

Mini Project Report On

**Expense Tracker** 

Submitted by Group id - 04

#### **Project Members**

KRISHNA HITNALIKAR (1032221979) RONAK PATIDAR (1032221964) HARSH KATIYAR (1032221852) AMAN KANTHALIA (1032221904)

> Under the Guidance of Prof. Prasad Purnaye Prof. Maithilee Patawar

School of Computer Engineering and Technology MIT World Peace University, Kothrud, Pune 411 038, Maharashtra - India 2024-2025

#### **Abstract**

The design and deployment of the long short-term memory neural networks are explored, which fall in the category of recurrent neural networks (RNNs) for general time-series forecasting. The system will be used to help people manage their budgets by predicting daily and weekly expenses from various categories such as food, transportation, entertainment, and other items that users might need.

In this day and age, with speed dominating society and consumption characterizing an era, it is tough for human beings to exercise financial discipline over their proper spending. Many people spend their savings too early, mostly because of unintended spending or lack of knowledge about where they spend their money.

The system essentially looks at historical financial data and makes future predictions about spending by individuals in real time. It provides users with day-to-day and weekly predictions of money to be spent and personalizes insights toward future expenses so that an individual can adjust his or her financial behavior proactively.

This approach of predicting will be placing individuals better in avoiding over-spending hence remain within their monthly budgeting and, therefore eventually maintain long-term financial stability.

This is even more applicable with the use of an LSTM because it will reflect not only the short-term fluctuations but also the long-term trends in spending behavior. It is, therefore, well suited to such a prediction task. Its recurrent structure enables it to keep track of relevant information from previous days' expenditure and generate predictions that take into account not only the patterns of recent days but also the dependency that exists between different types of spending.

It details all the processes involved in the lifecycle of an expense prediction system, starting with problem-domain analysis, where objectives were formulated. It also provides a comprehensive literature survey of the existing techniques in financial forecasting and enumerates the weaknesses of traditional methods, thereby pointing out the advantage of using machine learning models, especially LSTM networks. Such a system discusses technical requirements and design choices related to a model of data preprocessing, model architecture, training methodology, and evaluation metrics.

This system shall bridge this gap between these conventional budgeting techniques and modern technological advances, helping the user provide an intelligent tool through which he can manage his finances more efficiently. The prediction nature of the system helps the users approach their budgeting in a more proactive manner, thus providing a sense of financial discipline that offers great advantages in decisions made about daily expenditures.

With the advancement of machine learning, this path can be considered one of the possible further developments in using advanced neural networks like LSTM for personal finance management. This report illustrates some practical applications of such systems and demonstrates how artificial intelligence can apply to everyday financial challenges faced by people around the world.

### **Contents**

Abatroat	 T
ADSTRACT	 

	T		<u> </u>	
1	INTRODUCTION			
		1.1	Challenges in Personal Financial Management	
		1.2	Budgetary Limits and Financial Control	
		1.3	Expense Prediction System Using LSTM Networks	
2	SYATEM ARCHITECTURE			
		2.1	Architecture Figure	
3	LITERATURE SURVEY			
4	PROBLEM REQUIREMENTS			
		4.1	Project Scope	
		4.2	Project Aims	
		5.1	Screenshots	
6	NOVALTY			
7	CONCLUSION			
8	REFERENCES			

## Chapter 1: Introduction

#### 1.1 Challenges in Personal Financial Management

In today's fast-paced digital world, personal financial management has become increasingly challenging, particularly with the influx of large volumes of financial data. The growing complexity of financial transactions and consumption patterns results in significant difficulties for individuals aiming to maintain financial stability without incurring excessive debt. Big data analysis can play a crucial role in tackling these challenges by providing individuals with tools to better manage their financial resources.

With vast amounts of data generated through credit card transactions, bank accounts, digital payments, and online purchases, managing finances manually becomes overwhelming. Big data technologies help to analyze this wealth of data in real-time, offering deeper insights into spending behavior and patterns. However, people still face challenges such as irregular sources of income, unanticipated expenses, and the ever-changing financial decisions they must make. By analyzing these patterns and trends, big data solutions can provide actionable insights to address the unpredictability of personal spending, and help individuals adjust their financial plans accordingly.

#### 1.2 Budgetary Limits and Financial Control

Big data analysis provides powerful tools for monitoring and setting budgetary limits across different financial categories like groceries, transportation, and entertainment. By utilizing advanced analytics, individuals can gain detailed insights into their spending trends, enabling them to set precise budgetary limits based on their historical financial data.

The application of big data in expense prediction systems allows for dynamic adjustments to budgetary limits, based on real-time data feeds from various sources such as financial accounts, transactions, and spending behavior. The system can alert users when they are nearing their budgetary limits, helping them make real-time financial decisions. By utilizing data analytics and machine learning algorithms, big data platforms can identify areas where users are overspending and suggest areas where they could reduce expenses. This enhances financial control by providing an accurate, data-driven breakdown of expenses, empowering users to make smarter financial decisions while maintaining discipline over their finances.

#### 1.3 Expense Prediction System Using LSTM Networks

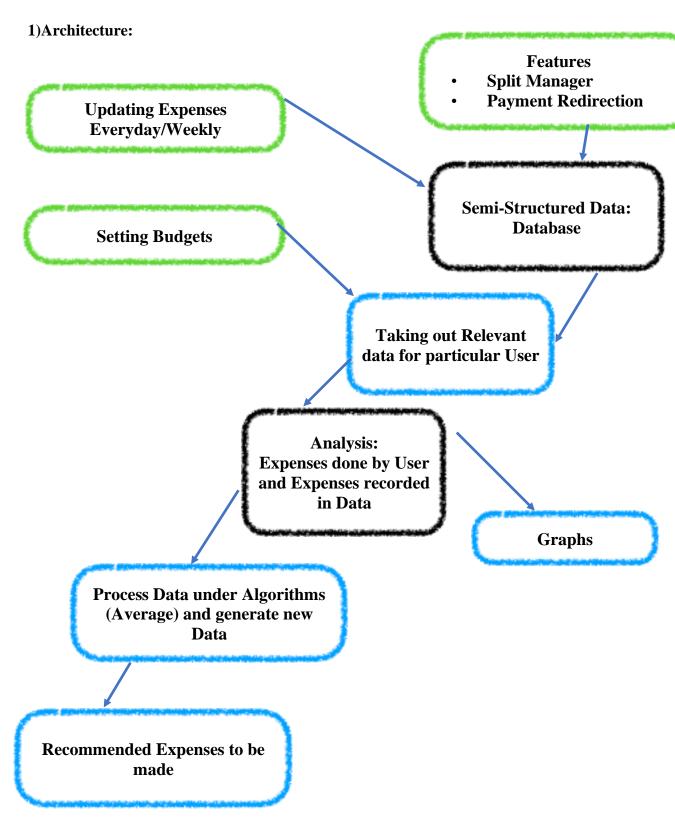
Big data plays a critical role in the development of expense prediction systems that use LSTM (Long Short-Term Memory) neural networks. As modern financial management generates huge volumes of transaction data, analyzing and predicting spending patterns becomes crucial for effective financial control. The combination of big data and LSTM models allows for the efficient processing of vast datasets and the accurate forecasting of individual expenses.

LSTM networks, a form of recurrent neural networks (RNN), are especially well-suited for time-series data, such as financial transactions, as they can capture both short-term and long-term dependencies in spending behavior. By integrating big data, the system can leverage large datasets, including historical spending data from diverse users, to create more robust and accurate expense predictions.

Big data allows the system to analyze multiple variables, such as personal income, spending categories, economic trends, and external factors (e.g., inflation rates) to provide more refined and predictive insights into future expenses. With this, users can better understand how much they will likely spend in the coming days or weeks, and the system can offer specific recommendations for budgeting.

Moreover, big data integration ensures that users' predictions are updated in real-time based on incoming financial data, making the system adaptive to changing spending patterns. This not only prevents overspending but also helps individuals adjust their financial strategies by providing clear insights into their future expenses. By using both LSTM networks and big data analysis, users are empowered with tools that enhance personal financial management and long-term planning.

# Chapter 2: SYSTEM ARCHITECTURE



### Chapter 3: Literature Survey

**Table 1: Literature Survey** 

S.N	1: Literature S	Year of	Algorithm/Methodol	Findings	Research Gaps
0	Title	Publicatio n	ogy	Findings	Research Gaps
1	SNAP	2021	Data Analysis	Effective in tracking expenses	Lacks predictive capabilities
2	Jupiter	2020	Statistical Methods	Efficient for expense management	Does not integrate machine learning or predictive analytics
3	Swindon Simplify	2019	Basic Expense Tracking	Simplifies expense tracking for users	No inclusion of predictive analytics
4	Finery	2022	Data Visualization	Strong visual representati on of expenses	No predictive recommendatio ns provided
5	Rather Go Commands Player	2023	Command-based tracking	Useful for tracking daily expenses	Not integrated with advanced machine learning methods
6	Expense Tracker	2021	Basic Expense Tracking	Effective in categorizing expenses	Lacks predictive analysis features
7	Expenditur e Manageme nt System	2022	Expense Management	Manages daily and monthly expenses	Does not use advanced predictive models

8	Expense	2021	Expense Management	Tracks daily	No predictive
	Tracker			expenses	analytics or
					forecasting
					tools

#### Literature Survey-

The domain of personal finance management has undergone substantial evolution, primarily driven by advancements in data analytics and machine learning (ML). Early systems were primarily focused on recording and categorizing expenses, offering users insights into their historical spending. However, such systems lacked the capability to predict future expenditures. With the development of machine learning models, particularly Long Short-Term Memory (LSTM) networks, there is now potential to forecast personal expenses based on historical spending data. This innovation opens up new possibilities for financial planning, offering users a more proactive approach to managing their budgets.

#### TAKINGS FROM RESEARCH PAPERS-

- [1]. SNAP (2021) was highly effective in tracking real-time expenses but identified a gap in its predictive capabilities, which would enhance its ability to provide users with foresight into future spending behavior.
- [2]. Jupiter (2020), utilizing statistical methods, showed the limitations of relying on such models for predicting future expenses, especially when confronted with non-linear or irregular financial patterns.
- [3]. Swindon Simplify (2019) underscored the importance of ease of use in expense tracking but highlighted a gap in the integration of machine learning techniques for future financial forecasting.
- [4]. Finery (2022) demonstrated the value of strong data visualization in helping users understand their historical expenses. However, the absence of predictive analytics in the system left users without forward-looking insights.
- [5]. Rather Go Commands Player (2023), with its command-based interface, showcased efficiency in daily tracking but lacked the sophistication required for predicting future financial behavior through machine learning models.
- [6]. Expense Tracker: A Smart Approach to Track Daily Expenses (2021) by Nutheti et al. demonstrated efficiency in daily tracking but did not explore predictive capabilities, leaving users reliant on retrospective tracking rather than proactive financial planning.

- [7]. Expenditure Management System (2022) by Gomathy presented a simplified approach to managing finances but did not include advanced predictive modeling, which could have enhanced its utility for users planning future expenditures.
- [8]. Expense Tracker (2021) by Garg et al. focused on categorizing expenses without providing the users with predictive insights, thus limiting its application in long-term financial planning.

## Chapter 4: Problem Requirements

#### 4.1 Project Scope

The scope of the project expands to designing a robust financial management tool capable of predicting daily and weekly expenditures by analyzing vast amounts of historical financial data using Big Data analysis and LSTM neural networks. The system will process large datasets, providing users with insights into their spending patterns, while also assisting in future cost projections.

This project leverages Big Data analysis to cover multiple expenditure categories such as food, transportation, entertainment, and more, allowing users to set budgetary limits for each. The system will handle real-time financial transaction data, which could include thousands or millions of records, ensuring scalability and performance efficiency. By utilizing min-max predictions, the system accounts for variable changes in spending behavior and provides a clear and comprehensive financial outlook for the user.

Python serves as the primary programming environment, utilizing libraries like TensorFlow for LSTM modeling, pandas for large-scale data manipulation, and Matplotlib/Seaborn for visualizing the results. The real-time financial data, stored in CSV format, serves as the data source. The system's hardware requirements, including at least 8 GB of RAM and a multi-core processor, are necessary to handle the computational load of Big Data processing and LSTM model training.

#### 4.2 Project Aims

- 1. Develop an Expense Prediction System using Big Data
- The focus is on designing an expense prediction system by leveraging Big Data and LSTM neural networks. This system will process large datasets containing historical financial transactions, providing users with daily and weekly expenditure predictions.

#### 2. Budget Constraints with Big Data

- The system enables users to apply budget limits to categories like food, transportation, and entertainment, helping them avoid overspending. Big Data analytics will provide real-time insights and alerts as users approach these budget limits, utilizing predictive algorithms.

#### 3. Big Data Analysis and Visualization

- Python libraries such as pandas, Matplotlib, and Seaborn are used for manipulating and visualizing extensive datasets. Big Data techniques will help produce transparent and insightful analyses on spending patterns, with visualizations reflecting trends and future predictions.

#### 4. Incorporation of Large-Scale Financial Data

- The system integrates a vast set of historical bank transaction data in CSV format. This real-time data will be processed using Big Data tools, allowing seamless integration with the Python data analytics stack.

#### 5. Min-Max Predictions using Big Data

- By leveraging Big Data, the system will calculate the standard deviation of the model's prediction errors, providing users with confidence intervals for spending variations. The large volume of data ensures more accurate forecasting for best-case and worst-case financial scenarios.

#### 6. Hardware and System Performance for Big Data Processing

- A minimum of 8 GB RAM and a multi-core processor are required to manage the large-scale data processing involved in training the LSTM model. These hardware specifications ensure efficient Big Data handling, which is essential for running real-time predictions and computations

.

## Chapter 6: Dataset Preparation

Data retrieved from Kaggle and data generated using faker python library and storing in format of Json file

Kaggle dataset Description

My Daily Expenses

#### Context

This is a dataset that consists of my monthly income and expenses. I've been tracking my finances for 4 months till now and I'll update it whenever I finish the consecutive months.

#### Content

There are categories and notes on which I've gained or spent my money on.

#### **Inspiration**

I've started to collect my own financial data to one day find what I've been wasting a lot of my money for. I'm sure I'll find it and hopefully learn from my mistakes.

Faker data set description

#### Content

This dataset includes daily expense records for an individual, detailing their name, email, phone number, transaction amount, date, and category (e.g., Food, Transportation, Miscellaneous). It spans from July 27 to August 4, 2024.

#### Context

It simulates daily spending across various categories, useful for testing expense-tracking applications, analyzing spending patterns, or budgeting.

#### **Inspiration**

Generated using the Faker library, it serves as mock data for testing financial tools, helping analyze and predict personal finance behaviors.

```
| Sustrict | Control | Con
```

Data convert into csv file for preprocessing because fetchable to python preprocessing code snippits

```
utarkash_spendings.csv
       200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 289.56, 2023-04-06, Food
       200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 193.87, 2023-04-06, Entertainment
       200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 436.01, 2023-04-07, Food
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 617.92, 2023-04-08, Utilities
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 26.32, 2023-04-08, Transportation
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 418.97, 2023-04-09, Food
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 91.06, 2023-04-09, Entertainment
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 18.04, 2023-04-09, Transportation
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 167.97, 2023-04-10, Transportation
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 350.76, 2023-04-10, Food
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 300.17, 2023-04-11, Food
     200,Uthkarsh Gola,joshinavya@yahoo.com,4333559941,256.08,2023-04-11,Food
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 119.83, 2023-04-11, Food
      200,Uthkarsh Gola,joshinavya@yahoo.com,4333559941,102.65,2023-04-12,Food
      200,Uthkarsh Gola,joshinavya@yahoo.com,4333559941,329.69,2023-04-12,Food
      200,Uthkarsh Gola,joshinavya@yahoo.com,4333559941,4271.67,2023-04-13,Housing
      200,Uthkarsh Gola,joshinavya@yahoo.com,4333559941,242.87,2023-04-13,Food
      200,Uthkarsh Gola,joshinavya@yahoo.com,4333559941,89.08,2023-04-13,Miscellaneous
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 263.15, 2023-04-14, Healthcare
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 290.32, 2023-04-14, Entertainment
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 1057.53, 2023-04-15, Utilities
      200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 291.35, 2023-04-15, Food
       200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 37.83, 2023-04-16, Healthcare
       200, Uthkarsh Gola, joshinavya@yahoo.com, 4333559941, 98.67, 2023-04-17, Food
```

Python code snippet used for reduction of data for getting desired columns like Data, Amount

```
decating_file.py >...
    import pandas as pd

# Read the CSV file

df = pd.read_csv('utarkash_spendings.csv', header=None, names=['ID', 'Name', 'Email', 'Phone', 'Amount',

# Convert Date column to datetime

df['Date'] = pd.to_datetime[df['Date'])

# Filter only 'Food' expenses

food_expenses = df[df['Category'] == 'Food'][['Date', 'Amount']]

# Group by Date and sum the food expenses

food_expenses_grouped = food_expenses.groupby('Date').sum().reset_index()

# Save to a new CSV file (Food expenses.csv', index=False)

# Group all expenses by Date and sum

total_expenses_grouped = df.groupby('Date').sum().reset_index()[['Date', 'Amount']]

# Save to a new CSV file (Total expenses)

total_expenses_grouped.to_csv('total_expenses.csv', index=False)

print("CSV files for food expenses and total expenses have been created.")
```

This data will be used for data training

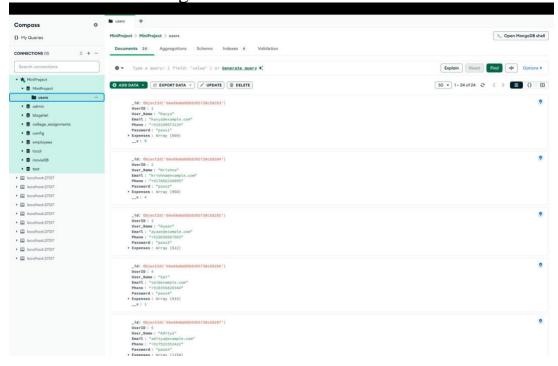
```
2023-07-05,622.59
    2023-07-06,600.54
   2023-07-07,796.34
70 2023-07-08,1122.52
71 2023-07-09,138.11
   2023-07-10,315.5
    2023-07-11,432.2
   2023-07-12,187.09
75 2023-07-13,1445.24
    2023-07-14,166.11
    2023-07-15,233.43
    2023-07-16,640.62
    2023-07-17,746.81
    2023-07-18,147.45000000000000
    2023-07-19,595.26
    2023-07-20,111.17
   2023-07-21,185.61
84 2023-07-22,230.63
85 2023-07-23,62.61
    2023-07-24,523.1
     2023-07-25,414.91
88 2023-07-26,692.15
89 2023-07-27,547.88
    2023-07-28,388.09
    2023-07-29,233.71
    2023-07-30,316.15
    2023-07-31,781.75
   2023-08-01,614.3
```

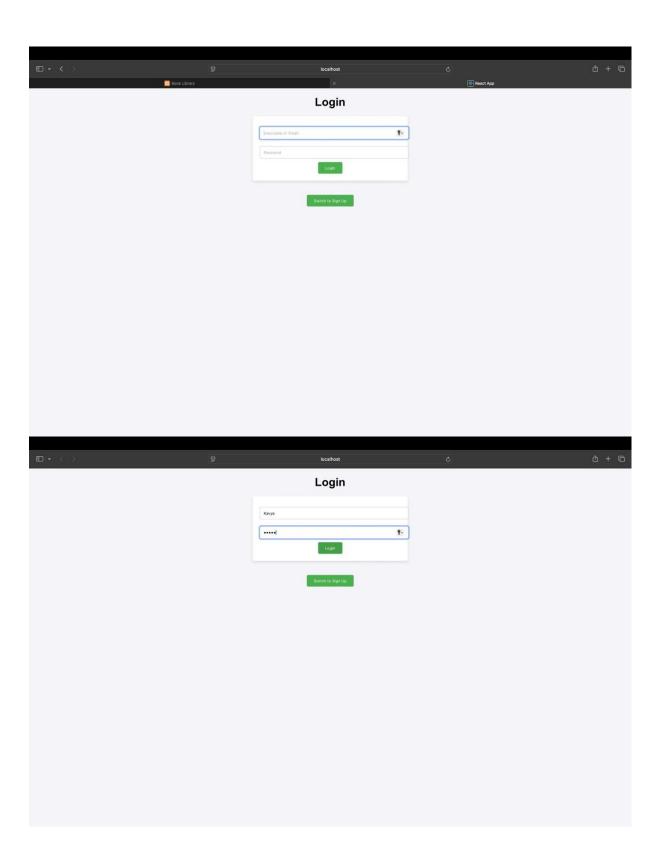
More refined data for weekly prediction of each category like food ,utilities This data will be use for training purposes

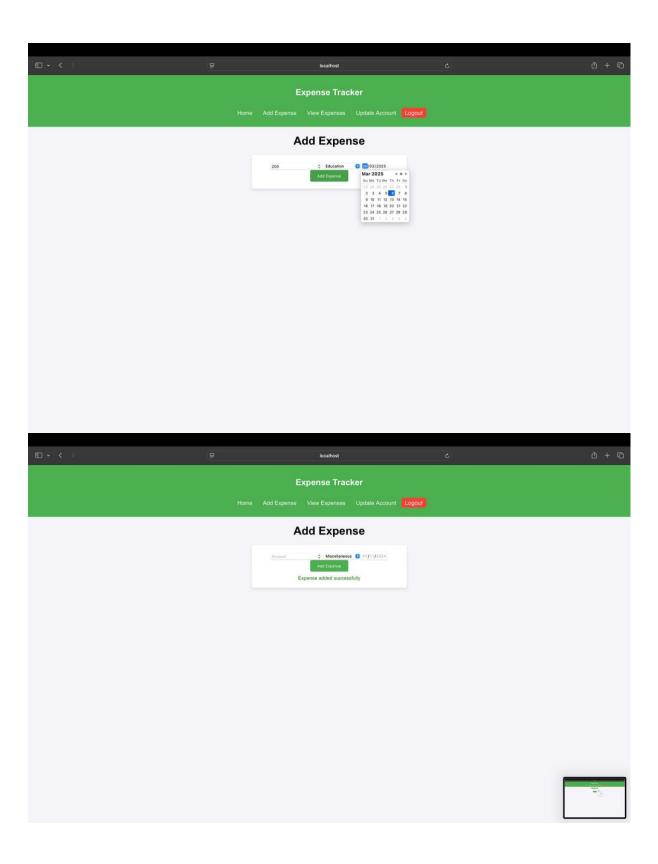
```
weekly_expenses.csv
     Week,Amount
     2023-14,725.5699999999999
     2023-15,2412.37
     2023-16,2030.63
     2023-17,1145.45
     2023-18,1786.48
     2023-19,1946.76
     2023-20,2127.66
     2023-21,1772.93
     2023-22,1202.47
    2023-23,1448.1799999999998
      2023-24,1598.3
13 2023-25,1924.39
    2023-26,1767.98
     2023-27,2402.17
    2023-28,1012.68000000000001
     2023-29,1532.13
     2023-30,1188.1
    2023-31,2125.1
     2023-32,1633.69
    2023-33,2023.9299999999999
    2023-34,1960.3799999999999
      2023-35,1475.32
     2023-36,1548.56
     2023-37,1507.43
      2023-38,468.16
     2023-39,1032.22
     2023-40,901.96
      2023-41,1814.68
```

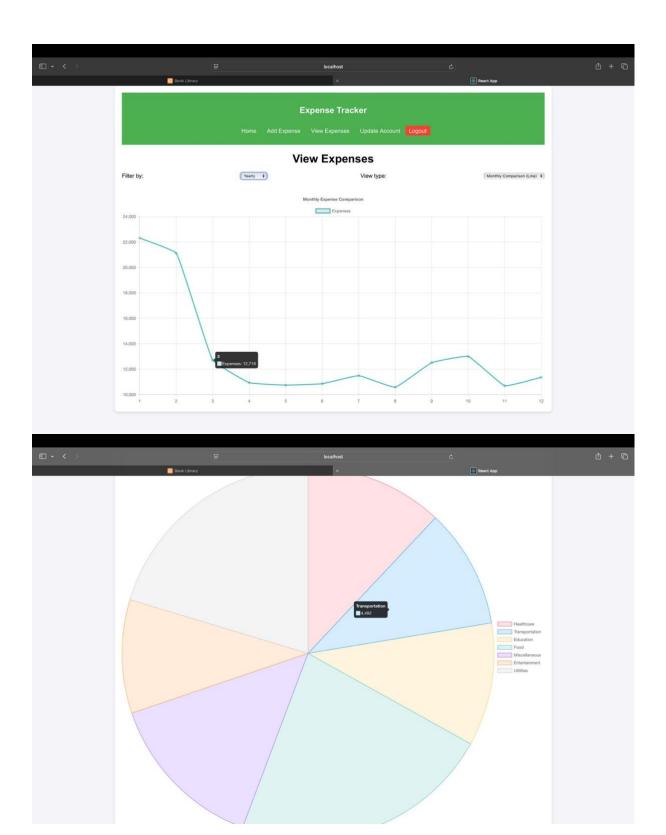
```
weekly_expenses.csv
      Week,Amount
      2023-14,725.5699999999999
      2023-15,2412.37
      2023-16,2030.63
      2023-17,1145.45
     2023-18,1786.48
     2023-19,1946.76
     2023-20,2127.66
     2023-21,1772.93
    2023-22,1202.47
2023-23,1448.179999999998
    2023-24,1598.3
2023-25,1924.39
2023-26,1767.98
    2023-27,2402.17
    2023-30,1188.1
2023-31,2125.1
    2023-32,1633.69
21 2023-33,2023.9299999999999
      2023-34,1960.379999999999
    2023-35,1475.32
    2023-36,1548.56
2023-37,1507.43
      2023-38,468.16
      2023-39,1032.22
     2023-40,901.96
     2023-41,1814.68
```

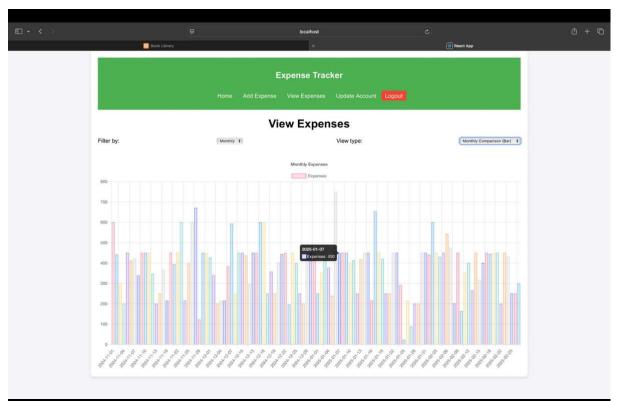
Data feeded to MongoDB schema

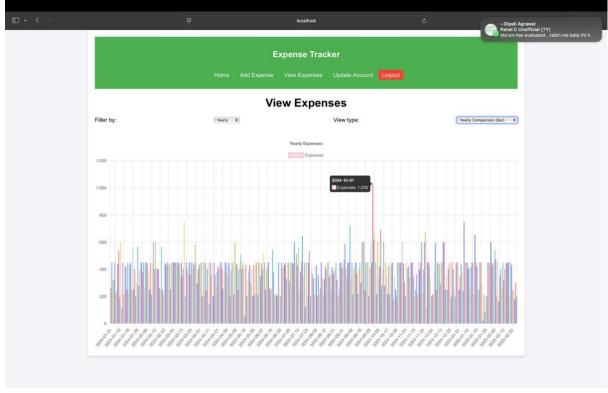












### Chapter 6: Novelty

This project brings a fresh and innovative approach to personal finance management by going beyond traditional expense tracking and introducing an expense prediction feature. While many existing applications focus primarily on recording, categorizing, and analyzing past expenses, this project takes a significant step forward by offering users the ability to forecast future spending. This predictive functionality is something that, to date, has not been widely implemented in similar applications, making it a truly novel contribution to the domain of personal finance management.

The key innovation in this project lies in the integration of a Long Short-Term Memory (LSTM) model, a specialized type of recurrent neural network (RNN) designed to handle time-series data. In this case, the LSTM model processes a user's historical spending patterns to predict their future expenditures. This is a marked improvement over simple trend analysis or budgeting tools that require users to manually input projected spending amounts. Instead, the model leverages past data to offer data-driven predictions, enabling users to anticipate their future spending behavior based on actual trends.

Another unique aspect of this project is its emphasis on personalized prediction. Most applications in the market today focus solely on aggregating past data into reports or charts, leaving the task of future financial planning entirely up to the user. In contrast, this system actively predicts what users are likely to spend in the coming days, weeks, or months, making it easier for them to prepare for upcoming financial obligations. This shifts the focus from reactive expense tracking to proactive financial planning, significantly enhancing the user's ability to manage their finances effectively.

In addition to prediction, this project is designed to help users manage their budgets more intelligently. For example, predictions can be tailored to provide insights into category-wise spending limits—such as food, entertainment, or travel—ensuring that users stay within their set budget while still maintaining flexibility. The LSTM model is also capable of identifying spending trends over time, such as recognizing increased spending on weekends or holidays, which allows the system to provide even more accurate and context-aware predictions.

No widely available expense management tools offer this combination of AI-powered predictions and dynamic budget suggestions. This project's predictive capability not only empowers users to make smarter decisions about their future expenses but also offers a competitive advantage over traditional expense tracking systems. The LSTM model's ability to learn from historical data and continuously refine its predictions ensures that the tool becomes more accurate and useful as it receives more data, making it a truly adaptive solution.

#### **Key Points:**

**Expense Prediction Capability:** Unlike existing applications that focus solely on tracking past expenses, this project predicts future expenditures, giving users a forward-looking perspective on their finances.

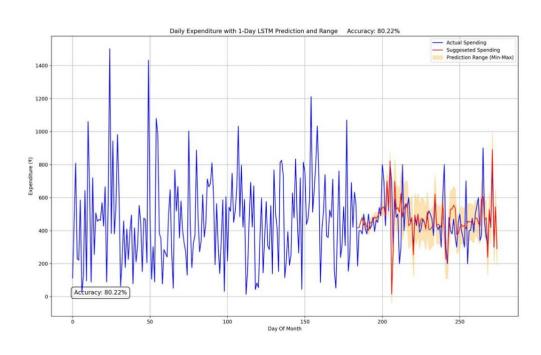
**Integration of LSTM Model:** The project uses an LSTM neural network to process timeseries data, allowing for more accurate predictions of future expenses based on historical spending patterns.

**Personalized Financial Insights:** The system provides personalized predictions tailored to each user's spending behavior, offering daily, weekly, or monthly expense forecasts.

**Category-wise Budget Suggestions**: Users receive recommendations for specific spending categories, ensuring that their expenditures align with their budget while adapting to changing financial trends.

**Proactive Budget Management:** By forecasting upcoming expenses, users can make informed decisions and adjust their budgets ahead of time, improving their overall financial planning.

**First-of-its-kind Solution:** No other widely used expense tracking tools have implemented an AI-based predictive model for future spending, positioning this project as a novel and valuable addition to the market.



# ← → 中 Q 左 🖰

### Chapter 7: Conclusion

Our project aims to revolutionize expense management by integrating machine learning for predictive analytics. By analyzing historical expense data and providing forecasts, the system offers valuable insights and recommendations for users. This approach not only addresses the limitations of traditional expense tracking methods but also introduces a novel solution that enhances financial planning and decision-making. The implementation of predictive analytics in expense management is a significant step forward, offering users a sophisticated tool to manage their finances more effectively.

### Chapter 8: References

#### Vancouver

ENT SYSTEM.

- EXPENDITURE MANAGEMENT SYSTEM [Internet]. ResearchGate. Available from: https://www.researchgate.net/publication/360620084\_EXPENDITURE\_MANAGEM
- Expense Tracker. Int J Res Appl Sci Eng Technol (IJRASET). 2021 Apr;9(4):1067.
- Finery. *Data Visualization*. 2022. Provides a strong visual representation of expenses, but lacks predictive recommendations.
- Rather Go Commands Player. *Command-based Tracking*. 2023. Useful for tracking daily expenses, but not integrated with advanced machine learning methods.
- SNAP. Data Analysis. 2021. Effective in tracking expenses; lacks predictive capabilities.
- Jupiter. Statistical Methods. 2020. Efficient for expense management; does not integrate machine learning or predictive analytics.
- Swindon Simplify. Basic Expense Tracking. 2019. Simplifies expense tracking for users; no inclusion of predictive analytics.
- Finery. Data Visualization. 2022. Strong visual representation of expenses; no predictive recommendations provided.
- Rather Go Commands Player. Command-based tracking. 2023. Useful for tracking daily expenses; not integrated with advanced machine learning methods.
- Expense Tracker. Basic Expense Tracking. 2021. Effective in categorizing expenses; lacks predictive analysis features.
- Expenditure Management System. Expense Management. 2022. Manages daily and monthly expenses; does not use advanced predictive models.
- Expense Tracker. Expense Management. 2021. Tracks daily expenses; no predictive analytics or forecasting tools.