Blake Lopez
ECE 532 - Project Update 2
12/01/2020

Hazardous Asteroid Classification

The LASSO algorithm with SMOTE has been completed. Each weight solved for in the notebook uses the ista_solve_hot function which takes an array of regularization parameters and returns weight for each regularization parameter. Then the regularization parameter with the lowest error on the pre over sampled data (before SMOTE is applied) is chosen.

A ranking method was used in this algorithm as well. However, instead of deleting features, the feature with the lowest weight was deleted. This is possible because the original features, which are all positive, have been scaled to be between 0 and 1. This ranking method identified seven features of importance. However, during validation it was found one feature was not necessary. This makes results of this ranking method less trustworthy. The six features suggested by this algorithm are "Absolute Magnitude", "Minimum Orbit Intersection", "Jupiter Tisserand Invariant", "Orbital Period", "Perihelion Distance", and "Aphelion Dist". These features fina an overall error of 13.2% with an error of just 0.6% for the hazardous class. These results are the best thus far.

Final testing of the algorithms completed thus far was conducted using the testing data that was separated earlier. The base least squared case with seven features classified each asteroid as being non-hazardous with an error of 15.8%. This is not surprising considering the imbalance of the class types. The initial guess of three features for the least squared with SMOTE had a prediction accuracy of 13.4% but had an hazardous error of 27.0%. The results from the least squared ranking technique with SMOTE and five features had the same results as the case without SMOTE despite promising results on the training data.

The LASSO algorithm with SMOTE has had the best results with six features giving an error of 19.6%. This error is higher than the initial guess of three features using SMOTE, but the error for the hazardous class had an error of 0%. This means that if the classification says the asteroid is safe, then one can confidently assume it is safe. The neural network algorithm has to be fully completed so final testing and comparisons will be made later.

Significant progress has been made on the project, but due to another class project I am behind schedule. Now that neural networks are being covered in class, the neural network algorithm will be completed. A revised timeline to finish up the project is below. The Github page for following the progress of this project is https://lopezbl.github.io/ECE532_Project/.

Blake Lopez
ECE 532 - Project Update 2
12/01/2020

Revised Project Timeline		
Date	**Tasks**	**Submission**
~~11/01/2020~~	~~Complete pre-processing of the data~~	~~None~~
~~11/08/2020~~	~~Complete the least squares algorithm for classification~~	~~None~~
~~11/16/2020~~	~~Complete the least squares algorithm for classification with SMOTE~~	~~None~~
~~11/17/2020~~	~~Update the Github page with progress~~	~~Update 1~~
~~11/24/2020~~	~~Complete the LASSO algorithm for classification with SMOTE~~	~~None~~
~~11/28/2020~~	~~Complete comparison of algorithms completed thus far~~	~~None~~
~~12/01/2020~~	~~Update Github page with progress~~	~~Update 2~~
12/04/2020	Complete the neural network algorithm for classification	None
12/05/2020	Complete final comparison and recommended classification	None
12/08/2020	Complete first draft of the final report	None
12/12/2020	Submit final draft of the final report	Final Report
12/17/2020	Evaluate two other projects	Peer Evaluations