

Twitterを用いたiPhoneの販売予測を立てるシステム

System to predict iPhone sales from Twitter

D-5

伊東 麟大 伊與田 光宏

Rinta ITO Mitsuhiko IYODA

千葉工業大学情報科学部

Department of Computer Science, Chiba Institute of Technology

1 はじめに

Twitterは、10代から60代まで、幅広い層から注目されているソーシャルメディア(SNS)である。現在もユーザーは急増しており、日本のアクティブユーザー数は4500万人を超えている。SNSには商品やコンテンツの感想等が投稿されていて、その投稿をマーケティングに活用する企業が増えている。

2 目的

本研究では、Twitterのツイートを用いて感情分析を行い、感情を考慮したiPhoneの販売予測システムの構成を目的とする。

3 システム構成

販売数を予測するまでの流れを下記に示す。

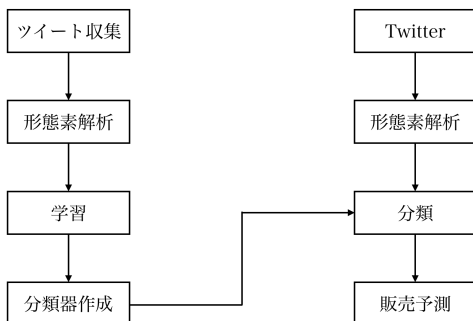


図1. フローチャート

3.1 GetOldTweets3

ツイートを取得する為、GetOldTweets3というものを用いる。現在Twitterの公式APIには時間制約があり、1週間以上の古いツイートを取得することができない。GetOldTweets3では1週間以上の古いツイートを取得することができる。GetOldTweets3を用いて、過去ツイートで指定したキーワードを含むツイートを多く収集する。

3.2 機械学習

本研究ではfastTextという機械学習ライブラリを用いる。fastTextとは、Facebookが開発した自然言語処

理を高速、高精度化するために有効な機械学習ライブラリである。fastTextはGitHubにてオープンソースとして公開している単語のベクトル化とテキスト分類をサポートするためのライブラリであり、fastTextという名の通り動作が軽く速いのが特徴である。

3.3 重回帰分析

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n + b \quad \text{--- ①}$$

yを目的変数、aを説明変数、xを回帰係数、bを切片とする。重回帰分析とは、n個の説明変数より目的変数を表す回帰式を求めることである。

4 推定方法

分類器に学習させる為のツイートを収集する。「iPhone ○○○○」のようにキーワードによる収集を行い学習させる。評価用ツイートを他に収集し、分類させる。ポジティブ、ネガティブツイートの数等のツイッターによるデータとネット上にあるiPhoneの売上高のデータを集め、2019年の日本の売上高を重回帰分析により予測する。

5 評価方法

訓練用ツイートを用いて作成した分類器で分類を行い、重回帰分析による既存のデータを求め結果の比較を行う。

6 結果

重回帰分析による決定係数、予測結果は以下の通りになった。

表1. 評価、予測結果

2019年 第一四半期	6650493528(\$)
決定係数	0.93

2019年 第1四半期の日本の売上は表1のような結果となった。決定係数は0.93となり、目安である0.5以上であったので説明変数の当てはまりが良いのではないかと考えられる。