The Role of Machine Learning in Predictive Trading: A Comparative Analysis of AI Trading Bots in the Context of Environmental Responsibility

FACULTY OF BEHAVIOURAL, MANAGEMENT AND SOCIAL SCIENCES

Calin Lazar

2300966 - c.lazar@student.utwente.nl

Master Business Administration: Financial Management

Master Thesis – 30.06.2024

Supervisor: DR. U. Ahmad

2nd Supervisor: DR. X. Huang

Acknowledgement

This thesis is written as part of my Master of Science in Business Administration with a specialisation in Financial Management at the University of Twente. I want to thank my professor, Dr. U. Ahmed, for assisting me during the process. His insight in the field and constructive feedback have been crucial in exploring the integration of environmental practices within AI trading algorithms. Additionally, I would like to thank Dr. X. Huang for the support and guidance along the way. I am deeply appreciative of their support and encouragement throughout this academic endeavour.

Abstract

This study explores the integration of environmental metrics within AI-driven trading strategies, focusing on their impact on long short-term memory (LSTM) neural networks. Recent advancements in artificial intelligence (AI) have revolutionised various sectors, including finance, where AI algorithms significantly enhance predictive trading. Despite the increasing emphasis on sustainable investment practices, the effectiveness of incorporating ESG scores into AI trading models remains largely underexplored. The primary objective of this research is to determine whether the inclusion of environmental scores can enhance the performance of LSTM predictive trading models. The study employs a comparative analysis using historical data from ten large US companies across diverse industries, integrating ESG scores and CO_2 reduction goals into the LSTM algorithm. The results indicate that contrary to expectations, environmental scores do not positively impact the performance of LSTM training. In some instances, including these factors even hindered the model's ability to predict market trends accurately.

Keywords: AI trading, ESG scores, LSTM neural networks, predictive trading, environmental impact, sustainable investment.

Table of Contents

Chapter 1: Introduction

1.1 Background information

1.1.1 Recent developments and context

In recent years, we have witnessed rapid advances in artificial intelligence technology, revolutionising standard practices in many sectors. In art, AI helped creators push the boundaries of creativity and provided many with the tools necessary to generate complex digital artworks (Cetinic & She, 2022), who otherwise would not possess the skill or knowledge required to create such pieces, as well as being a powerful assistant for more experienced artists (Jiang et al, 2023). In data science, AI algorithms and machine learning have greatly aided in data analysis and interpretation, helping support the rapid development of the industry. Meanwhile, in marketing, AI changed how companies approach consumer engagement through personalised experiences (Babatunde et al, 2024), resulting in higher customer satisfaction. In the financial sector, trading strategies have drastically adapted to the newly available AI technologies, which were also paired with the increasing computing power that new financial machines can output. In most cases, this resulted in better risk assessment and management, leading to more consistent returns (Gokhale et al, 2019). Also, from the consumer perspective, users have experienced a new type of customer service, one powered by AI, capable of delivering faster answers to their problems. Even if many upgrades are still necessary for a more extensive acceptance of AI based customer services (Zhao et al, 2022), this represents a massive leap in that direction. As we can see from these examples, organisations and society can greatly benefit from the developments of the ongoing AI revolution. The impact AI has is not negligible, and it may, in the future, become the most powerful tool available in many technical fields, such as investing. However, AI's impact is not reaching its full potential due to a lack of research and adequate information regarding its impact, particularly in the financial field, where its interactions with certain financial instruments and investment opportunities remain partially unknown and unexplored (Königstorfer & Thalmann, 2020). This study aims to better prepare investors and policymakers by equipping them with the information derived from this comparative analysis, which will hopefully bring light upon the impact specific metrics have on the performance of such models.

1.1.2 Implications for the scientific community

The rapid progression of Artificial Intelligence had two significant effects on existing scientific literature: it increased the interest of researchers worldwide, mainly stimulated by its great potential to benefit society, while at the same time, the available literature could not keep up with all the innovations in the field, leaving many areas undocumented and open for further exploration. This can be exciting for researchers as it creates new areas of study. However, at the same time, it could be detrimental to society as the maximum value that can be derived from this technology is not attained sooner. By not being adequately informed, investors and investment firms lose potential earnings, leading to market inefficiency (Stout, 2002). In theory, investors are less likely to make investments when not adequately informed, resulting in less potential earnings, fewer volumes on the market, and less commission for investment firms, hedge funds, pension funds, etc (Palmer, 2023). This is the actual danger of the research gap.

The rapid development in AI studies has been documented by a recent study from 2023, called the “AI Index Report,” published by the Stanford Institute for Human-Centred Artificial Intelligence (HAI in short), which highlights the significant developments in AI, showing rapid advancements in many fields (Stanford HAI, 2023). The study showed how the total global corporate investment in the field of AI research skyrocketed over the past years, peaking in 2021 at 276.14 billion US dollars, representing a staggering 985% increase since 2015 when the total sum only accumulated to 25.43 billion US dollars (Lynch, 2023). The study also shows that as a direct correlation to the increase in funding, as the number of papers on AI submitted to FAccT had increased tenfold in recent years. This indicates that the topic is relevant to corporations that finance such studies and to researchers who use this funding and conduct academic research on the subject, as they both see its added value as beneficial to themselves, their organisations, and society.

It is slowly becoming clear how the gap left by this rapid development in the literature covering AI integration and governance is hurting fields where the implications of technology could have a tremendous transformative impact, one of which is Finance. The potential in this industry may be one of the greatest, as it can revolutionise current practices. However, due to swift recent developments, this potential has yet to be fully reached, to the detriment of society, which could benefit significantly from better transparency, efficiency in the markets, inclusivity in financial services, and much more.

In order for the positive effect of this technology to reach the population as a large, this research gap must be filled, and this can generally only be done by understanding the exact boundaries of the gap itself and by further researching the topic, filling in the blanks in the current knowledge. Predictive analytics, for example, is a great place to start. As studies show, there is great potential for machine learning to aid in increasing prediction accuracy (Weng et al, 2018; Cao et al, 2021; Che et al, 2014). This is due to AI’s remarkable capacity to process large amounts of historical data in real time and deliver fast predictions that adapt to market conditions dynamically, helping traders make better decisions. Despite all these significant findings, the integration of environmental, social and governance scores into Long Short Term Memory (LSTM) models in financial price prediction is still largely unexplored, and there is a great lack of research on this particular topic. For example, Weng et al. (2018) explored the application of deep learning models in the prediction of stock prices, discovering and showcasing their great predictive ability, but did not explore the impact of any environmental scores and opted to stick to traditional trading means. Similarly to this, Cao et al. (2021) investigated the use of machine learning algorithms for trend identification, demonstrating machine learning's ability to detect profitable trading opportunities. Che et al. (2014) focused on the implications of recurrent neural networks (RNN) (LSTM is a type of RNN; more on this in the literature review section) for predicting time series data, having implications for forecasting the movements on financial markets. Despite these advancements, the fact that these studies, similarly to the majority of studies available, are only investigating traditional financial metrics and often overlook the impact of environmental factors, which in other fields of today’s academia is considered a very hot topic. The only study related to this topic is a very recent one from May 2024 by Lee et al., where public’s perception of ESG factors is taken into consideration as sentiment analysis. However, there are no studies available on the direct impact of ESG scores on the predictive performance of neural networks. Environmental scores and metrics can be seen as an assessment of a company's sustainability and ethical impact. In order to hasten the development of such algorithmic trading models capable of collecting and processing real-time data, a comprehensive understanding of the matter and environment these operate in is required. Through adequate research, newer models can build on existing data and reach the desired societal impact.

Through adequate research, newer models can build on existing data and reach the desired societal impact. This involves not only advancing the algorithms themselves but also understanding the regulatory, ethical, and practical implications of deploying such technologies in financial markets. Addressing these gaps will ensure that the integration of ESG scores into AI-driven predictive analytics is both effective and responsible, ultimately contributing to a more sustainable and equitable financial system.

1.1.3 AI trading bots and environmental practices

According to a report by the Economist, out of 31 trillion USD worth of analysed equities, more were managed by computers and algorithms rather than by human traders (Buchholz & Richter, 2019). Another study focused on the Forex market indicated that approximately 92% of trading was performed by trading algorithms instead of human traders (Smigel, 2023). This situation is very different from a few decades back when the financial scene was significantly different.

However, this switch to automated trading did not happen overnight. Over the past years, trading practices shifted, transitioning from manual trading to algorithmic and later AI-driven trading. Manual trading was characterised by physical trading floors, where traders made decisions based on intuition and fundamental analysis. Algorithmic trading emerged in parallel with the development and standardisation of the personal computer, allowing traders to use automated strategies and execute trades on predetermined criteria, a common example being the integration of buy/sell order limits or buy/sell stops. AI-driven trading utilises machine learning to analyse vast amounts of data, resulting in a more precise prediction of market movements and allowing traders to have algorithmic strategies that adapt as market conditions shift (Thakar, 2023). The complexity of the data comes in big part not only from the number of factors but also from how far back indicators and EAs can consider information when placing a trade, some going back days or even months for a single trade.

Due to all these applications, AI trading has become a very popular research and development topic. Some of these AI trading bots are called EA (expert advisors), and they can make use of algorithms to analyse vast amounts of financial data and make rapid trading decisions. Popular trading platforms, such as MT4 and MT5, have implemented EAs in their standard package, making it easier to set them up, especially when compared with the alternative of developing code from scratch designed to trade, such as the one used on a non-dedicated portal (Lima de Castro, 2021). When it comes to speed and accuracy of performing trades, EAs almost always outperform human traders in these particular fields. Others opt for more technical approaches by using other programming languages to develop these trading algorithms, such as Python. Due to the vast availability of community-developed packages, Python is a very powerful and up-to-date tool used by investors who specialise in this particular type of investing.

Some of these EAs are being implemented and modified by traders through the MT4 and MT5 platforms. These platforms do not possess the most user-friendly interface and can be challenging to use, as some require users to understand programming. Pre-made, ready-to-use EAs are also available for sale on the MT4 and MT5 marketplace; however, for the individual investor seeking to increase their retirement savings through steady compounding, this may still be too complicated and require further simplification. The recent developments made it easy for less experienced investors to leverage the benefits of AI trading without requiring any previous programming knowledge and little to no financial knowledge.

The benefit of these platforms derives from their ease of use. Finax and Peaks, two prevalent examples of such platforms on the Dutch market, claim they can offer investors above-average returns with automated trading techniques. Of course, like with any other investment, there is risk, and it is naturally in the interest of those investing to mitigate such risk. To do so, they must be well informed, and considering the gap found in the existing literature mentioned in the above paragraph, this may cause obstacles for those seeking to inform themselves on this topic.

One concern over the topic of AI is its impact on the environment. Studies show that AI development does not come at no cost to the environment, as such models produce large amounts of CO_2 emissions (Lynch, 2023). Social and environmental responsibility has been a big issue in Finance for many years, as investors pay more and more attention to the different practices that companies undertake before deciding to invest in them. Caring for the planet is a human sentiment that can positively impact investors' results, and considering that this is a collective concern, it may indeed have the power to shift the market. However, with more and more trading being done by machines, incapable of expressing sympathy to a decaying planet and only considering technical analysis in the making of trading decisions, one could only wonder if there is still a place for environmental awareness in modern trading, and whether or not can this still have a positive impact on the outcome and profitability of a trading strategy governed by such machine learning algorithms. This study is not going to investigate ways to reduce the carbon footprint caused by training AI models. The potential negative impact of AI on the environment was just brought up to help put things into perspective. If these models consume large amounts of energy during their training, this could perhaps be offset by AI leveraging environmental factors. Additionally, by informing the public on the efficiency or inefficiency of these factors, individuals interested in training their own models will not have to go through the same experimentation process that this study conducted, as they could simply refer to the findings of this paper. This would result in a reduction of wasted resources, so despite the results of this study, in theory, the study itself is contributing to a greener environment by informing future developers of the implications associated with ESG, CO_2 and temperature integration into neural networks.

Modern Finance increasingly recognises the importance of ethical and environmentally friendly practices and considers such factors when making investment decisions. ESG (environmental, social, and governance) scores and CO_2 reduction efforts (a popular metric in circular economics) reflect a company's sustainability and social responsibility, which can significantly impact stock performance (Zairis et al., 2024). Policymakers and investors use these metrics to guide investment decisions, favouring companies that contribute positively to environmental and social goals. Short-term ESG initiatives can attract investment and improve reputation, while in the long term, they are essential for sustainability and regulatory compliance. On the other hand, there is plenty of research showing that, in some cases, environmental factors have no direct impact on stock performance and, in some cases, even hinder a firm's activity by trying to attain a higher score (Ruan & Liu, 2021).

1.2 Research Goal and Research Question

1.2.1 ESG scores in AI trading platforms

The two platforms mentioned above, Finax and Peaks, have quite different approaches to trading when evaluating social impact factors. Finax does not consider ESG scores, while Peaks has this option available to its customers (Finax) (Peaks). Customers interested in safe environmental investment plans cannot choose Finax due to its lack of ESG score embedding. This can leave investors unsatisfied and firms without the funds they would otherwise have. There are options in the market, but are they better? Individually, one could take the past performances of Finax and Peaks, put them side by side, and compare the results to determine which approach is more suitable. However, since past performance is not guaranteed for future results, the applications of such a test would be hard to generalise and specific to these two platforms. Investment companies are also required by law to communicate this to their investors, under SEC Rule from Title 17, Chapter III, § 230.156 under the federal securities law of the Securities Act of 1933, section 17(a) and as well as section 10(b) of the Securities Exchange Act of 1934. Governments found it imperiously necessary to warn investors of the volatile nature of the financial markets, understanding that an uninformed investor is at a significantly higher risk. Even so, if the performances of the two firms are compared, these are not the only two available AI investment companies on the market, and more are bound to appear. Here again, we can see the detrimental effect of the current research gap in modern literature, where investors may be confused due to the lack of literature on the topic. To offer an unbiased solution to the problem, a new AI model will be developed for this paper. This new neural network algorithm will be tested on the same data in two parallel studies, with one significant difference between the two: embedding environmental factors. This will test whether or not using such factors can benefit the training of machine learning algorithms to offer a better return. While papers on the impact of environmental factors such as ESG scores or CO_2 emissions exist and studies on AI trading performance, the crossover of the two is mainly unexplored. It leaves many questions unanswered, as will be shown later in the literature review. One aspect that is interesting to investigate is the type of impact these factors have on the performance of machine learning algorithms. These factors are valuable to us humans, but they may not play such a significant role "in the eyes" of a machine, accounting only for statistically better results.

1.2.2 Bringing light upon the impact of environmental factors

Research indicates a growing interest in integrating ESG factors and CO2 measurements into financial decision-making. However, a specific study on how AI trading algorithms can incorporate sustainability metrics still needs to be explored. While studies have begun to investigate the broader implications of ESG factors and their impact on investment returns and risk assessment, their application within AI-driven trading strategies still requires further comprehensive analysis. This paper aims to bridge that gap by directly integrating ESG metrics into AI trading algorithms, an area yet to be thoroughly researched. Insights in this field can bring significant benefits to a wide range of stakeholders: managers (by offering a competitive advantage while allowing for sustainable growth), policymakers (as it promotes sustainable economic growth and helps in the formulation of policies by bringing light upon unexplored areas of an emerging field), investors (by providing a more informed decision-making process) and for the large public in general (encourages companies to adopt environmentally friendly practices, resulting in better living standards for all). This paper also aims to demonstrate the multifaced advantages of embedded ESG scores and CO_2 emission levels in AI-driven trading strategies and how implementing these factors can impact training and the predictive power of predictive machine learning models.

1.2.2 Research question

Considering the identified need for further exploration into the current use of AI trading bots, especially in regard to the integration of sustainability metrics, this paper aims to achieve the following primary objective: enhancing the incorporation of environmental practices in AI trading algorithms to improve financial decision-making and promote responsible investing. This objective addresses the ethical responsibility of facilitating sustainable investing by making it more accessible and allowing for more informed decision-making. With this in mind, we proceed to formulate the following research question:

How does the integration of environmental factors impact the performance of artificially intelligent predictive bots?

To analyse this interaction, a new AI model will be developed specifically for this study to explore the research question effectively. Instead of analysing existing Expert Advisors (EAs), this approach involves testing a new neural network algorithm on identical data sets, with and without incorporating environmental factors. As mentioned above, filling the research gap could result in saved resources and better use of the available ones and hopefully help new studies by answering the research questions of this paper. This dual-study methodology aims to directly assess the impact of sustainability metrics on the algorithm's ability to predict market trends and generate returns. Through this comparison, the research seeks to discover the impact of embedding ESG scores and CO_2 company policies into machine learning algorithms for financial trading. According to an article written by Deloitte, a multinational professional services network, 65% of investors take ESG scores into account, a substantial amount (Deloitte, 2022). However, it is important to bring to the attention of investors the fact that environmental scores may not have any impact (or even a negative one) on the performance of AI trading bots.

1.3 Structure

To properly manage these efforts, as mentioned above, this paper will follow the structure below. Section Two will provide a summary of the current literature and show how various theories and concepts have led to the development of this study's main research question. During chapter three, the methodology will be discussed alongside argumentation supporting the choice of statistical approach. In chapter four, issues concerning data will be presented, including how data will be collected and treated. In chapter five, the results of this study will be presented alongside their analysis. Lastly, chapter six summarises the entire work's findings, starting with a conclusion followed by practical implications and limitations as well as suggestions for further research.

Chapter 2: Literature Review

This chapter provides a comprehensive literature review on the integration of AI in trading, with a specific focus on the embedding of environmental practices, such as ESG, CO_2 emissions and the role of company climate goals. The literature review begins with an overview of AI's impact and role on trading strategies, exploring the transition from manual to algorithmic and, finally, AI-driven trading. Next, section 2.2 discusses the concept of ESG scores, highlighting their pros and cons while offering a better understanding of their role in financial decision-making. Section 2.3 explores the impact of CO_2 reduction goals on company performance, their impact on public opinion and how they could impact AI trading in the modern financial landscape. In section 2.4, we provide an overview of how reliable temperature goals are to firms and how they influence investors. Additionally, the literature review will also offer insight into what LSTM models are and briefly go through their general functionality, based on the book "Neural Networks from Scratch in Python" by Harrison Kinsely & Daniel Kukieła. This book also served as a theoretical basis for the development of the models used in this study. After the main factors that will be analysed in this study are presented, the general theoretical framework will be composed in section 2.5. Finally, the hypothesis of this study is formulated and explained, with the goal of further guiding the rest of the empirical analysis of the topic at hand.

2.1 AI in Trading

2.1.1 Efficient Market Hypothesis

One of the main overarching theories in the field of AI trading is that of the Efficient Market Hypothesis (EMH), which states that markets reflect all available information through their movement, making it challenging for human traders to achieve consistently positive returns due to this chaotic movement caused by the multitude of information. AI challenges this concept by identifying patterns that are not obvious to human observers. AI can have an overview of millions of entries and previous patterns simultaneously, something no human trader can be capable of. Human learning archives an understanding of such patterns through experience and long-term learning. Neural networks mimic this complex process and run it over many simulation scenarios, often over spans that defy what is possibly attainable within a human life (Downey, 2024).

There are three kinds of EMH: weak form, semi-strong form, and strong form, all varying based on how ineffective conventional trading forms are depending on current market conditions. In weak form, technical analysis is complicated, as the effect is already taking action, but it is possible. In this type of market, investors can use social and environmental practices, as they still significantly impact stock prices. In semi-strong, conditions worsen, and "chaos" intensifies, meaning that fundamentals analysis becomes ineffective, as the multitude of factors applying their forces on the markets is more vital than the news related to the release of various governmental financial statements. In theory, ESG scores and CO_2 measurements would become unusable to human investors in such conditions. The strong form, not even insider trading, can guarantee any consistent results, as market forces reach a level of impact on the price fluctuations that makes them hard to predict. All public and private information is reflected in the price in such an environment. ESG scores and CO_2 measurements can still guide a well-developed AI through the perilous financial landscape (Aminimehr et al., 2022). This is, of course, assuming that they have an impact to begin with. Studies showed that out of all models, LSTMs performed best in tests conducted on 21 years of market data. This theory, however, is shrouded in controversy, as critics argue that it overlooks market anomalies and the ability of skilled investors to beat the market (Inspired Economist, 2023).

2.1.2 Algorithmic Trading and Reinforced Learning

Building on theories described in the previous section, algorithmic trading and reinforcement learning represent a crucial point in the evolution of trading strategies, leveraging AI's capability to identify patterns and optimise adequate trading decisions that would typically be beyond human capacity. This approach incorporates advanced algorithms and machine learning techniques to adapt real-time trading strategies based on live market data in a dynamic way. This allows for a flexible plan that considers the cyclical nature of financial markets. Reinforced learning (RL), in particular, enables the integration of highly complex factors into trading strategies in real-time, facilitating the integration of complex factors. Reinforced learning means that the algorithm has to make decisions or take actions, and based on the outcomes (compared to a given optimal result) are used in order to avoid making mistakes in the future or to repeat favourable behaviour. ESG scores may not be considered complex factors as they rarely change, but when it comes to the way machines process information, every additional factor increases the processing time exponentially, especially when such environmental scores are not the only factor being taken into consideration. The unique synergy offered by algorithmic trading and reinforced learning also facilitates the transformative shift that is currently undergoing in the world of trading towards a more efficient, informed, and sustainable investment approach that aligns not only with financial performance goals but also with environmental and ethical aspects of trading, making the industry more appealing to a larger audience. Studies show that RL algorithms outperform baseline models and deliver profits even under unideal conditions (Zhang et al., 2020).

2.1.3 Algorithmic Efficiency in Predictive Trading

To reinforce the ideas mentioned in the previous section, we can take a deeper dive into the specific literature and find out that in the context of predictive trading, the role of AI extends from data mining and analysis (pattern identification) to real-time decision-making, allowing traders to act swiftly on market opportunities. The concept of algorithmic efficiency challenges the traditional hypothesis of EMH by highlighting that with data mining, real-time decision-making, and risk management optimisation, the unpredictability of markets disappears. With such a rapid reaction time, more information is beneficial, as it does not "confuse" the system as it would normally do to a traditional trader. Prior research has confirmed the efficiency of such predictive models in short timeframes, where fluctuations are heavily speculative and traditional price action works, meaning the market could fit in the weak-form EMH model. In such conditions, the predictive models were significantly more efficient than human traders (Syamala & Wadhwa, 2020). Environmental factors may not have a short-term evident effect (on 1-minute and 5-minute timeframes, for example) as their effects are typically long-term. For this reason, such factors are generally more effective for investing or swing trading. Swing is a type of trading in which the traders leave trades open for multiple days, even weeks, hoping to catch big moves in the market. This requires a significantly different approach to scalping (getting relatively more minor gains by correctly predicting short-term market movements). Scalpers typically leave their trades open for only a few minutes and use advantageous risk-to-reward ratios to ensure profitability. This study will uncover one interesting takeaway, investigating the efficiency of environmental factor integration within a one-hour time frame, a time frame often used by swing traders.

2.1.4 Adaptive Market Hypothesis

According to AMH, markets are not static but dynamic, evolving, "learning," and adapting over time. This means that markets get modelled to some degree by investor behaviour over time, and this factor should be considered when making financial decisions. If the current trend in many industries, not only Finance, is to be more environmentally aware, according to AMH, soon the markets will "learn" from these differences and will theoretically slowly start favouring companies that suit these new requirements made by investors wishing to help protect the environment. As traditional algorithms cannot change their strategy and adapt dynamically as they go, AI trading is left to take the lead. ESG score embedding into neural network models may be a solid option when considering AMH and its implications. This theory is central to the Efficient Market Dilemma, where the hypothesis is that you cannot beat the market due to the overwhelming amount of information. Until now, studies showed mixed results regarding the impact of bear vs bull markets on the theory of return predictability, suggesting that markets respond differently to certain conditions, underlining AMH's notion that markets adapt and evolve (Urquhart & McGroarty, 2016).

2.1.5 Stakeholder Theory in Financial Markets

Traditionally, it is managers' duty to ensure that they act in the best interest of the firm's shareholders, as managers are the stewards of the company, and owners bear the most risk (Menyah, 2013). When considering other approaches to running a business, some companies choose the path of stakeholder theory by taking into account other actors in the macro environment. This way, firms could be positively impacting society and the environment, as this may not only be the ethical thing to do but also beneficial, in the long run, as it promotes balance between profits and the well-being of all parties involved (Wallace, 2003). Firms could not pursue their long-term goals (five or ten years plus) if the communities where they operate were not around. There would not be any more workers to perform the tasks and no more customers to purchase the goods or services provided. By similar logic, many businesses would be unable to operate without a healthy environment where they could develop their activities, and this is also one of the reasons investors consider ESG scores.

Companies and financial institutions that consider the environment as a stakeholder and take it into account are acknowledging their responsibility to operate sustainably, reducing harm to ecosystems and biodiversity. If CO_2 emission data and ESG scores were considered in developing trading algorithms, it would become possible to have environmentally safe investment practices over time. Many investors may question the usefulness or practicality of such criteria; this study can enlighten us on these doubts to help future investors make informed choices.

2.1.6 Robo-advisors

The study "Robo-advisors: A substitute for human financial advice?" by Lukas Brenner and Tobias Meyll investigates the impact of robot advisors on investors' likelihood of seeking human financial advice (Brenner & Meyll, 2020). The study found a strong negative correlation between using robo-advisors and seeking human financial advice, suggesting robo-advisors as a preferable option for investors concerned about potential conflicts of interest between their own and those managing their funds. These findings show that robo-advisor offer a solid alternative to financial advice.

2.2 ESG scores in trading

To dive deeper into the topic and get a more comprehensive understanding, the following section will begin by highlighting the benefits and disadvantages of environmental practices. Studies (see section 2.2.1) have shown that companies with high ESG scores often enjoy better returns, demonstrating the growing relevance of ecological practices in Finance. On the other hand, there's studies showing the lack of immediate impact on stock prices and the lack of general impact on the stock prices, making it hard to use this information in predictive trading (see section 2.2.2).

2.2.1 Benefits of ESG scores

Thanks to such studies, more individuals consider ESG scores when making investment decisions, resulting in companies who follow such practices getting better access to capital. It also makes other firms consider engaging in an eco-friendlier approach. However, despite its popularity, the lack of mandatory ESG disclosure leads to inconsistent data quality and variances in ESG scores, impacting stakeholders across the board (Ostbirk, 2022). For this reason, when comparing different companies based on ESG scores, it is essential to have a reliable and consistent source. For this reason, it was decided that msci.com would serve as a reliable source of ESG and CO_2 rating scores for this study. All the companies analysed in this study are listed on the S&P 500, and music is considered a leader in providing critical decision support and accurate information, specialising in S&P 500 companies. In the paper Osbrik, the call for a universally accepted disclosure framework and mandatory reporting is emphasised, alongside the examination of potential regulatory approaches that would enhance the credibility and effectiveness of ESG metrics in promoting a sustainable future for trading.

MSCI provides ratings for all companies in the S&P500, alongside insights into the importance of these ratings and information on how they can be interpreted. Such insights can significantly enhance AI trading strategies by embedding these considerations into algorithms, potentially improving their predictive accuracy and financial returns while promoting ethical investment practices (ESG and performance). This integration underscores the evolving landscape of financial markets, where technology and sustainability converge to define the future of investing.

2.2.2 Drawbacks of ESG implementation

ESG scores do not come without a hitch, however. When considered as a variable, part of a broader investment strategy, their downsides must also be considered. Studies show that ESG scores do not contribute nearly as much as credit ratings to profitable outcomes. This is largely attributed to the lack of standardisation when it comes to ESG scores and to the lack of immediate impact on a company's financial performance. The inconsistency in ESG ratings also leads to companies receiving mixed signals on the topic, diminishing, therefore, the impact of their efforts to maintain good environmental, social and governance etiquette. Berg et al. suggest in their 2022 paper "The divergence of ESG Rating" that greater collaboration among agencies and companies, more transparency and cooperation could lead to a more standardised measurement approach (Berg et al., 2022). Until these ideas come to life, ESG score's impact on trading remains questionable. When it comes to AI algorithmic trading, more precisely, recurrent neural networks (RNN), these downside effects may get amplified, as without quality (consistent) data, the added benefit to training the network may be reduced.

Aside from having to deal with inconsistent data, investors are also faced with another challenge, that being the lack of a real link between traditional ESG and financial performance, making SRI (Socially Responsible Investing) ineffective in many cases, failing to deliver superior returns to shareholders. One additional drawback of ESG implementation into intelligent trading models is the general misconception that ESG scores are risk factors when, according to Porter et al., they should be rather seen as integral to competitive strategy. By doing so, investors have a pessimistic approach to ESG scores, not necessarily seeing a high score as a beneficial thing but more as the negation of a potential risk caused by a low score (Porter et al., 2019). Traditional trading is generally focused on short-term gains, as opposed to investing, where a generally longer-term perspective is used. Short-term ESG scores may lack an immediate impact, as mentioned earlier, leading to potential missed opportunities, especially if trades who are short-term oriented decide not to enter a trade in order to align with SRI, running the risk that the potential trade would have never experienced any impact directly correlated to a negative ESG rating, as the duration of that hypothetical position would not have stretched far enough to see a significant impact on its financial performance. In theory, a complex enough neural network, with sufficient training, could understand the effect (or lack of such) and act accordingly, but in practice, the additional information may slow down the system and reduce its efficiency, causing neurons to be busy with a variable that may not have a significant impact on the addressed time frame. Such issues may become obsolete with the implementation of quantum computers, which are capable of analysing data at very high speeds.

2.2.3 Transparency in ESG scores

Going more into detail about the measurements listed on msci.com, other concerns, aside from consistency, arise, such as transparency. To what degree all the data on MSCI is accurate and represents reality is heavily influenced by how this data is collected. The credibility of the ratings is different for ESG scores when the companies rate themselves, compared to when there is an external agency that goes inspect each company and rates it by a standardised list of criteria, assumably without the knowledge of the company, avoiding firm from “covering up” parts they would not want the public to know. Either of these scenarios is not what happens in reality, one being an untransparent way of measuring and another being an ideal setting where the public has a perfect idea of how environmentally friendly a company is. To a large extent, the impact ESG scores have on stock price is due to public opinion, as investors avoid unethical companies due to fear of backlash from the public and the public trying to pick products and services that are ESG friendly. If transparency is unsatisfactory, the public may stop trusting it, rendering it useless or inefficient.

MSCI has different standards for different industries and compares companies within an industry equally. This is why it is essential to use agency data, as it is impartial, and the agency has no interest in taking sides. MSCI uses company data, stakeholder data on the company, and additional alternative sources. Firms do not need to be entirely transparent, as MSCI does not base their ratings solely on the information the firm is giving to the public. The ESG scores on MSCI indicate the environmental risk investors are exposing themselves to by investing in this firm (Apiday, 2023). Therefore, these ratings are consistent and largely reliable, meaning they can be used as a trustworthy source of financial information that could aid financial decision-making. When it comes to emission goals, it becomes easier to quantify, as a company's goals are made public, and companies are rated based on what they make public related to their goals, not their actual performance.

2.3 CO_2 in trading

Alongside other environmental factors, CO_2 emission goals and awareness related to firm carbon footprint have gained significant relevance in the academic field in recent years. These metrics became a popular method to evaluate how environmentally responsible a company's practices are. For policymakers, this has become a more pressing issue, as seen in the European Union's initiatives to limit large corporations' emissions through taxation (EU Green Deal, 2021) (Fit for 55 - the EU's plan for a green transition - consilium, 2023). Similarly, in the United States, regulations such as the "Clean Air Act" govern CO_2 emission rules (EPA, 2024). These policies make it mandatory for firms to follow and abide by certain standards, no longer making it a preference to stay green.

Understanding these factors and how they come into play is becoming more relevant than ever in the context of trading, especially for investors with a long-term perspective. To align socially responsible investment practices with financial performance, it is essential to fully understand and research the impact of CO_2 emission goals. This will later help better understand how they impact trading outcomes for neural network models.

In the following sections, we will take a closer look at the potential benefits and drawbacks as well and try to understand how they impact AI trading models and how they could be implemented.

2.3.1 Benefits of CO_2 emission goals

Incorporating CO_2 reduction goals into a trading strategy comes with several benefits that can enhance financial performance and sustainability alike. The main significant advantage of CO_2 consideration is the potential for improved public perception or, contrary, the general public scepticism that could be assimilated with the lack of carbon footprint reduction goals. The presence of these factors may consolidate trends and it could potentially facilitate deep learning. Additionally, a positive outlook on the company could potentially lead to increased investments, as highlighted in the study "Impact of Environmental, Social and Governance (ESG) Factors on Stock Returns of Emerging Markets", which found a positive correlation between high environmental scores and stock performance in emerging markets (Bag & Mohanty, 2021).

Another benefit of CO_2 goals is that if these goals get reached, firms could be entitled to regulatory incentives often offered by governments in many jurisdictions. It is not uncommon for governments to offer incentives to companies who agree to commit to reducing their carbon footprint, therefore increasing the firm's capital and reducing the cost associated with trying to attain these goals. The US government provides approximately 6.4 billion USD annually to firms that implement and maintain CO_2 reduction measures. These incentives could, in some cases, offset the costs associated with achieving these goals, making it a more financially viable strategy (Dot funding programs and climate change). As observed in the Stanford Study mentioned in the introduction section of this paper, companies with robust carbon reduction strategies often experience enhanced operational efficiencies (Stanford HAI, 2023). By optimising energy use and reducing waste, firms can lower their operating costs, potentially leading to better profit margins.

2.3.2 Drawbacks of CO_2 goals

Whether or not a company actively incorporates CO_2 goals into its business models can significantly impact trading outcomes, offering both opportunities and challenges for traders who consider these factors. In most cases, for a firm to attain such goals requires much additional effort, requiring internal changes, coordination, and collaboration with suppliers and customers. Additionally, firms need to put in extra effort to make accurate measurements to determine their carbon footprint. In some jurisdictions, firms receive government incentives to compensate for these costs. Despite the incentives mentioned in the previous section, in specific industries, traditional practices may still be more cost-effective, even with the consideration of government grants. This is due to the fact that a lot of these goals cannot be attained without substantial investments in newer technologies, which often do not come at a low cost. This makes it harder for those who try to be environmentally friendly due to the need for proper legislation. The lack of a comprehensive regulatory framework following a national or internationally agreed-upon structure rather than a regional one is also having a great impact on the matter. Large firms often operate in multiple jurisdictions, and the lack of alignment among these can make it difficult to follow up with local laws when, in one jurisdiction, these goals are attainable because everyone has to follow them, and in others, it does not, due to lack of regulation. This can make it difficult for companies to adapt their practices and may end up taking the easier route, or if they decide to make the extra effort, they may run additional risks compared to their competitors, who are not following environmentally safe practices.

Another challenge arises from many companies' inability to measure all their emissions accurately. The lack of comprehensive measurement means that investors cannot compare accurate data. They would have issues differentiating between firms “playing the game” and those who do not, even if they are theoretically competitors in the same market (Shepherd, 2021). When a neural network tries to take CO_2 emission goals into consideration, it would at first take them as absolute, and if they prove not to be directly correlated to the firm’s performance, this would only be noticed epochs later in the training, slowing down the deep-learning process.

It is crucial for investors, human or machine, to have a straightforward way to differentiate among such inconsistencies in the industry to navigate the financial landscape better. Firms receiving millions of dollars in government grants as compensation for their efforts to reduce CO_2 emissions can represent a determining factor in decision-making. Long-term investors can consider these companies better equipped for the future as they protect their environment and care for their communities. On the other hand, the variability in emissions causes information asymmetry in the market, almost making it into a “market for lemons” (Levin, 2001).

2.3.3 CO_2 goals in trading

As seen from sections 2.2.1 and 2.2.2, CO_2 integration comes with both opportunities and challenges for investors, companies and policymakers. On one side, firms successfully integrating these carbon reduction strategies can benefit from enhanced public perception, regulatory incentives and added operational efficiency. On the other hand, we observe the drawbacks associated with CO2 goals, such as data inconsistency, high implementation costs, and regulatory disparities. For AI trading algorithms the risk could be the unnecessary overcomplication of the algorithm, leading to slower or less efficient outputs, while at the same time generating potential by adding a new metric that may create some corelations.

2.4 Temperature awareness and trading

Another increasingly important factor in the context of environmental sustainability and corporate responsibility is the level of awareness related to the organisation's impact on temperature. Climate change continues to impact global weather patterns (Canadell et al, 2023) at an alarmingly rapid rate, increasing the importance of temperature goals in corporate strategy (Newman & Noy, 2023). A way to measure a firm's impact on the planet's temperature is directly correlated with the amount of CO_2 emissions as a result of power consumption by projecting the future temperature impact on the planet's temperature over a set period of time. For this reason, companies are now setting temperature-related goals to reduce their negative impact on the environment and contribute to the ongoing global efforts to combat climate change. The following sections will explore the benefits and drawbacks of embedding temperature goals and temperature impact estimations into trading strategies and, finally, their implications for AI trading models.

2.4.1 Benefits of temperature goals

Similarly to the case of ESG scores and CO_2 emission goals, the main added benefit that comes with the implementation of temperature goals is the potential for improved public image, which could aid the consolidation of a potential uptrend or downtrend if we are talking about a lack of such goals aligning with a downtrend, figuratively speaking "adding gas to the fire". Looking at it from an optimistic perspective, companies that commit to temperature reduction goals are often seen as leaders in environmental responsibility, which can potentially enhance the organization's reputation, attracting investors interested in this type of investing. A study by Eccles, Ioannou, and Serafeim in 2014 found that companies with a strong ESG ethic (including temperature goals) tend to have better long-term financial performance. However, the study does acknowledge the potential involvement of a third variable, meaning that unobserved factors influence both sustainability performance and financial outcomes, posing real challenges in determining causal relationships between the ESG factors and the firm's success. It may very well be that firms that have sufficient resources to spend on ESG approaches are also having a sustainable business model that consistently generates cash flow, therefore giving the robustness of an uptrend. Such factors may not be picked up or interpreted by a neural network model, as the model is simply fed a value without any additional explanation. However, this shouldn't in any way hinder their efficiency because whether or not temperature goals impact the stock's performance, as long as the two values are correlated, the level of causality is relatively irrelevant to the model.

Temperature goals also benefit from regulatory incentives similar to those of carbon footprint reduction plans. Governments increasingly implement new regulations and incentives aimed at encouraging businesses to adopt temperature reduction measures and make them part of their business plan. The European Union's "Fit for 55" plan, mentioned in section 2.3, also includes this metric and aims to reduce greenhouse gas emissions by at least 55% by 2030, and hopes to achieve this through the cooperation of firms motivated by incentives awarded for reaching these goals. Compliance with such goals not only entitles firms to these rewards but also protects them from potential penalties. Additionally, reaching these goals can also help attain better operational efficiencies and contribute to cost savings. Lowering energy consumption where possible or switching to a more environmentally friendly method can lower utility bills and reduce reliance on fossil fuels. A report by the International Energy Agency indicates that if mass adoption, energy efficiency measures could save the global economy more than 500 billion USD annually (Iea, Executive summary – energy efficiency 2022 – analysis).

2.4.2 Drawbacks of temperature goals

One major challenge that arises from the high implementation cost, which may not be viable for firms with a smaller market cap. Achieving these temperature reduction goals often requires a substantial investment (McCollum et al, 2018), considering the fact that new technologies are often more costly than traditional methods and adapting the infrastructure to the new process may prove to be logistically challenging, especially for firms with limited available capital. A study by the World Bank (2021) indicates that the initial costs of transitioning to more energy-efficient systems can be, in many cases, too high for most SMEs. This can be particularly troublesome for companies in the EU or USA since companies here face stricter environmental regulations and compete on the market against products that were fabricated in countries where such terms do not apply, therefore competing at a lower cost. This results from the lack of internationally accepted standardised measurements and policies. Unlike financial metrics, environmental ones, including temperature goals, often lack consistency, especially when looking at a regional or even country level. As noted by Berg et al. (2022), the lack of consistency and the overall divergence of standardised measurement leads to confusion and misinterpretation among investors, policymakers and other stakeholders.

2.4.3 Temperature goals in trading

Due to all the similarities in both benefits and drawbacks of CO_2 emission goals and temperature goals, these two variables will be tested together in the model in order to create some more synergy and hopefully allow for better training results. An additional benefit of mixing these two is that we can test the effects of one added neuron and two added neurons alike.

2.5 Environmental practices in AI trading

To better understand the influence of AI on financial markets in relation to environmental scores, we can first take a look at Yadav's paper from 2019, "Impact of Artificial Intelligence on Trading in Financial Markets," which reveals a change in trend towards a more data-driven decision-making approach and more relevant algorithmic trading strategies. Since AI trading bots change their strategy in real-time on the basis of large quantities of data (something a human trader could only do through intuition), it resulted in unprecedented accuracy in market predictions, outperforming traditional trading methods in profitability. The paper lays the ground for further research in AI trading. It sets a direction for additional research to investigate more sophisticated trading algorithms considering more factors (Yadav, 2019). Some of these factors, such as in our case, can be Environmental, Social, and Governance scores or CO_2 reduction plans and commitments made public by companies. This approach aligns financial performance with sustainable and ethical investment practices.

To better understand the potential impact of environmental factors and how they are integrated into trading algorithms, it is essential to have a good understanding of previous literature on ESG factors in financial performance. For this reason, we will go (now in more detail) over the three articles previously mentioned in the introduction about the impact of ESG scores. Bag and Mohanty (2021) provide compelling evidence in their study of the impact of ESG factors on emerging markets and stock returns, suggesting a positive correlation between high ESG scores and stock performance. Our research will analyse the effect on already mature markets and see if the same principles apply to large enterprises. Similarly, MSCI has investigated the relationship between ESF scores and stock performance, again concluding that companies with strong ESG credentials may outperform their counterparts in the long term. Floros' study further supports this, pointing out the impact hearing behaviour can have on stock market returns. The paper points out that one of the potential drivers of better performance for firms with high ESG ratings may not necessarily come from the direct benefit of these practices but from the public opinion's views on the firm after being associated with these high scores. Regardless, the study confirms their impact and further convinces us to test the effect of environmental metrics on AI trading bot's performance. This means that by integrating these factors, trading strategies align with ethical and sustainable practices and the general trends set out by public opinion.

Prior research in the embedding of environmental metrics into AI trading algorithms has provided a foundational understanding, yet significant gaps remain. A notable example is the recent study published in Nature by Lee et al., investigating the impact of ESG sentiment analysis on the prediction accuracy of deep learning models (Lee et al, 2024). The study, conducted on an LSTM model attempting to predict movements of the S&P 500, demonstrated that sentiment analysis can enhance the performance of a predictive AI model. Building on this, our paper delves deeper into the interaction between environmental factors and deep learning models by focusing on a selection of individual stocks also belonging to the S&P index and focusing on a smaller time frame. While the paper of Lee et al. utilises ESG sentiment analysis derived from textual analysis of news and reports, our study takes a different approach by using actual ESG scores from a reliable source. By using a more quantitative measure, we are exploring a new interaction that offers a more stable indicator of a company's ESG impact, potentially reducing the volatility introduced by the market. On the other hand, the lack of variation in ESG scores may render the model slower than the dynamic approach used in Lee et al. The paper also makes use of error measurements to determine the model's performance, precisely the use of mean absolute percentage error. Similarly, this paper will make use of three error measurements in order to determine the performance of the model: MAE, MSE & RMSE. These three error metrics will finally be compared for additional synergy, and their alignment will further confirm the findings. Additionally, this paper will explore the interaction of two other environmental factors, CO_2 emission and temperature goals.

2.6 The link between Quantum Finance and AI

Quantum computers are guaranteed to revolutionise the world of finance. Over the years, the stock market has been heavily shifted by innovations in the field of information technology. The digital revolution changed the financial landscape, replacing the crowded rooms of stock exchanges and making tasks more digital. Quantum computers will be guaranteed to disrupt the stock markets. Highly complex algorithms and supper computers available today cannot even come close to 1% of the power of quantum computers working on the superposition principle. Traditional computers work with 0s and 1s, yes or no, following the binary system. Quantum computers, on the other hand, can have both 0 and 1 at the same time by operating at temperatures close to absolute 0, conditions in which particles are moving significantly slower, allowing for the so-called superposition. The implications for artificial intelligence in trading are tremendous. To better visualise the matter, let's use an analogy. Imagine the AI as a person trying to cross a labyrinth. The exit from the labyrinth is represented by profitable trading or accuracy in prediction. When AI is powered by a traditional computer, it must try every possible way before finding the right one. The AI will go on a path, and if that path is not the right one, the code will register it as bad and try a different route again. This analogy refers to how neural networks must undergo many epochs of training before finding the right weights for every neuron. However, an AI running on a quantum computer will try all the paths simultaneously, solving the equation (or finding the way out of the labyrinth) in the shortest amount of time possible. When we consider a system that takes into consideration vast amounts of data, we could end up with a trading system that doesn't take into consideration only one instrument but the whole market at once performs data mining, and analyses millions of factors every second, giving the best possible predictions with never before seen accuracy. In theory, such a powerful algorithm could solve the stock market, similarly to how modern software such as Stockfish solved chess. The impact on stocks may be up for discussion due to insider information. However, the impact on forex would be tremendous, especially if it were in the hands of banks or governments. Market makers have long been known for manipulating the market through techniques such as liquidity sweeps or filling fair value gaps. The level of market manipulation they could attain by using advanced AI algorithms paired with the sheer power of quantum computers may change the world of finance forever unless governments and policymakers act on it (Montanaro, 2016) (Orús et al., 2019).

Integrating quantum computing-powered AI into price prediction holds great potential and is likely to have a great transformative impact on the industry. As mentioned in the previous section, despite being powerful, traditional models are limited by the computational constraints imposed by the binary system, which restricts the speed at which they can process large amounts of data. This is similar to the red queen effect, where one needs to move very fast just to keep up with the others. Algorithms have to process all the available information within a given time frame, and the smaller the time frame, the faster these decisions need to be made. Before the traditional computer is done solving all the affiliated equations, new data has come out, and it needs to be analysed. Quantum computers, leveraging their ability to process multiple scenarios simultaneously through superposition, can overcome such boundaries, offering price predictions at never-before-seen speeds that drastically enhance the accuracy of the results in the process. When applied in real life, traditional AI often requires heavy simplification of the parameters in order to deliver timely predictions. This would not be the case for quantum-powered AI machines, meaning that quantum computers paired with AI may indeed be the final answer to the efficient market hypothesis, as mentioned in section 2.1.1. In short, predictions are more valuable the earlier you can have them and do not serve much of a purpose once the events they are attempting to predict have already happened (Rosenberg et al., 2015).

In summary, the fusion of quantum computing and AI represents a great leap forward in the world of finance and has many implications for predictive models. As this technology evolves and becomes more accessible, it will undoubtedly play a crucial role in shaping the future of finance.

2.8 LSTM models

Long Short-Term Memory (LSTM) is a type of recurrent neural network that is commonly used in deep learning and is generally characterised by the fact that it is very good at remembering the context of the situation and is modelled to analyse but at the same time, can forget with the help of forget gates (Hamad, 2023). This means that while using less of the available memory, the model can understand the general trend while forgetting bad patterns relatively quickly. Neural networks are composed of neurons, which take an input, process it with the aid of an equation, and generate an output. Recurrent neural networks work on the same principle; the only difference is that the output is recurrent because it will be taken into consideration in the next input. One issue, however, that arose with RNN is long-term dependency, which is particularly relevant when applied in trading. Traditional RNNs take into consideration older data alike, which, in the case of the stock market, is not always good. For example, if a company is going through different stages of organisational development (growth, maturity, decline), the patterns the prices follow may change. However, in such a scenario, a traditional RNN would still use data from years ago. What LSTM models do is keep only what the model deems to be essential from the past and, therefore, give more weight to recent events. In the figure below, you can see a graphical explanation where the first graph represents a normal neural network, the second a recurrent neural network, and the third represents an LSTM model, with the addition of the state to the cell.

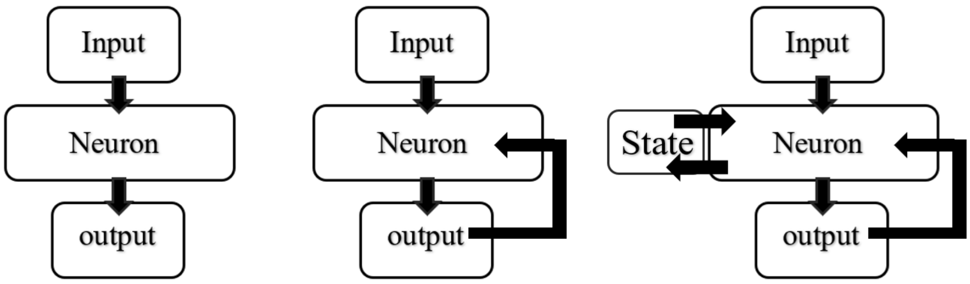


Figure 1

2.9 Theoretical framework

The theoretical framework of this study is based on recent advancements in LSTM models for stock price predictions, building on the work of Botunac et al. (2024) and Dahal et al (2023. These papers provide us with a comprehensive foundation that aids in the understanding of how different factors impact the performance of LSTM models in market forecasting scenarios.

In the study "Optimization of Traditional Stock Market Strategies Using the LSTM Hybrid Approach" by Botunac et al., the integration of traditional trading strategies in LSTM models was investigated, evaluating the performance enhancement of LSTM models. In this study, various performance metrics are used, such as the MAE, RMSE, and trading profitability. Various configurations and architectures of neural networks were tested prior to determining the optimal settings for stock price forecasting (Botunac et al., 2024). Our paper will take a similar approach, testing the impact of various environmental factors, while the results will be assessed via error measurement and by comparing the results of a simple trading simulation.

Dahal et al. provide a comprehensive analysis of LSTM and GRU (gated recurrent units) models in their study, investigating the significance of incorporating financial news sentiments alongside traditional trading indicators and stock features (open, close, high, low, volumes, etc.). This paper will not investigate the interaction of environmental factors with GRU models, leaving this open for further research. On the other hand, we will build on the paper's approach by incorporating LSTM models and analysing their performance (Dahal et al., 2023). GRU is a gating mechanism used in recurrent neural networks, first introduced in an extensive study in 2014. Functions are very similar to LSTM models, but some of it's features cause it to have fewer parameters. GRU also make use of hidden activation functions, where the hidden layer state is to be updated with a new hidden state (Cho et al., 2014). The paper was revolutionary as it introduced a new type of RNN and allowed for a multitude of further studies to follow based on its model.

Following the approaches used in the two cited papers, the model developed for this paper will be configured to process historical price data, following Botunac et al.'s method of optimising the data. This includes tuning the network architecture and hyperparameters through testing in order to attain a higher level of predictive accuracy. To align with the theoretical framework's strong emphasis on dynamic performance evaluation, the code will predict batches of five-time intervals, and based on the resulting error, once the data is shown to the model, the model will readjust its strategy, attempting to optimise for a smaller error. All other model choices are further described in section 3.5, "Model Choices".

2.10 Hypothesis Development

To cover the gaps left in the existing literature by the fast development of AI and to help attain its social benefits caused by the unexplored intersection of AI trading algorithms and environmental scores, the hypothesis of this paper aims to investigate how the integration of environmental metrics impacts AI predictive and trading performance and whether it can be a valid way to enhance financial performance and improve risk management. One way in which risk management can be measured is through the measurement of maximal drawbacks. This means that the EA will minimise it's losses. Furthermore, to fully assess the impact, this paper will investigate a few factors, starting with the direct impact on the model. Before assessing the impact of environmental scores on trading performance, we will first take a look at their impact on the deviation from reality in their predictive attempts. This will be tested by measuring MAE, MSE and RMSE. The aggregate error readings will be compared, and the one that generates the least amount of aggregate error will be considered to be the more effective predictive model. Based on this and on the reviewed literature, the following hypothesis will be tested in order to cover a larger spectrum of metrics:

H1: Integrating environmental factors into an LSTM predictive model has a positive impact on the model performance.

Since this study will investigate the impact of multiple environmental factors, H1 will be broken down into several smaller hypotheses:

h1a: Integrating ESG scores into an LSTM predictive model has a positive impact on the model performance.

h1b: Integrating CO_2 goals and temperature targets into an LSTM predictive model has a positive impact on the model's performance.

h1c: Integrating ESG scores, CO_2 goals and temperature targets into an LSTM predictive model has a positive impact on the model performance.

The impact will be tested, as mentioned earlier, by measuring the error, but the performance of a model can also be measured by its financial performance. For this reason, we will also test the financial performance in two different ways: maximum drawdown and final account balance as % return on the given testing period (November 2022 to May 2024). The overall return is important because this represents the quintessence of trading and every trader's goal: to generate returns. Additionally, the model will be tested for maximum drawdown because it is important to understand the level of risk associated with the returns, "as not all dollars are born equally". Attaining a high return with a high level of risk is often not sustainable in the long run, and by incorporating drawdown into the analysis, we can better understand how the models performed in the various scenarios. This period will represent approximately 30% of the available data. From here, we come with two more hypotheses to test; the first one is related to maximal drawdown:

H2: Integrating environmental factors into an LSTM predictive model reduces maximal drawdown in trading.

$H2_0$: Integrating environmental factors into an LSTM predictive model does not reduce maximal drawdown in trading.

And another one testing the total return after the testing period:

H3: Integrating environmental factors into an LSTM predictive model increases returns.

$H3_0$: Integrating environmental factors into an LSTM predictive model does not increase returns.

Similarly to H1, the hypothesises will be broken down into multiple smaller ones:

h2a: Integrating ESG scores into an LSTM predictive model reduces maximal drawdown in trading.

h2b: Integrating CO_2 goals and temperature targets into an LSTM predictive model reduces maximal drawdown in trading.

h2c: Integrating ESG scores, CO_2 goals and temperature targets into an LSTM predictive model reduces maximal drawdown in trading.

h3a: Integrating ESG scores into an LSTM predictive model increases returns.

h3b: Integrating CO_2 goals and temperature targets into an LSTM predictive model increases returns.

h3c: Integrating ESG scores, CO_2 goals and temperature targets into an LSTM predictive model increases returns.

Chapter 3: Methodology

3.1 Approach

In the Methodology section, the general approach to the study will be described along with some pre-requisite information necessary for the context of the Data section. The data from ten S&P 500 companies from different industries will be used to test the AI-powered algorithm that will be developed for this study, which will conduct a parallel analysis of the results of four different approaches on the same company data over the same time interval. Firstly, the standard neural network machine learning system will be tested, serving as a baseline, where the AI system will operate without considering environmental factors. Secondly, the same AI system will be tested with the implementation of ESG scores, allowing us to assess the impact of ESG scores on trading performance. Similarly, the network will be tested by implementing CO_2 emission goals alongside temperature impact metrics. Finally, both CO_2 and ESG scores will be implemented to see if the synergy of the two can benefit the results. Because these firms are in different industries, we can also make assumptions about the universality of the results and determine how they stand in different industries. This approach will help close the gap between AI trading algorithms and the role of environmental factors in the financial world.

3.2 Data collection

The data for this paper was taken from a premium Meta Trader account. The data was taken on a one-hour timeframe for 10 companies. This data was split into 2 sections: training and testing. The training data itself will also be divided into 3 sections: training, validation and testing. The double division of testing data was done in order to assure no overfitting. The other part of the used data consists of the ESG scores themselves. These have been collected from msci.com and had to be allocated to each individual time frame. This step is necessary because the RNN needs to work with arrays of equal size and shape. Additionally, the data needed to be normalised. This was done using the MinMaxScaler provided by the sklearn package. To simplify the way it works, the algorithm found two key values (the smallest and largest ones) within each dataset (in the closing price category being the lowest and highest value, for example) and used them to build a new scale. In our case, the smallest value becomes 0, and the largest becomes 1. In this way, the normalised data can be compared across the board, meaning the weights later on attributed by the network will be comparable, aiding in the learning process.

3.3 Algorithm description

This study will make use of Tensorflow's premade packages: Keras, Sequential, LSTM and Dense. These packages will be used to develop a simple machine-learning model that will serve as a benchmark for this study. The model will consist of one input layer. The size of this layer will vary among the four algorithms. The one basing its predictions solely on historical prices will have four neurons on the input layer. These will stand for open price, close price, high price and low price. The ESG model will incorporate five total neurons, while the CO_2 one will be paired with the temperature one, resulting in six input neurons on the first layer. The two environmental metrics were paired due to how closely related they are, and in the literature section, we could observe a similar positive and negative impact on trading efficiency. Due to these synergies, they have been paired, allowing for the testing of multiple types of input (four, five, six and seven neurons). The aforementioned four price parameters and an additional one for the CO_2 and ESG, respectively. The fourth model will take into consideration seven parameters: all the aforementioned ones and additionally a temperature impact one. Some of these values have a quite linear progression or a purely linear progression, depending on which one we are referring to. This may have both benefits and downsides. It may benefit the system by adding another source of information and by smoothing the line. If the value is not helpful, the system should, theoretically, over sufficient epochs, learn its inefficiency and simply reduce its attributed weight to virtually zero. However, there is the risk that these added parameters would hinder the learning process and, at best, slow it down. Even if the network eventually learns to ignore (in the event of a negative outcome), this will take more epochs to clear out than if the system never had this parameter to begin with. If the effect is positive, however, this may also take additional epochs to learn. In order to determine the number of epochs to be used, a small-scale experiment was conducted, and the results indicated that one hundred epochs were a good amount of training for the model, showing better performance compared to 10, 50 while still maintaining a manageable processing time, compared to 150 and 200, cases in which overfitting could be a risk. In order to assess the effectiveness of the newly embedded factors, an error analysis will be conducted. In order to address the first hypothesis related to the impact on the model's performance, measurement of error is crucial.

3.4 Model choices

In this section, the various choices that were made in relation to the construction and training of the model are explained. The code was developed in Python, a programming language known for its widely available premade and community-made packages. These packages are ideal and commonly used for the development of a machine-learning model. From these libraries, the ones used to develop the neural network are the TensorFlow and Keras libraries. These are very popular machine learning libraries that come with premade functions, aiding in the development of the model. One available alternative was the PyTorch library. This library offers a higher level of flexibility but also comes with an increased level of difficulty, and it's not as user-friendly as TensorFlow.

Before diving deeper into the explanation of all other choices made in the development of the model, it would help to visualise the model and explain it in a function. Bellow, in Figure 3, we can see a simplified version of the model, with four input neurons, 2 linear stacks of layers, and one output layer, resulting in the 5 predictions. The diagram would look slightly different for the models used in the training of the ESG, CO_2 and combined, having additional input layers. In Figure 3, we can see the complete diagram, showing the true complexity of the model.

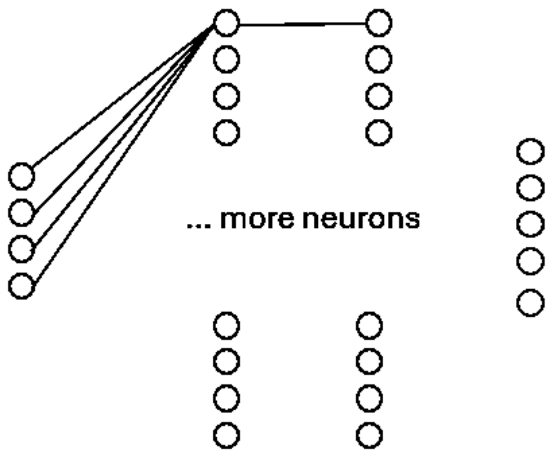


Figure 2

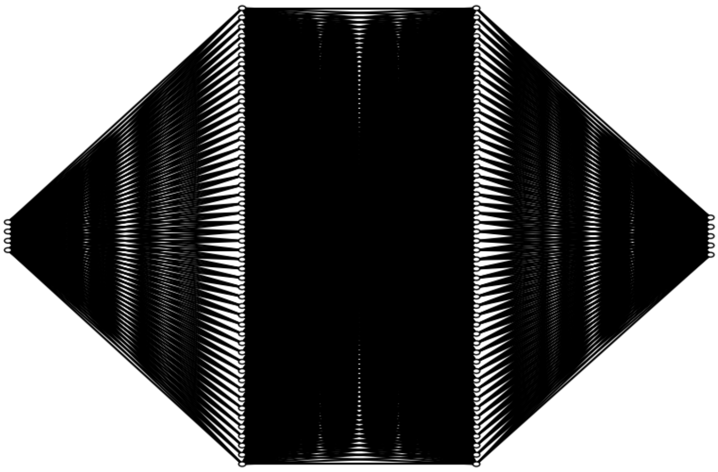


Figure 3

The core of any neural network model can be summarised with the aid of an equation (Thakur, 2018). In this special type of RNN, we are making use of gating units to control and direct the learning and forgetting process. Typical LSTM models (this one included) make use of three types of gates precisely: the input gate, the forget gate and the output gate (DeepAI, 2019). The input gate determines how much of the new information is being used and stored in the cell. The role of the forget gate is to determine what information should be discarded from the cell. Finally, in the output gate, the amount of information that passes from the cell state to the output. These gates serve as the sigmoid activation function in LSTM models, simply known as transfer functions. When graphed out, these functions form an “s shape” on the chart, as seen in the graph below:

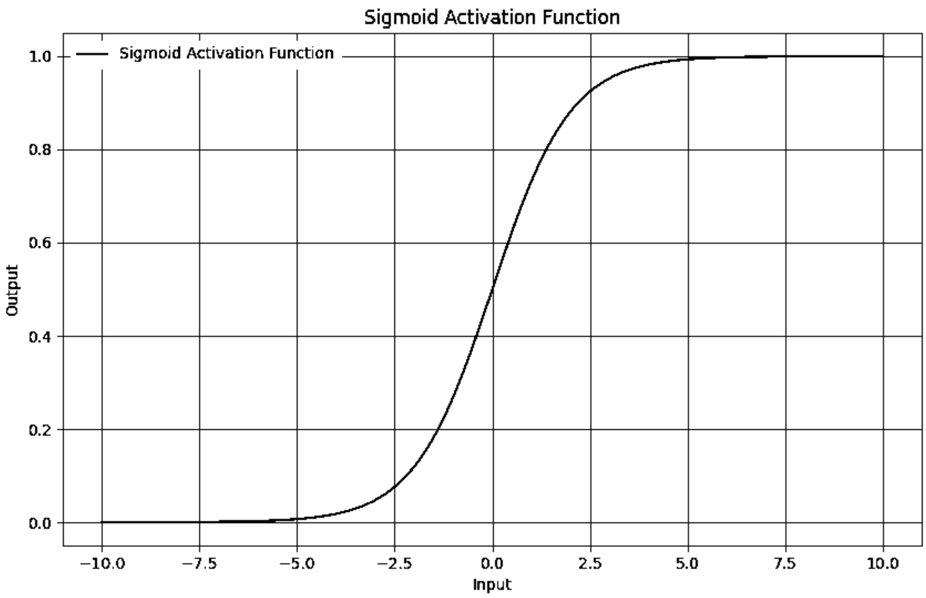


Figure 4

As you can see from the graph, the output values can only be values between 0 and 1. We are capable of getting values that are readable to us (prices of stocks at different times) thanks to the MinMaxScaler, by turning the prices into values contained within 0 and 1 and afterwards renormalising them after the model is done processing them. This is also the reason why, in Chapter 4, we decided to rescale the environmental score values to values between 1 and 100 to avoid the risk of having the network "die" and turn flat, resulting in no additional learning. This is because a value of 0 blocks any input, while a value of 1 allows any information to pass through the function. The MinMaxScaler package comes with premade functions, allowing for this normalisation to be done automatically.

The equations for general LSTM gates are the following (Thakur, 2018):

$$i_t = \sigma(w_i|h_{t-1}, x_t| + b_i)$$

$$f_t = \sigma\left(w_f|h_{t-1}, x_t| + b_f\right)$$

$$o_t = \sigma(w_o|h_{t-1}, x_t| + b_o)$$

$i_t \rightarrow input\ gate$

$f_t \rightarrow forget\ gate$

$o_t \rightarrow output\ gate$

$\sigma \rightarrow sigmoid\ activation\ function$

$w_x \rightarrow gate\ weight\ (x)$

$h_{t-1} \rightarrow privious\ block\ output$

$x_t \rightarrow input\ at\ current\ time$

$b_x \rightarrow biases\ (x)$

Since the model is designed to analyse simple linear stacks of layers, the sequential model was chosen alongside the Adam optimiser (adaptive moment estimation). This provided additional computational efficiency and allowed for adaptive learning in synergy with the long short-term memory, in contrast with the alternative of developing a simple recurring neural network (Kinsley & Kukeiela, 2020).

The sequential length was set to 64, and after experimenting with various increments (10, 32, 64, 100 and 134), this resulted in the best combination of model fit and the lowest rate of overfitting. In simpler terms, this means that the model will attempt to make a prediction for the next five trading hours based on the past 64. The model will attempt this prediction every five estimations and will learn from its own experiences how to better fit the model for future predictions by adjusting the weights.

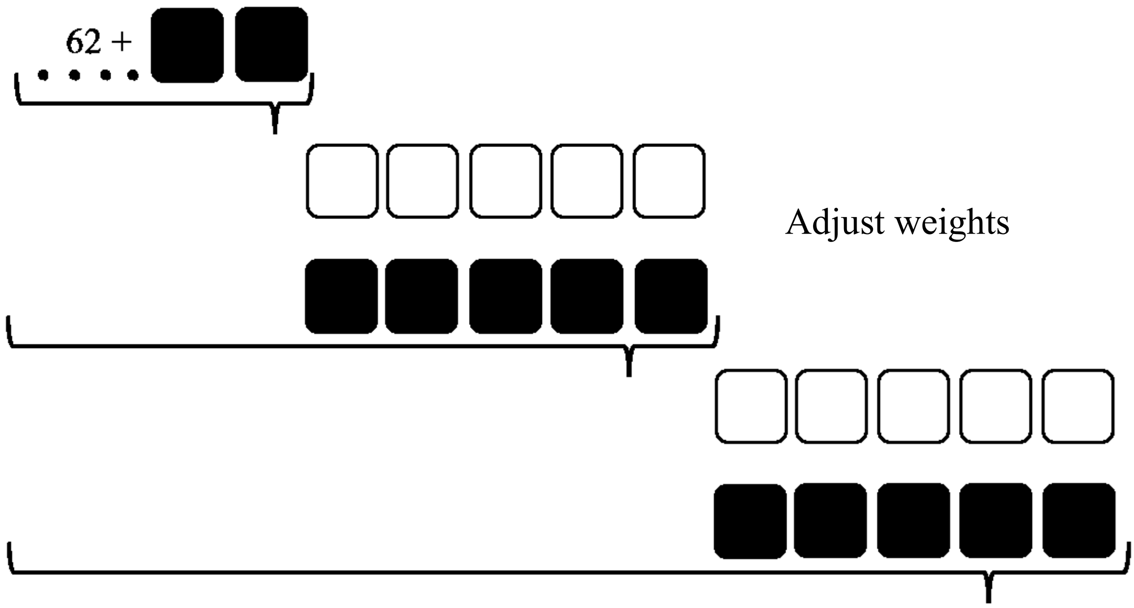


Figure 5

3.5 Measurement of error

In neural networks, measuring error is a crucial aspect of assessing a model's efficiency, as shown in the studies of Botunac et al. and Shahi et al. that formed the theoretical framework of this paper (Botunac et al., 2024) (Dahal et al.). A better-performing model has fewer errors, which is why the following section will analyse three types of error across the different configurations. These metrics are the mean absolute error, the root mean square error and the mean squared error. In this study, three different ways to measure error were employed: Mean Absolute Error (MAE), Mean squared Error (MSE) and Root Mean Squared Error (RMSE). These metrics serve as an excellent indicator of model performance, offering insight into the accuracy of the predictions and helping in comparing the different deep-learning models with the various parameter combinations. MAE is a measure of errors between a pair of observations expressing the same phenomenon, and it's calculated as the average absolute difference between the actual values and the predicted values. MAE provides a straightforward interpretation of how the predicted values performed compared to the actual values. Bellow, the general MAE formula is displayed (Chugh, 2024), where y_i represents the actual values and $\hat{y}_i$ represents the predicted values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MSE represents the average squared difference between the actual values and the predicted values. Therefore, we can understand that MSE is more sensitive to large errors compared to MAE due to the fact that the differences are squared, resulting in a much higher value, highlighting the discrepancies. By providing a clear penalty for larger errors this becomes ideal, especially in the context where we want as few deviations as possible in our RNN. Bellow, the MSE formula:

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

Finally, the last error metric employed in this study is the RMSE, which is the square root of the MSE, providing an error metric that is the same unit as the original data. It is commonly used in forecasting due to the balanced view it provides, incorporating both the mean and variance of the errors. Similar to MSE, high errors are penalised exponentially more, but the final values are more comparable since they are delivered on the original scale. Bellow the formula of RMSE is introduced:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

3.5 Trading simulation details

In order to answer H2 and H3, the model should be tested in a trading environment where its actual performance can be evaluated. For this purpose, a simple Python code was used to simulate trades on the available historical and predicted data. In order to understand the performance of the models, a brief explanation of the trading simulation's logic is crucial. First, for the purpose of the trading simulation, every model was given 1000 monetary units (USD, EUR, JPY are not relevant) and 100 leverage. The model was trained to place buy orders under certain conditions, and under opposite conditions, it placed sell orders. The model was preset to use a one to three risk-to-reward ratio, a standard setting in trading (Hayes) and was instructed to invest 1% of the account balance per trade. Every time a new prediction is made, the system considers the predicted values, and if the preset entry conditions are met, a new trade is made. Otherwise, the system will wait for a new opportunity. If the balance is sufficient, the simulation allows the model to place a trade and, every time interval after, checks if the SL or TP were hit. The stop loss was set at 1% of the entry price, and therefore, the take profit was set to 3%, maintaining that one to three risk-to-reward ratio.

The entry conditions were the following: if the average price of the five predictions is at least 3% higher than the last closing price, and the lowest prediction is not lower than 0.5% of the last closing price, the model will place a buy order, and the opposite for a sell order. By doing so, we ensure that the model only enters a trade when there is at least a significant move in the right direction and does not place contradicting trades.

The simulation keeps track of the number of lost and won trades, the maximum drawdown and the final balance. These metrics will be used in the results section to analyse the performance of the 40 LSTM models used in this study.

3.6 Limitations and areas of further research

While this study provides valuable insights into the effects of environmental metrics integration in AI trading models, several limitations should be considered, and certain areas remain open to further research. Addressing these limitations in further research can help validate the findings further and uncover new insights into the topic.

Further findings could potentially arise from the further investigation of in-depth industry-specific studies in the field of AI trading. It may be that the impact of environmental scores differs from industry to industry, and in certain industries, the public opinion's impact regarding environmental practices may be stronger than others, for example, in industries where consumer opinion and sentiment have a larger impact on the business. Industries like energy and manufacturing, which have a stronger impact on the environment than banking, for example, could have significantly different responses to environmental factors. This may also be due to regulatory differences among industries, and by focusing on specific industries, future research can identify industry-specific patterns and tailor AI models to better accommodate these differences.

Additionally, the use of different time frames might uncover new effects of ESG scores on LSTM models. Given the relatively static nature of environmental scores, it may be that the impact cannot be easily captured by NN models due to the lack of variance in the data for extended periods of time. The results of this study address mostly traders following day trading strategies and potentially swing traders with the help of deep learning methods. Further studies could explore the impact and relevance for short and long-term investors and how these scores affect the predictive power of models in those given conditions.

Additionally, other types of NN could be investigated, such as GRU or CNNs, with potential in pattern recognition of trends that show great potential in financial time series forecasting. Testing different types of NN could potentially uncover new interactions among the various factors that are fed to the model. Even the impact of combining ESG scores and traditional indicators, without price input, by generating buy and sell signals without predictive outputs could result in interesting findings. By exploring a wider range of ESG components, researchers could develop more robust solutions that better capture market trends and integrate a wider spectrum of environmental factors.

Chapter 4: Data

This study will make use of historical data of some of America's largest companies. The data used come from the period April 2020 – May 2024, covering four years of past financial data, time in which each company received five ESG scores. The cut-off point for the training and validation was October 2022, meaning that the data after this date was used in the testing and from that data, the results of this study were composed. Slightly more data was attributed to the training period in order to have a sample higher than the tested period. In Figure 6, we can see the complete list of companies used as subjects in this study:

Company name	Market Capital	Industry
Apple Inc (AAPL)	2.92 T	Consumer Electronics
Amazon.com Inc (AMZN)	1.81 T	Online Retailing
Bank of America Corporation (BAC)	261.71 B	Banking and Financial Services
Alphabet Inc (GOOGL)	1.86 T	Internet Services
Meta Platforms Inc (META)	1.19 T	Social Media
Netflix Inc (NFLX)	242.92 B	Streaming Entertainment
Nvidia Corporation (NVDA)	1.78 T	Semiconductor Technology
Tesla Inc (TSLA)	606.55 B	Electric Vehicle
Microsoft Corp (MSFT)	3.12 T	Computer Software
Pfizer Inc	155.62 B	Pharmaceutical

Figure 6: Source: Yahoo! Finance, date of data collection: 11.02.2024

The firm selection also provides a good overview of different industries, therefore helping to analyse the effectiveness of environmental factors in different fields.

These companies have their ESG scores listed on msci.com, alongside overviews of the company plans regarding carbon emissions, as shown below in the case of Pfizer Inc.:

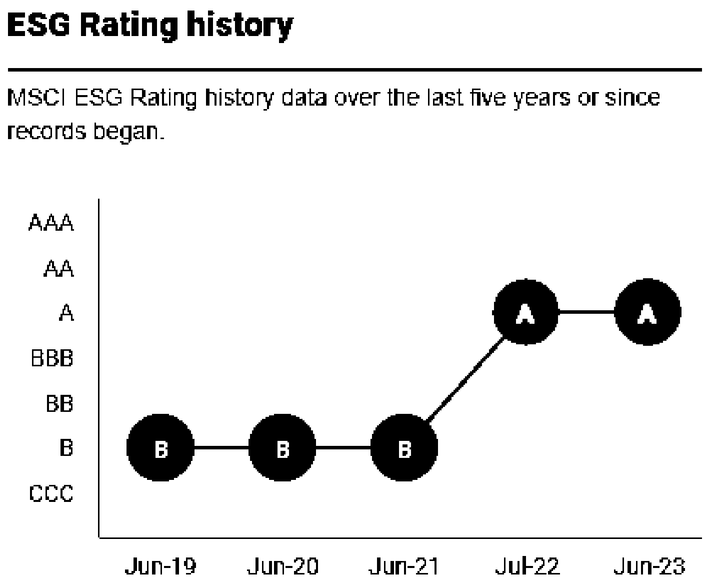


Figure 7: Source: msci.com, date of data collection: 15.01.2024

Decarbonization Target

Does PFIZER INC. have a decarbonization target?	YES
Does PFIZER INC. have a decarbonization target that is considered in the calculation of Implied Temperature Rise?	YES
Target Year	2040
Comprehensiveness % of company footprint covered by target	100.0%
Ambition Projected reduction per year to meet stated target**	-5.56% p.a.

Target data as of January 15, 2024*

Figure 8: Source: msci.com, date of data collection: 15.01.2024

The data used to train the model was partially acquired through a premium trading view account. It is intended for educational purposes only and is by no means intended for commercial use. All transactions were fictive, and the predictions were not used in actual trading; they were used only for the testing of the model. The companies were selected due to the data availability and because they have a combined market cap of nearly 14 trillion USD, a third of the S&P 500's total value, therefore being a representative sample. The rest of the data was taken from msci's website.

The values for ESG, CO_2 and temperature goals had to be normalised and, in some cases, had to be transformed from a string (ESG ratings, for example) into real numbers. ESG scores have been converted as shown in the table below:

ESG score	Assigned score for model training
CCC	1
B	16.6666
BB	33.3333
BBB	50
A	66.6666
AA	83.3333
AAA	100

Figure 9

The scale of 1 to 100 was chosen to the detriment of a -1 to 1 scale in order to maintain a certain level of alignment between prices and values. Prices never have negative values, and in our case, they are never 0. For this reason, the values of the ESG scores were transformed into a set of values that could easily blend with the others. Companies either have a CO_2 reduction plan or they don't, the reason for which, for the variable CO_2, a dichotomous variable was chosen, 1 or 100, again to maintain the same scale. The variable "T impact", representing the temperature impact the company's current behaviour is exerting on the environment, was treated in a similar manner to that of ESG but was given a linear structure similar to the carbon emission goals. In the table below, the exact figures are shown:

Temperature impact on environment (C°)	Assigned score for model training
3.2 ≤	1
1.9 – 3.1	25
1.5 – 2.0	75
≤ 1.5	100

Figure 10

It is to be noted that the only firm that had a rating of 1 in the metric "CO_2 emissions goals" was Tesla, while all other companies scored 100 since they had goals to reduce their carbon footprint.

This data will be fed to the model, and the results will be analysed in order to determine what impact environmental scores have on the training of the long-term-short-memory recurring neural network. The model will take the past five periods of the ESG rating and attempt to predict the prices for the next five hours. The data resulting from the four tests will be saved and categorised as "price" (referring to the fact that the data will only take in 4 inputs related to price), "ESG" (taking into consideration the previous four and adding ESG as the fifth), "CO2 target" (same as the previous, but replacing the ESG for CO_2 goals an adding the temperature impact to the calculation) and finally "combined" (for all seven metrics combined).

Chapter 5: Results

In this section, we will take a deep dive into the results of this study using an LSTM neural network for trading and attempt to determine the impact environmental factors have on the performance of this model. The results are organised into multiple sections, the first being the analysis of descriptive statistics. This section will provide a comprehensive summary of the predicted stock prices by analysing their mean, standard deviation, minimum values, maximum values and quartiles to better comprehend the distribution of the data.

This first section investigates the error measurements of the system, therefore addressing the first hypothesis, which is related to model performance. Section 5.3 will analyse the financial performance of the model, therefore addressing hypotheses two and three.

5.1 Descriptive statistics

Before getting into the detailed analysis of this paper's findings, let's first take a look at the descriptives of the data that's about to be analysed. All graphs generated from the analysed data, alongside graphs showing the performance of the predicted prices, can be found in the Annexes. When referring to a particular graph, it will be posted in the specific section for ease of reading. In the figure below, a brief overview of the results with a few metrics is provided that aims to give a better insight into the distribution and variability of the predicted stock prices. It is to be mentioned that each of the ten stocks compiled in the results below had approximately 3380 observations, all belonging to the secondary testing data set.

	MEAN	STD DEV	RANGE	SKEWNESS	OUTLIERS
PRICE	296.9	30.95	117.21	-0.26	0
ESG	296.28	30.4	116.22	-0.27	0.9
CO2 & T	295.62	30.58	116.64	-0.26	0.1
COMBINED	295.61	30.57	116.76	-0.29	1.1

Figure 11

Stock	**Configuration**	**Mean**	**Std Dev**	**Range**	**Skewness**	**Outliers**
AAPL	price	179.072	7.899	30.708	0.070	0
AAPL	ESG	178.215	7.681	30.096	0.059	0
AAPL	CO2 & T	179.162	8.006	31.133	0.077	0
AAPL	combined	178.161	7.737	29.945	0.069	0
AMZN	price	172.009	11.561	45.487	-0.690	0
AMZN	ESG	172.636	11.651	45.562	-0.689	0
AMZN	CO2 & T	172.029	11.429	46.048	-0.689	0
AMZN	combined	171.305	11.655	45.859	-0.689	0
BAC	price	35.501	2.001	7.906	0.022	0
BAC	ESG	35.446	2.076	8.132	0.018	0
BAC	CO2 & T	35.381	1.992	7.829	0.022	0
BAC	combined	35.607	2.076	8.144	0.015	0
GOOGL	price	148.769	10.486	44.256	0.568	0
GOOGL	ESG	149.659	10.577	44.381	0.574	0
GOOGL	CO2 & T	149.455	10.539	44.537	0.566	0
GOOGL	combined	149.660	10.627	44.453	0.570	0
META	price	455.915	51.155	184.664	-0.889	0
META	ESG	456.735	49.689	182.162	-0.892	0
META	CO2 & T	455.393	50.929	184.070	-0.892	0
META	combined	453.235	49.200	178.436	-0.890	0
MSFT	price	407.412	14.142	63.397	-0.964	0
MSFT	ESG	403.420	14.029	67.396	-1.005	9

MSFT	CO2 & T	407.573	14.138	63.410	-0.963	0
MSFT	combined	405.346	16.615	77.683	-1.146	10
NFLX	price	573.795	47.760	171.787	-0.941	0
NFLX	ESG	573.539	46.489	172.500	-0.946	0
NFLX	CO2 & T	575.003	48.029	173.180	-0.947	0
NFLX	combined	575.354	46.093	168.085	-0.966	0
NVDA	price	780.019	140.364	503.162	-0.714	0
NVDA	ESG	778.802	138.371	494.460	-0.718	0
NVDA	CO2 & T	766.143	136.817	495.382	-0.710	0
NVDA	combined	771.622	137.817	494.570	-0.711	0
PFE	price	27.563	0.972	4.678	-0.005	0
PFE	ESG	27.607	0.975	4.725	0.003	0
PFE	CO2 & T	27.555	1.002	4.760	-0.015	0
PFE	combined	27.502	0.983	4.798	-0.017	0
TSLA	price	188.920	23.125	116.091	0.908	0
TSLA	ESG	186.784	22.510	112.829	0.905	0
TSLA	CO2 & T	188.553	22.876	116.041	0.911	1
TSLA	combined	188.344	22.944	115.641	0.909	1

Figure 12

From the graphs above, we can determine that the central tendency of the mean values indicates that the values are fairly close to each other, indicating that the inclusion of ESG and CO_2 metrics does not significantly alter the price prediction by a significant amount. This is the potential result of the network assigning a small, close to zero or zero weight to certain parameters, blocking the information flow through the gate. The standard deviation indicates slight variations across different configurations, supporting the previous idea that prices are slightly deviating from the mean and that, in most cases, the various configurations did not significantly deviate from one another, resulting in comparable outputs. One exemption from this is Nvidia, which scored the highest standard deviation out of all stocks due to large price volatility. We will see in later sections that this is not the only case where Nvidia generated results that are different from those of the rest of the observations. To conclude, with the exception of a few cases, such as Microsoft combined, the introduction of environmental metrics does not appear to have a significant increase in variability.

Similarly to mean and standard deviation, the range of the values seems to remain rather constant. The general left skewness of the data indicates that the model had a preference for predicting lower prices. The large ranges observed belonged to high volatility stocks (NVDA, META), indicating broader price movements and their lack of significant variation from one configuration to another, which supports the fact that environmental factors do not drastically impact the model. When it comes to outliers, the number is generally minimal or null, but not across the board, with the exception of Microsoft, indicating that in its case, the implementation of environmental factors impacted the model, resulting in more extreme price predictions.

Bellow, two graphs are presented for a visual representation of the results. The rest of the graphs can be found in the annexe section.

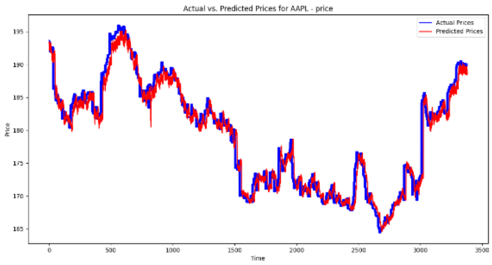
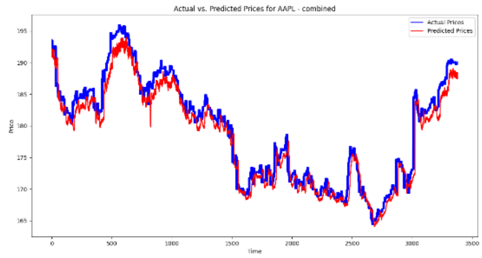



Figure 13: Side by side comparison of actual vs predicted prices in different model configurations – price on the left and combined on the right

The facts stated in the descriptive analysis can also be seen in the comparison of the side-by-side graph. The environmental factors graph on the right has more areas where the prediction is parallel with the actual prices, indicating a potential error in the loading of the sigmoid activation function or a potential bias negatively offsetting the graph (hence the left skewness of the data.

To briefly summarise the above, the similarity of the mean, std dev and range across the various configurations suggests that incorporating environmental metrics into the training of AI models, in our case LSTM models, does not significantly alter the variability of price predictions. The consistent skewness and the general lack of outliers suggest a robust model performance. Attached to the Annex section there are boxplots and histograms of all the tested models. From their review, we can determine that the data of the actuals and predicted follow a similar distribution, which is a good indication of the performance of the model.

5.2 Model Performance Evaluation

In this section, we will analyse the performance of the model, based on the error measurements. The results found in the different metrics align with one another, confirming the findings and indicating consistency in the model's evaluation procedure.

Row Labels	**Sum of RMSE**	**Sum of MSE**	**Sum of MAE**
combined	419.66	4,176.08	280.67
ESG	399.60	3,774.52	264.94
price	383.55	3,640.52	245.97
CO_2	408.07	4,534.17	270.15
Grand Total	**1,610.88**	**16,125.29**	**1,061.73**

Figure 14

5.2.1 Individual MAE analysis

The mean square error measures the average absolute difference between predicted and actual values. The best value was attained by the configuration "price", followed by ESG, CO_2 and lastly, combined, indicating that the more environmental metrics were added, the higher the error of the model. This can also be interpreted that while additional factors can provide more comprehensive predictions, they might also introduce more variability and potential error.

Figure 15 presents a heatmap of the individual error readings for the ten selected stocks, indicating consistency across most metrics. However, certain unique patterns will also repeat in the RSME and MSE readings. Overall, we can see that the price configuration tends to generate the lowest values, indicating that a simpler model may work better in our case.

Figure 15: Heatmap MAE values

Apple's most accurate predictions were for the price configuration, showing minimal error increases when additional metrics are included, indicating that environmental factors do not significantly impact the model's performance.

In Amazon's case, the CO_2 goals & temperature target metric configuration yielded the lower MAE, indicating that, in this case, the model performed better with the aid of this environmental factor. However, Amazon's prediction, when the model was exposed to ESG metrics, generated a slightly higher error than the rest. In the case of Bank of America, the results showed minimal error across the board, with neglectable slight variations indicating a minimal preference for simple price inputs. Similar to the Bank of America, Google's results were highly consistent across all configurations, with negligible MAE differences. In Microsoft's case, the predictions best performed on the price configuration, while ESG integration led to a noticeable increase in MAE. Netflix's predictions were relatively consistent, with the CO_2 and temperature configuration underperforming. Nvidia had the highest recorded average error from all observed companies. The carbon and temperature configuration underperformed compared to the rest, scoring the worst MAE score out of all configurations, 13,83. Pfizer had the best performance across all stocks, scoring some of the lowest error scores, with only the Bank of America coming close to its performance. In this case, all configurations scored a relatively similar score, indicating no significant impact caused by integrating environmental factors. Tesla scored fewer errors with the environmental metrics than the price one, and notably, Tesla was the only company that hasn't implemented carbon goals, all while having perfect ESG scores (AAA). In this case, we can see that the ESG had a positive impact.

5.2.2 MAE Aggregate

The aggregate results for the MAE values are the following:

Price: 30.75

ESG: 33.12

CO_2 & T: 33.76

Combined: 35.08

We can see that overall, the simple price configuration performed the best, and the more environmental scores we added, the higher the model error. Based on this metric, we would conclude that integrating environmental factors into an LSTM predictive model doesn't have a positive impact on the model's performance, therefore rejecting H1 and accepting the null hypothesis based on this metric alone. In the next sections, we will investigate whether or not we can confirm our findings with the other means of error measurement.

5.2.3 Individual RMSE analysis

RMSE stands for root mean squared error, and it measures the squared root of the average of squared differences between predicted and actual values (Chugh, 2024). Similarly to the MAE, a smaller value indicates a better-performing model. The main difference is that by squaring rather than subtracting the square root, we get a more comparable answer that is not so dependent on the original values, which may offer additional insight, considering that the stocks vary quite a bit in price from one another. The main findings of this section are that the price configuration still remains superior, with the combined one scoring the highest error, supporting the earlier findings.

Figure 16 illustrates a heatmap of the RMSE values, categorised per configuration and stock. Apple performed best under the price-only conditions with an RMSE score of 1.76. The combined configuration also yielded a relatively low error, suggesting that additional factors did not significantly deteriorate the performance. Amazon scored the lowest on carbon and temperature, suggesting that, in this particular case, environmental metrics had a positive impact on the model's training. On the other hand, the ESG model had the worst performance, suggesting that not all environmental factors were equally relevant in the given case. Bank of America Corporation displayed consistent results across all investigated metrics, suggesting that, in this particular case, the integration of environmental scores does not have a significant effect on the training of an LSTM model and did not significantly alter the deep learning process. Google, similarly to Amazon, displayed a better score on the carbon and temperature metric, with price performing the worst out of all metrics, indicating that in this particular case, environmental factors have a positive effect on the training of AI-based trading bots, the first fully positive response in the study. Meta, on the other hand, had the best performance on the price configuration, with all other metrics scoring relatively similar scores, slightly worse than the price metric, therefore contesting the first hypothesis. Microsoft had the best scores on price and carbon + T, with the other two configurations performing significantly worse. Netflix scored very similar RMSE scores on all four configurations, suggesting that, in this particular case, environmental factors do not significantly alter the performance of AI trading bots. Nvidia obtained once more the highest overall root mean squared error, with carbon and temperature scoring the worst, while the other metrics had comparable values, with price slightly outperforming the rest. PFE had the smallest measured error, confirming previous findings and suggesting that environmental factors do not significantly alter the performance of neural networks. It is to be noted, however, that even though the differences were relatively small, T and price scores were the best RMSE, with equally measured error. Finally, Tesla provided an interesting result, with all metrics scoring better than the price, suggesting once more that environmental scores have a positive impact on AI-powered trading bots. All these results may appear to be slightly contradicting one another since it seems that the results and the impact were, in most cases, company-specific, with very few having similar results as others. In order to obtain more insight into the data, the aggregate values will be calculated and analysed in the next section.

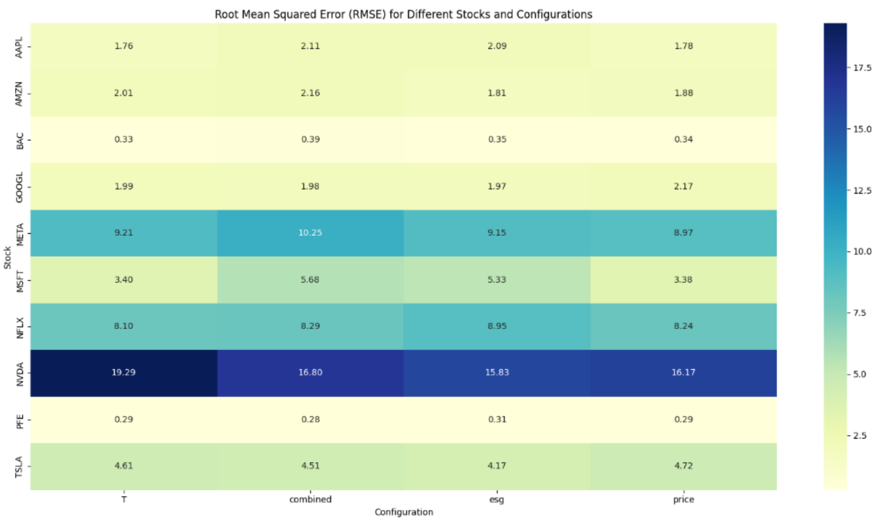


Figure 16: Heatmap RMSE values

5.2.4 RMSE Aggregate

The aggregate results for the RMSE values are the following:

Price: 46.16

ESG: 49.96

CO_2 & T: 50.41

Combined: 52.45

Despite the variations that appeared in the RMSE measurements compared to the MAE, the ranking of the best-performing configurations remained the same, confirming previous findings that ESG, CO_2 and temperature goals/scores do not have a positive effect on AI trading bots performance.

5.2.5 Individual MSE analysis

In the performance evaluation of neural network models, mean squared error offers a distinct perspective on the data compared to MAE or RMSE. Unlike MAE, which provides a pretty straightforward average, MSE squares the differences between predicted and actual values, offering more weight to larger errors helping us see the differentiation even better. Due to this particular characteristic, MSE is especially sensitive to outliers (Chugh, 2024). RMSE also considers the squared errors but returns the error value in the same units as the original data by taking the squared root of MSE. Therefore, by using MSE, we can highlight the differences found in the previous two sections about the error measurement of the model.

Similar to the previous sections, figure 17 displays an overview of the error measurement results.

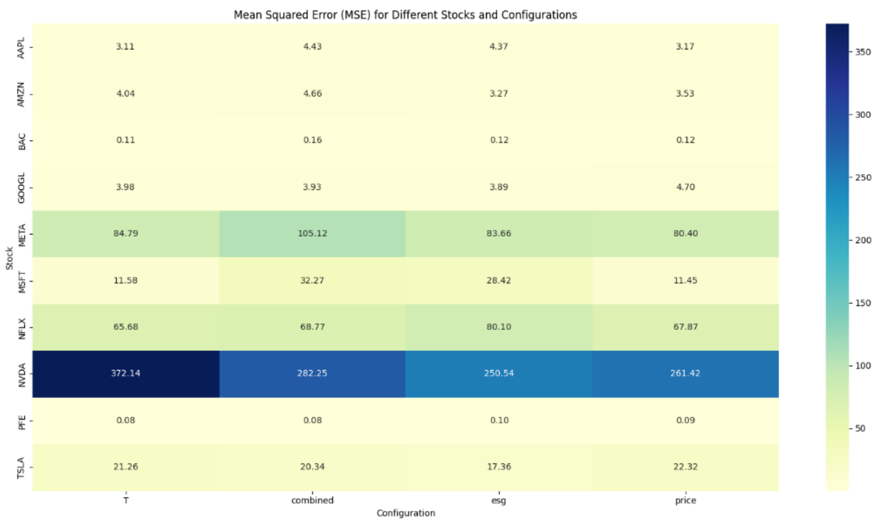


Figure 17: Heatmap MSE values

In Apple's case, the best performing configuration was the carbon goals, quickly followed by price and having combined and ESG performing slightly worse. This result was in line with the previous 2 measurements of error. Amazon's best score was ESG, with 3.27, outperforming the other models. The Bank of America maintained one of the overall best-performing models across the board, with very small errors in comparison to the rest. In the case of Google, all models performed similarly, indicating a lack of significant impact. Meta scored relatively high, indicating an ineffective model across the board, though fairly high error reading, particularly in the combined configuration, with price outperforming the rest. Microsoft's best performers were price, followed by T, without major differences between the two. The other two configurations, however, performed worse, indicating that their incorporation worsened the model's ability to learn. In the case of Netflix, all models performed similarly, with CO_2 & T slightly outperforming the rest. The only outlier was ESG, scoring significantly higher than the rest. Nvidia performed similarly to the previous error readings, having the highest margin of error. Once again, we can see that the ESG model performed slightly better than the other configurations. Pfizer maintained a low error across the board, while Tesla performed better under the ESG configuration.

5.2.6 MSE aggregate

The aggregate results for the MSE values are the following:

Price: 455.07

ESG: 471.83

Combined: 522.01

CO_2 & T: 566.77

The ranking changed slightly by having the combined score switch positions with the carbon and temperature configuration. However, this does not change the fact that the most efficient model only takes into consideration the price of the stock. Despite the slight variations in ranking, it is safe to assume that ESG scores, CO_2 goals and temperature impact have either harmed the model or displayed no significant impact. Based on these tests, we can conclude the following:

h1a: Integrating ESG scores into an LSTM predictive model has a positive impact on the model's performance.

We reject h1a, meaning that this study found out that integrating ESG scores into an LSTM predictive model does not positively impact the model performance.

h1b: Integrating CO_2 goals and temperature targets into an LSTM predictive model positively impacts the model's performance.

We reject h1b, meaning that this study found that integrating CO2 goals and temperature targets into an LSTM predictive model doesn't have a positive impact on the model performance.

h1c: Integrating ESG scores, CO_2 goals and temperature targets into an LSTM predictive model has a positive impact on the model performance.

We reject h1b, meaning that this study found out that integrating ESG scores, CO2 goals and temperature targets into an LSTM predictive model doesn't have a positive impact on the model performance.

Therefore:

We reject H1: Integrating environmental factors into an LSTM predictive model does not have a positive impact on the model performance.

5.3 Trading simulation results

In order to test hypotheses 2 and 3, the developed models had to be tested in a trading scenario. The results were positive, with all models resulting in profit after the trading period, with results ranging from 1021.31 (MSFT with full environmental configuration) to 5901.18 (TSLA with full environmental configuration). As it can be observed from this wide range, the implementation of environmental scores into neural trading algorithms had a wide range of results, indicating both potential for earning as well as risks. The results can be viewed in Figure 18.

Stock	Configuration	Final Balance	Total Wins	Total Losses	Maximum Drawdown
AAPL	price	1136.67	14	12	2.4%
AAPL	ESG	1131.44	13	10	2.0%
AAPL	CO2 & T	1127.14	12	12	2.4%
AAPL	combined	1127.94	13	11	2.0%
AMZN	price	1121.85	14	10	2.2%
AMZN	ESG	1156.05	15	10	2.0%
AMZN	CO2 & T	1115.92	21	18	3.9%
AMZN	combined	1064.30	13	14	3.7%
BAC	price	1215.73	14	11	4.4%
BAC	ESG	1249.17	16	8	2.2%
BAC	CO2 & T	1269.25	13	9	2.4%
BAC	combined	1172.19	16	12	4.4%
GOOGL	price	1465.72	30	10	2.6%
GOOGL	ESG	1424.69	21	9	2.2%
GOOGL	CO2 & T	1435.65	20	8	2.2%
GOOGL	combined	1420.80	20	10	2.4%
META	price	1659.69	44	41	3.9%

META	ESG	1703.86	60	56	5.7%
META	CO2 & T	1496.67	48	48	3.9%
META	combined	1518.53	66	70	7.2%
MSFT	price	1131.79	7	4	2.0%
MSFT	ESG	1070.93	28	31	5.7%
MSFT	CO2 & T	1085.25	6	8	2.0%
MSFT	combined	1021.31	37	56	12.5%
NFLX	price	1359.37	20	17	2.0%
NFLX	ESG	1361.13	52	54	8.2%
NFLX	CO2 & T	1321.11	17	17	2.2%
NFLX	combined	1434.16	27	18	2.6%
NVDA	price	4262.87	88	59	4.0%
NVDA	ESG	4398.78	87	52	3.9%
NVDA	CO2 & T	3568.06	120	112	8.0%
NVDA	combined	4670.80	104	74	7.1%
PFE	price	1232.97	18	14	3.9%
PFE	ESG	1222.65	19	16	3.9%
PFE	CO2 & T	1275.10	21	15	2.0%
PFE	combined	1393.65	22	10	2.0%
TSLA	price	4863.20	122	93	4.9%
TSLA	ESG	5388.25	97	50	3.0%
TSLA	CO2 & T	5568.93	121	81	4.9%
TSLA	combined	5901.18	126	84	4.9%

Figure 18: Trading simulation results

The model's success must also be attributed to the proper risk-to-reward ratio, as well as to the model itself. The maximum drawdown also varied extensively, ranging from 2% (MSFT with only price) to 12.5% (MSFT with full environmental configuration). In Figure 19, the data is reorganised for better visualisation.

	average drawdawn	dev	average losses	dev2	average wins	dev3	average final balance	dev4
price	3.2%	-16%	27.1	-14%	37.1	-9%	1944.98	-2%
ESG	3.9%	1%	29.6	-6%	40.8	1%	2010.69	1%
CO2 & T	3.4%	-12%	32.8	5%	39.9	-2%	1926.30	-3%
combined	4.9%	27%	35.9	15%	44.4	9%	2072.48	4%
average	3.8%		31.35		40.55		1,988.62	

Figure 19: Trading simulation summary results

From Figure 19, we can observe that the price metric outperformed the other configurations in terms of drawdown, indicating a more cautious performance. The general tendency to exaggerate trends could potentially come from an error related to an attributed bias, causing deviation of the predicted prices and making the slope go more in one direction. This could explain why we find the environmental factors at both ends of the spectrum, with the best and worst individual returns. This also explains why the drawdown for price was, on average, smaller. However, one interesting aspect is that the best average return was generated by the model using environmental scores, indicating that implementing these metrics into the training of neural network models. In the following paragraphs, a brief individual analysis of the performance of the various models will be presented.

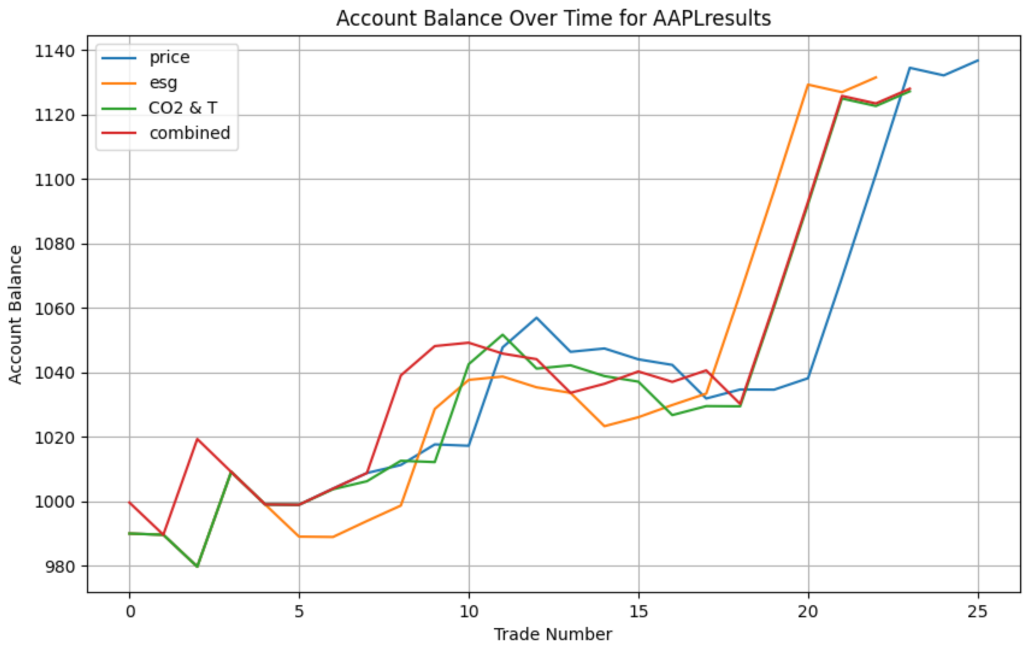


Figure 20

In Figure 20, we are presented with the financial performance of the models trained on Apple's stock data. In this case, we had some of the lowest amounts of trades placed, with only 22 to 25 trades being placed. This indicates that the required conditions were not met often, and the model was not able to find many opportunities for growth of 3%. However, despite this, the model managed to generate well over 11% returns over the tested period on all given configurations. All models performed in a comparable manner without any significant differences.

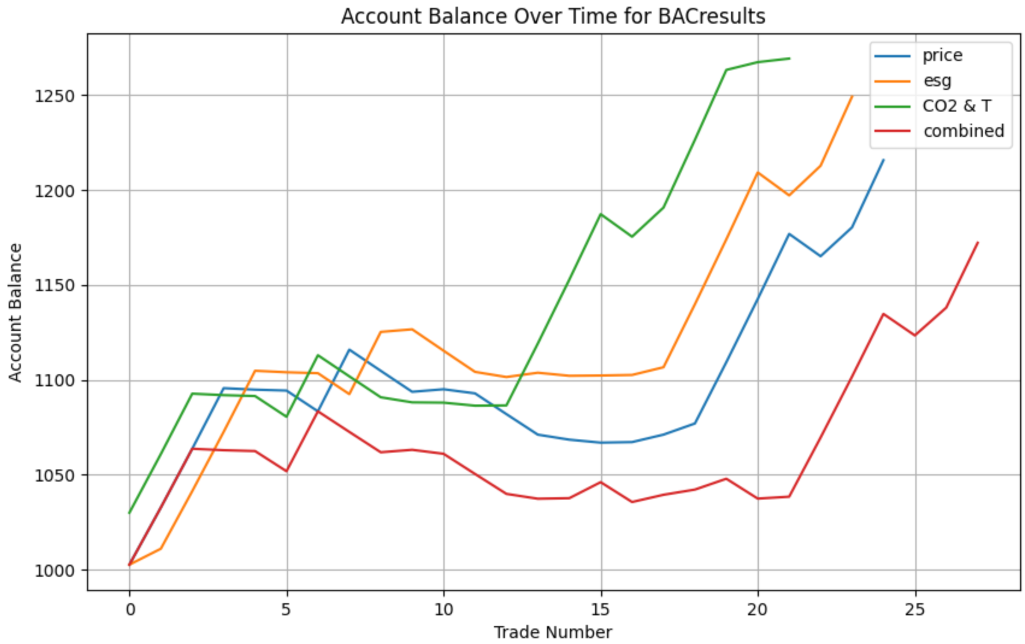


Figure 21

In Figure 21, we can see the performance of the models trained on the Bank of America Corporation, with somewhat varying results. In this particular case, carbon reduction goals & temperature impact had the best results, followed by ESG, indicating a strong impact of environmental factors.

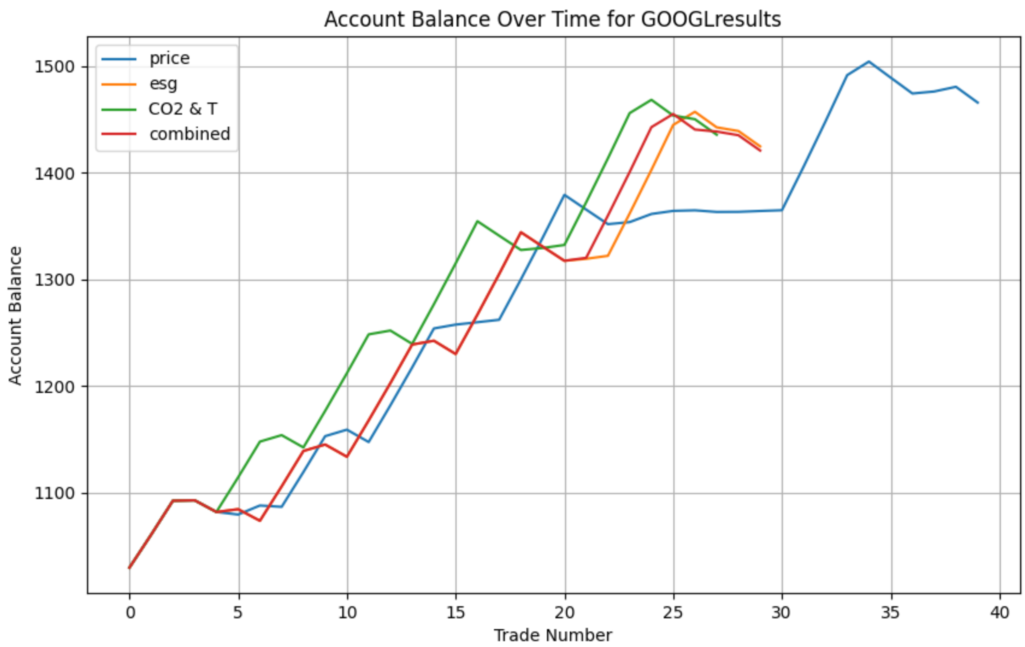


Figure 22

In the case of Google (figure 22), the traditional trading model, making use of only price, not only outperformed the other configurations but also placed nearly 30% more trades than the rest, indicating a high level of performance, potentially attributed to the very low error recorded by the model.

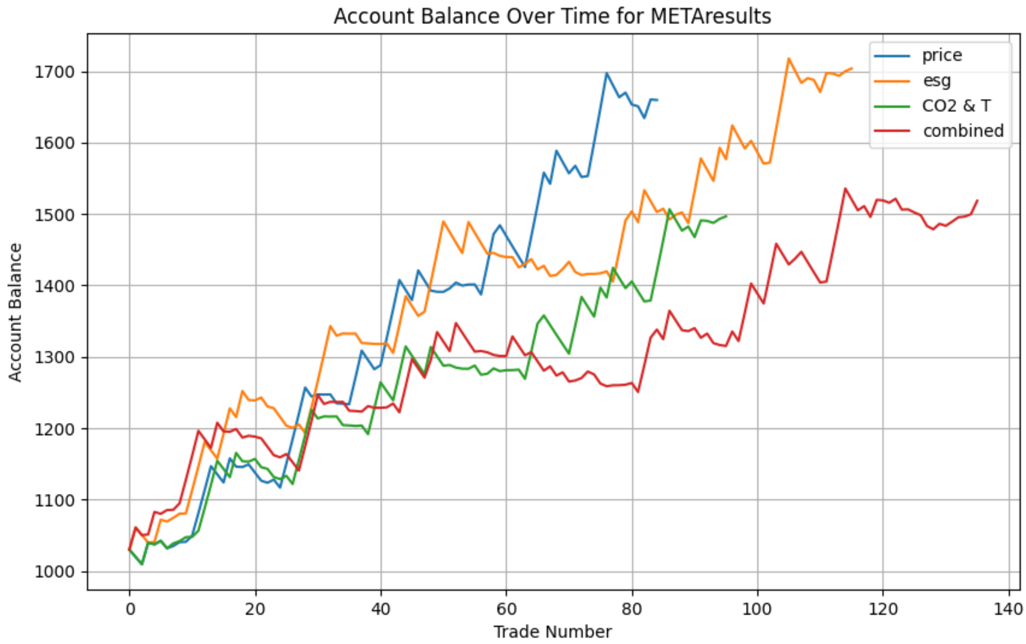


Figure 21

The stocks trained on Meta also had a very good return, ranging from 15% to 17%. Despite placing the most trades, the combined configuration underperformed compared to the others.

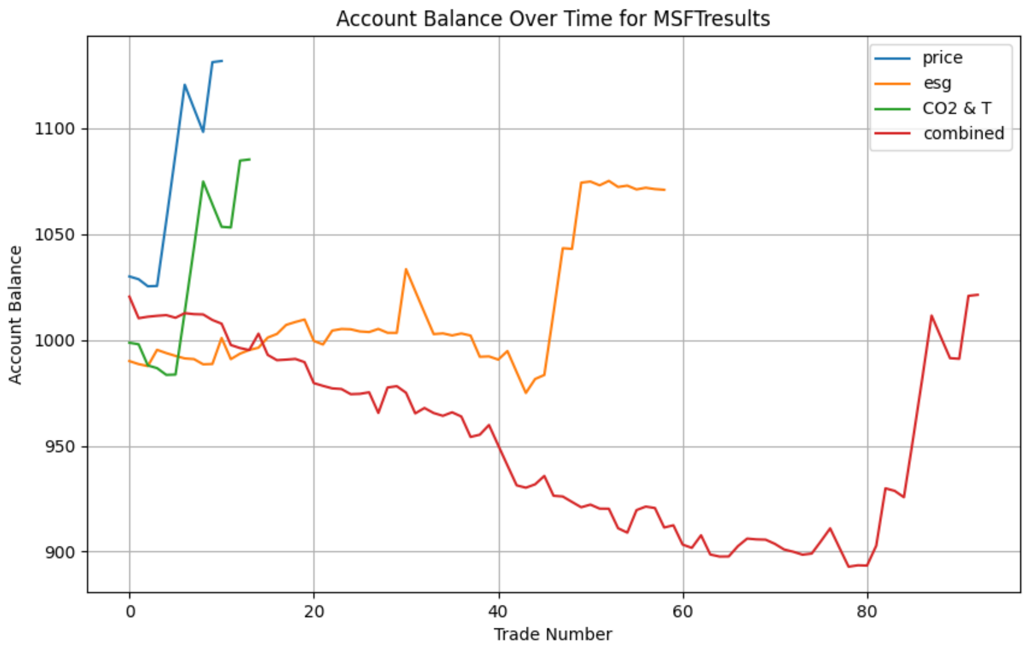


Figure 22

Perhaps some of the most unique results are the ones in the case of MSFT were the price metric significantly outperformed the environmental score metric. Despite placing less than 20% of the amounts of trades, it managed to attain the lowest drawback (out of all models) and attained the highest return of the MSFT trained models).

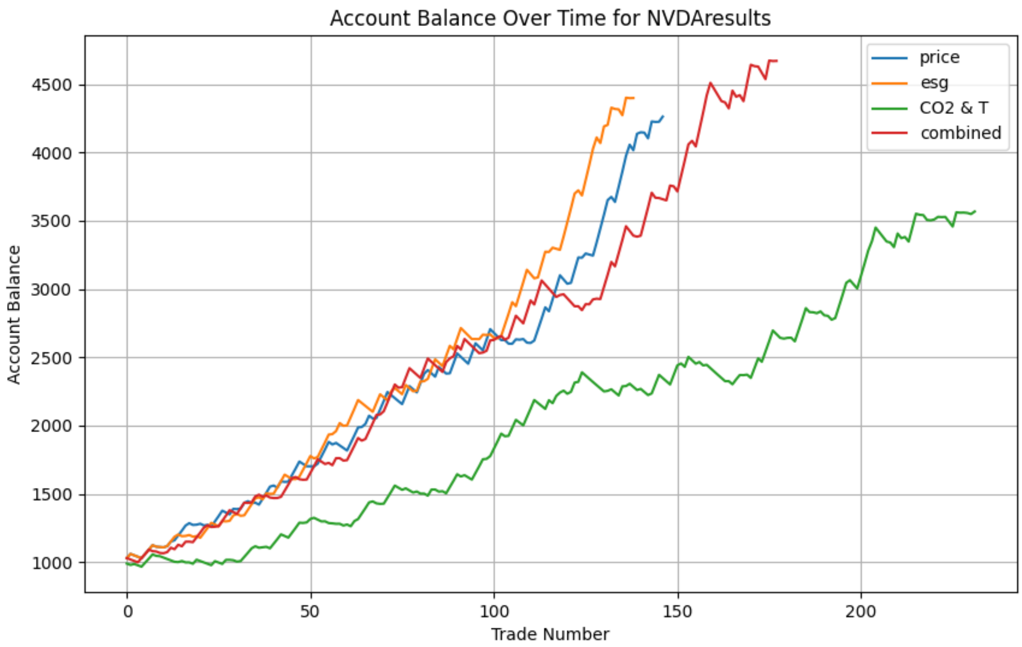


Figure 23

In the case of Nvidia, the combined metric performed the best, attaining a more than 45% return, nearly 5% more than the price configuration, indicating a strong correlation between the implementation of environmental scores and the performance of the model.

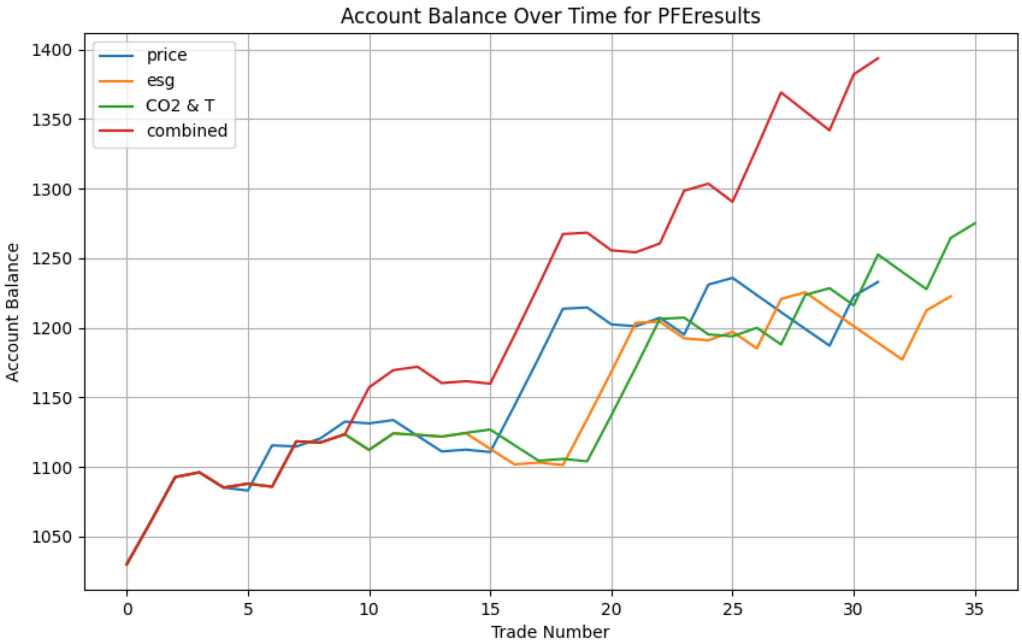


Figure 24

Pfizer, similarly to Nvidia, performed the best with the integration of environmental factors, with no significant differences between price, ESG, and carbon metrics.

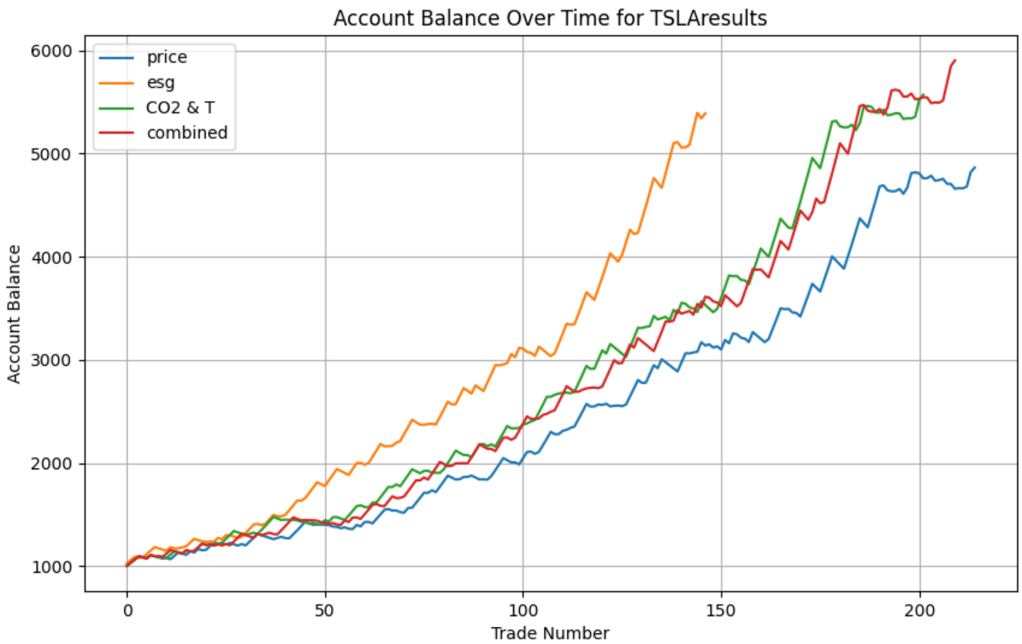


Figure 25

Finally, Tesla had the best results, scoring nearly 60% returns on the combined configuration. In this particular case we can see a great performance once more by the combined configuration. This further supports the fact that environmental factors in can have a positive impact on the trading outcome.

With these factors in mind, we can now accept/reject hypothesis two and three:

We reject h2a, meaning that the integration of ESG scores into an LSTM predictive models does not reduce maximal drawdown in trading.

We reject h2b, meaning that the integration of CO_2 goals and temperature targets into an LSTM predictive models does not reduce maximal drawdown in trading.

We reject h2c, meaning that the integration of ESG scores, CO_2 goals and temperature targets into an LSTM predictive models does not reduce maximal drawdown in trading.

We accept h3a, meaning that the integration of ESG scores into an LSTM predictive model increases returns.

We reject h3b, meaning that the integration of CO_2 goals and temperature targets into an LSTM predictive model does not increase returns.

We accept h3c, meaning that the integration of ESG scores, CO_2 goals and temperature targets into an LSTM predictive model increases returns.

Chapter 6: Conclusions

To summarise, the study explored the integration of environmental, social and governance scores, alongside CO_2 reduction goals and Temperature impact, into long-term memory neural networks and their impact on trading performance. The findings of this research indicate an intricate interaction between the incorporation of environmental metrics and the performance of LSTM models. Contrary to expectations, their impact on the accuracy of the model was not significant in most cases, while in some, it even resulted in a higher measured error, suggesting that these factors are not suitable for such models in the context of prediction accuracy. This suggests that the predictive power of these models is sensitive to the linear nature of these scores, slowing down the learning process. From a practical perspective, the findings highlight the challenges that arise from the embedding of environmental metrics. The fact that the results varied drastically from one stock to another could have various interpretations. It may be that the impact of these factors is industry-specific, with some being more affected than others. On the other hand, it may indicate that the actual scores and ratings may have an impact, as we saw in the case of Tesla, which attained the best results. This could either mean that a more complex model would be required to better capture and understand when these scores take effect and when they do not or that the increase is also dependent heavily on the stock movements. After all, Tesla experienced some of the largest price movements out of all analysed companies. The way the model is set up, it requires large movements in order to be effective, as the algorithm won't use the AI's predictions unless they signal strong movements. Despite these mixed results, the study highlights the growing importance of understanding the impact of environmental scores in financial decision-making. With more and more regulations and pressure from society to integrate environmentally friendly practices, it may be that these factors will continue to increase their impact over time. The study found that despite the fact that in certain scenarios, neural networks struggle to make use of environmental scores and adequately leverage them to their advantage, robust trading strategies can profitably incorporate them into their practice.

Furthermore, this paper highlights several limitations and areas for further investigation. The static nature of environmental scores poses challenges when it comes to their incorporation into deep learning models analysing complex and highly volatile data, as the two are hard to match. Future studies could explore their interaction with different types of neural networks and analyse their impact in order to bring even more light to the topic. Additionally, industry-specific studies may uncover hidden relationships between environmental practices and financial performance that are not apparent to humans but which patterns could be identified by deep learning.

In conclusion, while the integration of environmental metrics into AI trading models comes with significant challenges, it also shows high potential for enhancing sustainable investment practices. This study contributes to the literature on sustainable finance and AI by providing empirical evidence on the relationship between environmental factors and the performance of predictive trading models. By addressing the limitations as mentioned earlier and continuing to study AI models, future research can help unlock the true potential of integrating sustainability into financial decision-making, ultimately contributing to more responsible investment strategies.

References

Aminimehr, A., Raoofi, A., Aminimehr, A. et al. A Comprehensive Study of Market Prediction from Efficient Market Hypothesis up to Late Intelligent Market Prediction Approaches. Comput Econ 60, 781–815 (2022). https://doi.org/10.1007/s10614-022-10283-1

Apiday. (2023, December 29). MSCI ESG Rating: How are ESG ratings determined? MSCI ESG rating: How are ESG ratings determined? https://www.apiday.com/blog-posts/what-are-the-msci-ESG-ratings#:~:text=MSCI%20ESG%20Ratings%20calculate%20each,or%20reliance%20on%20government%20contracts.

Babatunde, S. O., Odejide, O. A., Edunjobi, T. E., & Ogundipe, D. O. (2024). The role of AI in marketing personalization: A theoretical exploration of consumer engagement strategies. International Journal of Management & Entrepreneurship Research, 6(3), 936-949.

Bag, D., & Mohanty, S. (2021). Impact of ESG factors on Emerging Market Stock Returns. The Journal of Impact and ESG Investing, 2(2), 138–147. https://doi.org/10.3905/jESG.2021.1.029

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG Ratings. Review of Finance, 26(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Brandl, B., & Hornuf, L. (2020). Where did FinTechs come from, and where do they go? The transformation of the financial industry in Germany after digitalization. Frontiers in Artificial Intelligence, 3, 8.

Botunac, I., Bosna, J., & Matetić, M. (2024). Optimization of Traditional Stock Market Strategies Using the LSTM Hybrid Approach. Information, 15(3), 136.

Brenner, L., & Meyll, T. (2020). Robo-advisors: A substitute for human financial advice? Journal of Behavioral and Experimental Finance, 25, 100275. https://doi.org/10.1016/j.jbef.2020.100275

Buchholz, K., & Richter, F. (2019, December 9). Infographic: Computers manage more stock than humans do. Statista Daily Data. https://www.statista.com/chart/20245/share-of-computerized-and-human-trading-in-us-equities/

Cao, S., Jiang, W., Wang, J. L., & Yang, B. (2021). From man vs. machine to man+ machine: The art and AI of stock analyses. Columbia Business School Research Paper.

Canadell, J. G., Monteiro, P. M., Costa, M. H., Cotrim da Cunha, L., Cox, P. M., Eliseev, A. V., ... & Zickfeld, K. (2023). Intergovernmental Panel on Climate Change (IPCC). Global carbon and other biogeochemical cycles and feedbacks. In Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change (pp. 673-816). Cambridge University Press.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. https://doi.org/ https://doi.org/10.48550/arXiv.1406.1078

Cetinic, E., & She, J. (2022). Understanding and creating art with AI: Review and outlook. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 18(2), 1-22.

Che, C., Huang, Z., Li, C., Zheng, H., & Tian, X. (2024). Integrating generative ai into financial market prediction for improved decision making. arXiv preprint arXiv:2404.03523

Commission proposes New Carbon Border Adjustment Mechanism and revision of the Energy Taxation Directive, as part of the EU Green deal. Taxation and Customs Union. (2021, July 14). https://taxation-customs.ec.europa.eu/news/commission-proposes-new-carbon-border-adjustment-mechanism-and-revision-energy-taxation-directive-2021-07-14_en

Chugh, A. (2024, January 18). Mae, MSE, RMSE, coefficient of determination, adjusted R squared-which metric is better?. Medium. https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e

Dahal, K. R., Pokhrel, N. R., Gaire, S., Mahatara, S., Joshi, R. P., Gupta, A., Banjade, H. R., & Joshi, J. (2023). A comparative study on effect of news sentiment on stock price prediction with Deep Learning Architecture. PLOS ONE, 18(4). https://doi.org/10.1371/journal.pone.0284695

DeepAI. (2019, May 17). Gated Neural Network. https://deepai.org/machine-learning-glossary-and-terms/gated-neural-network#:~:text=Input%20Gate%3A%20Determines%20how%20much,next%20layer%20or%20time%20step.

Deloitte. (2022, March 15). ESG Ratings: Do they add value? how to get prepared? https://www2.deloitte.com/ce/en/pages/about-deloitte/articles/ESG-ratings-do-they-add-value.html

Dot funding programs and climate change. U.S. Department of Transportation. (n.d.). https://www.transportation.gov/priorities/climate-and-sustainability/dot-funding-programs-and-climate-change#:~:text=Carbon%20Reduction%20Program%20%2D%20provides%20%246.4,alternative%20fuels%2C%20and%20port%20electrification.

Downey, L. (2024, February 11). Efficient market hypothesis (EMH): Definition and critique. Investopedia. https://www.investopedia.com/terms/e/efficientmarkethypothesis.asp

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management science, 60(11), 2835-2857.

Environmental Protection Agency. (2024). EPA. https://www.epa.gov/clean-air-act-overview

ESG and performance. MSCI. (n.d.). https://www.msci.com/ESG-101-what-is-ESG/ESG-and-performance

Finax. (n.d.). https://www.finax.eu/

Fit for 55 - the EU's plan for a green transition - consilium. (2023). https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55/

Gokhale, N., Gajjaria, A., Kaye, R., & Kuder, D. (2019). AI leaders in Financial Services. Deloitte Insights. https://www2.deloitte.com/us/en/insights/industry/financial-services/artificial-intelligence-ai-financial-services-frontrunners.html

Hamad, R. (2023, December 11). What is LSTM? introduction to long short-term memory. Medium. https://medium.com/@rebeen.jaff/what-is-lstm-introduction-to-long-short-term-memory-66bd3855b9ce

Hayes, A. (n.d.). Risk/reward ratio: What it is, how stock investors use it. Investopedia. https://www.investopedia.com/terms/r/riskrewardratio.asp

Iea. (n.d.). Executive summary – energy efficiency 2022 – analysis. IEA. https://www.iea.org/reports/energy-efficiency-2022/executive-summary

Inspired Economist. (2023, November 27). Efficiency market hypothesis (EMH): Understanding the pivotal theory in Finance. https://inspiredeconomist.com/articles/efficiency-market-hypothesis-emh/

Jiang, H. H., Brown, L., Cheng, J., Khan, M., Gupta, A., Workman, D., ... & Gebru, T. (2023, August). AI Art and its Impact on Artists. In Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (pp. 363-374).

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Julien, Kaffeetrinker, Domi99, Baeyens, T., & xx2. (2009, November 26). FANN2MQL Neural Network tutorial. MQL5. https://www.mql5.com/en/articles/1574

Kinsley, H., & Kukeiela, D. (2020). Neural Netowrks from Scratch in Python.

Königstorfer, F., & Thalmann, S. (2020). Applications of Artificial Intelligence in commercial banks– A research agenda for behavioral finance. Journal of behavioral and experimental finance, 27, 100352.

Kulikov, A., Alabed Alkader, N., Panaedova, G., Ogorodnikov, A., & Rebeka, E. (2023). Modelling Optimal Capital Structure in Gas and Oil Sector by Applying Simulation Theory and Programming Language of Python (Qatar Gas Transport Company). Energies, 16(10), 4067.

Lee, H., Kim, J. H., & Jung, H. S. (2024). Deep-learning-based stock market prediction incorporating ESG sentiment and technical indicators. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-61106-2

Levin, J. (2001). Information and the market for Lemons. The RAND Journal of Economics, 32(4), 657. https://doi.org/10.2307/2696386

Lima de Castro, P. A. (2021). mt5se: An Open Source Framework for Building Autonomous Trading Robots. Instituto Tecnologico de Aeron ´ Autica. https://doi.org/ https://doi.org/10.48550/arXiv.2101.08169

Lynch, S. (2023, April 3). 2023 state of AI in 14 charts. Stanford HAI. https://hai.stanford.edu/news/2023-state-ai-14-charts

Machado, M. R., & Karray, S. (2022). Assessing credit risk of commercial customers using hybrid machine learning algorithms. Expert Systems with Applications, 200, 116889. https://doi.org/10.1016/j.eswa.2022.116889

Masís, S. (2023). Interpretable Machine Learning with Python: Build explainable, fair, and robust high-performance models with hands-on, real-world examples. Packt Publishing Ltd.

McCollum, D. L., Zhou, W., Bertram, C., De Boer, H. S., Bosetti, V., Busch, S., ... & Riahi, K. (2018). Energy investment needs for fulfilling the Paris Agreement and achieving the Sustainable Development Goals. Nature Energy, 3(7), 589-599

Menyah, K. (2013). Stewardship Theory. In: Idowu, S.O., Capaldi, N., Zu, L., Gupta, A.D. (eds) Encyclopedia of Corporate Social Responsibility. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-28036-8_107

Montanaro, A. (2016). Quantum Algorithms: An overview. Npj Quantum Information, 2(1). https://doi.org/10.1038/npjqi.2015.23

MSCI ESG ratings methodology: Carbon Emissions Key ... (n.d.). https://www.msci.com/documents/1296102/34424357/MSCI+ESG+Ratings+Methodology+-+Carbon+Emissions+Key+Issue.pdf/bfc8304f-bf60-d4ad-07e4-9f72d2892f79?t=1666182592995

N., Floros, C. ESG performance, herding behaviour and stock market returns evidence from Europe. Oper Res Int J 23, 3 (2023). https://doi.org/10.1007/s12351-023-00745-1

Orús, R., Mugel, S., & Lizaso, E. (2019). Quantum Computing for Finance: Overview and Prospects. Reviews in Physics, 4, 100028. https://doi.org/10.1016/j.revip.2019.100028

Ostbirk, K. H. (2022). Environmental Social Go on mental Social Governance: Unco enhance: Uncovering Str erring Strengths, lengths, Weaknesses, and Misconceptions in ESG Disclosure and Rating. Senior Project Submitted to The Division of Social Studies of Bard College. https://doi.org/https://digitalcommons.bard.edu/senproj_s2022

Palmer, B. (n.d.). Institutional investors vs. retail investors: What's the difference? Investopedia. https://www.investopedia.com/ask/answers/06/institutionalinvestor.asp

Peaks. (n.d.). https://www.peaks.com/

Porter, M. E., Serafeim, G., & Kramer, M. (2019). Where ESG Fail. Institutional Investor. https://www.institutionalinvestor.com/article/2bswdin8nvg922puxdzwg/opinion/where-ESG-fails

Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. Journal of Marketing, 85(1), 131-151.

Rosenberg, G., Hanghnegahdar, P., Goddard, P., Carr, P., Wu, K., & Marcos López de Prado, M. (2015). Solving the Optimal Trading Trajectory Problem Using a Quantum Annealer. Computational Finance, 10(6). https://doi.org/ arXiv:1508.06182

Ruan, L., & Liu, H. (2021). Environmental, social, governance activities and firm performance: Evidence from China. Sustainability, 13(2), 767. https://doi.org/10.3390/su13020767

Shahi, T. B., Shrestha, A., Neupane, A., & Guo, W. (2020). Stock price forecasting with deep learning: A comparative study. Mathematics, 8(9), 1441.

Smigel, L. (2023, October 14). 79+ amazing algorithmic trading statistics (2024). Analysing Alpha. https://analyzingalpha.com/algorithmic-trading-statistics

Shepherd, D. (2021, October 13). BCG says that over 90% of firms don't measure emissions correctly. Bloomberg.com. https://www.bloomberg.com/news/articles/2021-10-13/over-90-of-firms-aren-t-measuring-emissions-correctly-bcg-says

Syamala, S. R., & Wadhwa, K. (2020). Trading performance and market efficiency: Evidence from algorithmic trading. Research in International Business and Finance, 54, 101283. https://doi.org/10.1016/j.ribaf.2020.101283

Stanford Institute for Human-Centred Artificial Intelligence (HAI). (2023). 2023 State of AI in 14 Charts. Retrieved from https://hai.stanford.edu/research/ai-index-2023

Thakar, C. (2023, July 7). The evolution of trading: Barter System to algo trading. Quantitative Finance & Algo Trading Blog by QuantInsti. https://blog.quantinsti.com/evolution-trading-barter-system-algo-trading/

Thakur, D. (2018, July 6). LSTM and its equations. Medium. https://medium.com/@divyanshu132/lstm-and-its-equations-5ee9246d04af

Urquhart, A., & McGroarty, F. (2016). Are stock markets efficient? Evidence of the adaptive market hypothesis. International Review of Financial Analysis, 47, 39–49. https://doi.org/10.1016/j.irfa.2016.06.011

Wallace, J. S. (2003). Value maximisation and stakeholder theory: Compatible or not? Journal of Applied Corporate Finance, 15(3), 120–127. https://doi.org/10.1111/j.1745-6622.2003.tb00466.x

Weng, B., Lu, L., Wang, X., Megahed, F. M., & Martinez, W. (2018). Predicting short-term stock prices using ensemble methods and online data sources. Expert Systems with Applications, 112, 258-273.

World Bank. (2021). ACCELERATING THE PRODUCTIVE USE OF ELECTRICITY

Yadav, M. (2019). Impact of Artificial Intelligence on Trading in Financial Markets. International Journal of Advanced Scientific Research and Management, 4(7), 32–36.

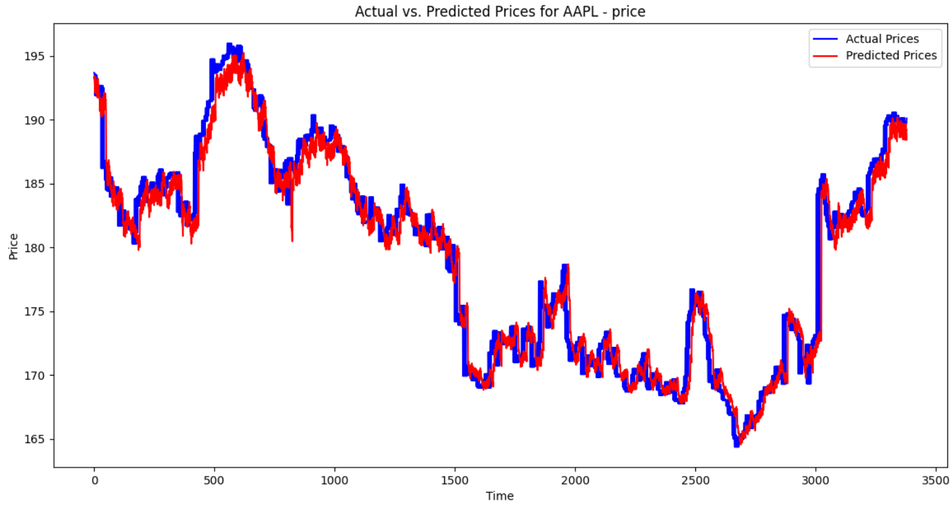
Zairis, G., Liargovas, P., & Apostolopoulos, N. (2024). Sustainable finance and ESG Importance: A systematic literature review and research agenda. Sustainability, 16(7), 2878. https://doi.org/10.3390/su16072878

Zhang, Z., Zohren, S., & Roberts, S. (2020). Deep reinforcement learning for trading. The Journal of Financial Data Science, 2(2), 25–40. https://doi.org/10.3905/jfds.2020.1.030

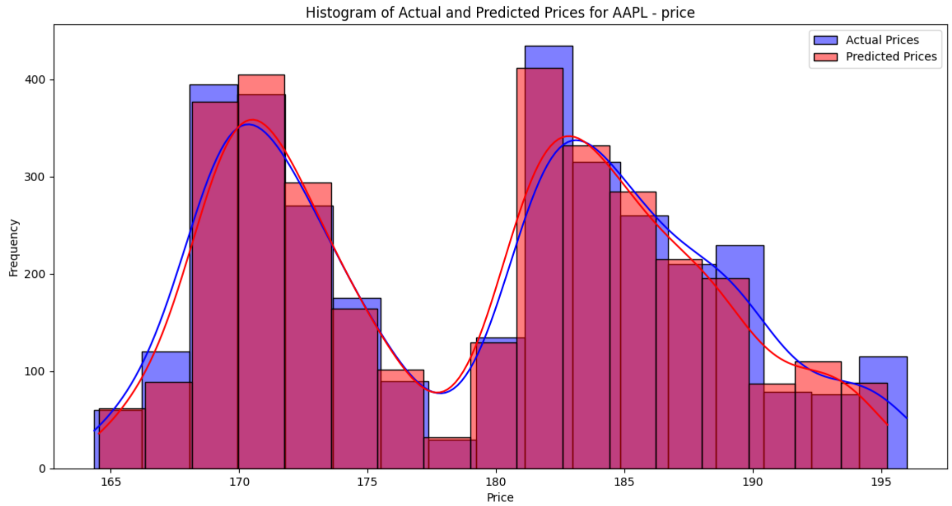
Appendix

Code of model:

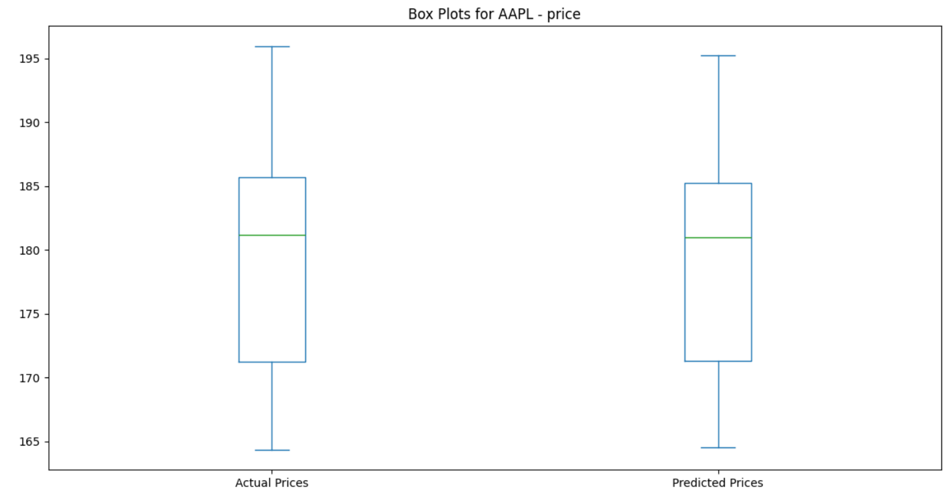
AAPL - price (Actual vs Predicted)



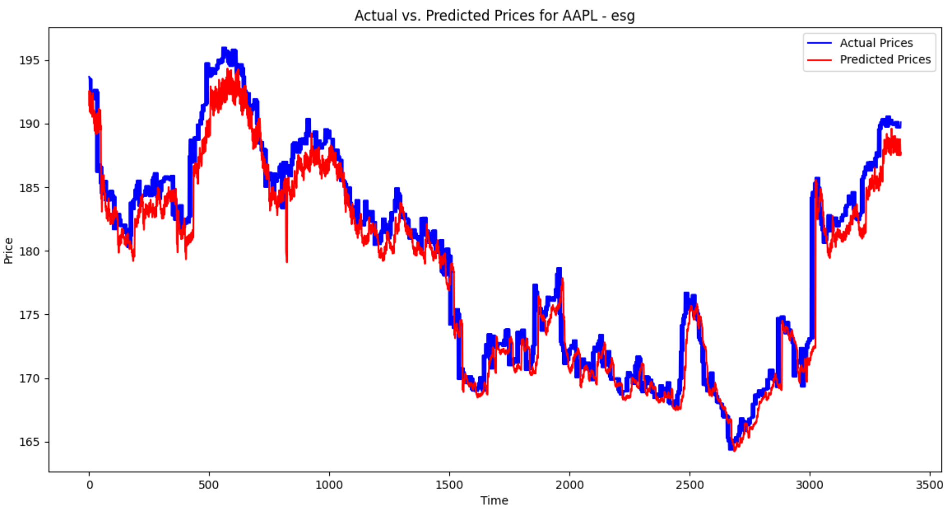
AAPL - price (Histogram)



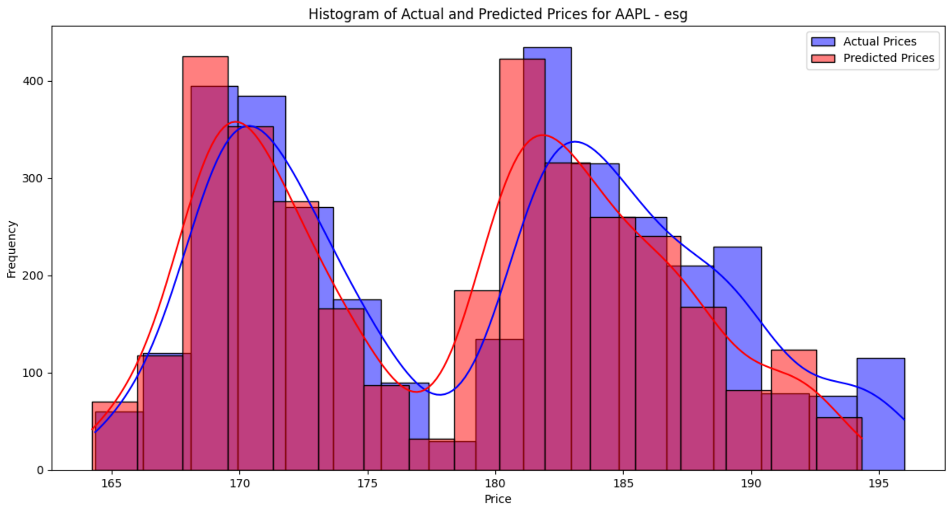
AAPL - price (Boxplot)



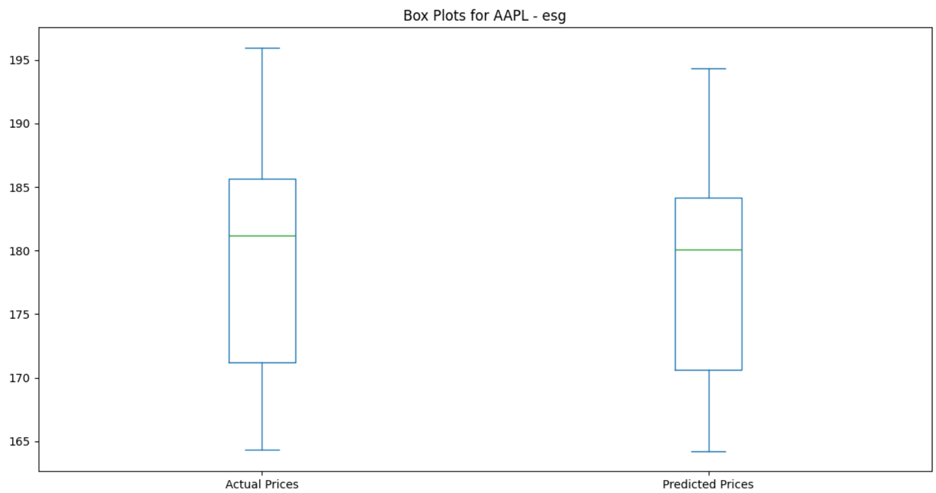
AAPL - ESG (Actual vs Predicted)



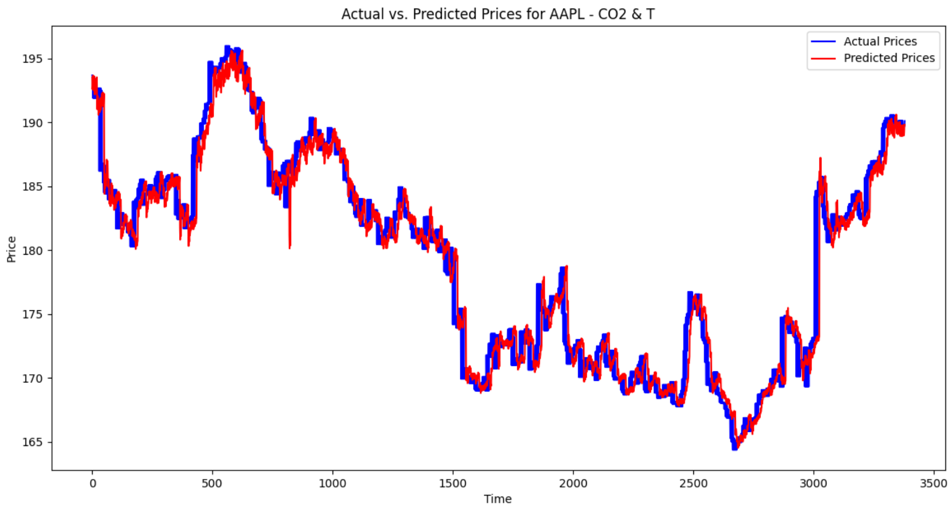
AAPL - ESG (Histogram)



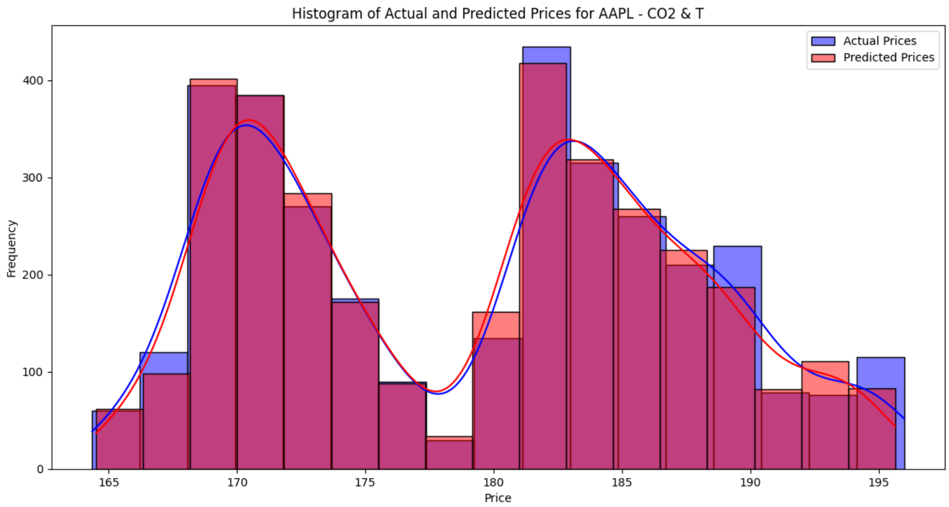
AAPL - ESG (Boxplot)



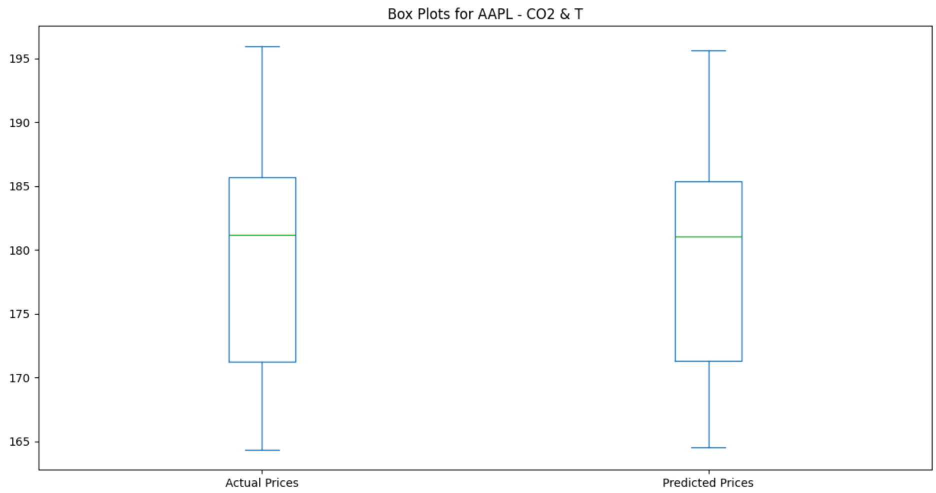
AAPL - CO2 & T (Actual vs Predicted)



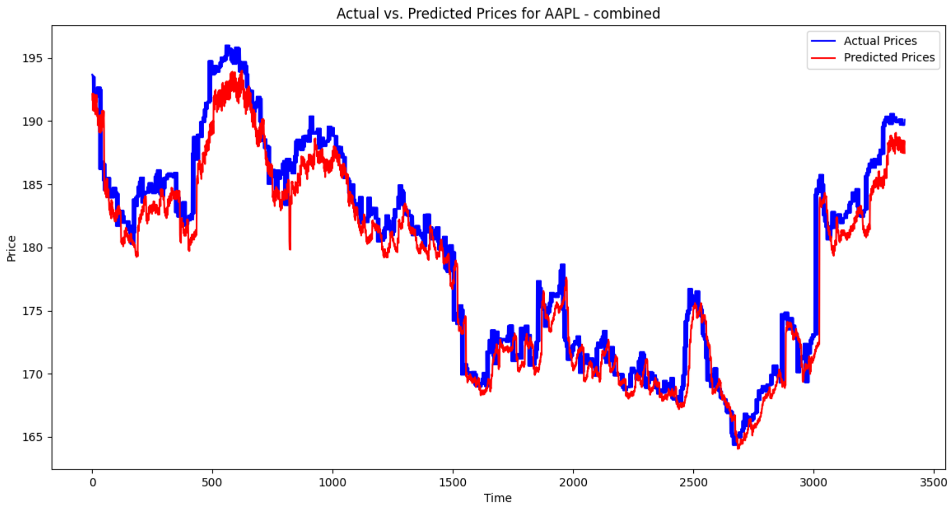
AAPL - CO2 & T (Histogram)



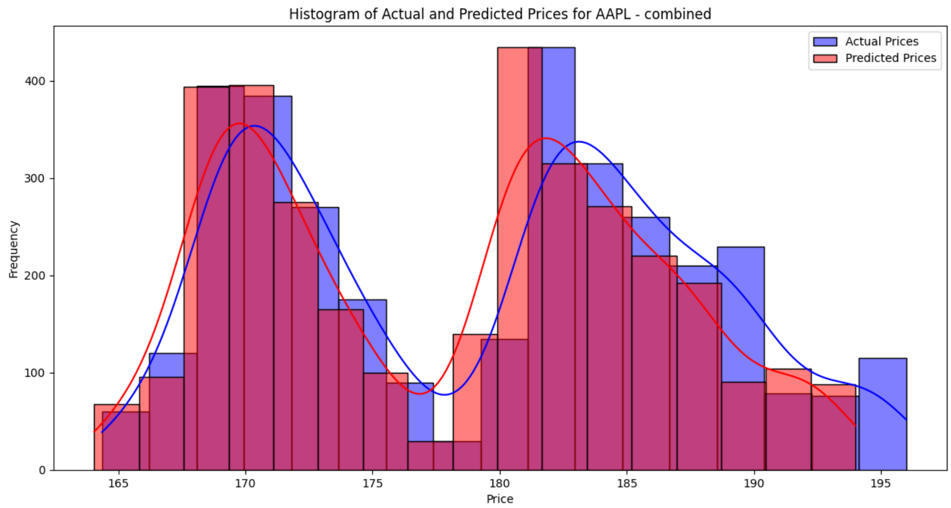
AAPL - CO2 & T (Boxplot)



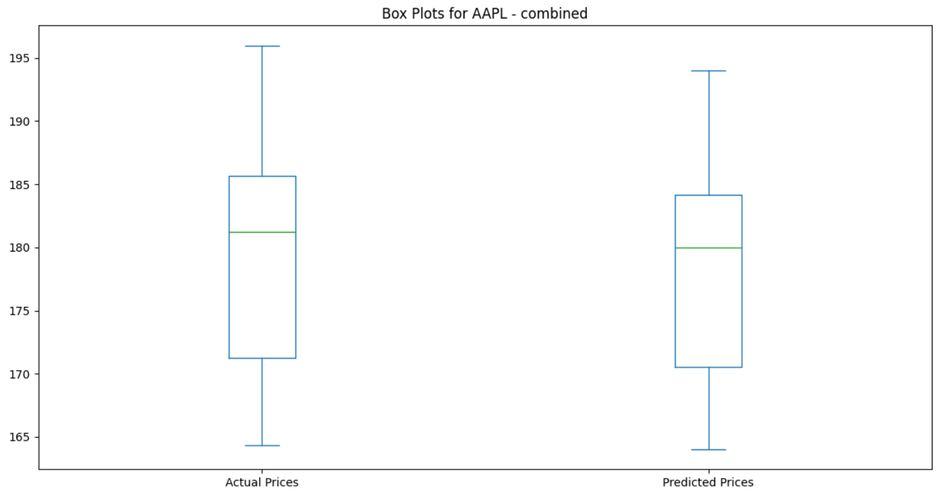
AAPL - combined (Actual vs Predicted)



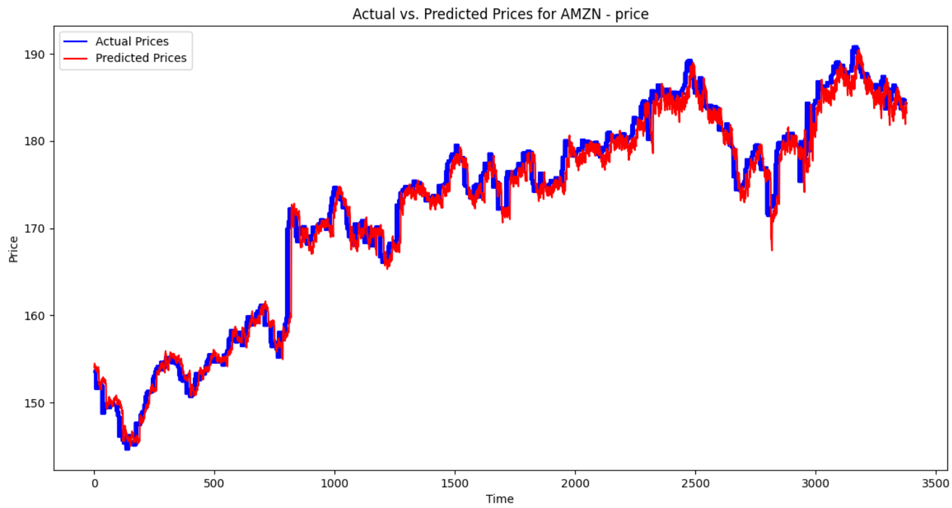
AAPL - combined (Histogram)



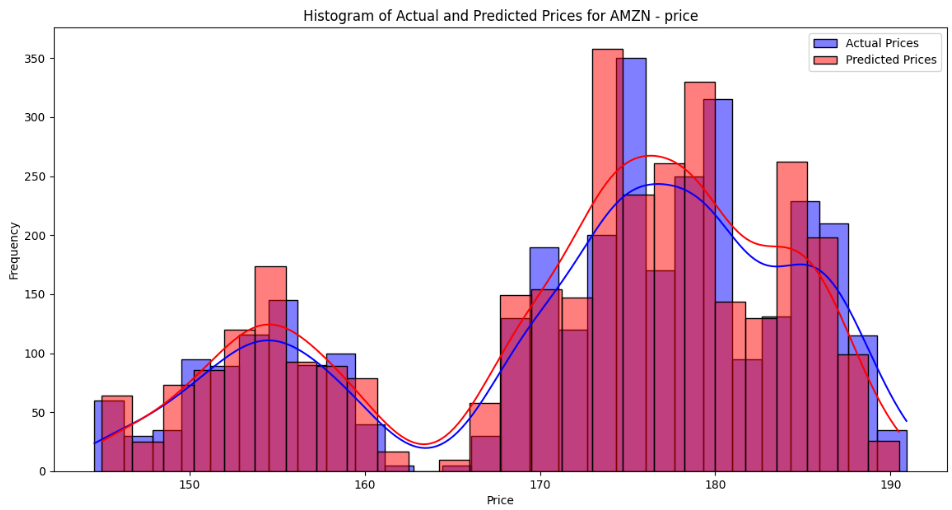
AAPL - combined (Boxplot)



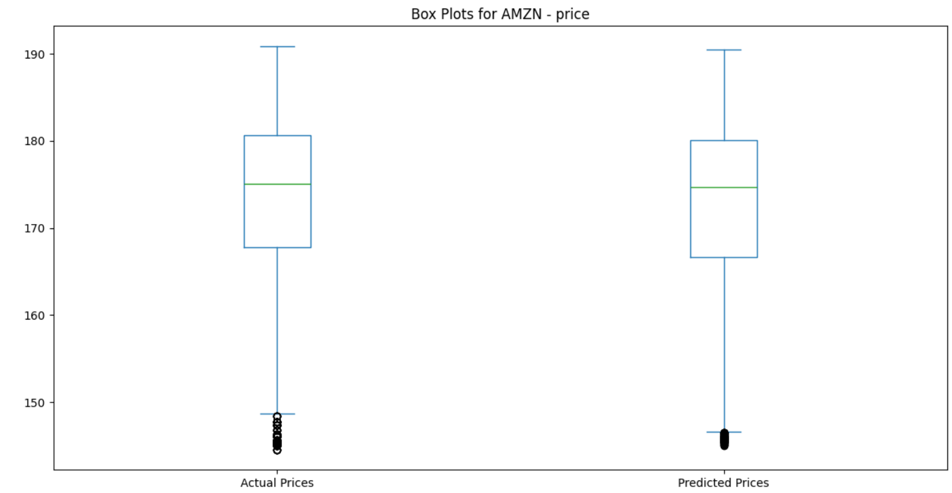
AMZN - price (Actual vs Predicted)



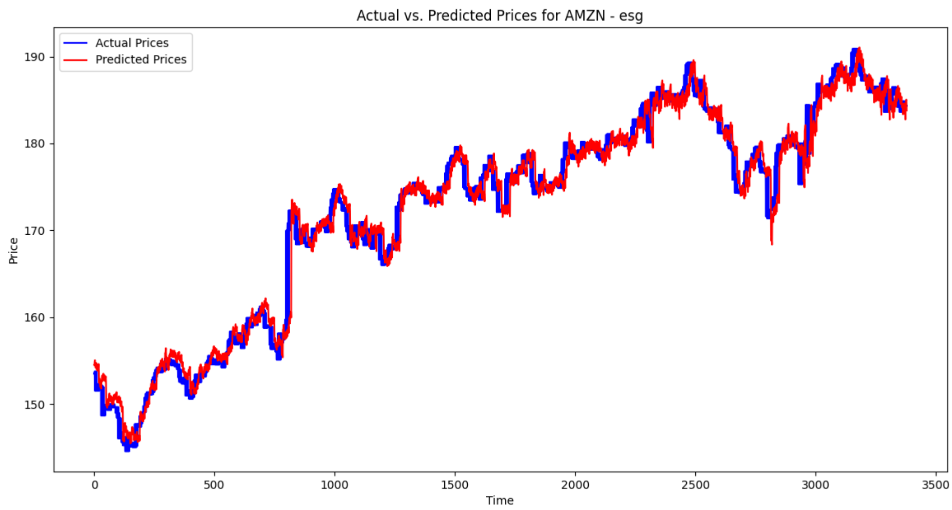
AMZN - price (Histogram)



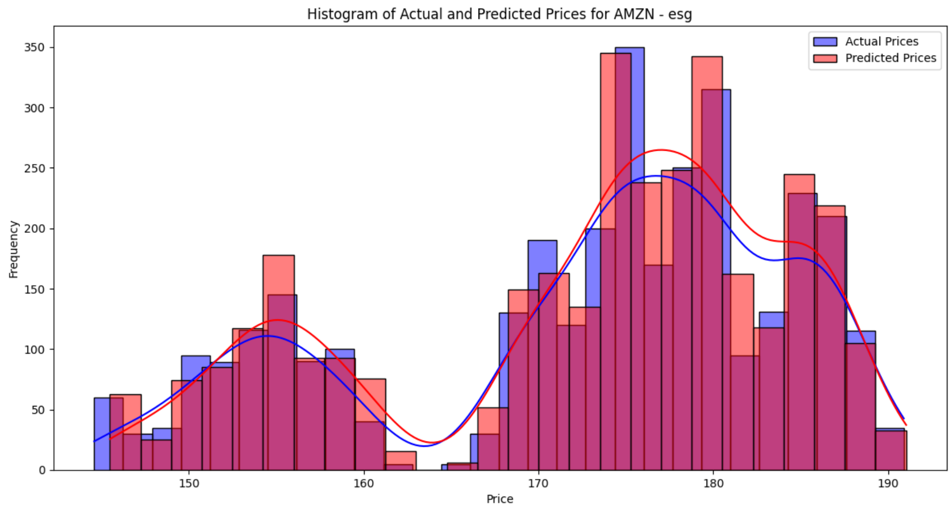
AMZN - price (Boxplot)



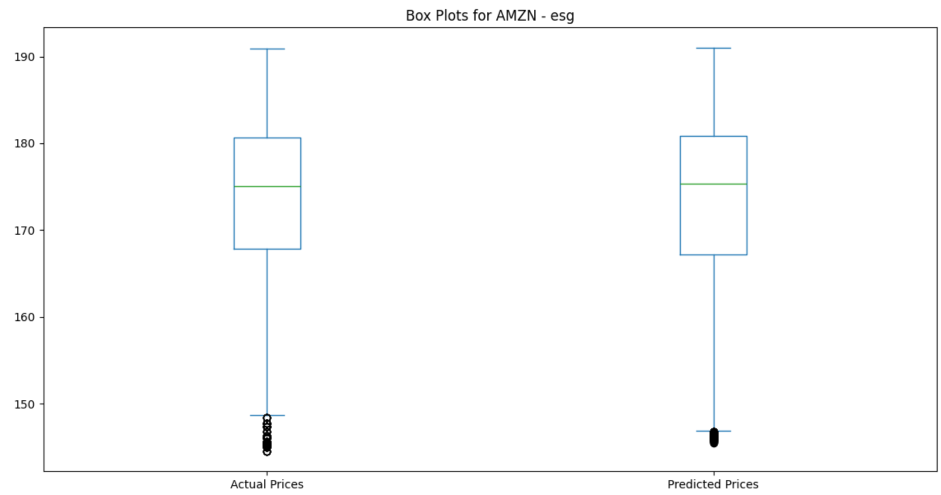
AMZN - ESG (Actual vs Predicted)



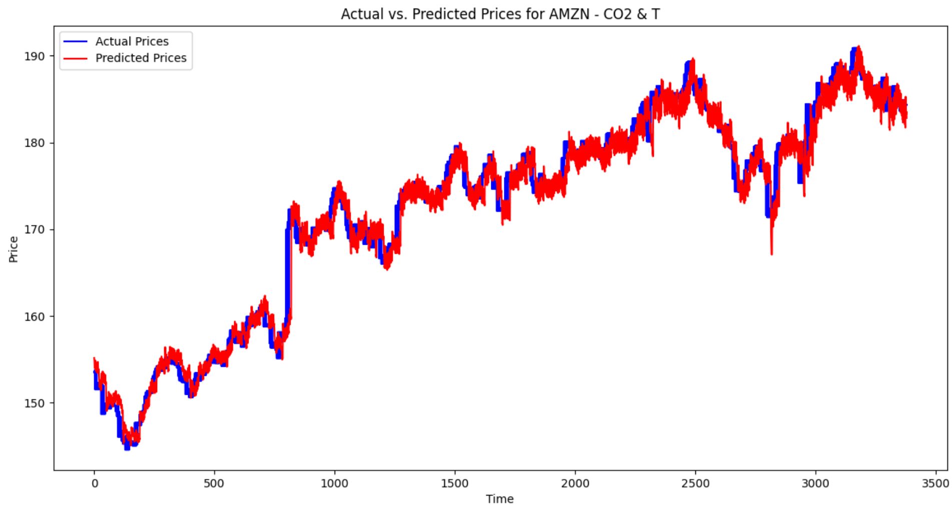
AMZN - ESG (Histogram)



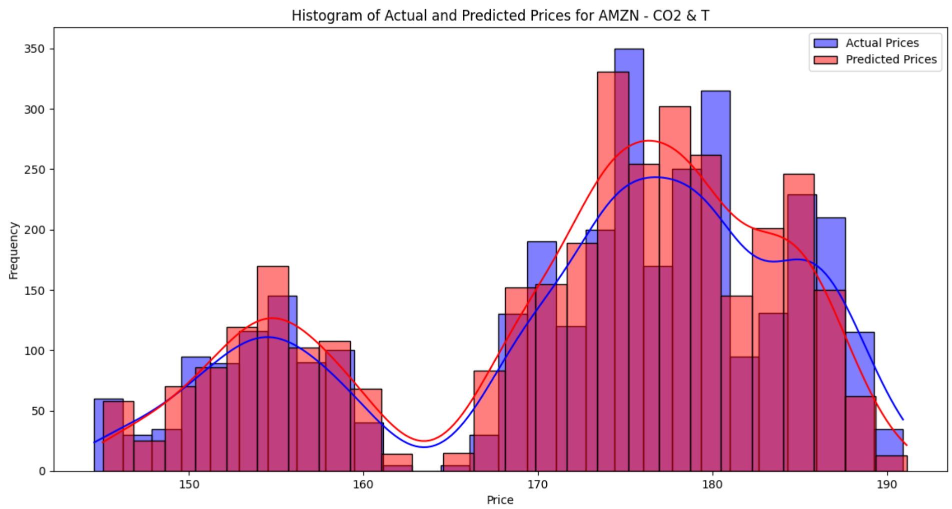
AMZN - ESG (Boxplot)



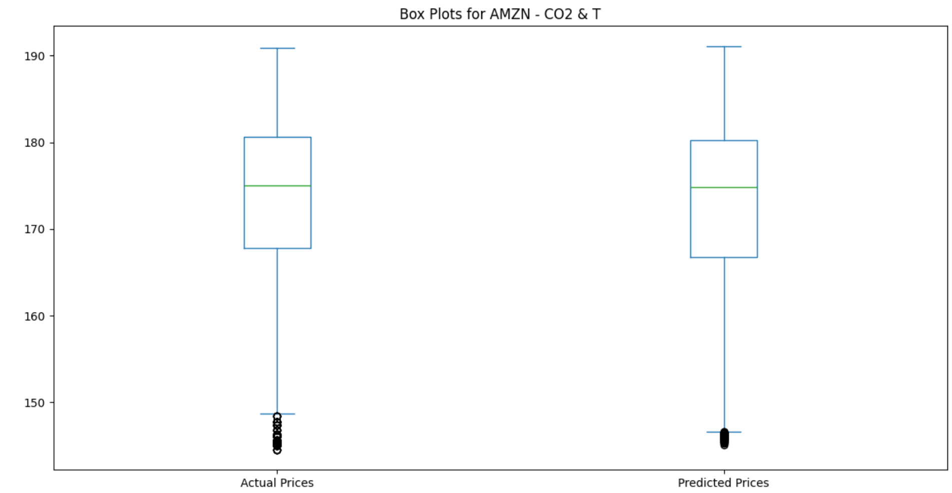
AMZN - CO2 & T (Actual vs Predicted)



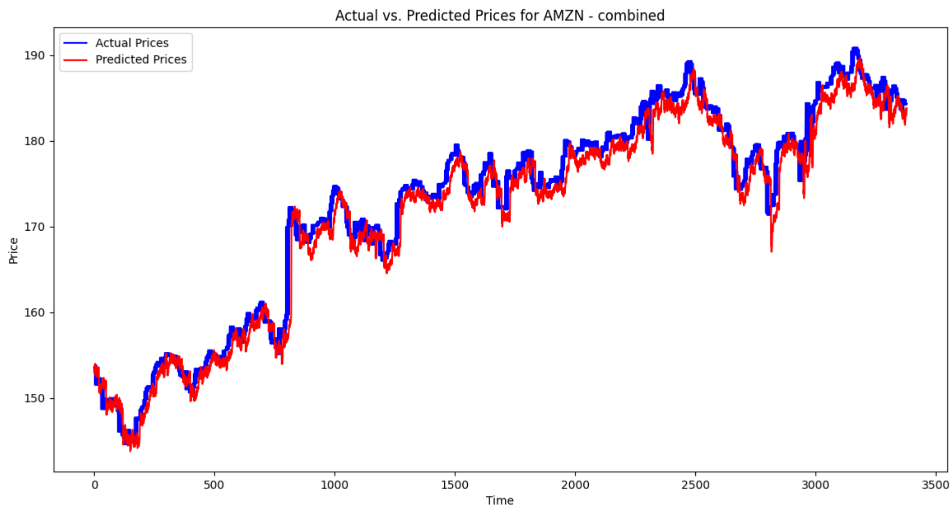
AMZN - CO2 & T (Histogram)



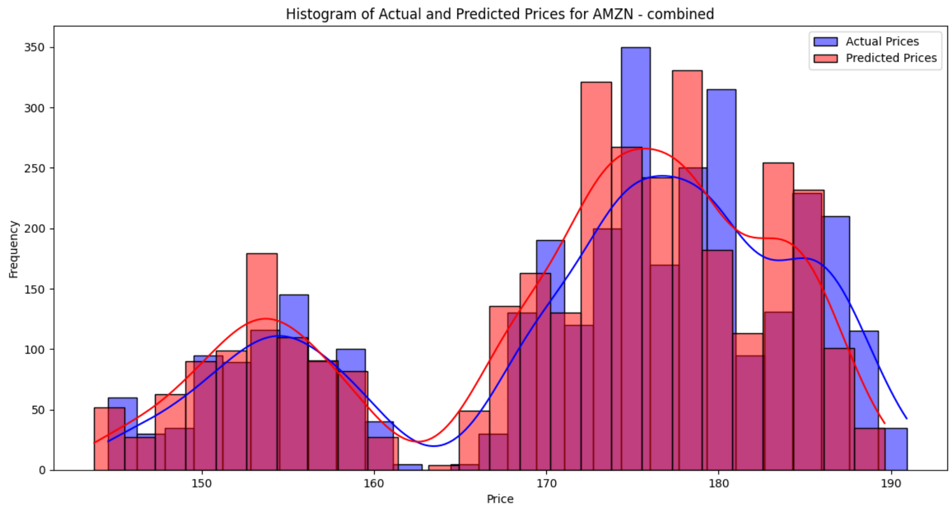
AMZN - CO2 & T (Boxplot)



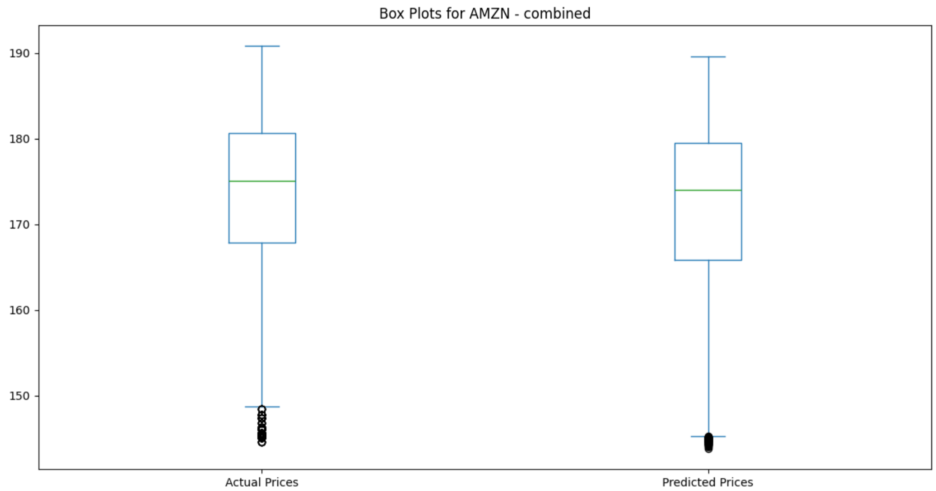
AMZN - combined (Actual vs Predicted)



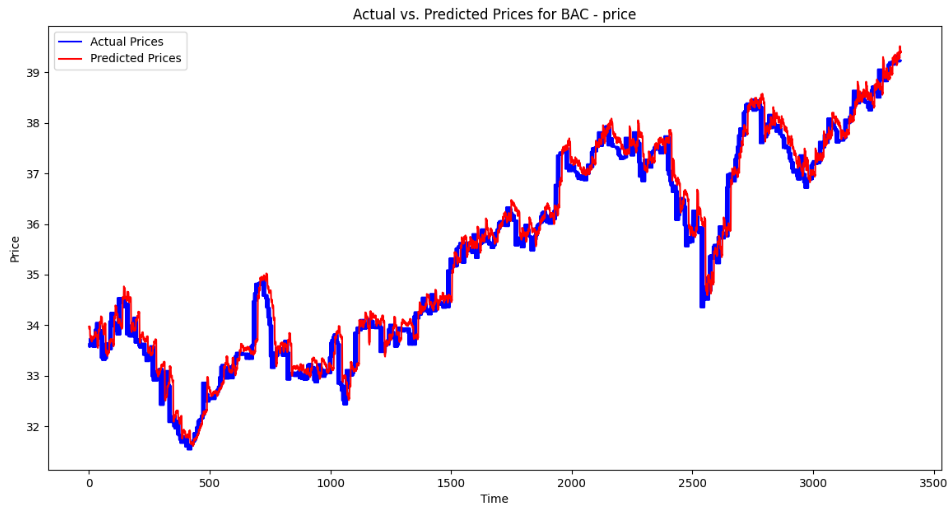
AMZN - combined (Histogram)



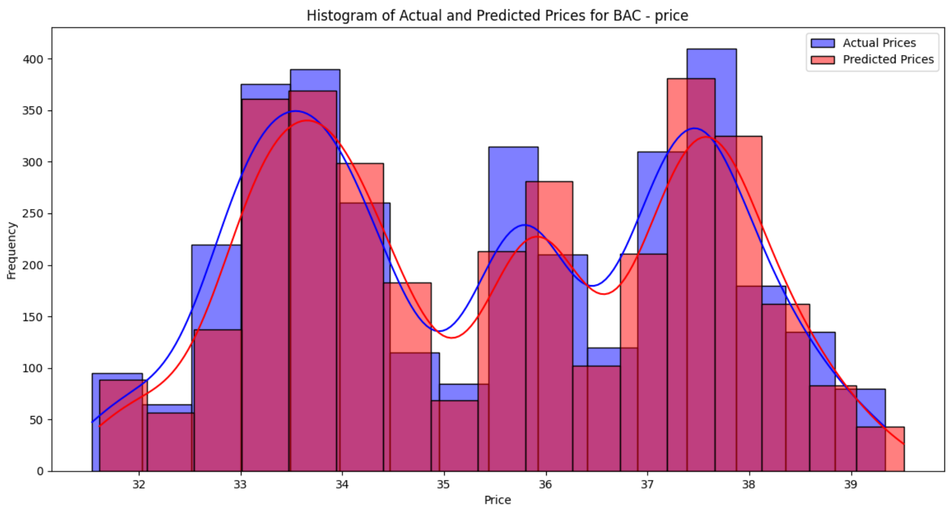
AMZN - combined (Boxplot)



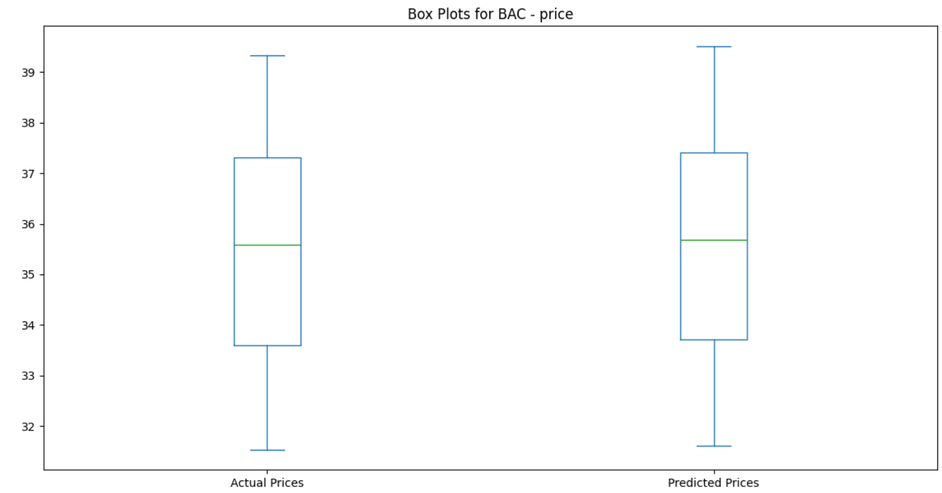
BAC - price (Actual vs Predicted)



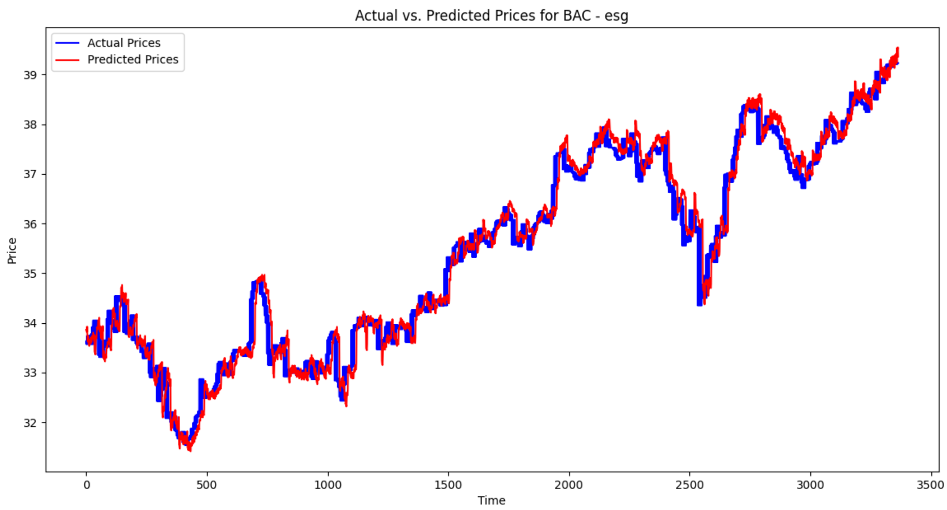
BAC - price (Histogram)



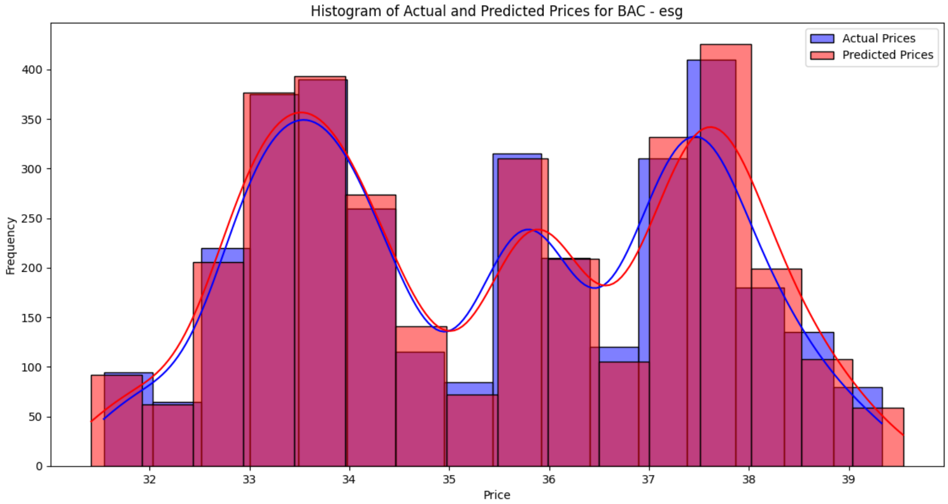
BAC - price (Boxplot)



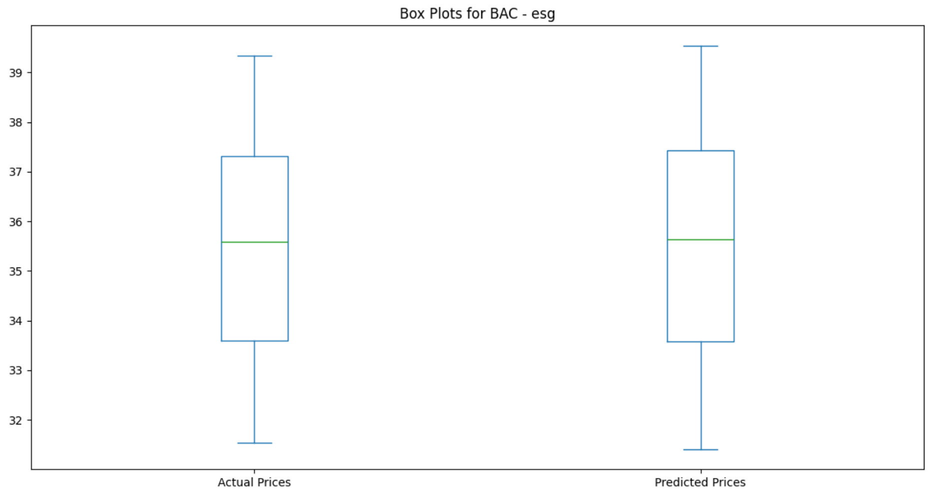
BAC - ESG (Actual vs Predicted)



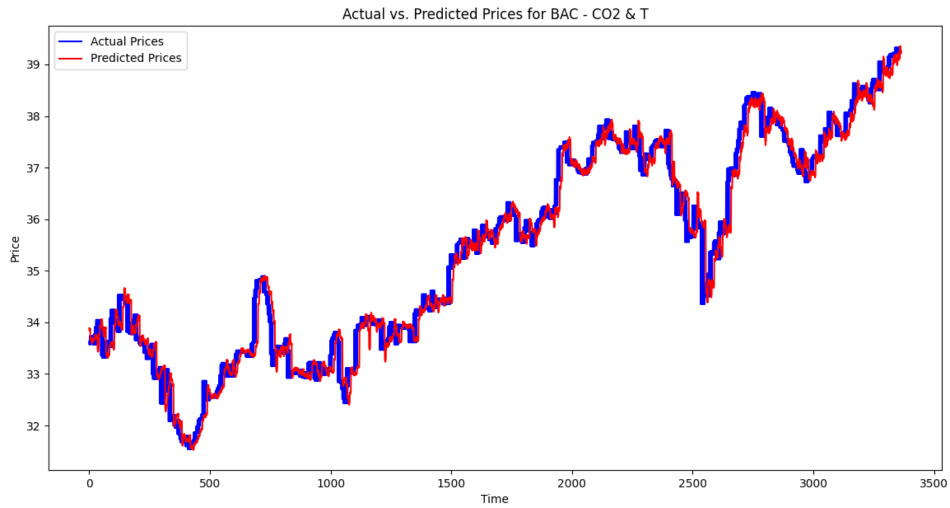
BAC - ESG (Histogram)



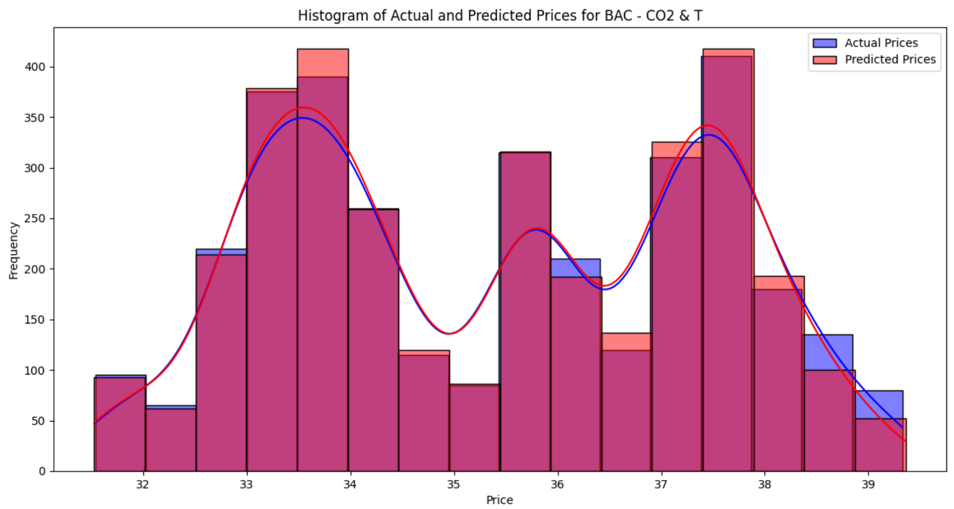
BAC - ESG (Boxplot)



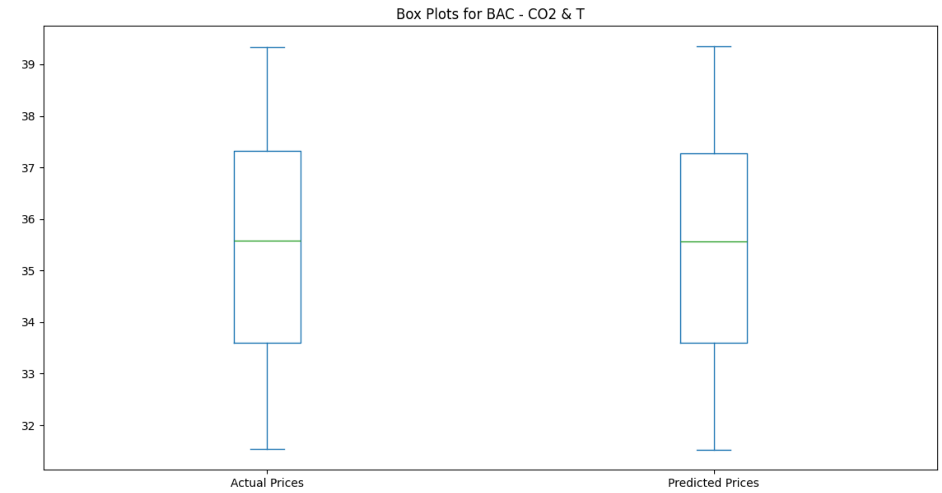
BAC - CO2 & T (Actual vs Predicted)



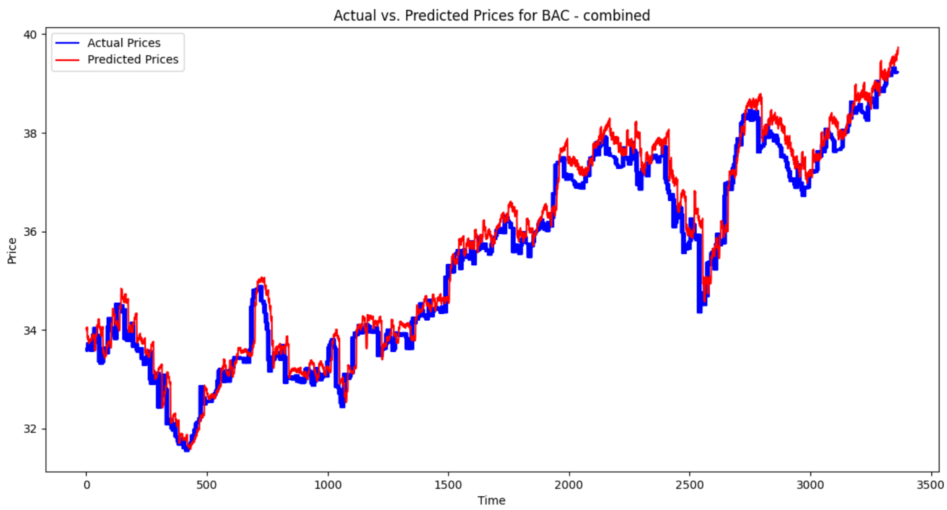
BAC - CO2 & T (Histogram)



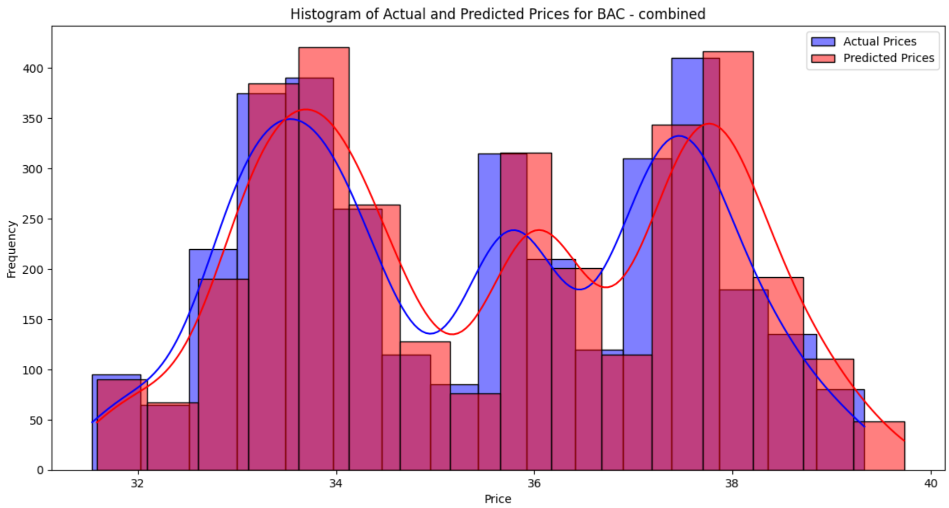
BAC - CO2 & T (Boxplot)



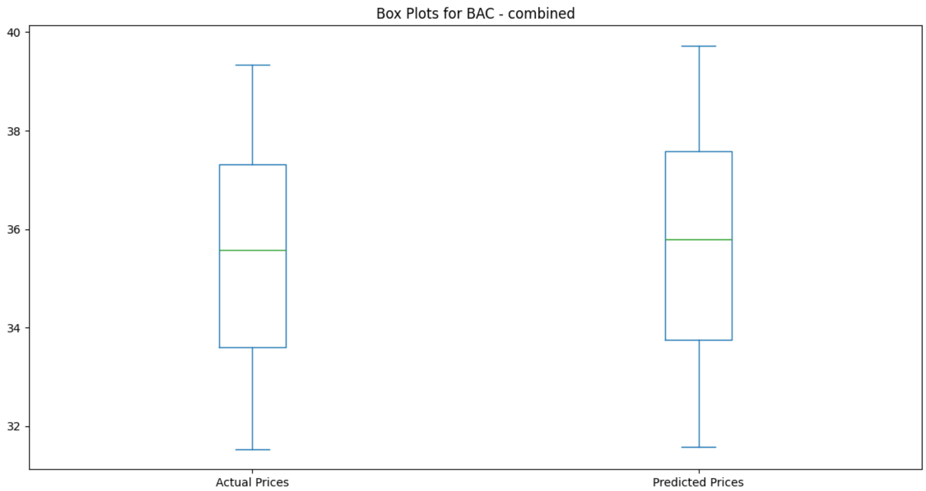
BAC - combined (Actual vs Predicted)



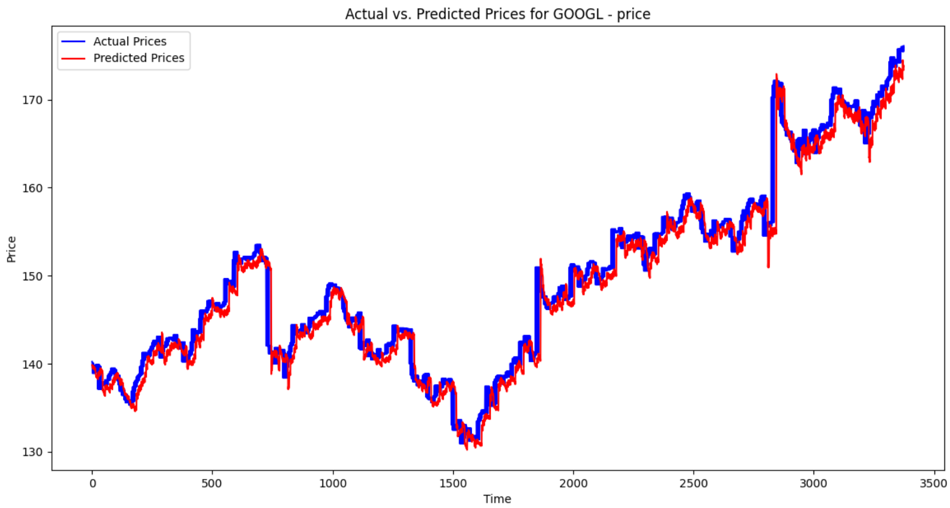
BAC - combined (Histogram)



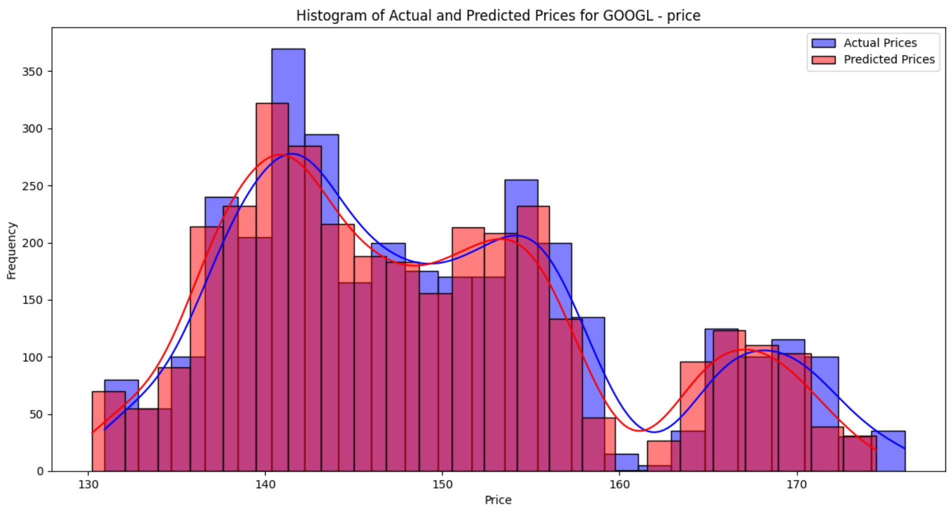
BAC - combined (Boxplot)



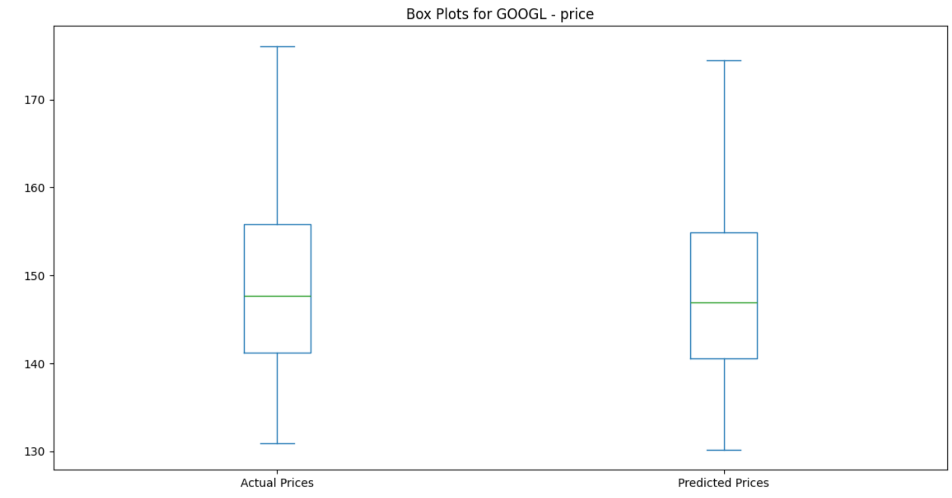
GOOGL - price (Actual vs Predicted)



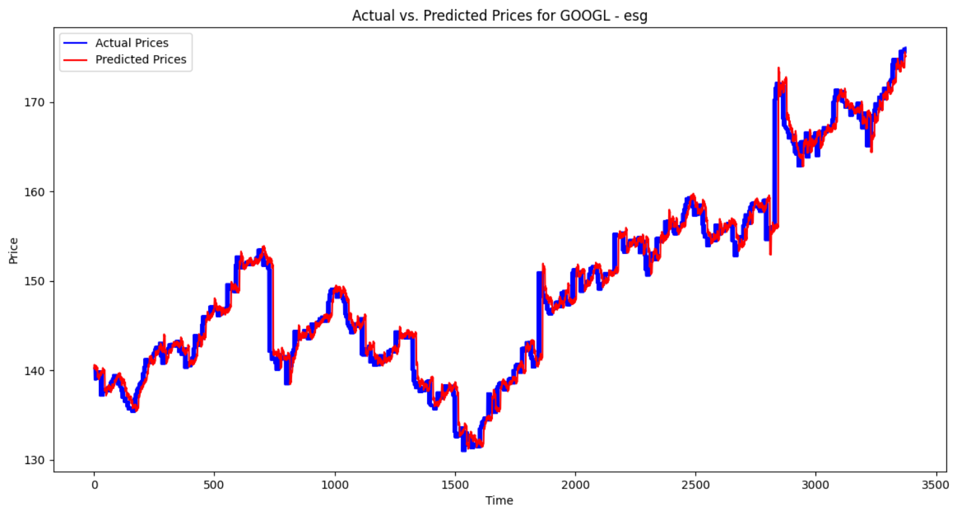
GOOGL - price (Histogram)



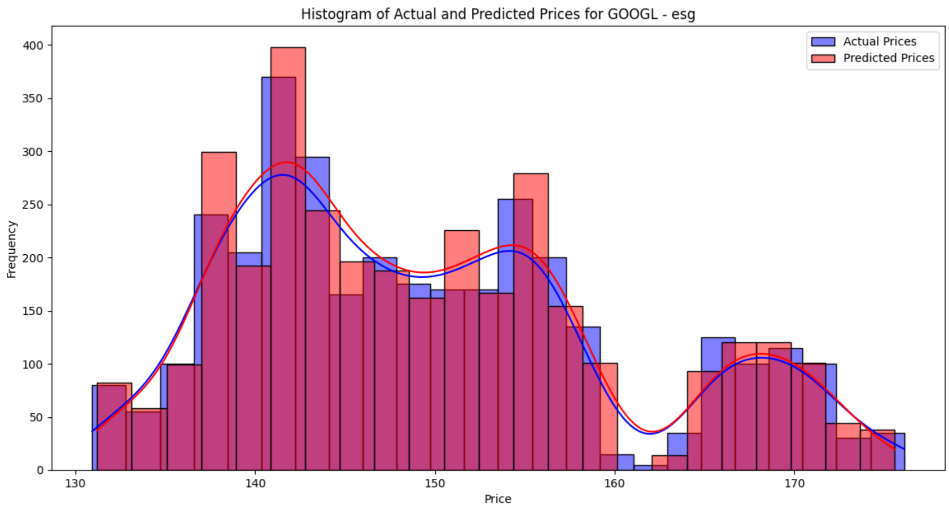
GOOGL - price (Boxplot)



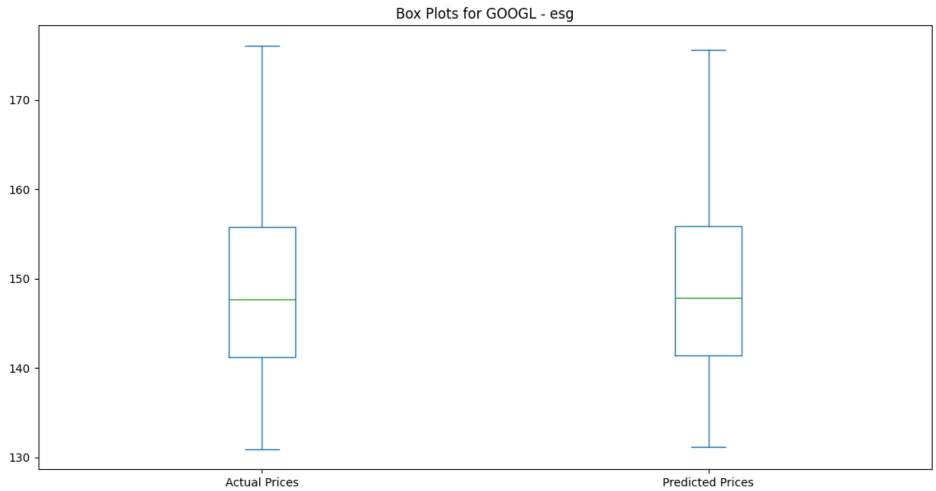
GOOGL - ESG (Actual vs Predicted)



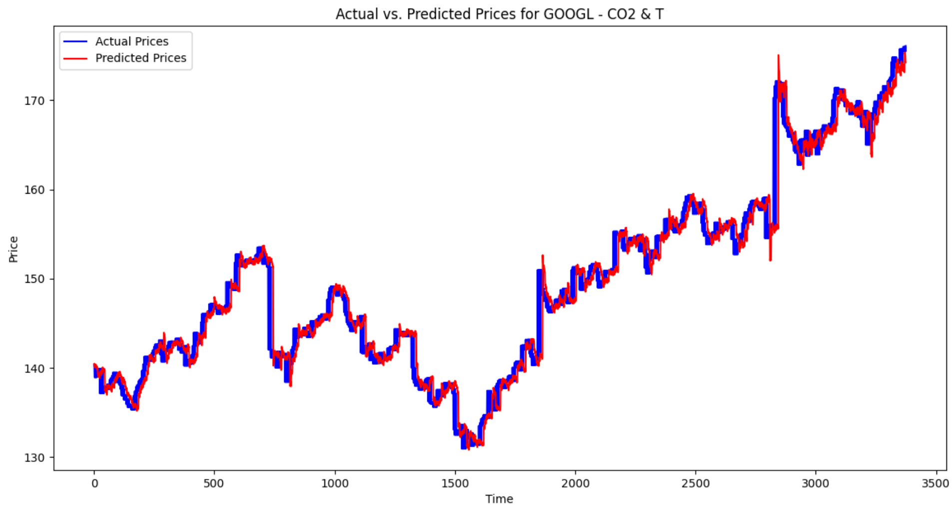
GOOGL - ESG (Histogram)



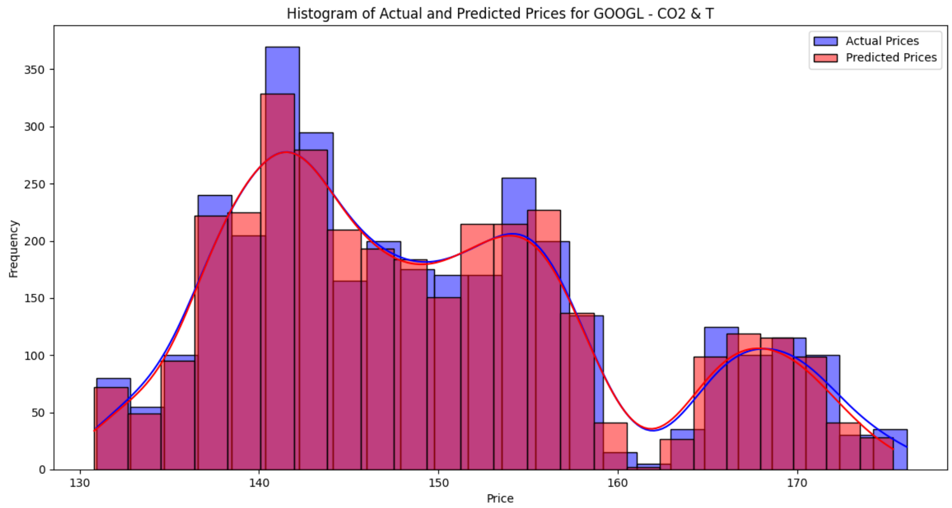
GOOGL - ESG (Boxplot)



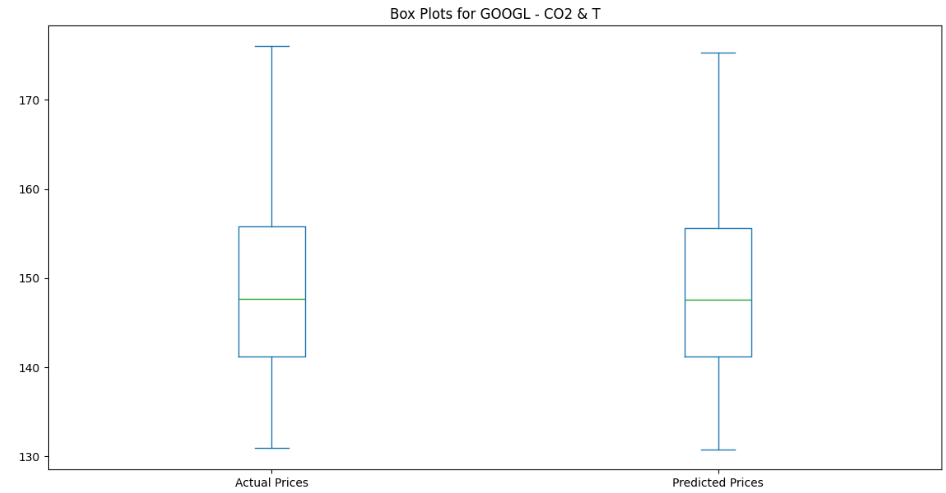
GOOGL - CO2 & T (Actual vs Predicted)



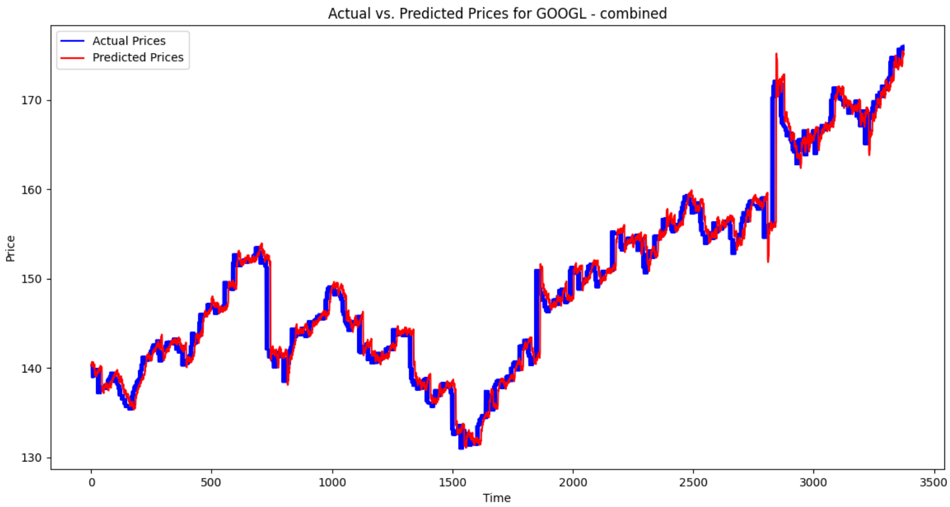
GOOGL - CO2 & T (Histogram)



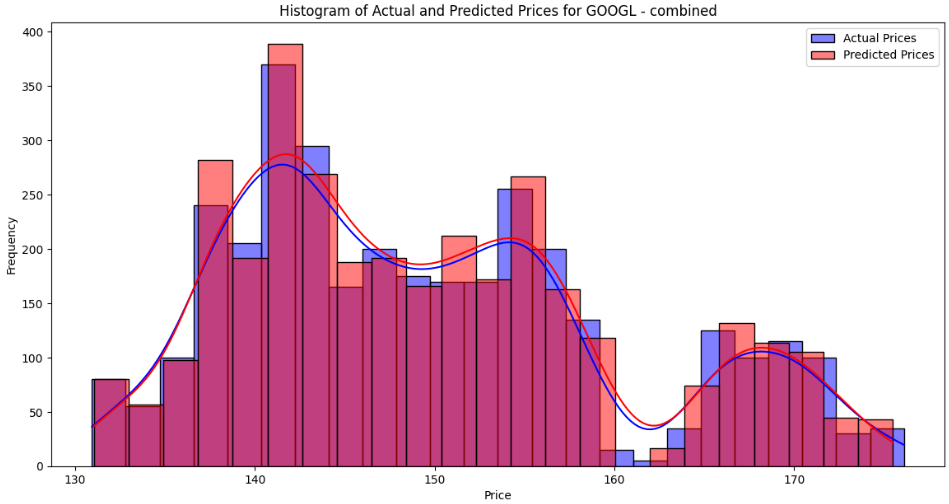
GOOGL - CO2 & T (Boxplot)



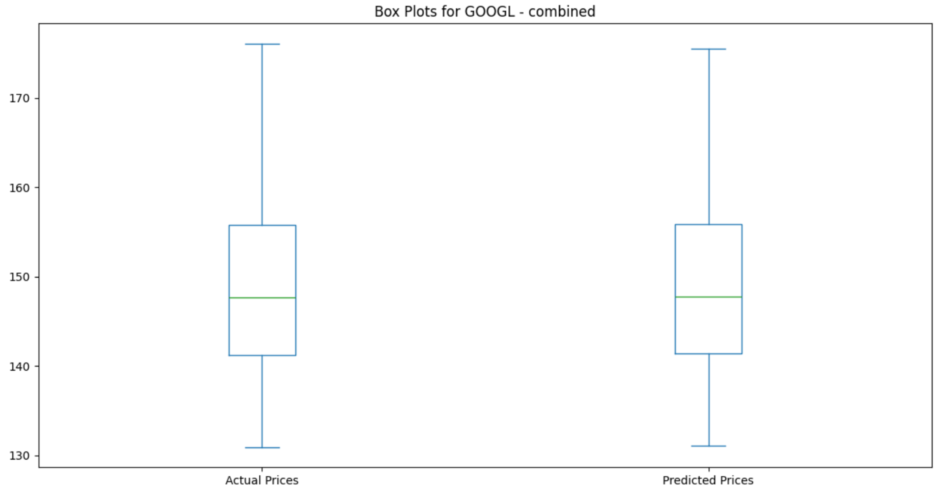
GOOGL - combined (Actual vs Predicted)



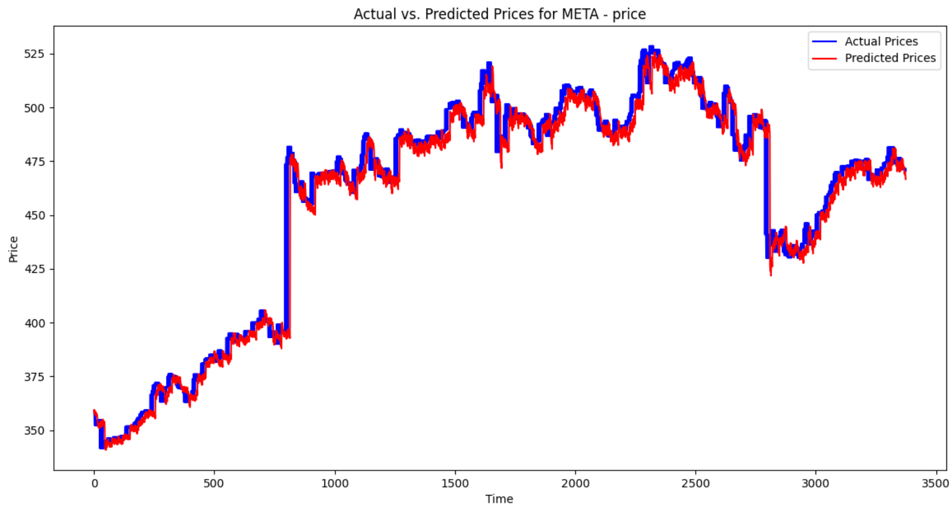
GOOGL - combined (Histogram)



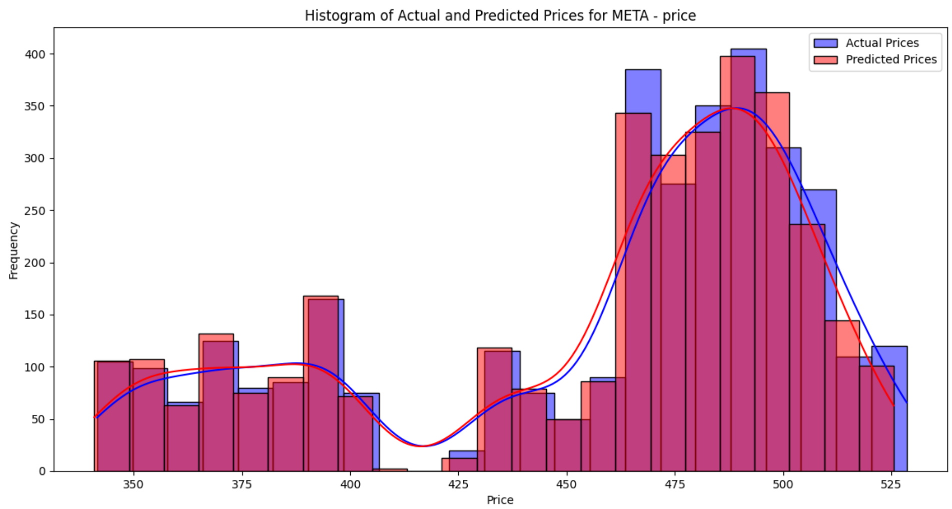
GOOGL - combined (Boxplot)



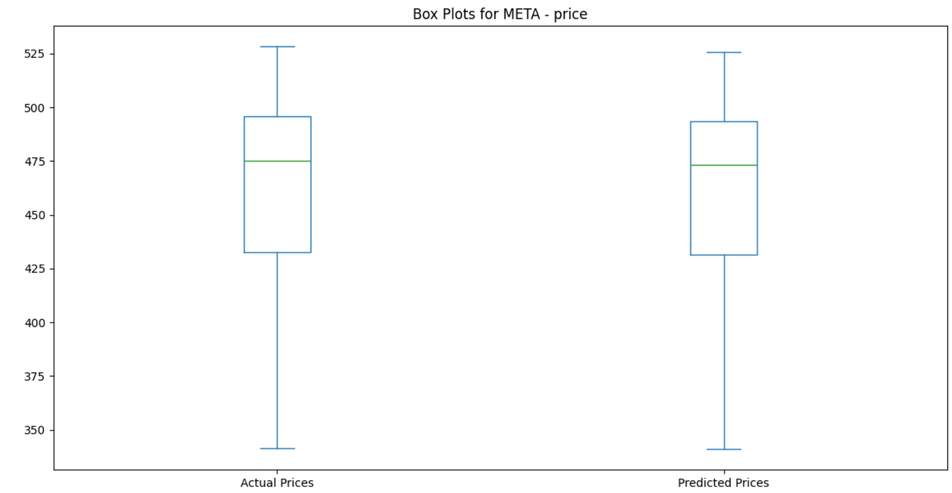
META - price (Actual vs Predicted)



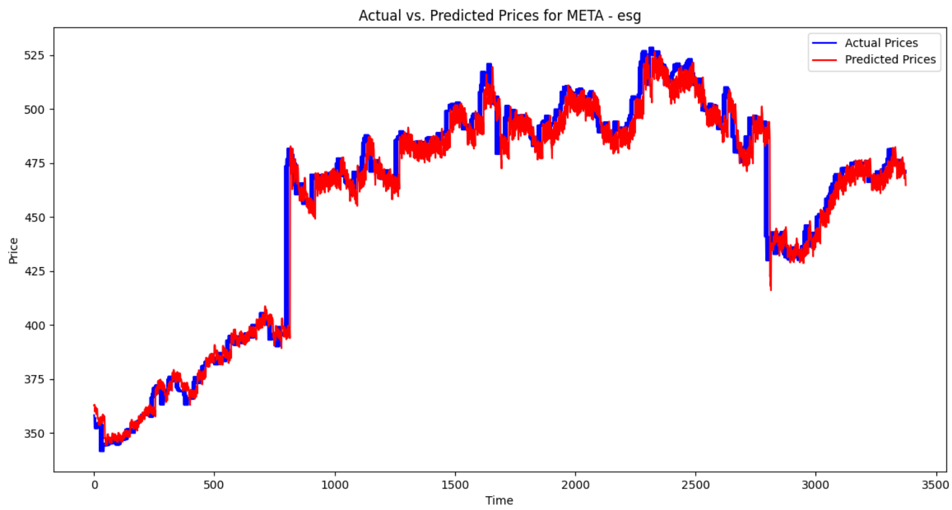
META - price (Histogram)



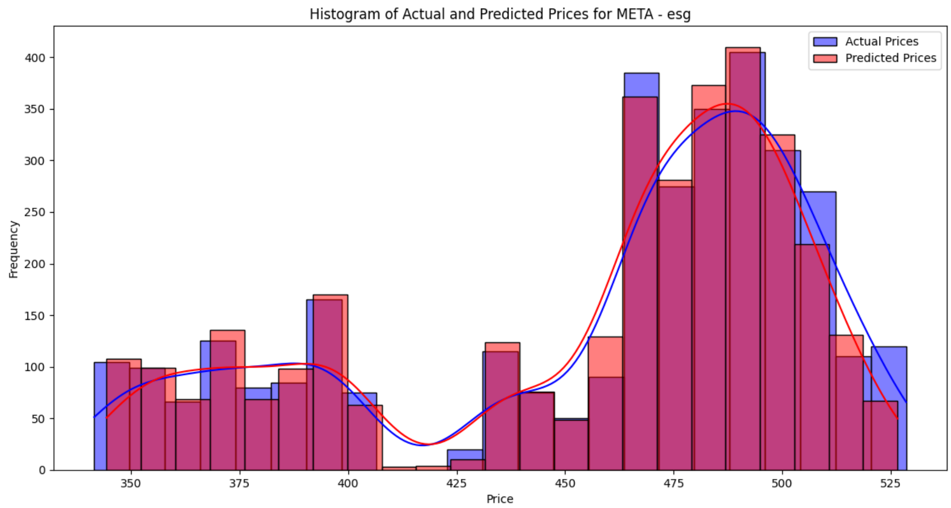
META - price (Boxplot)



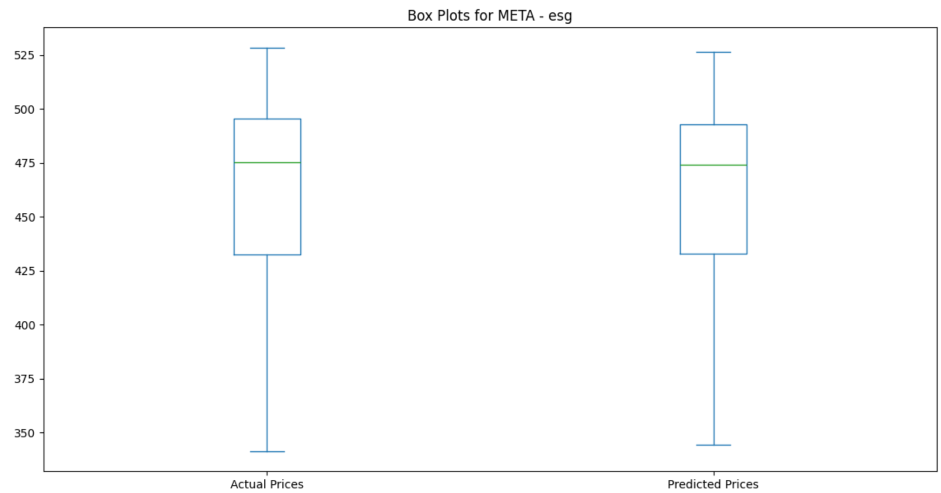
META - ESG (Actual vs Predicted)



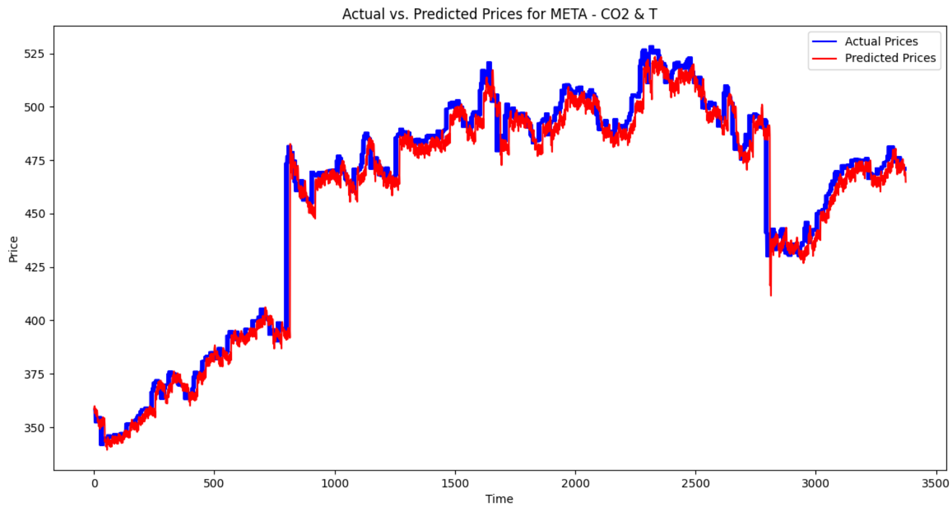
META - ESG (Histogram)



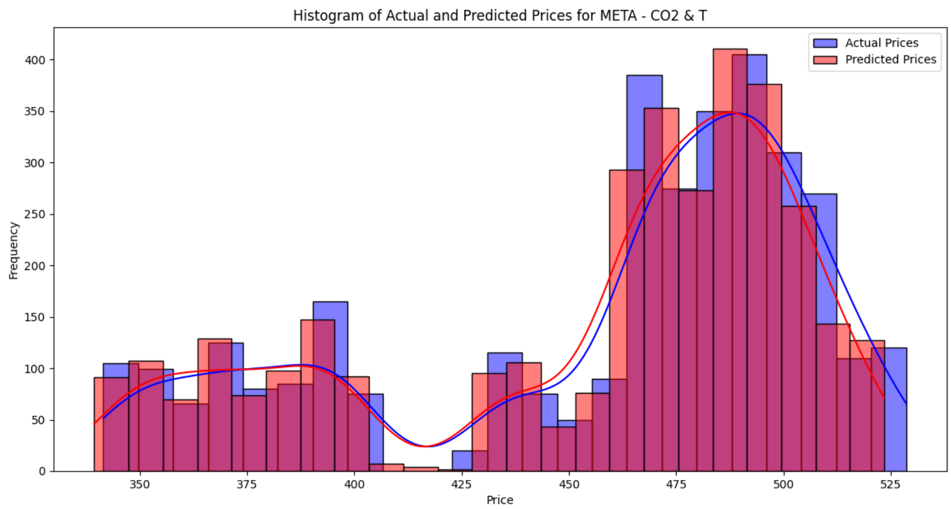
META - ESG (Boxplot)



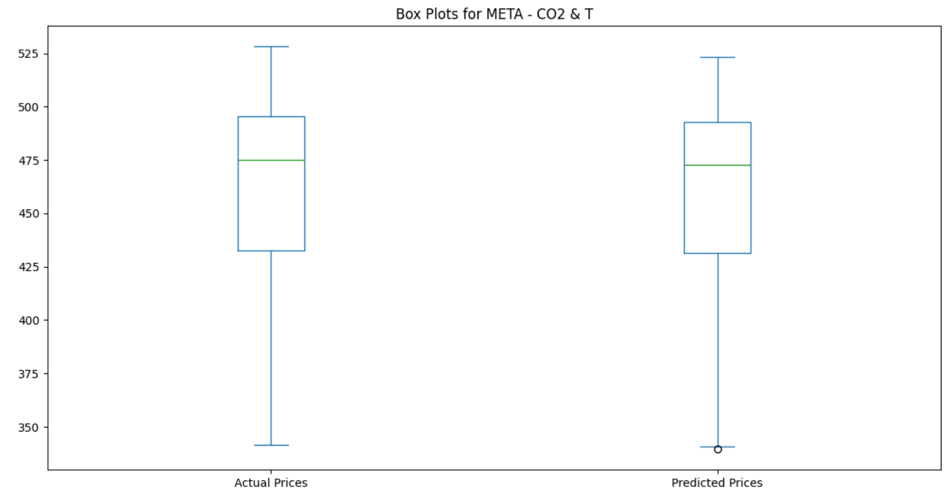
META - CO2 & T (Actual vs Predicted)



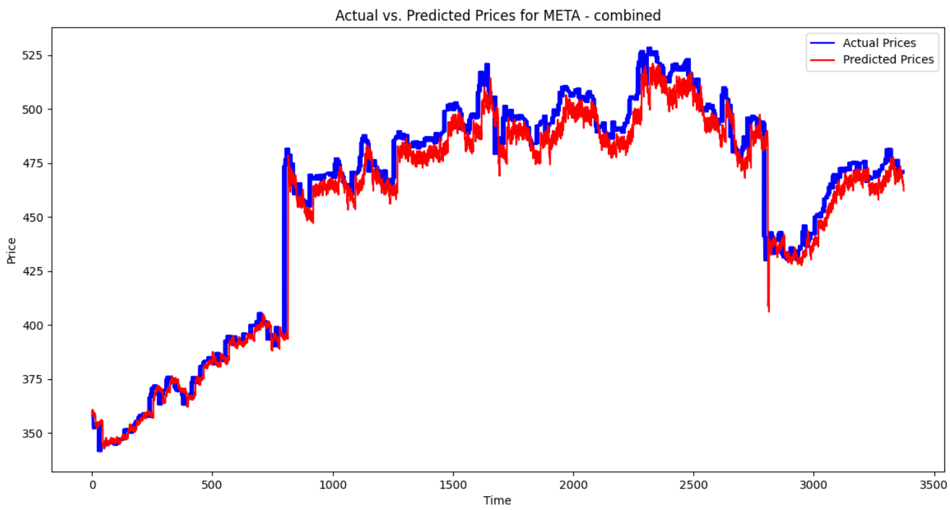
META - CO2 & T (Histogram)



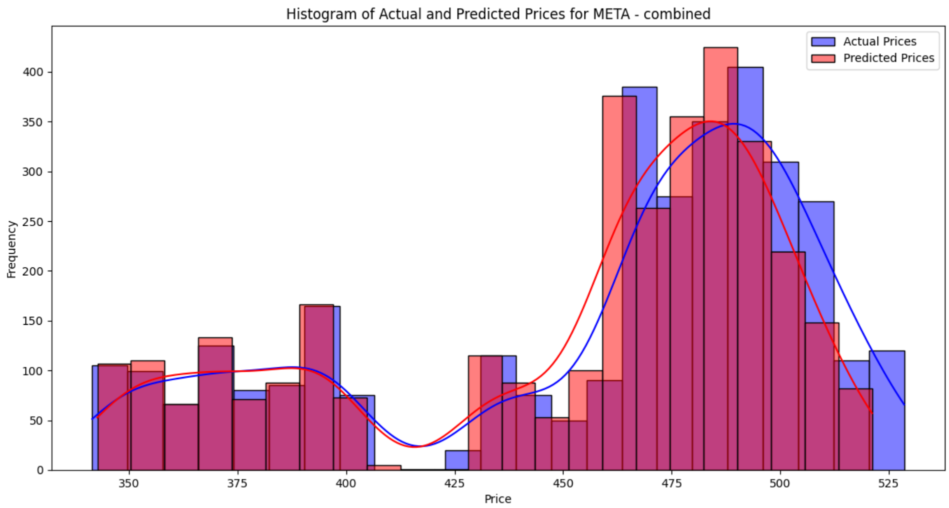
META - CO2 & T (Boxplot)



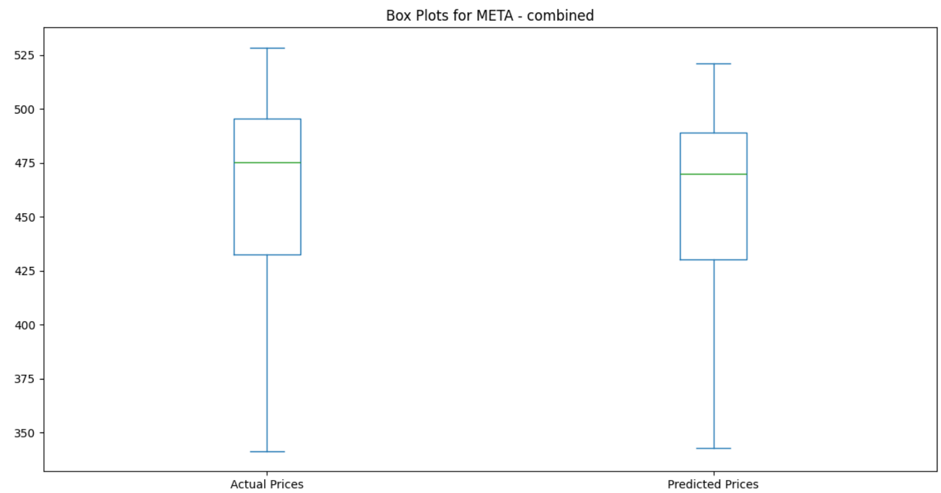
META - combined (Actual vs Predicted)



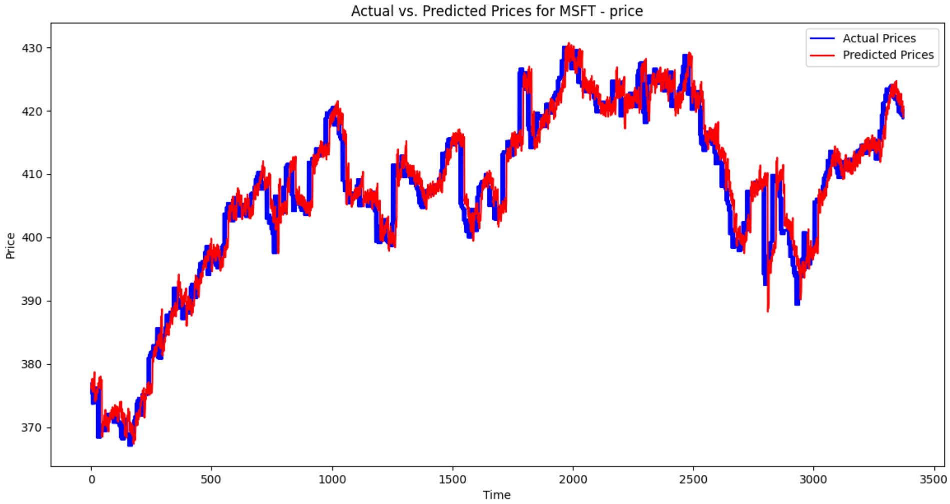
META - combined (Histogram)



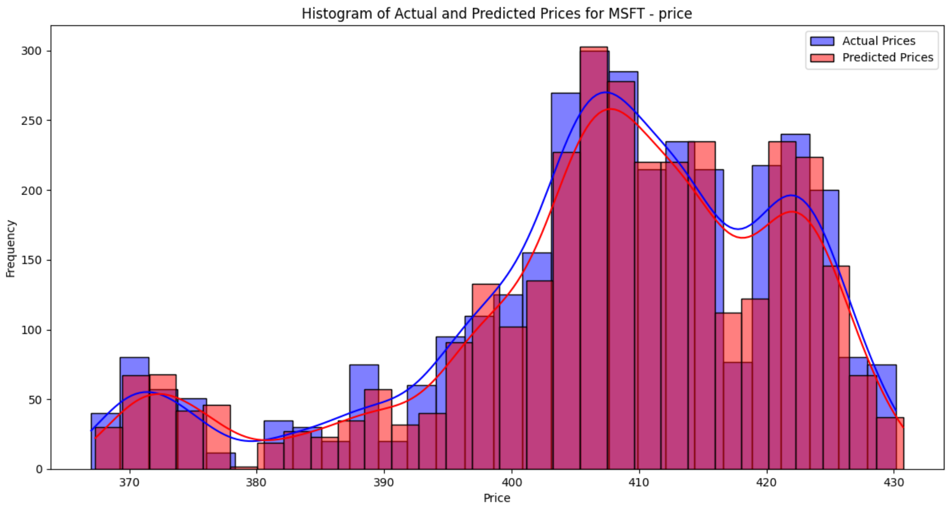
META - combined (Boxplot)



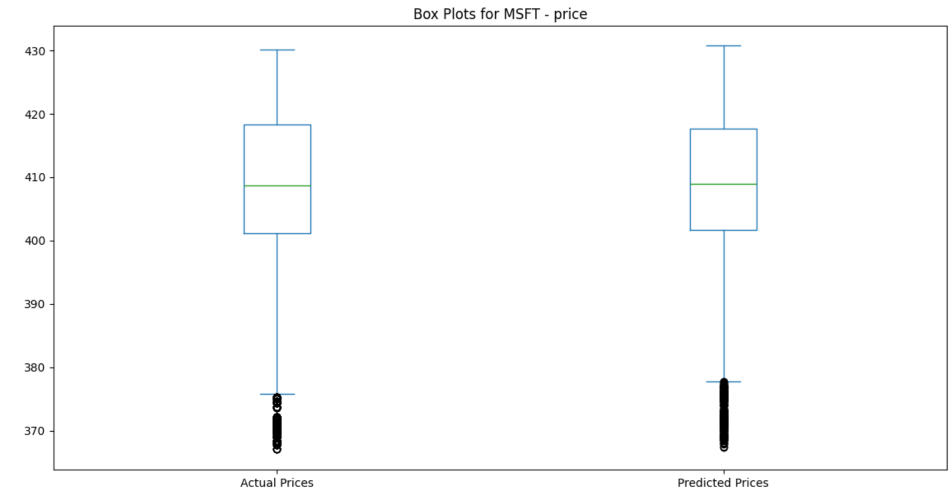
MSFT - price (Actual vs Predicted)



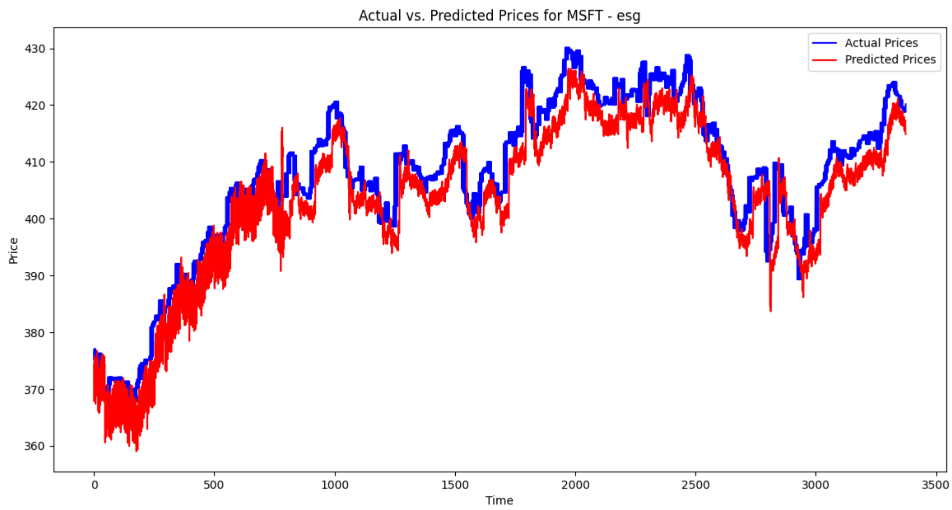
MSFT - price (Histogram)



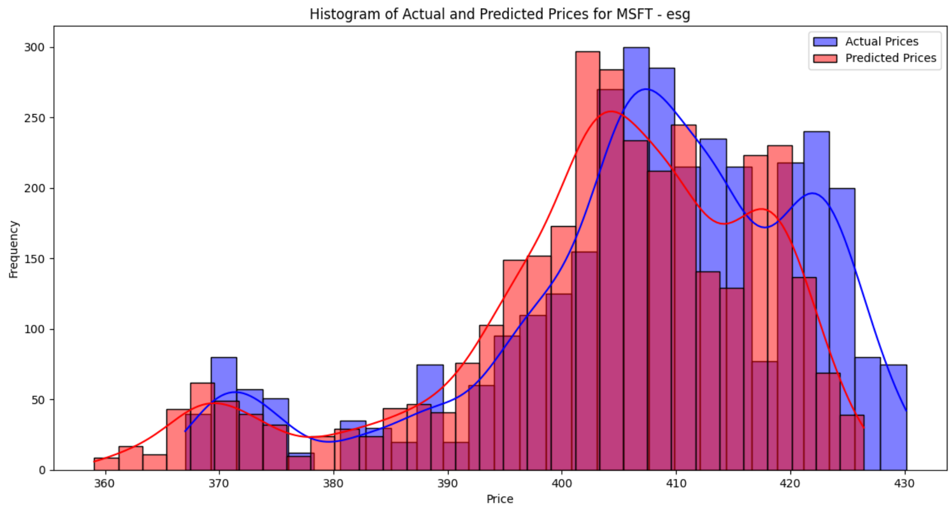
MSFT - price (Boxplot)



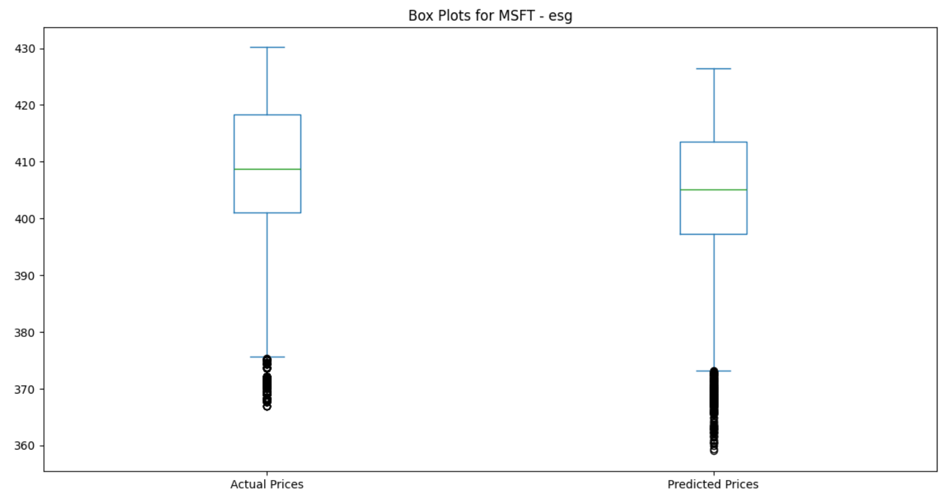
MSFT - ESG (Actual vs Predicted)



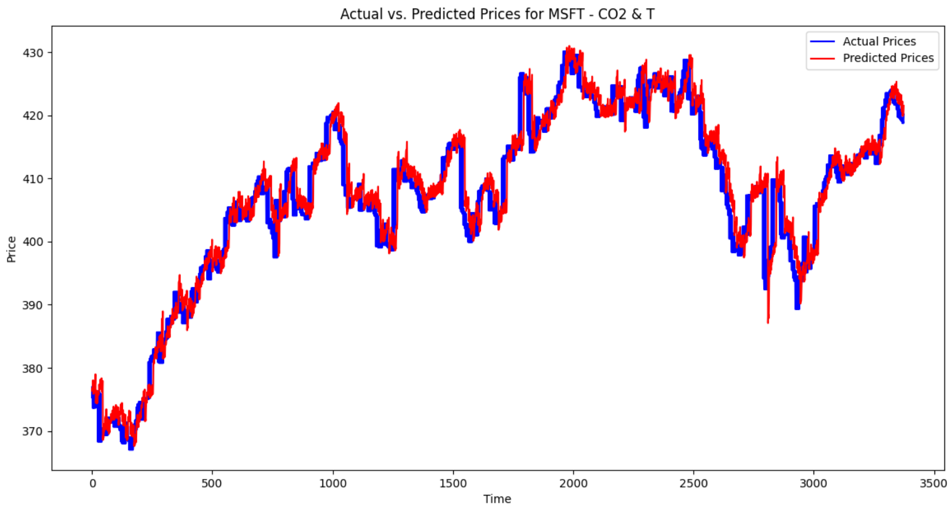
MSFT - ESG (Histogram)



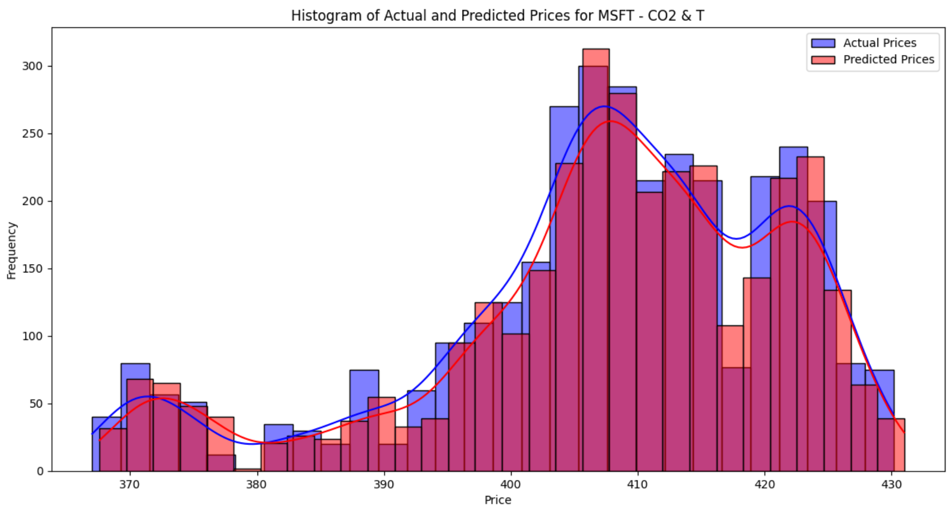
MSFT - ESG (Boxplot)



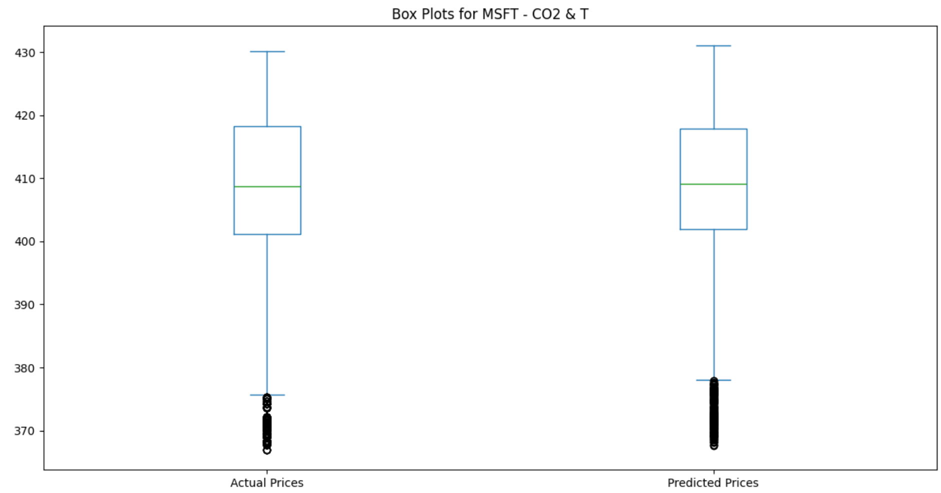
MSFT - CO2 & T (Actual vs Predicted)



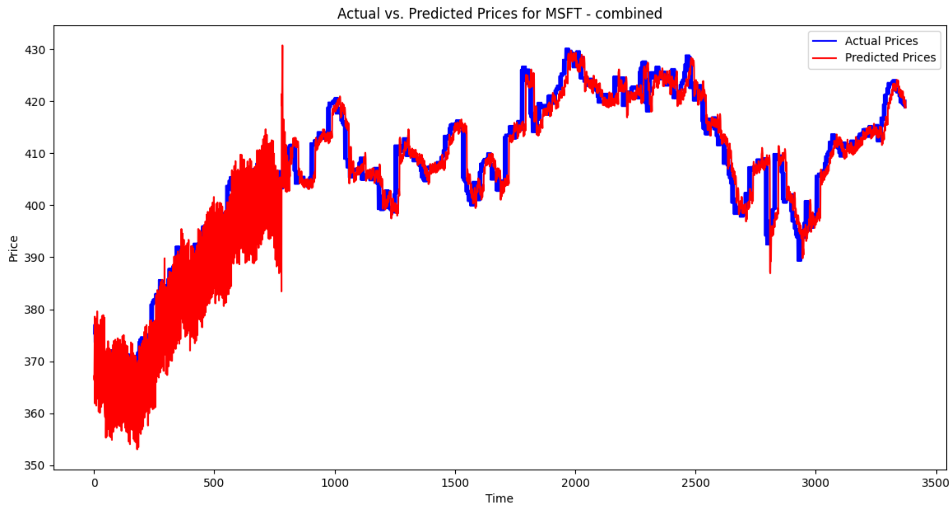
MSFT - CO2 & T (Histogram)



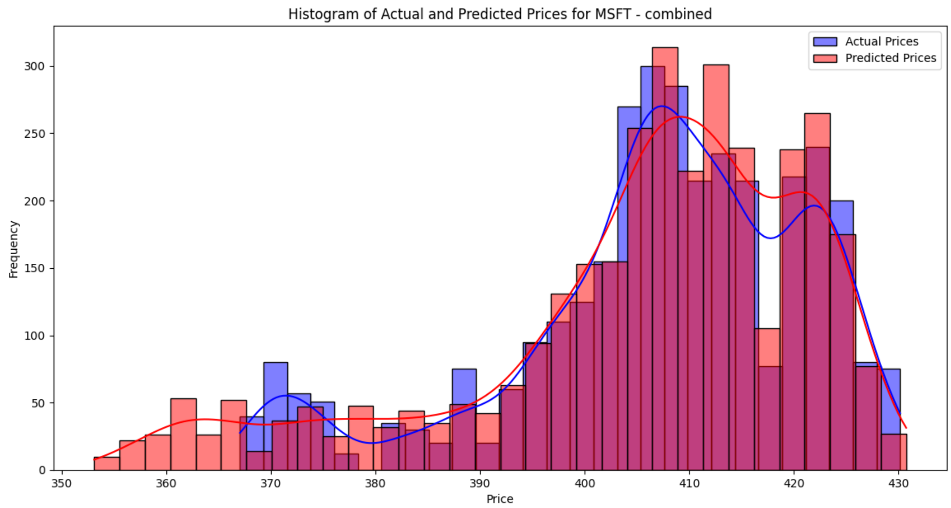
MSFT - CO2 & T (Boxplot)



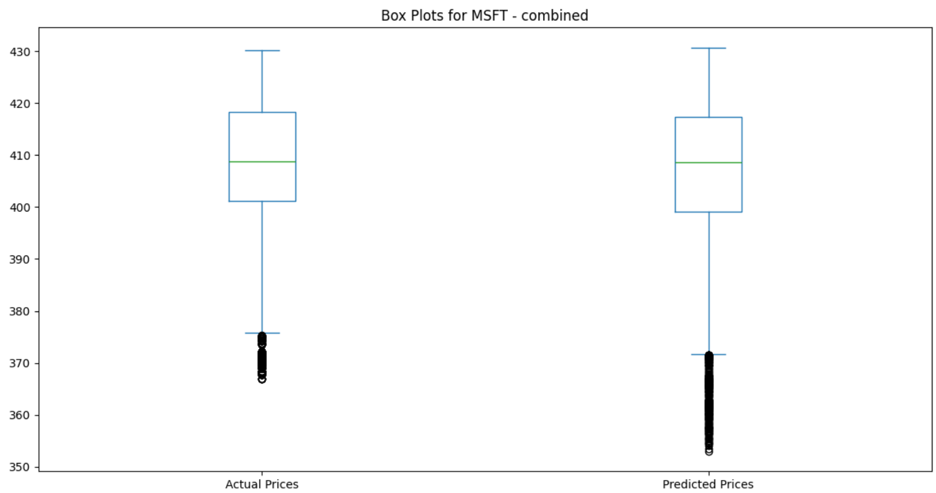
MSFT - combined (Actual vs Predicted)



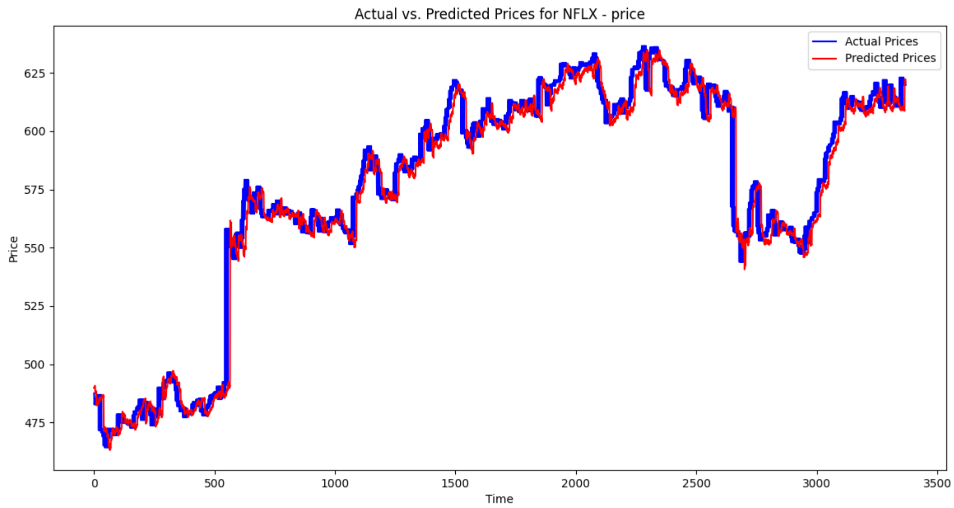
MSFT - combined (Histogram)



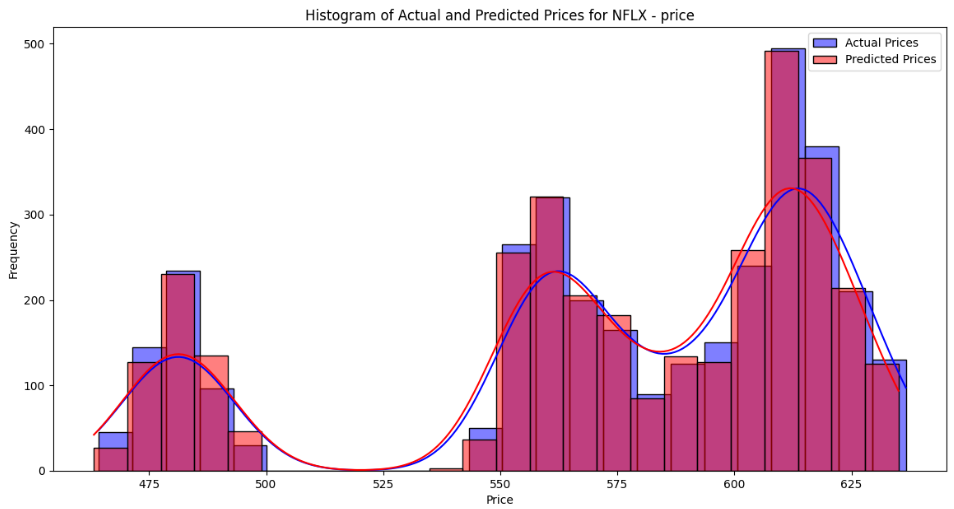
MSFT - combined (Boxplot)



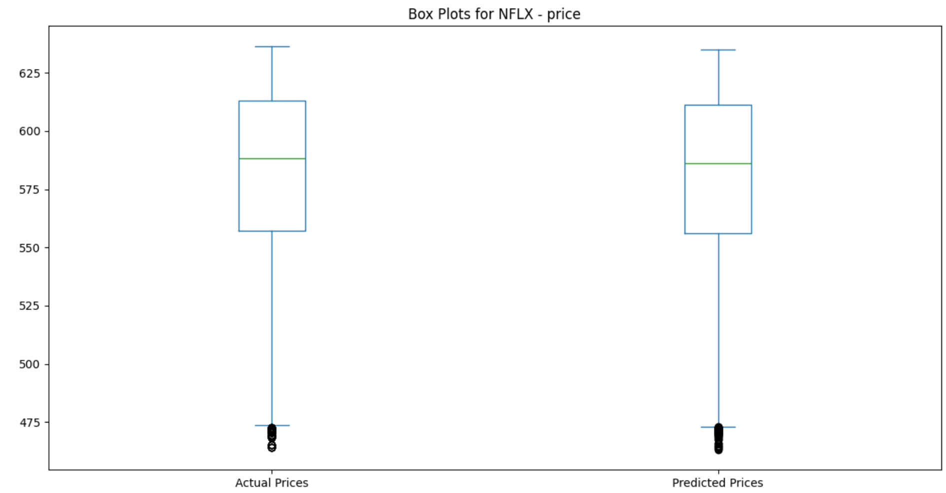
NFLX - price (Actual vs Predicted)



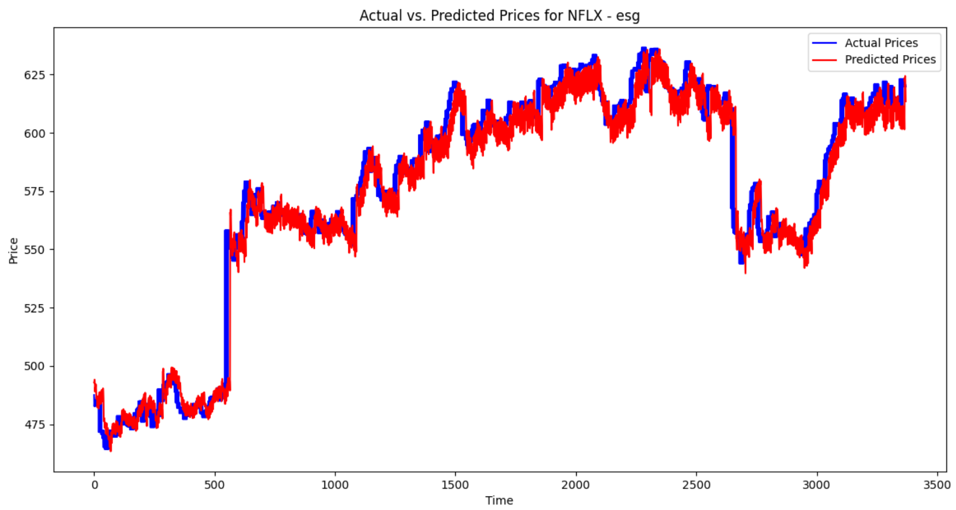
NFLX - price (Histogram)



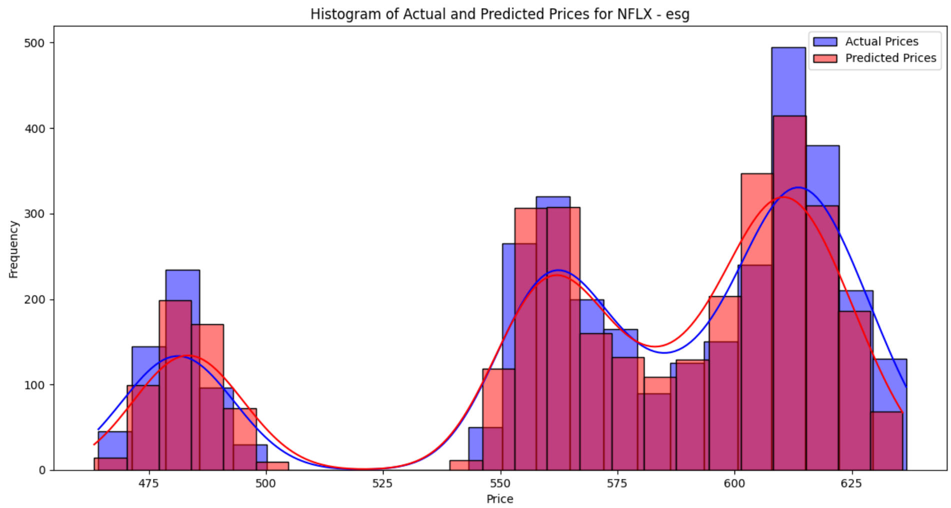
NFLX - price (Boxplot)



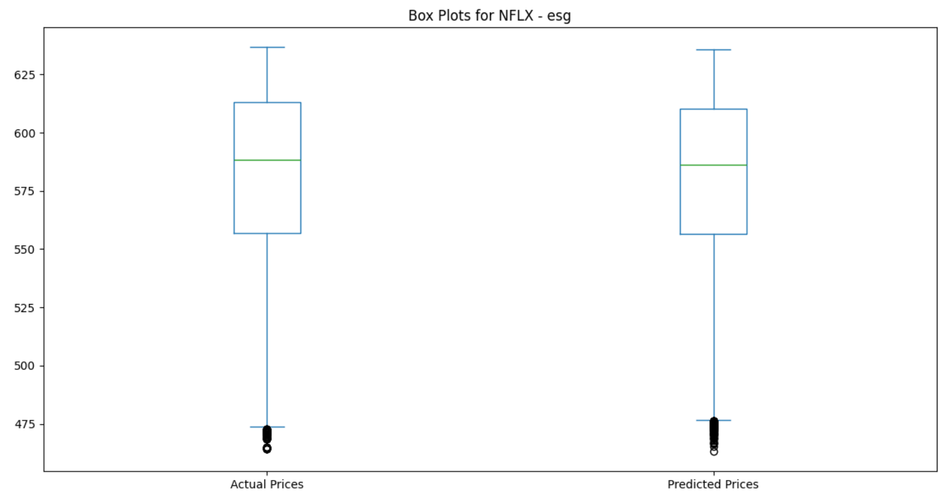
NFLX - ESG (Actual vs Predicted)



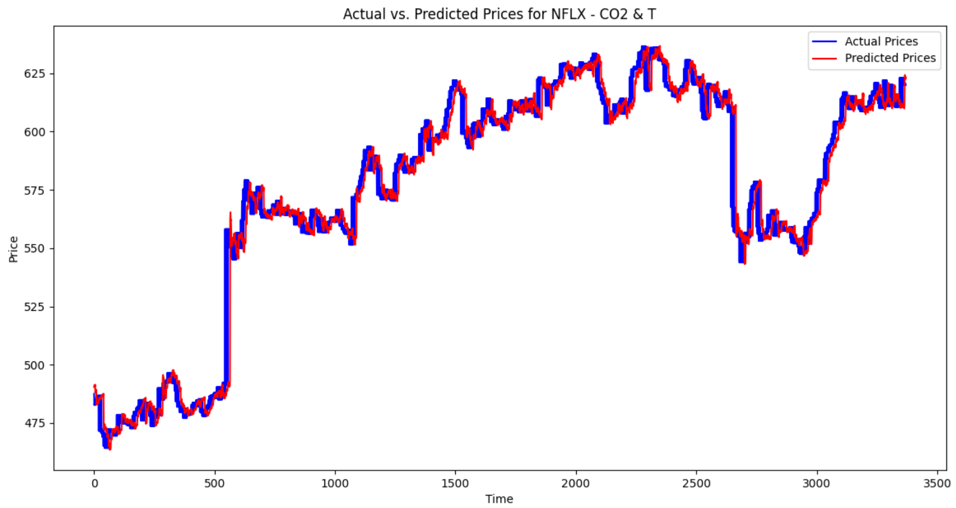
NFLX - ESG (Histogram)



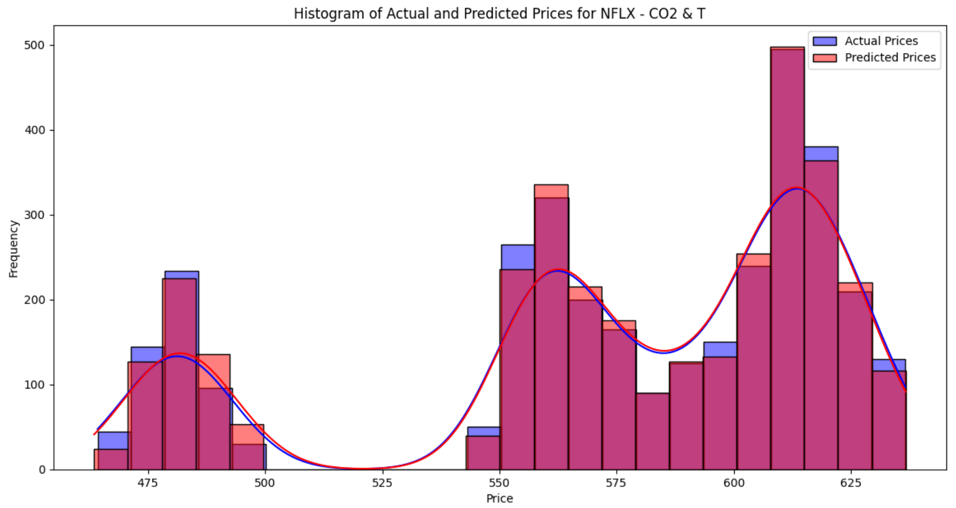
NFLX - ESG (Boxplot)



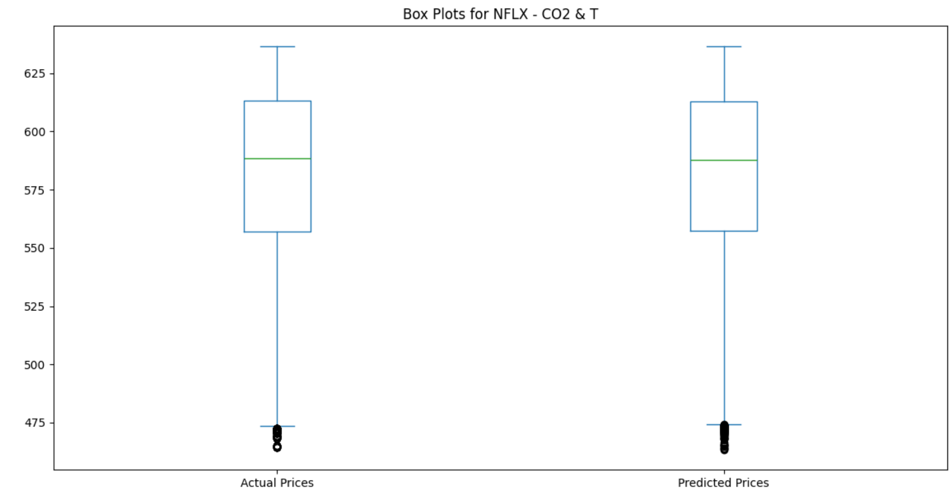
NFLX - CO2 & T (Actual vs Predicted)



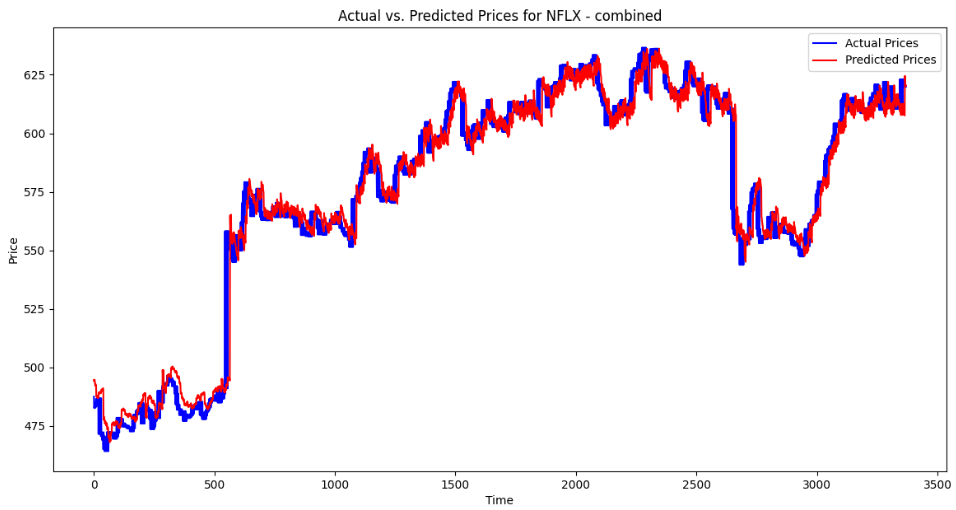
NFLX - CO2 & T (Histogram)



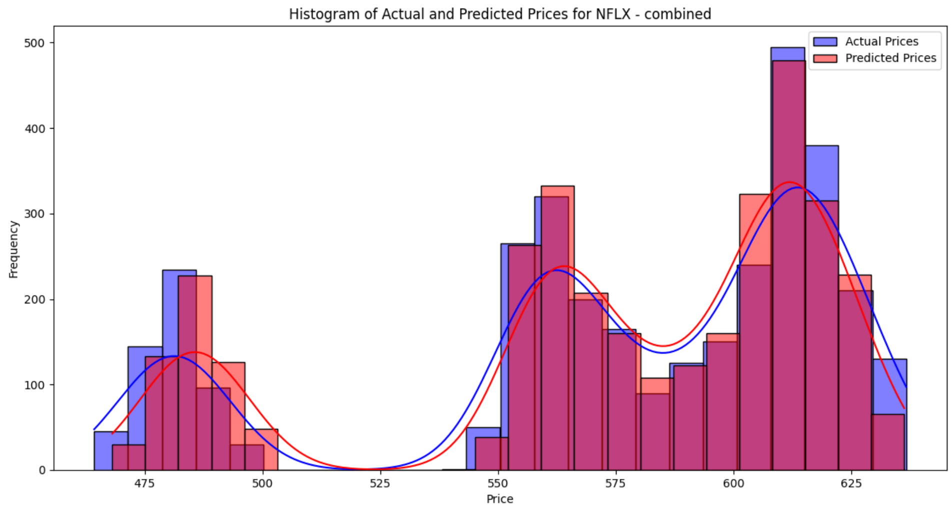
NFLX - CO2 & T (Boxplot)



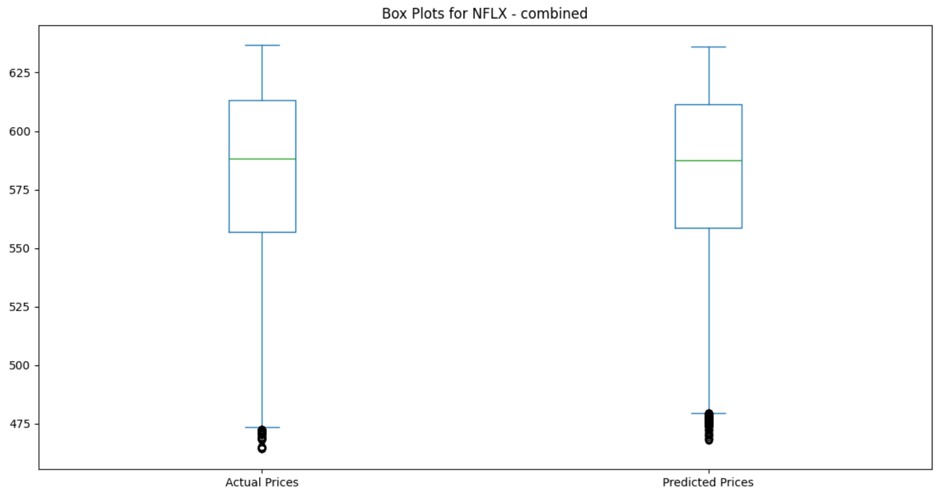
NFLX - combined (Actual vs Predicted)



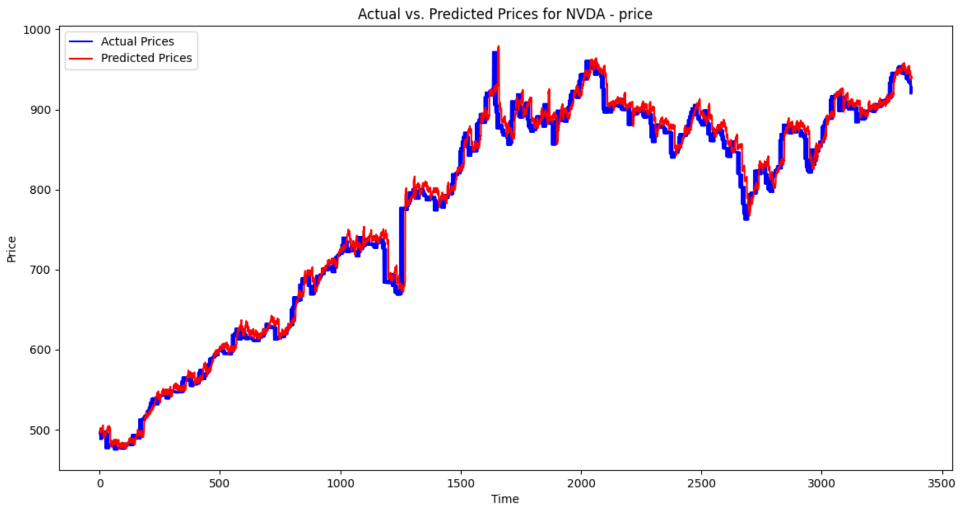
NFLX - combined (Histogram)



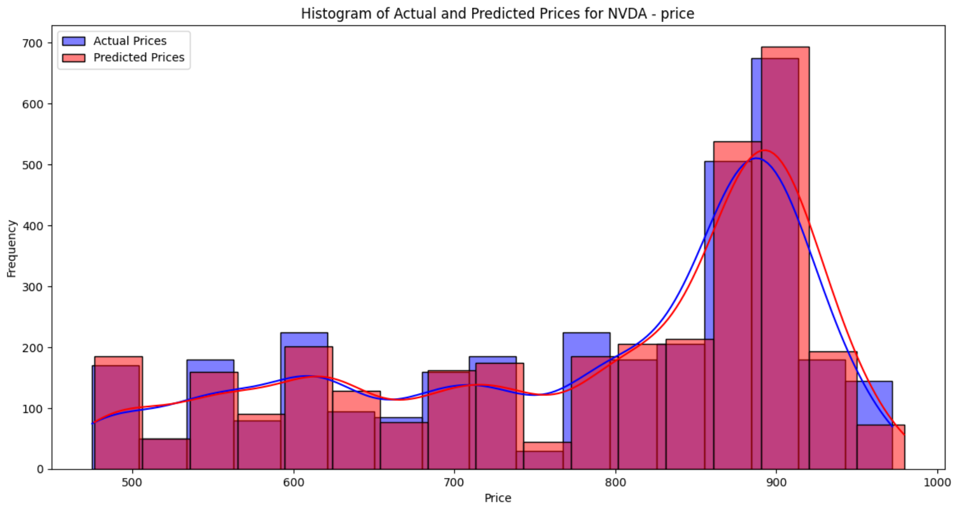
NFLX - combined (Boxplot)



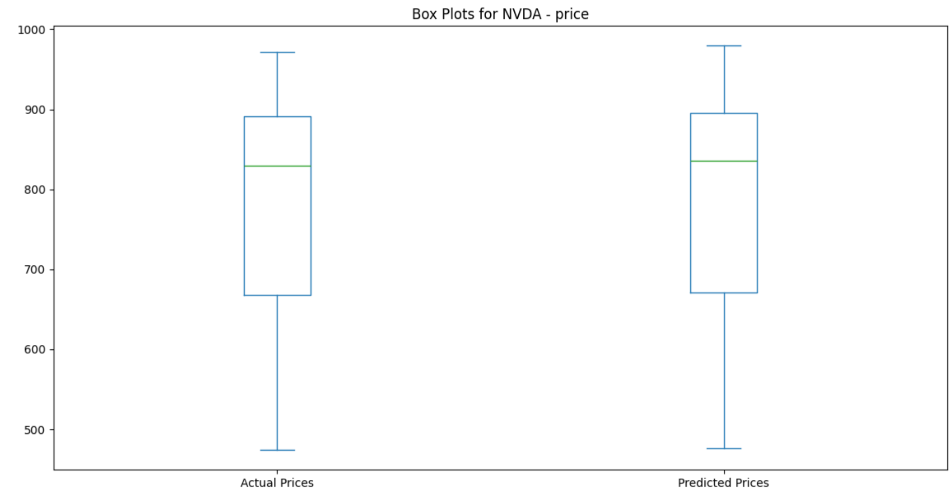
NVDA - price (Actual vs Predicted)



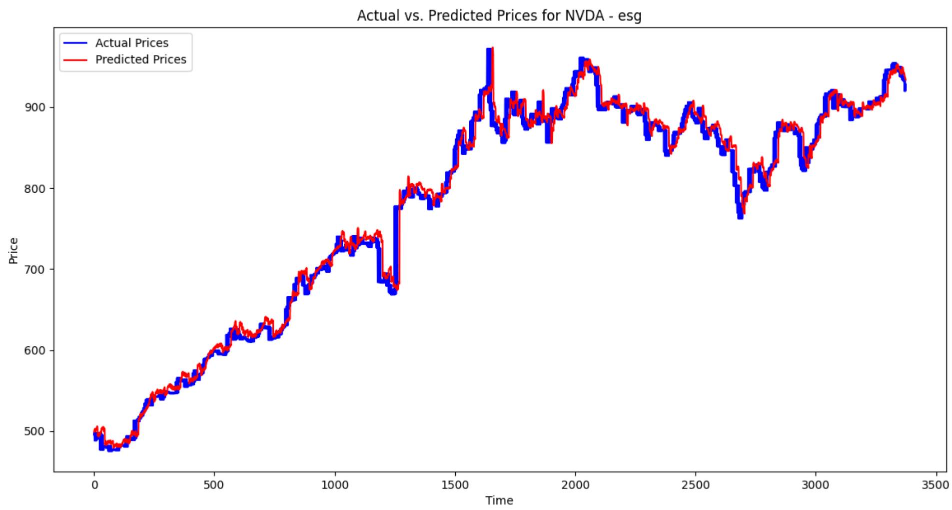
NVDA - price (Histogram)



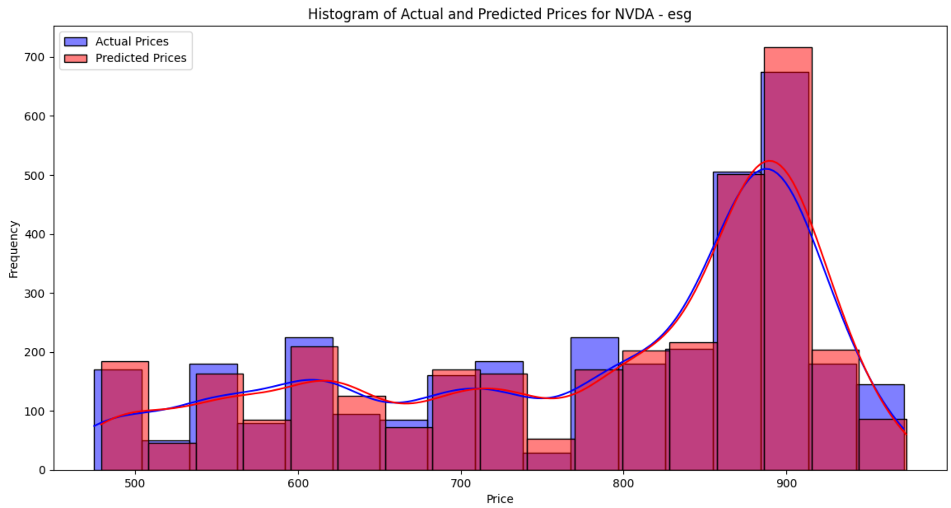
NVDA - price (Boxplot)



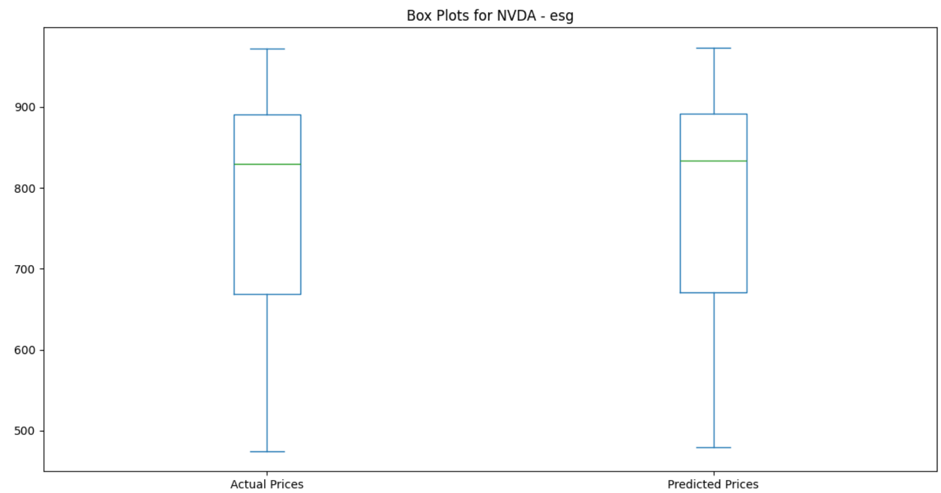
NVDA - ESG (Actual vs Predicted)



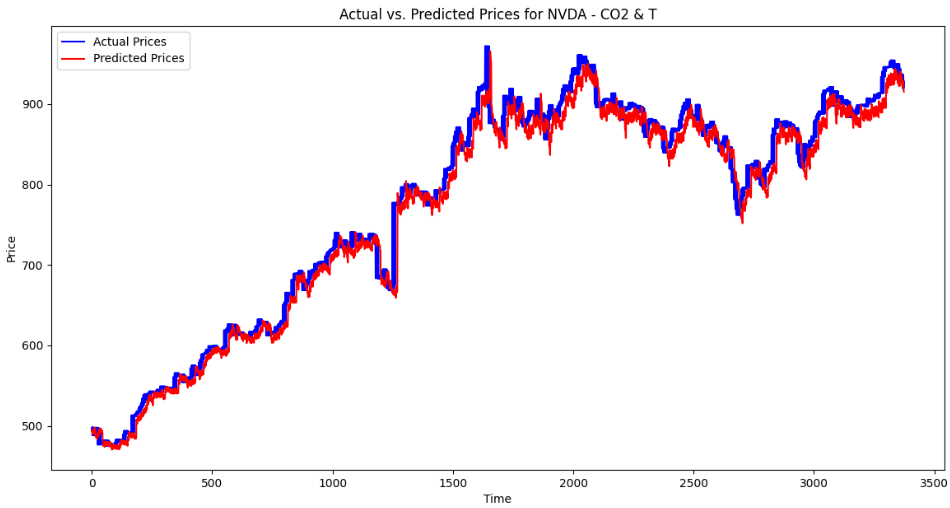
NVDA - ESG (Histogram)



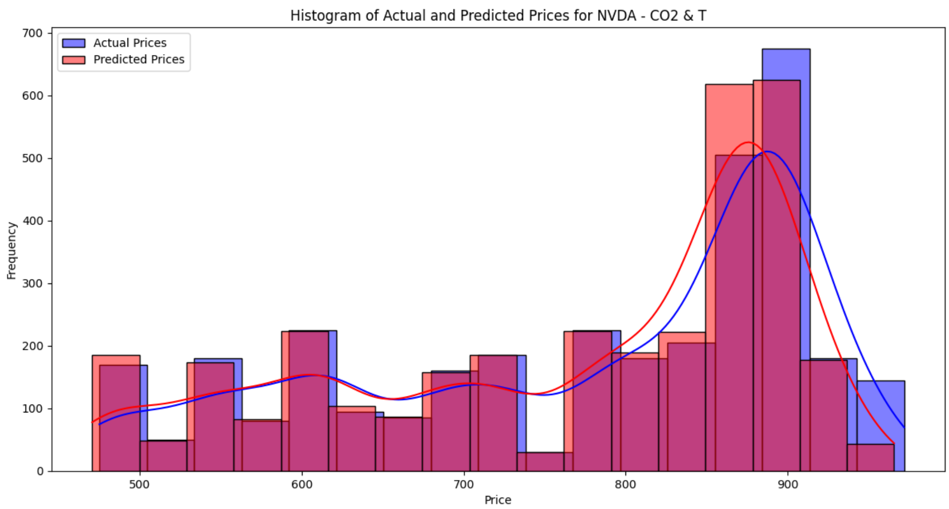
NVDA - ESG (Boxplot)



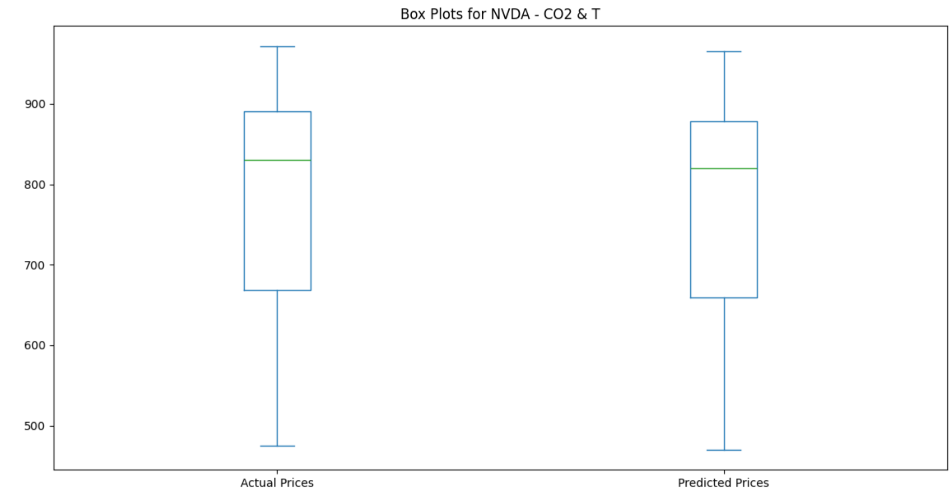
NVDA - CO2 & T (Actual vs Predicted)



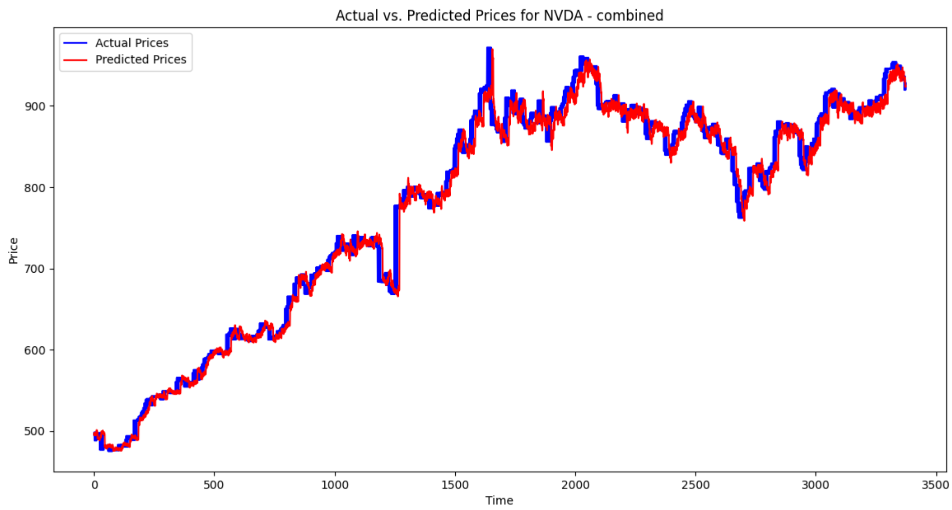
NVDA - CO2 & T (Histogram)



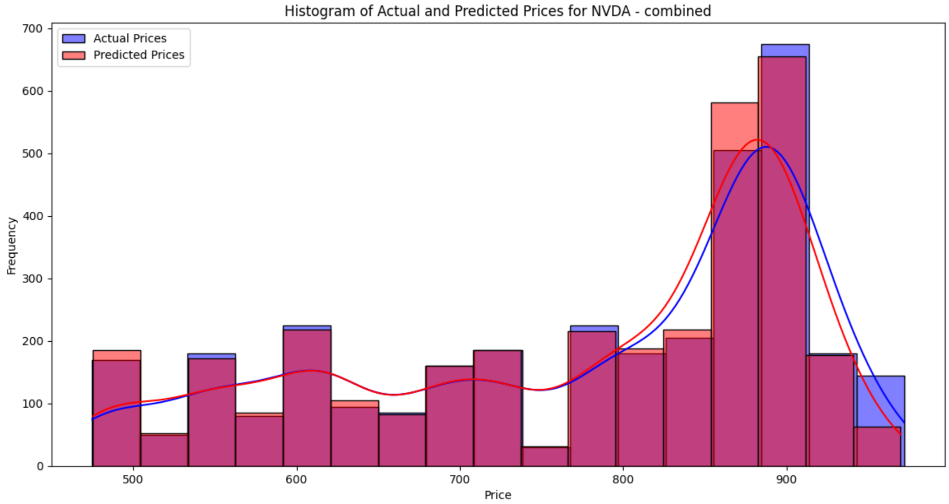
NVDA - CO2 & T (Boxplot)



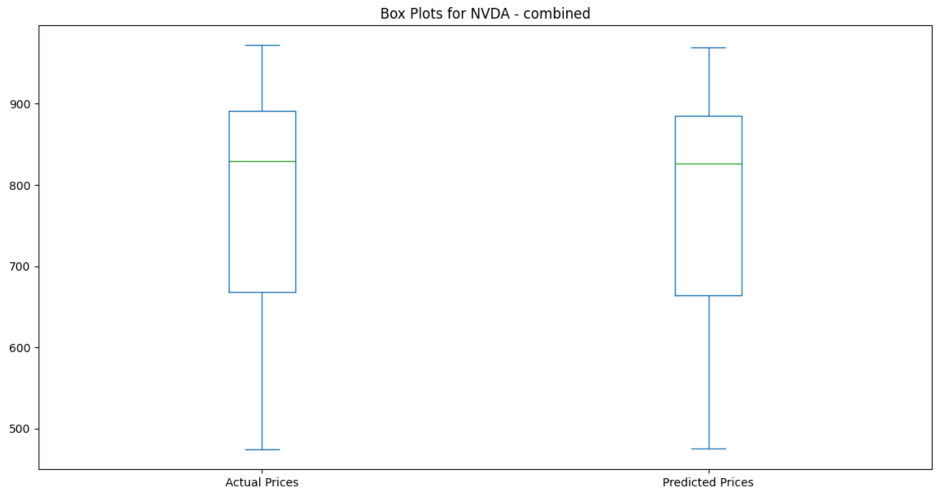
NVDA - combined (Actual vs Predicted)



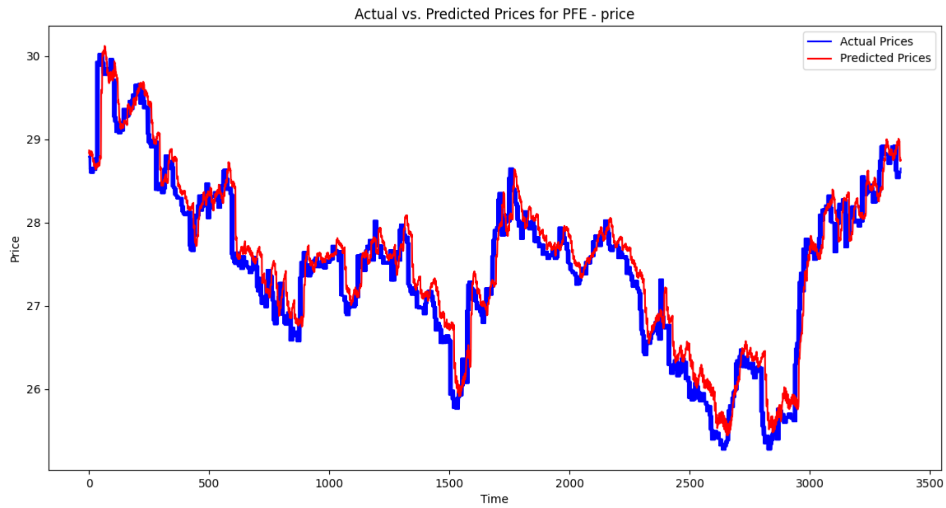
NVDA - combined (Histogram)



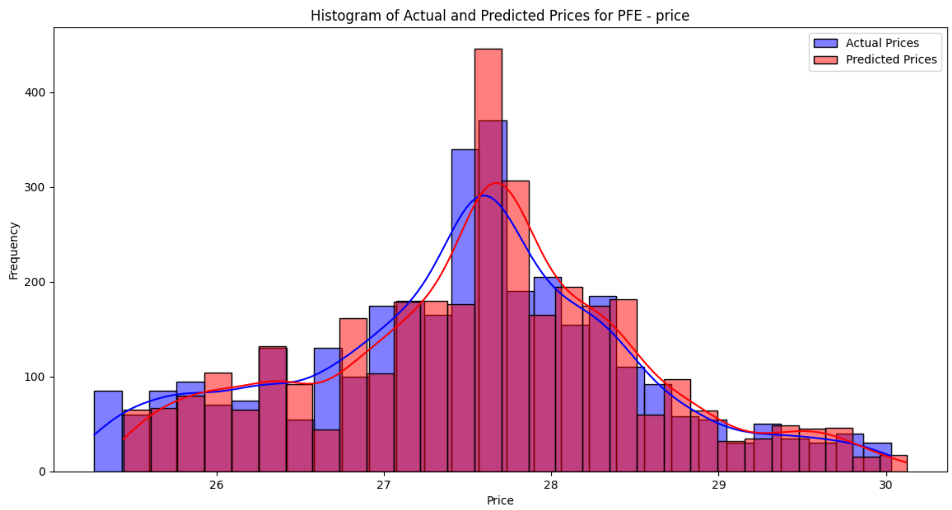
NVDA - combined (Boxplot)



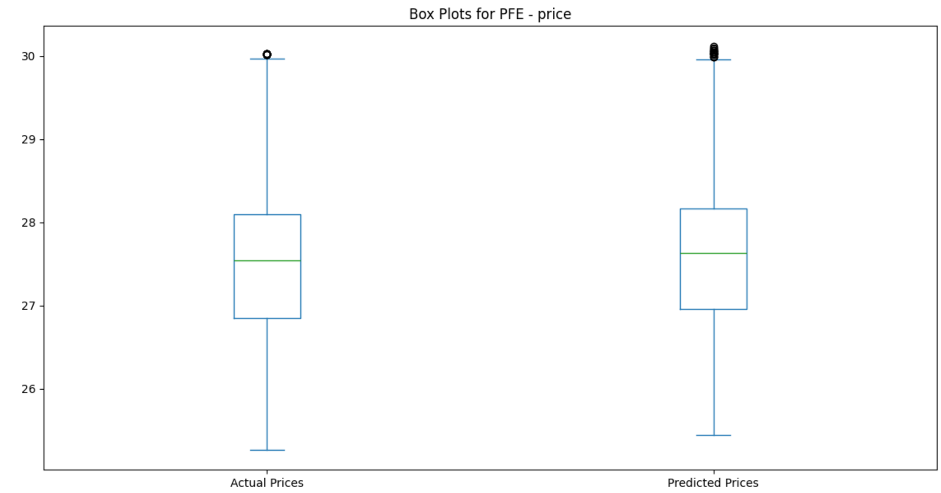
PFE - price (Actual vs Predicted)



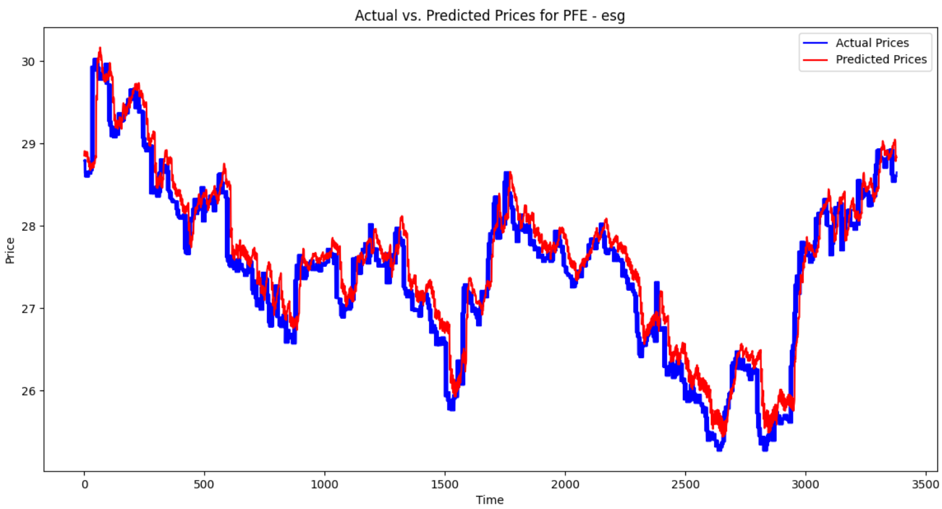
PFE - price (Histogram)



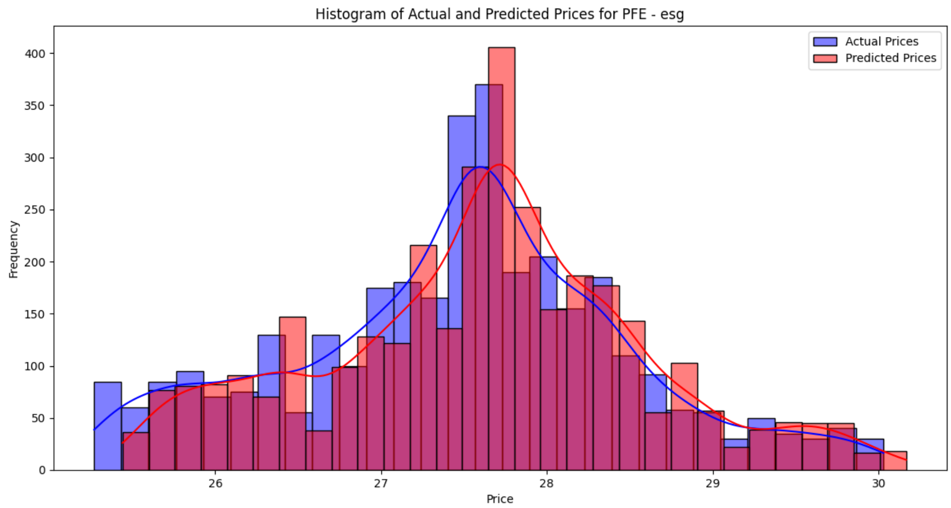
PFE - price (Boxplot)



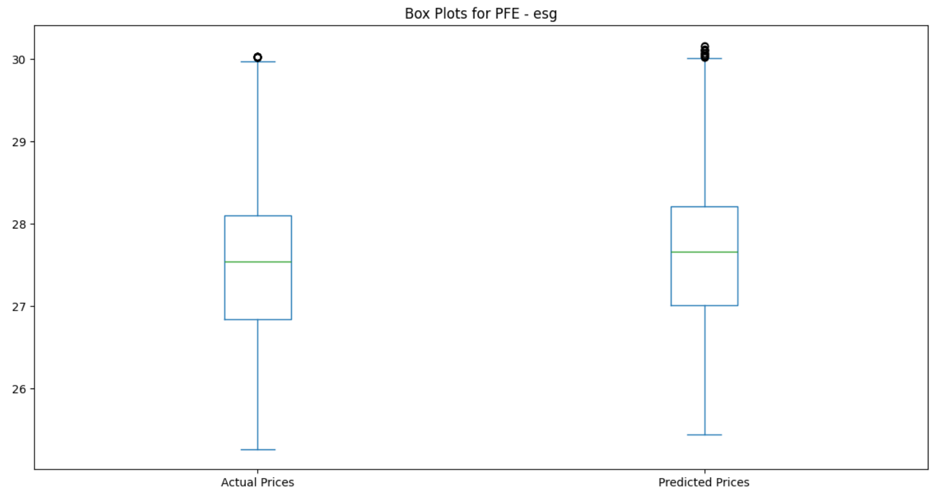
PFE - ESG (Actual vs Predicted)



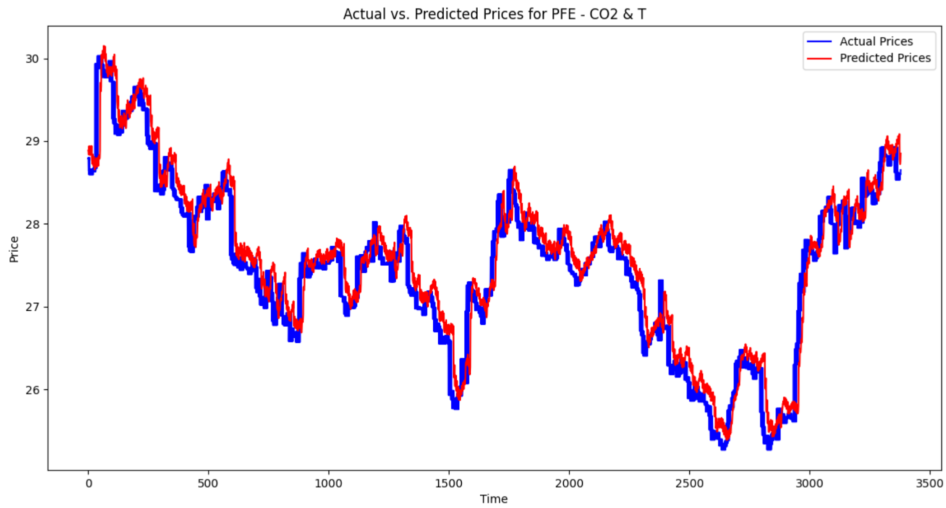
PFE - ESG (Histogram)



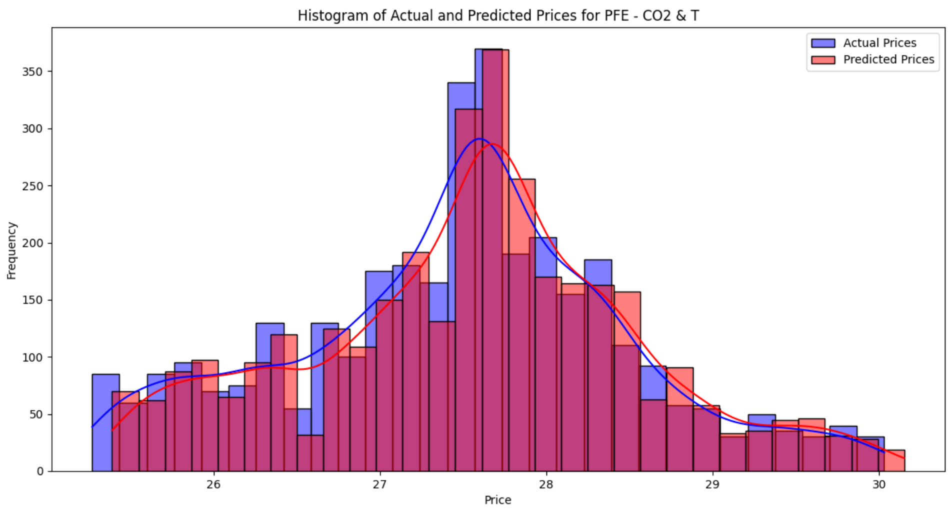
PFE - ESG (Boxplot)



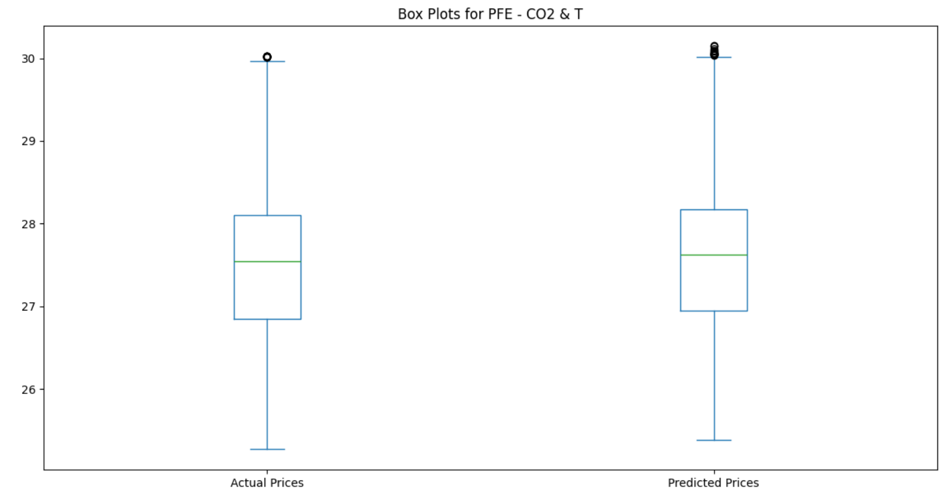
PFE - CO2 & T (Actual vs Predicted)



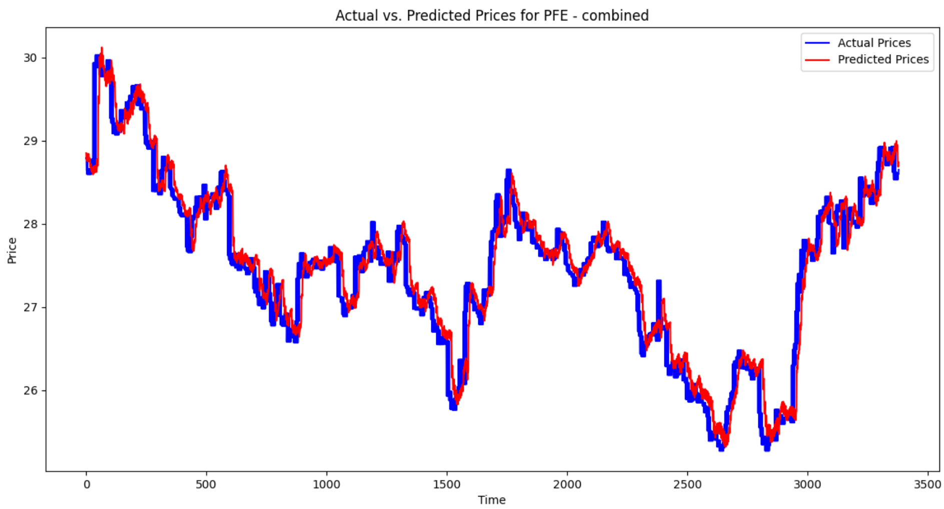
PFE - CO_2 & T (Histogram)



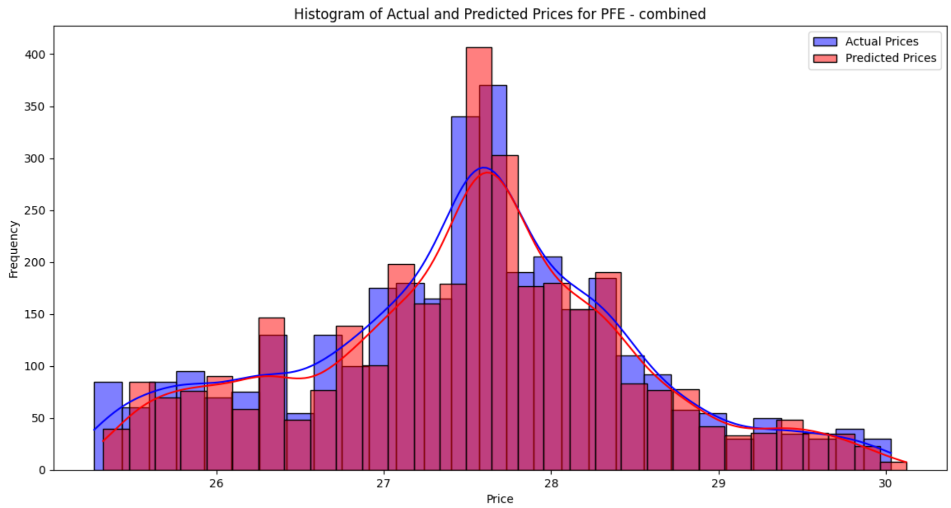
PFE - CO_2 & T (Boxplot)



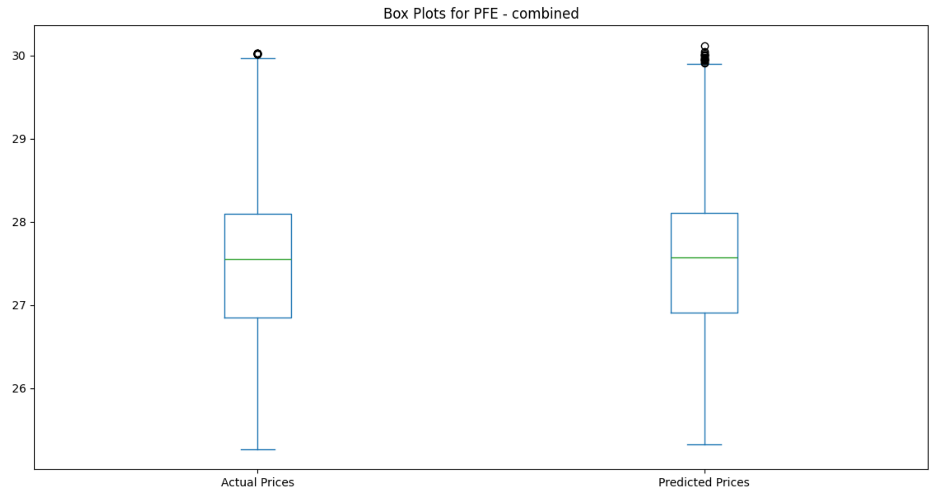
PFE - combined (Actual vs Predicted)



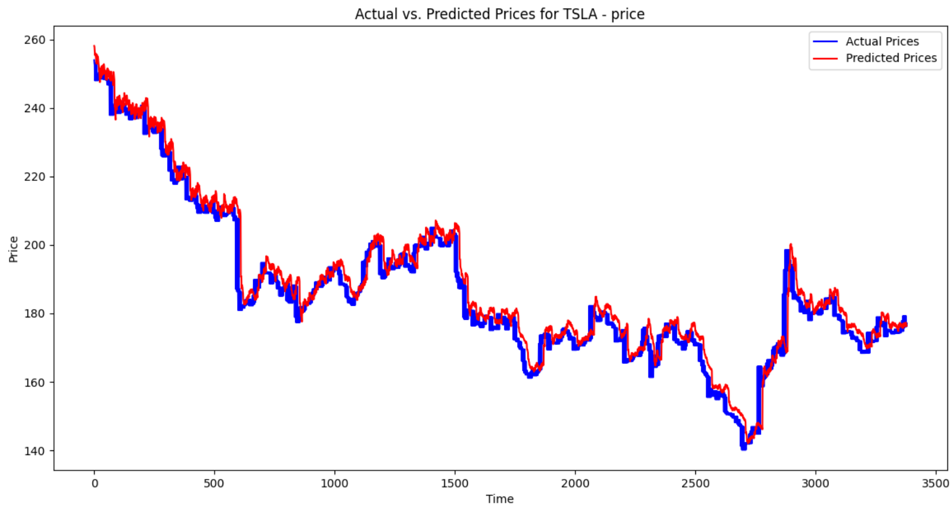
PFE - combined (Histogram)



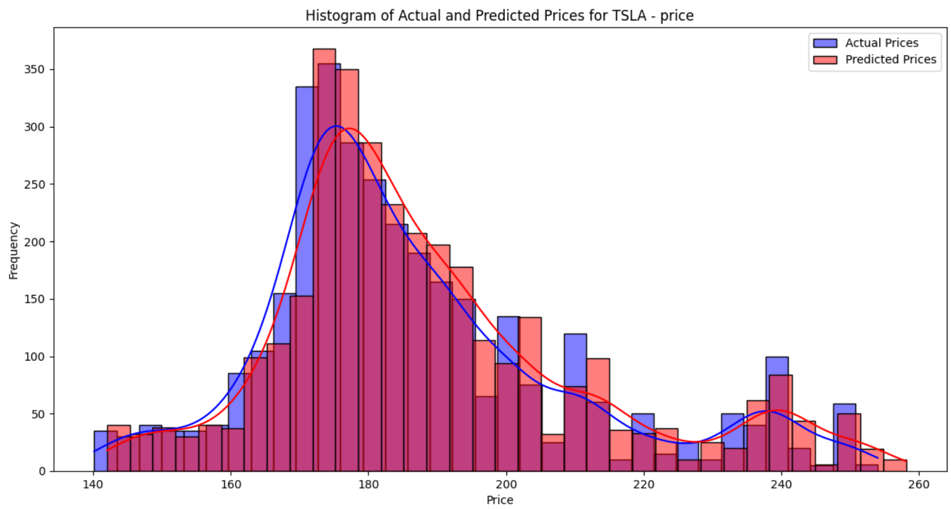
PFE - combined (Boxplot)



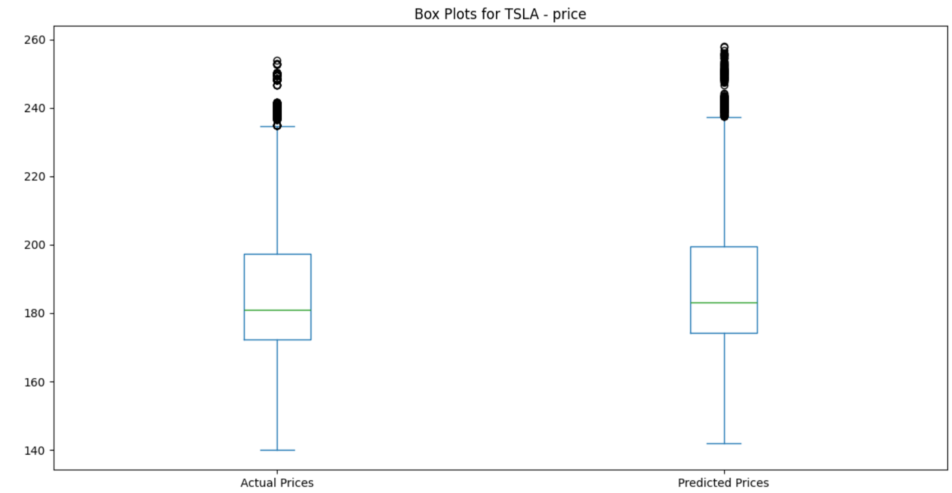
TSLA - price (Actual vs Predicted)



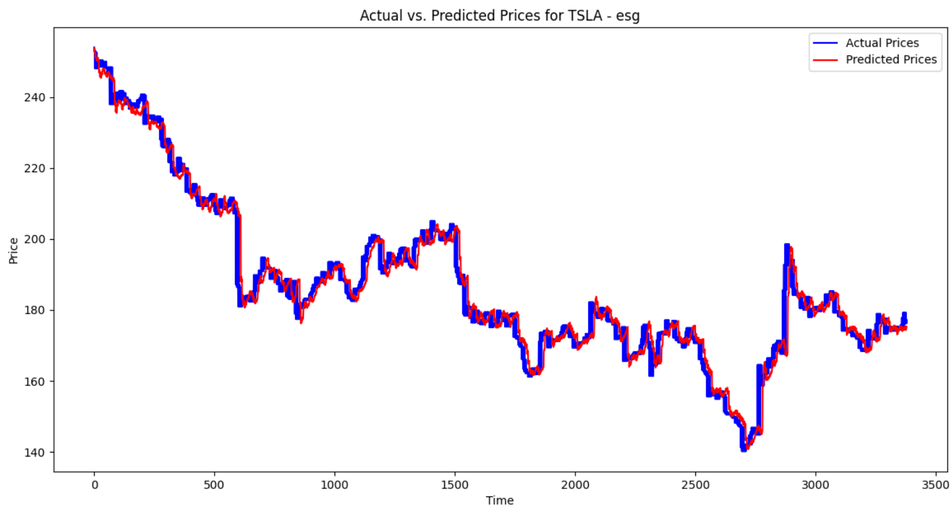
TSLA - price (Histogram)



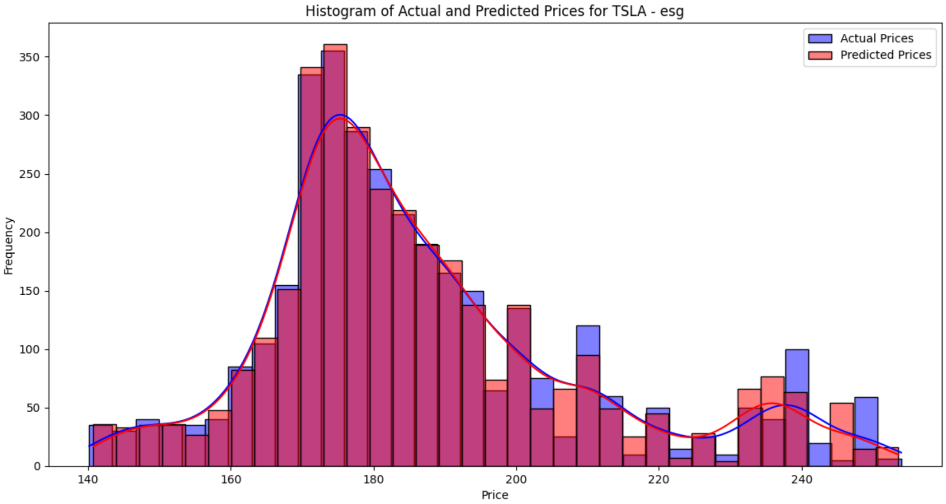
TSLA - price (Boxplot)



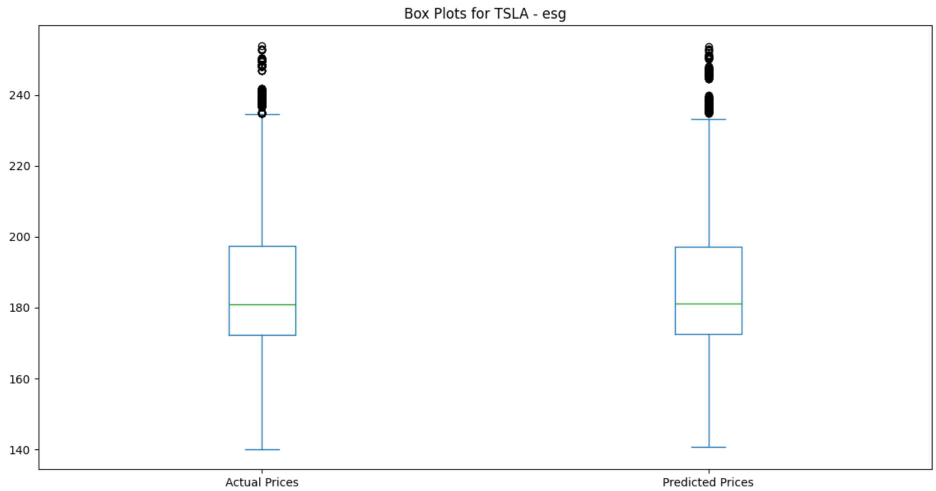
TSLA - ESG (Actual vs Predicted)



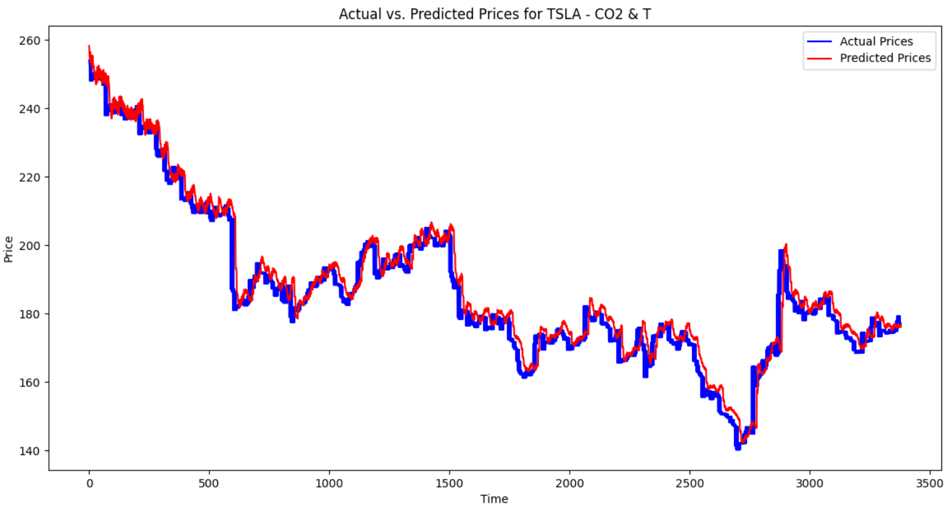
TSLA - ESG (Histogram)



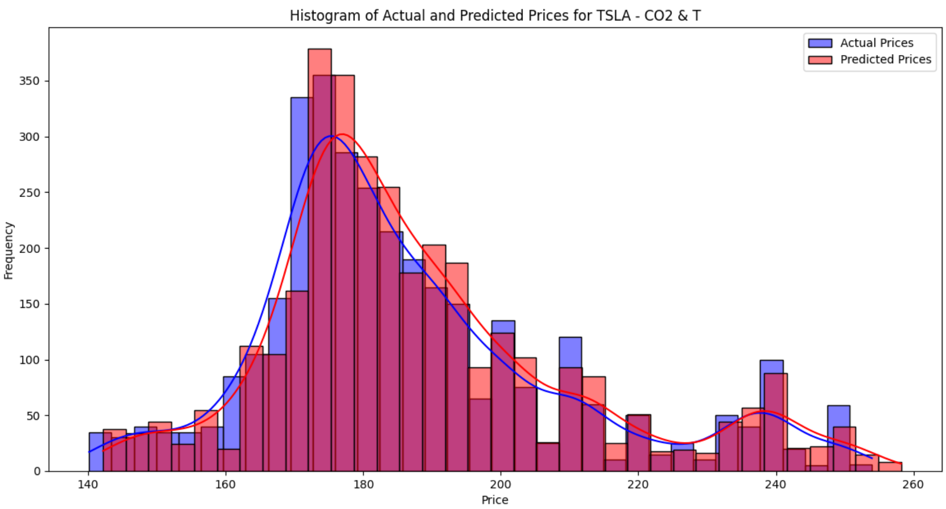
TSLA - ESG (Boxplot)



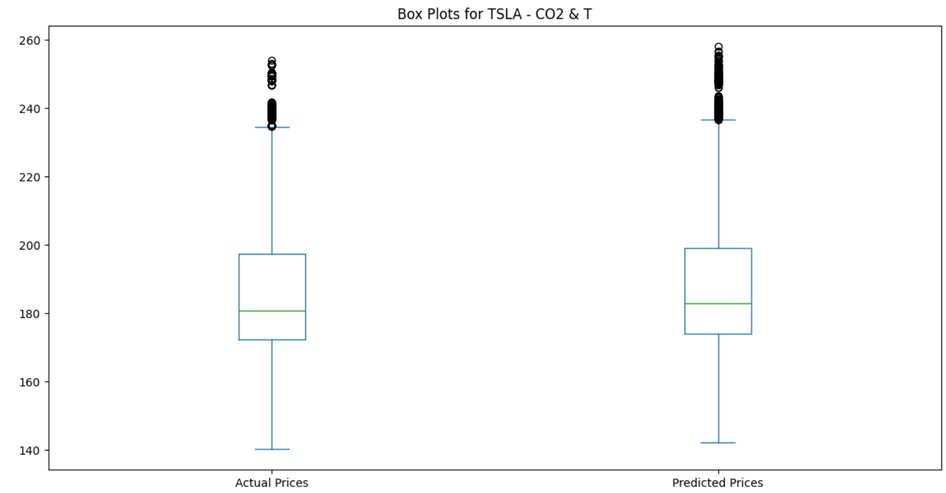
TSLA - CO2 & T (Actual vs Predicted)



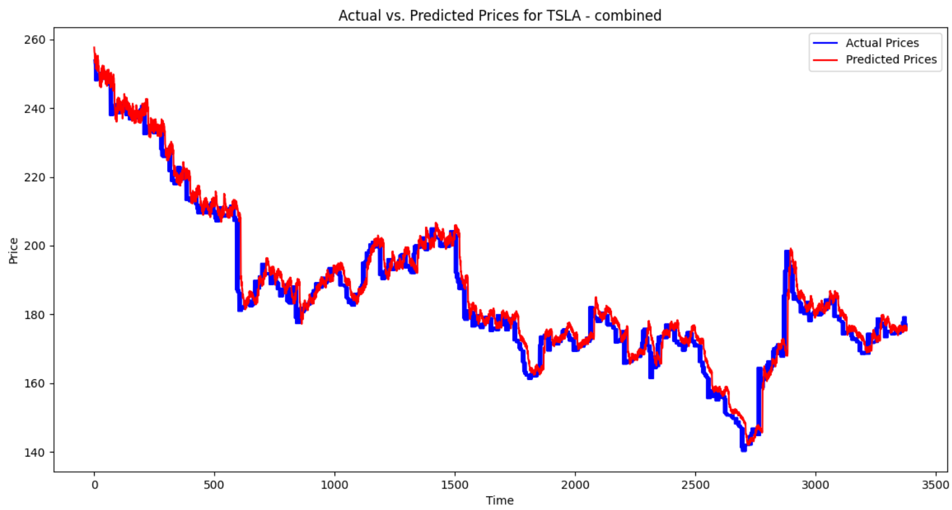
TSLA - CO2 & T (Histogram)



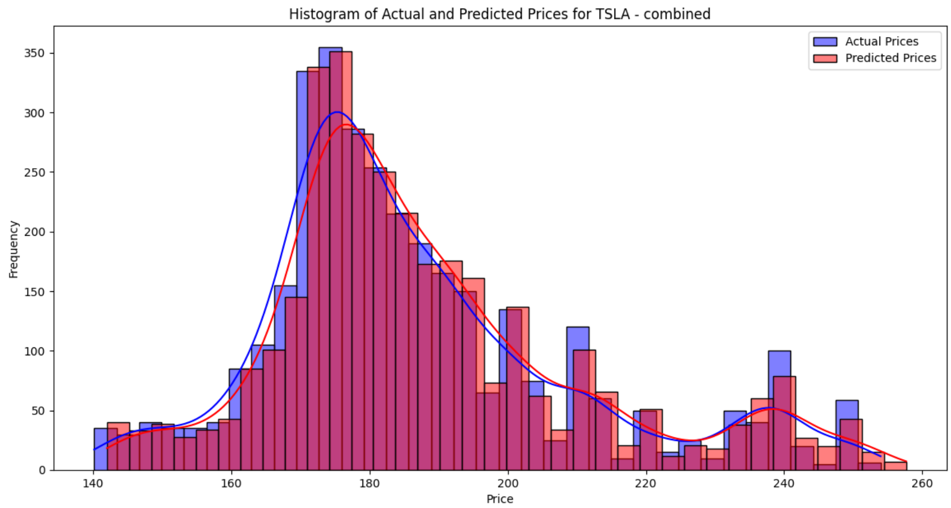
TSLA - CO2 & T (Boxplot)



TSLA - combined (Actual vs Predicted)



TSLA - combined (Histogram)



TSLA - combined (Boxplot)