Sentiment Polarity Classification of Figurative Language: Exploring the Role of Irony-Aware and Multifaceted Affect Features

Supplementary Material

Features in the Irony Detection Model

1. Structural features

Length Words - It refers to the amount of words in each tweet
Length Chars - It refers to the amount of characters in each tweet
Colon - It refers to the amount or ":" in each tweet
Exclamation - It refers to the amount or "!" in each tweet
Question - It refers to the amount or "?" in each tweet
PM - It refers to the sum of Colon, Exclamation, and Question
Verbs - It refers to the amount of words in each tweet labeled as verbs by the NLTK Part-of-speech tagger
Nouns - It refers to the amount of words in each tweet labeled as nouns by the NLTK Part-of-speech tagger
Adjectives - It refers to the amount of words in each tweet labeled as adjectives by the NLTK Part-of-speech tagger
Adverbs - It refers to the amount of words in each tweet labeled as adverbs by the NLTK Part-of-speech tagger
Uppercasechars - It refers to the amount of uppercase characters in each tweet
Total emoticons - It refers to the amount of emoticons in each tweet
Hashtags frequency - It refers to the amount of hashtags in each tweet
Mentions frequency - It refers to the amount of hashtags in each tweet
Retweet - It refers to the "RT" signal used to denote when a tweet was previously posted by other user.
counterValue - It refers to the frequency of counter-factuality terms
temporalValue - It refers to the frequency of temporal compression terms
similarity – It refers the degree of inconsistency in a tweet the Wu&Palmer semantic similarity measure. We used the WordNet::Similarity module available in: https://codegoogle.com/p/ws4j/

2. Affective features

a) Sentiment

AFINN
- afinnPos - It refers to the sum of the polarity values of words with positive score in Afinn
- afinnNeg - It refers to the sum of the polarity values of words with negative score in Afinn
- afinnValue - It refers to the difference between the afinnPos and afinnNeg

Hu&Liu (hl)
- hlPos - It refers to the amount of words in each tweet defined as positive in Hu&Liu
- hlNeg - It refers to the amount of words in each tweet defined as negative in Hu&Liu
- hlValue - It refers to the difference between the hlPos and hlNeg

General Inquirer (gi)
- giPos - It refers to the amount of words in each tweet defined as positive in General Inquirer
- giNeg - It refers to the amount of words in each tweet defined as negative in General Inquirer
- giValue - It refers to the difference between the giPos and giNeg

EmoLex
- emolexPos - It refers to the amount of words in each tweet defined as positive in EmoLex
- emolexNeg - It refers to the amount of words in each tweet defined as negative in EmoLex

SentiWordNet (swn)
- swnPos - It refers to the sum of the positive score for the words in each tweet
- swnNeg - It refers to the sum of the negative score for the words in each tweet
- swnValue - It refers to the difference betweet the swnPos and swnNeg

SenticNet (sn)
- snPolarityDimension - It refers to the sum of the polarity values for the words in each tweet that are defined in SenticNet
- snPolarityFormula - It refers to result of appliying the formula where the four dimensions in SenticNet are combined to determine a polarity value for each tweet.

Semantic Orientation (so)
- soValue - It refers to the sum of each word in the tweet according to the Semantic Orientation lexicon

EffectWordNet (ewn)
- ewnPositive - It refers to the amount of words in each tweet defined as positive in EffectWordNet

ewnNegative - It refers to the amount of words in each tweet defined as negative in EffectWordNet
ewnNull - It refers to the amount of words in each tweet defined as null in EffectWordNet

Subjectivity (subj)

subjPosStrong - It refers to the amount of words in each tweet defined as positive strong terms according to Subjectivity lexicon
subjPosWeak - It refers to the amount of words in each tweet defined as positive weak terms according to Subjectivity lexicon
subjNegStrong - It refers to the amount of words in each tweet defined as negative strong terms according to Subjectivity lexicon
subjNegWeak - It refers to the amount of words in each tweet defined as negative weak terms according to Subjectivity lexicon

b) Emotions

Dictionary of Affect in Language (dal)

dalPleasentness - It refers to the sum of the pleasentness values for the words in each tweet that are defined in the Dictonary of Affect in Language
dalActitude - It refers to the sum of the actitude values for the words in each tweet that are defined in the Dictionary of Affect in Language
dalImagery - It refers to the sum of the imagery values for the words in each tweet that are defined in the Dictionary of Affect in Language

SenticNet (sn)

snPleasantness - It refers to the sum of the pleasantness values for the words in each tweet that are defined in SenticNet
snAttention - It refers to the sum of the attention values for the words in each tweet that are defined in SenticNet
snSensitivity - It refers to the sum of the sensitivity values for the words in each tweet that are defined in SenticNet
snAptitude - It refers to the sum of the aptitude values for the words in each tweet that are defined in SenticNet

ANEW

anewValence - It refers to the sum of the valence values for the words in each tweet that are defined in ANEW
anewArousal - It refers to the sum of the arousal values for the words in each tweet that are defined in ANEW
anewDominance - It refers to the sum of the dominance values for the words in each tweet that are defined in ANEW

LIWC

liwcPos - It refers to the amount of words associated with positive emotions in LIWC
liwcNeg - It refers to the amount of words associated with negative emotions in LIWC
liwcValue - It refers to the difference between liwcPos and liwcNeg

EmoLex

emolexAnger - It refers to the amount of words associated with Anger
emolexAnticipation - It refers to the amount of words associated with Anticipation
emolexDisgust - It refers to the amount of words associated with Disgust
emolexFear - It refers to the amount of words associated with Fear
emolexJoy - It refers to the amount of words associated with Joy
emolexSadness - It refers to the amount of words associated with Sadness
emolexSurprise - It refers to the amount of words associated with Surprise
emolexTrust - It refers to the amount of words associated with Trust

EmoSenticNet (emoSN)

emoSNAnger - It refers to the amount of words associated with Anger
emoSNDisgust - It refers to the amount of words associated with Disgust
emoSNJoy - It refers to the amount of words associated with Joy
emoSNSadness - It refers to the amount of words associated with Sadness
emoSNSurprise - It refers to the amount of words associated with Surprise
emoSNFear - It refers to the amount of words associated with Fear

SentiSense

sentisenseAnger - It refers to the amount of words associated with Anger
sentisenseDisgust - It refers to the amount of words associated with Disgust
sentisenseJoy - It refers to the amount of words associated with Joy
sentisenseSadness - It refers to the amount of words associated with Sadness
sentisenseSurprise - It refers to the amount of words associated with Surprise
sentisenseFear - It refers to the amount of words associated with Fear
sentisenseAnticipation - It refers to the amount of words associated with Anticipation
sentisenseLike - It refers to the amount of words associated with Like
sentisenseLove - It refers to the amount of words associated with Love

Features in the Sentiment Analysis Model

1. Structural features

Length Words - It refers to the amount of words in each tweet
Length Chars - It refers to the amount of characters in each tweet
Colon - It refers to the amount or ":" in each tweet
Exclamation - It refers to the amount or "!" in each tweet
Question - It refers to the amount or "?" in each tweet
Comma - It refers to the amount or "," in each tweet
Semicolon - It refers to the amount or ";" in each tweet
PM - It refers to the sum of Colon, Exclamation, Question, Comma, and Semicolon in each tweet.
Verbs - It refers to the amount of words in each tweet labeled as verbs by the NLTK Part-of-speech tagger
Nouns - It refers to the amount of words in each tweet labeled as nouns by the NLTK Part-of-speech tagger
Adjectives - It refers to the amount of words in each tweet labeled as adjectives by the NLTK Part-of-speech tagger
Adverbs - It refers to the amount of words in each tweet labeled as adverbs by the NLTK Part-of-speech tagger
Uppercasechars - It refers to the amount of uppercase characters in each tweet
URL - It refers to the presence of an URL in each tweet
Positive emoticons - It refers to the amount of positive emoticons in each tweet
Negative emoticons - It refers to the amount of negative emoticons in each tweet

2. Twitter marks

Hashtags binary - It refers to the presence of hashtags in each tweet
Hashtags frequency - It refers to the amount of hashtags in each tweet
Mentions binary - It refers to the presence of mentions in each tweet
Mentions frequency - It refers to the amount of hashtags in each tweet
Retweet - It refers to the "RT" signal used to denote when a tweet was previously posted by other user.

3. Sentiment modifiers

Elongated words - It refers to the amount of words in each tweet where the same letter appears more than three times
Interjections - It refers to the amount of words in each tweet usually identified as interjections
Negations - It refers to the amount of words in each tweet usually identified as negations

4. Sentiment Analysis

a) AFINN

afinnPos - It refers to the sum of the polarity values of words with positive score in Afinn
afinnNeg - It refers to the sum of the polarity values of words with negative score in Afinn
afinnValue - It refers to the difference between the afinnPos and afinnNeg
afinnPos_normalized - It means the afinnPos dividing by the amount of words in each tweet
afinnNeg_normalized - It means the afinnNeg dividing by the amount of words in each tweet
afinnValue_normalized - It means the afinnVaue dividing by the amount of words in each tweet

b) Hu&Liu (hl)

hlPos - It refers to the amount of words in each tweet defined as positive in Hu&Liu
hlNeg - It refers to the amount of words in each tweet defined as negative in Hu&Liu
hlValue - It refers to the difference between the hlPos and hlNeg
hlPos_normalized - It means the hlPos dividing by the amount of words in each tweet
hlNeg_normalized - It means the hlNeg dividing by the amount of words in each tweet
hlValue_normalized - It means the hlValue dividing by the amount of words in each tweet

c) General Inquirer (gi)

giPos - It refers to the amount of words in each tweet defined as positive in General Inquirer
giNeg - It refers to the amount of words in each tweet defined as negative in General Inquirer
giValue - It refers to the difference between the giPos and giNeg

giPos_normalized - It means the giPos dividing by the amount of words in each tweet
giNeg_normalized - It means the giNeg dividing by the amount of words in each tweet
giValue_normalized - It means the giValue dividing by the amount of words in each tweet

d) SentiWordNet (swn)

swnPos - It refers to the sum of the positive score for the words in each tweet
swnNeg - It refers to the sum of the negative score for the words in each tweet
swnObj - It refers to the sum of the objective score for the words in each tweet
swnValue - It refers to the difference betweet the swnPos and swnNeg
swnPos_normalized - It means the swnPos dividing by the amount of words in each tweet
swnNeg_normalized - It means the swnNeg dividing by the amount of words in each tweet
swnObj_normalized - It means the swnObj dividing by the amount of words in each tweet
swnValue_normalized - It means the swnValue dividing by the amount of words in each tweet

e) EmoLex

emolexPos - It refers to the amount of words in each tweet defined as positive in EmoLex
emolexNeg - It refers to the amount of words in each tweet defined as negative in EmoLex
emolexValue - It refers to the difference between the emolexPos and emolexNeg
emolexPos_normalized - It means the emolexPos dividing by the amount of words in each tweet
emolexNeg_normalized - It means the emolexNeg dividing by the amount of words in each tweet
emolexValue_normalized - It means the emolexValue dividing by the amount of words in each tweet

f) SenticNet(sn)

snPolarityDimension - It refers to the sum of the polarity values for the words in each tweet that are defined in SenticNet
snPolarityFormula - It refers to result of applying the formula where the four dimensions in SenticNet are combined to determine a polarity value for each tweet.
snPolarityDimension_normalized - It means the snPolarityDimension dividing by the amount of words in each tweet
snPolarityFormula_normalized - It means the snPolarityFormula dividing by the amount of words in each tweet

g) Sentiment 140

sentiment140Value - It refers to the sum of the polarity values for the words in each tweet that are defined in Sentiment140
sentiment140Value_normalized - It means the sentiment140Value dividing by the amount of words in each tweet

h) NRC Hashtag Sentiment Lexicon

nrcHashtagValue - It refers to the sum of the polarity values for the words in each tweet that are defined in NRC Hashtag Sentiment Lexicon
nrcHashtagValue_normalized - It means the nrcHashtagValue dividing by the amount of words in each tweet

i) MPQA

mpqaPos - It refers to the amount of words in each tweet defined as positive in MPQA
mpqaNeg - It refers to the amount of words in each tweet defined as negative in MPQA
mpqaValue - It refers to the difference between the mpqaPos and mpqaNeg
mpqaPos_normalized - It means the mpqaPos dividing by the amount of words in each tweet
mpqaNeg_normalized - It means the mpqaNeg dividing by the amount of words in each tweet
mpqaValue_normalized - It means the mpqaValue dividing by the amount of words in each tweet

j) Pattern

patternValue - It refers to the "sentiment" function in Pattern. It provides a function to calculate a polarity value between -1.0 and +1.0 in a sentence.
patternValue_normalized - It means the patternValue dividing by the amount of words in each tweet

5. Categorical model of Emotions

a) EmoLex

emolexAnger - It refers to the amount of words associated with Anger
emolexAnticipation - It refers to the amount of words associated with Anticipation
emolexDisgust - It refers to the amount of words associated with Disgust
emolexFear - It refers to the amount of words associated with Fear
emolexJoy - It refers to the amount of words associated with Joy
emolexSadness - It refers to the amount of words associated with Sadness
emolexSurprise - It refers to the amount of words associated with Surprise
emolexTrust - It refers to the amount of words associated with Trust
emolexAnger_normalized - It means the emolexAnger dividing by the amount of words in each tweet
emolexAnticipation_normalized - It means the emolexAnticipation dividing by the amount of words in each tweet
emolexDisgust_normalized - It means the emolexDisgust dividing by the amount of words in each tweet
emolexFear_normalized - It means the emolexFear dividing by the amount of words in each tweet
emolexJoy_normalized - It means the emolexJoy dividing by the amount of words in each tweet

emolexSadness_normalized - It means the emolexSadness dividing by the amount of words in each tweet
emolexSurprise_normalized - It means the emolexSurprise dividing by the amount of words in each tweet
emolexTrust_normalized - It means the emolexTrust dividing by the amount of words in each tweet

b) LIWC

liwcPos - It refers to the amount of words associated with positive emotions in LIWC
liwcNeg - It refers to the amount of words associated with negative emotions in LIWC
liwcPos_normalized - It means the liwcPos dividing by the amount of words in each tweet
liwcNeg_normalized - It means the liwcNeg dividing by the amount of words in each tweet

c) EmoSenticNet (emoSN)

emoSNAnger - It refers to the amount of words associated with Anger
emoSNDisgust - It refers to the amount of words associated with Disgust
emoSNJoy - It refers to the amount of words associated with Joy
emoSNSadness - It refers to the amount of words associated with Sadness
emoSNSurprise - It refers to the amount of words associated with Surprise
emoSNFear - It refers to the amount of words associated with Fear
emoSNAnger_normalized - It means the emoSNAnger dividing by the amount of words in each tweet
emoSNDisgust_normalized - It means the emoSNDisgust dividing by the amount of words in each tweet
emoSNJoy_normalized - It means the emoSNJoy dividing by the amount of words in each tweet
emoSNSadness_normalized - It means the emoSNSadness dividing by the amount of words in each tweet
emoSNSurprise_normalized - It means the emoSNSurprise dividing by the amount of words in each tweet
emoSNFear_normalized - It means the emoSNFear dividing by the amount of words in each tweet

d)SentiSense

sentisenseAnger - It refers to the amount of words associated with Anger
sentisenseDisgust - It refers to the amount of words associated with Disgust
sentisenseJoy - It refers to the amount of words associated with Joy
sentisenseSadness - It refers to the amount of words associated with Sadness
sentisenseSurprise - It refers to the amount of words associated with Surprise
sentisenseFear - It refers to the amount of words associated with Fear
sentisenseAnticipation - It refers to the amount of words associated with Anticipation
sentisenseLike - It refers to the amount of words associated with Like
sentisenseLove - It refers to the amount of words associated with Love
sentisenseAnger_normalized - It means the sentisenseAnger dividing by the amount of words in each tweet
sentisenseDisgust_normalized - It means the sentisenseDisgust dividing by the amount of words in each tweet
sentisenseJoy_normalized - It means the sentisenseJoy dividing by the amount of words in each tweet
sentisenseSadness_normalized - It means the sentisenseSadness dividing by the amount of words in each tweet
sentisenseSurprise_normalized - It means the sentisenseSurprise dividing by the amount of words in each tweet
sentisenseFear_normalized - It means the sentisenseFear dividing by the amount of words in each tweet
sentisenseAnticipation_normalized - It means the sentisenseAnticipation dividing by the amount of words in each tweet
sentisenseLike_normalized - It means the sentisenseLike dividing by the amount of words in each tweet
sentisenseLove_normalized - It means the sentisenseLove dividing by the amount of words in each tweet

e) DepecheMood (dm)

dmAfraid - It refers to the amount of words associated with Afraid
dmAmused - It refers to the amount of words associated with Amused
dmAngry - It refers to the amount of words associated with Angry
dmAnnoyed - It refers to the amount of words associated with Annoyed
dmDontcare - It refers to the amount of words associated with Dont_Care
dmHappy - It refers to the amount of words associated with Happy
dmInspired - It refers to the amount of words associated with Inspired
dmSad - It refers to the amount of words associated with Sad
dmAfraid_normalized - It means the dmAfraid dividing by the amount of words in each tweet
dmAmused_normalized - It means the dmAmused dividing by the amount of words in each tweet
dmAngry_normalized - It means the dmAngry dividing by the amount of words in each tweet
dmAnnoyed_normalized - It means the dmAnnoyed dividing by the amount of words in each tweet
dmDontcare_normalized - It means the dmDontcare dividing by the amount of words in each tweet
dmHappy_normalized - It means the dmHappy dividing by the amount of words in each tweet
dmInspired_normalized - It means the dmInspired dividing by the amount of words in each tweet
dmSad_normalized - It means the dmSad dividing by the amount of words in each tweet

6. Dimensional models of Emotions

a) ANEW

anewValence - It refers to the sum of the valence values for the words in each tweet that are defined in ANEW

anewArousal - It refers to the sum of the arousal values for the words in each tweet that are defined in ANEW
anewDominance - It refers to the sum of the dominance values for the words in each tweet that are defined in ANEW
anewValence_normalized - It means the anewValence dividing by the amount of words in each tweet
anewArousal_normalized - It means the anewArousal dividing by the amount of words in each tweet
anewDominance_normalized - It means the anewDominance dividing by the amount of words in each tweet

b) Dictonary of Affect in Language (dal)

dalPleasantness - It refers to the sum of the pleasantness values for the words in each tweet that are defined in the Dictonary of Affect in Language
dalActitude - It refers to the sum of the actitude values for the words in each tweet that are defined in the Dictonary of Affect in Language
dalImagery - It refers to the sum of the imagery values for the words in each tweet that are defined in the Dictonary of Affect in Language
dalPleasantness_normalized - It means the dalPleasantness dividing by the amount of words in each tweet
dalActitude_normalized - It means the dalActitude dividing by the amount of words in each tweet
dalImagery_normalized - It means the dalImagery dividing by the amount of words in each tweet

c) SenticNet (sn)

snPleasantness - It refers to the sum of the pleasantness values for the words in each tweet that are defined in SenticNet
snAttention - It refers to the sum of the attention values for the words in each tweet that are defined in SenticNet
snSensitivity - It refers to the sum of the sensitivity values for the words in each tweet that are defined in SenticNet
snAptitude - It refers to the sum of the aptitude values for the words in each tweet that are defined in SenticNet
snPleasentness_normalized - It means the snPleasantness dividing by the amount of words in each tweet
snAttention_normalized - It means the snAttention dividing by the amount of words in each tweet
snSensitivity_normalized - It means the snSensitivity dividing by the amount of words in each tweet
snAptitude_normalized - It means the snAptitude dividing by the amount of words in each tweet

Additionally, two features belonging to the irony detection model output are added.

ironyHashtag - It refers to the presence of the hashtags #sarcasm or #not
ironyIDM - It refers to the result of the irony detection model to determine the presence of irony