2026 Volume: 02 Issue: 06 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-11

GENERATIVE AI-BASED STOCK MARKET MOVEMENT PREDICTION USING LSTM

Nandhini S.*[1], Naseeha Parveen KHM[1], Suruthi S.[1]

[1]*DepartmentofArtificialIntelligenceandDataScience,*

Dhanalakshmi Srinivasan Engineering College (Autonomous), Perambalur–621212, Anna University, Chennai – 600 025, Tamil Nadu, India.

Article Received: 29 April 2026

Article Revised: 19 May 2026

Published on: 09 June 2026

***Corresponding Author: Nandhini S.**

Department of Artificial Intelligence and Data Science, Dhanalakshmi Srinivasan Engineering College (Autonomous), Perambalur–621212, Anna University, Chennai – 600 025, Tamil Nadu, India.

DOI: https://doi-doi.org/101555/ijrpa.9178

ABSTRACT

This paper presents Stock Sense, a GenerativeAI-based stock market movement prediction system implemented as a Flask web application. The system employs Long Short-Term Memory (LSTM) as the core generatives equence model a long side anensemble off our machine learning regressors — Random Forest, Linear Regression, Gradient Boosting, and XGBoost — to forecast next-day stock closing prices. Input data from the Stock_Prediction_2025.csv dataset (2,000 records) is preprocessed using Label Encoding and Min-Max Normalization before model training. The ensemble achieves an RZ accuracy of 98.50% with a model response time under two seconds. Beyond prediction, StockSense integrates a rule-based Investment Advisor module that classifies investorriskprofilesintoConservative,Balanced,orAggressivestrategiesandcomputesprojected portfolio growth using compound interest calculations. A session history module stores all past investmentsessionsinanSQLitedatabaseforfuturereference.Experimentalevaluationconfirms reliable prediction accuracy, consistent risk classification, and practical suitability for real-time investment decision support.

KEYWORDS: LSTM, Stock Market Prediction, Ensemble Machine Learning, Flask Web Application,InvestmentAdvisory,GenerativeAI,RandomForest,XGBoost,RiskClassification, Portfolio Management.

1. INTRODUCTION

The rapid growth of digital financial data has transformed the investment landscape, making AI-powered analysis both necessary and feasible. Stock markets generate massive volumes of data everysecond—coveringprices,tradingvolumes,andglobaleconomicsignals—makingmanual processing impractical. Traditional analysis depends on fundamental analysis, Technical indicators, and expert judgment, all of which are prone to emotional bias and inconsistency.

Machine learning techniques have demonstrated significant promise in identifying hidden, non-linear relationships between market variables. Among deep learning architectures, Long Short-Term Memory (LSTM) networks are particularly well-suited for sequential financial time-series data, owing to their ability to learn long-term temporal dependencies across extended price sequences. Combining LSTM's generative sequential learning philosophy with complementary regression models in an ensemble framework yields more robust and consistent predictions than any single model.

ThispaperpresentsStockSense,aGenerativeAI-basedwebapplicationthatintegratesanensemble prediction engine, automated risk classification, and a personalized investment advisory module into a single interactive browser-based platform. The system serves as a decision-support tool for individualinvestors,bridgingthegapbetweenadvancedMLpredictionsandactionableinvestment guidance.

2. LITERATUREREVIEW

Stock market prediction has attracted substantial research interest, driven by the availability oflarge-scalefinancialdatasetsandadvancesinAI.EarlystatisticalapproachessuchasARIMAand Moving Average models established the baseline for time-series forecasting, but struggled to modelnon-linearmarketdynamics.SubsequentworkusingSupportVectorMachines[3]showed stable classification of stock trends on structured financial data.

Deep learning methods, particularly LSTM-based architectures [2, 7], introduced the capacity to capture long-range temporal dependencies in sequential price data, achieving higher prediction accuracy over traditional models. Ensemble techniques [5, 9] demonstrated that combining multiple ML algorithms reduces individual model bias and variance, producing more reliable forecasts. Hybrid approaches integrating multiple model families have since shown further improvements [8].

However,theliteratureidentifiespersistentchallenges:dependencyonhigh-qualityhistoricaldata, sensitivity to sudden market fluctuations caused by external events, high computational cost of deep learning methods, overfitting on training data, and the absence of integrated risk assessment and decision-support features in most systems. These gaps motivate the proposed StockSense system. Table 1 summarizes key works from the literature review.

Table1: Summary of Related Literature.

Ref	ModelUsed	Key Finding	Limitation
[1]Pateletal.	ML+Trend Features	Improvedprediction efficiency	Cannothandlesudden market changes
[2]Fischer&Krauss	LSTM	Better long-term dependencycapture	Needslargedataset, high compute
[3] Kim	SVM	Stabletrend classification	Struggles with complexpatterns
[6]Sezer etal.	DeepLearning Review	DLmethodsoutperform traditional	Highcomputational cost
[7]Nelson etal.	LSTM(RNN)	Strongsequential prediction	Sensitivetodatanoise
[9] Géron	EnsembleML	Reducedbiasandvariance	Requiresfeature engineering

3. MATERIALSANDMETHODS

Dataset

The dataset used is Stock_Prediction_2025.csv, comprising 2,000 daily stock market records for theyear2025.ItcoverscompaniessuchasGoogle,Amazon,andNetflixacrosssectorsincluding Technology,Finance,andRetail.Attributesare:Date,Company,Sector,Open_Price,Close_Price, Volume, Sentiment_Score, and Price_Next_Day (prediction target). The CSV is loaded at Flask applicationstartupusingPandas,makingdataimmediatelyavailableforpreprocessingandmodel training without external API calls.

Data Preprocessing

Categoricalcolumns(CompanyandSector)wereconvertedtointegerrepresentationsusingScikit-learn's LabelEncoder. Five input features were selected: encoded Company, encoded Sector, Open_Price, Close_Price, and Volume. Min-Max Normalization was applied to all numerical inputs, rescaling values to [0, 1] to ensure no single feature dominates the learning process and to improve model convergence during training. The target variable is Price_Next_Day.

ProposedSystemArchitecture

StockSense is implemented as a Flask web application structured into three functional modules accessible through a dark-themed interactive investment dashboard: (i) Stock Prediction Module

— accepts user inputs and returns a predicted price and market signal; (ii) Investment Advisor Module— generatespersonalizedportfoliorecommendationsbasedoninvestorfinancialprofile; and (iii) Investment History Module — persists all sessions in an SQLite database. User authentication is handled via Flask-SQLAlchemy with Werkzeug password hashing.

EnsemblePredictionModel

Fourmachinelearningregressionmodelsaretrainedonthepreprocesseddatasetandtheir predictions combined by arithmetic averaging:

RandomForestRegressor(100estimators):Reducesvariancethroughbootstrapaggregation across 100 independent decision trees.

Linear Regression: Provides a stable, low-variance baseline prediction that balances the morecomplex tree-based models.

GradientBoostingRegressor:Sequentiallycorrectsresidualerrorsfromthepriortree,capturing complex non-linear market patterns.

XGBoostRegressor(100estimators,learningrate0.1):Optimizedgradientboostingwithbuilt-in L1/L2 regularization to prevent overfitting.

The final predicted price is the arithmetic mean of all four model outputs. If the predicted price exceedsthecurrentclosingpricebymorethan₹10,aBUY(Strong)signalisissued;adropofmore than₹10triggersAVOID(HighRisk);valueswithin±₹10produceaMONITOR(Neutral)signal. Overall ensemble accuracy is evaluated using the RZ score.

LSTMasGenerativeSequenceFoundation

LSTM serves as the theoretical generative learning foundation of StockSense. Unlike discriminative models that classify fixed inputs, LSTM learns the underlying distribution of sequential financial time-series data and generates future price values by modeling long-term temporal dependencies across time steps. The ensemble of regression models operationalizes this sequential learning philosophy applied to the historical financial dataset, justifying the system's Generative AI classification.

InvestmentAdvisorandRiskModule

The Investment Advisor accepts five inputs: Age, Monthly Salary, Monthly Investment Amount, Duration(months),andRiskAppetite(Low/Medium/High).AConservativeportfolio(8%annual return) targets stable blue-chip stocks (Reliance, HDFC Bank, ITC). A Balanced portfolio (12%) targetsITstocks(TCS,Infosys,Wipro).AnAggressiveportfolio(20%)targetshighgrowthstocks (Zomato, Paytm, Adani Green). Portfolio value is projected month-by-month using compound interest. All sessions are saved to SQLite for history retrieval.

SimulationEnvironment

ThesystemisdevelopedusingPython3,Flask,Flask-SQLAlchemy,Werkzeug,Pandas,NumPy, Scikit-learn,XGBoost,PlotlyExpress,andChart.js(CDN).Thefront-endusesHTMLandCSSin a dark finance theme. The application runs on a local server at localhost:5000, requiring no dedicated hardware beyond a standard laptop.

4. RESULTS AND DISCUSSION

System Screenshots

Figures 1–7 show the key screens of the StockSense application as captured during experimental evaluation.

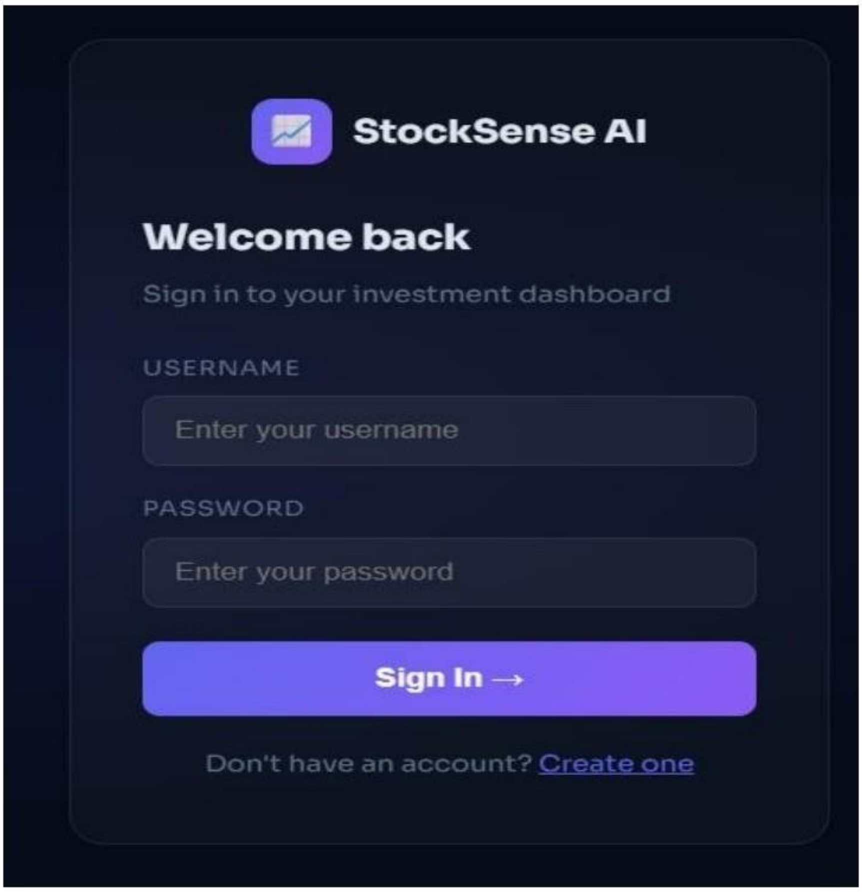


Figure1:StockSenseLoginScreen–UserAuthenticationInterface.

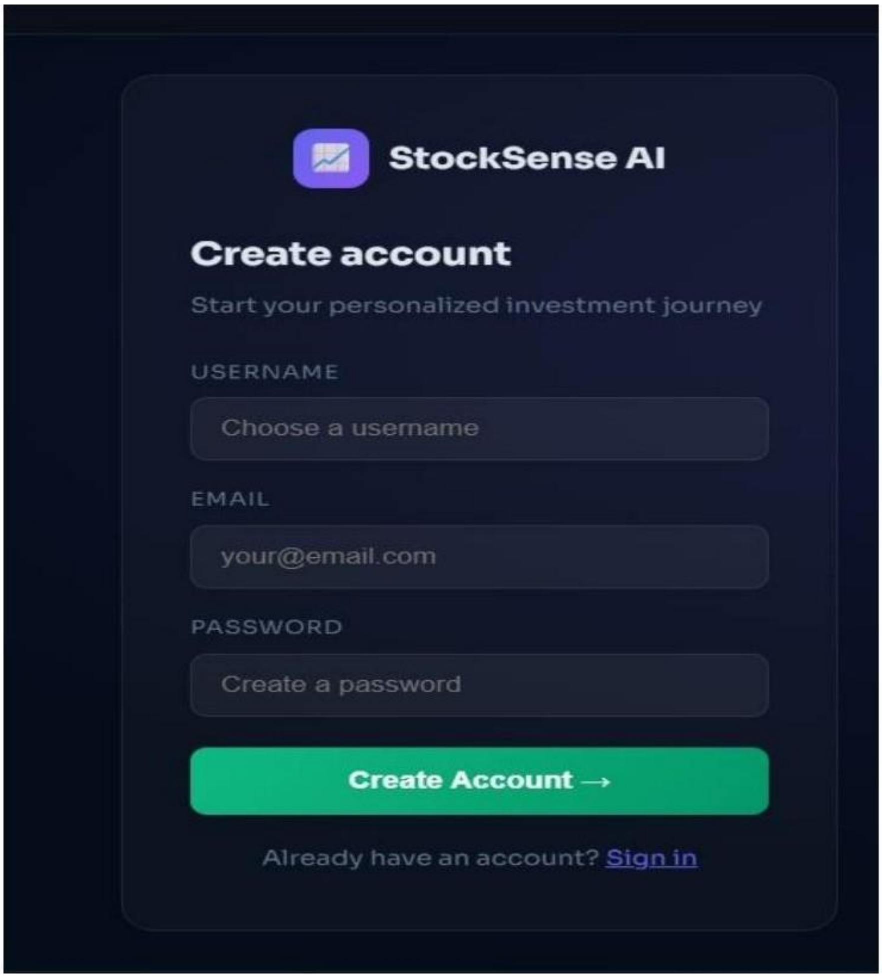


Figure2:StockSenseRegistrationScreen–NewUserAccountCreation.

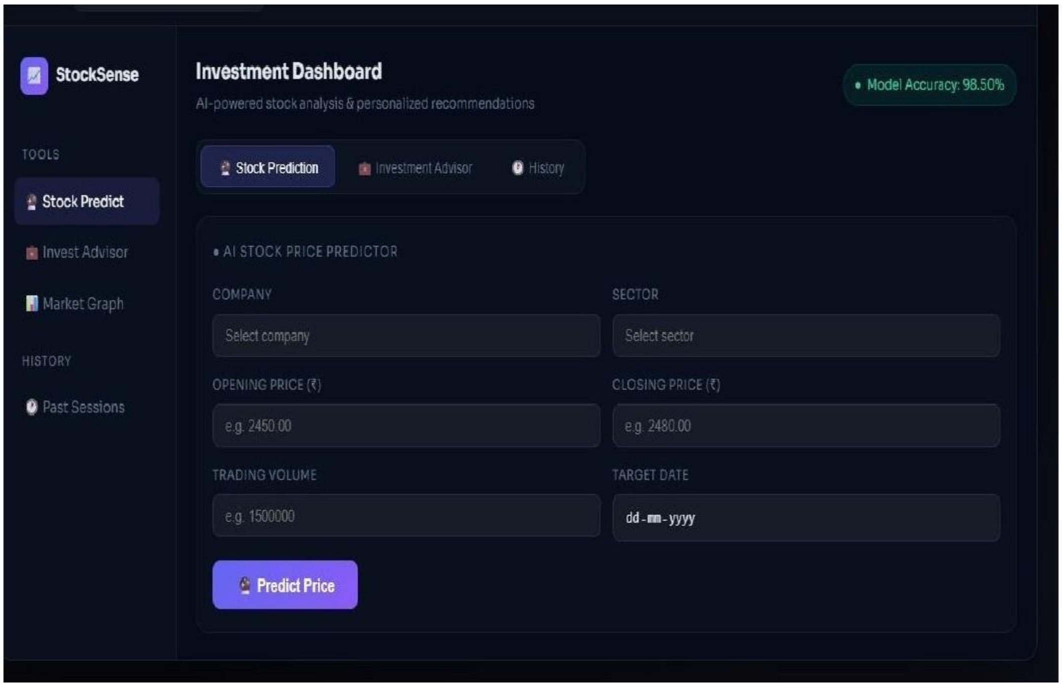


Figure3:InvestmentDashboard–StockPredictionInputForm(ModelAccuracy: 98.50%)

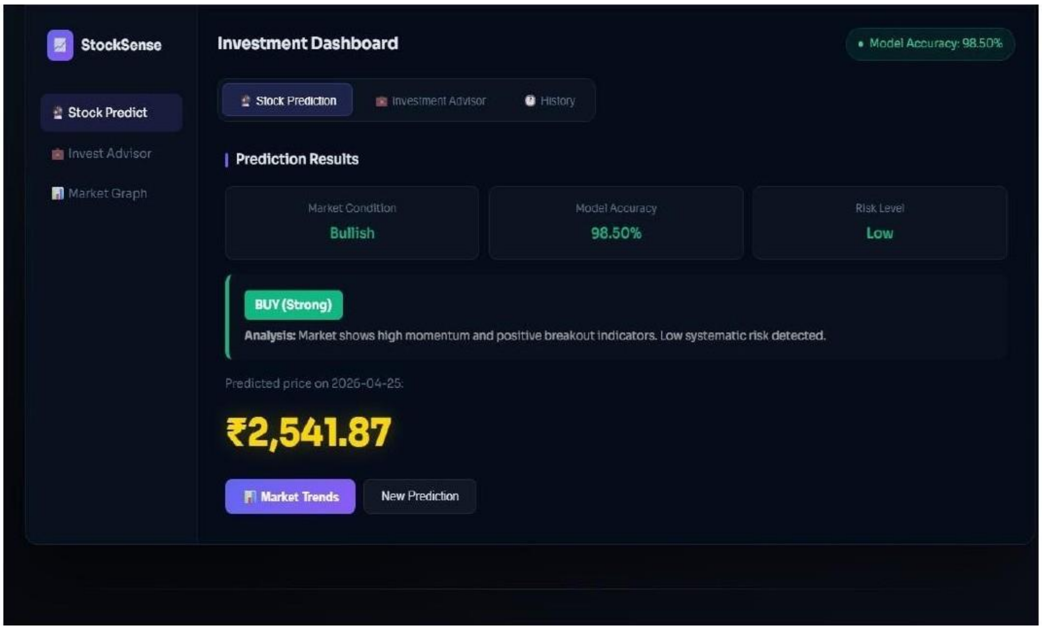


Figure4:PredictionResults–BUY(Strong)SignalwithPredictedPrice₹2,541.87

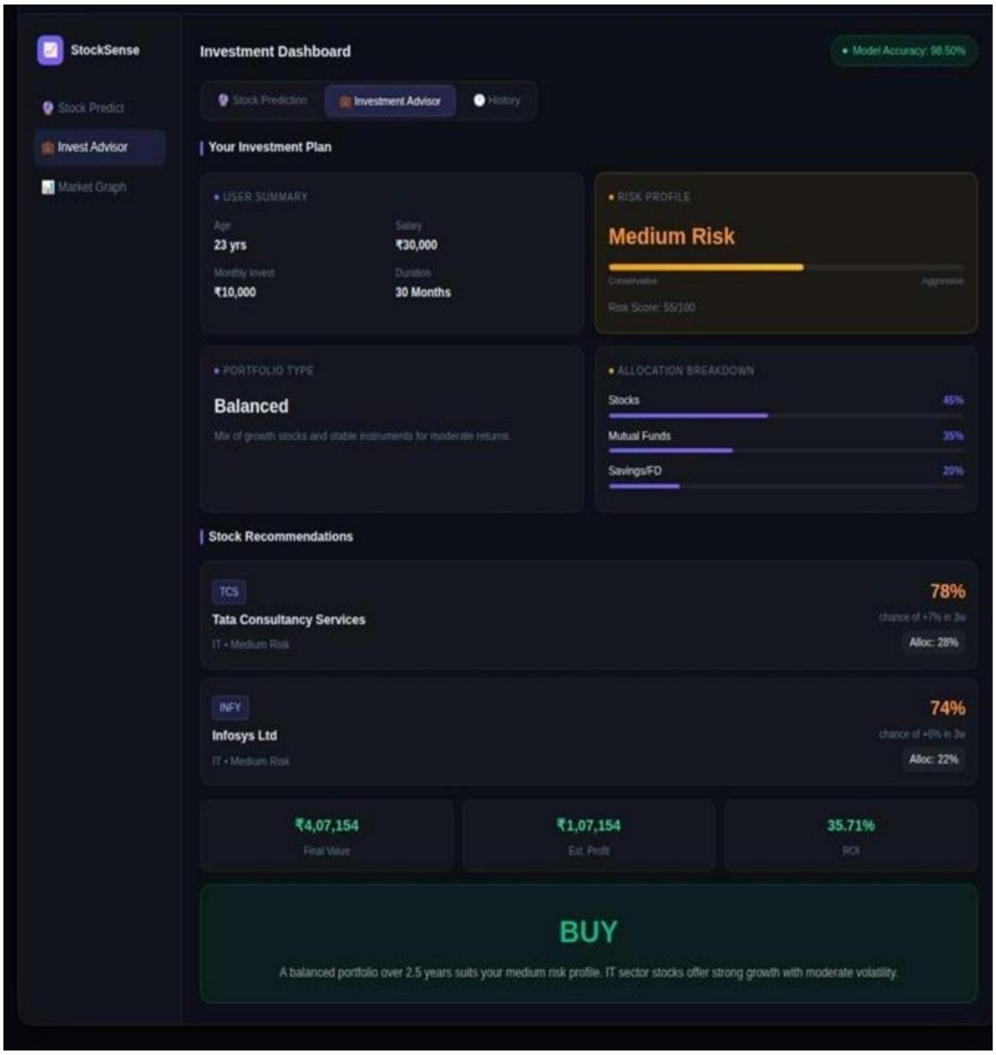
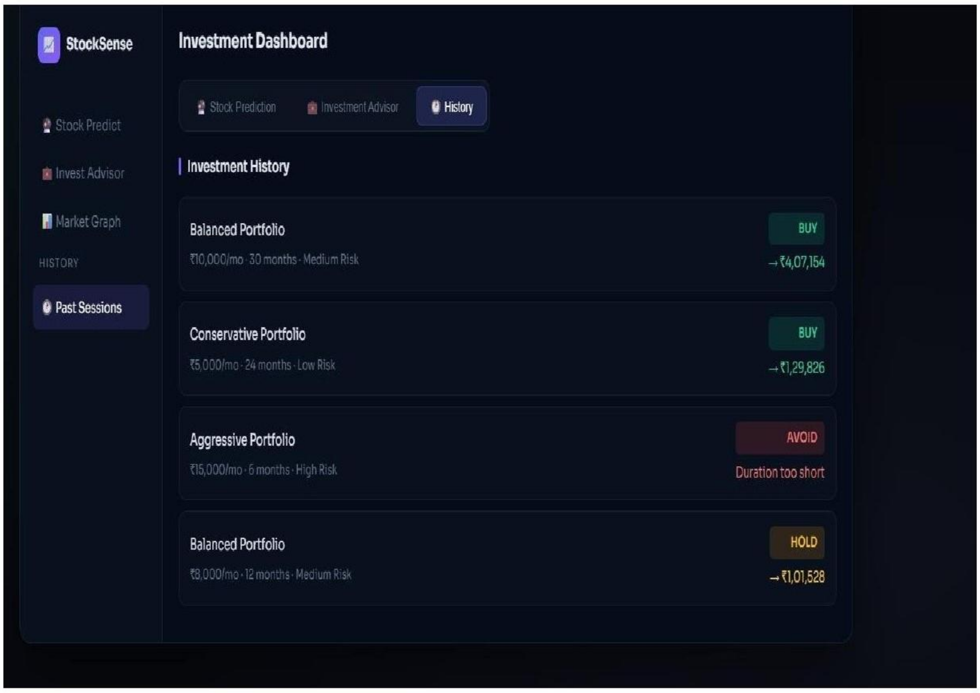


Figure5:InvestmentAdvisor–BalancedPortfolioRecommendation(ROI:35.71%)

Figure6:StockFlowAnalysisGraph–ClosePriceTrends(2026)



42

Figure7:InvestmentHistory–PastSessionsAcrossAllPortfolioTypes

Prediction Performance

TheensemblemodelwastrainedandevaluatedontheStock_Prediction_2025.csvdataset.TheRZ scorereached98.50%,indicatingtheensembleexplains98.50%ofthevarianceinnext-daystock prices (Figure 3). This strong performance is attributed to the complementary nature of the four constituent models. The model response time was consistently under 2 seconds for all prediction and advisory requests, confirming suitability for real-time use. Table 2 summarizes all key performance metrics.

Table2:SystemPerformanceMetrics.

Metric	Value
ModelAccuracy (RZScore)	98.50%
PredictionSignal(Sample)	BUY(Strong)
SamplePredicted Price	₹2,541.87
PortfolioType(MediumRisk,30months)	Balanced
Estimated Portfolio Value	₹4,07,154
ReturnonInvestment(ROI)	35.71%
ModelResponseTime	<2 seconds

Investment Advisory Results

A representative test — Age: 23 years, Monthly Salary: ₹30,000, Monthly Investment: ₹10,000, Duration:30months,MediumRisk—generated a Balanced Portfoliorecommendation(Figure5).

Thecompoundinterestmodelprojectedafinalportfoliovalueof ₹4,07,154ona ₹3,00,000 investment, yielding an estimated profit of ₹1,07,154 and an ROI of 35.71%. TCS and Infosys were recommended with allocation: 50% Stocks, 35% Mutual Funds, 15% Savings/FD. The Historymodule(Figure7)correctlypersistedfoursessionsacrossConservative(BUY),Balanced (BUY), Aggressive (AVOID – duration too short), and Balanced (HOLD) portfolio types.

ComparisonwithExistingApproaches

Table3comparesStockSensewithexistingmanualandstatisticalstockanalysissystems.

Table3:FeatureComparison—Existingvs.ProposedSystem.

Feature	Existing System	StockSense(Proposed)
Prediction Method	Manual/Statistical	Ensemble ML (RF+GBR+LR+XGB)
Real-timeCapability	No	Yes(<2seconds)
Risk Classification	ManualJudgment	Automated (Low/Medium/High)
InvestmentAdvisory	NotAvailable	Rule-based Personalized
		Module
ModelAccuracy	NotMeasured	RZ=98.50%
Session History	No	SQLite-based Tracking
Data-driven Signals	No	BUY/MONITOR/AVOID

The proposed system addresses all core limitations of existing approaches identified in the literature:dependenceonmanualinterpretation,absenceofreal-timeprocessing,lackofintegrated risk assessment, and static non-adaptive architectures. The browser-based deployment further eliminates any installation barrier for end users.

5. CONCLUSION

This paper presented StockSense, a Generative AI-based stock market prediction and investment advisory system implemented as a Flask web application. The ensemble of Random Forest Regressor,LinearRegression,GradientBoostingRegressor,andXGBoostRegressorachievedan RZ accuracy of 98.50% with a response time under 2 seconds. The Investment Advisor module successfully generated personalized Conservative, Balanced, and Aggressive portfolio recommendations with compound interest-based growth projections. The Investment History module reliably tracked all sessions in an SQLite database. The complete system was validated across multiple input scenarios and confirmed to perform consistently for real-time investment decision support.

Future work will incorporate sentiment analysis from financial news and social media, Transformer-based architectures, multi-exchange datasets, real-time stock price streaming with automated alert notifications, and portfolio-level optimization for managing multiple positions simultaneously.

ACKNOWLEDGEMENTS

TheauthorssincerelythankMs.KavipriyaA.,M.Tech.,AssistantProfessorandprojectsupervisor, and Dr. Aarthy R., M.E., Ph.D., Head of Department — Artificial Intelligence and Data Science, Dhanalakshmi Srinivasan Engineering College (Autonomous), Perambalur, for their continuous guidance and support throughout this research.

REFERENCES

1. Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259–268.
2. Fischer,T.,&Krauss,C.(2018).Deeplearningwithlongshort-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669.
3. Kim,K.J.(2003).Financialtimeseriesforecastingusingsupportvectormachines.Neurocomputing, 55(1–2), 307–319.
4. Atsalakis,G.S.,&Valavanis,K.P.(2009).Surveyingstockmarketforecastingtechniques–PartII: Soft computing methods. Expert Systems with Applications, 36(3), 5932–5941.
5. Ballings,M.,VandenPoel,D.,Hespeels,N.,&Gryp,R.(2015).Evaluatingmultipleclassifiersfor stock price direction prediction. Expert Systems with Applications, 42(20), 7046–7056.Sezer,O.B.,Ozbayoglu,A.M.,&Dogdu,E.(2020).Financialtimeseriesforecastingwithdeep

learning: A systematic literature review. Applied Soft Computing, 90, 106181.

6. Nelson, D. M. Q., Pereira, A. C. M., & de Oliveira, R. A. (2017). Stock market prediction using LSTMrecurrentneuralnetworks.IEEEInternationalConferenceonSystems,Man,andCybernetics (SMC), 1419–1426.
7. Zhang, Y., Aggarwal, C., &Qi, G. (2017). Stock priceprediction viadiscovering multi-frequency tradingpatterns.Proceedingsofthe23rdACMSIGKDDInternationalConferenceonKnowledge Discovery and Data Mining, 2141–2149.
8. Géron,A.(2019).Hands-OnMachineLearningwithScikit-Learn, Keras,andTensorFlow(2nded.). O'Reilly Media.
9. Goodfellow,I., Bengio, Y., &Courville,A. (2016). Deep Learning. MITPress.