Selection of location for Hospital Establishment in Pune - India

A Project Report

Submitted in Partial Fulfillment of Requirements for IBM Data Science Professional Certificate Program on Coursera

Submitted By – Swaroop Todankar

Date – 20 August 2019

 [1]

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

1. Introduction

With the increase in pollution and unhealthy lifestyle, more and more people are finding it necessary to make use of health services.

According to a report published by World Health Organization and World Bank, almost half of the world population does not have easy access to essential health services whether it may be walk-in clinics or hospital establishments. [2]

The primary factors for the establishment of these facilities are the location of the establishments and the funding required. Most countries have government funding allocated for such projects which are paid through taxes. The following question which arises is that – what are the factors that must be considered in order to select a location for such an establishment?

2. Business Problem

The main question is- If an organization is deciding to open a Hospital establishment in Pune, which location would be the best considering all the factors?

This report aims to put forward an analysis in selection of location for a hospital establishment using location data obtained from foursquare API. The algorithms used can help to identify the perfect location considering proximity to nearby similar facilities and the demand for the services.

3. People interested in the project – Target Audience

The target audience in this scenario are normal people who may need medical services. The main factors contributing to the demand are time constraint and location.

People in vicinity of medical services can use these, when needed without wasting any time in travelling. Further, in case of emergency situations, the proximity of such establishments can be very helpful, as a single second is very crucial in such cases.

4. Data required

The data required to build a model to suggest a location for the establishment is as follows:

4.1 Neighborhood Data: Data pertaining to neighborhoods of Pune city obtained from Wikipedia.

The data specifying the index and the neighborhoods of Pune city can be obtained through Wikipedia page: Neighborhoods of Pune city. This provides a table listing the neighborhoods. Web scraping techniques such as using the Beautiful Soup library or the Wikipedia library can be used to convert this html data into a pandas data frame. This is helpful for analysis with python in Jupyter Notebook.

4.2 Location Data: Location data of these neighborhoods obtained from geocoder library of google or location data available online.

The latitudes and longitudes of neighborhoods of Pune city are required in order to access the third step of the model preparation. The latitude and longitude (co-ordinates) can be obtained using the geocoder library of google.

An alternative to the above step, if the geocoder library becomes unreliable, is obtaining a geospatial file detailing the location co-ordinates from web directly.

4.3 Venue Data: Data of venues in these neighborhoods, obtained from Foursquare API

Using the location data obtained in second step of model preparation, the Foursquare API is used to obtain the venue data of these neighborhoods. The data is then cleaned and the data pertaining to hospital establishments of each neighborhoods is obtained. The data is then used for machine learning algorithms to perform exploratory analysis in order to obtain results and make inferences.

5. References

1. Proptiger, “Quick Update on Pune Smart City,” [Online]. Available: https://www.proptiger.com/guide/post/quick-update-on-pune-smart-city.
2. Who.int, “World Bank and WHO 2017 report,” [Online]. Available: https://www.who.int/news-room/detail/13-12-2017-world-bank-and-who-half-the-world-lacks-access-to-essential-health-services-100-million-still-pushed-into-extreme-poverty-because-of-health-expenses