The Scientific Temper (2024) Vol. 15 (2): 2207-2215
Doi: 10.58414/SCIENTIFICTEMPER.2024.15.2.33

E-ISSN: 2231-6396, ISSN: 0976-8653
https://scientifictemper.com/

RESEARCH ARTICLE

A framework for generating explanations of machine learning models in Fintech industry

Krishna P. Kalyanathaya[1*], Krishna Prasad K.[2]

Abstract

Artificial Intelligence is making significant inroads into various aspects of business and life bringing the transformation in many ways. The convolution of technology in finance is often called FINTECH rapidly growing area of transformation. In the FINTECH industry, AI can automate several financial processes and services such as fraud detection, customer services, credit assessment, price predictions, customer churning, trading services, risk management, underwriting, market forecasting. These processes and services are critical to financial sectors such as banking, insurance, currency, stock and commodity markets, wealth management, payment clearing houses, payment regulators etc. Regulations control these processes and should be transparent in their operations. AI models are inherently opaque in their outcomes and unable to be fully plugged into the financial processes and services. Explainable AI is the key area of research that can help to provide transparency to enable these AI models as fully operational business models to automate financial products and services. In this paper we will broadly outline the framework of explainable artificial intelligence (XAI) in finance sectors and services. We then look into one use case of credit assessment and develop an XAI framework to provide transparent outcomes from the AI models.

Keywords: Artificial intelligence, Machine learning, Explainable AI, Credit assessment, Fintech, Explainer dashboard, XAI.

Introduction

The finance domain is spread across a very broad field of applications ranging from banking, insurance, investments, commerce, manufacturing, and government, covering almost every part of our life and business. Fintech, short for "financial technology," is an application of technology to provide innovative financial products and services. It refers to a wide range of technological advancements that are designed to disrupt and enhance traditional financial services. Fintech is transforming the financial services landscape by challenging traditional models and creating more accessible, convenient, and cost-effective solutions for both consumers and businesses. It can reshape how financial transactions are conducted, how investments are managed, and the relationship of financial institutions with people. This can range from managing financial transactions to fighting fraud and applications include asset management, risk analysis, analytics, regulatory compliance, virtual assistance, and others. Traditional banking has started using disruptive fintech companies and their innovative solutions to reduce costs and address consumer pain points (Weber *et al.*, 2023).

The widespread adoption of artificial intelligence (AI) models in the finance sectors and services has highlighted challenges in their lack of transparency and interpretability. In order to address these concerns and facilitate the trustworthiness of AI-driven decision-making systems, this research presents a comprehensive framework for generating explanations of AI models specifically tailored for the financial industry.

The proposed framework builds upon existing techniques for model interpretability and explanation generation while tailoring them to the unique requirements of the fintech industry. It encompasses several key components, including input data preprocessing, model training, explanation generation, and model query engine.

The results of this research have the potential to

[1]College of Computer Science and Information Science, Srinivas University, Mangalore, India.

[2]Department of Cyber Security and Cyber Forensic, Institute of Engineering and Technology, Srinivas University, Mukka, Mangalore, India.

***Corresponding Author:** Krishna P. Kalyanathaya, College of Computer Science and Information Science, Srinivas University, Mangalore, India., E-Mail: krishna.prakash.kk@gmail.com

How to cite this article: Kalyanathaya, K. P., Prasad, K. K. (2024). A framework for generating explanations of machine learning models in Fintech industry. The Scientific Temper, **15**(2):2207-2215.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.2.33

Source of support: Nil

Conflict of interest: None.

Received: 05/04/2024 **Accepted:** 10/05/2024 **Published** : 15/06/2024

significantly benefit the fintech industry by enabling the adoption of AI models with enhanced interpretability and explainability. The framework's ability to generate transparent explanations will aid in identifying biases, improving model performance, complying with regulations, and fostering customer trust (Babaei *et al*., 2023).

In conclusion, this research addresses the critical need for generating explanations of AI models in the fintech industry through the evaluation of a comprehensive framework. By promoting transparency and interpretability, this framework aims to enhance trust, regulatory compliance, and decision-making processes within Fintech organizations.

Role of XAI in Fintech Industry

In our previous work (Kalyanathaya and Krishna Prasad, 2022), we have provided our definition of XAI as follows:

"Explainable AI is a technique of explaining how a Machine learning model performs the actions and makes predictions. It is aimed at explaining the rationale of the decision-making process (compare this with a judicial system that pronounces a judgment after thoroughly evaluating all the evidence submitted to the system). However, In the case of machine learning models, the evidence submitted to the model is evaluated by a complex algorithm pronouncing the outcome as a prediction. In this case, the ML model does not provide all the corroboration of evaluated evidence, and hence complex algorithm is a black-box (or opaque model)."

XAI, or eXplainable AI is reshaping the Fintech industry by making AI-driven decisions transparent, interpretable, and compliant. As the Fintech landscape continues to evolve, XAI stands as a crucial component in driving the industry toward a future where advanced AI and human understanding go hand in hand (Kalyanathaya and Krishna Prasad, 2022).

One of the primary challenges in adopting AI within the Fintech sector has been the "black-box" nature of many machine learning models. These models, while incredibly powerful at making predictions and decisions, often lack transparency. This lack of transparency poses significant issues when it comes to compliance, risk management, and customer trust.

This is where XAI steps in. By providing interpretable explanations for AI-driven decisions, XAI enables financial institutions to understand, validate, and communicate the rationale behind each decision. This is crucial for meeting regulatory requirements such as GDPR and for building trust with customers who demand transparency in how their financial data is being used.

In credit scoring, for example, XAI algorithms can show which variables and data points influenced a credit decision. This transparency ensures that discriminatory factors are eliminated, promoting fairness in lending. Similarly, in algorithmic trading, XAI can offer insights into why specific trades were executed, which is essential for compliance and risk management (Demajo *et al*., 2021).

Moreover, XAI enhances risk assessment and fraud detection. Traditional rule-based systems often struggle to keep up with the evolving tactics of fraudsters. AI-powered models, when paired with XAI, can provide insights into the patterns and features that led to the identification of a particular transaction as fraudulent. This knowledge empowers Fintech companies to adapt and stay ahead of new threats (De Lange *et al*., 2022).

While XAI offers a plethora of benefits, it's important to note that implementing XAI in Fintech isn't without challenges. Balancing the need for transparency with proprietary information protection, ensuring the accuracy of explanations, and managing the added complexity of XAI systems are some hurdles that companies must navigate.

Materials and Methods

As artificial intelligence (AI) models increasingly permeate the fintech industry, there is a growing need for interpretability and explainability to ensure transparency and trust in decision-making processes. This literature review explores the existing body of research and techniques related to the generation of explanations for AI models in the fintech industry (Demajo *et al*., 2021).

Explainability Techniques in the Fintech Industry

Within the fintech industry, the demand for explainability is driven by regulatory compliance, risk management, and customer trust. Research has focused on developing explainability techniques for specific fintech applications such as credit scoring, fraud detection, and algorithmic trading. These studies have explored the use of surrogate models, rule extraction algorithms, and visualizations to generate explanations that align with regulatory requirements, enhance model transparency, and facilitate human-understandable insights into AI-based financial decision-making (Gramespacher and Posth, 2021).

Domain-specific Challenges and Considerations

The complex nature of financial data, privacy concerns, regulatory constraints, and the need for interpretability in high-stakes decision-making necessitate tailored approaches. Researchers have emphasized the importance of incorporating domain-specific knowledge, financial regulations, and industry best practices into explainability frameworks. Adapting existing methods to account for financial risk assessment, fairness, robustness, and ethical considerations has been a key focus of recent studies.

Research Gap

Despite the increasing adoption of AI models in the fintech industry and the growing need for transparency and explainability, there exists a research gap in the development of a comprehensive framework specifically tailored to generating explanations of AI models in the fintech domain. While there have been studies on interpretability and explainability in AI models and some research on explainability techniques in specific fintech applications, there is a lack of a unified framework that addresses the unique challenges and requirements of the fintech industry (Weber *et al*., 2023).

The existing literature primarily focuses on general techniques for interpretability and explanation generation, often without considering the specific nuances of the fintech domain. These studies typically overlook important factors such as financial regulations, domain-specific knowledge, and the need to align with regulatory guidelines. Therefore, there is a need to bridge this gap by developing a framework that not only incorporates these factors but also provides accurate and compliant explanations for AI models in the fintech industry.

As shown in Table 1, the current research gap lies in the absence of a comprehensive framework for generating explanations of AI models in the fintech industry. This gap encompasses the need for a tailored framework that incorporates financial regulations and domain-specific knowledge and aligns with regulatory guidelines and evaluation of explanations in fintech applications, addressing ethical considerations, fairness, and bias.

These methods offer transparency by visualizing the decision-making process and feature importance.

Results and Discussions

The goal of this assessment is to demonstrate the type of explanations on the following research agenda (Kalyanathaya and Krishna Prasad, 2022):

- Build interpretability outside the model algorithm to explain the results
- Generate data-driven interpretations with a generic model to explain the results

Credit Assessment Model

Credit assessment is the process of evaluating an individual's or a company's credit worthiness to determine their ability to repay debt obligations. It is a crucial step in the lending and borrowing process as it helps financial institutions, such as banks, credit unions, or lending agencies, make informed decisions about extending credit to borrowers (Demajo and Dingli, 2021).

When a loan application is processed by a bank using AI/ML model:

- If the loan application is rejected, the customer (loan applicant) wants to know the reason for the rejection.

Table 1: Summary of XAI methods

XAI method	*Description*	*Applications*	*Types of explanations*
White box***(Weber *et al*., 2023)	Makes use of the inherent property of machine learning algorithm. No separate tool used here	Works with conventional ML models such as Decision tree and linear regression etc	Linear Regression, Decision Tree
LIME(Weber *et al*., 2023)	Local Interpretable Model-Agnostic Explanations (LIME) approximate the decision boundary around a specific instance and identify the most influential features for that instance.	Works with text or tabular data (numerical or categorical) or images. Works for any models developed using packages sci-kit learn, keras, Pytorch.	Feature importance. Generates local explanations, providing insights into individual predictions
SHAP(Weber *et al*., 2023)	Shapley additive explanations (SHAP) employs cooperative game theory to assign feature importance scores to individual features.	Works with Tree-based models, deep learning models	Feature Importance. Generates global and local interpretability to gain insights into the overall behavior and examination of individual predictions or instances.
Anchor(Ribeiro and Guestrin, 2018)	Anchor is a tool that generates if-then rules as explanations for AI model predictions. It focuses on identifying simple and understandable rules that hold for a particular prediction.	Similar to LIME. Works with any models developed using packages sci-kit learn, keras, Pytorch.	If-then-rules. Explanations of most influential feature values that contribute to a particular prediction while holding other features constant.
DeepLift(Shrikumar and Kundaj, 2017, July)	DeepLIFT is a tool specifically designed for deep learning models. It attributes the contribution of each input feature to the model's output by comparing the activations of the model with and without the feature.	Works with various deep learning models built with Neural Networks	Feature contribution. Explanations at the neuron level helping to understand the behavior of individual neurons in the deep network.
Explainer Dashboard(Oege Dijk, 2023)	Explainer dashboard is package used to build interactive dashboards for analyzing and explaining the predictions and workings of (scikit-learn compatible) machine learning models.	Works with scikit-learn, xgboost, and skorch (sklearn wrapper for tabular PyTorch models) and others.	Feature Importance, Feature Contribution, What-If Analysis

*** Makes use of the inherent property of machine learning algorithm. No separate tool is used here.

1	Loan ID	Customer	Loan Status	Current Lo	Term	Credit Sco	Annual Inc	Years in cu	Home Ownership	Purpose	Monthly D	Years of C	Months si	Number o	Number o	Current Cr	Maximum	Bankruptc	Tax Liens
2	14dd8831-	981165ec-	Fully Paid	445412	Short Term	709	1167493	8 years	Home Mortgage	Home Imp	5214.74	17.2	NA	6	1	228190	416746	1	0
3	4771cc26-	2de017a3-	Fully Paid	262328	Short Term			10+ years	Home Mortgage	Debt Cons	33295.98	21.1	8	35	0	229976	850784	0	0
4	4eed4e6a	5efb2b2b-	Fully Paid	99999999	Short Term	741	2231892	8 years	Own Home	Debt Cons	29200.53	14.9	29	18	1	297996	750090	0	0
5	77598f7b-	e777faab-	Fully Paid	347666	Long Term	721	806949	3 years	Own Home	Debt Cons	8741.9	12	NA	9	0	256329	386958	0	0
6	d4062e70-	81536ad9-	Fully Paid	176220	Short Term			5 years	Rent	Debt Cons	20639.7	6.1	NA	15	0	253460	427174	0	0
7	89d8cb0c-	4ffe99d3-	Charged Off	206602	Short Term	7290	896857	10+ years	Home Mortgage	Debt Cons	16367.74	17.3	NA	6	0	215308	272448	0	0
8	273581de-	90a75dde-	Fully Paid	217646	Short Term	730	1184194	< 1 year	Home Mortgage	Debt Cons	10855.08	19.6	10	13	1	122170	272052	1	0
9	db0dc6e1-	018973c9-	Charged Off	648714	Long Term			< 1 year	Home Mortgage	Buy House	14806.13	8.2	8	15	0	193306	864204	0	0
10	8af915d9-	af534dea-	Fully Paid	548746	Short Term	678	2559110	2 years	Rent	Debt Cons	18660.28	22.6	33	4	0	437171	555038	0	0
11	0b1c4e3d-	235c4a43-	Fully Paid	215952	Short Term	739	1454735	< 1 year	Rent	Debt Cons	39277.75	13.9	NA	20	0	669560	1021460	0	0
12	32c2e48f-	0de7bcdb	Fully Paid	99999999	Short Term	728	714628	3 years	Rent	Debt Cons	11851.06	16	76	16	0	203965	289784	0	0
13	fa096848-	aa0a6a22-	Fully Paid	541970	Short Term			10+ years	Home Mortgage	Home Imp	23568.55	23.2	NA	23	0	60705	1634468	0	0
14	403d7235-	11581f68-	Fully Paid	99999999	Short Term	740	776188	< 1 year	Own Home	Debt Cons	11578.22	8.5	25	6	0	134083	220220	0	0
15	01d878ae-	900c9191-	Fully Paid	99999999	Short Term	743	1560907	4 years	Rent	Debt Cons	17560.37	13.3	NA	10	1	225549	496474	1	0
16	2e841c8f-	2ac05980-	Fully Paid	234124	Short Term	727	693234	10+ years	Rent	Debt Cons	14211.24	24.7	46	10	1	28291	107052	1	0
17	7cbaa3fa-	3ec886e7-	Fully Paid	449020	Long Term			9 years	Own Home	Debt Cons	18904.81	19.4	NA	8	0	334533	428956	0	0
18	c9a16a9d-	abb4c446-	Charged Off	653004	Long Term			7 years	Home Mortgage	Debt Cons	14537.09	20.5	NA	9	0	302309	413754	0	0
19	24e8c8bd-	967e8733-	Fully Paid	666204	Long Term	723	1821967	10+ years	Home Mortgage	Debt Cons	17612.24	22	34	15	0	813694	2004618	0	0
20	c6be21f0-	c67b2cb5-	Fully Paid	66396	Short Term			10+ years	Rent	Debt Cons	9898.81	27.1	NA	23	1	9728	402380	1	0

Figure 1: Preview of credit assessment data

- If the loan application is rejected, what can the customer do to revise the credit assessment?
- The bank wants to ascertain whether the loan is approved or rejected correctly or not.
- The bank regulator wants to ascertain whether the loan applications are processed in full transparency or if there are no violations of regulations applicable to the loan approval process.

Currently, AI/ML models are unable to clearly provide evidences that business users or regulator can understand. A starting point to address these challenges is to provide insights into which features or variables the model considers most important in making decisions. Techniques like feature importance score from decision trees and permutation feature importance by using explainability tools can help in understanding feature contributions (Udaya Bhanu and Narayana, 2021).

In this assessment, we have used historical data(Zaur, 2017) with the following information shown Table 2.

Preview of the credit assessment data is shown in Figure 1.

Explainer Dashboard Tool

The explainer dashboard tool (Oege Dijk, 2023) serves as a crucial tool in demystifying AI/ML models used in credit assessment. By providing transparent insights, explanations, and interactive features, the dashboard empowers business users and regulators to make informed decisions while enhancing trust in the model's outcomes.

The explainer dashboard is designed to provide transparent and understandable insights into the decisions made by our AI/ML models in credit assessment. This dashboard aims to bridge the gap between technical model outputs and the needs of business users and regulators by offering clear explanations for each prediction.

This toolkit offers a convenient solution for swiftly deploying a web app dashboard that provides comprehensive explanations for a fitted machine-learning model compatible with sci-kit-learn. The interactive dashboard showcases a range of plots and insights, including model performance evaluations, feature importance, contributions of features to individual predictions, partial dependence plots, SHAP (interaction) values, and visualizations of individual decision trees.

Table 2: Dataset attributes and descriptions

No	*Column name*	*Description*
1	Loan ID	Unique ID for each loan application
2	Customer ID	Unique ID for each customer
3	Loan status	Status of the loan. Fully paid or charged off. This column determines the loan eligibility as approved or rejected.
4	Current loan amount	Outstanding loan amount
5	Term	Loan term, small term or long term.
6	Credit score	Numerical representation of an individual's credit risk
7	Annual income	Annual income of the loan applicant
8	Years in current job	No of years in the current job
9	Home ownership	Staying in own house or rented or mortgaged house
10	Purpose	Purpose of the loan such business loan, education loan or house loan or medical emergency etc.
11	Monthly debt	Monthly payment to debt.
12	Years of credit history	No of years of credit history
13	Months since the last delinquent	Months of past dues if any
14	Number of open accounts	Number of open accounts
15	Number of credit problems	Number of credit problems
16	Current credit balance	Current credit balance
17	Maximum open credit	The highest amount of credit available to the customer.
18	Bankruptcies	Bankruptcies YES or NO
19	Tax Liens	Any legal claim against the assets of the candidate

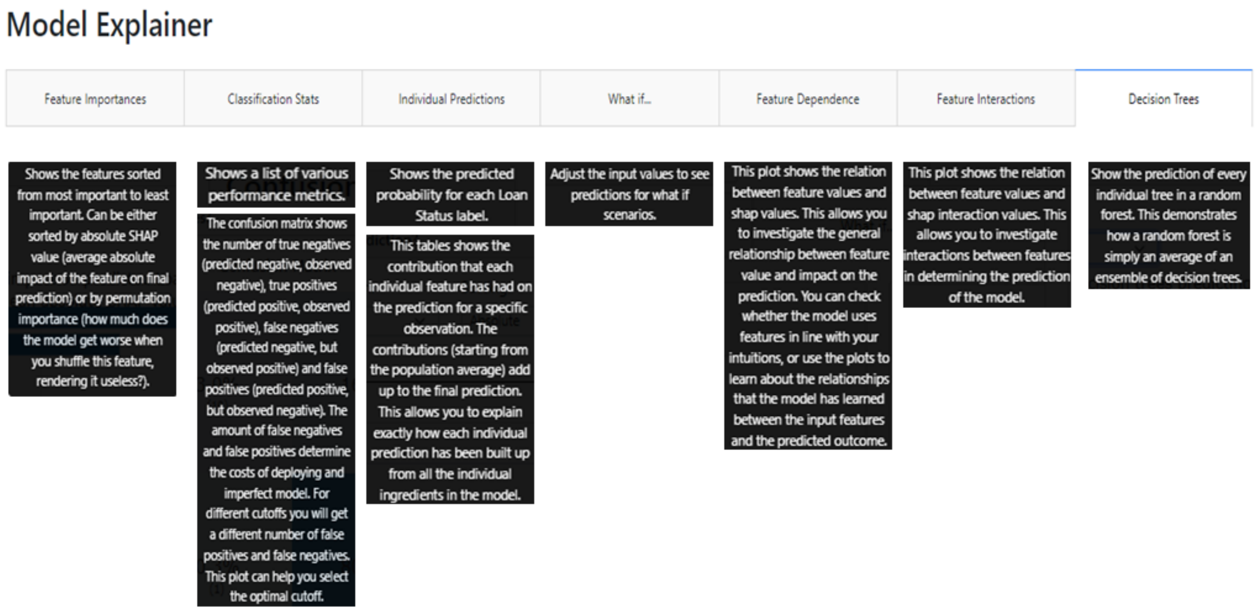

Figure 2 : Key features of explainer dashboard

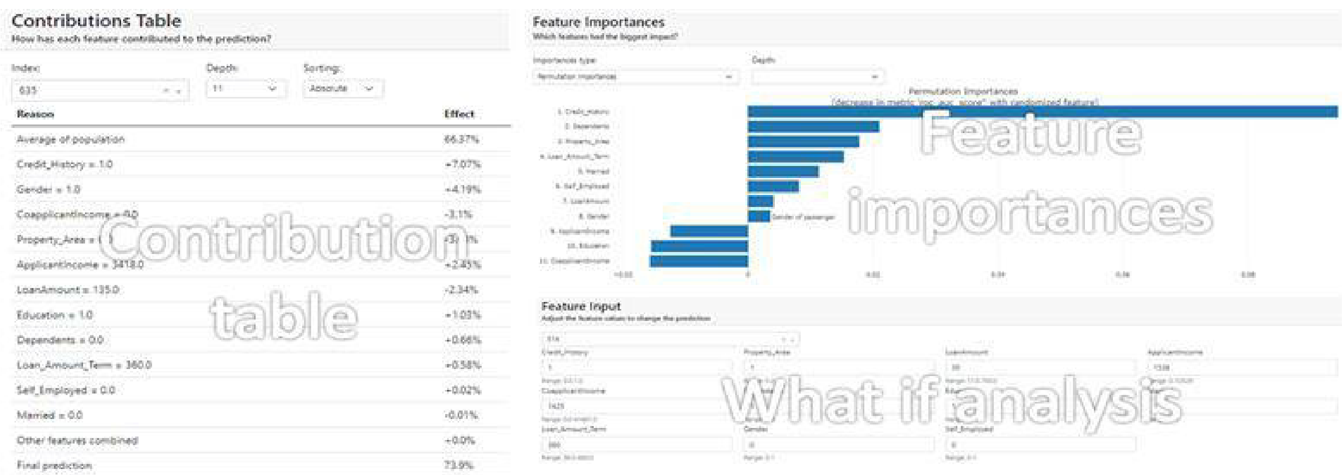

Figure 3: Various explainer dashboard visualization

Model Explainer Key Features

This package not only facilitates model understanding for data scientists but also bridges the gap for various stakeholders to engage with and comprehend model outcomes effectively. The model explainer visualization and key features of the explainer dashboard are shown in Figures 2 and 3.

Evaluation of Explainer Dashboard Tool for Credit Assessment

We have set up our experiments on XAI framework on Jupyter notebook to run the explainer dashboard. Our experiments used a random forest algorithm to build the model and evaluate the outcome with the explainer dashboard tool.

The random forest ML model with explainer dashboard launched with the following code in Figure 4.

Types of Explanations

After starting explainer dashboard, the URL can used to launch the dashboard in a web browser. The explainer dashboard will show the model explainer to provide the explanations on the credit assessment model.

Here, we will look into some of the prominent explanations that are suitable for answering credit assessment-related queries.

Feature Importance

Features are basically the information provided to the credit assessment model to determine whether credit can

```
In [16]: # import packages
         from sklearn.model_selection import train_test_split
         from sklearn.ensemble import RandomForestClassifier
         from explainerdashboard import ClassifierExplainer, ExplainerDashboard

In [17]: # credit assessment model
         X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33, random_state=42)
         model = RandomForestClassifier(n_estimators=50, max_depth=10).fit(X_train, y_train)

In [*]:  explainer = ClassifierExplainer(model, X_test, y_test, descriptions=feature_descriptions, labels=['Charged OFF', 'Fully Paid'])
         ExplainerDashboard(explainer).run()

         Detected RandomForestClassifier model: Changing class type to RandomForestClassifierExplainer...
         Note: model_output=='probability', so assuming that raw shap output of RandomForestClassifier is in probability space...
         Generating self.shap_explainer = shap.TreeExplainer(model)
         Building ExplainerDashboard..
         Detected notebook environment, consider setting mode='external', mode='inline' or mode='jupyterlab' to keep the notebook intera
         ctive while the dashboard is running...
         Warning: calculating shap interaction values can be slow! Pass shap_interaction=False to remove interactions tab.
         Generating layout...
         Calculating shap values...
         Calculating prediction probabilities...
         Calculating metrics...
         Calculating confusion matrices...
         Calculating classification_dfs...
         Calculating roc auc curves...
         Calculating pr auc curves...
         Calculating liftcurve_dfs...
         Calculating shap interaction values... (this may take a while)
         Reminder: TreeShap computational complexity is O(TLD^2), where T is the number of trees, L is the maximum number of leaves in a
         ny tree and D the maximal depth of any tree. So reducing these will speed up the calculation.
         Calculating dependencies...
         Calculating permutation importances (if slow, try setting n_jobs parameter)...
         Calculating pred_percentiles...
         Calculating predictions...
         Calculating ShadowDecTree for each individual decision tree...
         Reminder: you can store the explainer (including calculated dependencies) with explainer.dump('explainer.joblib') and reload wi
         th e.g. ClassifierExplainer.from_file('explainer.joblib')
         Registering callbacks...
         Starting ExplainerDashboard on http://[redacted]:8050
```

Figure 4: Launch explainer dashboard for random forest model

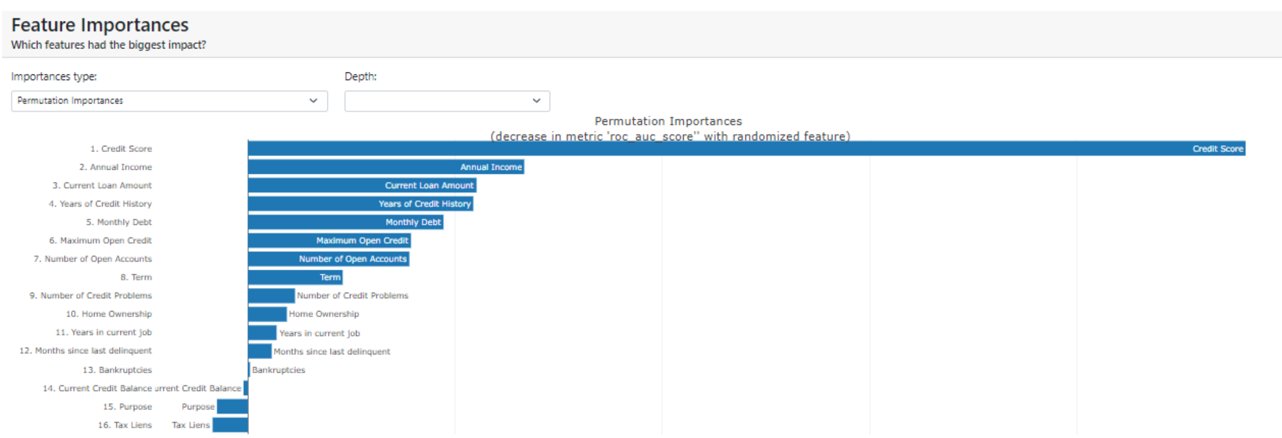

Figure 5A: Bar chart shows the importance of each feature in the credit assessment model

be approved or not. This information generally comes from customers or authorities (like banks or regulatory entities) when credit assessment is required.

The feature importance screen will provide the graphical visualization of features as in Figure 5A shown below. The bar chart shows each feature and its importance or value in determining credit assessment. Higher length of the bar, the more important feature.

We can quickly visualize the following features had a positive impact on assessing the credit approval of the customer. Here we list only the top 5 features influencing credit assessment.

- Credit score
- Annual income
- Current loan amount
- Years of credit history

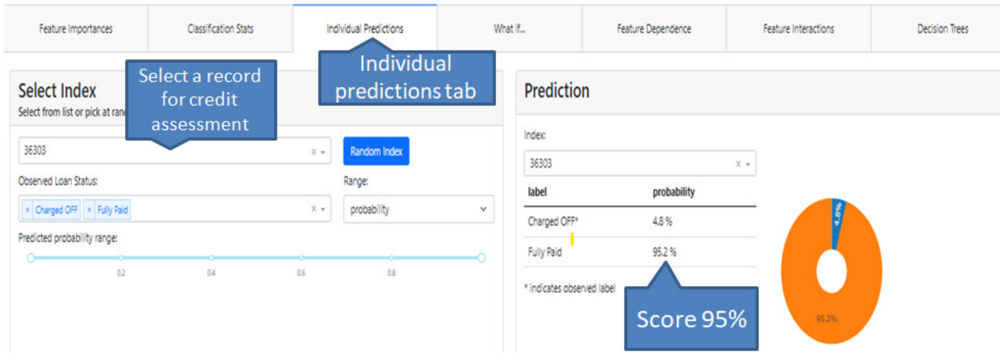

Figure 5B: Select a record for individual credit assessment

Contributions Table

How has each feature contributed to the prediction?

Index: 36303 — Selected record for credit assessment — Depth: 14 — Sorting: Importance

Reason	Effect
Average of population	78.12%
Credit Score = 736.0	+3.55%
Annual Income = 1626153.0	+3.03%
Current Loan Amount = 310134.0	-0.56%
Years of Credit History = 17.9	+1.24%
Term = 1.0	+1.14%
Current Credit Balance = 323076.0	+2.07%
Years in current job = 3.0	+0.79%
Monthly Debt = 19107.16	+2.55%
Maximum Open Credit = 515966.0	+1.07%
Number of Open Accounts = 8.0	+0.64%
Months since last delinquent = 51.0	+1.01%
Purpose = 3.0	+0.46%
Home Ownership = 2.0	+0.12%
Bankruptcies = 0.0	-0.1%
Other features combined	+0.13%
Final prediction	95.25%

Figure 6: Contribution table for the selected record of individual credit assessment

- Monthly debt

Similarly we can also visualize some features that had a negative impact on the credit approval. Here we list 3 features identified in the model.

- Current credit balance
- Purpose
- Tax liens

The feature importance will show the general influence of features in the credit assessment model. We can choose the individual predictions tab to determine predictions and the influence of features on each individual credit assessment. Here we select a particular record from credit assessment data and visualize the predictions and contribution of each feature for the prediction shown in Figure 5B.

Contribution of each feature in the prediction can be visualized as follows: From the contribution table shown in Figure 6, we can see that the features' credit score' and 'annual income' have a positive impact of collectively more than 6%. Similarly, there are some features with a negative impact on the predictions.

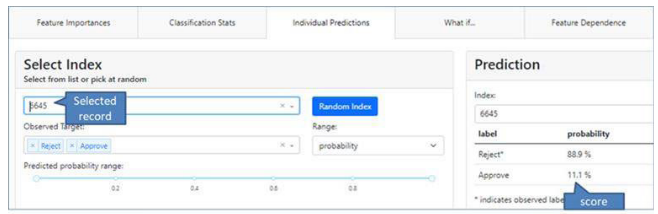

Figure 7: Individual prediction of credit assessment for the record with index 6645

What-If Analysis

Here, we will analyze explanations to credit assessment when the model rejects credit. We will select a record with a predicted credit approval score is very low. Let us select a record with index 6645, which has the individual prediction as shown below: Figure 7, the prediction of approval class probability is 11.1%. So, the credit request will be rejected. Now, we can look into the contribution of each feature to predict a low probability of approval of the credit as follows: From the feature contribution table shown in Figure 8, we can see that credit score value was negatively contributed by -64% due to abnormally high credit score.

We will further analyze what steps customers/authorities can take to improve the score so that credit can be approved. This is done through what-if analysis as follows: Now let us analyze what can customer/authority do to improve the score?

From the 'what-if' scenario screens in Figure 9, we can visualize the prediction and feature Input values. The credit score is abnormally high (7300) and, possibly due to wrong data entry during the application process. The normal value of good credit score will be in the range of 600 to 800 (may vary depending on the financial regulators of each country). The feature inputs can be modified and checked in a 'what-if' scenario as shown in Figure 10. After the credit score value is modified (new value 730), we can see prediction score has increased to 90.5%.

Table 3: Challenges in explainable AI

Description of challenge	*Approach the resolution*	*Type of explanations used by the explainer dashboard*
If the loan application is rejected, How can customer (loan applicant) know the reason for the rejection?	Use the contribution table in Individual predictions to identify top 5 features responsible for the approval score. Importance of the features on the score is indicated in percentage of effect. The features with negative values indicate the reason for the low approval score.	Feature Importance
If the loan application is rejected, what can customer do to revise the credit assessment?	Analyze the identified features to determine the correctness and possibility improving the values. We can apply improved values features and revalidate the credit assessment to revise prediction score.	What-if analysis

Hence, we can revalidate the loan application by making corrections (If it was due to wrong data entry).

Figure 8: Contribution of top 5 features in credit assessment for record 6645

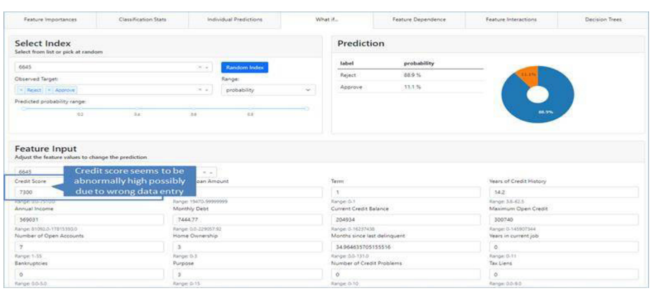

Figure 9: What-If scenario visualization for the record with index 6645

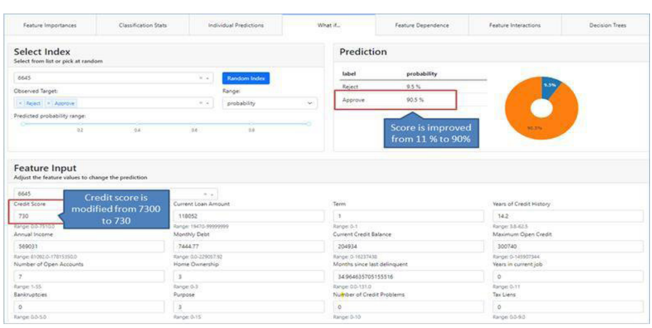

Figure 10: Demonstration of a 'what-if' scenario to improve the score from 11.1 to 90.5%

Conclusion

In this paper, we have identified four challenges of machine learning-based automated credit assessment and attempted to address two challenges using explainer dashboard as shown in Table 3.

The two challenges addressed here customer-oriented which will largely help to understand and revalidate the credit assessments.

The explainer dashboard tool does a great job in explaining the decisions predicted by the machine learning model. The various stakeholders can access the user-friendly web pages and visualize the various inputs that goes into the assessment of the machine learning model. It also provides the visualization of the contribution of each feature and allows the user to revalidate the inputs using what-if scenarios to ascertain the improved score. However, we need to further work on addressing the following challenges:

When a loan application is processed by bank using AI/ML model:

- How can the bank validate the loan is approved or rejected correctly or not?
- How can bank regulator ascertain the loan applications are processed in full transparency or there is no violations of regulations applicable to the loan approval process?

Acknowledgment

Nil

References

Babaei, G., Giudici, P., & Raffinetti, E. (2023). Explainable fintech lending. *Journal of Economics and Business*, 125-126. https://doi.org/10.1016/j.jeconbus.2023.106126

De Lange, P. E., Melsom, B., Vennerød, C. B., & Westgaard, S. (2022). Explainable AI for credit assessment in banks. *Journal of Risk and Financial Management, 15(12)*, 556. https://doi.org/10.3390/jrfm1512055

Demajo, L. M., Vella, V., & Dingli, A. (2021). An explanation framework for interpretable credit scoring. *International*

Journal of Artificial Intelligence and Applications (IJAIA), 12(1).

Gramespacher, T., & Posth, J. A. (2021). Employing explainable AI to optimize the return target function of a loan portfolio. *Frontiers in Artificial Intelligence, 4*, 693022.

Kalyanathaya, Krishna Prakash, & Krishna Prasad, K., (2022). A Literature Review and Research Agenda on Explainable Artificial Intelligence (XAI). *International Journal of Applied Engineering and Management Letters (IJAEML), 6(1)*, 43-59. DOI: https://doi.org/10.5281/zenodo.5998488

Oege Dijk, oegesam, Ray Bell, Lily, Simon-Free, Brandon Serna, rajgupt, yanhong-zhao-ef, Achim Gädke, Anamaria Todor, Evgeniy, Hugo, Mohammad Haizad, Tunay Okumus, & woochan-jang. (2023). *oegedijk/explainerdashboard: explainerdashboard 0.4.2: dtreeviz v2 compatiblity (v0.4.2)*. Zenodo. https://doi.org/10.5281/zenodo.7633294

Ribeiro, M. T., Singh, S., & Guestrin, C. (2018). Anchors: High-Precision Model-Agnostic Explanations. *Proceedings of the AAAI Conference on Artificial Intelligence, 32(1)*. https://doi.org/10.1609/aaai.v32i1.11491

Shrikumar, A., Greenside, P., & Kundaje, A. (2017, July). Learning important features through propagating activation differences. *International conference on machine learning (pp. 3145-3153)*. PMLR.

Udaya Bhanu, L., & Narayana, D. S. (2021). Customer loan prediction using supervised learning technique. *Mathematical Sciences Research Journal. An International Journal of Rapid Publication, 11(6)*, 403-7. https://doi.org/10.29322/ijsrp.11.06.2021.p11453

Weber, P., Carl, K. V., & Hinz, O. (2023). Applications of explainable artificial intelligence in finance—a systematic review of finance, information systems, and computer science literature. *Management Review Quarterly*, 1-41.

Zaur Begiev (2017). Bank Loan Dataset. Version 1. Retrieved January 15, 2024 from https://www.kaggle.com/datasets/zaurbegiev/my-dataset