Title: Foursquare Point-of-Interest Matching
Team: New Horizons
Team Members: Ling Zhou, Yu Cao, Tim Gorman

In this project we tackled a novel challenge posed by Foursquare on Kaggle: detecting true matches for geographic points-of-interest data. The competition description can be found here: https://www.kaggle.com/competitions/foursquare-location-matching. The datasets used for this challenge are from the associated kaggle competition; there are two files, train.csv and pairs.csv. We reduced the competition's original scope to produce a more tractable problem by focusing our modeling on the data in pairs.csv, turning a many-to-many classification problem into a binary classification problem. For this simplified problem we tested four different models: a logistic regression, a random forest, and two naïve models. We found that a random forest model did the best job predicting POI matches across four metrics. For future work we would expand the random forest model to other countries in pairs.csv and work toward modeling the complete training set in train.csv.