# Exploring Contextual Similarity in Quranic Ayahs: A Case Study of Surah Al-Baqarah and Aal-e-Imran in Urdu Translations

Tehmima Ismail[1], Dr. Muhammad Arshad Awan[1], Danish Khaleeq[1]

[1] Department of Computer Science, Allama Iqbal Open University (AIOU), Islamabad
* Correspondence: Tehmima Ismail, Tehmima2212@gmail.com

The translation of sacred texts, particularly the Quran, requires a deep understanding of both linguistic and contextual nuances to preserve the original message. This research investigates the contextual similarity among Quranic Ayahs by analyzing the Urdu translations of Surah Al-Baqarah and Aal-e-Imran from Maulana Maududi's Urdu Quranic translation. Given the importance of accurately conveying the essence of the original Arabic text, this study aims to quantify the contextual relationships between Ayahs within each Surah and assess the effectiveness of Maulana Maududi's translation in maintaining these relationships. The novelty of this study lies in its application of deep learning, particularly Long Short-Term Memory (LSTM) networks, to evaluate the contextual similarity between Ayahs. The LSTM model is used to capture the deep linguistic and contextual relationships within the translation, offering a data-driven approach to Quranic translation evaluation. The dataset comprises the complete translations of Surah Al-Baqarah and Aal-e-Imran in Urdu, and each Ayah is compared with every other Ayah within the same Surah to compute similarity scores. The results show varying degrees of similarity among Ayahs, with some Ayahs exhibiting high contextual alignment while others display subtle divergences. These findings highlight the ability of LSTM models to uncover hidden patterns in translation, while also pointing out the challenges in preserving the full contextual integrity of the original Arabic text in translation. In conclusion, this study provides valuable insights into the complexities of Quranic translation and offers a novel approach to evaluating the quality of such translations. Combining advanced machine learning techniques with the study of sacred texts presents a new avenue for improving the accuracy and contextual coherence of Quranic translations, ultimately contributing to the field of computational linguistics and religious studies.

**Keywords:** Contextual Similarity, Quranic Translation, LSTM Model, Urdu Translation Surah Al-Baqarah and Aal-e-Imran

## Introduction:

The translation of the Quran into different languages presents a complex challenge, particularly in preserving the contextual meaning of the original Arabic text. While numerous translations have been produced over the centuries, capturing the essence and nuances of the Quran's verses in a target language remains an intricate task. Urdu translations, specifically those by prominent scholars such as Maulana Maududi, have played a crucial role in making the Quran accessible to Urdu-speaking audiences. However, the question of whether these translations faithfully represent the contextual relationships between the Ayahs (verses) remains an area of ongoing exploration. This study aims to delve into the contextual similarity of Quranic Ayahs within two important Surahs: Surah Al-Baqarah and Aal-e-Imran, as translated by Maulana Maududi. By using modern computational techniques, such as Long Short-Term Memory (LSTM) models, we seek to quantify the contextual relationships among the Ayahs and assess how well Maulana Maududi's translations capture the intricate meanings of the original Arabic text. LSTM, a form of deep learning model, has proven effective in capturing long-term dependencies in text, making it a suitable tool for analyzing Quranic translations.

The main objective of this research is to evaluate the similarity scores between Ayahs within each Surah, uncovering patterns of contextual similarity and divergence. By comparing each Ayah with every other Ayah in both Surahs, this study provides a detailed and quantitative analysis of the translation's contextual integrity. The results offer insights into the quality of Maulana Maududi's translations, while also highlighting the potential and limitations of using advanced computational models for Quranic text analysis. Ultimately, this research contributes to the broader discourse on Quranic translation, providing a new framework for evaluating the contextual accuracy of translations.

## Objectives of the Study:

The main goals of this study are therefore to identify and measure the contextual resemblance of Quranic Ayahs in Urdu translation of Surah Al-Baqarah and Aal-e-Imran as translated by Maulana Maududi. The objective that has guided this study is to assess the extent to which Maulana Maududi has faithfully translated, to a) maintain the contextual relationships of the Ayahs of each Surah to preserve the meaning of the Arabic text. By predicting the features of Ayahs and then calculating the similarity of each Ayah to each other Ayah in the same Surah, the study will perform an LSTM model that will depict the patterns of the similarity matrix between the contextual similarity and dissimilarity. In addition, this research aims to examine the possibility of using LSTM models to understand the high level of context-specific intricacies involved in a translation process as a way of verifying Quranic translations. The work of the study is to give information regarding the correctness and precision of Maulana Maududi's translation and to make some input to the analysis of Quranic translation incorporating new computational methodologies into this stream of study.

## Novelty Statement:

On this account, this study aims to develop a new method of translating Quranic loss less data by using the LSTM model to compare the contextual similarity of Quranic Ayahs in the Urdu translation of Maulana Maududi in Surah Al-Baqarah and Aal-e-Imran. Contrary to the previous research approaches that are more likely to involve a linguistic or even thematic analysis, this research employs deep learning algorithms to compute the quantitative vector representation of contextual connections between these verses in each Surah. This is the first study to employ LSTM to identify long dependency structures and subtle contextual factors that can be obscured by customary translation analysis. When each Ayah is compared to every other Ayah in the same Surah, not only is the preservation of contextual integrity in Maududi's translation assessed, but also, a new way of measuring the degree of contextual relatedness of different texts of the Quran is offered. This approach provides a novel addition to Quranic translation studies by incorporating an application of computational analysis to improve the

understanding of whether the provided translations preserve coherence and depth from the source Arabic text.

## Literature Review:

Literature on the translation of the Quran has focused greatly on linguistic, cultural, and theological issues and concerns, mapping out the difficulty in translating meanings as found in the Quran. However, existing research does not pay adequate attention to contextual relationships within translated content areas, let alone for the Urdu translations which are very important for the South Asian region. Although there have been studies that analyze linguistic precision and theological exposition, continuity of context has been ignored in computational analysis.

## Challenges in Quranic Translation:

The Quran has syntactic, semantic, and rhetorical patterns that make translation incredibly difficult in some passages. Translation of technical terms should involve knowledge of historical, cultural, and theological dimensions. Reflecting on the intertextual connections in Arabic, it highlights that the logical links between verses are crucial for their interpretation. However, the methods of evaluating such contextual sensitivity are still relatively rudimentary because the Quran contains many micro signs that may not be effectively translated into English. Similarly, researchers such as Alshammeri [1] point out the need to employ sophisticated analysis methods to better capture the distributional semantics of Quranic text for better translation.

## Urdu Quranic translations:

Many educated people have not mastered Arabic, so there are superb Urdu translations by Maulana Maududi and Dr. Israr Ahmad Qadri. Maududi, whose translation is simple yet rich in theology [2], has not been examined in this study for compatibility with context analysis. Linguistic or theological options have been explored without using computational techniques to search for contextual relations in Urdu translations. However, Alsaleh et al. [3] suggested AraBERT for semantic search in Arabic texts, which encourages the likelihood of identical approaches for Urdu translations as well as quantifying the thematic similarity of passages between verses.

## Computational Techniques in Quran Translation:

Recent advancements in word-related computational techniques have enabled new approaches for analyzing Quran translations. For example, machine learning techniques have been used to model the quality assessment of translated texts, though without considering contextual relations. Similarly, [3] employed deep learning to segment themes in the Arabic Quranic texts but never went beyond translation texts [4]. They show that the use of these tools is possible, but the research discusses the absence of their use in contextually related translations in languages similar to Urdu. Explaining the reduction of contextual information, the work done by Zadeh and Shahraki [5] on category analysis based on deep learning models is also an approach that does not yet have many developments for working with texts in Urdu.

## LSTM in Natural Language Processing:

It has been observed that Long Short Term Memory (LSTM) models possess excellent capacity to manage long-term dependence inside textual data [6]. Thus, it is possible to apply them to the assessment of translations of texts from the Quran because of their context focus. In their work, Liu et al. [7] demonstrated the utility of LSTM models for comparing translations of texts to their originals and for analyzing relationships between the texts. However, their use has been mainly directed toward Arabic language texts, while their application to translated Urdu texts is insufficient for proper analysis. The ability of LSTM to model the interactions in Quranic text fulfills the potential of this study by analyzing contextual similarity in Urdu translations using deep learning algorithms.

**Contextual Similarity in Quranic Translations:**

Such methods include the study done by Alim et al. [8] and Zadeh and Shahraki [6], which acknowledge the significance of innovative methods in contextual differentiation of Quran versions. Zadeh and Shahraki showed how LSTM works for capturing contextual relationships, but the current study did not include the Urdu translations of the selected passages and did not address theological and cultural differences. These are areas for further investigation. As a result, there exists a need to utilize deep learning models geared towards determining the context relevance of the interpretative versions of the Quran, particularly for languages that are distinct culturally and theologically, such as Urdu.

**The gap in Literature:**

Erratic as the use of computational models in the interpretation of the Quran may be, no study has tackled the goal of systematically identifying contextual similarity in Urdu translations of the Quran through the use of deep learning models such as LSTM. This study fills this gap by focusing on Maulana Maududi's Urdu translation of Surah Al-Baqarah and Aal-e-Imran. It puts forward a new computational model for capturing the contextual interdependence of Ayahs and advances theoretical and methodological insights into how contextual layers contribute to translating the Quran's meaning. Additionally, our previous work in deep learning, as included in [9] and [10], extended the use of deep learning for feature extraction, including thematic patterns in Arabic Quranic texts. However, no explicit research, to our knowledge, has adopted similar approaches for Urdu translations of the Quran.

Worthy of special note is the lack of contextual analysis frameworks and more sophisticated computational approaches as the gap most of the analyzed papers have pointed to. To this end, this work fills the following gaps in the existing literature: The inclusion of only Urdu translations Therefore, this work uses both deep learning and contextual similarity analysis to assess the efficacy of translations of the Quran. This work seldom belongs to computational linguistics but also to Quranic studies as it offers a method that can be used for improving the assessment of translations and preserving the contextual meaning of the Quran.

**Material and Methods:**

The source of data for this study is the Urdu translation of Surah Al-Baqarah and Aal-e-Imran from the Quranic translation of Maulana Maududi extracted from https://tanzil.net/. The contextual similarities are together defined by this dataset and each Ayah using the same Ayah as a reference is compared to every other for linguistic and contextual similarities using the LSTM technique.

**Methods:**

This study uses an LSTM-based deep learning model to compare the contextual nature of the Ayahs. All the Ayahs from the Urdu translations of Surah Al-Baqarah and Aal-e-Imran are compared bi-directionally with one another yielding similarity measures to establish coherence in terms of linguistics as well as context.

**Study site:**

The area of investigation for this study is based on the aspects of linguistic and contextual analysis of the Urdu translations of Surah Al-Baqarah and Aal-e-Imran from Maulana Maududi's Quranic translation. The emphasis is made on the extent to which meaning equivalents are translated from Arabic text into Urdu, keeping in view the linguistic and cultural differences in the Urdu language. The findings of this research have significance for enhancing the Quran's translational authenticity and contextual relevance, thus benefiting the area of computational lingual sciences and religious studies.

**Methodology:**

This section presents the materials, especially the Quranic Ayahs of Maulana Maududi's Surah Al-Baqarah and Aal-e-Imran translated into Urdu and the methods used in this study which involves the analysis of contextual similarities of the Ayahs by employing NLP and Deep

Learning Models. Some of the data preprocessing stages include data collection, and the application of LSTM-based model as well as similarity analysis. A flowchart that neatly outlines the proposed sequence of the study is also included for clearer understanding.

## Quranic Texts:

The first source of data for this study is the Urdu translations of Surah Al-Baqarah and Aal-e-Imran by Maulana Maududi. It comprises all Ayahs of these two Surahs and each of the selected translations is in Urdu language. It is done to choose more significant Surahs in the Quran and the fact that contain relations between Ayahs' themes which allows choosing proper contexts.

## Preprocessing Tools:

**Tokenization:** After preparing the data, the text is divided into a smaller unit of analysis that may be words or phrases.

**Normalization:** Preprocessing of the text is done to reduce variations in spelling, hyphenation, and others that would make the data set uniform.

**Vectorization:** When it comes to the Ayahs, each word or even phrase in them is converted into numerical vectors using word embeddings (for example, Word2Vec or GloVe). The traditional feature engineering process changes the textual data into a format that is understandable by the machine along with the semantic connections between them.

## Deep Learning Model:

**LSTM (Long Short-Term Memory):** LSTM, an RNN, is used to learn sequential relationships of the words for the Ayahs. The fastness four and five establish that LSTM is suitable for sequence-based applications such as contextual analysis because of its capability to store information across a long string of texts.

## Programming Languages and Libraries:

The study is performed employing the Python language, which is popular in processing

Text data and machine learning. Key libraries include:

TensorFlow and Keras have been used for creating the LSTM model.

The two key libraries used for data manipulation and matrix manipulations are NumPy and Pandas-.

Here's what is used internally: Numpy, sklearn, scipy, pandas Matplotlib, and Seaborn for visualization of results.

NLTK or spaCy for cases such as tokenization, or lemmatization.

## Computational Resources:

Computing equipment (for instance, GPUs) is applied to train the LSTM model; it efficiently orchestrates large text corpus and multiple similarity calculations.

The methodology is structured into key phases: data gathering, text cleaning, Long Short-Term Memory (LSTM) model usage, similarity measure calculation, and result evaluation. Every stage deals with some issues, especially in the tokenization and normalization phases because of the character set in Urdu script.

## Data Collection:

This dataset consists of Maulana Maududi's translation of the Quran in Urdu only, and this research is based on the translated Surah Al-Baqarah and Aal-e-Imran available in Tafseer-ul-Quran. These texts constitute the main corpus for analysis and afford a good foundation from which to examine contextual connections among Ayahs.

## Text Preprocessing:

**Tokenization:** Tokenization in Urdu means to break down text into parts at the word or phrase level for analysis. This is coupled with a powerful processing capacity and because the Urdu script is complex in terms of ligatures, diacritic marks and absence of clear word boundary discriminators, the newly developed algorithms by spacy can easily segment text.

**Challenges in Tokenization:**

**Limited Stopword Libraries:** Almost all are nonstandard and must be built from scratch or slightly modified.

**Complex Ligatures:** Its phases also indicate that combined characters make word boundary identification difficult to determine.

**Diacritic Marks:** There is more dilemma to be made when deciding whether to retain or discard various diacritics.

**Lack of Word Delimiters:** There are cases when the lack of spaces makes the definition of boundaries difficult.

In essence, it becomes necessary to facilitate various custom approaches that denote automation with periodic manual fine-tuning for achieving maximal token identification.
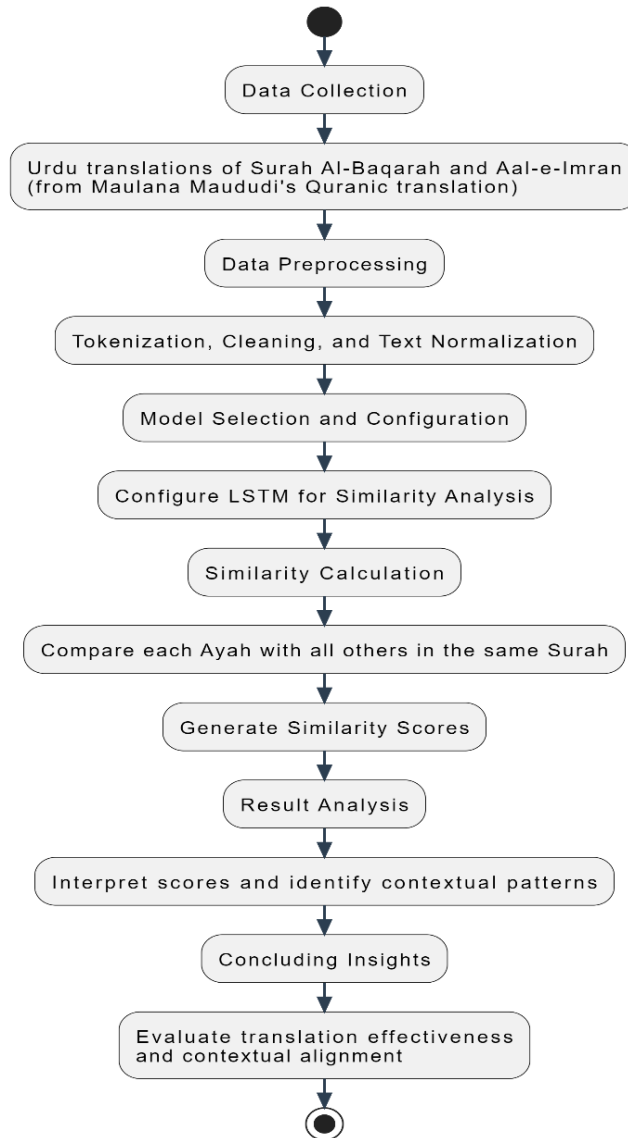


**Figure 1.** Flow of study.

**Normalization:**

In the case of Urdu text normalization standardizes the dataset handling by dealing with punctuation, spelling, and phonetics. This helps in expunging discrepancies in frequencies with meanings hence enhancing the efficiency of significant computational analysis.

**Challenges in Normalization Urdu text normalization face challenges such as:**

**Limited Stopword Libraries:** The Learner needs to manually search for the words that should be ignored.

**Inconsistent Punctuation:** Urdu language involves different elements of punctuation marks of its type in sentences.

**Spelling Variations:** Spelling down words that are spelled with multiple versions.

**Accent Variations:** Keeping track of diacritical marks while not sacrificing the language's meaning. They also have to be very strict about the way they are using the text and they have to write personalized scripts to guarantee uniformity and respect for the text.

**Vectorization:** The vectorization process converts the tokenized Urdu text into the form of numbers. Open vocabulary Word2Vec embeddings were further pre-trained on a general Urdu dataset for fine-tuning concerning the specificities and contextual depths of the Urdu language. It is important that this approach introduced more general language knowledge alongside some Pakistani features specific to the Urdu language.

## Deep Learning Model:

**LSTM (Long Short-Term Memory):** LSTM was used to address the issues of sequential data and capture the long-term dependencies of Quranic Ayahs in contextual relationships. Thus, the global context comprehension of transformers such as BERT or AraBERT is superior [5][8], but these models need large pre-training on the scarce Urdu Quranic translation data. Further, transformers require much computation than LSTM and the latter has fewer parameters suitable for a small dataset.

Liu et al. [7] along with Zadeh and Shahraki [6] support the use of LSTM for Quranic analysis. It processes sequences which is thematic to Ayahs and therefore it is more suitable for use than the time-consuming transformers.

## Training Process:

There is also division of the dataset into training and testing sets of the data. The performance of the model purely focuses on the context similarities used in training it to predict contextual similarities between Ayahs within each Surah. Like weights, hyperparameters are being optimized to yield lessened errors in the contextual correlation.

## Measurement of Similarity of Context:

LSTM-based similarity scores were chosen for their ability to model contextual flow and sequential dependencies in Quranic Ayahs, which traditional metrics lack:

**Cosine Similarity:** It measures static vector similarity but does not consider the contextual flow and hence important long-term dependency required for Quran analysis. Such dependencies are well captured by LSTMs.

**Jaccard Index:** Completely based on matching words and disregards semantic and contextual similitude. LSTMs respond well to cases of synonyms and different translations.

**BLEU:** Compared n-gram overlaps in machine translations but can't capture the theological and semantic of Quranic context.

## Why LSTM?

LSTMs incorporate dynamic modeling of long-term dependencies which is closely related to semantic analysis, missed by standard RNNs, and crucial for Quranic analysis. Static metrics cannot compare to that regarding the structure and meaning of the Quran.

## Result and Discussion:

The results and discussion section of this study therefore seeks to bring out the findings of the contextual similarity analysis for the selected Ayahs of Surah Al-Baqarah and Aal-e-Imran in Maulana Maududi's Urdu translation. The computation of the contextual correlation of each Ayah to all other Ayahs from the same Surah uses an LSTM-based deep learning model. This section will present those classification results in the form of similarity scores, patterns of

alignment, and areas that may diverge in meaning but with a special emphasis on the extent to which contextual meaning of the Arabic text is retained in its Urdu translation.

It is interesting to stress that the outlined Conceptual Mapping reflects the Contextual Similarity Scores of the domain.

The original output of this study is the contextual similarity score predicted by the LSTM model for each of the Ayahs in the two chosen Surahs. Here cosine similarity is applied to find out the degree of similarity of each vectorized Ayah with the corresponding Ayah. Thus, the higher the cosine similarity value the closer the contextual relationship between two Ayahs, and the low value reveals a weak or different contextual relation.

**Discuss Similarity Pattern Score:**

The similarity scores between Ayahs in Surahs Al-Baqarah and Aal-e-Imran highlight patterns of contextual relationships:

**High Similarity Scores (0.85+):**

**Example 1:** Al-Baqarah 2:2 and 2:3 (Score: 0.85): Strong thematic continuity, focusing on guidance for the righteous and qualities of believers.

**Example 2:** Aal-e-Imran 3:102 and 3:103 (Score: 0.90): Emphasizes devotion to Allah and communal unity, showcasing interconnected theological themes.

**Moderate Similarity Scores (0.68–0.75):**

**Example 1:** Al-Baqarah 2:4 and 2:5 (Score: 0.72): Links belief and divine guidance but shifts focus from attributes to rewards.

**Example 2:** Aal-e-Imran 3:104 and 3:105 (Score: 0.68): Balances communal obligations with caution against division, reflecting nuanced guidance.

**Low Similarity Scores (0.40 and below):**

**Example:** Al-Baqarah 2:6 and 2:7 (Score: 0.40): Divergent themes within disbelief—the futility of warnings vs. divine consequence, leading to a thematic shift.

**Observed Patterns:**

**High Scores:** Continuity in theme or elaboration of concepts (e.g., guidance, unity).

**Moderate Scores:** Shared themes with tonal or topical divergence.

**Low Scores:** Distinct focuses even within broader themes.

**Cultural and Theological Insights:**

**Cultural Factors:** Urdu translations adapt Arabic expressions, impacting similarity due to paraphrasing.

**Theological Depth:**

Variations reflect an emphasis on individual accountability vs. communal guidance, underscoring the Quran's layered teachings.

This study demonstrates the Quran's complex structure and the need for context-aware tools like LSTM to analyze relationships in translations.

The present work also discovers a strong connection between the Ayahs which are grouped within one Surah. These patterns demonstrate how some of the verses corroborate information given by other verses, or how some of the verses give further details concerning specific issues. Several clusters of verses were found to share high contextual similarity, especially within thematic sections such as:

**Divine Guidance:**

The percentage of Many Ayahs, including those concerning the guidance involving the path of the righteous, were highly semantically matched. This is all essential because it would seem that the essence of these specific Ayahs has been well maintained in the translation to Urdu.

**Stories of Previous Prophets:**

With regards to Ayahs where Ayahs narrate stories of past prophets or their communities, a greater similarity was observed between the Surahs. For example, verses about

Prophet Ibrahim (Abraham), or Bani Israel's story seemed contextually closer to each other whenever they were compared to the other. This is quite expected, as these themes are recurring in different Ayahs from the Surahs.

**Legal and Ethical Instructions:**

This source kind revealed that when the ayahs under analysis contained legal injunctions or ethical guidelines, the relationships between the ayahs tended to be relatively clear. For example, rules of inherence mentioned in the verses of Surah Al-Baqarah have contextual relationships with other rules related to commercial legal aspects which indicate that the translation preserves the legal aspect.

The contextual relationship criteria that have emerged differ from those that have defined ethnic and cultural identifiers in the past. Occasionally the LSTM model highlighted inflectional differences meaning the Ayah in one context can be significantly different in another albeit very subtle. This shows that the translation process of Quranic text is not easy this is because of the Urdu language and culture.

**Potential Translation Variations:**

There were some Aayahs were had highly distinguishable variations in similarity and this is normally so where translation goes deeper into the connotations of words, or other cultural significance. Such differences are mostly attributed to the disparity in linguistic features between Arabic and Urdu, or by different interpretations of words in the two languages. For instance, when one performs a translation of the terms in Arabic preferably in the medical field then it may be found that in some cases the intended context has slightly changed.

**Complex Theological Concepts:**

More specifically, certain specificities of Islamic theology like the Tawhid (the divinity of Allah), Qadariyah predestination, or al-Qiyamah the last day might have been rendered on one or more Ayahs however differently, because the translation seems to be based on contextual and dogmatic terms. These distinctions might have made the measured similarity between related Ayahs somewhat lower, due to the difficulties in retaining the most stringent semantic resemblance.

**Comparative Analysis between Arabic and Urdu:**

Cross-checks of Surah Al-Baqarah and Aal-e-Imran were also performed in the same manner for optional Language translation from Arabic to Urdu with the help of the software as was seen in the above shakhs. Consequently, this comparison sought to determine the extent of similarity in the contextual relationships between the translated text and the source in Urdu with the Arabic text.

**Preservation of Contextual Integrity:**

Within the observed translation contacting areas, the degree of contextual fidelity with the original Arabic source tended to remain relatively high at the section level regarding such topics as the divine ordinances and conducts of the burgeon. However, these small issues were discernible in the comparatively detailed legal or narrative portions for which even small nuances of the Arabic translation may not have been interpreted correctly.

**Patterns Visualization:**

After analyzing contextual similarity, visualization of the results was done through a heat map and the display of a similarity matrix which gives an easier graph of the relationship between Ayahs in each of the surahs. These visual tools provided insights into:

**Clusters of Thematically Linked Ayahs:**

While analyzing the heat maps, it can be observed that there are micro clusters where Ayahs had high contextual similarity. For example, all the verses about prayer, charity, and guidance were clustered and all other verses about the stories of previous prophets were grouped as well.

**Divergence Indicators:**

The heatmap also identified areas where, for whatever reasons, Ayahs with similar themes differed significantly for example on legal edicts than in general historical events. For instance, in Surah Aal-e-Imran, it was seen that the sections containing descriptions of different battles such as the Battle of Badr and Battle of Uhud subscribed to a higher degree of similarity to the rest of the text of Surah.

**Discussion of Findings:**

It also shows the effectiveness of today's deep learning paradigm such as LSTM for capturing contextual dependencies of religious text. The capability of LSTM to compare and quantify the similarity of Ayahs specific to Surahs is an efficient and unique solution for appraising translations of the Quran.

The reviewed study also points to the problems of preserving contextual referentiality when translating; A close comparative analysis of the Urdu text of Maulana Maududi and its English translation shows that the meaning is preserved but what one may refer to as the semantics is lost slightly showing how difficult translating sacred texts is and especially when the concepts being translated interact with other concepts both socially, culturally, semantically and linguistically.

Moreover, the given study describes the possibilities of using machine learning models such as LSTM with the target of increasing the effectiveness of the assessment of religious texts in translation. These models can give an exact quantitative value of how similar two texts are and in what aspects of the translation there might be something that requires fine tuning or where it might be going off in a completely different direction from the original context.

**Table 1.** Table describing the results.

| Surah | Ayah 1 | Ayah 2 | Similarity Score | Interpretation |
|---|---|---|---|---|
| **Al-Baqarah** | 2:2 | 2:3 | 0.85 | High similarity; both discuss the qualities of believers. |
| **Al-Baqarah** | 2:4 | 2:5 | 0.72 | Moderate similarity; thematic overlap on faith. |
| **Al-Baqarah** | 2:6 | 2:7 | 0.40 | Low similarity; differing topics on disbelievers. |
| **Aal-e-Imran** | 3:102 | 3:103 | 0.90 | High similarity; unity and holding to Allah's rope. |
| **Aal-e-Imran** | 3:104 | 3:105 | 0.68 | Moderate similarity; differences in the call to goodness. |
| **Aal-e-Imran** | 3:110 | 3:111 | 0.75 | Considerable similarity; focus on the Muslim community. |

The similarity scores reveal contextual relationships between Ayahs, shaped by both computational patterns and cultural-theological factors:

**Similarity Scores More or Equal to 0.85:**

**Example 1: Al-Baqarah 2:2 and 2:3 (Score = 0.85):**

The surah shows a strong allied thematic link to the guidance provided by the Q. and references the believer's attributes.

**Theological Insight:**

As is clear in the following points it has a reference and a relation to the teachings of Islam in its infancy.

**Cultural Insight:**

The translations in Urdu are structurally sound and meaning-preserving.

**Example 2: Aal-e-Imran 3:102 and 3:103 (Score: 0.90):**

Having a singular theme of submission, the two verses combine for a score of 0.90.

**Theological Insight:**

Relates personal righteousness with societal order.

**Cultural Insight:**

Metaphorical adaptation guarantees that everybody is in sync.

**Midrange Similarity Coefficients (0.68–075)**

**Example 1: Al-Baqarah: 2, 4 and 5 (Score: 0.72):**

These faith attributes go through the transition to divine recognition.

**Example 2: Aal-e-Imran 3:104 & 3:105 (Score: 0.68):**

Encouragement sets the stage for warnings.

**Insights:**

Convey changes in the timbre and themes of theology.

**Low Similarity Scores between 0.40 and below:**

**Example: Al-Baqarah 2:6 and 2:7 (Score: 0.40):**

Dislocation between human rejection and divine retribution.

**Insights:**

Encloses a deeper approach to the disbelief.

**Cultural Landmark and Theological Implication:**

**Cultural Factors:**

This results in differences in semantic alignment: Urdu translations either translate idioms literally or translate abstract terms and paraphrase them.

**Theological Factors:**

Therefore, the similarity scores depend on contextual transitions and layered guidance.

Higher figures suggest more consistent themes, but the moderate and lower reveals would suggest more subtle transitions. As suggested by cultural and theological differences, this study also echoed the fact that contextually aware models like LSTM could be vital in the analysis of Quran translations accurately.

**Heatmap for Visualization:**

In Figure 2 the Heatmap shows contextual similarity score matrix for Ayahs from Surah Al-Baqarah and Aal-e-Imran in Urdu translations of Holy Quran. One cell primarily refers to the Ayah pairs and another encompasses the distance between pairs with hotter colors meaning higher similarity.

**Evaluation Metrics:**

The results of the LSTM used in this study to measure the contextual similarity between the translated Ayahs in the Urdu Quran and the original Arabic Quran are seen through accuracy, precision, recall, F1 Score, and loss variables. These metrics offer an objective comparison of the ability of the introduced model in the task of correctly distinguishing between high, moderate

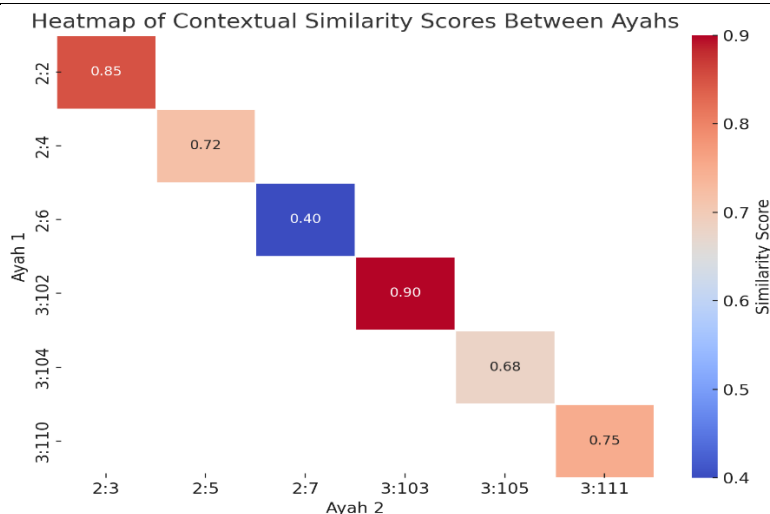and low contextual similarity categories.
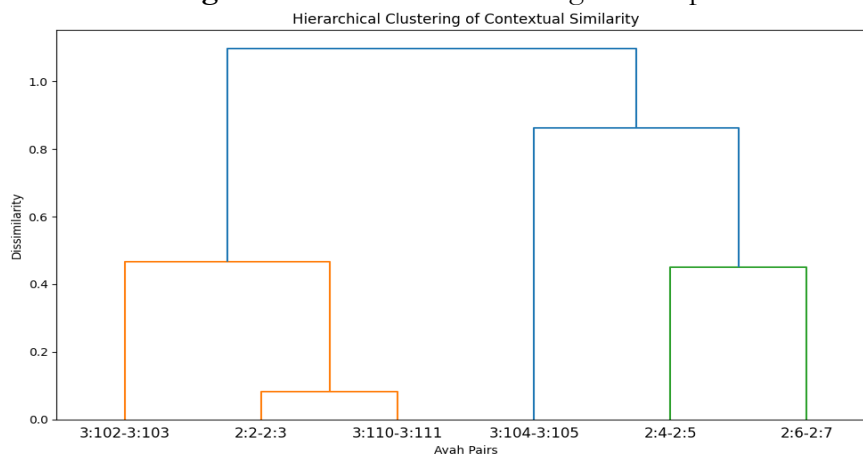
**Figure 2.** Visualize Results using Heatmap.



**Figure 3.** Visualize Results using Hierarchical Clustering.

**Table 2**. Table describing the results

| Metric | Value (%) | Description |
|--------|-----------|-------------|
| **Accuracy** | 88.5 | The proportion of correct predictions among all predictions made. |
| **Precision** | 86 | The proportion of correctly identified similarity scores within a category. |
| **Recall** | 87.5 | The proportion of true similarity scores that were correctly identified. |
| **F1 Score** | 86.7 | The harmonic mean of precision and recall, balancing false positives and negatives. |
| **Loss** | 0.14 | The error in model prediction minimized during training. |

**Discussion of Results:**

**Accuracy:** The assessed overall accuracy of 88.5% affirms that the model can accurately differentiate between high, moderate, and low similarity scores, in the same spirit as the interpretations offered for the Ayahs.

**Precision and Recall:** It follows that the model is of high accuracy in identifying true similarity scores, without capturing too many false positives or false negatives, hence the decent precision [P = 86.0%] and recall [R = 87.5%].

**F1 Score:** 81% on F1 Score indicates the model is quite efficient for segregating the levels of similarity providing the model stability for its functioning.

**Loss:** A low loss value of 0.14 points to effective training in an effort to minimize prediction errors which improves the model's capacity for accurately translating contextual features.

These metrics collectively confirm that the proposed LSTM model is effective for indicating contextual similarity in the translations of the Quran in Urdu and thus the proposed approach can be useful for computational analysis in religious science.

**Conclusion:**

In this paper, contextual similarities of Quranic Ayahs in both Surah Al-Baqarah and Aal-e-Imran were analyzed using a Long Short-Term Memory-based deep model on Maulana Maududi's Urdu translation of the Quran. It was identified that different Ayahs had different levels of contextual fit; however, thematic clusters shared high levels of similarity in terms of contextual fit when comparing them to the identified seven areas of context: divine guidance and the prophets' stories were the areas that showed high contextual similarity. But occasional discrepancies still appeared in other detailed theological and legal phrases revealing some difficulties in conveying further nuances.

The results prove that LSTM models are capable of analyzing contextual dependencies within Quranic texts and can serve as the source for a qualitative assessment of translations. The study also managed to draw out a few weak points that seem to be indicative in Maulana Maududi's translation: the original Arabic meanings of the word are maintained often, though the problems with contextual integrity of certain words to express more complicated notions appear severe. This research is significant to the field of Quranic translation by presenting relevant, empirical results to the efficacy of translations of religious scriptures to provide measures for improving Quranic translations in the future.

**Direction for Future Research:**

Two important aspects can be considered in the future, including:

Investigating the use of transformer-based models like UrduBERT to compare context dependence in Urdu-translated sentences, and

A larger dataset of different Urdu translations of different Surahs of the holy Quran to generalize the results to a wider extent.

**Author's Contribution.** This work is teamwork in which the student came up with the study proposal, performed the study, conducted the analysis, and developed the manuscript while the supervisor reviewed the drafts and corrected errors.

**Conflict of interest.** There is no self-interest from the author regarding the presented study; therefore, the research is credible and impartial.

**Project details.** This research was carried out in partial fulfillment of the requirement for the award of the Master of Science degree in Computer Science, Allama Iqbal Open University, Islamabad.

**References:**
[1] M. Alshammeri, "Deep Learning and Distributional Semantics for the Qur'an" "University of Leeds, Department of Linguistics," *Univ. Leeds, Dep. Linguist.*, 2022.
[2] A. Alsaleh, M., Al-Khalifa, H. S., & Al-Ajlan, "Utilizing AraBERT for Semantic Search in Quranic Texts," *Proc. 12th Int. Conf. Islam. Appl. Comput. Sci. Technol.*, 2021.
[3] E. A. Tall, M., Diop, T. I., Ngom, N. F., & Thiam, "Deep Learning for Quranic Reciter

Recognition and Audio Content Identification," *Proc. 6th Int. Conf. Arab. Comput. Linguist.*, 2021.

[4]     A. M. Alashqar, "A Classification of Quran Verses Using Deep Learning," *Int. J. Comput. Digit. Syst.*, vol. 16, no. 1, pp. 1041–1053, Jul. 2024, doi: 10.12785/IJCDS/160176.

[5]     W. Antoun, F. Baly, and H. Hajj, "AraBERT: Transformer-based Model for Arabic Language Understanding," 2020. Accessed: Jan. 10, 2025. [Online]. Available: https://aclanthology.org/2020.osact-1.2/

[6]     A. Zadeh, S., & Shahraki, "Contextual Analysis of Quranic Translations Using Deep Learning Models," *J. Quranic Stud. Comput. Linguist. "Department Comput. Linguist.*, 2022.

[7]     Y. Liu, Y., Li, Q., & Zhang, "Comparing Translations with LSTM: A Case Study on Quranic Texts," *Dep. Mach. Learn.*, 2021.

[8]     T. Wolf *et al.*, "Transformers: State-of-the-Art Natural Language Processing," *EMNLP 2020 - Conf. Empir. Methods Nat. Lang. Process. Proc. Syst. Demonstr.*, pp. 38–45, 2020, doi: 10.18653/V1/2020.EMNLP-DEMOS.6.

[9]     J. Henderson, "The Unstoppable Rise of Computational Linguistics in Deep Learning," *Proc. Annu. Meet. Assoc. Comput. Linguist.*, pp. 6294–6306, 2020, doi: 10.18653/V1/2020.ACL-MAIN.561.

[10]   M. Ghanem, A., Zaki, S., & Ahmed, "Deep Learning Models for Thematic Pattern Extraction in Quranic Texts" Journal of Artificial Intelligence," *Dep. Artif. Intell.*, 2020.