

小学校教師に対する Web ニュース推薦のための NIE 教材に適した記事判定法の検討

An Article Estimation Method for Web News Recommendation to Elementary School Teachers

関 伸也*¹, 安藤 一秋*²

Shinya SEKI*¹, Kazuaki ANDO*²

^{*1}香川大学大学院工学研究科

^{*1} Graduate School of Engineering, Kagawa University

^{*2}香川大学創造工学部

^{*2} Faculty of Engineering and Design, Kagawa University

Email: s18g471@stu.kagawa-u.ac.jp

あらまし：全国の小学校では、新聞を教材として活用する取り組み、NIE (Newspaper In Education) が実施されている。しかし、NIE の教材準備によって教師の負担が増加するという問題も生じている。本研究では、小学校教師の教材準備に関する負担を軽減し、NIE を継続的に実践できる環境を提供するため、NIE に活用しやすい Web ニュースを地域学習の教材集として推薦するシステムの構築を目的とする。本稿では、NIE 教材に適した Web ニュースを SVM (Support Vector Machine) を用いて判定する手法を検討する。

キーワード：NIE, Web ニュース, SVM, 小学校教師, 教材推薦, 記事推薦

1. はじめに

NIE (Newspaper in Education) は、新聞を教材として活用する取り組みのことであり、小学校や中学校を中心に、幅広い教育機関で実施されている⁽¹⁾。各地域の NIE 推進協議会は、NIE 実践報告書⁽²⁾を毎年発行している。小学校での実践報告によると、地元や周辺地域に関する記事を用い、地域学習として NIE を実践している例が多い。また、NIE を継続的に行うことで、児童の読解力・語彙力や社会への興味関心が向上したと述べられている。さらに、物事を多角的にとらえる力などの育成に効果があることも確認されている。

小学校での NIE では、各新聞社が発行している紙媒体の新聞や Web ニュースを使用している。しかし、これらの記事には、児童が学習していない漢字や、理解することが困難な表現などが使われており、児童にとって難解な文章であることが多い。そのため、日々発行される膨大な記事の中から、児童自身が自分の興味や課題に適した記事を探し出すことが困難という課題がある。また、教師にとっては、NIE に適した記事を探すことや、記事に関する教材研究によって、教材準備の負担が増加するといった課題がある。したがって、記事の推薦や難解な言葉の言い換えなど、児童に対する支援と、NIE に活用しやすい記事や関連・補足資料の検索・選択など、教師に対する支援が必要になる。そのため、新聞記事に関する読解支援⁽³⁾や児童に対する記事推薦⁽⁴⁾などの研究が行われている。

本研究では、NIE を実践する教師向けの支援に注目し、NIE に活用しやすい Web ニュース記事を地域学習の教材集として推薦するシステムの構築を目的

とする。NIE 教材としての価値が高い記事を推薦することで、教師が記事を探す負担を軽減し、NIE を継続的に実践するための支援が可能と考える。

本稿では、NIE 教材に適した記事を SVM (Support Vector Machine) で判定する手法について検討する。

2. 先行研究

我々の先行研究⁽⁵⁾では、NIE 教材に適した記事を判定する基準を検討するため、特定の新聞社が補助教材として提供している NIE ワークシートに注目し、採用されている記事のカテゴリと読解難易度の傾向を分析した。その結果、地域学習の NIE では、地元に関連する生物や行事、農作物などに関連する記事が採用されやすいこと、ワークシートに設定されている対象学年によって、採用されている記事のリーダビリティスコア⁽⁶⁾の傾向が異なること確認した。また、読解難易度に基づく記事判定手法を初期検討した結果、Bag of Words (BoW) に対し、リーダビリティスコアと習得済み漢字割合を素性に加えることで、判定性能の向上を確認した。表 1 に結果を示す。

表 1. 先行研究での実験結果

素性	precision	recall	f-measure
BoW	0.815	0.786	0.795
BoW+R+K	0.822	0.809	0.810

R: 読解難易度素性, K: 習得漢字率素性

初期検討では、記事内容を表す素性が BoW のみであったため、記事カテゴリの分析結果が活かされていない。そこで、本稿では、SVM の素性に形態素のカテゴリ情報を追加し、その有効性を確認する。形

態素のカテゴリ情報を用いることで、より記事内容を踏まえた判定が行えるようになり、また、地域特有語に対する汎化性能が向上することを期待する。

3. 記事判定 SVM の素性検討

3.1 形態素カテゴリ情報の取得

形態素のカテゴリ情報は、Juman++⁽⁷⁾の出力に含まれる名詞の意味カテゴリ 22 種を利用する。Juman++ の意味カテゴリー一覧を表 2 に示す。

表 2. Juman++ の意味カテゴリー

色	時間	自然物	組織・団体
人	動物	抽象物	植物-部位
数量	植物	形・模様	動物-部位
場所-機能	場所-その他	人工物-金銭	人工物-食べ物
場所-施設	人工物-その他	人工物-乗り物	場所-施設部位
場所-自然	人工物-衣類		

3.2 形態素カテゴリ素性

形態素カテゴリ素性には、以下を組み合わせたものを用いる。

- ・記事中の形態素カテゴリごとの出現有無
- ・記事中の形態素カテゴリごとの出現形態素数
- ・記事中の各形態素カテゴリが含まれる文数
- ・記事中の 2 種類の形態素カテゴリ (231 通り) が含まれる文数

4. SVM を用いた記事判定手法

記事判定は、記事が NIE 教材に適しているか、いないかの 2 値分類であるため、先行研究⁽⁵⁾と同様、2 値分類で高い性能を持つ SVM を用いる。

本稿では、SVM の素性として、先行研究⁽⁵⁾で最適な結果を得た、ベース素性、読解難易度素性、習得漢字率素性に加えて、形態素カテゴリ素性を組み合わせて、記事判定の性能を確認する。なお、予備実験の結果から、形態素カテゴリ素性には、4 つを組み合わせたものを利用する。

実験データには、先行研究⁽⁵⁾と同じ、NIE ワークシートから任意抽出した正例 (NIE 教材に適している記事) 122 記事と、負例 122 記事を用いる。

本稿では、判定結果の正確性を重視し、precision を最優先の評価指標とし、10 分割交差検証により、判定性能を評価する。

4.1 実験結果と考察

素性の組み合わせを変え、10 分割交差検証により、記事の判定実験を行った結果を表 3 に示す。表 3 の precision に着目すると、SVM の素性としてベース素性に読解難易度素性と形態素カテゴリ素性を加えた場合との判定性能が最も高いことがわかる。ベース素性のみと比較し、precision が 2.3 ポイント向上している。また、先行研究で最も判定性能が良かった組み合わせと比較しても 1.6 ポイント向上している。

さらに、形態素カテゴリ素性を追加した組み合わ

せと、追加していない組み合わせを比較すると、どの組み合わせでも判定性能が向上している。このことから、形態素カテゴリ素性が、NIE 教材に適した記事の判定に有効な素性であると考えられる。

表 3. 実験結果

素性	precision	recall	f-measure
BoW	0.815	0.786	0.795
BoW+C	0.830	0.784	0.803
BoW+R	0.817	0.794	0.800
BoW+R+C	0.838	0.784	0.807
BoW+K	0.809	0.801	0.801
BoW+K+C	0.830	0.784	0.803
BoW+R+K	0.821	0.801	0.806
BoW+R+K+C	0.828	0.784	0.803

R: 読解難易度素性, K: 習得漢字率素性,
C: 形態素カテゴリ素性

5. おわりに

本稿では、NIE 教材に適した記事を SVM で判定する手法について検討した。SVM の素性として、先行研究⁽⁵⁾で用いられた、ベース素性、読解難易度素性に、形態素カテゴリ素性を加えることで、先行研究の素性よりも、precision が 1.6 ポイント向上することを確認した。

今後は、判定精度のさらなる向上を目指し、ベース素性を BoW から、word2vec などの分散表現への変更や、記事文章の読みやすさを示す指標として、日本語教育語彙表⁽⁸⁾を用いることを検討している。また、判定した記事を教師に対して推薦するシステムの設計・構築を行う。

参考文献

- (1) 教育に新聞を, <http://nie.jp>, 2017 年 7 月 26 日確認
- (2) NIE 実践報告書, <http://nie.jp/report/panflet>, 2017 年 7 月 26 日確認
- (3) 河村宗一郎, 安藤一秋, “小学生を対象とした Web ニュース読解支援システムのための重要語抽出手法の検討”, JSAI2017 大会論文集, 1J1-5, 2017.
- (4) Shoya Tanaka, Kazuaki Ando, “Web News Recommendation for Elementary School Children using Degree of SNS Users’ Attention and Popular Search Queries among Children”, ACIS International Journal of Computer & Information Science, Vol.17, No.1, pp.17-23, 2016.
- (5) 関伸也, 安藤一秋, “小学校教師に対する Web ニュース推薦に向けた NIE ワークシートの分析と NIE 教材に適した記事判定法の検討”, JSAI2018 大会論文集, 1L1-5, 2018.
- (6) 李在鎬, “日本語教育のための文書難易度に関する研究”, 早稲田日本語教育学, 第 21 号, pp.1-16, 2016.
- (7) Juman++, <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN++>, 2018 年 6 月 8 日確認
- (8) 日本語教育語彙表, <http://jhlee.sakura.ne.jp/JEV.html>, 2018 年 6 月 8 日確認