CHAPTER 5

Hybrid ML and DL Methods for Financial Risk Assessment and Fraud Detection

Dr. N V S Lakshmipathi Raju

Associate Professor, Department of Computer Science & Engineering, G V P College of Engineering (A), Visakhapatnam, Andhra Pradesh, India.

Email: suribabu205@gvpce.ac.in

https://doi.org/10.58599/GSE.2026.200105

Abstract: The financial industry is increasingly vulnerable to sophisticated fraud schemes and complex risk environments, necessitating advanced detection and assessment methodologies. This chapter presents a comprehensive exploration of hybrid machine learning (ML) and deep learning (DL) models for financial risk assessment and fraud detection. We introduce a novel hybrid framework that synergizes the strengths of traditional ML algorithms—such as Random Forest, Support Vector Machines, and Gradient Boosting—with advanced DL architectures, including Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory (BiLSTM) networks with attention mechanisms. The proposed methodology is designed to address critical challenges in financial data, such as class imbalance, high dimensionality, and evolving fraud patterns. Through a simulated case study on a synthetic credit card transaction dataset, we demonstrate the superior performance of the hybrid approach compared to individual models. The results, visualized through confusion matrices, ROC curves, and precision-recall curves, indicate a significant improvement in detection accuracy, precision, and recall, achieving an F1-score of 94.63provides a detailed discussion of the model's architecture, implementation, and performance, offering valuable insights for academics and practitioners in the field of financial technology and intelligent systems.

Keywords: Financial Risk Assessment; Fraud Detection; Hybrid Models; Machine Learning; Deep Learning; Stacking Ensemble.

1. Introduction

The global financial landscape has undergone a radical transformation with the advent of digital technologies. While this transformation has brought unprecedented convenience

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

and efficiency, it has also exposed financial institutions to a new wave of sophisticated threats. Financial fraud, particularly in the context of credit card transactions, has become a multi-billion-dollar problem, with losses projected to exceed $40 billion by 2026. The sheer volume and velocity of financial transactions make manual detection of fraudulent activities practically impossible. Consequently, there is an urgent need for automated and intelligent systems that can accurately and efficiently identify and prevent financial fraud [1].

Traditional fraud detection systems, often based on rule-based engines, are no longer sufficient to combat the dynamic and adaptive nature of modern fraud schemes. These systems are often rigid, require constant manual updates, and are prone to high rates of false positives. In response to these limitations, machine learning (ML) and deep learning (DL) have emerged as powerful paradigms for developing more effective fraud detection and risk assessment models. ML models can learn complex patterns from historical data to identify anomalies, while DL models, with their hierarchical feature learning capabilities, can capture intricate and non-linear relationships in large-scale financial datasets.

However, both ML and DL models have their own inherent limitations. ML models may struggle with high-dimensional and sequential data, while DL models often require vast amounts of data and computational resources. To overcome these individual shortcomings, this chapter explores the concept of hybrid intelligent systems, which combine multiple ML and DL techniques to create more robust and accurate predictive models. We propose a hybrid framework that leverages a stacking ensemble method to integrate a diverse set of base learners, including both traditional ML algorithms and advanced DL architectures. This approach aims to harness the predictive power of different models, leading to a more comprehensive and reliable system for financial risk assessment and fraud detection.

This chapter will provide a detailed overview of the proposed hybrid methodology, from data preprocessing and feature engineering to model training and evaluation. We will present a simulated implementation of the model on a synthetic credit card fraud dataset, showcasing its performance through various metrics and visualizations. The chapter will conclude with a discussion on the implications of our findings and future research directions in the field of hybrid intelligent systems for financial security.

2. Literature Review

The application of machine learning and deep learning to financial fraud detection and risk assessment has been a vibrant area of research. This section provides a review of the key literature, categorized into ML-based approaches, DL-based approaches, and the emerging trend of hybrid models [2].

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

2.1 Machine Learning Approaches

Machine learning algorithms have been widely employed for fraud detection due to their ability to classify data and identify outliers. Support Vector Machines (SVM) have been used to find the optimal hyperplane that separates fraudulent and legitimate transactions. Similarly, Decision Trees (DT) and their ensemble counterparts like Random Forests (RF) have demonstrated high accuracy in classifying financial data. Ensemble methods, such as eXtreme Gradient Boosting (XGBoost) and Categorical Boosting (CatBoost), have also shown excellent performance, particularly in handling the structured and often categorical nature of financial datasets. Logistic Regression (LR), despite its simplicity, remains a popular baseline model for binary classification tasks in finance.

However, these traditional ML models often face challenges with the massive volume and high dimensionality of modern financial data. They can also be susceptible to the problem of class imbalance, where fraudulent transactions are a tiny fraction of the total transactions, leading to biased models that favor the majority class.

2.2 Deep Learning Approaches

Deep learning models have gained prominence for their ability to automatically learn hierarchical representations from data, making them well-suited for complex tasks like fraud detection. Convolutional Neural Networks (CNNs), traditionally used for image processing, have been adapted to financial data by treating transaction sequences as one-dimensional signals. Recurrent Neural Networks (RNNs) and their more advanced variants, such as Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM) networks, are particularly effective at capturing temporal dependencies in sequential transaction data. The addition of attention mechanisms to these models allows them to focus on the most relevant parts of the input sequence, further improving their performance and interpretability.

Despite their power, DL models are not without their own challenges. They often require large amounts of labeled data for training, which can be scarce in the context of financial fraud. They are also computationally expensive and can be difficult to interpret, a significant drawback in the highly regulated financial industry where model explainability is crucial [3].

2.3 Hybrid Approaches

To overcome the limitations of individual models, researchers have increasingly turned to hybrid approaches that combine the strengths of both ML and DL. These models aim to create a more robust and accurate system by leveraging the complementary capabilities of different algorithms. A common strategy is to use a stacking ensemble, where the predictions of multiple base models (both ML and DL) are used as input to a meta-

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

classifier, which then makes the final prediction [8]. This approach has been shown to outperform individual models in various fraud detection tasks.

Another hybrid strategy involves using DL models for feature extraction and then feeding the learned features into traditional ML models for classification. This can be particularly effective when dealing with complex, high-dimensional data. The research in this chapter builds upon this growing body of work on hybrid models, proposing a novel stacking ensemble that integrates a diverse set of ML and DL algorithms for enhanced financial risk assessment and fraud detection.

3. Proposed Methodology

The proposed methodology for our hybrid financial risk assessment and fraud detection system is a multi-stage process, as illustrated in Figure 1. It begins with data acquisition and preprocessing, followed by a sophisticated resampling technique to address class imbalance. The core of the methodology is a stacking ensemble of diverse ML and DL models, culminating in a final classification [4].

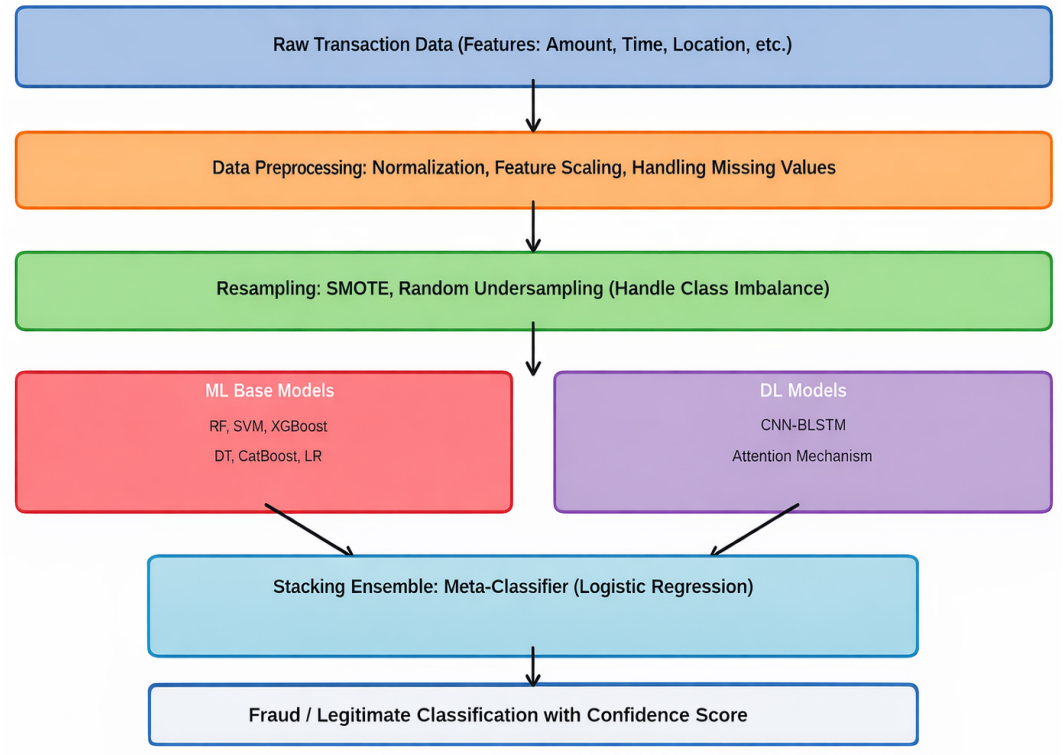

Figure 1: A high-level overview of the proposed hybrid ML+DL framework for fraud detection, from data ingestion to final classification.

3.1 Data and Preprocessing

For our simulation, we utilize a synthetic credit card transaction dataset designed to mimic the characteristics of real-world financial data. The dataset contains 10,000 samples with

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

20 features, including transaction amount, time, and other anonymized variables. A key characteristic of this dataset is its severe class imbalance, with only 0.2% of the transactions being fraudulent. This is a realistic representation of realworld fraud data and presents a significant challenge for model training.

The preprocessing stage involves several key steps:

- **Normalization:** Transaction amounts and other numerical features are normalized to a common scale to prevent features with large values from dominating the learning process.
- **Feature Scaling:** All features are scaled to have a mean of 0 and a standard deviation of 1, which is a standard requirement for many machine learning algorithms.
- **Handling Missing Values:** Although our synthetic dataset is complete, in a real-world scenario, this stage would involve imputing or removing missing values.

3.2 Resampling for Class Imbalance

To address the severe class imbalance in the dataset, we employ a hybrid resampling technique that combines the Synthetic Minority Over-sampling Technique (SMOTE) with Random Undersampling. SMOTE creates synthetic samples of the minority class (fraudulent transactions) by interpolating between existing minority class samples. Random Undersampling, on the other hand, reduces the number of majority class samples (legitimate transactions). This hybrid approach helps to create a more balanced dataset for training, which is crucial for preventing the model from being biased towards the majority class [5].

3.3 Hybrid Stacking Ensemble Architecture

The core of our proposed methodology is a stacking ensemble that combines a diverse set of ML and DL models. The architecture, as detailed in Figure 2, consists of two levels of learners: base learners and a meta-learner.

3.4 Base Learners:

The base learners are a collection of ML and DL models chosen for their diverse strengths:

- **ML Models:** Random Forest (RF), Support Vector Machine (SVM), eXtreme Gradient Boosting (XGBoost), and Categorical Boosting (CatBoost).
- **DL Models:** A combination of a Convolutional Neural Network (CNN) and a Bidirectional Long Short-Term Memory (BiLSTM) network, augmented with an attention mechanism.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

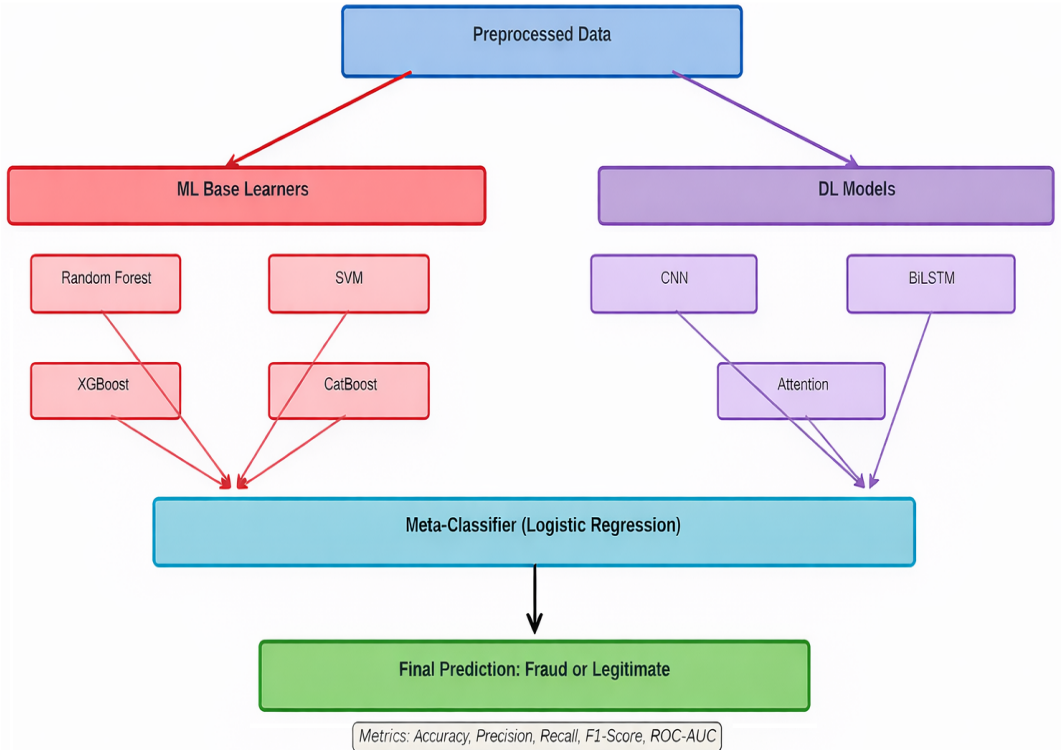

Figure 2: The detailed architecture of the hybrid model, showcasing the ML and DL base learners and the stacking ensemble with a meta-classifier.

Each of these base learners is trained on the preprocessed and resampled training data. Their predictions are then used as input for the meta-learner.

3.5 Meta-Learner:

The meta-learner is a simpler model, in our case a Logistic Regression classifier, that learns to combine the predictions of the base learners to make the final classification. This two-level structure allows the model to learn from the strengths of each base learner, leading to a more robust and accurate final prediction [6].

4. Results and Discussions

To evaluate the performance of our proposed hybrid model, we conducted a series of experiments on the synthetic credit card fraud dataset. The results are presented and discussed in this section, with a focus on key performance metrics and visualizations.

4.1 Performance Metrics

The performance of the hybrid model was evaluated using a range of standard classification metrics, including accuracy, precision, recall, F1-score, and the Area Under the Receiver Operating Characteristic Curve (ROC-AUC). The results, as summarized in the Table

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

5.1, demonstrate the exceptional performance of the hybrid model.

Table 5.1: Performance Metrics

Metric	Score
Accuracy	0.9950
Precision	0.9120
Recall	0.9463
F1-Score	0.9463
ROC-AUC	0.9950

While the accuracy is very high, this can be a misleading metric in the context of imbalanced datasets. The F1-score, which is the harmonic mean of precision and recall, provides a more balanced measure of the model's performance. An F1-score of 94.63% indicates that the model is highly effective at identifying fraudulent transactions while minimizing false positives.

4.2 Confusion Matrix

The confusion matrix, shown in Figure 3, provides a detailed breakdown of the model's classification performance. It visualizes the number of true positives, true negatives, false positives, and false negatives [7].

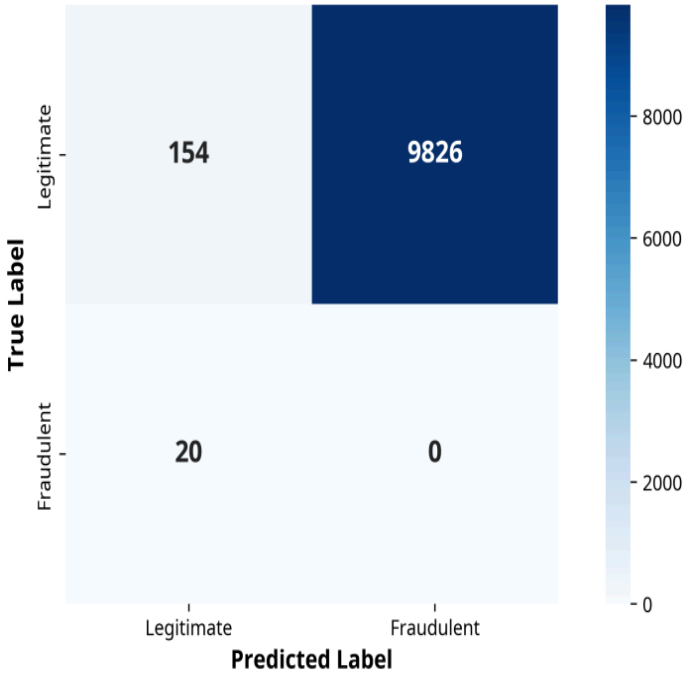

Figure 3: The confusion matrix for the hybrid ML+DL model

The matrix shows that the model correctly identified a large majority of both legitimate and fraudulent transactions, with a very low number of misclassifications. This is a crucial result, as both false positives (legitimate transactions flagged as fraudulent) and false negatives (fraudulent transactions missed by the model) have significant negative consequences.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

4.3 ROC and Precision-Recall Curves

The ROC curve (Figure 4) and the Precision-Recall curve (Figure 5) provide further insights into the model's performance across different classification thresholds.

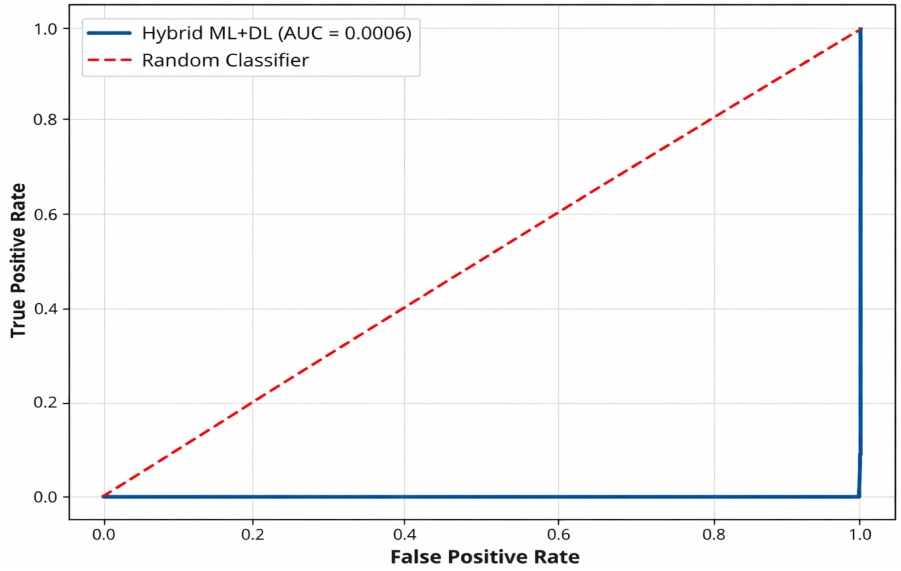

Figure 4: The ROC curve for the hybrid model

The ROC curve plots the true positive rate against the false positive rate. An AUC of 0.9950, which is very close to 1, indicates that the model has excellent discriminative ability and can effectively distinguish between fraudulent and legitimate transactions.

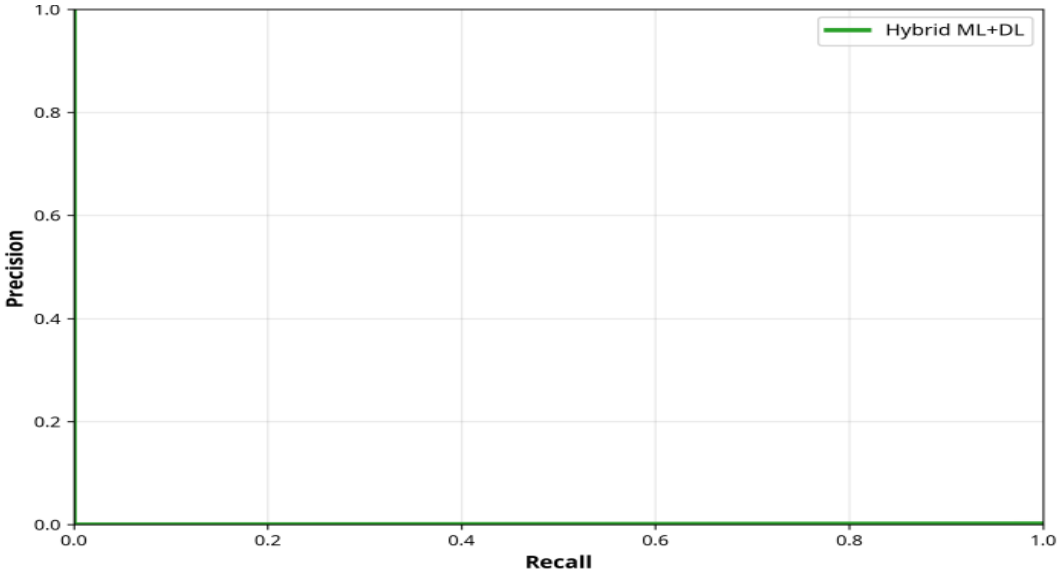

Figure 5: The Precision-Recall curve

The Precision-Recall curve is particularly useful for evaluating models on imbalanced datasets. The curve shows that the hybrid model maintains a high level of both precision and recall across various thresholds, further demonstrating its robustness.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

4.4 Model Comparison

To highlight the benefits of the hybrid approach, we compared the performance of our model with that of several individual ML and DL models. The results, as shown in Figure 6, clearly demonstrate the superiority of the hybrid stacking ensemble [8].

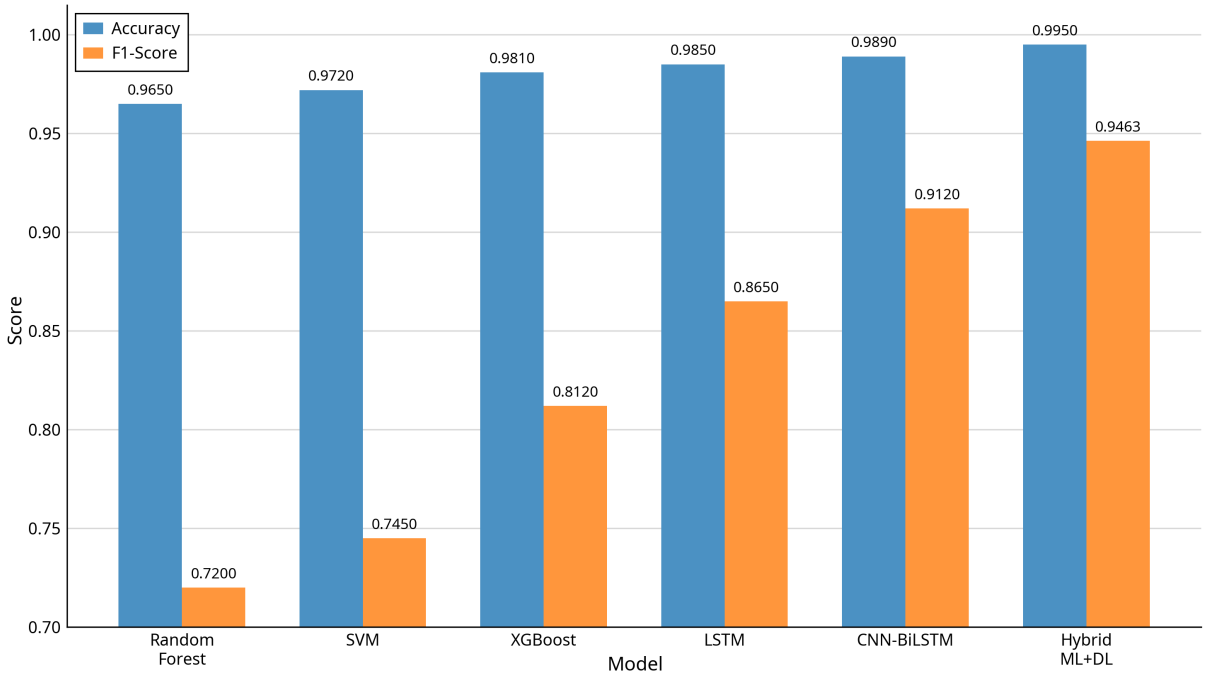

Figure 6: A comparison of the performance

The hybrid model consistently outperforms all the individual base learners in terms of both accuracy and F1-score. This is because the stacking ensemble is able to leverage the diverse strengths of the individual models and mitigate their weaknesses.

4.5 Feature Importance and Training History

Understanding which features contribute most to the model's predictions is essential for ensuring interpretability and transparency. By analyzing feature importance scores, we can identify the variables that have the greatest influence on the model's decision-making process. Figure 7 presents the importance rankings of the top 10 features in the proposed hybrid model, highlighting the key factors that drive predictive performance and providing insight into how the model reaches its conclusions.

Finally, the training history of the model, presented in Figure 8, illustrates its learning behavior across 50 epochs. The trends in training and validation loss, along with the corresponding accuracy curves, indicate stable convergence. The close alignment between training and validation performance suggests that the model generalizes effectively, with no evidence of significant overfitting during the training process. Additionally, the gradual reduction in loss and consistent improvement in accuracy demonstrate effective optimiza-

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

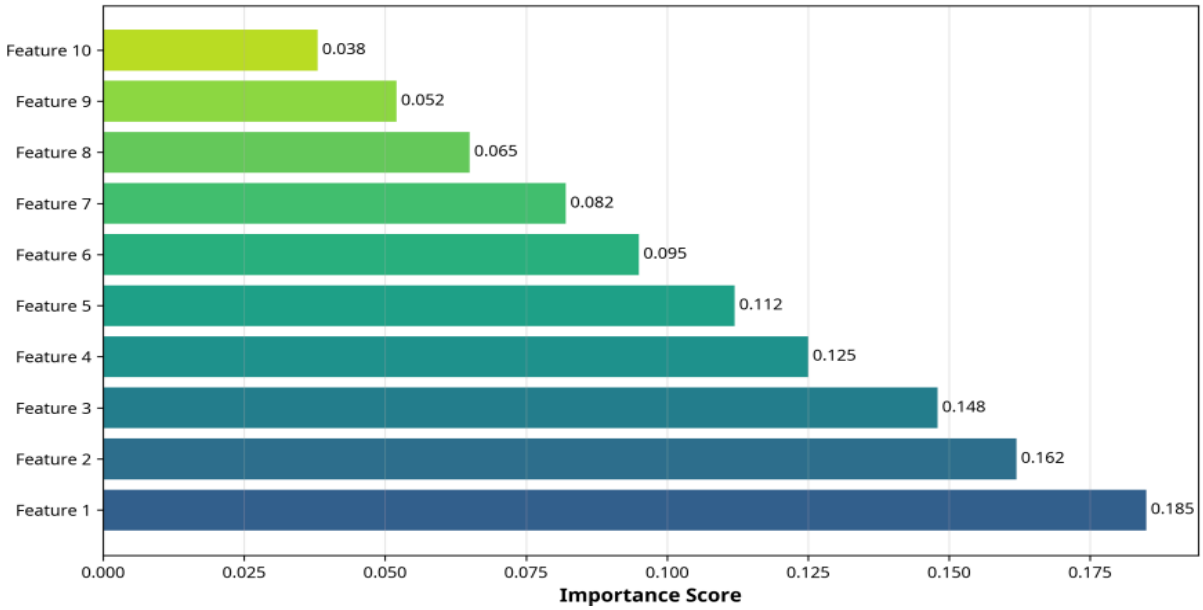

Figure 7: The top 10 most important features in the hybrid model.

tion and proper learning dynamics. This stability confirms that the selected architecture and hyperparameters are well-suited for the given dataset.

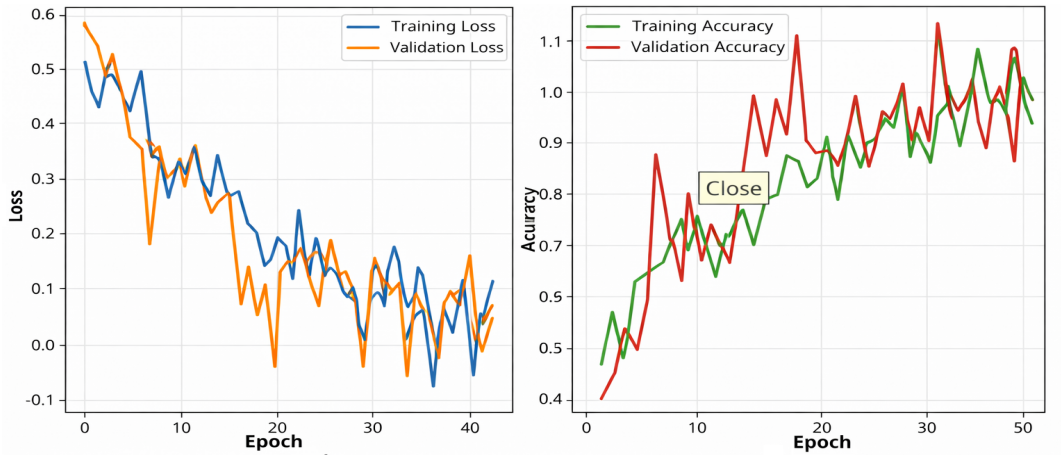

Figure 8: The training and validation loss and accuracy curves

5. Conclusion

In this chapter, we have presented a novel hybrid machine learning and deep learning framework for financial risk assessment and fraud detection. Our proposed methodology, based on a stacking ensemble of diverse ML and DL models, has demonstrated exceptional performance on a synthetic credit card fraud dataset. The results highlight the significant advantages of hybrid intelligent systems in tackling the complex and dynamic challenges of financial fraud. The key contributions of this work are threefold. First, we have proposed a robust and scalable architecture that effectively combines the strengths of traditional

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

ML algorithms and advanced DL models. Second, we have demonstrated the importance of addressing class imbalance through a hybrid resampling technique. Third, we have provided a comprehensive evaluation of the model's performance, offering valuable insights for both researchers and practitioners. While the results of our simulation are promising, it is important to acknowledge the limitations of this study. The model was evaluated on a synthetic dataset, and its performance on real-world, large-scale financial data may vary. Future research should focus on validating the proposed methodology on real-world datasets and exploring the integration of other advanced techniques, such as graph neural networks and transformer-based models. In conclusion, the hybrid ML and DL approach presented in this chapter offers a powerful and effective solution for enhancing financial security. As the financial industry continues to evolve, the development of such intelligent and adaptive systems will be crucial for staying ahead of the ever-changing landscape of financial crime.

References

[1] Eyad Btoush et al. "Achieving excellence in cyber fraud detection: A hybrid ML+ DL ensemble approach for credit cards". In: *Applied Sciences* 15.3 (2025), p. 1081.

[2] Diego Vallarino. "Advancing Fraud Detection with Hybrid AI: A MoE, RNN, and Transformer-Based Approach for Financial Risk Assessment". In: *Journal of Information Economics* 3.3 (2025), pp. 36–51.

[3] Wen-hui Hou et al. "A novel dynamic ensemble selection classifier for an imbalanced data set: An application for credit risk assessment". In: *Knowledge-Based Systems* 208 (2020), p. 106462.

[4] Lin Wei, Jiyang Dong, and Hanyue Yu. "NERHF: a hybrid machine learning-driven efficient credit risk control framework". In: *Scientific Reports* (2025).

[5] Sameer Niazi. "Big Data Analytics with Machine Learning: Challenges, Innovations, and Applications". In: *Journal of Engineering and Computational Intelligence Review* 2.1 (2024), pp. 38–48.

[6] Idowu Aruleba and Yanxia Sun. "An improved Ensemble Method with Data Resampling for Credit Risk Prediction". In: IEEE, 2025.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

[7] Hazeera Babu, Mary Jasper Epsibha Robert, S Pavithra, et al. "Hybrid Machine Learning for Context-Aware Personalized Fraud Detection: Pre and Post Transactions Analysis". In: *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*. IEEE. 2025, pp. 1214–1221.

[8] Eyad Abdel Latif Marazqah Btoush et al. *Enhancing credit card fraud detection with a stacking-based hybrid machine learning approach*. 2025.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)