IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 27, Issue 1, Ser. 3 (Jan. – Feb. 2025), PP 46-52
www.iosrjournals.org

Personal Finance Management Integrating Chatbot

Tanushree Dhananjay Bhamare
Computer Science & Business Systems (CSBS) Ramrao Adik Institute Of Technology
Navi Mumbai, India

Trupti Agarkar
Computer Science & Business Systems (CSBS) Ramrao Adik Institute Of Technology
Navi Mumbai, India Trupti.

Shreya Bharadwaj
Computer Science & Business Systems (CSBS) Ramrao Adik Institute Of Technology
Navi Mumbai, India

Abstract
Managing personal finances can be complex, balancing income and expenses against future financial goals. Our Personal Finance Management (PFM) tool, integrated with a machine-learning-powered chatbot, aims to assist users in tracking finances, making expense forecasts, and enabling smarter financial decisions. This system utilizes predictive models such as Linear Regression, Decision Trees, and Random Forest to analyze financial data and generate actionable insights through an intuitive chatbot and dashboard. It examines income patterns, expenditures, and savings to provide insights for improved financial health. The PFM emphasizes user engagement, data privacy, and prediction accuracy, empowering individuals to take greater control of their financial well-being. Keywords—Personal finance management (PFM), Machine learning, Chatbot, Financial management, Predictive modeling.

Date of Submission: 29-01-2025 Date of Acceptance: 09-02-2025

I. Introduction

In today's financial landscape, managing personal finances is challenging due to fluctuating expenses, multiple income streams, and various savings goals. This project addresses these complexities by creating a PFM system that leverages machine learning to predict future expenses and offers insights into financial health through a chatbot interface, allowing users to ask finance-related questions. Our system combines simple and complex machine learning algorithms for accurate predictions and insights based on historical financial data.

II. Ease Of Use

The PFM System is designed for ease of use, enabling users to track and improve their financial health effortlessly. Key features include:

- Chatbot for Real-time Insights: A conversational interface allows users to ask finance-related questions and receive instant, comprehensible responses.
- Financial Health Score: Calculates a score based on income stability, savings rate, and expense-to-income ratio, providing a snapshot of financial well-being.
- Income and Expense Tracking: Provides detailed breakdowns of income and spending across categories, displaying income-to-expense ratios for balanced budgeting.
- Expense Predictions: Forecasts expenses over monthly, quarterly, or yearly periods based on historical data, aiding short- and long-term budgeting.
- Graphical Budget Visualization: Displays monthly expenditure trends and category-based allocations, allowing users to track and adjust spending habits.
- Reporting: Offers customized reports on income, expenses, and financial ratios, helping users analyze spending behaviors.

III. Literature Survey

Current Status

There is a significant gap in financial literacy; while 19.6 percent of individuals have received formal financial education, 74.3 percent of respondents believe they are financially literate. Machine Learning (ML) has started to make strides in the financial sector, particularly in predictive analysis, offering promising accuracy in various areas:

- Expenditure predictions: 83.55% [1]
- Investment predictions: 86.7% [2]
- Retirement planning: 84.53% [2]

Existing Solutions

Several solutions have emerged to address the challenges in financial management:

- Financial Education Simulators: These tools are designed to enhance financial literacy by simulating real-world financial scenarios and providing interactive learning experiences [1].
- ML-Based Predictive Financial Models: These models leverage ML algorithms to predict market trends, individual expenditure, and investment outcomes with high accuracy [2].
- Digital Payment Analysis Tools: These tools analyze digital payment behaviors to detect fraud, predict spending habits, and provide insights into consumer financial behavior [4].
- Automated Financial Advisory Systems: These systems use AI to offer personalized financial advice, helping users make informed decisions regarding

Research Gaps

Despite the progress made in financial technology, there are still several gaps in existing research and solutions:

- Integration Gaps: AI and real-time financial analysis have not been well-integrated, limiting the potential for personalized financial education.
- Technical Gaps: There are challenges in real-time data processing, adaptive learning in financial chatbots, and ensuring privacy-preserving techniques in financial applications.

Uniqueness Of Our Project

Our project presents an integrated solution for AI-based financial management, combining real-time data analysis and personalized education. Key features of the project include:

- Advanced Features: Real-time budget tracking, financial recommendations, and secure data handling.
- User-Centric Design: The system uses an accessible natural language interface, adaptive learning algorithms, and behavior-based guidance to engage users effectively.
- Innovation Points: The project integrates a chatbot interface with an ML-based financial analysis model, offering personalized financial education through conversational AI. It utilizes a combination of ML models to help users manage their finances and adapt in real-time based on user behavior and learning.

This project aims to bridge significant gaps in financial literacy management by providing a practical, user-friendly solution that enhances financial education through intelligent systems.

IV. Units

Monetary values are in the data source currency (e.g., INR, USD). Predictive timelines are in months and years.

V. PFM System Related Equations

Major equations in our PFM System include:

- **Core Financial Health Metrics:**

$$\text{Savings Rate} = \frac{\text{Total Income} - \text{Total Expenses}}{\text{Total Income}} \times 100$$

$$\text{Expense to Income Ratio} = \frac{\text{Total Expenses}}{\text{Total Income}} \times 100$$

- Financial Health Score =
$(w_1 \times \text{Savings}) + (w_2 \times \text{Expense Management}) + \cdots$

- **Anomaly Detection:**

$$\text{Transaction Z-Score} = \frac{\text{Transaction Amount} - \text{Mean Amount}}{\text{Standard Deviation}}$$

- **Random Forest Model:**

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N}\hat{y_i}$$

- **Mean Squared Error (MSE):**

$$MSE = \frac{1}{n}\sum_{i=1}(y_i - \hat{y}_i)^2$$

VI. Implementation

The system is comprised of several phases starting with data preprocessing and going all the way to the final output, including the addition of a chatbot for the interactive relationship with the user and a financial analyzer for analytical purposes. The significant phases are as follows:

Feature Encoding

Converts all categorical variables involved in transactions and categories into numeric formats, which improves the compatibility of the data with machine learning models.

Data Structuring

Orders data in a time-series format, which facilitates better analysis and increases the likelihood of making accurate predictions over the trends seen in a given dataset.

Model Training

The financial analyzer makes use of several machine learning models while determining expenditures and financial health calculations.

- Model Selection: It encompasses models like Linear Regression for simple prediction, Decision Trees to capture complex, non-linear financial trends, and Random Forest to handle ensemble learning.
- Training: The above models are trained against the historical financial data, which has been partitioned 80% for training and 20% for testing, thus making sure that the past trends of spending were learned by these models, while tested on unseen data.
- Hyperparameter Tuning: This optimizes the performance of the model, especially for the Random Forest model, by hyper-tuning some parameters. Some of the parameters are the number of trees and the depth, thus the output is a very accurate model.
- Evaluation Metrics: In the development process, Mean Squared Error (MSE) and R-squared (R^2) values are calculated to evaluate the models. The model with the highest value of predictive accuracy is opted for deployment.

Financial Analyzer Module

The financial analyzer module comprises advanced analytics that are delivered to users with a data-driven perspective.

- Financial Health Score Calculation: Analyzer analyzes income stability, savings rate, and spending habits to give the overall Financial Health Score that mirrors wellness status.
- Expense Forecast: Using models learned by the analyzer, the system generates future expense forecasts on a monthly, quarterly, and yearly basis using historical patterns.
- Income-to-Expense Ratios: The system computes vital ratios—income to expense ratio—for the users to know if the financial situation is in equilibrium or not.
- Visualization and Reporting: It can design custom reports and visualizations; this includes spending trends for the last month, category break-ups, and budgetary recommendations. Bar and line graphics offer a better understanding of spending behavior or financial health.

Integration of Chatbot

It acts as the user interface, allowing direct interaction by enabling the user to engage with the analyzer's insights in real time and providing personalized responses.

- User Input Processing: The chatbot would process the input question from the user about his finances, like some questions on monthly expenditure, saving tips, and a few that are going to come as liabilities in the near future.
- Interaction with Financial Analyzer: In case the user asks for the same, like forecast of expense or financial health score, the interactions occur between the chatbot and analyzer. NLP would help the chatbot to respond naturally and create a great user experience.

- Personalized advice and alerts: The chatbot depends on prediction while giving recommendations and reminders as a proactive assistant to the users. For example, it may remind its user to save more or limit spending in a particular category.

Visualization
Both the analyzer and the chatbot depend on the visual data representation for clarity:
- Graphs and Charts: It presents income, expenses, savings rates, and financial health scores in graphical formats with easy user consumption of complex data.
- Comparative Analysis: It allows actual versus predicted expenses comparisons so that it will track the deviation when effective tracking of spending is developed.

This multi-layered implementation means that a Personal Finance Management System can search for strategic insights on both the higher level of abstraction and the lower, user-friendly interaction plane while constructing an integrated experience in financial planning.

VII. Result & Analysis

The Future Expense Prediction Project aimed to evaluate the effectiveness of different regression models, including Linear Regression, Decision Tree Regressor, and Random Forest Regressor, on a dataset containing personal finance information. This dataset included various attributes such as date, expense category, and income/expense status. The primary objective was to assess how well these models could predict future expenses based on historical data.

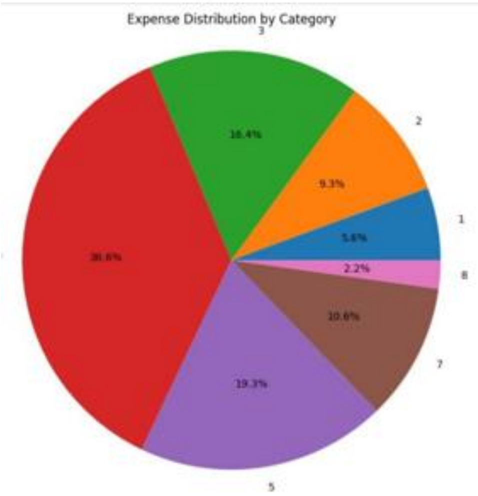

Fig. 1. Expense Distribution by Category Pie-chart (Output)

```
Unusual transactions detected:
     Date / Time  Mode  Category  Sub category  Income/Expense  Debit/Credit  \
28    2021-01-27     0         9           133               1       55000.0
64    2021-02-28     0         9           133               1       55000.0
74    2021-03-21     0         9           133               1       55000.0
85    2021-04-21     0         6           133               1       55000.0
97    2021-05-21     0         9           133               1       55000.0
112   2021-06-22     0         9           133               1       55000.0
120   2021-07-26     0         9           133               1       55000.0
125   2021-08-02     0         9           133               1       55000.0
140   2021-09-24     0         9           133               1       55000.0
144   2021-10-01     0         9           133               1       55140.0
161   2021-11-14     0         6           133               1       55240.0
222   2021-12-30     0         6           133               1       55530.0

      Year  Month  Day  Weekday  Quarter
28    2021      1   27        2        1
64    2021      2   28        6        1
74    2021      3   21        6        1
85    2021      4   21        2        2
97    2021      5   21        4        2
112   2021      6   22        1        2
120   2021      7   26        0        3
125   2021      8    2        0        3
140   2021      9   24        4        3
144   2021     10    1        4        4
161   2021     11   14        6        4
222   2021     12   30        3        4
Financial Health Score: 89.27/100
Savings Rate: 78.54%
Expense to Income Ratio: 21.46%

Financial Health Score: 89.27
Savings Rate: 78.54%
Expense to Income Ratio: 0.21

Available categories in the dataset:
['Allowance' 'Apparel' 'Education' 'Food' 'Household' 'Other' 'Salary'
 'Transportation' 'apparel' 'salary']
```

Fig. 2. Unusual Transaction detection & Financial Health Score,Savings Rate,Expense to Income Ratio Calculations (Output)

Models Evaluated
Three regression models were implemented and evaluated:
- Linear Regression
- Decision Tree Regressor
- Random Forest Regressor

Each model's performance was assessed using two key metrics:
- Mean Squared Error (MSE): Measures the average squared difference between predicted and actual values, with lower values indicating better performance.
- R-squared (R^2): Represents the proportion of variance in the dependent variable predictable from the independent variables. Values closer to 1 indicate a better fit.

Results Summary
Model Performance:
- The Decision Tree Regressor was identified as the best model with an R^2 score of 0.9610, indicating that approximately 96.1% of the variance in expenses could be explained by the model. This suggests excellent predictive capabilities for future expenses based on historical data.
- The Random Forest Regressor was also evaluated, with hyperparameter tuning yielding optimal parameters for better accuracy.

Financial Health Metrics:
- Financial Health Score: 89.27/100
- Savings Rate: 78.54%
- Expense to Income Ratio: 21.46%

These metrics indicate a strong financial position, where savings significantly exceed expenses.

Unusual Transactions Detected: The analysis identified several unusual transactions, highlighting potential areas for financial review. The detected transactions included consistent large debits across multiple months, which may warrant investigation to ensure they are legitimate and necessary.

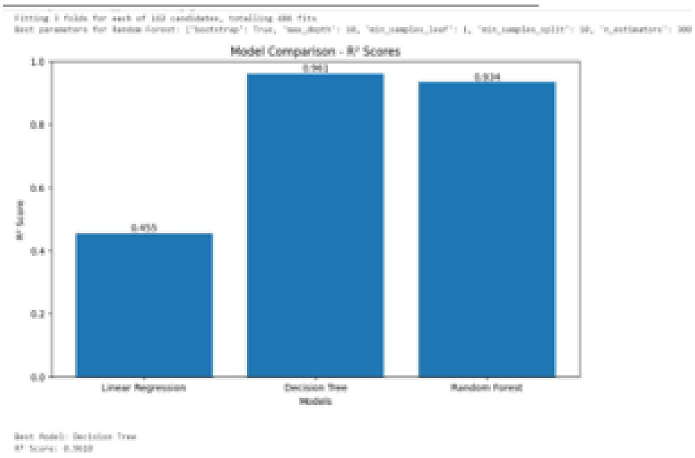

Fig. 3. Model Comparison : R^2 Scores Bar Graph (Output)

Key Observation
- Best Model Selected for Prediction: The Decision Tree is selected as the best model for prediction based on its highest R^2 value of 0.961.

Result & Analysis Comparitive analysis Table : Comparing models:

Model	*R² Value*	*Notes*
Linear Regression	*0.455*	*Performs well with linear relationships but limited accuracy.*
Decision Tree	*0.961*	*Excellent performance with non-linear relationships and minimal error.*
Random Forest	*0.934*	*Robust model, good accuracy, but slightly less effective than Decision Tree.*

(table 4.1 : Result & Analysis Comparitive analysis Table : Comparing models:)

Key Observation:

- ***Best Model Selected for Prediction: The Decision Tree is selected as the best model for prediction based on its highest R^2 value of 0.961.***

Fig. 4. Result & Analysis Comparative Analysis Table : Comparing models (Output)

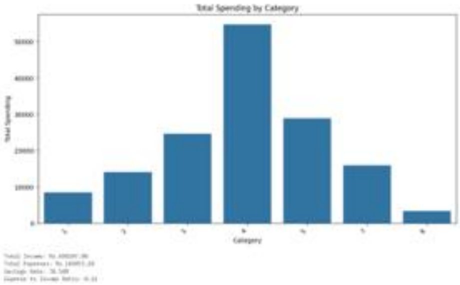

Fig. 5. Total Spending by Category Bar Graph & Total Income,Total Expenses Calculations (Output)

```
Top 3 Expense Categories:
Category
4    54810.00
5    28887.00
3    24607.76
Name: Debit/Credit, dtype: float64

Average Monthly Expense Prediction for 2022:
Rs.3997.33

Financial Insights and Recommendations:
1. Focus on reducing expenses in the top spending categories to save more.
2. Aim to increase your savings rate by cutting unnecessary expenses.
3. Consider creating a budget based on your expense predictions.
4. Look for opportunities to increase your income through investments or side jobs.
5. Start or increase investments to grow your wealth over time.
6. Regularly review and adjust your financial plan based on these insights.
```

Fig. 6. Top 3 Expense Categories, Average Monthly Expense Prediction, Financial Insights & Recommendations Calculations (Output)

VIII. Conclusion

In conclusion, this PFM tool offers an innovative solution for financial management by integrating machine learning algorithms with a chatbot interface. It helps users track their finances, predict future expenses, and take informed decisions to achieve their financial goals. The system is user-friendly, accurate, and capable of personalized financial education. As the field advances, incorporating further technologies like deep learning could enhance predictive power and usability even more.

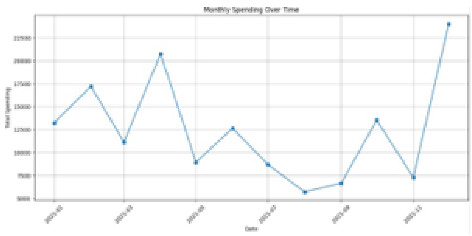

Fig. 7. Monthly Spending Over Time Graph (Output)

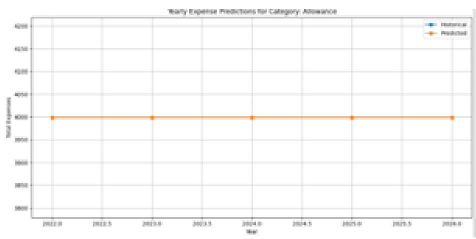

Fig. 11. Yearly Expense Predictions for Category: Allowance Graph (Output)

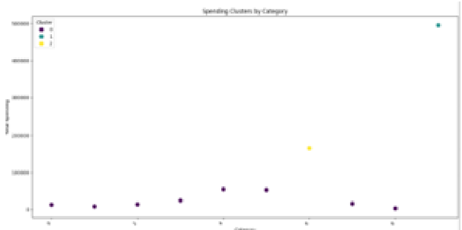

Fig. 8. Spending Clusters by Category Graph (Output)

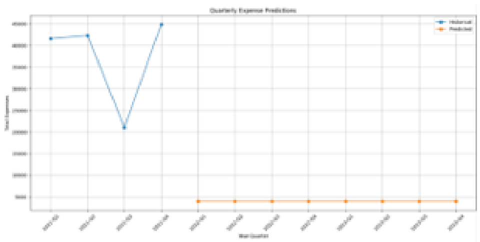

Fig. 12. Quarterly Expense Predictions Graph (Output)

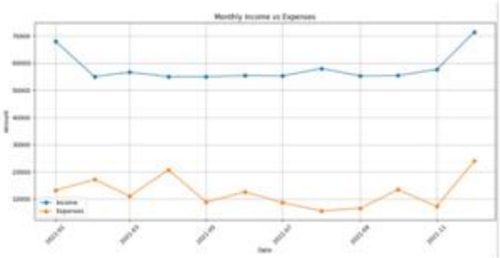

Fig. 9. Monthly Income vs Expenses Graph (Output)

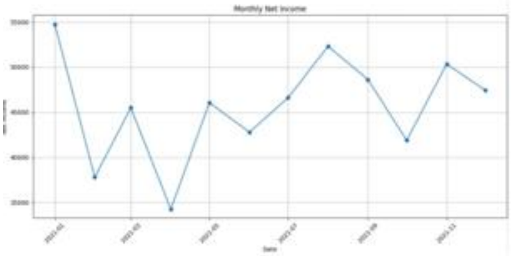

Fig. 10. Monthly Net Income Graph (Output)

```
Future Yearly Expenses (Decision Tree Model):
2022    168000.0
dtype: float64
```

Fig. 13. Future Yearly Expenses by Decision Tree Model Prediction (Output)

```
Financial analysis completed successfully!
Best model saved as 'best_financial_model.joblib'
```

Fig. 14. Financial Analysis Completed Successfully(Output)

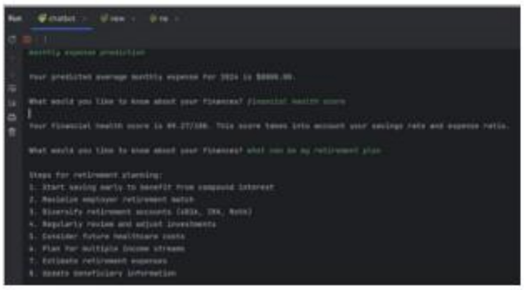

Fig. 15. Chatbot integration (Output)

Acknowledgments

We extend our heartfelt gratitude to the Department of Computer Science & Business Systems at Ramrao Adik Institute of Technology for their invaluable support and resources throughout this project. We would also like to express our sincere thanks to the Head of the Department of Computer Science & Business Systems and the Principal of Ramrao Adik Institute of Technology for their constant guidance, encouragement, and for providing us with an environment conducive to research and innovation.

References

[1] D. S. Antoniuk, T. A. Vakaliuk, V. V. Didkivskyi, O. Vizghalov, O. V. Oliinyk, And V. M. Yanchuk, “Using A Business Simulator With Elements Of Machine Learning To Develop Personal Finance Management Skills,” In Proc. Cosine@ Icteri, Oct. 2021, Pp. 59-70.

[2] R. Kalai, R. Ramesh, And K. Sundararajan, “Machine Learning Models For Predictive Analytics In Personal Finance,” In Modeling, Simulation And Optimization, B. Das, R. Patgiri, S. Bandyopadhyay, And V. E. Balas, Eds. Singapore: Springer Nature Singapore, 2022, Pp. 241–254.

[3] M. Sui, C. Zhang, L. Zhou, S. Liao, And C. Wei, “An Ensemble Approach To Stock Price Prediction Using Deep Learning And Time Series Models,” Preprints, Sep. 2024. [Online]. Available: Https://Doi.Org/10.20944/Preprints202409.2077.V1

[4] Xu Zhang, “Machine Learning Insights Into Digital Payment Behaviors And Fraud Prediction,” Preprints, July 2024.

[5] T. Sarkar, M. Rakhra, V. Sharma, And A. Singh, “An Empirical Comparison Of Machine Learning Techniques For Bank Loan Approval Prediction,” In Proc. 2024 Int. Conf. Communication, Computer Sciences And Engineering (Ic3se), 2024, Pp. 137-143, Doi: 10.1109/Ic3se62002.2024.10593355.