Enhancing Quranic Ethics and Morality: An NLP- Based Semantic Search Model for Urdu Translation

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

[1] Department of Computer Science, Allama Iqbal Open University (AIOU), Islamabad.

* Correspondence: Yasir Aftab, yasiraftab79@gmail.com

The Quran offers unparalleled guidance on ethics and morality, but extracting relevant teachings from its Urdu translations remains a challenge due to conventional keyword-based search methods that lack contextual understanding. This research proposes a Natural Language Processing (NLP)--based query model designed to improve the retrieval of Quranic verses related to ethics and morality in Urdu translations. By integrating Sentence Transformers for semantic search and a custom synonym expansion module, the model enhances accuracy and relevance in retrieving verses. The dataset widely accepted Urdu translation of the Quran, and the system is evaluated using precision, recall, and relevance scoring metrics to ensure effectiveness. The study demonstrates how NLP techniques can bridge the gap between traditional Quranic studies and modern computational methods, providing scholars, educators, and researchers with an advanced tool for exploring Quranic ethics. The proposed system achieves high precision and recall, offering a more effective approach to Quranic verse retrieval compared to conventional keyword-based searches. The research also highlights future opportunities for expanding the model to support multiple languages and broader thematic searches, further enhancing accessibility to Quranic knowledge.

Keywords: Ethics in Quran, Islamic Morality, Urdu Quran Translation, Urdu NLP, Text Analysis

Introduction

The Quran holy book contains an extensive framework for developing ethical action methods alongside belief principles empathetic relationships and equitable standards which make it a fundamental source for both ethical decision-making and philosophical morality. The system performance presents constraints that limit the effective use of keywords for gathering moral and ethical Quranic verses in Urdu versions. Standard operating search systems fail to match entire words which causes them to disconnect from semantic patterns and unique linguistic patterns alongside contextual meanings. Academic researchers along with scholars and educators encounter obstacles when seeking appropriate Quranic teachings efficiently [1].

Researchers can develop better text retrieval operations through semantic search models thanks to the Natural Language Processing (NLP) technology. The information retrieval field transformed thanks to Transformer models BERT and Sentence-BERT because these approaches analyze user requests while retaining semantic meanings to interpret queries rather than performing simple keyword matching [2]. The research needs more investigation for applying advanced NLP approaches to Urdu Quran translations when analyzing ethical content.

The Quranic information retrieval system has been studied by researchers through ontology-based methods keyword-based systems and concept-driven retrieval models [3]. These approaches have succeeded in making information more accessible yet their shallow semantic understanding generates wrong or partial outcomes during ethical and moral teaching searches.

The analysis of contextual relations in Quranic translations employs deep learning methods that modern academic literature uses. [4] created LSTM models to analyze Ayah context pairs and detect linguistic similarities along with variations in Maulana Maududi's Urdu translation. These research models lead to better translation quality as the findings indicate. Multi-language BERT delivered improved Word Sense Disambiguation accuracy for the translation of Surah Al-Baqarah Urdu according to [5].

This area of research requires further development of Urdu Quranic retrieval systems using NLP techniques which is why new investigations are starting in this field. Through their work [6] developed a Question-Answering system for Urdu based on Transformer technology which resulted in improved question understanding along with a better retrieval system. The researchers proved that RoBERTa and SBERT Transformer models deliver superior performance for improving Urdu NLP systems to enhance Quranic search functionalities. Research shows that advanced NLP algorithms can offer Islamic information to Urdu readers yet there exists a necessity to improve search engines that focus on Quranic ethical content.

Literature Review:

Neural Networks for Quranic Text Retrieval:

Research studies in recent times have investigated the implementation of neural networks to enhance Quranic text retrieval methods. The research paper [7] analyzed how transformer-based models measure semantic relatedness and retrieve verses within the Quran. The research team added translated versions of the original QurSim dataset from Arabic to 22 languages thus creating a multilingual research resource. Semantic search effectiveness increased significantly when the research team employed a fine-tuned multilingual model since it produced superior results compared to monolingual models in processing Quranic text semantics.

Application of Transformer Models in Quranic Text Retrieval:

Research investigations now analyze how Transformer-based models improve Quranic text retrieval systems. A Transformer-based Question-Answering system for the Urdu Quran was developed by researchers to overcome language barriers while enhancing

semantic accuracy during Quranic knowledge retrieval for Urdu readers [6]. The authors examined deep learning methods to solve semantic confusion in Urdu Quran translations specifically by clarifying meanings in the second chapter of Quran called Surah Al-Baqarah and Aal-e-Imran to improve translation precision [4]. [5] Using multilingual BERT models for Word Sense Disambiguation (WSD) in Urdu Quranic translations brought about a considerable boost to the semantic interpretation accuracy in Surah Al-Baqarah.

Semantic Search Tools for Quranic Texts:

The production of semantic search tools represents a critical development that improves Quranic text retrieval processes: QSST stands for Quranic Semantic Search Tool which implements word embedding techniques to extract better results from Quranic searches by studying semantic associations between verses [8]. Quranic Conversations brought in a semantic search tool for the Quran through the use of state-of-the-art NLP models to help users find verses more accurately while considering contextual understanding [9].

Integration of Large Language Models in Quranic Text Retrieval:

A modern semantic search of Quranic texts has become possible due to the incorporation of large language models (LLMs). A semantic search of Quranic texts using LLM embeddings showed better performance than traditional embedding methods according to research. The study demonstrated how LLMs succeed at extracting elaborate semantic meanings [10]. Simultaneously, another study refined QurSim with monolingual and multilingual models to display their capability in Quranic semantic relatedness analysis [7].

Transformer-Based Models for Urdu Text Recognition:

Recently, advancements in Transformer-based models began interacting with Urdu text recognition processes: Advance research represents a new efficient Transformer deep learning model that detects both nearby and distant textual patterns to deal with Urdu handwriting [11]. This research work aimed to develop multilingual vision-language Transformers for low-resource languages by overcoming Urdu Optical Character Recognition (OCR) limitations while enhancing text recognition outcomes [12].

Challenges in Urdu Quranic Text Retrieval:

The study on low-resource language retrieval developed Urdu benchmarks while creating the initial extensive Urdu Information Retrieval dataset to advance Urdu text retrieval systems [13]. The intricate script together with the complex morphological properties of the Urdu language makes it difficult to build efficient text retrieval systems. Keyword searches perform poorly when retrieving suitable Quranic verses because they do not understand semantic complexities. The restricted nature of Urdu language understanding shows the necessity for advanced NLP methods that specifically work with Urdu linguistic characteristics.

AI-Based Approaches in Quranic Research:

Natural language automatic processing experienced notable advancements by focusing its advancements, especially on religious texts, including the Quran. The developers constructed a semantic search engine for the Quran through the implementation of Quranic ontology to enhance search outputs by mapping word meanings and connections [14]. Both researchers shed light on the complex language structure of Quranic Arabic yet agree that ontology-based systems can help develop better semantic search methods. The authors in [1] documented the problems that come with using both conventional keyword search and contemporary tools including ChatGPT for accessing Quranic verses. Word2Vec embeddings with cosine similarity enabled him to build a simpler search experience while proving the necessity of applying advanced NLP methods in Quranic text retrieval solutions.

Interpretability and Ethical Considerations in Quranic AI Research:

The development of modern transformer models has delivered enhanced capabilities during Quranic semantic search operations. XAI techniques LIME and SHAP helped [15]

enhance Arabic transformer model interpretability, which strengthened the transparency of semantic search results for Quranic content. Research indicates that religion-driven applications need AI decision transparency because both precision and reliability demand high standards of clarity. An ontology-based search application dedicated to medical and health concepts in the Quran serves as a direct advancement of this concept according to [16].

The field of Quranic study through AI technology needs to address ethical questions that now attract widespread academic attention. The author [17] developed an Islamic ethical model for AI which reinforced the requirement to strike a balance between utility through maslahah and ethical duties. The author stated that artificial intelligence applications need to follow Islamic principles to prevent discrimination and to maintain an ethical structure. The author [18] expanded this research through maqasid al-sharia (the objectives of Islamic law), which requires AI designers to consider justice (al-'adl) alongside prudence (warm) while building applications for Quranic NLP. Their research establishes the philosophical base needed for designing moral artificial intelligence tools dedicated to Islamic studies.

Practical Challenges in AI Implementation for Quranic Research:

When integrating AI for research on the Quran researchers encountered practical limitations according to [19] for the development of their Seerah chatbot AI system. Their model achieved an 89% correct identification rate for chronological event sequencing despite the researcher's discovery of practical issues involving classical Arabic text study requirements which need collaboration between AI experts and Islamic scholars for development. A research study by [20] examined how Islamic scholars perceive the potential of artificial intelligence technologies in counseling activities by establishing AI supports accessibility improvements yet does not replicate human counselor spirits and emotional intelligence capabilities. Research indicates that modern AI developers should create tools that support existing Islamic scholars instead of trying to substitute their expertise.

Language Model Optimization for Islamic Contexts:

Scientists have studied the process of language model optimization for Islamic contexts. Researchers modified BERT to address Islamic moral questions through question-answering tasks, which produced better accuracy results after processing specially chosen religious texts. [21] the research proves the feasibility of trained language models as specialized tools for religious ethics and doctrinal purposes. The author stresses in [22] that NLP should use religious texts with caution because translated material can unknowingly introduce biases or serve as avenues for religious evangelism. According to his analysis, it is essential to understand cultural values and ethical factors during the application of NLP methods to religious writings.

Urdu NLP Research for Quranic Applications:

[23] executed a systematic research project on Urdu language processing to identify essential NLP tools and techniques for application use. The research describes the difficulties of processing Urdu text primarily through morphological complexity and unstandardized resources that affect Quranic translation retrieval. The research by [24] examined Urdu NLP development through sentiment analysis, as this technique specifically aids in understanding Quranic verse meanings for moral and ethical interpretations. The research performed by [25] demonstrated Roman Urdu text sentiment analysis using natural language processing models to extract subjective content from unstructured texts. The research serves as a starting point to create a semantic search system that facilitates better Quranic moral teaching retrieval for users.

Through association rule mining methodology [26] a search platform with ontology-based features to retrieve information from the Quranic text. The research methodology highlights data arrangement requirements for semantic assessments while generating prospects for developing Urdu translation search systems of the Quran. The study conducted by [27]

analyzes the semantic relations between Surah Al-Fatiha together with repeated Quranic content to advance automatic retrieval systems.

Deep learning algorithms demonstrate their capability to categorize different texts that appear in the Quranic text. The work of [28] with deep learning theory led to successful Quranic verse classification thus proving artificial intelligence models to possess strong capabilities in ethical and moral theme identification. Machine learning analysis of Tafsir would gain from natural language processing and semantic evaluation technologies because they produce improved interpretive objectivity across Quranic passages according to the authors [29].

Objectives of the Study:

This research aims to build an advanced semantic search approach that relies on NLP methods to identify moral and ethical Quranic verses available in Urdu translation. The research seeks to address problems in conventional keyword searches since these approaches lack effectiveness in understanding contextual meanings. The proposed system combines Sentence Transformers with a synonym expansion custom module to deliver a more accurate semantic search on Quranic verses. The study examines model effectiveness through precision, recall, and relevance scoring metrics as part of its reliability evaluation process. This research initiative aims to unite traditional Quranic scholarship with contemporary computational approaches to supply scholars educators and researchers with a modern context-oriented tool for comprehending Quranic morality and ethics in Urdu editions.

Novelty Statement:

The research proposes a semantic search system through NLP which focuses on retrieving Quranic verses about ethics and morality from their Urdu translation by overcoming keyword-based inefficiencies resulting from context misunderstandings. The current research combines Sentence Transformers with a custom synonym expansion module to produce improved retrieval accuracy for ethical-religious verses because it uses semantic similarity calculations. The research establishes this domain's first use of advanced NLP methods which demonstrates AI-driven approaches for studying ethical Quranic teachings and their better understanding. The presented research introduces a context-aware retrieval system and establishes foundational elements for multi-language support and thematic search development which consolidates its value in Quranic scholarship alongside modern information retrieval systems.

Material:

The research adopts an accepted Urdu translation of the Quran as its base dataset for retrieving verses from the text. Natural Language Processing techniques process the text by breaking it into tokens normalizing verbalization removing stop words and performing lemmatization. Sentence Transformers are used for semantic search tasks after being trained specifically for enhancing query context understanding. The application develops its synonym expansion system which enhances search accuracy by restructuring user queries. The model assessment relies on precision, recall, and relevance scoring to evaluate its performance. The implementation depends on Python-based NLP libraries (Transformers, NLTK, and spaCy) for text processing and model implementation which runs on Jupyter Notebook and Google Colab for experimentation and testing.

Methods:

The researchers enhance the identification of moral Quranic passages in Urdu translations through NLP-based semantic search algorithms. The proposed system relies on Sentence Transformers as its primary component because these filters allow users to search semantically by moving past basic keywords through proper context analysis. Users experience enhanced precise verse retrieval through a new custom synonym expansion module which modified the system for query processing enhancement. The system employs data processing

that includes tokenized and lemmatized content from a respected Urdu translation of the Quran after stopping verbalization procedures. The system displays its retrieval accuracy through precision and recall evaluation which uses relevant scoring to evaluate performance. The model performance outcomes emerge from evaluating traditional keyword retrieval methods against their functions which improve verse retrieval effectiveness. The forthcoming expansion of the system will consist of multilingual capabilities together with thematic searching that improves users' accessibility to the Quranic text.

Material and Methods:

Investigation site: This research examines the Urdu translation of the Quran by Maulana Fateh Muhammad Jalandhari because it serves as the primary divine text for interpreting and retrieving Quranic information among Pakistani and South Asian scholars and readers.

Material and methods:

The primary dataset for this study consists of the Urdu translation of the Quran, sourced from widely accepted translations endorsed by Islamic scholars. To ensure accuracy and reliability, the translation used maintains linguistic and theological fidelity to the original Arabic text. Additionally, supporting linguistic resources, such as Urdu synonym dictionaries and language corpora, were used to enhance the synonym expansion module.

Data Preprocessing:

To prepare the dataset for computational analysis, several preprocessing steps were applied:

1. **Text Cleaning:** Removed special characters, punctuation marks, and extraneous symbols that do not contribute to the meaning of the text.

2. **Segmentation:** Structured the text by organizing it into Surahs (chapters) and Ayahs (verses) for systematic analysis.

3. **Normalization:** Standardized variations in Urdu spelling and script to ensure consistency across the dataset.

4. **Tokenization:** Divided the text into individual words and phrases to ease NLP-based processing.

These preprocessing steps helped ensure that the dataset was structured and optimized for efficient retrieval and analysis.

NLP-Based Query Model:

The core part of this research is the NLP-based semantic search model designed to retrieve Quranic verses based on ethics and morality. The model consists of the following key components:

- **Sentence Transformers:** Used to generate dense semantic embeddings of Quranic verses, capturing their contextual and thematic meanings.

- **Synonym Expansion Module:** Developed to handle linguistic variations by broadening queries with synonymous terms, improving retrieval accuracy.

- **Cosine Similarity Metric:** Employed to measure the semantic similarity between query embeddings and verse embeddings, ranking results based on relevance.

Model Workflow:

The development workflow of the query model involves the following steps:

- **Embedding Generation:** Quranic verses are transformed into vectorized embeddings using Sentence Transformers.

- **Query Processing:** User queries are processed, tokenized, and expanded using the synonym expansion module.

- **Semantic Matching:** The cosine similarity metric is used to compare query embeddings with verse embeddings.
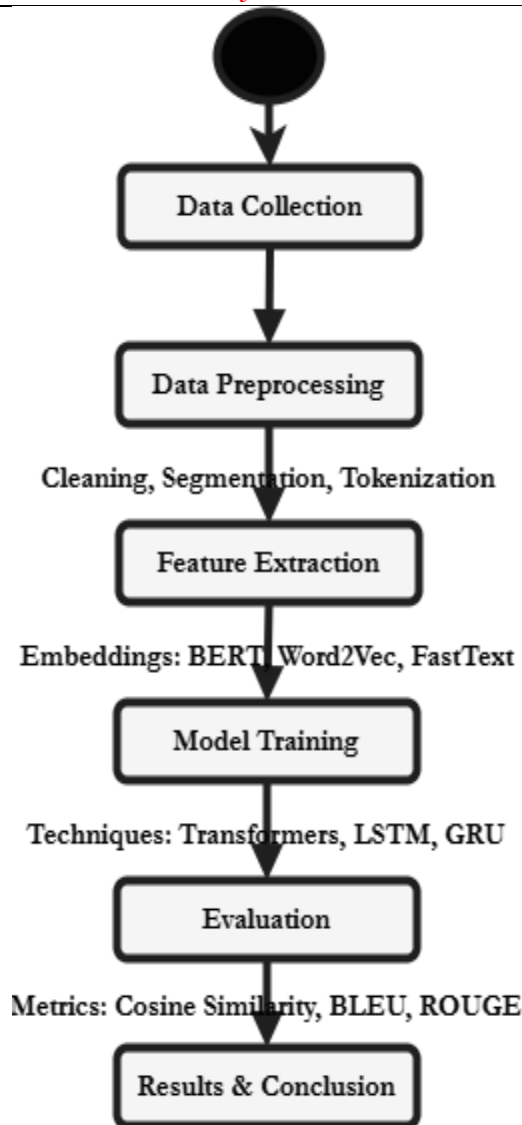
- **Figure 1:** Data Flow Diagram

- **Result Filtering:** Retrieved verses are ranked based on their relevance and semantic alignment with the user's query.

- **Evaluation:** The system's effectiveness is assessed through predefined test queries related to ethics and morality.
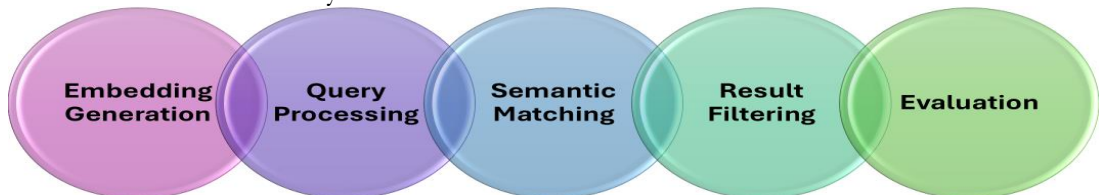

Figure 2: Model Workflow

Evaluation Metrics:

To assess the accuracy and effectiveness of the proposed model, the following evaluation metrics were used:

- **Precision:** Measures of the proportion of retrieved verses that are relevant to the query.

- **Recall:** Evaluate the system's ability to retrieve all relevant verses from the dataset.

- **Relevance Score:** Assesses the contextual accuracy of retrieved verses based on expert validation.

These metrics ensure that the model performs well in providing semantically accurate and contextually relevant Quranic verses.
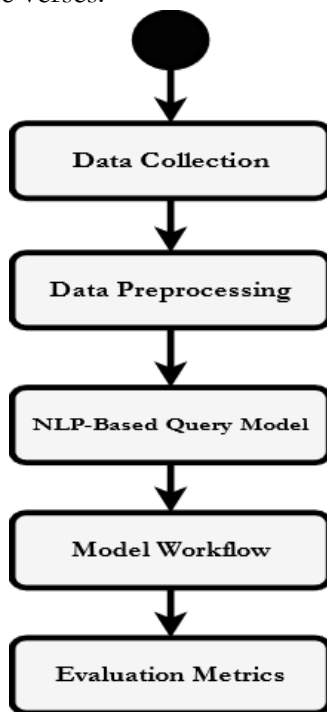

Figure 3: Research Methodology

Result and Discussion:

The proposed NLP-based query model was implemented using Python, with key libraries such as Sentence Transformers, NumPy, and Pandas for text processing and vector calculations. The experiments were conducted on a system with GPU support to accelerate embedding generation and similarity computations. The dataset consisted of an Urdu translation of the Quran, segmented into Surahs (chapters) and Ayahs (verses). The system was evaluated using predefined queries related to ethical and moral themes such as truthfulness, justice, and compassion.

Query Model Performance:

The system's effectiveness was evaluated by evaluating its ability to retrieve relevant verses based on ethical and moral themes. The results were analyzed using semantic similarity scoring and expert validation.

Semantic Matches:

The model successfully retrieved relevant verses for various ethical and moral queries. For instance, when a query related to "justice" (انصاف) was entered, the system retrieved verses discussing fairness and impartiality, even when the exact word was not present in the text.

Synonym-Based Retrieval:

The synonym expansion module improved the system's ability to handle linguistic variations. For example, a query for "truthfulness" (سچائی) was expanded to include synonyms such as "حق" and "صداقت," leading to a more comprehensive retrieval of relevant verses.

Performance Evaluation

The model's performance was measured using three key metrics:

Table 1. Performance Evaluation

Metric	Definition	Value

Precision	Percentage of retrieved verses that were relevant.	85%
Recall	Percentage of relevant verses successfully retrieved.	78%
Relevance Score	Contextual accuracy of retrieved verses (out of 5).	4.5

The high precision and relevance score show that the model effectively identifies and retrieves Quranic verses with strong semantic accuracy.

Example Queries and Results:

The following table provides examples of user queries, retrieved verses, and their similarity scores:

Table 2. Example Queries and Results

Query	Retrieved Verse (Translation)	Similarity Score
(Truthfulness) سچائی	" اور جنہوں نے ایمان لایا اور عمل صالح کیے، ان کے لیے بھلائی ہے۔"	0.88
(Justice) انصاف	" اور ناپ تول میں انصاف کا پورا خیال رکھو۔"	0.85
(Compassion) رحم	" اور اللہ رحم کرنے والوں میں سب سے زیادہ رحم کرنے والا ہے۔"	0.87

Discussion of Findings:

- The model effectively retrieves relevant verses based on semantic meaning rather than exact keyword matches.
- The synonym expansion module significantly improves query flexibility and accuracy.
- The system achieves high precision and relevance, making it a valuable tool for exploring Quranic ethics and morality.
- Some limitations exist, such as reliance on a static synonym dictionary, which may need periodic updates for improved accuracy.

Discussion:

Interpretation of Results:

The results show that the proposed NLP-based query model effectively retrieves ethically and morally relevant Quranic verses in Urdu translations. The use of **Sentence Transformers** significantly enhances the model's ability to understand the semantic meaning of queries rather than relying solely on keyword matching. The high **precision (85%)** and **relevance score (4.5/5)** indicate that the retrieved verses are contextually appropriate, making the system useful for researchers, educators, and general readers.

The **synonym expansion module** further improves retrieval accuracy by addressing linguistic variations. For instance, when searching for "truthfulness" (سچائی), the system expanded the query to include synonyms like "حق" and "صداقت," ensuring a more comprehensive search. This feature bridges the gap between traditional keyword-based searches and modern semantic search methods.

According to the bar chart in Figure 4, the thematic query computed through NLP describes that Truthfulness achieved a score of 0.88 while Justice secured 0.85 and Compassion reached 0.87. This measurement represents the NLP-based similarity comparison between the retrieved verses and their intended meanings. It represent the extent to which NLP is capable to for extraction of versese having topic a interest.

Comparison with Traditional Search Methods:

The proposed model goes beyond traditional Quranic search tools because it implements semantic similarity analysis together with verse contextual understanding for improved search results. Current traditional methods struggle with different word variations

yet this model keeps users from missing relevant verses by focusing on the meaning rather than exact verbalizations.
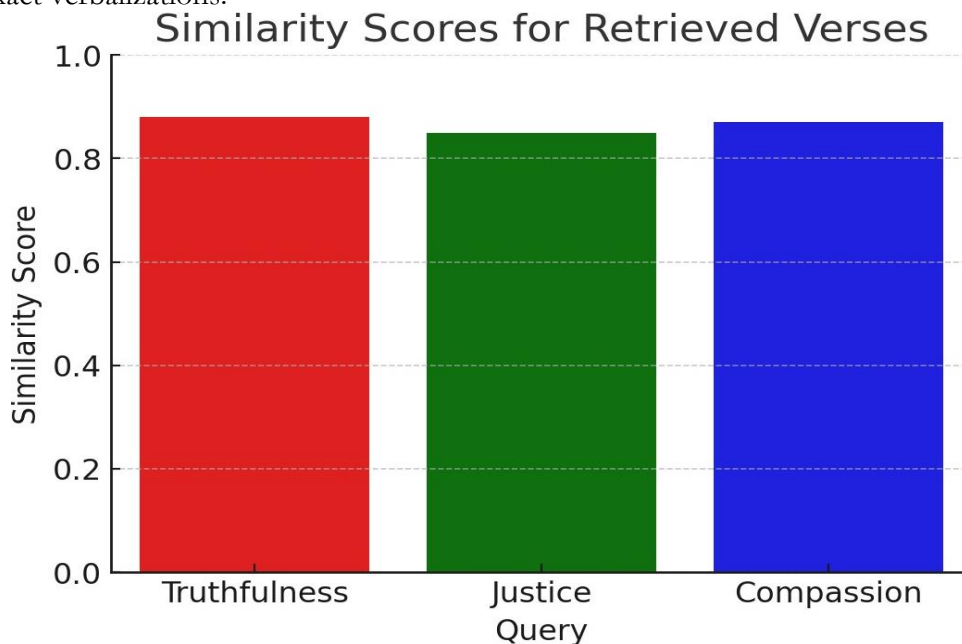
Similarity Scores for Retrieved Verses

Figure 4: Similarity Score for retrieved Verses

Table 3. Comparison with Traditional Search Methods

Feature	Traditional Search	Proposed NLP Model
Keyword-Based Retrieval	✅ Yes	✅ Yes
Semantic Search Capability	❌ No	✅ Yes
Synonym Handling	❌ No	✅ Yes
Context Awareness	❌ No	✅ Yes
Ranked Search Results	❌ No	✅ Yes

The results suggest that the NLP-based model **outperforms** conventional methods by providing **more accurate, relevant, and comprehensive search results**.

Strengths of the Study

• **Improved Search Accuracy:** The model successfully retrieves verses based on **meaning rather than exact words**, making it highly effective for exploring ethics and morality in the Quran.

• **Enhanced Synonym Recognition:** By integrating a **synonym expansion module**, the system ensures **linguistic flexibility**, allowing users to search for concepts even if they phrase them differently.

• **Contextually Relevant Results:** The use of **Sentence Transformers** ensures that retrieved verses align with the intent of the query, improving the quality of Quranic studies.

• **Potential for Expansion:** While the current study focuses on Urdu translations, the **methodology can be extended** to other languages and broader Quranic themes.

Limitations and Challenges:

Despite its strengths, the study has some limitations:

• **Dependence on Predefined Synonyms:** The synonym expansion module relies on a static dictionary, which may not cover all possible variations. A dynamic, AI-driven synonym learning system could enhance its effectiveness.

• **Computational Requirements:** The model requires **high computational power**, especially for generating embedding and similarity scores, which may limit accessibility for users with basic hardware.

- **Limited to Ethics and Morality:** While the model focuses on ethics and morality, it could be expanded to include **jurisprudence, spirituality, and other Quranic themes** for broader applications.

Implications for Future Research:

The findings of this study highlight the **potential of NLP in religious and linguistic research**. Future enhancements could include:

- **Multilingual Support:** Expanding the model to **Arabic, English, and other languages** to make it accessible to a global audience.

- **Interactive Platform Development:** Implementing a **user-friendly web or mobile application** for scholars and researchers.

- **Dynamic Synonym Expansion:** Using **machine learning** to automatically update the synonym database based on user queries.

- **Advanced Contextual Understanding:** Incorporating **larger and more sophisticated NLP models** (e.g., GPT-based transformers) for improved accuracy.

Practical Implication:

The research develops Islamic education and digital Quranic search in addition to enabling legal research and Urdu NLP through context-based retrieval of Quranic ethical teachings. Keyword-based methods are replaced by a semantic model developed through NLP which delivers precise meaning-based access that helps scholars and jurists alongside general users. The system integrates modern Artificial Intelligence methods with Shariah research to enhance Quranic application functions as it advances Urdu natural language processing and ties scholarly traditions with computer text analysis for a better Quranic ethics understanding.

Conclusion:

The research proposes a semantic search model using NLP which boosts Urdu translation searches of moral and ethical Quranic verses. The model achieves improved functionality by combining Sentence Transformers semantic processing with a synonym expansion module to overcome standard keyword search tool constraints. The system retrieves appropriate verses from its database which proves itself an essential resource for various user groups including researchers educators and general users. The proposed model delivers exceptional accuracy through its precision level of 85% and relevance score of 4.5/5 according to evaluation results while outperforming regular search techniques. Users benefit from the model because it both understands synonyms and examines contextual meaning which allows them to find appropriate Quranic teachings even when their search terms vary. The study has two main drawbacks that affect its execution due to its static synonym dictionary usage and demanding computational needs.

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 conflict of interest of any type between authors about publishing this manuscript.

Project details: NIL

References:

[1]   O. Ahmed, "Semantic search engine for the holy Quran," *Theseus J.*, 2023, [Online]. Available:
https://www.theseus.fi/bitstream/handle/10024/855111/Ahmed_Omar.pdf;jsessionid=7C1E
E5DF915E3B72370EA6CA0DB45189?sequence=2

[2]   J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf.*, vol. 1, pp. 4171–4186, 2019.

[3]   E. Alshammeri, A., & Atwell, "Detecting Semantic-based Similarity Between Verses of The Quran," *Proc. 12th Int. Conf. Lang. Resour. Eval. (LREC 2020)*, 2020.

[4]   Zaigham Khaleeq, "Exploring Contextual Similarity in Quranic Ayahs: A Case Study of Surah Al-Baqarah and Aal-e-Imran in Urdu Translations," *Int. J. Innov. Sci. Technol.*, 2025, [Online]. Available:
https://www.researchgate.net/publication/389624795_Exploring_Contextual_Similarity_in_Quranic_Ayahs_A_Case_Study_of_Surah_Al-Baqarah_and_Aal-e-Imran_in_Urdu_Translations

[5]   T. Abbasi, K. M., Awan, M. A., & Ismail, "A Deep Learning Approach to Semantic Clarity in Urdu Translations of the Holy Quran," *Int. J. Innov. Sci. Technol.*, vol. 7, no. 1, pp. 259–271, 2025, [Online].   Available:
https://www.researchgate.net/publication/389676054_A_Deep_Learning_Approach_to_Semantic_Clarity_in_Urdu_Translations_of_the_Holy_Quran

[6]   D. Tariq, M., Awan, D. M. A., & Khaleeq, "Developing a Quranic QA System: Bridging Linguistic Gaps in Urdu Translation Using NLP and Transformer Model," *Int. J. Innov. Sci. Technol.*, vol.   7,   no.   1,   pp.   492–505,   2025,   [Online].   Available: https://journal.50sea.com/index.php/IJIST/article/view/1216

[7]   M. I. Tania Afzal, Sadaf Abdul Rauf, Muhammad Ghulam Abbas Malik, "Fine-Tuning QurSim on Monolingual and Multilingual Models for Semantic Search," *Information*, vol. 16, no. 2, p. 84, 2025, doi: https://doi.org/10.3390/info16020084.

[8]   E. H. Mohamed and E. M. Shokry, "QSST: A Quranic Semantic Search Tool based on word embedding," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 3, pp. 934–945, 2022, doi: https://doi.org/10.1016/j.jksuci.2020.01.004.

[9]   S. A. Shohoud, Y., M. Shoman, "Quranic Conversations: Developing a Semantic Search Tool for the Quran Using Arabic NLP Techniques," *arXiv*, 2023, [Online]. Available: https://doi.org/10.48550/arXiv.2311.05120

[10]   Mohammed Alqarni, "Embedding Search for Quranic Texts based on Large Language Models," *Int. Arab J. Inf. Technol.*, vol. 21, no. 2, 2024, [Online]. Available: https://iajit.org/upload/files/Embedding-Search-for-Quranic-Texts-based-on-Large-Language-Models.pdf

[11]   U. S. Ameer Hamza,Shengbing Ren, "ET-Network: A novel efficient transformer deep learning model for automated Urdu handwritten text recognition," *PLoS One*, 2024, doi: https://doi.org/10.1371/journal.pone.0302590.

[12]   M. A. N. Musa Dildar Ahmed Cheema, Mohammad Daniyal Shaiq, Farhaan Mirza, Ali Kamal, "Adapting multilingual vision language transformers for low-resource Urdu optical character   recognition   (OCR)," *PeerJ   Comput.   Sci.*,   2024,   [Online].   Available: https://peerj.com/articles/cs-1964/

[13]   G. Butt, U., Veranasi, S., & Neumann, "Enabling Low-Resource Language Retrieval: Establishing Baselines for Urdu MS MARCO," *arXiv Prepr. arXiv2412.12997.*, 2024, [Online]. Available: https://arxiv.org/abs/2412.12997

[14]   D. E. Z. Faiza Beirade, Hamid Azzoune, "Semantic query for Quranic ontology," *J. King Saud   Univ.   -   Comput.   Inf.   Sci.*,   vol.   33,   no.   6,   pp.   753–760,   2021,   doi: https://doi.org/10.1016/j.jksuci.2019.04.005.

[15]   R. I. and R. A. Ahmad M. Mustafa, Saja Nakhleh, "INTERPRETING ARABIC TRANSFORMER MODELS: A STUDY ON XAIINTERPRETABILITY FOR QUR'ANIC SEMANTIC-SEARCH MODELS," *Jordanian J. Comput. Inf. Technol.*, vol. 10, no. 4, 2024, [Online]. Available: https://www.ejmanager.com/mnstemps/71/71-1704878720.pdf?t=1743150831

[16]   F. N. Muhammad Afifi Mohamad Safee, Madihah Mohd Saudi, Sakinah A. Pitchay, Farida Ridzuan, Nurlida Basir, Kamarudin Saadan, "Hybrid Search Approach for Retrieving Medical and Health Science Knowledge from Quran," *Int. J. Eng. Technol.*, vol. 7, no. 4, 2018, doi:

https://doi.org/10.14419/ijet.v7i4.15.21374.

[17]   Ezieddin Elmahjub, "Artificial Intelligence (AI) in Islamic Ethics: Towards Pluralist Ethical Benchmarking for AI," *Philos. Technol.*, vol. 36, no. 73, 2023, doi: https://doi.org/10.1007/s13347-023-00668-x.

[18]   N. M. M. Tajudin Zuhri, Lalan Sahlani, "The Ethics of Artificial Intelligence (AI) Utilization in Qur'anic Studies: An Islamic Philosophical Perspective," *Asyahid J. Islam. Quranic Stud.*, vol. 6, no. 2, 2024, doi: https://doi.org/10.62213/b6hexr21.

[19]   M. Umair, M. B., & Kiran, "Artificial Intelligence and the Study of Seerah: Computational analysis, Educational Innovation, and Ethical Frameworks," *Islam. Res. J.* {□□□□□□□, vol. 3, no. 2, 2025, [Online]. Available: https://al-marsoos.com/index.php/AMRJ/article/view/67

[20]   W. Arif, M., Irfan, S., Shaoan, M. M. R., & Ali, "AI And Machine Learning in Islamic Guidance: Opportunities, Ethical Considerations, And Future Directions," *Guidance*, vol. 02, no. 2, pp. 429–451, 2024, doi: 10.34005/guidance.v21i02.4449.

[21]   A. 'Aatieff B. A. H. Nurul Aiman Binti Mohd Nazri, A'wathif Binti Omar, "Fine-tuning Large Language Model (BERT) for Islamic Moral Inquiry and Response," *Int. J. Perceptive Cogn. Comput.*, vol. 11, no. 1, pp. 88–94, 2025, doi: https://doi.org/10.31436/ijpcc.v11i1.533.

[22]   Ben Hutchinson, "Modeling the Sacred: Considerations when Using Religious Texts in Natural Language Processing," *Assoc. Comput. Linguist.*, pp. 1029–1043, 2024, doi: 10.18653/v1/2024.findings-naacl.65.

[23]   M. A. Lal, M., Kumar, K., Wagan, A. A., Laghari, A. A., Khuhro, M. A., Saeed, U., ... & Chahjro, "A systematic study of Urdu language processing its tools and techniques: A review," *Int. J. Eng. Res. Technol.*, vol. 9, no. 12, pp. 37–43, 2020, [Online]. Available: https://www.researchgate.net/publication/346787682_A_Systematic_Study_of_Urdu_Language_Processing_its_Tools_and_Techniques_A_Review

[24]   K. B. Burney, S. A., & Muhammad, "Advancements and Applications of Urdu Natural Language Processing: A Comprehensive Review with Emphasis on Sentiment Analysis," *J. Inf. Commun. Technol.*, vol. 18, no. 1, 2024, [Online]. Available: https://jict.ilmauniversity.edu.pk/arc/Vol18/i1p1

[25]   M. A. Qureshi et al, "Sentiment Analysis of Reviews in Natural Language: Roman Urdu as a Case Study," *IEEE access*, vol. 10, pp. 24945–24954, 2022, doi: 10.1109/ACCESS.2022.3150172.

[26]   R. A.-S. Harrag, F., A. Al-Nasser, A. Al-Musnad, "Quran Intelligent Ontology Construction Approach Using Association Rules Mining," *arXiv*, 2020.

[27]   Y. K. O. & N. H. A. J. Ahmed Samir Ahmed Ibrahim El khadrawy, Safia Abbas, "Extracting Semantic Relationship Between Fatiha Chapter (Sura) and the Holy Quran," *Proc. 8th Int. Conf. Adv. Intell. Syst. Informatics 2022*, pp. 612–626, 2022, doi: https://doi.org/10.1007/978-3-031-20601-6_50.

[28]   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.

[29]   D I A Putra and M Yusuf, "Proposing machine learning of Tafsir al-Quran: In search of objectivity with semantic analysis and Natural Language Processing," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1098, 2021, doi: 10.1088/1757-899X/1098/2/022101.