

Exploring Educational Institutions in a neighbourhood

1. Introduction

1.1 Background

Educational Institution is a place where the young minds are shaped into smart individuals. These institutions have grown manifold in most prominent cities across the globe. Getting into the right institution means better knowledge and better outgrowth in career. But there are lot of challenges in determining the right institution. One such factor is the commuting distance. Today, considering the schools, there are variety of options to choose. But those options may not be preferred by the people. A person can easily find a set of school in their area. But it is just a cluster of data with which a person might not be satisfied. Therefore, an effective data model is required to predict a list of schools within a neighbourhood that matches the user preferences.

1.2 Problem

Imagine if people have the option to explore schools that are nearby to their **Home location** as well as has a **good rating**. Based on these inputs, they could get a variety of schools to choose from. The outcome of this project is to provide a list of options for the people which will ease up the process of finding the best school in their neighbourhood.

1.3 Interests

Longer the distance, longer is the travel time and larger the transportation costs. Hence, it is a fair advantage for the people to discover a school, which is nearby to their home ,and also provides quality education.

2. Data Acquisition

2.1 Data Sources

Dataset on the schools in a neighbourhood can be extracted using the Foursquare API request. From this data, name of the school, rating and the neighbourhood can be used. Another set of data is user input. A table with the list of user names, neighbourhood, approximate distance expected between school and home and rating is created manually.

2.2 Data Usage

A dataset with list of different user preferences on schools is converted to a table. The data will have the below contents.

Location	Distance from Home	Rating
Marble Hills, Manhattan	5kms	5
York, Manhattan	3 kms	4

The above dataset is prepared manually by myself. With this data, I get the required details to proceed with API request.

Initially, the latitude and longitude of each location is determined using geocode. These coordinates are used in the URL request. The list of schools from these coordinates are received from the Foursquare API request. Based on the schools listed, the corresponding user ratings are extracted. This list is passed on to another URL request which will extract the user reviews and tips data for each of the schools listed in the previous extract for a particular neighbourhood.

Further, the schools should be categorised based on the distance. Euclidean distance is used to calculate the distance between the home location and school. With this data, I will compare the distance provided by the user and the Euclidean distance to get the final set of schools.

Finally, the data retrieved in the previous step is combined with the first set of data that is extracted from Foursquare to form a final table. This table is further categorized based on the neighbourhood and the same is visualized in the Maps and is available for the User.