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

1. Which model is better at a 50% False positive rate?
Ans: 25 most frequent words is better at 50% False positive rate

2. Which model is better at a 10% False positive rate?
Ans: 25 words with most mutual information is better at 10% False positive rate

3. Which model is better at a 40% False negative rate?
Ans: 25 words with most mutual information is better at 40% False negative rate

4. What classification threshold would you use with the numFrequentWords = 25 model to achieve a 10% False positive rate?

Ans: Classification threshold = 0.22 results in 10% False positive rate

Tune the hyperperamters '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 hyperperparameters you found to achieve the result?

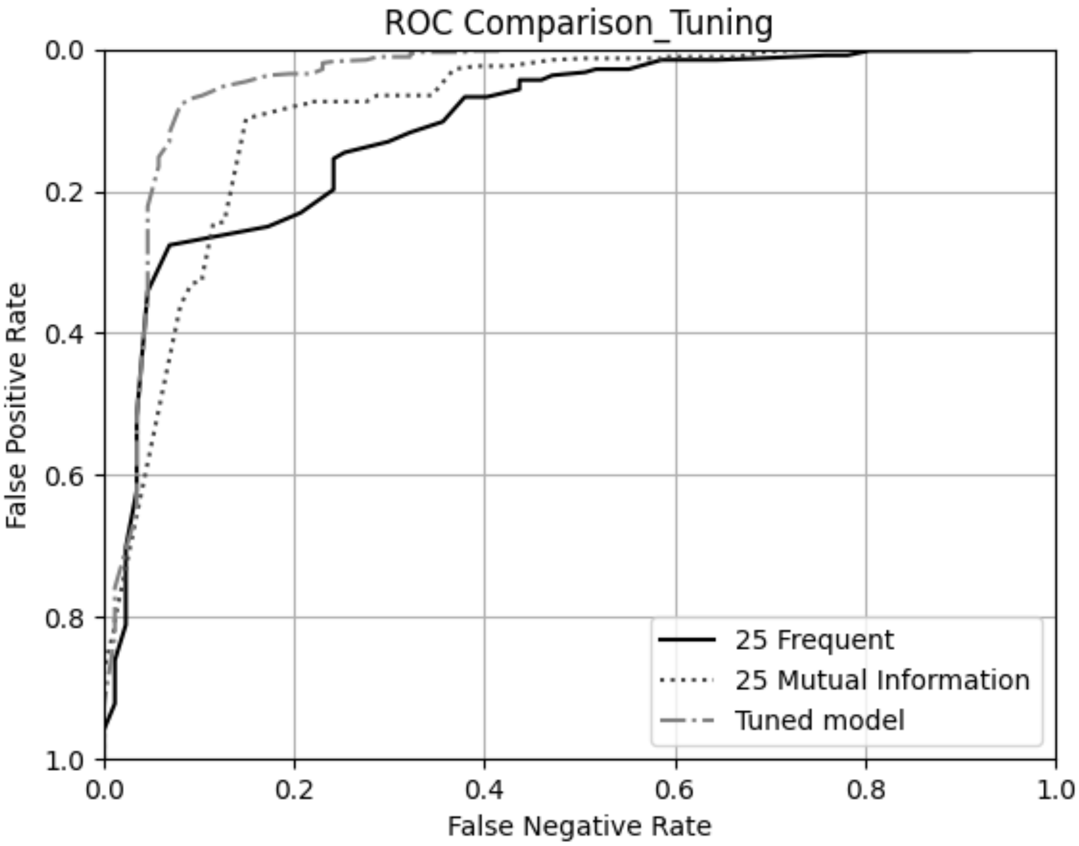
Ans:

1. *Tried increasing the number of frequent words, and saw no improvement at 50% FPR. There is slight improvement at 10% FPR, but it is not better than 25 Mutual Information.*
2. *Tried increasing the number of mutual information words and saw no improvement at 50% FPR. There is improvement at 10% FPR.*

I observed that using mutual information gives better result at 10% FPR, whereas using frequent words gives better results at 50% FPR. I tried both 'numMutualInformationWords' = 25 and 'numFrequentWords' = 100 features in a single model. This model gave an ROC plot that is better at 10% FPR, and no worse at 50% FPR.

Create an ROC plot showing:

* the model with 25 features by frequency;
* the model with 25 features by mutual inforamtion;
* and your new
model.


ROC Comparison_Tuning
False Positive Rate
False Negative Rate
25 Frequent
25 Mutual Information
Tuned model