

絵文字付きツイートを訓練事例として利用する分類器学習による ツイート中の主観同定*

徐 凌寒[†] 林 友超[†] 宇津呂 武仁[†]
筑波大学大学院 システム情報工学研究科[†]

1 はじめに

近年、ツイッター等の SNS が普及してきた。ツイッターの利用者の多くは、自身が配信するツイートの中に、様々な事柄に対する主観情報を含めることがよく知られている。ツイッターでは、文字数制限があるため、制限された文字数の範囲で適切に自分の感情を表すことが難しい。そこで、ツイッターでは、制限された文字数の範囲で感情を適切に表現するために、絵文字が利用される傾向にある。これに対して、自然言語処理の研究分野においては、絵文字を含めたテキストの分析手法が研究対象の一つとして注目されている。

そこで、本論文では、ツイート中の主観を自動分析することを研究の目的とする。そのための手段として、本論文では、主観を表す絵文字に着目し、絵文字を含むツイートを訓練データとして分類器を訓練する。ここで、本論文では、「ハッピー」を表す絵文字および「怒る」を表す絵文字を含むツイートのテキスト部分に含まれる主観の種類・内容を人手で分析したところ、約 20~40%程度の割合で、絵文字が表す主観とテキスト部分が表す主観が矛盾するツイートが混在することが分かった。そこで、本論文では、絵文字が表す主観とテキスト部分が表す主観が矛盾する割合が 40%程度となり絵文字を除外し、主観の矛盾するツイートの割合が約 20~35%程度となる絵文字のみを利用して、各絵文字の表す肯定的・否定的極性をそのままツイート自身の極性とみなして、ツイートの主観同定用の訓練事例として用いる。次に、本論文では、人手で中立の極性として判定したツイートを加えて、ツイートに対する肯定・否定・中立の三値分類器を訓練する。三値分類器によって、均等クラス分布ツイート集合、および、ツイート空間における主観極性分布に従ったツイート集合に対して、主観極性分類性能の評価を行う¹。

*Sentiment Analysis of Tweets by Classifier Learning utilizing Tweets with Emoji as Training Data

[†]Linghan Xu, Youchao Lin, Takehito Utsuro, Graduate School of Systems and Information Engineering, University of Tsukuba

¹本論文において、ツイート空間における主観極性分布に従った

2 絵文字を含むツイートにおける主観分析

本論文では、2016年11月13日から2017年6月23日までの約7ヶ月の期間において、Twitter Streaming API²を用いて収集したツイート³を分析対象とする。収集対象とする絵文字は、



の四種類とする。これらの各絵文字を含むツイートを収集した結果のツイート数を表1「収集総数」欄に示す。

次に、本論文では、収集されたツイートから無作為に抽出したサンプルに対して、

- (a) ツイート中のテキスト部分が表す主観と絵文字が表す主観が **合致**
- (b) ツイート中のテキスト部分が表す主観と絵文字が表す主観が **矛盾**

の二種類への分類を行った。分類結果の統計は表1に示す通りであり、四種類の絵文字の中では、特に、絵文字  を含むツイートにおいて、「テキスト部分と絵文字の主観が矛盾」となるツイートの割合が相対的に高かった。次に、「テキスト部分と絵文字の主観が矛盾」となるツイートのテキスト部分を人手で分析し、四種類に分類した。これらの統計および例を表2に示す。

3 均等クラス分布ツイート集合の三値分類

本節では、肯定・否定・中立の三値極性分類タスクを対象とし、特に、クラス分布を均等にしたツイート集ツイート集合を対象として分類器の評価を行う理由としては、ツイート空間から肯定・否定・中立の各極性のツイートを収集するタスクに対して訓練した分類器を適用した場合に、どの程度の性能が達成できるのかを見積もるためである。

²<https://dev.twitter.com/streaming/overview>

³二種類以上の絵文字を含むツイートのテキスト部分の主観が、ツイート中のどの絵文字の主観に従うのかは自明ではないため、本論文では、ツイートとしては、絵文字一種類のみを含み、その他のいずれの絵文字も含まないものを収集対象とする。

表 2: ツイートの主観と絵文字が矛盾する場合の分類

絵文字	分類	例	数(割合(%))	
ハッピー	😊	主観と絵文字が逆	朝からうざいって。だる 😊	6 (30.0%)
		主観なし	blog更新しています 😊	7 (35.0%)
		主観なしの広告	シリアルナンバー譲ります 😊 4枚あるので当選者は4人です!	3 (15.0%)
		日本語ではない	시키프레 시키프레 😊	4 (20.0%)
	😄	主観と絵文字が逆	仕事 全部 中途半端やりっぱなし にしてる奴の指示なんか 聞きたかねえよってな 😄	12 (29.3%)
		主観なし	わたしも長らく思っていました 😄	24 (58.5%)
		主観なしの広告	😊 メッセージリース をつけて#浅田真央 さんに届けよう!	3 (7.3%)
		日本語ではない	To read German in Katakana is both fun and struggle 😄	2 (4.9%)
怒る	😡	主観と絵文字が逆	可愛いすぎです 😡	25 (71.4%)
		主観なし	ぶんしゅかしよう 😡	10 (28.6%)
	😠	主観と絵文字が逆	あいらびゅ 😠	9 (45.0%)
		主観なし	ですよね 😠	10 (50.0%)
		日本語ではない	para d atormentar o otacu fedido caveira fudida 😠	1 (5.0%)

表 1: 絵文字を含むツイートにおける主観の分類

絵文字	ツイートの主観と絵文字が合致	ツイートの主観と絵文字が矛盾	調査総数	収集総数	
ハッピー	😊	100	20	120	332,573
	😄	50	41	91	37,576
怒る	😡	50	35	85	13,497
	😠	50	20	70	25,161

合を訓練・評価データとした場合について、ツイートの主観に対して肯定、否定、中立の三値極性分類を行う課題について述べる。

3.1 訓練・評価手順

本論文の分類器の訓練においては、まず、訓練データとして、前節における分析結果をふまえて、「テキスト部分と絵文字の主観が矛盾」の割合が最も高い絵文字😊を含むツイートは訓練データとしては用いないこととする。そして、残りの三種類の絵文字を含む全収集済みツイートのうち、前節で人手によって「テキスト部分と絵文字の主観が矛盾」の割合を調査する際に分析対象としたツイートを除外した残りのツイートから、分類器訓練のための訓練データを選定する。具体的には、まず、絵文字😊を含むツイートから3,000ツイートを無作為に選定する。次に、絵文字😄を含むツイートの集合、および、絵文字😡を含むツイートの集合の

和集合から3,000ツイートを無作為に選定する。一方、中立的極性のツイート集合については、2節において収集したツイート集合に対して、無作為に人手による極性判定を行い、中立と判定されたツイート3,000個を収集する。

各訓練事例のクラス値としては、

- 絵文字😊を含むツイートを肯定的極性クラス、
- 絵文字😄を含むツイート、および、絵文字😡を含むツイートを否定的極性クラス、
- 人手で中立に判定したツイートを中立的極性クラス、

とする。そして、これらの合計9,000ツイートを訓練・評価データとしてLIBSVMを適用し、10分割交差検定によって評価を行う。各訓練事例の素性表現を作成する際には、ツイート中のユーザー名(例:”@xxx”等の文字列)、URLs, “RT”の文字列、および、絵文字以外のテキスト部分の形態素を素性とする⁴。

3.2 評価結果

参照用の事例集合を、それぞれ、肯定的極性クラスの事例の集合、否定的極性クラスの事例の集合、中立クラスの事例の集合、および、三者の和集合(「全体」とした場合)についての評価結果を、図1のプロットに示す。この結果から、中立クラスの適合率は他のク

⁴形態素解析ツール MeCab (<http://taku910.github.io/mecab/>、辞書としてIPA辞書を用いる)によって求める。

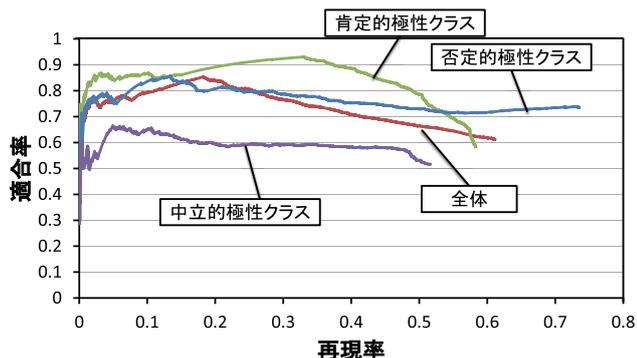


図 1: 均等クラス分布ツイート集合に対する三値分類の評価結果

クラスの適合率より低いことが分かる。また、再現率が 52%以下の場合においては、肯定的極性クラスの適合率は否定的極性クラスの適合率より高いが、再現率が 52%以上の場合においては、両者の関係は逆となる。

4 ツイート空間中のクラス分布に従うツイート集合の三値分類

本節では、前節と同様の肯定・否定・中立の三値極性分類タスクを対象とし、特に、Twitter Streaming API によって、ツイート空間における肯定・否定・中立クラス分布に従ったツイート集合を収集し、肯定、否定、中立の三値極性分類を行う課題について述べる。

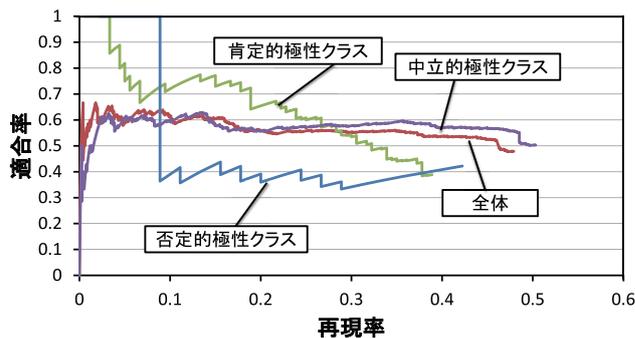
4.1 評価用ツイート集合

評価用ツイート集合としては、2017 年 1 月 19 日から 2017 年 1 月 22 日までの四日間の期間において、Twitter Streaming API を用いて収集した 22 万のツイートから、無作為に 1,000 ツイートを抽出し、人手で肯定、否定、中立の三値極性を判定する。その結果、180 の肯定的極性ツイート、45 の否定的極性ツイート、および、775 の中立的極性ツイートが収集された。この結果から、Twitter Streaming API によって、ツイート空間における肯定・否定・中立クラス分布に従ったツイート集合を収集した結果においては、中立的極性のツイートの数が圧倒的に多く、7 割以上を占めることが分かった。また、否定的極性のツイートの比率が 5%以下ということも分かった。

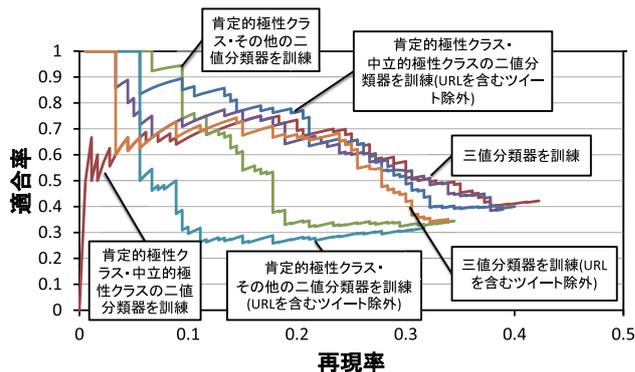
4.2 三値分類器による評価

3 節の評価実験において用いた訓練・評価用ツイートの全体を訓練用ツイート集合として三値分類器を訓練し、前節において作成した評価用ツイート集合に対して適用しその性能を評価する。

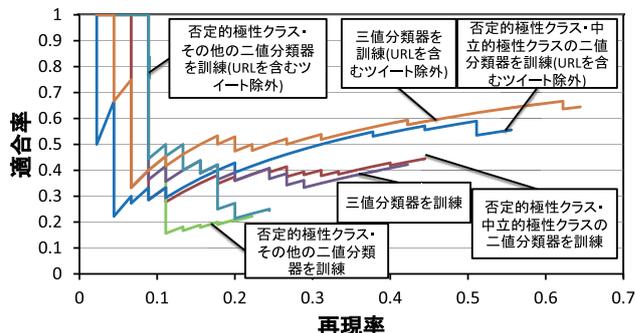
参照用の事例集合を、それぞれ、肯定的極性クラスの事例の集合、否定的極性クラスの事例の集合、中立的極性クラスの事例の集合、および、三者の和集合



(a) 三値分類器による評価結果



(b) 肯定的極性クラスを評価データとした場合の分類器の比較



(c) 否定的極性クラスを評価データとした場合の分類器の比較

図 2: ツイート空間中のクラス分布に従うツイート集合の三値分類の評価結果

(「全体」としての場合) についての評価結果を、図 2(a) のプロットに示す。

この結果から、肯定的極性クラスの性能は否定的極性クラスの性能を上回ることが分かる。また、再現率が 27%以下の場合、肯定的極性クラスの適合率は中立的極性クラスより高い。これは、肯定的極性クラスのツイートの中に、テキスト長が非常に短く、かつ、極性判定が極めて容易なツイート群が含まれており、これらのツイート群が原因となって、肯定的極性クラスのツイート全体の性能が相対的に高くなったからであると考えられる。

4.3 訓練事例集合別の性能比較

本節では、肯定的極性クラスおよび否定的極性クラスを対象として、訓練事例集合別に分類器を訓練した後、各評価対象極性クラスに対して適用した結果の性能を比較する。特に、本節では、URLを含むツイートにおいては広告の比率が高く、それらの広告ツイートにおいてはテキスト部分と絵文字の主観が一致しない場合が多いことをふまえて、URLを含むツイートを除外する場合と除外しない場合の二通りの訓練データを作成し、分類器の訓練・評価を行う。

4.3.1 肯定的極性クラス

以下の各分類器の評価を行った結果を図 2(b) に示す。

1. 肯定的・中立的極性クラスの各 3,000 ツイートを訓練事例として訓練した二値分類器
2. URL を含まないツイート (肯定的極性クラス 1,870+ 中立的極性クラス 2,023) を訓練事例として訓練した二値分類器
3. 肯定的極性クラスの 3,000 ツイート、および、否定的・中立的極性クラスの 3,000+3,000=6,000 ツイートを訓練事例として訓練した二値分類器
4. URL を含まないツイート (肯定的極性クラス 1,870+ 否定的・中立的極性クラス 2,448+2,023=4,471) を訓練事例として訓練した二値分類器
5. 4.2 節で評価した三値分類器
6. URL を含まないツイート (肯定的極性クラス 1,870+ 否定的極性クラス 2,448+ 中立的極性クラス 2,023) を訓練事例として訓練した三値分類器

図 2(b) の結果から分かるように、URL を含むツイートを除外して訓練した肯定的極性クラス・その他の二値分類器の性能は他の分類器の性能を下回っている。また、再現率が 10% 以下の場合においては、肯定的極性クラス・中立的極性クラスの二値分類器の性能が低下するが、それ以外の場合においては、他の分類器と同等の性能となる。全体を通しては、URL を含むツイートを除外して訓練した肯定的極性クラス・中立的極性クラスの二値分類器の性能が他の分類器の性能を上回っている。ただし、URL を含むツイートを除外して訓練することにより、三値分類器、肯定的極性クラス・その他の二値分類器とも性能が下がった。

4.3.2 否定的極性クラス

以下の各分類器の評価を行った結果を図 2(c) に示す。

1. 否定的・中立的極性クラスの各 3,000 ツイートを訓練事例として訓練した二値分類器

2. URL を含まないツイート (否定的極性クラス 2,448+ 中立的極性クラス 2,023) を訓練事例として訓練した二値分類器
3. 否定的極性クラスの 3,000 ツイート、および、肯定的・中立的極性クラスの 3,000+3,000=6,000 ツイートを訓練事例として訓練した二値分類器
4. URL を含まないツイート (否定的極性クラス 2,448+ 肯定的・中立的極性クラス 1,870+2,023=3,893) を訓練事例として訓練した二値分類器
5. 4.2 節で評価した三値分類器
6. URL を含まないツイート (肯定的極性クラス 1,870+ 否定的極性クラス 2,448+ 中立的極性クラス 2,023) を訓練事例として訓練した三値分類器

図 2(c) の結果から分かるように、URL を含むツイートを除外して訓練することにより、いずれの分類器においても性能が向上した。全体を通しては、URL を含むツイートを除外して訓練した三値分類器の性能が他の分類器の性能を上回っている。

5 関連研究

ツイート集合に対して人手で主観同定を行ったデータセットを蓄積し、分類器等の主観判定モデルを訓練する既存の研究事例においては、そのほとんどが、ツイート空間からツイート集合を収集する段階、あるいは、人手で主観判定を行う対象ツイートをサンプリングする段階において、肯定・否定極性のツイートが同等数程度収集できるように、何らかのフィルタリングを施すことが多い(例えば、SemEval の Sentiment Analysis in Twitter タスク (2013~2017) [1, 2])。

6 おわりに

本論文では、まず、絵文字を含むツイートにおいては、絵文字の表す主観とテキスト部分の表す主観が矛盾するものが一定規模混在しているという結論を得た。さらに、ツイート空間における主観極性分布を調査するとともに、絵文字を含むツイートと人手で主観を判定したツイートをを用いた三値分類モデルの訓練によって、ツイート空間における主観極性分布に従ったツイート集合に対して肯定、否定、中立三値極性分類を行い、その有効性を評価した。

参考文献

- [1] P. Nakov, S. Rosenthal, Z. Kozareva, V. Stoyanov, A. Ritter, and T. Wilson. SemEval-2013 task 2: Sentiment analysis in twitter. In *Proc. 7th SemEval*, pp. 312–320, 2013.
- [2] S. Rosenthal, N. Farra, and P. Nakov. SemEval-2017 task 4: Sentiment analysis in twitter. In *Proc. 11th SemEval*, pp. 502–518, 2017.