॥वसुधैव कुटुम्बकम्॥

SYMBIOSIS

INSTITUTE OF TECHNOLOGY, PUNE

Artificial Intelligence for Banking & Finance

User Transaction Forecasting

CA2 - Coding Based Assignment

Group Members

Kanan Bedi	21070126044
Aamya Bansal	21070126004

Prepared For:

Dr. Anupkumar Bongale

Table of Content

Statement of Need

Accurate transaction forecasting is essential for organisations to optimise their operations in today's competitive business environment. Financial planning, inventory control, resource allocation, and improving customer satisfaction all depend on accurate forecasting. Nevertheless, the intricacy of transaction data influenced by a number of variables, including seasonality, industry trends, and customer behaviour, is frequently missed by conventional forecasting techniques.

Precise transaction forecasting enables companies to:

<u>Improve Inventory Management</u>: By projecting future demand, you may avoid stockouts and overstock scenarios.

<u>Improve Financial Planning</u>: Make appropriate use of available resources and control cash flow.

<u>Enhance Customer Experience</u>: To boost satisfaction and loyalty, customise services depending on transaction trends.

Despite technological developments, a lot of businesses continue to use antiquated forecasting techniques, which lead to inefficiencies and lost opportunities. The goal of the Client Transaction Forecasting project is to use Long Short-Term Memory (LSTM) networks to create an advanced forecasting model. With this strategy, organisations will be able to analyse transaction data from the past more efficiently, identifying long-term linkages and increasing prediction accuracy.

Businesses may gain a competitive edge, react proactively to market swings, and promote growth in a dynamic environment by putting this advanced forecasting solution into practice.

Technical Functionality

Based on previous data, the Client Transaction Forecasting project forecasts future transaction volumes using sophisticated machine learning techniques, particularly Long Short-Term Memory (LSTM) networks. Among the technological features are:

Preprocessing of the data involves cleaning, normalising, and transforming raw transaction data into a format that is appropriate for training models. This procedure involves encoding category variables, scaling numerical characteristics, and handling missing values.

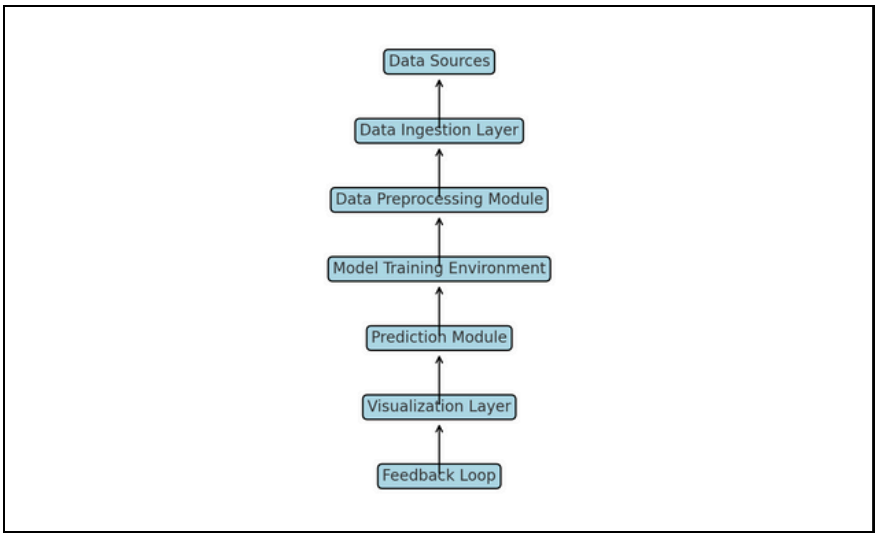
Feature engineering: From the transaction data, the system extracts pertinent characteristics including previous sales numbers, time-based indicators (day of the week, month, etc.), and seasonal trends. This facilitates the efficient capturing of underlying patterns by the LSTM model.

Model Training: Using the prepared dataset, the LSTM model is trained to identify patterns and generate precise predictions by taking use of the transaction data's sequential structure.

Forecast Generation: Following training, the model produces projections of future transaction volumes that may be examined and used to inform strategic choices.

Performance Evaluation: To guarantee accurate predictions for end users, the accuracy of the model is evaluated using a variety of metrics (e.g., Mean Absolute Error, Root Mean Squared Error).

User Interface Integration: To facilitate quick access and analysis, the forecasting findings can be presented through an intuitive user interface or connected into currently used business applications.



Architecture

The Client Transaction Forecasting system's architecture is made to support effective model training, forecast creation, and data processing. The following are the elements that make up the architecture:

1. Data Sources: Transaction Data: Databases, APIs, and CSV files are some of the sources from which historical transaction data is gathered.
 - External Data: Additional datasets (e.g., economic indicators, seasonal trends) may be integrated to enhance forecasting accuracy.

2. Data Ingestion Layer: This layer is responsible for extracting and loading data from the various sources into the system. It ensures data is available for preprocessing.

3. Data Preprocessing Module: This module cleans, normalizes, and transforms the raw data into a format suitable for analysis. It includes steps such as handling missing values, scaling, and feature engineering.

4. LSTM Model:
Model Training Environment The LSTM model, which handles sequential data, is the central component of the architecture. To identify patterns and trends, this model is trained using transaction data from the past.
Pipeline for Training: During the training phase, the data is divided into training and validation sets, the hyperparameters are optimised, and the model's performance is assessed.

5. Prediction Module: The model is used to produce transaction projections based on fresh or untested data once it has been trained. The forecasts can be produced in a number of ways, including tables and graphs.

6. Visualisation Layer: The forecasts and insights obtained from the model are presented in an intuitive interface (such as a dashboard or web application). Users are able to visualise trends, engage with the data, and make defensible judgements.

7. Feedback Loop: To continuously develop the model, a feedback mechanism is included in the architecture. New transaction data is added back into the training pipeline as it becomes available, enabling regular retraining and improved model accuracy.

Because of its modular architecture, which guarantees scalability, flexibility, and maintainability, businesses may better anticipate future needs and adjust to shifting transaction patterns.

Usage/Scope

Financial institutions and companies whose operations depend on precise transaction forecasts are the target market for the Client Transaction Forecasting system. Finance teams, marketing departments, and operational managers can use this technology to better manage inventories, forecast cash flow trends, and customise client engagement tactics. It helps businesses use past transaction data for strategic planning across a range of industries, such as retail, banking, and e-commerce. In order to improve decision-making processes, the scope includes projections based on transaction patterns, insights into seasonal fluctuations, and the identification of possible market possibilities.

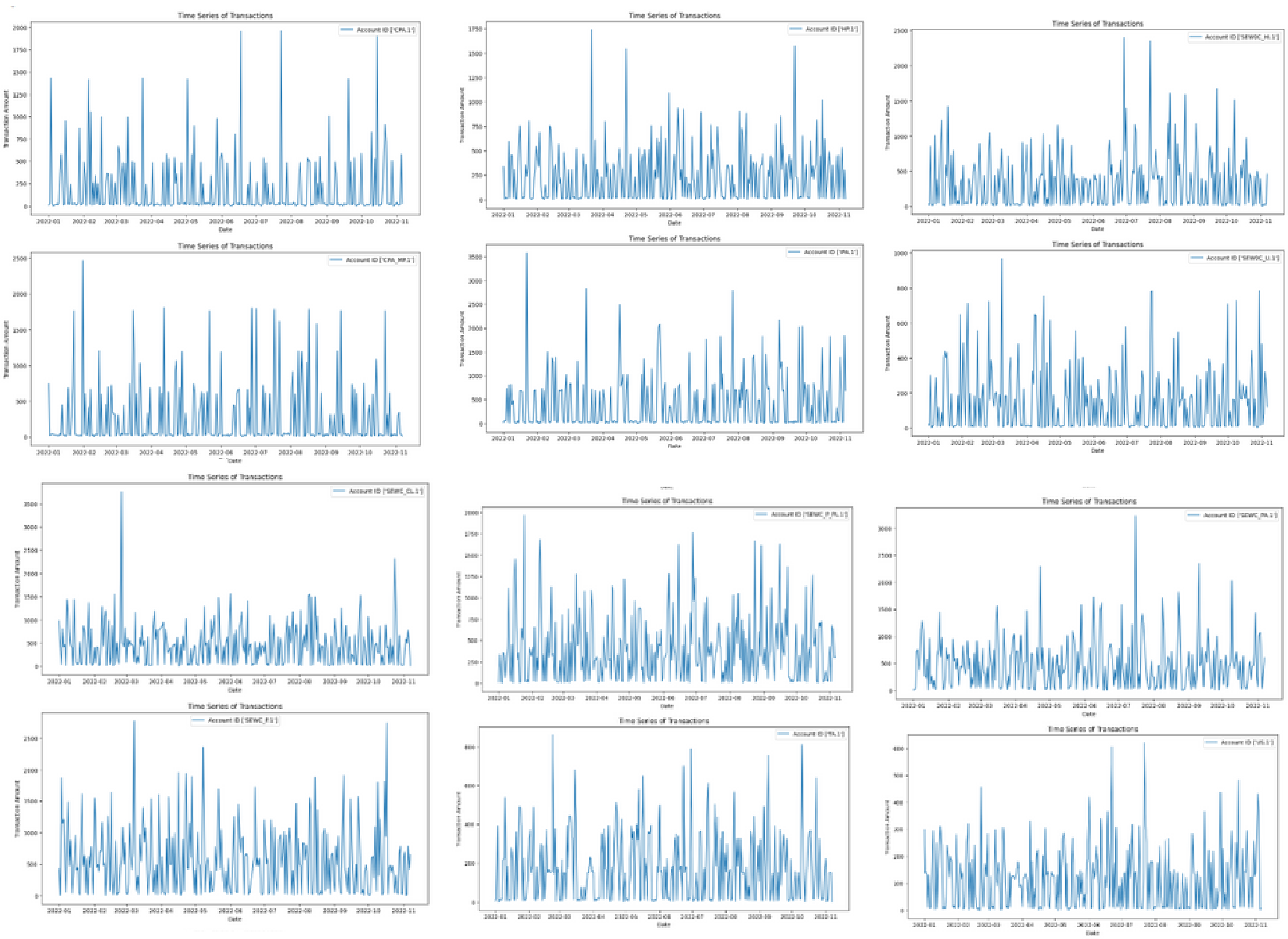
Impact/Overview

Businesses may now properly forecast future transaction volumes thanks to the Client Transaction Forecasting initiative. Businesses may increase profitability, reduce financial risk, and manage their resources more effectively by utilising cutting-edge machine learning technology. Because of the system's predictive capabilities, stakeholders may make well-informed decisions and react quickly to changes in the market and client expectations. In the end, this initiative promotes sustainable growth in a cutthroat market by enhancing operational effectiveness, customer happiness, and long-term financial stability.

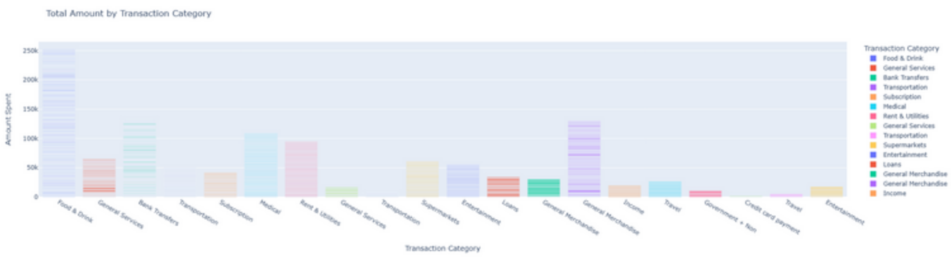
Github Repo Link

https://github.com/Kanan-Bedi/User-Transaction-Forecasting

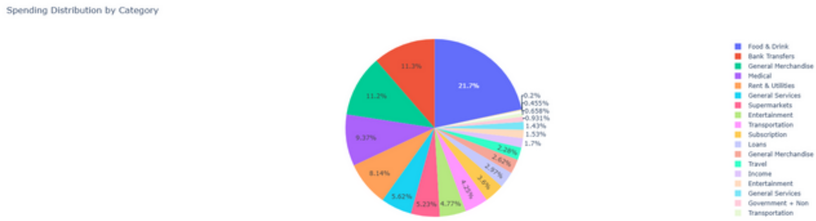
Graphs & Analysis



The graphs display the details of transaction amounts with reference to time variations representing periods of raised or lowered activity. All of the above mentioned graphs clearly depict peaks indicating periods of high volumes of transactions or a sudden spurt activity and sometimes a plateau or no activity for a certain period of time. These spikes are consistently observed to be irregular which mean that SM is periodic and not regular since the transaction activity is triggered by certain event(s). The variations can be seen in that some zones may have higher peak points than other zones for certain cycles of transactions. In general, all the analyzed data depict a situation of irregular and unpredictable transaction activity, although there are temporary peaks of activity.



Regional spending is illustrated through the bar chart below. “Food & Drink” tops the list concerning spending, spending at over 250, 000, followed by “Bank Transfers”, Medical and “General Merchandise” spending above 100,000. Others like Rent & Utilities, Travel, Credit Card Payment has a way smaller amount. However, it shows that food, services & medical are important categories most of the spending is in that category while entertainment and other are comparatively low.



It means the pie chart is explaining how much money is being spent on what kind of things. The most frequently occurring category is “Food & Drink” with 21.7%, and other categories are “Bank Transfers” with 11.3%, “General Merchandise,” with 11.2%, and “Medical” with 9.37%. Other categories that make up the remaining part of the total bill are somewhat smaller: “Supermarkets” is below 6%, “Transportation,” and “Subscriptions” too. More of it is spent on food expenses so many portions are left for other expenses.

```
Shape
(3727, 4)
Types
<pandas.io.formats.style.Styler object at 0x7ac6b86d8340>
Info
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 3727 entries, 2022-01-01 to 2022-11-06
Data columns (total 4 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Id_compte             3727 non-null   object
 1   amount_transaction    3727 non-null   float64
 2   category_transaction  3727 non-null   object
 3   Id_compte_encoded     3727 non-null   int64
dtypes: float64(1), int64(1), object(2)
memory usage: 145.6+ KB
None
Head & Tail
```

Head

Date	Id_compte	amount_transaction	category_transaction	Id_compte_encoded
2022-01-01 00:00:00	CPA.1CPA.1	10.340000	Bank TransfersBank Transfers	0
2022-01-02 00:00:00	CPA.1CPA.1CPA.1CPA.1CPA.1	31.490000	Bank TransfersBank TransfersBank TransfersBank TransfersBank Transfers	0
2022-01-03 00:00:00	CPA.1CPA.1CPA.1CPA.1CPA.1CPA.1	1430.340000	Bank TransfersBank TransfersBank TransfersBank TransfersBank TransfersGeneral Merchandise	0
2022-01-04 00:00:00	CPA.1CPA.1CPA.1CPA.1CPA.1CPA.1	32.900000	Bank TransfersBank TransfersBank TransfersBank TransfersBank TransfersBank Transfers	0
2022-01-05 00:00:00	0	0.000000	0	0

Tail

Date	Id_compte	amount_transaction	category_transaction	Id_compte_encoded
2022-11-02 00:00:00	US.1US.1US.1US.1US.1US.1US.1US.1US.1US.1	301.880000	Food & DrinkFood & DrinkBank TransfersTransportationBank TransfersFood & DrinkRent & UtilitiesTransportationBank TransfersTransportation	11
2022-11-03 00:00:00	US.1US.1US.1US.1US.1US.1US.1	431.920000	Bank TransfersBank TransfersFood & DrinkRent & UtilitiesBank TransfersFood & DrinkBank Transfers	11
2022-11-04 00:00:00	US.1US.1US.1US.1US.1US.1US.1US.1US.1	368.580000	Bank TransfersTransportationFood & DrinkBank TransfersBank TransfersTransportationFood & DrinkFood & DrinkGeneral Services	11
2022-11-05 00:00:00	US.1US.1	3.920000	Bank TransfersBank Transfers	11
2022-11-06 00:00:00	US.1US.1US.1US.1	5.380000	TransportationBank TransfersTransportationBank Transfers	11

NA Values

Number of NA Values

	Column	Number of NA
0	Id_compte	0
1	amount_transaction	0
2	category_transaction	0
3	Id_compte_encoded	0

This EDA presents a dataset of 3727 entries with transaction information of the 11 months of the year 2022. These are account ID, transaction amount, category and encoded account ID. No example of missing values can be seen in the given data sets. These are the first and last few transactions viewed for each given category; there are others such as Bank Transfers and Food & Drink with corresponding amount.

Conclusion

The Client Transaction Forecasting project showcases how cutting-edge machine learning methods can be applied to improve clients' decision-making processes. Businesses can optimise their operations and increase the accuracy of financial forecasts by utilising the system's ability to predict future transaction trends, which is achieved through the utilisation of a Long Short-Term Memory (LSTM) model and past transaction data. The solution's architecture is efficient and scalable, allowing it to be tailored to a range of business requirements. Because of this, this project not only helps clients with their current forecasting needs, but it also lays the groundwork for future improvements like adding real-time data integration and enhancing the model's functionality. In the end, this program gives clients the ability to lead growth, make wise judgements, and keep a competitive advantage in their particular industries.