

# Location of New Store Detection

**Mr. Krupen Sankhe<sup>1</sup>, Miss. Mrunali Shirke<sup>2</sup>, Mr. Yash Shinde<sup>3</sup>, Mr. Sareen Deore<sup>4</sup>**

<sup>1</sup> Student, Dept. Information and Technology, A. C. Patil college of engineering, Maharashtra, India

<sup>2</sup> Student, Dept. Information and Technology, A. C. Patil college of engineering, Maharashtra, India

<sup>3</sup> Student, Dept. Information and Technology, A. C. Patil college of engineering, Maharashtra, India

<sup>4</sup> Professor, Dept. Information and Technology, A. C. Patil college of engineering, Maharashtra, India

\*\*\*

**Abstract** - In the business field, selecting an appropriate location for a new opening store is important when entering an unexplored market. Opening a new store near the city center or the outskirts of town may lead to a huge difference in business profit, which is highly related to a business's success. To find a suitable location for a new store, retailers and business analysts should do lots of research, including surveying the local market, the number of competitors, the cost to invest in a new store, and so on. These business analyses cost a lot of time and money and will result in a huge loss if the new store location is decided improperly. Nowadays, since big data and machine learning techniques have been more developed and maturely applied in various business fields, I decided to bring up a method that helps select the best retail store location by the provided geographic data and linear regression algorithm. This may help business stakeholders and retailers save time and cost to get a better understanding of the local market and help with the better location of new retail store.

**Key Words:** linear regression, stakeholder

## INTRODUCTION

As the competition becomes more and more intense, many retail small store chain operators are eager to know how to evaluate new store locations quantitatively to support a store. Before setting up a new shop, various parameters such as the population of the city, infrastructure, logistics, etc. must take into consideration. Moreover, when a business plans to branch out to its operations in a new city, it also needs to consider these factors to ensure effective expansion. The Location of New Store Detection system has been developed to assist the user with finding the ideal location to open their new store This system requires the user to choose the city where they plan to set up their store from the available data set. Using Logistic regression, this

system then predicts whether a location is ideal for setting up a new store or not. This system has been developed using Django Framework with Python as the programming language.

## PROBLEM STATEMENT

Retailers are the most powerful actors in the distribution channels with proximity to consumers and the potential to create a market. store web site choice may be a strategic call, each in terms of client satisfaction and profit of the company to dynamical market conditions and intense competition. the site choice decision may be a long investment. retail store website choice may be a method that has got to be followed rigorously to pass sooner than the competition. In this method, electionrelated information assortment and scientific analysis, and decisions should be created within the light of these data., in process of taking decisions managers to use scientific methods to provide significant benefits in our  
NEW STORE LOCATION DETECTION SYSTEM helps in talking qualitative and quantitative helps the retailer make the wise guiding him with all the resources for the same.

## LITERATURE SURVEY

Over the period there is deep learning toward data processing and clustering in various application domains. In general, most of the work focuses on clustering similar data. the main aim of this work is to predict the best location for the new store. According to

Timothy Edwardo, (2020 International Conference on Information Technology Systems and Innovation (ICITSI)) the results, it can be concluded that the deep learning technique can be used in data analysis for a crucial business decision-making process, which translates to the opening of a new store in this context. It is proved by increasing percentages of stores that achieve

target prediction, which is 53.14% than the branch office method, which predicts 31.2% of stores that achieve target prediction. The model also predicts 28.32% of potential stores. Factors that have a significant impact on sales are rack size, store age, distance from a competitor, domain location, and store type. accuracy in producing this sales prediction is measured using RMSE, MAE, and MAPE which is 3023293.23,2318618.649, and 0.220483 respectively.

So-Hyun Park Park (2019) shows an analysis program and an analysis method using tracking the Mac address of a visiting visitor. presented the correlation analysis program first and then introduced the preference prediction program. They are programmed and introduce their analysis and prediction methods by introducing their input and output and processing methods.

The authors BaiWen JunYin Dong (Proceedings of the 2008 Winter Simulation Conference) have demonstrated how optimization and simulation techniques can be adopted in retail store sales prediction analysis. We develop a more elaborate analysis framework based on the GIS platform than the traditional gravity model. The technique provides the means of directly evaluating the impact of these factors. Besides the sales prediction of new stores, this framework can also provide an explicit estimate of the impact of new stores on existing competing stores, which is of more value for new site evaluation work of business development people. The real case has proved the effectiveness of the framework. The author Laurence Lin (2020) collected venue location data from Foursquare API as a basis to predict retail store popularity collected from Google Place API. They analyze the relationship between geographic features and prove its effectiveness to predict a retail store's popularity.

According to Emilio Macias (2020) data science, analysis, and machine learning techniques, this project aims to provide a solution for this problem and recommend the best for opening the low-cost supermarket

## SCOPE AND OBJECTIVE

Before opening a store in a new city, the entrepreneur needs to conduct extensive research about the infrastructure, population, logistical challenges they might face, etc. Businessmen seeking to expand their existing business into a new city often face the same troubles with finding the ideal spot that matches these parameters. This Location of New Store Detection helps users with assessing whether their chosen location is suitable for setting up their new stores. This web

application helps users get an idea of whether a given location would be conducive to the growth of a business or not.

## PROPOSED SYSTEM

The proposed system is user-friendly. Communication between client and system is flawless. All the information about the client and their demands is stored in a database. System and database are provided with high security. The proposed System consists following modules.

### Module 1: Signup

Users can at first signup with the use of email wherein email acts as username, the password will be sent over to the user's email address. As the user enters details at signup, they get added into the system's database which includes their username.

### Module 2: Login

Users can log in with credentials of email as username and the password sent on the email address as password respectively. These details are cross-checked with the database to verify and let the user access the system if the user has signed up. If the user hasn't signed up yet they are redirected to the signup page. If details are present in the system and the username and password used by the user are correct, then they are granted access to the system.

### Module 3: Input Data

The system asks for the location name as input. The user must provide the name of the city in which he wishes to open the new store.

### Module 4: Forget Password

If the user forgets their password or fails to find a password, they can have their password resent from this page. The user enters their registered email address which is verified if present in the database and if present, a new password will be sent.

### Module 5: View map

The user will click this button and will see a real-time map and will enter the city name where he wants to open the shop.

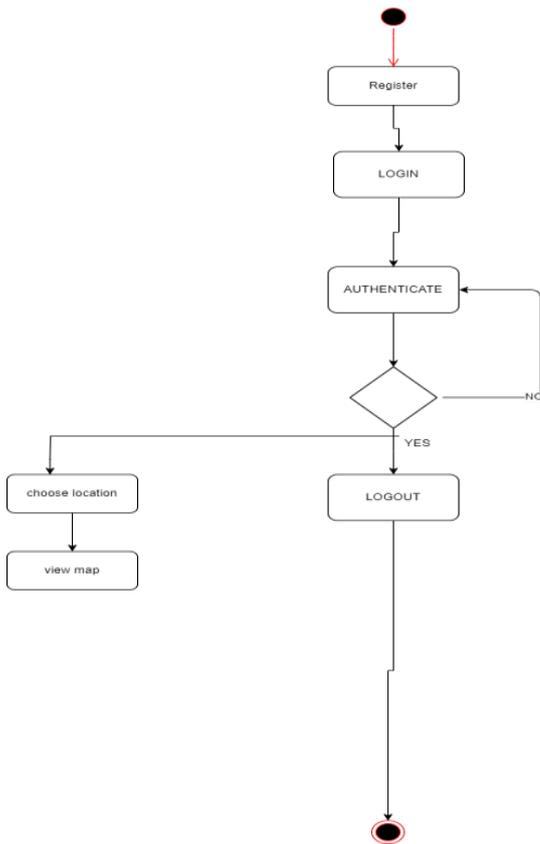


Fig 1: Block Diagram

The main aim of this project is to develop a system for giving the best location to open a new shop which will reduce the risk of opening a new shop without any knowledge of that location.

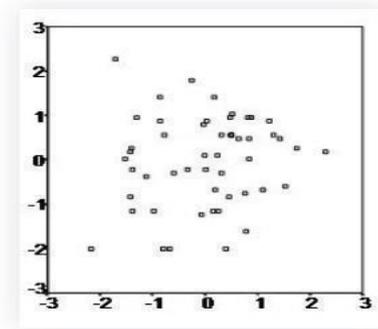
The application uses various libraries of python – pandas, NumPy, sci-kit learn, pickle and uses fundamental concepts of Machine Learning with Python 3. x for Data Pre-processing, Feature Engineering, Graphical Representations, and the use of Supervised Algorithm – Linear Regression to train and test the dataset for the model. After the model file is created, we generate the Django and webpages with the help of HTML5 and CSS, to provide with better User Interface and User Experience.

**ALGORITHM**

Linear Regression is thought of as a Machine Learning algorithmic program that enables us to map numeric inputs to numeric outputs, fitting a line into the data points.

Working:

Step1: analyzing the correlation and directivity of the information The first step allows the research worker to



formulate the model, i.e., that variable X encompasses a causal influence on variable Y and their relationship is linear.

Step 2: The second step of regression analysis is to suit the regression line. Mathematically least square estimation is employed to attenuate the unexplained residual. the basic plan behind this idea is illustrated within the following graph.

Step 3: measure the validity and quality of the model. When we fit a line through the scatter plot, the regression curve represents the estimated job satisfaction for a given age. However, the important observation may not fall precisely on the curve. we try and justify the scatter plot with an equation of  $y = b_0 + b_1x$ . the gap between the regression curve and the data point represents the unexplained variation, which is additionally known as the residual  $e_i$ . The method of least squares is employed to minimize the residual.

$$\sum e_i^2 = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - b_0 - b_1x_i)^2 \rightarrow \min \Rightarrow \hat{y}_i = b_0 + b_1x_i$$

The results of this equation would as an example be  $y_i = 1 + 0.1 * x_i$ . The last step for the linear regression analysis is the test of significance. linear regression uses 2 tests to check whether or not the found model and therefore the estimated coefficients are found within the general population the sample was drawn from.

Step 4: Our model is prepared.

**RESULTS AND DISCUSSION**

Considering the dataset, we have used for the model; we find the relationship amongst the various columns of the dataset trying to establish a relationship amongst them. On finding the most reliable data, we train and test the model on different algorithms to find precision,

accuracy, and recall scores. Here we found that amongst

Logistic Regression, linear regression, Decision Tree Classifier, Random Forest Classifier, and Support Vector Classifier linear regression gives the best results without any overfitting or underfitting issues. we have an accuracy of 99% in training data and an accuracy of 97% in testing data. These conditions negate both overfitting and underfitting issues and hence we get the best possible and accurate outcome as target (output) from the features (input).

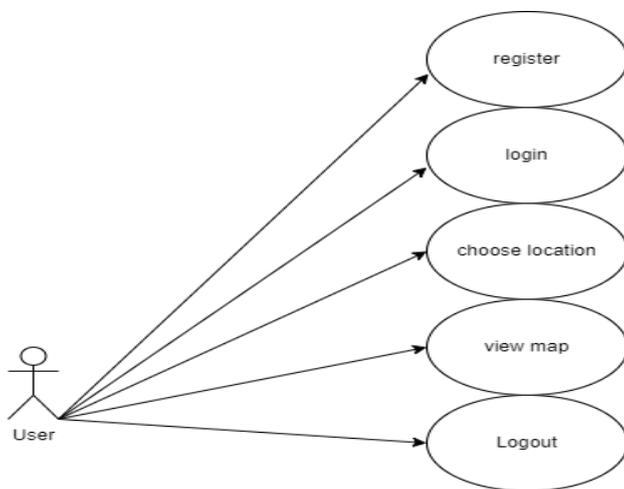


Fig 2: Use case Diagram

## CONCLUSIONS

This was our project of System design regarding the Location of new Store Detection developed in Python programming language. the development of this technique takes plenty of effort from us. we predict this system gave a great deal of satisfaction to us. though each task is not the time same to be perfect during this development field even additional improvement could also be attainable during this application. we have a tendency to learn numerous things and gained a lot of information regarding the development field.

## ACKNOWLEDGEMENT

We would like to acknowledge our college A. C. Patil College of Engineering, Principal, Head of Department, and management for the guidance, encouragement, and co-operation throughout the project.

## REFERENCES

1. So-Hyun Park Park Analysis using customer's Actual Data and Prediction Method Of store Preference .
2. K-Y Lin and J.J. quot : A deep Learning -Based Customer Forecasting Tool, &qout : IEEE Second International Conference On Multimedia Big Data, PP.198205,2016.
3. J. Han, M., and J.Pei, ; Data Mining: Concepts and techniques, ; Massachusetts, Morgan Kaufmann, 2011.
4. Tova, prediction Analysis in Determining New Marketing Store, 2020
5. Bai Wen Jun, Yin Jin Dong S Based Sales Forecasting on Retail Small Stores” 2008 Winter Simulation Conference.