

Features for regression model of **expressed satisfaction** in the support-seekers' replies to the received comment

Name	Features / Keywords / Process methods	Note
66 LIWC features	['function', 'pronoun', 'ppron', 'i', 'we', 'you', 'they', 'ipron', 'article', 'prep', 'auxverb', 'adverb', 'conj', 'negate', 'verb', 'adj', 'compare', 'interrog', 'quant', 'affect', 'posemo', 'negemo', 'anx', 'anger', 'sad', 'social', 'family', 'friend', 'female', 'cogproc', 'insight', 'cause', 'discrep', 'tentat', 'certain', 'differ', 'percept', 'see', 'hear', 'feel', 'bio', 'body', 'health', 'sexual', 'ingest', 'drives', 'affiliation', 'achieve', 'power', 'risk', 'focuspast', 'focuspresent', 'focusfuture', 'relativ', 'motion', 'space', 'time', 'work', 'leisure', 'home', 'death', 'informal', 'swear', 'netspeak', 'nonflu', 'filler']	Words frequency over the total number of words
Sentence count	from nltk.tokenize import sent_tokenize	Weighted by 3
Word count	from nltk.tokenize import TweetTokenizer	Weighted by 30
Unknown	['what to do', 'I don't know', 'not sure', 'what else to do', 'what's the point']	Weighted by number of words
Trial	['try', 'tried', 'I've', 'trying', 'therapist', 'med', 'therapy', 'doctor', 'pill', 'doc']	Binary
Lonely	['lonely', 'alone', 'loneliness', 'no friend', 'no one', '0 friend', 'No friend', 'nobody', 'no fucking friend', 'dont have friend', 'dont have any friend', 'don't have any friend']	
Sadness	['sad', 'tired', 'horrible', 'terrified', 'scared', 'afraid', 'frustrated', 'awful', 'pathetic', 'upset', 'fear', 'boring', 'Boring', 'miserable', 'tiring', 'feel down']	
Exciting mark	['!', ':)', '<', 'Wow', 'wow']	Weighted by 2
Feeling	feeling = ['feel']	
Negation	["not", "no", 'nobody', 'nothing', 'none', 'never', 'neither', 'nor', 'nowhere', 'hardly', 'scarcely', 'barely', "don't", "didn't", "isn't", 'wasn't', 'shouldn't', 'wouldn't', 'couldn't', 'doesn't']	
Thanks	['Thank', 'thank', 'it help', 'helpful', 'it does help', 'That helps', 'supportive']	Binary
Agreement	['means a', 'feel you', 'response', 'advice', 'you are right', 'Me too', 'the same feeling', 'feel similar', 'the same thing', 'exactly', 'Exactly']	

Note: the features and weights are chosen first by relevance and then by trial and error.

Features for classifier of **sought Informational Support (IS)** in the thread-starting post

Name	Features / Keywords / Process methods	Note
65 LIWC features	['function', 'pronoun', 'ppron', 'i', 'we', 'you', 'shehe', 'they', 'ipron', 'article', 'prep', 'auxverb', 'adverb', 'conj', 'negate', 'verb', 'adj', 'compare', 'interrog', 'number', 'quant', 'affect', 'posemo', 'negemo', 'anger', 'sad', 'social', 'friend', 'female', 'male', 'cogproc', 'insight', 'cause', 'discrep', 'tentat', 'certain', 'differ', 'percept', 'see', 'hear', 'feel', 'bio', 'body', 'health', 'sexual', 'ingest', 'drives', 'affiliation', 'achieve', 'power', 'reward', 'risk', 'focuspast', 'focuspresent', 'focusfuture', 'relativ', 'motion', 'space', 'time', 'work', 'leisure', 'home', 'informal', 'swear', 'filler']	Words frequency over the total number of words
Sentence count	from nltk.tokenize import sent_tokenize	Weighted by 8
Word count	from nltk.tokenize import TweetTokenizer	Weighted by 180
Url	re.findall('http[s]?://(?:[a-zA-Z] [0-9] [\$-_@.&+] [*\(\),](?:%[0-9a-fA-F][0-9a-fA-F]))+', doc)	Binary
POS tags	proper_nouns (NNP or NNPS); adj (JJ or JJR or JJS); cardinal_numbers (CD)	Weighted by total pos
Advice	['advice', 'option', 'tip', 'suggestion']	Weighted by number of words
Trial	['try', 'tried', 'I\ve', 'trying', 'therapist', 'med', 'therapy', 'doctor', 'pill']	
Unknown	['what to do', 'I don\t know', 'not sure', 'what else to do', 'what\s the point']	Weighted by number of words
Need help	['need help', 'any of you', 'anyone else', 'need support', 'feel the same', 'any feedback', 'help me']	Weighted by number of words
Extreme negativity	['suicide', 'suicidal', 'cut myself', 'hurt myself', 'die', 'dying', 'dead', 'death', 'can\t leave', 'self harm', 'kill', 'no hope']	Weighted by 4
Question mark	['?']	

Note: the features and weights are chosen first by relevance and then by trial and error.

Features for classifier of **sought Emotional Support (ES)** in the thread-starting post

Name	Features / Keywords / Process methods	Note
65 LIWC features	['function', 'pronoun', 'ppron', 'i', 'we', 'you', 'ipron', 'article', 'prep', 'auxverb', 'adverb', 'conj', 'negate', 'verb', 'adj', 'compare', 'interrog', 'quant', 'affect', 'posemo', 'negemo', 'anx', 'anger', 'sad', 'social', 'family', 'friend', 'female', 'cogproc', 'insight', 'cause', 'discrep', 'tentat', 'certain', 'differ', 'percept', 'see', 'hear', 'feel', 'bio', 'health', 'sexual', 'drives', 'affiliation', 'achieve', 'power', 'reward', 'risk', 'focuspast', 'focuspresent', 'focusfuture', 'relativ', 'motion', 'space', 'time', 'work', 'leisure', 'home', 'relig', 'death', 'informal', 'swear', 'netspeak', 'assent', 'nonflu']	Words frequency over the total number of words
Sentence count	from nltk.tokenize import sent_tokenize	Weighted by 9
Word count	from nltk.tokenize import TweetTokenizer	Weighted by 150
POS tags	proper_nouns (NNP or NNPS); adj (JJ or JJR or JJS); cardinal_numbers (CD)	Weighted by total pos
Lonely	['lonely', 'alone', 'loneliness', 'no friend', 'no one', '0 friend', 'No friend', 'nobody', 'no fucking friend', 'dont have friend', 'dont have any friend', 'don\'t have any friend']	Binary
Sadness	['sad', 'tired', 'horrible', 'terrified', 'scared', 'afraid', 'frustrated', 'awful', 'pathetic', 'upset', 'fear', 'boring', 'Boring', 'miserable', 'tiring', 'feel down']	Binary
Want a talk	['want to talk', 'talk to me', 'talks to me', 'Anyone else the same', 'Does anyone', 'Do anyone', 'needed to get it out', 'does everyone else', 'Does everyone else', 'need help', 'anyone else', 'any of you', 'feel the same', 'in the same situation', 'someone can relate', 'Anybody else', 'anybody else', 'someone to talk to', 'Has everyone ever felt', 'Is anyone else', 'someone please pray']	Binary
Feeling	['feel']	Weighted by 4
Extreme negativity	['suicide', 'suicidal', 'cut myself', 'hurt myself', 'die', 'dying', 'dead', 'death', 'can\'t leave', 'self-harm', 'kill myself', 'no hope', 'hopeless', 'don\'t want to live', 'not worth being alive', 'kills me', 'kill me', 'cutting myself', 'worthless', 'killing me', 'kill yourself', 'meaningless', 'self harm', 'end my life', 'meaningless']	Binary

Note: the features and weights are chosen first by relevance and then by trial and error.