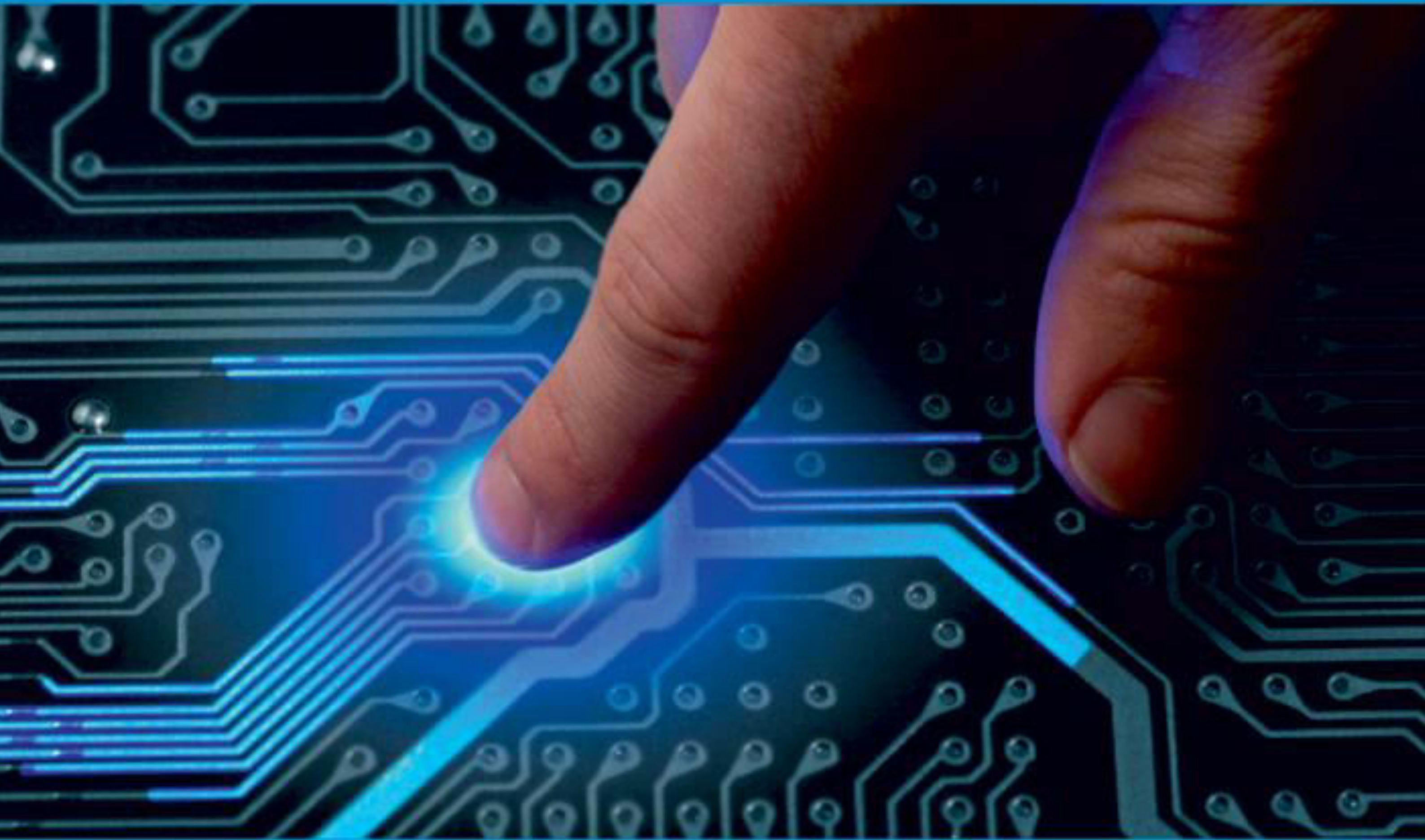
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 6, June 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1206008 |

FinExcellence: Financial Platform for Prediction of Financial Skills Using Machine Learning

Varun Madhav, Rohan Mathur, Shivam Mishra, Rohan Choudhary, Priya Singh

U.G. Student, Department of Computer Science & Engineering, JSS Academy of Technical Education, Noida, India

U.G. Student, Department of Computer Science & Engineering, JSS Academy of Technical Education, Noida, India

U.G. Student, Department of Computer Science & Engineering, JSS Academy of Technical Education, Noida, India

U.G. Student, Department of Computer Science & Engineering, JSS Academy of Technical Education, Noida, India

Assistant Professor, Department of Computer Science & Engineering, JSS Academy of Technical Education, Noida, India

ABSTRACT: The ability to analyze financial markets and make informed decisions is critical to successful business operations today. Our project introduces a new machine learning-trained platform to assess a person's ability to analyze past outcomes. The platform allows users to analyze and predict future business value using historical financial information such as balance sheets, income statements and cash flows. It provides a better understanding of the timeline of financial statements. Users predict the next market price, and the platform evaluates these predictions by comparing them with predictions from our machine learning models based on real market data. To measure accuracy, the platform provides insight into users' ability to evaluate their predictions by comparing them to predictions made by our models. It simplifies the forecasting process by providing accurate data and forecasts based on real forecast data. The platform bridges the gap between theoretical financial knowledge and practical applications, providing users with the knowledge and experience they need to make investment decisions. Target users include investors, educators, and analysts who want to improve their financial analysis skills with a hands-on, data-driven approach.

KEYWORDS: Financial markets, Machine learning, Market price prediction, Financial forecasting, Machine learning models

I. INTRODUCTION

Our platform is designed to improve users' market forecasting and forecasting capabilities by using machine learning algorithms such as our LSTM learning model to predict stock prices. Users can compare their predictions to our model's predictions to provide a benchmark for accuracy. This comparison not only allows users to understand their strengths and weaknesses but also reflects the difference between financial literacy and the investment layer, which currently does not exist in the financial world. This is especially useful for investors, educators, analysts, and recruiters. Users are encouraged to continue improving their financial analysis skills based on a comparison of the actual forecast with the forecast obtained from the trained LSTM model.

The platform offers access to comprehensive historical financial data, such as balance sheets, income statements, cash flow, and more, allowing users to make informed predictions by analyzing the company's financial history. Practicing with this information can improve users' analytical skills. Additionally, by integrating real-time decision-making into standard training, users should make instant decisions about buying, selling, or holding stocks based on their predictions. This real-world application allows users not only to learn but also to apply their knowledge by testing real materials. The platform fosters an environment of continuous learning and development. It includes a rich library of courses, such as guides, articles, and research papers that help users understand the market more deeply and improve their investment decisions. The availability of necessary tools and materials supports users in enhancing their financial analysis capabilities.

Users can track their performance over time, view accurate predictions, identify their strengths and weaknesses, and receive personalized recommendations for improvement. This comprehensive program encourages continuous learning and improvement in financial analysis, helping users enhance their skills and make more informed investment decisions.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1206008 |

II. RELATED WORK

1. **Stock Prediction Using LSTM Networks**
   Summary: This study demonstrates the accuracy of LSTM models in predicting stock prices by training on historical market data. LSTM models capture temporal dependencies and trends, leading to more accurate predictions compared to traditional methods like ARIMA and GARCH. The work highlights the potential of LSTMs in financial forecasting, making it highly relevant to our platform's predictive capabilities.

2. **DeepFinance: Harnessing Deep Learning for Financial Time Series Forecasting**
   Summary: This paper combines stacked autoencoders and LSTM networks to achieve superior prediction results for stock prices and volatilities. Autoencoders are used for feature extraction, while LSTM networks handle the sequence data of financial time series. This approach aligns with our platform's goal of enhancing financial forecasting accuracy through deep learning.

3. **Quantitative Investment Using Machine Learning**
   Summary: This example showcases how advanced machine learning, particularly LSTM methodology, can be used in empirical asset pricing to outperform traditional econometric models. Thanks to their ability to process large datasets and uncover deeper patterns, LSTMs align perfectly with the capabilities we facilitate on Optimyse.

4. **Predictive Modeling of Stock Prices Using LSTM**
   Summary: This research focuses on leveraging LSTM networks to predict business transactions. Through a comparative analysis with other neural network architectures, the study confirms LSTM's superiority due to its ability to capture long-term dependencies within financial data. This validation is pivotal for ensuring accurate product prediction on our platform.

5. **Evaluation of Machine Learning Models for Stock Price Prediction**
   Summary: This paper assesses various learning models, including LSTM, for predicting stock returns within the Chinese stock market. The findings reveal LSTM's superiority over traditional models in terms of prediction accuracy. The insights from this study regarding the application of LSTM models in financial forecasting directly inform the development and implementation of our platform.

6. **Fusion of Deep Learning Techniques for Enhanced Stock Prediction**
   Summary: Recent research aims to develop efficient mechanical trading systems using machine learning algorithms for estimating stock prices. This paper reviews AI and ML strategies, including LSTM, Hybrid LSTM, and CNN, assessing their accuracy and limitations for stock price forecasting, which aligns with our platform's goals.

7. **ML Algorithms for Forecasting Stock Prices.**
   Summary: This investigation explores stock prediction through the application of machine learning algorithms, utilizing regression and classification methods to anticipate closing prices and forecast market trends. The study's findings highlight the effectiveness of ML techniques in enhancing stock market predictions.

8. **Utilizing Machine Learning Classifiers for Anticipating Trends in Stock Market Exchanges**
   Summary: Applying machine learning classifiers to stock trends enhances prediction accuracy by analyzing data patterns with advanced algorithms. This study showcases the potential of ML classifiers in improving the forecasting of stock market trends.

9. **Comprehensive Review of Machine Learning Methods for Stock Price Prediction**
   Summary: This paper provides a comprehensive review of stock price prediction techniques leveraging artificial intelligence. It surveys diverse methods, including traditional ML, deep learning, neural networks, and graph-based approaches, examining challenges and outlining future prospects in this dynamic field.

10. **Fusion of Deep Learning Techniques for Enhanced Stock Prediction**
    Summary: Recent research aims to develop efficient mechanical trading systems using machine learning algorithms for estimating stock prices. This paper reviews AI and ML strategies, including LSTM, Hybrid LSTM, and CNN, assessing their accuracy and limitations for stock price forecasting, which aligns with our platform's goals.

11. **Total Value of Listed Domestic Companies (Current US$) Dataset**
    Summary: Market capitalization analysis assesses the total value of listed domestic companies' outstanding shares, offering insights into their size, market standing, and investor sentiment. This data significantly influences investment decisions and market trends, providing a valuable resource for financial analysis.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

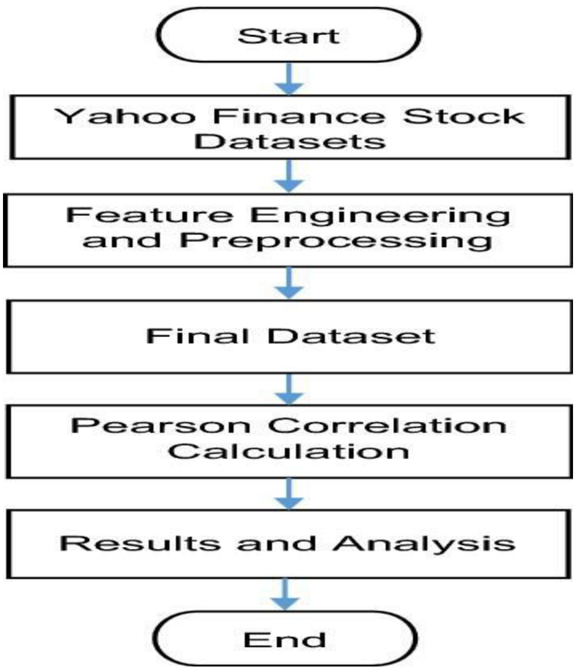
| DOI: 10.15680/IJIRCCE.2023.1206008 |

Ref	Model	Accuracy
	LSTM	MSE = 0.035 (avg)
[12]	Random Forest	EVS = -0.400594
[13]	LBL-LSTM (Proposed)	0.018 (Train), 0.027 (Test)
[14]	LSTM (Decentralized)	MSE = 0.0004
[15]	k-NN Regression	90%
[16]	SVM & Logistic Regression	87% to 90%
[17]	CNN-LSTM	MSE = 3.5 to 3.7

Table 1. Comparison of LSTM with different Algorithms

III. METHODOLOGY

3.1 Datasets:



The initial step involves accessing financial data from Yahoo Finance for selected companies. Key metrics like stock

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1206008 |

price, trading volume, and other financial indicators are extracted during this stage. This data serves as the primary source for subsequent analysis and forms the basis of the research. The primary steps are:

- **Feature Extraction and Preprocessing:** Advanced techniques are employed to extract relevant features from the raw data. This includes creating new variables, handling missing values, and applying normalization and scaling methods. The goal is to ensure data consistency and analytical readiness for further processing.
- **Final Dataset Preparation:** Post feature extraction and preprocessing, a refined and organized final dataset is obtained. This dataset contains carefully selected and organized data, ready for in-depth analysis and modeling. It serves as the foundation for subsequent analytical work and decision-making processes.
- **Statistical Analysis:** Statistical analysis, particularly Pearson correlation calculation, is conducted to understand the relationship between variables. Correlation between various financial indicators, such as market prices and products, is analyzed to derive insights.
- **Insights and Implications:** Insights gained through statistical analysis play a crucial role in guiding investment decisions, risk management strategies, and data optimization. These findings have significant implications for stakeholders, including investors, financial analysts, and policymakers. The information derived from the analysis can inform the improvement of forecasting models, investment strategies, and future research endeavors.

3.1.1 Market Data

In this study, we harnessed market data from Yahoo Finance, with a particular focus on the S&P 500 dataset, renowned for its comprehensive representation of the U.S. stock market. Comprising large-cap stocks across diverse sectors, this dataset served as a cornerstone for analyzing historical stock trends and current stock prices. By leveraging the S&P 500 dataset, we ensured the consistency and reliability of our analysis, facilitating robust evaluations of predictive models and assessments of users' analytical prowess. Our LSTM model was trained using stock prices spanning a five-year period, allowing us to compare its accuracy against other ML models.

3.1.2 Textual Data:

For textual data, we gathered information from various reputable sources, including financial news outlets such as The Wall Street Journal, Bloomberg, and CNBC, alongside online platforms like Google Finance and Yahoo Finance. These sources provided a rich tapestry of information and sentiment analysis, furnishing valuable insights into market sentiments and investor perceptions. By integrating textual data from these diverse sources, our aim was to capture a comprehensive understanding of market dynamics and sentiment trends, thereby augmenting the predictive capabilities of our LSTM model, which was initially trained on S&P 500 datasets. Furthermore, we augmented our dataset by incorporating data from prominent companies such as Meta, Google, and Apple, sourced from Yahoo Finance. This additional data enriched the training process of our LSTM model, enhancing its capacity to forecast stock prices accurately in future time periods.

3.2 Algorithms

3.2.1 Linear Regression:

Linear regression is a pivotal tool in financial market forecasting models due to its convenience and efficacy. Its applicability lies in establishing relationships between various factors and offering a method for predicting stock prices based on historical data.

- **Foundation of Linear Regression:** The principle of linear regression involves fitting an equation to data points, providing a simple yet effective method for understanding and predicting product patterns.
- **Ease of Interpretation and Use:** Its simplicity and ease of interpretation make linear regression ideal for users analyzing business transactions using historical financial data, such as company balance sheets and business indicators.
- **Predictive Ability:** Linear regression can establish a relationship between independent and dependent variables, serving as a significant predictor of future market prices.
- **Utilizing Historical Data:** By leveraging historical data, linear regression enables users to make informed investment decisions by predicting changes in market behavior.
- **Computational Efficiency:** Linear regression's computational efficiency and effectiveness in capturing the

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1206008 |

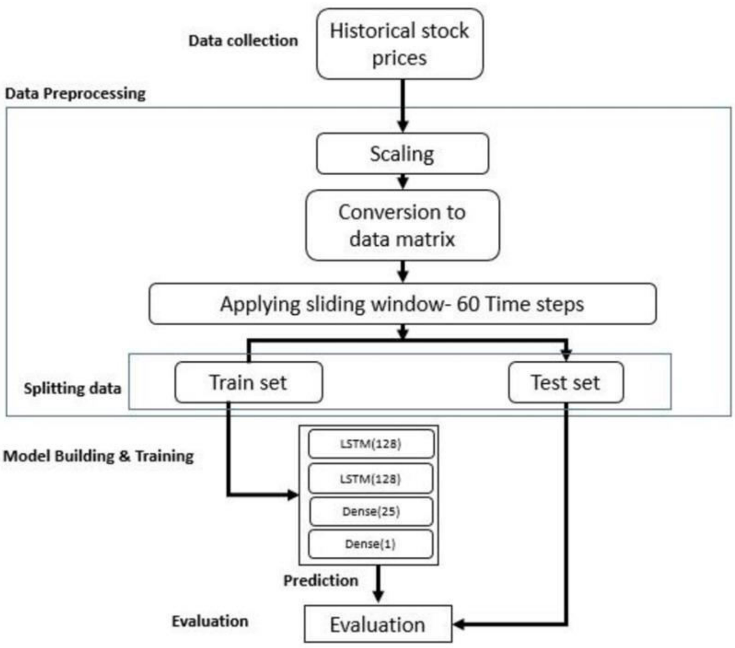
relationship between events align perfectly with assessing users' predictive abilities.

- **Value and Importance:** The value of linear regression, combined with its simplicity and predictive power, makes it an essential tool for understanding labor market trends and enhancing users' analysis and forecasting skills.

3.2.2 LSTM:

To predict stock prices using the short-term trend model (LSTM), the process begins by collecting historical stock prices of companies such as Apple, Meta, and indices such as the S&P 500.

- **Data Collection:** Historical stock prices of companies like Apple, Meta, and indices such as the S&P 500 are collected.
- **Scaling:** Stock prices are standardized to ensure they fall within a similar range, which is crucial for successful LSTM model analysis.
- **Data Transformation:** The original data is converted into matrix data to facilitate sequential data processing.
- **Sliding Window Application:** The 60-time step sliding window technique is applied to generate market values for LSTM models, enabling the learning of time patterns.
- **Model Architecture:** The LSTM model consists of two layers, each with 128 units, followed by two dense layers with 25 and 1 unit, respectively.
- **Training:** The models are trained to understand time-dependent expectations and patterns in the stock market.
- **Evaluation:** The estimated cost is compared to the actual cost to evaluate the performance of the model. This testing helps understand the accuracy and ability of the model and allows for adjustments to increase its predictive power



3.2.3.Random forest algorithm

Random forest algorithm is a non-linear model that integrates various decision trees into a forest. To understand random forest, it is important to consider two important concepts: testing and majority voting. Specifically, the training set for each decision tree is selected from the entire sample set. This article uses decision tree classification; Retrieving features in the feature matrix method can prevent overfitting of a decision tree. Finally, the ratio of misclassifications to the total sample is used as the out-of-bag misclassification rate of the random forest.

International Journal of Innovative Research in Computer and Communication Engineering

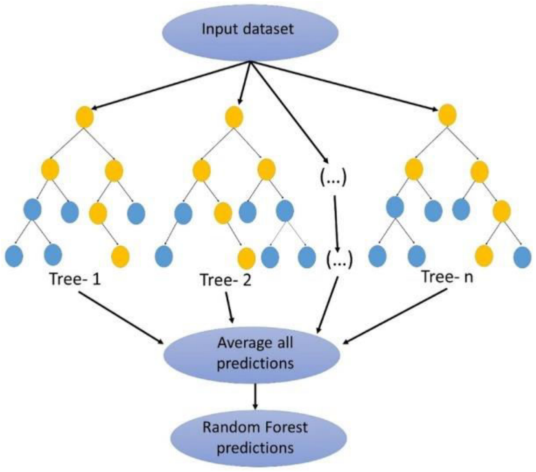
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1206008 |

In a random forest, each tree is created independently to be more diverse, and this is done through a technique called bootstrap sampling. [‘a’, 'b', 'c' ,'d', 'f']. The first pattern contains ['a', 'a', 'c', 'f', 'b'], while the second pattern contains ['d', 'a', 'f', 'CD'].

Use two independent bootstrapping samples to create separate decision trees that help build forest trees. To complete, use the random forest model to predict the market value where the predicted probability of each tree should be averaged and select the category with the highest probability based on the predictions. In this case, a random forest of five trees can be used for two satellite datasets:



From the above model, it can be seen that the order of these five trees is very different. All trees will make some errors because some of the training points shown here were not included in the training of these trees. This is mainly because a self-sampling random forest fits less than a tree, resulting in a more intuitive decision boundary. But when solving the real problem, we will use more trees (hundreds of thousands) to get a better intersection. Classifier.

The output category is determined by the mode of the categories output by each tree. The randomness is mainly reflected in two aspects:

(1) During the training of each tree, a dataset of size N is selected from all training samples, which may contain duplicates, for training purposes. This is known as bootstrap sampling.

(2) At each node, a subset of features is randomly selected to calculate the optimal segmentation method.

Therefore, this paper also used random forest regression to predict stock returns in the stock market

3.3. SYSTEM DESIGN AND IMPLEMENTATION:

- **Data Acquisition and Preprocessing:**
  Collect historical financial data from reliable sources like Yahoo Finance, including balance sheets, income statements, and market data for real companies. Preprocess the acquired data to handle missing values, normalize the data, and ensure consistency across different datasets. Use techniques such as linear interpolation and Min-Max scaling.

- **Feature Selection and Engineering:**
  Conduct feature selection to identify relevant predictors for stock price movement using statistical tests and domain knowledge. Incorporate domain-specific indicators like trading volume, price-to-earnings ratio (P/E), and market capitalization to capture fundamental aspects of the companies.

- **Model Development:**
  Experiment with various machine learning algorithms such as linear regression, random forest, and LSTM neural networks. Train the models using historical financial data, with a rolling window approach to capture temporal dependencies. Utilize scikit-learn for linear regression, ensemble method for random forest, and Keras library for LSTM networks.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 6, June 2024 ||
| DOI: 10.15680/IJIRCCE.2023.1206008 |

- **Model Evaluation and Validation:**
  Evaluate the performance of each model using metrics like mean squared error (MSE), mean absolute error (MAE), R-squared, accuracy, precision, recall, and F1 score. Perform cross-validation to assess generalization capability across different time periods. Conduct sensitivity analysis to investigate the impact of feature selection on model performance.

- **Validation and Testing:**
  Validate the predictive model using out-of-sample testing on unseen data. Conduct back testing simulations to evaluate the profitability and risk of trading strategies based on model predictions.

- **User Interface Development:**
  Develop an interactive user interface allowing users to access financial data of different companies. Provide visualizations of historical data through graphs for better analysis. Include features such as downloadable excel sheets for further data analysis.

- **Prediction Engine:**
  Implement a prediction engine that processes user input data and generates future predictions of stock values using trained LSTM models. Compare user predictions with model forecasts to provide feedback on prediction accuracy. Enable users to make decisions on buying, selling, or holding stocks based on predicted values.

- **Grading System:**
  Develop a grading system to compare user predictions with actual values from trained data models. Utilize accuracy metrics such as RMSE to quantify the difference between user predictions and actual values. Assign scores to user predictions based on accuracy, enabling users to track their performance over time.

- **Continuous Learning and Improvement:**
  Incorporate a feedback mechanism to gather user input on prediction accuracy and overall user experience. Periodically update LSTM models using fresh financial data to adapt to changing market conditions. Implement algorithmic enhancements based on research and industry best practices to improve prediction accuracy and reliability.

- **Real-time Data Integration:**
  Establish connections with financial markets and news sources and continuously monitor market developments, news updates, and other relevant events to provide timely recommendations to users. Integrate real-time data updates into predictive models and analysis to adapt to changing market conditions quickly

IV. EXPERIMENTAL RESULTS

Figures shows the Financial platform for stock predictions and financial forecasting.

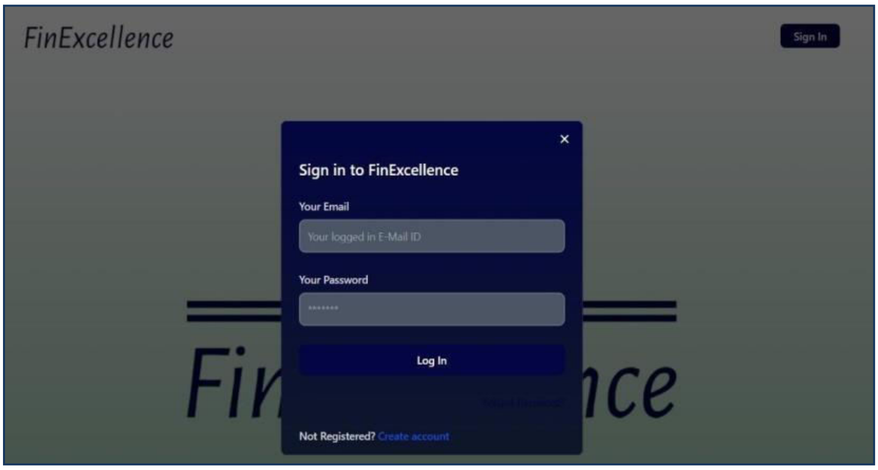


FIG 1: LOGIN PAGE

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1206008 |

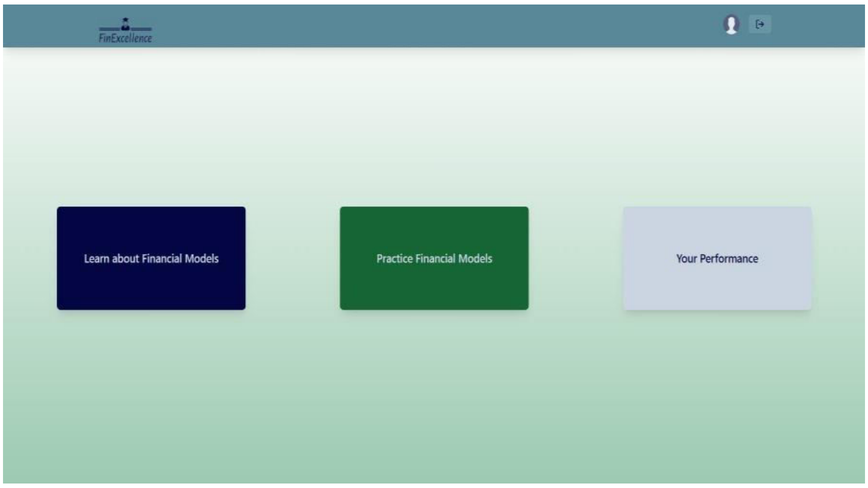


FIG2: HOME PAGE

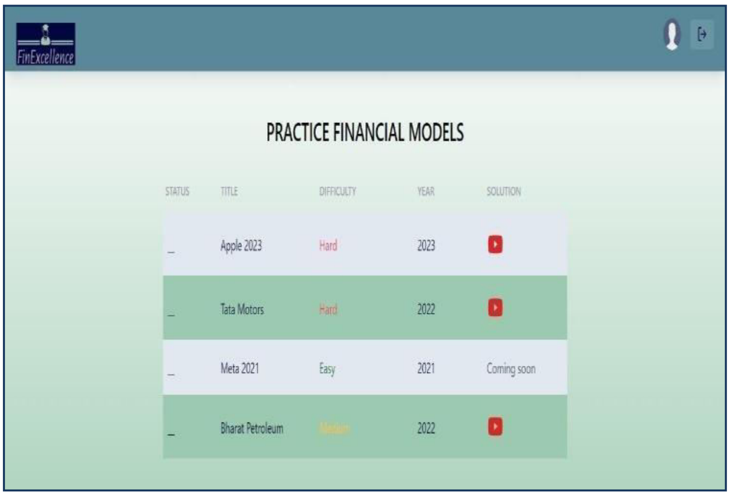


FIG3: QUESTIONS PAGE

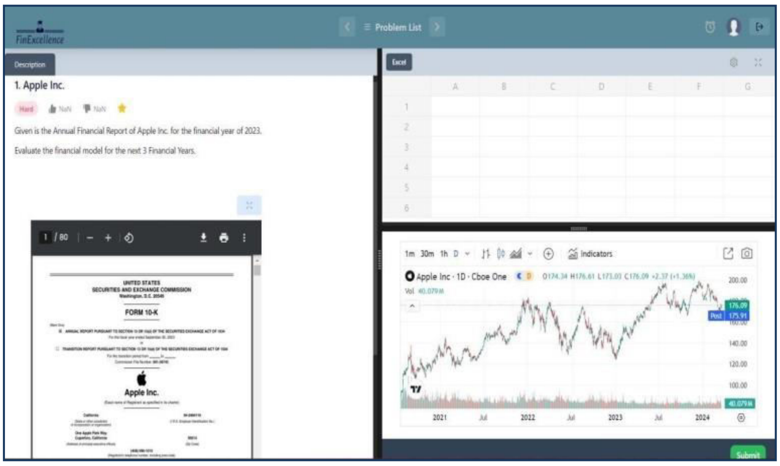


FIG 4: PROBLEMS PAGE

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1206008 |

FIG 5: ANSWER PAGE

V. CONCLUSION

By creating a new platform that applies our manually trained ML model for measuring user's performance who after analyzing the source materials provided to them make predictions of stock values is a noteworthy evolution in financial analysis of users. By using LSTM and linear regression amongst other high-end techniques, combined with proper testing and validation procedures the platform assists a user in the precise anticipation and evaluation of investment trends. The users are able to dive deeply into the company's history of revenues, balance sheet, and other important information, that will enable them to make effective predictions of the future values of business and markets. It also helps in bridging the gap between theoretical knowledge of finance on one hand and investment analysis on the other. In this way, citizens, investors, educators, and analysts will be able to benefit greatly from the site since it provides them with the tools to perform analysis on different aspects of life. Owing to the emerging realms of technological application, particularly in the field of finance, and the increasing complexity of financial markets, the utilization of innovative facilities such as machine learning, our platform acts as a promising prospect for redesigning the existing financial landscape

REFERENCES

[1] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669.
[2] Zhang, Y., & Tan, K. C. (2018). A deep learning framework for financial time series using stacked autoencoders and long short-term memory. Neurocomputing, 284, 284-291.
[3] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Data models, 2223-2273.
[4] Nelson, D. M., Pereira, A. C., & de Oliveira, R. A. (2017). Stock market's price movement prediction with LSTM neural networks. IEEE International Joint Conference on Neural Networks (IJCNN), 1419-1426.
[5] Chen, K., Zhou, Y., & Dai, F. (2019). A LSTM-based method for stock returns prediction: A case study of China stock market. IEEE International Conference on Big Data (Big Data), 2823-2824.
[6] Jaimin Shah, Darsh Vaidya, Manan Shah,A comprehensive review on multiple hybrid deep learning approaches for stock prediction,Intelligent Systems with Applications,Volume 16,2022
[7] Ali Khan, J. Predicting Trend in Stock Market Exchange Using Machine Learning Classifiers. Sci. Int. 2020, 28, 1363– 1367.
[8] Predicting Trend in Stock Market Exchange Using Machine Learning Classifiers,Ali Khan,J.,2020
[9] Machine Learning Approaches in Stock Price Prediction: A Systematic Review,Payal Soni,Yogya Tewari, Deepa Krishnan,2022
[10] Jaimin Shah, Darsh Vaidya, Manan Shah,A comprehensive review on multiple hybrid deep learning approaches for stock prediction,Intelligent Systems with Applications,Volume 16,2022
[11] Market Capitalization of Listed Domestic Companies (Current US$) Data. Available online: https://data.worldbank.org/ indicator/CM.MKT.LCAP.CD (accessed on 19 May 2021)
[12] Kompella, S., & Chakravarthy Chilukuri, K. C. C. (2020). Stock Market Prediction Using Machine Learning Methods. International Journal of Computer Engineering and Technology, 10(3), 2019.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 6, June 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1206008 |

[13] Gurav, U., & Kotrappa, D. S. (2020). Impact of COVID19 on stock market performance using efficient and predictive LBL-LSTM based mathematical model. International Journal on Emerging Technologies11 (4), 108-115.

[14] Bansal, G., Hasija, V., Chamola, V., Kumar, N., & Guizani, M. (2019, December). Smart stock exchange market: a secure predictive decentralized model. In 2019 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). IEEE.

[15] Ananthi, M., & Vijayakumar, K. (2021). Stock market analysis using candlestick regression and market trend prediction (CKRM). Journal of Ambient Intelligence and Humanized Computing, 12(5), 4819-4826.

[16] Parray, I. R., Khurana, S. S., Kumar, M., & Altalbe, A. A. (2020). Time series data analysis of stock price movement using machine learning techniques. Soft Computing, 24(21), 16509- 16517

[17] Bhattacharjee, I., & Bhattacharja, P. (2019, December). Stock Price Prediction: A Comparative Study between Traditional Statistical Approach and Machine Learning Approach. In 2019 4th International Conference on Electrical Information and Communication Technology (EICT) (pp. 1- 6). IEEE.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING