TVS CREDIT

e.p.i.c 7

ENRICH | PERFORM | INNOVATE | CHALLENGE

Sanskar Ughade - IIT Guwahati

#YourEPICMove

Analysis

Link: [Colab]

Data Assesment

- Assessed dataset shape
- Identified NULL values
- Data profiling

Descriptive Analysis & Visualization

- Calculated descriptive statistics
- Understood central tendency
- Visualized the feature distributions

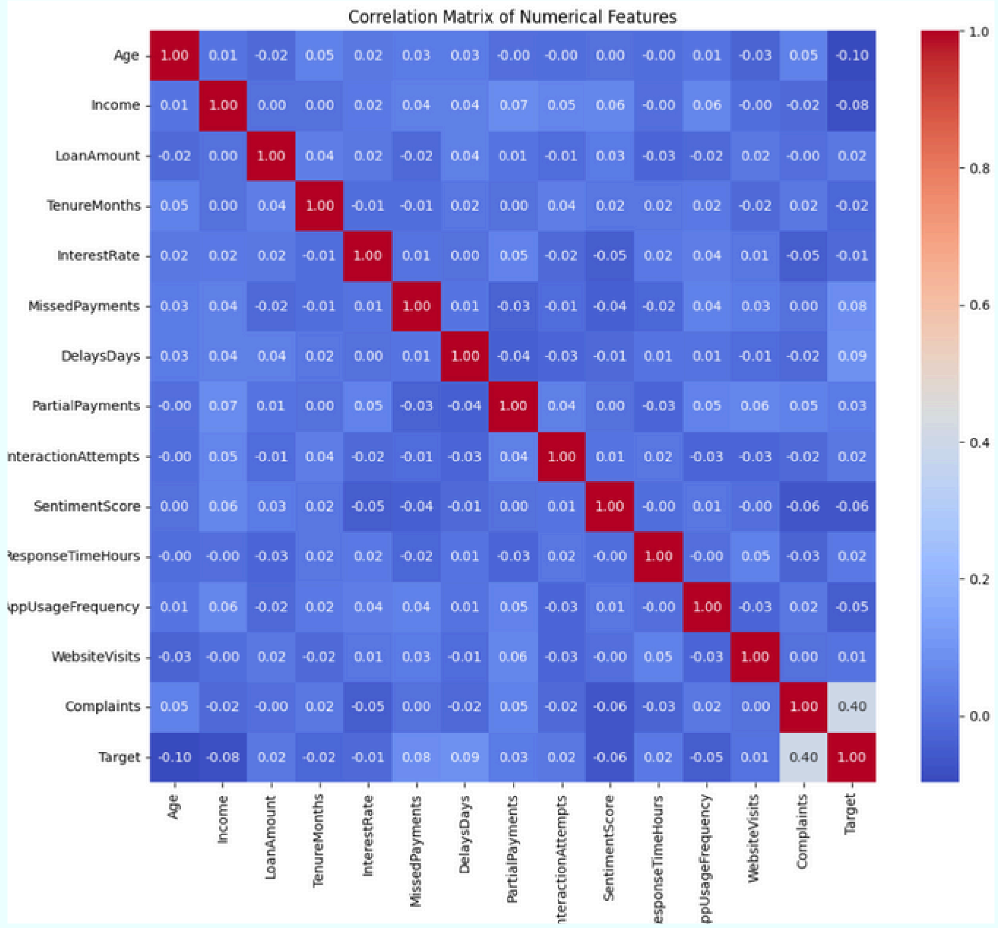
Categorical Analysis

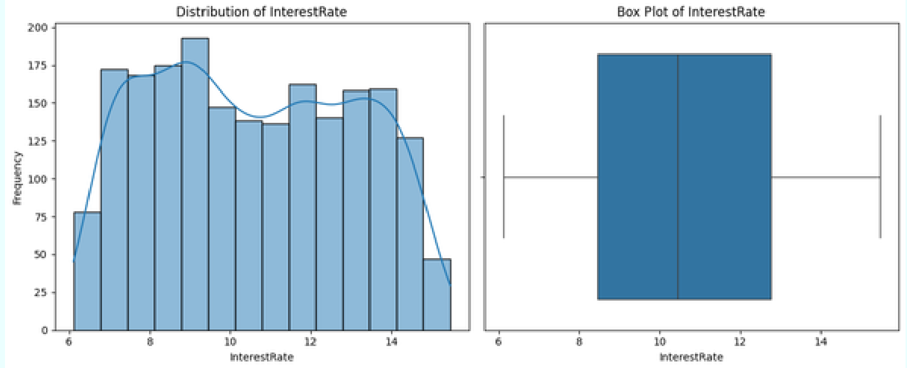
- Analyzed the distribution
- Visualized the freqeuncy among categories

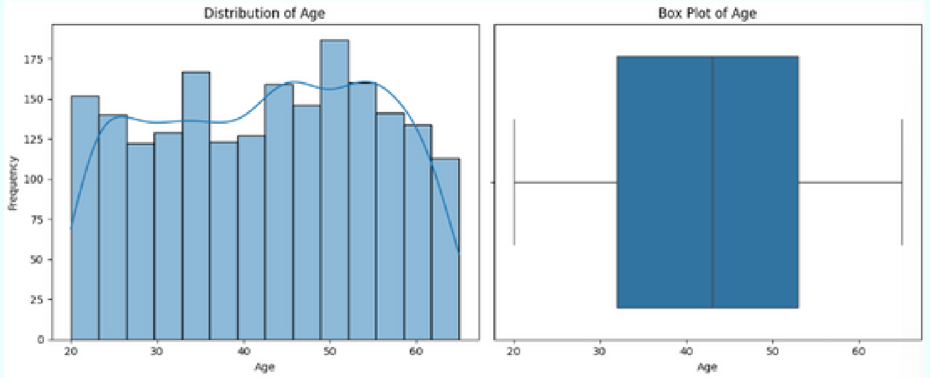
Correlation & Target Analysis

- Created a correlation Heatmap
- Analysed target distribution among categories (percentiles & count)
- Categorical target distribution visualization

Insights for Modeling

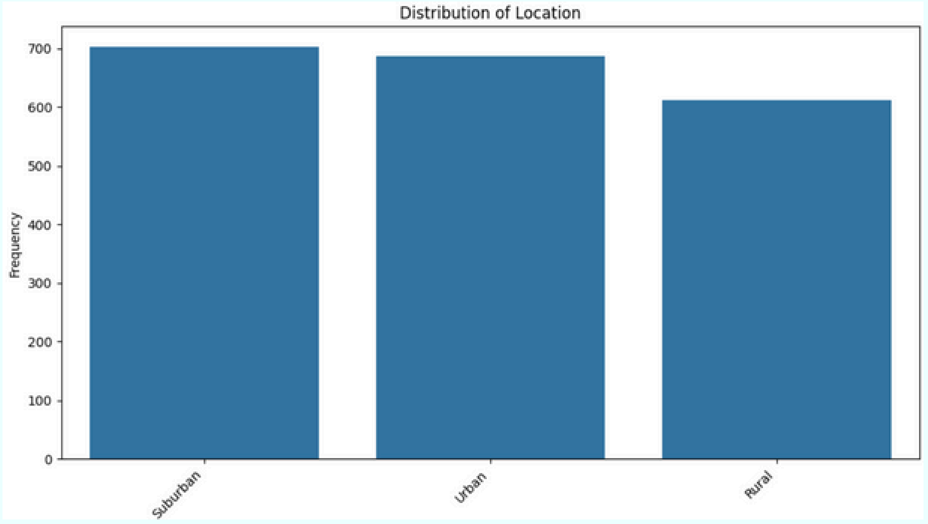


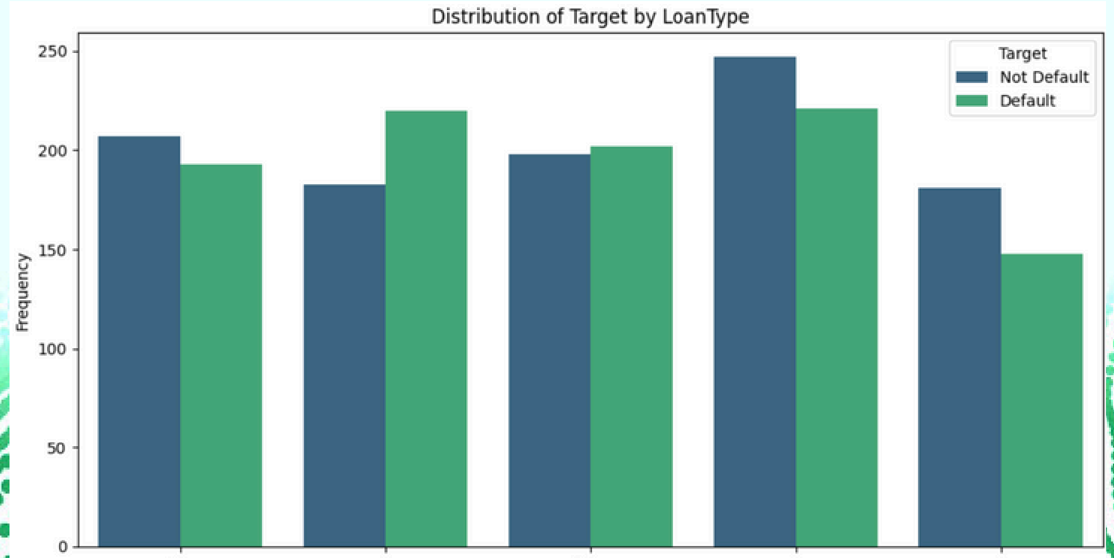




Target	0	1
LoanType		
Auto	247	221
Business	207	193
Education	183	220
Home	181	148
Personal	198	202

Location	
Suburban	702
Urban	686
Rural	612





Conclusion:
The target is strongly influenced by **complaints**, **delay days**, and **missed payments** (positive correlation), while **age** and **income** show a strong negative correlation.

These insights guide feature selection for predictive modeling.

#YourEPICMove

Prediction Modeling

- **One-hot encoded** categorical features into binary form.
- Created new features (eg Debt-to-Income Ratio, Delinquency Rate)
- Detected outliers with the **IQR technique**. Retained them to prevent dropping data. The chosen model could handle them.
- Normalised the dataset using **Min-Max Scaler** (range - 0 to 1)
- Split the data into training, validation and test sets.
- Chose the **Random Forest model** and performed hyperparameter tuning using **5-fold cross-validation** with GridSearchCV.
- Trained on full training set with optimal parameters.
- Assessed the model performance on the validation and test sets.
- Evaluated on validation and test sets using accuracy, precision, recall and **ROC-AUC**.
- Benchmarked against **XGBoost** and **LightGBM**.
- Fine-tuned tree threshold to improve metrics.
- Enhanced interpretability with **SHAP values** and dependence plots

Evaluation Metric

1. High precision

2. Low FN's(High Recall)

```
Random Forest Performance:
  Accuracy: 0.8175
  Precision: 0.8500
  Recall: 0.9122
------------------------------
XGBoost Performance:
  Accuracy: 0.8650
  Precision: 0.9081
  Recall: 0.8195
------------------------------
LightGBM Performance:
  Accuracy: 0.8600
  Precision: 0.9116
  Recall: 0.8049
------------------------------
Performance Summary:
XGBoost has the highest accuracy.
LightGBM has the highest precision.
Random Forest has the highest recall.
```

```
Accuracy on the test set (threshold=0.37): 0.8725
```

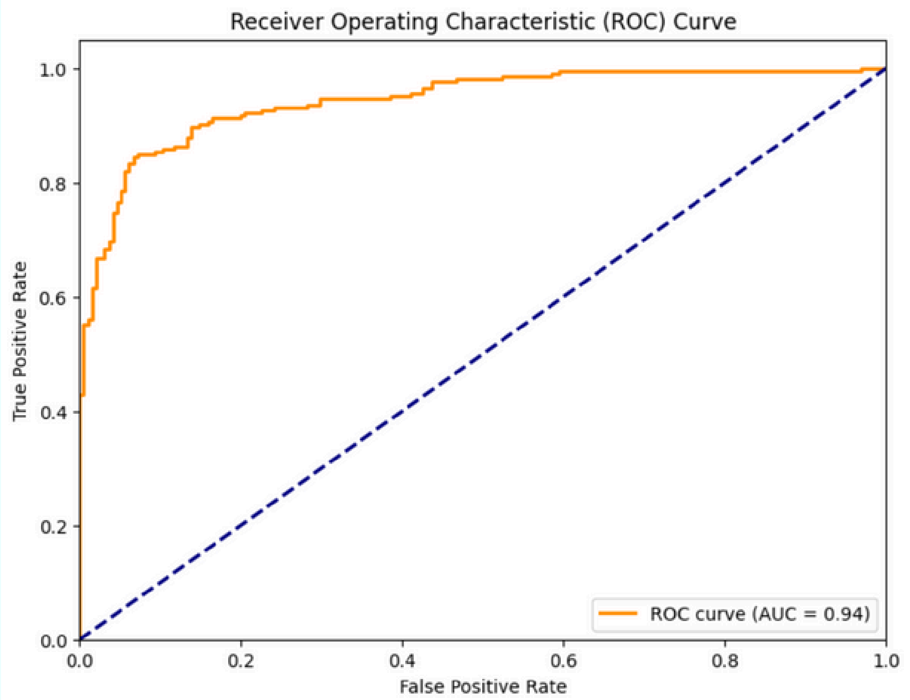
```
Accuracy on the Validation set (threshold=0.37): 0.8175
```

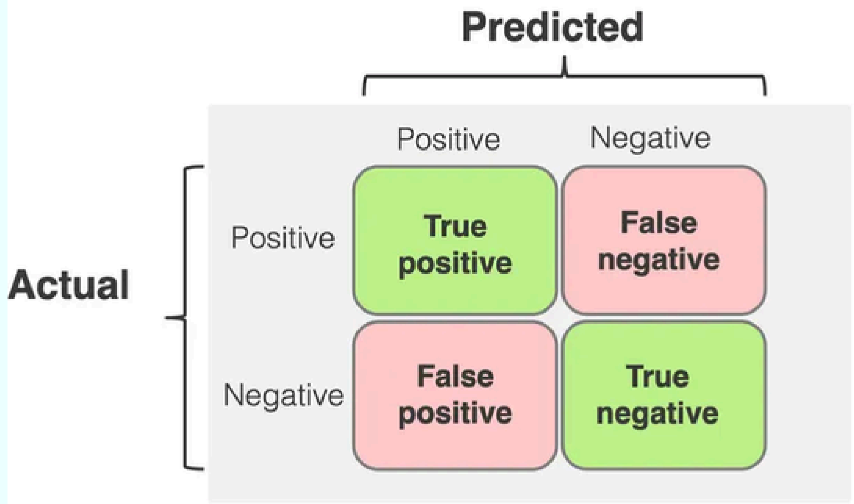
Test set

```
Precision: 0.8500
Recall: 0.9122
Confusion Matrix:
[[187  18]
 [ 33 162]]
```

Validation set

```
Precision: 0.7585
Recall: 0.9179
Confusion Matrix:
[[179  16]
 [ 57 148]]
```



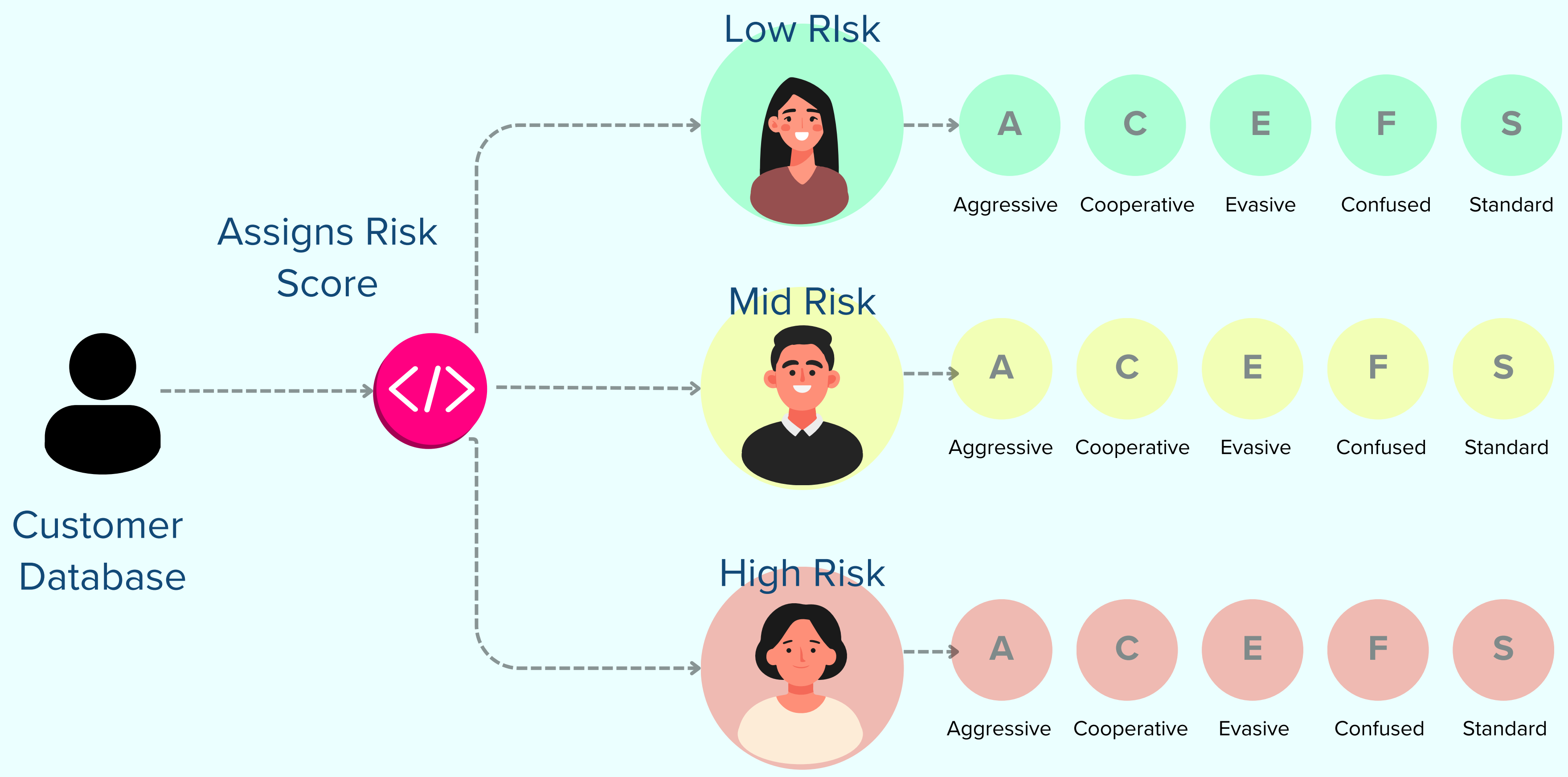


The evaluation criteria is chosen for the following Business Goals

1. Correctly identify actual defaulters as Default (**True Positives**)
2. Minimise cases where defaulters are predicted as Not Default (**False Negatives**), the riskiest scenario for the firm which means we in turn need to increase Recall
3. Prioritise **high precision** and **low False Negatives** in the confusion matrix

Recommendation Engine

TVS CREDIT e.p.i.c 7
ENRICH | PERFORM | INNOVATE | CHALLENGE



- Assigned a **risk score** to each customer based on the predictive model.
- Built a **rule-based engine** to generate personalised recommendations
- **Classified** customers into three categories, based on risk scores
  - Low Risk
  - Mid Risk
  - High Risk
- Defined **personas** to customers based on interaction history, complaints and other engagement signals
- Designed a **matrix to map strategies** for each risk persona combination
- Ran a script to generate sub-datasets based on customer behaviour and persona (e.g., aggressive customers, silent churners, loyal advocates etc.)

Note -

- **Problem Encountered** - Hardcoded the comparison thresholds, but the thresholds would only work on this dataset and not on any other dataset if there were variations in ranges.
- **Solution** - Shifted to percentile-based thresholds for features with a lower bound and upper bound percentile values. This made the thresholds dynamic, which adjusted to the dataset, leading to more robust recommendations.
- **Future Improvement** - In the future, the model can be trained to predict these thresholds automatically.

Dynamic Thresholds

```
sentiment_low = df['SentimentScore'].quantile(0.25)
sentiment_high = df['SentimentScore'].quantile(0.75)
complaints_high = df['Complaints'].quantile(0.75)
interaction_high = df['InteractionAttempts'].quantile(0.75)
response_high = df['ResponseTimeHours'].quantile(0.75)
website_visits_high = df['WebsiteVisits'].quantile(0.75)
```

#YourEPICMove

Recommendations were designed to leverage the chatbot for **friendly** and **empathetic** engagement with low-risk, cooperative customers, while **escalating** to human intervention as risk levels and customer aggression increases.

Strategy Matrix

Risk Level \ Behavior	Aggressive	Cooperative	Evasive	Confused	Standard
High	Senior Escalation & Pause Automated Contact	Empathetic Call from Human Agent to Offer Solutions also for settlement option	Assertive Human Follow-up & Multiple Channels	Problem-Solving Call from Human Agent	Firm Human Call to Discuss Payment
Medium	Immediate Handoff to Human Agent	Chatbot with Flexible Payment Options with loan restructuring	Automated Follow-up via SMS/Email	Ask Chatbot to help with Step-by-Step Payment Guide	Assertive Chatbot Reminder
Low	Formal but Automated Notice	Gentle Chatbot Nudge and early repayment option with discount etc	Proactive Chatbot Nudge with Direct Payment Link	Informative Chatbot to Explain Bill	Standard Automated Reminder

#YourEPICMove

ChatBot Bot Link

- Name of the bot : **TVS Sahahyak**
- Goal: To build a production-ready NBFC support chatbot that remembers users and fetches their financial data.
- Platform of Choice: Started with WhatsApp, switched to Telegram for faster setup.

Detailed Description

Database Design

- Customer Data → stores financial details.
- Conversation Memories → remembers chat history & key user info.

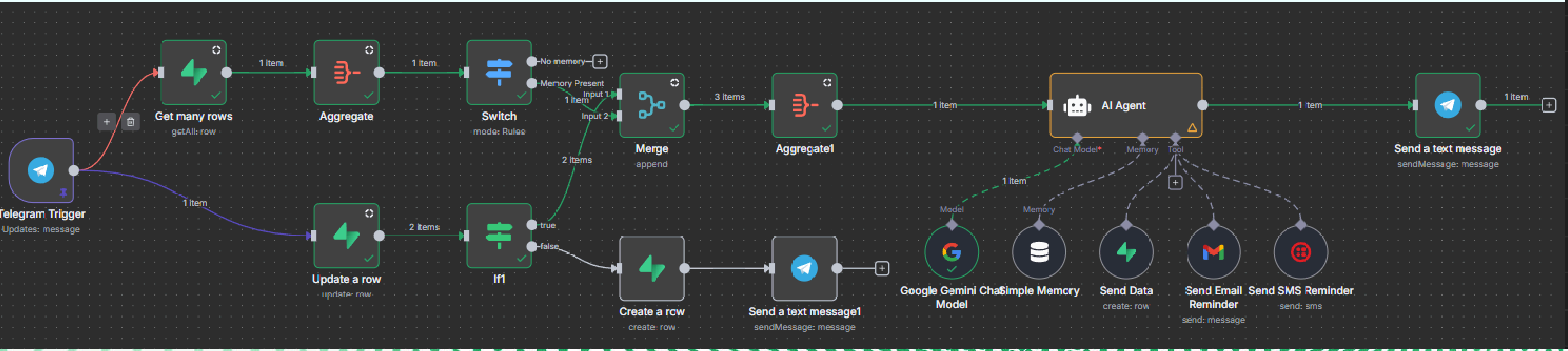
Tech-Stack

n8n Automation tool + Supabase for Database + Telegram for Chat integration

Conversational memory	
id	int8
created_at	timestamptz
Message	text
Sender First name	text
Sender Last name	text
Agent Response	json
Chat id	int8

Customer Data	
id	int8
created_at	timestamptz
First Name	text
Last Name	text
Loan amount	int8
Address	text
EMI	numeric
Intrest rate	numeric
No of missed payment	numeric
Credit Score	numeric
Customer Chat History	text
Chat id	numeric
Account No	int8
Loan type	text
Risk Score	numeric
Persona	text
Employement	text

n8n workflow



#YourEPICMove

TVS CREDIT e.p.i.c 7
ENRICH | PERFORM | INNOVATE | CHALLENGE

Features & USP

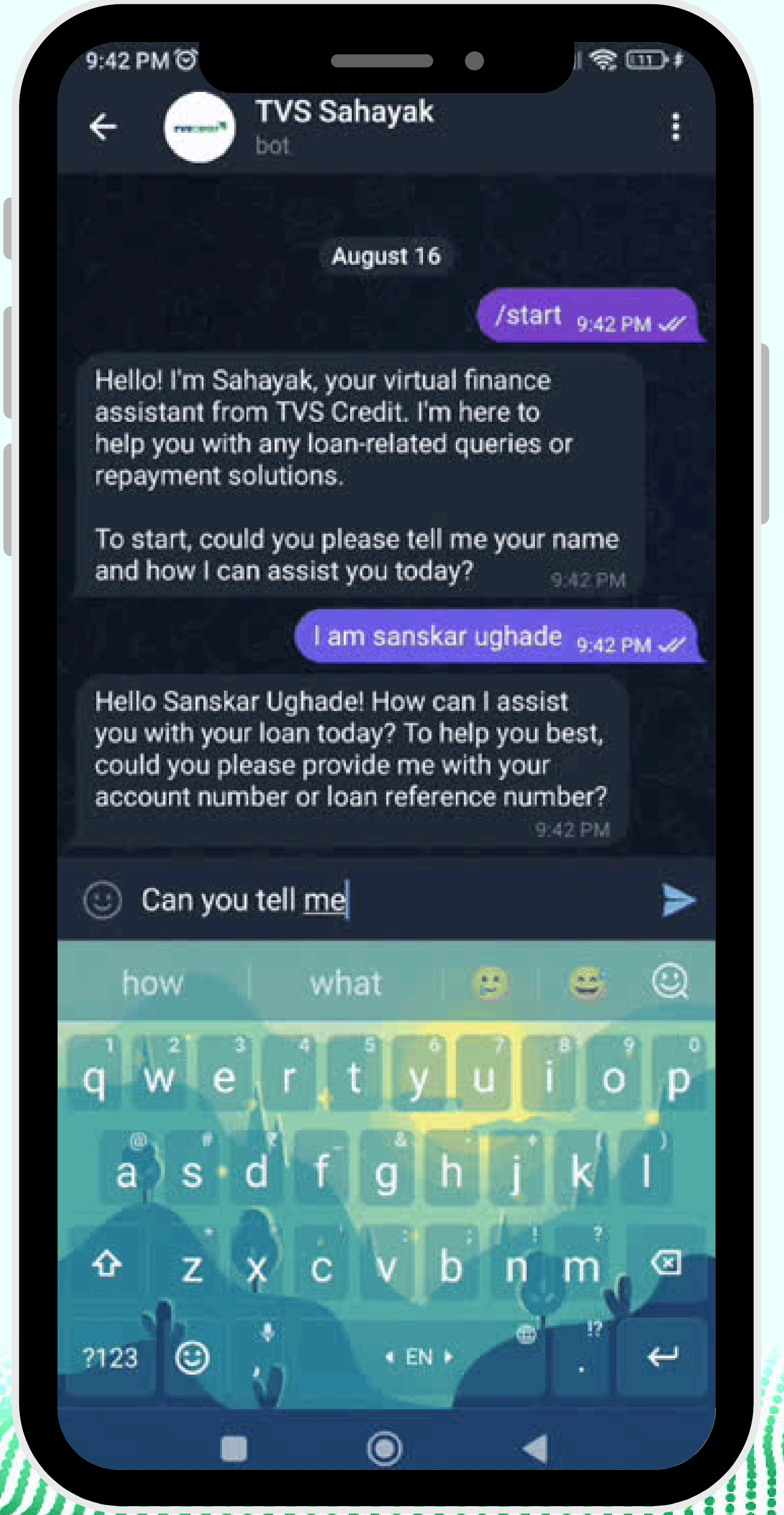
1. **Context Aware Responses** - Remembers past conversations to provide relevant answers without repeating questions.
2. **Direct Financial Data Access** - Instantly fetches customer information for faster resolution.
3. **Personalised Assistance** - Tailors replies based on user profile and past interactions.
4. **Seamless Onboarding** - Automatically identifies and saves new users without manual data entry.
5. **Robust Identification System** - Uses unique chat ID for accurate, consistent user recognition.
6. **Multi Case Handling** - Adapts to users with or without stored data and keeps workflow smooth.
7. **Optimized Workflow** - Built to eliminate duplicate responses and handle edge cases efficiently.
8. **Send Email and SMS reminders** - automate personalized emails with dynamic fields, reliable delivery, and opt-out handling.

Ensuring Data Privacy & Security

- Encrypt all data in transit (TLS 1.3) and at rest (AES-256)
- Mask or tokenise sensitive fields like account number & personal information before model input
- Enforce RBAC (role-based access) & strict access logging
- Keep PII (Personally Identifiable Information) within India to comply with DPDP Act 2023 & RBI guidelines)

Data-flow

- User Interaction → Chatbot UI captures queries (WhatsApp/Web/App).
- n8n Workflow Orchestration → Routes requests, manages logic & privacy.
- Mask/Tokenize PII → Replace sensitive info (e.g., account numbers) with tokens before sending to Gemini.
- Gemini API for Language → Handles natural conversation, never sees raw PII.
- Secure Backend (India Servers) → n8n maps tokens back to real data and fetches loan/account details.
- Merge Responses → Combine Gemini's reply with masked account details (e.g., account ending 7890).
- Compliance Ensured → PII stays in India (DPDP Act 2023 & RBI guidelines).
- Scalable → n8n workflows can scale via Docker/Kubernetes, handle multiple users in parallel.



#YourEPICMove

Scaling process

Scaling Guide

Hosting Setup → **Production Database** → **Queue-based Execution** → **Horizontal Scaling** → **File & Data Handling** → **Monitoring & Alerts** → **Reliability & Security**

Hosting Setup
Deploy on Cloud VM (AWS/GCP/Azure) or n8n Cloud.
For advanced scaling ➔ use Docker / Kubernetes for orchestration

Production Database
Switch from SQLite ➔ PostgreSQL for reliability & performance.

Queue-based Execution
Use Redis + Worker processes to run many workflows in parallel.

Horizontal Scaling
Add more workers as load increases.
Use Load Balancer for main servers (sticky sessions).

File & Data Handling
Store large files in Cloud Storage (e.g., S3).
Daily backups for database & workflows.

Monitoring & Alerts
Metrics dashboard (Prometheus + Grafana).
Track queue size, errors, worker health.

Reliability & Security
Secure API keys & credentials (encryption).
PgBouncer for DB stability under heavy load.

ROI Under Different Scenarios

Case	Default Reduction (₹100 Cr portfolio)	Infra Cost	Net Benefit
Worst	5%	₹50k/month	~₹4.9 Cr
Base	10%	₹25k/month	~₹9.7 Cr
Best	20%	₹10k/month	~₹19.9 Cr

Cost Benefit Analysis

Costs	Benefits
Development & Hosting - Cloud infra ₹15–25k/month, Dev & maintenance ₹5–10L/year	**Financial Savings** - 10–15% lower defaults → ₹10–15 Cr savings on ₹100 Cr portfolio
Integration & Training - System integration ₹2–3L Staff training ₹50k	**Operational Efficiency** - Automates 70–80% queries, frees staff for high-risk cases
Operational Costs - Monitoring & backups minor, optional enterprise scaling	**Customer Experience** - 24×7 support, faster resolutions, improved repayment rates
	Scalability & Readiness - Thousands of users handled without extra staff, extendable to WhatsApp/IVR/Email

#YourEPICMove

Appendix - I

Every month you people harass me. I told you I'll pay when I have money. Stop spamming or I'll complain to RBI. 01:05

I understand this must be frustrating for you, and I apologize if our communications have felt like harassment. That is certainly not our intention.

Given your situation, I want to assure you that we are here to help find a solution that works for you. I'll have a human agent reach out to you shortly to discuss your options and see how we can ease this burden. They will be able to talk through things in more detail and explore solutions that are tailored to your current financial circumstances. 01:05

Frustrated Complaint

I never took any loan from your company, why are you messaging me? 01:09

I understand your concern. It sounds frustrating to receive messages about a loan you don't believe you took. To help me investigate this, could you please provide your full name and any account number you might have with us, even if you don't recall taking a loan? This will help me look into our records and understand what's happening. 01:09

Frustrations & Suspicions

I lost my job, but I can pay half EMI for the next 3 months. Can you freeze interest for me? 01:01

I understand this is a difficult time for you, and I want to assure you that we're here to help. As I mentioned before, a human agent will be contacting you shortly to discuss your situation in detail. They will be able to explore options like freezing interest, restructuring your EMI, or other temporary solutions to ease your burden during these next few months. We'll work together to find the best way forward. 01:01

Request

I got a message saying I missed an EMI, but I paid on 1st Sept. Why are you showing me as default? 00:58

I understand your concern, Sanskar. It can be frustrating to receive a message about a missed EMI when you believe you've already made the payment.

To help me look into this, could you please tell me which loan account or type you're referring to? This will help me quickly check the payment status for the correct loan. 00:58

Confusion

Your executive was very rude last week. Anyway, tell me how much I need to pay this month. 00:58

I'm truly sorry to hear that you had a rude experience with one of our executives last week. That's not the impression we want to leave, and I sincerely apologize for any distress it caused.

Regarding your payment this month, your EMI amount is ₹20,000.

I want to make sure you feel supported. If managing this amount feels overwhelming right now, we can explore options like EMI restructuring to lower your monthly payments or even flexible micro-payment plans that might be easier to manage. My goal is to help you find a comfortable path forward. 00:58

Complaint and Dissatisfaction

Payment Reminder Mail

Payment Reminder For Home Loan Inbox ×

sanskarughadeuna@gmail.com
to me

11:44 AM (15 minutes ago)

Payment Reminder — TVS Credit

Hi Sanskar Ughade,

This is a friendly reminder that your loan payment for account ending with **41578962366** is due on **[Due Date, e.g., September 25, 2025]**.

Amount Due: **5112 ₹**

Due Date: **[Due Date, e.g., September 25, 2025]**

To make your payment securely online, please copy and paste the following link into your web browser:

Pay Now

If you have already made this payment, please disregard this message. It may take a couple of business days for the payment to reflect in our system.

Thank you for your prompt attention to this matter.
Sincerely,
TVS Credit

I have two loans with your company. One home loan and one bike loan. What's the EMI status of both? 00:53

Hello Sanskar! I can certainly help you with that.

I see that you have a Home Loan and a Car Loan with us. Could you please confirm if the 'Car Loan' is the one you are referring to as your 'Bike Loan'?

Based on our records:

For your Home Loan (Account ending in 66):
Your EMI is ₹5,112. It appears there are 3 missed payments on this loan.

For your Car Loan (Account ending in 52):
Your EMI is ₹20,000. It appears there are 5 missed payments on this loan.

I understand that managing multiple EMIs can sometimes be challenging. We're here to support you and find the best way forward. Would you be interested in exploring options like EMI restructuring to lighten your monthly load, or perhaps a one-time settlement offer if things feel overwhelming? We can certainly work together to find a solution that helps you breathe easier. 00:53

Loan Query

ScreenShots

Bot-response in different situations

I can't pay the full EMI of ₹15,000 this month. Can I pay ₹7,000 now and the rest later? 00:54

I understand this can be a challenging situation, and I'm here to help you find a comfortable solution. It's good that you reached out.

I see your EMI is ₹15,000 this month, and you're looking to pay ₹7,000 now. Just to confirm, is your usual EMI ₹15,000, or is it ₹20,000 as per our records for your car loan?

Could you also tell me about your current employment status? This information will help me explore the best possible options to support you. We can look into solutions like adjusting your EMI plan temporarily or exploring a different repayment structure that eases your financial burden this month. 00:54

Repayment Issue

I paid via UPI yesterday at 6:30 PM, but it's still showing unpaid. 01:05

I understand your concern, Sanskar. It can be frustrating when a payment doesn't reflect immediately.

To help me investigate this for you, could you please provide the UPI transaction ID and the exact amount you paid? Also, could you confirm your loan account number? This will help me track down the payment. 01:05

Payment Status Complaint

Appendix - II

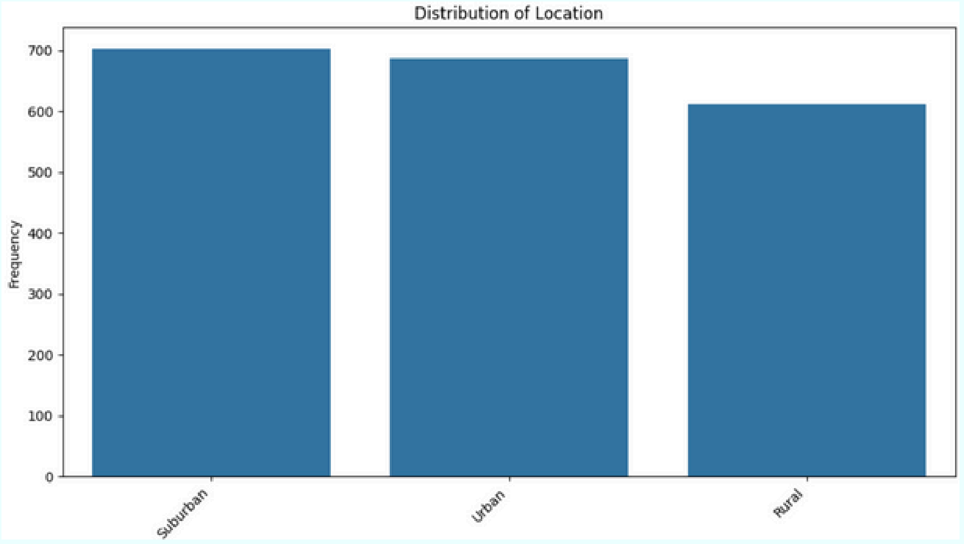
Analysis Category Distributions

Shape of Dataset
(2000, 19)

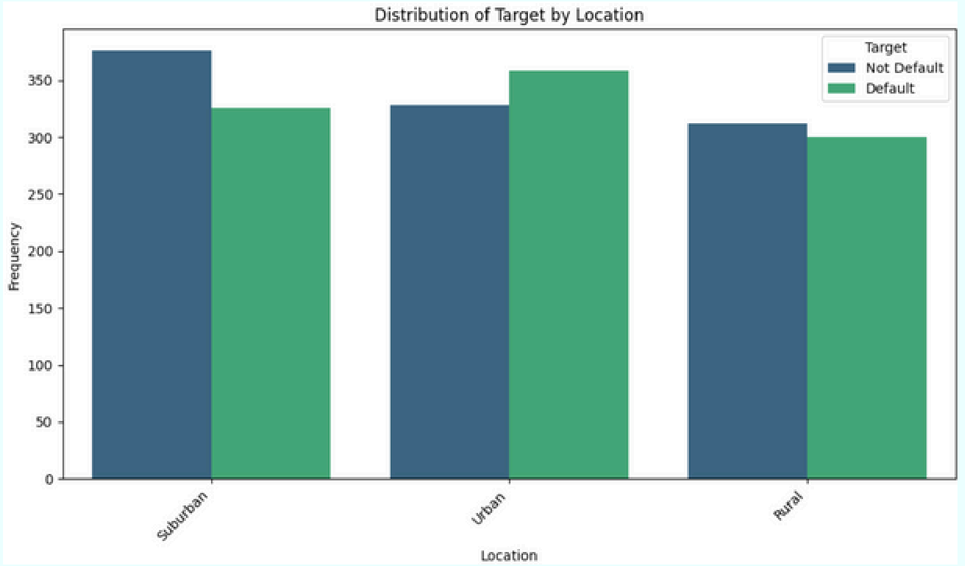
Missing Values
No Null Values

CustomerID	0
Age	0
Income	0
Location	0
EmploymentStatus	0
LoanAmount	0
TenureMonths	0
InterestRate	0
LoanType	0
MissedPayments	0
DelaysDays	0
PartialPayments	0
InteractionAttempts	0
SentimentScore	0
ResponseTimeHours	0
AppUsageFrequency	0
WebsiteVisits	0
Complaints	0
Target	0

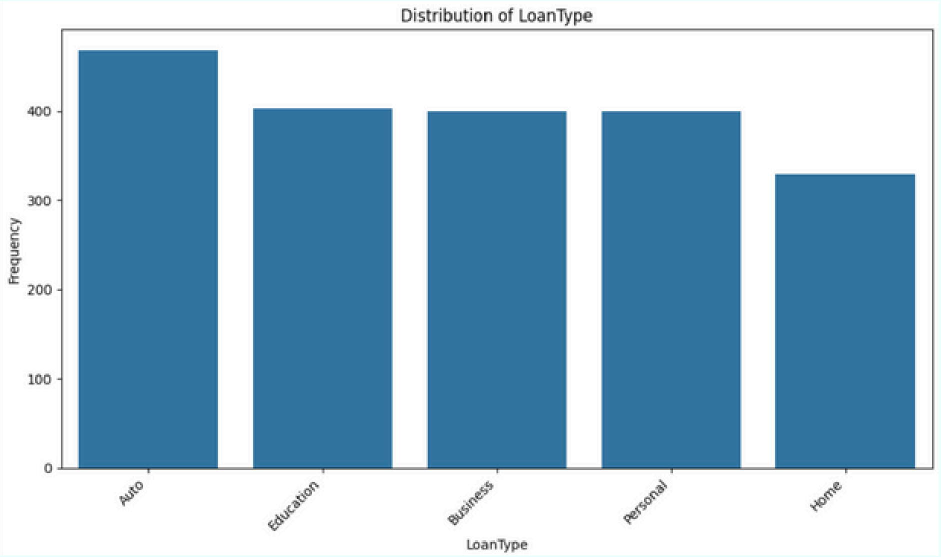
Location	
Suburban	702
Urban	686
Rural	612



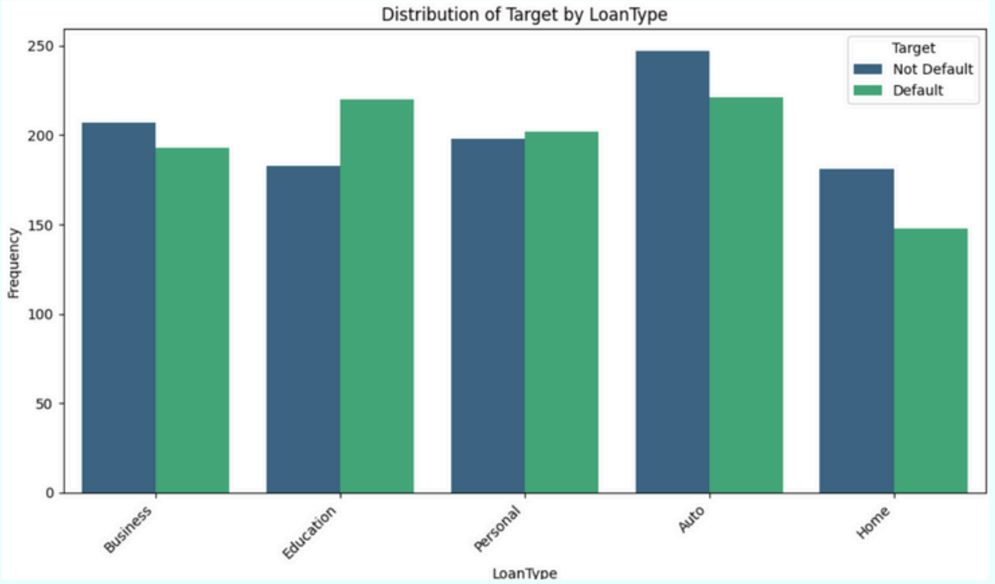
Target	0	1
Location		
Rural	312	300
Suburban	376	326
Urban	328	358



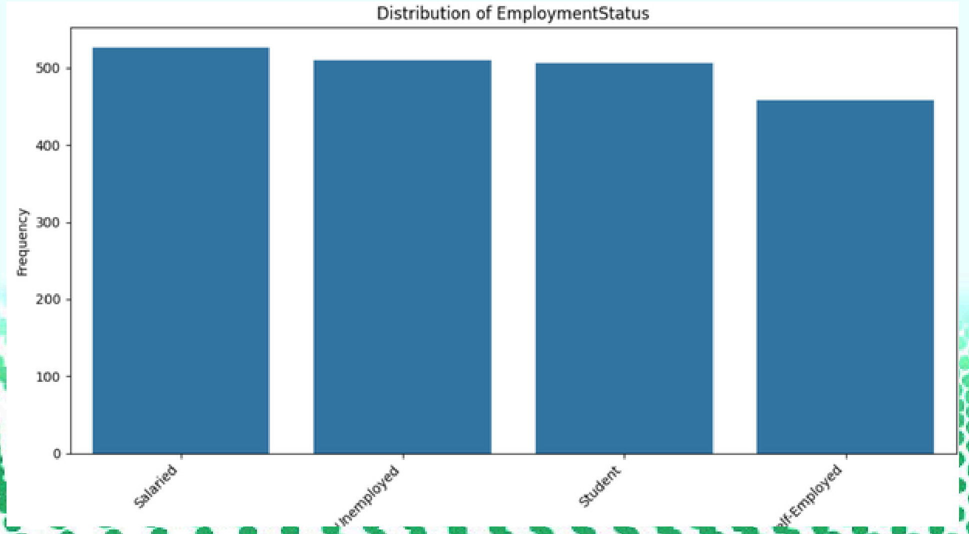
LoanType	
Auto	468
Education	403
Business	400
Personal	400
Home	329



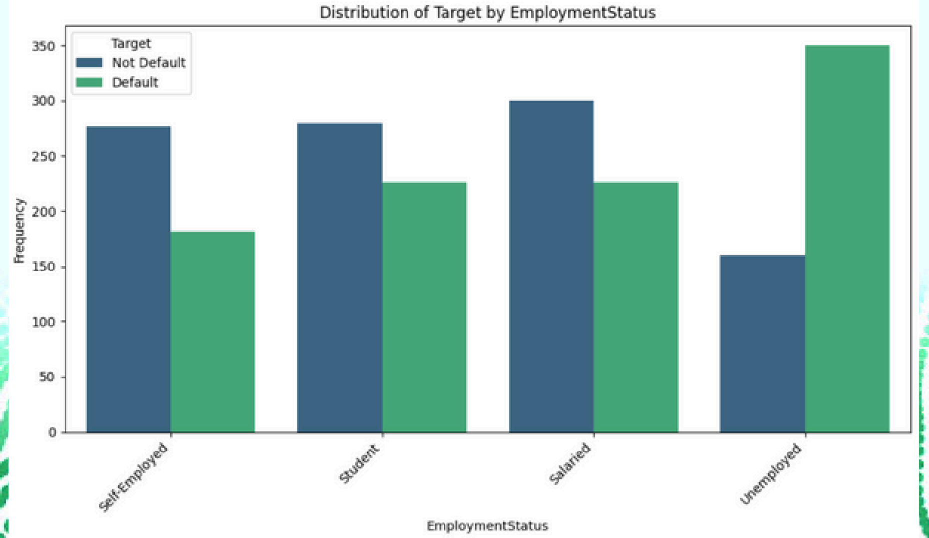
Target	0	1
LoanType		
Auto	247	221
Business	207	193
Education	183	220
Home	181	148
Personal	198	202



EmploymentStatus	
Salaried	526
Unemployed	510
Student	506
Self-Employed	458

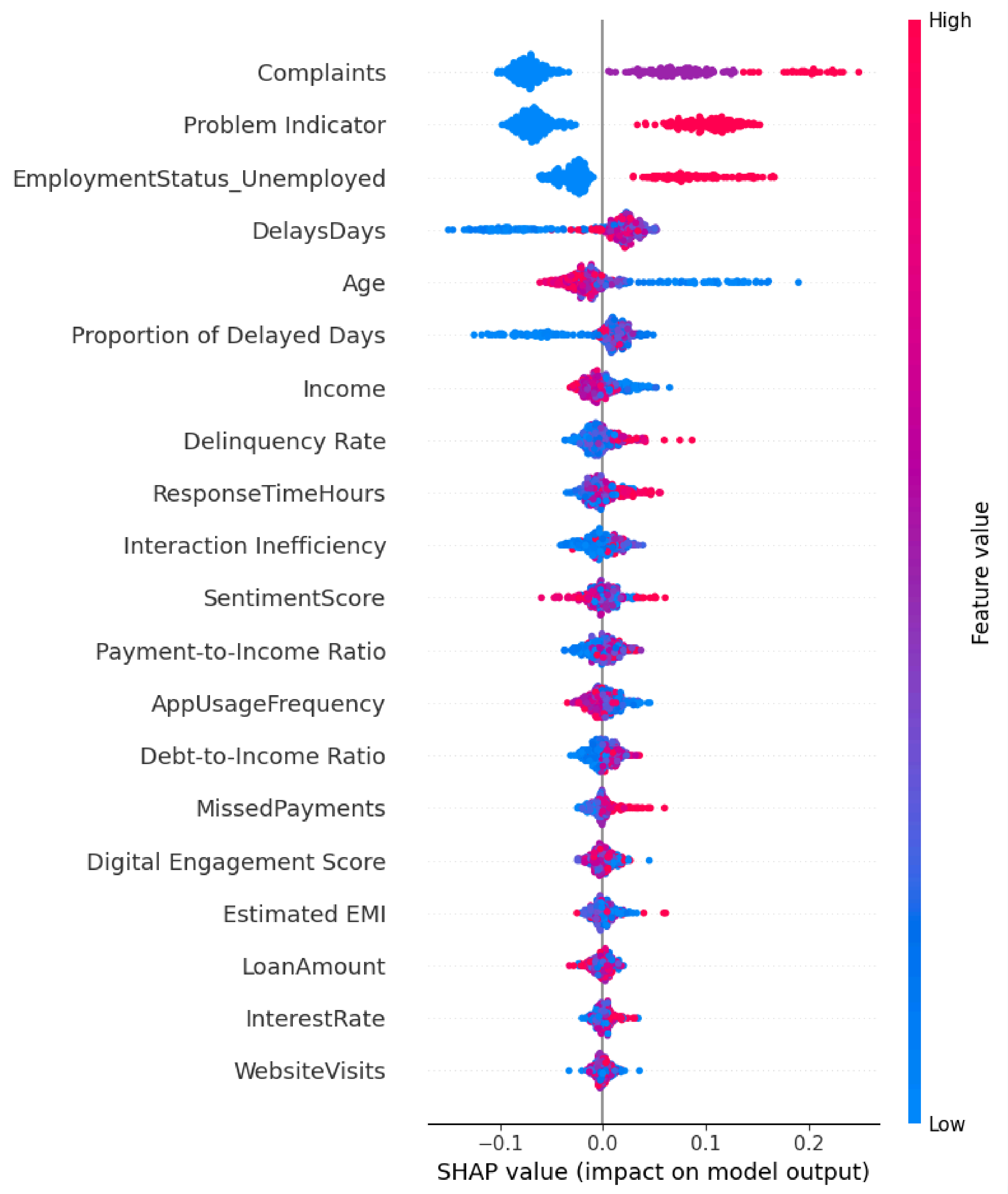


Target	0	1
EmploymentStatus		
Salaried	300	226
Self-Employed	276	182
Student	280	226
Unemployed	160	350



Appendix - III

SHAP Values



Links

- Analysis Notebook Collab | Drive
- Prediction Model Notebook Collab | Drive
- Recommendation Model Notebook Collab | Drive
- Chatbot (Sahayak) n8n workflow Drive
- Description of Building the Chatbot Doc
- Prediction with risk scores Link
- Ideations Link
- Recommmendation Strategy Output Link

#YourEPICMove

Thank You