MSc Dissertation Submission Cover Page

Student Name: MUIDEEN O. ABUBAKAR

Programme: MSc Banking and Digital Finance

Supervisor: Ramin Okhrati (PhD.)

Dissertation Title:

Sentiment Analysis on Sustainability Reports and its Relationship with Cost of Capital: A Case Study of Carbon-Intense Sectors in the United States

DECLARATION OF OWNERSHIP

- I confirm that I have read and understood the guidelines on plagiarism, that I understand the meaning of plagiarism and that I may be penalised for submitting work that has been plagiarised.
- I declare that all material presented in the accompanying work is entirely my own work except where explicitly and individually indicated and that all sources used in its preparation and all quotations are clearly cited.
- I have submitted an electronic copy of the project report through Moodle/Turnitin.

Should this statement prove to be untrue, I recognise the right of the Board of Examiners to recommend what action should be taken in line with UCL's regulations.

Signature: Muideen Abubakar	**Date: 15 September 2022**

University College London

MSc Banking and Digital Finance

Sentiment Analysis on Sustainability Reports and its Relationship with Cost of Capital: A Case Study of Carbon-Intense Sectors in the United States

15th September 2022

Dissertation submitted in partial fulfilment of the requirements for an MSc in Banking and Digital Finance, University College London

Institute of Finance and Technology
University College London

Abstract

Climate change remains a topical issue globally, and the Net Zero target is to limit the global temperature to 1.5^{O}c by 2050. To achieve this, GHG emissions need to drop by 45% by 2030, but it may be a pipe dream without the financial sector's active involvement. One critical role is to incentivise committed companies or penalise unrepentant polluters. To demonstrate their commitment, companies use their annual sustainability reports to inform investors how their activities impact the environment and their remediating initiatives. Meanwhile, it is crucial to investigate whether the contents of these reports have a significant relationship with the WACC demanded by investors. In addition, it is worth finding out if investors give high importance to the contents of these reports in their capital allocation and investment decisions.

This study adopts a Natural Language Processing (NLP) technique to determine the overall sentiments from sustainability reports and test the relationship between the sentiment scores and the WACC from 2017 to 2021. The FinBERT model was used for polarity derivation, focusing on the top five pollutive sectors in the US.

The findings reveal a mixed relationship between both variables among the examined sectors. The Real Estate and Agriculture sectors show a positive but weak correlation. Meanwhile, the top three pollutive sectors (Transportation, Energy/Electric power, and Industrials) demonstrate a negative but weak relationship. This underscores the huge attention and scrutiny accorded to these three sectors by investors. However, the weak relationship indicates that sustainability reports may not be a decisive factor in the investment and capital allocation process.

Acknowledgement

It has been a long stretch of rigour, but it was worth it. Having joined the program with no coding background and experience in data science, the challenging demand of the program has advanced my competencies and boosted my confidence for the career pursuits ahead.

However, I give gratitude to God for guiding me through those challenging times, and I thank myself for demonstrating resilience despite the numerous obstacles. My family also deserves credit for my success through this journey. Especially my father, a staunch advocate for quality education and upon whose immense support I continue to fly high.

Furthermore, I appreciate my supervisor for being very accessible throughout the program and during the thesis. His teaching method was considerate of the fact that the students possessed different levels of quantitative and programming backgrounds. Also, his soft-spoken demeanour and weekly guidance on the thesis established a warm atmosphere for me to reach out without hesitation when faced with challenges.

Finally, I thank my colleagues for finding me deserving of being their voice throughout the program. I also appreciate IFT for creating an avenue to share our opinions. The memories from this program would hang around for long.

Table of Contents

Abbreviations

AUM	Assets Under Management
BERT	Bidirectional Encoder Representations from Transformers
CAPM	Capital Asset Pricing Model
COE	Cost of Equity
ELMo	Embedded Language Modelling
EMH	Efficient Market Hypothesis
EPS	Earnings Per Share
ESG	Environmental, Social, Governance
GHG	GreenHouse Gas
GloVE	Global Vectors for Word Representation
LSTM	Long Short-Term Memory
MD&A	Management's Discussion and Analysis
MNLI	Multi Genre Natural Language Inference
NER	Named Entity Recognition
NLP	Natural Language Processing
ROA	Return on Asset
SEC	Securities and Exchange Commission
SQuAD	Stanford Question Answering Dataset
ULMFit	Universal Language Model Fine-Tuning
UN	United Nations
UNEP FI	United Nations Environment Programme Finance Initiative
WACC	Weighted Average Cost of Capital

1.0 Introduction

1.1 The Emergence of Sustainability in Finance

Sustainability is a term that spans across maintenance of the climate, land usage, water management and biodiversity. The concept of Sustainable Finance is not new, but it remains a fast-growing phenomenon in perhaps every aspect of finance. Putting a specific date and timeline for the start of its evolution may be contentious. However, it may be plausible to trace it to 1992, when the United Nations (UN) hosted the Earth Summit in Rio De Janeiro. Three decades ago, the UN launched the United Nations Environment Programme Finance Initiative (UNEP FI), aimed at birthing a financial system that supports the UN's sustainable development agenda through financing. Since then, notable initiatives have emerged within the financial services and capital market domain toward realising a sustainable environment. For instance, the Dow Jones Sustainability Index was launched in 1999, making leeway for investors seeking to identify and invest in companies that uphold sustainable business practices. Similarly, the UN launched the Principles for Responsible Investing in 2006, which serves as a guide for investment managers and aligns their investment choices to sustainable goals. As far as financial instruments are concerned, the European Investment Bank led the way in introducing a green bond in 2007; it was termed the Climate Awareness Bond.

Notwithstanding, one may wonder why finance is pivotal to the concept of sustainability. The relevance of finance in sustainability can be assessed from debt lending and equity investing perspectives. This is because they are businesses' primary sources of capital. For debt and equity, two critical considerations are the risk-return potential and intrinsic valuations. Hence, in a world where focus is increasingly given to climate change matters, investors have established that an institution's posture toward sustainability issues mirrors the attractiveness of such a firm. As such, business managers now agree that their response to climate change could make or mar their access to capital.

1.2 The Concept of Cost of Capital

Capital generally comes in two forms: debt or equity. Irrespective of what form of capital is favoured by a corporate, its managers must watch the cost of financing vis-à-vis its capital structure. In other words, a company needs to determine the optimal mix of capital that minimises its Weighted Average Cost of Capital (WACC). The WACC measures the cost of capital relative to the proportion of each source of financing to the overall capital structure. The WACC formula is given below.

$$\left(\frac{E}{V} * \ R_e\right) + \left(\frac{D}{V} * \ R_d * [1 - t]\right)$$

Where:
E = Equity value
D = Total debt outstanding
V = Equity value + Total debt outstanding
R_e = Cost of equity
R_d = Cost of debt
t = Tax rate (%)

The cost of debt is usually derived from the average rate on a company's existing debt while accounting for the tax benefit (1-t). The resulting tax shield points to why debt financing is cheaper. The cost of equity is the required rate of return by equity investors, and it also indicates the investor's assessment of a company's risk profile. Meanwhile, arriving at the cost of equity is not as direct as calculating the cost of debt. A company's R_e may be derived through two primary methods: Dividend Growth Model (DGM) or Capital Asset Pricing Model (CAPM). DGM applies to dividend-paying companies. It involves dividing a company's expected dividend per share by its price per share and adding a growth rate to the result. The growth rate in this instance, is the average growth in the dividend per share.

<u>Dividend Discount Model</u>

$$\frac{D_1}{P_o} + g$$

Where:
D_1 = Dividend Per Share over the next year

P_0 = Current market price of the stock
g = Dividend growth rate
NB: $D_1 = D_0 * (1 + g)$, where D_0 is the current Dividend Per Share

As indicated above, an alternative approach to the calculation of R_e is the CAPM. CAPM seeks to capture the inherent risk in a stock while adding a premium to compensate for taking such risk above a risk-free rate.

Capital Asset Pricing Model

$$R_f + \beta * (R_m - R_f)$$

R_f = Risk-free rate (typically a 10-yr treasury bill rate)
β = Measure of volatility/systematic risk
R_m = Market return

The foregoing illustrates the confluence between finance, sustainability, and a background to the cost of capital. In what follows, a background about Natural Language Processing (NLP), its emergence and its significance in finance will be assessed.

1.3 Natural Language Processing in Finance

Communication is in every aspect of human endeavour. Although the meaning of spoken or written words can be ambiguous, making its comprehension challenging for the audience. Hence, NLP emerged to deduce inferences from sentences, concisely summarise blocks of texts, and classify speeches into headings with the help of intelligent machines or software. Indeed, Johri et al., (2021) affirmed in their paper that machine translation emerged as one of the early aspects of NLP, where the aim was to automate programs that translate speeches from one language to the other. They further recorded that grammar could only be incorporated into machines from 1957 when Naomi Chomsky introduced *Syntactic Structures*. From there on, the field of NLP witnessed different shapes of evolution from the introduction of the idea of tokens in 1969 by Roger Schank to 1970, when Terry Winograd developed an NLP system that could respond to questions based on instructions provided by the user, and up to the 1980s when chatbots were first developed. However, worthy of note is that the early versions of NLP techniques were predominantly rule-based (symbolic techniques). A significant challenge that such techniques faced was the issue of context. Since the meaning of a word or speech largely depends on the context, the relevance and accuracy of those early NLP approaches were contested. Meanwhile, from 1983 onwards, most NLP researchers have become advocates of probabilistic models (stochastic techniques) in executing NLP tasks.

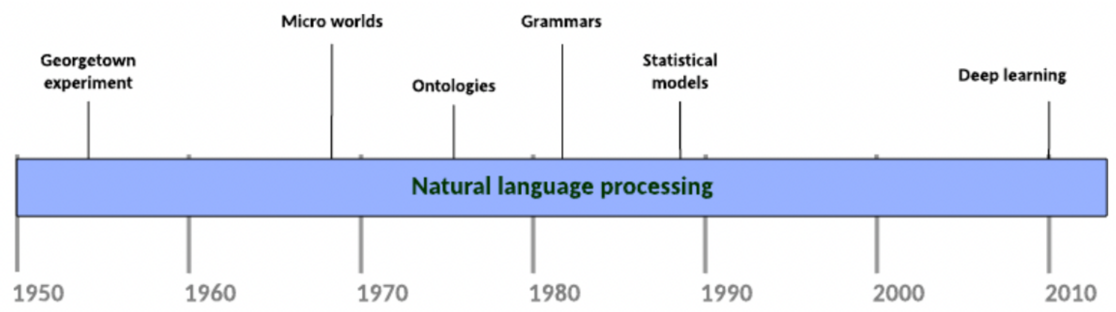


Fig 1.1: Timeline of Natural Language Processing (Jones, 2017)

NLP comprises seven popular techniques (listed below), all of which have their use cases in finance:

1. Keyword Extraction
2. Named Entity Recognition
3. Sentiment Analysis
4. Stemming and Lemmatisation
5. Summarisation
6. Text Classification
7. Topic Modelling

The size of unstructured financial data arguably outpaces that of structured data. Given its vastness, they are difficult to ignore. Especially in a fast-paced industry like finance and investments, where analysts continuously seek valuable signals from the slightest possible sources. Hence, NLP comes to the rescue in helping financial analysts derive the desired value from the unstructured data at their disposal. The relevance of NLP in finance spans risk analysis, trading, financial statement analysis, compliance, and client interaction.

As in the case of this research, where sentiment analysis is adopted to evaluate the associated climate risk of a company's operations, NLP can also be used to evaluate other forms of risk. For instance, a creditor could assess a debtor's likely credit attitude based on the behavioural information gathered from their social media activities. Similarly, Alvarado et al. (2015) assert that NLP provides an avenue to retrieve risk data using Named Entity Recognition (NER) methods on financial loan agreements.

The Efficient Market Hypothesis (EMH) posits that new information in the market is immediately absorbed into stock prices. It requires that traders act fast based on market news in their quest to generate alpha. However, given the enormity of financial news being released daily and globally, leveraging some NLP approaches will assist traders in summarising news contents. For instance, deriving sentiments from social media mentions of a stock, the resignation of a director or the acquisition of a company. Hence, traders can immediately make decisions based on the result of such NLP execution. Furthermore, NLP also demonstrates considerable relevance in conducting financial analysis. This is particularly useful during quarterly earning calls, where speech recognition techniques can transcribe speech into texts and derive emotion from the calls. Moreso, transcripts from such earning calls can be analysed for sentiment detection, and the result thereof can give an insight into a company's performance.

Just as within investments or trading domains, NLP also has its place in other areas of finance, such as retail banking. Client interaction is an avenue where NLP is predominantly adopted in retail banking. The advancement of chatbots in banking has evolved to enable client calls where callers speak with bots that could answer their queries verbally and transfer their calls to specialists when human intervention is necessary. The smooth verbal interaction between humans and bots demonstrates the sophistication of NLP models in decoding context and generating results.

1.4 Research Motivation

A Bloomberg article by Kishan (2022) revealed that sustainability-linked funds grew by 53% in 2021, with a global AUM of almost $3tn. Given the extreme climate conditions experienced globally, the transition to Net Zero continues to witness increased attention from investors, regulators, businesses, and other stakeholders. Finance plays a fundamental role in the transition toward limiting the global temperature to

1.5^{o}c. However, funding commitments made in fora like the COP26 are predominantly sourced from public finances. Meanwhile, private finance from institutional investors or lenders is also a formidable funding source. Hence, by their funding power, investors possess the prerogative to reward environmentally compliant companies with cheap capital or punish nonconformists and *greenwashers* with expensive capital. As such, this thesis firstly draws its motivation from the role investors could play in addressing climate change.

Furthermore, much of the earlier literature that has conducted related studies tends to lump ESG (Environmental, Social, Governance) as a single variable. Instead, this thesis is motivated by the severity of the prevailing climate change disasters to focus on only the 'E' element, i.e., environmental sustainability. Moreso, using ESG ratings is famous as an independent variable in previous studies. However, Berg et al. (2019) argue that ESG ratings are contentious. To contribute an alternative approach to the literature, this thesis seeks to adopt sentiment scores from sustainability reports instead of ESG ratings.

Lastly, the latest review period covered by most of the existing literature on similar subjects dates to 2019; this research will capture 2020 and 2021 data. This is to update the recency of the literature, given that recent events such as the COVID-19 pandemic and the prevailing energy crisis may have disrupted the historical rhetoric.

1.5 Research Objectives

This thesis investigates the direction and strength of the relationship between environmental/carbon reduction commitments and the cost of financing for companies operating in the emission-intense sectors or value chains in the United States. In addition, it seeks to establish if the cost of capital of such companies captures the climate risk inherent in their operations.

It also aims to discover the importance of sustainability reports in investments and capital allocation choices. Hence, this thesis aims to establish if debt and equity investors give a high weight to environmental factors in their investing criteria. The findings will indicate if investors reward sustainable companies with cheap capital or if a company's sustainability matters less.

1.6 Structure of the Thesis

Having discussed the background to this thesis, the motivations, and objectives, what follows will be a review of related literature in Chapter Two. In Chapter Three, the data used for the analysis will be discussed. This will cover the sources, the period, and the data manipulations that were done. In Chapter Four, the methodology adopted for this research will be evaluated. Also, the model used will be discussed alongside the steps taken to derive the result. In addition, the reasons why the thesis prefers the United States as the market of choice will be enumerated. However, Chapter Five will examine the results in detail while drawing inferences to possible reasons leading to each outcome. Graphical, tabular, and textual analysis will be adopted in this chapter. Chapter Six concludes the thesis, summarising the result and what can be concluded thereof. Likewise, ideas for future work will be mentioned.

2.0 Literature Review

2.1 Related Previous Work

The body of literature focusing on environmental sustainability and the broader ESG vis-à-vis its influence in finance and investment continues to expand. The scope of relevant papers represents a diverse attempt at studying ESG alongside numerous financial variables ranging from the accounting metrics of an institution to its public market valuation, risk metrics and other specific financial measures (e.g., revenue, cost of operations). Friede et al. (2015) aggregate all the research that studies how ESG relates to Financial Performance. Their findings revealed that almost 2,000 papers had been published on this subject, and the earliest research dates to 1982.

For a phenomenon that one may have considered nascent, having such an enormous volume of academic research lends credence to the importance of the topic. Therefore, additional studies must be conducted to unravel the many aspects of ESG in finance. The interest is not lopsided towards the academic realm but also the industry. Finance corporates and their clients/investors continue to direct funds to sustainability and social courses. As cited earlier, the total Assets Under Management (AUM) for ESG funds nears $3tn. A graphical illustration of the net flows and annual net assets is presented below.

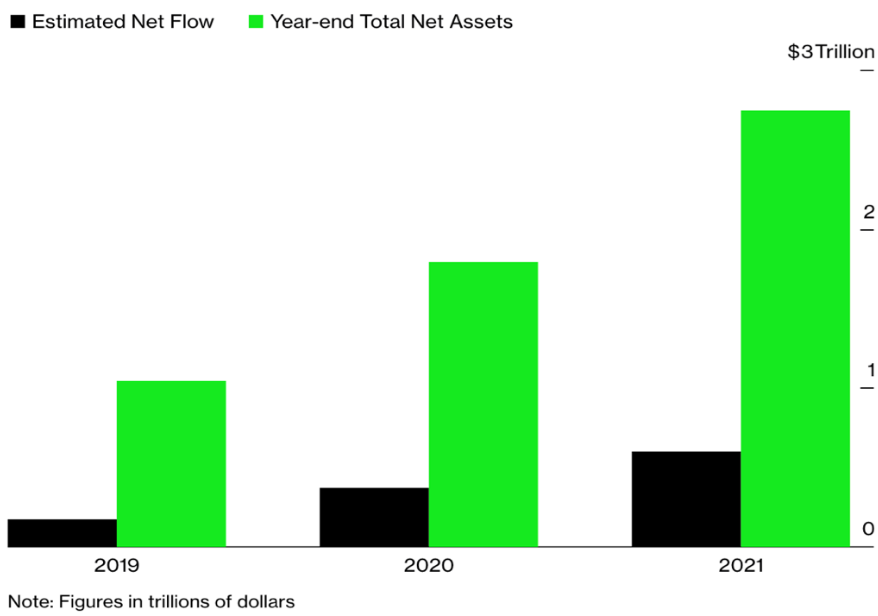


Fig 2.1: Sustainable Funds AUM and Net flows 2019 – 2021 (Kishan 2022)

Furthermore, one may be misled to think that the topic of ESG is predominantly a discourse relevant to only equities. Instead, the fixed-income market has also witnessed marked issuances of sustainability-themed instruments. The total issuance was $1.6tn in 2021 - an increase of 115.5% compared to 2020. Hence, the fact from the foregoing underscores the necessity to add to the body of ESG literature, especially as it deals with the cost of capital (cost of equity and cost of debt) of companies.

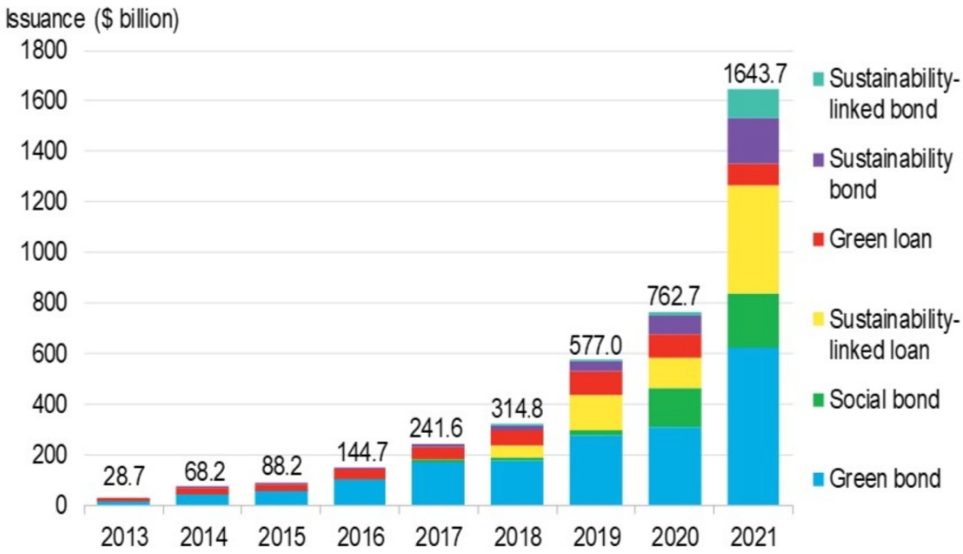


Fig 2.2: Issuance of Sustainable Funds 2013 – 2021 (Kishan 2022)

Expectedly, many related studies adopt different methodologies in their research. Similarly, authors focus on different countries, asset classes and any of the individual elements of E, S or G (or the three combined). Notwithstanding, of the 2,000 studies analysed by Friede et al. (2015), 90% of them concluded that ESG has a non-negative relationship with financial performance. The following paragraphs will review papers that have adopted different methodologies in studying how ESG relates to different financial metrics. Likewise, other literature that adopted Natural Language Processing (NLP) techniques, although a handful of them exists.

Whelan et al. (2021) conducted a meta-analysis of 1,000+ papers that studied the relationship between ESG and financial performance between 2015 and 2020. In their methodology, they split the literature between those focusing on investment-related variables (e.g., Sharpe ratio), papers studying corporate financial performance (e.g., ROE), and those focusing on carbon-related factors in relation to financial performance. Their findings were consistent with Friede et al. (2015). They established that 59% of the reviewed literature focusing on investment variables (e.g., Return on Equity) conclude on a positive relationship. In comparison, 58% of those using corporate financial performance variables (e.g., Net Profit) have positive outcomes.

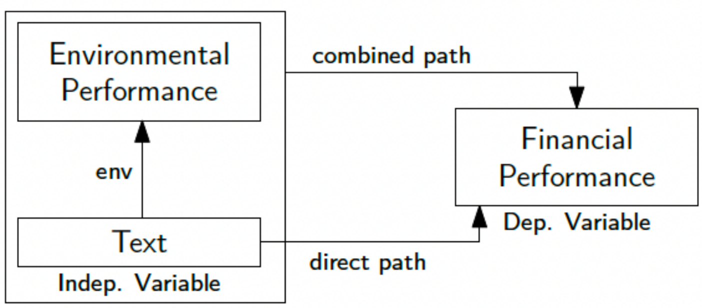
In their research, Ramirez et al. (2022) focus on Latin America (LatAm) to investigate the relationship between the cost of capital and ESG. The study covers 202 companies across 10 LatAm countries between 2017 and 2019. Their approach first considers the three ESG elements combined as a single variable. Thereafter, each of the three elements is separated, leading to four independent variables assessed against the cost of capital. Their findings reveal no relationship between the E and S criteria with the cost of capital but established that the G criteria demonstrated a negative relationship. Similarly, when taken as a whole, ESG has an inverse relationship with the cost of capital of public companies in LatAm.

Like Ramirez et al. (2022), Pellegrini et al. (2019) investigated the impact of ESG scores on the cost of equity and profitability of 182 sample firms in the oil and gas industry between 2002 and 2018. In their analysis, Return on Asset (ROA) was adopted as the measure of profitability and was compared with ESG scores derived from Thomson Reuters. However, their computation of the Cost of Equity (COE) did not follow the conventional CAPM approach. Instead, they adopt the Easton Model, where the cost of equity is implied from a firm's share price and Earnings Per Share (EPS). Their regression analysis delivers a result that is consistent with the literature. They find that for a 10% rise in ESG score, a company's COE declines by 1.34%. However, their study notes that the relationship between both variables is non-linear and U-shaped. They assert that the relationship between COE and ESG score remains negative up to a threshold determined by company size (measured by total assets). Thereafter, it turns positive. Meanwhile, their test of the relationship between ESG scores and ROA of oil and gas firms demonstrates a negative relationship. According to their findings, a 10% rise in ESG score causes a 0.45% decline in profitability.

Furthermore, Atan et al. (2018) conducted a study of 54 companies between 2010 and 2013 to investigate the impact of the combined and individual ESG factors on the performance of public companies in Malaysia. In their analysis, they took the Return on Equity to represent profitability, Tobin's Q was used a measure of the firm's value and WACC for cost of capital. Their approach was to conduct a panel regression between the four independent variables (ESG, E, S, G) and each of the three dependent variables by applying three statistic panel models – Random Effects, Pooled OLS, and Fixed Effects. However, unlike Pellegrini et al. (2019), who sourced their data from Thomson Reuters, the authors rely on the Bloomberg ESG database. Their analysis reveals that the combined and individual ESG scores do not influence Malaysian companies' ROE and Tobin's Q. Although the individual scores of E, S, and G do not also influence the WACC of Malaysian companies, they found that the combined ESG score does.

One would observe that all the literature cited thus far adopt conventional statistical approaches in their analysis, understandably so because most of the study on this subject focus on financial information. This indicates the neglect of textual or unstructured data and the dearth of advanced computational techniques in the literature as far as the subject of this thesis is concerned. Notwithstanding, the works of some authors (although few) that have adopted NLP techniques on similar topics will also be reviewed.

Felix et al. (2020) examined how the environmental and financial disclosures in financial statement reports demonstrate value to the investors vis-à-vis the company's financial performance. To achieve this, they train a deep learning model based on BERT (Devlin et al., 2018) and GLoVE embeddings (Pennington et al., 2014) to conduct a textual analysis of the Management's Discussion and Analysis (MD&A) section of annual and quarterly reports. Hence, they sought to determine if environmental disclosures contained in the MD&A section of company filings serve as a mediator for financial performance. The approach of their study is illustrated below.



<u>Fig 2.3: Schema of Study Approach by Felix et al. (2020)</u>

They leverage their NLP model to extract texts from the MD&A section. Having trained the model on a sentence and word basis, they expect it to predict financial performance based on the extracted texts (this refers to the direct path). Afterwards, the company's environmental performance is predicted from the extracted texts (env path). The two paths (direct and env) are combined to predict financial performance. If the combination of the paths results in a high F1 score, it will be inferred that environmental narratives complement financial narratives to predict financial performance. Upon analysing 554 US companies between 2014 and 2018, they concluded that textual data (financial and environmental) from the MD&A section of company fillings is not a good predictor of financial performance. Albeit they noted that there is evidence that NLP techniques can be used to determine the environmental performance of companies based on extracted texts from the MD&A section.

Given the potential bias inherent in companies' self-reported and often unaudited disclosures, Sokolov et al. (2021) attempted to develop an ESG index. They also adopted BERT and trained it with manually classified Twitter datasets. They sought to automate the classification of documents based on their contents under ten different sub-headings under E, S and G (Table 2.1). 1,468 unique labelled tweets were used in training the model, complemented with tuning hyperparameters and adding a hidden layer while adopting a binary cross entropy loss function as the classifier. In evaluating the performance of their model, the area under the ROC curve alongside the Precision-Recall curve was adopted and calculated for each classification, resulting in an average of 94.5.

Environmental	**Social**	**Governance**
Climate Change	Discrimination	Business Ethics
Product Safety	Health & Demographic Risk	Anti-Competitive Practices
Toxic Emissions & Waste	Supply Chain Management and Labour Standards	Corruption and Instability
	Data Security	

Table 2.1: ESG Index Classification (Sokolov et al., 2021)

One of the fascinating works concerning this subject was by KPMG in 2019[1]. They wanted to determine if a client's sustainability report reflects the positive and negative side of their environmental performance so that investors can make a sound decision. However, to appropriately assess the balance of a report, it needs to be measured quantitatively and objectively to prevent human subjectiveness. KPMG tackles this problem by gathering about 779 sustainability reports, and their methodology is split into three phases – i) Text extraction ii) Subject extraction iii) Sentiment analysis. They utilised the `PDFminer` package in python to extract the texts and perform the required text cleaning. They adopted unsupervised topic modelling to derive the summary subjects from these reports using Latent Dirichlet Allocation (LDA). LDA is a statistical model that extracts topics from a group of texts and models the topics as a multinomial distribution of words. However, one critical challenge is that such reports, even if predominantly negative, tend to be worded positively or in a more nuanced language that makes them sound neutral. Their first step was to manually label some sentences from the sample reports as either positive, negative, or neutral. Subsequently, they utilise 6 NLP tools or models[2] to predict the sentiments from these sentences to see if they agree with the manual labelling. The models' predictions were markedly divergent.

A major factor that may have caused the divergent predictions is these models' poor knowledge of context. To address this, the KPMG team also adopted the BERT model. The model was initially fine-tuned with manually labelled 3000 sentences and, upon testing, achieved good accuracy (Table 2.2).

Sentiment	Accuracy
Negative	71%
Neutral	94%
Positive	80%

Table 2.2: Prediction Accuracy of Fine-tuned KPMG BERT Model

Since the fundamental problem at hand is to objectively measure the balance of the sustainability reports of their clients, their concluding approach was to benchmark the

[1] Susanne Groothuis: Sentiment Analysis on Sustainability Reporting

[2] Heaven on Demand, Rosetta, Text-processor.com Stanford Sentiment Tree Bank, Treeblob, and Neural Network trained on 5000 movie reviews

sentiment score of a client's report against its peers. Hence, after deriving a sentiment score from a client's report, KPMG suggests to the client whether it needs to adjust its reporting based on the average sentiment score of its competitor's sustainability reports.

Having examined some of the relevant literature, it is apparent that a deep learning approach tends to have more credibility, the popular being BERT. There are mixed findings regarding the relationship between ESG and other financial variables. Regardless, as cited earlier, Whelan et al. (2021) find that about 58% of the literature agrees on a positive relationship. It is on the premise of these findings that a BERT-based model is favoured for this thesis, and the hypotheses are enumerated thus:

H_1: *There is a negative relationship between sentiment scores from sustainability reports and WACC*

H_2: *Sustainability reports are very influential in the investment decisions of debt and equity investors*

3.0 Data

3.1 Data Description

The fundamental data utilised in this study was the WACC and texts from companies' sustainability reports. The thesis focuses on the top 5 Greenhouse Gas emitting sectors in the US, according to the US Environmental Protection Agency (US EPA)[3]. These sectors and their respective contribution to the total emissions are illustrated below (the latest data at the time of this study was for 2020).

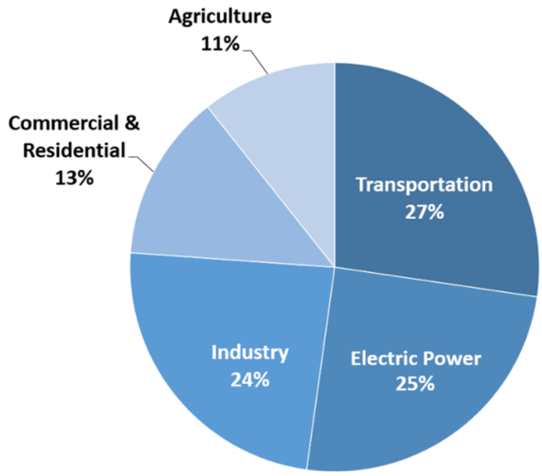


Fig 3.1: Total US Greenhouse Gas Emissions by Economic Sector, 2020

Following the above sector rankings, 34 companies were selected from all the sectors. This translates to 170 sustainability reports that were manually gathered. It should be noted that different companies adopt different titling for these reports. While some explicitly referred to theirs as Sustainability Report, others used terms such as ESG Report, CSR Disclosure, and the like. Notwithstanding, the contents of each report were vetted manually to confirm that they contained the kind of information relevant to this study. The details of the company selection method will be discussed in section 4.6. Meanwhile, the companies captured under this analysis are given below, alongside their corresponding WACC across the review period.

[3] EPA - Sources of Greenhouse Gas Emissions

Industry	Company	Ticker	Weighted Average Cost of Capital (%)				
			2017	2018	2019	2020	2021
Electric Power	Chevron Corp	CVX	6.94	9.00	7.27	7.92	10.04
	NextEra Energy Inc	NEE	4.41	5.83	4.52	6.36	7.94
	ConocoPhillips	COP	9.02	8.58	6.87	7.35	9.35
	Duke Energy Corp	DUK	3.94	4.74	3.57	4.95	6.05
	Southern Company	SO	4.34	4.60	3.74	5.61	6.89
	EOG Resources Inc	EOG	7.76	9.55	8.21	7.90	9.82
	Pioneer Natural Resources Co	PXD	8.17	10.25	8.54	8.12	9.31
	Dominion Energy Inc	D	4.70	5.00	3.93	4.95	5.90
	Schlumberger NV	SLB	7.41	8.58	8.00	7.04	9.16
Agriculture	Mondelez International Inc	MDLZ	8.15	7.52	6.40	6.01	7.07
	Archer-Daniels-Midland Co	ADM	7.51	7.78	6.18	6.15	7.64
	Ecolab Inc	ECL	8.11	8.96	6.97	8.79	10.2
	General Mills Inc	GIS	6.43	7.71	6.02	3.61	4.60
	Hershey	HSY	6.88	6.99	5.22	6.39	7.76
	Hormel Foods Corp	HRL	6.15	7.19	5.79	3.87	4.43
	Tyson Foods Inc	TSN	5.40	6.63	5.78	5.14	7.22
	Kellogg Co	K	5.45	6.36	5.04	4.55	5.29
Industry	Caterpillar Inc	CAT	8.07	9.92	7.83	6.05	7.09
	3M Co	MMM	8.20	11.09	8.16	6.56	7.45
	Illinois Tool Works Inc	ITW	9.37	9.62	8.03	8.25	9.49
	Freeport-McMoRan Inc	FCX	12.73	7.83	7.15	10.13	12.82
	Air Products and Chemicals Inc	APD	8.38	9.17	7.57	7.06	8.64
Commercial and Residential Real Estate	American Tower Corp	AMT	6.96	7.26	5.20	6.06	7.27
	Prologis Inc	PLD	7.08	7.64	6.45	8.21	10.10
	Equinix Inc	EQIX	6.02	6.90	5.38	6.54	7.79
	Welltower Inc	WELL	5.71	5.77	4.35	8.24	10.67
	Simon Property Group Inc	SPG	6.30	6.14	4.78	7.81	11.83
Transportation	Union Pacific Corp	UNP	7.72	9.38	7.32	7.67	8.92
	Lockheed Martin Corp	LMT	6.32	8.63	6.88	7.36	8.55
	Boeing Co	BA	9.48	10.89	7.36	11.64	13.76
	Northrop Grumman Corp	NOC	5.94	7.91	6.45	5.61	6.80
	CSX Corp	CSX	8.80	9.83	7.44	7.24	8.55
	Tesla Inc	TSLA	12.70	8.56	6.92	12.39	15.13
	FedEx Corp	FDX	9.01	10.66	9.17	4.70	7.64

Table 3.1: List of Companies, Industry and WACC

3.2 Data Sources

While the WACC data was obtained via the Bloomberg terminal, the sustainability reports were retrieved from the company's website. When the report for a specific year is no longer on the company's page, ResponsibilityReports – a repository of 11,448 ESG reports[4], serves as an alternative source.

3.3 Data Manipulation and Pre-processing

It was observed that some companies did not start issuing sustainability reports until 2018 (e.g., NextEra); hence they had no 2017 reports. Similarly, some may not have published a report for a particular year (e.g., Southern Co does not have a 2019 report) or were yet to release their 2021 report when this analysis was conducted. In those instances, since such companies published reports for at least four years within the review period, the sustainability report of the missing year was replaced with the annual report as a proxy. Regarding data pre-processing, the first step was utilising `pdfplumber` to extract text from the pdf files. Thereafter `punkt` (a sentence tokeniser from the `nltk` package) was downloaded to separate the extracted texts into sentences. Since some regular expressions are expected to be present in the extracted texts, a function was defined to decode utf-8 characters and replace them with a space. In addition, punctuation marks were eliminated, and the texts were converted to lowercase. The cleaned sentences were stored in a dataframe having three columns (Company, Year, Sentence), and each row represents an extracted sentence.

[4] ResponsibilityReports Homepage

4.0 Methodology

4.1 A Theoretical Background

Conducting sentiment analysis can prove to be a complex endeavour for several reasons. Two of such primary reasons include (1) when there is a need to capture context and semantics and (2) when the problem is domain-specific and would require special attention to domain-specific vocabularies. In such instances, determining the polarity from a body of texts will not be straightforward and would require further manipulation of the existing techniques or perhaps introducing a novel approach. In considering a similar challenge, FinBERT emerged (Araci, 2019). FinBERT is a pre-trained sentiment analysis model tailored for the finance domain. It relies on the BERT model by training it with a large financial corpus sourced from Reuters financial news articles, Financial PhraseBank and FiQA sentiment dataset. However, for an accurate description of the implementation of FinBERT, it is necessary to first describe the architecture of BERT – the foundation upon which FinBERT was developed.

4.1.1 BERT - Background

BERT is a deep language model that seeks to address the shortcomings of standard language models, which are predominantly unidirectional. Unidirectional models evaluate inputs only from left to right; hence, only the tokens that appear before the target sequence of tokens are evaluated. Therefore, this limits the capacity of such models to conduct tasks such as question answering. This is because context is required by evaluating the tokens from left to right and vice versa. Since contextual understanding is crucial, BERT addresses the challenge by introducing MLM and NSP[5]. MLM can be likened to a fill-in-the-gap exercise, whereby a sentence is input into the model, but a portion of it is concealed, and the model is tasked to predict the concealed portion, having learned the context in which the word appears.

[5] MLM – Masked Language Model, NSP – Next Sentence Prediction

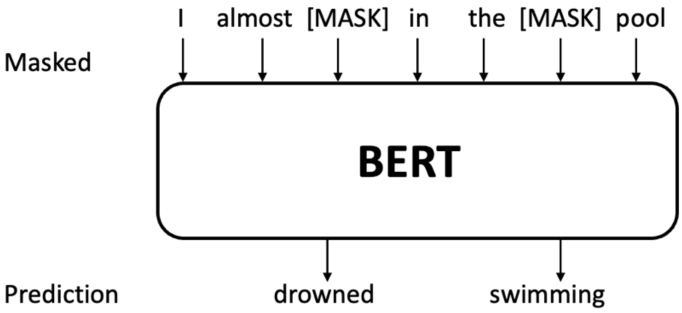


Fig 4.1: Simple illustration of MLM

Meanwhile, unlike MLM where the objective is to decipher word-to-word relationships, NSP seeks to detect the relationship on a sentence level. In other words, given sentence A, the model is tasked to predict sentence B.

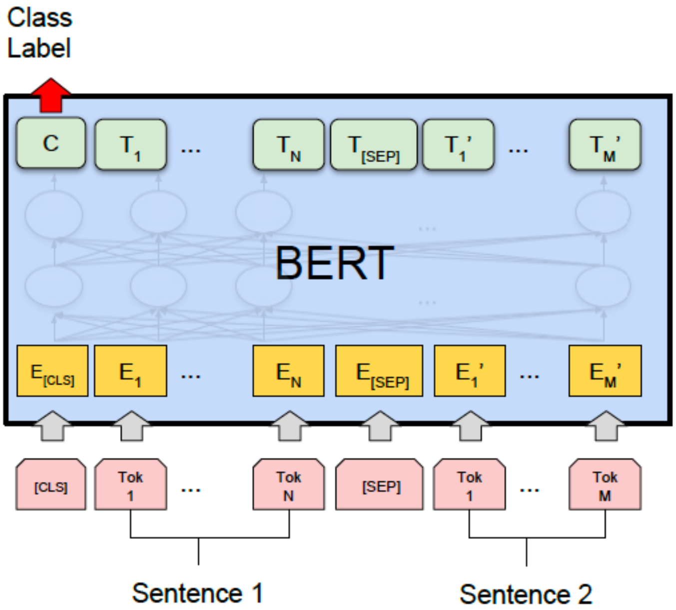


Fig 4.2: Illustration of BERT NSP Architecture (Devlin, Chang et al. 2018)

BERT is firstly pre-trained with unlabelled data (it was trained with 3.4mn words from Wikipedia and BooksCorpus), and subsequently fine-tuned with labelled data to configure the model for downstream tasks, e.g., tagging, question answering and sentiment analysis.

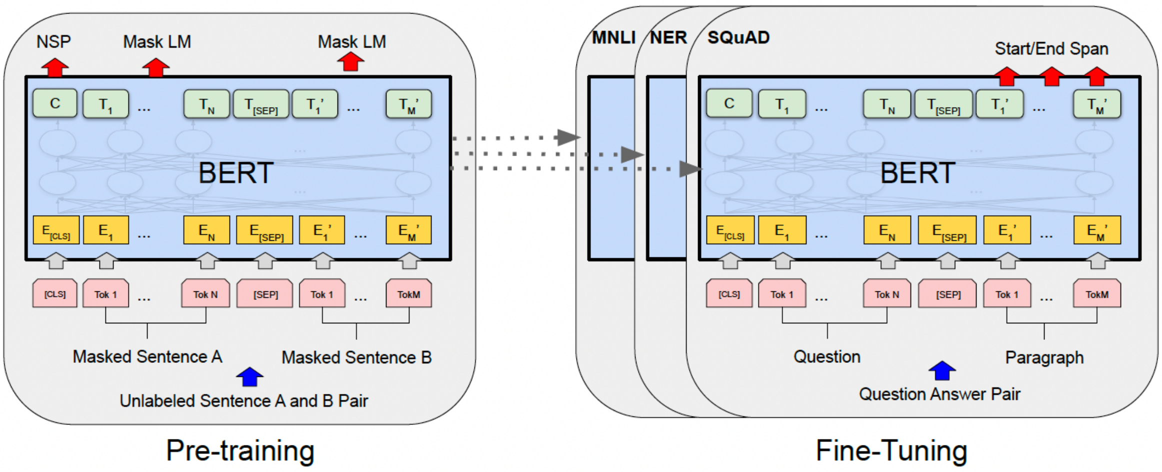


Fig 4.2: BERT Pre-training and Fine-tuning (Devlin et al., 2018)[6]

From the figure above, the pre-training process involves MLM and NSP discussed earlier. At the fine-tuning stage, the same architecture and parameters used in the pre-training are maintained (except the output layer), and the model is trained with labelled data.

4.1.2 FinBERT - Background

Notwithstanding the pre-training of BERT with BookCorpus and Wikipedia corpus, FinBERT includes an additional layer of corpus. Reuters TRC2 was used for the pre-training and Financial Phrasebank for the fine-tuning (Fig 4.3).

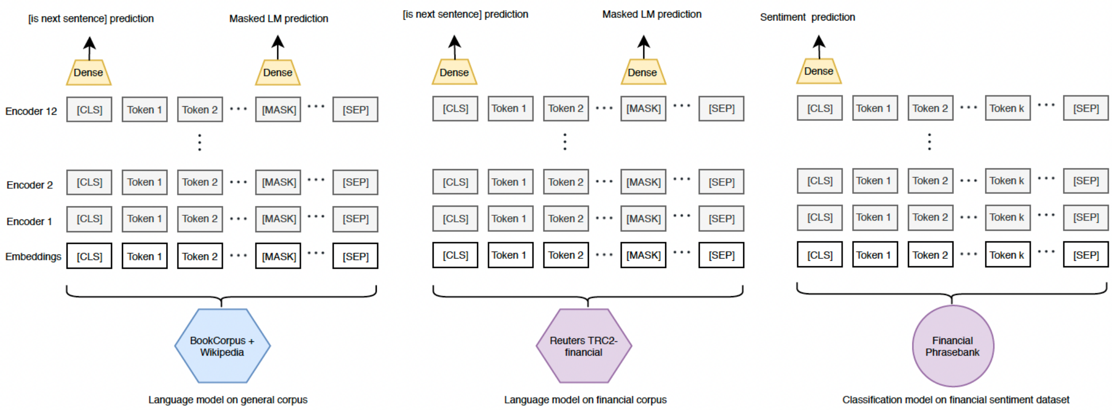


Fig 4.3: Pre-training and Fine-tuning process for FinBERT (Araci, 2019)

[6] CLS – Classification token, SEP – Separator token

The TRC2-Financial is a subset of TRC2 from Reuters made of 1.8mn news content published between 2008 and 2010. Following the filtering for financial vocabularies, the resulting corpus came to 46,143 news articles representing 29mn words and 400k sentences. Since the downstream task in the case of FinBERT is to conduct sentiment analysis, the model uses the Financial PhraseBank. It contains 4,845 sentences (60% was used as train set), manually classified as either positive, negative, or neutral.

In preparing the model for a classification exercise, a dense layer was added upon the hidden layer of the CLS token. Furthermore, one of the fundamental steps when fine-tuning was to design the model to avoid catastrophic forgetting (CF). Arora et al. (2019) describe CF as a situation where a model trained on one specific task loses its generalisation ability when it is fine-tuned on a second task since it has already lost the knowledge of the first task. Not guarding against CF will prevent the model from a continuous learning process since the previous information is lost upon learning something new.

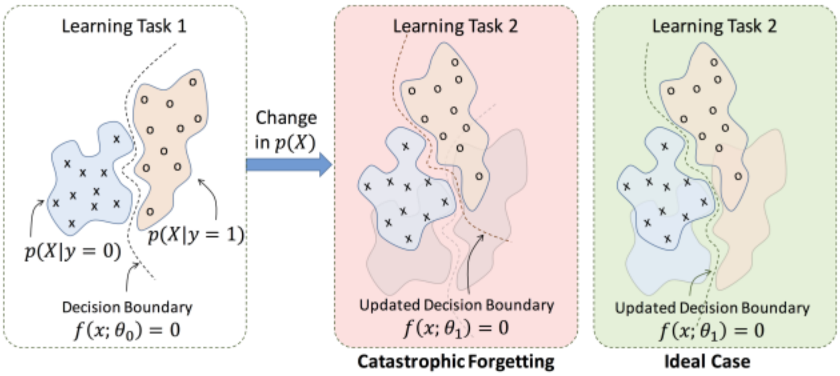


Fig 4.5: Catastrophic Forgetting in a Binary Classification Task (Kolouri et al., 2019)

By following the approach of Howard and Ruder (2018), the three techniques adopted to address the likelihood of CF are described below:

- **Gradual Unfreezing**: This involves the unfreezing (training) of each layer of the model starting from the last layer, and each unfrozen layer gets trained by the target task data with one iteration. The process continues backwards from the last layer up to the first layer.

- **Discriminative Fine-tuning**: This is based on the premise that since each layer of the model holds a different level of knowledge (with the lower layer having lower information), each layer should be tuned with a different learning rate. The parameters for discriminative fine-tuning can be inferred from the Stochastic Gradient Descent Formula:

$$\theta_t^l = \theta_{t-1}^l - \eta^l \,.\, \nabla_{\theta^l} J(\theta)$$

Where θ^l is the model at the l^{th} layer and L is the total layers in the model. η^l is the learning rate of the l^{th} layer and $\nabla_{\theta^l} J(\theta)$ is the gradient of the model's objective function. Having set a learning rate for the last layer, i.e., η^L, the learning rate of the other layers can be computed as $\eta^{l-1} = \eta^l/2.6$

- **Slanted Triangular Learning Rate**: When conducting task-specific training, we desire to have the model converge to a region of the parameter space at the start of the training. As such, the idea of STLR is to increase the learning rate as the epoch increases, up to a point when it is reduced gradually. The iteration point from where the learning rate is reduced is given as:

$$c = T \,.\, c\_fraction$$

T = (number of epochs x number of updates per epoch) and $c\_fraction$ is the fraction of iteration at which we increase the learning rate.

The fraction of the number of iterations to be increased or decreased is given as:

$$b = \begin{cases} \dfrac{t}{c}\,, & if\ t < c\ otherwise; \\ \\ 1 - \dfrac{t - c}{c\,.\,(\dfrac{1}{c\_fraction} - 1)} \end{cases}$$

The learning rate at iteration T is given as:

$$\eta_t = \eta_{max} \,.\, \frac{1 + b\,.\,(ratio - 1)}{ratio}$$

Where η_{max} is the maximum learning rate, and ratio is the difference between the lowest and maximum learning rate i.e., $\eta_{max} - \eta_{min}$

4.2 Accuracy of FinBERT

The FinBERT's accuracy was tested in two phases (Pre-training and Fine-tuning) and was based on three metrics – Macro F1 score, Accuracy score and the Cross-entropy loss. In the pre-training phase, these metrics compare in favour of FinBERT (Table 4.1) when assessed against LSTM, LSTM ELMo and ULMFit models, which were all trained similarly as the FinBERT. However, to substantiate the comparison, the size of the training set was gradually increased for each model. Based on this, a minimum of about 250 train size is required for the FinBERT and ULMFit to start performing well (Fig 4.6). However, LSTM still performs poorly despite a larger train size.

Model	Loss (%)	Accuracy (%)	F1 Score (%)
LSTM	81	71	64
LSTM with ELMo	72	75	70
ULMFit	41	83	79
FinBERT	37	86	84

Table 4.1: Accuracy of Pre-trained FinBERT (Araci, 2019)

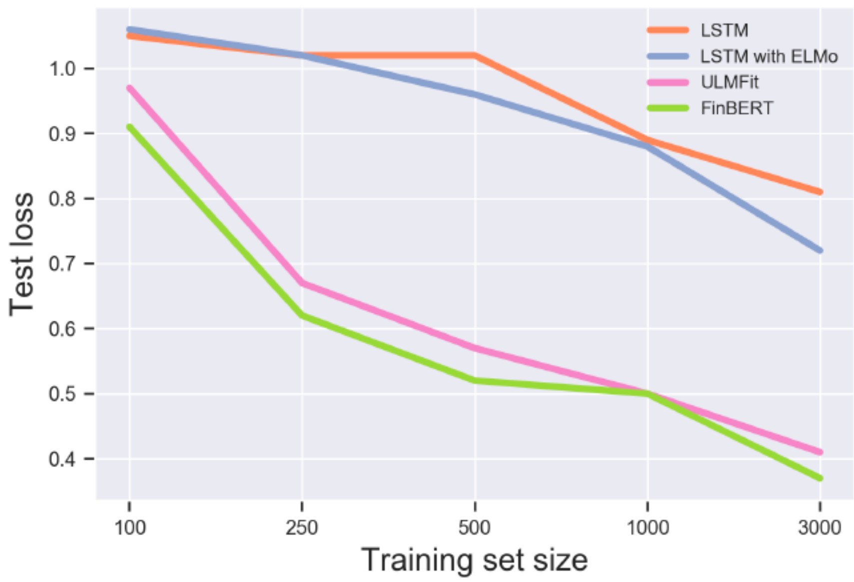


Fig 4.6: Cross-entropy Loss on Pre-training Set with Different Models (Araci, 2019)

Meanwhile, not much improvement was seen in the performance upon further fine-tuning using the Reuters RC2 for the downstream task. At this phase, the fine-tuned FinBERT model (accuracy – 86%) was compared with the default BERT (accuracy – 85%) and the FinBERT without fine-tuning (accuracy – 84%). The author posits that the unchanged performance is most likely due to a very good performance at the pre-training stage, such that there is no longer room for further improvement.

4.3 WACC as the Metric of Choice

WACC is a popular measure of the cost of capital and a relevant metric to multiple stakeholders. Meanwhile, its relevance is further elevated among investors keen on environmental sustainability. The choice of WACC as the dependent variable for this thesis is based on three motivations that are discussed below.

- **An indication of the climate risk inherent in a company's operations**: Due to the increasing attention to climate change effects, stakeholders are now looking closer at how a business's contribution to the climate may portend a risk to its operations. Hence, a company with poor environmental controls may risk legal consequences, regulatory fines, or bad publicity, which may transcend into dwindling revenue and future cash flow. Considering these potential risk factors, investors will be mindful of capturing the inherent climate risks when estimating a company's WACC towards determining the appropriate discount rate for valuation purposes. Moreover, if these climate risks are high, the investors will be compelled to demand a higher return for their investment.

- **A good metric to compare industry peers**: When comparing companies' sustainability credentials and its impact on financial performance, the WACC is a metric that can be adopted for such comparative analysis. Unlike other absolute metrics such as revenue or net profit, which the size of the company may distort, using the WACC as a comparative tool for industry peers puts the companies on a level playing field. Hence, investors can adequately evaluate their opportunity costs when assessing multiple investment opportunities.

- **A valuable tool for investors to control how companies respond to environmental sustainability**: Investors have now adopted moral suasion (otherwise termed 'active engagement') to persuade and support their investee companies to improve their sustainability scorecard. However, this method may take a long time before we begin seeing tangible changes. Meanwhile, investors can choose to exploit the WACC by demanding higher returns on debt or equity capital when polluting companies approach the capital market for financing. Hence, when financing becomes expensive for such companies, they could promptly consider improving their environmental profile.

4.4 Incorporating FinBERT with Cost of Capital

The choice of FinBERT hinges on the fact the dependent variable is a financial metric. Moreso, the model was trained with texts from financial reports. Meanwhile, financial reports do not only contain finance-related texts, they now also feature snippets of sustainability reports.

Furthermore, the existing alternatives fall short in several areas. Notably, the strength of the BERT model comes from its ability to randomly predict masked tokens irrespective of its position embedding, as opposed to being constrained to predicting just the next token. Moreover, Araci (2019) opines that due to the dearth of a large labelled dataset for the financial domain, the full potential of a neural network cannot be utilised for sentiment analysis.

Aside considering a deep learning method, the lexicon approach may have been considered. However, as indicated earlier, context remains crucial when conducting sentiment analysis. As a result, relying on just a dictionary-based approach that merely classifies a word into positive or negative will render using a lexicon approach grossly ineffective for our purpose.

Moreover, the choice of FinBERT also becomes compelling due to the absence of a comparable BERT model that uses a sustainability-focused corpus for fine-tuning at the time of writing this thesis.

4.5 The United States as the Economy of Choice

The economy of choice for this thesis is the United States of America. The consideration of data from public US companies in conducting this analysis was motivated by the following reasons:

- **Availability of data**: Given that the US is a major player in the global financial market and a key stakeholder in global climate discussions, easy access to relevant data is most probable in a market of this kind.

- **Largest and most developed capital market**: The US market has become a guidebook for other financial markets, even among the developed nations. Hence, the credibility of reported data is assured, thereby instilling a layer of confidence in the result of the analysis.

- **The country with the second highest global emissions**: The US comes after China regarding global GHG emissions (4,747mn mt)[7]. Hence, its massive contribution to global emissions makes it a convincing case study for this thesis.

- **Better environmental disclosure**: Public companies in the US are increasingly becoming open in terms of the environmental impact of their activities. Indeed, the US SEC proposed a rule in March 2021[8] which will mandate companies to report climate-related financial metrics in their 10-K reports. This aims to ensure the comparability of reports and help to check greenwashing.

4.6 Sampling Method

As previously commented, the thesis focuses on the top five GHG-emitting sectors in the US according to rankings by the US EPA. The approach to the company selection was to focus on the top 10 companies by market capitalisation by the end of July 2022. The list of companies under each sector was derived from Bloomberg, and the GICS sector classification[9] was adopted. However, the industry nomenclature used by the US EPA is slightly different. As such, minor adjustments were made to synchronise GICS and US EPA nomenclature towards ensuring that the list of companies is comprehensive, and some critical categories of institutions are not omitted. For instance, following the adjustments, Oil and Gas companies and Power/Energy firms were captured under the heading of 'Energy/Electric Power'.

Also, it was earlier commented that out of the top ten companies with a large market cap, only those with at least four years of sustainability reports between 2017 and 2021 were considered. Their annual financial report replaced the missing year's report. The decision to focus on the top 10 largest companies hinges on the fact that those sets of firms typically make up at least 50% of their overall industry market

[7] GreenHouse Gas Emission by Countries
[8] Enhancement and Standardization of Climate-Related Disclosures
[9] GICS Methodology

capitalisation. Meanwhile, each sector must have at least five sample companies. Otherwise, the next largest company would be considered in any sector in case there are less than five companies that have 4years+ of sustainability reports.

4.7 Deriving the Sentiment Score

Following the data manipulation procedure described in section 3.3, the next step was to deploy the FinBERT model to derive sentiment scores from the cleaned extracted texts. The model was cloned from the ProsusAI GitHub repository (see Appendix), and the path of the extracted texts for each industry links to a Google Drive (connecting the textual data files can be easily synchronised from here since the code was written via Google Colab). The prediction of the sentiment score[10] begins with tokenising the extracted texts into sentences by utilising the `sent_tokenize` package from `nltk`.

```
sentences = sent_tokenize(text)
```

The sentences are then split into batches, and the list of batch sentences is converted into features. The conversion is done with a user-defined function below.

```
def convert_examples_to_features(examples, label_list,
max_seq_length, tokenizer, mode='classification')
```

where;

`examples` = a list of the batched sentences/inputs
`label_list` = ['positive', 'negative', 'neutral']
`max_seq_length` = the maximum sequence length (64 was used in FinBert)
`tokenizer` = BertTokenizer

Each input gets an ID, attention mask and a type of input id (defined below). These three parameters are passed into the model to generate the logits and converted to probabilities by using the softmax activation function. The sentiment score is thereafter derived as the positive probability minus the negative probability.

```
logits = model(all_input_ids, all_attention_mask,
all_token_type_ids)[0]
logits = softmax(np.array(logits.cpu()))
sentiment_score = pd.Series(logits[:, 0] - logits[:, 1])
```

[10] Source: Line 581 – 625 of the GitHub FinBERT code demonstrates the prediction of the sentiment score

where;

`all_input_ids`: the numerical identity of each token
`all_attention_mask`: the zero-padding or truncation of the tokens to enable the model to decide which token to give attention to or ignore
`all_token_type_ids`: indicates if it is a CLS token or SEP token

Upon the deriving the sentiment scores per sentences in each sustainability report, we calculate the average of the sentiment scores. This is in line with the approach adopted by FinBERT (see Appendix for the link to the FinBERT GitHub repository). The resulting mean value is taken as the representative sentiment score for each report in a particular year (see Appendix for GitHub to thesis codes, containing the sentiment scores per company).

5.0 Results

5.1 Discussion of Results

A review of the predicted sentiment labels shows that majority of the sentences have a neutral tone (see Table 5.1). Upon inspecting the neutral sentences, it was observed that most of them lacked context, e.g., *24% reduction from a 2005 baseline*. Moreso, such sentences tend to be cliché because they appear multiple times in a company's sustainability report. For example - *Our vision of Building America involves protecting and strengthening this foundation*, appeared in the report of Union Pacific between 2017 and 2019. Given the foregoing, neutral sentences were taken as noisy data, which may distort the result. Hence, the sentiment scores of neutral sentences were ignored, only positive and negative sentences were considered in the analysis[11].

	Positive	Negative	Neutral
Agriculture	7,396	1,213	20,850
Electric Power/Energy	10,652	2,310	33,992
Industrials	6,025	1,031	19,762
Transportation	7,373	1,162	20,723
Real Estate	2,645	314	8,160

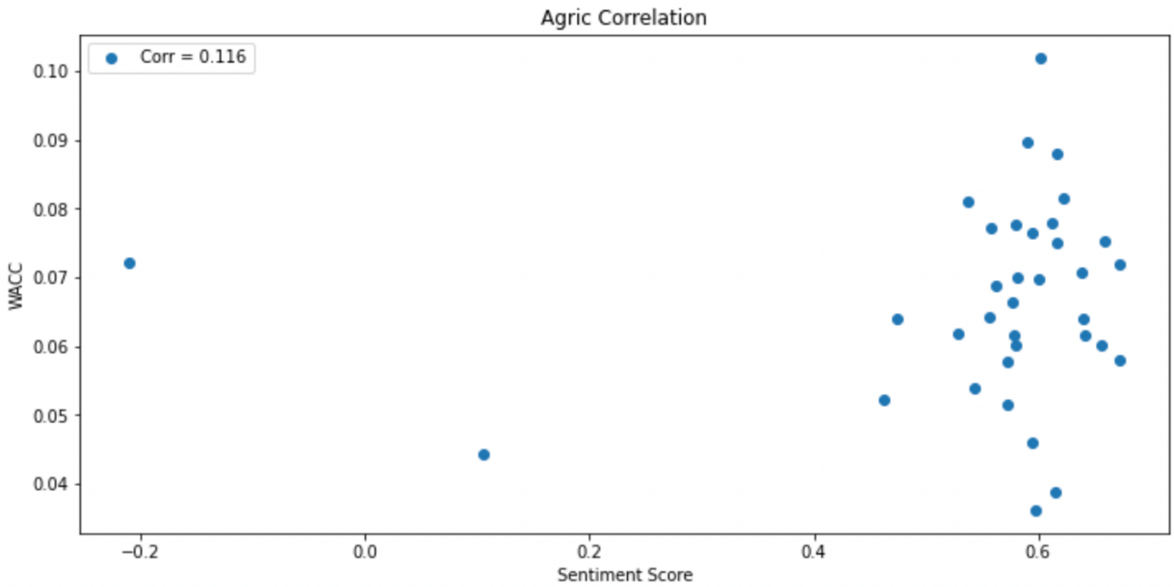
Table 5.1: Sentiment Distribution of Extracted Texts

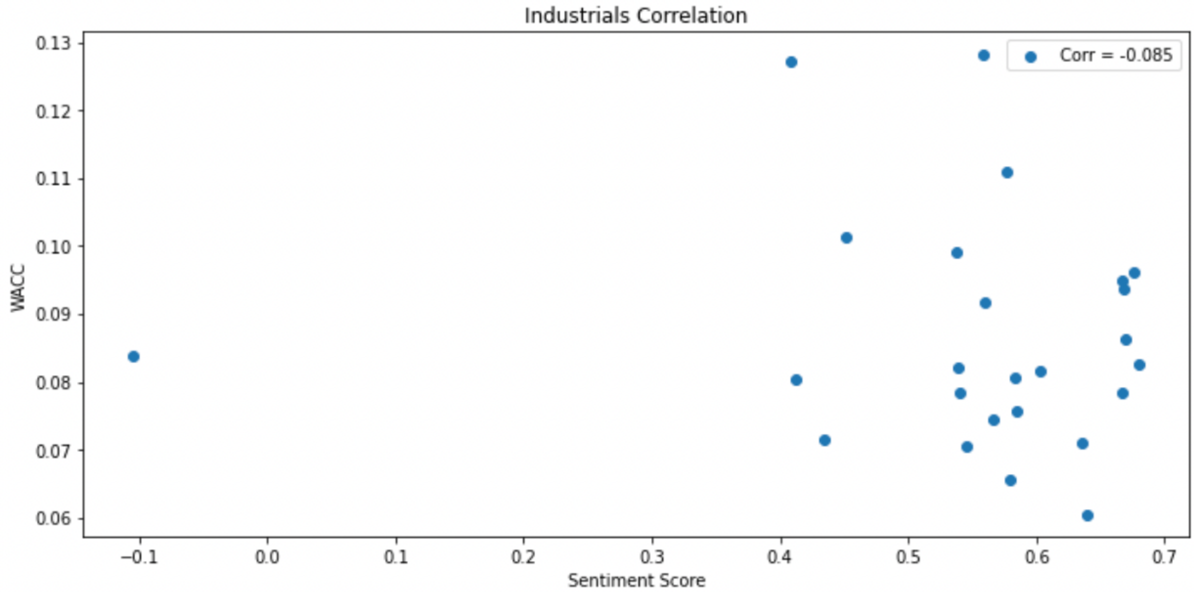
The result of the analysis was mixed. The sentiment scores of the Transportation, Energy/Electric Power, and Industrials sectors demonstrated a negative correlation with WACC as hypothesised. However, Agriculture and Real Estate had positive but weak correlation coefficients (see Fig 5.1).

Compared to other industries, transportation, energy, power, and industrial firms are arguably given more attention when environmental sustainability and climate change are discussed within the global financial domain. As such, it is not surprising to see a negative correlation for these industries. However, the positive correlation gotten for

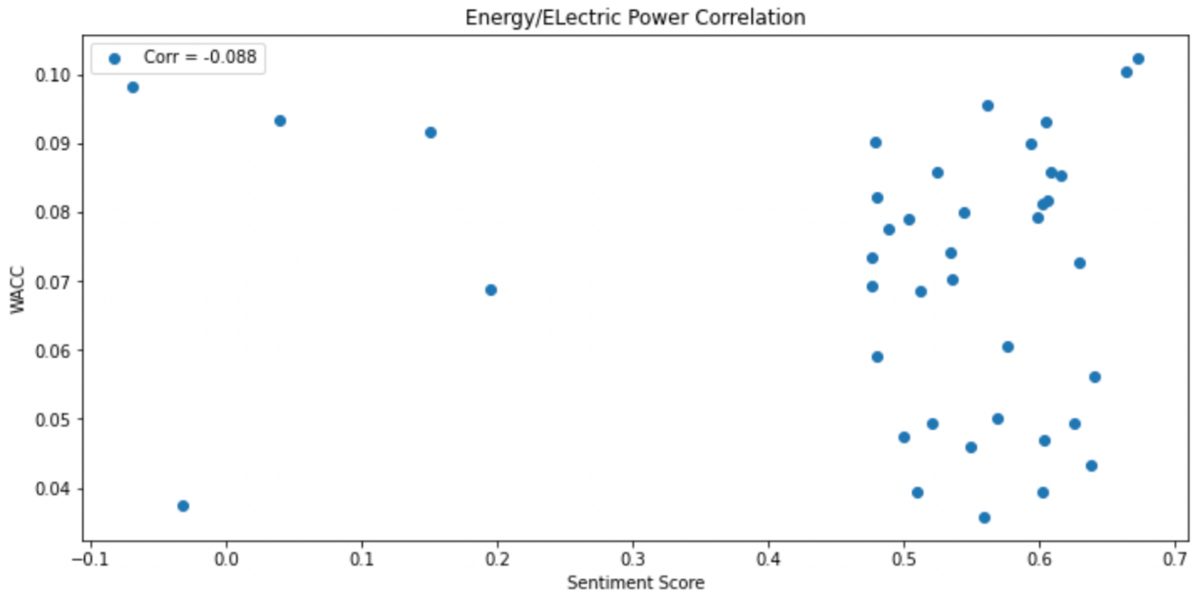
[11] The entire 2017 texts from NextEra Energy were classified as neutral, thereby leading to the total extinguish of NextEra in the result analysis

Agriculture, and Real Estate sectors leads us not to support the alternative hypothesis that there is a negative relationship between sentiment scores from sustainability reports and the cost of capital. This is further motivated by the weakness of the correlation coefficients across all the sectors. Furthermore, the weak correlation between their sentiment scores and WACC indicates that sustainability reports are not very influential in the final investment and capital allocation decisions. Hence, we do not also support the second hypothesis.

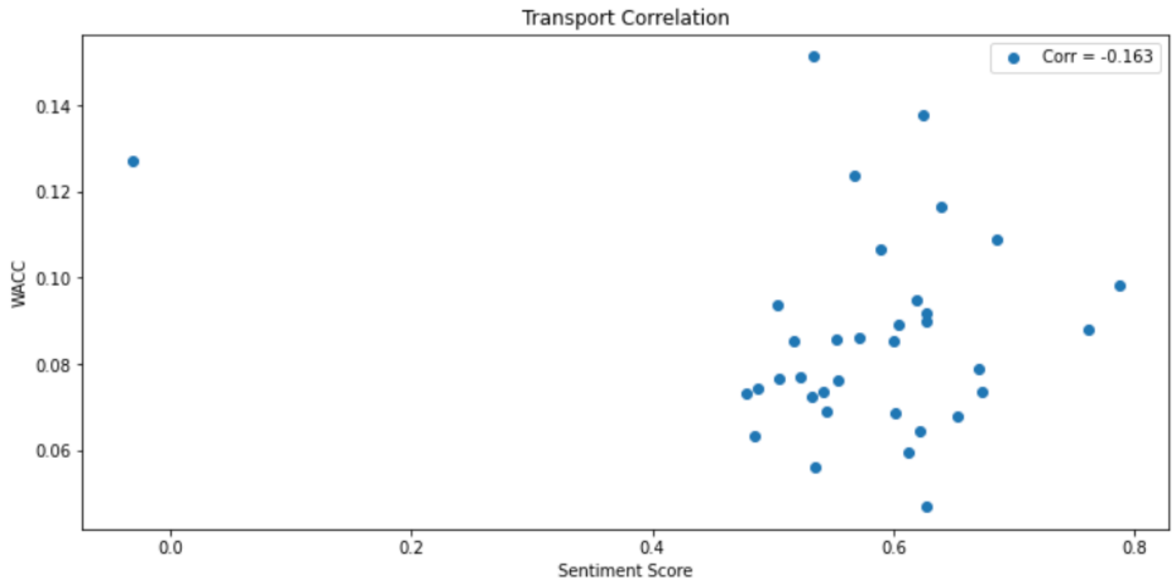




Industrials Correlation
Corr = -0.085
0.13
0.12
0.11
0.10
0.09
0.08
0.07
0.06
WACC
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Sentiment Score



Energy/ELectric Power Correlation
Corr = -0.088
0.10
0.09
0.08
0.07
0.06
0.05
0.04
WACC
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Sentiment Score



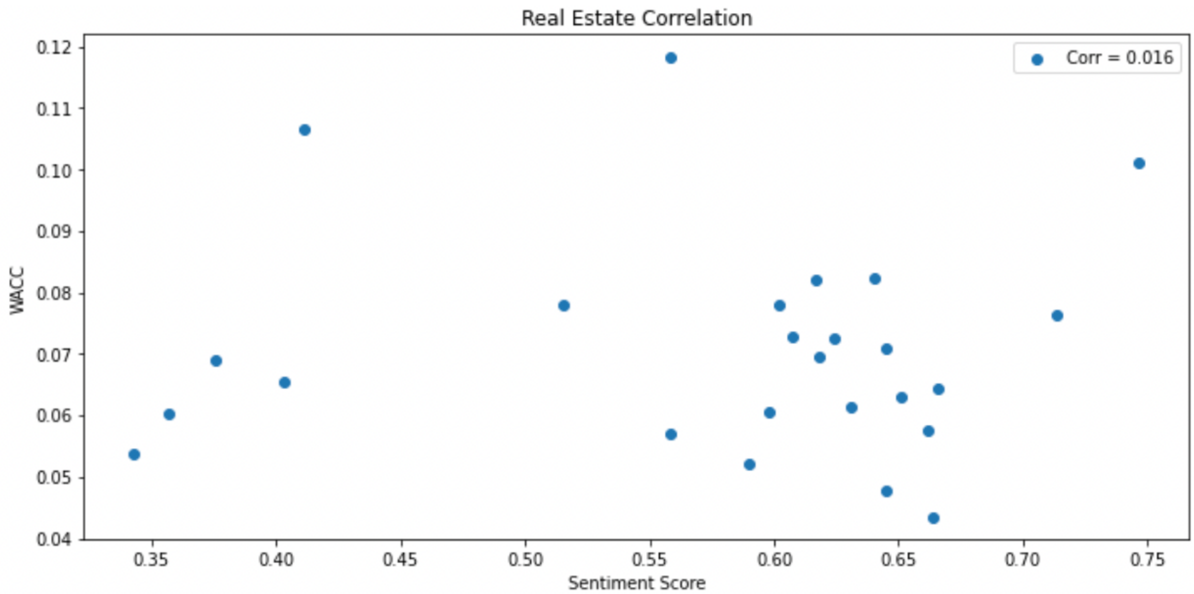


Fig 5.1: Correlation Analysis between Sentiment Scores and WACC

5.2 Limitation to the Results

Notwithstanding the impressive accuracy of FinBERT, some shortcomings abound. The model fails to accurately predict the sentiment when the text involves arithmetic comparison. For instance, one of the texts used in training FinBERT was – "Pre-tax loss totalled euro 0.3 million, compared to a loss of euro 2.2 million in the first quarter of 2005". This is a positive statement, but the model predicts it as Negative. Similarly, an example from the sustainability reports - *scope 3 emissions from employee business travel totalled 18,466 metric tons, a decrease from 18,603 in 2016.* This was predicted as negative, when indeed it is positive.

6.0 Conclusions and Future Work

6.1 Conclusions

This thesis seeks to investigate the relationship between the sentiments from sustainability reports and WACC based on the hypothesis that a negative relationship abounds between both. It also hypothesises that sustainable reports are very influential in the decision-making process of investors when making capital allocation choices. To examine these hypotheses, natural language analysis is conducted using the FinBERT model to derive the sentiment scores from the sustainability reports of 34 large-cap companies in the top five GHG-emitting sectors in the US. The review period spans from 2017 to 2021, and the WACC for these companies is sourced from Bloomberg. The resulting sentiment score is the positive minus the negative probabilities. However, neutral texts are regarded as noise and excluded from the analysis since they do not convey any adequate message and may interfere with the reliability of the result.

The outcomes across the five industries were divergent after testing the relationship between both variables. The Transportation, Energy/Electric Power and Industrial sector show a negative and weak relationship. Hence, it can be concluded that the sentiments from the sustainability reports from these sectors have an inverse relationship with their WACC. Contrarily, the Agriculture and Real Estate sectors show a positive and weak relationship. This leads us to conclude that sentiments from the reports of both sectors do not significantly correlate with WACC. In addition, considering the weak result across all the industries, this thesis also concludes that investors do not give very high importance to the contents of a sustainability report in the capital allocation or investment process. This conclusion corresponds with the assertion of Sokolov et al. (2021). They note that the management of companies believes that their sustainability disclosures do not have any importance in enabling investors to make an informed decision. Indeed, it is probable that investors discount these pieces of information since they are self-reported and are not governed by a standard. As such, their reliability in objectively assessing a company’s environmental performance may be questionable.

The findings of this thesis have two major contributions. Firstly, it signals to investors that they should begin to capture the environmental impact of a company’s operation as a risk factor, and the risk impact should reflect in their expected returns (cost of debt or cost of equity) from these companies. A higher or lower expected return will be a punishment or reward for errant and compliant companies respectively. In addition, this thesis believes and recommends that the debt capital market will be a faster transmission channel for this punish-reward mechanism. This is because companies tend to prefer debt financing over equity.

Secondly, the findings also highlight the need for investors to have a broad sectorial focus on climate change and not be lopsided. As much as some sectors are obvious culprits, other carbon-intense sectors should be put under equal scrutiny and called to accountability.

6.2 Future Work

While this thesis focuses on a select number of companies within the top five emitting sectors, it will be beneficial to have an additional study that will capture a wider sample size - company and sector-wise. Moreso, conducting a similar analysis on other markets (developed or emerging) will contribute immensely to the robustness of this topic. In addition, it will be insightful to know the outcome if FinBERT is used to analyse sentiments in relation to cost of debt and cost of equity as separate metrics. Furthermore, having a BERT model pre-trained with climate-related corpus might deliver a predictive result that beats the FinBERT. However, Howard and Ruder (2018) allude that it is expensive to train a model for domain-specific purposes.

7.0 References

Alvarado, J. C. S., K. Verspoor and T. Baldwin (2015). Domain Adaption of Named Entity Recognition to Support Credit Risk Assessment. Proceedings of the Australasian Language Technology Association Workshop.

Araci, D. (2019). FinBERT: Financial Sentiment Analysis with Pre-trained Language Models, University of Amsterdam.

Arora, G., A. Rahimi and T. Baldwin (2019). Does an LSTM forget more than a CNN? An empirical study of catastrophic forgetting in NLP. Proceedings of the 17th Annual Workshop of the Australasian Language Technology Association, Sydney, Australia, Australasian Language Technology Association.

Atan, R., M. Alam, J. Said and M. Zamri (2018). "The impacts of environmental, social, and governance factors on firm performance: Panel Study of Malaysian Companies." Management of Environmental Quality: An International Journal 29(2): 182-194.

Berg, F., J. F. Kölbel and R. Rigobon (2019). "Aggregate Confusion: The Divergence of ESG Ratings." Forthcoming Review of Finance.

Devlin, J., M.-W. Chang, K. Lee and K. Toutanova (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, Cornell University.

Felix, A., S. Henry and K. Roman (2020). A Computational Analysis of Financial and Environmental Narratives within Financial Reports and its Value for Investors. Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation.

Friede, G., T. Busch and A. Bassen (2015). "ESG and financial performance: aggregated evidence from more than 2000 empirical studies." Sustainable Finance & Investment 5(4): 210-233.

Howard, J. and S. Ruder (2018) "Universal Language Model Fine-tuning for Text Classification." DOI: https://doi.org/10.48550/arxiv.1801.06146.

Johri, P., M. Kathait, M. Sabharwal, A. T. Al-Taani and S. Suvanov (2021). Natural Language Processing: History, Evolution, Application and Future Work. Conference on Computing Informatics and Networks.: 365–375.

Jones, M. T. (2017). "Speaking out loud: An introduction to natural language processing." from https://developer.ibm.com/articles/cc-cognitive-natural-language-processing/.

Kishan, S. (2022). "ESG by the Numbers: Sustainable Investing Set Records in 2021." from https://www.bloomberg.com/news/articles/2022-02-03/esg-by-the-numbers-sustainable-investing-set-records-in-2021.

Kolouri, S., N. A. Ketz, X. Zou, J. L. Krichmar and P. K. Pilly (2019). Attention-Based Selective Plasticity.

United Nations, (2017). "The Evolution of Sustainable Finance." From https://www.unepfi.org/news/25th-anniversary/timeline/.

Pellegrini, C. B., R. Caruso and N. Mehmeti (2019). The impact of ESG scores on cost of equity and firm's profitability. New challenges in corporate governance: Theory and Practice: 38-40.

Pennington, J., R. Socher and C. Manning (2014). GloVe: Global Vectors for Word Representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, Association for Computational Linguistics.

Ramirez, A. G., J. Monsalve, J. D. González-Ruiz, P. Almonacid and A. Peña (2022). "Relationship between the Cost of Capital and Environmental, Social, and Governance Scores: Evidence from Latin America." Sustainability 14(9): 5012.

Sokolov, A., J. Mostovoy, J. Ding and L. Seco (2021). "Building Machine Learning Systems to Automate ESG Index Construction." The Journal of Impact and ESG Investing.

Whelan, T., U. Atz, T. V. Holt and C. C. CFA (2021). ESG and Financial Performance - Uncovering the relationship by aggregating evidence from 1000 plus studies from 2015 - 2020.

8.0 Appendix

GitHub to thesis codes: https://github.com/Muideen-Abubakar/Dissertation[12]

GitHub link to BERT: https://github.com/google-research/bert

GitHub link to FinBERT: https://github.com/ProsusAI/finBERT

[12] The notebook is split into two parts – the first part involves the pre-processing and sentence extraction, while the second part is the analysis and visualization of results