

企業レポート中の ESG に関する公約の検証可能性の判別

朱 博瑄

近年、投資、企業評価、企業経営などの分野において ESG（環境・社会・ガバナンス）が重要視されつつある。企業の ESG に関する取り組みや貢献などの「公約」は企業文書のテキスト情報から評価することができるが、公約が検証可能な時期を特定し、その実現可能性を「根拠」の性質によって検証することは、あたかも環境に配慮しているかのように見せかける企業のグリーンウォッシング行為の防止にも繋がり、ESG 評価において重要と考える。そこで本研究では企業文書から ESG に関する公約を自動抽出し、その検証可能性を根拠の性質によって自動評価する。これにより ESG 情報を用いた企業評価の支援を目指す。

本研究では、LLM（大規模言語モデル）を用いて自動で公約や根拠の抽出や分類を行う。その際、RAG（検索拡張生成）と In-context Learning（文脈内学習）の有効性を検証する。具体的には、検索対象データから分析対象データと類似した正解ラベル付きの複数の事例を検索し、LLM に入力するプロンプトに含め文脈内学習を行う。また、LLM を用いて検索対象データの抽出または分類結果の理由に関する解説文を自動生成し、検索対象データを拡張し LLM の文脈内学習を補強する。LLM による抽出または分類結果の出力には、プロンプトに同時に含める属性の情報を変えることで同時推定手法、公約・根拠同時推定手法、逐次推定手法の 3 手法を用い、他の属性の情報が文脈内学習に与える影響を分析する。

公約の検証可能性の判別のために、企業レポート中の文書を対象に、「公約の有無」、「公約が検証可能な時期」、「根拠の有無」、「根拠の性質」の 4 つの分類属性と、「公約の内容」、「根拠の内容」という 2 つの抽出属性の値を付与するアノテーションを行い、600 件のデータセットを構築した。4 つの分類属性においては、Fleiss の κ 係数 を用いた判定者間一致度が 0.6 を超え、作業員間の判定がある程度一貫していることを示した。

公約や根拠の抽出および分類実験では、3 つの出力手法の全てにおいて、提案手法を用いることで 6 つの属性の F 値のマクロ平均が向上し、RAG と解説文生成の有効性が確認できた。特に、公約・根拠同時推定手法では、RAG のみの場合と RAG と解説文有りの場合の共に、ベースライン手法との間で t 検定（有意水準 5%、両側検定）で有意差が認められ、また RAG と解説文有りの場合はマクロ平均が 0.691 と全ての実験結果の中で最も高かった。提案手法における失敗分析では、文脈全体を考慮した解説文生成の必要性を考察した。さらに 2 つの同時推定手法と逐次推定手法との間で「公約の内容」の抽出精度の結果を比較することで、公約の抽出または分類において根拠を考慮することの有効性を明らかにした。また、公約の検証可能性を評価するための「根拠の性質」の分類精度は最も高い F 値が 0.618 となり、一定の精度で分類できることを示した。

（指導教員 関 洋平）