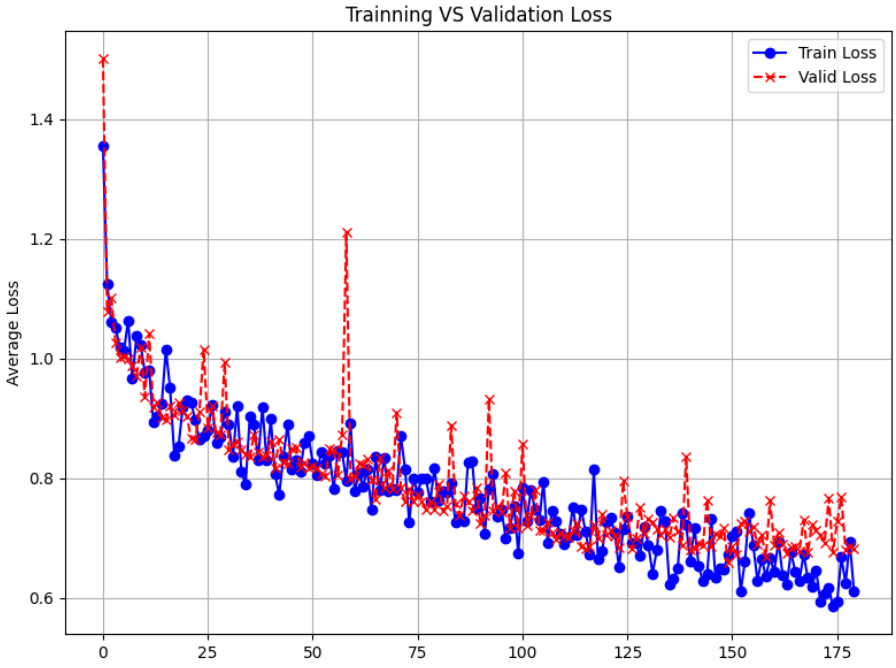
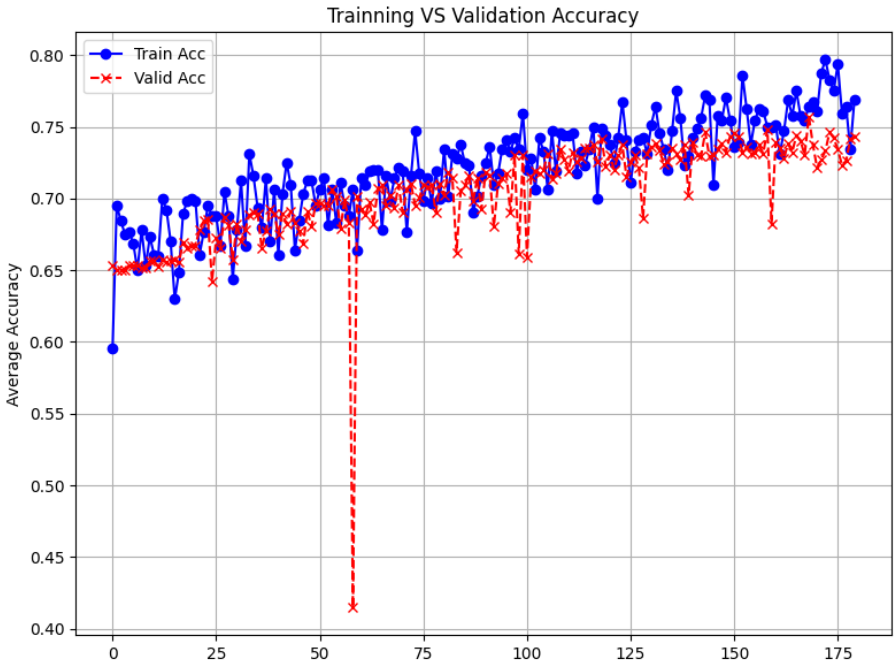
HAROKOPIO UNIVERSITY
DEPARTMENT OF INFORMATION & TELEMATICS

2nd Assignment Report: Machine Learning and Applications

Manousos Linardakis, it22064

Simple SND

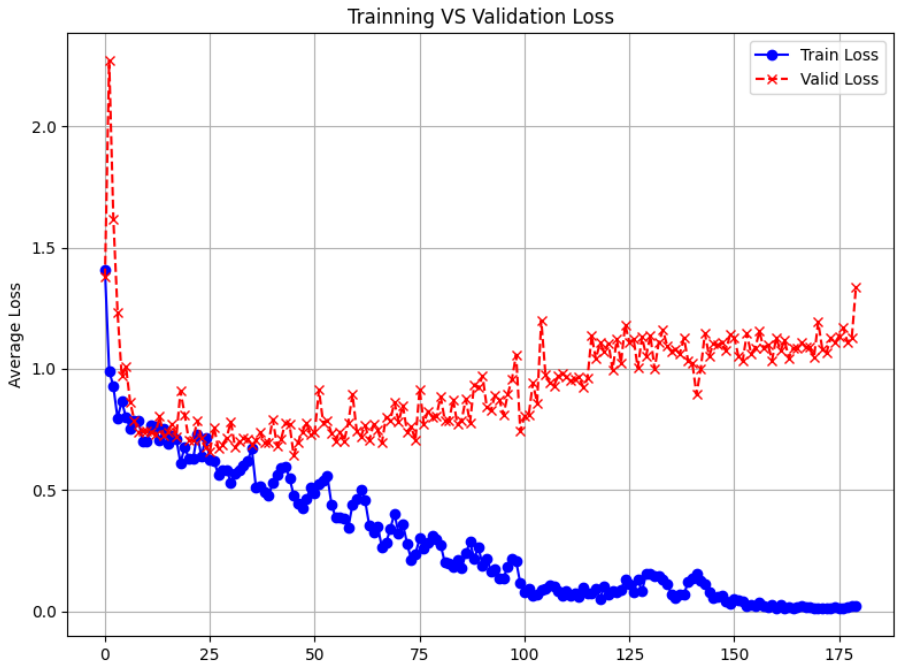

Trainning VS Validation Loss
Train Loss
Valid Loss
Average Loss
1.4
1.2
1.0
0.8
0.6
0
25
50
75
100
125
150
175

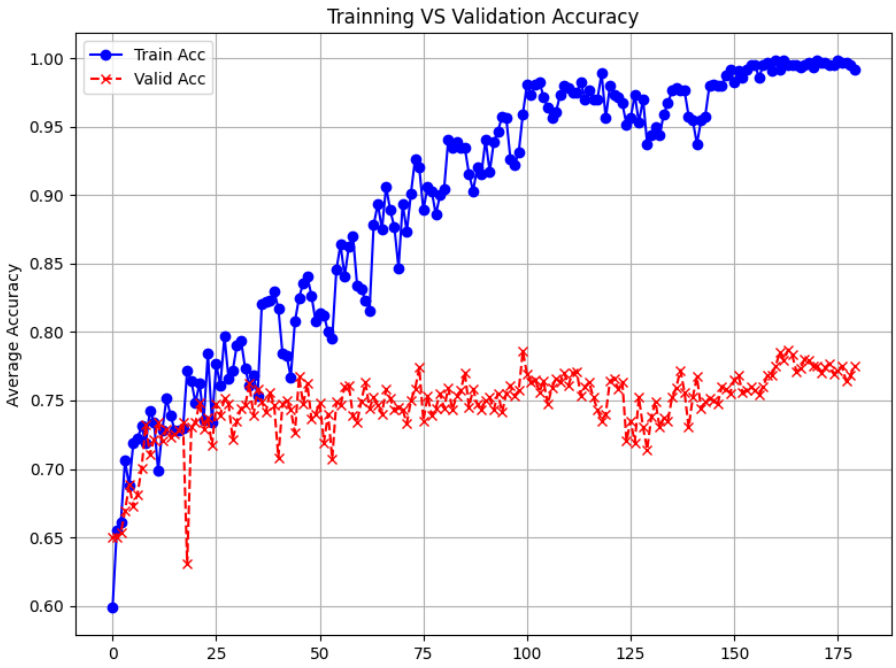

Trainning VS Validation Accuracy
Train Acc
Valid Acc
Average Accuracy
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0
25
50
75
100
125
150
175

```
===============
Average test loss: 0.711 | accuracy: 0.750 Confusion
Matrix:
[[   19    16     9     0    18    25     0]
 [   11    71     3     0    12    42     0]
 [    6    22    74     0    61   161     0]
 [    3    14     3     0     5    14     0]
 [    2     3     9     1 150     169     1]
 [    1    14    11     0 86 1926     1]
 [    0     6     0     0     3    18    15]]
===============
```

With the simple SND, we achieved a fairly good performance. From the confusion matrix we can see (looking at the main diagonal) that the model does a very good job of classifying the condition “Dermatofibroma” (6th line). It also does a good job of predicting the disease “Spilos” (2nd line). As the training experiments pass, the loss decreases and the accuracy increases. Also, no overfitting is observed, as the training and validation losses decrease at the "same" rate and do not deviate from each other.

Complex SND

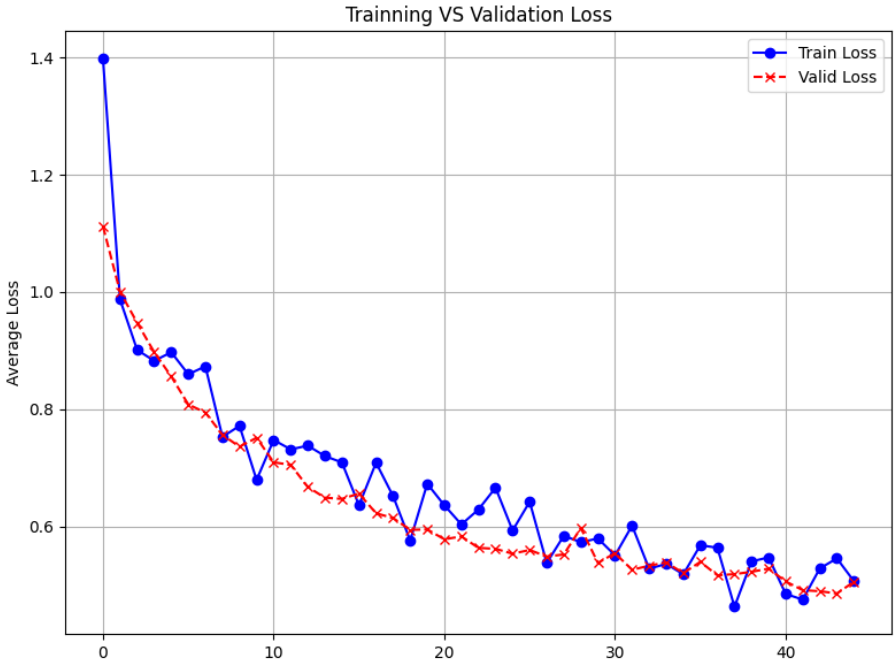


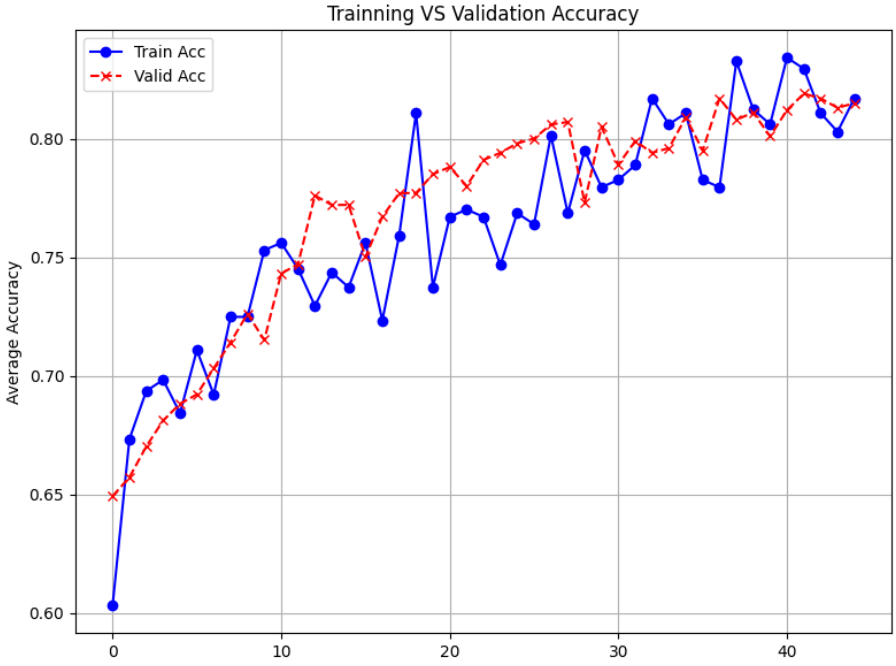



```
===============
Average test loss: 1,190 | accuracy: 0.781 Confusion
Matrix:
[[ 26 11 27   1   9  13   0]
 [ 9 81 14   3   4  26   2]
 [  8  15 180   2  20  97   2]
 [  3   7   3  16   0  10   0]
 [  7   3  38   2 140 144   1]
 [  5  36  46   0  71 1876   5]
 [  0   1   2   0   1   9  29]]
===============
```

In the complex SND, we see that the model did worse than the simple one. Immediately from the plot of Training VS Validation loss we see that it becomes overfit (towards the end, the validation curve deviates from the training, as the training decreases). This "handiness" is also observed by the average test loss, which is greater than the average test loss of the simple SND.

Transfer learning - transfer learning

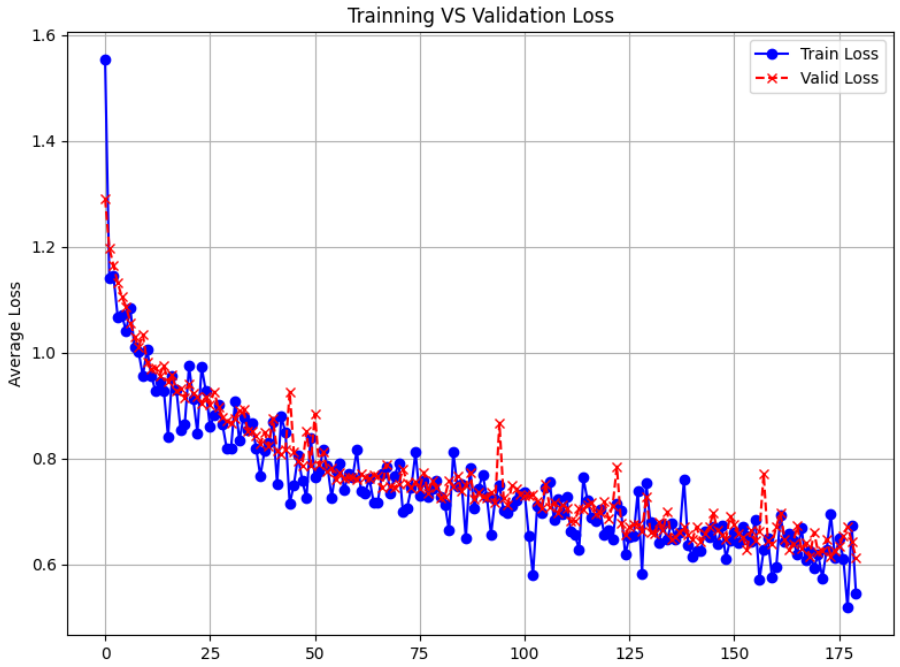


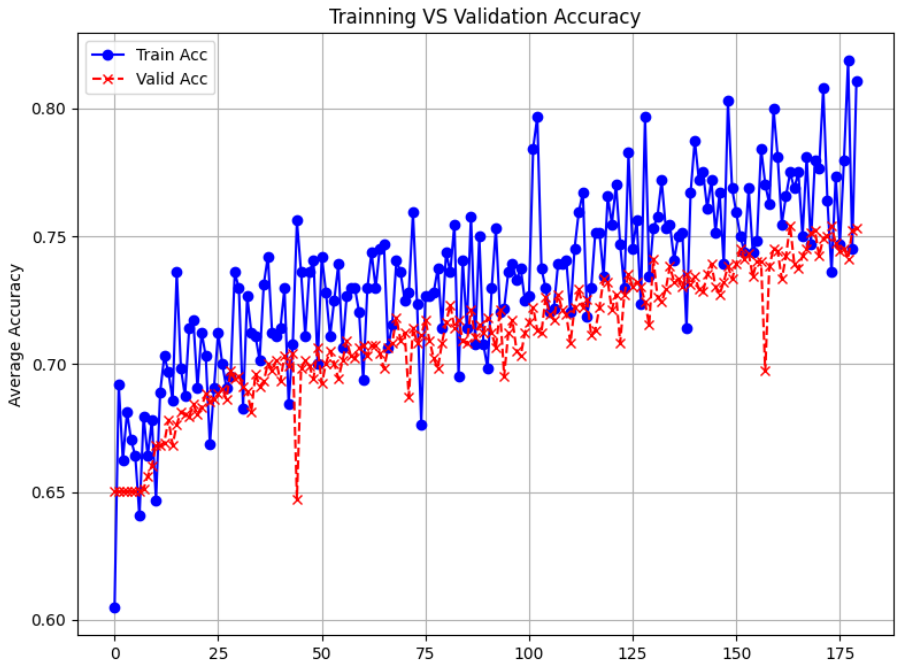


```
===============
Average test loss: 0.480 | accuracy: 0.828 Confusion
Matrix:
[[  48   6  17   0   6  10   0]
 [  12 104   7   0   5  11   0]
 [   6   8 222   2  11  75   0]
 [   3   2   9   2   2  18   3]
 [   9   3  45   0 118 160   0]
 [   2  11  42   1  22 1960   1]
 [   0   3   1   0   0   5  33]]
===============
```

By using ResNet34 we notice that the model performance improved even more (and in fewer epochs). In particular at the end we can see that the loss dropped below 0.5 (which is not observed in the previous models). There is also no overfitting of the data as the curves in Training VS Validation Loss do not deviate from each other (as in the 5th question). It is important to note that the average test loss decreased while the accuracy increased (compared to the previous SNDs). Also, from the confusion matrix we see that the SND has better predicted the BCC (3rd line) and VASC (7th line) classes compared to the previous queries. This behavior makes sense as it (resnet34) is already trained on millions of images giving it a significant head start for better classification over previous models.

Bonus: Utilization of demographic variables






```
===============
Average test loss: 0.660 | accuracy: 0.761 Confusion
Matrix:
[[   24    34    25     0     1     3     0]
 [   11    99    14     1     2     9     3]
 [    9    49   186     0     9    71     0]
 [    6    14     3     4     0    12     0]
 [   12    19    72     0    79   150     3]
 [   12    37    79     1    25 1877     8]
 [    0    11     1     2     1     9    18]]
===============
```

Using the SimpleModel of the 4th query, and adding the vector p with the demographic data, we notice that the loss in training and validation fell more than the simple model of the 4th query, so it has a better performance. This behavior can also be seen in testing, as the model achieved a lower avg loss and a higher avg accuracy than it did in the 4th question. Also, from the confusion matrix we see that the SND has better predicted the BCC class (3rd line) compared to the 4th query. This improvement is expected, as demographic characteristics are important information for predicting skin diseases/damages.