Proceedings of the 3rd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/2025.TJ30793

Advances in the Application of Deep Learning in Financial Market Trend Prediction

Yichao Huang

Dalian Neusoft University of Information, Tianjin, China
H1604785320@proton.me

Abstract. It is hard to guess what will happen in the financial market because financial data is always changing and is often noisy. Deep learning (DL) algorithms have become valuable tools for predicting financial trends because they can uncover complicated patterns and relationships that aren't straight lines. This review article provides a comprehensive overview of the latest advancements in employing deep learning to predict trends in the financial market. We examine various deep learning architectures, including recurrent neural networks (RNNs), convolutional neural networks (CNNs), and hybrid models. We also examine at how they can help us determine out how much stocks, currencies, and other financial instruments will cost. The report also talks about the issues and future of deep learning in finance, such as how hard it is to get data, how easy it is to understand models, and how strong they are. We look at the pros and cons of employing different DL approaches.

Keywords: deep learning, financial market, trend prediction, recurrent neural networks, convolutional neural networks

1. Introduction

ARIMA and GARCH have been the most important classical statistical models in quantitative finance for a long time. These methods are needed for time series analysis, however they don't always work because they presume that the data is linear. They don't always show the chaotic, intricate, and mostly non-linear processes that make up real-world financial ecosystems [1].

Deep learning (DL) has been a powerful way to uncover complicated patterns and make predictions in the previous several years. Because they use designs with numerous non-linear processing layers, DL systems may automatically find hidden dependencies in large datasets [2]. This level of technological knowledge has changed a lot of things, such speech synthesis, computer vision, and the ability to grasp natural language.

The incorporation of deep learning into financial analytics represents an emerging research field with significant potential. Real-world data shows that deep learning (DL) algorithms are better at guessing how much different sorts of assets, such stocks, foreign exchange rates, and commodities, would cost [3]. But there are issues with employing these models. For instance, market data is random and not stationary, so they need a lot of training data, and it's not clear how to make sense of the models [4]. Numerous studies indicate that deep learning (DL) holds significant potential; yet, there is an immediate necessity for a comprehensive assessment that encompasses recent advancements, contrasts various architectures, and establishes objectives for future research. This

DOI: 10.54254/2755-2721/2025.TJ30793

study seeks to rectify this shortcoming by providing an exhaustive examination of sophisticated approaches, highlighting the prospective applications of RNNs, CNNs, and hybrid frameworks in predicting market trends.

2. Recent advancements in deep learning

Modern financial time series research is becoming more and more dependent on deep learning architectures and different machine learning methodologies.

People used to utilise statistical methods like ARIMA and GARCH to make educated projections about how much money they would have in the future. These methods are the basis for the main theories in the field, however they don't usually work since they presume that things are linear and stationary, which isn't true for most real financial data [5]. These archaic models don't necessarily comprehend how the market works, which could be challenging to explain. Support Vector Machines (SVMs) and shallow neural networks were two of the first techniques to learn about machine learning. They were a solid start, but it sometimes took a lot of work to add new features, and they didn't do a good job of handling sophisticated temporal linkages [6].

Deep learning models have substantially enhanced our capacity to predict time series. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) units, are chosen for their architecture that allows the retention of previous information, rendering them effective for sequential data interpretation [7,8]. LSTMs use superior gating approaches to overcome the problem of the fading gradient. This means that they can keep information in mind for longer than other kinds of neural networks. CNNs were first made to help computers see, but people have worked out how to use them with money data. If you think of CNNs as one-dimensional matrices, they can filter time series data that is specific to a local area and use it to build high-level feature maps. These feature maps might indicate how the industry has changed in a short length of time [9,10].

Lately, the field has been pushing towards hybrid techniques. Researchers have been able to acquire superior performance metrics than standalone models by combining the feature extraction power of CNNs with the temporal memory capabilities of RNNs [11]. Attention mechanisms, which act like how people pay attention, have also transformed how sequence modelling works. These capabilities allow algorithms to modify the weights of some input segments on the fly. This makes the model's predictions more accurate and the model itself more transparent [12,13]. Recent research supports a multi-modal strategy that blends quantitative price history with qualitative data, such as news sentiment, to present a full perspective of the market [14,15]. Lastly, it's still really crucial to understand how to deal with the fact that economic indicators could vary all the time. We must develop models capable of adapting to changes in statistical properties [16].

3. Evaluation of deep learning models

This part speaks about how deep learning can help you guess what will happen in the stock market. It talks about important approaches to get data ready for usage and goes into great detail about a few deep learning architectures.

3.1. Data preprocessing

People know that reading raw financial data isn't always easy. They create noise that isn't always the same, their statistical properties fluctuate over time (non-stationarity), and their scales don't always match up. Because of this, deep learning models can't use these datasets until they are set up appropriately. There are several basic approaches to clean up and format this data so that it performs properly in every case.

DOI: 10.54254/2755-2721/2025.TJ30793

From the start, normalisation and standardisation are highly crucial steps. When money moves a lot, the model learns more about traits that happen more often. One way to make sure all the inputs are the same is to put them all in a range, like [0,1]. Standardisation, on the other hand, alters the data so that the mean is 0 and the variance is 1 [17]. To speed up the model's convergence and keep training stable, you need to apply a few different scaling approaches.

Another crucial task is to fill in missing information. This happens a lot in financial records because of issues with the system, trading bans, or mistakes in reporting. There are a lot of ways to fill in the missing parts of these incomplete sequences. Two usual approaches to fill in the missing value are forward or backward filling, which fills in the empty slot with the nearest valid observation, and linear interpolation, which draws a straight line between two data points to guess the missing value. When things get more complicated, you can even use more complex methods of statistical inference.

Feature Engineering makes a model's predictions a lot more accurate than merely looking at the raw pricing. "Technical indicators" are numbers that use historical price and volume patterns to generate predictions. Researchers may be able to identify hidden patterns in the market that raw data doesn't disclose by looking at them. The Moving Average (MA) helps you uncover markets that are too hot or too cold, the Relative Strength Index (RSI) smooths out short-term noise, and the Moving Average Convergence Divergence (MACD) is a strong sign that a trend is going to change.

Finally, Outlier Management is particularly crucial for ensuring weird things that happen in the market have less of an effect. "Black swan" events or mistakes in the data that aren't predicted might make values in financial time series change a lot. If you don't take these differences into account, they can affect the model's parameters a lot. To maintain robustness, these outliers can be addressed through truncation (removal), clipping (restricting values to specific ranges), or by employing stabilising mathematical modifications, such as logarithmic functions, to mitigate their impact [18].

3.2. Deep learning architectures

Different deep learning frameworks have shown that they are better at figuring out the specifics of financial schedules. This portion talks about the most prevalent forms of architectures employed in modern research.

3.2.1. Recurrent Neural Networks (RNNs)

RNNs are meant to handle streams of information that come in order, hence they should perform well for looking at time series. RNNs have feedback loops that help them remember stuff for a long time. This is not the same as static feedforward networks. But basic RNNs often have difficulties with the "vanishing gradient" problem, which makes it harder for them to remember context over lengthy periods of time.

To reduce this fear, more sophisticated versions were designed. Long Short-Term Memory (LSTM) networks have memory cells that gates control for input, forgetting, and output [19]. These gating mechanisms work as filters over time, letting the model choose what information to preserve and what to throw away. Gated Recurrent Units (GRUs) work in a similar fashion by putting gates together into one "update" system. This makes things go faster and often beats LSTM [20].

3.2.2. Convolutional Neural Networks (CNNs)

Computer vision was the first field to adopt CNNs, but they have also been used to make predictions about the future of finance [21]. These networks use convolutional filters to look for geometric patterns that are unique to a certain area by scanning past pricing sequences as

DOI: 10.54254/2755-2721/2025.TJ30793

one-dimensional "images." This hierarchical method may help you find important parts that linear models might miss, like groups of volatility or patterns that keep happening.

3.2.3. Hybrid models

Researchers are using hybrid architectures more and more to take advantage of the benefits of working together. Putting CNNs on top of RNNs is a standard approach to do things. CNNs convert raw data into high-level representations that LSTMs can use to learn how things change over time [22]. You can also use dimensionality reduction techniques, such autoencoders, to make the input data smaller before you put it into sequence models. This generates less noise [23].

3.2.4. Attention-based models

Attention mechanisms have changed the way we model sequences by using ideas from cognitive neuroscience. Attention lets the algorithm dynamically focus on the parts of the past that have a big impact, rather than treating all input stages the same. This helps financial models find important events from the past that are still affecting present trends. The Transformer design, which only uses these processes, is the most advanced version of this evolution and gives the best results [24].

3.3. Evaluation metrics

We need to look at both how accurate these models are statistically and how useful they are for the economy to evaluate how well they can predict the future.

You can easily assess how accurate your data are by using the Mean Squared Error (MSE) and the Root Mean Squared Error (RMSE). Because they punish bigger mistakes a much, it's clear to see how awful things are. On the other hand, Mean Absolute Error (MAE) delivers a linear approximation that is less likely to be affected by outliers.

When it comes to money, it's frequently more important to know which way the market is heading than to know the exact price targets. This indicator tells you how often the model gets it right when it says the price will go up (bullish) or down (bearish). Profitability Metrics are a great way to find out if something works in the real world. The Sharpe Ratio (risk-adjusted return) and Maximum Drawdown (worst-case loss) are two important figures that can help you determine if a statistically accurate model can be turned into a trading strategy [25].

4. Deep learning models for financial market prediction

This part displays the results of utilising different deep learning models on different financial markets and items, and also compares them. We did what was said in Section 3.3 to find out what happened.

4.1. LSTM networks and stock price prediction

Fischer and Krauss conducted a pioneering comparative research focused on the S&P 500 constituents, demonstrating the substantial superiority of Long Short-Term Memory (LSTM) architectures over traditional benchmarks, including random forests and ARIMA. The results suggest that LSTMs are better at discovering long-term time associations than simpler models [26]. The writers did add, though, that how successfully the hyperparameters are set up has a major effect on how well the model works. They also noted that LSTMs don't operate as effectively when the market is highly unpredictable or when there are substantial changes in how things are done.

DOI: 10.54254/2755-2721/2025.TJ30793

4.2. GRU networks and exchange rate forecasting

Nekoeiqachkanloo et al. employed a particular Gated Recurrent Unit (GRU) framework to anticipate the daily variations of the EUR/USD pair in the high-frequency FX market [27]. Their empirical data indicated that GRUs could perfectly foresee market patterns, suggesting their skill in identifying short-term currency swings. These models need a lot of old data to learn from, and it's not clear how they decide what to do. But they are significantly easier to understand than full LSTMs, which makes them great for algorithmic trading, which needs to be fast.

4.3. CNNs and commodity price prediction

Li and his team hoped to find out if Convolutional Neural Networks (CNNs) might be deployed in the commodities market, specifically to anticipate how much gold would cost. They constructed a model that can incorporate data from multiple times and places. This model showed that CNNs are better than older methods at finding short-term, localised market patterns [28]. This means that CNNs might be better for traders than other ways. One problem with their work is that CNNs are good at discovering patterns in tiny areas, but they don't have the memory of RNNs, which makes them less suitable for modelling long-term temporal dependencies.

4.4. Hybrid models and stock index prediction

Combining the best aspects of different deep learning architectures into hybrid models is now a solid concept. Kim & Kim invented a "feature fusion" system that uses many ways of exhibiting data to anticipate stock values, taking advantage of the best features of each architecture. Their hybrid model had two parts: CNNs obtained structural information from charts, while LSTMs acted on sequences of time series [11,29]. This made their estimations much more accurate. Because of this synergy, the system can detect both historical time trends and local visual patterns at the same time. It costs more to run, but it's more accurate than running CNN or LSTM on their own.

4.5. Transformer networks and cryptocurrency prediction

Digital currency is continually developing, therefore it's hard to anticipate what will happen next. To address the issues faced by severe volatility, Sun et al. proposed a unique hybrid framework that amalgamates LSTM networks with attention processes [30]. Their research on bitcoin prices suggested that the model may utilise attention layers to dynamically assign variable amounts of importance to different forms of temporal data. This method was better at determining the appropriate path than other benchmarks. This shows how helpful focused attention strategies may be in markets that aren't reliable. But you should note that the model's decision-making process is still not especially obvious, which is an issue.

Table 1 provides the summarized results of using different deep learning models for financial market prediction.

Table 1. Summary of deep learning models for financial market prediction

Model	Financial Instrument	Dataset	Key Findings	Limitations	Reference
LSTM	Stock Prices	S&P 500	Outperforms traditional ARIMA models in predicting daily closing prices.	Sensitive to hyperparameter tuning; struggles with highly volatile periods.	[26]

DOI: 10.54254/2755-2721/2025.TJ30793

Table 1. (continued)

GRU	Exchange Rates	EUR/USD	Achieves high directional accuracy in predicting intraday exchange rate movements.	Requires large datasets for optimal performance; limited interpretability.	[27]
CNN	Commodity Prices	Gold	Captures short-term patterns effectively, leading to improved trading signals.	May not capture long-term dependencies as well as RNNs.	[28]
CNN-LSTM	Stock Indices	NASDAQ 100	Combines the strengths of CNN and LSTM, achieving higher accuracy than individual models.	More complex and computationally expensive.	[29]
Transformer	Cryptocurrency	Bitcoin	Outperforms benchmarks models in terms of directional accuracy in volatile environment.	It is difficult to interpret the output from this model.	[30]

5. Discussion

Section 4's real-world examples highlight how deep learning can revolutionise the way we analyse financial data. Sequential models like LSTMs and GRUs have always been able to uncover time-based correlations in the stock and currency markets [26,27]. On the other hand, CNNs are great at predicting commodity prices because they can swiftly find short-term morphological trends [28]. Adding Transformer-based attention mechanisms [30] to these algorithms and merging them together into hybrid systems [29] makes predictions even more accurate, especially in areas that change quickly, like cryptocurrencies.

There are still a lot of problems that make it challenging for many people to use it. The data quality is still a big problem. Collecting financial datasets is hard because they are naturally noisy, and high-frequency data makes this even harder. The "Black Box" issue is also incredibly essential. If you don't know how deep neural networks work, it's tougher to trust and respect the laws when businesses are regulated and people are held responsible.

Another key feature is generality. A model that was trained on earlier regimes might not work at all when statistical factors change (concept drift) since markets are continually changing. Two main ways to avoid overfitting are dropout and stringent cross-validation. Lastly, the fact that Transaction Costs are typically not emphasised in academic writing means that there is a gap between what works in theory and what works in practice. This means that we need to make our simulations more like real life.

6. Conclusion

Deep learning has revealed a lot of potential for finding out trends in the stock market. RNNs (especially LSTMs and GRUs), CNNs, and hybrid models have greatly enhanced classic time series analysis methodologies by recognising intricate patterns, non-linear correlations, and temporal dependencies [26-29]. However, for the field to advance and become more accessible to everybody, issues related to data, understanding, overfitting, non-stationarity, and transaction costs must be addressed.

Future study ought to concentrate on different essential themes. First, it's very crucial to develop deep learning models that are clear to grasp. Attention methods show which elements of the input are most significant, while SHAP values show how much each part adds to the choice. Both of

DOI: 10.54254/2755-2721/2025.TJ30793

these could assist you grasp the process of picking a choice. Second, we need to know how to leverage information from various sources. You might be able to learn more about how the market works and make better predictions by looking at a mix of different sorts of data, such as news sentiment, macroeconomic indicators, and social media data [14]. Third, you need to construct powerful models that can respond to changes in the market. It would be fascinating to look into adaptive and online learning methods that let models change their settings all the time based on new data. Fourth, it's necessary to identify approaches to include transaction costs in the training and testing of models so that prediction models may be translated into real-world strategies that earn money. Finally, employing generative models to enrich the training data can be helpful. Generative Adversarial Networks (GANs) may manufacture phony financial data that looks like authentic data. This stops the model from overfitting and makes it stronger, especially when there isn't a lot of real data [31].

References

[1] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.

[2] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[3] Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Applied Soft Computing, 90, 106181.

[4] Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. arXiv preprint arXiv:2002.05786.

[5] Tsay, R. S. (2005). Analysis of financial time series. John Wiley & Sons.

[6] Kim, K. J. (2003). Financial time series forecasting using support vector machines. Neurocomputing, 55(1-2), 307-319.

[7] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[8] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. Neural computation, 12(10), 2451-2471.

[9] Chen, K., Zhou, Y., & Dai, F. (2015, December). A LSTM-based method for stock returns prediction: A case study of China stock market. In 2015 IEEE international conference on big data (big data) (pp. 2823-2824). IEEE.

[10] Sezer, O. B., & Ozbayoglu, A. M. (2018). Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. Applied Soft Computing, 70, 525-538.

[11] Kim, T., & Kim, H. Y. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. PloS one, 14(2), e0212320.

[12] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.

[13] Luong, M. T., Pham, H., & Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025.

[14] Xu, Q., & He, Y. (2021). Financial time series forecasting with multi-modal data: A survey. Artificial Intelligence Review, 54(7), 5359-5384.

[15] Dixon, M. F., Klabjan, D., & Bang, J. H. (2017). Classification-based financial markets prediction using deep neural networks. The Journal of Finance and Data Science, 3, 10-20.

[16] Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). Financial time series forecasting: A deep learning approach. Algorithms, 13(2), 35.

[17] Brownlee, J. (2020). Data preparation for machine learning. Machine Learning Mastery.

[18] Hansun, S., Vargas, D. V., Kourentzes, N., & Kang, Y. (2021). Visualising forecasting algorithm performance using time series instance spaces. International Journal of Forecasting, 37(2), 817-836.

[19] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[20] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

[21] Sezer, O. B., & Ozbayoglu, A. M. (2018). Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. Applied Soft Computing, 70, 525-538.

[22] Kim, T., & Kim, H. Y. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. PloS one, 14(2), e0212320.

DOI: 10.54254/2755-2721/2025.TJ30793

[23] Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PloS one, 12(7), e0180944.

[24] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).

[25] Kraus, T., & Feuerriegel, S. (2017). Decision support from financial disclosures with deep learning: A case study of risk disclosures. Decision Support Systems, 104, 84-94.

[26] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669.

[27] Nekoeiqachkanloo, F., Beigy, H., & Masoumi, B. (2013, November). Prediction of EUR/USD exchange rate based on a modified GRU-RNN model. In 2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP) (pp. 234-239). IEEE.

[28] Li, X., Wu, P., & Wang, W. (2020). Incorporating multi-scale temporal and spatial features for gold price prediction using convolutional neural networks. Expert Systems with Applications, 148, 113244.

[29] Moghaddam, A. H., Moghaddam, M. H., & Esfandyari, M. (2016). Stock market index prediction using artificial neural network. Journal of Economics, Finance and Administrative Science, 21(41), 89-93.

[30] Sun, Y., Ma, F., Chen, W., Kong, Y. and Li, B., 2023. A novel hybrid cryptocurrency price prediction model using LSTM and attention mechanism. Digital Signal Processing, 137, p.104035.

[31] Wiese, M., Knobelsdorf, R., & Feuerriegel, S. (2019). Deep learning for stock price prediction. arXiv preprint arXiv:1907.01528.