**Introduction to AI and ML**

**CSL236**

Project Report



Faculty name:Dr. Ruchika lalit Student name: Mudit Kumar

Roll No.: 20CSU322

Semester: 5th

Group:

**Department of Computer Science and Engineering**

**The NorthCap University, Gurugram- 122001, India**

**Session 2021-22**

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **S.No** |  | **Page No.** |
| **1.** | **Project Description** |  |
| **2.** | **Problem Statement** |  |
| **3.** | **Analysis**  **3.1 Hardware Requirements**  **3.2 Software Requirements** |  |
| **4.** | **Design**  **4.1 Data/Input Output Description:**  **4.2 Algorithmic Approach / Algorithm / DFD / ER diagram/Program Steps** |  |
| **5.** | **Implementation and Testing (stage/module wise)** |  |
| **6.** | **Output (Screenshots)** |  |
| **7.** | **Conclusion and Future Scope** |  |

**Project Description**

This project presents a method for the Prediction of house price of the Boston. I constructed three 3 regression models according to their accuracy and one Neural network . Then, prediction models are used to predict the . Python and Supervised Machine Learning Classification algorithms are used to construct these prediction models

**Problem Statement**

A real-state company wants to buy and sell the property at a cheaper rate so they decide to use machine learning. You are going to use various statistical analysis tools to build the best model to predict the value of a given house. Your task is to find the best price your client can sell their house at. The best guess from a model is one that best generalizes the data.

**We are going to cover the following steps:**

1. Load the Dataset (Import libraries and Load dataset)
2. Analyze Data (Descriptive Statistics, Unimodal and Multi-model Data Visualizations)
3. Validation Dataset
4. Evaluate Algorithms: Baseline
5. Evaluate Algorithms: Standardize Data
6. Algorithm Tuning
7. Ensemble Methods
8. Finalize Model
9. Summary
10. Reference

Goal:  Create a regression model that is able to accurately estimate the price of the house given the features.

**Analysis**

**3.1 Hardware Requirements**

We require one laptop

**3.2 Software Requirements**

Jupyter notebook,Tensorflow version 2.7 and above ,Nvidia-cuda

**Design**

**4.1 Data/Input Output Description:**

In this project we are taking various features such as age, tax ,crim etc as input which is processed with to predict perfect output Which is price of the house

**4.2 Algorithmic Approach / Algorithm / DFD / ER diagram/Program Steps**

**Steps of Proposed framework**

Broadly in physical world or realistic issues, there is no curb over the types of data. Some

dire pre-processing like removal of missing values, feature selection, etc. are always

required. Machine learning focuses on taking up contemporary techniques to process huge

amount of complex data with lower expense. The abstract view of proposed framework

has been represented in Figure 1. Figure 1 describes the framework of the prediction model

created to predict the price of house using 13 features,

processed by 3 different regressor models and 4 different neural network models , which give outputs with an acceptable

accuracy and precision percentage.

i Pre processing: Missing values are removed by replacing them by mean value

imputation.

ii Feature Selection: Feature are selected on the basis of their correlation with the Target(price)/labels

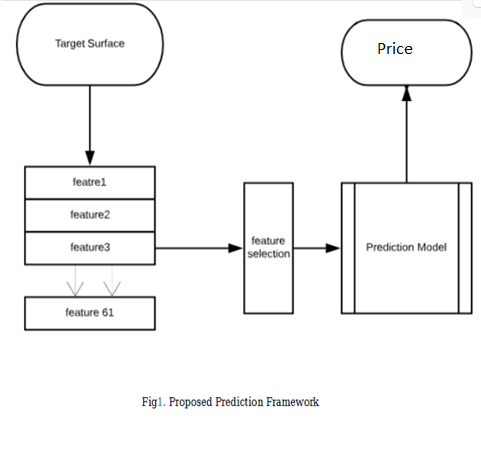
iii Prediction Model: Different ML classifiers are explored and implemented to find the best

possible solution. Tensorflow Neural network, being an neural network model has shown the highest

performance with 2.5% of Root\_mean\_squared\_error.

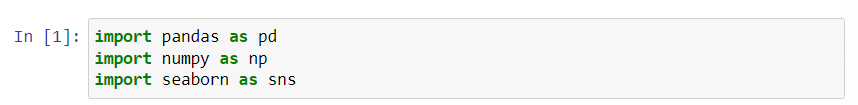
The outcome of this proposed framework helps to

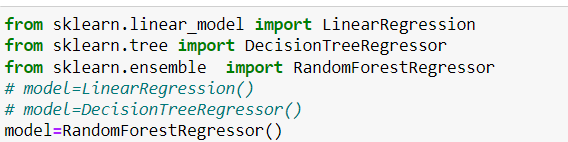
predict the price of the house .

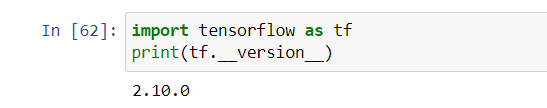


**Implementation and Testing (stage/module wise)**

**Importing the Dependencies**

****

****

****

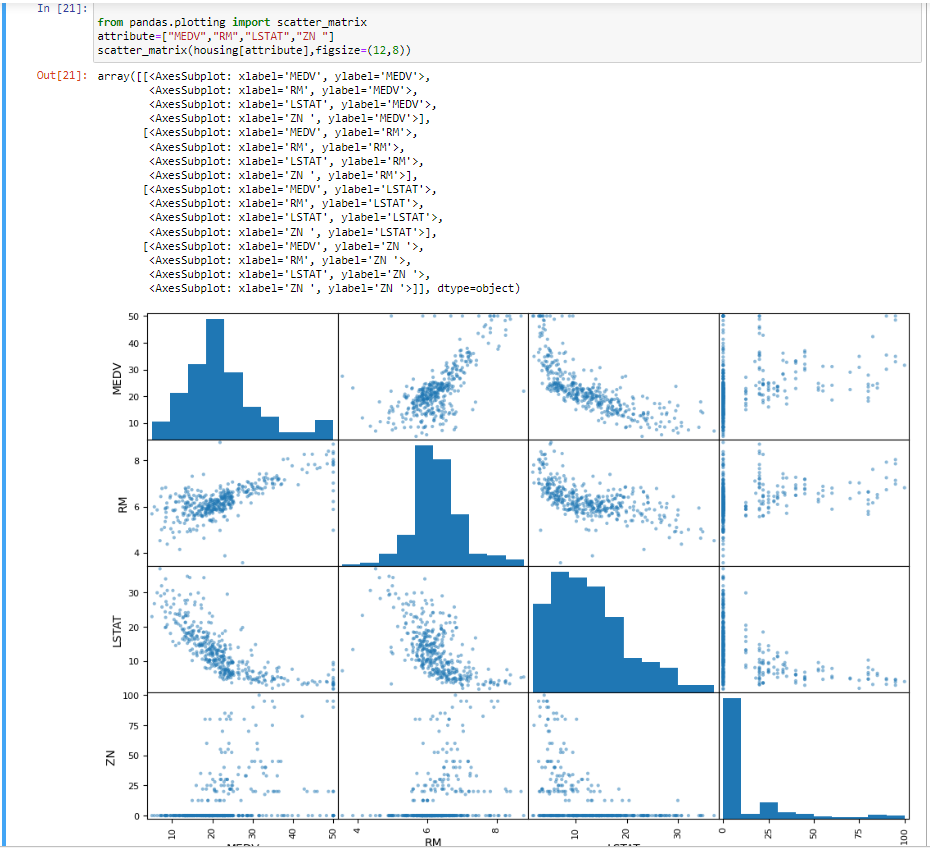
# Data Collection and Data Processing

# 

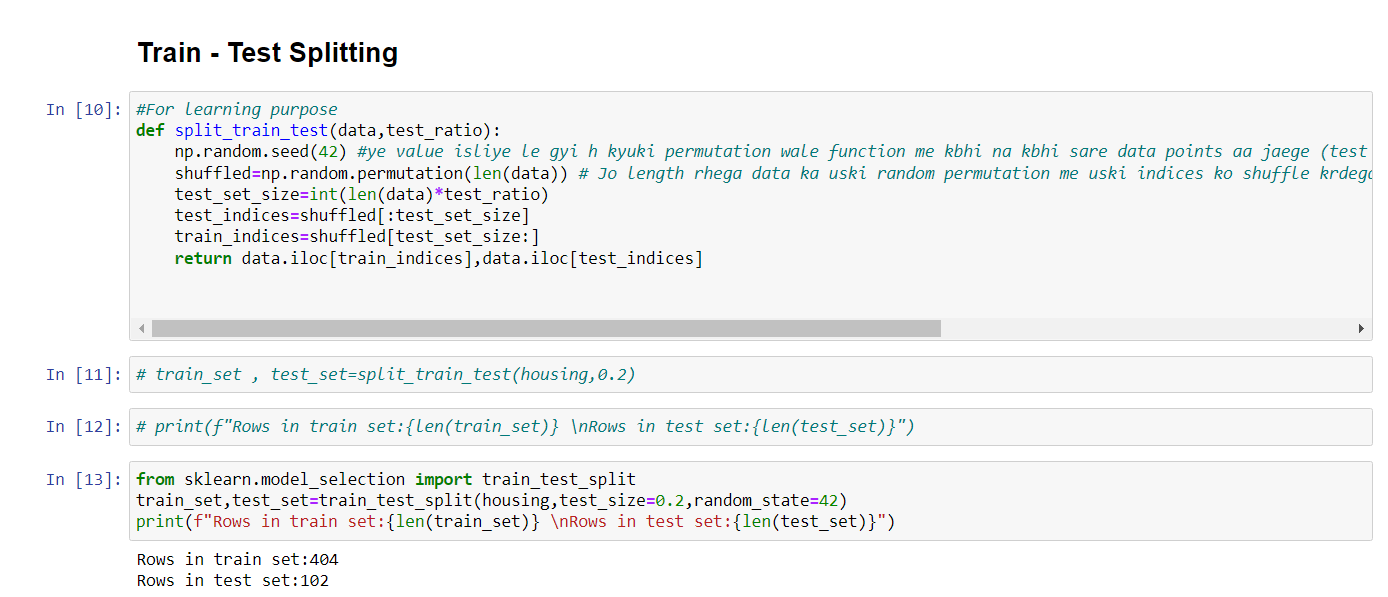
# 

# 

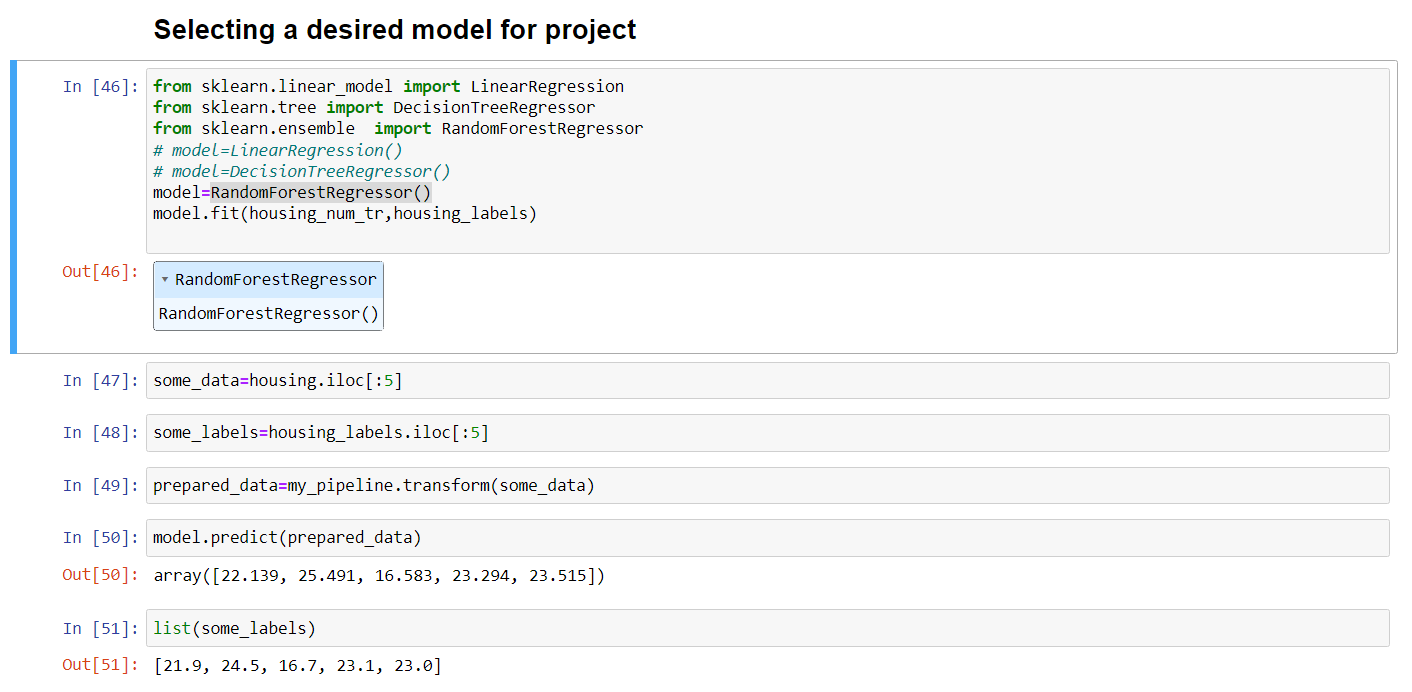
# 

****

# Training and Test data

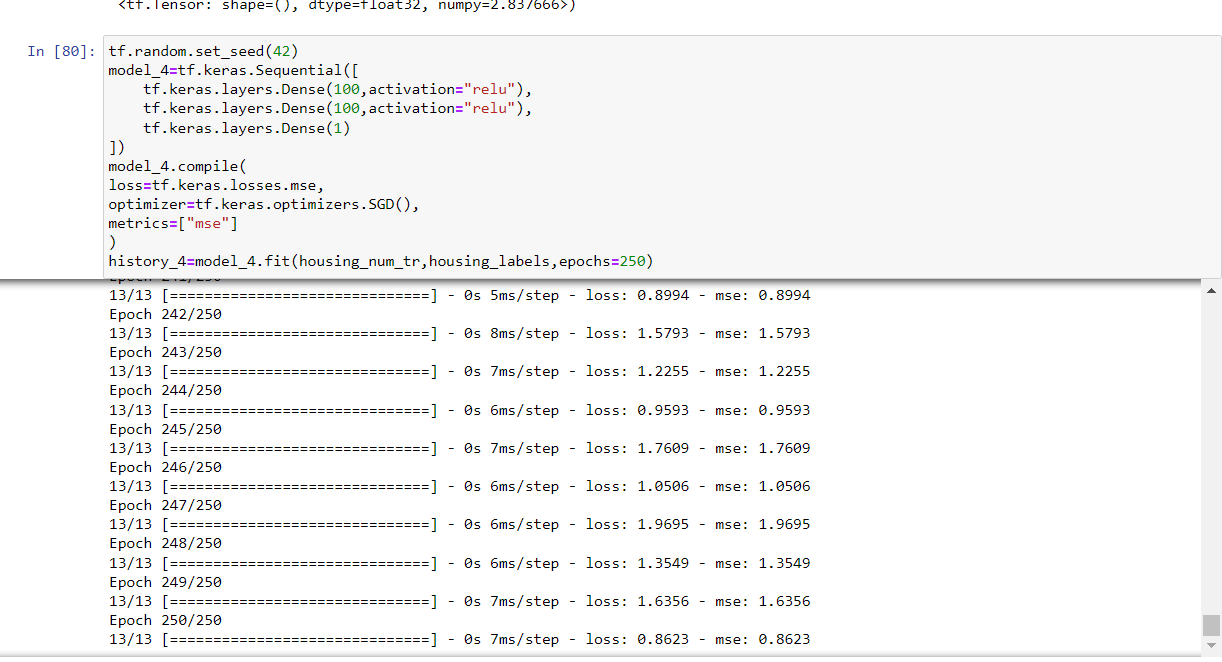
****

# Model Training \_1 --> RandomForestRegressor()

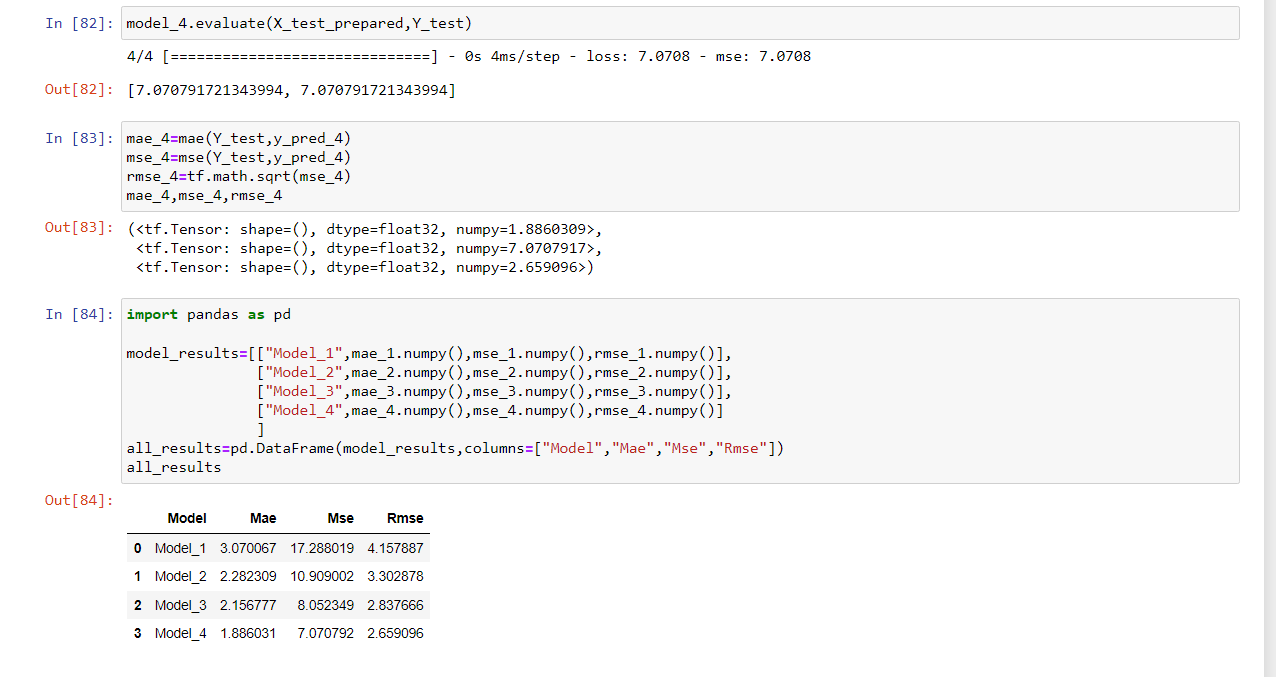
****

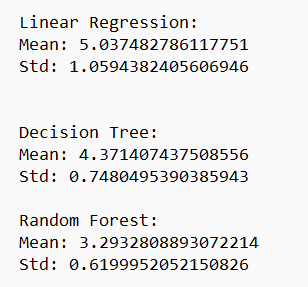
# Model Evaluation\_1

# Model\_2-Neural network using Tensorflow

****

**Model\_2\_Evaluation**

****

****

**Conclusion**

From the Exploratory Data Analysis, we could generate insight from the data. How each of the features relates to the target. Also, it can be seen from the evaluation of three models that Random Forest Regressor performed better than Linear Regression. But however our tweaked neural network model is able to perform much better in the end

**References used: 1.https://archive.ics.uci.edu/ml/machine-learning-databases/housing/housing.names - for dataset**

**2. https://www.tensorflow.org/**