

Assignment – 1

Title - Implement Boston housing price prediction problem.

Problem Statement - Implement Boston housing price prediction problem by Linear regression using Deep Neural network. Use Boston House price prediction dataset.

Objective -

To build a regression model to predict the price of houses.

Outcome -

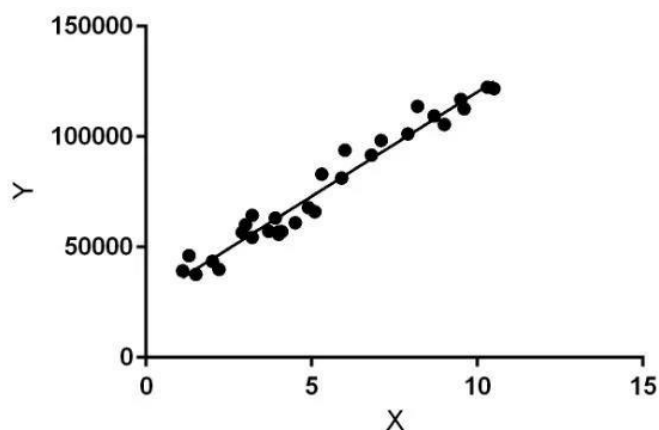
To understand the exploratory data analysis, split the training and testing data, Model Evaluation and Prediction by the linear regression on the Boston housing dataset.

Theory -

Linear Regression:

Linear Regression is a Supervised Machine Learning Model for finding the relationship between independent variables and dependent variable. Linear regression performs the task to predict the response (dependent) variable value (y) based on a given (independent) explanatory variable (x). So, this regression technique finds out a linear relationship between x (input) and y (output).

For example, if we have a dataset consisting of apartments properties and their prices in some specific area, Linear Regression algorithm can be used to find a mathematical function which will try to estimate the value of different apartment (outside of the dataset), based on its attributes.



Neural Network:

The basic unit of the brain is known as a neuron, there are approximately 86 billion neurons in our nervous system which are connected to 10^{14} - 10^{15} synapses. Each neuron receives a signal from the synapses and gives output after processing the signal. This idea is drawn from the brain to build a neural network.

Each neuron performs a dot product between the inputs and weights, adds biases, applies an activation function, and gives out the outputs. When a large number of neurons are present together to give out a large number of outputs, it forms a neural layer. Finally, multiple layers combine to form a neural network.

Algorithm:

1. **Import Libraries:** Install the required libraries and setup for the environment for the assignment. importing SciKit-Learn, Pandas, Seaborn, Matplotlib and Numpy.
2. **Importing Data and Checking out:** As data is in the CSV file, we will read the CSV using pandas read_csv function and check the first 5 rows of the data frame using head().
3. **Exploratory Data Analysis for House Price Prediction :** create some simple plot for visualizing the data.
4. **Get Data Ready For Training a Linear Regression Model:** now begin to train out the regression model. We will need to first split up our data into an X list that contains the features to train on, and a y list with the target variable, in this case, the Price column. We will ignore the Address column because it only has text which is not useful for linear regression modeling.
5. **Split Data into Train, Test:** Now split our dataset into a training set and testing set using sklearn train_test_split(). The training set will be used for training the model and testing set for testing the model. We are creating a split of 40% training data and 60% of the training set.
6. X_train and y_train contain data for the training model. X_test and y_test contain data for the testing model. X and y are features and target variable names.
7. **Creating and Training the LinearRegression Model:** import and create sklearn linear_model LinearRegression object and fit the training dataset in it.
8. **LinearRegression Model Evaluation:** Now evaluate the model by checking out its coefficients and how we can interpret them
9. **Predictions from our Linear Regression Model:** Find out the predictions of our test set and see how well it performs.
10. In the scatter plot, we see data is in a line form, which means our model has done good predictions.
11. **Regression Evaluation Metrics:** Here are three common evaluation metrics for regression problems.

Conclusion -

We have analyzed a Linear Regression Model which we help the real estate agent for estimating the house price.