

# 旅行スポットを対象としたツイート感情の遷移分析

11471084 根川 哲（灘本研究室）

あらまし：本研究では、Twitter からある旅行スポットを訪れた人のツイートを収集し、各スポットを訪れる前後のツイートからその人の感情の遷移分析を行う。各旅行スポットの特徴を発見することにより、気分にあった旅行スポットの推薦に利用する事が期待できる。

## 1. はじめに

マイクロブログが普及している現在、膨大な情報がインターネット上に投稿されている。その中でも Twitter は 1 つの投稿が 280 文字以内と短く、ユーザが体験したことをリアルタイムに気軽に投稿できる点が特徴である。その為、時系列的にツイートを見ることである時点でのユーザの感情を推定することが可能である。また、人々は旅行に行った際、その時の気分を気軽にツイートしている場合が多い。しかし、パワースポットであったり、癒されるスポットであったりと、旅行先によって旅行の目的が異なる。

そこで本研究では、ある旅行スポットを訪れた前後のツイートからその人の感情を抽出し、その感情の遷移を分析する。具体的には、スポットの特性に着目した分析とツイートしている人に着目した分析の 2 つの分析を行う。

スポットを対象とした分析では、6 箇所のスポットを対象とし、そのスポットを訪れた人のツイートから山本<sup>[1]</sup>らの提案する 8 軸の感情軸”喜・好”、”哀”、”怒・厭”、”安”、”怖”、”恥”、”昂”、”驚”を対象とする遷移分析を行う。人に着目した分析では、CRF を用いた機械学習により、ポジティブ、ネガティブ、ニュートラルの感情遷移分析を行う。2 種類の感情分析を行うことにより、スポットを訪れた人の感情の遷移と具体的にどの感情が遷移したのかを確認できる為、そのスポットの特徴を発見できると考える。

本研究により、旅行スポットの推薦や検索システムに役立てることが期待される。

## 2. ツイートの取得

旅行スポット名をクエリとし、Twitter API を用いて旅行スポットを訪れたユーザのツイートを取得する。取得するツイートは旅行スポットへ訪れた日から過去 800 件のツイートと、訪れた日から 1 週間のツイートとする。分析には 6 箇所の旅行スポットを用いた。

## 3. スポットに着目した感情の遷移分析

スポットの特性により感情の遷移がどのように行われているかの分析を行う。対象のスポットはパワースポットとして鈴虫寺を、テーマパークとして USJ とディズニーランドを、癒しのスポットとして海遊館と宮城蔵王キツネ村、大久

野島を用いる。

### 3.1 感情分類

スポットの特徴毎に微妙な感情も抽出したいと考え、スポットに着目した感情分析には、8 軸の感情軸による分類を行う。

### 3.2 結果と考察

分析対象としたスポット 6 箇所のうち鈴虫寺、USJ、東京ディズニーランド、大久野島の結果を図 1 に示す。横軸が 8 軸の感情の割合、縦軸が旅行へ「行く前」と「行った後」を示す。鈴虫寺では「安」が増え「哀」が減ったことから、鈴虫寺に行って説法を聞き、お願い事をする事によって前向きな気持ちになったと考えられる。USJ ではあまり感情の変化は見られず、同じテーマパークであるディズニーランドと似た結果となった。どちらのテーマパークも年間パスポートが存在しているので何度も訪れる人が多いと考えられる。大久野島では特に「安」の割合が大きく減り、「哀」の割合が大きく増えた。大久野島について調べてみると、かつて毒ガスを製造する工場がありその廃墟が今でも残っていることがわかった。うさぎの島で有名である大久野島を訪れることにより癒しの効果があるように思われたが、実際は大久野島の悲しい過去を目の当たりにすることによって、逆に「安」の感情は減り「哀」の感情が増えてしまうというのがこのスポットの特徴と言える。

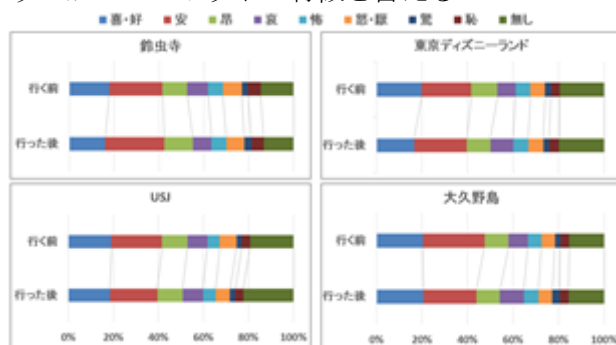


図 1 分析結果

## 4. 人に着目した分析

あるスポットを訪れた人の日ごろのツイートのタイプにより感情遷移に変化があるかを見る分析を行う。ここでは、その人のツイートの時系列変化を見るために、機械学習の CRF を用いる。また対象のスポットは 3 章の分析において、最も

感情の変化のあった大久野島とする。また、感情分類の際にはポジティブ、ネガティブ、ニュートラル3感情を対象とする。

#### 4.1 CRFによる学習

##### 教師データ

CRFの学習データには、アンケートにより感情毎に分類されたツイートを用いる。具体的には、10人の被験者に対し、無作為に抽出した30件のツイートを1セットとしてポジティブ、ネガティブ、ニュートラルの3感情の評価を行った。25セット、計750件のツイートに対し評価を行い、これを教師データとする。

##### 素性

素性には対象ツイート中に含まれる各語の原型、品詞、極性及び極性反転語の数を対象とする。極性反転語は国立国語研究所より公開されている分類語彙表増補改訂版データベース<sup>[2]</sup>内の反対の意味を持つ接続詞39語を対象とする。また、語の極性の導出には高村ら<sup>[3]</sup>が作成した単語感情極性対応表を用いる。ある語の極性  $x$  を単語感情極性対応表の値  $n$  に応じて決定する。具体的には  $n > 0.5$  のとき  $x = 2$ 、 $0.5 > n > 0$  のとき  $x = 1$ 、 $0 > n > -0.5$  のとき  $x = -1$ 、 $n < -0.5$  のとき  $x = -2$  とする。

#### 4.2 平常時のツイートタイプによる人分類

「日常に関するツイートの多い人」と「趣味に関するツイートの多い人」では旅行後の気分の変化に違いがでるのではないかと考え、ツイートを「日常に関するツイートの多い人」と「趣味に関するツイートの多い人」にタイプを分けるためにクラスタリングを行う。

この際、クラスタリング手法には Repeated Bisection 法を用いる。

#### 4.3 分析と考察

##### 4.3.1 仮説

分析をするにあたり以下の2つの仮説を立て、仮説が正しいかどうかを検証する。

**仮説(1)** 日常に関するツイートが多い人は挨拶などニュートラルなツイートが多いため、旅行へ行った後はポジティブなツイートが増える

**仮説(2)** 旅行に行く前より行った日の方がポジティブの割合が増え、その後の1週間もその割合が維持される。

##### 4.3.2 結果と考察

**仮説(1)** 日常に関するツイートが多い人10人中旅行後にポジティブな割合が減った人が7人、増えた人が3人だった。また、ポジティブの割合が増えた人の中でも著しく割合が増えた人はいなかった。ツイート内容を見てみると、行った日には大久野島へ行った感想のツイートが多く見られるものの、その翌日からは大久野島に関する

ツイートはほとんど見られなかった。その為、行った日はポジティブなツイートが多くなるものの、旅行後にそのスポットを振り返るようなツイートをする人は少ないということがわかった。このことより、仮説(1)は「偽」である。

**仮説(2)** 大久野島の結果の例を図2に示す。日常に関するツイートが多い人10人中7人、趣味に関するツイートが多い人10人中9人は旅行へ行った日の全体に占めるポジティブの割合が、旅行前後の全体に占めるポジティブの割合よりも高くなった。しかし、上記の16人全員、旅行後の全体に占めるポジティブの割合は旅行前と同程度まで下がっていた。これは、旅行に行った日はその楽しさからポジティブな感情は増えるものの、旅行後はその虚無感からポジティブな感情は減ってしまったと思われる。その為、旅行に行った日は気分転換になるものの、旅行後までは維持されないことがわかった。このことより、仮説(2)は「偽」である。

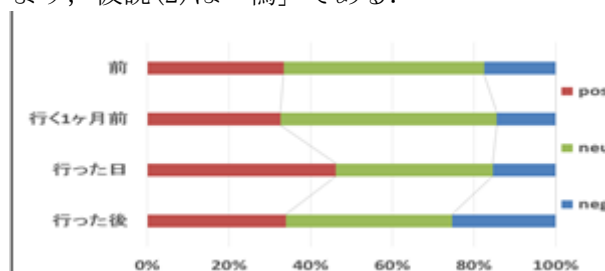


図2 大久野島を訪れる前後の感情遷移の例

#### 5. まとめと今後の課題

本研究では、Twitterから旅行スポットを訪れた人のツイートを収集し、各スポットを訪れる前後のツイートから感情の遷移分析を行った。具体的には、スポットの特徴に着目した分析と人に着目した分析を行った。その結果、旅行スポットの特徴により感情の変化に特徴があることが発見できた。また、人に着目した分析では、旅行に行った日は気分転換になるが、旅行後はその効果は持続しないことがわかった。今後の課題は対象とするスポットの数を増やし、違う切り口からグループ分けを行い、分析を行うことである。

##### 参考資料

[1] Y. Yamamoto, T. Kumamoto, A. Nadamoto, "Followee Recommendation Based on Topic Extraction and Sentiment Analysis from Tweets," iiWAS2015, pp. 215-225, 2015.

[2] 国立国語研究所(2004), "分類語彙表増補改訂版データベース(ver. 1.0)", 2015.

[3] 高村大地, 乾孝司, 奥村学, "スピンモデルによる単語の感情極性抽出", 情報処理学会論文誌ジャーナル, Vol147 No.02 pp. 627-637, 2006.