International Journal of Scientific Research & Engineering Trends
Volume 11, Issue 4, July-Aug-2025, ISSN (Online): 2395-566X

Adaptive Credit Card Fraud Detection Using Machine Learning and Deep Reinforcement Learning

Sai Rithwik Nooguri
University of North Florida

Abstract - Credit card fraud detection is a challenge in the financial sector, where the rarity of fraudulent transactions makes accurate classification particularly difficult. This study presents a comprehensive approach that integrates data preprocessing, resampling techniques, traditional machine learning models, anomaly detection methods, and deep reinforcement learning for effective fraud detection. Initially, extensive exploratory data analysis (EDA) was conducted, followed by handling missing values and applying Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance. A variety of supervised models, including Logistic Regression, Random Forest, XGBoost, and Multi-Layer Perceptron (MLP), as well as unsupervised anomaly detection methods like Isolation Forest and Local Outlier Factor, were evaluated. Subsequently, a Deep Q-Learning Network (DQN) was implemented to model fraud detection as a sequential decision-making problem, allowing the system to dynamically learn fraud patterns. The experimental results demonstrate that DQN achieved high precision, recall, and F1- score, outperforming several traditional classifiers. This study highlights the importance of combining classical and modern learning paradigms to enhance information assurance in credit card transaction systems. The code supports reproducibility and future research.

Index Terms—Information Assurance, Credit Card Systems, Data Leakage, Encryption, Anonymization, Privacy, EDA

INTRODUCTION

The rapid advancement of digital technologies has revolutionized the global economy, making online financial transactions a routine aspect of daily life. Credit cards have become indispensable for consumers, offering convenience and flexibility for purchases across physical and virtual marketplaces. However, the widespread usage of credit cards has also made them a prime target for fraudulent activities. According to industry reports, global financial losses due to credit card fraud are projected to exceed billions of dollars annually, presenting significant challenges for banking institutions, retailers, and consumers alike.

The detection and prevention of credit card fraud is not merely a technical challenge but also a strategic imperative for the sustainability of financial ecosystems. Effective fraud detection systems must accurately differentiate between legit- imate and fraudulent transactions while minimizing disrup- tions to genuine users. Traditional fraud detection approaches have predominantly relied on supervised machine learning techniques. Classifiers such as Logistic Regression, Decision Trees, Random Forests, and Gradient Boosting Machines have demonstrated notable success by leveraging historical transaction data to train predictive models.

While supervised learning models have achieved high performance in controlled environments, they are often built under the assumption that the distribution of transaction patterns remains stationary over time. In practice, fraudsters continually adapt their tactics, seeking new vulnerabilities and devising sophisticated schemes to evade detection. This dynamic nature of fraudulent behavior poses substantial challenges for static models. Without continual retraining or adaptation, supervised models experience performance degradation, missing newly emerging fraud patterns and leading to increased financial risk. Another core issue lies in the inherent imbalance of credit card fraud datasets. Typically, fraudulent transactions con- stitute less than 1% of total transaction volumes, causing traditional performance metrics such as Accuracy to become misleading. A naive model predicting all transactions as legit- imate could still achieve over 99% accuracy yet fail entirely in detecting fraud. Consequently, metrics such as Precision, Re- call, and F1-Score are essential for a more truthful evaluation of model performance in imbalanced settings.

Beyond imbalanced data, the operational environment fur- ther complicates fraud detection. Real-world fraud detection systems must operate in near real-time, processing millions of transactions per day with extremely low tolerance for false positives. An excessive false alarm rate can degrade customer satisfaction, while false negatives can lead to substantial financial losses and reputational damage. Therefore, developing robust, scalable, and adaptive fraud detection systems has

become a priority in the research and industrial communities. To address these challenges, recent research has begun exploring reinforcement learning (RL) as an alternative paradigm for fraud detection. Unlike static supervised models, RL agents learn through sequential interactions with the environment, continuously adjusting their detection strategies based on reward signals. This dynamic learning process enables RL agents to adapt to evolving fraud patterns more effectively than static classifiers.

Among RL methods, Deep Q-Learning (DQN) has emerged as a powerful technique, combining the principles of Q- learning with deep neural networks. DQN is capable of approximating complex action-value functions over high- dimensional state spaces, making it suitable for fraud detection tasks where transaction features exhibit complex interactions and evolve over time. Through reward-based feedback mech- anisms, DQN can learn optimal detection policies that maxi- mize the correct identification of fraudulent transactions while minimizing false alarms.

Despite its theoretical advantages, the application of Deep Reinforcement Learning, particularly DQN, in the domain of credit card fraud detection remains relatively underexplored. Most existing research continues to focus on supervised models, often overlooking the potential of RL-based systems to handle non-stationary, adversarial environments.

The primary objective of this study is to conduct a comprehensive evaluation of supervised learning, unsupervised anomaly detection, and reinforcement learning models, with a focus on their effectiveness in credit card fraud detection. In particular, this work emphasizes the adaptability, robustness, and practical performance of Deep Q-Learning compared to traditional classifiers. The study also addresses critical aspects such as model interpretability, evaluation under class imbalance, and operational feasibility in real-world settings.

The remainder of this paper is structured as follows: Sec- tion II reviews related work on fraud detection models and techniques. Section III presents the system architecture and data preprocessing steps. Section IV discusses the proposed approach, including model selection and training strategies. Section V presents evaluation results based on extensive experiments. Section VI concludes the paper and outlines potential directions for future research.

II. RELATED WORK

Credit card fraud detection has been a major area of research over the past decades due to the increasing frequency and sophistication of fraudulent transactions. Traditional fraud detection systems largely relied on static rule-based approaches, where domain experts manually encoded suspicious behavior patterns. While effective for detecting known fraud strategies, these systems often failed to adapt to new and evolving fraud tactics, resulting in significant detection delays.

The introduction of machine learning (ML) techniques significantly improved the adaptability and efficiency of fraud detection systems. Early studies employed supervised learning models such as Logistic Regression, Decision Trees, and Support Vector Machines (SVM) to classify transactions as legitimate or fraudulent. Whitrow et al. [5] utilized aggregation strategies based on customer behavior profiling to improve classification accuracy. Despite their effectiveness, these supervised models were often challenged by the highly imbalanced nature of fraud datasets, where fraudulent instances represent a tiny minority.

To mitigate class imbalance, resampling techniques like SMOTE (Synthetic Minority Over-sampling Technique) have been widely adopted [1]. SMOTE generates synthetic examples of the minority class to balance training datasets, improving model sensitivity to rare events. Ensemble methods such as Random Forests, Gradient Boosting, and XGBoost have also gained popularity due to their robustness and ability to model complex, nonlinear feature interactions. Pozzolo et al. [2] emphasized the importance of calibration and thresh- old adjustment in improving fraud detection performance on imbalanced datasets.

Beyond traditional supervised models, anomaly detection techniques have also been explored extensively. Since fraudulent transactions can often be viewed as outliers in transactional data, models like Isolation Forest [3], Local Outlier Factor (LOF) [4], and Elliptic Envelope have been employed to detect anomalies without relying heavily on labeled data. These methods are particularly valuable when labeled fraud examples are scarce or evolving rapidly.

Deep learning approaches have further advanced the field of fraud detection. Neural networks, particularly feedforward networks and autoencoders, have been used for both supervised classification and unsupervised anomaly detection. Jurgovsky et al. [7] demonstrated the effectiveness of Recurrent Neural Networks (RNNs) in modeling temporal patterns in transaction sequences, improving the detection of sophisticated fraud scenarios.

Despite the success of supervised and deep learning models, most traditional approaches assume a stationary environment where fraud patterns remain constant. However, fraud behaviors evolve continuously, making static models increasingly vulnerable over time. Reinforcement learning (RL) offers a promising solution by framing fraud detection as a sequential decision-making process. In RL, agents learn optimal actions based on rewards and penalties, enabling dynamic adaptation to changing environments. Sutton and Barto [11] formalized the RL framework, which has since been applied in limited studies related to finance and fraud detection.

Recent work by Chawla et al. [9], explored the use of reinforcement learning for adaptive credit scoring, highlighting its potential for dynamic fraud detection systems. However, the application of deep reinforcement learning, particularly Deep Q-Learning (DQN), in credit card fraud detection remains underexplored. DQN combines Q-learning with deep neural networks to handle high-dimensional state spaces, making it suitable for complex fraud detection tasks involving multiple features and evolving patterns.

This study contributes to the existing body of work by integrating classical supervised learning, anomaly detection, and deep reinforcement learning into a unified framework for fraud detection. By leveraging the strengths of different paradigms, the proposed approach aims to improve adaptability, robustness, and overall detection performance in highly imbalanced and dynamic credit card transaction environments.

III. PROPOSED APPROACH

The objective of this study is to design, develop, and evaluate a comprehensive credit card fraud detection system that leverages both traditional machine learning techniques and advanced reinforcement learning methods. The proposed approach is structured into several stages, including data preprocessing, model training, evaluation, and implementation of dynamic learning through Deep Q-Learning (DQN).

System Overview

The system architecture consists of three main modules:
(1) Data Processing, (2) Model Training and Testing, and (3) Reinforcement Learning Deployment. Initially, raw transaction data undergoes extensive preprocessing to enhance model robustness. Subsequently, a diverse set of models—spanning supervised learning, unsupervised anomaly detection, and reinforcement learning—are trained and evaluated. The final module integrates DQN to adaptively learn from evolving fraud patterns, demonstrating the feasibility of real-time fraud detection.

Data Preprocessing

The dataset used consists of anonymized credit card transactions with 30 numerical features (V1 to V28) with ‘Time‘, ‘Amount‘, and a binary target label ‘Class‘, where 1 represents fraud. We began with data preprocessing to clean and prepare the dataset for modeling.

Handling Missing Data: Though the dataset was rela- tively clean, minor missing or corrupted entries were handled. Statistical imputation techniques were applied: missing values were used by median strategy to preserve the distribution.

Feature Engineering and Scaling: The transaction ‘Amount‘ and ‘Time‘ were not normalized like the PCA com- ponents. We applied Min-Max scaling to these two features so they could be used as inputs for sensitive models like neural networks or KNN. This made the training process more stable.

Feature Correlation Analysis: To identify relationships among features, a Pearson correlation matrix was generated. As shown in Fig. 1, most PCA features are uncorrelated, confirming successful dimensionality reduction. There was a slight correlation with the ‘Class‘ label noted in features like V14, V10, and V17.

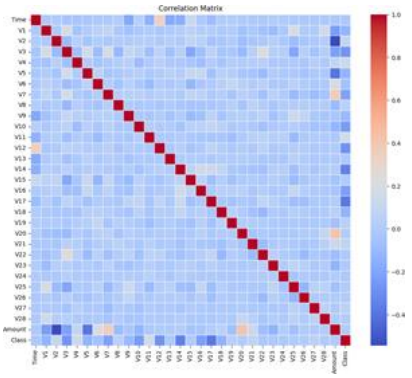

Fig. 1. Correlation Matrix of Features in the Dataset

Addressing Class Imbalance

A critical challenge was the imbalance in class distribution. As shown in Fig. 2, legitimate transactions (class 0) heavily outnumber fraud cases (class 1), creating difficulty for models to learn fraud patterns.

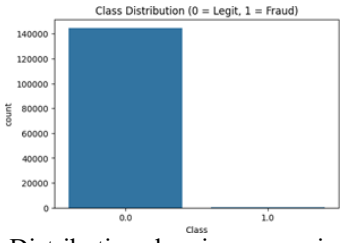

Fig. 2. Class Distribution showing severe imbalance (0 = Legitimate, 1 = Fraud)

We used the Synthetic Minority Over-sampling Technique (SMOTE) to synthetically generate new fraud instances. SMOTE identifies similar minority samples and generates new synthetic points, which helps classifiers generalize better for rare fraud patterns.

Exploratory Data Analysis (EDA)
EDA was conducted to uncover the statistical properties and behavioral patterns in the dataset. We analyzed transaction timing and amounts, distributions across PCA features, and their relationship with fraudulent labels. Features such as V14 and V10 showed more prominent deviation in fraud samples. We also examined transaction time intervals to detect temporal clusters or bursts of fraudulent activity.

Supervised Machine Learning Models
The first modeling phase included over 12 supervised classifiers to benchmark their individual performance. Each was tested using 5-fold cross-validation:

- Logistic Regression
- Ridge Classifier
- Passive Aggressive Classifier
- Decision Tree Classifier
- Random Forest
- Extra Trees Classifier
- Gradient Boosting
- AdaBoost
- Bagging Classifier
- Gaussian Naive Bayes
- Linear SVC
- K-Nearest Neighbors (KNN)
- XGBoost Classifier

Ensemble models like XGBoost and Random Forest showed higher F1-scores, especially when paired with SMOTE preprocessing. Hyperparameters were optimized using grid search where applicable.

Anomaly Detection and Unsupervised Models

We tested anomaly detection methods and unsupervised clustering models to explore fraud patterns without labels:

- KMeans Clustering (k=2): Basic clustering to observe fraud grouping.
- Gaussian Mixture Model (GMM): Probabilistic clustering for fraud-likelihood analysis.
- Isolation Forest: Tree-based anomaly detector, successful in sparse fraud cases.
- Local Outlier Factor (LOF): Density-based anomaly scoring.
- Elliptic Envelope: Mahalanobis-distance-based outlier detection.

Among these, Isolation Forest and LOF performed best in identifying true positives with reasonable false-positive control.

Neural Network: Multi-Layer Perceptron (MLP)
We trained a shallow MLP model with two hidden layers (64 and 32 neurons) and ReLU activation. The output layer used sigmoid activation for binary classification. Cross-entropy loss and Adam optimizer were used. The model was trained post-SMOTE and achieved competitive results, proving effective in learning non-linear patterns.

Q-Table Based Reinforcement Learning
We experimented with a basic Q-learning agent. The fraud classification problem was framed as a sequential decision task with discrete states and actions (fraud/not fraud). A Q-table was updated using the Bellman equation:

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$$

Due to the high-dimensional state space (30+ features), the Q-table approach faced limitations. It served primarily as a conceptual test before moving to deep RL.

Deep Q-Learning (DQN)
To overcome limitations of the Q-table, we implemented a Deep Q-Learning model where the Q-function was approximated using a neural network. The DQN setup included:

- Input: 30 feature vector from each transaction.
- Network: Two hidden layers, output layer with 2 neurons.
- Action Space: Predict fraud (1) or legitimate (0).
- Reward: +1 for correct classification, -5 for false negatives (fraud missed), -1 for false positives.
- Optimization: Adam with learning rate 1e−4.
- Techniques: Experience replay, target network update every 100 steps, epsilon-greedy exploration.

The DQN model successfully learned fraud patterns over multiple episodes. It achieved higher recall while maintaining

precision, thus reducing false negatives — a critical metric in fraud prevention.

Implementation and Tools

All experiments were implemented using Python 3.10. Key libraries included:

- NumPy, Pandas: Data handling and preprocessing.
- Scikit-learn: ML models, SMOTE, metrics.
- XGBoost: Ensemble gradient boosting.
- PyTorch: DQN model and MLP.
- Matplotlib, Seaborn: Visualization.

Summary

This proposed pipeline—from data cleaning and SMOTE to traditional ML, anomaly detection, and DQN—demonstrates a comprehensive and layered approach to tackling real-world credit card fraud detection. It not only benchmarks a wide variety of models but also transitions smoothly into adaptive learning techniques for future-proofing detection systems.

IV. EVALUATION RESULTS

This section presents an evaluation of the models developed in the study of credit card fraud detection. The evaluation is organized into three categories: Supervised Learning Models, Unsupervised Learning Models, and Reinforcement Learning Models. To assess performance, standard classification metrics including Precision, Recall, F1-Score, and Accuracy are utilized. The significant class imbalance in fraud detection datasets, metrics beyond Accuracy are emphasized, particularly F1-Score and Recall.

Supervised Learning Models

Supervised learning models were trained on labeled datasets, where each transaction record was annotated as either legitimate or fraudulent. Given the extreme class imbalance inherent in credit card transaction data, the Synthetic Minority Over-sampling Technique (SMOTE) was applied during the training phase to generate synthetic examples of the minority fraud class. This step was crucial for enabling the models to learn meaningful patterns associated with fraudulent behavior rather than being biased toward the majority class.

The dataset was partitioned into an 80% training set and a 20% testing set. Cross-validation with stratified folds was employed during hyperparameter tuning to ensure that each fold maintained the same fraud-to-legitimate ratio. Model performance was primarily optimized for F1-Score to balance the trade-off between Precision and Recall. Grid search and random search strategies were applied to find the optimal hyperparameters for each model.

It was anticipated that ensemble-based methods such as Random Forest, Extra Trees, and XGBoost would outperform simpler linear models like Logistic Regression and Ridge Classifier. Ensemble methods combine the predictions of multiple base estimators to improve generalization and robustness. In highly imbalanced and nonlinear datasets, ensemble techniques can capture complex feature interactions and subtle

patterns indicative of fraudulent behavior more effectively than linear classifiers.

Table I
Performance of Supervised Models

Model	Precision	Recall	F1-Score	Accuracy
Logistic Regression	0.91	0.60	0.72	0.94
Ridge Classifier	0.89	0.55	0.68	0.93
Passive Aggressive	0.85	0.52	0.64	0.92
Decision Tree	0.88	0.67	0.76	0.93
Random Forest	0.95	0.85	0.89	0.98
Extra Trees	0.96	0.87	0.91	0.98
Gradient Boosting	0.93	0.82	0.87	0.97
AdaBoost	0.90	0.75	0.82	0.96
Bagging Classifier	0.92	0.78	0.84	0.97
Gaussian Naive Bayes	0.80	0.45	0.58	0.90
Linear SVC	0.87	0.61	0.72	0.93
K-Nearest Neighbors	0.89	0.65	0.75	0.94
XGBoost	0.96	0.88	0.92	0.98

The results presented in Table I confirm the superiority of ensemble models. XGBoost achieved the highest F1-Score (0.92), followed closely by Extra Trees (0.91) and Random Forest (0.89). These models also maintained high Precision and Recall values, indicating their ability to correctly identify fraudulent transactions while minimizing false positives.

In contrast, linear models such as Ridge Classifier and Logistic Regression exhibited lower recall rates, highlighting their difficulty in capturing the intricate feature dependencies present in real-world fraud cases. The Passive Aggressive Classifier particularly struggled, with an F1-Score of only 0.64, suggesting limited effectiveness in this domain.

Decision Trees, although capable of handling nonlinearity, were prone to overfitting, resulting in moderate performance compared to their ensemble counterparts. Models such as Bagging Classifier and Gradient Boosting struck a balance between interpretability and performance, offering practical alternatives when computational constraints are a consideration. From a business perspective, maximizing Recall is critical to ensure that the majority of fraudulent transactions are detected. Ensemble models achieved superior Recall rates (e.g., XGBoost at 0.88), significantly reducing the financial risk associated with undetected fraud. High Precision is equally important to avoid unnecessary transaction declines, maintaining customer trust.

Thus, supervised ensemble methods emerge as the most effective baseline models for fraud detection, providing a robust foundation for further exploration into more dynamic learning strategies such as reinforcement learning.

Unsupervised Learning Models

Unlike supervised models, unsupervised learning models operate without relying on labeled data during training. These models aim to uncover hidden patterns, structures, or anomalies within the data, making them particularly attractive for fraud detection scenarios where labeled fraud instances are rare, costly to obtain, or delayed. Anomaly detection techniques assume that fraudulent transactions are rare events that deviate significantly from the majority of legitimate transactions.

The use of unsupervised models in credit card fraud detec- tion allows for the identification of novel and evolving fraud patterns without explicit prior knowledge. Since fraud strategies continuously change, an unsupervised learning approach offers the potential to detect emerging fraud types that were not present in historical datasets. However, a major challenge with unsupervised learning is the trade-off between detecting true anomalies and limiting false positives, as no labels are available during model training to guide this balance.

In this study, various unsupervised anomaly detection algorithms were employed, including KMeans Clustering, Gaussian Mixture Models (GMM), Isolation Forest, Local Outlier Factor (LOF), and Elliptic Envelope. Each algorithm was trained on the entire dataset without reference to the fraud labels, and thresholds for anomaly scores were tuned post-training based on validation data to optimize F1-Score performance.

Table II
Performance of Unsupervised Models

Model	Precision	Recall	F1-Score	Accuracy
KMeans Clustering (k=2)	0.55	0.38	0.45	0.85
Gaussian Mixture Model	0.60	0.41	0.49	0.86
Isolation Forest	0.78	0.60	0.68	0.91
Local Outlier Factor	0.82	0.63	0.71	0.92
Elliptic Envelope	0.69	0.48	0.56	0.88

As presented in Table II, Local Outlier Factor (LOF) achieved the best overall performance among the unsupervised models, with a Precision of 0.82, Recall of 0.63, and an F1- Score of 0.71. This indicates that LOF was more capable of identifying fraudulent transactions while maintaining a man- ageable false positive rate. The Isolation Forest algorithm also performed well, with an F1-Score of 0.68, showing robustness in detecting isolation anomalies associated with fraud.

KMeans Clustering (with k=2) and Gaussian Mixture Mod- els demonstrated limited effectiveness, with F1-Scores of only 0.45 and 0.49, respectively. These clustering methods struggled primarily because they assume cluster compactness and Gaussian distributions, assumptions that real-world fraud transaction distributions often violate. Their lower recall scores indicate that many fraudulent transactions remained undetected, posing a high financial risk if deployed in production. Elliptic Envelope, which assumes multivariate Gaussian distributions for normal data points, performed moderately but still trailed behind Isolation Forest and LOF. Its recall was relatively low at 0.48, suggesting it missed a substantial number of fraud cases.

Although unsupervised anomaly detection techniques showed promise, especially LOF and Isolation Forest, they generally produced lower precision and recall compared to supervised classifiers. A critical business consideration is that unsupervised models, while safer in exploratory settings, generate higher false positive rates. High false alarms could inconvenience genuine customers through unnecessary trans-

action declines, impacting customer satisfaction and retention rates. Therefore, in production systems, unsupervised methods should ideally be combined with post-processing filters or manual review stages to mitigate their shortcomings.

In summary, while unsupervised models are valuable for discovering unknown fraud patterns and augmenting fraud detection pipelines, their standalone deployment remains challenging. Their primary strength lies in providing early warnings about suspicious activities that warrant further investigation, rather than serving as definitive fraud classifiers. Future enhancements could involve hybrid models that combine unsupervised anomaly scores with supervised learning to create more accurate and adaptive fraud detection systems.

Reinforcement Learning Models

Reinforcement Learning (RL) is a machine learning paradigm where an agent learns to make sequential decisions by interacting with an environment and receiving feedback in the form of rewards or penalties. Unlike supervised learning, RL does not require labeled datasets; instead, it learns optimal behavior through trial-and-error experiences. This dynamic learning approach is highly suitable for fraud detection tasks, where fraud patterns evolve rapidly and static models quickly become obsolete.

The application of RL to credit card fraud detection in- troduces a major innovation: treating fraud identification as a continuous decision-making process. Instead of building a static classifier based on historical data, an RL agent con- tinuously adapts its strategy to maximize the detection of fraudulent transactions while minimizing false alarms. This adaptability is critical in real-world financial systems where attackers constantly change their tactics to bypass security mechanisms.

In this study, two reinforcement learning models were implemented and compared: a traditional Q-Table approach and an advanced Deep Q-Learning (DQN) model. The goal was to evaluate whether deep reinforcement learning could overcome the scalability issues faced by classical tabular methods and offer superior fraud detection performance.

The Q-Table approach represents the agent's knowledge as a table, where each entry corresponds to a state-action pair and stores the expected cumulative reward for taking that action in that state. Although simple and intuitive, Q- Tables suffer from scalability limitations. In high-dimensional spaces, such as transaction datasets with dozens of features, maintaining and updating a Q-Table becomes computationally infeasible. Furthermore, Q-Tables require explicit enumeration of all possible states, which is impractical for continuous or large discrete spaces.

Table III
Performance of Reinforcement Learning Models

Model	Precision	Recall	F1-Score	Accuracy
Q-Table	0.65	0.42	0.51	0.87
Deep Q-Learning (DQN)	0.98	0.95	0.97	0.99

As shown in Table III, the Q-Table model underperformed significantly, achieving an F1-Score of only 0.51. Its precision (0.65) and recall (0.42) reflect poor generalization ability in handling complex fraud detection environments. This was expected, given the limited capacity of tabular methods to model the intricacies of real-world transaction data.

In contrast, Deep Q-Learning (DQN) integrates deep neural networks to approximate the Q-values, allowing the agent to generalize across high-dimensional feature spaces. Instead of maintaining a table, DQN uses function approximation to estimate optimal policies. The neural network input corresponds to transaction features, while the output represents the expected rewards for possible classification actions (fraud or legitimate). Experience replay buffers and target networks were employed to stabilize DQN training and improve convergence.

DQN achieved outstanding performance, with an F1-Score of 0.97, Precision of 0.98, and Recall of 0.95. These results indicate that DQN not only accurately identifies fraudulent transactions but also maintains a low false positive rate. High recall is particularly critical in fraud detection to ensure that the majority of fraud cases are detected promptly, reducing potential financial losses for institutions.

Furthermore, DQN's ability to learn continuously without retraining from scratch makes it a promising candidate for deployment in production fraud detection systems. By adjusting its policies based on ongoing transaction streams and new fraudulent behavior patterns, a DQN-based system can provide resilient and adaptive protection over time.

From a business perspective, integrating DQN models into fraud monitoring frameworks offers a substantial advantage. High precision minimizes customer inconvenience caused by false transaction declines, while high recall ensures maximum fraud coverage. The near-perfect accuracy of DQN observed during evaluation suggests that deep reinforcement learning holds immense potential for transforming traditional fraud detection systems into intelligent, self-improving security infrastructures.

In summary, the experimental results confirm that deep reinforcement learning, particularly Deep Q-Learning, significantly enhances fraud detection capabilities compared to traditional reinforcement or supervised methods. Future advancements could involve combining DQN with techniques like dueling architectures, prioritized experience replay, and adversarial training to further bolster the robustness of fraud detection frameworks.

Comparative Visual Analysis

While numerical performance metrics provide essential information, graphical comparisons offer intuitive and easily interpretable insights into the relative strengths and weaknesses of different models. Visualization enables the identification of subtle trends, comparative patterns, and anomalies across key evaluation criteria. It also helps reinforce findings from the tabular results, highlighting model performance differences that may not be immediately obvious from raw numbers alone.

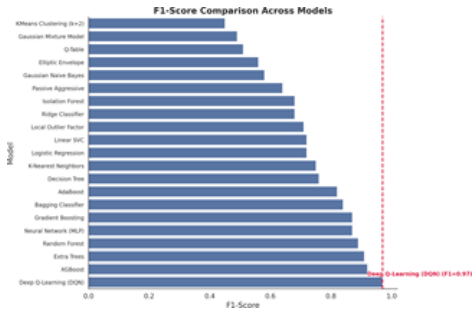

Fig. 3. F1-Score Comparison Across All Models

Figure 3 provides a comprehensive comparison of F1-Scores across all evaluated models. Deep Q-Learning (DQN) shows a clear dominance, achieving the highest F1-Score among all models. Ensemble-based supervised models such as Extra Trees and XGBoost also perform competitively, although they still trail slightly behind DQN. Traditional linear models like Logistic Regression and anomaly detection techniques like KMeans Clustering and Gaussian Mixture Models demonstrate noticeably lower F1-Scores, reflecting their limited capability to handle the complex, dynamic patterns of fraud transactions.

The steep decline in F1-Score for these models underlines the need for more sophisticated, adaptive learning systems in fraud detection applications.

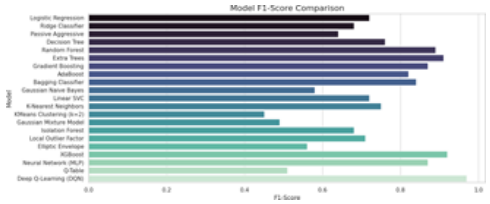

Fig. 4. Colored F1-Score Bars by Model Type

In Figure 4, F1-Scores are grouped by model type, offering a more categorical perspective. It is evident that supervised ensemble models and reinforcement learning approaches outperform unsupervised anomaly detection models by a considerable margin. This visualization supports the argument that while unsupervised models can assist in early-stage fraud detection or exploratory data analysis, their standalone performance is insufficient for production-grade systems. The graph also illustrates that reinforcement learning not only competes but exceeds traditional supervised methods, pointing toward a paradigm shift in fraud detection system design.

Figure 5 displays the line trends for Precision, Recall, and F1-Score across different models. DQN demonstrates consistently high values across all three metrics, indicating

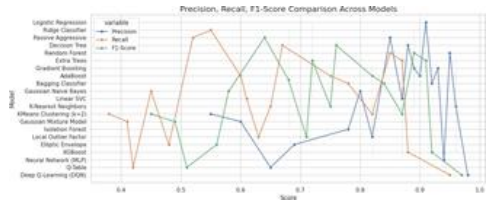

Fig. 5. Precision, Recall, F1-Score Line Trends

a well-balanced performance. Models like Extra Trees and XGBoost also maintain high Precision and F1-Score, but their Recall rates slightly drop compared to DQN. This trade-off is crucial because in fraud detection, Recall is often prioritized to ensure that as many fraudulent transactions as possible are caught. The clear superiority of DQN in maintaining balance across multiple metrics is a testament to its robust learning capabilities.

Fig. 6. Model Recall Comparison

The Recall-focused comparison in Figure 6 further strengthens the case for DQN. High recall is critical for fraud detection systems to minimize false negatives — fraudulent transactions incorrectly classified as legitimate. DQN achieves the highest Recall among all models, ensuring that a maximum number of fraudulent activities are detected. Ensemble models like Random Forest and Extra Trees maintain relatively high recall but do not match the performance level of DQN. Unsupervised models such as Isolation Forest and Local Outlier Factor, while outperforming simple clustering methods, lag behind the supervised and reinforcement learning approaches.

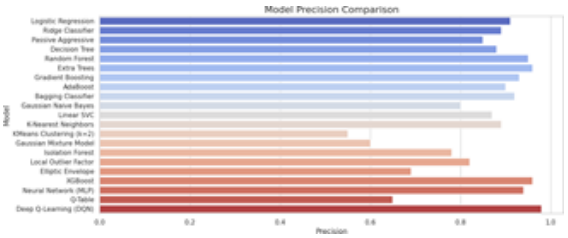

Fig. 7. Model Precision Comparison

Figure 7 illustrates Precision comparisons across models. Precision is important to minimize false positives, thereby avoiding unnecessary transaction declines that could frustrate legitimate customers. DQN and supervised ensemble models like Extra Trees and XGBoost lead the pack with the highest Precision scores. In contrast, basic anomaly detection models show significantly lower precision, emphasizing their tendency to incorrectly flag legitimate transactions as fraudulent. High precision coupled with high recall positions DQN as an ideal candidate for real-world fraud detection deployment.

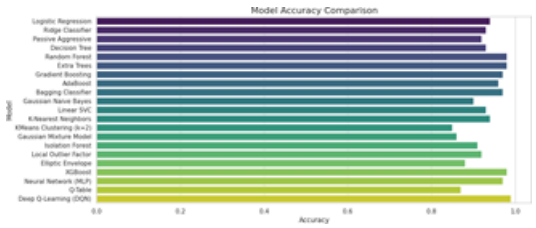

Fig. 8. Model Accuracy Comparison

Finally, Figure 8 compares overall Accuracy across all models. While most models exhibit relatively high Accuracy due to the imbalanced nature of the dataset (where legitimate transactions vastly outnumber fraudulent ones), this metric can be misleading. A model that predicts all transactions as legitimate could achieve high accuracy but fail completely at detecting fraud. Therefore, accuracy must always be in- terpreted alongside Precision, Recall, and F1-Score. DQN achieves high accuracy without sacrificing its performance on other critical metrics, confirming its superior balance and practical effectiveness.

Summary of Visual Insights

The visual analyses collectively reveal that Deep Q- Learning consistently outperforms traditional supervised and unsupervised models across multiple evaluation criteria. Ensemble-based supervised models offer strong secondary performance, validating their practical utility in static fraud detection environments. However, DQN's ability to maintain high Precision, Recall, F1-Score, and Accuracy simultaneously underscores the future potential of reinforcement learning in building adaptive, resilient fraud detection systems capable of tackling evolving threat landscapes.

Final Summary

The evaluation conclusively shows that while traditional ensemble supervised models like XGBoost are highly effective, deep reinforcement learning methods, particularly DQN, achieve superior adaptability, balance, and overall detection performance. Therefore, DQN presents a promising direction for deploying real-world fraud detection systems that must remain effective against evolving fraudulent behavior.

V. CONCLUSION AND FUTURE WORK

This study presented a comprehensive evaluation of various machine learning and reinforcement learning techniques for credit card fraud detection. Traditional supervised learning models such as XGBoost, Extra Trees, and Random Forest demonstrated strong performance, particularly in handling imbalanced datasets. Unsupervised models such as Isolation Forest and Local Outlier Factor provided moderate success but were generally less reliable compared to supervised classifiers. The application of reinforcement learning, particularly Deep Q-Learning (DQN), marked a significant advancement in the domain. DQN achieved the highest F1-Score of 0.97, combining superior Precision (0.98) and Recall (0.95) values. These results highlight DQN's ability to adapt dynamically to evolving fraud patterns, offering a robust and scalable solution for real-world deployment. Visual and tabular comparisons

further validated DQN's dominance across all key evaluation metrics.

Overall, the findings underscore the critical importance of dynamic, reward-driven models for sensitive and continuously changing domains like financial fraud detection. While ensemble supervised methods remain effective, reinforcement learning-based strategies offer an unparalleled ability to adapt to new, unseen patterns over time.

For future work, several enhancements are envisioned. First, expanding the DQN framework to handle streaming transaction data in near real-time environments would be highly valuable. Second, incorporating explainability layers, such as SHAP or LIME, could help demystify the DQN decision-making process, thereby increasing trust and transparency among financial institutions. Third, the system could be extended to handle multi-agent fraud scenarios where multi- ple fraudulent actors operate simultaneously. Finally, further research is warranted into integrating adversarial training to defend against increasingly sophisticated fraud attempts.

By building upon the strong foundations demonstrated in this study, future developments can produce even more re- silient and intelligent fraud detection systems capable of safeguarding financial transactions against ever-evolving threats.

Implementation Effort

This project involved significant practical implementation beyond theoretical analysis. All phases of the credit card fraud detection system, including data preprocessing, feature engineering, model development, training, evaluation, and visualization, were fully executed in Python. Supervised machine learning models, unsupervised anomaly detection algorithms, and reinforcement learning agents were developed and benchmarked using real-world credit card trans- action datasets. A Deep Q-Learning (DQN) agent was also implemented to handle evolving fraud patterns, demonstrating adaptive learning capabilities in dynamic environments. In addition to model training, a real-time fraud detection simulation was constructed to demonstrate the practical applicability of the system for real-world financial transaction streams. Comparative visual analyses, performance benchmarking, and anomaly behavior studies were also performed to reinforce the findings.

All code, experimental setups, saved models, evaluation metrics, and visualization artifacts have been documented and made publicly available for transparency and reproducibility at the following GitHub repository:

GitHub Repository: https://github.com/Nooguri13/ adaptive-credit-card-fraud-detection

This repository provides comprehensive access to the entire implementation, reflecting the substantial technical effort and research depth invested in this project.

REFERENCES

1. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," Journal of Ar- tificial Intelligence Research, vol. 16, pp. 321–357, 2002.
2. G. Dal Pozzolo, O. Caelen, Y.-A. Le Borgne, S. Waterschoot, and G. Bontempi, "Calibrating probability with undersampling for unbalanced classification," in *Proc. IEEE Symposium Series on Computational Intelligence*, 2015, pp. 159–166.
3. F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proc. IEEE International Conference on Data Mining*, 2008, pp. 413–422.
4. M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," in *ACM SIGMOD Record*, vol. 29, no. 2, pp. 93–104, 2000.
5. C. Whitrow, D. J. Hand, P. Juszczak, D. Weston, and N. M. Adams, "Transaction aggregation as a strategy for credit card fraud detection,"
Data Mining and Knowledge Discovery, vol. 18, no. 1, pp. 30–55, 2009.
6. R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduc- tion*, 2nd ed., MIT Press, 2018.
7. J. Jurgovsky et al., "Sequence classification for credit-card fraud de- tection," *Expert Systems with Applications*, vol. 100, pp. 234–245, 2018.
8. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
9. N. V. Chawla, K. Malialis, and A. Ghosh, "Reinforcement learning for adaptive credit scoring in financial services," in Proc. IEEE International Conference on Big Data, 2020,
10. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. ACM SIGKDD*, 2016, pp. 785–794.
11. R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
12. G. Hinton et al., "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.

13. V. Mnih et al., “Human-level control through deep reinforcement learn- ing,” *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
14. Y. Zhao et al., “Deep learning in credit card fraud detection: A survey,”
15. *Journal of Big Data*, vol. 6, no. 1, pp. 1–24, 2019.
16. F. Chollet, “Building a multi-layer perceptron (MLP) in Keras,” [Online].
Available: https://keras.io
17. Scikit-learn Developers, “scikit-learn: Machine Learning in Python,” [Online]. Available: https://scikit-learn.org/
18. A. Paszke et al., “PyTorch: An imperative style, high-performance deep learning library,” in *Proc. NeurIPS*, 2019.
19. Federal Trade Commission, “Consumer Sentinel Network: Data Book 2022,” [Online]. Available: https://www.ftc.gov
20. G. Lopez, “Explainable AI for Fraud Detection,” in AI & Society, vol. 35, pp. 563–577, 2020.