Team 128 - The Market Maven: Ensemble Modeling for Strategic Stock Insights

Jacob Wright	Robert Carlton	Alex Shropshire	Manav Kotharia
jwright383@gatech.edu	rcarlton7@gatech.edu	ashropshire6@gatech.edu	mkotharia1@gatech.edu

1 INTRODUCTION

The impact of stock price movements is of critical importance to developing sound economic policies, understanding global and individual investment trends, and securing the financial future of governments and individuals.

1.1 Motivation

Recent research trends in stock price prediction focus on a variety of topics including hyperparameter tuning, feature selection, and limited use of technical, fundamental, and non-financial data; yet no model has distinguished itself as the singular best.

Historically, using data science to predict stocks required manual model development, significant labor hours, and enterprise-grade budgets. Institutions have compiled vast data and built sophisticated methods to try to consistently outperform retail investors and competition.

1.2 Problem Definition

We found that the prevailing academic research limits input data to common indicators, focuses on optimizing a single preselected model type and fails to utilize ensemble modeling. We did not identify existing research that incorporated all three into a single project.

This highlights the need for innovative models that can combine the strengths of all three to deliver greater financial returns than existing models based on RMSE and Sharpe ratios and can yield insights into the drivers of its price predictions.

Our proposed model goes beyond prior research, building on the cited works for sentiment analysis, AutoML for self-constructing ensembles, predictive methods, and the visual display of financial data; while utilizing fundamental, technical, and atypical data to present predictions transparently through the PESTEL framework. Existing financial visualizations primarily consist of line graphs and/or candlestick charts. Our visualization incorporates the predicted stock price, insights into the data used for predictions, and comparisons to peer stock groups providing the investor with considerably more information and insight into the price prediction.

2 LITERATURE REVIEW

Hiransha, et al. [4] and Islam, et al. [6] both contrast the use of ARIMA and neural networks but come to different conclusions. Even today there remains an ongoing dispute over the relative efficacy of each model for time series prediction. These papers led us to search for a way to combine the strengths of multiple predictive domains, hoping to build a synergistic ensemble rather than find a single optimal method.

He, et al. [3] demonstrates methods for AutoML in finance. Their study provides insight into the evolution, current trends, and challenges in AutoML. Their work provided a strong basis for our application of AutoML best practices.
Carta, et al. [1] provides a finance-focused review of autoconfiguring ensembles. While this paper provides a strong basis for ensembles in finance, the authors fail to consider the potential of diverse, cross-category predictive models.

Morales-Hernandez, et al. [10] considers the use of multi-objective hyperparameter optimization. This is a topic of great importance in finance as it is necessary not only to maximize returns but also to minimize drawdown and volatility. This provides innovative competition for the traditional use of compound metrics like the Sharpe ratio in evaluating models.

Existing quantitative prediction strategies approach the price prediction problem from an analytics perspective but it's often clear that insight and nuance from the financial domain are deprioritized in favor of algorithmic sophistication.

In their survey of stock analysis, Nti, Adekoya, and Weyori identified two primary approaches: technical analysis focused on the movement of the stock price itself, which was utilized 66% of the time, and fundamental analysis focused on understanding the business and macroeconomic factors impacting the stock price.

When fundamental analysis is used, 98% of these methods also include social network sites and sentiment analysis as additional predictors[11],[8]. Technical analysis often relies on visually inspecting candlestick charts (with high, low, open, and close prices for a stock) and applying a set of 7,846 'signals' or 'rules' used to guide portfolio decisions [7].

Our review found that numerous statistical models, including logistic regression, ARIMA, GARCH, and linear regression, have been used to predict stock prices. In addition, a wide range of machine learning techniques have been applied to the challenge of stock price prediction. These include recurrent neural networks, LSTM networks, convolution neural networks, and reinforcement learning [5].

Ensemble methods were rarely used in these predictions - a promising area for further research - yet bagging has been used successfully to predict stock price direction [12],[14].

One challenge for any method predicting stock price direction is the variety and volume of trading and fundamental data that is available. Significant difficulty lies in separating useful data from noise. One approach to solving this is the use of feature selection techniques such as embeddings, filters, and wrappings [14].

The paper by S.S. Chen [2] considers the use of macroeconomic indicators, such as inflation rates and unemployment to predict market changes with a focus on recessions. The demonstrated efficacy of macroeconomic indicators for predicting market conditions lends credence to the use of economic changes as indicators in our model.

The works of S. Ko, et al. [9] and M. Raya, et al. [13] contribute valuable insights into various visualization methods. However, they primarily focus on the technical aspects of visualization without exploring innovative, user-centric approaches that could make complex financial data more intuitive. This highlights an opportunity for future research to contribute novel visualization methods that translate complex data into a user-friendly format. Our approach, detailed below, provides insight into the factors impacting the stock's predicted price while offering visual comparisons to similar stocks to increase transparency and understandability.

3 PROPOSED METHOD

3.1 Intuitive Reasoning

Contemporary predictive approaches fall into one or more of the following: technical analysis, fundamental analysis, or advanced quantitative methods. Similarly, novel data sources, often limited to Twitter sentiment analysis or political data, have been evaluated individually within academic literature, but seldom in conjunction with novel predictive methods. Furthermore, while ensemble models have been utilized within each category of financial analysis, there is limited peer-reviewed research into ensembles that span all of these categories.

a. We combine best practices from all three predictive approaches with atypical data in the form of a unique ensemble model. This is done by integrating fundamental data with contemporary quantitative methods for prediction, a largely unexplored practice introduced as 'quantamental investing' by JP Morgan in 2018 [15]. We take this a step further by incorporating the use of technical indicators for investing in our model.

b. Additionally, our project introduces a novel combination of indicators (see section 4.3) to quantify and perform PESTEL analysis on a user-specified stock. PESTEL, which stands for Political, Economic, Social, Technological, Environmental, and Legal, is a qualitative framework for business evaluation. This time-tested approach has been highly effective in the corporate sector, facilitating managers in evaluating the strengths and weaknesses of a business. However, our review of academic literature indicates that the proposed quantitative application of the model is unprecedented.

c. In terms of visual novelty, we have created a unique visualization of the prediction which shows the security's relative strength in each of the six areas of PESTEL. The ability to see the stock price broken down into these components presents a novel spin on factor-based investing while also building trust with the user via increased transparency.

d. Finally, our model combines these aspects using AutoML to form an optimal ensemble for any given stock. This facilitates our other areas of novelty as well as allows us to rapidly build and deploy models, helping our team build something that can compete with models developed by enterprise-scale models.

Combining all of the above methods is something that has not been attempted in peer-reviewed academic literature. This multifaceted approach gives our model a novel edge over contemporary methods.

3.2 Technical Details

Previous attempts at predicting stock direction or price have relied on fundamental indicators, technical indicators, or sentiment drawn from sites such as Twitter/X and Reddit to gauge perception of a given stock. Our model relies on a much broader set of predictors that include technical indicators, social sentiment, stock trader sentiment, company insider trades, and the stock transactions of US House and Senate representatives. This data, when combined with contemporary factors like ESG scores, yields a versatile dataset that stands as an original pool of potential predictors.

We have implemented these indicators using a combination of Financial Modeling Prep API calls and CSV data from sources like the Federal Reserve. We utilized PyCaret's AutoML to search for, train, tune, and ensemble models faster than manual implementations found in research. One key advantage of PyCaret is that we were able to test models that span multiple layers of complexity: (a) simple approaches (incl. ARIMA, auto-ARIMA, Sarima, & exponential smoothing), and (b) traditional ML approaches (incl. XGBoost, Random Forest). It also allowed us to quickly create and compare many model types, rank performance based on specific evaluation metrics, and blend multiple models together to derive a final ensemble algorithm. We've found this ensemble method underutilized in academic literature. True to its AutoML categorization, PyCaret provides a comprehensive suite of automatic cleaning and imputation features, low-code diagnostic visualizations, simple cross-validation, and the ability to self-tune hyperparameters given optimization metrics and grid search parameters.

3.3 Model Visualization

Our core prediction is visualized traditionally using a line chart showing the historical trend and our prediction with upper / lower bounds. Our visual novelty comes from the distinctive representation of the PESTEL framework which has been uniquely adapted for time series forecasting.

Political: We leveraged data on politicians' investment activities to gauge political sentiment. This is visualized as a top-down view of the Senate chamber seats with a needle leaning in the direction of the related political sentiment.

Economic: Our use of various economic indicators is depicted through a line chart as part of the PESTEL analysis.

Social: We extract information about peer companies to evaluate a stock's sentiment in relation to its industrial competitors using data from Twitter/X and StockTwits. The importance of this visualization is allowing the user to see a comparison of sentiment with competitors using an enhanced bar chart.

Technological: For the technological indicator we evaluate each company's research and development spending as a percentage of gross revenue and compare it with peers. This allows us to evaluate the company's technological advantage relative to its competitors.

Environmental: We integrate the use of ESG scores in our analysis as a proxy for the company's environmental strength. By doing this we capture a comprehensive view of the company's environmental, social and regulatory policies and practices.

Legal / Regulatory: This aspect is assessed through trends in bipartisan investments and divestments under the belief that unified spending signals a positive regulatory outlook.

Figure 1 - Project User Interface

4 EXPERIMENTS & EVALUATION

4.1 Experiment Goals

The experiments for our model are devised to comprehensively evaluate its risk-adjusted return, competitive strength against other SOTA models, and its viability for use in live trading. We also evaluate questions left unresolved from our literature review. First, does ARIMA outperform SOTA machine learning techniques, and if so, does the advantage depend on the time horizon of the prediction? Second, is there a singular best model for predicting stock prices or does each stock require a bespoke model to provide optimal results?

4.2 Experiment Setup and Core Ideas

To train our model, we divide data into training and testing sets using more recent data for testing. This is because the usefulness of financial data diminishes as time passes, suggesting more recent data to better reflect real market conditions. We allow the ensemble model to self-train on the training data and then use the testing data to evaluate our model's performance.

4.3 Evaluation

To gauge the performance of our model we use two metrics: RMSE and the Sharpe ratio to evaluate risk and return rather than a metric like MAPE which does not capture volatility.

RMSE is well suited to evaluating and comparing the performance of different models. In stock price prediction, large errors are undesirable and RMSE ensures that models are penalized more heavily for large mispredictions. Further, values are in the same units as the stock prices, making interpretation straightforward. Due to the squaring of errors, RMSE is more sensitive to outliers, which can be problematic as they may represent sudden market movements (e.g. a run on a particular stock) and thus are important to capture accurately. It's a clear and straightforward metric and often used by algorithms as part of their learning process.

But, RMSE is a model metric and does not consider risk and returns of the investment. The

Sharpe ratio indicates how much additional return an investor might get for holding a riskier asset.

$$S = \left(\frac{R_p - R_f}{\sigma_p}\right)$$

Figure 1 - Sharpe Ratio Formula, where R_p is return on portfolio, R_f is return on risk-free investment, and σ_p is the standard deviation of portfolio returns.

As seen in Figure 1 above, by dividing returns by standard deviation, the Sharpe ratio provides a risk-adjusted performance metric that allows us to balance competing objectives of risk and return.

4.4 Experiment Results

We tested our model across a variety of stocks within the technology sector to ensure robust performance across industries, sectors, and sizes. As we have trained models unique to each security, we did not have to adjust for survivorship bias. However, we did check to see if our model can predict trend reversals before they occur.

In order to effectively simulate real-world conditions when evaluating the cumulative returns of our algorithm, we used small percentage fees (0.05%) on each transaction to simulate slippage as the stock price is likely to change slightly when buying or selling a stock. Our tests simulate real-world conditions as well as possible to avoid the common pitfalls described in the literature.

Table 1 - 7 Day Scores

7 Days	Best	2nd	3rd
ARIMA	5	0	1
Bayesian Ridge	1	2	1
Gradient Boost	1	2	0
Elastic Net	0	2	1

Our experiments found that ARIMA appeared as the top model in 5 out of 7 tests using 7-day predictions. This lends credence to traditional methods in the debate over the efficacy of contemporary machine learning and artificial intelligence versus statistical methods. To evaluate the reliance of ARIMA on time horizon we ran the same simulations twice, once for 7 days and once for 30 days of predictions.

Table 2 - 30 Day Scores

30 Days	Best	2nd	3rd
Bayesian Ridge	3	4	0
ARIMA	3	0	1
Elastic Net	1	2	3
Random Forest	0	1	1

Using a scoring based on the weighted placement of each model during cross-validation, ARIMA scored 16 points on the 7-day prediction but dropped to 10 when predicting 30 days out. In the 30-day prediction, it was surpassed by Bayesian Ridge which achieved 17 points.

This indicates that current literature may be asking the wrong questions in regard to ARIMA versus machine learning. Although ARIMA is the strongest single model for short-term prediction, it does not dominate every short-term scenario. Our research found that there is no singular optimal model, but rather different models that are most performant for different scenarios with the time horizon having a significant impact on the efficacy of statistical models like ARIMA.

Regarding the efficacy of ensembles in comparison with individual models, we found that a blended ensemble was selected in six out of seven experiments with the outlier relying entirely on ARIMA. This selection was made based on the cross-validated performance of trained models subject to tuned weighting.

In terms of overall model performance, our blended model achieved an average cumulative return of 4.3337% across multiple 1-week test datasets with a return sample standard deviation of 0.015313%. This resulted in an annualized Sharpe ratio of 30.9 for our backtest which is incredibly strong but expected to drop significantly if the model is applied to live trading.

Risk aversion plays a key role in financial modeling. As previously discussed, we sought to optimize a model to avoid losses in addition to maximizing gains. Pursuant to this goal, our model was tuned to the point of correctly identifying approximately 80% of trend reversals in the test set and guessing the correct direction with 83% accuracy throughout testing. This suggests a high level of reliability in our forecasts to avoid catastrophic losses.

5 DISCUSSION

5.1 Conclusions

Our comprehensive testing across the technology sector stocks demonstrates the efficacy of our model. The performance proves that a combination of ensemble methods with a comprehensive dataset can create a highly performant model, innovatively applying PESTEL analysis to stock investing. We utilized RMSE during training to adjust for risk and employed AutoML to enable bespoke models for each stock.

To further the discussion in academic literature over the efficacy of ARIMA in contrast with machine learning we evaluated the use of ARIMA over different timeframes and securities. We found that, although ARIMA is the strongest model for short-term predictions, its predictive power falls off over time. Although contemporary literature seeks to address the problem of a 'best model' for stock price prediction, we conclude that there is no ideal model for the entire market and unique models must be tuned on a per-stock basis.

For all the advantages AutoML methods provide, many sacrifice the deeper customizations possible in manual implementations in favor of speed and standardization. AutoML may be considered too much of a black box or too prescriptive for certain research projects that demand adjustability at every step.

The innovative visualization of our model includes traditional elements enhanced by the interpretability of the PESTEL framework. This approach provides an intuitive and transparent means of return attribution applied to commonly used business metrics for interpretability.

5.2 Visualization

We developed a unique visualization (or dashboard) with 3 design goals:

1. Provide investors with the visualizations they expect, showing historical data and future stock price predictions. This type of visualization is common and easily interpretable, however, we've added the ability to compare the stock price predictions from the models we utilized.

2. Provide investors with a better understanding of the fundamental, technical, social, and other data used to predict the stock price in an easy-to-understand format and layout. Investors can see how each PESTEL component is used to predict future stock price behavior.

3. Provide investors with comparative data on similar stocks across the PESTEL components (e.g. political, economic, social, technology, environmental & legal) to help them better understand the stock price prediction in the context of other similar stocks.

In our visualization, all of these elements are presented to the investor when they select a specific stock to view its future predicted price and provide investors with a much more comprehensive view of their investment's predicted price, contributory factors, and context with similar stocks.

5.3 Future Research

Our research into stock price prediction has yielded interesting insights into data selection, model optimization, and the use of model selection using AutoML and PyCaret.

Opportunities to extend this research include determining if differing sets of input data for each stock might improve price predictions, understanding which type of data (technical, fundamental, social, or PESTEL) might have a greater impact on the predictions, and the impact of longer prediction time frames (moving from a 7-day window to 30 or 90 days).

6 TEAM CONTRIBUTIONS

All team members contributed equally.

7 REFERENCES

1. Carta, S., Corriga, A., Ferreira, A., Recupero, D. R., & Saia, R. (2019). A holistic auto-configurable ensemble machine learning strategy for financial trading. Computation, 7(4), 67.

2. Chen, S. S. (2009). Predicting the bear stock market: Macroeconomic variables as leading indicators. Journal of Banking & Finance, 33(2), 211-223.

3. He, X., Zhao, K., & Chu, X. (2021). AutoML: A survey of the state-of-the-art. Knowledge-Based Systems, 212, 106622.

4. Hiransha, M., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2018). NSE stock market prediction using deep-learning models. Procedia computer science, 132, 1351-1362.

5. Hu, Z., Zhao, Y., & Khushi, M. (2021). A survey of forex and stock price prediction using deep learning. Applied System Innovation, 4(1), 9.

6. Islam, M. R., & Nguyen, N. (2020). Comparison of financial models for stock price prediction. Journal of Risk and Financial Management, 13(8), 181.

7. Jiang, J., Kelly, B., & Xiu, D. (2023). (Re‐) Imag (in) ing price trends. The Journal of Finance, 78(6), 3193-3249.

8. Jin, Z., Yang, Y., & Liu, Y. (2020). Stock closing price prediction based on sentiment analysis and LSTM. Neural Computing and Applications, 32, 9713-9729.

9. Ko, S., Cho, I., Afzal, S., Yau, C., Chae, J., Malik, A., Beck, K., Jang, Y., Ribarsky, W. and Ebert, D.S. (2016), A Survey on Visual Analysis Approaches for Financial Data. Computer Graphics Forum, 35: 599-617.

10. Morales-Hernández, A., Van Nieuwenhuyse, I., & Rojas Gonzalez, S. (2023). A survey on multi-objective hyperparameter optimization algorithms for machine learning. Artificial Intelligence Review, 56(8), 8043-8093.

11. Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). A systematic review of fundamental and technical analysis of stock market predictions. Artificial Intelligence Review, 53(4), 3007-3057.

12. Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). Efficient stock-market prediction using ensemble support vector machine. Open Computer Science, 10(1), 153-163.

13. M. Raya, D. Srinivasan, V. M, A. Adedoyin and M. Sathiyanarayanan, "Visualizing, Comparing and Forecasting Stock Market Prediction," 2022 IEEE Delhi Section Conference (DELCON), New Delhi, India, 2022, pp. 1-7

14. Sivri, M. S., & Ustundag, A. (2024). An adaptive and enhanced framework for daily stock market prediction using feature selection and ensemble learning algorithms. Journal of Business Analytics, 7(1), 42-62.

15. Slimmon, A., & Delany, L. (2018). Quantamental Investing–The future is now. Investment Insight, Morgan Stanley Investment Management, Solutions & Multi-Asset, May 2018.