

SNS におけるレコメンドシステムに関する研究 -コールドスタート問題の緩和・解決-

関野 裕通^{††}

† 東京電機大学情報環境学部

和田 雄次[†]

†† 東京電機大学大学院情報環境学研究所

1. はじめに

SNS において、レコメンドシステムと呼ばれる機能が用いられる場合があるが、ユーザの利用履歴データが少ない、または無い場合は、適切なユーザがレコメンドされる可能性は低くなってしまふ。これをコールドスタート問題という。本稿では SNS サイトにおけるユーザ推薦において、コールドスタート問題を緩和・解決する手法の発見を目標とする。

2. IQ 指向推薦方式

栗原の提案した[1]、「文章を読む速度」と「IQ」を属性として用いる推薦方式であり、コールドスタート問題を解決するための手法として考案され、IQ と学習教材の難易度の関係と、文章を読む速度と学習教材の文章量の関係を示し、学習教材コンテンツの推薦において有効な推薦を行える可能性があることを示した。しかし、ユーザ推薦における効果は検証されていないため、IQ 指向推薦方式を用いたユーザ推薦がコールドスタート問題の緩和・解決に用いることができないかを検証する。

3. 検証実験の流れ

各被験者に IQ と文章を読む速度を測定してもらい、測定したIQと文章を読む速度を用いて被験者毎に類似度測定を行う(第一測定)。その後、各被験者に問題を解いてもらい、その結果を利用履歴として扱う。利用履歴データのみを用いて、被験者毎に第一測定で最も類似していた被験者との類似度を測定する(第二測定)。

第一測定と第二測定において、類似度を測定する際に用いたものはユークリッド距離[2][3]であり式 1 で求めた。この式で 算出される値は 0 から 1 の範囲となり、1 に近いほど類似しており、0 に近いほど似ていない。

$$d(p, q) = \frac{1}{1 + \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2}}$$

…式 1

解答してもらった問題は数学の問題で、ベクトル・集合・確率・図形の四分野である。第二測定で利用履歴データのみを用いるのは、IQ と文章を読む速度を用いたコールドスタート状態の推薦が成功しているかどうかを、利用履歴データのみを用いた推薦の結果から判断するためである。

収集したデータの項目毎に偏差値を求め、求めた偏差値を用いて各測定を行った。偏差値を利用する理由は、デ

ータを比較する際に共通の尺度を用いるべきだと考えたからである。また、今回の実験の被験者は 5 名である。

4. 実験結果

表 1. 第一測定の類似度

被験者No	1	2	3	4	5
1		0.059	0.047	0.09	0.052
2	0.059		0.036	0.049	0.041
3	0.047	0.036		0.031	0.267
4	0.09	0.049	0.031		0.033
5	0.052	0.041	0.267	0.033	

表 2. 第二測定の類似度

被験者No	1	2	3	4	5
1		0.029	0.025	0.05	0.071
2	0.029		0.042	0.035	0.025
3	0.025	0.042		0.033	0.021
4	0.05	0.035	0.033		0.048
5	0.071	0.025	0.021	0.048	

表 3. 実験結果

被験者No	第一測定 最類似ユーザ	第一測定 準類似ユーザ	第二測定 最類似ユーザ	第二測定 準類似ユーザ
1	4	2	5	4
2	1	4	3	4
3	5	1	2	4
4	1	2	1	5
5	3	1	1	4

表 1 は第一測定の類似度、表 2 は第二測定の類似度を示したものである。表 1 と表 2 を基にして、第一測定と第二測定それぞれ最も似ていたユーザと二番目に似ていたユーザ(準類似ユーザ)を示したものが表 4 である。第一測定、第二測定共に最類似ユーザを推薦できていたのは 4 番の被験者のみであるが、3 番の被験者以外は第一測定で最類似、準類似だったユーザのどちらかが第二測定の最類似、準類似ユーザに含まれている。そのため、この場合においてはおおむね推薦が上手くいっていると考えられる。

5. おわりに

被験者の数が多い場合にも同様の結果が出るか確かめる必要がある。

参考文献

[1] 栗原 隆平”Web 教材データベースからの教材推薦サービスに関する研究”, 情報処理学会研究報告, コンピュータと教育研究会報告, 2013-CE-118(2), 1-7, 2013-02-01
 [2] Wikipedia ユークリッド距離 ※2015 年 1 月現在
<http://ja.wikipedia.org/wiki/%E3%83%A6%E3%83%BC%E3%82%AF%E3%83%AA%E3%83%83%E3%83%89%E8%B7%9D%E9%9B%A2>
 [3] Toby Segaran ”集合知プログラミング” pp.10-11, 2008-07