# Slide List

1. Introductory slide – names, etc.
2. Phishing statistics slide (See speech shorthand)
3. Dataset Page – Web Page Phishing Detection Dataset. Get stats from shorthand, have a few screenshots from Kaggle
4. Genetic Algorithm – Just a title and the GA image (See github repo)
5. Slide for datacleaning steps, have dot point with each of these titles – Pearson correlation, analysis of correlation heatmap, use of fine-grained feature selection, scaling methods
6. Pearson Correlation bar graph – see miniproject.ipynb on git repo for more details
7. Correlation Heatmap – see miniproject.ipynb on git repo
8. Fine-grained feature selection – one-liners for PCA and feature mask. Have some nice pictures
9. Scaling methods – list robust, normalizing, standardizing and no scaling
10. KNN slide 1 – Introduce KNN. Mostly pictures. Equation for Euclidean Distance
11. KNN Slide 2 – expectations for testing + training times, expectations for dimensionality reduction and scaling methods
12. Decision Tree slide 1 – Introduce DT. Mostly pictures. Equation for Entropy and Thresholding
13. Decision Tree Slide 2 – expectations for testing + training times + accuracy, expectations for dimensionality reduction and scaling methods
14. Random Forest slide 1 – Introduce RF. Mostly pictures. Say that it’s an ensemble of decision trees
15. Random Forest Slide 2 – expectations for testing + training times + accuracy, expectations for dimensionality reduction and scaling methods
16. KNN Results Slide – Show testing and training times, accuracy and the GA’s final refined model (Discuss in group, cos the GA’s final model might have to be tinkered to look good on a slide)
17. DT Results Slide – Show testing and training times, accuracy and the GA’s final refined model (Discuss in group, cos the GA’s final model might have to be tinkered to look good on a slide)
18. RF Results Slide – Show testing and training times, accuracy and the GA’s final refined model (Discuss in group, cos the GA’s final model might have to be tinkered to look good on a slide)
19. Results Comparison slide – Table showing comparison of each model’s training + testing times, as well as its accuracy
20. Conclusion slide – that’s all folks