

Survey Of AI Based Stock Market Prediction

Nitin Naik¹, Dr. Sapana Rathod²

¹Research Scholar, Sunrise University, Alwar

²Associate Professor, Sunrise University, Alwar

Abstract

The main aim of this study is to investigate and assess data to derive relevant findings. Substantial effort is invested in the background to collect data through techniques that guide us toward the specific data needed to address the core research question. This chapter explores various options for data analysis techniques available for our data and identifies the most suitable technique for analysing the data. The research delves into the analysis of stock market forecasting in India through AI application technology. We critically evaluate the effectiveness, challenges, and future directions of AI in forecasting stock prices and market movements, aiming to provide insights and guidance for researchers, practitioners, and investors.

I Introduction

The quest for accurate stock market predictions has always been at the forefront of financial research. The inherently volatile and chaotic nature of the stock market, influenced by countless factors including economic indicators, company performance, political events, and investor sentiment, makes prediction a challenging task. With the advent of AI, new horizons have been explored in predicting stock market trends with greater accuracy. This paper surveys the landscape of AI-based stock market prediction, highlighting the progression from early machine learning models to sophisticated deep learning and NLP approaches. A sample size of 365 respondents was selected for this study's data analysis. The data analysis approach involves percentage analysis, confirmatory factor analysis, descriptive statistics, and structural equation modeling (SEM). Percentage analysis, descriptive statistics, and scale reliability testing are executed utilizing the SPSS V26 software tool, offering a quantitative foundation for further investigation. Structural equation modeling enhances the research by validating measurement scales and assessing the complex interrelationships among variables, utilizing AMOS V23, confirming a robust appraisal of the fitness models, and providing valuable insights for the dynamics influencing stock market forecasting in the context of Artificial Intelligence technology in India. Also, to predict the most influential component, the recommended approach of D-CNN is employed, utilizing Python.

In this study Portfolio diversification (PD), Risk and return trade-offs (RRT), Market efficiency (ME), Technology advancement (TA), and Investor decision-making (IDM) act as independent variables; and Stock market forecasting (SMF) act as a dependent variable.

II Sample Description

Below tables in the sample description below provide a detailed breakdown of the sample characteristics and responses to key questions related to stock market participation and the role of AI-based predictions in investment decisions. The "Measures" column outlines the demographic parameters considered, while the "Constructs" column specifies the subcategories within each measure. The "f" column represents the frequency or count of respondents, and the "%" column indicates the percentage of the total sample each category comprises.

2.1.1 Age Description

The provided table outlines the distribution of survey respondents based on age groups and presents the percentage representation of each category within the sample. The age groups are categorized as below 25, 26-35, 36-45, and 46 and above. Among the respondents, 3.8% fall into the age category

below 25, while the majority, constituting 24.1%, belong to the 26-35 age group. The 36-45 age range encompasses 42.2% of the respondents, indicating a significant portion of the sample falls within the middle-age bracket. Respondents aged 46 and above makeup 29.9% of the total sample. This breakdown provides a clear overview of the age distribution within the surveyed population, enabling a better understanding of the demographic composition and potentially offering insights into age-related patterns or preferences that may be relevant to the study's objectives.

Table 2.1: Age Distribution

Measures	Constructs	<i>f</i>	%
Age	below 25	14	3.8
	26- 35	88	24.1
	36-45	154	42.2
	46 and above	109	29.9

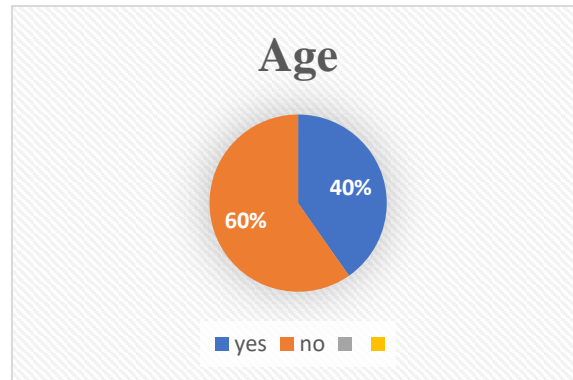


Chart 2.1: Age Distribution

Table 4.2 presents the distribution of survey respondents based on their occupation. The occupations are classified into four groups: Employees, Self-employed individuals, Retired individuals, and Others. Among the respondents, 6.8% are categorized as Employees, indicating individuals engaged in formal employment. A significant portion of the sample, constituting 20.0%, falls under the Self-employed category, reflecting those who are their employers or engage in entrepreneurial activities. The Retired category comprises the largest proportion, with 47.7% of respondents indicating that they are no longer in active employment. The remaining 25.5% fall into the others category, which may include individuals with various miscellaneous occupational statuses not covered by the specified categories. This occupation distribution offers valuable insights into the diversity of occupational backgrounds within the surveyed population, shedding light on the different segments of the workforce or non-working population represented in the study.

Table 2.2: Occupation Distribution

Measures	Constructs	<i>f</i>	%
Occupation	Employees	25	6.8
	Self-employed	73	20.0
	Retired	174	47.7
	Others	93	25.5

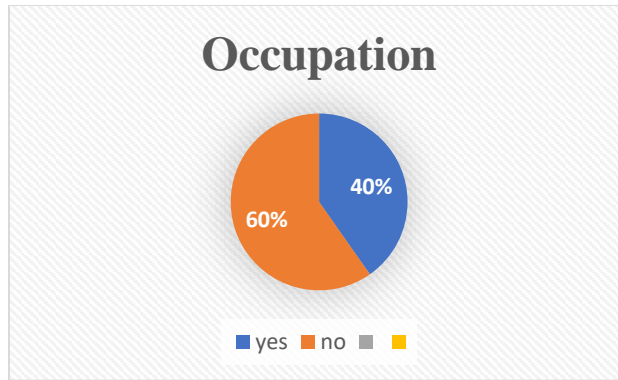


Chart 2.2: Occupation Distribution

Table 4.3 illustrates the distribution of survey respondents based on their income levels. The income levels are segmented into four groups: Below 2 Lakhs/annum, 3 to 5 lakhs/annum, 5 to 10 lakhs/annum, and above 10 Lakhs/annum. Among the respondents, 7.7% fall into the Below 2 Lakhs/annum category, indicating individuals with lower annual income levels. The 3 to 5 lakhs/annum bracket constitutes 26.6% of the sample, representing a substantial portion of respondents with moderate income levels. The majority of respondents, accounting for 40.3%, report an income range of 5 to 10 lakhs/annum. The 10 Lakhs/annum above category includes 25.5% of the respondents, indicating a significant proportion with higher income levels. This income-level distribution provides a comprehensive overview of the financial diversity within the surveyed population, offering information about the study participants' financial backgrounds.

Table 2.3: Income-Level Distribution

Measures	Constructs	<i>f</i>	%
Income level	Below 2 Lakhs/ annum	28	7.7
	3 to 5 lakhs/ annum	97	26.6
	5 to 10 lakhs/ annum	147	40.3
	above 10 Lakhs/ annum	93	25.5

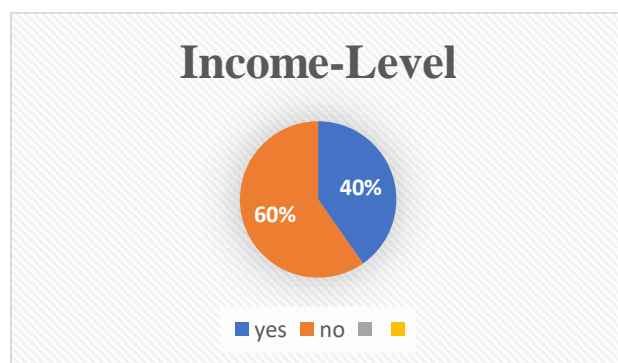


Chart 2.3: Income-Level Distribution

Table 4.3 outlines the distribution of responses to stock market-related questions. The first set of questions pertains to the frequency of stock market participation, where respondents are categorized based on their engagement: Daily (5.5%), Weekly (25.8%), Monthly (40.8%), and Occasionally (27.9%). This breakdown offers insights into the varied levels of involvement and activity among the surveyed individuals in the stock market. The second set of questions focuses on the importance placed on AI-based stock market predictions in investment decisions. Responses include Very

Important (5.5%), Important (24.4%), Neutral (32.9%), Not Important (28.8%), and Not at all Important (8.5%). This distribution provides a nuanced understanding of the perceived significance of AI-based predictions in guiding investment choices. The third set of questions explores respondents' openness to using deep learning-based forecasting models for investment decisions, with responses categorized as Yes (37.5%), No (21.9%), and Maybe (40.5%). This breakdown showcases the diverse attitudes toward adopting advanced technological tools in the context of investment strategies. Finally, the fourth question investigates respondents' willingness to share additional data for a more personalized AI-based forecasting model. The responses include Yes (40.3%) and No (59.7%), indicating the level of comfort or reluctance in providing additional information for a more tailored predictive model. Collectively, this distribution of responses to stock market-related questions provides a comprehensive overview of the participants' attitudes, behaviors, and preferences regarding stock market engagement and the integration of AI and deep learning in their investment decision-making processes.

Table 2.4: Stock Market-related Question Distribution

Measures	Constructs	<i>f</i>	%
How often do you participate in the stock market?	Daily	20	5.5
	Weekly	94	25.8
	Monthly	149	40.8
	Occasionally	102	27.9
How much importance do you give to AI-based stock market predictions in your investment decisions?	Very Important	20	5.5
	Important	89	24.4
	Neutral	120	32.9
	Not Important	105	28.8
	Not at all Important	31	8.5
Are you open to using deep learning-based forecasting models for your investment decisions?	Yes	137	37.5
	No	80	21.9
	Maybe	148	40.5
Would you be willing to share additional data (such as portfolio composition) for a more personalized AI-based forecasting model?	yes	147	40.3
	no	218	59.7

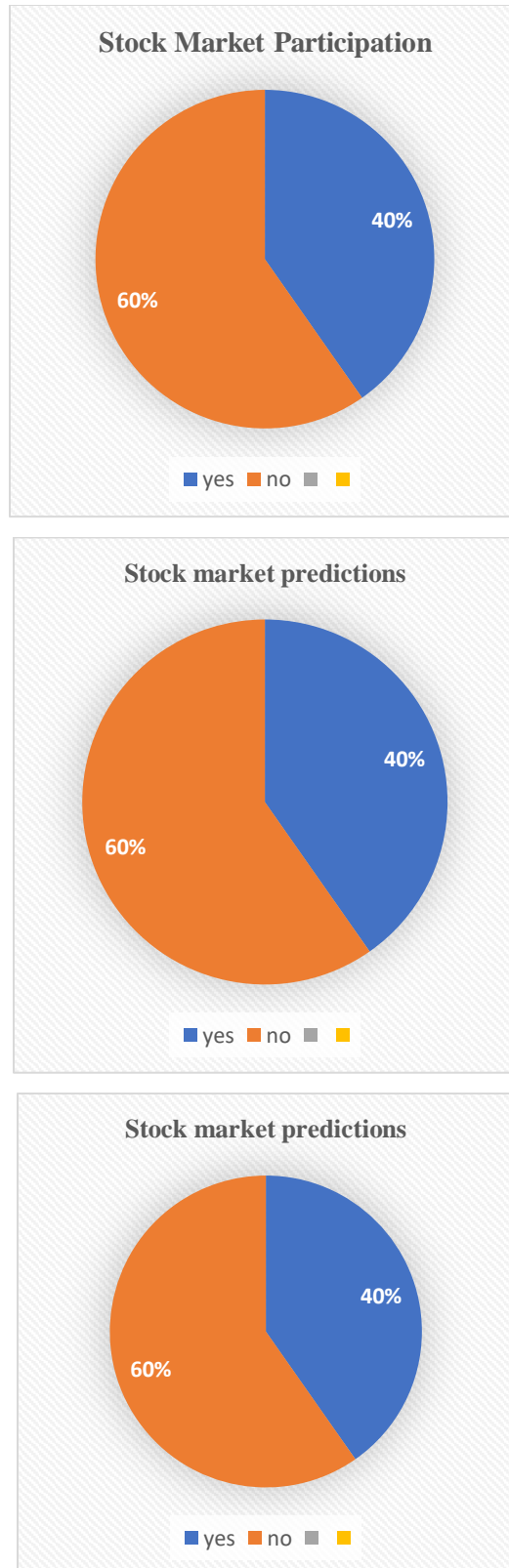


Chart 2.4: Stock Market Prediction Distribution

III Future Directions

The future of AI in stock market prediction lies in addressing the current limitations and exploring new methodologies. Potential directions include the development of more robust and interpretable

models, integration of alternative data sources, and leveraging advancements in AI research such as transformer models and reinforcement learning.

IV Conclusion

AI-based stock market prediction has evolved dramatically, offering promising results and insights into the complex dynamics of financial markets. While challenges remain, the ongoing advancements in AI and computational power, along with the increasing availability of data, suggest a bright future for AI in finance. This survey aims to serve as a foundation for further research and application in this exciting and rapidly evolving field.

V References

- 1.Kumar, M., & Thenmozhi, M. (2006). Forecasting stock index movement: A comparison of support vector machines and random forest. *Indian Institute of Capital Markets 9th Capital Markets Conference Paper*.
- 2.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.
- 3.Dixon, M., Klabjan, D., & Bang, J. H. (2016). Classification-based financial markets prediction using deep neural networks. *Algorithmic Finance*, 6(3-4), 67-77.
4. 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.
5. Zhang, L., Aggarwal, C., & Qi, G. J. (2017). Stock price prediction via discovering multi-frequency trading patterns. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
6. Ballings, M., Van den Poel, D., Hespeels, N., & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications*, 42(20), 7046-7056.
- 7.Hair Jr, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, 109, 101-110. <https://doi.org/10.1016/j.jbusres.2019.11.069>