Smart Expense Tracker

Dr. Nita Thakare[1], Astha Murmare[2], Dikshika Nandanwar[3], Sanika Mandlekar[4], Sanchi Muneshwar[5], Devesh Rahangdale[6]
[1,2,3,4,5,6] *Computer Technology, Priyadarshini College of Engineering*

***Abstract*—Managing personal finances in daily life is a vital yet often neglected aspect of our life. Individuals perform number of transactions daily for food, travel, shopping, etc yet most fails to track their expenses. The Smart Expense Tracker is an intelligent web-based system designed to simplify personal management through advance Machine Learning and AI. This technique helps users in analyzing spending behaviour. generating visual dashboards, budget alerts.**
The backend is powered by Python (Flask), utilizing pandas for efficient data processing and analysis.
System leverages MySQL database for secure data storage. The model supports cloud storage and ensures data accessibility and synchronization across various devices. This project aims to simplify personal finance management improving one's decision making through real -time analytics. This reates a simple and smart system to manage their money in easier way.

***Index Terms*—Expense Tracking, Machine Learning, Financial Management, Budget Alerts, Data Visualization, Flask, AI Automation.**

I. INTRODUCTION

In this digital era, tracking personal expenses manually is both time-consuming and inaccurate.
People often lose tracking small daily expenses, which leads to poor financial planning. Traditional methods such as notebooks or spreadsheets require manual effort and discipline.
The application is designed to make management easier by providing a digital platform where users can manage, add, delete, and categorize expenses.
The Smart Expense Tracker system introduces a technology, intelligent expense management solution. It uses Artificial Intelligence (AI) and Machine Learning (ML) algorithms to categorize expenses into relevant groups such as Food, Transport, Utilities, Shopping, etc automatically.
This automation saves time and enhances accuracy in financial management. Through visualization techniques like pie charts, bar charts application provides a dashboard that will help users understand their daily, weekly financial habits. User can also set monthly spending limits, and the system triggers smart alerts when budget threshold is reached. Sytem's architecture integrated secure backend and cloud integration providing accessibility and scalability making it suitable for students, working professionals. The system encourages financial discipline by helping users analyse and visualize the impact of their choices, promoting awareness of unnecessary expenditures and supports long-term saving goals.

II. LITERATURE SURVEY

1. AI- Based Expense Categorization

Recent studies have explored the use of supervised learning algorithms to classify user expenses automatically. For instance, Aishwarya and Hemalatha (2023) proposed a Smart Expense Tracker system that utilized Random Forest and Neural Network models to categorize expenses based on transaction text. Their model achieved improved accuracy and efficiency in expense classification compared to manual categorization.

2. Machine Learning and Predictive Analytics

Beyond classification, several researchers have applied predictive analytics to forecast user spending patterns and future financial needs. Patil et al. (2023) developed a finance management app using Linear Regression and Random Forest to predict monthly expenditure based on historical data. Their system provided recommendations for budgeting and saving, making it more interactive and proactive.

3. Cloud and Web-Based Expense Systems

Modern expense management systems increasingly integrate with cloud platforms for scalability and data security. Singh and Bhosale (2023) implemented a cloud-enabled expense tracker that used Firebase for real-time data storage and synchronization. Their study emphasized that using a centralized cloud system ensures multi-device accessibility and enhances user experience through seamless updated.

4. Visualization and User Engagement

Data Visualization plays an essential role in improving user understanding and engagement. Gupta et al. (2023) developed a personal finance tool that used Plotly and Matplotlib to generate interactive dashboards, allowing user to visualize category-wise spending, detect trends, and make better budgeting decisions. Studies show that visual analytics increases user awareness and long-term adoption of financial management applications.

III. OBJECTIVES

Automated Expense Management: To enable users to add daily income and expenses, which are automatically categorized using machine learning models such as Naïve Bayes or Decision Tree algorithms.

Spending Pattern Analysis: To analyse and visualize spending trends through charts, graphs, and monthly summaries, helping users understand where and how their money is spent.

Budget Planning and Smart Alerts**:** To allow users to set monthly budgets and receive alerts or notifications when nearing or exceeding their spending limits, promoting financial discipline.

AI-Driven Financial Insights: To incorporate predictive analytics and personalized recommendations based on user spending behaviour, making financial management proactive rather than reactive.

Data Security and Cloud Backup: To ensure data privacy through secure authentication and encryption, along with cloud-based backup for real-time synchronization and accessibility across devices.

IV. METHODOLOGY

The development of Smart Expense Tracker follows a systematic and organized process to ensure accurate tracking, analysis, and prediction of user expenses.

1. Data Collection

In this stage of data collection, expense data is gathered from various sources such as user input, uploaded receipts, and digital payment records. Users can manually enter transaction details like amount, date, and description, or use automated methods. All collected data is stored in a structured format within a database to create a reliable dataset for machine learning model training.

2. Data Preprocessing

Raw data often contains noise, missing values, or inconsistencies, so preprocessing is necessary to make it suitable for analysis. The data is cleaned by removing duplicates and unnecessary symbols, and text descriptions are standardized using Natural Language Processing (NLP) techniques. Features such as transaction amount, date, and description are extracted. The data set is then split into training and testing subsets for model development.

3. Machine Learning Model Development

Machine learning model development trains the system to automatically categorize expenses into specific groups such as Food, Travel, and Utilities. Supervised learning algorithms like Naive Bayes, Decision Tree, and Random Forest are applied to the labelled dataset using Scikit-learn library. The model learns to recognize spending patterns from transaction descriptions and predict the category for new entries. Once trained, the model is saved for future use during real-time prediction.

Naive Bayes Algorithm: -

1. Load the training dataset.
2. Preprocess all text by converting to lowercase, removing stopwords, and extracting keywords.
3. Identify all unique categories in the dataset.
4. Calculate the prior probability for each category.
5. Calculate the likelihood of each word appearing in each category.
6. Store all probabilities in the model.
7. Accept a new user input description.

8. Preprocess the text similar to training.
9. Calculate the probability of the description belonging to each category.
10. Select the category with the highest probability.
11. Return the predicted category.

Decision Tree Algorithm: -
1. Load the preprocessed training dataset.
2. Define the feature set (keywords, amount, merchant, etc.).
3. Calculate the impurity (Gini or Entropy) for each feature.
4. Select the feature that provides maximum information gain.
5. Split the dataset into subsets based on the selected feature.
6. Recursively repeat the process on each subset to grow the tree.
7. Stop when all records in a node belong to the same category or when no further split is possible.
8. Accept new user input and extract its features.
9. Traverse the tree from the root node.
10. Follow branches based on matching feature conditions.
11. Stop at a leaf node.
12. Return the category stored in the leaf node.

PERFORMANCE METRIC	NAÏVE BAYES	DECISION TREE
Accuracy	85%	87%
Precision	83%	88%
Training Time	Very low	Moderate
Recall	81%	87%
Best use case	Text-only classification	Text+ numeric data classification

4. Validation and Testing
The trained models are validated and tested to evaluate accuracy and reliability. The model's performance is measured using metrics such as accuracy, precision, recall, and F1-score. Cross-validation ensures that the model performs consistently across different subsets of data. Misclassified transactions are analyses to identify possible improvements in preprocessing or model tuning. This stage ensures that the final model provides accurate and dependable expense categorization in real-world use.

5. System Deployment
The final model is integrated into a user-friendly application that operates in real-time. The trained model is integrated into a web-based application using the Flask framework. The backend connects to a secure database for data storage, while the frontend allows users to input transactions and view categorized results. The complete system is deployed on cloud platforms for real-time access. The deployed system provides automated expense tracker, visual analytics, and budget alerts to the end users.

V. COMPONENTS USED

Hardware Components
1.Computer / Laptop – Runs the software
2.Operating System – Windows, macOS, or Linux

Software Components
1.Frontend Technologies – HTML (Hyper Text Markup Language), CSS (Cascading Style Sheets), JavaScript.

2.Backend Technologies – Python
Machine Learning
1.Scikit-learn – Traditional ML models
2.Pandas / NumPy – Data manipulation Method
Database
1.MySQL – Used to store structured database.

REFERENCES

[1] Kulshrestha et al. (2025) –Proposed using Predictive Analytics to forecast future budgets for better financial planning.
[2] R. et al. (2025) - Simplified expense management via a web application to record, organize, and analyze spending for improved user control.
[3] Manasa, S., & Shivashankar (2025) - Developed a personalized finance management web platform that uses Linear Regression for tracking, predictions, and alerts.
[4] Chauhan et al. (SMARTEXPENSE) (2024) - Used a Convolutional Neural Network (CNN) for deep learning-based anomaly detection to identify irregular spending patterns with high accuracy.

[5] Aishwarya, S., & Hemalatha, S. (2023) - Developed a smart expense tracker using machine learning (Random Forest, Neural Network).
[6] Patil et al. (MIT ADT University) (2023) - Created a finance manager application using ML algorithms to enable budget tracking and personalized dashboards for better savings.
[7] Jain, A., & Patel, P. (2023) - Implemented SMS-based expense tracking using Regex-based filtering and a Support Vector Machine (SVM) model for classification.
[8] Srivastava, V. (2023) - Developed a budget advice system that uses K-Means Clustering + Natural Language Processing (NLP) for dynamic and personalized budget recommendations.
[9] IEEE Access / conference paper (2022) - Improved automatic categorization of transactions and receipts using a CNN/RNN hybrid for OCR and BERT-based text classification.
[10] ACM/Conference (2021) - Focused on extracting structured fields (merchant, amount, date) from receipts and mapping them to categories using OCR, a rule-based semantic parser, and a transformer.