Finwise: A Backend Architecture and Multilingual Voice Interface for Personal Finance Management

Gulam Abdul Qadir[1], Krishna Dave[2], Grishma Parshivnikar[3], Arpan Sorade[4], Dr. Tabassum H Khan[5]

Department of Artificial Intelligence G H Raisoni College of Engineering and Management, Nagpur, India

DOI: https://doi.org/10.51244/IJRSI.2026.1311000019

Received: 04 November 2026; Accepted: 09 November 2026; Published: 19 November 2026

ABSTRACT

Most financial applications presume you know English language well enough to be able to read, write and understand it. Many do not have this ability; some may prefer not to use the written form. Elderly smartphone users, first time smartphone users and those who speak regional languages typically revert to using a notebook or their memory for tracking their income, expenses, savings, debt (loans), monthly loan payments (EMI) and credit card accounts. FinWise is an android application which brings all of your income, expenses, savings, loans/EMIs, Credit Cards onto single dashboard. In this research paper we discuss two components of this application: the backend and the voice assistant.The Backend handles Authentication, Transaction Storage, Cloud Sync with Offline SQLite Cache and Encrypted Transport.The Voice Assistant accepts voice input from user in three languages i.e., Hindi, Marathi and English identifies the intent behind the command given by user and performs action based on that e.g., logs an Expense or Reads out a Saving Tip.This Research Paper details major Design Decisions made while designing these two components, describes how they integrate with each other through Categorization Module and Recommendation Module of Application. It also provides information about Testing conducted till date. Till date Functional Testing & End-to-End Testing has been completed but accuracy and latency measurement will follow.

Keywords : Personal Finance, Voice Assistant, Multilingual Speech Interface, Indian Languages, Backend Architecture, Financial Inclusion, Android

INTRODUCTION

Most people in our country track their money with so many different apps like for expenses they use different app for income they use different app and for savings they use different app. Then there comes a different problem in India estimated 19.1% of people in India above age seven are illiterate. So, for solving these 2 problems mostly we have built FinWise an AI powered app that have unified dashboard, tracks – expenses, loans, EMIs and cards, Its AI suggestion system suggests you how and where you should put your money based on your income and expenses. Then comes the crazy part as for the illiterate people it has the AI voice assistant which logs expenses and some other details like savings, expenses, etc. with just your voice and also for the illiterate people the voice assistant also supports the regional languages but for now it have only 3 languages supported currently (English, Hindi, Marathi).

II. RELATED WORK

A. AI in Personal Finance Apps

Most tools out there currently only do 1 thing at a time. IJERT, April 2026 paper gives us the ML based expense categorisation and recommendation system that uses the OCR receipt scanning [1]. Research by Sai Deepak Talasila in IRJET vol. 1, 2024 proposes MyFinanceAI which is given to 1,000-user pilot to do their financial planning for 6 months which resulted in 22% higher monthly savings and reducing their financial stress [2]. AI recommendation systems learn users' behaviour and then suggest them suitable financial products, Investments, or actions according to article on springer papers in volume 81 published on 03 March,2022 [3].

B. Speech and Language Technology for Indian Languages

There are few tools available for the languages for speech model like whisper a large multilingual recognition model which is good but sometimes gets weaker in low resources languages then there is another one Bhashini a Indian government model which is used for the project which is also very suitable for initial stages of the project currently we our using and a manual voice recognition which is built with code but Bhashini is preferred for this project which we will use for further development of this project. We used Bhashini model for our project because it contains more Indian regional languages than whisper and also, it is an Indian model. It helps the illiterate people to also manage their finances and increase their savings while reducing their financial stress.

III. Problem Definition and Objectives

When most of the people talk about the finance apps it is just like they are talking about an app which does only a thing at a time like one app for managing expenses, another for savings and other apps for doing only a particular thing and also most of the apps have the slow manual logging hard-to-track loans, EMIs, interest and card dues.

Our objective is to reduce the financial stress of the people and give them the seamless experience for managing their finance so for that we have built an app that have unified dashboard to do all loggings to manage expenses, saving, income, etc. at the same time and it have the voice assistant which solves the problem of manually logging everything to the app and giving users a new experience to manage it.

IV. System Architecture

FinWise is layered into several parts. Input arrives by manually typing or by using voice assistant. Then the input is processed by the processing layer and detects what is it asked for. A core engine then routes each request to the modules for expense analysis, budgeting and goal planning. Results go into the encrypted database, and after that output arrives in the form of insights, alerts and reports as voice or text. User reacts and then its reaction flow back to the model so suggestions improve overtime. These are the layers in the FinWise which is used for the whole system to work.

The android app comes in the middle. As it hosts the dashboard and voice interface, have a local SQLite database for the app to be used offline, and calls the cloud services for authentication, synchronisation and backup. TLS is used for the protection of traffic between phone and cloud. The backend described comes sits between the app, the cloud store and the AI modules. Then the Touch and Voice go through it and an expense spoken aloud or hand typed data passes the same validation and land in the same table.

V. Backend Design

A. Services and Data Model

Below is the Table I which contains all the services listed with all the information for each and every services. Mainly the logging of the data and the voice assistant come under the main backend operations for the current design.

TABLE I [Backend services and responsibilities]

Service	Responsibility
Authentication	Sign-in, session handling and per-user data separation
Transactions	Record income and expenses with amount, date, purpose and category
Credit	Track loans, EMIs with interest, card dues and due dates
Savings and dashboard	Aggregate balance, monthly spending and savings progress
AI gateway	Pass transactions to categorisation and recommendation modules and return results

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue XI November 2026

Sync and backup	Reconcile the local SQLite cache with cloud storage
Voice handler	Uses Bhashini API to process the voice of the other person and log the data and store it

B. Storage, Sync and Offline Use

Every entry gets registered on SQLite database. A user who has the poor internet connection records its transaction which stores locally then after that when the phone gets connected to the internet all the data gets synced to the cloud. Supabase will be used in further development and for the data which is entered at the last time will be added to the cloud.

C. Security and Privacy

Financial records and voice data are sensitive so for security it comes with that at the initial stages and for that FinWise have the good security TLS that covers data that is in transit as well as encryption that covers the data which is in rest. Voice recording of a person is stored for less than 5 min and then it gets deleted it is not manual so the user can't control the data that is got into it.

VI. Multilingual Voice Assistant

A. Design Goals

The project guide gave two instructions: app should be easy to use, hence everyone from student to working professionals from literate to illiterate people. So for that anything is done in fewer steps, voice assistant accepts many types of phrases and people can register each and everything in their regional language.

B. Processing Pipeline

Below there is a figure that shows how the processing happens when you give the voice command than app records the audio then using speech-to-text it gets transcribed then the transcribe gets cleaned because Hindi and Marathi mix in English words and write numbers in several ways. "Thousand", "1000" and their Marathi and Hindi must result to the same amount. After that intent and entity extraction decides what person wants to extract the details then action router calls out the matching backend service then reply comes in the form of text and spoken aloud.

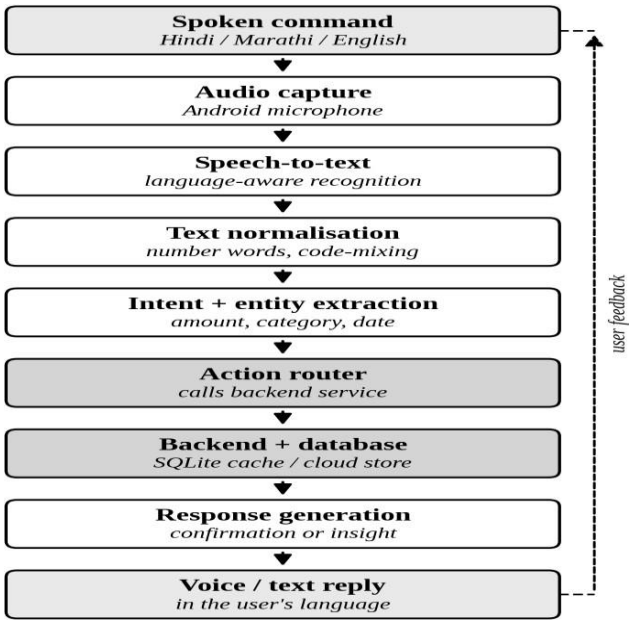

Fig. 1. Processing pipeline of the FinWise voice assistant

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue XI November 2026

C. Supported Intents

For this day the intent set is small on purpose to test all the things. A list of intent given below is tested in all the three languages and the table (Table II) below lists all the intent, details and example.

TABLE II [SUPPORTED INTENTS AND EXAMPLE UTTERANCES]

Intent	Details needed	Example utterance
Log expense	Amount, purpose, date	"Aaj 200 rupaye chai par kharch kiye" (Hindi)
Check balance	None	"Majhyakade kiti shillak aahe?" (Marathi)
Monthly spending	Month, optional category	"How much did I spend on food this month?"
Hear advice	None	"Mujhe bachat ki salah batao" (Hindi)

If the required details are missing then the voice assistant will ask for the follow-up instead of random guessing it.

D. Handling Variation in Speech

Three kinds of variation in our design take's place. Sometimes in sentence language mixing comes like people drops "Petrol" or "Bill" in the middle of Hindi and Marathi sentences so we cannot make the model assume one language utterance then for second comes the number of words spoken last comes the accent and background noise because they get to the recognition stage before any logic described. Mentor asked to insert intent recognition that survives natural phrasing that is why assistant matches the meaning and key entities not just fixed sentences.

VII. IMPLEMENTATION AND RESULTS

A. Implementation Setup

FinWise is an installable android app build using Kotlin and Gradle for android it is an .apk file that when opened user gets prompted to enter their name after that it gets to the main homepage where there is navigation bar that have: dashboard, income and expense entry, credit tracking, settings and the voice assistant. Its version is currently v1.0.0 released with the initials not with all the features.

B. Results

FinWise currently solves the initial problems it is the working project with working - expenses, income and savings log; voice assistant comes with the 3 languages initially English, Hindi and Marathi also it is currently working for 4 understandable commands; A recommendation system for savings of your monthly income based on your expenses of all the month; SIP goal planner which calculates your return based on the no. of years of investment, amount you are investing, expected annual return and risk profile (calculated as equity ÷ debt) which have 3 levels conservative, Moderate and Aggressive; It have the loan calculator which is used to give you the exact amount of EMIs to pay.

DISCUSSION

The results so far show us that the architecture works. The user says a word and the computer reads it from the microphone to the database, then replies back in the user's language. However, what we don't know yet is if Marathi will be as effective as Hindi or English. Speech literature has indicated that Marathi may not be as effective as Hindi or English [4]. This is the first thing we need to find out with this evaluation.

VIII. LIMITATIONS AND FUTURE WORK

Next steps include addition of more Indian languages; better handling of accents and code-mixing in phrasal structures; enhanced intent-Recognition capabilities; receipt scanning for entry of expenses; providing personal finance recommendations for fixed deposits, gold and stocks. On device (voice) processing will allow complete offline operation. The assistant needs further protection of voice logging and user-financial data. Most importantly a field study with real target users is needed.

CONCLUSION

This research has described the backend and voice assistant functionalities of Finwise, a mobile application using android which includes all the user's income, expense, savings and loan payments into a single screen. By routing both touch input and voice input through the same set of APIs, the financial data is ensured to be coherent as well as allowing the voice input capabilities to have full access to all other components of the application. The use of a local SQLite database allows the user to create new transactions when there is no network connection; however, this is done in conjunction with secure encryption (AES) and transport layer security (TLS) to ensure the data stored locally remains protected until synced with cloud servers. Tests were conducted to demonstrate functional operations from start-to-finish in the three languages supported by Finwise: Hindi, Marathi and English. However, these test results did not assess the speed at which the system can recognize user requests nor did they assess how accurately the system recognizes those requests. As such, assessing the systems' ability to provide timely responses (response time), as well as its overall reliability (accuracy) will be assessed in future testing stages.

ACKNOWLEDGMENT

The authors thank their project guide, Dr. Tabassum H. Khan, and the Department of Artificial Intelligence at G H Raisoni College of Engineering and Management for their guidance and support during this project.

REFERENCES

1. S. D. Talasila, “AI-driven personal finance management: Revolutionizing budgeting and financial planning,” International Research Journal of Engineering and Technology (IRJET), vol. 11, no. 07, pp. 397–406, Jul. 2024. [Online]. Available: https://www.irjet.net/archives/V11/i7/IRJET-V11I755.pdf
2. M. Sharaf, E. E.-D. Hemdan, A. El-Sayed, and N. A. El-Bahnasawy, “A survey on recommendation systems for financial services,” Multimedia Tools and Applications, vol. 81, no. 12, pp. 16761–16781, May 2022, doi: 10.1007/s11042-022-12564-1.
3. A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, “Robust speech recognition via large-scale weak supervision,” in Proc. 40th International Conference on Machine Learning (ICML), PMLR, vol. 202, 2023, pp. 28492–28518.
4. AI4Bharat / Bhashini, Bhashini API Documentation and ULCA Platform for Indian Language Speech, Translation, and OCR Models, National Language Translation Mission, Ministry of Electronics and Information Technology (MeitY), Government of India. [Online]. Available: https://bhashini.gitbook.io/bhashini-apis
5. D. Kakwani, A. Kunchukuttan, S. Golla, Gokul N. C., A. Bhattacharyya, M. M. Khapra, and P. Kumar, “IndicNLPSuite: Monolingual corpora, evaluation benchmarks and pre-trained multilingual language models for Indian languages,” in Findings of the Association for Computational Linguistics: EMNLP 2020, 2020, pp. 4948–4961, doi: 10.18653/v1/2020.findings-emnlp.445.
6. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), Minneapolis, MN, USA, 2019, pp. 4171–4186, doi: 10.18653/v1/N19-1423.