

NIE 支援に向けた新聞社説における事実・意見判別手法の検討

川原拓実¹ 安藤一秋¹¹ 香川大学創造工学部

{s22t312, ando.kazuaki}@kagawa-u.ac.jp

概要

近年、新聞を活用した教育 (NIE) が多くの教育機関で実施され、その有用性が報告されている。一方で、教員が適切な教材記事を選定する負担が大きいという課題がある。そこで本研究では、NIE において頻繁に行われる「読み比べ」に着目し、教育に適した記事ペアを推薦するシステムの構築を目指す。本稿では、システム構築の基盤として、特に意見と事実の判別が困難な「社説」を対象とした事実・意見分類モデルについて検討する。言語学的知見に基づくルールベースモデルと、大規模言語モデルを用いた機械学習モデルを構築し、社説記事の事実・意見分類性能を比較検証した。その結果、機械学習モデルの F1 値は事実文で 0.9243、意見文で 0.8291 を示し、文脈を考慮した学習の有効性が確認された。

1 はじめに

近年、新聞を教材として活用する教育である NIE (Newspaper in Education) が、全国の教育機関で広く実施されている [1, 2]。小学校の NIE 実施報告書 [3] によると、NIE の実践は、児童生徒の表現力や思考力、主体性などの能力の育成に効果があることが示されている。特に小学校では、「新聞記事の読み比べ」実践が多く、メディアリテラシーの育成に重点が置かれている。

しかし、「読み比べ」を実践するためには、教員が複数の新聞社から適切な記事ペアを選出する必要がある。教員は、膨大な記事の中から NIE に適したトピックを探し出し、意見や視点の異なる記事をペアリングするために内容を精読する必要がある。この教材準備の負担は極めて大きく、実践現場において深刻な課題になっている。

そこで本研究では、小学校教員による記事選定の支援を目的とし、小学校での NIE に適した異なる新聞社の記事ペア (読み比べ記事) を推薦するシステムの構築を目指す。小学校における NIE では、一般

記事と社説が利用されているが、一般記事に比べて筆者の主張が顕著に表れる「社説」は、視点や意見の対比を学ぶ読み比べ教材として重要度が高く、事実と意見の判別モデルを構築する上でもより有用な情報源といえる。

そこで本稿では、システム構築の第一段階として、各社の主張が最も明確に表れる「社説」に着目し、社説テキストから事実と意見を分類するモデルについて検討する。社説では、文末のモダリティによって明示的な意見が述べられる場合だけでなく、事実を記述する単語の選択に書き手の主観的な評価が含まれる場合も多い。そのため、単純なキーワードの一致による判別は困難であり、文脈を考慮した高精度な分類手法が必要とされる。そこで、ルールベース手法と機械学習手法の比較実験を通じて、社説における事実・意見分類に最適な手法とその課題を明らかにする。

2 関連研究

意見抽出は、多くの分野で研究されている。新聞ドメインにおいては、韓ら [4] が社説から書き手の価値観を推定するため、意見文を抽出・連結する手法を提案している。また、LLM を用いて YouTube のコメントから主要意見を抽出する手法を提案する研究 [5] も見られる。しかし、本研究の目的は、価値観分析や主要意見の抽出ではなく、教育現場における記事の読み比べ支援にある。効果的な読み比べを実現するためには、その基盤として、記事を構成するテキストを「客観的な事実」と「筆者の意見」に正確に区別して扱うことが不可欠である。そこで本稿では、意見のみならず、事実・意見双方の分類精度に着目して、構築したモデルを評価する。

3 読み比べ実践の分析

小学校で NIE を実践する教師に読み比べに適した記事を推薦するシステムの仕様を決定するため、まずは小学校で実践される読み比べについて分析

する。

過去 10 年間の NIE 実践報告書が、Web 上のデータベース [6] として公開されている。このデータベースを利用して、小学校における「読み比べ」授業で 10 年間に実施された 120 回のうち、分析可能な 59 回の授業について「手法」と「ねらい」を分析する。

分析の結果、「手法」としては「同一の話題を比較する」が最多で 35 回を占め、小学校の NIE において中核的な位置づけであることを確認した。また、「ねらい」としては、「多面的・多角的に物事を捉える力を養う」が最多で 14 回を占め、同一の話題について多様な視点に触れることを通じて、そのような能力を育成することが目的であることを確認した。以上の結果から、本研究では、メディアリテラシーの向上に向け、多面的・多角的な視点の獲得に資する「同一の話題を比較する」実践に着目する。

また、この実践を工夫して取り組んだ結果、教育に価値がある事例も報告されている [7]。一般記事を用いた「読み比べ」の実践例が多くみられる一方、社説を活用した実践例も一部存在する。本研究では、一般記事と社説の両方を活用した「読み比べ」に適した記事推薦システムの実現を目的とするが、本稿では、まず、多角的な視点の提示において最も重要な役割を果たし、かつ新聞社の意見が最も明確に反映される社説に注目する。

4 記事推薦システムの概要

NIE における読み比べ実践の分析結果に基づき、本研究では異なる新聞社の記事群から、読み比べに適した記事ペアを自動抽出・推薦するシステムの構築を目指す。図 1 に本システムの全体の構成図を示す。前述の通り、「事実」と「意見」の正確な分離は、後続の「話題一致判定」および「視点対比判定」の精度を左右する基盤技術である。そこで本稿では、システム構築の第一段階として、「事実・意見分類モデル」に焦点を当て、社説を対象とした分類モデルについて検討する。

5 データセットの構築

本研究は、最終的に複数紙への適用を想定しているが、本稿ではシステム構築の第一段階として、毎日新聞の社説を用いてデータセットを構築し、社説テキストに対する事実・意見分類モデルについて検討する。以降、正解ラベルを付与する際のアノテ

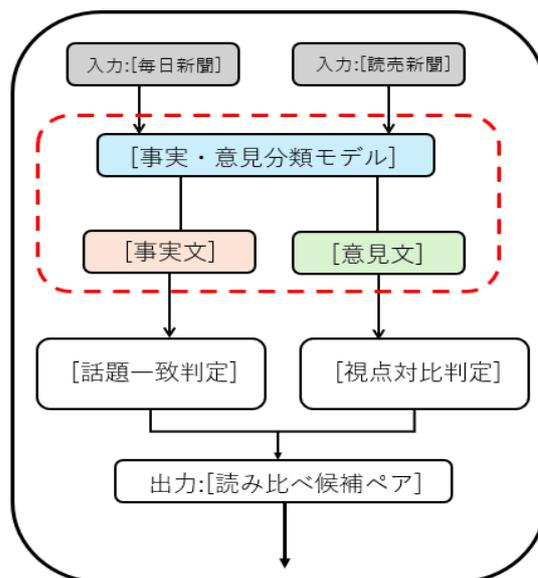


図 1 提案システムの全体処理フロー

ションと、事実・意見文の分析について述べる。

5.1 アノテーション

2024 年版毎日新聞記事データ集 [8] の社説計 641 記事に対して、それぞれの本文を一文単位に分割し、合計 16,756 文を抽出した。そのうち、75 記事 1,859 文に対して、「事実文」または「意見文」のラベル付与を実施した。各ラベルの定義は以下の通りである。

- 事実文 (Fact) : 客観的な出来事や状況を記述するなど、書き手の主観を含まない文
- 意見文 (Opinion) : 筆者の推測や評価、感情など、書き手の主観的な要素を含む文

アノテーションは、基本的に一人の作業員（筆者）によって実施した。しかし、社説には事実と意見の境界が曖昧な文も存在する。そこで、一個人の主観による偏りを防ぎ、データセットの客観性を担保するため、判断が困難な事例については第三者（研究室のメンバー 2 名）と協議し、合意形成を経た上で最終的なラベルを付与した。最終的に、事実文 1,314 文、意見文 545 文からなるデータセットが構築できた。

5.2 事実・意見の分布分析

構築したデータセットを用いて、社説記事内における事実文と意見文の出現傾向を分析した。図 2 に、記事の先頭から末尾にかけて各記事ごとのラベ

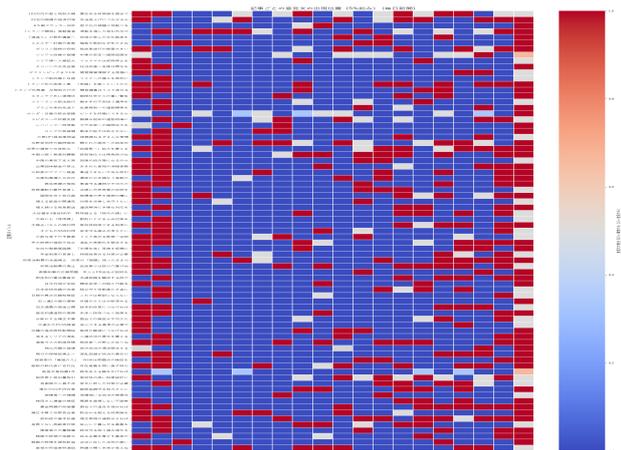


図2 社説における事実・意見文の分布

ル分布のヒートマップを示す。

図より、一般的に意見が述べられるとされる記事の冒頭や末尾だけでなく、途中にも意見文が見られた。この結果は、位置情報だけで単純に事実文と意見文を分離できる構造ではないことを示唆している。したがって、特定の位置に存在する文を抽出するルールベース手法では不十分であり、文脈を考慮して文単位で分類するモデルの構築が不可欠である。

6 事実・意見分類手法の検討

6.1 ルールベース手法

新聞記事、特に社説に現れる意見文においては、文末表現や特定の評価語に書き手の主観が表れやすい。そこで本手法では、詳細な意見表現辞書を構築し、それを用いたパターンマッチングで分類する。

6.1.1 辞書の構築

意見表現辞書の作成にあたっては、国立国語研究所の「分類語彙表 (WLSP)」[9]を利用した。

構築の手順として、まず分類語彙表から主観的評価・感情・思考・意思など、意見文に含まれやすい意味カテゴリを選択した。次に、各カテゴリに含まれる語をシード（種）としてシソーラス[10]を用いて類義語を検索し、語彙を拡張した。このプロセスを経て、最終的に1,042語からなる意見表現辞書を構築した。

6.1.2 判定アルゴリズム

対象文において、単に意見表現辞書内の語の有無を確認するだけでなく、それが「記者の意見」な

か「他者の意見の引用」なのかを区別する必要がある。そこで、(1) 鍵括弧で囲まれた文は除外する、(2) 辞書内の語とのマッチングは文末から5トークン以内限定する、という2つの制約を適用する。そして、これらの制約を満たす文を「意見文」、それ以外を「事実文」として分類する。

6.2 機械学習手法

単語の表面的な出現パターンだけでなく、文脈を考慮して事実・意見文を分類することを目的に、日本語事前学習モデル「Sarashina-Embedding-v1-1B」[11]を採用する。本手法では、Decoderモデルを用いるため、文末の[EOS]トークンの隠れ状態を全結合層に入力し、「事実」と「意見」で二値分類する。データセットでファインチューニングすることで、社説特有の言語表現への適用を図る。

7 評価実験

7.1 実験設定

本研究で構築したデータセット（社説75記事、1,859文）には、事実文1,314文と意見文545文が含まれている。データの偏りを排除し、モデルの汎化性能を評価するために、5分割交差検証で評価する。各検証において、学習データのみ事実文をアンダーサンプリングし、意見文と同数になるよう学習させる。評価用データは、全事例に対する予測結果を得るため、アンダーサンプリングは行わず元の分布のまま使用する。

評価指標には、Precision, Recall, F1値を用い、全データの予測結果を統合したマイクロ平均として算出する。各分割の学習において、検証データに対する意見クラスのF1値が最大となったエポックのモデルを採用する。なお、最適化アルゴリズムには、AdamWを用い、エポック数を10、バッチサイズを8、学習率を $2e-5$ 、最大系列長を512に設定して実験する。

7.2 評価結果

ルールベース手法と機械学習手法の比較結果を表1に示す。

実験の結果、機械学習手法は、ルールベース手法と比較して、ほぼすべての評価指標において大幅な性能向上を示した。特に、意見文において、ルールベース手法ではF1値が0.3480にとどまっていたの

表 1 事実・意見分類の精度比較

手法	クラス	Prec.	Rec.	F1
ルール	事実	0.7474	0.9163	0.8232
	意見	0.5565	0.2532	0.3480
機械学習 (Sarashina)	事実	0.9430	0.9064	0.9243
	意見	0.7936	0.8679	0.8291

に対して、機械学習手法では0.8291まで向上し、本タスクにおける機械学習手法の有効性が示された。

一方で、機械学習手法では、事実文のF1値が0.9243であるのに対し、意見文は0.8291と約0.1ポイント低い値となっている。また、事実文のPrecisionが高い水準に達しているのに対し、意見文のPrecisionは事実文よりも低く、モデルが意見文と判定したものの、実際には事実文であった誤分類が一定数存在することを示唆している。

この結果は、社説において事実文と意見文の境界が必ずしも明確ではなく、特に意見文の判定において、機械学習手法でも識別が困難な事例が存在することを示唆している。

8 エラー分析

実験結果に基づき、各手法の挙動と誤分類の傾向について分析する。

8.1 ルールベース手法の限界

まず、機械学習手法の優位性を明らかにするために、ルールベース手法における誤分類事例を分析する。ルールベース手法において意見文のRecallが低かった主な原因は、辞書マッチングによる表層的な判定の限界にある。例えば、「～と言わざるを得ない」や「～と言うほかない」といった二重否定を含む強調された表現は、書き手の強い主張を示す。また、「～の悪あがきぶりだった」という表現における「悪あがき」は、形式上動作の様態を示す名詞にすぎないが、文脈上は筆者による強い批判の意図を含むと考えられる。これらの事例では、ルールベース手法は文末の特定語彙に依存するため、文脈依存性の高い意見表現を認識できなかった。一方、機械学習手法は、文脈全体からこれらの表現を筆者の評価として捉え、正しく「意見文」とであると分類できていた。このことから、社説の分類においては、単語の有無だけでなく、文脈情報の考慮が不可欠であることが確認された。

8.2 機械学習手法の誤分類事例

機械学習手法の評価結果(表1)において、意見クラスのRecallよりもPrecisionは相対的に低い値となった。これは、モデルが事実文を過剰に意見文と誤分類していることを示している。事実文を意見文と誤分類した123件の事例を分析した結果、以下の3つの言語的特徴が誤分類の要因として考えられる。

第一に、「社会的期待・懸念の記述」である。例えば、「健康に影響が出ることも懸念されている」のように、「期待される」や「懸念される」といった受動態表現を含む文である。これらは筆者個人の主張ではなく、社会的な一般論や客観的な状況報告であるが、モデルは「懸念」など意見性を帯びた語彙に強く反応し、意見文として抽出する傾向が見られた。

第二に、「状況の評価的描写」である。例えば、「人的基盤の強化への目配りは乏しかった」といった文である。これらは客観的な事実の描写として記述されているが、「乏しい」や「深刻だ」といった強い形容表現が含まれるため、モデルがこれを書き手の主観的评价として捉えたと考えられる。

第三に、「事象の背景分析」である。例えば、「譲歩したのは、法案を成立させられないからだ」のように、事象に対する因果関係を説明する文である。「～したためだ」などの論理的な説明構造は、社説における意見展開の形式と類似しており、事実としての背景説明であっても意見文と混同される事例が確認された。

9 おわりに

本稿では、小学校でのNIEにおける読み比べ記事推薦システムの実現に向け、社説記事に対する事実・意見文の分類手法について検討した。社説を用いたデータセットを構築し、ルールベース手法と機械学習手法の分類性能を比較した結果、機械学習手法はルールベース手法に比べてF1値で大幅な向上を示し、特に文脈情報の利用が意見文の抽出に不可欠であることを実証した。

今後は、評価実験で明らかになった誤分類事例への対応に加え、他紙へのドメイン適応性について検証する。最終的には、構築した分類手法を用いて記事間のトピック一致判定や視点对比分析を実施し、教育現場で実用可能な記事推薦システムの実現を目指す。

参考文献

- [1] NIE の実践指定校一覧と校数の推移. <https://nie.jp/selected/>. (引用日 2026-01-06).
- [2] 尾高泉. 民主主義に参加する力を育む－メディアが実践するリテラシー教育－. 情報と科学の技術, Vol. 74, No. 2, pp. 46–51, 2024.
- [3] NIE の学習効果調査結果. <http://nie.jp/research/survey/>, 2024. (引用日 2026-01-06).
- [4] 韓毅弘, Junjie SHAN, 西原陽子. 意見文の抽出と連結による社説記事の書き手の価値観推定. 知能と情報 (日本情報ファジィ学会誌), Vol. 36, No. 4, pp. 757–764, 2024.
- [5] 小森敦也, 砂山渡, 服部峻. Chatgpt を用いた youtube コメントからの主要意見の抽出. In **The 38th Annual Conference of the Japanese Society for Artificial Intelligence**, 2024.
- [6] 新聞を活用した教育実践データベース. <https://nie.jp/report/?gp=1&ky=新聞読み比べ&fw=>. (引用日 2026-01-06).
- [7] 本間工太郎. 社会的事象を多面的・多角的に考え, 「思考力, 判断力, 表現力」を高める方策の工夫. 教育実践研究, Vol. 29, pp. 25–30, 2019.
- [8] 日外アソシエーツ株式会社. 毎日新聞データ集 2024 版 (著作権者: 毎日新聞社). <https://www.nichigai.co.jp/dcs/index5.html>, 2024.
- [9] 分類語彙表 (WLSP). <https://github.com/masayu-a/WLSP>. (引用日 2026-01-06).
- [10] 連想類語辞典: 日本語シソーラス. <https://renso-ruigo.com/>. (引用日 2026-01-06).
- [11] Sarashina-embedding-v1-1b. <https://huggingface.co/sbintuitions/sarashina-embedding-v1-1b>. (引用日 2026-01-06).