Korean Twitter Emotion Classification Using Automatically Built Emotion Lexicons and Fine-Grained Features

Hyo Jin Do and Ho-Jin Choi
School of Computing
Korea Advanced Institute of Science and Technology
Daejeon, Republic of Korea
{hjdo, hojinc}@kaist.ac.kr

Abstract

In recent years many people have begun to express their thoughts and opinions on Twitter. Naturally, Twitter has become an effective source to investigate people's emotions for numerous applications. Classifying only positive and negative tweets has been exploited in depth, whereas analyzing finer emotions is still a difficult task. More elaborate emotion lexicons should be developed to deal with this problem, but existing lexicon sets are mostly in English. Moreover, building such lexicons is known to be extremely labor-intensive or resource-intensive. Finer-grained features need to be taken into account when determining finer-emotions, but many existing works still utilize coarse features that have been widely used in analyzing only the polarity of emotion. In this paper, we present a method to automatically build fine-grained emotion lexicon sets and suggest features that improve the performance of machine learning based emotion classification in Korean Twitter texts.

1 Introduction

Nowadays, people freely express their thoughts on microblogs, and Twitter is known to be one of the popularly used services. In 2014, 500 million tweets were sent per day by 316 million monthly active users across the globe[1]. Not surprisingly, Twitter has been actively mined in the field of computer science to investigate public opinion (Diakopoulos and Shamma, 2010; Kim et al., 2014; O'Connor et al., 2010), get real-time information (Doan et al., 2012), and even forecast future events (Bollen et al., 2011). All such research shows Twitter's potentials in the analysis of human thought and behavior. In particular, researchers are showing interest in the analysis of human emotions presented in Twitter messages. Many studies have been done to classify sentiments (positive and negative) in tweets. Going further, researchers are currently trying to analyze fine-grained emotions beyond polarity. Fine-grained emotion analysis is known to be more challenging than sentiment analysis because it must identify subtle differences between emotions. Dealing with emotions in an individual Twitter post is even more difficult because of its short length with the frequent use of informal words and erroneous sentence structures. Elaborate emotion lexicons should be used to deal with the problem, but non-English speaking countries have difficulties using existing lexicon sets because they are mostly in English. Further, building such lexicons is known to be extremely labor-intensive or resource-intensive that can be a burden to under-resourced countries. Moreover, a set of features that achieves the best performance in fine-grained emotion classification should be exploited that is particularly attuned to tweets written in specific language.

Our goal in this paper is to classify Korean Twitter messages into fine-grained emotions. The emotion types are Ekman's six basic emotions (Ekman, 1992) and it is known to be the most frequently used in the field of computer science for emotion mining and classification (Bann and Bryson, 2012). For this goal, we employed machine learning algorithms

[1] https://about.twitter.com/company

29th Pacific Asia Conference on Language, Information and Computation: Posters, pages 142 - 150
Shanghai, China, October 30 - November 1, 2015

with fine-grained features including an emotion lexicon feature. Specifically, we addressed the following problems:

1. *Emotion lexicon construction.* Is there any simple and automatic method to generate emotion lexicons particularly attuned to the Twitter domain without using other lexical resources?

2. *Feature engineering.* What is the best set of features that can effectively show the subtle distinctions between finer-grained emotions expressed in Korean Twitter texts?

We propose an emotion lexicon construction method and features to address the problems above. Our main contributions are the following:

1. *Emotion lexicon construction.* We propose the weighted tweet frequency (weighted TwF) method, a simple and automatic way to build emotion lexicon lists directly from an annotated corpus without using other resources. The method will be useful for many countries where relevant resources are not available.

2. *Feature engineering.* We propose a set of fine-grained and language-specific features that improves the overall performance of machine learning based emotion classification in Korean Twitter texts.

3. *Resource and Dataset* Our study is unique because emotion analysis on Korean Twitter texts has rarely been addressed before. In addition, we built an annotated dataset, emotion lexicon sets, and other resources. Since finding related datasets and resources in Korean is difficult, we believe our work can contribute to future related studies.

The rest of this paper is organized as follows. Section 2 overviews related work, and in Section 3, we introduce our annotated dataset. In Section 4, we present our emotion lexicon construction method and in Section 5, we describe features designed for classification. We provide experimental results and analysis in Section 6 and conclude in Section 7.

2 Related Work

There have been extensive studies on sentiment analysis that classify expressions of sentiment into positive and negative emotions. In the last few years, researchers have started to explore finer granularity of emotion because simple division of polarity may not suffice in many real-world applications. There are two main approaches to emotion analysis, one is a lexicon based approach and the other is a machine learning based approach. The lexicon based approach utilizes a dictionary of words annotated with their emotional orientation and simply counts the words or aggregates the according values presented in texts. In contrast, the machine learning based approach performs classification using machine learning algorithms based on carefully designed features. Roberts et al. (2012) tried to classify seven emotions in the Twitter domain with binary support vector machine(SVM) classifiers, and Balabantaray et al. (2012) also used SVM classifiers with features including WordNet-Affect emotion lexicons.

A large number of existing emotion lexicon sets were built manually such as WordNet-Affect (Strapparava and Valitutti, 2004) and Linguistic Inquiry and Word Count (Pennebaker et al., 2001). Crowdsourcing is often utilized to obtain a large volume of human annotated lexicon sets such as the NRC Word-Emotion Association Lexicon (Mohammad and Turney, 2013). Non-English speaking countries like Korea have difficulties building emotion lexicons without human labor because existing lexicons and crowd-sourcing platforms are mostly available in English. To deal with the difficulties, one popular approach is to build lexicons upon other resources. For example, AffectNet (Cambria and Hussain, 2012) was constructed using ConceptNet(Liu and Singh, 2004) and WordNet-Affect (Strapparava and Valitutti, 2004). Another popular choice for building lexicon sets automatically is translating existing lexicon lists written in English. Those built by Remus et al. (2010) and Momtazi (2012) are examples. We also propose an automatic method that does not require lexical resources and translation.

Very few attempts have been made so far to analyze emotions in Korean text. Cho and Lee (2006) identified eight emotions in Korean song lyrics with manually annotated word emotion vectors. Lee et al.

(2013) classified Korean tweets into seven emotions and achieved 52% accuracy when using the multinomial naïve Bayes algorithm with morpheme features. The only publicly available Korean emotion lexicons we found were a set of 265 terms of nine emotion types, manually built by Rhee et al. (2008). Our work differs from the aforementioned Korean studies because we automatically construct larger emotion lexicon sets and introduce fine-grained features that are particularly attuned for Korean Twitter texts.

3 Korean Twitter Emotion Analysis (KTEA) Dataset

A Twitter dataset annotated by emotion types is essential in the machine learning based approach for the purpose of training. To build the corpus, we collected random Korean Twitter messages using Twitter streaming API. We removed tweets with RT, URL links, and replies. After the collection process, a corpus can be annotated either manually by human annotators or automatically by distant labels (Go et al., 2009; Wicaksono et al., ; Lee et al., 2013). In our case, we manually annotated the corpus. Each tweet was labeled by three annotators, producing three emotion labels per tweet. Consequently, we constructed a Korean Twitter Emotion Analysis (KTEA) dataset[2], which contains 5,706 valid tweets labeled by seven types of emotions - Ekman's six emotions and no emotion(neutral). Using the dataset, we constructed emotion lexicon sets as described in Section 4 and trained our machine learning algorithm as presented in Section 5.

4 Constructing Emotion Lexicons

4.1 Our Approach - Weighted Tweet Frequency

We built emotion lexicons automatically from the annotated corpus without using other lexical resources. For the construction, we utilized part of our KTEA dataset, which is the set of tweets, each of which was labeled as representing one of Ekman's six emotion types (disregarding the neutral case) by at least one annotator. Table 1 shows the number of tweets we used per emotion for the purpose of lexicon construction.

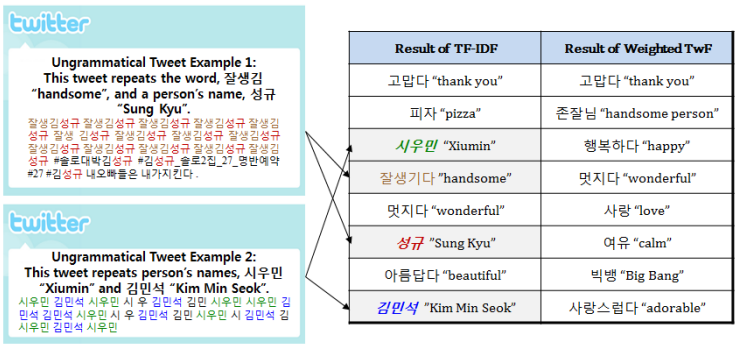
Emotion	Number of Tweets
Happiness	770
Sadness	1377
Anger	903
Disgust	694
Surprise	475
Fear	228
Total	4447

Table 1: The number of tweets we used to generate emotion lexicons using weighted TwF approach

To generate emotion lexicons, we propose the weighted tweet Frequency (weighted TwF) method. First, we aggregated tweets of the same emotion label in one document (d), producing six documents (D) of tweets as a result. Using the six documents, we calculated the weighted TwF for each term (t) that appeared in the documents. The weighted TwF is expressed in Equations 1, 2, and 3. Consequently, we generated six emotion lexicon lists, one list for each emotion type. Each lexicon has a weighted TwF value which shows the strength of the corresponding emotion, i.e., the higher the value is, the stronger the emotion is. The basic idea is similar to the concept of term frequency - inverse document frequency (TF-IDF)[3], for which the occurrences of a term are counted and a penalty is given if the term appears in several documents. However, TF-IDF is not appropriate for our task because the structures of tweets are often highly ungrammatical, and there are many tweets with meaningless terms, which are sometimes excessively repeated in one tweet. In such cases, the meaningless terms produce high term frequency, which results in erroneous emotion lexicons. As illustrated in Figure 1, when term frequency (TF-IDF) is used, we can see some words (that are names in this example), such as 시우민 "Xiumin", 성규 "Sung Kyu", and 김민석 "Kim Min Seok", ranked high in the happiness lexicon list. This is because there are few tweets that excessively repeat those names. Similar kinds of unstructured tweets are frequent in Twitter, and we can disregard such cases by using the tweet frequency defined in Equation 1. It counts the number of tweets instead as true emotion lexicons appear across many tweets, not in a few erroneous tweets.

[2] goo.gl/Gu0GNw

[3] https://en.wikipedia.org/wiki/Tf-idf



Result of TF-IDF	Result of Weighted TwF
고맙다 "thank you"	고맙다 "thank you"
피자 "pizza"	존잘님 "handsome person"
시우민 "Xiumin"	행복하다 "happy"
잘생기다 "handsome"	멋지다 "wonderful"
멋지다 "wonderful"	사랑 "love"
성규 "Sung Kyu"	여유 "calm"
아름답다 "beautiful"	빅뱅 "Big Bang"
김민석 "Kim Min Seok"	사랑스럽다 "adorable"

Figure 1: Example of top-ranked happiness lexicons generated from TF-IDF and weighted TwF

Another reason why TF-IDF is not suitable is the log term in IDF, which is trivial due to the small number of documents. Thus, we used a simple weighting scheme instead as in Equation 2. We set the weight to zero when a lexicon appeared in all the emotion documents in order to remove lexical items that appear very frequently but without any emotions, for example, 있다 "is", and 나 "I".

- d : A document with tweets of same emotion
- D : Total set of ds
- t : Target term
- n : The number of ds where t appears

$$\text{Normalized Tweet Frequency} = \frac{\text{Number of tweets in } d \text{ where } t \text{ appears}}{\text{Total number of terms in the } d} \quad (1)$$

$$\text{Weight} = \begin{cases} \frac{1}{n} & n < |D| \\ 0 & n = |D| \end{cases} \quad (2)$$

$$\text{weighted Tweet Frequency (weighted TwF)} = \text{Normalized Tweet Frequency} \times \text{Weight} \quad (3)$$

Automatic methods of building emotion lexicons have been studied in many works. There are two widely used methods, namely, a thesaurus-based approach (Section 4.2) and a translation-based approach (Section 4.3).

4.2 Thesaurus-Based Approach

The thesaurus-based method builds emotion lexicon lists using synonyms. Using a small set of emotion seed words, this method looks for synonyms using a thesaurus and adds them to the emotion lexicon lists. Due to the lack of a large and representative Korean thesaurus, we combined various publicly available resources, namely, Dong-a's Prime dictionary[4], Naver dictionary[5], a Korean thesaurus[6], and Wise-WordNet[7]. First, seed words – happiness, sadness, anger, disgust, surprise, and fear – were translated into Korean using Dong-a's Prime English-Korean Dictionary. Then, we extended the emotion lexicon sets to include derivatives and synonyms using various resources and thesauruses. Since the resources were not perfect, there were many erroneous synonyms. Thus, for the last step, we manually removed the unreasonable ones. The detailed procedure is summarized in Table 2.

4.3 Translation-Based Approach

There are many lexical resources in English for emotion analysis. This method translates such resources to a specific language, in our case, Korean. Among many lexical resources, we chose WordNet-Affect (Strapparava and Valitutti, 2004) as it is one of the popular and typical emotion lexicon sets used in emotion analysis, and it is freely available. WordNet-Affect contains WordNet synonyms and is manually annotated by Ekman's six emotions. We translated the WordNet-Affect list using Google Translate[8]. We employed the Google service as it is the most widely used translator and its performance is known to be fairly accurate. However, there were

[4]http://www.dongapublishing.com/entry/index.html
[5]dic.naver.com
[6]http://www.wordnet.co.kr/
[7]Software Research Laboratory, ETRI
[8]https://translate.google.co.kr/

Thesaurus-Based Approach
Seed words
happiness, sadness, anger, disgust, surprise, fear (each seed word constructs according emotion lexicon list)
Step 1. Translate seed words to Korean using Dong-a's Prime dictionary
Step 2. Add derivatives using NAVER dictionary
Step 3. Using Korean thesaurus, add synonyms of each word
Step 4. Using WiseWordNet, add primary synonyms of each word
Step 5. Leave only exclusive words for each emotion and remove duplicates within list
Step 6. Manually remove unreasonable or misleading emotion words

Table 2: Procedure of making emotion lexicons using thesaurus-based approach

some erroneous translations since the Korean translator is not perfect. Thus, we manually modified and removed problematic words and duplicates.

4.4 Comparison

In this section, we explain the qualitative aspects of our lexicon construction method in comparison with other approaches. The advantages of our emotion lexicon sets built by weighted TwF approach are the following:

1. As the wordlist is constructed based on real Twitter messages, the method generates Twitter-specific lexicons that include slang, swear words, and ungrammatical words. *Example: 존잘님 "slang for handsome person", 조아 "ungrammatical word for like"*

2. Our method discovers topics that are closely related to some particular emotions. *Example: 야자 "night school study" (sadness - many students feel sad when they are forced to study at night in school)*

3. It is possible to discover keywords that particularly appear in a specific time range. The method automatically updates the lexicons to include newly-coined words, which are essential for emotion analysis in Twitter domain. *Example: 빅뱅 "Big Bang" (happiness - a famous Korean singer Big Bang released a new album at the time we constructed the emotion lexicons)*

We show the effectiveness of our weighted TwF approach by comparing it with the popular thesaurus-based and translation-based approaches. Table 3 compares the three approaches. These approaches can automatically generate emotion lexicons. To be specific, using the thesaurus-based approach, we are able to construct emotion wordlists easily and automatically by using only a small set of seed words. The translation-based approach also translates the existing emotion lexicons automatically using translators. However, the thesaurus-based approach is heavily dependent on lexical resources like dictionaries and thesauruses. A well-built thesaurus is not likely to be available in many non-English speaking countries. Additionally, translation-based approach requires a reliable translator. In comparison, our weighted TwF approach is based on statistics, which are independent of lexical resources and translators; thus, it would be very useful for under-resourced countries. Moreover, we observed that the thesaurus- and translation-based approaches generate a lot of erroneous words due to errors of resources and translators. Hence, manual removal of those words was necessary to achieve accurate results. In contrast, our approach generates lexicons with strength values that show how accurately the word may belong to an emotion type. Even though erroneous words are included in the list, they are likely to be ignored due to the low weighted TwF value. Lastly, our lexicon sets are particularly attuned to the Twitter domain; they include slang, jargon, ungrammatical words, and newly-coined words, whereas most other approaches do not.

Approach	Weighted TwF	Thesaurus	Translation
Automatic?	O	O	O
Resource-free?	O	X(thesaurus)	X(translator)
No manual work?	O	X	X
Twitter-specific?	O	X	X

Table 3: Comparison of our weighted TwF approach with thesaurus-based and translation-based approaches

5 Machine Learning with Fine-Grained Features

Our goal is to classify Korean Twitter messages according to one of the following six emotions, happiness, sadness, anger, disgust, surprise, and fear. We used a machine learning algorithm to classify each Twitter message represented by a feature vector. We

first explain features that we propose in this work and explain our machine learning classification.

5.1 Fine-Grained Features

Feature engineering is very important in machine learning. Features that have been traditionally used in emotion analysis are lexicons and punctuations. Positive and negative emoticons such as :) and :(have also been used in some research. However, more fine-grained and language-specific features are necessary to distinguish finer granularity of emotions. To come up with some effective features, we worked with the following ideas:

- Emoticons and symbols may express specific types of fine-grained emotions
- Some alphabet letters may convey emotions
- Exclamation words may appear in surprise messages
- Swear words may appear in angry messages

We explain how we designed the features according to the ideas we presented above with some examples.

Fine-Grained Emoticons and Symbols Emoticons and symbols are important in analyzing online language because many people express their feelings using them. We constructed a list of emoticons for each fine-grained emotion type that are used in Korea as well as general emoticons widely used in Eastern and Western countries. We constructed a dictionary of emoticons and symbols with the aid of various website articles. Moreover, we included Emojis which have become increasingly popular on Twitter since mobile devices adopted them. We sorted each emoticon and symbol into one of the six emotions using the explanations written in the websites. One interesting aspect of the dictionary is that it utilizes regular expressions to incorporate various mutations of emoticons. For example, Korean emoticons often use various or extended particles to represent the mouth of a face. In the case of a smiling face (^^), people use various mutation of such emoticons such as (^_^),(^____^),(^.^),(^ㅅ^),(^3^). In other words, similar to languages, emoticons also have informal versions of similar patterns. Thus, we incorporated such common cases with regular expressions. Part of the dictionary of emoticons and symbols is shown in Figure 2.

Emotion	Happiness	Sadness	Anger	Disgust	Surprise	Fear
Examples	=) \^([^\s]*)\^	ㅠㅠ T([^\s]*)T	>:(o`Д′o	−^− 0π0	\ (◎o◎) / ！ O_o	~_~; -_-;

Figure 2: Example of fine-grained emoticon-emotion dictionary

Korean Letters	Meaning
ㅋ , ㅎ	Laughing with Happiness
ㅠ , ㅜ	Crying with Sadness
ㄷㄷ	Shaking with Fear
ㅂㄷㅂㄷ	Trembling with Anger

Table 4: List of Korean letters closely related to emotions

Korean Emotion Letters Language-specific feature are important in performing emotion analysis for a specific language. Koreans use certain Korean letters to show emotions, so we took certain letters into account that are listed in Table 4. 'ㅋ' and 'ㅎ' are often used to indicate laughter, while 'ㅠ' and 'ㅜ' indicate crying. Also, sequences of letters, such as 'ㄷㄷ' and 'ㅂㄷㅂㄷ', are often used to express fear and anger, respectively. We counted and normalized the number of such emotion letters and added them as a language-specific feature.

Exclamations of Surprise According to Merriam-Webster dictionary, the definition of an exclamation is "a sharp or sudden cry, a word, phrase, or sound that expresses a strong emotion[9]". We assumed that exclamations are often used in tweets expressing surprise, such as 맙소사 "oh no", 앗 "oh dear", and 우와 "wow". We searched various websites and collected examples to make a list of exclamations of surprise. As a result, we constructed a list of 45 surprise exclamation words. We then counted the number of occurrences of such words in tweets and added them as a feature.

Swear Words We observed that swear words are often used in angry tweets; therefore, we assumed that there occurrence is a strong clue to identify tweets expressing anger. We constructed our own list of Korean swear words by combining numerous related

[9]http://www.merriam-webster.com/dictionary/exclamation

resources and websites. As a result, a list of 227 Korean swear words was built. For each tweet, we counted the number of occurrences of swear words and added them as a feature.

Consequently, we designed a feature vector based on the conventional features as well as the features we presented above. To sum up, we considered the following features for classification:

1. Emotion lexicons (weighted TwF)
2. Emotion lexicons (thesaurus+translation)
3. Punctuation (? ! !? . , ∼)
4. Fine-grained emoticons and symbols
5. Korean emotion letters
6. Exclamations of surprise
7. Swear words

5.2 Machine Learning Based Classification

Before constructing a machine learning classifier, we applied the synthetic minority oversampling technique (SMOTE) (Chawla et al., 2002) to the training set, a well-known oversampling method, which is known to be more effective than the plain oversampling method with replication. We preferred an oversampling method to an undersampling method since our dataset is highly imbalanced, and undersampling removes too many instances. SMOTE generates synthetic instances of the minority class by choosing a random point for each line segment between randomly selected neighbors from k nearest minority neighbors. As a result of applying SMOTE to our training set, we could make a balanced dataset, which is favored for most machine learning algorithms. We compared several machine learning algorithms for classification, including support vector machine (SVM), multinomial logistic regression, random forest, J48, naive Bayes, and zeroR.

6 Experimental Results and Analysis

We performed experiments using WEKA[10] to evaluate 1) our weighted tweet frequency method and 2) the performance of machine learning based classification using the feature vector we engineered.

Dataset For training and testing the machine learning algorithms, we used 899 Twitter messages from our KTEA dataset, which contains tweets for which three annotators all agreed on the emotion type, excluding neutral. We performed 5-cross validation.

Performance Measure We used precision, recall and F-measure to evaluate the classification performance for each emotion type. Also, the weighted average of each measure was computed to determine the overall performance of unbalanced test dataset.

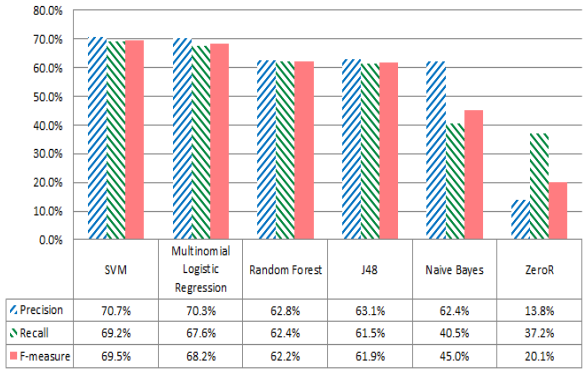
Weighted Tweet Frequency We investigated the performance of our lexicon building method, weighted tweet Frequency, and compared it with the performance of the thesaurus- and translation-based methods. We found that the lexicons based on the thesaurus- and translation-based approaches suffer from low coverage due to the lack of reliable words produced by the Korean resources and translator. Therefore, we combined the lexicon lists produced by the thesaurus- and translation-based approaches to make a larger emotion lexicon list. In other words, we compared our approach (weighted TwF) against the combined approach (thesaurus+translation). The precision, recall, and F-measure of using SVM is shown in Figure 3. The F-measure of our approach is higher than that of the thesaurus+translation approach. The precision of the thesaurus+translation approach is relatively high due to the manual removal of erroneous words from the lists. However, its recall is very low because it does not contain Twitter-specific words. Furthermore, our approach, used together with the thesaurus+translation approach, achieves the best performance.

Machine Learning Based Classification First, we investigated the most appropriate machine learning algorithm for classification. We tested various machine learning algorithms: SVM, multinomial logistic regression, random forest, J48, naive Bayes, and zeroR. Figure 4 shows the results. SVM produced the best precision, recall, and F-measure compared to the others.

We conducted another experiment to evaluate how well the features we proposed improved the performance of SVM. As shown in Figure 5, the best performance was observed when all the features were combined and the overall F-measure was about 70%. Emotion lexicon and punctuation features achieved an F-measure of about 64%. Adding the exclamation of surprise feature improved the classification of the surprise emotion by a 12% F-measure. Further adding Korean emotion letters

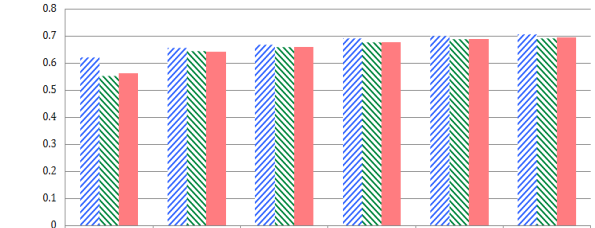
[10] http://www.cs.waikato.ac.nz/ml/weka/

Figure 3: Precision, recall, and F-measure of using our weighted TwF, thesaurus+translation, and the two approaches combined. The best accuracy was observed when all the approaches were used.



	SVM	Multinomial Logistic Regression	Random Forest	J48	Naive Bayes	ZeroR
Precision	70.7%	70.3%	62.8%	63.1%	62.4%	13.8%
Recall	69.2%	67.6%	62.4%	61.5%	40.5%	37.2%
F-measure	69.5%	68.2%	62.2%	61.9%	45.0%	20.1%

Figure 4: Precision, recall, and F-measure achieved by using different machine learning algorithms. SVM generated the best performance.



	A	B	C	D	E	F
Precision	0.621	0.656	0.669	0.69	0.701	0.707
Recall	0.553	0.644	0.659	0.675	0.688	0.692
F-measure	0.563	0.644	0.66	0.677	0.691	0.695

Feature Sets

A. Emotion Lexicon (Weighted TwF, Thesaurus+Translation)
B. A + Punctuation
C. B + Exclamation of Surprise
D. C + Alphabet
E. D + Emoticon and Symbol
F. E + Swear Words

Figure 5: Precision, recall, and F-measure using combined set of features. Using all features achieved the best performance which is about 70% F-measure

helped to classify sadness, increasing the classification score from 67% to 72.1%, while the value for fear increased from 76.3% to 79.5%. Moreover, combining emoticon and symbol feature particularly improved the classification for happiness increased from 73% to 76.3%. Lastly, we added the swear word feature. As expected, it increased the classification of anger from 54.4% to 56.5%. Overall, we found that our fine-grained features helped the analysis of fine-grained emotions, and we believe that improving the feature resources will further improve the overall performance.

7 Conclusion

We proposed a machine learning based classification method that sorts Korean Twitter messages into six emotion types using carefully designed features. Emotion analysis research in under-resourced countries can benefit from our emotion lexicon building method as we automatically construct lexicons without any help from other resources and tools. In addition, we suggested several fine-grained features to improve classification performance. We believe that our research, the KTEA dataset, and resources represent a significant step forward in Korean Twitter emotion analysis studies, which have been rarely addressed before.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (No. 2010-0028631).

References

R. C. Balabantaray, Mudasir Mohammad, and Nibha Sharma. 2012. Multi-class twitter emotion classification: A new approach. *International Journal of Applied Information Systems*, 4(1):48–53.

E. Y. Bann and J. J. Bryson. 2012. The conceptualisation of emotion qualia: Semantic clustering of emotional tweets. In *Proceedings of the 13th Neural Computation and Psychology Workshop*. World Scientific.

Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8.

Erik Cambria and Amir Hussain. 2012. *Sentic Computing: Techniques, Tools, and Applications*, volume 2. Springer Science & Business Media.

Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. Smote: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357.

Young Hwan Cho and Kong Joo Lee. 2006. Automatic affect recognition using natural language processing techniques and manually built affect lexicon. *IEICE Transactions on Information and Systems*, 89(12):2964–2971.

Nicholas A. Diakopoulos and David A. Shamma. 2010. Characterizing debate performance via aggregated twitter sentiment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1195–1198. ACM.

Son Doan, Bao-Khanh Ho Vo, and Nigel Collier. 2012. An analysis of twitter messages in the 2011 tohoku earthquake. In *Electronic Healthcare*, volume 91, pages 58–66. Springer.

Paul Ekman. 1992. An argument for basic emotions. *Cognition & Emotion*, 6(3-4):169–200.

Alec Go, Richa Bhayani, and Lei Huang. 2009. Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, 1:12.

Jeongin Kim, Dongjin Choi, Myunggwon Hwang, and Pankoo Kim. 2014. Analysis on smartphone related twitter reviews by using opinion mining techniques. In *Advanced Approaches to Intelligent Information and Database Systems*, pages 205–212. Springer.

Cheolseong Lee, Donghee Choi, Seongsoon Kim, and Jaewoo Kang. 2013. Classification and analysis of emotion in korean microblog texts. *Journal of Korean Institute of Information Scientists and Engineers: Databases*, 40(3):159–167.

Hugo Liu and Push Singh. 2004. Conceptnet — a practical commonsense reasoning tool-kit. *Journal of BT Technology*, 22(4):211–226.

Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3):436–465.

Saeedeh Momtazi. 2012. Fine-grained german sentiment analysis on social media. In *Proceedings of the 8th International Conference on Language Resources and Evaluation*. European Language Resources Association (ELRA).

Brendan O'Connor, Ramnath Balasubramanyan, Bryan R Routledge, and Noah A Smith. 2010. From tweets to polls: Linking text sentiment to public opinion time series. *Proceedings of the International Conference on Web and Social Media*, 11(122-129):1–2.

James W. Pennebaker, Martha E. Francis, and Roger J. Booth. 2001. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71:2001.

Robert Remus, Uwe Quasthoff, and Gerhard Heyer. 2010. Sentiws - a publicly available german-language resource for sentiment analysis. In *Proceedings of the 7th International Conference on Language Resources and Evaluation*. European Language Resources Association (ELRA).

June Woong Rhee, Hyun Joo Song, Eun Kyung Na, and Hyun Suk Kim. 2008. Classification of emotion terms in korean. *Korean Journal of Journalism and Communication Studies*, 52(1):85–116.

Kirk Roberts, Michael A. Roach, Joseph Johnson, Josh Guthrie, and Sanda M. Harabagiu. 2012. Empatweet: Annotating and detecting emotions on twitter. In *Proceedings of the 8th International Conference on Language Resources and Evaluation*. European Language Resources Association (ELRA).

Carlo Strapparava and Alessandro Valitutti. 2004. Wordnet affect: an affective extension of wordnet. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, volume 4, pages 1083–1086.

Alfan F. Wicaksono, Clara Vania, Distiawan T. Bayu, and Mirna Adriani. Automatically building a corpus for sentiment analysis on indonesian tweets. *28th Pacific Asia Conference on Language, Information and Computing*.