

**HOUSING: PRICE PREDICTION**

**Submitted By**

**Praveena S**

**Acknowledgement**

* B N Lashmi, G H Raghunandhan

A conceptual overview of data mining

2011 National Conference on Innovations in

Emerging Technology

* R Manjula

Real estate value prediction using multivariate regression models

Materials Science Engineering Conference Series, volume 263, issue 4

* A Varma

House Price Prediciton using Machine Learning and Neural Networks

2018 Second International Conference on Inventive communications and computational Technologies, p.1936

* Arietta Sean M

City forensics Using visual elements to predict non-visual city attributes

IEEE transactions on visualization and computer graphics, volume 20, p.2624 - 2633

**Introduction**

Accurately estimating the value of real estate is an important problem for many staeholders including house owners, house buyers, agenrs, creditors and investors. It is also a difficult one. Though it is common knowledge that factors such as the size, number of rooms and location affect the price, there are many other things at play. Additionally prices are sensitive to changes in market demand and the peculiarities of each situation, such as when a property needs to be urgently sold.

The sales price of a property can be predicted in various ways, but is often based on regression techniques. All regression techniques essentially involve one or more predictor variables as input and a single target variable as output.

In this paper, we compare different machine learning methods performance in predicting the selling price of hours based on a number of features such as the area, the number of bed and bathrooms and the geographical position.

**Analytical Problem Framing**

In this paper, the data is selected from US datasets, which can be downloaded from the internet. Housing value of Boston suburb can be measured through the data of 13 features. These features include per capita crime rate by town, proportion of nonretail business acres per town and index of accessibility to radial highways.

Housing value of Boston suburb is analyzed and forecast by linear regression and gradient boosting regression. After getting rid of missing samples from original dataset. 400 samples are treated as training data and 52 samples are treated as test data.

**Models\Development and Evaluation**

In this paper we choose the type of machine learning prediction that is suitable to our problem. We eant to determine if this is regression problem or a classification problem. In this project we want to predict the price of a house given information about it. The price we want to predict is a continuous value, it can be any real number. This can be seen by looking at the target variable in our dataset Sale Price.

That means that the prediction type that is a appropriate to our problem is regression.

Linear Regression models the relationship between the target variable and the independent variables. It fits a linear with coefficients to the data in order to minimize the residual sum of squares between the target variable in the dataset and the predicted values by the linear approximation.

**Conclusion**

The research question for this study is to study how well house prices can be predicted by using Linear Regression and Gradient Boosting Regressor.