

ユーザ評価データのマイニング結果に基づく

スマートフォンアプリの特徴分析

プロジェクトマネジメントコース 矢吹研究室 0942112 増田 知之

1. 研究背景

世界中でスマートフォン利用者が年々増加している。スマートフォンのプラットフォームには、Androids や iOS, Symbian, Research In Motion, Bada, Microsoft などがあるが、その中でも特に、Android と iOS の利用者が多い。

Android と iOS のスマートフォン上で動作するアプリケーション（以下アプリ）は、Google が運営するアプリ配布サイトである Google play, Apple が運営するアプリ配布サイトである App store で配布されている。この2つのストアでは、2013年11月までにおよそ195万ものアプリが配布されている[1]。

これらのアプリは、それぞれのストアでランキングによって順位付けされている。ランキングで上位のアプリはユーザに好印象を与えるアプリであり、下位のアプリはユーザに悪印象を与えると考えられる。しかし、ランキング上位のアプリが必ずしもユーザに好かれているわけではない。そのようなアプリでも、ユーザからの評価値（星の数）が少ないものはある。

2. 研究目的

本研究の目的はユーザの評価に基づいてアプリを分類することである。アプリに対してユーザが付与する評価点（星の数）の分布は、アプリによってさまざまである。例えば、評価点1と5のみが同数あるアプリと、評価点3のみのアプリは、平均評価点は同じだが、その性質は大きく異なると思われる。アプリの性質の違いが最も際立つような評価点の解析方法を見出し、その解釈によってアプリを分類することを目指す。

3. 研究方法

アプリの無料・有料ランキング及び評価データを毎日18時に取得データをする。取得したデータをもとに、データマイニング手法を適用することで、アプリを分類し、その結果を考察することによってアプリの評価指標を見出す[2]。

具体的には、アプリにつけられた評価値（星）をその値（星の数）ごとに集計し、主成分分析を試みる。

4. 結果

iOS のトップセールアプリの評価値を、星1つの割合から星5つの割合まで、5つの数値で表現し、主成分分析を行った結果を図1に示す。第1主成分は星1つの星5つの割合への重みが大きく、第2主成分はレビューの影響度によって分散されているものであった。第1主成分スコアは評価値の平均とよく相関しているが、第2主成分スコアは評価値の平均と相関していない。

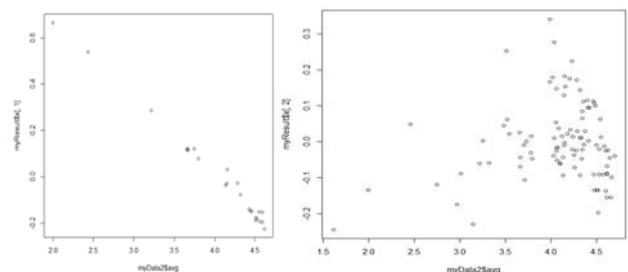


図1 平均評価点と主成分スコアの相関。
横軸は平均評価値。縦軸は主成分スコア
(左が第1・右が第2主成分スコア)

5. 結論

スマートフォンアプリの評価値を主成分分析した結果、アプリの評価は星1つと星5つを重視する見方とそれぞれの星の影響度に分けられることがわかった。評価値の平均とよく相関する第1主成分スコアはアプリ評価の新たな視点とはならないが、評価値の平均と相関しない第2主成分は、アプリ評価の新しい視点になることが期待できる。

ユーザ評価データをマイニングすることによって、個々のユーザからのアプリへの評価や、評価値の単純な平均、アプリのランキング以外にも、アプリを評価する指標があり得ることを示した。このような視点を活用することによって、開発者が、ターゲットとなるユーザからの評価を開発に生かすやすくなることが期待される。

参考文献

- [1] Taisyo. App Store のアプリ数、95万本に到達か?!
<http://taisyo0.com/2013/09/25/19997.html>
2014.1.20.
- [2] 金明哲. Rによるデータサイエンスデータ解析の基礎から最新手法まで一、森北出版. p.320.
2007