

Article

Understanding Public Opinion towards ESG and Green Finance with the Use of Explainable Artificial Intelligence

Wihan van der Heever ¹, Ranjan Satapathy ², Ji Min Park ³ and Erik Cambria ^{1,*}

¹ College of Computing and Data Science, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore; wihan001@e.ntu.edu.sg

² Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, #16-16 Connexis, Singapore 138632, Singapore; satapathy_ranjan@ihpc.a-star.edu.sg

³ Nanyang Business School, Nanyang Technological University, Gaia, 91 Nanyang Avenue, Singapore 639956, Singapore; kate.park@ntu.edu.sg

* Correspondence: cambria@ntu.edu.sg

Abstract: This study leverages explainable artificial intelligence (XAI) techniques to analyze public sentiment towards Environmental, Social, and Governance (ESG) factors, climate change, and green finance. It does so by developing a novel multi-task learning framework combining aspect-based sentiment analysis, co-reference resolution, and contrastive learning to extract nuanced insights from a large corpus of social media data. Our approach integrates state-of-the-art models, including the SenticNet API, for sentiment analysis and implements multiple XAI methods such as LIME, SHAP, and Permutation Importance to enhance interpretability. Results reveal predominantly positive sentiment towards environmental topics, with notable variations across ESG categories. The contrastive learning visualization demonstrates clear sentiment clustering while highlighting areas of uncertainty. This research contributes to the field by providing an interpretable, trustworthy AI system for ESG sentiment analysis, offering valuable insights for policymakers and business stakeholders navigating the complex landscape of sustainable finance and climate action. The methodology proposed in this paper advances the current state of AI in ESG and green finance in several ways. By combining aspect-based sentiment analysis, co-reference resolution, and contrastive learning, our approach provides a more comprehensive understanding of public sentiment towards ESG factors than traditional methods. The integration of multiple XAI techniques (LIME, SHAP, and Permutation Importance) offers a transparent view of the subtlety of the model's decision-making process, which is crucial for building trust in AI-driven ESG assessments. Our approach enables a more accurate representation of public opinion, essential for informed decision-making in sustainable finance. This paper paves the way for more transparent and explainable AI applications in critical domains like ESG.

Keywords: explainable AI; interpretability; climate change; ESG; aspect-based sentiment analysis; LIME; SHAP; SenticNet

MSC: 68T50



Citation: van der Heever, W.; Satapathy, R.; Park, J.M.; Cambria, E. Understanding Public Opinion towards ESG and Green Finance with the Use of Explainable Artificial Intelligence. *Mathematics* **2024**, *12*, 3119. <https://doi.org/10.3390/math12193119>

Academic Editor: Mudassir M. Rashid

Received: 26 August 2024

Revised: 18 September 2024

Accepted: 2 October 2024

Published: 5 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

This research seeks to explore public perceptions and the value placed on climate change and green energy. By examining public opinion, we aim to uncover trends, patterns, and correlations over time, which will guide future studies on climate change and environmental issues. It also aims to understand how the public perceives and values the efforts of companies in the environmental factors of ESG credit evaluation. In analyzing public opinion data, key areas where companies need to improve their performance, address any public concerns, and win customers' trust can be identified [1]. This study draws on relevant sources that examine the linkage between ESG performance and credit ratings, as well as the significance of ESG factors to the public [2].

The first outcome of this article is to analyze different aspects of public concerns towards climate change and green energy based on explainable artificial Intelligence (XAI) and opinion mining techniques. The aspects include but are not limited to:

1. **Cost:** the cost of transitioning to clean energy sources, such as solar and wind power, as well as the cost of implementing policies and regulations to reduce greenhouse gas emissions may be a concern for some members of the public.
2. **Reliability:** the public may be concerned about the reliability and stability of clean energy sources, as they may not be as consistent as traditional fossil fuels.
3. **Adequate access:** clean energy may not be accessible to everyone, some people may not have access to the necessary infrastructure or resources to take advantage of clean energy sources and may be concerned about the lack of access.
4. **Inconvenience:** some individuals may be concerned about the inconvenience of having to change their lifestyle to adapt to clean energy sources, such as the inconvenience of having to charge electric cars or the potential changes to the way they heat their homes.
5. **Health concerns:** some people may be concerned about the potential health risks associated with certain clean energy technologies, such as wind turbines or solar power plants.
6. **Political will:** some people may be concerned about the political will of the government to tackle the issue of climate change and implement policies to support the transition to clean energy.

The second outcome of this research article is analyzing different aspects of public concerns towards companies' efforts, related to environmental factors of ESG credit evaluation. We focus on the environmental factors of ESG evaluation that were defined by S&P Global Ratings, <https://www.spglobal.com/ratings/en/> (accessed on 23 April 2024), while analyzing the factors from public perceptions. The analyzed factors include:

1. **Climate transition risks:** the potential impacts of the transition to a low-carbon economy on a company's creditworthiness.
2. **Physical risks:** the potential impacts of extreme weather events, natural disasters, and other physical risks on a company's operations and financial performance.
3. **Natural capital:** the value of natural resources and ecosystems, and how a company's operations may affect them.
4. **Waste and pollution:** the potential negative impacts of a company's waste and pollution on human health and the environment.
5. **Other environmental factors:** a company's overall environmental performance and adherence to environmental regulations and standards.

An automatic system that can quantify public opinions about the above aspects in different time ranges is delivered. The system is accurate in performance and explainable in terms of the decision-making process. We also deliver a dataset that is used for training this system.

Both above-mentioned outcomes are valuable for companies and policymakers to better understand public perceptions and concerns related to climate change, green energy and environmental performance, which can inform their strategies and policies to address these concerns more. This value is made more potent with the trustworthy nature of the explainability in which the system provides reliable results. Recent advances in natural language processing have led to improved methods for analyzing ESG-related texts. For instance, refs. [3,4] introduced ESG-BERT, a multi-task learning approach for ESG text analysis. The field of ESG sentiment analysis employs a variety of tools and techniques, including methods such as Support Vector Machines (SVM) and Random Forests, which are used for their interpretability and efficiency, especially with smaller datasets. Advanced models similar to that of LSTM, BERT, and other Transformer-based architectures are employed for their ability to capture complex language patterns and contexts.

Combinations of traditional ML and deep learning techniques aim to balance interpretability with performance, whereas tailored models like ESG-BERT are designed to

capture domain-specific nuances in ESG language [5]. While such approaches have made significant strides, there remains a need for more comprehensive and interpretable methods that can capture the nuances of ESG sentiments, as is explicated in Table 1 below.

Table 1. Comparison of different approaches to ESG sentiment analysis.

Approach	Examples	Strengths	Limitations
Traditional Machine Learning	SVM [6], Random Forests [7]	<ul style="list-style-type: none"> • Interpretable results • Effective for smaller datasets • Computationally efficient 	<ul style="list-style-type: none"> • May struggle with complex language patterns • Limited ability to capture context • Less effective for large-scale data [8]
Deep Learning	LSTM [9], BERT [10], Transformer-based models [11]	<ul style="list-style-type: none"> • Can capture complex language patterns • Effective for large datasets • State-of-the-art performance on many NLP tasks [12] 	<ul style="list-style-type: none"> • Often seen as “black box” models [13] • Require large amounts of data • Computationally intensive
Hybrid Approaches	Combination of traditional ML and deep learning [14]	<ul style="list-style-type: none"> • Balances interpretability and performance • Can leverage strengths of multiple approaches • Potential for improved generalization [15] 	<ul style="list-style-type: none"> • May be complex to implement • Potential for overfitting if not carefully designed • Requires expertise in multiple modeling paradigms
Specific ESG-focused Models	ESGBERT [3], FinBERT [16]	<ul style="list-style-type: none"> • Tailored for ESG-specific language and concepts • Can capture domain-specific nuances • Potentially higher accuracy on ESG tasks [4] 	<ul style="list-style-type: none"> • May have limited generalizability outside ESG domain • Requires specialized training data • Potentially less flexible for evolving ESG concepts

In this study, we address this need by developing a novel multi-task learning framework. While recent work such as [3,4] have applied multi-task learning to ESG text analysis, our framework extends this approach by integrating aspect-based sentiment analysis, co-reference resolution, and contrastive learning with multiple explainable AI techniques (LIME, SHAP, and Permutation Importance), thereby offering a more comprehensive and interpretable analysis of ESG sentiments. The diagram in Figure 1 exhibits the design of the framework. This diagram illustrates the process that begins with data collection and preprocessing of Twitter data related to ESG topics. The core of the framework consists of three main components working in parallel: multi-task learning (including aspect-based sentiment analysis and co-reference resolution), explainable AI techniques (LIME, SHAP, and Permutation Importance), and the SenticGCN for neurosymbolic sentiment analysis.

The outputs from these components feed into a contrastive learning module, which creates a sentiment-aware embedding space. Finally, the framework produces interpretable results and actionable insights for stakeholders in the ESG domain. This approach allows us to capture the complex, multi-faceted nature of public opinion on ESG issues while maintaining interpretability, a crucial factor in the domain of sustainable finance and climate action, as well as accomplishing the outcomes mentioned above.

The remainder of this paper is organized as follows: Section 2 delves into the methodology, detailing our novel multi-task learning framework that combines aspect-based sentiment analysis, co-reference resolution, and contrastive learning; Section 3 reveals the results, consisting of an overview of the data, aspect-based sentiment analysis findings, XAI analyses, contrastive learning visualizations, and practical implications for policymakers and business stakeholders; Section 4 discusses the limitations and challenges of the current model, addressing data source constraints, model limitations, and explainability

trade-offs; finally, Section 5 explores future research directions and potential applications of our framework in ESG and green finance domains.

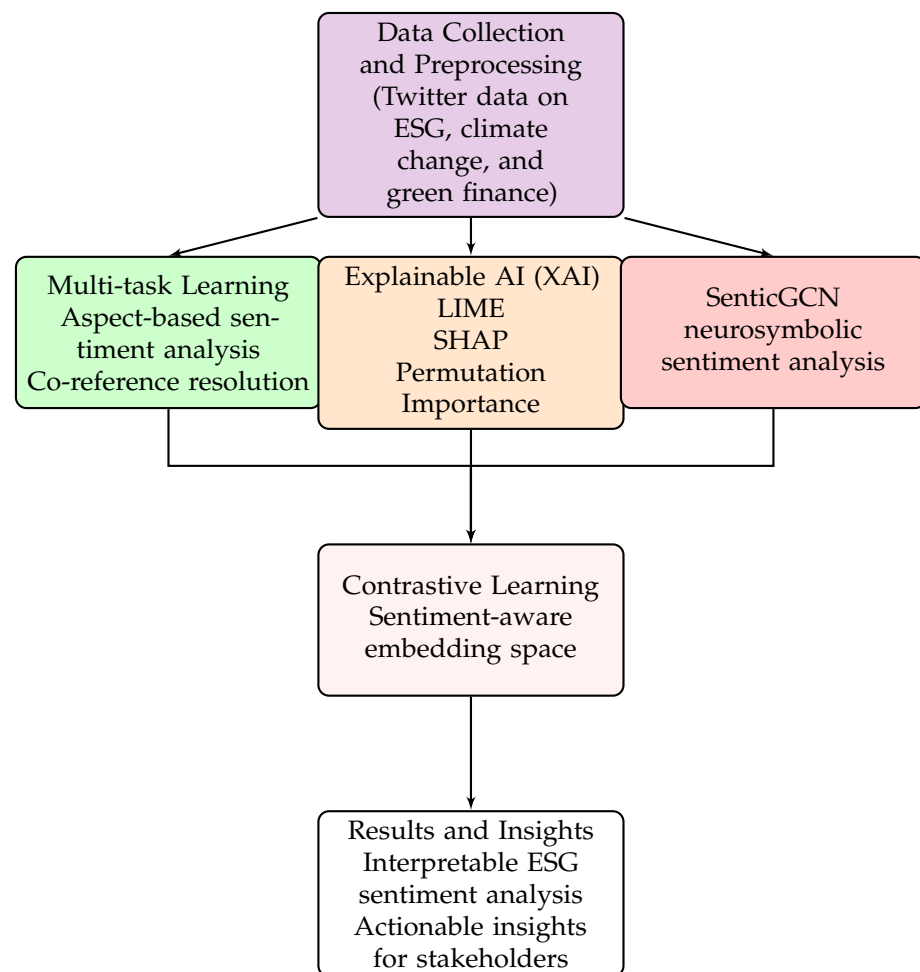


Figure 1. Research framework for ESG sentiment analysis using explainable AI.

2. Materials and Methods

We developed a dataset for aspect-based sentiment analysis on ESG environmental factors, climate change and energy. The dataset is mainly sourced from Twitter. The dataset contains four types of labels: target terms, aspect terms, opinion terms, and sentiment polarities. We study public opinion on the aspect level because a Tweet may express different opinions on different aspects, e.g., “The natural gas company has a factory not far from my home. That factory exploded last year and destroyed many nearby public facilities. Although they invest a lot of money in dealing with pollutant discharge every year, this factory is too risky to us.” In the short text, “natural gas company” is a target; “factory” is an aspect with the opinion of “risky” and the sentiment polarity of “negative”; “dealing with pollutant discharge” is another aspect with the opinion of “invest a lot of money” and the sentiment polarity of “positive”. Noticeably, both aspects belong to the target “natural gas company”, although the latter aspect is connected to the target by referring to “they”. In practice, a model needs to address this to achieve comprehensive opinion mining for the target company. Compared to aspect-based sentiment analysis (ABSA) for typical product/service reviews, the aspect and opinion expressions in our research domain can be more diverse, because the aspects and opinions are grounded in different industries. Considering the contextual information surrounding the aspects is crucial for accurate sentiment polarity prediction. This involves analyzing the linguistic patterns, modifiers, and negation cues that influence the sentiment expressed towards an

aspect. The objective is to detect and categorize them so that they can be mapped to the topics we are interested in.

Thus, the challenges of this task can be summarized as follows:

1. Identifying multiple entities, e.g., targets, aspects, and opinions from diverse industries and climate change topics.
2. Sentiment polarity prediction for aspects.
3. Co-reference resolution for target and aspect detection.
4. Categorizing different aspects by research topics.

Firstly, we developed a multi-task learning and meta-learning-based framework for identifying targets, extracting aspects, and mining opinions. Meta-learning has the advantage of learning multiple tasks with fewer data so that we can achieve better performance given a certain amount of annotated data. The meta-learning approach addresses several challenges specific to ESG analysis, such as the diversity of topics, the scarcity of specialized data, and the need for adaptability in a rapidly evolving field. It is worth noting that while this component allowed us to achieve competitive performance with only one labeled example per class for new ESG-related concepts, it is part of a larger, multi-faceted approach. We also boost multi-task learning by effectuating a soft-parameter sharing mechanism and a loss weight balancing method. The intuition is that the identified targets, aspects, and opinions have certain dependent relationships in a text. Sharing the learned dependent features of each task likely supports the learning of other tasks. However, the learning progress of different tasks may be different. Balancing loss weights of different tasks in multi-task learning can balance the learning progress so that the multi-task learning model can achieve global optima.

We then developed an explainable neurosymbolic framework for sentiment classification based on an improved SenticNet knowledge base. SenticNet is one of the largest and most state-of-the-art sentiment analysis-oriented knowledge bases for neurosymbolic sentic computing [17–19]. The explainable neural framework, namely Sentic Graph Convolutional Network (SenticGCN for short), allows us to access the decision-making process in a human-interpretable way, e.g., understanding what input tokens contribute to the prediction significantly while simultaneously achieving certain generalization abilities in learning feature representations. This approach is similar to that of [20] in the manner of use of a Long Short-Term Memory (LSTM) neural network and contrasts with the work of [21], which utilizes a dual-transformer network instead.

The symbolism-based knowledge representation allows us to access the sentiment basis and reasoning process for predicting a sentiment label. Thus, their combination takes advantage of explainability from different aspects. Regarding the soft-parameter sharing mechanism mentioned earlier, our approach integrates affective knowledge from SenticNet into the graph construction process for GCNs, enhancing word dependencies with affective information. This differs from traditional soft-parameter sharing in that it dynamically adjusts the graph structure based on sentiment-related information, rather than sharing neural network parameters. As to the loss weight balancing method, we employ a weighted loss function (standard cross-entropy loss with L_2 regularization) to give different importance to various classes, which is especially useful in imbalanced datasets, while integrating semantic information see Figure 2.

Thereafter, we created a co-reference resolution model to process texts so that the referral expressions can be mapped to the referred entities. We incorporated common-sense knowledge to boost model performance [22]. Implementing co-reference resolution techniques can help in identifying and linking multiple mentions of the same target or aspect within the tweets. This involves resolving pronouns, anaphoric expressions, and other referring expressions to their antecedents [23]. Co-reference resolution ensures that all mentions of a specific target or aspect are properly identified and grouped together. The intuition is that common-sense knowledge can provide additional information to link pronouns and entities [24]. For example, large ESG investments are often initiated by the

parent company rather than its factories or subsidiaries. Thus, “they” refers to the “natural gas company” rather than the “factory” in the aforementioned example text.

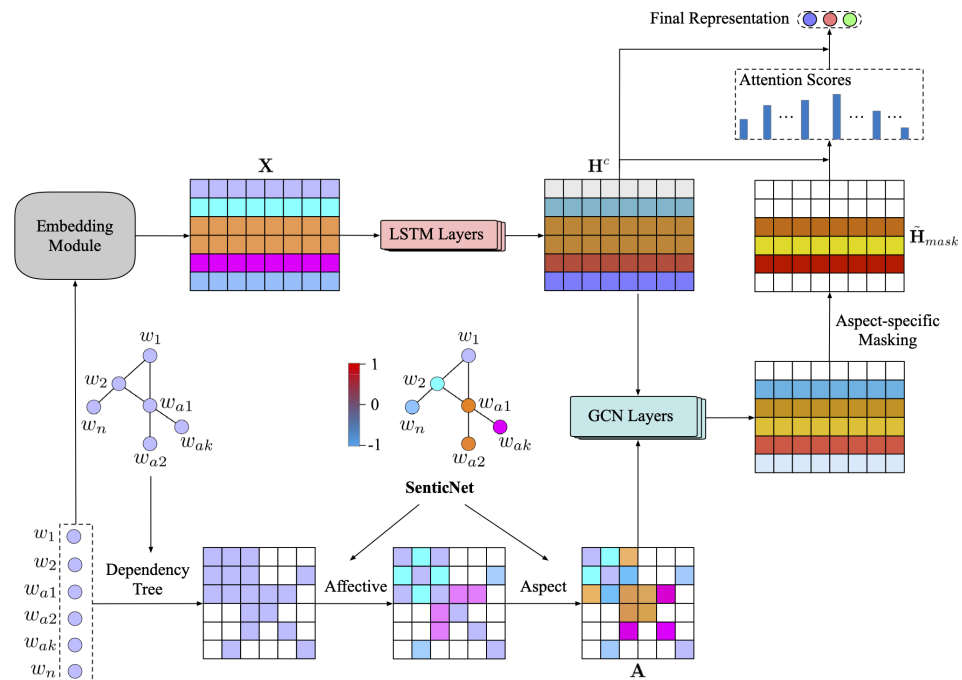


Figure 2. Visual depiction of the inner workings of the SenticGCN model [17].

In addition to this thorough analysis, we employed several explainable artificial intelligence methods to augment the results. These methods, including Local Interpretable Model-agnostic Explanations (LIME), SHapley Additive exPlanations (SHAP), and Permutation Importance (PI), make up a comprehensive XAI strategy that enhances the interpretability, reliability, and actionability of the sentiment analysis model in the nuanced domain of ESG data [25,26]. It does so by providing a robust, global view of feature importance that complements the local explanations of LIME and the unified approach of SHAP [27]. The key idea behind LIME is to approximate the complex model with a simpler, interpretable model (e.g., a linear model) locally around the prediction you want to explain [28].

SHAP is based on cooperative game theory, specifically the Shapley value concept, which assigns a value to each feature that represents its contribution to the final prediction [27]. Permutation Importance measures the importance of a feature by calculating the increase in the model’s prediction error after permuting the feature’s values [7]. Both LIME and Permutation Importance approaches are model-agnostic; however, where LIME might fail to globally capture non-linear and interaction effects across the whole dataset, Permutation Importance compensates for this [29]. SHAP excels at both local and global interpretability by detailing explanations for specific predictions and supplying a grasp of the important features in the entire dataset, respectively [30]. These three mechanisms can greatly add to the model’s reliability as they aid in making complex machine learning models more transparent and boost trust in the decision-making ability of the model, especially in context specific areas, for example in finance, healthcare and legal systems [31]. The XAI methods are also proficient at detecting potential biases and overfitting since they can identify reliance on irrelevant features or expose unexpectedly high importance-dependencies [32]. Finally, we commenced with a contrastive learning method to categorize aspects into our research topics. The main advantage of contrastive learning is that it can be more effective at learning useful representations of the data than traditional supervised learning methods [33]. The intuition is that we can compare the similarity of different aspects and the anchor research topics. Contrastive learning can better distinguish hard negative samples, e.g., the diverse language of aspect terms [34].

To summarize, our proposed approach combines multi-task learning, explainable sentiment analysis, co-reference resolution, and contrastive learning to achieve the objective of comprehensive aspect-based sentiment analysis. We contribute novel algorithms in meta-learning, multi-task learning, neurosymbolic AI, explainable AI, common-sense reasoning, aspect-based sentiment analysis, co-reference resolution, and a new dataset in climate change and ESG research domains [35].

Our preliminary works encompass a developed tool for diverse affective computing tasks, e.g., concept parsing, subjectivity detection, polarity classification, intensity ranking, emotion recognition, aspect extraction, personality prediction, sarcasm identification, depression categorization, toxicity spotting, and engagement measurement [17]. We have published several works for neurosymbolic-based explainable sentiment analysis [36] and contrastive learning [37]. These works are necessary for achieving explainable aspect-based sentiment analysis, which is key to obtaining real-time interpretable insights from social media data.

Our novel approach integrates aspect-based sentiment analysis (ABSA), co-reference resolution, and contrastive learning into a single, cohesive framework. The process begins with data preprocessing, including tokenization, parsing, and named entity recognition to identify ESG-related entities and aspects. We then apply ABSA to identify specific ESG topics and classify the sentiment for each identified aspect using a fine-tuned BERT model. This produces a list of aspects and their associated sentiments for each post.

Next, we employ co-reference resolution, applying a neural model to identify and link mentions of the same entity across the text. This step integrates with the ABSA results to consolidate sentiments for entities mentioned multiple times or in different ways, refining our aspect-sentiment pairs. The framework then utilizes contrastive learning to generate embeddings for each aspect-sentiment pair. This involves creating positive pairs (aspects with similar sentiments) and negative pairs (aspects with different sentiments) and training a contrastive loss function to organize these in the embedding space.

The integration of these techniques is iterative and synergistic. Outputs from ABSA and co-reference resolution feed into the contrastive learning model, whose embeddings in turn refine the ABSA model's classifications. This iterative process allows each component to inform and improve the others, resulting in a more nuanced and accurate analysis. The final output aggregates results from all three components, providing a comprehensive sentiment analysis for each post that includes overall sentiment, sentiment per ESG aspect, and relationships between different aspects based on the contrastive embedding space.

Our framework's novelty lies in this iterative and integrative nature. ABSA provides granular sentiment analysis for specific ESG aspects, while the co-reference resolution enhances accuracy by correctly attributing sentiments to their respective entities. Contrastive learning creates a rich, sentiment-aware embedding space that captures the relationships between different ESG aspects and their associated sentiments. This integrated approach allows us to capture complex, multi-faceted sentiments in ESG-related discussions, handle the challenges of informal social media text, and provide interpretable results that can be traced back to specific aspects and textual elements. Finally, we apply an explainability layer using LIME, SHAP, and Permutation Importance to generate explanations for both individual predictions and global model behavior, enhancing the interpretability of our results.

3. Results

3.1. Overview of the Data

The original dataset consists of a total of 20,961 tweets, all of which relate, to some degree, to either climate change, green finance, ESG or similar topics. The data were collected from Twitter from the time period January 2014 to March 2024 (122 months), using the Twitter API, adhering to the platform's terms of service and data usage policies. Only textual tweets, with time information available, were collected. The dataset was prepared using a semi-supervised method where first the 50 most frequent topics (according

to TopicBERT [38]) and concept words (<https://github.com/SenticNet/concept-parser>, accessed on 3 August 2024) were extracted. Using these words as aspects, the tweets were filtered and annotated with these aspects. Along with this, the sentiment was annotated using the SenticNet API [17].

3.2. Aspect-Based Sentiment Analysis

There are approximately 9000 positive instances, whereas there are roughly 5200 neutral and about 4000 negative instances. This is highlighted in Figure 3 below.

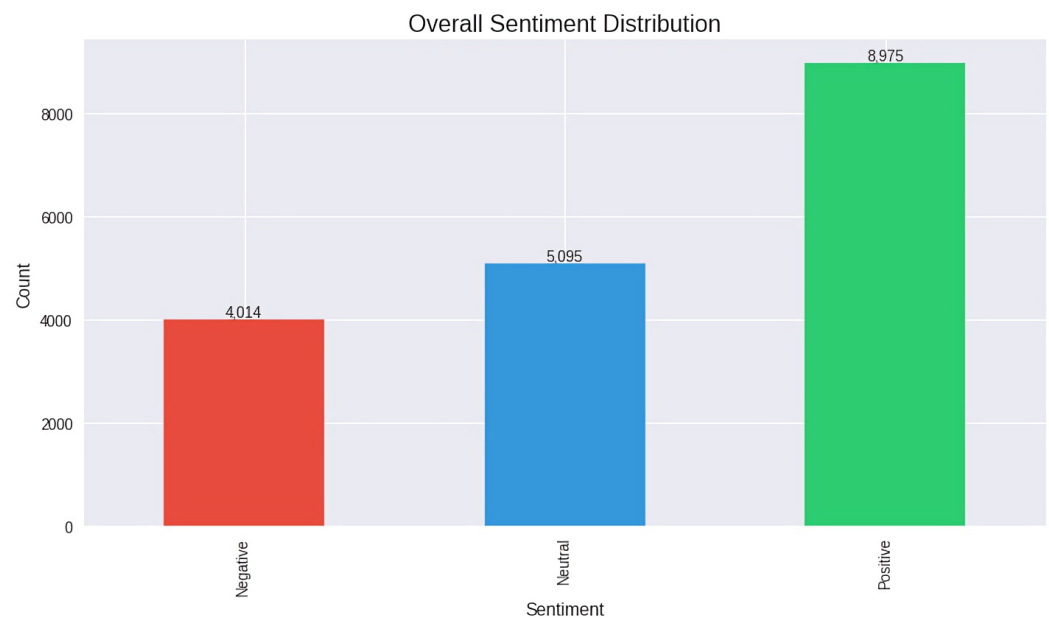


Figure 3. Illustration of the sentiment distribution when the sentiments of the aspects are aggregated.

The chart above clearly shows that positive sentiment dominates the dataset. This suggests that the majority of the analyzed aspects or topics have been perceived favorably in the corpus. While positive sentiment is the most common, there's a significant presence of neutral and negative sentiments, indicating a diverse range of opinions or reactions in the dataset, which is crucial for balanced aspect-based sentiment analysis. This also alludes to a variegated opinion amongst the public on the topics of climate change, ESG and green finance. The balanced distribution (though skewed towards positive) indicates that the sentiment analysis model is capable of distinguishing between different sentiment polarities effectively.

Environment has the highest volume of sentiments, predominantly positive (50.4%). The Social and Governance categories have significantly fewer entries compared to Environment and Other. The Other category, which refers to any non-ESG topics, has a more balanced distribution but still leans positive (40%). Negative sentiments are consistently the lowest across all categories, except in Other where they are exceeded by the percentage of neutral sentiments. Overall, Figure 4 portrays a strong positive bias in ESG-related sentiments, particularly in environmental topics, while governance issues appear under-represented. Our analysis revealed significant variations in sentiment across different ESG categories. Some of the most notable findings include:

1. Environmental (E) Category:
 - Highest overall positive sentiment (50.4%),
 - Key positive aspects: "sustainable", "clean energy", "renewable",
 - Areas of concern: "climate crisis", "pollution",
 - Observed trend: increasing positive sentiment towards renewable energy initiatives.
2. Social (S) Category:

- More balanced sentiment distribution compared to Environmental,
 - Key positive aspects: “diversity”, “employee wellbeing”, “community engagement”,
 - Areas of concern: “workplace safety”, “data privacy”.
 - Observed trend: growing emphasis on corporate social responsibility.
3. Governance (G) Category:
 - Lowest overall sentiment scores among the three categories,
 - Key positive aspects: “transparency”, “ethical leadership”,
 - Areas of concern: “executive compensation”, “corruption”,
 - Observed trend: increasing demand for corporate accountability.
 4. Cross-category Observations:
 - Environmental topics generated the most engagement and strongest sentiments,
 - Governance issues, while fewer in number, often elicited more negative sentiments.
 - Social topics showed the most variability in sentiment, likely due to the diverse range of issues covered.
 5. Temporal Variations:
 - Environmental sentiment showed more volatility, often correlating with major climate events or policy announcements.
 - Governance sentiment remained relatively stable over time.
 - Social sentiment demonstrated a gradual positive trend, possibly reflecting increased corporate focus on social issues.

These variations highlight the complex and multifaceted nature of public sentiment towards ESG issues. They underscore the importance of refined, category-specific approaches in ESG analysis and strategy development.

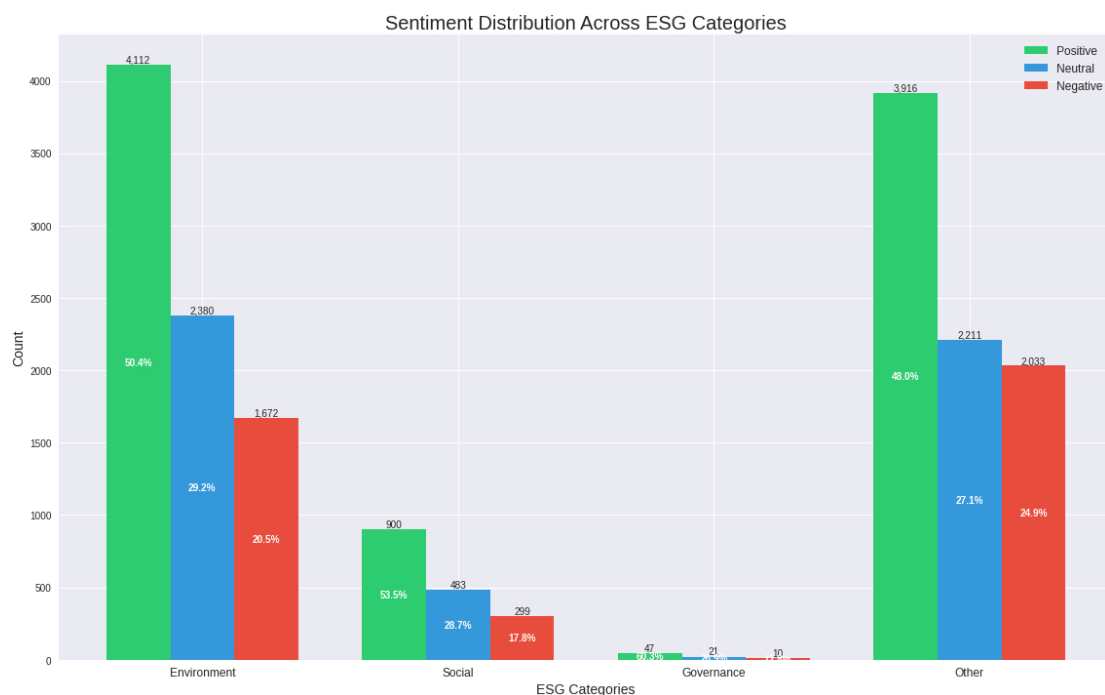


Figure 4. Sentiment distribution depicted by the categorization of Environment, Social, Governance and Other.

3.3. XAI Analyses

3.3.1. SenticGCN

The SenticNet API goes beyond simple sentiment analysis by considering the semantic and contextual information of words and concepts. It captures the polarity (positive or negative) and the intensity of sentiments associated with specific concepts. By combining

3.3.3. SHAP

SHAP, in contrast to LIME, provides a global view of feature importance across the entire dataset [27]. This global perspective helps us identify the most significant ESG aspects or sentiment drivers across all the tweets.

The most notable difference in the SHAP analysis is that it shows a wider range of importance scores (0.549 to 0.889) compared to LIME (0.139 to 0.266), emphasized by the red dotted lines in Figures 5a and 6a (LIME has a mean importance of 0.18, whereas SHAP has a mean importance of 0.65). The high importance of “crisis” aligns with risk assessment in sentiment analysis. The presence of both positive (“growth”, “positive”) and negative (“waste”, “no”) terms suggests the model captures a range of sentiments.

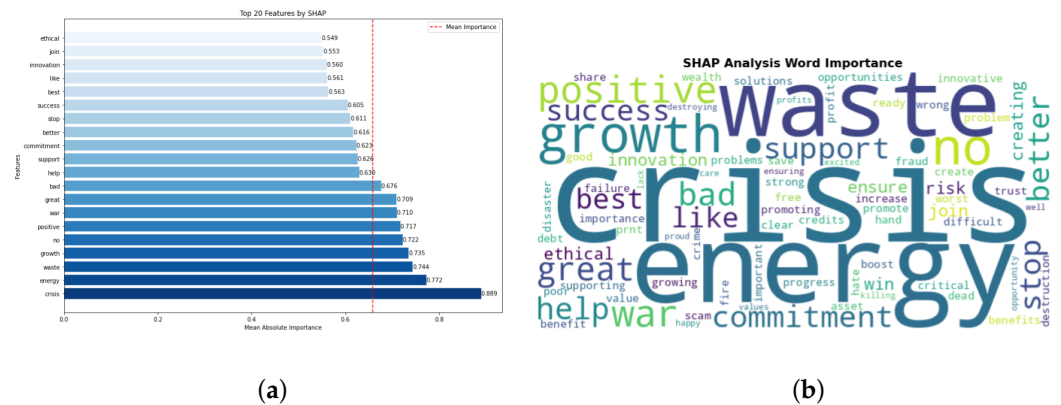


Figure 6. Visualizations of the SHAP analysis on the ESG dataset. (a) Top 20 features according to the SHAP analysis. (b) Word cloud of the most important terms according to the SHAP analysis.

3.3.4. Comparing LIME versus SHAP

It is a valuable exercise to frame these two interpretability methods against each other [41]. We do so in a simple comparative visualization, as seen in the figure below.

The SHAP analysis emphasizes more positive and ESG-related terms, while LIME highlights more crisis-related terms. While LIME focuses on more negative or risk-related terms like “war” and “stop”, SHAP seems to capture a broader range of sentiments and topics, for example, the prominent positive terms like “success” and “growth”. From Figure 7a, it is evident that SHAP generally shows higher importance scores than LIME for most features. This could indicate that the LIME approach is somewhat more prudent in attributing importance to features than that of the SHAP approach. We see that the word “crisis” is the most important for the SHAP analysis, where “war” is the LIME analysis’ top feature. The scatter plot in Figure 7b indicates a correlation of 0.42 between LIME and SHAP scores, suggesting a moderate positive relationship. This is to be expected as the local (specific) predictions should be in line with the global expectation from the data. The comparison of LIME and SHAP analyses reveals their complementary nature. LIME excels in providing detailed, instance-specific explanations, which is crucial for understanding individual predictions. On the other hand, the SHAP approach offers a more comprehensive and consistent global view of feature importance. The moderate correlation (0.42) between their scores suggests that while they often agree on feature importance, they also capture different aspects of the model’s behavior [42]. This complementary approach provides a more holistic understanding of the model, balancing local interpretability with global feature importance, and enhancing the overall explainability of the sentiment analysis model. This difference highlights the value of using multiple XAI techniques to gain a comprehensive understanding of model behavior.

tweet. This helps us understand which aspects or keywords are propelling the model's predictions for the sentiment.

3.3.6. Permutation Importance

This Permutation Importance analysis contributes to the explainability of the ESG sentiment model, offering a clear picture of which linguistic elements most strongly influence the model's sentiment predictions.

Figure 9 shows which words or terms have the most significant impact on the model's predictions for ESG sentiment. Many top features are directly related to ESG concepts, such as "greenwashing", "green", "responsible", "climate", and "sustainability", indicating the model's strong focus on ESG-related terms. However, there's a notable emphasis on environmental terms, suggesting the model may be particularly sensitive to environmental aspects of ESG. This corroborates with the previous results as well. The presence of "business" and "investing" highlights the model's consideration of corporate and financial aspects in ESG sentiment. The PI analysis picks up on words like "transition" and "responsible" that indicate the model's attention to companies' efforts in moving towards more sustainable practices. As an XAI technique, permutation importance reveals not just which features are important, but quantifies their impact on the model's performance, enhancing interpretability. This visualization provides an easily understandable representation of the model's decision-making process, which is valuable for communicating with non-technical stakeholders in the context of climate change, green finance and ESG.

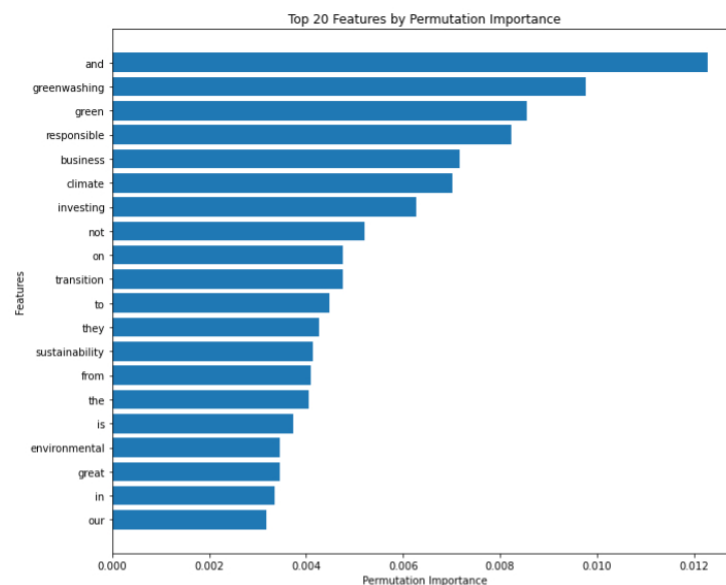


Figure 9. Portrayal of the top 20 features in the dataset according to Permutation Importance.

We used these techniques in combination to provide a comprehensive, multi-level interpretation of our model. LIME offered granular, instance-level explanations. SHAP provided a bridge between local and global interpretations. The amalgamation of LIME and SHAP delivered a balanced perspective of which features are highly contributing to the model. Permutation Importance gave a high-level view of feature importance. By triangulating insights from these three techniques, we were able to validate the model's decision-making process across different levels of granularity. It also enabled us to identify potential biases or unexpected patterns in the model's behavior. In addition to this, we provide stakeholders with clear, interpretable insights into the factors driving ESG sentiment in social media discussions.

3.4. Contrastive Learning

Contrastive learning is a technique used to learn representations by comparing similar and dissimilar samples. The main idea is to push similar samples closer together in the embedding space while pushing dissimilar samples apart. We created positive pairs (samples with the same sentiment) and negative pairs (samples with different sentiments), then utilized a simple encoder (e.g., a small neural network) to create embeddings for each sample. After defining a contrastive loss function that encourages similar samples to have similar embeddings and dissimilar samples to have different embeddings, we trained the model to minimize this loss.

A t-Distributed Stochastic Neighbor Embedding (t-SNE) plot is a visualization technique used to represent high-dimensional data, like the different word embeddings, in a lower-dimensional space, typically 2D or 3D [15]. We use this because it enables us to see patterns, clusters, and relationships between data points with more ease, allowing visual inspection of the model's decision boundaries and making the model's behavior more interpretable to human analysts. The t-SNE plot of the embeddings gathered from the contrastive learning process is depicted below in Figure 10. The clusters exhibit distinct sentiments, which is a positive indicator that the contrastive learning model has successfully captured sentiment-related features in the embeddings. There is an explicit separation between positive (green) and negative (red) sentiments, demonstrating the model's ability to differentiate extreme sentiments. Neutral sentiments (blue) are more spread out and often overlap with both positive and negative clusters. The distribution of neutral sentiments reflect the nuanced nature of neutral sentiments in ESG contexts, where they might share characteristics with both sentiment polarities. The representation continuity (where there is a gradual transition from positive to neutral to negative sentiments) aids in explaining the model's decision-making process for borderline cases.

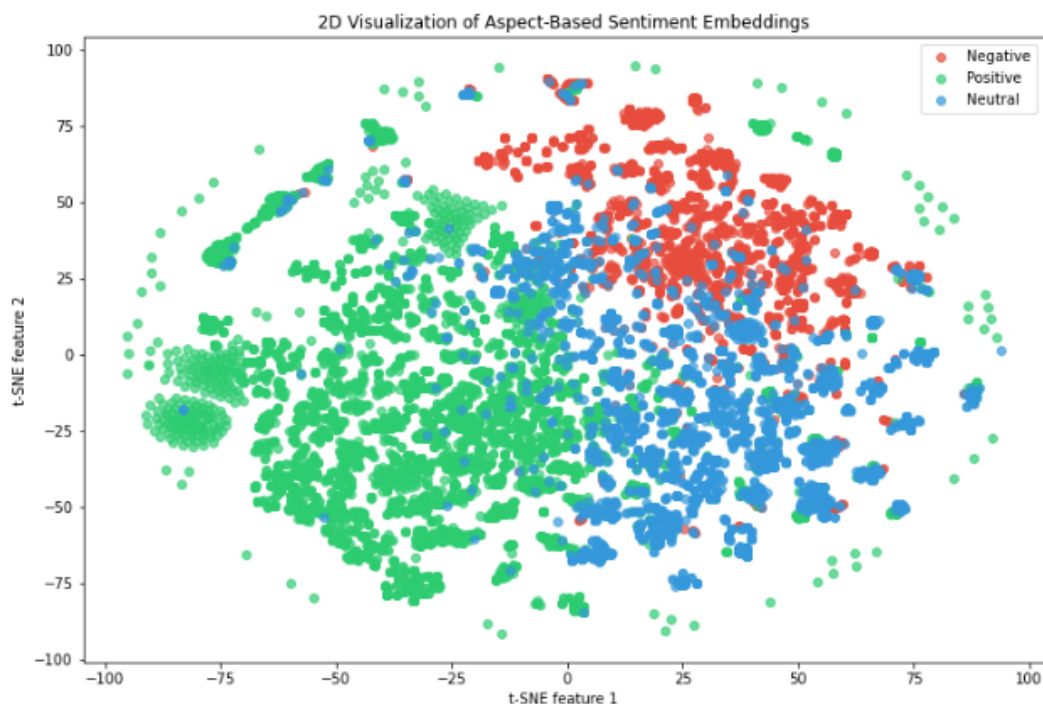


Figure 10. t-SNE plot of the contrastive learning analysis on the ESG dataset. The embeddings were obtained from the aspect terms and the overall sentiment calculated by the SenticGCN model.

Contrastive learning has helped the model learn meaningful features that distinguish between different sentiments without explicit labeling. This unsupervised aspect enhances the model's ability to capture nuanced sentiment indicators in ESG tweets. The visualization, as depicted in Figure 10, provides crucial insights into the sentiment distribution

and clustering of ESG-related topics. This t-SNE plot represents the high-dimensional embeddings of our data in a 2D space, allowing us to visually inspect the relationships between different sentiments and topics. The plot shows a clear clustering of sentiments, with positive (green), negative (red), and neutral (blue) sentiments forming distinct groups. This clustering indicates that our model has successfully learned to differentiate between various sentiment polarities in the context of ESG topics. While there are distinct clusters, we also observe a continuum between sentiments, particularly between neutral and both positive and negative sentiments. This gradual transition reflects the nuanced nature of ESG sentiments, where opinions are not always strictly polarized. The positive sentiment cluster appears to be the largest and most dense, followed by negative, then neutral. This corroborates our earlier findings about the overall sentiment distribution in the dataset and suggests a generally positive public perception of ESG initiatives.

Areas where different sentiment clusters overlap are particularly interesting. These regions represent topics or aspects where public opinion is divided or where the sentiment is more ambiguous. For example, the overlap between positive and neutral clusters might represent ESG initiatives that are viewed favorably by some but with indifference by others. The visualization also reveals outliers and sub-clusters within each sentiment group. These could represent unique or emerging ESG topics that elicit specific sentiments.

In order to better understand these uncertain regions, we map the aspect terms onto the t-SNE cluster centroids and plot them, as shown in Figure 11.

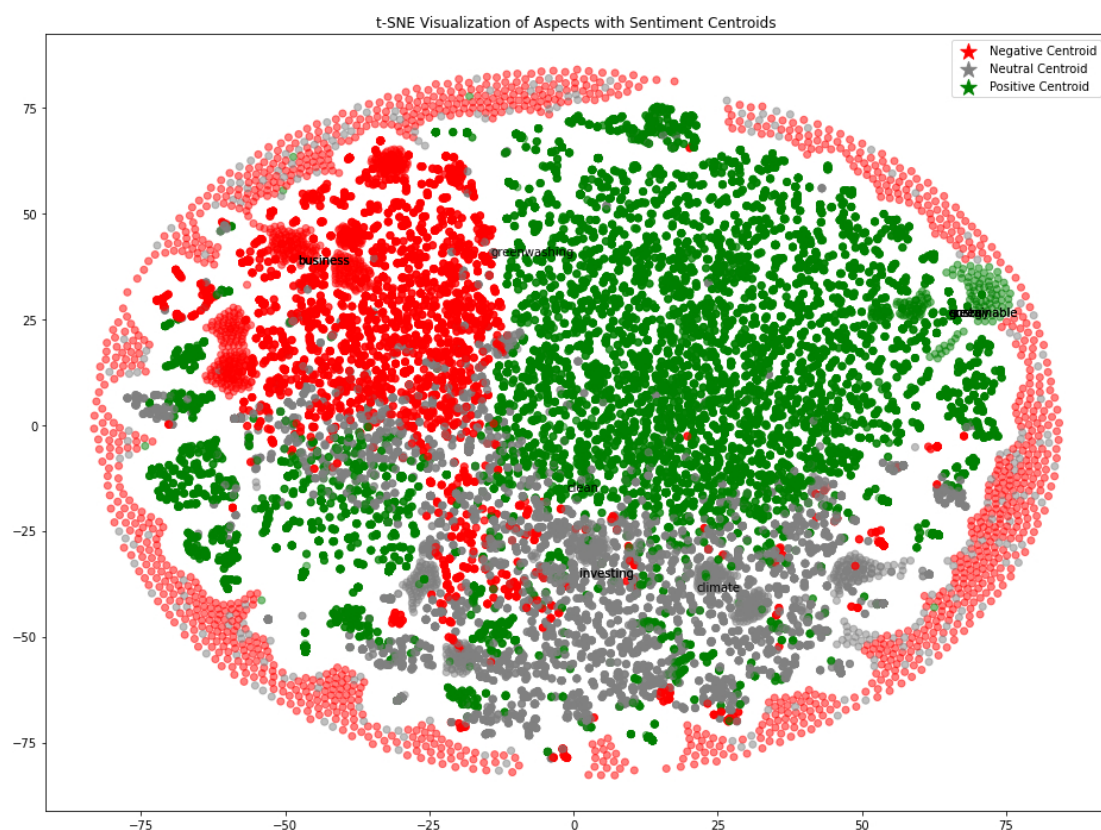


Figure 11. Centroid embedding t-SNE plot from the contrastive learning approach with some important aspect terms superimposed.

From the above figure, we see that the centroids of the positive sentiments correspond to aspects such as “eco”, “sustainable” and “clean”. This is to be expected. The negative centroids include the aspects “greenwashing” and “business”, indicating that the public do not have a positive view of businesses in the context of ESG-related topics. Aspects such as “investing” and “climate” are the neutral sentiment centroids, which can be interpreted as a mix of sentiments (opposing opinions) on these aspects. There are significant areas of

overlap between sentiments, particularly between neutral and both positive and negative. This suggests nuanced or context-dependent aspects that may not have a clear-cut sentiment. The positive sentiment cluster appears to be the largest and most dense, followed by negative, then neutral. This could indicate a bias in the dataset or a tendency for ESG topics to elicit stronger sentiments. The non-circular shapes of the clusters indicate that the sentiment space is not uniformly distributed, reflecting the complexity of ESG sentiment analysis. The importance of this visualization lies in its multifaceted ability to enhance our understanding and application of the model. Firstly, it serves to validate the model's performance, as the clear clustering confirms that our contrastive learning approach has successfully captured meaningful sentiment patterns in the data. Beyond validation, the visualization excels in identifying nuanced sentiments; by showcasing the continuum and overlaps between sentiment clusters, it highlights areas where public opinion is more complex or divided, providing a more accurate representation of the often ambiguous nature of ESG sentiments. Furthermore, the visualization acts as a guide for further analysis, with outliers and sub-clusters potentially indicating interesting topics or aspects that warrant deeper investigation.

This can lead researchers to uncover new insights or emerging trends in ESG sentiment. From a practical standpoint, the visualization holds significant value for stakeholders, offering an intuitive understanding of the sentiment landscape that can directly inform strategy and policy decisions in the ESG domain. Lastly, by providing a visual representation of the model's internal representations, the visualization enhances the interpretability of our AI approach, adding another layer of explainability. This is crucial in the field of ESG, where transparency and understanding of AI-driven insights are paramount for building trust and making informed decisions.

This contrastive learning visualization, therefore, not only demonstrates the effectiveness of our approach but also provides valuable insights into the complex landscape of public sentiment towards ESG topics. It serves as a powerful tool for both validating our methodology and deriving actionable insights for stakeholders in the ESG domain.

3.5. Practical Implications for Policymakers and Business Stakeholders

The insights derived from our explainable AI approach to ESG sentiment analysis have significant practical implications for both policymakers and business stakeholders. Here, we elaborate on how these findings can influence decision-making in ESG and green finance:

3.5.1. Implications for Policymakers

Sentiment trends can guide the development of new policies or the refinement of existing ones. For instance, strong positive sentiment toward renewable energy could support policies promoting green technology, whereas areas with predominantly negative sentiment might indicate where additional regulation or public education is needed. Policymakers can also use sentiment analysis to prioritize ESG initiatives based on public interest and concern, ensuring efficient allocation of resources. Understanding public sentiment can help in crafting more effective communication strategies about ESG policies, addressing concerns, and highlighting positively perceived initiatives. On a global scale, sentiment analysis can inform international policy harmonization efforts, highlighting areas where global consensus exists or where regional differences need to be addressed. Additionally, rapid shifts in sentiment could serve as an early warning system for emerging ESG issues, allowing policymakers to proactively address potential crises [43].

3.5.2. Implications for Business Stakeholders

For business stakeholders, the implications of this research are equally profound. Companies can align their ESG strategies with public sentiment, focusing on areas that resonate positively with the public and addressing aspects that generate concern [44]. This alignment can lead to more effective risk management, as sentiment analysis can help

identify potential reputational risks related to ESG factors, allowing businesses to proactively manage these risks [45]. The insights gained from sentiment analysis can also drive innovation in products and services that align with positive ESG sentiments, potentially creating new market opportunities. For investors, understanding public sentiment can inform investment decisions, potentially predicting market trends related to ESG factors [46]. This comprehension of public sentiment can also guide companies in their ESG reporting and disclosure practices, focusing on aspects that the public deems most important.

3.5.3. Collaborative Opportunities

The implications of this research extend beyond individual sectors, offering opportunities for collaboration between public and private entities. Shared understanding of ESG sentiments can facilitate more effective public–private partnerships in addressing ESG challenges [47]. Both policymakers and businesses can use these insights to develop educational initiatives that address misconceptions and promote an understanding of ESG concepts.

By leveraging the granular, interpretable insights provided by our explainable AI approach [1,48,49], policymakers and business stakeholders can make more informed, data-driven decisions in the rapidly evolving landscape of ESG and green finance. This can lead to more effective policies, better-aligned business strategies, and ultimately, progress toward sustainable development goals [50].

4. Limitations and Challenges

While our study provides valuable insights into public sentiment towards ESG and green finance using explainable AI techniques, it is important to acknowledge several limitations and challenges:

1. **Data Source Limitations:**
 - Social media bias: our primary data source (Twitter) may not be representative of the entire population, potentially skewing results towards users of this platform.
 - Language constraints: although we used multi-language sentiment analysis, the majority of our data was in English, potentially underrepresenting non-English speaking populations.
2. **Model Limitations:**
 - Context interpretation: despite using advanced NLP techniques, our model may sometimes struggle with complex context, sarcasm, or culturally specific references.
 - Evolving language: the rapid evolution of ESG-related terminology poses a challenge for keeping the model up-to-date.
3. **Temporal Challenges:**
 - Short-term fluctuations: public sentiment can be highly volatile, especially in response to current events, which may not reflect long-term trends.
 - Historical data limitations: our analysis may be biased towards more recent data due to the increasing prevalence of ESG discussions on social media platforms.
4. **Explainability Trade-offs:**
 - Complexity vs. Interpretability: while we strive for model explainability, there is an inherent trade-off between model complexity (which can capture nuanced patterns) and ease of interpretation.
 - Feature importance variability: different explainability methods (LIME, SHAP, Permutation Importance) sometimes yield slightly different results, which can be challenging to reconcile.
5. **Sentiment Ambiguity:**

distinguishing between a truly neutral sentiment and lack of opinion was sometimes challenging, potentially affecting our analysis of less polarizing topics.

6. Scope Limitations:

- Geographic bias: despite efforts to maintain a global perspective, our data may be skewed toward regions with higher social media engagement on ESG topics.
- Industry-specific refinement: while we aimed for a broad understanding of ESG sentiment, industry-specific nuances may not be fully captured in our general model.

7. Validation Challenges:

- establishing a reliable “ground truth” for sentiment in ESG contexts is challenging, making it difficult to definitively validate our model’s accuracy.

These limitations and challenges highlight areas for potential improvement in future research. They also underscore the importance of interpreting our results within the context of these constraints. Despite these limitations, we believe our study provides valuable insights into public sentiment towards ESG and green finance and contributes significantly to the development of explainable AI in this domain.

5. Discussion

Future research directions for us include alternative topic modeling techniques, such as Latent Dirichlet Allocation (LDA) or Non-negative Matrix Factorization (NMF), to identify aspects that can help in categorizing them into broader research topics [8]. Topic modeling algorithms discover latent themes or topics within a collection of texts based on the co-occurrence patterns of words. Another possible direction is aligning the identified aspects with an established ESG taxonomy or framework, which can facilitate their categorization into predefined research topics. Techniques like dependency parsing or rule-based approaches can be employed to capture the contextual nuances. This could be especially useful for naturally occurring categorizations, such as the ESG-specific context.

If labeled data are available, a supervised classification model can be trained to automatically categorize aspects into research topics. This requires a dataset where aspects are manually assigned to their corresponding research topics by an expert in the field of ESG, which can then be used to train a classifier to predict the topic categories for new or unseen aspects. Alternatively, the focus of the current model can shift from social media (capturing public sentiment, understanding grassroots movements, and identifying emerging trends) to news articles, which could potentially be more reliable, in-depth, and context-rich.

Moreover, future research directions could explore the integration of this methodology with large language models for more nuanced text understanding, or the application of these techniques to other domains requiring high levels of transparency and explainability in AI decision-making.

There is a myriad of potential future applications, including, but not limited to, real-time ESG risk assessment for investment portfolios, automated ESG reporting and compliance monitoring, predictive modeling of public sentiment towards new sustainability initiatives and personalized ESG strategy recommendations for corporations.

Author Contributions: Writing—original draft, W.v.d.H.; Writing—review & editing, R.S. and J.M.P.; Supervision, E.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research/project is supported by the Ministry of Education, Singapore under its MOE Academic Research Fund Tier 2 (STEM RIE2025 Award MOE-T2EP20123-0005) and by the RIE2025 Industry Alignment Fund—Industry Collaboration Projects (IAF-ICP) (Award I2301E0026), administered by A*STAR, as well as supported by Alibaba Group and NTU Singapore.

Data Availability Statement: Dataset available upon request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

XAI	eXplainable Artificial Intelligence
ESG	Environment, Social and Governance
NLP	Natural Language Processing
ABSA	Aspect-Based Sentiment Analysis
SenticGCN	SenticNet Graph Convolutional Network
SVM	Support Vector Machines
LSTM	Long Short-Term Memory
BERT	Bidirectional Encoder Representations from Transformers
LIME	Local Interpretable Model-agnostic Explanations
SHAP	SHapley Additive exPlanations
API	Application Programming Interface
t-SNE	t-Distributed Stochastic Neighbor Embedding
LDA	Latent Dirichlet Allocation
NMF	Non-negative Matrix Factorization

References

1. Yeo, W.J.; van der Heever, W.; Mao, R.; Cambria, E.; Satapathy, R.; Mengaldo, G. A comprehensive review on financial explainable AI. *arXiv* **2023**, arXiv:2309.11960.
2. Du, K.; Xing, F.; Mao, R.; Cambria, E. Financial sentiment analysis: Techniques and applications. *ACM Comput. Surv.* **2024**, *56*, 1–42. [[CrossRef](#)]
3. Mehra, S.; Louka, R.; Zhang, Y. Esgbert: Language model to help with classification tasks related to companies environmental, social, and governance practices. *arXiv* **2022**, arXiv:2203.16788.
4. Montariol, S.; Martinc, M.; Pelicon, A.; Pollak, S.; Koloski, B.; Lončarski, I.; Valentinčič, A.; Šuštar, K.S.; Ichev, R.; Žnidaršič, M. Multi-task Learning for Features Extraction in Financial Annual Reports. In *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 7–24.
5. Chaturvedi, I.; Ong, Y.S.; Tsang, I.; Welsch, R.; Cambria, E. Learning word dependencies in text by means of a deep recurrent belief network. *Knowl.-Based Syst.* **2016**, *108*, 144–154. [[CrossRef](#)]
6. Cortes, C. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
7. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
8. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: Berlin/Heidelberg, Germany, 2017.
9. Gandhi, U.D.; Malarvizhi Kumar, P.; Chandra Babu, G.; Karthick, G. Sentiment analysis on twitter data by using convolutional neural network (CNN) and long short term memory (LSTM). *Wirel. Pers. Commun.* **2021**, 1–10. [[CrossRef](#)]
10. Devlin, J. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
11. Vaswani, A. Attention is all you need. *arXiv* **2017**, arXiv:1706.03762.
12. Cambria, E. *Understanding Natural Language Understanding*; Springer: Berlin/Heidelberg, Germany, 2024; ISBN 978-3-031-73973-6.
13. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A survey of methods for explaining black box models. *ACM Comput. Surv.* (CSUR) **2018**, *51*, 1–42. [[CrossRef](#)]
14. Ghorbanali, A.; Sohrabi, M.K.; Yaghmaee, F. Ensemble transfer learning-based multimodal sentiment analysis using weighted convolutional neural networks. *Inf. Process. Manag.* **2022**, *59*, 102929. [[CrossRef](#)]
15. Wang, Y.; Huang, H.; Rudin, C.; Shaposhnik, Y. Understanding how dimension reduction tools work: An empirical approach to deciphering t-SNE, UMAP, TriMAP, and PaCMAP for data visualization. *J. Mach. Learn. Res.* **2021**, *22*, 1–73.
16. Araci, D. FinBERT: Financial Sentiment Analysis with Pre-trained Language Models. *arXiv* **2019**, arXiv:1908.10063.
17. SenticNet API. 2024. Available online: <https://sentic.net/api> (accessed on 19 August 2024).
18. Cambria, E.; Mao, R.; Han, S.; Liu, Q. Sentic parser: A graph-based approach to concept extraction for sentiment analysis. In *Proceedings of the 2022 IEEE International Conference on Data Mining Workshops (ICDMW)*, Orlando, FL, USA, 28 November–1 December 2022; pp. 413–420.
19. Mao, R.; Du, K.; Ma, Y.; Zhu, L.; Cambria, E. Discovering the cognition behind language: Financial metaphor analysis with MetaPro. In *Proceedings of the 2023 IEEE International Conference on Data Mining (ICDM)*, Shanghai, China, 1–4 December 2023; pp. 1211–1216.
20. Wu, H.; Huang, C.; Deng, S. Improving aspect-based sentiment analysis with Knowledge-aware Dependency Graph Network. *Inf. Fusion* **2023**, *92*, 289–299. [[CrossRef](#)]
21. Tang, H.; Ji, D.; Li, C.; Zhou, Q. Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online, 5–10 July 2020; pp. 6578–6588.
22. Clark, K.; Manning, C.D. Deep reinforcement learning for mention-ranking coreference models. *arXiv* **2016**, arXiv:1609.08667.

23. Lee, K.; He, L.; Zettlemoyer, L. Higher-order coreference resolution with coarse-to-fine inference. *arXiv* **2018**, arXiv:1804.05392.
24. Raghunathan, K.; Lee, H.; Rangarajan, S.; Chambers, N.; Surdeanu, M.; Jurafsky, D.; Manning, C.D. A multi-pass sieve for coreference resolution. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Cambridge, MA, USA, 9–11 October 2010; pp. 492–501.
25. Adadi, A.; Berrada, M. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [[CrossRef](#)]
26. Du, K.; Xing, F.; Mao, R.; Cambria, E. An evaluation of reasoning capabilities of large language models in financial sentiment analysis. In Proceedings of the IEEE Conference on Artificial Intelligence (IEEE CAI), Singapore, 25–27 June 2024; pp. 189–194.
27. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017, Volume 30.
28. Ribeiro, M.T.; Singh, S.; Guestrin, C. “Why should i trust you?” Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144.
29. Molnar, C. *Interpretable Machine Learning*; Lulu. com: Morrisville, NC, USA, 2020.
30. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [[CrossRef](#)]
31. Gunning, D.; Aha, D. DARPA’s explainable artificial intelligence (XAI) program. *AI Mag.* **2019**, *40*, 44–58.
32. Lipton, Z.C. The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue* **2018**, *16*, 31–57. [[CrossRef](#)]
33. Chen, T.; Kornblith, S.; Norouzi, M.; Hinton, G. A simple framework for contrastive learning of visual representations. In Proceedings of the 37th International Conference on Machine Learning, PMLR 119, Vienna, Austria, 12–18 July 2020; pp. 1597–1607.
34. Khosla, P.; Teterwak, P.; Wang, C.; Sarna, A.; Tian, Y.; Isola, P.; Maschinot, A.; Liu, C.; Krishnan, D. Supervised contrastive learning. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 18661–18673.
35. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [[CrossRef](#)] [[PubMed](#)]
36. Cambria, E.; Zhang, X.; Mao, R.; Chen, M.; Kwok, K. SenticNet 8: Fusing emotion AI and commonsense AI for interpretable, trustworthy, and explainable affective computing. In Proceedings of the International Conference on Human-Computer Interaction (HCI), Washington, DC, USA, 29 June–4 July 2024.
37. Nguyen, T.; Luu, A.T. Contrastive learning for neural topic model. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 11974–11986.
38. Zhou, Y.; Liao, L.; Gao, Y.; Wang, R.; Huang, H. TopicBERT: A topic-enhanced neural language model fine-tuned for sentiment classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *34*, 380–393. [[CrossRef](#)] [[PubMed](#)]
39. Garreau, D.; Luxburg, U. Explaining the explainer: A first theoretical analysis of LIME. In Proceedings of the 26nd International Conference on Artificial Intelligence and Statistics, PMLR, Online, 26–28 August 2020; pp. 1287–1296.
40. Zafar, M.R.; Khan, N. Deterministic local interpretable model-agnostic explanations for stable explainability. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 525–541. [[CrossRef](#)]
41. Vij, A.; Nanjundan, P. Comparing strategies for post-hoc explanations in machine learning models. In Proceedings of the Mobile Computing and Sustainable Informatics: Proceedings of ICMCSI 2021, Patan, Nepal, 27–28 January 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 585–592.
42. Gujarati, D.N. *Essentials of Econometrics*; Sage Publications: Thousand Oaks, CA, USA, 2021.
43. Flammer, C.; Toffel, M.W.; Viswanathan, K. Shareholder activism and firms’ voluntary disclosure of climate change risks. *Strateg. Manag. J.* **2021**, *42*, 1850–1879. [[CrossRef](#)]
44. Dremptic, S.; Klein, C.; Zwergel, B. The influence of firm size on the ESG score: Corporate sustainability ratings under review. *J. Bus. Ethics* **2020**, *167*, 333–360. [[CrossRef](#)]
45. Giese, G.; Lee, L.E.; Melas, D.; Nagy, Z.; Nishikawa, L. Foundations of ESG investing: How ESG affects equity valuation, risk, and performance. *J. Portf. Manag.* **2019**, *45*, 69–83. [[CrossRef](#)]
46. Consolandi, C.; Phadke, H.; Hawley, J.; Eccles, R.G. Material ESG outcomes and SDG externalities: Evaluating the health care sector’s contribution to the SDGs. *Organ. Environ.* **2020**, *33*, 511–533. [[CrossRef](#)]
47. Ma, M.; Wang, N.; Mu, W.; Zhang, L. The instrumentality of public-private partnerships for achieving Sustainable Development Goals. *Sustainability* **2022**, *14*, 13756. [[CrossRef](#)]
48. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bannetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [[CrossRef](#)]

-
49. Ong, K.; van der Heever, W.; Satapathy, R.; Cambria, E.; Mengaldo, G. FinXABSA: Explainable finance through aspect-based sentiment analysis. In Proceedings of the 2023 IEEE International Conference on Data Mining Workshops (ICDMW), Shanghai, China, 4 December 2023; pp. 773–782.
 50. Cihon, P.; Schuett, J.; Baum, S.D. Corporate governance of artificial intelligence in the public interest. *Information* **2021**, *12*, 275. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.