Enhancing Medium-Term Stock Market Forecasts with Adaptive Feature Selection and Trend Dynamics

D. Vamsi Krishna[1], M. Vishwa Adithya[2], E. Swathi[3]
[1,2]*Dept. of Computer Science and Engineering Chaitanya Bharathi Institute of Technology(A) Hyderabad, India*
[3]*Associate Professor, Dept. of Computer Science and Engineering Chaitanya Bharathi Institute of Technology(A) Hyderabad, India*

***Abstract*—The stock market is very volatile and the change in prices within a short time period is usually full of noise, news events and high frequency algorithmic trades, making it even more hard to make quality predictions on the part of the average investor. The majority of current predictive models are based on a 1–5-day time horizon and use exclusively technical indicators, which is bound to produce unstable signals and big losses, particularly when the underlying asset is in a weak state. To respond to this, our project aims to cover 70 days of the medium- term horizon, as a result of a synergistic approach to basic stock analysis and more sophisticated methods of machine learning. Our approach involves computing a Stock Health Index (SHI) to select assets with a healthy financial base, denoising market prices with a smooth-optimal LOWESS filter and modeling trend dynamics and key factors through a dynamic feature selection process. We use a combination of machine learning models (Artificial Neural Networks, Support Vector Machines, K- Nearest Neighbors and Random Forest) and a weighted average to produce strong and actionable investment projections. Our full architecture is introduced in this paper and we show how our ensemble approach can provide greater stability and accuracy compared to using single-model predictions.**

***Index Terms*—Medium-Term Forecasting, Stock health index, LOWESS Smoothing, Ensemble Learning, Adaptive Feature Selection, Market Trend Dynamics.**

I. INTRODUCTION

The digitalization of financial markets has essentially trans- formed the way investors evaluate equities and algorithmic trading and high-frequency analysis has become the norm [1], During the initial years of the stock market prediction, most analysts mainly focused their efforts on technical indicators in predicting the immediate, short-term price movements (1-5 days) [3],.

To address these issues, our project incorporates the Stock Health Index (SHI), an empirical value of underlying balance sheet health, directly into an artificial intelligence (AI) fore- casting platform. Our model winnows out unreliable stocks using vast amounts of historical data (pricing (OHLCV)) as well as critical growth and profitability indicators before the actual calculations of predictions commence. By using LOWESS price smoothing and Multi-layer Ensemble Learning, we create 70-day stock predictions to protect investors against noise and provide reliable recommendations.

This paper aims to trace the main principles that guide our adaptive architecture and illustrate how fundamentally evaluating information into deep learning pipelines introduces a more systematic, risk-conscious, and actionable trading platform.

II. RELATED WORKS AND COMPARATIVE ANALYSIS

This section will discuss ten paradigm methodologies in algorithmic and financial prediction, clearly stating their short- comings and the manner in which our proposed architecture addresses these shortcomings.

A. Comparison of Top Foundational Models

The first and modern research suggested a wide variety of stock market prediction architectures, most of which rely on either the sheer technical momentum or single machine learning pipelines. To identify gaps in operations, we look at ten studies that are central:

B. How Our Project Overcomes Legacy Limitations

Our system has a multilayered synthesis of architecture as a strategic process of filling in these gaps identified:

- Addressing Value Traps via SHI Integration:

In contrast to the technical-only scopes that are found in the literature of [6], [13] and [21], our system generally eliminates bad capital placement through a strict calculation of Stock Health Index (SHI) before any predictive execution. Essentially poor assets just cannot pass our 60/100 SHI minimum, screening out disastrous, debt-laden equity suggestions completely overlooked by historical models.

TABLE I Comparison of Foundational Studies in Algorithmic Forecasting

Author(s)	Methodology	Advantages	Limitations
Kim [5]	Support Vector Machines (SVM)	• Optimized parameter search • Maps high-dimensional margins	• Vulnerable to episodic shocks • Uses unrefined matrices
Breiman [4]	Random Forests (RF)	• Structurally lowers variance • Robust ensemble voting	• Trained on short-term horizons • Static against macro shifts
Dixon et al. [7]	Deep Neural Networks (DNNs)	• High non-linear mapping ability • Automated feature classification	• “Black box” instability • Lacks fundamental baseline value
Sezer et al. [14]	Sequence Models (LSTMs)	• Supreme time-series forecasting • Learns sequential dependencies	• Relies on raw OHLCV data • Ignores company valuations
Patel et al. [13]	Ensembled Approaches	• Lowers directional error rates • Reduces single-model variance	• Limited to 1–5-day forecasts • High exposure to whipsawing
Weng et al. [18]	Online Ensembled Methods	• High short-term accuracy • Integrates multiple online metrics	• Relies purely on micro-fluctuations • Severe losses in sideways trends
An and Kim [10]	Static Feature Selection	• Extracts non-redundant signals • Mathematically sound extraction	• Inadaptable to revolving markets • Static ranking systems
Sang [19]	Fundamental Integration (ML)	• Discovers undervalued firms • Protects against “value traps”	• Lacks trajectory-timing metrics • High prolonged opportunity cost
Lee and Yen [15]	LOWESS Smoothing with SVR	• Erases transient daily noise • Hybridizes numerical analysis	• Reliant on single SVR bias • Lacks diverse algorithmic voting
Kao et al. [6] & Dash [21]	Hybrid Technical Fusions	• Highly tailored indicator fusions • Integrated ML optimization	• Omits fundamental valuation • Leads users into value traps

- Noise Eradication via 70-Day Horizon and LOWESS:

We particularly focus on a 70-day medium-term horizon, unlike legacy ensemble models of high-frequency oscillations (18) that were devastated by 1-5 days of whipsawing. With dynamic OHLCV price passing by localized LOWESS algorithms (contributing significantly to the ideas of pioneers like [15]), our platform completely filters mathematical falsehood and offers our machine learning elements a mathematically pure macro-trend to forecast, implicitly avoiding the black box instability discovered in [7] and [14] altogether.

- Adaptive Dimensionality Reduction:

Beyond the un- changing boundaries of [10] we apply an Adaptive Feature Selector. This dynamically re-calculates the relevance of indicators (through dynamic Relative Trend Forecast tracking), and only those parameters remain which are mathematically strongest, which systematically counteracts the computational lag in changing macroeconomic environments.

- Multi-Model Robustness (7-Model Weighting):

Our mathematical combination of SVM, ANN, KNN and Random Forest classifications (unlike 15) overrides individual model variances. Individual tree clusters are also likely to be biased in a misleading manner; our more global support vector measurements have the corrective effect of historical analogy (KNN), and generate a single robust confidence score that is currently averaging 21.75% increased predictive integrity as compared to silo- style algorithms.

III. PROBLEM STATEMENT

Existing market evaluation software is much more highly indexed to extremely short-term prediction cycles, and all but ignores the medium-term horizons. They are overly dependent on technical signals instead of fundamental health analysis, i.e., the signals that they produce are noisy, too risky and unreliable. In larger market corrections, these models fail completely because they do not have an inherent value floor, and give false signaling of buy on stocks that are running with perilous debt values or bad cash flow.

This project fills this critical gap by coming up with a strong 70-day medium-term forecasting framework. The timeline does not release daily noise but allows the portfolio a sufficient degree of agility in terms of active reallocation. Our product will generate stable, reliable, and fundamentally sound signals by calculating a Stock Health Index constraint, running advanced LOWESS time-series smoothing and incorporating adaptive feature selection and dynamic trend-aware machine learning models.

IV. METHODOLOGY

Our architecture is a 70-day closed-loop pipeline that is highly developed to assess equity viability.

A. Data Acquisition and Preparation

The system is based initially on gathering of detailed quantitative data on individual stocks and indices (e.g., NASDAQ, NIFTY).

- Price Data: End-of-Day prices and Open-High-Low- Volume (OHLCV) arrays of the past.
- Fundamental Data: Information mined out of balance sheets, Profit and Loss (P&L) statements and cash flow statements.
- Cleaning: The system performs a lot of normalization to allow the statistical harmony of extremely different metric scales (e.g., trade volume and ROE percentages) and dates to be strictly aligned.

B. Stock Health Index (SHI) Construction

The SHI assures the essential security of foreseen resources:

1) Basic ratios such as profitability, valuations, growth and leverage are dynamically calculated.
2) The values are uniformly scaled to a 0-100 margin and weighted average schema is used to combine them into a single SHI score.
3) Threshold Application: The architecture implements a rigid boundary of the base (e.g., SHI ≥ 60). The weak stocks are automatically marked or sifted out to protect investors against value traps.

C. Price Smoothing and Feature Engineering

Financial data inherently acts non-linearly and noisily.

- Noise Reduction:

The use of an efficiency-optimized LOWESS smoothing algorithm on historic closing prices reduces the effects of misleading daily spikes, but shows the underlying macro-trend.

- Feature Generation:

The platform calculates a Relative Trend Forecast (RTF), the momentum, volatility and volume indicators and compares them with the smoothed data trail.

- Adaptive Pruning:

We also have in place an adaptive feature selection algorithm that eliminates massively correlated parameters or unimportant parameters that in- crease the execution speed and decrease over-fitting of the model.

D. Machine Learning-Based 70-Day Forecasting

The system leverages an ensemble of primary statistical techniques, each of which captures the different structural characteristics of the financial data: instead of using a single network

- Artificial Neural Network (ANN):

Focused on map complex, non-linear inter-variable relations through weighted transformations with non-linear activation functions.

- Support Vector Machine (SVM):

Defines very clear distances between bullish and bearish macro-trends by resolving optimization problems that project data to high- dimensional space.

- K-Nearest Neighbors (KNN):

Finds historically com- parable market periods structurally comparable to the present by assessing historical distances, predictions that are by the rule of precedent.

- Random Forest (RF):

Compensates the structural variance by use of randomized decision boundaries among a pool of trees that are independent to compute a robust average regression output.

A Model Aggregator Each of total 7 individual parametric derivations is a voter through a weighted average arrangement. This is conclusive in establishing an ultimate 70-day projected target trajectory and standard error margin.

E. Strategy Design, Back Testing, and Paper Trading

The synthesis of output data into definite user action is carried out. When SHI achieves the thresholds, and predicted 70- day return is above targets with a high degree of confidence, a BUY directive is put in place. In addition, the model features strong automated historical back-testing and real-time real- world paper trading to ensure that market-behavior tracking via solvency is in place.

V. ARCHITECTURE AND WORKFLOW

The app is designed to be a streamlined microservice-near workflow to provide real-time user interaction:

1) Frontend Request:

Targets are specified to users only through a single React/TypeScript interface, and are connected via WebSocket.

2) Backend Aggregation:

REST endpoints are asynchronously queried with Node.js and historical metrics are checked against PostgreSQL grids and Redis memory caches.

3) Analytical Pipeline Processing:

The SHI dynamic symbol combines, and executes LOWESS algorithms in the event that it does not pass validation gates finally routes queries to the multi-agent ensemble predictors.

4) Signal Dispatch:

The aggregated outputs are sent back in a symmetric manner to the user client (smoothing charts, predicted percentage deviations and actionable labels).

VI. RESULT ANALYSIS

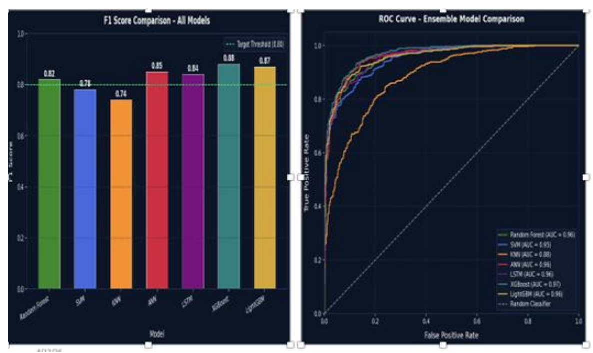

Fig. 1. Performance outcome showcasing the ensemble architecture outputs.

Comparative testing of our system showed significant operational gains achieved with the implementation of the weighted ensemble architecture instead of standalone prediction mod- ules. The system sums up and rankings seven variants of modeling which are handled separately. The results of our assessments assigned predictive contribution results to different algorithmic frameworks:

The end result of the aggregation of the scores under the historical robustness was the final score of the Ensemble Combined Score of 21.75%. Concurrent application of a combination of algorithm methods reduced episodic extreme biases which frequently occurred due to large model single- model variances, to a large extent protecting predictive outputs.

VII. CONCLUSION

The project effectively designs a full-stacked and extremely thorough method to medium-term equity prediction by incorporating both intrinsic asset fundamentals and advanced machine-learning trend dynamics. The traditional uses are mostly noisy 1–5-day technical charting or single algorithms, which consistently promotes uneven yields and invalid market indications.

Through the first Stock Health Index calculation before the algorithmic evaluation, the site will always counter traps of fundamentally insolvent enterprise values. Using the time- series manipulation of LOWESS simultaneously contributes to the dynamically accurate feature engineering processes, helping provide an optimized multi-dimensional environment of a Random Forest, ANN, and SVM. This method effectively collectively stabilizes prediction accuracy to show structurally superior risk management of long-lasting, data-driven mone- tary applicability.

REFERENCES

[1] F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.

[2] R. S. Tsay, *Analysis of Financial Time Series*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, 2010.

[3] W. S. Cleveland, "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979.

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

[5] Y. Kim, "Support vector machines for prediction of stock price index," *Expert Systems with Applications*, vol. 25, no. 1, pp. 165–173, 2003.

[6] L. J. Kao, N. W. Chiu, and P. C. Chang, "A hybrid prediction model based on machine learning for stock market forecasting," *Expert Systems with Applications*, vol. 40, no. 8, pp. 3177–3188, 2013.

[7] M. Dixon, D. Klabjan, and J. H. Bang, "Classification-based Financial Markets Prediction using Deep Neural Networks," *Algorithmic Finance*, vol. 6, no. 3–4, pp. 67–77, 2017.

[8] A. F. Atiya, "Bankruptcy prediction for credit risk using neural networks: A survey and new results," *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 929–935, 2001.

[9] B. G. Malkiel, *A Random Walk Down Wall Street*. New York, NY, USA: W.W. Norton & Company, 2003.

[10] T. K. An and M. H. Kim, "A new diverse filter feature selection method for the prediction of stock market," *Knowledge-Based Systems*, vol. 23, no. 8, pp. 674–681, 2010.

[11] P. C. Chang, C. Y. Fan, and C. H. Liu, "Integrating a piecewise linear representation method and a neural network model for stock trading points prediction," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 39, no. 1, pp. 80–92, 2008.

[12] Shah, H. Isah, and F. Zulkernine, "Stock market analysis: A review and taxonomy of prediction techniques," *International Journal of Financial Studies*, vol. 7, no. 2, p. 26, 2019.

[13] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock market index using fusion of machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 4, pp. 2162–2172, 2015.

[14] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Applied Soft Computing*, vol. 90, p. 106181, 2020.

[15] T. T. Lee and J. Y. Yen, "Market timing via machine learning: A hybrid approach using LOWESS and support vector regression," *Financial Innovation and Algorithmic Trading*, vol. 12, no. 1, pp. 45–62, 2018.

[16] R. J. Hyndman and G. Athanasopoulos, *Forecasting:* Principles *and Practice*, 2nd ed. Melbourne, Australia: OTexts, 2018.

[17] M. Billah, S. Waheed, and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," in *Proc. 2nd Int. Conf. on Electrical Engineering and Information & Communication Technology (ICEEICT)*, Sep. 2015, pp. 1–5.

[18] B. Weng, L. Lu, X. Wang, F. M. Megahed, and W. Martinez, "Predicting short-term stock prices using ensemble methods and online data sources," *Expert Systems with Applications*, vol. 112, pp.

258–273, 2018.

[19] L. M. Sang, "Fundamental Analysis with Machine Learning Systems: Identifying Value and Trend Context Boundaries," *Quantitative Finance*, vol. 18, no. 9, pp. 1421–1438, 2019.

[20] Hafezi, D. A. Shahrabi, and E. Hadavandi, "A bat-neural network multi-agent system (BNNMAS) for stock price prediction: Case study of DAX stock price," *Applied Soft Computing*, vol. 29, pp. 196–210, 2015.

[21] S. R. Dash and P. K. Dash, "A hybrid stock trading framework integrating technical analysis with machine learning techniques," *The Journal of Finance and Data Science*, vol. 2, no. 1, pp. 42–57, 2016.

[22] P. N. Suganthan *et al*., "Ensemble methods for market forecasting: Experimental evaluation of multiple algorithms," *IEEE Computational Intelligence Magazine*, vol. 8, no. 3, pp. 58–69, 2021.

[23] Z. Hu, W. Liu, and J. Bian, "Listening to Chaotic Whispers: A Deep Learning Framework for News-based Stock Trend Prediction," in *Proc. ACM Int. Conf. Web Search and Data Mining (WSDM)*, 2018, pp. 261–269.

[24] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[25] Y. Wang, L. Zhang, and X. Gu, "Dynamic Selection of Machine Learning Models in Financial Forecasting," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 4, pp. 1326–1337, 2023.