CSECU-DSG at SemEval-2025 Task 6: Exploiting Multilingual Feature Fusion-based Approach for Corporate Promise Verification

Tashin Hossain[1], Abu Nowshed Chy[2],
[1,2]Department of Computer Science and Engineering,
University of Chittagong, Chattogram-4331, Bangladesh
[1]tashin.hossain.cu@gmail.com, [2] nowshed@cu.ac.bd

Abstract

Trust and commitment are fundamental to personal, professional, and legal interactions, particularly in an era where digital platforms facilitate communication and transactions. Ensuring the authenticity and fulfillment of promises has become a critical concern for individuals and organizations, necessitating a robust verification mechanism. To address this challenge, SemEval-2025 Task 6 introduced a promise verification task, encompassing five different languages, which involves analyzing multi-industrial reports, including corporate disclosures, ESG reports, and legal contracts etc. In response to this challenge, we propose a multilingual promise verification framework that integrates textual analysis, contextual understanding, and probabilistic assessment to evaluate the validity and fulfillment of given promises. Our approach leverages a feature fusion of LASER and USE, employing a Bi-LSTM neural network architecture combined with an MLP for the identification of promises and supporting evidence within documents. Experimental evaluations conducted using the ML-Promise dataset demonstrate that our system achieves competitive performance across multiple languages.

1 Introduction

In today's society, corporate, governmental, and public personalities' pledges have the power to shape public perception, influence stakeholder trust, and determine institutional reputation. These organizations' promises of social responsibility, environmental responsibility, and ethical governance are important markers of their legitimacy and accountability. However, the quantity and magnitude of such commitments make it extremely difficult to confirm whether they are actually being fulfilled. It is now more important than ever to be able to verify a promise, especially in the world of business, where corporations commonly make grand claims about their impact on Environmental, Social, and Governance (ESG) measures.

In order to tackle this problem, SemEval-2025 introduces ML-Promise, the first multilingual dataset created especially for promise verification (Chen et al., 2025). This dataset allows for a cross-cultural analysis of corporate promises verifying four different evaluation criteria, described in Table 1.

In recent years, the application of Natural Language Processing (NLP) in ESG analysis and sustainability reporting has become increasingly common. For instance, Gutierrez-Bustamante et al. (Gutierrez-Bustamante and Espinosa-Leal, 2022) employed Latent Semantic Analysis (LSA) and Global Vectors (GloVe) for word representation to assess the alignment of sustainability reports with the Global Reporting Initiatives (GRI) framework. Gorovaia et al. (Gorovaia and Makrominas, 2024) employed text analysis of corporate social responsibility (CSR) reports to identify reporting inconsistency between violator companies that violate environmental infractions and non-violator companies. Moreover, the ESGReveal system, introduced by Zou et al. (Zou et al., 2025), blends large language models (LLMs) with retrieval-augmented generation (RAG) to retrieve structured ESG details from ESG reports.

In this paper, we illustrate our insights accumulated from experimenting on this task. We proposed feature fusion based neural architecture, to identify the promise and evidence statements. Here, we utilized the Language-Agnostic SEntence Representations (LASER) and Universal Sentence Encoder (USE) embedding for the feature fusion.

The structure of this paper is as follows: Section 2 provides a detailed explanation of our proposed framework. In Section 3, we present the experimental setup along with a comparative performance analysis. In Section 4, we provide our insights and explanability of the models on the task. Lastly, we conclude the paper in Section 5, discussing poten-

Proceedings of the The 19th International Workshop on Semantic Evaluation (SemEval-2025), pages 1849–1858
July 31 - August 1, 2025 ©2025 Association for Computational Linguistics

Label	Description	Possible Values
Promise Identification	Does the statement contain a promise?	Yes / No
Supporting Evidence	Is there evidence backing the claim?	Yes / No
Clarity of Promise	Is the promise clear, unclear, or misleading?	Clear / Not Clear / Misleading
Timeline Verification	When can the promise be verified?	<2 years, 2-5 years, >5 years, Other

Table 1: Promise Evaluation Criteria

tial future directions.

2 System Overview

We shape the corporate promise verification task as a sequence classification task and employ an ensemble of embedding models, LASER, and USE for the feature extraction process, then combine these two embedding layers, and feed them into the Bi-LSTM + MLP layer to get the desired label. The framework of our system is depicted in Figure 1.

2.1 Data Preprocessing Techniques

Preparing documents for analysis is an essential part of NLP, which turns unstructured text data into understandable text. First, we extract the text from documents using PyPDF2 (Li et al., 2023). Following that, we eliminating noise, such as extraneous punctuation, special characters, HTML tags, or unrelated information, that could impair model performance.

2.2 LASER Embedding

Laser (Language-Agnostic SEntence Representations) (Schwenk and Douze, 2017; Artetxe and Schwenk, 2019; Schwenk and Li, 2018) is a unified model that generates high-quality sentence embeddings for over 90 languages. We utilized this multilingual feature embedding model for the feature extraction. After text preprocessing, we gather the 1024-dimensional feature embedding from this model to feed into our final system.

2.2.1 Universal Sentence Encoder (USE)

Universal Sentence Encoder (Cer et al., 2018) is a language-independent, fixed-dimensional vector form of representing text that embeds semantic meaning in a language-independent, domain-independent, and task-independent manner. We extracted the 1024-dimensional feature embedding from the sentence encoder to feed into our system.

2.3 Bi-LSTM + MLP

Bidirectional Long Short-Term Memory (Bi-LSTM) (Liu and Guo, 2019; Zhang and Rao, 2020; Deng et al., 2021) is a complex form of LSTM for improving sequential information processing with both past and future contextual information capture. Unlike a traditional LSTM, processing one direction sequentially, BiLSTM consists of two LSTM layers in parallel, one processing in sequence direction and one in the reverse direction. Output of both directions is then combined, and both past and future context can be utilized in prediction by the model. For our proposed framework, we concatenate features coming from the LASER and USE model, feed them into the Bi-LSTM layer for training, and then go through the MLP layer before getting the final predictions.

2.4 Model's Prediction

Following the process of feature fusion, the model is subsequently directed towards an additional feed-forward layer, which ultimately results in the generation of probabilities for each distinct category. Finally, we utilize Equation 1 to get the final predictions.

$$argmax(f(x)) = x \in X \tag{1}$$

where $f(x)$ denotes the probabilities of the output layer, X denotes the number of classes and x is the highest probability index.

3 Experiments and Evaluations

3.1 Dataset Description

In the following subsections, we will overview the dataset for the promise verification task.

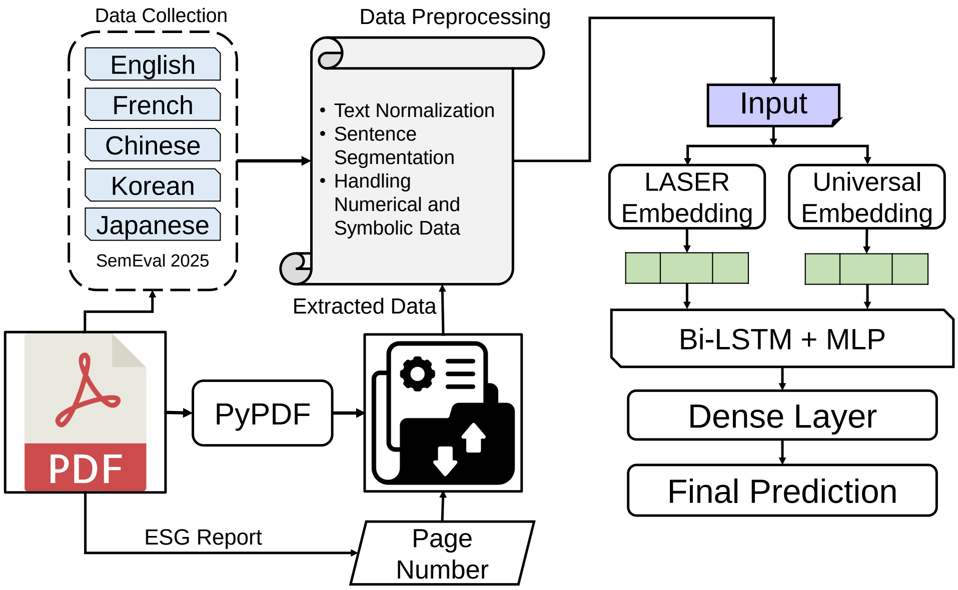

Figure 1: Proposed Framework.

3.1.1 Dataset Overview

SemEval 2025 Promise Verification Task (Chen et al., 2025) The ML-Promise dataset is a multilingual dataset designed to analyze and verify corporate promises made in ESG reports. These reports often contain commitments related to sustainability, ethical governance, and social responsibility. However, companies may exaggerate or misrepresent their claims (a practice known as greenwashing). The ML-Promise dataset aims to provide structured data to assess the credibility of such corporate commitments.

The dataset consists of 3,010 instances collected from ESG reports published in five different languages which are depicted in Figure 2.

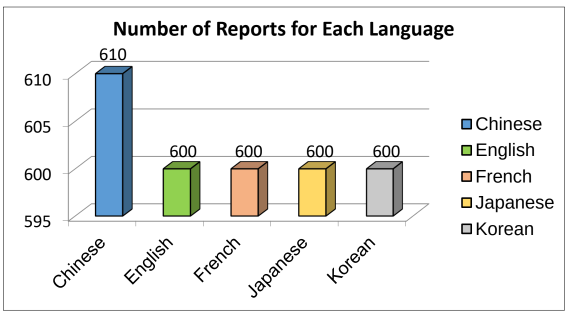

Figure 2: Data distribution among different languages.

3.1.2 Statistical Analysis of Dataset Labels

A quantitative analysis of the dataset reveals key trends in corporate ESG reporting among different countries and different industries.

Promise Identification Rates The proportion of statements identified as corporate promises varies significantly across languages, as shown in Figure 3.

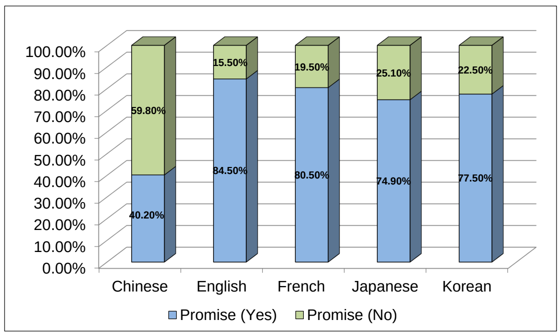

Figure 3: Promise identification rate among different languages.

English and French companies tend to make more explicit promises than Chinese firms. Notably, the Chinese dataset has the lowest promise rate at 40.2%, suggesting that Chinese reports may be more vague or general in nature, potentially lacking clear commitments or measurable objectives.

Supporting Evidence Availability The availability of supporting evidence varies significantly across languages, as depicted in Figure 4. English and Chinese companies rarely provide supporting evidence, with only 20.1% of statements backed by tangible proof. In contrast, French, Japanese, and Korean firms are more likely to include supporting documents, with Korean companies leading at 75.6%, followed by French (71.6%) and Japanese

(66.4%).

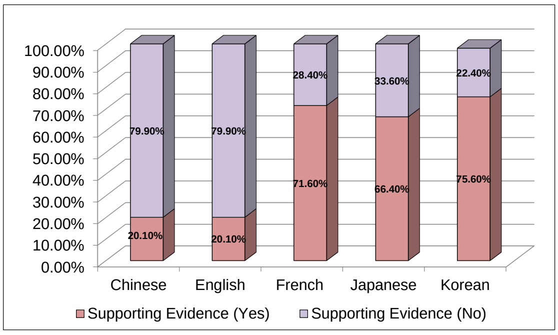

Figure 4: Evidence availability rate among different languages.

Clarity of Promise-Evidence Pair The clarity of the promise-evidence pair varies across languages, as shown in Figure 5. Korean corporate reports demonstrate the highest clarity, with 94.8% of statements being clearly supported and almost no misleading claims. In contrast, English and Japanese reports have a relatively higher rate of misleading claims, around 4%. While Chinese reports show a strong clarity rate (64.7%), they contain no misleading statements, whereas French reports exhibit a lower misleading rate of 1.5%, maintaining a balanced level of clarity.

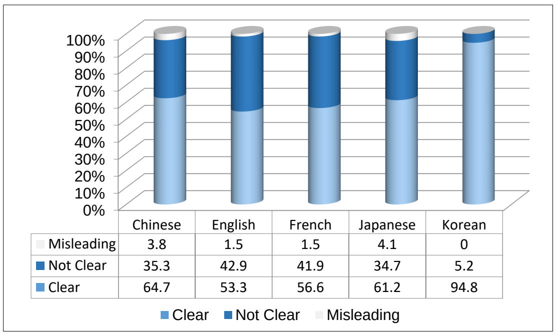

Figure 5: Clarity of promise-evidence rate among different languages.

Differences in Short-Term vs. Long-Term Promises The distribution of short-term and long-term promises varies across languages, as shown in Figure 6. Korean (45.5%) and Chinese (37.5%) firms make the most short-term commitments (<2 years), indicating a stronger focus on immediate actions. In contrast, English reports have the highest proportion of unclear timelines (75%), suggesting a tendency toward vague or long-term commitments without specific verification periods. French and Japanese firms demonstrate a more balanced distribution across different timeframes, while Korean reports contain the fewest long-term (>5 years) commitments. Overall, English and French firms

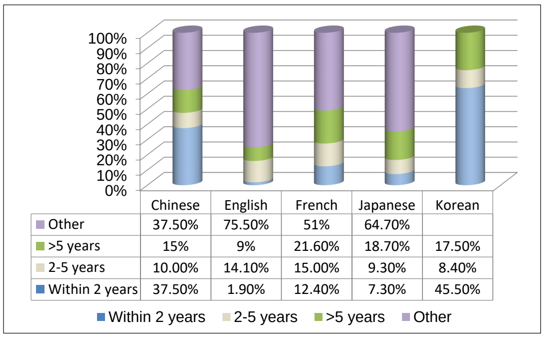

Figure 6: Promise verification timeline among different languages.

offer more specific commitments, whereas Chinese reports are less expected to include detailed commitments. Korean firms are the highest in both supporting evidence and clarity, whereas English and Japanese reports include more misleading information. Taiwanese and Korean firms focus on short-term action in their commitment timelines, whereas English firms have the largest proportion of unclear or unverifiable timelines.

3.2 Experimental Setup

For our system, we employ the feature fusion based neural network architecture. The configuration of the system is provided in Table 2.

Settings of the Proposed System
1. *Sentence Embedding*: LASER, USE
2. *Embedding Dimension*: 1024
3. *Optimizer*: Adam, AdamW
4. *Learning_rate*: 1e-7 to 7e-5
5. *Epochs*: 10 to 30
6. *Batch Size*: 16, 32, 64

Table 2: System settings.

3.3 Result Analysis

This section includes some experimental analysis to support our proposed Fused Bi-LSTM+MLP system. Before finalizing the methods architecture, we first conducted several tests based on the training data of the promise verification dataset. However, to estimate the performance of our promise verification system, we utilized the F1-score as a primary evaluation measure. The results of our proposed model's performance on two tasks are displayed in Table 3 and the detailed experimental baseline score are shown in Figure 7.

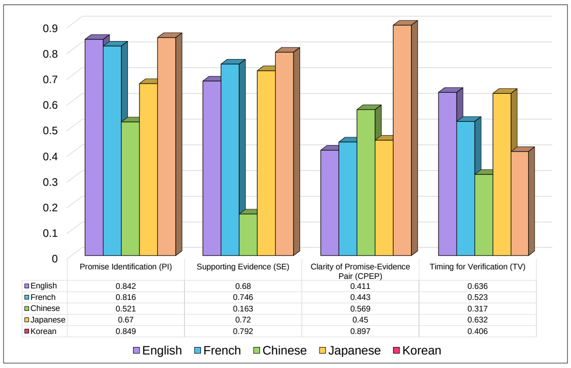

Figure 7: Experimental results on different languages.

We have shown the performance of our experimented models on the SemEval 2025 promise verification dataset. The Figure 7 demonstrates the results on four different categories in five different languages.

While doing experimental analysis, it is evident that the scoring pattern among different languages, such as English, Chinese, Korean, French, and Japanese also matches the data analysis pattern for the promise identification and evidence verification task. In the dataset overview section, we noticed that English and French language ESG reports had a greater ratio of promise which was also reflected in the scoring. Furthermore, our system shows competitive performance for each of the languages except Korean language.

4 Discussion

To maintain clarity and focus, supplementary analyses—including topic modeling to identify key themes in corporate promises, sentiment analysis aimed at detecting greenwashing in ESG reports, and word frequency analysis for promise detection—are provided in Appendix 5. In this section, we focus on core evaluation aspects, including explainability analysis using SHAP and a detailed error analysis of the evidence identification task, to further interpret model performance and understand its limitations.

4.1 Explainability Analysis using SHAP

To better understand the decision-making process of the model, we apply SHAP (SHapley Additive exPlanations) (Lundberg, 2017) to describe feature contributions. SHAP dissects the effect of individual words or phrases on the prediction of the model, making it transparent to decisions. The bar chart in Figure 8 illustrates the top 20 most important words influencing the evidence prediction model, ranked by their mean absolute SHAP values. The higher the SHAP value, the greater the word's impact on the model's decision-making process. Words like "risk", "employees", "group", and "2022" have the strongest influence, suggesting they play a crucial role in determining whether a piece of text contains supporting evidence. Other significant terms, such as "burberry", "emissions", "ethics", and "supply," indicate the model's focus on themes related to corporate responsibility, financial matters, and ethical concerns. This visualization helps in understanding which words contribute most to the model's predictions, providing insights into its decision logic.

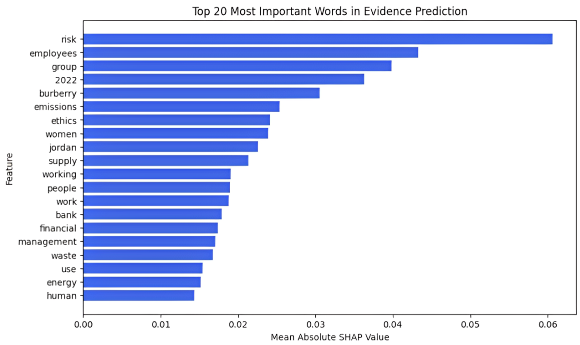

Figure 8: Word impact toward evidence prediction.

4.2 Error Analysis

4.2.1 Error Analysis for Promise Identificatin Task

The Table A6 presents an error analysis for the promise identification task, comparing sample texts with their predicted and true labels. The second prediction is correct, as the model accurately identified the commitment to reducing energy intensity and decarbonizing electricity usage, making it a valid promise. The presence of explicit terms like "We will reduce our energy intensity" strengthens the model's decision.

However, the first prediction is incorrect, as the model classified a disclaimer about forward-looking statements as a promise. The statement contains phrases like "we cannot guarantee their realization" and "undertakes no duty to update such information except as required under applicable law", which indicate caution rather than a commitment. The misclassification suggests that the model might have mistaken forward-looking language for

Language	PromiseVerification	EvidenceVerification
English	0.8113	0.7114
Chinese	0.6360	0.6973
French	0.7225	0.5850
Japanese	0.9300	0.5650
Korean	0.2340	0.2160

Table 3: Task scores for different languages.

a definitive obligation, leading to an overestimation.

This error highlights a key challenge in promise verification—distinguishing between actual commitments and general disclaimers or legal protections. Future improvements should focus on context-aware training, refining the model's ability to differentiate between legally binding statements and non-committal language, thereby improving classification accuracy.

4.2.2 Error Analysis for Evidence Idenfication Task

The Table A7 presents an error analysis for the evidence identification task, where the model determines whether a given text contains supporting evidence for a claim or commitment. Among the three examples, two were correctly classified, while one was misclassified. The first sample, discussing collaboration and environmental reporting, was correctly classified as evidence since it explicitly mentions engagement with stakeholders, regulators, and Indigenous communities. Additionally, it provides quantifiable data, stating that in 2022, Canada Nickel had zero instances of environmental non-compliance, fines, or violations, making it a strong supporting evidence statement. The second sample, detailing the risk assessment process, was also correctly classified as not containing evidence since it describes a procedure rather than providing specific data or verifiable reports. However, the third sample, outlining the responsibilities of the ESG Committee, was misclassified as containing evidence when it actually does not. While the text discusses accountability in areas such as health and safety, climate change, and social matters, it does not present concrete proof, such as compliance data or measurable outcomes. The misclassification suggests that the model may be over-relying on governance-related terms rather than distinguishing between general policy descriptions and verifiable evidence. Future improvements should focus on refining the model's ability to differentiate between commitments, procedural descriptions, and factual evidence, ensuring more accurate classification.

5 Conclusion and Future Work

In this paper, we introduced a feature fusion based neural architecture framework for corporate promise verification. Among them, for the feature fusion, we levaraged the feature embedding coming from LASER and Universal Sentence Encoder. Then, the Bi-LSTM with MLP framework trained with after the feature integration to get the predictions for promise and evidence identification subtasks.

Our future plan amalgamates working with the clarity of promise and verification timeline subtasks, as well as focus on the better model by incorporating multilingual Retrieval-Augmented Generation (RAG) with Large Language Model (LLM) based framework for the corporate promise verification task.

References

Mikel Artetxe and Holger Schwenk. 2019. Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond. *Transactions of the association for computational linguistics*, 7:597–610.

Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, et al. 2018. Universal sentence encoder. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 169–174.

Chung-Chi Chen, Yohei Seki, Hakusen Shu, Anais Lhuissier, Juyeon Kang, Hanwool Lee, Min-Yuh Day, and Hiroya Takamura. 2025. Semeval-2025 task 6: Multinational, multilingual, multi-industry promise verification. In *Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval-2025)*, pages 2426–2436, Vienna, Austria. Association for Computational Linguistics.

Jianfeng Deng, Lianglun Cheng, and Zhuowei Wang. 2021. Attention-based bilstm fused cnn with gating mechanism model for chinese long text classification. *Computer Speech & Language*, 68:101182.

Nina Gorovaia and Michalis Makrominas. 2024. Identifying greenwashing in corporate-social responsibility reports using natural-language processing. *European Financial Management*.

Marcelo Gutierrez-Bustamante and Leonardo Espinosa-Leal. 2022. Natural language processing methods for scoring sustainability reports—a study of nordic listed companies. *Sustainability*, 14(15):9165.

Jing Li, Peizhang Wang, Lu Jia, Run Mao, Qian Li, Yongle He, Yi Sun, and Pinwang Zhao. 2023. Design and implementation of an automated pdf drawing statistics tool based on python. In *Sixth International Conference on Computer Information Science and Application Technology (CISAT 2023)*, volume 12800, pages 1686–1690. SPIE.

Gang Liu and Jiabao Guo. 2019. Bidirectional lstm with attention mechanism and convolutional layer for text classification. *Neurocomputing*, 337:325–338.

Scott Lundberg. 2017. A unified approach to interpreting model predictions. *arXiv preprint arXiv:1705.07874*.

Holger Schwenk and Matthijs Douze. 2017. Learning joint multilingual sentence representations with neural machine translation. *arXiv preprint arXiv:1704.04154*.

Holger Schwenk and Xian Li. 2018. A corpus for multilingual document classification in eight languages. *arXiv preprint arXiv:1805.09821*.

Yunxiang Zhang and Zhuyi Rao. 2020. n-bilstm: Bilstm with n-gram features for text classification. In *2020 IEEE 5th information technology and mechatronics engineering conference (ITOEC)*, pages 1056–1059. IEEE.

Yi Zou, Mengying Shi, Zhongjie Chen, Zhu Deng, ZongXiong Lei, Zihan Zeng, Shiming Yang, Hongxiang Tong, Lei Xiao, and Wenwen Zhou. 2025. Esgreveal: An llm-based approach for extracting structured data from esg reports. *Journal of Cleaner Production*, 489:144572.

A Supplementary Analyses

A.1 Topic Modeling: Identifying Key Themes in Corporate Promises

Corporate ESG reports encompass a range of diverse themes, ranging from climate change to social responsibility. To reveal the dominant subjects of corporate commitments, **Latent Dirichlet Allocation (LDA)** is applied, which identifies key themes underlying the text. Table A4 presents the five most salient topics from the English dataset. Each topic comprises key terms that provide a hint of the primary areas of emphasis of corporate commitments, enabling a systematic interpretation of ESG priorities across industries.

Topic	Top Keywords
Topic 1	Equality, fashion, commitments, liquidity, Arabi, 500, LVMH, Kering, SMEs, care
Topic 2	Energy, term, low, products, circular, supply, fashion, waste, carbon, emissions
Topic 3	Staff, disability, men, like, LVMH, heritage, European, spill, pragmatic, collaborate
Topic 4	Middle, emergency, East, dialogue, year, trade, create, bank, best, Jordan
Topic 5	Impact, environmental, safety, business, bank, management, information, employees, group, risk

Table A4: Identified topics and keywords from Topic Modeling.

- **Topic 1 (Equality and Corporate Commitments):** This topic is centered around fashion industry commitments, corporate liquidity, and equality. The presence of terms like LVMH, Kering (major fashion companies), and SMEs suggests a focus on sustainability and inclusivity in the fashion sector.
- **Topic 2 (Energy and Sustainability):** The words carbon, waste, emissions, and circular supply indicate that this topic deals with environmental sustainability, particularly in reducing carbon footprints, managing waste, and promoting circular economies.
- **Topic 3 (Workforce and Inclusion):** This topic relates to diversity, disability inclusion, and employee well-being, as seen through words like staff, disability, heritage, and collaborate. The presence of LVMH again suggests a connection to the fashion industry's employment policies.
- **Topic 4 (Crisis Response and Economic Stability):** The keywords Middle East, emer-

gency, trade, and bank suggest a focus on corporate responses to crises, possibly related to humanitarian aid, financial support, or economic stability in specific regions.

- **Topic 5 (Risk and Environmental Impact):** This theme revolves around corporate risk management, environmental responsibility, and workplace safety, as indicated by words like impact, safety, business, employees, and risk management.

Thus, the LDA key ESG theme analysis reveals that corporate commitments are centered around sustainability, social responsibility, crisis management, and labor inclusion. Certain industries, like fashion and finance, seem to be prominent in these commitments. Further, though some of the topics prioritize long-term sustainability objectives, others prioritize short-term economic and social issues.

A.2 Sentiment Analysis: Detecting Greenwashing in ESG Reports

The analysis seeks to identify greenwashing behavior in ESG reports through the sentiment of corporate commitments. Greenwashing is said to happen when firms utilize excessively positive language to present a false picture of their sustainability initiatives without the presence of considerable evidence. This was evaluated by using the VADER (Valence Aware Dictionary and Sentiment Reasoner) tool to measure the sentiment polarity of statements. Every statement was given a sentiment score from -1 (most negative) to +1 (most positive). Statements scoring above 0.8 on the sentiment scale and having no evidence to support them were considered possible instances of greenwashing. The sentiment distribution, as the Figure 9 shows, indicates that the majority of statements are concentrated towards the positive side of the scale, specifically towards a sentiment score of 1. This shows that firms tend to use very optimistic language in their ESG commitments. Yet, the existence of some neutral and negative statements indicates variability in the sentiment tone. The clustering of statements with high positive sentiment scores identifies the potential risk of greenwashing, particularly if such statements are not supported by concrete evidence, calling for a more critical analysis of firms' ESG commitments.

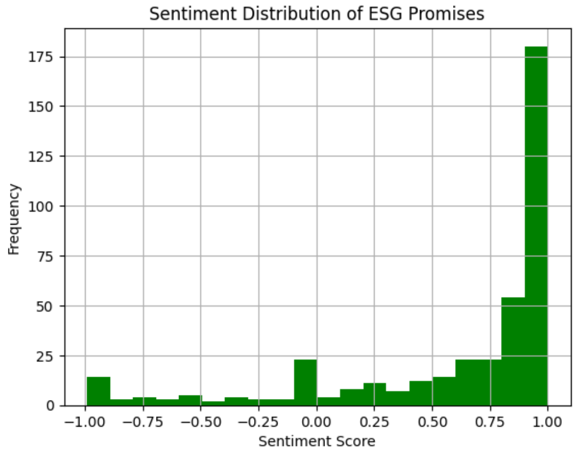

Figure 9: Sentiment analysis in detecting greenwashing in ESG reports.

A.3 Word Frequency Analysis for Promise Detection

The word frequency Table A5 for the promise detection english dataset shows the main differences between promise and non-promise statements.

In promise statements, the most common words are risk (231), employee (190), impact (130), business (128), information (125), emission (115), support (108), and program (104). These words indicate a strong focus on commitment, responsibility, and organized actions, especially in risk management, environmental impact, and business programs. Words such as training (89), community (89), and opportunity (83) also point to a commitment to employee development and corporate social responsibility (CSR).

In contrast, in non-promise statements, although risk (137) remains the most frequent word, other high-frequency words such as management (64), governance (24), compliance (22), report (19), and policy (32) suggest a more regulatory or descriptive rather than commitment-oriented tone. In addition, the occurrence of words such as privacy (16), client (15), and chief (15) suggests a focus on general policy and leadership arrangements rather than specific action. The occurrence of impact (17) in both groups suggests that impact assessment is addressed in both promise-driven and neutral contexts, but with different implications.

This analysis points out that promise language often involves active words like support, program, training, and action, whereas non-promise language is more policy-oriented or descriptive and focuses on management, compliance, and gover-

Promise	**Frequency**	**Non-Promise**	**Frequency**
Risk	231	Risk	137
Employee	190	Management	64
Impact	130	Information	40
Business	128	Policy	32
Information	125	Data	30
Emission	115	Business	26
Support	108	Governance	24
Program	104	ESG	23
Work	93	Compliance	22
Community	89	Employee	20
Training	89	Report	19
Management	88	Group's	19
Environmental	87	Conduct	18
Service	85	Climate	18
Policy	84	Impact	17
Opportunity	83	Ensures	17
Year	80	Privacy	16
Product	79	Service	15
Action	77	Chief	15
People	76	Client	15

Table A5: Most Frequent Words in Promise vs. Non-Promise Statements.

nance.

A.4 Additional Tables from Error Analysis

Sample Text	Predicted Label	True Label
Certain statements in this report are forward-looking statements that involve a number of risks and uncertainties that could cause actual results to diï¬€er materially. These statements are made under the â€œSafe Harborâ€ provisions ofthe U.S. Private Securities Litigation Reform Act of 1995. Forward-looking statements may be marked by such terms as ...	Yes	Yes
We will reduce our energy intensity by leveraging our expertise and strength in product technologies, manufacturing process know-how, and energy savings while we continue to grow our business. On the energy supply side, the paths we follow to decarbonize the electricity we use are, in order of priority, installing distributed solar on the rooftops of our factories, signing renewable power purchase agreements (PPAs), and purchasing green electricity from the spot market. Most of our manufacturing facilities are...	Yes	Yes

Table A6: Error analysis for promise identification task.

Sample Text	Predicted Label	True Label
Collaboration We work with stakeholders, regulators, and Indigenous communities to understand and address concerns, obtain local expertise on environmental conditions and land and resource use, and discuss baseline/monitoring programs, potential impacts, and proposed mitigation measures. These eï¬€orts are supported ...	Yes	Yes
Risk Assessment Prior to conducting any activities that may have an impact on the environment, a risk assessment compliant with our Responsible Exploration Policy is conducted by our environmental team to determine ...	No	No
ENVIRONMENTAL, SOCIAL AND GOVERNANCE (ESG) COMMITTEE Oversees fulïllment of responsibilities relating to health and safety, Indigenous relations, climate change, and environmental and social matters, and advocates for integration of sustainability into Board governance ...	Yes	No

Table A7: Error analysis for evidence identification task.