Jeffrey Li, Valari Graham
Dr. Mazidi
CS 4375

a. How SVM works, and how SVM kernels work, your impression of the strengths and weaknesses of SVM

Support Vector Machine (SVM) is a supervised learning algorithm that can be applied to both regression and classification models. The data points are plotted and a hyperplane is used to separate the two classes of data points. The hyperplane has 2 margin lines and the data points that fall on the margins are called support vectors. These points determine the best hyperplane, which maximizes the distance between the support vectors and the hyperplane. One issue that may come up is that the points are not linearly separable, which is where the kernels come in. Although the kernel is normally linear, a polynomial or radial kernel can be applied to better separate the data, making SVM very versatile. SVM has many strengths, including being applicable to both regression and classification, being useful even in the presence of outliers, and being useful in higher dimensions. However, SVM takes a long time for larger datasets and there is a good chance it will not perform better than the linear model.

b. How Random Forest works, how the other 2 algorithms you used work compared to the simple decision tree, your impression of the strengths and weaknesses of these ensemble techniques

Random forest is a popular machine learning algorithm based on decision trees using the ensemble method. The random forest algorithm can be used for both classification and regression. If there are enough trees in the forest, then the classifier won't lead to overfitting the model. Other advantages of the random forest model is that it can handle missing values and can be modeled for categorical values. It's more accurate than the decision tree algorithm.
Furthermore, the XGBoost is another popular machine learning algorithm. Gradient boosting is a supervised learning algorithm that uses weak learners, such as decision trees, to predict a target variable. It uses the gradient descent algorithm to minimize the loss when adding new models. A prominent advantage of XGBoost is the fact that the hyper-parameters can be tuned.

Lastly, Adaboost uses the principle of learners growing sequentially. Each subsequent learner is grown from previously grown learners. An advantage of Adaboost is that it is less prone to overfitting. This is due to the fact that the input parameters are not optimized together.