SMAI Assignment-2 Report
Submitted by: Ayush Kumar Dwivedi (2008802002)

Q-1: K Nearest Neighbors:

Command to run the script:

```
python q-1.py
Enter the dataset: 1: Robot1, 2: Robot2, 3: Iris
```

Part-1

1. **Robot1:**

----------Self made kNN model----------
Accuracy: 0.72
Recall: 0.7243589743589745
Precision: 0.7333333333333334
F1 Score: 0.7181964573268922

----------Scikit-Learn kNN model for verification----------
Accuracy_sk: 0.72
Recall_sk: 0.72
Precision_sk: 0.7461538461538462
F1 Score_sk: 0.7227053140096619

2. **Robot2:**

----------Self made kNN model----------
Accuracy: 0.92
Recall: 0.9198717948717949
Precision: 0.9198717948717949
F1 Score: 0.9198717948717949

----------Scikit-Learn kNN model for verification----------
Accuracy_sk: 0.92
Recall_sk: 0.92
Precision_sk: 0.92
F1 Score_sk: 0.92

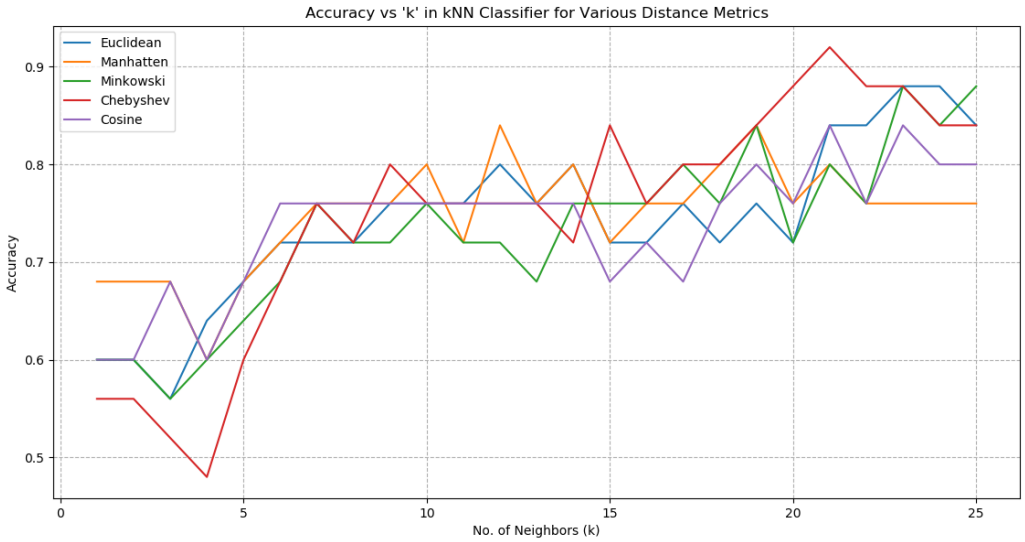
3. **Iris:**

----------Self made kNN model----------
Accuracy: 0.9629629629629629
Recall: 0.9583333333333334
Precision: 0.9696969696969697
F1 Score: 0.9619047619047619

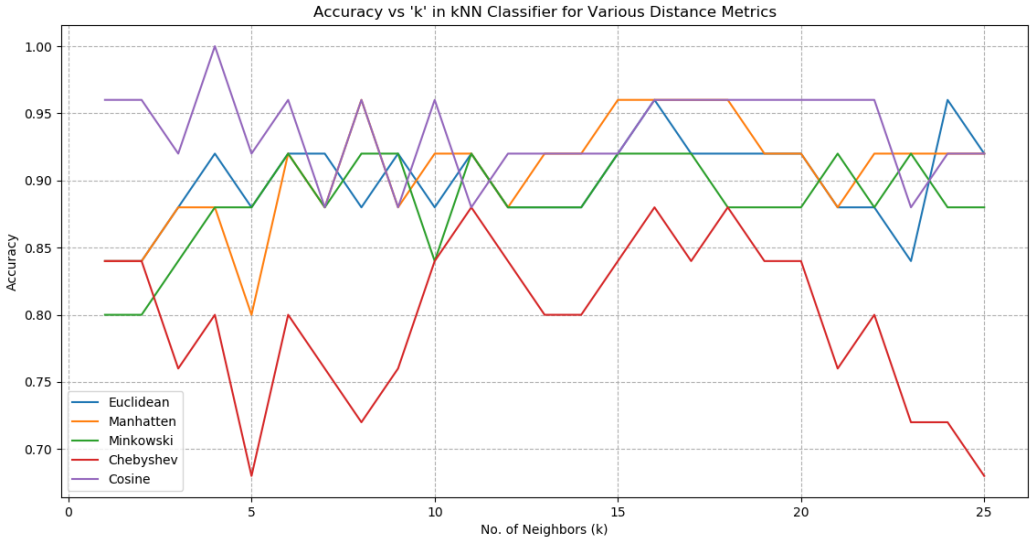
----------Scikit-Learn kNN model for verification----------
Accuracy_sk: 0.9629629629629629
Recall_sk: 0.9629629629629629
Precision_sk: 0.9675925925925926
F1 Score_sk: 0.9633156966490299

Part-2 : Accuracy vs K for all three dataset

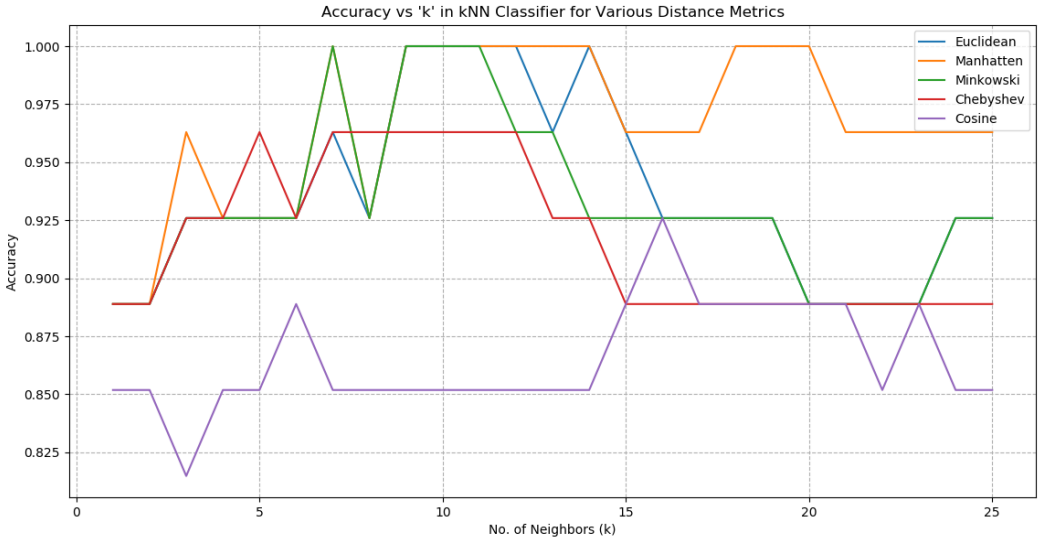
Robot1:



Robot2:



Iris:



Q-2 : Naïve Bayes Classifier

Validation Accuracy: 0.8955555555555555

Q-3 : Naïve Bayes Classifier

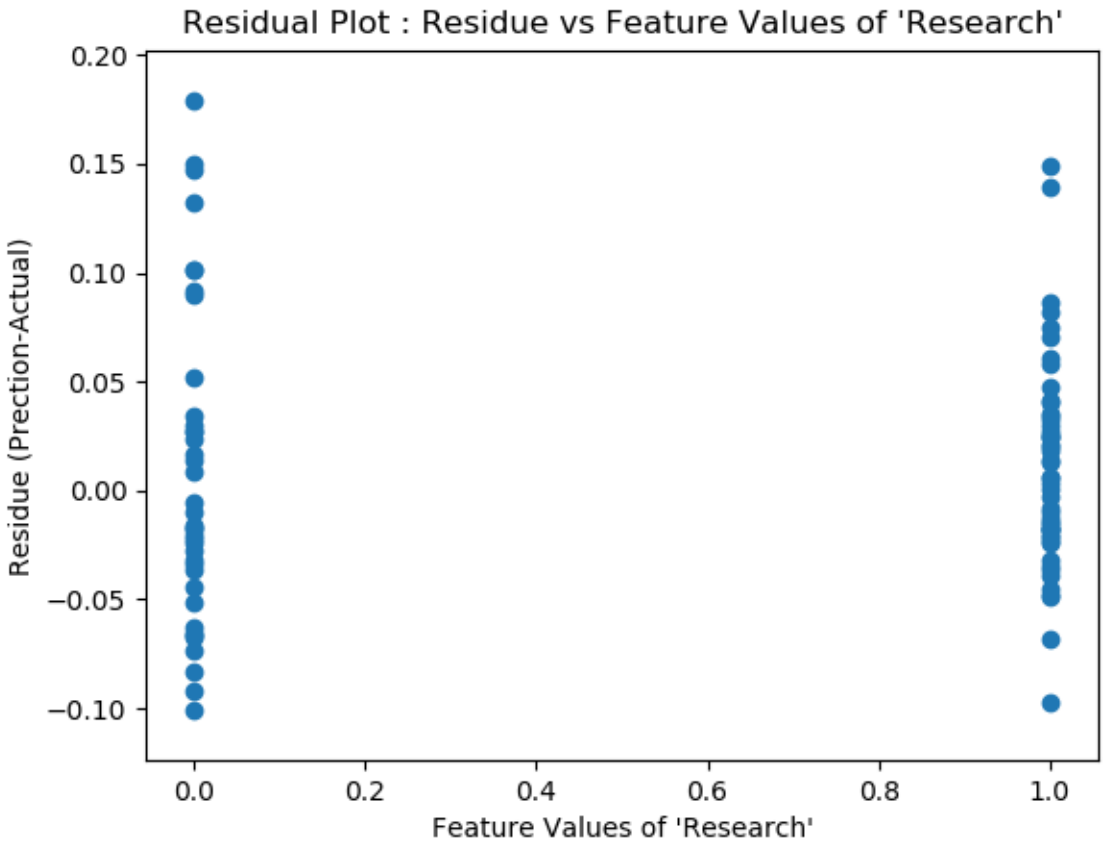
Part 2:

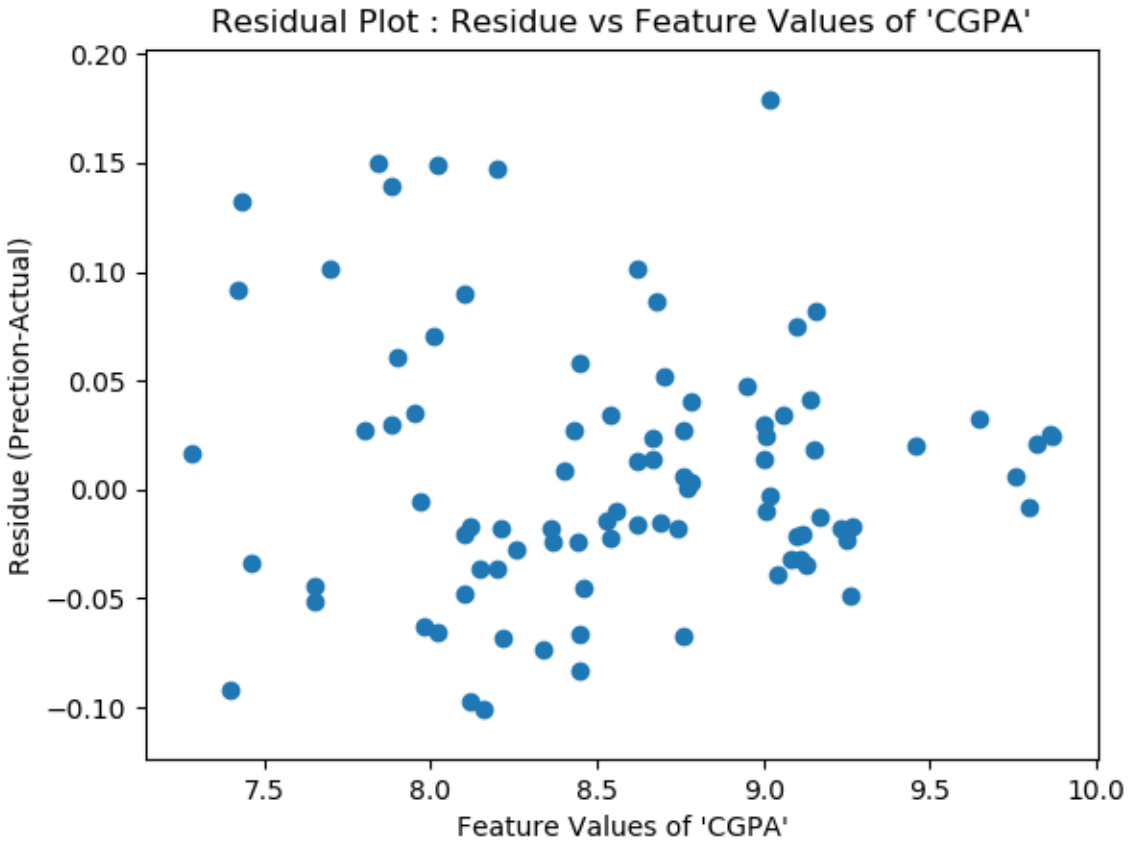
Validation MAE : 0.045025178666509735
Validation MSE : 0.0035022166040901007
Validation MPE : -0.02325964891178607

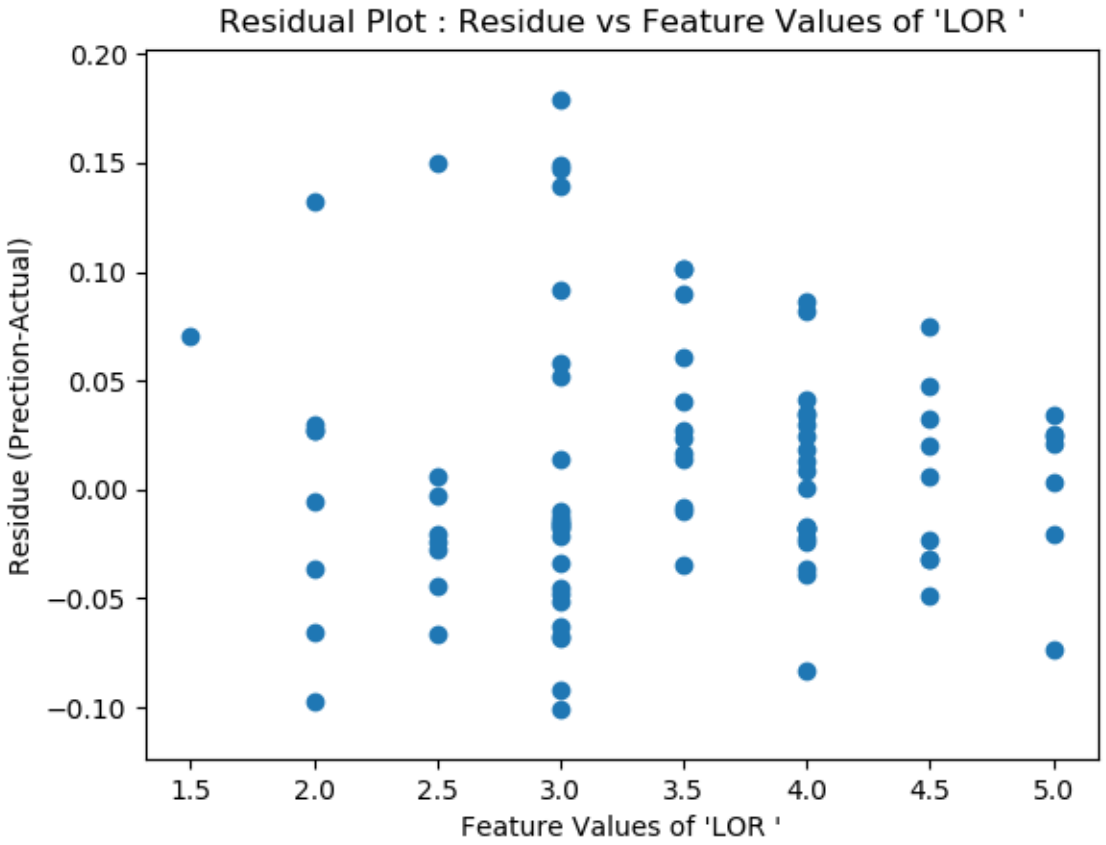
Since MSE is used as the cost function, it has the least error.

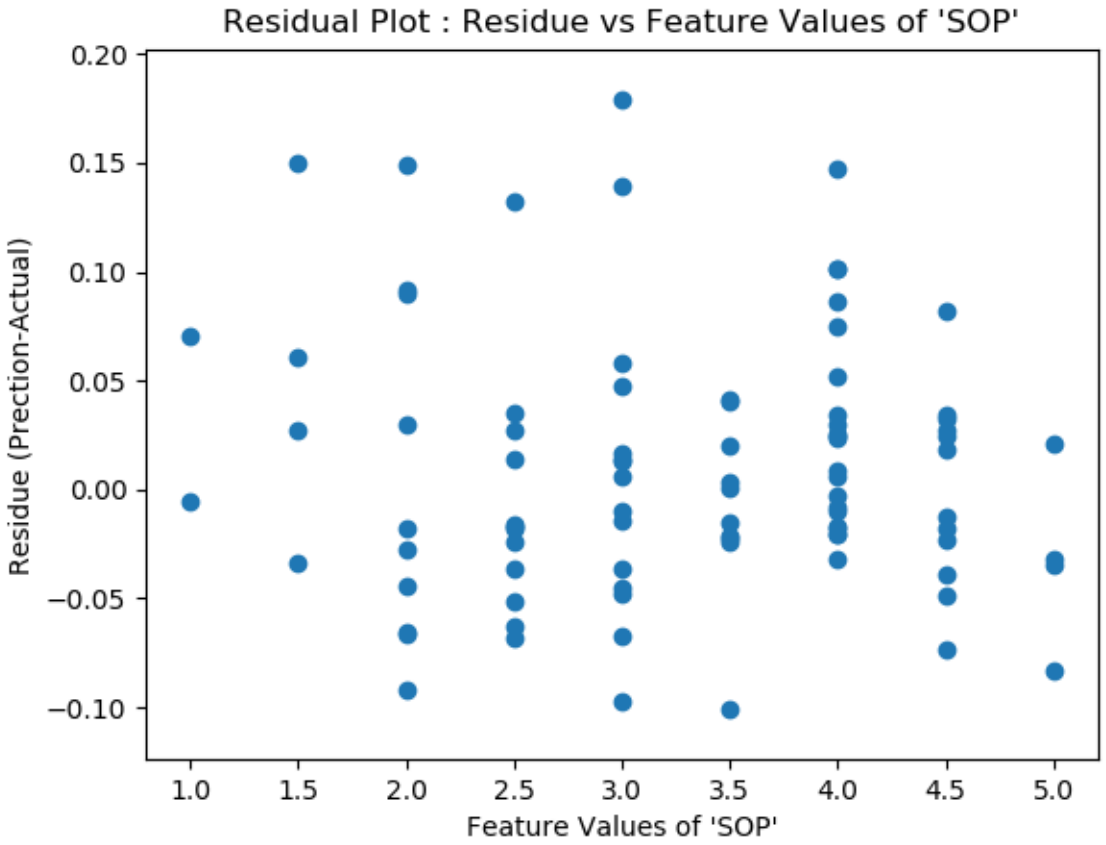
Part 3:

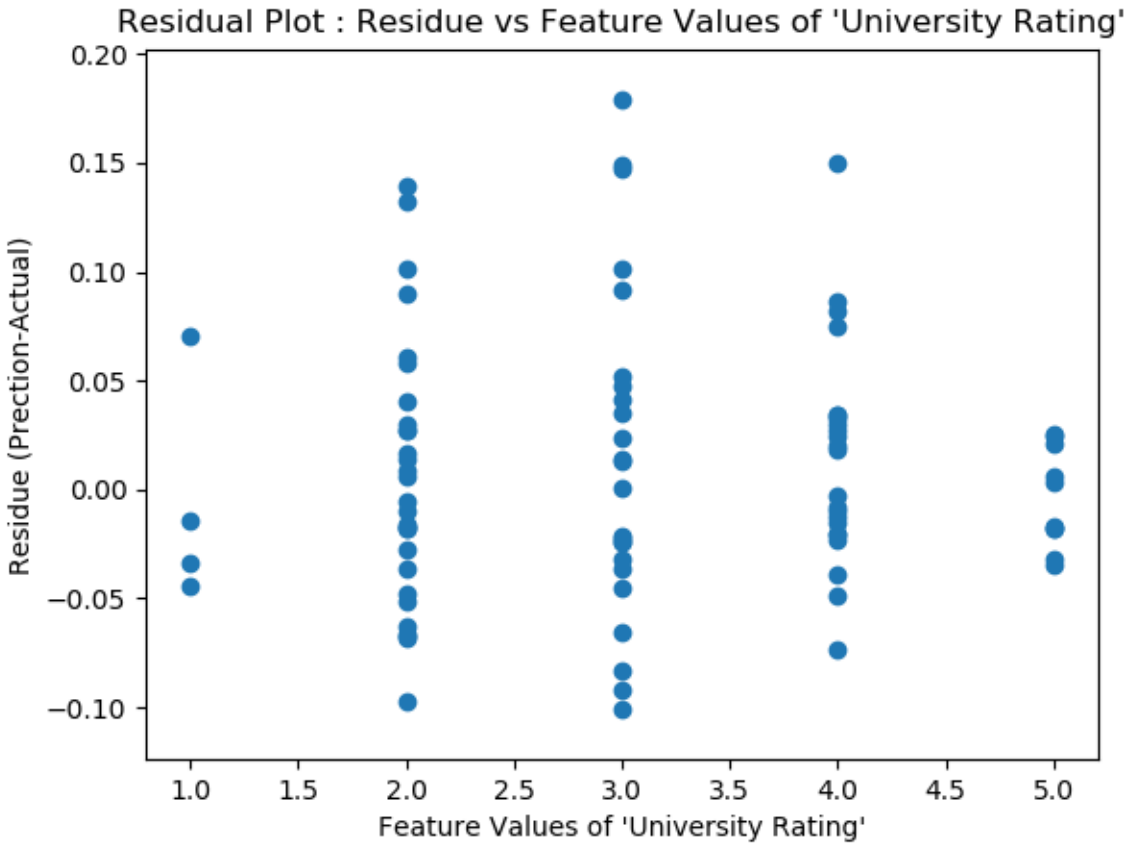
For good model check that the residue are randomly scattered around zero for the entire range of fitted values. When the residuals center on zero, they indicate that the model's predictions are correct on average rather than systematically too high or low. Regression also assumes that the residuals follow a normal distribution and that the degree of scattering is the same for all fitted values.

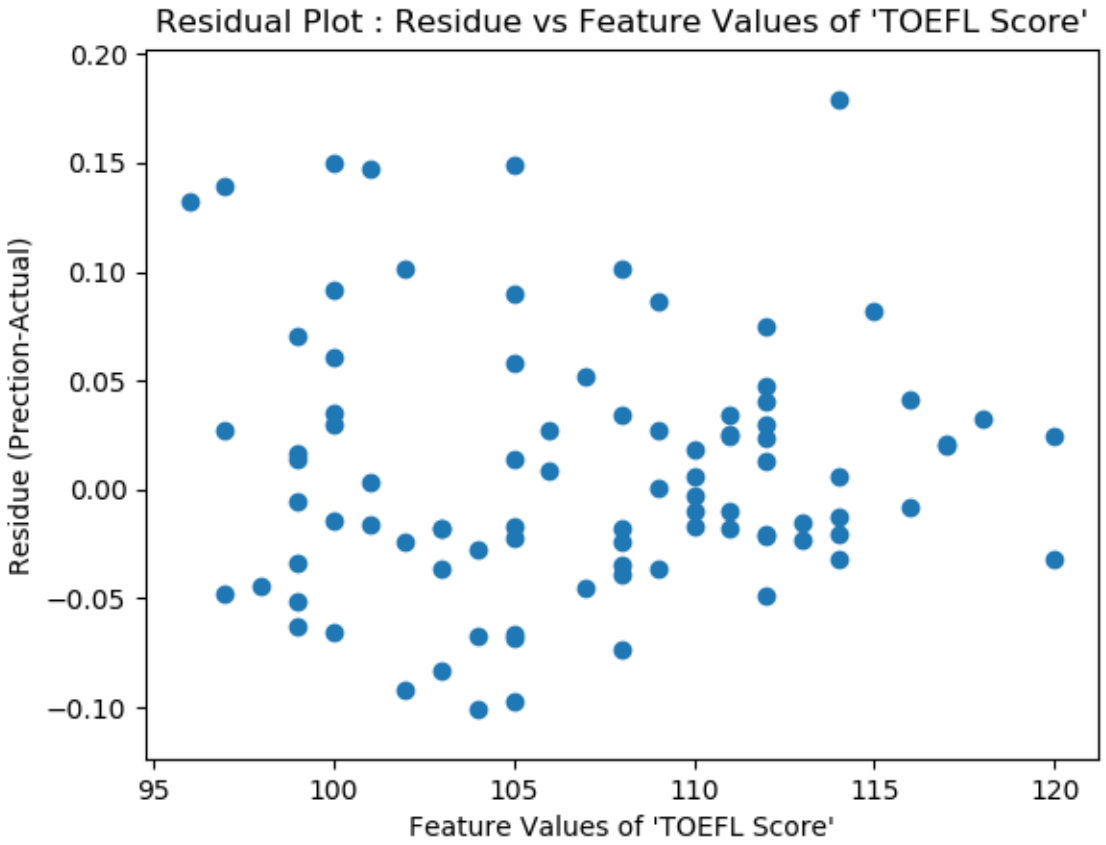

Residual Plot : Residue vs Feature Values of 'Research'
Residue (Prection-Actual)
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
0.0
0.2
0.4
0.6
0.8
1.0
Feature Values of 'Research'

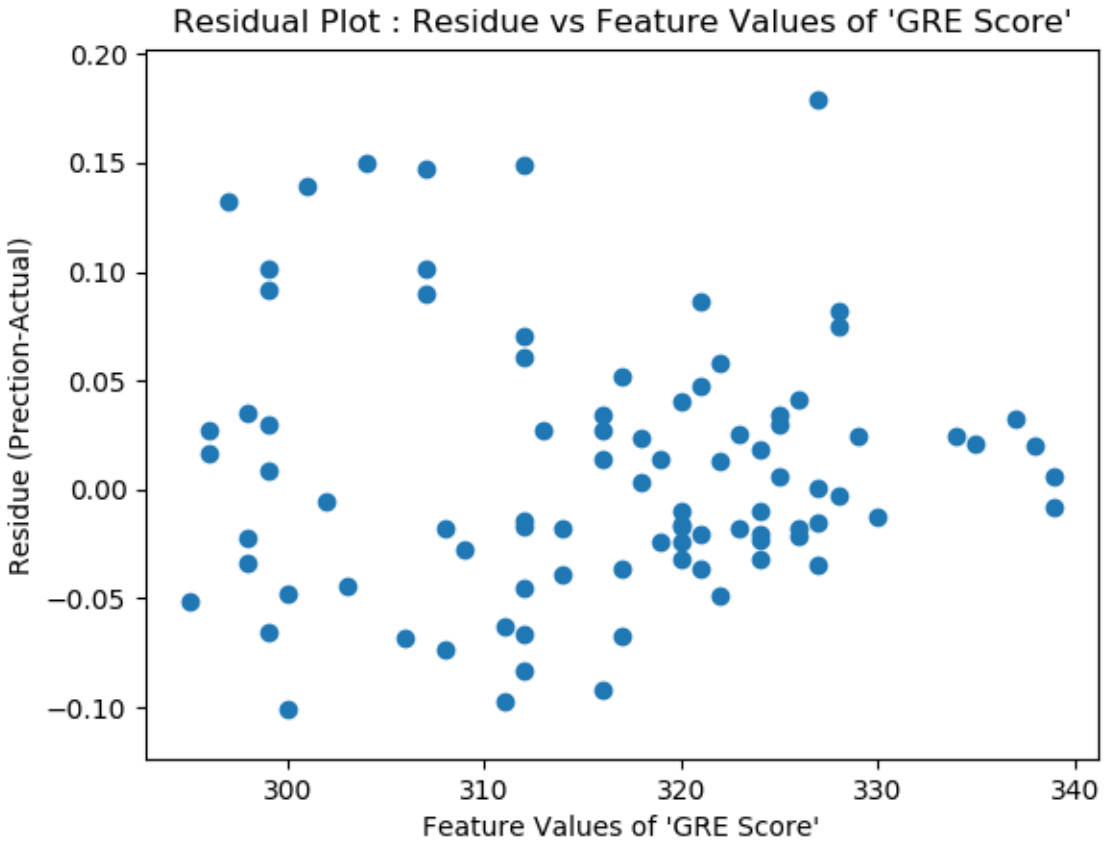

Residual Plot : Residue vs Feature Values of 'CGPA'
Residue (Prection-Actual)
Feature Values of 'CGPA'
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
7.5
8.0
8.5
9.0
9.5
10.0


Residual Plot : Residue vs Feature Values of 'LOR '
Residue (Prection-Actual)
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Feature Values of 'LOR '


Residual Plot : Residue vs Feature Values of 'SOP'
Residue (Prection-Actual)
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Feature Values of 'SOP'


Residual Plot : Residue vs Feature Values of 'University Rating'
Residue (Prection-Actual)
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Feature Values of 'University Rating'


Residual Plot : Residue vs Feature Values of 'TOEFL Score'
Residue (Prection-Actual)
Feature Values of 'TOEFL Score'
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
95
100
105
110
115
120


Residual Plot : Residue vs Feature Values of 'GRE Score'
Residue (Prection-Actual)
Feature Values of 'GRE Score'
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
300
310
320
330
340