

---

**RESEARCH ARTICLE**

## Machine-Learning-based English Quranic Translation: An Evaluation of ChatGPT

Ismail Dahia<sup>1</sup> ✉ and Mohamed Belbacha<sup>2</sup>

<sup>1</sup>Laboratory of Applied Communication in Context, Mohamed First University, Oujda, Morocco

<sup>2</sup>Multidisciplinary Faculty of Nador, Mohamed First University, Oujda, Morocco

**Corresponding Author:** Ismail Dahia, **E-mail:** [Ismail.dahia@ump.ac.ma](mailto:Ismail.dahia@ump.ac.ma)

---

**ABSTRACT**

Translating the Holy Quran into different languages is a crucial task. It is a daunting one, given the linguistic complexity and rich cultural context of this holy scripture. The recent advancements in technology, especially in artificial intelligence, brought about a pressing need to explore the potential of machine learning-based tools in easing the translation of religious texts, the Holy Quran in particular. This study explores the potential of adopting ChatGPT, a large language model (LLM) powered by machine learning, in English Quranic translation. The study aims to assess the accuracy, adequacy, and cultural sensitivity of ChatGPT's English Quranic translation and identify potential challenges and limitations of using this tool in this domain. The study employed a comparative analysis approach, evaluating ChatGPT's translation of the El Fatiha Chapter against five human-generated translations using BLEU (Bilingual Evaluation Understudy) and METEOR (Metric for Evaluation of Translation with Explicit ORdering) machine translation metrics. The study also utilized an interpretation of the chapter by El Qurtubi to ensure that ChatGPT has the entire linguistic and cultural context of the chapter. The study's findings indicated that ChatGPT's English translation of El Fatiha achieved a moderate level of accuracy and adequacy with mean BLEU and METEOR scores of 0,48 and 0,78, respectively. However, there were areas where ChatGPT's translation could be improved. The use of ML-based tools such as ChatGPT in English Quranic translation offers promising benefits as it can aid translators in the translation process.

**KEYWORDS**

Quranic; translation; ChatGPT; BLEU; METEOR; adequacy; linguistic

**ARTICLE INFORMATION**

**ACCEPTED:** 15 July 2024

**PUBLISHED:** 15 August 2024

**DOI:** 10.32996/ijllt.2024.7.8.17

---

### 1. Introduction

Translating the Holy Quran is often described as a daunting challenge due to its rhetoric and culturally loaded nature. Muslim scholars and translators often argue that true translation is impractical, resulting in the distortion of the original meaning of the scripture (Von Denffer, 1983). Given that over 80% of Muslims do not speak Arabic (World Population Review, 2023), it is a necessity to achieve an error-free translation that captures the essence of the original script so that the non-Arabic-speaking population experiences the true sense of the Holy Quran.

In order to address the challenges mentioned above, scholars have thought about benefiting from technological breakthroughs, including machine-learning-based tools, a paradigm-shift technology, in translating the Holy Quran. The incorporation of machine learning represents a pioneering step toward Quranic renditions as it may provide translations that are culturally sensitive and linguistically precise. Although machine learning has been employed in the study of the Quran (Al Anazi and Shahin, 2022; Al Ghamdi and Khan, 2022; Alkhateeb, 2020; Khan et al., 2019; Adeleke, 2018), it was only confined to the evaluation of previously rendered translations of the Holy Quran. Few studies (Hamed et al., 2021) have looked at the capabilities of machine-learning-based tools and how they could help in the rendition of the Holy Quran, thus addressing the previously mentioned challenges inherent to the translation of the Holy Quran.

Hence, the primary aim of the study is to explore the effectiveness of integrating a machine learning-based tool, ChatGPT, into the translation process. Five scholarly recognized translations of El-Fatiha Chapter (The Opening) were given to ChatGPT to exploit ChatGPT's ability to learn and adapt, which are some of the machine-learning-based tools' capabilities (Taecharungroj, 2023). A scholarly recognized interpretation of the El-Fatiha Chapter by El Al-Qurtūbī (2003) was provided to ChatGPT to have an understanding of the linguistic and cultural context of the Chapter, making the translation as accurate and adequate as possible. After ChatGPT generated its own translation, a comparative analysis was conducted to examine the linguistic accuracy and cultural adequacy of ChatGPT. The comparison included ChatGPT's translated version of the El Fatiha Chapter and the five versions it learned from. Hence, this research aims to answer the following questions:

- (1) How accurate and adequate is ChatGPT in translating the Holy Qur'an?
- (2) To what extent can machine-generated translation convey rhetorical and cultural elements present in the Holy Qur'an?
- (3) What are the challenges and limitations encountered when using ChatGPT for translating the Holy Quran?

Investigating the power of machine-learning-based tools such as ChatGPT, Google Bard AI, Chatsonic, etc., can help scholars and translators make informed decisions about how the inherent challenges of Quranic Translation can be addressed and overcome.

## 2. Literature Review

The study of the Holy Quran holds immense significance. This section is divided into three main parts: challenges of Quranic translation, machine translation, and machine learning and Quranic translation.

### 2.1. Challenges of Quranic Translation

The importance of the study of the Holy Quran lies in its being the most sacred scripture for two billion people across the globe, whether by Arabic-speaking countries, which represents about 20% of the Muslim population, or by non-Arabic-speaking countries, representing 80% worldwide (World Population Review, 2023). This indicates a growing need for the translation of the Holy Quran, ensuring that the divine message is effectively conveyed and understood. Indonesia, a non-Arabic-speaking country, is the country that hosts the largest number of Muslims in the world, representing about 13%.

Although it may seem at first evident that it is practically possible to render translations of the Quran for non-Arabic speakers to understand the word of Allah and His teachings, there are a number of Muslim scholars (Arberry, 1957; Pickthall, 1971; Von Denffer, 1983) who believe otherwise. For instance, Arberry (1957) and Pickthall (1971) perceive the Holy Quran as being untranslatable. In the introduction to his work of translating the Holy Quran, *The Koran Interpreted*, Arberry argues that the uniqueness of the Holy Quran lies in its eloquent and rhetorical language, which is impossible to render in any language. This explains the titling of his translation as *The Koran Interpreted* rather than "The Koran". In the same vein, Von Denffer (1983) posits in his book *Ulum al-Qur'an: An Introduction to the Sciences of the Qur'an* that it is generally accepted among Muslim scholars that rendering the Holy Quran into another language without distorting the original meaning is an impossible task for the translator. Fazlur Rahman (1988) shares the same view that the Quran, as a divinely inspired text that is inextricably linked to the Arabic language, is difficult to render into any language.

Other studies (Tibawi, 1962; Raof, 2001; Qassem, 2021) also discuss the challenges of rendering the Holy Quran and thus argue for its untranslatability. For example, Tibawi (1962) contends that Arabic, being an exceedingly rich language with a concise and metaphorical vocabulary, is unparalleled in its depth; thus, any attempt to translate it is doomed to failure. Likewise, Raof (2001) provides examples of the aspects that are inherent to Quranic Arabic but alien to other languages. He emphasizes that Quranic discourse is not just a mere sequence of words; rather, it represents a linguistic texture with rich and diverse elements. These elements encompass not only syntactic and semantic aspects but also rhetorical and cultural features. The scholars above have all argued for the non-equivalence of Quranic translation, which is the most cited challenge whenever a translation is conducted.

The large number of Quranic translations further explains the complexity of the task of rendering the Holy Quran, indicating that it is almost impossible to have a universally agreed-upon translation that is error-free while preserving the rhetoric and idiomatic nature of the Holy Quran. This is one of the reasons why some translators have recourse to a whole description and annotations in order to convey the meaning and power of single Quranic words. This often results in the distortion of the eloquence of the Quranic text. This is emphasized by Larson (1997), who argues that a direct translation is not always possible or effective. Translators often need to adapt the language to fit the target audience and culture. This may involve using different words, phrases, or even sentence structures to convey the same meaning.

Since it can be argued that all existing translations of the Holy Quran fall short of fully preserving the original meaning and the essence of the script, these translations can all be labeled as poor (Menacere, 1999). In view of these inherent challenges that come with translating the Holy Quran, more research must be conducted to address these challenges. While the studies above discussed

the issues associated with the translation of the Holy Quran, they did not propose well-defined and effectively grounded solutions to resolve or, at least, minimize the inaccuracies found in most Quranic translations.

## **2.2. Machine Translation**

Machine translation has substantially evolved since its inception. The rise of artificial intelligence has brought about profound changes in how MT is viewed and used. Since the aim of the current study is to delve into the extent to which MT can aid in rendering the Holy Quran, it is crucial to look at the various definitions proposed by different scholars, as this will make us understand the nature as well as the scope of MT's use.

The term 'machine translation' was originally devised by Warren Weaver in 1949 (Wang et al., 2020). According to Hutchins (1986), machine translation (MT) is the automated conversion of texts or speeches from one language to another, emphasizing its automatic nature and distinguishing it from human-based translations. Similarly, Wilks (2009) and Kay (1982) define MT as the utilization of computer algorithms for translating natural languages. Likewise, Vauquois (1998) characterizes MT as a technology enabling computers to translate natural language text across different languages. This process involves the analysis of the text on lexical, syntactic, and semantic levels to produce an adequate and meaningful translation.

Despite its potential usefulness, Kay (1982) notes that human-aided MT tends to outperform fully automated translations. This perspective aligns with Melby's (1998) observations on the limitations of machine translation, stating that MT is most effective when dealing with less complex and highly repetitive textual content encountered in everyday language use. This view is also supported by Somers et al. (1990), who contend that MT systems have traditionally operated on the assumption that source texts inherently encapsulate almost all the information required for adequate translation. However, this is not true, given the fact that the two languages have distinct linguistic and cultural is highly conspicuous. For Somers et al. (1990), these differences between languages significantly affect the quality of translation performed by MT systems. A more recent definition was provided by Forcada (2018), who defines MT as the application of computer-assisted software to translate a text from a source language (SL) to a target language (TL). He emphasizes the importance of distinguishing between MT and MT-aided translation, where professionals use various computer-based tools in their translation processes.

The different definitions above provide both a theoretical and practical framework for approaching the use of AI-based tools in translating the Holy Quran. This also helps with the sensitivity required for translating religious texts. For instance, Melby's perspective on the limitations inherent to MT helps scholars and translation enthusiasts set realistic goals regarding the effectiveness of AI in rendering the Holy Quran.

## **2.3. Machine Learning and English Quranic Translation**

Since the very first translation of the Holy Quran by Salman, the Persian (Yahaghi, 2002), numerous scholars have attempted to provide more elaborate translations that were linguistically and culturally adequate. As the world advances, more innovative technological tools emerge. Thanks to the fact that several scholars try to make the best of these tools to have more effective Quranic translations, a number of translation problems could be overcome. For Kay (1982), the growing demand for translation and the challenges faced by human translators could only be effectively addressed through technology. These challenges, according to him, can barely be overcome by "better language teaching, greater incentives for translators, or improved administrative procedures, worthy though these goals undoubtedly are." (p.74)

Machine learning is a great example of such technological advancements. The incorporation of such advancements into different fields, such as medicine, teaching, religion, and translation, has become a necessity. Given the different challenges encountered by most Muslim and non-Muslim translators and scholars in studying the Holy Quran, numerous scholars have begun considering the incorporation of machine-learning-based tools or artificial intelligence-based tools (AI) to make the study of the Holy Quran easy and time-saving. This is emphasized by Adeleke (2018), who posits that the automation of the study of the Quran, including verse analysis, Quranic translation, and other related studies, is an essential requirement for the improvement and understanding of Quranic studies.

The following are some of the advantages of machine learning suggested by Khanzode and Sarode (2020):

- It makes fewer errors.
- It saves time.
- It has the ability to multi-task.
- It has a high success rate.
- It can discover and learn things.

There are a number of scholars (Adeleke, 2018; Khan et al, 2019; Alkhateeb, 2020; Al Ghamdi and Khan,2022; Al-Anazi and Shahin, 2022) who drew the attention to the advantages above to study different aspects of the Quran. For instance, Al Ghamdi and Khan (2022) utilized computational methods to compare different Quranic translations. On the other hand, other scholars such as Alkhateeb (2020), Al-Anazi and Shahin (2022), and Khan et al. (2019) adopted AI to study how it can help in recognizing Quran reciters using different speech recognition models. Adeleke et al. (2018) conducted a study to automatically label Quranic verses using a machine learning approach. They aimed to classify Quranic verses into three predefined categories: faith, worship, and etiquette. To achieve this, they employed four machine learning classifiers: Support Vector Machine (SVM), Naive Bayes (NB), J48, and k-Nearest Neighbors (k-NN). The results revealed that all four classifiers performed well, achieving an accuracy of over 80%. Notably, the Naive Bayes algorithm outperformed the others, achieving the highest accuracy of 93.9% and an Area Under the Curve (AUC) score of 0.964. Adeleke et al.'s study demonstrates the effectiveness of using ML-based models to study and understand different aspects of the Holy Quran.

Although a growing body of research has dealt with the use of technology in Quranic translation, few studies have dealt with AI-powered Quranic translation. For instance, Hamed et al. (2021) used machine learning in translating the Holy Quran. However, the language studied was Italian. Furthermore, no interpretations were used in the machine learning models (LSTM and GRU), which is something that the current study utilized. As for Al-Sukhni et al. (2016), while they evaluated ML-based translations of the Holy Quran, this evaluation was based on Google Translate and Bing. It should be noted that the previously-mentioned ML-based tools are translation-specific, unlike the recently developed ML tools such as ChatGPT, Google Bard, Chatsonic AI, etc.

Therefore, the study aims to examine ChatGPT's efficacy as a machine-learning tool in translating the Holy Quran into English. El Fatiha is the used chapter. The selection of El Fatiha as the focal chapter was based on its notable significance in the Holy Quran. Due to the myriad shortcomings often found in almost all the renderings of the Holy Quran, it is the author's aim to examine how effective a machine-generated translation of the Holy Quran is. ChatGPT's capacity for accurately and contextually translating the sacred text is explored by evaluating and understanding how machine-generated translation compares in terms of accuracy, context, and adequacy to human-generated translations.

### **3. Methodology**

#### **3.1. Sampling Procedures and Data Collection Instruments**

Given the impracticality of empirically studying the English translation version of the whole Holy Qur'an within the scope of one study, the El Fatiha Chapter (The Opening) is used as the main sample from which conclusions are drawn regarding the effectiveness of ChatGPT in rendering the Chapter into English. To achieve this aim, an interpretation of the Chapter by Al-Qurtubi and five widely-used and scholarly recognized translations are used. The five translators included in the study, listed in chronological order from oldest to most recent, are:

1. Abdullah Yusuf Ali (1934), *The Holy Qur'an: Text, Translation and Commentary*.
2. Thomas Ballantyne Irving (1985), *The Qur'an: First American Version; Translation and Commentary*.
3. Mohammed Marmaduke Pickthall (1960), *The Meaning of the Glorious Koran*.
4. Muhammad Taqi Din Al-Hilali and Muhammad Muhsin (1996), *The Noble Quran: English translation of the meaning and commentary*.
5. Muhammad A.S. Abdel Haleem (2004), *The Qur'an*, Oxford University Press.

The main rationale behind the selection of these translators is to provide ChatGPT with a comprehensive and diverse understanding of Quranic translations, considering different linguistic styles, interpretations, and cultural contexts over time. By utilizing these varied translations, the objective is to expose ChatGPT to a wide range of linguistic styles, cultural contexts, and interpretive methods present in the five Quranic translations above. The goal is also to enable ChatGPT to generate its own translated version of the El Fatiha chapter. ChatGPT's translation drew on the linguistic and cultural insights it gained from the above respective translations. ChatGPT is also part of the main data collection methods as it is used to generate its own translated version of the Chapter after comparing the five translations above, learning from them, and then providing its own translation.

Below is the translated version of the El Fatiha Chapter provided by ChatGPT

' In the name of Allah, the Most Compassionate, the Most Merciful. Praise be to Allah, the Lord of all realms, The Most Compassionate, the Most Merciful. Master of the Day of Judgment. You alone we worship, and You alone we ask for help. Guide us on the Straight Path, The path of those who have received Your grace; not the path of those who have brought down wrath upon themselves, nor of those who have gone astray.

### **3.2. Data Collection and Data Analysis Procedures**

After generating its own translated version of the Quranic Chapter, ChatGPT's translated version was compared with five human translations, as Table 1. shows. BLEU (Bilingual Evaluation Understudy) and METEOR (Metric for Evaluation of Translation with Explicit ORdering), two widely used metrics for machine translation evaluation, are adopted. This comparative analysis aims to evaluate the strengths and limitations of the machine-learning approach to Quranic translation.

#### **3.2.1. BLEU and METEOR Metrics**

Since ChatGPT is at the center of the current study, it is of paramount importance to measure how close and adequate ChatGPT's translation is to the five translations. The rationale behind using these machine translation evaluation metrics lies in their ability to judge the closeness of a machine translation to one or more reference translations (Papineni et al., 2002; Banerjee & Lavie, 2005). Furthermore, these automated metrics are crucial tools for data analysis as they provide quantitative, objective measures of translation quality, thus enabling us to assess the performance of machine learning models (Banerjee, 2005).

According to (Papineni et al., 2002), the way BLEU works is by measuring the precision of machine-powered translations by comparing them to a set of human translations referred to as reference translations. A higher BLEU score indicates a closer match to the reference translations, signifying improved translation quality. METEOR, on the other hand, is more sophisticated and comprehensive in its scores. This is because it takes into account synonyms, stemming, and word order. In addition to that, it provides a comprehensive assessment of the fluency and naturalness of the translations (Lavie et al., 2004).

In the current study, ChatGPT's translation is compared to the five reference translations to calculate both BLEU and METEOR scores. The results showed the similarities and differences between ChatGPT's output in comparison with the reference translations in terms of vocabulary, syntax, and overall fluency.

#### **3.2.2. Description of Data Analysis Tools**

The evaluation of ChatGPT's translation was carried out using a Python-based package called pymeteor and an NLTK (Natural Language Toolkit) library's BLEU calculation function. This package offers a meteor function that calculates METEOR scores when reference and candidate translations are provided. Prior to calculating the BLEU and METEOR scores, both the reference and candidate translations were tokenized using Python. The tokenization of the translations involved breaking down the text into smaller units called tokens so as to ease the calculation of both BLEU and METEOR scores, which rely on matching tokens between the reference and candidate translations (Papineni et al., 2002; Banerjee & Lavie, 2005). For each reference translation, the METEOR function was invoked, independently assessing ChatGPT's output against that specific reference. The process was repeated five times. This approach allowed for individualized evaluation against each reference.

As for the BLEU scores, the NLTK library's BLEU calculation function was utilized to independently assess ChatGPT's output against that specific reference. The NLTK library's BLEU calculation considered the precision of matching n-grams (individual words) between ChatGPT's output and the reference translation, producing BLEU scores. These scores measure the degree of overlap between the candidate and reference translations.

## **4. Results and Discussion**

As Figure 1. indicates, the BLEU and METEOR scores for each translation range from 0.309 to 0.699 for BLEU and 0.528 to 0.890 for METEOR. The highest BLEU score is 0.699, which is achieved by Thomas Irving's translation. The highest METEOR score is 0.890, which is also achieved by Thomas Irving's translation.

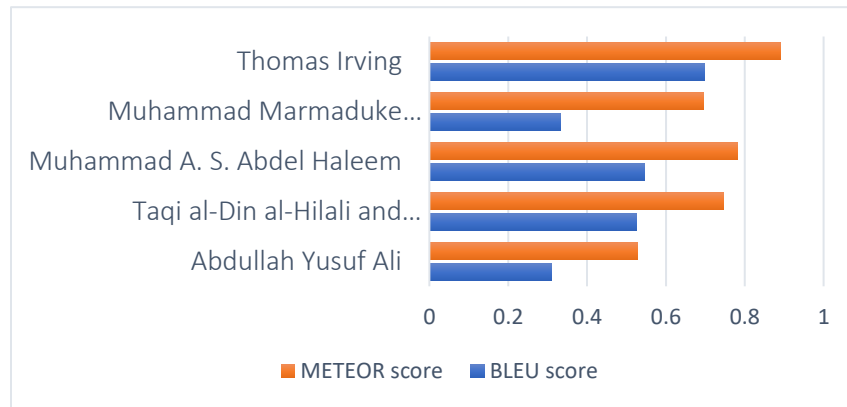


Figure 1: BLEU and METEOR scores of ChatGPT's translation

In analyzing the El-Fatiha chapter translation performed by ChatGPT, we can also see in Figure 1. that its results differ from those achieved by the human translator. The METEOR and BLEU scores show that the closest translation to ChatGPT was that of Thomas Irving, with METEOR and BLEU scores of 0.89 and 0.69, respectively. The first remark concerns ChatGPT's use of "realms" in rendering the word (العالمين) 'Al-'Ālamīn'. The meaning of the word "realm" found in the dictionary does not match, even metaphorically, that of "world". According to Oxford Learner's Dictionaries (n.d.), the word "realm" describes a particular area of activity, interest, or knowledge. It can refer to a domain or sphere, such as the realm of science or the realm of literature. Additionally, it may denote a kingdom or a country ruled by a monarch. On the other hand, "world" typically refers to the entire planet Earth, including all its countries and peoples. In a broader sense, it can refer to various aspects of human life or the entirety of existence. The latter is consistent with El Qurtubi's (2003) interpretation of the word as he argues that "world" is used to denote all that exists. It is worth noting that the full meaning of the word was and still is the subject of much controversy among Quran interpreters (El Qurtubi, 2003). For him, interpretations vary from those who claim that the word encompasses everything that has a soul to others who believe that it denotes all that exists. In all cases, it is evident that ChatGPT failed to capture the meaning of Arabic since the word it opted for does not capture the intended meaning of the original word. The five translations, however, have all succeeded in conveying the intended meaning through the use of different phrasings (e.g., the Lord of all beings, Lord of worlds, the Lord of all that exists).

Furthermore, while Irving used "The Compassionate" and "The Merciful" ('الرَّحِيمُ، الرَّحْمَنُ') 'Ar-Raḥmān, Ar-Rahīm' to describe Allah's attributes, ChatGPT used "The Most Compassionate" and "The Most Merciful". Although both translations indicate how merciful and compassionate Allah is, ChatGPT's version laid more emphasis on these attributes by using the superlative "Most", which is used to indicate the highest degree of the qualities of compassion and mercy, which can't be surpassed. This emphasis was also present in other translations, such as those of El Hilali and Khan and Yusuf Ali, who used "The Most Gracious, The Most Merciful". Other translations used similar wording to describe these qualities, ranging from "The Beneficent and The Merciful" to "The Lord of Mercy, The Giver of Mercy".

The METEOR score of 0.89 could be explained by the closeness of wording and phrasing between ChatGPT's translation and Irving's. This is because, as mentioned before, METEOR scores are sensitive to differences in word choice and sentence structure. This applies to the other translations that have received lower METEOR scores, such as those of Yusuf Ali.

As for (إِيَّاكَ تَعْبُدُ وَإِيَّاكَ تَسْتَعِينُ) 'Iyyāka na'budu wa iyyāka nasta'in', while Irving laid emphasis on the verse "It is You we worship; it is You we ask for help" by using a cleft sentence whose aim is to emphasize a particular part of a sentence, ChatGPT's use of "You alone" also emphasizes Allah as the one being worthy of worship and the one all beings ask help from by using "alone" and topicalization. The topicalization was achieved by moving the object of the sentence "You" to the front. This is similar to Pickthall and El Hilali and Khan, who used "alone" and "only" respectively. On the other hand, Abdel Haleem and Irving shared the same translation of the verse by using the it-cleft sentence structure.

In translating (أَهْدِنَا الصِّرَاطَ الْمُسْتَقِيمَ) 'Ihdinā aṣ-ṣirāṭa al-mustaqīm', ChatGPT's translation of the verse was identical to the translations by Irving, El Hilali and Khan, and Abdel Haleem, all of whom translated the verse as "Guide us on the Straight Path." On the other side, Pickthall and Yusuf's translations differed from the others by rendering the Arabic word (أَهْدِنَا) into "show". It seems that when ChatGPT was presented with the five translations, it opted for the closest word that conveyed the Quranic meaning (guide). This best aligns with the Quranic meaning of the word since (أَهْدِنَا) reflects the concept of seeking guidance and being led along the right path. Conversely, Pickthall and Yusuf's translations of the word imply revealing or indicating but do not necessarily convey the sense of being guided or directed toward the right path.

ChatGPT's translation of (صِرَاطَ الَّذِينَ أَنْعَمْتَ عَلَيْهِمْ غَيْرِ الْمَغْضُوبِ عَلَيْهِمْ وَلَا الضَّالِّينَ) 'Siraṭa alladhīna an'amta 'alayhim ghayri l-maḡḏūbi 'alayhim walā ḍ-ḍāllīn.' verse is identical to those of Irving and Abdel Haleem's. This is consistent with the METEOR scores of ChatGPT's translation, 0.890 and 0.781 respectively, and its closeness to theirs. This is also true for BLEU scores (0.69 and 0.56 respectively) as Figure 1. shows. In contrast, Hilali and Khan's translation is identical to that of Pickthall's while Yusuf's translation is different from all the previous translations because of his use of archaic terms in his translation such as "Thou", "hast" and "Thy". It should be noted that while Abdel Haleem, Hilali and Khan, and Irving retained the same translation of the word (صِرَاط) as "the path", Hilali and Khan's translation of the word changed from "the path" to "the way". This also explains why ChatGPT's translation had the lowest METEOR score of 0.528 and BLEU score of 0,30, when compared to that of Yusuf's, indicating potential issues with precise word choice and semantic accuracy. Oxford Learner's Dictionaries (n.d.) define way in different manners. According to Oxford, it primarily denotes a method, style, or manner of doing something, while also referring to a particular mode of behavior or the typical way of life for a specific group. Additionally, it can indicate a route or road taken to reach a destination. The word path is defined as a way or track created by the action of people walking, suggesting a physical route or trail. However, it extends beyond this physical meaning to convey a plan of action or a strategic approach to achieving a goal. In a similar vein,

Overall, it is evident that the machine-generated translations achieved a moderate level of accuracy and adequacy. The mean BLEU score ( $\bar{x} = 0.48$ ) indicates that the translations were able to capture about 48% of the n-grams (matching words) found in the reference translations. This suggests that the translations were generally accurate in terms of conveying the meaning of El Fatiha chapter. The METEOR mean score of ( $\bar{x} = 0.72$ ) suggests that the translations were also relatively fluent, meaning that they read smoothly and naturally in English. This is a promising result, as it indicates that ChatGPT is not simply producing literal word-for-word translations but is able to capture some of the nuances of the chapter in question.

The results of this study suggest that machine learning-based tools have the potential of rendering the Holy Quran. The BLEU scores of the ChatGPT's translation were relatively small compared to METEOR scores, which were relatively high. As mentioned above, a mean BLEU score of 0.48 (less than 50% similarity to the chapter in question) indicates that ChatGPT's translation's quality could be improved. However, it should be noted that given that BLEU scores are generally based on matching n-grams between the reference translation (human translations) and machine translation (Papineni et al., 2002; Banerjee & Lavie, 2005), the scores only show the word for word matching and does not account for other translation criteria such as the cultural aspect of the text under examination. This means that a high or low BLEU score does not necessarily mean that the ChatGPT's translation of El Fatiha is accurate or not, as there are other variables that are unaccounted for. The METEOR scores, on the other hand, capture more linguistic and cultural nuances (Lavie et al., 2004). Based on this, a mean METEOR score of 0.72 suggests that ChatGPT's translation was not only accurate in terms of capturing the overall meaning of the El Fatiha chapter but also producing fluent and natural-sounding language. As mentioned previously, the Holy Quran contains rich imagery, symbolism, and cultural references, making it a challenge to translate accurately.

Overall. While there were some problems with word choice in ChatGPT's translation, it is evident that the general intended meaning was captured. There was little ambiguity in the translation. This might have been due to El Qurtubi (2003) interpretation that was provided to ChatGPT prior to translating the Chapter, without which the overall accuracy and adequacy would have been different, thus affecting the overall BLEU and METEOR scores. This is because the interpretation provided it with the necessary context required to understand the nuances of the Chapter. With regard to cultural adequacy, it is important to note that the ML-based translation took into consideration the cultural context of the chapter by using language that resonates with the religious and cultural sensibilities of the Chapter. This was achieved through the effective use of key religious terms (e.g., In the name of Allah, Praise be to Allah) that hold significant importance in Islamic tradition. This is helpful as it ensures that the text is not only linguistically accurate but also culturally adequate and accessible to readers.

The fact that ChatGPT was able to achieve a high METEOR score suggests that it is capable of handling these complexities and producing translations that are faithful to the original scripture. This, however, does not mean that machine learning tools can replace human translators, as the human touch is always needed. For Giammarresi and Lapalme (2016), the quality of a given translation will never be achieved by MT. They also posit that such limitations are inherent to the translation process in general. This inability to comprehend the deeper meaning of language hinders the ability of machines to produce translations that are truly equivalent to the original texts, especially religious ones. It should be noted that Giammarresi and Lapalme (2016) discussed the impossibility of achieving more accurate translations than those performed by human translators, but they did not mention how the task of translation is further complicated when dealing with more nuanced and sensitive texts like religious text. This indicates that no matter how advanced AI gets, it cannot exceed humans' potential.

As mentioned earlier, the successful training of a given model depends heavily on the amount of data it was fed (Chai et al., 2022). In the case of ChatGPT, only five translations were used, and this might have had a big impact on the translation it provided, thus affecting both the BLEU and METEOR scores. This means that further research might be focused on training ML models whose

sole purpose is to automate the translation of the Holy Quran, thus producing more readable and adequate translations. In other words, the inherent challenges associated with rendering the Holy Quran require machine learning models that are specifically tailored to the task of Quranic translation.

## 5. Conclusion

The translation of the Holy Quran presents a challenge for many translators, given the linguistic complexity and the rhetorical nature of the scripture. The advent of ML has offered promising tools to assist in Quranic translation. This study explored the potential of ChatGPT, a large language model powered by ML, in Quranic translation. ChatGPT's translation of El Fatiha, the first chapter of the Quran, was analyzed and evaluated to assess its accuracy, fluency, and cultural sensitivity.

The findings of the study showed that ChatGPT has the potential to help with the translation of the Holy Quran. They also showed that BLEU scores were relatively small compared to METEOR scores. The closest translation to ChatGPT's was that of Thomas Irving's, with BLEU and METEOR scores of 0.69 and 0.89, respectively. While the BLEU scores might indicate room for improvement in terms of word-to-word correspondence, the overall METEOR scores, especially in comparison to Thomas Irving's translation (89% similarity) and the overall METEOR mean score of 0.72 (72% similarity), highlight the potential of ML-based tools, especially ChatGPT, in performing translations that capture the essence of the scripture with a greater degree of linguistic and cultural adequacy. This also highlights the proficiency of the METEOR metric in evaluating machine translations. These high METEOR scores also confirm previous findings on the effectiveness of this ML evaluation metric compared to BLEU (Saadany and Orasan, 2021). It should be borne in mind that the aim of the study was not to evaluate the effectiveness of BLEU and METEOR metrics. Instead, the main objective was to investigate the extent to which ML-based tools such as ChatGPT can aid in rendering the Holy Quran. While the findings showed that total dependence on such tools might not guarantee an error-free translation of the Holy Quran, these tools can still offer promising results as far as enhancing the translation. However, one should bear in mind that human oversight and review are required and remain essential as they can ensure the quality and integrity of the ML-based translation. This will also ensure that the message of the Holy Quran is not distorted and misunderstood. This is consistent with Giammarresi and Lapalme (2016), who posit that MT's inability to comprehend complex language will never be capable of matching the original meaning. This is further complicated when religious texts are concerned.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Acknowledgments:** We acknowledge that no people, no grants, and no funds have been received by any funding organizations.

## References

- [1] Abdel Haleem, M. A. S. (2004). *The Qur'an*. Oxford, UK: Oxford University Press.
- [2] Abdul-Raof, H. (2001). *Arabic Stylistics: A Coursebook*. Otto Harrassowitz Verlag.
- [3] Adeleke, A., Samsudin, N., Mustapha, A., & Khalid, S. A. (2018). Automating Quranic Verses Labeling Using Machine Learning Approach. *Indonesian Journal of Electrical Engineering and Computer Science*, 10(1), 925-931.
- [4] Al-Anazi, M. M., & Shahin, O. R. (2022). A Machine Learning Model for the Identification of the Holy Quran Reciter Utilizing k-nearest Neighbor and Artificial Neural Networks. *Inf. Sci. Lett.*, 11(4), 1093-1102.
- [5] Al-Ghamdi, N. M., & Khan, M. B. (2022). Assessment of Performance of Machine Learning Based Similarities Calculated for Different English Translations of Holy Quran. *International Journal of Computer Science & Network Security*, 22(4), 111-118.
- [6] Al-Hilali, T., & Khan, M. (1996). *The Noble Quran: English Translation of the Meanings and Commentary*. Madinah: King Fahd Complex for the Printing of the Holy Quran.
- [7] Ali, A. Y. (1934). *The Holy Qur'an: Translation and Commentary*. Lahore, Pakistan: Sh. Muhammad Ashraf.
- [8] Alkhateeb, J. H. (2020). A Machine Learning Approach for Recognizing the Holy Quran Reciter. *International Journal of Advanced Computer Science and Applications*, 11(7).
- [9] Al-Qurtūbī, M. B. A. (2003). *Tafsir Al-Qurtubi: Classical Commentary of the Holy Qur'an* (A. Bewley). Dar Al-Taqwa Ltd.
- [10] Al-Sukhni, E., Al-Kabi, M. N., & Alsmadi, I. M. (2016). An Automatic Evaluation for Online Machine Translation: Holy Quran Case Study. *International Journal of Advanced Computer Science and Applications*, 7(6), 118-123.
- [11] Andreyev, N. D. (1962). Linguistic Aspects of Translation. *Proceedings of the Ninth International Congress of Linguists*, Cambridge, 27-31.
- [12] Arberry, J. A. (1955). *The Koran Interpreted*. George Allan & Unwin LTD. London. The Macmillan Company. New York
- [13] 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).
- [14] Chai, C., Wang, J., Luo, Y., Niu, Z., & Li, G. (2022). Data Management for Machine Learning: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, 35(5), 4646-4667.
- [15] Font, J. E., & Costa-Jussa, M. R. (2019). Equalizing Gender Biases in Neural Machine Translation with Word Embeddings Techniques. arXiv preprint arXiv:1901.03116. <https://doi.org/10.48550/arXiv.1901.03116>
- [16] Forcada, L. M. (2018). Machine translation today. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of Translation Studies* (pp. 215-223). Benjamins Publishing Company.



- [17] Giammarresi, S., & Lalpalmé, G. (2016). Computer Science and Translation: Natural Languages and Machine Translation. In Y. Gambier & L. van Doorslaer (Eds.), *Border Crossings: Translation Studies and other Disciplines* (pp. 205-224). John Benjamins Publishing Company. <http://www.pbs.org/wgbh/globalconnections/mideast/questions/types/index.html>
- [18] Hamed, H., Helmy, A. M., & Mohammed, A. (2021, May). Deep Learning Approach for Translating Arabic holy Quran into Italian Language. *In 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)* (pp. 193-199). IEEE.
- [19] Irving, T. B. (1985). *The Qur'an: The First American Version; Translated and Commentary*. Quran Archive - Texts and Studies on the Quran. Retrieved from <http://quran-archive.org/explorer/thomas-ballantyne-irving/1985?page=1>
- [20] Kay, M. (1982). Machine Translation. *American Journal of Computational Linguistics*, 8(2), 74-78.
- [21] Khan, R. U., Qamar, A. M., & Hadwan, M. (2019). Quranic Reciter Recognition: A Machine Learning Approach. *Advances in Science, Technology and Engineering Systems Journal*, 4(6), 173-176.
- [22] Khanzode, K. C. A., & Sarode, R. D. (2020). Advantages and Disadvantages of Artificial Intelligence and Machine Learning: A Literature Review. *International Journal of Library & Information Science (IJLIS)*, 9(1), 3.
- [23] Larson, M. L. (1997). *Meaning-based Translation: A Guide to Cross-Language Equivalence*. University Press of America.
- [24] Lavie, A., Sagae, K., & Jayaraman, S. (2004). The Significance of Recall in Automatic Metrics for MT Evaluation. *In Machine Translation: From Real Users to Research: 6th Conference of the Association for Machine Translation in the Americas, AMTA 2004, Washington, DC, USA, September 28-October 2, 2004. Proceedings 6* (pp. 134-143). Springer Berlin Heidelberg.
- [25] Melby, A.K. (1996). *The Possibility of Language*. Warner John Benjamins Publishing Company.
- [26] Ménacère, K. M. (1999). Linguistic Acrobatics: Translating a Web of Cultural Connections. *Babel*, 45(4), 345-354.
- [27] OpenAI. (2023). ChatGPT (October 14 version) [Large language model]. <https://chat.openai.com/chat>
- [28] Oxford Learner's Dictionaries (n.d.). Oxford University Press. Retrieved May 7, 2023, from <https://www.oxfordlearnersdictionaries.com/>
- [29] 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).
- [30] Pickthall, M. M. W. (1930). *The Meaning of the Glorious Koran: An Explanatory Translation*. London, UK: George Allen & Unwin.
- [31] Qassem, M. (2021). Translations of stylistic Variations in the Noble Qur'an into English: Stylistic and Exegetical Perspectives. *Journal of Intercultural Communication Research*, 50(3), 304-322.
- [32] Saadany, H., & Orasan, C. (2021). BLEU, METEOR, BERT Score: Evaluation of Metrics Performance in Assessing Critical Translation Errors in Sentiment-Oriented Text. arXiv preprint arXiv:2109.14250.
- [33] Somers, H., Tsujii, J. I., & Jones, D. (1990). Machine Translation Without a Source Text. In COLING 1990 Volume 3: Papers presented to the 13th International Conference on Computational Linguistics.
- [34] Taecharunroj, V. (2023). "What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing*, 7(1), 35.
- [35] Tibawi, A.L (1962). Is The Qur'an Translatable? Early Muslim Opinion. Harvard University: Cambridge, Massachusetts. First read at the XXVth International Congress of Orientalists on Friday 12th August 1960, at the University of Moscow, and later published in *The Muslim World*, Volume 52, pages 1-16.
- [36] Vauquois, B. (1968). A Survey of Formal Grammars and Algorithms for Recognition and Transformation in Machine Translation, IFIP Congress-68 (Edinburgh), pp. 254-260.
- [37] Von, D. A. (1983). 'Ulum al-Qur'an: An Introduction to the Sciences of the Qur'an. The Islamic Foundation. UK.
- [38] Wilks, Y. (2008). *Machine Translation: Its Scope and Limits*. Springer Science & Business Media. Springer <https://archive.org/details/machinetranslati0000wilk>
- [39] World Population Review (2023). Muslim population by country 2023. World Population Review. <https://worldpopulationreview.com/country-rankings/muslim-population-by-country>
- [40] Yahaghi, M. J. (2002). An Introduction to Early Persian Qur'anic Translations. *Journal of Qur'anic Studies*, 4(2), 105-109.