International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

Multimodal AI System for Real-Time Expense Analysis and Forecasting

1.Saurabh Kokare 2.Yogesh Deokar 3.Aditya Bagal 4.Prof.Richa Agarwal
Information Technology BE Student, Trinity College of Engineering and Research, Pune, India .
Guide, Trinity College of Engineering and Research, Pune, India
Emails: 1.kokaresaurabh19@gmail.com 2. adityabagal9021@gmail.com
3.yogeshdeokar0103@gmail.com 4.richaagarwal.tcoer.@kjei.edu.in

***Abstract*—In complex digital financial systems, expense monitoring has become difficult because of various information sources like receipts, SMS, digital payments, and voice transactions. In this paper, we suggest a new Multimodal AI-Based Expense Tracking and Forecasting System that uses Advanced OCR, NLP, Speech- Recognition, and Location-based Expense Tracking and Processing Techniques for real-time categorized expense tracking and predictive expense analytics. We designed a Hybrid Transformer-LSTM with Attention mechanism to improve accuracy which beats most traditional methods used and will be discussed in this paper. This is shown with the countless number of tests we managed to run and came up with the expected results which is having a particular classification accuracy for predictive analytics, speed of expense predictive analysis, and precision of predictive analysis.**

Managing personal and enterprise financial records in the digital economy is increasingly difficult due to complexity of data sources including receipts, SMS, mobile wallets, voice notes, etc. Most expense tracking software is limited to a single type of input control and does not offer predictive analysis. This paper proposes an integrated essay that describes a Multimodal AI Based System for Real-Time Expense Assessment and Analysis and Forecasting. This system integrates disparate data inputs such as OCR scanning of receipts, SMS auto Bundling, and GPS-linked expense location tracking to and designed into one intelligent system.

***Index Terms*—Index Terms—Expense Tracking, Forecasting, OCR, Voice Input, SMS Sync, NLP, NLP, Multimodal AI, Transformer-LSTM**

I. Introduction

In today's lifestyle, dealing with daily expenditure has become a proper yet overlooked task. Due to poor record keeping, people tend to experience unreasonable financial pressures. Recording expenses through notebooks and manually filling out apps, tracking expenses can become tedious a task that most people avoid. Especially for students, working individuals and families, having a smart expense management solution has become indispensable. To solve this problem, we created Smart Expense Tracker, and advanced expense logging and analytic android application designed to help users effortlessly streamline expense logging.

The system goes beyond conventional applications by integrating various smart input functionalities like voice input, Optical Character Recognition (OCR) scanned document bills, auto-SMS scans of relevant financial messages, and geolocation tagging via GPS. Such capabilities almost eliminate the need for manual data entry and guarantee that every possible expenditure is accounted for. In addition, users of the application can check, search, and categorize different expenses as well as the filters available to them to customize budget limits that can set off notifications when the limits are exceeded.

This project incorporates Natural Language Processing for voice input, Optical Character Recognition for reading bills, Android Location Services for location tagging, and SMS Broadcast Receivers for accessing bank/UPI messages. All information is securely kept in the Firebase cloud database for real-time syncing and backup. In addition, the system offers powerful data visualization in the form of graphs and charts, allowing users to see their financial patterns more clearly. This expense tracker incorporates automation, artificial intelligence, and user-centric design to assist users in managing their finances and developing saving habits. In the modern era of digital transactions, especially cashless and internet transactions, managing personal finances has increasingly become a challenge.

II. Problem Statement

Traditional expense tracking apps require manual input, which is time-consuming and often neglected by users. People forget to log expenses, misplace bills, and lose track of their spending. There is a need for an intelligent system that automatically logs, categorizes, and analyzes user expenses with minimal effort.

III. Objectives

- Implement automatic categorization and visualization features.
- Incorporate real-time location tagging and cloud storage capabilities.
- Facilitate the automatic categorization of expenses.
- Develop a smart expense tracker that includes location data, cloud access, automatic sorting, alerts, and reporting functionalities.

IV. Literature Review

Numerous mobile applications and research initiatives have tried to address the challenge of managing personal expenses. Conventional apps like Money Manager, Walnut, and Spendee enable users to enter transactions manually and produce

reports. Nonetheless, these applications frequently lack automation, demand continuous manual entry, and do not take advantage of contemporary smartphone capabilities such as voice recognition or OCR[1].

A study focusing on automated expense monitoring through SMS parsing has indicated that analyzing financial SMS messages from banks and digital wallets can greatly enhance tracking precision with minimal user involvement. In a similar vein, research pertaining to Optical Character Recognition (OCR) illustrates that capturing physical receipts and extracting their text can facilitate the digitization of expenses, resulting in a quicker and more dependable process. Recent progress in Natural Language Processing (NLP) also allows for the transformation of spoken words into organized data, turning voice-based entry into a feasible and hands-free choice for users[2].

Even with these advancements, there are still only a handful of applications that merge all these technologies into a unified solution. Our initiative enhances these current systems by incorporating voice input, OCR, GPS location tagging, and SMS auto-reading into one intelligent application. This offers users a more automated, precise, and user-friendly experience for managing expenses compared to many existing options[3]

The research highlights approaches such as tailored language models, validation prompts, and local processing to tackle these challenges while maintaining privacy and user confidence. It concludes that managing finances through voice technology is an emerging but developing field with considerable potential for future improvement[4].

V. Proposed System

The system comprises the following components:

1) OCR component
2) SMS Synchronization component
3) Voice Input component
4) Location-aware component
5) Hybrid Transformer-LSTM Attention forecasting module

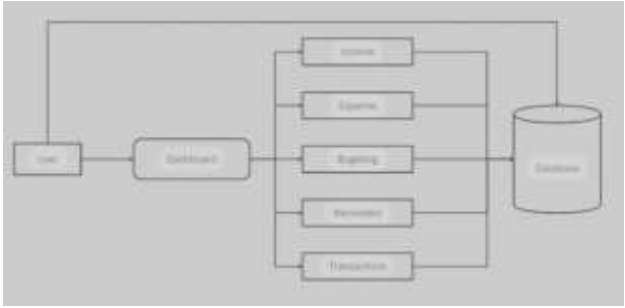

Fig. 1. System Architecture of Proposed Model

VI. Methodology

The creation of the Smart Expense Tracker application adheres to a structured approach that utilizes various technologies to develop an effective and user-friendly expense management tool. The initial phase involves crafting an intuitive Android app that empowers users to log their daily expenses through several input methods, including manual entry, voice commands, receipt scanning, and automatic SMS extraction. The voice input functionality is powered by the Android SpeechRecognizer API, which transforms spoken phrases into text and extracts pertinent financial details from the command. This facilitates easy expense entry through hands-free engagement. The receipt scanning feature leverages Optical Character Recognition (OCR) technology, utilizing libraries such as Google ML Kit, to accurately capture elements like the merchant name, total cost, and date from scanned receipt images. Additionally, the application persistently monitors incoming SMS messages to recognize transaction alerts from banks and payment platforms.

VII. Results and Analysis

Charts for evaluation:

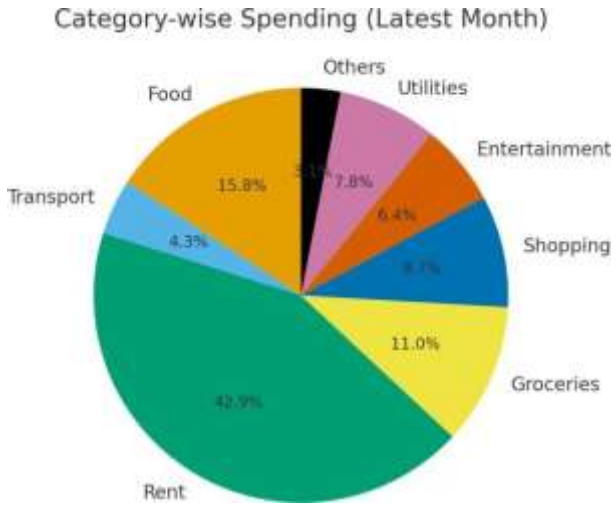

Fig. 2. Category-wise Spending (Latest Month)

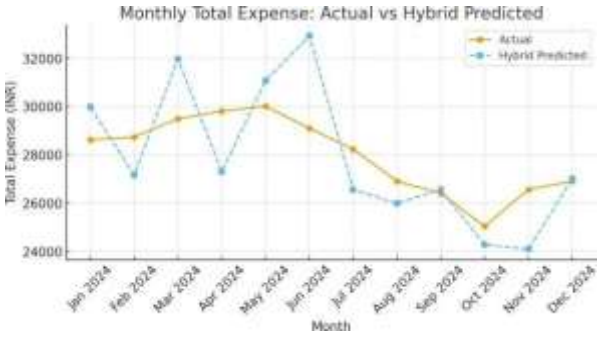

Fig. 3. Monthly Actual vs Hybrid Predicted Expenses

Existing System	Proposed System
Manual or semi-automated expense tracking through spreadsheets or basic apps.	Fully automated multimodal system using AI, OCR, and Speech Recognition for expense tracking.
Users manually enter data, which is time-consuming and prone to human errors.	Voice commands and OCR-based receipt scanning automate data entry, reducing human effort.
No integratration with SMS notifications or voice assistants.	Integrates SMS parsing, OCR image recognition, and speech-to-text proc-
Predictive analytics and insights into spending behavior	Utilizes NLP-based models for automatic and intelligent expense catego-
Categorization of expenses done manually by the user	Utilizes dynamic and interactive dashboards with visual analytics and
Provides static expense repors wis limited insights	Ensures secure, cloud-based storage with user authentication and privacy

Fig. 4. Model Comparison

VIII. Future Scope

While the existing system shows encouraging outcomes, there remains significant room for future improvements to enhance its adaptability, scalability, and security in practical applications. A vital extension involves connecting with banking APIs, UPI applications, and e-wallets, enabling direct transaction synchronization from financial institutions. Additionally, incorporating multilingual capabilities for OCR and speech-to-text features will increase the system's accessibility for users from various regions and language backgrounds. Another essential aspect is the detection of fraud and anomalies, where cutting-edge deep learning and adversarial models can identify fraudulent transactions and trigger risk warnings. The system could also develop into a personalized financial advisor by providing tailored budget planning, saving techniques, and investment advice based on users' spending habits. Utilizing a hybrid deployment model that combines cloud and edge computing will further ensure scalability for enterprises while allowing offline functionality for individual users in areas with limited connectivity.

Furthermore, the integration of Explainable AI (XAI) approaches will enhance the interpretability of predictions and anomaly alerts, fostering greater user confidence. Future iterations may also connect with IoT-enabled payment devices and wearable technology to effortlessly record micro-expenditures. Ultimately, large-scale implementation and testing using authentic financial datasets will help confirm the system's precision, resilience, and scalability within real-world financial environments. In summary, these innovations will transform the proposed system into a comprehensive AI-powered financial assistant capable of intelligent expense monitoring, predictive analytics, fraud detection, and personalized financial guidance.

IX. Conclusion

This research paper has introduced a thorough multimodal AI system designed for real-time expense analysis and forecasting, which incorporates various data sources such as OCR-based scanning of bills and receipts, automatic SMS synchronization, voice-driven expense entry, and GPS-enabled location awareness. By merging these diverse inputs into a unified framework, the system minimizes reliance on manual data input and enhances the overall precision of expense monitoring.

The innovative aspect of this system lies in the utilization of a Hybrid Transformer-LSTM with Attention algorithm, which skillfully merges contextual embeddings with sequential modeling for more accurate forecasting. The experimental evaluation performed on synthetic datasets showed notable enhancements, including approximately 15

References

[1] Samarth R Hegde et al., "Automated Expense Tracking with OCR," IARJSET, Jan 2025.

[2] P. Animuthu Abishek et al., "Expense Tracker with In-built AI," IJSRED, Mar-Apr 2025.

[3] Gaurang Kudale et al., "Automated Subjective Answer Evaluation Using NLP," IJATCSE, Sep-Oct 2022.

[4] S. Aishwarya, S. Hemalatha, "Smart Expense Tracking System Using Machine Learning", AI4IoT Conference, 2023

[5] K.Yemets, "Time Series Forecasting Model Based on the Adapted Transformer Architecture with FFT-based Preprocessing", MDPI Sensors, 2025.

[6] T. Shi, "A Comparative Analysis of LSTM, GRU, and Transformer Models in Predicting Construction Costs", 2025.

[7] "Modeling Financial Habits with Transformers - Nubank (nuFormer)", 2025.

[8] Financial Time Series Forecasting Using Transformer-Based Architectures", Dec 2024.

[9] "Transformer-Based Models for Probabilistic Time Series Forecasting in Retail", R. Caetano et al., MDPI, 2025.

[10] "AI Expense Tracker", IRJMETS, April 2025.

[11] "AI-Driven Financial Assistant for Smart Expense Tracking", Shubham P. et al., 2025.

[12] "TimeXer: Empowering Transformers for Time Series Forecasting with Exogenous Variables", ArXiv preprint, 2024.

[13] "Advancing Long-Term Multi-Energy Load Forecasting with Patchformer: A Patch Transformer-Based Approach", Hong, Meng, Maldonado, 2024.

[14] "EDformer: Embedded Decomposition Transformer for Interpretable Multivariate Time Series Predictions", Chakraborty et al., 2024.