

# CRYPTOCURRENCY TRADING WITH ENSEMBLE MACHINE LEARNING ALGORITHM



Mr. Siwat Assavakijphanich

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Science in Insurance  
Department of Statistics  
FACULTY OF COMMERCE AND ACCOUNTANCY  
Chulalongkorn University  
Academic Year 2021  
Copyright of Chulalongkorn University

การซื้อขายสกุลเงินดิจิทัลด้วยการเรียนรู้ของเครื่องที่รวมการคาดคะเนจากหลายแบบจำลอง



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาการประกันภัย ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2564

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title                                    CRYPTOCURRENCY TRADING WITH ENSEMBLE  
    MACHINE LEARNING ALGORITHM  
By    Mr. Siwat Assavakijphanich  
Field of Study                                    Insurance  
Thesis Advisor                                   ISARIYA SUTTAKULPIBOON, Ph.D.

---

Accepted by the FACULTY OF COMMERCE AND ACCOUNTANCY,  
Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of  
Science

..... Dean of the FACULTY OF  
    COMMERCE AND  
    ACCOUNTANCY  
(Associate Professor Wilert Puriwat, D.Phil.)

THESIS COMMITTEE

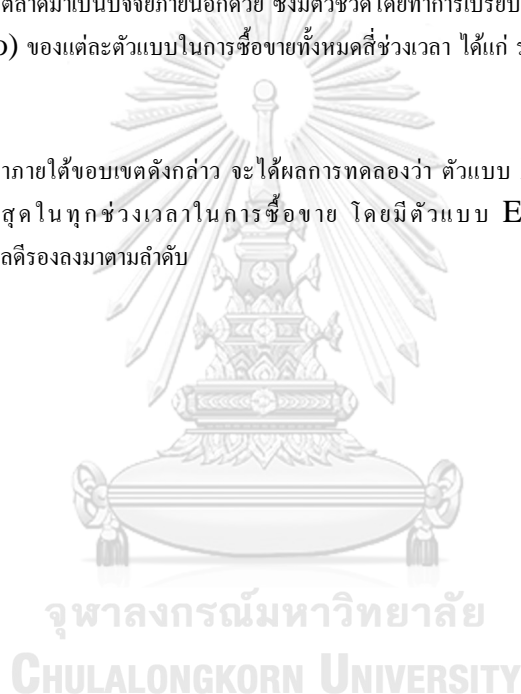
..... Chairman  
(Associate Professor ASSADAPORN SAPSOMBOON,  
Ph.D.)  
..... Thesis Advisor  
(ISARIYA SUTTAKULPIBOON, Ph.D.)  
..... Examiner  
(Assistant Professor VITARA PUNGPAPONG, Ph.D.)  
..... External Examiner  
(Assistant Professor Sampan Nettayanun, Ph.D.)

  
จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

ศิวัช อัสวักิจพานิช : การซื้อขายสกุลเงินดิจิทัลด้วยการเรียนรู้ของเครื่องที่รวมการคาดคะเนจากหลายแบบจำลอง.  
 ( CRYPTOCURRENCY TRADING WITH ENSEMBLE MACHINE  
 LEARNING ALGORITHM) อ.ที่ปรึกษาหลัก : อ. ดร.อิสริยะ สัตกุลพิบูลย์

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความแม่นยำของกลยุทธ์สำหรับการซื้อขายเหรียญคริปโตเคอเรนซี (Cryptocurrency) ที่ได้จากตัวแบบ XGBoost, ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนสำหรับการลดถอย, ตัวแบบโครงข่ายประสาทเทียม และ การรวมตัวแบบด้วยวิธี Equally Weighted Forecast Combinations, วิธี Adaptive Regression by Mixing และ วิธี Aggregation of Forecasts Through Exponential Reweight โดยมีเหรียญคริปโตเคอเรนซีทั้งหมดได้แก่ เหรียญ Bitcoin, เหรียญ Ethereum, เหรียญ Cardano, เหรียญ Binance Coin, เหรียญ Ripple, เหรียญ Polygon, เหรียญ Uniswap, เหรียญ Doge Coin, เหรียญ Chainlink และ เหรียญ Polkadot โดยนำราคาปิดของ ตลาดหุ้นทั้งหมดสิบตลาดทั่วโลก และ จำนวนของเหรียญในตลาดมาเป็นปัจจัยภายนอกด้วย ซึ่งมีตัวชี้วัด โดยทำการเปรียบเทียบค่าของความแม่นยำและอัตราส่วนชาร์ป (Sharpe Ratio) ของแต่ละตัวแบบในการซื้อขายทั้งหมดในช่วงเวลา ได้แก่ รายหนึ่งชั่วโมง, รายหกชั่วโมง, รายวัน และ รายอาทิตย์

จากการศึกษาภายใต้ขอบเขตดังกล่าว จะได้ผลการทดลองว่า ตัวแบบ XGBoost ให้ค่าความแม่นยำและอัตราส่วนชาร์ป สูงที่สุดในทุกช่วงเวลาในการซื้อขาย โดยมีตัวแบบ Equally Weighted Forecast Combinations ให้ผลดีรองลงมาตามลำดับ



สาขาวิชา            การประกันภัย  
 ปีการศึกษา        2564

ลายมือชื่อนิติ .....  
 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

# # 6280416326 : MAJOR INSURANCE

KEYWORD Cryptocurrency Trading, Ensemble Algorithm, XGBoost, Support  
D: Vector Machine, Long Short-Term Memory

Siwat Assavakijphanich : CRYPTOCURRENCY TRADING WITH  
ENSEMBLE MACHINE LEARNING ALGORITHM. Advisor: ISARIYA  
SUTTAKULPIBOON, Ph.D.

In this project, we aimed to use ensemble machine learning algorithms to trade ten cryptocurrencies along with attempting to add more external factors. Cryptocurrency included in this project were Cardano (ADA), Binance Coin (BNB), Bitcoin (BTC), DOGE, DOT, Ethereum (ETH), LINK, Polygon (MATIC), Uniswap (UNI), and Ripple (XRP). Furthermore, ten external factors, which are ten major stock indices, were added to the algorithm. All machine learning algorithms in this project are used to trade for four trading circumstances, an 1-hour interval, six-hour interval, daily interval and weekly interval. There are six machine learning models in this project which will be separated as based and ensemble models. XGBoost (XGB), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM) are based, models. Next, three based models are used to combine with three ensemble methods which are Equally Weighted Forecast Combinations (EW), Adaptive Regression by Mixing (ARM), and Aggregation of Forecasts Through Exponential Reweight (AFTER). Models are evaluated under two criterias, accuracy and Sharpe Ratio. As a result, XGBoost outperformed other models for all trading intervals while Equally Weighted Forecast Combination ranked second in both accuracy and Sharpe Ratio.



Field of Study: Insurance

Student's Signature

Academic Year: 2021

Advisor's Signature

Year:

.....

## ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my thesis advisor Isariya Suttakulpiboon, Ph. D. for the support throughout the research, for his patience motivation and immense knowledge. His guidance helped me in all the time of reseach and writing of this thesis. Futhermore, I would not have solved problems without your advice. I could not have imagined having a better advisor and mentor for my Master Degree study.

Besides my advisor, I would like to thank my thesis commitees: Assoc. Prof. Assadaporn Sapsomboon, Ph.D. , Asst. Prof. Vitara Pungpapong, Ph.D. ,and Asst. Prof. Sampan Nettayanun, Ph.D. for their insightful comments, research suggestions and encouragements. I would like to thank professors of the faculty of commerce and accoutancy, Chulalongkorn University for an opportunity to study in the Insurance program, and knowledge about insuarnce, and statistics.

Last, I would like to thank my family for supporting me since I was born and giving encouragement throughout the research, and my friends that giving guidnace and encouragement.

# TABLE OF CONTENTS

	<b>Page</b>
.....	iii
ABSTRACT (THAI) .....	iii
.....	iv
ABSTRACT (ENGLISH).....	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
List of Figures.....	1
List of Tables.....	2
Introduction.....	5
Literature Review.....	9
Cryptocurrency overview.....	9
Cryptocurrency Algorithmic Trading & Performance.....	12
The new algorithmic trading approach.....	16
Methodology.....	18
Data Collection & Preparation.....	18
Data Input & Model Training.....	19
Machine Learning-Based Trading Algorithms.....	20
Support Vector Machines (SVM).....	20
Recurrent Neural Network (RNN).....	22
Long Short-Term Memory (LSTM).....	24
Extreme Gradient Boosting (XGBoost).....	27
Model Combining Method.....	28
Equally Weighted Forecast Combinations.....	29
Adaptive Regression by Mixing.....	29
Aggregation of Forecasts Through Exponential Reweight.....	30

Model Evaluation.....	32
Result .....	33
Primary Result Evaluation .....	38
Visualizing Result.....	72
Discussion .....	79
Comparison with previous work.....	82
Research Limitation.....	83
Future Work Suggestion .....	84
REFERENCES .....	86
Appendix.....	88
VITA.....	133





## List of Figures

Figure 1: Recurrent Neural Network structure .....	23
Figure 2: Long-Short Term Memory Structure .....	25
Figure 3: Scatter plot of Accuracy versus Sharpe Ratio of training set for hourly trade .....	73
Figure 4 : Scatter plot of Accuracy versus Sharpe Ratio of testing set for hourly trade .....	74
Figure 5: Scatter plot of Accuracy versus Sharpe Ratio of training set for six-hour trade.....	74
<i>Figure 6: Scatter plot of Accuracy versus Sharpe Ratio of training set for six-hour trade .....</i>	<i>75</i>
Figure 7: Scatter plot of Accuracy versus Sharpe Ratio of training set for daily trade	75
Figure 8: Scatter plot of Accuracy versus Sharpe Ratio of testing set for daily trade.	76
Figure 9: Scatter plot of Accuracy versus Sharpe Ratio of training set for weekly trade .....	76
Figure 10: Scatter plot of Accuracy versus Sharpe Ratio of testing set for weekly trade .....	77
Figure 11: Example of a “good trend” prediction during research.....	84

## List of Tables

Table 1: Summary of Cryptocurrency Hourly Prices From 1 <sup>st</sup> June 2019 to 31 <sup>st</sup> May 2021.....	34
Table 2: Summary of Cryptocurrencies Arithmetic Hourly Returns 1 <sup>st</sup> June 2019 to 31 <sup>st</sup> May 2021 .....	35
Table 3: Correlation Table of Dataset.....	37
Table 4: Accuracy of all models and coins for hourly trading .....	39
Table 5: Accuracy of all models and coins for six-hour trading.....	40
Table 6: Accuracy of all models and coins for daily trading.....	42
Table 7: Accuracy of all models and coins for weekly trading .....	43
Table 8: Model performance of training set for hourly trade in business term .....	48
Table 9: Model performance of testing set for hourly trade in business term .....	51
Table 10: Model performance of training set for six-hour trade in business term .....	54
Table 11 : Model performance of testing set for six-hour trade in business term .....	58
Table 12: Model performance of training set for daily trade in business term.....	61
Table 13: Model performance of testing set for daily trade in business term.....	64
Table 14: Model performance of testing set for weekly trade in business term .....	68
Table 15: Model performance of testing set for weekly trade in business term .....	71

**Abbreviations in the report are represented as**

C1	as	Bitcoin
C2	as	Ethereurm
C3	as	Cardano
C4	as	Binance Coin
C5	as	Ripple
C6	as	Chainlink
C7	as	Uniswap
C8	as	Polkadot
C9	as	Dogecoin
C10	as	Polygon
F1	as	CAC 40 Index
F2	as	Nasdaq Composite
F3	as	DAX performance Index
F4	as	Hang Seng Index
F5	as	Bovespa Index
F6	as	Dow Jones Industrial Index
F7	as	Russell Index
F8	as	S&P 500 Index



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

F9	as	Euro Stoxx 50
F10	as	UTX Index
V1	as	Bitcoin's Volume
V2	as	Ethereum's Volume
V3	as	Cardano's Volume
V4	as	Binance Coin's Volume
V5	as	Ripple's Volume
V6	as	Chainlink's Volume
V7	as	Uniswap's Volume
V8	as	Polkadot's Volume
V9	as	Dogecoin's Volume
V10	as	Polygon's Volume
M1	as	XGBoost
M2	as	Support Vector Machines
M3	as	Long Short-Term Memory
M4	as	Equally Weighted Forecast Combinations
M5	as	Adaptive Regression by Mixing
M6	as	Aggregation of Forecasts Through Exponential Reweight

## Introduction

Because of its growing popularity in the past ten years, the stock trading algorithm has been studied by many researchers using different approaches. For example, (Shen et al., 2012) established a stock market prediction model with a machine learning model called Support Vector Machine (SVM). They predicted the next-day stock trend of NASDAQ, S&P500, and DJIA. The resultant trading model proved to be as accurate as any other benchmarks. In another instance, Zhong et al. (2019) utilized multiple machine learning algorithms in Deep Neural Networks and Artificial Neural networks to forecast the daily return of stocks. They predicted the daily return of the SPDR S&P500 ETF based on 60 financial and economic features. The experiment was successful; the algorithms worked for all stocks with at least 60% accuracy and 90% in some instances. Another example is (Dash & Dash, 2016). They employed their own computational efficient functional link artificial neural network (CEFLANN) model, a complex trading algorithm of stocks, which included technical factors for predicting the NASDAQ, S&P, and DJIA stock market. And while the model utilized a complicated technical analysis, the output it produced was a simple trading strategy consisting of h buy, hold and sell signals. Additionally, (Strader et al., 2020) reviewed and devised research directions of the stock trading algorithm since it had first originated. Artificial neural networks work best for predicting numerical values, while Support Vector Machines are ideal for solving the classification problems. Most investors tend to keep the information about their algorithms secret, making the stock trading area a zero-sum game. The reason they do so is to prevent equality in an investment firm, but most models are standardized.

Cryptocurrency trading algorithms can be divided into two categories: machine learning and non-machine learning. To illustrate, (Madan et al., 2015), (Colianni et al., 2015), (Żbikowski, 2016) and are three examples of the early stage of cryptocurrency trading algorithms. The first one, (Madan et al., 2015), stated that the non-machine learning model was only suited for the short run. The early stages of their models focused on Bitcoin (BTC), and the machine learning models were relatively simple, such as Random Forest and Support Vector Machines. Most researchers are looking for machine learning-based algorithms because non-machine learning models cannot keep up with the high-frequency data of cryptocurrency. The models of non-machine learning in this stage can perform well for only a short run of data, whereas their machine learning-based counterparts not only require less time to compute but are also more precise. As for (Lahmiri & Bekiros, 2019), (Koker & Koutmos, 2020), and (Sebastião & Godinho, 2021) they explored later stages of cryptocurrency trading algorithms, Recurrent Neural Network (RNN) and Long-Short Term Memory (LSTM), which are more complex and diverse than previous ones. They also predicted other cryptocurrency coins besides Bitcoin (BTC). In this stage, LSTM and RNN perform well with BTC, but not with other coins.

While the prediction of cryptocurrency price using machine learning-based models might be precise, there are some major issues that prevent it from being accepted as a tool for predicting stock. First, there is not a single well-standardized cryptocurrency coin price prediction model due to each coin's own unique fundamentals. At present, researchers are still seeking a new machine learning model that works as perfectly as possible. Currently, there are too many machine learning models that have been introduced into the field of cryptocurrency price prediction. A

few examples are Support Vector Machines, Recurrent Neural Network, Long-Short Term Memory, and Random Forest. As a result, none of the models is precise enough to predict most coins. Some models run smoothly only for forecasting the price of Bitcoin, while some only for Ethereum. More specifically, owing to a lack of consensus in the field of machine learning-based models for cryptocurrency algorithms, the best method that everyone agrees upon does not exist.

This is the opposite of the stock trading algorithm, which has been widely used and has a standardized model. The second problem with the prediction of cryptocurrency price using machine learning-based models is the relatively low accuracy of the cryptocurrency price predictor, compared to stock trading predictors. Prior study demonstrated that the machine learning model for cryptocurrency price prediction was accurate only for the short run, while the stock trading algorithm worked well for both the short and long run.

One way to improve the accuracy of machine learning models is by combining multiple machine learning models. An example of a study that used such an approach is that of (Lyu & Nikora, 1991), which combined software reliability prediction using the equally weighted linear combination model. In contrast, there are not many combined trading models for machine learning, especially ones with stock and cryptocurrency trading. The trading algorithm is standardized for stock and does not need other combination methods, as (Strader et al., 2020) mentioned. On the other hand, the cryptocurrency trading algorithm is neither standardized nor accurate; therefore, expanding the field of research on combining methods is one way to increase the accuracy of the model.

With the option of combining multiple machine learning models available, this new method is supposed to improve accuracy and solve the issue of standardization. The three models selected in this research, Long Short-Term Memory (LSTM), Support Vector Machines (SVM), Extreme Gradient Boosting (XGBoost), are going to be combined with three different combination methods: Equally Weighted Forecast Combinations, Adaptive Regression by Mixing, and Aggregation of Forecasts Through Exponential Reweight. Our objective is to create a model with more than 75 percent accuracy and to increase the accuracy of the combined model to be more than that of LSTM, SVM, and XGBoost for all cryptocurrency coins.





## Literature Review

### Cryptocurrency overview

A cryptocurrency is a digital currency with a unique security system called cryptography. Thanks to this security system, it is highly difficult to counterfeit a cryptocurrency, allowing users to conduct safer transactions than relying on conventional financial institutions. Cryptography was also the reason the currency generated attention among investors since its foundation in 2009.

There has been an extremely dramatic rise in the cryptocurrency market, especially in recent years. Bitcoin, the originator of all cryptocurrencies, has risen over 4000 percent since 2017 and as of April 2021, it has a combined market value of over 1,100 billion dollars. Due to its unique fundamentals, the cryptocurrency has attracted the interest of a significant number of investors since 2020. The number of global crypto users exceeded 101 million in the third quarter of 2020, compared to 5 million in 2016, or a rise of more than 2000 percent.

Despite its extraordinarily high value, the price of cryptocurrency is extremely volatile and responds to many external factors. For example, in one of the earliest incidents of the market, the price of the DOGE coin plunged by nearly 30 percent after Elon Musk had mentioned it in a television broadcast called Saturday Night Live. The latest announcement of Tesla concerning Bitcoin is another instance that illustrates how cryptocurrency can be influenced by random external factors. In another example, Tesla stated on 14th June 2021 that the company would continue to accept BTC as a currency. As a result, the price of BTC surged by 9.8% solely due to one incident. Unlike that of other trading assets, the value of cryptocurrency can drastically rise and fall in a short period of time. Furthermore, cryptocurrency is an

alternative for diversifying portfolio risk. According to (Chuen et al., 2017), cryptocurrency can be considered an option because of the fact that its correlation value between traditional assets is consistently low. Cryptocurrency's average return is also higher than that of conventional investments.

In trading, investors seek ways to trade efficiently, and one of those ways is finding factors impacting trading assets, including cryptocurrency. Common factors are supply and demand, company-related factors, investor sentiment, and interest rates. Another example of factors is Volatility and significant stock indices, which are more complex ones.

Factors that influence cryptocurrency, like other assets, can be divided into two parts: internal and external factors, according to (Poyser, 2017). Internal factors are transaction cost, reward system, mining difficulty (Hash rate), coin circulation, and forks. Those factors are due to cryptocurrency's fundamentals themselves, which directly impact the prices of the currency and changes over time. On the other hand, though the external factors do not shift over time, they exert a lower impact on price changes. Attractiveness, exchange rates, and restrictions are examples of external factors.

Investors and researchers are examining factors impacting cryptocurrency with different approaches. (Sovbetov, 2018) tried to find those factors by applying the ARDL technique and document findings. He conducted an experiment by testing five cryptocurrency coins: Bitcoin, Ethereum, Dash, Litecoin, and Monero. The author noted that trading volume and Volatility affect all coins selected in the short- and long run. He also observed that the S&P500, one of the major stock indices, has a

negligible impact on prices. (Sovbetov, 2018) concluded the study by suggesting that many more factors still remain to be discovered.

Moreover, (Liu & Tsyvinski, 2021) performed an experiment on Bitcoin regarding factors predicting cryptocurrency returns. The authors also attempted to propose their hypothesis that the behavior of cryptocurrencies is driven as a unit of account, similar to that of other currencies. Consequently, it can be concluded from the research that cryptocurrency returns can be predicted by only two factors related to its market: momentum and investors' attention. According to (Sovbetov, 2018) , Volatility is the only factor from the internal factors that can be exploited and also has the most impact on predicting prices. Other internal factors cannot be categorized as a predictor because of their complexity and limited resources; conversely, although the information regarding Volatility is easily accessible, it must be calculated before actual use.

Multiple researchers are experimenting with different models for calculating Volatility. According to (Katsiampa, 2017), numerous GARCH models were used to compile the value of Volatility. In the end, using AR-CGARCH is the best option because the model has the highest accuracy out of all GARCH models. Nevertheless, (Charles & Darné, 2019) later countered (Katsiampa, 2017). They replicated and reanalyzed GARCH models with modern techniques. The result demonstrated that all GARCH models are unsuitable for predicting the Volatility of cryptocurrency, indicating that there is no proper method for calculating Volatility.

External factors affecting cryptocurrency prices that can be utilized at the moment are the exchange rate and stock markets' indices. Stock markets' indices

include the major ones, which are S&P500 (SPX), Dow Jones Industrial Average (DJI), and Nasdaq Composite (IXIC). Only these two factors, the exchange rate and stock markets' indices, can be used thanks to their numerical value. If investors do not use numerical data as a predictor, the result will be both too complicated and inaccurate.

### **Cryptocurrency Algorithmic Trading & Performance**

Cryptocurrency trading predictions and other trading predictions originated in mathematical models such as regression. Experienced and skilled investors even rely on machine learning, rather than regression, as a predictor because of its superior speed and accuracy. Machine learning algorithms could also be applied to stock trading, as shown by many studies.

Using regression as a cryptocurrency predictor is outdated. Madan et al. (2015) tested the regression model for forecasting cryptocurrency trading. The model performed well but only in low-frequency data. Regression cannot be used in the long run or high-frequency data because it takes too long to compute. Furthermore, the mathematical model of regression is not sophisticated enough to solve a time series like data.

Though the fundamentals of cryptocurrency may pose difficult problems for machine learning researchers in the long run, machine learning algorithms for cryptocurrency are new methods for forecasting or building trading algorithms despite the currency's unstable performance. In consequence, numerous researchers (Madan et al., 2015; Colianni et al., 2015; Lahmiri; Zbikowski, 2016; Bekiros, 2019; Koker

and Koutmos, 2020, and Sebastial and Godinho, 2021) are presently exploring the field of algorithmic trading with the intention of applying it to cryptocurrency.

The algorithmic trading researchers have been performing experiments on machine learning models, starting with ones that were comparatively simple and scaling up in complexity after. Madan et al., 2015; Colianni et al., 2015; Lahmiri; Zbikowski, 2016; Sebastial and Godinho, 2021 experimented on a low complex model such as Support Vector Machines, while Bekiros, 2019; Koker, and Koutmos, 2020 were on more complex machine learning models such as Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM).

Madan et al. (2015) were early publishers who studied machine learning for cryptocurrency. Their framework was to test whether machine learning will eventually outperform the regression model. The selected machine learning models were Random Forest and Support Vector Machines. Due to the limited access to data, only Bitcoin was chosen as the testing coin, as this research aimed to use different frequencies of dataset such as daily, 10 minutes, and 10 seconds. The result was that low-frequency data, such as daily prices, could predict the price change with an accuracy of 98.7%, but in the long run like high-frequency data, machine learning could not perform at the high level of around 50-55%. This approach also did not include external factors in the algorithm.

Colianni et al. (2015) later considered using Twitter direct texts called "tweet" as an external factor in predicting Bitcoin's price. The method is also known as sentiment analysis. Support Vector Machines, Logistic Regression, and Naïve Bayes were selected as models for Bitcoin. Experiments were conducted on two sets of data,

hourly and daily data, for 21 days. The result was that both the day-to-day and hour-to-hour predictions achieved more than 90% accuracy. The publishers also stated that applying an algorithm to a frequency higher than hour-to-hour would be excessively complicated because the massive number of tweets would render sentiment analysis inefficient. Also, more advanced machine learning should be applied to create a better algorithm.

Zbikowski (2016) invented new models for enhancing cryptocurrency trading algorithms: Exponential Moving Average, Box Support Vector Machines, and Volume-Weighted Support Vector Machines. The objective of the research was to prove that the three new models were suitable for cryptocurrency trading algorithms. A high-frequency data (15minutes) was selected on a 24-day span, and the currency was Bitcoin (BTC). The result was that all models invented could enhance previous successful ones, primarily Support Vector Machines. In this research, two out of three models were the Support Vector Machines model that had been upgraded so that it could perform under more extreme circumstances like a high-frequency dataset. As Madan et al. (2015) stated, plain Support Vector Machines could perform only on the low-frequency dataset: the daily dataset.

Lahmiri and Bekiros (2019) explore one of the most advanced machine learning models: Long Short-Term Memory (LSTM). LSTM computes time-series data like cryptocurrency prices at a high level . A longer data span (8 years) was tested with this model, with the purpose being to find its advantages and disadvantages. The result was that LSTM is by far the most accurate model because

of its complexity. Nonetheless, the researchers also noted that the more complex a model is, the longer it takes to compute or operate data.

Though four studies that explored these subject yielded promising results, none of them considered relevant factors impacting cryptocurrency price changes. All authors suggested that in the future, algorithms should be improved by expanding with new factors as a parameter. Also, implementing new factors should help reveal the limitations of each model.

As mentioned in the previous section concerning factors impacting cryptocurrency prices, exploring more factors creates algorithmic trading for cryptocurrency. Koker and Koutmos (2020) and Sebastial and Godinho (2021) reintroduced less complex machine learning algorithms while experimenting with factors instead of experimenting with new models without factors.

The goal of Koker and Koutmos (2020) was to prove that cryptocurrency prices follow a random walk process with factors. Multiple coins, such as Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), Ripple (XRP), and Monero (XMR), were tested to expand the range of trading algorithms. The result was not as expected; cryptocurrency prices do not follow the random walk process while applying factors in a model.

Sebastial and Godinho (2021) also aimed to prove the same point as Koker and Koutmos (2020) but with different models, such as Random Forest and Support Vector Machines. The settings of this research consisted of 4 years of data span and three different coins: Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC). As previously noted, the experiment was conducted using low-frequency data, since less complex

models limit low-frequency datasets. The result was unsatisfactory, similar to that of Koker and Koutmos (2020). All models selected performed well only for Bitcoins (BTC), as opposed to Ethereum (ETH) and Litecoin (LTC). The authors also proposed that for future works, data partitioning and parameter settings can improve machine performance. besides factors.

To summarize the overview of current models, the summary Table 1A, shown in the appendix, is created to display the apparent limitations of each model. First, most of the successful trading algorithms are suitable for Bitcoin (BTC). However, cryptocurrency is not solely about Bitcoin, but it also encompasses thousands of other trading coins. Consequently, to increase the success of algorithms, models need to be able to function properly with other coins as well. Second, more complex models such as RNN or LSTM are not successful when factors are applied. Lahmiri and Bekiros (2019) stated that models were successful only when factors were not applied. Last, when factors are applied to trading algorithms, models are unable to perform in high-frequency circumstances. Both Koker and Koutmos (2020) and Sebastiao and Godinho (2021) used daily data to train the model, and the result went well only for Bitcoins. After reviewing the literature, algorithmic trading is proposed to improve cryptocurrency trading strategies.

### **The new algorithmic trading approach**

A new method is proposed to explore more into the field of cryptocurrency algorithms. Three machine learning models, Long Short-Term Memory (LSTM), Support Vector Machines (SVM), Extreme Gradient Boosting (XGBoost), are used with factors in addition to sentiment analysis. These models are combined with three combination methods in this research: Adaptive Regression by Mixing, Equally



Weighted Forecast Combinations, and Aggregation of Forecasts Through Exponential Reweight. Last, each of the ten cryptocurrency coins is individually chosen for the prediction, with external factors being the other nine coins not selected, together with the ten major stock indices.

There are several reasons for proposing a new method to fill research gaps. First, models are not generalized for all cryptocurrency coins; each model selected has its advantage. Though LSTM has a complex structure and high accuracy compared to others, they can perform well only under low-frequency data. In contrast, SVM has a less complex structure, but it can perform in high-frequency datasets like 60 minutes. Meanwhile, XGBoost is suitable for low-frequency dataset, and when computing the model itself can properly deal with missing dataset. As a result, unlike stock trading algorithms, all three models are not generalized. Second, no research has been done regarding combining machine learning models for cryptocurrency. (Choudhry & Garg, 2008) and (Shen et al., 2012) stated that "combined" machine learning algorithms yield a satisfactory result with high accuracy. Therefore, the combined model might help fill the gap caused by the low accuracy of the cryptocurrency trading algorithm. Last, besides proposing new models, expanding the scope of data is also an objective of this research. Existing research was tested mainly on the originator of cryptocurrency, Bitcoin, because in the past, other coins were not nearly as popular and their prices tended to rise or fall according to that of Bitcoin. However, some coins do not follow Bitcoin now, so we should find a generalized model that is capable of predicting all coins. Therefore, this research will deal with predicting ten cryptocurrency coins.

## Methodology

This research aims to create new algorithmic approaches for cryptocurrency trading and find the best models for each cryptocurrency coin. To do so, three machine learning models, Support Vector Regression (SVM), Extreme Gradient Boosting (XGBoost), and Long Short-Term Memory (LSTM), are chosen as "base models." Base models are turned into new models by combining them with three methods: Equally Weighted Forecast Combinations, Adaptive Regression by Mixing, and Aggregation of Forecasts Through Exponential Reweight. Data collected will be used as the input for predicting prices, and the output will be the price at the next timestamp of a specific cryptocurrency coin. A simple short-long trading strategy will be formulated to test the accuracy of the models, and the Sharpe ratio will also be used to evaluate the machine learning model. The detailed methodology is separated into four sections: Data Collection & Preparation, Model, Machine Learning Based Trading Algorithms, Model Combining Method, and Model Evaluation.

### Data Collection & Preparation

Data Collection & Preparation was the preliminary step of creating machine learning-based algorithms. Ten cryptocurrency coins were collected under the hourly frequency in Python. The data was collected from 1st June 2019 to 31st May 2021. The 10 Cryptocurrencies that were chosen were the top 10 according to [www.coinmarketcap.com](http://www.coinmarketcap.com), where coins are ranked by "Market Cap," which is calculated by multiplying a coin's price with its current supply. Ten coins selected are Bitcoin(BTC), Ethereum (ETH), Tether (USDT), Cardano (ADA), Binance Coin (BNB), Ripple (XRP), Dogecoin (DOGE), USD Coin (USDC), Polkadot (DOT) and

Uniswap (UNI). Coins' trading fees vary day to day, so we assume the coins' trading fee to be 0.1% for all coins traded in each transaction, according to Binance.

The other essential data for machine learning are factors influencing the predicted values. Besides ten cryptocurrency prices, ten major stock indices would be incorporated as factors for predicting the prices. The ten major stock indices, ranked by market cap as of May 2021, were NASDAQ-100, S&P 500, Hang Seng Index, FTSE 100, Dow Jones Industrial Average, DAX30, Russell 2000, CAC40, Euro STOXX50, and Bovespa Index. A reason behind selecting stock indices is that evidence was founded on the cointegration relationship between cryptocurrency and major stock indices, as Dirican and Canoz (2017) stated. All factors' values were collected from 1st June 2019 to 31st May 2021. For this research, all of the data collected was used as factors for the machine learning model. For example, only Bitcoin's price was predicted in the Bitcoin prediction algorithm, and the nine cryptocurrency coins and ten major stock indices would be used as factors for Bitcoin.

The ten cryptocurrency coins' prices would be split into two datasets for training and testing machine learning models. The data was split by a ratio of 70% and 30% to increase the efficiency of the model, according to Breiman and Spector (1992). The training data is 70% of the dataset from 1st June 2019 to 30th September 2020, while the test set, or the remaining 30%, is from 1st October 2020 to 31st May 2021.

### **Data Input & Model Training**

After data preparation, datasets, the ten cryptocurrency coins' prices, and external factors will be input to a separate machine learning model. The machine learning model uses three main types of data: Images, Text, and categorical data

(.CSV), but in this research, only the categorical data will be used. Categorical data or a CSV file is a spreadsheet that contains all datasets prepared, but since our models can train with only one coin at a time, there will be ten separate CSV files for each coin.

After inputting the data for each cryptocurrency coin, the next step of the research was model training. All models and coin prediction were trained separately, which meant that all coins and models were independent of each other. The dataset would undergo training, but the time elapsed for each model was not the same. How much time the process required depended on the complexity of the machine learning model.

The machine learning progress will be operated under Python version 3.8.0 with multiple libraries, numpy, pandas, seaborn, matplotlib, sklearn, xgboost, keras and tensorflow. Libraries imported in this research are divided into three purposes. First, numpy and pandas are imported for cleaning and operating with out dataset. Second, seaborn and matplotlib are libraries for plotting figures for the research. Last, sklearn, xgboost, keras and tensorflow are imported for machine learning purposes.

### **Machine Learning-Based Trading Algorithms**

#### **Support Vector Machines (SVM)**

*Support Vector Machines* is a classifier formally defined by a separating hyperplane. Not only is the model suitable for data classification, but it can also be used for regression analysis. According to (Ince & Trafalis, 2008), the Support Vector model consists of Kernel Function, which acts as translation function in the matrix as

$$K(u, v) = \phi(u)^T \phi(v) \quad (1)$$

Commonly used Kernel Functions are Linear, Polynomial, Radial Basis Function, Gaussian, and Sigmoid. For this trading algorithm, Polynomial Function, Gaussian Function, and Radial Basis Function are selected.

- Polynomial  $K(x_i, x_j) = (x_i^T y_j + 1)^d$  (2)

- Radial Basis Function  $K(x_i, x_j) = e^{-\gamma|x_i-x_j|^2}, \gamma > 0$  (3)

- Gaussian  $K(x_i, x_j) = e^{-\frac{1}{2\sigma^2}|x_i-x_j|^2}, \gamma > 0$  (4)

while  $d$ ,  $\gamma$ ,  $a$ , and  $b$  are the parameters of the Kernel function, which is commonly adjusted manually, depending on the situation.

The predicting model of SVM is defined as

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \langle x_i, x_j \rangle + b \quad (5)$$

Consider selected Kernel functions; the functions can be applied to the predicting model in place of  $\langle x_i, x_j \rangle$  Of equation (5). The selected Kernel functions will act as translators for the non-linear dataset. As a result, the proper Support Vector Regression function is defined as

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K \langle x_i, x_j \rangle + b \quad (6)$$

For forecasting cryptocurrency prices, Support Vector Regression was selected by using Polynomial Function, Gaussian Function, and Radial Basis Function as the Kernel Function under two circumstances which are

- C from the value of 1 to 10 (Increasing by 1)
- Epsilon ( $\epsilon$ ) value of 0.01 to 0.10 (Increasing by 0.01)

The reason behind selecting this model was that SVM itself does not depend on an input space's dimension. It means that SVM can be used to map lower-dimensional data into higher-dimensional data. After running the model, by considering MSE or mean square error, which is defined as

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

A model with the lowest value of MSE was selected for creating trading strategies. The trading strategies were short/long with the following criteria.

- Let  $y_{t+1} = x_{t+1} - c$  for each hour price, by having  $y_{t+1}$  as the price after one hour after deducting transaction costs and  $x_{t+1}$  as the price of the previous hour, and  $c$  as the transaction fee of any cryptocurrency coins value of 0.1%.
- If  $y_{t+1} > 0$ , use a long strategy for a specific cryptocurrency coin at the time  $t+1$ .
- If  $y_{t+1} \leq 0$ , use a short strategy for a specific cryptocurrency coin at the time  $t+1$ .

#### Recurrent Neural Network (RNN)

A recurrent Neural Network or RNN is an artificial neural network commonly used in natural language processing (NLP). RNN is designed to recognize a sequential dataset like cryptocurrency. According to (Mikolov et al., 2010), the model uses its complexity and performance to predict the likely scenario. RNN uses an algorithm called backpropagation to update the weights of the networks. The model consists of 3 components: input, hidden layer, and output, as follows.

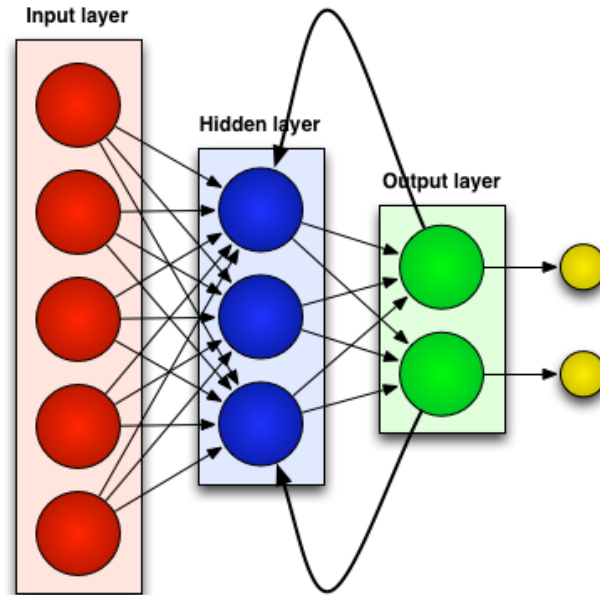


Figure 1: Recurrent Neural Network structure

The hidden layer of RNN is the model's core because the number of hidden layers varies according to the model's performance. In a hidden layer, four values are calculated at Timestep(t)

$$H_t = \sigma(U * X_t + W * H_{t-1}) \quad (8)$$

$$y_t = \sigma(V * H_t) \quad (9)$$

$$J^t(\theta) = -\sum_{j=1}^{|M|} y_{t,j} \log y_{t,j} \quad (10)$$

$$J(\theta) = -\frac{1}{T} \sum_{t=1}^T \sum_{j=1}^{|M|} y_{t,j} \log y_{t,j} \quad (11)$$

Where

- $H_t$  is a hidden layer time step

- $\sigma$  is an activation function
- $U$  is the weight vector for the Hidden layer
- $W$  is the same weight vector for different time stamps
- $V$  is the weight vector for the output layer
- $y$  is word vector for Output word
- $t$  is the time elapsed
- $\sigma$  is SoftMax function

### Long Short-Term Memory (LSTM)

Like RNN, Long Short-Term Memory or LSTM is a type of artificial neural network, but the latter has a relatively more complex structure. In RNN, there are three components, which are input, output, and hidden layer, whereas LSTM has six components or LSTM cells. In other words, LSTM is an upgrade version of RNN. According to Chen et al. (2019), the components are Forget Gate, Candidate layer, Input Gate, Output Gate, Hidden state, and Memory state. LSTM cell diagram at the time step  $t$  is



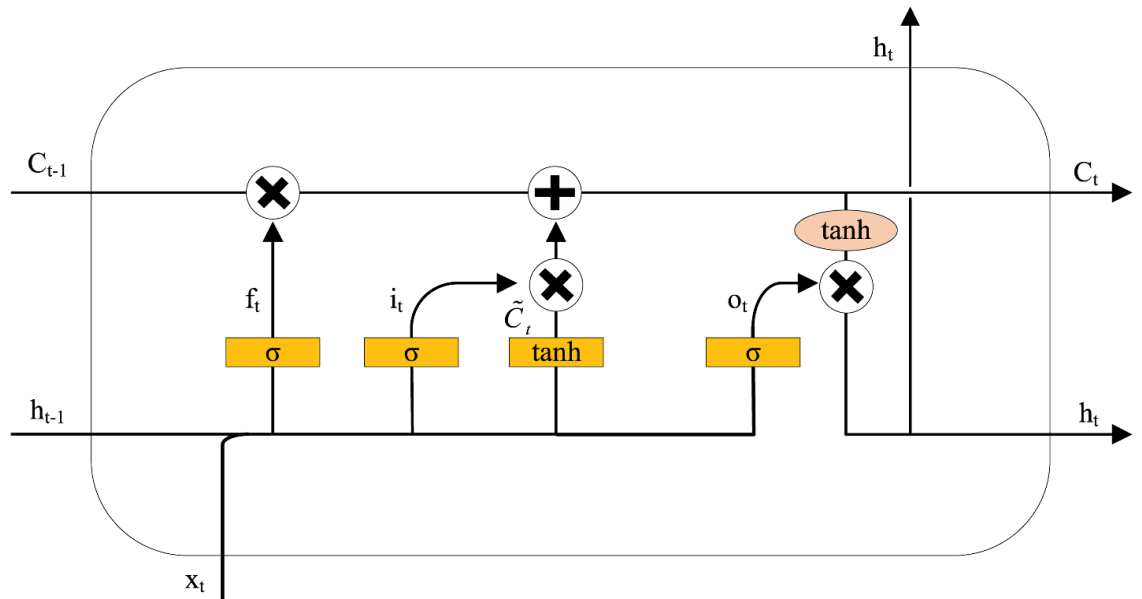


Figure 2: Long-Short Term Memory Structure

Where

- $X_t$  is input vector
- $H_{t-1}$  is previous cell output
- $C_{t-1}$  is previous cell memory
- $H_t$  is current cell output
- $C_t$  is current cell memory
- $f$  is the forget gate
- $I$  is an input gate
- $O$  is an output gate

From the labeled LSTM cell, some components are functions as

$$f_t = \sigma(X_t * U_f + H_{t-1} * W_f) \quad (13)$$

$$\underline{C}_t = \tanh(X_t * U_c + H_{t-1} * W_c) \quad (14)$$

$$I_t = \sigma(X_t * U_i + H_{t-1} * W_i) \quad (15)$$

$$O_t = \sigma(X_t * U_o + H_{t-1} * W_o) \quad (16)$$

$$C_t = f_t * C_{t-1} + l_t * \underline{C}_t \quad (17)$$

$$H_t = O_t * \tanh(C_t) \quad (18)$$

For cryptocurrency trading algorithms, Long Short-Term Memory (LSTM) is used under backpropagation while applying learning rate at 0.001, 0.01, and 0.1 respectively, with hidden layers of 2. Similar to RNN, selecting a lower learning rate can make LSTM perform at a better level. However, it also requires more time and more computer performance. The advantage of using LSTM is its complexity. LSTM is complex and suitable for sequential data or time series. MSE is a criterion for finding the best model; the lower MSE a trading strategy has, the better. By considering MSE or mean square error, which is defined as

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (19)$$

A model with the lowest value of MSE is selected for creating trading strategies. The trading strategies are short/long with the following criteria.

- Let  $y_{t+1} = x_{t+1} - c$  for each hour price, by having  $y_{t+1}$  as the price after one hour after deducting transaction costs and  $x_{t+1}$  as the price of the previous hour, and  $c$  as the transaction fee of any cryptocurrency coins value of 0.1%.
- If  $y_{t+1} > 0$ , use a long strategy for a specific cryptocurrency coin at the time  $t+1$ .

- If  $y_{t+1} \leq 0$ , use a short strategy for a specific cryptocurrency coin at the time  $t+1$ .

#### Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting or XGBoost is an ensemble machine learning model that implemented from gradient boosting. XGboost contains features that surpass Gradient boosting which are regularization, sparse aware, parallelization and cache optimization. According to (Chen & Guestrin, 2016), XGboost contains an objective equation of

$$L^{(t)} = \sum_{i=1}^n [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega f_t(x_i) \quad (20)$$

The general parameter of the XGBoost model is gmtree booster while setting up other parameters as default values to test. The model is unlike LSTM or SVM that XGBoost can adjust the parameter after computing models. The model will be adjusted with parameters eta (0.01-0.3) and max\_depth (3-10).

After adjusting, MSE will be used as a criteria for finding the best model; the lower MSE a trading strategy has, the better. By considering MSE or mean square error, which is defined as

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (21)$$

A model with the lowest value of MSE is selected for creating trading strategies. The trading strategies are short/long with the following criteria.

- Let  $y_{t+1} = x_{t+1} - c$  for each hour price, by having  $y_{t+1}$  as the price after one hour after deducting transaction costs and  $x_{t+1}$  as the price of the previous hour, and  $c$  as the transaction fee of any cryptocurrency coins value of 0.1%.
- If  $y_{t+1} > 0$ , use a long strategy for a specific cryptocurrency coin at the time  $t+1$ .
- If  $y_{t+1} \leq 0$ , use a short strategy for a specific cryptocurrency coin at the time  $t+1$ .

### Model Combining Method

Relying on a single forecasting model has many limitations. Forecasting values are sometimes unstable or uncertain, which means that if values in the dataset are changed, the forecasting values will also change significantly. (Armstrong, 2001) stated that combining forecasts is especially useful when uncertain which method is the most accurate. Furthermore, the result from the study showed that combined methods were sometimes more accurate than even their most accurate components.

To eliminate the instability and the uncertainty of forecasting models, combined forecasts are invented and defined as

$$y_i^c = \sum_{j=1}^M w_j y^{(j)}, i = 1, 2, \dots, n \quad (22)$$

where

- $y_i^c$  is the combined forecasted value
- $y^{(j)}$  is predicted value of selected models  $j, j = 1, 2, \dots, M$
- $w_j$  is the weight distributed to each model

- $j$  is single forecasting model of  $j, j = 1, 2, \dots, M$

The combination methods selected for testing with cryptocurrency forecasting were Adaptive Regression by Mixing, Equally Weighted Forecast Combinations, and Aggregation of Forecasts Through Exponential Reweight. All of these models were proposed by (Yang, 2003). Using these three techniques, our method merges all previous single models (SVM, XGBoost, and LSTM) into all combined forecast methods.

#### Equally Weighted Forecast Combinations

Equally Weighted Forecast Combinations or EW is a method that equally distributes weights of forecasting model of  $j$  as

$$w_j = \frac{1}{M} \quad (23)$$

#### Adaptive Regression by Mixing

Adaptive Regression by Mixing or ARM method is a combined forecast method where errors of based models are normally distributed. The data for this method is equally split into two sets. The first is for estimating parameters, while the second set is used to estimate the accuracy. Finally, parameters and accuracy are used for weight calculation.

Detailed processes of ARM are divided into four sections: first, equally split data while

$$Z^{(1)} = (X_i, Y_i)_{i=1}^n \quad (24)$$

$$Z^{(2)} = (X_i, Y_i)_{i=n/2+1}^n \quad (25)$$

Next,  $Z^{(1)}$  is used for estimating the new coefficient of model  $j$ ; the new model is defined as

$$y_i^{\wedge(j)} = f_{j,Z^{(1)}}(x'_i, \theta_{j,Z^{(1)}}^{\sim}), i = 1, 2, \dots, \frac{n}{2}, j = 1, 2, \dots, M \quad (26)$$

, and  $Z^{(2)}$  is used for calculating the accuracy of the previous model.

Next, an overall measure of discrepancy is calculated by

$$D_j = \sum_{i=n/2+1}^n \left( y_i - f_{j,Z^{(1)}}(x'_i, \theta_{j,Z^{(1)}}^{\sim}) \right)^2 \quad (27)$$

Lastly, the New weight for model  $j$  is calculated from

$$w_j = \frac{(\sigma_{j,Z_1}^{\wedge})^{-n/2} e^{\left( \frac{-(\sigma_{j,Z_1}^{\wedge})^{-2} D_j}{2} \right)}}{\sum_{q=1}^M (\sigma_{q,Z_1}^{\wedge})^{-n/2} e^{\left( \frac{-(\sigma_{q,Z_1}^{\wedge})^{-2} D_q}{2} \right)}} \quad (28)$$

### Aggregation of Forecasts Through Exponential Reweight

Aggregation of Forecasts Through Exponential Reweight or AFTER method is for combining forecasts where errors are typically distributed where using Error Variance Estimates  $(\sigma_{it}^{\wedge 2})$  where

$$\sigma_{it}^{\wedge 2} = \sum_{\tau=1}^{t-1} \frac{e_{i\tau}^2}{(t-1)} \quad (29)$$

Let  $\pi = \{\pi_i: i \geq 1\}$  as prior weights where all of them are equals when  $\pi_i = 1$  and  $w_{i1} = \pi_i$

For  $t \geq 2$ , weight calculation can be done by

$$w_{it} = \frac{w_{i1} \left( \prod_{\tau=1}^{t-1} \sigma_{it}^{-1} \right) \exp \left( -\frac{1}{2} \sum_{\tau=1}^{t-1} \frac{e_{it}^2}{\sigma_{it}^2} \right)}{\sum_{j=1}^m \left( w_{j1} \left( \prod_{\tau=1}^{t-1} \sigma_{jt}^{-1} \right) \exp \left( -\frac{1}{2} \sum_{\tau=1}^{t-1} \frac{e_{jt}^2}{\sigma_{jt}^2} \right) \right)} \quad (30)$$

Where  $i = 1, 2, \dots, m$  and  $\tau \geq 1$

A model with the lowest value of MSE is selected for creating trading strategies. The trading strategies are short/long with the following criteria.

- Let  $y_{t+1} = x_{t+1} - c$  for each hour price, by having  $y_{t+1}$  as the price after one hour after deducting transaction costs and  $x_{t+1}$  as the price of the previous hour, and  $c$  as the transaction fee of any cryptocurrency coins value of 0.1%.
- If  $y_{t+1} > 0$ , use a long strategy for a specific cryptocurrency coin at the time  $t+1$ .
- If  $y_{t+1} \leq 0$ , use a short strategy for a specific cryptocurrency coin at the time  $t+1$ .

Applying the same trading strategies as the previous models is the critical point in this research. The machine learning models needed to be under the same circumstances as trading strategies so that they could be compared to each other after combining.

### Model Evaluation

The six models achieved, which are Long Short-Term Model (LSTM), Support Vector Machines (SVM), Extreme Gradient Boosting (XGBoost), Adaptive Regression by Mixing, Equally Weighted Forecast Combinations, and Aggregation of Forecasts Through Exponential Reweight will be compared to each other using two values: Sharpe Ratio and accuracy. Sharpe Ratio is for measuring the performance when models are performing trading strategies. The formula of Sharpe Ratio is

$$\frac{R_p - R_f}{\sigma_p} \quad (31)$$

where:

- $R_p$  is the return of the trading
- $R_f$  is the risk-free rate
- $\sigma_p$  is the standard deviation of excess return

The other value indicating the best model is accuracy. Accuracy is for the evaluation of machine learning models and has the following definition.

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (32)$$

After gathering all of Sharpe Ratio's accuracy and accuracy, the comparison chart will be devised to illustrate the best model. The models with the highest Sharpe Ratio or accuracy are collected. The method is applied individually to each cryptocurrency coin.



## Result

The purpose of this research is to compare different trading algorithms on ten cryptocurrency coins namely Bitcoin, Ethereum, Cardano, Binance Coin, Ripple, Polygon, Chainlink, Uniswap, Doge, and Polkadot in Binance. Six trading algorithms are applied - XgBoost, Support Vector Machine, Long-Short Term Memory, Equally Weighted Forecast Combinations, Adaptive Regression by Mixing and Aggregation of Forecasts Through Exponential Reweight. The dataset includes ten cryptocurrency coin prices, ten major stock indices and coins' trading volumes from 1<sup>st</sup> June 2019 to 31<sup>st</sup> May 2021 as external factors. Data summary and the main analyses are as follows.

Coins	Obs.	Mean	SD.	Min	Q1	Media n	Q3	Max	Skewnes s	Kurtosi s
C1	1735 3	17993.12	15975.6 3	8585.99	8786	10249	18410	64577	1.59	1.01
C2	1735 3	622.40	787.26	274.72	186	247	557	4333	2.18	4.41
C3	1735 3	0.27	0.45	0.09	0.05	0.08	0.15	2.42	2.15	3.49
C4	1735 3	75.88	139.45	32.95	17.13	23.33	32.84	684.22	2.74	6.66
C5	1735 3	0.37	0.30	0.44	0.22	0.27	0.39	1.94	2.88	8.13
C6	1735 3	10.41	11.09	1.04	2.44	4.28	13.51	52.27	1.48	1.37
C7	1289 2	16.98	13.99	23.26	4.30	9.28	30.71	49.28	0.48	-1.37
C8	1654 4	0.03	0.10	0.33	0.00	0.00	0.00	0.74	3.92	15.52
C9	6091	16.26	13.45	28.24	3.61	9.17	29.13	44.64	0.42	-1.38
C10	1735 3	0.11	0.30	0.02	0.01	0.02	0.03	2.57	4.67	23.80
Factors	Obs.	Mean	SD.	Min	Q1	Media n	Q3	Max	Skewnes s	Kurtosi s
F1	1735 3	5413.19	560.37	5241.46	4965	5528	5861	6484	-0.44	-0.49
F2	1735 3	10251.08	2197.98	6978.02	8210	9630	12455	14042	0.26	-1.42
F3	1735 3	12866.62	1342.60	11792.8 1	1222 8	12945	13577	15520	-0.37	0.65
F4	1735 3	26640.75	1897.63	26893.8 6	2504 0	26645	28418	31085	-0.10	-0.96
F5	1735 3	105045.0 2	13025.2 7	97020.0 0	9869 7	10453 0	11588 2	125561	-0.72	0.25
F6	1735 3	28008.91	2988.52	24819.7 8	2635 5	27650	29591	34778	0.16	0.38

Coins	Obs.	Mean	SD.	Min	Q1	Median	Q3	Max	Skewness	Kurtosis
F7	1735 3	1681.31	324.84	1469.98	1490	1578	1892	2360	0.58	-0.55
F8	1735 3	3299.28	429.66	2744.45	2984	3226	3585	4233	0.46	-0.47
F9	1735 3	3468.36	338.44	3300.22	3290	3499	3697	4071	-0.53	0.23
F10	1735 3	6691.76	647.40	7022.61	6112	6719	7302	7687	-0.24	-1.09
Volum e	Obs.	Mean	SD.	Min	Q1	Median	Q3	Max	Skewness	Kurtosis
V1	1735 3	42731.92	47756.6 2	16438.0 0	1635 1	27719	56831	345986 5	22.95	1520.6 1
V2	1735 3	21799.55	31252.3 8	16449.0 0	5722	10123	24700	860483	5.19	61.15
V3	1735 3	9980.58	22772.4 4	2757.00	866	2163	7535	526289	5.72	56.08
V4	1735 3	16056.23	34663.1 2	6089.00	2496	4941	12098	220604 5	17.54	937.17
V5	1735 3	13745.11	27658.4 2	8238.00	2263	4072	11808	586341	5.99	64.48
V6	1735 3	6910.54	9392.34	1470.00	1253	3768	8817	259667	4.39	48.73
V7	1289 2	12309.72	17474.6 1	43259.0 0	3022	6840	14817	650293	7.56	166.66
V8	1654 4	13745.22	56993.6 8	32285.0 0	60	186	1168	152348 2	8.67	112.62
V9	6091	8246.47	9238.54	21476.0 0	3172	5601	9985	194970	4.98	50.79
V10	1735 3	5223.96	22575.4 2	1154.00	310	690	1765	624965	10.09	145.28

Table 1: Summary of Cryptocurrency Hourly Prices From 1<sup>st</sup> June 2019 to 31<sup>st</sup> May 2021

Coins	Obs.	Mean	SD.	Min	Q1	Median	Q3	Max	Skewness	Kurtosis
C1	1735 3	0.00%	0.88%	-0.21%	-0.28%	0.01%	0.31%	14.81 %	-1.99	65.00
C2	1735 3	0.01%	1.07%	-0.91%	-0.38%	0.01%	0.42%	13.05 %	-2.03	42.95
C3	1735 3	-0.03%	3.20%	0.00%	0.00%	0.00%	0.00%	33.33 %	-2.43	56.40
C4	1735 3	0.01%	1.21%	0.00%	-0.44%	0.00%	0.47%	11.23 %	-1.89	41.34
C5	1735 3	-0.01%	1.79%	0.00%	0.00%	0.00%	0.00%	21.74 %	-0.75	16.70
C6	1735 3	0.01%	1.57%	-0.12%	-0.64%	0.00%	0.66%	35.21 %	-0.22	36.89
C7	1289 2	-4.15%	30.35%	0.00%	17.78 %	-0.01%	15.07 %	55.85 %	-0.86	0.91
C8	1654 4	0.01%	1.80%	0.00%	-0.41%	0.00%	0.41%	41.18 %	0.56	84.55
C9	6091	0.01%	2.11%	0.00%	-0.93%	0.04%	0.95%	32.04 %	-0.14	22.12
C10	1735 3	0.01%	2.05%	1.33%	-0.72%	0.00%	0.73%	26.21 %	-6.40	310.87
Factors	Obs.	Mean	SD.	Min	Q1	Median	Q3	Max	Skewness	Kurtosis

Coins	Obs.	Mean	SD.	Min	Q1	Media n	Q3	Max	Skewnes s	Kurtosi s
F1	1735 3	0.00%	0.27%	0.00%	0.00%	0.00%	0.00%	7.74%	-9.98	649.79
F2	1735 3	0.00%	0.31%	0.00%	0.00%	0.00%	0.00%	9.15%	-5.63	450.36
F3	1735 3	0.00%	0.28%	0.00%	0.00%	0.00%	0.00%	9.89%	-7.88	633.91
F4	1735 3	0.00%	0.22%	0.00%	0.00%	0.00%	0.00%	4.81%	-2.57	176.31
F5	1735 3	0.00%	0.38%	0.00%	0.00%	0.00%	0.00%	12.21 %	-13.24	763.57
F6	1735 3	0.00%	0.29%	0.00%	0.00%	0.00%	0.00%	12.93 %	2.91	634.03
F7	1735 3	0.00%	0.37%	0.00%	0.00%	0.00%	0.00%	8.59%	-9.76	539.34
F8	1735 3	0.00%	0.28%	0.00%	0.00%	0.00%	0.00%	11.98 %	3.05	562.14
F9	1735 3	0.00%	0.27%	0.00%	0.00%	0.00%	0.00%	8.46%	-10.22	686.04
F10	1735 3	0.00%	0.24%	0.00%	0.00%	0.00%	0.00%	10.87 %	5.28	502.48
Volum e	Obs.	Mean	SD.	Min	Q1	Media n	Q3	Max	Skewnes s	Kurtosi s
V1	1735 3	- 10.25%	217.27 %	-42.43%	25.85 %	-2.80%	17.74 %	99.76 %	-70.75	5492.1 9
V2	1735 3	- 10.43%	125.05 %	-62.85%	30.13 %	-2.88%	20.60 %	98.22 %	-70.98	6127.5 0
V3	1735 3	- 16.97%	201.78 %	-76.39%	42.20 %	-3.37%	27.08 %	98.94 %	-90.80	9838.7 6
V4	1735 3	- 12.92%	230.71 %	110.55 %	32.47 %	-2.18%	22.21 %	99.39 %	-80.96	7512.5 9
V5	1735 3	- 12.36%	173.00 %	-67.88%	31.83 %	-2.89%	21.17 %	98.87 %	-78.32	6961.7 9
V6	1735 3	- 15.95%	117.46 %	-63.33%	41.87 %	-3.57%	26.66 %	99.50 %	-40.49	2422.8 7
V7	1289 2	- 74.36%	300.88 %	0.00%	92.34 %	-0.93%	50.48 %	98.35 %	-15.13	460.24
V8	1654 4	- 27.01%	138.22 %	0.00%	49.18 %	-2.88%	31.35 %	97.97 %	-29.75	2030.2 4
V9	6091	- 17.65%	143.17 %	0.00%	42.45 %	-2.99%	27.52 %	99.07 %	-34.14	1515.6 9
V10	1735 3	- 21.46%	149.23 %	4.55%	50.19 %	-3.72%	30.25 %	99.71 %	-40.48	2307.2 0

Table 2: Summary of Cryptocurrencies Arithmetic Hourly Returns 1<sup>st</sup> June 2019 to 31<sup>st</sup> May 2021

Table 1 shows the summary statistics of the entire dataset while Table 2 shows the Arithmetic hourly returns. The total hourly observation is 17,353 samples except

C7, C8 and C9 which do not have the same number of samples. This happens because three coins entered Binance market later from other coins along with their volumes. Cryptocurrency's data are skewing to the right which means that training dataset and testing data set are having significant different of values. Range of volume of cryptocurrency increased over time by a significant value which mean that cryptocurrency market drew attention from investors over time. In term of Arithmetic return cryptocurrency has high relatively high hourly return especially C7. For factors like stock indices prices, hourly change does not change much since prices are in daily, so when converted into hourly they do not have a price change for a day. Kurtosis for all values in this research are positively high value except C7, C9, F1, F2, F4, F7, F8 and F10 which has negative value. To make a summary easier, plots between value and time of coins and factors are included in Appendix.

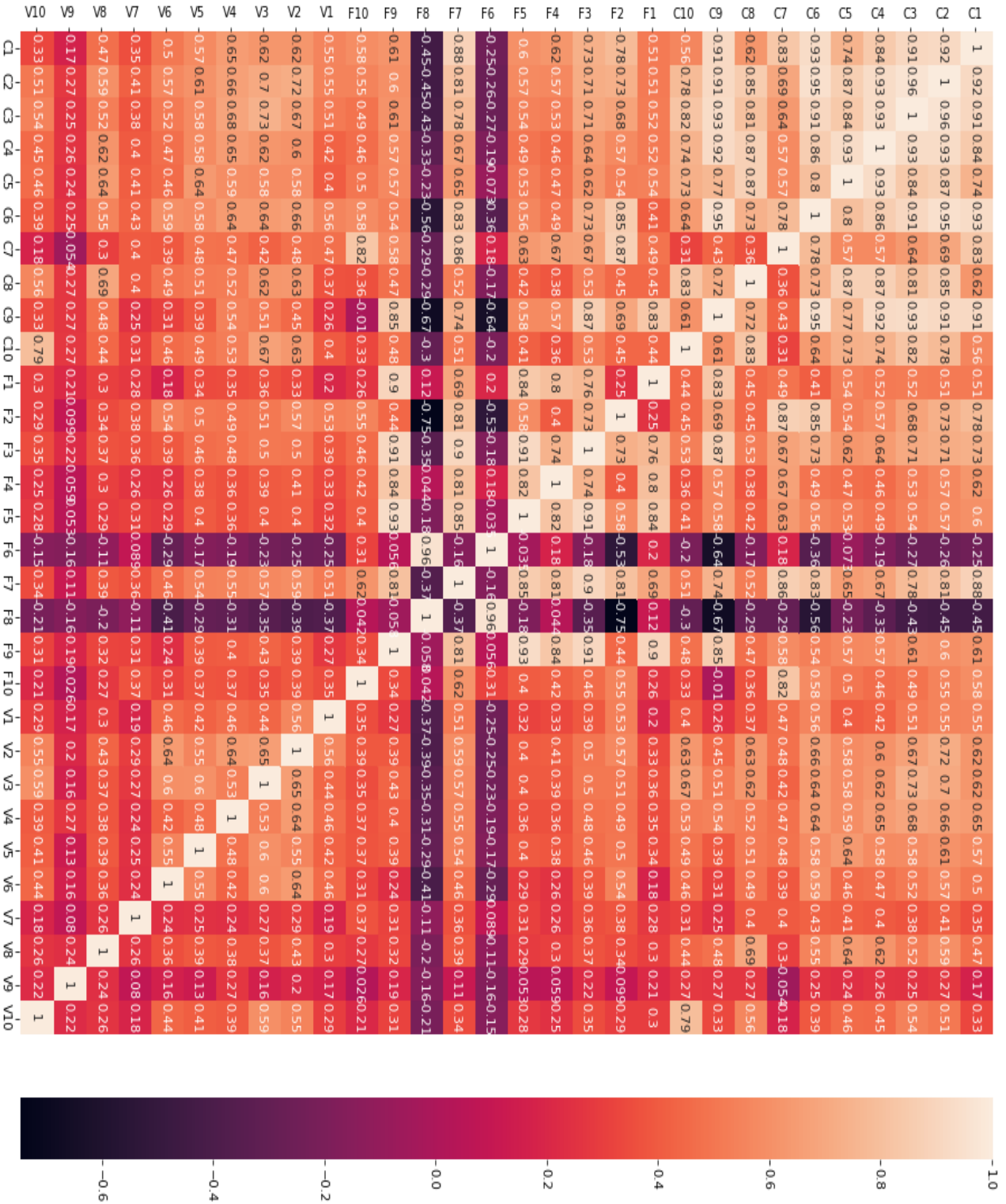


Table 3: Correlation Table of Dataset

From table 3, cryptocurrency coins from C2 to C10 has high correlation value with C1 which is Bitcoin, so if the trend of C1 changes, other coins values might have a change to change. For factors, all factors except F6 and F8 are having high correlation value with cryptocurrency coins which we will be predicting. F6 and F8 are not correlating with other data expect themselves because they're stock indices from the same country which is the United States of America. Volumes of each cryptocurrency coin are use as a feature in this research by using only the volume of each specific coin while predicting. The reason to not includes other volumes because each volume for example, C1 and V1 have high correlation between them. As a result, along with ten stock indices and other cryptocurrency that are not predicting, volume of specific cryptocurrency is added as a feature in this experiment.

### **Primary Result Evaluation**

After importing dataset into six machine learning algorithms, the predicted values of each coin and each algorithm are used for creating trading algorithm. Our trading algorithm is a simple signal algorithm. Both original and predicted dataset are used to find their return of each respective hour. If the predicted and actual sign of returns are similar, it means that the algorithm correctly predicts. An accuracy of models is calculated from the number of corrected signals divided by number of total predictions. Signals that are collected into four different trade timing which hourly, six hours, daily and weekly. There are four tables in this section which are divided for each trade timing.

		Accuracy hourly trade (%)									
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
M1	Train	51.02%	50.41%	84.88%	52.73%	72.95%	49.24%	66.90%	62.96%	79.20%	49.99%
	Test	50.15%	52.42%	50.03%	51.32%	55.61%	52.62%	48.03%	50.49%	49.27%	48.80%
M2	Train	50.84%	50.22%	75.47%	48.79%	51.09%	49.42%	50.26%	48.97%	65.05%	50.98%
	Test	50.74%	50.51%	47.34%	47.34%	48.40%	50.48%	47.90%	47.24%	27.74%	50.03%
M3	Train	49.68%	49.61%	65.71%	50.00%	51.09%	49.80%	50.57%	51.27%	54.93%	50.32%
	Test	51.66%	49.16%	51.89%	50.44%	48.40%	50.64%	49.43%	50.59%	59.12%	47.15%
M4	Train	50.85%	49.26%	67.96%	49.36%	48.19%	48.94%	64.92%	60.28%	71.53%	49.38%
	Test	51.20%	50.44%	51.89%	50.41%	48.21%	50.81%	46.75%	50.92%	43.80%	47.12%
M5	Train	50.49%	49.71%	76.01%	49.32%	55.04%	49.39%	53.37%	52.51%	60.73%	50.75%
	Test	49.62%	49.78%	53.21%	49.52%	54.23%	49.87%	47.99%	48.51%	49.05%	49.34%
M6	Train	50.65%	49.59%	50.15%	49.91%	50.53%	49.80%	50.51%	62.66%	76.00%	50.29%
	Test	49.39%	49.95%	40.21%	49.36%	48.70%	50.97%	49.17%	50.49%	49.64%	47.02%

Table 4: Accuracy of all models and coins for hourly trading

From table 4, C3, C9 and C5 are in first, second and third in overall accuracy of training set where all coins perform similarly in the testing set. There is no significant difference to determine which coin is doing worse than others in this trade since other coins beside C3, C5 and C9 are performing slightly around 50 percent in both train and test set.

In term of model performance, M1 performed significantly better than others in training set, but the testing set are similar to each other. M4 also has good outcome for training set especially for C3, C7 and C8 which are almost as good as M1.

Performances in training set and testing of C1, C2, C4 and C10 are not difference which are around 50 percent, but they are different for C3, C5, C6, C7, C8 and C9. For, C3, C5, C7, C8 and C9, accuracies dropped significantly from training set to testing especially M2 of C9 where an accuracy dropped from 65 to 27 percent. The reason why an accuracy dropped for these coins are that the training dataset were not as variate as the testing dataset. As a result, machine learning models cannot catch

trends for the testing set. C6 is the only coin that have testing accuracy better than training accuracy.

Ensemble algorithms which are M4, M5 and M6 are better than M2 and M5 in term of accuracy performance, but M1 outperformed them for training set. For testing set, there are not significant differences in accuracy to determine which model is better than another. This situation occurred because ensemble algorithms are using weights from M1, M2 and M3 to make a prediction. For an hourly trading, it can be determined that M1 is significantly better than M2 and M3 which means that M2 and M3 were dragging the accuracy of M1 down while performing in ensemble algorithms.

		Accuracy 6 hour trade (%)									
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
M1	Train	79.26%	51.40%	93.21%	60.66%	77.61%	60.91%	85.33%	85.12%	89.04%	57.41%
	Test	56.51%	62.44%	58.98%	55.35%	79.74%	51.73%	72.44%	73.91%	68.52%	50.25%
M2	Train	53.29%	49.42%	93.09%	53.05%	50.62%	52.02%	52.47%	54.26%	65.30%	52.67%
	Test	52.22%	55.35%	58.98%	55.35%	48.43%	49.26%	46.79%	50.00%	33.33%	50.25%
M3	Train	52.47%	50.08%	63.42%	51.48%	50.62%	50.53%	50.72%	50.88%	52.51%	52.02%
	Test	50.08%	47.94%	51.24%	50.58%	48.43%	50.91%	55.13%	48.37%	68.52%	51.40%
M4	Train	72.84%	59.51%	66.30%	57.65%	58.35%	58.77%	80.86%	78.35%	79.00%	56.71%
	Test	57.50%	60.63%	51.24%	50.58%	53.05%	54.53%	80.86%	65.22%	59.26%	51.57%
M5	Train	50.10%	48.58%	73.47%	49.84%	52.27%	49.57%	55.50%	56.18%	54.41%	52.34%
	Test	48.45%	50.56%	51.35%	50.56%	52.27%	51.09%	55.50%	50.22%	54.41%	49.84%
M6	Train	79.05%	48.58%	53.05%	58.56%	50.37%	50.58%	50.56%	84.57%	88.13%	51.98%
	Test	56.51%	49.42%	42.83%	47.12%	48.43%	52.72%	50.56%	73.37%	68.52%	50.08%

Table 5: Accuracy of all models and coins for six-hour trading

From table 5, C7, C8 and C9 are performing well in both training and testing for several model. C2, C4, C6 and C10 are not performing well in both train and test set. Furthermore, C2 and C5 has better testing accuracy than training accuracy where



testing set of those coins are better than 60 percent. The other coins are in 50 percent range for training and testing set.

In term of model performance, M1 performed significantly better than others in training set, but the testing set are similar to each other except for testing set of M1 for C2, C5, C7, C8 and C9. M4 and M6 also performed as good as M1 for C7, C8 and C9.

Performances in training set and testing of C1, C4 and C10 are not difference which are around 50 percent, but they are different for C2, C3, C5, C6, C7, C8 and C9. For, C3 and C9, accuracies dropped significantly from training set to testing especially M2 of C9 where an accuracy dropped from 65 to 33 percent. The reason why an accuracy dropped for these coins are that the training dataset were not as variate as the testing dataset. As a result, machine learning models cannot catch trends for the testing set. Several models of C2 and C5 are having better test accuracy than training set. C7 and C8 has similar training and testing set accuracy where some of them have good high accuracy in both training and testing set, for example, M4 of C7.

Ensemble algorithms which are M4, M5 and M6 are better than M2 and M5 in term of accuracy performance, but M1 outperformed them for training set. For testing set, there are not significant differences in accuracy to determine which model is better than another. This situation occurred because ensemble algorithms are using weights from M1, M2 and M3 to make a prediction. For a six-hour trading, it can be determined that M1 is significantly better than M2 and M3 which means that M2 and M3 were dragging the accuracy of M1 down while performing in ensemble algorithms.

		Accuracy daily trade (%)									
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
M1	Train	90.72%	52.27%	90.53%	64.77%	83.14%	71.78%	93.38%	93.13%	91.49%	62.31%
	Test	58.33%	72.73%	57.58%	59.85%	82.58%	53.79%	91.18%	77.50%	72.73%	53.03%
M2	Train	57.77%	47.73%	88.26%	52.27%	46.02%	55.30%	59.56%	61.88%	70.21%	56.44%
	Test	53.79%	60.61%	57.58%	59.85%	43.18%	54.55%	47.06%	57.50%	27.27%	53.03%
M3	Train	52.46%	51.52%	62.12%	50.19%	46.02%	48.30%	47.79%	50.00%	55.32%	52.46%
	Test	50.76%	46.21%	46.97%	50.76%	43.18%	50.00%	50.00%	52.50%	63.64%	49.24%
M4	Train	80.49%	67.23%	64.77%	59.28%	71.59%	61.74%	83.09%	80.00%	82.98%	61.17%
	Test	56.82%	64.39%	46.97%	50.76%	64.39%	57.58%	83.09%	82.93%	63.64%	49.24%
M5	Train	47.58%	48.18%	66.67%	48.79%	47.58%	46.97%	55.29%	61.00%	55.17%	51.52%
	Test	43.94%	55.45%	52.42%	48.79%	47.58%	46.06%	55.29%	58.00%	55.17%	46.06%
M6	Train	89.77%	48.18%	55.11%	61.36%	46.21%	48.48%	47.79%	92.50%	89.36%	52.46%
	Test	56.82%	46.97%	47.73%	44.70%	43.18%	53.03%	47.79%	77.50%	63.64%	52.27%

Table 6: Accuracy of all models and coins for daily trading

From table 6, C7, C8 and C9 are performing well in both training and testing for several model. C2, C4, C6 and C10 are not performing well in both train and test set. C2 is the only coin that contain a better testing accuracy in M1 while the other coins are in 50 percent range for training and testing set. The significant change in daily trade is that C1 has higher overall accuracy and some outstanding accuracy in training set.

In term of model performance, M1 and M4 performed significantly better than others in training set, but the testing set are similar to each other except for testing set of M1 for C2, C5, C7, C8 and C9. M4 and M6 also performed as good as M1 for C7, C8 and C9.

Performances in training set and testing of C4 and C10 are not difference which are around 50 percent, but they are different for C1, C2, C3, C5, C6, C7, C8 and C9. For, C1, C3, C8 and C9, accuracies dropped significantly from training set to testing especially M2 of C9 where an accuracy dropped from 70 to 27 percent. The

reason why an accuracy dropped for these coins are that the training dataset were not as variate as the testing dataset. As a result, machine learning models cannot catch trends for the testing set. Only M1 of C2 is having better test accuracy than training set. C7 and C8 has similar training and testing set accuracy where some of them have good high accuracy in both training and testing set, for example, M1 of C7.

Ensemble algorithms which are M4, M5 and M6 are better than M2 and M5 in term of accuracy performance, but M1 outperformed them for training set. For testing set, there are not significant differences in accuracy to determine which model is better than another. This situation occurred because ensemble algorithms are using weights from M1, M2 and M3 to make a prediction. For a daily trading, it can be determined that M1 is significantly better than M2 and M3 which means that M2 and M3 were dragging the accuracy of M1 down while performing in ensemble algorithms.

		Accuracy weekly trade (%)									
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
M1	Train	97.22%	61.11%	84.72%	69.44%	94.44%	79.17%	94.44%	95.45%	100.00%	77.78%
	Test	55.56%	72.22%	66.67%	66.67%	100.00%	44.44%	50.00%	80.00%		72.22%
M2	Train	58.33%	55.56%	79.17%	52.78%	43.06%	66.67%	61.11%	68.18%	100.00%	62.50%
	Test	66.67%	61.11%	66.67%	66.67%	55.56%	61.11%	50.00%	100.00%		72.22%
M3	Train	50.00%	55.56%	51.39%	40.28%	43.06%	44.44%	66.67%	68.18%	50.00%	51.39%
	Test	72.22%	50.00%	61.11%	55.56%	55.56%	38.89%	75.00%	40.00%		44.44%
M4	Train	75.00%	76.39%	54.17%	51.39%	83.33%	69.44%	83.33%	90.91%	83.33%	69.44%
	Test	55.56%	77.78%	61.11%	55.56%	83.33%	72.22%	83.33%	80.00%		44.44%
M5	Train	44.44%	48.89%	57.78%	48.89%	46.67%	44.44%	81.82%	53.85%	75.00%	44.44%
	Test	62.22%	57.78%	60.00%	57.78%	46.67%	62.22%	81.82%	53.85%		44.44%
M6	Train	95.83%	48.89%	55.56%	63.89%	44.44%	44.44%	61.11%	90.91%	83.33%	52.78%
	Test	55.56%	55.56%	38.89%	50.00%	61.11%	55.56%	61.11%	80.00%		50.00%

Table 7: Accuracy of all models and coins for weekly trading

From table 7, C7, C8 and C9 are performing well in both training and testing for several model. C2, C4, C6 and C10 are not performing well in both train and test set. C1, C2, C5, and C8 are coins that contain a better testing accuracy while the other coins are in 50 percent range for the testing set.

In term of model performances, M1 and M4 have outstanding performances compared to other models. More than half of coins has high accuracy in both training and testing set of M1 and M4.

Performances in training set and testing of C4 and C10 are not difference which are around 50 percent, but they are different for C1, C2, C3, C5, C6, C7, C8 and C9. For, C1, C3, C6 and C7, accuracies dropped significantly from training set to testing especially M6 of C3 where an accuracy dropped from 55 to 38 percent. The reason why an accuracy dropped for these coins are that the training dataset were not as variate as the testing dataset. As a result, machine learning models cannot catch trends for the testing set. Several models of C1, C2, C5 and C8 are having better test accuracy than training set. C7 and C8 has similar training and testing set accuracy where some of them have good high accuracy in both training and testing set, for example, M4 of C7.

Ensemble algorithms which are M4, M5 and M6 are better than M2 and M5 in term of accuracy performance, but M1 outperformed them for training set. For testing set, there are not significant differences in accuracy to determine which model is better than another. This situation occurred because ensemble algorithms are using weights from M1, M2 and M3 to make a prediction. For a weekly trading, it can be determined that M1 is significantly better than M2 and M3 which means that M2 and

M3 were dragging the accuracy of M1 down while performing in ensemble algorithms.

In conclusion, as trading time increase, accuracies are increasing especially in training set. In an hourly trading, there was not a model that significantly better than others in term of overall accuracy, because the testing outcome from the models are similar to each other's, but for six-hour, daily and weekly trading, significant difference in testing set outcome were spotted. In our literature review, there were researchers that have predicted value of cryptocurrency coins for hourly, daily and weekly. The results show that daily and weekly are yielding better performances than hourly prediction because the hourly prediction is too extreme for prediction.

Sharpe Ratio is our second measurement for this project. The ratio is calculated by dividing an expected percent return by the standard deviation of each coin and model individually. These upcoming tables will also be divided into 4 different signal timing which are hourly, 6-hour, daily and weekly. To make our evaluation more impactful for investor, we decided to calculate the Variance, percent profit from trading, Value at Risk and Tailed Value at Risk. To make our four tables shorten, we decided to split each table into train and test set.

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C1	0.00%	7E-05	0.81%	0.00	40.43%	-0.83%	-1.39%
M2	C1	-0.09%	7E-05	0.81%	-0.11	40.09%	-1.03%	-1.93%
M3	C1	-0.09%	7E-05	0.81%	-0.12	39.22%	-1.10%	-1.88%
M4	C1	-0.01%	7E-05	0.81%	-0.02	40.26%	-0.86%	-1.46%
M5	C1	-0.08%	7E-05	0.85%	-0.10	39.86%	-0.97%	-1.86%
M6	C1	0.10%	7E-05	0.81%	0.12	51.20%	-0.74%	-1.29%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C2	-0.01%	2E-05	0.39%	-0.02	93.98%	-0.16%	-0.89%
M2	C2	-0.09%	1E-04	0.99%	-0.09	42.72%	-1.40%	-2.37%
M3	C2	-0.08%	1E-04	0.99%	-0.08	42.91%	-1.41%	-2.31%
M4	C2	-0.09%	1E-04	0.99%	-0.09	42.28%	-1.38%	-2.30%
M5	C2	-0.09%	9E-05	0.97%	-0.09	42.11%	-1.31%	-2.23%
M6	C2	-0.08%	1E-04	0.99%	-0.08	42.87%	-1.41%	-2.31%
M1	C3	-0.11%	9E-05	0.96%	-0.11	96.17%	0.00%	-2.98%
M2	C3	-0.01%	9E-08	0.03%	-0.20	95.71%	0.00%	-0.14%
M3	C3	-0.08%	8E-04	2.81%	-0.03	36.90%	-0.10%	-9.11%
M4	C3	-0.13%	8E-04	2.76%	-0.05	41.29%	-0.10%	-9.11%
M5	C3	-0.08%	8E-04	2.90%	-0.03	53.93%	-0.10%	-19.19%
M6	C3	-0.12%	1E-03	3.46%	-0.03	3.43%	-0.10%	-10.01%
M1	C4	-0.10%	8E-08	0.03%	-3.41	3.95%	-0.13%	-0.15%
M2	C4	-0.10%	8E-08	0.03%	-3.41	3.95%	-0.13%	-0.15%
M3	C4	-0.10%	1E-04	1.01%	-0.10	43.45%	-1.44%	-2.38%
M4	C4	-0.11%	1E-04	1.01%	-0.11	42.80%	-1.43%	-2.35%
M5	C4	-0.13%	1E-04	1.03%	-0.12	42.08%	-1.40%	-2.53%
M6	C4	-0.12%	1E-04	1.01%	-0.11	42.92%	-1.46%	-2.47%
M1	C5	-0.16%	2E-04	1.50%	-0.10	57.37%	-3.55%	-4.52%
M2	C5	-0.10%	3E-04	1.70%	-0.06	8.76%	-3.55%	-4.85%
M3	C5	-0.10%	3E-04	1.70%	-0.06	8.76%	-3.55%	-4.85%
M4	C5	-0.22%	3E-04	1.70%	-0.13	6.24%	-3.95%	-4.85%
M5	C5	-0.11%	2E-04	1.47%	-0.08	13.17%	-0.71%	-4.67%
M6	C5	-0.11%	3E-04	1.71%	-0.06	7.77%	-3.67%	-4.97%
M1	C6	-0.01%	6E-05	0.80%	-0.01	88.22%	-0.90%	-1.95%
M2	C6	-0.09%	2E-04	1.41%	-0.07	44.79%	-2.03%	-3.31%
M3	C6	-0.09%	2E-04	1.41%	-0.07	45.28%	-2.07%	-3.35%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M4	C6	-0.09%	2E-04	1.41%	-0.06	44.45%	-2.00%	-3.26%
M5	C6	-0.10%	2E-04	1.36%	-0.08	44.36%	-1.97%	-3.18%
M6	C6	-0.09%	2E-04	1.41%	-0.07	45.26%	-2.07%	-3.35%
M1	C7	0.67%	3E-04	1.71%	0.39	64.53%	-1.63%	-2.40%
M2	C7	-0.05%	4E-04	1.88%	-0.03	46.83%	-2.89%	-4.06%
M3	C7	-0.10%	4E-04	1.88%	-0.05	47.46%	-2.99%	-4.42%
M4	C7	0.59%	3E-04	1.75%	0.34	61.91%	-1.80%	-2.58%
M5	C7	-0.05%	4E-04	2.09%	-0.02	49.41%	-3.32%	-4.70%
M6	C7	-0.10%	4E-04	1.88%	-0.05	47.40%	-2.99%	-4.42%
M1	C8	0.44%	2E-04	1.54%	0.28	61.19%	-1.63%	-2.24%
M2	C8	-0.08%	3E-04	1.64%	-0.05	45.14%	-2.54%	-3.60%
M3	C8	-0.06%	3E-04	1.64%	-0.04	47.63%	-2.50%	-3.81%
M4	C8	0.37%	2E-04	1.58%	0.23	56.83%	-1.75%	-2.39%
M5	C8	-0.04%	3E-04	1.63%	-0.02	47.12%	-2.53%	-3.58%
M6	C8	0.43%	2E-04	1.56%	0.28	59.05%	-1.65%	-2.27%
M1	C9	1.69%	1E-03	3.60%	0.47	62.37%	-0.10%	-2.97%
M2	C9	0.02%	2E-03	4.02%	0.01	23.65%	-5.10%	-9.95%
M3	C9	-0.10%	1E-03	3.83%	-0.02	31.42%	-5.13%	-9.56%
M4	C9	1.63%	1E-03	3.63%	0.45	39.91%	-1.80%	-3.01%
M5	C9	0.14%	2E-03	4.10%	0.04	18.57%	-5.57%	-10.47%
M6	C9	1.66%	1E-03	3.62%	0.46	40.46%	-0.10%	-2.97%
M1	C10	-0.18%	3E-04	1.82%	-0.10	26.64%	-2.94%	-4.63%
M2	C10	-0.10%	3E-08	0.02%	-5.86	2.83%	-0.10%	-0.10%
M3	C10	-0.09%	3E-04	1.77%	-0.05	45.96%	-2.17%	-3.70%
M4	C10	-0.10%	3E-04	1.77%	-0.06	45.50%	-2.12%	-3.65%
M5	C10	-0.10%	4E-04	1.94%	-0.05	45.92%	-2.15%	-3.92%
M6	C10	-0.10%	4E-04	1.94%	-0.05	45.92%	-2.15%	-3.92%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	Average	0.22%	3E-04	1.32%	-0.26	59.49%	-1.19%	-2.41%
M2	Average	-0.07%	3E-04	1.25%	-0.99	35.45%	-1.88%	-3.05%
M3	Average	-0.09%	4E-04	1.79%	-0.06	38.90%	-2.25%	-4.54%
M4	Average	0.19%	4E-04	1.74%	0.05	42.15%	-1.72%	-3.50%
M5	Average	-0.06%	4E-04	1.83%	-0.06	39.65%	-2.00%	-5.63%
M6	Average	0.15%	4E-04	1.84%	0.04	38.63%	-1.63%	-3.80%

*Table 8: Model performance of training set for hourly trade in business term*

From table 8, C9, C7 and C8 produced the highest return in respective order and have positive return where other coins are producing negative return slightly below zero percent. Furthermore, when adding risk into consideration or by observing Sharpe Ratio which is a relation between return and risk of coins, C7, C9, and C8 have the highest Sharpe Ratio in respective order.

In term of model performance, M1 is the best while trading C2, C6, C7, C8 and C9. M3 is the best while trading C3, C4, C5 and C10, and M6 is the best while trading C1. On the other hand, if we consider an average performance of each model, M4 is the best model while trading all cryptocurrency coins.

Ensemble algorithms are performing better than base models in term of overall performance. During this trade, M1 and M3 which are the base models are among the best while trading a specific coin. The aim of ensemble algorithm is to bring out upsides of each base model to make the better model, but it can be worse if the base



models has only outstanding model. As a result, from having 2 good base models and one bad, ensemble algorithms have better overall performances than base models.

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C1	-0.04%	4E-05	0.66%	-0.06	87.59%	-0.85%	-1.89%
M2	C1	-0.09%	1E-04	1.07%	-0.08	45.49%	-1.74%	-2.52%
M3	C1	-0.09%	1E-04	1.07%	-0.09	45.75%	-1.70%	-2.68%
M4	C1	-0.09%	1E-04	1.07%	-0.09	45.46%	-1.67%	-2.68%
M5	C1	-0.10%	8E-05	0.89%	-0.11	41.68%	-1.42%	-2.19%
M6	C1	-0.11%	1E-04	1.07%	-0.10	43.52%	-1.62%	-2.70%
M1	C2	-0.01%	2E-05	0.48%	-0.03	91.94%	-0.55%	-1.38%
M2	C2	-0.08%	2E-04	1.39%	-0.06	47.79%	-2.20%	-3.31%
M3	C2	-0.10%	2E-04	1.39%	-0.07	45.06%	-2.18%	-3.27%
M4	C2	-0.07%	2E-04	1.39%	-0.05	46.91%	-2.17%	-3.23%
M5	C2	-0.10%	1E-04	1.19%	-0.08	45.54%	-1.88%	-3.06%
M6	C2	-0.11%	2E-04	1.39%	-0.08	44.83%	-2.20%	-3.36%
M1	C3	-0.11%	2E-04	1.54%	-0.07	51.11%	-2.36%	-3.88%
M2	C3	-0.14%	4E-04	2.03%	-0.07	34.39%	-3.03%	-4.54%
M3	C3	-0.07%	4E-04	2.03%	-0.03	35.32%	-3.08%	-4.67%
M4	C3	-0.07%	4E-04	2.03%	-0.03	35.32%	-3.08%	-4.67%
M5	C3	-0.09%	5E-04	2.32%	-0.04	23.03%	-3.31%	-6.41%
M6	C3	-0.17%	4E-04	2.02%	-0.08	33.71%	-3.23%	-5.08%
M1	C4	-0.10%	3E-04	1.83%	-0.05	44.33%	-2.67%	-3.89%
M2	C4	-0.10%	3E-04	1.83%	-0.05	44.33%	-2.67%	-3.89%
M3	C4	-0.11%	4E-04	1.94%	-0.05	46.91%	-2.96%	-4.76%
M4	C4	-0.11%	4E-04	1.94%	-0.05	46.91%	-2.96%	-4.76%
M5	C4	-0.10%	2E-04	1.44%	-0.07	44.62%	-2.08%	-3.50%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
						%		
M6	C4	-0.14%	4E-04	1.94%	-0.07	46.54%	-3.09%	-4.77%
M1	C5	-0.16%	2E-04	1.50%	-0.10	57.37%	-3.55%	-4.52%
M2	C5	-0.10%	3E-04	1.70%	-0.06	8.76%	-3.55%	-4.85%
M3	C5	-0.10%	3E-04	1.70%	-0.06	8.76%	-3.55%	-4.85%
M4	C5	-0.22%	3E-04	1.70%	-0.13	6.24%	-3.95%	-4.85%
M5	C5	-0.11%	2E-04	1.47%	-0.08	13.17%	-0.71%	-4.67%
M6	C5	-0.11%	3E-04	1.71%	-0.06	7.77%	-3.67%	-4.97%
M1	C6	0.00%	1E-05	0.38%	0.00	98.26%	0.00%	-1.41%
M2	C6	-0.10%	3E-04	1.86%	-0.06	47.47%	-2.80%	-4.17%
M3	C6	-0.11%	3E-04	1.86%	-0.06	47.33%	-2.74%	-4.29%
M4	C6	-0.09%	3E-04	1.86%	-0.05	47.70%	-2.72%	-4.20%
M5	C6	-0.11%	3E-04	1.65%	-0.07	46.22%	-2.59%	-3.92%
M6	C6	-0.06%	3E-04	1.86%	-0.03	48.03%	-2.72%	-3.98%
M1	C7	-0.07%	7E-04	2.57%	-0.03	47.32%	-3.77%	-5.55%
M2	C7	-0.13%	7E-04	2.59%	-0.05	45.03%	-3.86%	-5.44%
M3	C7	0.02%	7E-04	2.59%	0.01	46.81%	-3.73%	-5.14%
M4	C7	0.59%	3E-04	1.75%	0.34	61.91%	-1.80%	-2.58%
M5	C7	-0.05%	4E-04	2.09%	-0.02	49.41%	-3.32%	-4.70%
M6	C7	-0.10%	4E-04	1.88%	-0.05	47.40%	-2.99%	-4.42%
M1	C8	0.02%	8E-04	2.84%	0.01	48.38%	-3.54%	-5.54%
M2	C8	-0.09%	8E-04	2.85%	-0.03	45.13%	-3.65%	-5.35%
M3	C8	-0.02%	8E-04	2.84%	-0.01	48.48%	-3.81%	-5.61%
M4	C8	-0.01%	8E-04	2.84%	0.00	48.38%	-3.57%	-5.70%
M5	C8	-0.12%	5E-04	2.21%	-0.05	45.41%	-3.11%	-4.74%
M6	C8	0.02%	8E-04	2.84%	0.01	47.73%	-3.54%	-5.54%
M1	C9	-0.18%	7E-04	2.57%	-0.07	27.47%	-3.98%	-6.88%
M2	C9	-0.03%	7E-04	2.65%	-0.01	23.08%	-3.33%	-5.68%
M3	C9	-0.31%	7E-04	2.64%	-0.12	21.98%	-4.90%	-7.12%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M4	C9	-0.18%	7E-04	2.65%	-0.07	21.98%	-3.43%	-6.20%
M5	C9	0.14%	2E-03	4.10%	0.04	18.57%	-5.57%	-10.47%
M6	C9	-0.24%	7E-04	2.65%	-0.09	22.71%	-5.14%	-6.95%
M1	C10	-0.07%	2E-04	1.43%	-0.05	77.13%	-2.40%	-3.79%
M2	C10	0.04%	3E-04	1.64%	0.02	100.00%	0.00%	0.00%
M3	C10	-0.20%	9E-04	2.98%	-0.07	45.59%	-4.88%	-7.05%
M4	C10	-0.20%	9E-04	2.98%	-0.07	45.62%	-4.88%	-7.05%
M5	C10	-0.10%	5E-04	2.19%	-0.05	46.39%	-3.40%	-5.34%
M6	C10	-0.20%	9E-04	2.98%	-0.07	45.39%	-4.89%	-7.05%
M1 Average		-0.07%	3E-04	1.58%	-0.05	63.09%	-2.37%	-3.87%
M2 Average		-0.08%	4E-04	1.96%	-0.04	44.15%	-2.68%	-3.98%
M3 Average		-0.11%	5E-04	2.10%	-0.06	39.20%	-3.35%	-4.94%
M4 Average		-0.04%	4E-04	2.02%	-0.02	40.64%	-3.02%	-4.59%
M5 Average		-0.07%	5E-04	1.95%	-0.05	37.40%	-2.74%	-4.90%
M6 Average		-0.12%	4E-04	2.03%	-0.06	38.76%	-3.31%	-4.88%

Table 9: Model performance of testing set for hourly trade in business term

From table 9, C7, C8 and C2 produced the highest return in respective order and have positive return where other coins are producing negative return slightly below zero percent. Furthermore, when adding risk into consideration or by observing Sharpe Ratio which is a relation between return and risk of coins, C7, C8, and C6 have the highest Sharpe Ratio in respective order.

In term of model performance, M1 is the best while trading C1, C2, C4, C6 and C8. M4 is the best while trading C3 and C7, and M2, M5 and M9 are best at

trading C10, C9 and C5 respectively. On the other hand, if we consider an average performance of each model, M4 is the best model while trading all cryptocurrency coins.

Ensemble algorithms are performing worse than base models in term of overall performance. During this trade, M1 which is the base models is among the best while trading a specific coin. The aim of ensemble algorithm is to bring out upsides of each base model to make the better model, but it can be worse if the base models has only outstanding model. As a result, from having only a good base models and two bad, ensemble algorithms have worse overall performances than base models.

Model	Coins	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C1	0.73%	0.00	1.55%	0.47	74.34%	-0.62%	-1.06%
M2	C1	0.01%	0.00	1.76%	0.00	48.89%	-2.09%	-3.97%
M3	C1	-0.07%	0.00	1.76%	-0.04	47.36%	-2.28%	-4.10%
M4	C1	0.58%	0.00	1.62%	0.36	67.87%	-0.94%	-1.83%
M5	C1	-0.11%	0.00	1.87%	-0.06	45.95%	-2.20%	-4.24%
M6	C1	0.73%	0.00	1.55%	0.47	74.09%	-0.62%	-1.07%
M1	C2	0.20%	0.00	2.15%	0.09	53.67%	-2.52%	-4.45%
M2	C2	0.07%	0.00	2.16%	0.03	49.09%	-2.76%	-4.62%
M3	C2	-0.02%	0.00	2.17%	-0.01	47.78%	-2.94%	-4.76%
M4	C2	0.21%	0.00	2.15%	0.10	52.59%	-2.43%	-4.21%
M5	C2	0.00%	0.00	2.19%	0.00	48.52%	-2.82%	-4.65%
M6	C2	-0.02%	0.00	2.17%	-0.01	47.78%	-2.94%	-4.76%
M1	C3	0.28%	0.00	4.96%	0.06	8.48%	-3.66%	-11.41%
M2	C3	0.10%	0.00	4.97%	0.02	7.41%	-6.34%	-12.20%
M3	C3	-0.17%	0.00	4.98%	-0.03	7.00%	-6.77%	-14.29%
M4	C3	0.03%	0.00	4.98%	0.01	8.11%	-6.06%	-13.40%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
						%		
M5	C3	-0.16%	0.00	5.44%	-0.03	3.23%	-0.10%	-20.23%
M6	C3	0.08%	0.00	4.98%	0.02	8.48%	-5.66%	-13.53%
M1	C4	0.27%	0.00	2.14%	0.13	57.37%	-2.56%	-4.57%
M2	C4	-0.05%	0.00	2.17%	-0.02	49.67%	-3.11%	-5.28%
M3	C4	-0.11%	0.00	2.17%	-0.05	48.39%	-3.27%	-5.17%
M4	C4	0.21%	0.00	2.15%	0.10	54.65%	-2.61%	-4.43%
M5	C4	-0.11%	0.00	2.23%	-0.05	46.28%	-2.90%	-5.07%
M6	C4	0.23%	0.00	2.15%	0.11	54.98%	-2.56%	-4.48%
M1	C5	0.80%	0.00	2.86%	0.28	23.76%	-3.54%	-4.62%
M2	C5	-0.11%	0.00	3.00%	-0.04	14.37%	-4.86%	-6.90%
M3	C5	-0.11%	0.00	3.00%	-0.04	14.37%	-4.86%	-6.90%
M4	C5	0.71%	0.00	2.88%	0.25	23.02%	-3.55%	-5.08%
M5	C5	-0.11%	0.00	2.56%	-0.04	10.80%	-4.86%	-6.27%
M6	C5	-0.10%	0.00	3.00%	-0.03	14.50%	-4.86%	-6.94%
M1	C6	0.71%	0.00	2.83%	0.25	58.57%	-3.05%	-4.60%
M2	C6	0.13%	0.00	2.94%	0.04	49.67%	-3.96%	-6.12%
M3	C6	-0.05%	0.00	2.94%	-0.02	48.23%	-4.37%	-6.47%
M4	C6	0.53%	0.00	2.88%	0.18	56.47%	-3.40%	-5.27%
M5	C6	-0.15%	0.00	2.87%	-0.05	47.53%	-4.26%	-6.79%
M6	C6	-0.05%	0.00	2.94%	-0.02	48.27%	-4.37%	-6.47%
M1	C7	2.40%	0.00	2.97%	0.81	84.05%	-1.25%	-2.17%
M2	C7	0.54%	0.00	3.83%	0.14	51.04%	-4.52%	-6.94%
M3	C7	0.01%	0.00	3.88%	0.00	49.44%	-6.08%	-8.82%
M4	C7	2.27%	0.00	3.07%	0.74	79.59%	-1.63%	-2.60%
M5	C7	0.67%	0.00	4.25%	0.16	53.71%	-5.26%	-7.66%
M6	C7	0.01%	0.00	3.88%	0.00	49.44%	-6.08%	-8.82%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C8	2.03%	0.00	2.61%	0.78	83.49%	-1.14%	-1.90%
M2	C8	0.36%	0.00	3.34%	0.11	51.69%	-4.33%	-5.93%
M3	C8	0.03%	0.00	3.37%	0.01	48.85%	-4.66%	-7.64%
M4	C8	1.71%	0.00	2.84%	0.60	76.59%	-2.14%	-3.22%
M5	C8	0.42%	0.00	3.52%	0.12	52.28%	-4.41%	-6.25%
M6	C8	2.01%	0.00	2.63%	0.76	82.95%	-1.18%	-2.01%
M1	C9	4.52%	0.00	6.50%	0.70	59.36%	-0.10%	-2.57%
M2	C9	1.52%	0.01	7.82%	0.19	36.53%	-8.80%	-14.58%
M3	C9	0.71%	0.01	7.94%	0.09	31.51%	-11.48%	-14.98%
M4	C9	4.28%	0.00	6.67%	0.64	55.25%	-1.85%	-3.31%
M5	C9	1.31%	0.01	8.20%	0.16	25.74%	-10.41%	-13.29%
M6	C9	4.52%	0.00	6.50%	0.70	59.36%	-0.10%	-2.57%
M1	C10	0.41%	0.00	3.67%	0.11	47.28%	-4.02%	-6.38%
M2	C10	0.01%	0.00	3.70%	0.00	45.30%	-4.47%	-8.77%
M3	C10	-0.09%	0.00	3.70%	-0.03	47.61%	-5.23%	-9.75%
M4	C10	0.01%	0.00	0.07%	0.20	60.26%	-0.06%	-0.13%
M5	C10	-0.12%	0.00	3.51%	-0.03	46.08%	-4.75%	-8.45%
M6	C10	0.00%	0.00	0.07%	-0.01	49.51%	-0.09%	-0.16%
M1 Average		1.24%	0.00	3.22%	0.3671	55.04%	-2.25%	-4.37%
M2 Average		0.26%	0.00	3.57%	0.0487	40.37%	-4.53%	-7.53%
M3 Average		0.01%	0.00	3.59%	-0.0108	39.05%	-5.19%	-8.29%
M4 Average		1.06%	0.00	2.93%	0.3178	53.44%	-2.47%	-4.35%
M5 Average		0.16%	0.00	3.66%	0.0174	38.01%	-4.20%	-8.29%
M6 Average		0.74%	0.00	2.99%	0.1985	48.94%	-2.85%	-5.08%

Table 10: Model performance of training set for six-hour trade in business term

From table 10, C9, C8 and C7 produced the highest return in respective order and have positive return where other coins are producing positive return slightly above zero percent. Furthermore, when adding risk into consideration or by observing Sharpe Ratio which is a relation between return and risk of coins, C9, C8, and C7 have the highest Sharpe Ratio in respective order.

In term of model performance, M1 is the best while trading C1, C3, C4, C5, C6, C7, C8 and C9. M4 is the best while trading C2 and C10. Furthermore, if we consider an average performance of each model, M1 is the best model while trading all cryptocurrency coins.

Ensemble algorithms are performing worse than base models in term of overall performance. During this trade, M1 which is the base models is among the best while trading a specific coin. M4 was close having as good as M1 for overall performance since M4 is the ensemble algorithm with the highest weight of M1. The aim of ensemble algorithm is to bring out upsides of each base model to make the better model, but it can be worse if the base models has only outstanding model. As a result, from having only a good base models and two bad, ensemble algorithms have worse overall performances than base models.

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C1	0.32%	0.00	2.28%	0.14	55.35%	-2.94%	-4.61%
M2	C1	0.09%	0.00	2.31%	0.04	50.91%	-3.52%	-5.04%
M3	C1	0.01%	0.00	2.32%	0.00	47.94%	-3.26%	-5.12%
M4	C1	0.44%	0.00	2.26%	0.20	55.68%	-2.67%	-3.96%
M5	C1	-0.07%	0.00	1.91%	-0.04	45.23%	-3.18%	-4.39%
M6	C1	0.29%	0.00	2.29%	0.12	53.38%	-3.11%	-4.40%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C2	0.43%	0.00	3.04%	0.14	59.31%	-4.47%	-7.19%
M2	C2	0.19%	0.00	3.07%	0.06	52.22%	-4.88%	-6.86%
M3	C2	-0.13%	0.00	3.09%	-0.04	45.47%	-5.05%	-7.15%
M4	C2	0.53%	0.00	3.02%	0.17	57.99%	-3.95%	-6.55%
M5	C2	-0.07%	0.00	2.42%	-0.03	47.86%	-3.95%	-5.72%
M6	C2	-0.12%	0.00	3.09%	-0.04	46.79%	-5.29%	-7.44%
M1	C3	0.28%	0.00	4.96%	0.06	8.48%	-5.26%	-7.68%
M2	C3	0.10%	0.00	4.97%	0.02	7.41%	-5.26%	-7.68%
M3	C3	-0.17%	0.00	4.98%	-0.03	7.00%	-5.97%	-9.32%
M4	C3	0.03%	0.00	4.98%	0.01	8.11%	-5.97%	-9.32%
M5	C3	-0.16%	0.00	5.44%	-0.03	3.23%	-7.25%	-10.44%
M6	C3	0.08%	0.00	4.98%	0.02	8.48%	-6.53%	-9.66%
M1	C4	0.36%	0.00	4.49%	0.08	53.87%	-5.70%	-9.92%
M2	C4	0.36%	0.00	4.49%	0.08	53.87%	-5.70%	-9.92%
M3	C4	-0.01%	0.00	4.52%	0.00	49.26%	-6.65%	-10.74%
M4	C4	-0.01%	0.00	4.52%	0.00	49.26%	-6.65%	-10.74%
M5	C4	-0.15%	0.00	2.91%	-0.05	48.19%	-4.84%	-7.07%
M6	C4	-0.43%	0.00	4.51%	-0.09	45.47%	-7.64%	-12.78%
M1	C5	0.46%	0.00	2.13%	0.21	16.31%	-0.10%	-4.48%
M2	C5	-0.09%	0.00	2.20%	-0.04	10.71%	-4.10%	-6.01%
M3	C5	-0.09%	0.00	2.20%	-0.04	10.71%	-4.10%	-6.01%
M4	C5	0.47%	0.00	2.13%	0.22	16.80%	-0.10%	-4.83%
M5	C5	-0.11%	0.00	2.56%	-0.04	10.80%	-4.86%	-6.27%
M6	C5	-0.09%	0.00	2.20%	-0.04	10.71%	-4.10%	-6.01%
M1	C6	0.11%	0.00	3.77%	0.03	51.07%	-5.91%	-8.59%
M2	C6	0.13%	0.00	3.78%	0.03	49.09%	-5.55%	-8.47%
M3	C6	-0.13%	0.00	3.78%	-0.03	50.08%	-6.63%	-9.50%



Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M4	C6	0.34%	0.00	3.76%	0.09	54.04%	-5.53%	-7.94%
M5	C6	-0.03%	0.00	3.44%	-0.01	49.70%	-5.14%	-7.96%
M6	C6	0.00%	0.00	3.78%	0.00	52.22%	-6.73%	-9.50%
M1	C7	2.31%	0.00	5.49%	0.42	71.15%	-3.39%	-9.29%
M2	C7	-0.01%	0.00	6.00%	0.00	46.15%	-8.48%	-14.50%
M3	C7	0.66%	0.00	5.95%	0.11	54.49%	-6.76%	-9.99%
M4	C7	2.27%	0.00	3.07%	0.74	79.59%	-1.63%	-2.60%
M5	C7	0.67%	0.00	4.25%	0.16	53.71%	-5.26%	-7.66%
M6	C7	0.01%	0.00	3.88%	0.00	49.44%	-6.08%	-8.82%
M1	C8	1.58%	0.00	5.18%	0.31	72.28%	-6.63%	-11.21%
M2	C8	0.22%	0.00	5.44%	0.04	48.91%	-6.98%	-11.72%
M3	C8	0.51%	0.00	5.41%	0.09	47.83%	-6.59%	-9.12%
M4	C8	0.96%	0.00	5.34%	0.18	64.13%	-6.98%	-11.25%
M5	C8	0.09%	0.00	3.77%	0.02	48.48%	-5.88%	-7.87%
M6	C8	1.53%	0.00	5.19%	0.30	71.74%	-6.63%	-11.21%
M1	C9	2.25%	0.00	3.72%	0.60	53.70%	-3.32%	-4.57%
M2	C9	0.01%	0.00	4.41%	0.00	33.33%	-9.71%	-11.92%
M3	C9	-0.12%	0.00	4.41%	-0.03	38.89%	-9.71%	-11.92%
M4	C9	2.05%	0.00	3.83%	0.54	50.00%	-3.18%	-4.57%
M5	C9	1.31%	0.01	8.20%	0.16	25.74%	-10.41%	-13.29%
M6	C9	2.25%	0.00	3.72%	0.60	53.70%	-3.32%	-4.57%
M1	C10	0.79%	0.00	6.75%	0.12	49.26%	-8.48%	-11.83%
M2	C10	0.79%	0.00	6.75%	0.12	49.26%	-8.48%	-11.83%
M3	C10	-0.16%	0.00	6.33%	-0.03	51.40%	-8.87%	-15.54%
M4	C10	-0.42%	0.00	6.81%	-0.06	50.41%	-10.92%	-17.24%
M5	C10	-0.18%	0.00	4.65%	-0.04	48.19%	-7.59%	-11.67%
M6	C10	-0.53%	0.00	6.80%	-0.08	48.93%	-10.92%	-17.13%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	Average	0.89%	0.00	4.18%	0.21	49.08%	-4.62%	-7.94%
M2	Average	0.18%	0.00	4.34%	0.03	40.19%	-6.27%	-9.40%
M3	Average	0.04%	0.00	4.30%	0.00	40.31%	-6.36%	-9.44%
M4	Average	0.67%	0.00	3.97%	0.21	48.60%	-4.76%	-7.90%
M5	Average	0.13%	0.00	3.95%	0.01	38.11%	-5.84%	-8.23%
M6	Average	0.30%	0.00	4.04%	0.08	44.09%	-6.03%	-9.15%

*Table 11 : Model performance of testing set for six-hour trade in business term*

From table 11, C9, C7 and C8 produced the highest return in respective order and have positive return where other coins are producing positive return slightly above zero percent. Furthermore, when adding risk into consideration or by observing Sharpe Ratio which is a relation between return and risk of coins, C9, C7, and C8 have the highest Sharpe Ratio in respective order.

In term of model performance, M1 is the best while trading C3 and C8. M4 is the best while trading C1, C2, C5, C6 and C7. M1 has tied with M2 while trading C4 and C10, and with M6 while predicting C9. Furthermore, if we consider an average performance of each model, M1 and M4 are tied for the best model while trading all cryptocurrency coins.

Ensemble algorithms are performing slightly better than base models in term of overall performance. During this trade, M1 and M3 which are the base models are among the best while trading a specific coin. The aim of ensemble algorithm is to bring out upsides of each base model to make the better model, but it can be worse if the base models has only outstanding model. As a result, from having 2 good base

models and one bad, ensemble algorithms have better overall performances than base models.

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C1	2.18%	0.00	3.08%	0.71	87.50%	-0.32%	-0.58%
M2	C1	0.25%	0.00	3.82%	0.07	55.49%	-5.31%	-9.83%
M3	C1	0.19%	0.00	3.82%	0.05	49.81%	-4.91%	-7.81%
M4	C1	1.82%	0.00	3.32%	0.55	77.84%	-1.60%	-3.06%
M5	C1	0.08%	0.00	4.07%	0.02	45.45%	-5.12%	-8.79%
M6	C1	2.17%	0.00	3.09%	0.70	86.93%	-0.36%	-0.63%
M1	C2	0.99%	0.00	4.44%	0.22	60.80%	-5.39%	-8.95%
M2	C2	0.68%	0.00	4.51%	0.15	54.73%	-5.58%	-9.22%
M3	C2	0.14%	0.00	4.57%	0.03	51.14%	-7.45%	-10.31%
M4	C2	1.02%	0.00	4.44%	0.23	60.04%	-5.29%	-8.46%
M5	C2	0.05%	0.00	4.44%	0.01	50.30%	-6.85%	-9.41%
M6	C2	0.14%	0.00	4.57%	0.03	51.14%	-7.45%	-10.31%
M1	C3	1.21%	0.01	8.02%	0.15	17.23%	-10.09%	-20.62%
M2	C3	0.62%	0.01	8.10%	0.08	14.02%	-10.09%	-19.68%
M3	C3	-0.24%	0.01	8.14%	-0.03	13.45%	-15.50%	-24.66%
M4	C3	0.09%	0.01	8.13%	0.01	14.77%	-14.99%	-23.73%
M5	C3	-0.84%	0.01	8.74%	-0.10	5.76%	-20.08%	-27.78%
M6	C3	0.26%	0.01	8.13%	0.03	16.29%	-16.38%	-24.83%
M1	C4	1.21%	0.00	4.22%	0.29	63.64%	-4.85%	-8.54%
M2	C4	0.11%	0.00	4.41%	0.03	51.14%	-6.67%	-10.82%
M3	C4	-0.10%	0.00	4.42%	-0.02	49.05%	-6.69%	-10.32%
M4	C4	0.81%	0.00	4.32%	0.19	58.14%	-5.07%	-8.23%
M5	C4	-0.28%	0.00	4.49%	-0.06	48.18%	-7.47%	-10.86%
M6	C4	1.03%	0.00	4.27%	0.24	60.23%	-4.89%	-7.94%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C5	2.87%	0.00	4.83%	0.59	48.30%	-0.10%	-4.22%
M2	C5	-0.39%	0.00	5.67%	-0.07	22.54%	-9.07%	-13.50%
M3	C5	-0.39%	0.00	5.67%	-0.07	22.54%	-9.07%	-13.50%
M4	C5	2.76%	0.00	4.90%	0.56	46.78%	-0.10%	-4.11%
M5	C5	-0.25%	0.00	4.25%	-0.06	20.00%	-5.66%	-9.42%
M6	C5	-0.41%	0.00	5.67%	-0.07	22.35%	-9.07%	-13.50%
M1	C6	2.73%	0.00	5.65%	0.48	70.83%	-4.69%	-7.84%
M2	C6	1.09%	0.00	6.21%	0.18	54.36%	-7.95%	-11.52%
M3	C6	-0.17%	0.00	6.33%	-0.03	47.16%	-10.33%	-14.18%
M4	C6	1.80%	0.00	6.03%	0.30	60.61%	-7.08%	-10.32%
M5	C6	-0.41%	0.00	6.04%	-0.07	46.06%	-10.56%	-15.44%
M6	C6	-0.16%	0.00	6.33%	-0.02	47.35%	-10.33%	-14.18%
M1	C7	7.15%	0.00	6.81%	1.05	92.65%	-0.58%	-0.89%
M2	C7	3.29%	0.01	9.36%	0.35	58.82%	-8.93%	-13.07%
M3	C7	-0.71%	0.01	9.95%	-0.07	47.06%	-14.58%	-24.26%
M4	C7	6.19%	0.01	7.71%	0.80	82.35%	-3.73%	-6.69%
M5	C7	3.30%	0.01	10.60%	0.31	55.29%	-10.41%	-15.21%
M6	C7	-0.75%	0.01	9.95%	-0.07	47.06%	-14.58%	-24.26%
M1	C8	5.04%	0.00	5.46%	0.92	92.50%	-0.68%	-1.00%
M2	C8	2.47%	0.00	7.05%	0.35	61.25%	-6.87%	-9.24%
M3	C8	0.84%	0.01	7.45%	0.11	48.75%	-8.27%	-11.29%
M4	C8	3.92%	0.00	6.34%	0.62	78.75%	-5.00%	-7.38%
M5	C8	2.38%	0.01	8.10%	0.29	60.00%	-8.09%	-11.68%
M6	C8	4.96%	0.00	5.54%	0.90	91.88%	-0.74%	-1.71%
M1	C9	11.41%	0.03	16.97%	0.67	68.09%	-1.30%	-2.38%
M2	C9	8.52%	0.03	18.64%	0.46	51.06%	-13.32%	-16.41%
M3	C9	-0.05%	0.04	20.59%	0.00	38.30%	-21.98%	-57.40%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M4	C9	11.24%	0.03	17.09%	0.66	65.96%	-2.45%	-3.34%
M5	C9	6.66%	0.05	23.19%	0.29	37.93%	-19.99%	-21.33%
M6	C9	11.41%	0.03	16.97%	0.67	68.09%	-1.30%	-2.38%
M1	C10	1.99%	0.01	8.63%	0.23	50.00%	-6.83%	-10.88%
M2	C10	0.50%	0.01	8.85%	0.06	45.45%	-9.41%	-19.63%
M3	C10	0.45%	0.01	8.86%	0.05	51.52%	-12.93%	-20.64%
M4	C10	0.05%	0.00	0.16%	0.31	67.61%	-0.11%	-0.25%
M5	C10	-0.07%	0.01	7.97%	-0.01	47.88%	-12.10%	-18.75%
M6	C10	0.00%	0.00	0.17%	-0.03	48.79%	-0.24%	-0.41%
M1 Average		3.68%	0.01	6.81%	0.53	65.15%	-3.48%	-6.59%
M2 Average		1.72%	0.01	7.66%	0.16	46.89%	-8.32%	-13.29%
M3 Average		0.00%	0.01	7.98%	0.00	41.88%	-11.17%	-19.44%
M4 Average		2.97%	0.01	6.24%	0.42	61.29%	-4.54%	-7.56%
M5 Average		1.06%	0.01	8.19%	0.06	41.69%	-10.63%	-14.87%
M6 Average		1.86%	0.01	6.47%	0.24	54.01%	-6.53%	-10.01%

Table 12: Model performance of training set for daily trade in business term

From table 12, C9, C8 and C7 produced the highest return in respective order and have positive return where other coins are producing positive return slightly above zero percent. Furthermore, when adding risk into consideration or by observing Sharpe Ratio which is a relation between return and risk of coins, C8, C9, and C7 have the highest Sharpe Ratio in respective order.

In term of model performance, M1 is the best while trading C1, C3, C4, C5, C6, C7, C8 and C9. M4 is the best while trading C2 and C10. Furthermore, if we

consider an average performance of each model, M1 is the best model while trading all cryptocurrency coins.

Ensemble algorithms are performing worse than base models in term of overall performance. During this trade, M1 which is the base models is among the best while trading a specific coin. M4 was close having as good as M1 for overall performance since M4 is the ensemble algorithm with the highest weight of M1. The aim of ensemble algorithm is to bring out upsides of each base model to make the better model, but it can be worse if the base models has only outstanding model. As a result, from having only a good base models and two bad, ensemble algorithms have worse overall performances than base models.

Model	Coins	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C1	0.83%	0.00	4.67%	0.18	56.82%	-7.11%	-10.72%
M2	C1	0.72%	0.00	4.70%	0.15	50.76%	-6.33%	-9.77%
M3	C1	0.06%	0.00	4.76%	0.01	49.24%	-8.70%	-10.50%
M4	C1	1.20%	0.00	4.59%	0.26	55.30%	-5.67%	-7.87%
M5	C1	-0.31%	0.00	4.02%	-0.08	42.73%	-6.63%	-8.70%
M6	C1	0.53%	0.00	4.73%	0.11	56.06%	-8.26%	-9.12%
M1	C2	2.09%	0.00	5.90%	0.35	71.21%	-9.30%	-12.88%
M2	C2	1.29%	0.00	6.15%	0.21	59.09%	-10.05%	-12.75%
M3	C2	-0.51%	0.00	6.29%	-0.08	44.70%	-11.84%	-15.76%
M4	C2	1.75%	0.00	6.03%	0.29	62.88%	-10.44%	-13.02%
M5	C2	0.67%	0.00	5.46%	0.12	54.85%	-8.09%	-11.60%
M6	C2	-0.62%	0.00	6.29%	-0.10	46.21%	-11.84%	-15.76%
M1	C3	1.40%	0.01	8.13%	0.17	50.00%	-10.57%	-15.32%
M2	C3	1.40%	0.01	8.13%	0.17	50.00%	-10.57%	-15.32%
M3	C3	-0.46%	0.01	8.27%	-0.06	43.18%	-15.58%	-19.62%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
						%		
M4	C3	-0.46%	0.01	8.27%	-0.06	43.18%	-15.58%	-19.62%
M5	C3	0.35%	0.01	7.60%	0.05	31.82%	-11.22%	-18.37%
M6	C3	-0.91%	0.01	8.24%	-0.11	45.45%	-17.41%	-21.96%
M1	C4	2.07%	0.01	10.55%	0.20	59.85%	-12.06%	-17.58%
M2	C4	2.07%	0.01	10.55%	0.20	59.85%	-12.06%	-17.58%
M3	C4	0.16%	0.01	10.77%	0.01	50.76%	-17.29%	-25.56%
M4	C4	0.16%	0.01	10.77%	0.01	50.76%	-17.29%	-25.56%
M5	C4	-0.30%	0.01	7.65%	-0.04	46.67%	-10.67%	-16.43%
M6	C4	-1.76%	0.01	10.66%	-0.17	44.70%	-24.22%	-34.29%
M1	C5	1.87%	0.00	3.01%	0.62	41.67%	-0.10%	-4.11%
M2	C5	-0.40%	0.00	3.59%	-0.11	19.70%	-5.37%	-10.77%
M3	C5	-0.40%	0.00	3.59%	-0.11	19.70%	-5.37%	-10.77%
M4	C5	1.70%	0.00	3.12%	0.55	39.39%	-1.27%	-3.96%
M5	C5	-0.25%	0.00	4.25%	-0.06	20.00%	-5.66%	-9.42%
M6	C5	-0.40%	0.00	3.59%	-0.11	19.70%	-5.37%	-10.77%
M1	C6	0.73%	0.01	8.22%	0.09	53.79%	-11.27%	-18.03%
M2	C6	1.34%	0.01	8.14%	0.16	53.79%	-10.91%	-16.15%
M3	C6	-0.05%	0.01	8.27%	-0.01	49.24%	-14.48%	-19.46%
M4	C6	1.59%	0.01	8.10%	0.20	56.82%	-11.06%	-16.88%
M5	C6	-0.10%	0.01	7.47%	-0.01	45.76%	-12.68%	-15.58%
M6	C6	0.64%	0.01	8.24%	0.08	52.27%	-12.02%	-17.71%
M1	C7	8.00%	0.01	11.67%	0.69	88.24%	-5.22%	-5.50%
M2	C7	-0.36%	0.02	14.29%	-0.03	47.06%	-25.57%	-55.17%
M3	C7	1.31%	0.02	14.21%	0.09	47.06%	-17.54%	-19.91%
M4	C7	6.19%	0.01	7.71%	0.80	82.35%	-3.73%	-6.69%
M5	C7	3.30%	0.01	10.60%	0.31	55.29%	-10.41%	-15.21%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M6	C7	-0.75%	0.01	9.95%	-0.07	47.06%	-14.58%	-24.26%
M1	C8	4.29%	0.01	7.39%	0.58	77.50%	-8.51%	-9.12%
M2	C8	1.73%	0.01	8.43%	0.21	57.50%	-10.08%	-11.97%
M3	C8	0.86%	0.01	8.57%	0.10	52.50%	-14.51%	-17.24%
M4	C8	7.41%	0.03	16.29%	0.46	82.93%	-8.47%	-9.12%
M5	C8	0.40%	0.01	8.40%	0.05	57.00%	-12.30%	-18.24%
M6	C8	4.29%	0.01	7.39%	0.58	77.50%	-8.51%	-9.12%
M1	C9	7.58%	0.01	8.41%	0.90	63.64%	-3.44%	-3.44%
M2	C9	-1.43%	0.01	11.57%	-0.12	27.27%	-22.34%	-22.34%
M3	C9	1.39%	0.01	11.54%	0.12	36.36%	-14.40%	-14.40%
M4	C9	7.03%	0.01	8.93%	0.79	54.55%	-3.44%	-3.44%
M5	C9	6.66%	0.05	23.19%	0.29	37.93%	-19.99%	-21.33%
M6	C9	4.98%	0.01	10.36%	0.48	54.55%	-14.40%	-14.40%
M1	C10	4.09%	0.02	15.55%	0.26	53.03%	-15.70%	-19.20%
M2	C10	4.09%	0.02	15.55%	0.26	53.03%	-15.70%	-19.20%
M3	C10	-1.00%	0.03	16.10%	-0.06	49.24%	-27.95%	-37.17%
M4	C10	-1.00%	0.03	16.10%	-0.06	49.24%	-27.95%	-37.17%
M5	C10	-1.53%	0.01	10.11%	-0.15	45.15%	-18.91%	-28.85%
M6	C10	-0.34%	0.03	16.12%	-0.02	52.27%	-25.03%	-37.17%
M1 Average		3.30%	0.01	8.35%	0.40	61.57%	-8.33%	-11.59%
M2 Average		1.04%	0.01	9.11%	0.11	47.80%	-12.90%	-19.10%
M3 Average		0.13%	0.01	9.24%	0.00	44.20%	-14.77%	-19.04%
M4 Average		2.56%	0.01	8.99%	0.32	57.74%	-10.49%	-14.33%
M5 Average		0.89%	0.01	8.87%	0.05	43.72%	-11.66%	-16.37%
M6 Average		0.57%	0.01	8.56%	0.07	49.58%	-14.16%	-19.46%

Table 13: Model performance of testing set for daily trade in business term



From table 13, C9, C8 and C7 produced the highest return in respective order and have positive return where other coins are producing positive return slightly above zero percent. Furthermore, when adding risk into consideration or by observing Sharpe Ratio which is a relation between return and risk of coins, C9, C8, and C7 have the highest Sharpe Ratio in respective order.

In term of model performance, M1 is the best while trading C1, C2, C3, C4, C5, C8, C9, and C10. M4 is the best while trading C6 and C7. Furthermore, if we consider an average performance of each model, M1 is the best model while trading all cryptocurrency coins.

Ensemble algorithms are performing worse than base models in term of overall performance. During this trade, M1 which is the base models is among the best while trading a specific coin. M4 was close having as good as M1 for overall performance since M4 is the ensemble algorithm with the highest weight of M1. The aim of ensemble algorithm is to bring out upsides of each base model to make the better model, but it can be worse if the base models has only outstanding model. As a result, from having only a good base models and two bad, ensemble algorithms have worse overall performances than base models.

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C1	7.12%	0.01	7.40%	0.96	95.83%	0.01%	-0.16%
M2	C1	3.23%	0.01	9.82%	0.33	55.56%	-11.71%	-14.51%
M3	C1	0.27%	0.01	10.37%	0.03	48.61%	-19.41%	-26.90%
M4	C1	5.86%	0.01	8.46%	0.69	73.61%	-4.03%	-7.74%
M5	C1	0.08%	0.00	4.07%	0.02	45.45%	-5.12%	-8.79%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M6	C1	6.88%	0.01	7.63%	0.90	94.44%	-0.08%	-3.03%
M1	C2	6.30%	0.01	12.03%	0.52	77.78%	-7.05%	-16.82%
M2	C2	5.70%	0.02	12.34%	0.46	62.50%	-9.86%	-13.74%
M3	C2	-0.11%	0.02	13.66%	-0.01	51.39%	-24.38%	-45.10%
M4	C2	6.31%	0.01	12.03%	0.52	69.44%	-8.59%	-12.12%
M5	C2	-1.41%	0.01	11.53%	-0.12	44.44%	-24.99%	-34.00%
M6	C2	0.04%	0.02	13.66%	0.00	52.78%	-24.38%	-45.10%
M1	C3	6.59%	0.03	18.10%	0.36	37.50%	-21.83%	-25.08%
M2	C3	4.44%	0.04	18.76%	0.24	30.56%	-25.08%	-25.08%
M3	C3	2.46%	0.04	19.15%	0.13	27.78%	-25.10%	-36.25%
M4	C3	2.89%	0.04	19.08%	0.15	29.17%	-25.10%	-36.25%
M5	C3	3.01%	0.03	17.98%	0.17	22.22%	-30.95%	-41.81%
M6	C3	3.59%	0.04	18.96%	0.19	33.33%	-25.10%	-36.25%
M1	C4	5.74%	0.01	10.79%	0.53	68.06%	-7.50%	-9.61%
M2	C4	1.56%	0.01	12.16%	0.13	51.39%	-20.70%	-31.85%
M3	C4	-0.85%	0.02	12.26%	-0.07	37.50%	-17.29%	-28.84%
M4	C4	1.00%	0.01	12.23%	0.08	48.61%	-16.91%	-28.71%
M5	C4	-0.13%	0.01	11.81%	-0.01	46.67%	-24.44%	-34.49%
M6	C4	5.24%	0.01	11.05%	0.47	62.50%	-9.42%	-12.54%
M1	C5	9.49%	0.01	12.00%	0.79	79.17%	-0.10%	-0.10%
M2	C5	-1.02%	0.02	15.38%	-0.07	36.11%	-22.58%	-46.07%
M3	C5	-1.02%	0.02	15.38%	-0.07	36.11%	-22.58%	-46.07%
M4	C5	8.71%	0.02	12.59%	0.69	75.00%	-1.56%	-9.41%
M5	C5	-0.62%	0.01	9.27%	-0.07	35.56%	-18.90%	-23.98%
M6	C5	-0.82%	0.02	15.39%	-0.05	37.50%	-22.58%	-46.07%
M1	C6	11.18%	0.02	15.07%	0.74	79.17%	-7.25%	-8.84%
M2	C6	8.52%	0.03	16.76%	0.51	66.67%	-13.02%	-18.10%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M3	C6	0.44%	0.04	18.87%	0.02	44.44%	-27.94%	-36.51%
M4	C6	8.73%	0.03	16.65%	0.52	69.44%	-11.92%	-15.07%
M5	C6	-1.47%	0.03	16.63%	-0.09	44.44%	-27.95%	-39.22%
M6	C6	0.44%	0.04	18.87%	0.02	44.44%	-27.94%	-36.51%
M1	C7	19.94%	0.03	18.52%	1.08	94.44%	-1.26%	-1.26%
M2	C7	12.34%	0.06	24.59%	0.50	61.11%	-23.62%	-23.62%
M3	C7	9.94%	0.07	25.73%	0.39	66.67%	-46.32%	-46.32%
M4	C7	17.99%	0.04	20.54%	0.88	83.33%	-7.48%	-7.48%
M5	C7	22.51%	0.06	24.15%	0.93	81.82%	-7.48%	-7.48%
M6	C7	9.82%	0.07	25.78%	0.38	61.11%	-46.32%	-46.32%
M1	C8	19.23%	0.06	23.76%	0.81	95.45%	-0.40%	-0.91%
M2	C8	14.70%	0.07	26.95%	0.55	68.18%	-13.40%	-14.03%
M3	C8	5.80%	0.09	30.36%	0.19	68.18%	-94.71%	-108.06%
M4	C8	18.92%	0.06	24.02%	0.79	90.91%	-3.09%	-3.48%
M5	C8	18.99%	0.11	33.88%	0.56	53.85%	-14.03%	-14.03%
M6	C8	16.34%	0.07	25.94%	0.63	90.91%	-27.26%	-31.91%
M1	C9	79.40%	1.81	134.47%	0.59	83.33%	-0.10%	-0.10%
M2	C9	79.40%	1.81	134.47%	0.59	83.33%	-0.10%	3.60%
M3	C9	59.53%	2.14	146.28%	0.41	50.00%	-39.43%	-39.43%
M4	C9	79.40%	1.81	134.47%	0.59	83.33%	-0.10%	3.60%
M5	C9	98.05%	2.84	168.60%	0.58	75.00%	-0.10%	3.60%
M6	C9	79.40%	1.81	134.47%	0.59	83.33%	-0.10%	3.60%
M1	C10	12.99%	0.07	26.25%	0.49	61.11%	-18.51%	-18.51%
M2	C10	5.62%	0.08	28.90%	0.19	55.56%	-66.39%	-66.39%
M3	C10	11.66%	0.07	26.91%	0.43	55.56%	-18.51%	-18.51%
M4	C10	0.22%	0.00	0.46%	0.49	76.39%	-0.25%	-0.25%
M5	C10	2.74%	0.05	22.33%	0.12	48.89%	-66.39%	-66.39%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M6	C10	2.74%	0.05	22.33%	0.12	48.89%	-66.39%	-66.39%
M1 Average		17.80%	0.21	27.84%	0.69	77.18%	-6.40%	-8.14%
M2 Average		13.45%	0.22	30.01%	0.34	57.10%	-20.65%	-24.98%
M3 Average		8.81%	0.25	31.90%	0.15	48.62%	-33.57%	-43.20%
M4 Average		15.00%	0.20	26.05%	0.54	69.92%	-7.90%	-11.69%
M5 Average		14.18%	0.32	32.02%	0.21	49.83%	-22.03%	-26.66%
M6 Average		12.37%	0.21	29.41%	0.33	60.92%	-24.96%	-32.05%

*Table 14: Model performance of testing set for weekly trade in business term*

From table 14, C9, C8 and C7 produced the highest return in respective order and have positive return where other coins are producing positive return slightly above zero percent. Furthermore, when adding risk into consideration or by observing Sharpe Ratio which is a relation between return and risk of coins, C8, C9, and C1 have the highest Sharpe Ratio in respective order.

In term of model performance, M1 is the best while trading all cryptocurrency coins except trading C10. M4 is the best model for trading C10 weekly.

Ensemble algorithms are performing worse than base models in term of overall performance. During this trade, M1 which is the base models is among the best while trading a specific coin. M4 was close having as good as M1 for overall performance since M4 is the ensemble algorithm with the highest weight of M1. The aim of ensemble algorithm is to bring out upsides of each base model to make the better model, but it can be worse if the base models has only outstanding model. As a

result, from having only a good base models and two bad, ensemble algorithms have worse overall performances than base models.

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C1	4.41%	0.02	13.78%	0.32	55.56%	-18.11%	-18.11%
M2	C1	7.87%	0.01	12.02%	0.65	66.67%	-10.16%	-10.16%
M3	C1	5.93%	0.02	13.16%	0.45	72.22%	-18.11%	-18.11%
M4	C1	3.76%	0.02	13.99%	0.27	55.56%	-18.11%	-18.11%
M5	C1	1.33%	0.01	11.07%	0.12	60.00%	-25.11%	-28.12%
M6	C1	1.52%	0.02	14.46%	0.11	55.56%	-19.55%	-19.55%
M1	C2	33.25%	0.27	52.30%	0.64	72.22%	-37.29%	-37.29%
M2	C2	33.25%	0.27	52.30%	0.64	72.22%	-23.23%	-23.23%
M3	C2	11.17%	0.38	61.50%	0.18	44.44%	-27.97%	-27.97%
M4	C2	11.17%	0.38	61.50%	0.18	44.44%	-8.72%	-8.72%
M5	C2	-5.53%	0.17	40.74%	-0.14	44.44%	-27.69%	-29.80%
M6	C2	15.67%	0.37	60.44%	0.26	50.00%	-24.46%	-24.46%
M1	C3	8.63%	0.03	16.50%	0.52	72.22%	-31.87%	-31.87%
M2	C3	9.91%	0.02	15.72%	0.63	61.11%	-31.87%	-31.87%
M3	C3	-1.86%	0.04	18.72%	-0.10	50.00%	-37.28%	-37.28%
M4	C3	12.52%	0.02	13.59%	0.92	77.78%	-37.28%	-37.28%
M5	C3	1.45%	0.02	15.78%	0.09	53.33%	-28.50%	-29.57%
M6	C3	2.19%	0.03	18.65%	0.12	55.56%	-87.69%	-87.69%
M1	C4	10.70%	0.07	25.54%	0.42	61.11%	-58.97%	-58.97%
M2	C4	10.70%	0.07	25.54%	0.42	61.11%	-58.97%	-58.97%
M3	C4	6.49%	0.07	27.02%	0.24	55.56%	-58.97%	-58.97%
M4	C4	6.49%	0.07	27.02%	0.24	55.56%	-58.97%	-58.97%
M5	C4	4.25%	0.04	19.99%	0.21	51.11%	-44.71%	-58.80%
M6	C4	-8.34%	0.07	26.57%	-0.31	38.89%	-118.00%	-118.00%

Model	Coin	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M1	C5	18.22%	0.14	37.56%	0.49	66.67%	-0.10%	-0.10%
M2	C5	18.22%	0.14	37.56%	0.49	66.67%	-19.31%	-19.31%
M3	C5	11.89%	0.16	40.18%	0.30	55.56%	-19.31%	-19.31%
M4	C5	11.89%	0.16	40.18%	0.30	55.56%	-19.31%	-19.31%
M5	C5	2.04%	0.07	25.83%	0.08	57.78%	-18.90%	-23.98%
M6	C5	-11.05%	0.16	40.57%	-0.27	50.00%	-19.31%	-19.31%
M1	C6	6.63%	0.00	5.70%	1.16	88.89%	-42.98%	-42.98%
M2	C6	-1.06%	0.01	8.91%	-0.12	44.44%	-22.37%	-22.37%
M3	C6	-1.06%	0.01	8.91%	-0.12	44.44%	-42.98%	-42.98%
M4	C6	3.53%	0.01	8.15%	0.43	72.22%	-17.04%	-17.04%
M5	C6	-0.62%	0.01	9.27%	-0.07	35.56%	-19.23%	-23.09%
M6	C6	-0.26%	0.01	8.96%	-0.03	50.00%	-17.60%	-17.60%
M1	C7	-0.79%	0.04	18.99%	-0.04	44.44%	-3.57%	-3.57%
M2	C7	7.35%	0.03	17.40%	0.42	61.11%	-4.50%	-4.50%
M3	C7	-4.68%	0.03	18.44%	-0.25	38.89%	-37.53%	-37.53%
M4	C7	10.56%	0.02	15.54%	0.68	72.22%	-7.48%	-7.48%
M5	C7	6.54%	0.03	17.51%	0.37	62.22%	-7.48%	-7.48%
M6	C7	6.20%	0.03	17.89%	0.35	55.56%	-46.32%	-46.32%
M1	C8	9.41%	0.04	18.95%	0.50	50.00%	-0.60%	-0.60%
M2	C8	7.42%	0.04	20.10%	0.37	50.00%	0.40%	0.40%
M3	C8	-7.38%	0.04	20.17%	-0.37	75.00%	-28.14%	-28.14%
M4	C8	17.99%	0.04	20.54%	0.88	83.33%	-0.60%	-0.60%
M5	C8	22.51%	0.06	24.15%	0.93	81.82%	-20.73%	-20.73%
M6	C8	9.82%	0.07	25.78%	0.38	61.11%	-0.60%	-0.60%
M1	C9	Lack of Data for 1 week trade						
M2	C9	Lack of Data for 1 week trade						
M3	C9	Lack of Data for 1 week trade						
M4	C9	Lack of Data for 1 week trade						

Model	Coins	Expected Return	Variance	Standard Deviation	Sharpe Ratio	% Profit	Value at Risk	Tailed Value at Risk
M5	C9	Lack of Data for 1 week trade						
M6	C9	Lack of Data for 1 week trade						
M1	C10	13.36%	0.02	12.36%	1.08	80.00%	-0.10%	0.00%
M2	C10	13.56%	0.01	12.09%	1.12	100.00%	-0.10%	0.00%
M3	C10	-12.89%	0.02	13.18%	-0.98	40.00%	-0.10%	0.00%
M4	C10	13.36%	0.02	12.36%	1.08	80.00%	-0.10%	0.00%
M5	C10	4.10%	0.03	17.61%	0.23	53.85%	-0.10%	3.60%
M6	C10	13.36%	0.02	12.36%	1.08	80.00%	-0.10%	0.00%
M1 Average		11.53%	0.07	22.41%	0.56	65.68%	-21.51%	-21.50%
M2 Average		11.91%	0.07	22.40%	0.51	64.81%	-18.90%	-18.89%
M3 Average		0.85%	0.09	24.59%	-0.07	52.90%	-30.04%	-30.03%
M4 Average		10.14%	0.08	23.65%	0.55	66.30%	-18.62%	-18.61%
M5 Average		4.01%	0.05	20.22%	0.20	55.57%	-21.38%	-24.22%
M6 Average		3.23%	0.09	25.08%	0.19	55.19%	-37.07%	-37.06%

Table 15: Model performance of testing set for weekly trade in business term

From table 15, C2, C8 and C5 produced the highest return in respective order and have positive return where other coins are producing positive return slightly above zero percent. Furthermore, when adding risk into consideration or by observing Sharpe Ratio which is a relation between return and risk of coins, C10, C8, and C3 have the highest Sharpe Ratio in respective order.

In term of model performance, M1 is the best while trading C1, C2, C3, C4, C5, C8, C9, and C10. M4 is the best while trading C6 and C7. Furthermore, if we

consider an average performance of each model, M1 is the best model while trading all cryptocurrency coins

Ensemble algorithms are performing worse than base models in term of overall performance. During this trade, M1 and M2 which are the base models are among the best while trading a specific coin. M4 was close having as good as M1 and M2 for overall performance since M4 is the ensemble algorithm with the highest weight of M1.

In conclusion, return has increased when trading at lower frequency. On the other hand, Sharpe ratio has not increased significantly because when expected return goes up, Sharpe Ratio is going up along with the return.

In term of model performance, M1 is the best for obtaining high expected return and high Sharpe Ratio despite trading in all 4 situations. Ensemble algorithms especially M4 are producing positive value of both return and Sharpe Ratio, but they are not as good as M1 except during daily trade situation. Madan et al., 2015 stated that the best time to trade a cryptocurrency is the daily trade since their report showed that an hourly data has relatively low accuracy comparing to a daily data. Furthermore, we cannot exactly compare a weekly trading to the others because there was too less data to make a conclusion even it has high return and Sharpe Ratio.

### **Visualizing Result**

Last section for result is to visualize it. The figure of train and test set of each trading signal are plotted with Accuracy vs. Sharpe Ratio scatter plot. By doing this certain method, we can make our observation for our best or the most generalized model less complex than observing only raw result data. For this section, the model



with Sharpe Ratio lesser than one or Accuracy lesser than 50 percent will be crossed out.

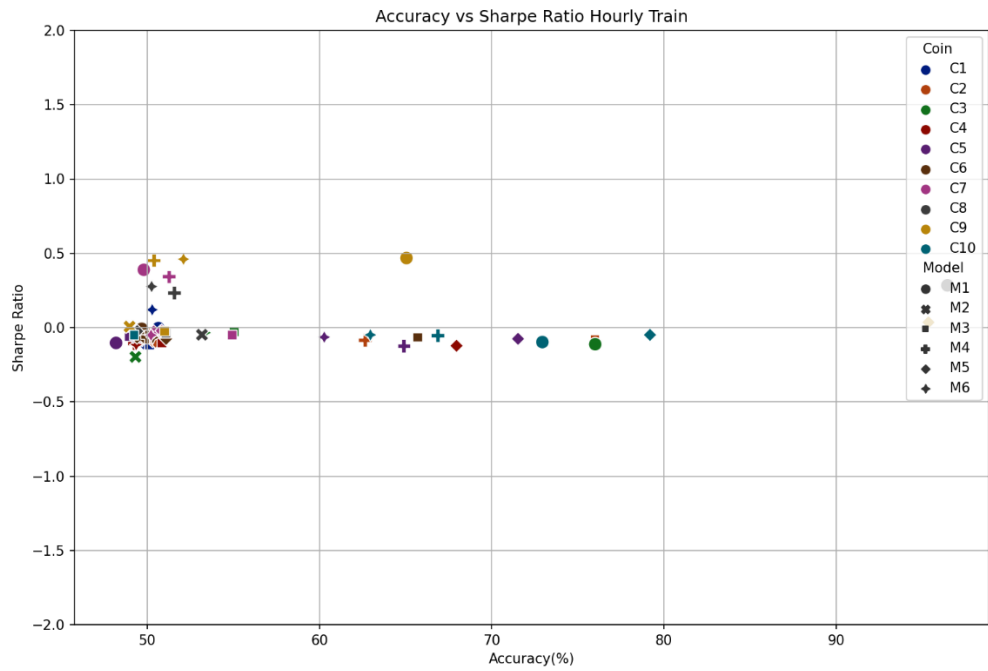


Figure 3: Scatter plot of Accuracy versus Sharpe Ratio of training set for hourly trade

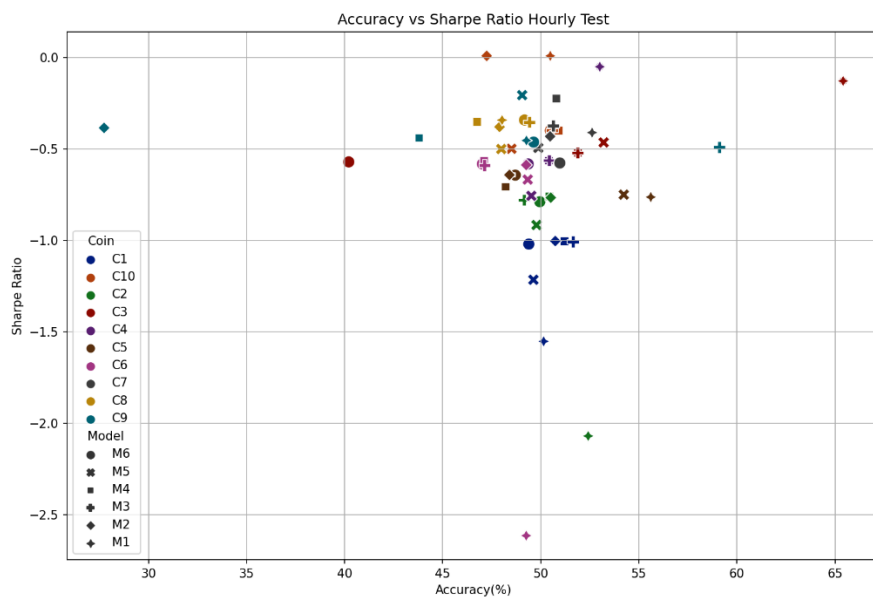


Figure 4 : Scatter plot of Accuracy versus Sharpe Ratio of testing set for hourly trade

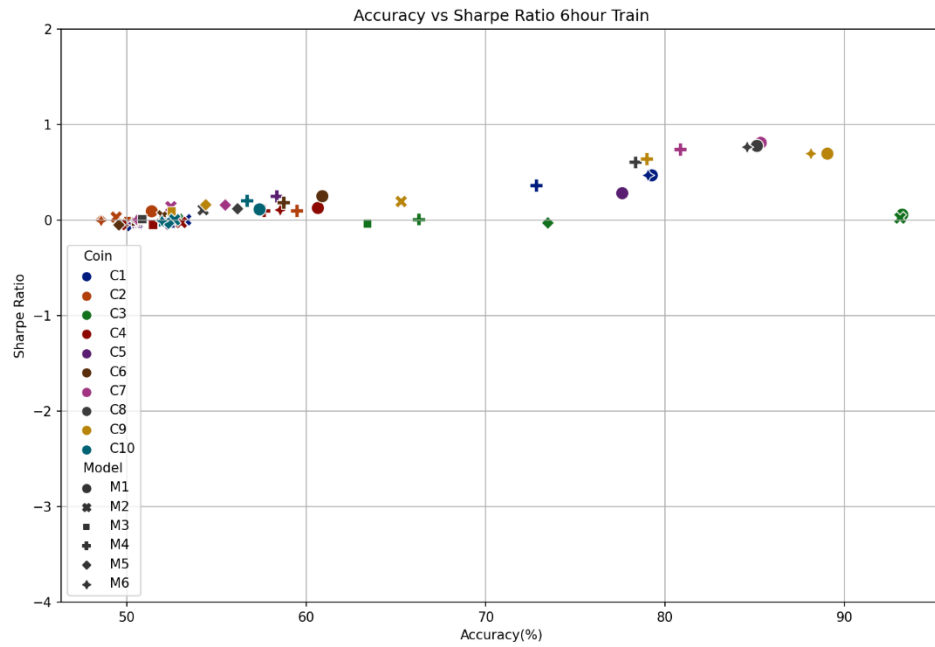


Figure 5: Scatter plot of Accuracy versus Sharpe Ratio of training set for six-hour trade

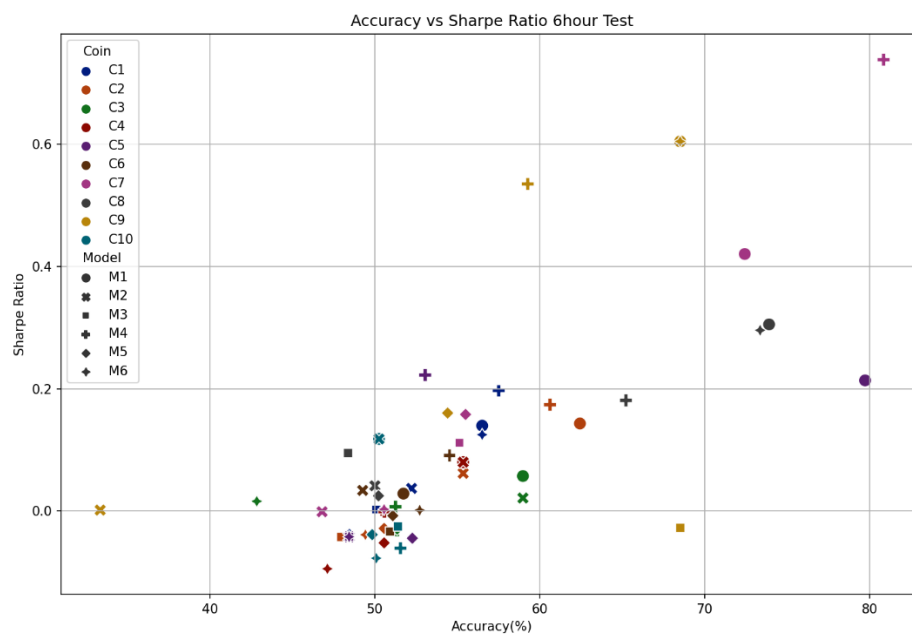


Figure 6: Scatter plot of Accuracy versus Sharpe Ratio of training set for six-hour trade

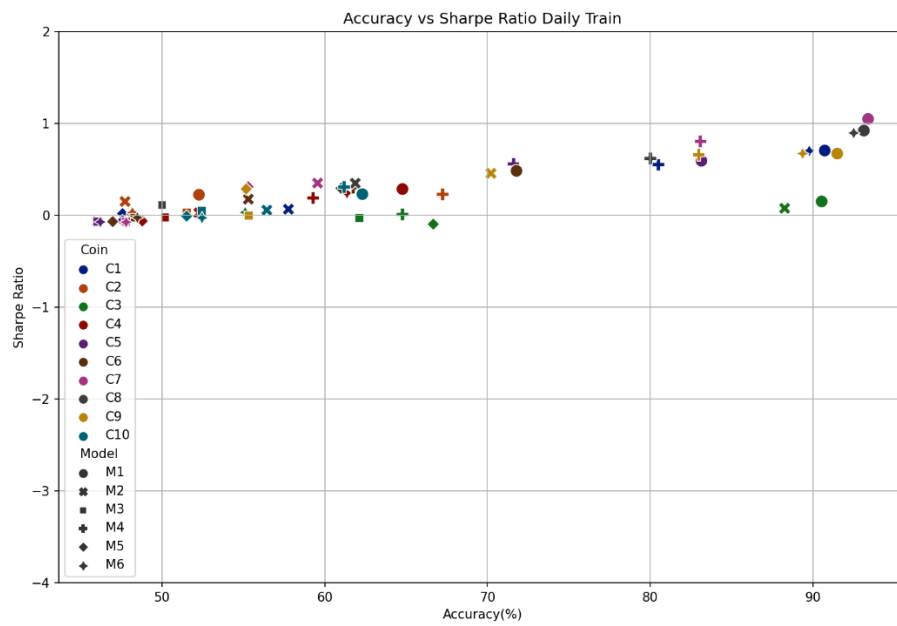


Figure 7: Scatter plot of Accuracy versus Sharpe Ratio of training set for daily trade



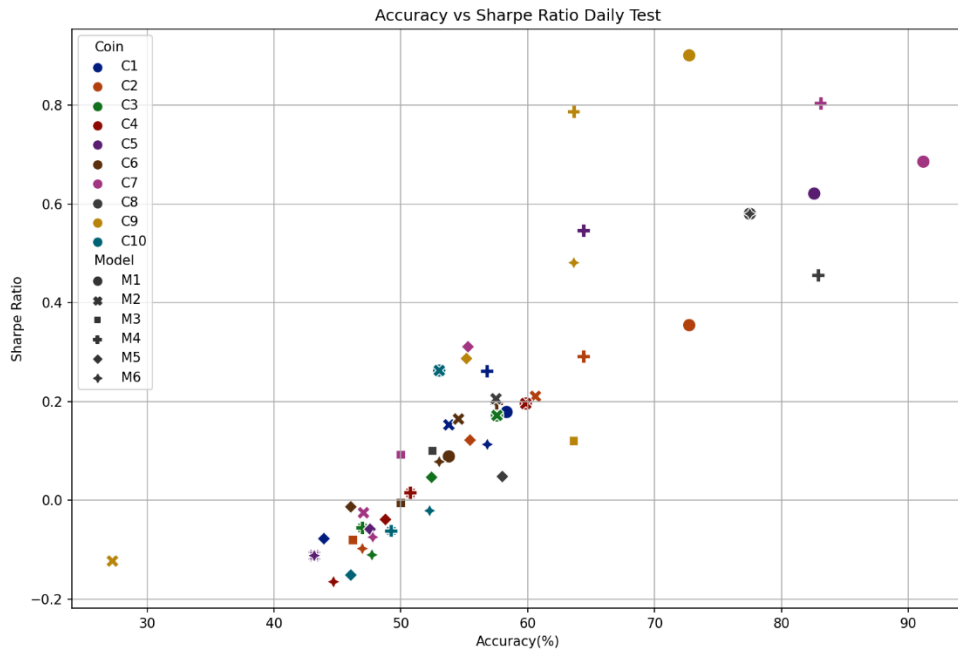


Figure 8: Scatter plot of Accuracy versus Sharpe Ratio of testing set for daily trade

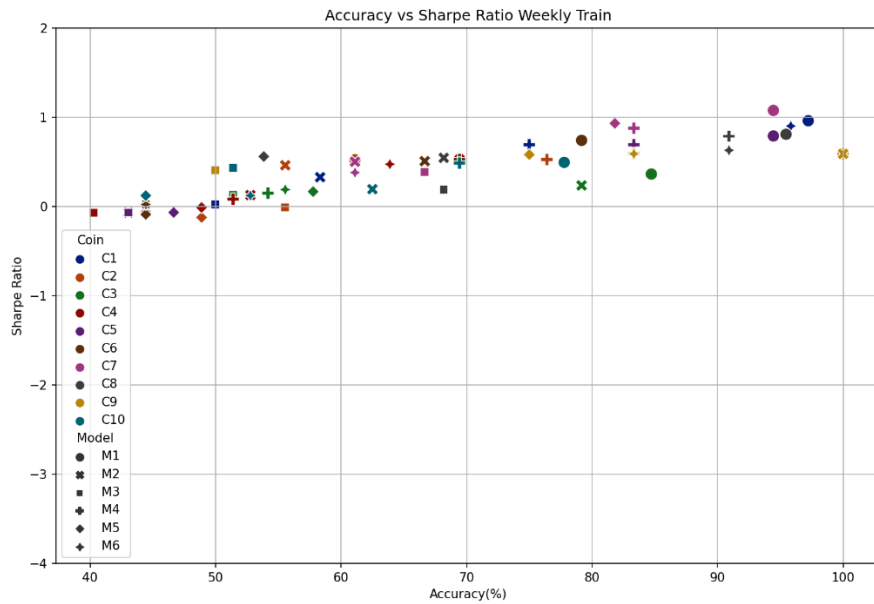


Figure 9: Scatter plot of Accuracy versus Sharpe Ratio of training set for weekly trade

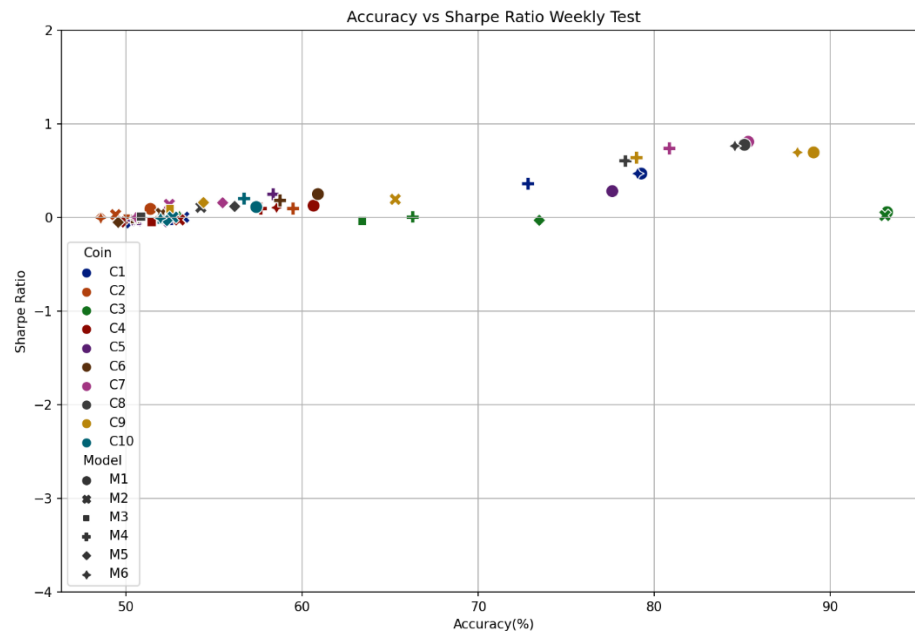


Figure 10: Scatter plot of Accuracy versus Sharpe Ratio of testing set for weekly trade

According to figure 4, 5, 6, 7, 8, 9 and 10, model 1, model 4 and model 6 are three of the best models in all of trading circumstances while C7, C8 and C9 are giving the best outcome in this research. Model 1 is the most outstanding model because outputs from trading cryptocurrency coins are mostly in the fourth quartile comparing to other models that has only one or two output that are in the fourth quartile.

In term of trading circumstances, accuracies change slightly when the trade timing increases. Accuracies are mostly in the 50 to 60 percent area for all trading circumstances, but the number of outputs that have outstanding accuracies are increasing. According to our literature review, it shows that predicting cryptocurrency daily has better accuracy than the accuracy from predicting hourly.

For Sharpe Ratios, in hourly trading, our Sharpe Ratios are not passing the criteria of above zero, but since the six-hour trading, there are significant change for Sharpe Ratios. In a hourly trading especially in testing set, there are few outcomes that passes the zero percent criteria. The testing set of hourly trading has only two outputs that fits our criteria. On the other hand, since the six-hour trading, there are only few results that do not fit our criteria. This means that overall Sharpe Ratio is increasing when the trade timing has increased.



## Discussion

Cryptocurrency trading has emerged as an investing choice for recent years. Its price has been dramatically risen since 2017. Cryptocurrency has a high return, but it also has very high risk to invest. To prevent a loss from trading cryptocurrency, investors are leaning toward cryptocurrency trading algorithm because a computer makes fewer mistakes compare to human.

Not like stock trading algorithm, a cryptocurrency trading algorithm still have several limitations. First, an algorithm does not work with high frequency trading. Second, a model itself is not generalized. To be more specific, a trading algorithm does not work when trading for multiple cryptocurrency coins which means that a model that works on Bitcoin will not work on other coins like Ethereum. Third, an algorithm has relative low accuracy.

A way to improve accuracy is using ensemble algorithm. Lyu and Nikora, 1991 had used an ensemble method to combine software prediction, and the accuracy increased. No cryptocurrency trading algorithm has gone through ensemble algorithm, so this is an opportunity to expand a field of research. The research aim to make cryptocurrency trading algorithm more accurate and generalized by using ensemble algorithms.

To create an ensemble algorithm, based models should be selected. There are 3 models that will be used as based models, Support Vector Machines, XGBoost and Long Short-Term Memory. According to Madan et al., 2015; Colianni et al., 2015; Lahmiri; Zbikowski, 2016; Sebastial and Godinho, 2021, they have explored the Support Vector Machine and found out that this model works with a low frequency

dataset and does not consume a lot of time while computing output. Madan et al., 2015 also explore a random forest model and said that a more structure of random forest should be able to fill the gap that random forest only works with low frequency data. As a result, XGBoost, an upgrade version of random forest algorithm is selected as one of based models due to its fast computing time and high accuracy. Last, a Long Short-Term memory is selected as the last based model. Lahmiri and Bekiros ,2019 stated that Long Short-Term memory is suitable for high frequency data, but the only downside is that it has very long computing due to its architecture complexity.

The methodology contains 3 parts, data collection, machine learning model, and model evaluation. There will be 10 cryptocurrency coins that are separately going through 3 based machine learning models and 3 ensemble machine learning models. First, data that were collected are closing price and volume of cryptocurrency and ten major stock indices closing price which all of them are in hourly data. Next, a single cryptocurrency will go through six machine learning models with 20 features which are 10 major stock indices, 9 closed prices of cryptocurrency that are not currently predicting and a volume of specific cryptocurrency. The machine models that were used as based models are XGBoost, Support Vector Machine and Long-Short term memory, and ensemble algorithms that will combine 3 based models are Equally Weighted Forecast Combinations, Adaptive Regression by Mixing and Aggregation of Forecasts through Exponential Reweight. Last, six models of 10 cryptocurrency coins will be evaluated under 2 factors, accuracy and Sharpe Ratio of returns.

The outputs from machine learning are undergoing 4 trading situations which are hourly trading, six-hour trading, daily trading and weekly trading to compare



between models even the dataset is hourly dataset. An accuracy is calculated from the correct prediction signal divided by a total signal, and Sharpe Ratio can be calculated from dividing the expected return by standard deviation of the return. In this research, a risk-free rate is neglected from the Sharpe Ratio calculation because our trade timings are too frequent for implying a risk-free rate. The results are evaluated separately by accuracy and Sharpe Ratio then they will be merged by creating scatter plots for each trading situations.

By comparing accuracies separately in 4 trading intervals, M1 was the best model where M4 came in second for all 4 intervals. M1 was performing best in training set, but there was a huge gap between train and test accuracies. On the other hand, M4 was the consistent model between train and test set.

For the Sharpe Ratio comparison, M1 was the best model where M4 came in second for all 4 trading intervals. It has the highest expected return while also having high risk. M4 was almost as good as M1 in term of expected return, but it has higher risk than M1. As a result, M1 has higher Sharpe Ratio than M4.

After merging into a scatter plot, model 1, model 4 and model6 are three of the best models in all of trading circumstances while C7, C8 and C9 are giving the best outcome in this research. Model 1 is the most outstanding model because outputs from trading cryptocurrency coins are mostly in the fourth quartile comparing to other models that has only one or two output that are in the fourth quartile.

In term of trading circumstances, accuracies change slightly when the trade timing increases. Accuracies are mostly in the 50 to 60 percent area for all trading circumstances, but the number of outputs that have outstanding accuracies are

increasing. According to our literature review, it shows that predicting cryptocurrency daily has better accuracy than the accuracy from predicting hourly.

Ensemble algorithms are performing worse than base models in term of overall performance. During the trade, M1 was only based model that outstanding while M2 and M3 were performing under average. The aim of ensemble algorithm is to bring out upsides of each base model to make the better model, but it can be worse if the base models has only outstanding model. As a result, from having only a good base models and two bad, ensemble algorithms have worse overall performances than base models.

### **Comparison with previous work**

The previous works in literature review that can be compared with our result are Madan et al., 2015, Colianni et al., 2015 and Lahmiri and Bekiros, 2019. They stated the accuracy from their experiment while others where comparison with no accuracy stated. There are no Sharpe Ratio stated in the literature reviews.

First comparison is the data frequency or trading frequency. Our results show that while trading hourly our accuracy is around 50-52 percent overall, and it has around 70 percent over daily trading. Accuracies from previous research were around 90-95 percent during daily trade, and 50-55% during hourly trade. Our result in hourly trade is not differ from the previous research, but it was significantly different with an hourly trade. The projected accuracy was 85-90 percent to be similar as the literature which is 15 percent more than our outcome. In conclusion, trade timing is not the only factor that affect our accuracy.

Second, features adding into machine learning algorithms can be compared. Madan et al., 2015 and Lahmiri and Behkiros, 2019 used previous closing price to predict the next hour price, which means that the model did not contain feature. On the other hand, Colianni et al., 2015 used tweet on twitter as a feature. Our research uses 21 features compared to one feature of Colianni et al., 2015. The result was different by 20 percent accuracy.

### **Research Limitation**

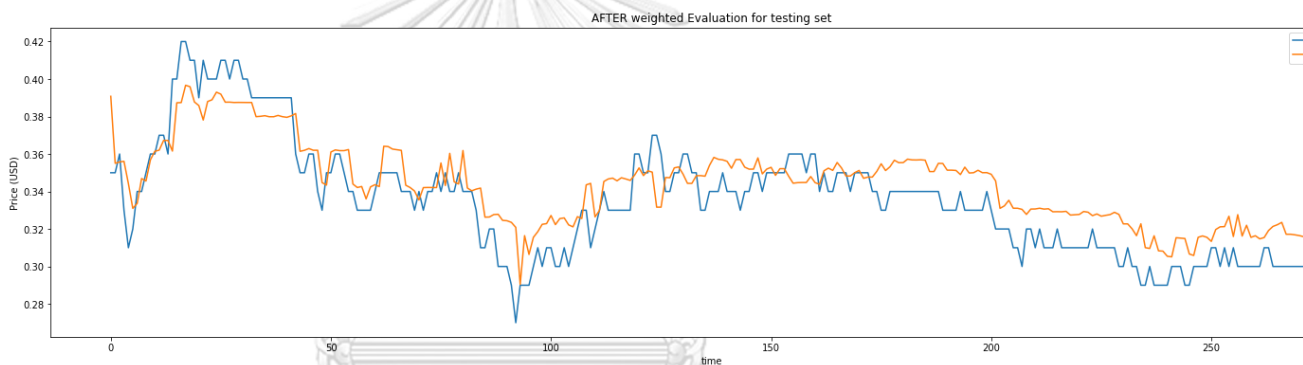
There are four main errors found in this research. The first problem is acquiring the dataset. The dataset for this research is an hourly dataset which is not mostly provided on most software, especially stock indices data. During this research, we tried to collect stock indices data using Bloomberg Terminal, which is widely used and has plenty of data, but in the end, we could not acquire hourly data of stock indices. Bloomberg Terminal can provide hourly data for six months timeframe even with the most expensive plan on the software, but our research required to use of hourly data for two years span. We have to assume that stock prices are equal for the whole day.

Second, the timeframe is not probable for predicting some cryptocurrency coins. Half of the dataset is a linear dataset for some coins like MATIC, so most predictions are in linear value or even cause the straight-line prediction. This error occurred because, at the beginning of the dataset, which is 2019, several coins have significant value to trade, but most cryptocurrency coins were just created.

Third, there is too much dataset. Our data is at 21 by 17362 cells in a CSV file. By having a colossal dataset and many features for machine learning, the models will

sometimes undergo an underfitting situation where the model cannot capture the relationship between the input and output variables. We have tried to reduce features of the machine learning model to only previous hour prices to predict upcoming hour prices instead of our current method. As a result, accuracy slightly increased, and no underfitting situations occurred.

Last, our evaluation method is simple. In this research, several models' outcomes have an excellent trend for candidates for a generalized trading algorithm, but the accuracy is low, as shown in Figure 4.



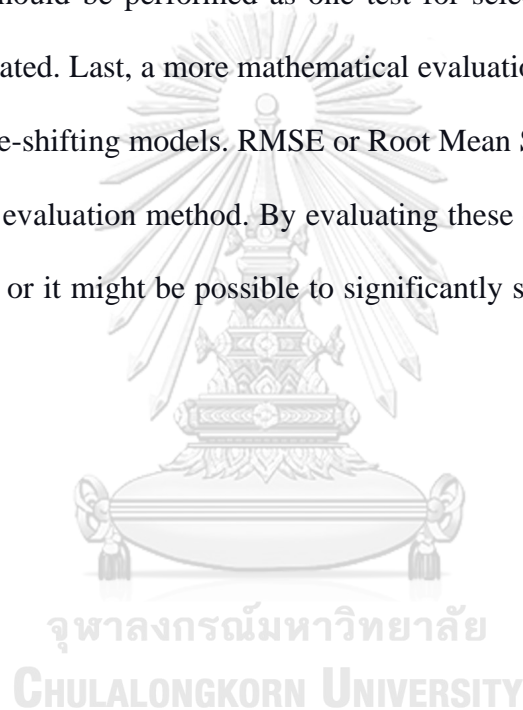
*Figure 11: Example of a “good trend” prediction during research*

From the figure 11, this is the result of the AFTER method of DOGE. The result shows that trend of the predicted value is similar to the original one, but it has 49.64 percent accuracy. A little phase shifting caused it. The phase-shifting problem was causing the decreasing value of accuracy. Furthermore, we cannot confirm that one model is better than another by using our accuracy calculation method.

### **Future Work Suggestion**

To further improve the accuracy of the machine learning algorithm, eliminating errors in this experiment will help. An improvement for three main errors

is using a newer dataset, reducing features, and exploring more evaluation methods. First, the dataset's timeframe should be more up to date so that all cryptocurrency coins selected are widely used. This will prevent the linear dataset problem. Second, to eliminate underfitting circumstances, using fewer features is one of the solutions by plotting the correlation heat map and selecting only high correlation features during the exploratory data analysis session. To be more complicated selecting features, a significance test should be performed as one test for selecting features as (Horel & Giesecke, 2020) stated. Last, a more mathematical evaluation method can improve the evaluation of phase-shifting models. RMSE or Root Mean Square Error is an example of a mathematical evaluation method. By evaluating these errors, this adjustment will increase accuracy, or it might be possible to significantly see the differences between each model.



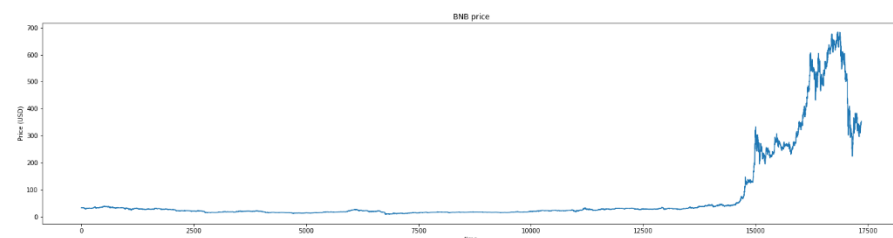
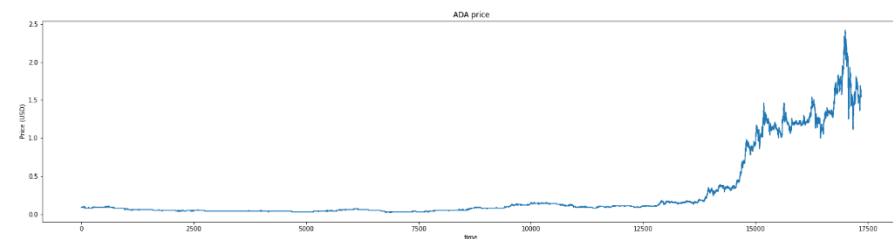
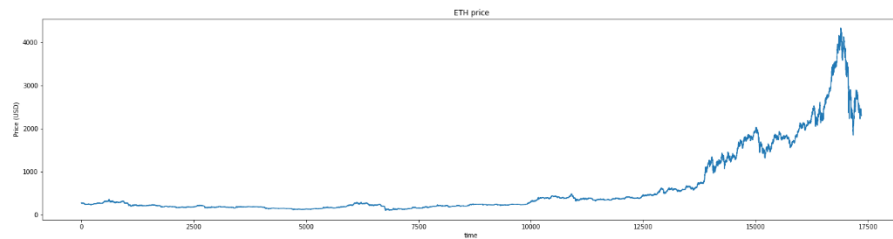
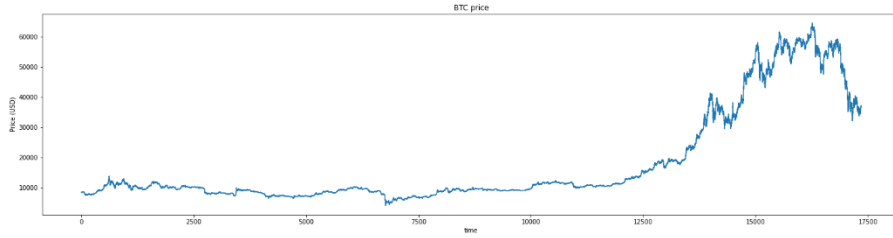
## REFERENCES

- Armstrong, J. S. (2001). Combining forecasts. In *Principles of forecasting* (pp. 417-439). Springer.
- Charles, A., & Darné, O. (2019). Volatility estimation for Bitcoin: Replication and robustness. *International Economics*, 157, 23-32.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining,
- Choudhry, R., & Garg, K. (2008). A hybrid machine learning system for stock market forecasting. *World Academy of Science, Engineering and Technology*, 39(3), 315-318.
- Chuen, D. L. K., Guo, L., & Wang, Y. (2017). Cryptocurrency: A new investment opportunity? *The Journal of Alternative Investments*, 20(3), 16-40.
- Colianni, S., Rosales, S., & Signorotti, M. (2015). Algorithmic trading of cryptocurrency based on Twitter sentiment analysis. *CS229 Project*, 1-5.
- Dash, R., & Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The Journal of Finance and Data Science*, 2(1), 42-57.
- Horel, E., & Giesecke, K. (2020). Significance tests for neural networks. *Journal of Machine Learning Research*, 21(227), 1-29.
- Ince, H., & Trafalis, T. B. (2008). Short term forecasting with support vector machines and application to stock price prediction. *International Journal of General Systems*, 37(6), 677-687.
- Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters*, 158, 3-6.
- Koker, T. E., & Koutmos, D. (2020). Cryptocurrency Trading Using Machine Learning. *Journal of Risk and Financial Management*, 13(8), 178.
- Lahmiri, S., & Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons & Fractals*, 118, 35-40.
- Liu, Y., & Tsyvinski, A. (2021). Risks and returns of cryptocurrency. *The Review of Financial Studies*, 34(6), 2689-2727.
- Lyu, M. R., & Nikora, A. P. (1991). A heuristic approach for software reliability prediction: the equally-weighted linear combination model. ISSRE,
- Madan, I., Saluja, S., & Zhao, A. (2015). Automated bitcoin trading via machine learning algorithms. URL: <http://cs229.stanford.edu/proj2014/Isaac%20Madan>, 20.
- Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., & Khudanpur, S. (2010). Recurrent neural network based language model. Interspeech,
- Poyser, O. (2017). Exploring the determinants of Bitcoin's price: an application of Bayesian Structural Time Series. *arXiv preprint arXiv:1706.01437*.
- Sebastião, H., & Godinho, P. (2021). Forecasting and trading cryptocurrencies with machine learning under changing market conditions. *Financial Innovation*, 7(1), 1-30.
- Shen, S., Jiang, H., & Zhang, T. (2012). Stock market forecasting using machine learning algorithms. *Department of Electrical Engineering, Stanford University, Stanford, CA*, 1-5.

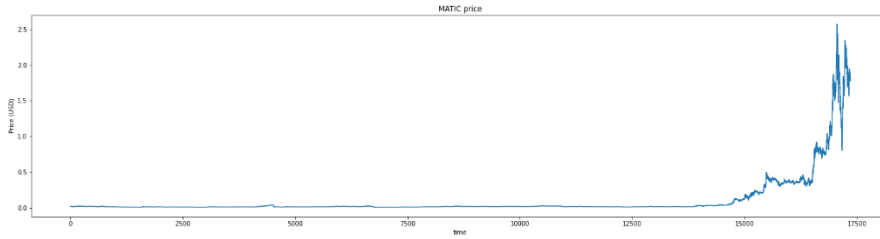
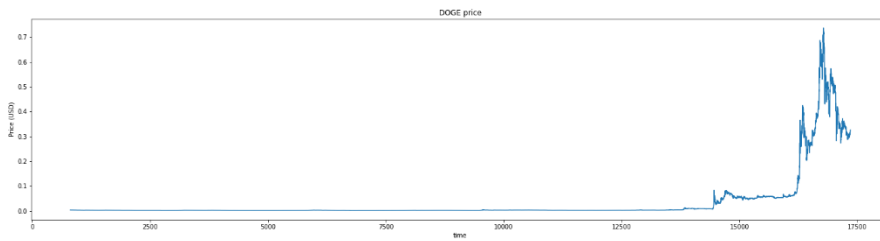
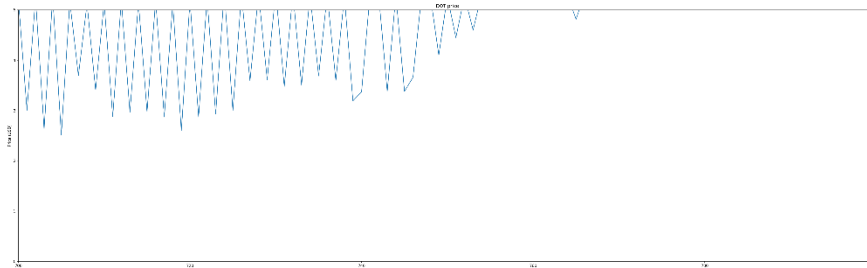
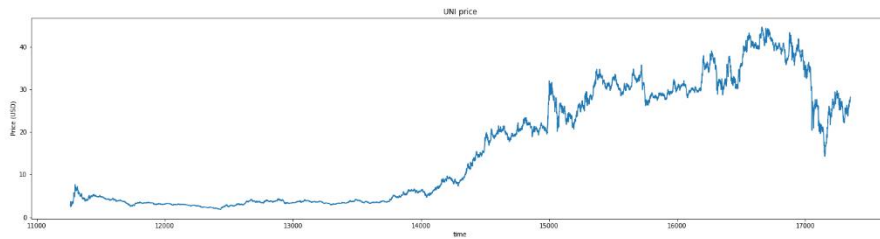
- Sovbetov, Y. (2018). Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, bitcoin, and monero. *Journal of Economics and Financial Analysis*, 2(2), 1-27.
- Strader, T. J., Rozycki, J. J., Root, T. H., & Huang, Y.-H. J. (2020). Machine learning stock market prediction studies: Review and research directions. *Journal of International Technology and Information Management*, 28(4), 63-83.
- Yang, Y. (2003). Regression with multiple candidate models: selecting or mixing? *Statistica Sinica*, 783-809.
- Żbikowski, K. (2016). Application of machine learning algorithms for bitcoin automated trading. In *Machine intelligence and big data in industry* (pp. 161-168). Springer.



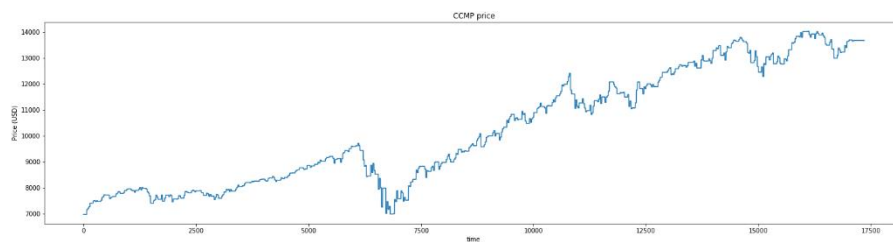
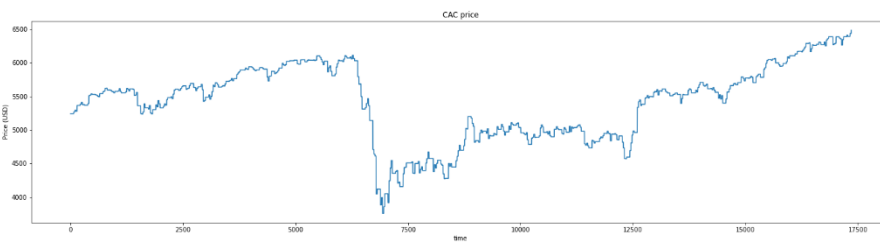
# Appendix

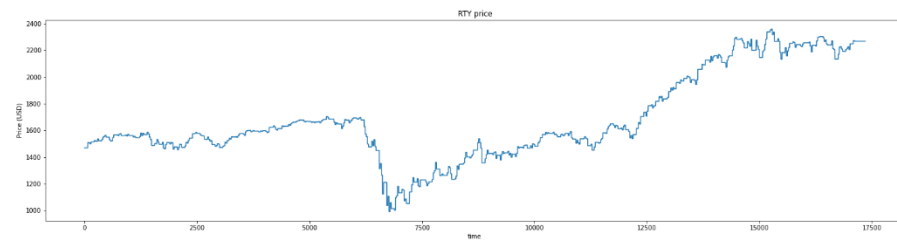
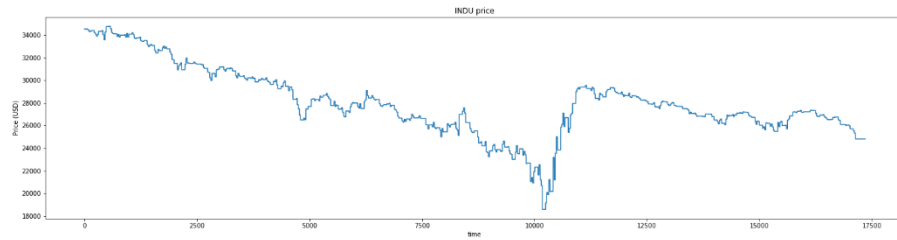
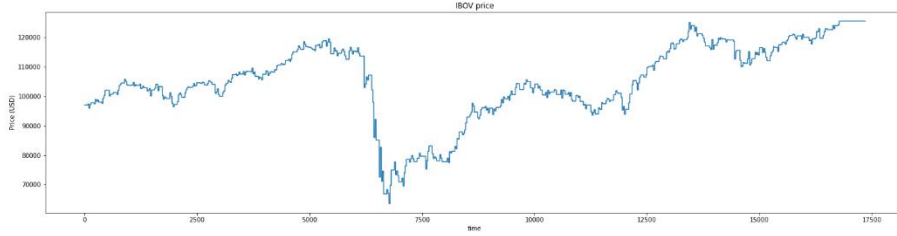
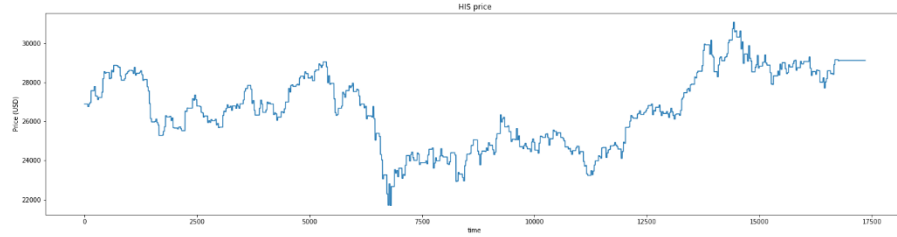
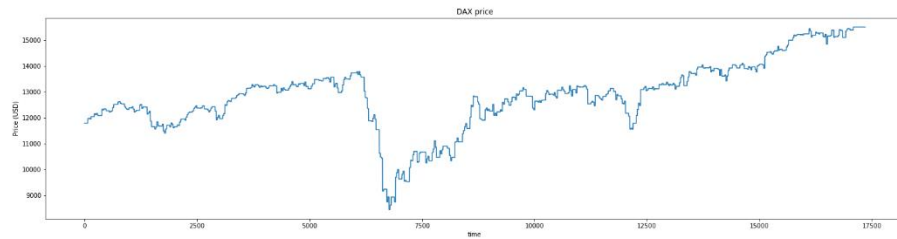


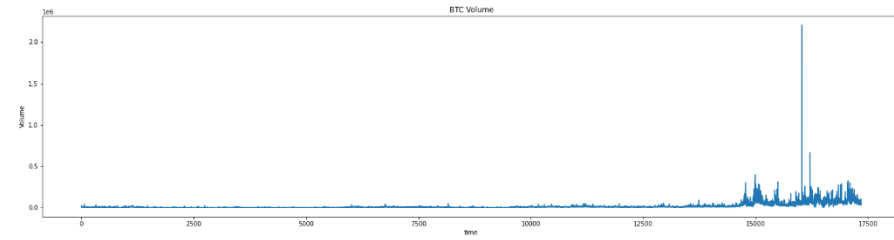
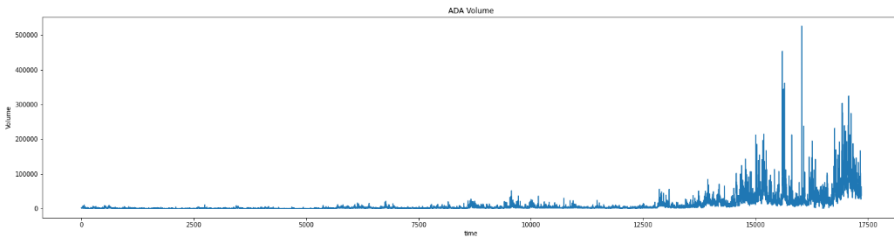
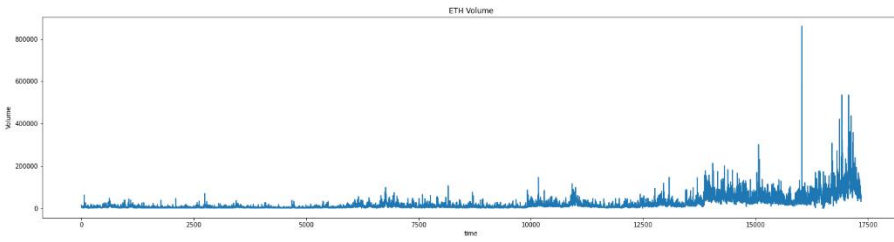
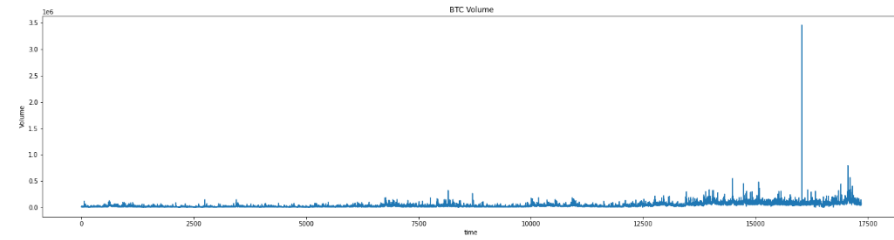
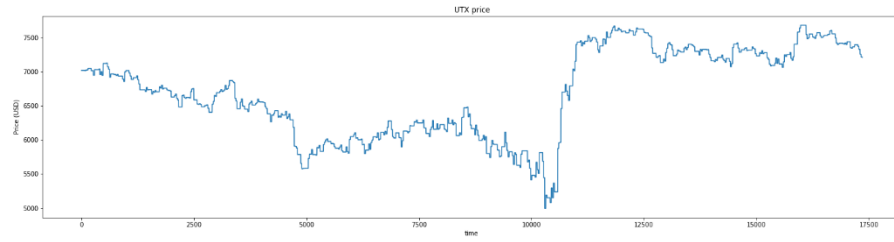
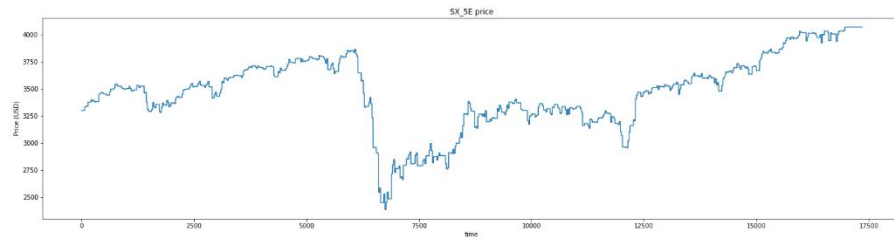


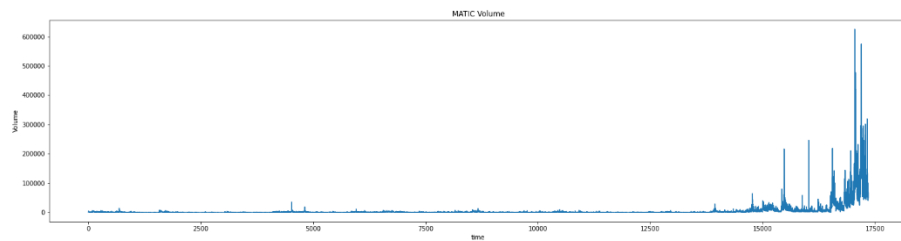
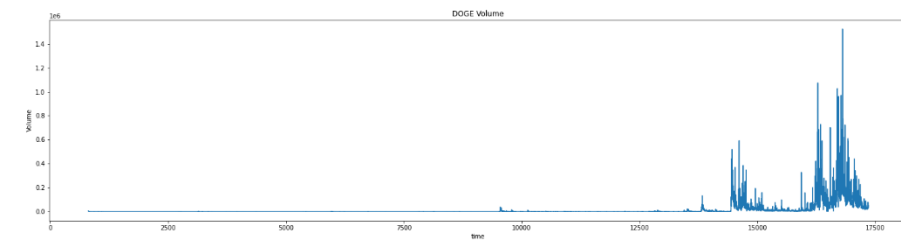
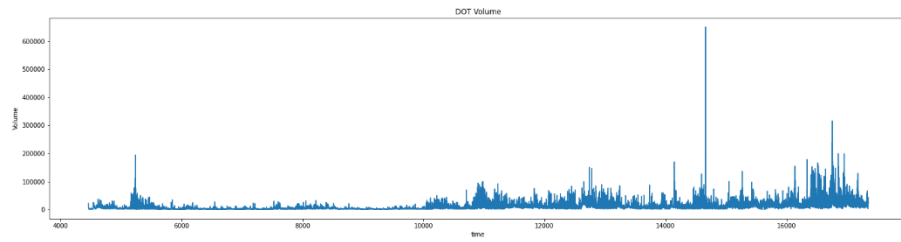
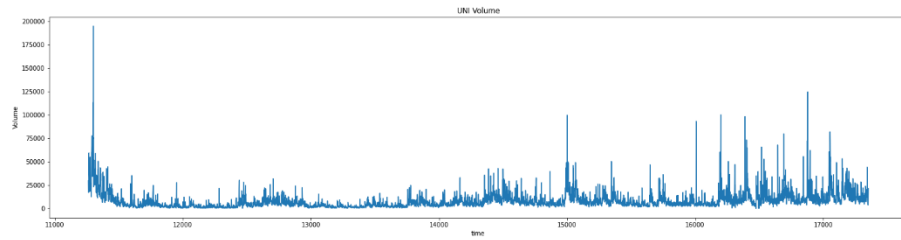
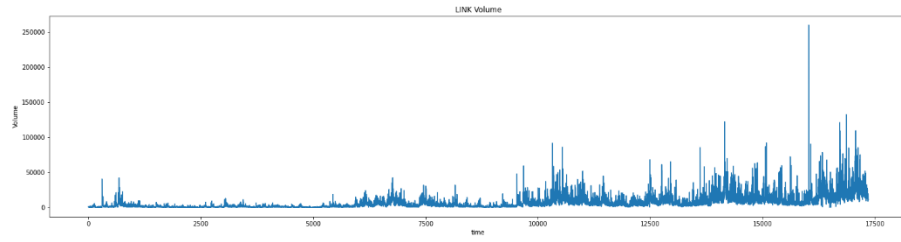
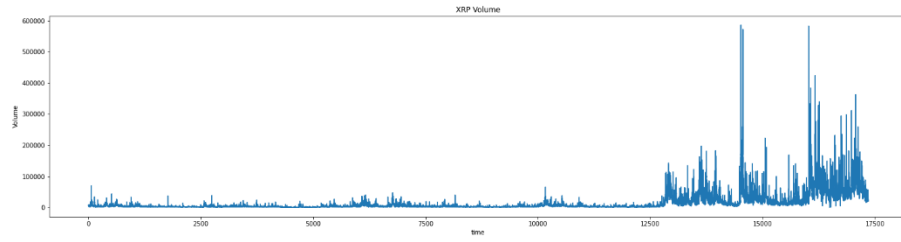


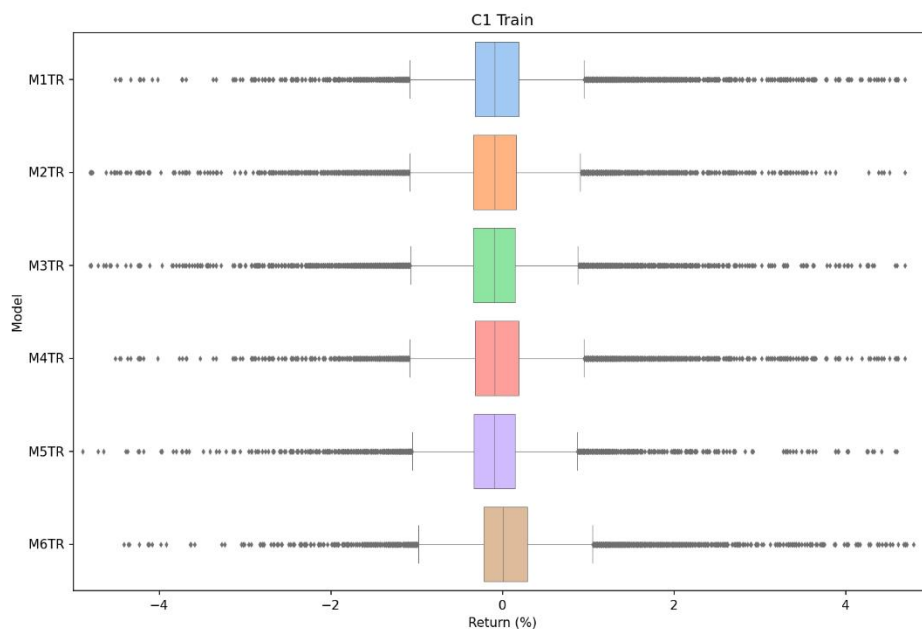
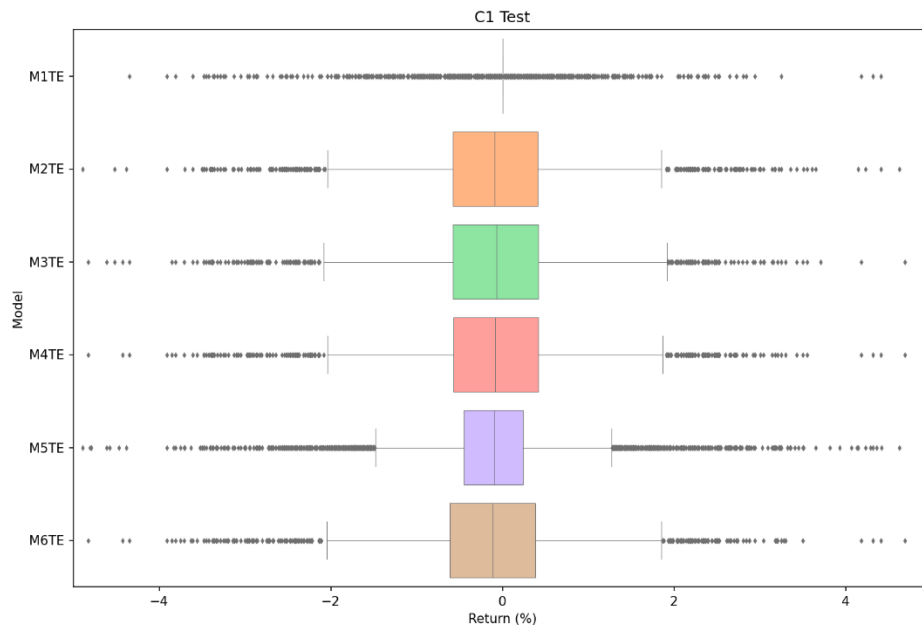
СНИИ АІ ОНСКОРН UNIVERSITY

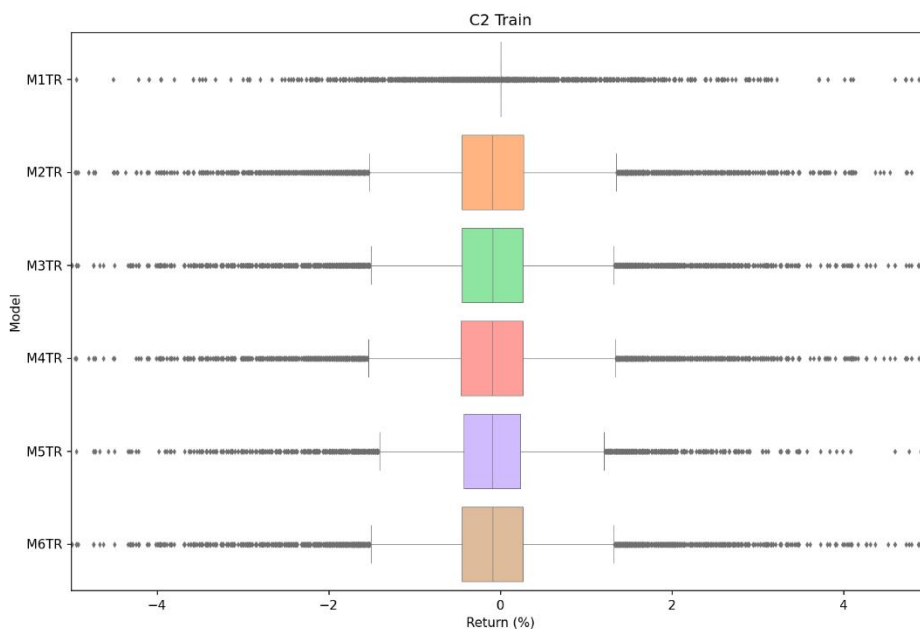
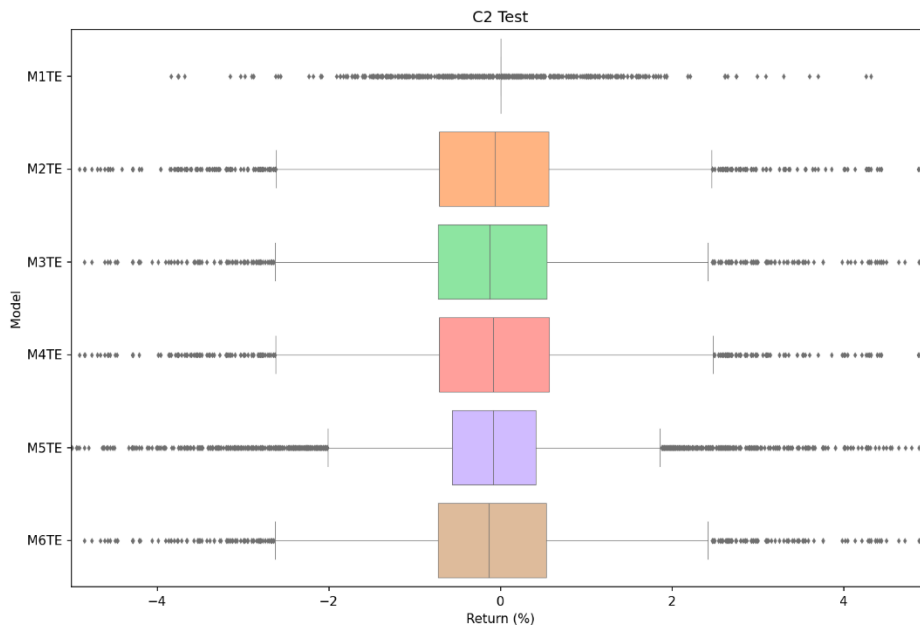


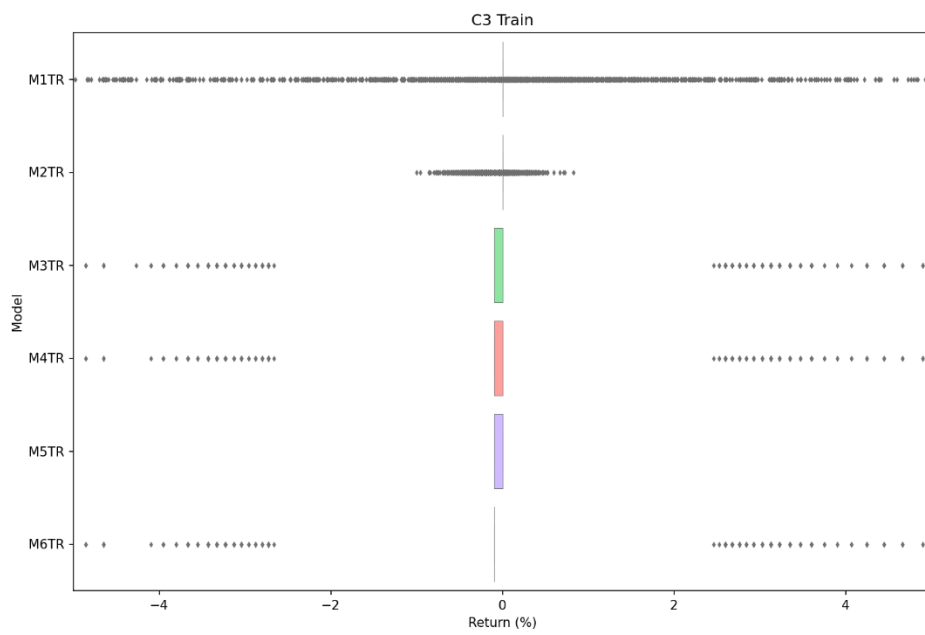
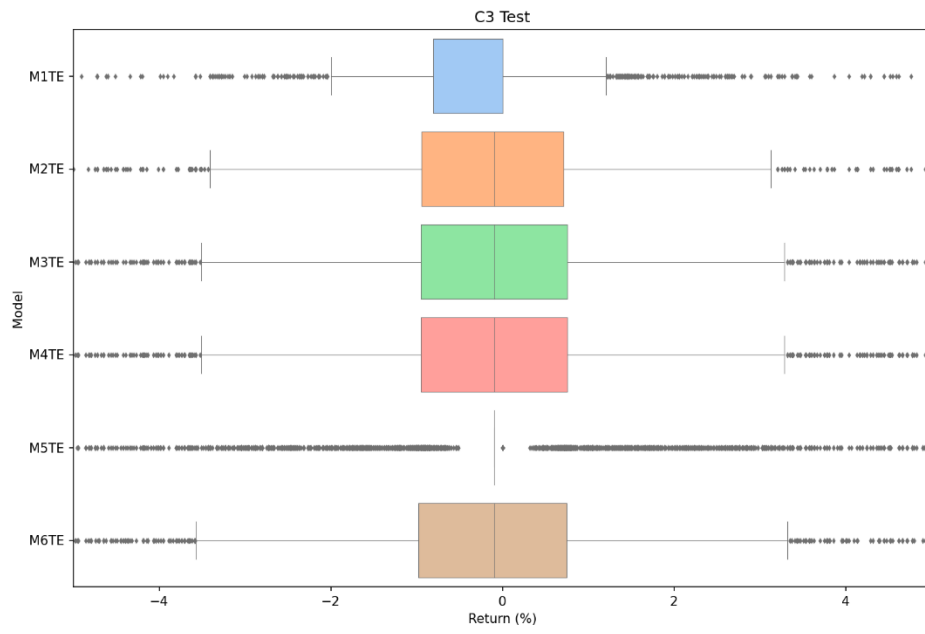


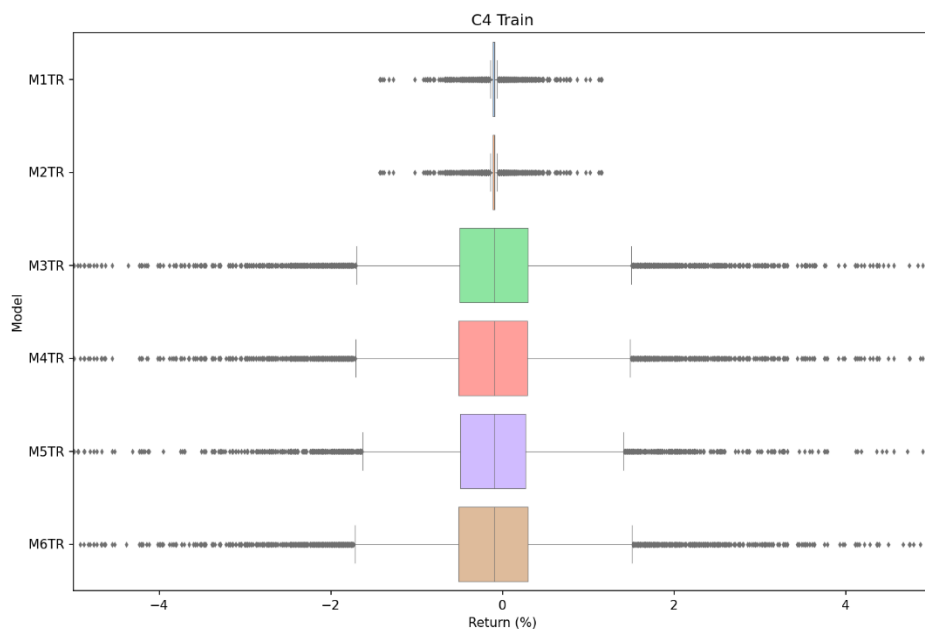
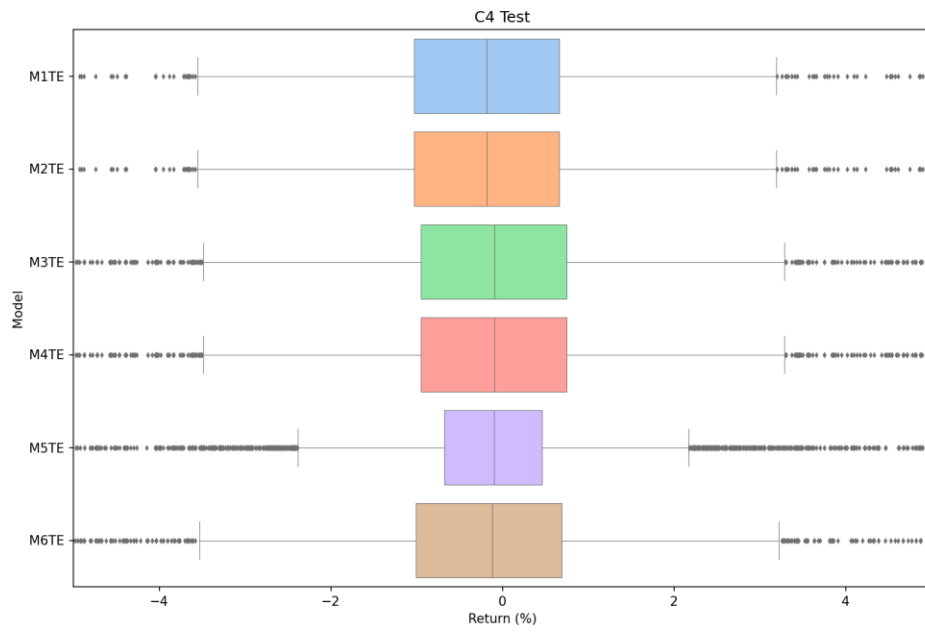




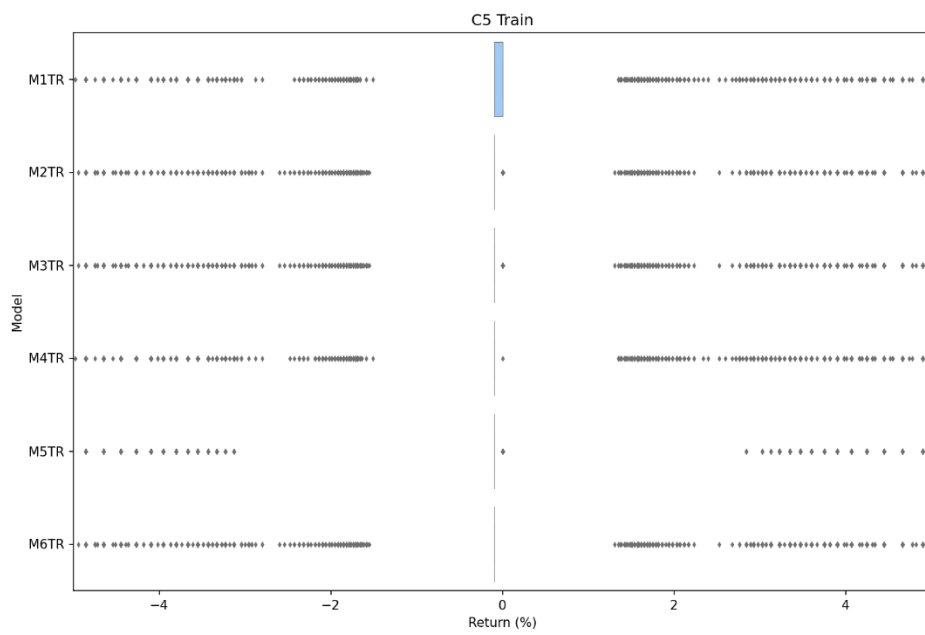
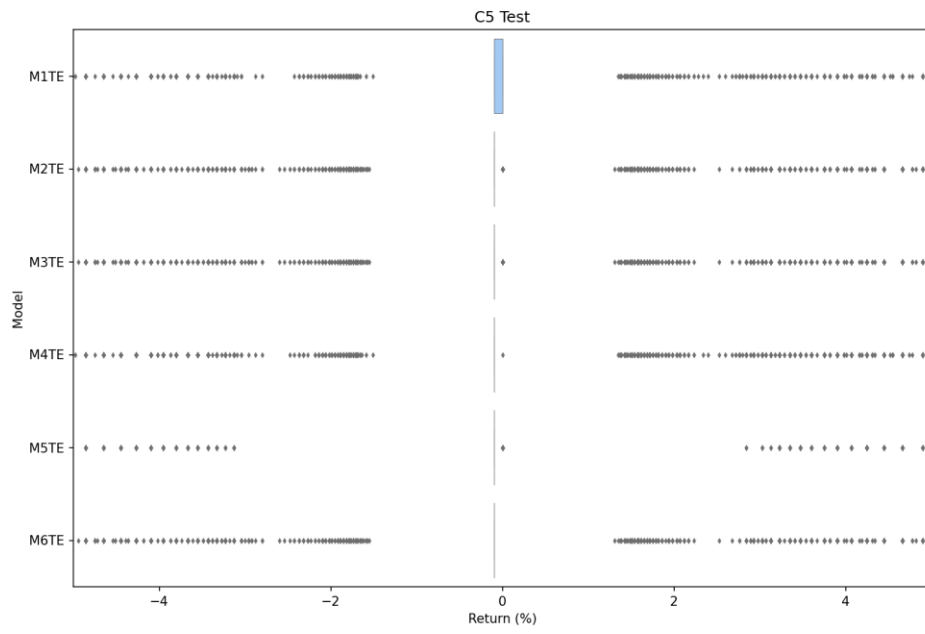


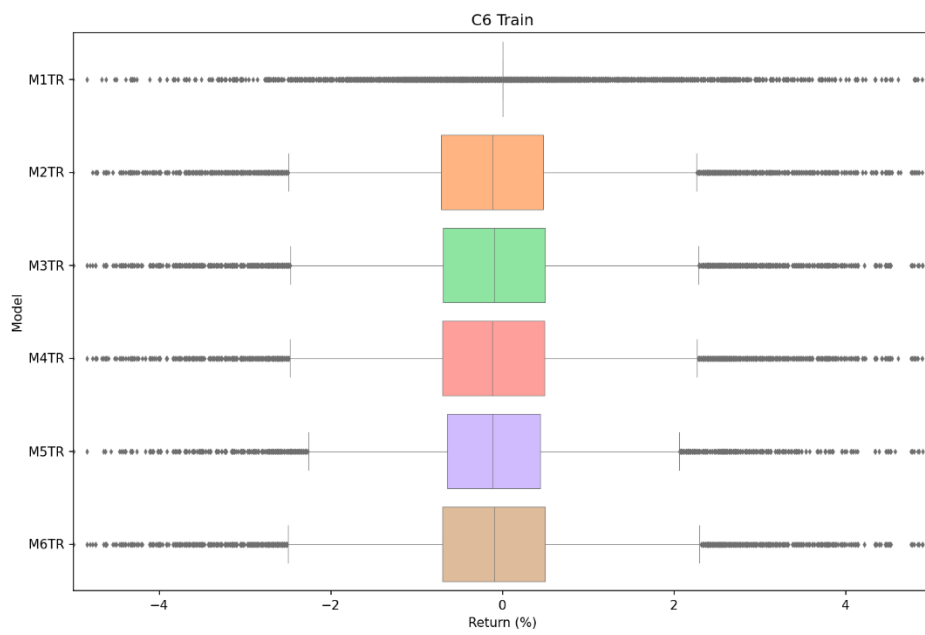
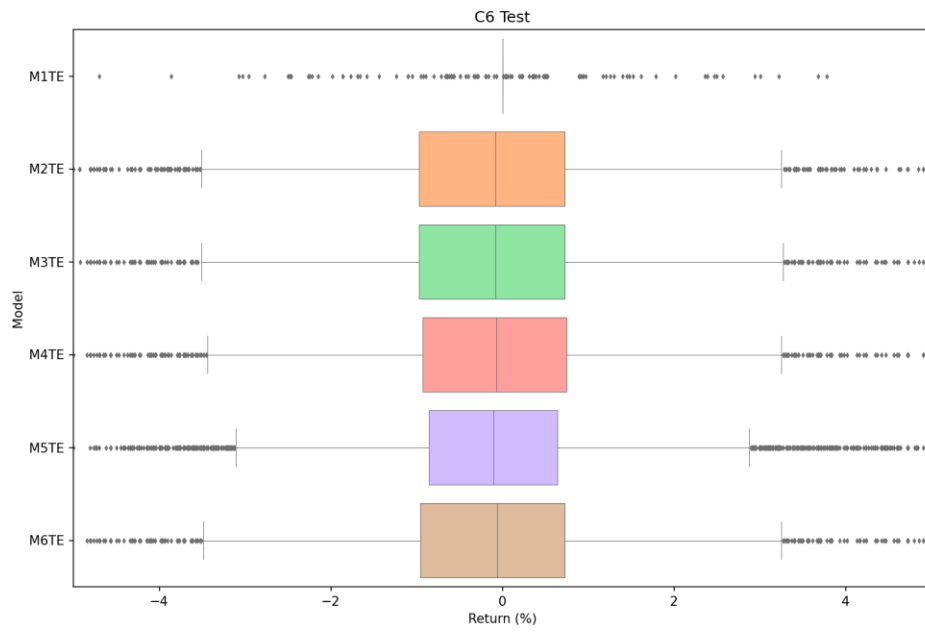


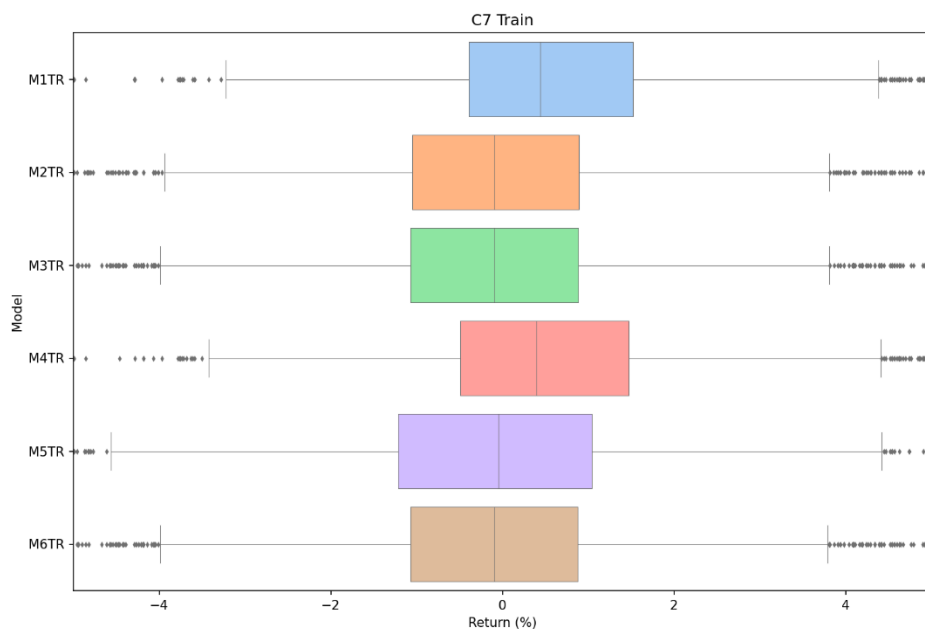
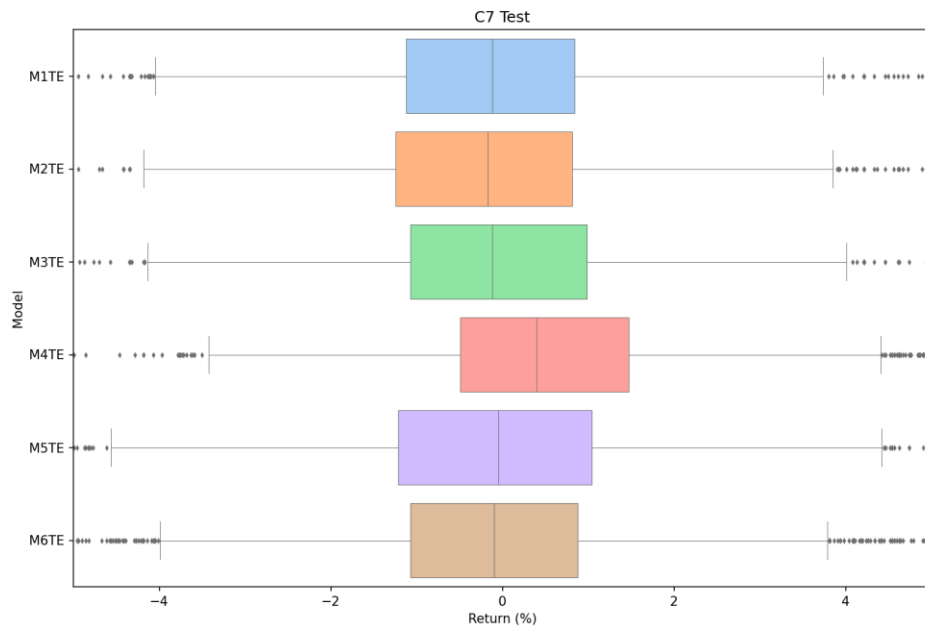


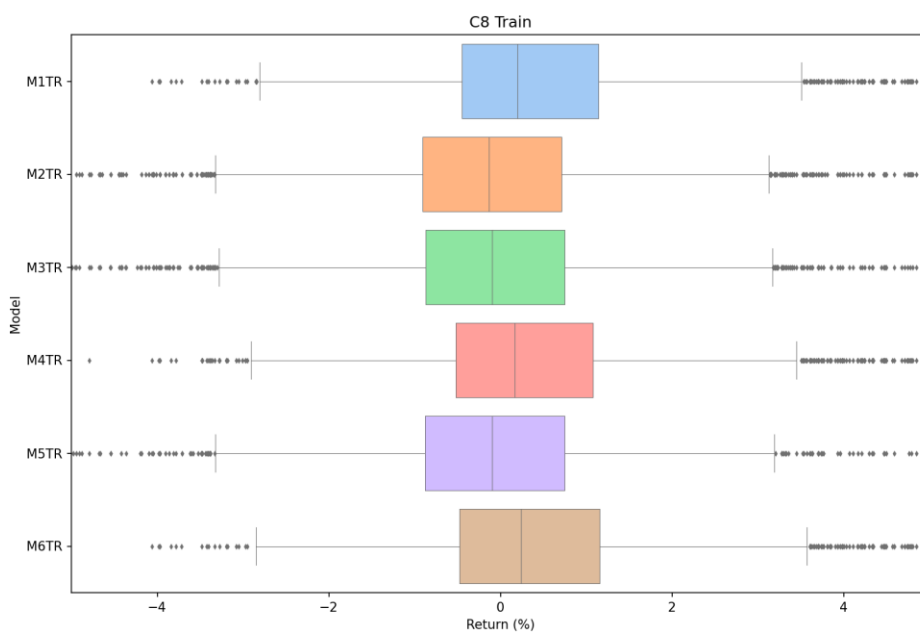
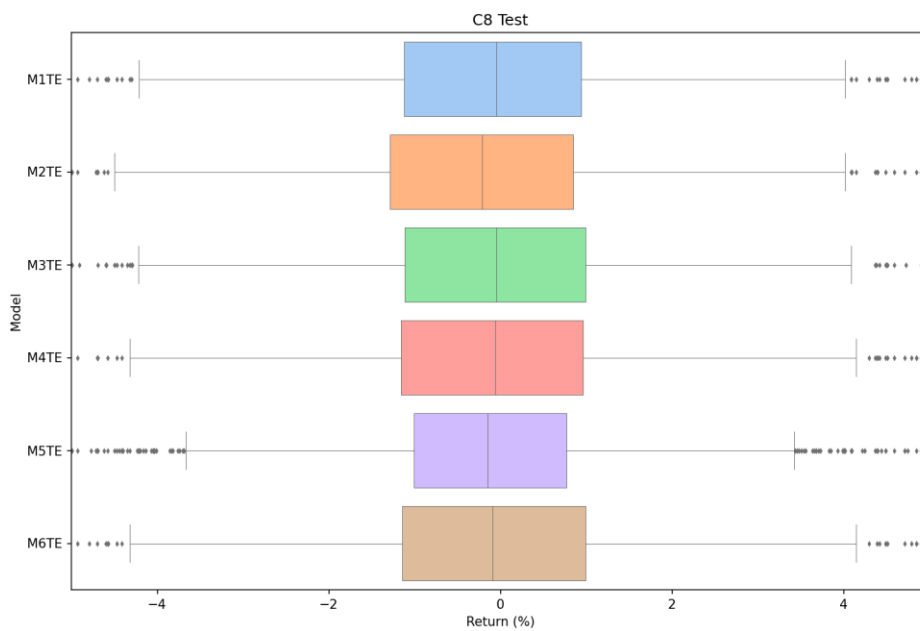


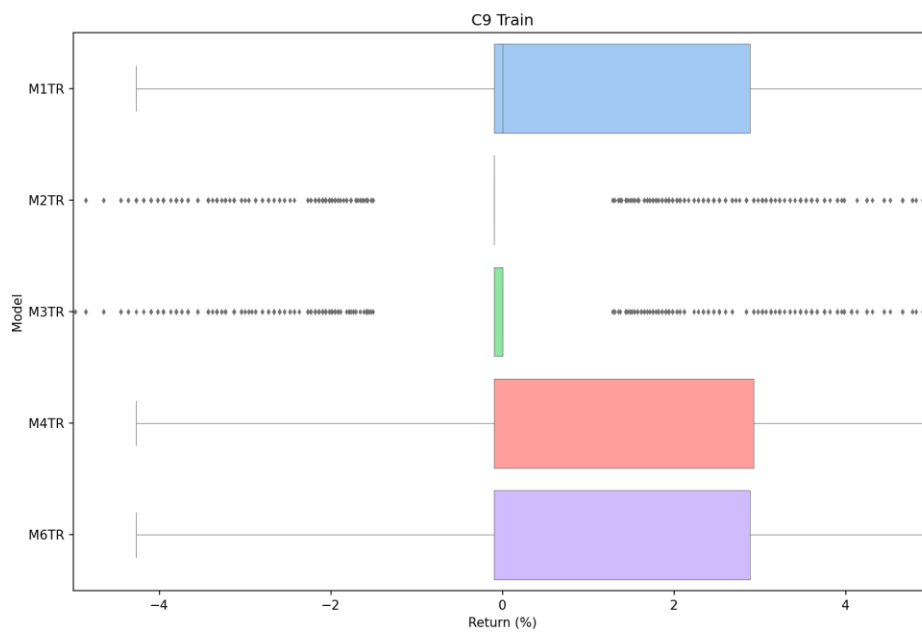
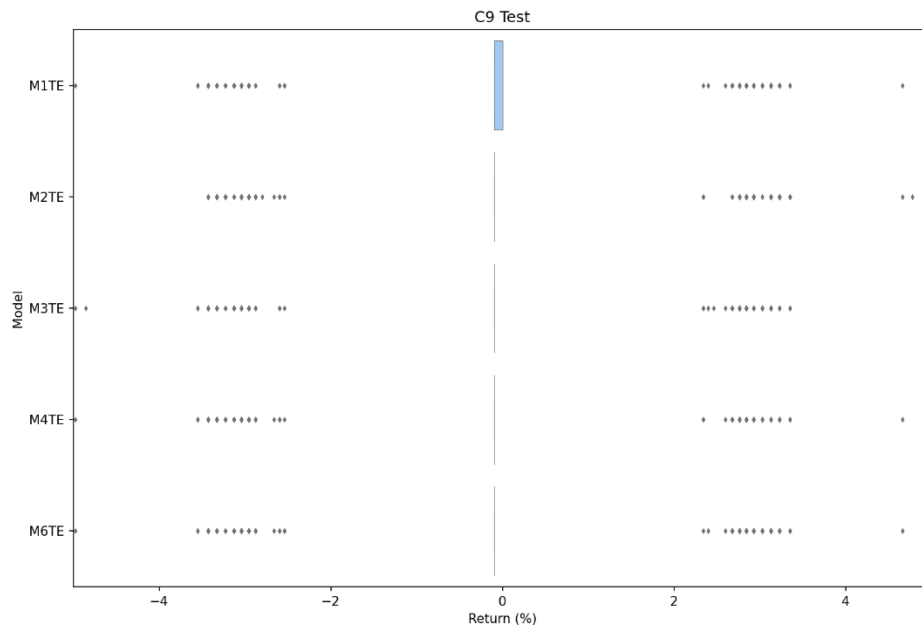


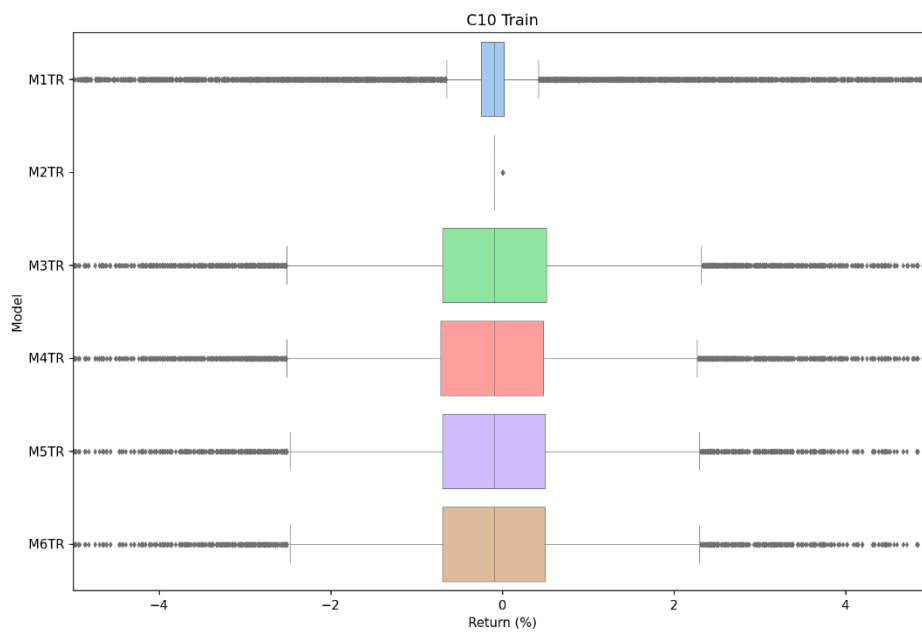
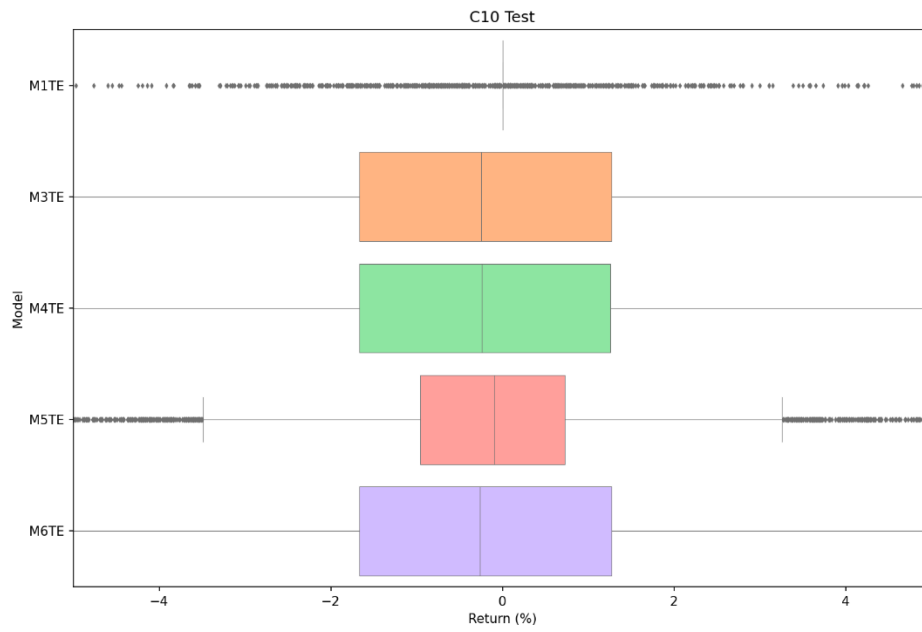


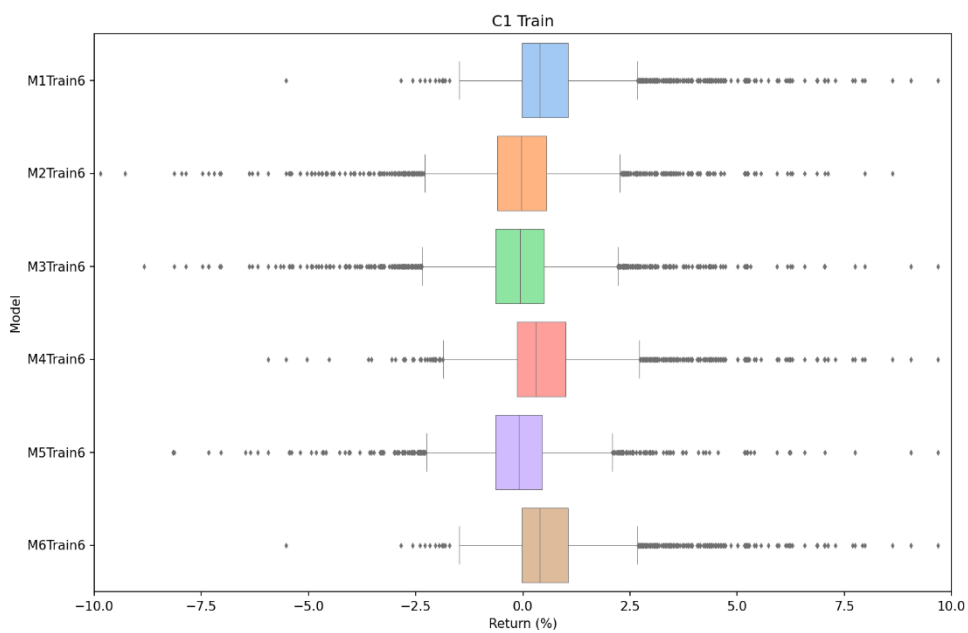
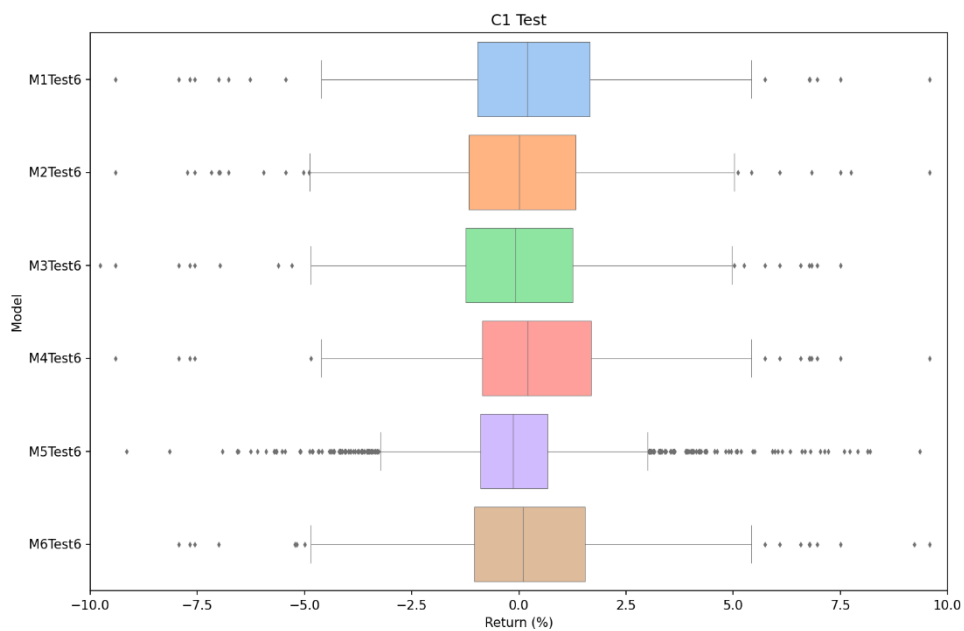


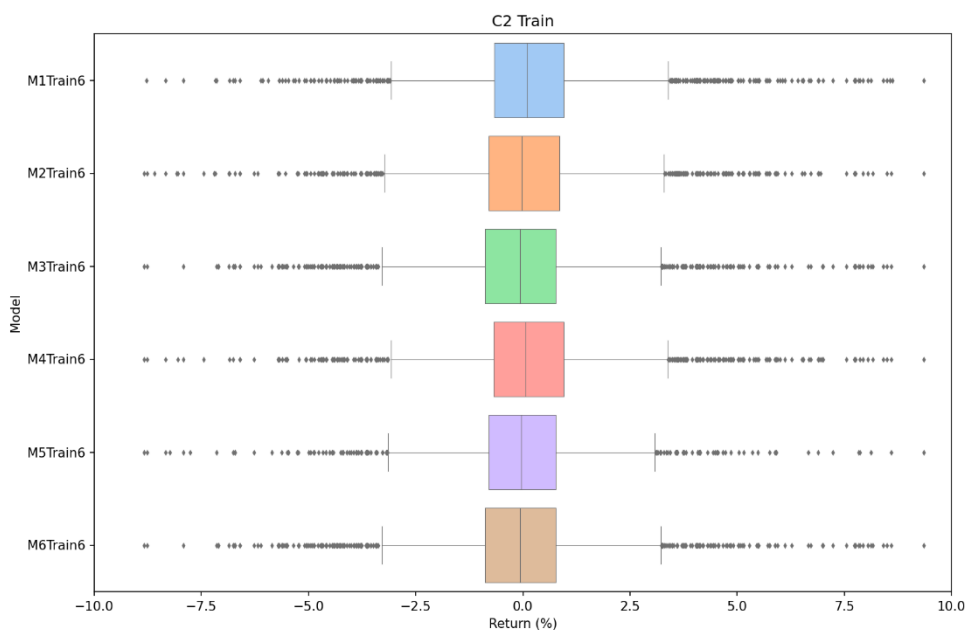
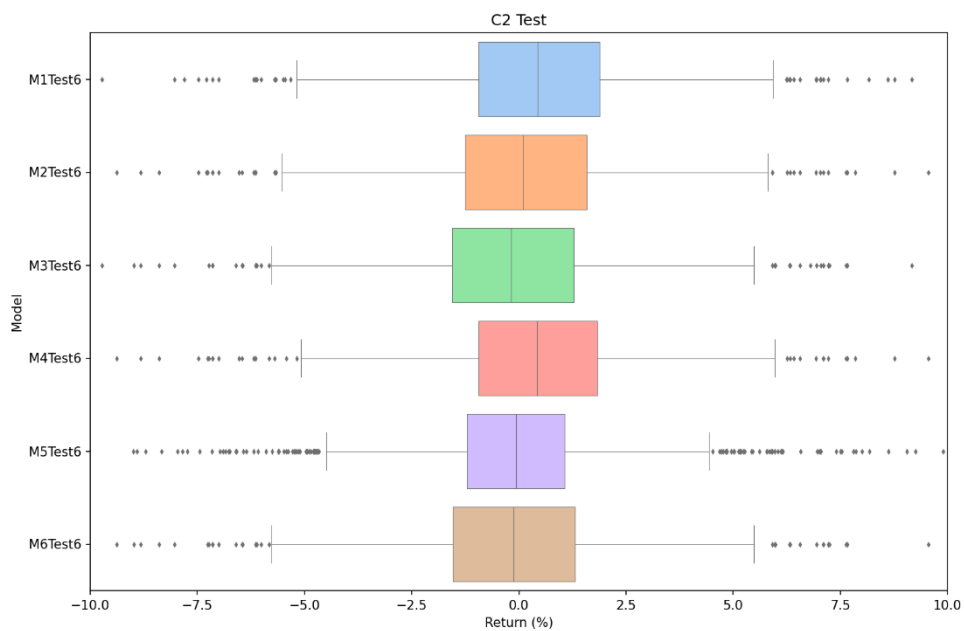




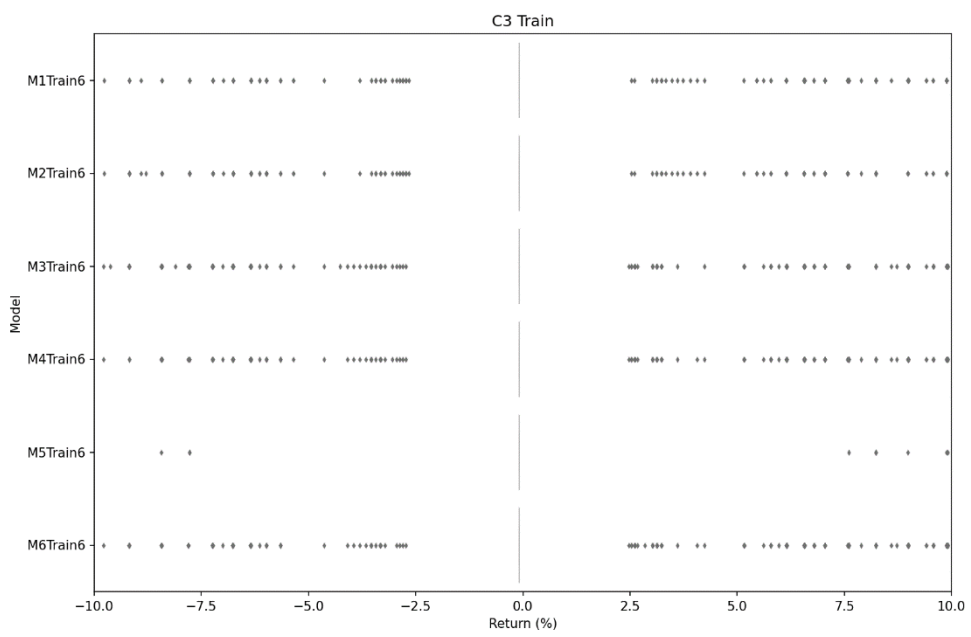
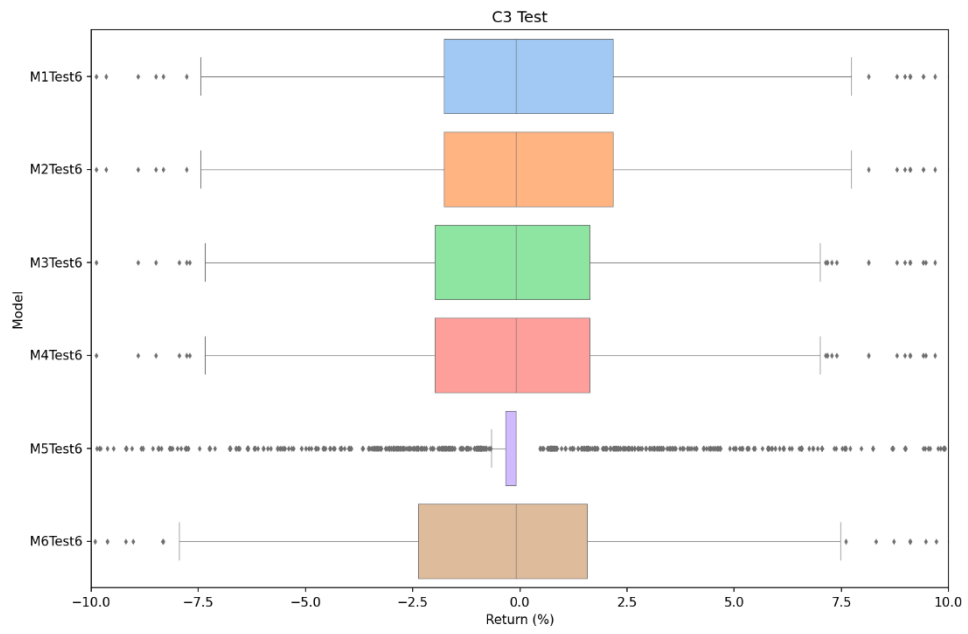


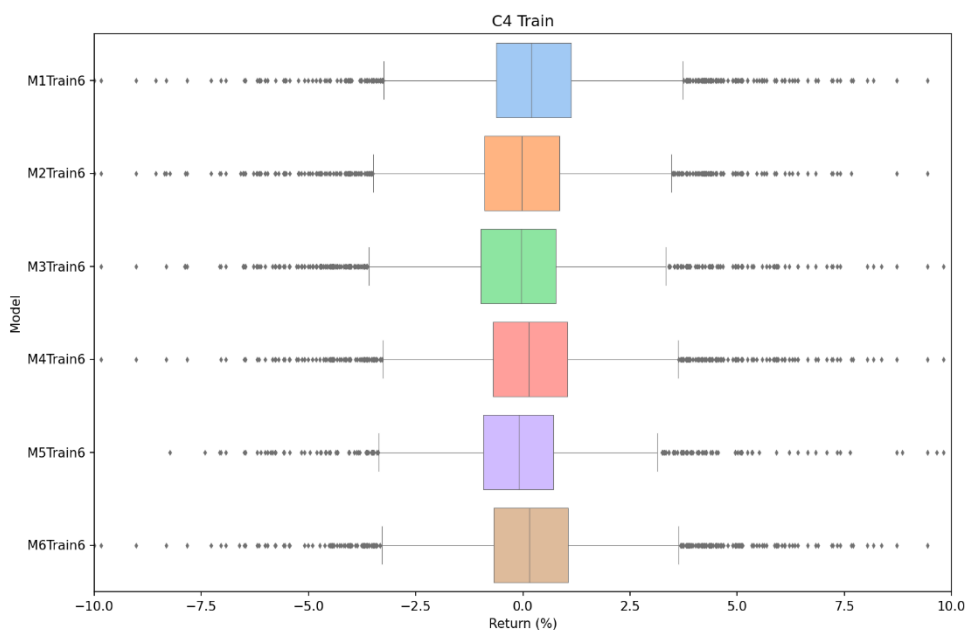
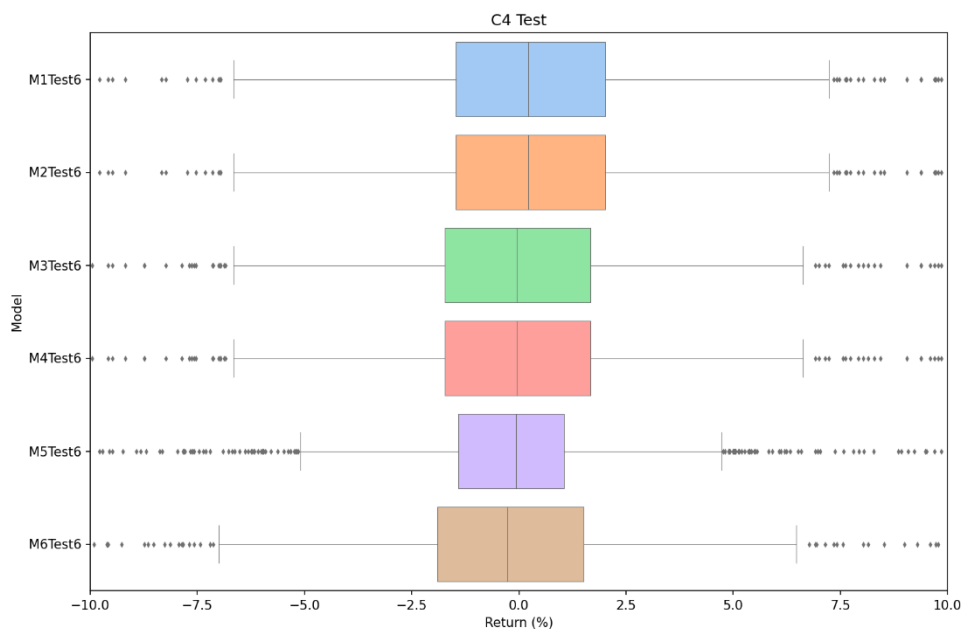


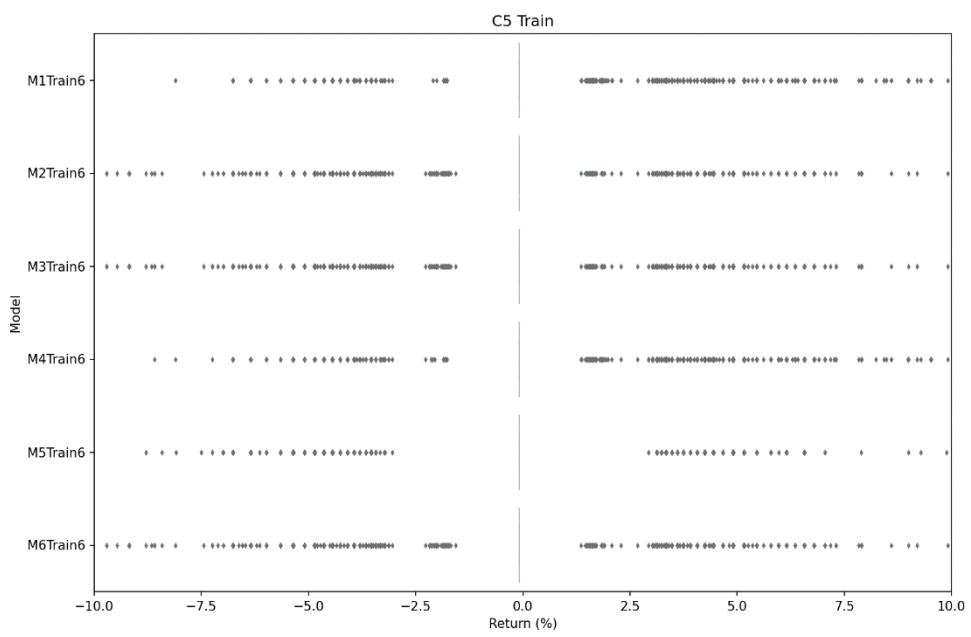
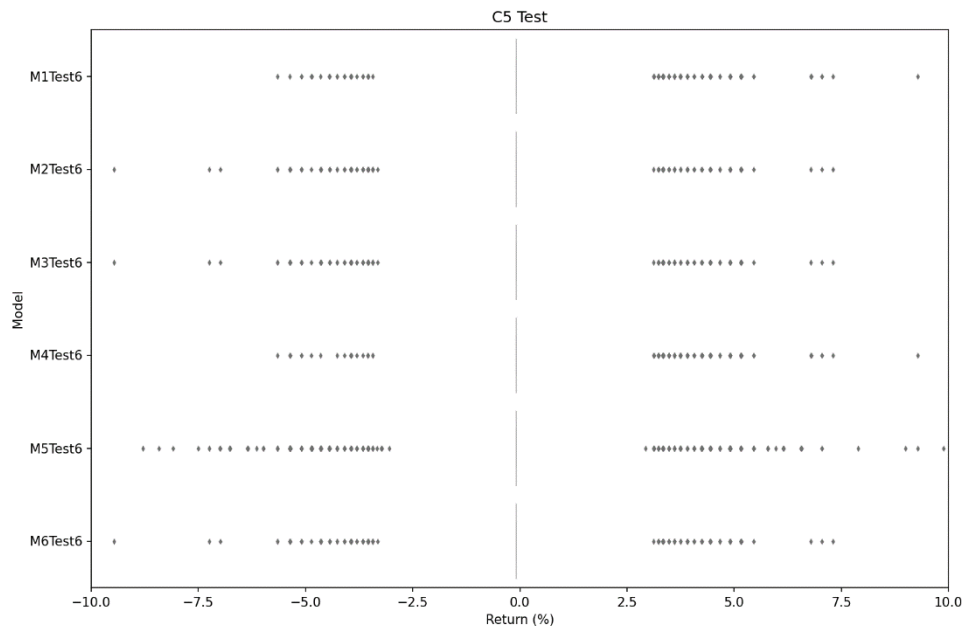


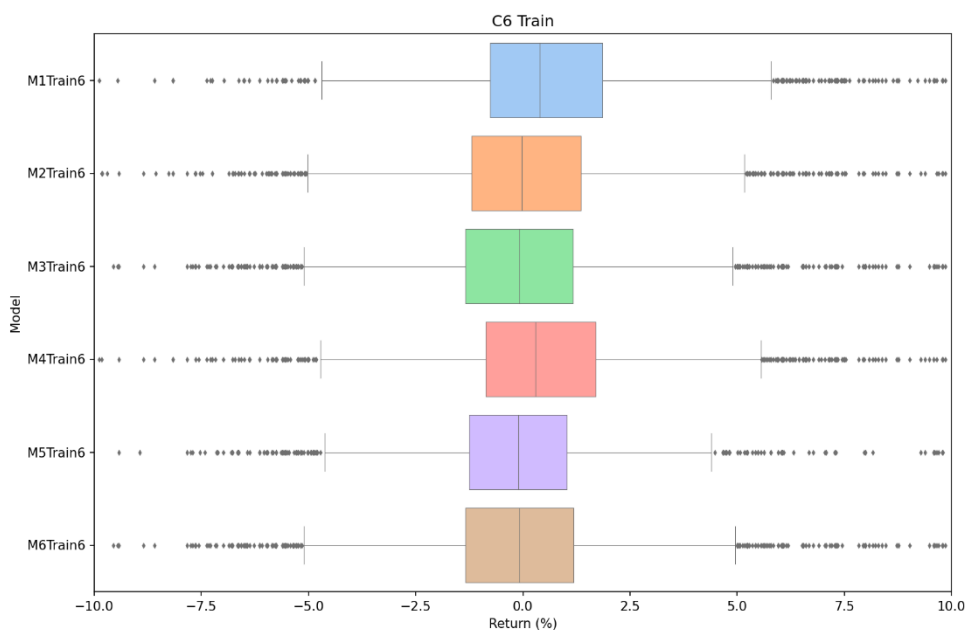
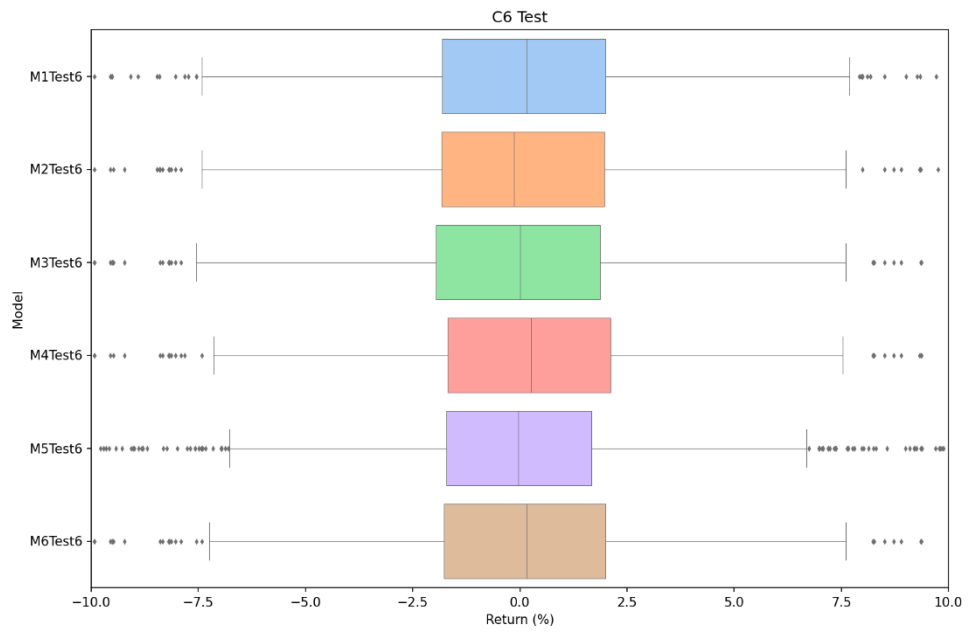


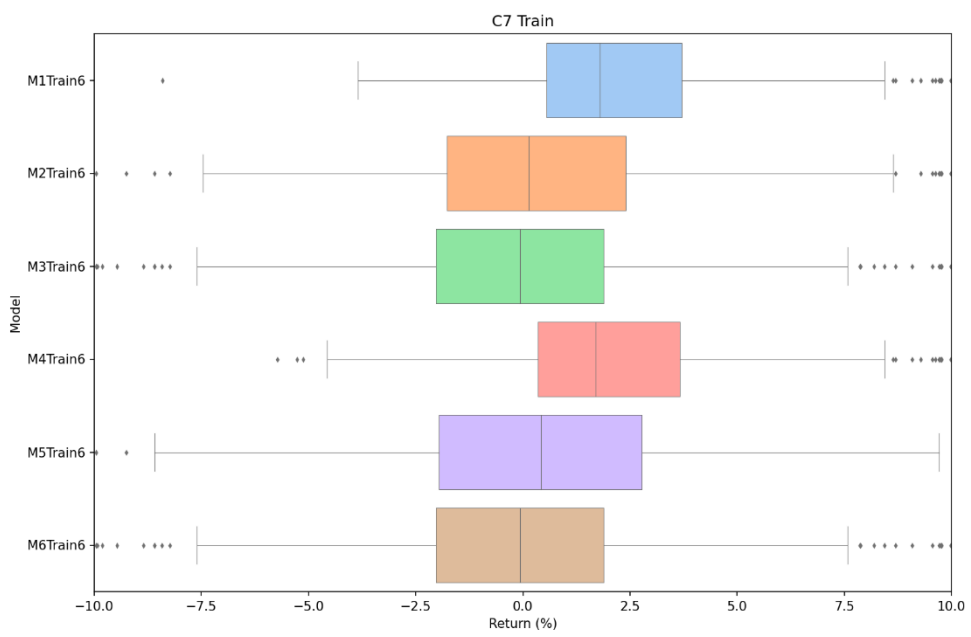
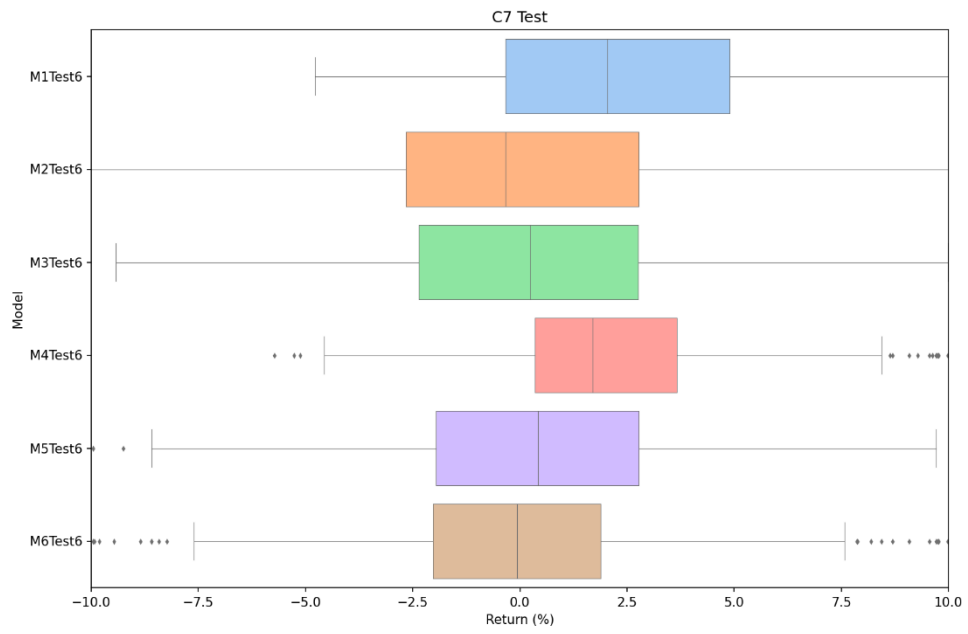


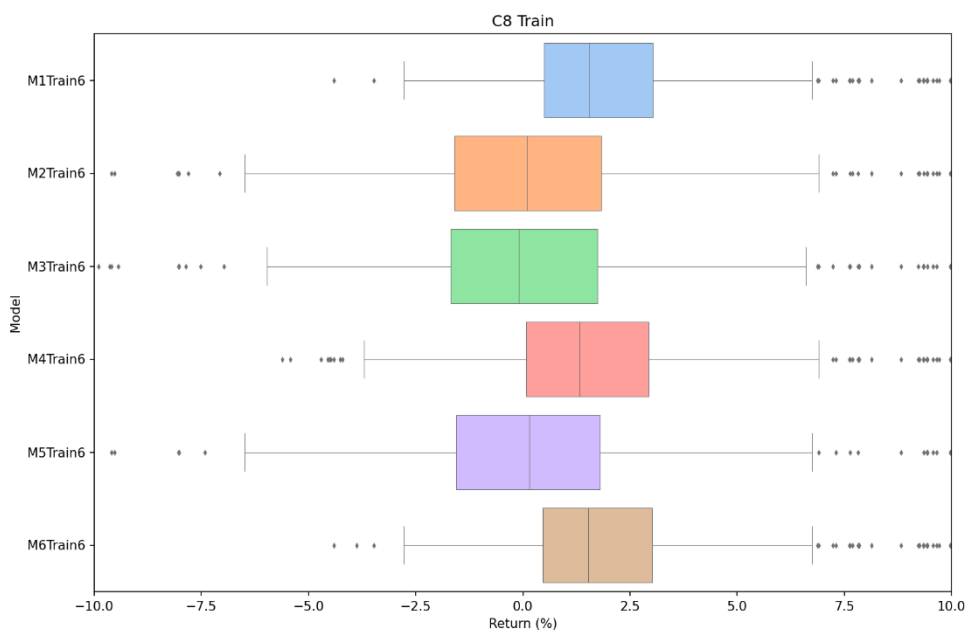
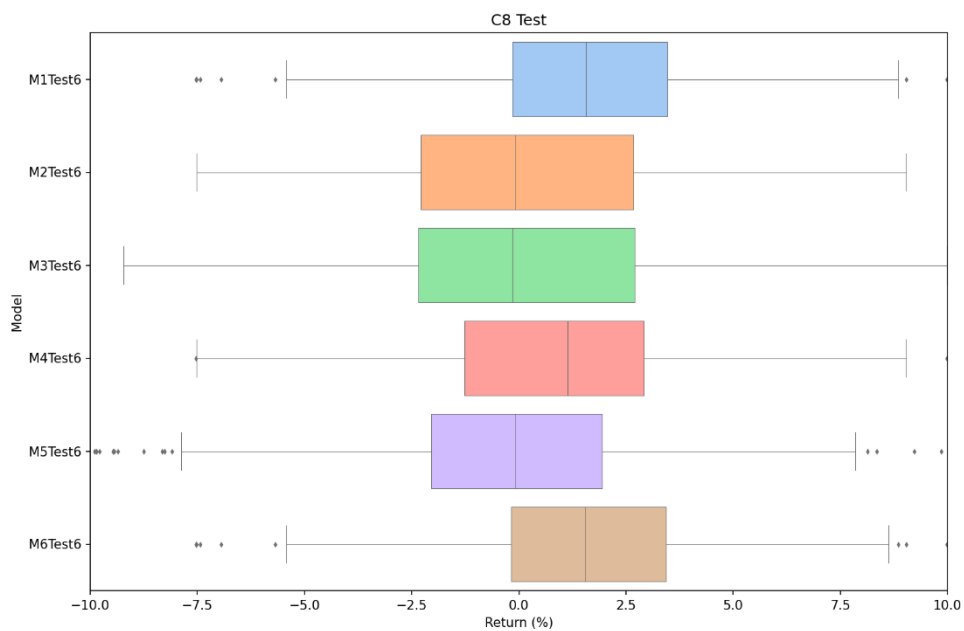


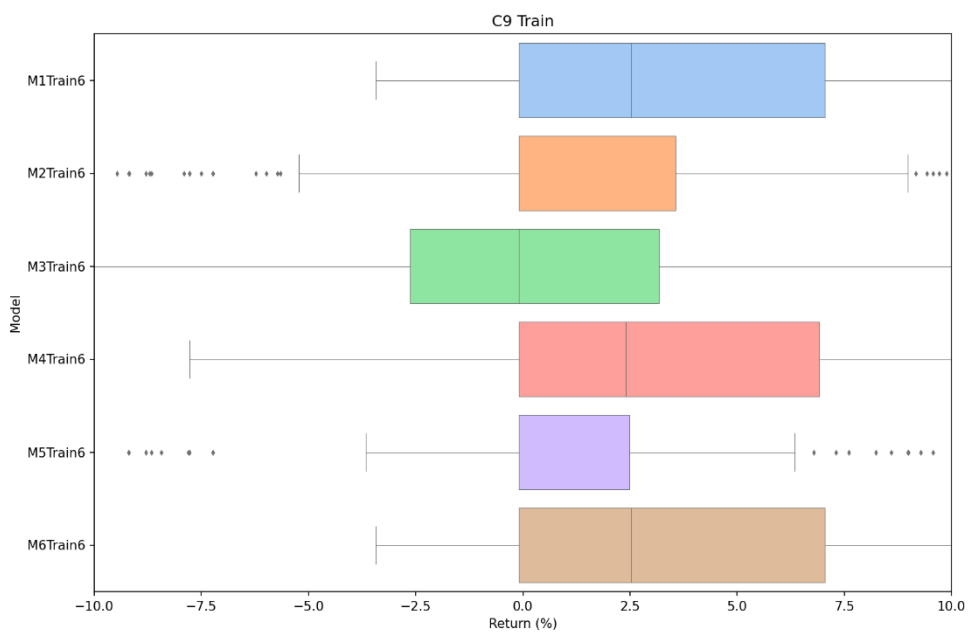
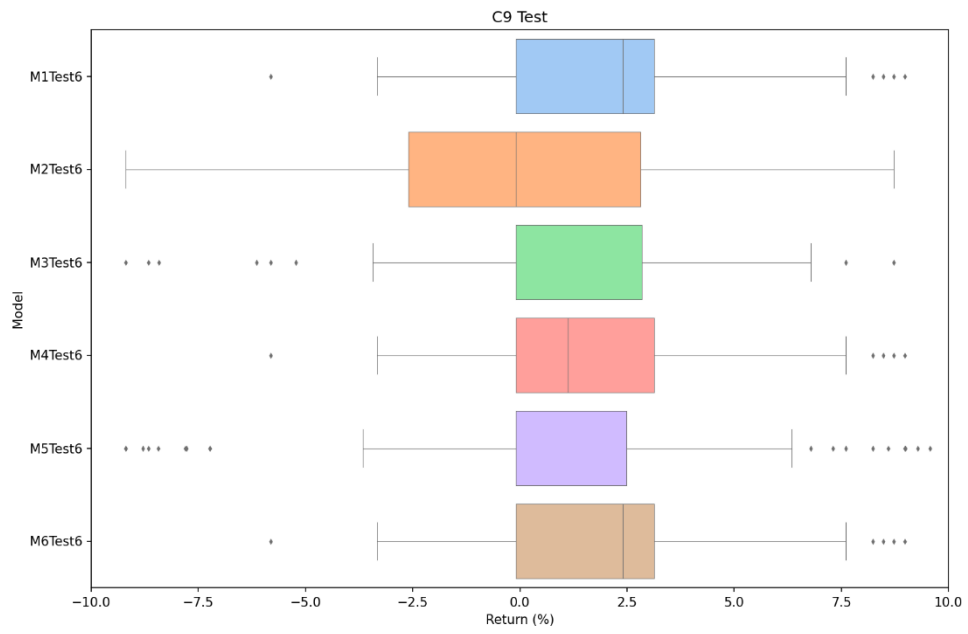


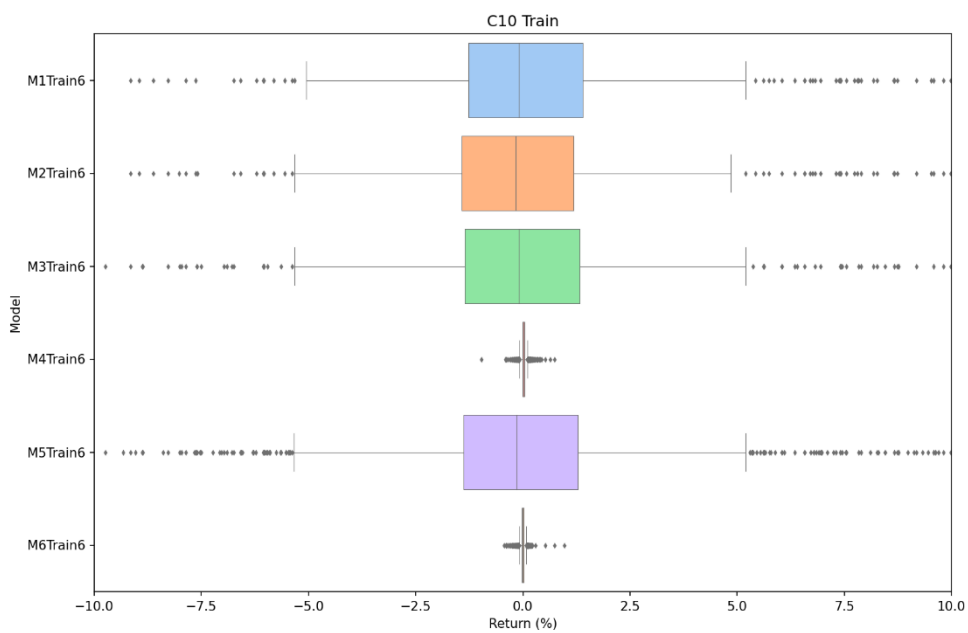
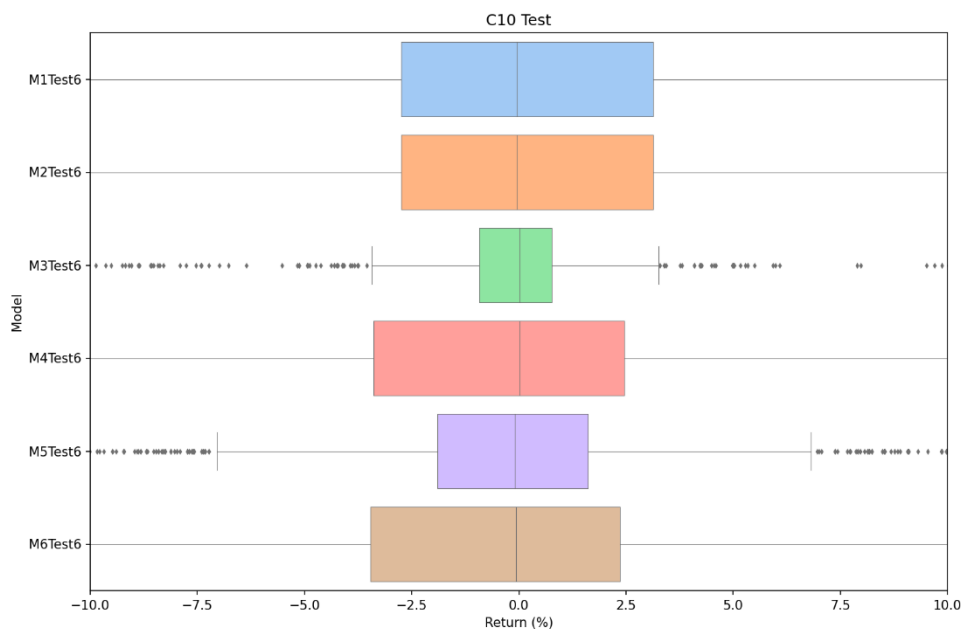




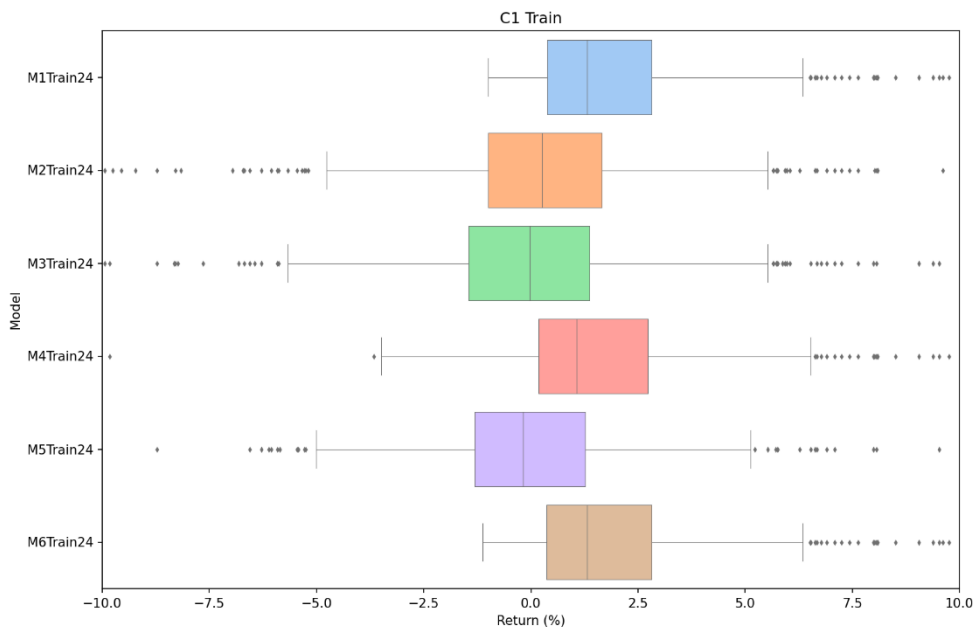
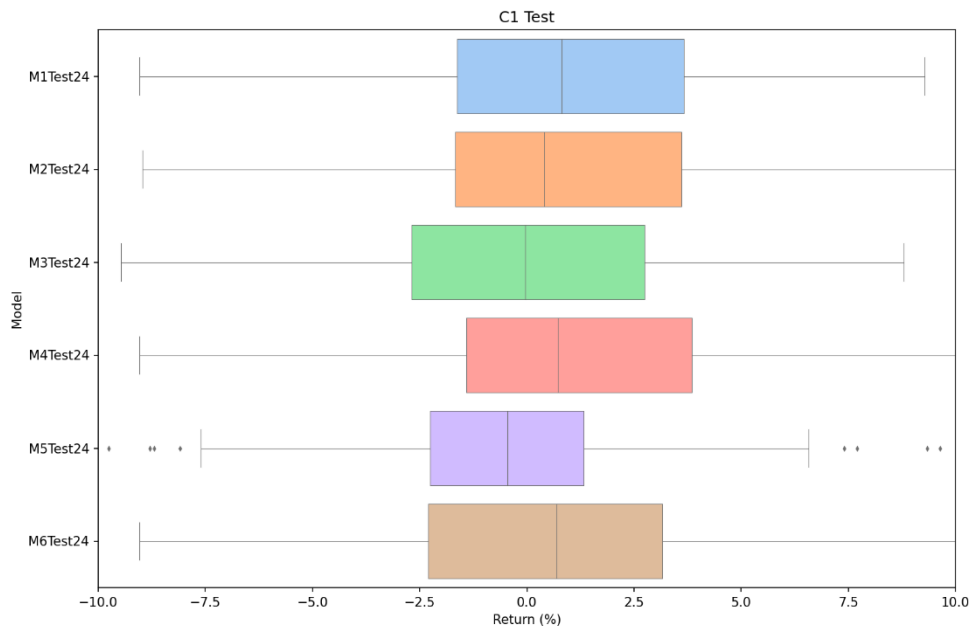


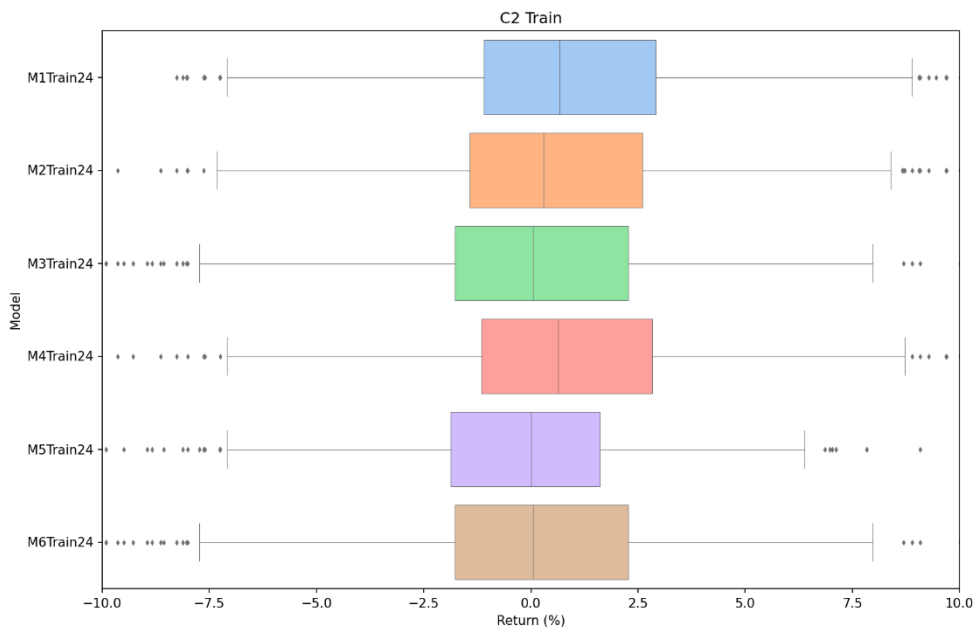
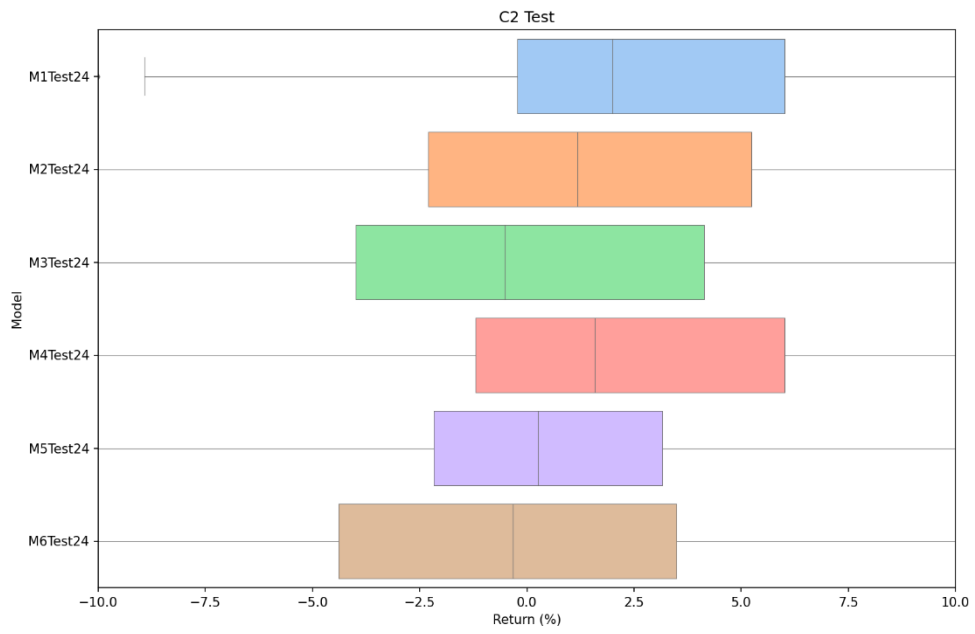


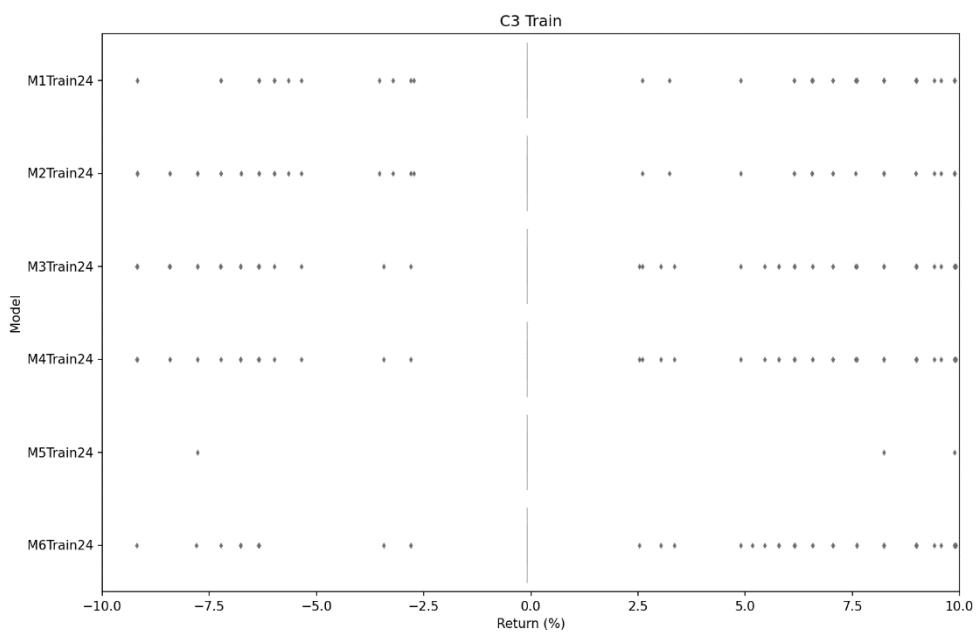
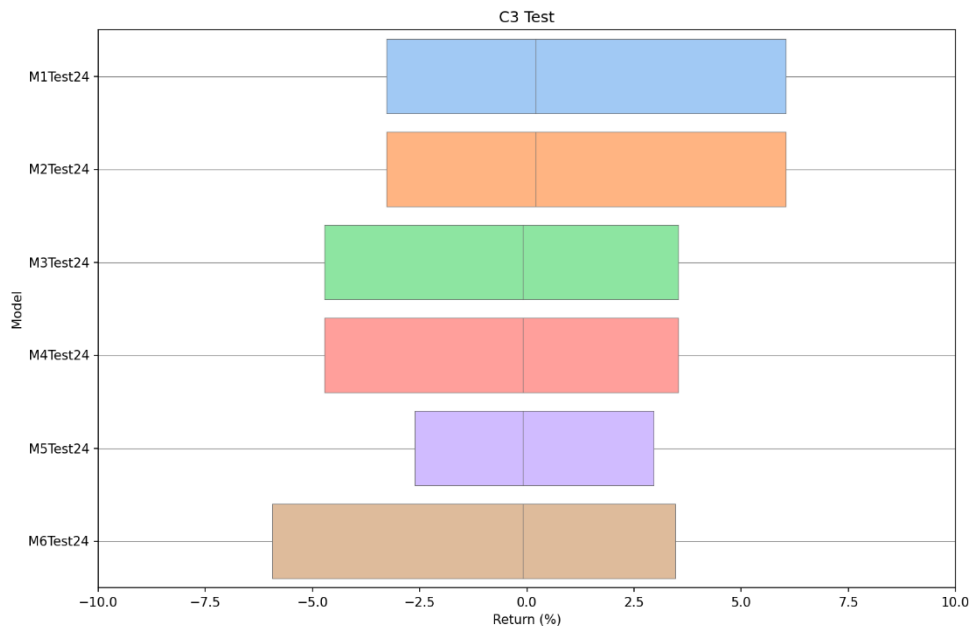


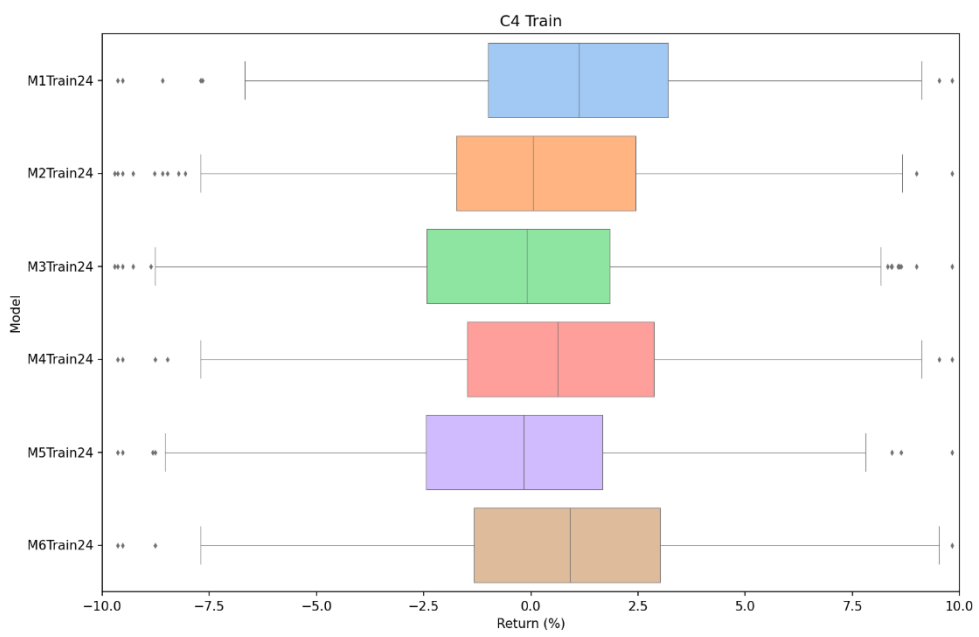
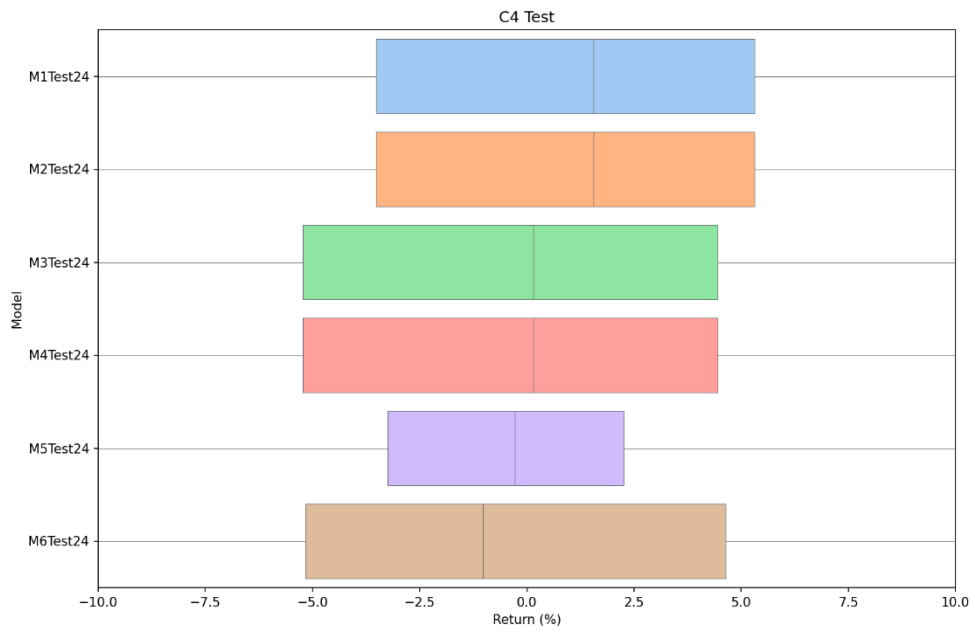


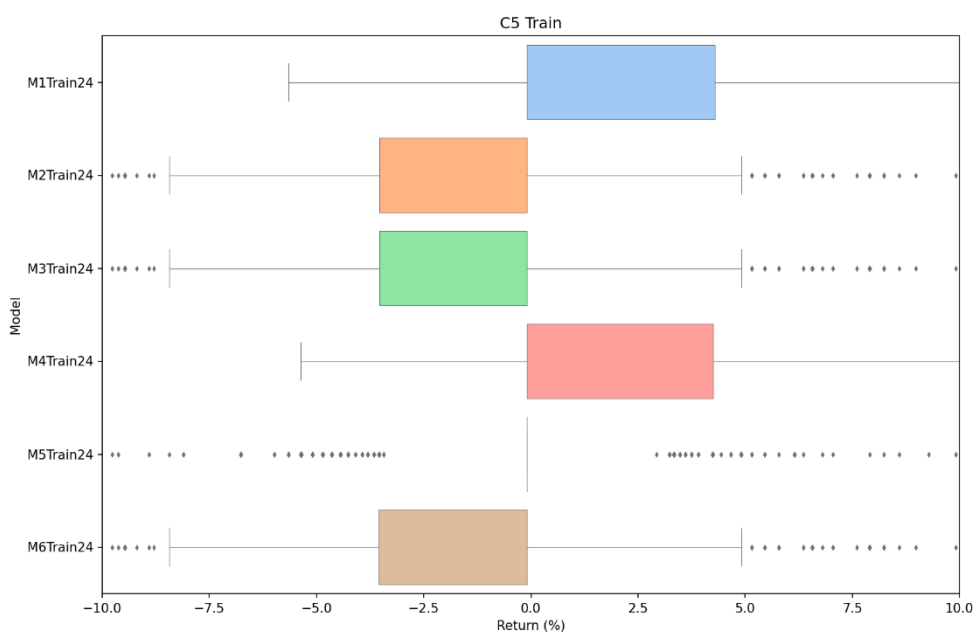
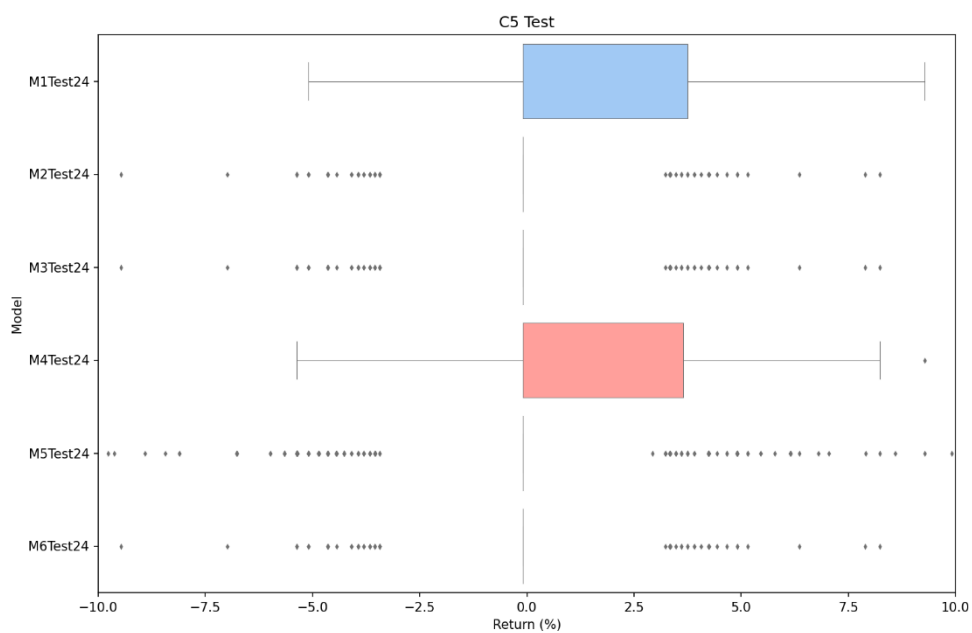


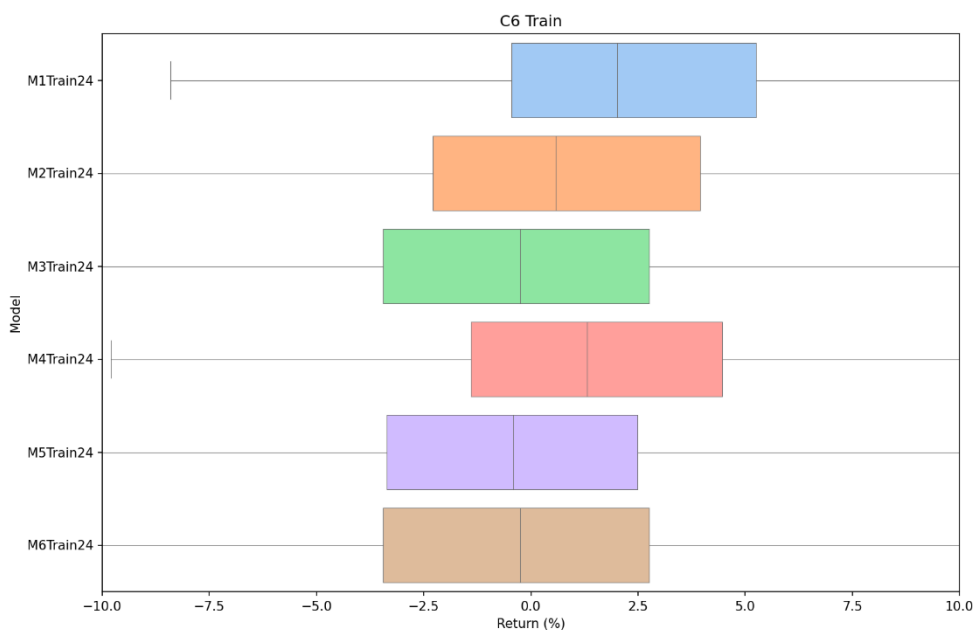
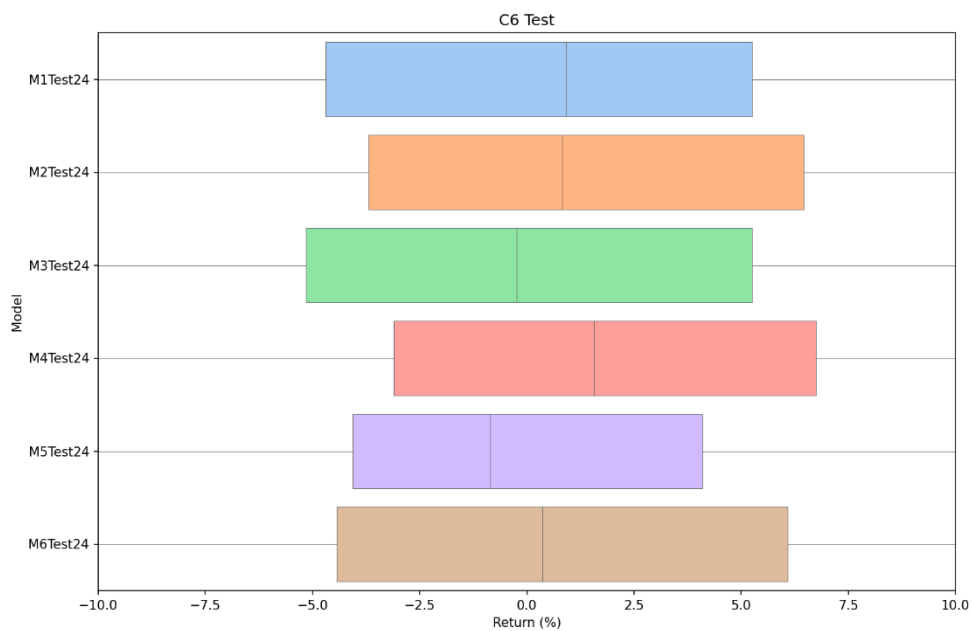


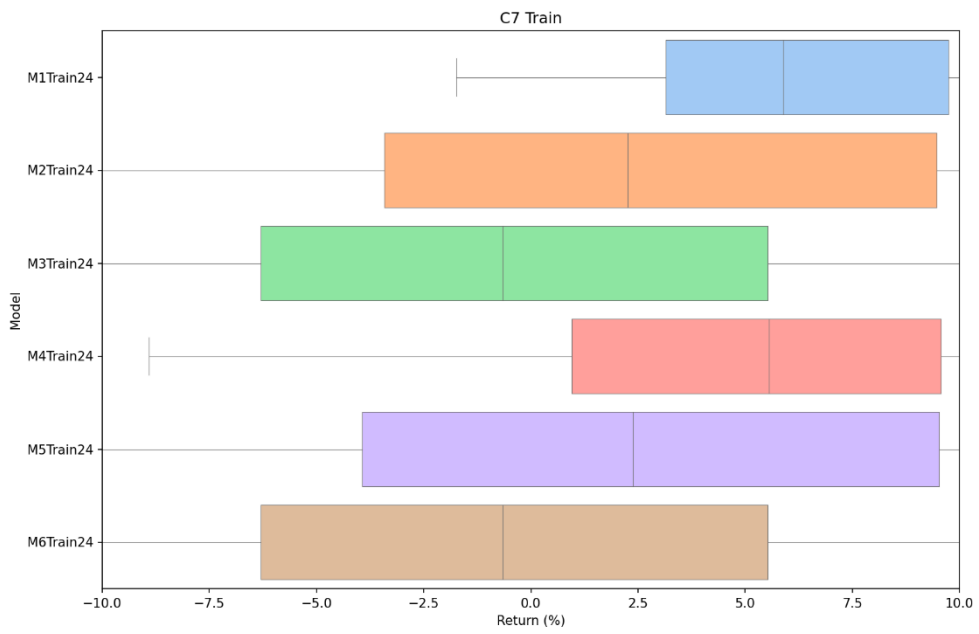
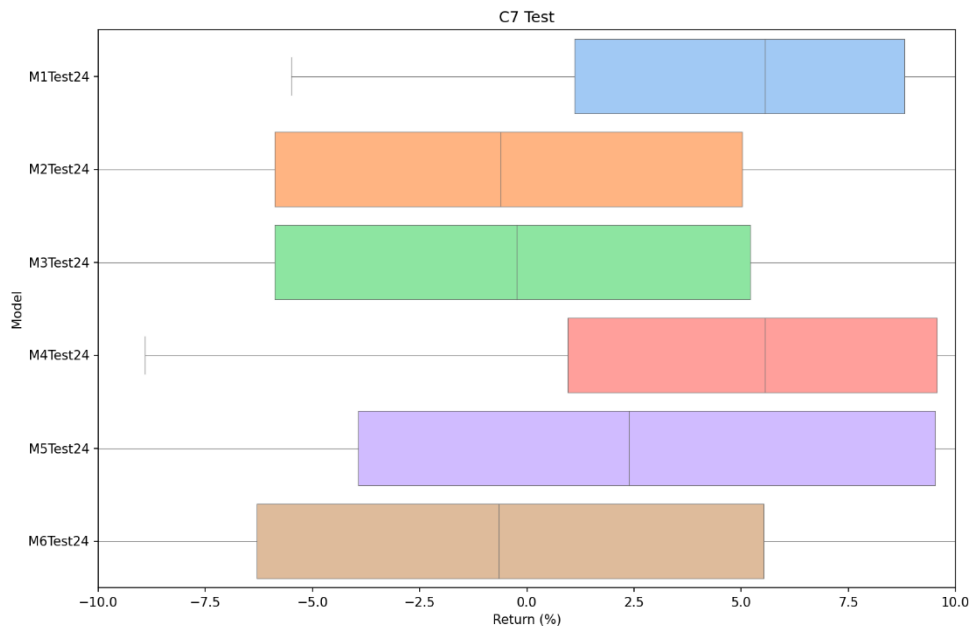


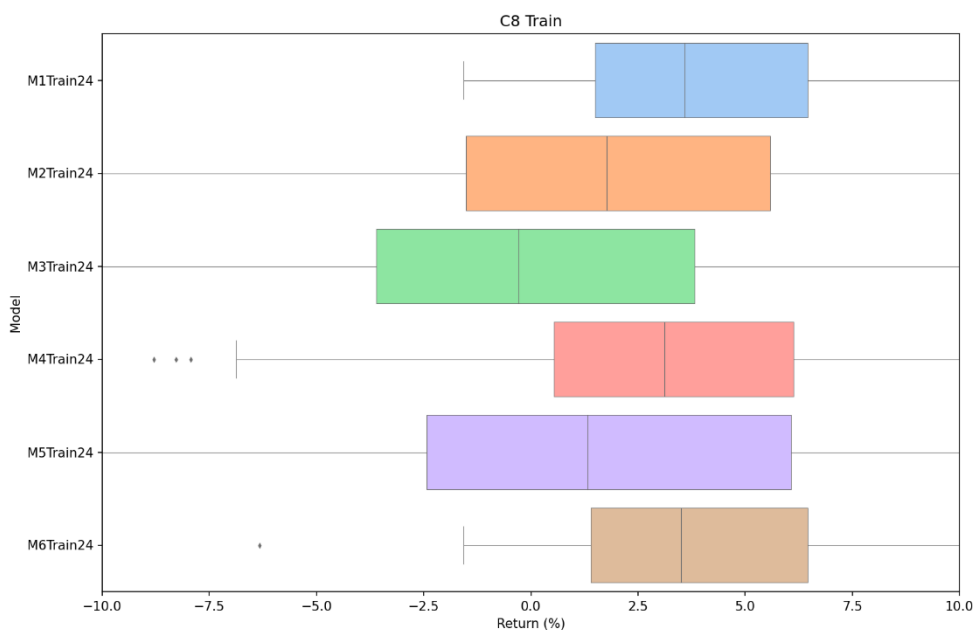
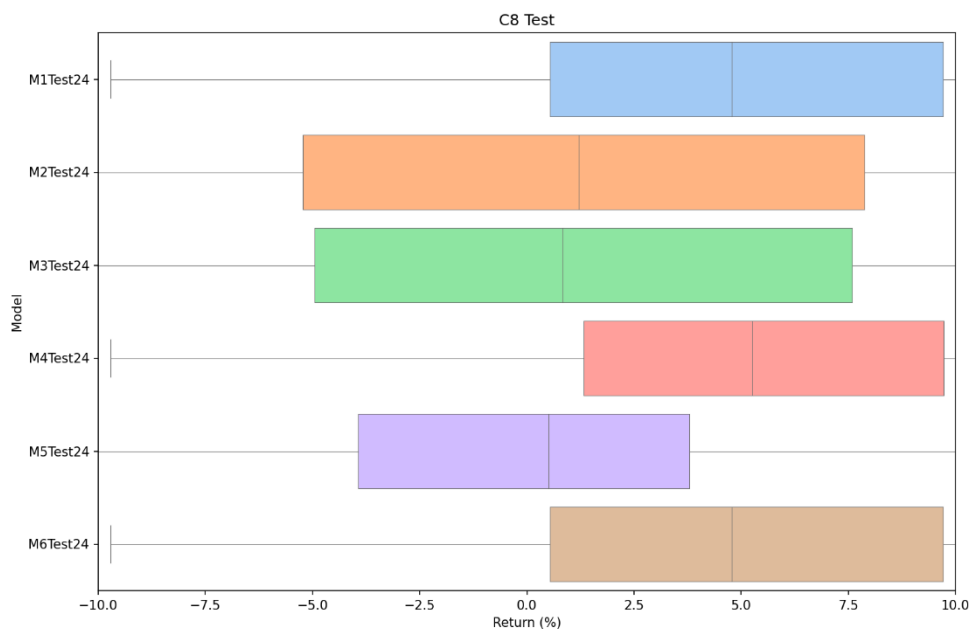




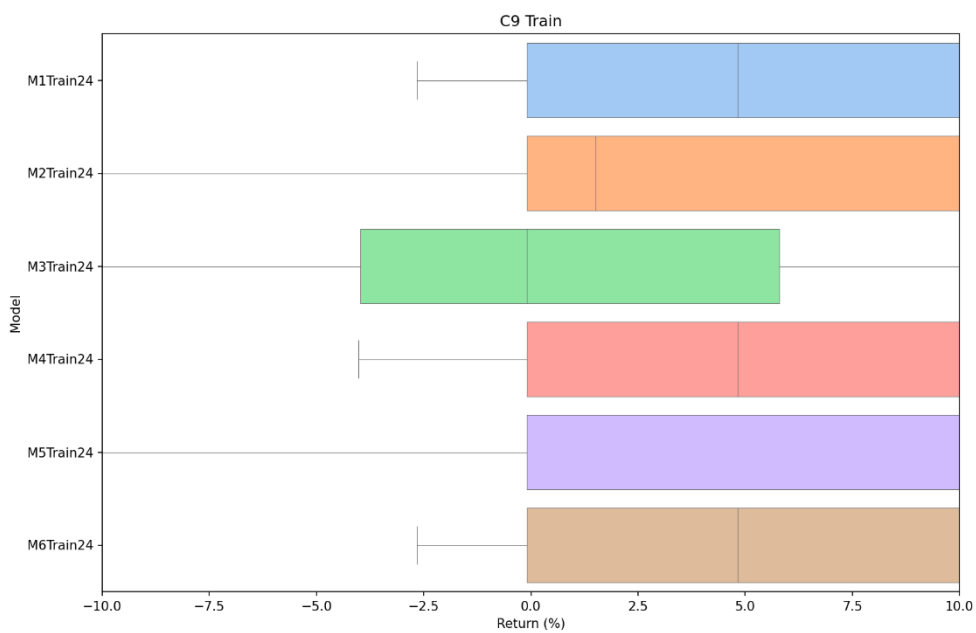
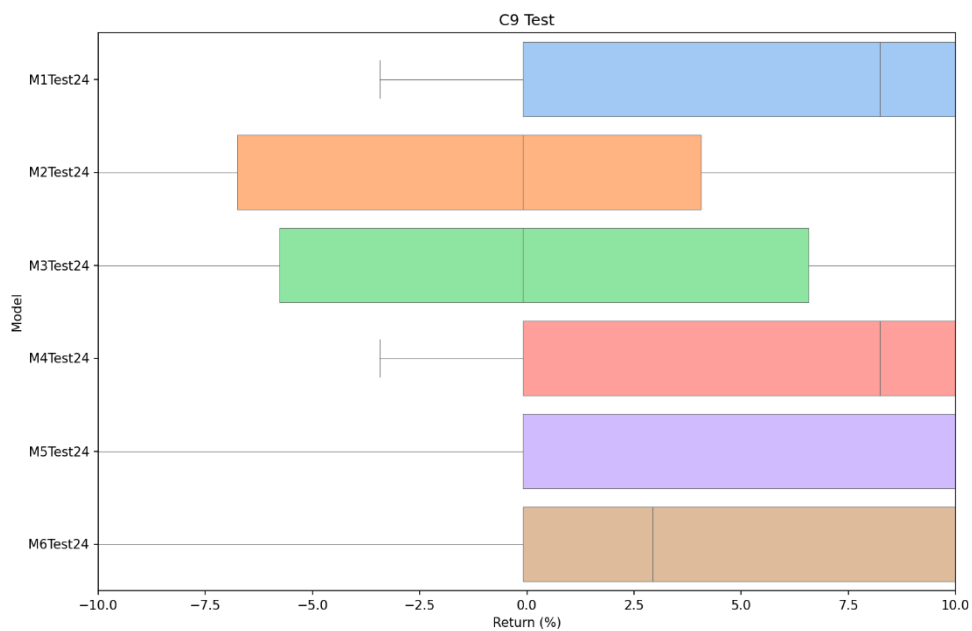


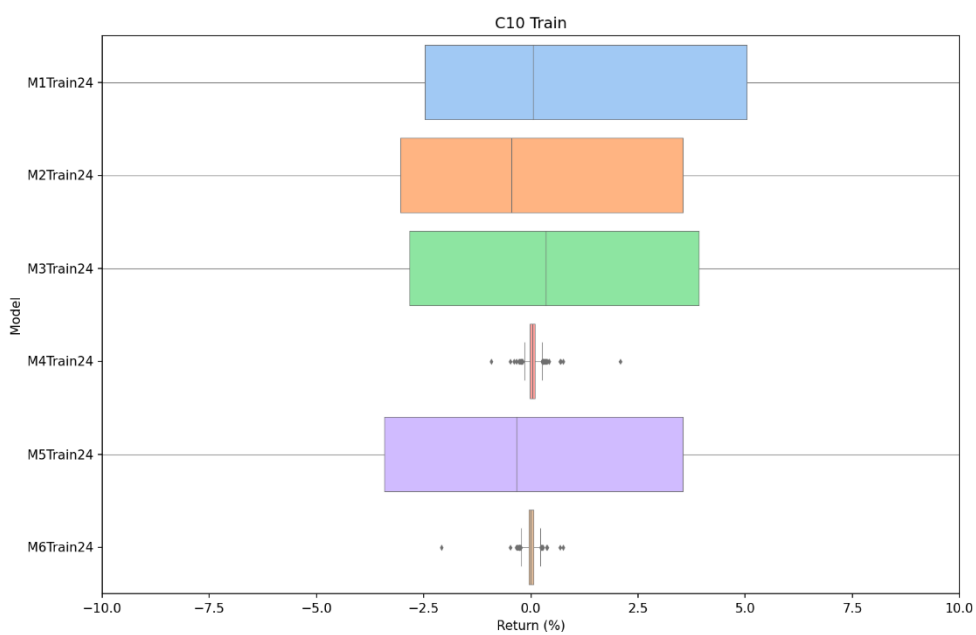
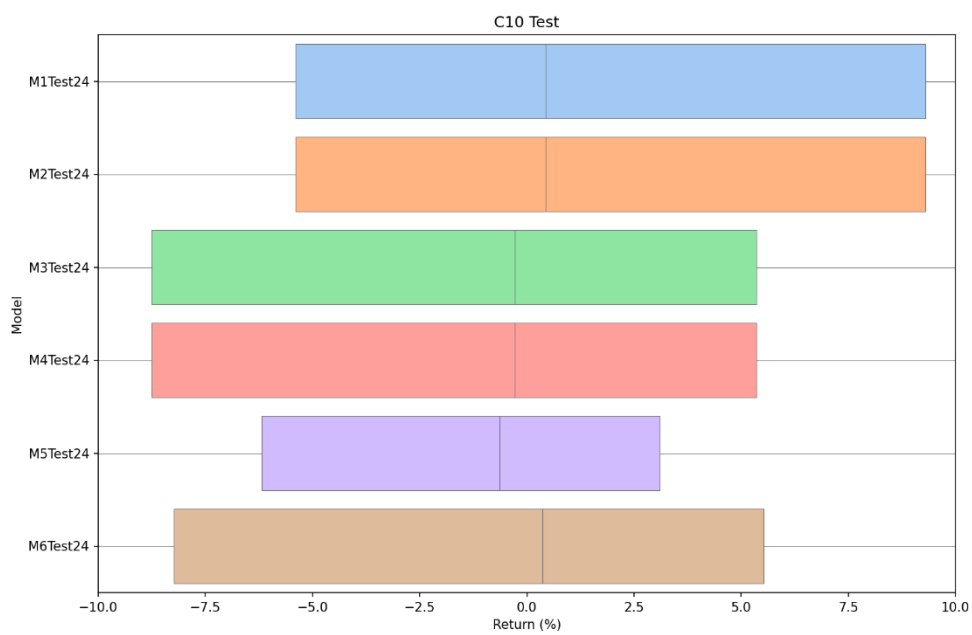


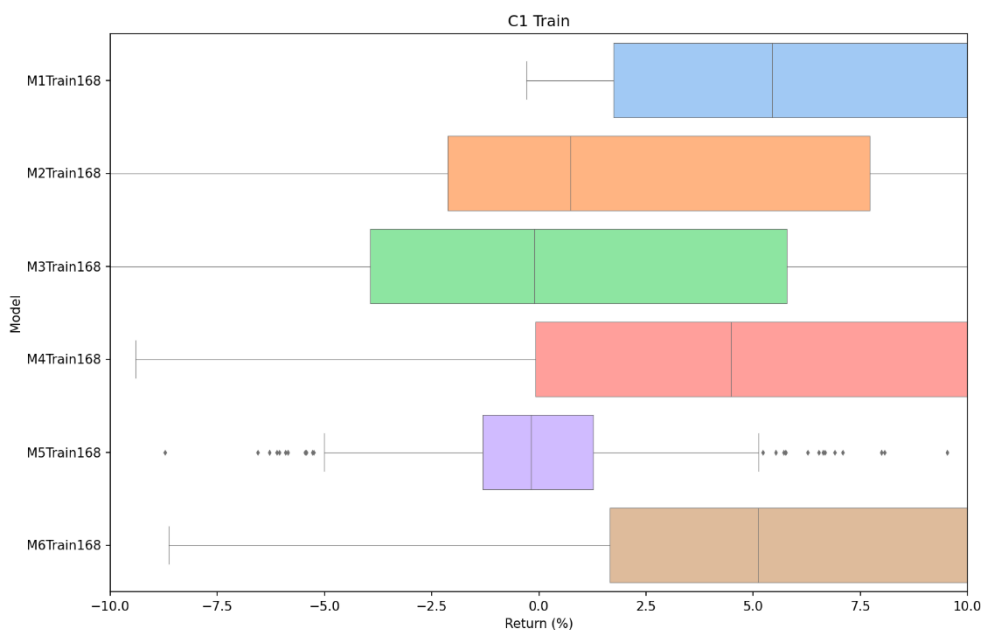
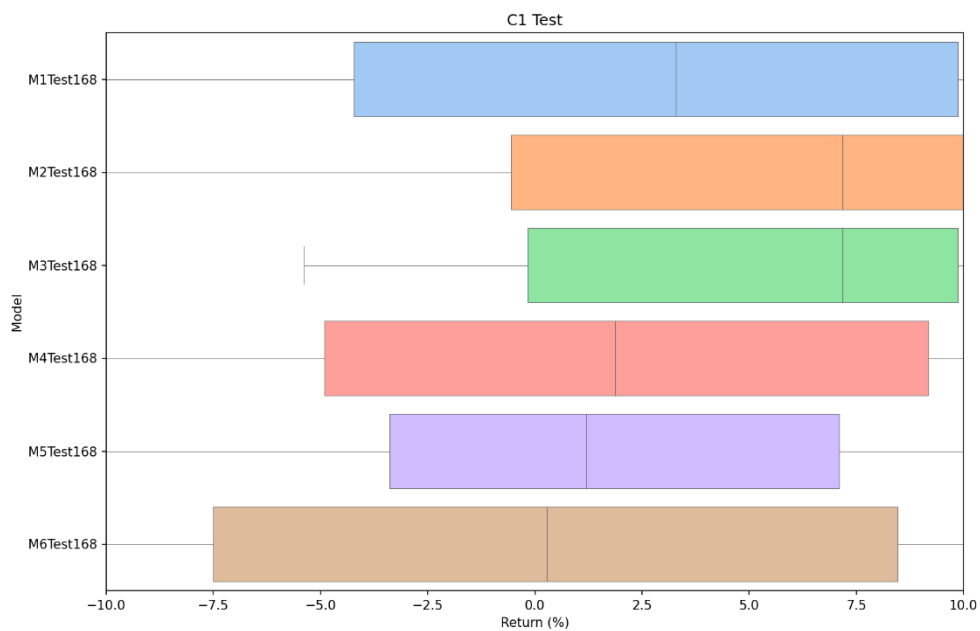


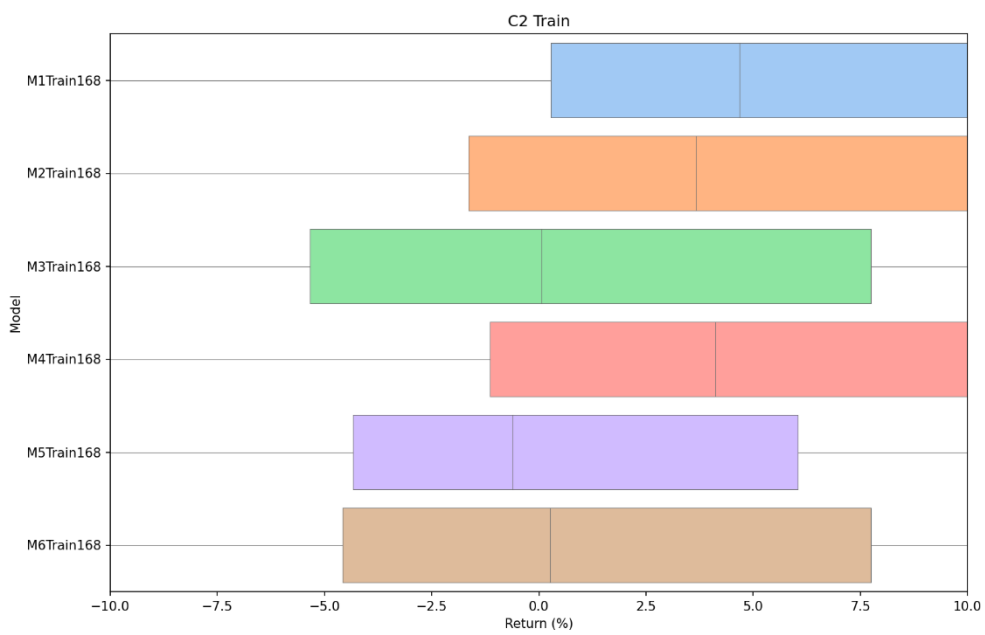
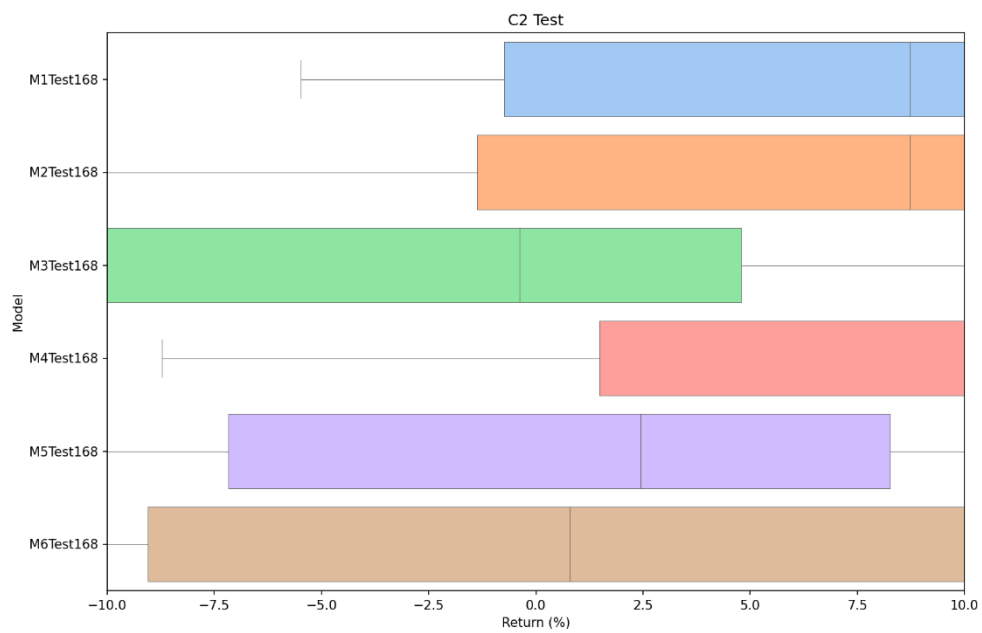


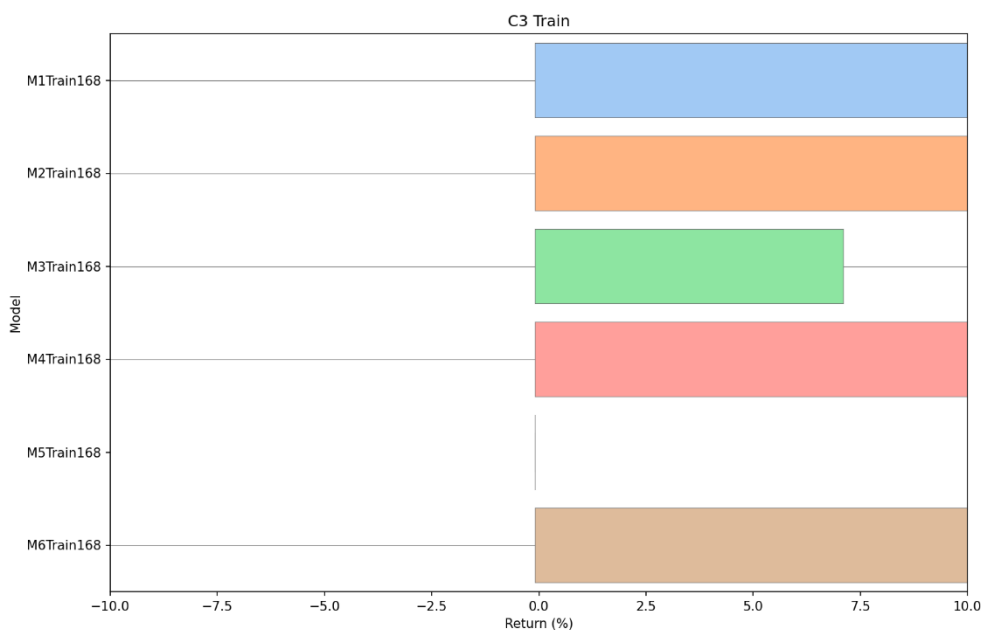
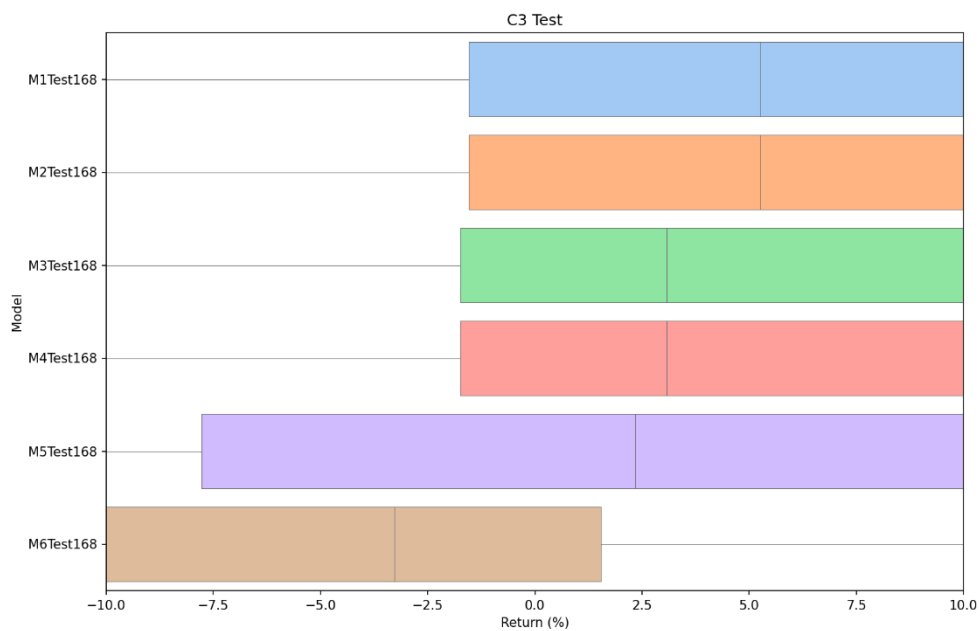


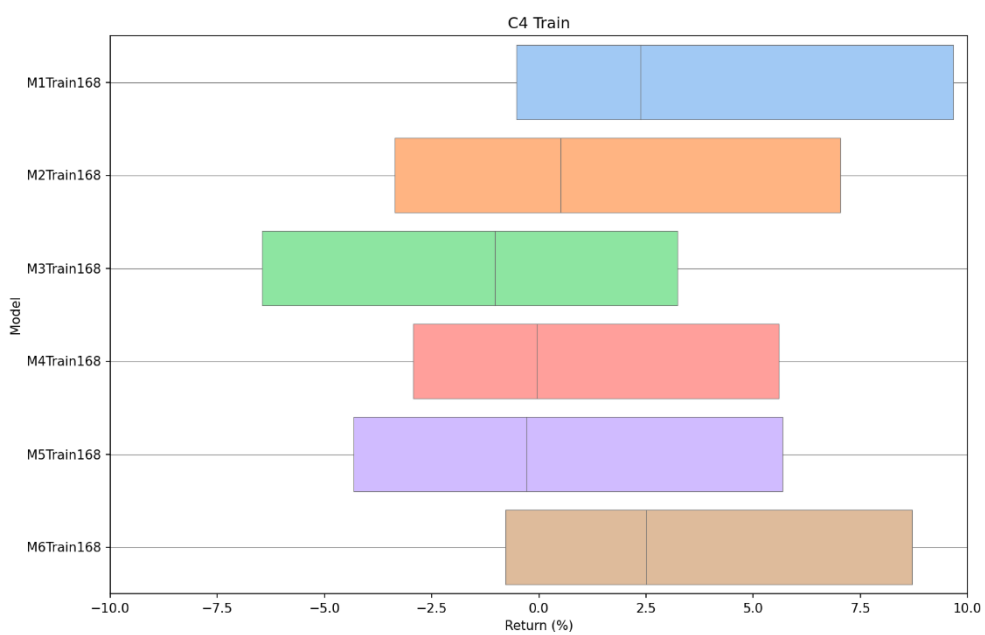
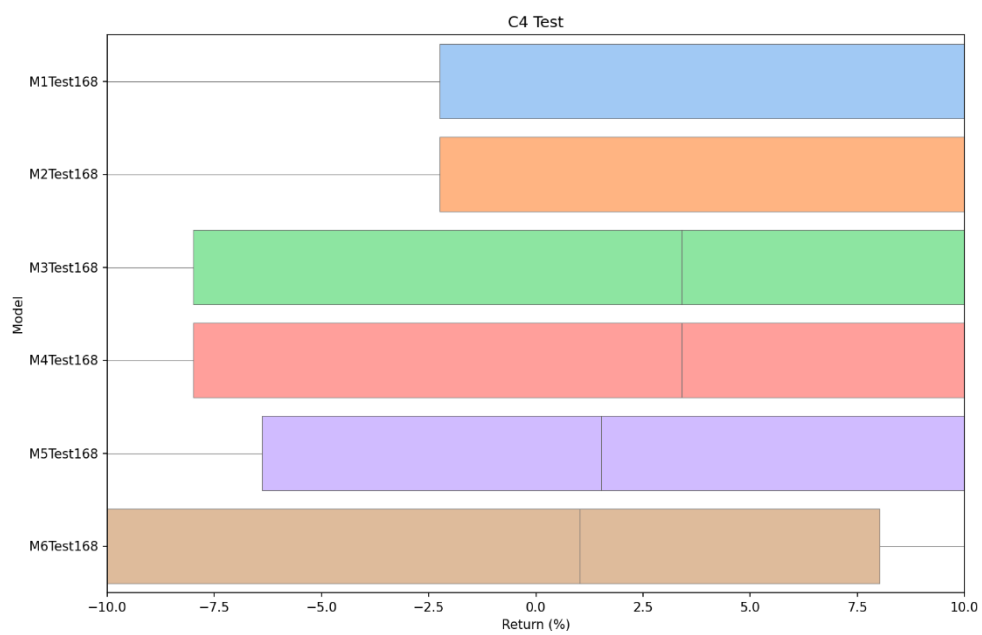


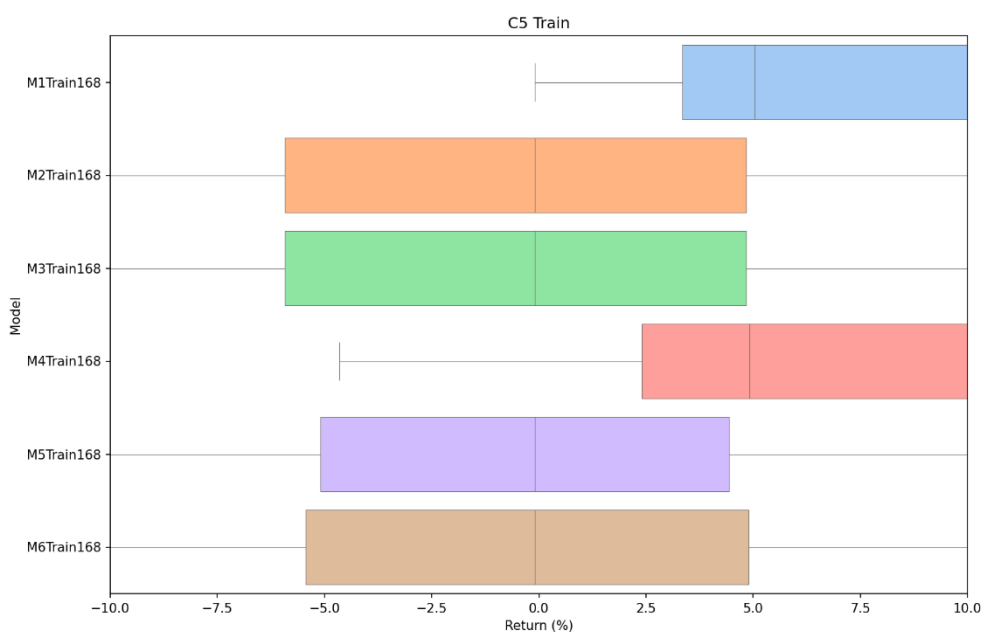
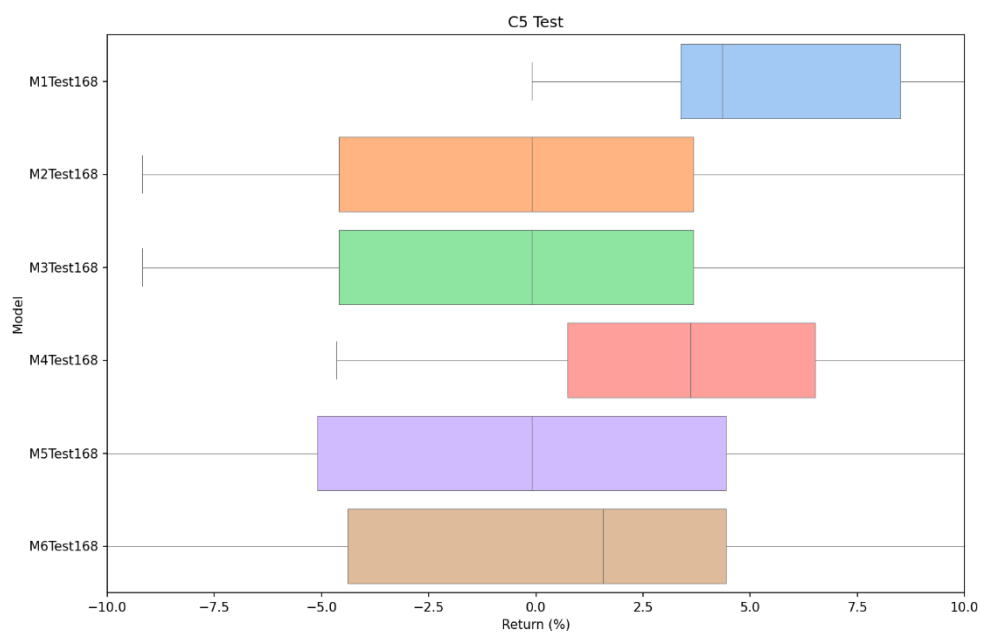


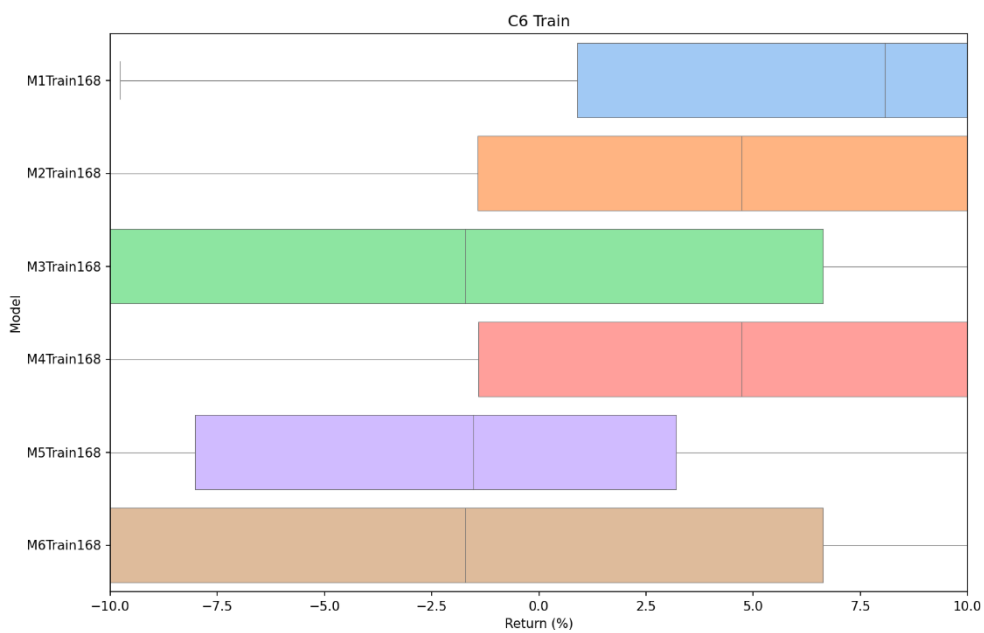
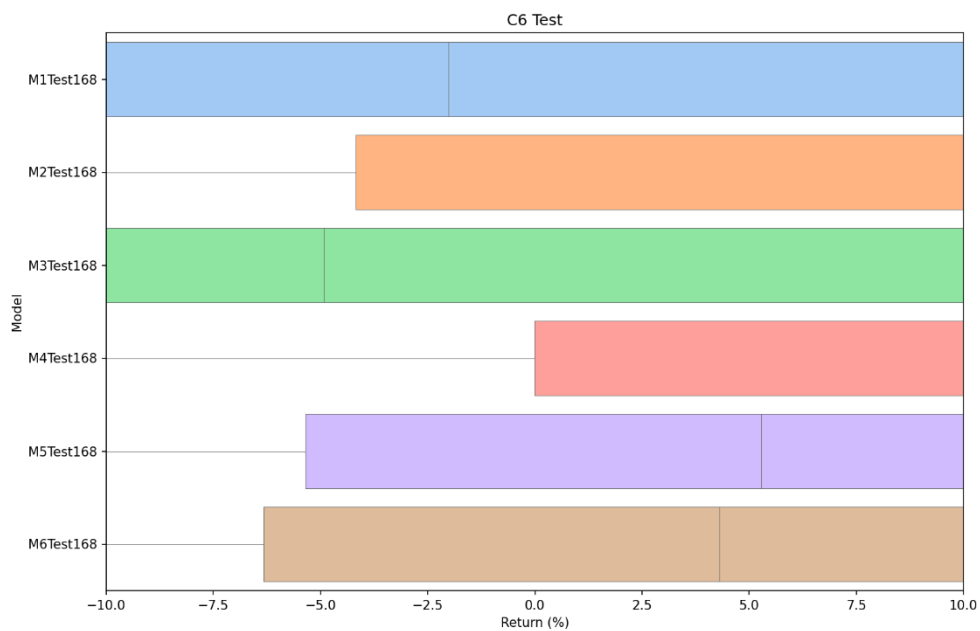




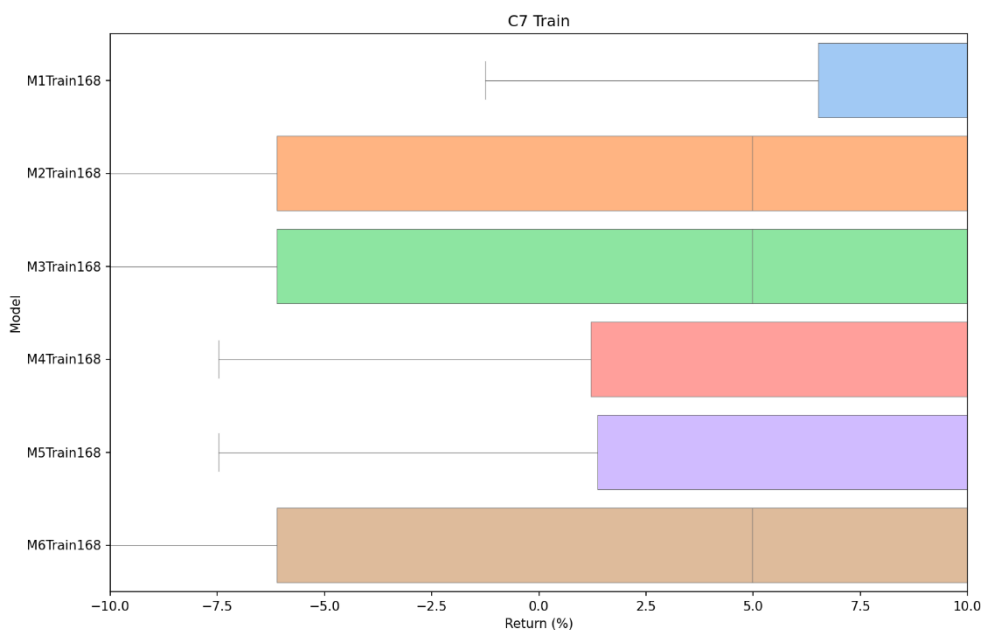
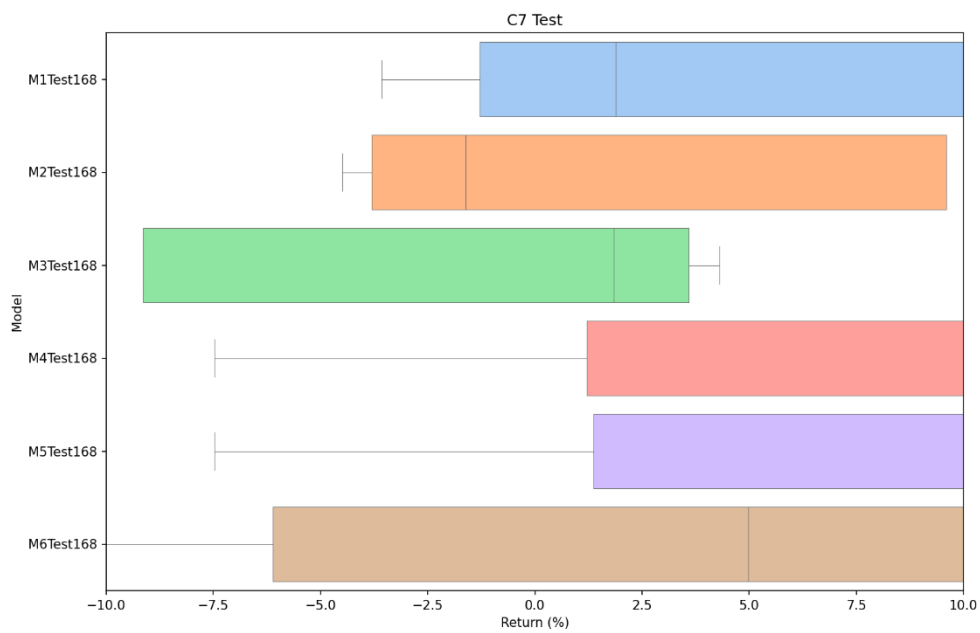


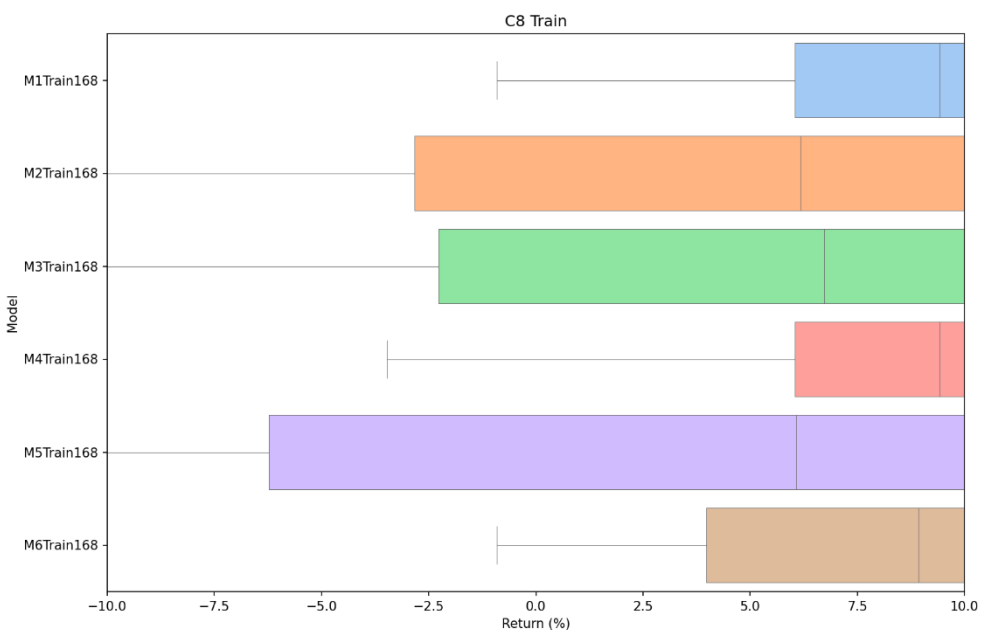
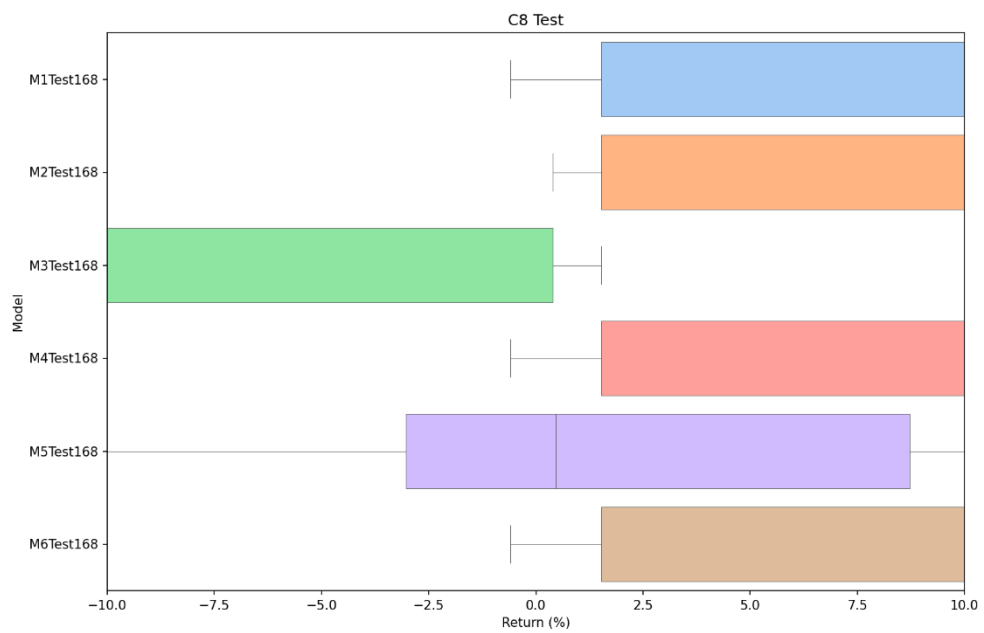


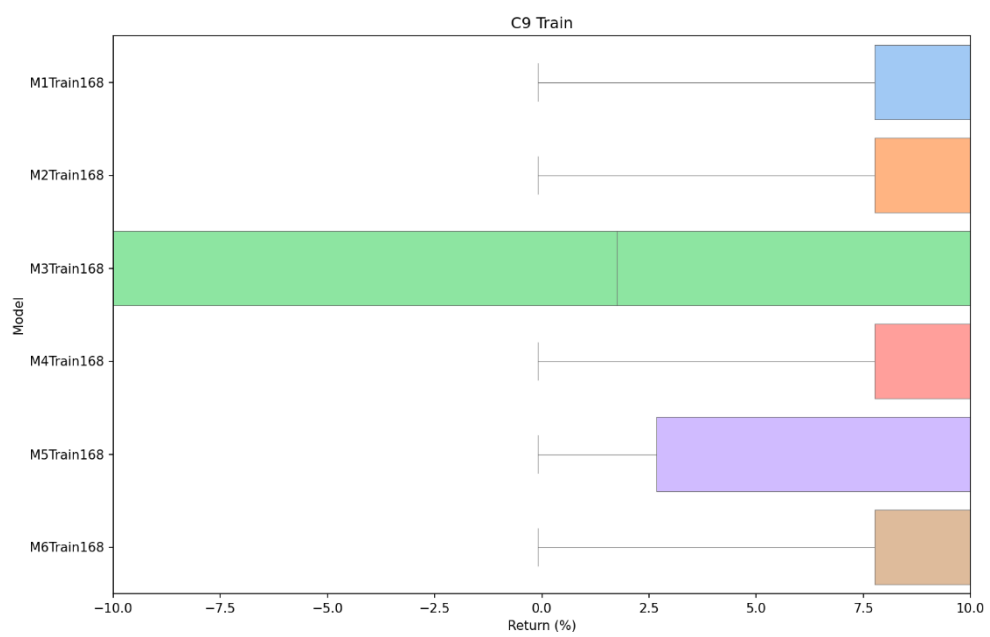
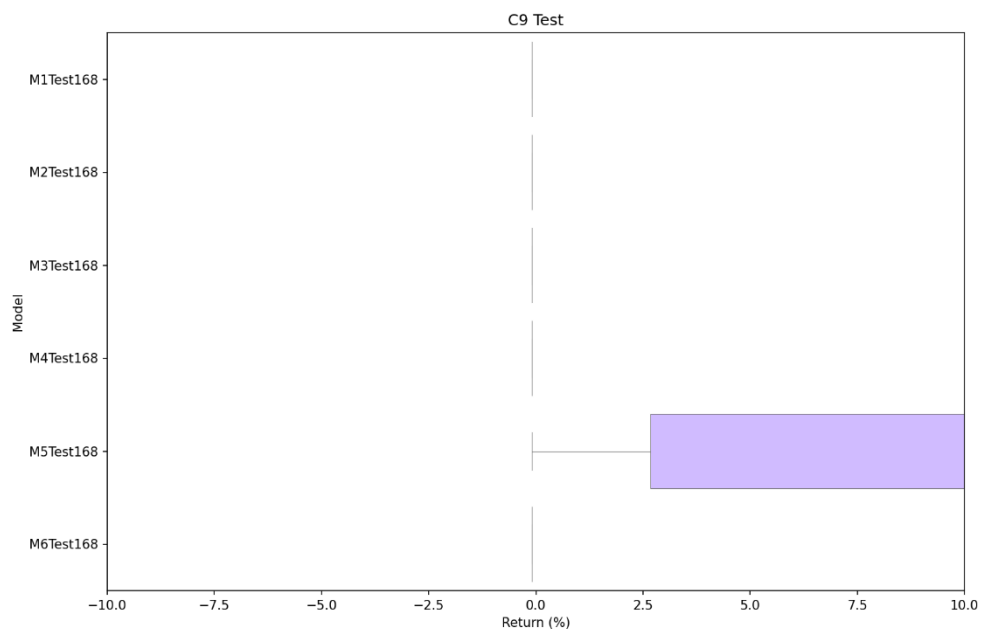


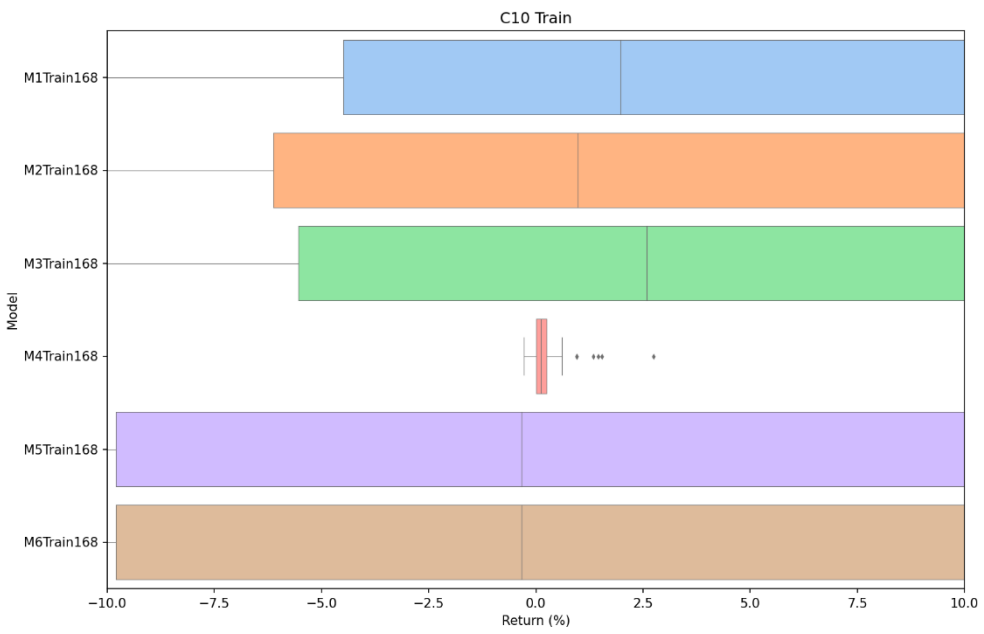
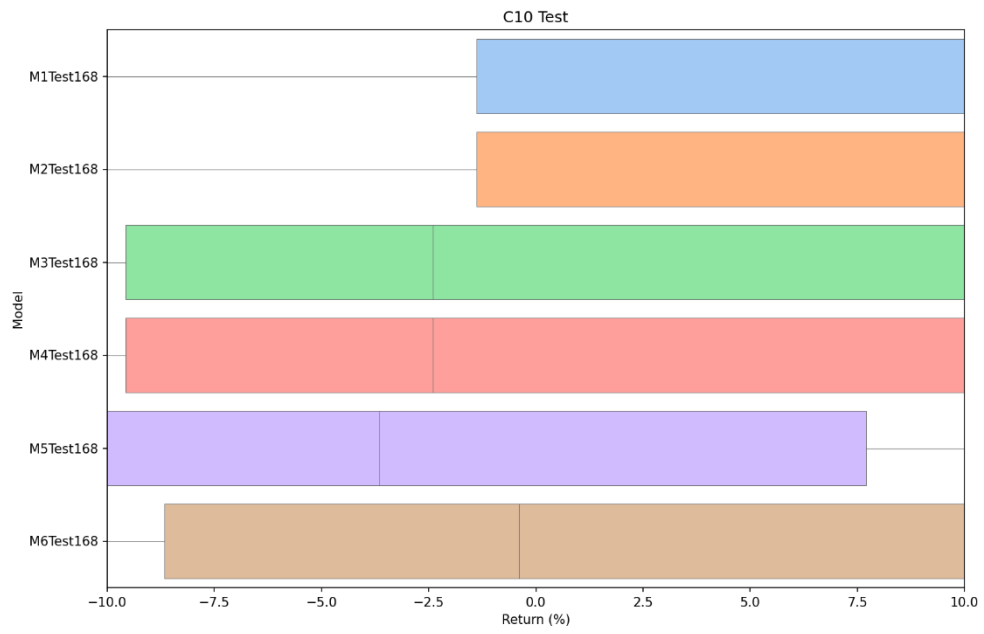












## VITA

**NAME** ศิวัช อัสวกิจพานิช

**DATE OF BIRTH** 24 กรกฎาคม 2540

**PLACE OF BIRTH** รพ.เมโย

**INSTITUTIONS ATTENDED** มัธยมศึกษา รร.โยธินบูรณะ  
ปริญญาตรี คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

**HOME ADDRESS** 12 ซ. นาคนิवास 41 ถ. นาคนิवास แขวงลาดพร้าว เขตลาดพร้าว กรุงเทพฯ 10230



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY