

ユーザーレビューを用いたゲームソフトの特徴分析

早稲田大学 *入田凌 永田靖 大野高裕

1. 研究背景と目的

Web 上にはメーカー側から提供される販売促進情報のほかに、実際に商品・サービスを利用した消費者の感想や批評といった評価情報であるユーザーレビューが大量に発信されている。商品・サービスを購入する際にはこれらのユーザーレビューを参考にする人は多い。そして今日、新型コロナウイルスの影響で外出しにくい状況が続いている。自宅で過ごす時間が増えたため、家庭用ゲーム機の需要が急増し自宅でゲームを楽しむ人が増えている。

本研究では、EC (Electric Commerce) サイトから家庭用ゲーム機のゲームソフトを対象にユーザーレビューを収集し、テキストマイニングと多変量解析法を用いて分析する。ゲームソフトのパッケージや販売促進情報からは得られない特徴を得ること、ゲームソフト選びを支援することを目的としている。

2. 先行研究

中谷・星野[1]は、商品の価値はその物としての特徴や評価によってのみ決まるものではなく、その商品を買った人の経験によって得られる価値(経験価値)が重要であると考え、この経験価値を用いるゲームソフトの推薦システムを提案している。

清水[2]は、中谷・星野[1]の経験価値をもとにゲームソフトの評判を抽出し、経験価値に重みづけを行ったうえで多変量解析法を用いてゲームソフトの評判分析をしている。

3. データ

3.1 取得したゲームソフトのレビュー

レビュー数の多い 2020 年販売本数ランキング上位 20 作品(以下人気作品)とレビュー数の少ない 3 作品のレビューを用いる。EC サイト「Amazon.co.jp」[3]から 2020 年 12 月 25 日までに投稿されたレビューをすべて取得した。

3.2 抽出する特徴量

特徴量として中谷・星野[1]の経験価値より「喜び・驚き・困惑・怒り・哀しみ・疲労・不安・不満・生活・他者・視覚・聴覚」を抽出する。それに加え、ゲーム固有の特徴量として「操作性・内容量・戦略性・運・奥深さ・始めやすさ」があると考え、操作性は「操作難易度」、内容量は「内容の多さ」、戦略性は「戦略性の有無」、運は「運要素の有無」、奥深さは「内容の濃さ」、始めやすさは「初心者でも楽しめるか」を表す特徴量

となるように抽出する。

3.3 経験価値の抽出

経験価値の抽出は各ゲームソフトの人気レビュー上位 100 件を用いる。各経験価値を表す単語を清水[2]を参考に定める。経験価値を表す単語の出現件数を調査し、それを標準化したもの得点とする。経験価値とそれを表す単語の一部を表 1 に示す。

表 1 経験価値とそれを表す単語 (一部)

経験価値	経験価値を表す単語
喜び	喜び, 幸せ, 満足, 心地よい
驚き	驚き, 衝撃, ショック, びっくり
困惑	困惑, 悩み, 迷う, 困る, 不安
怒り	怒り, 腹が立つ, イライラ
哀しみ	哀しみ, 涙, 泣く
疲労	疲れる, 疲労, 嫌気
不安	心配, 不気味, おびえる
不満	酷い, くだらない
生活	トレーニング, 運動, ダイエット
他者	友達, 家族, 恋人
視覚	映像, グラフィック, ヴィジュアル
聴覚	音楽, サウンド, BGM, 効果音

3.4 ゲーム性の抽出

操作性・内容量・奥深さ・始めやすさの特徴量を抽出するため、各ゲームソフトのレビューすべてを用いる。そして指定した単語組を用いて共起分析を行う。単語組には正負の値(以下ポイント)を当てる。単語組にポイント当てる際に表 2 の基準で割り当てた。例として操作性で用いる単語組とポイントを表 3 に示す。

表 2 ゲーム性とポイントを当てる基準

ゲーム性	1	-1
操作性	簡単	難しい
内容量	多い	少ない
奥深さ	濃い	薄い
始めやすさ	始めやすい	始めにくい

表 3 操作性で用いる単語組みとポイント (一部)

単語組		ポイント
操作	簡単	1
操作性	悪い	-1

また得点付けをする際に、共起回数が多いほど、単語間の距離が近いほど大きい値を示すスコアを用いる。単語組が*i*単語隣で共起した回数を*t_i*、単語組が*i*単語隣同士で共起したときのスコア係数を*a_i*とするとスコアは式(1)で表す。スコア係数は単語間の距離が近ければ高い値を示すよう*a₁*から*a₅*まで順に 1, 0.5, 0.333, 0.25, 0.2 である。

$$\text{スコア} = \sum_{i=1}^5 t_i \times a_i \quad (1)$$

以下の手順で得点とする。

1. 指定した単語組のスコアを得る。
2. スコアとポイントの積をすべての単語組で算出し合計する。
3. 合計したものをレビュー100件当たりの値に変換し、標準化したものをそのゲーム性の得点とする。

戦略性・運の抽出は、それらのゲーム性を表す単語を用いる。ゲーム性を表す単語のレビュー100件当たり出現件数を調査し、標準化したものを得点とする。ゲーム性を表す単語を表4に示す。

表4 ゲーム性とそれを表す単語

ゲーム性	ゲーム性を表す単語
戦略性	戦略、戦術、戦略的、戦術的、作戦、計画
運	運、運ゲー、確率

3.5 単語の妥当性の検証

特徴量を抽出するために指定した単語が特徴量をうまく表せているかを検証する。一部のゲームソフトのレビューより、指定した単語や単語組が含まれるレビューを実際に読み、特徴量に沿った意味である割合を検出率として表5に表す。

表5 各特徴量の検出率

操作性	内容量	戦略性	運	奥深さ	始めやすさ
70.97	86.25	85.79	83.05	93.33	85.71

操作性の検出率がほかの特徴量より低いですが、検出率は最低でも7割を超えているので指定した単語に問題がないと考えられる。

4. 分析結果

4.1 主成分分析による特徴の把握

多数ある特徴量を要約してゲームソフトの特徴をみるために人気作品の経験価値12変数とゲーム性6変数を併せた18変数を用いて主成分分析(出発行列:相関係数行列)を行った。

表6 主成分分析・固有値と寄与率

主成分番号	主成分1	主成分2	主成分3	主成分4	主成分5	主成分6	主成分7	主成分8	主成分9
固有値	4.636	3.103	2.475	1.783	1.254	1.049	0.845	0.752	0.601
寄与率	0.258	0.172	0.137	0.099	0.070	0.058	0.047	0.042	0.033
累積寄与率	0.258	0.430	0.567	0.666	0.736	0.794	0.841	0.883	0.917

主成分番号	主成分10	主成分11	主成分12	主成分13	主成分14	主成分15	主成分16	主成分17	主成分18
固有値	0.453	0.439	0.238	0.148	0.087	0.063	0.045	0.027	0.002
寄与率	0.025	0.024	0.013	0.008	0.005	0.003	0.003	0.002	0.000
累積寄与率	0.942	0.966	0.979	0.988	0.992	0.996	0.998	1.000	1.000

固有値が1以上であり、累積寄与率も0.794とおよそ8割ある第6主成分までを用いて個体の特徴をみる。

表7 第1主成分から第6主成分の因子負荷量

	主成分1	主成分2	主成分3	主成分4	主成分5	主成分6
聴覚	0.696	-0.267	-0.210	-0.047	-0.327	-0.195
視覚	0.679	0.392	-0.406	-0.003	-0.327	0.232
他者	-0.274	-0.578	0.439	0.051	0.091	0.053
生活	0.166	0.309	0.597	0.388	0.240	-0.447
不満	0.815	-0.197	-0.218	-0.240	0.104	0.209
不安	0.532	0.607	-0.108	0.447	-0.073	-0.044
疲労	0.834	-0.021	0.334	0.045	-0.049	-0.236
哀しみ	0.330	-0.528	-0.390	-0.264	-0.049	-0.377
怒り	0.542	-0.515	-0.043	-0.303	0.374	0.083
困惑	0.483	0.164	0.421	0.117	0.366	0.479
驚き	0.884	0.193	0.029	0.327	-0.038	0.060
喜び	0.504	-0.397	0.526	-0.222	0.163	-0.199
始めやすさ	-0.186	0.481	-0.315	-0.299	0.562	0.040
奥深さ	-0.179	-0.459	-0.381	0.546	0.132	-0.218
運	-0.276	-0.422	-0.331	0.649	0.044	0.148
戦略性	0.252	-0.587	-0.464	0.174	0.206	0.153
内容量	0.077	-0.523	0.508	0.328	-0.189	0.328
操作性	0.249	0.188	-0.333	0.299	0.486	-0.169

表7に第1主成分から第6主成分の因子負荷量を示す。表6よりそれぞれの主成分を読み解いた結果をもとに、各主成分の意味づけを行う。ただし、意味づけにくい主成分については無理に意味づけを行うと個体の特徴を読みづらくなる可能性があるため、そのまま読み取った特徴を主成分の意味として扱う。まず第1主成分の特徴は経験価値の値が全体的に大きいことより、「経験価値の充実度」とする。第2主成分の特徴は、他人とのつながりが薄く、疲れにくい、始めやすいことより、「始めやすく一人で淡々とする度合い」とする。第3主成分の特徴は、視覚的な刺激が薄く他人とのつながりがあり生活環境にも刺激を与えるが疲れやすく内容は量が多いという点である。この主成分に関しては特徴をそのまま意味とする。第4主成分の特徴は、運要素の強い奥深いゲームでドキドキした刺激を感じるゲームを表す主成分であることより、「運を楽しむ度合い」とする。第5主成分の特徴は、視覚・聴覚的な刺激が薄く操作が間で始めやすいことより、「単純さの度合い」とする。第6主成分の特徴は、生活環境への刺激が薄く困惑を感じる事が多く内容量もあるゲームであるという点である。この主成分に関しても特徴をそのまま意味とする。

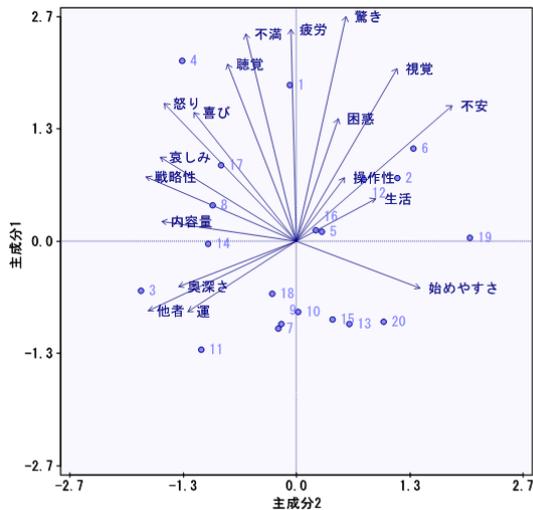


図1 第1主成分と第2主成分のバイプロット図

図1は第1主成分と第2主成分のバイプロット図である。これより、第1主成分において特に高い値を示しているのは「4. ファイナルファンタジーVII リメイク」、「1. あつまれ どうぶつの森」である。2020年度の販売台数の上位ソフトで上位である理由は経験価値の豊富さにあると考えられる。第2主成分において特に高い値を示しているのは、「19. バイオハザード RE:3」、「2. リングフィットアドベンチャー」、「6. マリオカート 8 デラックス」である。よってこれらは、比較的始めやすく一人でやるようなゲームソフトであることがわかる。

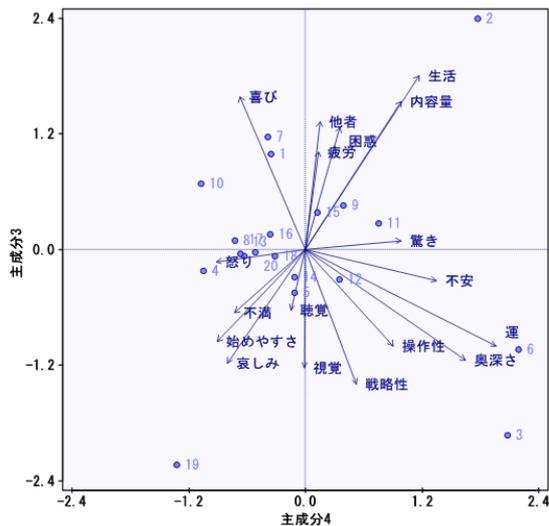


図2 第3主成分と第4主成分のバイプロット図

図2は第3主成分と第4主成分のバイプロット図である。これより、第3主成分において特に高い値を示しているのは「2. リングフィットアドベンチャー」である。自宅で運動を行うソフトの特徴が反映されていると考えられる。第4主成分において特に高い値を示しているのは、「6. マリオカート 8 デラックス」、「3. 桃太郎

電鉄 ～昭和 平成 令和も定番！～」である。よってこれらは、運要素のあるドキドキを楽しむゲームだとわかる。

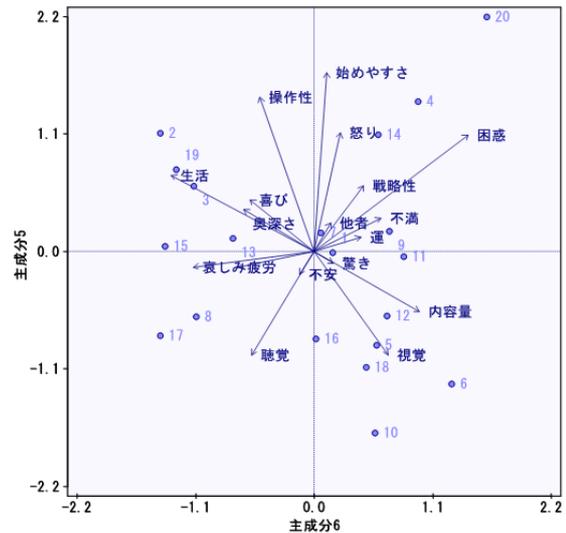


図3 第5主成分と第6主成分のバイプロット図

図3は第5主成分と第6主成分のバイプロット図である。これより、第5主成分において特に高い値を示しているのは「20. eBASEBALL パワフルプロ野球 2020」である。コマンドゲームの単調さが反映されていると考えられる。第6主成分において特に高い値を示しているのは、「20. eBASEBALL パワフルプロ野球 2020」、「6. マリオカート 8 デラックス」、「4. ファイナルファンタジーVII リメイク」である。よってこれらは、生活環境への刺激が薄く困惑を感じる事が多く内容量もあるゲームであるとわかる。

4.2 クラスタ分析による特徴の把握

前節の主成分分析では主成分ごとに特徴的なゲームソフトをピックアップしたが、いずれの主成分でもあまり特徴的でないソフトも存在したため、それらをクラスターに分割することでクラスターごとの特徴を読み取ることを目的としている。

5.2節で行った主成分分析で得られた第6主成分までの主成分得点を用いてクラスター分析(ワード法)を行った結果を示す。

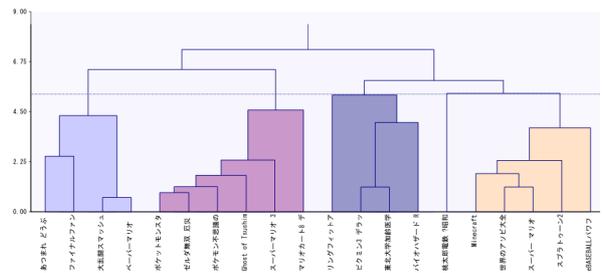


図4 クラスタ分析によるデンドログラム

図4はクラスター分析を行った結果のデンドログラム

である。図4の破線でクラスターを区切ることで、6つのクラスターを得ることができる。ここではクラスターを左から順に第1、第2、第3、第4、第5クラスターと名付ける。

表8 クラスターごとに見る主成分得点

	第1主成分	第2主成分	第3主成分	第4主成分	第5主成分	第6主成分
第1クラスター	1.3258	-0.8200	0.2078	-0.6963	0.0023	-0.3510
第2クラスター	0.1052	0.3998	-0.1662	0.1110	-1.0603	0.6037
第3クラスター	-0.2763	1.0575	0.1223	-0.0165	0.5203	-1.2248
第4クラスター	-0.5760	-1.8210	-1.9110	2.0650	0.6180	-1.1280
第5クラスター	-0.8506	-0.3056	0.3168	0.0238	0.7306	0.7618

表8はクラスターごとに主成分得点の平均値を示す。第4クラスターは桃太郎電鉄のみで前節の分析で特徴を見られたので、ここでは第4クラスター以外を検討する。第1クラスターは、販売本数上位のソフトのクラスターであり第1主成分で特定の高い値を、第4主成分で特定の低い値を示していることから、経験価値の充実した運要素の少ないゲームソフトであることがわかる。次に第2クラスターは、様々なジャンルのシリーズもののソフトのクラスターであり、第5主成分で特定の低い値を示し、第6主成分で高い値を示すことから、単純ではないが内容量は十分であるゲームソフトであることがわかる。続いて第3クラスターでは、個性のあるゲームソフトのクラスターであるが、第2主成分で特定の高い値を示し第6主成分で特定の低い値を示すことから、内容量は少ないが一人で始めやすいゲームソフトであることがわかる。最後に第5クラスターは、対人戦や協力プレイをするゲームソフトで、第1主成分で特定の低い値を示し第3主成分、第5主成分、第6主成分で特定の高い値を示すことから、経験価値は少ないが比較的万人受けするゲームソフトだとわかる。

4.3 レビュー件数の少ないゲームソフトの検証

本研究で行った経験価値とゲーム性の抽出をレビュー件数の少ないゲームソフトで行った場合を検証する。

表9 レビュー件数の少ないゲームソフトの標準化前の得点

ソフト名	操作性	内容量	戦略性	運	奥深さ	始めやすさ	喜び	驚き	困惑
キャプテン翼 RISE OF NEW CHAMPIONS	0.209	0.000	0.000	0.000	0.000	0.000	21.250	2.500	6.250
マリオカートライブホームサーキット	-4.750	-0.625	0.000	2.500	-1.750	-2.500	27.500	2.500	0.000
BAYONETTA&VANQUISH	8.500	0.000	0.000	0.000	0.000	0.000	15.000	0.000	10.000

ソフト名	怒り	哀しみ	疲労	不安	不満	生活	他者	理想	総括
キャプテン翼 RISE OF NEW CHAMPIONS	0.000	0.000	2.500	3.750	5.000	0.000	8.750	3.750	12.50
マリオカートライブホームサーキット	5.000	0.000	10.000	5.000	10.000	0.000	10.000	20.000	5.000
BAYONETTA&VANQUISH	0.000	0.000	0.000	5.000	15.000	0.000	0.000	10.000	0.000

表9より、レビュー件数の少ないゲームソフトでは得点0が多く特徴の抽出が十分行えていないことがわ

かる。次に人気ソフトも含めて標準化を行い人気ソフトと得点を比較する。一部の人気ソフトとレビュー件数の少ないゲームソフトの得点を図5に示す。

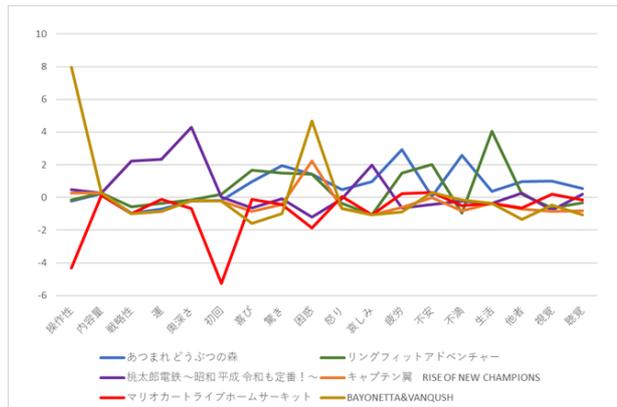


図5 人気ソフトとレビュー件数の少ないゲームソフトの得点

図5より、レビュー件数の少ないゲームソフトの得点がほかの人気ソフトと比べて著しく高い値を示したり低い値を示したりすることがわかる。よって本研究で行ったようにしてユーザーレビューからゲームソフトの特徴を抽出する場合、十分なレビュー件数がないと特徴をうまく抽出できないことがわかる。

5. まとめと今後の課題

ゲームソフトのパッケージや販売促進情報からは得られない特徴を得ることができた。この結果はゲームソフト選択を支援できると考えられる。

今後の課題としては、レビュー件数の少ないゲームソフトの特徴が抽出しづらくほかのゲームソフトとの比較がうまく行えないことが上げられる。レビュー件数の少ないマイナーなゲームソフトにも対応できるよう拡張することも必要であると考えられる。

参考文献

- 1) 中谷知博. 星野准一:経験的価値の分類に基づくゲーム推薦システム, 情報理学会研究報告, No.129, p.49-56(2008).
- 2) 清水航:レビューの語の重みを考慮したテキストマイニングによるゲームの評判分析, 法政大学大学院工学研究科システム工学専攻修士論文(2010).
- 3) Amazon.co.jp, https://www.amazon.co.jp/ref=nav_logo, 最終閲覧日 2020/12/25.