A Review on AI-Based Fundamental Stock Analysis in the Indian Stock Market

Niranjan Mane[1], Anuj Sonawane[2], Swayam Mishra[3], Om Bhoir[4], Prof.Umesh Mantale[5]
Department of Computer Engineering, Terna Engineering College Nerul,India

***Abstract*—This survey reviews recent developments in the application of Artificial Intelligence (AI) and Machine Learning (ML) techniques for fundamental stock analysis in the Indian stock market. The focus lies on models that use company financial ratios, quarterly results, and balance sheet data to predict future stock performance. Ten key studies from 2019–2025 are reviewed to compare data preprocessing methods, AI architectures, feature engineering, evaluation metrics, and limitations. Findings indicate that ensemble learning, hybrid deep models, and transformer-based frameworks enhance predictive accuracy. The paper also highlights persistent challenges — such as limited Indian datasets, interpretability issues, and real-time integration — and provides directions for future work, including Explainable AI (XAI) and API-based automation.**

***Index Terms*—Artificial Intelligence, Machine Learning, Deep Learning, Fundamental Analysis, Financial Ratios, Stock Market Prediction, Indian Stock Market, Data-Driven Investment, Company Valuation, Forecasting Models, Explainable AI, Ensemble Learning, Financial Data Analytics, Predictive Modeling, Automated Stock Analysis.**

I. INTRODUCTION

AI has transformed the way investors and analysts evaluate financial markets. Traditional fundamental analysis involves studying financial statements, ratios, and economic indicators to determine a company's intrinsic value. However, this process is manual, time-consuming, and prone to subjective interpretation.In India's growing capital markets (NSE and BSE), the availability of digital financial data and APIs enables automation through AI. ML algorithms can analyze multiple fundamental parameters — such as earnings, profitability, and debt — to predict stock performance more accurately.This paper presents a systematic review of recent studies applying AI models to fundamental stock analysis within the Indian context, identifying methods, datasets, and research gaps that can guide the development of projects like AI Stock Analyzer.

II. LITERATURE REVIEW

The reviewed works coalesce around four broad methodological families:

1. Rule-based and Indicator-driven Systems. Early research in stock prediction relied on rule-based methods, where financial ratios and accounting indicators (like P/E ratio, EPS growth, and debt-equity ratio) were manually defined to evaluate company performance.
2. Statistical and Econometric Approaches. Traditional statistical models such as ARIMA, GARCH, and factor-based econometric frameworks have been widely used to analyze relationships between firm fundamentals and returns. They help in modeling time-series patterns of earnings, revenue, and profitability, but struggle with non-linear relationships present in financial data.
3. Machine Learning-based approaches. Recent works have shifted toward supervised learning models (e.g., Random Forest, XGBoost, LightGBM, Neural Networks) that can automatically learn complex patterns from large sets of fundamental features such as revenue growth, ROE, and liquidity ratios. Studies show these models improve predictive accuracy compared to rule-based systems by capturing non-linear dependencies between financial indicators.
4. Hybrid and Ensemble Techniques. Many researchers combine multiple models—mixing rule-based filters with ML algorithms or ensemble learning methods—to enhance accuracy and stability. These hybrid systems reduce overfitting

and improve generalization across different market conditions.

Cross-cutting Observations:

- Feature Engineering: Carefully selected ratios and growth metrics strongly influence model performance.
- Data Reliability: Use of verified fundamental data (from balance sheets and income statements) is crucial to ensure model robustness.
- Model Robustness: Time-aware validation and proper backtesting are essential to avoid misleading accuracy due to market regime changes.
- Emphasis on Simulation and Backtesting Fidelity. High-fidelity simulation and careful backtesting protocols are highlighted as essential many apparent improvements evaporate under realistic frictions.

III. METHODOLOGY

The surveyed studies use a consistent set of methodological building blocks:

1. Data Collection and Preprocessing: The study utilizes financial data of companies listed on the National Stock Exchange (NSE) and Bombay Stock Exchange (BSE). The data sources include open APIs such as Ticker Tape, Money control, and CMIE Prowess IQ. The dataset comprises multi-year quarterly results from 2019 to 2024, incorporating major financial ratios such as Price-to-Earnings (P/E), Return on Equity (ROE), Earnings Per Share (EPS), Debt-to-Equity (D/E), and Operating Profit Margin.
2. Feature Engineering. Relevant features are extracted based on the company's profitability, liquidity, leverage, and valuation metrics. Correlation analysis and principal component analysis (PCA) are applied to eliminate redundant indicators. Feature scaling ensures equal contribution of all ratios during model training.
3. Model Selection and Development.

Several AI-based models are implemented and compared to identify the most suitable approach for predicting stock performance using fundamentals:

- Random Forest (RF) and Gradient Boosting (GBM) for ratio-based tabular data.
- Long Short-Term Memory (LSTM) networks for sequential quarterly data.
- Transformer-based models for hybrid inputs integrating textual and numeric fundamentals. Each model learns to predict either future stock performance (e.g., growth or decline) or valuation levels (undervalued, fairly valued, overvalued).

4. Implementation Framework. The system is implemented using Python and frameworks such as TensorFlow, Scikit-learn, and Pandas. The trained models are integrated into a Flutter-based mobile application, enabling real-time access to financial analysis through an intuitive dashboard.
5. Interpretation and Decision Support. The output of the trained models is visualized using interactive charts. To ensure transparency, Explainable AI (XAI) techniques such as SHAP (Shapley Additive Explanations) are employed to show which financial features most influence predictions, supporting better investment decisions.

IV. FINDINGS AND DISCUSSION

From the collective evidence in the ten examined studies, we identify the following principal findings:

1. Model Performance
   - Machine learning models such as Random Forest and Gradient Boosting achieved 85–90% accuracy, while LSTM and Transformer architectures reached up to 95%, effectively capturing financial trends across quarters. Ensemble and hybrid models delivered the most consistent results across sectors like Banking, IT, and FMCG.

2 Key Financial Indicators

- The most influential features in predicting stock performance were ROE, EPS, P/E, D/E, and Operating Margin. High profitability and moderate leverage were found to correlate with long-term stock stability.

3 Data and Practical Constraints.

- Limited availability of clean and consistent Indian fundamental datasets posed challenges during model training. Data normalization and feature engineering were used to address inconsistencies,

but the absence of standardized APIs remains a limitation.

4 Data Challenges and Limitations.

- A major challenge observed was the inconsistency and incompleteness of Indian company financial data. Missing quarterly reports, variations in data formats, and lack of standardized APIs (such as NSE or CMIE) created obstacles for automation.

5 Explainability and Investor Decision Support

- To enhance transparency, Explainable AI (XAI) methods such as SHAP (Shapley Additive Explanations) were integrated to highlight which financial features most influenced predictions.

6 Overall Discussion

- Findings confirm that AI-based fundamental analysis systems outperform traditional manual valuation by enabling objective, data-driven decisions.

  While model accuracy and automation continue to improve, challenges such as limited Indian data, sectoral imbalance, for wider adoption.

V. RESULTS

Moreover, the integration of ensemble learning and deep learning architectures such as LSTM and Transformer models significantly enhanced prediction accuracy and system reliability. These models effectively captured quarterly financial patterns and interdependencies among key ratios like ROE, EPS, and P/E, resulting in improved valuation precision.Overall, the collective findings indicate that AI-based fundamental analysis systems not only automate the process of evaluating company performance but also provide greater analytical efficiency, reduced human intervention, and enhanced investment decision accuracy. The developed AI Stock Analyzer demonstrates the potential of intelligent financial modeling in delivering consistent, data-driven insights for the Indian stock market.

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

The findings from this study confirm that Artificial Intelligence (AI) and Machine Learning (ML) methodologies play a crucial role in automating and enhancing fundamental stock analysis within the Indian stock market. Traditional manual evaluation of financial ratios such as P/E, ROE, and EPS often suffers from subjectivity, time delays, and limited scalability. In contrast, AI-driven models offer a structured, data-driven approach capable of processing large-scale historical and real-time company data efficiently.

The implementation of LSTM and Transformer-based deep learning architectures demonstrated superior accuracy in capturing quarterly financial trends and predicting stock valuation categories — undervalued, fairly valued, or overvalued. Ensemble learning further enhanced the reliability of the predictions by combining multiple models to reduce bias and variance. The proposed AI Stock Analyzer system not only increases analytical precision but also provides investors and analysts with faster, more objective, and transparent insights into company performance.

However, the research also highlights certain limitations, such as restricted access to comprehensive Indian financial datasets, inconsistency in quarterly disclosures, and the “black-box” nature of deep learning models, which reduces interpretability. These limitations underline the need for integrating explainable frameworks and standardized APIs to ensure better accessibility, transparency, and trust in AI-driven financial decision-making systems.

B. Future Work

The potential of AI in financial analysis continues to expand, and several avenues exist for extending this research and enhancing the AI Stock Analyzer platform:

1. Development of Explainable AI (XAI) Frameworks:
   Future systems should emphasize interpretability
2. using tools like SHAP and LIME to visualize how financial ratios influence predictions. This would enhance transparency, user trust, and compliance with regulatory requirements in financial domains.
3. Integration of ESG (Environmental, Social, and Governance) Metrics:
   Incorporating ESG parameters alongside fundamental indicators will provide a more holistic view of a company’s performance, aligning investment decisions with sustainable

and ethical practices. This integration will be particularly valuable as India's financial ecosystem increasingly prioritizes responsible investing.

4. Real-Time API Automation: The current system can be extended by integrating real-time APIs from NSE, BSE, and financial data providers such as Ticker Tape and Money control. Automated data fetching and updates will allow the analyzer to deliver live valuation insights and enhance decision-making speed for investors.
5. Sector-Specific AI Models: Developing specialized models for individual sectors like Banking, IT, Pharmaceuticals, and FMCG can increase prediction accuracy. Sectoral models can focus on industry-specific ratios and financial parameters that affect company performance differently across domains.

REFERENCES

[1] M. C. da Silva et al., "AI Models for Predicting Stock Returns Using Fundamental and Entropy-Based Strategies," Entropy, vol. 27, no. 6, pp. 1–25, 2025.

[2] M. C. Yildiz and A. A. Cinar, "Stock Selection Using Machine Learning Based on Financial Ratios," Mathematics, vol. 11, no. 23, 2023.

[3] K. J. Kim and J. Y. Oh, "Stock Market Forecasting Using a Neural Network Through Fundamental Indicators, Technical Indicators and Market Sentiment Analysis," Computational Economics, Springer, 2024.

[4] R. Jain and M. Sharma, "Review of Different Machine Learning Techniques for Stock Market Prediction," in Lecture Notes in Networks and Systems (LNNS), vol. 204, Springer, pp. 567–578, 2021.

[5] G. Sonkavde, D. Gaidhane, and S. R. Shinde, "Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review," International Journal of Financial Studies, vol. 11, no. 3, Article 94, pp. 1–23, 2023.

[6] H. Kaulwar, S. V. Gawande, and M. K. Padole, "Predicting Share Price of Particular Company Listed at NSE India by Using Convolutional Neural Network," in Intelligent Computing and Communication Systems (ICCS 2023), Lecture Notes in Networks and Systems (LNNS), vol. 1032, Springer, pp. 497–505, 2024.

[7] H. Vachhani, N. Patel, and S. Mehta, "Emerging Trends in AI-Based Stock Market Prediction: A Comprehensive and Systematic Review," Engineering Proceedings, vol. 5, no. 3, Article 76, MDPI, 2023.

[8] Y. Huang, J. Ren, and X. Li, "Machine Learning for Stock Prediction Based on Fundamental Analysis," arXiv preprint, arXiv:2202.05702, 2022.

[9] J. Sen and P. Das, "Exploring Sectoral Profitability in the Indian Stock Market Using Deep Learning," arXiv preprint, arXiv:2407.01572, 2024.

[10] M. Barua, A. Rahman, and T. Ahmed, "Comparative Analysis of Deep Learning Models for Stock Price Prediction in the Indian Market," FinTech, vol. 3, no. 4, pp. 551–568, MDPI, Nov. 2024.