

顔文字の役割に着目した ツイートの文の感情抽出手法の提案

◎山本 湧輝(甲南大学)

若井 祐樹(甲南大学)

熊本 忠彦(千葉工業大学)

灘本 明代(甲南大学)

はじめに

Twitter

- 140文字以内の短文を投稿
- 気軽に発言できる



かつとる

まあ永夜抄が全体見ても一番わかりやすく**楽しい**からなあ
開く

1分



リアルリザードンらしいよおーうち

でもああいうとこでわいわいするの**楽しい**よねえ(´-`-´)´
開く

1分



あぐれっしも(又の名をサリ)

ぴよんだむと赤風さんが後ろにいるのにここにいて もう なに なん
か **楽しい**。

開く

1分



ちよりんぬ。

やっぱり学校は**楽しい**のっ(*w`*)! でもねむーい。動画みたいっ
開く

1分



綾瀬ハル

@tempus_anni **楽しい**疲れですw

● 会話を表示

1分

気軽にツイート



今思っていること
や意見を投稿する



その人の感情が出
やすい

感情を含んだツイート

顔文字なし

こんな天気の良い日はわくわくするな



楽しい

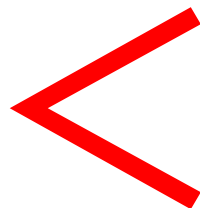
顔文字あり

こんな天気の良い日はわくわくするな(*´▽`*)



より楽しい

顔文字なしの
感情



顔文字ありの
感情

感情を含んだツイート

顔文字なし

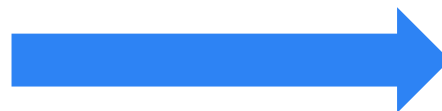
むかつくからブロックしよ



怒ってる

顔文字あり

むかつくからブロックしよ \ (^o^)/



あまり
怒ってない

顔文字なしの
感情

≠

顔文字ありの
感情

目的

ポイント

顔文字が入るとツイートの感情が変化する



文だけではなく...

顔文字を考慮したツイートの感情抽出

期待できること

- 映画の要約
- おすすめの本を推薦

ホントに悲しい(*^^*)

文

顔文字

1 文から感情値の抽出

2 顔文字の感情抽出

3 文と顔文字の関係から役割を考慮

文から感情値の抽出



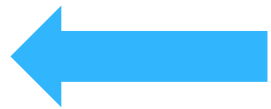
Twitter

こんな天気の良い日はわくわくする



形態素解析

こんな/天気/の/いい日/わくわく/する



感情語辞書

	喜	好	安	哀	厭	怖	怒	恥	昂	驚
わくわく	0.76	0	0	0	0	0	0	0	0.3	0
いい	0.25	0.52	0.21	0	0	0	0	0	0	0



各感情値の合計を求める

	喜	好	安	哀	厭	怖	怒	恥	昂	驚
感情値	1.01	0.52	0.21	0	0	0	0	0	0.3	0

感情軸と感情語辞書

感情軸

- 中村明が構築した10軸を用いる

喜 好 安 哀 厭 怖 怒 恥 昂 驚

感情語辞書

- 若井らが構築した感情語辞書を用いる

感情軸

喜		好		安	
楽しみだ	0.85	にこやかだ	0.9	安心	0.72
大切だ	0.72	アルバム	0.45	心地よい	0.23
クローバー	0.32	愛犬	0.72	安堵	0.12

感情値

ホントに悲しい(*^^*)



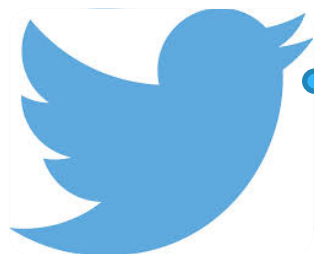
顔文字

1 文から感情値の抽出

2 顔文字の感情抽出

3 文と顔文字の関係から役割を考慮

顔文字から感情の抽出



Twitter

こんな天気のいい日は
わくわくする
な(*´▽`*)



こんな天気のいい日はわくわくするな(*´▽`*)

顔文字辞書

顔文字

感情

(*´▽`*)

喜

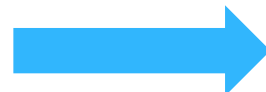
顔文字辞書構築の手順

1. 教師データとする顔文字を集める
2. 10人により教師データの顔文字の感情を決定
3. 教師データの顔文字をパーツに分類
4. パーツごとの感情を決定
5. 4で抽出したパーツをもつ顔文字400個を収集
6. 400個の顔文字の感情をパーツに基づき決定
7. 顔文字辞書に感情の決定した顔文字を追加

教師データとする顔文字

出現頻度が高い

顔文字を含む17,647ツイート



顔文字を100個収集

実験に使用した顔文字の例

顔文字	回数	割合
(^o^)	2832	0.160
(^^)	935	0.053
(>_<)	917	0.052
(´·_·`)	742	0.042
(^-^)	527	0.030
(´▽`)	504	0.029
(´·ω·`)	499	0.028
(´·ω·`)	498	0.028
(´·ω·`)	496	0.028
(´;ω;`)	413	0.023
(¯▽¯)	389	0.022
(^ω^)	361	0.020

顔文字辞書構築の手順

1. 教師データとする顔文字を集める
2. 10人により教師データの顔文字の感情を決定
3. 教師データの顔文字をパーツに分類
4. パーツごとの感情を決定
5. 4で抽出したパーツをもつ顔文字400個を収集
6. 400個の顔文字の感情をパーツに基づき決定
7. 顔文字辞書に感情の決定した顔文字を追加

教師データの顔文字の感情

- 目的
「教師データの顔文字の感情を決定する」
- 顔文字
「ツイッターでよく使われている顔文字100個」
- 被験者
「10人」

実験方法

顔文字を見てもらい当てはまると思う感情を選ぶ
(順位の高い順に1→2→3)

実験の例

顔文字	喜	哀	怒	怖	好	厭	昂	驚	安	恥
(^o^)	1位				2位				3位	
(^^)	1位				2位				3位	
(>_<)				3位				2位		1位

顔文字辞書構築の手順

1. 教師データとする顔文字を集める
2. 10人により教師データの顔文字の感情を決定
3. 教師データの顔文字をパーツに分類
4. パーツごとの感情を決定
5. 4で抽出したパーツをもつ顔文字400個を収集
6. 400個の顔文字の感情をパーツに基づき決定
7. 顔文字辞書に感情の決定した顔文字を追加

パーツごとの感情を決定

顔文字のパーツに着目



- 顔文字を「**頬**」「**眉**」「**目**」「**口**」の4パーツに分ける
- 教師データの顔文字の感情を元にパーツごとの感情を決定

目	感情		
^^	喜	好	安
><	哀	厭	驚
..	哀	厭	恥
。。	驚		
≧≦	喜	好	昂

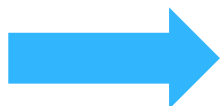
顔文字辞書構築の手順

1. 教師データとする顔文字を集める
2. 10人により教師データの顔文字の感情を決定
3. 教師データの顔文字をパーツに分類
4. パーツごとの感情を決定
5. 4で抽出したパーツをもつ顔文字400個を収集
6. 400個の顔文字の感情をパーツに基づき決定
7. 顔文字辞書に感情の決定した顔文字を追加

顔文字辞書の構築

結果

- 教師データ(100個)+パーツに着目して収集した顔文字(400個)



計500個の顔文字を顔文字辞書に追加

登録されている顔文字の例

感情軸	単語		
喜	(^o^)	(^^)	(^-^)
好	(^o^)	(' ▽ `)	(≧ ▽ ≦)
安	(' ▽ `)	(· ω ·)	(· ▽ ·)
哀	(>_<)	(' · _ · `)	(T o T)
厭	(' · _ · `)	(' · _ · `)	(' ; ω ; `)
怖	(' ; ω ; `)	(((((; ° Д °))))))	(; ° Д °)
怒	(` Д ´)	(` · ω · ´)	(` Д ´)
恥	(>_<)	(^ _ ^ ;)	(" ω ")
昂	(` Д ´)	(≧ ▽ ≦)	(` · ω · ´)
驚	(° Д °)	(¯ Д ¯ ;)	(° _ °)

ホントに悲しい(*^^*)



文



顔文字

1 文から感情値の抽出

2 顔文字の感情抽出

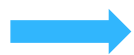
3 文と顔文字の関係から役割を考慮

文と顔文字の関係

同じ顔文字でも文によって使われ方が異なる

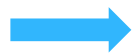
顔文字

(^O^)



めっちゃ楽しい(^O^)

感情を強めている



悲しいよ~(^O^)

感情を弱めている



ポイント

文と顔文字の関係から全体の感情値を算出する



顔文字の役割

役割の種類

強調

顔文字が文の感情を強めている関係

特徴	例
文の感情＝顔文字の感情	めっちゃ嬉しい(*^^*) ほんと辛い(´・ω・｀)

弛緩

顔文字が文の感情を弱めている関係

特徴	例
文の感情↔顔文字の感情	むかつくからブロックしよ＼(^o^)/

自嘲

顔文字があることで、自分を嘲笑うようなもの

特徴	例
否定形+(笑), (^o^)	全然課題終わってないよ(^o^)

顔文字を考慮した感情抽出の流れ

こんな天気の良い日はわくわくするな(*´▽`*)



文の感情

- ツイートを形態素に分ける
- 顔文字辞書を用いて感情値を抽出



顔文字を含む場合

顔文字の感情

- 顔文字辞書とのマッチングにより顔文字の感情を抽出

文の感情

喜



役割を決定



顔文字の感情

喜

$$\text{ツイートの感情値} = \text{文の感情値} \times \text{役割ごとの重み}$$

役割と重みの決定方法

目的

役割と役割ごとの重みを決定する

実験方法

同じ文に対して顔文字が付いている時と付いていないときを評価してもらう

- 被験者 10代～50代の男女10人ずつ計100人
- 質問数 200ツイート

例

今親元と離れたばっかだから心配ごと多い



差



今親元と離れたばっかだから心配ごと多い(´▽`)

顔文字がつくことで感情がどのように変化するかを調べる

実験の結果

- 重み F_i は下記の式で算出

$$F_i = 1 + \frac{y_i - x_i}{\max |y_i - x_i|}$$

i : 文と顔文字の感情の組み合わせ
 F_i : 文と顔文字の感情の組み合わせごとの重み
 y_i : 顔文字を含む感情の合計値
 x_i : 顔文字を含まない感情の合計

重み		文の感情									
		喜	哀	怒	怖	好	厭	昂	驚	安	恥
顔文字 の感情	喜	1.34	0.00	0.45	0.76	1.23	0.78	1.04	0.68	1.13	0.91
	哀	0.59	1.34	0.42	0.93	0.61	1.04	0.83	1.05	1.03	0.98
	怒	0.10	0.56	1.62	1.02	0.85	1.07	0.98	0.70	0.75	1.13
	怖	0.35	0.56	1.63	1.03	0.75	1.32	1.05	0.57	0.73	1.25
	好	1.26	0.12	0.39	0.83	1.23	0.77	1.00	0.65	1.10	1.05
	厭	0.59	1.23	0.42	0.93	0.61	1.10	0.83	1.05	1.03	0.98
	昂	0.93	0.49	1.32	0.72	1.21	0.80	1.04	0.82	0.73	1.35
	驚	0.24	0.43	0.64	0.81	0.58	0.89	1.01	1.67	0.75	0.73
	安	1.16	0.15	0.40	0.88	1.24	0.78	0.91	0.76	1.24	1.11
	恥	1.06	0.45	0.53	0.60	1.41	1.01	1.04	0.77	0.98	1.41

実験の結果

重み		文の感情									
		喜	哀	怒	怖	好	厭	昂	驚	安	恥
顔文字の感情	喜	1.34	0.00	0.45	0.76	1.23	0.78	1.04	0.68	1.13	0.91
	哀	0.59	1.34	0.42	0.93	0.61	1.04	0.83	1.05	1.03	0.98
	怒	0.10	0.56	1.62	1.02	0.85	1.07	0.98	0.70	0.75	1.13
	怖	0.35	0.56	1.63	1.03	0.75	1.32	1.05	0.57	0.73	1.25
	好	1.26	0.12	0.39	0.83	1.23	0.77	1.00	0.65	1.10	1.05
	厭	0.59	1.23	0.42	0.93	0.61	1.10	0.83	1.05	1.03	0.98
	昂	0.93	0.49	1.32	0.72	1.21	0.80	1.04	0.82	0.73	1.35
	驚	0.24	0.43	0.64	0.81	0.58	0.89	1.01	1.67	0.75	0.73
	安	1.16	0.15	0.40	0.88	1.24	0.78	0.91	0.76	1.24	1.11
	恥	1.06	0.45	0.53	0.60	1.41	1.01	1.04	0.77	0.98	1.41

重み ≥ 1.00



強調

重み < 1.00



弛緩

- 「昂」「安」「恥」の軸に関してはあまり値が変化しなかった



この感情は他の感情に引っ張られると考えられる

結果の例

テンション上がってきた

感情

- 喜
- 昂

- アンケートの結果「喜」「昂」の二つの感情に決定した

テンション上がってきた(^O^)

感情

- 喜 → 強調
- 昂 → そのまま

- 喜の感情に顔文字の影響がいき昂には影響していなかった



昂が他の感情に引っ張られると考えられる

顔文字を考慮した感情抽出の流れ

重み		文の感情									
		喜	哀	怒	怖	好	厭	昂	驚	安	恥
顔文字の感情	喜	1.34	0.00	0.45	0.76	1.23	0.78	1.04	0.68	1.13	0.91
	哀	0.59	1.34	0.42	0.93	0.61	1.04	0.83	1.05	1.03	0.98
	怒	0.10	0.56	1.62	1.02	0.85	1.07	0.98	0.70	0.75	1.13
	怖	0.35	0.56	1.63	1.03	0.75	1.32	1.05	0.57	0.73	1.25
	好	1.26	0.12	0.39	0.83	1.23	0.77	1.00	0.65	1.10	1.05
	厭	0.59	1.23	0.42	0.93	0.61	1.10	0.83	1.05	1.03	0.98
	昂	0.93	0.49	1.32	0.72	1.21	0.80	1.04	0.82	0.73	1.35
	驚	0.24	0.43	0.64	0.81	0.58	0.89	1.01	1.67	0.75	0.73
	安	1.16	0.15	0.40	0.88	1.24	0.78	0.91	0.76	1.24	1.11
恥	1.06	0.45	0.53	0.60	1.41	1.01	1.04	0.77	0.98	1.41	

文の感情

喜



強調



顔文字の感情

喜

ツイートの感情値

=

文の感情値

×

1.34

評価実験

目的

顔文字を考慮した感情抽出の有用性を測る

被験者 3人
設問数 顔文字付きの100ツイート

	適合率	再現率	F値
顔文字を考慮しない	0.24	0.37	0.29
顔文字を考慮する	0.38	0.44	0.40

顔文字を考慮した方が「適合率」「再現率」「F値」すべて向上し良い結果となった しかし...

F値	昂	恥	驚
顔文字を考慮しない	0.42	0.33	0.31
顔文字を考慮する	0.37	0.33	0.38

「昂」「恥」「驚」はあまりいい値を示さなかった。
→この感情は他の感情に引っ張られると考えられる

評価実験

目的

顔文字を考慮した感情抽出の有用性を測る

被験者 3人
設問数 顔文字付きの100ツイート

- 顔文字を考慮した場合ツイートが正しい役割になっているか

	喜	好	哀	哀	怒	怖	
強調	0.54	0.63	0.67	弛緩	0.86	0.97	0.90

- 強調より弛緩のほうが全体的に精度は良かった
- 弛緩において「怒」、「怖」などの感情に関しては特によい結果となった

顔文字は「怒」「怖」などの感情とはとても印象が離れているため

結果の例

ほんとムカつく

感情

• 怒

- 実験の結果「怒」の感情に決定した

ほんとムカつく(^O^)

感情

• 怒 → とても弱められる

- 文の怒りの感情が顔文字で打ち消されていた



顔文字は「怒」「怖」の軸とはかけ離れているため大きく感情が打ち消されていた

まとめ

- ユーザ実験を行い感情軸ごとの顔文字辞書を構築した
- 顔文字の役割ごとの重みを決定した
- 文と顔文字の関係からツイートの感情抽出を行った

今後の課題

- 顔文字辞書の拡張
- 顔文字の影響を考慮して感情軸の見直し