**Pembersihan Data**

1. **Oil-Spill Dataset**
   1. Tanpa Pembersihan
      1. Logistic Regression
         * Sensitivity = 0.9898648648648649
         * Specificity = 0.14285714285714285
         * Precision = 0.9606557377049181
         * NPV = 0.4
         * Accuracy = 0.9516129032258065
      2. MLP
         * Sensitivity = 0.9833333333333333
         * Specificity = 0.0
         * Precision = 0.9672131147540983
         * NPV = 0.0
         * Accuracy = 0.9516129032258065
   2. Penghapusan single value
      1. Logistic Regression
         * Sensitivity= 0.9898648648648649
         * Specificity= 0.14285714285714285
         * Precision= 0.9606557377049181
         * NPV= 0.4
         * Accuracy= 0.9516129032258065
      2. MLP
         * Sensitivity= 0.9832214765100671
         * Specificity= 0.0
         * Precision= 0.9606557377049181
         * NPV= 0.0
         * Accuracy= 0.9451612903225807
   3. Penghapusan few values
      1. Logistic Regression
         * Sensitivity= 0.9898648648648649
         * Specificity= 0.14285714285714285
         * Precision= 0.9606557377049181
         * NPV= 0.4
         * Accuracy= 0.9516129032258065
      2. MLP
         * Sensitivity= 0.9831081081081081
         * Specificity= 0.0
         * Precision= 0.9540983606557377
         * NPV= 0.0
         * Accuracy= 0.9387096774193548
   4. Penghapusan low variance
      1. Logistic Regression
         * Sensitivity= 0.9898648648648649
         * Specificity= 0.14285714285714285
         * Precision= 0.9606557377049181
         * NPV= 0.4
         * Accuracy= 0.9516129032258065
      2. MLP
         * Sensitivity= 0.9833887043189369
         * Specificity= 0.0
         * Precision= 0.9704918032786886
         * NPV= 0.0
         * Accuracy= 0.9548387096774194
   5. Menggunakan Variance treshold
      * + \*Tercantum dalam file .ipnyb
2. **Iris Dataset**
   1. Tanpa Pembersihan
      1. Logistic Regression
         * Iris-Setosa
           1. Sensitivity = 1.0
           2. Specificity = 1.0
           3. Precision = 1.0
           4. NPV = 1.0
           5. Accuracy = 1.0
         * Iris-Versicolor
           1. Sensitivity = 0.9047619047619048
           2. Specificity = 0.9655172413793104
           3. Precision = 0.95
           4. NPV = 0.9333333333333333
           5. Accuracy = 0.94
         * Iris-Virginica
           1. Sensitivity = 0.9230769230769231
           2. Specificity = 0.9459459459459459
           3. Precision = 0.8571428571428571
           4. NPV = 0.9722222222222222
           5. Accuracy = 0.94
      2. MLP
         * Iris-Setosa
           1. Sensitivity = 1.0
           2. Specificity = 1.0
           3. Precision = 1.0
           4. NPV = 1.0
           5. Accuracy = 1.0
         * Iris-Versicolor
           1. Sensitivity = 0.625
           2. Specificity = 0.9705882352941176
           3. Precision = 0.9090909090909091
           4. NPV = 0.8461538461538461
           5. Accuracy = 0.86
         * Iris-Virginica
           1. Sensitivity = 0.9444444444444444
           2. Specificity = 0.8125
           3. Precision = 0.7391304347826086
           4. NPV = 0.9629629629629629
           5. Accuracy = 0.86
   2. Penghapusan Data Terduplikasi
      1. Logistic Regression
         * Iris-Setosa
           1. Sensitivity = 1.0
           2. Specificity = 1.0
           3. Precision = 1.0
           4. NPV = 1.0
           5. Accuracy = 1.0
         * Iris-Versicolor
           1. Sensitivity = 0.625
           2. Specificity = 0.9705882352941176
           3. Precision = 0.9090909090909091
           4. NPV = 0.8461538461538461
           5. Accuracy = 0.86
         * Iris-Virginica
           1. Sensitivity = 0.9444444444444444
           2. Specificity = 0.8125
           3. Precision = 0.7391304347826086
           4. NPV = 0.9629629629629629
           5. Accuracy = 0.86
      2. MLP
         * Iris-Setosa
           1. Sensitivity = 1.0
           2. Specificity = 1.0
           3. Precision = 1.0
           4. NPV = 1.0
           5. Accuracy = 1.0
         * Iris-Versicolor
           1. Sensitivity = 0.75
           2. Specificity = 1.0
           3. Precision = 1.0
           4. NPV = 0.8529411764705882
           5. Accuracy = 0.8979591836734694
         * Iris-Virginica
           1. Sensitivity = 1.0
           2. Specificity = 0.8529411764705882
           3. Precision = 0.75
           4. NPV = 1.0
           5. Accuracy = 0.8979591836734694

Analisa :

Dengan membersihkan data kita akan meningkatkan produktifitas dan kualitas informasi yang model kita prediksi. Dapat dilihat melalui data-data diatas bahwa secara rata-rata terdapat kenaikan nilai Sensitivity, Specificity, Precision, NPV, maupun Accuracy. Contohnya pada model MLP data Iris-Virginica, Accuracy yang sebelumnya 0.86 meningkat menjadi 0.8979591836734694.