RESEARCH & INNOVATION DIVISION
IJIST

A Deep Learning Approach to Semantic Clarity in Urdu Translations of the Holy Quran

Kashif Masood Abbasi[1], Dr. Muhammad Arshad Awan[1],Tehmima Ismail[1]
[1]Department of Computer Science, Allama Iqbal Open University (AIOU), Islamabad.
* Correspondence: Kashif Masood Abbasi, kashif1870@gmail.com

Citation| Abbasi. K. M, Awan. M. A, Ismail. T, "A Deep Learning Approach to Semantic Clarity in Urdu Translations of the Holy Quran", IJIST, Vol. 7 Issue. 1 pp 259-271, Feb 2025
DOI | https://doi.org/10.33411/ijist/202571259271
Received| Jan 03, 2025 **Revised|** Jan 28, 2025 **Accepted|** Feb 04, 2025 **Published|** Feb 06, 2025.

The Holy Quran holds profound significance from both religious and linguistic perspectives yet its Urdu translations face difficulties in preserving the original meaning because of ambiguous words which create interpretation challenges for speakers and listeners. This research tackles translation ambiguity in the Urdu translations of the Holy Quran authored by Maulana Abul A'ala Maududi and Fateh Muhammad Jalandhry by applying Word Sense Disambiguation methods with deep learning algorithms. A model based on multilingual BERT identifies ambiguous word senses for Surah Al-Baqarah in particular. The dataset features Surah Al-Baqarah's complete Urdu translation together with a Sense Inventory that contains 3 to 8 senses for 50 frequently used Urdu ambiguous words which are collected from GitHub repository. Sequence classification frameworks within BERT receive contextual embeddings during fine-tuning. The evaluation framework includes the determination of F1 scores alongside confusion matrix analysis and classification report assessment. The model achieved an F1-score of 0.82 when identifying the most frequent sense while reaching an average F1-score of 0.62 across eight predefined sense labels. A sense prediction system functions to improve word sense matching thereby leading to more precise translations. The proposed research makes significant contributions to computational linguistics and Quranic studies by delivering an expandable method that solves word sense ambiguity while offering important insights to help translators and scholars improve their understanding of how context affects meaning within translated texts.

Keywords: Word Sense Disambiguation (WSD); Urdu Quran Translation; Multilingual BERT; Deep Learning in Linguistics; Natural Language Processing.

Introduction:

The Quran, revered as the holy scripture of Islam, holds immense credibility among over 1.8 billion Muslims worldwide. Therefore, understanding it in its true essence is essential for both religious practice and scholarly study. However, the translation of the Quran in Urdu faces the challenge of word sense ambiguity; as some words have multiple meanings. Urdu, like many languages, contains numerous words with several meanings depending on the context of their use, which might confuse the readers. Traditional translation approaches rely on linguistic expertise, but they often struggle to resolve ambiguities systematically, leading to variations in meaning across different translations.

Word Sense Disambiguation (WSD) is a technique in Natural Language Processing (NLP) where the real intention of a specific word is identified based on the surrounding context. Recent advancements in NLP and Deep Learning have provided powerful tools for addressing lexical ambiguity through automated WSD techniques. In particular, transformer-based models such as BERT have demonstrated remarkable success in contextual word understanding. However, limited research has been conducted on applying these models to the Urdu translation of the Holy Quran. This study aims to bridge this gap by leveraging a multilingual BERT model to enhance clarity in Urdu translations through WSD. By focusing on ambiguous words in Surah Al-Baqarah, this research seeks to develop a computational approach that predicts the most contextually appropriate sense of a word, thereby improving translation consistency and interpretability.

The scope of this study is confined to two widely recognized Urdu translations: Maulana Abul A'ala Maududi's translation and Fateh Muhammad Jalandhry's translation. A manually curated sense inventory, derived from an online source, is used to provide reference meanings for ambiguous words. The model is trained and evaluated on a dataset containing occurrences of these words in different contexts within Surah Al-Baqarah. While this research demonstrates the potential of deep learning in Quranic translation studies, future work may explore broader datasets, additional translations, and advanced techniques such as data augmentation to further refine the model's performance. By integrating deep learning with linguistic analysis, this study contributes to the ongoing efforts to enhance the accessibility and accuracy of Quranic translations. The findings have implications for scholars, translators, and computational linguists working on religious texts, providing a novel framework for disambiguating complex linguistic structures in Urdu.

The remainder of this paper is organized as follows: The introductory portion includes the study's objectives, novelty statement, literature review, and research gap. The Material and Methods section discusses the entire research technique. The study's findings are then thoroughly explained in the Results and Discussions section.

Objectives of the Study:

The specific objectives of the study are to develop a WSD model that identifies and interprets the meanings of polysemantic words in Surah Al-Baqarah of the Holy Quran translated into Urdu by Maulana Abul A'ala Maududi and Fateh Muhammad Jalandhry and to evaluate the model's performance using the accuracy and F1-score. The precise aim of the current research is to enhance the interpretative quality of Urdu translations of the Holy Quran, which must stay as close to the Arabic original while at the same time being more comprehensible to the Urdu readers. This work not only provides a research contribution to the field of Quranic studies but also provides a starting stone to solve many text translation and interpretation problems based on advanced natural language processing procedures.

Novelty Statement: This research presents a new approach to improving the clarity of the Urdu translation of the Holy Quran by mitigating the word sense ambiguity using deep

learning methods. This work, therefore, uses a multilingual BERT model to eliminate the semantic ambiguity of words in the Urdu translations of Surah Al-Baqarah done by Maulana Abul A'ala Maududi and Fateh Muhammad Jalandhry. The work also presents a carefully designed dataset consisting of 3–8 senses for ambiguous Urdu words and employs contextual embeddings for sense differentiation. In other words, the study establishes a benchmark of WSD for religious text in the Urdu language. Consequently, the results enrich the theoretical background of computational linguistics and Quranic studies, providing scholars and translators with a practical solution for increasing the efficiency of sacred texts translation, as well as a helpful tool to maintain contextual integrity for all specialists interested in Quran translation and interpretation.

Literature Review:

WSD is closely related to NLP, especially its sub-discipline that addresses the field of polysemy, according to which different words in context can have various meanings. This becomes more critical when the word in question, such as those in the Urdu language, has multiple connotations attached to it.

The enhanced WSD model from BERT is comparatively more accurate than the benchmark algorithms and contributes towards highlighting the role of WSD in the improvement of many downstream tasks such as sentiment analysis [1].A novel unsupervised graph-based algorithm for Hindi WSD was used in a study[2] to discuss the effectiveness of this approach for WSD. A study[3] introduced the Urdu Natural Language Toolkit (UNLT), which addressed the lack of an NLTK equivalent for languages with little textual resource provision like Urdu. Another study[4] involved applying, evaluating, and comparing deep learning techniques for Urdu WSD tasks using specific neural network architectures. An approach [5]for WSD was introduced to enhance the accuracy of the Chinese WSD by establishing the integrality of the present RegNet with Efficient Channel Attention. Recent WSD algorithms that have advanced the field and contributed to state-of-the-art solutions were discussed in detail in a study [6]. A study [7]investigated the differences between pre-trained and custom-trained word embedding models for word sense disambiguation. A recent study[8] focuses on enhancing Word Sense Disambiguation (WSD) by leveraging gloss (sense definitions) within supervised neural systems. It proposes three BERT-based models that utilize context-gloss pairs and fine-tune BERT on the SemCor3.0 corpus. The results demonstrate that the proposed approach outperforms state-of-the-art systems on several English WSD benchmark datasets.

Research gap:

Numerous research studies focus on Word Sense Disambiguation (WSD) for widely spoken languages such as English, Chinese, and Arabic, but limited work has been done on Urdu translations of the Quran. Previous approaches to translating the Quran into Urdu face several challenges, including the limited availability of WSD systems trained on Quranic corpora, the complexity of Urdu's rich morphological structure, which leads to multiple word meanings, the presence of idioms and scriptural styles that differ from modern usage, and the influence of human judgment in translation. When discussing Quranic translations, where there are a multitude of ambiguous words, WSD is revealed as a very significant stage needed to clarify these ambiguities. WSD entails identifying the relevance of a certain meaning out of the several meanings of the word under consideration. However, challenges remain, such as the time required to retrieve data or access databases. Advancements in NLP and deep learning models, such as BERT, which utilizes contextual embeddings, offer promising solutions to address these issues. This research utilizes these advancements to try to solve the problem of obscurity in the Urdu translations of the holy Quran with special reference to Surah Al-Baqarah in an effort to improve the accuracy of the translations.

This research therefore seeks to fill these gaps by implementing deep learning

techniques for the construction of a WSD system designed specifically for disambiguating words in well-known Urdu translations of the holy Quran, especially the Surah Al-Baqarah.

Material and Methods:

The primary data for this study include Urdu translations of Surah Al-Baqarah and a sense inventory containing 3 to 8 possible senses of 50 frequently used Urdu ambiguous words. From these translations, an annotated dataset required for training and performance evaluation has been developed and provided in correspondence to the sense inventory. The operational technique encompasses a supervised deep-learning approach with linguistic analysis. For the WSD context, the mBERT, fine-tuned in many languages including Urdu, was employed.

Methodology:

This section overviews the method used to investigate the phenomenon of the difficulty of Word Sense Disambiguation (WSD) in Urdu translations of the Holy Quran. First, some existing WSD techniques were explained. Then, the methodology used in this research is described.

WSD techniques can be categorized into three primary approaches: supervised learning methods, unsupervised learning methods, and knowledge-based methods. Furthermore, hybrid approaches are also employed in which some features of these methods are combined to deal with the problems of WSD effectively.

Supervised Learning:

Supervised WSD approaches, as discussed in a study [9], require training data with each word instance assigned to its right sense. These methods apply machine learning to build models that learn contextual representations and semantic associations. The first step is feature extraction, where feature values are extracted; the next step is model training, where a machine learning model is built, and the final step is sense prediction, where new senses are predicted. Some notable supervised learning techniques include the Naïve Bayes Classifier, Support Vector Machines (SVM), and Maximum Entropy Model.

Unsupervised Learning:

Unsupervised learning methods operate without the need for labeled data, making them highly scalable and particularly useful for languages or domains with limited resources[10]. These approaches rely on distributional semantics, whereby words having similar meanings are often used in similar contexts. Some prominent unsupervised methods are: Clustering and Latent Semantic Analysis (LSA).

Knowledge-based Methods:

Another type of WSD technique is the knowledge-based method, which utilizes external knowledge sources like dictionaries, thesauri, or ontologies in identifying the correct sense of the identified word. A study [11] describes the use of knowledge-based methods. These methods use both the words' meanings and the context in which they are used to find the correct meaning. Key techniques include the Lesk Algorithm and WordNet.

Hybrid Approach:

The integration of WSD strategies in the Hybrid approach entails the enhancement of various techniques such as; supervised methods, unsupervised methods, and knowledge-based methods. These approaches take advantage of having a labeled dataset for supervised learning, while at the same time embracing the benefits of the unsupervised methods where they search for structures and patterns in the unlabeled data, and also the knowledge-based methods where external knowledge brings in semantic information for use in an information retrieval process. The hybrid approach is suggested in a study [12] for Arabic WSD for information retrieval. Figure 1 illustrates the flow of the study.

Data Collection and Pre-processing: As the first activity in preparation, the datasets and

sense inventory that would be used in the training and evaluation of the WSD model were obtained. To acquire the Urdu translations of Surah Al-Baqarah in this research, data was obtained from tanzil.net [13], a widely available and authoritative source for translations of the Quran in Urdu and other languages.

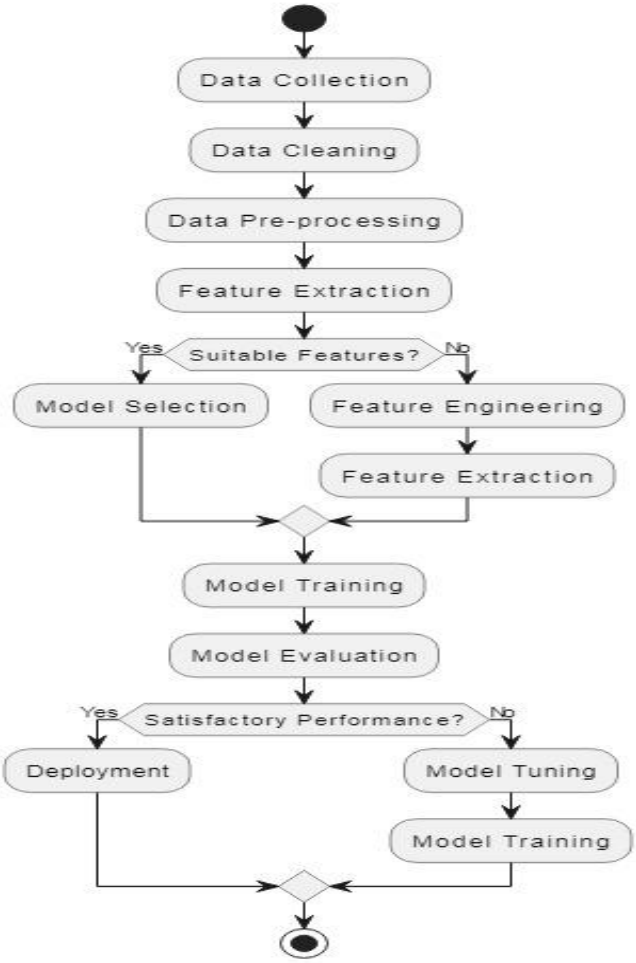

Figure 1. Flow of study

Two well-known and widely used Urdu translations of the Holy Quran by Maulana Abul A'ala Maududi and Fateh Muhammad Jalandhry were selected. These translations were chosen based on their linguistic richness, theological significance, and widespread acceptance. Maududi's translation is known for its explanatory nature, making it more detailed but potentially introducing interpretative ambiguity. Jalandhry's translation is more literal, staying closer to the original Arabic structure, making it useful for comparison in disambiguation tasks. Both are widely recognized and accepted among scholars, ensuring credibility and practical relevance for Urdu-speaking readers. The sense inventory used for the Word Sense Disambiguation (WSD) task for Urdu translation was collected from the web [14] and is freely available and contains 50 ambiguous words that are often used in the Urdu language. This sense inventory was created in a study [15] and the senses for ambiguous words were taken from a popular Urdu Lughat dictionary. Urdu Lughat is a comprehensive Urdu dictionary created by the Dictionary Board in Karachi, Pakistan, and is freely available for research purposes through an online interface. The dictionary contains approximately 120,000 unique words with multiple senses, synonyms, glosses, and descriptive examples. Three resources IndoWordNet, CLE UrduWordNet, and Urdu Lughat were considered to create the sense inventory. Manual inspection highlights the fact that the most suitable resource for the generation of sense inventory was Urdu Lughat. Every word was linked to 3 to 8 different

meanings. To create a training dataset, the collection of natural language sentences is provided together with the ambiguous word and correct sense marked from the inventory. Thus, the identification of unique sentences to be used in a separate testing dataset was also conducted. To optimize context consideration in text, ambiguous words in the sentences were encircled with <target> and </target> to directly point at it.

Programming Languages and Libraries:

The study is performed employing the Python language, which is popular in processing text data and machine learning. The WSD model implementation used various libraries to make processing and model training and evaluation tasks more efficient in this research. 'Pandas' served as the data manipulation tool for Excel file management and together with 'Matplotlib' the project created visualization elements including loss curves and confusion matrix heatmaps. The deep learning framework 'PyTorch' enabled model development and training while serving as the platform to fine-tune BERT-based models through custom data pipeline functionality. Through the 'Transformers' library users gained access to BERT pre-trained models alongside the required tokenizers to utilize context-based embeddings. The model evaluation metrics including confusion matrices and classification reports were processed with 'Scikit-learn' while 'Openpyxl' allowed smooth reading of Excel files stored in '.xlsx' format. A group of libraries functioned together to enable the full development of the proposed WSD system.

Feature Extraction:

The sense inventory used in this research was obtained from an online source and originally contained 50 ambiguous words, each with 3 to 8 possible senses. However, during the evaluation phase, it was observed that 16 of these words either did not appear or occurred too infrequently in the test dataset. To ensure meaningful learning and evaluation, the sense inventory was refined to include only the 34 words that had sufficient representation in the test data. This adjustment improved model reliability by focusing on words with enough contextual occurrences for effective sense disambiguation. After these adjustments, the final sense inventory consisted of 34 ambiguous words, ranging from 3 to 8 different senses/meanings. These senses represent the different possible meanings of each word and are essential for the WSD task, where the goal is to identify the most contextually appropriate sense. Thus, by reducing the inventory to only those words and senses that are most appropriate and encountered more often, the model became more efficient in terms of computation and achieved higher accuracy in the task of word sense disambiguation in the Urdu translations of the Holy Quran.

Model Selection:

The research employed BERT (Bidirectional Encoder Representations from Transformers) for its ability to generate context-aware embeddings. The "Bert-base-multilingual-cased" model was chosen as it optimizes for the uniqueness of Urdu. The tokenizer was set to convert sentences to input IDs and attention masks that do not exceed 128 in length. For the fine-tuning of the BERT model, sequence classification was used, and the number of output labels was set equal to the maximum number of the identified senses in the inventory (8). Table 1 given below shows the hyperparameters that were selected to train the model.

Table 1: Hyperparameter Selection

Hyperparameter	Value
Learning Rate	2×10^{-5}
Batch Size	16
Number of Epochs	9

Hyperparameter	Value
Optimizer	AdamW
Dropout Rate	0.1
Maximum Sequence Length	128 tokens

We selected the above hyperparameters because they yielded the best performance for our BERT-based WSD model. Through multiple experiments, we observed that increasing the number of epochs beyond 9 did not improve model accuracy; instead, it led to overfitting, reducing generalization. The "AdamW optimizer" with an appropriately tuned learning rate as discussed in [1] ensured stable convergence, while a batch size of 16 provided an optimal balance between computational efficiency and training stability. The dropout rate of 0.1 effectively prevented overfitting, and the sequence length of 128 tokens was chosen to efficiently process Quranic verses while maintaining contextual integrity. These choices align with best practices in BERT fine-tuning for NLP tasks ensuring a well-generalized and high-performing model.

Model Training:

The training phase involved fine-tuning the BERT model on the annotated dataset. A custom input/output data handler class was designed to encode inputs with their respective tokens and assign integer labels to the corresponding senses. Both the training and testing datasets were prepared in the form of PyTorch DataLoaders to process data in batches. The training of the model was done with the AdamW optimizer at a learning rate of 2×10^{-5} for 9 epochs. Cross-entropy was used in order to quantify the deviation of the predicted and actual sense labels. Loss values were tracked to check the stability of the convergence during each epoch. The training was performed in a GPU environment to speed up the computations. Table 2 shows the number of sentences used for training against each ambiguous word.

Table 2: Training Dataset Specifications.

Words	No. of Sentences	Words	No. of Sentences
دل	96	رنگ	85
نظر	90	روشنی	104
طور	54	زبان	90
کم	63	زندہ	89
بجلی	75	سفر	64
بند	82	سمجھ	62
بھول	70	سوال	85
پاس	111	شامل	57
پانی	117	شکر	81
تیار	97	صحیح	79
حصہ	89	عمر	84
خاص	89	قسم	71
خون	69	کار	87
درمیان	67	کتاب	88
دور	61	کہنا	117
دیکھ	118	مل	114
دین	54	ملک	84

Model Evaluation:

The trained model was evaluated to assess its performance in predicting the correct senses of ambiguous words. To present a broad picture of model performance, different measures were used to compare the results. For the evaluation criterion, the Weighted F1-score was used, averaging the F1-score across different sense classes. Additionally, a class-wise evaluation

was conducted, presenting precision, recall, F1-score, and a detailed classification report. An added visual to the model was the confusion matrix so as to predict and evaluate the predictive scale for the potential improvement of the model. Thus, by integrating these basic metrics, the evaluation procedure delivered more comprehensive insights into the prospects and drawbacks of the model. The results not only substantiated the integration of the proposed methodology but also provided insight into the difficulties encountered in WSD for the Urdu translation of the Quran. The list of the top 5 ambiguous words and the number of sentences from Surah Al-Baqarh in which they appear are given below in Figure 2.

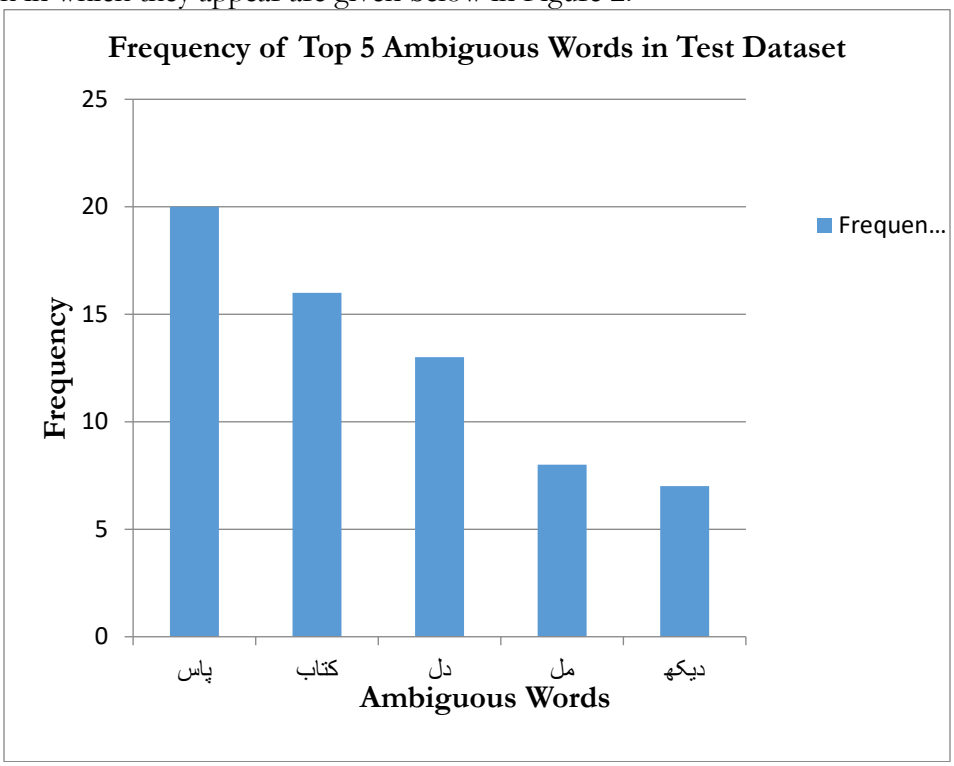

Figure 2. Frequency of Ambiguous Words in Test Dataset.

Sense Prediction:

A custom sense prediction framework was developed to enable the practical application of the trained model. In the case of input sentences, they were preconditioned to contain <target> and </target> around the ambiguous word. These sentences were preprocessed and then changed into input tensors. The most likely sense label was also identified by the model, as were the sense meanings from the sense inventory. Showing the predicted sense along with all the possible senses of the target word helped in interpreting and translating the text to an appropriate degree of precision.

Result and Discussions:

The BERT-based multilingual sequence classification model (mBERT) was trained with a training corpus for 9 epochs. During the training, the AdamW optimizer was used, with the learning rate equal to 2e-5. Loss values were recorded at the end of each epoch to evaluate the model's learning progress and convergence status. The detailed loss values are presented in the Table 3:

Table 3.Training Loss per Epoch.

Epoch	Loss
1/9	1.4919
2/9	1.0170
3/9	0.8071
4/9	0.7021

5/9	0.5887
6/9	0.4819
7/9	0.3726
8/9	0.2758
9/9	0.2349

The loss function gradually declines as shown in Figure 3. This shows that training is helpful in the context of the examined problem. Largely, after encompassing all the epochs, the loss rate was greatly minimized, which showcased the smooth enhancement of the model's learning ability of the relationship between the inputted sentences and the sense labels.

The effectiveness of this model was tested using a new set of test sentences that had not been used in any of the related works before. Precision, recall, F1-score, and support values were determined for the eight sense labels investigated in the study.

The results illustrated that Sense 0 and Sense 4 obtained higher support and achieved higher precision, recall, and F1 scores in the dataset. The overall good levels shown by the average F1 score of 0.62 suggest that the model better represents the most frequent senses in its output. Compared with the other senses, Sense 3, Sense 5, and Sense 6 were found to perform poorly, due to issues caused by dataset imbalance. The lack of predictions in Sense 7 requires more training samples for the underpopulated senses. A confusion matrix was computed, accompanied by a graphical representation to illustrate the model's performance in distinguishing related senses as shown in Figure 4. The evaluation was conducted on a separate test set containing previously unseen sentences.

Notably, the correct predictions are concentrated around the diagonal, which suggests that the model is capable of giving the right sense for the material tested in most cases. Off-diagonal elements indicate the particular occasions when the model failed to classify the correct sense properly in some cases, simultaneously using quite similar or complementary senses.

Detailed Analysis of the Confusion Matrix:

Through the confusion matrix, we can identify major misclassification patterns that demonstrate the model's problems in discriminating between different word senses. The model accurately classified 35 examples of Sense0 yet misidentified these instances 10 times, most frequently mislabeling them as Sense1, because of semantic affiliation issues. The prediction model misidentified over half of Sense1 instances since it correctly classified 16 out of 33 examples but wrongly marked 16 examples as Sense0 because of semantic relatedness. The classifier for Sense2 achieved 14 correct matches out of 22 but often mixed up Sense0 and Sense3 possibly due to weak distinctness in training samples and shared word patterns. The complete misclassification of Sense3, Sense5, Sense6, and Sense7 stands out as a major concern. Sense3 presented itself 4 times in the dataset yet the algorithm failed to properly classify any of these instances which were mistakenly grouped under Sense0, and Sense2. Sense5 and Sense6 received no correct predictions which further demonstrate the issues of extreme class imbalance because they prevent the model from learning to foresee underrepresented senses.

The detailed classification report is provided below in Table 4:

Table 4: Classification Report.

Sense	Precision	Recall	F1-Score	Support
Sense0	0.53	0.78	0.63	45
Sense1	0.57	0.48	0.52	33
Sense2	0.78	0.64	0.70	22
Sense3	0.00	0.00	0.00	4
Sense4	0.96	0.72	0.82	32

Sense5	0.00	0.00	0.00	3
Sense6	0.00	0.00	0.00	1
Sense7	0.00	0.00	0.00	0
Accuracy			0.63	140
Macro avg	0.35	0.33	0.33	140
Weighted avg	0.65	0.63	0.62	140

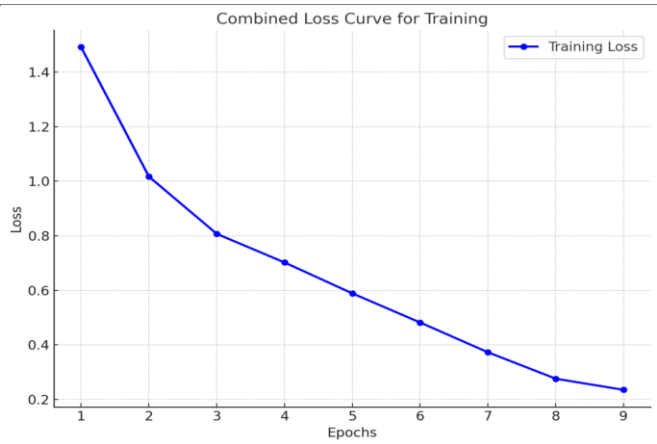

Figure 3. Combined Loss Curve for Training.

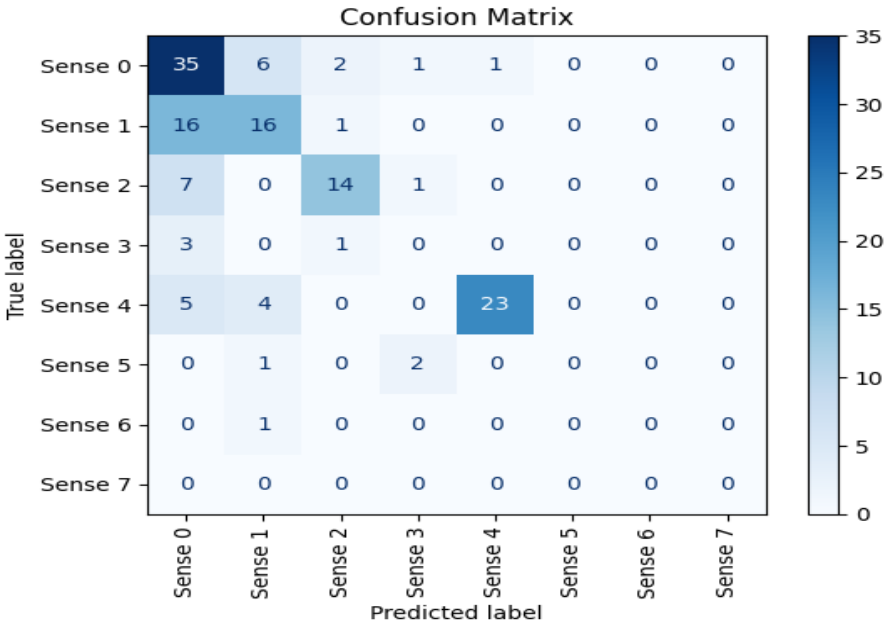

Figure 4. Confusion Matrix.

The model's complete inability to identify any occurrences of Sense7 is mainly due to the reason that there are no sentences in the training dataset in which this sense appears. While Sense4 achieved relatively good results with 23 correct classifications out of 32.The model demonstrates significant weaknesses in understanding three distinct aspects of the context: low-resource senses along with imbalanced classes and sense pairs that share contextual similarities. The model shows preference toward common senses which causes reduced performance for infrequent senses when generalizing. The current misclassification situations stress the requirement to use data augmentation and class weighing methods in the future, as due to time constraints these methods were not applied in this study,to improve training sets for underrepresented senses and develop class rebalancing strategies and improved contextual

representations to reduce semantic overlap between similar senses. The focus of future research should be on creating advanced methods for sense-specific feature extraction and domain-adaptive fine-tuning to overcome identified challenges. The sense prediction module's output is shown in Figure 5.

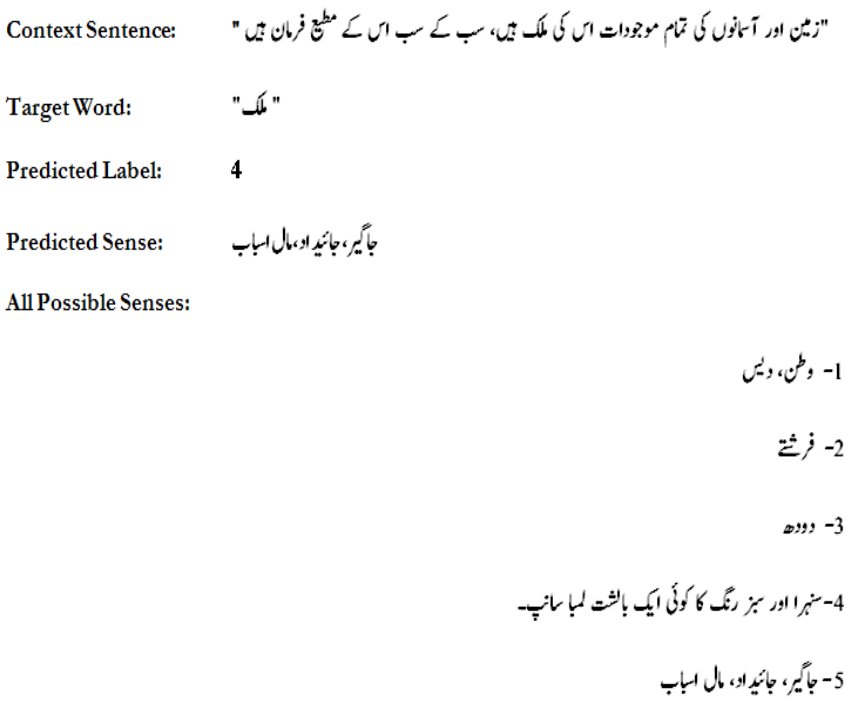

Figure 5. Sense Prediction Module.

An analysis of this study establishes the multilingual BERT model's capability to perform Word Sense Disambiguation effectively in the Urdu translations of the Holy Quran's Surah Al-Baqarah. Vital findings demonstrated the system achieved the highest frequent sense F1-score of 0.82 and overall presented an average weighted F1-score of 0.62 across predefined sense labels when verifying contextual ambiguity in words. The study's findings validate previous research demonstrating deep learning models' effectiveness for Word Sense Disambiguation but provide unique support for Urdu language processing, where computational linguistics lacks attention. The addition of a Sense Inventory alongside the system led to improved disambiguation processing, which demonstrates that linguistic resources aid in boosting translation accuracy.

The model achieved robust performance but its dependence on embeddings generated from previous training and its usage of a small dataset should be given additional research attention. Future research should increase the size of the dataset, assess different network topologies, and test the approach against various Quranic chapters as well as equivalent linguistic sections. This research brings a groundbreaking model to permanently address Urdu translations' persistent ambiguity in word meanings while providing important perspectives to Urdu translators alongside computational linguists and Quranic scholars working to achieve more accurate translations.

Conclusion:

Our study sought to improve Urdu translations of the holy Quran by handling word sense ambiguity with a multilingual BERT model. The analysis of Surah Al-Baqarah helped solve Urdu's multiple meaning and context-driven problems. The WSD model improved translation clarity dramatically through its use of a BERT model trained on context-derived data. The application's success with deep learning proves that BERT helps solve WSD problems in Urdu even when resources are limited. Our study improves translation

methodology while enhancing Quranic research and demonstrates how cutting-edge NLP aligns with precise religious text translation. This research offers several key contributions to the fields of computational linguistics, Quranic studies, and Urdu language processing. The next step in this study could be the expansion of the proposed Word Sense Disambiguation (WSD) beyond the current scope to encompass other Surahs of the Holy Quran. This could involve applying the same process across other surahs of the Quran; handling with word sense ambiguity for the comprehensive text. In the current setup, the multilingual BERT model has been employed, and the use of other contextual features could add more value.

Acknowledgement: I am grateful to my supervisor Dr. Muhammad Arshad Awan for their helpful guidance and support during my Master's thesis project. I am deeply thankful to my Computer Science professors at AIOU and my family members who supported me all this time. I thank all people who took part in this study and supported the research process.

Author's Contribution: This work is actually a teamwork in which the students contributed as a whole. They developed and executed the research project by creating a WSD methodology and implementing a BERT-based Word Sense Disambiguation model to improve Urdu translations of the Holy Quran. They built the dataset for ambiguous words and senses as part of data preparation before training and testing the WSD model to verify it achieved the study goals 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: This research was carried out in partial fulfillment of the requirement for the award of the Master of Science degree in Computer Science at Allama Iqbal Open University, Islamabad.

References:

[1] S. K. and R. Nassar, "EnhancedBERT: A feature-rich ensemble model for Arabic word sense disambiguation with statistical analysis and optimized data collection," *J. King Saud Univ.-Comput. Inf. Sci*, vol. 36, no. 1, p. 101911, 2024.

[2] and T. J. S. P.Jha, S. Agarwal, A. Abbas, "A novel unsupervised graph-based algorithm for Hindi word sense disambiguation," *SN Comput. Sci*, vol. 4, no. 5, p. 675, 2023.

[3] and P. R. J. Shafi, H. R. Iqbal, R. M. A. Nawab, "UNLT: Urdu Natural Language Toolkit," *Nat. Lang. Eng*, vol. 29, no. 4, pp. 942–977, 2023.

[4] and M. S. A.Saeed, R. M. A. Nawab, "Investigating the feasibility of deep learning methods for Urdu word sense disambiguation," *Trans. Asian Low-Resour. Lang. Inf. Process*, vol. 21, no. 2, pp. 1–16, 2021.

[5] and X. Y. G. C. X. Zhang, Y. L. Shao, "Word sense disambiguation based on RegNet with efficient channel attention and dilated convolution," *IEEE Access*, vol. 11, pp. 130733–130742, 2023.

[6] and N. K. R. N. Tyagi, S. Chakraborty, A. Kumar, "Word sense disambiguation models emerging trends: A comparative analysis," *J. Phys. Conf. Ser.*, vol. 2161, no. 1, p. 012035, 2022.

[7] and N. H. M. F.Ullah, A. Saeed, "Comparison of pre-trained vs custom-trained word embedding models for word sense disambiguation," *Adv. Distrib. Comput. Artif. Intell. J.*, vol. 12, no. 1, pp. e31084–e31084, 2023.

[8] and X. H. L. Huang, C. Sun, X. Qiu, "GlossBERT: BERT for word sense disambiguation with gloss knowledge," *arXiv Prepr.*, 2019.

[9] R. Saidi, et al., "WSDTN: A Novel Dataset for Arabic Word Sense Disambiguation," *Int. Conf. Comput. Collect. Intell.*, 2023.

[10] M. Alian and A. Awajan, "Arabic Word Sense Disambiguation Using Sense Inventories," *Int. J. Inf. Technol.*, vol. 15, no. 2, pp. 735–744, 2023.

[11] B. Abdelaali and Y. Tlili-Guiassa, "Swarm Optimization for Arabic Word Sense

Disambiguation Based on English Pre-Trained Word Embeddings," *"5th Int. Symp. Informatics its Appl. (ISIA)", IEEE,* 2022.

[12] M. A. Abderrahim and M. E.-A. Abderrahim, "Arabic Word Sense Disambiguation for Information Retrieval," *Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 21, no. 4, pp. 1–19, 2022.

[13] "Tanzil Project, Tanzil Quran Text, Available: https://tanzil.net/trans/.".

[14] "GitHub repository, Available: https://github.com/alisaeed007/UAW-WSD-18-Corpus.".

[15] P. Saeed, A., Nawab, R. M. A., Stevenson, M., & Rayson, "A sense annotated corpus for all-words Urdu word sense disambiguation," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 18, no. 4, pp. 1–19, 2019.