

セレンディピティを考慮したコミック推薦システムの検討

白石 誠† 下川原 英理† 山口 亨†

東京都立大学システムデザイン学部情報科学科†

1. はじめに

現在、年間に多くのコミックが発売されており、コミック市場は年々拡大している。そのため大量のコミックを適した形でユーザーに推薦することが求められている。現在の推薦システムとして、大きく協調フィルタリングと内容ベースフィルタリングの2つがあげられ[1]、推薦精度を重視する研究が多く見られた。しかし、近年、推薦精度が高い商品が必ずしもユーザーを満足させるものではないとの見方があり、推薦制度以外の指標としてセレンディピティが挙げられる[2]。セレンディピティとは偶然、予想外のものを発見することを指し、情報推薦においてユーザーの興味の幅を広げるための重要な指標である。本研究ではこのセレンディピティを考慮し、推薦を実現するものとする。

コミックの推薦を実現するにあたり、コミックの表紙画像のみで分類するものとする。カテゴリや内容などで分類する方法も挙げられるが、内容ベースフィルタリングだと推薦する商品に偏りが出てしまい、セレンディピティな推薦が期待できない。また協調フィルタリングを活用した推薦もあるが、新製品や新規ユーザーの情報が十分にない対象に対して、十分な推薦が行えないコールドスタート問題や評価されていない商品の推薦が行えないなどの問題が挙げられる。

コミックの表紙のみで購入をする「表紙買い」などの経験した人は多い。そのためコミックの表紙の好みに応じて画像を分類することで、本来のユーザーの興味を推測しつつ、思いがけない商品を推薦できると仮説を立てた。本研究ではコミックの表紙の好みに着目し、その上で、セレンディピティが高い推薦が行えるか検討する。

2. 提案手法

コミック画像を分類し、推薦する方法として事前にトレーニングされた学習モデルである VGG-16 を使って画像の特徴量を抽出し、クラスタリングを行った。VGG-16 とは「ImageNet」と呼ばれる大規模画像データセットで学習された CNN モデルであり、入力画像を 1000 のクラスに分類することが可能である。16 層の CNN であり[3]、朴らが 2019 年に行ったコミックの内容推定の研究ではコミックの表紙画像に対して、VGG-16 を活用すること内容推測において高い精度が得られた[4]。構図がシンプルであり、汎用性が高いモデルであるため本研究でも採用した。

今回はこの VGG-16 の最終レイヤーである、予測レイヤーを削除することで、特徴ベクトルを取得し、特徴ベクトルの類似性に基づいて画像をクラスタリングした。クラスタリングの手法としては、最も一般的である、K-means 法を活用した。クラスタリングで分けたグループごとに被験者に提示し、選択されたクラスタグループ内で再クラスタリングを行うことでより類似したクラスタグループに分けるものとした。

3. データベース

収集した画像は楽天ブックス API を活用して出版社 25 社から収集した。収集したコミック画像合計 27259 枚を活用して実験を行った。

4. 実験方法

本実験では提案手法の比較手法として Amazon を利用した。Amazon を選定した理由は、代表的な推薦システム、特に協調フィルタリングを利用したものであるためである。また奥ら[5]の研究でも一般的な推薦手法として Amazon を利用しており、本研究でも採用した。実験の大まかな流れは以下に示す。

A Proposal for a Cover-based Comic Recommendation System
Considering Serendipity Induction

†Makoto Shiraiishi, Graduate School of Computer Science,
Tokyo Metropolitan University

‡{Eri Shimokawara, Toru Yamaguchi}, Department of
Computer Science, Tokyo Metropolitan University

表1 アンケートの内容

質問番号	質問	回答項目
Q1	この本は知っていましたか?	知っていた/聞いたことはある/知らなかった
Q2	システムから提示される前から、私はこの本に興味があった。	1:思う-5:全く思わない(5段階)
Q3	システムから提示されて初めて私はこの本に興味があることに気づいた。	1:思う-5:全く思わない(5段階)
Q4	私はこの本がなぜシステムから提示されたかを理解できた。	1:思う-5:全く思わない(5段階)
Q5	私自身、自力ではこの本は見つけられなかったと思う。	1:思う-5:全く思わない(5段階)
Q6	この絵柄が好きだ。	1:思う-5:全く思わない(5段階)

- (1) 画像を複数提示された中から好きな画像を 1 つ選んでもらう。
 - (2) (1) で選んだ画像によって新たに提示された画像の中から 1 つ選んでもらう。(Amazon では「この商品を買った人はこんな商品も買っています」の項目の中で選択してもらい、ものとする)
 - (3) (2) の試行を合計 2 回行う。
 - (4) 最後に提示された画像の中から好きな画像を 1 つ選んでもらい、アンケートに回答する。
- アンケートの内容は奥らが行ったセレンディピティ指向情報推薦実験[5]を参考に新たに絵柄の好みの質問を追加した(表 1)。

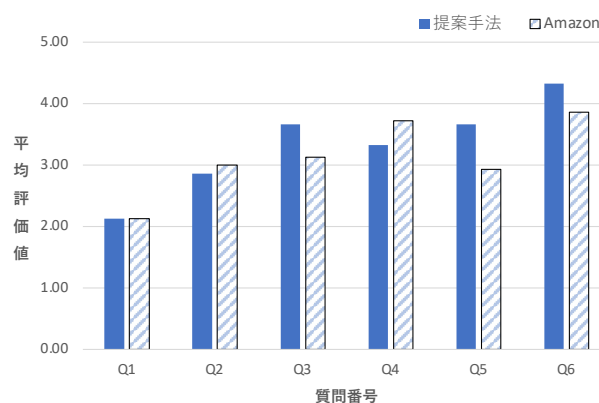


図1 Q1 から Q6 の実験結果の評価平均値

5. 実験結果

本実験では、被験者に提案手法を利用してもらうことで、意外性のあるコミックの発見ができ、セレンディピティの高い推薦ができるか検証した。被験者は 21 歳-24 歳の男女 15 名である。被験者には提案手法と Amazon の両方利用してもらい、それぞれアンケートに回答してもらった。図 1 ではそれぞれの質問に対して、点数を割り振りその評価平均値を求めた。図 1 の結果より、まず Q1 について述べる。提案システムと Amazon で「知らなかった」7 人、「聞いたことはある」3 人、「知っていた」5 人とそれぞれ同数で差は見られなかった。これは提案システムは Amazon と同等の新規性がある商品提案ができたことがわかる。しかし、新規性だけではセレンディピティが高い推薦とは言えない。次に Q3 について述べる。それぞれの評価平均値は提案システムが 3.67、Amazon が 3.13 であった。提案システムが Amazon より、意外性のある推薦ができた。Amazon と同等の新規性に加え、意外性のある推薦が行えたことから提案システムがセレンディピティの高い推薦が行えることが示唆された。

6. まとめ

本研究では、コミックの表紙に着目して、セレンディピティが高い推薦が行えるか検証した。実験により提案手法が既存の手法である Amazon よりセレンディピティが高い推薦が行えることが示唆された。

参考文献

- [1] 神寫敏弘. 推薦システムのアルゴリズム (2) 人工知能, 2008, 23. 1: 89-103.
- [2] 奥健太. セレンディピティ指向情報推薦の研究動向. 知能と情報, 2013, 25. 1: 2-10.
- [3] Very Deep Convolutional Networks for LargeScaleImageRecognition, KarenSimonyan, Andrew Zisserman ICLR 2015
- [4] 朴炳宣; 松下光範. 深層学習を用いたコミックの内容推定. 日本画像学会誌, 2019, 58. 5: 528-537.
- [5] 奥健太; 服部文夫. セレンディピティ指向情報推薦のためのフュージョンベース推薦システム. 知能と情報, 2013, 25. 1: 524-539.