International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. , No. (Mon-20..)

http://dx.doi.org/10.12785/ijcds/XXXXXX

Credit Card Fraud Detection using Reinforcement Learning

Mrs. Smita Mahajan[1], Dr. Shrikrishna Kolhar[1], Ayushi Patil[1], Ms. Shreya Mahajan[1], Ms. Jinal Menpara[1] and Mr. Amay Mahajan[2]

[1]*Symbiosis Institute of Technology, Symbiosis International (Deemed University), Lavale, Pune, Maharashtra, India, 412115*
[2]*Mirae Asset Global Investments Newyork, USA, 10036*

Received Mon. 20, Revised Mon. 20, Accepted Mon. 20, Published Mon. 20

Abstract: Financial transactions are still plagued by credit card fraud, which poses a serious threat to both individuals and businesses. The evolution of fraud techniques frequently outpaces the ability of antiquated methods to detect them. This research uses reinforcement-based learning, more especially Deep Q-Learning, to examine credit card fraud detection. The first steps in this approach involved processing the dataset to extract features that would help with data normalization and classification. Subsequently, a DQN architecture that was appropriate for detecting credit card fraud was created and included parameters that would self-adjust over the course of several training sessions. After receiving training, DQN was able to distinguish between real and fraudulent transactions with an accuracy score of 90.54% on the testing set. To sum up, the findings suggest that the application of reinforcement learning, especially Deep Q-Learning, appears to be a practical and trustworthy technique for identifying credit card fraud. The constant learning process built on transaction practices makes it easier to predict how wrongdoing will change over time while maintaining transaction security. The current study adds to the body of knowledge on fraud prediction techniques by offering financial institutions, and businesses targeted advice and insights to help them effectively combat fraudulent activity.

Keywords: Credit card fraud detection, Reinforcement Learning, Deep Q Network, Q-Learning, Experience-Replay

1. INTRODUCTION

In recent years, fraudulent use of credit cards, which results in massive losses for cardholders as well as financial institutions, has led to the loss of billions annually [1]. It is interesting to note that credit card theft is not only common in the world of finance but also one that raises much anxiety. Attempts to find ways of preventing such big losses caused by theft of cash have been concentrated upon recently. Digitization has increased the usage of electronic payment methods of great importance to businesses and customers. One of the reasons why this has been successful is that there are various ways through which people can make payments such as m-commerce and p-commerce systems. Hackers continue coming up with different mechanisms through which they can compromise sophisticated technological infrastructures. The growth of electronic payment options worldwide has increased the ease of access to banking services by consumers and businesses, with credit card transactions being popular [2]. By means of convenience, nonetheless, there is a price to pay concerning the increasing likelihood that people are going to cheat the user in some way. Financial institutions, traders, and consumers perspectively face numerous issues with credit cards when thinking about different forms of fraud that could comprise unauthorized transactions, identity theft, and account takeover, all of which result in fraud. There is an increasing urgency for the development of smart fraud prevention systems that learn over time to check losses associated with online payment transactions and to ensure that electronic payment systems remain credible [3]. Conventional fraud detection approaches are mainly dependent on rule-based systems and statistical models, which frequently find it difficult to keep up with changing methodologies used by fraudsters. When it comes to this, advanced machine learning methods have surfaced as a potential method of boosting fraud detection capacities [4]. When these methods are considered, one really interesting paradigm is Reinforcement Learning (RL) because it looks for ways of making good decisions by interacting with surroundings on its own.

This research has investigated how to develop a credit card fraud detection system using Reinforcement Learning. The conventional approach to fraud detection solutions in financial institutions involves framing the issue as a classification challenge and focusing on enhancing the models' fraud recall rates[5]. When considered within the context of fraudulent activity, RL algorithms have been found to be able to change their approach according to new trends, thereby improving upon their real-time detection strategies [6]. With this method, the system's capacity to continuously get better in performance, as well as its ability to resist up-

to-date fraudulent tricks, is high, something that leads to a decreased number of wrongly labeled activities and boosts overall efficiency in performing fraud detection roles.

This research's main purpose is to investigate the possibility of incorporating RL methods within current fraud prevention systems. The main goal is to come up with a model centered on reinforcement learning which can detect fraud effectively using transactional data, feature engineering as well as reward based training without sounding too many false alarms. The number of machine learning applications in fraud detection is expanding rapidly. So, this paper adds to this number by providing information on the potential of Reinforcement Learning to address new problems that come with credit card fraud [7]. This work adds to the expanding body of knowledge literature on machine learning applications in fraud detection by providing insights into the potential of Reinforcement Learning to handle the developing issues of credit card fraud.

This study is set up in the following manner. Section II describes the broad overview of research done around fraud detection previously. It outlines various methodologies and results obtained by some of the significant papers in the spectrum of credit card fraud detection. In the methodology section (section III), a brief description of the dataset that has been used is given. The proposed methodology of credit card fraud detection using RL is explained in detail along with the insights for hyperparameters used during the experiment. The results section (Section IV), the comparative study of different parameters and its effect on resultant accuracy has been presented. The conclusion section (Section V) finally brings the paper to a close and highlights its future work.

2. Literature Review

This paper discusses the damages that the financial sector experiences exclusively on credit card theft. Since it is billions of dollars annually [8]. With the aim of eliminating this issue, authors have used Q-Credit Card Fraud Detector Architecture, consisting of deep learning, auto-encoders, and neural agents to distinguish a true transaction as opposed to a fraudulent one and got 98.1% results. Predictions are done through the Q-learning method. Simulations of the model using a computer can be observed in terms of how promoting the classification of the fraud category is and how fast it reacts.

Vima et al. touch on the problem the hackers are having with taking advantage of holes in the digital payment infrastructure and calls for more advanced and better fraud detection methods. The prime goal of classic fraud detection has traditionally been the highest recall rate possible; however, this could possibly lead to other subpar results. Resistance to data inequality, adaptability to changes in fraud trends, and a balance for their rates of fraud and decline are the models for good designing of systems against fraud. The suggested method formulates an artificial intelligence model for fraud detection as a sequential decision-making problem with utility maximization in which a reward is used. This strategy covers the model's assessment versus other classifiers and the model's performance by using a public pecuniary fraud dataset with Deep Q-learning and got 99.88 % accuracy. The system will be focused on the resolution of the most urgent issues in its upcoming version [9].

In an online commerce setting, the paragraph focuses on the aggravations and, respectively, the mitigating solutions relating to the instantaneous detection of credit card fraud. It draws attention to the fact that Internet fraud through credit cards has increased over a period of time, and several varieties have evolved [10]. This is because each person carries a different person and dynamic transaction patterns, making the real-time identification operation a difficult, if not impossible, one. The study to be carried out aims to examine the core function of deep reinforcement learning in the real-time detection of fraud of credit cards. The validation performance of 97% was achieved by the model named Deep Q network, which was trained on and evaluated with the dataset from the Kaggle. In the event necessary without additional people's training, the system itself will be able to make adjustments over time by drawing on previous experience.

In view of the fact that cyber crimes and cyber attacks happen more and more often, this particular research work stresses the importance of cyber security in the banking industry. Credit card fraud on a global scale is a huge security issue, and the usual method of detecting it is prone to inaccuracies, labor-intensive, and slow. The study has reviewed 181 research publications that were published between 2019 and 2021 to investigate machine learning and deep learning algorithms with respect to credit card cyber fraud detection, and SMOTE gives the best result with an accuracy of 99.95%. The research paper encourages the use of best practices by summarizing the approaches and explaining why their application for academics and the banking sector is important. Finally, it recommends more study areas since they reflect the good and bad sides of the current fraud detection systems. The objective of this thorough analysis is to determine what researchers and business experts have in detecting cyber fraud [11].

This study by Dang et al. evaluates the recent achievements made in deep reinforcement learning (DRL) and machine learning (ML) algorithms and also examines the problem of imbalanced datasets in credit card fraud detection systems [12]. It mentions the process of balancing the dataset by using resampling techniques such as SMOTE and ADASYN, then utilizing ML algorithms on the balanced dataset to build credit card fraud detection systems. This paragraph talks about the tests with empirical data and performance indicators that are used to rate models. It concludes that the models acquired an accuracy rate higher than 99% when original datasets were re-sampled with ADASYN and SMOTE before they were divided into training and testing sets. So, ML models show a drop in performance in the case of limited computing data resources methods, especially for the logistic regression, which is very precise and scores ADASYN in the F1 sphere. Due to that, the model has only one type of vehicle state representation, and the efficiency

is not perfect. It is not enough.
The research work by El Bouchti et al. describes how financial institutions have grown to additionally leverage increasingly sophisticated technologies like deep learning and machine learning to predict risks, spot fraud, and more effectively facilitate the division of their clients. The field of machine learning partakes widely in Deep Reinforcement Learning (DRL), which introduces a sequential behavior in animals based on the study of animal learning in Markov decision processes. One of these has posited a linkage between the brain processes related to the reward system and the DRL algorithm's functioning, the latest research in financial risk analysis and fraud detection [13]. This research describes how financial institutions have grown to additionally leverage increasingly sophisticated technologies like deep learning and machine learning to predict risks, spot frauds, and more effectively facilitate the division of their clients. The field of machine learning partakes widely in Deep Reinforcement Learning (DRL) introducing in animals a sequential behavior which had been based on the study of animals learning in Markov decision processes. One of his has posited a linkage between the brain processes related to the reward system and the DRL algorithm's functioning, the latest research in financial risk analysis and fraud detection. There are used DRL that will be introduced, two financial applications will show, and How they could be employed are discussed. According to this study, the use of data abnormality detection in critical domains such as cyber security, banking and health care is examined [14]. The causes of the anomalies and their attributes are diverse since the anomalies are rare and unpredictable events in the datasets. Consequently, it is almost impossible to generate training data for every possible abnormality class. In Particular, Reinforcement Learning (RL) techniques are shown in this research article as a strategy for imparting complex behavior in high-dimensional environments. Unlike traditional methods, the RL approach assumes data instead of missing the anomalies and inferring issues. The research focuses on the application of RL that addresses the inherent drawbacks of traditional anomaly detection strategies. The authors have used the Deep RL (Meta-AAD) approach and got 98% accuracy. The research explores both the demand and drawbacks as well as the benefits of using RL for anomaly detection.
This work done by Shen et al. highlights two shortcomings of downstream fraud alert systems – first, such systems use human procedures impromptu and have not yet been well adopted, and, secondly, they apply machine learning models to the fraud detection systems [15]. Using anomaly detection to find and explore abnormal items in the data is described in Sec. Machine learning is considered to be one of the most practical approaches despite the fact that it has its drawbacks when applied to large unlabeled datasets. In anomaly detection, Deep Reinforcement Learning (DRL) methods perform better than existing supervised and unsupervised models. In this work, a Systematic Literature Review (SLR) on DRL methods is provided that is used for detecting anomalies across various domains.This review concludes that DRL offers promise for further study within the field while also suggesting some advice or guidelines to those who would want to do more work in this area [16].
The section addresses the anomaly detection issue that is common across many significant applications which include a few annotated examples of anomalies and a greater number of unannotated examples. Existing techniques typically involve either applying unsupervised learning over non-annotated information sets or concentrating exclusively on certain limited instances of labeled anomalies, which may not be the entire universe of possible anomalies [17].
There are important tests on real datasets that show the model outperforms quite significantly. As proof, they employed 48 actual datasets with the Deep Q-learning with Partly Labeled Anomalies (DPLAN) approach, with accuracy rates ranging from 23% to 98%.
The article titled "Q-Credit Card Fraud Detector (Q-CCFD) for Imbalanced Classification using Reinforcement Learning" examines how to use reinforcement learning (RL) methods to address the problem of uneven categorization in credit card fraud detection. Using an unbalanced dataset made up of credit card users' transactions, the study tackles the problem of credit card fraud detection. The Q-CCFD system integrates a Q-learning algorithm with Deep Learning, Auto-Encoder, and Neural Agents as Artificial Intelligence approaches. This work is important because it provides a new alternative for combining present artificial intelligence methods with reinforcement learning strategies, which are used in combating the difficult issue of detecting credit card fraud [8].
According to the paper "RDQN: Ensemble of Deep Neural Network with Reinforcement Learning in Classification Based on Rough Set Theory for Digital Transactional Fraud Detection", there is a big worry about fake deals in financial services [18].
In this way, after doing some research on the previously used approaches for fraudulent transaction detection, a gap in the existing fraud detection mechanism was seen. These methodologies have given promising results on the benchmark datasets using traditional machine learning pipelines and several other approaches. Still, when modern fraudsters and their seamless attacks on financial institutions are observed, an innovative approach to deal with the ever-changing world of technology has to be established. Therefore, after observing multiple methodologies to deal with the problem, it was found that Reinforcement Learning (RL) would give promising results for this problem. RL is equipped with its dynamic nature to explore the search space and deal with the real time scenarios, without much supervision. Hence, in this paper, a novel approach to credit card fraud detection using DQN in Reinforcement Learning has been proposed.

3. Proposed methodology

This section discusses how the dataset used for this research work is collected; it describes the implementation Overview, the stepwise discussion about the algorithm, the training process, and the selection of hyperparameters.

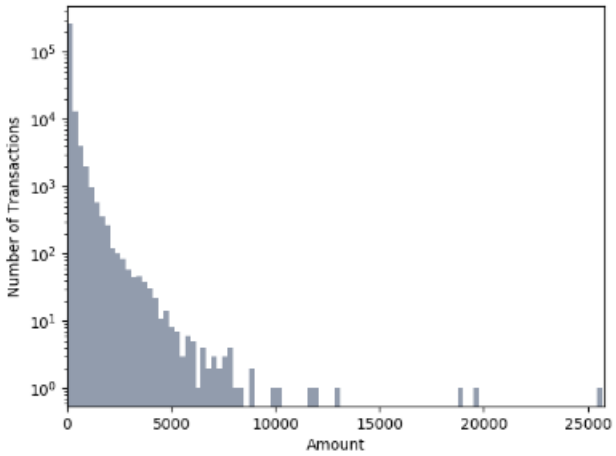

Figure 1. Distribution of dataset parameters

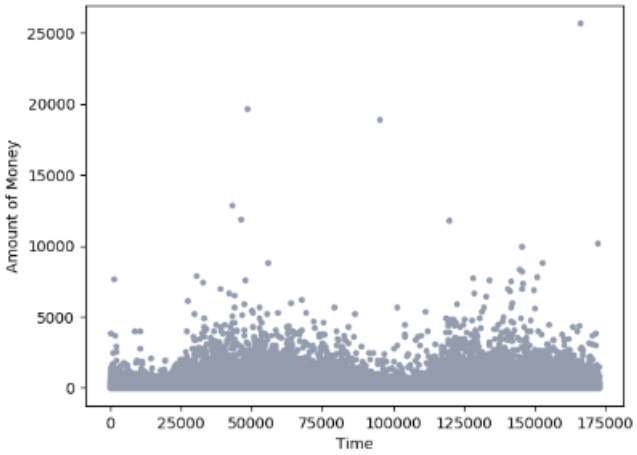

Figure 2. Distribution of parameters

A. *Dataset Description and Collection*

The dataset, which represents credit card transactions made in two days by European cardholders in September 2013, shows a significant imbalance between legitimate and fraudulent transactions - 492 out of 284 of which 807 transactions (MLG-Kaggle, 2015) [19]. This kind of big business is known for its fraudulent activity. The Principal Component Analysis (PCA) transformation has been used to conceal numerical variables in the database. The original attributes of the data are protected by these 28 features, which go by the numbers V1, V2,..., and V28. They are irreversible and hold users' personal data. The two distinct attributes that PCA hasn't changed are amount and time. In addition, there is a basic value called Class that is a significant variable in the database. The money involved in each transaction is indicated by the feature named Amount. The average transaction size in this data set [20] is $ 88.3 %. Figure 1 and Figure **??** illustrate how only a tiny number of transactions approximate the greatest value discovered, with the majority of the data concentrated at very small values near zero. However, as can be seen in Figure 3, the money for each transaction is represented, with some amounts differing from the others.

These transactions, which involve a substantial money transfer, are referred to as outliers in this context. Information now available demonstrates that scammers regularly moved tiny sums of money to carry out their theft. The class characteristic, which shows whether or not the transactions are fraudulent, determines the value of the target variable, which takes a value of 0 in the absence of fraud and 1 in the case of fraud. This feature demonstrates that a minimum amount of fake cases exist, accounting for 0.17 % of the total data. While 99.83 % of cases are not fake. The data is found to be highly unbalanced, necessitating the selection of suitable metrics in order to partition the data and ensure that the system is trained effectively.

B. *Implementation Overview*

The speed at which electronic commerce technology is developing has led to a significant surge in the usage of credit cards. The class characteristic, which shows whether or not the transactions are fraudulent, determines the value of the target variable, which takes a value of 0 in the absence of fraud and 1 in the case of fraud. For many years, supervised machine learning models have consistently produced fraud detection results that are at the cutting edge of the field. In this study, a novel deep Q-network layout for deep reinforcement learning the agent is provided, utilizing Experience Replay and value function approximation in conjunction with an OpenAI Gym environment tailored to the specific needs. In accordance with input batches, the deep Q-agent performs classification action using the epsilon-greedy policy. Following that, the agent receives rewards from the OpenAI environment determined by how well it evaluates its behaviors. The agent has a recall of this entire encounter. After a batch is completed, the deep Q-agent takes a sample of memory from its experience buffer, computes the loss, changes the weights using back-propagation, and updates the Q-value using the Q-network. The results show that the model has achieved state-of-the-art performance and has been able to identify fraudulent transactions and those that are not.
To implement this approach in practice, first, a custom environment using openai Gym was created. To keep it clean and simple, a new conda virtual environment was initialized. Once all the packages required to run the project are installed, this environment can be shared using an 'environment.yml' file to replicate the environment onto any other machine as well. Installation of the gym library is required once the virtual environment has been created. The gym atmosphere has been termed "gym-fraud." The folder includes an 'envs' directory that will include information on the environment and an initialization file that is used to register each gym environment. This file maps an ID to the environment's entry point. Information for spreading initialization of the gym-fraud environment may be found

```
BUFFER_SIZE = int(1e5)  # replay buffer size
BATCH_SIZE = 64         # minibatch size
GAMMA = 0.99            # discount factor
TAU = 1e-3              # for soft update of target parameters
LR = 5e-4               # learning rate
UPDATE_EVERY = 4        # how often to update the network
EPSILON = 0.8           # probability of choosing on-policy action
```

Figure 3. Hyperparameters

in the setup.py file. Another initialization file in the envs directory is used to import environments from individual class files into the gym. The necessary methods comprise the skeleton environment that has been included in the $fraud_env.py$ file. The main file to be run is the iPython notebook which contains code to use the environment and train the DQN algorithm for effective classification of fraudulent transactions. To run the project, first start by installing the 'gym-fraud' custom environment package that was created and create the 'fraud-v0' environment instance. After this, load the credit card transaction data and preprocess that using data normalization. The algorithm hyperparameters are as shown in Figure 3. The goal in this research study is to investigate the approach used in the context of reinforcement learning tasks to implement and assess the Deep Q-Learning (DQL) algorithm. One of the main drawbacks of Q-learning is that it becomes impractical when dealing with large state spaces since the size of the Q-table grows exponentially with the number of states and actions. In these situations, the procedure grows computationally costly and necessitates a large amount of memory in order to store the Q-values. Envision a game where each state consists of 1000 actions. One million cells in a table would be required. And when you compare that to chess or go, that is a relatively small state space. Furthermore, because Q-learning is unable to deduce the Q-value of a new state from a previous one, it cannot be applied to unknown states.
This brings up two points: First, the memory required to store and update the table increases with the number of states. Secondly, it would be unfeasible to devote the time required to examine each state in order to produce the required Q-table. An alternate strategy to address this problem is to integrate deep neural networks and Q-learning. This method is called Deep Q-Learning (DQL), the working is as explained in Figure 4. For every (state, action) pair, the neural networks in DQL serve as the approximation of the Q-value. The state is fed into the neural network, which generates the Q-values for every action that can be taken. Figure 5illustrates the DQL method explained above [21].

C. Algorithm Details

The algorithm used for this project is widely known as Deep Q-Learning (DQL), which uses Deep Q-Network to function [22]. By employing the Bellman equation to repeatedly improve the estimates, Q-learning seeks to discover the ideal action-value function (Q). Experience replay

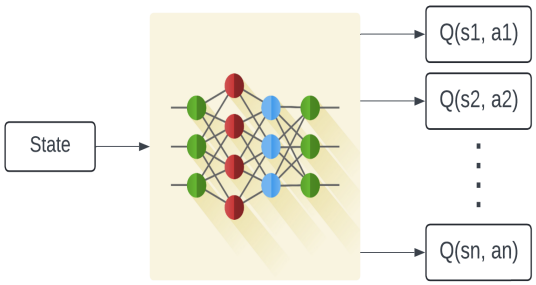

Figure 4. Deep Q-Learning

helps stabilize training by randomly sampling transitions from the replay memory, reducing correlations between consecutive updates [23]. The stepwise algorithm is as explained here.

1) Initialization

- Training Data (D) The algorithm begins with a set of training data, denoted as D. Each data point consists of a state (z) and an associated action (p).
- Episode Number (N)
  The total number of episodes for training.
- Replay Memory (M)
  Initialize a replay memory with a capacity of N. This memory stores transitions (state, action, reward, next state, done flag) encountered during training.
- Value Function (Q)
  Initialize the value function Q with random weights θ. The estimated cumulative payoff (Q) for performing a certain action in a particular condition is described as:
- Simulation Environments ϵ
  Set up the simulation environments for training.

2) Training Loop

For each episode ($n = 1$ to N):

- Shuffle the training data D.
- Initialize the current state (s_1) using the first data point (z_1).
- For each time step ($t = 1$ to T):
  - Choose an action a_t based on an exploration strategy (e.g., ϵ-greedy). With probability ϵ, select a random action; otherwise, use the current policy $\pi(s_t)$ to determine the action.
  - After taking action in the surroundings, you will get rewarded (r_t) and move on to the next state (s_{t+1}).
  - Replay memory M should include the transition ($a_t, s_t, r_t, s_{t+1}, \text{done}_t$).

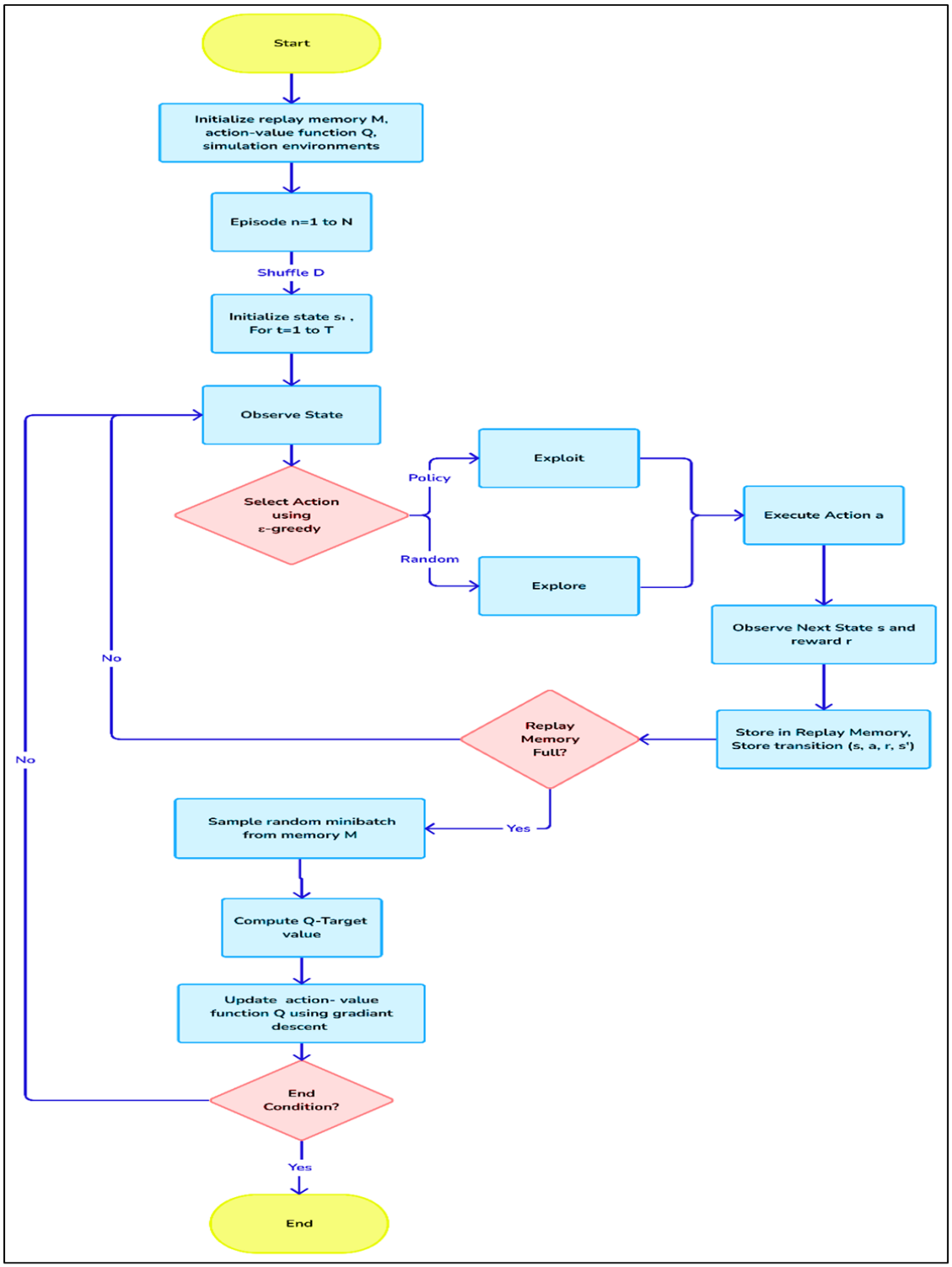

Figure 5. Deep Q Learning Algorithm with experience replay

- A batch of transitions from $M(s_j, a_j, r_j, s_{j+1}, \text{done}_j)$ should be sampled.
- Compute the target Q-value:
  - If done_j is True (indicating the end of an episode), set $y_j = r_j$.
  - Otherwise, set

$$y_j = r_j + \gamma \cdot \max Q(s_{j+1}, a'; \theta) \quad (1)$$

    where a' represents the action that maximizes Q in the next state.
- For updating the Q-network weights θ, perform a gradient descent step. The loss function is

$$L(\theta) = (y_j - Q(s_j, a_j; \theta))^2 \quad (2)$$

- If done_t is True, break out of the loop (end the episode).

3) Hyper parameter Tuning

1) Learning Rate
The learning rate is a hyperparameter that determines the magnitude of the step to be made while updating the model's parameters at each iteration of the optimization process, which is often gradient descent. In essence, it indicates how quickly or slowly a model learns from the input. The learning rate is crucial since it might have an impact on how well the optimization method converges. An excessively low learning rate might cause the model to stall out at local minima or take longer to converge [24]. However, if the learning rate is set too high, the optimization process can never converge and instead overshoot the minimum. Thus, a decent learning rate is essential for this, and we've experimented with different learning rates to observe how the results change.
2) TAU
The "soft update of target parameters" (or TAU) is a machine learning approach that is mostly used in deep reinforcement learning with neural networks as well as other machine learning domains [25]. The process of updating target parameters gradually or smoothly over time, as opposed to abruptly, is known as the "soft update" of goal parameters. Typically, to do this, the current target parameters and primary parameters are added together using a weighted method, where the weight corresponds to the level of updating. The weight is sometimes set via a so-called "tau" parameter that sets the pace at which the target parameters are updated. Soft updates can eliminate oscillations or divergence in training and stabilize the training process. This is often applied to deep reinforcement learning algorithms, including deep Q-learning networks and deep deterministic policy gradients, to stabilize and facilitate their convergence.

4) Termination

The algorithm continues training for N episodes, updating the Q-values based on experience replay.

4. Results

The accuracy score of several models that have been proposed for categorizing fraudulent and non-fraudulent transactions is shown in Table 1. After a significant amount of training, the model was able to distinguish between fraudulent and non-fraudulent transactions with 90.54% accuracy. A straightforward technique that merely compares the total number of completed transactions to the total transactions that the model correctly recognizes is used to determine the accuracy of the model. More specifically, the formula is the ratio of all transactions now under consideration to all transactions that the model correctly predicted [26]. The accuracy (Acc) of the model may be stated mathematically as the equation 3.

$$\text{ACC} = \frac{\text{Number of correctly predicted transactions}}{\text{Total Number of Transactions}} \quad (3)$$

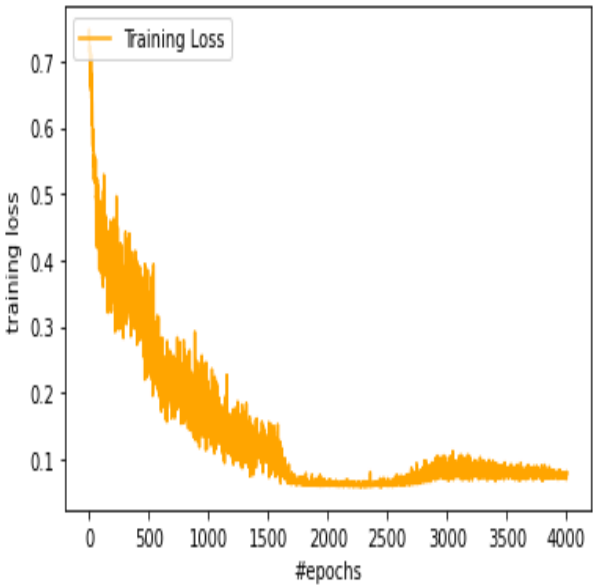

Figure 6. Loss vs epochs

The results demonstrate that artificial neural networks outperform other models when faced with a classification problem, such as 99% accuracy in detecting credit card fraud.. The random forest model and logistic regression technique seem promising for this sample. Their false positive rate is minimal, and their genuine positive rate is high. This model showed state-of-the-art performance on a highly skewed credit-card fraud data set, successfully categorizing fraudulent as well as non-fraudulent events for 94.50% times. The precision of this nature opens up a lot of room for the examination of how much potential there is in reinforcement learning applications to categorization problems and in making decisions, even as enhanced performance could be realized if better rewards were developed as well as more adaptive tuning of hyperparameters in the

TABLE I. Comparative study of results using different learning rate

LR	Buffer size	Batch size	TAU	Gamma	Epsilon	Accuracy
5.00E-04	100000	64	1.00E-03	0.99	0.2	90.55
5.00E-03	100000	64	1.00E-03	0.99	0.2	90.57
2.00E-04	100000	64	1.00E-03	0.99	0.2	90.52

TABLE II. Comparative study of results using different learning rate

TAU	Buffer size	Batch size	LR	Gamma	Epsilon	Accuracy
1.00E-03	100000	64	5.00E-04	0.99	0.2	90.55
5.00E-03	100000	64	5.00E-04	0.99	0.2	90.53
0.9	100000	64	5.00E-04	0.99	0.2	90.56

instant case in order to improve its accuracy. Figure 6 shows how training loss is changing with increasing epochs.

Additionally, the comparative analysis of the outcomes after adjusting hyperparameters such as learning rate and TAU is displayed in Tables 1 and 2. The findings below show that the design is not significantly affected by modifications to these two parameters.

5. CONCLUSION

This research proposed a new approach to the problem of fraudulent credit card transactions using Deep Q-Network. The benchmark dataset for credit card fraud transactions was used in the system's construction. Detecting 90.54% of all fraudulent transactions thereby, the suggested model could keep up with different patterns of money transactions, adjusting to it every time. The suggested method is adaptable and scalable, making it appropriate for near-real-time identification of fraud in credit card transactions. By expanding this system to cover diverse operating system platforms, financial institutions can easily prevent fraudulent transactions, thus improving the financial security of citizens as well as businesses.
In the future, intricate deep reinforcement learning architectures will be explored, which involve strategies such as graph neural networks and attention mechanisms for credit card fraud detection. Subsequent research endeavors may employ diverse, sophisticated deep reinforcement learning algorithms, such as Double Deep Q-network and Dueling DQN, to conduct comprehensive performance evaluations. Pertaining to different datasets, the subsequent investigations evaluate the functionality of the algorithm.

REFERENCES

[1] K. J. Barker, J. D'amato, and P. Sheridon, "Credit card fraud: awareness and prevention," *Journal of financial crime*, vol. 15, no. 4, pp. 398–410, 2008.

[2] W. Ming-Yen Teoh, S. Choy Chong, B. Lin, and J. Wei Chua, "Factors affecting consumers' perception of electronic payment: an empirical analysis," *Internet Research*, vol. 23, no. 4, pp. 465–485, 2013.

[3] B. Vyas, "Java in action: Ai for fraud detection and prevention," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 58–69, 2023.

[4] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: a review of anomaly detection techniques and recent advances," *Expert systems With applications*, vol. 193, p. 116429, 2022.

[5] F. Carcillo, Y.-A. Le Borgne, O. Caelen, Y. Kessaci, F. Oblé, and G. Bontempi, "Combining unsupervised and supervised learning in credit card fraud detection," *Information sciences*, vol. 557, pp. 317–331, 2021.

[6] R. Nian, J. Liu, and B. Huang, "A review on reinforcement learning: Introduction and applications in industrial process control," *Computers & Chemical Engineering*, vol. 139, p. 106886, 2020.

[7] S. Bhatore, L. Mohan, and Y. R. Reddy, "Machine learning techniques for credit risk evaluation: a systematic literature review," *Journal of Banking and Financial Technology*, vol. 4, no. 1, pp. 111–138, 2020.

[8] L. Zhinin-Vera, O. Chang, R. Valencia-Ramos, R. Velastegui, G. E. Pilliza, and F. Quinga-Socasi, "Q-credit card fraud detector for imbalanced classification using reinforcement learning." in *ICAART (1)*, 2020, pp. 279–286.

[9] S. Vimal, K. Kayathwal, H. Wadhwa, and G. Dhama, "Application of deep reinforcement learning to payment fraud," *arXiv preprint arXiv:2112.04236*, 2021.

[10] A. Qayoom, M. A. Khuhro, K. Kumar, M. Waqas, U. Saeed, S. ur Rehman, Y. Wu, and S. Wang, "A novel approach for credit card fraud transaction detection using deep reinforcement learning scheme," *PeerJ Computer Science*, vol. 10, p. e1998, 2024.

[11] E. A. L. M. Btoush, X. Zhou, R. Gururajan, K. C. Chan, R. Genrich, and P. Sankaran, "A systematic review of literature on credit card cyber fraud detection using machine and deep learning," *PeerJ Computer Science*, vol. 9, p. e1278, 2023.

[12] T. K. Dang, T. C. Tran, L. M. Tuan, and M. V. Tiep, "Machine learning based on resampling approaches and deep reinforcement learning for credit card fraud detection systems," *Applied Sciences*, vol. 11, no. 21, p. 10004, 2021.

[13] A. El Bouchti, A. Chakroun, H. Abbar, and C. Okar, "Fraud detection in banking using deep reinforcement learning," in *2017*

Seventh International Conference on Innovative Computing Technology (INTECH). IEEE, 2017, pp. 58–63.

[14] P. Michalski, "Anomaly detection in the context of reinforcement learning," 2021.

[15] H. Shen and E. Kurshan, "Deep q-network-based adaptive alert threshold selection policy for payment fraud systems in retail banking," in *Proceedings of the First ACM International Conference on AI in Finance*, 2020, pp. 1–7.

[16] K. Arshad, R. F. Ali, A. Muneer, I. A. Aziz, S. Naseer, N. S. Khan, and S. M. Taib, "Deep reinforcement learning for anomaly detection: A systematic review," *IEEE Access*, vol. 10, pp. 124 017–124 035, 2022.

[17] G. Pang, A. van den Hengel, C. Shen, and L. Cao, "Toward deep supervised anomaly detection: Reinforcement learning from partially labeled anomaly data," in *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*, 2021, pp. 1298–1308.

[18] C. G. Tekkali and K. Natarajan, "Rdqn: ensemble of deep neural network with reinforcement learning in classification based on rough set theory for digital transactional fraud detection," *Complex & Intelligent Systems*, vol. 9, no. 5, pp. 5313–5332, 2023.

[19] M. L. G. ULB, "Credit card fraud detection," 2018, accessed: 2024-06-10. [Online]. Available: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

[20] S. Georgieva, M. Markova, and V. Pavlov, "Using neural network for credit card fraud detection," in *AIP Conference Proceedings*, vol. 2159, no. 1. AIP Publishing, 2019.

[21] A. Soshin, "Q-learning vs deep q-learning vs deep q-network," 2021, accessed: 2024-06-10. [Online]. Available: https://www.baeldung.com/cs/q-learning-vs-deep-q-learning-vs-deep-q-network

[22] J. Fan, Z. Wang, Y. Xie, and Z. Yang, "A theoretical analysis of deep q-learning," in *Learning for dynamics and control*. PMLR, 2020, pp. 486–489.

[23] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, "Prioritized experience replay," *arXiv preprint arXiv:1511.05952*, 2015.

[24] G. D. Magoulas, M. N. Vrahatis, and G. S. Androulakis, "Improving the convergence of the backpropagation algorithm using learning rate adaptation methods," *Neural Computation*, vol. 11, no. 7, pp. 1769–1796, 1999.

[25] T. Kobayashi and W. E. L. Ilboudo, "T-soft update of target network for deep reinforcement learning," *Neural Networks*, vol. 136, pp. 63–71, 2021.

[26] F. Itoo, Meenakshi, and S. Singh, "Comparison and analysis of logistic regression, naïve bayes and knn machine learning algorithms for credit card fraud detection," *International Journal of Information Technology*, vol. 13, no. 4, pp. 1503–1511, 2021.

Smita Mahajan Dr. Smita Mahajan, an Assistant Professor at Symbiosis Institute of Technology in Pune, India, holds a PhD. in Computer Science Engineering. She is passionate about advancing AI and ML techniques, with research interests in developing targeted language models and chatbots, applying reinforcement learning to real-world problems, and using AI in healthcare, digitization, time series analysis, Industry 4.0, sensor networks, medical imaging, and sustainability. Dr. Mahajan is dedicated to exploring the intersection of AI and various applied domains to drive innovation and practical applications.

Shrikrishna Kolhar received a B.E. degree in Electronics and Telecom-munication Engineering at Savitribai Phule Pune University, Pune, India. He received an M.E. degree in Electronics Digital Systems at Savitribai Phule Pune University, Pune, India. He received his PhD degree in Electronics and Telecommunication Engineering at Symbiosis International (Deemed University), Pune, India. He is an Assistant Professor at Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India. His areas of interest include digital image processing, medical image processing, pattern recognition, computer vision, machine learning, and deep learning.

Ayushi Patil She received a B.E. degree in Computer Science at Savitribai Phule Pune University, Pune, India. She is currently pursuing M.Tech in Artificial Intelligence and Machine Learning at Symbiosis International (Deemed University), Pune, India. Her areas of interest include Reinforcement Learning, Natural Language Processing, and Image Processing.

Shreya Mahajan She received B.E. degree in Information Technology at Savitribai Phule Pune University, Pune, India. She is currently pursuing an M.Tech in Artificial Intelligence and Machine Learning at Symbiosis International (Deemed University), Pune, India. Her areas of interest include Reinforcement Learning, Natural Language Processing, and Image Processing.

Jinal Menpara She received a B.Tech degree in Computer Science at U.V. Patel College of Engineering, Ganpat University, Gujrat, India. She is currently pursuing M.Tech in Artificial Intelligence and Machine Learning at Symbiosis International (Deemed University), Pune, India. Her areas of interest include Reinforcement Learning, Natural Language Processing, and Image Processing.

Amay Mahajan He received his Bachelor's in Computer Engineering from Savitribai Phule Pune University. He received his Masters of Quantitative Finance from Rutgers University, Newark, New Jersey, USA. He is currently working as a Jr. Risk Analyst at Mirae Asset Global Investments, an Asset Management firm based in New York City. His area of interest includes Risk Management, Portfolio Management, and Financial security valuation (securities including but not limited to Equities, Fixed Income, Derivatives, Asset Backed Securities).