Feature extraction from text and images

Total points 4

1. TF-IDF is applied to a matrix where each column represents a word, each row represents a document, and each value shows the number of times a particular word occurred in a particular document. Choose the correct statements. **1 point**

- [ ] TF normalizes sum of the column values to 1
- [ ] IDF scales features inversely proportionally to a number of word occurrences over documents
- [ ] TF normalizes sum of the row values to 1
- [ ] IDF scales features proportional to the frequency of word's occurrences

2. What of these methods can be used to preprocess texts? **1 point**

- [ ] Plumbing
- [ ] Levenshteining
- [ ] Lowercase transformation
- [ ] Plumping
- [ ] Stemming
- [ ] Lemmatization
- [ ] Stopwords removal

3. What is the main purpose of Lemmatization and Stemming? **1 point**

- ( ) To induce common word amplification standards to the most useful for machine learning algorithms form.
- ( ) To reduce significance of common words.
- ( ) To reduce inflectional forms and sometimes derivationally related forms of a word to a common base form.
- ( ) To remove words which are not useful.

4. To learn Word2vec embeddings we need ... **1 point**

- [ ] Labels for the documents in the corpora
- [ ] Text corpora
- [ ] GloVe embeddings
- [ ] Labels for each word in the documents in the corpora

Coursera Honor Code Learn more

- [ ] I, **Chinmay kumar Das**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Submit Save draft