

# Twitter におけるツイートの感情の予測

## Prediction of human emotion by applying to operations of tweets on Twitter

D-5

勝山 亮† 青木 誠† 木村 敦†

Ryo Katsuyama Makoto Aoki Atsushi Kimura

宮保 憲治† 島田 尊正† 新津善弘† 冬爪 成人†

Noriharu Miyaho Takamasa Shimada Yoshihiro Niitsu Narito Fuyutsume

†東京電機大学 †芝浦工業大学

Tokyo Denki University Shibaura Institute of Technology

### 1. はじめに

近年、スマートフォンが普及している中で、Twitter で情報を発信するユーザーが増えつつある。しかしながら、匿名による誹謗中傷や犯罪に関わるツイートが増えることにより、新たな社会的問題を起こしている。

筆者らは、Twitter 投稿時の情動を記録する Android アプリケーション[1]を独自に開発し、そのアプリケーションを用いて、記事投稿をした時の感情を評価するための情報を追加することにより、Twitter 情報に含まれる感情推定を行う実験を行った。

本稿では、Android アプリケーションによって得られたツイートデータを活用して、Azure Machine Learning (以下 Azure ML)[2]による機械学習で推定を行う方法により得られた結果を述べる。

### 2. データの収集

実験には学生計 30 名が参加し、開発したクライアント用アプリケーションを各自の Android 端末にインストールし、そのアプリケーションを通じて普段通りに Twitter を利用させた。開発したアプリケーションを用いて、毎回のツイート時に感情値を入力し、感情種別毎の値をアプリケーションにより記録した。記録する感情のカテゴリは、Plutchik の提唱した「感情の輪」[3]に基づいて設定した。

実験協力者には 14 日間アプリケーションを使用させ、14 日後までに累積したデータを解析した。結果的に全ての被験者から、合計で 456 件のデータを収集し、CSV ファイル形式で保存した。実験者がツイートした時に表示される記事例を表 1 に示す。表 1 に示す 8 つのデータが、実験参加者が実際に評価した正解データである。

表 1 データの運用例

記事	日時	感情1	感情2	感情3	感情4	文字数	改行の数
今日は暖かいですねー	2015年11月11日 14時50分51秒	1	2	2	3	10	0

### 3. Azure ML による機械学習

実験で収集した全てのデータを、Azure ML による機械学習で解析した。本実験では、被験者のデータを事前に収集しているために、教師あり学習と等価である。教師あり学習の代表的な手法である線形回帰、ニューラルネットワーク、ディビジョンツリーの 3 つの方法で検証した。配置したモジュール(線形回帰)を図 1 に示す。図 1 の Linear Regression モジュールは線形回帰を表す。Train により、正解付きのデータを機械に学習させて、推定モデルを作成した。Score により、Train にて生成された推定データを使って推定値を得た。Evaluate は、Score により比較された結果、学習済みのモデルがどのくらいの精度を達成できたかを数値化したものである。Convert to CSV により、機械学習したデータを CSV ファイルに変換した。

取得した全 456 件のデータのうち、75%を教師データとして活用し、残りの 25%を解析データとして活用した。

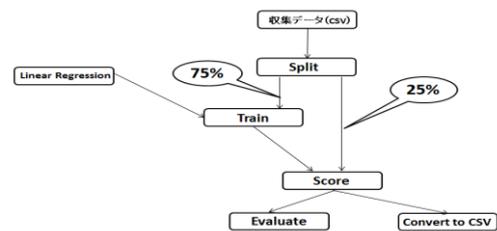


図 1 Azure ML のモジュール図 (線形回帰)

### 4. 推定結果

データ収集後、機械学習に活用した 2 つのツイートの感情値を表 2 に、さらにこれらのツイートに対して機械学習によって予測したデータ(抜粋)を表 3 に示す。表 2, 3 に示した感情値は表 1 の感情 1~4 に相当する。表 3 の Scored Labels の値は線形回帰による機械学習によって推定されたデータである。表 2 の感情値と表 3 の Scored Labels の値を比較すると、感情値 4 のツイート 1 に示すように、値には差が存在するが、多くは推定データと近似することが判明した。

表 2 ツイートの感情値

記事	感情値1(喜び-悲しみ)	感情値2(受容-嫌悪)	感情値3(期待-驚き)	感情値4(恐れ-怒り)
ツイート1	5	5	5	5
ツイート2	2	2	2	3

表 3 機械学習によって予測されたデータ

記事	Scored Labels			
	感情値1(喜び-悲しみ)	感情値2(受容-嫌悪)	感情値3(期待-驚き)	感情値4(恐れ-怒り)
ツイート1	4.76	4.56	4.24	3.6
ツイート2	2.32	2.99	2.49	2.52

### 5. まとめ

本稿で示した機械学習の手法により、おおよその感情推定が可能であることが検証できた。今後は、感情値に差があるデータの特徴を精査し、精度を一層高めるために、新たな教師データの追加や適用範囲の拡大化を進める予定である。

### 参考文献

- [1] 青木他, “Twitter におけるツイートの感情予測システム 2016 年電子情報通信学会発表予定”.
- [2] Microsoft, “Microsoft Azure Machine Learning,” Available: <https://studio.azureml.net..>
- [3] Plutchik, “Plutchik, R The Nature of Emotions,” Available: <http://www.fractal.org/Bewustzijns-Besturings-Model/Nature-of-emotions.htm..>