慶應義塾大学学術情報リポジトリ
Keio Associated Repository of Academic resouces

Title	ESG関連の株主提案が企業のESG活動に与える影響について
Sub Title	
Author	河東, 宗平(Kato, Shuhei) 林, 高樹(Hayashi, Takaki)
Publisher	慶應義塾大学大学院経営管理研究科
Publication year	2022
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2022年度経営学 第3960号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40003001-00002022-3960

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

慶應義塾大学大学院経営管理研究科修士課程

学位論文（　　2022　年度）

論文題名

ESG 関連の株主提案が企業の ESG 活動に与える影響について

主　査	林　高樹　　教授
副　査	高橋 大志　教授
副　査	武田 史子　教授
副　査	

氏　名	河東　宗平

論 文 要 旨

所属ゼミ	林高樹 研究会	氏名	河東 宗平

（論文題名）
ESG 関連の株主提案が企業の ESG 活動に与える影響について

（内容の要旨）

昨今、物言う株主と呼ばれるアクティビストの台頭が目立つ。あらゆるステークホルダーにとってアクティビストの存在は無視できない状況になってきている。
また ESG 投資を始めとした、ESG に関する企業や市場の注目が高まっており、機関投資家だけでなく一般消費者も企業の ESG 活動に関心を寄せる時代に突入した。

アクティビストが企業に対して働きかけをする際、株主総会にて株主提案文書（Shareholder Proposals）を提案する形をとる場合が多い。株主総会での決議の後一定の賛成を得られれば株主の正式な要求ということで、企業は提案内容にコミットする努力義務が発生する。投資家心理としては、投資した企業の価値を高めるため、より企業活動を活発にコミットさせる株主提案の内容を探ることがメリットとなり得る。本研究は、ESG に関する株主提案文書のテキスト情報が、企業の ESG 活動にどのような要素が影響を与えているのかを探る探索的な研究である。

これまで、株主提案に関する研究は多くされてきたが、ESG に関する株主提案文書自体に自然言語処理技術を用いテキスト情報を定量化し、ESG に関する企業の取り組みをターゲットとして設定した研究は管見の限りではまだ存在せず、その点は新規性と考える。
分析に使用したデータとしては、企業の ESG 活動については代理変数として KBS で契約をしている REFINITIVE EIKON の ESG スコアを、ESG に関する株主提案のテキスト情報については、量的な視点から米国証券取引委員会(SEC)のデータベースに掲載のある株主提案のテキストを使用した。テキストデータの情報は先行研究に倣い BERT にて自然言語処理して Specificity と Uniqueness の 2 つの指標に集約し、これらと業界や年次などをダミー変数として分析用データセットに含めた。

分析に使用した統計的方法論は ESG スコアを目的変数、その他を説明変数とする回帰分析である。回帰分析の結果は、Specificity（文書中における固有表現割合）の指標が有意で ESG スコアに対して正の相関があることが分かった。また BERTopic によるトピック分析も実施し、ESG に関する株主提案文書の内容に関するトピック抽出も試みた。

目次

1　はじめに
1.1　研究の背景
1.2　研究の目的
1.3　株主提案に関する規制
1.4　ESG に関する株主提案の現状
2　先行研究
2.1　ESG スコアを用いた研究について
2.2　株主提案文書を用いた研究について
2.3　自然言語処理技術について
2.3.1 固有表現抽出
2.3.2 文章ベクトルの取得
2.3.3 テキスト情報からトピック抽出
2.4　本研究の貢献について
3　仮説
4　研究の調査方法
4.1　使用するデータセット
4.1.1 Proxy Monitor について
4.1.2 REFINITIVE EIKON について
4.2　分析手法
4.2.1 リサーチデザイン
4.2.2 自然言語処理技術を応用した分析
4.2.2.1 BERT による文書ベクトルの取得
4.2.2.2 固有表現抽出技術を用いた分析
4.3　使用する変数のまとめ
5　分析結果
5.1　データの可視化について
5.2　重回帰分析の結果について
6　結論
6.1 まとめ
6.2 課題と限界そして今後の展望について
6.2.1 課題と展望について
6.2.2 限界について
謝辞
参考文献
付録

はじめに

1.1 研究の背景

株主価値経営とは、企業経営のひとつの考え方であり、企業価値を高めていくことによって株主価値の最大化を目指す経営である。上場企業において、株価は理論的には株主価値と等しくなると考えることもできるため、株主価値＝株価ととらえることもできる。つまり、株主価値経営は、企業の株価の最大化を目指す経営とも言える。もちろん企業には様々な形があり、多くのステークホルダーが関係している、それぞれにとってあるべき企業の姿があるため、決して株主価値経営が正解というわけではない。物言う株主と呼ばれるアクティビストの企業への提案が世界的に増えている背景には、記述した株主価値の最大化を目論んだ活動であることが一般的に理解されている。M&A Cloud によると、

> アクティビストは、投資先企業の一定割合以上の株式を保有することで、投資先企業の企業価値向上を大義名分とした株主権利行使を行います。具体的な権利行使としては、経営陣との対話・交渉、株主提案権の行使、各種提案（増配や自社株買いなどの株主還元策、非効率な事業売却や遊休資産の売却、経営陣刷新等)、場合によっては会社提案議案の否決に向けた委任状勧誘等が挙げられます。

とあり、アメリカにおいて先駆けて始まった活動である。またアメリカでは、ISS（Institutional Shareholder Services, Inc.）やグライスルイスなど議決権行使助言会社と呼ばれる、機関投資家に対して、株主総会における議決権行使について助言を行う機関が存在している。

近年日本でも、海外投資家が日本企業に対して、海外水準の株主権利を要求するようになり、アクティビストの存在感が増している。例えば、東芝の取締役選任が一部アクティビストによる選任や、香港系アクティビストによる東京ドームの買収危機なども記憶に新しい。

アクティビストの複数の活動において株主提案があげられるが、米国における ESG（環境・社会・企業統治）の視点で企業に経営などの改善を求める株主提案が増加している。米国において 2022 年の ESG 関連の株主提案数が過去最多となった。日経新聞によると、

> 米非営利団体（NPO）アズ・ユー・ソウによると、今年の米国での ESG 関連

の株主提案（4 月 21 日時点）は 583 件と 21 年（499 件）から 17%増えた。不平等の解消など「社会（S)」問題をテーマにした提案が 18%増の 304 件、「環境（E)」関連は 46%増の 161 件だった。「企業統治（G)」の提案は 44 件と 44%減った。

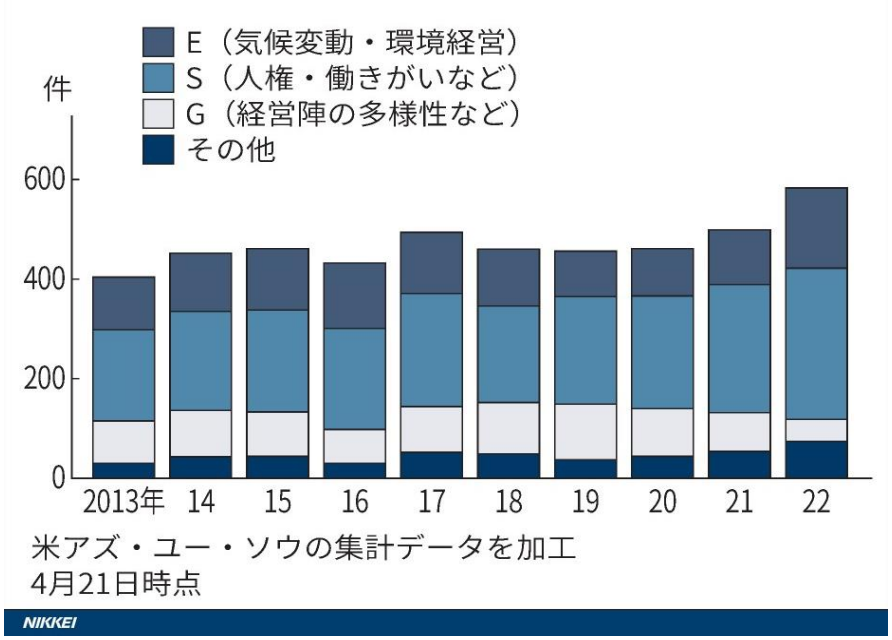


図 1　時系列でみた米国の ESG 関連の株主提案数
（日本経済新聞 2022 年 5 月 27 日）

とあり、ESG に関する株主提案が年々増加していることがわかる。今後 ESG に関する株主提案は世界的により重要な要素となっていくと考えられる。

湯山(2019)では、ESG についてこう説明されている。

ESG とは、環境（Environment)・社会（Society)・ガバナンス（Governance）の 3 つの頭文字をとったものである。ESG 要素を考慮した投資を「ESG 投資」といい、2006 年に国際連合の責任投資原則（PRI: Principles for Responsible Investment）の中で提唱された後、特に注目を集めてきた投資手法である。ESG はそれぞれ具体的にどのようなものかというと、例えば、E は地球温暖化対策、S は働き易さ、女性従業員の活躍、G は取締役構成などを示すものとされる。

1.2 研究の目的

本研究の目的は、「企業の ESG 活動に影響を与える、株主提案文書におけるテキスト情報の特徴を探索する」ことである。特徴が明らかになることで、企業価値を高める株主提案文書を提案できる可能性は高くなり、ひいては企業の株主価値最大化へ貢献出来うると考える。

株主提案という性質上、株主総会において株主の議決をとるプロセスを経る。その際に、提案内容がいかに企業へポジティブな影響を与えるのか、その点においてはテキスト情報を介して株主へ伝達することとなり、株主提案文書のテキスト情報は大きな役割をもつと言える。その意味では、株主提案文書のテキスト情報において、特徴が明らかになるベネフィットはアクティビストや株主らにとって大きいと考える。

1.3 株主提案に関する規制

一口に株主提案と言っても、各国によって株主総会における議決権を行使した株主提案の規制は異なる。今回は研究対象である、米国と比較として日本における違いを確認する。株式の保有要件については、日本では議決権の 1%以上または議決権の 300 個以上の株式が条件として必要となっている。米国では、議決権の 1%以上または市場価格 2000 ドル以上の株式が必要となっている。両国において、株式の保有要件は高くないハードルと考えられる。

一方で、株式提案における提案内容に関する規制は大きく異なる。日本では法令や定款に違反しない限り提案内容は強く制限されないが、米国では配当額を具体的に要求したり、取締役の選解任を求めたりすることはできない。議案の数についても日本において制限はないが、米国では 1 株主 1 議案までに限られている。

また日本の株主提案内容で最も多いのは経営陣の選任、解任となり、続いて定款変更である。これも日米の規制による違いが出ている。Yeh (2014)では、米国の株主は取締役会の解任なしに定款の変更ができないが、日本の株主は取締役会の同意なしに定款の変更ができる、と述べられている。

表 1 日米の株主提案の違い
（日本経済新聞 2021 年 5 月 28 日）

日米の株主提案の違い		
	日本	米国
株式の保有要件	議決権の1%以上または議決権300個以上の株式	議決権の1%以上または市場価格2000ドル以上の株式
認められない内容	・法律や定款違反 ・10%以上の賛成が得られなかった提案と実質的に同一の内容で、前回提案から3年たっていないもの	・取締役選解任 ・具体的な配当額に関する議案 ・過去5年以内に提出され、一定の賛成を得られなかった議案と同一の内容

NIKKEI

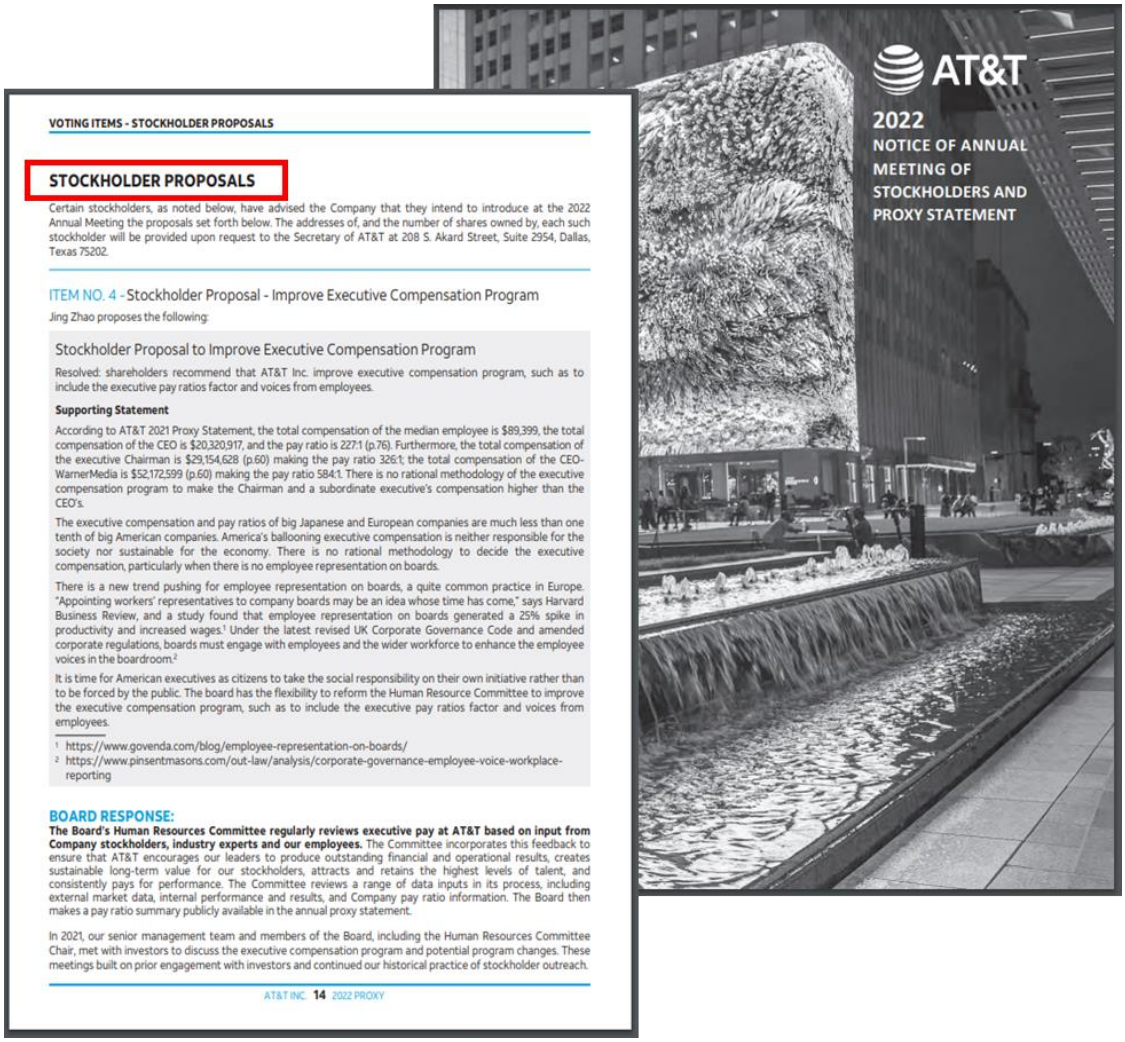
1.4 ESG に関する株主提案の現状

世界的にアクティビストの動きが活発になっている事は自明であるが、ESG に関する株主提案の現状はどうであろうか。米国においては、前述の ESG に関する株主提案数をみると 17%増となり、機関投資家が上場企業の ESG に対する監視の目を強めている事がわかる。

これまで ESG 投資の旗振り役として活動してきたブラックロック社に対し、一部関係者から「運用成績を犠牲にしている」との批判も出ている。

一方で、ウクライナ侵攻によって、ロシアでの事業活動を中止・撤退を求める動きなどの例にもみられるように、企業に社会的責任を問う声は一段と強まっており、今後もアクティビストによる株主提案は一層加熱する傾向にあると考える。

2022 年度の AT&T の株主提案文書を一例として下記示す。Notice of Annual meeting of Stockholders and Proxy Statement 内に Stockholder Proposals という項目があり、株主提案の内容が記載されている。



VOTING ITEMS - STOCKHOLDER PROPOSALS

STOCKHOLDER PROPOSALS

Certain stockholders, as noted below, have advised the Company that they intend to introduce at the 2022 Annual Meeting the proposals set forth below. The addresses of, and the number of shares owned by, each such stockholder will be provided upon request to the Secretary of AT&T at 208 S. Akard Street, Suite 2954, Dallas, Texas 75202.

ITEM NO. 4 - Stockholder Proposal - Improve Executive Compensation Program

Jing Zhao proposes the following:

Stockholder Proposal to Improve Executive Compensation Program

Resolved: shareholders recommend that AT&T Inc. improve executive compensation program, such as to include the executive pay ratios factor and voices from employees.

Supporting Statement

According to AT&T 2021 Proxy Statement, the total compensation of the median employee is $89,399, the total compensation of the CEO is $20,320,917, and the pay ratio is 227:1 (p.76). Furthermore, the total compensation of the executive Chairman is $29,154,628 (p.60) making the pay ratio 326:1; the total compensation of the CEO-WarnerMedia is $52,172,599 (p.60) making the pay ratio 584:1. There is no rational methodology of the executive compensation program to make the Chairman and a subordinate executive's compensation higher than the CEO's.

The executive compensation and pay ratios of big Japanese and European companies are much less than one tenth of big American companies. America's ballooning executive compensation is neither responsible for the society nor sustainable for the economy. There is no rational methodology to decide the executive compensation, particularly when there is no employee representation on boards.

There is a new trend pushing for employee representation on boards, a quite common practice in Europe. "Appointing workers' representatives to company boards may be an idea whose time has come," says Harvard Business Review, and a study found that employee representation on boards generated a 25% spike in productivity and increased wages.[1] Under the latest revised UK Corporate Governance Code and amended corporate regulations, boards must engage with employees and the wider workforce to enhance the employee voices in the boardroom.[2]

It is time for American executives as citizens to take the social responsibility on their own initiative rather than to be forced by the public. The board has the flexibility to reform the Human Resource Committee to improve the executive compensation program, such as to include the executive pay ratios factor and voices from employees.

[1] https://www.govenda.com/blog/employee-representation-on-boards/
[2] https://www.pinsentmasons.com/out-law/analysis/corporate-governance-employee-voice-workplace-reporting

BOARD RESPONSE:

The Board's Human Resources Committee regularly reviews executive pay at AT&T based on input from Company stockholders, industry experts and our employees. The Committee incorporates this feedback to ensure that AT&T encourages our leaders to produce outstanding financial and operational results, creates sustainable long-term value for our stockholders, attracts and retains the highest levels of talent, and consistently pays for performance. The Committee reviews a range of data inputs in its process, including external market data, internal performance and results, and Company pay ratio information. The Board then makes a pay ratio summary publicly available in the annual proxy statement.

In 2021, our senior management team and members of the Board, including the Human Resources Committee Chair, met with investors to discuss the executive compensation program and potential program changes. These meetings built on prior engagement with investors and continued our historical practice of stockholder outreach.

AT&T INC. 14 2022 PROXY

図 2　AT&T の株主提案例 [1]

[1] https://investors.att.com/~/media/Files/A/ATT-IR-V2/documents/att-2022-proxy.pdf

2. 先行研究

2.1 ESG スコアを用いた研究について

ESG スコアを活用した研究は、SDGs 達成や株式流動性との関係性や、投資パフォーマンスとの関係性などの研究がされている。また、株主提案文書を用いた研究は、米国において株主提案と市場価値に関連性があるのかの検討や、国内における株主提案が行われやすい因子の探索の研究がされている。

Sasaki(2020)は、SDGs 達成と ESG スコアは高度に相関していることを示しており、ESG スコアは SDGs の達成を予想する指標になることを示した。

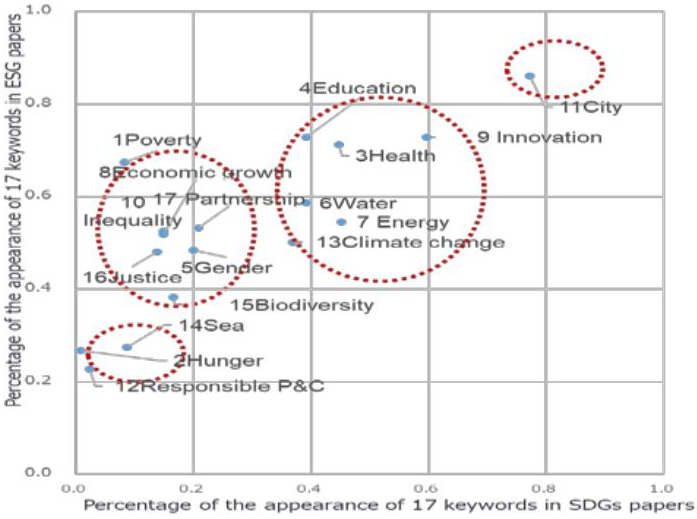


図 3　Percentage of the appearance of 17 keywords in SDGs and ESG research（Sasaki2020 より抜粋）

青木（2016)によると、欧米を中心に ESG 情報を含む非財務情報の開示が義務付けられており、日本でも 2014 年に金融庁が日本版スチュワードシップ・コード（機関投資家に向けた行動規範）を公表するなど、中長期的な企業価値の向上を図るにはこういった取り組みが不可欠であったと述べられている。

Armuji（2016)は、より具体的な内容に踏み込み、ESG 活動の取組が企業の経済的なパフォーマンスにどう影響しているか、マレーシアとシンガポールの企業を対象に ESG スコアを用い両国の差異を研究している。畠田(2021)は、ESG スコアが高い場合には株式流動性が高いことが示されているが、一方で ESG スコアと投資パフォーマンスが相関しているわけではないことも報告している。

Kim (2014)は、企業の社会的責任（CSR）とコーポレートガバナンスがその企業の信用格付けに与える影響について研究している。説明変数として、CSRの指標であるCGIスコア（Corporate governance index）や、コーポレートガバナンスに関する指標等を用い、目的変数として企業の信用格付けレートを用いている。回帰分析の結果としては、両因子が信用格付けレートに有意な影響を及ぼしていることが明らかとなり、また両回帰係数の符号は正であった。似た視点として林(2018)は、企業の情報開示がESGスコアに影響を及ぼすという研究を実施している。企業からの開示情報のみに基づいて行うESG評価では、ESG評価を正しく行うことが困難である。そのため広範な規制が必要であると述べている。

2.2 株主提案文書を使用した研究について

Cunat(2010)によると、米国における株主提案文書に関する研究では、企業ガバナンスに関わる株主提案がなされた場合、短期の市場リターンはプラスになるとされている。一方、Gordon(1993) では、機関投資家が株主提案を提出するということは経営陣との事前交渉がうまくいっていないと判断し、市場の短期リターンがマイナスという結果になっていると述べられている。

Walkling (1996)は、米国の企業において、株主提案の数が多い企業は、企業規模が大きく、また機関投資家の保有比率が高く、財務パフォーマンスが低いことを示している。

Yeh (2014)は、日本の上場企業の株主提案を受けた135社を対象に、株主総会決議が、企業のパフォーマンスや価値を向上できるのか検証した。新しい示唆として、懲戒的な性質の株主提案が経営陣の堅固な姿勢を軽減し、企業価値を改善することを明らかにした。また、ROA（総資産利益率）が低く、企業内部者の持株比率が低く外国人株主の比率が高い企業ほど株主提案がなされることを、示している。またYeh (2017) では、2004-2013年の日本における株主提案を分析し、法的に強力な議決権の行使が経営に影響を与えるかを研究しており、株主提案の種類は企業特性によって異なり、大株主による議決権行使は企業にプラスの影響を与えることを示している。

久多里 (2018)は、日本国内の株主提案について傾向を分析しており、日本における株主提案数は増加傾向にあり、提案内容は定款の変更が大部分で、それ以外の内容は多岐にわたっていると述べている。

Stuart（2022）は、日本企業においてアクティビストと株主への反応を分析している。市場の反応はアクティビストによって異なり、株主投票反対はアクティビストの存在に依存していることを示している。

2.3 自然言語処理技術について

2.3.1 固有表現抽出

まず自然言語処理とは私達が日常書いたり喋ったりしている日本語や英語どの言語のことであり、自然言語処理は、この自然言語を処理する技術の総称である。

固有表現抽出（Names Entity Recognition）とは、情報抽出といわれる分野の中でも基礎的技術の一つとされている。情報抽出・情報検索のワークショップである IREX(Information Retrieval and Extraction Exercise)では以下の固有表現が定義されている。

表 2 IREX による固有表現の定義

種類		抽出対象
ORGANIZATION	組織名	会社名や団体名といった組織の名称
PERSON	人名	役職等も含む個人を表す名前や呼称
LOCATION	地名	都道府県名や観光スポットなど場所を表す名称
DATE	日付	具体的な年月日や、昨日／今日／明日といった相対的な表現も含む日付を表す名称
TIME	時間	日付表現と同様に時間を表す名称
MONEY	金額	1 万円、4,000 円など金額を表す表現
PERCENT	割合	1/2、50%、2 割など割合を表す表現

内元ら(2000)によると、固有表現を抽出する方法には大きく分けると、人手で作成した規則に基づく方法と学習に基づく方法があるとされている。金野(2020)でも述べられているが、前者のやり方だとマニュアルで固有の表現に対してラベルを割り当てる必要があり、膨大な作業工数が必要となる。また新たな固有表現や語彙が現れる度に更新しなくてはならないため、現実的ではない。そのため、サポートベクトルマシーンを利用した研究や、深層学習を用いた研究なども進んでいる。

2.3.2 文章ベクトルの取得

テキスト情報から文書ベクトルを作成するにあたり、本研究では BERT（Bidirectional Encoder Representations from Transformers）を用いた。BERT は Devlin(2019)によって紹介されたモデルとなり、Transformer をベースとしたモデルである。BERT モデルの特徴として、注目する単語の前後の文脈の両方（双方向）について学習している点にあると青嶋(2019)で述べられている。

BERT は双方向の Transformer を 12 段接続し、2 種類の事前学習を行い、その後に様々なタスクに合わせたファインチューニングを行うことが特徴である。

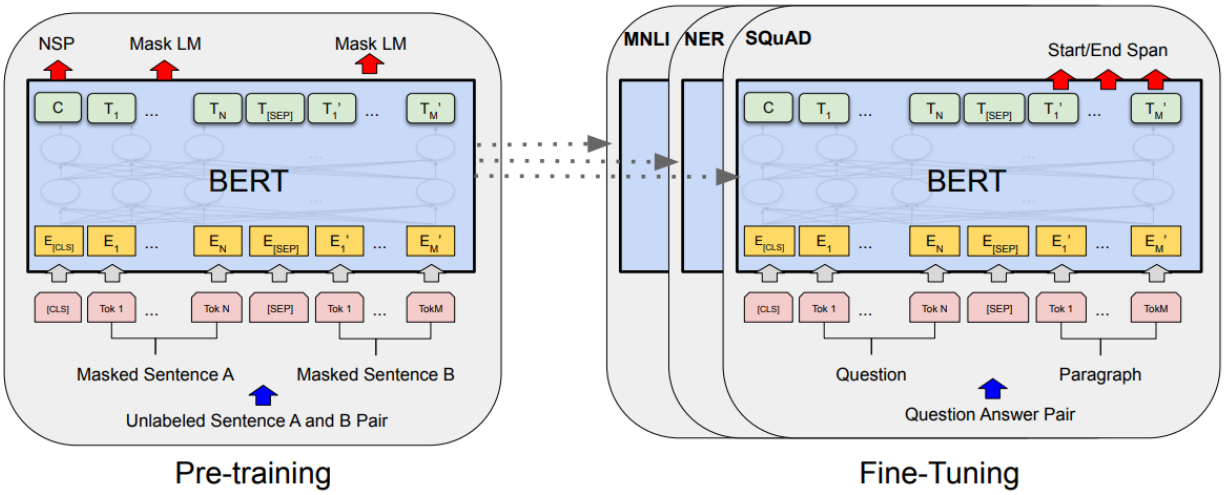


図 4　BERT のモデル概要（Delvin, 2019）より抜粋

BERT で出来る事の一つとして、文脈を考慮した文書ベクトルが獲得できる点が挙げられる。金野(2021)は、有価証券報告書のテキストデータに自然言語処理技術を用いて「Specificity」「Stickiness」「Uniqueness」の 3 つの説明変数となる指標を作成し、目的変数に株価データを使用し、両者の反応を検証した。

2.3.3 テキスト情報からトピック抽出

トピック分析の先行研究について、Grootendorst(2022)は、文書の潜在的なトピックを発見するのにトピックモデルは有用なツールであると述べており、

また最近の研究では、トピックモデルをクラスタリングタスクとしてアプローチすることの可能性を示している。

加藤(2021)は、分散表現を用いたトピック分析である BERTopic を用い、二つのクオリティペーパーにおける肥満に関する記事を対象とし、異なる二紙間、及び異なる時点間の類似性や変化を分析している。図 7 のように、トピックの数とトピックに属する主要単語が抽出される。

成（2021）は、新型コロナ不満アンケートデータに対して BERTopic にて時系列に分けトピック分析を行い、コロナ禍における人々の不満の特徴を調べた。

表 3　トピックモデルによるクラス分類の例（加藤 2021 より抜粋）

番号	トピックに属する主要単語	新聞
0	sugar_drinks_sugary_cocacola	Guardian
1	fat_blame_weight_stigma	NewYorkTimes
2	fat_eating_duchess_shriver	Guardian
3	drug_gene_genetic_pill	Guardian
4	pupils_sport_childhood_children	Guardian
5	schools_milk_vending_lunch	NewYorkTimes
6	uk_nhs_years_obese	Guardian
7	genes_fat_brain_hormone	NewYorkTimes
8	diabetes_disease_dementia_insulin	NewYorkTimes
9	ads_food_ofcom_watchdog	Guardian
10	food_americans_kfc_eating	NewYorkTimes
11	soda_tax_bloomberg_mayor	NewYorkTimes
12	smoking_government_nhs_alcohol	Guardian
13	children_kids_child_parents	NewYorkTimes
14	cancer_cancers_smoking_breast	Guardian
15	birth_babies_pregnancy_breastfed	NewYorkTimes
16	cyclists_healthy_cities_buildings	Guardian
17	takeaways_schools_takeaway_shops	Guardian
18	industry_manufactures_asda_advertising	Guardian
19	sugery_gastric_disability_bariatric	Guardian

2.4 本研究の貢献について

前述した通り、これまで株主提案やESGに関する先行研究はあるものの、ファイナンス分野の他の研究テーマに比較して相対的に新しいテーマと言える[1]。

管見の限りでは、ESGスコアとESGに関連する株式提案文書のテキスト情報の双方を同時に取り上げてその関係性を調べた研究はまだない。従って、ESGに関する株主提案文書のテキストデータに対して、自然言語処理技術（BERT）を適用することで含まれる情報を定量化し、ESGに関する企業の取り組みとの関係性を調査した点は新規性があると考える。

今日のESGに対する社会的関心の高さ、また、アクティビスト等の株主の経営への積極的な関与の社会的動向を踏まえると、今後ますますESGに関する株主提案の重要性は増していくと考えられる。先行研究が希少な分野における探索研究を行うことにより得られる実証的知見は、学術的、社会的にも意義のあるものとなることが期待される。

---

[1] 本論文の執筆時点2月22日時点におけるGoogle Scholarの検索結果：″shareholder proposals″ AND ″ESG rating″のように両者を組み合わせた研究はGoogle Scholarで検索してみると、237件ヒットし、被引用数が30本以上の論文は5本となる。″shareholder proposals″ AND ″ESG score″だと179件ヒットし、被引用数が30本以上の論文は3本となる。

3. 仮説

改めて記述すると、研究の目的は「企業の ESG 活動に影響を与える、ESG に関する株主提案文書におけるテキスト情報の特徴を探索する」ことである。また詳しくは研究の調査方法の項で詳細を記述するが、企業の ESG 活動の代理変数として ESG スコアを用いる。

本研究は探索的研究であるが、ESG に関する株主提案文書と企業の ESG 活動の関係性は、株主の意向の総意が株主提案となるため、ESG に関する株主提案の場合、企業の ESG 活動に影響を与えることが予想される。すなわち、初期仮説として、テキスト情報から自然言語処理で指標化した変数と ESG スコアにおいて反応がある、言い換えると、上記指標化した変数を説明変数、ESG スコアを目的変数とする回帰モデルにおいて、これらの回帰係数が有意であると予想される。

また今回は探索的な研究と位置づけ、目的変数と説明変数の反応だけでなく、株主提案文書内容についてもテキスト情報をトピック分析し特徴を探る。

4. 研究の調査方法

4.1 使用するデータセット

今回使用するデータとして、大きく 2 種類のデータがある。1 つ目は、米国における ESG に関する株主提案のテキストデータである。2 つ目は、企業の ESG に関する取組を代理変数とした ESG スコアである。

米国における ESG に関する株主提案のテキストデータは、Web サイトの Proxy Monitor より、2006 年度から 2021 年度の 16 年分収集した。

企業の ESG に関する取組を代理変数としての ESG スコアは、REFINITIVE EIKON より 2006 年度から 2021 年度の 16 年分収集した。またその際、ESG スコアだけでなく、企業毎の売上、従業員数、業界データも同時に収集した。最終的に、784 行のデータセットとなっている。

表 4　使用するデータセット概要（著者作成）

データ	内容	形式
複数年分の ESG 関連の株主提案文書	2006 - 2021 年「Proxy Monitor」に掲載されている企業の株主提案文書 ※カテゴリーは「Corporate Governance」「Social Policy」を選択	テキスト

複数年分のESGスコア等	REFINITIVE EIKONデータベース内のESGのスコアリング情報等（売上、従業員数、業界）	数値データ

4.1.1 Proxy Monitorについて

Proxy Monitor [1]のデータベースは、2006年から2022年までの年次株主総会で公開された、米国の大手上場企業 250 社の株主が提出した委任状提案のデータベースである。

同データベースは、年、会社、業界、提案者、提案のタイプなどで検索することができ、今回は「Corporate Governance」「Social Policy」を選択し株主提案文書をウェブ上から入手した。

Proxy Monitorのデータベース上に表示されているリンクをクリックすると、各株主提案が記載された実際の委任状データに遷移することができる。それらのオリジナルのデータベースは米国証券取引委員会（Securities and Exchange Commission）のものとなる。

また具体的入手方法については、米国証券取引委員会のオリジナルデータへ1社毎に遷移し、委任状の中で株主提案のテキストデータ部分のみを抽出した。米国証券取引委員会のオリジナルデータのHTMLの構造が、全てpタグもしくはbタグにて羅列式で記述しているため、下記プログラムの設計にて年代毎にスクレイピングを実施した。（ガイドラインに基づき 1 秒間に 10 リクエスト以下のアクセス規約を遵守）

Table of Contents（目次）内に、大文字小文字問わず次の用語（Shareholder Proposal(s), Stockholder proposal(s)）がある場合その項をスクレイピングするスタート地点と設定し、終了地点は目次内の次の項とする。キーワードの設定が出来た後、本文のテキスト情報を読み込み、タイトル部分でスターティングキーワードが出てきた時、テキストデータの抽出を開始し、エンディングキーワードが出てきた時にテキストデータの抽出を終了する。

上記課題点としては、そもそもTable of Contentsがない場合、キーワードの設定ができないため抽出はできない。また、Table of Contentsの項目名と本文中のタイトル名が異なる場合、キーワードが合致したいため抽出ができない。この点は今後スクレイピングの設計を改善することで、精度をあげる余地が残っている。

[1] https://www.proxymonitor.org/

4.1.2 REFINITIVE EIKON について

今回企業の ESG に関する取組を示す指標として、Refinitiv 社の ESG スコアを代理変数として用いる。ESG スコアを提供している機関として、REFINITIVE 以外に Bloomberg 社、MSCI 社があげられるが、REFINITIVE 社の ESG スコアデータを用いる理由は、既に慶応義塾大学経営管理研究科が REFINITIVE 社と契約をしており、一定の条件の元、同データにアクセスできるためである。

Refinitiv 社は LSEG（ロンドン証券取引所）のグループ企業として、金融市場のデータとインフラストラクチャを提供する世界有数のプロバイダーである。Refinitiv ESG スコアデータは、公開情報源（年次報告書、NGO ウェブサイト、CSR レポート等）から標準化されている数値データとなる。

Refinitiv ESG スコアデータで使用している指標について、Governance ピラーの中に「Shareholders」カテゴリーの項目があり、同カテゴリーの中で定義に「shareholder proposal（株主提案）」含む項目が以下の 3 つあるが、これらはいずれも ESG スコアの算出には用いない項目であることを Refinitiv 社に確認済みである。つまり、株主提案の有無や内容は Refinitiv 社の ESG のスコアには直接的には影響しない。

- ✓ TR.PolicyShareholderEngagement
- ✓ TR.AdvanceNoticeShareholderProposal
- ✓ TR.AdvanceNoticePeriod

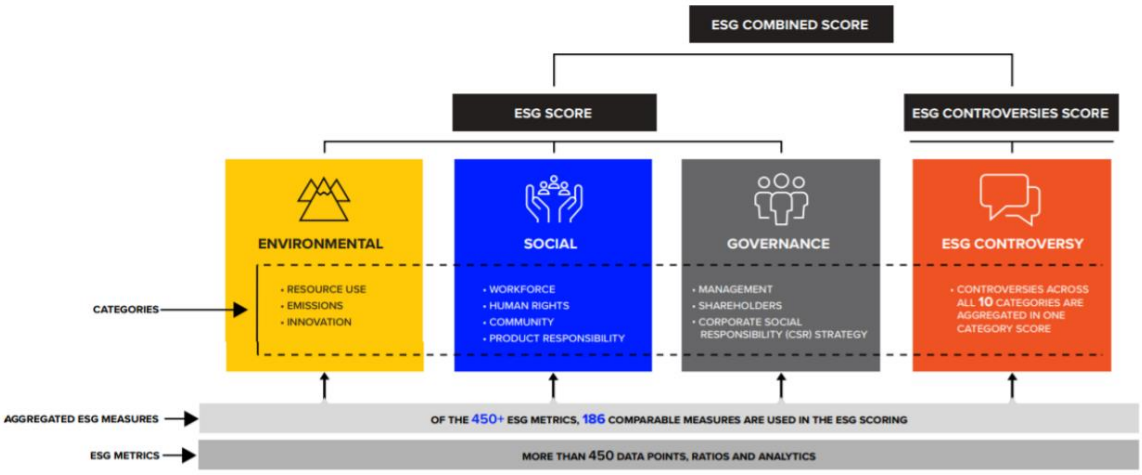


図 5　ESG スコアの概要
（REFINITIVE EIKO より引用）

4.2 分析手法

4.2.1 リサーチデザイン

本研究の分析は大きく、次の 2 ステップに分かれる。

1. 自然言語処理技術を用いる事で ESG に関する米国株主提案のテキスト情報を定量化し、説明変数を作成する。
2. 前のステップで作成した変数と企業規模や業種など企業の特徴を表す変数用いて回帰分析を行い、テキスト情報と企業の ESG 活動の反応の関係性を分析する。

テキスト情報定量化についてであるが、本研究では金野(2021)でも採用された Specificity と Uniqueness の 2 つの指標を採用することにする。その理由は以下である。

Specificity は文書中の固有表現の割合であり、Specificity の割合が高いということは、該当の企業にとってより当事者意識が高く、具体的な内容が多いということであり、株主の賛同を得られやすく企業活動に影響を与えやすい文書であると考える。また Uniqueness は、テキストデータにおける業界平均からの距離の指標となり、業界の傾向から近しいのか、遠く独特な内容なのか、その点は株主の賛同に影響を与え、企業活動に影響を与える文書の要素であると考える。

並行し、ESG に関する株主提案文書のコンテンツについてもトピック分析を実施する。ステップ 1.2 を図で表現したのが図 6 である。

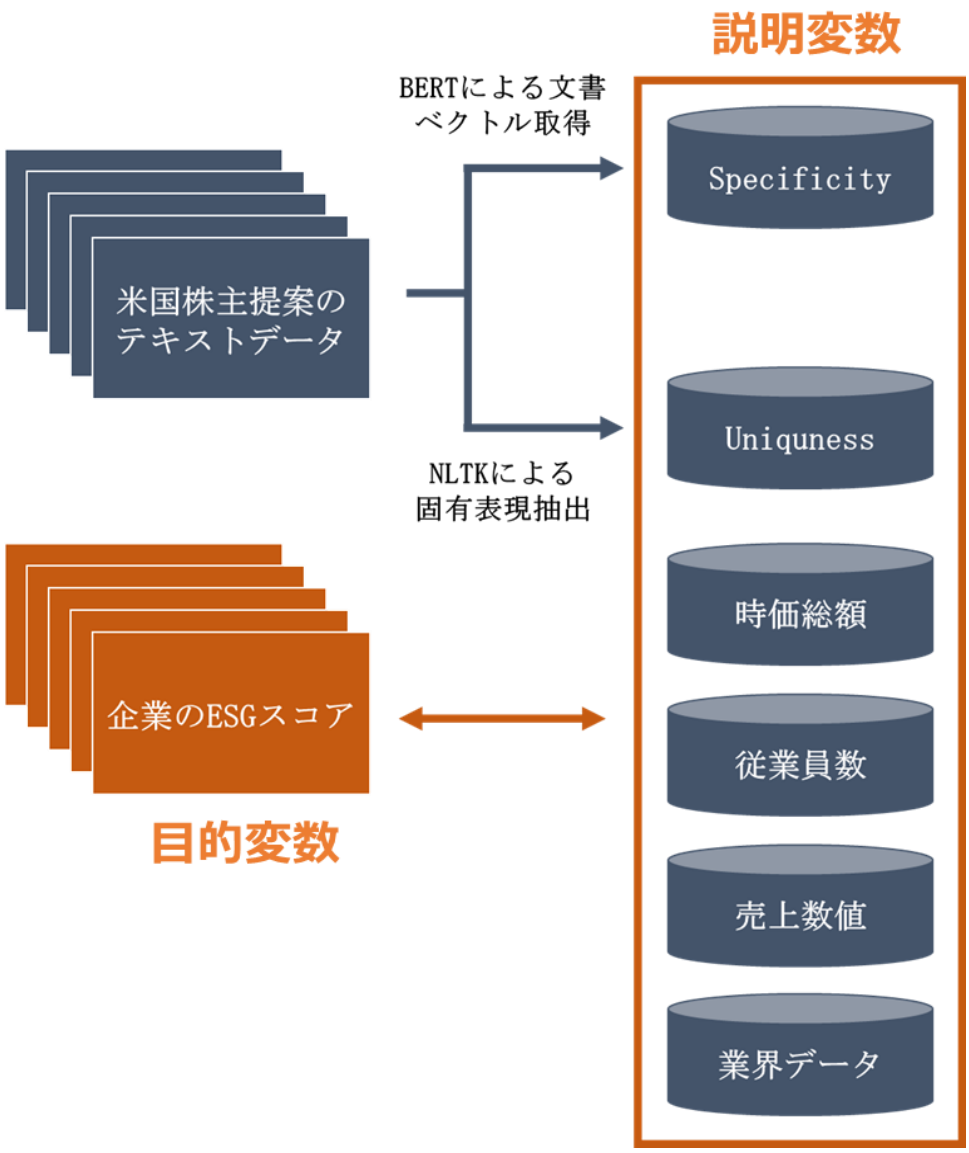


図 6　1-2 におけるリサーチデザインの概要

上記ステップ前のデータ整備の部分を簡単に記載する。Proxy Monitor よりスクレイピングしたテキスト情報と REFINITIVE EIKON からダウンロードした企業毎の ESG スコアを結合する作業が分析前のデータ整備として必要である。今回は、Excel にて vlookup 関数を用い両者を結合した。企業名にて完全一致にて合致させたため、一部欠落しているデータがある。

両データの結合後、株主提案のされていない企業や年度の欠損値を削除し、また回帰分析後の係数を比較するため Mkt_Value, Employee, Sales, Specificity の 4 つの変数を対数化し、その後 Mkt_Value, Employee, Sales, Specificity, Uniqueness の 5 つの変数を正規化した。

業界と株主提案が実施された年次データについては、業界や年次による反応の違いをコントロールするために、ダミー変数を導入した。

4.2.2 自然言語処理技術を応用した分析

4.2.2.1 BERT による文書ベクトルの取得

株主提案文書は、業種ごとにリーダー企業をベンチマークとして記載される事が考えられる。そのため、業界内で同内容の株主提案が提出されるなどしばしばある。そこで、各企業の株主提案の文書ベクトルを算出した後、業種ごとに文書ベクトルの平均ベクトルを計算する。この平均ベクトルが、業界内での典型的な記載内容であると仮定し、各企業の文書ベクトルと平均ベクトルの距離を計算する。その文書が業界内の典型的な株主提案からどれだけ離れているか、つまりどれだけ独自性があるかを表す指標として、Uniqueness が金野(2020)で定義されている。本研究では、ある業種に所属する企業 i の t 年の文書ベクトルを以下表す。

$$x_{i,t} = \left(x_1, x_{2,\ldots,} x_{384}\right)^t$$

その業種の t 年の平均ベクトルは以下のように表す。

$$\hat{x}_{i,t} = \left(\hat{x}_1, \hat{x}_{2,\ldots,} \hat{x}_{384}\right)^t$$

この時、Uniqueness（i,t）は以下のように計算される。

$$Uniqueness_{i.t} = 1 - \left(\frac{x_{i,t} \cdot \bar{x}_t}{\|x_{i,t}\| \|\bar{x}_t\|}\right) \tag{1}$$

4.2.2.2 固有表現抽出技術を用いた分析

先行研究でもふれたが、Specificity の指標を計算するために固有表現抽出技術を用いる。固有表現抽出を行うにあたって、Python で使用可能なオープンライブラリ spaCy の English モデル en_core_web_sm-3.4.1 を使用した。このライブラリよる固有表現抽出で抽出できる固有表現のカテゴリーは以下である。

表 5　固有表現のカテゴリー表 [1]

TYPE	DESCRIPTION
PERSON	People, including fictional.
NORP	Nationalities or religious or political groups.
FAC (FACILITY)	Buildings, airports, highways, bridges, etc.
ORG (ORGANIZATION)	Companies, agencies, institutions, etc.
GPE	Countries, cities, states.
LOC (LOCATION)	Non-GPE locations, mountain ranges, bodies of water.
PRODUCT	Vehicles, weapons, foods, etc. (Not services)
EVENT	Named hurricanes, battles, wars, sports events, etc.
WORK_OF_ART	Titles of books, songs, etc.
LAW	Named documents made into laws.
LANGUAGE	Any named language.
DATE	Absolute or relative dates or periods.
TIME	Times smaller than a day.
PERCENT	Percentage, including “%”.
MONEY	Monetary values, including unit.
QUANTITY	Measurements, as of weight or distance.
ORDINAL	“first”, “second”, etc.
CARDINAL	Numerals that do not fall under another type.

[1] https://catalog.ldc.upenn.edu/docs/LDC2013T19/OntoNotes-Release-5.0.pdf より抜粋

次に、このライブラリを Python で使用し、テキスト情報の一部を固有表現抽出した例を次に示す。

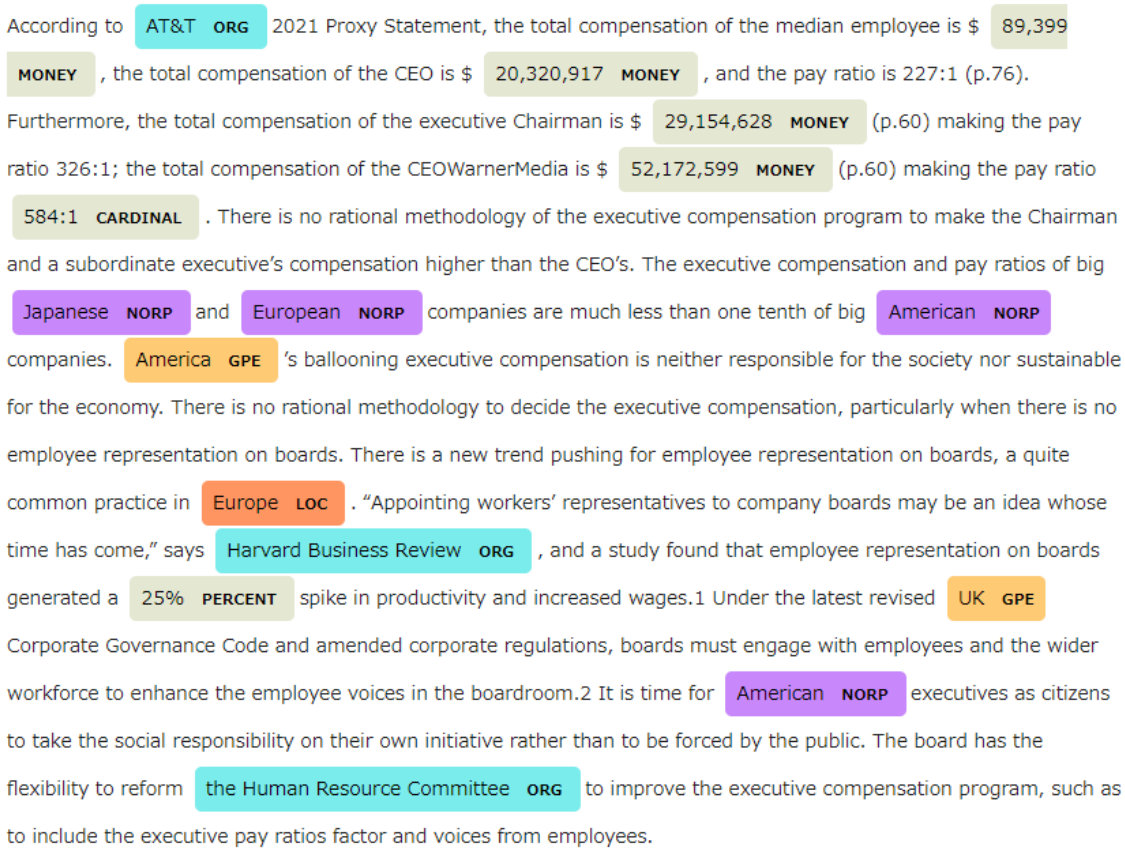


図 7　spaCy による固有表現抽出イメージ

概ね正しく抽出できている事が確認できる。この抽出によって、文書に含まれている全単語数と固有な表現の単語数を得ることができる。その結果を用いて金野(2020)で定義されている、Specificity を用いる。

ある企業 i の t 年度の文書の総単語数が Total(i, t)、総固有表現数が Ent(i, t)の時、Specificity(i, t)は以下のように計算する。

$$Specificity_{i.t} = \frac{Ent_{i,t}}{Total_{i,t}} \tag{2}$$

4.3 使用する変数のまとめ

株主提案のテキストデータから作成した 2 つの指標と、16 分類の年次データ及び 12 種の業種データをダミー変数として加え、31 個の説明変数を持つ加工データが完成した。

記号	被説明変数	補足
	ESG スコア	

説明変数	ソースデータ内のデータ項目	変数の説明
X1	Mkt_Value	時価総額
X2	Employee	従業員数
X3	Sales	売上規模
X4	Specificity	固有表現割合
X5	Uniqueness	独自性（1-コサイン類似度）
X6	Administrative and Support and Waste Management and Remediation Services	業界のダミー変数
X7	Finance and Insurance	業界のダミー変数
X8	Health Care and Social Assistance	業界のダミー変数
X9	Information	業界のダミー変数
X10	Manufacturing	業界のダミー変数
X11	Mining, Quarrying, and Oil and Gas Extraction	業界のダミー変数
X12	Professional, Scientific, and Technical Services	業界のダミー変数
X13	Retail Trade	業界のダミー変数
X14	Transportation and Warehousing	業界のダミー変数
X15	Utilities	業界のダミー変数
X16	Wholesale Trade	業界のダミー変数
X17	2007	年次のダミー変数
X18	2008	年次のダミー変数
X19	2009	年次のダミー変数
X20	2010	年次のダミー変数
X21	2011	年次のダミー変数

X22	2012	年次のダミー変数
X23	2013	年次のダミー変数
X24	2014	年次のダミー変数
X25	2015	年次のダミー変数
X26	2016	年次のダミー変数
X27	2017	年次のダミー変数
X28	2018	年次のダミー変数
X29	2019	年次のダミー変数
X30	2020	年次のダミー変数
X31	2021	年次のダミー変数

ダミー変数を作成するにあたり、ベースラインとなる変数があるため上記表は実際の変数数よりマイナス 1 の数となる。今回の場合は、年次は「2006 年」で業界は「Accommodation and Food Services」がベースラインにあたる。

5. 分析結果

5.1 データの可視化について

まず各変数についての基本統計量の確認をする。その際、次項に対数変換と正規化の前処理を実施する前後にて統計量の詳細を示す。

データ個数は共に、784 個となる。Mkt_Value、Employee、Sales、Specificity に関してはヒストグラムの示す通り、サンプルが偏っているため、本研究では対数変換を行うこととする。また変数間のスケールが大きく異なるため、スケールを合わせるため、正規化を行った。

続いて各変数について、基本統計量を確認する。Year は株主提案文書が提出された年となり、2006 年から 2021 年までの幅となる。ESG_Score は会見年度末におけるその年の ESG_Score となる。最大値が 94.9、最小値が 8.4 となり、平均は 62.5 となっている。続いて、テキスト情報から作成した Specificity は、最大値が 0.53、最小値が 0 となっており、平均値は 0.09 であった。平均値は陸(2021)の中で実施されていた統合報告書の固有表現割合の数値とも概ね近い値となっている。

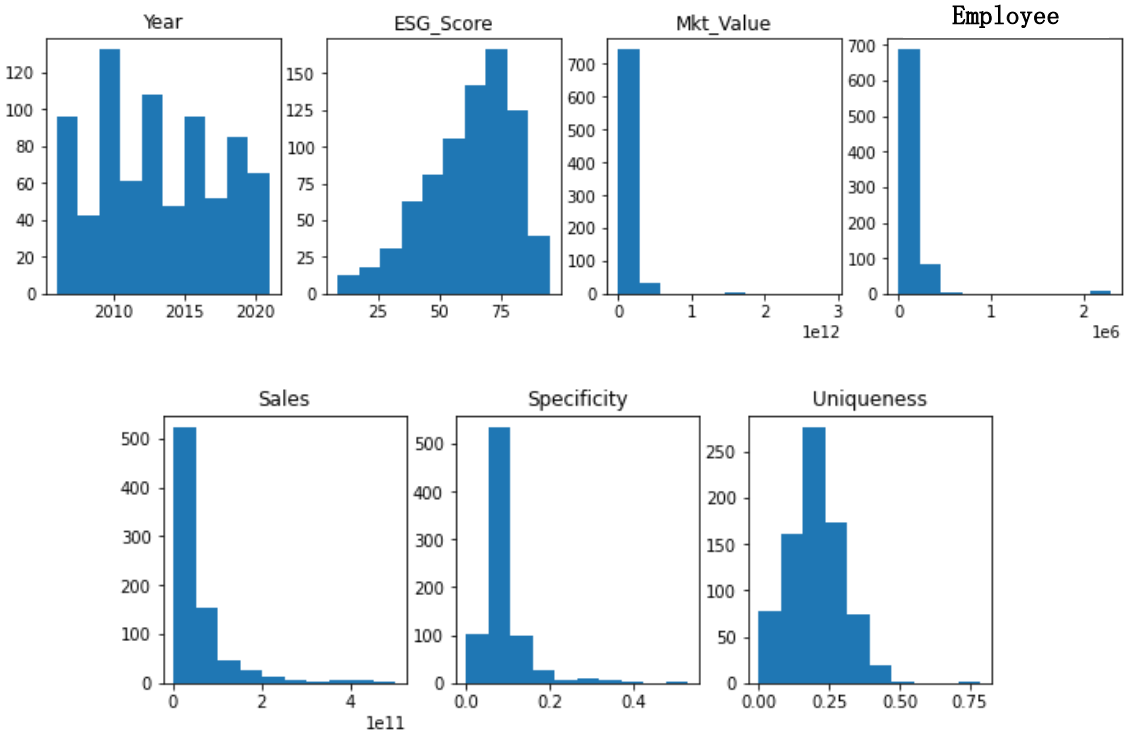


図 8　データセットの概要（データ前処理前）

	Year	ESG_Score	Mkt_Value	Employee	Sales
count	784.000000	784.000000	7.840000e+02	7.840000e+02	7.840000e+02
mean	2012.956633	62.458366	9.742475e+10	1.186434e+05	5.595771e+10
std	4.354139	17.780844	1.891881e+11	2.428435e+05	7.460502e+10
min	2006.000000	8.414749	6.030055e+08	2.075000e+03	2.198000e+09
25%	2009.000000	50.860095	1.999977e+10	2.765500e+04	1.401375e+10
50%	2013.000000	66.422578	4.238071e+10	5.716900e+04	2.492800e+10
75%	2017.000000	76.089280	1.222448e+11	1.335000e+05	7.040050e+10
max	2021.000000	94.931304	2.901645e+12	2.300000e+06	5.003430e+11

	Specificity	Uniqueness
count	784.000000	7.840000e+02
mean	0.090295	2.040791e-01
std	0.054690	9.698025e-02
min	0.000000	-1.192093e-07
25%	0.065524	1.451744e-01
50%	0.082562	1.963122e-01
75%	0.101271	2.626064e-01
max	0.530303	7.884542e-01

図 9　基本統計量の概要（データ前処理前）

続いて、本データセットにおける年次及び業界に区切り特徴を確認する。株主提案が提出された ESG スコアを持つ企業数を年次別でみた数は年によって若干のばらつきはあるものの、概ね一定であることが分かる。

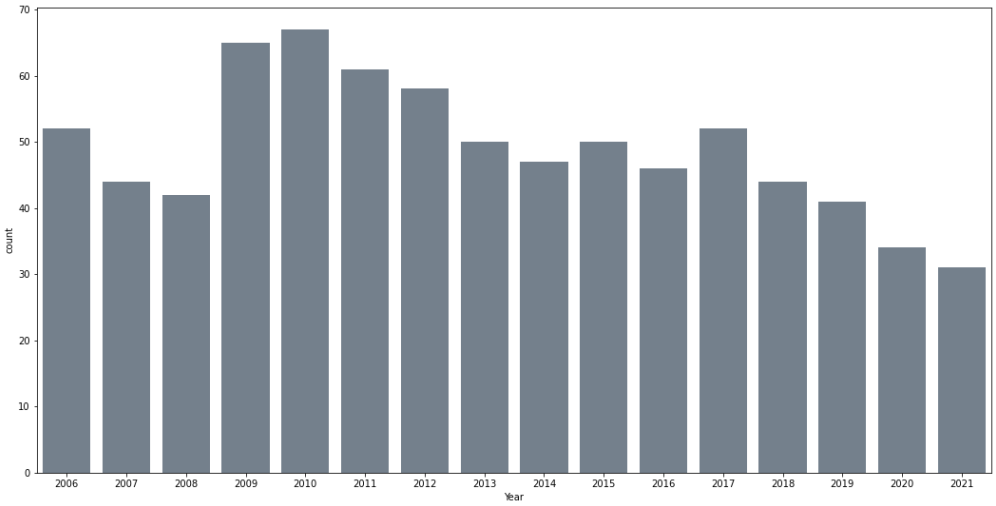


図 10　2006 年から 2021 年における株主提案がされた ESG スコアを持つ企業数

次に年次別に、各社の ESG スコアの分布を箱ひげ図にて確認すると、年を経るごとに ESG スコアの平均値及び分布もまとまり非線形ながらも年々昇していることが分かる。従って、この影響をコントロールするためにも年次をダミー変数とししして回帰式に入れることは妥当であると考える。

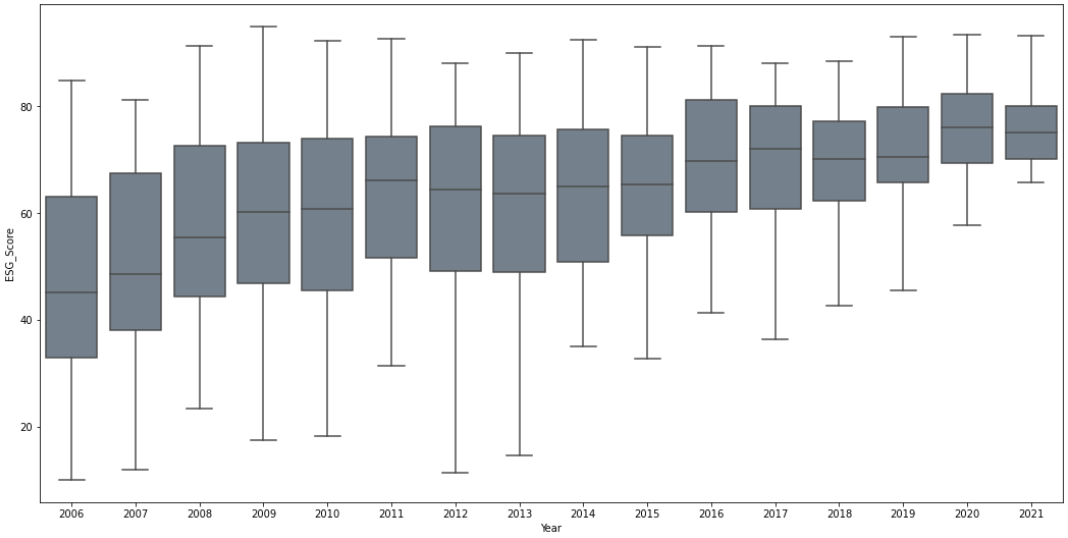


図 11　2006 年から 2021 年における ESG スコアの箱ひげ図

続いて業界別の特徴を確認する。業界別の株主提案がされた ESG スコアをもつ企業数は、「Manufacturing」が最も数が多く、次に「Finance and Insurance」という特徴が分かった。

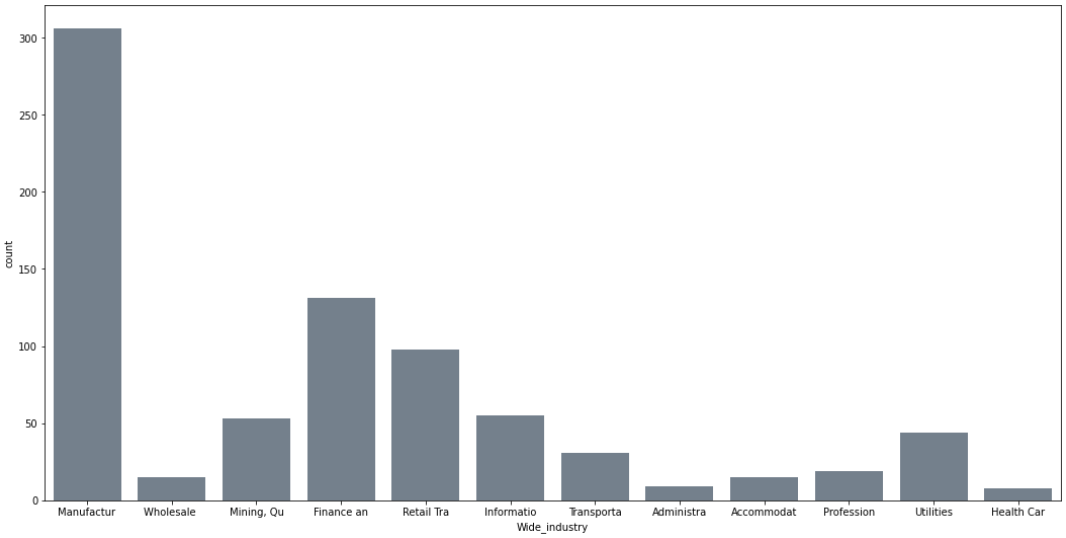


図 12　業界別の株主提案がされた ESG スコアを持つ企業数

次に、業界別の ESG スコアの特徴を確認する。概ね平均値は近しいものの業界によって分布のばらつきがありそうである、しかし先ほどの結果より企業数の数が 50 以下と少ない業界もあることに注意は必要である。

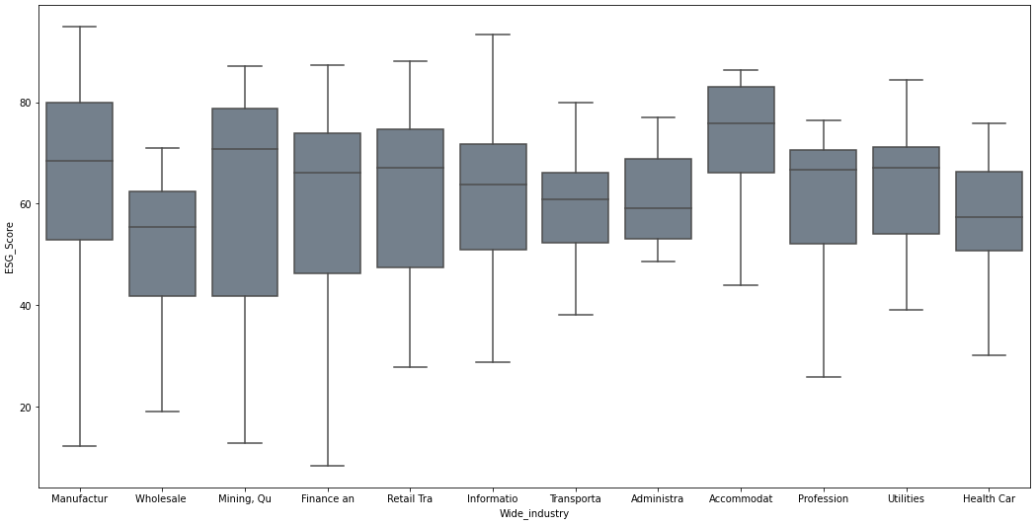


図 13　業界別の ESG スコアの箱ひげ図

次に、対数変換を行った後のヒストグラムと、正規化を行った後のヒストグラム及び基本統計量を示す。

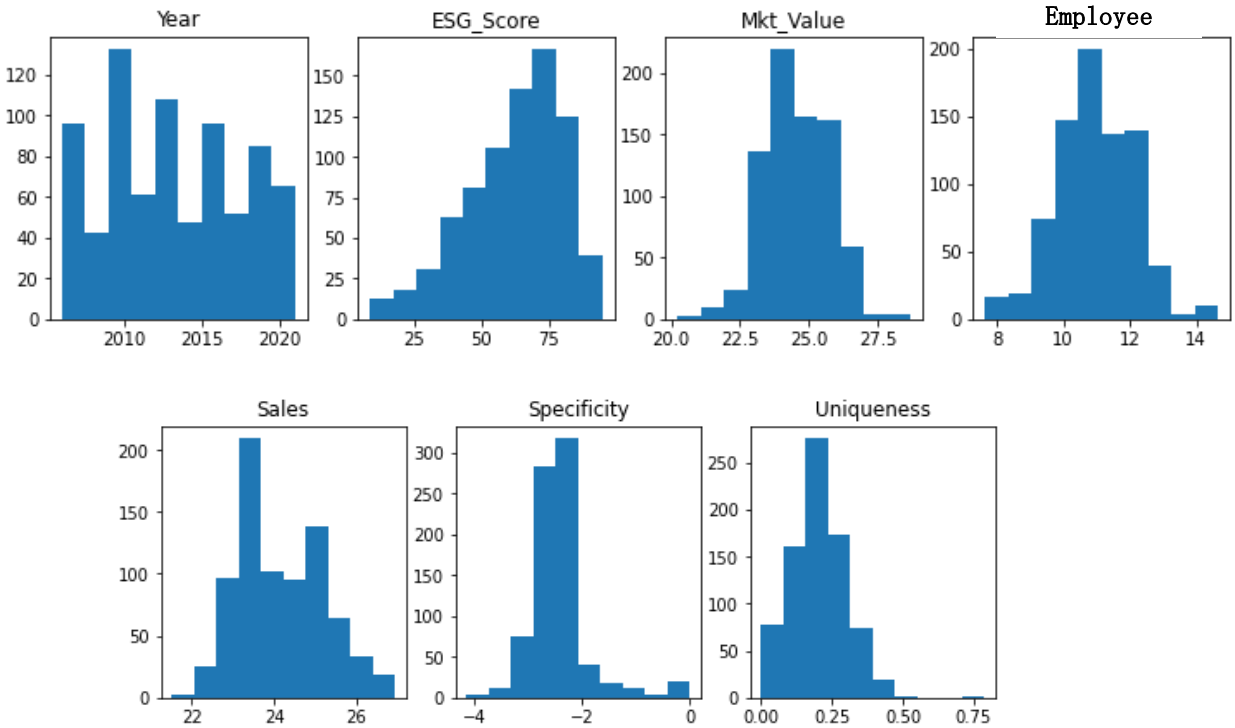


図 14　対数変換後のデータセットの概要

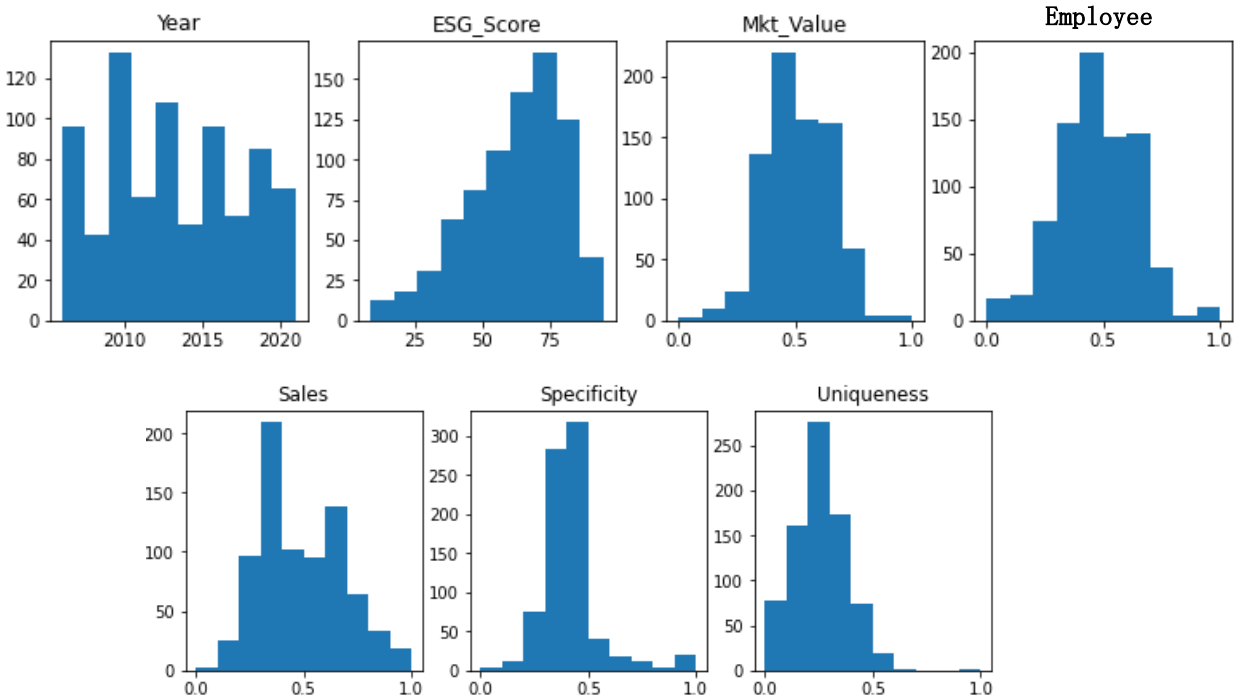


図 15　正規化後のデータセットの概要

	Year	ESG_Score	Mkt_Value	Employee	Sales
count	784.000000	784.000000	784.000000	784.000000	784.000000
mean	2012.956633	62.458366	0.513936	0.473918	0.488611
std	4.354139	17.780844	0.141394	0.166747	0.192588
min	2006.000000	8.414749	0.000000	0.000000	0.000000
25%	2009.000000	50.860095	0.412974	0.369409	0.341300
50%	2013.000000	66.422578	0.501543	0.472998	0.447413
75%	2017.000000	76.089280	0.626482	0.593966	0.638691
max	2021.000000	94.931304	1.000000	1.000000	1.000000

	Specificity	Uniqueness
count	784.000000	784.000000
mean	0.418880	0.258835
std	0.139206	0.123000
min	0.000000	0.000000
25%	0.348200	0.184125
50%	0.405981	0.248984
75%	0.454370	0.333065
max	1.000000	1.000000

図 16　基本統計量の概要（データ前処理後）

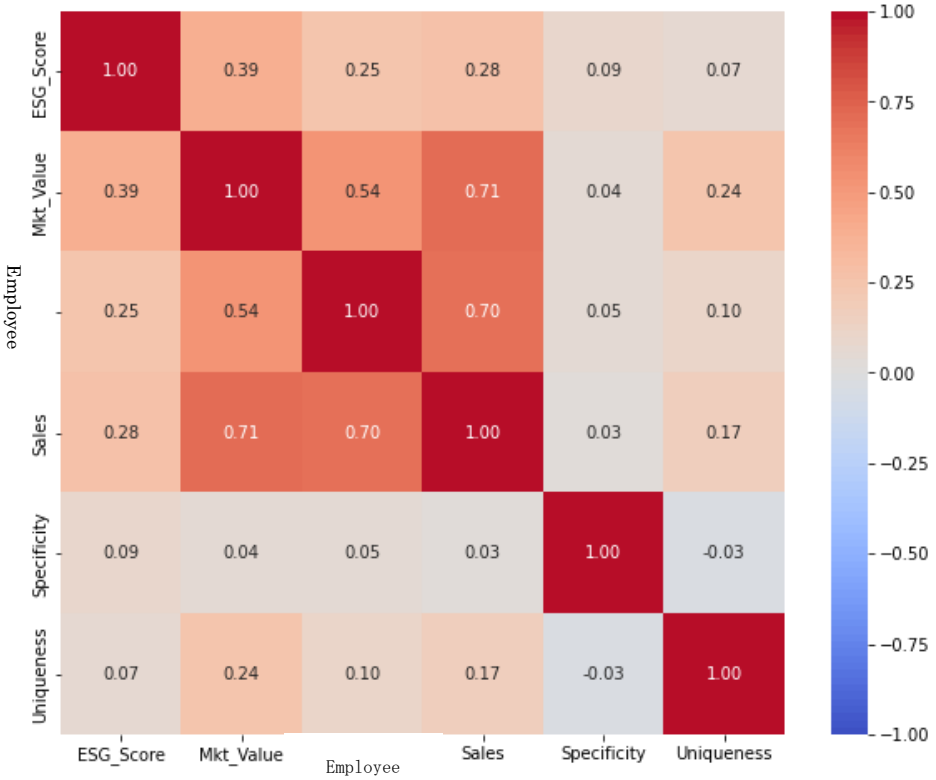


図 17　相関ヒートマップ

相関ヒートマップ及び散布図を確認してみると、当然のことながら、Sales、Mkt_Value、Employee の相関は特徴的である。一方で、赤枠部分の ESG スコアとテキスト情報から作成した、Specificity と Uniqueness の相関はほとんどみられない。

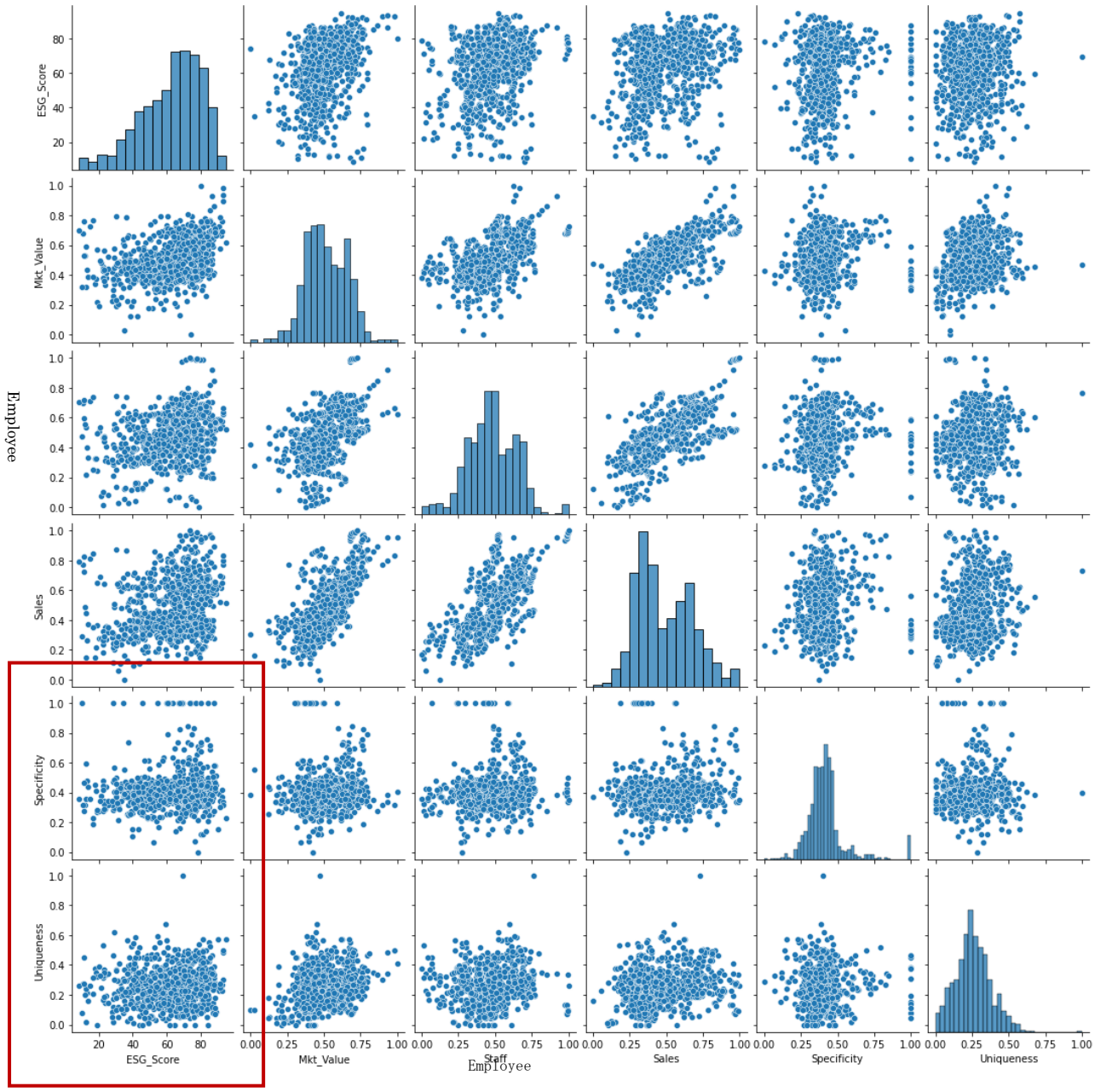


図 18　各変数の総散布図

BERTopic によるトピック分析について

BERTopic では「I」や「and」などの単語はあまり気にしなくてよいと、他モデルに比べると言われている。しかし、本研究のテキストデータではその点が顕著に表れてしまったため、NLTK における下記ストップワーズ 179 個を除去した。

['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', "you're", "you've", "you'll", "you'd", 'your', 'yours', 'yourself', 'yourselves', 'he', 'him', 'his', 'himself', 'she', "she's", 'her', 'hers', 'herself', 'it', "it's", 'its', 'itself', 'they', 'them', 'their', 'theirs', 'themselves', 'what', 'which', 'who', 'whom', 'this', 'that', "that'll", 'these', 'those', 'am', 'is', 'are', 'was', 'were', 'be', 'been', 'being', 'have', 'has', 'had', 'having', 'do', 'does', 'did', 'doing', 'a', 'an', 'the', 'and', 'but', 'if', 'or', 'because', 'as', 'until', 'while', 'of', 'at', 'by', 'for', 'with', 'about', 'against', 'between', 'into', 'through', 'during', 'before', 'after', 'above', 'below', 'to', 'from', 'up', 'down', 'in', 'out', 'on', 'off', 'over', 'under', 'again', 'further', 'then', 'once', 'here', 'there', 'when', 'where', 'why', 'how', 'all', 'any', 'both', 'each', 'few', 'more', 'most', 'other', 'some', 'such', 'no', 'nor', 'not', 'only', 'own', 'same', 'so', 'than', 'too', 'very', 's', 't', 'can', 'will', 'just', 'don', "don't", 'should', "should've", 'now', 'd', 'll', 'm', 'o', 're', 've', 'y', 'ain', 'aren', "aren't", 'couldn', "couldn't", 'didn', "didn't", 'doesn', "doesn't", 'hadn', "hadn't", 'hasn', "hasn't", 'haven', "haven't", 'isn', "isn't", 'ma', 'mightn', "mightn't", 'mustn', "mustn't", 'needn', "needn't", 'shan', "shan't", 'shouldn', "shouldn't", 'wasn', "wasn't", 'weren', "weren't", 'won', "won't", 'wouldn', "wouldn't"]

上記既存のストップワーズに加え、株主提案文書自体の特徴を表す単語「shareholder(s)」「stockholder(s)」「proposal(s)」「the」「board」「annual」「director(s)」「mr」を追加し、除外した。

トピックは図のように 10 つ分類された、上位 5 つ関して特徴を推定する。一つ目のトピックは、proxy, meeting, company, vote というキーワードとなり 168 個カウントされている、「株主総会」に関するトピックであると推定す

る。二つ目のトピックは、company, political, vote, report というキーワードであり、121 個カウントされている。トピックとしては「政策、提案、投票」に関するものだと推定する。三つ目のトピックは、company, vote, majority, voting となり、115 個カウントされている、「投票の最大派閥」に関するものだと推定する。四つ目のトピックは、compensation, proxy, executive, company とでており 108 個カウントされている、「役員報酬」のトピックと推定する。五つ目のトピックは、proxy, executive, meeting, committee と 108 個出ており、「経営陣」に関するトピックであると推定する。

表 6　BERTopic による分類とキーワードの抽出

	Topic	Count	Name
0	0	168	0_proxy_meeting_company_vote
1	1	121	1_company_political_vote_report
2	2	115	2_company_vote_majority_voting
3	3	108	3_compensation_proxy_executive_company
4	4	108	4_proxy_executive_meeting_committee
5	5	81	5_meeting_company_proxy_shares
6	6	73	6_company_executive_committee_compensation
7	7	37	7_sign_cisco_please_vote
8	8	16	8_separation_relating_presented_properly
9	9	16	9_street_meeting_date_must

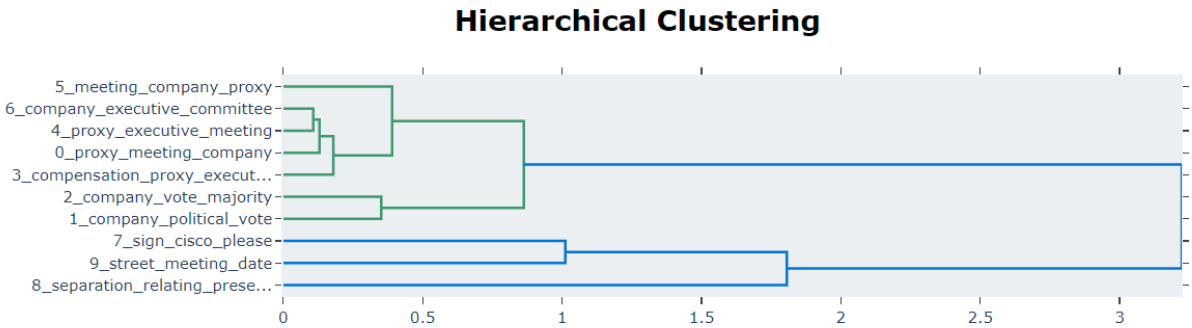


図 19　トピックのクラスタリング分類

図 20　BERTopic によるトピック分類間の位置関係

トピック 3 の「経営陣の役員報酬」に関するトピックは企業のガバナンス領域に該当する可能性があるが、ESG に関するトピックはあまり抽出されず、株主提案という要素が濃くでた結果となった。

5.2 重回帰分析の結果について

これらの変数を重回帰分析した結果が次項である。年次と業界はダミー変数として年次による効果と業界による効果をコントロールするために入れているため、枠に囲まれた x1,2,3,4,5 について注目し考察をしてみる。

有意水準 5%以下で、有意な結果として出ているのは、X1 の Mkt_Value(= 時価総額)、x2 の Employee (=従業員数),x4 の Specificity(=固有表現割合)であった。次に回帰係数は、X1 の Mkt_Value(= 時価総額)が「32.8569」、x2 の Employee (=従業員数)が「17.4753」 ,x4 の Specificity(=固有表現割合)が「10.1377」という結果となり、全て正の値であるため、これらの変数の増加は, ESG スコアにプラスの影響を与える可能性がある。

X1 の Mkt_Value(= 時価総額)について、この変数が 1 増加すると、ESG スコアは 32.8569 増加する傾向にあるという事である。増加幅は多少異なるにせよ、x2 の Employee (=従業員数)についても同様の解釈となる。

時価総額や従業員数が大きくなるとより企業として注目され、ガバナンス等が効いてくる（効かせるようになる）ことや、CSR 活動などに力をいれる余裕が出てきて ESG スコアが上がるという解釈が出来るであろう。

次に、x4 の Specificity(=固有表現割合)の変数が 1 増加すると、ESG スコアが 6.1962 増加する点であるが、株主提案文書内のテキスト情報において固有表現割合の割合が高い方が、ESG スコアが高くなる傾向があることを示している。あえて解釈をすると、固有表現つまり自分達に関する具体的な事象に関する株主提案は株主総会で可決され、それが企業の ESG 活動に正の影響を与えうるという事である。

一方で、Uniqueness(= 独自性)は有意水準 10%以下でも有意な結果として出ていない。回帰係数は、マイナス 5.0136 となっており、負の影響を及ぼす解釈となる。つまり、有意な結果であれば、Uniqueness(= 独自性)の変数が 1 増加すると、ESG スコアは 5.0136 減少する傾向があるということになる。業界水準から乖離しすぎる ESG に関する株主提案内容の場合、ESG スコアを落とすことになり得るのかもしれない。

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.283
Model:                            OLS   Adj. R-squared:                  0.253
Method:                 Least Squares   F-statistic:                     9.564
Date:                Wed, 28 Dec 2022   Prob (F-statistic):           5.78e-37
Time:                        17:16:29   Log-Likelihood:                -3238.1
No. Observations:                 784   AIC:                             6540.
Df Residuals:                     752   BIC:                             6689.
Df Model:                          31
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         36.1007      5.315      6.792      0.000      25.667      46.534
x1            32.8569      6.667      4.928      0.000      19.768      45.945
x2            17.4753      5.083      3.438      0.001       7.496      27.454
x3             6.1962      5.439      1.139      0.255      -4.481      16.874
x4            10.1377      4.104      2.470      0.014       2.081      18.195
x5            -5.0136      4.969     -1.009      0.313     -14.768       4.741
x6             1.2722      6.578      0.193      0.847     -11.640      14.185
x7             3.5555      4.246      0.837      0.403      -4.780      11.891
x8             5.9861      6.817      0.878      0.380      -7.396      19.368
x9             3.4613      4.540      0.762      0.446      -5.451      12.374
x10            4.8116      4.119      1.168      0.243      -3.275      12.898
x11            5.5028      4.559      1.207      0.228      -3.448      14.453
x12            3.2504      5.368      0.606      0.545      -7.287      13.788
x13            6.2525      4.313      1.450      0.148      -2.215      14.720
x14            6.1314      4.909      1.249      0.212      -3.505      15.768
x15            1.6892      4.645      0.364      0.716      -7.430      10.808
x16            3.3222      5.680      0.585      0.559      -7.829      14.474
x17          -11.2761      3.259     -3.460      0.001     -17.674      -4.879
x18           -1.2560      3.205     -0.392      0.695      -7.547       5.035
x19          -16.5466      2.934     -5.639      0.000     -22.307     -10.786
x20          -18.4956      2.963     -6.241      0.000     -24.313     -12.678
x21           -9.0181      3.098     -2.911      0.004     -15.099      -2.937
x22           -7.4923      3.037     -2.467      0.014     -13.454      -1.530
x23           -3.9359      3.143     -1.252      0.211     -10.106       2.234
x24          -22.9647      3.284     -6.993      0.000     -29.412     -16.518
x25          -12.7745      3.192     -4.002      0.000     -19.041      -6.508
x26           -1.8221      3.257     -0.559      0.576      -8.216       4.572
x27           -5.6187      3.177     -1.769      0.077     -11.856       0.618
x28           -4.2358      3.214     -1.318      0.188     -10.545       2.073
x29          -12.6903      3.550     -3.574      0.000     -19.660      -5.720
x30          -10.1348      3.499     -2.897      0.004     -17.003      -3.266
x31           -0.9906      3.745     -0.265      0.791      -8.342       6.361
==============================================================================
Omnibus:                       52.901   Durbin-Watson:                   1.376
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               62.543
Skew:                          -0.644   Prob(JB):                     2.62e-14
Kurtosis:                       3.508   Cond. No.                         39.2
==============================================================================
```

図 21　重回帰分析の結果

6.結論

6.1 まとめ

市場のESGに対する関心の高まり、そしてアクティビストの台頭から、ESGスコアとESGに関する株主提案文書間の関係性を探るというテーマで研究を行った。米国の784件のデータを用いて探索的分析したところ、Specificity(=固有表現割合)の回帰係数が正となり有意な結果が得られつつも、強い相関があるとは言い切れない結論となった。またUniqueness（=独自性）の回帰係数が負となったが、有意な結果とはならなかった。

6.2 課題と限界そして今後の展望について

6.2.1 課題と展望について

本研究において探索的分析を進める中で、少なくとも以下の点が分析の精緻化を図る上での課題として残された。

1点目として、そもそもESGスコアを構成している要素の問題である。ESGスコアの作成にあたり、株主提案文書のテキスト情報を使っていないと確認はとっているものの、本研究で使用した時価総額や従業員数などのその他の変数については少なからず要素としてESGスコアの構成に影響を与えていると考える。具体的には今回得られた実証的知見の整合性を、他のESGスコア（ブルームバークなど）を使用し相違点があるか確認することが挙げられる。

2点目として、分析に採用したサンプルの代表性に関する問題である。今回はスクレイピングに関わるテクニカルな理由により784件のサンプル（全2121件の36.9%)を用いることになった。よって、今回得られた結果がバイアスを含んでいる可能性があり、結果を直ちに一般化することはできないことに注意が必要である。今後は、スクレイピング方法を改良することによりサンプルサイズを大幅に増やせる可能性がある。今回の分析に使用したデータは米国証券取引委員会（Securities and Exchange Commission）より株主提案文書のテキスト情報を抽出したものであるが、前述したとおりHTMLの構造上完全にスクレイピングが出来ていない点が挙げられる。スクレイピングのプログラムを改善するか、マニュアルにて株主提案文書を抽出する方法が考えられるが、研究の再現性を考慮すると前者を改善するのが良いと考える。米国証券取引委員会（Securities and Exchange Commission）のデータについての補足であるが、フェアアクセスポリシーとなり誰でもデータにアクセスできる。ただし、1秒間に10リクエスト以下というアクセス規制がある。また、サイト内の自動ツ

ールでのクローリングは禁止されているため、米国の全てのデータを抽出するサイト内でページ遷移するようなプログラムによるスクレイピングは出来ない。今回は、Proxy Monitor にて整理された ESG 関連の株主提案文書のリストがあり、文書毎の各 URL に直接遷移する形で取得をして都度離脱をしているため、サイト内のページ間クローリングに該当しないようにした。

3 点目として、テキスト情報の指標化に関する網羅性の欠如である。今回は Specificity と Uniqueness の 2 つの指標化を実施したが、他の代替となる指標を用いた分析が必要である。今回の分析に使用した説明変数は数が少なく、“探索的”分析としては不十分なものであった。また、BERT も次々に改良型が提案されており、それらの最新の技術を採用することも可能である。

4 点目として、目的変数である ESG Score と、株主提案文書の関係性を明らかにする目的を達成するためには、分析に使用する指標や適用する分析方法など多面的により有効なアプローチを探る必要がある。例えば、今回の研究では、それぞれの株主提案文書を業界毎の平均文書との距離で指標を作成したが、ESG に関する株主提案として模範的な文書があれば、そこからの距離を使って指標を作成することが考えられる。具体的な模範的な文書としては、国のガイドラインやもしくは議会の議事録などが考えられる。

また、BERTopic においてクラスタリング方法（手法や距離の選択）や BERT モデルをより最新のモデルに変更することにより的確な分析結果が得られる余地がある。

5 点目として、予想していた結果が出なかった点、つまり想定していた初期仮説と異なる結果が得られた点である。研究目的に対してリサーチデザインの設定が妥当でなかった可能性がある。この点は、本研究のテーマ、すなわち、ESG スコアと株主提案の関係性理解に関するより本質的な点であり、今後研究を発展させていく上で十分に検討を図る必要がある。具体的には、予想していた逆の因果が現れている可能性があり、ESG スコアが高い企業であれば、ESG に関する企業活動について株主とコミュニケーションが良く出来ている、よって、その特徴はテキスト化された指標（Specificity や Uniqueness など）に反映されるはずである、と考えられる。具体的なアプローチとしては、今回説明変数に用いた Specificity や Uniqueness を左辺の目的変数にそれぞれ設定し、右辺に説明変数として ESG スコアを採用する分析アプローチである。

6 点目として、ESG 関連以外の株主提案を受けている企業の ESG スコアデータの活用についてである。今回は ESG 関連の株主提案に絞ったデータセットにて分析を進めてきたが、ESG 関連以外の株主提案のある企業も含め仕分けすることが出来れば、両グループ間での係数の差を比較することで ESG 関連の株主提案を受けている企業の ESG スコアへの効果を評価できる可能性がある。ま

た、株主提案を受けていない企業のESGスコアデータを含めた、全体のESGスコアデータの経年傾向を可視化することで、新たな発見がある可能性がある。

6.2.1 限界について

日米など国ごとの差を比較する視点も研究の意義があると考えるが、ESG関連の日本の株主提案数は毎年10件程度であり、そのほとんどがガバナンスに関するものであるため、データセットの整備は難しい。

謝辞

本研究を行う上で、指導教官である林高樹教授には非常に多くのご指導を頂きました。1年という短い期間ではありましたが、ゼミナールを通して、多変量解析法の基本的な統計学から、ビジネス界で活用されているデータサイエンスまで、また修士論文の執筆にあたり学術論文の調べ方や書き方などご指導いただきました。最終的に修士論文として形にできたのは林先生の面倒見の良い、熱心なご指導のおかげでございます。また副査を引き受けて下さった、高橋大志教授と武田史子教授には、リサーチデザインに関する鋭いご指摘とアドバイスを賜りました。改めて、この場をお借りして皆さまには感謝申し上げます。そして、林高樹研究室のメンバー（ローデンバーグ君、駒井君）と一緒の研究室になれたこと嬉しく思います。サブゼミでのアドバイスやOB後援会のサポートなど、たくさん助けて頂きましたこと、感謝いたします。M44の同級生の方々にも精神的に支えられ、KBSでの2年間をやり抜くことが出来たと感じております、心より感謝申し上げます。本当にありがとうございました。

参考文献（アルファベット順）

青木 崇 (2016), 業価値の向上を目指す日本企業の情報開示のあり方とESG活動, 論集 第69巻第1・2号.

青嶋 智久 (2019), 日本語BERTモデルを用いた経済テキストデータのセンチメント分析, The 33rd Annual Conference of the Japanese Society for Artificial Intelligence.

Manel Allaya (2022) , Do Dual-Class Shares Matter for Annual Report Readability?, ISAFE 2022.

Vincente Cunat (2010), The Vote Is Cast: The Effect of Corporate Governance on Shareholder Value. The Journal of Finance, American Finance Association, Volume67, Issue5.

Jacob Devlin (2019), BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv:1810.04805.

Stuart L. Gillan (2022), Heterogeneity in shareholder activism: Evidence from Japan, Pacific-Basin Finance Journal,Volume 77, February 2023, 101891

Lilli A. Gordon (1993), Information, Ownership Structure, and Shareholder Voting: Evidence from Shareholder-Sponsored Corporate Governance Proposals, The Journal of Finance, American Finance Association. Volume48,Issue2.

Maarten Grootendorst (2022), BERTopic: Neural topic modeling with a class-based TF-IDF procedure, arXiv:2203.05794v1 [cs.CL] 11 Mar 2022.

畠田 敬 (2021), 日本におけるESG開示スコアがもつ含意, 神戸大学経済経営学会 国民経済雑誌,223(5):47-69.

林 寿和 (2018), 開示情報量と企業による印象操作が ESG 評価に及ぼす影響についての一考察：シミュレーションモデルを用いて, 日本経営倫理学会誌第 25 号.

J Karpoff (1996), Corporate governance and shareholder initiatives: Empirical evidence, Journal of Financial Economics, volume 42, p. 365 - 395, Posted: 1996.

加藤 弘祐,(2021), クオリティペーパーを対象とした分散表現に基づくトピック分析, フードシステム研究第 28 巻 4 号 p. 328-333.

Dong-young Kim (2014), Effects of Corporate Social Responsibility and Governance on Its Credit Ratings, Hindawi Publishing Corporation Scientific World Journal Volume 2014, Article ID 305452, 6 pages.

金野 有真(2020), 有価証券報告書のテキスト情報が投資家に与えるインパクトの定量的分析, 慶應義塾大学大学院経営管理研究科, 修士学位論文, 2020 年度経営学第 3711 号.

九多里 桐子(2018), わが国における株主提案の現状, 北九州市立大学「商経論集」第 53 巻第 1.2.3.4 号.

Hiroshi Sasaki (2020), An investigation of the relationship between SDG commitment level and ESG scores of all listed companies in Japan, 日本情報経営学会 第 80 回大会予稿集.

Indarawati Tarmuji(2016), The Impact of Environmental, Social and Governance Practices (ESG) on Economic Performance: Evidence from ESG Score, International Journal of Trade, Economics and Finance, Vol. 7, No. 3, June 2016.

陸 詩 瓊 (2021), A Text-Mining Analysis on Non-Financial Reporting Practices and the Quality of Disclosure, 慶應義塾大学大学院経営管理研究科, 修士学位論文, 2021 年度経営学 第 3888 号.

成 誠 (2021), テキストマイニングによる「新型コロナ不満情報」の特徴抽出と特徴分析, 慶應義塾大学大学院経営管理研究科. 修士学位論文. 2021 年度経営学第 3849 号.

Tsung-ming Yeh (2014), Large Shareholders, Shareholder Proposals, and Firm Performance: Evidence from Japan. Corporate Governance: International Review 22-4 312-329, 2014.

Tsung-ming Yeh (2017), Determinants and consequences of shareholder proposals: The cases of board election, charter amendment, and profit disposal. Journal of Corporate Finance 45 (2017) 245-261.

内元 清貴(2000), 最大エントロピ-モデルと書き換え規則に基づく固有表現抽出, 自然言語処理 7 巻 2 号 p.63-90.

湯山 智教(2019), ESG 投資のパフォーマンス評価を巡る現状と課題, みずほ証券株式会社・株式会社日本投資環境研究所発行「資本市場リサーチ」2019 年冬季第 50 号特別号 P.85-P.112.

付録

展望部分で記載した部分について出来る部分実行した結果を記載する。作成した Specificity と Uniqueness を目的変数として設定し、ESG スコアを右辺に置き、ESG スコアが単位当たり変化したときのそれぞれの指標の変化についても特徴を探った結果を記述する。その際に使用した ESG スコアの分布は下記となる。

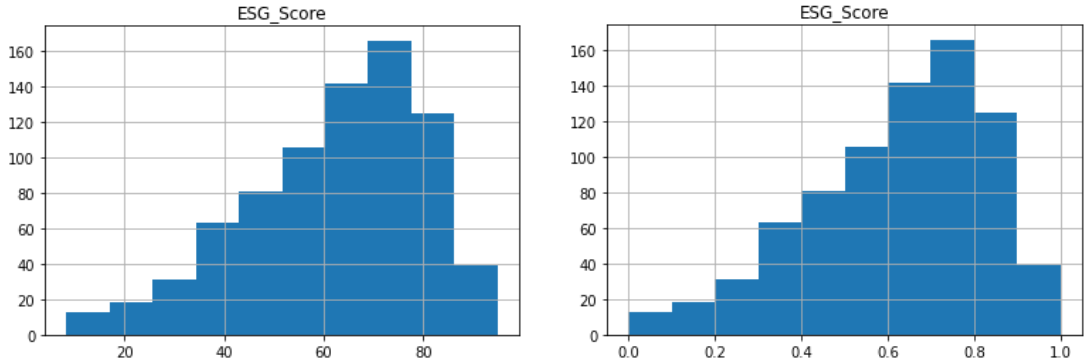


図 26　ESG スコアの分布について

回帰分析の結果は、図 27,28 となり、x1 が共に ESG スコアとなる。Specificity を目的変数として回帰させた時、有意水準 10%以下にて有意であり、ESG_Score の係数は「0.0687」となり正の関係であることが分かる。一方で、 Uniqueness を目的変数として回帰させた時、有意水準 10%以下でも有意とはならなかったが、係数をみてみると「-0.0233」となり、負の関係である。指標として、Uniqueness は適切でない可能性が考えられる事が分かった。

記号	被説明変数	補足
	Specificity / Uniqueness	5 の説明変数と連動

x	説明変数	
1	ESG_Score	ESG スコア
2	Mkt_Value	時価総額
3	Employee	従業員数
4	Sales	売上規模
5	Specificity / Uniqueness	目的変数と連動

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.084
Model:                            OLS   Adj. R-squared:                  0.046
Method:                 Least Squares   F-statistic:                     2.215
Date:                Mon, 09 Jan 2023   Prob (F-statistic):           0.000189
Time:                        11:57:04   Log-Likelihood:                 468.19
No. Observations:                 784   AIC:                            -872.4
Df Residuals:                     752   BIC:                            -723.1
Df Model:                          31
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.3255      0.046      7.019      0.000       0.234       0.417
x1             0.0687      0.028      2.470      0.014       0.014       0.123
x2             0.0518      0.060      0.865      0.387      -0.066       0.169
x3             0.0116      0.045      0.257      0.797      -0.077       0.101
x4            -0.0777      0.048     -1.617      0.106      -0.172       0.017
x5            -0.0686      0.044     -1.562      0.119      -0.155       0.018
x6             0.0825      0.058      1.419      0.156      -0.032       0.197
x7             0.0660      0.038      1.759      0.079      -0.008       0.140
x8             0.0672      0.060      1.114      0.266      -0.051       0.186
x9             0.0815      0.040      2.034      0.042       0.003       0.160
x10            0.0685      0.036      1.883      0.060      -0.003       0.140
x11            0.0961      0.040      2.389      0.017       0.017       0.175
x12            0.1050      0.047      2.218      0.027       0.012       0.198
x13            0.0590      0.038      1.547      0.122      -0.016       0.134
x14            0.0564      0.043      1.299      0.194      -0.029       0.142
x15            0.0847      0.041      2.067      0.039       0.004       0.165
x16            0.0874      0.050      1.742      0.082      -0.011       0.186
x17            0.0208      0.029      0.717      0.474      -0.036       0.078
x18           -0.0388      0.028     -1.368      0.172      -0.094       0.017
x19            0.0227      0.026      0.858      0.391      -0.029       0.075
x20           -0.0005      0.027     -0.018      0.985      -0.053       0.052
x21            0.0321      0.028      1.165      0.244      -0.022       0.086
x22           -0.0482      0.027     -1.790      0.074      -0.101       0.005
x23           -0.0715      0.028     -2.580      0.010      -0.126      -0.017
x24            0.0363      0.030      1.211      0.226      -0.023       0.095
x25            0.0270      0.029      0.946      0.345      -0.029       0.083
x26            0.0415      0.029      1.441      0.150      -0.015       0.098
x27            0.0694      0.028      2.473      0.014       0.014       0.124
x28           -0.0305      0.028     -1.072      0.284      -0.086       0.025
x29            0.0068      0.032      0.214      0.831      -0.055       0.069
x30           -0.0193      0.031     -0.620      0.535      -0.080       0.042
x31           -0.0054      0.033     -0.164      0.870      -0.070       0.060
==============================================================================
Omnibus:                      346.570   Durbin-Watson:                   1.602
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1990.880
Skew:                           1.938   Prob(JB):                         0.00
Kurtosis:                       9.777   Cond. No.                         41.0
==============================================================================
```

図 27　重回帰分析の結果
(Specificity を目的変数に設定)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.194
Model:                            OLS   Adj. R-squared:                  0.161
Method:                 Least Squares   F-statistic:                     5.830
Date:                Mon, 09 Jan 2023   Prob (F-statistic):           3.26e-20
Time:                        11:57:04   Log-Likelihood:                 615.41
No. Observations:                 784   AIC:                            -1167.
Df Residuals:                     752   BIC:                            -1018.
Df Model:                          31
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1921      0.039      4.919      0.000       0.115       0.269
x1            -0.0233      0.023     -1.009      0.313      -0.069       0.022
x2             0.1831      0.049      3.720      0.000       0.086       0.280
x3             0.0085      0.038      0.226      0.822      -0.065       0.082
x4             0.0403      0.040      1.009      0.313      -0.038       0.119
x5            -0.0471      0.030     -1.562      0.119      -0.106       0.012
x6            -0.0524      0.048     -1.087      0.278      -0.147       0.042
x7             0.0007      0.031      0.022      0.982      -0.060       0.062
x8            -0.0228      0.050     -0.456      0.648      -0.121       0.075
x9            -0.0020      0.033     -0.059      0.953      -0.067       0.063
x10           -0.0063      0.030     -0.208      0.835      -0.066       0.053
x11           -0.0115      0.033     -0.343      0.732      -0.077       0.054
x12            0.0017      0.039      0.044      0.965      -0.076       0.079
x13            0.0217      0.032      0.687      0.492      -0.040       0.084
x14           -0.0386      0.036     -1.071      0.284      -0.109       0.032
x15           -0.0260      0.034     -0.762      0.446      -0.093       0.041
x16            0.0027      0.042      0.064      0.949      -0.079       0.084
x17            0.0045      0.024      0.186      0.853      -0.043       0.052
x18           -0.0075      0.024     -0.320      0.749      -0.054       0.039
x19           -0.0004      0.022     -0.019      0.985      -0.044       0.043
x20           -0.0266      0.022     -1.196      0.232      -0.070       0.017
x21            0.0487      0.023      2.136      0.033       0.004       0.093
x22            0.0065      0.022      0.289      0.773      -0.037       0.050
x23           -0.0346      0.023     -1.502      0.134      -0.080       0.011
x24           -0.0356      0.025     -1.436      0.151      -0.084       0.013
x25           -0.0590      0.024     -2.504      0.012      -0.105      -0.013
x26            0.1022      0.024      4.330      0.000       0.056       0.149
x27           -0.0169      0.023     -0.724      0.469      -0.063       0.029
x28           -0.0508      0.024     -2.158      0.031      -0.097      -0.005
x29           -0.0531      0.026     -2.030      0.043      -0.105      -0.002
x30           -0.0286      0.026     -1.111      0.267      -0.079       0.022
x31           -0.1065      0.027     -3.917      0.000      -0.160      -0.053
==============================================================================
Omnibus:                      127.425   Durbin-Watson:                   1.616
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              318.510
Skew:                           0.856   Prob(JB):                     6.86e-70
Kurtosis:                       5.612   Cond. No.                         41.9
==============================================================================
```

図 28　重回帰分析の結果

(Uniqueness を目的変数に設定)