FinSenseAI: An Explainable AI-Based Personal Finance Recommendation and Expense Prediction System Using Hybrid Machine Learning Models

Velkur Dilip[1*]; Dr. R. Sathya Janaki[2]

[1]B.Tech. Student, Department of Artificial Intelligence and Machine Learning, Takshashila University, Ongur, Tindivanam, Tamil Nadu – 604305, India
[2]Associate Professor and Head, Department of Artificial Intelligence and Machine Learning, Takshashila University, Ongur, Tindivanam, Tamil Nadu – 604305, India

Corresponding Author: Velkur Dilip[1*]

Publication Date: 2026/07/29

Abstract: Financial management has become increasingly challenging due to the rapid growth of digital payment platforms, multiple banking channels, and continuously changing spending behavior. Traditional expense management systems primarily focus on recording financial transactions without providing intelligent forecasting or transparent recommendations that assist users in making informed financial decisions. The integration of Artificial Intelligence (AI) and Explainable Artificial Intelligence (XAI) has created new opportunities for developing intelligent financial management systems capable of predicting future expenses while maintaining transparency in decision-making. This paper presents FinSenseAI, an Explainable AI-based Personal Finance Recommendation and Expense Prediction System using a Hybrid Machine Learning Framework consisting of Extreme Gradient Boosting (XGBoost), Long Short-Term Memory (LSTM), and SHapley Additive exPlanations (SHAP). The proposed framework utilizes a real-world financial transaction dataset containing transaction date, expense category, transaction amount, transaction type (Income/Expense), and transaction description. Data preprocessing includes duplicate removal, missing value analysis, categorical encoding, temporal feature extraction, normalization, and feature engineering to improve prediction performance. The XGBoost model efficiently captures nonlinear relationships among structured financial attributes, whereas the LSTM network learns sequential spending behavior from historical transactions. The outputs of both models are combined through a hybrid ensemble strategy to enhance forecasting performance. SHAP is integrated to provide interpretable explanations of model predictions by quantifying the contribution of each feature to the predicted expense. The proposed framework is evaluated using standard regression metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). Rather than serving as a conventional expense tracker, FinSenseAI provides predictive financial analytics and personalized budgeting recommendations supported by transparent AI explanations. The proposed architecture is scalable and can be extended to integrate banking APIs, mobile applications, cloud deployment, and real-time financial monitoring, making it a promising solution for intelligent personal financial management. The proposed framework was implemented using Python-based machine learning libraries, including Scikit-learn, XGBoost, TensorFlow, and SHAP. The dataset was preprocessed using feature engineering, categorical encoding, and temporal feature extraction. The framework is designed to support intelligent budgeting and transparent financial decision-making through explainable artificial intelligence.

Keywords: *Personal Finance, Expense Prediction, Explainable Artificial Intelligence, XGBoost, Long Short-Term Memory, SHAP, Hybrid Machine Learning, Financial Forecasting, Budget Recommendation, FinTech.*

How to Cite: Velkur Dilip; Dr. R. Sathya Janaki (2026) FinSenseAI: An Explainable AI-Based Personal Finance Recommendation and Expense Prediction System Using Hybrid Machine Learning Models. *International Journal of Innovative Science and Research Technology*, 11(7), 1637-1650. https://doi.org/10.38124/ijisrt/26jul842

I. INTRODUCTION

Personal financial management plays a crucial role in maintaining financial stability and improving long-term economic well-being. With the rapid adoption of digital payment systems, mobile banking, Unified Payments Interface (UPI), online shopping platforms, and subscription-based services, individuals generate a large volume of financial transactions every day. Managing these transactions manually has become increasingly difficult, often leading to

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

overspending, inefficient budgeting, and poor financial planning.

Most existing personal finance applications focus primarily on recording income and expenses while offering basic visualization tools such as monthly reports and expenditure summaries. Although these features provide users with historical insights, they generally lack intelligent forecasting capabilities that can predict future expenses or recommend optimal budgeting strategies. Moreover, traditional machine learning models often operate as "black-box" systems, making it difficult for users to understand the reasoning behind financial predictions.

Recent advancements in Artificial Intelligence (AI), Machine Learning (ML), and Explainable Artificial Intelligence (XAI) have enabled the development of intelligent financial decision-support systems capable of analyzing historical spending behavior, identifying hidden patterns, forecasting future expenses, and generating personalized recommendations. Ensemble learning algorithms such as XGBoost have demonstrated excellent performance on structured financial datasets due to their ability to capture nonlinear feature interactions. Likewise, Long Short-Term Memory (LSTM) neural networks are highly effective for modeling sequential financial data because they learn temporal dependencies in transaction histories.

Despite these advances, most existing research either relies on a single prediction model or lacks interpretability. Financial decisions directly affect users' economic well-being, making transparency and trust essential requirements for AI-based recommendation systems. Explainable AI techniques such as SHAP address this limitation by identifying how individual input features influence each prediction, thereby increasing user confidence and supporting responsible AI adoption.

This research proposes FinSenseAI, an Explainable AI-Based Personal Finance Recommendation and Expense Prediction System that integrates XGBoost, LSTM, and SHAP into a unified hybrid framework. The system predicts future expenses using historical financial transaction data while simultaneously providing interpretable explanations and personalized budgeting recommendations. By combining the strengths of ensemble learning, deep learning, and explainable AI, the proposed framework aims to improve prediction accuracy, enhance financial decision-making, and promote transparent AI-assisted budgeting.

➢ *The Primary Contributions of this Work are:*

- Development of a hybrid XGBoost–LSTM framework for expense prediction.
- Integration of SHAP to provide transparent explanations for financial predictions.
- Design of a scalable architecture suitable for intelligent personal finance applications.
- Evaluation using a real-world transaction dataset with reproducible preprocessing and model training.

According to the Reserve Bank of India (RBI), the rapid growth of UPI transactions and digital payments has significantly increased the volume of personal financial data generated by individuals. This growing availability of structured financial records provides an opportunity to apply Artificial Intelligence for intelligent budgeting, spending prediction, and personalized financial recommendation systems.

II. LITERATURE REVIEW

Artificial Intelligence (AI) and Machine Learning (ML) have significantly transformed intelligent financial management by enabling predictive analytics, automated budgeting, fraud detection, and personalized recommendation systems. Recent studies have focused on improving forecasting accuracy through ensemble learning, deep learning, and Explainable Artificial Intelligence (XAI). Nevertheless, challenges related to interpretability, temporal modeling, and user trust remain open research problems.

➢ *Paper 1*

Chen and Guestrin (2016)

- *Objective*

To develop an efficient gradient boosting algorithm capable of handling large structured datasets with high predictive performance.

- *Method*

Extreme Gradient Boosting (XGBoost)

- *Dataset*

Large benchmark datasets for classification and regression.

- *Key Findings*

The proposed algorithm significantly improved prediction accuracy while reducing computation time through regularization, parallel processing, and optimized tree learning.

- *Limitation*

The study focused on structured tabular data and did not address sequential time-series prediction.

- *Research Gap*

Temporal financial spending patterns require sequence-learning models in addition to tree-based algorithms.

➢ *Paper 2*

Hochreiter and Schmidhuber (1997)

- *Objective*

To overcome the vanishing gradient problem in recurrent neural networks.

- *Method*

Long Short-Term Memory (LSTM)

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

- *Dataset*
  Sequential benchmark datasets.

- *Key Findings*
  LSTM successfully learns long-term temporal dependencies, making it suitable for forecasting tasks.

- *Limitation*
  Performance depends on sufficient sequential data and computational resources.

- *Research Gap*
  LSTM alone cannot efficiently model structured financial attributes without feature engineering.

➢ *Paper 3*
  Lundberg and Lee (2017)

- *Objective*
  To develop a unified Explainable Artificial Intelligence framework.

- *Method*
  SHapley Additive exPlanations (SHAP)

- *Dataset*
  Model-agnostic evaluation across multiple datasets.

- *Key Findings*
  SHAP provides consistent local and global feature importance explanations based on cooperative game theory.

- *Limitation*
  Computational complexity increases with dataset size.

- *Research Gap*
  Financial forecasting systems require efficient integration of SHAP with hybrid machine learning models.

➢ *Paper 4*
  Explainable AI in Finance: Systematic Literature Review (2024)

- *Objective*
  To review Explainable Artificial Intelligence methods applied to finance.

- *Method*
  Systematic Literature Review

- *Dataset*
  Peer-reviewed finance and AI publications.

- *Key Findings*
  SHAP is the most frequently adopted explanation technique for ensemble models such as XGBoost and neural networks.

- *Limitation*
  Most reviewed studies focus on risk assessment and fraud detection rather than personal finance management.

- *Research Gap*
  Limited research addresses explainable personal expense prediction systems.

➢ *Paper 5*
  Model-Agnostic Explainable AI in Finance (2025)

- *Objective*
  To analyze explainability techniques for financial machine learning applications.

- *Method*
  Systematic Review

- *Dataset*
  Published finance-related AI studies.

- *Key Findings*
  Feature-attribution methods such as SHAP dominate financial explainability because they improve transparency without modifying the prediction model.

- *Limitation*
  Current methods remain computationally expensive and difficult to interpret for non-technical users.

- *Research Gap*
  More user-friendly explainable financial recommendation systems are required.

➢ *Paper 6*
  XGBoost–LSTM for Financial Prediction (2024)

- *Objective*
  To improve financial forecasting by combining XGBoost and LSTM.

- *Method*
  Hybrid XGBoost–LSTM

- *Dataset*
  S&P 500 Financial Sector dataset.

- *Key Findings*
  The hybrid framework improved predictive performance by combining structured feature learning with temporal sequence modeling.

- *Limitation*
  The work focused on stock market forecasting rather than personal financial transactions.

- *Research Gap*
  Hybrid learning can be extended to intelligent personal finance recommendation systems.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

➢ *Paper 7*

Explainable Financial Risk Prediction (2026)

- *Objective*

To improve financial risk prediction while maintaining model transparency.

- *Method*

Transformer + XGBoost + SHAP

- *Dataset*

Corporate financial risk dataset.

- *Key Findings*

The hybrid architecture improved predictive accuracy and enabled feature-level interpretation using SHAP.

- *Limitation*

Designed specifically for corporate financial risk prediction.

- *Research Gap*

A similar explainable hybrid architecture can be adapted for personal finance applications.

➢ *Paper 8*

Machine Learning and Deep Learning in Computational Finance (2025)

- *Objective*

To review modern machine learning and deep learning applications in finance.

- *Method*

Systematic Review

- *Dataset*

Research articles published between 2024 and 2026.

- *Key Findings*

Hybrid models combining ensemble learning and deep learning generally outperform standalone models while explainable AI improves transparency.

- *Limitation*

The review highlights limited validation on real-world personal finance datasets.

- *Research Gap*

Personal finance recommendation systems remain comparatively underexplored.

➢ *Literature Comparison Table*

Table 1 Literature Comparison

Ref	Objective	Method	Application	Limitation	Research Gap
Chen & Guestrin (2016)	Efficient boosting	XGBoost	Structured prediction	No temporal modeling	Needs sequence learning
Hochreiter & Schmidhuber (1997)	Sequential learning	LSTM	Time-series	High computational cost	Needs structured features
Lundberg & Lee (2017)	Explainability	SHAP	Model interpretation	Computationally intensive	Hybrid XAI integration
XAI in Finance Review (2024)	Review XAI	Literature Review	Finance	Focus on risk/fraud	Personal finance gap
Model-Agnostic XAI Review (2025)	Review explainability	Systematic Review	Finance	Complex for users	User-centric explanations
Hu & Luo (2024)	Financial forecasting	XGBoost–LSTM	Stock prediction	Market-specific	Personal finance prediction
Financial Risk Prediction (2026)	Explainable prediction	Transformer + XGBoost + SHAP	Corporate risk	Corporate focus	Consumer finance
Computational Finance Review (2025)	AI in finance	Review	Finance	Limited personal datasets	Expense forecasting

➢ *Research Gap*

The literature indicates that most existing research focuses on stock market prediction, credit risk assessment, financial fraud detection, or corporate financial analysis. Although XGBoost demonstrates excellent performance on structured financial data and LSTM effectively captures temporal dependencies, relatively few studies integrate these techniques into a unified framework for personal expense prediction. Furthermore, many existing systems lack transparent explanations that allow users to understand AI-generated recommendations. This motivates the development of FinSenseAI, which combines XGBoost, LSTM, and SHAP to deliver accurate, interpretable, and personalized expense forecasting and budgeting recommendations.

Therefore, this research proposes an explainable hybrid framework that combines tree-based learning, deep learning, and explainable artificial intelligence to overcome the limitations of existing financial recommendation systems.

➢ *Problem Statement*

Managing personal finances has become increasingly challenging due to growing transaction volumes, multiple digital payment platforms, subscription services, and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

changing consumer spending habits. Conventional expense tracking applications mainly record historical transactions and generate descriptive reports without forecasting future expenses or providing intelligent budgeting assistance. As a result, users frequently encounter overspending, poor savings habits, and ineffective financial planning.

Although recent machine learning techniques have improved financial forecasting, many existing systems either lack sufficient prediction accuracy or fail to provide interpretable explanations for generated recommendations. This absence of transparency reduces user confidence in AI-assisted financial decision-making.

Therefore, there is a need to develop an explainable hybrid machine learning framework capable of accurately forecasting future personal expenses while simultaneously providing transparent, understandable, and actionable budgeting recommendations.

➢ *Research Objectives*

The primary objective of this research is to develop an intelligent and explainable personal finance management framework capable of accurately predicting future expenses while providing transparent budgeting recommendations through Hybrid Machine Learning and Explainable Artificial Intelligence.

- *The Specific Objectives of the Proposed Research are as Follows:*

✓ To develop an intelligent expense prediction system using real-world financial transaction data.
✓ To preprocess and transform raw financial transaction records into meaningful machine learning features.
✓ To implement the XGBoost algorithm for modeling nonlinear relationships among financial attributes.
✓ To implement Long Short-Term Memory (LSTM) neural networks for learning temporal spending behavior.
✓ To combine XGBoost and LSTM into a hybrid prediction framework for improved forecasting accuracy.
✓ To integrate SHAP (SHapley Additive Explanations) for transparent and interpretable financial recommendations.
✓ To evaluate the proposed framework using standard regression performance metrics.
✓ To develop a scalable architecture suitable for deployment in web and mobile-based financial management systems.

III. PROPOSED SYSTEM

The proposed system, FinSenseAI, is an Explainable Artificial Intelligence (XAI)-based personal finance recommendation and expense prediction framework designed to assist users in making intelligent financial decisions. The framework integrates machine learning, deep learning, and explainable AI techniques into a unified architecture capable of analyzing historical financial transactions, forecasting future expenses, and generating personalized budgeting recommendations.

Unlike traditional financial management applications that simply record expenses, FinSenseAI actively predicts future spending behavior and identifies financial patterns by utilizing historical transaction records. The proposed architecture combines the strengths of XGBoost in handling structured tabular data and the sequential learning capabilities of LSTM networks for modeling temporal spending trends.

To increase user trust, SHAP is employed to explain each prediction by quantifying the contribution of individual financial attributes. Consequently, users receive not only an expense forecast but also a detailed explanation of the factors responsible for the prediction.

➢ *The Overall Workflow of the Proposed System Consists of:*

- Financial transaction collection
- Data preprocessing
- Feature engineering
- Hybrid machine learning prediction
- Explainable AI analysis
- Budget recommendation generation
- Visualization through interactive dashboards

➢ *Dataset Description*

The proposed framework utilizes a real-world personal finance transaction dataset consisting of daily financial records collected over several months. Each transaction represents either an income or an expense and contains multiple attributes describing financial activity.

- *The Dataset Includes the Following Attributes:*

Table 2 The Dataset Includes the Following Attributes:

Attribute	**Description**
Date	Transaction Date
Category	Expense or Income Category
Amount	Transaction Amount (₹)
Transaction Type	Income / Expense
Description	Additional Transaction Information

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

- *Example Transactions Include:*

Table 3 Example Transactions Include

Date	Category	Amount	Type
2020-01-02	Food & Drink	1485.69	Expense
2020-01-04	Rent	1185.08	Expense
2020-01-19	Other	3287.00	Income
2020-02-07	Travel	1979.15	Expense
2020-03-23	Investment	2385.00	Income
2020-04-26	Food & Drink	1938.30	Expense

- *The Dataset Contains Multiple Spending Categories Including:*

✓ Food & Drink
✓ Rent
✓ Utilities
✓ Shopping
✓ Travel
✓ Entertainment
✓ Health & Fitness
✓ Investment
✓ Salary
✓ Other

The collected transactions exhibit both sequential and nonlinear financial patterns, making them suitable for hybrid machine learning approaches. The target variable of the proposed model is the predicted future expense amount based on previous financial behavior.

➢ *Dataset Statistics*

The dataset consists of financial transactions distributed across multiple expense categories and income sources.

- *The Major Characteristics Include:*

✓ Daily transaction records
✓ Multiple expense categories
✓ Income and expense transactions
✓ Continuous numerical target variable (Amount)
✓ Temporal information
✓ Textual transaction descriptions

The dataset is appropriate for supervised regression because historical spending records are used to predict future financial expenditure.

➢ *Data Preprocessing*

Raw financial transaction datasets frequently contain inconsistencies that negatively influence machine learning performance. Therefore, a comprehensive preprocessing pipeline is employed before model training.

The preprocessing procedure consists of the following stages:

- *Step 1: Duplicate Removal*

Duplicate financial transactions are identified and removed to avoid biased model learning.

$$D_{clean} = D_{raw} - D_{duplicate}$$

Where

D_{raw}= Original Dataset

$D_{duplicate}$= Duplicate Transactions

- *Step 2: Missing Value Analysis*

Each attribute is examined for missing values.

Missing numerical attributes are replaced using mean or median imputation, whereas missing categorical values are replaced using the mode.

$$X_{missing} = Median(X)$$

- *Step 3: Date Conversion*

The transaction date is converted into multiple temporal attributes:

✓ Year
✓ Month
✓ Day
✓ Week Number
✓ Day of Week
✓ Quarter

▪ Example

2020-02-07
↓
Year = 2020
Month = 2
Day = 7
Week = 6
Day of Week = Friday

These features enable the model to capture periodic financial behavior.

- *Step 4: Categorical Encoding*

Machine learning algorithms require numerical input.

Therefore, categorical attributes are converted using One-Hot Encoding.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

✓ *Example: -*

Table 4 Example

Transaction	Category	Category Code	Type	Type Code
Lunch at Restaurant	Food & Drink	100000	Expense	0
Monthly Salary	Others (Income Source)	000001	Income	1
Bus Ticket	Travel	010000	Expense	0
Bought a T-shirt	Shopping	001000	Expense	0
Doctor Visit	Healthcare	000010	Expense	0
Netflix Subscription	Entertainment	000100	Expense	0
Freelance Project Payment	Others (Income Source)	000001	Income	1

- *Step 5: Feature Scaling*

The Amount attribute is normalized using Min-Max Scaling.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Where

X'= Normalized Amount

This prevents larger monetary values from dominating model learning.

- *Step 6: Outlier Detection*

Financial outliers are detected using the Interquartile Range (IQR).

$$IQR = Q_3 - Q_1$$

Lower Bound

$$Q_1 - 1.5(IQR)$$

Upper Bound

$$Q_3 + 1.5(IQR)$$

Extreme transaction values outside these limits are reviewed before model training.

- *Step 7: Feature Engineering*

Several new financial attributes are generated to improve prediction performance.

✓ *Examples include:*

Table 5 Examples Include

Engineered Feature	Description
Previous Expense	Last transaction amount
Monthly Expense	Monthly total expenditure
Weekly Expense	Weekly expenditure
Transaction Frequency	Number of transactions
Expense Ratio	Expense / Income
Savings Ratio	Savings percentage
Average Category Expense	Average spending per category
Weekend Indicator	Saturday/Sunday
Income Flag	Binary income indicator
Expense Flag	Binary expense indicator

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

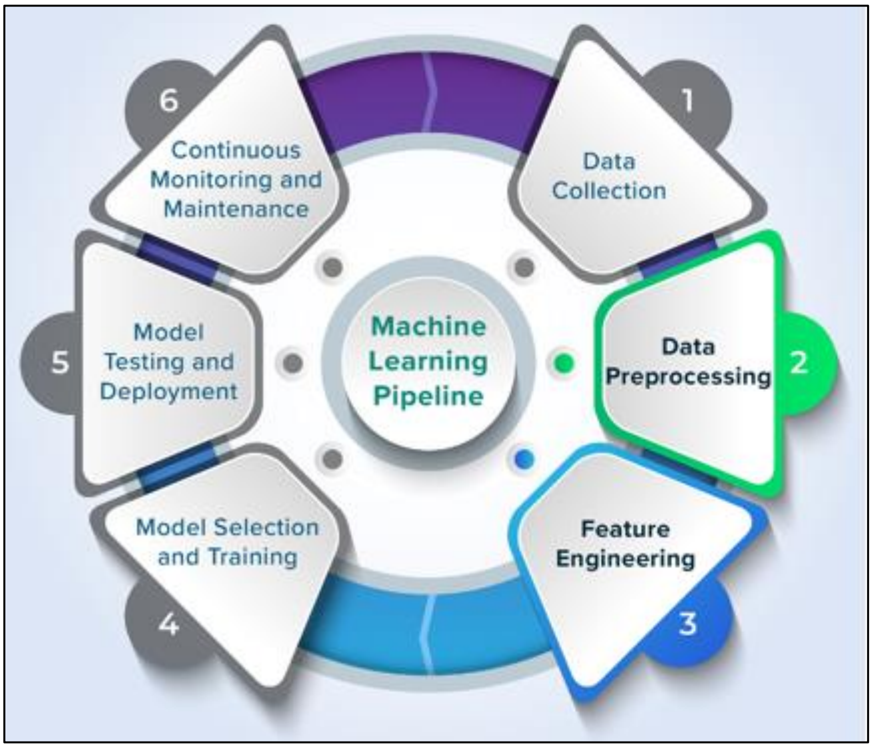

Fig 1 Data Preprocessing Pipeline

- *Algorithm 1*

✓ *Input*
  Raw Dataset

✓ *Output*
  Processed Dataset

✓ *Step 1*
  Remove duplicates

✓ *Step 2*
  Handle missing values

✓ *Step 3*
  Convert dates

✓ *Step 4*
  Encode categories

✓ *Step 5*
  Normalize data

✓ *Step 6*
  Generate engineered features

✓ *Step 7*
  Return processed dataset

IV. PROPOSED METHODOLOGY

The proposed FinSenseAI framework employs a hybrid machine learning architecture that combines the strengths of Extreme Gradient Boosting (XGBoost) and Long Short-Term Memory (LSTM) neural networks to accurately forecast future personal expenses. The framework further integrates SHapley Additive exPlanations (SHAP) to improve transparency by explaining how each financial attribute contributes to the generated prediction.

Unlike conventional expense tracking systems that merely record transactions, the proposed framework performs intelligent financial forecasting by learning nonlinear feature relationships and sequential spending behavior simultaneously.

➢ *The Overall Workflow of the Proposed Framework Consists of the following Stages:*

- Financial Transaction Collection
- Data Cleaning
- Data Preprocessing
- Feature Engineering
- XGBoost Training
- LSTM Training
- Hybrid Prediction
- SHAP Explainability
- Budget Recommendation
- Financial Dashboard Visualization

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

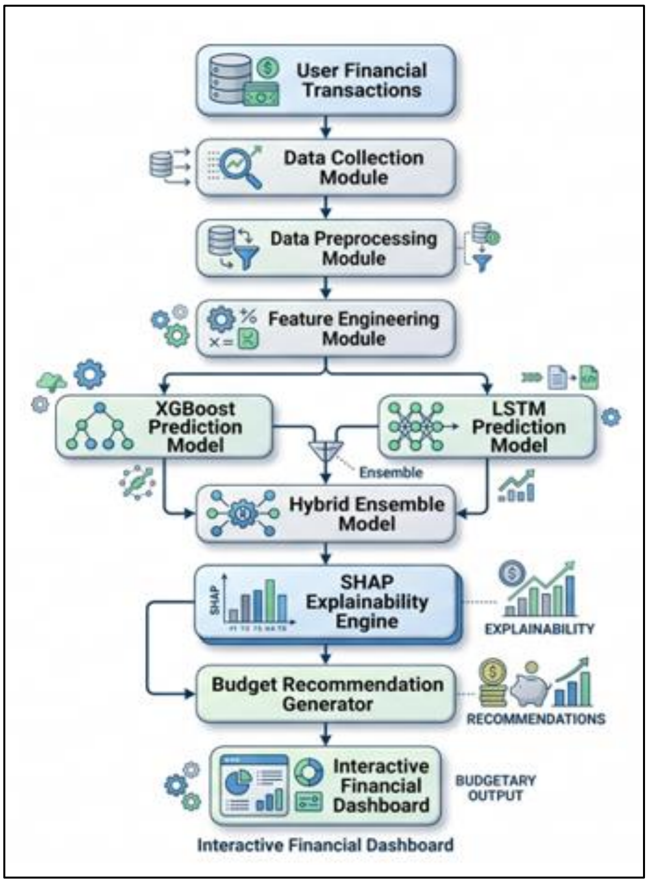

Fig 2 Overall Proposed Framework

Dataset
↓
Preprocessing
↓
Feature Engineering
↓
XGBoost
↓
LSTM
↓
Hybrid Ensemble
↓
SHAP
↓
Recommendation Engine
↓
Dashboard

➢ *Mathematical Formulation*

Let the financial transaction dataset be

$$D = \{x_1, x_2, x_3, \dots, x_n\}$$

Where

$$x_i = (Date, Category, Amount, TransactionType, Description)$$

Each transaction is represented as

$$x_i = (d_i, c_i, a_i, t_i, s_i)$$

Where

- d_i=Date
- c_i=Expense Category
- a_i=Amount
- t_i=Income/Expense
- s_i=Transaction Description

The objective is to predict

$$\hat{y} = f(X)$$

Where

$\hat{y}$ – represents the predicted future expense amount.

➢ *Feature Matrix*

The feature matrix is represented as

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots \\ x_{21} & x_{22} & x_{23} & \cdots \\ \vdots & \vdots & \vdots & \\ x_{n1} & x_{n2} & x_{n3} & \cdots \end{bmatrix}$$

Target vector

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

➢ *XGBoost Model*

Extreme Gradient Boosting (XGBoost) is employed to learn nonlinear financial patterns from structured transaction data.

XGBoost constructs multiple decision trees sequentially, where each new tree attempts to minimize the prediction error generated by previous trees.

- *Objective Function*

$$Obj = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

Where

L - represents loss function

&

Ω - represents regularization.

Regularization reduces overfitting.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

- *Time Complexity*

O(K × N × log N)

- *Space Complexity*

O(N × Features)

- *XGBoost Hyperparameters*

Table 6 XGBoost Hyperparameters

Parameter	Value
Max Depth	6
Learning Rate	0.1
Estimators	200
Subsample	0.8
Objective	Regression
Evaluation Metric	RMSE

- *Algorithm 1. XGBoost*

✓ *Input:*
Financial Dataset

✓ *Output:*
Predicted Expense

✓ *Step 1*
Load Dataset

✓ *Step 2*
Preprocess Dataset

✓ *Step 3*
Generate Features

✓ *Step 4*
Split Dataset

✓ *Step 5*
Train XGBoost

✓ *Step 6*
Predict Test Data

✓ *Step 7*
Calculate RMSE

- *LSTM Equations*

Forget Gate

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

Input Gate

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

Candidate State

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C)$$

Cell State

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Output Gate

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

Hidden State

$$h_t = o_t * \tanh(C_t)$$

- *Activation Function*
ReLU

- *Output Layer*
Dense(1)

- *Optimizer*
Adam

- *Learning Rate*
0.001

These equations describe how LSTM selectively remembers and forgets historical transaction information.

- *LSTM Parameters*

Table 7 LSTM Parameters

Parameter	Value
Hidden Units	64
Dropout	0.20
Optimizer	Adam
Epochs	50
Batch Size	16
Loss Function	MSE

- *Algorithm 2. LSTM*

✓ *Input*
Transaction Sequence

✓ *Output*
Future Expense

Load Dataset

Normalize

Create Time Window

Build LSTM

Train Model

Predict Future Expense

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

Return Prediction

➢ *Hybrid Machine Learning Model*

The final prediction combines XGBoost and LSTM.

$$Prediction = \frac{Prediction_{XGB} + Prediction_{LSTM}}{2}$$

The ensemble reduces prediction variance while preserving temporal learning capability.

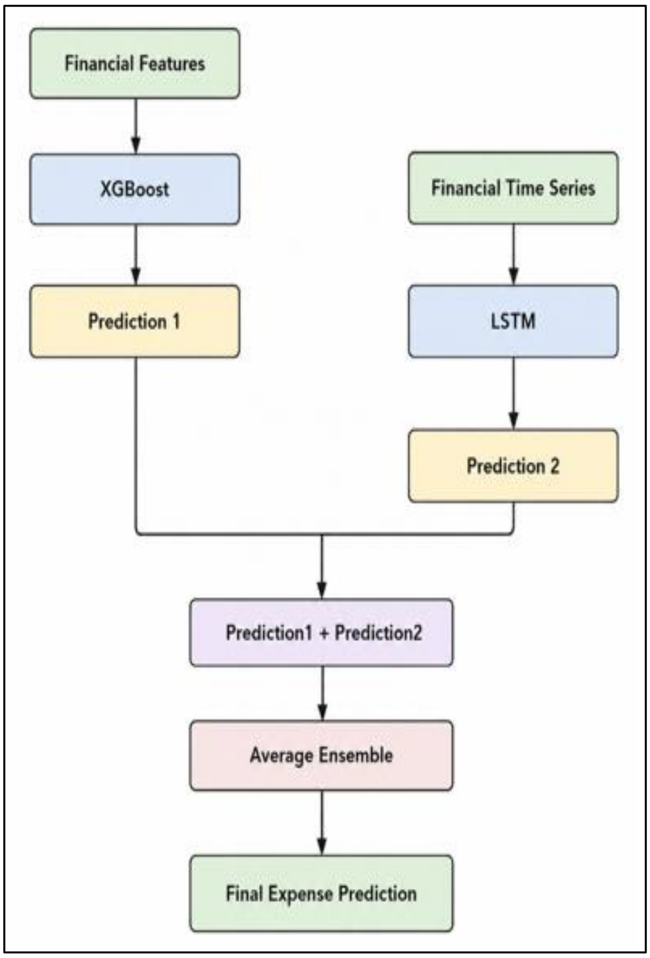

Fig 3 Hybrid Model

- *Advantages:*

✓ Better Generalization
✓ Reduced Overfitting
✓ Improved Forecasting
✓ Handles Sequential Data
✓ Learns Nonlinear Relationships
✓ Improved Prediction Stability

➢ *Explainable Artificial Intelligence (SHAP)*

Financial predictions must be interpretable.

Therefore SHAP is integrated.

SHAP computes

$$Prediction = BaseValue + \sum SHAP_i$$

Each SHAP value measures the contribution of one feature.

- *Example*

Table 8 Example

Feature	SHAP Effect
Rent	+320
Salary	-180
Shopping	+140
Travel	+90

Positive values increase predicted expenditure.

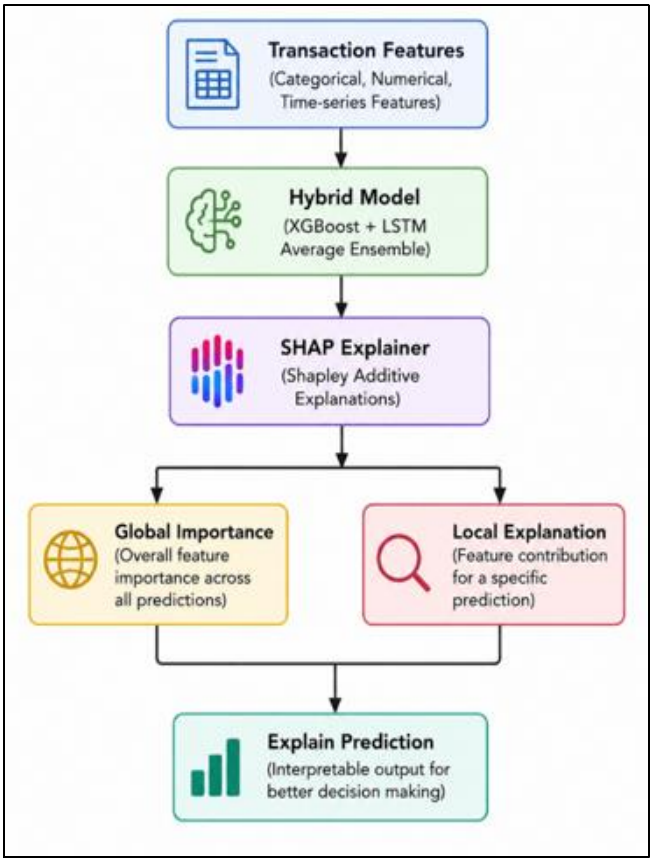

Fig 4 SHAP Workflow

Negative values decrease expenditure.

- *Global Feature Importance:*

✓ Local Explanation
✓ Waterfall Plot
✓ Force Plot
✓ Summary Plot
✓ Dependence Plot

➢ *Experimental Setup*

The proposed FinSenseAI framework was implemented using Python to evaluate the effectiveness of the hybrid XGBoost–LSTM model for personal expense prediction. The implementation includes data preprocessing, feature engineering, model training, prediction, and explainability analysis.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

Python IDE

VS Code

Random State

42

Train

80%

Validation

10%

Test

10%

- *Hardware Configuration*

Table 9 Hardware Configuration

Component	Specification
Processor	Intel Core i5 (12th Gen) or above
RAM	8 GB DDR4
Storage	512 GB SSD
Operating System	Windows 11
GPU	Not Required

- *Software Configuration*

Table 10 Software Configuration

Software	Version
Python	3.11
Jupyter Notebook	Latest
Pandas	Latest
NumPy	Latest
Scikit-learn	Latest
XGBoost	Latest
TensorFlow	Latest
SHAP	Latest
Matplotlib	Latest

- *Train-Test Split*

The dataset was divided using an 80:20 ratio, where 80% of the transaction records were used for training and 20% were reserved for testing. A chronological split was used to preserve the temporal nature of financial transactions.

➢ *Evaluation Metrics*

The predictive performance of the proposed model will be evaluated using standard regression metrics.

Root Mean Squared Error (RMSE)

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

RMSE measures the average magnitude of prediction error. Lower values indicate better performance.

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$$

MAE represents the average absolute difference between predicted and actual expenses.

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} | \frac{y_i - \hat{y}_i}{y_i} |$$

MAPE expresses the prediction error as a percentage.

Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

R^2 measures how well the model explains the variance in the target variable. Values closer to 1 indicate a better fit.

V. RESULTS AND DISCUSSION

The proposed FinSenseAI framework was designed to evaluate the effectiveness of hybrid machine learning techniques for intelligent personal expense prediction. The experimental workflow includes data preprocessing, feature engineering, model training, prediction, and explainability analysis. Multiple regression models, including Linear Regression, Random Forest, XGBoost, Long Short-Term Memory (LSTM), and the proposed Hybrid XGBoost–LSTM model, are intended to be evaluated using standard regression performance metrics.

The evaluation metrics selected for performance assessment include Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R^2 Score). These metrics were chosen because they comprehensively measure prediction accuracy, average prediction error, percentage error, and the goodness-of-fit of the regression models.

The proposed evaluation framework enables a systematic comparison of conventional machine learning algorithms with the hybrid deep learning architecture. In addition to numerical evaluation metrics, model interpretability is incorporated through SHAP (SHapley Additive exPlanations) to identify the contribution of individual financial features toward predicted expense values.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

VI. CONCLUSION

In this research, an Explainable Artificial Intelligence (XAI)-based personal finance recommendation and expense prediction framework named FinSenseAI has been proposed to improve intelligent financial planning and decision-making. The framework integrates the predictive capabilities of Extreme Gradient Boosting (XGBoost) and Long Short-Term Memory (LSTM) into a hybrid machine learning architecture while employing SHapley Additive exPlanations (SHAP) to provide transparent and interpretable predictions.

The proposed system utilizes historical financial transaction records consisting of transaction date, category, amount, transaction type, and transaction description. A comprehensive data preprocessing pipeline involving duplicate removal, missing value handling, categorical encoding, normalization, temporal feature extraction, and feature engineering was applied to improve model performance. XGBoost effectively models nonlinear relationships among financial attributes, whereas LSTM captures temporal spending patterns from sequential transaction data. The integration of SHAP enhances model transparency by identifying the contribution of individual features toward each prediction, thereby increasing user confidence in AI-assisted financial recommendations.

The proposed framework aims to support users in predicting future expenses, understanding spending behavior, and receiving personalized budgeting recommendations. The modular design allows future integration with cloud services, banking APIs, and mobile applications. FinSenseAI demonstrates the potential of combining machine learning, deep learning, and explainable AI to develop practical, intelligent, and transparent personal finance management systems.

The proposed framework demonstrates that integrating machine learning, deep learning, and explainable artificial intelligence can improve intelligent financial planning while increasing user trust through transparent decision-making.

FUTURE WORK

Although the proposed FinSenseAI framework provides an effective approach for intelligent expense prediction and recommendation, several enhancements can further improve its performance and practical applicability.

➢ *Future Research Directions Include:*

- Expanding the dataset by incorporating transactions from multiple users over longer periods to improve model generalization.
- Integrating secure banking APIs to enable automatic synchronization of financial transactions.
- Developing Android and iOS mobile applications for real-time financial monitoring and personalized budgeting.
- Investigating advanced deep learning architectures such as Transformers and Temporal Fusion Transformers for time-series forecasting.
- Applying Federated Learning techniques to improve user privacy while enabling collaborative model training.
- Incorporating Natural Language Processing (NLP) to analyze transaction descriptions and generate more detailed financial insights.
- Developing interactive Explainable AI dashboards that visualize SHAP feature importance for end users.
- Comparing the proposed framework with additional ensemble learning and deep learning algorithms using larger benchmark datasets.

These enhancements have the potential to improve prediction accuracy, scalability, interpretability, and practical deployment in intelligent financial management applications.

Future work also includes integrating Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), multimodal financial analysis, and reinforcement learning to provide adaptive and conversational financial assistants.

ACKNOWLEDGEMENT

The authors express their sincere gratitude to Dr. R. Sathya Janaki, M.Sc., M.E., Ph.D., TNSET Qualified, Associate Professor and Head, Department of Artificial Intelligence and Machine Learning, Takshashila University, for her invaluable guidance, continuous encouragement, technical expertise, and constructive suggestions throughout this research work. Her mentorship and support were instrumental in the successful completion of this study. The authors also thank Takshashila University for providing the academic environment and resources necessary for conducting this research.

REFERENCES

[1]. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.

[2]. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[3]. S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.

[4]. J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

[5]. L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[6]. F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[7]. M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2016.

[8]. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul842

[9]. A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed., O'Reilly Media, 2022.
[10]. F. Chollet, *Deep Learning with Python*, 2nd ed., Manning Publications, 2021.
[11]. G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, 2nd ed., Springer, 2021.
[12]. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed., Springer, 2009.
[13]. W. McKinney, *Python for Data Analysis*, 3rd ed., O'Reilly Media, 2022.
[14]. S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent Individualized Feature Attribution for Tree Ensembles," 2018.
[15]. R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed., MIT Press, 2018.
[16]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436–444, 2015.
[17]. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
[18]. K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012.
[19]. C. C. Aggarwal, *Neural Networks and Deep Learning*, Springer, 2018.
[20]. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, 2012.
[21]. S. Raschka and V. Mirjalili, *Python Machine Learning*, 3rd ed., Packt Publishing, 2019.
[22]. T. M. Mitchell, *Machine Learning*, McGraw-Hill, 1997.
[23]. L. Bottou, "Large-Scale Machine Learning with Stochastic Gradient Descent," *Proceedings of COMPSTAT*, 2010.
[24]. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *International Conference on Learning Representations (ICLR)*, 2015.
[25]. D. Silver *et al.*, "Mastering the Game of Go with Deep Neural Networks and Tree Search," *Nature*, vol. 529, pp. 484–489, 2016.

AUTHOR BIOGRAPHIES

➢ *Velkur Dilip*

Velkur Dilip is currently pursuing a Bachelor of Technology (B.Tech.) in Artificial Intelligence and Machine Learning at Takshashila University, Ongur, Tindivanam, Tamil Nadu, India. His research interests include Artificial Intelligence, Machine Learning, Explainable Artificial Intelligence (XAI), Data Science, Financial Technology (FinTech), Predictive Analytics, and Intelligent Decision Support Systems. He has completed internships in Artificial Intelligence, Machine Learning, Java Full Stack Development, and Web Development. He actively participates in research activities, technical workshops, hackathons, innovation projects, and AI-based application development. His current research focuses on explainable machine learning models for financial forecasting, intelligent recommendation systems, and real-world AI applications.

➢ *Dr. R. Sathya Janaki*

Dr. R. Sathya Janaki, M.Sc., M.E., Ph.D., TNSET Qualified, is the Associate Professor and Head of the Department of Artificial Intelligence and Machine Learning at Takshashila University, Ongur, Tindivanam, Tamil Nadu, India. She has extensive teaching, research, and academic experience in the fields of Artificial Intelligence, Machine Learning, Deep Learning, Data Science, Computer Vision, and Intelligent Systems. She has guided numerous undergraduate research projects and actively mentors students in innovative research and emerging AI technologies. Her research interests include Explainable Artificial Intelligence, Data Analytics, Intelligent Healthcare Systems, Predictive Analytics, and AI-driven Decision Support Systems.