

Research Paper on A Novel Approach for Business Store Site Selection

Mayur Shewale

Abstract: *For Entrepreneurs, location plays a major role in determining success. There are many different factors that affect the location selection for a store. Some of the factors are geography, demography, open spaces, weather, population, availability of raw material and many more. The selection of store location is one of the most significant decisions in store-based retailing as good locations are key elements for attracting customers to the outlets. The problem of site selection is a very complex and has multi-criteria decision making. This paper surveys the existing work done in this area and proposes a novel machine learning model based on neighborhood of the localities to solve this problem. The purpose of choosing neighborhood data is as it is more feasible to do market analysis judging by the characteristic of the existing shops in that area. The model is based on the use of an original dataset collected for Pune city scraped from different websites and collected using Foursquare API. The model is scaled to predict location for businesses at higher level abstraction.*

Keywords--Business, location selection, web scraping, decision making, machine learning.

I. INTRODUCTION

With increase in the population, demands are also increasing. As demands are unlimited, and demands become habits and customs this is the reason a huge number of opportunities in the market for business owner are being created on daily basis. This leads to an intense competition. So, business owner needs to consider many different factors to compete the competitors surrounding them. The factors such as geography, demography, open spaces, weather conditions, population, availability of resources and many more are needed to be bought into consideration while starting a new business as they highly affect to the growth of a particular business. Among all other factors mentioned above, location plays a very crucial role within the business. Selection of a

site should be based on a systematic approach as these are the building blocks of a business on basis which the future of that business totally depends. According to our survey many people did research on this topic of appropriate site selection for a store, and different people considered many different factors as mentioned above and used various regression techniques, algorithms and neural networks to get optimal results. This paper presents a method to solve the issue of appropriate site selection for a store, considering neighborhood as a major factor as location decision relates to the entire physical structure of that particular outlet. As we get an idea from the neighborhoods that what are the existing businesses running in that particular

II. LITERATURE SURVEY

The survey was conducted on store locations in different localities of Pune city. For this survey-based research paper our main aim was to study what are the key factors behind start of any new business that is how store locators choose locations to open a store in order to maintain their existence in this competitive world as well as maximize profit at the same time. It is well known that the location selection is vital for successful operation of businesses in stores. Especially in case of retail stores the location plays major role because for retailers' availability of resources at their store location is mandatory. This is the main reason a good site selection strategy is thought to be an effective means to reduce cost and obtain high benefits in business [9]. There are various studies on deriving the best parameters for location selection for opening new business in store at selected location. Regarding location analysis researches are also done for a store in global context [7]. Some of the researches that has been taken in this literature are as follows: In a work by Divaries, Cosmas, Jaravaza in 2013 using trade area analysis proved that different locations have different trade area characteristics [11]. In 2018 another study was completed on site selection of retail shop by Luyao Wang, Hong Fan and Yankun Wang, this study had provided a new method that selects site for opening a new business, which fills the gap in the site selection for small retail shops. The two-step model, including the spatial accessibility estimation process with gravity model and the market potential

evaluation process with BP-PCA model, makes the site selection convincing and near reality [9]. In 2019 a study was conducted by MansiKarna to understand multi criteria problem like how chain stores select most convenient locations, here

Analytical Hierarchy Process was used which considered qualitative as well as quantitative approach in decision making [7]. After going through many different research papers and according to the researches done previously, we came to know that for different location selection many criteria comes into picture and all of them vary from each other. Each of those criteria have unique quality which is somehow required for selecting location of a store. Through this we came up with the fact that a single model cannot explain all such criteria as each them is unique in their own contexts. So, it is very necessary to understand that factors such as demography, geography, population, open spaces, high traffic areas, occupies the major background in the location selection process of store along with business and market strategies.

III. PROPOSED MODEL

After a thorough survey of existing paper, we came up with a novel model for which we planned to do survey on Pune city. To build this model we needed the data of all the localities of Pune city. As the appropriate data was not available on any medium so we came up with an idea to scrape the data from a website i.e. to use 'Web Scraping'. We found the website <https://www.mapsofindia.com/pune/localities/> [1] where a list of all the 96 different localities of Pune were listed. So the data was scraped using an online platform namely 'Parse Hub' <https://www.parsehub.com/> [2] which is a powerful & open source web scraping tool that scrapes the data from any given URL. The data collected was then imported in a 'CSV' i.e. comma separated value format. We used 'Python' a robust

open source programming language which supports multiple libraries and is very efficient to use. Dataset was imported into Python using ‘Pandas’ library for further processing. Our next aim was to get the geographical co-ordinates of all the localities of the Pune city which was done using ‘Geopy’. Geopy is a library which helps to convert an address into a Latitude & Longitude values. Therefore, we retrieved the Latitudes and Longitudes of the localities and also found the Geographical Co-ordinates of the Pune city. Further ‘Folium’ library was used which is a map rendering library in python which provides a leaflet of a map using geographical co-ordinates and latitude, longitude values.

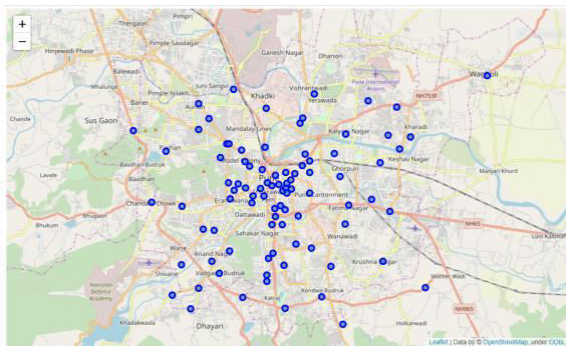


Figure 1:Localities in Pune City

This above image is map of all the localities of Pune city. We have highlighted them using blue coloured dots. Now we used Foursquare application programming interface [3] which explores the neighbor hoods of a particular locality using latitude and longitude based on given radius and given limitations by using GIS (Geographical Information System) tool. GIS tools presents geographical data as this tool is designed to locate all the given points on the earth surface and map all of them according to their positions on the earth surface itself. By using this tool we were able to get the most accurate datapoints. Further by providing a 5000meter radius we got datapoints i.e. all neighbor hoods in the specific radius.

To generate this information, we need to create an API request URL to Foursquare. We proceeded with Data Wrangling and by using ‘One Hot Encoding’ we labelled the data into binary

categories resulting into whether a venue exist or not. Then we proceed with creating a new data frame specially for stores in the locality. The next was to create clusters of multiple stores existing in all localities of Pune city, for this approach we applied an Unsupervised Machine Learning Algorithm Namely K-Means, this algorithm creates clusters for every unique data point based on centroids. Further we used Folium which is a library that supports python. With the help of this library we generated maps showing clusters.

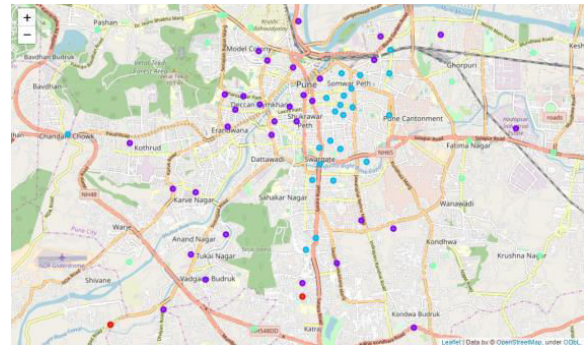


Figure 2: Clustering of Super Markets

The architecture of our model describes the methodology which we used above to propose our model which is as follows:

CONCLUSION

In this particular study we tried to understand the major

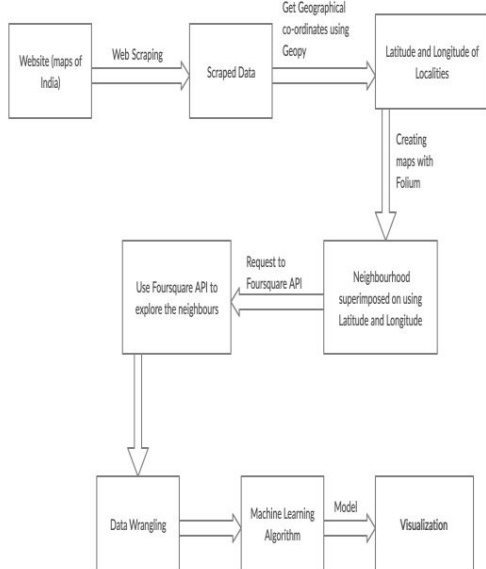


Figure 3: Architecture

role which neighborhood plays in order to select a particular location to open a new store or any new business at any site. As mentioned above we explored neighborhoods of all localities of Pune city listed in the created dataset. With this we came to know that by exploring neighborhoods it was much easier to do market analysis of each and every locality of the city. As this gave us accurate details of what all are the types of businesses currently running in all localities. Through which we got an idea of all the existing competitors of all types of businesses, this made decision making easier with respect to the formation of any new business. In the above proposed model, we formed clusters by taking a specific venue category that is Supermarkets into consideration. In all 5 clusters were formed the highest was of 4 and the lowest was 0. Our future scope is to include population

which we will collect from census data that will take place in the year 2021 into our model, so as to propose a model which will answer to questions such as “Which type of business should be opened and in which area?” and also “At which location should I start my business so that I will earn more profit?”. We hereby conclude that location has a major roleplaying and is one of the major factors on which a business is dependent. A great location affects the overall business in a long-term which provides consistency and stability to the business in terms of market profitability and brings faster marketability to specific elements in the business in short term, whereas a location which is bad will result in an unsuccessful business. Hence location plays a major impact and would result in whether a business in a particular neighborhood will be profitable or not.

REFERENCES

- [1] Maps of India <https://www.mapsofindia.com/>
- [2] Parsehub <https://www.parsehub.com/>
- [3] Foursquare <https://foursquare.com/>
- [4] Google Maps <https://www.google.com/maps/>
- [5] Hikmit Erbiyik, “Retail store location selection problem with multiple analytical hierarchy process of decision making an application in Turkey”, Published by Elsevier Ltd., The 8th International Strategic Management Conference, 2012
- [6] Dongdong Ge and Luhui Hu, “Intelligent site selection for bricks and mortar stores”, Published in Modern Supply Chain Research and Applications. Published by Emerald Publishing Limited, 2019. <http://creativecommons.org/licenses/by/4.0/legalcode>
- [7] Mansi Karna and Anusha Rai, “A study on selection of Location by

retail chain: Big Mart”, Published in International Journal of

Research - *Granthaalayah*, 7(1), 383-395, 2019.

<https://doi.org/10.5281/zenodo.2561093>.

[8] DragoPupavac, “Choice of location for retail businesses” Submitted

to Special Issue on Profit-Driven Analytics January 17, 2018

[9] Luyao Wang, “Site Selection of Retail Shops Based on Spatial

Accessibility and Hybrid BP Neural Network”, Published by

International Journal of Geo-Information, 29 May 2018.

[10] Joshua K, “Retail Site Selection: A New, Innovative Model for

Retail Development”. *Cornell Real Estate Review*, 7(1), 1-26.

Retrieved from <http://scholarship.sha.cornell.edu/crer/vol7/iss1/17>.

[11] DivariesCosmas, “The Role of Store Location in Influencing

Customers’ Store Choice”, Published by Journal of Emerging

Trends in Economics and Management Sciences (JETEMS)

4(3):302-307 (ISSN: 2141-7016).

[12] Rogers, D.S., “Retail location analysis in practice”, *Research*

Review, Vol. 14 No. 2, pp. 73-78, 2017.

[13] DmytroKaramshuk and AnastasiosNoulas, “Geo-Spotting: Mining

Online Location-based Services for Optimal Retail Store

Placement”, Published in University of Cambridge and School of

Mathematical Sciences, Queen Mary University of London, UK.

[14] Shmoys and Tardos, “Approximation algorithms for facility

location problems”, *Proceedings of the 29th Annual ACM*

Symposium on Theory of Computing, pp. 265-274.

[15] Evaluating the Neighbourhood When Choosing a Business Facility,

<https://www.bizfilings.com/toolkit/research-topics/officehr/>

[evaluating-the-neighborhood-when-choosing-a-business-facility](https://www.bizfilings.com/toolkit/research-topics/officehr/evaluating-the-neighborhood-when-choosing-a-business-facility)

[16] Location needs of various business types

<https://www.inc.com/encyclopedia/site-selection.html>

[17] Site Selection Wikipedia

https://en.wikipedia.org/wiki/Site_selection

[18] Choosing a retail store location

[https://www.thebalancesmb.com/choosing-a-retail-store-location-](https://www.thebalancesmb.com/choosing-a-retail-store-location-2890245)

[2890245](https://www.thebalancesmb.com/choosing-a-retail-store-location-2890245)