# K-Nearest Neighbors Algorithm for Short-to-Medium Term Directional Stock Price Forecasting: An Analysis of Thailand's Banking Sector

Passawan Noppakaew*, Parit Wanitchatchawan, Kanchana Phuhoy, Natthasorn Seubwong
Faculty of Science-Department of Mathematics, Silpakorn University, Nakhon Pathom, 73000, Thailand

***Abstract*—This research explores the efficacy of a parsimonious K-Nearest Neighbors (KNN) framework for short-to-medium term stock direction forecasting, focusing specifically on the banking sector within Thailand's SET50 index. Prior preliminary analysis, aimed at determining the optimal prediction horizon, indicated that a 60-day forecast yielded the most effective results, establishing the scope of this study as medium-term prediction. The objective of this analysis is to determine if a 60-day directional movement can be effectively captured using a minimalist feature set limited to the current day's Opening Price and the previous day's 14-day Simple and Exponential Moving Averages. Employing a rolling-window validation methodology on seven key banking stocks during H1 2025, the KNN model demonstrated significant predictive capability. The average accuracy across the selected banking stocks reached 82.0%, with standout performance for TISCO and SCB. While results varied across stocks, our findings substantiate the theoretical and practical sufficiency of a simplicity-first approach. The research demonstrates that in high-noise emerging markets, feature sparsity and instance-based logic serve as an essential defense against overfitting, providing institutional practitioners with a transparent and robust alternative to complex methodologies.**

***Keywords*—Stock trend prediction; K-Nearest Neighbors; technical analysis***

## I. Introduction

The intricate, volatile, and non-linear nature of financial markets presents a formidable and long-standing challenge [1]. Accurately forecasting stock price movements has been a desirable objective for both investors and academics [2], [3], as marginal improvements in predictive accuracy can offer significant financial advantages. Based on the limitations of traditional linear models for forecasting stock movements, alternative approaches have been explored. Machine learning (ML) which is a sub-field of artificial intelligence (AI) has emerged as a powerful paradigm for identifying the complex latent structures within financial time-series data [4], [5], [6]. See also [7], [8], [9], [10] for alternative approaches using deep learning techniques.

Unlike conventional models, ML algorithms excel at identifying non-linear patterns and hidden correlations within high-dimensional datasets [11]. This capability has led to widespread applications across various domains, from natural language processing and computer vision to intelligent recommendation engines [12]. The research efforts within ML-based financial forecasting can be broadly divided into two main approaches: one focusing on price prediction and the other on market directional forecasting. The first approach treats the problem as a regression task, attempting to forecast the exact future value or closing price of a stock. While academically rigorous, this method encounters significant challenges due to the high level of stochastic noise in market data. The second approach, which is the focus of this study, frames the problem as a classification task, predicting the directional movement, such as upward or downward trends, of the stock price. This classification approach offers distinct practical advantages. For a trader or investor, their decision is often binary, i.e., buy or sell, not the prediction of an exact price point. A model that correctly classifies the direction of a trend, even if it misses the magnitude, provides a clearer and more actionable signal for formulating trading strategies. The focus on direction rather than price is often more robust and less susceptible to the noise that plagues regression models.

Among the suite of ML tools, the K-Nearest Neighbors (KNN) algorithm, a non-parametric and instance-based learning method, has attracted significant attention due to its simplicity, transparency, and effectiveness [13]. KNN operates on the principle of feature similarity, classifying a new data point based on the majority class of its $k$ nearest neighbors in the feature space. Due to its reliance on local similarity rather than global parameter estimation, KNN-based frameworks are effectively utilized to mitigate overfitting in noisy financial time series, particularly when coupled with strategic feature selection [14]. Its utility has been validated in diverse fields, such as the trading system proposed by Teixeira and de Oliveira [15], who utilized a high-dimensional feature space of 22 technical indicators to achieve significant results. However, recent advancements in emerging market research suggest a strategic shift toward binary directional classification as a more robust objective than point-based price estimation [6]. Specifically, in the context of the Stock Exchange of Thailand (SET), evidence indicates that focusing on market direction allows for higher predictive stability despite the inherent volatility of the Thai banking sector [6].

Consequently, this study diverges from the feature-rich precedent by investigating a minimalist, direction-oriented hypothesis. Rather than adopting the extensive 22-indicator framework, the current research examines whether effective directional forecasting for the Thai banking sector is achievable through a deliberately restricted feature set, thereby enhancing both model transparency and practical interpretability. This study investigates the hypothesis that a simple KNN classifier can effectively perform directional forecasting even with minimal input data. The analytical focus is narrowed specifically to

*Corresponding author.

stocks within the Banking sector, as listed in Thailand's SET50 Index for the H1 2025 period, namely BBL, KBANK, KTB, KTC, SCB, TISCO, and TTB. This sector presents a particularly interesting case for study due to several factors: bank stocks often exhibit a degree of sensitivity to macroeconomic indicators, their performance can be interrelated, reflecting broader economic health, and they represent a significant component of the overall market index, making their predictability valuable for investors. Within this focused scope, the objective is to predict whether the closing price 60 days in the future will be up (higher) or down (lower) relative to the current day's opening price. This prediction utilizes a deliberately minimalist feature set consisting of: 1) the current day's Opening Price, 2) the 14-day Simple Moving Average (SMA) from the previous day, and 3) the 14-day Exponential Moving Average (EMA) also from the previous day. The use of opening prices serves to enhance economic realism by simulating decision-making at the start of the trading session. By evaluating this model specifically on these key banking stocks, this research aims to determine if the minimalist KNN approach provides a practical and effective basis for directional forecasting within this vital market sector. To distinguish the unique contributions of this study from existing KNN-based forecasting literature, this study pivots toward a parsimonious framework that challenges the prevailing necessity of model complexity in high-noise environments. Methodologically, this study demonstrates that a minimalist feature set can maintain high predictive accuracy by prioritizing robustness over complexity. Such an approach serves as a strategic defense against the overfitting typically encountered in algorithmically dense architectures within high-noise environments. Beyond its empirical setting, the choice of the Thai banking sector serves as a strategic economic bellwether. Grounded in the theoretical consensus that financial systems are primary engines of national stability [16], [17], this research leverages the heightened sensitivity of these securities to macroeconomic shifts as a rigorous stress test for model reliability [18]. Finally, the research offers significant practical utility by providing a transparent, instance-based logic for institutional practitioners. Unlike complex models, the KNN framework offers superior interpretability, allowing users to substantiate forecasts through historical analogies–a vital requirement for real-world trading deployment.

The remainder of this study is organized into four primary sections to provide a comprehensive overview of the research framework: Section II details the research methodology, specifically focusing on the data selection process within the Thai banking sector and the implementation of the deliberately minimalist feature set. Section III reports the empirical results and predictive performance across the selected constituents. Finally, Section IV provides a synthesis of key findings, summarizing the practical utility of the parsimonious model and offering strategic recommendations for subsequent research directions.

## II. METHODOLOGY

This study employs a quantitative methodology to ascertain the effectiveness of a KNN model for financial market classification. The process involves data acquisition from a major emerging market index, precise variable features, and a model training and validation process.

### A. Dataset and Scope

Instead of serving as a nominal constraint for this study, the focus on the Thai banking sector represents a strategic decision to test the proposed model within a pivotal economic bellwether [17]. In an emerging market such as Thailand, the banking industry acts as a direct proxy for broader national stability, characterized by its acute sensitivity to monetary policy shifts, interest rate volatility, and macroeconomic fluctuations [16], [18]. This inherent sensitivity creates a high-variance and information-rich environment which is an ideal test for evaluating whether medium-term technical indicators can maintain predictive reliability under shifting conditions. To uphold the analytical reliability of this research and shield the model from the stochastic noise inherent in low-volume securities, this study deliberately focuses on banking constituents within the SET50 Index. This approach ensures the study is grounded in the market's most liquid and institutionally significant securities.

For this study, the banking constituents were identified from the official SET50 list [19] for the H1 2025 period starting from January 1, 2025, to June 30, 2025, namely BBL, KBANK, KTB, KTC, SCB, TISCO, and TTB, as shown in Table I. Concentrating on this sector allows a focused examination of the model's performance within a group of stocks often influenced by similar macroeconomic factors.

TABLE I. THE BANKING SECTOR IN SET50 INDEX (H1 2025: JAN 1 - JUN 30, 2025)

| No. | Symbol | Company Name |
|---|---|---|
| 1 | BBL | BANGKOK BANK PUBLIC COMPANY LIMITED |
| 2 | KBANK | KASIKORNBANK PUBLIC COMPANY LIMITED |
| 3 | KTB | KRUNG THAI BANK PUBLIC COMPANY LIMITED |
| 4 | KTC | KRUNGTHAI CARD PUBLIC COMPANY LIMITED |
| 5 | SCB | SCB X PUBLIC COMPANY LIMITED |
| 6 | TISCO | TISCO FINANCIAL GROUP PUBLIC COMPANY LIMITED |
| 7 | TTB | TMBTHANACHART BANK PUBLIC COMPANY LIMITED |

All necessary historical time-series data (including daily Open, High, Low, Close, and Volume) for these selected banking stocks were programmatically retrieved using the Yahoo Finance (yfinance) library. To establish an appropriate prediction horizon for the main study, a preliminary experiment was conducted. This involved evaluating the KNN model's average prediction accuracy across various forward-looking horizons ($n$ days), ranging from 3 to 120 days, while keeping the feature set constant. The results of this sensitivity analysis indicated that model balances between achieving near-maximal predictive performance and maintaining a forecast window appropriate for short-to-medium-term trading strategies at $n = 60$ days (details in Appendix A, Appendix B, and Appendix C). Therefore, a 60-day prediction horizon was adopted for the final evaluation. The final out-of-sample prediction (testing) phase, using this 60-day target, was conducted on data from January 1, 2025, to June 30, 2025.

### B. Study Design and Variable Engineering

This research is designed as a binary classification task. The objective is to predict the directional movement of a stock's price over a fixed, 60-day forward-looking horizon.

*1) Target variable (dependent variable):* For any given trading day $t$, the target variable $Y(t)$ is defined as the future price direction. It is assigned a binary value based on a comparison between the closing price on day $t$ and the closing price 60 days later ($t + 60$):

- Class 1 (Up) if $Close_{t+60} > Close_t$,
- Class 0 (Down) if $Close_{t+60} \leq Close_t$,

where, $Close_t$ represents the closing price on the prediction day (day $t$) and $Close_{t+60}$ is the closing price 60 days later. This definition is designed to directly model the practical trading decision of determining whether buying a stock at today's closing price will yield a higher closing price 60 days from now.

*2) Feature set:* Unlike models that rely on a high-dimensional feature space, this study adopts a deliberately minimal approach. Intricate feature engineering is sidestepped in favor of utilizing fundamental and widely recognized technical indicators common in financial analysis. Specifically, the feature set is constructed solely from the current day's Opening Price, alongside the previous day's 14-day Simple Moving Average (SMA) and 14-day Exponential Moving Average (EMA). Further details on these moving averages are provided in [20]. This deliberate simplicity allows for a focused evaluation of the core predictive power inherent in these basic trend metrics.

The resulting feature set $X(t)$, used to predict the outcome for day $t$ consists of only three variables:

- Open Price ($Open_t$) is the opening price on the day of prediction.
- SMA-14 ($SMA_{14(t-1)}$) is the 14-day Simple Moving Average, calculated using closing prices up to the previous trading day ($t - 1$).
- EMA-14 ($EMA_{14(t-1)}$) is the 14-day Exponential Moving Average, also calculated using data up to the previous trading day ($t - 1$).

The 14-day lookback period for the moving averages was selected to align with the empirically determined optimal prediction horizon (see Appendix A, Appendix B, and Appendix C) and reflects common practice in short-to-medium-term technical analysis. Fine-tuning the moving average periods was considered beyond the scope of this study, which focuses on the effectiveness of the chosen minimalist feature types. This specific feature construction is critical, as it ensures that only information known at the time of prediction (the morning of day $t$) is used, thereby preventing any lookahead bias.

*3) Model training and validation:* The KNN algorithm was selected as the core classifier. To simulate realistic market conditions and ensure model robustness, a rolling-window validation procedure was implemented.

For each of the 7 stocks, and for each day in the test period (Jan-Jun 2025), the model was trained as follows:

- Data Window: A historical data window of approximately 1 year was retrieved relative to the prediction day (e.g., for a prediction on June 1, 2025, data from June 1, 2024, to May 31, 2025, will be used).
- K-Value Optimization: This historical window was split into a training set and a testing (validation) set (3:1 ratio). The KNN model was run iteratively to find the optimal hyperparameter $k$ (the number of neighbors) that yielded the highest accuracy on this validation set.
- Prediction: The full historical window was then used as the training set, and the optimized $k$ and the model created by the historical data window ($\approx$ 1 year) was used to make the single directional prediction, which is class 0 or 1, for the target day.

This process was repeated for every trading day in the test period, ensuring the model was always retrained on the most recent data available before making a new prediction. The final accuracy, precision and recall were calculated as the percentage of correctly predicted days over the entire 6-month test period. Minimalist approach provides a practical and effective basis for KNN-based directional forecasting.

## III. Results

This section presents the empirical findings from applying the KNN classification model, configured with the minimalist feature set and a 60-day prediction horizon, to the selected banking stocks within Thailand's SET50 Index (H1 2025). The model's performance is evaluated on a per-stock basis using standard classification metrics, focusing primarily on prediction accuracy, recall, and precision, over the designated test period. The predictive performance of the KNN model across the selected banking stocks is detailed in Table II and visualized in Fig. 1. Overall, the model demonstrates considerable effectiveness for most stocks within this sector, although performance varies significantly between individual securities.

TABLE II. KNN Model Performance Metrics for SET50 Banking Stocks (H1 2025).

| Stock Symbol | Accuracy (%) | Precision (%) | Recall (%) |
|---|---|---|---|
| BBL | 61.0 | 40.5 | 44.7 |
| KBANK | 83.9 | 93.4 | 86.7 |
| KTB | 80.5 | 84.9 | 92.8 |
| KTC | 84.7 | 72.3 | 100.0 |
| SCB | 89.8 | 93.7 | 95.4 |
| TISCO | 94.9 | 96.3 | 98.1 |
| TTB | 78.8 | 91.3 | 80.2 |
| **Average** | **82.0** | **81.8** | **85.4** |

Both SCB and TISCO exhibit strong performance with high accuracy, precision and recall. Especially, TISCO emerges as the top performer, achieving exceptionally high scores across all metrics. This suggests that the model is highly reliable and effective at predicting the 60-day directional movement for these specific stocks. Whereas, several other banks, including KBANK, KTB, and KTC, also show robust results, generally achieving accuracies above 80%. Especially, KTC displays a perfect recall score, indicating the model successfully identified all actual upward movements within the test period for this stock. However, this comes with a lower precision compared to KBANK and SCB, implying a higher rate of false positives. KTB also shows a strong recall paired with moderate precision. TTB presents good precision, but

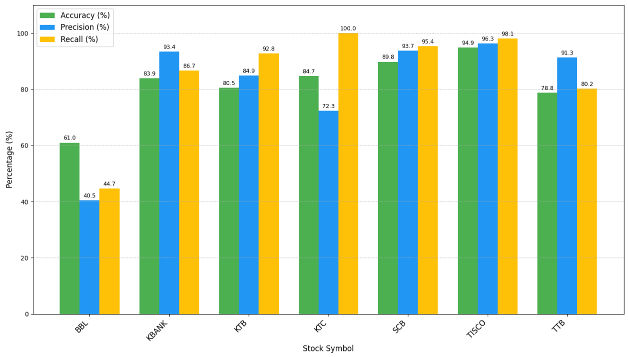


Fig. 1. Bar chart of KNN performance metrics for banking stocks.

lowers recall compared to the top performers. In contrast, BBL shows significantly weaker results across all aspects of the performance metrics. The model struggles considerably with predicting this stock's movement compared to the others in the sector.

Considering the average performance across the seven banking stocks, the KNN model with the selected features appears generally proficient for this sector. The high average recall suggests the model is good at capturing upward trends, while the respectable average precision indicates that when it predicts an upward movement, it is correct about 82% of the time on average.

## IV. Conclusion

This research set out to explore the viability of a simplified approach to predict stock price direction within the Thai financial market. The study specifically investigated whether a KNN classifier, equipped with a deliberately minimalist set of technical indicators (Opening Price, 14-day SMA, 14-day EMA), could effectively forecast medium-term (60-day) upward and downward movements for stocks in the SET50 banking sector. The selection of the 60-day horizon was itself informed by a preliminary sensitivity analysis aiming for a balance between predictive power and practical trading relevance.

The study showed that the KNN model demonstrated considerable predictive capability, achieving an average accuracy of 82.0% across the sector. Performance was particularly strong for certain stocks like TISCO (94.9% accuracy) and SCB (89.8% accuracy), suggesting the model can be highly effective in specific cases. The findings indicate that even without complex feature engineering or computationally intensive algorithms, a basic KNN approach can capture meaningful predictive signals for medium-term directional changes, at least within this homogeneous sector. The high average recall (85.4%) further suggests a particular aptitude for identifying potential upward trends. However, the results were not uniform. The model's significantly lower performance for BBL underscores that predictive success, even with a consistent methodology, can be highly stock-specific.

The robust accuracy achieved through this minimalist framework underscores the theoretical and practical sufficiency of a simplified KNN approach, even in the absence of complex algorithmic benchmarks. These findings suggest that in high-noise environments like the SET50 banking sector, feature sparsity and low model complexity may actually serve as a primary defense against the overfitting typically observed in more elaborate architectures. By demonstrating that significant market signals can be effectively captured without the inherent computational overhead of complex models, this study provides a compelling case for a parsimonious modeling framework. Furthermore, the instance-based logic of KNN offers superior interpretability—a vital requirement for practitioners who prioritize understanding the historical analogies driving a prediction as much as the forecast itself.

Future research could productively expand upon this work by comparing the minimalist KNN approach against other classifiers (e.g., Logistic Regression, Naive Bayes) and applying the methodology to diverse market sectors and international markets. In addition, future research will investigate adaptive methods for selecting the optimal prediction horizon and moving average periods. Despite its limitations, this study affirms the potential utility of simple, interpretable models like KNN in the challenging domain of financial market prediction.

## Acknowledgment

The authors wish to thank Department of Mathematics, Faculty of Science, Silpakorn University. We would also like to thank anonymous reviewers for valuable comments and suggestions.

## References

[1] E. 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] P. C. Tetlock, *Giving content to investor sentiment: The role of media in the stock market*, The Journal of Finance, vol. 62, no. 3, pp. 1139–1168, 2007.

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

[4] W. Huang, Y. Nakamori, and S. Y. Wang, *Forecasting stock market movement direction with support vector machine*, Computers & Operations Research, vol. 32, no. 10, pp. 2513–2522, 2005.

[5] S. Basak, S. Kar, S. Saha, L. Khaidem, and S. R. Dey, *Predicting the direction of stock market prices using tree-based classifiers*, The North American Journal of Economics and Finance, vol. 47, pp. 552–567, 2019.

[6] K. Saetia and J. Yokrattanasak, *Stock Movement Prediction Using Machine Learning Based on Technical Indicators and Google Trend Searches in Thailand*, International Journal of Financial Studies, vol. 11, no. 1, Art. no. 5, 2023.

[7] E. W. Saad, D. V. Prokhorov, and D. C. Wunsch, *Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks*, IEEE Transactions on Neural Networks, vol. 9, no. 6, pp. 1456–1470, 1998.

[8] G. S. Atsalakis and K. P. Valavanis, *Surveying stock market forecasting techniques – Part II: Soft computing methods*, Expert Systems with Applications, vol. 36, no. 3, pp. 5932–5941, 2009.

[9] 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.

[10] A. S. Baimuruev and A. B. Zharmagambetov, *Review of artificial intelligence methods and applications*, Journal of Physics: Conference Series, vol. 1889, no. 2, p. 022014, 2021.

[11] S. B. Kotsiantis, *Supervised machine learning: A review of classification techniques*, Informatica, vol. 31, pp. 249–268, 2007.

[12] M. I. Jordan and T. M. Mitchell, *Machine learning: Trends, perspectives, and prospects*, Science, vol. 349, pp. 255–260, 2015.

[13] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ: Pearson, 2010.

[14] S. Lahmiri and S. Bekiros, *Randomness, informational entropy, and volatility interdependencies among major world markets: A k-nearest neighbors approach*, Chaos, Solitons & Fractals, vol. 132, Art. no. 109509, 2020.

[15] L. A. Teixeira and A. L. I. de Oliveira, *A method for automatic stock trading combining technical analysis and nearest neighbor classification*, Expert Systems with Applications, vol. 37, no. 10, pp. 6885–6890, 2010.

[16] R. G. King and R. Levine, *Finance and Growth: Schumpeter Might Be Right*, The Quarterly Journal of Economics, vol. 108, no. 3, pp. 717–737, 1993.

[17] R. Levine, *Financial Development and Economic Growth: Views and Agenda*, Journal of Economic Literature, vol. 35, no. 2, pp. 688–726, 1997.

[18] L. Menkhoff and C. Suwanaporn, *10 Years after the crisis: Thailand's financial system reform*, Journal of Asian Economics, vol. 18, no. 1, pp. 4–20, 2007.

[19] The Stock Exchange of Thailand (SET). (2025). *Constituents List SET50 & SET100*. [Online]. Available: https://www.set.or.th/en/market/information/securities-list/constituents-list-set50-set100 [Accessed: June 14, 2025].

[20] J. J. Murphy, *Technical Analysis of the Financial Markets*. New York, NY: New York Institute of Finance, 1999.

## APPENDIX

Before proceeding with the main analysis, it was crucial to empirically justify the selection of the prediction horizon ($n$). An arbitrary choice could weaken the study's validity. Therefore, a sensitivity analysis was conducted to evaluate the KNN model's performance across various forward-looking timeframes.

### A. Methodology

The core experimental setup mirrored the main study's design, with only the prediction horizon ($n$) being varied.

- Horizons Tested: A range of plausible short-to-medium-term horizons was selected for comparison. The parameter $n$ was defined using the following intervals: $3, 5, 7, 10, 14, 21, 30, 45, 60, 90, 120$ days.
- Target Variable Definition: For each tested horizon $n$, the binary target variable $Y(t)$ was redefined as:
  - Class 1 (Up) if $Close_{t+n} > Close_t$,
  - Class 0 (Down) if $Close_{t+n} \leq Close_t$.
- Constant Factors: The minimalist feature set (current Open Price, previous day's SMA-14, previous day's EMA-14) and the KNN algorithm remained constant across all tests.
- Optimal $k$ Selection: For each stock and each prediction day, the optimal number of neighbors ($k$) was determined. This involved:
  - Splitting the 1-year historical data starting from January 1, 2024, to December 31, 2024 into a 3:1 training/validation set.
  - Iterating through $k$ values from 1 to 30.
  - Selecting the $k$ that yielded the highest accuracy on the validation set for that specific stock and that specific training window.
- Performance Evaluation: Using the optimized $k$, predictions were made for the target day within the rolling window procedure. This entire process was repeated for all 50 stocks across the full test period (October 1, 2024, to December 31, 2024) for each of the selected horizons ($n = 3, 5, ..., 120$). The performance metric was the average prediction accuracy, precision and recall calculated across all 50 stocks for each horizon.

### B. Results

The key metrics tracked were average accuracy, precision, and recall. The results are presented in Table III and visualized in Fig. 2.

TABLE III. AVERAGE PERFORMANCE METRICS VS. PREDICTION HORIZON ($n$ DAYS).

| **Prediction Horizon ($n$ days)** | **Average Metrics (%)** | | |
|---|---|---|---|
| | **Accuracy** | **Precision** | **Recall** |
| 3 | 61.0 | 56.0 | 47.0 |
| 5 | 65.0 | 65.0 | 54.0 |
| 7 | 68.0 | 67.0 | 66.0 |
| 10 | 69.0 | 67.0 | 64.0 |
| 14 | 70.0 | 70.0 | 67.0 |
| 21 | 73.0 | 74.0 | 70.0 |
| 30 | 75.0 | 75.0 | 74.0 |
| 45 | 75.0 | 78.0 | 73.0 |
| 60 | 77.0 | 78.0 | 78.0 |
| 90 | 77.0 | 79.0 | 78.0 |
| 120 | 79.0 | 80.0 | 80.0 |

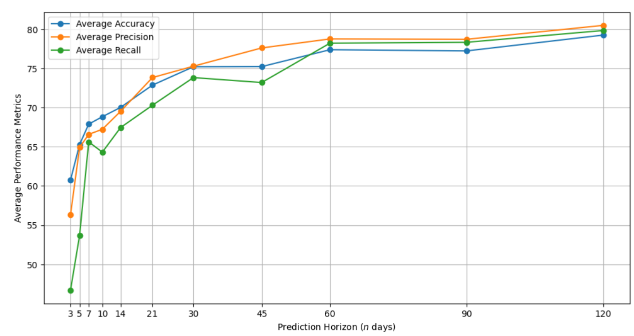


Fig. 2. Average performance metrics vs. prediction horizon ($n$ days) up to 120 days.

As depicted in Fig. 2, all three performance metrics generally exhibit an upward trend across the entire tested range. Notably, the rate of improvement decelerates considerably after approximately $n = 45$ days. Average Precision, in particular, begins to plateau around $n = 60$ days, showing only marginal increases thereafter up to $n = 120$. While the highest average accuracy $\approx 79\%$ was observed at $n = 120$, the gain compared to $n = 60$ which has the average accuracy $\approx 77\%$ is relatively small, especially considering the doubling of the forecast period from approximately two months to four months. Moreover, $n = 120$ is beyond the typical scope of short-to-medium-term trading. Therefore, a horizon of $n = 60$ days (approximately two months) was chosen for this study. This decision represents a deliberate trade-off. It captures the vast majority of the predictive performance achievable while maintaining a forecast window that remains relevant and actionable for medium-term position adjustments.

### C. Conclusion

Based on this empirical evidence, the 60-day prediction horizon ($n = 60$) was selected as the most suitable timeframe for the main classification task in this study. The sensitivity analysis in Fig. 2 supports this choice, indicating significantly diminishing improvement rates in performance gains for horizons substantially longer than 60 days.