

Coursera Capstone – “The Battle of the Neighborhoods”

Comparative analysis using location data via Foursquare

Problem Statement

In today's world, it is a headache for parents to decide which school is good for their children. It is very important for a child to have his/her education at a better place so that he/she should be influenced by bad things nearby. The mission of this project is to use Foursquare location data and regional clustering of venue information to determine what might be the 'best' school in York for a child. My proposal, then, is an analysis of the neighborhoods of schools in York for the consideration of safe places for children. The objective is to evaluate the neighborhoods of schools and analyse whether there is any venue nearby which can be harmful for a child.

Data Overview

The data that we will use for this analysis is a combination of a CSV file that has been prepared for the purposes of the analysis from multiple sources (schoolslst.csv) and the location/venue information in foursquare. Schoolslst.csv has a list of schools in York along with their latitude and longitude. Using foursquare APIs, we can get the details and category of the neighborhood venues and using this we can classify the schools whether it is safe or unsafe for a child using SVM linear classification algorithm.