

1. 目的

本研究の目的は、Twitterを対象とし、形態素解析ツール及び単語感情極性対応表を利用して、ユーザーの感情がポジティブかネガティブかを判定し、人物像の推定に役立つ分析を行うことである。

3年後期の情報セミナーで学んだPHPとMySQLの基礎を活用し、アプリケーション開発を行った。

2. システム概要

図1はシステム構成を表している。認証システムを導入することで、アプリケーションと連携したユーザーのみが機能を使えるようにした。

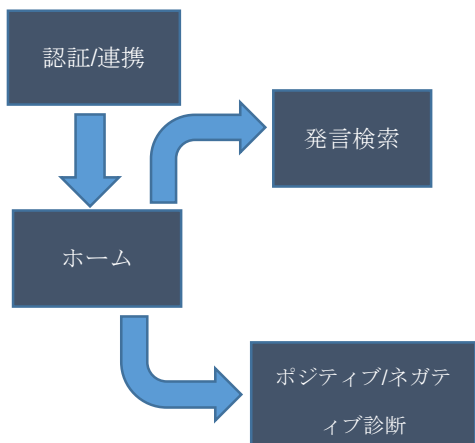


図1 システム構成

発言検索機能はTwitter APIを利用したもので、ユーザーのアイコン、ユーザーID、ツイートの表示などがされる。

3. ユーザーの感情分析

分析の方法として、まず、形態素解析ツールであるIgo-phpを利用して、文を単語ごとに分解する。次に、PN Tableと呼ばれる単語感情極性対応表を利用して、文に含まれる単語と、対応表に含まれる単語が一致した場合に、評価システムに対応表の評価値を返す仕組みになっている。

辞|0.896496| 世|世界|-0.785992|

図2 単語の評価値

文に含まれる単語と、対応表に含まれる単語が一致した場合に、図2に示すように、単語の右に

その評価値を表示している。



図3 トランプ大統領の感情グラフ

図3はトランプ大統領の感情グラフである。横軸がツイートの日付、縦軸がツイートの評価値を示す。高ければポジティブ寄り、低ければネガティブ寄りである。

ツイートの頻度や発言内容、グラフなどから総合的に対象の感情を分析した。

「with **great** negotiating **strength** because of our **tremendous** economy」

「It was a **true** honor」「with our **GREAT** @USForcesJapan」

細かい分析の一例として、上にトランプ大統領のツイートの一部を示した(図3の☆の部分に当たる)。赤字で示した単語は全て大きなポジティブ評価となっている。トランプ大統領は何かを称えるとき、その素晴らしさを強調する単語を頻繁に付けるということがわかった。

4. 課題

発言の文に含まれる単語と、データベースに登録した単語が一致するか毎回総当たりで調べているため、処理が遅く、リアルタイムでの分析ができなくなってしまった。速くするためには、処理の部分のプログラムを改善する必要がある。