Shared Task: Detecting greenwashing signals through a comparison of ESG reports and public media

Background

In the past two years, greenwashing has been at the top of the social agenda of companies, and there is no sign of relinquishing its position. Greenwashing occurs when companies over-report positive data about their sustainability efforts while downplaying the negative impacts of their operations. The reason for this is the subjective nature of most ESG (Environmental, Social, and Governance) information. Public companies are required to report on their sustainability efforts. However, when investors and ESG rating providers focus on self-reported data (ESG reports and marketing communications), they often receive a distorted and overly positive picture of the company's ESG performance. Because of this information asymmetry, even investors who want to invest sustainably can be misled in their investment decisions.

Data from external media can help address this problem. To a large extent, third-party content providers have no interest in promoting a particular company's sustainability efforts - so by considering data from a variety of media outlets, we can form a more objective and even critical picture of a company.

Goals

In this task, we want to:

- Better understand the nature of greenwashing through large-scale text analysis
- Investigate whether sentiment analysis can shed light into greenwashing risks
- Analyze whether specific ESG topics (modeled in terms of the Sustainable Development Goals) are more prone to greenwashing

Organizers

The task is organized as a collaborative task between industry and academia by:

[Dr. Janna Lipenkova](), CEO, Equintel GmbH, Germany
[Susie Xi Rao](), Ph.D. candidate / Researcher at ETH Zurich
[Dr. Guang Lu](), Lecturer for Data Science, Lucerne University of Applied Sciences and Arts

Task

To conduct this task, we provide participants with a comprehensive ESG dataset covering company ESG reports as well as public media targeting a wide range of stakeholders (including investors, NGOs, regulators, society, etc.). Additionally, we provide a dataset with the descriptions of the 25 Sustainable Development Goals. The task is to develop approaches to identify gaps and inconsistencies between company-reported data and "external" data that may indicate greenwashing. This can be done at the level of ESG sentiment and/or the SDGs.

This is an application-oriented task. While there are no specific requirements for the NLP algorithms to be used, we suggest to focus on the following three aspects:

- Better understand the nature of greenwashing and try to quantify its degree
- Prototype NLP approaches to detect greenwashing using public documents reflecting different stakeholders
- Visualize potential indicators of greenwashing as well as the reliability of these indicators in a credible and verifiable way

Participants

Bachelor, Master and PhD students at Universities (of Applied Sciences) as well as NLP engineers and Data Scientists from industry.

Dataset description

The dataset contains ~11.000 English-language ESG documents for DAX companies. It contains both company reports and third-party data, which is distinguished by the "internal" field. The list of fields is as follows::

- symbol: stock symbol of the company
- company: company name
- date: publication date of document
- title: document title
- content: document content
- datatype: document type
- internal: is this a report by company (1) or a third-party document (0)
- domain (optional): Web domain where the document was published
- url (optional): URL where the document can be accessed
- esg_topics (optional): ESG topics extracted from the data using our internal NLP

The data can be loaded into a dataframe using the following Python snippet:

```
import pandas as pd

df = pd.read_csv('esg_documents_for_dax_companies.csv', delimiter = '|',
index_col = 0)
```

Environment requirement

For the use of LLMs, we recommend using computing resources with access to GPU.

Description of the 5 stages

Stage 1: Exploratory data analysis, preprocessing and cleaning

In this step, participants get a first overview of the dataset and learn to prepare it for the subsequent NLP analyses. This can involve the following statistical analyses:

- Average length by datatype
- Word segmentation and word frequencies
- Number of documents by company
- Using TFIDF to find the most characteristic words by company
- Timeseries of ESG topic distributions to analyse patterns over time

This exploration should also shed light on possible steps of data preprocessing and cleaning, such as:

- Word segmentation
- Cleaning of table of contents, contact details etc.
- Cleaning of numbers and special symbols

Outputs:

- EDA notebook with visualizations
- Notebook with data cleaning steps

Stage 2: Data annotation

In this stage, the participants use an LLM of their choice to annotate the data for sentiment (0 = negative, 0.5 = neutral, 1 = positive) to produce training data for the sentiment analysis in stage 3. In terms of LLM choice, we recommend either using GPT-3 (free, limited access) or an LLM pre-trained using multi-task learning, such as T5 or T0. In terms of the level of granularity, we recommend splitting the texts into sentences and annotating the data on sentence level. The steps can be as follows:

- Manual annotation of ca. 200 of the documents/sentences as "gold standard", sampled randomly from the full dataset
- Setup of 2-3 LLMs to test for annotation
- Experimentation with different prompting strategies (zero-shot/few-shot) and testing on the “gold standard”

Outputs:

- Manually annotated "gold standard" dataset of ca. 200 documents/sentenceAutomatically annotated dataset of all documents
- Description of prompting strategy

Stage 3: Sentiment analysis & comparison between internal and external data

In this stage, participants create a train/dev/test split (recommended proportion: 70%/15%/15%) of the data annotated in stage 2 and train a sentiment analysis classifier. The classifier should output scores on a continuous scale between 0 and 1. Participants then compare the average sentiment of internal vs. external data about a company. They sort the companies based on the difference between internal and external sentiment and do a manual follow-up research to see if the companies with the biggest gap have been explicitly involved in greenwashing during the considered timeframe.

Outputs:

- Training notebook
- Precision, recall and accuracy scores for train and test datasets

Stage 4: Alignment with Sustainable Development Goals

In this stage, participants use either an LLM or sentence embeddings to determine the relevance of specific SDGs to the different companies. The SDGs are described in the supplementary SDG dataset. For the LLM approach, we recommend formulating a prompt for directly querying the relevance of a specific SDG description for the documents about a company. For the sentence embeddings approach, participants can use an embedding library such as Laser or Sentence-BERT to embed both the SDG descriptions and the documents related to a company and compare them in terms of similarity.

Stage 5: Submission & report

In this stage, participants produce a report which describes the methodology and the outputs for the different stages.

References

[1] Naderer, Brigitte, Desirée Schmuck, and Jörg Matthes. '2.3 Greenwashing: Disinformation through Green Advertising. *Commercial communication in the digital age: Information or disinformation* 105 (2017): 120.

[2] SESAMm: How Organizations Are Using NLP To Detect Greenwashing, 2022, retrieved on April 12, 2023.

[3] Noyes, Lydia et al. A Guide to Greenwashing and How to Spot It, 2022, retrieved on April 12, 2023.

[4] Nemes, Noémi et al. An Integrated Framework to Assess Greenwashing, *Sustainability 14*(8), 2022.

[5] The Greenwashing Files, retrieved on April 12, 2023.