

**FUTURE OF LOAN APPROVALS WITH EXPLAINABLE AI**A. DURGA DEVI MADAM<sup>1</sup>, KATHULA KARUNA<sup>2</sup>,<sup>1</sup>Assistant professor , MCA DEPT, Dantuluri Narayana Raju College, **Bhimavaram, Andharapadesh****Email:-** adurgadevi760@gmail.com<sup>2</sup>PG Student of MCA, Dantuluri Narayana Raju College, **Bhimavaram, Andharapadesh****Email:-** kkaruna8996@gmail.com**ABSTRACT**

Widespread adoption of automated decision-making by artificial intelligence (AI) is witnessed due to specular advances in computation power and improvements in optimization algorithms especially in machine learning (ML). Complex ML models provide good prediction accuracy; however, the opacity of ML models does not provide sufficient assurance for their adoption in the automation of lending decisions. This paper presents an explainable AI decision- support system to automate the loan underwriting process by belief-rule-base (BRB). This system can accommodate human knowledge and can also learn from historical data by supervised learning. The hierarchical structure of BRB can accommodate factual and heuristic rules. The system can explain the chain of events leading to a decision for a loan application by the importance of an activated rule and the contribution of antecedent attributes in the rule. A business case study on automation of mortgage underwriting is demonstrated to show that the BRB system can provide a good trade-off between accuracy and explainability. The textual explanation produced by the activation of rules could be used as a reason for the denial of a loan. The decision- making process for an application can be comprehended by the significance of rules in providing the decision and contribution of its antecedent attributes.

**1 INTRODUCTION**

Underwriting skill is learnt through several months of training and exchange of knowledge by senior underwriters. This task requires underwriters to be fairly analytical, very organized, and accurate to give informed decision to approve or reject a loan application. Underwriters concurrently analyze a large quantity of information to find affordability, repayment history and collateral. Furthermore, sometimes they are required to change the process due to a shift in regulatory and compliance standards, investor requirements, and customer demands (Krovvidy, 2008).

New technology and strong machine learning (ML) algorithms have opened the doors for a straightthrough loan application process. Artificial intelligence (AI) systems can execute rules and process customers' information in a few milliseconds. Financial institutions have recognized the benefits of AI and are using it in a different subset of the underwriting process and are keen to test and implement newly introduced digital innovation. AI systems are expected to replicate human decision-making skills. However, even today transformation of various algorithmic concepts into training data could be very challenging to solve every instance of the problem for a range of lending products. It may not be able to

solve a tiny subset of the problem (Aggour, Bonissone, Cheetham, & Messmer, 2006).

## 1. Explainable AI (XAI) in Financial Decision Making:

Authors: Miller, Tim. et al.

Summary: This seminal work introduces the concept of Explainable AI and its importance in the financial domain. It highlights the need for interpretable models, especially in applications like loan underwriting, where transparency is critical.

## 2. Interpretable Machine Learning for Credit Scoring: A Case Study on Peer- to- Peer Lending:

Authors: Ribeiro, Marco Tulio. et al.

Summary: The paper discusses the application of interpretable machine learning models in credit scoring, emphasizing the importance of understanding model predictions. The study showcases how interpretability can be achieved without sacrificing predictive performance.

### 3 IMPLEMENTATION STUDY

#### Existing System:

In the current landscape of loan approvals in the financial services industry, traditional methods typically rely on manual review processes and rule-based systems to assess applicants' creditworthiness. These methods often involve subjective assessments by loan officers based on limited information such as credit scores, income levels, and employment history. While these approaches have been effective to some extent, they may suffer from inefficiencies, biases, and lack of scalability. Additionally, the use of manual processes can introduce delays and inconsistencies in decision-making, leading to suboptimal outcomes for both lenders and borrowers. As a result, there is a growing interest in leveraging artificial intelligence (AI) and machine learning (ML) techniques to automate and improve the loan approval process.

#### Disadvantages:

- Limited Scalability
- Inefficiency
- Lack of Transparency

#### Proposed System & alogirtham

The proposed system for loan approvals in the financial services industry aims to overcome the limitations of

existing methods by leveraging advanced artificial intelligence (AI) and machine learning (ML) techniques, while ensuring transparency and fairness through explainable AI (XAI). The system utilizes AI algorithms to analyze a wide range of data sources, including traditional credit bureau information, alternative data sources, and non-traditional data such as social media activity or transaction history..

#### 4.1 Advantages:

Predicting traffic routes offers several advantages that can significantly enhance transportation efficiency and convenience:

**1. Reduced Congestion:** By predicting traffic patterns, authorities can optimize traffic flow, suggesting alternative routes to drivers before congestion builds up. This reduces overall traffic congestion and minimizes delays.

**2. Time Savings:** Efficient route prediction helps drivers choose the fastest routes based on real-time traffic conditions. This saves commuters time and reduces fuel consumption and emissions associated with idling in traffic.

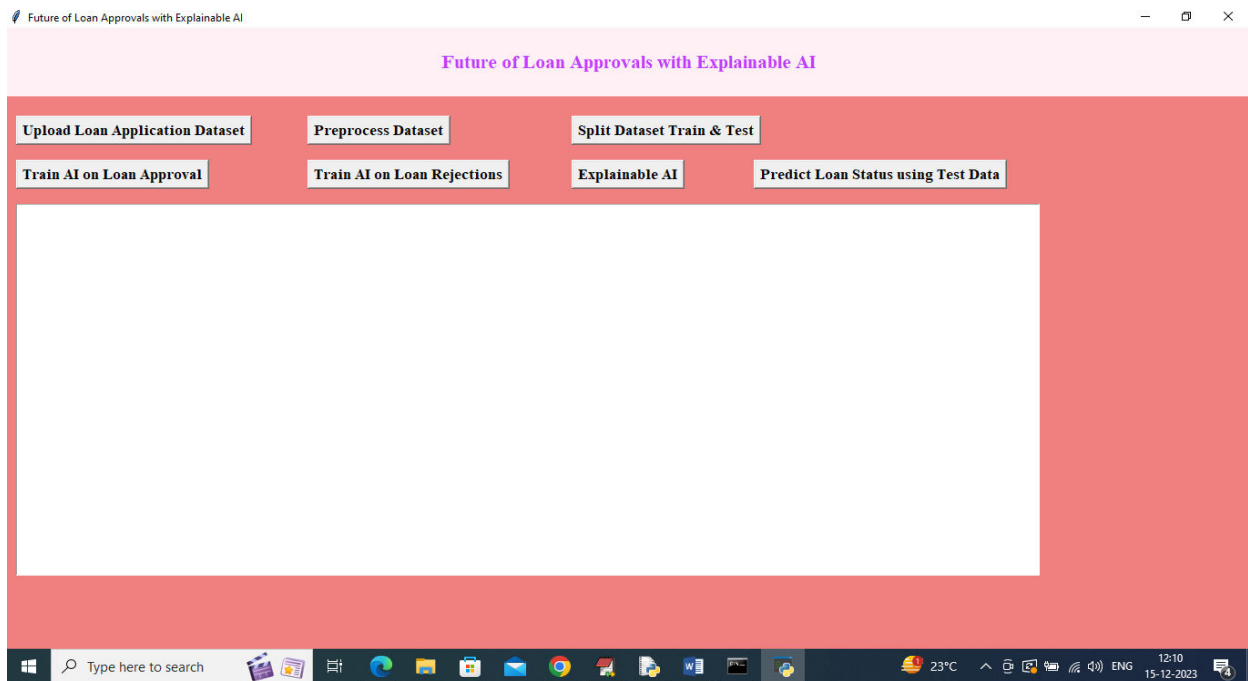
### IMPLEMENTATION

- 1) Upload Loan Application Dataset: using this module we will upload dataset to application and then application will read entire dataset and then find all class labels for loan and reject reason and plot them in a graph
- 2) Pre-process Dataset: dataset contains missing value and both numeric and non-numeric data so by employing label encoder class will convert all data into numeric format and then normalized all dataset values to make it clean.
- 3) Split Dataset Train & Test: using this module will split Dataset in to train and test where application using 80% dataset for training and 20% for testing

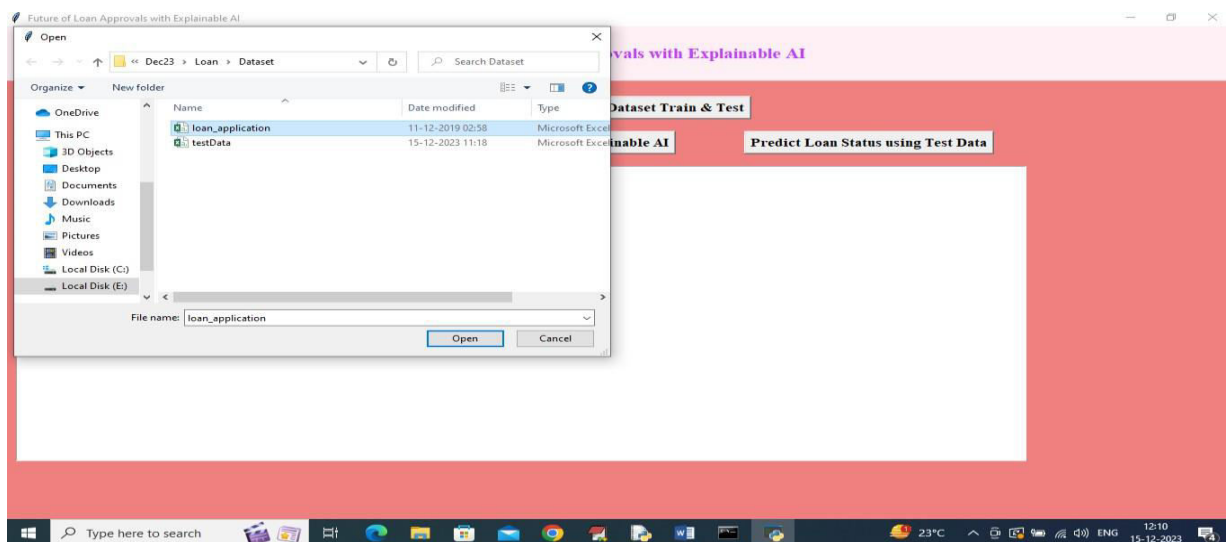
## 5 RESULTS AND DISCUSSION

### 1.1 SCREENSHOTS

To run project double, click on `run.bat` file to get below screen



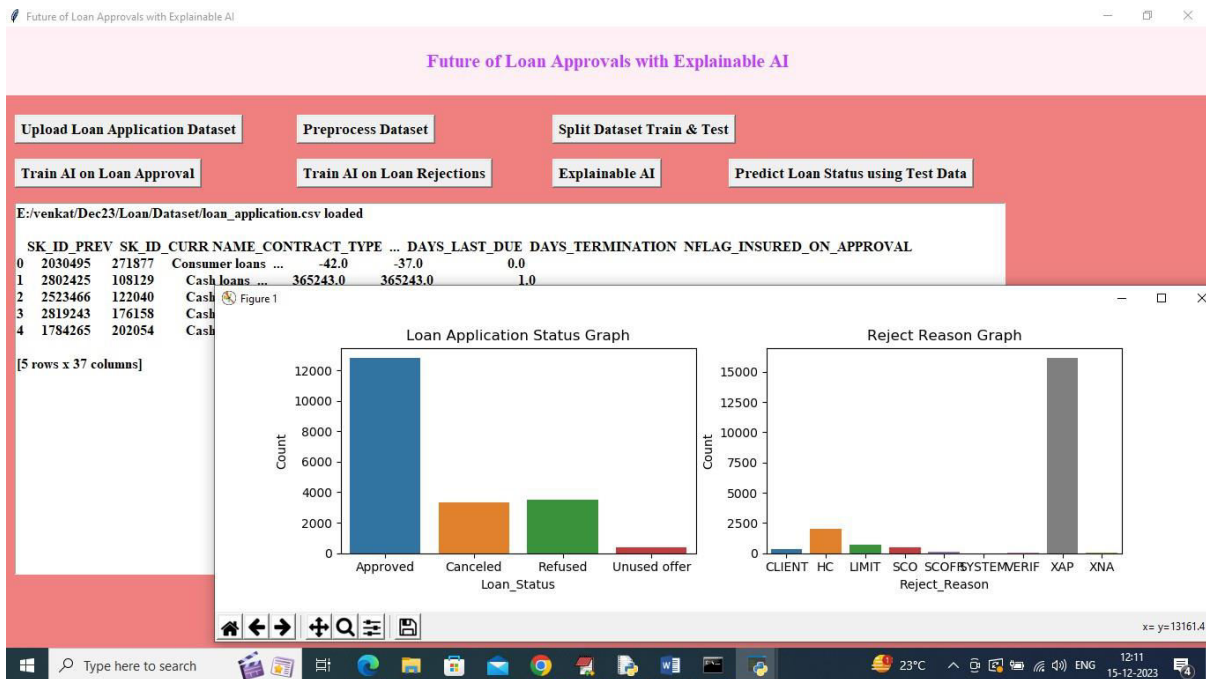
In above screen click on **Upload Loan Application Dataset** button to upload dataset and then will get below output



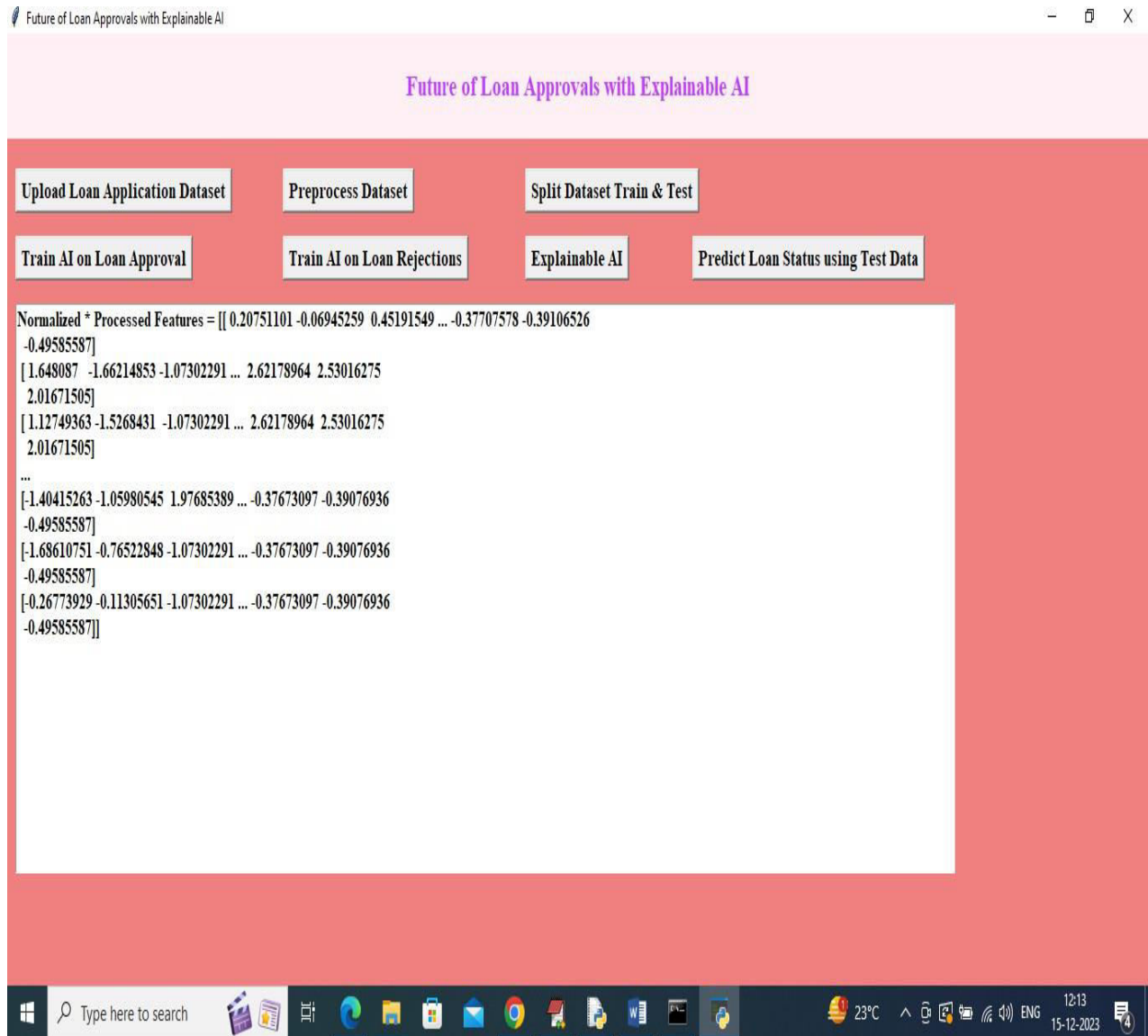
In above screen selecting and uploading **loan\_application.csv** file and then click on **Open** button to load dataset and get below output

SK_ID	PREV_SK_ID	CURR_SK_ID	NAME	CONTRACT	TYPE	DAYS_LAST_DUE	DAYS_TERMINATION	NFLAG	INSURED	ON_APPROVAL	AMT	APP	AMT_CREI	AMT_DOV	AMT_GOC	WEEKDAY	HOUR_AP	FLAG	LAS	NFLAG	LA_RATE	DO_RATE	INTI_RATE	INTI_RATE	NAME_CA	NAME_CC	DAYS_DEC	NAME_PA	CODE_REJ	NAME
2030495	271877	Consumer	loans	...	...	-42.0	-37.0	0.0			17145	17145	0	17145	SATURDA	15	Y			1	0	0.182832	0.867336	XAP	Approved	-73	Cash thro	XAP		
2802425	108129	Cash loans	...	...	...	365243.0	365243.0	1.0			607500	607500	0	607500	THURSDA	11	Y			1				XNA	Approved	-164	XNA	XAP	Unacc	
2523466	122040	Cash	loans	...	...	15060.74	112500	136444.5			112500	112500	0	112500	TUESDAY	11	Y			1				XNA	Approved	-301	Cash thro	XAP	Spous	
2819243	176158	Cash	loans	...	...	47041.34	450000	470790			450000	450000	0	450000	MONDAY	7	Y			1				XNA	Approved	-512	Cash thro	XAP		
1784265	202054	Cash	loans	...	...	31924.4	337500	404055			337500	337500	0	337500	THURSDA	9	Y			1				Repairs	Refused	-781	Cash thro	HC		
1383531	199383	Cash	loans	...	...	23703.93	315000	340573.5			315000	315000	0	315000	SATURDA	8	Y			1				Everyday	Approved	-684	Cash thro	XAP	Famil	
2315218	175704	Cash	loans	...	...	0	0	0			0	0	0	0	TUESDAY	11	Y			1				XNA	Canceled	-14	XNA	XAP		
1656711	296299	Cash	loans	...	...	0	0	0			0	0	0	0	MONDAY	7	Y			1				XNA	Canceled	-21	XNA	XAP		
2367563	342292	Cash	loans	...	...	0	0	0			0	0	0	0	MONDAY	15	Y			1				XNA	Canceled	-386	XNA	XAP		
2579447	334349	Cash	loans	...	...	0	0	0			0	0	0	0	SATURDA	15	Y			1				XNA	Canceled	-57	XNA	XAP		
1715995	447712	Cash	loans	...	...	11368.62	270000	335754			270000	335754	0	270000	FRIDAY	7	Y			1				XNA	Approved	-735	Cash thro	XAP	Unacc	
2257824	161140	Cash	loans	...	...	13832.78	211500	246397.5			211500	246397.5	0	211500	FRIDAY	10	Y			1				XNA	Approved	-815	Cash thro	XAP	Unacc	
2330894	258628	Cash	loans	...	...	12165.21	148500	174361.5			148500	174361.5	0	148500	TUESDAY	15	Y			1				XNA	Approved	-860	Cash thro	XAP	Unacc	
1397919	321676	Consumer	loans	...	...	7654.86	53779.5	57564			53779.5	57564	0	53779.5	SUNDAY	15	Y			1	0	0		XAP	Approved	-408	Cash thro	XAP	Unacc	
2273188	270658	Consumer	loans	...	...	9644.22	26550	27252			26550	27252	0	26550	SATURDA	10	Y			1	0	0		XAP	Approved	-726	Cash thro	XAP	Unacc	
1232483	151612	Consumer	loans	...	...	21307.46	126490.5	119893			126490.5	126490.5	0	126490.5	TUESDAY	7	Y			1	0.103971			XAP	Approved	-699	Cash thro	XAP	Unacc	
2163253	154602	Consumer	loans	...	...	4187.34	26955	27297			26955	27297	1350	26955	SATURDA	12	Y			1	0.051324			XAP	Approved	-1473	Cash thro	XAP	Unacc	
1285768	142748	Revolving	loans	...	...	9000	180000	180000			180000	180000	0	180000	FRIDAY	13	Y			1				XAP	Approved	-336	XNA	XAP	Unacc	
2393109	396305	Cash	loans	...	...	10181.7	180000	180000			180000	180000	0	180000	THURSDA	14	Y			1				XNA	Approved	-700	Cash thro	XAP	Unacc	
1173070	199178	Cash	loans	...	...	4666.5	45000	49455			45000	49455	0	45000	SATURDA	16	Y			1				Everyday	Refused	-584	XNA	HC		

In the above screen explains the what about in loan\_application data set.



In above screen dataset loaded and in text area can see few records from dataset and in first graph x-axis represents loan status and y-axis represents Number of Records available in that loan status class label. In second graph x-axis represents rejection reason and y-axis represents records size and in dataset we have both numeric and non-numeric values so to convert to numeric data then click on Pre-process Dataset button to get below output



Future of Loan Approvals with Explainable AI

Upload Loan Application Dataset    Preprocess Dataset    Split Dataset Train & Test

Train AI on Loan Approval    Train AI on Loan Rejections    Explainable AI    Predict Loan Status using Test Data

Normalized \* Processed Features = [[ 0.20751101 -0.06945259 0.45191549 ... -0.37707578 -0.39106526  
-0.49585587]  
[ 1.648087 -1.66214853 -1.07302291 ... 2.62178964 2.53016275  
2.01671505]  
[ 1.12749363 -1.5268431 -1.07302291 ... 2.62178964 2.53016275  
2.01671505]  
...  
[-1.40415263 -1.05980545 1.97685389 ... -0.37673097 -0.39076936  
-0.49585587]  
[-1.68610751 -0.76522848 -1.07302291 ... -0.37673097 -0.39076936  
-0.49585587]  
[-0.26773929 -0.11305651 -1.07302291 ... -0.37673097 -0.39076936  
-0.49585587]]

In above screen dataset converted to numeric format and then click on  Split Dataset Train & Test  button to split dataset into train and test and then will get below output

Future of Loan Approvals with Explainable AI

Future of Loan Approvals with Explainable AI

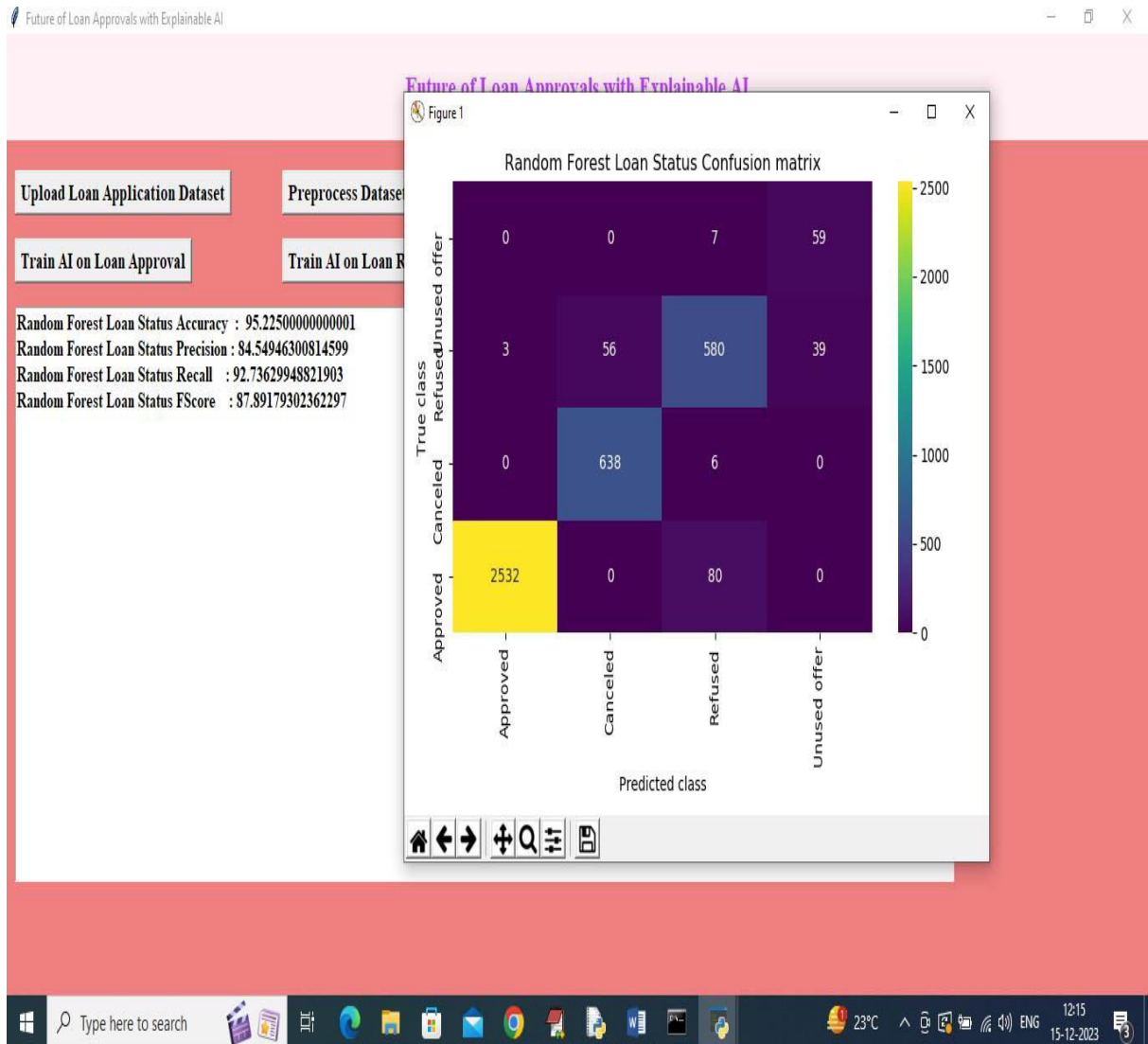
Upload Loan Application Dataset    Preprocess Dataset    Split Dataset Train & Test

Train AI on Loan Approval    Train AI on Loan Rejections    Explainable AI    Predict Loan Status using Test Data

Total records found in dataset = 20000  
Total features found in dataset = 35  
80% dataset for training : 16000  
20% dataset for testing : 4000

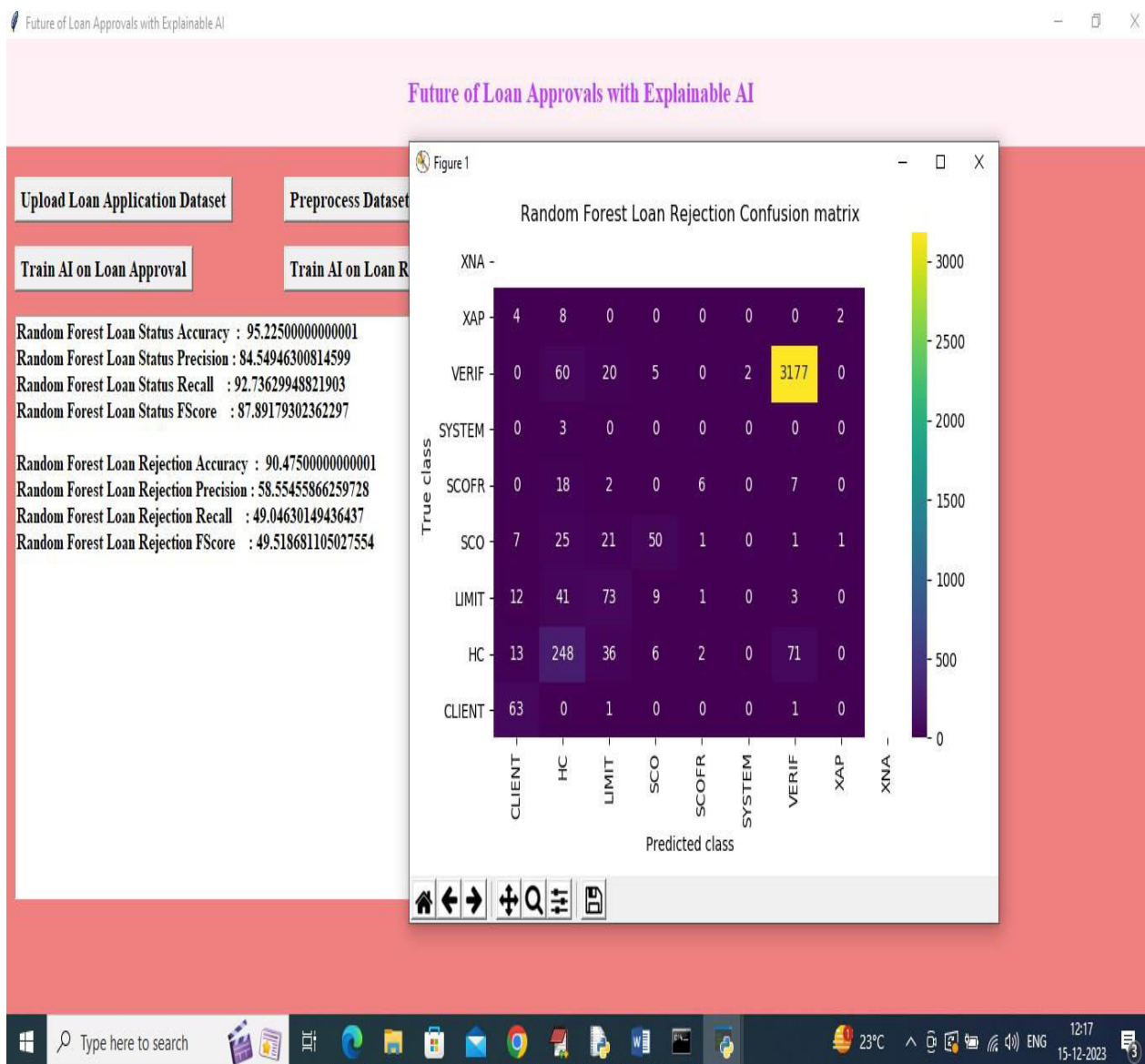
Windows taskbar: Type here to search, 23°C, 12:14, 15-12-2023

In above screen can see dataset size with total number of features and then can see TRAIN and TEST size and now click on [Train AO on Loan Status Approval](#) button to train AI and get below output

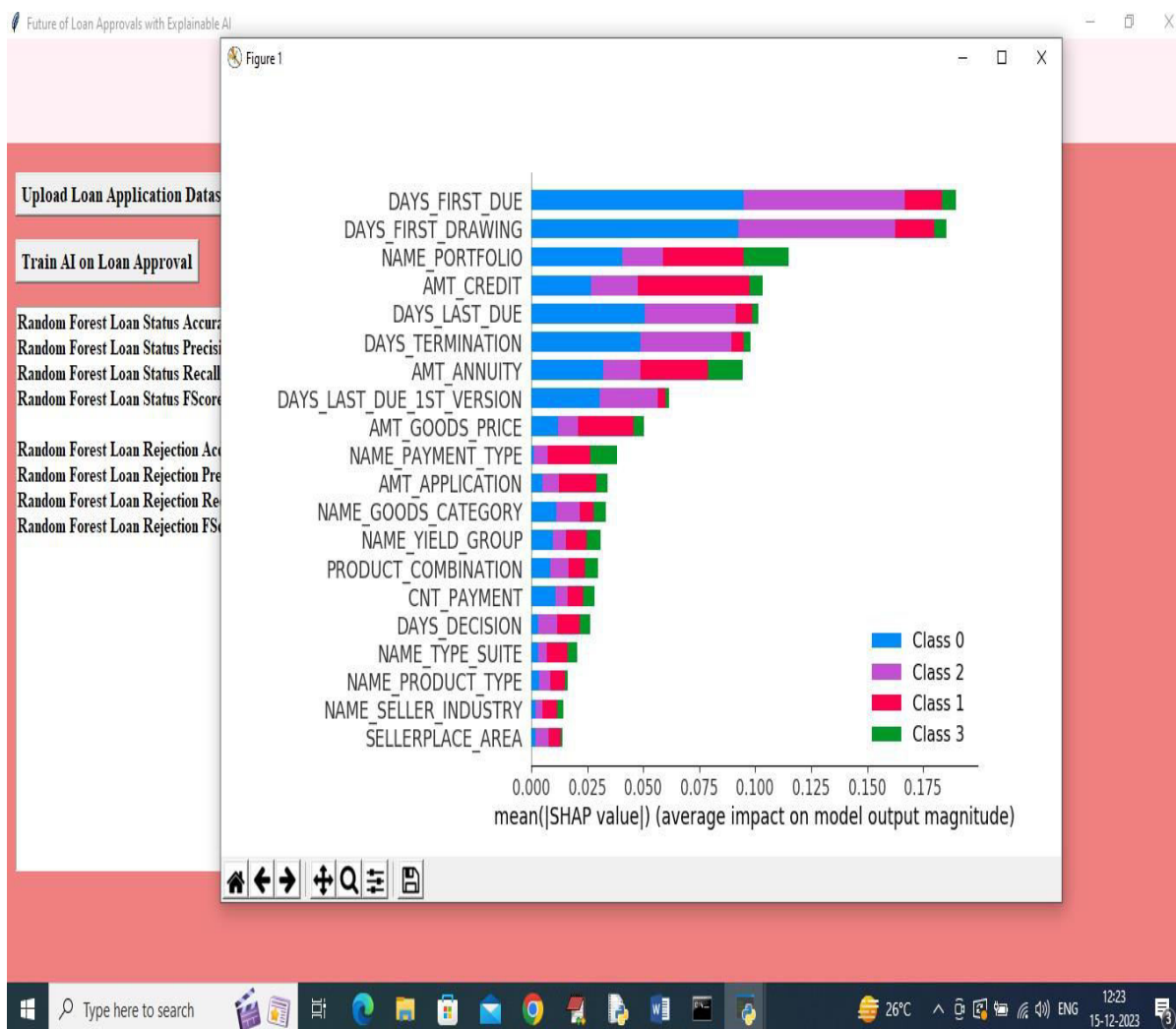


In above screen AI Random Forest got 95% accuracy on Loan STATUS and can see other metrics also. In above confusion matrix graph x-axis represents LOAN STATUS Predicted Labels and y-axis represents TRUE labels and all boxes in diagonal contains correct prediction count and remaining blue boxes contains incorrect prediction count which are very few. Now click on Train AI on Loan Rejections button to train AI on rejection reason and get below output

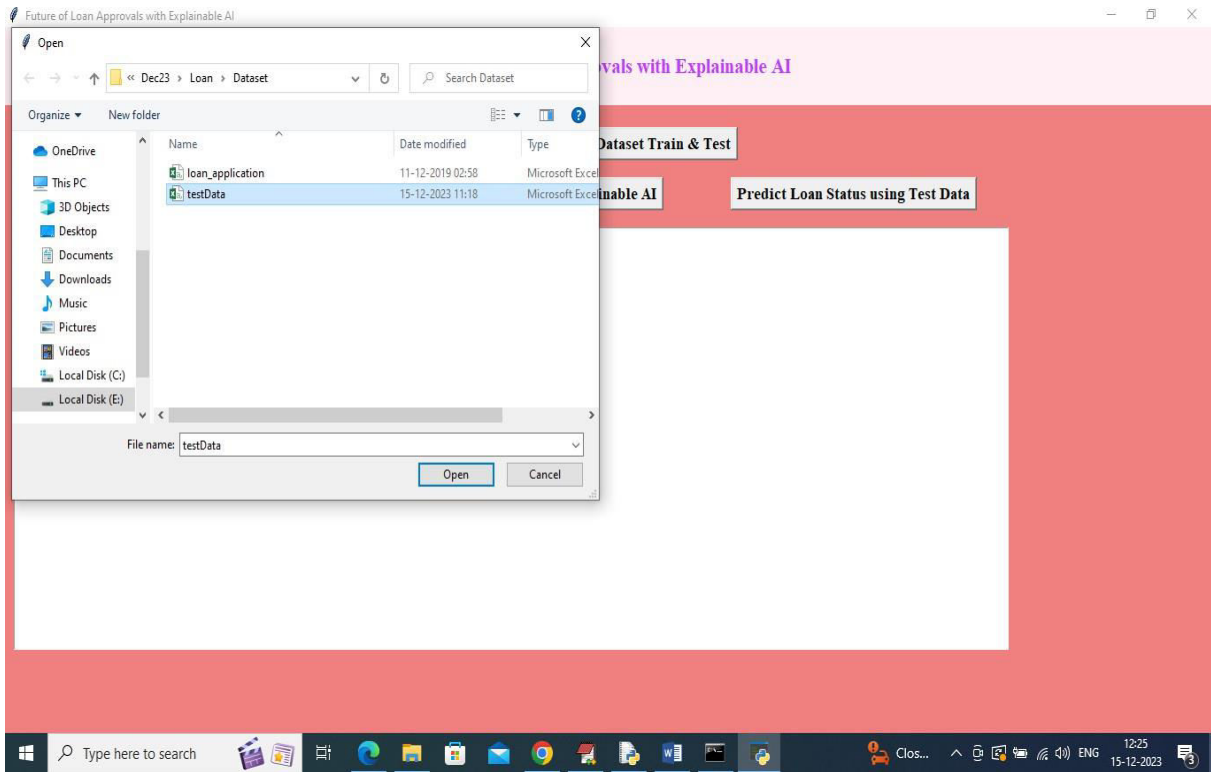




In above screen AI on REJECTION got 90% accuracy and in confusion matrix graph x-axis represents Rejection Reason Predicted Labels and y-axis represents True label and in diagonal boxes we can see correct prediction count and remaining boxes contains incorrect prediction count. Now click on Explainable AI button to get below features explanation on prediction

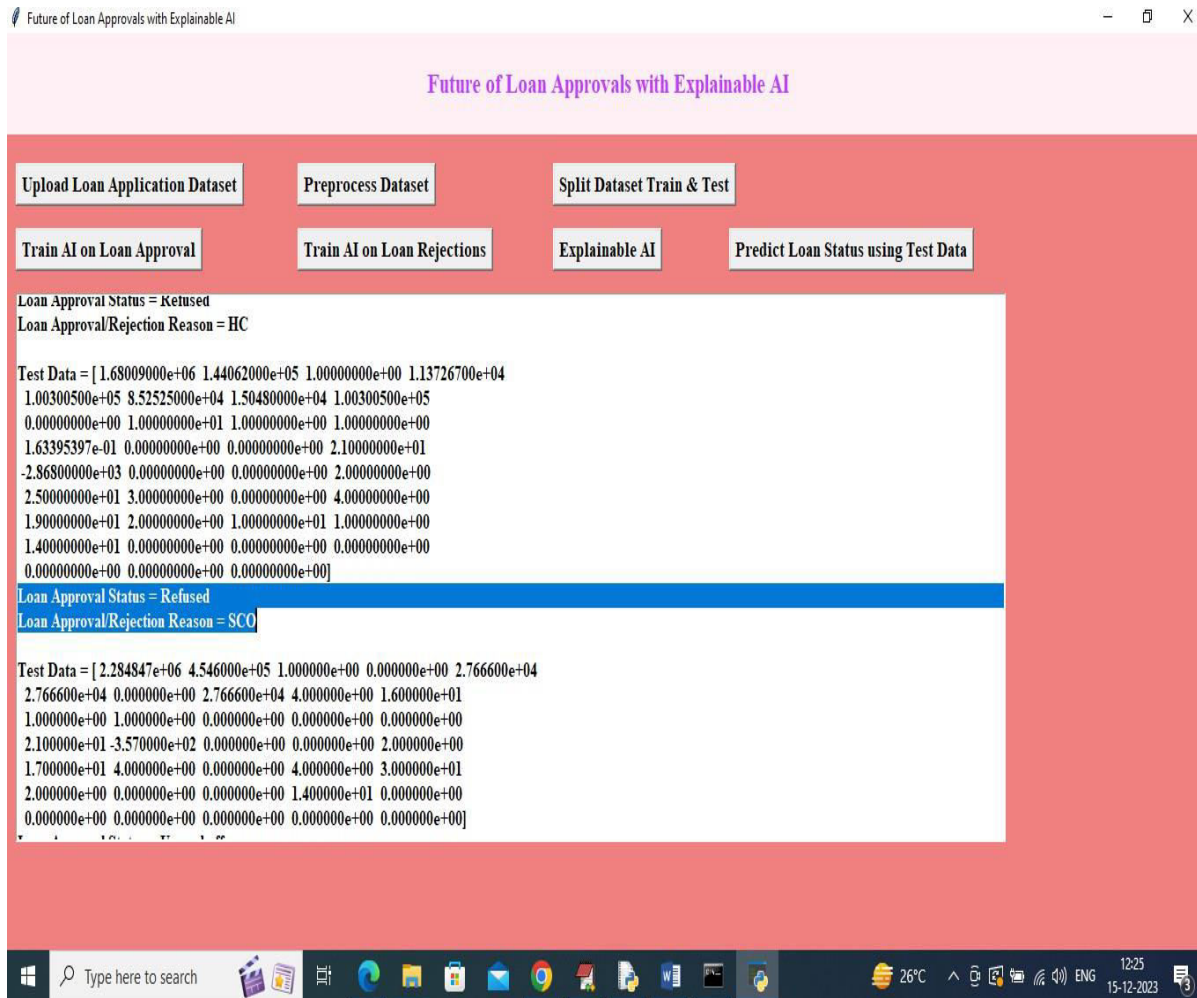


In above SHAP explanation screen in each bar we can see 4 different colours and each colour represents one class label and based on colour percentage we can say which feature names is contributing how much to predict that class label. Now close above graph and then click on [Predict Loan Status using Test Data](#) button to upload test data and then will get below prediction

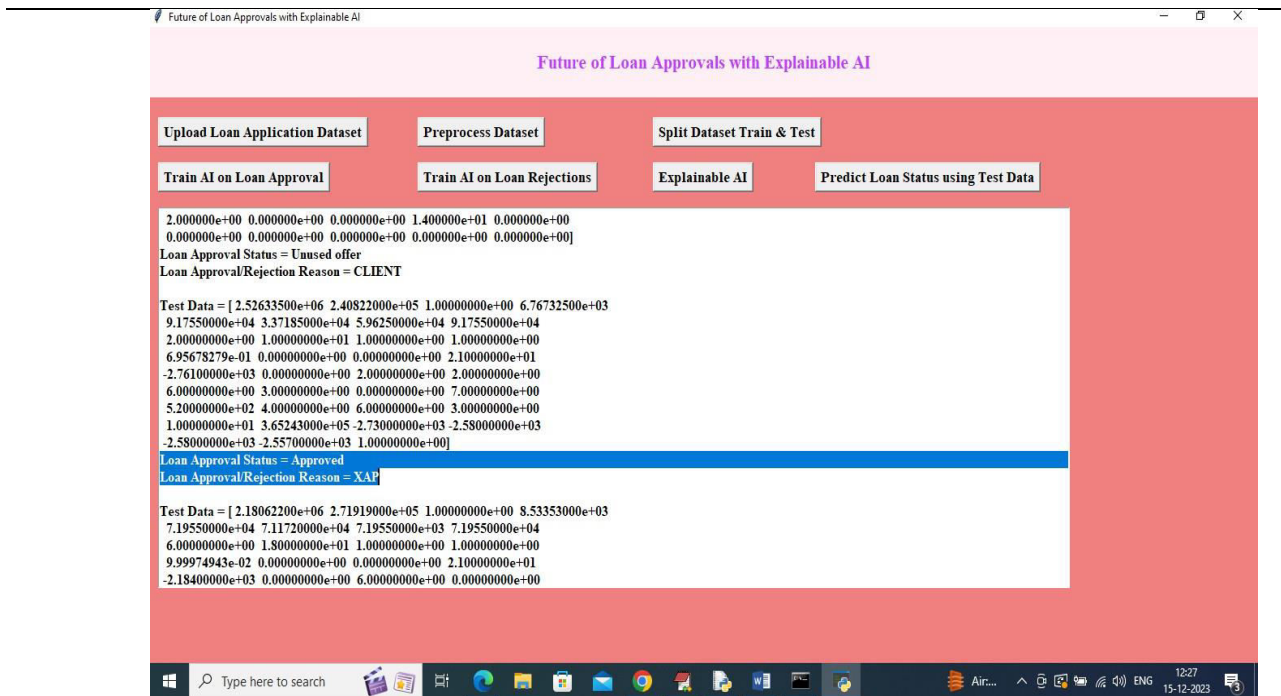


In above screen selecting and uploading testData.csv file and then click on **Open** button. In the below screen explains the what about in testData.csv data set.

SK_ID	PR	SK_ID_CUI	NAME_CC	AMT	ANN	AMT_APP	AMT_CREI	AMT_DOV	AMT_GOC	WEEKDAY	HOUR	AP	FLAG	LA	NFLAG	LA	RATE_D	DOV	RATE_I	INTI	NAME_CA	NAME_CC	DAYS	DEC	NAME_PA	CODE	REJ	NAME
1583704	315664	Cash loans	0	0						WEDNESD	15	Y	1							XNA	Refused	-430	XNA	HC				
2352627	266560	Revolving loans	0	0						WEDNESD	7	Y	1							XAP	Canceled	-323	XNA	XAP				
1348316	104826	Revolving	22500	450000	450000				450000	MONDAY	14	Y	1							XAP	Refused	-183	XNA	HC				
1680090	144062	Consumer	11372.67	100300.5	85252.5	15048	100300.5	FRIDAY	10	Y	1	0.163395							XAP	Refused	-2868	Cash thro	SCO					
2284847	454600	Consumer loans	27666	27666	0	27666	THURSDA	16	Y	1	0								XAP	Unused of	-357	Cash thro	CLIENT					
2526335	240822	Consumer	6767.325	91755	33718.5	59625	91755	SATURDA	10	Y	1	0.695678							XAP	Approved	-2761	Cash thro	XAP	Famil				
2180622	271919	Consumer	8533.53	71955	71172	7195.5	71955	WEDNESD	18	Y	1	0.099997							XAP	Approved	-2184	Cash thro	XAP	Spous				
1776835	449198	Cash loans	0	0						WEDNESD	14	Y	1							XNA	Canceled	-166	XNA	XAP				
1140216	297314	Consumer	15969.06	94459.05	100687.5	4.05	94459.05	TUESDAY	8	Y	1	4.38E-05							XAP	Refused	-2736	Cash thro	SCO	Childr				
1127838	183458	Consumer	3973.095	17959.05	13945.5	4459.05	17959.05	THURSDA	12	Y	1	0.263865							XAP	Approved	-2762	Cash thro	XAP	Spous				

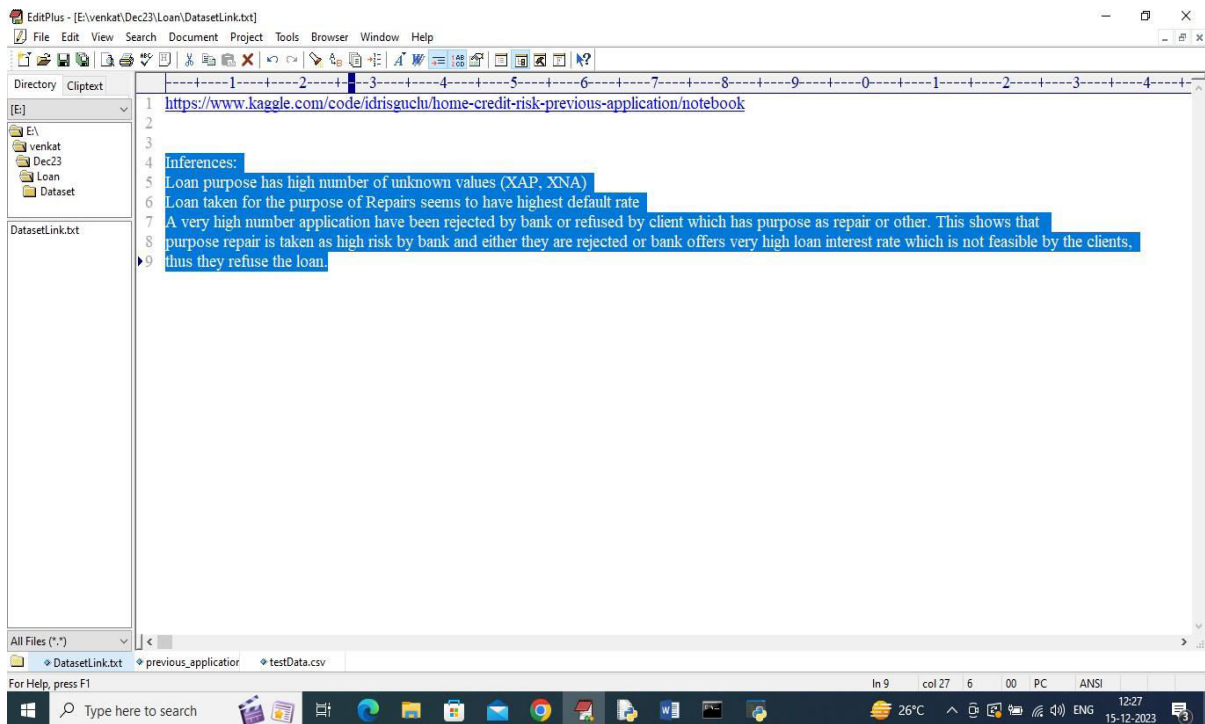


In above screen in square bracket, we can see test data and then in blue colour selected line next to TEST data we can see LOAN STATUS prediction and REASON details. Scroll down above output to view all predictions



In above screen can see other prediction output.

For Reason Rejected code you can read below description



In above screen read blue color selected text to know about REJECTED REASON codes

---

## 6. CONCLUSION AND FUTURE WORK

### CONCLUSION

In this paper, we presented the methodology to develop the belief-rule-based (BRB) system as an explainable AI decision-support-system to automate the underwriting process of lend loans. Unlike blackbox models, the BRB system can explicitly accommodate expert knowledge and can also learn from data by supervised learning, though the acquisition of expert knowledge can be a time-consuming and labor-intensive task.

The decision-making process in this system can be explained by the importance of rules activated by a data point representing a loan application and by the contribution of attributes in activated rules. Through a business case study, we have demonstrated that the proposed AI decision-support-system provides a good trade-off between prediction accuracy and explainability. The importance of activated rules and their attributes in the rules help to understand the reasoning behind the decisions. The textual explanations initiated by the chain of events in the factual-rule- base to the heuristic-rule-base could be sent to rejected applicants as reasons for denying their loan applications.

## 7. REFERENCES

1. Abellán, J., & Castellano, J. G. (2017). A comparative study on base classifiers in ensemble methods for credit scoring.,. *Expert Systems with Applications*, 73, 1-10.
2. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138-52160.
3. Aggour, K. S., Bonissone, P. P., Cheetham, W. E., & Messmer, R. P. (2006). Automating the underwriting of insurance applications. *AI magazine*, 27(3), 36-36.
- Aitken, R. (2017). 'All data is credit data': Constituting the unbanked. *Competition & Change*. 21(4), 274-300.
4. Akinyokun, O. C. (2015). Fuzzy logic-driven expert system for the diagnosis of heart failure disease. *Artif. Intel. Research*, 12-21.
5. Ala'raj, M., & Abbod, M. F. (2016). Classifiers consensus system approach for credit scoring. *Knowledge-Based Systems*, 104, 89-105.
6. Bensic, M., Sarlija, N., & Zekic- Susac, M. (2005). Modelling small- business credit scoring by using logistic regression, neural networks and decision trees. *Intelligent Systems in Accounting, Finance & Management: International Journal*, 13(3), 133-150.
7. Bijak, K., & Thomas, L. C. (2012). Does segmentation always improve model performance in credit scoring? *Expert Systems with Applications*, 39(3), 2433-2442.
8. Casalicchio, G., Molnar, C., & Bischl, B. (2018). Visualizing the feature importance for black box models. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, (pp. 655-670).
9. Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and

---

subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11(Jul), 2079- 2107.

10. Chen, W., Ma, C., & Ma, L. (2009). Mining the customer credit using hybrid support vector machine technique. *Expert systems with applications*, 36(4), 7611-7616.