

BERT を用いた ESG 関連文章の GRI スタンダード分類

Classification of ESG-related sentences into GRI standards using BERT

若本 亮佑^{1*1} 鵜飼 和渡^{1*2} 里中 裕輔^{1*3}

高木幸雄^{1*4} ファイサル ハディプトラ^{1*5}

Ryosuke Wakamoto¹, Kazuto Ukai¹, Yusuke Satonaka¹

Yukio Takagi¹, Faisal Hadiputra¹

¹ 株式会社電通国際情報サービス

¹ Information Services International-Dentsu, Ltd.

Abstract: 投資や融資の分野において、各企業が公開する ESG の情報が重要な判断材料となり始めている。しかし、今日の日本の ESG 情報開示は基準やフォーマットが統一されていないため、企業間の ESG 情報を統一された指標で比較することには時間と労力を要する。ここで、企業が公開するオープンデータを、既存の ESG 情報を開示する枠組みに則って整理することができれば、投融資における ESG リスクの判断をより迅速かつ正確に行うことができる。そこで、本研究では、企業が公開している ESG に関連する文章を ESG 情報開示枠組みの 1 つである GRI スタンダード (Global Reporting Initiative standards) に則って分類するモデルを作成する。具体的には、まず、企業が公開する GRI スタンダード対照表と、これらに紐づけられた統合報告書や Web ページを基にアノテーションを行うことで ESG 文章のデータセットを作成する。次に、日本語に特化した独自の BERT を基にファインチューニングを行うことで、GRI スタンダードに準拠した分類モデルを構築する。このとき、GRI スタンダードは項目が複数の粒度で枝分かれしているため、分類する粒度を変えた 2 つのモデルを併用する手法を提案する。

1. はじめに

環境 (Environment)・社会 (Social)・ガバナンス (Governance) に配慮した企業への投資を行う「ESG 投資」は年々世界的に広がっており[1]、企業は事業を推進していく中で ESG に対して取り組むことを求められている。投資や融資の分野において、ESG の情報開示に対応していないことは潜在的な経営リスクとみなされることから、ESG 情報は企業の持続性を評価する上で重要な判断材料となり始めている。

そのため、一部企業は ESG 情報を開示する枠組み (以降、ESG 情報開示枠組みと呼称する。) に則り、各社の ESG 情報を開示している。代表的な ESG 情報開示枠組みには、TCFD 報告書のほか、GRI スタンダード (Global Reporting Initiative Standards) や SASB スタンダード (Sustainability Accounting Standards Board Standards)、国際統合報告フレームワーク (The International <IR> Framework) などが存在し、これらの開示・発行を行う企業数は年々増加している[2]。また、自主的な開示のほか、法整備によ

って開示を求める動きもある。例えば、2022 年 4 月以降、一部上場会社は TCFD (Task Force on Climate-related Financial Disclosure, 気候関連財務情報開示タスクフォース) 提言に基づいたリスク情報の開示を実質的に義務付けられる。このように、企業に対して ESG に関する情報開示を求める動きが広がりを見せており、ESG 情報開示枠組みを活用した企業の ESG 情報の発信は今後も活発に行われていくものと考えられる。

しかし、今日の ESG 情報開示枠組みはフォーマットが統一されておらず、開示内容や基準も異なるため、企業間の ESG 情報を統一された指標で比較することは難しい。これによって、例えば、金融機関においては事業法人との対話に向けた事前準備に多大な時間と労力を要するなどの現状があることから、すべての企業の ESG 情報が、特定の 1 つの観点でまとめられていることにはニーズがある。このような課題に対し、企業が公開するオープンデータを既存の ESG 情報開示枠組みに則って整理することができれば、投融資における ESG リスクの判断をより迅

速かつ正確に行うことができる。

そこで、本研究では ESG 情報開示枠組みに焦点を当て、企業が公開している ESG に関連する文章を、ESG 情報開示枠組みの 1 つである GRI スタンダードの項目に紐づける AI モデルの作成を行う。

具体的には、まず、企業が公開する GRI スタンダード対照表と、これらに紐づけられた統合報告書や Web ページを基にアノテーションを行うことで ESG 文章のデータセットを作成する。次に、日本語に特化した独自の BERT を基にファインチューニングを行うことで、GRI スタンダードの項目に紐づける分類モデルを構築する。このとき、GRI スタンダードは項目が複数の粒度で枝分かれしているため、最も大きな分類項目として「E (環境), S (社会), G (ガバナンス), その他」(以降、大項目と呼称する。)の分類モデルを提案するとともに、「E」について最も細かい項目 (以降、小項目と呼称する。)で分類を行うモデルを併用する手法を提案し、それぞれについてベースラインとなるモデルの性能との比較を行う。

2. 関連研究

本研究では、企業のオープンデータを既存の ESG 情報開示枠組みに則って整理することを目的とするため、1) 広範な表現を持つ企業情報を分析対象データとすること、2) 分析対象データについて、GRI スタンダードのうち細分化された項目への分類を行うこと、について取り組んだ。

本研究と同様、ESG 情報に対して BERT を適用し、文章を抽出する研究としては[3]が挙げられる。[3]では、有価証券報告書の文章に対して GRI スタンダードに準拠してラベリングを行い、BERT を用いた ESG の分類を行っている。[3]と本研究は、文章分類タスクである点や使用するモデルについて共通しているが、ESG の分類についてマルチラベルデータとしている点異なる。他にも、分析対象を統合報告書や Web ページといった表現の幅がより広いものを選択している点、GRI スタンダードの小項目についての分類を実施している点についても異なっている。本研究で有価証券報告書を用いなかった理由は、小項目について有価証券報告書の参照回数が少ないことや、有価証券報告書内の「E」に該当する文章は定量的な記述が多く、自然言語処理に不向きであると判断したためである。

他にも、本研究で取り扱う ESG 情報の抽出に関連する研究として[4]がある。[4]は統合報告書を対象としており、データセットをルールベースで生成し、そのデータで学習させたモデルを用いて、ページ単位で ESG に関連した文章を抽出している。[4]は最

終的な分類目的を小項目に設定し、そのためのパイプラインを想定している点が本研究と異なる。また、Web ページを追加で使用していることによって単語の分布が異なっている点や、実験において使用しているモデルなどについても異なっている。

その他の ESG 情報の活用として、ESG スコアに着目した研究[5]もある。[5]は、ESG 格付け企業が公開しているスコアを基に、企業の信用格付けについて予測を行った。

3. 提案手法

3.1 データセット

本研究で使用する 24 社分のオープンデータの一覧を表 1 に示す。本研究では、GRI スタンダード対照表に基づき、それに基づく統合報告書、WEB ページ、CSR 報告書の各ページを主な ESG 情報として紐づけを行い、データセットの作成を行った。このとき、GRI スタンダード対照表のデータ形式が HTML であれば自動で、PDF であれば手動で GRI 項目に該当するページの紐づけを行い、PDF の情報は「Google Cloud Vision API」を用いて文章データを抽出した。

「Google Cloud Vision API」では認識された文章ブロック単位で情報が出力されるため、ページ単位で結合し各ページの文章データとした。

3.2 大項目分類タスク

本研究では、企業が公開する ESG 文章に対し、最終的に GRI スタンダード内の細分化された小項目への分類を行う。しかし、GRI スタンダード (2016 年版) の小項目は 148 個存在するため、各社のデータについてラベリングが必要であることを考慮すると、ESG 情報を 148 クラスの小項目で分類した場合に十分な学習効果が期待できない。そのため、ESG 情報を「E, S, G, その他」の大項目で分類した後、小項

表 1. 本研究で使用する企業のオープンデータ一覧

	データ形式	会社数	ページ数
GRI スタンダード 対照表	PDF / HTML	24	24
WEB ページ	HTML	17	791
統合報告書	PDF	13	1232
CSR 報告書	PDF	2	160

目へ分類するパイプラインを想定し、独立した分類タスクとした。

本章では、大項目分類に用いるデータセットとモデルについて述べる。

3.2.1 大項目分類用データセット

3.1 で作成したデータを基に、大項目分類用のデータセットの作成を行った。このとき、文章データは 1) 「Google Cloud Vision API」を用いて抽出した文章データ、2) クローリング後の HTML から抽出した文章データ、のいずれかであり、各企業が公開している PDF の視覚的情報や HTML 構造が多様であることを考えると、これらはそのまま学習データとして用いることができない可能性がある。そのため、以下のような処理を用いて文章データの整形を行った。

- ・ヘッダー/フッターの除去
- ・句読点を含まない文章の削除
- ・数値の割合が 50%を超える文章の削除
- ・10 文字以下の文章の削除

上記の処理後、句点区切りの文章に分割し、大項目ラベル付けを行った。このとき、本研究における GRI スタンドアードの大項目のラベル定義を表 2 に示す。各ページ文章は複数の GRI 項目に紐づけられる可能性があるため、表 2 に従い、分割前のページ文章で紐づけられていた GRI 項目に該当する大項目をマルチラベルでセットした。

3.2.2 大項目分類モデル

3.2.1 で作成した文章データセットについて、機械学習モデルを用いて大項目の分類を行う。

本研究における ESG 情報の分類は、文字数やデータ数を考慮し、句点区切りの文章で行っている。しかし、句点区切りの文章とした場合、本来ページと紐づけられている GRI 項目について意図した情報が必ずしも含まれているとはいえないため、人の目で見ただけに正しいラベルとなっていない可能性がある。そのため、使用する分類モデルはラベル誤りに対す

表 2. 本研究における大項目のラベル定義

ラベル	GRI スタンドアード中の 対応する項目番号
E	300 番台
S	400 番台
G	102-18 ~ 102-39
その他	上記に該当しない項目

るロバスト性が求められる。

ここで、自然言語処理において高い性能を持つモデルとして、Devlin らにより提案された BERT (Bidirectional Encoder Representations from Transformers) [6]がある。BERT は大規模な自然言語データセットを事前学習させることによって、分類や Question Answering といった様々な自然言語処理タスクに応用可能かつ、いずれのタスクでも高い精度を誇るモデルである。

BERT は事前学習によって、ラベル誤りへの一定のロバスト性を獲得する[7]ため、BERT に基づくモデルで実験を行うこととした。本研究では、軽量化され、BERT と同等以上の性能を持つ ALBERT[8]を用いる。

ここで、大項目分類タスクは以下の 2 つのモデルを用いて比較を行う。

1. 日本語事前学習済み ALBERT
2. 1.+ESG 情報再事前学習済み ALBERT

1.では日本語 Wikipedia で事前学習させた ALBERT をベースとし、最終層に全結合層を加え、大項目分類用データセットを用いて fine-tuning を行った。

また 2.では、日本語 Wikipedia で事前学習させた後に「3.1 で取得した ESG 情報」を用いて再度事前学習させた ALBERT をベースとし、最終層に全結合層を加え、大項目分類用データセットを用いて fine-tuning を行った。

3.3 小項目分類タスク

ESG 文章データの大項目分類が行われたことを想定し、小項目への分類を行う。

本研究では実験的に、大項目のうち「E」のみを使用して、「E」の小項目に特化した分類モデルの構築を行った。本章では、小項目分類に用いるデータセットとモデルについて述べる。

3.3.1 小項目分類用データセット

3.2.1 で作成したデータのうち、「E」の文章データについて小項目分類用のデータセットの作成を行った。

このとき、小項目が求める報告内容には定量的なものが存在する。本研究で扱うデータの特性上、定量的な情報は図やグラフで表されることが多いが、これらの情報は文章データとして取得できないため、正しいラベルのデータでも区別が難しい。そのため、「E」の小項目について、ラベルの除去と統合の 2 つの前処理を実験的に行うこととした。小項目の元ラベル及び除去対象、前処理後の小項目のラベル定義を表 3 に示す。項目で定められた報告内容が定量的かつ、本データセットにおいて項目の記載数が少な

表 3. 本研究における小項目のラベル定義

小項目ラベル	除去の有無	前処理後ラベル
301-1	-	301-x
301-2	-	
301-3	-	
302-1	-	302-x
302-2	-	
302-3	-	
302-4	-	
302-5	-	
303-1	-	303-1
303-2	-	303-2
303-3	-	303-3
303-4	-	303-4
303-5	-	303-5
304-1	-	304-x
304-2	-	
304-3	-	
304-4	-	
305-1	-	305-x
305-2	-	
305-3	-	
305-4	-	
305-5	-	
305-6	○	
305-7	-	
306-1	-	306-1
306-2	-	306-2
306-3	-	306-3
306-4	-	306-4
306-5	○	-
307-1	-	307-1
308-1	-	308-x
308-2	-	

いと判断した項目を除去対象とし、学習対象のラベルから除くこととした。さらに、小項目を抽象化した GRI 項目において、定量的な小項目が大半を占める場合、その項目が包有する複数の小項目ラベルを統合し新たなラベルとした。

3.3.2 小項目分類モデル

3.3.1 で作成した文章データセットについて、機械学習モデルを用いて小項目の分類を行う。ベースとする小項目モデルは、3.2 と同様の理由からベースのモデルは ALBERT とした。

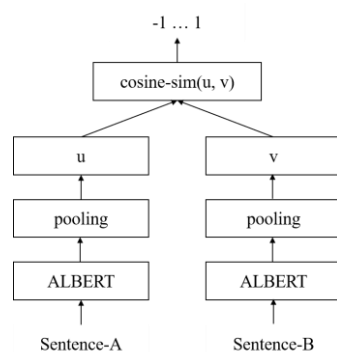


図 1. ALBERT をベースとした Sentence-BERT のアーキテクチャ

ここで、本研究が、1) 広範な表現を持つ企業情報を分析対象データとすること、2) 分析対象データについて、GRI スタンドのうち細分化された項目への分類を行うこと、を考慮すると、3.2.2 と同様の ALBERT に全結合層を加えた分類モデルを用いることは望ましくない。なぜなら、GRI スタンドを始めとした ESG 情報開示枠組みはいずれも改訂を続けており、これにより小項目は高頻度で増減する可能性があるためである。3.2.2 で用いたモデルは分類する小項目数を固定して学習させる必要があるため、改訂ごとに整理されたデータを大量に収集する必要があり、再学習も含めモデルの維持に多大な労力を要する。

そのため、小項目分類における学習モデルは Sentence-BERT[9] も用いることとした。Sentence-BERT は BERT をベースに構築されており、複数の文章を入力として文章の類似度をベクトルから計算する学習モデルである。本研究で用いた、ALBERT をベースとした Sentence-BERT の構造を図 1 に示す。

図 1 より、Sentence-BERT は入力として Sentence-A、Sentence-B を用いて、各入力文章を ALBERT でエンコードしたベクトルについてコサイン類似度を取り、教師ラベルの正否を基に学習を行う。このとき、Sentence-A は「3.3.1 のデータセットの文章」、Sentence-B は「GRI 項目の報告内容とその詳細を結合した文章」とし、Sentence-A と Sentence-B が対応している場合は教師ラベルを 1、それ以外の教師ラベルを 0 として与えた。テスト時は、Sentence-BERT を用いて、「3.3.1 のデータセットの文章」のうちテストデータとした各文章を特徴量 u に、すべての GRI 項目の報告内容をそれぞれ特徴量 v にエンコードした。このとき、すべての v に対して u とのコサイン類似度を計算し、その類似度が閾値を超えた場合には、 u の元文章 (Sentence-A) が v の元となる GRI 項目の報告内容について記述しているものとみなす。その

ため、Sentence-BERT においても分類結果はマルチラベルとなる。

上記の学習方式をとることにより、小項目について内容の変更や増減があった場合にも、ベクトル化した ESG 情報と GRI 項目の報告内容について、類似度を計算するだけで分類可否の判断が可能になる。

ここで、小項目分類タスクは以下の 4 つのモデルを用いて比較を行う。

1. 日本語事前学習済み ALBERT
2. 1.+ESG 情報再事前学習済み ALBERT
3. 日本語事前学習済み Sentence-BERT
4. 3.+ESG 情報再事前学習済み Sentence-BERT

1., 2. は 3.2.2 と同様のアーキテクチャ・学習方法に小項目分類用データセットを適用したものである。

また、3. では ALBERT を日本語の Wikipedia で事前学習させ、Sentence-BERT に組み込んだ後、小項目分類用データセットを用いて fine-tuning を行った。

4. では、日本語 Wikipedia で事前学習させた後に「3.1 で取得した ESG 情報」で再度事前学習させた ALBERT を Sentence-BERT に組み込み、小項目分類用データセットを用いて fine-tuning を行った。

4. 実験と結果

4.1 実験条件

3. の各手法の実験にあたって、3.1 のデータセットを基に、3.2.1 の大項目分類用データセット、3.3.1 の

表 4. 学習データ数

モデル	タスク	訓練データ数	検証データ数	テストデータ数
ALBERT	大項目分類	19486	1083	1083
	小項目分類	4387	241	255
Sentence-BERT	小項目分類	108447	1095	498

小項目分類用データセットを作成した。このとき、各モデル・タスクに用いた学習データ、検証データ、テストデータの数を表 4 に示す。小項目分類について、ALBERT と Sentence-BERT は同じデータを用いているが、Sentence-BERT は 2 つの文章データ (ESG 関連文章と GRI 項目の報告内容) について正負の組み合わせを行ったものを入力データとするため、GRI の各項目との組み合わせを網羅したことによりデータ数が ALBERT のデータセットと比べ増加している。

実験に用いる性能指標は多クラス正解率 (以降、Multi-class Accuracy と呼称する.) とした。これは、それぞれの小項目/大項目ラベルが正しく正解と予測される平均的な確率を表す。加えて、各 GRI 項目単位での Accuracy の算出を行ったほか、実用的な指標として完全正解率 (以降、All-match Accuracy と呼称する.) を使用した。本研究で用いた文章データはラベルにノイズを含んでいることが想定されるため、平均的な正解率が高いことだけでなく、全てのラベルが正解しているデータの割合が高いことも性能評価の観点の 1 つとした。

また、モデル学習時のパラメータ、バッチサイズは全てのモデルで 2 とし、エポック数は大項目分類タスクでは 4、小項目分類タスクでは 10 とした。テスト時に Sentence-BERT で識別したとみなすコサイン類似度の閾値は 0.65 として実験を行った。

4.2 結果

4.2.1 大項目分類タスク

大項目分類タスクの実験結果を表 5 に示す。表 5 より、日本語事前学習済みモデルに比べ、ESG をさらに学習させたモデルの性能のほうが高く高い結果となった。各項目について比較すると、「E, S, G」のいずれも性能は向上しており、「その他」についてのみ性能が低下する結果となった。このことから、事前学習によって専門的な単語表現の学習が行

表 5. 大項目分類モデルの実験結果

モデル	E Accuracy (%)	S Accuracy (%)	G Accuracy (%)	Others Accuracy (%)	Multi-class Accuracy (%)	All-match Accuracy (%)
ALBERT <ja>	87.81	81.16	84.40	75.07	82.11	49.58
ALBERT <ja+esg>	88.64	82.46	85.87	74.05	82.76	50.88

<ja> … 日本語事前学習済みモデル、<ja+esg> … 日本語事前学習後、ESG 情報再事前学習済みモデル

表 6. 小項目分類モデルの実験結果

モデル	301-x	302-x	303-1	303-2	303-3	303-4	303-5	304-x	305-x	306-1	306-2	306-3	306-4	307-1	308-x	Multi-class	All-match
	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.	Acc.
	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
ALBERT <ja>	71.09	75.00	86.72	89.06	80.08	84.77	86.33	93.75	73.83	88.67	76.17	88.28	89.06	82.81	82.42	83.20	27.34
ALBERT <ja+esg>	72.66	78.52	88.28	92.58	83.98	85.55	85.94	92.97	75.78	89.84	79.69	89.84	88.28	86.33	83.59	84.92	31.25
Sentence- BERT <ja>	76.10	78.31	86.75	90.36	80.52	84.74	83.73	91.97	75.90	85.74	80.32	88.76	86.95	84.54	85.14	83.99	36.14
Sentence- BERT <ja+esg>	78.51	79.72	87.35	91.37	83.94	87.15	86.14	91.77	78.92	86.35	83.13	91.77	89.36	87.15	86.75	85.96	38.96

Acc. … Accuracy, <ja> … 日本語事前学習済みモデル, <ja+esg> … 日本語事前学習後, ESG 情報再事前学習済みモデル

われたことにより、「その他」と比較して文章表現が限定的な「E, S, G」の分類性能が上がったものと考えられる。しかし、これらの結果は優位水準 5%を満たしておらず、有意差を確認することができなかった。

4.2.2 小項目分類タスク

小項目分類タスクの実験結果を表 6 に示す。表 6 より、ALBERT について<ja>と<ja+esg>を比較すると、4.2.1 と同様に、日本語事前学習済みモデルに比べ、ESG 情報を追加学習させたモデルが性能的に上回る結果となった。また、これらの結果については優位水準 5%を満たしており、有意差を確認することができた。Sentence-BERT についても同様に、ESG 情報を追加学習させたモデルの性能が高く、かつ有意差が確認された。そのため、小項目分類タスクにおいては一般的な日本語の事前学習だけでなく、ESG の専門的な用語を含んだ事前学習がより効果的であることが確認できた。この理由としては、小項目分類タスクが、大項目分類タスクと比べてより専門的な単語や文章表現の識別を必要とされるためと考えられる。

ここで、ALBERT と Sentence-BERT について比較すると、表 6 より、<ja>、<ja+esg>を問わず、Sentence-BERT の All-match Accuracy , Multi-class Accuracy のほうが高い結果となった。このことから、実用的なメリットだけでなく性能面においても Sentence-BERT が適しているといえる。また、大項目分類、小項目分類ともに各 GRI 項目の分類性能にばらつきがみられたことから、エンコードされた文章の特徴分布の大きさが GRI 項目ごとに大きく異なる可能性が

ある。そのため、クラス間の特徴量の距離を一定に抑えることができる Quadruplet Loss[10]などの学習方法を導入することで、更なる性能の向上が期待できる。

5. おわりに

本研究では、企業が公開している ESG に関連する文章からデータセットを作成し、GRI スタンドールの大項目及び「E」の小項目への分類を行った。

実験の結果より、ESG の専門的な用語を含んだ追加の事前学習は有効であり、より専門的な単語表現を識別において必要とする小項目分類タスクにおいて、その効果が顕著であることを示した。また、Sentence-BERT を用いた分類アーキテクチャが ALBERT より高い性能を持ち、かつ実用面での利点が高いことを示した。

今後の課題としては、ページ単位での評価が挙げられる。本研究では文章単位での評価を行ったが、GRI 項目は複数の文章を含む「ページ」と紐づけされた情報であるため、各文章の推論結果から複合的にページでの評価を行うことが求められる。

また、同時にデータの質の確保も重要な課題である。本研究で作成したデータセットは 1)文章単位のラベル付けによって生じるラベルのノイズ、2) データ特性による文章中のノイズ、を含んでいるため、性能向上にはデータ数及びバリエーションの拡充に加え、より質の良い教師データの確保が必要となる。特に、図表によって定量的に表される場合があるような一部の小項目については、データの質が低い可能性が高く改善の余地が大きい。

また、これらの課題の改善に加え、今後は、Sentence-BERT の予測結果が類似度となる特性を利用し、ESG 情報開示枠組みと企業の ESG 文章を用いた ESG 評価のスコアリングについても行いたい。

参考文献

- [1] Global Sustainable Investment Alliance: Global Sustainable Investment Review 2020, (2021)
<http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf>
- [2] ニッセイアセットマネジメント株式会社: GPIF 委託調査研究 ESG に関する情報開示についての調査研究 報告書 《概要》, (2019)
https://www.gpif.go.jp/investment/research_2019.pdf
- [3] 土橋諒太, 中田和秀: BERT を用いた有価証券報告書からの ESG 関連文抽出, 第 26 回 金融情報学研究会, (2021)
- [4] 河村康平, 高野海斗, 酒井 浩之, 永並健吾, 中川慧: 機械学習を用いた統合報告書の ESG 関連ページの推定, 第 27 回 金融情報学研究会, (2021)
- [5] 東出卓朗: Random Forest を用いた ESG 情報による信用格付予測モデル構築の提案と CDS 市場分析への応用可能性の検討, 先物・オプションレポート, Vol.30, No.12, (2018)
- [6] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv preprint arXiv:1810.04805, (2018)
- [7] 仁木裕太, 坂地泰紀, 和泉潔, 松島裕康: 再事前学習した BERT を用いた金融文書中の因果関係知識有無の判別, 第 34 回 人工知能学会全国大会, (2020)
- [8] Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut: Albert: A lite bert for self-supervised learning of language representations, arXiv preprint arXiv:1909.11942, (2019)
- [9] Nils Reimers and Iryna Gurevych: Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, arXiv preprint arXiv:1908.10084, (2019)
- [1 0] Weihua Chen, Xiaotang Chen, Jianguo Zhang, and Kaiqi Huang: Beyond Triplet Loss: A Deep Quadruplet Network for Person Re-Identification, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 403-412, (2017)