

# Customer Based Mapping of Real Estate Using Data Mining Techniques

<sup>1</sup>Harshvadan Rode, <sup>2</sup>Akshay Vardekar, <sup>3</sup>Atharva Patil, <sup>4</sup>Prof. Sejal D'Mello

<sup>1</sup>Student, <sup>2</sup>Student, <sup>3</sup>Student, <sup>4</sup>Professor  
Department of Information Technology,  
Atharva College of Engineering, Mumbai, India

**Abstract:** Now-a-days people looking to buy a new house tend to be more cautious and careful with their budgets and market policies. The existing system involves computation of house prices without the crucial prediction about future market tendency and rate increase. Aim of this project was to develop a real estate web application using Microsoft ASP.NET and SQL 2008. The real estate system permit the functionality for users, allowing them to search for properties by features, pricing or address. It also provides functionality for the seller vendor by authorizing them to log into the system and add and include new advertisements or delete and remove the existing ones. For this each user is provided with a login account and password. Along with this, when the user will quest for the property, initial property price and predicted property price will be displayed. By examining previous market trends and price sets, and also upcoming developments future prices will be predicted. For the price value prediction we will be applying classification algorithm. The functioning this project comprises a website which accepts customer's specifications and then operates on the application of data mining. This application will aid customers to lend in an estate without approaching an agent or broker. It also reduces the hazards involved in the transaction.

**Index Terms:** Data mining, house price forecasting, prediction, linear regression, real estate

## 1. Introduction

A colossal amount of knowledge has been assembled on text mining for Information Retrieval (IR) since a long time. Using automated text mining algorithms to devise knowledge from natural language texts provides diverse challenges but also offer exclusive possibilities. Natural language texts is one of the most common and natural forms of saving information. This can be easily perceived by a human but it is a great challenge for machines to acquire meaning from this data input. However, machines do bid an important asset over human capacity: computing power. This implies that computers can detect patterns, which are non-trivial frequencies, within data faster, authentic and more precise than their human counterpart, which can only be done if the structure design of the data is known. Natural language does consist implicit grammatical structure, but these structures are deeply convoluted and spread across different languages.

## 2. Related Work: To determine Customer Interest in Real Estate with the help of Data Mining Algorithms

With a huge volume of amorphous assets and documents, the Housing industry has gained a high competition in business world. The data mining algorithms in this housing industry provides an additional weightage to the developers by processing those data, predicting further price trends in future and thus guiding the real estate customers to make proper knowledge-driven decisions. In this paper, the main attention is on data mining techniques and its fundamentals to build a model which not only recognizes the most appropriate locality for a customer according to their requirements, and also ranks them. This is used to forecast a suitable location by ranking methodology. It analyses a set of locations selected by the customer.

### 2.1. Drawbacks

It does not forecast the future rates of the real estate property mentioned by the customer. Because of this, the risk in investing money in a house increases. To reduce this error, customers keen to approach a real estate broker which again increments the price of the process. This leads to the transformation of the current system.

## 3. Proposed System

Everything manual nowadays is transforming into digital automated systems. The main aim of this project is to forecast the real estate prices so as to reduce the issues faced by the customer willing to buy real estate property. The current scenario is that the customer meets a real estate broker to manage his/her money and then the broker put forward for consideration some appropriate estates for his financial investments. But this method is risky as the broker may put forward for consideration wrong real estate properties and thus resulting to loss of the customer's money. The manual method which is currently used in the market is out dated and has high risk. So as to overcome this fault, there is a need for an updated and automated system. Data mining techniques can be used to help customers to invest in a proper real estate according to customer's mentioned needs. Also the new system will be efficient with respect to cost and time. The administrator will add features of real estate properties into the system and depending

on this features the system will predict the estimated price with the help of data mining techniques. When user searches property, the list of property will be displayed to the customer along with the forecasted price.

#### 4. Linear Regression:

Regression is a statistical measure used for forecasting. It discovers the bonding firmness between dependent variable (known as label attribute) and other changing variables also called as independent variables (known as regular attribute). Regression provides with continuous value of the dependent variable that is beneficial for forecasting.

Mathematically, the linear regression formula is given by,

$$y = a + bx$$

Where:

$$a = \frac{\sum y - b \sum x}{n}$$

$$b = \frac{n \sum (xy) - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$

Linear regression method is used to forecast the relationship (line) amongst data points. There can be various (linear or nonlinear) ways to describe the relationship. In the linear model, it is dependent on the intercept and the slope.

#### 5. Working of the System

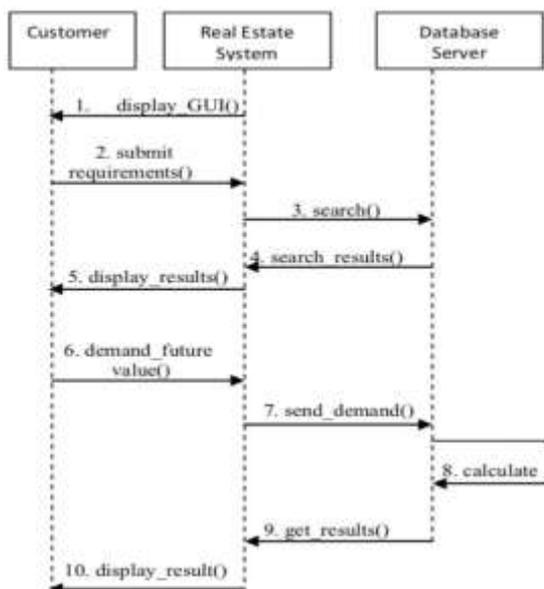


Fig. Working of the system

The diagram above explains the working of the Real Estate Customer Mapping system. The proposed system is supposed to be a website with 2 objects: the web interface for the customer which will predict the value of a real estate property the customer wishes to purchase and the Database server. The database server also includes the computational mechanism described in the algorithm. When the customer first login into the website they are displayed with an interface where in they can enter their real estate requirements such as the type of house, the location of the real estate property etc. A data index searching then provides with customer desired outputs consisting of matching properties. Now, if the customer wants to have a look at that particular real estate property price in future they can enter the date of future wherein they wish to see the price of that property. The system will identify the date and categorize it in the quarters. The algorithm then will determine the value of price and give the output in the form of predicted price back to the customer.

#### 6. Future Scope

Real estate is a growing factor and there is bright future scope to it. User can give input to find desired residence suitable to his needs and will get the ideal search result best compatible to him. User just needs to take a right advice from the experts in the field which the website will provide after analysis. A major future scope is adding estate database of more cities which will give the users to search more estates and reach a precise decision.

## 7. Conclusion

The today's real estate world has become very difficult to store such huge data and extract them for using in different purposes like business requirements. The system makes use of such data in the most efficient way. Data mining Algorithm is optimally used by the systems. Thus such Data mining algorithm provides important tools to fulfil customers needs by increasing the accuracy of estate choice and hence reducing the risk of investing in an estate. Many new and added features can be developed to make these systems accepted worldwide. A factor like recession that affect the house prices can also be added. All these details and information will support the desired system to run on a larger level.

## 8. Acknowledgements

We are using this golden opportunity to express our gratitude to one and all who have supported us throughout this research. We are very much thankful for the aspiring guidance, invaluable constructive criticisms and friendly advice during the process of conducting research. We are also thankful for the support and encouragement of members of the management of Atharva College of Engineering, Malad West. We are highly indebted to Head of Department of Information Technology, Prof. Nileema Pathak and our guide Prof. Sejal D'Mello for their guidance and constant supervision as well as for providing necessary information regarding the research & also for their support and co-operation in completing the project on research. We are highly obliged to the laboratory technicians who have given their time, energy and attention for their invaluable support.

## References

- [1] Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, 2015. Introduction to Linear Regression Analysis
- [2] Byeonghwa Park, Jae Kwon Bae (2015). Using machine learning algorithms for housing price prediction, Volume 42, Pages 2928-2934
- [3] Real Estate Price Prediction with Regression and Classification, CS 229 Autumn 2016 Project Final Report
- [4] <http://www.99acres.com/property-rates-and-pricetrends-in-mumbai>
- [5] Robin A Dubin 1998 Predicting Housing Prices using Multiple Listings Data Journal of Real Estate Finance and Economics vol.17(1) pp 35-59

