Software Quality & Project Management

Lab Report: 4

CRN: 75765

Title: MDL Model Testing and Evaluation

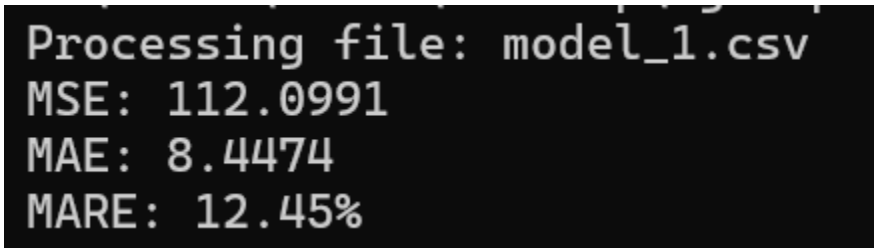
Liam Reid 100875207

Github Link: https://github.com/Liam-Reid1/SOFE3980U-Lab4

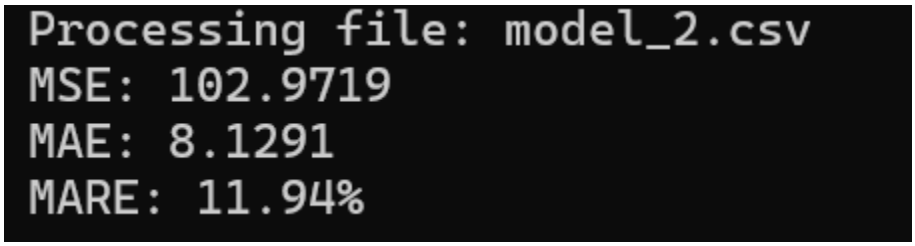
Task 1:

```
C:\Users\liamr\Desktop\mgmt project\lab4\SOFE3980U-Lab4\SVCR>java -jar target/SVCR-1.0.0-jar-with-dependen
106.15388        113.56206
129.09142        147.8016
34.36562         29.701094
54.359077        55.832047
79.5487          78.70409
41.90821         31.97168
123.00964        111.9446
104.46343        99.56216
50.432903        58.677044
123.52526        131.32625
```

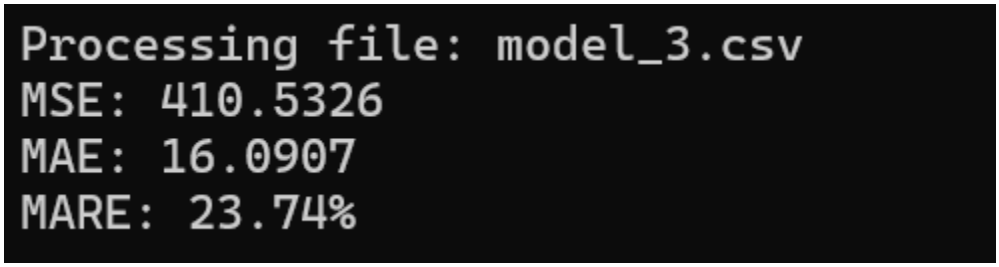
Model_1:



Model_2:



Model_3:



Comparing the three models, model_2 has the lowest MSE, MAE, and MARE score, making it the model with the lowest error.

Task 2:

```
C:\Users\liamr\Desktop\mgmt project\lab4\SOFE3980U-Lab4\SVBR>java -jar target/SVBR-1.0.0-jar-with-dependenci
0        0.3280905
1        0.70533866
0        0.5030721
1        0.5299625
1        0.575525
0        0.30359086
0        0.001
1        0.7970452
1        0.906946
1        0.999
```

Model_1:

```
Processing file: model_1.csv
BCE: 0.3844
Accuracy: 0.8441
Precision: 0.8459
Recall: 0.8461
F1 Score: 0.8460
AUC-ROC: 0.8441
Confusion Matrix:
TP: 4283  FP: 780
FN: 779  TN: 4158
```

Model_2:

```
Processing file: model_2.csv
BCE: 0.3404
Accuracy: 0.8931
Precision: 0.8992
Recall: 0.8884
F1 Score: 0.8938
AUC-ROC: 0.8932
Confusion Matrix:
TP: 4497  FP: 504
FN: 565  TN: 4434
```

Model_3:

```
Processing file: model_3.csv
BCE: 0.3122
Accuracy: 0.9546
Precision: 0.9555
Recall: 0.9548
F1 Score: 0.9551
AUC-ROC: 0.9546
Confusion Matrix:
TP: 4833  FP: 225
FN: 229  TN: 4713
```

Model_3 has the highest accuracy, precision, recall, F1 Score, and AUC-ROC, making it the model with the best performance. Additionally its BCE score is the lowest,making it an optimal score.

Task 3:

```
3       0.13846475      0.08112068      0.5465031       0.08852483      0.14538662
3       0.05339846      0.14028792      0.50836426      0.1183168       0.17963256
3       0.22510053      0.09468151      0.34942052      0.2268822       0.10391525
5       0.20582186      0.20312473      0.10914556      0.11105233      0.37085554
3       0.24865267      0.16397181      0.2927781       0.08857448      0.20602293
4       0.041044727     0.20974156      0.124620296     0.52514243      0.09945098
4       0.11933829      0.14353032      0.13375178      0.52552944      0.07785014
3       0.2717753       0.13703002      0.38883603      0.10864721      0.09371145
3       0.12810154      0.121750504     0.29039726      0.19095209      0.26879862
2       0.18415536      0.5249311       0.15088637      0.092125826     0.04790138
```

```
Cross-Entropy (CE): 1.0077
Confusion Matrix:
505 35 35 28 44
148 1906 139 136 130
197 238 2886 202 237
145 144 126 1944 139
33 37 33 32 501
```

Discussion:

In machine learning, accuracy, precision, and recall play roles in understanding the model's performance. Each of these scores contributes to the problem depending on the application.

Accuracy:
This measures the overall correctness of a model. It provides a general rating of how well the model is performing. However, accuracy may be misleading depending on the datasets. For example, in a medical test for a rare disease where 99% of patients do not have the disease, a model predicting 'no disease' for everyone would have 99% accuracy but fail to detect actual cases. In this case, accuracy alone is not a reliable metric.

Precision:
This measurement tells us the proportion of the predicted positive cases are actually positive. This can be very important when it is imperative to avoid false positives. For example, in email spam detection, marking a legitimate email as spam (false positive) can result in lost important messages. Therefore, a high-precision model is preferred to minimize false positives.

Recall:
Measures the proportion of actual positives correctly identified. This can be very important when missing a positive instance is dangerous. For example, high recall is essential to detect all possible cases of a disease, even at the cost of some false positives.