iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** VI **Month of publication:** June 2024

DOI: https://doi.org/10.22214/ijraset.2024.63375

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VI June 2024- Available at www.ijraset.com

Credit Card Fraud Detection with LIME and SHAP

Pavuluri Swetha[1], Challa Abhishek[2], Alapaty Shivasmika[3]
[1, 2, 3]*Student, Electronics and Communication Engineering, Maulana Azad National Institute of Technology, Bhopal, India*

Abstract: ***The rate of credit card theft has increased in recent years due to the widespread adoption of advanced technology and worldwide communication networks. Scammers are always searching for novel techniques to engage in unlawful activities, irrespective of the fact that credit card fraud results in billions of dollars in losses for individuals and financial institutions annually. Therefore, the absence of fraud detection technologies significantly hampers the successful functioning of banks and other financial organisations. The objective of the project is to create a credit card fraud detection system using machine learning models such as XGBoost and Decision Tree. The accuracy of these models will be assessed using LIME and SHAP models, which provide explanations for the chosen models. Ultimately, a comparison examination of the machine learning models will be conducted.***
Keywords: Template, Scribbr, IEEE, format

I. INTRODUCTION

Multiple types of credit cards, each with its own set of characteristics (e.g., expiration date or spectrum of expiration dates), may be applied to a customer's account. At this time, automated teller machines (ATMs), online banking, and financial institutions utilise the card data. One crucial element for each individual is a unique card number. In addition to the plastic card being safeguarded, the credit card number is also protected.

A surge in credit card fraud has occurred concurrently with the exponential growth of credit card transactions. Credit card fraud occurs when an individual uses their card to accomplish a purchase without physically possessing the necessary funds. Many individuals rely on data mining and other statistical tools to surmount this obstacle. In order to identify fraudulent credit card transactions, artificial intelligence, machine learning, and pattern matching are frequently employed (Pawar, Patil, Martin, & Chaudhari, 2017). Genetic algorithms, a variety of evolutionary algorithm, aim to identify and implement the most effective methods for preventing fraud. The development of a reliable and secure electronic payment system is imperative for the purpose of identifying fraudulent transactions.

As a result of the simplicity with which sensitive data can be transmitted via the Internet and other global communication networks, credit card larceny has increased. It is remarkable that con artists are never satisfied until they discover a new way to take more money; credit card fraud costs individuals and businesses an astounding $19 billion annually. (Pawar, Patil, Martin, & Chaudhari, 2017). Therefore, financial organisations like banks need fraud detection systems to survive.

A. Aims, Objectives

1) The objective of this project is to construct a credit card fraud detection system utilising machine learning models (e.g., XGBoost, Decision Tree, etc.) and subsequently validate its output using LIME and SHAP models (e.g., to assess its explainability). In conclusion, the machine learning models will be contrasted.
2) The aims of the report are outlined below: Conduct a comprehensive literature review.
3) Implement LIME and SHAP alongside ML models.
4) Conduct a comparative analysis.

template Credit cards offer convenience and ease of use, but can also be used by deceitful individuals to accumulate large sums of money. Fraud artists often steal a user's confidential PIN, unaware of the transaction's nature. Detection systems aim to minimize unauthorised use by observing large transactions and distinguishing valid from fraudulent ones. (Meshram & Bhanarkar, 2012)

DIFFICULTIES IN DETECTION OF CREDIT CARD FRAUD

Fraud detection systems face several challenges, including the scarcity of experimental and real-world data, noise, and overlapping data.

To ensure reliability, fraud detection systems must process a small proportion of fraudulent credit card transactions, account for both the funds lost to fraud and those required to detect it, and consider the decision-making layer's actions in response to fraudulent activity detection (Ahmed & Shamsuddin, 2021) (Chaudhary, Yadav, & Mallick, 2012).

Credit card fraud is a significant challenge for modern businesses, with over 10 million illicit money transfers occurring in 2000. Identity thieves often employ constant credit card larceny, with incidents increasing by 21% in 2008. Despite measures to regulate a 1/120th of 1% of all transactions conducted with the intention of facilitating fraud, the financial loss would still amount to billions of dollars (Chaudhary & Mallick, 2019).

Credit card fraud can be committed through physical theft of the card or unauthorised access to the merchant's financial information required to process a legitimate transaction. The primary account number (PAN) is often imprinted directly on the card, while the data is stored on the reverse side on a magnetic stripe (Dong, Huang, Lehane, & Ma, 2020).

Credit card fraud can be perpetrated through various methods, including application deception, unauthorised purchases made with lost or stolen cards, fabrication or modification of cards, and money theft through the interception of credit card transactions. Cybercriminals are also increasingly committing online fraud through cloning websites and creating fraudulent merchant websites (Awoyemi, Adentumbi, & Oluwadare, 2017).

II. LITERATURE REVIEW

Computerized commercial payment systems have led to unlawful access to financial data, causing apprehension in contemporary culture. Online transactions, resulting in financial theft and customer loyalty loss, pose a risk of job loss (Fiore, Santis, Perla, Zanetti, & Palmieri, 2019). This includes enterprises, clients, financial establishments, and merchants. The realm of fraudulent activities requires the adoption of diverse processes to avoid the acquisition of money or assets by deceitful means.

Academics are developing a technique to detect and prevent fraudulent operations, which are causing substantial financial losses. Several methodologies have already been proposed and evaluated in the relevant domain. Several of these are further upon in the subsequent sections. In previous studies, it was demonstrated that techniques such as LR, RF, Decision Tree, Gradient Boosting (GB), and Support Vector Machines (SVM) are effective. In article (Mishra & Ghorpade, 2018), The classifiers GB, LR, RD, SVM, and an ensemble of classifiers were employed, yielding a recall rate over 91% on a European dataset utilising the aforementioned approaches. High accuracy and recall were achieved only after performing under sampling on the data and rebalancing the dataset.

Two approaches that have demonstrated potential in the identification of fraudulent activities are outlier detection and k-nearest neighbours (KNN) (Malini & Pushpa, 2017; Navamani, Phil, & Krishnan, 2018). Supporting evidence suggests that they are capable of enhancing fraud detection rates while reducing false alarm occurrences. In addition, a published investigation found that the KNN method performed exceptionally well when compared to other conventional algorithms (Awoyemi, Adentumbi, & Oluwadare, 2017).

However, unlike the papers mentioned before, this study did a comparison between many traditional techniques and deep learning methodologies. (Kazemi & Zarrabi, 2017). Each of the investigated methodologies yielded results that were around 80% accurate compared to the real value. The researchers (Dhankhad, Far, & Mohammed, 2018) A comparative analysis of several methodologies was conducted utilising a dataset sourced from Europe. The aforementioned classifiers comprised RF, GB, LR, DT, KNN, XGBoost (XGB), MLP, and a layering classifier composed of multiple machine learning classifiers. All of the methodologies demonstrated exceptional accuracy, surpassing 90%, due to the meticulous preparation of the data. The stacking classifier had the most efficacy, with a 95 percent efficiency and a 95 percent recall rate. In study (Wang, Wang, Z Ye, Cai, & Pan, 2018), An assessment was conducted on a neural network using the European dataset, and the outcomes were encouraging. The experiment utilised backpropagation neural networks, with optimisation performed using the Whale approach. Pussirirat and Yan (2020) conducted a study that compared three different datasets. We discovered the potential of algorithms like MLP to prevent and detect credit card fraud by combining Auto-encoder and Restricted Boltzmann Machine techniques.

Using the autoencoder method, (Al-Shabi, 2019) By providing a solution, we addressed the issue of handling divergent data throughout the credit card identification procedure. Its performance outperforms logistic regression in the setting of an unbalanced dataset. In terms of detecting credit card fraud, (Singh, Ranjan, and Tiwari, 2021) The Random Forest (RF) ensemble classification algorithm delivers exceptional results by combining oversampling and under sampling strategies. Oversampling strategies were used to attain parity between the occurrences of the majority and minority groups. Under sampling techniques were used to reduce the bulk of class samples in order to establish an equitable distribution of classes within the dataset.

In their study on credit card fraud detection, (Zou, Zhang, & Jiang, 2019) used a denoising autoencoder (DAE), a neural network approach, to reduce noise. This method has the potential to expand sample size, classify the sampled dataset, and eliminate noise. This method was created to address the issue of an uneven dataset.

The Imblearn module's 'SMOTE' method is used to remove unwanted noise from the datasets. Finally, the denoised and oversampled datasets are fed into a deep fully connected neural network. We successfully found the optimal recall efficiency of 84% and accuracy of 97.93% by running a battery of tests. 2019 (Zou, Zhang, & Jiang).

Roy et al.'s 2018 research, published the same year, concluded that deep learning, a technology, generated findings similar to existing fraud detection systems such as LR and Gradient Boosted Trees. Deep learning, on the other hand, is capable of producing amazing results since it involves complicated network structures and necessitates careful parameter adjustment. During a span of 80 million pre-documented credit-card transactions, (Roy, et al., 2018). The researchers concluded that the LSTM model was significantly more effective than the conventional Auto-encoder Neural Network (ANN). Moreover, they discovered that the magnitude of the network and the computational resources at hand are two additional variables that exert an influence on performance. They initially obtained data properties using an ANN, as stated by Lin and Jiang (2020). In doing so, the dimensionality of the data was decreased).

Local vs Global Interpretability

Linear regression methods forecast individual data points using beta coefficients, demonstrating global faithfulness. Individual differences are accounted for in causal analysis by "average" causal evaluation. Because of the function's linear and monotonic behaviour, local fidelity explanations are more accurate in explaining a single data point or variance owing to several sources (Kopitar, Cilar, Kocbek, & Stiglic, 2020).

When is it necessary to provide explanations of models?

Developing explanations necessitates a substantial commitment of time and energy. Therefore, it is crucial to determine the circumstances in which explanations must be given. (Doshi-Velez, et al., 2017) Formulated a generalisation on three distinct scenarios that necessitate the provision of explanations. Consumers may be entitled to an explanation in some situations, such as when the decision has a significant impact on the users as opposed to the decision-makers.

Furthermore, people want answers when they express curiosity about the dependability of the technology, without casting any doubt on the results. (Doshi-Velez, et al., 2017), Except for the basic circumstances, the text examined the role of explanation from a legislative perspective. They emphasised the necessity of generating explanations to facilitate legally sound judgements.

III. IMPLEMENTATION

DATASET

Digital payments increase security concerns due to cybercrime. Over 5 million records are hacked daily, highlighting fraud in Card-Present and Card-not-Present transactions. The Data Breach Index aims to improve understanding and reduce fraud risks in digital transactions.

Distance_from_home: The distance between the user's home and the location where the transaction took place.

distance_from_last_transaction: The distance between the current transaction's location and the prior transaction's location in space.

ratio_to_median_purchase_price: This measure indicates the relationship between a transaction's purchase price and the median purchase price. It provides useful information about the transaction's financial background.

repeat_retailer: A binary variable that indicates if the transaction originated from the same store as a previous transaction.

used_chip: A binary variable indicating whether or not the transaction was completed.

The credit card contained a chip.

used_pin_number: A binary flag that indicates whether or not the transaction was authenticated using a PIN number.

online_order: A binary variable that indicates whether the transaction is an online order.

Fraud: A binary classification that indicates if the transaction was discovered as fraudulent (1 for fraud, 0 for non-fraud).

This dataset is a valuable resource for researchers looking into the various trends and circumstances that contribute to fraudulent transactions in the ever-changing world of digital payments. (Credit Card Fraud Dataset)

A. Random forest

The Random Forest Classifier in scikit-learn accommodates categories and numerical target variables. It requires dividing the dataset into a training and test set to ensure model information and performance evaluation on unseen data, preventing overfitting.

1) LIME Black Box model interpretation

In order to determine which features, cause the algorithm to veer towards the result, this specific row is chosen. According to LIME, the ratio to the median purchase price, the utilised PIN number, the distance from home, and the distance from the last transaction are the main factors that help the model estimate that the transaction is not fraudulent.

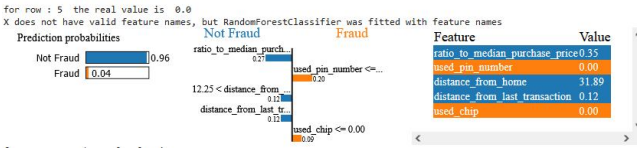


Figure 1: DT - LIME row 5

In below figure, the LIME black box model for decision tree is given where the feature importance for data row 15 where the real value is fraud and the prediction is fraud. The features which indicate whether the prediction is fraud, Ratio to median purchase having value 5, distance from the last transaction is 6.85, the customer has not used the pin number. These are the features which tells the machine learning model the transaction is fraud.

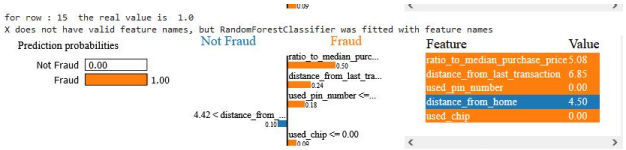


Figure 2: DT – LIME row 15

2) SHAP Black Box model interpretation

In the below figure, the transaction is not fraud, the features ratio to median purchase price, distance from home and distance from last transaction and online orders value, these values represent and tells the model the prediction is not fraud.

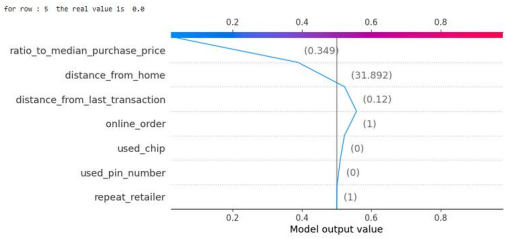


Figure 3: DT - For row 5, real value 0

In the below figure, the transaction is fraud, Due to the features ratio to median purchase price, online order, distance from home, used pin number, distance from last transaction, these are the features due to which the transaction is fraud.

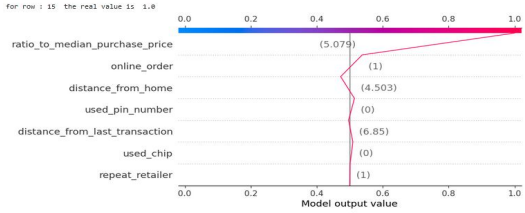


Figure 4: DT - For row 15, real value 1

B. Logistic regression model

1) Lime Black Box model interpretation

In the below figure the lime modules explains prediction of the Logistic regression model, the model predicts for 5th row transaction as not fraud, the reason behind the prediction of the not fraud transaction is explained by lime module, which is ratio to median purchase price be 0.35, distance from home to be 31.89. The pin number and chip has not been used in the transaction. So these are the main features which predicts transaction as not fraud, and are clearly explained by the lime model

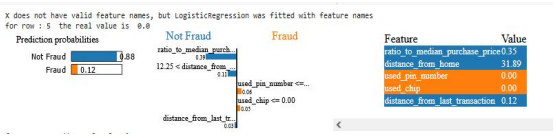


Figure 5: RF – LIME row 5

In the below figure the lime module explains the prediction of logistic regression model for the 15th row the prediction done by the model is fraud, the lime explains that the transaction is fraud because, the ratio to median purchase price is the main reason and the distance from home is 4.50 these values indicate toward the fraud transaction

Figure 6: RF – LIME row 15

2) SHAP Black Box model interpretation

In the below figure the shap module is explaining the reasons behind the prediction of the logistic regression for the 5th row, the model predicts the transaction as not fraud and the shap explains that the ratio to median purchase price, distance from home, online order and the distance from the last transaction are the top 4 features which contributes toward the models prediction as not fraud.

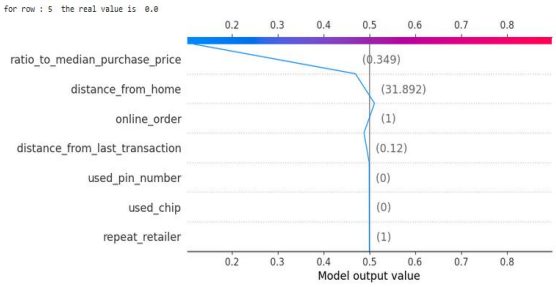


Figure 7: RF - For row 5

In the below figure the SHAP module explains the prediction of logistic regression model for the 15th row the prediction done by the model is fraud, the lime explains that the transaction is fraud because, the ratio to median purchase price is the main reason and the distance from home is 4.50, online order and used chip are the feature values indicates toward the fraud transaction

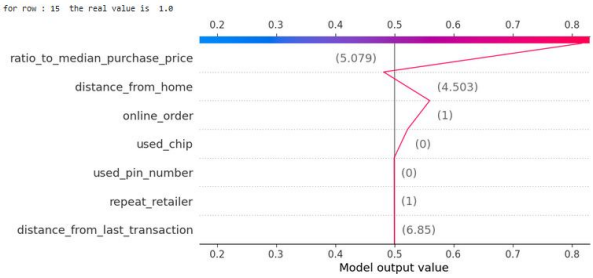


Figure 8: RF - For row 15, real value 0

C. Decision Tree Model

1) Lime Black Box model interpretation

In the below figure the lime modules explains prediction of the decision tree model, the model predicts for 5th row transaction as not fraud, the reason behind the prediction of the not fraud transaction is explained by lime module, which is ratio to median purchase price be 0.28, The pin number and chip has not been used in the transaction. So, these are the main features which predicts transaction as not fraud, and are clearly explained by the lime model.

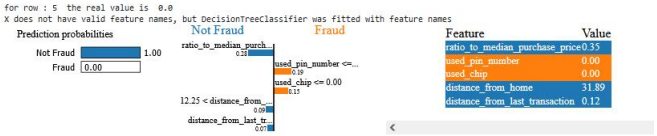


Figure 9: Decision Tree – LIME row 5

In the below figure the lime module explains the prediction of decision tree model for the 15th row the prediction done by the model is fraud, the lime explains that the transaction is fraud because, the ratio to median purchase price is the main reason and the distance from home, used pin number is 0 and used chip is 0

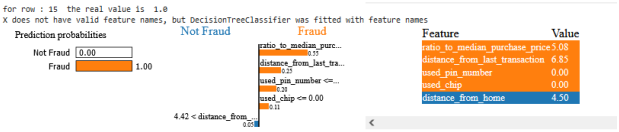


Figure 10: Decision Tree – LIME row 15

2) SHAP Black Box model interpretation

In the below figure the shap module is explaining the reasons behind the prediction of the Decision tree classifier for the 5th row, the model predicts the transaction as not fraud and the shap explains that the ratio to median purchase price, distance from home, distance from last transaction, online order, used chip and used pin number are the features which contributes toward the model's prediction as not fraud.

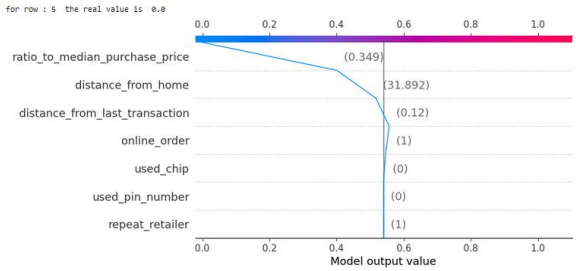


Figure 11: XGB – SHAP row 25

The SHAP module predicts fraud in a 15th-row decision tree model, based on factors such as the transaction's fraud ratio, distance from home, online order, and used chip.

IV. RESULTS

A. Understand how lime and shapely work

The research reveals the transparency of fraud detection models like Logistic Regression, Random Forests, and Decision Trees. It shows that fraudulent transactions are prevalent, with factors like distance to residence, median purchase price, and internet transactions affecting the conclusion's non-fraudulent nature. The interpretability of these methodologies improves understanding of the models' decision-making processes, suggesting that machine learning models can be more transparent and responsible in fraud detection. Two methods for condensing model simulation results are Lime and Shapley.

B. Models Evaluation Results

The graphs below show the methods used to forecast the performance of several machine learning algorithms.

Model Comparison

Model	F Value	Accuracy	Remarks
Random Forest	0.99	0.99	Baseline Model
Logistic Regression	0.95	0.95	Baseline Model
Decision Tree	0.97	0.97	Baseline Model

Figure 12: Accuracy score and F1 score of all models

F1 Value and Accuracy were used to assess the effectiveness of three discrete models in fraud detection: Decision Tree, Random Forest, and Logistic Regression. The study evaluated the effectiveness of three discrete models in fraud detection: Decision Tree, Random Forest, and Logistic Regression. The Random Forest model performed well, with an F Value of 0.99 and 99% accuracy rate. Decision Tree and Logistic Regression were the most effective, with logistic regression trailing behind. Further research and innovation could improve these models' fraud prediction capabilities.

V. CONCLUSION

In conclusion, comparative testing of the three fraud detection models (Random Forest, Logistic Regression, and Decision Tree) revealed that they were effective at detecting fraudulent transactions. The Random Forest model's excellent F Value of 0.99 and 99% accuracy rate revealed its efficiency in detecting fraud. The Logistic Regression and Decision Tree models likewise performed admirably, with accuracy rates of 97% and 95%, respectively, and F Values of 0.97 and 0.95.

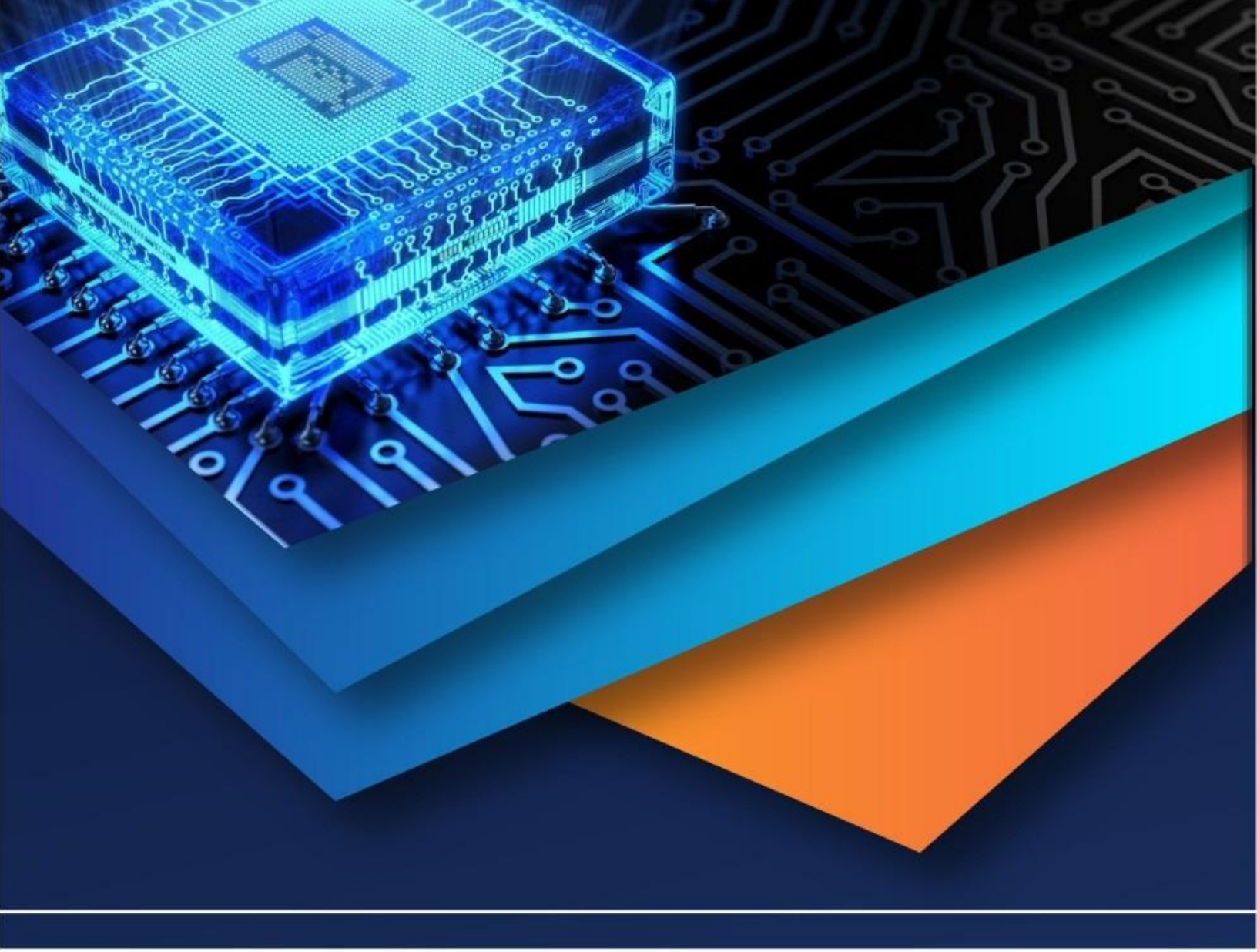
It is possible to gain essential insights into the complexities of decision-making and the pivotal elements that influence projections by using analytical tools such as LIME and SHAP. Consistent interpretability techniques that highlight crucial aspects such as distance from home, ratio to median purchase price, lack of a PIN number and chip usage, and so on, lead to a better understanding of the models' selection criteria.

Although these models provide solid foundations, more research and development could result in even more sophisticated and specialised models, particularly when it comes to addressing the unique complications connected with fraud detection. In conclusion, our research contributes significantly to the ongoing work of constructing reliable and understandable fraud detection programmes. These models are critical for assuring the security of diverse domains and the integrity of financial transactions.

REFERENCES

[1] Adewumi, A. O., & Akinyelu, A. A. (2018). A survey of machine-learning and nature-inspired based credit card fraud detection techniques. International Journal of System Assurance Engineering Management, 8(2), 937-953.
[2] Ahmed, F., & Shamsuddin, R. (2021). A Comparative Study of Credit Card Fraud Detection Using the Combination of Machine Learning Techniques with Data Imbalance Solution. In 2021 2nd International Conference on Computing and Data Science (CDS) (pp. 112-118). Stanford, CA, USA: IEEE.
[3] Al-Shabi, M. (2019). Credit card fraud detection using autoencoder model in unbalanced datasets. Journal of Advances in Mathematics and Computer Science, 33(5), 1-6.
[4] Arya, V., Bellamy, R., Chen, P., Dhurandhar, A., Hind, M., Hoffman, S., . . . Mourad, S. (2019). One Explanation Does Not Fit All: A Toolkit and Taxonomy of AI Explainability Techniques. arXiv preprint arXiv:1909.03012.
[5] Awoyemi, J. O., Adentumbi, A. O., & Oluwadare, S. A. (2017). Credit card fraud detection using Machine Learning Techniques: A Comparative Analysis. Computing Networking and Informatics (ICCNI) 2017 International Conference , (pp. 1-9).
[6] Chaudhary, K., & Mallick, B. (2019). Credit Card Fraud: The study of its impact and detection techniques. International Journal of Computer Science and Network (IJCSN), 1(4), 31-35.
[7] Chaudhary, K., Yadav, J., & Mallick, B. (2012). A review of Fraud Detection Techniques: Credit Card. International Journal of Computer Applications, 45(1), 39-44. Retrieved 8 17, 2022, from https://ijcaonline.org/archives/volume45/number1/6748-8991
[8] Dhankhad, S., Far, B., & Mohammed, E. A. (2018). Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study. IEEE International Conference on Information Reuse and Integration (IRI), (pp. 122-125).
[9] Dong, W., Huang, Y., Lehane, B., & Ma, G. (2020). XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring. Automation in Construction, 114, 103155.
[10] Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., . . . Weller, A. (2017). Accountability of AI under the law: The role of explanation. arXiv:1711.01134.
[11] Du, M., Liu, N., & Hu, X. (2018). Techniques for Interpretable Machine Learning. arXiv:1808.00033.
[12] Farrugia, D., Zerafa, T., Cini, C., Kuasney, B., & Livori, K. (2021). A Real-Time Prescriptive Solution for Explainable Cyber-Fraud Detection Within the iGaming Industry. SN Computer Science, 2(3), 1-9.
[13] Fiore, U., Santis, A. D., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences, 479, 448-455.
[14] Geetha, R., Sivasubramanian, S., & Kaliappan, M. e. (2019). Cervical cancer identification with synthetic minority oversampling technique and PCA analysis using random forest classifier. Journal of Medical Systems, 43, 286.
[15] Hao, M., Hejiang, S., Junjie, L., & Shen, W. (2020). Developing window behavior models for residential buildings using XGBoost algorithm. Energy and Buildings, 205, 109564.
[16] Jackins, V., Vimal, S., Kaliappan, M., & Lee, M. Y. (2021). AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. The Journal of Supercomputing volume, 77, 5198–5219.
[17] Ji, Y. (2021). EXPLAINABLE AI METHODS FOR CREDIT CARD FRAUD DETECTION. Master Degree Project in Informatics with a specialization in Data Science, (pp. 1-49).
[18] Kalaiselvi, N., Rajalakshmi, S., & Padmavathi, J. (2019). Credit card fraud detection using learning to rank approach. International Conference on Computation of Power Energy Information and Communication (ICCPEIC) , (pp. 191-196).
[19] Kazemi, Z., & Zarrabi, H. (2017). Using deep networks for fraud detection in the credit card transactions. Knowledge-Based Engineering and Innovation (KBEI) 2017 IEEE 4th International Conference , (pp. 630-633).
[20] Kopitar, L., Cilar, L., Kocbek, P., & Stiglic, G. (2020). Local vs. Global Interpretability of Machine Learning Models in Type 2 Diabetes Mellitus Screening. Artificial Intelligence in Medicine: Knowledge Representation and Transparent and Explainable Systems , 108-119.
[21] Lakshmi, S. V., & Kavilla, S. D. (2018). Machine Learning For Credit Card Fraud Detection System. International Journal of Applied Engineering Research, 13(24), 16819-16824.
[22] Malini, N., & Pushpa, M. (2017). Analysis on Credit Card Fraud Identification Techniques based on KNN and Outlier Detection. Advances in Electrical Electronics Information Communication and Bio-Informatics (AEEICB) Third International Conference, (p. 25).

[23] Meshram, P. L., & Bhanarkar, P. (2012). Credit And ATM Card Fraud Detection Using Genetic Approach. International journal of engineering research and technology, 1(10). Retrieved 8 17, 2022, from https://ijert.org/research/credit-and-atm-card-fraud-detection-using-genetic-approach-ijertv1is10375.pdf
[24] Mishra, A., & Ghorpade, C. (2018). Credit Card Fraud Detection on the Skewed Data Using Various Classification and Ensemble Techniques. IEEE International Students' Conference on Electrical Electronics and Computer Science (SCEECS) (pp. 1-5). IEEE.
[25] Mishra, P., Biancolillo, A., Roger, J. M., Marinie, F., & Rutledge, D. N. (2020). New data preprocessing trends based on ensemble of multiple preprocessing techniques. TrAC Trends in Analytical Chemistry, 132, 116045.
[26] Modi, H., Lakhani, S., Patel, N., & Patel, V. (2013). Fraud Detection in Credit Card System UsingWeb Mining. International Journal of Innovative Research in Computer and Communication Engineering, 1(2), 175-179. Retrieved 8 17, 2022, from http://ijircce.com/upload/2013/april/5_v1204035_fraud_o.pdf
[27] Mohammed, E., & Far, B. (2019). Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study. IEEE Annals of the History of Computing.
[28] Navamani, C., Phil, M., & Krishnan, S. (2018). Credit Card Nearest Neighbor Based Outlier Detection Techniques. International Journal of Computer Techniques, 5(2), 52-60.
[29] Nayak, A. (2019, December 22). Idea Behind LIME and SHAP. Retrieved from Towards data science: https://towardsdatascience.com/idea-behind-lime-and-shap-b603d35d34eb
[30] Nelli, F. (2018). Python Data Analytics. Berkeley, CA: Apress .
[31] Nicodeme, C. (2020). Build confidence and acceptance of AI-based decision support systems-Explainable and liable AI. In 2020 13th International Conference on Human System Interaction (HSI) (pp. 20-23). Tokyo, Japan: IEEE.
[32] Pawar, A., Patil, V., Martin, S., & Chaudhari, M. S. (2017). Credit Card Fraud Detection Using Hidden Markov Model. Imperial journal of interdisciplinary research, 3(4). Retrieved 8 17, 2022, from http://ijesc.org/upload/93822e4dd508bdd6f0b9644c745036ca.credit card fraud detection using hidden markov model.pdf
[33] Poduska, J. (2018, December 5). SHAP and LIME Python Libraries: Part 1 - Great Explainers, with Pros and Cons to Both. Retrieved from Domino data lab: https://www.dominodatalab.com/blog/shap-lime-python-libraries-part-1-great-explainers-pros-cons
[34] Pumsirirat, A., & Yan, L. (2020). Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine. International journal of advanced computer science and applications, 9(1), 18-25.
[35] Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. (2019). Credit Card Fraud Detection Using AdaBoost and Majority Voting. IEEE Access, 9, 14277 - 14284.
[36] Roy, A., Sun, J., Mahoney, R., Alonzi, L., Adams, S., & Beling, P. (2018). Deep learning detecting fraud in credit card transactions. In 2018 Systems and Information Engineering Design Symposium (SIEDS) (pp. 129-134). Charlottesville, VA, USA: IEEE.
[37] Selvani, L., & Kavila, D. (2018). Machine Learning For Credit Card Fraud Detection System. International Journal of Applied Engineering Research, 13(24), 16819-16824.
[38] Shiyang, X. (2018). Random Forest for Credit Card Fraud Detection. IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), (pp. 101-115).
[39] Singh, A., Ranjan, R., & Tiwari, A. (2021). Credit Card Fraud Detection under Extreme Imbalanced Data: A Comparative Study of Data-level Algorithms. Journal of Experimental & Theoretical Artificial Intelligence, 1-28.
[40] Somvanshi, M., Chavan, P., Tambade, S., & Shinde, S. V. (2017). A review of machine learning techniques using decision tree and support vector machine. International Conference on Computing Communication Control and automation (ICCUBEA). Pune, India: IEEE.
[41] Wang, C., Wang, Y., Z Ye, L. Y., Cai, W., & Pan, S. (2018). Credit card fraud detection based on whale algorithm optimized BP neural network. 13th International Conference on Computer Science & Education (ICCSE), (pp. 1-4).
[42] Wu, T., & Wang, Y. (2021). Locally Interpretable One-Class Anomaly Detection for Credit Card Fraud Detection. arXiv preprint arXiv:2108.02501.
[43] Zoldi, S. (2017). Explainable AI in Fraud Detection - A Back to the Future Story. Retrieved from FICO: https://www.fico.com/blogs/explainable-ai-fraud-detection-back-future-story
[44] Zou, J., Zhang, J., & Jiang, P. (2019). Credit card fraud detection using autoencoder neural network. arXiv:1908.11553.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY