ETF Forecast: Application of Innovative Machine Learning Models in the Field of ETF Forecasting

Tong Zhang[a]
Department of Business, University of New South Wales, Kensington, Sydney, Australia

Keywords: ETF Prediction, Machine Learning Models, Innovative Algorithm, Financial Technology.

Abstract: With global exchange-traded fund (ETF) assets exceeding $15 trillion, ETFs are becoming increasingly significant in the global financial system. Specifically, because of the unpredictability of the outside world in recent years, investors have been more interested in low-cost, low-risk, and highly transparent investments, which is reflected in the increasing scale of passive investment ETFs and the increase in research in the field of ETF prediction. This study examines how machine learning models are now being used in the field of ETF prediction and chooses innovative machine learning models from the previous two years to review. Including the superposition of common models, the combination of traditional financial models and deep learning models, etc., the results show that these innovative combinations can significantly improve the effectiveness of ETF predictions. Researchers who wish to develop innovative machine learning models for ETF prediction will benefit from this study's understanding of current findings and possible research directions.

1 INTRODUCTION

On January 23, 1993, Standard & Poor's Depositary Receipt (SPDR) launched the first ETF in the United States, named after its underlying index. SPDR S&P 500 is the largest open-end index fund in the world (Liebi, 2020). Like mutual funds and index portfolios, exchange-traded funds (ETFs) were created by assembling a group of stocks. Existing ETFs are a further extension of index funds, with advantages including ease of trading, tax benefits, and significant cost-effectiveness (Joshi & Dash, 2024). From the late 1990s to the early 21st century, ETFs began to expand globally. Between 2003 and 2023, the total assets under ETF control increased from 204.3 billion to 11,507 billion USD (Joshi & Dash, 2024; Statista Research Department, 2024).

Since the first ETF appeared on the market in 1993, the ETF market has developed rapidly. In 2003, there were only 276 ETFs in the world. As of 2023, this number has grown to 10,319 (Statista Research Department, 2023). The rapid development of ETFs in recent years demonstrates the further intensification of global investors' interest in ETF investment, a trend that began after the 2008 financial crisis (EPFR, 2021). ETFs are popular among investors for their high liquidity, low fees and lower volatility compared to stocks (Liebi, 2020; Joshi & Dash, 2024). A 2022 survey of 60 executives worldwide by PricewaterhouseCoopers (PwC) pointed out that the strong resilience and growth potential of ETFs in resisting risks have been further highlighted during COVID-19 (PricewaterhouseCoopers, 2021). With the contribution of unprecedented capital inflows, a substantial number of new entrants and diverse product innovations and distribution opportunities, the activity, diversity and innovation of the ETF market are constantly increasing, and it has become an important part of the global asset and wealth management field.

Today, nearly 50% of investment in the United States comes from ETFs, and the size of the ETF market in the United States accounts for about 70% of the overall ETF market size, leading the world (Joshi & Dash, 2024). In particular, in August 2019, passive investment exceeded active investment in the US ETF market for the first time (Joshi & Dash, 2024).

This further shows that investors are increasingly pursuing stable returns, which has promoted the

[a] https://orcid.org/0009-0008-2667-0287

Zhang, T.
ETF Forecast: Application of Innovative Machine Learning Models in the Field of ETF Forecasting.
DOI: 10.5220/0013700100004670

In *Proceedings of the 2nd International Conference on Data Science and Engineering (ICDSE 2025)*, pages 497-501
ISBN: 978-989-758-765-8

further in-depth application of machine learning algorithms ETF performance forecasting.

Stock price forecasting is the central challenge at the nexus of computer science and finance (Yüksel, 2023). The well-established efficient market theory suggests that financial markets efficiently process information, ensuring that key data is swiftly and accurately incorporated into stock valuations. Consequently, investors struggle to achieve consistent returns above the market average without engaging in manipulative practices. While this perspective may seem pessimistic, it underscores the inherent difficulty of predicting market movements. Furthermore, the financial sector plays a vital role in economic growth, attracting extensive research and innovation from scholars worldwide.

2 EXISTING LITERATURE

ETF, as an investment tool that has flourished in the past few years, has attracted the attention of an extensive amount of researchers and investors. At present, most of the research in this field is to verify the model validity of a single machine learning model for a single or a few ETF markets.

2.1 Previous Literature

To validate the LSTM model's reliability in forecasting the return direction and particular ETF prices, Horst employed ETFs from five distinct marketplaces (Horst, 2022).

The study found that the Root Mean Squared Error (RMSE) values of different industries vary greatly, so the accuracy of the model is strongly correlated with the choice of industry, with the real estate industry performing the best and the energy industry performing the worst. Adding or reducing the number of ETFs will also greatly affect the prediction results.

Gowani and Kanjiani also chose the LSTM model and tested it in 9 different industries and more than 2,200 Vanguard industry ETFs (Gowani & Kanjiani, 2024). The difference is that the researchers chose R-squared value to evaluate the effectiveness of the model. The results became more optimistic, with R-squared values exceeding 0.68 in all industries and reaching 0.9095 in the energy industry, showing its strong predictive ability.

Astuy et al. selected Extreme Gradient Boosting (XGB), Neural Prophet model, Extra Trees (ET), and K-nearest Neighbors (KNN) for three ETFs (Sagarzazu Astuy, 2022). The results show that the prediction effects of different models vary greatly. When making investment decisions, it is advisable to employ multiple models and incorporate diverse predictive outcomes to enhance accuracy and robustness.

Between January 1, 2006, and October 31, 2023, Wang & Zhang chose ten distinct ETFs from the S&P 500 (Wang & Zhang, 2024). The model selection includes Decision Tree (DT), Gaussian Naive Bayes (GNB) and Neural Networks (NN).

The concentration of this research is on forecasting the performance of several ETFs in the future. Based on the forecasts, trading strategies are then developed, and the actual returns are monitored. Below market performance, that is, the short-selling trading strategy can obtain returns that exceed the return of holding the S&P 500 alone.

Through the investigation of existing research, it is found that machine learning models are widely used in this field, whether predicting the direction of returns, specific prices or market performance, but most of them use original models, and different models have different performances in different markets. Therefore, this study pays special attention to the innovation and universality testing of the model, hoping to improve the effectiveness of prediction.

2.2 Innovative Algorithm Model

Chang et al. took the “Yuanta/P-shares Taiwan Top 50 ETF (Exchange Traded Fund)” (ETF50) as the research object(Chang et al., 2024). ETF 50 is an open-end exchange-traded index fund in the Taiwan market, which aims to provide investors with investment returns similar to the performance of the stocks of the top 50 companies in Taiwan by market value. The trading code is 0050. ETF50's holdings cover multiple industries and fields such as semiconductors, financial insurance, and electronic components. Among them, TSMC, as the leading company in global semiconductor manufacturing, plays a pillar role in Taiwan's economy.

The typical stock prediction model analysis method relies heavily on past data and assumes that future market patterns will stay within a known range(Chang et al., 2024). However, in recent years, this initial assumption has been broken by unforeseen disruptive events, including the public health crisis around 2020, the economic dispute between China and America in 2018, and the American financial turmoil in 2011.

These events have disrupted and significantly shattered production activities and affected the regular operation of numerous industries, underlining the gravity of the situation. Therefore, the researchers adopted a new calibration process "Short-Term Bias Compensation" (STBC), aiming to further calibrate LSTM model predictions. This method could minimize the volatility influence impact of sudden or extreme events on forecasting accuracy. Additionally, fuzzy rules evolved via genetic algorithms (GA) are employed for optimizing trading strategies.

The dataset covered ETF trading data from 2003 to 2020, and black swan events are deliberately included to assess model precision. By calculating daily forecast errors and comparing the daily anticipated amounts with the actual data, the prediction's short-term bias (STB) was first estimated. The anticipated price for the following day will be automatically adjusted if the STB rises above a predetermined level. The STBC parameters are further optimized by using genetic algorithms (GAs). Simultaneously, trading strategies and stock buying and selling signals were determined by genetic fuzzy systems (GFS). According to the findings, STBC can cut the prediction error by over 90%.

Shih et al. focused on exploring how to combine traditional financial models (such as the Fama-French three-factor model, the capital asset pricing model (CAPM)) and artificial neural network architectures applicable to handling time series, including LSTM, ANN, GRU, CNN and their variants, to obtain better results in ETF daily return prediction (Shih et al., 2024). The multi-factor market model and the Taiwan Economic Journal (TEJ) provided the data for this Python-based analysis. Daily returns of six ETFs listed in Taiwan: Yuanta Taiwan 50 (0050), Yuanta Mid-Cap 100 (0051), Yuanta Electronics (0053), Yuanta S&P Custom China 50 (0054), Yuanta MSCI Taiwan Financial (0055), and Yuanta Taiwan Dividend Plus (0056) were chosen as the dependent variables. The time range spans from 2010 to 2020.

The study is divided into three parts. First, the Fama-French three-factor model and the deep learning algorithm's daily return prediction effects are compared. The contrast of the linear traditional model and the nonlinear artificial neural network is one of the distinctive study features.

The findings demonstrated that the nonlinear ANN combined with the CAPM, the Fama-French three-factor model and the Fama-French five-factor model yielded significantly stronger predictive power than regression-based approaches. First and second place went to the ANN combination with the Fama-French three-factor and Fama-French five-factor models, respectively. Other commonly used artificial neural networks, LSTM and GRUs, were further studied. The prediction effects of these two models were compared with the CAPM, the Fama-French three-factor model and the Fama-French five-factor model. The study demonstrated that the Fama-French three-factor model in association alongside any artificial neural network performed better than the Fama-French five-factor model and the CAPM in conjunction with other models. Out of all the combinations that contained the Fama-French three-factor model, the combination of LSTM and it yielded the lowest mean absolute error (MAE) value.

Furthermore, the Fama-French three-factor model's MAE values were all lower than ANN's when combined with LSTM and GRU. Moreover, the study began to add other variables to explore better prediction results. The added factors include Momentum Factor, Investment Factor, Profitability Factor, Dividend Yield Factor, Long-Term Reversal, and Short-Term Reversal. The results showed that the combination of the Fama-French three-factor model performs well regardless of whether the Short-Term Reversal factor is added. The researchers also added CNN but found that the prediction error increased after adding CNN except in 2016 and 2019. Therefore, the model using hybrid or stacked networks has not effectively improved the ability to explain daily prices. The study pointed out that the combination of the Fama-French three-factor model and LSTM was the most effective way to predict daily returns, providing excellent prediction accuracy.

Using the historical return data of its constituent equities, this study (Piovezan, de Andrade Junior, & Ávila, 2024) suggested a machine learning-based prediction strategy for the direction of ETF returns. The study used information from five reference markets to compare the outcomes with buy-and-hold and naive prediction strategies. Regression models (linear regression, ridge regression, extreme gradient boosting (XGBoost), lightweight gradient boosting (LightGBM), and classification models (logistic regression, support vector machine (SVM), naive Bayes (GaussianNB), K-nearest neighbour (KNN), and random forest) were among the machine learning models that were employed. Each model was studied and applied to five underlying indexes. BOVA11, SPY, DAXEXx, MAXIS, and ISF ETFs reflect the Ibovespa, S&P500, DAX, NIKKEI, and FTSE indexes. The 12 largest constituent stocks that make up each index were selected, including their daily closing prices from January 1, 2012 to January 25, 2022, and the entire data set contains about 2,500 trading days. The historical data of these constituent stocks are used as the input data of the model to predict the return direction of the ETF on the following trading day.

Python 3.6.5 is the programming language, and Anaconda is the integrated development environment (IDE) used for programming. In order to ensure the robustness of the algorithm, the NumPy function was first used to calculate the logarithmic return and preprocess the data.

The data set was divided in half, taking into account the previous 1,000 trading days, and the distribution ratio between testing and training was set at 60/40%. The daily ETF return (t+1) was reversely distributed back to the returns of its 12 constituent stocks on the day before (t) to train the algorithm.

This enabled the algorithm to ascertain the relationship between the ETF's current day returns and the returns of its constituent equities the day before. A binary classification (0, 1) represented the ETF's return on the current day (t+1). If the predicted return was greater than zero, it was classified as 1, and if it was less than 0, it was classified as 0. Next, a machine learning model was imported to implement this method. Sharpe index, profit factor and maximum drawdown were used as financial metrics, and mean square error, root mean square error and mean absolute error are used as error metrics.

The results showed that among the two control strategies of buy and hold and naive forecast, SPY achieved the highest return, which can be explained by the higher linearity of growth of SPY constituents, while BOVA and NIKKEI had lower returns, and DAX and FTSE had downward returns. The machine learning model performed better than buy and hold on the Sharpe ratio and profit factor, except for BOVA in the classification model and FTSE in the classification model. The classification model and LSTM model showed the lowest drawdown value in general.

According to an analysis of the computational efficiency of errors and scores, most regression models had better scores and fewer prediction mistakes than naive forecasts. Machine learning models typically outperformed the naïve control approach and occasionally outperformed the buy and hold strategy when evaluating logarithmic returns. The highest-performing models were KNN and SVM, whereas the worst-performing models were LGBM regression and LGBM classifier. In conclusion, machine learning is an optional tool for financial data prediction. By forecasting the direction of returns, most models outperform buy and hold strategies regarding returns, errors, and scores.

3 CONCLUSIONS

This study selected a relatively novel model combination and research method in ETF prediction using machine learning Algorithms. This paper paid special attention to the superposition and combination of models, as well as the universality of models, and tried to find better ways to improve the effectiveness of ETF prediction. The results showed that STBC can reduce nearly 90% of prediction errors, and the combination of Fama-French three-factor model and LSTM can significantly improve the accuracy of prediction. It was worth noting that the effectiveness of different models in different ETF markets varies, so when choosing, it was still necessary to pay attention to their adaptability to specific ETFs. Currently, there are few studies on innovative models in the field of ETF forecasting, and their universality needs to be further proven. Most studies only focus on the application of one or several models to individual ETFs. Nevertheless, in the field of stock prediction, the number of similar innovative studies is relatively rich, which may be related to the difficulty of obtaining data sets. However, relatively speaking, the ETF market is still in the development stage, and ETF is still a young investment tool. It is reasonable that there are few related studies. Therefore, the researcher suggests that more public data sets belonging to ETFs should be established in the future to help other researchers reduce the difficulty of obtaining data sets and save a certain amount of data preparation time. Furthermore, the algorithmic process of deep learning models is still in a "black box" state. Researchers can visualize the model training process in future research or choose to use related Explainable AI (XAI) methods to improve the transparency and interpretability of the algorithm.

This study focuses on innovation and comprehensiveness beyond conventional methods, and opens up new ideas for researchers in the field of ETF prediction. This paper hope that researchers will fill the gap in this research field in the future, continue to explore innovative combinations and optimizations of machine learning models, and improve prediction accuracy.

REFERENCES

Chang, R. I., Wang, C. H., Wei, L. C., & Lu, Y. F. (2024). LSTM with short-term bias compensation to determine trading strategy under black swan events of Taiwan ETF50 stock. Applied Sciences, 14(18), 8576.

EPFR. (2021, November 11). Data reports global ETF assets exceed $10 trillion. EPFR. Retrieved from https://epfr.com/insights/press-release/epfr-data-reports-total-global-etf-assets-exceed-10-trillion/

Gowani, R., & Kanjiani, Z. (2024). Advanced LSTM neural networks for predicting directional changes in sector-specific ETFs using machine learning techniques. arXiv preprint arXiv:2409.05778.

Horst, J. (2022). Machine learning for exchange-traded fund price predictions (Bachelor's thesis, University of Twente).

Joshi, G., & Dash, R. K. (2024). Exchange-traded funds and the future of passive investments: A bibliometric review and future research agenda. Future Business Journal, 10(1), 17.

Liebi, L. J. (2020). The effect of ETFs on financial markets: A literature review. Financial Markets and Portfolio Management, 34(2), 165-178.

Li, H., Shen, Y., & Zhu, Y. (2018, November). Stock price prediction using attention-based multi-input LSTM. In Asian Conference on Machine Learning (pp. 454-469). PMLR.

Piovezan, R. P. B., de Andrade Junior, P. P., & Ávila, S. L. (2024). Machine learning method for return direction forecast of exchange-traded funds (ETFs) using classification and regression models. Computational Economics, 63(5), 1827-1852.

PricewaterhouseCoopers. (2021). ETFs 2026: The next big leap. PwC. Retrieved from https://www.pwc.com/gx/en/industries/financial-services/publications/etf-2026-the-next-big-leap.html

Sagarzazu Astuy, E. (2022). Estimación del retorno y riesgo esperado de un Exchange Traded Fund (ETF) aplicando modelos de machine learning (ML).

Shih, K.-H., Wang, Y.-H., Kao, I.-C., & Lai, F.-M. (2024). Forecasting ETF performance: A comparative study of deep learning models and the Fama-French three-factor model. Mathematics, 12(19), 3158.

Statista Research Department. (2023). Number of exchange-traded funds (ETFs) worldwide from 2003 to 2023. Statista. Retrieved from https://www.statista.com/statistics/278249/global-number-of-etfs/

Statista Research Department. (2024). Worldwide ETF assets under management 2003-2023. Statista. Retrieved from https://www.statista.com/statistics/224579/worldwide-etf-assets-under-management-since-1997/

Wang, J., & Zhang, G. (2024). Predicting relative performance of sector exchange-traded funds using machine learning. Journal of Business & Behavioral Sciences, 36(1).

Yüksel, B. (2023). Comparative analysis of LSTM model in predicting ETF stock prices for different sectors (Master's thesis, Marmara Universitesi, Turkey).