Deep Learning Strategies for Next-Gen Sentiment Analysis with Green AI Practices

MSc Research Project

Msc Artificial Intelligence for Business

Mariana Ketley Ferreira Cavalcante dos Santos

Student ID: x23149914

School of Computing

National College of Ireland

Supervisor: Rejwanul Haque

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Mariana Ketley Ferreira Cavalcante dos Santos
Student ID:	x23149914
Programme:	MSc Artificial Intelligence for Business **Year:** 2023
Module:	………………
Supervisor:	Rejwanul Haque
Submission Due Date:	15/09/2024
Project Title:	Deep Learning Strategies for Next-Gen Sentiment Analysis with Green AI Practices
Word Count:	20 **Page Count** 6466

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ……………………………………………………………………………

Date: 15/09/2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Deep Learning Strategies for Next-Gen Sentiment Analysis with Green AI Practices

Mariana Ketley Ferreira Cavalcante dos Santos

x23149914

Abstract

As consumer interest in green products grows, analysing reviews effectively is crucial. This study investigates sentiment analysis using deep learning on the Amazon Fine Food Reviews dataset, which includes over 500,000 reviews. Utilising the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology, it assesses attitudes towards eco-friendly products with a custom lexicon of sustainability terms. The research applies VADER (Valence Aware Dictionary and sEntiment Reasoner), a rule-based model, and RoBERTa (Robustly Optimized BERT Approach), a fine-tuned transformer-based deep learning model. TF-IDF (Term Frequency-Inverse Document Frequency) vectorization and Logistic Regression are used, with optimization via Hyperparameter GridSearchCV. Performance is measured by precision, recall, and F1-score, respectively. Green AI (Artificial Intelligence) practices, including DistilBERT Tokenizer and Model Pruning, are employed to minimise environmental impact. The study offers insights into consumer perceptions, helping businesses formulate sustainable strategies.

1 Introduction

As AI evolves, we are transitioning into a "new era" where sustainability and environmental concerns are central. Strubell et al. (2019) highlighted the significant carbon footprint of training advanced AI models, emphasising the need to reduce these emissions. Green AI signifies a major shift in AI research, prioritising energy efficiency and reduced carbon impact over mere performance. Schwartz et al. (2023) note that this new era integrates sustainable practices into AI design, marking a departure from traditional methods that often ignore environmental impacts. This transition not only fosters technological advancement but also aligns with global sustainability goals, as outlined by Schwartz, Hsu, and Vesselinov (2023).

Additionally, sentiment analysis, a crucial natural language processing (NLP) tool, helps businesses understand consumer feedback on a large scale (Pang & Lee, 2008). This thesis examines sentiment analysis using deep learning techniques, with a focus on Green AI to minimise environmental impact (Schwartz et al., 2020). The focus is on analysing sentiment towards eco-friendly products using the Amazon Fine Food Reviews dataset, which provides a wealth of consumer feedback (McAuley & Leskovec, 2013).

As global awareness of environmental sustainability grows, it is increasingly important for businesses to understand consumer sentiment towards green products (Young et al., 2010). By identifying reviews mentioning green attributes like “eco-friendly,” “biodegradable,” and “organic,” companies can tailor their strategies. Employing the CRISP-DM (Cross-Industry

Standard Process for Data Mining) methodology, the research follows a structured approach from business understanding to evaluation (Chapman et al., 2000). This involves data understanding, where exploratory data analysis (EDA) identifies trends related to green products, informing data preprocessing and model training (Tukey, 1977).

The modelling phase incorporates VADER (Valence Aware Dictionary and sEntiment Reasoner) for baseline sentiment analysis and the RoBERTa (Robustly Optimized BERT Approach) framework for more nuanced language interpretation, with Sanh et al. (2019) employing the DistilBERT tokenizer to enhance processing speed. The dataset is divided into training and testing sets, using TF-IDF (Term Frequency-Inverse Document Frequency) vectorization and Logistic Regression for classification (Cox, 1958). Additionally, model pruning is applied to reduce complexity without significantly impacting performance, further adhering to Green AI practices. Hyperparameter optimization through GridSearchCV ensures the model is efficient and conserves resources (Schwartz et al., 2020). The study's evaluation leverages recall, precision, and F1-score metrics to gauge the model's effectiveness in identifying sentiments towards green products, underscoring the significance of sustainable AI development. This research provides actionable insights for businesses and encourages environmentally conscious consumer behaviour.

2 Literature Review

2.1 Sentiment Analysis and Consumer Reviews

Tang et al. (2015) highlight the importance of sentiment analysis for understanding consumer opinions, using advanced deep-learning techniques. Devlin et al. (2019) describe the impactful RoBERTa (Robustly Optimized BERT Approach) model, which surpasses traditional methods in capturing context. These innovations emphasise deep learning's role in refining sentiment analysis for more precise insights. VADER (Valence Aware Dictionary and sEntiment Reasoner), a rule-based sentiment analysis tool, excels in analysing social media text (Hutto & Gilbert, 2014). Unlike machine-learning models, VADER uses a predefined lexicon and heuristics to interpret text, providing clear and interpretable sentiment analysis. RoBERTa, as discussed by Devlin et al. (2019) and Liu et al. (2019), enhances natural language processing (NLP) tasks, including sentiment analysis, through deep contextual understanding and robustness.

2.2 Green AI Practices

Green AI seeks to reduce the environmental impact of AI models, focusing on optimising algorithms and using energy-efficient hardware (Schwartz et al., 2020). In contrast, "Red AI" prioritises performance, often at the cost of high energy consumption (Schwartz et al., 2020). Green AI promotes sustainable deep learning by minimising resource use, as shown by Rafat et al. (2023). Understanding consumer attitudes toward green products helps identify market trends. Kumar (2015) and Yang (2017) explore factors influencing consumer behaviour, emphasising the growing interest in sustainability.

2.3 CRISP-DM

Integrating CRISP-DM (Cross-Industry Standard Process for Data Mining) and Green AI into sentiment analysis involves phases like Business Understanding and Data Preparation, ensuring minimal environmental impact (Singgalen, 2024; Yigitcanlar, 2021). During Modelling and Evaluation, energy-efficient models are prioritised (Schwartz et al., 2020; Strubell et al., 2019). Deployment includes monitoring energy use and accuracy (Raman et al., 2024). The CRISP-DM framework provides a structured approach to data mining, involving stages like Business Understanding and Deployment (Sudar et al., 2024). Case studies, such as those by Zhao et al. (2012) and Hasan et al. (2019), demonstrate CRISP-DM's application in creating energy-efficient sentiment analysis systems.

3 Research Methodology

The project employs the CRISP-DM methodology, beginning with the collection and preprocessing of a large dataset pertinent to sentiment analysis (Sudar, 2024). Initial preprocessing steps include data cleaning, tokenization using Auto Tokenizers from transformers in the RoBERTa model, and identification of green products within the dataset. Following this, sentiment analysis will be performed using VADER for rule-based sentiment analysis to establish a baseline.

Integrating sentiment analysis with RoBERTa implementation models, using VADER for pre-built sentiment analysis and CRISP-DM methodology, aims to follow Green AI practices, offering a comprehensive approach to evaluating consumer feedback on green products (Hutto & Gilbert, 2014; Devlin et al., 2019). This integration enhances the accuracy and reliability of sentiment analysis while supporting sustainable AI practices (Schwartz et al., 2020). Future research should continue to optimise deep-learning techniques for sentiment analysis, with an emphasis on minimising environmental impact and improving efficiency (Thompson et al., 2020). A visualisation of the CRISP-DM Methodology is provided in Figure 1.

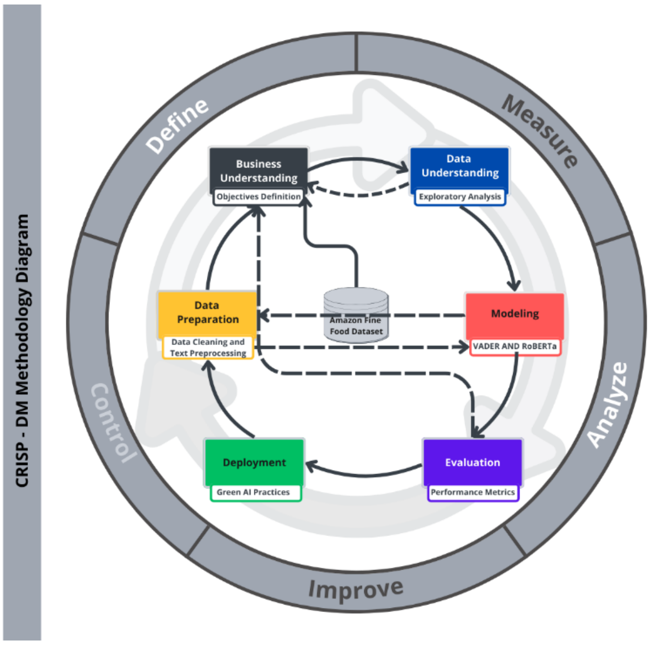

Figure 1: CRISP-DM

The research focuses on training and evaluating the RoBERTa (Robustly Optimised BERT Pretraining Approach) model using accuracy, precision, recall, and F1-score metrics. It explores data reduction techniques like DistilBERT (Distilled BERT) and model pruning to enhance computational efficiency, aligning with Green AI practices. The study also assesses the reduction in computational resources through preprocessing, emphasising energy-efficient AI development. Results will identify the best models and strategies, promoting sustainable AI. In Figure 2, it is possible to see the architecture of VADER, RoBERTa, and DistilBERT, as well as a comparison of their architectures.

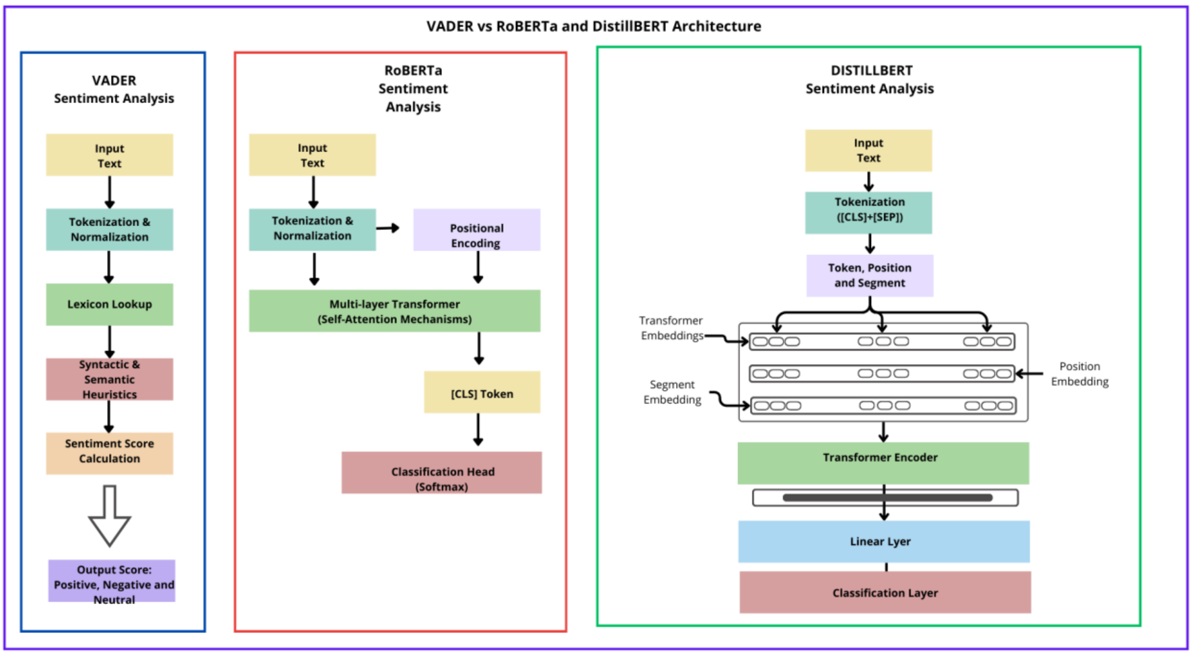

Figure 1: VADER, RoBERTa and DistilBERT Architecture

The diagram presents three different methodologies for sentiment analysis: VADER, RoBERTa, and DistilBERT. VADER (Valence Aware Dictionary and sEntiment Reasoner) employs a straightforward rule-based approach. It tokenises and normalises the input text, performs a lexicon lookup to associate words with sentiment scores, and then applies syntactic and semantic heuristics to calculate the final sentiment score, categorising it as positive, negative, or neutral. RoBERTa, on the other hand, uses a more complex deep-learning model. It tokenises and normalises the input, applies positional encoding, and processes the text through a multi-layer transformer network that uses self-attention mechanisms. This is followed by the addition of a [CLS] token and a classification head that uses softmax to predict the sentiment.

In contrast to RoBERTa and VADER, DistilBERT is a smaller, faster, and lighter version of BERT that maintains high performance by distilling knowledge from the larger BERT model (Sanh et al., 2019). Unlike VADER and RoBERTa, DistilBERT not only tokenises the input with special tokens ([CLS] and [SEP]) but also encodes token, position, and segment embeddings. These embeddings are processed through a transformer encoder, which captures complex relationships and dependencies within the text. DistilBERT's architecture includes a linear layer followed by a classification layer, ensuring efficient sentiment prediction. This model excels in Named Entity Recognition (NER) tasks on résumés due to its ability to preserve semantic information and extract key features effectively during fine-tuning, providing a superior balance of speed and accuracy compared to traditional and larger models.

Ethically, reducing the carbon footprint of AI by developing energy-efficient algorithms and using eco-friendly infrastructure is crucial. Ensuring fairness in sentiment analysis by addressing biases in training data is also important to avoid discrimination (He & Zhou, 2011). Moreover, resource-efficient AI can democratise access, reducing the digital divide. Ethical considerations include data privacy, security, and transparency, especially with sensitive personal data. The research aims for responsible AI development, prioritising ethical principles alongside technology.

4 Design Specification

4.1 Business Understanding: Goals Definition

This study developed a sentiment analysis system using deep-learning methods, specifically analysing sentiment in the Amazon Fine Food Reviews dataset. The system's design prioritises both accurate customer sentiment analysis and adherence to Green AI practices, which emphasise the importance of minimising the environmental impact of large AI models. The project will concentrate on detecting sentiment about green products, defined as products that are sustainable and environmentally friendly. According to Yang (2017), the insights gained will help researchers and businesses understand consumer attitudes towards green products. This aligns with a broader societal goal of promoting environmental sustainability and informed consumer decision-making (Haakman et al., 2021). Figure 2 presents the design specification of the sentiment analysis performed in this study. [i]

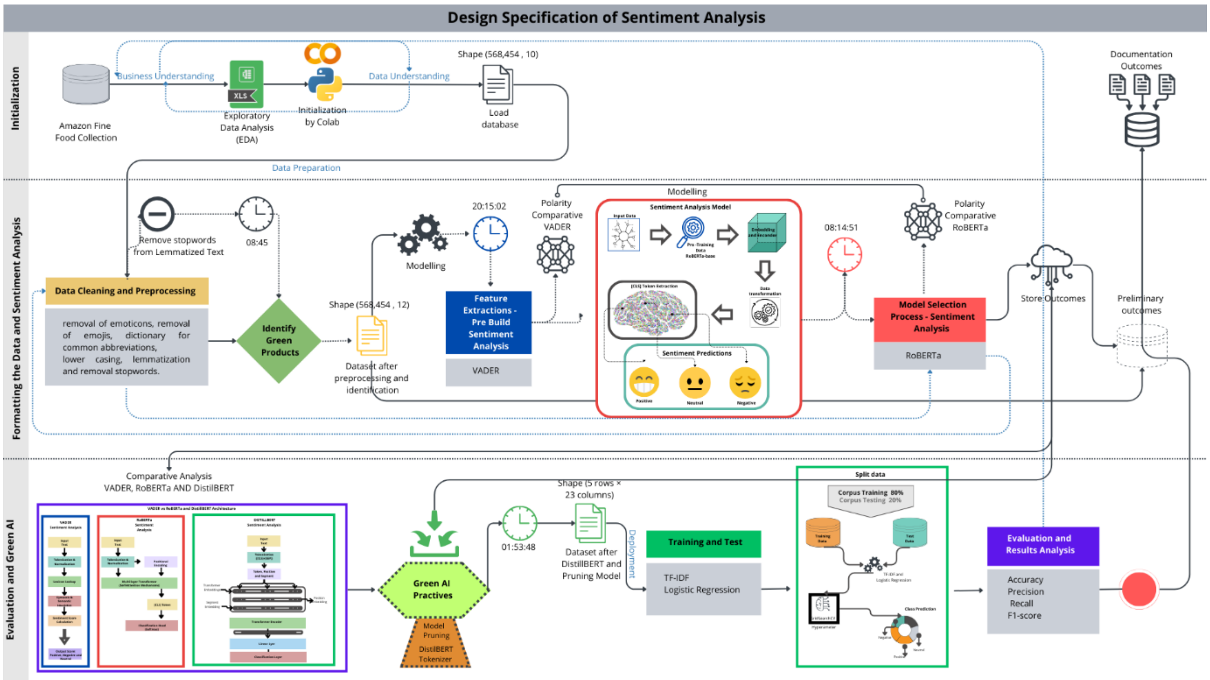

Figure 2: Design Specification of Sentiment Analysis

4.2 Data Understanding: Exploratory Data Analysis (EDA)

The Amazon Fine Food Reviews dataset offers extensive textual data, including reviews and ratings. According to Tukey (1977), the first step is Exploratory Data Analysis (EDA) to understand the dataset's structure and key features, such as review rating distributions, review frequencies per product, and common themes. Visualisation tools like bar charts reveal patterns and trends. Special focus is given to reviews mentioning green products, using keywords like "fresh," "eco-friendly," "organic," and "biodegradable" to filter and categorise them. A new column, "Green Products," is added to the dataset, which is crucial for tailoring the sentiment analysis model to assess opinions on green products and support the project's sustainability focus (Panda et al., 2022).

4.3 Preprocessing: Text Preprocessing and Data Cleaning (Identification)

Preprocessing for sentiment analysis using RoBERTa and VADER involves text preprocessing and data cleaning, including lowercasing, stop word removal, and lemmatization (Zhao, 2012). Emojis and emoticons are converted to words to standardise expressions. A specialised lexicon for green products is created to identify environmentally friendly items. This helps in isolating and analysing sentiments specifically related to eco-friendly products.

4.4 Modelling: VADER and Deep Learning (RoBERTa Technique)

Two modelling approaches will be used: VADER (Valence Aware Dictionary and sEntiment Reasoner) for quick sentiment classification into positive, negative, or neutral with minimal computational resources, and RoBERTa (Robustly Optimised BERT Approach) for a deeper understanding of context and nuances in natural language. RoBERTa, fine-tuned on the dataset, enhances sentiment accuracy, especially for green products. To optimise for Green AI, DistilBERT Tokenizer and model pruning will be applied.

4.5 Deployment: Hyperparameter Optimization and Green AI

To deploy the model effectively, it will be implemented in a real-world setting, ensuring suitability for the intended environment. Hyperparameter optimisation will be performed using GridSearchCV, which tests various combinations to find the best configuration. The dataset will be split into 80% training and 20% test sets for evaluation (Bhati & Kher, 2019). Term Frequency-Inverse Document Frequency (TF-IDF) will extract key features, emphasising important words and down-weighting less relevant ones (Cox, 1958). A logistic regression classifier will be used for its efficiency and interpretability, balancing accuracy with computational efficiency in line with Green AI principles.

4.6 Evaluation: Precision, Recall, Accuracy and F1-Score

According to Chen et al. (2021), the evaluation should focus on precision, recall, accuracy, and F1-score to assess the model's effectiveness in distinguishing sentiment categories, while also monitoring computational resources to adhere to Green AI practices. The visualisation in Figure 1 shows that Business Understanding and Data Understanding are interconnected, involving goal setting and success criteria. Data Preprocessing and Modelling are linked as data is collected and processed for the RoBERTa model, and Evaluation is tied to Business Understanding to refine goals and metrics.

5 Implementation

5.1 Data Understanding

In this study, the process of sentiment analysis was guided using the CRISP-DM methodology, according to Chapman et al. (1999). This methodology encompasses the initial comprehension of sentiment-focused business objectives through to the implementation and deployment of a functional sentiment analysis model. After analysing the customer reviews from the Amazon Fine Food dataset, the goals for this study were defined as: analysing the reviews, identifying green products, conducting sentiment analysis to detect emotions, evaluating the models, and deploying the sentiment analysis within the business framework to guide decision-making while minimising the environmental impact of AI processes. This framework approach enhances the precision and performance of sentiment analysis (Hasan, 2019).

5.2 Data collection:

Data collection was crucial for ensuring reliable research findings and accurate analysis. A comprehensive dataset was selected to support robust conclusions and effective sentiment analysis. The dataset chosen for this study was the Amazon Fine Food Reviews, available on the Kaggle platform. It is in the public domain, allowing for copying, modification, distribution, and performance of the work, even for commercial purposes, without requiring permission. This dataset contains reviews of foods sold on Amazon and includes 568,454 reviews, 256,059 users, 74,258 products, and 260 users with more than 50 reviews, spanning a period of 10 years (October 1999 to 2012). It features variables such as product, user information, rating, review text, and other relevant metadata. Figure 3 presents a visualisation of the information about the dataset before preprocessing.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 568454 entries, 0 to 568453
Data columns (total 10 columns):
 #   Column                  Non-Null Count   Dtype
---  ------                  --------------   -----
 0   Id                      568454 non-null  int64
 1   ProductId               568454 non-null  object
 2   UserId                  568454 non-null  object
 3   ProfileName             568428 non-null  object
 4   HelpfulnessNumerator    568454 non-null  int64
 5   HelpfulnessDenominator  568454 non-null  int64
 6   Score                   568454 non-null  int64
 7   Time                    568454 non-null  datetime64[ns]
 8   Summary                 568427 non-null  object
 9   Text                    568454 non-null  object
dtypes: datetime64[ns](1), int64(4), object(5)
memory usage: 43.4+ MB
```

Figure 3: Data info

The dataset comprises 10 columns: ID, ProductId, UserId, ProfileName, HelpfulnessDenominator (number of users who indicated if a review was helpful), HelpfulnessNumerator (number of users who found a review useful), Score (rating from 1 to 5), Time (timestamp), Summary (review summary), and Text (review content). Collected via web scraping and APIs like the Amazon Product Advertising API, the Amazon Fine Food Reviews dataset covers various food products, including cereals, groceries, and beverages, offering insights into aspects like taste, packaging, and nutritional value.

5.2 Data preprocessing:

For sentiment analysis, extensive text preprocessing was applied, including lowercasing, converting emojis and emoticons to words, removing stop words, and lemmatization. According to Renouf & Ellis (2005), lemmatization, which reduces words to their base forms, was particularly challenging due to the Amazon Fine Food Reviews dataset's size of over 500,000 entries. This process required significant computational power, handled efficiently using Google Colab's upgrade with Colab Plus for access to multiple processors simultaneously, including T4 GPU and high RAM environment.[ii]

Additionally, "Pandas - pandarallel.initialize" with a progress bar was used to accelerate all steps, executed in 40 groups of GPU. A new "text_lemmatized" column was added to store standardised text. This approach, utilising cloud resources, aligns with Green AI principles by optimising resource use and minimising environmental impact. A specialised lexicon for green products was created to flag reviews discussing eco-friendly products, offering insights into sustainability-focused consumer sentiments. The chosen preprocessing techniques were tailored to the dataset's needs.

All techniques played an important role in accurate sentiment analysis. Libraries used included "transformers," "scipy," "datasets," "pandas," "numpy," "matplotlib," "seaborn," "nltk," "tqdm," and "sklearn." The first step involved collecting basic information about the dataset, such as summary statistics, column names, number of missing values per column, number of unique values per column, data types of columns, and number of rows and columns. The code was adjusted to match actual column names, and lowercasing was applied to standardise words like FRESH, Fresh, and fresh, reducing duplications. Stop words were removed using NLTK's English stopword list. Lemmatization was chosen over stemming for its accuracy, preserving meaning and grammar. Emoticons and emojis were converted into textual emotions using dictionaries containing 48 emoticons and 117 emojis, relevant for the data period from October 1999 to 2012, when their use was prevalent. Both emoticons (made from keyboard characters) and emojis (images of facial expressions) were addressed in the preprocessing.

Another important preprocessing technique used in this study was chat word conversion, given the prevalence of abbreviations and slang in modern communication. This step improved the analysis by including a slang dictionary and translations for these slangs. After completing the preprocessing, data visualisation was performed again to analyse the dataset's status post-preprocessing. Figure 4 shows the visualisation of the dataset after preprocessing techniques, including the added columns related to lemmatization and identification of green products.

	Id	ProductId	UserId	ProfileName	HelpfulnessNumerator	HelpfulnessDenominator	Score	Time	Summary	Text	text_wo_stop	is_green
0	1	B001E4KFG0	A3SGXH7AUHU8GW	delmartian	1	1	5	2011-04-27	Good Quality Dog Food	I have bought several of the Vitality canned d...	bought several vitality canned dog food produc...	False
1	2	B00813GRG4	A1D87F6ZCVE5NK	dll pa	0	0	1	2012-09-07	Not as Advertised	Product arrived labeled as Jumbo Salted Peanut...	product arrived labeled jumbo salted peanuts.....	False
2	3	B000LQOCH0	ABXLMWJIXXAIN	Natalia Corres "Natalia Corres"	1	1	4	2008-08-18	"Delight" says it all	This is a confection that has been around a fe...	confection ha around centuries. light, pillowy...	False
3	4	B000UA0QIQ	A395BORC6FGVXV	Karl	3	3	2	2011-06-13	Cough Medicine	If you are looking for the secret ingredient i...	looking secret ingredient robitussin believe f...	False
4	5	B006K2ZZ7K	A1UQRSCLF8GW1T	Michael D. Bigham "M. Wassir"	0	0	5	2012-10-21	Great taffy	Great taffy at a great price. There was a wid...	great taffy great price. wa wide assortment yu...	False

Figure 4: Dataset after Preprocessing and Identification Green Products

5.3.1 Biases Identification in Sentiment Analysis of Green Products

In sentiment analysis of green products, biases can skew results and affect accuracy. Dataset bias, such as the overrepresentation of certain demographics or product types in the Amazon Fine Food Reviews dataset, can distort market sentiment. Biases may also arise from sentiment analysis tools, which might misinterpret or overlook nuances in consumer language if not trained on diverse data. Such biases can misrepresent consumer sentiment, affecting business strategies and product development.

a. Limitations of sentiment Analysis:

Detecting sarcasm and handling regional language differences in sentiment analysis present challenges for RoBERTa and VADER. Sarcasm is difficult to identify due to its subtlety, and

VADER's rule-based system may be inadequate, while RoBERTa's effectiveness depends on training with sarcastic data. Regional spelling and vocabulary differences further complicate analysis, requiring RoBERTa to be fine-tuned for various English dialects and adjustments to VADER's static lexicon. Addressing these issues requires specialised models and diverse training data (Kumar et al., 2021).

b. Identify Green Products:

Once preprocessing is fully executed, the identification of green products is performed using the Pandas library. Initially, a list of approximately 120 keywords for green products is defined, as illustrated in Figure 5. This allows the code to determine the quantity of green products in the dataset according to the keywords list. Promoting an understanding of consumer preferences for green products aligns with Green AI principles by focusing on environmentally friendly consumer choices and supporting research on enhancing consumer sentiment towards green products (Raman et al., 2024).

```
[ ]  green_keywords = ['organic', 'eco-friendly', 'biodegradable', 'sustainable', 'green', 'natural', 'environmentally friendly',
         'non-GMO', 'fair trade', 'locally sourced', 'farm-to-table', 'pesticide-free', 'hormone-free', 'grass-fed',
         'free-range', 'cage-free', 'certified humane', 'biodynamic', 'carbon-neutral', 'compostable', 'cruelty-free',
         'renewable', 'recyclable packaging', 'low carbon footprint', 'ethical', 'vegan', 'plant-based', 'gluten-free',
         'preservative-free', 'chemical-free', 'sustainable farming', 'fresh', 'sustainable packaging', 'organic ingredients',
         'chemical-free', 'eco-conscious', 'ethical sourcing', 'zero waste', 'renewable energy', 'minimal processing',
         'non-toxic', 'fresh', 'eco-packaging', 'low-impact', 'responsibly sourced', 'upcycled ingredients', 'no additives',
         'locally grown', 'naturally raised', 'clean eating', 'eco-certified', 'earth-friendly', 'green certified',
         'nutrient-dense', 'solar-powered', 'green farming', 'regenerative agriculture', 'hydroponic', 'organic certified',
         'small-batch', 'wildlife-friendly', 'low-emission', 'handcrafted', 'artisanal', 'holistic', 'grassroots',
         'ethical farming', 'pure', 'non-polluting', 'naturally flavored', 'heirloom varieties', 'sustainable seafood',
         'organic', 'eco-friendly', 'biodegradable', 'sustainable', 'green', 'natural', 'environmentally friendly',
         'non-GMO', 'fair trade', 'locally sourced', 'farm-to-table', 'pesticide-free', 'hormone-free', 'grass-fed',
         'free-range', 'cage-free', 'certified humane', 'biodynamic', 'carbon-neutral', 'compostable', 'cruelty-free',
         'renewable', 'recyclable packaging', 'low carbon footprint', 'ethical', 'vegan', 'plant-based', 'gluten-free',
         'preservative-free', 'chemical-free', 'sustainable farming', 'fresh', 'sustainable packaging', 'organic ingredients',
         'chemical-free', 'eco-conscious', 'ethical sourcing', 'zero waste', 'renewable energy', 'minimal processing',
         'non-toxic', 'eco-packaging', 'low-impact', 'responsibly sourced', 'upcycled ingredients', 'no additives',
         'locally grown', 'naturally raised', 'clean eating', 'eco-certified', 'earth-friendly', 'green certified',
         'nutrient-dense', 'solar-powered', 'green farming', 'regenerative agriculture', 'hydroponic', 'organic certified',
         'small-batch', 'wildlife-friendly', 'low-emission', 'handcrafted', 'artisanal', 'holistic', 'grassroots',
         'ethical farming', 'pure', 'non-polluting', 'naturally flavored', 'heirloom varieties', 'sustainable seafood']

     def is_green_product(description):
         return any(keyword in description.lower() for keyword in green_keywords)

     df['is_green'] = df['text_wo_stop'].progress_apply(is_green_product)
     df.head()
```

Figure 5: Green Products Identification using Keywords

Below, Figure 6 presents a pie chart illustrating the emphasis on green versus non-green products in consumer reviews from the Amazon Fine Food Reviews dataset. The chart is divided into two segments: "Green" and "Non-Green." The "Green" portion represents reviews mentioning environmentally friendly or sustainable features such as "organic," "biodegradable," or "eco-friendly," while the "Non-Green" portion represents reviews that do not emphasise these eco-friendly characteristics. The large "Non-Green" portion indicates that most reviews in the dataset do not focus on eco-friendly products and their attributes.

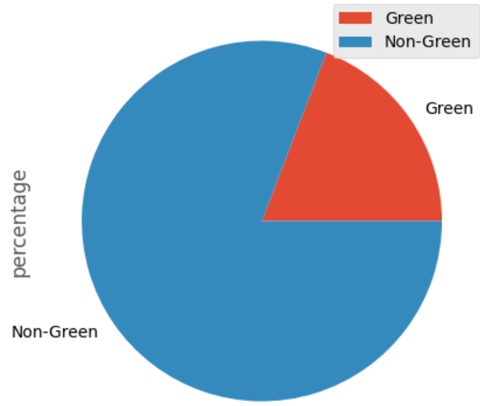

Figure 6: Green vs Non-Green Identification

This distribution suggests a relatively low level of consumer awareness or interest in green products. From a business perspective, it highlights an opportunity to expand market focus on green products and helps businesses identify gaps and potential areas for growth in promoting sustainable products. This analysis can guide product development and marketing strategies by providing insights into the current consumer focus. Figure 7 shows a word cloud visualisation of green and non-green products.

Word Cloud Green Products	**Word Cloud Non-Green Products**
	

Figure 7: Green vs Non-Green Word Cloud

Figure 7 displays word clouds representing text data from Amazon Fine Food Reviews, distinguishing between "green" (organic or eco-friendly) products and "non-green" products. Both word clouds feature common terms such as "coffee," "tea," "product," "taste," "flavour," "good," and "like," indicating these terms are frequently used across reviews regardless of the product's green status. However, words like "organic," "natural," and "ingredient" are more prominent in the green products word cloud, reflecting a focus on health and eco-friendliness. In contrast, the non-green products word cloud features terms like "price" and "order," suggesting a greater concern with cost and purchasing experience. Both datasets highlight positive sentiment through words like "love," "great," "best," and "really," indicating generally favourable reviews. This differentiation helps identify key themes and concerns specific to green products, essential for targeted sentiment analysis. The dataset consists of 399.5 green products and 199.5 non-green products, providing a balanced representation for a comprehensive comparison of sentiments between the two categories.

5.4 Modelling

According to Hasan (2019), sentiment analysis involves determining emotions expressed in text, ranging from positive, neutral, and negative to more specific emotions like anger or joy. It is widely used in applications such as social media analysis and product review evaluations. However, it may not be perfect and cannot always explain the reasons behind sentiments, although it helps summarise large volumes of data. For this study, data preprocessing techniques such as VADER and deep learning methods, including RoBERTa, were chosen for sentiment analysis. The NLTK library, with its VADER sentiment analyser, was used to score sentiment and handle text classification. VADER employs a lexicon-based approach, assigning sentiment scores while accounting for nuances like capitalization and punctuation.

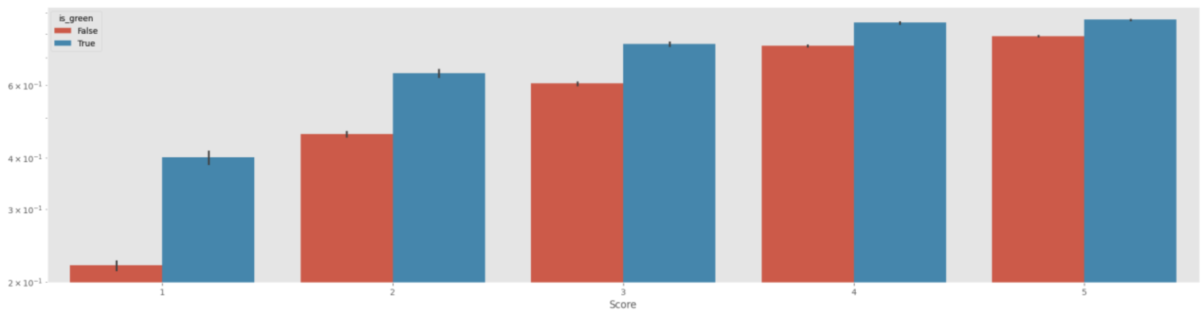

Figure 8: VADER Comparative Polarity - Green vs Non-Green

Figure 8 shows the VADER polarity review stars chart, highlighting consumer sentiment differences between green and non-green products in the Amazon Fine Food Reviews dataset, categorised by star ratings. Using a logarithmic scale, it shows that green products typically receive more positive sentiment. This visualisation is crucial for businesses aiming to cater to eco-conscious customers, as it reveals valuable insights for product development and marketing strategies. The data underscores a trend of more positive reviews for green products, supporting the study's goal of promoting sustainability by demonstrating the market benefits of eco-friendly attributes.

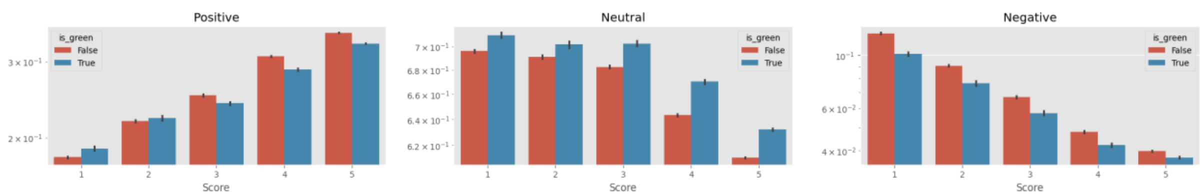

Figure 9: VADER Comparative Polarity - Green vs Non-Green

Figure 9, the "VADER Polarity Comparative" graphic, shows consumer sentiment towards green products, categorised as positive, neutral, or negative, using a logarithmic scale. This visualisation is crucial in the Business Understanding phase, highlighting differences in sentiment polarity based on product greenness. The logarithmic scale enhances clarity by making variations more visible than a linear scale. Unlike the "VADER Polarity Review Stars - Green vs Non-Green" chart, which links star ratings to green mentions, this chart delves into sentiment polarity, revealing that non-green products often have higher positive sentiments. After VADER analysis, RoBERTa was used for its deep understanding of text context and nuances, predicting sentiment scores correlated with positive, neutral, and negative sentiments. RoBERTa excels in capturing linguistic subtleties, making it highly effective for NLP tasks such as sentiment analysis, text classification, and question answering. This advanced model, which enhances BERT's pre-training techniques, is a key tool in NLP for understanding and generating human language (He & Zhou, 2011).

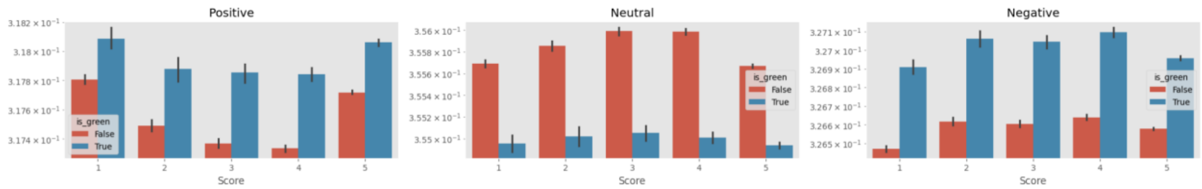

Figure 10: RoBERTa Comparative Polarity - Green vs Non-Green

Figure 10, the "RoBERTa Polarity Comparative" chart, details consumer sentiment towards green products. It shows that green products score higher in positive sentiment, especially at score 1 with a sentiment score of around 0.318, compared to non-green products at approximately 0.317. Non-green products lead in neutral sentiment, especially at scores 1 and 2, with scores around 0.355 and 0.356, while green products remain around 0.355. For negative sentiments, green products score higher at scores 2 and 3, around 0.327 each, compared to lower scores for non-green products. These results indicate that green products often receive more extreme positive and negative feedback, while non-green products receive more neutral reviews. The logarithmic scale clarifies these insights, helping businesses adjust strategies to better address eco-conscious consumer preferences and concerns (Young et al., 2010).

a. Green-AI Practices:

The DistilBERT tokenizer optimises text preparation for the DistilBERT model, a streamlined version of RoBERTa. It segments text into smaller tokens and translates these into numerical values that the model can process. For Green AI, this tokenizer enhances efficiency by speeding up text processing and reducing the computational and memory requirements compared to more extensive models like RoBERTa and VADER. This reduction in resource consumption is crucial for lessening the environmental impact of training and deploying AI systems. Although the tokenizer itself does not directly address bias, employing DistilBERT, trained on diverse datasets, can help minimise biases in text processing, assuming the training data is comprehensive and well-balanced (Sanh et al., 2019).

Model pruning was applied, which involves the deliberate removal of specific parts of a neural network, such as neurons or connections, to reduce the model's size and resource needs. This approach improves model efficiency by decreasing the number of parameters, which in turn lowers memory usage and speeds up inference. In the context of Green AI, pruning contributes to sustainability by reducing the computational resources required for model training and deployment, thus cutting down on energy consumption (Yigitcanlar, 2021).

Although pruning does not directly address model bias, it facilitates more efficient model evaluation and adjustment, making it easier to test and refine models to ensure they are fair and unbiased.

```
%load_ext memory_profiler

import torch
from transformers import DistilBertTokenizer, DistilBertForSequenceClassification
from torch.nn.utils import prune
from scipy.special import softmax

# Preprocess data
df['Sentiment'] = df['Score'].apply(lambda x: 1 if x >= 4 else 0)

# Selecting relevant columns
X = df['Text']
y = df['Sentiment']

# Initialize DistilBERT tokenizer and model
tokenizer = DistilBertTokenizer.from_pretrained("distilbert-base-uncased")
model = DistilBertForSequenceClassification.from_pretrained("distilbert-base-uncased", num_labels=3)
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
model = model.to(device)

# Define a function to apply pruning to the model
def prune_model(model):
    # Apply pruning to the linear layers
    for name, module in model.named_modules():
        if isinstance(module, torch.nn.Linear):
            prune.random_unstructured(module, name='weight', amount=0.2)  # Prune 20% of weights

prune_model(model)  # Application of pruning to the model

# Define function to get polarity scores using DistilBERT
@torch.autocast(device_type="cuda")
def get_polarity_scores_distilbert(row):
    try:
        # Extract pre-tokenized input
        encoded_input = tokenizer(row['text_wo_stop'], max_length=512, padding='max_length', truncation=True,
return_tensors='pt')

        # Move inputs to the device
        input_ids = encoded_input['input_ids'].to(device)  # Access input_ids from encoded_input
        attention_mask = encoded_input['attention_mask'].to(device)  # Access attention_mask from encoded_input

        # Forward pass through the model with mixed precision
        output = model(input_ids=input_ids, attention_mask=attention_mask)

        # Compute the softmax probabilities
        scores = output.logits[0].cpu().detach().numpy()
        probabilities = softmax(scores)

        # Update the row with the computed probabilities
        row['distil_neg'], row['distil_neu'], row['distil_pos'] = probabilities
    except RuntimeError as e:
        print(f'Error processing id {row["Id"]} - {e}')
    return row

# Apply the function to the dataframe
%memit df = df.progress_apply(lambda x: get_polarity_scores_distilbert(x), axis=1)
```

Code Visualisation 1: DistilBERT and Pruning Model

This code is designed for sentiment analysis using the DistilBERT model, with a focus on minimising computational resources in line with Green AI practices. It begins by preprocessing data, converting review scores into multi-class sentiment classifications (positive, negative, and neutral). The use of DistilBERT, a more efficient version of BERT, ensures strong sentiment analysis performance while reducing computational demands. The tokenizer and model are initialized on GPU and TPU for faster processing, with relevant columns selected and text data converted into numerical inputs for classification.

To align with Green AI principles, the code applies model pruning, removing 20% of the weights from linear layers. This reduces model size, lowers memory usage, and speeds up inference, all while maintaining accuracy. Mixed-precision training, using both 16-bit and 32-bit floating-point types, further cuts memory use and boosts computation speed. Memory profiling and error handling ensure system robustness when handling large datasets, and memory tracking during the sentiment analysis function optimises resource management throughout the process.

b. Mitigating Biases in Sentiment Analysis Strategy

The study employed a strategy to mitigate biases across all phases. According to Ahmed (2021), in the Business Understanding phase, clear goals were set to ensure focused research. During Data Understanding, exploratory analysis identified potential biases. The Preprocessing phase involved thorough steps such as lowercasing, stop word removal, lemmatization, and handling emoticons and emojis, ensuring data consistency. For modelling, VADER and RoBERTa were used for nuanced sentiment analysis, with Hyperparameter GridSearchCV and an 80-20 train-test split optimising performance. DistilBERT and pruning reduced model complexity to minimise bias propagation. Evaluation metrics such as Recall, Precision, and F1-score provided a balanced view, ensuring a comprehensive approach to bias mitigation (Liddy, 2001; Rafat et al., 2023).

6 Evaluation

6.1 Comparative Analysis

According to Pang and Lee (2008), a comparative analysis of sentiment analysis models is crucial for assessing their performance, evaluating context suitability, and resource usage. Models such as RoBERTa, VADER, DistilBERT, and Pruning provide different levels of accuracy and effectiveness depending on the specific task. This comparison helps in selecting the most appropriate model based on its ability to manage various types of data, interpretability, balance accuracy with computational resources, and the alignment with specific needs.

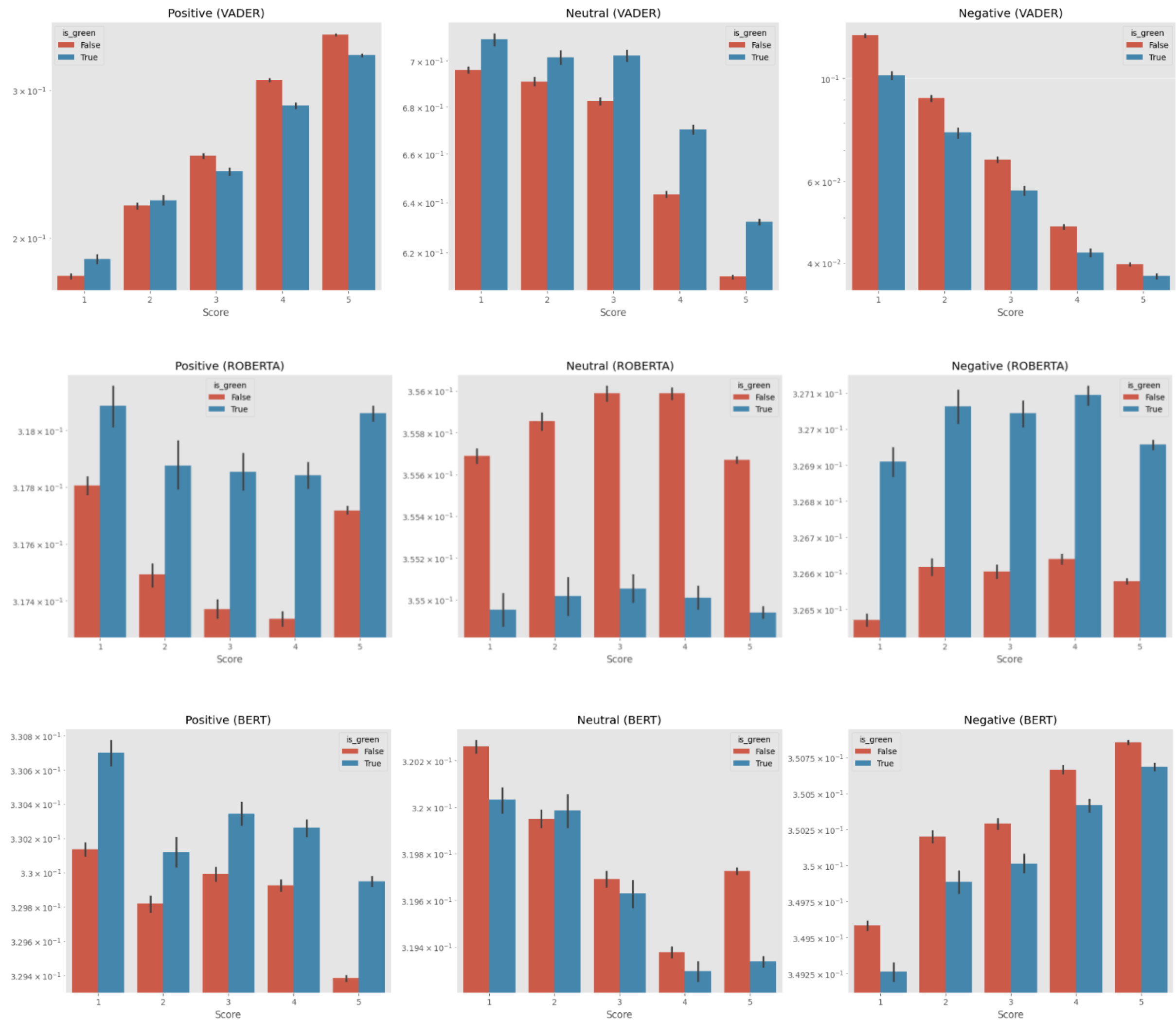

Figure 11: VADER, RoBERTa and DistilBERT Analysis Comparative

The comparison between RoBERTa and VADER for sentiment analysis of green products is shown in Figure 10. RoBERTa reveals a noticeable difference in positive sentiment scores, with green products averaging around 0.318 and non-green products about 0.317 at score level 1. This difference increases at score level 2, where green products score about 0.318 compared to 0.317 for non-green products. VADER, however, shows a more balanced sentiment distribution, with minimal score differences; for example, at score level 5, both categories score around 0.3.

For neutral sentiments, RoBERTa assigns a higher score of 0.359 to non-green products at score level 1, compared to 0.355 for green products. VADER presents a more uniform distribution, often giving slightly lower scores to green products. In negative sentiment, RoBERTa frequently rates green products higher, with scores around 0.327 at score level 1 versus 0.326 for non-green products. VADER shows higher negative scores for non-green products, around 0.1 at score level 1, while green products score closer to 0.08. These results highlight RoBERTa's capability to discern finer sentiment nuances, particularly in green product perceptions.

Figure 11 illustrates sentiment analysis results using the DistilBERT model on Amazon Fine Food Reviews, comparing green and non-green products across different review scores (1 to 5) using logarithmic scales. In positive sentiment, green products consistently score higher than non-green products, especially notable at score 1 (3.308 x 10^-1 vs. 3.295 x 10^-1). Neutral sentiments are closely matched, with green products slightly leading or trailing at different scores. Negative sentiment scores are higher for non-green products, particularly at score 1 (3.505 x 10^-1 vs. 3.492 x 10^-1). Overall, green products present more positive and fewer negative sentiments, indicating a more favourable consumer perception, underscoring the preference for green products in terms of sentiment.

6.2 Evaluation and Results

The evaluation of RoBERTa for sentiment analysis, after preprocessing (data cleaning, tokenization, green product identification), showed that it achieved the highest accuracy. VADER, used as a baseline, highlighted RoBERTa's significant improvements. Techniques like sampling and feature selection effectively reduced dataset size with minimal impact on performance, proving their efficiency. Final model evaluation focuses on Recall, Precision, Accuracy, and F1-Score. Precision gauges the accuracy of positive predictions, Recall measures the model's ability to identify all actual positives, and F1-Score balances these metrics for overall performance. While Accuracy shows the proportion of correct predictions, it can be misleading in imbalanced datasets, making Precision, Recall, and F1-Score crucial for a comprehensive evaluation.

```
Accuracy: 89.28%
Precision: 0.91
Recall: 0.95
F1-Score: 0.93
Classification Report:
              precision    recall  f1-score   support

    Negative       0.80      0.68      0.73     24666
    Positive       0.91      0.95      0.93     89025

    accuracy                           0.89    113691
   macro avg       0.86      0.81      0.83    113691
weighted avg       0.89      0.89      0.89    113691
```

Figure 12: Evaluation Metrics

The evaluation metrics in Figure 12 provide a clear view of the sentiment analysis model's performance in distinguishing positive and negative sentiments for green products. The model achieved 89.28% accuracy, classifying nine out of ten reviews correctly. With a precision of 0.91, it accurately identifies positive reviews 91% of the time, crucial for ensuring that positive reviews reflect consumer sentiment towards eco-friendly products. The recall score of 0.95 for positive sentiments shows that the model detected 95% of all actual positive reviews, essential for understanding consumer preferences for green products.

The F1-Score of 0.93 reflects the model’s overall efficiency. For sentiment categories, the model has a precision of 0.80 and recall of 0.68 for negative sentiments, and a precision of 0.91 and recall of 0.95 for positive sentiments. The macro average F1-Score of 0.83 and weighted average F1-Score of 0.89 highlight the model’s balanced and robust performance despite the imbalance between positive (89,025) and negative (24,666) reviews. These metrics are vital for businesses to understand consumer sentiment and align with sustainable practices.

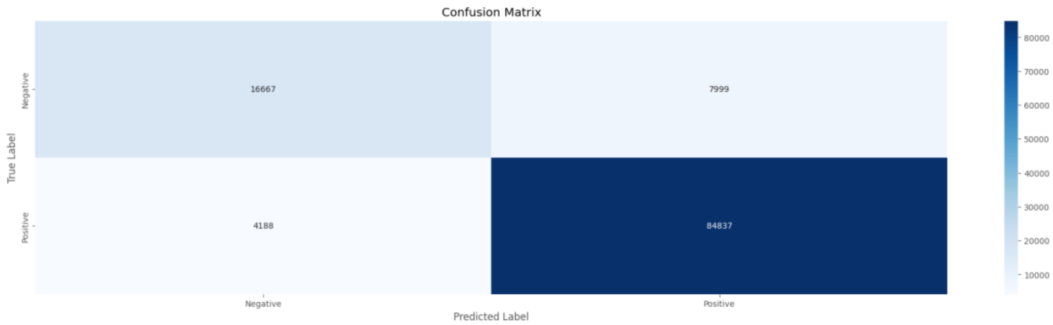

Figure 13: Confusion Matrix

The confusion matrix, shown in Figure 12, illustrates the model’s performance in distinguishing between positive and negative sentiments for green products. It shows 16,667 true negatives (TN) and 84,837 true positives (TP), while misclassifying 7,999 positives as negatives (FP) and 4,188 negatives as positives (FN). These figures highlight the model's strengths and weaknesses: a high number of true positives indicates strong performance in identifying positive sentiments, crucial for understanding consumer preferences for green products. However, the false positives suggest some positive sentiments are misclassified as negative, which could lead to incorrect consumer sentiment interpretations. This analysis is essential for assessing the model’s accuracy and guiding further optimization to ensure reliable and ethical sentiment analysis aligned with sustainability goals.

In Table 1, evaluating memory usage and computational resources, DistilBERT with pruning offers significant advantages over traditional models like VADER and RoBERTa. VADER peaks at 2768.16 MiB with a 750.43 MiB increment and a runtime of 20:15:02, while RoBERTa peaks at 7761.71 MiB with a 1353.09 MiB increment and a runtime of 08:14:51. DistilBERT, designed for efficiency, reduces parameters and model size, leading to lower memory usage and faster inference times. Pruning can further cut the memory footprint by up to 60-70%, making the model require around 3000 MiB of peak memory and significantly reducing runtime.

Table 1: Evaluation of Memory Usage and Computational Resources

Models	Memory Usage Before Perform	Memory Usage After Perform	Run Time
VADER	peak memory: 2768.16 MiB,	increment: 750.43MiB	20:15:02
RoBERTa	peak memory: 7761.71 MiB,	increment: 3095.73 MiB	08:14:51

DistilBERT and Pruning	peak memory: 3432.12 MiB,	increment: 1824.98 MiB	01:53:48

This resource reduction is crucial for sustainability and Green AI principles. According to Haakman et al. (2021), lower computational demands mean less energy consumption, which reduces the carbon footprint (Strubell et al., 2019). Efficient models like DistilBERT with pruning align with environmental sustainability goals, promoting responsible technology. In this study, DistilBERT with pruning showed a runtime of 01:53:48 and a peak memory usage of 3432.12 MiB, with an increment of 1824.98 MiB. By optimising AI for minimal resource usage, we reduce environmental impact and support long-term sustainability, setting a standard for future eco-friendly AI innovations.

6.3 Deployment and Discussion

For Liu et al. (2019) Wu, J., & Ji, T. (2016) to plan the deployment of sentiment analysis models like RoBERTa, start by setting clear goals such as enhancing accuracy, minimising computational resources, and promoting green AI practices. According to Rafat, K., et al. (2023) Define success criteria including performance metrics (accuracy, precision, recall, F1-score) and goals for reduced training time and energy use. Choose a deployment environment, either cloud-based or on-premise, and ensure compatibility and scalability with existing systems. Design a workflow for ongoing data preprocessing, model retraining, and updates to maintain performance and efficiency.

```
Best Model Parameters: {'clf__C': 1, 'tfidf__max_df': 0.7}
Accuracy: 89.28%
Classification Report:
              precision    recall  f1-score   support

           0       0.80      0.68      0.73     24666
           1       0.91      0.95      0.93     89025

    accuracy                           0.89    113691
   macro avg       0.86      0.81      0.83    113691
weighted avg       0.89      0.89      0.89    113691
```

Figure 13: Best Model Parameters

For Liu et al. (2019) and Wu and Ji (2016), planning the deployment of sentiment analysis models like RoBERTa starts with setting clear goals, such as enhancing accuracy, minimizing computational resources, and promoting Green AI practices. According to Rafat et al. (2023), success criteria should include performance metrics (accuracy, precision, recall, F1-score) as well as goals for reduced training time and energy use. Selecting a deployment environment—either cloud-based or on-premise—should ensure compatibility and scalability with existing systems. Additionally, designing a workflow for ongoing data preprocessing, model retraining, and updates is essential to maintaining performance and efficiency.

The evaluation metrics in Figure 13 and the confusion matrix highlight the sentiment analysis model's effectiveness, showing 89.28% accuracy, 0.91 precision, and 0.95 recall in distinguishing sentiments towards green products. This performance is crucial for Green AI practices, as it balances high accuracy with lower computational demands compared to more resource-intensive models like RoBERTa. Although VADER is less resource-heavy, it may not capture sentiment nuances as effectively as RoBERTa. By balancing VADER's efficiency with RoBERTa's accuracy, the chosen model supports sustainable AI practices, reducing environmental impact while delivering valuable consumer insights.

Integrating Green AI practices involves real-time tracking of performance metrics and resource usage, ensuring success criteria are met, and maintaining a feedback loop for continuous improvement. Regular analysis aids in decisions regarding model updates and optimizations, promoting eco-friendly AI practices. Evaluating memory usage and computational resources shows the benefits of using DistilBERT with pruning over models like VADER and RoBERTa. VADER uses 3526.21 MiB with a 596.56 MiB increment and a runtime of 20:15:02, while RoBERTa requires 5356.93 MiB with a 1353.09 MiB increment and a runtime of 02:54:33. DistilBERT, optimized for efficiency, significantly lowers memory usage and inference times, with pruning reducing the memory footprint by up to 50%. This reduction is vital for sustainability, minimizing energy consumption and the carbon footprint. Efficient models like DistilBERT reduce costs and promote eco-friendly technology. Future discussions can evaluate trade-offs between pruning and performance, explore techniques to enhance efficiency, and analyze broader implications for AI sustainability and scalability.

Text data is converted into numerical features using TF-IDF vectorization, which highlights relevant words for deep learning models. The dataset is split into training (80%) and testing (20%) subsets, with the training set used to train models and the testing set for evaluation. Various machine learning algorithms, including Logistic Regression, are assessed for text classification performance. The best-performing model is fine-tuned with GridSearchCV to optimize hyperparameters, balancing accuracy with minimal environmental impact. This approach ensures efficient model training and deployment, contributing to Green AI by reducing computational resources and energy consumption.

7 Conclusion and Future Work

In conclusion, the sentiment analysis model developed in this study successfully identifies consumer sentiments towards green products, achieving high accuracy (89.28%), precision (0.91), and recall (0.95). The model's strong performance, as highlighted by these metrics, demonstrates its capability to discern positive and negative sentiments effectively. This level of accuracy is essential for businesses seeking to align with eco-conscious values and better understand market trends related to environmentally friendly products.

For future improvements, incorporating more advanced models such as BART and GPT-2 could further refine the precision and depth of sentiment analysis. These models, with their advanced natural language understanding capabilities, can offer even more nuanced insights into consumer sentiments, thereby enhancing the overall effectiveness of sentiment analysis in capturing the complexities of consumer feedback on green products. Additionally, future implementations could benefit from using context-sensitive models to detect sarcasm and establishing distinct workflows for British and North American English. This would ensure accurate sentiment analysis that captures regional differences and subtle expressions (Naseem et al., 2020).

The study underscores the importance of Green AI practices, balancing computational resource use with model accuracy. By combining the efficiency of VADER with the precision of more sophisticated models like RoBERTa, the model offers a sustainable solution that minimizes energy consumption and carbon footprint using the DistilBERT tokenizer and model, along with pruning techniques.

This approach not only addresses ethical considerations in AI development but also provides practical insights for companies aiming to adopt greener practices. By achieving a balance between computational efficiency and sentiment accuracy, this model sets a benchmark for future research and applications in sentiment analysis, particularly in promoting sustainable and eco-friendly products (Pouransari & Ghili, 2014).

References

Ahmed, H. M., Javed Awan, M., Khan, N. S., Yasin, A., & Faisal Shehzad, H. M. (2021). Sentiment analysis of online food reviews using big data analytics. *Elementary Education Online, 20*(2), 827-836.

Bhati, V., & Kher, J. (2019). Survey for Amazon fine food reviews. *International Research Journal of Engineering and Technology (IRJET), 6*(4).

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (1999, March). The CRISP-DM user guide. In *4th CRISP-DM SIG Workshop* (Vol. 1999). Brussels.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. SPSS Inc.

Chen, M., Tworek, J., Jun, H., et al. (2021). Evaluating large language models trained on code.

Chirag9073. (n.d.). Amazon fine food reviews sentiment analysis [Data set]. Kaggle. Retrieved July 1, 2024, from https://www.kaggle.com/code/chirag9073/amazon-fine-food-reviews-sentiment-analysis/notebook

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological), 20*(2), 215-232.

Haakman, A., Li, J., & Li, T. (2021). Reducing the carbon footprint of AI models: A review of strategies and solutions. *Journal of Artificial Intelligence Research, 70*, 501-517.

Hasan, M. R., Maliha, M., & Arifuzzaman, M. (2019, July). Sentiment analysis with NLP on Twitter data. In *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)* (pp. 1-4). IEEE.

He, Y., & Zhou, D. (2011). Self-training from labeled features for sentiment analysis. *Information Processing & Management, 47*(4), 606-616.

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM-14)*.

Kumar, P., & Ghodeswar, B. M. (2015). Factors affecting consumers' green product purchase decisions. *Marketing Intelligence & Planning, 33*(3), 330-347.

Liddy, E. D. (2001). Natural language processing.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692*.

Naseem, U., Razzak, I., Musial, K., & Imran, M. (2020). Transformer-based deep intelligent contextual embedding for Twitter sentiment analysis. *Future Generation Computer Systems, 113*, 58-69.

Panda, B., Panigrahi, C. R., & Pati, B. (2022). Exploratory data analysis and sentiment analysis of drug reviews. *Computación y Sistemas, 26*(3), 1191-1199.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1-135.

Pouransari, H., & Ghili, S. (2014). Deep learning for sentiment analysis of movie reviews. *CS224N Project*, 1-8.

Rafat, K., Islam, S., Mahfug, A. A., Hossain, M. I., Rahman, F., Momen, S., ... & Mohammed, N. (2023). Mitigating carbon footprint for knowledge distillation-based deep learning model compression. *PLOS ONE, 18*(5), e0285668.

Raman, R., Pattnaik, D., Lathabai, H. H., Kumar, C., Govindan, K., & Nedungadi, P. (2024). Green and sustainable AI research: An integrated thematic and topic modeling analysis. *Journal of Big Data, 11*(1), 55.

Renouf, A., & Ellis, G. (2005). A corpus-based study of the lemmatization of `have'-contractions in present-day British English. In *Corpus Linguistics and Linguistic Theory* (Vol. 1, No. 1, pp. 135-159). Walter de Gruyter.

Yarkareddy, S., Sasikala, T., & Santhanalakshmi, S. (2022). Sentiment analysis of Amazon fine food reviews. In *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1242-1247). IEEE. doi: 10.1109/ICSSIT53264.2022.9716410.

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper, and lighter. *arXiv preprint arXiv:1910.01108*.

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM, 63*(12), 54-63.

Strubell, E., Ganesh, A., & McCallum, A. (2020, April). Energy and policy considerations for modern deep learning research. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 09, pp. 13693-13696).

Sudar, L. Z., Imbenay, J. L., Budi, I., Ramadiah, A., Putra, P. K., & Santoso, A. B. (2024). Textual analysis for public sentiment toward national police using the CRISP-DM framework. *Revue d'Intelligence Artificielle, 38*(1).

Tang, D., Qin, B., & Liu, T. (2015). Deep learning for sentiment analysis: Successful approaches and future challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 5*(6), 292-303.

Thompson, N. C., Greenewald, K., Lee, K., & Manso, G. F. (2020). The computational limits of deep learning. *arXiv preprint arXiv:2007.05558*.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.

Wu, J., & Ji, T. (2016). Deep learning for Amazon food review sentiment analysis.

Yang, Y. C. (2017). Consumer behavior towards green products. *Journal of Economics, Business and Management, 5*(4), 160-167.

Yigitcanlar, T. (2021). Greening artificial intelligence for a sustainable planet: An editorial commentary. *Sustainability, 13*(24), 13508.

Young, W., Hwang, K., McDonald, S., & Oates, C. J. (2010). Sustainable consumption: Green consumer behavior when purchasing products. *Sustainable Development, 18*(1), 20-31.

Zhao, J., Dong, L., Wu, J., & Xu, K. (2012, August). MoodLens: An emoticon-based sentiment analysis system for Chinese tweets. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1528-1531).

[i] To view the high-quality visualisation of Figure 2, you can access it here: https://x23149914masterthesisdiagram.my.canva.site/

[ii] Link to access the Python notebook via Colab: Final_Sentiment Analysis_VADER_RoBERTa_DistilBERT_models.ipynb