

# オンライン上の対話における聞き手の感情予測

長谷川貴之<sup>†</sup> 鍛冶 伸裕<sup>††</sup> 吉永 直樹<sup>††</sup> 豊田 正史<sup>††</sup>

<sup>†</sup> 東京大学大学院情報理工学系研究科 〒113-8654 東京都文京区本郷 7-3-1

<sup>††</sup> 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

E-mail: †{hasegawa,kaji,ynaga,toyoda}@tkl.iis.u-tokyo.ac.jp

あらまし 円滑な対話コミュニケーションを行うためには聞き手の感情を考慮することが重要である。しかし、これまでの対話研究では発話者の感情を推定することに留まっており、聞き手の感情を予測することは試みられてこなかった。本論文では、オンライン上の対話において聞き手の感情を予測する統計的手法を提案する。我々はマイクロブログ上の大規模な対話データから発話に発話者の感情を自動的にタグ付けした教師付きコーパスを作成する。このコーパスを作成するために少数の感情表現を手がかりとしたタグ付け規則を利用する。提案手法では、直前の対話相手の発話だけでなく、それ以前の2人の対話からも素性を作成する。実験を通して直前の対話相手の発話のみから作成した素性を用いた手法よりも提案手法の予測精度が高くなることを示す。

キーワード 対話, 感情, 分類

Takayuki HASEGAWA<sup>†</sup>, Nobuhiro KAJI<sup>††</sup>, Naoki YOSHINAGA<sup>††</sup>, and Masashi TOYODA<sup>††</sup>

<sup>†</sup> Graduate School of Information Science and Technology, the University of Tokyo Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8654 Japan

<sup>††</sup> Institute of Industrial Science, the University of Tokyo Komaba 4-6-1, Meguro-ku, 153-8505 Japan

E-mail: †{hasegawa,kaji,ynaga,toyoda}@tkl.iis.u-tokyo.ac.jp

## 1. はじめに

人間が円滑な対話コミュニケーションを行うためには相手の感情を考慮することが重要である。例えば、相手の感情を害するような発言を避けるなどといったことは誰もが日常的に行っていることであろう。ある発話によって聞き手に喚起される感情を計算機で予測することが可能になれば、コミュニケーション支援に有用である。例えばオンラインのコミュニケーションやカスタマサービスなどにおいて、聞き手が気分を害するような発話を自動的に検出し、書き手に対して警告するようなシステムへ応用が想定できる。

従来、感情に着目した研究としては、テキスト/発話の書き手/話し手の感情を推定する研究が多い[1]~[3]。受け手側である読み手/聞き手の感情を推定する研究もなされているものの、ニュースや人の物語などのテキストを対象としており[4],[5]、対話における聞き手の感情を予測する研究はない。対話は、対話の履歴が存在するという点で、ニュースのようなインタラクションのないまとまった量で与えられるテキストとは異なる性質を持つ。この性質により、対話特有の素性を考慮した手法を適用できる。

そこで本論文では、対話において聞き手の感情を予測するタ

発話	感情
A: 一緒に夕食にいかない?	
B: すみません。私は 38 度の熱があるため行けません。	
A: え!? そうなの!? すぐに良くなるといいですね。	驚き
B: ありがとう。そういってくれて嬉しいよ。	喜び

表 1 感情付き対話の例: 最初の列は 2 ユーザーによる一連の発話であり、2 列めは発話から推定された発話者の感情を示している。

スクに対し、実現するための方法を提案する。このタスクでは、システムは与えられた対話履歴に対して聞き手の感情を出力する。本研究では、単純化のため、対話は 2 ユーザーによって行われると考える(表 1)。感情予測の対象となるユーザーをターゲットユーザー、その対話相手に対話パートナーと呼ぶ。対話パートナーの発話で終わる対話履歴が与えられたとき、システムはその対話履歴からターゲットユーザーの感情を予測する。例えば、表 1 における「え!? そうなの!? すぐに良くなるといいですね」までの発話が対話履歴として与えられればターゲットユーザー B の感情として喜びを出力とする。

我々は上記のタスクに対して統計的手法の適用を試みる。近年では、ソーシャルネットワークサービス(SNS)の普及に伴い、SNS 上に大量の対話データが蓄積されつつあり、統計的学

ユーザー数	672,937
対話数	311,541,839
発話の異なり数	1,007,403,858
対話数 / ユーザー数	463.0
発話数 / ユーザー数	1497.0
発話数 / 対話数	3.2

表 2 Twitter から抽出した対話の統計

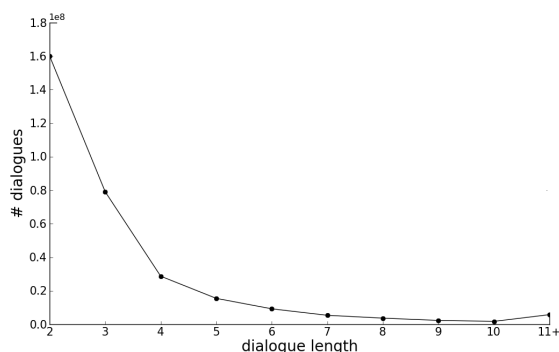


図 1 対話数と対話の長さの関係

習手法の適用可能性が広がりつつある。そこで、そうした SNS 上の対話データの 1 つである Twitter のデータに対して自動的に感情タグを付与することにより、統計的学習に必要な大規模な訓練コーパスの構築を行った。その上で、聞き手の感情を予測するために一般的な分類器を用いた手法を提案する。ここでの我々の貢献は、ニュースのようなインタラクションのないテキストでは利用できない素性の有効性を確認したことである。対話パートナーの直前の返答だけでなく、ターゲットユーザーの直前の発話、さらには対話パートナーのさらに前の発話というように対話履歴から素性を作成する。それらに加えて、それぞれの発話から推定された発話者の感情も素性として利用した。

評価実験では、日本語のツイートから約 3 億の対話を自動的に構築し、そのコーパスを利用して聞き手の感情を予測するための分類器を学習した。評価実験において、提案手法は対話履歴を用いない手法よりも精度を向上させることができた。

以下、本稿の構成を述べる。2 節では、訓練コーパスに関する定義とその作り方を説明をする。3 節では、対話における聞き手の感情を予測する手法について述べる。4 節では、評価実験について述べる。5 節では、関連研究について述べる。そして 6 節にて本稿をまとめる。

## 2. 感情タグ付き対話コーパス

教師あり学習を適用する上では、どのように大規模かつ信頼性の高い訓練コーパスを作るかが問題となる。本節では、SNS から大規模な感情タグ付き対話コーパスを構築する方法について述べる。はじめに、我々の対話コーパスのソースとして、Twitter から発話を集める。次に、感情表現を手がかりとすることで発話に発話者の感情を自動的にタグ付けする。

感情	感情表現
怒り	いらいら、腹が立つ、怒る
期待	楽しみ、期待する、わくわく
嫌悪	嫌う、うんざり、不快
恐れ	恐れる、心配、怖い
喜び	嬉しい、幸せ、感激
悲しみ	悲しい、寂しい
驚き	驚く、びっくり、(・・;) )
受容	安心、頼りになる、ほっとする

表 3 感情表現の例

### 2.1 Twitter の対話の取得

まず Twitter REST API<sup>(注1)</sup> を利用することにより発話 (ツイート) をクロールした。クロールされたデータは、2011/3 ~ 2012/12 までの期間に、77 万ユーザーによってツイートされた 55 億の発話で構成されている。次に、Twitter 特有の表現を削除、または変更することで、クロールされた発話をクリーニングする。具体的には、引用を示すマークである RT や QT を含むツイートは削除し、URL の文字列を含むツイートでは URL 文字列を「URL」に置き換える。Twitter にはボット (特定のキーワードに反応し自動的に応答するプログラム) と呼ばれるユーザーが存在し、ボットの発話をクロールしてしまうと発話に偏りが起こり感情予測や応答生成に悪影響を及ぼしてしまうため、ツイートのユーザー名に「bot」が含まれるツイートは削除した。また、ツイートがあるユーザーへの返答になっている場合、先頭または末尾にユーザー名が含まれる場合がある。これらは、感情予測器に悪影響を及ぼすため、先頭と末尾に現れる @user\_name は削除する。それ以外の位置にある @user\_name はユーザー名を比較し、発話者のユーザー名であれば「私」、対話相手のユーザー名であれば「あなた」、そのどちらでもない場合は「彼」に置き換える。

次に、前処理した発話から 2 ユーザーによる連続した返答のやりとりを対話として抽出する (表 1)。具体的には、Twitter REST API によって提供される 'in\_reply\_to\_status\_id' のフィールドを利用することによって、ある発話に対して応答している発話を同定する。

表 2 に抽出された対話の統計をまとめた。また、図 1 は対話の長さごとに発話数を描画した図である。ほとんどの対話 (98.2%) は高々 10 回で成り立っていた。

### 2.2 感情表現に基づく発話者の感情タグ付与

このようにして得られた対話データ中の発話に対して、人手で用意した感情表現を手がかりとすることで発話に発話者の感情をタグ付けする。本研究では、Plutchik [6] が定めた基本感情である 8 つの感情 (怒り、期待、嫌悪、恐れ、喜び、悲しみ、驚き、受容) を感情カテゴリとして採用し、それぞれの感情カテゴリにつき平均 10 個の感情表現を手で用意した。表 3 に使用した感情表現の例を示した。

教師あり学習においてはタグ付けの精度が重要となるため、

(注 1): <https://dev.twitter.com/docs/api/>

感情表現を含み、なおかつ、以下の2つの条件を満たす発話をタグ付与の対象とする。

- (1) 感情表現が自立語を修飾していない
- (2) 感情表現が否定、仮定、命令、疑問、譲歩、引用の表現を伴っていない

例えば、「私は怒った父親が怖い」には怒りのタグは付与しない(1つ目の条件により排除される)。この文では、発話者の父親が怒っており、発話者は恐れ of 感情を表現しているため、怒りのタグを付与することは不適切であり、代わりに恐れをタグを付与すべきである。また、「明日は遠足だから雨が降ったら悲しいなあ」に悲しみのタグは付与しない(2つ目の条件により排除される)。この文は可能性に基づく言及となっており、悲しみのタグを付与することは不適切である。条件に使われた表現は、否定(ない、ぬ)、仮定(たら)、疑問(?)などである。

表4に感情タグが付いた発話数とその精度を示す。タグ付けの精度は無作為に選んだ各感情カテゴリ100発話を2人の作業者によって調査した。2人の判定結果のκ係数は0.85となり、ほぼ完全な一致となった。表のタグ付け精度は2人の判定結果の平均であり、嫌悪以外の全カテゴリで95%を超えていた。

### 3. 提案手法

本節では、対話における聞き手の感情を予測する手法を述べる。我々はこのタスクを多クラス分類問題と捉え、2節で述べた感情タグ付き対話コーパスから分類器を学習する。このシステムの入力は対話パートナーの発話で終わる対話履歴であり(例: 表1における最初の2発話)、システムの出力はPlutchikの感情カテゴリの1つである(例: 表1における驚き)。

このタスクとニュースや人の物語を読んだ人の感情を推定する研究[4],[5]との主な違いは、我々のタスクでは対話における感情を扱っている点であり、対話履歴を考慮した手法を適用することができる。そこで、本論文では対話履歴から素性を作成し、それをを用いた手法を提案する。対話履歴からユーザーの過去の状態や過去に起きたイベントを推定することができれば、対話における感情予測に貢献すると考えられる。

本手法で用いる素性は単語 n-gram、発話者の感情である。我々は単語 n-gram 素性 ( $n \leq 3$ ) をそれぞれの発話から抽出する。抽出された n-gram は特定の感情を喚起させるイベントや行為(例: 表1における‘38度の熱’)、発話のスタイルやトーン(例: すみません、そうなの!?)を捉えることができる。それぞれの発話から得られた n-gram 素性は別々に扱う。

n-gram 素性だけではスパースになりやすいため、本手法ではそれぞれの発話から推定した発話者の感情を素性として加える。パートナーの感情はターゲットユーザーの感情に強く影響を与えることが報告されている[7]ため、直前のパートナーの発話から推定された感情素性は、出力であるターゲットユーザーの感情の予測に貢献すると期待できる。さらに対話中にユーザーの感情が変化しない限り、過去のターゲットユーザーの感情がと予測される感情と同一になるであろう。

この感情素性を得るために、2節で述べた感情表現を利用したタグ付け手法を用いる。もし発話中に感情表現がなかった場

感情	精度	発話数
怒り	95.0%	190,555
期待	99.0%	2,548,706
嫌悪	93.0%	475,711
恐れ	96.0%	2,671,222
喜び	95.0%	2,725,235
悲しみ	97.0%	712,273
驚き	97.0%	975,433
受容	97.5%	359,482

表4 感情がタグ付けされた発話の精度と数

合には、各感情カテゴリごとに学習した2値分類器を利用することで発話者の感情を推定する。この2値分類器にはPassive Aggressiveを利用する。2節で得られた感情タグが付与された発話を正例とし、単語 n-gram のみを素性として分類器を学習した。学習データの負例は、正例と同じ数の発話をツイートコーパスから無作為に選んだ。そのまま学習すると人手で選んだ感情表現に大きな重みが与えられてしまうため、感情表現を取り除いてから学習した。こうして8つの2値分類器が学習される。しかし、分類器が複数あるため、複数ラベルが出力される場合がある。このような場合は、Plattのシグモイドフィッティング[8]によって2値分類器のマージンを確率に変換し、その確率が高い2値分類器の出力を採用する。どの分類器もラベルを出力しなかった場合は感情なしのラベルをつける。

### 4. 評価実験

#### 4.1 実験設定

感情タグ付き対話コーパスから感情タグがついた発話までを対話として抽出する(例: 表1における最初の2発話)。感情カテゴリごとに訓練データのサイズが大きく異なるため、怒り以外の感情カテゴリを怒りの発話ペアのサイズに揃えた。各感情ともに19,000対話から17,000を訓練データ、10,000をパラメータチューニング用の開発データ、10,000をテストデータに利用した。分類器としては support-class passive aggressive アルゴリズム[9]を実装した多クラス分類器 opal<sup>(注2)</sup>を用いた。

#### 4.2 実験結果

本研究では、各感情において次の素性セットで分類器を学習し、各感情ごとに精度、再現率、 $F_1$ を比較した。

- $D_{t-1}$  対話パートナーの最も新しい発話から作成した素性
- $D_{t-1}$  &  $T_{t-1}$  対話パートナーとターゲットユーザーの最も新しい発話から作成した素性
- +  $D_{t-4}$  &  $T_{t-4}$  対話パートナーとターゲットユーザーの最も新しい4つの発話から作成した素性

表5に n-gram 素性だけを用いた手法(n-gram)と感情素性を利用した手法(n-gram + emotion)の実験結果を示す。対話パートナーの発話から作成した素性に加えて、ターゲットユーザーの直前の発話から作成した素性を考慮することで、全ての

(注2): <http://www.tkl.iis.u-tokyo.ac.jp/~ynaga/opal/>

感情	n-gram									n-gram + emotion								
	$D_{t-1}$			$D_{t-1} \& T_{t-1}$			$+D_{t-4} \& T_{t-4}$			$D_{t-1}$			$D_{t-1} \& T_{t-1}$			$+D_{t-4} \& T_{t-4}$		
	精度	再現率	F <sub>1</sub>	精度	再現率	F <sub>1</sub>	精度	再現率	F <sub>1</sub>	精度	再現率	F <sub>1</sub>	精度	再現率	F <sub>1</sub>	精度	再現率	F <sub>1</sub>
怒り	0.405	0.463	0.432	0.445	0.512	0.476	0.453	0.510	0.480	0.438	0.474	0.455	<b>0.467</b>	<b>0.519</b>	<b>0.491</b>	0.456	0.511	0.482
期待	0.406	0.548	0.466	0.432	0.567	0.490	0.431	0.570	0.491	0.444	0.555	0.493	0.466	0.574	0.514	<b>0.467</b>	<b>0.575</b>	<b>0.515</b>
嫌悪	0.378	0.353	0.365	0.410	0.380	0.394	0.405	0.395	0.400	0.393	0.390	0.392	<b>0.426</b>	<b>0.398</b>	<b>0.412</b>	0.420	0.396	0.408
恐れ	0.375	0.301	0.334	0.402	0.316	0.354	0.400	0.316	0.353	0.403	0.382	0.392	<b>0.423</b>	0.393	<b>0.408</b>	0.422	<b>0.394</b>	<b>0.408</b>
喜び	0.394	0.449	0.420	0.414	0.480	0.444	0.420	0.477	0.447	0.419	0.470	0.433	<b>0.443</b>	<b>0.485</b>	<b>0.458</b>	0.432	0.476	0.453
悲しみ	0.408	0.355	0.380	0.437	0.377	0.405	0.444	0.383	0.411	0.418	0.372	0.394	<b>0.443</b>	0.394	<b>0.417</b>	0.440	<b>0.395</b>	<b>0.417</b>
驚き	0.352	0.269	0.305	0.388	0.314	0.347	0.394	0.315	0.350	0.388	0.302	0.339	<b>0.408</b>	0.344	0.373	<b>0.408</b>	<b>0.347</b>	<b>0.375</b>
受容	0.399	0.399	0.399	0.419	0.417	0.418	0.424	0.419	0.421	0.421	0.395	0.408	0.439	<b>0.413</b>	<b>0.426</b>	<b>0.440</b>	0.405	0.422
平均	0.390	0.392	0.388	0.418	0.420	0.416	0.421	0.423	0.419	0.416	0.418	0.415	<b>0.438</b>	<b>0.440</b>	<b>0.437</b>	0.436	0.438	0.435

表5 n-gram 素性と感情素性を利用した評価実験の結果

		予測した感情									
		怒り	期待	嫌悪	恐れ	喜び	悲しみ	驚き	受容	合計	
正解の感情	怒り	<b>5186</b>	488	<i>1178</i>	675	499	766	627	581	10000	
	期待	352	<b>5736</b>	425	426	<i>1363</i>	532	600	566	10000	
	嫌悪	<i>1519</i>	596	<b>3981</b>	1141	591	909	657	606	10000	
	恐れ	956	740	1015	<b>3933</b>	564	719	909	<i>1164</i>	10000	
	喜び	478	<i>1736</i>	456	360	<b>4847</b>	590	718	815	10000	
	悲しみ	<i>1073</i>	987	917	665	959	<b>3936</b>	746	717	10000	
	驚き	817	1102	753	956	<i>1318</i>	773	<b>3441</b>	840	10000	
	受容	725	937	620	<i>1146</i>	1045	656	738	<b>4133</b>	10000	
	合計	11106	12322	9345	9302	11186	8881	8436	9422	80000	

表6 全ての対話履歴を利用した手法の分類結果の混同行列。最も多く選んだカテゴリを太文字、最も多く間違えたカテゴリを斜体で示している。

感情カテゴリにおいて分類精度が向上した。さらに、感情素性が分類精度の向上に貢献していることも確認された。n-gram 素性のみを利用した手法では、素性の抽出に用いる対話の履歴を増やすことで分類精度は向上しているが、分類に大きく寄与している発話は2ユーザーの最も新しい発話であった。対話の多くは2ツイートで成り立っており、 $D_{t-1}$  と  $T_{t-1}$  以外の発話から素性を作成しても分類精度には寄与しなかったと考えられる。その一方で感情素性を加えた手法では、履歴を増やすことで分類精度が低下した。これにより、直前の発話 ( $D_{t-1}$  と  $T_{t-1}$ ) 以外から推定される感情は感情予測の精度向上には寄与しないことがわかる。

表6に最も分類精度が高かった分類器(表5の  $D_{t-1}$  &  $T_{t-1}$ )の混同行列を示す。最も多く選んだカテゴリを太文字、最も多く間違えたカテゴリを斜体で示している。この表からいくつかの典型的な混同する感情カテゴリのペアを見て取れる。分類器は嫌悪と怒り、喜びと期待、恐れと受容を正しく分類することが難しかった。この原因は、これらの感情カテゴリが本質的に類似しているためである。Plutchikの分類では、怒りと嫌悪、期待と喜び、恐れと受容がそれぞれ隣接し、これらの組は類似した感情とされており、同様の結果となった。また、驚きは期待や喜びなどのポジティブな感情に誤分類される傾向が強かった。驚きという感情は、驚きと喜び、驚きと期待というように他の感情と同時に共起しやすいため、ポジティブな感情に誤分類されやすかったと考えられる。

### 4.3 分類に寄与した素性

最後に分類器の素性について述べる。分類器のモデルにおいて最も大きな重みを得ていたのは、どの感情カテゴリにおいてもコーパス作成に利用した感情表現であった。 $D_{t-1}$ 、 $T_{t-1}$ に現れる感情表現が上位であったことから、ターゲットユーザーの感情は自身の1つ前の発話と同じ感情である傾向が強くなり、また対話パートナーの感情に影響されていると考えることができる。感情表現以外に大きな重みを得ていたフレーズの上位10件を表7に示した。表中の(t)はターゲットユーザーの発話、(d)は対話パートナーの発話から作成された素性であることを表している。感情素性は感情の名前を中括弧{ }で囲んだ。

この結果から典型的な聞き手の感情を喚起するフレーズを見てとれる。受容においては、無事、なで(なでなで)、大丈夫といった相手を安心させたり、慰めるフレーズに大きな重みが与えられていた。また、お誕生日おめでとう、素敵ですといった相手を祝福するフレーズが感情を喚起させることがわかる。典型的ではないが特徴的である感情が怒りであり、対話パートナーが民主党関連の話題を出すとターゲットユーザーに怒りの感情が喚起されることがわかる。これはツイートを収集した期間における民主党に対する感情が強く反映されている。

### 4.4 誤り分析

誤分類した発話から各カテゴリごとに25発話づつ無作為に選び、その原因について調査した。誤り数が20以上の原因を表8に示した。上位2つは分類器だけでなく人間が誤る原因と

怒り	期待	嫌悪	恐れ	喜び	悲しみ	驚き	受容
東電 (d)	{期待} (d)	{嫌悪} (d)	{恐れ} (d)	フォローさせ (d)	別れ (d)	{驚き} (t)	なでなで (d)
{怒り} (t)	{期待} (t)	梅雨 (d)	ホラー (d)	お誕生日 (d)	お別れ (d)	{驚き} (d)	なで (d)
国会 (d)	レコーディング (d)	やだ (d)	レーシック (d)	はじめまして (d)	{悲しみ} (t)	偶然 (d)	無事です (d)
ぶんぶくり (d)	発売 (d)	{嫌悪} (t)	発熱 (d)	誕生日おめでとう (d)	最終回 (d)	衝撃 (d)	大丈夫です (d)
民主党 (d)	完成 (d)	湿度 (d)	心霊 (d)	お誕生日 (d)	ドロリッチ (d)	大丈夫ですか (d)	無事 (d)
許せ (d)	いよいよ (d)	ジメジメ (d)	嘔吐 (d)	さんの絵 (d)	閉店 (d)	すご (d)	ので大丈夫 (d)
クソ (t)	リハーサル (d)	雨 (d)	不気味 (d)	日おめでとう (d)	亡くなり (d)	バツリ (d)	{受容} (d)
(-)(d)	収録 (d)	梅雨入り (d)	大丈夫ですか (d)	素敵でし (d)	引退 (d)	偶然 (t)	まかせろ (d)
税金 (d)	発売です (d)	湿度 (d)	大丈夫か (d)	素敵でした (d)	ドロ (d)	ばったり (d)	任せ (d)
態度 (d)	リハ (d)	蒸し暑い (d)	余震 (d)	拝見 (d)	リッチ (d)	まさかの (t)	大丈夫です! (d)

表7 分類器の重みのうち、感情表現を取り除いた上位10素性

原因	対話数
情報不足	64
複数感情の喚起の可能性	44
過去の発話の感情を考慮したことによる副作用	26

表8 タグ付け誤りの原因と集計

なり、その他は人間なら判断できるが分類器では誤分類の原因となる。

#### a) 情報不足

与えられた発話から得られる情報量が少ないために、多くの対話で誤分類していた。人間によって情報不足が原因で感情予測が難しいと判断されたケースは3種類ある。

- (1) 少ない文字数、又は突然の話題の展開
- (2) URLのウェブページに関する知識の欠如
- (3) ユーザーの背景知識の欠如

1つ目の原因は、発話の文字数が少なく、n-gram素性が少なくなったことである。もし作成されたn-gram素性の中に大きな重みがついた素性がないならば、分類器が正しく判断することは難しくなる。発話の長さだけでなく、感情を喚起しない対話、例えば日常会話のような対話が与えられたときも分類器は正しく判断できない。なぜなら日常会話の後に感情的な発話が続く対話では、突然の話題の展開をすることが多いからである。このような対話は正確に予測することは不可能であるため、提案手法でも正しく予測できる必要はない。次の例は正解が驚きとなる対話である。この例では、あいさつ後に前日の話題を話している。人間でも「おかえり」の後に前日の地震の話をすることは推測できない。

$T_{t-1}$  ただいまぁー  
 $D_{t-1}$  おかえりい~  
 $T_t$  おかえりありがとうございます とんぴさんもお疲れさまでした 笹 昼間の地震、びっくりでしたね。大丈夫でしたか?

2つ目の原因は、URLのウェブページの内容がわからないことである。例えば、ウェブページが怒っている人が書いたブログであれば、喚起される感情も怒りである可能性が高い。しかし、URLという字面だけではそのような情報を得ることはでき

ない。それだけでなく、ウェブページの内容について対話がない場合も字面だけでは情報が不足してしまう。次の例は正解が怒りとなる対話である。この例では対話パートナーがブログの内容に関してターゲットユーザーをからかっている。

$T_{t-1}$  ブログを更新しました。「おしゃれっ(\* )」 url  
 $D_{t-1}$  とりあえず、背筋伸ばさなきゃな...とか思いましたっ。ピシッ  
 $T_t$  うむッ、よるしいd( ) 背筋伸ばしてからのぉ~、鯛焼きべろんちょ(\* 艸)笑。くおらぁーッ( `´)

3つ目の原因は、ユーザーの背景知識がないことである。例えば、「明日、台風くるらしいよ」と言われたとき、台風が来ることを知っている人であれば驚かず、台風が来ることを知らない人は驚くであろう。しかしながら、分類器はユーザーの背景知識については知らないため、正確な出力を出すことが難しい。次の例は分類器が怒りに分類したが、正解は驚きとなる対話である。この例では漫画の連載が終わったことをターゲットユーザーに伝えている。

$T_{t-1}$  予定なんもなくてヤル気もなかったから抜け殻のようにダラダラ。  
 $D_{t-1}$  刃牙終わったよ。  
 $T_t$  わぁーびっくり!読み直さな

#### b) 複数感情の喚起の可能性

複数の感情が正解となりうる対話において誤分類してしまっていた。複数感情を喚起される可能性があり、感情予測が難しいと判断されたケースは2種類ある。

- (1) 複数感情のうち、どの感情が喚起されるかは受け手によって異なる
- (2) 複数の感情が同時に喚起される

1つ目は、複数の感情(例:喜びと期待)が本質的に類似している場合、複数の感情のどちらも喚起されうることである。本研究では正解を1人のユーザーの感情から作成しているため、異なる個人ごとに異なる感情を喚起されうる対話であっても正解は1つとなっており、それが問題となっていた。しかし、複数個人の感情に対応できるデータセットを用意するには多人数に

同じ対話をしてもらい、そのときの感情を集計する必要があるため、とてもコストが高くなってしまふ。低コストでそのようなデータセットを作ることは現状では難しく、これは今後の課題である。次の例は分類器が喜びに分類し、正解が期待であった対話である。この例では、誕生日を祝う文脈で、対話パートナーがターゲットユーザーに対し、何かすることを仄めかしている。最後の発話の後、実際のデータではターゲットユーザーに期待の感情が喚起されているが、人によっては喜ぶ人もいると考えられる。

D<sub>t-2</sub> 誕生日おめでと~\*° \*.....:\*( \* \*)' \*... ..\*° \*  
 T<sub>t-1</sub> ありありありがとぉんん !!  
 D<sub>t-1</sub> なんが面白いことしないとねえ(=´ ` )人(´ ` =)  
 T<sub>t</sub> おおっwなにしてくれんだろw楽しみ

2つ目は、複数の感情（例：驚きと喜び）が同時に喚起される可能性があることである。例えば、我々は親しい人間の受賞報告を受けるとき、喜びと驚きの感情を持ち合わせるだろう。本研究では短いツイートには基本的に1つの感情しかないと仮定しているため、複数の感情を同時に喚起される対話で問題となっていた。次の例は分類器が喜びに分類し、正解が驚きであった例である。この例では、対話パートナーがターゲットユーザーを祝福すると同時に、誕生日が同じであることを伝えている。人が見れば明らかに驚きかつ喜んでいるが、本研究では正解が1ラベルのみのため、複数ラベルをつけることができなかった。

T<sub>t-1</sub> おはようございます。今日は僕の母ちゃんが頑張った日。父ちゃんが名前決めてくれた日。朝から生まれたときの話聞いている。ありがとう。愚痴も言われてます w  
 D<sub>t-1</sub> おはよう！まさか誕生日が同じだとはw おめでとぅ \ (^o^ ) / 呪呪呪 笑  
 T<sub>t</sub> ありがとうございます° +. \ (´ ` \*)° ほんとにw ビックリしました ww

### c) 過去の発話の感情を考慮したことによる副作用

過去の発話の感情が強く影響したために誤分類してしまうことがあった。提案手法では、ターゲットユーザーは対話パートナーの感情の影響を受けやすい、ターゲットユーザーの感情は過去の感情と同じになりやすいという2つの仮説に基づき、感情素性を利用した。しかし、全ての対話でこのような現象が起こるわけではなく、感情素性に大きな重みが付いたことが原因で間違ってしまうことがあった。次の例は、対話パートナーのD<sub>t-1</sub> から恐れ感情素性が作成されるが、実際のデータではターゲットユーザーが安心している例である。

D<sub>t-1</sub> みなさんありがとうございます(´ ; ;´) 久しぶりに怖かったです...  
 T<sub>t</sub> 津波の心配は無いみたいだから、ちょっと安心。

## 5. 関連研究

従来、感情に着目した研究としては、書き手/話し手の感情にもとづいてテキスト/発話を予測する試みが行われている [1] ~ [3]。これに対し、我々は、書き手/話し手ではなく、その読み手/聞き手の感情に基づく予測を試みている。

テキストの書き手の感情ではなく、読み手の感情を推定する研究については、Lin ら [4] や Socher ら [5] の研究がある。これらの研究では、テキストはまとまった量で与えられ、インタラクションのない性質を持つテキストである。我々の研究は、これらの研究とは対話に焦点を当てている点で異なり、そのようなタスク設定では抽出できない対話履歴から作成した素性を用いている。また、Tokuhisa ら [10] は、あるイベント（例：クリスマスプレゼントをもらう）によって喚起される感情（例：嬉しい）をウェブテキストから獲得している。これらの研究は、我々が行うタスクと問題設定が近いが、対話における感情を扱っているわけではない。しかしながら、得られたイベントと感情のペアは、我々のタスクにおいても利用することが出来る。

Kim ら [7] は対話データにおける感情の移り変わりを調査した。Kim らは発話者の感情を自動推定し、その移り変わりに関する定量的調査を行っている。

## 6. おわりに

本稿では、対話における聞き手の感情を予測するタスクを提案し、その手法を提案した。対話履歴と発話者の感情を素性に加えることで、分類精度を向上させることができた。今後の課題は、多クラス分類以外の方法で聞き手の感情を予測することである。感情の組み合わせや感情の強さを考慮した手法を適用出来れば、より正確な感情予測が可能になる。

## 文 献

- [1] Moataz El Ayadi, Mohamed S. Kamel, and Fakhri Karray. Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, Vol. 44, pp. 572–587, 2011.
- [2] Dipankar Das and Sivaji Bandyopadhyay. Analyzing emotional statements ? roles of general and physiological variables. In *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology*, SAHP' 11, pp. 59–67, Chiang Mai, Thailand, November 2011.
- [3] Alexandra Balahur, Jesús M. Hermida, and Andrés Montoyo. *Detecting Implicit Expressions of Sentiment in Text Based on Commonsense Knowledge*. Association for Computational Linguistics, Portland, Oregon, June 2011.
- [4] Kevin Lin and Hsin-Hsi Hsin-Yih. Ranking reader emotions using pairwise loss minimization and emotional distribution regression. In *Proceedings of EMNLP*, pp. 136–144, 2008.
- [5] Richard Socher, Jeffrey Pennington, Eric H. Huang, Andrew Y. Ng, and Christopher D. Manning. Semi-supervised recursive autoencoders for predicting sentiment distributions. In *Proceedings of EMNLP*, pp. 151–161, 2011.
- [6] Robert Plutchik. A general psychoevolutionary theory of emotion.

In *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion*, pp. 3–33. New York: Academic, 1980.

- [7] Suin Kim, JinYeong Bak, and Alice Haeyun Oh. Do you feel what I feel? social aspects of emotions in Twitter conversations. In *Proceedings of ICWSM*, pp. 495–498, 2012.
- [8] H.T. Lin, C.J. Lin, and R.C. Weng. A note on platt 's probabilistic outputs for support vector machines. *Machine Learning*, Vol. 68, No. 3, pp. 267–276, 2007.
- [9] Shin Matsushima, Nobuyuki Shimizu, Kazuhiro Yoshida, Takashi Ninomiya, and Hiroshi Nakagawa. Exact passive-aggressive algorithm for multiclass classification using support class. In *Proceedings of SDM*, pp. 301–314, 2010.
- [10] Ryoko Tokuhisa, Kentaro Inui, and Yuji Matsumoto. Emotion classification using massive examples extracted from the Web. In *Proceedings of COLING*, pp. 881–888, 2008.