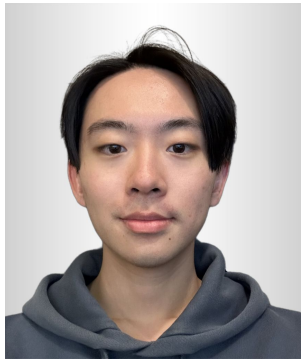
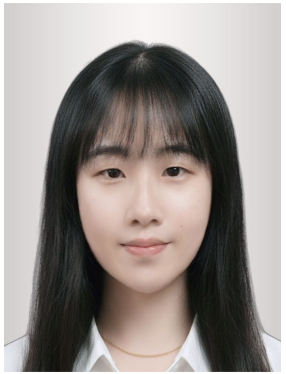
SentiPromiseESG: Sentiment Analysis of Sustainability Promises Across Industries




Wei-Chen Huang[1] Hsin-Ting Lu[2] Wen-Ze Chen[2] Yu-Han Huang[3] Prof. Min-Yuh Day[2]*

Department of Statistics[1], Graduate Institute of Information Management[2],
Department of Accounting[3], National Taipei University, New Taipei City, Taiwan
myday@gm.ntpu.edu.tw*

Table of Contents

Introduction

Research Background

- Sustainability reporting has become a key channel for firms to communicate their **environmental**, **social**, and **governance** (ESG) performance and manage stakeholder expectations (KPMG, 2022).

- Domain-specific models such as **FinBERT** (Araci, 2019) have been applied to ESG topic classification and sentiment detection.

- Sentiment-based indicators have been used to measure **reporting quality** and **environmental messaging strategies** (Barbeito-Caamaño & Chalmeta, 2020).

Research Motivation

- **Greenwashing Concerns**: Firms often use **symbolic commitments**, **optimistic tone**, or **promotional language** to project responsibility without implementing corresponding operational changes (Delmas & Burbano, 2011).

- **Lack of Verifiability**: Unsubstantiated sustainability claims undermine their credibility and complicate external evaluation (Lublóy et al., 2025).

- **Technical Limitations**: Model performance can vary substantially, and LLMs may be misled by **repeated keywords** or **loosely related text spans** (Barbeito-Caamaño & Chalmeta, 2020).

Research Objective

Main Objective:

- To propose a sentiment-aware analytical framework for identifying and evaluating sustainability promises.

Research Question:

- Is sentiment associated with the presence of sustainability promises?
- Is sentiment associated with the provision of verifiable evidence?
- Is sentiment associated with the distribution of ESG promise types (social vs. environmental, governance vs. environmental)?
- Is the sentiment of evidence text associated with the clarity of the evidence?

Literature Review

The Rise of ESG Disclosure and Greenwashing Practices

- ESG reports have gradually become a primary source of information for stakeholders to assess corporate sustainability performance (Kim et al., 2023).

- This discretion may lead to **selective presentation of favorable information** and **the downplaying of negative aspects**, which creates information asymmetry (Xu et al., 2025).

Large Language Model Driven ESG Text Analysis and its Challenges

- **Application of LLMs (Birti et al., 2025):**
  - **Sentiment classification**
  - **Identification of corporate commitments**
  - **Detection of greenwashing practices**

- **Cross-linguistic Challenges:**
  - Cross linguistic analysis is difficult because ESG reports across languages contain diverse and extensive content, which makes automated commitment verification highly challenging (Turk et al., 2025).

Applications of NLP in Commitment Identification and Sentiment Analysis

- **Existing NLP Approaches (Schimanski et al., 2024):**
  - **Keyword based rule retrieval**
  - **Fine tuning BERT models**
  - **Using GPT with few shot prompting for information extraction**

- **Challenges - Contextual Ambiguity (Sun et al., 2024):**
  - Because of repeated or ambiguous keywords, algorithms often struggle to determine which matched segment is most relevant when the same term appears in different contexts within a report (Sun et al., 2024).

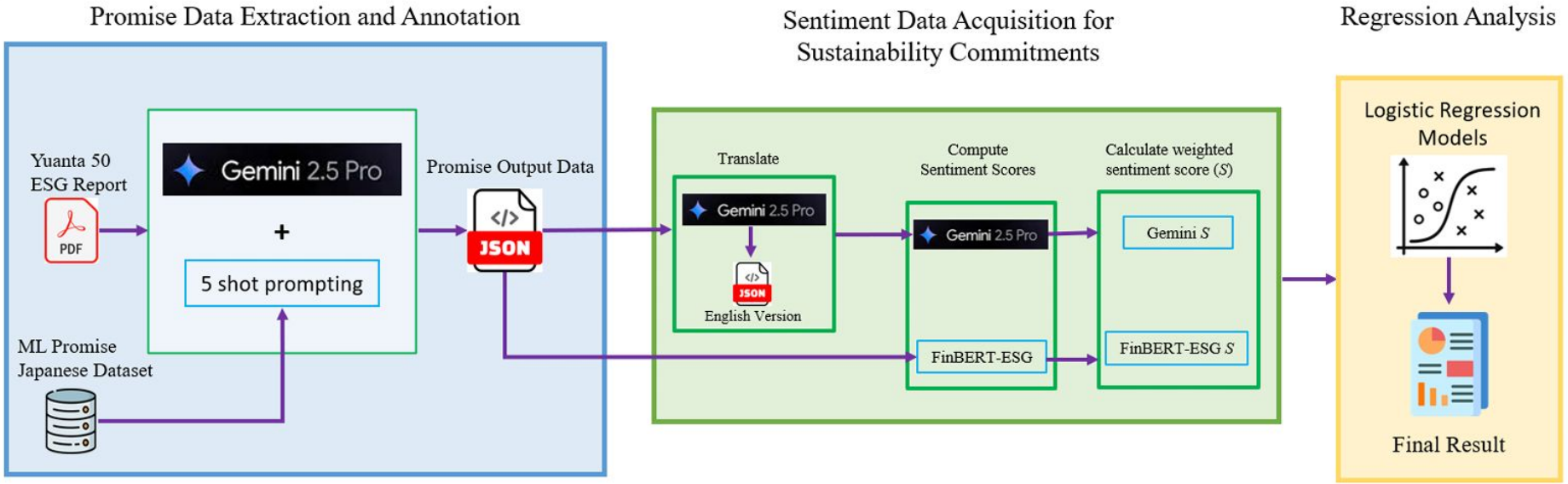
Research Gap

- Voluntary disclosure leads to **selective reporting** and **vague commitments**.
- General models struggle with fine-grained semantics and complex texts.
- Current NLP lacks context understanding and evidence verification.

Research Method

System Architecture

- Overall Workflow of the SentiPromiseESG Analysis Framework



weighted sentiment score: $S = 1 \cdot p_{pos} + 0 \cdot p_{neu} - 1 \cdot p_{neg}$

PromiseEval Dataset Label

Promise Status

- Yes
- No

Evidence Status

- Yes
- No

Evidence Quality

- Clear
- Not Clear
- Misleading
- N/A

Verification Timeline

- Already
- Within 2 years
- Between 2–5 years
- More than 5 years
- N/A

Ref.: Seki, Y., Shu, H., Lhuissier, A., Lee, H., Kang, J., Day, M.-Y., & Chen, C.-C. (2024). ML-Promise: A Multilingual Dataset for Corporate Promise Verification. arXiv preprint arXiv:2411.04473.

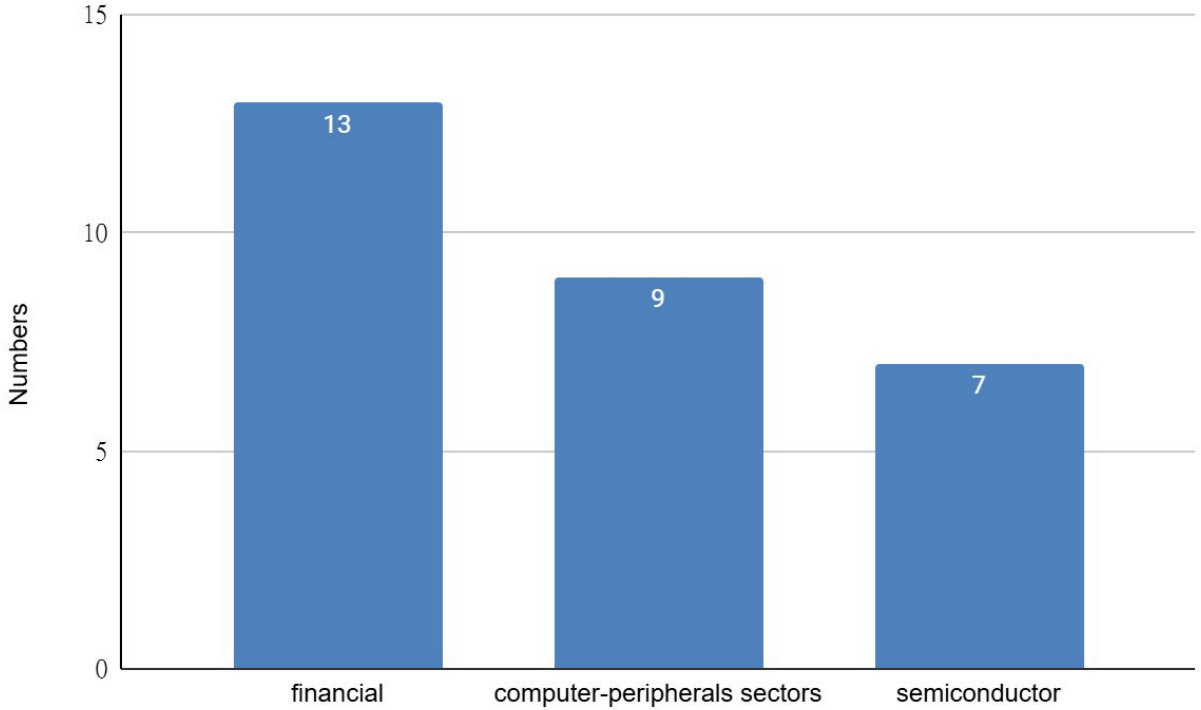
Dataset Example

data	The UMC Group has established a stakeholder engagement mechanism to identify ...
URL & page_number	https://www.umc.com/upload/media/07_Sustainability/72_Reports_and_Results… / 19
promise_status	Yes
promise_string	Understand the reasonable expectations and needs of stakeholders and appropriately respond to ...
verification_timeline	already
evidence_status	Yes
evidence_string	The United Microelectronics Corporation (UMC) Group has established a stakeholder...
evidence_quality	Not Clear
ESG Type	G
Data Sentiment	"negative": 0.0053, "neutral": 0.6799, "positive": 0.2967
Promise Sentiment	"negative": 0.0149, "neutral": 0.894, "positive": 0.0822
Evidence Sentiment	"negative": 0.0063, "neutral": 0.0066, "positive": 0.5481

Dataset Overview

Sample Size : 15345

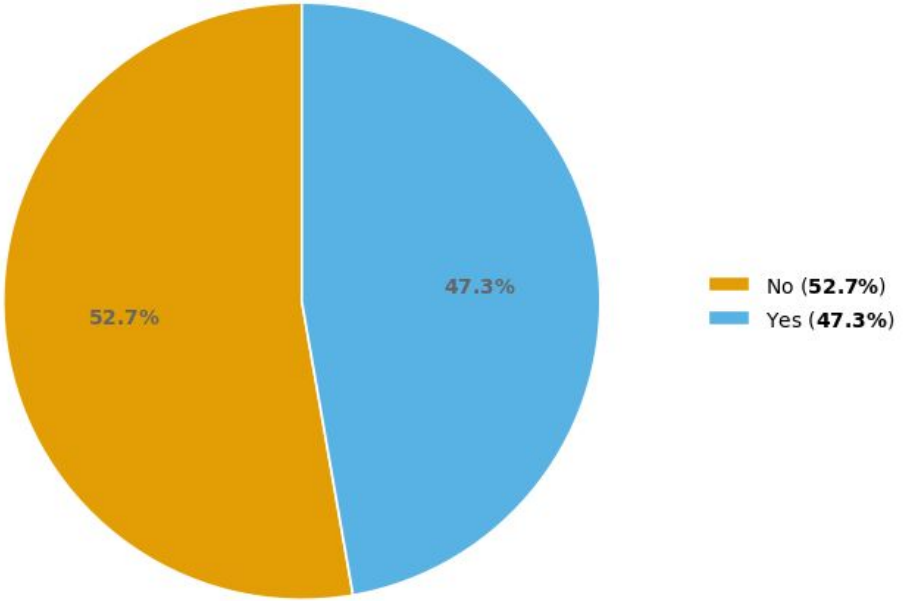
- ESG reports from 29 large Taiwanese firms in the semiconductor,financial, and computer-peripherals sectors



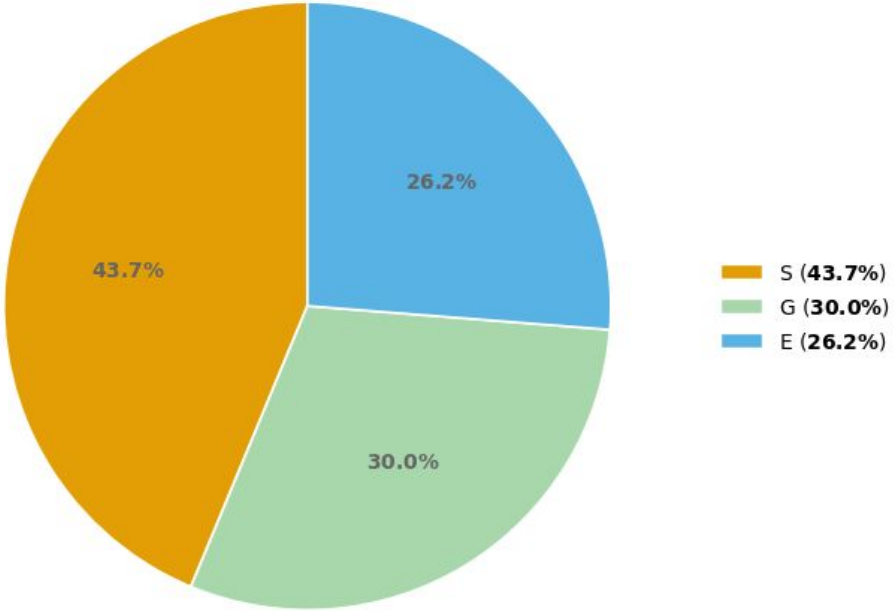
Dataset Overview

- Label distribution of the **SentiPromiseESG Dataset**

Sample Size : 15345



Promise Status

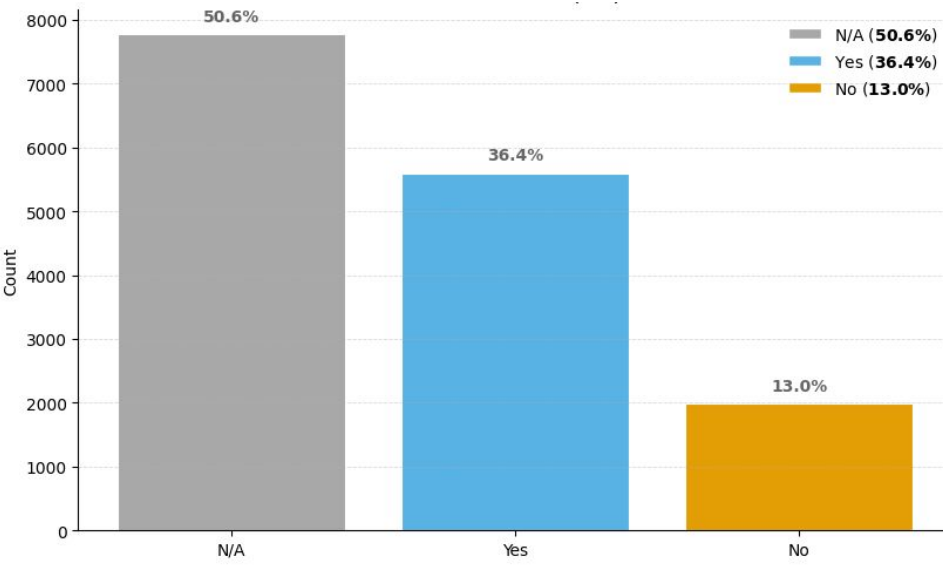


ESG Type

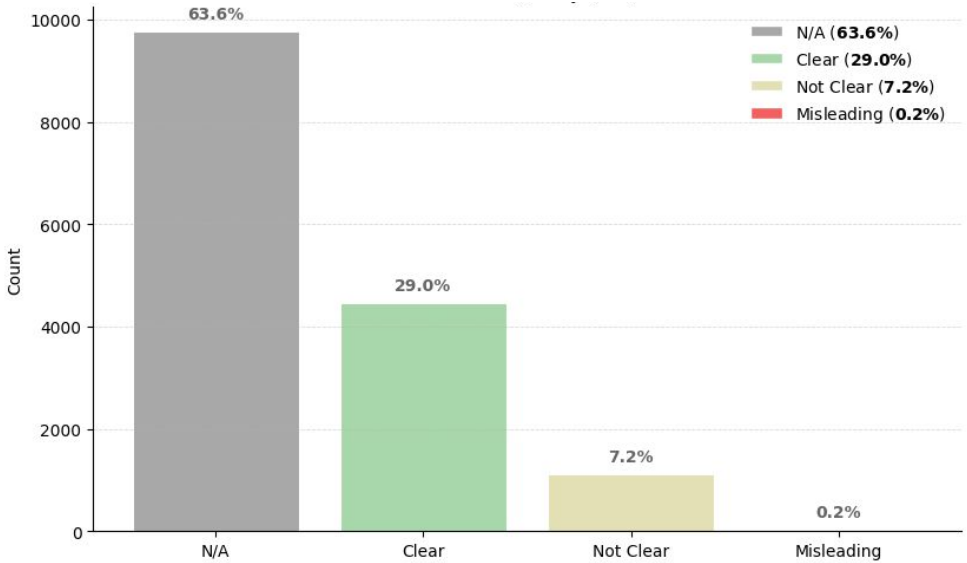
Dataset Overview

- Label distribution of the **SentiPromiseESG Dataset**

Sample Size : 15345



Evidence Status

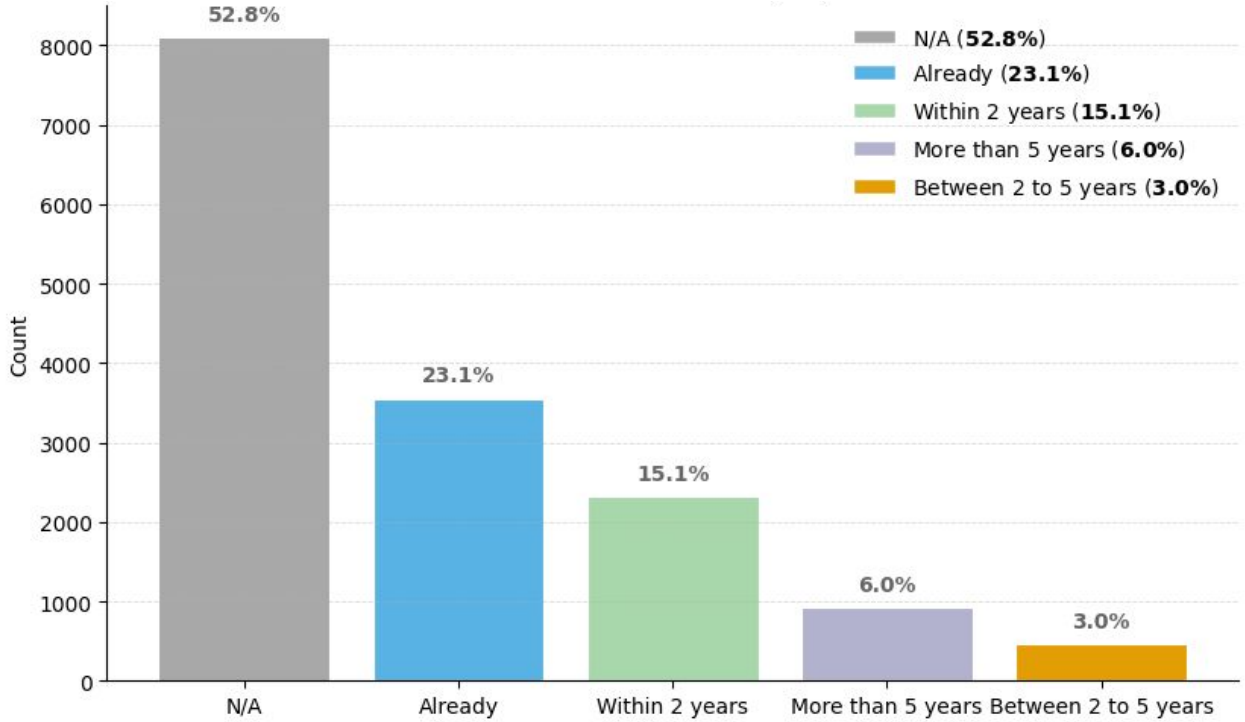


Evidence Quality

Dataset Overview

- Label distribution of the **SentiPromiseESG Dataset**

Sample Size : 15345



Experiment Results and Analysis

Results and Analysis

- Regression Analysis Results for the **SentiPromiseESG** Dataset (Gemini 2.5 Pro)

Outcome (Model)	Predictor	β	OR	p	Sig.
Promise	DSW	0.9767	2.656	<.001	***
Evidence Status	DSW	-0.4667	0.627	< .001	***
ESG Type (S vs E)	DSW	1.4258	4.161	<.001	***
ESG Type (G vs E)	DSW	-1.4917	0.225	<.001	***
Evidence Quality	ESW	0.0199	1.020	.856	n.s.

Note1. *** $p < .001$, ** $p < .01$, * $p < .05$; n.s. = not significant.
Note2. DSW = data_sentiment_weighted; ESW = evidence_sentiment_weighted.

Results and Analysis

- Cross-Industry Regression Analysis Results for the **SentiPromiseESG** Dataset (Gemini 2.5 Pro)

Outcome (Model)	Industry	β	OR	p	Sig.
Promise	Semiconductor	0.9130	2.492	<.001	***
	Financial	0.8997	2.459	<.001	***
	Computer Periph.	1.0974	2.996	<.001	***
Evidence Status	Semiconductor	-0.4667	0.627	.0044	**
	Financial	-0.5140	0.598	.0012	**
	Computer Periph.	-0.5140	0.831	.2974	n.s.
ESG Type (S vs E)	Semiconductor	1.4474	4.252	<.001	***
	Financial	1.7474	5.740	<.001	***
	Computer Periph.	0.9635	2.621	<.001	***
ESG Type (G vs E)	Semiconductor	0.290	0.290	<.001	***
	Financial	-1.7016	0.182	<.001	***
	Computer Periph.	-1.2228	0.294	<.001	***
Evidence Quality	Semiconductor	0.0666	1.069	.7862	n.s.
	Financial	-0.2255	0.798	.2494	n.s.
	Computer Periph.	0.3082	1.361	.1553	n.s.

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; n.s. = not significant.

Results and Analysis

- Regression Analysis Results for the **SentiPromiseESG** Dataset (FinBERT-ESG)

Outcome (Model)	Predictor	β	OR	p	Sig.
Promise	DSW	0.3626	1.437	<.001	***
Evidence Status	DSW	-0.0341	0.967	.464	n.s.
ESG Type (S vs E)	DSW	1.5173	4.560	<.001	***
ESG Type (G vs E)	DSW	-0.5849	0.557	<.001	***
Evidence Quality	ESW	0.1399	1.150	.014	*

Note1. *** $p < .001$, ** $p < .01$, * $p < .05$; n.s. = not significant.
Note2. DSW = data_sentiment_weighted; ESW = evidence_sentiment_weighted.

Results and Analysis

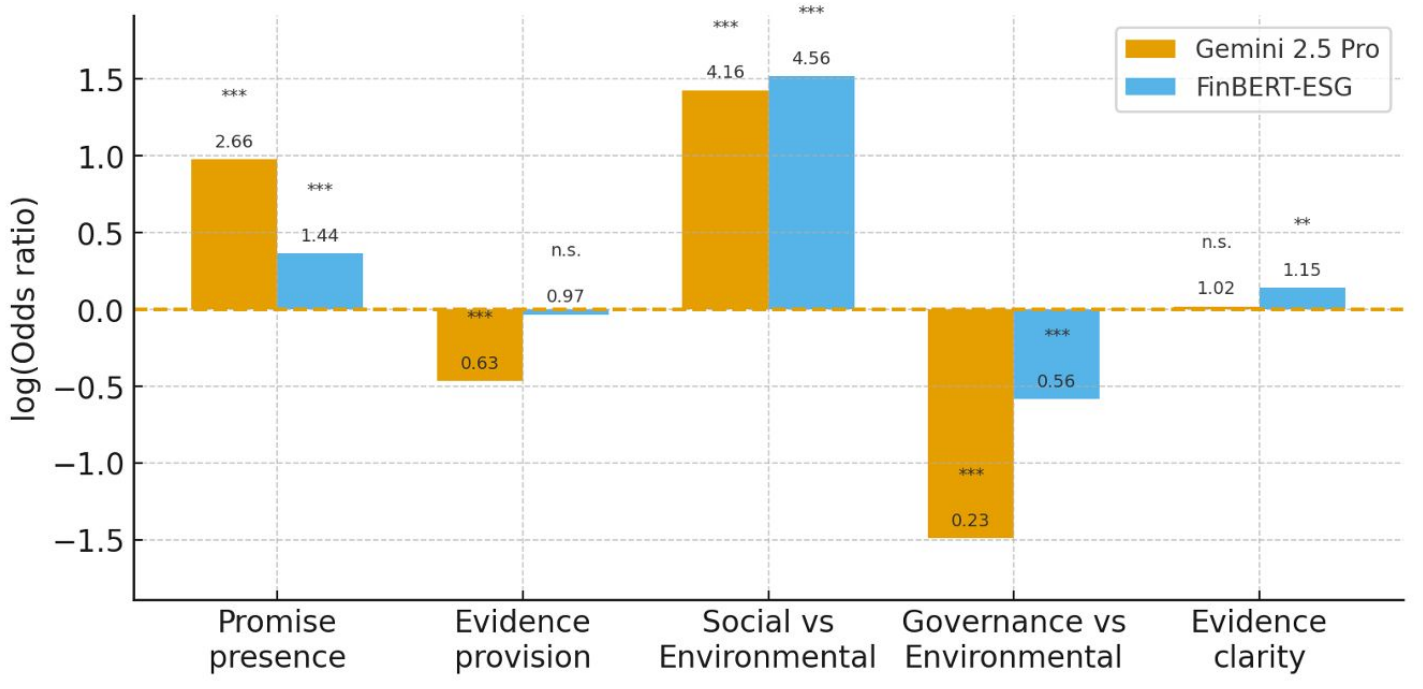
- Cross-Industry Regression Analysis Results for the **SentiPromiseESG** Dataset (FinBERT-ESG)

Outcome (Model)	Industry	β	OR	p	Sig.
Promise	Semiconductor	0.2247	1.252	<.001	***
	Financial	0.4116	1.509	<.001	***
	Computer Periph.	0.4430	1.557	<.001	***
Evidence Status	Semiconductor	0.1431	1.154	.0857	n.s.
	Financial	-0.1297	0.878	.0902	n.s.
	Computer Periph.	0.0539	1.055	.5844	n.s.
ESG Type (S vs E)	Semiconductor	1.9210	6.828	<.001	***
	Financial	1.2526	3.499	<.001	***
	Computer Periph.	1.5610	4.764	<.001	***
ESG Type (G vs E)	Semiconductor	-0.7512	0.472	<.001	***
	Financial	-0.5037	0.604	<.001	***
	Computer Periph.	-0.4071	0.666	<.001	***
Evidence Quality	Semiconductor	-0.0814	0.922	.5021	n.s.
	Financial	0.1859	1.204	.0385	*
	Computer Periph.	0.1690	1.184	.1292	n.s.

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; n.s. = not significant.

Results and Analysis

- Sentiment effects on all outcomes for the **SentiPromiseESG** Dataset



Note. *** p < .001, ** p < .01, * p < .05; n.s. = not significant.

Conclusion

Major Findings

- Integrates LLM labeling with dual sentiment models (Gemini 2.5 Pro, FinBERT-ESG) to test how tone relates to ESG promise behavior.
- Sentiment is consistently positive and significant for promise presence across models and industries.
- Sentiment–evidence provision effects are not robust across models/industries.
- Sentiment shifts ESG-type distribution: more positive tone in Social; Governance is more neutral/less positive.
- Evidence sentiment has limited impact on evidence clarity (only modest FinBERT-ESG effects).

Research Contributions

- Builds a scalable, reproducible pipeline from annotation to regression with two sentiment models.
- Releases **SentiPromiseESG (15,345 labeled sentences)** for greenwashing and ESG text research.

Managerial Implications

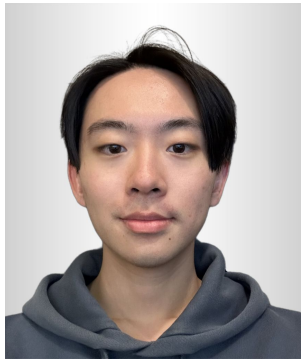
- The findings suggest that positive tone reflects communication style rather than the underlying credibility of a commitment.
- Analysis shows that sentiment clearly shapes ESG narratives, with firms using more positive language for social-related commitments compared to governance-related statements.

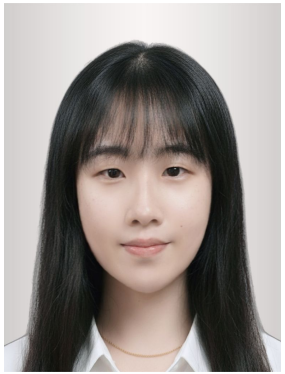
AI Cup：VeriPromiseESG 2026 ESG Promise VerificationCompetition

AI Cup 2026 - ESG Promise Verification Competition

Q &A

SentiPromiseESG: Sentiment Analysis of Sustainability Promises Across Industries




Wei-Chen Huang[1] Hsin-Ting Lu[2] Wen-Ze Chen[2] Yu-Han Huang[3] Prof. Min-Yuh Day[2]*

Department of Statistics[1], Graduate Institute of Information Management[2],
Department of Accounting[3], National Taipei University, New Taipei City, Taiwan
myday@gm.ntpu.edu.tw*