Assignment 1 – Update the vocabulary of the patient sentiment classification demo

Task:

Augment the RNN patient sentiment classification notebook as follows:

1) Analyse for 2 new patient records where patient note contains words not in the vocabulary, how it is handled. Print the index of the tokens.
2) Modify the vocabulary to add these words and a few training records with these words and see if the accuracy of the prediction improves for these new records

Submit:

Link to completed notebook with the outputs in it

Assignment 2 – Word Sense: Embeddings

Task:

Discover how different families of embedding models treat the polysemous word **"bank."**

You will analyse embeddings created by models Spacy (en_core_web_md), word2vec for **polysemous** words - *finance* sense, the *river-edge* sense, or both.

Take following 2 sentences :

"I sat on a river bank",

"I went to bank to deposit money"

Task:

1) print the embeddings for the word "bank" by using different embedding models such as word2vec, spacy
2) Retrieve the 20 nearest neighbors of the token "bank" using cosine similarity.

Submit:

Link to completed notebook with the outputs in it