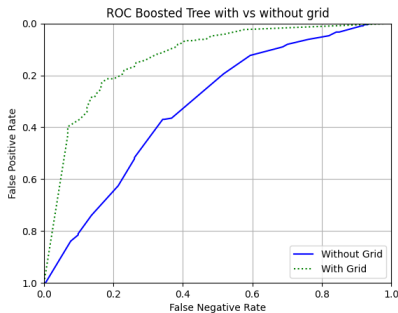
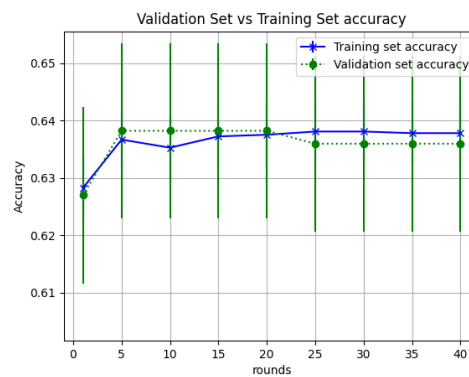
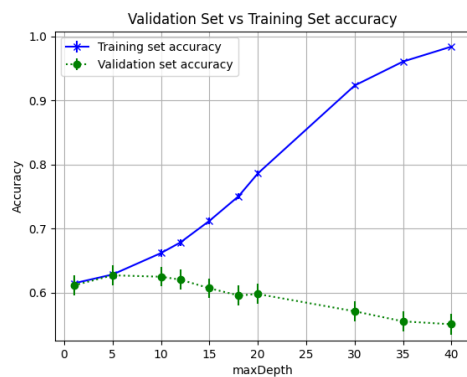


4 points - A short writeup using precise technical terms describe if the new features improved validation-set accuracy. No more than 4 charts and ~200 words. Make sure to include:

- \* error bounds,
- \* hyperparameter tuning visualization,
- \* training vs validation accuracy (and some bias variance discussion),
- \* ROC with and without the features.

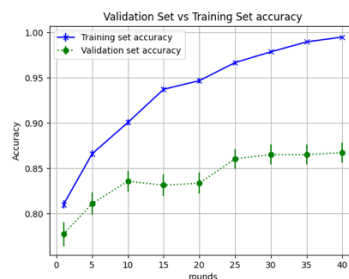
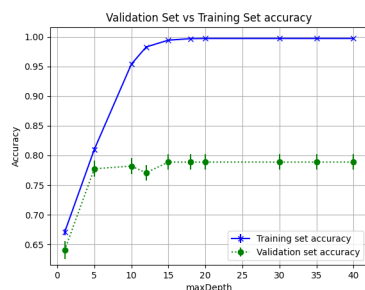


### Accuracy without 3x3 grids:



I trained the models by sweeping the hyperparameter maxDepth followed by tuning the hyperparameter rounds for the boosted tree. The best validation set accuracy without splitting the image into grids of 3 x 3 pixels was 62.69% (50% confidence interval 62.15% - 63.23%) at maxDepth = 5 before reaching a point of diminishing returns. Boosting did not help increase the accuracy with 50% confidence levels when the image was not split into 3x3 grids. The model ran into variance issues when trying to tune the parameter number of rounds because there was no improvement in the accuracies.

### Accuracy with 3x3 grids:



Splitting the image into 3x3 pixel grids and fitting it into a boosted tree gave best validation set accuracy of 86.06% (50% confidence interval 84.96% – 87.16%) at max depth = 5 and rounds = 25. Therefore, splitting the image into 3x3 grids improved the validation set accuracy by at least 21.73% with one sided 50% confidence interval. The ROC curve also shows the significant improvement in the area under curve as a result of splitting the image into 3x3 grids. This model shows variance issues after maxDepth = 5 and number of rounds = 25.