center piloted an in-house developed Chabot across four UK colleges to respond user inquiries by

mimicking Ada's customized design. This automates responses for frequently asked questions, directing only complex inquiries to staff. Early pilots indicate considerable time savings. (Jisc, 2022)

UC Office of the President's Intelligent Process Automation: This University of California system-wide initiative focuses on robotic process automation, machine learning and AI for streamlining administrative tasks. use cases include automated student communications for enrollment, graduation etc. and simplifying procurement workflows. (UC Presidential Working Group on AI, 2021)

Ellucian Ethos Platform: Ethos integrates AI and analytics into Ellucian's student information systems for higher education. It provides predictive insights on student engagement, success and outcomes while optimizing workflows like admissions. Clients have achieved up to 35% efficiency gains using Ethos. (Ellucian, 2023)

6. Standards for AI Adoption in Higher Education Records Management:

The adoption of standards and best practices is critical for ensuring the ethical, transparent and accountable use of AI for records management in higher education. Some of the relevant standards include:

ISO Records Management Standards: ISO 15489 provides guidance for records management principles and ISO 23081 deals with metadata for records. Compliance with these standards enables the effective application of AI for records classification, retention and disposition. (ISO, 2016)

NIST AI Risk Management Framework: NIST outlines processes for AI trustworthiness covering areas like data governance, model documentation, bias testing etc. Adhering to such risk management best practices is vital when deploying AI algorithms for sensitive tasks like student record processing. (Tabassi, 2023)

Diversity, Equity and Inclusion Principles: DEI guidelines, like the ones from ACM and OECD, provide standards for mitigating algorithmic bias and ensuring just outcomes, especially for marginalized communities. Following such principles helps develop ethical AI systems for higher education. (OECD, 2019)

EU Ethics Guidelines for Trustworthy AI: Formulated by the EU's High-Level Expert Group, these principles champion values-based AI that is lawful, ethical and robust. Implementing transparency, oversight and accountability measures as recommended facilitates responsible use of AI for records management. (EU, 2019)

Compliance with pertinent records management, archival, risk management and ethical AI standards allows higher education institutions to responsibly unlock the advantages of AI while addressing associated risks and limitations.

7. Obstacles to AI Adoption in Higher Education Records Management:

While AI has immense potential to transform higher education records management, there are also notable obstacles impeding rapid and widespread adoption:

Data Privacy Concerns: Apprehension persists about the privacy risks with allowing AI systems to access student records containing sensitive personal data like grades, disabilities, finances etc. Developing solutions compliant with regulations like FERPA is key. (Popenici & Kerr, 2017)

Bias and Fairness Challenges: Since AI algorithms can perpetuate and amplify existing societal biases, their use for administrative decision-making raises concerns about equitable treatment, especially for marginalized student groups. Ongoing biases need redressal. (Green et al., 2022)

Lack of Skills and Understanding: Most higher education records management staff lack fluency with AI technologies. Substantial training and change management is imperative for successful adoption within administrative teams. (Webb et al., 2021)

Costs and Infrastructure Requirements: Migrating paper-based processes to AI-managed electronic systems requires major technology investments in digitization, storage, software and integration. Ongoing costs like vendor fees and subscriptions are also considerable. (Khan & Al-Yasiri, 2016)

Resistance to Workforce Disruption: As AI takes over repetitive administrative tasks, higher education staff may perceive job loss risks and hence resist transitioning to automated systems, hampering adoption. Leadership must proactively address such workforce concerns. (Clay, 2018)

Algorithmic Transparency Issues: The 'black box' nature of some AI systems obscures how they arrive at outputs. Such opacity heightens accountability risks surrounding administrative decisions. Human oversight and explainable AI methods are essential. (Robertson & Wagner, 2021)

Cybersecurity Vulnerabilities: Like other software systems, AI applications used for records management can be susceptible to hacking, ransomware and data leaks. Robust cybersecurity protections are paramount. (Horizon, 2022)

Overcoming these barriers necessitates multifaceted strategies encompassing ethics, skills development, change management, infrastructure upgrades, cybersecurity investments and governance reforms. With deliberate efforts, AI's benefits can be secured while addressing the risks and limitations impeding adoption.

8.**The Results:**

- ✓ The study showed that artificial intelligence can be exploited in several records management processes.
- ✓ Artificial intelligence has the ability to improve and develop the process of making good use of documents.
- ✓ Artificial intelligence provides several techniques that can be used to control and manage a wealth of information, processing and managing documents.
- ✓ The reliance of higher education institutions on artificial intelligence techniques to manage their documents contributes to helping administrative employees organize their administrative tasks.
- ✓ The use of artificial intelligence in document management enables higher education institutions to develop their administrative side and provide better services, whether to their administrative staff or to students.
- ✓ Artificial intelligence can be used to manage the documents of higher education institutions in several processes, such as registering students and controlling their files. Ease of retrieving information and documents, which helps in decision-making, organizing programs and completing routine work.
- ✓ Exploiting technology to develop the administrative side of higher education institutions helps in developing their outcomes and increases their visibility.
- ✓ The study proves that using AI in RM is important for higher education institutions development especially at this digital age.

9. Recommendations:

- ➢ All institutions in general and higher education institutions in particular should place great importance on developing the administrative aspect and working to keep pace with the changes of the times.
- ➢ Using modern technology to develop university administration and services.
- ➢ Ensure to give greater importance to the field of document management and work to improve it.

- Working on using artificial intelligence technology in document management for higher education institutions in order to develop their administrative aspect
- It is necessary to work on training administrators and document specialists to deal with artificial intelligence technology to develop administrative services
- Ensuring keeping pace with the development of the times and making good use of modern technology for the benefit of higher education institutions
- Making good use of artificial intelligence technology to carry out document management tasks more effectively to provide services to students
- Ensuring the proper management of information and documents of the higher education institution in order to preserve its institutional memory
- Invest in digitization initiatives to convert paper-based records into digital formats that can be incorporated into AI systems. This builds the data foundation needed for AI-enabled records management.
- Start with pilot projects focused on targeted use cases like automating student data management. Validate value delivery before expanding AI to other processes.
- Leverage cloud solutions and vendors to offset upfront infrastructure investments. However, retain ownership of data and AI models.
- Comply with standards like ISO records management principles, SAA guidelines for AI-based archiving, and AI ethics best practices.
- Promote explainable and transparent AI where administrators can understand how algorithms arrive at outputs affecting students.

Conclusion:

In conclusion, this analytical theoretical examination underscores AI's vital role in transforming records management in higher education. The study systematically demonstrated how AI-enabled enhancements in document processing, administrative functions, and decision-making are instrumental for improving efficiency, accuracy and overall efficacy of records management. Also, it has stated some initiatives and the standards of using AI, in addition to the obstacles that may be encountered when applying AI in records management. The observed advantages, spanning accelerated workflows to enriched student experiences, validate AI's revolutionary impact in modernizing higher education record management. The study has reached some results such as:

- Artificial intelligence can be applied to many records management processes to improve efficiency and consistency.

- AI has the capability to enhance how institutions organize, utilize and control their information resources.

- By leveraging machine learning and natural language processing, AI can facilitate automated categorization, digitization, upkeep, storage, finding and evaluation of records.

As institutions tackle surging data volumes, AI emerges as an indispensable instrument not just for managing records at scale but also for fostering a more connected, responsive and insightful administrative apparatus. With concerted efforts to promote ethical, accountable AI and address adoption barriers, higher education can harness the full potential of artificial intelligence to revolutionize records management practices.

However, we must also remain cognizant of the pragmatic challenges involved in integrating AI within higher education record systems, including data privacy risks, algorithmic bias issues, infrastructure costs, workforce disruption, and ethical concerns surrounding transparency and accountability. A measured approach is imperative, embracing AI's benefits while proactively addressing its limitations through comprehensive strategies encompassing governance, skills development, culture change, infrastructure upgrades and cybersecurity investments. With diligent efforts to promote responsible and ethical AI adoption while managing associated risks, higher education institutions can fully utilize artificial intelligence to transform record management practices and unlock new heights of efficiency, insights and student experiences in the future, Although further research is required, AI technologies exhibit immense capacity to transform records management in higher education through improved governance of information flows.

Bibliography:

1. Clay, J. (2018, June 20). Guide to the intelligent campus. *Intelligent Campus*. https://intelligentcampus.jiscinvolve.org/wp/2018/06/20/guide-to-the-intelligent-campus/
2. Ellucian. (2023). *Ethos Open Software | Ellucian*. https://www.ellucian.com/solutions/ellucian-ethos
3. EU. (2019, April 8). *Ethics guidelines for trustworthy AI | Shaping Europe's digital future*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
4. Fayaz Ahmad, S., Alam, M., Rahmat, M., Mubarik, M., & Hyder, S. (2022). Academic and Administrative Role of Artificial Intelligence in Education. *Sustainability*, *14*, 1101. https://doi.org/10.3390/su14031101
5. Green, E., Singh, D., Chia, R., & Actis, A. (2022). *Ai ethics and Higher Education Good Practice and Guidance for Educators, Learners, and Institutions Education Ethics* (p.270). https://www.globethics.net/documents/10131/26882157/GE_EE_10_isbn9782889314423.pdf
6. Horizon. (2022, April 18). *2022 EDUCAUSE Horizon Report | Teaching and Learning Edition*. EDUCAUSE Library. https://library.educause.edu/resources/2022/4/2022-educause-horizon-report-teaching-and-learning-edition
7. ISO. (2016, April 26). *ISO 15489-1:2016*. ISO. https://www.iso.org/standard/62542.html

8. Jisc. (2022, December 5). *An introduction to chatbots*. Jisc. https://beta.jisc.ac.uk/guides/an-introduction-to-chatbots
9. Kelley Connect. (2020, February 27). *A.I. and the Impact It Has on Document Management in Education*. Kelley Connect. https://kelleyconnect.com/artificial-intelligence-and-the-impact-it-is-having-on-document-management-in-education/
10. Khan, N., & Al-Yasiri, A. (2016). Framework for Cloud Computing Adoption: A Roadmap for Smes to Cloud Migration. *International Journal on Cloud Computing: Services and Architecture*, *5*(5/6), 01–15. https://doi.org/10.5121/ijccsa.2015.5601
11. Klutka, J., Ackerly, N., & Magda, A. J. (n.d.). *ARTIFICIAL INTELLIGENCE IN HIGHER EDUCATION Current Uses and Future Applications*. Wiley. https://universityservices.wiley.com/wp-content/uploads/2020/12/201811-AI-in-Higher-Education-TLH-with-new-bage.pdf
12. Mashilo Thomas, M. (2021). *Utilising artificial intelligence technology for the management of records at the Council for Scientific and Industrial Research in South Africa* [Thesis]. https://uir.unisa.ac.za/handle/10500/28800
13. Mashilo Thomas, M. (2022). Artificial intelligence for the improvement of records management activities at the Council for Scientific and Industrial Research. *Journal of the South African Society of Archivists*, *55*, 16–26. https://doi.org/10.4314/jsasa.v55i.2
14. MIT Libraries. (n.d.). *From The Libraries: Digital Initiatives*. Retrieved November 10, 2023, from https://web.mit.edu/fnl/vol/134/seidman.htm
15. OECD. (2019). *OECD Legal Instruments*. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
16. Parycek, P., Schmid, V., & Novak, A. (2023). *Artificial Intelligence (AI) and Automation in Administrative Procedures: Potentials, Limitations, and Framework Conditions / Journal of the Knowledge Economy*. https://doi.org/10.1007/s13132-023-01433-3
17. Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, *12*(1), 22. https://doi.org/10.1186/s41039-017-0062-8
18. Reitz, J. M. (2014). *ODLIS*. Online Dictionary for Library and Information Science. https://odlis.abc-clio.com/odlis_a.html
19. Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoupova, T., & Stuart, K. (2019). More human than human? Artificial intelligence in the archive. *Archives and Manuscripts*, *47*(2), 179–203. https://doi.org/10.1080/01576895.2018.1502088
20. Tabassi, E. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (error: 100-1; p. error: 100-1). National Institute of Standards and Technology (U.S.). https://doi.org/10.6028/NIST.AI.100-1
21. UC Presidential Working Group on AI. (2021). *Recommendations to Guide the University of California's Artificial Intelligence Strategy*. University of California. https://www.ucop.edu/ethics-compliance-audit-services/compliance/presidential-working-group-on-artificial-intelligence.html
22. Webb, M., Attewell, S., Moule, T., Nicholson, H., & Fraz, A. (2021, April 26). *Artificial intelligence (AI) in tertiary education—Jisc*. Jisc. https://beta.jisc.ac.uk/reports/artificial-intelligence-in-tertiary-education
23. Alasmari, J. S. . (2023). The Dynamics of Verbal and Non-Verbal Linguistic Communication in The Saudi Sports Community. *Arts for Linguistic & Literary Studies*, *5*(4), 539–569. https://doi.org/10.53286/arts.v5i4.1676
24. Omer, N. I. M. (2024). Maintaining Meaningful Human Interaction in AI-Enhanced Language Learning Environments: A Systematic Review . *Arts for Linguistic & Literary Studies*, *6*(3), 533–552. https://doi.org/10.53286/arts.v6i3.2083

25. Al-Ghobesi, A. A. H. (2025). Risks of Relying on Artificial Intelligence in Learning Arabic Language Sciences Through the Meta Application. *Arts for Linguistic & Literary Studies, 7*(1), 396–419. https://doi.org/10.53286/arts.v7i1.2420

26. Ahmed, M. R. A. (2025). Accreditation and Quality Assurance: Exploring Impact and Assessing Institutional Change in the US and Saudi Arabian Higher Education Institutions. *Arts for Linguistic & Literary Studies, 7*(1), 626–639. https://doi.org/10.53286/arts.v7i1.2419

مجلـــة العلوم الهندسية والتقنية
مجلة علمية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

A Web Mining Approach for Evaluation of Quality Assurance at University of Science and Technology

Ahmed Sultan Al-Hegami[1], Nathem Saeed Ahmed Al-Hddad[2], Hussein Alkhader Alsaeedi[3]

[1]Faculty of Computers and Information Technology, University of Sana'a, Sana'a, Yemen

[2]Department of Computer Science, University of Science and Technology, Aden, Yemen

[3]Department of Computer Science, University of Science and Technology, Sana'a, Yemen

alhegami@su.edu.ye n.alhaddad@aden.ust.edu.ye alhussein1977@gmail.com

ABSTRACT

In today's digital landscape, there is an abundance of websites containing vast amounts of information. However, the challenge lies in extracting valuable knowledge from this sea of data. Knowledge serves as the foundation for making informed decisions, making it essential to develop efficient mining approaches. This research aims to address this problem by proposing a novel framework. The framework will focus on analyzing diverse datasets, particularly those related to educational, administrative, and student activities at the University of Science and Technology. Through an in-depth analysis of this data, the quality of services provided by the university will be evaluated. The ultimate goal is to make optimal decisions that will enhance the university's overall performance. This includes selecting the best teaching staff and improving various other services to ensure a higher level of efficiency and effectiveness.

Keywords: Data mining, Rules Discovery, Quality Assurance, Web mining, Web Content, Web Usage, KDD.

1. INTRODUCTION

Computer science evolved during the last years of the last century, dramatically and rapidly, besides the emergence of many technologies which served humans such as: database, data warehouse, and data integrated, also environment communication, networks, Internet, etc., that contribute to the rise of information.

There are many websites that involve a lot of data but the problem lies in the fact that taking a decision in getting the information and finding out knowledge is quite difficult. For this reason, the mining approach is required to find a proper solution to this problem.

In this study, the content of the various websites of University of Science and Technology will be analyzed in an attempt to find a solution for the problem of the research which is the evaluation of services quality offered by the university.

Accordingly, the available data will be analyzed to evaluate the educational, administrative and students' activities and services because they are the primary services for any educational institution. Algorithms data mining are used for that to create a general role for prediction process with any new case.

After that, we are going to depend on the role that will occur during the contact with a new process so that we realize from the beginning what is good or not. By this, we will predict the quality of primary services in the university. A new framework will be to extract the patterns from the historical data. The resulted patterns will be the basis to create a general role to data with any new case and predict its result.

2. PROBLEM STATEMENT

The study deals with the following issues:

Discovering knowledge from the large volume of data is difficult and not effectively utilized.

The discovery of uninteresting patterns included repeated, conformed, generalized, and specialized knowledge from data mining process.

3. RELATED WORK

In [3], the authors introduce novel possibilities for leveraging robust graph representations in conjunction with popular machine learning algorithms. They demonstrate how graph distance can be used to assess graph similarity and adapt well-known algorithms like k-means clustering and k-nearest neighbors' classification to operate on graphs. The effectiveness of these techniques is evaluated in the field of web content mining, where multiple approaches for representing web document content using graphs are introduced.

In (Ahmed, 2025) [4], the authors compare different web page representations and highlight the impact of data collection and classification processes on the results. They find that a binary representation yields the best performance in binary classification scenarios, but note that other representations may be more effective as the number of classes increases.

In [5], the authors address the challenges of accessing consumer opinions prior to the internet era. They focus on customer reviews as a valuable source of information and propose a visual analysis system for comparing consumer opinions across multiple products. They introduce a supervised pattern discovery method to automate the identification of product features from reviews.

In [7], the authors present a system that allows analysts to interactively correct errors of an automatic system, making the process more efficient than manual tagging. Future work includes improving automatic techniques, studying the strength of opinions, and extracting useful information from other opinion sources.

In [10], the authors propose a solution for concise and conceptual audience metrics in web analytics. They discuss techniques for mining web pages generated by web servers and introduce term-based metrics. They leverage term grouping algorithms and concept hierarchies from ontologies to aggregate the metrics into concept-based metrics using OLAP tools.

In [6], the authors develop indicators of the quality of service provided by Egyptian universities through electronic means. They aim to assure students that the university provides high-quality decisions through the internet as part of the educational experience.

In [8], the researchers divide evaluation criteria for educational websites into content and technical aspects. They propose nine criteria to help teachers evaluate websites for personal resource use and classroom implementation, distinguishing between informative and well-designed websites and those that are uninformative and poorly designed.

In [10], the authors discuss the usage of the internet as a source of information and emphasize the importance of evaluating information skills that will be utilized throughout one's lifetime.

4. The NWMA Approach

4.1. The NWMA Approach Framework

The main core of this research is building a good framework to discover the patterns which will lead to evaluate the university services.

This evolution process will support the university to improve its performance. Figure (4.1) shows the framework of the proposed approach. The proposed approach consists of five main steps as follows:

1. Settings the parameters and creating of nodes.

2. Classification of parameters.
3. Computing of Attributes.
4. Generate partial rule.
5. Produce of final results.

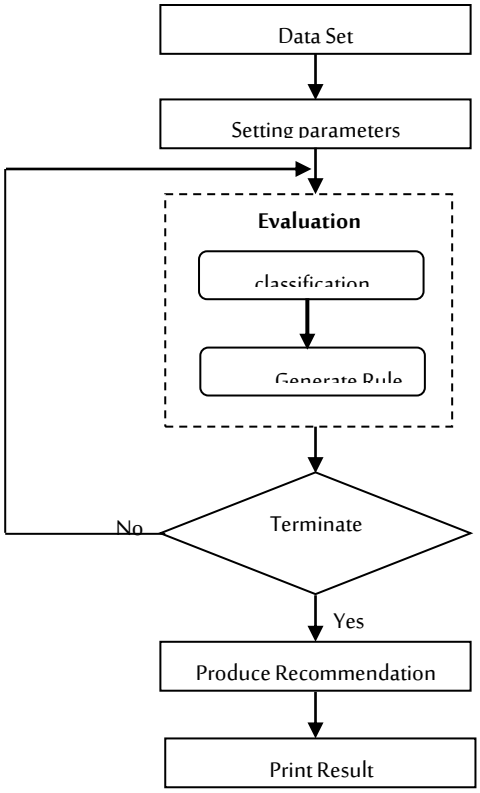


Figure (2): Framework.

4.1.1 Setting parameters

There are many factors on which I will depend in evaluation of the educational services in university of science and technology.

Such factors vary with each other in the degrees of importance.

There are primary and secondary standards. So, the degree of significance has been recorded for every standard.

The objective of this is:

- Achieving flexibility:

To be able to decrease and increase the standards of evaluation and this facilitates the decrease and increase the percentage significance so that it does not affect the total percentage.

- Obtaining accuracy:

All the standards do not have equal values in measuring performance.

However, there is a great significance for some standards of the decision makers.

The following table shows every group with its significance so the total degree is the full percentage %100.

The parameters are classified into two classifications:

4.1.2 Classification of Parameters:

We will classify parameters into two classes:

- **Main Attributes:**

In table (1) shows that:

Student Activities	Administrative services	Educational services
%25	%25	%50

Table (1): Main Attributes.

- **Secondary Attributes**

In table (2) shows that:

Educational services							Administrative services				Student Activities			
The Experience	The Age	Dealing with Students	Using Technology	Library Hours	Using LMS	Official Schedule	Performance of Procedures	Application of Regulations	Dealing with Customer	Quality Systems	Utilizing The Activity	Carrying Out The Activities	Student's interaction with The	Achieving Objective
%25							%25				%50			

Table (2): Secondary Attributes.

5 Computing of Attributes:

The following notations are used in the algorithms:

MC=Main Class

SA=Secondary Attributes

CP=Class Degree

SP=Secondary Attributes Degree

WC=Class Weight

WS=Secondary Attributes Weight

1.The following law is used to compute the main class degree:

$$CP = \sum_{i=1}^{n} (\text{SA}) * \boldsymbol{100/Wc}$$

where,

CP=Class Degree

WC=Class Weight

SA=Secondary Attributes

n=the number of Secondary Attributes

2.The following law is used to compute the secondary attributes value:

$$SA = \sum_{i=1}^{n} (\text{SP}) * \boldsymbol{Wc/Ws}$$

where,

SA=Secondary Attributes

SP=Secondary Attributes Degree

WC=Class Weight

WS=Secondary Attributes Weight

n=the number of secondary attributes degree

5.1. Implementation and Experimentation

The Decision Rules Generated from the NWMA approach:

5.1.1 First: Educational Services

1.If teacher age ="$\leq$30" then bad performance

2.If teacher age = "30 .. 40" then good performance

3.If teacher age ="> 40" and teacher expertise ="small then bad performance

4.If teacher age ="> 40" and teacher expertise ="medium" then good performance

5.If teacher age ="> 40" and teacher expertise ="large" then good performance

We used the above rules to building the model that shown in the next figure:

6 Second: The Administrative Services

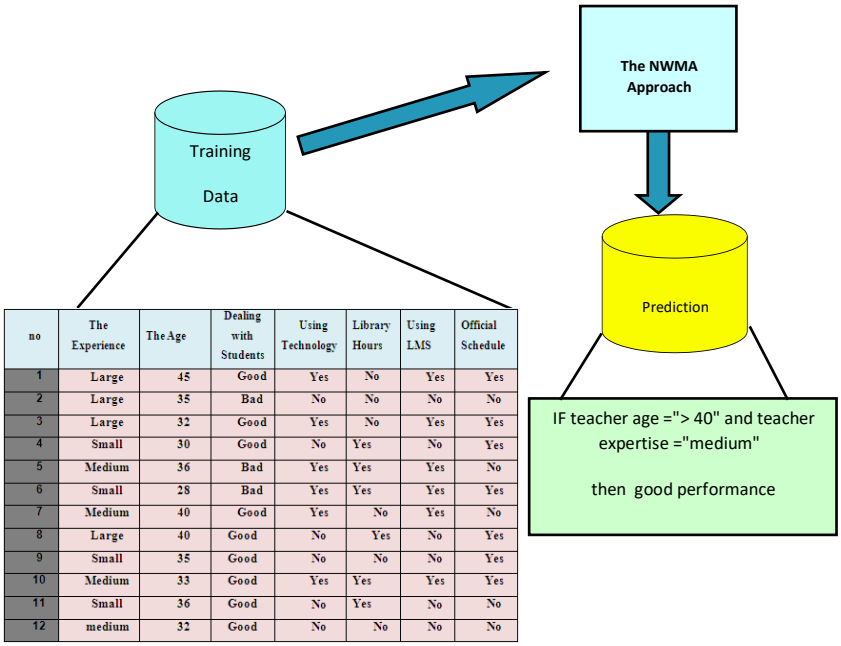
1.If Dealing with Students ="low" then bad performance

2.If Dealing with Students ="medium" then good performance

3.If Dealing with Students="high" and applied of regulations="low" then bad performance

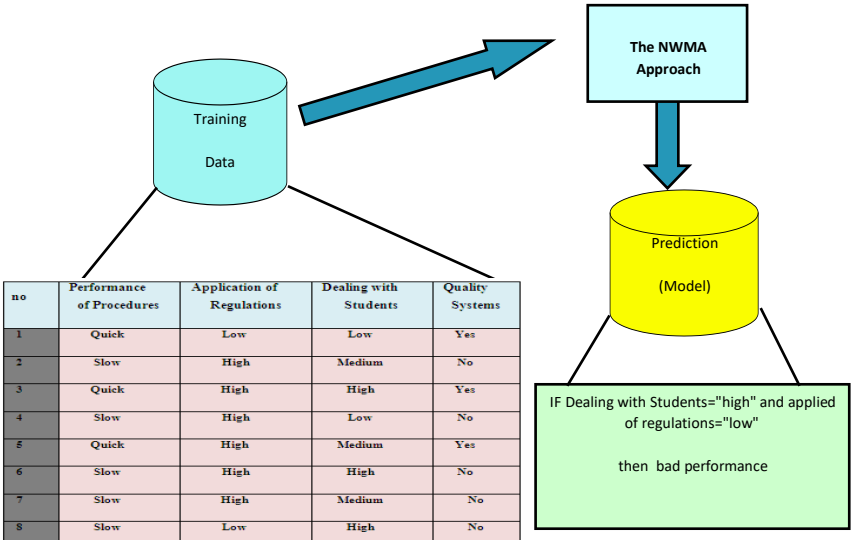
4.If Dealing with Students="high" and applied of regulations="high" then good performance

We used the above rules to building the model that shown in the next figure:



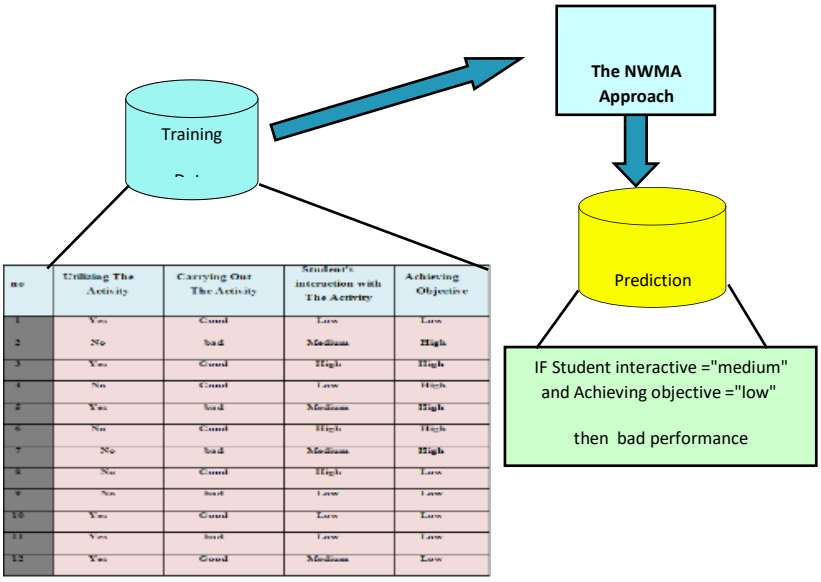
no	The Experience	The Age	Dealing with Students	Using Technology	Library Hours	Using LMS	Official Schedule
1	Large	45	Good	Yes	No	Yes	Yes
2	Large	35	Bad	No	No	No	No
3	Large	32	Good	Yes	No	Yes	Yes
4	Small	30	Good	No	Yes	No	Yes
5	Medium	36	Bad	Yes	Yes	Yes	No
6	Small	28	Bad	Yes	Yes	Yes	Yes
7	Medium	40	Good	Yes	No	Yes	No
8	Large	40	Good	No	Yes	No	Yes
9	Small	35	Good	No	No	No	Yes
10	Medium	33	Good	Yes	Yes	Yes	Yes
11	Small	36	Good	No	Yes	No	No
12	medium	32	Good	No	No	No	No

Figure (3): Module building for The Administrative Services



no	Performance of Procedures	Application of Regulations	Dealing with Students	Quality Systems
1	Quick	Low	Low	Yes
2	Slow	High	Medium	No
3	Quick	High	High	Yes
4	Slow	High	Low	No
5	Quick	High	Medium	Yes
6	Slow	High	High	No
7	Slow	High	Medium	No
8	Slow	Low	High	No
9	Quick	Low	Low	No
10	Slow	Low	Low	Yes
11	Quick	Low	Low	Yes
12	Quick	Low	Medium	Yes

Figure (4): module building for The Educational Services.



no	Utilizing The Activity	Carrying Out The Activity	Student's interaction with The Activity	Achieving Objective
1	Yes	Good	Low	Low
2	No	bad	Medium	High
3	Yes	Good	High	High
4	No	Good	Low	High
5	Yes	bad	Medium	High
6	No	Good	High	High
7	No	bad	Medium	High
8	No	Good	High	Low
9	No	bad	Low	Low
10	Yes	Good	Low	Low
11	Yes	bad	Low	Low
12	Yes	Good	Medium	Low

Figure (5): Module building for Students Activities Services

7 Third: Students the Activities Services

1.If Student interactive="low" then bad performance

2.If Student interactive ="medium" and Achieving objective ="low" then bad performance

3.If Student interactive ="medium" and Achieving objective ="high" then good performance

4.If Student interactive ="high" and Achieving objective ="low" then good performance

5.If Student interactive ="high" and Achieving objective ="high" then good performance

We used the above rules to building the model that shown in the next figure:

6.2. Prediction

- **Prediction for The Educational Services**

Now, we will use the new record as attesting data form prediction. So we want to predict the performance of the following the teacher by using the next:

The Experience = " Medium "

The Age = "43"

Dealing with Students ="Good"

Using Technology = "No"

Library Hours = "Yes"

Using LMS =" No"

Official Schedule= "No"

Performance ="?"

We cannot use the rule no. 1

And We cannot use the rule no. 2

And We cannot use the rule no. 3

But we can use the rule no.4

IF teacher age ="> 40" and teacher expertise ="medium"

then good performance

If 43>40 (TRUE) and "medium" ="medium"(TRUE) then good performance

The use of the model in prediction is shown in the in the following figure:

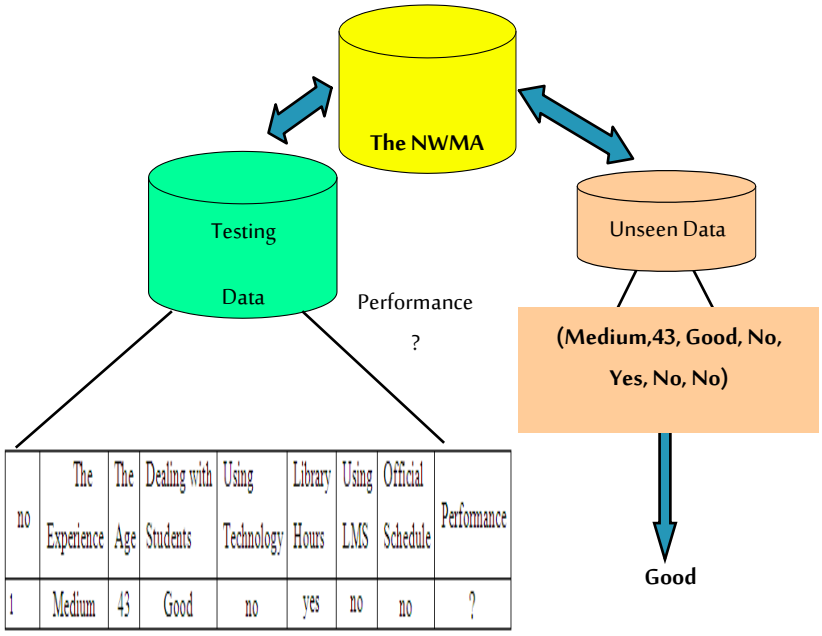


Figure (6): Using Model in Prediction for the Educational Services

- **Prediction for the Administrative Services**

Now we will use the new record as attesting data form prediction. So we want to predict the performance of new management by using the next:

Performance of Procedures ="Quick"

Application of Regulations ="low"

Dealing with Students ="high"

Quality Systems ="yes"

Performance ="?"

We cannot use the rule no. 1

And We cannot use the rule no. 2

But we can use the rule no.3

IF Dealing with Students="high" and applied of regulations="low"

then bad performance

If high=high (TRUE) and "low" ="low"(TRUE) then bad performance

The use of the model in prediction is shown in the in the following figure:

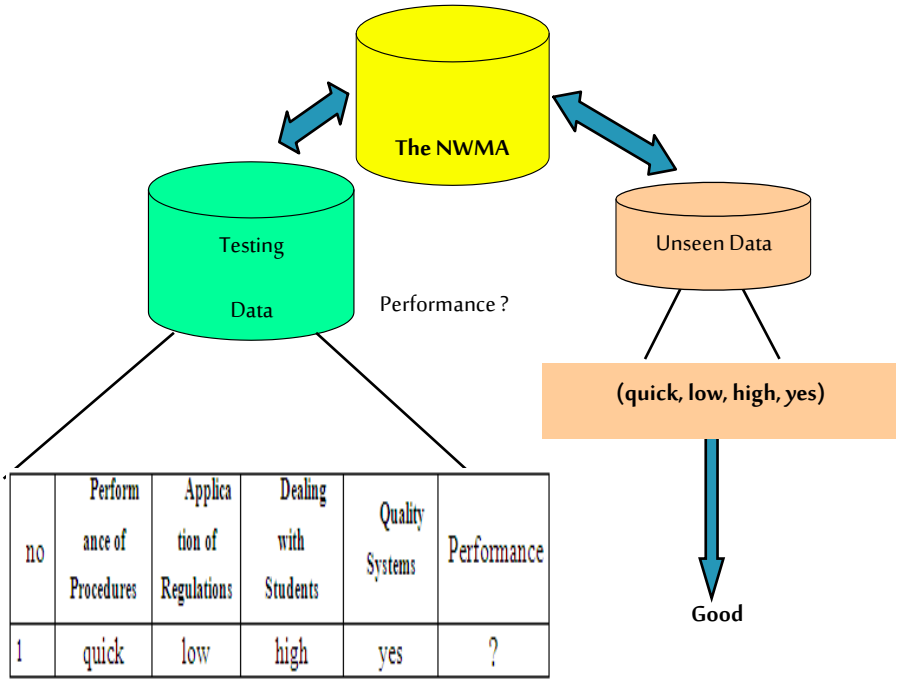


Figure (7): The Use of Model in Prediction for The Administrative Services

- **Prediction for Students the Activities Services**

Now we will use the new record as attesting data form prediction. So we want to predict the performance of new activity by using the next:

Utilizing The Activity = "No"

Carrying Out The Activity ="bad"

Student's interaction with The Activity= " Medium"

Achieving Objective = "low"

Performance ="?"

We cannot use the rule no. 1

But we can use the rule no.2

IF Student interactive ="medium" and Achieving objective ="low" then bad performance

If medium=medium (TRUE) and low="low"(TRUE) then bad performance

The use of the model in prediction is shown in the in the following figure:

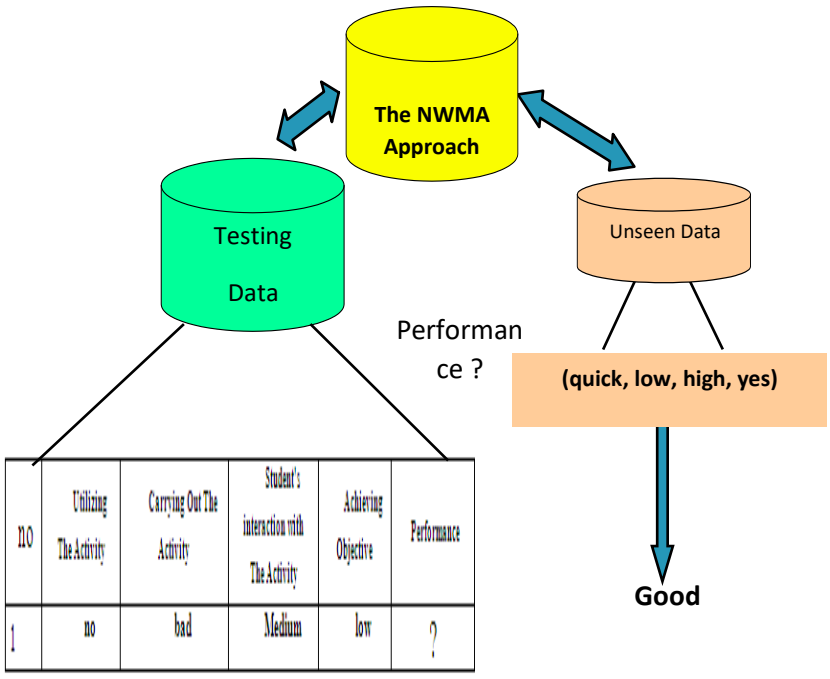


Figure (7): The Use of Model in Prediction for Students' Activities Services

7.1. Conclusions:

In this research, we proposed a web mining approach for evaluation of quality assurance at University of Science and Technology.

We implemented a framework using C++ programming Language. The framework is experimented and evaluated using real dataset and results have been presented.

The Performance of services in UST depended on educational services.

7.2. Future work:

Future work should consist of more experiments with other dataset, as well as more elaborated experiments to optimize several parameters of the algorithm. In addition, the framework can be implemented and experimented with large dataset by programming it in any programming language.

References:

[1] Julio Ponce And Adem Karahoca, Data Mining and Knowledge Discovery in Real Life Applications.

[2] Olfa Nasraoui, Myra Spiliopoulou, Jaideep Srivastava, Bamshad Mobasher, Brij Masand, Advances in Web Mining and Web Usage Analysis.

[3] Amento, L. Terveen, and W. Hill. "Does Authority Mean Quality? Predicting Expert Quality Ratings of Web Documents. In Proc. of the 23rd ACM SIGIR Conf. on Research and Development in Information Retrieval", 2000, pp. 296–303.

[4] Ayres, J. Gehrke, T. Yiu, and J. Flannick. "Sequential Pattern Mining Using Bitmaps In Proc. of the Eighth ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD'02)", 2002, pp. 429–435.

[5] Baeza-Yates and B. Ribeiro-Neto. "Modern Information Retrieval. Addison-Wesley", 1999.

[6] Barbar, Y. Li and J. Couto. COOLCAT. "an Entropy-Based Algorithm for Categorical Clustering. In Proc. of the 11th Intl. Conf. on Information and knowledge management (CIKM'02)", 2002, pp. 582–589.

[7] Berendt, B. Mobasher, M. Nakagawa, and M. Spiliopoulou. "The Impact of Site Structure and User Environment on Session Reconstruction in Web Usage Analysis. In Proc. of the KDD'02 WebKDD Workshop", 2002.

[8] Blum, and S. Chawla. "Learning from Labeled and Unlabeled Data Using Graph Mincuts. In Proc. of Intl. Conf. on Machine Learning (ICML'01)", 2001, pp.19–26.

[9] Brunk and M. J. Pazzani. "An Investigation of Noise-Tolerant Relational Concept Learning Algorithms. In Proc. of the 8th Intl. "Workshop on Machine Learning", 1991, pp. 389–393.

[10] Buchner and M. D. Mulvenna. "Discovering Internet Marketing Intelligence through Online Analytical Web Usage Mining. In Proc. of the ACM SIGMOD Intl. Conf. on Management of Data (SIGMOD'99)", 1999, pp. 54–61.

[11] Buehrer, S. Parthasarathy, and A. Ghoting. "Out-of-core frequent pattern mining on a commodity PC. In Proc. of the ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD'06)", 2006, pp. 86 – 95.

[12] Burges. "A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, 2(2)", 1998, pp. 955–974.

[13] Cooley, B. Mobasher, and J. Srivastava. "Data Preparation for Mining World Wide Web Browsing Patterns. Knowledge and Information Systems, 1(1)", 1999, pp. 5–32.

[14] Dhillon. "Co-Clustering Documents and Words Using Bipartite Spectral Graph Partitioning. In Proc. of the 7th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD'01)", 2001, pp. 269–274.

[15] Dhillon, S. Mallela, and D. S. Modha. Information-Theoretic Co-Clustering. In Proc. of The Ninth ACM SIGKDD Intl. Conf. "on Knowledge Discovery and Data Mining (KDD'03)", 2003, pp. 89–98.

[16] Domingos, and M. J. Pazzani. "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. Machine Learning", 1997,29(2–3), pp. 103–130.

[17] http://www.flazx.com

[18] http://www.ust.edu.ye

[19] http://www.aden.ust.edu.ye

[20] http://www.taiz.ust.edu.ye

[21] http://www.hoda.ust.edu.ye

[22] http://kdd.ics.edu

[23] Ahmed, M. R. A. (2025). Accreditation and Quality Assurance: Exploring Impact and Assessing Institutional Change in the US and Saudi Arabian Higher Education Institutions. *Arts for Linguistic & Literary Studies*, *7*(1), 626–639. https://doi.org/10.53286/arts.v7i1.2419

مجلـــة العلـوم
الهندسية والتقنية
مجلة علمية نصف سنوية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

Journal of Engineering and Technological Sciences – JOEATS

A Scientific refereed journal issued by the Faculty of Engineering and the Faculty of Computers and Informatics, Thamar University

Journal of Engineering and Technological Sciences - JOEATS published by the Faculty of Engineering and the Faculty of Computers and Informatics, Thamar University JOEATS publishes Engineering and Computers and Information Systems science articles the area of Architecture, Civil, Electrical, Mechanical, Communications, Urban planning and design, Mechatronics, Computer and information technology, Environment and renewable energy and other engineering fields.

ADVISORY BOARD

EDITORIAL SECRETARY

Address

Faculty of Engineering
Thamar University
Thamar city, Republic of Yemen

Contacts

joeats-tu@tu.edu.ye

Tel. 777296603

Scientific Journal

Published biannual by the Faculty of Engineering and the Faculty of Computers and Informatics, Thamar University

Steps to publish in the journal

The publishing process in the journal proceeds according to the following steps:

1. The paper should be characterized by originality and sound scientific methodology.
2. The paper should not have been previously published or submitted for any publication to another party, and the researcher has to submit a written undertaking for that.
3. Papers should be written in a sound language, taking into account the rules of punctuation and accuracy of forms - if any - in (Word) format.
4. A copy of the research to be published shall be delivered to the office of Journal of Engineering and Technological Sciences – JOEATS, Thamar University, or it shall be sent to the editor-in-chief of the journal at the e-mail: joeats-tu@tu.edu.ye
5. After receiving the research, it will be confirmed by sending an email to the first author.
6. The initial review is made of the quality of the research and whether it matches the requirements of publication and quality in the journal.
7. When the research is initially accepted for publication in the journal, a request to pay the publishing fees is sent by e-mail.
8. After receiving the non-refundable publishing fees (cash or bank transfer), the research evaluation and arbitration phase begins.
   The research is sent for evaluation and arbitration by (two arbitrators).
9. According to the decision of the arbitrators, the editor-in-chief takes the decision and it is sent to the writer - the writers**. The decision taken has three cases:**

- Acceptance of publication without modification - or acceptance of publication with minor - basic modifications (required to be amended) or refusing to publish the research:
- After the author or authors make the required amendments and send them to the arbitrators and approval by the arbitrators and the acceptance of the publication of the research, the decision to accept the publication is sent to the writer along with the copyright form attached. The writer signs the copyright form and sends it back by e-mail to the editor-in-chief within 6 days.

- **Publication fee**

  - The writer and writers are required to pay the non-refundable amounts due (see the table below) as a fee for publishing the research, and this is done only after the editor-in-chief informs the writer or writers that the research is appropriate and adheres to the requirements of the journal. -Bank transfer.

Fees	Characteristics of the first researcher
20000 Y.R	Faculty members / Thamar University
30000Y.R	Yemeni faculty members / universities, bodies and organizations
150 USD	oreign researchers from countries and external organizations

- **Research submission guide:**

  - Journal of Engineering and Technological Sciences - JOEATS published by the Faculty of Engineering and the Faculty of Computers and Informatics, Thamar University JOEATS publishes Engineering and Computers and Information Systems science articles the area of Architecture, Civil, Electrical, Mechanical, Communications, Urban planning and design,

Mechatronics, Computer and information technology, Environment and renewable energy and other engineering fields, In one of the two languages (Arabic or English), and the research is prepared and received electronically in MS WORD format. The guide for submitting research shows the following requirements for publication:

- **Research Components:-**

- The research must contain the following: The first paper (a title of the research - the name of the author - the writers and their titles, a summary of the research (less than 250 words) and key words). It starts from the second paper (contents of the research, which includes the introduction, the research problem, research objectives, research background, methodology, analysis, results, and recommendations, if any, and references including tables, figures, charts, and illustrations images). The entire article should not exceed (8000) words.

- **Stages of evaluation and arbitration:**

- Each research is evaluated initially by the editorial board. Suitable research is sent for publication to (two independent arbitrators) at least, and based on the recommendation of the arbitrators and in consultation with the editorial board, the editor-in-chief decides one of three decisions: Accepting publication without modification or accepting publication with secondary-fundamental modifications (requests to amend it) or refuses to publish the research. joeats-tu@tu.edu.ye

- **Research delivery:**

- Researchers wishing to publish their research in the Journal of Engineering and Technological Sciences - JOEATS can submit their research by e-mail: joeats-tu@tu.edu.ye
- After receiving the research by e-mail, the researcher will be notified of receiving his research by e-mail, and after the researcher has paid the publishing fees, the stage of arbitration of the research will begin by independent arbitrators of different nationalities, for more information about the method, form and size of the research, researchers will be provided with a guide to submitting research via e-mail, or it can be obtained on the following website of the journal

- **Copyrights:-**

1. Copyright of transmitted materials If the research includes excerpts from works protected by authors' rights, the author (researcher) shall obtain permission from the copyright owners and provide proof when submitting the copyright form of his research.
2. Transfer of copyright: A copyright transfer form (researcher) for all copyrights (sent to the researcher) shall be attached to the Journal of Engineering and Technological Sciences - JOEATS after being signed by the researcher. All requests to reproduce or republish the research in whole or in part, and the researcher sends them to the editor-in-chief of the journal via e-mail: joeats-tu@tu.edu.ye

مجلــة العلوم الهندسية والتقنية
مجلة علمية نصف سنوية محكمة - تصدر عن كلية الهندسة وكلية الحاسبات والمعلوماتية – جامعة ذمار

Contents

JOURNAL OF ENGINEERING AND TECHNOLOGICAL SCIENCES - JOEATS

Scientific Journal Biannual

issued by the Faculty of Engineering and the Faculty of Computers and Informatics, Thamar University

Volume: 3، Issue: 1