City, University of London

Data-driven techniques to estimate the growth of small and medium-sized enterprises using ESG metrics and news-based sentiments.

Submitted in partial fulfillment of requirements
for the degree of

Masters of Science - Data Science(FT)(2022-23)

by

Siddharth Madan
Student ID: 220055323

Guide:

Dr. Daniel Chicharro

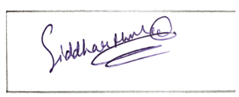
City, University of London

City, University of London

(Department of Computer Science)

DECLARATION

By submitting this work, I declare that this work is entirely my own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the assessment instructions and any other relevant programme and module documentation. In submitting this work I acknowledge that I have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. I also acknowledge that this work will be subject to a variety of checks for academic misconduct



Siddharth Madan
Student ID: 220055323

Date:10th December, 2023
Place: London, United Kingdom

Abstract

This research work explores three very different questions from one another and yet all of those are aligned for the interest of investors and fiduciaries. Individuals or institutions looking to invest in small and medium-sized enterprises might get perplexed by the limited research-based methods to resort to in order to rely on and eventually utilise them to their benefit and the benefit of society in general. Environmental, Social, and Governance conscientious decision-making trend is picking up worldwide and our work is going to provide insight into the popular interest of many in the modern day. We have shown the possibility of ESG ratings given by standard rating agencies to be considered as a serious metric to gauge the growth scope of a firm.

Apart from this, we will discuss the prospects of a conventional time series prediction model like ARIMA and its relevance with a limited amount of data and its prediction capabilities and limitations. This would again enable the investors to make an educated investment decision.

Alongside both these aspects we will also be studying the implementation of sentiment analysis on news articles about the firms in which investors are keen on knowing. This would serve as another great piece of insight into the prospects of growth of a company.

Contents

List of Figures

Chapter 1

Introduction

This chapter aims to present a brief introduction to the background and motivation behind this research endeavour. We would also elaborate on the beneficiaries of this work, specifics of the objectives and goals, our work plan, changes to the originally proposed project, and the report structure as a whole.

1.1 Motivation and Background

Environmental, Social, and Governance (ESG) investing has become an increasingly important concept in the financial world, reflecting a growing recognition that investment decisions can have broad implications beyond simple financial returns. This section delves into the various facets of ESG investing, its relationship with traditional investment theories, the financial implications, and the challenges and prospects it poses for investors and fiduciaries.

ESG investing, an evolution of socially responsible investing (SRI), has roots stretching back several decades. Initially, SRI focused on excluding investments in sectors like tobacco or weapons. Over time, this evolved into a more comprehensive approach that considers a wide range of environmental, social, and governance issues. Events such as the global financial crisis of 2008 and the rising awareness of climate change have further propelled ESG into mainstream investing.

In ESG investing, the environmental aspect examines a company's impact on the Earth, including its carbon footprint and waste management. The social component considers how the company treats people, including labor practices and community engagement. Governance involves the company's leadership, executive pay, audits, and shareholder rights. Investors increasingly recognize that companies adhering to robust ESG standards can mitigate risks and potentially enhance long-term returns.

Modern Portfolio Theory (MPT) posits that investment returns should be evaluated relative to their risk, with diversification as a key element. ESG investing often grapples with the balance between adhering to ESG criteria and maintaining diversification. The theory suggests that constraining the investment pool based on ESG criteria might limit opportunities, potentially impacting returns (Malkiel, 1984) [1].

The financial performance of ESG investments has been a subject of much debate. Some studies suggest that ESG investments can match or even outperform traditional investments. Key performance indicators include absolute returns, risk-adjusted returns, and the Sharpe ratio. However, the varied methodology in these studies makes it challenging to draw definitive conclusions (Busch, et. al, 2015)[2].

ESG investing often incurs additional costs, such as higher management fees and the expense of thorough ESG screening. Measuring the social and governance impact is complex and can involve significant costs in data collection and analysis. Additionally, ESG criteria can limit investment opportunities, potentially leading to lower diversification and higher risk.

Fiduciaries, tasked with managing investments on behalf of others, face unique challenges in ESG investing. They must balance the pursuit of financial returns with the achievement of social goals, often within the framework of complex legal and ethical considerations. This balance can be particularly challenging when clients have specific ESG preferences that may not align with optimal financial returns.

The future of ESG investing appears promising, driven by increasing investor awareness and regulatory changes. Technological advancements, such as AI and data analytics, are improving the ability to assess and integrate ESG factors into investment decisions. However, the evolving nature of ESG criteria and the dynamic regulatory landscape pose

Figure 1.1: This schematic diagram showcases some of the key aspects with the framework of the Environmental, Social, and Governance umbrella to be considered by rating agencies in order to further produce their own ESG ratings upon standardisation and normalisation of all the parameters.

ongoing challenges.

Various case studies illustrate the successes and challenges of ESG investing. For instance, investment funds that prioritized renewable energy have seen significant growth, aligning financial returns with environmental impact. Conversely, some ESG-focused funds have struggled, highlighting the complexity of integrating ESG criteria into investment strategies effectively (CFA Institute and PRI, 2018)[3].

ESG investing represents a paradigm shift in the investment world, intertwining financial performance with broader societal impacts. While it poses challenges in terms of costs,

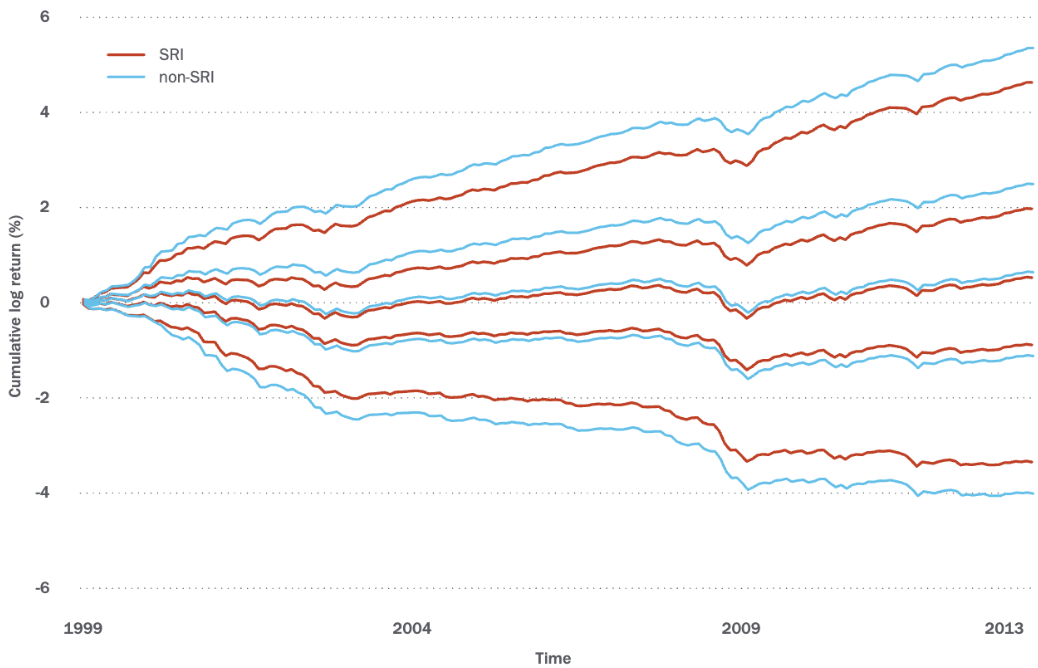


Figure 1.2: This plot is from (Zu, Thomas and Zvingelis, 2014)[4], and is a log distribution of total fund returns and suggests that non-SRI has outperformed (but not by much!) SRIs for an investment period of 15 years.

measurement, and risk management, it also offers opportunities for investors to align their financial goals with their values. As the world continues to grapple with environmental and social challenges, ESG investing will likely play an increasingly important role in shaping corporate behaviors and investment strategies. The key for investors and fiduciaries lies in striking a balance between achieving desirable social outcomes and ensuring robust financial returns.

Authors in (Zu, Thomas, and Zvingelis, 2014) [4], analyse Socially Responsible Investment (SRI) funds and compare them to conventional non-SRI funds, revealing intriguing insights into the financial performance of socially conscious investing. A key finding is the similarity in average performance between SRI and non-SRI funds, challenging the notion that ethical investment strategies compromise financial returns. As seen in Figure 1.2, the non-SRI funds (in blue) outperform the SRI funds but not by much, and also, another plot from (Zu, Thomas and Zvingelis, 2014)[4] extends its analysis to cross-sectional return distributions, uncovering that SRI funds typically demonstrate a more concentrated return profile around the median. This suggests a more predictable range of outcomes for SRI funds, indicating hardly any difference in the performance of SRIs and non-SRIs at the 5th decile (which is the median) and hence, alluding that the Good and the Gold may go together. Interestingly, in scenarios of below-median returns, particularly during bear markets, SRI funds tend to outperform their non-SRI counterparts.

Further, the study investigates the potential reasons behind the performance differences, hypothesizing that the distinct investment universes accessible to SRI and non-SRI funds could be a factor. This hypothesis is tested, and it's found that the SRI holdings universe exhibits greater homogeneity in risk factor exposures compared to non-SRI holdings, providing a possible explanation for the observed performance patterns.

To summarise, this paper emphasizes the significant and persistent differences in cross-sectional performance between SRI and non-SRI funds. These differences are particularly notable in terms of total return, risk-adjusted return, and risk exposures. This comprehensive analysis is crucial for investment professionals seeking to understand how socially responsible investing compares with traditional investment strategies, not just in average performance metrics, but also in terms of investment universe characteristics and risk factor exposures.

The research article "Ownership of ESG Characteristics" (Bateman and Goldberg, 2023)[5], provides an in-depth exploration of the concept of ownership of Environmental, Social, and Governance (ESG) characteristics within the framework of socially responsible investing. The authors not only provide a methodological guide but also an ethical inquiry into the implications of owning ESG characteristics in investment portfolios i.e. the concept of ownership in SRI is expanded beyond financial stakes to include the ethical and practical aspects of a company's activities. The article also elaborates that investors, by virtue of their investments, own a share in the ESG attributes of the companies. This ownership encompasses a wide range of ESG factors, from carbon emissions and board diversity to human rights policies.

This article introduces methodologies to quantify an investor's ownership of different ESG characteristics. These include metrics for physical characteristics like carbon emissions, ratios such as the percentage of women on corporate boards, and binary characteristics like the presence or absence of a human rights policy. The approach involves aggregating company-specific ESG data to the portfolio level and comparing it against a benchmark to ascertain active ownership.

The authors delve into the ethical dimensions of owning ESG characteristics, suggesting that investors bear some responsibility for the activities of the companies they invest in. This responsibility could manifest in various ways, such as shareholder activism, where investors use their stake to influence company behavior, or through decisions to invest in companies with positive ESG characteristics.

A critical distinction made in the paper is between ownership and impact. Ownership reflects an investor's share in the ESG characteristics of a company, while impact refers to the ability to influence a company's policies or practices. The study highlights that ownership is a zero-sum game—when one investor divests, another acquires that ownership, whereas impact can lead to real changes in company behavior.

The article provides practical examples and scenarios to illustrate the application of their equations for measuring ownership. These examples help to understand the practical aspects of ESG ownership and its potential implications for investment strategies. The authors include a case study of a hypothetical investment in the S&P 500 Fossil Fuel Free Index to demonstrate how these methodologies can be applied in realistic settings.

The distinction between ownership and impact is particularly insightful, underscoring the need for investors to actively engage in influencing corporate behaviors for positive change. This work not only provides tools and methodologies for investors to assess their ethical responsibilities but also encourages a deeper understanding of the broader impact of their investments beyond financial returns.

All the discussion above essentially establishes the importance of ESG characteristics in any portfolio and the implications of using it as a criterion for the investors or their fiduciaries. This now takes us to the next step of using this as a criterion by quantifying the ESG characteristics of a firm and quantifying it in a way that can be standardised and normalised for comparing various companies.

Well, this task is commonly undertaken by popular financial news media outlets such as Bloomberg, Morningstar, MSCI, etc. Each of these rating agencies has its own rating criteria and mostly provides the ratings for publicly listed companies and not for unlisted small and medium-sized enterprises (SMEs). This brought us to explore some eccentric ways to acquire ESG performance for SMEs. We have discussed more about the datasets that we have used and the ratings we have chosen to consider for the purpose of this study in the subsections to follow.

In this research project, we have also considered the analysis of sentiment about certain small and medium-sized companies from online news articles. Sentiment analysis of companies from news articles is an increasingly vital tool in the modern business landscape.

One of the primary applications of sentiment analysis of companies in news articles is in the field of market prediction and investment strategies. News articles often contain vital information that can influence investor perceptions and market movements. (Loughran and McDonald, 2011)[6] in their work on financial sentiment analysis demonstrate how the tone of financial news can significantly impact stock prices and trading volumes. Furthermore, (Tetlock, 2007)[7] highlights that quantifying the tone of news stories can predict stock market fluctuations. By analyzing sentiments in news articles, investors and analysts can gauge public opinion and market trends, thereby making more informed investment decisions.

Sentiment analysis is also crucial in brand management and public relations. News articles often shape the public perception of a company. The work of Pang and Lee (2008)[8] on opinion mining and sentiment analysis underscores how understanding public sentiment through news analysis can inform companies about their brand image and reception in the market. This is particularly relevant in crisis management, as identified by Coombs (2007)[9], who notes the importance of monitoring news media to manage and mitigate negative publicity effectively.

Another key application of sentiment analysis is in strategic decision-making and competitive analysis. Companies can use sentiment analysis to monitor not only their own public perception but also that of their competitors. Authors (Brynjolfsson, Hitt, and Kim, 2011)[10] in their study on data-driven decision-making indicate how companies leveraging big data, including news sentiment, gain competitive advantages. Understanding the market positioning and strategies of competitors through sentiment analysis can provide valuable insights for strategic planning.

Sentiment analysis of news articles can provide companies with insights into consumer needs and market gaps, thus informing innovation and product development. As Davenport and Harris (2007)[11] argue in their work on competitive strategy and business analytics, companies that effectively utilize data (including sentiment analysis) can innovate more effectively and tailor their products and services to meet specific consumer demands.

Additionally, sentiment analysis is crucial in risk management. By analyzing news sentiments, companies can identify potential risks and threats in their operating environment. As noted by Aven (2016)[12] in his work on risk assessment and decision analysis, understanding the sentiment and tone of news articles can help in anticipating potential crises and planning appropriate risk management strategies.

Sentiment analysis of companies from news articles plays a crucial role in various aspects of business operations and strategy. From market prediction and investment strategies to brand management, and from strategic decision-making to innovation and risk management, the insights derived from sentiment analysis are invaluable. The ability to effectively analyze and interpret these sentiments can provide investors with a significant competitive edge in today's data-driven business environment.

1.2 Beneficiaries

The research project will enable investment firms and fund managers to make educated decisions that will put their companies and individuals in the best possible position by comparing and committing to this procedure. This research project has the potential to yield real benefits for SMEs by enabling them to make data-driven decisions, lower risks, increase competitiveness, and align with ESG issues. These findings could lead to improved operational efficiency, stable finances, and long-term growth for SMEs across a range of industries.

Many asset managers, pension funds, and investment firms that use performance-based growth indicators and ESG scores to rank companies as part of their investment strategy would be interested in the task of accurately predicting growth and ESG ratings of a company that has not been publicly listed. SMEs can use cutting-edge machine learning algorithms to gain a competitive edge in their respective industries. Precise performance projections aid businesses in optimising processes, adapting to dynamic market circumstances, and preserving a competitive advantage. With the help of this research, SMEs will be able to outperform their rivals and make data-driven decisions.

1.3 Objectives

There are three major objectives of this research endeavour:

1. **Can we predict the growth of a company using the ESG ratings assigned to those companies by some standard rating agency?** Predicting the growth of a company using its ESG performance as a firm as criteria is a topic of growing interest in the field of sustainable finance and investment. Our work is based on the premise that ESG ratings, which are increasingly recognized as significant indicators of a company's operational and financial performance, can offer insights into its future growth potential.

2. **How can sentiment analysis of the online news articles about those companies provide additional information for the investors to make an educated investment decision?** The prospect of using sentiment analysis of online news articles as a predictor of the growth of SMEs is a compelling concept for us to explore, drawing from the fields of data science, finance, and business analysis. This idea hinges on the assumption that public sentiment, as reflected in news media, can provide valuable insights into the potential growth trajectories of SMEs.

3. **Possibility of providing yet another additional information about the growth projections based on popular and tested methods like ARIMA?** The Autore-

gressive Integrated Moving Average (ARIMA) model is a popular statistical approach used for time series forecasting, which is used for predicting a company's turnover based on its historical data. It effectively captures and models various patterns in time series data such as trends, cycles, and seasonality, which are common in company turnover data.

1.4 Work Plan

The work plan is majorly same as the original plan in the proposal (in Appendix), with some minor changes, and will be discussed further. The literature review was more extensive than initially planned because the scope of this research encapsulates a very broadly discussed topic for the past decade or so, which made it imperative for us to scale the bounds of our approach. The papers that we referred to during our literature survey will be referenced throughout this report as per the necessity of context.

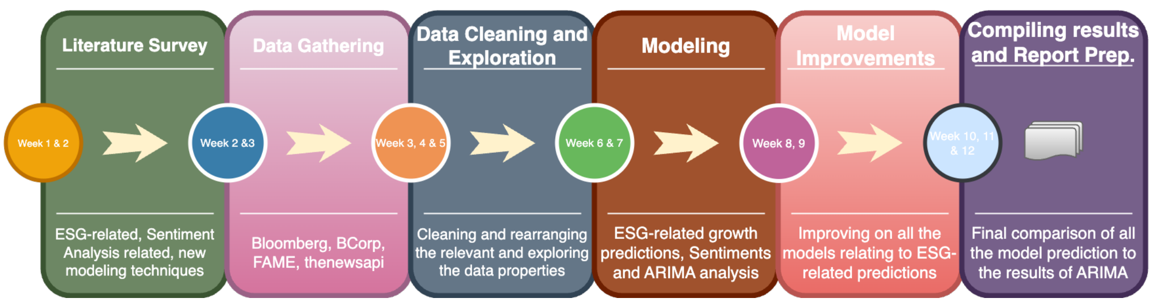


Figure 1.3: This schematic diagram showcases the work flow through the course of the project duration.

In the next step, we emphasised all our efforts in gathering credible data on both the ESG and financial performances of SMEs. This was particularly a challenge and we proposed in our project proposal to take the aid of transfer learning in order to generate the ESG ratings for the SMEs in consideration, it was realised during the period of literature review that it would end up in misalignment of our main goal i.e. checking the predictability of some neural networks and machine learning algorithms for growth of certain companies when provided with their ESG ratings. We are grateful to Equitably.AI for providing a way out by providing us with two datasets i.e. data from the Bloomberg terminal for over two thousand listed companies which included their financial performance and the

ESG ratings. They also provided us with an ESG ratings dataset for approximately four thousand SMEs which were certified by BCorp. Equitably.AI encouragingly provided this dataset for academic purposes and they would be keen to learn from the results of our research work. Next, we targeted only UK-based SMEs for further analysis and acquired the financial performance for approx. fifty SMEs from the FAME database using access to it from City's online databases. For the sentiment analysis of the online news for these fifty targeted SMEs we used free API from 'thenewsapi' in order to gather the relevant articles for all these companies.

In the next step as seen in Figure 1.3, we performed data cleaning and dimensionality reduction in order to process and make our data ready to be fed into neural networks by zeroing in on features and the target variable to be considered for our purpose here, more on this will be discussed in the sections to follow.

In the next two steps, we focused on building a bunch of logical and working algorithms for all the three objectives that we've mentioned in the subsection above. Improvements were made by trying to improve the performance of ESG-related prediction algorithms.

In the final step we compiled all the results and compared our ARIMA predictions to the predictions of both our other approaches i.e. ESG-related growth predictions and growth predictions using sentiment analysis from news article sentiments.

1.5 Changes to the project

As mentioned in the previous section, it was discovered during the literature review phase that using transfer learning to generate the ESG ratings for SMEs would be in conflict with our primary objective, which was to determine whether certain neural networks and machine learning algorithms could accurately predict the growth of specific companies based on their ESG ratings. We then made a deliberate change in our project methodology from what was previously proposed i.e. instead of generating our own ESG ratings with the aid of transfer learning and using listed companies' ESG ratings to estimate similar ratings for SMEs, we have taken the route of targeting the SMEs with some ESG ratings available i.e. those companies with BCorp membership which in turn publishes the ESG performance of its member companies. And then eventually use these BCorp ESG ratings to predict the growth in the revenue of the SMEs in consideration.

We additionally added Sentiment Analysis of news articles of those SMEs in consideration as a method to predict the growth of those companies.

1.6 Report Structure

Chapter 1: Introduction

Introduction about the concept of Environment, Social, and Governance aspects of any firm. Details about the prospects ESG considerations in investing entails. Discussions about previous works and the scope of socially responsible investments. Defining the explicit objectives for this research project and brief about the beneficiaries of this work. Work plan about the sequencing of all the steps undertaken to accomplish the objectives of this work and also the changes from the original proposed project.

Chapter 2: Context

In this chapter, we describe some more of the research articles we reviewed before the commencement of the project work. We emphasise understanding the context of this research problem even further and highlight some key challenges that could limit our approach.

Chapter 3: Methodology and Execution

This chapter lays out the methodology of each and every step towards executing our project goals. This includes the nuances of gathering datasets from various resources and then processing the data, reducing its dimensions, identifying the flaws in the approach, and then applying improvements.

Chapter 4: Results

This chapter explains the results that we obtained from the methodology followed in the previous chapter. We then compare the results with the growth predictions of our ARIMA model.

Chapter 5: Critical Discussion

In this chapter, we elaborate on the results and their implications in the context of the objectives mentioned in Chapter 1. We then highlight whether the particular objectives were achieved or not.

Chapter 6: Conclusion and Reflection

In this final chapter, we discuss the conclusions of this research endeavour and its implications for the beneficiaries, and its academic relevance.

Chapter 2

Context

In this chapter, we review a few more of the research articles we examined before starting the project. We stress the need to comprehend the context of this study problem even further and point out some significant obstacles that can prevent us from taking particular approaches.

2.1 Introduction

ESG ratings provide an assessment of a company's performance in environmental, social, and governance aspects. Authors in (Rau and Yu, 2023)[13] discuss that these ratings, which emerged in the early 1980s, serve as tools for investors to screen firms based on their ESG performance.

This paper also highlights that there has been a significant increase in the availability of ESG data. This surge is in response to the growing demand for information about companies' ESG practices. Several ESG data providers compile these ratings. They collect data and aggregate overall ESG scores as well as separate scores for each of the Environmental, Social, and Governance pillars (Rau and Yu, 2023)[13].

Authors assert that despite the growth in ESG data provision, there are notable issues, particularly concerning data quality and the divergence of ESG ratings. Initially, ESG data were sourced from public resources like financial reports and company websites. However, with increasing disclosure requirements, more firms are publishing annual CSR reports. While this enhances the availability of ESG data, it also raises concerns about the data's quality and reliability. Their work also notifies the potential for ”greenwashing,” where the reported ESG metrics might not accurately reflect the company's actual practices. Another significant problem is the discrepancy in ESG ratings provided by different data providers. This article points out that there are considerable variations among these providers in terms of coverage, metrics, criteria, and methodologies used for ESG rating.

To summarise this article (Rau and Yu, 2023)[13] which presents a literature review and asserts that, while ESG ratings are crucial for assessing a company's sustainability performance, some inherent challenges lay in ensuring the reliability and consistency of these ratings. It underscores the need for cautious interpretation of ESG data due to potential quality issues and the lack of standardization in rating methodologies across different providers.

This article provided us insight and an in-depth understanding of the intricacies of ESG ratings and some crucial ideas to support our objective i.e. leveraging ESG ratings in order to check for the possibilities of predicting the growth of a company.

Usage of sentiment analysis from news sources to predict the performance of small and medium-sized enterprises (SMEs) is increasingly relevant in the interconnected global market. A key study that delves into this area is titled ”Sentiment Correlation in Financial News Networks and Associated Market Movements” (Wan, X., Yang, J., Marinov, S., et al., 2021)[14].

This article reveals that sentiment towards one company, as reflected in media coverage, can influence not only its own market performance but also affect the sentiment and performance of other companies in the same or different sectors. This finding is crucial as it underlines the interconnected nature of the market and the potential broader impact of sentiment on related companies.

This study leverages advanced NLP techniques to algorithmically analyze large-scale unstructured data like financial news articles. This approach allows for the construction of networks of news and sentiment to monitor risk and volatility, while also studying the correlation of sentiment with market movements. Our work here follows a similar approach as mentioned in this (Wan, X., Yang, J., Marinov, S., et al., 2021)[14] article.

2.2 SMEs and Private Equity Investment Scenario

In this research work we have focused our research efforts on SMEs located in the United Kingdom (UK) where, the businesses classified as small and medium-sized enterprises (SMEs) are those with fewer than 250 employees (even some of the companies have much higher employee base than 250, but all are unlisted and private enterprises), a gross asset value of no more than 12.5 million British pounds, and a turnover of no more than 25 million British pounds (Statista Research Department, 2022)[15]. Authors in [15] highlight that SMEs contribute significantly to the UK economy given the huge portion of the country's enterprises that fit into this group and for a variety of reasons, including to meet growing customer demand, diversify their product offerings, or bolster their competitive positions, businesses must expand. They also rightly indicate that businesses frequently depend on outside funding to expand, and specify that every year, at least 25% of SMEs have expansion plans.

In their article[15], authors mention that Bank loans, building society loans, bank overdrafts, leasing or hiring, credit card loans, and government or local government grant programmes are the most popular forms of financing. They indicate that Twenty percent or so of SMEs had rejections for every application they submitted as applications for external financing are not always granted.

In Figure 2.1, from (Statista Research Department, 2022)[15], we see the number of SMEs for each corresponding field for the year 2022 in the UK. In this plot, it is apparent that the labor-intensive sectors such as Construction, Wholesale, and retail, and service industries such as healthcare and administrative services. This is obviously the case for almost all countries since there are hardly any multi-national companies in these sectors and local SMEs function well in these particular sectors due to lesser margins for supporting a large management structure. In our work, we have focused majorly on SMEs from sectors like finance, food and accommodation services, Information and Communication, news media

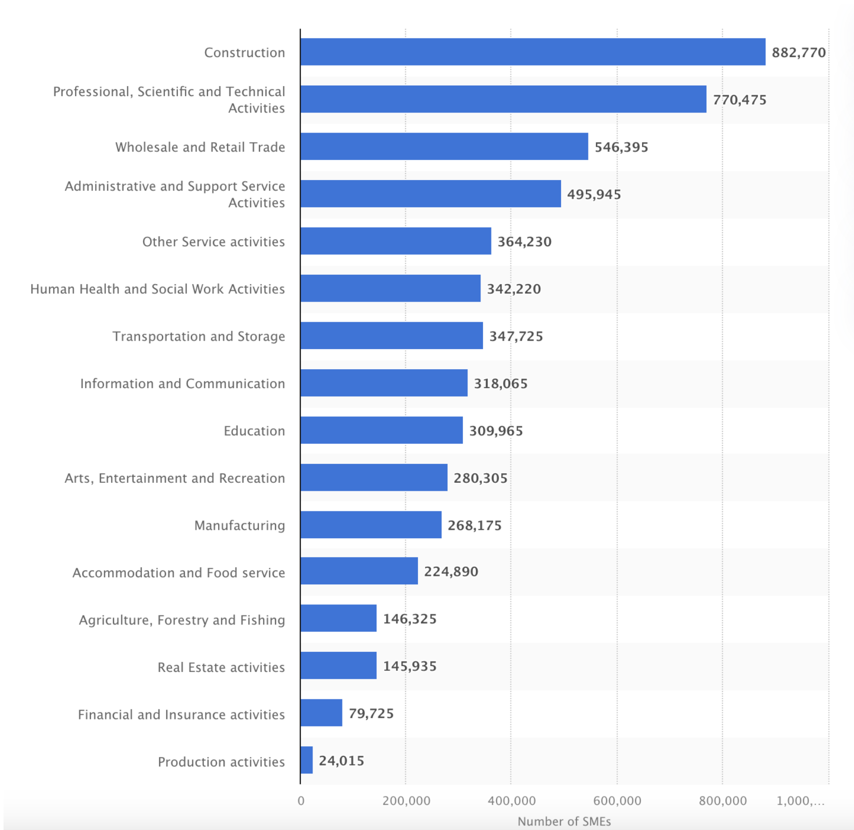


Figure 2.1: This plot from (Statista Research Department, 2022)[15], shows the number of SMEs (in the UK) in the respective fields as mentioned in the year 2023

groups, etc., specifics of which will be discussed in the next chapter.

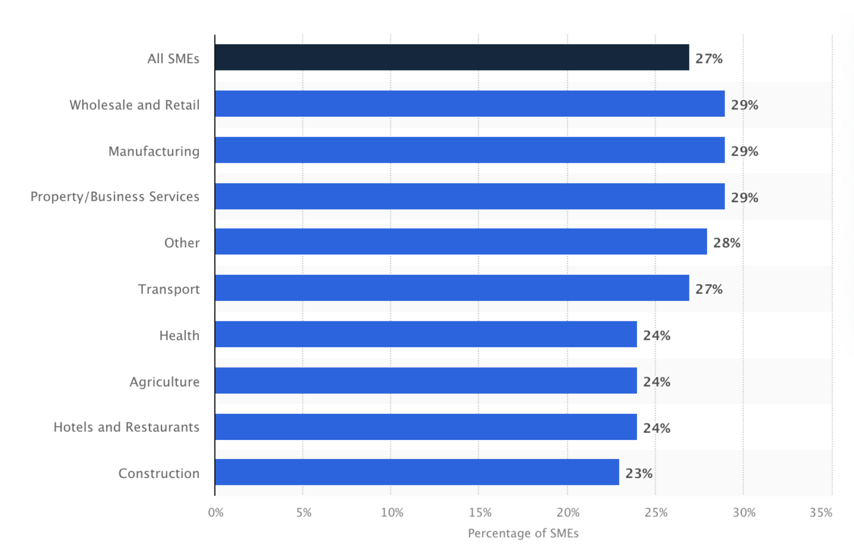


Figure 2.2: This plot from (Statista Research Department, 2022)[15], shows percentage of SMEs that achieved growth in the last 12 months in the United Kingdom in 2022

In Figure 2.2, also from (Statista Research Department, 2022)[15], we can see the percentage growth of various sectors of SMEs (in the UK) in 2022. The overall average growth percentage for all the SMEs is indicated to be about 27% and the range lies from 23% growth in construction to 29% in retail, manufacturing, and property services. These growth results are for the year 2022 when compared to the last year i.e. 2021 and might indicate the surge due to post COVID-19 resurgence in commercial activities.

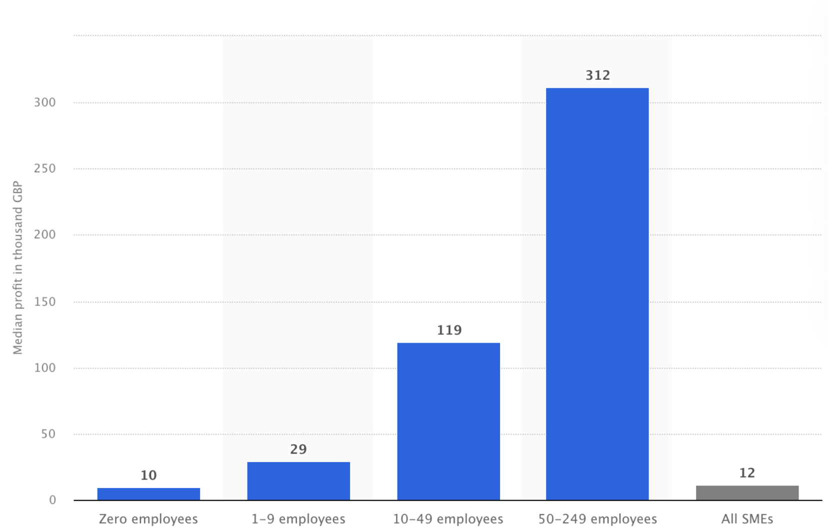


Figure 2.3: This plot from (Statista Research Department, 2022)[15], shows median profit made by SMEs (in the UK) in 2022, by enterprise size.

Figure 2.3, again from (Statista Research Department, 2022)[15], depicts the median profit made by SMEs of different enterprise sizes in the UK for 2022. This plot showcases a hyperbolic relation between the company size and the median growth throughout the year. The first bar shows the median growth for zero employees or those enterprises with self-employed personnel and the last bar shows the average median profit in GBP for all SMEs and the absolute value is much closer to the self-employed enterprises, indicating a much larger pool of SMEs with 1 to 9 employees.

All these figures and statistics are essential in understanding the SME market in the UK and its growth dynamics. In light of all these insights, we will be able to provide a broader picture to investors and fiduciaries to gauge a better investment decision.

2.3 Predicting Growth with various methods

So far we have elaborated on the importance of predicting growth of SMEs for all the interested investors and private equity stakeholders, but in this section we will discuss the specifics of the methods we would like to choose in our forecasting modeling. Authors in the research article titled "An optimized model using LSTM network for demand forecasting" (Abbasimeher, Shabani, and Yousefi, 2020)[16] presents a detailed methodology for demand forecasting using Long Short-Term Memory (LSTM) neural networks.

Authors in [16] compare their model's performance with popular time series data handling models: ETS (Exponential Smoothing), ARIMA, SVM, ANN, and single-layered LSTM. In their work, they claim to have developed a configured LSTM algorithm that considers temporal dependencies in a time series data. Their proposed method automatically selects the best predicting model by considering various combinations of LSTM hyperparameters for a given time series with the aid of grid search method.

The authors focus on the application of this model in the context of a furniture company's demand forecasting. This article (Abbasimeher, Shabani, and Yousefi, 2020)[16] includes comprehensive steps such as data preparation, hyperparameter optimization, and model training. Key aspects of the methodology involve data cleaning, normalisation, and the generation of a list containing combinations of hyperparameters for the LSTM network. The LSTM network's application is optimised through a grid search method to enhance forecasting performance.

Authors in [16] assert that their algorithm has the ability to capture non-linear patterns in time series data, while also considering the inherent characteristics of non-stationary time series data. Hence, they show that their configured LSTM has performed better than all the other method considered. The metrics used by the authors to assess the results are RMSE and SMAPE.

We referred to "Financial sentiment analysis using FinBERT with application in predicting stock movement" by (Jiang and Zeng, 2023)[17] in order to survey recent research in the domain of sentiment analysis on news in finance. In this research article the authors develop a deep neural network(DNN) based on FinBERT algorithm application on financial news of 6000+ stocks and additionally concatenated that sentiment analysis with the stock price data from 'yfinance' API and applied LSTM model on that data to eventually compare it with conventional models like ARIMA, LSTM (on both textual and numerical), FinBERT, and BERT. In the conclusion of this research article the authors have indicated that their model seems to underpredict in the long-run, but for short-term predictions their algorithm outperforms conventional algorithms like ARIMA which has its limitations with non-linear relations in the time series, and LSTM which is only considering either textual or numerical data at one time. The article (Jiang and Zeng, 2023)[17] also suggests a possibility of improvement with fine-tuning the FinBERT algorithm with their whole dataset or at least an augmented dataset.

We will discuss the details of FinBERT and LSTM modelling techniques in the chapters to follow.

2.4 Data Gathering

As mentioned in the previous chapter, we were provided a couple of datasets by Equitably.AI the first dataset is from Bloomberg terminal about more than 2,000 listed companies, including their ESG ratings and financial performance. Additionally, they gave us access to a dataset of ESG ratings for about 4,000 BCorp-certified SMEs. Equitably.AI kindly provided this dataset for scholarly use, and they expressed interest in studying the findings of our investigation.

Next, we narrowed our focus to solely UK-based SMEs for the purpose of this research project. Using the access to the FAME database provided by City's online databases, we were able to obtain the financial performance data for about fifty SMEs. We also used a free API platform 'thenewsapi' to collect pertinent articles for each of these fifty targeted SMEs in order to perform sentiment analysis on their online news.

Chapter 3

Methods

In this chapter, we elaborate on various methods of data collection and then perform exploratory data analysis. After this, we move on to explain the methodology of each predicting model we used with details of the workings of each modeling technique.

3.1 Datasets and Exploratory Data Analysis

To initiate our research with the provided data we take the dataset from the Bloomberg terminal of the listed companies. The Total Number of Unique Companies in this dataset is 2991, with 4 years of data for each i.e. from 2015 to 2018. We then in the step of data cleaning remove all the rows with null values and "unnamed" columns and drop the MSCI rating column due to many missing values. We will be anyway not be using the MSCI ratings for the purpose of this project because of the incompleteness of the MSCI data, and since this dataset is from Bloomberg terminal and it is their ratings we will be considered for training all our models.

The main columns for consideration in this project are EBITDA (Earnings Before Interest, Taxes, Depreciation, and Amortization), ENVIRON DISCLOSURE SCORE (Environmental disclosure score), SOCIAL DISCLOSURE SCORE (Social disclosure score), GOVNCE DISCLOSURE SCORE (Governance disclosure score), ESG DISCLOSURE SCORE (Overall ESG disclosure score).

We then clean the dataset by eliminating all the rows with null values in any of the cells

Column name	Mean	Min.	Max.	Std. Deviation
ESG	39.650	4.958	80.991	13.512
Environment	32.536	0.775	82.945	17.882
Social	38.854	3.125	89.062	15.738
Governance	56.531	3.571	83.928	13.512
EBITDA(in Millions)	2929.947	-3362.961	80342.00	5830.311

Table 3.1: Basic statistics of the key columns in our dataset

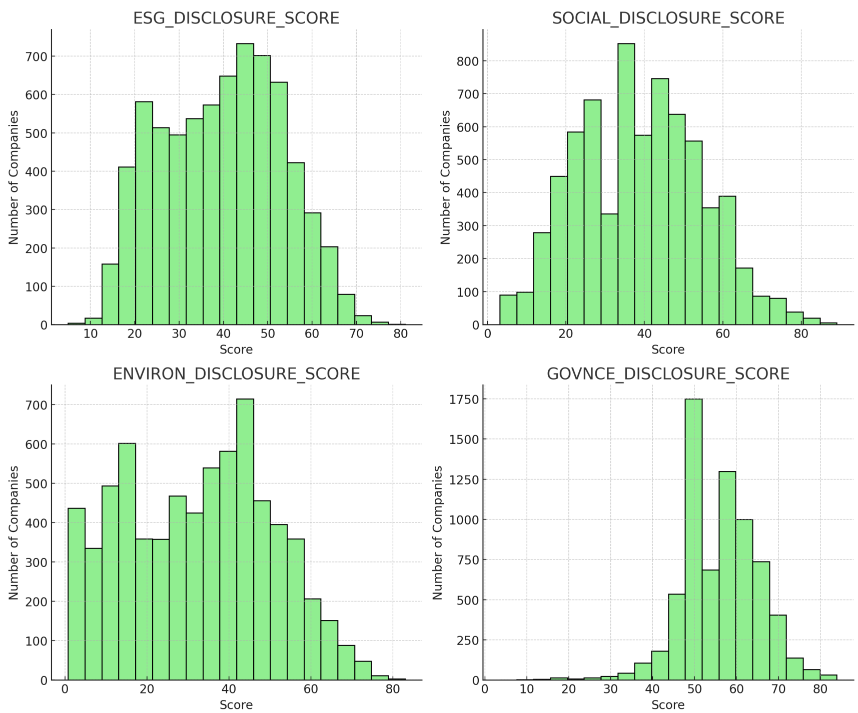


Figure 3.1: This plot showcases the distribution of all ESG disclosure scores (overall, social, environmental, and governance) across the Bloomberg listed companies dataset.

in these significant columns after deciding which are the most important to take into account for all upcoming modeling. Table 3.1 shows the basic statistics of all these key columns that we have decided to consider for all future models. All these values changed throughout the course of 4 years of data length for each individual company. Figure 3.1, visualises the distribution of all the ESG metrics for all the 2055 publicly listed companies from the Bloomberg dataset. We will be discussing more about the implications of all these distributions in the following chapters.

We have chosen this column in order to study the correlation of ESG ratings (and individual Environment, Social, and Governance Scores) with the EBITDA of those companies. We will be considering Environment, Social, and Governance Scores, and overall ESG ratings given by Bloomberg using their confidential criterion in which all these listed companies comply with certain data transparency and also indulge in providing all the necessary details as per the requirements of Bloomberg ESG research norms. Using these details as features of the neural networks and machine learning algorithms we will try to predict our target variable i.e. EBITDA.

Moving on to our second dataset which is the BCorp members data, this data consists of a list of SMEs from various countries around the world and to be specific there are 4547 SMEs in total. The total number of columns in this dataset is 127, such as company name, date of certification, description (about the company's bio), number of employees (range), etc, From this big pool of SMEs, we filtered the total number of SMEs in the UK which turned out to be 687 in total. Then taking these names we manually searched for at least 5 to 10 years of financial performance data among these 687 companies on FAME's database using its access from City's online databases available for students and alumni.

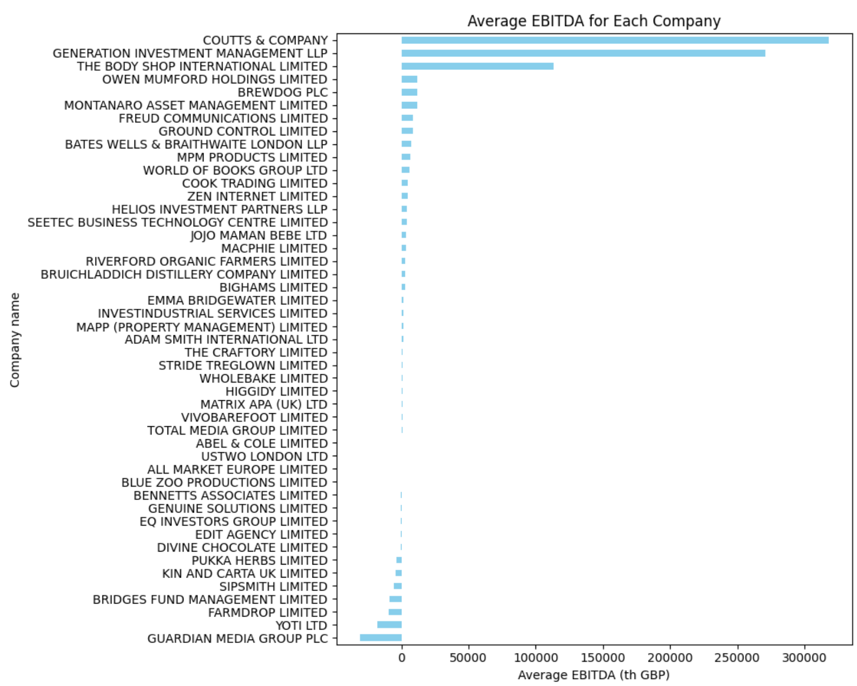


Figure 3.2: This plot visualizes the average EBITDA in thousands of GBP for all the SMEs, and is represented in ascending order from bottom to top.

Figure 3.2 above, shows the average EBITDA in thousands of GBP, there is a stark difference in the span of EBITDA across all the companies, the implications of this aspect will be discussed in the following chapters.

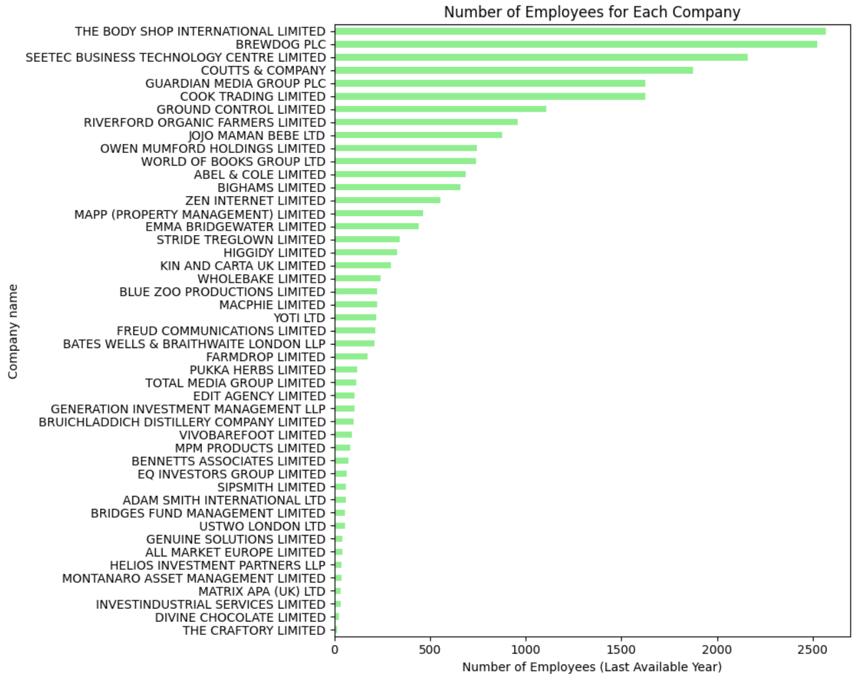


Figure 3.3: This plot visualizes the number of employees for each of the SMEs, again in ascending order from bottom to top.

Similar to the previous plot, Figure 3.3 elaborates on certain factual information about each of the SMEs in consideration. This figure depicts the distribution of the number of employees for each of the corresponding SMEs. As one would expect high correlation between Figure 3.2 and 3.3 but this doesn't seem to be entirely the case here, as many companies with fewer employees have supposedly higher average revenue. This might be the case due to a plethora of reasons and will be looked upon in the chapters to follow.

We were able to acquire a substantial amount of data for about 47 companies (50) and the dataset ranges from as early as 2010 to 2022 in most cases (a few years less in the case of some companies). It was extremely challenging to narrow down these SMEs because they almost have no obligation to provide their ESG metrics to any rating agency since they are not publicly listed enterprises. It was equally hard to gather financial performance data for these companies as we manually searched for almost all of the 687 UK-based companies. It eventually condensed to 47 companies, many of the companies did not even provide their financial performance consistently, possibly due to a huge slump in their revenue during the COVID-19 period another apparent thing noticeable was many of those 687 SMEs were under solvency procedure or already solvent. We had to rule out such companies from the scope of this work, as it would not be possible for us to put all those companies in the same bracket as those that were still up and running.

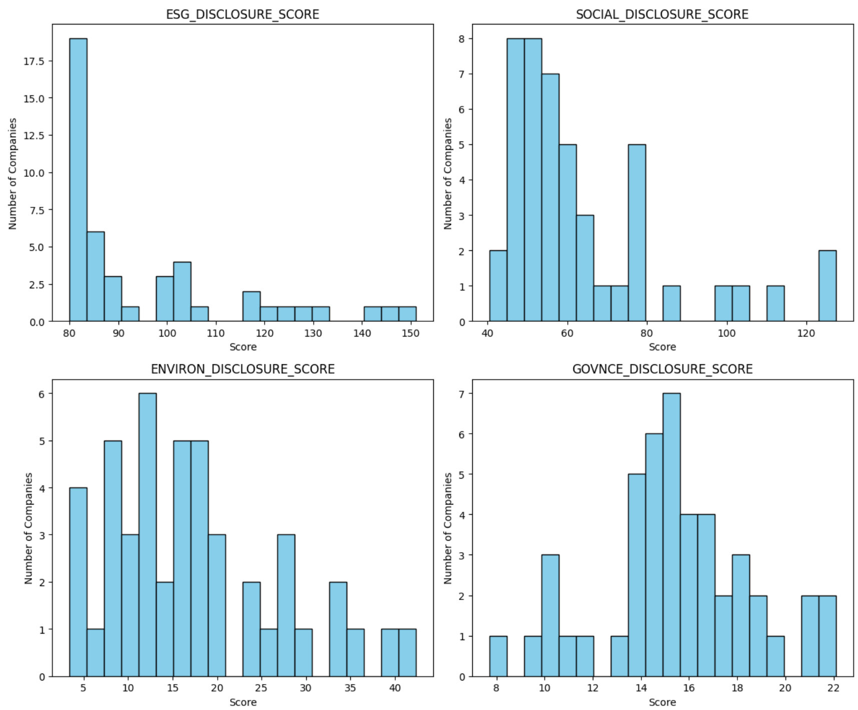


Figure 3.4: This plot showcases the distribution of ESG metrics for all the 47 BCorp member SMEs that we've considered for this project.

From the BCorp members' data, we identified 6 main columns representing the ESG ratings for our work here. Those columns are Impact Area Customer, Impact Area Worker, Impact Area Community, Impact Area Environment, Impact Area Governance, and Overall. We then manually combined these impact area customer, worker, and community to make it one column and named it SOCIAL DISCLOSURE SCORE (Social disclosure score), this was done to maintain continuity with the training dataset i.e. Bloomberg terminal data for listed companies. For the same reason, we changed the names of impact area environment, impact area governance, and Overall columns to ENVIRON DISCLOSURE SCORE (Environmental disclosure score), GOVNCE DISCLOSURE SCORE (Governance disclosure score), and ESG DISCLOSURE SCORE (Overall ESG disclosure score) respectively.

There were many other columns that were missing a lot of values and since we had the overall ESG score, all their minute weights would be captured in the ESG DISCLOSURE SCORE column. This also allowed us to circumvent the necessity to research the nitty-gritty of both the ESG rating methodologies from Bloomberg and from BCorp. Studying both of the methodologies would be crucial in order to understand the impact and the depth of insight for each of these SMEs and listed companies, but it is something that lies beyond the scope of our work here.

Finally, we will now discuss about fetching data using "thenewsapi" about news-related data from various news sources online. To begin, we import the requests library used for making HTTP requests in Python. For formatting Company Names we replace spaces with "%20" in each company name (for URL formatting). In order to perform data Fetching in a loop over company names to fetch data using the API obtained from "thenewsapi", we construct a URL with a company name and page number. We then extract and append relevant information from the API response to the lists and end up creating DataFrame from the lists filled with data. Our initial approach combines API data retrieval, and data processing using pandas.

	Unnamed: 0	Unnamed: 0.1	Company	Title	UUID	Published At	Categories	Relevance Score	Description
0	0	0	Coutts Company	Coutts CEO Peter Flavel Resigns	04f0a771-62f3-422c-b53f-a40b0fa56670	2023-07-27T13:11:22.000000Z	['politics', 'general']	32.443060	Coutts chief executive Peter Flavel has resign...
1	1	1	Coutts Company	Coutts chief Peter Flavel steps down	0334ed7f-4c8c-4ecf-a361-91392b9cf168	2023-07-27T13:24:29.000000Z	['general', 'business']	31.877266	Exit from upmarket lender 'by mutual consent' ...
2	2	2	Coutts Company	Coutts-Trotter permanently takes on NSW treasu...	3a3bf900-7850-4968-89fc-1f97bbd41fd4	2023-05-22T00:26:41.000000Z	['politics']	31.383171	Michael Coutts-Trotter, former secretary of NS...
3	3	3	The Body Shop International	The Body Shop Celebrates Canada's Ban on Cosme...	35b6a48d-5521-4194-b360-b61fe2d6f069	2023-06-30T16:30:22.000000Z	['entertainment']	45.267220	The Body Shop, Cruelty Free International, and...
4	4	4	The Body Shop International	Natura & Co to Sell The Body Shop	4352226d-3096-4434-8e93-18a59fdd024d	2023-11-14T13:15:51.000000Z	['entertainment']	45.224884	International private equity group Aurelius wi...

Figure 3.5: This is what the first few rows of our fetched data look like.

Figure 3.5 is the representation of the first few rows of our dataset (concatenated after a couple of fetching stretches due to the maximum free hits allowed for one account on this API is about 100 hits only).

```
count    453.000000
mean      27.368547
std       11.032695
min        1.637818
25%       20.222445
50%       27.409771
75%       34.444400
max       77.963090
Name: Relevance Score
```

Figure 3.6: Some basic stats of the Relevance Score of each news article in our dataset.

As we can see in the second last column in the head of our data frame from Figure 3.5, we eventually explored some basic stats to understand the relevance score's distribution and that of our fetched dataset. Figure 3.6 displays some of the basic statistics of the relevance score of all the articles in our dataset. The maximum and the minimum values are 77.96 and 1.63 respectively which means that the dataset has a varied range of pertinence in context to the SMEs. And understandably the mean of the relevance score is around 27.36 which is also indicating the significance of the dataset to be reasonable.

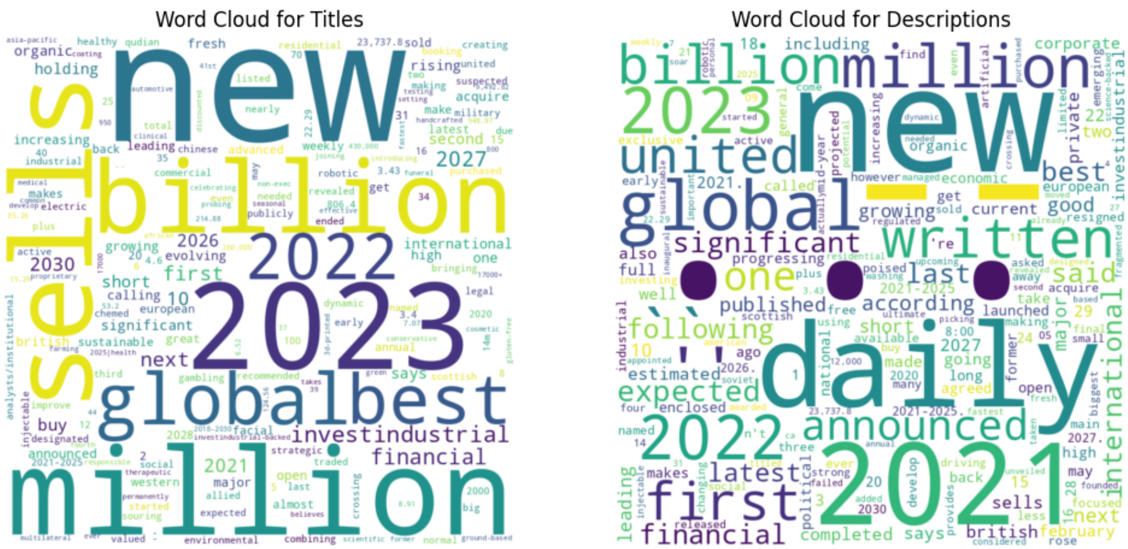


Figure 3.7: Word Clouds of the most commonly used words in the title and the description of each news article.

To explore this acquired data further we make Word Clouds taking the aid of python's "nltk" library and various packages in that library. Figure 3.7 represents the two-word clouds that we made from both the titles and descriptions columns of our dataset. The title word cloud highlights words like billion, million, globalbest, sells, and years (2022, 2023, 2026, 2030, etc.) which indicate the exact relevance that we might be looking for in order to determine the sentiment about the growth or performance of any company in a news article. Similarly, in the word cloud for the descriptions column, we can see a lot of such words: new, global, million, united, announced, years, etc., indicating a solid correlation with our objectives.

Carrying forward this preliminary research about the dataset we have plotted the most frequently occurring words in our dataset (barring the nouns, pronouns, prepositions,

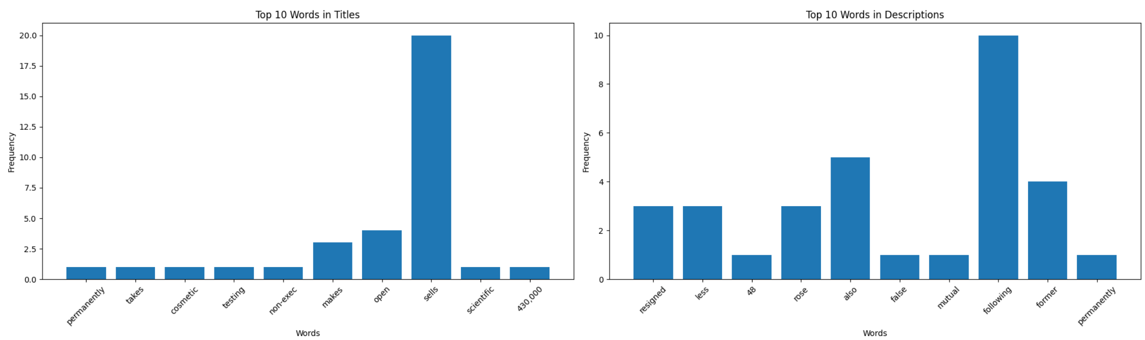


Figure 3.8: Frequency of the top 10 most common words in titles and description of each article.

etc.). As we can see in Figure 3.8, sells is the most frequently used word in the titles of the articles, and words like resigned, rose, former, etc. are among the top 10 frequently used words in the description of the articles. This is again a sign of appropriateness of the fetched news article data about SMEs to some extent.

3.2 Implementation of our first prediction model

We will be discussing the implementation of our first prediction model in this section. We have considered the data from the Bloomberg Terminal i.e. the dataset of 2055 publicly listed companies with their ESG metrics as features to train on and their EBITDA to be predicted by our configured model. We have used Long-Short Term Memory (LSTM) neural network which is a form of Recurrent Neural Network (RNN) that is preferred for time-series predictions (Greff et al., 2017)[18]. The primary distinction between an LSTM and an RNN is its ability to map input and output data suitably and preserve long-range time dependency information (Greff et al., 2017)[18]. Since the LSTM network architecture has gates and cells to regulate information flow, it differs from the traditional perceptron architecture. The LSTM is comprised of an input gate, a forget gate, an internal state (cell memory), and an output gate, as shown in Figure 3.9.

Where all the notations in Figure 3.9 from (Abbasimehr et. al., 2020)[16] are:

$x(t_i)$: The input value at time point t_i

$h(t_{i-1})$ and $h(t_i)$: The output value at time points $t-1$ and t

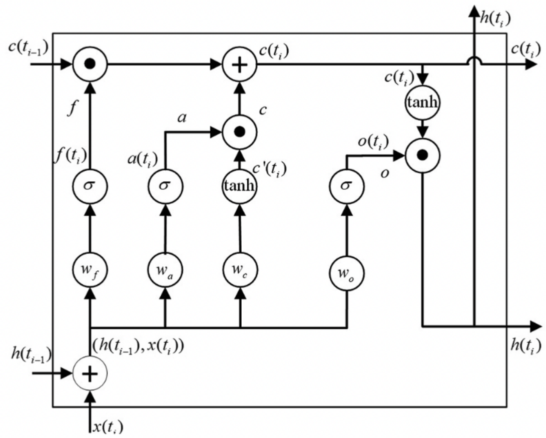


Figure 3.9: Architecture of LSTM model from (Greff et. al., 2017)[18]

$c(t_{i-1})$ and $c(t_i)$: Cell states at time points $t-1$ and t

$b = b_a, b_f, b_c, b_o$: Biases of input gate, forget gate, internal state, and output gate

$\vec{w_1} = w_a, w_f, w_c, w_o$: Weight matrices of input gate, forget gate, internal state and output gate

$\vec{w_2} = w_{ha}, w_{hf}, w_{hc}, w_{ho}$: Recurrent weights

$\vec{a} = a(t_i), f(t_i), c(ti), o(t_i)$: Output results for input gate, forget gate, internal state, and output gate.

So the functioning of LSTM goes like this: the forget gate $f(t_i)$ accepts $x(t_i)$ and $h(t_{i-1})$ as input to assess the information to be updated in $c(t_{i-1})$ using a sigmoid activation (Greff et al., 2017)[18]. The equations representing the updating and flow of learning in an LSTM algorithm are:

$$a(t_i) = \sigma(W_{ax}x(t_i) + W_{ah}h(t_{i-1}) + b_a) \tag{3.1}$$

$$f(t_i) = \sigma(W_{fx}x(t_i) + W_{fh}h(t_{i-1}) + b_f) \tag{3.2}$$

$$c(t_i) = f(t_i) \circ c(t_{i-1}) + a(t_i) \circ \tanh(W_{cx}x(t_i) + W_{ch}h(t_{i-1}) + b_c) \tag{3.3}$$

$$o(t_i) = \sigma(W_{ox}x(t_i) + W_{oh}h(t_{i-1}) + b_o) \tag{3.4}$$

$$h(t_i) = o(t_i) \circ \tanh(c(t_i)) \tag{3.5}$$

in these above equations σ and tanh represent the sigmoid and hyperbolic tangent activation functions respectively; and $\circ$ represents the dot product operator. In general, the LSTM algorithm learns by taking the subsequent actions:

- Uses Equations (3.1) to (3.5) to calculate the LSTM output - (learning Forward).
- Calculates the error between each layer's input and resultant data.
- The input gate, cell, and forget gate all allow the influence of the error's reverse propagation.
- An optimisation technique is used to adjust each gate's weight based on the error term.

After a certain number of iterations, the aforementioned four-step procedure is repeated to get the ideal weight and bias values (Abbasimehr et. al., 2020)[16]. This is the basic architecture of the LSTM algorithm that we have used for our prediction model. As mentioned earlier, we have used data from the Bloomberg terminal of listed companies which is 4 years of data for each company and we defined a function to split the data into training and validation sets. We split the data based on the "Year" column for each company, all years except the last one will be used for training. The last year's data was used for validation/testing. Then we go on to define the features for the LSTM algorithm to train on i.e. environment, Social, Governance and overall ESG disclosure scores. After that step we perform a grid search on all the hyperparameters i.e. batch size, epoch, units, and dropouts. We found the best set of hyperparameters to be: units = 50, dropout rate = 0.3, batch size = 16, and epochs = 50.

Then our next approach starts by extracting features and target variables ('EBITDA') from the training data. After that, these features and target variables are scaled within a range of 0 to 1 using MinMaxScaler. For neural network models, this scaling is essential since it facilitates faster and more reliable training. A create dataset function is then used to transform the dataset into a format appropriate for supervised learning. Based on a given look-back parameter, this function generates sequences of the input data (X) and the corresponding output (y). The look-back parameter determines how many previous time steps are used to predict the next time step. Upon running these commands we get a final training loss of 0.00455, validation Mean Squared Error of 0.00404, and validation Root Mean Squared Error of 0.0635. After this step, we save our trained model in the form of a ".h5" file on our Google Drive.

Now we need to make sure that the testing data i.e. the SMEs data from BCorp and FAME is also in similar shape and the columns are compatible with the training dataset. We then encountered a concern, in this SME dataset even though we have EBITDA values for 10 years for most companies the ESG metrics are available only for the year when those companies were last certified by BCorp. Hence, we need to fill/impute all those empty cells in order to be able to test this dataset on our trained model. In order to deal with this problem we considered using Gaussian Mixture Model (GMM) to impute the missing values of the ESG metrics for all the 47 SMEs. GMMs or jittering is a form of Gaussian noise augmentation, which is a commonly used method to impute the missing values based on the Gaussian (Normal) distribution of the existing values in each specified column (McCaw et. al., 2022)[19].

After imputing all these missing values we then upload the updated CSV file for testing our trained LSTM algorithm. After establishing all the requirements before testing this data i.e. defining the features and the target variable we then finally test this SME data on the trained model. The following are the results that we obtained: Mean Squared Error: 1522326767.999794 and Root Mean Squared Error: 39017.00613834683. We also generate the plots for each and every individual company in order to visualise the predicted and the actual values curves of EBITDA against each other. From these plots, it is apparent that our LSTM algorithm is not predicting the EBITDA values at all. We will discuss this more elaborately in the next chapter, but we then decided to explore the prospects of feature engineering in order to utilise the power of neural networks and machine learning algorithms appropriately.

3.3 Prediction Models with first model improvements.

In order to implement the first attempt to improve the model performance we referred to (Brealey, et. al., 2011)[20]. In this textbook, we read about the method to assess the growth of any firm by calculating the Compound Annual Growth Rate (CAGR) for various financial metrics. CAGR is a standard financial formula used to determine the annual growth rate of an investment over a specified period. While this specific formula is applied to EBITDA (Earnings Before Interest, Taxes, Depreciation, and Amortization) in our context, the underlying CAGR formula remains consistent across various financial metrics. The formula for EBITDA CAGR for our purpose would look like this (Brealey, et. al., 2011)[20]:

$$\text{EBITDA\_CAGR} = \left(\frac{\text{End Value of EBITDA}}{\text{Start Value of EBITDA}}\right)^{\frac{1}{n}} - 1$$

where n is the total number of years for which we have EBITDA values, and this is going to change if we have different lengths of EBITDA data for each company. But this is going to suffice the purpose for us. We have calculated one individual EBITDA CAGR number for each company in both the listed companies dataset and the SMEs dataset. We are doing so because we want to reduce each company's stats to one single row in both datasets. So for the Bloomberg dataset, we have reduced the 4 years of data (4 rows for each unique company) to 1 single row by also averaging the ESG metrics for all four years for each unique company. This will make the length of our Bloomberg data to be 2055 rows i.e. number of unique companies in the dataset. Similarly, for SMEs, we compute CAGR for each of the 47 SMEs, and for their ESG metrics, we don't have to average it because in the case of these BCorp member companies, we just have one single year of data and we simply consider those values. In this way, we have eliminated the need to impute the missing ESG values for SMEs dataset using GMMs or any other method.

3.3.1 LSTM implementation on EBITDA CAGR

Before the direct implementation of the LSTM algorithm here again we will now normalise the features i.e. all the ESG metrics in both the datasets of Bloomberg's listed companies and BCorp's SMEs. We will normalise both datasets separately in order to preserve the scaling of both the separate company scale. We have used the MinMax Normalisation for normalising. Then we move on to apply a Grid Search on a bunch of hyperparameters to tune the best set of combinations for our final training LSTM algorithm. Upon the Grid Search, the best set of hyperparameters are units = 50, dropout rate = 0.3, batch size = 16, and epochs = 100.

We then eventually train our new LSTM algorithm with EBITDA CAGR as our target. After training the algorithm we save it again in a ".h5" file like previously to call it as a "model" upon which the BCorp and FAME SMEs data could be tested. In the next steps, we do exactly that and eventually calculate the MSE, RMSE, etc. to check the accuracy of our predictions. We also plot our predictions against the actual values which will be shown in the next chapter.

3.3.2 SVM implementation on EBITDA CAGR

Support Vector Machine (SVM) is a machine learning technique for tackling regression and classification problems that is based on statistical learning theory. SVM maps the data points into a higher-dimensional space where they become separable using kernels like linear, polynomial, and Radial Basis Functions (RBF) for data that are not linearly separable. An SVM selects a hyperplane with the maximum margin between data points, assuming a set of data points from two classes (Han et. al., 2011)[21].

In order to implement the SVM algorithm we take the normalised datasets for both test and train purposes i.e. normalised listed companies' data (ESG) and normalised SMEs data (ESG). Again we have considered the EBITDA CAGR to be the target variable and the features are also the same i.e. all four ESG metrics. We have considered the Radial Basis Function or "rbf" to be our kernel consideration for the final training model. The rest of the implementation procedure is pretty similar to the LSTM algorithm as we save this trained SVM model into a ".joblib" file and then load the testing SME dataset to test the trained model on it and then make predictions. We generate various error evaluation parameters and visualisations to check and compare the accuracy of prediction the results of which are explained in the next chapter.

3.4 Prediction models with second improvement

After seeing the impact of EBITDA CAGR on both of our algorithms we want to try a new way to possibly improve our prediction results. In the previous section, we normalised both our datasets separately, ESG parameters (all four columns) and EBITDA CAGR. But this time around we realised that it actually makes sense to normalise the ESG metrics for both listed companies and unlisted SMEs separately because those ratings are provided by two separate agencies and just represent a subjective representation of the ESG performance of a company and the scale for both are different and would seem more appropriate to normalise separately.

Instead we this time considered the possibility of normalising EBITDA CAGR for both datasets in one scale. It is important to note that we are not normalising the column EBITDA in the pre-processed data. But we are normalising the compound annual growth rate of the EBITDA of each company and that is comparable for all the companies no matter what their scale is, and moreover, our neural networks and machine learning algorithms need similar and balanced metrics to learn and eventually predict. By this modification, we are enabling both our algorithms to learn comparable target variables while under training and predict something which is on the same normalised scale as it expects.

In order to do so, we combine both datasets by concatenating them into one data frame and then normalising using MinMax scaling normalisation to normalise the whole EBITDA CAGR column for this new data frame which consists of both SMEs and listed companies. After achieving normalisation we then separate both the data frames again and label them as the first part and the second part, representing SMEs data and listed companies data respectively.

3.4.1 LSTM implementation on second improvement

The methodology of this LSTM algorithm implementation is exactly the same as the previous one. The major differences obviously being the input data frames for both testing and training. Next, we adjust a number of hyperparameters using a Grid Search in order to find the optimal set of combinations for our last training LSTM algorithm. The optimal set of hyperparameters, as determined by the Grid Search, are units = 50, dropout rate = 0.2, batch size = 16, and epochs = 100.

Finally, we use EBITDA CAGR (normalised on the combination of both datasets) as our target for training our new LSTM algorithm. We resave the algorithm after it has been trained in a ".h5" file, just like before, and refer to it as a "model" that will be used to test the BCorp and FAME SMEs data. We carry out just that in the following steps, finally calculating the MSE, RMSE, etc. to verify the precision of our forecasts. In the following chapter, we will discuss our prediction plots and compare the results to all the previous and new models.

3.4.2 SVM implementation of second improvement

Even in the case of SVM, the methodology of implementation of the algorithm is just as same as its previous implementation and only the input data files are different. Yet again, the EBITDA CAGR has been chosen as the target variable, and all four ESG metrics are included in the features. For the final training model, we opted for the Radial Basis Function, or "rbf," as our kernel. After saving the trained SVM model into a ".joblib" file, we load the testing SME dataset to test the trained model on it before making predictions. The rest of the implementation process is much the same as the LSTM and previous SVM algorithm approaches. To evaluate and assess prediction accuracy, we produce some error evaluation metrics and visualisations, the outcomes of which are discussed in the following chapter.

3.5 Prediction model with the third improvement

The third and final attempt to improve the prediction of our algorithms we choose to calculate the percentage change, which is often used to assess volatility, is a fundamental concept in finance and economics (Berk and DeMarzo, 2019)[22]. The percentage change formula is:

$$\text{Percentage Change} = \left(\frac{\text{New Value} - \text{Old Value}}{\text{Old Value}}\right) \times 100\%$$

This method is frequently used to calculate the rate at which a variable changes over time, such as the EBITDA of a business from year to year. Volatility is a widely accepted approach in financial analysis for evaluating volatility using percentage change. A statistical indicator of the distribution of returns for a certain securities or market index is called volatility. The standard deviation or variation between returns from the same investment or market index is frequently used in the finance industry to quantify volatility (Berk and DeMarzo, 2019)[22].

One can directly evaluate the volatility of EBITDA by calculating the percentage changes in EBITDA from year to year. This is simply a measurement of how much EBITDA fluctuates. In order to calculate the percentage change i.e. a measure of volatility we take the above-mentioned formula from (Berk and DeMarzo, 2019)[22] and plug the EBITDA values from both the pre-processed datasets separately and independently calculate EBITDA Volatility and replace the EBITDA columns in both of these input datasets.

3.5.1 LSTM implementation on EBITDA Volatility

Before implementing the LSTM algorithm we calculate the EBITDA Volatility for both the datasets as mentioned above. The methodology for implementing the LSTM algorithm again is the same as in the previous section. We calculate the best set of hyperparameters using Grid Search on a few hyperparameters and we get the following to be the best parameters epochs = 100, batch size = 20, optimizer = Adam, neurons = 50, and dropout rate = 0.2. We train our model on these best set of hyperparameters in the next step and save the model in the ".h5" file. Then to test the prediction capability of this new trained model we upload the SME data with EBITDA Volatility column as the target variable and other normalised ESG columns as features. And eventually just like previously we generate a few metrics to visualise and numerically evaluate the performance of our prediction algorithm and will be discussed in the next chapter.

3.5.2 SVM implementation on EBITDA Volatility

Similar to the LSTM algorithm implementation we take the input of both the datasets with EBITDA Volatility as a column in them. For the training stage of our SVM implementation, we begin with hyperparameter Grid Search for the best set of hyperparameters. For C - the regularization parameter we take the following values: (0.1, 1, 10); for gamma - the kernel coefficient the following: (scale, auto) and for kernel we take: (RBF, linear, poly, sigmoid). The gird search yielded the following best hyperparameters: C = 0.1, gamma = scale, and kernel = RBF.

After this, we train our SVM algorithm on the above-mentioned best hyperparameters using the training dataset with EBITDA Volatility as the target variable and ESG metrics as features. Upon saving this trained model as a ".joblib" file we then test our model with the test dataset and generate the metrics to compare and visualise the model predictions, discussed in the next chapter.

3.6 Prediction of Growth with ARIMA

The ARIMA (AutoRegressive Integrated Moving Average) model is a widely used statistical method for forecasting time series data, especially in the context of finance. Understanding ARIMA requires a grasp of its three main components (Hamilton, 1994)[23]:

• AutoRegressive (AR): This aspect of the model captures the relationship between a variable's current value and its past values. In financial time series, this can help model the momentum or mean-reversion effects often observed in asset prices or returns. The AR part uses parameters (denoted as p in ARIMA(p,d,q)) to indicate how many past time points are used for prediction.

• Integrated (I): Integration refers to differencing the time series data, which means subtracting the previous value from the current value. This process helps in making the time series stationary, a key requirement for many time series forecasting methods. A stationary time series is one whose properties do not depend on the time at which the series is observed, thereby having constant mean and variance over time. The order of differencing is denoted by d in ARIMA(p,d,q).

• Moving Average (MA): It refers to the relationship between the time series and residual errors from a moving average model applied to lagged observations. The parameter q in ARIMA(p,d,q) represents the size of the moving average window. One major drawback of ARIMA is that it assumes that a provided time series is linear (Adhikari and Agrawal, 2013)[24].

3.6.1 ARIMA implementaion

For the implementation of ARIMA, we begin with loading the SMEs data with EBITDA values for each of the 47 companies from the FAME database. After loading this file to our Python notebook we check for missing years in the dataset as it would affect the ARIMA model because of its requirement of continuous time series. We use dropna() to deal with any missing values (if any). In the next step, we proceed to perform the Augmented Dickey-Fuller (ADF) test. The ADF test is a statistical test used to determine whether a given time series is stationary. Stationarity is a key assumption in many time series models, including ARIMA. The function adfuller() from a time series analysis library (statsmodels in Python) is used to perform the ADF test on the 'EBITDA' column of our data frame. The result of adfuller() function is stored in result[].

The function constructs a text string result text [] that reports the ADF statistic, p-value, and critical values. The p-value is particularly important if it's below a threshold (commonly 0.05), it suggests that the time series is stationary. This means there's enough statistical evidence to reject the null hypothesis of non-stationarity. Depending on the p-value, the function adds a conclusion to result text [], stating whether the data is stationary or not.

Next, we move on to plotting the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) for a time series, which in this case is the 'EBITDA'. These plots are critical in time series analysis, especially for determining the parameters of ARIMA models. We based on ACF and PACF decide to consider the configuration of (1,0,0) for the ARIMA parameters (p,d,q). Then we run our ARIMA model and obtain various statistics to evaluate its performance. All the performance metrics and the plots to visualise the performance of the ARIMA algorithm will be discussed in the next chapter.

3.7 Sentiment Analysis of the news articles data

We have a cleaned and processed dataset from "thenewsapi" and we just need to apply the models to compute sentiment analysis for the all these news articles corresponding to the 47 SMEs.

3.7.1 VADER algorithm implementation

VADER (Valence Aware Dictionary and Sentiment Reasoner) is a sentiment analysis tool particularly well-suited for dealing with social media text, including short phrases, emojis, and slang. It's unique in its approach to understanding the contextual nuances of emotional language.

VADER contains a lexicon (a list of words and their sentiment intensities). It calculates a compound score that summarizes the overall sentiment of a text. This score is a normalized, weighted composite score computed by summing the valence scores of each word in the lexicon, adjusted according to the rules, and then normalized to range between -1 (most extreme negative) and +1 (most extreme positive) (Hutto and Gilbert, 2014)[25].

In order to implement the VADER algorithm we just upload our news article dataset in the Python notebook. Then, we use the SentimentIntensityAnalyzer from the nltk.sentiment module. NLTK (Natural Language Toolkit) is a suite of libraries and programs for symbolic and statistical natural language processing (NLP) for English written in Python. We then apply sentiment analysis to the "Description" column of the dataset which has the descriptions corresponding to the article and the company it is referring to. We define a function to compound the score of positive, negative, or neutral sentiment of all the articles and then we plot the distribution of the compound scores of all the articles in the dataset. We then aggregate the compounded score and group it by company name to get a final sentiment score for each company. We then also plotted a table to visualise the number of positive, negative, or neutral articles for each of the companies in the dataset. We will discuss more about the results in the next chapter.

3.7.2 FinBERT implementation

BERT Algorithm: BERT is a groundbreaking algorithm developed by Google, primarily used for natural language processing tasks. It understands the context of a word in a sentence by looking at the words that come before and after it. This makes it highly effective for understanding the nuances of language.

FinBERT Adaptation: FinBERT is an adaptation of the original BERT model, specifically fine-tuned for financial texts. This adaptation involves training the BERT model on a large corpus of financial texts, such as company reports, financial news, and analyst reports. By doing so, FinBERT becomes more adept at understanding the jargon, expressions, and nuances typical in financial contexts (Jiang and Zeng, 2023)[17].

We took this pre-trained model to test it on our dataset and the implementation part was almost very straightforward. First, we load the FinBERT model. Then we create a sentiment analysis pipeline using the loaded model. Then we define a function analyze sentiment () that takes a piece of text (news article) and returns the sentiment label (Positive, Negative, Neutral). Eventually, we group the resulting DataFrame by company names and counts of the occurrences of each sentiment. The final table (sentiment counts) shows the number of positive, negative, and neutral articles for each company.

Chapter 4

Results

In this Chapter, we succinctly present all the results of our models and compare the results of one algorithm to another to highlight the insights one might infer from the comparisons.

4.1 Prediction of the first LSTM Algorithm

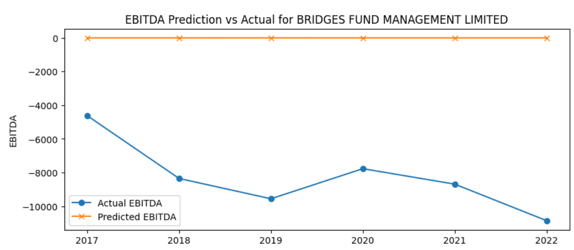


Figure 4.1: This plot represents prediction vs actual EBITDA for one of the 47 SMEs.

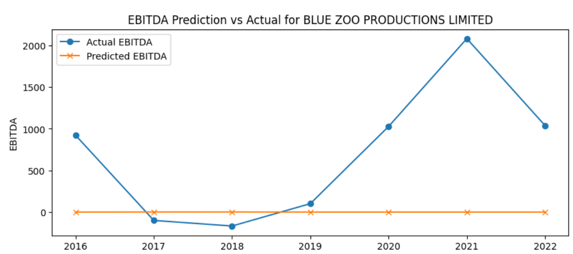


Figure 4.2: This is another plot showing prediction vs actual EBITDA values for one of the SMEs.

Both plots represent the abysmal performance of our first prediction model. Figure 4.1 highlights a company called Bridges Fund Management which had shown losses in its EBITDA since 2017 and our model is not predicting anything and simply giving a straight line at zero. At first glance, it might seem that our model isn't able to capture the negative EBITDA trends and hence we have shown another plot i.e. Figure 4.2 which is the prediction plot of another company called Blue Zoo Productions Limited and it is evident that the algorithm seems to show no learning or prediction even with the positive EBITDA trend. Following are the two error metrics that we have generated to get a context on the performance of this first prediction algorithm:

Mean Squared Error: 1522326767.999794

Root Mean Squared Error: 39017.00613834683

4.2 Prediction of first improvements - EBITDA CAGR

Next, we evaluate the results of our first attempt to improve the performance of our LSTM prediction algorithm. In this case, as discussed above we have considered independently normalised features and computed EBITDA CAGR values for both train and test datasets. We could certainly anticipate some improvement in the prediction from using this methodology because these normalised values for all the considered columns make it easier for the LSTM neural network to learn and compare the desired pattern. However, the straight line of the predicted values shows the inability of the LSTM algorithm to grab the trend and make predictions.

The error metrics for this algorithm are:

Mean Squared Error: 0.06558830316584847,

Root Mean Squared Error: 0.2561021342469611,

R-Squared: -2.290248054378423

Interestingly, when we apply SVM to the same approach we can see some movement from the norm; and these seem to be the first signs of any algorithm responding to the EBITDA trend. Here, the SVM algorithm isn't really predicting any close enough values

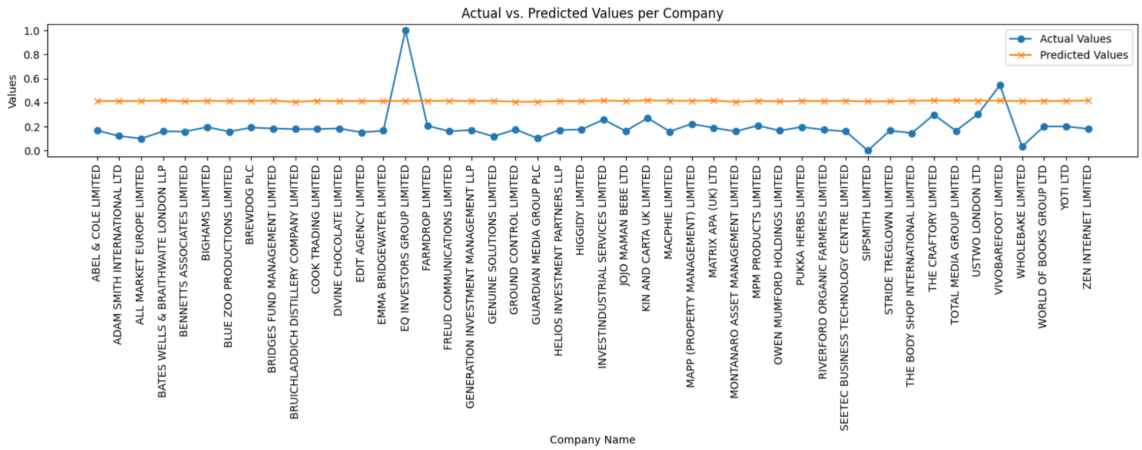


Figure 4.3: This is a plot representing the performance of the LSTM algorithm with EBITDA CAGR as the target variable.

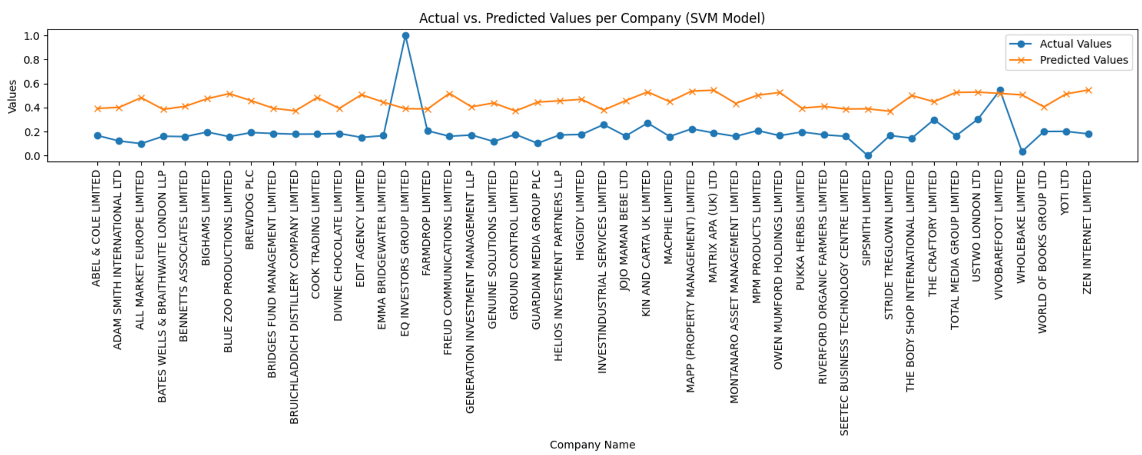


Figure 4.4: This plot represents the performance of the SVM algorithm with EBITDA CAGR for all 47 companies.

but still seems to be moving with the pattern. The following are the performance metrics of the SVM algorithm:

Mean Squared Error: 0.08746729748395268

R-Squared Score: -3.3878114157123758

Mean Absolute Error: 0.28057562828469124

4.3 Prediction of second improvement - Normalised CAGR

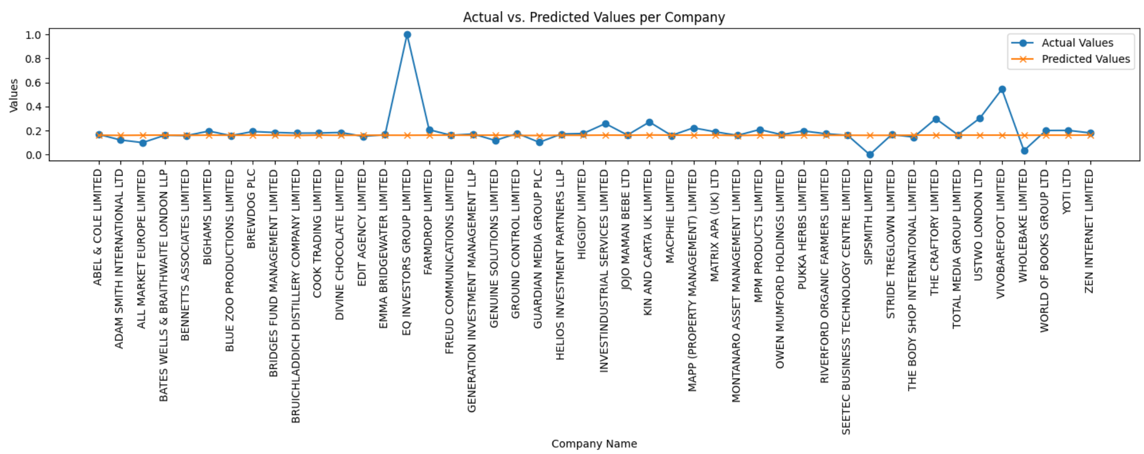


Figure 4.5: The prediction plot of the LSTM algorithm with the second improvement.

We now examine the LSTM algorithm in light of our second attempt to improve the results of predictions. Here, we are using EBITDA CAGR which is normalised with both SMEs data and listed companies' data in one dataframe.

The prediction curve even though is better aligned to the actual values curve and one possible leeway seemingly is the presence of some outliers in the dataset. EQ Investors Group is one such outlier that outperforms all the other SMEs and makes the prediction curve look impotent. The following are the error metrics for this algorithm:

Mean Squared Error: 0.021284209119420545,

Root Mean Squared Error: 0.14589108649749835,

R-Squared: -0.06772586366621214.

Figure 4.6 depicts the prediction performance of the SVM algorithm on the second improvement in the methodology. The following are the error metrics of the SVM algorithm with the second improvement:

Mean Squared Error: 0.02137342761354317

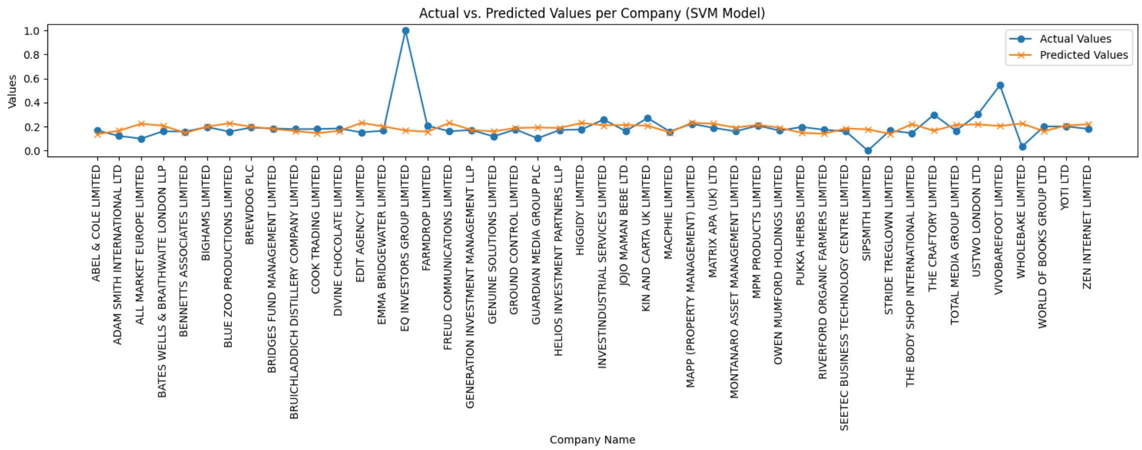


Figure 4.6: The prediction plot of the SVM algorithm with the second improvement.

R-Squared Score: -0.07220152415035774

Mean Absolute Error: 0.070247533558113O3

From these metrics, we can compare the SVM and LSTM models and observe that LSTM only slightly trumps the performance of the SVM algorithm. We will be discussing the implications of all these results in the following chapters.

4.4 Prediction of third improvements - EBITDA Volatility

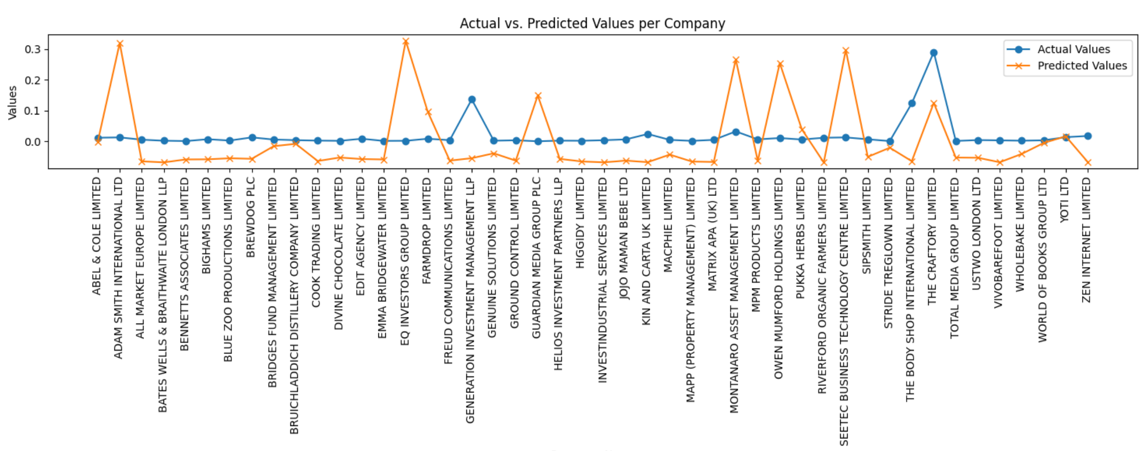


Figure 4.7: The prediction plot of the LSTM algorithm with the third improvement.

As we have discussed in the previous chapter in order to improve our prediction results we attempted to feature engineer the target column by calculating its percentage change or volatility. Figure 4.7 is the result of the LSTM algorithm implementation on EBITDA Volatility. As much as we wanted to say that maybe the third time is the charm and to some extent, we can endorse that claim since the plot itself is doing a lot as compared to all our previous algorithm predictions. We have some statistics to endorse our claim as well, since:

Mean Squared Error (MSE): 0.014175202768339

Root Mean Squared Error (RMSE): 0.11905966054184347

are evidently the best among all the other algorithm performance metrics.

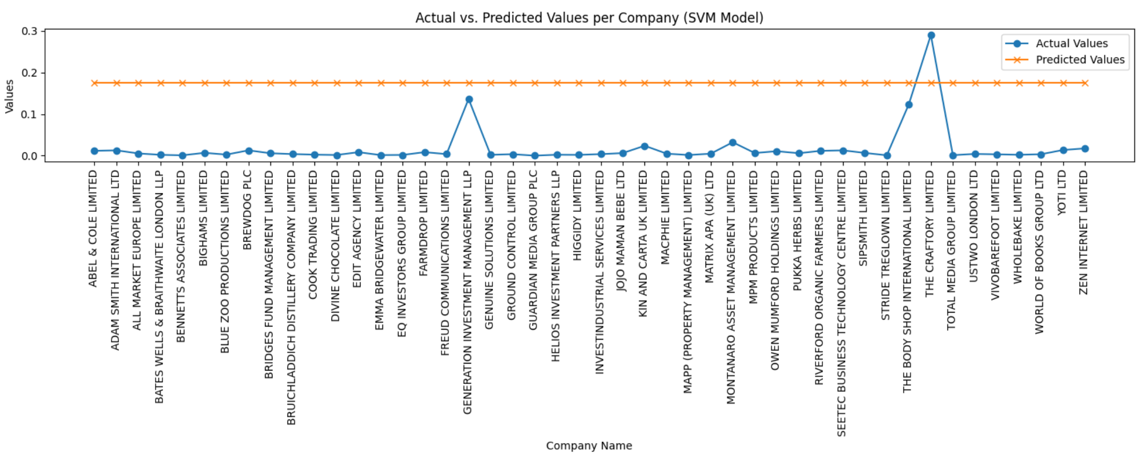


Figure 4.8: The prediction plot of the SVM algorithm with the third improvement.

For the case of the SVM algorithm implementation on EBITDA Volatility datasets, the results are amusingly off from what was expected. From Figure 4.8 we can evaluate the cluelessness of the predicted values curve. There is only one positive in this algorithm's defense, the scale of the plot is extremely small and most actual values are near the zero line as compared to the predicted values. Also, due to the extremely small-scale nature of the actual and prediction values the errors are reasonably acceptable as compared to other algorithms:

Mean Squared Error: 0.027133368378709118

Mean Absolute Error: 0.1625066160608054

4.5 Prediction of Growth with ARIMA

Using ARIMA for forecasting the EBITDA values for the next two years was the objective and has been appropriated accordingly. Figures 4.9 to 4.14 showcase a sequential depiction of growth predictions of all 47 SMEs in consideration for this study.

The model used is ARIMA(1, 0, 0). This indicates an AutoRegressive model of order 1 (one lag), with no differencing (I=0), and no moving average component (MA=0).

Coefficients (coef):

• const: Represents the constant term in the model. The estimated value is 3371.5568 with a standard error of 2257.769. The z-value (1.493) and P-value (0.135) suggest that this constant term is not statistically significant at conventional levels (e.g., 0.05).

• ar.L1: Is the coefficient for the first lag of the AR(1) model. It's estimated at 0.1650 with a standard error of 1.169. The z-value and P-value indicate that this coefficient is not statistically significant.

Model Fit Statistics:

Log Likelihood: The log-likelihood, at -122.920, measures how well the model fits the data. Higher values indicate a better fit.

AIC (Akaike Information Criterion): 251.840. It's a measure of the model's quality and complexity. Lower AIC values indicate a better model.

BIC (Bayesian Information Criterion): 253.535. Like AIC, it measures model quality, penalizing more complex models.

HQIC (Hannan-Quinn Information Criterion): 251.492. Another criterion for model selection, balancing goodness of fit and complexity.

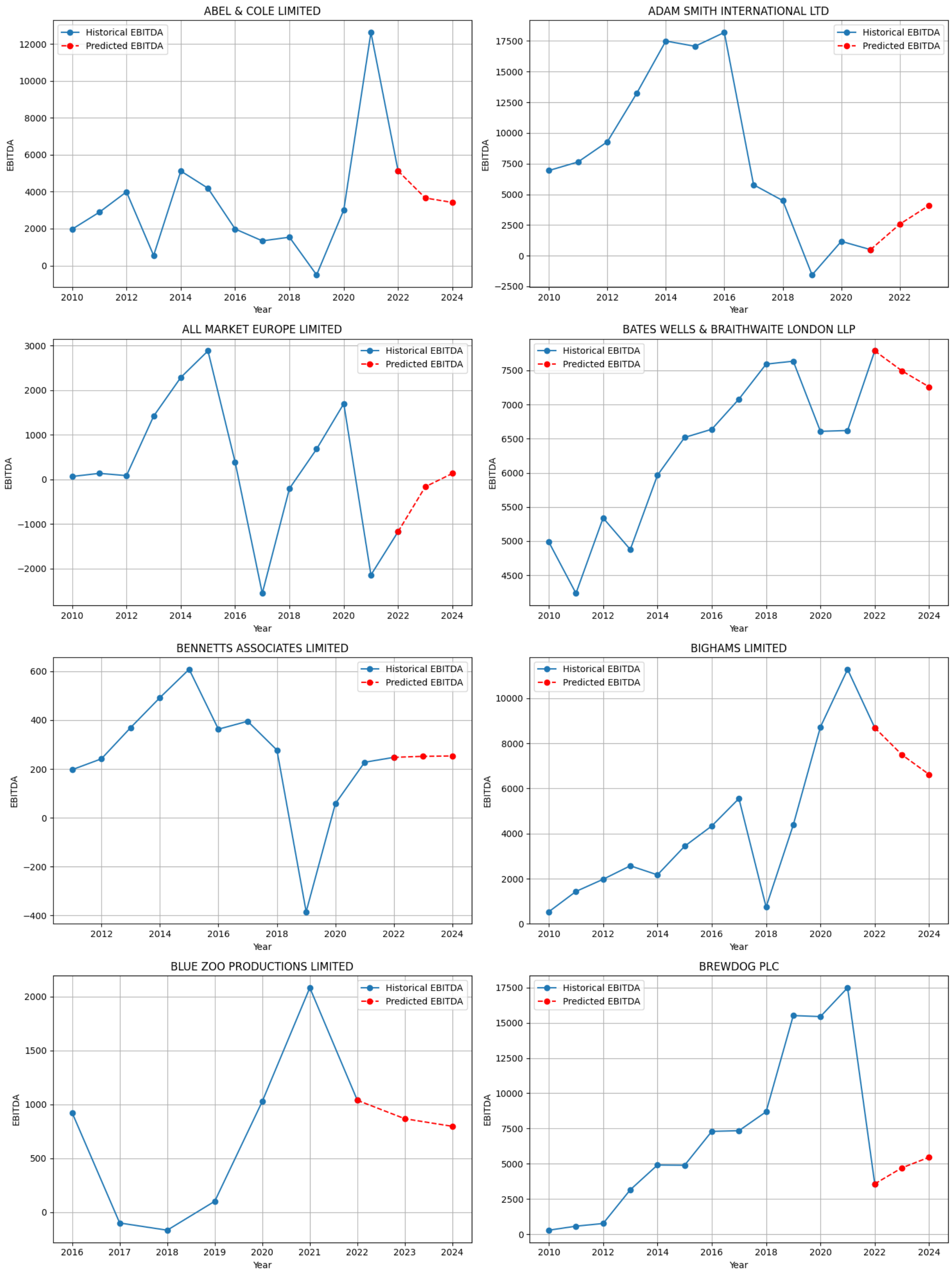


Figure 4.9: Growth prediction plots using ARIMA for companies 1 to 8.

Covariance Type: The type used is 'opg', indicating the optimizer used for estimating the covariance matrix of the parameters.

Figure 4.10: Growth prediction plots using ARIMA for companies 9 to 16.

sigma2: This is the estimated variance of the error term (residuals) of the model, with a value of about 1.015e+07. The standard error and its confidence interval suggest that

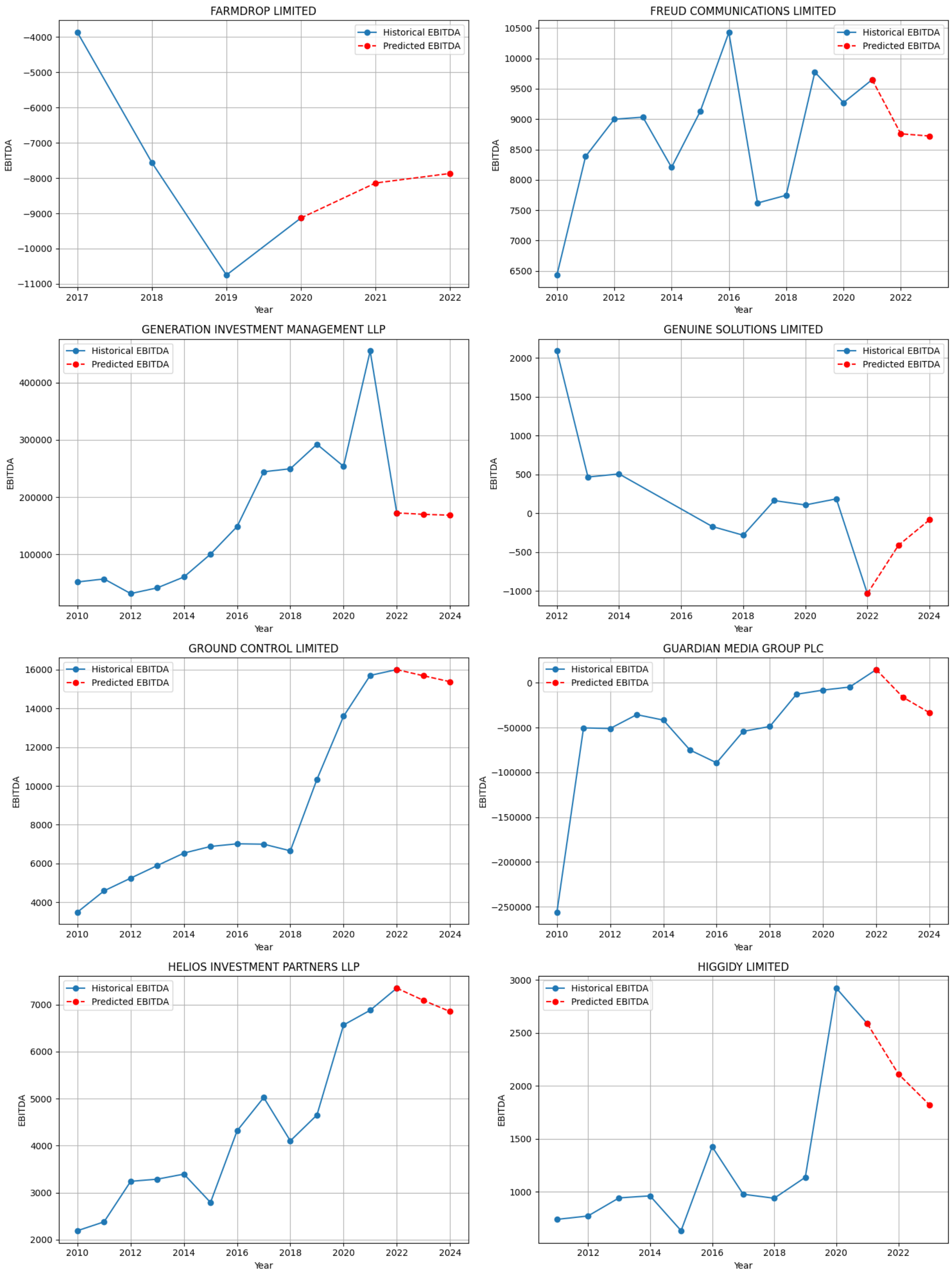


Figure 4.11: Growth prediction plots using ARIMA for companies 17 to 24.

this estimate is statistically significant.

Figure 4.12: Growth prediction plots using ARIMA for companies 25 to 32.

Diagnostic Tests:

Figure 4.13: Growth prediction plots using ARIMA for companies 33 to 40.

- Ljung-Box Test: Tests for autocorrelation in residuals. The Prob(Q) value of 0.83 suggests no significant autocorrelation.

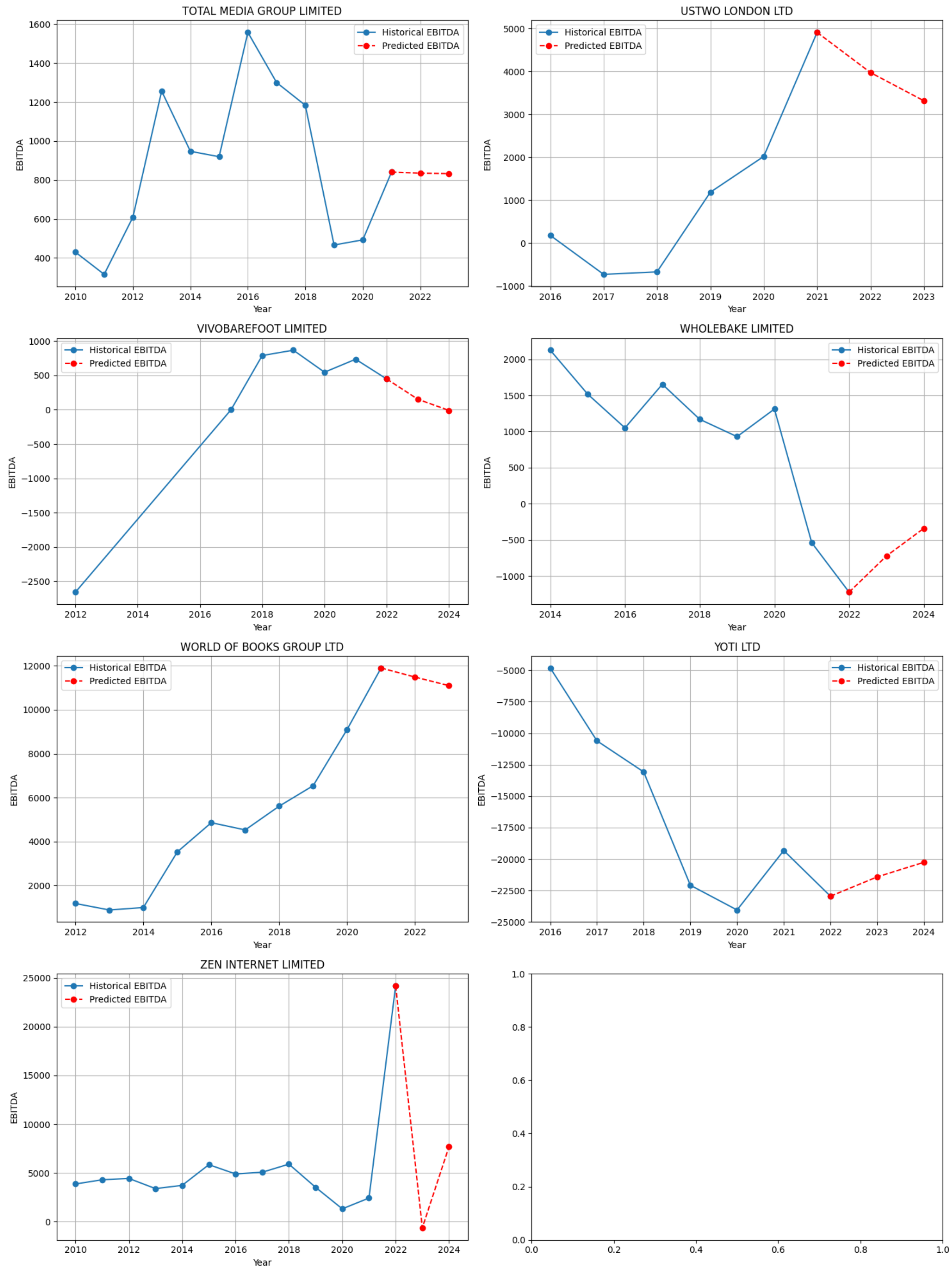


Figure 4.14: Growth prediction plots using ARIMA for companies 41 to 47.

• Jarque-Bera Test: Tests for normality of residuals. The Prob(JB) of 0.00 indicates that the residuals are not normally distributed.

• Heteroskedasticity Test: The Prob(H) value of 0.06 suggests that there might be some heteroskedasticity in the model, although this is borderline significant.

• Skew and Kurtosis: Skewness of 1.90 and kurtosis of 6.58 also support the Jarque-Bera test's finding of non-normal residuals.

4.6 Sentiment analysis - VADER

Company Sentiments

	Negative	Neutral	Positive
Abel Cole	0	4	8
Adam Smith International	3	4	5
All Market Europe Ltd	0	3	9
Bates Wells	2	7	3
Bennetts Associates	1	1	5
Bighams	0	0	1
Blue Zoo Productions Ltd	0	0	2
BrewDog	1	5	9
Bridges Fund Management	0	4	8
Bruichladdich Distillery Co. Ltd	0	4	1
Cook Trading Ltd	1	7	4
Coutts Company	5	8	5
Divine Chocolate Ltd	1	1	4
EQ Investors Group Limited	0	4	8
Edit Agency Limited	1	7	4
Emma Bridgewater Ltd	1	2	2
Farmdrop Ltd	0	1	0
Freud Communications Limited	4	3	4
Generation Investment Management LLP	1	8	6
Genuine Solutions	0	2	10
Ground Control Ltd.	2	7	6
Guardian Media Group	6	7	5
Helios Investment Partners	0	1	11
Investindustrial	0	4	8
JoJo Maman Bebe	2	5	5
Kin and Carta Europe	3	7	2
MAPP Ltd	2	4	0
MPM Products Ltd	1	4	7
Matrix APA (UK) Ltd.	1	1	1
Montanaro Asset Management	1	5	6
Owen Mumford Ltd	1	2	9
Pukka Herbs	2	3	7
Riverford Organic Farmers Ltd	0	2	4
Sipsmith	0	4	8
Stride Treglown	0	1	2
The Body Shop International	1	9	8
The Craftory	1	2	7
Total Media Group Ltd	1	5	6
VIVOBAREFOOT	3	3	6
World of Books Group	2	4	6
Yoti	1	4	7
Zen Internet Ltd	0	9	3
ustwo	1	3	8

Figure 4.15: The results of the VADER algorithm on news data.

From Figure 4.15 we can glean the results of the VADER algorithm and the basic statics for these results are:

Total Negative sentiments: 52

Total Neutral sentiments: 171

Total Positive sentiments: 230

4.7 Sentiment analysis - FinBERT

Company Sentiments Analyzed by FinBERT

	Negative	Neutral	Positive
Abel Cole	0	11	1
Adam Smith International	1	10	1
All Market Europe Ltd	0	3	9
Bates Wells	0	11	1
Bennetts Associates	0	5	2
Bighams	0	1	0
Blue Zoo Productions Ltd	0	1	1
BrewDog	2	12	1
Bridges Fund Management	2	9	1
Bruichladdich Distillery Co. Ltd	0	5	0
Cook Trading Ltd	1	11	0
Coutts Company	5	13	0
Divine Chocolate Ltd	1	4	1
EQ Investors Group Limited	0	10	2
Edit Agency Limited	0	11	1
Emma Bridgewater Ltd	1	3	1
Farmdrop Ltd	0	1	0
Freud Communications Limited	0	10	1
Generation Investment Management LLP	1	14	0
Genuine Solutions	0	10	2
Ground Control Ltd.	1	10	4
Guardian Media Group	1	15	2
Helios Investment Partners	0	12	0
Investindustrial	0	12	0
JoJo Maman Bebe	2	10	0
Kin and Carta Europe	1	8	3
MAPP Ltd	0	6	0
MPM Products Ltd	0	6	6
Matrix APA (UK) Ltd.	0	2	1
Montanaro Asset Management	2	8	2
Owen Mumford Ltd	0	5	7
Pukka Herbs	0	8	4
Riverford Organic Farmers Ltd	0	2	4
Sipsmith	0	10	2
Stride Treglown	0	3	0
The Body Shop International	2	15	1
The Craftory	0	10	0
Total Media Group Ltd	2	8	2
VIVOBAREFOOT	2	7	3
World of Books Group	1	9	2
Yoti	1	10	1
Zen Internet Ltd	0	11	1
ustwo	0	10	2

Figure 4.16: The results of the FinBERT algorithm on news data.

From Figure 4.16 we can evaluate the results of the FinBERT algorithm and the basic statics for these results are:

Total Negative sentiments: 29

Total Neutral sentiments: 352

Total Positive sentiments: 72

4.8 Comparison of VADER and FinBERT

From the data, we can observe the following differences:

VADER detected more negative and positive sentiments compared to FinBERT. FinBERT detected a significantly higher number of neutral sentiments. The interpretation of which tool is "better" depends on the context and the expected outcome of the sentiment analysis.

If the news articles are expected to have more pronounced sentiments (either positive or negative), then VADER's results may seem more aligned with this expectation as it identifies stronger sentiments more frequently. If, however, the articles are of a more technical or balanced nature (which could be common in financial or specialized news), FinBERT's higher neutral counts may be more appropriate as it might detect the subtleties and complexities in the language better. FinBERT's strength lies in its training on financial texts, which can contain nuanced sentiments that are neither clearly positive nor clearly negative. VADER's lexicon and rule-based approach may lead to stronger sentiment scores for certain words which could explain the higher number of positive sentiments detected.

In conclusion, neither tool is universally "better," but rather, each has strengths that make it more suitable for different types of text. The choice between VADER and FinBERT should be made based on the specific nature of the text being analyzed and the goals of the sentiment analysis task.

Chapter 5

Discussion

In this chapter, we discuss the outcomes of various stages in our project and their impact on the course that we followed to culminate this research endeavour.

5.1 Prospects of prediction of growth using ESG metrics/ratings

Our very first objective of this project was to answer the question of whether can we predict the growth of a company using the ESG ratings (metrics). Now that we have concluded the research part we can claim that this objective has been partially achieved. We endorse this claim by emphasizing that we were able to gather as much ESG-related data for these SMEs as possible and then also narrowed down their financial metrics from the FAME database and then started to apply one algorithm after the other with multiple improvements and yet we can't say for that all the objectives are achieved with certainty.

We started with a straightforward approach of the implementation of the LSTM algorithm keeping all the features and target variables in tandem for both training and testing datasets. The first result we obtained was extremely terrible in terms of accuracy and prediction. This led us to dig deep into finance-based literature review even further to try and figure out a feature that could be engineered in a way to solve our purpose and enable this prediction process to exercise a fair chance.

We hopped on to three different ways of improving our prediction algorithms. EBITDA CAGR seemed to be a perfect candidate for us to be able to implement first. The results were encouraging and we went on to improve the scope of EBITDA CAGR by providing it a second upgrade in the form of normalising it with both dataframe concatenated together. This further improved the performance of both SVM and LSTM algorithms.

But we didn't settle there, we went on to explore yet another concept of finance and investment banking called Volatility. This eventually gave us our best-performing model in LSTM for the target variable EBITDA Volatility.

We gleaned a lot of in-depth insights about the meaning and utility of ESG ratings for a numerous number of companies of various scales and backgrounds. The amount of data that we could gather with the constraints of time and financial resources was an achievement but also an impediment in itself since it proved to not be sufficient. This was sort of anticipated since the very beginning of this project and was highlighted in a couple of papers in the literature review part of this report.

We will discuss and evaluate the shortcomings of the project endeavour and the learning to glean from those critical self-reflections in the next chapter.

5.2 How can ARIMA suffice an educated investment decision?

From the results shown in the previous section we can infer that implementation of ARIMA algorithm was successfully implemented and the predictions for the EBITDA for all 47 SMEs were successfully completed as well. ARIMA is a time test modeling technique for many different time series applications. ARIMA has its own limitations and challenges and we have also encountered a similar mixed bag while using ARIMA as a modeling technique.

ARIMA can be a powerful tool to predict the targeted variable in a time series given enough time steps and the data should establish a linear relationship in order for ARIMA to suffice its utility. We will be discussing more about the specifics of ARIMA and its shortcomings in the next chapter.

5.3 Viability of Sentiment Analysis for an educated investment decision

As we mentioned in the previous chapter sentiment analysis can be a powerful tool for assessing the news sentiment or the public image of a particular company. We in this study have focused our scope only on SMEs which were unlisted companies and yet some of the companies were middle-sized to be featured periodically in news articles. We were able to gather a sufficient amount of data for most of these SMEs and then we were also able to successfully implement the sentiment-related pre-trained algorithms without much fuss to extract valuable information for investors and their fiduciaries.

From Figures 4.15 and 4.16, it is evident that we can quantify the overall positive, negative, or neutral sentiment for each individual company and hence this objective should be considered as achieved too.

Chapter 6

Critical Evaluation and Conclusion

During the course of this chapter, we will evaluate the shortcomings of our work and in general the learnings and takeaways from this research endeavour.

6.1 Review of ESG-related prediction models

We have already emphasised the importance of credible data in at least a couple of chapters prior to this. Although we had taken both of our datasets from standard rating agencies the opinion about the veracity of these self-reported ESG metrics face constant skepticism from many scholar and researchers.

The fact that ESG assessment demands a sufficient investment of time, human effort, and capital to maintain and update the streamline of all the latest data, has made it challenging to accept the ESG ratings as the right indicator of the growth of a company.

Can we use ESG to predict the growth of a company?

From our rigorous research effort, we can endorse the claim of many scholars that the current information about the ESG performance of any company let alone the unlisted or listed companies should be taken with a pinch of salt. This opinion is based entirely on the inferences gleaned while participating in this research endeavour.

6.2 Review of ARIMA as choice to suffice investment decisions

In order to critically evaluate the shortcomings of our ARIMA model's performance we can consider the following aspects:

With only 13 observations, the sample size is quite small. Small samples can lead to unreliable estimates and make it difficult to find statistically significant results.

Non-normal residuals (Jarque-Bera Test) indicate outliers or skewed data. If the EBITDA data contains outliers or is inherently skewed, it might affect the normality of residuals. If the variance of the error terms changes over time, it can lead to heteroskedasticity. This is common in financial time series data, like EBITDA. If the model doesn't fit the data well, it can manifest as heteroskedastic residuals.

Limited Data Points i.e. with only 13 observations, the estimates of the model parameters can be unstable and not generalize well. There's a higher risk of overfitting with smaller datasets, meaning the model might capture noise instead of the underlying pattern. Another reason could be if the data has a seasonal pattern not accounted for, it could affect the model's performance.

Can we use ARIMA to suffice informed investment decisions?

To address these issues, we might consider expanding the dataset, trying different model specifications, checking for and handling outliers, and ensuring that the data preprocessing steps are appropriately applied. Also, testing other types of models, like SARIMAX with different parameters or other time series models, could provide better results.

6.3 Review of Sentiment Analysis as a viable option for investors to consider

Sentiment analysis tools like VADER and FinBERT can face challenges or even fail to predict the right sentiment under certain scenarios. Here are some of the common issues they might encounter:

Both methods may struggle with text that uses sarcasm or irony because the literal meaning of the words may be positive or neutral while the intended meaning is negative. Sentences that are ambiguous or have multiple interpretations can lead to incorrect sentiment predictions.

Language that is heavily domain-specific, including technical jargon or industry lingo, can be misinterpreted if the tool is not specifically trained on data from that domain. Cultural references, idioms, or phrases that are not part of the training data can lead to incorrect sentiment analysis. Creative writing, poetry, or any text with a non-standard use of language can pose a challenge to sentiment analysis tools.

Sentiment analysis tools can become outdated if they are not continuously trained with new data. Language evolves, and models might not recognize newer expressions of sentiment

Can Sentiment Analysis serve as a viable option for investors to rely on?

Each tool's performance will vary depending on how well it has been trained to handle these issues. FinBERT, for example, might perform better with financial texts even in some of these challenging scenarios due to its specialized training data. VADER might perform adequately in social media contexts where expressions are more direct and less nuanced. The key to improving accuracy is selecting the right tool for the job and providing it with domain-specific training data when possible.

From our research-based opinion, we can certainly endorse this statement that sentiment analysis can serve as a crucial tool to assess the prospects of a firm and would be a vital tool to any investor as an additional set of information to engage and educate from in order to make educated investment decisions.

Bibliography

[1] B. G. Malkiel. “Random Walk Down Wall Street. 2nd ed. New York, NY: WW Norton.” In: (1984). URL: https://search.worldcat.org/title/50919959.

[2] Timo Busch, Alexander Bassen, and Gunnar Friede. “ESG and Financial Performance: Aggregated Evidence from More than 2000 Empirical Studies, Journal of Sustainable Finance & Investment”. In: *Journal of Sustainable Finance & Investment* (2015). DOI: 10.1080/20430795.2015.1118917.

[3] CFA Institute and Principles for Responsible Investment. “ESG Case Studies ESG Integration Framework”. In: (2018). DOI: 978-1-942713-57-9. URL: https://rpc.cfainstitute.org/en/research/surveys/guidance-case-studies-esg-integration-survey-report.

[4] Jianan Du, Brandon Thomas, and Janis Zvingelis. “How and Why SRI Performance Differs from Conventional Strategies.” In: (2014). URL: https://www.investpmc.com/insights/white-papers/how-and-why-sri-performance-differs-conventional-strategies.

[5] M.E. Bateman and L.R. Goldberg. “Ownership of ESG characteristics.” In: (2023). URL: https://doi.org/10.1057/s41260-023-00335-y.

[6] T. Loughran and B. McDonald. “When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks.” In: (2011). DOI: TheJournalofFinance,66(1),35-65..

[7] P. C. Tetlock. “Giving content to investor sentiment: The role of media in the stock market.” In: (2007). DOI: TheJournalofFinance,62(3),1139-1168.

[8] B. Pang and L. Lee. “Opinion mining and sentiment analysis.” In: (2008). DOI: FoundationsandTrendsinInformationRetrieval,2(1-2),1-135..

[9] W. T. Coombs. "Ongoing crisis communication: Planning, managing, and responding." In: (2007). DOI: SagePublications..

[10] E. Brynjolfsson, L. M. Hitt, and H. H. Kim. "Strength in numbers: How does data-driven decision-making affect firm performance?" In: (2011). URL: https://dx.doi.org/10.2139/ssrn.1819486.

[11] T. H. Davenport and J. Harris. "Competing on analytics: The new science of winning." In: (2007). URL: https://hbr.org/2006/01/competing-on-analytics.

[12] T. Aven. "Risk assessment and risk management: Review of recent advances on their foundation." In: (2016). DOI: EuropeanJournalofOperationalResearch,253(1),1-13.. URL: https://hbr.org/2006/01/competing-on-analytics.

[13] P. Raghavendra Rau and Ting Yu. "A survey on ESG: investors, institutions and firms." In: (2023). DOI: ChinaFinanceReviewInternational,ISSN:2044-1398. URL: https://www.emerald.com/insight/content/doi/10.1108/CFRI-12-2022-0260/full/html#abstract.

[14] X. Wan et al. "Sentiment correlation in financial news networks and associated market movements." In: (2021). URL: https://doi.org/10.1038/s41598-021-82338-6.

[15] Statista Research Department. "Financing of small and medium enterprises in the UK - statistics facts." In: (2021). URL: https://www.statista.com/topics/5875/financing-of-smes-in-the-uk/#topicOverview.

[16] Hossein Abbasimehr, Mostafa Shabani, and Mohsen Yousefi. "An optimized model using LSTM network for demand forecasting." In: (2020). DOI: https://doi.org/10.1016/j.cie.2020.106435.

[17] Tingsong Jiang and Andy Zeng. "Financial sentiment analysis using FinBERT with application in predicting stock movement." In: (2023). DOI: 10.48550/arXiv.2306.02136.

[18] K. Greff et al. "LSTM: A Search Space Odyssey, in IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 10, pp. 2222-2232". In: (2017). DOI: 10.1109/TNNLS.2016.2582924.

[19] Z.R. McCaw, H. Aschard, and H. Fitting Julienne. “Gaussian mixture models on incomplete data. BMC Bioinformatics 23, 208 (2022)”. In: (2022). DOI: https://doi.org/10.1186/s12859-022-04740-9.

[20] Richard A. Brealey, Stewart C. Myers, and Franklin Allen. “Principles of Corporate Finance”. In: (2011). DOI: https://doi.org/10.1186/s12859-022-04740-9.

[21] J. Han, J. Pei, and H. Tong. “Data mining: concepts and techniques.” In: (2011).

[22] Jonathan Berk and Peter DeMarzo. “Corporate Finance - 5th Global Edition”. In: (2019). DOI: 9781292304151.

[23] James D Hamilton. “Time Series Analysis. Princeton University Press”. In: (1994). DOI: https://doi.org/10.2307/j.ctv14jx6sm.

[24] R. Adhikari and R. K. Agrawal. “An introductory study on time series modeling and forecasting.” In: (2013). DOI: arXiv:1302.6613.

[25] C. Hutto and E. Gilbert. “VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proceedings of the International AAAI Conference on Web and Social Media”. In: (2014). DOI: https://doi.org/10.1609/icwsm.v8i1.14550.

Acknowledgements

I take this opportunity to show my gratitude towards all of the members of Equitably.AI especially Dr. Tillman Weyde and John-Paul Danee for all the valuable discussions and moral support. I would also like to acknowledge the role of Dr. Daniel Chicharro in making this research work possible as he was extremely encouraging and supportive, and I am extremely thankful to him for showing faith in my abilities.

Regards,
Siddharth Madan