IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

Fraudulent Banking Transaction Classification Using Deep Learning Algorithm.

P. Manikandaprabhu[1], S. Prasanna[2], K. Sivaranjan[3], R. Senthilkumar[4]
Assistant Professor, Department of Computer Science Engineering[1]
Students, Department of Computer Science and Engineering [2,3,4]
Anjalai Ammal Mahalingam Engineering College, Thiruvarur, India

Abstract: *Financial fraud is a significant problem in the banking industry, and detecting fraudulent transactions is a critical task for banks to protect their customers and maintain trust in the financial system. Traditional rule-based approaches for detecting fraud often rely on pre-defined thresholds and heuristics, which can be circumvented by sophisticated fraudsters. As a result, machine learning techniques have gained increasing attention in recent years for their ability to automatically learn from data and adapt to changing fraud patterns .In this project, we propose a novel approach for classifying fraudulent banking transactions using a deep learning algorithm. Our approach leverages the power of deep neural networks to automatically extract meaningful features from transaction data, and then use these features to accurately classify transactions as either fraudulent or legitimate.*

Keywords: Fraud Prediction

I. INTRODUCTION

In recent years, there has been a significant increase in the volume of financial transactions due to the expansion of financial institutions and the popularity of web-based e-commerce. Fraudulent transactions have become a growing problem in online banking, and fraud detection has always been challenging [However, technology can be a tool to combat fraud. To prevent further possible fraud, it is important to detect the fraud right away after its occurrence . Fraud can be defined as wrongful or criminal deception intended to result in financial or personal gain. There are two mechanisms, fraud prevention and fraud detection, that can be exploited to avoid fraud-related losses. Fraud prevention is a proactive method that stops fraud from happening in the first place. On the other hand, fraud detection is needed when a fraudster attempts a fraudulent transaction. Fraud detection in banking is considered a binary classification problem in which data is classified as legitimate or fraudulent . Because banking data is large in volume and with datasets containing a large amount of transaction data, manually reviewing and finding patterns for fraudulent transactions is either impossible or takes a long time. Therefore, deep learning-based algorithms play a pivotal role in fraud detection and prediction . Deep learning algorithms and high processing power increase the capability of handling large datasets and fraud detection in a more efficient manner. Deep learning algorithms and deep learning also provide fast and efficient solutions to real-time problems. An ideal fraud detection system should detect more fraudulent cases, and the precision of detecting fraudulent cases should be high, i.e., all results should be correctly detected, which will lead to the trust of customers in the bank, and on the other hand, the bank will not suffer losses due to incorrect detection

1.1Multilayer Perceptron

The Multilayer Perceptron (MLP) algorithm is a type of artificial neural network (ANN) that is widely used for various machine learning tasks, including classification, regression, and pattern recognition. MLP is a feedforward neural network, meaning that information flows in one direction, from input nodes through hidden layers to output nodes, without loops or feedback connections. The basic building block of an MLP is the artificial neuron or node, also known as a perceptron. Each neuron receives input from multiple sources, applies a weighted sum of these inputs, passes the sum through an activation function, and produces an output. MLPs consist of multiple layers of interconnected neurons, with an input layer that receives input data, one or more hidden layers that

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

perform computations, and an output layer that produces the final output. Training an MLP involves feeding labeled training data through the network, adjusting the weights of the connections between neurons based on the error between the predicted output and the actual output. This process is repeated for multiple iterations until the model's performance converges to a satisfactory level. Overall, MLP is a versatile and widely used algorithm in the field of machine learning, known for its ability to learn complex patterns and make accurate predictions. Its applications span across various domains, and it continues to be an important tool in the arsenal of machine learning practitioners.

II. BANKING FRAUD PREDICTION

Banking fraud prediction is a complex and sophisticated process that involves the use of advanced technologies and data analysis techniques to identify, prevent, and mitigate fraudulent activities in the banking industry. With the increasing digitization of banking services and the ever-evolving landscape of financial crimes, fraud detection has become a critical area of focus for banks and financial institutions. One of the key elements of banking fraud prediction is data collection. Banks collect vast amounts of data from various sources, including transaction records, customer information, account activity, and external data such as market trends, historical fraud patterns, and social media feeds. This data is gathered and consolidated into a comprehensive dataset for analysis. Data analysis is a crucial step in the banking fraud prediction process. Advanced data analysis techniques, such as machine learning algorithms, statistical analysis, and data visualization, are applied to the collected data to uncover patterns, trends, and anomalies that may indicate potential fraudulent activities. These techniques enable banks to identify unusual or suspicious behaviors, such as abnormal transaction patterns, changes in customer behavior, or discrepancies in account activity, which could be indicative of fraud. Risk assessment is another important element of banking fraud prediction. Banks assess various risk factors associated with different types of transactions, customers, and accounts to determine their likelihood of being fraudulent. Risk scores are assigned to transactions or accounts based on predetermined rules or machine learning models, and transactions or accounts that exhibit high-risk scores are flagged for further investigation. Early detection and prevention are the primary goals of banking fraud prediction. By analyzing data and assessing risks in real-time or near-real-time, banks can proactively identify potential fraudulent activities and take appropriate action to prevent financial losses. This may include blocking suspicious transactions, freezing accounts, or triggering alerts to fraud investigators for further investigation.Banking fraud prediction also involves continuous monitoring and improvement. Fraudsters are constantly coming up with new techniques and strategies to circumvent detection, so banks need to regularly update their models and techniques to stay ahead of emerging threats. This may involve incorporating new data sources, refining algorithms, or adopting new technologies, such as artificial intelligence and machine learning, to enhance fraud detection capabilities.

III. LITERATURE SURVEY

S.NO	TITLES	AUTHOR	YEAR	MERITS	DEMERITS
1	Ecommerce fraud detection through fraud islands and multi-layer machine learning model	J. Nanduri, Y.-W. Liu, K. Yang, and Y. Jia	2020	Improved accuracy, Scalability, Flexibility.	Training data limitations, High in Cost.
2	Elucidation of big data analytics in banking	M. S. Delgosha, N. Hajiheydari, and S. M. Fahimi	2019	Improved customer experience, Better Decision making	Data quality and privacy concerns, High in Cost.
3	Analyzing credit card fraud detection based on machine learning models	R. Almutairi, A. Godavarthi, A. R.Kotha, and E. Ceesay.	2021	Real-time fraud detection, Scalability, Cost effective	Lack of transparency, Overreliance on historical data

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

4	Machine learning based credit card fraud detection	E. F. Malik, K. W. Khaw, B. Belaton, W. P. Wong, and X. Chew	2020	Enhanced customer experience, Improved accuracy	Data quality and privacy concerns, Regulatory compliance

IV. DEEP LEARNING ALGORITHM

Deep learning algorithms are a cutting-edge form of artificial neural networks that have revolutionized the field of machine learning. These algorithms are designed to automatically learn complex patterns and representations from large amounts of data, making them ideal for handling big data and solving complex problems. What sets deep learning algorithms apart is their ability to learn hierarchical representations through multiple interconnected layers of neurons. This hierarchical structure allows them to process data in a highly flexible and adaptive manner, automatically capturing relevant features and patterns at different levels of abstraction. One of the key advantages of deep learning algorithms is their ability to learn directly from raw data, without relying on handcrafted features or rules. This makes them highly effective in tasks where traditional machine learning approaches may fall short, such as image recognition, speech recognition, natural language processing, and playing complex games. Deep learning algorithms can automatically extract features, learn representations, and make predictions with minimal human intervention, making them highly efficient and scalable for large-scale applications..The impact of deep learning algorithms extends beyond traditional machine learning tasks. They have been successfully applied in fields such as healthcare, where they can assist in disease diagnosis and drug discovery, as well as in autonomous vehicles, where they enable advanced perception and decision-making capabilities. Deep learning algorithms are also being used in robotics, speech synthesis, and virtual assistants, among many other applications. In conclusion, deep learning algorithms are powerful and versatile tools that have revolutionized the field of machine learning. Their ability to automatically learn complex patterns from data, adapt to changing environments, and achieve state-of-the-art performance in various domains has made them indispensable for many real-world applications. As research and development in deep learning continue to advance, we can expect to see even more groundbreaking applications and advancements in artificial intelligence.

IV. PROPOSED SYSTEM

Here we use advanced analytics techniques such as deep learning to analyze large amounts of data and to detect fraudulent patterns that are not easily visible through traditional methods .These will help banks to stay ahead from fraudsters and ensure the long-term effectiveness of the system. The system will utilize a deep learning algorithm, such as a convolutional neural network (CNN) or a recurrent neural network (RNN), to learn from a large dataset of labeled banking transaction data.

The system will first preprocess the transaction data, including features such as transaction amount, transaction type, transaction location, and transaction time, to prepare it for input into the deep learning model. The deep learning algorithm will then be trained on this preprocessed data, using labeled examples of both fraudulent and legitimate transactions to learn the underlying patterns and representations of fraudulent transactions.

V. EXISTING CHALLENGES

- One of the biggest issues with traditional fraud detection systems such as Rule-based method is large number of legitimate transactions are flagged as fraudulent, causing inconvenience for customers and wasting resources for banks.
- Another issue with traditional fraud detection methods is the actual fraudulent transactions go undetected, resulting in financial losses for the bank and its customers.
- The existing system may also incorporate real-time transaction monitoring and anomaly detection techniques to identify suspicious transactions in real-time and trigger appropriate actions, such as alerts to fraud analysts or transaction blocking. This can enable timely detection and prevention of fraudulent transactions

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

VI. RELATED WORKS

There have been many works related Fraud prediction. Here are some of them:

- Several research studies have been conducted on fraudulent banking transaction classification using deep learning algorithms. These studies have explored various approaches and techniques to improve the accuracy and efficiency of fraud detection in the banking industry.
- One common approach is the use of convolutional neural networks (CNNs) for feature extraction and classification of transaction data. CNNs have been shown to be effective in automatically learning relevant features from transaction data, such as transaction amount, type, and time, which can be indicative of fraudulent activities. Researchers have proposed different architectures of CNNs, including one-dimensional (1D) CNNs for time-series data and two-dimensional (2D) CNNs for transaction data represented as images, to capture different types of information in the data.
- Recurrent neural networks (RNNs) have also been employed in related work for fraudulent transaction classification, as they can model sequential dependencies in transaction data. Long short-term memory (LSTM) and gated recurrent unit (GRU) are popular RNN variants that have been used for detecting fraudulent activities in time-series transaction data.
- Additionally, researchers have explored the use of ensemble methods, such as stacking and boosting, to combine the predictions of multiple deep learning models for improved fraud detection accuracy. Transfer learning, where pre-trained deep learning models are fine-tuned on the target task of fraudulent transaction classification, has also been investigated to leverage knowledge from related tasks and domains.
- Other techniques that have been employed in related work include data augmentation, feature engineering, and hyperparameter tuning to optimize the performance of deep learning models. Some studies have also incorporated unsupervised learning approaches, such as autoencoders and variational autoencoders, for anomaly detection in transaction data.
- Furthermore, researchers have investigated the use of real-time monitoring and anomaly detection techniques, such as rule-based systems, clustering algorithms, and outlier detection methods, to complement deep learning models and improve the timeliness of fraud detection in dynamic banking environments.
- Overall, the related work in fraudulent banking transaction classification using deep learning algorithms demonstrates the potential of these approaches in accurately detecting and preventing fraudulent activities in the banking industry. However, further research and development are needed to continually advance the state-of-the-art and address the challenges and limitations associated with real-world banking data, evolving fraud patterns, and interpretability of deep learning models.

Deep Learning Models for Fraud Detection:

- "Fraud Detection in Credit Card Transactions Using Deep Learning" - This study proposed a deep learning model based on a combination of CNN and LSTM for fraud detection in credit card transactions. The model achieved high accuracy in detecting fraudulent activities.
- "Deep Autoencoders for Fraud Detection in Banking Transactions" - This study proposed the use of deep autoencoders for feature extraction and anomaly detection in banking transactions, achieving high accuracy in detecting fraudulent activities.

Real-time Fraud Detection:

- "Real-Time Fraud Detection in Credit Card Transactions Using Deep Learning" - This study developed a real-time fraud detection system using LSTM-based deep learning model integrated into a commercial credit card payment gateway, achieving low false positive rates in real-time.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

- "Real-Time Fraud Detection in Mobile Banking Transactions Using Deep Learning" - This study focused on real-time fraud detection in mobile banking transactions and proposed a deep learning model based on LSTM, which was integrated into a mobile banking application for real-time fraud detection.

VII. METHODOLOGY

The methodology for fraudulent banking transaction classification using deep learning algorithm typically involves the following steps:

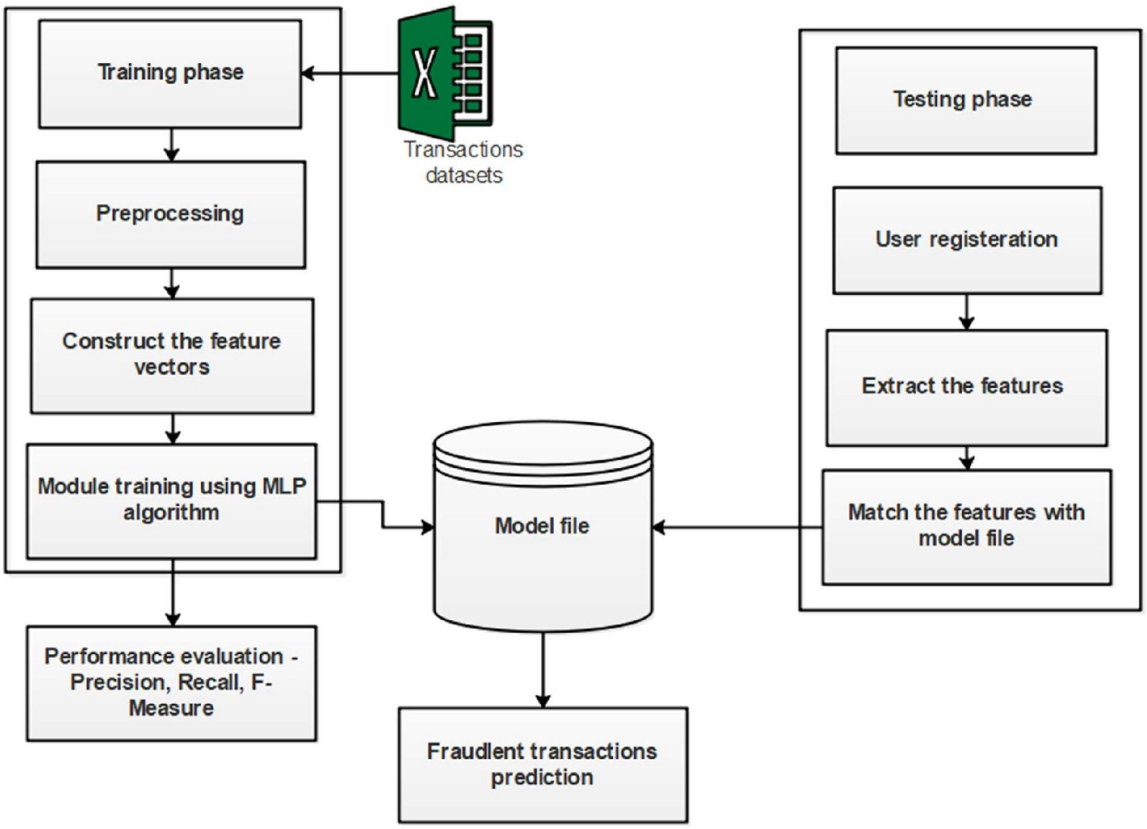
- Data Collection and Preprocessing: The first step is to collect the relevant banking transaction data, including both legitimate and fraudulent transactions. The data may consist of various features such as transaction amount, transaction type, time stamp, and customer information. The data is then preprocessed, which may involve data cleaning, feature extraction, and normalization to prepare it for input into the deep learning model.
- Data Splitting: The preprocessed data is then split into training, validation, and test sets. The training set is used to train the deep learning model, the validation set is used for hyperparameter tuning and model selection, and the test set is used to evaluate the final model's performance.
- Deep Learning Model Development: Next, a deep learning model is developed using appropriate algorithms such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), or their variants like Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRUs). The model architecture is designed to capture relevant patterns and features from the transaction data, with appropriate activation functions, loss functions, and optimization algorithms.
- Model Training and Evaluation: The deep learning model is trained on the training set using the labeled data, and the model's performance is evaluated on the validation set using metrics such as accuracy, precision, recall, F1-score, and area under the Receiver Operating Characteristic (ROC) curve. The model is fine-tuned by adjusting hyperparameters such as learning rate, batch size, and number of layers, based on the validation set performance.
- Model Testing and Performance Assessment: Once the model is trained and fine-tuned, it is evaluated on the test set to assess its performance in real-world scenarios. The performance metrics are calculated, and the model's accuracy in detecting fraudulent transactions is assessed.
- Model Deployment: If the model meets the desired performance criteria, it can be deployed in a real-world banking environment for automated fraud detection. The model can be integrated into the banking system to classify incoming transactions in real-time and flag potential fraudulent transactions for further investigation.

VIII. SYSTEM ARCHITECTURE

The architecture of a heart disease diagnosis system using machine learning algorithms typically includes the following components:

- **Data preprocessing:** The data preprocessing component prepares the collected data for use in machine learning algorithms by cleaning, normalizing, and extracting relevant features.
- **Feature selection:** The feature selection component identifies and selects the most informative features that are most relevant to the diagnosis of heart disease.
- **Deep learning algorithms:** This component includes deep learning algorithms such as neural Network ,DT, and LR that are trained on the selected features to predict the Fraudulent .
- **Model evaluation:** The model evaluation component assesses the accuracy and performance of the trained models using metrics such as accuracy, sensitivity, specificity, and precision.
- **Model selection:** The model selection component selects the best-performing model based on the evaluation metrics.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023



Detailed Design:

MULTI-LAYER PECEPTRON ALGORITHM:

The multilayer perceptron (MLP) is a type of artificial neural network that consists of multiple layers of interconnected nodes, and is commonly used for classification and
regression tasks. Here are the steps involved in training an MLP algorithm:

- **Data preparation**: The first step is to prepare the data for training by dividing it into training, validation, and testing sets, and preprocessing the data by normalizing or standardizing it.
- **Model architecture**: Define the number of layers, nodes, and activation functions for the MLP. The input layer corresponds to the input features, the output layer corresponds to the output variable, and the hidden layers are intermediate layers that
- perform transformations on the data.
- **Weight initialization**: Initialize the weights of the MLP randomly, typically from a Gaussian distribution.
- **Forward propagation**: Feed the input data through the input layer, and then propagate it through the hidden layers using a set of weights and biases, applying activation functions to produce the output of each node.
- **Error calculation**: Calculate the error between the predicted output and the actual output using a loss function such as mean squared error.
- **Backward propagation**: Propagate the error backward through the network, adjusting the weights and biases in each layer using gradient descent to minimize the error.
- **Update weights**: Update the weights using the calculated gradients and a learning rate that determines the size of the weight updates.
- **Repeat**: Repeat steps 4-7 for a certain number of epochs or until convergence is reached.
- **Evaluation**: Evaluate the performance of the trained model on the testing set, and make any necessary adjustments to the model or data preprocessing steps to improve performance.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

IX. CONCLUSION

In conclusion, MLP (Multilayer Perceptron) algorithm is a powerful tool for detecting fraud in behaviors. Its ability to learn patterns and make predictions based on previous data makes it an ideal candidate for detecting fraudulent behavior. By training an MLP model on a dataset of known fraudulent and non- fraudulent behavior, the model can then be used to predict the likelihood of new behavior being fraudulent. The accuracy of the MLP model depends on the quality and quantity of the data used for training. Therefore, it is crucial to use a comprehensive dataset that covers a wide range of fraudulent behaviors to ensure the model can recognize the patterns and characteristics of fraudulent behavior accurately.

X. ACKNOWLEDGMENT

I'd want to use this chance to offer my sincere gratitude to everyone who helped me with the research project report. Throughout the course of my project work, I sincerely appreciate their unwavering direction, priceless constructive feedback, and helpful advice. Their frank and perceptive opinions on numerous project-related issues have been quite beneficial. Additionally, I want to express my gratitude to Principal Dr. S. N. Ramaswamy and the AAMEC management for their unwavering encouragement and support. I owe a great debt of gratitude to my department head, Dr. K. Velmurugan, and my mentor, assistant professor Mr.P. MANIKANDAPRABHU, for their persistent supervision, steadfast oversight, and providing of the necessary information throughout the project

REFERENCES

[1]. Kolli, Chandra Sekhar, and Uma Devi Tatavarthi. "Fraud detection in bank transaction with wrapper model and Harris water optimization-based deep recurrent neural

[2]. network." Kybernetes 50.6 (2020): 1731-1750.

[3]. Minastireanu, Elena-Adriana, and Gabriela Mesnita. "An Analysis of the Most Used Machine Learning Algorithms for Online Fraud Detection." Informatica Economica 23.1 (2019).

[4]. Tiwari, Pooja, et al. "Credit card fraud detection using machine learning: a study." arXiv preprint arXiv:2108.10005 (2021).

[5]. Chen, Joy Iong-Zong, and Kong-Long Lai. "Deep convolution neural network model for credit-card fraud detection and alert." Journal of Artificial Intelligence 3.02 (2021): 101-112.

[6]. Yang, Wensi, et al. "Ffd: A federated learning based method for credit card fraud detection." Big Data–BigData 2019: 8th International Congress, Held as Part of the Services Conference

[7]. Federation, SCF 2019, San Diego, CA, USA, June 25–30, 2019, Proceedings 8. Springer International Publishing, 2019.

[8]. Al-Shabi, M. A. "Credit card fraud detection using autoencoder model in unbalanced datasets." Journal of Advances in Mathematics and Computer Science 33.5 (2019): 1-16.

[9]. John, Hyder, and Sameena Naaz. Credit card fraud detection using local outlier factor and isolation forest. Int. J. Comput. Sci. Eng 7.4 (2019): 1060-1064.

[10]. Abdelrahman, Osama, and Pantea Keikhosrokiani. "Assembly line anomaly detection and root cause analysis using machine learning." IEEE Access 8 (2020): 189661-189672.

[11]. Maniraj, S. P., et al. "Credit card fraud detection using machine learning and data science" International Journal of Engineering Research 8.9 (2019): 110-115.

[12]. Chen, Joy Iong-Zong, and Kong-Long Lai. "Deep convolution neural network model for credit-card fraud detection and alert." Journal of Artificial Intelligence 3.02 (2021): 101-112.

[13]. Tiong, Leslie Ching Ow, and HeeJeong Jasmine Lee. "E-cheating Prevention Measures: Detection of Cheating at Online Examinations Using Deep Learning Approach--A Case Study." arXiv preprint arXiv:2101.09841 (2021).