

Case Study

Market movement prediction using chart patterns and attention mechanism

Witawat Norasaed¹ · Thitirat Siriborvornratanakul¹

Received: 10 August 2023 / Accepted: 26 December 2023

Published online: 09 January 2024

© The Author(s) 2023 [OPEN](#)

Abstract

The new concept of forecasting market movements is introduced to fill the gap in stock price forecasting that uses a fixed time window. The candlestick data are grouped into waves before being input into the Transformer model to capture many different chart patterns and provide longer price movement patterns for better feature extraction. This work is divided into two main parts. The first part is wave extraction, for which we propose a new algorithm for grouping candlestick data into waves. The second part deals with Transformer modeling. The result of this work has shown that the Transformer model trained using wave series can help traders reduce their losses and choose a suitable trading strategy based on the prediction results. The model trading result of the test set has earned about 19.8% (using fixed 1 contract size per order) which is better than using MACD signal which can give earning of 12.5%.

Keywords Deep learning · Transformer · Attention · Chart pattern analysis · Market trend prediction

1 Introduction

The research area in price prediction [1, 8, 9] is crucial for the research community as it addresses fundamental challenges in financial markets, providing valuable insights for investors, policymakers, and financial institutions. Advancements in this field not only contribute to more accurate risk assessment and investment strategies but also enhance our understanding of market dynamics, fostering the development of robust financial models. In particular, predicting financial markets has long been a challenging research topic due to the complexity of the market, associated with many factors such as the global economy, natural events, and market news. Despite this difficulty, the stock chart (e.g., tradingview.com) is one of the technical analysis tools used in various ways to support trader's decisions (Nazário, Silva, Sobreiro, & Kimura, [12]) as we often see a professional trader workstation with many monitors for multi-timeframe charts. The stock chart is also adapted as a feature for machine learning trading algorithms, employing various forms such as images of daily candlesticks (Cohen, Balch, & Veloso, [4]), three-bar candlesticks [3], and images of stock price line charts [10].

Chart pattern analysis, a technique for stock analysis, utilizes past price movements' patterns in the stock chart to predict future market movements. For instance, the "triangle pattern" is a combination of waves, involving two lower highs and two higher lows. Elliott Wave Theory [2, 7], published by Ralph Nelson Elliott in his book "The Elliott Wave Principle" in 1938, supports the use of historical patterns to predict the future price, emphasizing the stock market's tendency to move in repeating wave patterns. These chart patterns aid in developing trading strategies by assessing the probability of successful trades based on historical price patterns. Detecting chart patterns involves three categories [17]. Firstly,

✉ Thitirat Siriborvornratanakul, thitirat@as.nida.ac.th; Witawat Norasaed, witawat.nor@stu.nida.ac.th | ¹Graduate School of Applied Statistics, National Institute of Development Administration (NIDA), Bangkok, Thailand.



rule-based techniques [14, 16] use a mathematical algorithm with a set of numerical rules to capture a pattern from a stock chart. Secondly, template matching [11, 19] identifies chart patterns based on the similarity between the template chart and stock charts. Finally, machine learning (ML) and deep learning (DL) techniques have been used to propose several methods for identifying chart patterns and extracting chart representation data, such as clustering model [3], time series prediction [5, 6, 20], convolutional neural networks (CNN) [4, 10]. Machine learning and deep learning are the preferred methods to extract abstract features from the stock charts, as the rule-based and template-matching techniques have common limitations in extracting multiple chart patterns. Recent studies demonstrate that deep learning is superior to machine learning in its performance [15], especially the Transformer model [5, 6, 18, 20].

Most studies on time series have used either candlesticks or the entire stock chart as input, with a fixed window size (e.g., [5, 6, 10, 20]). However, we propose that the input window size should vary depending on each period's trend to capture different chart patterns. In this paper, instead of using price data directly, we experiment with the concept of using a series of waves as chart pattern representations in combination with other technical indicators. We propose a method for grouping price data into a series of waves, which are then input into a Transformer-based attention mechanism to find the knowledge of historical patterns and predict the next wave's characteristics (price movement range). The prediction result focuses on the short-term market trend rather than the price to avoid market fluctuations. This work is divided into two main parts: wave extraction (Sect. 3) and Transformer modelling (Sect. 4).

2 Related works

This paper focuses on applying chart representation to a prediction model. Thus, these two main topics are explored and summarized in this section.

2.1 Chart pattern representation

The key focus is to develop a feature similar to the concept of chart pattern recognition but at a lower level. There are various proposed methods for detecting chart patterns, which can be divided into three categories [17].

- (1) Rule-based techniques: This method employs a mathematical algorithm with a set of numerical rules to capture patterns from stock charts. For instance, Savin et al. [14] proposed a head and shoulders pattern identification technique using a mathematical algorithm to identify peaks of the head and shoulders pattern with a specific head and shoulders ratio. Sim et al. [16] used the differential price as a criterion for the chart representation.
- (2) Template matching: This technique identifies chart patterns based on the similarity between the template chart and stock charts. The template chart is a numerical representation, e.g., a weighted matrix of a given chart pattern [11] and a Haar wavelet transform of the stock chart [19].
- (3) Machine learning techniques: Several methods use machine learning to identify chart patterns and extract chart representation data. For example, [3] uses a clustering model to group stock charts and candlestick patterns into clusters for further analysis. [4, 10] Use a convolutional neural network (CNN) to extract features from stock charts instead of using numerical time series.

2.2 Prediction model

The transformer model is chosen for this work because it has proven to be suitable for time series problems compared to the original recurrent neural networks (RNN). In recent years, the attention mechanism has been introduced into financial market forecasting in various forms.

Zhang and Zhang [20] experimented with price prediction using long short-term memory (LSTM), LSTM-Attention, and Transformer by predicting the next day's price using 30 days' price as input. This experiment confirmed that the attention mechanism (accuracy score of 0.601 for LSTM-Attention and 0.582 for Transformer) performed better than LSTM (accuracy score of 0.556). It was suggested that other factors besides the time feature should be included to obtain better results.

Daiya et al. [5] proposed a combination of convolution neural networks (CNN), natural language processing (NLP), and attention mechanisms (called Att-DCNN and Att-biNTN respectively) to predict the next day's price using 40 days of data. The result showed that the Att-DCNN achieved approximately a 3% improvement in accuracy, rising from 68.81 to

72.23% over the baseline (RNN-based) model. According to Daiya et al. [6], a Transformer was constructed on top of the previous Att-DCNN and another Transformer-Encoder which was utilized to encode news information. The output of this Transformer was then passed to the softmax layer to predict whether the price would increase or decrease.

3 Wave extraction

This section aims to strategically group price data into wave series, capturing the essence of small trends. Recognizing that chart patterns are constructed from these nuanced trends, and considering the indeterminate length of each pattern, our approach focuses on predicting entire upcoming trends. This aligns with practical trading considerations, given the perpetual fluctuations in the market. Merely predicting the next candle price on the chart lacks meaningful context. Figure 1 illustrates the envisioned outcome of this process, which directly correlates with the ultimate objective of our work—predicting the movement of the next wave to formulate effective trading strategies for distinct movement patterns. To discern price movement, we employ straightforward two-line trend lines that delineate the channel of the price movement. The upper line signifies the high price, while the lower line represents the low price. As each wave concludes, the model forecasts both upper and lower lines, providing a guiding framework for the impending wave. This method enhances the robustness of our predictive approach, ensuring a more comprehensive understanding of market dynamics.

3.1 Data

The raw data in this study is XAUUSD (gold spot per U.S. dollar) public historical 1 min bar chart data from histdata.com (Apr 2009–Jul 2021). The data is in Open-High-Low-Close (OHLC) format, which means that one data point contains 5 columns (date and OHLC). The 1 min time frame is combined into common time frames for other frequencies, namely 5-, 15-, 30-, and 60 min time frames. In total, the 4,358,731 data of the 1-minute time frame are converted into 872,498 data for the 5 min time frame, 292,367 data for the 15 min time frame, 146,353 data for the 30 min time frame, and 73,699 data for the 60 minute time frame.

The data are divided into 3 sets by date: Training set 80% of the data (Apr 2009–Dec 2018), Validation set 10% of the data (Jan 2019–Apr 2020), and Test set 10% of the data (May 2020–Jul 2021). This splitting ratio of 80:10:10 is used to contribute majority of our data to the training data.

3.2 Capture trend

To create a price movement range for each small trend, two linear lines are utilized to group the price data into waves. These waves serve as both features and targets for prediction. The decision is made to employ the linear least squares method for fitting two linear lines. Specifically, the upper linear line is fitted to the high of each candle, while the lower linear line is fitted to the low of each candle. The channel progresses candlestick by candlestick, persisting as long as the price remains within the same trend. Figure 2 provides an illustrative example of this algorithm in a 15 min timeframe. The loss from the linear least square fit is measured at each step, serving as a key parameter to determine the end of the trend.

Fig. 1 A handcrafted example (Credit: tradingview.com). The yellow dashed line is the desired outcome of Wave Extraction



Fig. 2 Example of a wave extraction algorithm to capture the price trend in the 15 min time frame. The trend, which is represented by blue straight lines, starts at step (a) and the algorithm continues to capture this trend in step (b). Then, the trend is stopped in step (c) when the loss is above the cut-off. After step (c), the algorithm will start to capture the next trend



Step a.



Step b.



Step c.

3.3 End of trend

The end of a trend is gauged using two weighted parameters: the loss of the least square error of the two lines and the number of candlesticks in the trend. This dual-parameter approach aims to strike a balance between promoting the length of the trend and mitigating the error of the linear least square fit. Through experimentation, it was observed that the loss of least squares at the beginning of the trend often fluctuated due to low data points, leading to undesirable truncation when the error spiked from the previous step. The inclusion of the number of candlestick wedges helps smooth the overall loss at the trend's inception, thereby improving the overall trend fit. The weighted parameters have been meticulously fine-tuned, accompanied by a sample visualization of trend fit from various periods of the training dataset. Figure 3 presents an example of wave extraction from the same period data using different weighted parameters. The color variations in Fig. 3 are solely for visualization purposes and hold no specific meaning for the model.

The wave extraction algorithm was applied to our price data across different time frames, transforming the raw candlestick data into a concise sequence of waves. This process significantly reduces the volume of data, as depicted in Table 1.

4 Transformer model

This section will focus on the development of Transformer models, starting with data preparation and ending with model evaluation. Since several studies have shown that the Transformer model has better performance in predicting the stock market [5, 6, 20], our work mainly focuses on tuning the Transformer model without comparison with other types of time series models. In this study, we used an NVIDIA GeForce GTX 960M as the main resource and used TensorFlow framework version 2.6.0, which took about 2–3 hours to train 700 k trainable parameters on 16,000 training data.

4.1 Preparing data

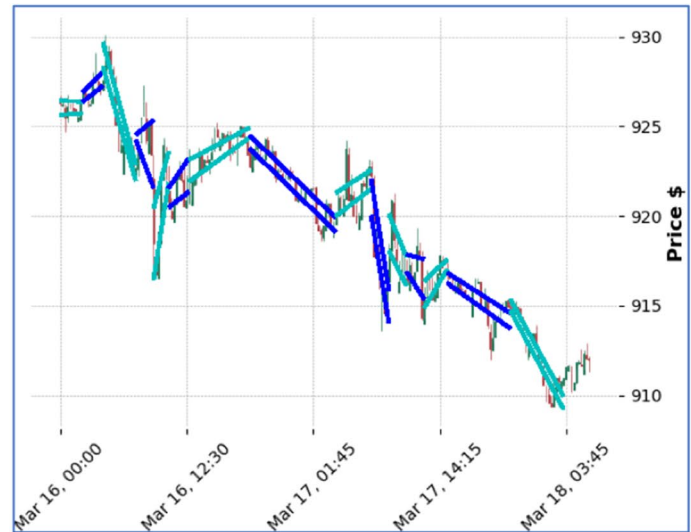
After dividing the data into wave series, we prepare the features of each wave for the model inputs. There are two types of features: price data and financial indicators.

- Price data features: The characteristics of the price data consist of the number of candles in a wave, the high and low of the last candle, the average of the high and low of a wave, and the slope of the upper and lower lines. There are 7 feature columns in total, with the last 2 features used as prediction targets for the time series regression.
- Financial features: In our experiment, we use the Technical Analysis Library [13] to compute all financial indicators. The indicators in the library are divided into four groups according to their objectives, namely volume, volatility, trend, and momentum. Of these four groups, we can only use three indicator groups because our raw data does not contain volume information. The nine popular indicators (16 features) in our work are:

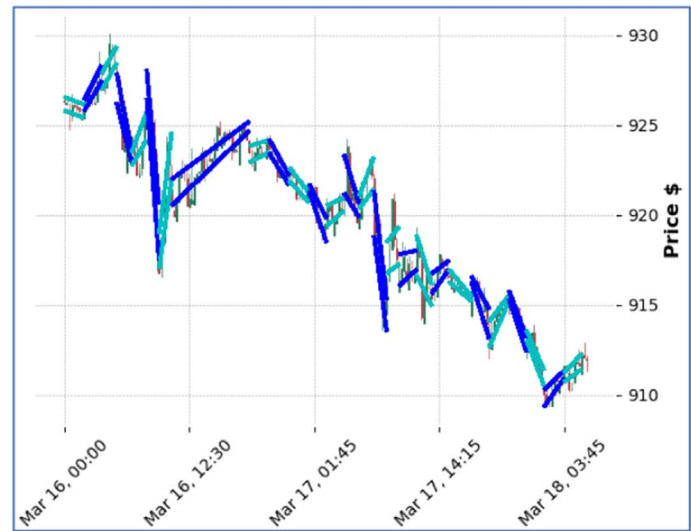
- 1) Moving Average Convergence Divergence (MACD) (trend indicator)
- 2) Exponential Moving Average (EMA) (trend indicator)
- 3) Simple Moving Average (SMA) (trend indicator)
- 4) Average Directional Movement Index (ADX) (trend indicator)
- 5) Relative Strength Index (RSI) (momentum indicator)
- 6) Stochastic Oscillator (momentum indicator)
- 7) Williams %R (momentum indicator)
- 8) Bollinger Bands (volatility indicator)
- 9) Keltner Channel (volatility indicator)

We represent each wave by two sets of data: the indicators of the last candle of the wave (16 features) and the average of the indicators of the wave (16 features). In summary, there are total 39 features: 7 features are from price data features and 32 features are from financial features.

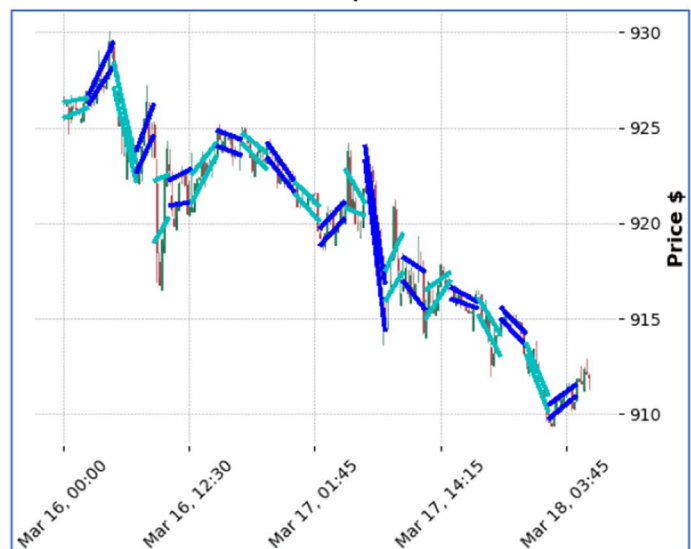
Fig. 3 Example results of this wave extraction algorithm using the different weighted parameters. In these examples, example (a) uses the highest cut-off value, so it promotes the continuity of the trend. On the other hand, examples (b) and (c) contain more waves due to the small cut-off value. The difference in line color is only for visualization purposes only



Example a.



Example b.



Example c.

Table 1 The amount of data after applying the wave extraction algorithm

Timeframe	Raw data	Wave data
5 min	872.498	39.929
15 min	292.367	20.225
30 min	146.353	12.846
60 min	73.699	8.061

4.2 Model architecture

The basic structure of our Transformer model comes from Att-DCNN, which was proposed by Daiya et al [5] and adopted the dilated convolutions and skip connection from the WaveNet architecture. Then the model was further modified to work with our input and prediction goals. The major changes were made in the input and the output of the model. The input of the model was adjusted to take 39 features and 24 sequence windows. The output layer was changed from softmax to linear for the regression of trend slopes. The modification was not only limited to the structure but also to hyperparameters and activation functions. Figure 4 shows a transformer block in this model. Transformer model is employed as the backbone of this architecture, which is structured around individual Transformer blocks. Each block incorporates a multi-head attention mechanism, followed by layer normalization. Additionally, within each block, there are five dilated convolutional layers with a kernel size of 3 and dilation rates of 1, 2, 2, 2, and 4.

A notable feature of this model lies in the integration of dilated convolutions within the Transformer structure. The use of dilated convolutions, each with a distinct dilation rate, allows the model to effectively capture information across different receptive fields. This enhances the model's ability to discern intricate patterns and dependencies within the input data. Following the Transformer encoder's processing, the results undergo global average pooling, serving as a mechanism to distil the encoded information. Subsequently, this pooled representation is fed into a Multi-Layer Perceptron (MLP) for the final output. These elements collectively contribute to the model's proficiency in capturing long-range dependencies and intricate patterns within the input data. The final integration with an MLP enhances the model's ability to generate meaningful predictions, making it well-suited for the complexities associated with financial market prediction tasks.

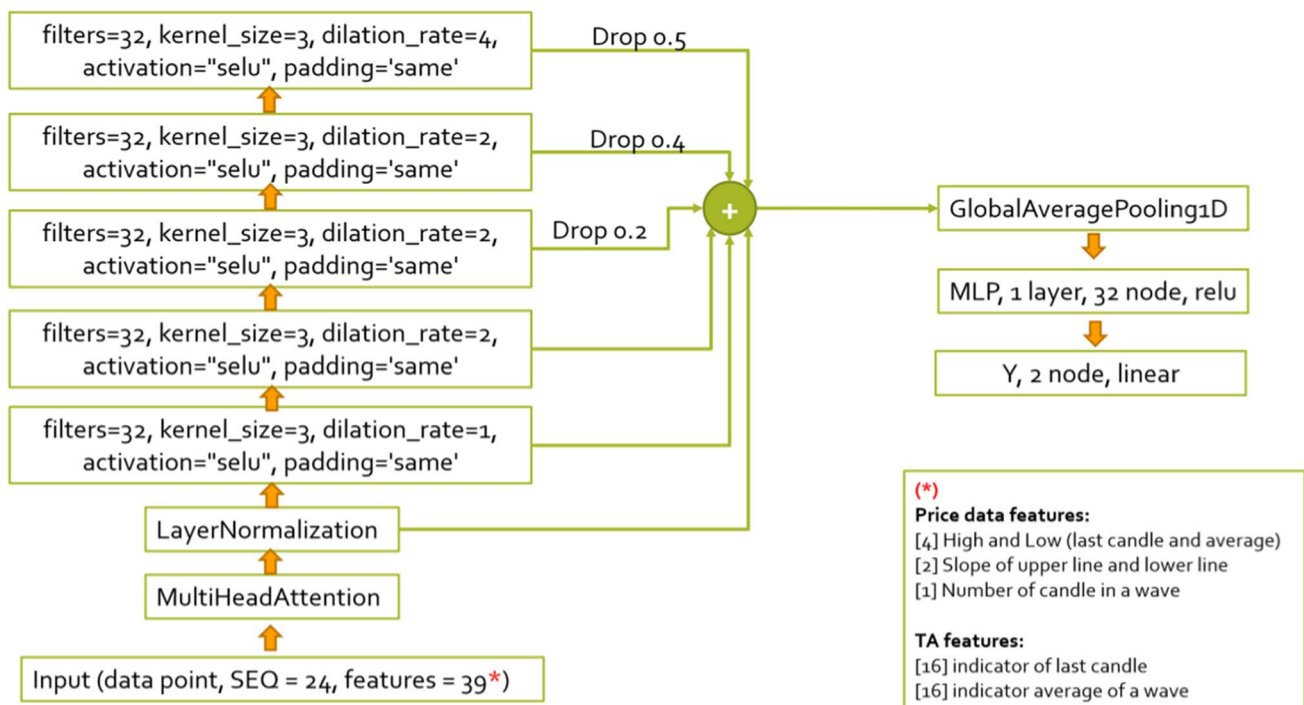


Fig. 4 Our transformer block structure

4.3 Model tuning and evaluation

The regression model was chosen in this paper to incorporate knowledge of the magnitude of market movements into trading strategy and money management, which could increase the overall profit rate, rather than just knowing the possibility of an upswing or downswing. However, unlike the straightforward interpretation offered by the accuracy score of a classification model, assessing the regression model's performance solely based on mean absolute error (MAE) presents challenges. To address this, we introduced three key metrics to comprehensively evaluate the model's effectiveness.

1) Up/down ratio

This metric quantifies the balance between uptrend and downtrend predictions by dividing the number of uptrend predictions by the number of downtrend predictions. A balanced model should ideally yield a ratio close to 1, reflecting an equal proficiency in predicting both uptrends and downtrends. If the model can predict an uptrend and a downtrend equally, we could say that the model has no bias in predicting highs and lows. The determination of trend predictions involves the slope of the upper and lower lines, as outlined in Table 2. The threshold number at 0.022 is the standard deviation of the slope of the train data set divided by two. This metric provides valuable insights into the model's bias or impartiality in predicting market highs and lows.

$$\text{up - down ratio} = \frac{\text{number of uptrend predictions}}{\text{number of downtrend predictions}} \quad (1)$$

2) Loss rate

The inclusion of the Loss Rate as a model evaluation metric is pivotal for understanding the downside risks associated with the model, aiding in the effective management of trading positions for each order. This metric is calculated by determining the percentage of instances where the model predicts an uptrend while the actual trend is down, and vice versa. By quantifying these instances, the Loss Rate provides crucial insights into the model's performance in adverse scenarios, contributing valuable information for risk management in trading positions.

$$\text{loss rate up trend}(\%) = \frac{\text{number of actual down trend}}{\text{number of predicted up trend}} * 100 \quad (2)$$

$$\text{loss rate down trend}(\%) = \frac{\text{number of actual up trend}}{\text{number of predicted down trend}} * 100 \quad (3)$$

$$\text{loss rate}(\%) = \frac{1}{2}(\text{loss rate up trend} + \text{loss rate down trend}) \quad (4)$$

3) Sideways (SW) precision

We put a lot of emphasis on sideways, given that a sideways trend represents a substantial portion of overall market trends and holds particular relevance in high-frequency trading. This metric gauges the model's precision in identifying sideways trends, which are prevalent and often exploited in various high-frequency trading strategies, including ping-pong trading and grid trading.

$$\text{sideways precision} = \frac{\text{number of correct SW prediction}}{\text{number of SW prediction}} \quad (5)$$

Table 2 The definition of trend prediction

Trend prediction	Upper line slope	Lower line slope
Uptrend	> 0.022	> 0.0
Downtrend	< 0.0	< - 0.022
Sideway	< = 0.022	> = - 0.022

Table 3 presents examples of model optimization along with corresponding evaluation matrices. The model underwent variations in crucial hyperparameters, including the number of multi-head attention, transformer blocks, and sequences. Notably, the model's performance exhibited sensitivity to hyperparameters, particularly the sequence window, which directly influenced the training data knowledge. Since we used training data during the 2009–2018 period which was the major uptrend for gold prices, the model tended to bias uptrend prediction when we used too short and too long sequence. During model optimization, we tried to increase the complexity of the model by increasing the number of features, blocks, heads, and filters. At the beginning of the increasing complexity, the performance was better, but eventually, it reached the limit, which happened to result in a bad score for one of the matrices. The highlighted model in green stands out as the best-performing model according to our evaluation metrics. This model adopts the transformer structure depicted in Fig. 4, featuring four stacked transformer blocks, each comprising eight multi-head attention units. Results from the test data set indicate that this selected model has the lowest prediction loss rate among the considered models, with a balanced ratio of upward and downward trend predictions at 1.36. Additionally, it achieves the second-best precision in sideways trend prediction, showcasing superior overall performance compared to other models.

See Fig. 5 for examples of prediction results. The top image exemplifies the model's ability to capture a robust downward trend, while the bottom image demonstrates accurate prediction of a sustained upward trend. Both instances showcase the model's proficiency not only in identifying trends but also in gauging their magnitudes, offering valuable insights for potential trading strategies. A straightforward simulation focusing solely on trend predictions over the test dataset indicates the model's capacity to generate profitable trades. With a fixed contract size of 1 per trade, the trading profit reaches 19.8% of the initial margin for 50,000 THB. This surpasses the performance of the MACD signal under identical assumptions, which achieves a profit of only 12.5%. The notable advantage of utilizing a series of trends as input and predicting the next trend output is its ability to capitalize on profit opportunities during sideways movements. In contrast, the traditional MACD signal proves profitable only when prices exhibit strong upward or downward trends.

The current performance limitations may stem from various factors, with one potential explanation being the ease of trend capture through the linear line. Despite this, it's noteworthy that the model's performance still surpasses that of traditional candlestick series. To assess the comparative effectiveness, we conducted an experiment comparing the wave series and candle series. Resource constraints led to the input limitation of candle data up to a window of 60 sequences.

Table 3 The examples of model tuning

Time-frame	Model	MAE	Up/Down ratio	Loss rate (%)	SW precision (%)
15	MACD RSI EMA BB indicators ff=4, head=4, block=4, SEQ=24, MPL=relu	0.5729	0.42	47.92	63.5
15	MACD RSI EMA BB indicators ff=64, head=4, block=4, SEQ=24, MPL=relu	0.0933	0.87	45.84	64.7
15	MACD RSI EMA BB indicators ff=32, head=4, block=4, SEQ=24, MPL=selu	0.0690	0.34	41.09	66.4
15	MACD RSI EMA BB indicators ff=32, head=4, block=4, SEQ=16, MPL=selu	0.0371	No down	48.28	63.2
15	Use MACD RSI EMA BB indicators ff=32, head=4, block=4, SEQ=32, MPL=selu	0.3660	2.77	47.46	63.5
15	Use all indicators (refer to Sect. 4.1) ff=32, head=4, block=4, SEQ=24, MPL=selu	0.0666	0.07	39.75	63.8
15	Use all indicators (refer to Sect. 4.1) ff=32, head=8, block=4, SEQ=24, MPL=selu	0.0484	1.36	40.86	64.7
15	Use all indicators (refer to Sect. 4.1) ff=40, head=8, block=8, SEQ=24, MPL=selu	0.0940	0.56	46.57	63.2
15	Use all indicators (refer to Sect. 4.1) ff=32, head=8, block=4, SEQ=24, MPL=relu	0.0416	0.64	41.35	64.3
5	Use all indicators (refer to Sect. 4.1) ff=32, head=8, block=4, SEQ=24, MPL=selu	0.0183	2.65	53.31	64.7
60	Use all indicators (refer to Sect. 4.1) ff=32, head=8, block=4, SEQ=24, MPL=selu	0.8283	0.75	45.15	49.5

The row with bold font represents our best model from the experiment

Fig. 5 Examples of prediction results. The red lines are the predicted tilt as a result of 24 waves entered and the superimposed blue lines are the actual tilt of that wave series

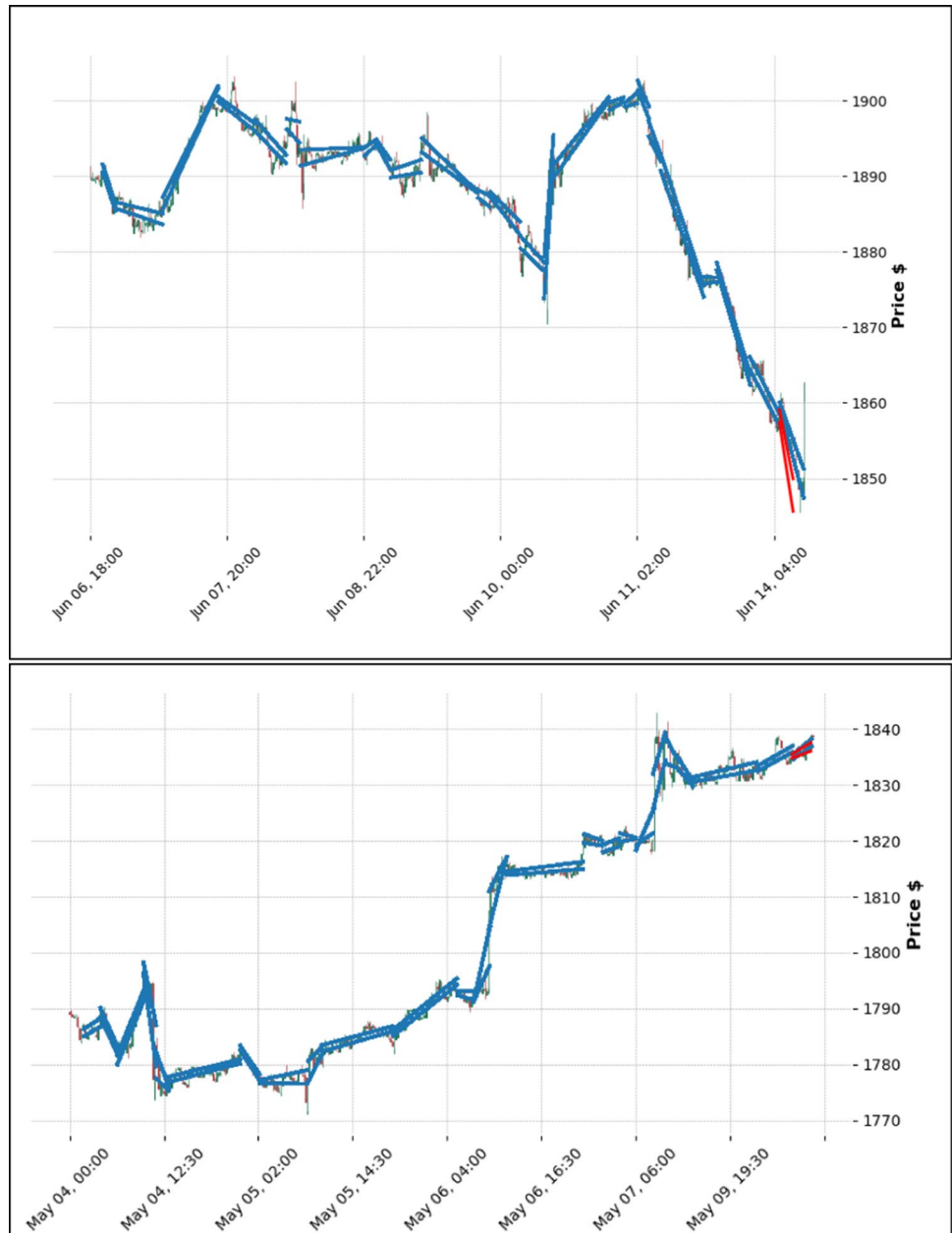


Table 4 The model performance comparing between waves input and candle input

Time-frame	Model	MAE	Up/Down ratio	Loss rate (%)	SW precision (%)
15	Wave series ff=32, head=8, block=4, SEQ=24, MPL=relu	0.0484	1.36	40.86	64.7
15	Traditional series ff=32, head=8, block=4, SEQ=60, MPL=relu	0.0265	8.68	42.91	54.5

Despite these limitations, Table 4 reveals that the wave series outperforms the candle series in three out of four evaluation metrics. This emphasizes the superiority of the wave series in capturing and predicting market trends, even within the constraints of input window limitations.

5 Discussion

The success of our model in predicting market movements through wave series data marks a significant stride in aiding traders to minimize losses and make informed decisions. Our model's superiority over traditional candlestick data, particularly in three out of four metrics, substantiates the efficacy of the chosen wave series approach.

The special strength lies in the utilization of wave series constructed with linear lines, capturing both trend and magnitude. The chosen linear lines method is both easy to understand and powerful, contributing to the model's effectiveness in recognizing complex market patterns. This method can capture the whole pattern despite the varied time window of pattern. However, the main challenge is resource demands, particularly the need for extensive input data due to an extended and dynamic time window. This limits its applicability to assets with substantial trading histories, and the observed significant data reduction underlines these resource challenges.

Moving forward, optimizations should explore methods for grouping wave data, experiment with data from multiple time frames, and consider additional financial indicators. Hyperparameter tuning, especially for the sequence window, requires careful consideration to balance model complexity and performance. Despite these challenges, the model holds practical implications for trading strategies and risk management, outperforming traditional candlestick series and demonstrating profitability in simulated trading scenarios.

6 Conclusion and future works

This study introduces a novel approach to predicting market movements by leveraging the wave series as opposed to traditional candlestick data. The current findings demonstrate the model's potential to assist traders in minimizing losses and making informed decisions regarding suitable trading strategies based on predicted market moves. However, a primary limitation lies in the model's high demand for input data due to the extended and dynamic input time window. The Transformer model can only be effectively trained on stocks with substantial historical price data, providing a requisite volume for training. For instance, our raw data from a 15 min time window spanning a decade was condensed to a mere 20,225 data points after transformation into waves. Due to time and resource constraints, several potential avenues for further model optimization in future works include exploring alternative methods for grouping wave data, incorporating data from multiple time frames within a single model, and integrating additional financial indicators. Addressing these aspects could enhance the model's efficiency and applicability across a broader spectrum of financial instruments, thereby advancing its potential for practical implementation in diverse market scenarios.

Author contribution All authors contributed to this article equally.

Funding The authors did not receive support from any organization for the submitted work.

Data availability The datasets generated and/or analyzed during the current study are from the XAUUSD (gold spot per U.S. dollar) public historical 1 min bar chart data from histdata.com (Apr 2009–Jul 2021).

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Altan A, Karasu S, Bekiros S. Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques. *Chaos, Solitons Fractals*. 2019;126:325–36.
2. Atsalakis GS, Dimitrakakis EM, Zopounidis CD. Elliott wave theory and neuro-fuzzy systems, in stock market prediction: the WASP system. *Expert Syst Appl*. 2011;38(8):9196–206.
3. Bureš M. *Machine learning in algorithmic trading*. Czechia: Charles University; 2021.
4. Cohen N, Balch T, Veloso M. (2020). Trading via image classification. In *Proceedings of the First ACM International Conference on AI in Finance* (pp. 1–6).
5. Daiya D, Wu MS, Lin C. (2020). Stock movement prediction that integrates heterogeneous data sources using dilated causal convolution networks with attention. In *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 8359–8363). IEEE.
6. Daiya D, Lin C. (2021). Stock Movement Prediction and Portfolio Management via Multimodal Learning with Transformer. In *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 3305–3309). IEEE.
7. Frost AJ, Prechter RR, Collins CJ. *Elliott wave principle: key to market behavior*. Wiley; 1999.
8. Karasu S, Altan A. Crude oil time series prediction model based on LSTM network with chaotic Henry gas solubility optimization. *Energy*. 2022;242:122964.
9. Karasu S, Altan A, Bekiros S, Ahmad W. A new forecasting model with wrapper-based feature selection approach using multi-objective optimization technique for chaotic crude oil time series. *Energy*. 2020;212:118750.
10. Kim T, Kim HY. Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PLoS ONE*. 2019;14(2):e0212320.
11. Leigh W, Frohlich CJ, Hornik S, Purvis RL, Roberts TL. Trading with a stock chart heuristic. *IEEE Trans Syst Man Cyber-Part A: Syst Humans*. 2007;38(1):93–104.
12. Nazário RTF, Silva JLe, Sobreiro VA, Kimura H. A literature review of technical analysis on stock markets. *Q Rev Econ Finance*. 2017;66:115–26.
13. Padiál, D. L. (2018). *Technical Analysis Library in Python*. Retrieved from <https://github.com/bukosabino/ta>.
14. Savin G, Weller P, Zvingelis J. The predictive power of “head-and-shoulders” price patterns in the US stock market. *J Financ Economet*. 2007;5(2):243–65.
15. Sezer OB, Gudelek MU, Ozbayoglu AM. Financial time series forecasting with deep learning: a systematic literature review: 2005–2019. *Appl Soft Comput*. 2020;90:106181.
16. Sim KH, Sim KY, Bong N. (2018). Dynamic Time Interval Data Representation in Scalable Financial Time Series Pattern Recognition. In *Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence* (pp. 120–125).
17. Tsinaslanidis P, Guijarro F. What makes trading strategies based on chart pattern recognition profitable? *Expert Syst*. 2021;38(5):e12596.
18. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. *Adv Neural Inform Process Syst*. 2017;30:5998–6008.
19. Wu KP, Wu YP, Lee HM. (2014). Stock Trend Prediction by Using A-Means and AprioriAll Algorithm for Sequential Chart Pattern Mining. *Journal of Information Science & Engineering*, 30(3).
20. Zhang, S., and Zhang, H. (2020). Prediction of Stock Closing Prices Based on Attention Mechanism. In *2020 16th Dahe Fortune China Forum and Chinese High-educational Management Annual Academic Conference (DFHMC)* (pp. 244–248). IEEE.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.