Assignment#03_SVR_DT_RF_Regression

Dataset Description:

This dataset is a Housing dataset where each record in the dataset shows information about a Boston suburb. There are 13 features in the dataset. These features include factors such as crime rate, average number of rooms, proximity to employment centers, property tax rates, and pollution levels. The objective is to predict the median value of homes in $1000s (MEDV)using these input features.

Independent Variables/Features:

The independent variables used in this polynomial regression are:

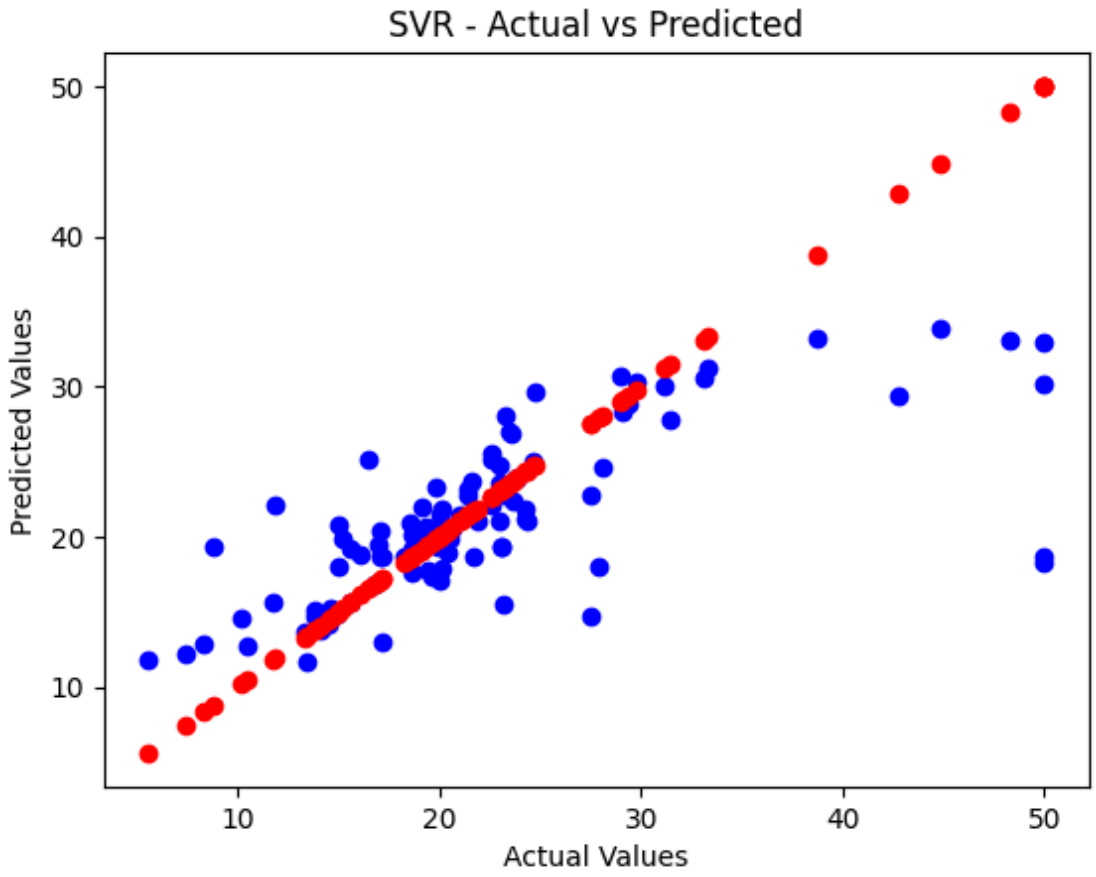
CRIM: Per capita crime rate by town.
ZN: Proportion of residential land zoned for large lots.
INDUS: Proportion of non-retail business acres per town.
CHAS: Charles River dummy variable (1 if tract bounds river; 0 otherwise).
NOX: Nitrogen oxide concentration (parts per 10 million).
RM: Average number of rooms per dwelling.
AGE: Proportion of owner-occupied units built before 1940.
DIS: Weighted distances to employment centers.
RAD: Index of accessibility to radial highways.
TAX: Full-value property tax rate per $10,000.
PTRATIO: Pupil-teacher ratio by town.
B: Proportion of people of African American descent.
LSTAT: Percentage of lower status population.

Dependent/Output Variable:

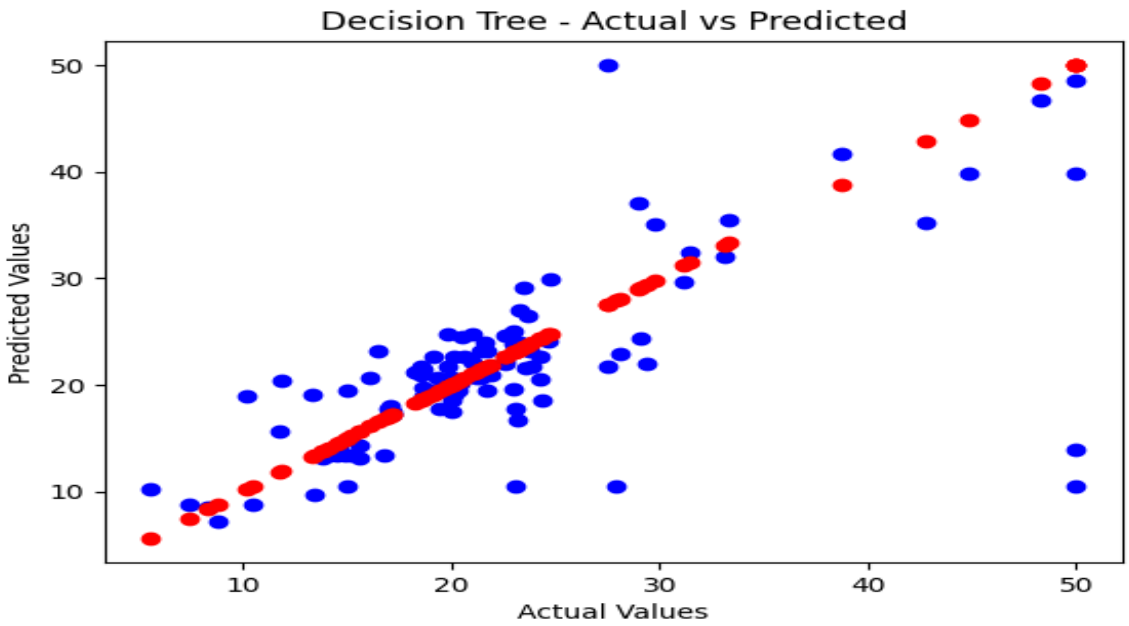
The output variable is the MEDV (median value of owner-occupied homes in $1000s).

Charts for SVR:

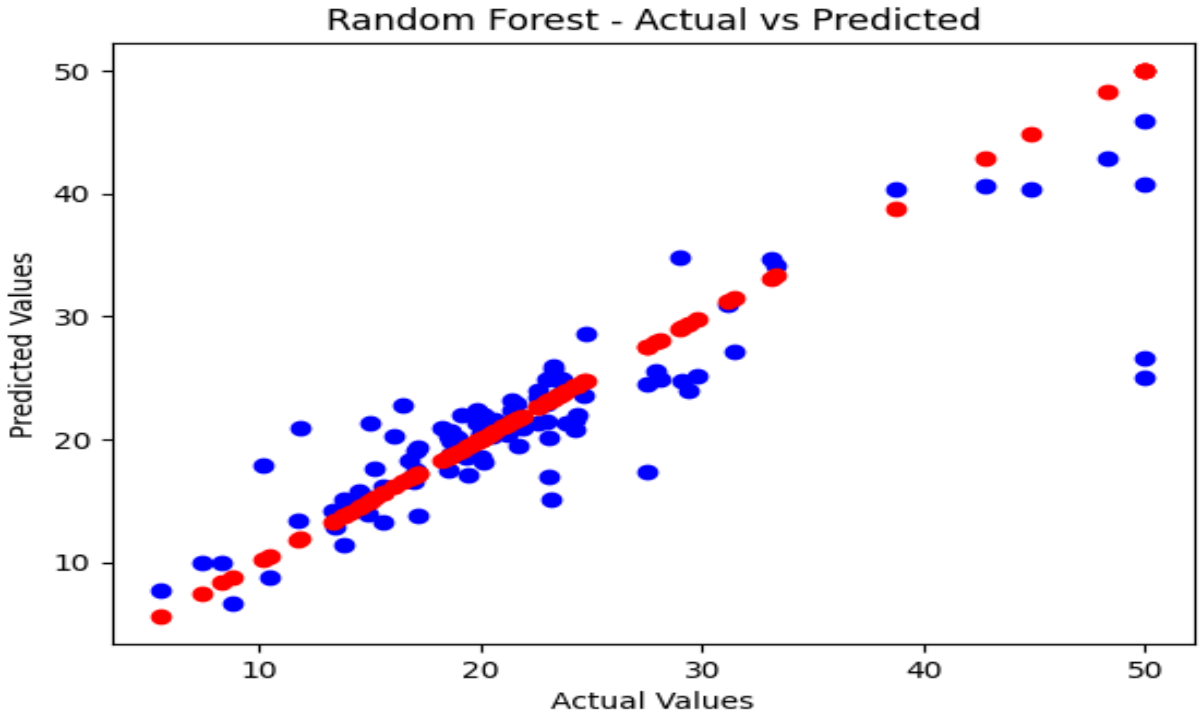
#Red dots are actual values and Blue dots are predicted values



Charts for Decision Tree Regression:



Charts for Random Forest Regression:



Prediction of Values:

[0.02, 20, 5.0, 0, 0.55, 6.0, 50, 3.0, 2, 300, 15, 380, 5] are Input Values for the 13 features from CRM to LSTAT in that order respectively.

SVR Prediction for this input: [26.50341582]
Random Forest Prediction for this input: [21.1]
Decision Tree Prediction for this input: [23.2]