Tasks

1. Lexical Features:
(a) Total number of words per post
(b) Total number of sentences per post
(c) Total number of words per sentences
(d) Frequency of dictionary words
(e) Frequency of word extensions (not usual word, like 'hiii!', 'hahahaha', 'heyyyyy')
(f) Lexical diversity
(g) n-grams collection (up to trigram) along with their frequencies
(h) average hashtag per post

2. Syntactical Features:
(a) Total number of beginning of sentences characters capitalised
(b) Number of punctuation per sentence
(c) words with all capital letters and their frequencies
(d) alphanumeric words with their frequencies
(e) special characters, digits, exclamation, punctuation and question marks with frequencies
(f) No of upper case letters
(g) Average usage of quotations

3. Special features:
(a) Emoji collection along with their frequencies
(b) Number of Emojis per word & post
(c) Mentions of other users & URLS with frequency
(d) Hashtags collections with their frequencies