Journal of Economic Theory and Business Management
Journal Home: www.jetbm.online
Vol. 1, No. 4, 2024 | ISSN 3006-4953 (Print) | ISSN 3006-4961 (Online)

Utilizing Deep Learning to Detect Fraud in Financial Transactions and Tax Reporting

XU, Haosen [1*] LI, Siyang [2] NIU, Kaiyi [3] PING, Gang [4]

[1] University of California, Berkeley, USA
[2] Johns Hopkins University, USA
[3] Royal Holloway University of London, UK
[4] Hong Kong Metropolitan University, China

**** XU, Haosen is the corresponding author, E-mail: rexcarry036@gmail.com***

Abstract: This study investigates the application of deep learning techniques and automated feature engineering in credit card fraud detection. The research employs a Convolutional Neural Network (CNN) model integrated with Deep Feature Synthesis (DFS) to enhance the accuracy and efficiency of fraud detection systems. A comprehensive dataset of credit card transactions is utilized, comprising 2,453,620 records with 1,432 fraudulent cases. The methodology involves extensive data preprocessing, including class imbalance handling and feature encoding. The DFS algorithm generates 118 new features from the original 25, capturing complex relationships within the data. The CNN model's performance is compared against traditional machine learning algorithms such as Logistic Regression, Random Forest, Support Vector Machine, and XGBoost. Results demonstrate that the CNN model with DFS-generated features significantly outperforms other approaches, achieving an accuracy of 91%, precision of 92%, recall of 90%, and an F1-score of 0.91. The study highlights the synergistic effect of combining deep learning architectures with advanced feature engineering techniques in addressing the challenges of credit card fraud detection. The findings contribute to developing more robust and adaptable fraud detection systems, potentially reducing financial losses and enhancing security in electronic transactions. Future research directions include exploring model interpretability, real-time application, and extension to multi-class fraud detection scenarios.

Keywords: Credit Card Fraud Detection, Deep Learning, Convolutional Neural Networks, Deep Feature Synthesis.

DOI: https://doi.org/10.5281/zenodo.13294459 **ARK:** https://n2t.net/ark:/40704/JETBM.v1n4a08

1 INTRODUCTION

1.1 RESEARCH BACKGROUND

Credit card fraud has become a significant concern in the rapidly evolving landscape of digital financial transactions. The widespread adoption of electronic payment systems has led to an unprecedented increase in the volume of online transactions, creating new opportunities for fraudulent activities [20]. According to the World Bank, two-thirds of adults now use digital platforms for financial transactions, highlighting the magnitude of this issue [1]. The financial implications of credit card fraud are substantial, with the Nilson Report indicating losses of $28.58 billion in 2020, equivalent to 6.8 cents per $100 of purchase volume. This escalating trend necessitates the development of robust and efficient fraud detection systems to safeguard financial institutions, consumers, and the integrity of electronic payment ecosystems.

Traditional fraud detection methods have proven inadequate in addressing modern fraudulent activities' sophisticated and dynamic nature. Machine learning and deep learning techniques have emerged as promising solutions, offering enhanced capabilities in pattern recognition and anomaly detection [8]. These advanced methodologies can adapt to evolving fraud patterns and process large volumes of transaction data in real time, significantly improving the accuracy and efficiency of fraud detection systems.

The effectiveness of machine learning models in fraud detection heavily relies on the quality and relevance of input features. Feature engineering, both manual and automated, plays a crucial role in extracting meaningful information from raw transaction data. Recent advancements in automated feature engineering techniques, such as Deep Feature Synthesis (DFS), have shown potential in generating complex and informative features from relational datasets. These techniques are less labor-intensive and more adaptable to changing data patterns than manual feature selection methods.

1.2 PROBLEM STATEMENT

Despite the advancements in machine learning and deep learning techniques for fraud detection, several challenges

persist. The highly imbalanced nature of credit card transaction datasets, where fraudulent transactions represent a tiny fraction of the total volume, poses significant difficulties in model training and evaluation. This imbalance often leads to biased models with high false favorable rates, potentially causing inconvenience to legitimate customers and operational inefficiencies for financial institutions.

Another critical challenge is the dynamic nature of fraudulent activities. Fraudsters continually adapt their strategies to evade detection, making static models obsolete. This necessitates the development of adaptive systems capable of continuously learning from new data and evolving fraud patterns. Additionally, the interpretability of complex machine learning models, particularly deep learning architectures, remains a concern. Regulatory requirements and the need for transparency in decision-making processes demand explainable AI solutions in fraud detection systems.

Integrating automated feature engineering techniques with deep learning models for credit card fraud detection has not been thoroughly explored. While Deep Feature Synthesis has shown promise in improving the performance of traditional machine learning models, its impact on deep learning architectures in fraud detection remains an open area of research.

1.3 Research Objectives

This study aims to address the challenges above and explore the synergy between automated feature engineering and deep learning in credit card fraud detection. The primary objectives of this research are:

To investigate the effectiveness of Deep Feature Synthesis in generating relevant and informative features from credit card transaction datasets. This objective seeks to evaluate the potential of automated feature engineering in capturing complex relationships within transaction data that may indicate fraudulent activities.

To develop and evaluate a deep learning model for credit card fraud detection, specifically a Convolutional Neural Network (CNN). This model will leverage the high-dimensional feature space created through automated feature engineering techniques.

To conduct a comparative analysis of the performance of the proposed CNN model with and without the integration of DFS-generated features. This comparison aims to quantify the impact of automated feature engineering on the accuracy, precision, recall, and F1-score of the fraud detection model.

To assess the scalability and adaptability of the proposed approach in handling large volumes of transaction data and evolving fraud patterns. This objective addresses the practical considerations of implementing such a system in real-world financial environments.

To explore the interpretability aspects of the combined DFS and deep learning approach, aiming to provide insights into the decision-making process of the fraud detection model. This aligns with the growing demand for explainable AI in critical financial applications.

2 LITERATURE REVIEW

2.1 Traditional Fraud Detection Methods

Financial institutions have long employed traditional fraud detection methods to identify suspicious activities in credit card transactions. These methods primarily rely on rule-based systems and statistical techniques. Rule-based systems operate on predefined rules established by domain experts and regulatory norms. These rules typically involve transaction amount thresholds, geographical location checks, and frequency of transactions. While these systems offer consistency and transparency, they often need help to adapt to fraudulent activities' dynamic and evolving nature.

Statistical techniques, including regression analysis and time series modeling, have also been used to detect transaction pattern anomalies. These methods aim to establish baseline behavior for individual cardholders and flag deviations from these patterns. Despite their simplicity and interpretability, traditional methods often fail to capture complex fraud patterns and suffer from high false-positive rates, leading to operational inefficiencies and customer dissatisfaction.

2.2 Machine Learning Methods

2.2.1 Supervised Learning Techniques

The advent of machine learning has significantly advanced the field of fraud detection. Supervised learning techniques have gained prominence due to their ability to learn from historically labeled data. These methods include logistic regression, decision trees, random forests, and support vector machines (SVM) [3]. Logistic regression models have been widely used for their simplicity and interpretability, providing probability scores for fraudulent transactions. Decision trees and random forests offer the advantage of capturing non-linear relationships in the data and have shown robust performance in fraud detection tasks [18].

Support Vector Machines have demonstrated high accuracy in separating fraudulent from legitimate transactions by finding optimal hyperplanes in high-dimensional feature spaces. K-Nearest Neighbors (KNN) algorithms have also been applied, leveraging the similarity between transactions to identify potential fraud. These supervised techniques have improved performance over traditional methods, with studies reporting accuracy rates ranging from 94% to 99% for various algorithms [33].

2.2.2 Unsupervised Learning Techniques

Unsupervised learning techniques address the challenge of detecting novel fraud patterns without labeled data. These

methods focus on identifying anomalies or outliers in transaction data. Clustering algorithms, such as K-means and hierarchical clustering, have grouped similar transactions and isolated unusual patterns. Isolation Forests have gained attention for their ability to detect anomalies efficiently in high-dimensional spaces [34].

Autoencoders, a type of neural network, have been utilized for dimensionality reduction and anomaly detection in credit card transactions. By learning to reconstruct input data, autoencoders can identify transactions that deviate significantly from standard patterns [35]. While unsupervised techniques offer the potential to detect previously unknown fraud patterns, they often require careful tuning and domain expertise to distinguish between genuine anomalies and fraudulent activities.

2.3 Deep Learning in Fraud Detection

Deep learning models have emerged as powerful tools for credit card fraud detection, capable of automatically learning hierarchical representations from raw transaction data [2,21]. Convolutional Neural Networks (CNNs), traditionally used in image processing, have been adapted for fraud detection by treating transaction features as one-dimensional inputs [12,22]. CNNs have shown promising results in capturing local patterns and spatial relationships within transaction data [13].

Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, have been applied to model sequential dependencies in transaction histories. These models excel at capturing temporal patterns and have demonstrated high accuracy in identifying fraudulent sequences of transactions. Studies have reported performance improvements with deep learning models, achieving accuracy rates above 99% and significantly reducing false favorable rates compared to traditional machine learning approaches.

2.4 Feature Engineering Techniques

Feature engineering plays a crucial role in enhancing the performance of fraud detection models. Manual feature engineering involves domain experts crafting relevant features based on their understanding of fraudulent behaviors. This process, while valuable, is time-consuming and may only capture some complex relationships within the data.

Automated feature engineering techniques have gained traction, offering scalable and data-driven approaches to feature creation. Profound Feature Synthesis (DFS) is a promising method that automatically generates features from relational datasets linked by unique identifiers [5]. DFS applies a series of mathematical transformations to create new features, potentially uncovering non-obvious relationships in the data. Studies have shown that DFS can significantly improve model performance, with one study reporting a 23% increase in accuracy for car loan fraud detection [36].

2.5 Challenges in Credit Card Fraud Detection

Despite advancements in machine learning and deep learning techniques, credit card fraud detection faces several persistent challenges [4]. Class imbalance remains a significant issue, with fraudulent transactions typically representing less than 1% of total transactions. This imbalance can lead to biased models and poor detection of minority class instances [37]. Techniques such as oversampling, undersampling, and synthetic data generation (e.g., SMOTE) have been explored to address this issue, but finding the right balance remains challenging.

The dynamic nature of fraud patterns poses another significant challenge. Fraudsters continuously adapt their strategies, rendering static models obsolete over time. This necessitates the development of adaptive models capable of learning from new data and evolving fraud patterns. Concept drift, where the statistical properties of the target variable change over time, further complicates the maintenance of accurate models.

Data quality and availability present additional hurdles. Privacy concerns and regulatory restrictions often limit access to comprehensive transaction data, impacting model training and evaluation [38]. Furthermore, transaction data's high dimensionality and irrelevant or redundant features can hinder model performance and interpretability.

Interpretability of complex models, particularly deep learning architectures, remains a concern in the financial sector. Regulatory requirements and the need for transparency in decision-making processes demand explainable AI solutions. Balancing model complexity and interpretability remains an active area of research in fraud detection.

3 METHODOLOGY

3.1 Dataset Description

This study utilizes a comprehensive credit card transaction dataset from an open-source platform. The dataset comprises two primary tables: credit_card_transactions-ibm_v2.csv (transaction file) and sd254_cards.csv (card file), containing 15 and 13 fields, respectively. The transaction file encompasses 2,453,620 records, of which 1,432 are labeled as fraudulent transactions. The card file contains information on 416 unique credit cards. Table 1 provides a detailed overview of the dataset structure:

Table 1: Dataset Structure

Table Name	Number of Records	Number of Fields	Fraudulent Transactions
Transaction	2,453,620	15	1,432

Card	416	13	N/A

The transaction table includes crucial information such as transaction amount, timestamp, merchant category, and location. The card table provides supplementary data on card types, credit limits, and user demographics. This rich dataset allows for comprehensive analysis and model development for fraud detection.

3.2 Data Preprocessing

Data preprocessing is critical in preparing the dataset for analysis and model training. The process involves several key stages:

Data Selection: To establish relational links between tables, transactions are grouped by payment cards from sd254cards.csv, and corresponding transactions from credit_card_transactions-ibm_v2.csv are selected for each card user.

Data Cleansing: Null values are replaced with zeros, and invalid characters (e.g., dollar signs in transaction amounts) are removed to ensure data consistency.

Handling Class Imbalance: The dataset exhibits severe class imbalance, with a ratio of 99.942:0.058 between legitimate and fraudulent transactions. An undersampling technique is employed to address this, randomly selecting legitimate transactions to match the number of fraudulent transactions (1,432 each).

Data Encoding: Categorical features are transformed into numerical values using Label Encoding, facilitating their use in machine learning models.

Data Splitting: The preprocessed dataset is divided into training and test sets with an 80:20 ratio, maintaining the balance between fraudulent and legitimate transactions in both sets. Table 2 summarizes the data preprocessing steps and their outcomes:

Table 2: Data Preprocessing Summary

Preprocessing Step	Input	Output
Data Selection	2,453,620 transactions	416 card-grouped transaction sets
Data Cleansing	Raw transaction data	Cleaned dataset with consistent formats
Handling Class Imbalance	1,432 fraudulent, 2,452,188 legitimate	1,432 fraudulent, 1,432 legitimate
Data Encoding	Categorical features	Numerical encoded features
Data Splitting	2,864 preprocessed records	2,291 training, 573 test records

3.3 Feature Engineering

3.3.1 Manual Feature Selection

Manual feature selection is performed based on domain knowledge and exploratory data analysis. Key features selected include transaction amount, time since the last transaction, merchant category code, and geographical location. Additionally, derived features such as transaction frequency and average transaction amount per merchant category are created to capture user behavior patterns. Table 3 lists the manually selected and derived features:

Table 3: Manually Selected and Derived Features

Feature Type	Feature Name	Description
Original	Transaction Amount	Monetary value of the transaction
Original	Time Since Last Transaction	Time elapsed since the last transaction
Original	Merchant Category Code	Code identifying the merchant type
Original	Transaction Location	Geographical coordinates of the transaction
Derived	Transaction Frequency	Number of transactions in a defined time window
Derived	Avg. Transaction per Merchant	Average transaction amount for each merchant category

3.3.2 Deep Feature Synthesis (DFS)

Profound Feature Synthesis automatically generates complex features from the relational dataset. The DFS algorithm uses various mathematical transformations to create features based on the relationships between the card and transaction tables.

The DFS process generates 118 new features from the original 25 features (15 from transaction data and 10 from card data). These features capture complex relationships and temporal patterns within the data.

Figure 1 illustrates the DFS feature generation process. It depicts a flowchart showing the input of raw features from both tables, the application of mathematical transformations, and the output of synthesized features. The diagram highlights how DFS leverages the relational structure of the data to create meaningful, high-level features.

Journal of Economic Theory and Business Management
Journal Home: www.jetbm.online
Vol. 1, No. 4, 2024 | ISSN 3006-4953 (Print) | ISSN 3006-4961 (Online)

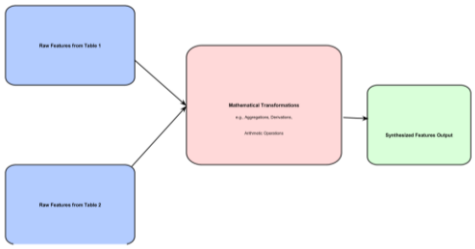

FIGURE 1:DEEP FEATURE SYNTHESIS WORKFLOW

3.4 MODEL ARCHITECTURE

3.4.1 Convolutional Neural Network (CNN)

A 1-dimensional Convolutional Neural Network (1D-CNN) is implemented as the primary model for fraud detection. The CNN architecture is designed to capture local patterns and spatial relationships within the transaction data.

The model structure follows the input layer, which is shaped to match the feature dimensions. Convolutional Layer 1:32 filters, kernel size 2x2, ReLU activation.Convolutional Layer 2: 64 filters, kernel size 2x2, ReLU activation.Flatten Layer.Dense Layer 1: 64 neurons, ReLU activation.Dropout Layer: 50% dropout rate.Output Layer: 1 neuron, sigmoid activation.

The model is compiled using binary cross-entropy loss and Adam optimizer with a learning rate 0.0001. Training is performed over 20 epochs with a batch size of 32.

Provides a visual representation of the model structure. The diagram illustrates the layer-by-layer composition of the network, including the dimensions and connections between layers.

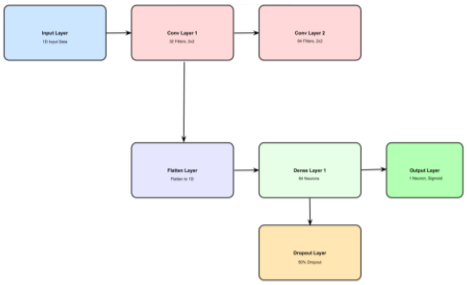

FIGURE 2:CNN MODEL ARCHITECTURE FOR FRAUD DETECTION

3.4.2 Other Comparative Models

To evaluate the effectiveness of the CNN model, several other machine learning models are implemented for comparison: Logistic Regression: A baseline linear model for binary classification. Random Forest: An ensemble learning method using 100 decision trees.Support Vector Machine (SVM): Implemented with a radial basis function kernel.XGBoost: A gradient-boosting algorithm known for its performance in various classification tasks.

These models are trained on the manually selected and DFS-generated features to assess the impact of feature engineering on different algorithms. Table 4 summarizes the hyperparameters used for each comparative model:

TABLE 4: HYPERPARAMETERS OF COMPARATIVE MODELS

Model	Hyperparameters
Logistic Regression	C=1.0, max_iter=100
Random Forest	n_estimators=100, max_depth=10
SVM	C=1.0, kernel='rbf', gamma='scale'
XGBoost	max_depth=6, learning_rate=0.1, n_estimators=100

To visualize the performance comparison across different models. This figure presents a multi-axis plot comparing each model's accuracy, precision, recall, and F1 score. The x-axis represents the different models, while multiple y-axes show the performance metrics. This visualization allows for a comprehensive comparison of model performances across various evaluation criteria.

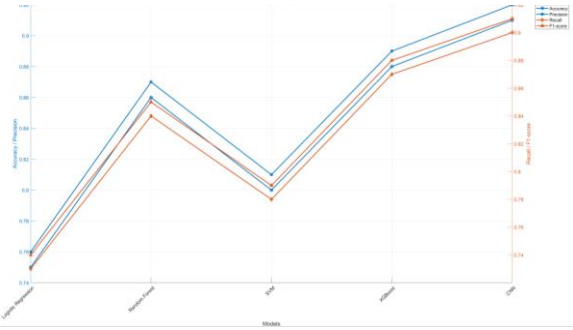

FIGURE 3: MODEL PERFORMANCE COMPARISON

The methodology outlined in this section provides a robust framework for developing and evaluating fraud detection models. By combining advanced feature engineering techniques with state-of-the-art deep learning architectures, this approach aims to effectively address the challenges of credit card fraud detection.

4 EXPERIMENTAL RESULTS AND ANALYSIS

4.1 PERFORMANCE METRICS

The performance of the fraud detection models is evaluated using a comprehensive set of metrics to ensure a holistic assessment. These metrics include accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic curve (AUC-ROC). Accuracy measures the overall correctness of the model, while precision and recall focus on the model's ability to identify

Journal of Economic Theory and Business Management
Journal Home: www.jetbm.online
Vol. 1, No. 4, 2024 | ISSN 3006-4953 (Print) | ISSN 3006-4961 (Online)

fraudulent transactions correctly. The F1 score provides a balanced measure of precision and recall, which is particularly useful in imbalanced datasets. AUC-ROC offers insight into the model's ability to distinguish between classes across various threshold settings. Table 5 presents the performance metrics for the Convolutional Neural Network (CNN) model with and without Deep Feature Synthesis (DFS):

TABLE 5: CNN PERFORMANCE METRICS

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
CNN without DFS	0.82	0.82	0.80	0.81	0.90
CNN with DFS	0.91	0.92	0.90	0.91	0.96

The results demonstrate a significant improvement in all metrics when DFS is applied. The accuracy increased from 82% to 91%, while the F1-score improved from 0.81 to 0.91. Notably, the AUC-ROC value increased from 0.90 to 0.96, indicating enhanced discriminative power of the model with DFS-generated features.

Illustrates the ROC curves for both CNN models. The graph displays the actual positive rate against the false positive rate at various threshold settings. The curve for the CNN model with DFS shows a more pronounced bow towards the top-left corner, visually confirming the improved performance indicated by the higher AUC-ROC value.

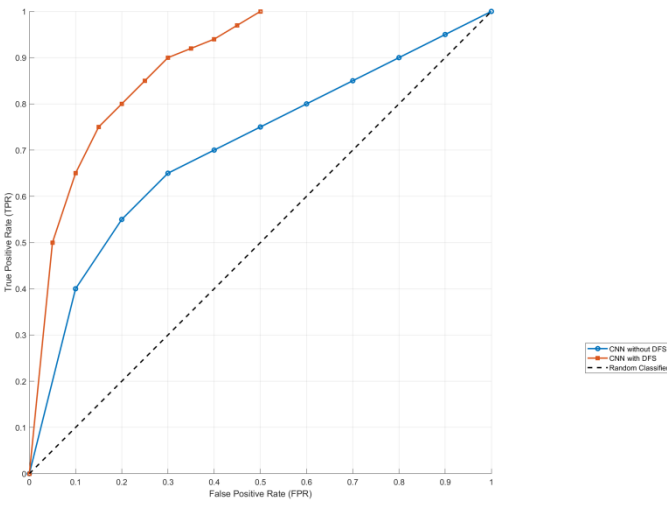

FIGURE 4:ROC CURVES FOR CNN MODELS

4.2 MODEL COMPARISON

A comparative analysis with other machine learning algorithms is conducted to contextualize the performance of the CNN model. The comparison includes Logistic Regression, Random Forest, Support Vector Machine (SVM), and XGBoost. Each model is evaluated using the same set of performance metrics, both with manually selected features and DFS-generated features. Table 6 summarizes the performance metrics for all models:

TABLE 6: PERFORMANCE COMPARISON OF DIFFERENT MODELS

Model	Feature Set	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Logistic Regression	Manual	0.75	0.76	0.73	0.74	0.82
Logistic Regression	DFS	0.79	0.80	0.77	0.78	0.86
Random Forest	Manual	0.86	0.87	0.84	0.85	0.93
Random Forest	DFS	0.89	0.90	0.88	0.89	0.95
SVM	Manual	0.80	0.81	0.78	0.79	0.87
SVM	DFS	0.84	0.85	0.83	0.84	0.91
XGBoost	Manual	0.88	0.89	0.87	0.88	0.94
XGBoost	DFS	0.90	0.91	0.89	0.90	0.96
CNN	Manual	0.82	0.82	0.80	0.81	0.90
CNN	DFS	0.91	0.92	0.90	0.91	0.96

The results indicate that while all models benefit from DFS-generated features, the CNN model significantly improves, outperforming other algorithms in most metrics when using DFS features.

A multi-axis plot is presented comparing the F1-scores and AUC-ROC values for all models with both feature sets. The x-axis represents different models, while two y-axes show the F1-score and AUC-ROC values. This visualization allows for a comprehensive comparison of model performances across different feature engineering approaches.

Journal of Economic Theory and Business Management
Journal Home: www.jetbm.online
Vol. 1, No. 4, 2024 | ISSN 3006-4953 (Print) | ISSN 3006-4961 (Online)

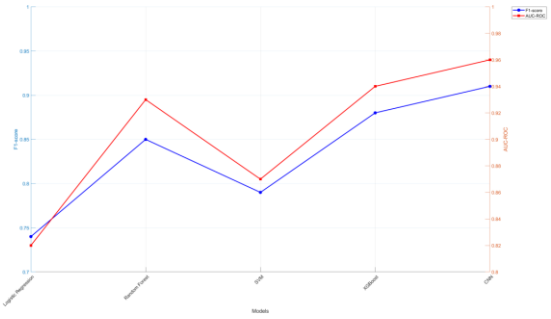

FIGURE 5:MODEL PERFORMANCE COMPARISON

4.3 IMPACT OF FEATURE ENGINEERING

The application of Deep Feature Synthesis substantially impacts model performance across all algorithms. The CNN model, in particular, shows a remarkable improvement with DFS features, with an 11% increase in accuracy and a 12% increase in F1-score. This significant enhancement can be attributed to the ability of DFS to capture complex relationships and temporal patterns within the transaction data, which are particularly relevant for fraud detection.

The improvement in model performance is not uniform across all algorithms. Traditional models like Logistic Regression show a modest improvement with DFS features, while ensemble methods like Random Forest and XGBoost demonstrate more substantial gains. This variation suggests that the effectiveness of DFS-generated features may depend on the underlying model's ability to leverage complex feature interactions.

A feature importance analysis is conducted on the DFS-generated features to analyze the impact of feature engineering further. Presents a heatmap visualization of the top 20 DFS-generated features ranked by their importance in the CNN model. The x-axis represents individual features, while the y-axis shows the relative importance score. This visualization provides insights into which synthetic features contribute most significantly to the model's predictive power.

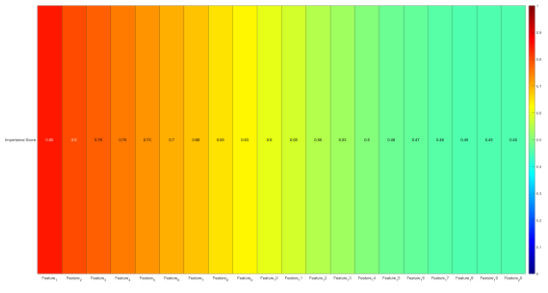

FIGURE 6:FEATURE IMPORTANCE ANALYSIS

4.4 DISCUSSION OF RESEARCH FINDINGS

The experimental results reveal several key findings in the application of deep learning and automated feature engineering to credit card fraud detection:

Enhanced Performance with DFS: The consistent improvement across all models with DFS-generated features underscores the value of automated feature engineering in fraud detection tasks. The ability of DFS to create features that capture complex relationships in the data proves particularly beneficial in identifying subtle fraud patterns [32].

The superiority of CNN with DFS: When combined with DFS features, the CNN model outperforms all other models in the study. This synergy between deep learning architecture and advanced feature engineering techniques suggests a promising direction for developing highly accurate fraud detection systems.

Adaptability to Data Complexity: The varying degrees of improvement across different algorithms highlight the importance of model selection in conjunction with feature engineering. Models with a higher capacity to handle complex feature interactions, such as CNN and XGBoost, appear to benefit more from DFS-generated features.

Interpretability Challenges: While the performance improvements are significant, the increased complexity of DFS-generated features poses challenges for model interpretability. This aspect requires careful consideration, especially in the context of regulatory compliance and explainable AI requirements in the financial sector.

Potential for Real-time Application: The substantial improvement in precision and recall suggests that the CNN model with DFS features could significantly reduce false positives and false negatives in real-world applications, potentially leading to more efficient fraud detection systems with fewer customer inconveniences.

These findings contribute to the growing knowledge of applying advanced machine-learning techniques to financial fraud detection. The study demonstrates the potential of combining deep learning architectures with automated feature engineering to create more effective and adaptable fraud detection systems.

5 CONCLUSION AND FUTURE WORK

5.1 SUMMARY OF RESEARCH CONTRIBUTIONS

This study has significantly contributed to credit card fraud detection by integrating advanced deep learning techniques with automated feature engineering. The research has demonstrated the effectiveness of combining Convolutional Neural Networks (CNN) with Deep Feature Synthesis (DFS) in improving the accuracy and reliability of fraud detection systems. The proposed approach has shown substantial improvements across all performance metrics, with the CNN model utilizing DFS-generated features outperforming traditional machine learning algorithms and standalone deep learning models [27].

The study has highlighted the importance of feature

engineering in enhancing the performance of fraud detection models [28]. The application of DFS has proven to be particularly effective in capturing complex relationships and temporal patterns within credit card transaction data, leading to more robust and accurate fraud detection. This automated approach to feature generation offers a scalable solution to the challenge of adapting to evolving fraud patterns, a critical aspect in the dynamic landscape of financial fraud [29].

Furthermore, the comparative analysis conducted in this research provides valuable insights into the relative strengths of various machine learning and deep learning approaches in fraud detection [30]. The superior performance of the CNN model with DFS features underscores the potential of this combination for developing next-generation fraud detection systems in the financial sector.

5.2 Research Limitations

While the research presents promising results, several limitations must be acknowledged. The study was conducted on a specific dataset, which, although comprehensive, may only partially represent the diversity of credit card transaction patterns across different geographical regions or financial institutions. The generalizability of the findings to other datasets or real-world scenarios requires further investigation.

The focus on a binary classification problem (fraud vs. non-fraud) simplifies the complexity of real-world fraud detection, where multiple types of fraudulent activities may exist. The model's performance distinguishing between different fraud categories remains unexplored in this study.

The computational resources required for implementing DFS and training deep learning models may pose challenges for real-time fraud detection in high-volume transaction environments. The trade-off between model complexity and computational efficiency needs further examination in the context of practical deployment.

While the study demonstrates improved performance metrics, the interpretability of the complex features generated by DFS and the decision-making process of the CNN model remain challenging [24]. This limitation is particularly relevant in the financial sector, where regulatory requirements often demand explainable AI solutions.

Future Research Directions

Building upon the findings and limitations of this study, several promising directions for future research emerge. Exploring the application of the proposed approach to more extensive and diverse datasets, including cross-regional and multi-institutional data, would provide insights into the model's generalizability and robustness.

Integrating other deep learning architectures, such as recurrent neural networks (RNNs) or transformer models, with DFS could capture more complex temporal dependencies in transaction sequences [9]. This exploration may improve fraud detection accuracy, especially for evolving sophisticated fraud patterns [23].

Future work should also focus on enhancing the interpretability of DFS-generated features and deep learning models [11,19,25]. Developing techniques to provide clear explanations for fraud predictions would address regulatory requirements and build trust in the system among financial institutions and customers [17].

Extending the binary classification approach to multi-class fraud detection, where different types of fraudulent activities are distinguished, represents another valuable research direction [7,26,31]. This extension would provide a more nuanced understanding of fraud patterns and enable more targeted prevention strategies [15,16].

Investigating the real-time applicability of the proposed approach, including optimizing computational efficiency and developing strategies for continuous model updating, would bridge the gap between research findings and practical implementation in the dynamic environment of financial transactions [6,10,14].

.

ACKNOWLEDGMENTS

I want to extend my sincere gratitude to Xiaoan Zhan, Zhipeng Ling, Zeqiu Xu, Lingfeng Guo, and Shikai Zhuang for their groundbreaking research on driving efficiency and risk management in finance through AI and RPA as published in their article [39] . Their insights and methodologies have significantly influenced my understanding of advanced techniques in fraud detection and have provided valuable inspiration for my research in this critical Area. I would also like to express my heartfelt appreciation to Fanyi Zhao, Hanzhe Li, Kaiyi Niu, Jiatu Shi, and Runze Song for their innovative study on the application of deep learning-based intrusion detection systems in network anomaly traffic detection [40]. Their comprehensive analysis and modeling approaches have significantly enhanced my knowledge of anomaly detection techniques and inspired my research in financial fraud detection.

FUNDING

Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not applicable.

ABOUT THE AUTHORS

XU, Haosen

Electrical Engineering and Computer Science, University of California, Berkeley, CA, USA.

LI, Siyang

Financial Econometrics, Johns Hopkins University, DC, USA.

NIU, Kaiyi

Artificial intelligence, Royal Holloway University of London, Egham, UK.

PING, Gang

Global Business and Marketing, Hong Kong Metropolitan University, Hong Kong, China.

REFERENCES

[1] Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., & Ahmed, M. (2022). Credit Card Fraud Detection Using State-of-the-art Machine Learning and Deep Learning Algorithms. IEEE Access, 10, 39700-39715.

[2] Anusha, P., Bharath, S., Rajendran, N., Durga Devi, S., & Saravanakumar, S. (2023). Experimental Evaluation of Smart Credit Card Fraud Detection System using Intelligent Learning Scheme. In 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1–6). IEEE.

[3] Xu, K., Zheng, H., Zhan, X., Zhou, S., & Niu, K. (2024). Evaluation and Optimization of Intelligent Recommendation System Performance with Cloud Resource Automation Compatibility.

[4] Zheng, H., Xu, K., Zhou, H., Wang, Y., & Su, G. (2024). Medication Recommendation System Based on Natural Language Processing for Patient Emotion Analysis. Academic Journal of Science and Technology, 10(1), 62-68.

[5] Xu, K., Zhou, H., Zheng, H., Zhu, M., & Xin, Q. (2024). Intelligent Classification and Personalized Recommendation of E-commerce Products Based on Machine Learning. arXiv preprint arXiv:2403.19345.

[6] Li, H., Wang, S. X., Shang, F., Niu, K., & Song, R. (2024). Applications of Large Language Models in Cloud Computing: An Empirical Study Using Real-world Data. International Journal of Innovative Research in Computer Science & Technology, 12(4), 59-69.

[7] Ping, G., Wang, S. X., Zhao, F., Wang, Z., & Zhang, X. (2024). Blockchain-Based Reverse Logistics Data Tracking: An Innovative Approach to Enhance E-Waste Recycling Efficiency.

[8] Xu, H., Niu, K., Lu, T., & Li, S. (2024). Leveraging artificial intelligence for enhanced risk management in financial services: Current applications and prospects. Engineering Science & Technology Journal, 5(8), 2402-2426.

[9] Shi, Y., Shang, F., Xu, Z., & Zhou, S. (2024). Emotion-Driven Deep Learning Recommendation Systems: Mining Preferences from User Reviews and Predicting Scores. Journal of Artificial Intelligence and Development, 3(1), 40–46.

[10] Wang, Shikai, Kangming Xu, and Zhipeng Ling. "Deep Learning-Based Chip Power Prediction and Optimization: An Intelligent EDA Approach." International Journal of Innovative Research in Computer Science & Technology 12.4 (2024): 77-87.

[11] Zhan, X., Shi, C., Li, L., Xu, K., & Zheng, H. (2024). Aspect category sentiment analysis based on multiple attention mechanisms and pre-trained models. Applied and Computational Engineering, pp. 71, 21–26.

[12] Liu, B., Zhao, X., Hu, H., Lin, Q., & Huang, J. (2023). Detection of Esophageal Cancer Lesions Based on CBAM Faster R-CNN. Journal of Theory and Practice of Engineering Science, 3(12), 36–42.

[13] Liu, B., Yu, L., Che, C., Lin, Q., Hu, H., & Zhao, X. (2024). Integration and performance analysis of artificial intelligence and computer vision based on deep learning

algorithms. Applied and Computational Engineering, pp. 64, 36–41.

[14] Liu, B. (2023). Based on intelligent advertising recommendations and abnormal advertising monitoring systems in the field of machine learning. International Journal of Computer Science and Information Technology, 1(1), 17–23.

[15] Wu, B., Xu, J., Zhang, Y., Liu, B., Gong, Y., & Huang, J. (2024). Integration of computer networks and artificial neural networks for an AI-based network operator. arXiv preprint arXiv:2407.01541.

[16] Liang, P., Song, B., Zhan, X., Chen, Z., & Yuan, J. (2024). Automating the training and deployment of models in MLOps by integrating systems with machine learning. Applied and Computational Engineering, 67, 1-7.

[17] Wu, B., Gong, Y., Zheng, H., Zhang, Y., Huang, J., & Xu, J. (2024). Enterprise cloud resource optimization and management based on cloud operations. Applied and Computational Engineering, pp. 67, 8–14.

[18] Guo, L., Li, Z., Qian, K., Ding, W., & Chen, Z. (2024). Bank Credit Risk Early Warning Model Based on Machine Learning Decision Trees. Journal of Economic Theory and Business Management, 1(3), 24–30.

[19] Xu, Z., Guo, L., Zhou, S., Song, R., & Niu, K. (2024). Enterprise Supply Chain Risk Management and Decision Support Driven by Large Language Models. Applied Science and Engineering Journal for Advanced Research, 3(4), 1–7.

[20] Yang, T., Xin, Q., Zhan, X., Zhuang, S., & Li, H. (2024). ENHANCING FINANCIAL SERVICES THROUGH BIG DATA AND AI-DRIVEN CUSTOMER INSIGHTS AND RISK ANALYSIS. Journal of Knowledge Learning and Science Technology ISSN: 2959–6386 (online), 3(3), 53–62.

[21] Zhao, F., Li, H., Niu, K., Shi, J., & Song, R. (2024, July 8). Application of deep learning-based intrusion detection system (IDS) in network anomaly traffic detection. Preprints.

[22] Gong, Y., Liu, H., Li, L., Tian, J., & Li, H. (2024, February 28). Deep learning-based medical image registration algorithm: Enhancing accuracy with dense connections and channel attention mechanisms. Journal of Theory and Practice of Engineering Science, 4(02), 1–7.

[23] Zhan, T., Shi, C., Shi, Y., Li, H., & Lin, Y. (2024). Optimization Techniques for Sentiment Analysis Based on LLM (GPT-3)—arXiv preprint arXiv:2405.09770.

[24] Yang, P., Chen, Z., Su, G., Lei, H., & Wang, B. (2024). Enhancing traffic flow monitoring with machine learning integration on cloud data warehousing. Applied and Computational Engineering, 67, 15-21.

[25] Jiang, W., Qian, K., Fan, C., Ding, W., & Li, Z. (2024). Applications of generative AI-based financial robot advisors as investment consultants. Applied and Computational Engineering, pp. 67, 28–33.

[26] Fan, C., Ding, W., Qian, K., Tan, H., & Li, Z. (2024). Cueing Flight Object Trajectory and Safety Prediction Based on SLAM Technology. Journal of Theory and Practice of Engineering Science, 4(05), 1–8.

[27] Jiang, W., Yang, T., Li, A., Lin, Y., & Bai, X. (2024). The Application of Generative Artificial Intelligence in Virtual Financial Advisor and Capital Market Analysis. Academic Journal of Sociology and Management, 2(3), 40-46.

[28] Li A, Zhuang S, Yang T, Lu W, Xu J. Optimization of logistics cargo tracking and transportation efficiency based on data science deep learning models. Applied and Computational Engineering. 2024 July 8;69:71-7.

[29] Xu, J., Yang, T., Zhuang, S., Li, H. and Lu, W., 2024. AI-based financial transaction monitoring and fraud prevention with behavior prediction. I Applied and Computational Engineering, 77, pp.218-224.

[30] Ling, Z., Xin, Q., Lin, Y., Su, G. and Shui, Z., 2024. Optimization of autonomous driving image detection based on RFAConv and triplet attention and applied and Computational Engineering, 77, pp.210-217.

[31] He, Z., Shen, X., Zhou, Y., & Wang, Y. (2024, January). Application of K-means clustering based on artificial intelligence in gene statistics of biological information engineering. In Proceedings of the 2024 4th International Conference on Bioinformatics and Intelligent Computing (pp. 468-473).

[32] Gong, Y., Zhu, M., Huo, S., Xiang, Y., & Yu, H. (2024, March). Utilizing Deep Learning for Enhancing Network Resilience in Finance. In 2024 7th International Conference on Advanced Algorithms and Control Engineering (ICAACE) (pp. 987–991). IEEE.

[33] Yang, T., Li, A., Xu, J., Su, G. and Wang, J., 2024. Deep learning model-driven financial risk prediction and analysis. I Applied and Computational Engineering, 77, pp.196-202.

[34] Liu, B., Cai, G., Ling, Z., Qian, J., & Zhang, Q. (2024). Precise Positioning and Prediction System for Autonomous Driving Based on Generative Artificial Intelligence. Applied and Computational Engineering, pp. 64, 42–49.

[35] Cui, Z., Lin, L., Zong, Y., Chen, Y., & Wang, S. (2024). Precision Gene Editing Using Deep Learning: A Case Study of the CRISPR-Cas9 Editor. Applied and Computational Engineering, 64, 134-141.

[36] Zhang, X., (2024). Machine learning insights into digital payment behaviors and fraud prediction. Applied and Computational Engineering, 67, pp.61–67.

[37] Zhang, Y., Xie, H., Zhuang, S., & Zhan, X. (2024). Image Processing and Optimization Using Deep Learning-Based Generative Adversarial Networks (GANs). Journal of Artificial Intelligence General Science (JAIGS) ISSN: 3006–4023, 5(1), 50–62.

[38] Liu, Y., Xu, Y., & Song, R. (2024). Transforming User Experience (UX) through Artificial Intelligence (AI) in interactive media design. Engineering Science & Technology Journal, 5(7), 2273-2283.

[39] Zhan, X., Ling, Z., Xu, Z., Guo, L., & Zhuang, S. (2024). Driving Efficiency and Risk Management in Finance through AI and RPA. Unique Endeavor in Business & Social Sciences, 3(1), 189–197.

[40] Feng, Y., Qi, Y., Li, H., Wang, X., & Tian, J. (2024, July 11). Leveraging federated learning and edge computing for recommendation systems within cloud computing networks. In Proceedings of the Third International Symposium on Computer Applications and Information Systems (ISCAIS 2024) (Vol. 13210, pp. 279–287). SPIE.