1. Linear classifier
   1. Logistic regression
      1. Lbfgs
      2. Accuracy
   2. 20 good/bad
   3. PCA
      1. Computation time
   4. Advantage/disadvantage
2. KNN
   1. KNN
      1. Accuracy
   2. Computation
      1. # data point
      2. Which set
   3. Advantage/disadvantage
   4. Optimal K
   5. Split after TF-IDF
3. Decision tree
   1. Test accuracy
   2. Optimal depth
   3. 2 advantage
4. Validation
5. Cross validation
   1. 5 fold
   2. Confusion matrix
6. New feature
   1. Why
   2. Re-optimize