Nirmit Shah 20622139

1. **What you think about the dataset?**
   a. I think the dataset is pretty thorough. There are multiple forms of verbs and their tenses and many different sentence types. There are also some unique ones that aren't complete sentences like "the Alaskan Knight" so it would be interesting to see how the algorithm handles that. As "Knight" is not usually a proper noun but in this case it is if it's a book title or similar. There is a good mix of formal sentences and very casual ones that should capture most of the sentences tried in the whole space. There are also different lengths of sentences so it will test how the network works with different sized inputs.

2. **Please explain your observation different embedding size and different numbers of hidden layers.**
   a. (embedding size, hidden layers, num epochs) = (accuracy, training time in seconds)
   b. BaseLine: (1,1,1) = (.15, .253)

Embedding Size	Hidden Layers	Num Epochs
(10,1,1) = (.54, .322)	(1,2,1) = (.17, .322)	(1,1,2) = (.21, 473)
(20,1,1) = (.62, .386)	(1,3,1) = (.18, .382)	(1,1,4) = (.23, .897)
(40,1,1) = (.66, .565)		(1,1,8) = (.31, 1.717)
(80,1,1) = (.72, 1.196)		

   c. Increasing numepochs and embedding size greatly increase the training time.
   d. (25,1,8) = (.77, 2.965)
   e. (25,2,8) = (.77, 3.932)
   f. (25,3,8) = (.77, 5.064)
   g. From d,e,f, it looks like numHiddenLayers doesn't have much of an impact on the accuracy but does have an impact on the training time.
   h. (30, 1, 6) = (.78, 2.621)

3. **How good do you think the token-based feedforward model is? Why?**
   a. I think there are some numerous shortcomings to the token-based model. In one of my datasets the word "Email" at the very beginning was capitalized and the model classified it as a proper noun. Sometimes "Email" can be a proper noun, but it has to be figured out from context which this model doesn't use. However, the token based model does have some good qualities. For some words like "year" or "into" that are usually

only seen as nouns or prepositions, the model got those words correctly. The model also has some problems with plurality and figuring out of a word is a plural or the singular verb (I take walks vs He walks). Sometimes, it marked the word "officers" as a VBN when it should be an NNS. I feel like the token-based model is pretty good at basic classification for the time it takes to train it.

4. **Please suggest an alternative model that solves the drawbacks you observe.**
    a. I think context is a major thing that is lacking with this model. Currently it only looks at one word at a time to determine part of speech and that isn't the best way to determine part of speech because the same word could be different parts of speech depending on the context. A different model would be one that incorporates the parts of speech of some of the previous words and some of the following words. This gives more context and lets the model make a more 'guided' decision and reduce language bias.