

放送中に投稿された Tweet を用いた TV 番組の推薦

野津 貴之† 若原 徹†

法政大学 情報科学部†

1. まえがき

TV 番組の推薦では現在、電子番組表の情報を用いることで TV 番組間のキャストやスタッフ、制作会社、番組紹介文などの類似度を評価して推薦するもの[1]や、履歴情報を用いることで TV 番組の選択支援を行うもの[2]、などがある。しかし、現在放送中の TV 番組の類似度を評価し推薦するものは少ない。また、近年 Twitter や Facebook などの SNS サイトでの人気・ロコミから流行した邦画、アニメ映画、TV 番組が続出しているため、SNS サイトへの投稿に関心が寄せられている。

本研究ではその SNS の一つである Twitter に着目し、TV 番組の放送中にその TV 番組に対して投稿された tweets を用いて TV 番組の類似性を分析し、ユーザが選択した TV 番組から別の TV 番組を複数推薦する手法を提案する。提案手法では、まず、放送中の数十種類の TV 番組を対象に、Twitter のストリーミング API を使用することで、各 TV 番組に対して投稿された tweets を全て収集する。次に、収集した tweets を用いて、Tweet を投稿したユーザの TV 番組間での視聴の共起確率、各ユーザの tweet 数を用いた関心度の類似度、tweets に含まれる感情語から算出した感情分布の類似度を算出する。最後に、それらを用いた TV 番組の推薦手法を複数提案する。アニメにカテゴリ化された TV 番組を対象に、被験者を用いた評価実験を行い、提案手法の有効性を示した。

2. データの収集

Twitter で用いられている各 TV 番組固有のハッシュタグを含む tweets を、Twitter のストリーミング API を使用し収集する。次に、収集した tweets の中で、推薦に用いるデータとして不適切な tweets を排除する。まず、tweets を収集する段階で Retweet を排除した。また、Twitter にはトレンドワードを自動で Tweet するスパムアカウントが多く存在する。TV 番組固有のハッシュタグは Twitter のトレンドワードとなる場合も多く、スパムアカウントがそのハッシュタグを Tweet する場合がある。また、スパムアカウントは Twitter にスパムアカウントだと判断されないために、短時間で多くの Tweet をすることは少ない。よって、1 つの TV 番組に対する tweet 数が 3 以下のユーザを排除することで、スパムアカウントの tweets を排除した。実験で収集対象とする TV 番組は、Twitter で多くのユーザから多くの tweet データを得ることが可能なものとして、電子番組表でアニメにカテゴリ化されている TV 番組とした。2017 年 11 月 30 日から 12 月 6 日の 1 週間に放送された 71 番組を収集対象とし、有用な tweet 数は、7,388 ユーザによる 190,960 tweets であった。

TV Program Recommendation Based on Real-time Tweets

†Takayuki Nozu, †Toru Wakahara

†Faculty of Computer and Information Sciences,
Hosei University

3. 提案手法

収集した tweets を用いて、各々の TV 番組の類似性を分析する。推薦の指標として Tweet を投稿したユーザの TV 番組間での視聴の共起確率、各ユーザの tweet 数を用いた関心度の類似度、tweets に含まれる感情語から算出した感情分布の類似度を算出する。

3.1. 視聴の共起確率

ある TV 番組を見たユーザが別の TV 番組を見る確率を算出し、確率の高い TV 番組ほど類似性も高いとして推薦を行う。TV 番組 X を見たユーザの集合を X、TV 番組 Y を見たユーザの集合を Y としたとき、X を見たユーザが Y を見る条件付き確率 $P(Y|X)$ は式(1)となる。ただし、 $\#(S)$ は集合 S の要素数を表す。本研究では、この確率を視聴の共起確率と呼ぶことにする。

$$P(Y|X) = \frac{\#(X \cap Y)}{\#(X)} \quad (1)$$

3.2. 関心度の類似度

ユーザの TV 番組に対する tweet 数をその TV 番組への関心度とみなし、TV 番組間の関心度の類似度を求め、類似度の高い TV 番組を推薦する。TV 番組 X と TV 番組 Y の類似度 s を求めるとき、X と Y のどちらも見ている n 人のユーザの、X に対する tweet 数を (x_1, x_2, \dots, x_n) 、Y に対する tweet 数を (y_1, y_2, \dots, y_n) として、ピアソン相関を用いて式(2)から類似度を得る。 \bar{x}, \bar{y} は標本平均を表す。

$$s = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

3.3. 感情分布の類似度

Twitter に投稿された tweets を対象として感情の分析を行う研究は多く行われている。本研究では、各 TV 番組について投稿された tweets に、10 種類の感情軸（喜、怒、哀、怖、恥、好、厭、昂、安、驚）に分類される感情語とのマッチングを行い、加算していくことで各 TV 番組に 10 次元の感情分布ベクトルを作成する。TV 番組間の感情分布ベクトルのコサイン類似度を求めることで、類似度の値が大きい TV 番組を推薦する。

まず、10 種類の感情軸（喜、怒、哀、怖、恥、好、厭、昂、安、驚）に基づく感情表現辞典[3]から、それぞれの感情軸について 100~200 語を選定し、合計 1,513 語から成る感情語データベースを作成した。このデータベースを用いて、感情語が Tweet とマッチする毎に感情分布ベクトルの対応要素を加算していき、Tweet を投稿したユーザ毎に当該 TV 番組に対する感情分布ベクトルを作成する。ユーザの全ての tweets の感情分布ベクトルの処理が完了したら正規化を行い、全ての正規化されたユーザの感情分布ベクトルを加算したベクトルを正規化したものを、当該 TV 番組の感情分布ベクトルとする。

しかし、使用した1,513語の感情語データベースでは、2.で示した収集した有用な190,960 tweetsの内、約4.6%の8,691 tweetsしかマッチしなかった。本研究と同じくTwitterに投稿されたtweetsから感情語を抽出した研究[4]では、感情語とマッチしたTweetの割合は10~15%ほどであった。そこで、感情語データベースの拡張を行い、感情語データベースとマッチするTweetの増加を試みた。

まず、日本語WordNet[5][6]を用いて感情語データベースにある感情語の同義語を、感情語データベースに追加した。これにより、感情語データベースが8,444語に拡張され、約29%の55,905 tweetsがマッチした。次に、Mecab[7]を用いて感情語データベースにある感情語の形態素解析を行った上、ひらがな表記、カタカナ表記を追加した。これにより、感情語データベースが20,431語に拡張され、約52%の100,181 tweetsがマッチした。最後に、Tweetによく含まれているが感情語データベースにはない感情語を16語手動で追加した。これにより、感情語データベースが20,447語に拡張され、約64%の121,323 tweetsがマッチした。この最終的な20,447語の感情語データベースを用いてTV番組の感情分布ベクトルを算出した。

感情分布ベクトルの類似度は、コサイン類似度を用いることで算出し、類似度の高いTV番組を推薦する。

3.4. 関心度と感情分布の組み合わせ

3.2.で示した関心度の類似度と、3.3.で示した感情分布の類似度で得られた推薦ランクを足し合わせたランクで順位付けをすることで、関心度の類似度と感情分布の類似度の2つを考慮した推薦を行う。

4. 実験

実験では、2.で示した実験用tweetsを使用し、3.で示した4種類の推薦指標を用いてTV番組推薦システムを作成した。推薦システムは、被験者が好きなTV番組を選択することで、4種類の推薦指標によって推薦されるTV番組が3つずつ表示される。

そして、各推薦指標に対して、推薦された3つのTV番組が被験者に合っていたかどうかを、被験者それぞれが1~10点の10段階で評価を行う。また、被験者に「このようなTV番組の推薦システムが実際に存在したら利用したいか」という質問を出して、「はい」か「いいえ」で回答を選択するアンケートも行った。被験者となるユーザは、大学生21名であった。2017年12月8日に被験者全員を集めて、同一環境、同一時間で比較評価実験を行った。

表1に実験結果の各推薦指標の評価を示す。ただし、全ユーザの評価の平均値を各推薦指標の評価とした。

表1 推薦手法の全体評価

推薦指標	視聴の共起確率	関心度の類似度	感情分布の類似度	関心度と感情分布組み合わせ
評価	4.21	5.72	4.92	5.03

実験の結果、関心度の類似度を用いた推薦手法の評価が最も高い5.72となり、視聴の共起確率が最も低い4.21という評価になった。

最後に、アンケートには、「はい」と回答した被験者が20名(約95%)、「いいえ」と回答した被験者が1名(約5%)だった。

5. 考察

表1より、TV番組間でのユーザのtweet数の相関を評価する、関心の類似度を用いた推薦手法が、TV番組の放送中に投稿されたtweetsを用いたTV番組の推薦に最も適した推薦手法だと考えられる。一方、視聴の共起確率を用いた推薦手法の評価が最も低い値になった。これについては、視聴の共起確率で推薦されるTV番組に、番組表の前後の番組や、同じ日に放送されているTV番組を推薦する可能性が高くなっているため、放送時間の類似性が大きく反映してしまっていることが要因と考えられる。また、感情分布の類似度を用いた推薦手法の評価は、関心度の類似度を用いた推薦手法より低くなった。これについては、感情分布のコサイン類似度を計算した場合、いずれも高い類似度を得て、大きな差が出なかったことがマイナスに作用したと考えられる。さらに、3.3.で示したように、tweetsによく含まれている感情語を追加することでマッチした数が大きく増加していることから、まだtweetsの感情語抽出が十分には行われていないと予想される。最後に、アンケート結果から、少なくとも今回被験者に選んだ10代後半から20代前半の学生に対してはTV番組の推薦システムの需要があると言える。

6. むすび

本研究では、TV番組の放送中に投稿されたtweetsを用いてTV番組の類似性を分析し、ユーザが選択したTV番組から別のTV番組を複数推薦する手法について、3種類の指標：視聴の共起確率、関心の類似度、感情分布の類似度を提案した。21名の学生を被験者として比較評価実験を行った結果、TV番組間でのユーザのtweet数の相関に着目した、関心の類似度によるTV番組の推薦手法が最も高い評価を得た。今後の課題として、tweetsによく含まれている感情語を感情語データベースにさらに追加することで、感情分布の類似度を用いた推薦手法の有効性を向上できるかどうかの検討がある。

文 献

- [1] 山田一郎, 宮崎勝, 住吉英樹, 古宮弘智, 田中英輝, “ランダムウォークを利用した番組類似性評価”, 情報処理学会研究報告, vol. 2012-NL-207, no. 12, pp. 1-7, July 2012.
- [2] 隆朋也, 渡辺尚, 樽口秀昭, “履歴情報を用いたTV番組選択支援エージェント”, 情報処理学会論文誌, vol. 42, no. 12, pp. 3130-3143, Dec. 2001.
- [3] 中村明, 感情表現辞典, 東京堂出版, 1993.
- [4] 若井祐樹, 熊本忠彦, 灘本明代, “映画に対する実況ツイートの感情抽出手法の提案”, 情報処理学会研究報告, vol. 2013-DBS-158, no. 16, pp. 1-6, 2013.
- [5] 日本語WordNet, <http://compling.hss.ntu.edu.sg/wnja/> (2017-12-06 アクセス)
- [6] F. Bond, T. Baldwin, R. Fothergill, K. Uchimoto, “Japanese SemCor: A Sense-tagged Corpus of Japanese,” Proc. of the 6th International Conference of the Global WordNet Association (GWC-2012), Matsue, 2012.
- [7] MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://taku910.github.io/mecab/> (2017-12-06 アクセス).