Bachelor Project - Results log

Timo Wahl (s3812030)

March 2021

Replication results

First results

Date: 3-3-2021
Results from forming three multiple layer perceptron classifiers as described in the paper. The first classifier consists of 1 hidden layer with 12 nodes. The second classifier consists of two hidden layers with 24 and 6 nodes respectively. The third classifier consists of three hidden layers with 24, 10 and 3 nodes respectively.

In type A (multiple fail) the classifiers are quite good at getting a good accuracy. The accuracy being close to 92 percent on average. The second and third classifiers perform slightly better at an average accuracy close to 93 percent.
In type B (single fail) the classifiers are still quite good at getting a decent accuracy. The accuracy being close to 85 percent on average. The second and third classifiers perform slightly worse (both 83 percent) when compared to the first classifier (86 percent).

Concluding: The percentages for the classifiers do not mimic the paper, as the percentages in the paper are much higher for type A (99.25, 98.9, 98.75 for classifier 1, 2 and 3) and lower for type (72.25, 76.67, 74.33 for classifier 1,2,3 in the paper). It is interesting to see that, like in the paper, the accuracy percentages do drop when comparing type A to type B. However, as the percentages do not directly mimic the paper, questions remain as to why that is and how that can be fixed. The results are also not very stable, as they are only recording over 600 data points (4 way split over 2400), therefore a loop over 2400 data points is likely a better option to get more stable results.

```
MLPClassifier(hidden_layer_sizes=12, max_iter=3000)
[[267  48]
 [ 35 250]]
              precision    recall  f1-score   support

           0       0.88      0.85      0.87       315
           1       0.84      0.88      0.86       285

    accuracy                           0.86       600
   macro avg       0.86      0.86      0.86       600
weighted avg       0.86      0.86      0.86       600

MLPClassifier(hidden_layer_sizes=(24, 6), max_iter=3000)
[[250  65]
 [ 39 246]]
              precision    recall  f1-score   support

           0       0.87      0.79      0.83       315
           1       0.79      0.86      0.83       285

    accuracy                           0.83       600
   macro avg       0.83      0.83      0.83       600
weighted avg       0.83      0.83      0.83       600

MLPClassifier(hidden_layer_sizes=(24, 10, 3), max_iter=3000)
[[260  55]
 [ 50 235]]
              precision    recall  f1-score   support

           0       0.84      0.83      0.83       315
           1       0.81      0.82      0.82       285

    accuracy                           0.82       600
   macro avg       0.82      0.82      0.82       600
weighted avg       0.83      0.82      0.83       600
```

Figure 1: Example results for 3-3-2021 for the type A condition. From the figure the confusion matrix, and classification report for each of the three classifiers are visible.

Second results

Date: 17-3-2021

Results show that the percentages for the normal dataset are often quite close to those of the paper, but ultimately fail to get within an insignificant distance of those percentages from the paper. For example the percentages for training on the multiple fail condition and testing on it show good results coming within 2-4 percent of the required percentages. However other major differences can be found for example the net trained on the single fail condition returns too low results for both the single fail and multiple fail test sets. Another major difference is that in the paper, the three layered network does not converge, which it does for me.

```
-----------------------------------------------------------------------------------
ITERATIONS: 20
===================================================================================
Training on multiple fail, testing on multiple fail final result
Single layer neural network accuracy: 0.9518499999999998
Double layer neural network accuracy: 0.9525750000000001
Triple layer neural network accuracy: 0.9475000000000001
===================================================================================
Training on multiple fail, testing on single fail final result
Single layer neural network accuracy: 0.809375
Double layer neural network accuracy: 0.8125499999999999
Triple layer neural network accuracy: 0.80395
===================================================================================
Training on single fail, testing on multiple fail final result
Single layer neural network accuracy: 0.950675
Double layer neural network accuracy: 0.9460749999999998
Triple layer neural network accuracy: 0.934075
===================================================================================
Training on single fail, testing on single fail final result
Single layer neural network accuracy: 0.9054249999999998
Double layer neural network accuracy: 0.8961749999999998
Triple layer neural network accuracy: 0.8736
```

Figure 2: Example results for 17-3-2021 for all the conditions for the normal dataset. Averaged over 20 iterations

The results in figures 3 - 5 show the output when testing on the age and distance dataset (with a neural network trained on single or multiple fail conditions). All the graphs somewhat align with those from the paper; generally showing the same important characteristics, however as was the case with figure 2, the specifics of the graphs differ when compared to those from the paper.
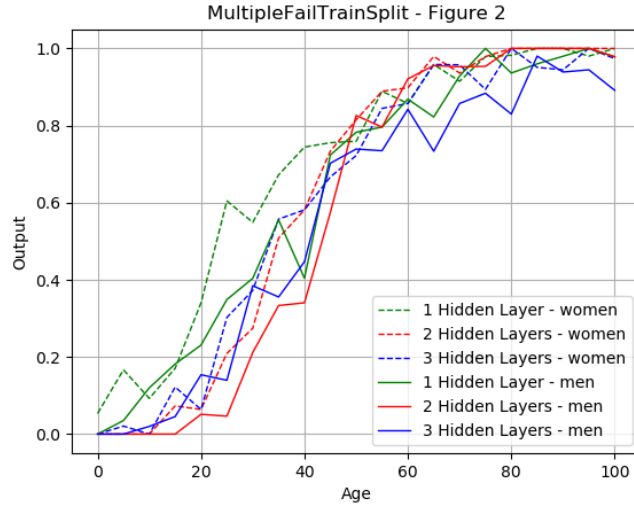
Figure 3: Multiple fail trained network results when testing on the age dataset
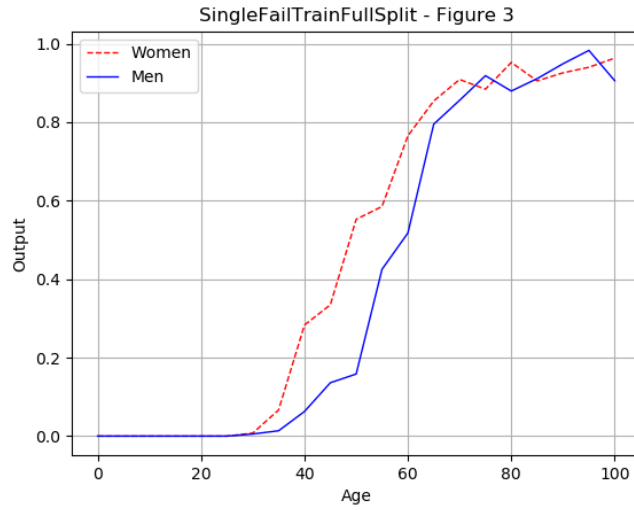

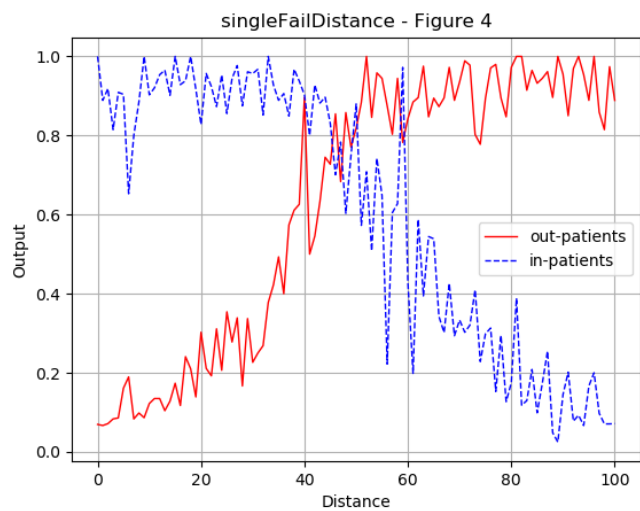Figure 4: Single fail trained network results when trained on the age dataset

Figure 5: Single fail trained network results when trained on the distance dataset