

Sentiment Analysis on Korean Blog Comments using Korean Emotional Word Dictionary

Junwon Lee
Machine Learning
Fall 2019

Introduction

Sentiment analysis is a process of categorizing texts based on the writers' emotions or subjective opinions. Common sentiment analysis focuses on the positivity and negativity of the text.

South Korea has been undergoing through rapid modernization, and this has brought several contemporary social issues to public discussions, such as gender equality and high suicide rates. This uprise catalyzed the pollution of the reply culture in social media, where flammers (known as Akpeulleo in Korean) publicly express hate toward particular individual or groups who are even slightly involved with these issues through blog comments with anonymous ID's.

In 2011, the Korea Advanced Institute of Science technology (KAIST) published the Korean Emotional Word Dictionary (KEWD), which consists of 8698 words with clear sentiment from various sources.

The Goal

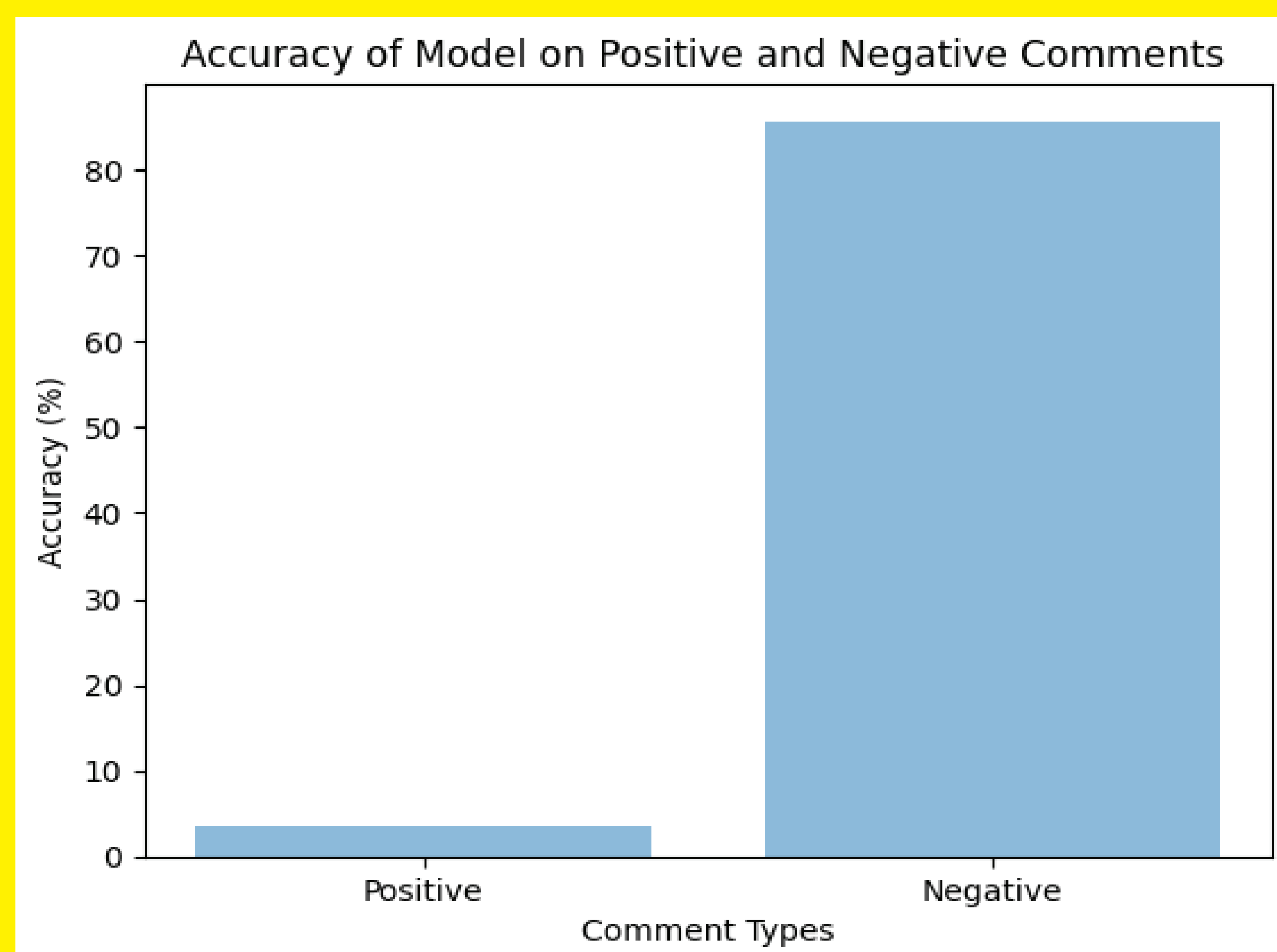
Construct a Naive Bayes model based on the Korean Emotional Word Dictionary, and use the model to predict the sentiment of Korean blog comments.

Experiment

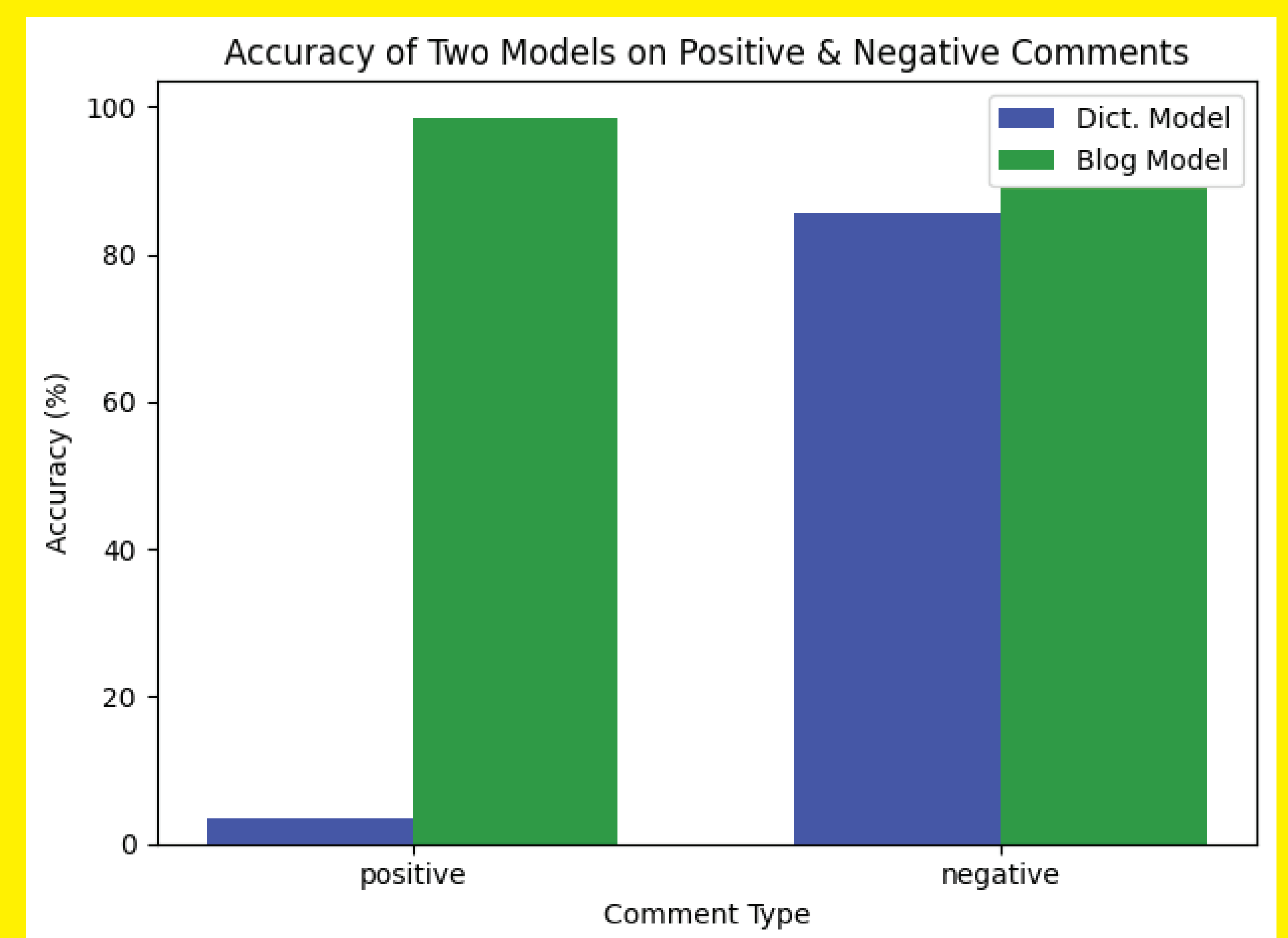
Bag of words is constructed from the KEWD. To do this, each word in the Dictionary is converted into vectors of features, where the vector consists of counts of difference features. Using this vector Naive Bayes model was fitted using this bag of words.

Then, with the same model, I measured the accuracy of the model by predicting the sentiment of 2000 blog comments from the Korean Radical Anti-male Website. I compared the performance of the model with the Naive Bayes model created from the words from the blog comments.

Evaluation and Results



The model was highly successful (~80% accuracy) in predicting the negative comments, but not successful (~4%) in predicting the positive comments.



Compared to the model based on the comments, the model based on the dictionary performed about the same as identifying the true negative comments, while it is not even close to successful in identifying the positive comments.

Which Words are Important in Each Model?

Model Based on the KEWD

Positive	Negative
만족 (Satisfaction)	사악 (Evil)
견고 (Solidity)	식상 (Fed up with)
기대 (Expectation)	곤란 (Embarrassing)
충실 (Devotion)	바보 (Idiot)
훌륭 (Excellence)	부족 (Unsatisfaction)

Model Based on the Blog Comments

Positive	Negative
있어요 (possess in honorific)	이기 (selfishness)
나는 (I)	멋지노 (fancy in casual language)
싶어 (I want to)	ㅋㅋㅋ (Korean slang for lol)
그는 (He in honorific)	ㅇㄱㄹㅇ (Korean slang for "this is real")
어때 (What do you think)	붏나 (Korean slang for really)

Analysis

After understanding both models' performances, I found out that each model was focusing on different features of the comments to evaluate the sentiment. While the model based on the Dictionary emphasized on the adjectives and nouns in the sentences, the model based on the blog comments focused on the first-person pronouns, honorifics or polite language (known as jondaemal), and slangs which are prominent in blog comments but barely covered in the Dictionary. This could possibly explain why the model based on the comments performed better.

Conclusion

The model constructed from Korean Emotional Word Dictionary was effective in categorizing negative sentences primarily because it identified negative nouns and adjectives in the comments. However, it was unsuccessful in correctly categorizing positive comments. The primary cause was the inappropriate metric that the model used on the blog comments. In order to create a more accurate model to classify the comments, future corpus should contain variations of Korean language beyond the adjectives and nouns, such as honorifics and slangs.