Educational Administration: Theory and Practice
2024,30(9) 973-978
ISSN:2148-2403
https://kuey.net/

Research Article

Impact of Artificial Intelligence on the Quality of Arabic Language

Dr.Umm Kulthum Hakum Dawood bin Yahya[1*], Dr. Muhammad Madiani[2]

[1]Department of Jurisprudence and its Principles, Faculty of Sharia, King Khalid University, Email: obenyahia@kku.edu.sa
[2]College of Information Technology, United Arab Emirates University

Citation: Dr.Umm Kulthum Hakum Dawood bin Yahya, et.al (2024), Impact of Artificial Intelligence on the Quality of Arabic Language, Educational Administration: Theory and Practice, 30(9), 973-978
Doi: 10.53555/kuey.v30i9.9897

ARTICLE INFO

ABSTRACT

This research aims to study the impact of the generative artificial intelligence model, ChatGPT, on the quality of the Arabic language by analyzing the texts it generates in terms of accuracy, coherence, and linguistic diversity, comparing them with naturally written human texts.

The study also addresses the major challenges the model faces in processing the Arabic language, such as grammatical errors, structural inaccuracies, and reliance on limited data sources.

Due to the nature of the research, a descriptive, analytical, and experimental methodology has been adopted. Experiments will be conducted on a sample of texts generated by ChatGPT and compared with natural texts using specialized linguistic analysis tools.

Additionally, the results will be evaluated by a group of Arabic language experts to assess text quality, adherence to grammatical and morphological rules, as well as vocabulary diversity and semantic coherence.

Keywords: Artificial intelligence, Large language models, Arabic language quality, generative AI, Natural language processing.

Introduction

In recent years, the world has witnessed significant advancements in the field of artificial intelligence, with generative AI emerging as a groundbreaking development in the ability of computational systems to understand and generate natural language. This type of AI relies on large language models, which are deep neural networks trained on vast amounts of text sourced from books, articles, and digital content available on the internet. Among the most prominent of these models are ChatGPT and DALL-E from OpenAI, as well as Gemini from Google.

Large language models are distinguished by their ability to generate creative texts, summarize information, and answer questions in a manner that simulates human responses. While these models have revolutionized language processing in multiple languages, they still face significant challenges when dealing with Arabic, a language known for its richness and complexity. This is largely due to the limited quality and quantity of Arabic-language data used in training compared to languages such as English.

Despite efforts to develop models that support the Arabic language, such as the "Aleem" and "Jais" models, the texts generated in Arabic often exhibit lower accuracy and linguistic richness compared to naturally written texts.

Research Significance

The significance of this research lies in its attempt to shed light on the impact that large language models can have on the Arabic language. It also aims to identify the challenges associated with this process and propose solutions to improve the quality of AI-generated texts, thereby narrowing the gap between human-written and machine-generated Arabic texts.

Research Questions

- To what extent do large language models affect the quality of the Arabic language?
- What are the main challenges in training large language models for Arabic?
- How can the quality of data used to train Arabic language models be improved?

Copyright © 2024 by Author/s and Licensed by Kuey. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

Research Objectives
- Analyze the impact of large language models on the quality of Arabic texts.
- Discuss the challenges related to data collection and model training for Arabic.
- Provide recommendations to enhance language models for the development of Arabic linguistic processing.

Previous Studies:
Numerous studies have explored generative AI in foreign languages, with most focusing on languages such as English and French due to the availability of large datasets and the ease of processing these languages using AI technologies.
In contrast, research on generative AI in Arabic remains limited. This is largely due to the unique characteristics of the Arabic language, including its complex grammatical structures, rich morphological system, and the variations between Modern Standard Arabic (MSA) and regional dialects. These linguistic differences make it difficult to generalize the findings of studies conducted on other languages to Arabic. Consequently, there is a need for specialized research that considers the structural particularities of Arabic and the challenges of processing it through AI models.

Among the studies that intersect with this research in both scope and specificity are:
- Applications of Artificial Intelligence and Their Impact on Developing Creative Linguistic Skills in Outstanding High School Students (Abdel Qader, et al., 2023):
This recent study examines the impact of AI technologies on enhancing the creative linguistic skills of outstanding students. The study was conducted on a sample of 30 students from the third-year secondary level at Al-Azhar institutes. A pretest was administered before applying the AI-based learning intervention, followed by a posttest to measure linguistic creativity. The results indicated statistically significant differences in favor of the post-test, demonstrating the effectiveness of AI applications in developing these skills. The study recommended integrating AI applications into the educational process and focusing on developing students' creative linguistic abilities across different educational levels.
- The Impact of Artificial Intelligence on Teaching the Arabic Language (Fitriana, 2023):
This research explored the influence of artificial intelligence on Arabic language education, highlighting both the benefits and challenges associated with integrating AI into the learning process.
The study employed a descriptive-analytical approach, providing a comprehensive understanding of the phenomenon. It also included examples of AI applications used in Arabic language instruction.
While this study serves as a strong foundation for understanding the relationship between AI and Arabic language teaching, it requires further experimental analysis and expanded research to incorporate the latest advancements in the field.
- Artificial Intelligence Applications and Their Role in Developing Smart Content from the Perspective of Faculty Members (Al-Shammari, 2023):
It aimed to explore the role of AI applications in developing intelligent educational content by analyzing faculty members' perspectives on the effectiveness of these technologies in enhancing education quality and meeting learners' needs.
The study adopted a descriptive-analytical approach, focusing on how AI is used to create educational content that aligns with modern teaching requirements. It also discussed the benefits of AI, such as curriculum enhancement, adaptive learning, and personalized educational tools, as well as the challenges associated with its implementation, including high costs, algorithm complexity, and reduced student-teacher interaction.
- The study, conducted by researchers Zaki and Ali (2024):
The study explored the integration of AI technologies in teaching the Arabic language, with a particular focus on developing personalized and adaptive learning materials tailored to students' needs. It assessed five major AI tools:
ChatGPT, Gemini, Co-Pilot, JAIS, and Diffit, evaluating them based on their ability to generate linguistically accurate and pedagogically appropriate texts for intermediate Arabic learners.
The study ultimately reached several key findings, which are:

First: Challenges Facing AI in Arabic
Overuse of nominal sentences.
Frequent errors in verbs and pronouns.
Irregular repetition of vocabulary without considering the context.
Influence of English on generated texts, leading to unnatural expressions and grammatical mistakes.

Second: Advantages of the Studied Tools
Providing diverse content and different types of questions.
Ability to generate engaging content for learners.
- The Use of Artificial Intelligence to Develop Arabic Language Skills in Learning (Mulyanto, et al., 2024)
The research examined the role of artificial intelligence in improving Arabic language education, focusing on vocabulary development, grammar, and conversational proficiency. It employed an analytical approach to study the impact of AI tools such as ChatGPT, JAIS, and Gemini in supporting learners.

The findings indicated that these technologies enhanced the learning experience, yet faced challenges such as linguistic repetition, grammatical errors, and English influence on Arabic phrasing. Despite its significance, the study lacked quantitative analysis and experimental comparisons with traditional teaching methods. Ultimately, the research recommended further experimental studies to ensure the effectiveness of AI in language skill development.

- The Reality of the Arabic Language in AI Applications: A Descriptive Applied Study (Al-Amairah, 2025):

The study explores the relationship between artificial intelligence and the Arabic language, focusing on how core linguistic knowledge is represented to support language comprehension.

Using a descriptive-applied methodology, the research highlights the challenges facing the Arabic language, including structural changes, the rise of mixed Arabic-English language usage, and the impact of machine translation on classical Arabic.

Research Methodology

The study adopts the descriptive-analytical approach, as it is suitable for assessing the development of generative AI in the Arabic language. It focuses on the linguistic characteristics of Arabic, including grammar and sentence structures, and their impact on the performance of language models.

Additionally, the applied approach was used to examine the challenges faced by language models in understanding and generating Arabic texts. This was conducted through experiments on ChatGPT, evaluating the accuracy, diversity, and linguistic coherence of its outputs.

Section One: Theoretical Framework

First: Generative Artificial Intelligence

Generative AI is a type of artificial intelligence that utilizes machine learning techniques and deep neural networks to simulate human capabilities in creating new and original content, such as texts, images, and videos. AI models can generate outputs of the same type as the input (e.g., text-to-text conversion) or different types (e.g., text-to-image or text-to-video transformation) (Saudi Authority for Data and Artificial Intelligence (SDAIA, 2023, p. 8).

Some of the key factors behind the success of generative AI include:

- The availability of vast amounts of data on the internet.
- Advancements in computational power.

Second: Large Language Models (LLMs)

Large Language Models (LLMs) are advanced AI systems that rely on deep neural networks and deep learning techniques to process Natural Language Processing (NLP) tasks.

These models are trained on massive amounts of textual data to generate human-like text, understand context, answer questions, and produce new content.

Examples of such models include:

GPT-4 by OpenAI, BERT by Google, Jais, an Arabic language model from Mohamed bin Zayed University of Artificial Intelligence, LLaMA by Meta.

Key Features of Large Language Models (LLMs):

- LLMs rely on the transformer architecture, as seen in models like GPT (Generative Pre-trained Transformer).
- These models excel at translation, text summarization, content creation, and sentiment analysis.
- They are trained using large-scale data from various sources, such as books, articles, and websites.

Reports indicate that LLMs were initially developed with a primary focus on the English language, making their performance relatively less efficient in other languages due to limited training data. AI-driven applications like chatbots, content moderation, and sentiment analysis perform significantly better in English than in the 7,000+ other global languages (Lost in Translation: Large Language Models in Non-English Content Analysis (May 22, 2023).

Supporting this claim, Statista reports that approximately 52% of all online textual data is in English, while other languages share the remaining percentage. The Arabic language accounts for only 0.6% of the total available online data (https://www.statista.com/statistics/262946/most-common-languages-on-the-internet/)

Third: Text Quality

Text quality refers to how effectively a text achieves its intended purpose with clarity and impact. The key quality criteria include:

Vocabulary Diversity: The use of a wide range of words and expressions.

Accuracy: Adherence to correct grammar and spelling rules.

Coherence: Logical flow and smooth transition of ideas.

Additionally, quality depends on the credibility and impact of the information, especially in persuasive or interactive texts. A well-written text is easy to understand, engaging, and conveys ideas clearly without unnecessary complexity, making it more effective in fulfilling its purpose.

Fourth: Natural Language
According to the Data and AI Glossary published by the Saudi Data and AI Authority (SDAIA), the term "natural language" is defined as:
"A human language that can be expressed through text, speech, or sign language" (Saudi Data and Artificial Intelligence Authority, 2022, p. 88).
Language models face challenges when processing Arabic compared to other languages. Despite efforts to develop specialized Arabic language models, the main challenges include:

Complexity of the Arabic Language:
Arabic is a rich and intricate language, characterized by a complex morphological and grammatical system compared to other languages. This complexity makes it difficult for AI tools to accurately understand Arabic.

Lack of Data:
AI faces significant challenges when dealing with low-resource languages, as the amount of available training data is limited and often of low quality. Language models typically rely on two main sources for training:
Wikipedia: Often based on translations from English.
News articles: Sometimes lacking accuracy and mixing formal Arabic (Fusha) with colloquial expressions.
In some cases, machine translation is used to fill the data gap, but translation errors can further degrade the quality of the models rather than improve them.

Dialectal Variations:
The existence of various Arabic dialects alongside Modern Standard Arabic (MSA) makes it even more challenging for language models to understand and process all contextual variations accurately.
Despite these challenges unique to the Arabic language, several research initiatives have emerged to develop specialized Arabic language models, such as:

ALAM:
The Saudi Data and Artificial Intelligence Authority (SDAIA) announced that the ALAM model achieved first place globally as the best generative Arabic language model according to the standardized Arabic MMLU evaluation. This achievement highlights its advancements in Natural Language Processing (NLP). The announcement was made during the Global AI Summit, with the participation of SDAIA leadership, reaffirming Saudi Arabia's commitment to digitally enhancing the Arabic language as part of its Vision 2030 goals.
ALAM was developed at the National Center for Artificial Intelligence by national experts, in collaboration with 160 government entities. The model was trained on 500 billion Arabic text units and released in multiple versions, ranging from 7 to 70 billion parameters. Additionally, IBM has made ALAM available on its watsonx.ai platform, while Microsoft has integrated it into its Azure services, reflecting global confidence in its capabilities for Arabic text analysis and intelligent content generation (https://www.sba.sa/ar/nmwdhj-lawm-alafdl-almyana-dmn-fyth-byn-alnmadhj-altwlydyt-fy-alalm-ballght-alrbyt).

AraGen Leaderboard Model:
The AraGen Leaderboard is a framework designed to redefine the evaluation of large Arabic language models (LLMs). It was launched by Inception, a subsidiary of G42, in collaboration with Mohamed bin Zayed University of Artificial Intelligence (MBZUAI). The model introduces an innovative metric tailored for the Arabic language, called "3C3H," which evaluates AI models across six key dimensions: correctness, completeness, conciseness, helpfulness, honesty, and harmlessness (https://inceptionai.ai/2024/12/04/rethinking-llm-evaluation-with-3c3h-aragen-benchmark-and-leaderboard/index.html).

Jais-13B-Chat Model:
The Jais-13B-Chat is a bilingual (Arabic-English) generative language model with 13 billion parameters. It was developed through a collaboration between the Inception Institute of Artificial Intelligence (IIAI), Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), and Cerebras Systems.
The model is based on the Transformers architecture and utilizes ALiBi positional encoding, enhancing its ability to handle long text sequences and improving contextual accuracy.
Key features of the model include: the ability to process complex tasks, handling nuances of the Arabic language effectively, and applications in smart assistants, customer service, text analysis, and content generation, with a focus on providing accurate and reliable responses in both Arabic and English (https://huggingface.co/inceptionai/jais-13b-chat).

Section Two: Applied Framework
We created a small dataset sourced from ChatGPT, collecting approximately one million questions extracted from available Arabic-language data.

To achieve this, we examined all freely available Arabic-language datasets that we could download, totaling over 141 million sentences, or approximately 2.3 billion words. From this vast collection, we filtered only those sentences ending with a question mark.
From this subset, we selected a sample of questions that begin with an interrogative word, such as: Where, Why, How, When, Who, How much/many, etc. We then submitted these questions to an AI program, recorded their responses, and analyzed them.
For basic syntactic analysis, we used the Stanford POS Tagger (https://nlp.stanford.edu/projects/arabic.shtml) to classify parts of speech, allowing us to determine the most frequently occurring words at the beginning of AI-generated sentences.
Key Comparisons Between AI-Generated and Natural Texts:
First, as shown in the table below, although AI-generated text tends to have longer sentenced (18 words on average) compared to natural text (16 words on average), its lexical diversity is lower.
This means that the AI model relies on a preferred vocabulary, using the same words whenever possible.
As a result, in natural text, a new word appears every 29 words on average. In contrast, in AI-generated text, a new word appears only every 34 words.

	Natural text	AI-generated text
Average number of words in a sentence	16.38	18.07
Lexical diversity	33.89	27.93

Secondly, when examining word sequence repetition in both natural and AI-generated texts, we find that AI-generated sentences frequently begin with a preposition.
Most commonly, these are (in) or (from).
The preposition “in” is often used to introduce historical timeframes (e.g., in the 1960s”).
The preposition “from” frequently appears due to the phrase “It is important…” at the beginning of sentences.
This pattern suggests a strong influence from English, where sentences often begin with a preposition or the phrase “It is important…”
Thirdly, in natural text, there is a balance between verbs and nouns at the start of sentences. However, AI-generated text tends to favor verbs, particularly past tense verbs conjugated for the third-person singular (he/she/it).
This tendency may arise due to a lack of broader context in AI responses, as our dataset consists of independent questions.
When context is missing, the third-person singular becomes the most neutral and widely applicable choice.
Fourthly, natural text exhibits greater syntactic diversity than AI-generated text.
When measuring the ratio of unique five-word sequences to the total number of sentences, the natural text has a higher diversity score (94.72%), compared to AI-generated text (87.70%).

Conclusion:

Generative AI plays a significant role in shaping the future of the Arabic language. However, it faces challenges related to data quality and accuracy. By improving data sources and enhancing training techniques, it is possible to boost the efficiency and effectiveness of Arabic language models, making them better at representing and processing the language.

Recommendations

1. Developing rich Arabic language datasets to support more effective model training.
2. Enhancing the accuracy of machine translation between Arabic and other languages.
3. Creating more advanced virtual assistants capable of understanding Arabic in different contexts.
4. Supporting Arabic e-learning platforms with more accurate AI-driven content.
5. Conducting comparative studies on the impact of AI on Arabic versus other languages. This would provide a broader perspective on how AI technologies evolve in supporting Arabic compared to languages like English or French.
6. Addressing cultural and linguistic challenges in AI-based Arabic learning, including the differences between Modern Standard Arabic (MSA) and dialects, and their impact on the accuracy of AI models in comprehension and text generation

Acknowledgement:

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Group Project under grant number. (RGP.2/ 309/ 45)

References

1. Abdel Qader, Abdel Razek Mukhtar Mahmoud Rashwan, Ahmed Mohamed Ali and Abdel Gawad, Ahmed Abdel Fattah Abdel Wahab (2023). Applications of Artificial Intelligence and Their Impact on Developing Creative Linguistic Skills in Outstanding High School Students. Journal of the College of Education, Vol. 39(1), pp. 109-135. https://search.emarefa.net/detail/BIM-1460590
2. Fitriana, Saif (2023). The Impact of Artificial Intelligence on Teaching the Arabic Language. International Conference for Graduate Students in Arabic Language Teaching, Literature, and Linguistics (ICON-POSTALL). https://prosiding.arab-um.com/index.php/ICON-POSTALL
3. Al-Shammari, Najwa Mohammed (2023). Artificial Intelligence Applications and Their Role in Developing Smart Content from the Perspective of Faculty Members. Journal of Educational and Psychological Sciences, 7(41), 26-46. https://doi.org/10.26389/AJSRP.N160823
4. Zaki, M., & Ali, A. (2024). Integrating Artificial Intelligence Technologies in Teaching Arabic: A Study on Adaptive Educational Materials. Presented at the International Conference on Artificial Intelligence and Linguistics, Sharjah, October 17-18, 2024, American University of Sharjah & Emirates Center for Research and Studies.
5. Mulyanto, D., Zaki, M., & Redo, A. M. A. (2024). The Use of Artificial Intelligence to Develop Arabic Language Skills in Learning. An-Nidzam: Jurnal Manajemen Pendidikan dan Studi Islam, Darussalam Gontor University, 11(1), 18-32
6. Al-Amairah, M. I. (2025). The Reality of the Arabic Language in AI Applications: A Descriptive Applied Study. Dirasat: Human and Social Sciences, 25(2), 459-471.
7. Saudi Authority for Data and Artificial Intelligence (SDAIA). (November 2023). Generative Artificial Intelligence.
8. Lost in Translation: Large Language Models in Non-English Content Analysis (May 22, 2023), Center for Democracy & Technology (CDT).
9. https://www.statista.com/statistics/262946/most-common-languages-on-the-internet/
10. Saudi Data and Artificial Intelligence Authority (SDAIA) (2022) Data and Artificial Intelligence Dictionary in Arabic and English.
11. https://www.sba.sa/ar/nmwdhj-lawm-alafdl-almyana-dmn-fyth-byn-alnmadhj-altwlydyt-fy-alalm-ballght-alrbyt.
12. https://inceptionai.ai/2024/12/04/rethinking-llm-evaluation-with-3c3h-aragen-benchmark-and-leaderboard/index.html.
13. https://huggingface.co/inceptionai/jais-13b-chat.
14. https://nlp.stanford.edu/projects/arabic.shtml.