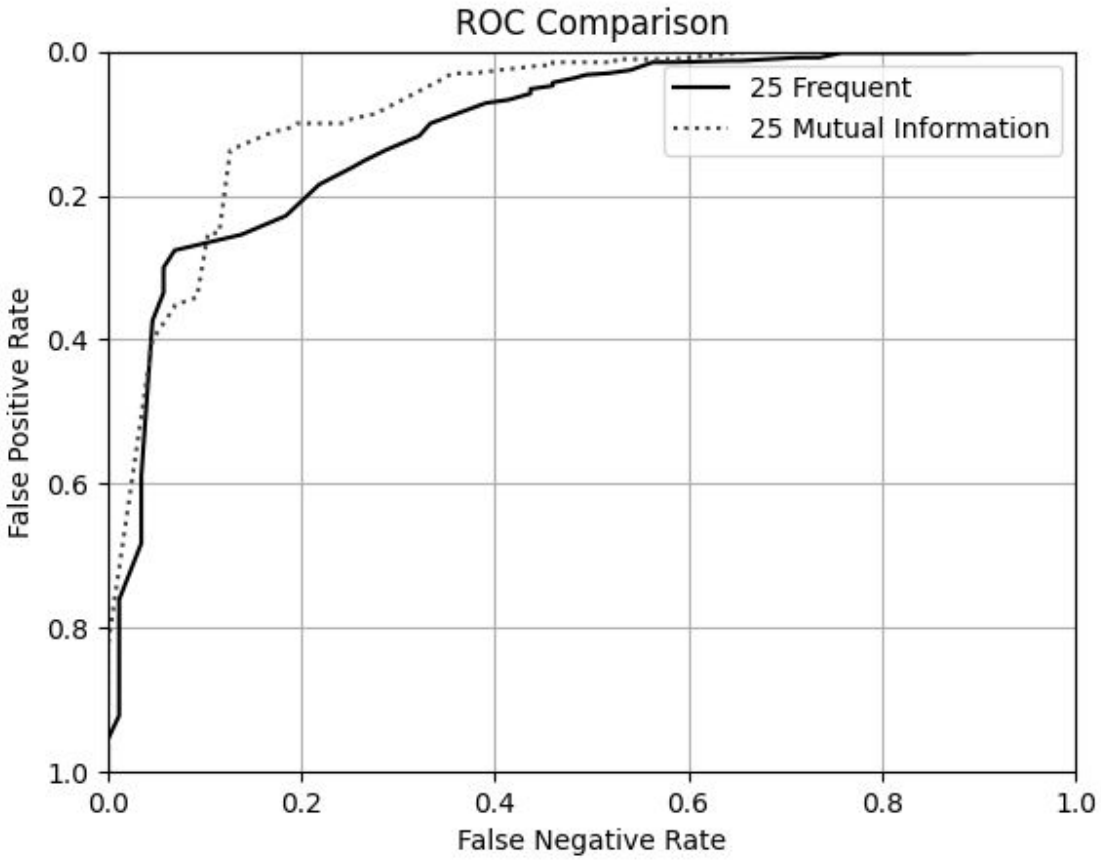
Answer the following questions about the models produced by the framework code:



Which model is better at a 50% False positive rate? Mutual Information
Which model is better at a 10% False positive rate? Mutual Information
Which model is better at a 40% False negative rate? Mutual Information

What classification threshold would you use with the numFrequentWords = 25 model to achieve a 10% False positive rate? For a 10% FPR is a threshold of 0.2.

Tune the hyperparameters 'numMutualInformationWords' and 'numFrequentWords' until you find a new model that is better than both of the models in the framework at both the 50% and the 10% false positive rate.

In a few sentences, describe the strategy you used to find the better hyperparameters and the hyperparameters you found to achieve the result?

I first tried number of features near 25 (e.g. +/- 5 to 10). It was unclear how much better these models were at the 50% and 10% FPR. Then I just doubled the number of features and voila.

Create an ROC plot showing:

*** the model with 25 features by frequency**
*** the model with 25 features by mutual information**
*** and your new model**