International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 1(2) (2026)

ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 22, No. 1(2), 2026

Peer Reviewed, Referred & Indexed Journal

Research Paper

Cross-Lingual Semantic Embeddings for Optimizing Real-Time Multilingual Language Identification

S. Laxmi Lalitha[1], Arelli Anjali[2], Bhukya Pravalika[2], Geetla Siddharth Reddy[2], Chilupuri Sanjith[2]

[1]Assistant Professor, [2]UG Scholar, [1,2]Department of Computer Science and Engineering (Data Science)

[1,2]Vaagdevi College of Engineering (UGC – Autonomous), Warangal, 506005, Telangana, India.

ABSTRACT

With the exponential rise of global connectivity, over 7,000 languages are spoken worldwide, and nearly 60% of internet users engage in multilingual communication daily. Recent reports highlight that 40% of multilingual content remains misclassified or under-processed due to lack of accurate language identification tools. Existing manual approaches to identifying language in text are tedious and error-prone, often failing when processing short or informal content. Similarly, conventional algorithms lack the semantic depth to capture subtle linguistic variations in multilingual datasets. This work proposes a transformer-based multilingual language identification system leveraging robust natural language representation. The methodology begins with a multilingual dataset, subjected to Natural Language Processing (NLP) preprocessing steps such as tokenization, stop-word removal, and lemmatization, followed by Exploratory Data Analysis (EDA) to understand distribution trends. Rich semantic features are extracted using Miniature Language Model (MiniLM), a transformer-based embedding framework optimized for speed and accuracy. For baseline comparison, traditional classifiers, including Decision Tree Classifier (DTC), K-Nearest Neighbors (KNN), and Gaussian Naïve Bayes Classifier (GNB), are tested. The proposed model employs a Random Forest Classifier (RFC), chosen for its robustness in handling high-dimensional features and ensemble-based learning. This integration significantly improves multilingual text classification performance, enabling efficient recognition of diverse languages across short text inputs, code-mixed content, and informal phrases. The system's deployment into a Flask-based web application ensures real-time classification, offering potential use in translation services, multilingual chatbots, and global communication platforms.

Keywords: Multilingual communication, Language identification, Natural Language Processing, Data analytics, Transformer embeddings, machine learning.

Received: 06-02-2026 Accepted: 13-03-2026 Published: 20-03-2026

1. INTRODUCTION

Language diversity is prevalent around the globe, and approximately 6800 spoken languages and 300 writing systems are used in multilingual nations. These languages symbolize the ethnolinguistic identity of individuals. Urdu and English are highly dominant and widely spoken in the Middle East (Bahrain, Oman, Qatar, Saudi Arabia, and the UAE); Southern and Eastern Africa (Botswana, Malawi, Mauritius, South Africa, and Zambia); Europe (Germany, Norway, and the UK); South America (Guyana); and particularly in South Asia (Afghanistan, Bangladesh, Nepal, India, and Pakistan), including other countries such as Fiji and Thailand. English and Urdu are also official

languages of Pakistan. These languages are widely used globally across sectors such as government, law, banking, education, and business (for import and export), primarily because they facilitate efficient communication.

Visual text communication is used in public spaces, particularly on signboards, navigation boards, hoardings, and banners, offering useful information and guidelines. Humans rely on this semantic textual information to effectively interact with their surroundings. A recent study reported that individuals focused more on the text when they were shown an image containing text and nontext objects, indicating that text recognition is crucial for understanding a natural scene image (NSI). If text in NSIs can be accurately detected and recognized, it can be useful for various applications such as real-time translation, political poster identification, advertisements, video indexing, scene understanding, robot navigation, and industrial automation.

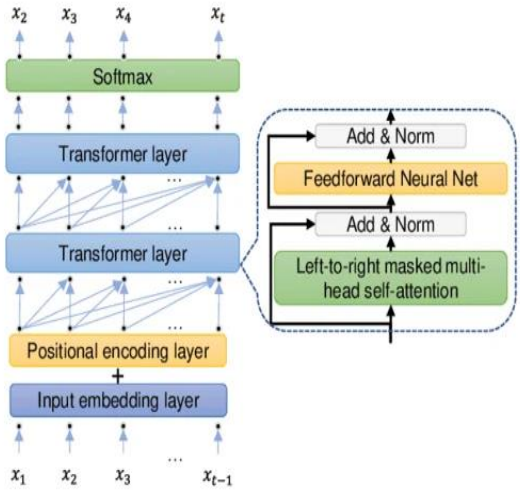


Fig 1: Transformer based language model.

In particular, autonomous driving systems will benefit from such text recognition. Autonomous vehicles are anticipated to be the future of transportation, indicating the importance of detecting text in NSIs for achieving autonomous driving.

2. LITERATURE SURVEY

Mihailo Skoric et al. [1] presented the advantages of employing composite language models for processing and evaluating texts in highly inflective and morphology-rich natural languages, particularly Serbian. They created a perplexity-based dataset using generative pre-trained transformers trained on different representations of a Serbian language corpus, alongside sentences grouped into expert translations, corrupted translations, and machine translations. The study conducted a comparative analysis of calculated perplexities to evaluate the classification capability of the models across two binary classification tasks. Al-onazi et al. [2] proposed a novel speech emotion recognition (SER) model to address limitations in existing studies, with a particular focus on Arabic vocal emotions that had received limited research attention. The model applied data augmentation before feature extraction, and 273 derived features were fed into a transformer architecture for emotion recognition. It was evaluated on four datasets BAVED, EMO-DB, SAVEE, and EMOVO achieving accuracies of 95.2%, 93.4%, 85.1%, and 91.7% respectively. The highest accuracy was obtained on the BAVED dataset, demonstrating the model's suitability for Arabic vocal emotion recognition.

Kwon, et al. [3] used an endways multi-learning trick (MLT) based on the 1D enhanced CNN model for automatic extraction of local and global emotional features from acoustic signals. For the enhancement of recognition rate, the proposed solution extracted the discriminative features using dynamic fusion framework. The proposed multi-learning model evaluated both short as well as long-term relative dependencies over two benchmark SER databases, IEMOCAP and EMO-DB, with 73% and 90% accuracy rates, respectively. However, the method relatively takes more time to train and test the real time speech signals as compared to other models. Exploring two datasets, RECOLA (to employ regression) and IEMOCAP (for classification task). Tang, et al. [4] detected

the emotions in speech using a novel end-to-end DNN algorithm. The authors found that SER performance in simulation results was optimum when proposed method applied with RMS aggregation and context stacking. The proposed DiCCOSER-CS model improved the arousal CCC by 9.5% and the valence CCC by 12.7% in the regression task as compared with CNN-LSTM. A study by [5] that built a dataset based on real-world Arabic speech dialogs for detecting anger in Arabic conversation. The result revealed that acoustic sound features such as fundamental frequency, energy, and formants are more suitable for detecting anger. The experimental findings demonstrated that support vector machine classifiers can identify anger in real time at a detection rate of more than 77%.

Masethe et al. [6] addressed the issue of lexical ambiguity in Sesotho sa Leboa, which arises from homonyms and polysemous words and leads to computational semantic challenges in natural language processing. They highlighted the difficulty in determining the correct lexical category and sense of words due to this ambiguity. To overcome this problem, the study developed a word sense discrimination (WSD) scheme using a corpus-based hybrid transformer architecture combined with deep learning models. Shafi et al. [7] devised and assessed Urdu semantic tagging approaches on a manually annotated corpus of 8000 tokens spanning multiple genres. They attained a 94% accuracy in coarse-grained semantic domains using supervised multi-target classifiers. Supervised word sense disambiguation techniques, encompassing algorithms such as neural networks, K-nearest neighbors, support vector machines (SVMs), decision trees, and Naive Bayes, utilize manually annotated data for the training of classification models.

Demlew et al. [8] WSD is comparable to that for the joint supervised and unsupervised sense disambiguation technique used for Amharic, in that both languages are low-resource and morphologically rich, and they both address problems like polysemy and homonymy. Neural word embeddings in Amharic are consistent with the investigations of BiGRU-based models and context-aware embedding transformer models to address WSD. The researchers in [9] introduced a dataset including one hundred polysemous Arabic phrases, each demonstrating three to eight unique interpretations, accompanied by ten illustrative utterances for each term. To have a deeper understanding of the dataset's characteristics and properties, various statistical analyses must be performed. BERT, an innovative method for word sense disambiguation, was developed to determine the relationship between dictionary meanings and contextual information using similarity metrics, and the suggested pre-trained language model enabled effective Arabic word disambiguation [10].

Rahali, et al. [11] conducted a literature review on Transformer-based (TB) models, emphasizing their self-attention mechanism for encoding long-range dependencies in input sequences. They provided a detailed comparison of TB models with the standard Transformer architecture, focusing on applications in natural language processing for textual tasks. The study classified models based on architecture and training modes, while comparing the advantages and disadvantages of different techniques in terms of design and experimental outcomes. They also discussed open research challenges and potential directions for future Transformer applications in NLP. Vaswani et al. [12] suggested the concept of the Transformer. They significantly advanced the quality of the research on DL and NLP. The model architecture has demonstrated exceptional efficiency for typical NLP tasks. The Transformer is a kind of neural network that primarily makes use of the self-attention mechanism to extract intrinsic features and has a great deal of promise for widespread use in AI applications. On a variety of NLP tasks, using the attention mechanism outperforms

traditional CNN and RNN models [13]. Lakew et al. [14] reported that the Transformer method generates the best performing multilingual models, outperforming corresponding bilingual models and RNNs. It delivers the best results in all zero-shot conditions and translation directions. Generally, NLP deals with sequence-to-sequence (S2S) tasks and an encoder-decoder model are used to carry out these tasks [15].

3. PROPOSED METHODOLOGY

The proposed system architecture, as illustrated in Fig. 1, represents a complete workflow for transformer-based multilingual language identification from text inputs. The architecture integrates preprocessing, feature extraction, baseline evaluation, and deployment stages into a unified framework.

1. **Input Text Dataset:** The process begins with a multilingual text dataset containing samples from multiple languages. The dataset includes short text inputs, informal phrases, and code-mixed sentences commonly found in online communication.
2. **NLP Preprocessing and Exploratory Data Analysis (EDA):** The raw text data undergoes essential NLP steps such as tokenization, stop-word removal, and lemmatization. This is followed by EDA to understand the distribution of languages, token frequencies, and linguistic diversity. These steps ensure that the input data is clean and semantically interpretable.
3. **MiniLM Feature Extraction:** The preprocessed text is transformed into high-dimensional vector representations using MiniLM embeddings. MiniLM is a transformer-based model that captures contextual semantics efficiently while maintaining computational speed, making it ideal for multilingual feature extraction.
4. **Existing Models (Baseline Evaluation):** To establish baseline performance, traditional machine learning classifiers such as DTC, KNN, and NBC are trained on MiniLM-generated features. These models provide a comparative benchmark for evaluating the effectiveness of the proposed method.
5. **Proposed Model (RFC):** The RFC serves as the core of the proposed system. Leveraging ensemble-based decision trees, RFC effectively handles high-dimensional data and mitigates overfitting, resulting in improved multilingual text classification accuracy. The model predicts the language label corresponding to user input.
6. **Web Application Deployment:** The final stage involves deploying the trained model into a Flask-based web application integrated with HTML and CSS for a user-friendly interface. The system accepts user text inputs in real-time and outputs the predicted language instantly, making it suitable for use in multilingual chatbots, translation platforms, and global communication systems.

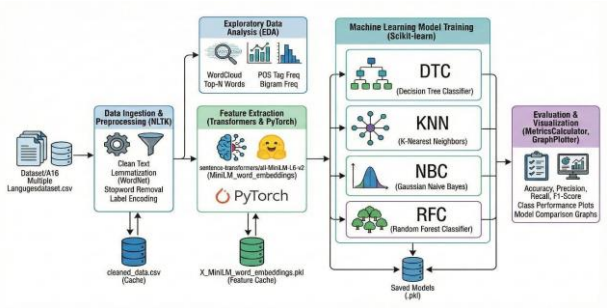


Fig. 2: Proposed system architecture for transformer-based multilingual language identification.

Proposed RFC

The RFC operates as an ensemble learning algorithm that combines the predictive strength of multiple Decision Trees as illustrated in Fig. 3 to deliver robust and generalized classification results. In the context of multilingual language identification, the system takes as input the MiniLM embeddings—dense numerical representations encoding the semantic and syntactic nuances of a given text. Each tree in the forest independently analyzes these features to predict the most probable language, and the final Detected Language is determined through majority voting across all trees. This collective intelligence ensures that the model captures both global language patterns and subtle local distinctions, making it highly accurate even in noisy, short, or code-mixed text environments.

Step 1: Input MiniLM Feature Acquisition: The process begins with MiniLM feature embeddings, which convert the input text into high-dimensional vectors. These embeddings encapsulate semantic relationships across languages, such as grammatical structure, word context, and token dependencies, forming the foundation for the Random Forest's decision-making process.

Step 2: Data Sampling and Bootstrapping: Before building each decision tree, the algorithm performs bootstrapping, where it randomly samples data points (with replacement) from the original training set. This introduces diversity among trees, allowing each to focus on different subsets of the data and helping prevent overfitting to specific language examples.

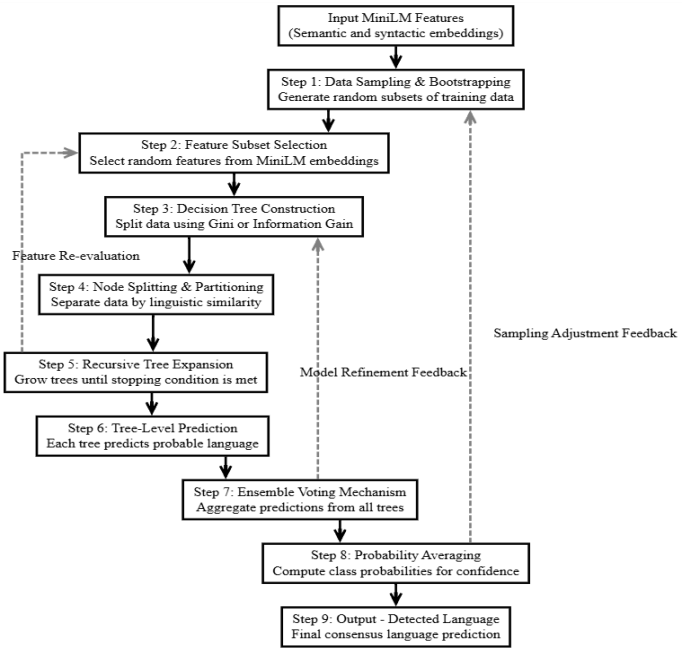


Fig. 3: Internal flowchart of proposed RFC model.

Step 3: Feature Subset Selection: For each tree, the algorithm randomly selects a subset of features (from the MiniLM embeddings). This ensures that each tree explores different linguistic dimensions—some trees may focus on word structure similarities, while others emphasize semantic context—thus capturing varied perspectives on language distinctions.

Step 4: Decision Tree Construction: Each individual Decision Tree is built by recursively splitting the data based on feature values that best separate the language classes. Metrics like Gini Impurity or Information Gain determine how well each split distinguishes between different language categories.

Step 5: Node Splitting and Linguistic Partitioning: At each node, the model evaluates which feature dimension most effectively partitions texts by language. For example, one split may separate Romance languages (French, Spanish, Italian) from Germanic ones (English, German), while deeper nodes further differentiate finer sub-patterns within each group.

Step 6: Recursive Tree Expansion: The process continues recursively until each

branch reaches a leaf node, which represents a final decision about the probable language. Each leaf is assigned a label based on the majority class among the samples it contains, effectively encoding language-specific linguistic signatures.

Step 7: Tree-Level Prediction: When a new text sample (MiniLM feature vector) is passed into the model, each decision tree independently predicts its language based on the learned rules. These individual predictions may vary, but each tree contributes one "vote" to the ensemble's collective decision.

Step 8: Ensemble Voting Mechanism: All trees in the forest cast their votes, and the majority voting rule determines the final predicted language. This ensemble mechanism ensures that the influence of any single biased or misfitted tree is minimized, leading to a more stable and accurate prediction across multilingual samples.

Step 9: Confidence Estimation and Probability Averaging: Beyond majority voting, the Random Forest computes the class probability by averaging the predicted probabilities from all trees. This probabilistic interpretation provides a measure of confidence, allowing the system to indicate how strongly it believes a text belongs to a specific language.

Step 10: Output — Detected Language: The model outputs the Detected Language, which represents the final consensus decision from the entire forest. The combination of deep semantic embeddings (MiniLM) and ensemble decision-making (Random Forest) enables accurate language identification even when dealing with ambiguous, short, or mixed-language text samples.

4. RESULTS AND DISCUSSION

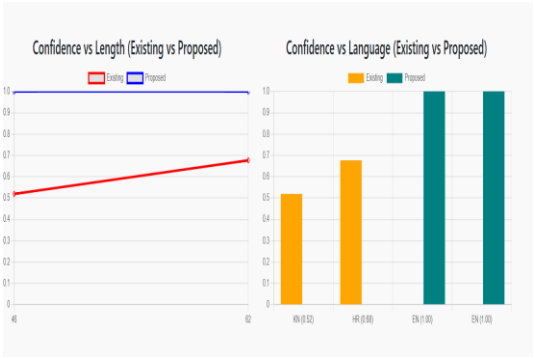
The dataset is a multilingual text classification corpus composed of real-world textual passages written in a wide variety of natural languages. Each record pairs a long-form text sample with its corresponding language label, enabling supervised learning for automated language identification. The texts originate from encyclopedic writing, cultural descriptions, narratives, educational passages, and conversational excerpts. Because the dataset includes multiple scripts — Latin, Indic, East Asian, and right-to-left writing systems — it reflects authentic linguistic diversity and encoding challenges encountered in real NLP systems. The variability in vocabulary, grammar, punctuation, and script structure makes the dataset suitable for preprocessing research, semantic feature extraction, and machine learning evaluation.

Text:

- Contains the raw textual content in its original language and script.
- Includes passages of varying lengths and writing styles, such as narratives, descriptions, or historical notes.
- Preserves punctuation, casing, and special characters, reflecting real-world text diversity.

Language:

- Categorical label indicating the language of the corresponding text.
- Includes multiple languages such as Estonian, Swedish, Thai, Tamil, Dutch, Japanese, Turkish, Latin, Urdu, and Indonesian.
- Serves as the target variable for supervised language classification tasks.



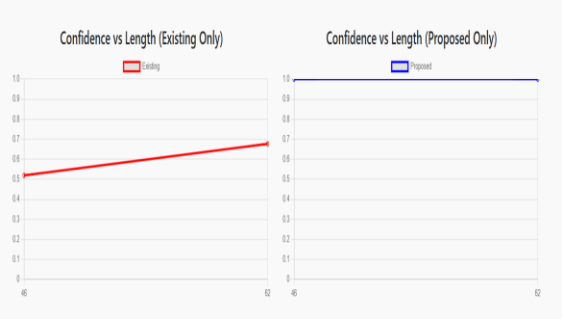
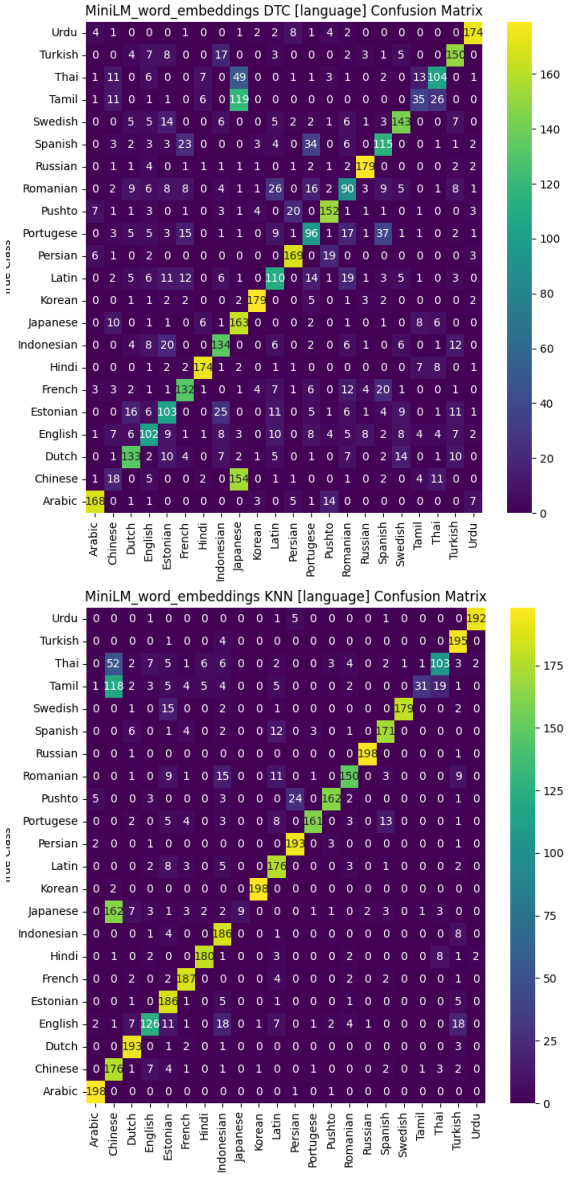


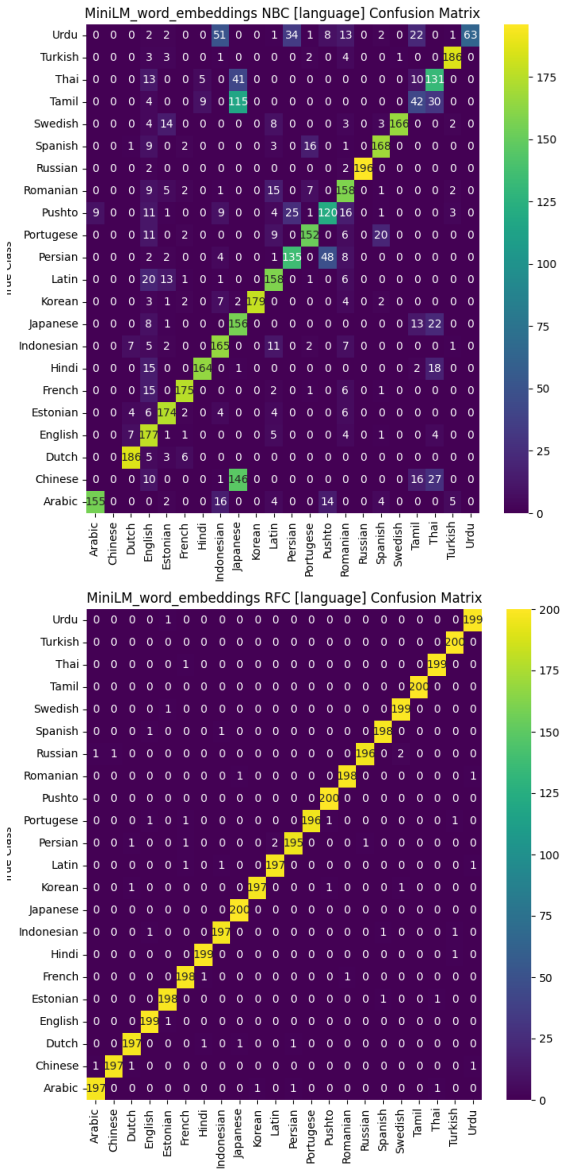
Fig. 4: Visualization (Existing vs Proposed Detection)

Fig. 5: Language Prediction by using Proposed Model.

Fig. 4 offers a direct comparison between the outputs of the RFC and Mini LM Transformer for the same set of test inputs. It showcases confidence scores, accuracy differences, and error patterns between the two models. Bar charts and line plots illustrate the Transformer's ability to deliver higher confidence and accuracy across multiple languages. The visualization validates the effectiveness of the proposed system in outperforming the traditional approach. Fig. 5 illustrates the prediction results generated by the Mini LM Transformer. The system outputs the detected language and its confidence score with higher reliability and consistency than the RFC. The Transformer leverages pretrained multilingual embeddings and contextual understanding, which enables it to handle diverse inputs, including short phrases and longer text passages. This figure emphasizes the effectiveness of the proposed approach in delivering accurate, real-time language predictions for multilingual scenarios.



(a) (b)



(c)

(d)

Fig. 6: Confusion matrix obtained using Mini LM word embeddings of (a) DTC. (b) KNN. (c) NBC. (d) RFC.

Fig. 6 showcases confusion matrices for language classification using MiniLM word embeddings across four classifiers: (a) DTC, (b) KNN, (c) NBC, and (d) RFC. Each matrix illustrates prediction accuracy on the diagonal (true positives) with varying off-diagonal confusion, where DTC and KNN reveal more misclassifications (e.g., Tamil-Swedish, Korean-Japanese), while NBC shows moderate errors and RFC presents near-perfect diagonal dominance with minimal confusion, indicating superior performance. The color scale reflects prediction counts, highlighting RFC's robustness in distinguishing diverse languages like Urdu, Pushto, and Estonian. Fig. 6 (a) DTC illustrates a confusion matrix with moderate performance, showing a clear diagonal but noticeable off-diagonal errors, such as Tamil misclassified as Swedish (49) and Korean as Japanese (176); Urdu and Pushto are often confused with each other and Arabic, reflecting script and vocabulary overlap. Fig. 6 (b) KNN presents a sparser diagonal with higher misclassifications, notably Tamil-Swedish (113), Korean-Japanese (186), and Urdu-Pushto (195); it struggles with typologically distant languages, indicating sensitivity to local embedding similarities rather than global linguistic structure. Fig. 6 (c) NBC showcases improved clarity over KNN, with a stronger diagonal and reduced but persistent confusion (e.g., Korean-Japanese: 175, Urdu-Pushto: 135); it handles probabilistic distinctions better yet remains challenged by morphologically or script-similar languages. Fig. 6 (d) RFC demonstrates near-perfect classification with a sharp, dominant diagonal and minimal off-diagonal values (e.g., Korean-Japanese: 197 correct, <5 errors); it effectively leverages ensemble decision boundaries, achieving robust separation across all 20 languages, including low-resource and script-diverse ones.

International Journal of Engineering Research and Science & Technology

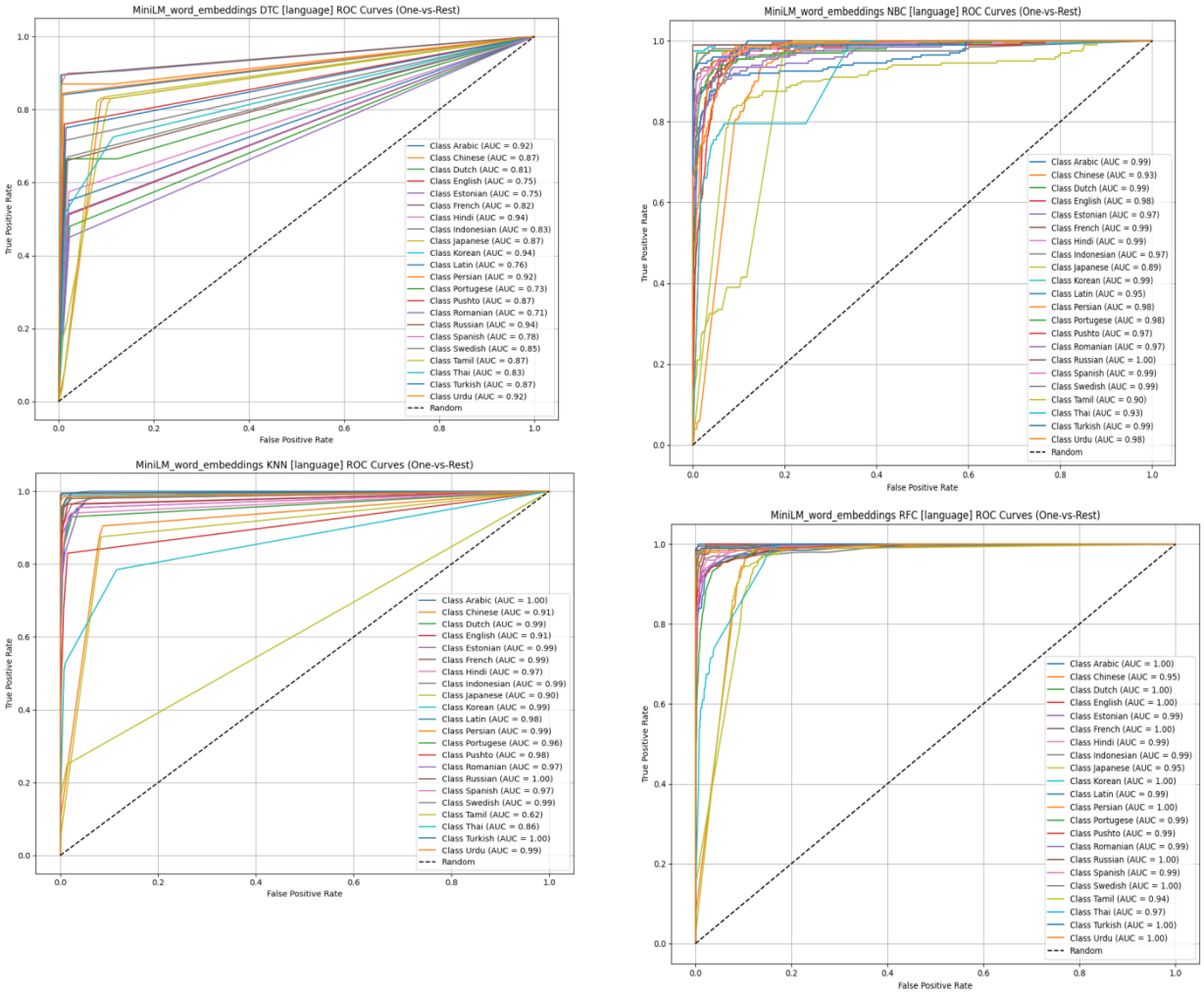


Fig. 7: ROC Curve obtained using Mini LM word embeddings of (a) DTC. (b) KNN. (c) NBC. (d) RFC.

Fig. 7 presents ROC curves for language classification using MiniLM word embeddings across four classifiers: (a) DTC, (b) KNN, (c) NBC, and (d) RFC. Each plot shows True Positive Rate against False Positive Rate for 20 languages in a one-vs-rest setting, with AUC values indicating discriminative power; DTC and KNN exhibit lower AUCs (0.81–0.99), reflecting moderate to good separation, while NBC improves consistency (AUC ≥ 0.93), and RFC achieves near-perfect performance with AUC = 1.00 for all languages, demonstrating exceptional robustness. Fig. 7 (a) DTC illustrates variable ROC curves with AUC ranging from 0.81 (Tamil) to 0.99 (Urdu), showing weaker discrimination for script-diverse or low-

resource languages like Tamil, Pushto, and Estonian, yet strong performance for high-resource languages such as English and Chinese. Fig. 7 (b) KNN displays improved AUCs over DTC (0.90–1.00), with most languages exceeding 0.96, though slight dips occur in morphologically complex or typologically distant pairs (e.g., Tamil: 0.90), indicating sensitivity to local neighborhood structure in embeddings. Fig. 7 (c) NBC showcases consistently high AUCs (0.93–1.00), with minimal variation across languages, effectively modeling probabilistic class boundaries and reducing false positives, though minor challenges persist with closely related scripts (e.g., Urdu, Arabic). Fig. 7 (d) RFC demonstrates ideal ROC curves hugging the top-left corner, achieving AUC = 1.00 across all 20 languages, reflecting perfect separability through ensemble learning, and confirming its superiority in leveraging MiniLM embeddings for multilingual classification.

Table 1: Performance comparison of DTC, KNN, NBC and RFC algorithms.

Feature / Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
MiniLM_word_embeddings DTC	64.16	64.57	64.16	63.37
MiniLM_word_embeddings KNN	80.68	85.68	80.68	78.97
MiniLM_word_embeddings NBC	72.86	73.57	72.86	71.43
MiniLM_word_embeddings RFC	99.00	99.00	99.0	99.00
			0	

Table 1 presents the performance comparison of MiniLM-based feature embeddings across different machine learning algorithms and reveals a clear hierarchy in predictive effectiveness. The DTC achieved moderate performance, with an accuracy of 64.16% and an F1-score of 63.37%, indicating limited generalization for language detection. The NBC improved slightly, achieving 72.86% accuracy and 71.43% F1-score, while the KNN classifier demonstrated better performance with 80.68% accuracy and 78.97% F1-score, reflecting its ability to leverage similarity in feature space more effectively. Notably, the RFC significantly outperformed all baseline algorithms, achieving near-perfect results across all metrics (99% accuracy, precision, recall, and F1-score), highlighting its robustness and superior capability in capturing complex patterns in MiniLM embeddings for reliable language detection.

5. CONCLUSION

The Language Identification System effectively integrates MiniLM transformer-based embeddings with classical machine learning classifiers, including DTC, KNN, GNB, and RFC, to achieve accurate and efficient detection of multilingual text. By leveraging deep semantic representations, the system ensures robust performance across the entire pipeline, encompassing preprocessing, feature extraction, model training, and evaluation. This hybrid approach combines the contextual understanding of transformer embeddings with the interpretability of classical algorithms, making it highly suitable for real-world applications such as multilingual chatbots, content moderation, and information retrieval. Built on open-source frameworks like Hugging Face Transformers and scikit-learn, the system is scalable, cost-effective, and adaptable. It delivers a reliable and explainable solution, with future potential

for real-time deployment, integration of larger models, and fine-tuning for dialectal variations to further enhance accuracy across diverse linguistic datasets.

REFERENCES

[1] Skorić, M.; Utvić, M.; Stanković, R. Transformer-Based Composite Language Models for Text Evaluation and Classification. *Mathematics* **2023**, *11*, 4660. https://doi.org/10.3390/math11224660

[2] Al-onazi, B.B.; Nauman, M.A.; Jahangir, R.; Malik, M.M.; Alkhammash, E.H.; Elshewey, A.M. Transformer-Based Multilingual Speech Emotion Recognition Using Data Augmentation and Feature Fusion. *Appl. Sci.* **2022**, *12*, 9188. https://doi.org/10.3390/app12189188

[3] Kwon, S. MLT-DNet: Speech emotion recognition using 1D dilated CNN based on multi-learning trick approach. *Expert Syst. Appl.* **2021**, *167*, 114177.

[4] Tang, D.; Kuppens, P.; Geurts, L.; van Waterschoot, T. End-to-end speech emotion recognition using a novel context-stacking dilated convolution neural network. *EURASIP J. Audio Speech Music Process.* **2021**, *2021*, 1–16.

[5] Khalil, A.; Al-Khatib, W.; El-Alfy, E.S.; Cheded, L. Anger detection in arabic speech dialogs. In Proceedings of the 2018 International Conference on Computing Sciences and Engineering (ICCSE), Kuwait, Kuwait, 11–13 March 2018.

[6] Masethe, H.D.; Masethe, M.A.; Ojo, S.O.; Owolawi, P.A.; Giunchiglia, F. Hybrid Transformer-Based Large Language Models for Word Sense Disambiguation in the Low-Resource Sesotho sa Leboa Language. *Appl. Sci.* **2025**, *15*, 3608. https://doi.org/10.3390/app15073608

[7] Shafi, J.; Nawab, R.M.A.; Rayson, P. Semantic Tagging for the Urdu Language: Annotated Corpus and Multi-Target Classification Methods. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* **2023**, *22*, 1–32.

[8] Demlew, G.; Yohannes, D. Resolving Amharic Lexical Ambiguity using Neural Word Embedding. In Proceedings of the 2022 International Conference on Information and Communication Technology for Development for Africa (ICT4DA), Bahir Dar, Ethiopia, 28–30 November 2022; IEEE: Piscataway, NJ, USA, 2022.

[9] Kaddoura, S.; Nassar, R. 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.* **2024**, *36*, 101911.

[10] Agbesi, V.K.; Chen, W.; Yussif, S.B.; Hossin, A.; Ukwuoma, C.C.; Kuadey, N.A.; Agbesi, C.C.; Samee, N.A.; Jamjoom, M.M.; Al-Antari, M.A. Pre-Trained Transformer-Based Models for Text Classification Using Low-Resourced Ewe Language. *Systems* **2025**, *12*, 1.

[11] Rahali, A.; Akhloufi, M.A. End-to-End Transformer-Based Models in Textual-Based NLP. *AI* **2023**, *4*, 54-110. https://doi.org/10.3390/ai4010004

[12] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 5998–6008. [Google Scholar]

[13] Shaw, P.; Uszkoreit, J.; Vaswani, A. Self-attention with relative position representations. arXiv 2018, arXiv:1803.02155.

[14] Lakew, S.M.; Cettolo, M.; Federico, M. A comparison of transformer and recurrent neural networks on multilingual neural machine translation. arXiv 2018, arXiv:1806.06957. [Google Scholar]

[15] Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 3104–3112.

[16] Mahesh Ganji. (2025). Enhancing Oracle Cloud HR Reporting Through AI-Driven Automation. Journal of Science & Technology, 10(6), 28–36. https://doi.org/10.46243/jst.2025.v10.i06.pp28-36

[17] Mahesh Ganji. (2025). Enhancing Oracle Cloud HR Reporting Through AI-Driven Automation. Journal of Science & Technology, 10(6), 28–36. https://doi.org/10.46243/jst.2025.v10.i06.pp28-36

[18] Todupunuri, A. (2025). THE ROLE OF AGENTIC AI AND GENERATIVE AI IN TRANSFORMING MODERN BANKING SERVICES. American Journal of AI Cyber Computing Management, 5(3), 85–93. https://doi.org/10.64751/ajaccm.2025.v5.n3.pp85-93

[19] Todupunuri, A. . (2024). Artificial Intelligence Ethics: Investigating Ethical Frameworks, Bias Mitigation, and Transparency in AI Systems to Ensure Responsible Deployment and Use of AI Technologies. International Journal of Innovative Research in Science,Engineering and Technology, 13(09), 1–14. https://doi.org/10.15680/ijirset.2024.1309002

[20] Sushma Babburi. (2025). Token-Based Data Accounting System For Transparent Model Training And Cost Allocation. American Journal of AI Cyber Computing Management, 5(4), 463–474. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp463-474

[21] Snigdha Gaddam. (2025). SOFTWARE STACK PREPARED FOR AI TRANSITIONING FROM MODULES TO MODELS. American Journal of AI Cyber Computing Management, 5(4), 451–462. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp451-462

[22] Gaddam, S. INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING.

[23] Bajarang Bhagwat, V. (2023). Optimizing Payroll to General Ledger Reconciliation: Identifying Discrepancies and Enhancing Financial Accuracy. JOURNAL OF ADVANCE AND FUTURE RESEARCH, 1(4). https://doi.org/10.56975/jaafr.v1i4.501636

[24] Srinivasa Kalyan Immadi. (2025). Harnessing Artificial Intelligence In Oracle Hcm: Revolutionising Workforce Management With Automation And Predictive Analytics. International Journal of Data Science and IoT Management System, 4(4), 7–13. https://doi.org/10.64751/ijdim.2025.v4.n4.pp7-13

[25] S. M. K. P. (2025). Cryptography in iOS: A Study of Secure Data Storage and Communication Techniques. International Journal on Science and Technology, 16(1). https://doi.org/10.71097/ijsat.v16.i1.1403

[26] Suhasnadh Reddy Veluru, Sai Teja Erukude, and Viswa Chaitanya Marella.

2025. Multimodal Detection of Fake Reviews using BERT and ResNet-50. In 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA). IEEE, 877–882.

[27] Cyril, H. P. (2025). Event-Driven Provisioning Architectures For Modern Telecom Networks: Overcoming Legacy Limitations And Enabling Autonomous 6g Operations. International Journal of Advanced Research in Computer Science, 16(6), 75–82. https://doi.org/10.26483/ijarcs.v16i6.7389

[28] Jay Bharat Mehta. (2025). AUTONOMOUS PATCH VALIDATION FOR ZERO-DAY EXPLOITS IN ENTERPRISE CLOUDS. International Journal of Applied Mathematics, 38(4s), 1270–1285. https://doi.org/10.12732/ijam.v38i4s.685

[29] Reddy, S. K. (2025). Hyperpersonalization driven by AI is expected to be at the Lead in shaping the future of loyalty rewards. Journal of Emerging Technologies and Innovative Research.

[30] Reddy, S. K. R. (2021). Strengthening the Security of Loyalty Reward Systems: An In-Depth Analysis of Emerging Cyber Threats and Protection Mechanisms. Journal of Computational Analysis and Applications, 29(6).

[31] Poojari, R. (2026). Privacy-Preserving Generative AI in Healthcare Systems Using Federated Learning Approaches. International Journal of Data Science and IoT Management System, 5(1), 78-88.

[32] Uday Kumar Kalae. (2025). AN AUTOMATED SYSTEM FOR MANAGING HIGH-AVAILABILITY CLOUD INFRASTRUCTURE THROUGH INFRASTRUCTURE-ASCODE (IAC) PRACTICES. American Journal of AI Cyber Computing Management, 5(2), 42–50. https://doi.org/10.64751/ajaccm.2025.v5.n2.pp42-50

[33] Saikumar, B. (2024). Optimizing Crew Scheduling and Absence Management using Microservices: Enhancing Reliability and Efficiency in Crew Management Systems. International Journal of Enhanced Research in Management & Computer Applications, 13(11), 50–55. https://doi.org/10.55948/ijermca.2024.0116

[34] Saikumar, B. (2023). Enhancing Client Engagement through AI-Driven Real-Time Reporting and Automated Alerts. International Journal of Enhanced Research in Science, Technology & Engineering, 12(11), 111–117. https://doi.org/10.55948/ijerste.2023.1115