

決算短信における見通し文と結果文のアラインメント

平松 悠太¹ 小川 泰弘² 外山 勝彦¹

¹ 名古屋大学大学院情報学研究科 ² 名古屋市立大学データサイエンス学部
hiramatsu.yuta.m0@es.mail.nagoya-u.ac.jp ogawa@ds.nagoya-cu.ac.jp
toyama@is.nagoya-u.ac.jp

概要

企業が公開する決算短信には、事業の結果について述べる文のほかに、将来の業績や事業計画などを見通す文が含まれる。事業の見通しとその結果を比較できれば、企業ごとに見通しの傾向を知ることができる。それにより、投資家は企業について、より正確に予測できる。しかし、投資初心者が決算短信を分析することは容易でない。本研究では、投資初心者の支援のために、時系列に並んだ決算短信の間で、それらに出現する見通し文と結果文をアラインメントする。具体的には、文の類似度や大規模言語モデル (LLM) を単独で用いる手法に加えて、それらを併用する手法を検証する。その結果、LLM を用いる手法が最も有効であることを明らかにした。

1 はじめに

企業が公開する財務資料は、中・長期投資家にとって株式売買のための重要な情報源となる。特に、決算短信や有価証券報告書には企業の財務状況や業績などが記述される。その中でも決算短信は、決算後 45 日以内の公開が求められる。ところが、日本の上場企業の約 7 割が 3 月決算であると言われており [1]、多数の決算短信が同時期に一度に公開されることになる。近年、日本政府は投資を推進している [2] が、初心者の投資家にとって、多数の決算短信から必要な情報を素早く抽出することは容易でない。したがって、そのような作業を自動化することは投資家に対して有効な支援になる。

先行研究では、決算短信における重要な情報として、業績予測文や見通し文を挙げている。北森ら [3] は業績予測文を「業績要因を含む今後の業績予測を示す文」と定義し、例文「音楽分野音楽制作が好調であることなどにより、分野全体の売上高は 10 月時点の想定を上回る見込みです」を挙げている。さらに、平松ら [4] は日本や世界全体の景気を予測す

る文を景気予測文、業界全体の業績を予測する文を業界予測文、自社の事業について計画を述べる事業計画文とそれぞれ定義している。各文の例を付録の表 4 に示す。さらに、業績予測文にそれらの文を加えて見通し文と定義し、連続する年度の決算短信の間で見通し文どうしをアラインメントしている。

しかし、見通しの傾向は企業ごとに異なるため、それだけでは十分でない。例えば、楽観的な見通しを示す企業があれば、堅実な見通しを示す企業もある。この傾向の差は投資家が業績をより正確に予測するための妨げとなるため、見通しとその結果を企業ごとに比較して、傾向を分析することが必要である。そこで、本研究では、業績や事業について過去の実績を述べる文を新たに結果文と定義し、見通し文と結果文をアラインメントする。

アラインメントには平松ら [4] の手法を用いる。具体的には、文の類似度と LLM をそれぞれ単独で用いる手法に加えて、それらを併用する手法がある。しかし、アラインメントの対象は、平松らでは見通し文どうしの間であるが、本研究では見通し文と結果文の間である。そのため、平松らの手法が本研究の対象でも有効であるかどうかを検証する。

なお、決算短信は四半期単位で公開されるものであるが、本稿では年度単位で比較し、四半期単位の比較には今後取り組む予定である。

さらに、本稿では見通し文のみ人手で抽出し、見通し文以外の文 (非見通し文) を結果文の候補として与える。その理由は、非見通し文のほとんどは結果文であり、人手ですべての結果文を非見通し文から抽出することはコストが高いからである。

2 先行研究

本節では、財務資料を対象としたアラインメントの先行研究について述べる。

2.1 有価証券報告書における差分抽出

鈴木ら [5] は、時系列に並んだ有価証券報告書において、ゲシュタルトパターンマッチングによる類似度を用いて文をアラインメントし、差分を抽出するタスクを提案している。鈴木らがアラインメントした例を付録の表 5 に示す。なお、鈴木らがアラインメントした文の組には主語が異なる文があるが、そのような組は記述の対象も異なる。本研究が目的とする投資家への情報支援において、見通しとその結果を比較するためには、主語が異なる文をアラインメントすることは好ましくない。

2.2 決算短信における見通し文のアラインメント

平松ら [4] は、連続する年度の決算短信において、見通し文をアラインメントする手法を提案している。具体的には、文の類似度と LLM をそれぞれ単独で用いる手法と、それらを併用する手法の三つである。併用手法では、文の類似度を用いて再現率重視でアラインメントした後、LLM に適切なアラインメントのみを選択させている。平松らがアラインメントした例を付録の表 6 に示す。しかし、見通しの傾向は企業ごとに異なるため、見通し文どうしのアラインメントだけでは十分でない。

3 見通し文と結果文のアラインメント

本節では、結果文を定義した後、見通し文と結果文における適切なアラインメントについて述べる。

3.1 結果文

本研究では、世界や日本全体の景気、業界全体の業績、自社の業績、自社の事業について企業が結果を述べる文を新たに結果文と定義する。結果文が見通し文と異なる点は、前者が過去の事実について記述し、後者が未来の予測について記述する点である。以下に結果文の例を示す。

- 当期における世界経済としては、感染症の影響により個人消費が低迷しました。
- 外食業界でも感染防止により客足が遠のき、厳しい環境が続きました。
- 当社の売上高は前期比 10%減の 120 百万円、営業利益は前期比 3%減の 15 百万円でした。
- 当社ではテイクアウトに注力し、新メニューの開発や人材・設備投資に取り組みました。

3.2 見通し文と結果文の適切なアラインメント

本稿では、平松ら [4] と同様に、以下の二つの条件を満たすものを適切なアラインメントと定義する。

1. 記述の対象が同じである。
2. アラインメントの粒度が可能な限り細かい。

まず、1 の例として、文「感染症の影響で飲食分野の売上が減少する見込みです」と文「SNS の認知度低下のため飲食分野の売上が低迷しました」を考える。この 2 文が記述する対象はどちらも飲食分野の売上であるため、アラインメントすることが適切である。一方、文「感染症の影響で売上が低迷する見込みです」と文「感染症の影響で売上原価が上昇しました」はどちらも感染症について言及しているが、記述の対象は売上と売上原価で異なるため、アラインメントすることは適切でない。

次に、2 について、本稿では平松ら [4] と同様に一対複数のアラインメントを許す。しかし、一対一と一対複数のアラインメントがどちらも可能な場合、前者を優先する。例えば、二つの短信にそれぞれ売上高と営業利益について記述する文が出現した場合、それぞれ別々にアラインメントする。

4 提案手法

本研究では、平松ら [4] が提案した三つの手法が見通し文と結果文のアラインメントにおいても有効かどうかを検証する。なお、本研究では PDF 形式¹⁾の決算短信から、平松ら [6] の方法で各文のテキストデータを抽出した。²⁾

4.1 手法 1：文の類似度を用いる手法

手法 1 では、文の類似度を用いる。具体的には、連続する年度の決算短信 A, B について、A に出現するすべての見通し文と B に出現するすべての非見通し文の間で類似度を計算する。次に、A の各見通し文に対して、B の非見通し文のうち類似度の上位 k 文をそれぞれアラインメントする。なお、見通し文に対して非見通し文をアラインメントする場合と、その逆の場合では結果が変わる。そのため、A から B, B から A の両方向で実験する。文の類似度は以下の三つの方法で計算する。

- 1) 決算短信は従来 PDF 形式のみで公開されてきたが、2024 年 6 月末に終了する事業年度の決算に係る決算短信からは HTML 形式での提出が求められる。しかし、現時点では、決算短信を HTML 形式で取得できる企業の割合は十分でない。
- 2) 決算短信の取得には東証上場会社情報サービスを用いた。

1. 文を形態素に分割して自立語を抽出し、形態素の Jaccard 類似度を用いる方法 (Morph)
2. 文を事前学習モデル BERT に入力し、最終層の入力を BERT 埋め込みとして獲得して、そのコサイン類似度を利用する方法
3. ゲシュタルトパターンマッチングを利用した Python の関数である `diffib.SequenceMatcher` の出力を文の類似度とする方法 [5]

4.2 手法 2 : LLM を用いる手法

手法 2 では、LLM を利用する。具体的には、決算短信 A に出現する見通し文のリストと同 B に出現する非見通し文のリストをともに与えて、適切なアラインメントの集合を出力させる。LLM を利用する理由は、1 文単位のアラインメントではなく、文脈を考慮してアラインメントできると考えられるからである。例えば、文「新型コロナウイルスの影響により、純資産は 10% 増となる見込みです」と文「新型コロナウイルスの影響により、流動負債は 10% 増となりました」は記述の対象が異なるためアラインメントすべきでないが、文の類似度でそれを判断することは容易ではない。しかし、後者の文の前後に純資産の結果を述べる文があれば、純資産について見通す文とその結果を述べる文がアラインメントされる。そのため、前述の 2 文はアラインメントしないと判断することが容易になると期待できる。

4.3 手法 3 : 文の類似度と LLM の併用手法

手法 3 では、文の類似度と LLM を併用する。具体的には、次の二段階でアラインメントする。

- (1) 手法 1 を用いて、再現率重視でアラインメントの候補を挙げる。
- (2) (1) の出力を手法 2 (LLM) における入力とし、適切なアラインメントを選択する。

なお、(1) では手法 1 の k の値を大きくすることにより、再現率重視のアラインメントを行う。

4.4 評価データセット

見通し文と結果文のアラインメントを評価するためのデータセットは、知る限りでは存在しない。そこで、人手によるアノテーションを行い、評価データセットを新たに作成する。具体的には、まず、同一企業かつ連続する年度の決算短信から、人手で見通し文を抽出する。さらに、適切なアラインメント

表 1 見通し文と非見通し文の文数

	前年度分		次年度分	
見通し文	2019	83	2020	119
	2021	194	2022	198
	合計	277	合計	317
非見通し文	2019	1,512	2020	1,541
	2021	1,329	2022	1,427
	合計	2,841	合計	2,968

表 2 LLM に与えるパラメータ

パラメータ名	値
シード値	12345678
temperature (温度)	1e-30

を人手でアノテーションする。

5 実験

評価実験を行い、手法の有効性を検証した。

5.1 評価指標

本稿では、各手法の性能を F1 値のマイクロ平均とマクロ平均で評価する。なお、一对複数のアラインメントは一对一に展開してから評価する。

5.2 実験に使用したデータセット

本実験では、4.4 節で述べた評価データセットを企業 30 社の通期決算短信計 60 件から作成した³⁾。なお、そのうち 15 社は、前年度分として 2019 年度、次年度分として 2020 年度のものを用い、残る 15 社はそれぞれ 2021 年度、2022 年度のものを用いた。データセット中の見通し文と非見通し文の数を表 1 に示す。アラインメント数は 163 組であり、一对一のアラインメントに展開すると 415 組となった。

5.3 実験設定

手法 1 のうち、BERT 埋め込みの獲得には東北大学が公開する日本語 BERT⁴⁾を使用した。なお、本実験では BERT のファイン・チューニングは行っていない。次に、手法 2 では、LLM として GPT-4 Turbo と GPT-4o を利用し、各モデルを用いてそれぞれ 3 回ずつ出力させた。API に与えたパラメータを表 2 に示す。プロンプトには以下の要素を記述した⁵⁾。

3) データセットは GitHub において公開予定である。

https://github.com/nnu257/alignment_dataset

4) <https://www.nlp.ecei.tohoku.ac.jp/news-release/3284/>

5) プロンプトの全文は Github で公開している。

<https://github.com/nnu257/Alignment/>

表3 評価実験の結果 (マイクロ平均)

手法	モデル	direc	k	Re.	Pr.	F1
手法1	Morph	f	1	0.24	0.36	0.29
	BERT 埋め込み	f	1	0.20	0.30	0.24
	SequenceMatcher	f	1	0.21	0.32	0.26
	Morph	f	2	0.34	0.26	0.30
	Morph	b	1	0.42	0.06	0.11
	Morph	b	2	0.55	0.04	0.08
手法2	GPT-4 Turbo	-	-	0.48	0.49	0.48
	GPT-4o	-	-	0.30	0.53	0.38
手法3	Morph+GPT-4 Turbo	f	1	0.39	0.53	0.45
	Morph+GPT-4 Turbo	f	2	0.38	0.50	0.43
	Morph+GPT-4 Turbo	f	3	0.41	0.52	0.46
	Morph+GPT-4 Turbo	f	4	0.42	0.47	0.45
	Morph+GPT-4 Turbo	f	5	0.43	0.46	0.45
	Morph+GPT-4 Turbo	f	6	0.42	0.46	0.44

- ペルソナの設定
- 見通し文と結果文の説明
- 適切なアラインメントの定義
- アラインメントの例 (5-shot)

最後に、手法3の(1)には、類似度を用いる手法1において最も性能が高いMorphを用いた。

5.4 結果

実験結果のうち、マイクロ平均の値を表3に、マクロ平均の値を付録の表7にそれぞれ示す。なお、表中のdirecはアラインメントの方向であり、見通し文から結果文へのアラインメントはf、その逆はbである。さらに、手法2の結果はLLMで3回出力した結果の平均値を記述した。

まず、手法1ではMorphの性能が最高であった。さらに、アラインメントの性能やkの値について比較すると、見通し文から結果文へアラインメントすることが有効であり、どちらの方向においてもkの値によって性能は大きく変化しなかった。次に、手法2は手法1の性能を大きく上回っており、GPT-4 Turboを用いると性能が特に向上した。最後に、手法3はいずれのkの値においても、手法2を下回った。各手法における結果の比較を付録Bに示す。

5.5 考察

まず、手法1では表層的な手法の性能が他の手法の性能を上回った。これは、決算短信が発表のつど、一から作られるのではなく、以前の決算短信を

参考にして作られるために、共通する言い回しが出現しやすいという性質を示唆している。特に、見通し文や結果文は互いを参考にして記述内容が決定されている可能性がある。

次に、見通し文から結果文の方向でアラインメントする手法の性能が高い理由は、見通し文と結果文の出現数が原因だと考えられる。評価データにおいて、決算短信1件には見通し文が約5文、結果文が約50文それぞれ含まれ、見通し文1文に対して結果文は約2文がアラインメントされている。一方、結果文には見通し文がアラインメントされない文が多い。したがって、見通し文に対して、類似度が高い結果文をアラインメントする方向が有効である。

さらに、手法2の性能が手法1の性能を上回った理由は、LLMが文脈を考慮したからだと考えられる。実際に、Morphのアラインメントには、類似度は高いが、記述する事業セグメントが異なる文の組があった。一方、手法2では事業セグメントを考慮して正しいアラインメントがされていた。

最後に、平松ら[4]の結果とは異なり、手法3の性能は手法2の性能を上回らなかった。その理由は、平松らのアライメントの対象は見通し文どうしの間であるが、本研究では見通し文と結果文の間であり、文のカテゴリが異なるためであると考えられる。カテゴリが異なる文の表層的な類似度は低いため、それを利用する手法1の有効性が低下する。平松らの研究では手法1と手法2の性能が同等であったことから、併用により両手法の強みを活用できたが、本研究では手法1の性能が低いため、併用は有効ではなかったと考えられる。

6 まとめ

本研究では、見通し文と結果文のアラインメントに取り組んだ。具体的には、平松ら[4]が見通し文どうしのアラインメントに利用した手法の性能を検証した。その結果、LLMを単独で用いる手法の性能が最も高いことを明らかにした。

本稿では、文を決算短信から抽出して利用しているが、決算短信には見出しや段落などの情報もある。そこで、より正確なアラインメントのために、それらを用いたアラインメント手法を提案することは今後の課題である。

謝辞

本研究は JSPS 科研費 JP23K25155 の助成を受けた。

参考文献

- [1] ニッセイ基礎研究所. 3月期決算の会社が多いというのは本当か、またその理由は?, (2024-12 閲覧). <https://www.nli-research.co.jp/report/detail/id=61217?pno=2&site=nli>.
- [2] 内閣官房. 資産所得倍増プラン, (2024-12 閲覧). https://www.cas.go.jp/jp/seisaku/atarashii_sihonsyugi/pdf/dabiplan2022.pdf.
- [3] 北森詩織, 酒井浩之, 坂地泰紀. 決算短信 PDF からの業績予測文の抽出. 電子情報通信学会論文誌, Vol. J100-D, No. 2, pp. 150–161, 2017.
- [4] 平松悠太, 小川泰弘, 外山勝彦. 決算短信における見通し文のアラインメント. 電子情報通信学会技術研究報告, Vol. 124, No. 173, pp. 69–74, 2024.
- [5] 鈴木雅弘, 坂地泰記, 和泉潔. 時系列に並んだ金融文書からの差分抽出タスクの提案. 人工知能学会全国大会論文集, Vol. 37, p. 1E3GS602, 2023.
- [6] 平松悠太, 小川泰弘, 外山勝彦. 決算短信における業績予測文のブートストラップを用いた抽出と極性付与. 電子情報通信学会技術研究報告, Vol. 123, No. 176, pp. 57–62, 2023.

A 例文

表4 見通し文の例

分類	文
業績予測文	売上高は大幅未達の見込みです。
業界予測文	小売業界におきましては、人手不足に伴う人件費の上昇により、厳しい経営環境が続くと予想されます。
景気予測文	わが国経済は、新型コロナウイルスの流行により、先行きは不透明な状況にあります。
事業計画文	生産能力の強化のための投資を行い、市場ニーズへの対応力向上を推進します。

表5 鈴木ら [5] がアラインメントした文対の例 (1-A と 1-B, 2-A と 2-B をそれぞれアラインメントしている)

番号	文
1-A	そのため、インターネット広告、展示会への出店等を通じて知名度向上に努めてまいりました。
1-B	そのため、新たなテレビCMを始め、交通広告、インターネット広告、展示会への出店等を通じて、安否確認サービスの知名度向上に努めてまいりました。
2-A	純資産は、前連結会計年度末に比べて [NUM] 百万円の増加となりました。
2-B	流動負債は、前連結会計年度末に比べて [NUM] 百万円の増加となりました。

表6 平松ら [4] がアラインメントした文対の例 (3-A, 3-B に対して 3-C をアラインメントしている)

番号	文 (一部略)
3-A	オペレーション改革に取り組みます。
3-B	さらに、生鮮部門の小型加工センター稼働や母店製造供給など将来を見据えた取り組みを一日も早く軌道に乗せ、人手不足やコスト増の状況においても、安定した利益を確保できる仕組みづくりを目指してまいります。
3-C	さらに、前事業年度に引き続き将来を見据えた取り組みも進め、当社の強みである店舗主導主義において、標準化された仕組みを土台に各店舗の独自性が最大限発揮できるようなオペレーションの再構築を進め、業務効率の改善と標準化に取り組み、より一層の生産性向上に注力してまいります。

B アラインメントの出力例

見通し文と結果文のアラインメント例を表8に示す。4-Aは売上高と利益の両方について記述しているが、4-Bは売上高、4-Cは利益についてそれぞれ単独で記述している。そのため、4-Aに対して4-Bと4-Cの2文をアラインメントすべきである。実験の結果、手法2は正解のアラインメントを出力したが、手法1と手法3では正解とは異なる文、すなわち、事業セグメント別の売上や利益について述べる文をアラインメントした。その原因は、表層的な類似度を用いていることだと考えられる。

表7 評価実験の結果 (マクロ平均)

手法	モデル	direc	k	Re.	Pr.	F1
手法1	Morph	f	1	0.32	0.40	0.34
	BERT 埋め込み	f	1	0.27	0.38	0.31
	SequenceMatcher	f	1	0.29	0.37	0.31
	Morph	f	2	0.43	0.31	0.34
	Morph	b	1	0.53	0.10	0.15
手法2	Morph	b	2	0.68	0.08	0.11
	GPT-4 Turbo	-	-	0.56	0.61	0.56
手法3	GPT-4o	-	-	0.39	0.58	0.43
	Morph+GPT-4 Turbo	f	1	0.50	0.66	0.54
	Morph+GPT-4 Turbo	f	2	0.46	0.59	0.49
	Morph+GPT-4 Turbo	f	3	0.52	0.62	0.54
	Morph+GPT-4 Turbo	f	4	0.52	0.56	0.51
	Morph+GPT-4 Turbo	f	5	0.57	0.55	0.52
	Morph+GPT-4 Turbo	f	6	0.57	0.56	0.54

表8 見通し文と結果文におけるアラインメントの例

番号	文 (一部略)
4-A	業績見通しにつきましては、連結受注高は20,200百万円、連結売上高は20,000百万円、営業利益2,800百万円、経常利益2,860百万円、親会社株主に帰属する当期純利益1,940百万円を予定しております。
4-B	当連結会計年度における当社グループの業績は、連結受注高は18,448百万円、連結売上高は19,231百万円となりました。
4-C	利益面では、積極的な人材投資、IT投資等により、営業利益は1,934百万円、経常利益は2,012百万円、親会社株主に帰属する当期純利益は1,726百万円となりました。