**Outlier Detection for Non-Normal 1D data**

Data Size : 50000+1000

Outliers : 2.225 % (1135) ( Outliers are added at only one tail i.e. Right tail)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm | Parameters | TP | FP | TN | FN | Precision | Recall |
| IQR | No | 1135 | 161 | 49704 | 0 | 0.997 | 1.0 |
| Isolation Forest | Contamination = 0.02225 | 1065 | 59 | 49806 | 70 | 0.997 | 0.938 |
| Contamination = ‘auto’ | 1135 | 3601 | 46264 | 0 | 0.929 | 1.0 |
| Local Outlier Factor | Contamination = 0.02225, n\_neighbors = ‘auto’ | 57 | 1078 | 48787 | 1078 | 0.958 | 0.050 |
| Contamination = 0.02225,n\_neighbors = 100 | 288 | 847 | 49018 | 847 | 0.967 | 0.254 |
| Contamination = 0.02225, n\_neighbors = 500 | 463 | 672 | 49193 | 672 | 0.974 | 0.408 |
| Contamination = 0.02225, n\_neighbors = 1000 | 900 | 235 | 49630 | 235 | 0.991 | 0.793 |
| Contamination = 0.02225, n\_neighbors = 2000 | 1134 | 1 | 49864 | 1 | 0.999 | 0.999 |
| One Class SVM | Kernel = ‘linear’, Nu = 0.02225 | 135 | 49865 | 0 | 1000 | 0.003 | 0.119 |
| Kernel = ‘rbf’, Nu = 0.02225 | 1083 | 51 | 49814 | 52 | 0.998 | 0.954 |

Data Size : 50000+2500

Outliers : 5.019 % (2635) ( Outliers are added at both the tails i.e. left and right equally )

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm | Parameters | TP | FP | TN | FN | Precision | Recall |
| IQR | No | 2635 | 69 | 49796 | 0 | 0.998 | 1.0 |
| Isolation Forest | Contamination = 0.05019 | 2611 | 16 | 49849 | 24 | 0.999 | 0.991 |
| Contamination = ‘auto’ | 2635 | 974 | 48891 | 0 | 0.981 | 1.0 |
| Local Outlier Factor | Contamination = 0.05019,neighbors = ‘auto’ | 169 | 2466 | 47399 | 2466 | 0.906 | 0.064 |
| Contamination = 0.05019,n\_neighbors = 100 | 542 | 2093 | 47772 | 2093 | 0.920 | 0.206 |
| Contamination = 0.05019, n\_neighbors = 500 | 1330 | 1305 | 48560 | 1305 | 0.950 | 0.505 |
| Contamination = 0.05019, n\_neighbors = 1000 | 1598 | 1037 | 1037 | 48828 | 0.960 | 0.606 |
| Contamination = 0.05019, n\_neighbors = 2000 | 2634 | 1 | 49864 | 1 | 0.999 | 0.999 |
| One Class SVM | Kernel = ‘linear’, Nu = 0.05019 | 2538 | 93 | 49772 | 97 | 0.996 | 0.963 |
| Kernel = ‘rbf’, Nu = 0.05019 |  |  |  |  |  |  |

Data Size : 50000+5000

Outliers : ( Outliers are added at only one tail i.e. Left tail )

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm | Parameters | TP | FP | TN | FN | Precision | Recall |
| IQR | No |  |  |  |  |  |  |
| Isolation Forest | Contamination = 0.02225 |  |  |  |  |  |  |
| Local Outlier Factor | Contamination = 0.02225, n\_neighbors = ‘auto’ |  |  |  |  |  |  |
| Contamination = 0.02225,n\_neighbors = 100 |  |  |  |  |  |  |
| Contamination = 0.02225, n\_neighbors = 500 |  |  |  |  |  |  |
| Contamination = 0.02225, n\_neighbors = 1000 |  |  |  |  |  |  |
| Contamination = 0.02225, n\_neighbors = 2000 |  |  |  |  |  |  |
| One Class SVM | Kernel = ‘linear’, Nu = 0.02225 |  |  |  |  |  |  |
| Kernel = ‘rbf’, Nu = 0.02225 |  |  |  |  |  |  |

Data Size : 50000+10000

Outliers : ( Outliers are added at both tails i.e. left tail and right tail )

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm | Parameters | TP | FP | TN | FN | Precision | Recall |
| IQR | No |  |  |  |  |  |  |
| Isolation Forest | Contamination = 0.02225 |  |  |  |  |  |  |
| Local Outlier Factor | Contamination = 0.02225, n\_neighbors = ‘auto’ |  |  |  |  |  |  |
| Contamination = 0.02225,n\_neighbors = 100 |  |  |  |  |  |  |
| Contamination = 0.02225, n\_neighbors = 500 |  |  |  |  |  |  |
| Contamination = 0.02225, n\_neighbors = 1000 |  |  |  |  |  |  |
| Contamination = 0.02225, n\_neighbors = 2000 |  |  |  |  |  |  |
| One Class SVM | Kernel = ‘linear’, Nu = 0.02225 |  |  |  |  |  |  |
| Kernel = ‘rbf’, Nu = 0.02225 |  |  |  |  |  |  |