令和二年度  
プロジェクトデザインⅢ プロジェクトレポート

YouTube動画の再生数を予測するモデルの構築

金沢工業大学 工学部 情報工学科

4EP4-35 鶴瀬 和輝

指導教員 元木 光雄 准教授

令和3年1月21日提出

論文概要

　YouTubeに動画を投稿し,広告収入を得て生活を担うYouTuberという職業が存在します. YouTubeには動画を視聴するだけでなく,視聴した動画に対して評価やコメントをすることが可能であり,広告収入は動画コンテンツの評価値や再生数に応じて変動します.動画の再生数を増やすために,色々な施策やノウハウなどがメディアで紹介されていますが,実際にはどのような指標が再生数を増やすのに適しているのかは分かっていません.そこで本研究ではYouTube APIのメタデータとLightGBMを使用した動画の再生数を予測するモデルを構築し,モデルから再生数を増やすための指標を見つけることを目的とする.

目次

[第1章 序論 1](#_Toc31413590)

[第2章 準備 3](#_Toc31413591)

[2.1 グラフ理論の用語 3](#_Toc31413592)

[2.2 極大クリーク列挙 3](#_Toc31413593)

[2.3 データ研磨 4](#_Toc31413594)

[第3章 提案手法 6](#_Toc31413595)

[3.1 ツイートの取得と都道府県による分類 6](#_Toc31413596)

[3.2 共起ネットワークの生成 6](#_Toc31413597)

[3.3 都道府県ごとの特徴的なトレンド抽出 8](#_Toc31413598)

[第4章 実験による評価 9](#_Toc31413599)

[4.1 二値化とデータ研磨の閾値 9](#_Toc31413600)

[4.2 都道府県ごとの特徴的なトレンド抽出 11](#_Toc31413601)

[第5章 結論 13](#_Toc31413602)

[参考文献 14](#_Toc31413603)

[謝辞 15](#_Toc31413604)

[付録A ソースコード 16](#_Toc31413605)

[A.1 ツイートの取得と都道府県による分類 16](#_Toc31413606)

[A.1.1 search\_data.py 16](#_Toc31413607)

[A.1.2 sorting\_data.py 17](#_Toc31413608)

[A.2 共起ネットワークの生成 19](#_Toc31413609)

[A.2.1 Morphological\_analysis.py 19](#_Toc31413610)

[A.2.2 clique\_enumeration.py 20](#_Toc31413611)

[A.3 都道府県ごとの特徴的なトレンドの抽出 23](#_Toc31413612)

[A.3.1 create\_comparison 23](#_Toc31413613)

[A.3.2 comparison\_maximalClique 23](#_Toc31413614)