

ユーザーの期待する反応に応じたツイート分類

Tweet classification based on expected response by user

植田 智明*¹ 折原 良平*¹ 清 雄一*¹ 田原 康之*¹ 大須賀 昭彦*¹
 Tomoaki Ueda Ryohei Orihara Yuichi Sei Yasuyuki Tahara Akihiko Ohsuga

*¹ 電気通信大学大学院情報システム学研究所
 Graduate School of Information Systems, University of Electro-Communications

In recent years, the response from another user against the intentions and expectations of the original twitterer may cause discomfort and stress, which is a social phenomenon known as "SNS fatigue". Such responses typically arise from mutual misunderstanding between users. In order to resolve the problem, it is important to know responses that is expected by the original twitterer. In this paper, we propose a classification method of tweets based on the response that users expect, and experimentally evaluate the method.

1. 背景と目的

1.1 背景

近年, Twitterをはじめとするマイクロブログや Facebook などの SNS において, "SNS 疲れ" [yamakami 12] が問題となっている。"SNS 疲れ" の定義には様々なものがあるが, 本研究では, [加藤 13] の言葉を借り, 「SNS を利用する中で利用経験に基づいた何らかの否定的感情を抱き, サイト利用を控えたり, 退会したりした経験」と定義する。ソーシャルネットワーキングサービスの特性上, 面識のない人とのやり取りや文字数制限による説明の不足, ユーザーやトピックに関する前提知識の有無, 一連のメッセージのうち一部分の拡散などにより, メッセージの発信者の意図と受信者の解釈のずれが生じ, "話の噛み合わない" や "冗談が通じない" といった状態が否定的感情を生むこととなる。また, このような発信者と受信者のずれは, 炎上や友人関係の不和といった社会的不利益につながりかねない。

本研究においても, 事前調査として Twitter ユーザー 145 名を対象にアンケートを実施したところ, 85% が他のユーザーからの投稿や返信によって, 否定的な感情を感じた経験があると回答した。その具体的な内容として, 図 1 に示すように, "他のユーザーへの誹謗中傷" や "TL (タイムライン) 上での言い争い", "リプライなどにおける論点のずれたやり取り" といったものが否定的な感情を生む原因になっているとの知見が得られた。特に "論点のずれたやり取り" は 34% と比較的多く, 誹謗中傷や言い争いの原因になりうると思われるため, こうした "論点のずれの発生" を検知し, 不利益を被る前に防ぐことが必要である。

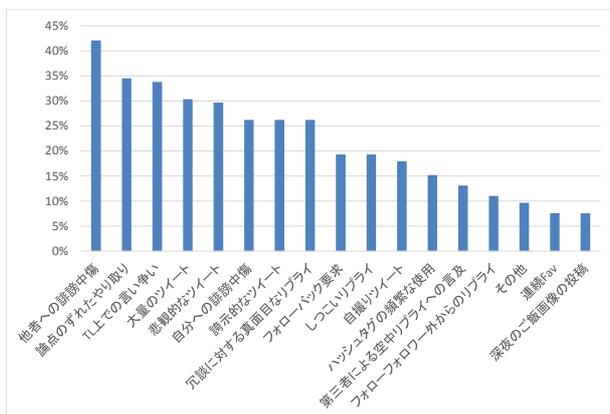
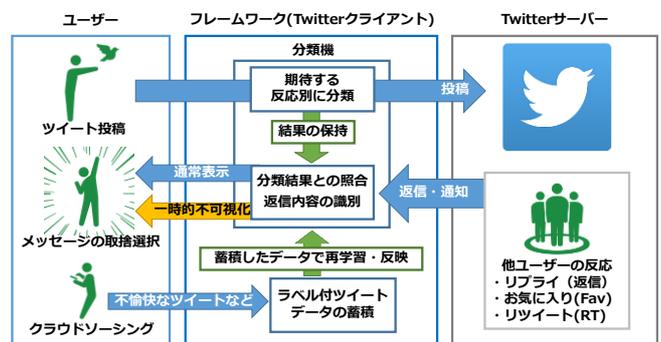


図 1 Twitter 上における具体的なストレス経験の要因

1.2 本研究の目的

本研究の最終目的は, このような発信者と受信者の解釈のずれを機械学習によって検知し, リプライ (ツイートに対する返信) メッセージを一時的に不可視化し, メッセージを見るか見ないかは元の発信者の意思に委ねるという形での "SNS 疲れ" につながるストレスの軽減を行うことを狙いとしている。また, 図 2 に示すようなシステムとしての実装を視野に入れている。

本論文では, その初期段階として, リプライ形式でないツイートに対する発信者の期待する反応の分類に関する手法の提案と評価を行った。



2. 関連研究

Twitter や SNS を対象にした研究は様々存在するが, ここでは, SNS 疲れの軽減やメッセージによる聞き手の感情予測, 有用な情報の抽出, ツイートの分類研究について示す。

[来迎 14]らは, SNS やコミュニケーションツールにおいて, "気疲れ" を起こす原因の一つとして, 送信者の返信への期待感と受信者の相手のメッセージに対して返信しなければならないという義務感に着目し, 一定時間で消えるメッセージによる, 受信者の義務感を軽減しようとする試みの一環として, メッセージアプリのプロトタイプを実装した。評価項目として, アプリの確認回数や返信に対する時間の短さ, メッセージの内容が短文や返事のみが変わるといったことが想定されており, 今後の本研究の評価としても有用であると考えられる。

[長谷川 14]らは, メール等のオンライン上の対話における発話行為が相手の感情にどのような影響を与えるのかという観点で, 感情予測と喚起する応答文の生成を行った。この研究の応用として, メールなどにおける気分を害するような発話のフィル

タリングといったものが挙げられており、本論文の次段階にあたる、発信に対するリプライの分析において参考になると考えられる。

[Zhao 13]らは、Twitter におけるツイートの中から、Information Needs(質問ツイート)を抽出し、時系列データとして解析を行った。Information Needs の抽出には、特徴選択手法である Bi-Normal Separation (BNS)を使用することで、特徴次元数を削減し、素性別に作成された Support Vector Machine (SVM) 分類器を AdaboostDiv という独自のブースト手法によって、精度を高めた。この研究では、Information Needs を「?’を含むもの」という前提を置いているが、これは英語圏特有の文化であり、日本語では、疑問文に対し、「?’を付与しなければならない」という文法規則は無いため、Information Needs の定義をそのまま適用することは難しい。しかし、分類手法として、SVM や BNS といった手法は強力であり、本研究における分類器作成の参考とした。

[原 14]らは、Twitter などのマイクロブログにおけるインフォーマルな書き込みの分類手法の提案を行った。インフォーマルな書き込みとは、災害情報やニュースといった公的な情報ではなく、発信者自身の心情や状態といったメッセージを指す。分類器としては、Parametric Mixture Models(PMM)と SVM を組み合わせたものとしている。こうした書き込みには人々の行動や心理が記述されており、期待する反応の分析に参考となると考えられる。

3. 提案手法

3.1 発信者の期待する反応

発信者の意図と受信者の解釈のずれを検知するために、発信者の期待する、リプライ・Fav(お気に入り)・RT(リツイート)といった他のユーザーからの反応をあらかじめ発信の段階で分類しておき、受信者が元の発信者に対して送ったリプライの内容とのマッチングにより、外れている場合、ずれが生じていると判定するというアプローチが考えられる。このアプローチの利点は、単純にリプライとして送られてきたメッセージだけを解析する場合と比べ、発信したツイートに対する他のユーザーからの反応があった時点で、不愉快になりそうかどうかがある程度分かるため、リアルタイム性が求められる Twitter との相性が良い点である。また、返信メッセージのみを解析し、不愉快につながるかを判定することは TPO に左右されるため、単一のテキストを利用するだけでは困難だと考えられる。

発信者は何かしらの目的を持って発信するものであると考えられる。例えば、何か知りたい情報があれば、その情報が得られるように、質問形式のツイートをし、フォロワーを増やしたいという意図があれば、面白いユーザーになることを目指して、Fav や RT をたくさんしてもらえようという出来事を面白おかしくツイートする。また、発信者は単に発信することを目的にしている場合がある。例としては、「お腹がすいた」や「眠い」といった他のユーザーに対して、作用しないツイートや「○○なう」といった近況報告などが挙げられる。このようなツイートは、他ユーザーからの反応を望まず、実際に発信者が芸能人であるといった特殊な場合を除いて RT されることはまずない。

発信者の期待する反応別に分類することは、Twitter やその他 SNS からの有用性の高い情報を抽出という観点からも有意義なことであり、スパムフィルタやユーザーの意見・関心の抽出、エンターテインメント性の高い画像やテキスト抽出といったものに応用できると考えられる。

3.2 分類体系

本研究では、表 1 に示すように発信者の期待する反応に基づいて、8 つの項目を定義した。各項目を順に追って説明する。

表 1 発信者の期待する反応

ツイートの種類	小分類	期待する反応
ひとりごと		無反応
宣伝・ニュース (AN)	宣伝・告知	RT
	情報・ニュース	
感想・コメント・主張	感想	Fav
	コメント	
	主張	
Topic-providing (TP)	画像有り	Fav・RT
	画像無し	
Information Needs	質問	リプライ
	募集	リプライ・RT
挨拶	おはよう	リプライ・Fav
	おやすみ・ただいま等	
チェーンタグ	RT系	すべて
	ふぁぼ系	
	リプライ系	
	フォロー型	
	複合型	

(1) ひとりごと

ひとりごとツイートは自身の身の回りの出来事や予定、現在地、身体状態、愚痴、思考をまとめるためのメモ書き、非リプライ形式の特定の他ユーザーへの呼びかけ、言及といった内容のツイートである。他の項目と比べ、自己完結性や他のユーザーが知りえない領域への言及となり、局所的なコンテキストを持つ。そのため、他ユーザーからのリプライや Fav, RT は期待しないものと推察される。

(2) 宣伝・ニュース(AN)

AN ツイートは、店舗やサービスの宣伝、社会的な出来事のニュースといった、Fav で個人的な備忘録としてとどめておいてもらうこと以上に、RT してもらうことで、たくさんの人に伝えたい意図が含まれているものである。ニュースの発信源と内容における主体となる対象は異なる場合が多く、リプライは期待しないものと考えられる。

(3) 感想・コメント・主張

感想・コメント・主張ツイートは、ユーザーが TL 上の話題や社会情勢・TV 番組といった、大域的な話題に対する感想や意見、立場を明らかにするものである。具体的には、ツイートにハッシュタグを付与し、テレビやラジオ番組への感想や実況を述べるものや、AN ツイートの引用 RT, TP ツイートを RT した後に感想を述べるといったものが挙げられる。こういったツイートは積極的な議論よりも、同意や共感といった Fav を期待するものと考えられる。リプライによる批評や批判、RT による拡散は炎上のきっかけとなりうるため、原則期待しないものと考えられる。

(4) Topic-providing(TP)

TP ツイートは、非日常的な場面の画像や面白いと思われる稀有な体験談、会話内容を TL 上に発信すること、いわば TL への話題提供を行うツイートで、Fav や RT をしてもらうことを意識したものである。性質上、フォロワーでないユーザーからの的外れなリプライに晒される機会が多くなることや内容を脚色している場合も多いと考えられるため、リプライは期待しないものと考えられる。

(5) Information Needs(IN)

IN ツイートは、TL や特定ユーザーへの質問となっており、知識の正誤確認や物事の詳細・おすすめ、特定個人の情報を尋ねるものとなり、他者の見解や歌詞から曲名、画像をもとに、正

式名称を教えてもらうといった例が見受けられる。こういった性質上、発信者はフォロワーの中にこたえてくれる人がいるという期待の上で、発信していることから、Fav や RT は期待しておらず、回答や提案のリプライを期待しているものと考えられる。

(6) 募集

募集ツイートは、AN ツイートと同様にたくさんの人に見てもらいたいという意図があるが、さらに情報や条件に該当する人や興味を持った人からの返信を期待しているものとしている。AN ツイートの特性に加え、リプライも期待しているものと考えられる。

(7) 挨拶

挨拶ツイートは、実世界で使われる挨拶を TL に対して発信するものであり、たびたび略されることが多い。挨拶は特に相手からの返信・反応を期待していると考えられることから、リプライや Fav を期待しているものと考えられる。挨拶ツイートに対して、挨拶で返すやりとりもたびたび見受けられる。

(8) チェーンタグ

チェーンタグツイートはハッシュタグに期待している反応(リプライ, Fav, RT, フォロー)を明示または、”繋がりたい”といったフォローを催促するといった内容を含むものである。このようなツイートは、フォロワーが RT したり、ハッシュタグをコピーしてつぶやいたりすることによって広がっていく。このようなツイートでは、ハッシュタグ内に書かれている反応を欲していると考えられる。

3.3 分類手法

本研究では、ツイートの分類に際し、教師あり学習手法の一つである Support Vector Machine(SVM)を利用し、先述のツイートの種類別に分類器を作成した。

(1) 教師ラベル付け

Twitter の StreamingAPI を用いて取得した日本語のツイートに対し、公式クライアントや bot 以外の一定数のユーザーが使用しているサードパーティ製のクライアントからツイートされた 22010 件に対してラベル付けを行った。ラベル付けに際して、表 2 に示す各項目の表現や特徴に基づいて、ラベルを付与した。また、ラベル付けの結果を表 3 に示す。

表 2 分類作業における手がかり表現・特徴

ツイートの種類	分類における手がかり表現・特徴
ひとりごと	体の状態を示す語や現在地、予定などの場所を示す語、詠嘆から始まるツイート、1単語といった参照関係が不明なもの
宣伝・ニュース (AN)	宣伝としては日時や場所といった情報や告知が含まれているものニュース記事のURLを含み見出しや概要が記述されているもの
感想・コメント・主張	”かわいい”や”面白い”といった物事に対する評価を示す形容詞を含むものや”好き”など対象を評価するもの、番組名や放送局のハッシュタグを付与しているもの、”>RT”といった引用を示すもの、社会情勢に関するニュースを受けて、明示された物事に対して論じているもの
Topic-providing (TP)	画像有りのものに関しては、短文や”これはwww”といった画像に関する説明を含むもの 画像無しに関しては、”#方言でだからダメって言ったでしょ”といったユーザーに対するお題のハッシュタグを含むものや”I”で会話形式をとるもの、診断メーカーなどのサービスによるツイートといったもの
Information Needs	”どこ”や”〜か”、”〜なの?”といった質問文の形式をとるもの ”おすすぬめありますか”といった他者の意見を募るもの
募集	”募集”や”〜な人いませんか”などの返信を募るキーワードを含むツイート
挨拶	”おはよう”、”おやすみ”、”いただきます”、”ただいま”、”おやすみ”といった挨拶表現や”おは”などの略語を含むツイート
チェーンタグ	ハッシュタグに”フォロー”や”リプライ”、”Fav”、”RT”といったキーワードを含むもの

(2) 前処理

URL などの英字文字列による素性数の肥大化を防ぐため、前処理として、URL(http://~)と@(ユーザー名)の正規化を行った。URL は”URL”、ユーザー名に関しては、”@USERNAME”変換することとした。

表 3 ラベル付け結果

ツイートの種類	個数	割合
ひとりごと	14409	0.655
宣伝・ニュース	1688	0.077
感想・コメント・主張	1153	0.052
Topic-providing	2345	0.107
Information Needs	259	0.012
募集	501	0.023
挨拶	1171	0.053
チェーンタグ	465	0.021
複数ラベル	19	0.001
合計	22010	1.000

(3) 素性

素性には N-gram(N=1,2,3)を使用した。N-gram を採用した理由は、予備実験において、日本語形態素解析器 kuromoji[Atilika 12]との比較において、N-gram がより精度よく分類できたことと、[Zhao 13]においても、N-gram が最も精度よく分類できたとの結果を受けたためである。

(4) 特徴選択

N-gram では、形態素解析器などの方式と比べ、冗長性の高い素性が多く発生し、特徴次元数が大きくなってしまいうる欠点がある。そこで用いられるのが特徴選択手法であり、代表的なものとして、Information Gain や TF-IDF(IDF)などがあげられる。こうした特徴選択手法は、次元の削減だけでなく、精度の向上も見込むことができるため有用である。本研究では、Bi-Normal Separation(BNS)[Forman 08]を用いた。BNS は従来手法よりも、精度良く特徴的な素性を選別できるとされているため採用した。BNS は以下の式で定義されている。

$$\|F^{-1}(tpr) - F^{-1}(fpr)\|$$

$$tpr = tp / (tp + fn)$$

$$fpr = fp / (fp + tn)$$

ただし、 F^{-1} は逆正規累積分布関数。tp は正例のサンプルケースのうち、ある素性を含む数、tn は含まない数であり、fp, fn はその負例の場合となる。

(5) 不均衡データへの対応

表 3 に示した”Information Needs”は他の分類に比べて絶対数が少なく、事前実験においても、識別がうまくいかなかった。そこで学習の際、負例のサンプリングを行うことで、負例を半数程度に減らし、正例と負例のバランスを調整した学習を行うこととした。また、”ひとりごと”に関しても、正例が過半数を超えてしまうことから、正例と負例を入れ替えて学習すると手法はあるものの、先述の Information Needs と同様に、正例側をサンプリングし、負例側が過半数となるようバランス調整を行うこととした。

4. 評価実験

4.1 実験方法

作成した 8 つの分類器に対し、ラベル付けを行った 22010 件のうち、単一のラベルのみがついた 21991 件のデータを用いて、10 重交差検定を行った。評価指標として、Precision, Recall, F 値を用いた。

4.2 実験結果

実験結果を表 4 に示す。灰色の網掛け部分は、不均衡データへの対応のため、学習時に正例、もしくは負例に対するサンプリングを行ったものであり、10 重交差検定を 10 回行った結果の平均を示している。F 値(BNS 使用)に関して、斜字で強調して

いるものは BNS を使用しない F 値から、精度の向上が見られた値であることを示している。

表 4 実験結果

ツイートの種類	Precision	Recall	F 値	F 値(BNS 使用)
ひとりごと	0.87464	0.88185	0.87830	0.87307
宣伝・ニュース	0.78708	0.66020	0.71808	0.74901
感想・コメント・主張	0.75764	0.09646	0.17114	0.16470
Topic-providing	0.63049	0.25428	0.36240	0.58789
Information Needs	0.36016	0.32714	0.34218	0.31512
募集	0.93618	0.58950	0.72345	0.77661
挨拶	0.91749	0.81802	0.86491	0.87169
チェーンタグ	0.88872	0.70113	0.78386	0.83937

4.3 考察

F 値 0.7 を超える高い精度を出し、BNS を用いることでさらに精度が向上した宣伝・ニュース、募集、挨拶、チェーンタグツイートはある種の“キーワード”を持つため他と比べて高い精度が得られたと考えられる。また、BNS による特徴選択によって、正例と負例のどちらにも出現する特徴的でない素性が排除されることによって、精度がさらに向上したことから、BNS の有用性がうかがえる。

ひとりごとに関しては、全分類中最も高い F 値を出したものの、BNS を利用することによって精度が低下してしまった。ひとりごとでは、「あっ」や「え?」といった詠嘆や「あああああ」といった一文字の連続などツイートを含むことから、意味を持たない素性が多く発生してしまう。こうしたツイートを N-gram で 1~3 文字の素性に変えてしまうと、正例・負例の出現数の差が大きくなり、ひらがなの素性を特徴的でないものとして、ストップワードとして切り捨ててしまうために、精度の向上が見られなかったのではないかと推察される。この対策として、一文字または単語の連続といったものは情報量の少ないツイートであるとして、前処理の段階で別の処理による分類を行い、SVM を使って学習・分類させないといったことがあげられる。

感想・コメント・主張に関しては、Recall が著しく低いことがわかる。分類結果の真陽性を見ると、“>RT”という引用符やハッシュタグ、“わかる”といった感想ツイートが占めており、“好き”や“面白い”といった感想や社会情勢に関する主張や考察に関するツイートは全く検知できていなかった。N-gram では、品詞を考慮することができず、また、物事に対する考察などのコメントや社会情勢を踏まえての主張は発信者個人の特有の言い回しなどが感想に比べて多いため、分類項目を通しての特徴が見出しにくいのではないかと考えられる。対策として、感想・コメント・主張という 3 つの小分類のそれぞれに対し、独立した分類器を作成し、それぞれの特徴に対する分類性能を追求するといったことが考えられる。

Topic-providing に関して、画像有りや画像無しで特徴とする素性に大きく差が出たことが、精度低下の要因となっていると考えられる。TP ツイートでは 1 つのツイートで内容が完結している場合が多く、画像有りでは画像の情報量が大きいため、数文字程度の説明や感想にとどまり、画像無しでは、長文となる傾向にあることが分かった。こうした特徴を踏まえると、精度の向上のためには、画像の有無で分類器を分ける必要があるといえる。

Information Needs に関して、本研究では“?”に依存しない定義をしているため、精度が低いものとなっている。絶対数が少ないことが一番の要因であると考えられるため、本論文では対象外としている、リプライ形式の IN ツイートも含めることで、改善されると考えられる。

5. おわりに

本論文では、リプライ形式でないツイートに対し、ユーザーの期待する反応別の分類手法の提案と評価を行った。Twitter におけるストレス経験のうちの一つである、論点のずれたやり取りの原因を発信者の意図と受信者の解釈のずれと仮定した。そのずれの検知のためには、ユーザーの期待する反応を知る必要があり、反応別の分類の定義を示した。分類の定義に基づいてラベル付けを行い、BNS による特徴選択の後、SVM による分類を行った。評価実験により、各分類項目の精度と改善点が明らかになった。本研究における目的である、“SNS 疲れ”軽減のための発信者の期待する反応を分類する見込みが立ったと同時に、提案手法の応用として、広告・ニュース分類器のスパムフィルタとしての利用やユーザーの属性の推定といったものが挙げられる。今後の課題として、本論文では対象から外したリプライ形式のツイートへの対応や教師ラベルの一般性を保証するための複数人でのラベル付け、新語への対応やより一般的なコーパス作成のためのクラウドソーシングのプラットフォーム開発といったものが挙げられる。

謝辞

本研究は JSPS 科研費 24300005, 26330081, 26870201 の助成を受けたものです。

本研究を遂行するにあたり、研究の機会と議論・研鑽の場を提供して頂き、御指導頂いた国立情報学研究所/東京大学 本位田 真一 教授をはじめ、活発な議論と貴重な御意見を頂いた研究グループの皆様へ感謝致します。

参考文献

- [Zhao 13] Zhe Zhao, Qiaozhu Mei: Questions about Questions: An Empirical Analysis of Information Needs on Twitter, Proceedings of the 22nd international conference on World Wide Web (WWW2013), pp.1545-1555, 2013.
- [Yamakami 12] Toshihiko Yamakami: Towards understanding SNS fatigue: exploration of social experience in the Virtual World, 2012 7th International Conference on Computing and Convergence Technology (ICCT), pp. 3-5, 2012.
- [Forman 08] George Forman: BNS feature scaling: an improved representation over tf-idf for svm text classification, CIKM '08 Proceedings of the 17th ACM conference on Information and knowledge management, pp.263-170, 2008.
- [加藤 13] 加藤千枝: 「SNS 疲れ」に繋がるネガティブ経験の実態: 高校生 15 名への面接結果に基づいて(研究), 社会情報学 2(1), pp.31-43, 2013.
- [来迎 14] 来迎直裕, 小笠原直人, 佐藤隆, 布川博士: 消えるメッセージによる義務感を軽減するコミュニケーションツール, 情報処理学会研究報告. EC, エンタテインメントコンピューティング 2014-EC-31(1), pp.1-6, 2014.
- [長谷川 14] 長谷川貴之, 鍛冶伸裕, 吉永直樹, 豊田正史: オンライン上の対話における聞き手の感情の予測と喚起, 人工知能学会論文誌 29 巻 1 号 SPI-J, pp.90-99, 2014.
- [原 14] 原正和, 浅井拓海, 高橋寛幸, 但馬康宏, 菊井玄一郎: マイクロブログのインフォーマルな書き込みに対する自動分類, 情報処理学会研究報告. MPS, 数理モデル化と問題解決研究報告 2014-MPS-97(25), pp.1-2, 2014.
- [Atilika 12] 日本語形態素解析エンジン「kuromoji」
<http://www.atilika.com/ja/products/kuromoji.html>.