Chapter

Utterance Emotion Estimation by Using Feature of Syntactic Pattern

Kazuyuki Matsumoto

Abstract

Emotion has been defined as basic emotions by various researchers, however, there are not many studies describing the relation between emotion and language patterns in detail based on statistical information. There are various languages all over the world, and even a language of the same country has different writing styles/expressions depending on which language media is used or who is a writer/speaker, which is thought to make it difficult to analyze the relation of emotion and language patterns. The author has been engaged in constructing and analyzing emotion corpora in some domains based on different sources. From the analysis results, emotion expressions started to become more understood that they have differences and tendencies according to the attributes of the writers and the speakers. In this chapter, I focused on the differences detected in the attributes of the writer/speaker with respect to language patterns; in usage tendencies or combinations of words, unknown expressions (slangs), sentence patterns, non-verbal expressions (emoji, emoticon, etc.) with relevant emotions, then introduce the outcome of the analytical survey on a large scale corpus obtained from a social networking service.

Keywords: emotion estimation, kansei robotics, slang, emoji

1. Introduction

In the research field of psychology, cognitive linguistics, it has been analyzed and studied about emotion and language [1, 2]. With regard to the relation between basic emotions and language, Fischer [3] performed cluster analysis and created a systematic chart based on emotion categories (emotion word) that can be expressed by language.

In the field of natural language processing, especially, sentiment analysis, a lot of researchers have been engaged in a study on the relationship between language patterns and emotion [4–8]. However, there are various languages all over the world, and language pattern varies depending on language media or writers. For this reason, there are no dictionaries describing language patterns and emotion cyclopaedically.

In the studies by Matsumoto [5] and Tokuhisa [9], they related language pattern dictionaries and occurred emotions. Mera et al. [10] proposed a framework to calculate degrees of positive/negative by using an emotion calculation formula for each case frame pattern. Because most of the methods proposed in these studies were assumed to be applied to “ideal” and “grammatical” sentences, they might not be effective for sentences on Internet.

Matsumoto et al. [11] proposed a method to estimate emotion in utterances including grammatically incorrect expressions such as Internet slangs. In the case of such casual expressions, it is thought to be more effective to take a method by machine learning based on a large scale natural language corpus than to register the knowledge into a dictionary. However, it is difficult to obtain a large scale corpus with labels, and it costs high to make such a corpus. Matsumoto et al. proposed a method to extract features based on word distributed representations as a robust method for unknown expressions. Their method converts words into distributed representation vectors and quantizes them with unsupervised clustering. They demonstrated that the method is robust to unknown expressions compared to existing methods.

After describing emotion estimation methods based on: dictionary, pattern and corpus, we introduce such important elements in corpus-based emotion estimation as gender differences and use of emoji expressions. Then we propose a deep learning-based method that uses a syntactic pattern as a feature combining the corpus-based method and the pattern-based method.

Section 2 introduces the emotion expression dictionary used in our previous research, Section 3 describes the emotion estimation by sentence patterns, and Section 4 explains the corpus-based emotion estimation method. Section 5 analyses emotion estimation with elements of gender and emojis. Section 6 propose a method based on syntax patterns, and Section 7 summarizes this chapter.

2. Emotion expression dictionary

Dictionaries collecting emotion expressions or evaluation expressions already exist [12–14]. These dictionaries defined emotional kinds that can be expressed with the words or phrases as classification categories and are registered them words or phrases. WordNet-Affect is a database created by extending WordNet thesaurus (conceptual database). A part of the information registered in WordNet-Affect is shown in **Table 1**.

There is a study that converted WordNet-Affect into Japanese language [15]. The evaluation polarity dictionary and the Japanese appraisal evaluation expression dictionary are language resources available for reputation analysis or opinion analysis,

A-Labels	Examples
EMOTION	*noun anger#1, verb fear#1*
MOOD	*noun animosisy#1, adjective amiable#1*
TRAIT	*noun aggressiveness#1, adjective competitive#1*
COGNITIVE STATE	*noun confusion#2, adjective dazed#2*
PHYSICAL STATE	*noun illness#1, adjective all in#1*
HEDONIC SIGNAL	*noun hurt#3, noun suffering#4*
EMOTION-ELICITING SITUATION	*noun awkwardness#3, adjective out of danger#1*
EMOTIONAL RESPONSE	*noun cold sweat#1, verb tremble#2*
BEHAVIOR	*noun offense#1, adjective inhibited#1*
ATTITUDE	*noun intolerance#1, noun defensive#1*
SENSATION	*noun coldness#1, verb feel#3*

Table 1.
A-labels and corresponding example synsets.

and they include words with annotation of emotion polarity; positive/negative. To analyze emotion of a sentence written in Japanese, an emotion expression dictionary including Japanese emotion expressions is necessary. It is also necessary to correspond linguistic resources to each language for emotion analysis written in foreign language. Because a framework of linguistic resource might be different according to the kind of language, it is difficult to make a unified dictionary.

In the case of Japanese language, the "Emotion Expression Dictionary" by Nakamura [16] is often referred to and often used in the field of natural language processing. However, many of the expressions included in the emotion expression dictionary are written words appeared in novels, therefore, there are some expressions that are rarely used as colloquial expressions. The Emotion synonym dictionary [17] also includes a few colloquial expressions, listing up the expressions which are thought to be useful for writing novels, scenarios and dramatic dialogs. Currently, as there are no dictionaries that cover practical language expressions such as colloquial expressions, such expressions or patterns are usually extracted from linguistic corpora.

As representative databases with registration of sentence patterns related to emotion expressions, there are EDR electronic dictionary [18], GoiTaikei: A Japanese Lexicon [19], and Kyoto University Case Frame [20]. However, because these linguistic resources are focused on semantic relations, emotion information is not annotated to these databases.

Using dictionaries has an aspect that known knowledge defined by human can be effectively used, however, it is often insufficient when it comes to dealing with things that are greatly related to human sensibilities such as emotions. While some words or expressions always give us unchangeable meanings or impressions, others change their meanings or impressions with the times. For example, the fairness and common sense toward the attributes such as race, religion and gender have changed significantly between decades before and today, so that this issue has been often referred to as one of the problems of artificial intelligence in recent years. Also, as language itself changes, dictionaries need to be updated constantly. In the form of a Wikipedia dictionary, some errors or old information are corrected or updated by being exposed to many people on the Web. However, such descriptions in the Wikipedia dictionary are based on the sensibility of the majority of people, it may not be possible to appropriately estimate the emotions of people with different sensibilities, so there is a limit to emotion estimation with just dictionaries.

3. Relation of sentence patterns and emotion

This section explains the relation of sentence patterns and emotion from the viewpoint of natural language processing by introducing the studies by Matsumoto [5] and Tokuhisa [9]. Matsumoto et al. [5] focused on the emotion occurrence condition for each sentence pattern to estimate emotion in dialog. They also constructed a dictionary that was registered emotion expressions to consider emotion values of each word. The emotion values mean the strength level of expressing each emotion.

Their study used a sentence pattern database that was extended the emotion calculation formula proposed by Mera et al. [10]. However, because they targeted basic sentence patterns, the method has the same problem with the existing method such as lack of versatility and it is weak to spoken expressions. The "Japanese Lexicon" [19] introduces a sentence pattern of each word. In the example of "Crying, "the sentence patterns are:

- N1 ga N2 wo Warau (N1 laughs at N2)

N1 and N2 are nouns. The emotion expressed by the sentence can differ depending on the noun applicable to N1 and N2. Referring to the example sentence: "Jiro cries over his debt," "debt" generally has a negative image. However, the emotion generated in this sentence can be affected by the speaker's attitude to "Jiro." These patterns were necessary to be annotated rules manually. **Figure 1** shows the case frame pattern of "N1 *ga* N2 *de/ni Naku*."

The following table (**Table 2**) shows some examples of sentence patterns and emotion occurrence rules. These information are saved as XML format on account of readability. **Figure 2** shows the emotion occurrence event sentence pattern database with XML format.

Matsumoto et al. [21] also extracted emotion occurrence event sentence patterns from a corpus. The following describes a flow of automatic extraction by Matsumoto et al. showing by example.

Step 1. The inputted sentence is analyzed by dependency parser. "CaboCha [22]"was used as the dependency parser.

First, according to the result of dependency parsing the last segment of the sentence is judged as a predicate. When a segment relates to the predicate and the end of the segment is either case particle or binding particle of "*ga*," "*ha*," "*wo*," "*ni*," "*he*," "*de*," "*to*," "*kara*," "*made*" or "*yori*" is extracted as surface case.

Step 2. The noun included in the obtained surface case element is annotated the semantic attributes based on "A Japanese Lexicon."

If the semantic attributes of the noun cannot be obtained, the basic form of the noun will be set into the surface case slot without annotating semantic attributes. The segment independent from the segment of predicate is not judged as case

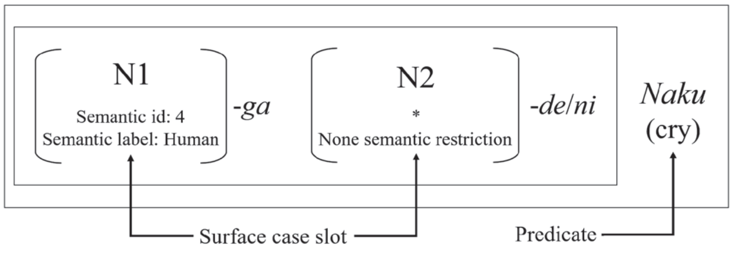

Figure 1.
Case frame pattern of "N1 ga N2 de/ni Naku".

Sentence Patern			Sbj	FVN	EA
English	**Case Pattern**	**Predicate**			
N1 cries over N2	N1[3]-*ga*, N2[*]-*ni/de*	*Naku*	N1	$f_{N2} \le 0$	Sorrow
				$f_{N2} > 0$	Joy
N1 is angry at N2	N1[4]-*ga*, N2[*]-*wo*	*Okoru*	N1	N/A	Anger
N1 laughts at N2	N1[4]-*ga*, N2[*]-*wo*	*Warau*	N1	$f_{N2} \ge 0$	Joy
				$f_{N2} < 0$	Contempt
N1 worries about N2	N1[3]-*ga*, N2[*]-*wo*	*Ureeru*	N1	N/A	Anxiety
N1 is flustered by N2	N1[4]-*ga*, N2[1000]-de, *Ochitsuki-wo*	*Ushinau*	N1	N/A	Surprise
N1 subdues N2's pride	N1[4]-*ga*, N2[4]-*no*, *Hana-wo*	*Oru*	N2	N/A	Shame
N1 discommodes N2	N1[3]-*ga*, N2[4]-*ni*, *Meiwaku-wo*	*Kakeru*	N2	N/A	Hate
N2 is filled with N1	N1[3]-*ga*, N2[4,41,238]-*ni*	*Komiageru*	N2	N/A	N1's *EA*

Table 2.
Example of sentence pattern and emotion occurrence rule.

DOI: http://dx.doi.org/10.5772/intechopen.96597

```
   <!--  感情文型パタン辞書ファイル  -->
 ▼<ESPatternDic name="Jepattern" jname="日本語感情文型パタン">
   ▼<EptItem id="0">
      <pred>愛読|する</pred>
      <read>アイドク|スル</read>
     ▼<emotionInfo>
        <emtag level="3.00" type="main">like</emtag>
      </emotionInfo>
      <image>3</image>
     ▼<e_pattern>
       ▼<case id="0" symbol="N1">
          <type>が</type>
          <sem>4</sem>
        </case>
       ▼<case id="1" symbol="N2">
          <type>を</type>
          <sem>92,011,101,037</sem>
        </case>
      </e_pattern>
    </EptItem>
   ▼<EptItem id="1">
      <pred>愛読|する</pred>
      <read>アイドク|スル</read>
     ▼<emotionInfo>
        <emtag level="3.00" type="main">like</emtag>
      </emotionInfo>
      <image>3</image>
     ▼<e_pattern>
       ▼<case id="0" symbol="N1">
          <type>が</type>
          <sem>4</sem>
        </case>
       ▼<case id="1" symbol="N2">
          <type>を</type>
          <sem>4</sem>
        </case>
      </e_pattern>
    </EptItem>
```

Figure 2.
XML format of sentence pattern database.

element. Because such sentence segment might be important element for deciding emotion attributes, it is extracted as modifier element. The obtained sentence pattern will be as 'EPT.'

Step 3. The set of emotion attributes 'E' annotated to the inputted sentence is decided as emotion attribute of 'EPT.' The combinations of 'EPT' and 'E' obtained from Step1 to Step3 are registered to the emotion occurrence sentence pattern DB. **Figure 3** shows an example of extraction process when "*Watashi wa odoroki no amari me wo shirokuro saseta*." is inputted.

This study automatically extracted sentence patterns from the emotion labeled corpus, created and evaluated the sentence pattern database. As the result of the cross-validation experiments for eight kinds of emotion estimations from sentences expressing emotions based on the corpus-derived sentence pattern database, approx. 42% emotion estimation accuracy was obtained.

Tokuhisa et al. [23] statistically analyzed the valency pattern of each sentence pattern, and proposed a method for emotion inference. Tokuhisa et al. [24] constructed and evaluated the dialog corpus by annotating emotion tags focusing on facial expressions of characters from manga comics.

Their study mainly target the utterances in dialogs, the target data are utterances not by actual persons but by fictional persons. Although these data are simulated

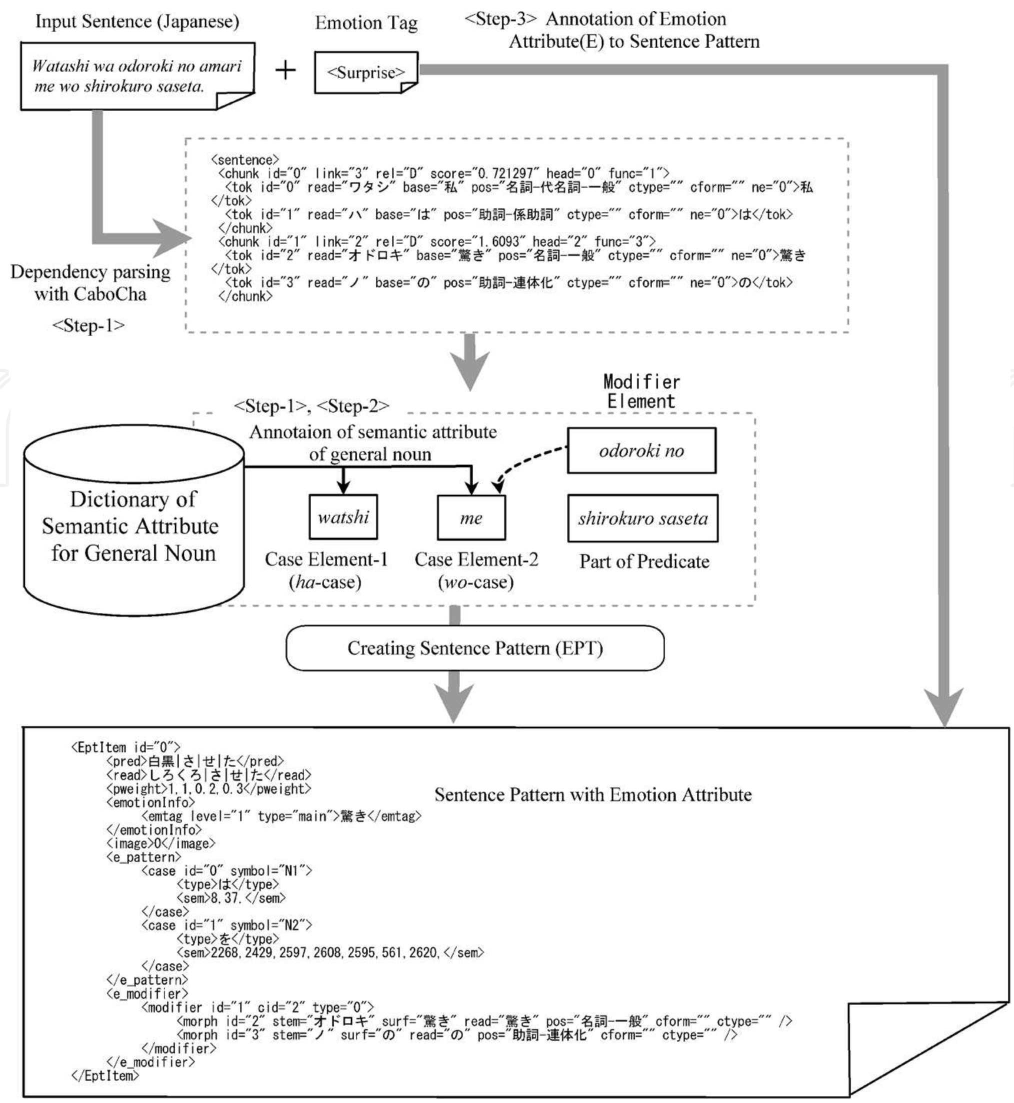

Figure 3.
*Flowchart of creation of sentence pattern (*EPT*).*

real dialog, it is considered that there exist some bias by the authors and generality might be lacking.

It is difficult to register entire colloquial expressions into a dictionary by strictly typifying their sentence patterns, besides, there are few challenging studies that try to annotate emotion that is subjective and sensitive to the sentence patterns. However, I thought that it would not be impossible to extract a relation between emotion and language patterns by studying thoroughly the recent corpus-based methods.

4. Corpus-based emotion analysis method

This section describes a corpus-based emotion analysis method by referring to the related literatures. The corpus annotated with emotion tags is defined as the emotion corpus. We would like to introduce existing studies that created and evaluated emotion analysis models based on statistical information and machine learning using emotion corpora.

4.1 Japanese-English parallel corpus [Minato et al.]

Minato et al. [25, 26] annotated emotion tags on word and sentence units included in Japanese and English parallel corpora. The completed corpus included

1,190 Japanese-English sentences. Based on the statistic results of the tagged words and sentences, they proposed and evaluated an emotion estimation method. They further considered the relevance between the two languages. Overview of their corpus is shown in **Table 3**.

The annotation to the corpus was made by the author, and evaluation by some examinees were not conducted. Matsumoto et al. [27] conducted an questionnaire on this corpus to several examinees and analyzed precision and recall between the tags annotated by the author and the tags annotated by some examinees. Because all examinees were Japanese people, they evaluated only Japanese sentences (1190 sentences). They calculated reliability of the annotation of the emotion tag by multi annotators. Reliability of tag annotation was calculated based on the match frequency among the three operators (initial tag annotator and two examinees). In their study, they proposed a method to reconstruct an emotion corpus by annotating reliability values. Reliability of tag is calculated with the Eq. (1). $\sum W(Tag_x)$ shows the sum of the weight of the tags annotated by the corpus creator and the two examinees

$$\text{Reliability}(Tag_x) = \sum W(Tag_x) \times \left(\frac{\text{Number of Matched Evaluators}}{\text{Total Number of Evaluators}}\right)^2 \tag{1}$$

They calculated the importance for each emotion category by calculating reliability of tag annotation. For calculation they used the weight of emotion tags according to the reliability as weight for emotion category instead of using simple word frequency. The calculation is based on the TFIDF method. Eq. (2) shows the weight of emotion category.

$$W_j^i = \alpha_i \times \sum RW_j^i \times \log \frac{N}{cf_j} \tag{2}$$

$$\alpha_i = \frac{1}{\sqrt{\sum_{m=1}^{l} \left(\frac{\sum RW_j^i}{cf_m}\right)^2}} \tag{3}$$

$\sum RW_j^i$ shows the sum of weight at emotion category 'E_i' in corpus. The 'cf_m' shows the number of emotion category tagged all sentences which included the word

Description	Frequency(J/E)	
Total # of Sentences	1190	1190
Total # of Words (Unique)	14202(2131)	11235(2409)
Average Words per Sentence	11.93	9.44
Total # of Emotion Sentences	601	601
Total # of Emotion Words (Unique)	1220(610)	1249(894)
Total # of Emotion Idioms (Unique)	274(231)	259(248)
Total # of Modifier Words (Unique)	108(70)	39(35)
Total # of Negative Words (Unique)	88(26)	31(15)
Average of Emotion Words/Idioms per Sentence	1.26	1.27
Average of times an emotion word did not carry emotion	0.22	0.15

Table 3.
Corpus statistics.

'w_m'. 'N' shows the total number of emotion category, and 'l' shows the total unique frequency of word. 'α_i' is normalization coefficient which is calculated with Eq. (3).

4.2 Corpus-based method using N-gram [Mishina et al.]

Mishina et al. [28] extracted word n-gram features from the emotion corpora, and proposed an emotion estimation method using the similarity score RECARE which was improved from BLEU often used for translation evaluation. The target emotion categories were four kinds; "anger", "joy", "hate", "hope". The problems of the method are; i) necessary to calculate similarity with all sentences in the corpus, and ii) the estimation accuracy affected by the corpus quality because the method is a simple example-based method.

4.3 Corpus creation and analysis [Quan et al.]

Quan et al. [29] constructed a large size of Chinese weblog emotion corpus "Ren-CECps," and analyzedthe corpus. In Ren-CECps, emotion tags were annotated to sentence, word, paragraph, and article units by some test subjects, and the corpus was analyzed from various viewpoints. The annotation to the corpus required hands, and as the size becomes larger and the corpus includes richer information, the higher annotation costs. There is a demerit that because the target are weblog articles, if there is bias in the writers, that will affect the quality of the corpus.

5. Analysis of emotion expressions according to gender

5.1 The emotion labeled corpus divided according to the users' attributes

In the study of Matsumoto et al. [30], they targeted the tweet sentences posted on Twitter and targeted each tweet for emotion estimation. Therefore, they needed to annotate emotion tags on each tweet sentence. The emotion estimation model was generated with the following steps:

1. The attribute labeled user account list is created from the accounts of popular users whose user attributes are known.

2. The tweets are collected for each user by using the attribute labeled user account list.

3. The four annotators manually annotate emotion tags on the collected tweets.

4. The emotion estimation model is created by extracting features from the tweets and by using a machine learning method.

In Step 4, the feature is extracted. First, the tweet sentence is split into word units by morphological analysis. Then, the words are converted into the distributed representations. They used another corpus to train the distributed representations. For about one year, they continued collecting tweets randomly; then, based on these tweets, they constructed a tweet corpus. They converted the corpus into the word-splitting format and used the text in this format for training the distributed representations.

Then, they annotated the emotion tags on the tweet sentences. Emotion tags annotated to the tweets are as follows:

- Positive emotions: "Joy," "Hope," "Love," "Relief," "Reception"
- Negative emotions: "Anger," "Hate," "Sorrow," "Fear," "Surprise," "Anxiety"
- Other emotions: "No emotion" and "Unclassified"

The total number of the emotion categories is 13. Some examples of the labeled tweets and their user attributes are shown in **Table 4**. The numbers of tweets for each emotion tag are shown in **Table 5**. As shown in **Table 5**, I found that there is bias in the numbers of tweets for each emotion.

In their chapter, they reported that emotion estimation accuracy increase by training the emotion corpus which is prepared for each attributes.

However, one thing to keep in mind when estimating emotion based on the corpus is who to annotate the corpus is. If the annotators' attributes and sensibilities are biased, a biased emotion estimation model would be built by learning the biased corpus. Such model cannot infer appropriate emotions according to the attributes of the authors or the speakers of the object sentence for estimation. To clarify the issue that attributes affect emotion estimation, the next subsection analyses what emotional expressions are used depending on gender based on the corpus.

5.2 Analysis of emotion expressions for each gender

In this section, I analyze emotion expressions by targeting on an emotion labeled corpus that are divided by gender. By investigating appearance frequency of emotion expressions included in the emotion expression dictionary and the kinds of the emotion labels annotated to the tweets including each emotion expression, I analyze appearance tendency of each expression according to gender by TF-ICF.

Emotion Tag	Tweet	Attribute	
		Sex	**Job**
Joy	At that time, we had a good match!	Male	Athlete
Hope	We did good! We should take a good rest and get together in good condition in the next live.	Male	Comedian
Sorrow	Coordinate plans for clothes are ruined by rain.	Female	Musician
Anxiety	All of the roads are in heavy traffic jam... I'm afraid if I could make it to the soccer game...	Female	Athlete

Table 4.
Labeled tweets and attributes of the users.

Joy	**Hope**	**Love**	**Relief**	**Reception**
6020	2135	487	67	15
Anger	**Hate**	**Sorrow**	**Fear**	**Surprise**
150	71	921	78	425
	Anxiety	**No emotion**	**Unknown**	
	301	943	46	

Table 5.
Annotation frequency of each emotion tag.

The formula of TF-ICF calculation is shown as Eq.(4) and Eq.(5). TF means Term Frequency, ICF means Inverse Category Frequency. TF_i^e shows word frequency appeared in emotion category e. ICF_i^e shows the value which is the number of emotion categories divided by the number of emotion categories including word i.

$$TF - ICF_i^e = TF_i^e \times ICF_i^e \tag{4}$$

$$ICF_i^e = \log \frac{|C|}{|\{C : t_i \in C\}|} \tag{5}$$

The TF-ICF calculation results for each gender are shown in **Table 6**. In this table, only top 10 expressions and TF-ICF scores are displayed.

From the analysis result, there are not significant difference between male and female. It is cause that only the general expressions are treated in the emotion expression dictionary for expression extraction.

	Joy					Sorrow			
	Male		Female			Male		Female	
Rank	Word	TF-ICF	Word	TF-ICF	Rank	Word	TF-ICF	Word	TF-ICF
1	*Minna*	309.4	*Minna*	383.3	1	*Konndo*	48.6	*Konndo*	48.6
2	*Ureshii*	188.8	*Ureshii*	323.0	2	*Itsuka*	19.5	*Kawaii*	31.1
3	*Shikkari*	153.7	*Daisuki*	190.7	3	*Kawaii*	17.5	*Itsuka*	31.1
4	*Chotto*	81.7	*Chotto*	118.7	4	*Wa*	17.5	*Mamonaku*	15.6
5	*Omoshiroi*	75.9	*Omoshiroi*	64.2	5	*Shimijimi*	17.5	*Yowamushi*	13.6
6	*Daisuki*	75.9	*Mina*	52.5	6	*Benn*	11.7	*Sugosu*	13.6
7	*Ooyorokobi*	56.4	*Zenzen*	48.6	7	*Tsumaranai*	9.7	*Benn*	9.7
8	*Deru*	52.5	*Zenbu*	42.8	8	*Mamonaku*	9.7	*Shizuka*	9.7
9	*Zenbu*	46.7	*Shikkari*	42.8	9	*Negau*	7.8	*Wa*	9.7
10	*Zenzen*	44.8	*Mattaku*	40.9	10	*Sugosu*	5.8	*Tsumetai*	7.8
	Love					**Anxiety**			
	Male		**Female**			**Male**		**Female**	
Rank	**Word**	**TF-ICF**	**Word**	**TF-ICF**	**Rank**	**Word**	**TF-ICF**	**Word**	**TF-ICF**
1	*Iru*	527.3	*Ii*	439.8	1	*Shi*	3712.8	*Shi*	3537.7
2	*Aru*	467.0	*Aru*	381.4	2	*Mashi*	2006.0	*Mashi*	1605.3
3	*Ii*	393.1	*Iru*	361.9	3	*Nai*	1109.0	*Nai*	1012.0
4	*Naru*	305.5	*Naru*	262.7	4	*Teru*	492.1	*Teru*	643.2
5	*Omou*	155.7	*Omou*	118.7	5	*Tanoshimi*	330.8	*Tanoshimi*	325.0
6	*Taisetsu*	70.1	*Taisetsu*	95.3	6	*Omoi*	321.6	*Ki*	235.5
7	*Hitsuyo*	37.0	*Kanjiru*	35.0	7	*Kore*	202.4	*Kore*	196.9
8	*Kanjiru*	27.2	*Atsui*	33.1	8	*Ki*	177.1	*Kansha*	151.8
9	*Atsui*	23.4	*Machi*	21.4	9	*Ganbari*	173.3	*Suki*	141.4
10	*Machi*	15.6	*Hitsuyo*	19.5	10	*Suki*	169.1	*Shiawase*	132.3

Table 6.
TF-ICF calculation results for each gender.

Joy				Sorrow			
Male		Female		Male		Female	
!	12.565865	!	12.496449	すみません	5.386413	さようなら	6.069605
	12.001341		11.384511	申し訳	5.071520	悔し	6.019864
°	10.443845	^	11.331308	残念	4.998112	メグ	4.512232
・	10.362423	‼	11.190660	相馬	4.815891	迷彩	4.512232
た	10.357401	📘	10.528542	配信	4.690341	・・・・・・・・	4.046404
の	10.321575	📙	10.528542	T	4.098353	イジグ	4.046404
て	9.947300	📗	10.528542	遅れ	3.664781	iPhone	4.027549
児嶋	9.903452	📕	10.528542	連携	3.611918	婆	3.714398
、	9.136461	#	10.465068	皮	3.611918	個	3.649186
飛	9.104408	の	10.240669	負け	3.587754	むくみ	3.611918
に	8.744838	た	10.208031	チャンス	3.583067	(´・(	3.501875
v	8.529375	て	10.022494	店員	3.501875	焼け	3.501875
弦	8.184875	^	9.891287	△	3.501875	ヘルペス	3.465736
o	7.917915	。	9.863882	誠に	3.501875	ごめん	3.337754
^	7.879624	:	9.731163	お祈り	3.501875	残念	3.301151
ます	7.809569	、	8.848594	動け	3.501875	悲しい	3.216726
まし	7.748093	に	8.566338	冥福	3.501875	痛	3.164503
～	7.747877	o	8.515148	亡くなり	3.501875	～(￣◇￣;)	3.092760
競馬	7.607281	୨୧	8.505089	潰し	3.501875	ディズニー	3.034803
#	7.522663	!	8.382803	喫し	3.465736	こぼし	3.034803
が	7.497108	っ	8.344959	痛く	3.452185	豆腐	3.015480
・	7.433959	絢香	8.331537	申し上げ	3.447335	02	3.008155
ありがとう	7.418473	`	8.317766	児嶋	3.301151	mayu	3.008155
で	7.351571	♡	8.261423	サーブ	3.243721	巡礼	3.008155
を	7.336861			痛み	3.143043	どい	3.008155
團	7.319106			マメ	3.143043	老け	3.008155
氣志	7.319106			悔しい	3.136576	剥け	3.008155
				ジャンケン	3.092760	久野	3.008155
				マイル	3.092760	ローブ	3.008155
				ラリー	3.092760	四国	3.008155

Anger				Anxiety			
Male		Female		Male		Female	
ヨー	4.512232	こら	3.865949	怖い	9.831017	予告編	3.611918
くたばれ	4.377344	パンティー	3.008155	つまらない	7.520387	怖い	3.454141
お嬢様	3.034803	被り	3.008155	顔面	4.512232	ストレス	3.034803
覚悟	3.008155	たれ	3.008155	怖	4.512232	渋滞	3.034803
例外	3.008155	出棺	3.008155	偏差	4.512232	下痢	3.008155
虫除け	3.008155	ラブホ	3.008155	闇	3.611918	ドンキ	3.008155
大江	2.626406	じん	2.407946	kunugifujihiko	3.501875	行為	3.008155
うそ	2.626406	たか	2.407946	眠い	3.243721	漏らし	3.008155
変態	2.626406	怒	2.407946	弦	3.015480	腫れ	2.876821
ひどい	2.626406	あやまる	2.407946	しだ	3.008155	心配	2.815267
うるさい	2.407946	シワ	2.407946	チャト	3.008155	地震	2.653141
しつこい	2.407946	(`	2.319570	ださ	3.008155	GP	2.626406
詐欺	2.319570	(-_-#)	2.079442	千賀	3.008155	花粉	2.514435
つく	2.297662	尾行	2.023202	耳鳴り	3.008155	面	2.514435
運	2.122513	ドヤ	2.023202	地震	2.954859	恐怖_	2.514435
パスタ	2.023202	ぶつぶつ	2.023202	怖かっ	2.626406	(;	2.462382
陰気	2.023202	マスコミ	2.023202	グチ	2.626406	怖く	2.462382
噛ま	2.023202	首都	2.023202	台風	2.486280	もしか	2.407946
眠たい	1.885826	ばか	2.023202	拾え	2.407946	しゅん	2.407946
blueriver	1.885826	526	1.969906	便乗	2.407946	未知	2.407946
蚊	1.885826	ayupon	1.969906	DM	2.407946	シワ	2.407946
バス停	1.750937	蚊	1.969906	不審	2.407946	ソワソワ	2.319570
アホ	1.750937	怒っ	1.750937	MEKURU	2.407946	こわ	2.319570
解放	1.750937	スキー	1.750937	乗り換え	2.407946	体調	2.299528
		装着	1.750937	付けよ	2.407946		

Table 7.
A part of TF-ICF calculation without limitation of emotional expression.

In addition, the results shown in **Table 7** were obtained by TF-ICF calculation without limitation of emotional expression.

It is found there are expressive differences of gender as seen from this result. For example, in both of gender, the symbols often be used in emotion: "Joy". Above all, female often use emoji. On the other hand, even though, comparatively, female use genial emotional expressions in emotion: "Anger", male often use radical expressions. It is considered that there are specific emotional expressions for each gender, and those express the gender difference of emotional expression.

Difference of emotional expressions by gender might decrease the estimation accuracy of the learned emotion estimation model due to gender bias. In order to avoid this, it would be useful to prepare an emotion estimation model for each gender or attribute, or to replace the expressions related to attributes with common expressions. In any case, it is clear that some sort of breakthrough is needed to maintain the fairness of machine learning.

5.3 Analysis of emoji

In this subsection, I analyze the appearance tendency of non-verbal expressions such as emoji according to gender. We analyzed usage trend of emoji from the total 59,009 tweets which were collected separately from the emotion corpus for each gender.

The results are shown in **Figures 4** and **5**. In this figure, the horizontal axes shows Emoji type, the vertical axes shows frequency of use. In the graph of male, emojis with over 20 frequency are shown, and in the graph of female, emojis with over 100 frequency are shown. The types of emoji were set 4 classes; expression, emotion, exclamation and other. **Table 8** shows the result of emoji types and frequencies by counting emojis appeared over 10 times. As seen from this result, females had tendancy to use more emojis than male, and female often used emoji expressing expressions or emotions. As was expected that females would use more rich emotion expressions in their tweets, it was obvious from this usage trend of emoji. On the other hand, males used more exclamation marks than other types of emoji.

This result indicates that not only emotional expressions but also nonverbal expressions such as emojis have sufficient influence on emotion estimation. In addition to emojis, Japanese language has emoticons and ASCII art to convey various emotions. Globally, nonverbal expressions play important roles in

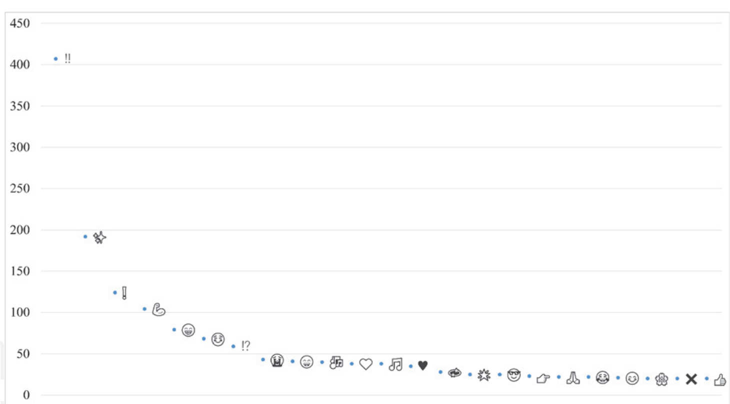

Figure 4.
Trend of emoji by male.

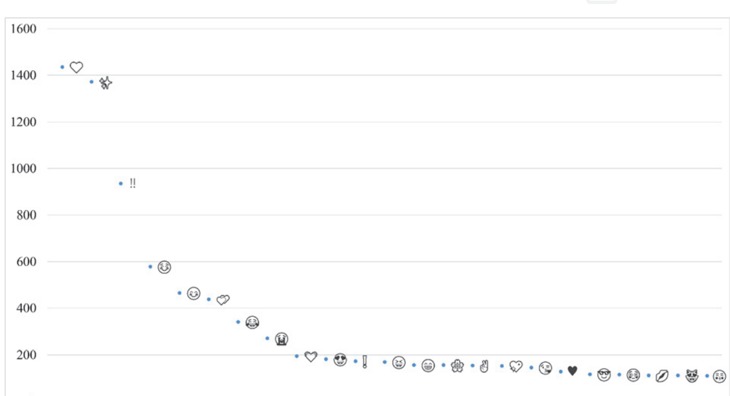

Figure 5.
Trend of emoji by female.

DOI: http://dx.doi.org/10.5772/intechopen.96597

Gender	Type	Freq.	Rate
Male	Other	742	0.402
	Exclamation	600	0.325
	Expression	358	0.194
	Emotion	146	0.079
Female	Expression	3,536	0.32
	Other	3,312	0.3
	Emotion	3,002	0.272
	Exclamation	1,186	0.107

Table 8.
Emoji types and frequencies (over 10 times).

communication on the Web. From this, it is important to understand nonverbal expressions in order to estimate emotions.

6. Emotion estimation from feature of syntactic pattern by deep learning

6.1 Creation of emotion estimation by deep neural networks

We train language patterns that show emotions by using a deep learning method. We use syntactic patterns obtained from the parsing results by the Japanese dependency and case structure analyzer as features for learning. We use KNP [31] as the Japanese dependency and case structure analyzer. KNP is a syntactic, case and reference analyzer developed by Kyoto University. This system uses a noun case frame dictionary constructed by 7 billion web text.

As preprocessing of KNP, it is necessary to annotate morphological features on word unit by using a morphological analyzer. In this study, I make this annotation of morphological features by the morphological analyzer Juman [32]. As seen in **Figure 6**, sentences are analyzed by KNP.

As the result of analysis, the features are annotated on morpheme level and chunk level. The analysis result consists from "Clause layer", "Tag layer", "Morpheme layer". In this study, the features are extracted from the "Tag layer". For training, I use the features that have been annotated on chunk level to associate syntactic patterns with emotions. The examples of features annotated on chunk level are shown in **Table 9**.

The training data are the utterances annotated with emotion tags by manual. These utterances are used in the study by Matsumoto et al. [33], the source sentences are bilingual (Japanese-English). Because these sentences were used as

```
# S-ID:1 KNP:4.20-CF1.1 DATE:2021/03/09 SCORE:-11.50197
* 1D <文頭><連体修飾><用言:形><係:連格><レベル:B-><区切:0-5><ID:（形判連体）><連体節><状態述語><正規化
+ 1D <文頭><連体修飾><用言:形><係:連格><レベル:B-><区切:0-5><ID:（形判連体）><連体節><状態述語><正規化
素晴らしい すばらしい 素晴らしい 形容詞 3 * 0 イ形容詞イ段 19 基本形 2 "代表表記:素晴らしい/すばらしい"
* -1D <文末><時制-過去><句点><体言><用言:判><レベル:C><区切:5-5><ID:（文末）><裸名詞><係:文末><提題受:
+ -1D <文末><時制-過去><句点><体言><用言:判><レベル:C><区切:5-5><ID:（文末）><裸名詞><係:文末><提題受:
景色 けしき 景色 名詞 6 普通名詞 1 * 0 * 0 "代表表記:景色/けしき カテゴリ:抽象物" <代表表記:景色/けしき
でした でした だ 判定詞 4 * 0 判定詞 25 デス列タ形 33 NIL <表現文末><かな漢字><ひらがな><活用語><付属>
。 。 。 特殊 1 句点 1 * 0 * 0 NIL <文末><英記号><記号><付属>
EOS
```

Figure 6.
Analysis results by KNP.

NE:ARTIFACT	NE:MONEY	Wikpedia hypernym: Film of USA	Counter:cm
NE:DATE	NE:ORGANIZATION	Wikipedia hypernym: Area	Counter:dollar
NE:LOCATION	NE:PERCENT	Wikipedia hypernym:Personal Name	Counter:meter
NE:MONEY	NE:PERSON	Wikipedia hypernym:Novel	Counter:person

Table 9.
Example of features.

examples for English composition, it is easy to extract syntactic patterns from sentences. As a preliminary experiment, I confirm emotion estimation accuracy by cross-validation. The breakdown of the five kinds of experimental corpora are shown in **Table 10**.

As the training, I use bi-directional LSTM (bi-LSTM) [34] which is extended LSTM (Long Short-Term Memory) [35]; a kind of recurrent neural networks. LSTM is suited to learning sequences. It enables efficient learning by memorizing and deleting past inputs. **Figure 7** shows the neural network structure using bi-LSTM. I use two LSTM layers.

In this study, I create a feature vector by chunk unit, and input the feature vector from the beginning of a sentence for training. The maximum number of chunks was set as 30 based on the maximum number of the chunks in the corpora.

Name	Cps1	Cps2	Cps3	Cps4	Cps5
# of Sentences	1190	1235	1554	1097	1054
# of Words	12548	8791	14980	20334	11860
# of Words per 1 Sentence	10.5	7.1	9.6	18.5	11.3
# of Vocabulary	2569	1753	2973	3890	2355
# of Clauses	6331	3789	6601	9231	5503
# of Clauses per 1 Sentence	5.3	3.1	4.2	8.4	5.2

Table 10.
Statistic of emotion tagged corpora.

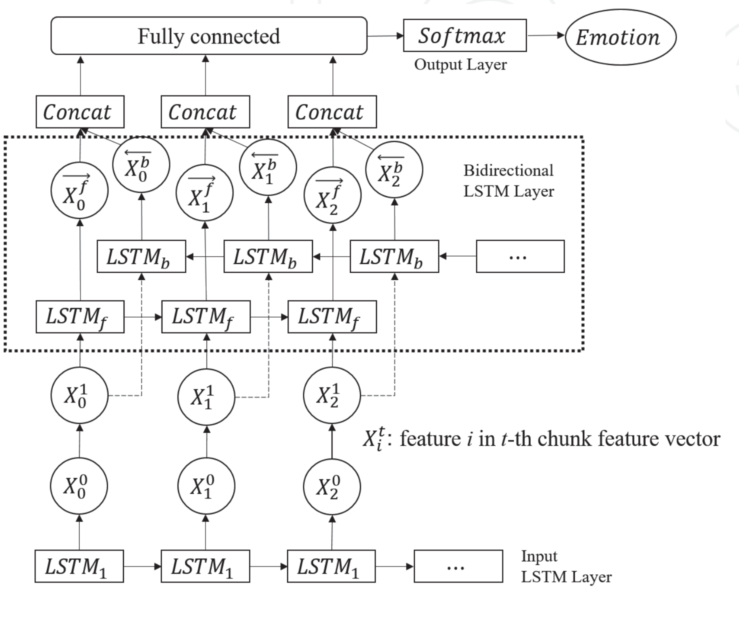

Figure 7.
Neural networks using bidirectional LSTM.

Emotion	Cps1	Cps2	Cps3	Cps4	Cps5
Joy	36.90%	65.50%	27.60%	59.50%	44.90%
Anger	43.10%	47.10%	19.40%	52.10%	38.40%
Sorrow	36.80%	41.20%	25.40%	32.00%	23.90%
Surprise	27.50%	58.50%	13.90%	24.60%	63.20%
Neutral	67.40%	8.70%	73.60%	46.00%	74.80%
Average	42.34%	44.20%	31.98%	42.84%	49.04%

Table 11.
F-measures of the preliminary experimental results.

Table 11 shows the result of the preliminary experiment. Averaged F-measure was 32–49%. The cause of this was thought to be the bias of emotion tags.

6.2 Experiment

I apply the emotion estimator trained syntactic features using bi-LSTM to the tweet sentences for each gender and evaluate the estimator by calculating accuracy. The architecture of the neural networks using bi-LSTM is shown in **Figure 8**. The tweet corpus shown in **Table 12** was used for the experiment.

We compare the result of the proposed method and the emotion estimation result based on emoji. The dictionary registered emojis and their expressing emotions is constructed as the Emoji Emotion Dictionary. The emoji emotion vectors of the emojis that are not in the dictionary are estimated. Emoji emotion vector of each emoji is obtained by calculating similarity with the seed emojis included in the emoji emotion dictionary and by acquiring emotion categories and similarities of top 5 similar seed emojis. The cosine similarity between the emoji distributed representations is used as the similarity of emojis. Eq. (6), (7), (8) shows the calculation of an emoji emotion vector.

$$EV_{e_i} = \left(ew_{e_i}^1 ew_{e_i}^2 \ldots ew_{e_i}^j \ldots ew_{e_i}^n\right) \tag{6}$$

$$\begin{aligned} EV_{avg} &= \frac{1}{|EM_{topN}|} \sum_{em_{e_i} \in EM_{topN}} \left(sim_{e_i} \times EV_{e_i}\right) \\ &= \left(ew_{avg}^1 ew_{avg}^2 \ldots ew_{avg}^j \ldots ew_{avg}^n\right) \end{aligned} \tag{7}$$

$$emotion = \arg\max_x ew_{avg}^x \tag{8}$$

Eq.(6) shows emotion vector EV_{e_i} of emoji em_{e_i}. Emoji emotion vector is a weighted mean of the emotion vectors of the top N similar seed emojis using similarity sim_{e_i} with seed emojis. $ew_{e_i}^j$ shows the weight of emotion category j. Eq.(7) is the formula to calculate the mean emoji emotion vector from the top similar N emoji set EM_{topN}. The estimated emotion is outputted as the emotion category x with the maximum weight value ew_{avg}^j of the mean vector by Eq.(8). The averaged emoji emotion vector is outputted by calculating emojis including in the sentences as the emotion estimation result. In this study, N value is set as 5 to estimate emotions.

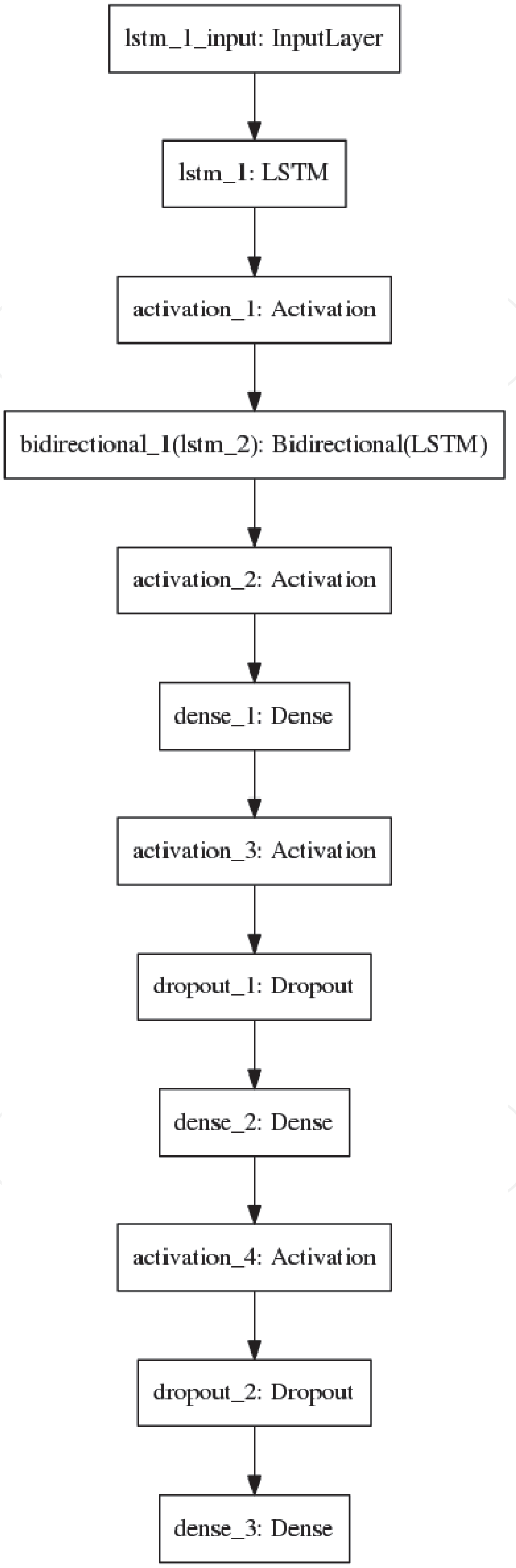

Figure 8.
bi-LSTM neural networks architecture.

	Joy	Anger	Sorrow	Surprise
Number of Sentences	995	418	997	219
Avg. Number of Clauses	8.3	7.5	7.9	7.3
Total Number of Words (Uniq.)	18599 (4766)	7163 (2177)	16523 (4015)	3670 (1337)
Avg. Number of Words	18.7	17.1	16.6	16.8
Total Number of Emojis (Uniq.)	1979 (161)	869 (68)	1581 (93)	379 (46)
Avg. Number of Emojis	1.99	2.08	1.59	1.73

Table 12.
Number of tweet sentences for each emotion.

	Proposed Method	Emoji-based Method
Joy	54.3%	46.3%
Anger	24.3%	4.0%
Sorrow	56.1%	24.6%
Surprise	30.6%	100.0%
Average	41.3%	43.7%

Table 13.
Comparison between the accuracies of the proposed method and the emoji-based method.

6.3 Experimental results

Because neutral tags were not annotated to the target tweet corpus, the accuracies for 4 emotion categories were calculated: "Joy," "Anger," "Sorrow," "Surprise." The experimental result is shown in **Table 13**. The highest accuracy was found for "Sorrow, "and the second highest was for "Joy. "The lowest accuracy was 24.3% and that was obtained for "Anger".

On the other hand, the overall accuracy was 43.7% by the emoji-based method, which was better than by the bi-LSTM based proposed method. However, the accuracy for "Anger" was low; 4% although the accuracy for "Surprise" was 100%. The primal reason is that the varieties of "Surprise" seed emoji were smaller than other kinds of emotions. It is also because that the number of tweets expressing "Surprise" with emoji was relatively scarce.

This result shows that using the syntax pattern enables effective emotion estimation using deep learning even with a small amount of learning data. It is thought that a more accurate model can be realized by flexibly changing dictionary knowledge depending on the domain or the speaker of the target sentence.

7. Conclusions

This chapter introduced our study on "emotion analysis on Japanese language" in the research field of the existing natural language processing and linguistic resources. Most of the existing approaches tried to associate emotions and language patterns, however, if language patterns express different emotions depending on the words consisting of the sentences, the rules for millions of combinations must be described.

It will be effective to analyze emotions based on corpora by annotating emotions on the corpora. In this chapter, various features were annotated on sentences by using a syntactic parser and feature vectors were generated by clause unit. The emotions of the tweet sentences were estimated by training the features using bi-LSTM neural networks.

It was also shown that the capability to development emotions from language patterns by using “emoji” as non-verbal expression. From the experimental results, the emoji-based method was found to be effective to tweet sentences including emoji. Because the amount of the emotion labeled data is limited and the existing dictionary and corpus-based methods cannot cover emotion expressions that are colloquially and depended on users’ attributes, improvement of estimation accuracy is limited. Because emojis are non-verbal emotion expressions that can be used for all users, and the emoji expressions are not depended on the kind of languages, it is a hopeful key of emotion analysis in future.

In addition, syntax pattern might not be correctly extracted from the casual sentences that are often seen in dialogs on SNS. In that case, general-purpose neural language models such as BERT [36] and GPT-3 [37] will be useful. Future developments in language models might eliminate the necessity of human-defined linguistic knowledge such as syntactic patterns, however, methods such as fine tuning are still effective to build emotional estimation models satisfying the needs of all the people from large data. In that case, dictionary knowledge and syntax patterns will play effective roles in improving accuracy and presenting the basis for judgment.

Acknowledgements

This work was partially supported by JSPS KAKENHI Grant Numbers JP20K12027.

Author details

Kazuyuki Matsumoto
Tokushima University, Tokushima, Japan

*Address all correspondence to: matumoto@tokushima-u.ac.jp

References

[1] Kircanski, K, Lieberman, MD, Craske MG. "Feelings into words: contributions of language to exposure therapy" Psychol Sci. 2012 Oct 1;23(10):1086–91. doi: 10.1177/0956797612443830. Epub 2012 Aug 16.

[2] Kuperman, V, Estes, Z, Brysbaert, M, Warriner, AB. "Emotion and language: Valence and arousal affect word recognition", J Exp Psychol Gen. 2014 Jun;143(3):1065–1081. doi: 10.1037/a0035669. Epub 2014 Feb 3.

[3] Fischer, KW, Shaver, P, Carnochan, P. "A skill approach to emotional development: From basic- to subordinate-category emotions. In W. Damon (Ed.), Child development today and tomorrow, pp.107–136, (1989).

[4] Matsumoto, K, Ren, F, Kuroiwa, S. "Emotion Estimation System Based on Emotion Occurrence Sentence Pattern," International Conference on Intelligent Computing (ICIC 2006): Computational Intelligence. pp 902–911, 2006.

[5] Matsumoto, K, Mishina, K, Ren, F, and Kuroiwa, S. "Emotion Estimation Algorithm based on Emotion Occurrence Sentence Pattern," Journal of Natural Language Processing. 2007; **14**(3):239-271, https://doi.org/10.5715/jnlp.14.3_239. (in Japanese)

[6] Žižka, J and Dařena, F. "Automatic Sentiment Analysis Using the Textual Pattern Content Similarity in Natural Language," TSD 2010, LNAI 6231, pp. 224–231, 2010.

[7] Han, Y and Kim, KK. "Sentiment analysis on social media using morphological sentence pattern model," 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA), INSPEC Accession Number: 17009997.

[8] Su, S and Lynn, KT. "Extracting Product Features and Opinion Words Using Pattern Knowledge in Customer Reviews," The ScientificWorld Journal Volume 2013, Article ID 394758, 5 pages http://dx.doi.org/10.1155/2013/394758.

[9] Tokuhisa, M, Murakami, J, Ikehara, S. "Construction and Evaluation of Text-Dialog Corpus with Emotion Tags Focusing on Facial Expression in Comics," Conference: Knowledge-Based Intelligent Information and Engineering Systems, 10th International Conference, KES 2006, Bournemouth, UK, October 9–11, 2006, Proceedings, Part III, DOI10.1007/11893011_91.

[10] Mera, K. and Ichimura, T. "Emotion Analyzing Method Using Physiological State," Proceedings of KES 2004: Knowledge-Based Intelligent Information and Engineering Systems, pp. 195–201.

[11] Matsumoto, K, Yoshida, M, Xiao, Q, Luo, X and Kita, K. "Emotion recognition for sentences with unknown expressions based on semantic similarity by using Bag of Concepts," 12th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2015, pp. 1394–1399, 2015.

[12] WordNet Affect: http://wndomains.fbk.eu/wnaffect.html

[13] Kobayashi, N, Inui, K, Matsumoto, Y, Tateishi, K. "Collecting Evaluative Expressions for Opinion Extraction," Journal of Natural Language Processing 12(3), pp. 203–222, 2005.

[14] Sano, M. (2010) Japanese Dictionary of Appraisal —attitude— (JAppraisal Dictionary ver1.0), published by Gengo Shigen Kyokai.

[15] Torii, Y, Das, D, Bandyopadhyay, S, Okumura, M. "Developing Japanese WordNet Affect for Analyzing

Emotions," Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, ACL-HLT 2011, pages 80–86, 2011.

[16] Nakamura, A. 1993. Emotional Expression Dictionary (in Japanese). Tokyodo shuppan.

[17] Ackerman, A and Puglisi, B. "The Emotion Thesaurus: A Writer's Guide to Character Expression, " Createspace Independent Pub. 2012.

[18] EDR Electronic Dictionary Version 1 Technical Guide, EDR TR2–003, 1995.

[19] Goitaikei – A Japanese Lexicon (in Japanese), Ikehara, S., Miyazaki, M., Shirai, S., Yokoo, A., Nakaiwa, H., Ogura, K., Oyama, Y., and Hayashi, Y. 1997. Iwanami Shoten

[20] Kawahara, D and Kurohashi, S. "Case Frame Compilation from the Web using High-Performance Computing," In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC2006), 2006.

[21] Matsumoto, K, Minato, J, Tsuchiya, S and Ren, F. "An Algorithm for Extracting Emotion Occurrence Sentence Pattern from Corpus," Proceedings of 2008 AIAC Symposium, 109–123, Tokushima, Mar. 2008.

[22] Kudo, T and Matsumoto, Y. "Fast Methods for Kernel-Based Text Analysis," Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL2003), pp. 24-31, 2003.

[23] Tokuhisa, M, Tanaka, T, Ikehara, S and Murakami, J. "Emotion Reasoning Based on Valency Patterns–a prototype annotation of causal relationships—" Human and Artificial Intelligence Systems — From Control to Autonomy, Proceedings of the 4th International Symposium on Human and Artificial Intelligence Systems (HART2004), International Series on Natural and Artificial Intelligence. 2004;**1**:534-539.

[24] Tokuhisa, M, Ikehara, S and Murakami, J. "Construction of Text-dialog Corpus with Emotion Tags Focusing on Facial Expression in Comics," Journal of Natural Language Processing. 2007;**14**(3):193-217. (in Japanese)

[25] Minato, J, Matsumoto K, Ren, F and Kuroiwa, S. "Corpus-based analysis of Japanese-English of emotional expressions." In Proceedings of IEEE International Conference on Natural Language Processing and Knowledge Enginerring, pp. 413–418, 2007.

[26] Minato, J, Matsumoto K, Ren, F, Tsuchiya, S and Kuroiwa, S. "Evaluation of emotion estimation methods based on statistic features of emotion tagged corpus." International Journal of Innovative Computing, Information and Control, Vol. 4, No. 8, pp. 1931–1941, 2008.

[27] Matsumoto, K, Minato, J, Tsuchiya, S and Ren, F. "An Approach for Evaluating Emotion Tagged Corpus," In Proceedings of IADIS International Conference Interfaces and Human Computer Interaction 2008, pp. 258–262, Amsterdam, Jul. 2008.

[28] Mishina, K, Tsuchiya, S, Suzuki, M and Ren, F. "An Improvement of Example-based Emotion Estimation Using Similarity between Sentence and each Corpus," Journal of Natural Language Processing. 2010;**17**(4):91-110.

[29] Quan, C and Ren, F. "Sentence Emotion Analysis and Recognition Based on Emotion Words Using Ren-CECps," International Journal of Advanced Intelligence. 2010;**2**(1):105-117.

[30] Matsumoto, K, Yoshida, M, Kita, K, Wu, Y and Iwasa, F. "Effect of Users Attribute on Emotion Estimation from

Twitter," Proceedings of the 2nd IEEE International Conference on Computer and Communications, Vol.3, pp. 1186-1190, Oct. 2016.

[31] Kawahara, D and Kurohashi, S. "A Fully-Lexicalized Probabilistic Model for Japanese Syntactic and Case Structure Analysis," In Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL2006), pp.176–183, 2006.

[32] Japanese Morphological Analyzer: http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN

[33] Matsumoto, K, Ren, F, Yoshida, M and Kita, K. "Refinement by Filtering Translation Candidates and Similarity Based Approach to Expand Emotion Tagged Corpus," Knowledge Discovery, Knowledge Engineering and Knowledge Management. Communications in Computer and Information Science (CCIS). 2017;**631**:260-280.

[34] Schuster, M and Kuldip KP. "Bidirectional recurrent neural networks." IEEE Transactions on Signal Processing. 1997;**45**(11):2673-2681.

[35] Hochreiter, S and Schmidhuber, J., "Long short-term memory," Neural Computation. 1997;**9**(8):1735-1780

[36] Devlin, Y, Chang, MW, Lee, K and Toutanova, K. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019;**1**:4171-4186.

[37] Brown, T, Mann, B, Ryder, N, et al. "Language Models are Few-Shot Learners." In Proceedings of Advances in Neural Information Processing Systems 33 (NeurIPS 2020), 2020.