

漫画作品の支持率と登場人物に関わる定量的な特徴の関係 —勾配ブースティング決定木を用いた作品の評価予測—

西崎光太郎

本研究の目的は、漫画読者の作品に対する感覚的評価に影響を与える要因を解明する事である。現在、日本の漫画は国際的にも広く認知されており、雑誌や単行本以外に限らずアニメや映画等、複数のメディアに展開されるなど、文化的かつ経済的にも多様な利益を生む日本の一大文化産業となっている。そのため、今後の漫画業界には継続して質の高い作品を制作する事が求められると考えられる。しかし、漫画に関する既存の学術研究には、漫画の人気要因に焦点を当てた研究が少なく、ある漫画作品の支持率とその作品が持つ漫画的特徴の関係については明らかになっていない。そこで、本研究では漫画家や漫画編集部が作品を制作する上で特に重要視する「キャラクター」という要素に着目し、マンガ大賞における高評価作品と低評価作品の定量的な分析から、キャラクターに関する数的な特徴を用いて機械学習による分類モデルを構築する。そして、構築した分類モデルの予測精度とロジスティック回帰分析の結果から、読者の作品評価に影響を与える特徴の妥当性を検証する。

再帰的特徴量削減と Forward Selection を用いて特徴量を選択しつつ、10分割の交差検証を20回繰り返した結果、キャラクター毎の媒介中心性とセリフに関する特徴量を用いて作品に対する評価を80%以上の精度で予測できる事が分かった。特に、媒介中心性が高い5人の登場人物に着目した場合においては、2番目と3番目に高い媒介中心性の比、主人公及び1番目と5番目に媒介中心性が高い登場人物の有向発話率（文字数）、そして4番目に媒介中心性が高い登場人物の有向セリフ率（吹き出し数）が作品の評価に影響している事が分類モデルの予測精度から明らかになった。このときの分類モデルの正解率は82.65%であり、以上の特徴量を用いて逐次投入法によるロジスティック回帰分析を行った結果、2・3個目の特徴量と5個目の特徴量が作品評価に有意に影響する事が分かった。また、有向発話率（文字数）と媒介中心性、発話率（文字数）と登場頻度という4つの各指標においてその値が最も高い登場人物5人に着目した場合、1番目に有向発話率（文字数）が高い登場人物の発話率（吹き出し数）と4番目に媒介中心性が高い登場人物の有向セリフ率、3番目に高い媒介中心性の比と4番目に登場頻度が高い登場人物の発話率（吹き出し数）が作品評価に影響を与えている事が明らかになった。このときの分類モデルの正解率は84.85%であり、ロジスティック回帰分析によって1~3個目の特徴量が作品評価に対して有意に影響する事が分かった。

今後の課題として、分類モデルの予測精度と信頼性を上げる為に機械学習に用いるデータを増やす必要がある。また、本研究で得た知見を実際の制作現場で活用するには利便性の面で問題があるため、より実用的な方法論を確立する必要がある。

（指導教員 真榮城哲也）