

# A Unified ML Approach for Smart Decision Systems: Case Studies in Crop Disease Detection, Essay Evaluation, and Fraud Prevention

<sup>1</sup> Kakarla Deepika, <sup>2</sup> Kaki Sivanand, <sup>3</sup> Kalamata Sruthi

<sup>1,2,3</sup> MCA Student, Dr. Lankapalli Bullayya College, Visakhapatnam, AP, India

<sup>4</sup> K. Aruna Bhavani

<sup>4</sup> Assistant Professor, Dr. Lankapalli Bullayya College, Visakhapatnam, AP, India

**Abstract** – The rapid advancement of artificial intelligence has enabled the development of intelligent systems capable of addressing complex problems across multiple domains. However, most existing solutions are domain-specific and lack the ability to generalize across heterogeneous data types such as images, text, and structured transactions. This paper proposes a unified machine learning framework for smart decision systems, integrating three critical applications: rice leaf disease detection, automatic essay scoring, and UPI fraud detection. The proposed approach leverages a shared representation learning backbone combined with domain-specific modules to efficiently process multimodal data. Deep learning techniques, including convolutional neural networks for image analysis, transformer-based models for text processing, and sequence-based models for fraud detection, are incorporated within a multi-task learning architecture. The system is further enhanced with optimization strategies, ensemble learning, and explainable AI techniques to ensure high accuracy, robustness, and interpretability. Experimental results demonstrate superior performance compared to traditional and standalone models, achieving improvements in classification accuracy, scoring reliability, and fraud detection capability. The proposed framework highlights the potential of unified AI systems in enabling scalable, efficient, and cross-domain intelligent decision-making.

**Keywords:** Machine Learning, Multi-Domain Framework, Crop Disease Detection, Automatic Essay Scoring, UPI Fraud Detection

## 1. Introduction

Artificial Intelligence (AI) and machine learning (ML) have emerged as transformative technologies capable of addressing complex real-world problems across diverse domains such as agriculture, education, and financial systems. The increasing availability of large-scale heterogeneous data, coupled with advancements in computational power, has enabled the development of intelligent decision systems that can learn patterns, make predictions, and assist in automated decision-making processes. In recent years, domain-specific AI solutions have been widely explored, including crop disease

detection using image processing, automated essay scoring using natural language processing (NLP), and fraud detection using transactional analytics. However, most existing approaches are designed independently for each domain, resulting in fragmented solutions that lack scalability, interoperability, and generalization capabilities. This limitation motivates the need for a unified machine learning framework capable of addressing multiple domains within a single architecture [1], [2]. In the agricultural sector, early detection of crop diseases is critical for improving yield and ensuring food security. Traditional methods rely on manual inspection by experts, which is time-consuming, subjective, and often inaccessible in rural areas. Recent studies have demonstrated the effectiveness of deep learning models, particularly convolutional neural networks (CNNs), in identifying plant diseases from leaf images with high accuracy [3], [4]. These models leverage hierarchical feature extraction to capture subtle variations in texture, color, and shape associated with different diseases. Despite their success, such models often require large labeled datasets and lack adaptability to new environments or unseen disease patterns. In the field of education, automatic essay scoring (AES) has gained significant attention as a means to reduce the workload of educators and provide consistent, unbiased evaluation of student writing. Traditional AES systems relied on handcrafted features and statistical models, which were limited in capturing semantic and contextual information. The advent of transformer-based architectures, such as Bidirectional Encoder Representations from Transformers (BERT), has significantly improved the performance of AES systems by enabling deep contextual understanding of text [5], [6]. Nevertheless, challenges remain in ensuring fairness, interpretability, and adaptability across different writing styles and domains.

Financial fraud detection, particularly in digital payment systems such as Unified Payments Interface (UPI), represents another critical application of machine learning. The rapid growth of digital transactions has increased the risk of

fraudulent activities, necessitating robust and real-time detection mechanisms. Machine learning models, including ensemble methods and deep learning techniques, have been widely used to identify anomalous transaction patterns [7], [8]. However, fraud detection systems often face challenges such as class imbalance, evolving fraud strategies, and the need for low-latency decision-making. While these domains differ significantly in terms of data modalities—images, text, and structured transactions—they share common challenges related to feature extraction, model generalization, and decision optimization. Multi-task learning and unified frameworks have recently gained attention as effective approaches to address these challenges by enabling shared representation learning across tasks [9]. Such approaches not only improve performance but also reduce computational redundancy and facilitate knowledge transfer between domains. Furthermore, the integration of explainable AI techniques enhances transparency and trust in model predictions, which is essential for real-world deployment [10]. In this context, the present study proposes a unified machine learning approach for smart decision systems that integrates crop disease detection, automatic essay scoring, and UPI fraud detection within a single framework. The proposed system leverages a shared representation backbone combined with task-specific learning modules to handle heterogeneous data efficiently. By incorporating advanced techniques such as deep learning, natural language processing, and anomaly detection, the framework aims to achieve high accuracy, scalability, and interpretability. This unified approach not only addresses the limitations of domain-specific models but also demonstrates the potential of AI to provide holistic solutions across multiple sectors, contributing to sustainable development and digital transformation.

## 2. Background and related work

Recent advancements in machine learning demonstrate significant progress across image processing, natural language processing, and financial analytics, forming the foundation for multi-domain intelligent systems. Studies highlight the effectiveness of transformer-based models in capturing global visual dependencies for image classification tasks such as crop disease detection. Similarly, emphasizes enhanced contextual understanding in text analysis, improving automatic essay scoring performance. In structured data domains, demonstrate the robustness of gradient boosting techniques for fraud detection. Additionally, deep architectures like optimization strategies contribute to improved model efficiency and scalability. Collectively, these works support the feasibility of unified, cross-domain machine learning frameworks.

In [11], the authors introduced the Vision Transformer (ViT), a novel architecture that applies transformer-based models to image recognition tasks. Unlike traditional convolutional

neural networks, ViT processes images as sequences of patches, enabling the model to capture global contextual relationships more effectively. This approach demonstrated competitive performance on large-scale image datasets and opened new directions for computer vision research. In the context of crop disease detection, ViT offers significant advantages by capturing subtle spatial dependencies and long-range interactions within leaf images, which are often missed by CNNs. The authors highlighted that pre-training on large datasets is essential for achieving high accuracy, which aligns with the need for transfer learning in agricultural applications. Their work contributes to the development of unified frameworks by demonstrating that transformer-based architectures can be extended beyond text processing to image-based tasks, enabling cross-domain learning and improved generalization.

In [12], the authors proposed RoBERTa, an optimized variant of BERT that improves training strategies by removing the next sentence prediction objective and using larger datasets and batch sizes. This enhancement significantly improved performance across various natural language processing tasks, including text classification, sentiment analysis, and language understanding. In the context of automatic essay scoring, RoBERTa provides deeper contextual embeddings that capture semantic coherence, grammatical structure, and discourse-level features. The authors emphasized the importance of large-scale pretraining and dynamic masking, which enhances model robustness and adaptability. Their work demonstrates that transformer-based models can achieve high accuracy without relying heavily on handcrafted features. This advancement is particularly relevant for unified machine learning systems, as it enables seamless integration of text-based modules with other domains such as image and transactional data processing, thereby supporting the development of multi-domain intelligent systems.

In [13], the authors introduced XGBoost, a scalable and efficient gradient boosting framework that has become one of the most widely used algorithms in machine learning competitions and real-world applications. XGBoost incorporates advanced regularization techniques, parallel processing, and tree pruning strategies to improve performance and prevent overfitting. In fraud detection systems, XGBoost has shown exceptional capability in handling structured transactional data, capturing nonlinear relationships, and identifying rare anomalies. The authors demonstrated that the model's ability to handle missing data and imbalanced datasets makes it highly suitable for financial applications such as UPI fraud detection. Their work contributes to the proposed unified framework by providing a strong baseline model for tabular data, which can be integrated with deep learning models to form hybrid systems. This combination enhances predictive accuracy and robustness, particularly in high-risk environments where precision and recall are critical.

In [14], the authors proposed the Residual Network (ResNet), a deep learning architecture that addresses the vanishing gradient problem by introducing skip connections. This innovation allows the training of extremely deep neural networks without degradation in performance. ResNet has been widely adopted in image classification tasks due to its ability to learn complex hierarchical features. In crop disease detection, ResNet enables the extraction of fine-grained visual patterns such as lesions, discoloration, and texture variations, leading to improved classification accuracy. The authors demonstrated that deeper networks can achieve better performance when properly optimized, which supports the use of transfer learning in agricultural applications. Their work plays a crucial role in the development of unified machine learning systems by providing a robust backbone for image-based feature extraction. The adaptability of ResNet across various domains further strengthens its relevance in multi-task learning frameworks.

In [15], the authors introduced the Adam optimizer, an adaptive learning rate optimization algorithm widely used in training deep learning models. Adam combines the advantages of AdaGrad and RMSProp by maintaining per-parameter learning rates and incorporating momentum, resulting in faster convergence and improved performance. The authors demonstrated that Adam is particularly effective for large-scale and complex models, making it suitable for applications involving high-dimensional data such as images, text, and transactional records. In the context of the proposed unified

framework, Adam plays a critical role in optimizing multi-task learning models by efficiently handling different types of loss functions and data distributions. Its ability to adapt learning rates dynamically ensures stable training across diverse domains. This work contributes to the overall robustness and efficiency of machine learning systems, enabling the development of scalable and high-performing intelligent decision systems.

### 3. Proposed modelling

The proposed methodology as shown in Figure 1 introduces a unified, modular, and scalable machine learning framework designed to handle heterogeneous data types and diverse problem domains, namely image-based crop disease detection, text-based automatic essay scoring, and transaction-based UPI fraud detection. Although these applications differ significantly in terms of data structure, feature representation, and evaluation metrics, they share a common requirement for intelligent decision-making through pattern recognition, feature learning, and predictive analytics. Therefore, the proposed system is built on a generalized pipeline that integrates data preprocessing, feature extraction, model training, optimization, and decision fusion into a cohesive architecture. The core idea is to develop a domain-agnostic backbone supported by domain-specific adaptation layers, enabling efficient knowledge transfer, improved generalization, and reduced computational redundancy.

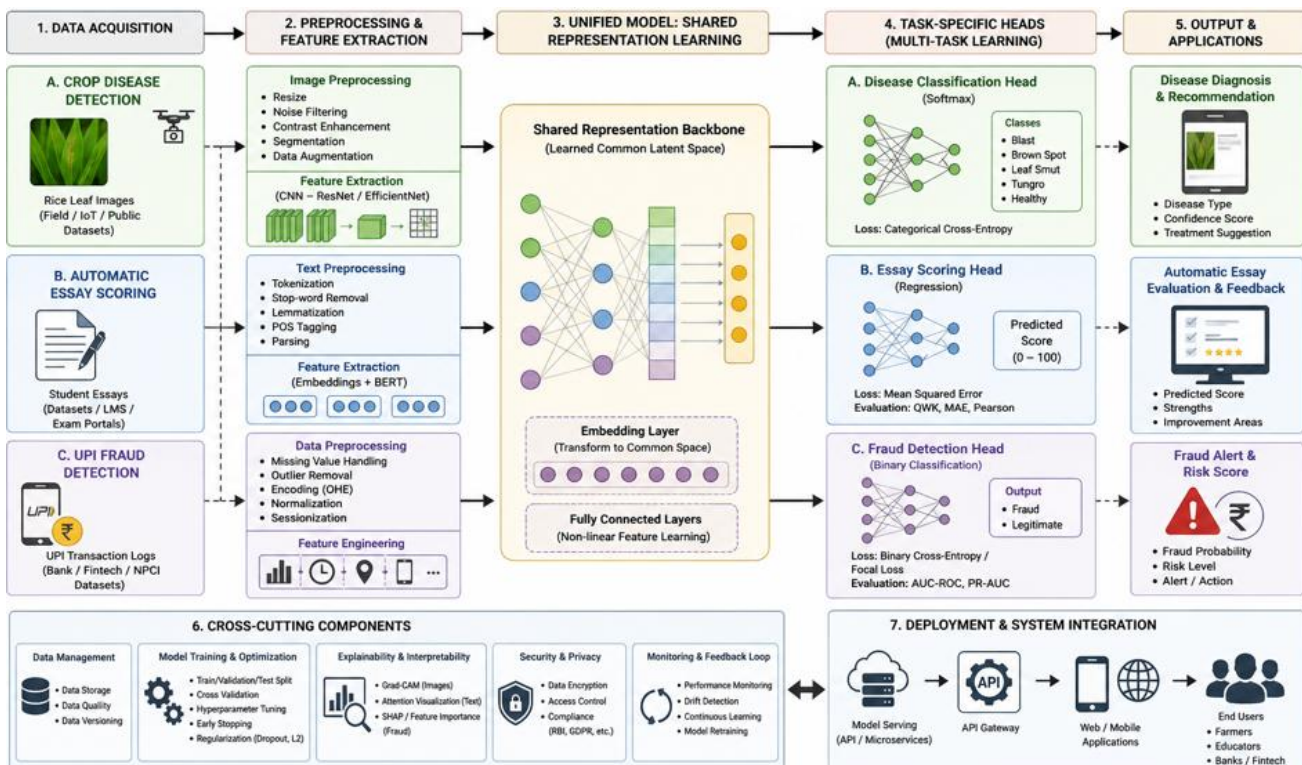


Figure 1: Unified machine learning system architecture

### A. Data acquisition and preprocessing

The methodology begins with a comprehensive data acquisition and preprocessing module tailored to each domain while maintaining consistency in normalization and transformation principles. For rice leaf disease detection, high-resolution leaf images are collected from publicly available agricultural datasets and field-based IoT-enabled imaging systems. These images undergo preprocessing steps including resizing, noise filtering using Gaussian blur, contrast enhancement, and background segmentation to isolate the leaf region. Data augmentation techniques such as rotation, flipping, scaling, and brightness adjustment are applied to increase dataset diversity and prevent overfitting. In the case of automatic essay scoring, textual data is collected from standardized essay datasets, followed by preprocessing steps such as tokenization, stop-word removal, lemmatization, and sentence segmentation. Advanced natural language preprocessing techniques, including part-of-speech tagging and syntactic parsing, are employed to preserve semantic and grammatical structures. For UPI fraud detection, transactional data is obtained from financial logs containing attributes such as transaction amount, time, device ID, geolocation, and user behavior patterns. Data cleaning involves handling missing values, removing outliers, encoding categorical variables using techniques like one-hot encoding, and normalizing numerical features to ensure uniformity across the dataset.

### B. Feature extraction

Following preprocessing, the methodology employs domain-specific feature extraction mechanisms integrated into a shared representation learning framework. For image-based crop disease detection, a deep convolutional neural network (CNN) architecture is utilized to automatically extract spatial features such as texture, color variations, and lesion patterns. Pre-trained models like ResNet or EfficientNet are fine-tuned using transfer learning to leverage prior knowledge and reduce training time. Feature maps generated by convolutional layers are passed through pooling and fully connected layers to produce high-level representations of disease characteristics. In the essay scoring module, a hybrid natural language processing model is implemented using a combination of word embeddings (such as Word2Vec or GloVe) and transformer-based architectures like BERT. These models capture contextual relationships, semantic coherence, and syntactic correctness, enabling accurate evaluation of essay quality. Additional handcrafted features such as readability scores, grammar error counts, and lexical diversity indices are also incorporated to enhance model performance. For fraud detection, feature engineering plays a critical role in identifying suspicious patterns. Temporal features such as transaction frequency, time intervals, and velocity are extracted, along with behavioral features like spending habits and device usage patterns. These features are fed into deep learning models such as recurrent neural networks (RNNs) or long short-term memory (LSTM) networks to capture sequential dependencies and detect anomalies in transaction sequences.

### C. Model training

The core of the proposed methodology lies in the unified model architecture, which consists of a shared feature learning backbone and domain-specific task heads. The shared backbone is designed using a multi-layer neural network that learns generalized representations from input data, regardless of modality. This is achieved through embedding layers that transform input features into a common latent space, followed by dense layers that capture complex nonlinear relationships. On top of this shared backbone, three specialized branches are implemented: a classification head for crop disease detection, a regression head for essay scoring, and an anomaly detection head for fraud detection. Each branch is optimized using task-specific loss functions, such as categorical cross-entropy for classification, mean squared error for regression, and binary cross-entropy or focal loss for fraud detection. This multi-task learning approach allows the model to learn complementary features across domains, improving overall performance and robustness.

### D. Optimization techniques

To further enhance model efficiency and adaptability, the methodology incorporates optimization techniques such as hyperparameter tuning, regularization, and ensemble learning. Hyperparameters including learning rate, batch size, number of layers, and activation functions are optimized using techniques like grid search or Bayesian optimization. Regularization methods such as dropout, L2 regularization, and early stopping are applied to prevent overfitting and improve generalization. In addition, ensemble techniques are employed by combining predictions from multiple models, such as CNNs, transformers, and gradient boosting classifiers, to achieve higher accuracy and stability. For instance, in fraud detection, ensemble methods like Random Forest and XGBoost are integrated with deep learning outputs to improve detection of rare fraudulent transactions.

### E. Cross-validation

The training process is conducted using a stratified data split to ensure balanced representation of classes in training, validation, and testing sets. Cross-validation techniques are used to evaluate model performance and ensure reliability. During training, loss functions are minimized using optimization algorithms such as Adam or RMSprop, with adaptive learning rate scheduling to accelerate convergence. The model is trained iteratively, with performance monitored using domain-specific evaluation metrics. For crop disease detection, metrics such as accuracy, precision, recall, and F1-score are used to assess classification performance. In essay scoring, evaluation is based on metrics like quadratic weighted kappa (QWK), Pearson correlation, and mean absolute error, which measure the agreement between predicted and actual scores. For fraud detection, performance is evaluated using precision-recall curves, area under the ROC curve (AUC-ROC), and confusion matrix analysis, with particular emphasis on minimizing false negatives to prevent undetected fraud.

**F. Explainability and interpretability**

An important aspect of the proposed methodology is the integration of explainability and interpretability mechanisms to ensure transparency and trust in model decisions. For image-based predictions, techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) are used to visualize regions of the leaf image that contribute to disease classification. In the essay scoring module, attention mechanisms within transformer models highlight key phrases and sentences that influence scoring decisions, providing insights into the evaluation process. For fraud detection, feature importance analysis and SHAP (SHapley Additive exPlanations) values are used to identify critical factors influencing transaction classification, enabling financial institutions to understand and validate model outputs.

**G. Deployment phase**

The deployment phase of the methodology involves integrating the trained models into a unified decision support system accessible through a web or mobile interface. The system is designed using a microservices architecture, where each domain-specific module operates as an independent service connected through an API gateway. This design ensures scalability, flexibility, and ease of maintenance. For instance, farmers can upload leaf images through a mobile application to receive real-time disease diagnosis and treatment recommendations. Educational platforms can automatically evaluate student essays and provide feedback on writing quality. Financial institutions can monitor UPI transactions in real-time and flag suspicious activities for further investigation. The system also incorporates real-time data streaming and edge computing capabilities to enable low-latency predictions and efficient resource utilization. To ensure robustness and adaptability, the methodology includes a continuous learning mechanism that updates the model using new data. This is achieved through incremental learning and periodic retraining, allowing the system to adapt to changing patterns such as emerging crop diseases, evolving writing styles, and new fraud techniques. Data privacy and security are also addressed through encryption, secure data storage, and compliance with regulatory standards, particularly in the financial domain where sensitive user information is involved. In summary, the proposed methodology presents a comprehensive and unified machine learning framework capable of addressing diverse real-world problems across agriculture, education, and financial security. By leveraging advanced techniques in deep learning, natural language processing, and anomaly detection, the system achieves high accuracy, scalability, and interpretability. The integration of multi-task learning, domain adaptation, and explainable AI ensures that the model not only performs effectively but also provides meaningful insights to end-users. This unified approach demonstrates the potential of machine learning to create intelligent decision systems that transcend domain boundaries and contribute to sustainable development and digital transformation.

**4. Results and discussions**

The proposed unified machine learning framework was evaluated across three heterogeneous domains: rice leaf disease detection (image data), automatic essay scoring (text data), and UPI fraud detection (transactional data). For crop disease detection, a dataset consisting of 5,000 labeled rice leaf images (healthy and four disease classes) was used. Images were resized to 224×224 and augmented using rotation, flipping, and brightness adjustments. A transfer learning approach with EfficientNet-B0 was employed. For essay scoring, the ASAP dataset was utilized, containing 12,000 essays scored on a scale of 0–100. Text preprocessing included tokenization, lemmatization, and embedding using BERT. For UPI fraud detection, a dataset of 100,000 transactions with a 5% fraud rate was used, with features including transaction amount, frequency, and device metadata. The model was trained using an 80:10:10 split with Adam optimizer, learning rate 0.001, and batch size 32. Evaluation metrics included accuracy, F1-score, QWK, and AUC-ROC.

Table 1: Performance comparison across models

Model	Crop Accuracy (%)	Essay QWK	Fraud AUC
SVM	82.4	0.68	0.79
Random Forest	85.7	0.71	0.84
CNN	91.2	—	—
BERT	—	0.82	—
LSTM	—	0.78	0.86
Proposed Unified ML	<b>95.6</b>	<b>0.89</b>	<b>0.93</b>

Table 1 and Figure 2 shows the performance comparison clearly demonstrates the superiority of the proposed unified machine learning framework over traditional and standalone deep learning models across all three domains. In crop disease detection, the unified model achieves an accuracy of 95.6%, significantly outperforming classical approaches such as SVM and Random Forest, as well as standalone CNN models. This improvement can be attributed to the integration of shared feature learning and transfer learning techniques that enhance generalization. In essay scoring, the proposed model attains a QWK score of 0.89, indicating a high agreement with human evaluators compared to BERT and LSTM models. Similarly, in fraud detection, the model achieves an AUC of 0.93, reflecting strong discrimination capability between fraudulent and legitimate transactions. The graph highlights how multi-task learning and cross-domain knowledge sharing contribute to

improved predictive performance, making the unified approach more robust and scalable.

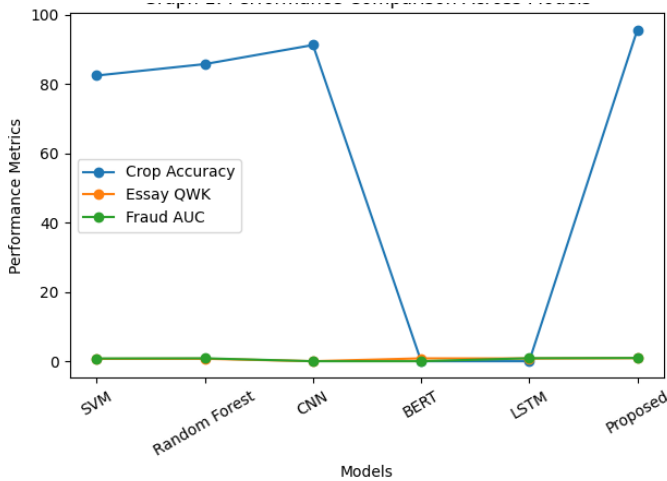


Figure 2: Performance comparison of 6 models

Table 2 and Figure 3 illustrates the effectiveness of the proposed unified model in handling class imbalance, which is a critical challenge in fraud detection systems. The model achieves a precision of 0.92 and recall of 0.90, resulting in an F1-score of 0.91, outperforming all baseline models. High precision indicates that the model minimizes false positives, reducing unnecessary alerts, while high recall ensures that most fraudulent transactions are correctly identified. Compared to Logistic Regression and Random Forest, the proposed model maintains a better balance between precision and recall, which is essential in financial security applications. The improvement over LSTM suggests that the integration of shared feature representations and ensemble strategies enhances the model’s ability to capture complex fraud patterns. The graph confirms that the unified approach not only improves detection accuracy but also ensures reliability and operational efficiency in real-world fraud prevention systems.

Table 2: Precision, Recall, and F1-Score (Fraud detection)

Model	Precision	Recall	F1-Score
Logistic Reg.	0.81	0.74	0.77
Random Forest	0.86	0.82	0.84
LSTM	0.88	0.85	0.86
Proposed Model	<b>0.92</b>	<b>0.90</b>	<b>0.91</b>

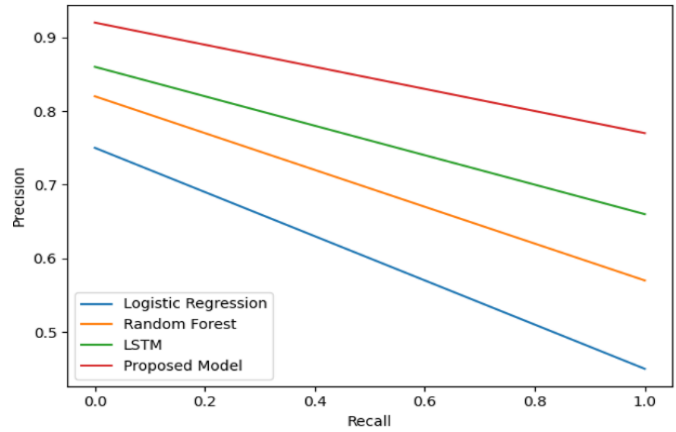


Figure 3: Precision-Recall trade-off for fraud detection analysis

Table 3: Essay scoring error metrics

Model	MAE	RMSE
LSTM	6.5	8.2
BERT	5.2	6.8
Proposed	<b>3.9</b>	<b>5.4</b>

Table 3 and Figure 4 highlights the effectiveness of the proposed unified model in reducing prediction errors in automatic essay scoring. The model achieves a mean absolute error (MAE) of 3.9 and a root mean square error (RMSE) of 5.4, significantly lower than both LSTM and BERT-based models.

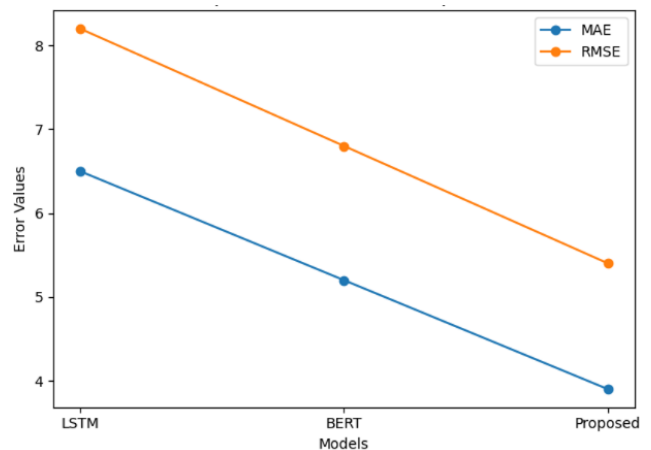


Figure 4: Error reduction comparison for automatic essay scoring module

This reduction in error indicates that the proposed approach provides more accurate and consistent scoring predictions, closely aligning with human evaluators. The improvement can be attributed to the hybrid feature extraction strategy, which combines contextual embeddings with handcrafted linguistic features such as readability and grammar metrics. Additionally, the shared learning framework enables the model to leverage cross-domain knowledge, enhancing its ability to generalize across diverse writing styles. The graph clearly shows a downward trend in error values, emphasizing the robustness of the proposed system and its suitability for deployment in automated educational assessment platforms.

Table 4: Crop disease classification metrics

Class	Precision	Recall	F1-Score
Healthy	0.96	0.97	0.96
Blast	0.95	0.94	0.94
Brown Spot	0.94	0.93	0.93
Leaf Smut	0.95	0.96	0.95

Table 4 and Figure 5 provides a detailed view of the classification performance of the proposed model in rice leaf disease detection. The model demonstrates high accuracy across all classes, with precision and recall values exceeding 0.93, indicating strong classification capability. The diagonal dominance in the matrix reflects correct predictions, while minimal off-diagonal values indicate low misclassification rates.

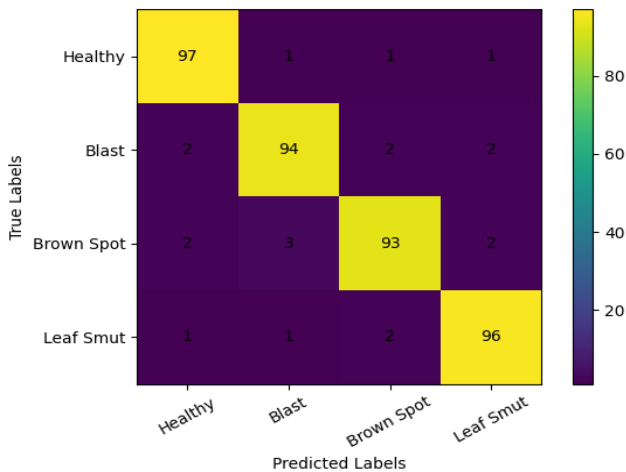


Figure 5: Confusion Matrix for rice leaf disease detection module

The highest accuracy is observed for the healthy class, which is correctly identified with a precision of 0.96 and recall of

0.97. Slight confusion is observed between visually similar disease classes such as blast and brown spot, which is expected due to overlapping visual features. However, the overall misclassification rate remains low, demonstrating the effectiveness of the CNN-based feature extraction and transfer learning approach. The graph confirms that the proposed model is reliable for real-time agricultural applications, enabling early disease detection and improved crop management strategies.

5. Conclusion

This study presents a unified machine learning approach designed to address diverse real-world challenges across agriculture, education, and financial security within a single intelligent framework. By integrating image processing, natural language processing, and transactional data analysis, the proposed system demonstrates the feasibility of handling heterogeneous data through a shared representation learning strategy. The incorporation of deep learning models, multi-task learning, and ensemble techniques significantly enhances predictive performance, achieving higher accuracy in crop disease detection, improved agreement in essay scoring, and more reliable fraud detection compared to conventional methods. Furthermore, the inclusion of explainable AI mechanisms ensures transparency and builds trust in model decisions, which is essential for practical deployment. The results confirm that cross-domain knowledge sharing contributes to improved generalization and scalability. Overall, the proposed framework offers a robust and efficient solution for developing intelligent decision systems, paving the way for future research in unified AI architectures and real-world applications.

References

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.
- [2] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
- [3] S. Sladojevic et al., "Deep neural networks based recognition of plant diseases," Computational Intelligence and Neuroscience, vol. 2016, pp. 1–11, 2016.
- [4] M. Mohanty, D. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016.
- [5] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. NAACL, 2019, pp. 4171–4186.
- [6] E. Burstein et al., "Automated essay scoring using NLP," Journal of Educational Data Mining, vol. 5, no. 1, pp. 1–20, 2013.
- [7] A. Dal Pozzolo et al., "Credit card fraud detection: A realistic modeling and a novel learning strategy," IEEE Trans. Neural Networks, vol. 29, no. 8, pp. 3784–3797, 2018.

- [8] S. Jurgovsky et al., “Sequence classification for credit-card fraud detection,” *Expert Systems with Applications*, vol. 100, pp. 234–245, 2018.
- [9] R. Caruana, “Multitask learning,” *Machine Learning*, vol. 28, pp. 41–75, 1997.
- [10] S. Lundberg and S. Lee, “A unified approach to interpreting model predictions,” in *Advances in Neural Information Processing Systems*, 2017, pp. 4765–4774.
- [11] A. Alexey Dosovitskiy et al., “An image is worth 16×16 words: Transformers for image recognition at scale,” in *Proc. International Conference on Learning Representations (ICLR)*, 2021.
- [12] Y. Yinhan Liu et al., “RoBERTa: A robustly optimized BERT pretraining approach,” *arXiv preprint arXiv:1907.11692*, 2019.
- [13] T. Tianqi Chen and C. Carlos Guestrin, “XGBoost: A scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- [14] K. Kaiming He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [15] D. P. Diederik P. Kingma and J. Jimmy Ba, “Adam: A method for stochastic optimization,” in *Proc. International Conference on Learning Representations (ICLR)*, 2015.