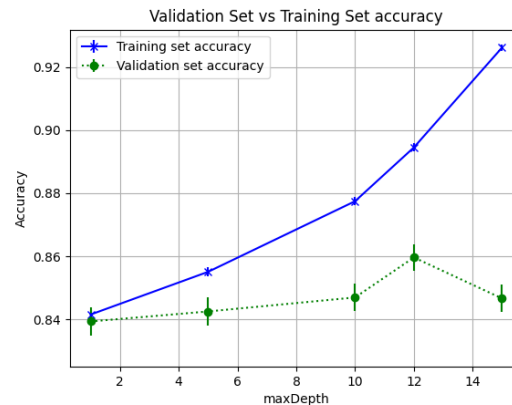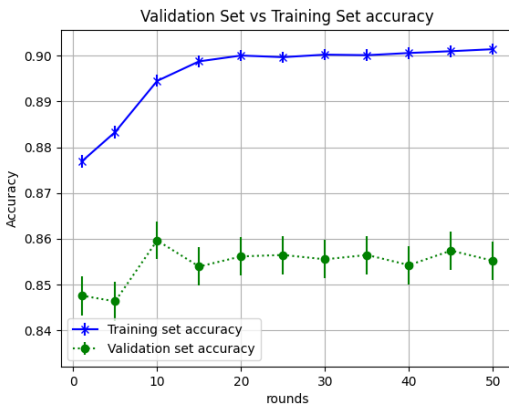6 points - A short writeup describing the tuning process. The bare minimum text to clearly indicate the following elements so the TA can find the required pieces and verify they were done correctly. Include at least one chart showing a parameter sweep demonstrating tuning boosted trees. Include an ROC curve of the best versions of logistic regression, decision trees, and boosted trees you found when trained on the full training set and evaluated on the test data. Include the parameter settings and accuracy of those three best models -- and error bounds.

Parameter tuning boosted trees:

I tuned the hyperparameter rounds first and the maxDepth on the boosted tree. I started observing overfitting at rounds = 10 when there was a sudden drop in validation set accuracy. When tuning MaxDepth, I observed overfitting at maxDepth=12. Therefore I stopped at those hyper parameters for boosted trees.
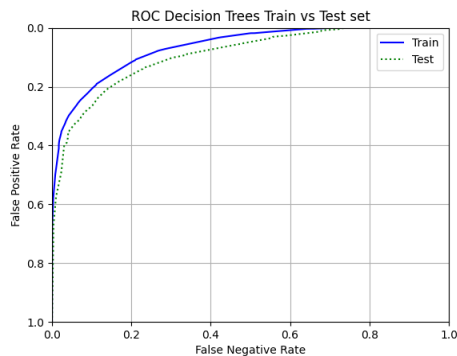
ROC Curves
1. **Decision Trees**
   Parameter Settings: MaxDepth = 12, with numeric features
   Train set accuracy: 87.69% (50% confidence interval 87.55% - 87.83%)
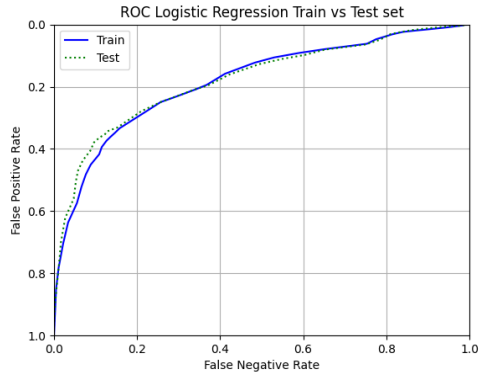   Test set accuracy: 84.75 (50% confidence interval 84.33% - 85.18%)

2. Logistic Regression

Parameter: convergence=0.0001, stepSize=0.001, with normalized numeric features
Train set accuracy: 76.09% (50% confidence interval 75.91% - 76.45%)
Test set accuracy: 75.44%(50% confidence interval 74.93% - 75.95%)

3. Boosted Trees

Parameter: Maxdepth = 12, Rounds = 10, with numeric features
Train set accuracy: 89.44% (50% confidence interval 89.31% - 89.57)
Test set accuracy: 75.12%(50% confidence interval 74.61% - 75.64%)