

情報可視化論 最終課題

吉見 圭祐 (227x127x)

1. Introduction

近年、コロナ化であることも相まって、動画配信サービスの需要は非常に高いものとなっている。その中でも、特に Netflix は日本だけでも会員数 500 万人を超える巨大市場として知られている。本レポートでは、そのような巨大な市場の動向を直感的に理解するためのインタラクティブな可視化システムを提案する。

2. Method

データには、Kaggle で公開されているデータセット (Latest Netflix data with 26+ joined attributes, n.d.)を使用する。

各年代における動向を理解するために、二つの図を用いる。一つ目は Wordcloud である。Wordcloud とは、文章中で出現頻度が高い単語を複数選び出し、その頻度に応じた大きさで図示する手法である。この手法を用いて、各年において公開された作品のうち、ジャンルごとの占める割合を可視化する。二つ目は Internet Movie Database(インターネットムービーデータベース:IMDb)で公開されている 10 点満点の IMDb スコアをヒストグラムにしたものである。これらによって、各年における人気のジャンル、全体的な作品の質を直感的に読み取ることが可能となる。

3. Result

作成したシステムを用いた分析の結果として、2015 年～2021 年の間で人気のジャンルはそれほど大きく変化していないということが判明した。コメディとドラマがどの年でも大部分を占めており、特に最新のデータになるほどその傾向が強くなっていた。次いでアクション、スリラーのジャンルが人気だった。IMDb スコアは年ごとに大きな違いはなく、平均的に 7 点程度が最頻値となっていた。

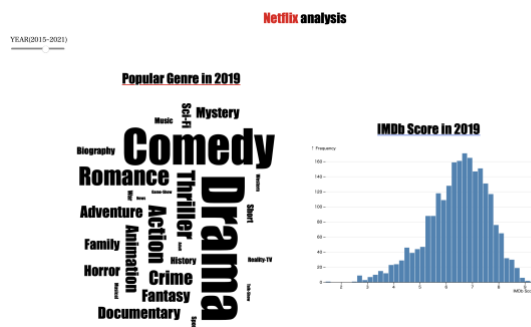


Fig1. Snapshot of visualization system