Journal of Interdisciplinary Qur'anic Studies

Journal of Interdisciplinary Qur'anic Studies Vol.4, Issue 2, December 2025, 71-91

Sentiment and Tone Analysis of the Holy Qur'an Using Natural Language Processing

Mahdi Rafiei

Master's Degree in Computer Engineering, Department of Computer Engineering, Shahed University, Tehran, Iran.

Ehsan Khadangi [1]

Assistant Professor, Department of Computer Science, Shahed University, Tehran, Iran.

Article History: Received 1 March 2025; Accepted 16 May 2025

ABSTRACT: Original Paper

This article investigates sentiment and tone analysis of the verses of the Holy Qur'an using advanced Natural Language Processing (NLP) techniques. By leveraging deep learning transformer models such as AraBERT and MARBERT, distinct models were designed and implemented for sentiment analysis (positive, negative, and neutral) and multi-label tone analysis. These models aim to identify emotional and tonal patterns within the sacred text of the Qur'an. Evaluation results demonstrate satisfactory accuracy and F1 scores in detecting these patterns. Furthermore, an analysis of the Qur'an's text reveals a balanced distribution of tones across its chapters (surahs). This research underscores the potential of NLP as a powerful tool for analyzing complex and multifaceted religious texts, paving the way for future studies in this domain.

KEYWORDS: The Qur'an, Sentiment analysis, Tone analysis, Natural Language Processing (NLP), Deep learning, Transformers.

1. Corresponding Author. Email Address: khadangi@shahed.ac.ir

http://dx.doi.org/10.37264/JIQS.V4I2.5

1. Introduction

The Holy Qur'an, the divine scripture of Islam, serves not only as a source of spiritual and ethical guidance but also as one of the most prominent literary and religious texts in the world due to its unique linguistic and rhetorical features. Since its revelation, the Qur'an has captivated scholars, exegetes, and researchers from various disciplines due to its eloquence, rhetoric, and tonal diversity. One of its striking characteristics is the variety of emotional and affective tones in its verses, expressed through diverse tones such as mercy, admonition, fear, glad tidings, and others. This tonal variety plays a crucial role in conveying divine messages to audiences with diverse backgrounds and emotional states.

Tone in religious texts, particularly the Qur'an, refers to the emotional or affective quality conveyed by the text to its audience, shaped by the choice and arrangement of words. It plays a pivotal role in transmitting spiritual messages and emotionally impacting readers. Qur'anic verses employ various tones to evoke emotions such as hope, fear, joy, or caution, each leaving a distinct impression on the reader or listener (Eissa 2023). For instance, verses emphasizing divine mercy and forgiveness instill tranquility and hope, while those addressing punishment or warning evoke fear and a sense of responsibility. Analyzing these tones not only deepens the understanding of the verses' meanings but also has applications in fields such as religious education, Qur'anic exegesis, and religious psychology (Goel & Arsiwala 2024). The challenge of translating the Qur'an from Arabic to other languages underscores the importance of studying its tone (Yari & Firouziyan Pour Esfahani 2025). Islamic theology considers the Qur'an miraculous and inimitable, asserting that its text should not be detached from its original Arabic, as a single Arabic word can carry multiple meanings depending on context. The significance of tone in the Qur'an includes:

- Conveying Meaning and Emotion: Tone is vital for conveying the profound spiritual and religious meanings of the Qur'an. It helps readers grasp both the literal meaning and the implicit concepts, symbols, assumptions, and ideals embedded in the text (Hezarkhani & Ashrafi 2023).
- Moral and Didactic Impact: As an educational and exhortative text, the Qur'an uses tone to critique certain ideas and behaviors while encouraging readers to pursue ethical goals (Hezarkhani & Ashrafi 2023). The didactic tone, consistent throughout Qur'anic narratives, reflects God's intent to transform and elevate the reader (Eissa 2023).
- Engaging and Comprehending the Reader: Tone is a key factor in the narrative's appeal and its ability to connect with the audience. Proper tone comprehension is essential for readers to understand the text's structure and the relationships between its sections. It also serves as a critical clue for uncovering the theme and intent of a textual work (Eissa 2023). Using NLP, these emotional and tonal patterns can be

systematically and automatically identified.

The primary objective of this research is to design and implement two models for analyzing the sentiment and tone of Qur'anic verses using advanced NLP techniques. The first model focuses on identifying general sentiments (positive, negative, and neutral), while the second addresses multi-label tone analysis, such as mercy, reverence, command, and warning. By employing transformer-based models, preprocessing tools, and data augmentation techniques, this study aims to achieve high accuracy in detecting these patterns. Additionally, the sentiment and tone of the Qur'anic text is analyzed using the developed models.

The article is organized as follows: Section 2 reviews the research background on NLP and religious text analysis. Section 3 details the research methodology, including data, preprocessing, data augmentation, and model training processes. Sections 4 and 5 present the evaluation results of the models and the analysis of model output. Finally, Section 6 provides a discussion and conclusion of the findings.

2. Literature Review

The analysis of religious texts using NLP has emerged as a compelling topic in recent years, bridging computer science and religious studies (Bengueddach 2025). Comprehensive surveys on Arabic Sentiment Analysis (ASA) have focused on deep learning methods and highlighted challenges in processing complex Arabic texts like the Qur'an (Shi & Agrawal 2025). This section reviews studies related to Arabic text processing, deep learning concepts and transformer models, sentiment and tone analysis in religious texts, and the application of deep learning models in Arabic text analysis (Kusal et al. 2023).

2.1. Natural Language Processing in Qur'anic Text

The Arabic language, with its unique features such as diacritics, diverse forms of the letter *alif* (e.g., *alif maqṣūrah, alif mamdūdah, hamzah*), and complex grammatical and rhetorical structures, poses significant challenges for NLP (Abdul-Mageed et al. 2021). These characteristics render standard NLP models designed for languages like English less effective for Arabic text processing. To address these challenges, Arabic text must undergo preprocessing steps such as diacritic removal, normalization of *alif* forms, and elimination of unnecessary characters to prepare it for deep learning model training. Additionally, models and tools specifically fine-tuned for Arabic are required.

Recent studies have explored the structure of Qur'anic surahs (El-Affendi et al. 2025). One prominent theory is the *Thematic Unity Theory*, which posits that each surah revolves around a central theme. A key branch of this theory, the *Introduction and Exposition Theory*, suggests that God introduces the theme of each surah in its opening verses, elaborates on it through various methods such as stories, signs, comparisons, and predictions, and concludes in the final verses. Khadangi et al. (2022) investigated these theories using NLP techniques, calculating the similarity of Qur'anic roots using three methods: TF-IDF, Word2Vec, and root accompaniment. Their results confirmed that the studied surahs exhibit internal conceptual coherence, focusing on one or a few related themes. Additionally, comparing the similarity between the introduction and body of surahs validated the Introduction and Exposition Theory for many surahs. Furthermore, by analyzing the similarity between surahs relative to their order and revelation sequence, they concluded that the Qur'an's surah arrangement is relatively structured.

Computational text-mining algorithms have also been proposed for conceptualizing Qur'anic verses, providing a foundation for advanced tone and sentiment analysis (Azari et al. 2020).

2.2. Deep Learning and Transformer Models

Deep learning, a subset of machine learning, employs multi-layered artificial neural networks to model complex data patterns. It has made significant strides in recent years, particularly in NLP and computer vision, due to its ability to handle large and complex datasets. In NLP, earlier models such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs) were widely used, and in specific applications to religious texts like the Qur'an, models such as Support Vector Machines (SVMs) have been employed for multi-label classification of translated English verses, which can also be useful for tone analysis (Prabowo et al. 2019). However, the introduction of the transformer architecture in 2017 marked a significant breakthrough in NLP.

Transformers, introduced in the seminal paper *Attention Is All You Need* by Vaswani et al. (2017), revolutionized sequence data modeling. Unlike RNNs, which process data sequentially with limited parallelization, transformers enable fully parallel processing, significantly speeding up model training. The self-attention mechanism allows the model to directly compare all words in a sequence, effectively capturing their relationships. Transformers consist of two main components: an encoder and a decoder, each comprising multiple layers. The encoder consists of six identical layers, each with two sub-layers: a multi-head attention layer and a feed-forward neural network. The attention layer enables the model to assess the importance of different words in a sequence relative to a specific word. The decoder also consists of six layers but includes an additional sub-layer for multi-head attention over the encoder's output.

The introduction of transformers significantly improved performance in various NLP tasks, such as machine translation. Their simple yet highly efficient architecture has made them a cornerstone of modern deep learning models, with models like BERT, GPT, and T5 building directly on this framework. Models like MARBERT, optimized specifically for Arabic, leverage this architecture and demonstrate strong capabilities in understanding semantic and emotional relationships in Arabic texts (Abdul-Mageed et al. 2021).

Recent studies have applied transformer models to Qur'anic text analysis (Khadangi and Shabani 2023). Research indicates that Qur'anic surahs exhibit internal structure and organization, with each surah pursuing a specific purpose. While each surah focuses on a distinct theme, the Qur'an identifies 114 broad themes, and the notable similarity between adjacent surahs underscores their deliberate sequential arrangement. Khadangi and Shabani (2023) proposed a model comprising embedding and autoencoding components. The embedding phase used BERT to represent meaning and themes, while the autoencoding phase clustered data using soft labeling. Their findings revealed high semantic correlation between proximate surahs, decreasing with greater distance.

Alam et al. (2025) addressed the proliferation of Qur'anic content on social media, which poses challenges in verifying the authenticity of verses. They developed a method using transformer-based models (BERT-Base-Arabic, AraBERT, and MARBERT) trained on a dataset of authentic and fabricated verses. Among the models, MARBERT, designed for Arabic dialects, achieved the best performance with an F1-score of 94%.

2.3. Sentiment and Tone Analysis in Religious Texts and the Qur'an

Sentiment and tone analysis in religious texts like the Qur'an is challenging due to linguistic and semantic complexities (al-Ayyoub et al. 2019). Religious texts often feature composite tones, where a single verse may convey multiple emotions or tones (e.g., mercy and warning) simultaneously. This complexity renders traditional sentiment analysis methods, typically designed to identify a dominant emotion, less effective. Previous studies, such as Abu Farha and Magdy (2021), demonstrated that deep learning models can effectively identify emotional patterns in Arabic texts, but advanced methods are still needed for multi-tone analysis. For instance, Qur'anic verses addressing mercy and forgiveness may evoke hope and tranquility, while those concerning punishment induce fear and caution. This complexity necessitates models capable of predicting multiple labels simultaneously. Most studies in this area have focused on translations of the Qur'an (Gaanoun & Alsuhaibani 2025).

Islamic scholars have undertaken extensive efforts to translate the Qur'an into various languages, with English being a prominent target language. Eissa (2023) analyzed the tone of translated Qur'anic verses, finding that 63% exhibited a non-ironic tone. Among thematic components, there was a near balance between entity and concept in sentiment analysis, with 72% of textual elements showing unipolarity. Notably, verses with positive polarity outnumbered those with negative polarity. Thematic sentiment analysis revealed a prevalence of non-polar sentiments, followed by positive, negative, neutral, strong positive, and strong negative polarities.

This study approached sentiment analysis from two perspectives: global text-level analysis, which found 72% of verses to be unipolar and 1% ironic, and thematic sentiment analysis, which identified themes and varying levels of polarity. Entity-based sentiment analysis determines emotions toward specific entities (e.g., a person or group) within a text.

Karami et al. (2023) highlighted the exponential growth of textual data (books, blogs, and articles) and the time-consuming nature of manual analysis. They emphasized the importance of automated sentiment analysis for identifying writing styles and target audiences, particularly for the Qur'an as a divine and miraculous text. Their study developed a multi-label sentiment analysis model for English translations of the Qur'an using transformer-based models like RoBERTa and BiLSTM, achieving 77% accuracy. Their proposed model integrates RoBERTa's language understanding capabilities with syntactic features. It comprises two sub-models: one processes the raw text using RoBERTa, while the other incorporates dependency tags or parts of speech, processed via BiLSTM. The outputs are combined through a dense classification layer.

Sentiment analysis has also been applied to other religious texts. Vora et al. (2024) analyzed sentiments in selected Bible passages across five translations, noting variations in vocabulary and emotional tones such as humor, optimism, and compassion. They used a BERT-based model fine-tuned on the SenWave dataset for section-by-section sentiment analysis. The study used three English Bible translations: the King James Version (widely used), the Lamsa Version (translated from Aramaic), and the Simple English Version. The authors reviewed recurrent neural network models like LSTM and highlighted recent advancements in transformer models. They employed a BERT-based model for sentiment analysis, fine-tuned on the SenWave dataset.

3. Research Methodology

The methodology includes two distinct models for sentiment analysis (positive, negative, and neutral) and multi-label tone analysis. Both models follow similar steps, including text preprocessing, data augmentation, tokenization, and model training, but differ in label types and evaluation metrics.

3.1. Data

The dataset comprises approximately 1,000 Qur’anic verses, labeled in two ways with assistance from domain experts, prior studies such as (Hezarkhani & Ashrafi 2023), and large language models. The first labeling categorizes sentiment as positive (360 verses), negative (460 verses), or neutral (130 verses). The second labeling categorizes tone as “glad tidings and affectionate” (300 verses), “reverent and awe-inspiring” (365 verses), “commanding and authoritative” (370 verses), or “warning, reproachful, and fear-inducing” (410 verses). Some verses have multiple tones. Four tone labels were chosen to balance data volume and model performance. Approximately 100 verses were labeled by experts, while the rest were labeled using a majority-voting approach based on expert-labeled verses and outputs from large language models (GPT, Grok, Gemini, and Perplexity) (Zhang & Takada 2025).

3.2. Text Preprocessing and Vectorization

In NLP research, particularly for Arabic texts, preprocessing is critical for standardizing, cleaning, and preparing textual data to enhance machine learning model performance. Preprocessing steps include removing diacritics (e.g., *fatḥah*, *kasrah*, and *ḍammah*) to simplify text due to limited data volume, eliminating special characters like *hamzah* to reduce unnecessary variation, standardizing *alif* forms, removing numbers, and eliminating extra spaces. To address class imbalance, the neutral class was augmented using oversampling.

Text vectorization converts preprocessed text into numerical representations for machine learning models. This involves tokenization and transforming texts into numerical vectors. Using the MARBERT tokenizer, texts were split into tokens, with a maximum length of 128 tokens. Longer texts were truncated, and shorter texts were padded to ensure uniform input. To handle tokenization errors, a fallback mechanism tokenized empty text to produce valid vectors. Two types of vectors were generated: one mapping tokens to unique IDs in the BERT vocabulary and a binary mask distinguishing real tokens from padding.

Vectorization ensures text is machine-readable, with fixed-length inputs (128 tokens), and padding improving computational efficiency. MARBERT's tokenizer, optimized for Arabic, effectively handles the language's complex structures.

3.3. Proposed Model

This study employs deep learning with transformer-based models (AraBERT and MARBERT) for tone analysis, with a BiLSTM-based model for comparison. AraBERT and MARBERT, pre-trained for Arabic NLP, were selected for their ability to understand Arabic linguistic and semantic structures. The models were configured with three sentiment labels (positive, negative, and neutral) and four tone labels (glad tidings, reverent, commanding, and warning). Inputs include numerical vectors and attention masks from the vectorization stage. A comparison of the Transformer models used is shown in Table 1.

Class imbalance, particularly in the neutral class, was addressed by weighting the loss function. The neutral class weight was increased by 1.5 to balance attention to underrepresented classes. Label smoothing (0.1) was applied to improve generalization by reducing overconfidence in predictions.

Table 1. Comparison of AraBERT and MARBERT Models

Feature	AraBERT	MARBERT
Language Domain Focus	Modern Standard Arabic (MSA)	Arabic Dialect (DA) and MSA
Train data	61 GB (6.2 billion tokens)	128 GB (15.6 billion tokens)
Data Sources	Formal texts (books, news, web)	Arabic tweets
NSP Objective	Yes	None
Main Application	Formal and academic texts	Informal content and social media

3.3.1. Optimization Settings

AdamW optimizer: An improved version of Adam with a learning rate of 2e-5 and weight decay of 0.2 to prevent overfitting.

Learning rate schedule: A linear schedule with warm-up, gradually increasing the learning rate in the first 10% of training steps, then decreasing it linearly for stable convergence.

Dropout: A dropout rate of 0.3 in hidden and attention layers to reduce overfitting.

3.3.2. Early Stopping and Cross-Validation

Early stopping prevented overfitting, with a patience of 2 epochs and a minimum improvement delta of 0.01. The best model, based on weighted F1-score, was saved. Five-fold cross-validation was used, with 85% of data for training/validation and 15% for testing. Data were split into five subsets, with one subset used for validation and the others for training in each fold. The average performance across folds provided a robust estimate, with the best model selected for final evaluation.

3.3.3. Training Process

Models were trained with early stopping, using 70% of data for training and 15% for validation. In each epoch, data was processed in batches, with loss calculated using a weighted loss function and label smoothing. Gradients were updated, and the model was evaluated on the validation set for metrics like loss, accuracy, recall, and F1-score.

4. Model Evaluation

To evaluate each of the developed models, experiments were conducted on each model. The performance of each model was assessed using metrics such as accuracy, precision, recall, and F1-score to address one of the research questions concerning the models' effectiveness on the dataset created from Qur’anic verses. The evaluations revealed differences in the models' performance.

In training the models using the dataset of verses labeled as positive, negative, and neutral, the model developed with the BiLSTM algorithm achieved its best performance in epoch 7, reaching an accuracy of 93.15%, precision of 93.22%, recall of 93.15%, and an F1-score of 93.12%. The model developed using MARBERT achieved its best performance in epoch 4, with an accuracy of 97.96%, precision of 97.97%, recall of 97.96%, and an F1-score of 97.95%. Additionally, the F1-score for the positive label was 98% with a precision of 99%, for the negative label an F1-score of 98% with a precision of 99%, and for the neutral label an F1-score of 94% with a precision of 95%. The model developed using AraBERT achieved its best performance in epoch 5, with an accuracy of 94.9%, precision of 95%, recall of 94.9%, and an F1-score of 94.91%. The F1-score for the positive label was 93% with a precision of 92%, for the negative label an F1-score of 96% with a precision of 98%, and for the neutral label an F1-score of 95% with a precision of 92% (Figure 1).

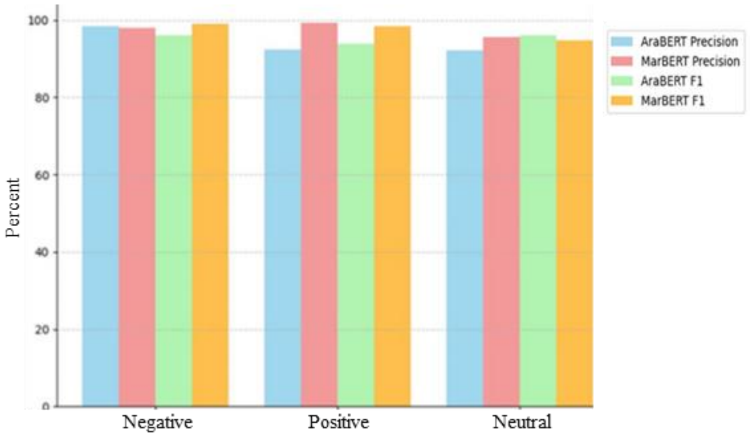

Figure 1. Comparison of the evaluation of two Transformer models on sentiment classes

Among the evaluated models, MARBERT outperformed the others across all evaluation metrics used, although the performance of the other models was also satisfactory and acceptable. The confusion matrices for the transformer models are presented in Figures 2 and 3, demonstrating their strong performance. Table 2 provides a summary of the performance of the evaluated models.

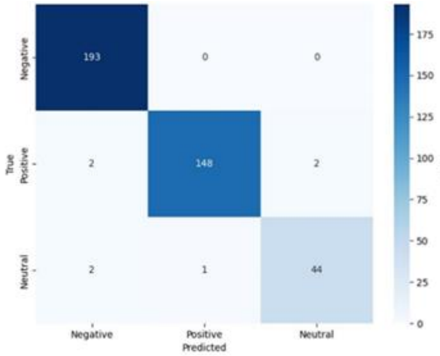

Figure 2. Confusion Matrix of the MARBERT Sentiment Model

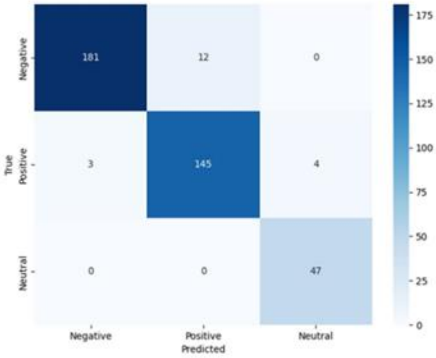

Figure 3. Confusion Matrix of the AraBERT Sentiment Model

Table 2. Sentiment Analysis Model Performance

Model	Accuracy	Precision	Recall	F1-Score
BiLSTM	93.15%	93.22%	93.15%	93.12%
MARBERT	97.96%	97.97%	97.96%	97.95%
AraBERT	94.9%	95%	94.9%	94.91%

We also employed probabilistic classification methods to determine the probability (percentage) of each verse and surah being positive or negative. For this purpose, machine learning algorithms including Logistic Regression (LR) and Support Vector Machines (SVMs) were used. The LR model achieved an accuracy of 86.46% and an F1-score of 84.69%, whereas the SVM model obtained an accuracy of 72.92% and an F1-score of 71.27%.

In training the models using the dataset of verses labeled with four tone categories, the model developed with the BiLSTM algorithm achieved its best performance in epoch 8, reaching an accuracy of 81.54%, precision of 92.72%, recall of 92.82%, and an F1-score of 92.73%. The model developed using MARBERT achieved its best performance in epoch 5, with an accuracy of 91.01%, precision of 94.99%, recall of 96.30%, and an F1-score of 95.58%. The model developed using AraBERT achieved its best performance in epoch 6, with an accuracy of 91.80%, precision of 95.98%, recall of 95.79%, and an F1-score of 95.87% (Table 3). Among the evaluated models, MARBERT and AraBERT demonstrated superior performance across the evaluation metrics used. However, the performance of the BiLSTM model was also satisfactory and acceptable. The performance of the transformer models, broken down by tone classes, is presented in Figure 4.

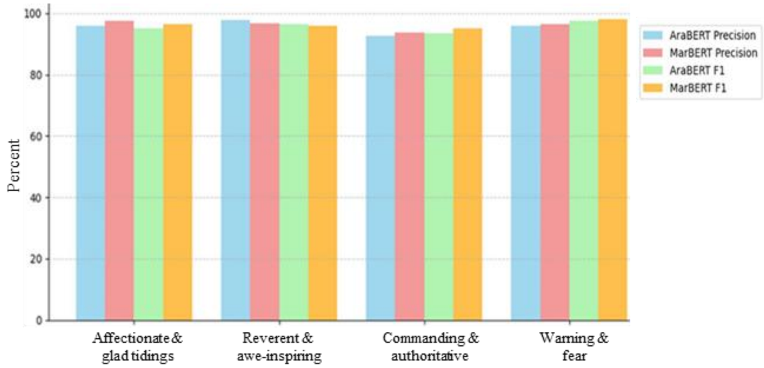

Figure 4. Comparison of the evaluation of two Transformer models on sentiment classes

The confusion matrices for the MARBERT and AraBERT models are shown in Figures 5 and 6.

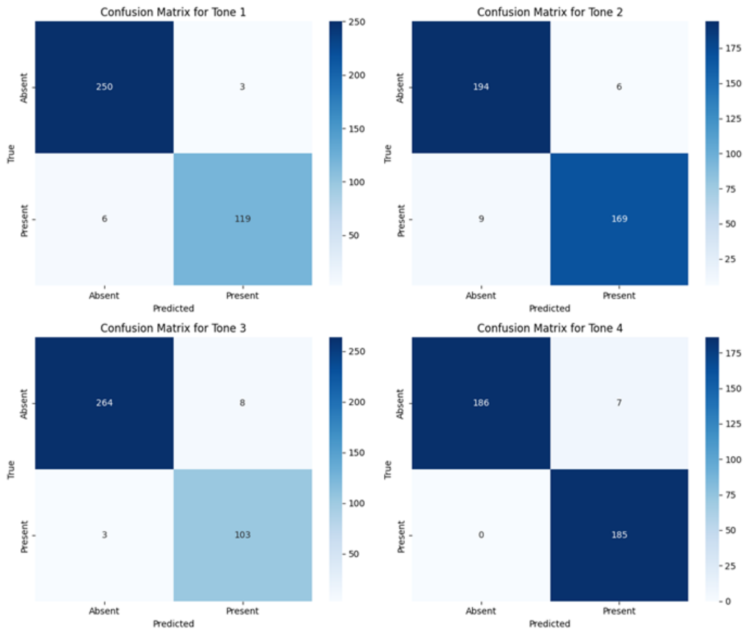

Figure 5. Confusion Matrix of the MARBERT Tone Model (Tone 1: Affection, Tone 2: Sanctification and Reverence, Tone 3: Authoritative and Directive, Tone 4: Warning and Reproach)

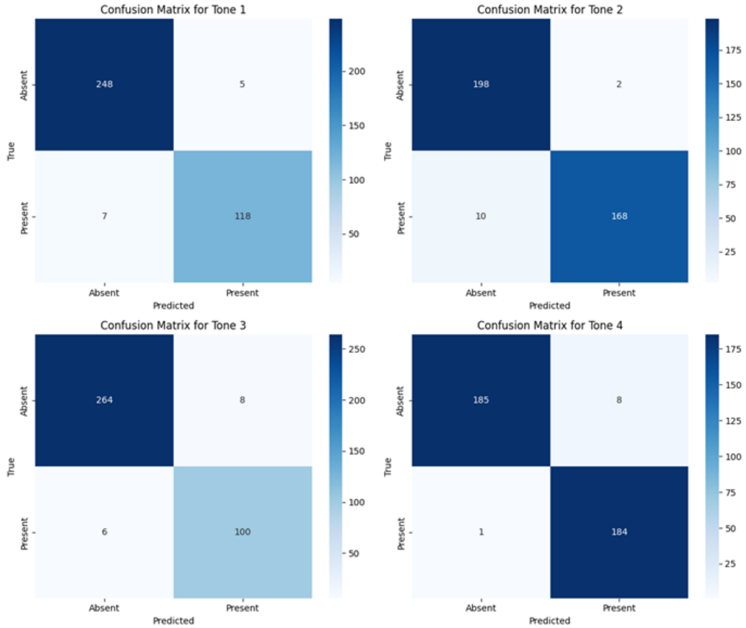

Figure 6. Confusion Matrix of the AraBERT Tone Model (Tone 1: Affection, Tone 2: Sanctification and Reverence, Tone 3: Authoritative and Directive, Tone 4: Warning and Reproach)

Table 3. Tone Analysis Model Performance

Model	Accuracy	Precision	Recall	F1-Score
BiLSTM	81.54%	92.72%	92.82%	92.73%
MARBERT	91.01%	94.99%	96.30%	95.58%
AraBERT	91.80%	95.98%	95.79%	95.87%

In conclusion, it can be inferred that the MARBERT models demonstrated superior performance in analyzing the sentiment and tone of the dataset created from the verses of the Qur'an, making them suitable as reference models for examining the sentiment and tone of the Qur'anic text.

Additionally, as observed in Table 4, the models presented in this article achieved a higher F1-score compared to other studies evaluated.

Table 4. Comparison of the Performance of Developed Transformer Models with Other Studies in the Domain of Qur'anic Text Analysis

Model	Accuracy	F1-Score
MARBERT (This Study)	91.01%	95.58%
AraBERT (This Study)	91.80%	95.87%
MARBERT (Alam et al. 2025)	93.73%	93.80%
AraBERT (Alam et al. 2025)	57.99%	70.00%
RoBERTa (Karami et al. 2023)	77.00%	-

5. Results and Discussion

The sentiment and tone analysis of the Qur'an was conducted using the MARBERT model. Analysis revealed that approximately 37.8% of the Qur'anic text is positive, 45.5% negative, and 16.7% neutral, as depicted in Figure 7.

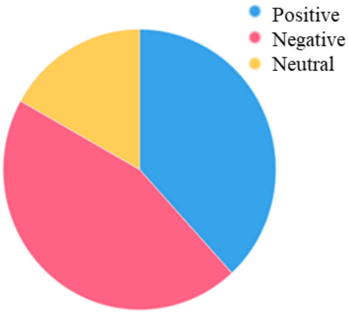

Figure 7. Sentiment Distribution across the Entire Qur'an

As shown in Figure 8, in many surahs, negative sentiments (represented in red) predominate, reaching over 80–90% in some cases. This indicates that a significant portion of the Qur'anic verses conveys warnings, admonitions, or serious messages regarding the consequences of human behavior. Despite the prevalence of negative sentiments, positive sentiments (represented in green) are present in nearly all surahs, reflecting the Qur'an's hopeful, encouraging, and inviting aspects, which consistently accompany its warnings. The neutral segment (represented in gray) constitutes 10–20% of most surahs, indicating verses that are informational or descriptive, lacking strong emotional undertones and focusing on conveying rulings, events, or descriptions. The initial surahs of the Qur'an, particularly longer ones such as al-Baqarah, Āl ʿImrān, and al-Nisāʾ, exhibit a more balanced mix of positive and negative sentiments. In contrast, shorter surahs towards the end of the Qur'an show a pronounced tendency toward either negative or positive sentiments, with significant fluctuations in their proportions. Surahs such as al-Ghāshiyah, al-Ḍuḥā, al-Ḥadīd, and al-Fatḥ are notable for their higher proportion of positive sentiments, emphasizing themes of divine mercy, blessings, and guidance. It can be concluded that, alongside its warnings and admonitions, the Qur'an consistently incorporates hopeful, affectionate, and inviting messages. Moreover, a significant portion of its verses lacks direct emotional weight, focusing instead on conveying rulings and narratives. This diversity and balance in tone and sentiment are among the prominent features of the Qur'anic text, underscoring the comprehensiveness of its divine message.

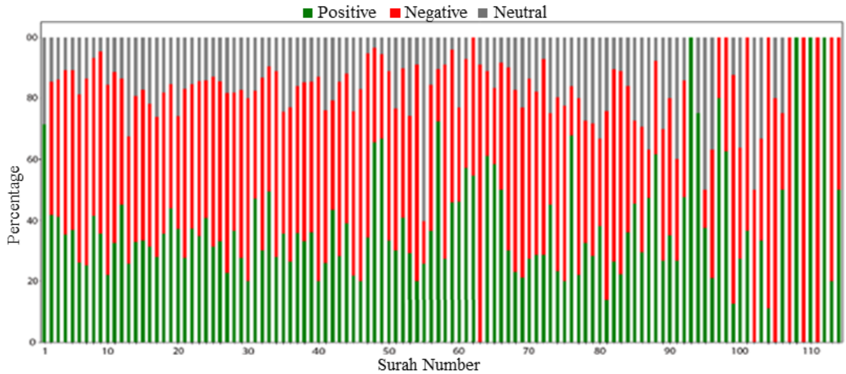

Figure 8. Sentiment Distribution across the Surahs of the Qur'an

In a surah-by-surah analysis, among the surahs of the Qur'an with more than 20 verses, the surahs al-Ḥadīd, al-Insān, al-Ghāshiyah, al-Fatḥ, and al-Burūj exhibit the highest tendency toward positive sentiment, respectively. Conversely, the surahs Yūnus, al-Qamar, Sabaʾ, and al-Aḥqāf show the highest tendency toward negative sentiment. In Figures 9 and 10, the surahs with the highest rates of positive and negative sentiment classes are displayed.

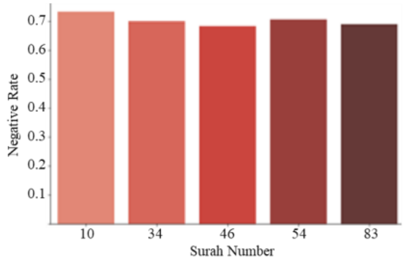

Figure 9. Surahs with the Highest Negative Sentiment

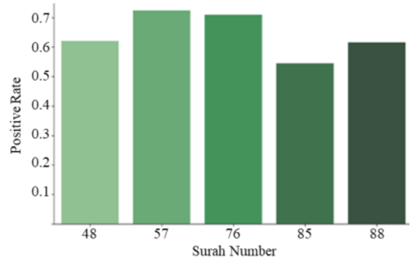

Figure 10. Surahs with the Highest Positive Sentiment

Figures 11, 12, 13, and 14 illustrate the probability percentages of each surah being positive or negative, which were computed using probabilistic classification based on the Logistic Regression (LR) and Support Vector Machine (SVM) algorithms.

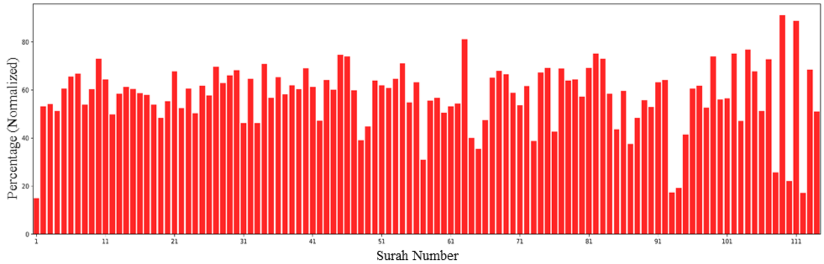

Figure 11. Normalized negative probability per surah (LR)

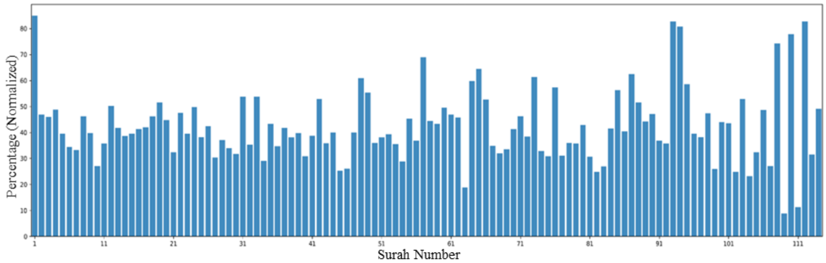

Figure 12. Normalized positive probability per surah (LR)

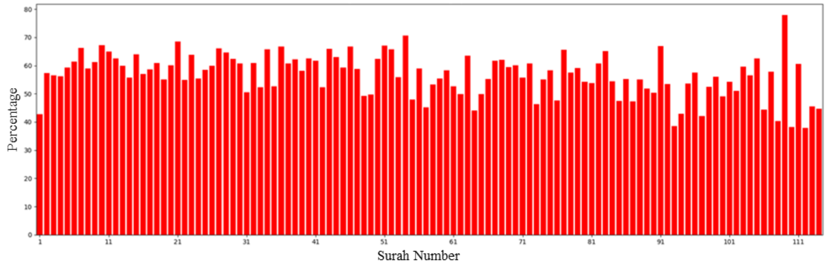

Figure 13. Negative probability per surah (SVM)

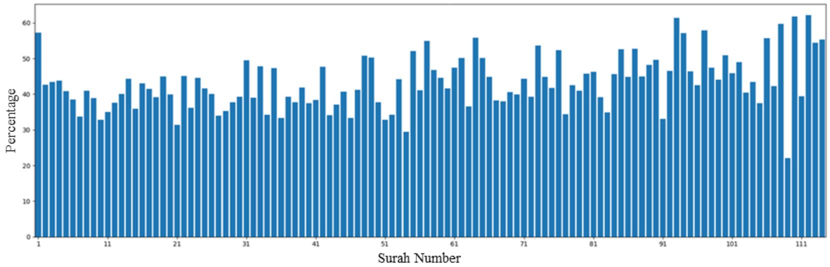

Figure 14. Positive probability per surah (SVM)

In the analysis of the tone of the Qur'an using the MARBERT model, which has the capability to classify Qur'anic text into four classes—"Good News and Affectionate," "Sanctification," "Authoritative and Directive," and "Warning, Reproachful, and Frightening"—approximately 19.8% of the verses are Good News and Affectionate, 23.6% of the verses have a Sanctification tone, 20.5% are Authoritative and Directive, and 36% of the verses exhibit a Warning, Reproachful, or Frightening tone. The chart of this distribution can be seen in Figure 15.

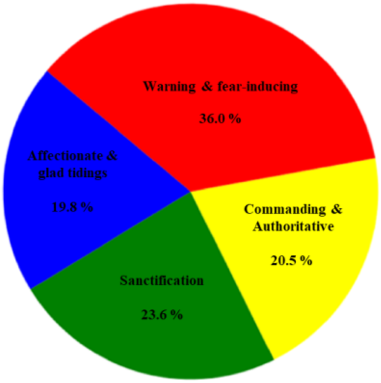

Figure 15. Tone Distribution in the Qur'an

Figure 16 illustrates the distribution of the four different tones across the 114 surahs of the Qur'an based on the tone analysis model discussed earlier. The horizontal axis represents the surah numbers from 1 to 114, while the vertical axis shows the contribution of each tone in each surah (ranging from 0 to 2, as each verse can exhibit more than one tone). According to this figure, the distribution of tones across the various surahs of the Qur'an is highly diverse, with no single tone uniformly dominating the entire text. This diversity reflects the rhetorical styles and varied themes of the Qur'an, shaped by the contexts and purposes of each surah. It should also be noted that some verses can exhibit more than one tone. In surahs like Al-Munāfiqūn, Al-Qamar, and Al- Qāriʿah, a predominant tone of warning and admonition is observed, while in surahs like al-Muzzammil and al-

Ghāshiyah, an affectionate tone is more prevalent. Surahs like al-Takwīr and al-Raḥmān feature a predominant tone of sanctification, reverence, and awe, whereas surahs like al-Tawbah and al-Mujādilah are characterized by a prevailing authoritative and directive tone. In Figures 17 to 20, the surahs with the highest rates of each tone are specified.

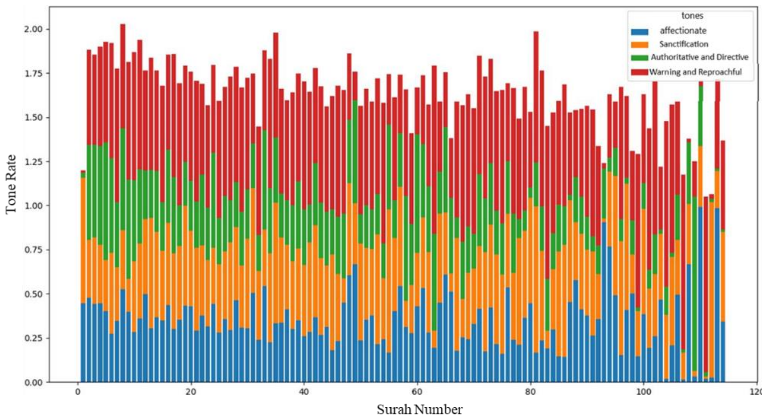

Figure 16. Tone Distribution across the Surahs of the Qur'an (Each Verse Can Have More Than One Tone)

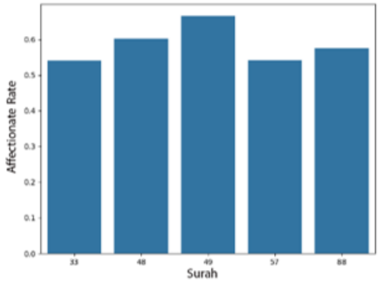

Figure 17. Surahs with the Highest Affectionate Tone

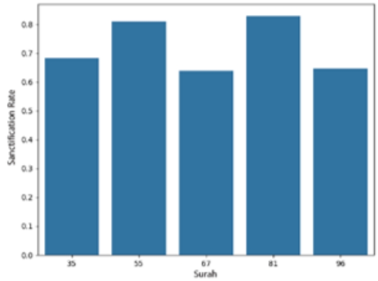

Figure 18. Surahs with the Highest Sanctification Tone

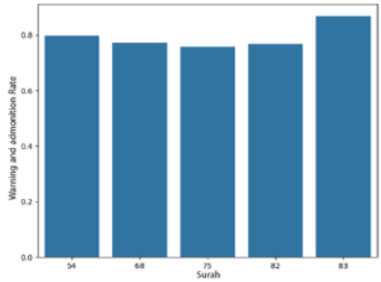

Figure 19. Surahs with the Highest Warning Tone

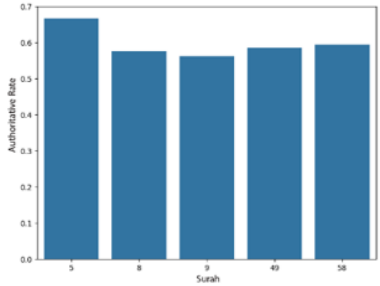

Figure 20. Surahs with the Highest Authoritative Tone

Table 5 displays a number of Qur'anic verses along with their English translations and the sentiment and tone detected by the model for each.

Table 5. Examples of Qur’anic Verses with Detected Sentiment and Tone

Ayah	Translation	Label
وَعَدَ اللَّهُ الَّذِينَ آمَنُوا وَعَمِلُوا الصَّالِحَاتِ لَهُمْ مَغْفِرَةٌ وَأَجْرٌ عَظِيمٌ (المائده/9)	God has promised forgiveness and a great reward to the righteously striving believers (Q. 5:9).	Positive/ Affection
وَعَدَ اللَّهُ الْمُؤْمِنِينَ وَالْمُؤْمِنَاتِ جَنَّاتٍ تَجْرِي مِنْ تَحْتِهَا الْأَنْهَارُ خَالِدِينَ فِيهَا وَمَسَاكِنَ طَيِّبَةً فِي جَنَّاتِ عَدْنٍ وَرِضْوَانٌ مِنَ اللَّهِ أَكْبَرُ ذَلِكَ هُوَ الْفَوْزُ الْعَظِيمُ (التوبه/72)	God has promised the believers gardens wherein streams flow and wherein they will live forever in the excellent mansions of the garden of Eden. What is more important than all this for them is that God is pleased with them. Such is the supreme triumph (Q. 9:72).	Positive/ Affection
وَالَّذِينَ كَفَرُوا وَكَذَّبُوا بِآيَاتِنَا أُولَئِكَ أَصْحَابُ الْجَحِيمِ (مائده 10)	However, the unbelievers who have called Our revelations lies will have hell for their dwelling (Q. 5:10).	Negative/ Warning
هُوَ اللَّهُ الَّذِي لَا إِلَهَ إِلَّا هُوَ الْمَلِكُ الْقُدُّوسُ السَّلَامُ الْمُؤْمِنُ الْمُهَيْمِنُ الْعَزِيزُ الْجَبَّارُ الْمُتَكَبِّرُ سُبْحَانَ اللَّهِ عَمَّا يُشْرِكُونَ (الحشر/23)	He is the only Lord, the King, the Holy, the Peace, the Forgiver, the Watchful Guardian, the Majestic, the Dominant, and the Exalted. God is too exalted to have any partner (Q. 59:23).	Neutral/ Sanctification
الَّذِي خَلَقَ سَبْعَ سَمَاوَاتٍ طِبَاقًا مَا تَرَى فِي خَلْقِ الرَّحْمَنِ مِنْ تَفَاوُتٍ فَارْجِعِ الْبَصَرَ هَلْ تَرَى مِنْ فُطُورٍ (الملک/3)	It is He who has created seven heavens, one above the other. You can see no flaw in the creation of the Beneficent God. Look again. Can you see faults? (Q. 67:3)	Neutral/ Sanctification
يَا أَيُّهَا الَّذِينَ آمَنُوا كُتِبَ عَلَيْكُمُ الصِّيَامُ كَمَا كُتِبَ عَلَى الَّذِينَ مِنْ قَبْلِكُمْ لَعَلَّكُمْ تَتَّقُونَ (البقرة/183)	Believers, fasting has been made mandatory for you as it was made mandatory for the people before you, so that you may have fear of God (Q. 2:183).	Neutral/ Authoritative
الزَّانِيَةُ وَالزَّانِي فَاجْلِدُوا كُلَّ وَاحِدٍ مِنْهُمَا مِائَةَ جَلْدَةٍ وَلَا تَأْخُذْكُمْ بِهِمَا رَأْفَةٌ فِي دِينِ اللَّهِ إِنْ كُنْتُمْ تُؤْمِنُونَ بِاللَّهِ وَالْيَوْمِ الْآخِرِ وَلْيَشْهَدْ عَذَابَهُمَا طَائِفَةٌ مِنَ الْمُؤْمِنِينَ (النور/2)	Flog the fornicatress and the fornicator with a hundred lashes each. Let there be no reluctance in enforcing the laws of God, if you have faith in God and the Day of Judgment. Let it take place in the presence of a group of believers (Q. 24:2).	Negative/ Authoritative/ Warning

As observed in Table 5, verse 72 of Surah al-Tawbah, in which Allah promises Paradise to the believers, has a tone of good news and affection. In contrast, the second verse of Surah al-Nūr, which addresses the punishment of sinners, carries a warning and frightening tone. Additionally, verses such as the third verse of Surah al-Mulk, which describe the attributes or power of Almighty God, exhibit a tone of sanctification, reverence, and awe. Verses such as verse 183 of Surah al-Baqarah, which issue commands to the believers, have an authoritative and directive tone.

6. Conclusion

This study examines the effectiveness of transformer-based deep learning models in analyzing the sentiment and tone of the text of the Qur’an. Three different models were implemented: a BiLSTM-based model, the MARBERT model, and the AraBERT model, all of which were fine-tuned to identify the sentiment and tone associated with each verse. Despite the inherent complexities of the Arabic language and the subtle differences among the verses, these models demonstrated remarkable performance. In particular, the MARBERT model, which was utilized to analyze the

Qur'anic text, stood out. Based on these analyses, we can conclude that transformer-based models can accurately detect the sentiments and tones associated with individual verses. The findings indicate that fine-tuned transformer models such as MARBERT and AraBERT have a stronger understanding of the meaning and structure of the text and can be effectively used to analyze the sentiment and tone of religious texts.

Acknowledgements

This paper has been extracted from the thesis prepared by Mahdi Rafiei under the supervision of Dr. Ehsan Khadangi in fulfilment of the requirements for the degree of Master of Science in Computer Engineering. The authors extend their gratitude to the department of computer engineering, Shahed University and the Research Deputy for supporting the research.

Declarations

Funding: No funding was received for conducting this study.
Conflict of Interest: The authors declare no competing interests.

References

Abdul-Mageed, M., Elmadany, A., & Nagoudi, E. M. B. (2021). ARBERT & MARBERT: Deep bidirectional transformers for Arabic. In *Proceedings of ACL 2021*. https://aclanthology.org/2021.acl-long.551

Abu Farha, I., & Magdy, W. (2021). A comparative study of effective approaches for Arabic sentiment analysis. *Information Processing & Management, 58*(2), Article 102438. https://doi.org/10.1016/j.ipm.2020.102438

Alam, K. A., Khalid, M., Ali, S. A., Mahmood, H., Shafi, Q., Haroon, M., & Haider, Z. (2025). Automated authentication of Qur'anic verses using BERT-based language models. In *Proceedings of the New Horizons in Computational Linguistics for Religious Texts* (59–66). Abu Dhabi, UAE. Association for Computational Linguistics.

Al-Ayyoub, M., Khamaiseh, A. A., Jararweh, Y., & Al-Kabi, M. N. (2019). A comprehensive survey of Arabic sentiment analysis. *Information Processing & Management, 56*(2), 320–342. https://doi.org/10.1016/j.ipm.2018.07.006

Azari, S., Shokrollahifar, M., & Lessani Fesharaki, M. A. (2020). Computational conceptualization of Qur'anic verses using text mining knowledge. *Zehn, 21*(82), 81–108.

Bengueddach, A. (2025). Multi-dimensional NLP analysis of the Qur'an: Structure, semantics, and emotion across qirāʾāt. *QURANICA: International Journal of Quranic Research,* 17(1). https://doi.org/10.22452/quranica.vol17no1.12

Eissa, M. E. (2023). Sentiment analysis of a translated text of the Qur'ān using the Pareto principle. *Islamic Studies,* 62(3), 379–402. https://doi.org/10.52541/isiri.v62i3.2706

El-Affendi, M. A., Al-Rajhi, L., Al-Khudhair, K., & Al-Harbi, H. (2025). NLP approaches to the Bible, Qur'an, and Bhagavad Gita. In *Proceedings of the First Workshop on Computational Linguistics for Religious Texts (CLRel 2025)* (1–10). https://aclanthology.org/2025.clrel-1.1

Gaanoun, K., & Alsuhaibani, M. (2025). Sentiment preservation in Qur'an translation with artificial intelligence approach: Study in reputable English translation of the Qur'an. *Humanities and Social Sciences Communications, 12*, Article 222. https://doi.org/10.1057/s41599-024-04181-0

Goel, P., & Arsiwala, R. (2024). A comparative study of religious scriptures using natural language processing. *Journal of Student Research*, 13(2). https://doi.org/10.47611/jsrhs.v13i2.6471

Hezarkhani, M., & Ashrafi, A. (2023). Analyzing the element of tone in the single-verse short short stories of the Qur'an. *Journal of Culture-Communication Studies*, *24*(63), 141–163. https://doi.org/10.22083/jccs.2023.388953.3720

Karami, M., Talebpour, A., Tajabadi, F. and Hajimohammadi, Z. (2023). Emotion Detection from the Text of the Qur'an Using Advance Roberta Deep Learning Net. *Journal of Interdisciplinary Qur'anic Studies*, 2(1), 147–165. https://doi.org/10.37264/jiqs.v2i1june2023.8

Khadangi, E., & Shabani, M. (2023). Study of the organization of the Qur'anic surahs using the similarity-based approach in deep learning. *Journal of Interdisciplinary Qur'anic Studies, 2*(2), 69–85. https://doi.org/10.37264/JIQS.V2I2December2023.4

Khadangi, E., Fazeli, M. M., & Naghavi, M. (2022). The study on Qur'anic surahs' structured-ness and their order organization using NLP techniques. *Journal of Interdisciplinary Qur'anic Studies, 1*(2), 29–56. https://doi.org/10.37264/jiqs.v1i2.3

Kusal, S., Patil, S., Choudrie, J., et al. (2023). A systematic review of applications of natural language processing and future challenges with special emphasis on text-based emotion detection. *Artificial Intelligence Review, 56*, 14215–14266. https://doi.org/10.1007/s10462-023-10439-0

Prabowo, S. A., et al. (2019). An implementation of support vector machine on the multi-label classification of English-translated Qur'anic verses. *Journal of Computer Science,* 15(12), 1752–1758. https://doi.org/10.3844/jcssp.2019.1752.1758

Sarwar, M. (1981). *The Holy Qur'an: Arabic text and English translation*. Englewood: The Islamic Seminary Inc.

Shi, Z., & Agrawal, R. (2025). A comprehensive survey of contemporary Arabic sentiment analysis: Methods, challenges, and future directions. In *Findings of the Association for Computational Linguistics: NAACL 2025* (3760–3772), Albuquerque, New Mexico. Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.findings-naacl.208

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *arXiv*. https://arxiv.org/abs/1706.03762

Vora, M., Blau, T., Kachhwal, V., Solo, G., & Chandra, R. (2024). Large language model for Bible sentiment analysis: Sermon on the Mount. *arXiv*. https://doi.org/10.48550/arxiv.2401.00689

Yari, A., & Firouziyan Pour Esfahani, A. (2025). The Interaction of Human Translators, Machine Translation, and Artificial Intelligence in Qur'anic Translation. *Pazhouhesh Dini*, 24(50), 71–94.

Zhang, Y., & Takada, S. (2025). Applying LLMs to active learning: Toward cost-efficient cross-task text classification without manually labeled data. *International Journal of Intelligent Systems*, Article 6472544, 1–14. https://doi.org/10.1155/ijis.6472544