Week starting 17/10/22 Status Report

What I achieved

* Setup anaconda environments and scikit-learn properly
* Completed scikit learn and NLTK tutorial
* Completed text classification with Kera tutorial
* Completed text classification with Spacy tutorial
* Completed scikit learn working with data tutorial
* Read through and tried parsing enron script

Time Spent - ~9 hours

Issues

The scikit-learn work with data tutorial had some issues in it that I had to work around, and I did not get to parse all the examples but I still read through everything

The enron script failed to parse the pre-processing section as I think it may use some older version of the environments that are imported.

Questions

Do you have an updated/ error free version of the enron script, or should I try and come up with my own solution?

What is the frequency were the category was assigned to each message in the enron .cats files? As in n3 in the categories.txt file where it explains the categories.