AS YOU SOW
30 YEARS 1992 – 2022
2020 Milvia St Suite 500
Berkeley, CA 94704
www.asyousow.org
BUILDING A SAFE, JUST, AND SUSTAINABLE WORLD SINCE 1992

Via email to rule-comments@sec.gov

August 15, 2022

Vanessa A. Countryman
Secretary Securities and Exchange Commission
100 F Street NE Washington, DC 20549-1090

RE: Investment Company Names (File No. S7-16-22)

Dear Ms. Countryman:

On behalf of *As You Sow* I welcome the opportunity to provide this comment letter in response to the Notice of Proposed Rulemaking addressing "Investment Company Names" (File No. S7-16-22) ("Proposal"). *As You Sow* is the nation's leading shareholder advocacy group, with a 30-year track record promoting environmental and social corporate responsibility. Our issue areas include climate change, ocean plastics, environmental health, racial justice, workplace diversity, and executive compensation, among others.

In 2015, we developed the Invest Your Values platform to rate and rank mutual funds and ETFs on seven ESG criteria including fossil fuels, deforestation, gender equality, civilian firearms, the prison industrial complex, military weapons, and tobacco. We update this platform every month with new fund and ETF holdings. This gives us a unique expertise on fund naming: we see every mutual fund and ETF that has "ESG" in its name and/or prospectus, many of which fail our ESG-related criteria.

Two Attachments:
I have attached to this comment letter a written report (Attachment 1) and a slide deck (Attachment 2) of a paper we published in January 2022, in partnership with the University of California San Diego (UCSD) Rady School of Business. This paper looked at U.S.-domiciled equity funds (open-end mutual funds and exchange traded funds) with "ESG" in their names.

The study scored funds using the "Invest Your Values" ESG ratings, including how the funds define ESG in their prospectuses. Critically, the study found that, of 94 funds with "ESG" in the fund name, 60 earned a D or F grade on at least one of seven ESG issues. In other words, almost two-thirds of "ESG"-named funds carry significant ESG risk on at least one ESG issue. We then applied natural language processing (NLP) to extract common definitions from the fund prospectuses. The purpose was to establish "truth in labeling" of mutual funds and ETFs and help eliminate misleading marketing, fund naming, and prospectus language. We applied the DistillBERT model using NLP methods to identify whether funds had true indicia of ESG purpose. A key finding was that the language in the prospectuses was not an accurate predictor of whether an "ESG"-named fund earned a D or F grade on an ESG issue. The report also provides recommendations and suggested improvements on appropriate phrases that likely should appear on ESG named fund prospectuses.

Comments:
We support the Securities and Exchange Commission's (SEC) efforts to clarify fund names and minimize the use of misleading or deceptive fund names. A fund name communicates to an investor important

information about the fund's characteristics, investment style, or theme. Thus, it is within the SEC's core mission to protect investors by strengthening the Names Rule.

The amendments in the Proposal expand the Names Rule's 80% investment policy to require that funds with terms suggesting that a fund focuses on a particular investment strategy must actually prioritize such considerations in its investment decision-making. We are pleased that the Proposal applies to all funds and does not propose separate sustainable or ESG funds standards.

We support the provision that any terms used in fund names that suggest an investment focus are actually consistent with those terms' plain English meaning or established industry use. This will also mitigate instances where a fund technically conforms to the 80% rule but contradicts the fund name with the remaining portion of the holding. An example of this is a fund with "fossil-fuel free" in the name that includes fossil fuel holdings in the 20% basket. This outcome is unlikely to occur where "fossil free" plays a central role in the fund's investment strategy.

We recommend, however, that the SEC clarify the language in section (d) of the proposal, "Use of ESG terms in fund names." The term "generally no more significant" is too vague, and "one or more ESG factors" is overly broad. Both may be difficult to implement. We recommend the following language for section (d):

> *Use of ESG terms in fund names*. Any fund that uses ESG terms in its name must satisfy the requirements as described in section (a)(2) of this section. A fund using ESG terms in its name will be considered materially deceptive and misleading if the fund fails to meet either criteria below:

- the fund considers ESG factors but such ESG factors are not the principal purpose of the fund's investment strategy;
- the fund does not satisfy the 80% threshold as described in section (a)(2).

Thank you for considering these comments,

Sincerely,

Andrew Behar
CEO

MGTF 490 Capstone Project:

Identify "Greenwashing" Funds using NLP in Firms' Prospectuses

Final Report

Min Yi Li | Qianchen Zheng | Hao-Che Hsu | Yin Zhu

1、Executive Summary

The purpose of our research is to determine how funds with "ESG" in their name define ESG in the prospectus. We will use natural language processing (NLP) to extract common definitions in order to create a list of terms or phrases that accurately define ESG funds. If enforced by the U.S. Securities and Exchange Commission (SEC), this will contribute to "truth in labeling" of mutual funds and ETFs and help eliminate misleading marketing, fund naming, and prospectus language. We intend to apply the DistillBERT model using NLP methods that are able to identify whether funds are "real" ESG funds or just greenwashing funds based on the prospectus and other grading information provided by firms. Finally, as prospectuses ultimately proved ineffective as a tool to distinguish good from bad ESG funds, we will provide As You Sow with recommendations and suggested improvements on appropriate phrases that likely should appear on ESG named fund prospectuses.

Keywords: Natural Language Processing (NLP), ESG, prospectus, Greenwashing fund, DistillBERT

2、Introduction/Background

Greenwashing Funds: In recent years, attention given to environmental issues has increased and the demand for environmentally sustainable products has been growing. Many companies are also facing increasing pressure from the public and government agencies, with society demanding that companies reduce their impact on the environment. The term "greenwashing" has entered into the public discussion. Greenwashing refers to the practice of companies that produce advertisements, events, or products under the guise of being environmentally friendly, when in fact, the company's products are not environmentally friendly. For example, a greenwashing company might make environmental claims such as their products are made from recycled materials or biodegradable, while these claims are only partially true or complete fraud. Such bad social conduct causes consumers' suspicion towards all green claims, and impedes the development of green industries.

Natural Language Processing (NLP): Natural language processing is a form of artificial intelligence that analyzes human language and is centered on helping machines understand human language. It is an interdisciplinary academic field combining linguistics and computer science, which has progressively become well-known in recent years. So far, natural language processing has been developed and applied in diverse fields, such as customer service, advertising, and others. Take business applications as an instance. Natural language processing can execute sentiment analysis for thousands of comments on social media to evaluate customer satisfaction, thereby calibrating subsequent marketing strategies. As the technology matures, natural language processing likely will keep making progress and become even more widely adopted.

3、Literature Review

- ***Why 'greenwashing' is an issue for sustainable investments—and how to avoid it [Alicia Adamczyk] [April 2021]***

Based on research from Morningstar, ESG funds captured $51.1 billion of net new money from investors in 2020, a record and more than double the prior year. ESG-themed investment products have attracted significant "market hype" in the sustainable investing field. However, there is no precise definition of "sustainable" products, which facilitates the greenwashing phenomenon. "Rather than taking meaningful action against climate change, the financial industry is simply greenwashing investments, or making false claims about the sustainability of their products, to make money off of a popular trend", Tariq Fancy, former BlackRock chief investment officer said. Therefore, the author proposed some reminders for investors to align their values through investment and avoid greenwashing at the same time. For example, investors should conscientiously read the prospectus about the fund objective before making an investment decision. In addition, some financial institutions provide ESG rating systems for reference. As proof, As You Sow created the Invest Your Values search tools, which investors can use to learn more about their investments. Here is the link to the tool: https://www.asyousow.org/invest-your-values

- ***Vanguard misrepresenting on its climate credentials to clients? [September 2021]***

Think tank Universal Owner released a report in September 2021 about Vanguard Group's misconduct on its investment practices. Universal Owner is a Europe-based mission-driven organization that believes the financial sector has a critical role in combating climate change and other systemic risks. The members consist of market experts and academic researchers, working at the nexus of finance, climate, and data. Vanguard, as the second-largest asset manager in the world, claimed that its investment strategy would stick to SDGs guidelines, especially on the mitigation of climate change. (Sustainable Development Goals have been adopted by the United Nations as a call action to protect the planet. In addition, the 17 SDGs stress economic development must balance social, economic and environmental sustainability.) In reality, according to the report, Vanguard still holds stakes in fossil fuel companies that derive a significant percentage of their revenue from thermal coals and tar sands. Apparently, Vanguard's ESG funds were not reflective of a move toward decarbonization and served as a greenwashing example. Universal Owner has proposed multiple guidelines and recommendations to realign its portfolio holdings, such as divesting these companies, transitioning from brown to green assets, reorienting the stewardship, etc.

- ***Big Data Shakes up ESG investing [Deutsche Bank Research] [October 2018]***

As environmental awareness rises, ESG has gained tremendous popularity in recent years. At the same time, artificial intelligence as well as big data progressively become mature with the advancement of technology. Deutsche Bank incorporated ESG issues and artificial intelligence, thereby forming a novel investment strategy. Deutsche Bank developed a new α-Dig system, which learned to read through news articles and reports, summarizing some particular words, e.g., “settlement”, “resolve”, “agreeing to pay”, which are regarded as positive ESG indicators. In addition, the models show that companies with these positive ESG indicators tend to outperform the broader European index within six months. The models show that, on average, companies outperform by two percentage points after the announcement of a litigation settlement, but this does not happen immediately. The possible reason is that investors are very inefficient at digesting ESG information and, as a result, the gains and losses from the release of such data take time to manifest. Another application of the α-Dig system is to detect greenwashers. They applied machine learning and algorithms and natural language processing techniques to infer context and understanding from company information that is increasingly subject to greenwashing. For the training session, the machine learning algorithm assessed company commitments and detected carbon-related discussions in sustainability reports, thereby identifying several relevant keywords. Afterwards, they created a ranking system to identify which companies are more ESG-focused with the combination of linguistic features, such as quantitative words, passive words, etc. According to the research team, the result is quite significant for making a prediction of the likelihood that a company would fulfill its sustainability commitment in the next year given its prior year’s sustainability report.

4、*Data Selection*

Our analysis was based on 94 funds shared with us by our sponsor, As You Sow. For a complete list of the funds, see the Appendix.

As You Sow 94's partial fund list

Fund Name	Ticker	Asset Manager	Shareclass Type	Inception Date
Pioneer Balanced ESG Fund	AOBLX	Amundi US	Open-end mutual fund	12/19/1991
BlackRock Advantage ESG Emerging Markets Equity Fund	BLZIX	BlackRock/iShares	Open-end mutual fund	8/18/2020
BlackRock Advantage ESG International Equity Fund	BRZIX	BlackRock/iShares	Open-end mutual fund	8/18/2020
BlackRock Advantage ESG U.S. Equity Fund	BIRKX	BlackRock/iShares	Open-end mutual fund	3/28/2016
BlackRock LifePath ESG Index 2030 Fund	LENIX	BlackRock/iShares	Open-end mutual fund	8/18/2020
...	...	...	...	...
Xtrackers MSCI USA ESG Leaders Equity ETF	USSG	Xtrackers	ETF	3/6/2019
Xtrackers S&P 500 ESG ETF	SNPE	Xtrackers	ETF	6/25/2019
Xtrackers S&P MidCap 400 ESG ETF	MIDE	Xtrackers	ETF	2/23/2021
Please see complete list of 94 funds in Appendix				

As You Sow provided a list of funds that contain ESG in their fund name, which will serve as our "proof of concept" sample. First, we gathered the prospectuses for each of the 94 funds from the fund companies' official websites and Bloomberg. Next, as we noticed, a prospectus would contain a description of the company's business, biographies of management members, financial statements, etc. Many sections are unrelated to our topic. In addition, some prospectuses are long and some are short, and we want to keep the contents that we put into the model for each fund a similar length. Therefore, with cautious estimations, we decided to mainly focus on the contents of funds' investment objectives, principal strategies and principal risks. As most companies lay out their goals, valuation standards, their potential risks and difficulties in these sections, we believe these are the most important parts. We realize some information from other sections could potentially be related to ESG, such as the management board, where we can find if leaders had any prior experience with ESG. However, those details are much less informative and are hard to discern. In order to avoid unnecessary disruptions and to bring uniformity to our data set, we decided to eliminate them.

5、Empirical Methodology & Analysis

The main focus is to perform NLP on the prospectuses of funds with ESG in their name, which leads to a detailed analysis on [1]tokenization, [2]stemming, [3]lemmatization, and sentiment analysis. After the prospectus cleaning process, we read the 94 fund prospectuses as separate text files in Python. The first step will be the tokenization process by using the NLTK package in Python; here, we break text in the paragraph from the fund's prospectus into smaller chunks of words or sentences. A token is a single entity that is built into a sentence or paragraph. Lemmatization and stemming would be the next steps to reduce the words into original or based words which will be linguistically correct "lemmas" and chop off the derivational affixes. We also apply post-tagging to each tokenized word so that it will identify the grammatical group of each tokenized word—whether it is noun, adjective, pronoun, adverb, etc. This process will allow us to look for relationships within a sentence. Then, we apply the F distribution to break down chunks of words to observe the most common 10-12 words in the fund's prospectus and filter out stop-words (words such as "the" ,"a," "however," etc.) and remove any noise from the tokenization and plot the distribution graph.

[1] Tokenization: the process of substituting a sensitive data element with a non-sensitive equivalent, referred to as a token, an identifier that maps back to the sensitive data through a tokenization system.
[2] Stemming: the process of reducing inflected (or sometimes derived) words to their word stem, base or root form—generally a written word form.
[3]Lemmatization: the process of grouping inflected forms together as a single base form.

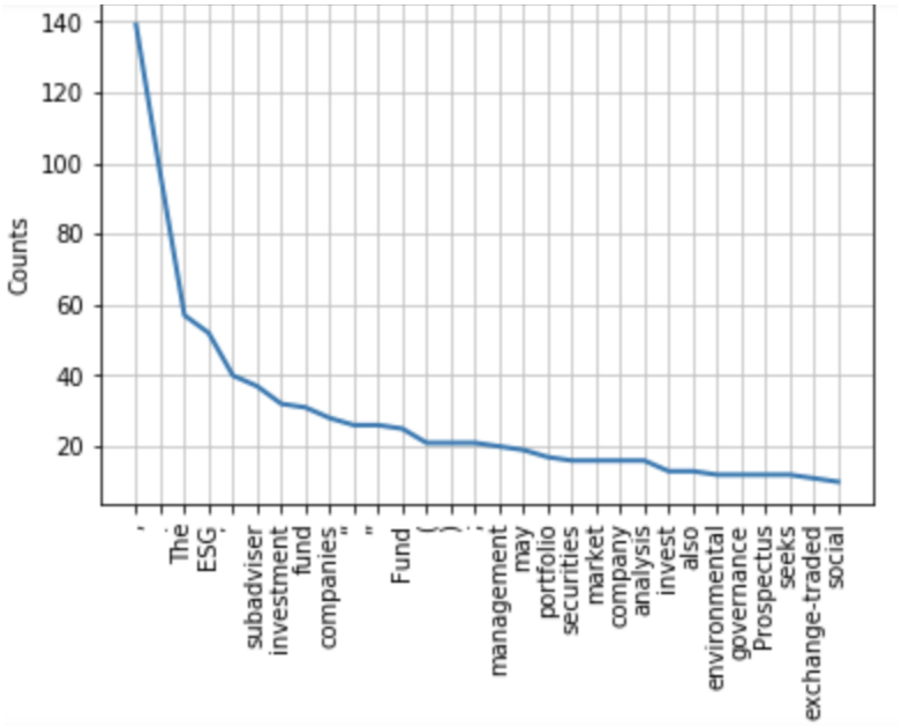


Figure 4.1: F distribution of Common 20 words (Prospectus sample No.42)

Next, as we break down information in a prospectus from paragraphs into sentences and tokenized words, we will find the frequency of each ESG key term provided by As You Sow (listed below) and newer added-on ESG key terms generated by the team:

- As You Sow provided ESG key terms: “Carbon”, “Climate”, “Divestment”, “Engagement”, “Environmental”, “ESG”, “Ethical”, “Exclusions, “Fossil”, “Green”, “Impact”, “Integration”, “Moral”, “PRI”, “Religious”, “Responsible”, “SDG”, “Social”, “SRI”, “Sustainable”, “Governance”
- Team Alpha-generated terms: “Alcohol”, “Gambling”, “Tobacco”, “Nuclear”, “Power”, “Energy”, “Thermal”, “Fuel”, “Coal”, “Oil”, “Gas”, “Weapons”, “Waste”, “Firearms”, “Ammunition”, “Minority”, “Emissions”, “Diversity”, “Gambling”, “Anti-corruption”, “Labor”, “Human rights”, “Community “,

“Quality monitoring”, “Gender”, “Community”, “Monitoring”, “Score”, “Consistent”, “Standard”, “Valuation”, “Criteria”

As You Sow provided a list of ESG named funds along with their ESG grades in seven primary areas of interest:

1. Fossil Fuels
2. Deforestation
3. Gender equality
4. Civilian Firearms
5. Prison Industrial Complex
6. Military Weapons
7. Tobacco

Grades range from A to F. In this case, we divided the 94 funds into two categories: good and bad. A good fund is identified to contain an A/B/C rating grade in all 7 grading areas. A bad fund is identified to contain D/F rating grades in any 7 grading areas. Overall, there are 34 good funds and 60 bad funds. To preliminarily differentiate the prospectuses of good and bad funds, we start the analysis by finding the frequency of the glossary terms (Figure 4.2, 4.3).

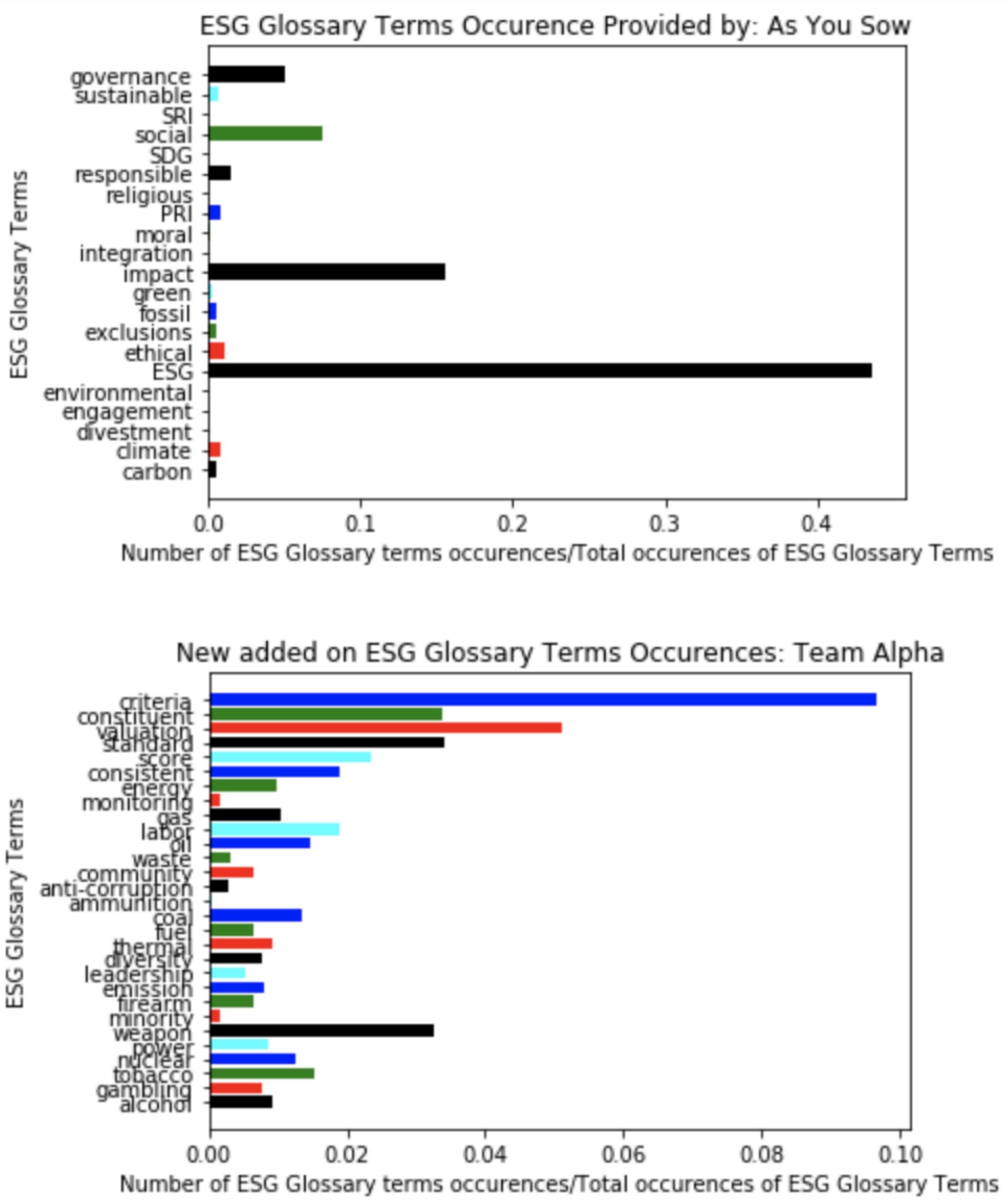
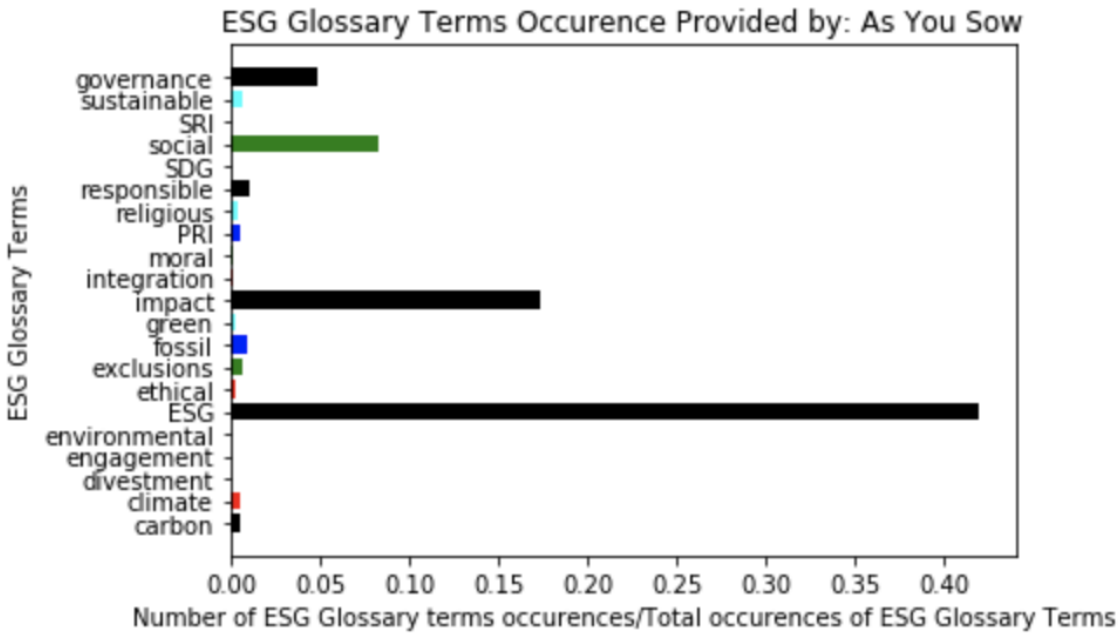


Figure 4.2: ESG Glossary terms Occurrences (Good Rated funds)



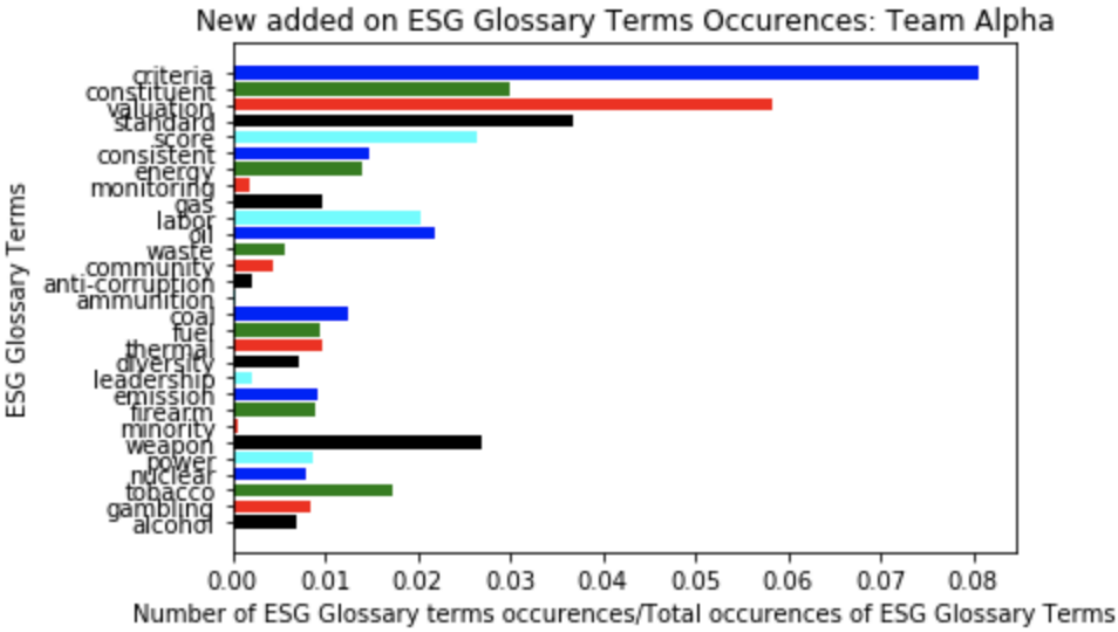


Figure 4.3 ESG Glossary terms Frequency Bar Plot (Bad Rated funds)

The results shown in Figures 4.2 and 4.3 are quite similar on both good and bad funds. The most frequent words are *ESG*, *criteria*, *standard*, *score*, among others. Based on the single word frequency results, we cannot tell the difference between a good and bad prospectus, and that the language used to distinguish good from bad funds may not be so discriminative. In this case, we generated typical “wiggle” terms that we found as we extracted those sentences and counted their occurrences in general.

Below is a list of generated typical “Wiggle” terms:

- Wiggle terms: “may consider”, “seek”, “believe”, “pursue”, “only”, “most”, “help”, “always”, “possibly”, “would”, “could”, “used”, “may”, “might”

Next step, we decided to try a new approach on seeking the differences between the prospectuses across different funds. To be specific, we extracted common phrases in both good and bad funds’ prospectuses and sought to analyze them for insights. These defensive phrases appear to obscure a manager’s objectives with respect to ESG. To target these phrases, we use the NLTK package to extract sentences that contain both ESG glossary and wiggle terms, and export them into individual text files. Next, we further extract some controversial and defensive phrases through the contents of the prospectus that have been greatly shortened by the NLP process. The results, shown in Table 4.4 below, were not conclusive.

May not exhibit positive ESG characteristics	Good	26%	Bad	30%
Fund may underperform other funds that do not have an ESG focus	Good	21%	Bad	33%
ESG characteristics may not be the same companies selected by other index providers that use similar ESG screens	Good	12%	Bad	15%
ESG practices may shift into and out of favor	Good	6%	Bad	7%
ESG characteristics/performance may change over time	Good	12%	Bad	7%
This ESG policy/investment strategy/criteria may result in the Funds foregoing opportunities to buy certain securities	Good	35%	Bad	27%
These factors may include, but are not limited to (ESG factors)	Good	3%	Bad	5%
ESG information and scores across third party data providers, may be inaccurate or incomplete	Good	15%	Bad	10%
Investors may differ in their views of ESG characteristics	Good	0%	Bad	17%
The manager seeks to fully integrate ESG criteria into the stock selection	Good	18%	Bad	13%

Table 4.4 Common Defensive ESG phrases Frequency (%) in Good & Bad Funds

We figured out that the common defensive phrases were found among both good rated funds and bad rated funds. We believe this is because companies generally want fewer constraints on their investment strategies and decisions. However, in many cases, it's not good for the investors, since investors *want* to find the funds' strategies and ESG standards on picking securities through their prospectus—and this function will be meaningless if it's too flexible for the company to change it to fit its needs. On the other hand, we think this could be due in part to the hypothesis that funds tend to use some prospectus templates while writing their own. To support this notion, we find some particular phrases that appear frequently among funds from different companies.

Last, we find some phrases are heavily used by bad rated funds, such as "Investors may differ in their views of ESG characteristics" or the "Fund may underperform other funds that do not have an ESG focus" (Table 4.4). We think these phrases could be used to identify deceptive funds as they tend to excuse themselves in advance for their bad performance and/or lack of adherence to ESG standards.

To be mindful, there are some limitations. The results are not very conclusive due to the small and unproportionate sample size. Also, the selection process is quite subjective and may not be very accurate. In addition, these selected phrases are not fixed. Some sentences with similar wording and meaning would also be counted. At the same time, we believe our preliminary findings may offer investors a starting point to engage with fund management teams for a deeper discussion on specific ESG-related practices.

- ***Sentiment Analysis using DistilBERT and Huggingface***

BERT stands for Bidirectional Encoder Representatives from transformers. It became a popular pretrained language model starting in 2018 after the release of an academic paper on the topic.[4] BERT is an open-sourced model code that broke several records for difficult language-based tasks. It is a pre-trained model on massive datasets that enables NLP to use this free powerhouse. DistillBERT, obviously as the name suggests, is a distilled version of the BERT model. DistillBERT is a Huggingface transformer model, which is smaller and faster than the BERT model (and more suitable for very large datasets). It was pre-trained on the raw text only compared to the BERT model. Our objective is to apply the DistillBERT model to predict whether an ESG named fund is good or bad. To emphasize, the ratings we applied sorted the original 94 funds into 60 bad funds and 34 good funds.

By random shuffling, we split the dataset into training for *80%* and testing for *20%*. We then apply a tokenizer from the transformer DistillBERT model to the trained, test dataset, which will give two lists: input_ids (numerical representation for the sequence which the model will use), attention_mask (values to attend or not: 0 or 1). Then, we turn the label and encodings into datasets so that data can be batched easily, which we will train. Next, the most important step is to fine tune the DistillBERT model so that the trainer will be expecting and able to define the TrainingArgument and instantiate the Trainer. The output will show two lists: steps and training loss (Figure 4.5). The training loss can be interpreted as model error or mean square error. In this case, the lower the training loss, the better our model performed. Even though the training loss goes slightly up and down, in the long term, the training loss increases over time, which implies the model is *not* learning over time.

[4] Devlin, Jacob, Ming-Wei Chang, Kenton Lee and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Full article here: 1810.04805.pdf (arxiv.org)

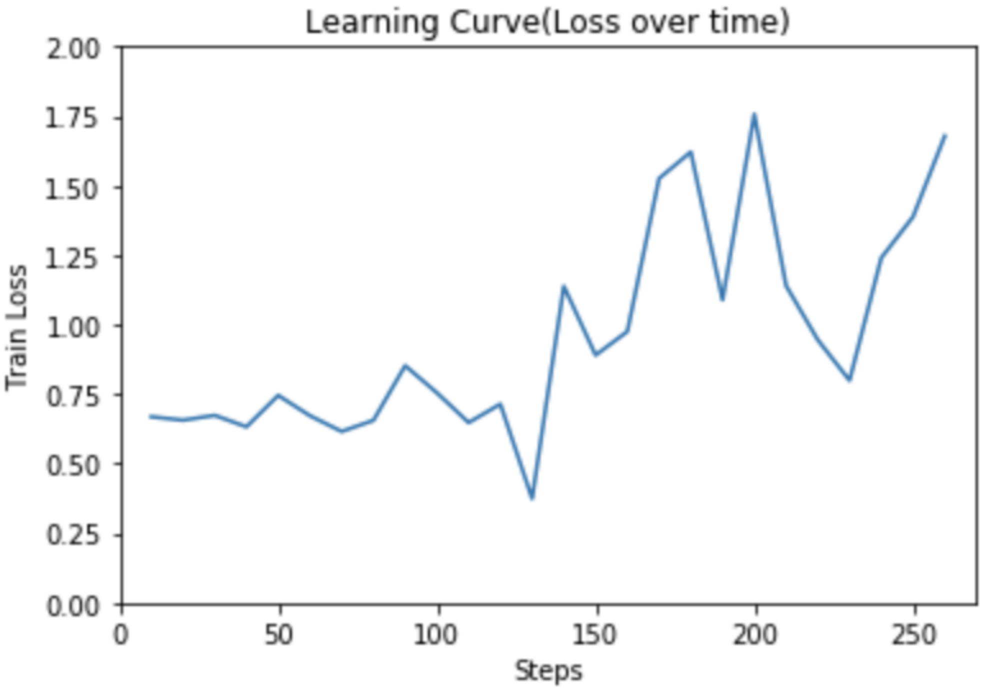


Figure 4.5 Learning Curve (Training Loss over time)

Finally, we conduct fine tuning with native PyTorch on testing datasets. The output will be the prediction of good/bad funds presented as: probability of bad fund, probability of good fund, true label of fund (0 = bad, 1 = good), and its fund ticker (Table 4.6). Since we split the dataset into 80% for training, 20% for testing, 19 funds will be randomly selected for testing. For example, Xtrackers MSCI USA ESG Leaders Equity ETF[‘USSG’], which is a good fund with all ratings above D/F grade has categorised into bad funds with probability of 0.9879, but only probability of 0.0121 to be categorised into good funds. Similar results applied to other true-label good funds but were categorised as bad funds. According to the model prediction output in Table 4.6, (the second and third columns show the probability of being a bad fund and the probability of being a good fund) we observed the model categorized every fund as a bad fund perhaps because of the larger or disproportional sample size for bad funds (60 bad funds vs 34 good funds in total, 13 bad funds vs 6 good funds in testing datasets) or the model simply cannot tell the difference between good and bad funds. Based on our observations, the DistillBERT model cannot give an accurate prediction on good/bad funds based on the fund prospectuses. At this point, we validated that the model is unable to identify greenwashing funds by only processing fund prospectuses.

```
tensor([0.9843, 0.0157], grad_fn=<SelectBackward0>) tensor([0]) ['IQSI']
tensor([0.9879, 0.0121], grad_fn=<SelectBackward0>) tensor([1]) ['USSG']
tensor([0.9868, 0.0132], grad_fn=<SelectBackward0>) tensor([1]) ['CACG']
tensor([0.9868, 0.0132], grad_fn=<SelectBackward0>) tensor([0]) ['NULV']
tensor([0.9881, 0.0119], grad_fn=<SelectBackward0>) tensor([0]) ['BRZIX']
tensor([0.9861, 0.0139], grad_fn=<SelectBackward0>) tensor([1]) ['IGIAX']
tensor([0.9771, 0.0229], grad_fn=<SelectBackward0>) tensor([0]) ['GSIFX']
tensor([0.9875, 0.0125], grad_fn=<SelectBackward0>) tensor([1]) ['TPYAX']
tensor([0.9879, 0.0121], grad_fn=<SelectBackward0>) tensor([0]) ['RAFE']
tensor([0.9886, 0.0114], grad_fn=<SelectBackward0>) tensor([0]) ['LEWIX']
tensor([0.9872, 0.0128], grad_fn=<SelectBackward0>) tensor([0]) ['LRGE']
tensor([0.9887, 0.0113], grad_fn=<SelectBackward0>) tensor([0]) ['LENIX']
tensor([0.9864, 0.0136], grad_fn=<SelectBackward0>) tensor([0]) ['ESML']
tensor([0.9872, 0.0128], grad_fn=<SelectBackward0>) tensor([0]) ['CFCV']
tensor([0.9744, 0.0256], grad_fn=<SelectBackward0>) tensor([0]) ['GAGVX']
tensor([0.9874, 0.0126], grad_fn=<SelectBackward0>) tensor([1]) ['SNPE']
tensor([0.9873, 0.0127], grad_fn=<SelectBackward0>) tensor([1]) ['BCAMX']
tensor([0.9878, 0.0122], grad_fn=<SelectBackward0>) tensor([0]) ['XJH']
tensor([0.9871, 0.0129], grad_fn=<SelectBackward0>) tensor([0]) ['ESGE']
```

Table 4.6 DistillBERT model prediction on good/bad funds

6、Recommendations & Improvement

Based on the above empirical methods, we can hardly tell the difference between the prospectuses for good and bad funds. Since every issuer has its own writing style or shares similar templates among companies, it is quite difficult to identify the specific phrases of good fund prospectuses. We can still summarize the contents that the ESG fund prospectus should include. We browsed through the 34 good prospectuses and boiled down some recommended characteristics based on our judgment. The characteristics are delineated as follows and corresponding examples would also be provided for each.

First, a good ESG prospectus should disclose detailed information about the ESG criteria and references from third-party authorities. Some funds with poor ratings may not have precise language about the ESG criteria they use. For good funds, for example, the issuers will clearly state their requirements for each type of industry involved in environmental pollution and social issues, rather than vaguely saying that ESG characteristics would be different in the view of each investor. In addition, the issuers will clearly state their investment style, such as positive screening or negative screening. To the extent that an ESG-related fund is reluctant to disclose the ESG selection standard in their portfolio, they are likely trying to set the stage for non-compliance with ESG requirements.

Here, we offer an example from the SPDR® S&P 500® ESG ETF prospectus that uses what we view as imprecise language with respect to ESG. The following was extracted from the ETF's principal investment strategy section:

"*The Index is designed to measure the performance of securities meeting certain sustainability criteria (criteria related to ESG factors), while maintaining similar overall industry group weights as the S&P 500 Index. Securities eligible for inclusion in the Index comprise all constitu*

• *Have involvement with tobacco-related products and services, based on certain levels of production or revenue, or hold certain ownership stakes in a company involved in these products and services, as determined by Sustainalytics;*

• *Are involved in controversial weapons, including cluster weapons, landmines, biological or chemical weapons, depleted uranium weapons, white phosphorus weapons, or nuclear weapons, or hold certain ownership stakes in a company involved in these activities, as determined by Sustainalytics;*

• *Have a United Nations Global Compact ("UNGC") score in the bottom 5% of all UNGC-scored companies globally, as determined by Arabesque;*

• *Have an S&P DJI ESG Score, as assigned by SAM, that falls within the worst 25% from each Global Industry Classification Standard (GICS) industry group among the combined constituents of the S&P Global LargeMidCap Index and the S&P Global 1200 Index;*

• *Generate greater than 5% of their revenue from thermal coal extraction or electricity generation, as determined by Sustainalytics; or*

• *Do not have (i) Sustainalytics coverage for tobacco-, controversial weapons- and thermal coal-related involvement; (ii) a UNGC score determined by Arabesque; or (iii) an S&P DJI ESG Score.*

UNGC scores provided by Arabesque implement quantitative models and data to arrive at a company score based on the normative principles of the UNGC: human rights, labor rights, the environment, and anti-corruption. S&P DJI ESG Scores are assigned by SAM, an ESG scoring business unit of S&P Global Inc. (an affiliate of the Index Provider (defined below)), using its Corporate Sustainability Assessment, which is an annual evaluation of a company, based on ESG factors that SAM determines are financially *material to the company, relative to its industry peer companies as determined by SAM.*"

Second, we were surprised to find that some of the good prospectuses mentioned that they would not simply select the targets with high ESG scores to include in their portfolios. Instead, they would use their power to raise the attention of companies to ESG issues. Before investing in such funds, we believe investors would be well served to ask the managers for information regarding their success engaging with specific companies to improve the companies' ESG practices and, subsequently, their scores.

From our perspective, those funds that aggressively promote ESG are likely to be far from greenwashing. The following passage is extracted from the Boston Common ESG Impact International Fund prospectus:

"We use our voice as a shareowner to raise environmental, social, and governance issues with the management of select portfolio companies through a variety of channels. These may include engaging in dialogue with management, participating in shareholder proposal filings, voting proxies in accordance with our proxy voting guidelines, and participating in the annual shareholder meeting process."

Third, the sentences or paragraphs should convey the intent of adding ESG in the investment thesis in the first place. We know that lots of greenwashing funds tend to consider multiple factors during the stock selection process and dilute the importance of ESG. Therefore, we think that the truly ESG fund prospectus should regard ESG as their core value. Here, we share an example from the DWS ESG Core Equity Fund prospectus"

"Prior to considering financial information, the security selection process evaluates an issuer based on Environmental, Social and Corporate Governance (ESG) criteria."

Fourth, monitoring is a vital part of portfolio management. Especially for the field of ESG, the performance of a company and the evaluation standard of individual ESG characters could change frequently. Thus, we are aware that the ratings could vary in a short period of time. The manager should monitor the ESG performance of the components periodically for the reconstitution process. Moreover, the manager should track and adjust the structure of the fund on an ongoing basis. In other words, those who stick to monitoring techniques likely tend to pay more attention to ESG compliance and embody the essence of ESG. The Gabelli ESG Fund prospectus and prospectus for its TrueShares ESG Active Opportunities ETF offer good examples:

"The Adviser will monitor each holding on a regular basis to ensure its compliance with the Fund's guidelines." - extracted from Gabelli ESG Fund prospectus

"The portfolio is then monitored by the Adviser and Sub-Adviser and the weightings are adjusted regularly with a focus on each company's ESG Rating and Relative Value." - - extracted from TrueShares ESG Active Opportunities ETF prospectus

7、Conclusion

In our research, we have studied the ESG named funds in the list provided by As You Sow from different aspects. The entire study is based on the ESG scores provided by As You Sow for the funds in the list. The following are a couple of results we have used to derive different dimensions depending on the study:

1） Using the NLTK package in Python, each prospectus will be tokenized, stemmed and lemmatized, and then the occurrence pattern and frequency of ESG key terms will be analyzed. The frequency of keywords in all prospectuses are quite similar, which makes it hard to distinguish the possibility of greenwashing. Based on this method, it is hard to differentiate between funds by observing differences in the frequency of keywords.

2） The funds in the list are sorted into different categories based on their high or low ratings in the seven ESG main areas of engagement. Upon comparative analysis, it is found that defensive phrases appear more frequently in the poorly rated funds. We presume that these poor ESG performers tend to excuse poor ESG performance in advance—which may be a marker to identify funds we'd consider "suspicious."

3） The DistillBERT model is used to find linguistic patterns in the prospectus and analyze how ESG funds differ under different ratings. The learning curve of the model is on an upward trend, which does not produce significant results and implies the model is not learning over time. Overall, the DistillBERT model is unable to predict accurately whether the fund is good or bad based on the fund prospectuses. So using two different NLP approaches, we arrive at the same conclusion: we cannot identify good funds from bad funds using prospectuses.

4） To evaluate whether a fund is suspected of greenwashing, a convenient entry point is to consider the ESG score rating of each company held in the fund's portfolio. We studied the prospectuses of highly rated funds in the list provided by As You Sow and identified some of what we consider good attributes in the prospectuses of the most highly rated ESG funds. Namely, these funds offer disclosure of compliance with ESG indices and criteria, continuous tracking of the portfolio and potential changes to ESG eligibility, and positive public orientation towards environmental protection.

After performing the above analysis, we conclude that the linguistic pattern of the prospectus of the fund has a relatively low correlation with its ESG rating. We take into account some limitations of our research: the small sample size and the disproportionate number of ESG funds scored at each level, as well as the fact that some of the keywords used for the analysis were chosen with some subjectivity. Meanwhile, considering the instability of the results of the current analysis, the prospectus should *not* be considered as an only criterion for discriminating greenwashing. A prospectus contains information about the company management team, financial performance, exposure risk and other related information that investors would like to know. Since the content that needs to be included is not limited to the underlying asset composition of the portfolio, other information is easy to obscure our research results. Moreover, if other materials, such as fact books, annual reports or sustainability reports, would be included in the study, the research might have been able to reach more effective, definitive findings.

8、 *References & Works Cited*

Adamczyk, Alicia. "Why 'Greenwashing' Is an Issue for Sustainable Investments-and How to Avoid It." *CNBC*, CNBC, 23 Apr. 2021, https://www.cnbc.com/2021/04/23/what-to-know-about-greenwashing-in-sustainable-investments.html.

Mahtani, Sahil, Jim Reid and Luke Templeman. "Big Data Shakes up ESG Investing." *Deutsche Bank Research*, 4 Oct. 2018, https://www.dbresearch.com/PROD/RPS_EN-PROD/Big_data_shakes_up_ESG_investing/RPS_EN_DOC_VIEW.calias?rwnode=PROD0000000000435639&ProdCollection=PROD0000000000478852.

Universal Ownership Research Report, "Vanguard and Universal Ownership", September. 2021 https://www.universalowner.org/utility-report

Rating reference on mutual funds

Fossil fuel grade

https://fossilfreefunds.org/how-it-works

Deforestation grade

https://deforestationfreefunds.org/how-it-works

Gender equality grade

https://genderequalityfunds.org/how-it-works

Civilian firearm grade

https://gunfreefunds.org/how-it-works

Prison industrial complex grade

https://prisonfreefunds.org/how-it-works

Military weapons grade

https://weaponfreefunds.org/how-it-works

Tobacco grade

https://tobaccofreefunds.org/how-it-works

Appendix

Complete list of As You Sow 94 funds

Fund Name	Ticker	Asset Manager	Shareclass Type	Inception Date	Rating

Pioneer Balanced ESG Fund	AOBLX	Amundi US	Open-end mutual fund	12/19/1991	Bad
BlackRock Advantage ESG Emerging Markets Equity Fund	BLZIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock Advantage ESG International Equity Fund	BRZIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock Advantage ESG U.S. Equity Fund	BIRKX	BlackRock/iShares	Open-end mutual fund	3/28/2016	bad
BlackRock LifePath ESG Index 2030 Fund	LENIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock LifePath ESG Index 2035 Fund	LEJIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock LifePath ESG Index 2040 Fund	LEKIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock LifePath ESG Index 2045 Fund	LEHIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock LifePath ESG Index 2050 Fund	LEGIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock LifePath ESG Index 2055 Fund	LEEIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock LifePath ESG Index 2060 Fund	LEZIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
BlackRock LifePath ESG Index 2065 Fund	LEWIX	BlackRock/iShares	Open-end mutual fund	8/18/2020	Bad
iShares ESG Advanced MSCI EAFE Index ETF	DMXF	BlackRock/iShares	ETF	6/16/2020	Bad
iShares ESG Advanced MSCI EM ETF	EMXF	BlackRock/iShares	ETF	10/6/2020	Bad
iShares ESG Aware MSCI EAFE ETF	ESGD	BlackRock/iShares	ETF	6/28/2016	Bad
iShares ESG Aware MSCI EM ETF	ESGE	BlackRock/iShares	ETF	6/28/2016	Bad
iShares ESG Aware MSCI USA ETF	ESGU	BlackRock/iShares	ETF	12/1/2016	Bad
iShares ESG Aware MSCI USA Small-Cap ETF	ESML	BlackRock/iShares	ETF	4/10/2018	Bad
iShares ESG MSCI EM Leaders ETF	LDEM	BlackRock/iShares	ETF	2/5/2020	Bad
iShares ESG MSCI USA Leaders ETF	SUSL	BlackRock/iShares	ETF	5/7/2019	Good
iShares MSCI USA ESG Select ETF	SUSA	BlackRock/iShares	ETF	1/24/2005	Good
iShares® ESG Advanced MSCI USA ETF	USXF	BlackRock/iShares	ETF	6/16/2020	Good
iShares® ESG Screened S&P 500 ETF	XVV	BlackRock/iShares	ETF	9/22/2020	Good
iShares® ESG Screened S&P Mid-Cap ETF	XJH	BlackRock/iShares	ETF	9/22/2020	Bad
iShares® ESG Screened S&P Small-Cap ETF	XJR	BlackRock/iShares	ETF	9/22/2020	Bad
Boston Common ESG Impact International Fund	BCAIX	Boston Common	Open-end mutual fund	12/29/2010	Good
Boston Common ESG Impact U.S. Equity Fund	BCAMX	Boston Common	Open-end mutual fund	4/30/2012	Good
Coho Relative Value ESG Fund	CESGX	Coho	Open-end mutual fund	11/27/2019	Good
DWS ESG Core Equity Fund	MIDVX	DWS	Open-end mutual fund	8/1/2005	Good
DWS ESG International Core Equity Fund	DURAX	DWS	Open-end mutual fund	11/11/2014	Bad
Dana Epiphany ESG Equity Fund	ESGIX	Dana Investment	Open-end mutual fund	2/13/2008	Good
Dana Epiphany ESG Small Cap Equity Fund	DSCIX	Dana Investment	Open-end mutual fund	11/3/2015	Bad

Direxion MSCI USA ESG - Leaders vs. Laggards ETF	ESNG	Direxion Funds	ETF	2/5/2020	Bad
Ecofin Global Water ESG Fund	EBLU	Ecofin	ETF	2/14/2017	Good
First Trust TCW ESG Premier Equity ETF	EPRE	First Trust	ETF	5/25/2021	Bad
Fisher Investments Institutional Group ESG Stock Fund for Retirement Plans	QDVSX	Fisher Investments	Open-end mutual fund	12/13/2019	Good
FlexShares STOXX Global ESG Select Index Fund	ESGG	Flexshares Trust	ETF	7/13/2016	Bad
FlexShares STOXX US ESG Select Index Fund	ESG	Flexshares Trust	ETF	7/13/2016	Good
ClearBridge All Cap Growth ESG ETF	CACG	Franklin Templeton Investments	ETF	5/3/2017	Good
ClearBridge Dividend Strategy ESG ETF	YLDE	Franklin Templeton Investments	ETF	5/22/2017	Bad
ClearBridge Large Cap Growth ESG ETF	LRGE	Franklin Templeton Investments	ETF	5/22/2017	Bad
Clearbridge Focus Value ESG ETF	CFCV	Franklin Templeton Investments	ETF	5/27/2020	Bad
Gabelli ESG Fund	ESGHX	Gabelli	Open-end mutual fund	6/1/2007	Good
Glenmede Responsible ESG U.S. Equity Portfolio	RESGX	Glenmede	Open-end mutual fund	12/22/2015	Bad
Goldman Sachs ESG Emerging Markets Equity Fund	GEBAX	Goldman Sachs	Open-end mutual fund	5/31/2018	Good
Goldman Sachs International Equity ESG Fund	GSIFX	Goldman Sachs	Open-end mutual fund	12/1/1992	Bad
Goldman Sachs U.S. Equity ESG Fund	GAGVX	Goldman Sachs	Open-end mutual fund	11/30/2009	Bad
Gotham ESG Large Value Fund	GESGX	Gotham	Open-end mutual fund	12/28/2018	Bad
Hartford Schroders ESG US Equity ETF	HEET	Hartford Mutual Funds	ETF	8/10/2021	Bad
Horizon ESG Defensive Core Fund	HESAX	Horizon Investments	Open-end mutual fund	1/8/2020	Good
IQ Candriam ESG International Equity ETF	IQSI	IndexIQ	ETF	12/16/2019	Bad
IQ Candriam ESG US Equity ETF	IQSU	IndexIQ	ETF	12/16/2019	Good
Inspire Tactical Large Cap ESG ETF	RISN	Inspire	ETF	7/15/2020	Bad
Integrity ESG Growth & Income Fund	IGIAX	IntegrityVikingFunds	Open-end mutual fund	1/3/1995	Good
Invesco Real Assets ESG ETF	IVRA	Invesco	ETF	12/18/2020	Bad
Invesco US Large Cap Core ESG ETF	IVLC	Invesco	ETF	12/18/2020	Good
John Hancock ESG International Equity Fund	JTQAX	John Hancock	Open-end mutual fund	12/14/2016	Good
John Hancock ESG Large Cap Core Fund	JHJAX	John Hancock	Open-end mutual fund	6/6/2016	Good
Kennedy Capital ESG SMID Cap Fund	KESGX	Kennedy Capital Management	Open-end mutual fund	6/28/2019	Bad
KraneShares MSCI China ESG Leaders Index ETF	KESG	KraneShares	ETF	7/29/2020	Bad
Matthews Asia ESG Fund	MISFX	Matthews Asia Funds	Open-end mutual fund	4/30/2015	Good
AVDR US LargeCap ESG ETF	AVDG	New Age Alpha	ETF	12/29/2020	Bad
PIMCO RAFI ESG U.S. ETF	RAFE	PIMCO	ETF	12/18/2019	Bad
LGBTQ100 ESG ETF	LGBT	Procure ETF Trust	ETF	5/17/2021	Bad
SPDR® S&P 500® ESG ETF	EFIV	SPDR State Street Global Advisors	ETF	7/27/2020	Good

Sit ESG Growth Fund	IESGX	Sit	Open-end mutual fund	7/1/2016	Bad
Stance Equity ESG Large Cap Core ETF	STNC	Stance	ETF	3/12/2021	Bad
Nuveen ESG Emerging Markets Equity ETF	NUEM	TIAA Investments/Nuveen	ETF	6/6/2017	Bad
Nuveen ESG International Developed Markets Equity ETF	NUDM	TIAA Investments/Nuveen	ETF	6/6/2017	Bad
Nuveen ESG Large-Cap ETF	NULC	TIAA Investments/Nuveen	ETF	6/3/2019	Good
Nuveen ESG Large-Cap Growth ETF	NULG	TIAA Investments/Nuveen	ETF	12/13/2016	Good
Nuveen ESG Large-Cap Value ETF	NULV	TIAA Investments/Nuveen	ETF	12/13/2016	Bad
Nuveen ESG Mid-Cap Growth ETF	NUMG	TIAA Investments/Nuveen	ETF	12/13/2016	Bad
Nuveen ESG Mid-Cap Value ETF	NUMV	TIAA Investments/Nuveen	ETF	12/13/2016	Bad
Nuveen ESG Small-Cap ETF	NUSC	TIAA Investments/Nuveen	ETF	12/13/2016	Bad
Nuveen Winslow Large-Cap Growth ESG Fund	NWCAX	TIAA Investments/Nuveen	Open-end mutual fund	5/15/2009	Good
Touchstone Global ESG Equity Fund	TEQAX	Touchstone	Open-end mutual fund	12/19/1997	Good
Touchstone International ESG Equity Fund	TPYAX	Touchstone	Open-end mutual fund	12/3/2007	Good
Trillium ESG Global Equity Fund	PORIX	Trillium Mutual Funds	Open-end mutual fund	3/30/2007	Good
Trillium ESG Small/Mid Cap Fund	TSMDX	Trillium Mutual Funds	Open-end mutual fund	8/31/2015	Bad
TrueShares ESG Active Opportunities ETF	ECOZ	TrueShares	ETF	2/28/2020	Good
Trend Aggregation ESG ETF	TEGS	Tuttle	ETF	5/7/2020	Bad
Vanguard ESG International Stock ETF	VSGX	Vanguard	ETF	9/18/2018	Bad
Vanguard ESG U.S. Stock ETF	ESGV	Vanguard	ETF	9/18/2018	Good
Vanguard Global ESG Select Stock Fund	VESGX	Vanguard	Open-end mutual fund	6/5/2019	Good
WisdomTree Emerging Markets ESG Fund	RESE	WisdomTree	ETF	4/7/2016	Bad
WisdomTree International ESG Fund	RESD	WisdomTree	ETF	11/3/2016	Bad
WisdomTree U.S. ESG Fund	RESP	WisdomTree	ETF	2/23/2007	Bad
Xtrackers MSCI EAFE ESG Leaders Equity ETF	EASG	Xtrackers	ETF	9/5/2018	Bad
Xtrackers MSCI Emerging Markets ESG Leaders Equity ETF	EMSG	Xtrackers	ETF	12/4/2018	Bad
Xtrackers MSCI USA ESG Leaders Equity ETF	USSG	Xtrackers	ETF	3/6/2019	Good
Xtrackers S&P 500 ESG ETF	SNPE	Xtrackers	ETF	6/25/2019	Good
Xtrackers S&P MidCap 400 ESG ETF	MIDE	Xtrackers	ETF	2/23/2021	Bad
Xtrackers S&P SmallCap 600 ESG ETF	SMLE	Xtrackers	ETF	2/23/2021	Bad

ATTACHMENT 2

AS YOU SOW

Capstone Project Final Presentation

Team Alpha

Min Yi Li

Hao-Che Hsu

Yin Zhu

Qianchen Zheng

2021.12.01

Table of Contents

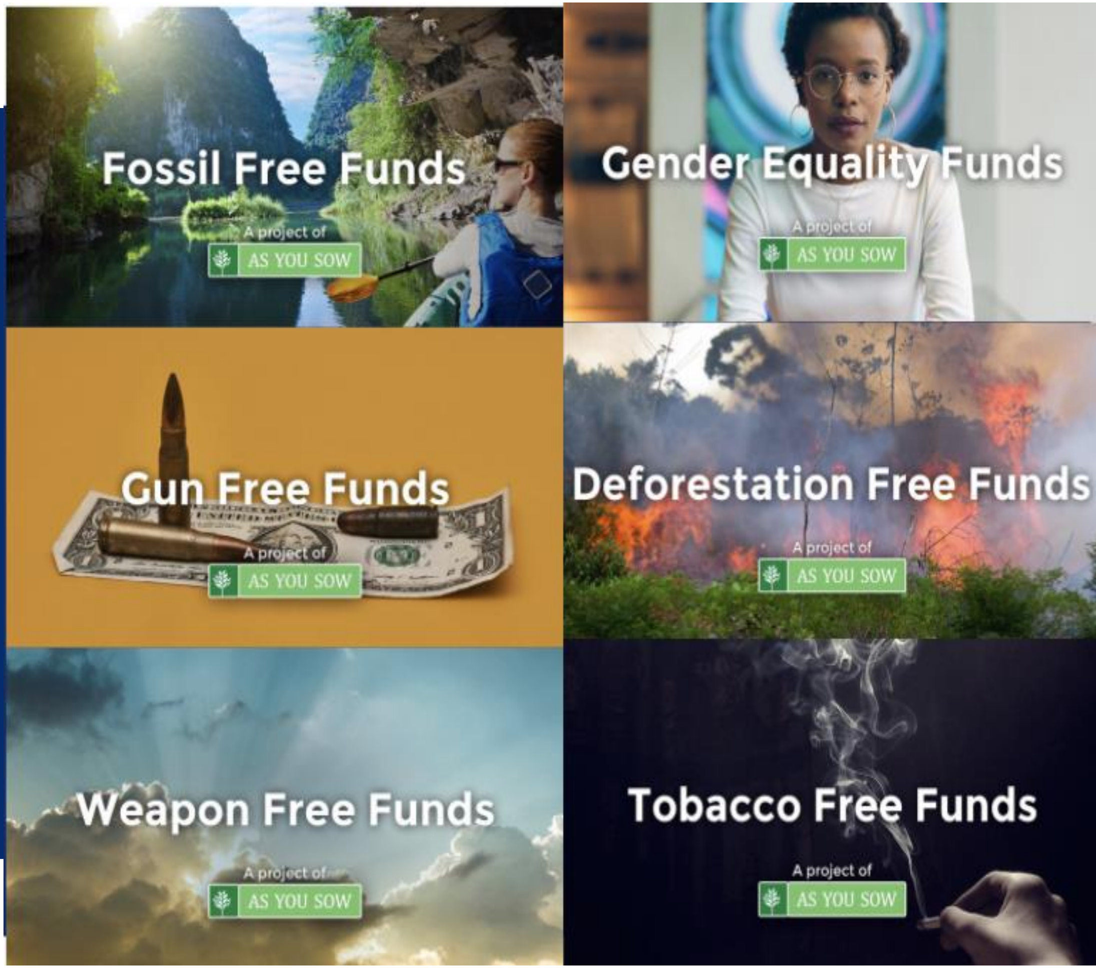
1 Introduction & Background

- Greenwashing Funds: the practice of companies that produce advertisements, events, or products under the guise of being environmentally friendly, when in fact, the company's products are not environmentally friendly

- Natural Language Processing (NLP):a form of artificial intelligence that analyzes human language and is centered on helping machines understand human language

2 ESG Fund Ratings

- Divided 94 funds into good and bad fund categories
- Scales: Fossil fuel grade, Deforestation grade, Gender equality grade, Civilian firearms grade, Prison industrial complex grade, Military weapons grade, Tobacco grade
- Good Funds: do not contain any D/F rating in all scales
- Bad Funds: contain D/F rating in any 7 scales
- Good funds: 34
- Bad funds: 60



3

Key Terms

- ESG glossary:
  “Alcohol”, “Gambling”, “Tobacco”, “Nuclear”, “Power”, “Energy”, “Thermal” ,“Fuel”, “Coal”, “Oil”, “Gas”, “Weapons”, “Waste”, “Firearms”, “Ammunition”, “Minority”, “Emissions”, “Diversity”, “Gambling”, “Anti-corruption”,“Labor”,“Human rights”, “Community “, “Quality monitoring”, “Gender”

- Wiggle term:
  “May”, “Seek”, “Believe”, “Pursue”, “Only”, “Most”, “Help”, “Always”, “Possibly”, “Would”, “Could”, “Used”, “May consider”, “Might”

4 Empirical Method & Analysis

Natural Language Processing

- ➢ Use NLTK package to perform text analysis
- ❑ Tokenization - yes
- ❑ Remove stopwords - yes
- ❑ Lemmatization and Stemming - yes
- ❑ Post-Tagging - yes
- ❑ F distribution(10 common words) - yes
- ❑ Sort Good/Bad rating funds - yes
- ❑ ESG terms frequency - yes
- ❑ Plot occurrences - yes
- ❑ Extract phrases(ESG & Wiggle terms) - yes

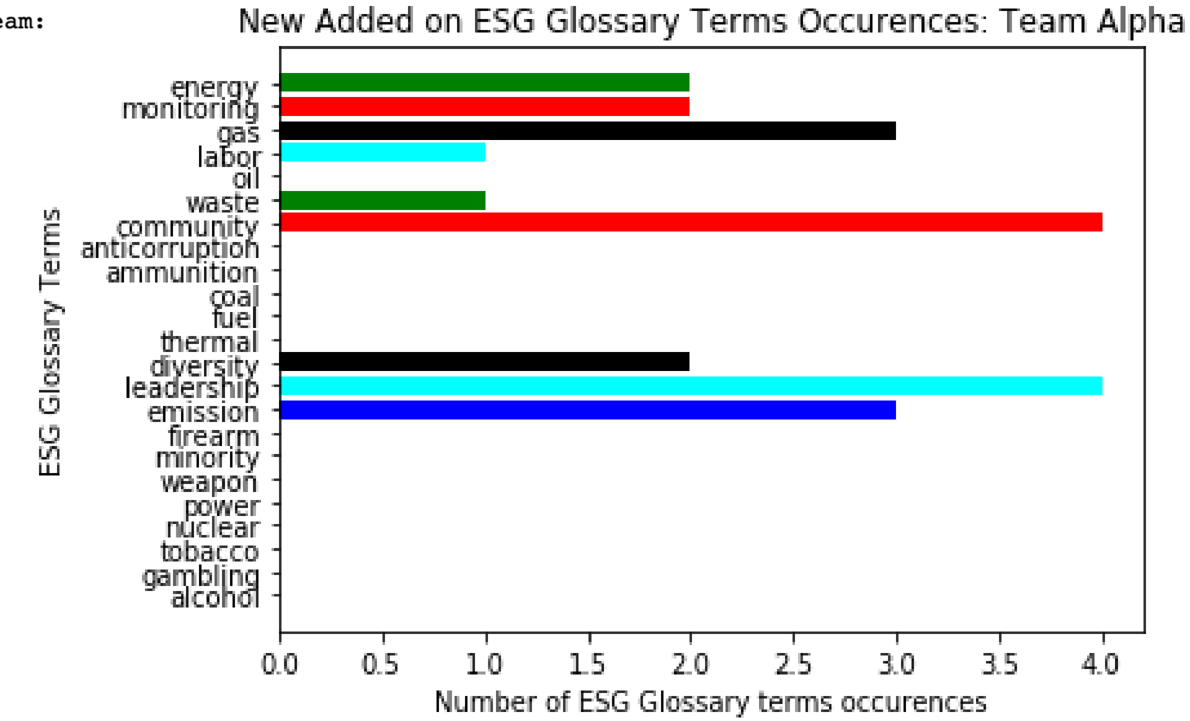
```
import nltk
import numpy as np
import matplotlib.pyplot as plt
from nltk.tokenize import sent_tokenize
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
from nltk.stem.wordnet import WordNetLemmatizer
from nltk.stem.porter import PorterStemmer
from nltk.tag import pos_tag
from nltk.probability import FreqDist


def nlp(text):
    tokenized_word=word_tokenize(text) #tokenized prospectus text file into words
    #print(tokenized_word)
    filtered_words = [word for word in tokenized_word if word not in stopwords.words('english')]
    #remove stopwords from tokenized words in prospectus
    tokenized_sent=sent_tokenize(text)
    #print(tokenized_sent)
    #tokenized prospectus text file into sentences

    #Stemming: reduces words to their word root word or chops off the derivational affixes
    ps = PorterStemmer()
    stemmed_words=[]
    for w in tokenized_sent:
        stemmed_words.append(ps.stem(w))
    #print("Stemmed Sentence:",stemmed_words)

    #Lemmatization: reduces words to their base word, which is linguistically correct lemmas.
    lemmatizer = WordNetLemmatizer()
    #stem = PorterStemmer(text)
    lemmatized_output = ' '.join([lemmatizer.lemmatize(w) for w in tokenized_word])
    #print(lemmatized_output)
```

```
#Post-Tagging:identify the grammatical group of a given word.
#Whether it is a NOUN, PRONOUN, ADJECTIVE, VERB, ADVERBS, etc. based on the context.
#POS Tagging looks for relationships within the sentence and assigns a corresponding tag to the word.
post_tagging = pos_tag(tokenized_word)
#Print(post_tagging)

fdist = FreqDist(filtered_words) #apply f distribution after tokenization and removal of stopwords
fdist.most_common(12) #find the most common words in the prospectus
print("The most common words in this prospectus are: " + str(fdist.most_common(10)))
fdist.plot(12,cumulative=False) #plot the f distribution
plt.show()
```

4 Empirical Visualization

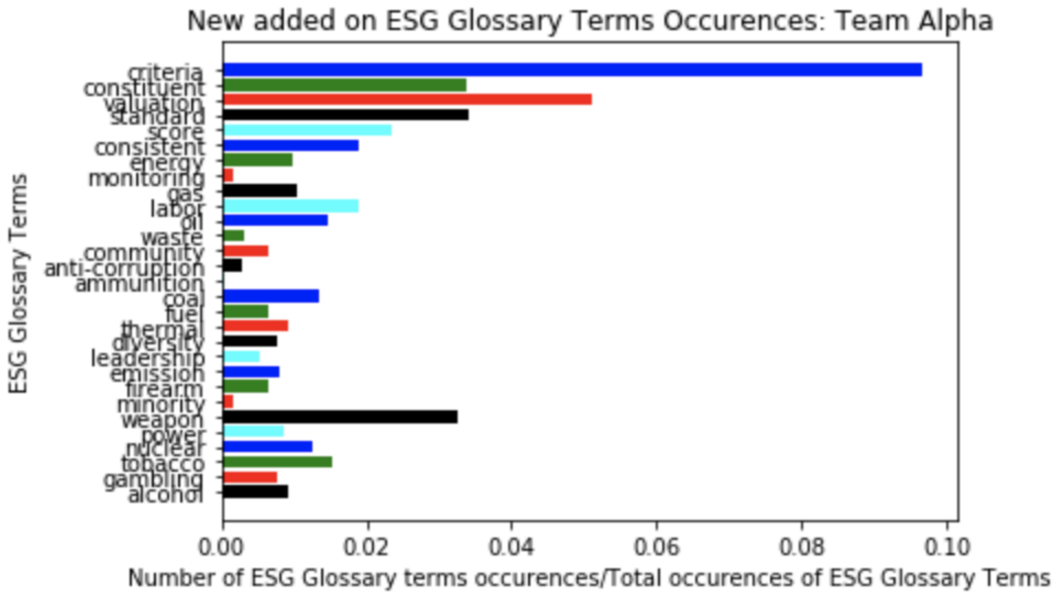
ESG Glossary Terms Occurrences: Team Alpha (Prospectus Sample No.42)

```
Below are the occurence of new added on ESG glossary terms created by the team:
no occurrence on term: alcohol
no occurrence on term: gambling
no occurrence on term: tobacco
no occurrence on term: nuclear
no occurrence on term: power
no occurrence on term: weapon
no occurrence on term: minority
The occurrence times of the term leadership is: 4
The occurrence times of the term diversity is: 2
no occurrence on term: firearms
The occurrence times of the term emission is: 3
no occurrence on term: thermal
no occurrence on term: oil
no occurrence on term: fuel
no occurrence on term: coal
no occurrence on term: ammunition
no occurrence on term: anti-corruption
The occurrence times of the term community is: 4
The occurrence times of the term waste is: 1
The occurrence times of the term labor is: 1
The occurrence times of the term gas is: 3
The occurrence times of the term monitoring is: 2
The occurrence times of the term energy is: 2
```

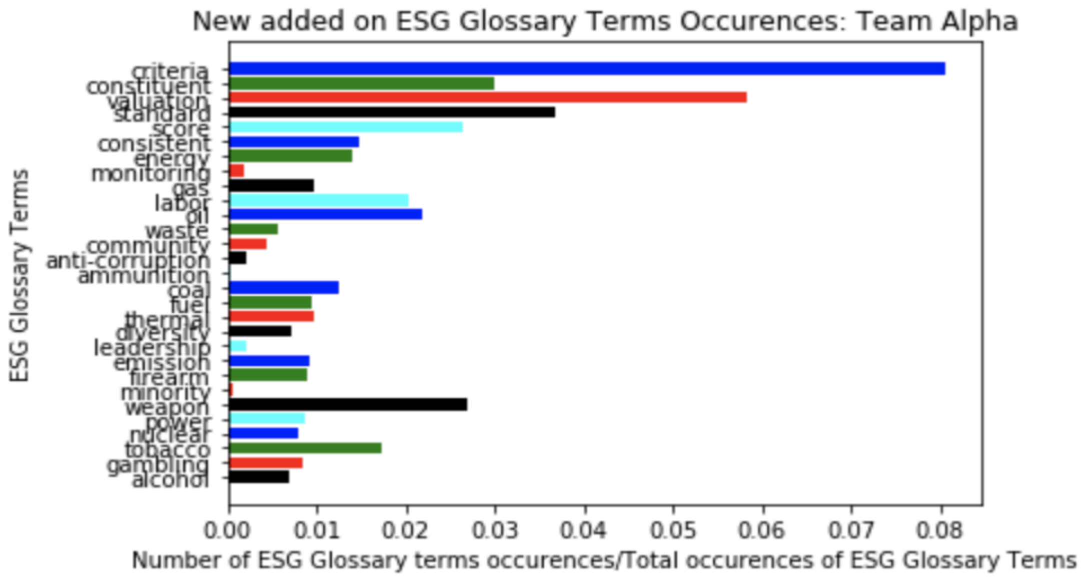


4 Empirical Visualization

Good Rating funds ESG Glossary Terms frequency(%) visualization



Bad Rating funds ESG Glossary Terms frequency(%) visualization



6 Empirical Method

ESG & Wiggle term Phrase Frequency in Good/Bad rating funds

- Both good and bad rating funds has similar phrase
- Some phrases are heavily used by bad rating funds
- Limitation: the small and unproportionate sample size, subjective selection

May not exhibit positive ESG characteristics	Good	26%	Bad	30%
Fund may underperform other funds that do not have an ESG focus	Good	21%	Bad	33%
ESG characteristics may not be the same companies selected by other index providers that use similar ESG screens	Good	12%	Bad	15%
ESG practices may shift into and out of favor	Good	6%	Bad	7%
ESG characteristics/performance may change over time	Good	12%	Bad	7%
This ESG policy/investment strategy/criteria may result in the Funds foregoing opportunities to buy certain securities	Good	35%	Bad	27%
These factors may include, but are not limited to (ESG factors)	Good	3%	Bad	5%
ESG information and scores across third party data providers, may be inaccurate or incomplete	Good	15%	Bad	10%
Investors may differ in their views of ESG characteristics	Good	0%	Bad	17%
The manager seeks to fully integrate ESG criteria into the stock selection	Good	18%	Bad	13%

4 BERT Model Analysis

BERT & DistillBERT Model

- ➢ BERT stands for Bidirectional Encoder Representatives from Transformers.
- ➢ DistillBERT: huggingface transformer model (smaller & faster than the BERT model)
- ➢ Objective: Apply DistillBERT to predict whether an ESG named fund is good or bad
  - ○ Random Shuffling(80% training, 20% testing)
  - ○ Apply tokenizer(input_ids, attention_mask)
  - ○ Fine-tune DistillBERT(define TrainingArgument, instantiate Trainer)
  - ○ Report Steps and Training Loss
  - ○ Make predictions

```
import torch

class capstone(torch.utils.data.Dataset):
    def __init__(self, data, tokenizer):
        self.data = data
        self.encodings = tokenizer([i['text'] for i in self.data], truncation=True, padding=False)

    def __getitem__(self, idx):
        item = {key: torch.tensor(val[idx]) for key, val in self.encodings.items()}
        item['labels'] = torch.tensor(1 if self.data[idx]['label']=='good' else 0)
        item['ticker'] = self.data[idx]['ticker']
        return item

    def __len__(self):
        return len(self.data)

train_dataset = capstone(train, tokenizer)
test_dataset = capstone(test, tokenizer)
```

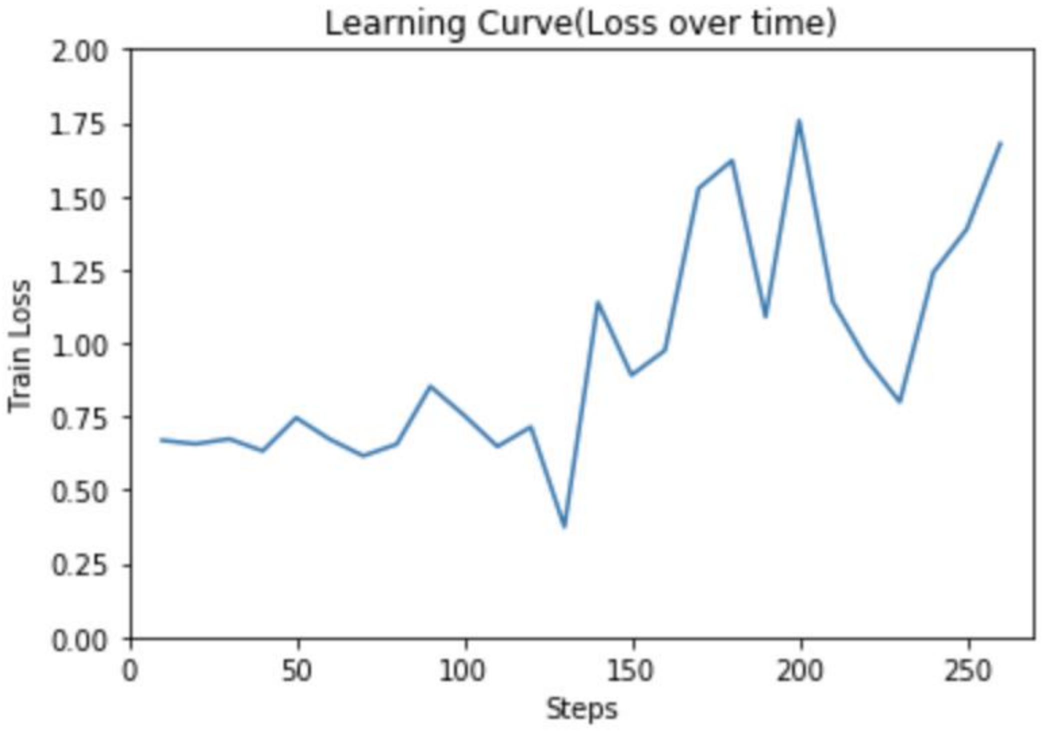
```
from transformers import DistilBertForSequenceClassification, Trainer, TrainingArguments

training_args = TrainingArguments(
    output_dir='./results',          # output directory
    num_train_epochs=10,              # total number of training epochs
    per_device_train_batch_size=1,  # batch size per device during training
    per_device_eval_batch_size=1,   # batch size for evaluation
    warmup_steps=500,                # number of warmup steps for learning rate scheduler
    weight_decay=0.01,               # strength of weight decay
    logging_dir='./logs',            # directory for storing logs
    logging_steps=10,
)

model = DistilBertForSequenceClassification.from_pretrained("distilbert-base-uncased")

trainer = Trainer(
    model=model,                         # the instantiated 🤗 Transformers model to be trained
    args=training_args,                  # training arguments, defined above
    train_dataset=train_dataset,         # training dataset
    eval_dataset=test_dataset             # evaluation dataset
)
```

```
from torch.utils.data import DataLoader
from transformers import DistilBertForSequenceClassification, AdamW


model.eval()

test_loader = DataLoader(test_dataset, batch_size=1)

for batch in test_loader:
    input_ids = batch['input_ids']
    attention_mask = batch['attention_mask']
    labels = batch['labels']
    outputs = model(input_ids, attention_mask=attention_mask, labels=labels)
    print(torch.nn.functional.softmax(outputs['logits'], dim=1)[0] , labels, batch['ticker'])
```

4 DistillBERT Model Analysis

Learning Curve

- ➢ Training Loss: Model Error or Mean Square Error
- ➢ Training Loss: The **lower**, the **better**
- ➢ Learning Curve: **Upward** Trend ----> Model **not learning**



DistillBERT Prediction

- ➢ fine tuning with native PyTorch on testing datasets
- ➢ prediction of good/bad funds presented as: probability of bad/good fund, true label of fund (0 = bad, 1 = good), fund ticker
- ➢ Conclusion: **Unable** to predicted accurately on good/bad funds based on prospectuses

```
tensor([0.9843, 0.0157], grad_fn=<SelectBackward0>) tensor([0]) ['IQSI']
tensor([0.9879, 0.0121], grad_fn=<SelectBackward0>) tensor([1]) ['USSG']
tensor([0.9868, 0.0132], grad_fn=<SelectBackward0>) tensor([1]) ['CACG']
tensor([0.9868, 0.0132], grad_fn=<SelectBackward0>) tensor([0]) ['NULV']
tensor([0.9881, 0.0119], grad_fn=<SelectBackward0>) tensor([0]) ['BRZIX']
tensor([0.9861, 0.0139], grad_fn=<SelectBackward0>) tensor([1]) ['IGIAX']
tensor([0.9771, 0.0229], grad_fn=<SelectBackward0>) tensor([0]) ['GSIFX']
tensor([0.9875, 0.0125], grad_fn=<SelectBackward0>) tensor([1]) ['TPYAX']
tensor([0.9879, 0.0121], grad_fn=<SelectBackward0>) tensor([0]) ['RAFE']
tensor([0.9886, 0.0114], grad_fn=<SelectBackward0>) tensor([0]) ['LEWIX']
tensor([0.9872, 0.0128], grad_fn=<SelectBackward0>) tensor([0]) ['LRGE']
tensor([0.9887, 0.0113], grad_fn=<SelectBackward0>) tensor([0]) ['LENIX']
tensor([0.9864, 0.0136], grad_fn=<SelectBackward0>) tensor([0]) ['ESML']
tensor([0.9872, 0.0128], grad_fn=<SelectBackward0>) tensor([0]) ['CFCV']
tensor([0.9744, 0.0256], grad_fn=<SelectBackward0>) tensor([0]) ['GAGVX']
tensor([0.9874, 0.0126], grad_fn=<SelectBackward0>) tensor([1]) ['SNPE']
tensor([0.9873, 0.0127], grad_fn=<SelectBackward0>) tensor([1]) ['BCAMX']
tensor([0.9878, 0.0122], grad_fn=<SelectBackward0>) tensor([0]) ['XJH']
tensor([0.9871, 0.0129], grad_fn=<SelectBackward0>) tensor([0]) ['ESGE']
```

7 Recommended Phrases

- **Detailed Disclosure:**

  “Securities eligible for inclusion in the Index comprise all constituents of the S&P 500 Index except for companies that: Generate 5% or greater of their revenue from thermal coal extraction or power generation, as determined by Sustainalytics; or Do not have (i) Sustainalytics coverage for tobacco-, controversial weapons- and thermal coal-related involvement; (ii) a UNGC score determined by Arabesque; or (iii) an S&P DJI ESG Score. ”

  (extracted from SPDR® S&P 500® ESG ETF prospectus)

- **Active in ESG Promotion:**

  “We use our voice as a shareowner to raise environmental, social, and governance issues with the management of select portfolio companies through a variety of channels. These may include engaging in dialogue with management, participating in shareholder proposal filings, voting proxies in accordance with our proxy voting guidelines, and participating in the annual shareholder meeting process.” (extracted from Boston Common ESG Impact International Fund prospectus)

7 Recommended Phrases

- **ESG Priority:**

  "Prior to considering financial information, the security selection process evaluates an issuer based on Environmental, Social and Corporate Governance (ESG) criteria." (extracted from DWS ESG Core Equity Fund prospectus)

- **Monitoring:**

  "The Adviser will monitor each holding on a regular basis to ensure its compliance with the Fund's guidelines." (extracted from Gabelli ESG Fund prospectus)

  "The portfolio is then monitored by the Adviser and Sub-Adviser and the weightings are adjusted regularly with a focus on each company's ESG Rating and Relative Value." (extracted from TrueShares ESG Active Opportunities ETF prospectus)

8

Conclusions

- The linguistic pattern of prospectus has a relatively **LOW CORRECTION** with its own accompanying ESG rating

- The **Limitations** of Our Study:

1. Small sample size
2. The disproportionation of ESG funds scored at each level
3. Subjectivity that could affect our finding

- Prospectus should not be considered as an **Only** criteria for discriminating greenwashing

THANKS

Team Alpha 2021.12.01