1. Status and Stopping Point

There is no stopping point in the assignment. All the planned methods implemented and discussed in the analysis.

1. Addition Functions
2. Feature engineering and data transformation to make training time feasible.
3. Long time simulations and comparison between the effect of estimators and tree depth
4. Easy and challenging parts.

The challenging part is thinking about features that can represent the text in a way that make training time acceptable. Also, running the algorithm overnights to get all possible combination for in-depth analysis. Easy part was implementing the algorithm using the tools provided by sickit learn.