

Tokens

Tokens w/ POS tags

Lightly modified, single-piece tweet

Sentences

Lemmas

Negation-handled

lemmas

Conflated and filtered

lemmas

Normalized tweet

Tokens

Tokens w/ POS tags

Filtered bigrams

Sentence tokenization

Word tokenization

POS tagging

Lemmatization

Negation handling

Simple modifications like expanding contractions or removing unnecessary space

Conflation or removal of

stop words, numbers, users, hashtags, URLs, etc.