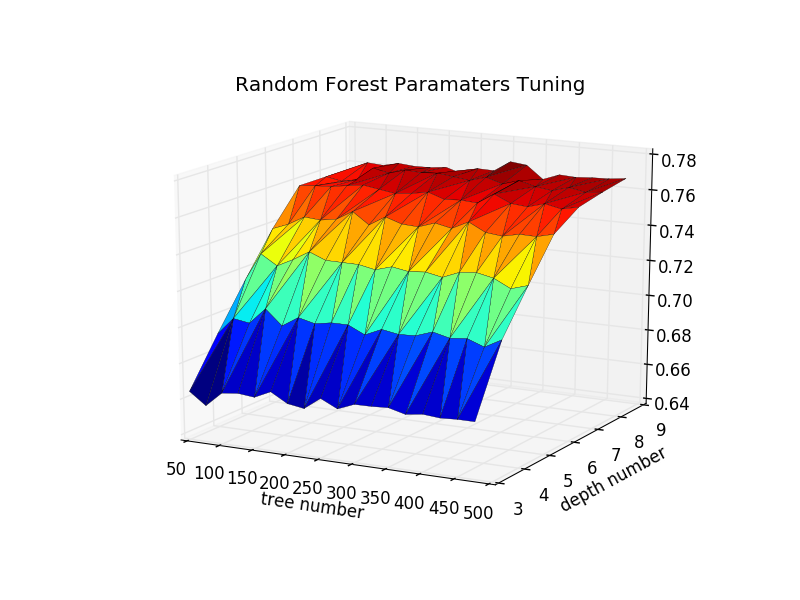
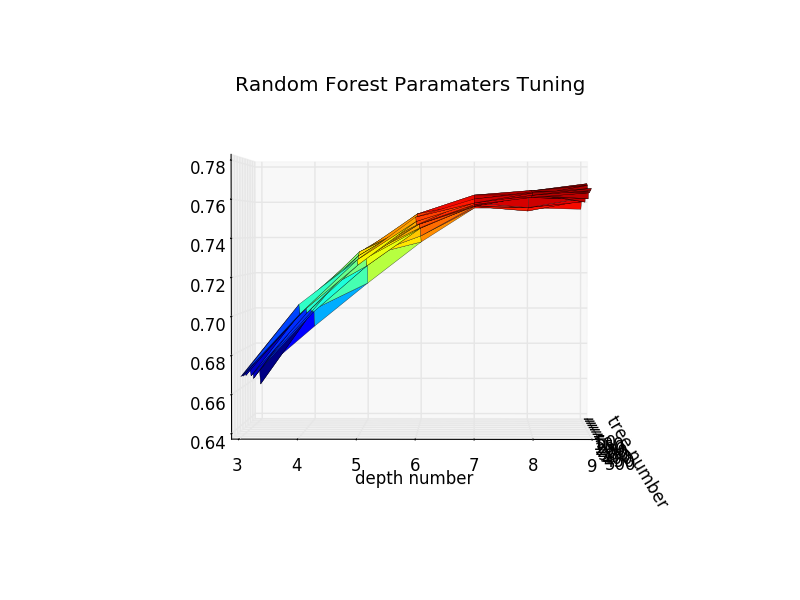
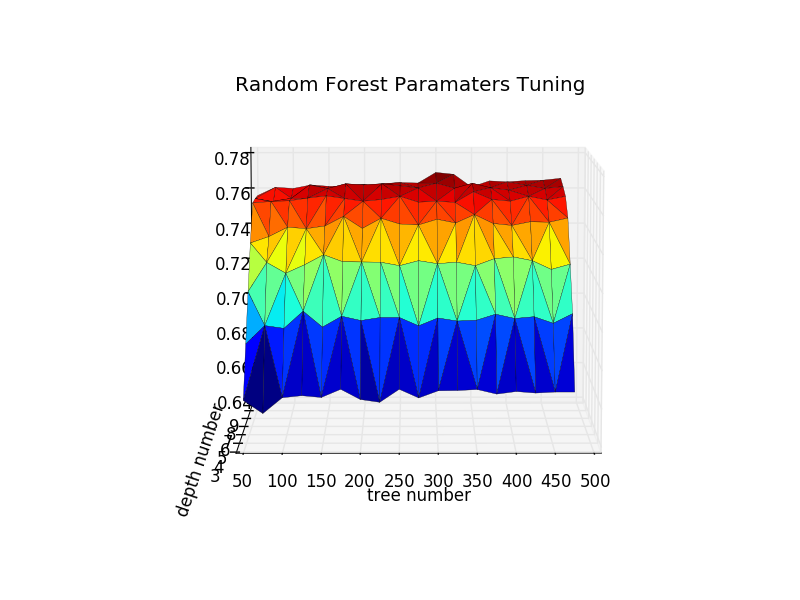
In terms of Random Forest, there are a few parameters can be tuned, and the number of tree and the maximum depth of each tree are two of the most important parameters. In order to tune these two parameters, we use the 10-fold cross validation on the training set.

The z-axis denotes the mean accuracy on the validation set, and the x-axis and y-axis denote the tree number and depth number respectively, shown as the following figure. The tree number ranges from 50 to 500 with 50 step size, and depth number ranges from 3 to 9 with 1 step size.





From the above figure, we know that as the depth number increases, the mean accuracy on the validation set increases as well, but when it becomes flat after the depth number is 7. However, as the depth number increase, the running time of the algorithm increases. In terms of accuracy and efficiency, we think 7 may be the best parameter for our model.



From the above figure, we know that the tree number does not have a crucial influence on our problem. Commonly speaking, the more trees in the random forest, the more stable of the model. However, considering the efficiency of our model, we choose tree number as 250.