

Feature	Definition
personalcount	Number of first person pronouns in tweet.
is_notenglish	Number of words contained in a list of most common words in tweets rated as not english.
tagnoun	Number of nouns in tweet divided by the total number of words.
sis_noise	Ratio of a measure of similarity between tweet and noise corpus by its uncertainty.
sis_signal	Ratio of a measure of similarity between tweet and targeted community corpus by its uncertainty.
bigrams_noise	Weighted number of bigrams in tweet contained in corpus of noise tweets.
is_english	Ratio of number of words in English by number of words in foreign languages.
common_noise	Weighted number of words in tweet similar to common words in noise corpus.
common_signal	Weighted number of words in tweet similar to common words in targeted community corpus.
wordscount	Number of words in tweet.
ncharacters	Number of characters in tweet.