Lukas van der Watt

Assignment 2

Questions:

1. Results Regular model

Text

Description automatically generated

Graphical user interface, chart, histogram

Description automatically generated

Graphical user interface, chart, application, histogram

Description automatically generated

Graphical user interface, text

Description automatically generated

Augmentation model:

Graphical user interface, text, application

Description automatically generated

A performance (accuracy) of 76.5% was achieved with the augmentation model. A training sample size of 1000, validation sample of 500 and test sample of 500 was used. Compared to the model without augmentation it performed better by about 6%

1. Text

   Description automatically generatedRegular Model:

Graphical user interface

Description automatically generated

Augmentation Model: Graphical user interface, text, application, email

Description automatically generated

An accuracy of 82.2% can be seen in the augmentation model of in question 2. The training sample size was increased to 1500 instead of 1000. With the increase in training sample, we also see an increase in accuracy on the test set. The model can generalize better.

3) Regular Model:

Text

Description automatically generated

A picture containing table

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generatedAugmentation Model:

The augmentation model had an accuracy of 83.2% which outperformed the augmentation models in questions 1 & 2. It had a training sample size of 2000 which was the highest of the previous questions. It can therefore be seen that a larger training sample will allow for better accuracy on the test set once the model has been trained. Note: epochs of the code was only run to 30 due to execution time and computing power restrictions. A higher accuracy is therefore very achievable with more epochs run.

1. Graphical user interface, text, application, email

   Description automatically generatedDue to the immense amount of time it took to run just one epoch and all the interruptions that occurred and collab’s requirement to then restart the chunk I was not able to finish this question. However this pretrained model will perform better than the “from scratch method” done earlier because it was trained on thousands of data points and will have a bigger training sample. The trend to notice in this assignment is that a bigger training sample will allow for a better accuracy on the test set.

