

BERT を用いたカーボンプライシング関連 論文の分析

Short Review
2023年7月

投資工学研究所
伊藤 央峻

1. はじめに

二酸化炭素（CO₂）の排出削減を促進する手法の1つにカーボンプライシングがある。CO₂の排出に価格付けすることで、排出量の削減に経済的なインセンティブを与えることを目的としている。主要な2つのアプローチとして炭素税と排出量取引があり、気候変動問題の解決のための重要な政策手段として各国で導入が進んでいる。

カーボンプライシングについて、多くの学術論文が発表されている。統合評価モデルを使った炭素価格や経済的インパクトの推定、制度に関する調査・分析、カーボン削減に関する技術的な話題等様々である。Wang *et al.* (2019) によるとカーボンプライシングのモデル分析に関する論文は毎年数百本に上り、本数は増加する傾向にあった。こうした論文の理解には経済や気候等に関する多様な専門知識が必要となる。

近年、大量のテキストデータを使ってトレーニングした大規模言語モデル (Large Language Model, LLM) を用いて人間の理解を補助する研究が盛んである。本研究では LLM の一種である BERT を用いてカーボンプライシング論文の内容把握・分析を試みた。BERT は他の高性能な LLM と比べ規模は小さいものの、目的に応じたファインチューニング（データを使った再学習による微調整）で特定タスクに適したモデルを作ることが可能で、高い性能を実現できる。

以下、2章で先行研究や本研究の概要を説明し、3、4章で各分析の結果を説明する。

2. BERT を用いたカーボンプライシング関連論文の分析の概要

2.1 先行研究

2.1.1 大規模言語モデルと BERT について

大規模言語モデルは、自然言語処理（NLP）の分野で使用される人工知能（AI）モデルを指す。大量のテキストデータを学習して言語の統計的なパターンや構造を捉え、その知識を活用してテキスト生成、質問応答、文章要約、翻訳といった様々なタスクを実行する。

BERT は、2018年に Google が提案した LLM で、Transformer と呼ばれる深層学習モデルをベースにしている。文脈を考慮しつつ単語の意味や関係性を把握することで、様々な自然言語処理のタスクにおいて優れた結果を出すことで知られている。

BERT では事前学習とファインチューニングの2つのステップで学習が行われる。事前学習では、

Wikipedia 全文等の大規模なテキストコーパスを使って、教師なし学習を行う。この過程で、文の中の単語の意味や文脈を理解するための単語の埋め込み表現を獲得する。その後、質問応答や文書分類といった用途に応じてファインチューニングを実施する。

2.1.2 環境関連分野の BERT 活用事例

環境関連分野において BERT を活用した先行研究について述べる。

鈴木ら（2021）は BERT を用いて、企業等の活動事例や課題に関する文章を SDGs の枠組みに沿って分類する分類器を作成した。河南・関（2023）は様々な評価会社が算出する ESG スコアについて根拠が不明瞭になりがちである点に着目し、根拠となった文章の抽出を目的として、有価証券報告書の文章に ESG 分類（E・S・Gのどの分野か）及び極性（ポジティブ・ネガティブ）を付与する分類器を作成した。川崎（2023）は、BERT を用いて有価証券報告書の「事業等のリスク」に関する文章を TCFD 提言の気候変動リスク 6 項目に分類する分類器を作成した。

以上のように、環境関連分野の何らかの正解データをもとに BERT で分類器を作成し分析するという研究は多くなされているが、カーボンプライシングの関連論文については同様の分析は無いことから、本研究で検証し結果を報告する。

言語モデルの研究では特定分野の文章を集中的に学習することでその分野について、より高い性能を発揮するという結果が報告されている。例えば、環境分野の文章を集中的に学習させたモデルとして ClimateBERT¹がある。4 章ではこのモデルを用いてトピック分析を行う。

2.2 本研究の流れ

本研究では下記の順序でカーボンプライシング関連論文のテキスト分析を進める。

- ① 学習用データの収集
- ② 人手によるラベル付け
- ③ BERT による文章分類器の作成・分析

2.2.1 学習用データの収集

カーボンプライシングに関連する学術論文をインターネット上で収集し、学習データとした。具体的には、2021 年 12 月末までに公表された論文について、SSRN 等の論文サイトで「climate」「the social cost of carbon」等のカーボンプライシングに関連するキーワードで検索し、得られた約 1000 本の論文の要旨（Abstract）をテキストデータ化した。

¹ DistilRoBERTa モデルをベースに気候関連のニュースや研究論文や企業開示等、計 200 万パラグラフ以上を追加で学習した。気候関連の分類タスク等で、DistilRoBERTa モデル対比で精度改善がみられた (<https://www.chatclimate.ai/climatebert>)。

2.2.2 人手によるラベル付け

各論文について作業員 16 人で分担してラベル付けを行った。分類の基準及び各ラベルのデータ数は図表 1 に示した。

ラベル 1～3 はカーボンプライシング（炭素排出に対する価格付け）に関連した論文、ラベル 4～5 は関連していない論文と判断する。ラベル 1～5 の 5 分類及びカーボンプライシング関連論文か否かの 2 分類の 2 つの基準でのラベルが作成されたことになる。

各ラベルについて詳細を説明する。炭素価格を直接算出する内容の論文をラベル 1 とする。ラベル 1 に該当しないもののうち、炭素価格そのものではないが、コスト・回避損失等の代替的な価値を算出した論文をラベル 2 に割り当てた。ラベル 3 は、炭素価格等の算出はしていないものの、なんらかの炭素価格を用いた分析をしている論文を割り当てた。ラベル 4 は、カーボンプライシングに関する内容は無いが、気候変動モデル等に関する論文を割り当てた。ラベル 5 には、ラベル 1～4 に該当しない無関係の論文を割り当てた。

図表 1 分類の基準とデータ数

ラベル	基準	CP関連	データ数
ラベル 1	カーボンをプライシングしている	○	99
ラベル 2	カーボンプロキシプライシング	○	239
ラベル 3	なんらかのカーボン価格を置いて分析している	○	118
ラベル 4	カーボンプライシングとは関係ないが、気候モデルの部分の情報	×	243
ラベル 5	無関係の論文	×	257

(出所) 日興リサーチセンター作成

2.2.3 BERT による文章分類器の作成・分析

前項までで用意した学習用データを用いて事前学習済み BERT²をファインチューニングし、カーボンプライシング論文の分類器（5 分類・2 分類の 2 種）を作成する。学習用データのうち 9 割を学習に使い、1 割をテストデータとして使用する。次章で分類精度を確認し、SHAP 値によって分類の妥当性を評価する。さらに 4 章では BERT の文章ベクトルを利用したトピック分析手法である BERTopic を用いた分析を行う。

3. 分析結果

3.1 分類器の分類精度

3.1.1 5 分類の結果

正解率（Accuracy）は 42%となった。ランダム選択の場合（20%）を上回っており、ある程度分類

² <https://huggingface.co/bert-base-uncased>

できているものの、高い数値ではない。その要因に、作業員 16 人が各文章を 1 回ずつしか判断していない（ダブルチェック等をしていない）ため、ラベル付けにノイズが含まれる可能性がある。

そこで、テストデータを確認し、BERT と作業員のラベルの違いを検証した（図表 2 参照）。作業員がカーボンライジング関連でない（ラベル 4~5）と判断した論文のうち、BERT がカーボンライジング関連である（ラベル 1~3）とした 9 本の論文に絞り再確認した。その結果、6 本の論文のラベルを修正した³ところ、正解率は 46%に改善した。今回は一部データの改善にとどまったが、全てのデータのラベル付けを見直すことでさらなる精度改善が見込まれる⁴。

図表 2 分類精度（5 分類）⁵

	改善前	改善後
正解率 (Accuracy)	0.42	0.46
適合率 (Precision)	0.40	0.44
再現率 (Recall)	0.36	0.41
F1スコア	0.37	0.42

(出所) 日興リサーチセンター作成

3.1.2 2 分類の結果

結果を図表 3 に示した。正解率は 72%となり、ランダム時（50%）を上回った。さらに前項と同様に、作業員がカーボンライジング関連でないと判断した論文のうち BERT がカーボンライジング関連であるとした 17 本の論文を再確認した。その結果、10 本の論文のラベルを修正したところ、正解率は 79%に改善した。学習データ数が比較的少ないため、小さな修正であっても改善幅は大きかった。

図表 3 分類精度（2 分類）

	改善前	改善後
正解率 (Accuracy)	0.72	0.79
適合率 (Precision)	0.63	0.81
再現率 (Recall)	0.71	0.81
F1スコア	0.67	0.81

(出所) 日興リサーチセンター作成

³ 教師データを効率的に改善する「能動学習 (Active Learning)」にあたる操作。データ修正後は再度ランダムにデータをシャッフルし、学習データとテストデータを 9 : 1 の割合で分けて精度を再検証した。

⁴ climateBERT を用いて同様の実験をした場合でも特段の精度改善はみられなかった。

⁵ ここでの適合率、再現率、F1 スコアはラベルごとに算出した値を平均した値（マクロ平均値）である。

3.2 SHAP 値確認による可視化

3.2.1 SHAP について

SHAP (SHapley Additive exPlanations) は、機械学習モデルの特徴量の重要度を評価するための手法である。ゲーム理論の Shapley value に基づいており、個々の特徴量がモデルの予測結果にどの程度貢献しているかを定量化する。特徴量の重要度を可視化し、モデルの解釈性を向上させるために用いられる⁶。特に BERT 等の言語モデルでの文章分類においては、各ワードが予測結果にどの程度寄与したかを把握できる。

3.2.2 SHAP 値の確認

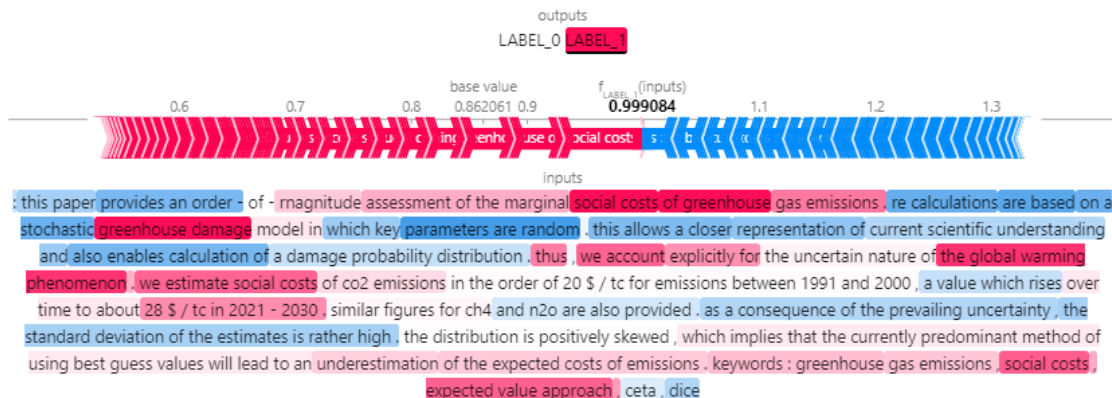
2 分類における SHAP 値を図表 4、5 にて確認した。図表 4、5 は作業者と BERT が共にカーボンプライシング関連論文であると分類した文章について、SHAP 値をテキスト及び数直線上で示したものである。正の寄与度を持つ特徴量が赤、負の寄与度を持つ特徴量が青で表示されている。例えば「social costs of greenhouse」や「carbon price」等のワードが赤字で表記されており、より濃い色となっている。これはカーボンプライシング関連論文であると BERT が分類した際の寄与度の高さを示す。

また、数直線上に表示された base value はテキストが全てマスクされた（隠された）状態でモデルに入力した場合の出力であり、 $f_{LABEL_1}(inputs)$ はテキストを全てマスクせずにモデルに入力した場合の出力（最終予測値）である。SHAP 値は各単語が存在する場合と存在しない場合の予測値の変化から寄与度を計算したものである。

なお、テストデータ（約 100 本の論文テキスト）において SHAP 値の絶対値が上位の単語を確認したところ、「carbon」「price」「cost」「tax」等が上位となっており、これらの単語が BERT の判断に強く寄与していた。これらの単語は作業者がカーボンプライシング関連か否かを判断する際にも重要な単語であり、こうした方法でもカーボンプライシングに関連する単語が BERT の判断に寄与していることを確認した。

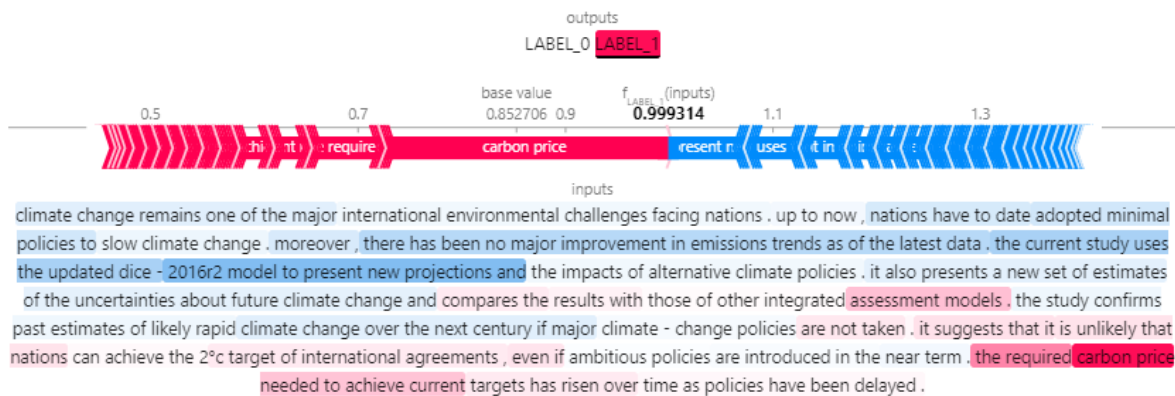
⁶ 機械学習アルゴリズム等のブラックボックスになりがちな工程について人間が理解しやすくする一連のプロセス・メソッドを「Explainable AI (XAI)」と呼ぶ。SHAP は XAI の手法の 1 つ。

図表 4 SHAP 値の確認①



(出所) 日興リサーチセンター作成

図表 5 SHAP 値の確認②



(出所) 日興リサーチセンター作成

4. BERTopic によるトピック分析

4.1 トピック分析及び BERTopic について

トピック分析は、大量のテキストデータから主題やトピックを抽出し、関連する文書や文章をグループ化するための手法で、テキストデータの可視化に役立つ。主要な手法の一つである LDA (Latent Dirichlet Allocation、潜在的ディリクレ配分法) 等、教師なし学習の手法を用いることで、ラベリングの手間をかけずにデータを可視化できる。

本研究では、カーボンプライシングに関する知識やトレンドの把握、トピックごとの論文の特定による効率化等を目的として、BERTopic を用いたトピック分析を行った。

BERTopic⁷はトピック分析手法の一つで、単純な単語の数や共起関係ではなく、BERT を用いて文脈を考慮したトピック抽出やグループ化を行う点が特徴となっている。具体的な手順としては、まず BERT で分析対象の文章をベクトルに変換する。ベクトルはそのままだと高次元でクラスタリングには向かな

⁷ ライブラリが公開されている (<https://github.com/MaartenGr/BERTopic>)。

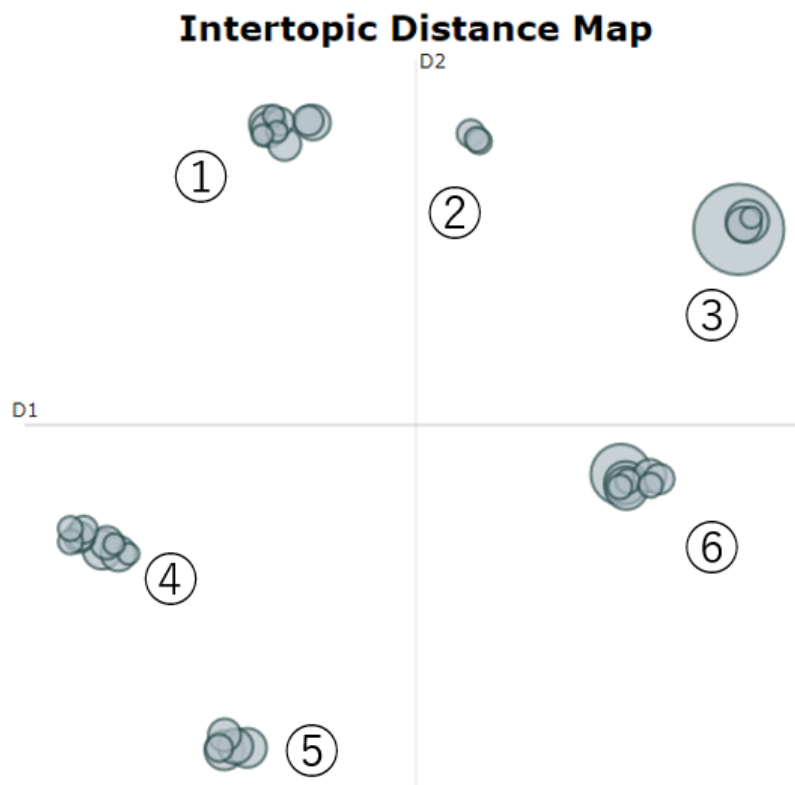
いため、次元削減を行う。その後クラスタリングし、各クラスタについて c-TF-IDF⁸によって代表的なワードを選び、トピックとする。

BERTopic では各工程で使用するアルゴリズムを細かく指定できる。本研究では BERTopic のデフォルトの設定に従い、次元削減に UMAP⁹、クラスタリングに HDBSCAN¹⁰を用いた。また、ベクトル変換に用いるモデルには ClimateBERT を用いた。

4.2 分析結果

2章で作業者がカーボンプライシングに関連していると判定した 466 本の論文に対し、BERTopic でトピック分析した結果を図表 6 に示した。その結果、39 のクラスタが作成され、それらは大きく 6 つのトピックに分かれた。

図表 6 BERTopic によるカーボンプライシング関連論文のトピック分析



(出所) 日興リサーチセンター作成

⁸ ある単語の重要性を示す統計量として、文章内の単語の頻度 (TF) と全文章における希少性 (IDF) を掛け合わせた TF-IDF がある。c-TF-IDF はこの指標をクラスタごとの計算に対応させたもの。

⁹ UMAP は 2018 年に提案された比較的新しい次元削減手法で、従来手法よりも高速かつ情報の損失が少ないのが特徴。

¹⁰ HDBSCAN は密度ベースのクラスタリング手法である DBSCAN を階層型クラスタリングに拡張した手法。様々な形状のクラスタを見つけることができ、どのクラスタにも属さないノイズとなっている文章を外れ値として検出できる。次節では外れ値を除いて分析している。

6 トピックの主な単語等を図表 7 にまとめた。トピック①には「DICE」「SCC」等の単語が頻出しており、統合評価モデル¹¹に関するトピックと考えられる。トピック②にはグローバルな規制や政治対応に関する単語が多かった。トピック③には CO₂ の削減や吸収に関する技術的な論文が多い。トピック④には気候変動による環境への影響等に関する単語が多かった。トピック⑤はラムゼーモデルなど伝統的な経済モデルを用いた論文が多い。トピック⑥には EU-ETS 等の排出権取引市場に関する分析が多かった。

2 章の人手によるラベル付けではカーボンプライシング関連論文の分類について「カーボンをプライシングしている」「カーボンプロキシプライシング」「なんらかのカーボン価格を置いて分析している」の3分類を考えたが、BERTopic を用いることで6種のより具体的な分類を見つけることができた。

図表 7 主要6トピック

トピック	主な単語	個数
① 統合評価モデル	DICE, SCC, IAMS, model, climate	52
② 政治・外交・規制	EPA, global, regulatory, environmental, principle	13
③ CO ₂ 削減・吸収技術	CO ₂ , capture, energy, electricity, fuels, renewable, technology	77
④ 気候変動の被害	warming, ecosystem, wetland, farm, agriculture, inequality	55
⑤ 伝統的な経済モデル	ramsey, aversion, policy, population, optimal, risk, damage	40
⑥ EU-ETSなど排出権取引	EU, ETS, leakage, lobbying, trading, market, tax	73

(出所) 日興リサーチセンター作成

5. おわりに

本稿では BERT を用いたカーボンプライシング関連論文の分析を行った。収集したデータを人手によりラベル付けし、学習データとして用いることで文章分類器を作成した。分類器はランダム時を上回る一定の精度を示した。

また、SHAP により、分類根拠を可視化し、カーボンプライシングに関するワードがモデルの判断に用いられていることを確認した。さらに BERTopic によるトピック分析を実施し、カーボンプライシング論文を6つの主要なトピックに分けることができることを示した。

課題として、データ数の少なさ、ラベリングにおけるノイズによる精度低下等が挙げられる。これらの問題への対処として、データの追加及びラベルの見直し等が必要になる。また、LLM の中でも BERT よりも規模の大きい GPT-3¹²等のモデルを用いることによっても、性能向上の可能性がある。以上を今後の課題としたい。

¹¹ 経済と気候の相互作用を1つにまとめたモデルで、米イェール大学のウィリアム・ノードハウス教授が1990年代に初めて開発し、後にノーベル経済学賞を受賞した。

¹² モデルの性能を示すパラメータ数では、BERT がおよそ1億~3億であるのに対して、GPT-3 が1750億となっている。ただし、こうした大規模なモデルの多くはインターフェースのみが公開されカスタマイズが難しい等の課題もある。

参考文献

川崎正勝（2023）, 「BERT を用いた気候変動リスク文の抽出と可視化」 <https://www.nikko-research.co.jp/library/12405/>

河南直希, 関洋平（2023）, 「ESG スコアの判断の根拠となるテキストの抽出」
https://www.jstage.jst.go.jp/article/jsaisigtwo/2023/FIN-030/2023_106/_article/-char/ja/

鈴木かの子, 松井孝典, 川久保俊, 増原直樹, 岩見麻子, 町村尚（2021）, 「BERT モデルを用いたSDGs に関するマルチラベル文書分類器の構築とマッチングシステムの開発」
https://www.jstage.jst.go.jp/article/pjsai/JSIAI2021/0/JSIAI2021_4H3GS11d01/_article/-char/ja/

Wang, P., Deng, X., Zhou, H. & Yu, S. (2019) , “Estimates of the social cost of carbon A review based on meta-analysis”
<https://www.sciencedirect.com/science/article/abs/pii/S0959652618334589>

Webersinke, N., Kraus, M., Bingler, J.A., & Leippold, M. (2021) , “ClimateBert: A Pretrained Language Model for Climate-Related Text” <https://arxiv.org/pdf/2110.12010.pdf>

(END)