Comparative Sentiment Analysis of Urdu Qur'anic Translations: A Machine Learning Approach

Nimra Nawaz*[1], Muhammad A. Awan[1]

***Abstract*—The Holy Quran holds profound religious and linguistic significance, guiding human civilization. Translating its emotional nuances into Urdu is challenging due to language structure, cultural context, and varied translation approaches. Accurate emotional tone analysis in Urdu translations is crucial to preserving the original text's theological and emotional meanings. This study introduces a new method to evaluate emotional tones in eight Urdu translations of Surah An-Naba and Surah An-Naazi'aat. Each verse was manually classified as positive, cautionary (negative), or neutral. Sentiment analysis techniques were applied, and machine learning models (SVM, Logistic Regression, Random Forest, MLP, and LSTM) were trained. Logistic Regression and SVM achieved the highest accuracy (78% and 72% respectively) on Maududi's translation. The analysis revealed significant variations in emotional expression across different Urdu translations. Comparing these findings with English translations of Surah An-Naba, Urdu versions showed distinct sentiment trends; Maududi's translation had a higher positive sentiment rate (48.8%), contrasting with stronger cautionary and neutral aspects in English translations. This research contributes to computational linguistics and Qur'anic studies by demonstrating the value of Natural Language Processing (NLP) tools in maintaining the religious, emotional, and linguistic accuracy of Qur'anic translations. Future work will expand the dataset and test additional translations. The methodology can later be implemented in other fields that are more technical.**

Index Terms*— *Comparative Analysis, Machine Learning, Natural Language Processing (NLP), Sentiment Analysis, Text analysis, Urdu Qur'anic Translations.

I. Introduction

The Holy Quran is considered the final word in divine guidance, having profound theological and linguistic depth. It is considered sacred not only for its spiritual teachings but also for its rhetorical flair and linguistic intricacy. Translating the Quran from Arabic into other languages, like Urdu, is a major challenge.

Manuscript received on 12 March 2025.
[1] Department of Computer Science, Allama Iqbal Open University.
*Corresponding author: nimranawaz847@gmail.com

Such challenges are caused by structural variations between languages, cultural differences, and the interpretative aspect of translation, which may influence the emotional tone communicated in the translated text. Considering the significance of the maintenance of both the theological and emotional purpose of the Qur'anic verses, it becomes mandatory to examine how various translations portray these emotional tones. The emotional tone of Qur'anic verses is crucial in shaping a reader's understanding and perception of the message. The Qur'anic text contains neutral, positive, and cautionary thoughts that should be accurately translated to preserve the intended meaning.

In this study, a Machine Learning Approach, the emotional expressions in Surah An-Naba and Surah An-Naazi'aat are analyzed and contrasted. The emotional tone of verses in eight well-known Urdu translations by Maulana Maududi, Ahmed Ali, Muhammad Hussain Najafi, Ahmed Raza Khan, Syed Zeeshan Haider Jawadi, Tahir ul Qadri, Muhammad Junagarhi, and Fateh Muhammad Jalandhry is examined using sentiment analysis techniques.

Natural Language Processing (NLP) techniques provide a systematic approach to analyzing sentiment in textual data. This research combines machine learning and deep learning models, including Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), Multilayer Perceptron (MLP), and Long Short-Term Memory (LSTM), to classify each verse's emotional tone into three categories: positive, cautionary (negative), and neutral. Sentiment annotation was conducted manually before training the models to ensure accurate classification. The comparative analysis examines how different translators convey emotions and how translation approaches influence sentiment representation.

Furthermore, this study extends its findings by comparing the results with [1], who explored the emotional tone of Surah An-Naba using English translations. The comparison highlights differences between Urdu and English translations, revealing distinct sentiment trends. Notably, Maududi's Urdu translation exhibits a higher proportion of positive sentiment (48.8%), while English translations such as Saheeh International and Pickthall emphasize cautionary

(58%) and neutral (27%) sentiments. These findings underscore the need for computational tools to enhance the accuracy and consistency of Qur'anic translations while preserving theological meaning and emotional depth.

This research contributes to the field of computational linguistics and Qur'anic studies by developing a sentiment analysis framework tailored to Qur'anic translations. The research intends to improve knowledge regarding translation differences and their effects on sentiment detection. The proposed framework requires expansion by including new Surahs, plus evaluating various translation methods to make it more precise in future research directions.

II. Objective of the Study

The objectives of this study are to identify and classify the emotional tone of verses in Surah An-Naba and Surah An-Naazi'aat across various Urdu translations using sentiment analysis techniques. The study examines the matching and unique distribution patterns of sentiment across these translations without ignoring the outcome regarding emotional interpretation. This research assesses Urdu translation sentiment classification through a comparison with the results obtained using ten English translations of Surah An-Naba for detecting cross-language emotional variations [1].

III. Background and Novelty Statement

The Holy Quran, as the ultimate source of divine guidance for Muslims, is a religious scripture and a text of immense linguistic, philosophical, and moral significance. The sacred verses use amplifying language to portray aspects of faith and morality, together with divine justice and human accountability, and moral responsibility. Translating Qur'anic content requires complete preservation of the original Arabic emotional content. In Qur'anic verses, the emotional expression serves a purpose to trigger particular spiritual and moral, and intellectual reactions from readers. Different language translations of the Quran face specific problems due to the act of interpreting language while adapting to various cultural backgrounds.

Urdu, spoken by approximately 160 to 180 million people worldwide, including around 17 million native speakers in Pakistan and over 50 million in India, with significantly more using it as a second language, is a widely spoken language in the Muslim world that features various translations of the Quran, each demonstrating subtle differences in language choice and interpretive nuance.

This study introduces a novel comparative analysis of emotional tone in multiple Urdu translations of the Quran, focusing on Surah An-Naba and Surah An-Naazi'aat. This investigation deviates from previous studies, which studied English translations of sentiment by applying machine learning and deep learning analysis to Urdu translations while exploring interpretive and linguistic differences. This research generates exceptional cross-linguistic insights about Qur'anic interpretation through its result comparison with Refaee's study, which examined English translations [1].

IV. Literature Review

The literature on sentiment analysis in religious texts has gained important traction in recent years. Researchers have explored various methodologies for analyzing the emotional tone of sacred scriptures, including the Bible and the Quran, to understand how linguistic and cultural factors affect interpretation. The following subsection contains essential findings from studies that examine sentiment analysis together with Qur'anic translations and Natural Language Processing applications for religious and Arabic texts. Sentiment analysis is a field of study that examines sentiment in various contexts, including the Quran, Twitter, Arabic dialects, and social media chats. It has been applied in various languages, including Arabic and Emirati dialects, and has been used in various scenarios, like the COVID-19 pandemic.

A. Sentiment Analysis in Arabic Text

Sentiment analysis is a widely adopted methodology in NLP to assess the emotional tone of a text. In recent years, scholars have utilized sentiment analysis to analyze religious texts to investigate their emotional and linguistic influence. For instance, Salloum and his research members provide a broad review of several studies related to Arabic text mining, focusing on the Holy Quran sentiment analysis and web documents, highlighting the importance and challenges of text mining in extracting information from unstructured text [2]. Karami and his team focus on enhancing emotion detection models, particularly in the context of the Qur'an, by including syntactic features to improve accuracy and generality [3]. This investigation showcases practical applications in pastoral care and counseling within Islamic communities and challenges common perceptions by revealing joy and happiness as predominant emotions in the Qur'anic verses. Alqarni & Rahman focus on Arabic sentiment analysis related to the COVID-19 pandemic in Saudi Arabia [4]. They collected datasets from Riyadh, Dammam, and Jeddah in two different periods and compared the proposed technique with state-of-the-art techniques. Elhassan and his team explore the effectiveness of various word embeddings and data set setups in Arabic sentiment analysis using deep learning models and achieve high accuracy on the HARD dataset

with fastText word embedding [5]. Vora, with his team of researchers, focuses on sentiment analysis of selected chapters of the Bible, specifically the Sermon on the Mount, using different translations to explore the varying sentiments expressed [6].

B. Sentiment Analysis and Qur'anic Translation

The application of sentiment analysis specifically to the Quran has been explored by several scholars. Refaee explores sentiment analysis of one chapter of the Quran using 10 English translations. He reveals differences between translations in conveying negative, positive, and neutral sentiments, highlighting the importance of accurate translations for understanding the Quran, with suggestions for improving translations and further exploring Qur'anic sentiment [1]. Our research builds upon Refaee's work by taking the analysis further to include Urdu translations to make a comparative assessment of sentiment differences in different linguistic contexts.

C. Challenges in Sentiment Analysis of Urdu Texts

Urdu sentiment analysis faces several challenges due to the richness of the language in terms of morphology (different forms), changes in words, and script-related issues. Sial and his team of researchers proposed a framework for sentiment analysis of Urdu news articles, mentioning the challenges involved in correct sentiment classification owing to the scarcity of linguistic resources [7]. To overcome these issues, our research utilizes both machine learning algorithms (e.g., Support Vector Machine, Logistic Regression, and Random Forest) as well as deep learning models (e.g., Neural Networks, Multilayer Perceptron, and LSTM) to examine sentiment in Qur'anic Urdu translations.

D. Machine Learning and Deep Learning Models in Sentiment Analysis

Traditional machine learning models such as Support Vector Machine (SVM), Logistic Regression (LR), and Random Forest (RF) have been widely used in sentiment analysis due to their effectiveness in text classification tasks. More recently, deep learning approaches such as Long Short-Term Memory (LSTM) and Multilayer Perceptron (MLP) models have shown promising results in sentiment classification of complex textual data. Study by Zhou and Zhu have also demonstrated that deep learning models outperform traditional classifiers in handling nuanced sentiment variations [11].

Duwairi & Qarqaz discuss sentiment analysis in Arabic text using machine learning classifiers on a dataset of tweets and comments, highlighting the importance of sentiment analysis and the need for larger datasets and semi-supervised learning [8]. Both machine learning and deep learning methods are used in this research to compare the performance of various models in analyzing Urdu Qur'anic translations.

V. Significance and Research Contribution

While previous research has primarily focused on sentiment analysis in English translations of the Quran, this study fills a crucial gap by applying sentiment analysis to multiple Urdu translations of Surah An-Naba and Surah An-Naazi'aat. By comparing sentiment classification results across different translators, this research provides valuable insights into how translation choices influence emotional tone. Additionally, by benchmarking our findings against Refaee's study on English translations, we contribute to a broader cross-linguistic understanding of Qur'anic sentiment analysis [1].

VI. Materials and Methods

The source of data for this study is the Urdu translation of Surah An-Naba and An-Naazi'aat. This study uses eight different Urdu translations of the Quran from renowned scholars, including Maulana Maududi, Ahmed Ali, Muhammad Hussain Najafi, Ahmed Raza Khan, Syed Zeeshan Haider Jawadi, Tahir ul Qadri, Muhammad Junagarhi, and Fateh Muhammad Jalandhry. The translations were extracted from https://tanzil.net/ and processed for sentiment analysis. Manual Sentiment annotation was performed by carefully reading each verse and categorizing verses into positive, cautionary (negative), and neutral sentiments. Moreover, machine learning and deep learning models, including SVM, LR, RF, MLP, and LSTM, were utilized to examine and categorize the emotional tone in such translations.

The study employs a sentiment analysis approach to classify the emotional tone of verses in Surah An-Naba and Surah An-Naazi'aat across eight Urdu translations. The methodology involves data collection from Tanzil.net, followed by text preprocessing to remove noise (normalization), stopword removal, and tokenization to normalize the content. Manual sentiment annotation was conducted to label each verse as positive, cautionary (negative), or neutral sentiments. SVM, LR, RF, MLP, and LSTM models of machine learning and deep learning were trained over the annotated data to examine the sentiment differences across various translations. The models were evaluated using accuracy, precision, recall, and F1-score to determine and carefully assess the performance of the models in terms of sentiment classification.

VII. Methodology

An analysis of text sentiment has been used to classify

the emotional content of verses in Surah An-Naba and Surah An-Naazi'aat from eight different Urdu versions. The methodology consists of a structured pipeline including data collection, text preprocessing, sentiment annotation, feature extraction, model selection & training, and evaluation and analysis.

After completing all the steps, a comparison (comparative analysis) was conducted across the eight Urdu translations. Moreover, a comparative analysis was performed among these Urdu translations and the English translations by Refaee to examine cross-linguistic emotional tone differences. The overall steps followed in this study are illustrated in the flow diagram, as shown in Fig 1.

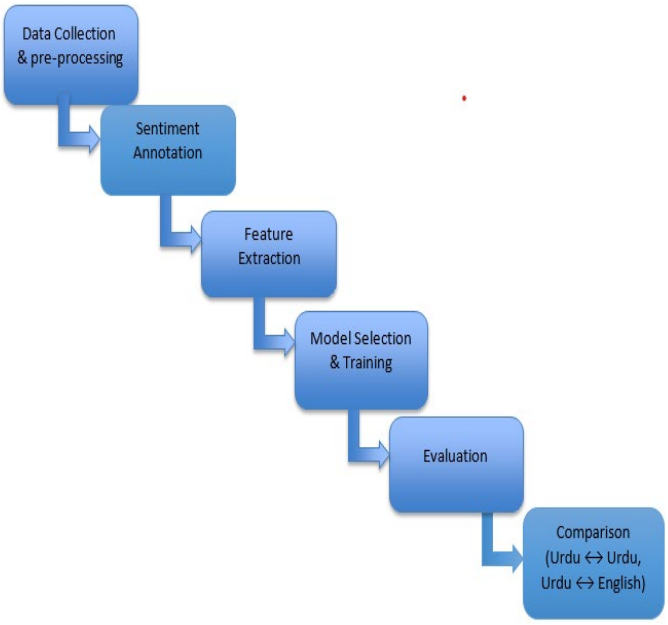


Fig 1: Flow of Study

A. *Data Collection*

The translations were extracted from Tanzil.net, a reliable source for Qur'anic translations in multiple languages, including Urdu. The dataset was prepared to facilitate the emotional tone analysis task by annotating the verses of Surah An-Naba and Surah An-Naazi'aat with sentiment labels. The dataset includes translations by eight different scholars, certifying linguistic diversity and interpretational tones. The data collected was stored in Excel and text format for further processing.

B. *Text Preprocessing*

To prepare the dataset for sentiment analysis, the process in the preprocessing involved the following steps:

- The noise removal step eliminates punctuation, special characters, numbers, and extra spaces using regular expressions (RegEx).
- In stopword removal, meaningless texts, like all common words that do not contribute to sentiment classification, were thrown away.
- Tokenization segments the text into meaningful units while preserving verse integrity.

In this study, verse-level tokenization was performed, in which each verse translation was converted into a single token. To handle the grammatical complexity of the Urdu language, these actions were required.

C. *Sentiment Annotation*

Qur'anic scholars and native Urdu speakers carefully annotated the datasets to categorize the verse translations into three emotional categories: positive, cautionary (negative), and neutral. Due to the religious and sensitive content of the Qur'anic text, sentiment annotation was manually done by domain-literate human annotators or experts. Instead of inter-annotator agreement measures like Cohen's Kappa, which has the potential to simplify the theological complexity of the verses, we depended on expert consensus to provide loyal and respectful sentiment labeling. This helped maintain the interpretive integrity of the Qur'anic content concerning various Urdu translations.

Positive sentiments contain verses that express hope, reward, or reassuring tones. Cautionary (Negative) sentiments contain verses that highlight warnings, punishments, or threats. Neutral sentiments are assigned to verses that emphasize the oneness of Allah, His absolute authority, and the dependence of all creations on Him. It also includes those verses in this category that do not exhibit strong emotions towards positive or cautionary sentiments. Labels 1, 0, and 2 are used for labeling. 1 is used for positive, 0 is for cautionary, and 2 is for neutral sentiments. To enable reproducibility, all the annotated datasets employed in this research, such as sentiment labels for every verse in various eight Urdu translations, are made publicly available. Researchers and readers can access them through the following link:

https://drive.google.com/drive/folders/1MnLJLWDtRvB9LIPzKmiJKiW3yzh-kceR?usp=sharing

A screenshot of the annotated dataset used in this study is also provided in Fig 2 for reference. During the annotation process, several challenges were encountered, particularly related to class imbalance.

In Class Imbalance, model performance was impacted by the neutral category's dominance, necessitating careful treatment throughout training. The Synthetic Minority Over-Sampling Technique (SMOTE) was used to address

class imbalance, specifically for the under-represented Neutral class.

Surah	AyahNumber	AyahText	Label
78	1	کس چیز کی بابت وہ آپس میں سوال کرتے ہیں	2
78	2	اس بڑی خبر کے متعلق	2
78	3	جس میں وہ اختلاف کر رہے ہیں	2
78	4	ہرگز ایسا نہیں عنقریب وہ جان لیں گے	2
78	5	پھر ہرگز ایسا نہیں عنقریب وہ جان لیں گے	2
78	6	کیا ہم نے زمین کو فرش نہیں بنایا	1
78	7	اور پہاڑوں کو میخیں	1
78	8	اور ہم نے تمہیں جوڑے جوڑے پیدا کیا	1
78	9	اور تمہاری نیند کو راحت کا باعث بنایا	1
78	10	اور رات کو پردہ پوش بنایا	1
78	11	اور دن کو روزی کمانے کے لیے بنایا	1
78	12	اور ہم نے تمہارے اوپر سات سخت (آسمان) بنائے	1
78	13	اور ایک جگمگاتا ہوا چراغ بنایا	1
78	14	اور ہم نے بادلوں سے زور کا پانی اتارا	1
78	15	تاکہ ہم اس سے اناج اور گھاس اگائیں	1
78	16	اور گھنے باغ اگائیں	1
78	17	بے شک فیصلہ کا دن معین ہو چکا ہے	2
78	18	جس دن صور میں پھونکا جائے گا پھر تم گروہ درگروہ چلے آؤ گے	0
78	19	اور آسمان کھولا جائے گا تو (اس میں) دروازے ہوجائیں گے	0
78	20	اور پہاڑ اڑائے جائیں گے تو ریت ہو جائیں گے	0
78	21	بے شک دوزخ گھات میں لگی ہے	0

Fig 2: Dataset by Ahmed Ali

D. *Feature Extraction*

Two methods were used in the study for extracting features. Term Frequency-Inverse Document Frequency (TF-IDF) analysis is used to convert text into numerical features. The Scikit-learn TfidfVectorizer, which is used in the TF-IDF approach, performed well with SVM and other applicable classifiers like Random Forest, MLP, and Logistic Regression.

The second method used word embedding averaging to generate sentence vectors using FastText Word Embedding with cc.ur.300.bin. This method was implemented by Gensim, a powerful Python package for natural language processing and unsupervised topic modelling that offers effective tools for loading and manipulating pre-trained word embeddings. To represent text with improved contextual comprehension, the resulting vectors were subsequently fed into LSTM and neural networks.

E. *Deep Learning and Machine Learning Models*

For sentiment classification, both ML and DL models are used. ML models (Logistic Regression, Support Vector Machine, and Random Forest) require less training time and are successful in text classification when combined with the feature extraction technique of TF-IDF. On the other hand, Deep learning models (Multi-Layer Perceptron and Long Short-Term Memory) capture deep contextual relationships in text data.

F. *Training Process*

For training models, fast text word embedding was used as an input feature. Because it is already trained in large text corpora and does not need additional training. Instead of training from scratch, load the pretrained. FastText Urdu model (cc.ur.300.bin) and use it to extract word vectors. There is also a division of the dataset into training and testing sets of the data. Hyperparameter tuning, grid search, and other optimization techniques were used to enhance model performance. Training is conducted until loss stabilization to ensure meaningful learning of sentiment patterns.

G. *Evaluation Metrics*

The model's performance was assessed using accuracy, precision, recall, and F1 score. Label verses from both Surahs (Surah An-Naba and Surah An-Naazi'aat) made up the test dataset. Three sentiment labels, positive, cautionary (negative), and neutral, were used to assess the models.

VIII. Class Imbalance Issue

The dataset significantly affects the neutral category due to class imbalance because the neutral category has fewer verses compared to positive and cautionary verses. This class imbalance negatively affects the performance of the neutral category. For handling this issue, SMOTE (Synthetic Minority Over-sampling Technique) is used to balance the dataset. Additionally, some verses which fall into two categories at the same time, such as being both Neutral and Positive or Neutral and Cautionary, are a challenge for the models.

IX. Tools, Libraries, and Programming Environment

Python was used as the main programming language because of its extensive support for deep learning and machine learning libraries. This study was developed and operated mostly using Google Colab (https://colab.research.google.com/). Google Colab is widely utilized in the model's testing, training, and assessment stages because of its user-friendly interface and free computing resources for experimenting and visualizing various configurations.

Key libraries used in this study include Pandas, NumPy, and Regex for preprocessing and data handling; Scikit-learn for TF-IDF feature extraction and implementing machine learning models; and Gensim for utilizing FastText embeddings. Deep learning models were implemented using TensorFlow, Keras, and PyTorch. To address class imbalance, SMOTE (Synthetic Minority Over-sampling Technique) was applied. For visualization and evaluation of results, Matplotlib, Seaborn, and Scikit-learn's metrics module were employed.

X. Computational Resources

The development of ML and DL models occurred through a process using Google Colab alongside Python programming language capabilities. The free GPU/TPU support from Colab substantially shortened computational tasks' execution time in a way that improved processing efficiency. The preinstalled libraries, along with Google Drive integration in this platform, allowed them to work with their data smoothly.

XI. Results and Analysis

The classification report, summarizing the performance of all models across the eight Urdu translations of the Surah An-Naba and An - An-Naazi'aat, is presented in the TABLE I to TABLE IVand Fig 3 to Fig 6.

A. Models Performance on Individual Translations

1) Maulana Abul A'ala Maududi Translation

TABLE I
Accuracy, Precision, Recall, and F1-Score for Each Model in the Maududi Dataset

Model	Accuracy	Precision	Recall	F1- Score
LR (Logistic Regression)	0.78	0.79	0.78	0.77
SVM (Support Vector Machine)	0.78	0.79	0.78	0.77
RF (Random Forest)	0.61	0.61	0.61	0.61
MLP (Multi-Layer Perceptron)	0.78	0.75	0.78	0.76
LSTM (Long Short-Term Memory)	0.66	0.64	0.65	0.64

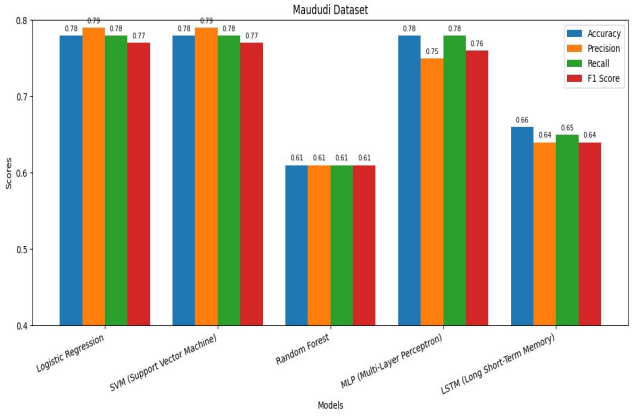


Fig 3: Models' Performance on the Maududi Dataset

2) Ahmed Ali Translation

TABLE II
Accuracy, Precision, Recall, and F1-Score for Each Model in the Ali Dataset

Model	Accuracy	Precision	Recall	F1- Score
LR (Logistic Regression)	0.72	0.50	0.54	0.51
SVM (Support Vector Machine)	0.67	0.45	0.50	0.47
RF (Random Forest)	0.61	0.42	0.46	0.43
MLP (Multi-Layer Perceptron)	0.67	0.52	0.50	0.46
LSTM (Long Short-Term Memory)	0.72	0.50	0.54	0.51

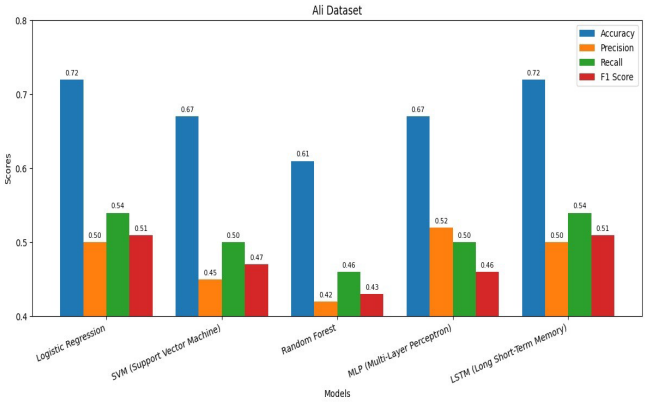


Fig 4: Models' Performance on the Ali Dataset

3) Fateh Muhammad Jalandhry Translation

TABLE III
Accuracy, Precision, Recall, and F1-Score for Each Model in the Jalandhry Dataset

Model	Accuracy	Precision	Recall	F1- Score
LR (Logistic Regression)	0.56	0.57	0.56	0.55
SVM (Support Vector Machine)	0.56	0.61	0.56	0.55
RF (Random Forest)	0.61	0.68	0.61	0.59
MLP (Multi-Layer Perceptron)	0.56	0.56	0.56	0.56
LSTM (Long Short-Term Memory)	0.56	0.56	0.56	0.56

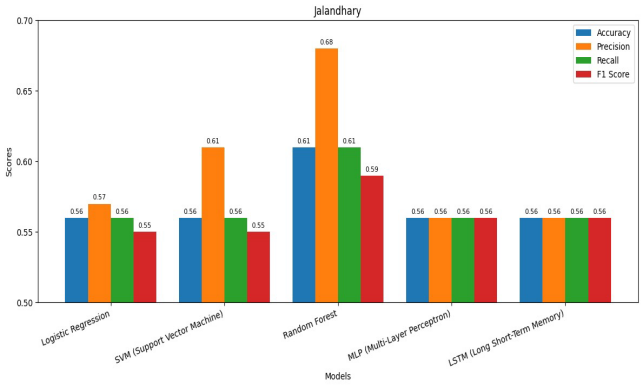


Fig 5: Models' Performance on the Jalandhry Dataset

4) Tahir ul Qadri Translation

TABLE IV
Accuracy, Precision, Recall, and F1-Score for Each Model in the Qadri Dataset

Model	Accuracy	Precision	Recall	F1- Score
LR (Logistic Regression)	0.67	0.70	0.67	0.62
SVM (Support Vector Machine)	0.50	0.63	0.50	0.54
RF (Random Forest)	0.61	0.62	0.61	0.61
MLP (Multi-Layer Perceptron)	0.61	0.62	0.61	0.61
LSTM (Long Short-Term Memory)	0.22	0.25	0.23	0.22

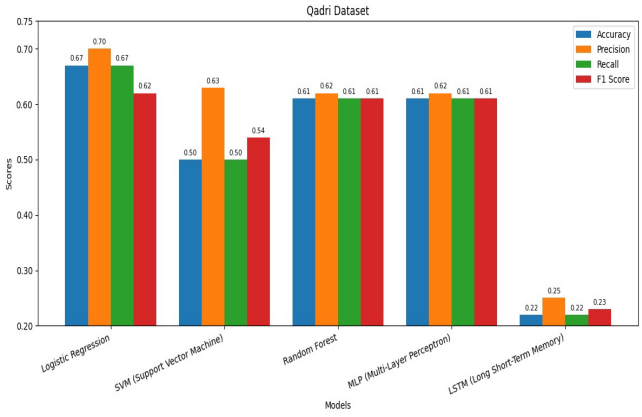


Fig 6: Models' Performance on the Qadri Dataset

B. Comparison between Eight Urdu Translations

This section showcases performance results from sentiment classification executed on eight different Urdu versions of Surah An-Naba, along with An-Naazi'aat. The research makes use of five different models, which include Logistic Regression (LR), along with Support Vector Machine (SVM) and Random Forest (RF) and Multi-Layer Perceptron (MLP), and Long Short-Term Memory (LSTM). An evaluation of these models exists in the TABLE V to TABLE VIII through their assessment of Accuracy, Precision, Recall, and F1-Score.

TABLE V
Comparison of Accuracy among Models and Translations

Dataset	LR	SVM	RF	MLP	LSTM
Maududi	0.78	0.78	0.61	0.78	0.66
Ahmed Raza	0.6	0.1	0.67	0.72	0.67
Ahmed Ali	0.72	0.67	0.61	0.67	0.72
Jalandhry	0.56	0.56	0.61	0.50	0.56
Qadri	0.67	0.50	0.61	0.50	0.22
Jawadi	0.67	0.56	0.78	0.83	0.61
Junagarhi	0.44	0.39	0.50	0.56	0.67
Najafi	0.67	0.67	0.67	0.61	0.50

TABLE VI
Comparison of Precision among Models and Translations

Dataset	LR	SVM	RF	MLP	LSTM
Maududi	0.79	0.79	0.61	0.75	0.69
Ahmed Raza	0.58	0.57	0.60	0.76	0.67
Ahmed Ali	0.50	0.45	0.42	0.52	0.47
Jalandhry	0.57	0.61	0.68	0.48	0.54
Qadri	0.70	0.63	0.62	0.53	0.57
Jawadi	0.78	0.57	0.89	0.88	0.84
Junagarhi	0.42	0.38	0.47	0.54	0.49
Najafi	0.65	0.45	0.76	0.56	0.59

TABLE VII
Comparison of Recall among Models and Translations

Dataset	LR	SVM	RF	MLP	LSTM
Maududi	0.78	0.78	0.61	0.78	0.74
Ahmed Raza	0.61	0.61	0.67	0.72	0.69
Ahmed Ali	0.54	0.50	0.46	0.50	0.47
Jalandhry	0.56	0.56	0.61	0.50	0.54
Qadri	0.67	0.50	0.61	0.50	0.54
Jawadi	0.75	0.55	0.71	0.85	0.79
Junagarhi	0.44	0.39	0.50	0.56	0.0.51
Najafi	0.62	0.50	0.62	0.58	0.59

A. Discussion on Model's Performance

Logistic Regression, Support Vector Machine, and Multi-Layer Perceptron delivered stable results when tested on various translations, especially within the Maududi, Ahmed Ali, and Jawadi datasets. The Random Forest model showed inconsistent results across translations. Its best

outcome (0.78% accuracy) appeared in Jawadi, yet its worst performance (0.61% accuracy) emerged in Maududi. Results of the LSTM model displayed diverging levels of accuracy; also, LSTM displayed high instability during Qadri dataset translation (0.22 accuracy).

TABLE VIII
Comparison of F1 Score among Models and Translations

Dataset	LR	SVM	RF	MLP	LSTM
Maududi	0.77	0.77	0.61	0.76	0.73
Ahmed Raza	0.58	0.58	0.62	0.72	0.68
Ahmed Ali	0.51	0.47	0.43	0.46	0.44
Jalandhry	0.55	0.55	0.59	0.44	0.49
Qadri	0.62	0.54	0.61	0.51	0.54
Jawadi	0.73	0.51	0.75	0.86	0.79
Junagarhi	0.43	0.38	0.48	0.54	0.49
Najafi	0.60	0.47	0.67	0.56	0.57

B. Effect of Translation Differences

The classification accuracy improved with some translations because they maintained linguistic consistency, although other translations demonstrated increased performance variation. Emotional tone distributions in the Jawadi translation achieved optimal accuracy rates of 0.83 with MLP and 0.78 with Random Forest. The Junagarhi translation achieved the least successful sentiment classification rates because its vocabulary appeared to create problems for sentiment analysis.

The difference in emotional tone seen between various Urdu translations can be explained by a number of factors. These range from the differences in the way translators understand particular Arabic words, their respective theological persuasions, linguistic tastes, and the periods in which the translations were made. These factors affect word selection and stylistic expression and, in turn, the way sentiment is communicated. This resonates with central ideas in translation studies, wherein semantic representation and cultural context play an integral part in structuring meaning in the process of translation.

1) High Performance for Positive and Cautionary Sentiments

The result shows that the Positive and Cautionary sentiment category values of precision, recall, and F1-score were greater with all the models. The models could identify Positive contexts successfully, for example, verses which highlight mercy, reward, and heaven, and Cautionary contexts, for example, verses which highlight penalty, warning, and hellfire. This indicates that the strong correlation between these sentiments and their related verses allowed for effective classification.

2) Low Performance for Neutral Sentiment

Analysis of the "Neutral" sentiment showed inadequate results since its dataset contained limited examples. The misinterpretation of content happened due to the overlap between the Cautionary and Neutral sentiments. The application of the SMOTE technique for balancing classes produced limited enhancements to the results. The diverse assignments of certain verses created problems for the trained systems because they existed in multiple classification categories simultaneously.

The distribution of classified sentiments across the categories Positive, Cautionary, and Neutral is presented in Fig 7.

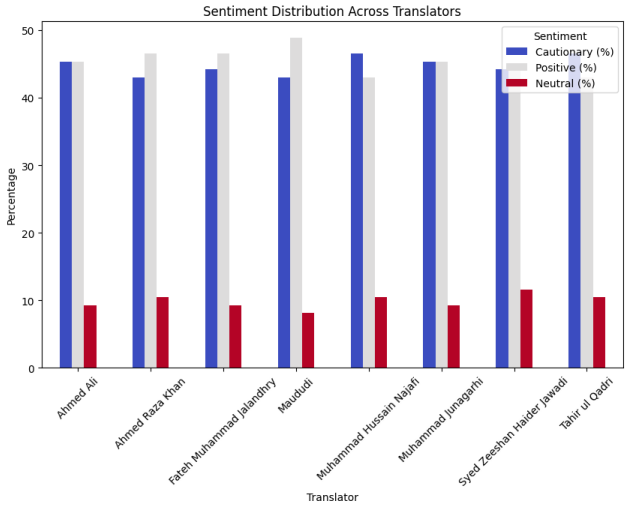


Fig 7: Positive, Cautionary & Neutral Sentiments

3) Challenges with Imbalanced Sentiments

The dataset contained fewer Neutral samples, which affected the performance of models and also the neutral sentiment category. Neutral sentiment category performs poorly, and it also affects the performance of all models because its number of ayahs is less compared to the other two categories, which is a challenge for models. SMOTE is used for handling this issue, but further improvement can also be made in the future, like a more balanced sentiment category dataset should be obtained that contains the same number of positive and negative, and cautionary statements. Advanced contextual embedding models created specifically for this task should be employed.

C. Why Comparison between Urdu Translations

The translation of Surahs by different scholars during different times resulted in various expressive approaches based on their interpretation methods. The emotional message of the text is altered depending on the translator's interpretive approach. Each translator, including Abul A'ala Maududi, Ahmed Ali, and Muhammad Hussain Najafi, presents their distinctive wording patterns affecting

sentiment analysis, and Ahmed Raza Khan and Syed Zeeshan Haider Jawadi demonstrate dissimilar emotional tonalities. The same verse can present various emotional meanings because different translations express them as either neutral or charged with either positive or cautionary messages. We can evaluate language expression changes between translations when analyzing Surah emotions through sentiment analysis because different word choices radically affect emotional characterization.

D. Comparison of Sentiment Analysis between Urdu and English Translations

This section explains the comparison of sentiment analysis between Urdu and English. The emotional tone analysis of eight Urdu translations and the ten English translations examined by Refaee were compared [1]. Sentiments are divided into three categories by the analysis: cautionary, positive, and neutral sentiment. The variation in sentiment distribution between the two languages is presented in TABLE IX below. The following abbreviations are used in the TABLE IX: U-Trans (Urdu Translators), CS (Cautionary Sentiment), PS (Positive Sentiment), NS (Neutral Sentiment), E-Trans (English Translators), CS (Cautionary Sentiment), PS (Positive Sentiment), NS (Neutral Sentiment).

TABLE IX
Comparison between English & Urdu

U-Trans	CS %	PS %	NS %	E-Trans	CS %	PS %	NS %
Ahmed Ali	45	45	10	Saheh	58	16	26
Ahmed Raza	43	47	10	Yusuf Ali	47	35	18
Jalandhary	44	47	9	Asad	57	36	7
Najafi	47	43	10	Shaker	37	37	26
Maududi	43	49	8	Maududi	56	38	6
Junagarhi	45	44	9	Hilaii & Khan	53	43	4
Jawadi	44	44	12	Khattab	50	23	27
Qadri	47	43	10	Pickthall	45	43	12
-	-	-	-	Qaribullah	38	38	24
-	-	-	-	Sarwae	45	44	11

E. Discussion on Findings

The results indicate that English translations had higher cautionary and neutral sentiments. On the other hand, Urdu translations had higher positive sentiment. This difference could result from variation in translation interpretation, cultural influences, and linguistic structure. A comparative visualization of Positive, Cautionary, and Neutral sentiments between Urdu and English translations is presented in Fig 8 to Fig 10.

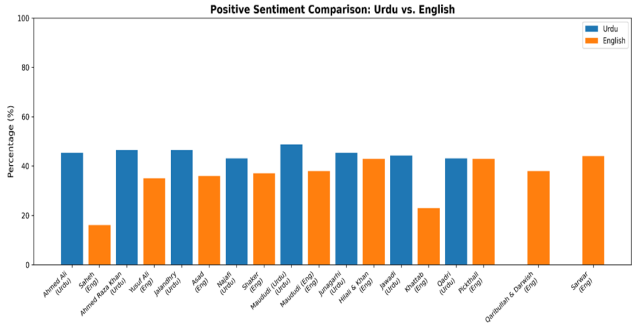


Fig 8: Positive Sentiment

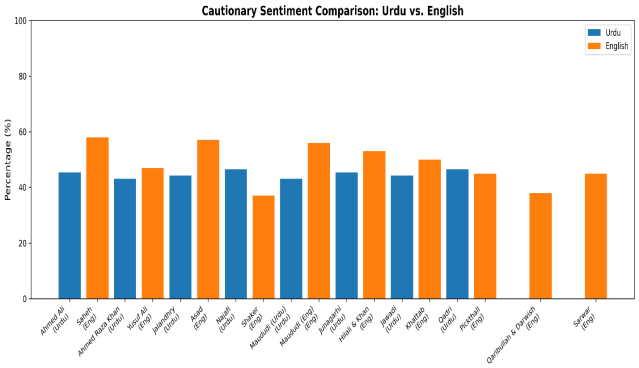


Fig 9: Cautionary Sentiment

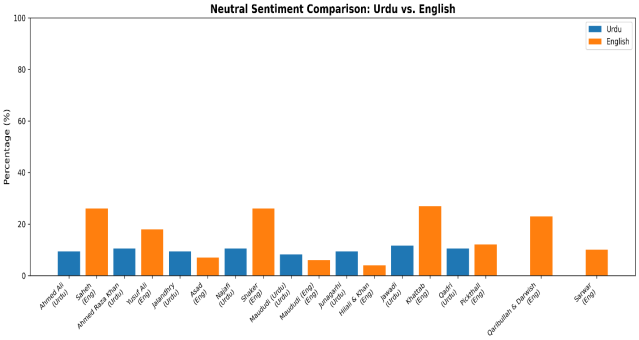


Fig 10: Neutral Sentiment

F. Discussion on Positive, Cautionary, and Neutral sentiment

The Urdu Maududi version demonstrates the highest positive sentiment at 48.8%, whereas the English versions of Sarwar and Pickthall have comparatively higher positive sentiment (44%). The cautionary tone of the English Saheh translation reaches 58%, whereas Urdu Najafi and Qadri translations present 46.5% cautionary sentiment.

The extent of neutral sentiments in English translations ranges between 4% and 27%, whereas Urdu translations stay between 8% and 11.6%. The English translation reaches its highest point of neutral sentiment through Khattab at 27%, and Shaker attains 26%. The Urdu Jawadi translation demonstrates the highest neutral sentiment level compared to other Urdu versions at 11.6%.

XII. Conclusions

This study analyzed the sentiment analysis through Urdu versions of Surah An-Naba and Surah An-Naazi'aat translated by Maulana Abul A'ala Maududi, Muhammad Hussain Najafi, Syed Zeeshan Haider Jawadi, Fateh Muhammad Jalandhry, Muhammad Junagarhi, Ahmed Raza Khan, Tahir ul Qadri, and Ahmed Ali. Machine learning models (Logistic Regression, Support Vector Machine, and Random Forest Classifier) and deep learning models (Multi- Layer Perceptron and Long Short-Term Memory) within sentiment analysis determined the three possible classifications of verses, which could be positive or cautionary, or neutral.

The research showed that positive and cautionary sentiments were successfully detected, but the identification of neutral sentiments proved difficult because of class distribution imbalances and sentiment overlap. The application of SMOTE balancing methods did not successfully increase the classification accuracy rates for neutral verses.

A comparative analysis of different Urdu translations emphasized differences in emotional tone across scholars' understandings, indicating how linguistic style and word choice affected sentiment classification. The examination of results through English translations enabled researchers to verify whether shifts in sentiment appeared because of language structure, together with translation preferences. These results contribute to understanding the effect of translation on sentiment expression in religious texts, helping refine sentiment analysis techniques for Qur'anic studies.

XIII. Directions for Future Work

In order to improve class imbalance, future research can investigate transformer-based models like UrduBERT, which are pretrained on extensive Urdu corpora and better capture the emotional and contextual subtleties of text compared to traditional models. Although this study focused specifically on Surah An-Naba and Surah An-Naazi'aat, the approach and methodology can be extended to other parts of the Quran. Expanding the study to additional Surah's beyond Surah An-Naba and An-Naazi'aat, incorporate sophisticated language aspects, and enhancing datasets can help make the results more general and provide a better picture of how emotions are conveyed in Qur'anic translations.

Future work will involve analyzing a broader range of Surahs to support more generalizable conclusions and to explore a wider spectrum of emotional tone variations across different translations. The emotional tone of every verse was manually labeled with three categories of sentiment: positive, cautionary, and neutral. This reduced taxonomic system was chosen deliberately to trade interpretability for feasibility, especially considering the intricacy of the religious text and constraints on annotation resources. Although not capturing all theological subtlety, these three categories were adequate to capture the grand emotional intentions present in the verses being analyzed.

Future research will extend this framework to incorporate a more nuanced taxonomy of emotions to more accurately capture the broad range of theological and emotional language in the Qur'an. Furthermore, addressing the class imbalance issue and model training can be improved by enriching the manually annotated dataset. The evaluation of multiple languages through multilingual analysis demonstrates how sentiment expressions differ between translations of the Quran.

References

[1] Refaee, M. (2023) Sentiment Analysis of one Chapter of Holy Quran for Ten Versions of English Translation.

[2] Salloum, S. A., AlHamad, A. Q., Al-Emran, M., & Shaalan, K. (2018). A survey of Arabic text mining. Intelligent natural language processing: Trends and applications, 417-431.

[3] Karami, M., Talebpour, A., Tajabadi, F., & 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).

[4] Alqarni, A., & Rahman, A. (2023). Arabic tweets-based sentiment analysis to investigate the impact of COVID-19 in KSA: a deep learning approach. Big Data and Cognitive Computing, 7(1), 16.

[5] Elhassan, N., Varone, G., Ahmed, R., Gogate, M., Dashtipour, K., Almoamari, H., . . . Hussain, A. (2023). Arabic sentiment analysis based on word embeddings and deep learning. Computers, 12(6), 126.

[6] Vora, M., Blau, T., Kachhwal, V., Solo, A. M., & Chandra, R. (2024). Large language model for Bible sentiment analysis: Sermon on the Mount. arXiv preprint arXiv:2401.00689.

[7] Sial, M. H., Javed, A., & Shahzad, K. (2021). Sentiment analysis for Urdu: Challenges and a machine learning-based framework. International Journal of Computational Linguistics, 9(4), 120-135.

[8] Duwairi, R. M. and I. Qarqaz (2014). Arabic sentiment analysis using supervised classification. 2014 International Conference on Future Internet of Things and Cloud, IEEE.

[9] Karami, M., Talebpour, A., Tajabadi, F., & 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).

[10] Sial, M. H., Javed, A., & Shahzad, K. (2021). Sentiment analysis for Urdu: Challenges and a machine learning-based framework. International Journal of Computational Linguistics, 9(4), 120-135.

[11] Zhou, X. and Y. Zhu (2024). Unsupervised Feature Classification-Based Sentiment Analysis of Online Social Network Texts.