Language Identification For Multilingual Machine Translation

Maddu Karthik[1], Mundru Vaishnavi[2], Kasu Swathi[3], Muppalla Sankar[4], Mundru Chandra Sekhar Rao[5]

[2,3,4,5]UG Student,ECE,Chalapathi Institute Of Engineering&Technology Guntur-Andhra Pradesh,India

[1]Assistant Professor ECE,Chalapathi Institute Of Engineering&Technology Guntur-Andhra Pradesh,India

Abstract**—**Language identification plays a vital role in multilingual machine translation systems, as it determines the source language of a given text before translation is performed. Accurate identification is essential to ensure that the appropriate translation model is selected, thereby improving the quality and reliability of the translated output. In this project, we propose an efficient language identification approach using machine learning rechniques to enhance multilingual translation performance. The system utilizes supervised learning algorithms such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Random Forest to classify input text into its corresponding language. Text data is preprocessed using techniques like cleaning, tokenization, and TF-IDF-based feature extraction with n-grams to capture linguistic patterns effectively. The trained models are evaluated using performance metrics such as accuracy, precision, recall, and F1-score.

Keywords— Language Identification, Multilingual Machine Translation, Natural Language Processing (NLP), Machine Learning, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest, TF-IDF, N-grams, Text Classification, Feature Extraction, Language Detection, Supervised Learning.

I. INTRODUCTION

In today's digitally connected world, communication across different languages has become increasingly important. With the rapid growth of the internet, social media, and global information exchange, there is a rising demand for systems that can automatically understand and translate multiple languages. Multilingual machine translation systems play a key role in breaking language barriers, enabling users to access information and communicate effectively regardless of their native language.

A fundamental step in any multilingual translation system is language identification, which involves determining the language of a given input text. Accurate language identification is crucial because the performance of the translation system depends heavily on selecting the correct language model. If the input language is incorrectly identified, it can lead to poor translation quality and misinterpretation of information.

Language identification is a challenging task due to the diversity of languages, variations in writing styles, and the presence of short or noisy text data. The problem becomes even more complex in real-world scenarios where code-switching occurs, meaning multiple languages are used within the same sentence or text. Traditional rule-based and statistical methods often struggle to handle such complexities, especially when dealing with limited or ambiguous data.

Recent advancements in machine learning and natural language processing (NLP) have provided more effective solutions for language identification. By leveraging techniques such as feature extraction, n-grams, and supervised learning algorithms,

Doi:10.48047/ijiee.2026.16.2.12

modern systems can learn patterns from large datasets and accurately classify languages. Algorithms like K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Random Forest have shown promising results in handling multilingual text classification tasks.

In this project, we propose a machine learning-based approach for language identification that integrates seamlessly with a multilingual machine translation system. The system preprocesses text data, extracts meaningful features using TF-IDF and n-grams, and applies multiple classification algorithms to identify the language. Once the language is detected, the text is translated into English, enabling effective cross-language communication.

The main objective of this work is to improve the accuracy and efficiency of language identification, thereby enhancing the overall performance of multilingual machine translation systems. The proposed approach is designed to handle diverse text inputs and provide reliable results in practical applications.

II. REVIEW &LITERATURE SURVEY

Language identification has been widely studied as a fundamental task in natural language processing (NLP), particularly for its role in multilingual machine translation systems. Over the years, researchers have proposed various approaches ranging from traditional statistical methods to advanced deep learning models.

Early work in language identification primarily relied on rule-based and statistical techniques, where languages were distinguished using character frequencies, word distributions, and handcrafted linguistic rules. One notable contribution is the work by Marco Lui and Timothy Baldwin, who introduced *langid.py*, an efficient and widely used language identification tool. Their approach demonstrated strong performance across multiple domains without requiring extensive preprocessing, making it suitable for real-world applications.

With the advancement of machine learning, researchers began exploring supervised learning approaches for language classification. Techniques such as K-Nearest Neighbors (KNN) and Support Vector Machines (SVM) became popular due to their ability to handle high-dimensional textual data. Studies have shown that SVM, in particular, provides high accuracy in text classification tasks because of its capability to find optimal decision boundaries in complex feature spaces.

Further improvements were achieved through feature extraction techniques such as n-grams and TF-IDF representations. The work by Armand Joulin et al. introduced efficient text classification methods that leverage simple yet powerful features, significantly improving performance while maintaining computational efficiency. These methods are highly effective in capturing language-specific patterns, especially in short text data.

Recent developments have focused on deep learning and neural network-based approaches. Researchers like Tom Kocmi and Ondřej Bojar proposed neural network models that use character-level representations to improve language identification accuracy, particularly in multilingual and code-switching scenarios. Similarly, advanced language models such as BERT (by Jacob Devlin et al.) and GPT (by Alec Radford et al.) have demonstrated remarkable capabilities in understanding contextual information, leading to improved performance in language-related tasks.

In the context of multilingual machine translation, studies by Guillaume Lample et al. and Emmanouil Antonios Platanios et al. highlight the importance of accurate language identification as a preprocessing step. Their work shows that incorrect language detection can significantly degrade translation quality, emphasizing the need for robust identification systems.

Despite these advancements, several challenges remain. Handling code-switching, identifying low-resource languages, and dealing

Doi:10.48047/ijiee.2026.16.2.12

with short or noisy text continue to be difficult problems. Research by King and Abney explored weakly supervised methods to address mixed-language documents, while other studies have focused on improving model generalization across diverse datasets.

In this project, we build upon these existing works by implementing a machine learning-based approach that combines effective feature extraction with multiple classification algorithms, including KNN, SVM, and Random Forest. The proposed system aims to provide high accuracy, robustness, and practical usability in real-world multilingual translation applications.

III. RESEARCH METHODOLOGY

The proposed system for language identification in multilingual machine translation follows a structured and systematic approach to ensure accurate and efficient performance. The process begins with the collection of a multilingual dataset that contains text samples from various languages along with their corresponding labels. This dataset forms the basis for training and evaluating the machine learning models. Care is taken to include diverse and representative data so that the system can generalize well across different languages and writing styles.

Once the dataset is collected, it undergoes a preprocessing stage to improve data quality and consistency. This step involves removing missing values, eliminating special characters, and cleaning noisy or irrelevant information from the text. The text is then standardized to a uniform format, which helps in reducing inconsistencies and improving the effectiveness of subsequent processing steps. Proper preprocessing plays a crucial role in enhancing the overall performance of the model.

After preprocessing, the text data is transformed into a numerical format through feature extraction techniques. In this project, Term Frequency–Inverse Document Frequency (TF-IDF) is used along with character-level n-grams ranging from one to three characters. These techniques help capture important linguistic patterns and structural characteristics of different languages, making it easier for the machine learning models to distinguish between them. The resulting feature vectors represent the textual data in a form suitable for classification.

The dataset is then divided into two parts: a training set and a testing set. Typically, 80% of the data is used for training the models, while the remaining 20% is reserved for testing. This division ensures that the models are evaluated on unseen data, providing a fair assessment of their performance and generalization capability.

In the model training phase, three supervised machine learning algorithms are implemented: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Random Forest. Each algorithm learns patterns from the training data and builds a predictive model for language classification. KNN classifies text based on similarity to neighboring data points, SVM identifies optimal decision boundaries in high-dimensional space, and Random Forest improves accuracy by combining multiple decision trees.

Following training, the models are evaluated using performance metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive understanding of the model's effectiveness. Additionally, confusion matrices are used to visualize the classification results and identify any misclassifications. A comparison of all three algorithms is carried out to determine the best-performing model.

Once the optimal model is selected, it is used for real-time language detection. When a user inputs text, the system preprocesses and converts it into feature vectors, which are then passed to the trained model to predict the language. After identifying the language, the system integrates a translation module that converts the detected text into English, enabling smooth multilingual communication.

Finally, all components of the system are integrated into a user-friendly interface that allows users to upload datasets, train models, compare performance, and perform language

Doi:10.48047/ijiee.2026.16.2.12

detection and translation. This complete methodology ensures that the system is both efficient and practical for real-world multilingual applications.

IV. EXISITING SYSTEM

Existing systems for language identification in multilingual machine translation primarily rely on traditional approaches such as rule-based and statistical methods. Rule-based systems use predefined linguistic rules, including grammar patterns, dictionaries, and language-specific characteristics, to identify the language of a given text. While these systems can work effectively for well-structured and clearly defined inputs, they are often difficult to maintain and update, especially when dealing with a large number of languages or evolving linguistic patterns.

Statistical methods, on the other hand, analyze features such as character frequency, word distribution, and n-gram patterns to determine the language. These approaches provide better flexibility compared to rule-based systems and can handle a wider range of languages. However, their performance is often limited when dealing with short text inputs, noisy data, or informal language commonly found in real-world applications such as social media and messaging platforms.

Another major limitation of existing systems is their inability to effectively handle code-switching, where multiple languages are used within the same sentence or document. Most traditional models assume that a text belongs to a single language, which leads to incorrect predictions in mixed-language scenarios. Additionally, these systems often struggle with low-resource languages, as they require large amounts of labeled data for accurate identification.

Furthermore, many existing language identification systems are not well integrated with modern multilingual machine translation frameworks. Incorrect language detection can result in selecting inappropriate translation models, which significantly degrades translation quality and leads to miscommunication. These limitations highlight the need for more advanced, flexible, and accurate approaches to language identification that can adapt to real-world multilingual environments.

V. PROPOSED METHODOLOGY

The proposed methodology aims to develop an efficient and accurate language identification system that enhances the performance of multilingual machine translation. Unlike traditional approaches, this system leverages machine learning techniques to automatically learn patterns from multilingual text data and provide reliable language detection, even in challenging scenarios.

The methodology begins with the collection of a diverse multilingual dataset containing text samples from various languages. This dataset is carefully prepared to include different writing styles and linguistic variations, which helps the model generalize better in real-world applications. The collected data then undergoes preprocessing, where unnecessary symbols, missing values, and noise are removed. The text is cleaned and standardized to ensure consistency and improve model performance.

Following preprocessing, the text data is transformed into a numerical representation using feature extraction techniques. In this system, TF-IDF (Term Frequency–Inverse Document Frequency) combined with character-level n-grams is used to capture important language patterns. This approach effectively represents the structural and statistical properties of different languages, making it easier for machine learning models to distinguish between them.

The processed dataset is then divided into training and testing sets, typically using an 80:20 ratio. The training data is used to build predictive models, while the testing data is used to evaluate

Doi:10.48047/ijiee.2026.16.2.12

their performance. Three supervised machine learning algorithms are employed in this system: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Random Forest. Each algorithm is trained on the dataset to classify text into its corresponding language based on learned features.

After training, the models are evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score. A comparative analysis is conducted to determine the most effective algorithm. Among the implemented models, SVM generally provides higher accuracy due to its ability to handle high-dimensional data efficiently, while Random Forest offers robustness and stability.

Once the best-performing model is selected, it is used for real-time language detection. When a user inputs text, the system processes it using the same feature extraction techniques and predicts the language using the trained model. To enhance usability, the system is integrated with a translation module that converts the detected language into English, enabling seamless multilingual communication.

Overall, the proposed methodology provides a robust, scalable, and practical solution for language identification by combining effective preprocessing, feature extraction, and machine learning techniques. This approach significantly improves accuracy and overcomes the limitations of traditional systems, making it suitable for real-world multilingual applications.

VI. BLOCK DIAGRAM

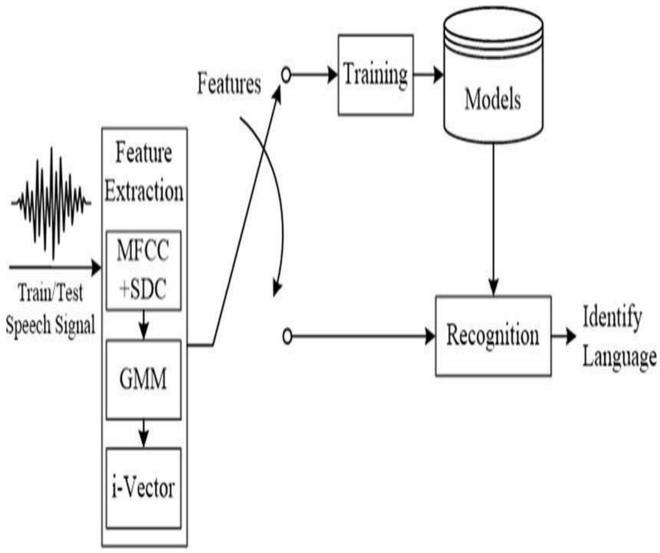

Fig. 6.2. Block Diagram

VII. RESULTS AND OUTCOMES

The proposed system for language identification and multilingual machine translation was successfully implemented and evaluated using a multilingual dataset. The performance of the system was analyzed by training three machine learning algorithms, namely K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Random Forest. Each model was tested using unseen data to ensure a fair evaluation of its accuracy and generalization capability.

Doi:10.48047/ijiee.2026.16.2.12

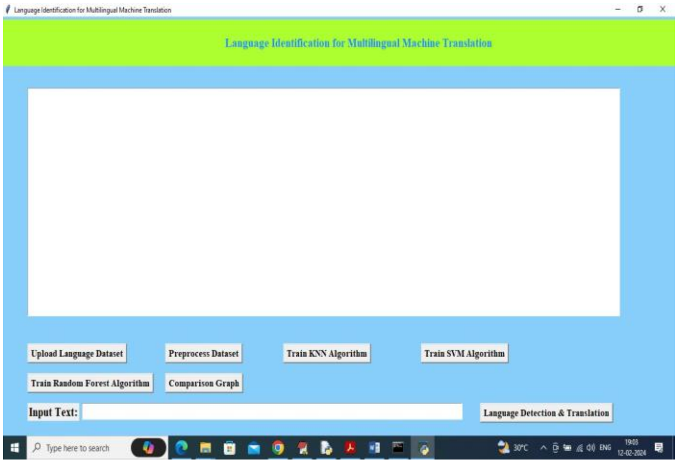

Fig:7.1:Output1

The experimental results indicate that all three algorithms were able to classify languages with reasonable accuracy; however, their performance varied based on the nature of the data and feature representation. Among the models, the Support Vector Machine (SVM) achieved the highest accuracy, demonstrating its effectiveness in handling high-dimensional textual data. KNN showed moderate performance due to its dependency on similarity measures and sensitivity to data distribution, while Random Forest provided stable and reliable results with balanced performance across evaluation metrics.

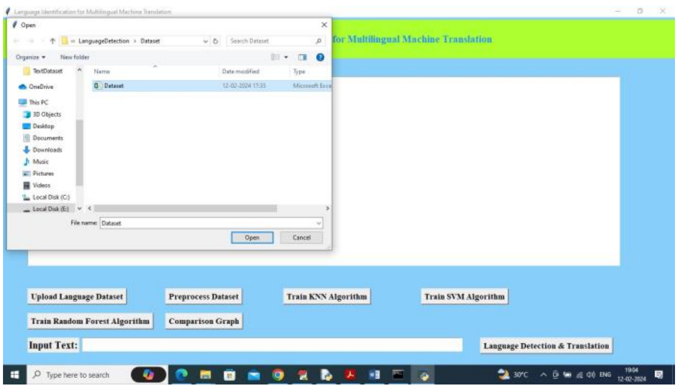

Fig:7.2:Output 2

Performance metrics such as accuracy, precision, recall, and F1-score were used to evaluate the models comprehensively. The results showed that the system is capable of accurately identifying languages even when the input text is short or contains minor noise. Confusion matrices further illustrated that most predictions were correctly classified, with only a few misclassifications occurring between linguistically similar languages.

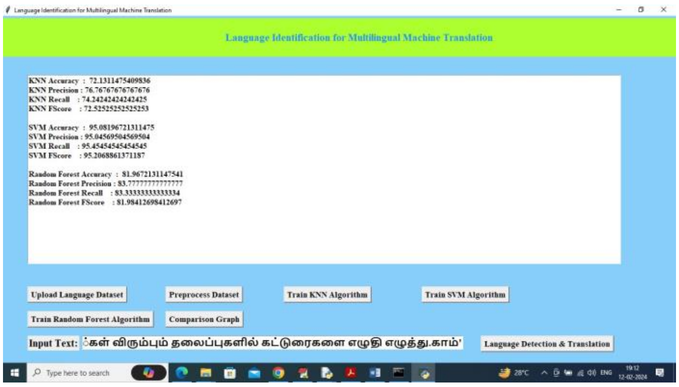

Fig:7.3:Output 3

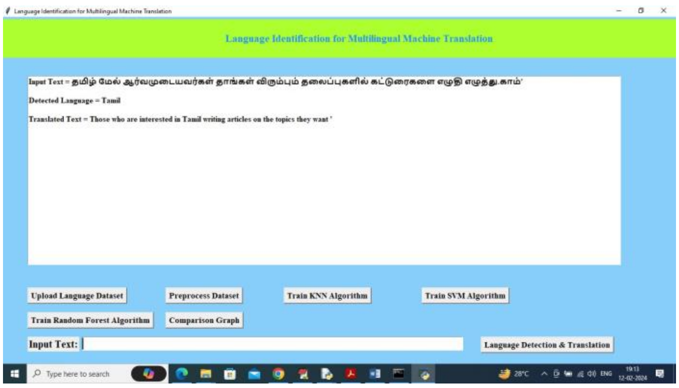

Fig:7.4:Output 4

In addition to language identification, the system successfully integrated a translation module that converts the detected language into English. This feature enhances the practical usability of the system by enabling users to not only identify the language but also understand the content through translation. The graphical comparison of algorithms provided clear insights into their relative performance, helping in selecting the most suitable model for deployment.

Doi:10.48047/ijiee.2026.16.2.12

Fig:7.5:Output 5

Fig:7.6: Output 6

Overall, the outcomes of this project demonstrate that the proposed machine learning-based approach is effective, reliable, and suitable for real-world multilingual applications. The system achieves high accuracy, supports multiple languages, and provides a user-friendly interface for language detection and translation, thereby improving communication across language barriers.

VIII.CONCLUSION

In this project, a comprehensive system for language identification in multilingual machine translation has been designed and implemented using machine learning techniques. The study emphasizes the importance of accurate language detection as a critical preprocessing step that directly influences the performance and reliability of translation systems. By addressing this fundamental requirement, the proposed system contributes to improving communication across multiple languages in real-world applications.

The system utilizes effective preprocessing methods, TF-IDF feature extraction, and character-level n-grams to represent textual data in a meaningful numerical form. These features enable the machine learning models to capture the unique linguistic patterns of different languages. The implementation of supervised learning algorithms such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Random Forest provides a comparative framework to evaluate different classification approaches. Among these, SVM demonstrated superior performance due to its ability to handle high-dimensional data and complex decision boundaries, while Random Forest offered robustness and stability.

One of the key achievements of this work is the ability of the system to handle diverse input conditions, including short text and moderately noisy data. The integration of a translation module further enhances the practicality of the system by not only identifying the language but also converting it into English, making it useful for end users. This combination of language detection and translation creates a complete pipeline that can be applied in applications such as chat systems, content analysis, and multilingual information retrieval.

Despite the promising results, certain challenges still exist. The system may face difficulties when dealing with highly mixed-language inputs (code-switching), very low-resource languages, or extremely ambiguous text. These challenges highlight the need for further improvements in model architecture, dataset diversity, and contextual understanding. Incorporating deep learning models such as neural networks or transformer-based architectures could

Doi:10.48047/ijiee.2026.16.2.12

significantly enhance performance in such complex scenarios.

Future work can focus on expanding the dataset to include more languages and dialects, improving real-time processing capabilities, and integrating advanced natural language processing techniques. Additionally, optimizing the system for deployment in web or mobile applications can increase its accessibility and usability.

In conclusion, the proposed system demonstrates that machine learning-based language identification is an effective and practical solution for enhancing multilingual machine translation. By achieving high accuracy and providing a user-friendly implementation, this work contributes to the development of intelligent systems that support seamless communication across language barriers, thereby addressing one of the key challenges in today's globalized digital world.

REFERENCES

1. Vellela, S. S., & Balamanigandan, R. (2024). Optimized clustering routing framework to maintain the optimal energy status in the wsn mobile cloud environment. Multimedia Tools and Applications, 83(3), 7919-7938.
2. Vellela, S. S., & Balamanigandan, R. (2023). An intelligent sleep-awake energy management system for wireless sensor network. Peer-to-Peer Networking and Applications, 16(6), 2714-2731.
3. Vellela, S. S., & Balamanigandan, R. (2024). An efficient attack detection and prevention approach for secure WSN mobile cloud environment. Soft Computing, 28(19), 11279-11293.
4. Vellela, S. S. (2023). Enhanced speckle noise reduction in breast cancer ultrasound imagery using a hybrid deep learning model. Ingénierie des Systèmes d'Information.
5. Polasi, P. K., Vellela, S. S., Narayana, J. L., Simon, J., Kapileswar, N., Prabu, R. T., & Rashed, A. N. Z. (2026). Data rates transmission, operation performance speed and figure of merit signature for various quadurature light sources under spectral and thermal effects. Journal of Optics, 55(1), 633-643.
6. Praveen, S. P., Nakka, R., Chokka, A., Thatha, V. N., Vellela, S. S., & Sirisha, U. (2023). A novel classification approach for grape leaf disease detection based on different attention deep learning techniques. International Journal of Advanced Computer Science and Applications (IJACSA), 14(6), 2023.
7. Vellela, S. S., Rao, M. V., Mantena, S. V., Reddy, M. J., Vatambeti, R., & Rahman, S. Z. (2024). Evaluation of Tennis Teaching Effect Using Optimized DL Model with Cloud Computing System. International Journal of Modern Education and Computer Science (IJMECS), 16(2), 16-28.
8. Vellela, S. S., & Krishna, A. M. (2020). On Board Artificial Intelligence With Service Aggregation for Edge Computing in Industrial Applications. Journal of Critical Reviews, 7(07).
9. Madhuri, A., Jyothi, V. E., Praveen, S. P., Sindhura, S., Srinivas, V. S., & Kumar, D. L. S. (2024). A new multi-level semi-supervised learning approach for network intrusion detection system based on the 'goa'. Journal of Interconnection Networks, 24(supp01), 2143047.
10. Raju, V. V. K., Bhavani, Y. V. K. D., Nandikonda, P., Kareemunnisa, F. N. U., Brahmeswara, K. B., & Sindhura, S. (2026). Iterative and Statistical Analytical Review of Predictive Modeling Approaches in Educational Systems: A Comprehensive Benchmark of AI-Driven Methods. International Journal of Innovative Technology and Interdisciplinary Sciences, 9(1), 490-522.
11. Biyyapu, N., Veerapaneni, E. J., Surapaneni, P. P., Vellela, S. S., & Vatambeti, R. (2024). Designing a modified feature aggregation model with hybrid sampling techniques for network intrusion detection. Cluster Computing, 27(5), 5913-5931.

Doi:10.48047/ijiee.2026.16.2.12

12. Praveen, S. P., Vellela, S. S., & Balamanigandan, R. (2024). SmartIris ML: harnessing machine learning for enhanced multi-biometric authentication. Journal of Next Generation Technology (ISSN: 2583-021X), 4(1).
13. Vuyyuru, L. R., Purimetla, N. R., Reddy, K. Y., Vellela, S. S., Basha, S. K., & Vatambeti, R. (2025). Advancing automated street crime detection: a drone-based system integrating CNN models and enhanced feature selection techniques. International Journal of Machine Learning and Cybernetics, 16(2), 959-981.
14. Vellela, S. S., Roja, D., Purimetla, N. R., Thalakola, S., Vuyyuru, L. R., & Vatambeti, R. (2025). Cyber threat detection in industry 4.0: Leveraging GloVe and self-attention mechanisms in BiLSTM for enhanced intrusion detection. Computers and Electrical Engineering, 124, 110368.
15. Vellela, S. S., Pushpalatha, D., Sarathkumar, G., Kavitha, C. H., & Harshithkumar, D. (2023). Advanced intelligence health insurance cost prediction using random forest. ZKG International, 8.
16. Vellela, S. S., Babu, B. V., & Mahendra, Y. B. (2024). IoT-based tank water monitoring systems: enhancing efficiency and sustainability. International Journal for Modern Trends in Science and Technology, 10(02), 291-298.
17. Vellela, S. S., Varshini, K., Jeevana, M., Kadheer, S. K., & Kumar, T. P. (2024). Iot based smart irrigation and controlling system. IoT Based Smart Irrigation and Controlling System, International Journal for Modern Trends in Science and Technology, 10(02), 77-85.
18. Vellela, S. S., Chaganti, A., Gadde, S., Bachina, P., & Karre, R. (2022). A Novel Approach for Detecting Automated Spammers in Twitter. Mukt Shabd, 11, 49-53.
19. Vellela, S. S., Narapasetty, S., Somepalli, M., Merikapudi, V., & Pathuri, S. (2022). Fake News Articles Classifying Using Natural Language Processing to Identify in-article Attribution as a Supervised Learning Estimator. Mukt Shabd Journal, 11.
20. Vellela, S. S., Vineeth, S., & Suresh, V. (2024). IoT Based ICU Patient Monitoring System. IoT Based ICU Patient Monitoring System, International Journal for Modern Trends in Science and Technology, 10(02), 265-273.
21. Vellela, S. S., & Balamanigandan, R. (2025). Designing a Dynamic News App Using Python. Available at SSRN 5250912.
22. Vellela, S. S., Rao, M. V., Krishna, C. V. M., Rao, T. S., & Dasthavejula, R. (2026). Piezoelectric and Shape-Memory Materials for Actuators and Energy Harvesting in Mechanical, Electronics, and Biomedical Engineering Using AI-Based Design. In Advanced Materials for Biomedical Devices (pp. 195-206). CRC Press.
23. Vellela, S. S., Singu, K., Kakarla, L. S., Tadikonda, P., & Sattenapalli, S. N. R. (2025). NLP-Driven Summarization: Efficient Extraction of Key Information from Legal and Financial Documents. Available at SSRN 5250908.
24. Vellela, S. S., Anusha, P., Vullam, N. R., Jala, J., Bellapu, V. S., & Vindhya, A. S. (2025, October). Quantum Cryptography and Key Distribution for Secure Communication in the Post Quantum World. In 2025 International Conference on Sustainable Communication Networks and Application (ICSCN) (pp. 619-624). IEEE.
25. Roja, D., Jidugu, S. K., Rao, T. S., Vuyyuru, L. R., Vellela, S. S., & Ranjani, B. S. (2025, December). High-Fidelity Image Synthesis using Enhanced Generative Adversarial Networks with Attention Mechanisms. In 2025 International Conference on NexGen Networks and Cybernetics (IC2NC) (pp. 885-890). IEEE.
26. Vellela, S. S., Vuyyuru, L. R., Jidugu, S. K., Rao, M. P., & Srinivas, B. R. (2025, November). The Impact Of Quantum Computing On Blockchain Security And Quantum Resistant Protocols. In 2025 2nd

Doi:10.48047/ijiee.2026.16.2.12

International Conference on Intelligent Systems for Cybersecurity (ISCS) (pp. 1-6). IEEE.

27. Yanamadala, N., & Vellela, S. S. (2025, June). Ensuring Authenticity and Confidentiality in Images using SHA-ECC Fusion. In 2025 Second International Conference on Networks and Soft Computing (ICNSoC) (pp. 684-689). IEEE.
28. Vellela, S. S. (2024). A Comprehensive Review of AI Techniques in Serious Games: Decision Making and Machine Learning.
29. Burra, R. S., APCV, G. R., & Vellela, S. S. (2024). Strategic Insights: Unleashing the Power of Big Data Analytics for Credit Investigation and Risk Mitigation in Commercial Banking. International Journal of Progressive Research in Engineering Management and Science, 4(01), 458-464.
30. Vellela, S. S., Purimetla, N. R., Vindhya, A. S., Vullam, N. R., Srinivas, B. R., & Vuyyuru, L. R. (2025, October). Design and Simulation of Quantum Error Correction Codes for Scalable Quantum Architectures. In 2025 7th International Conference on Innovative Data Communication Technologies and Application (ICIDCA) (pp. 1570-1575). IEEE.
31. Vellela, S. S., Purimetla, N. R., Rao, P. V., Daniel, V. A. A., Koppolu, H. K. R., & Janani, B. (2025). AI-Enabled Wearable Hemodynamic Monitoring System for Early Identification of Thrombotic Events. Vascular and Endovascular Review, 8(16s), 321-336.
32. Venkatesh, N., Maheswari, S., & Triveni, P. (2024). Harnessing IoT for Real-Time Plant Health Monitoring: Challenges and Opportunities.
33. Reddy, B. V., Kumar, A. H., Gopi, C., Prasad, Y. V. D., Vellela, S. S., & Roja, D. (2025, April). Machine learning based automated liver fibrosis stage diagnosis with prediction. In 2025 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE) (pp. 1-6). IEEE.
34. Rao, M. V., Sreeraman, Y., Mantena, S. V., Gundu, V., Roja, D., & Vatambeti, R. (2024). Brinjal Crop yield prediction using Shuffled shepherd optimization algorithm based ACNN-OBDLSTM model in Smart Agriculture. Journal of Integrated Science and Technology, 12(1), 710-710.
35. Haritha, K., Geethika, N. S., Venkateswarlu, K., Kumar, R. H., & Ramakrishna, Y. Enhancing Public Safety with AI & ML-Based CCTV Surveillance.
36. Haritha, K., Prakash, P. B., Pravallika, D., Venkatesh, K., & Venkatesh, G. Enhancing Object Detection in Autonomous Vehicles Under Low-Light Conditions Using Federated Learning and YOLOv5.
37. Ram, C. S., Vellela, S. S., Sravanthi Javvadi, D. V., Rashid, S. Z., & Madhumathi, S. M. (2025). Integrated Robotic–Imaging Platforms in Endovascular Surgery: Current Capabilities and Future Directions. Vascular and Endovascular Review, 8(16s), 285-298.
38. Roja, D., Navya, G., Srujana, B. S., Mamatha, P., & Sai, C. Y. K. Deep Learning for Hotel Reviews: A Framework for Sentiment Classification and Fake Review Detection.
39. Pakalapati, S., Rani, C. J., Vellela, S. S., Thanuja, N., & Bindu, M. N. H. (2025, November). Progressive GAN-based Framework for Realistic Image Generation and Style Transfer. In 2025 5th International Conference on Evolutionary Computing and Mobile Sustainable Networks (ICECMSN) (pp. 474-479). IEEE.
40. Balamanigandan, R., Vellela, S. S., Gorintla, S., Vuyyuru, L. R., Thanuja, N., & Rao, T. S. (2025, September). Quantum-Enhanced Data Security for Electronic Health Records: A Framework for Post-Quantum Cryptography in Healthcare Systems. In 2025 6th International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1924-1929). IEEE.
41. Roja, D., Amulya, P., Nagasai, M., Prasad, D. D., & Babu, A. V. Machine Learning-Based

Doi:10.48047/ijiee.2026.16.2.12

Early Diagnosis of Fish Diseases via Water Quality Data.

42. Sai, M. B., & Vellela, S. S. (2025, December). Hybrid ML Driven Multi-Cloud Service Work Load Prediction For Financial Systems. In 2025 1st International Conference on Advancement in Futuristic Technologies (ICAFT) (pp. 1-6). IEEE.
43. Kareemunnisa, D., Haritha, K., Ranjani, B. S., Venkateswarlu, K., & Bindu, M. N. H. DUAL-STAGE PRIVACY PROTECTION FOR GRAPH NEURAL NETWORKS AGAINST INFERENCE ATTACKS.
44. Mandava, R., Haritha, K., Vellela, S. S., Purimetla, N. R., Mohan, B. K., & Harinadh, T. (2025, June). Analysing User Perceptions of Trust in Financial Systems Using Explainable AI. In 2025 Second International Conference on Networks and Soft Computing (ICNSoC) (pp. 26-30). IEEE.

Doi:10.48047/ijiee.2026.16.2.12