Room for improvements

1. Include sklearn classification\_report
2. Include F1 scores: this is a good metric to measure specificity and sensitivity at the same time
3. When presenting, show the test and train together
4. For the final model, ensure that the difference in test and train score is less than 3-5% difference to ensure that model is not overfit to the train dataset