**W207 Final Project Notes and Ideas for Future Work**

**Grading scale**

Prob Detection (10pts)

Sensible Methods (25pts)

Feature Engineering (20pts)

Error Analysis (15pts)

Notebook Presentation (10pts)

Good Results (10pts)

In Class Presentation (10pts)

**General Notebook Structure**

* Make it a story!
* Make it fun to read!
* Use visualizations rather than table dumps
* Summarize small choices like types of scaling, focus on big things like what it does
* Show what worked and what didn’t work, discuss why
* [ ] Discuss tree types, possibly add photos?
* [ ] Discuss what evaluation measures we will be using (accuracy, precision, recal, f1-score) and why we chose those over others.

**Data Exploration**

* [ ] Correlation matrix? Maybe not in final report but could be nice for learning
* [ ] Graph features, and colorize points by label to view groupings

**Feature Engineering**

* [ ] Combine distance to hydrology, Euclidean distance?
* [ ] Remove soil types 7 and 15 because they have no
* Check out making everything discrete?

**Models etc.**

* [ ] Make final validation function
* [ ] Replace support with accuracy in cross validation function
* [ ] Make function to compare how much a feature improves the model?