# ストーリーコンテンツに対する日本語レビュー中のネタバレ判別に関する研究

**A Spoiler Detection Method for Japanese-written Reviews of Story Contents**

加守田 侑
情報処理領域

## 1. はじめに

オンラインショッピングサイトの多くには利用者が商品に関して自由に記述できる商品レビュー（以下，レビュー）が存在する．レビューに記載されている他者の意見は商品購入の際に参考となる．しかし，商品が漫画や小説などストーリーコンテンツの場合はネタバレがレビューに記載されている場合がある．レビューに含まれるネタバレを読むことで，実際に商品を見た際の楽しみが減ってしまうので問題となる．

本研究ではレビューを 1 文単位に分割し，ネタバレを含んだ文を判別することを目的とする．日本語が持つ独特の文法や言い回しに対応するため，単語の分散表現と Long short-term memory (LSTM) [1] を利用して判別を行う手法を提案する．

## 2. 関連研究

ストーリーコンテンツに関するネタバレを判別する研究として，田島ら [2] は Twitter に投稿されたアニメ作品のネタバレツイートを判定する研究を行っている．田島らは手動によるアニメ作品のジャンル分けを行い，各ツイートに含まれる特定の品詞の単語と文節の係り受け関係を特徴として，Support Vector Machine (SVM) で各ジャンルごとに学習および判別を行っている．しかし，この手法ではジャンルごとにネタバレか否かのタグ付けされたデータが必要となり汎用性がないと考えられる．

英語のレビューを対象にネタバレ（あらすじ）を判別する研究として，岩井らの研究 [3] がある．岩井らは Amazon.com 上のストーリーコンテンツに関する英語のレビューを対象に 1 文単位であらすじの判別を行っている．岩井らの提案手法は，あらすじとの相互情報量の大きい単語を主な特徴として Naive Bayes で判別を行っている．岩井らは英語のレビューの場合，あらすじとの相互情報量の大きい単語として代名詞が多く含まれていることを述べている．しかし日本語のレビューの場合，主語の省略が多いため代名詞による判別が難しいと考えられる．また，あらすじとの相互情報量の大きい単語はタグ付けされた訓練データに含まれている単語しか考慮できず，未知語の多い文章の場合は判別の特徴として利用できないと考えられる．

## 3. 提案手法

本研究では，レビューの各文に含まれる単語を fastText [4] を用いて分散表現に変換し，LSTM によってネタバレを含む文の判定を行う．また fastText の訓練データとしてネタバレ判別実験に使用しないレビューを用いることで，ネタバレ判別実験の訓練データに依存しない分散表現を学習する手法を提案する．提案手法の流れを図 1 に示す．

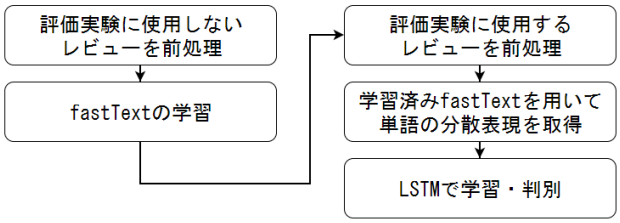


図 1　提案手法の流れ

### 3.1. 前処理

まず，レビューを 1 文単位に分割するために，判別実験で使用するレビューを大まかに確認し，以下の 10 個の記号と改行コードで分割を行う．

。？！ ？！ . . . ♪ * ★ ☆

次に，各文を形態素解析器 MeCab 0.996 を用いて単語単位に分割する．また，MeCab の辞書として新語辞書 mecab-ipadic-NEologd v0.0.5 を用いる．単語単位に分割された文からネタバレを表す品詞と考えられる名詞・動詞・形容詞・形容動詞のみを抽出し，原型に直して判別に利用する．

### 3.2. 単語の分散表現取得

単語の分散表現の取得には fastText を用いる．単語の分散表現とは単語の意味を実数値のベクトルで表現したものである．fastText は同じ文脈に出現する単語は同じ意味を持つという分布仮説に基づいて学習を行う教師なし学習器である．fastText は単語のサブワードの分散表現も同時に学習するため，未知語であっても学習済みのサブワードが含まれていれば分散表現を推測することができる．本研究では単語の分散表現を 100 次元として fastText の学習を行った．

また，fastText の訓練データにはネタバレ判別実験に使用しない商品のレビューを用いた．これはネタバレ判別実験に使用するレビューに依存しない単語の分散表現を学習するためである．

### 3.3. LSTM による学習・判別

文を構成している単語の列を時系列データとしてとらえ，リカレントニューラルネットワークの一種である LSTM をネタバレの判別器として利用する．文に含まれる各単語の分散表現を順に入力し，学習および判別を行う．

使用するのは 3 層のニューラルネットワークである．入力層が単語の分散表現に対応する 100 次元，中間層は LSTM のユニットを持つ 50 次元，出力層は 2 次元の全結合層とする．入力層は 0.2，中間層は 0.5 の割合で Dropout を行う．出力層では softmax 関数を適用し，出力層の各次元をそれぞれネタバレである確率とネタバレでない確率としている．また，各層の出力には Batch Normalization を適用している．

## 4. 評価実験

本研究では日本語のレビューデータセットとして，楽天株式会社が提供している楽天市場の商品に関するレビューデータセットを使用する．

### 4.1. データセット作成

楽天市場の漫画カテゴリに含まれる商品の中で，5 件以上のレビューが含まれる商品をランダムに 100 商品選択する．選択した商品からそれぞれ 5 件のレビューをランダムに抽出する．抽出した計 500 件のレビューを 1 文単位に分割し，各文がネタバレを含む文か否かを研究室の学生 3 名が各々の判断で判定した．2 人以上がネタバレと答えた文をネタバレを含む文とし，結果としてネタバレを含む文が 182 文，ネタバレを含まない文が 1167 文得られた．これをネタバレ判別実験に使用するデータセットとする．また，ネタバレ判別実験に使用しないレビューは fastText の学習に利用する．

### 4.2. 学習・判別

ネタバレ判別実験に使用するデータセットは小規模なので，10 分割交差検定によって学習および判別を行う．ニューラルネットワークは学習回数を 200 回，loss 関数にクロスエントロピー，最適化アルゴリズムに Adadelta を用いる．判別実験に使用するデータセットは 1 : 6 の不均衡データセットであるため，loss 関数計算時にクラスごとにデータ数に応じた重みを掛ける．

また，同じ商品レビューに対してネタバレの判別を行っている岩井らの手法を比較に用いる．岩井らの手法は英語のデータセットでのみ評価しているため，今回使用した日本語のデータセットに適用して結果を比較する．評価指標として適合率・再現率・F 値を使用する．

- 適合率 $= \frac{\text{正しく判別されたネタバレ文の数}}{\text{ネタバレと判別した文の数}}$
- 再現率 $= \frac{\text{正しく判別されたネタバレ文の数}}{\text{ネタバレ文の総数}}$
- F 値 $= \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$

### 4.3. 結果

岩井らの手法と提案手法の実験結果を表 1 に示す．英語のデータセットに対して行われた岩井らの手法の結果と比較すると日本語のデータセットでは岩井らの手法と提案手法の両方で大きく精度が下がっている．原因の一つとして，判別実験用のデータセットのデータ数が少なかったことがあげられる．また，日本語の場合は主語の省略や形態素解析時の誤りなど多岐にわたる要因により英語の場合に比べ精度が下がると考えられる．

表 **1**　ネタバレ判別実験結果

| 手法 | 適合率 | 再現率 | F 値 |
|---|---|---|---|
| [3]（英語・元論文） | 0.80 | 0.77 | 0.78 |
| [3]（日本語） | 0.39 | 0.40 | 0.39 |
| 提案手法（日本語） | **0.47** | **0.67** | **0.55** |

次に日本語のデータセットの結果を比較すると，提案手法は全ての指標で岩井らの手法よりも精度が高くなっている．これは 2 節で述べたように，岩井らの手法では代名詞の省略に対応できないため精度が低くなったと考えられる．一方，提案手法は単語の分散表現を利用したことにより代名詞に頼らず，文に出現する単語の意味を用いて判別できたことが，精度が高くなった一因だと考えられる．また，提案手法の結果では再現率が高くなっているため，より多くのネタバレを判別できていることがわかる．

## 5. まとめ

本研究では日本語のレビューデータに存在するネタバレが含まれる文の判別を，fastText から得られた単語の分散表現と LSTM を用いて行う手法を提案した．結果として，日本語のデータセットに対して関連研究の手法を適用した場合と比べ提案手法は精度が高くなり優位性が見られた．

今後の課題を述べる．アンケート結果を確認すると，筆者の見解ではネタバレではないと思われる文がネタバレと判定されていた．そのため，提案手法で作成された判別器をベースに利用者ごとにネタバレのタグ付けを行い学習させることで利用者に特化した判別器が作成され精度も向上すると考えられる．また，レビューに含まれるネタバレ省く際，文脈が崩壊し，利用者がストレスを感じることになる．そこで，文脈の崩壊度を定義し崩壊度が閾値以上の場合は，文脈がつながるように接続詞などで文章を修正する処理が必要となる．

## 謝辞

本研究ではデータセットとして，楽天株式会社が国立情報学研究所の協力により研究目的で提供している「楽天公開データ」を利用させていただきました．深く感謝いたします．

## 参考文献

[1] F. A. Gers, J. A. Schmidhuber, and F. A. Cummins, “Learning to Forget: Continual Prediction with LSTM”, Neural Computation, vol. 12, no. 10, pp. 2451–2471, Oct. 2000.

[2] 田島一樹, 中村聡史, “Twitter におけるアニメのネタバレツイート判定手法の提案”, 第 8 回データ光学と情報マネジメントに関するフォーラム (DEIM), 2016.

[3] 岩井秀成, 土方嘉徳, 西田正吾, “レビューの文脈一貫性を用いたあらすじ文判定手法”, 情報処理学会論文誌・データベース, vol. 7, no. 2, pp. 11–23, Jun. 2014.

[4] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, “Enriching Word Vectors with Subword Information”, Transactions of the Association of Computational Linguistics, vol. 5, pp. 135–146, 2017.