

FannyLo_KNNRegressor

April 10, 2023

```
[1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import sys
```

```
[2]: dataset=pd.read_csv('Volumetric_features.csv')
dataset = dataset.drop('S.No', axis = 1)
dataset
```

```
[2]:      Left-Lateral-Ventricle  Left-Inf-Lat-Vent  Left-Cerebellum-White-Matter  \
0                22916.9                982.7                15196.7
1                22953.2                984.5                15289.7
2                23320.4                1062.1                15382.1
3                24360.0                1000.5                14805.4
4                25769.4                1124.4                16331.1
...                ...                ...                ...
4221             27065.6                532.4                12425.1
4222             28408.8                912.7                14024.8
4223             34467.9                1659.6                12744.5
4224             31627.5                1334.4                15883.2
4225             14879.4                704.2                11346.6
```

```
      Left-Cerebellum-Cortex  Left-Thalamus  Left-Caudate  Left-Putamen  \
0                55796.4                6855.5                2956.4                4240.7
1                55778.6                6835.1                3064.2                4498.6
2                55551.2                7566.0                3231.7                4456.2
3                54041.8                8004.6                3137.3                4262.2
4                54108.6                6677.4                2964.4                4204.6
...                ...                ...                ...                ...
4221             51042.9                6354.8                3822.6                4490.5
4222             43103.5                6060.7                3114.2                3731.0
4223             54924.8                6256.7                3573.4                3526.6
4224             57148.2                6982.4                4475.8                4464.4
4225             50468.5                6935.4                3258.5                3751.5
```

```
      Left-Pallidum  3rd-Ventricle  4th-Ventricle  ...  \
```

| | | | | |
|------|--------|--------|--------|-----|
| 0 | 2223.9 | 2034.4 | 1572.5 | ... |
| 1 | 2354.1 | 1927.1 | 1650.5 | ... |
| 2 | 1995.4 | 2064.7 | 1522.1 | ... |
| 3 | 1983.4 | 2017.7 | 1570.3 | ... |
| 4 | 2409.7 | 2251.8 | 1601.1 | ... |
| ... | ... | ... | ... | ... |
| 4221 | 2019.4 | 1256.2 | 2037.6 | ... |
| 4222 | 1937.4 | 1669.9 | 2124.9 | ... |
| 4223 | 2189.9 | 3063.1 | 2511.9 | ... |
| 4224 | 2317.8 | 3809.0 | 3133.5 | ... |
| 4225 | 2226.5 | 1898.4 | 2505.5 | ... |

| | rh_supramarginal_thickness | rh_frontalpole_thickness | \ |
|------|----------------------------|--------------------------|---|
| 0 | 2.408 | 2.629 | |
| 1 | 2.417 | 2.640 | |
| 2 | 2.374 | 2.601 | |
| 3 | 2.366 | 2.639 | |
| 4 | 2.381 | 2.555 | |
| ... | ... | ... | |
| 4221 | 2.505 | 2.666 | |
| 4222 | 2.385 | 3.008 | |
| 4223 | 2.028 | 2.995 | |
| 4224 | 2.491 | 2.865 | |
| 4225 | 2.474 | 3.150 | |

| | rh_temporalpole_thickness | rh_transversetemporal_thickness | \ |
|------|---------------------------|---------------------------------|---|
| 0 | 3.519 | 2.009 | |
| 1 | 3.488 | 2.111 | |
| 2 | 3.342 | 2.146 | |
| 3 | 3.361 | 2.056 | |
| 4 | 3.450 | 2.052 | |
| ... | ... | ... | |
| 4221 | 2.915 | 2.243 | |
| 4222 | 3.572 | 2.040 | |
| 4223 | 3.706 | 1.928 | |
| 4224 | 3.456 | 2.317 | |
| 4225 | 3.691 | 2.337 | |

| | rh_insula_thickness | rh_MeanThickness_thickness | BrainSegVolNotVent.2 | \ |
|------|---------------------|----------------------------|----------------------|---|
| 0 | 2.825 | 2.33635 | 1093846 | |
| 1 | 2.720 | 2.34202 | 1099876 | |
| 2 | 2.684 | 2.31982 | 1097999 | |
| 3 | 2.700 | 2.29215 | 1070117 | |
| 4 | 2.574 | 2.30397 | 1075926 | |
| ... | ... | ... | ... | |
| 4221 | 2.683 | 2.29264 | 1108782 | |
| 4222 | 2.866 | 2.30156 | 960586 | |

| | | | |
|------|-------|---------|---------|
| 4223 | 2.610 | 2.19622 | 1033357 |
| 4224 | 2.900 | 2.43580 | 1073339 |
| 4225 | 2.787 | 2.43420 | 992086 |

| | eTIV.1 | Age | dataset |
|------|-------------|-----|---------|
| 0 | 1619602.965 | 85 | 1 |
| 1 | 1624755.130 | 85 | 1 |
| 2 | 1622609.518 | 86 | 1 |
| 3 | 1583854.236 | 87 | 1 |
| 4 | 1617375.362 | 89 | 1 |
| ... | ... | ... | ... |
| 4221 | 1561822.106 | 79 | 9 |
| 4222 | 1530179.480 | 79 | 9 |
| 4223 | 1604323.353 | 84 | 9 |
| 4224 | 1620891.799 | 80 | 9 |
| 4225 | 1513076.040 | 86 | 9 |

[4226 rows x 140 columns]

```
[3]: y=dataset['Age']
      X=dataset.loc[:,dataset.columns!="Age"]
      X
```

| [3]: | Left-Lateral-Ventricle | Left-Inf-Lat-Vent | Left-Cerebellum-White-Matter | \ |
|------|------------------------|-------------------|------------------------------|---|
| 0 | 22916.9 | 982.7 | 15196.7 | |
| 1 | 22953.2 | 984.5 | 15289.7 | |
| 2 | 23320.4 | 1062.1 | 15382.1 | |
| 3 | 24360.0 | 1000.5 | 14805.4 | |
| 4 | 25769.4 | 1124.4 | 16331.1 | |
| ... | ... | ... | ... | |
| 4221 | 27065.6 | 532.4 | 12425.1 | |
| 4222 | 28408.8 | 912.7 | 14024.8 | |
| 4223 | 34467.9 | 1659.6 | 12744.5 | |
| 4224 | 31627.5 | 1334.4 | 15883.2 | |
| 4225 | 14879.4 | 704.2 | 11346.6 | |

| | Left-Cerebellum-Cortex | Left-Thalamus | Left-Caudate | Left-Putamen | \ |
|------|------------------------|---------------|--------------|--------------|---|
| 0 | 55796.4 | 6855.5 | 2956.4 | 4240.7 | |
| 1 | 55778.6 | 6835.1 | 3064.2 | 4498.6 | |
| 2 | 55551.2 | 7566.0 | 3231.7 | 4456.2 | |
| 3 | 54041.8 | 8004.6 | 3137.3 | 4262.2 | |
| 4 | 54108.6 | 6677.4 | 2964.4 | 4204.6 | |
| ... | ... | ... | ... | ... | |
| 4221 | 51042.9 | 6354.8 | 3822.6 | 4490.5 | |
| 4222 | 43103.5 | 6060.7 | 3114.2 | 3731.0 | |
| 4223 | 54924.8 | 6256.7 | 3573.4 | 3526.6 | |
| 4224 | 57148.2 | 6982.4 | 4475.8 | 4464.4 | |

| | | | | |
|------|---------|--------|--------|--------|
| 4225 | 50468.5 | 6935.4 | 3258.5 | 3751.5 |
|------|---------|--------|--------|--------|

| | Left-Pallidum | 3rd-Ventricle | 4th-Ventricle | ... | \ |
|------|---------------|---------------|---------------|-----|---|
| 0 | 2223.9 | 2034.4 | 1572.5 | ... | |
| 1 | 2354.1 | 1927.1 | 1650.5 | ... | |
| 2 | 1995.4 | 2064.7 | 1522.1 | ... | |
| 3 | 1983.4 | 2017.7 | 1570.3 | ... | |
| 4 | 2409.7 | 2251.8 | 1601.1 | ... | |
| ... | ... | ... | ... | ... | |
| 4221 | 2019.4 | 1256.2 | 2037.6 | ... | |
| 4222 | 1937.4 | 1669.9 | 2124.9 | ... | |
| 4223 | 2189.9 | 3063.1 | 2511.9 | ... | |
| 4224 | 2317.8 | 3809.0 | 3133.5 | ... | |
| 4225 | 2226.5 | 1898.4 | 2505.5 | ... | |

| | rh_superiortemporal_thickness | rh_supramarginal_thickness | \ |
|------|-------------------------------|----------------------------|---|
| 0 | 2.648 | 2.408 | |
| 1 | 2.660 | 2.417 | |
| 2 | 2.597 | 2.374 | |
| 3 | 2.604 | 2.366 | |
| 4 | 2.597 | 2.381 | |
| ... | ... | ... | |
| 4221 | 2.457 | 2.505 | |
| 4222 | 2.497 | 2.385 | |
| 4223 | 2.407 | 2.028 | |
| 4224 | 2.700 | 2.491 | |
| 4225 | 2.746 | 2.474 | |

| | rh_frontalpole_thickness | rh_temporalpole_thickness | \ |
|------|--------------------------|---------------------------|---|
| 0 | 2.629 | 3.519 | |
| 1 | 2.640 | 3.488 | |
| 2 | 2.601 | 3.342 | |
| 3 | 2.639 | 3.361 | |
| 4 | 2.555 | 3.450 | |
| ... | ... | ... | |
| 4221 | 2.666 | 2.915 | |
| 4222 | 3.008 | 3.572 | |
| 4223 | 2.995 | 3.706 | |
| 4224 | 2.865 | 3.456 | |
| 4225 | 3.150 | 3.691 | |

| | rh_transversetemporal_thickness | rh_insula_thickness | \ |
|---|---------------------------------|---------------------|---|
| 0 | 2.009 | 2.825 | |
| 1 | 2.111 | 2.720 | |
| 2 | 2.146 | 2.684 | |
| 3 | 2.056 | 2.700 | |
| 4 | 2.052 | 2.574 | |

| | | |
|------|-------|-------|
| ... | ... | ... |
| 4221 | 2.243 | 2.683 |
| 4222 | 2.040 | 2.866 |
| 4223 | 1.928 | 2.610 |
| 4224 | 2.317 | 2.900 |
| 4225 | 2.337 | 2.787 |

| | rh_MeanThickness_thickness | BrainSegVolNotVent.2 | eTIV.1 | dataset |
|------|----------------------------|----------------------|-------------|---------|
| 0 | 2.33635 | 1093846 | 1619602.965 | 1 |
| 1 | 2.34202 | 1099876 | 1624755.130 | 1 |
| 2 | 2.31982 | 1097999 | 1622609.518 | 1 |
| 3 | 2.29215 | 1070117 | 1583854.236 | 1 |
| 4 | 2.30397 | 1075926 | 1617375.362 | 1 |
| ... | ... | ... | ... | ... |
| 4221 | 2.29264 | 1108782 | 1561822.106 | 9 |
| 4222 | 2.30156 | 960586 | 1530179.480 | 9 |
| 4223 | 2.19622 | 1033357 | 1604323.353 | 9 |
| 4224 | 2.43580 | 1073339 | 1620891.799 | 9 |
| 4225 | 2.43420 | 992086 | 1513076.040 | 9 |

[4226 rows x 139 columns]

```
[4]: from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test = train_test_split(X,y,test_size=0.2,
→random_state=142)
x_train
```

```
[4]: Left-Lateral-Ventricle Left-Inf-Lat-Vent Left-Cerebellum-White-Matter \
2215 7887.3 206.4 14316.9
3708 7186.4 301.8 17914.4
57 10543.1 362.5 14313.3
4180 25067.3 819.8 12510.4
514 14770.7 315.2 12315.1
... ... ...
450 41251.8 513.9 12478.4
443 12105.8 572.6 11270.4
2187 8012.3 280.9 15634.6
1616 4014.0 131.0 14073.1
277 11448.4 692.4 16122.7
```

| | Left-Cerebellum-Cortex | Left-Thalamus | Left-Caudate | Left-Putamen \ |
|------|------------------------|---------------|--------------|----------------|
| 2215 | 43127.6 | 6706.9 | 2648.1 | 3822.5 |
| 3708 | 65322.8 | 9979.2 | 3413.6 | 5584.1 |
| 57 | 44585.8 | 5849.7 | 3362.7 | 4425.9 |
| 4180 | 52180.4 | 5916.2 | 2906.3 | 3375.2 |
| 514 | 45624.3 | 6867.6 | 3551.2 | 3624.0 |

| | | | | |
|------|---------|--------|--------|--------|
| ... | ... | ... | ... | ... |
| 450 | 57563.2 | 5842.8 | 3858.2 | 4575.0 |
| 443 | 46891.8 | 6223.7 | 2945.9 | 3671.8 |
| 2187 | 52776.8 | 9648.5 | 3353.4 | 4710.6 |
| 1616 | 56284.7 | 7894.7 | 3571.2 | 5597.3 |
| 277 | 51780.4 | 7569.9 | 3047.0 | 3831.4 |

| | | | | | |
|------|---------------|---------------|---------------|-----|---|
| | Left-Pallidum | 3rd-Ventricle | 4th-Ventricle | ... | \ |
| 2215 | 1485.3 | 1169.4 | 1463.3 | ... | |
| 3708 | 2400.1 | 1069.7 | 2833.7 | ... | |
| 57 | 1875.8 | 1092.6 | 1216.8 | ... | |
| 4180 | 1792.9 | 1484.3 | 1658.4 | ... | |
| 514 | 1920.9 | 1175.4 | 2041.7 | ... | |
| ... | ... | ... | ... | ... | |
| 450 | 1854.4 | 2233.4 | 1534.6 | ... | |
| 443 | 2074.2 | 1270.0 | 1712.0 | ... | |
| 2187 | 2108.2 | 1012.3 | 1734.4 | ... | |
| 1616 | 1953.9 | 791.4 | 1379.8 | ... | |
| 277 | 1617.4 | 1054.5 | 1441.8 | ... | |

| | | | |
|------|-------------------------------|----------------------------|---|
| | rh_superiortemporal_thickness | rh_supramarginal_thickness | \ |
| 2215 | 2.632 | 2.527 | |
| 3708 | 2.827 | 2.389 | |
| 57 | 2.655 | 2.238 | |
| 4180 | 2.671 | 2.307 | |
| 514 | 2.361 | 2.217 | |
| ... | ... | ... | |
| 450 | 2.421 | 2.126 | |
| 443 | 2.698 | 2.682 | |
| 2187 | 2.883 | 2.482 | |
| 1616 | 3.032 | 2.763 | |
| 277 | 2.448 | 2.176 | |

| | | | |
|------|--------------------------|---------------------------|---|
| | rh_frontalpole_thickness | rh_temporalpole_thickness | \ |
| 2215 | 2.199 | 3.764 | |
| 3708 | 2.988 | 3.485 | |
| 57 | 2.301 | 3.131 | |
| 4180 | 3.302 | 3.629 | |
| 514 | 2.645 | 2.821 | |
| ... | ... | ... | |
| 450 | 2.638 | 3.956 | |
| 443 | 2.622 | 3.953 | |
| 2187 | 3.055 | 3.473 | |
| 1616 | 2.742 | 3.519 | |
| 277 | 2.165 | 3.771 | |

| | | |
|---------------------------------|---------------------|---|
| rh_transversetemporal_thickness | rh_insula_thickness | \ |
|---------------------------------|---------------------|---|

| | | |
|------|-------|-------|
| 2215 | 2.332 | 2.750 |
| 3708 | 2.385 | 2.884 |
| 57 | 2.404 | 2.750 |
| 4180 | 2.469 | 2.929 |
| 514 | 2.008 | 2.613 |
| ... | ... | ... |
| 450 | 2.169 | 2.614 |
| 443 | 2.877 | 3.070 |
| 2187 | 2.445 | 2.654 |
| 1616 | 2.649 | 3.156 |
| 277 | 1.942 | 3.079 |

| | rh_MeanThickness_thickness | BrainSegVolNotVent.2 | eTIV.1 | dataset |
|------|----------------------------|----------------------|-------------|---------|
| 2215 | 2.34946 | 912547 | 1359253.430 | 4 |
| 3708 | 2.40574 | 1292198 | 1713954.995 | 9 |
| 57 | 2.29054 | 961822 | 1406281.699 | 1 |
| 4180 | 2.38743 | 973793 | 1441879.638 | 9 |
| 514 | 2.21423 | 1065788 | 1527516.270 | 1 |
| ... | ... | ... | ... | ... |
| 450 | 2.22891 | 1084538 | 1681641.995 | 1 |
| 443 | 2.50184 | 995258 | 1528786.788 | 1 |
| 2187 | 2.34528 | 1275413 | 1792900.541 | 4 |
| 1616 | 2.56356 | 1114795 | 1469161.228 | 2 |
| 277 | 2.23453 | 1108659 | 1541852.963 | 1 |

[3380 rows x 139 columns]

```
[5]: from sklearn.neighbors import KNeighborsRegressor
neigh = KNeighborsRegressor(n_neighbors=2)
```

```
[6]: neigh.fit(X, y)
```

```
[6]: KNeighborsRegressor(n_neighbors=2)
```

```
[7]: from sklearn.model_selection import cross_val_predict # For K-Fold Cross-Validation
      from sklearn.metrics import r2_score # For find accuracy with R2 Score
      from sklearn.metrics import mean_squared_error # For MSE
      from math import sqrt # For squareroot operation

      y_pred_train = neigh.predict(x_train)
      y_pred_test = neigh.predict(x_test)

      accuracy_train = r2_score(y_train, y_pred_train)
      print("Training R2 for Regression Model: ", accuracy_train)

      accuracy_test = r2_score(y_test, y_pred_test)
```

```

print("Testing R2 for Regression Model: ", accuracy_test)

RMSE_train = sqrt(mean_squared_error(y_train, y_pred_train))
print("RMSE for Training Data: ", RMSE_train)

RMSE_test = sqrt(mean_squared_error(y_test, y_pred_test))
print("RMSE for Testing Data: ", RMSE_test)

```

```

Training R2 for Regression Model:  0.8844393019224387
Testing R2 for Regression Model:  0.8919469364118608
RMSE for Training Data:  6.830984707273609
RMSE for Testing Data:  6.549976088599814

```

```

[8]: true_val = y_train
     pred_val = y_pred_train

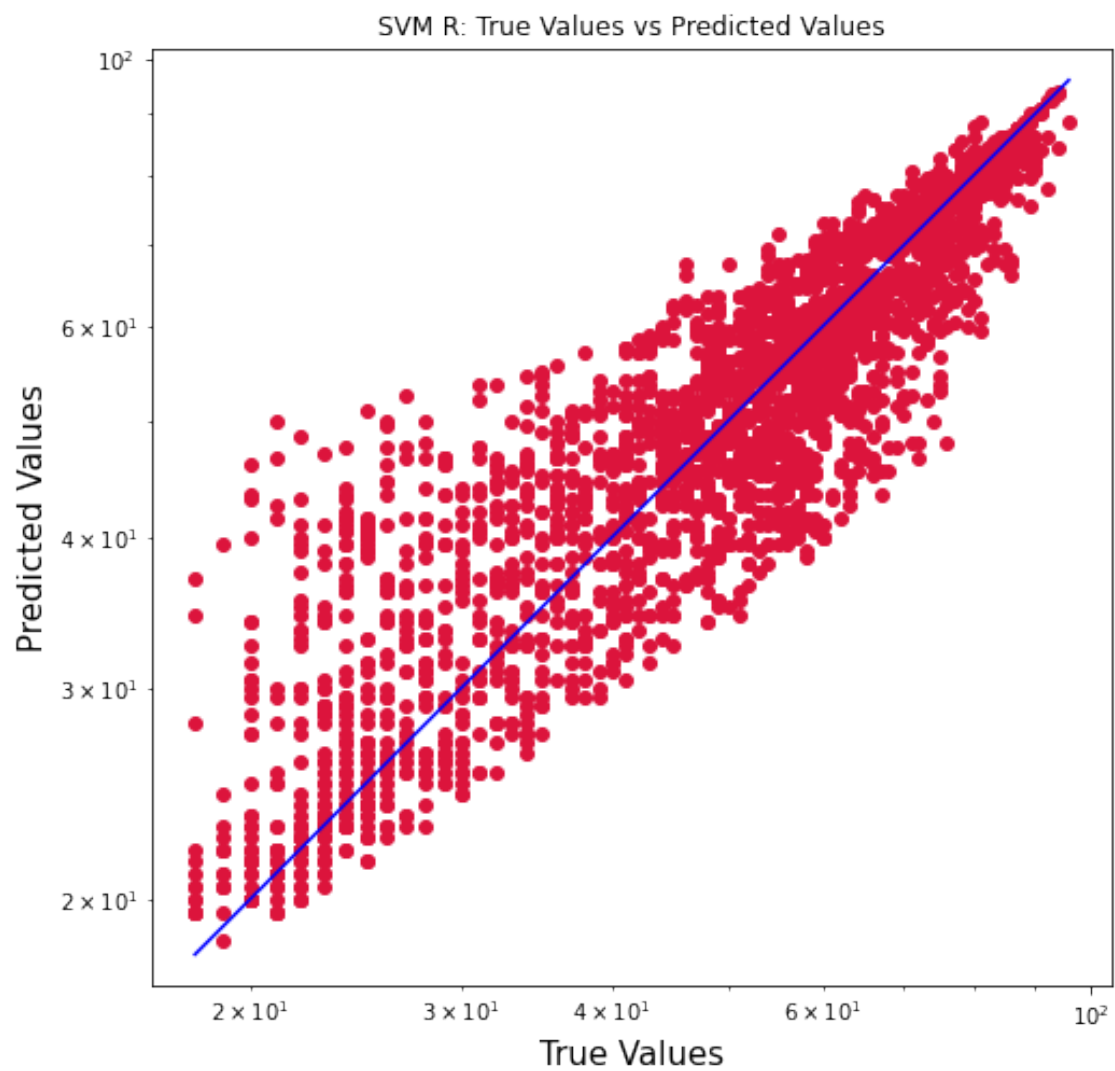
```

```

[9]: plt.figure(figsize=(8,8))
     plt.scatter(true_val, pred_val, c='crimson')
     plt.yscale('log')
     plt.xscale('log')

     p1 = max(max(pred_val), max(true_val))
     p2 = min(min(pred_val), min(true_val))
     plt.plot([p1, p2], [p1, p2], 'b-')
     plt.xlabel('True Values', fontsize=15)
     plt.ylabel('Predicted Values', fontsize=15)
     plt.title("SVM R: True Values vs Predicted Values")
     plt.axis('equal')
     plt.show()

```

[]: