SBI HACKATHON 2025

PROTOTYPE PRESENTATION

INTELLIGENT LOAN RISK ASSESSMENT & TRACKING SYSTEM

TEAM IDATEN:
PRIYANSHU MAURYA
VISHNU SINGH
ANUBHAV SAHA
AADITYA SHAHI

THE PROBLEM STATEMENTS

01

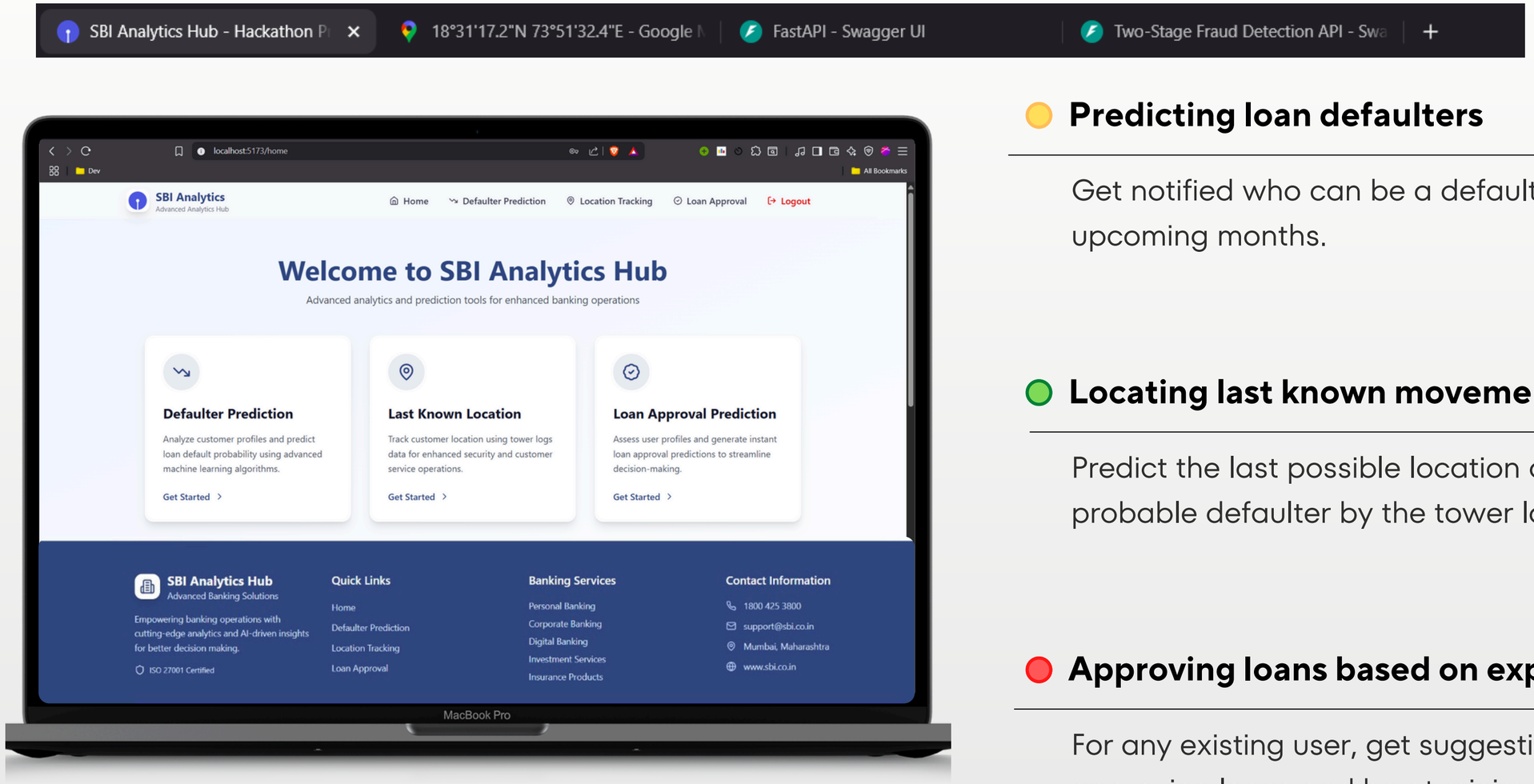
Detect Fraudulent Activity:
Predict fraudulent accounts/transactions using historical data
Ensure model explainability and resilience to adversarial mimicry

02

Infer Defaulters' Last Location:

- Predict final tower/region using device/tower logs
- Visualize and justify prediction

Introducing Our Prototype



Predicting loan defaulters

Get notified who can be a defaulter in upcoming months.

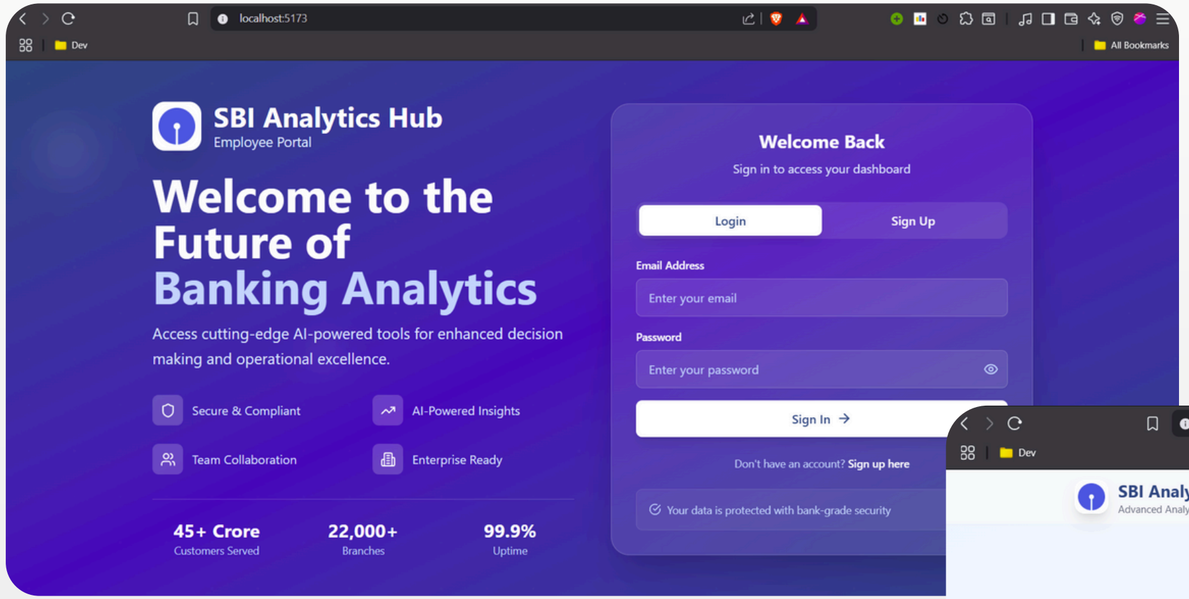
Locating last known movement of defaulters

Predict the last possible location of the defaulter/ probable defaulter by the tower logs of the user

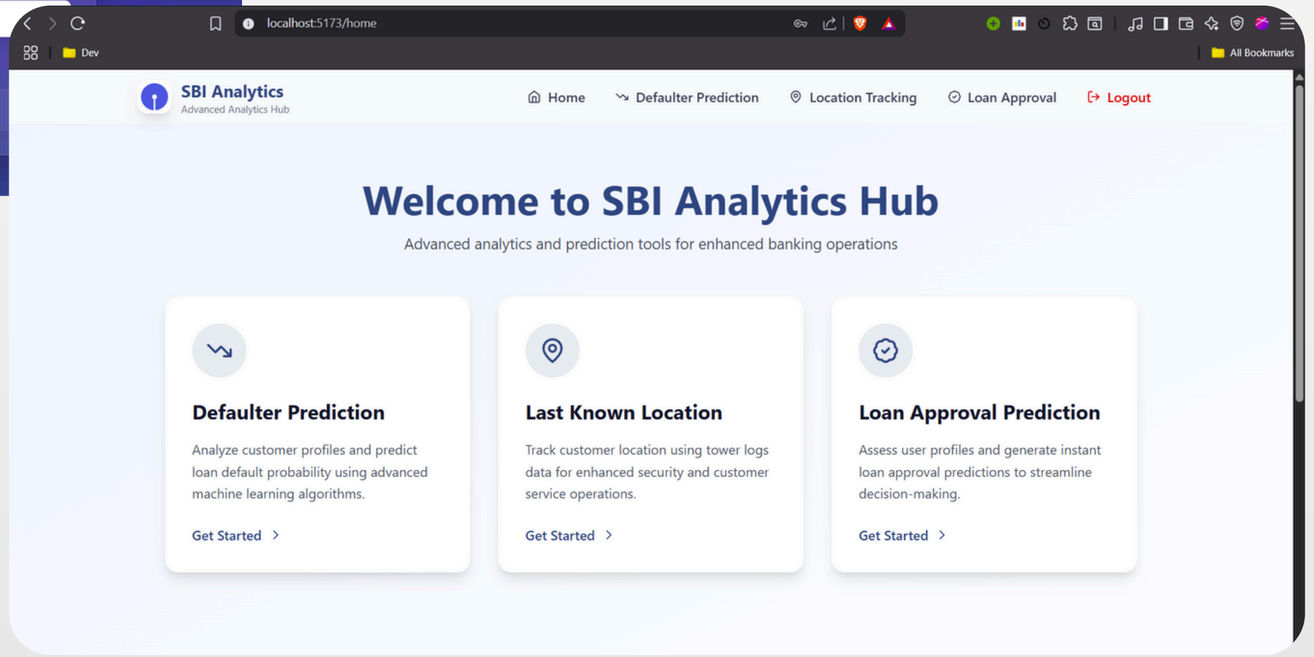
Approving loans based on explainable AI

For any existing user, get suggestion from AI for approving loans and be at minimum risk

User Flow / System Access



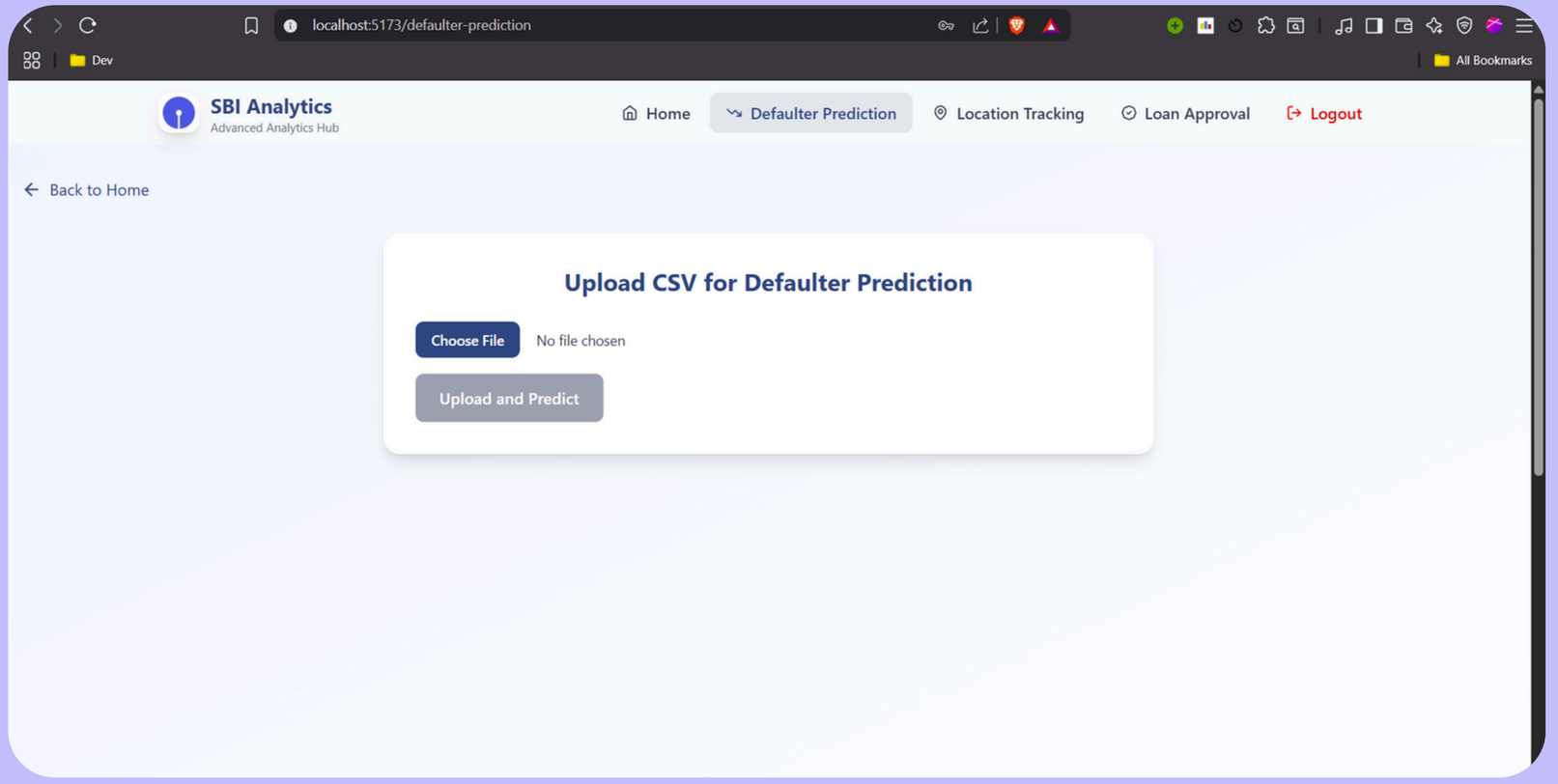
Login/Signup



Module 1

Defaulter Prediction

ASKS FOR A .CSV FILE OF USERS AND THEIR BANKING DATA (FEATURES)



Module 1

Defaulter Prediction

0/1 PREDICTIONS

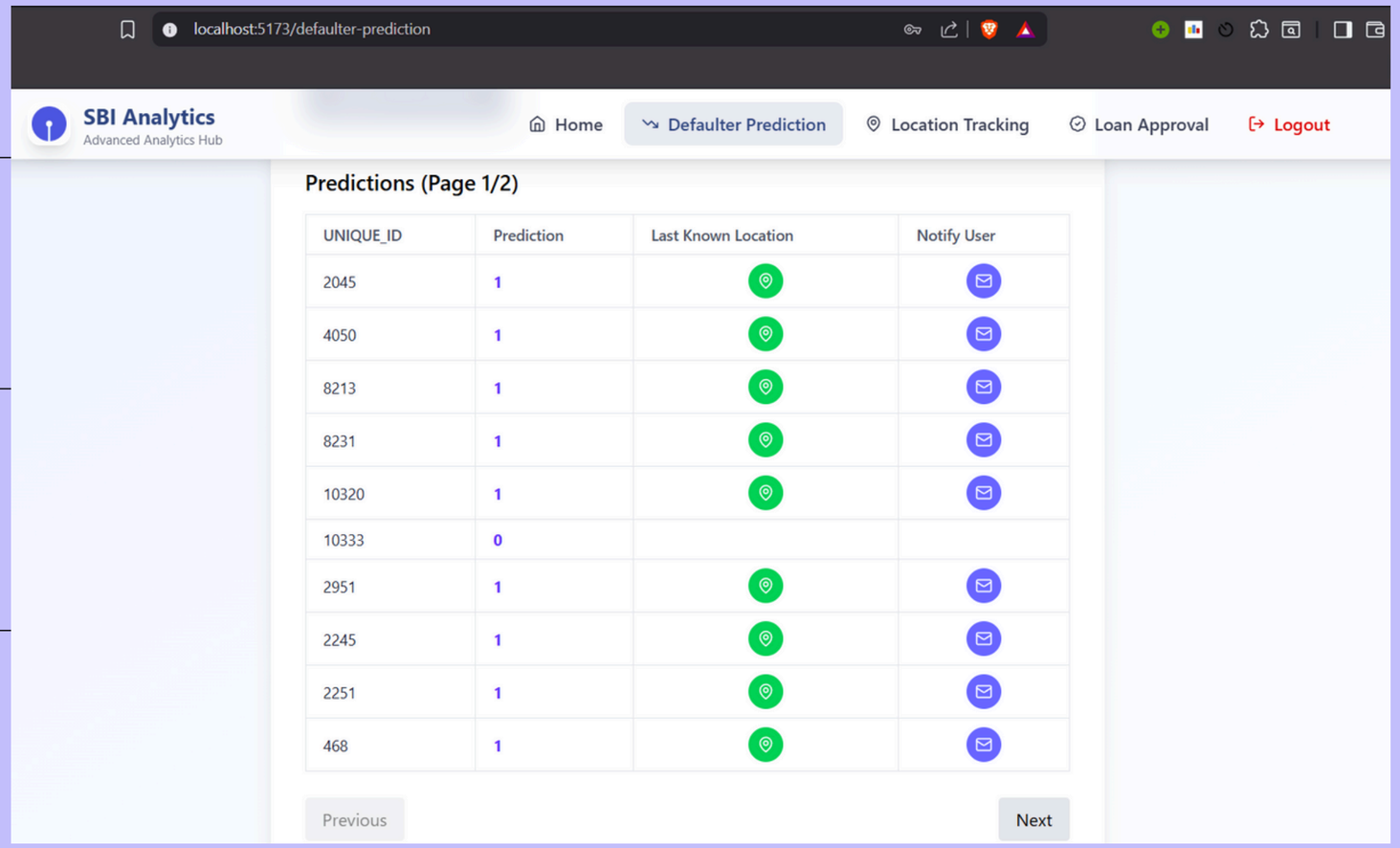
The predictions are generated from an ML API that we have hosted through the backend.

Predicting last known location

The location button takes you to the window that predicts the last known location of the (probable or) defaulter

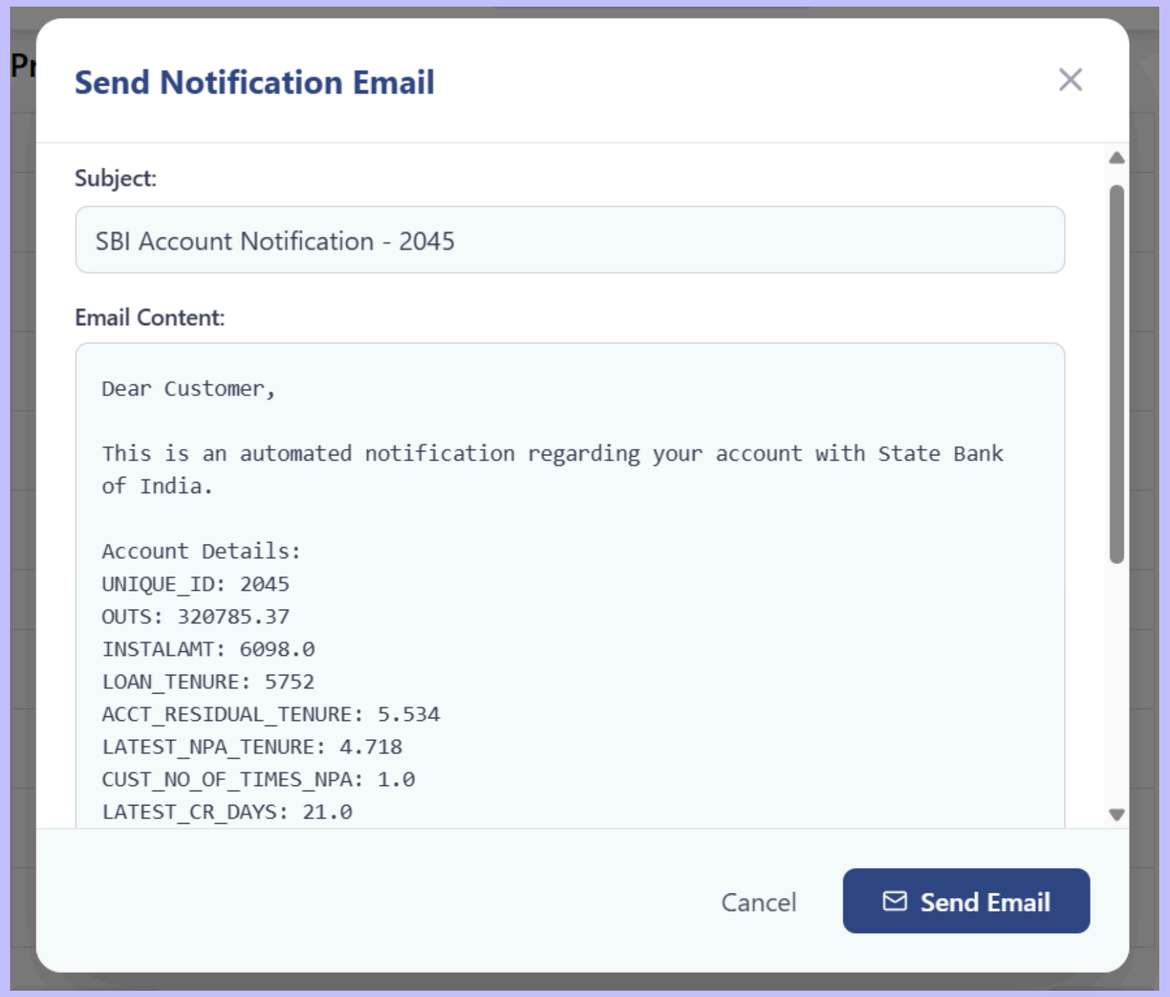
Notify User

If you find that the user has been genuine and might have been flagged defaulter this time, SBI can notify the user via automated mail just by a click.



Module 1

Defaulter Prediction

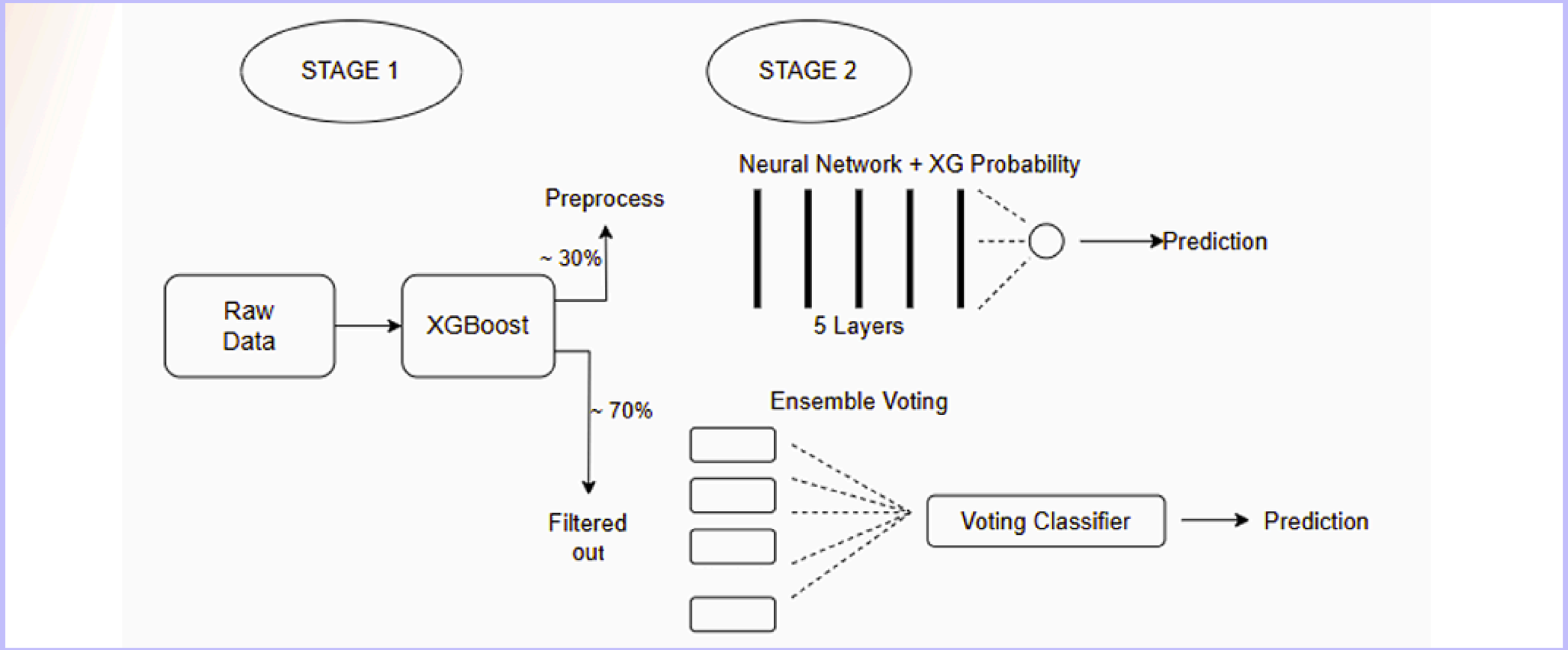


Automated email notification

The email contains important metrices and account details that might be responsible for the delay by the user and just by a click the email will be fetched from the database and will be sent to the user.

Module 1

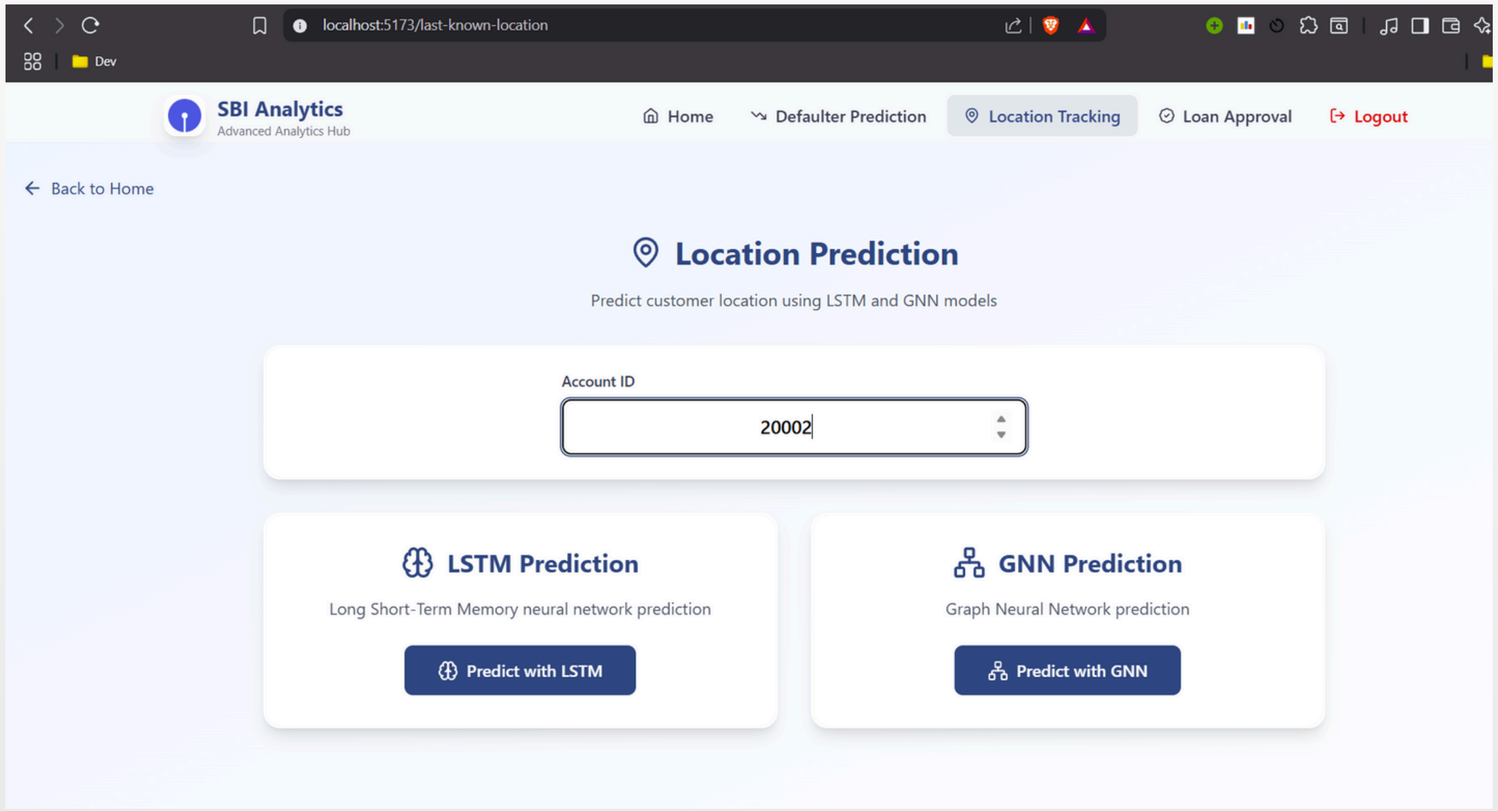
Approach Explaination



Module 2

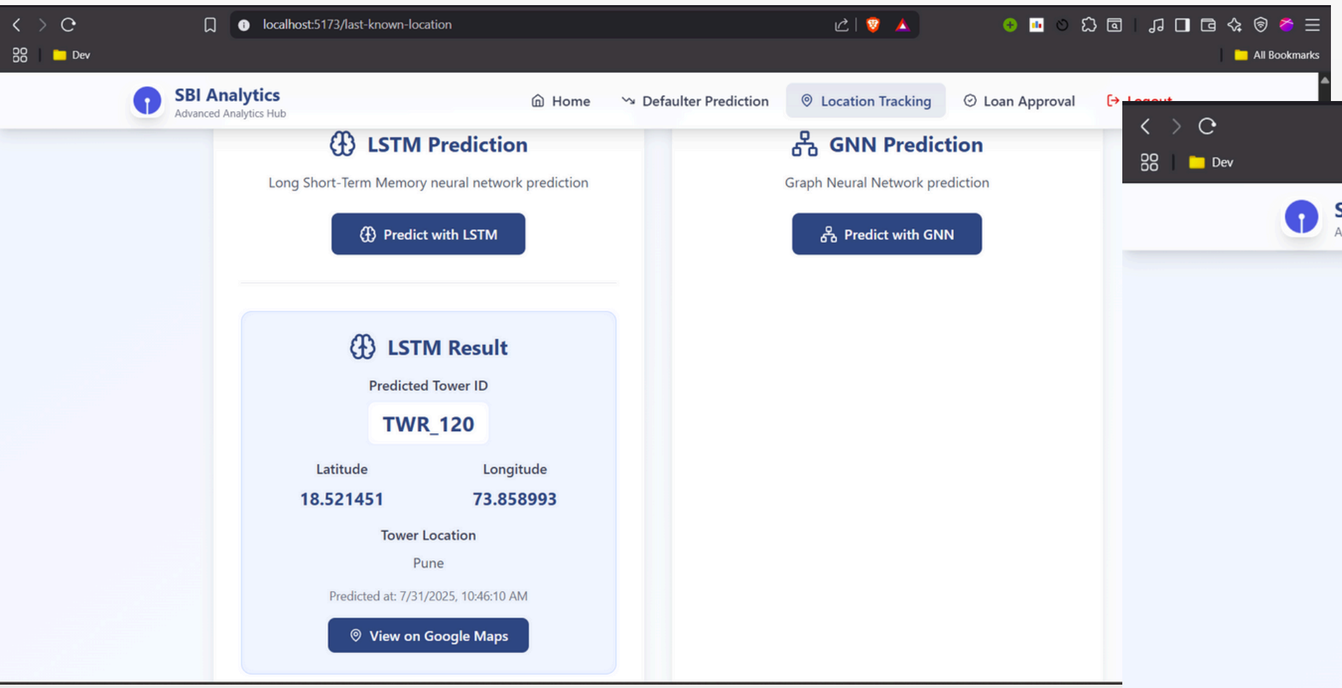
Last Known Location Finder

Asks unique Account Id and then predicts the location via 2 models

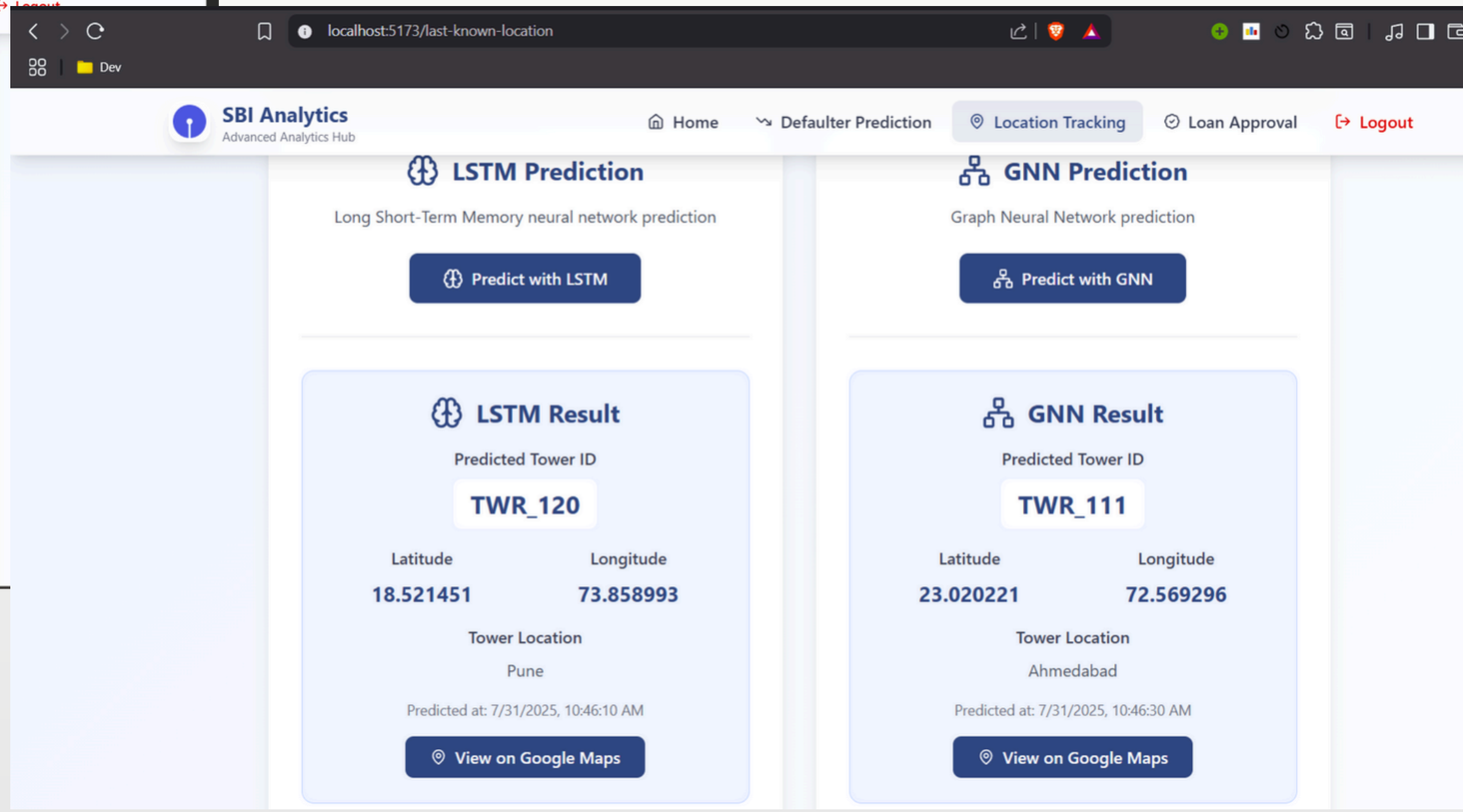


Module 2

Last Known Location Finder



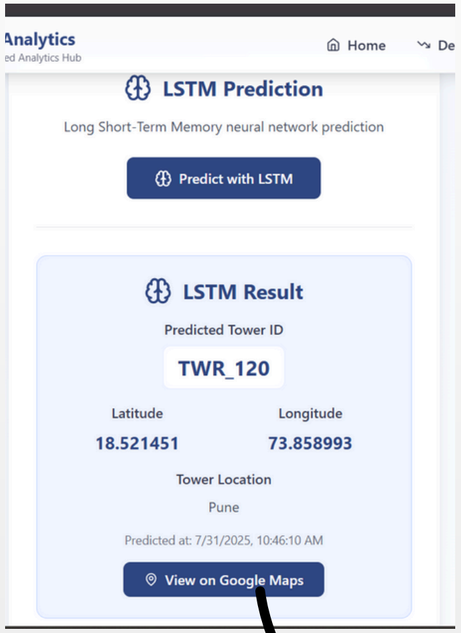
LSTM Prediction



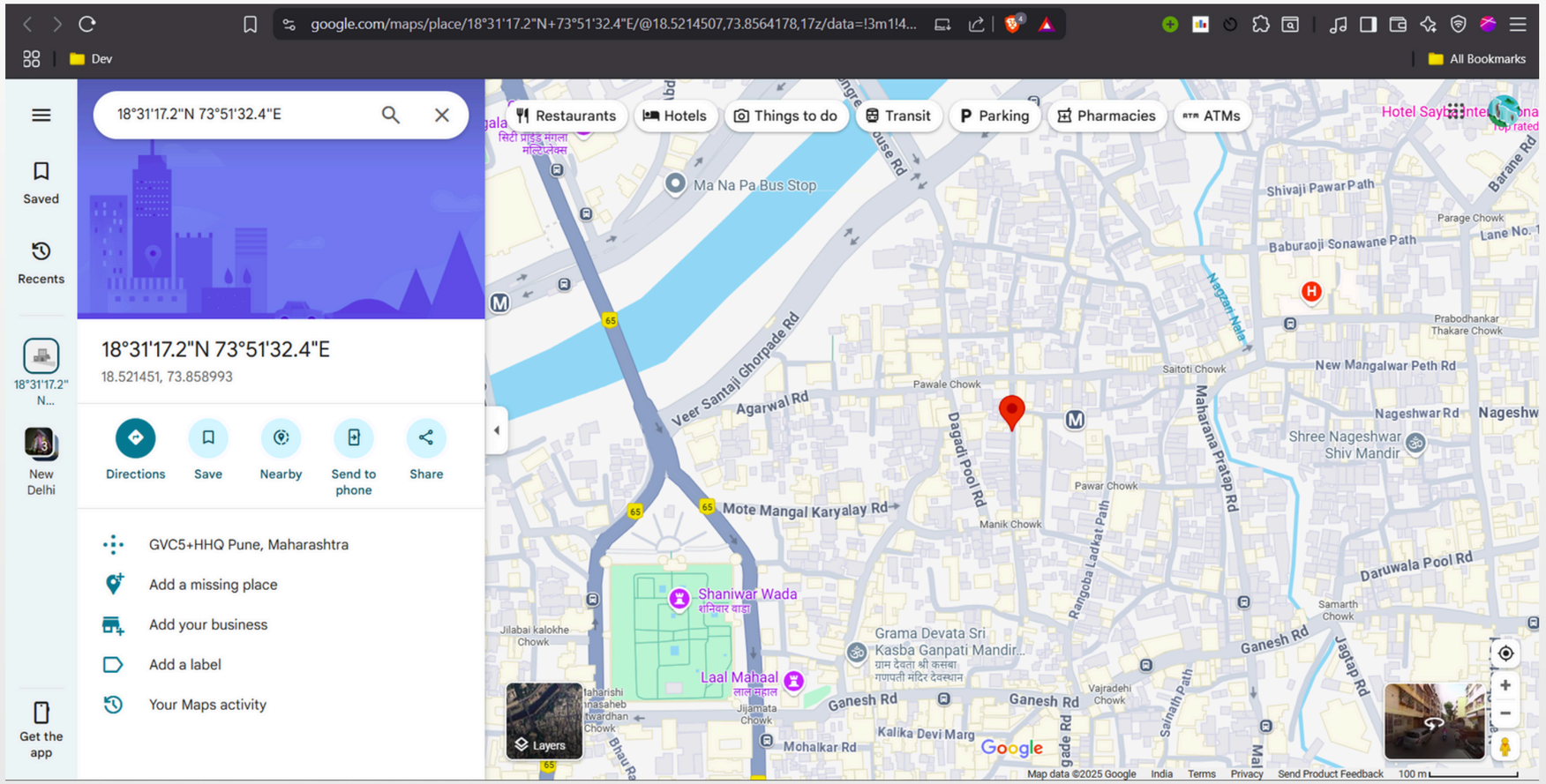
LSTM AND GNN Predictions

Module 2

Last Known Location Finder

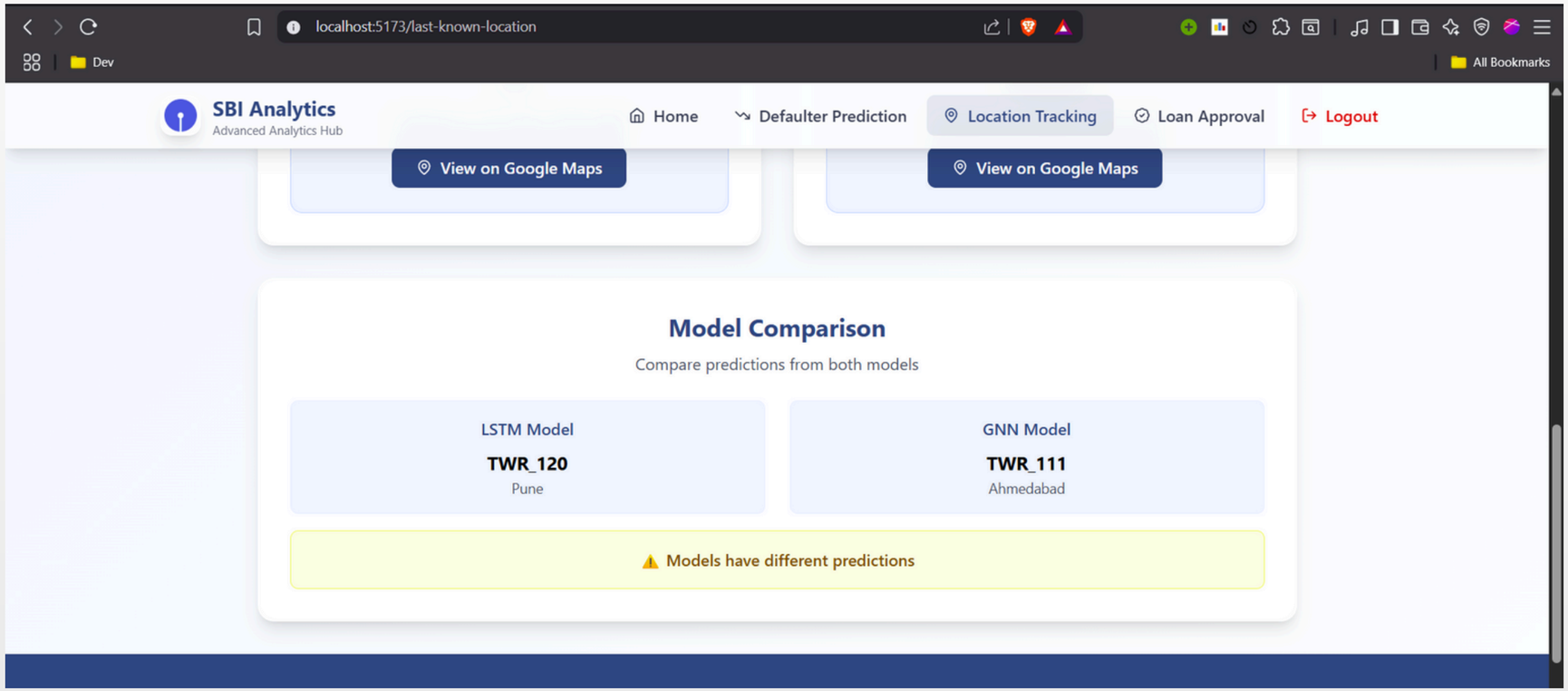


Click



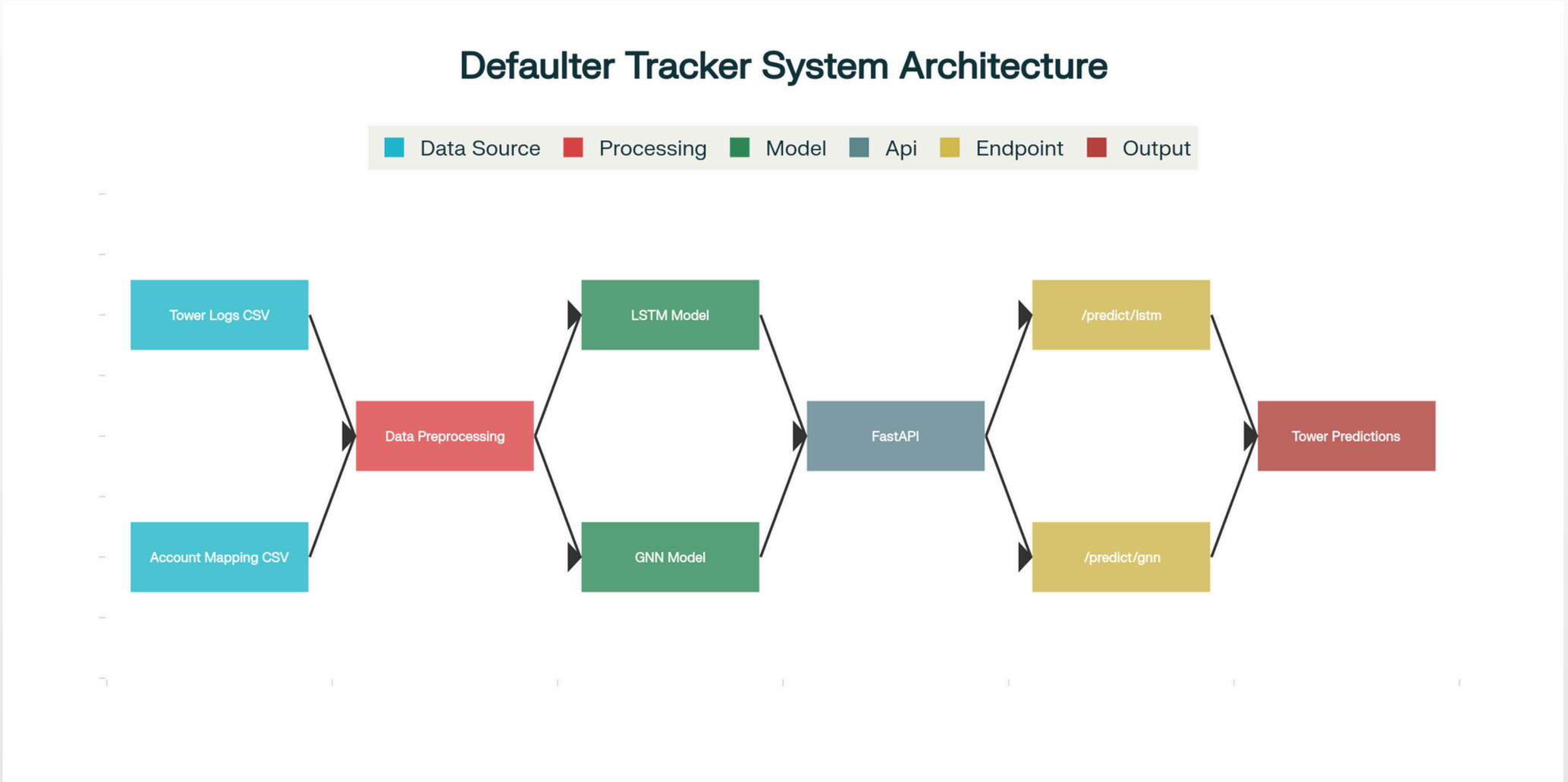
Module 2

Last Known Location Finder



Model Comparisons

Approach Explanation



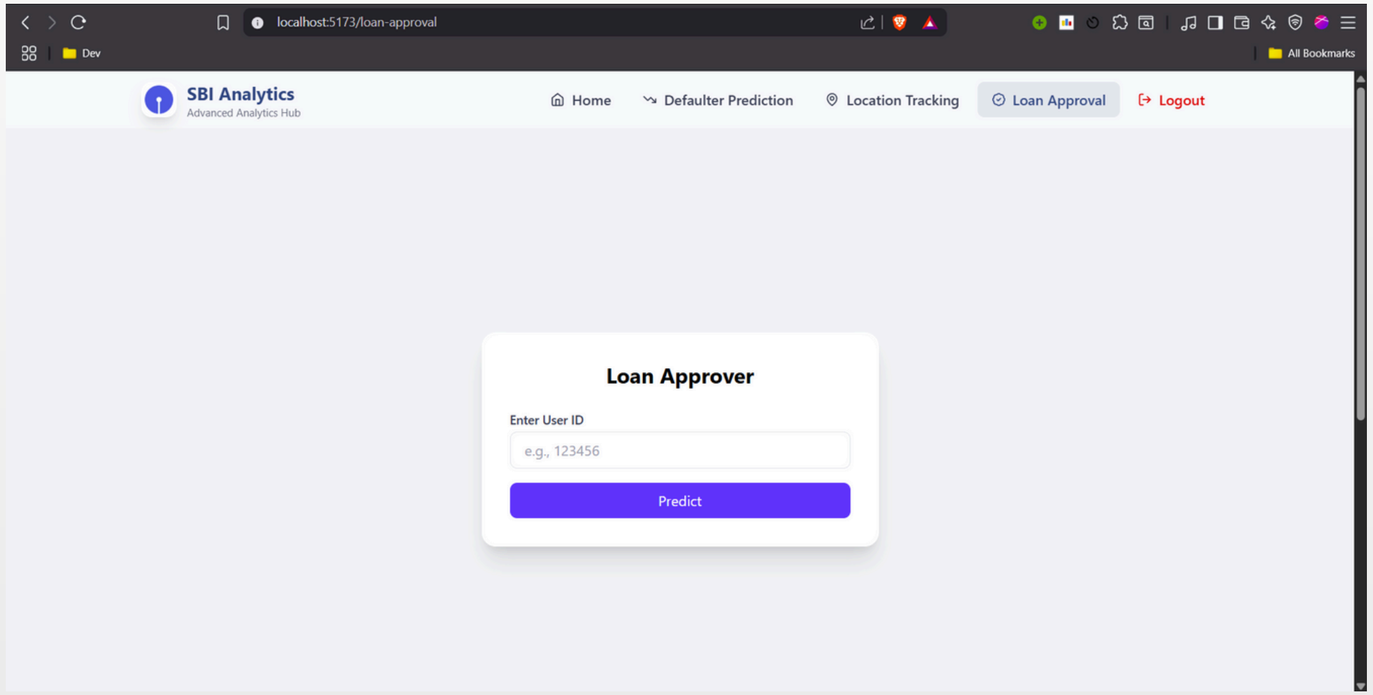
Approach Explanation

LSTM vs GNN Tower Prediction Comparison

Category	LSTM	GNN
Model Type	Recurrent Neural Network	Graph Neural Network
Input Data	Sequential tower IDs (last 5)	Tower coordinates and connections
Pattern Recognition	Temporal patterns	Spatial relationships
Architecture Details	Time-series sequence processing	GraphSAGE with 2 layers
Prediction Method	Next tower in sequence	Graph-based neighbor prediction
Strengths	Captures movement history, temporal dependencies	Spatial awareness, network topology
Use Cases	Route prediction, habitual patterns	Geographic clustering, proximity-based prediction

Module 3

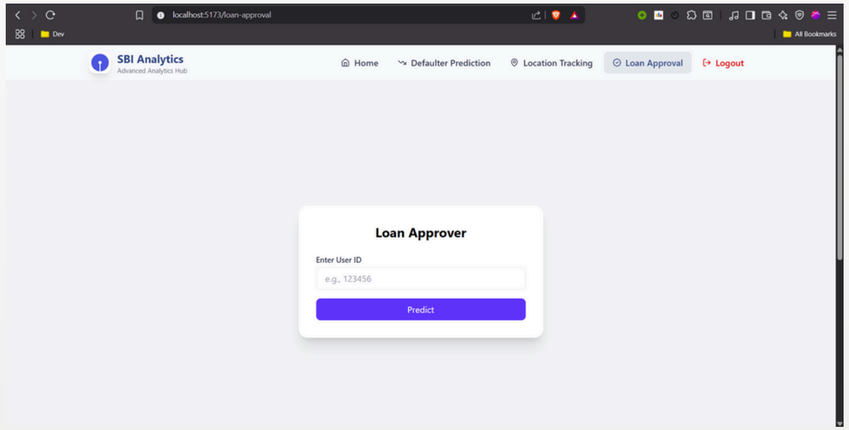
Loan Approval Predictor

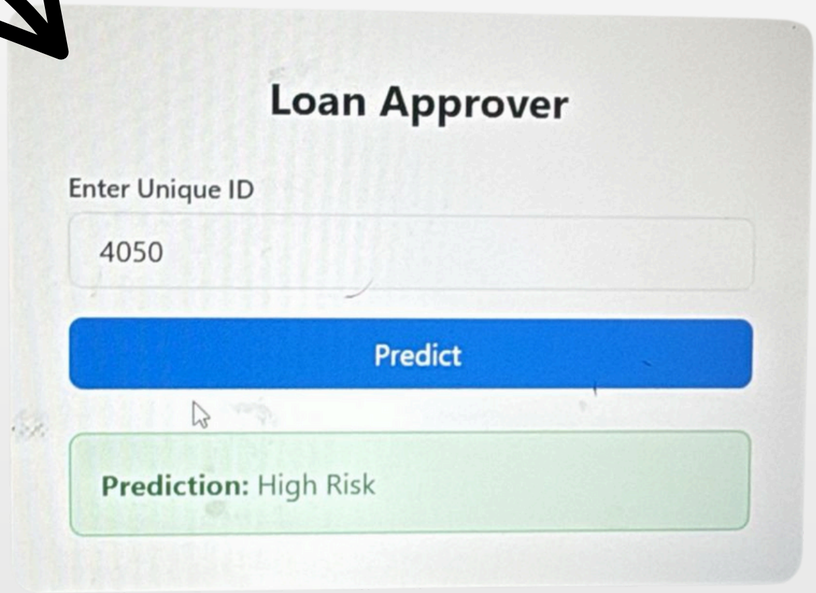


For Transparent Loan Decisioning

Asks for the Unique User Id in the database

Module 3
Loan Approval Predictor

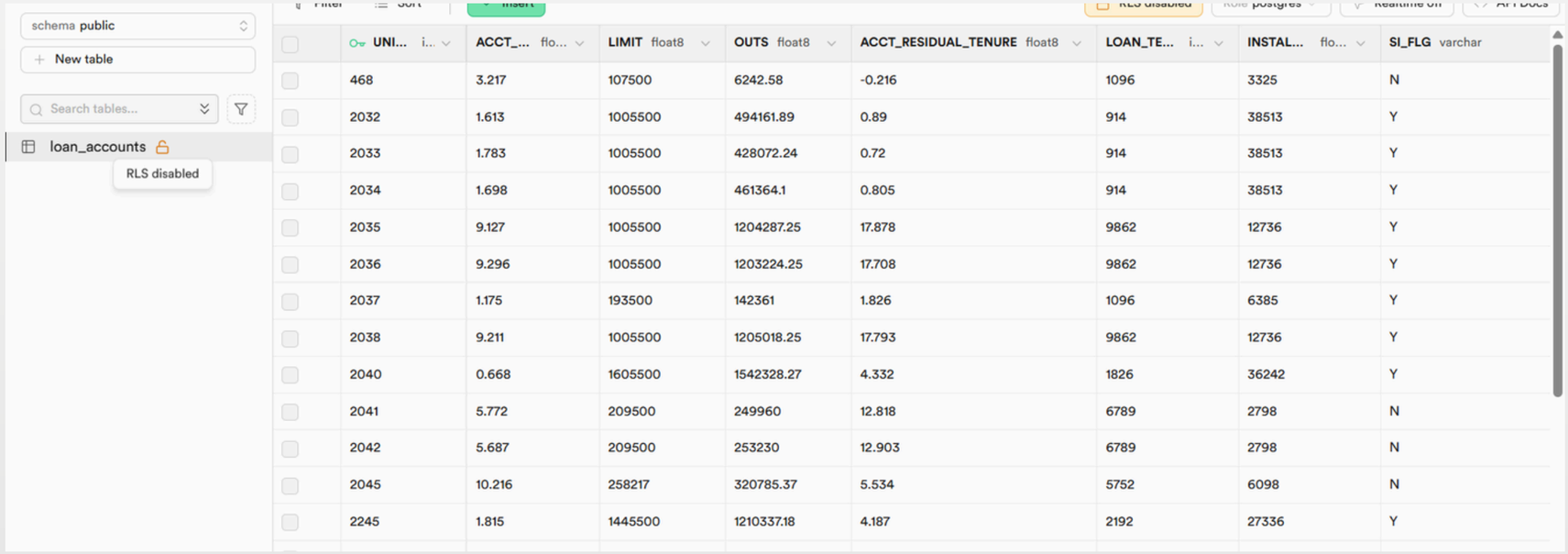




- **Enter Unique ID** → Risk assessment
- **Output**: "Risky"/"Not Risky" with justification
- **Features:** SHAP-based explainability

How it works?

- It works based on an API call to our cloud based Postgre SQL database service called Supabase
- The API can be used for performing CRUD operations on the database.
- Why cloud based Data base? Supabase has all cybersecurity features inbuilt to prevent hackers from accessing the database.(The recent tea app breach has exposed the vulnerabilities in firebase)
- Supabase has the option to directly upload our csv files in the database via their webpage.
- Our API returns target and Unique id as response in JSON format
- Based on the target value one can get the defaulter/non defaulter values without having to rerun the ML model api everytime.



UNI... i...	ACCT_... flo...	LIMIT float8	OUTS float8	ACCT_RESIDUAL_TENURE float8	LOAN_TE... i...	INSTAL... flo...	SI_FLG varchar
468	3.217	107500	6242.58	-0.216	1096	3325	N
2032	1.613	1005500	494161.89	0.89	914	38513	Y
2033	1.783	1005500	428072.24	0.72	914	38513	Y
2034	1.698	1005500	461364.1	0.805	914	38513	Y
2035	9.127	1005500	1204287.25	17.878	9862	12736	Y
2036	9.296	1005500	1203224.25	17.708	9862	12736	Y
2037	1.175	193500	142361	1.826	1096	6385	Y
2038	9.211	1005500	1205018.25	17.793	9862	12736	Y
2040	0.668	1605500	1542328.27	4.332	1826	36242	Y
2041	5.772	209500	249960	12.818	6789	2798	N
2042	5.687	209500	253230	12.903	6789	2798	N
2045	10.216	258217	320785.37	5.534	5752	6098	N
2245	1.815	1445500	1210337.18	4.187	2192	27336	Y

Tech Stack Used

- **Frontend:** React.js, TypeScript, Tailwind CSS, Vite
- **Backend:** FastAPI (Python)
- **Database:** Supabase (PostgreSQL)
- **Machine Learning Models:**
  XGBoost, Random Forest, Logistic Regression
  LSTM – For sequential modeling of user movement
  GNN (Graph Neural Network) – For mobility graph-based defaulter tracking
- **Data Processing & Preprocessing:**
  pandas, numpy, scikit-learn, KNNImputer, StandardScaler, imblearn (SMOTE)
- **Visualization:** matplotlib, seaborn

Privacy and Security

Implemented:

- Login with username and hashed passwords to prevent unauthorised access.
- Cloud based database which can prevent certain IP addreses from accesing the database.
- Route Protection to prevent unauthorised access.
- Database protected with a password.
- Immune to SQL injections.

Future updates:

- API rate limiting to prevent DDOS attacks.
- Zero trust Authentication to prevent AI based malware tools.
- Two Factor Authentication.

What this means for SBI?

- Cloud based DB allows a huge volume of data to be stored without performance issues on host device.
- Any authorised SBI employee can use this app at any location to track defaulters.
- Ability to manually change status of a customer in case of errors.
- Accurate predictions of potential defaulters with minimal false positives.
- Ability to track defaulters for better loan recovery.
- Instantly get reasons for why a customer is flagged as defaulter.
- FastAPI + Supabase ensures a scalable, secure, and cloud-compatible infrastructure for future deployments.

Future scope

- Dynamically improve the training model as new customer data is added and fine tune it for location based trends in defaulters.
- Develop a companion app to help recovery agents view predicted defaulter locations and activity patterns.
- Adapt the system for use in insurance fraud, credit card anomaly detection, and rural loan defaulter analytics.
- Incorporate transformer-based models and spatio-temporal GNNs for improved accuracy and generalization.
- Customer Behaviour Analytics dashboard.

Thank You

We Are Open To Questions Now