Cureus Journal of Computer Science

Open Access Original Article

Explanation of Prediction of Machine Learning Models Using Local Interpretable Model-Agnostics Explanation and SHapley Additive exPlanations With Case Study of Credit Card Defaulters

Received 04/29/2025
Review began 05/15/2025
Review ended 08/12/2025
Published 09/01/2025

DOI:
https://doi.org/10.7759/s44389-025-05988-x

Indu Arora [1]

1. Department of Computer Science and Applications, Mehr Chand Mahajan D.A.V. College for Women, Chandigarh, IND

Corresponding author: Indu Arora, indu.arora@mcmdavcwchd.in

Abstract

Artificial intelligence (AI) models, including machine learning and deep learning, play a crucial role in decision-making and predictive analysis across various fields. Such models often act as black boxes, as the basis of prediction is not provided. At times, understanding the basis of prediction is essential to assess the reliability of the model. To address this issue, the Explainable AI (XAI) technique is found to be useful. XAI helps in understanding how and why a model makes a certain prediction. This research paper focuses on feature-based techniques of XAI - SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) - which help in identifying the key features influencing the decisions of prediction models. These techniques are applied to two models, Random Forest and Decision Tree, developed in Python using a dataset from a case study titled Credit Card Defaulter, which identifies good or bad credit applicants. Two random scenarios are used to check the reliability of the prediction process. It has been observed that out of 12 features, the features code_gender, flag_own_car, flag_own_reality, and cnt_fam_members contributed the most to predicting good or bad credit applicants. The SHAP and LIME techniques used in explanation of prediction enhance the trust on the models used and identify the major contributing factors influencing the predictions.

Categories: AI applications, Explainable AI, Machine Learning (ML)
Keywords: artificial intelligence, xai, lime, shap, machine learning

Introduction

Explainable artificial intelligence (XAI) encompasses a set of processes and methods that help in understanding the justification of the prediction or decision made using artificial intelligence (AI) models, including machine learning and deep learning models. XAI helps in finding factors that influence the prediction. There are four types of XAI techniques: data explainability, model explainability, feature-based techniques and example-based techniques [1]. The objective of this paper is to explain the use of feature-based techniques of LIME (Local Interpretable Model-Agnostics Explanation) and SHAP (SHapely Additive exPlanation) with the help of a case study, Credit Card Defaulter, in identifying the most important features such as code_gender, flag_own_car, flag_own_reality and cnt_fam_members, which have impacted the prediction. This paper also compares the results of the SHAP and LIME, feature-based XAI techniques, in Random Forest and Decision Tree AI models.

Materials And Methods

Background of XAI

AI is a set of technologies and techniques that enable computers to perform tasks that would typically require human interaction. AI enhances the intelligence and capabilities of any application [2]. AI consists of two words: Artificial, which means non-natural, and Intelligence, which refers to the ability to understand and think. AI is basically the study of making computers work like the human brain. The major focus of AI is on learning, reasoning, and self-correction. Machine Learning (ML), a subset of AI, comprises algorithms that enable systems (computers) to learn from data. ML uses statistical and mathematical methods to improve the learning experience [3]. Deep Learning (DL), a further subset of ML, makes use of neural networks. DL models provide more accuracy than ML models and are of two types: black box and white box [4].

Black box models arrive at a decision without providing detailed explanations or processes. These models are generally more accurate than white box models, but such models cannot be interpreted easily. They are used to model complex scenarios with deep and non-linear interactions between data. DL models, boosting models, and random forest models are examples of black box models. Neural networks used in DL

How to cite this article
Arora I (September 01, 2025) Explanation of Prediction of Machine Learning Models Using Local Interpretable Model-Agnostics Explanation and SHapley Additive exPlanations With Case Study of Credit Card Defaulters. Cureus J Comput Sci 2 : es44389-025-05988-x. DOI https://doi.org/10.7759/s44389-025-05988-x

are difficult to understand and interpret. DL models are built over layers and nodes and also use weights and biases, which are not easy to interpret when used in AI modelling.

White box models are explainable and describe the processes followed in reaching the decision. White box models help to understand how these models behave, process data, and give weightage to variables and how the predictions are made. Linear tree, regression tree and decision tree are few models that fall into this category. Since these models help in understanding the purpose of rationale and decision-making process in such a way that an average person can understand it, these are popularly known as XAI [5-6]. XAI strengthens the reliability of results produced by ML or DL models or other AI techniques. XAI describes AI model, its impact, biases, accuracy and transparency. XAI is being used in many business and research applications for better decisions [7-8]. XAI can be defined as "a collection of procedures and techniques that enable ML algorithms to produce results with explanation and reliability". Figure *1* shows the working of ML and ML with XAI.

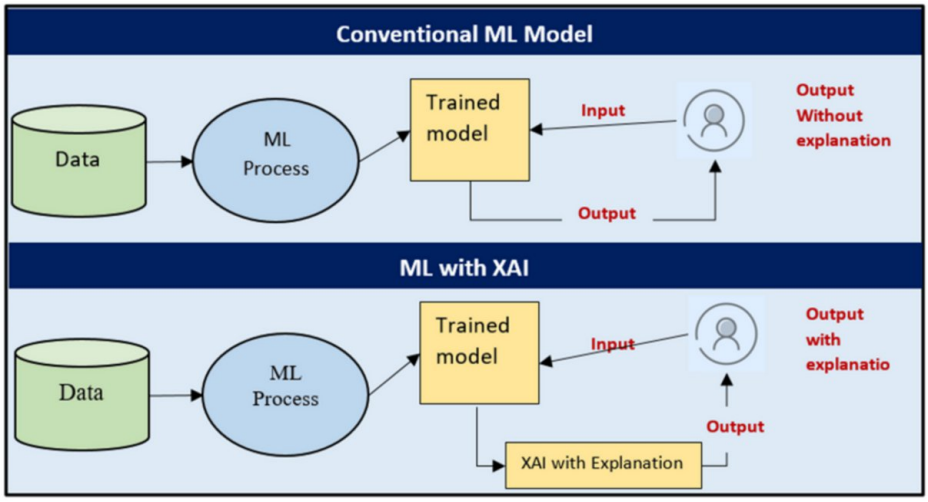

FIGURE 1: Working of ML and ML with XAI

ML, Machine Learning; XAI, Explainable Artificial Intelligence

Major benefits of XAI are improved decision-making, increased trust and acceptability, reduced risk and liability, optimized AI and increased adoption of AI.

XAI works on four principles: Explanation, Meaningful, Explanation Accuracy and Knowledge Limits. The first principle, Explanation, provides information on reasoning of the output [9]. The second principle, Meaningful, gives sufficient understanding to the user on explanation. The third principle, Explanation Accuracy, ensures the accuracy in meaning and explanation of the output produced by XAI. The fourth principle, Knowledge Limit, provides information on the answers that could be judged.

XAI helps in identifying important features used in prediction and ranks them. XAI also provides information on most relevant factors that influence predictions. Lastly, XAI helps in providing graphical and visual representation. XAI is being used in various areas like health sector, financial section, education, legal, e-commerce, entertainment, surveillance and self-driving cars [10]. There are certain challenges and limitations in using XAI for decision-making; for example, sensitive data is revealed. So, its security and privacy become an issue. It also adds to the complexity of AI models [11].

The performance of any application is improving with the use of AI, but alongside performance, the explanation of decisions made using AI is also crucial [12]. The challenge arises when AI provides a prediction but is unable to explain the reasoning behind it. Cartolano et al. [13] use XAI in the area of smart agriculture to infer the behavior of the AI system.

The XAI techniques and methods are broadly classified into two categories: ante-hoc and post-hoc methods. The difference between these two methods is based on explainability factor. If the model is fundamentally explainable, it is known as ante-hoc method. On the other hand, in post-hoc method, explainability is achieved by XAI model after training the data.

There are four types of XAI techniques: Data Explainability, Model Explainability, Feature-Based Techniques and Example-Based Techniques. Data Explainability techniques provide insights into data using various visualization techniques like Exploratory Analysis and Dimension Reduction Techniques.

The purpose of Model Explainability techniques is to provide clear, understandable and interpretable explanation of the model. These techniques are classified based on White Box and Black Box models. Linear model and Decision Tree models are a few examples of White Box models, while Explainable Neural Networks, Support Vector Machines and Tree Ensemble are a few examples of Black Box Models. Feature-based techniques provide information about features that have major contribution in making decision in a model. Techniques like LIME and SHAP are a few examples of feature-based techniques. Example-based techniques provide additional information on accuracy and reliability of model by taking example of an instance in the dataset. Contrastive explanation method and anchor explanation come under this category. Contrastive explanation method gives local explanation for black box model. Anchor explanation is a specific rule or condition that, when satisfied, causes the model to make the same prediction. The current research paper focuses on LIME and SHAP - feature-based techniques of XAI.

The first technique, LIME, explains each individual prediction and provides understanding of any Black Box ML model. Macro Ribeiro developed LIME in 2016. It has two components: Local and Model-agnostic [14]. The first component, Local, provides explanation of some part of an ML model. It gives explanations that are locally faithful within the boundary of data being explained. Local assumes that all models are linear. It tries to fit single observation (data), followed by checking the behaviour of a global model on locally created model [15]. The second component, Model-agnostic, provides insight into the function of a model irrespective of its structure. It gives some understanding or intuition of prediction about an ML model. Model-agnostics methods are produced by post-hoc interpretation techniques on local level to explain individual data, at the regional level to describe subgroups and at global level to explain global space.

The second technique, SHAP, is used to understand and debug the model and is a powerful package. SHAP gives various values that are used to get explanation on prediction. This technique provides data on contribution of each feature of the model in prediction [16-17]. The model may be Linear Regression, Decision Tree, Random Forests or Neural Networks [18]. SHAP also helps to know the aggregate trend on multiple predictions. The SHAP values have several properties such as additivity, local accuracy and missingness. Its additive property finds the contribution of each feature and then sums up the value to know the overall contribution in prediction. The Local Accuracy property finds the difference between the actual model output and the expected output for a given input. This leads to the accuracy of the model. The Missingness property i.e., Zero SHAP value, deals with missing features or irrelevant features [19]. Both the Feature-based techniques, LIME and SHAP, have been used in next sections with the help of a case study to identify the major contributing factors influencing the predictions regarding good or bad applicants for credit card.

Credit card defaulter: A case study

To explain the concept of XAI, a case study of Credit Card Record is used. Whenever an applicant applies for a credit card, a credit score is calculated by financial institutes before issuing credit cards. Credit score is a common risk control method that uses personal information and data of applicants. Many financial institutes use credit scores to decide loan disbursement or issue of credit cards. Credit Information Bureau (India) Limited maintains credit history of customers in India and is widely used in credit scoring model, while Fair Isaac Corporation is used globally in credit scoring mode.

The dataset is available at Kaggle [20]. Based on the data, it predicts the probability of an applicant being defaulter on credit card borrowings. Credit score is based on historical data. The data of credit records is used to develop an ML model for the current research work. It further employs SHAP and LIME techniques to explain the predictions for more transparency in decision-making. It consists of two data files, application_record.csv and credit_record.csv, having 21 columns. The names of columns/features are given in Table *1*, which are self-explanatory.

Data File: Application_record.csv			
Id	amt_income_total	flag_mobil	Id
code_gender	name_income_type	flag_work_phone	months_balance
flag_own_car	name_education_type	flag_phone	Status
flag_own_realty	name_housing_type	flag_email	
cnt_children	days_birth	occupation_type	
name_family_status	days_employed	cnt_fam_members	

TABLE 1: Columns of Dataset

The major features, of the datasets, whose values need understanding, are given below.

1) The column months_balance in the credit_record.csv is the starting point and is relative. Its values start from 0 and go in negative. A value of 0 represents current month, -1 indicates the previous month, and so on.

2) The status column in the credit_record.csv has different values like 0, 1, 2, 3, 4, 5, C and X. A value of 0 means 1-29 days overdue, 1 indicates 30-59 days overdue, 2 signifies 60-89 days overdue, 3 represents 90-119 days overdue, 4 means 120-149 days overdue and 5 signifies overdue or bad debts including write-offs for more than 150 days. A value 'C' means the loan was paid off for that month and 'X' means there was no loan for the month.

3) The column days_birth of application_record.csv counts backward, i.e., 0 for current day, -1 for yesterday and so on.

4) The column days_employed of application_record.csv is the start date of employment. It counts backwards from current day (0). If the value is positive, it means the person is currently unemployed; if negative, it means, the person has been employed for some time.

Methodology adopted

To explain the concept of XAI, ML models such as Logistic Regression, Support Vector Machine, Decision Tree and Random Forest are used in the case study of Credit Card Defaulter. This study follows a structured approach to identify the key features that influence the predictions of these ML models using SHAP and LIME techniques. The methodology adopted consists of the following phases to identify the key features:

1) Preprocessing of Data

2) Preparing CSV File

3) Building ML Models

4) Explanation of Prediction Using LIME and SHAP

Preprocessing of Data

The data is preprocessed to clean it from various aspects, such as handling missing values, taking numeric values for the features, etc. SQL Server is used for preprocessing, although tools like spreadsheets and Python can also be used.

Step 1: This step loads the data into SQL Server database. A database named 'creditcards' is created. Two tables, application_record and credit_record, are created. The downloaded data in CSV form is loaded into these respective tables. There are 438,557 records in table application_record and 1,048,575 records in credit_records. It has been found that there are a total of 36,456 records that are common in both the tables based on ID field. Therefore, these records are extracted and saved in another table named application_record_processed.

Step 2: This step creates additional columns. The following additional columns are created in the table

application_record_processed for better values of features:

1) Age: It is used to store the age of the customer and is calculated using the days_birth column.

2) Service_length: It is used to store the length of employment of the applicant and is calculated using days_employed.

3) Income_range: It is used to store income group.

4) Final_status: It is used to store the label whether the credit card holder is a good or bad applicant.

Step 3: The values of additional columns are evaluated and stored as per the formula given below:

1) The age is calculated by applying the formula absolute(days_birth)/365.25. The calculated age is rounded and stored in the column age.

2) The value of service_length is created using the formula absolute(days_employed)/365.25.

3) The income_range is set to income group as per the following assumed criteria based on the minimum and maximum values of amt_income_total.

a) If the value of amt_income_total is less than 5,00,000, it is 1.

b) If the value of amt_income_total is between 5,00,000 and 10,00,000, it is 2.

c) If the value of amt_income_total is between 10,00,000 and 15,00,000, it is 3.

d) In all remaining cases, it is set to 4.

Step 4: This step converts non-numeric data to numeric. The data of some of the important features of the dataset is in non-numeric form. For the purpose of training data, it is desired to convert the non-numeric data to numeric data. The numeric data of features required for training data is shown in Table *2*.

Code_gender		**flag_own_car**		**flag_own_reality**	
Numeric Code	Value	Numeric Code	Value	Numeric Code	Value
1	Male	1	Yes	1	Yes
2	Female	0	No	0	No
name_qualification_type		net_income_type		name_family_status	
Numeric Code	Value	Numeric Code	Value	Numeric Code	Value
1	Academic Degree	1	Commercial Associate	1	Civil Marriage
2	Higher Education	2	Pensioner	2	Married
3	Incomplete Higher Education	3	State Servant	3	Separated
4	Lower Secondary	4	Student	4	Single
5	Secondary/Special Secondary	5	Other	5	Widow
name_housing_type					
Numeric Code	Value	Numeric Code	Value	Numeric Code	Value
1	Co-op apartment	3	Municipal Apartment	5	With parents
2	House/apartment	4	Office Apartment	6	Rented Accommodation

TABLE 2: Conversion of Text Data to Numeric

The values of the feature occupation_type are also converted to numeric from text data type. There are 18 unique values, and they are assigned numeric values from 1 to 18. Since there are 18 different

occupations, whole number random values between 1 and 18 are used to replace blank occupation type. Other techniques such as distribution-based interpolation can also be used.

Step 5: This step handles null values. In the dataset, it is found that 31.05% of the values in the occupation_type feature are null. These null values cannot be assigned mean, minimum or maximum values. Therefore, the null values are replaced with random values generated between 1 and 18. All other features contain non-null values.

Step 6: This step labels the data. The dataset does not contain any label, so a label customer_type is created, and its value is set to either 1 or 0. An assumption is made that an installment made within 60 days is considered acceptable. For all other cases, it is considered bad debt. Based on this assumption, every record is labeled.

Preparing CSV File

After applying the preprocessing steps, the data is saved in a CSV format.

Building ML Model

The CSV file is loaded into memory, and ML models were developed. Prior to model development, outliers and imbalanced data were handled. Eighty percent of the data was used for training and 20% for testing. A summary of the models is provided in Table *3*. The accuracies of the trained and test data are shown in Figures *2* and 3, respectively. Based on accuracy, Decision Tree and Random Forest models were used to explain the predictions.

Model	Train_Accuracy	Test_Accuracy
Logistics Regression	0.499581	0.501677
Support Vector Machine	0.522836	0.517612
Decision Tree	0.993271	0.976593
Random Forest	0.993271	0.985819
XGBClassifier	0.965003	0.959363
AdaBoostClassifier	0.752621	0.756252

TABLE 3: Summary of Models' Accuracy

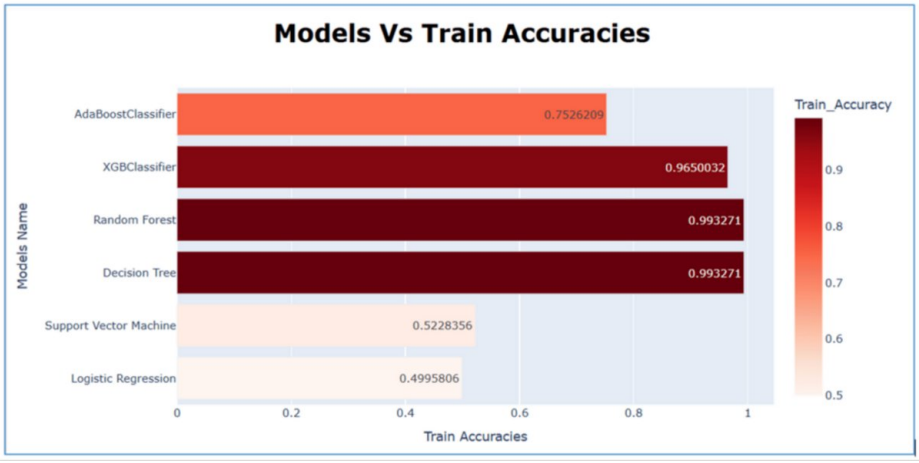

FIGURE 2: Accuracy of Trained Data in Different Models

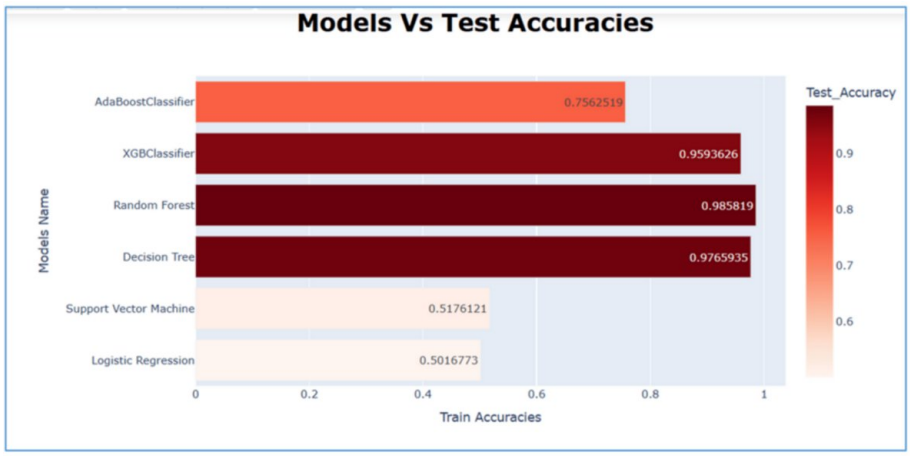

FIGURE 3: Accuracy of Test Data in Different Models

Results

A machine with an i7 processor and 16 GB RAM is used for the experimental setup. Preprocessing on the dataset is done using SQL Server. The preprocessed data is then exported to a CSV file. Jupyter Notebook, the interactive environment for developing Python applications, and the libraries Pandas, NumPy, Matplotlib, and Seaborn are used for the implementation of ML models. The same setup is further used for the explanation of predictions using SHAP and LIME techniques.

Explanation of prediction using LIME

LIME is used to explain the top features' contribution in arriving at a decision based on the input values supplied to the model. In this section, first, two sets of data are identified from the dataset, and then the same data is used to explain the prediction.

Values Features for Explanation

Two scenarios are used to identify the top features that led to decision. The values of each of features in both the scenarios are randomly selected, as shown in Table *4*.

Feature	Scenario I	Scenario II
code_gender	2	2
flag_own_car	0	0
flag_own_realty	1	1
amt_income_total	2,25,000	1,35,000
name_education_type	2	5
name_family_status	1	2
name_housing_type	2	2
cnt_fam_members	2	2
age	36	47
occupation_type	5	3
service_length	2	16

TABLE 4: Values of Features for Explanation

Explanation Using Classification

The Random Forest technique was applied first on both the scenarios. Classification mode is used to explain the prediction. The Prediction probability and Local explanation of Scenario 1 are depicted in Figures *4* and 5. It shows 87% probability of good application and 13% probability of bad application for giving credit. Similarly, in case of Scenario II, the Prediction probability and Local explanation are shown in Figures *6* and 7.

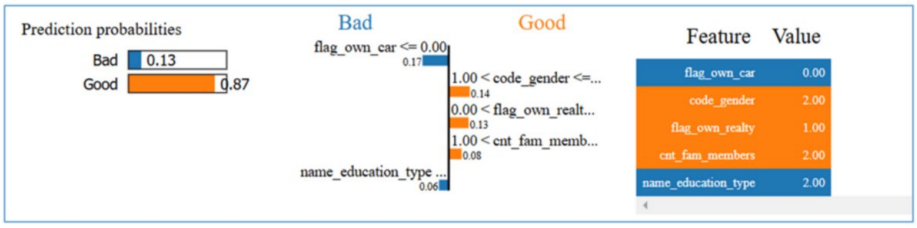

FIGURE 4: Prediction Probabilities of Scenario I

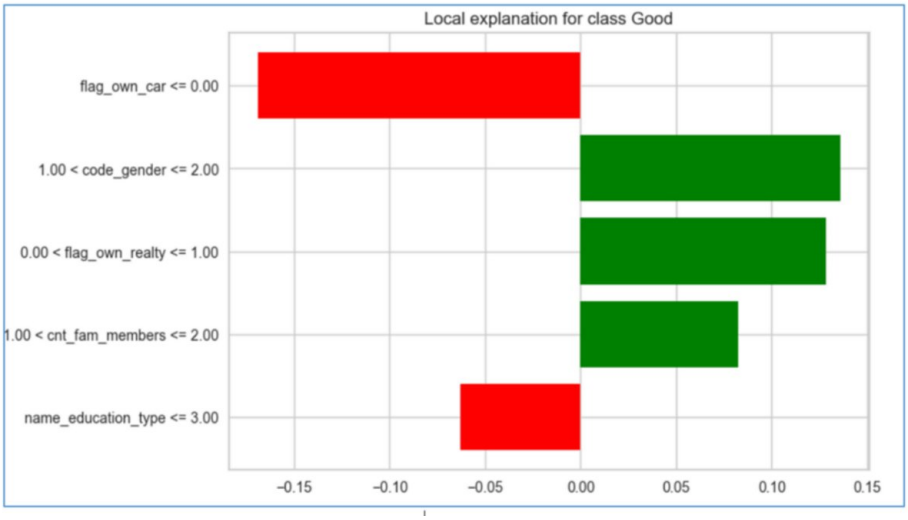

FIGURE 5: Local Explanation of Scenario I

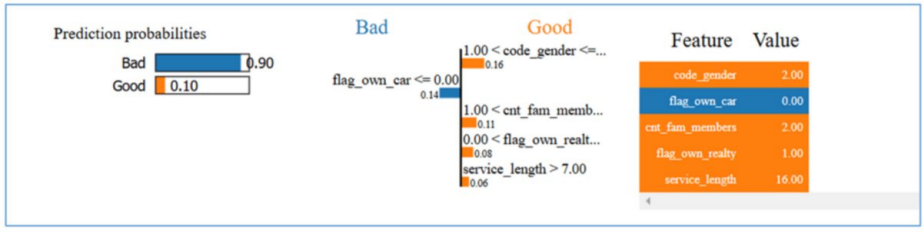

FIGURE 6: Prediction Probabilities of Scenario II

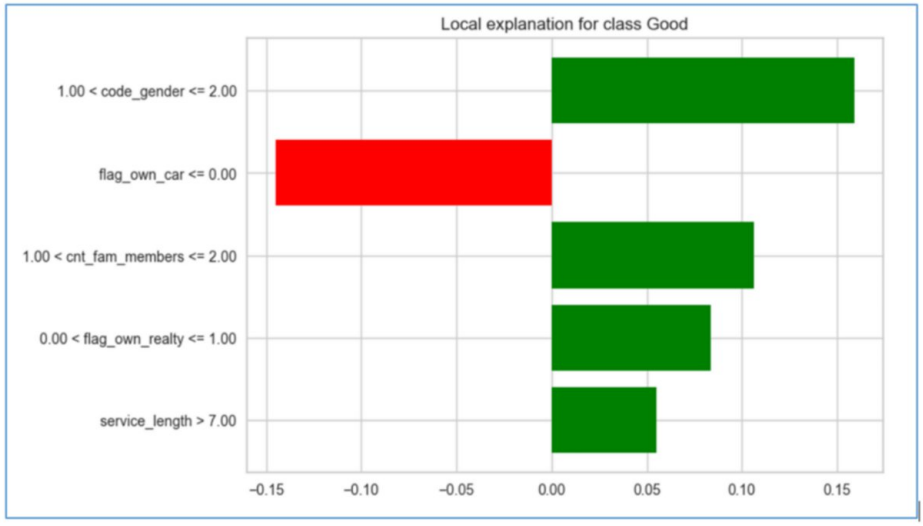

FIGURE 7: Local Explanation of Scenario II

After Random Forest model, the second developed model of Decision Tree was also used for explanation using LIME technique. Table *5* shows the summary of results of both the techniques.

Scenario/Model	Random Forest		Decision Tree	
	Bad (in Percentage)	Good (in Percentage)	Bad	Good
Scenario I	13	87	0	100
Scenario II	90	10	90	10

TABLE 5: Summary of Results

Explanation of prediction using SHAP

SHAP technique is used in both the scenarios and models, namely Random Forest and Decision Tree. First, a summary of Random Forest model is explained, which explains the contribution of each feature in the model, as given in Figure *8*. In the second stage, the contribution of each feature in both the scenarios is found, as shown in Figures *9* and 10. Waterfall plot is used to explain the contribution of features of a single instance. The red-colored SHAP values push value towards prediction of good applicants while blue-colored SHAP values push towards bad applicants. The SHAP values are computed based on Game Theoretic Approach that calculates contribution of each player in overall result. Positive values impact positively while negative values impact negatively on decision. The absolute values show the impact. The final prediction f(x) is the sum of base value plus some of all the SHAP values. Base value is average value, and it is same for all instances.

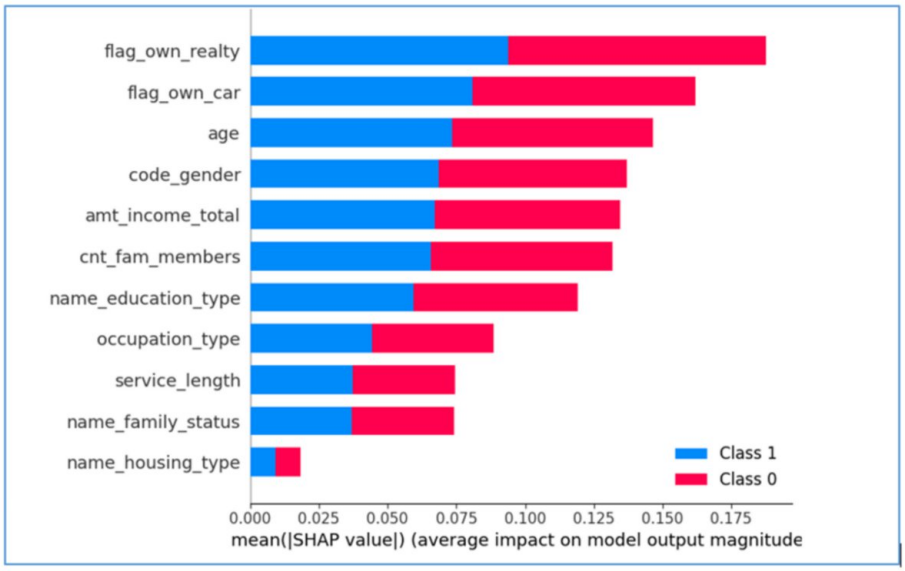

FIGURE 8: Summary of Contribution of All the Features on Model's Output (Random Forest)

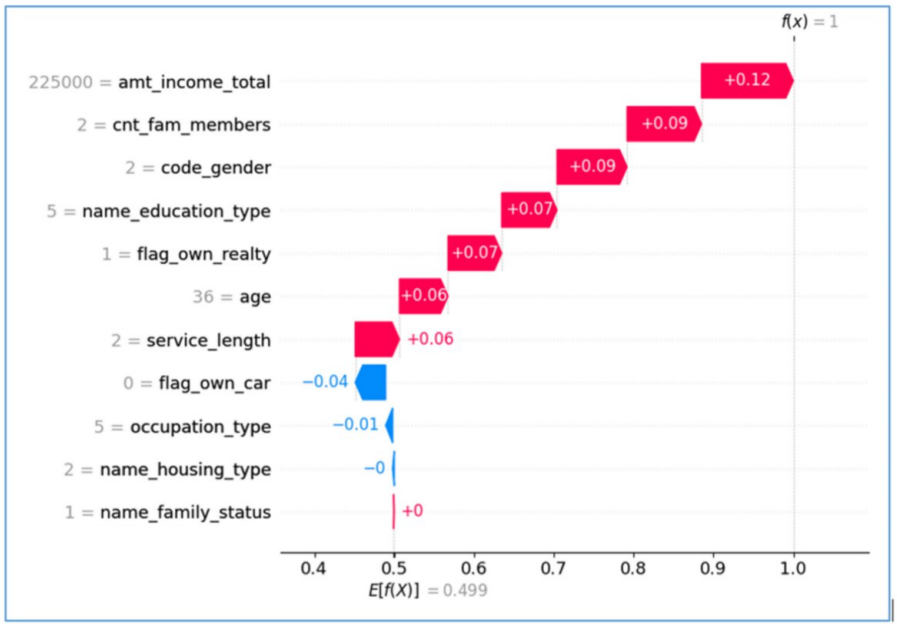

FIGURE 9: Contribution of Each Feature in Random Forest Model (Scenario I)

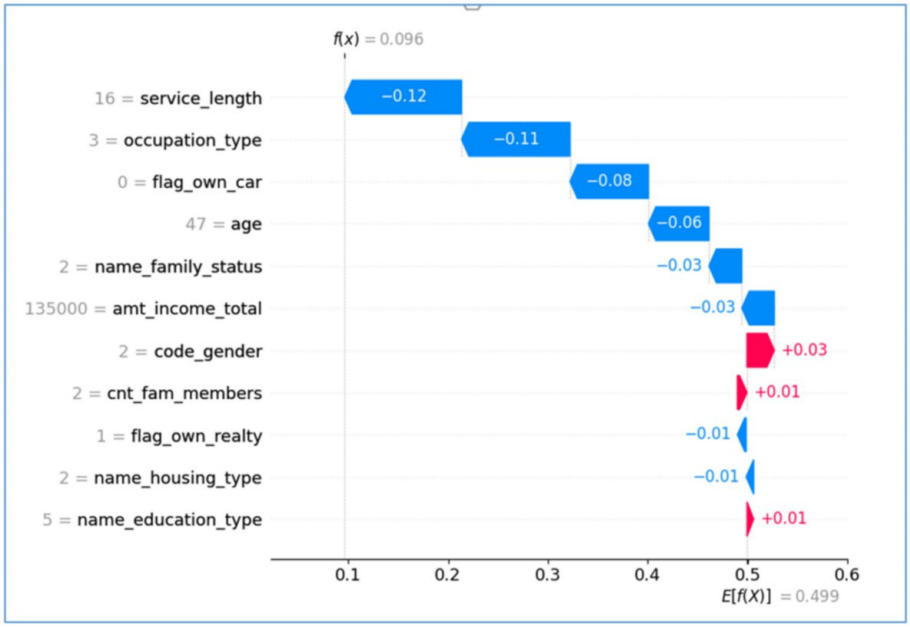

FIGURE 10: Contribution of Each Feature in Random Forest Model (Scenario II)

Similarly, SHAP is applied on Decision Tree model for both the scenarios. The summary of contribution of all the features, contribution of features in prediction of Scenario I and contribution of features in prediction of scenario II is given in Figures *11*, *12* and 13, respectively.

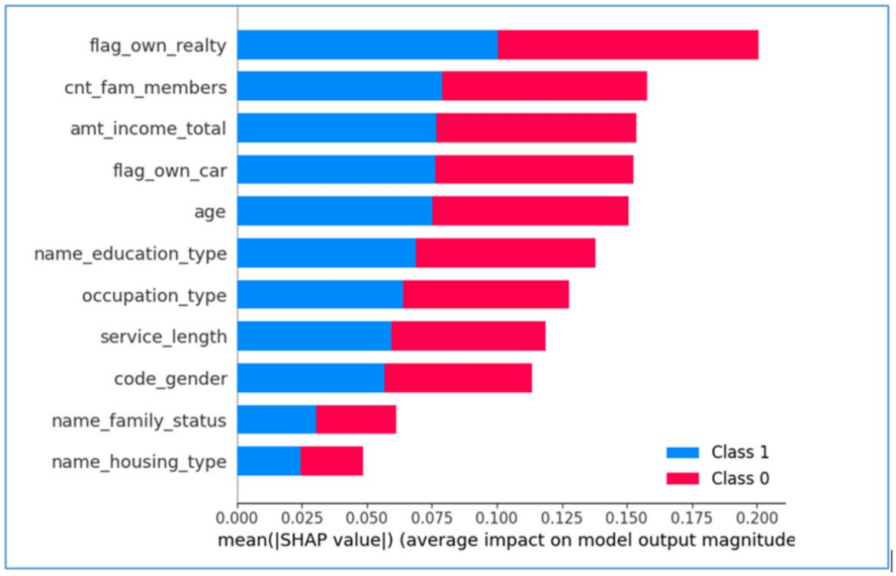

FIGURE 11: Summary of Contribution of All the Features on Model's Output (Decision Tree)

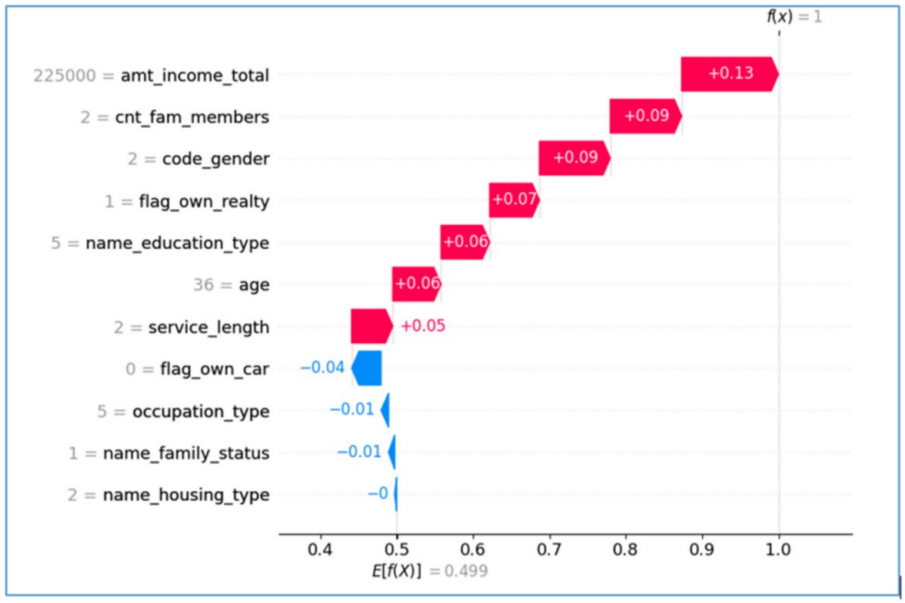

FIGURE 12: Contribution of Each Feature in Decision Tree Model (Scenario I)

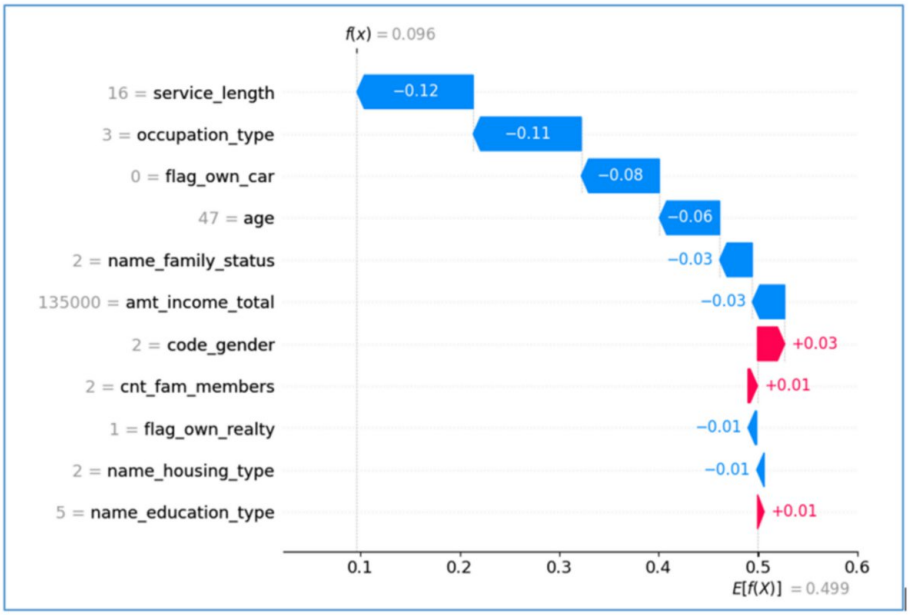

FIGURE 13: Contribution of Each Feature in Decision Tree Model (Scenario II)

Discussion

Table *6* shows that contribution of features in prediction varies from model to model and technique to technique. While analyzing Table *6*, it is found that there are eight different cases in which contribution of different features is shown. It has been observed that the features code_gender, flag_own_car, flag_own_reality and cnt_fam_members contribute the most while predicting the bad or good credit applicants.

The contribution of each feature in prediction is different for Random Forest and Decision Tree models when LIME and SHAP XAI techniques are used. For example, in Scenario I, the feature flag_own_car has topped in contribution in prediction when Random Forest model with LIME is used, and the feature code_gender contributed the most when the Model Decision Tree with LIME is used. The proposed

techniques are found helpful in understanding the role of features in decision-making. It is clear from the Table 6 that the feature flag_own_car has the highest contribution in both the scenarios when Random Forest is used with LIME, while in Decision Table models with LIME, the feature code_gender topped in contribution.

The techniques used in explanation of prediction enhance the trust on the model and identify the major contributing factors. Hence, XAI tools such as SHAP and LIME help in making AI models transparent, interpretable and trustworthy.

Features	LIME (Classification Mode)				SHAP (Waterfall Mode)			
	Scenario I		Scenario II		Scenario I		Scenario II	
	Random Forest	Decision Tree	Random Forest	Decision Tree	Random Forest	Decision Tree	Random Forest	Decision Tree
code_gender	2	1	2	1	3	3	-	-
flag_own_car	1	2	1	2	-	-	4	3
flag_own_realty	3	4	3	4	5	4	-	-
amt_income_total	-	-	-	-	1	1	5	-
name_education_type	5	-	-	-	4	5	-	-
name_family_status	-	5	5	-	-	-	-	5
name_housing_type	-	-	-	-	-	-	-	-
cnt_fam_members	4	3	4	3	2	2	-	-
age	-	-	-	5	-	-	3	4
customer_type	-	-	-	-	-	-	-	-
occupation_type	-	-	-	-	-	-	1	2
service_length	-	-	-	-	-	-	2	1

TABLE 6: Comparison of LIME and SHAP in Identifying Top Features Used in Prediction

LIME, Local Interpretable Model-Agnostics Explanation; SHAP, SHapely Additive exPlanation

Conclusions

The paper focuses on feature-based XAI techniques, specifically LIME and SHAP, which provide valuable insights into the working of ML models and the way they make predictions, enhancing their transparency and trust. LIME focuses on local explanations, providing insights into the way a specific model prediction was made for a particular input, while SHAP offers both local and global explanations, explaining individual predictions and the overall model behaviour. It uses a case study on credit card default prediction to demonstrate how these methods identify key features that influence model predictions. The paper also compares the interpretability results of LIME and SHAP in understanding model behaviour and indicates the use of these models in improving model transparency, interpretability and trust by identifying key contributing factors. Future works include the usability of SHAP and LIME in AI models other than Random Forest and Decision Tree.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Indu Arora

Acquisition, analysis, or interpretation of data: Indu Arora

Drafting of the manuscript: Indu Arora

Critical review of the manuscript for important intellectual content: Indu Arora

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

This work would not have been possible without the support and guidance of Dr. Manish Arora, Scientist ‘F’, NIELIT Deemed to be University. The encouragement and insightful feedback were instrumental in accomplishing this task.

References

1. Ali S, Abuhmed T, El-Sappagh S, et al.: Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Information Fusion. 2023, 99:1566-2535. 10.1016/j.inffus.2023.101805
2. Sarker IH: AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. SN Computer Science. 2022, 3:158. 10.1007/s42979-022-01043-x
3. Kühl N, Schemmer M, Goutier M, Satzger G: Artificial intelligence and machine learning. Electronic Markets. 2022, 32:2235-2244. 10.1007/s12525-022-00598-0
4. Subramanian R, Moar RR, Singh S: White-box machine learning approaches to identify governing equations for overall dynamics of manufacturing systems: A case study on distillation column. Machine Learning with Applications. 2021, 3:100014. 10.1016/j.mlwa.2020.100014
5. Retzlaff CO, Angerschmid A, Saranti A, Schneeberger D, Röttger R, Müller H, Holzinger A: Post-hoc vs ante-hoc explanations: XAI design guidelines for data scientists. Cognitive Systems Research. 2024, 86:101243. 10.1016/j.cogsys.2024.101243
6. Dwivedi R, Dave D, Naik H, et al.: Explainable AI (XAI): Core ideas, techniques, and solutions. ACM Computing Surveys. 2023, 55:1-33. 10.1145/3561048
7. Ishikawa S, Todo M, Taki M, et al.: Example-based explainable AI and its application for remote sensing image classification. International Journal of Applied Earth Observation and Geoinformation. 2023, 118:103215. 10.1016/j.jag.2023.103215
8. Zodage P, Harianawala H, Shaikh H, Kharodia A: Explainable AI (XAI): History, basic ideas and methods. International Journal of Advanced Research in Science Communication and Technology. 2024, 560-568. 10.48175/ijarsct-16988
9. Phillips PJ, Hahn CA, Fontana PC, Yates AN, Greene K, Broniatowski DA, Przybocki MA: Four Principles of Explainable Artificial Intelligence. National Institute of Standards and Technology, Gaithersburg, MD; 2021. 10.6028/NIST.IR.8312
10. Srinivasu PN, Sandhya N, Jhaveri RH, Raut R: From blackbox to explainable AI in healthcare: existing tools and case studies. Mobile Information Systems. 2022, 2022:8167821. 10.1155/2022/8167821
11. Explainable AI: The Challenges and Limitation of XAI, E-Spin. (2025). https://www.e-spincorp.com/explainable-ai-the-challenges-and-limitation-of-xai/.
12. Kalasampath K, Spoorthi KN, Sajeev S, Kuppa SS, Ajay K, Maruthamuthu A: A literature review on applications of explainable artificial intelligence (XAI). IEEE Access. 2025, 13:41111-41140. 10.1109/access.2025.3546681
13. Cartolano A, Cuzzocrea A, Pilato G: Analyzing and assessing explainable AI models for smart agriculture environments. Multimedia Tools and Applications. 2024, 83:37225-37246. 10.1007/s11042-023-17978-z
14. Gaspar D, Silva P, Silva C: Explainable AI for intrusion detection systems: LIME and SHAP applicability on multi-layer perceptron. IEEE Access. 2024, 12:30164-30175. 10.1109/ACCESS.2024.3368377
15. Rao S, Mehta S, Kulkarni S, Dalvi H, Katre N, Narvekar M: A study of LIME and SHAP model explainers for autonomous disease predictions. 2022 IEEE Bombay Section Signature Conference (IBSSC), Mumbai, India. 2022, 1-6. 10.1109/IBSSC56953.2022.10037324
16. Salih AM, Raisi-Estabragh Z, Boscolo Galazzo I, Radeva P, Petersen SE, Lekadir K, Menegaz G: A perspective on explainable artificial intelligence methods: SHAP and LIME. Advanced Intelligent Systems. 2025, 7:2400304. 10.1002/aisy.202400304
17. Nohara Y, Matsumoto K, Soejima H, Nakashima N: Explanation of machine learning models using shapley additive explanation and application for real data in hospital. Computer Methods and Programs in Biomedicine. 2022, 214:106584. 10.1016/j.cmpb.2021.106584
18. Zhang S, Li X: A comparative study of machine learning regression models for predicting construction duration. Journal of Asian Architecture and Building Engineering. 2023, 23:1980-1996. 10.1080/13467581.2023.2278887
19. Matthews S, Hartman B: mSHAP: SHAP values for two-part models. Risks. 2022, 10:3. 10.3390/risks10010003
20. Credit Card Approval Prediction: A Credit Card Dataset for Machine Learning. https://www.kaggle.com/datasets/rikdifos/credit-card-approval-prediction.