

Machine Learning

Aaron

Ex1

- Clean Away "Signature Words"
- TfIdf It
- Accessing TfIdf Features

Ex2

- Number of Features and Overfitting
- Accuracy of Your Overfit Decision Tree
- Identify the Most Powerful Features and Use TfIdf to Get the Most Important Word
- Remove, Repeat
- Accuracy of the Overfit Tree

Ex3

- Explained Variance of Each PC
- F1 Score vs. No. of PCs Used

Ex4

- Your First (Overfit) POI Identifier
- Deploying a Training/Testing Regime

Ex5

- Applying Metrics to Your POI Identifier
- Unpacking Into Precision and Recall
- Recall of Your POI Identifier
- How Many True Positives
- Precision
- Recall