ADVANCED MACHINE LEARNING

ASSIGNMENT -2

REPORT

I have two approaches, for training the model; starting from scratch or using a pretrained model. In terms of performance I measure accuracy and validation loss along with the number of epochs it takes to train.

When I initially trained the model with 1000 training samples, 500 validation samples and 1000 test samples I achieved an accuracy of 71.6% with a validation loss of 1.86 after 30 epochs. On the hand using a pretrained model gave me an accuracy of 98.3% with a significantly lower validation loss of 0.01 after just 20 epochs.

To further improve my results I implemented data augmentation techniques by increasing my training data to include 2000 samples while keeping the number of validation and test samples. This led to an accuracy boost to 78.3% although there was an increase in validation loss to 0.95 after training for around 100 epochs.

Next I experimented with increased training data where I doubled the number of training samples while maintaining the number for validation and testing purposes (2000 each). Despite this effort my accuracy dropped slightly to 71.7% with a validation loss of 2.19 after training for another set of approximately equal length epochs compared to trials.

Finally using training data consisting of four times many samples (4000) for training as before while maintaining consistency in the other sets (500 each) I managed to achieve better results again. My accuracy rose to around 73.5%. There was still some increase in validation loss which reached up to about2.49 at this point after another set length equal length epoch compared to prior experiments.

Through iterations and adjustments in both size and model initialization strategies my best results were obtained when utilizing a pretrained model with increased amounts of optimized training data resulting in an impressive 99.6% accuracy and a relatively low validation loss of 0.07, after training for approximately 20 epochs.

In my findings it appears that using a pretrained network yields results compared to training a network from scratch. While increasing the training sample may improve test accuracy it also leads to an increase, in validation loss. I've noticed that data augmentation performs well when training a network from scratch but not with a pretrained network. In my opinion the optimal model would involve using a pretrained network, with increased training data resulting in a 99.4% test accuracy and minimal validation loss of 0.04 with fewer epochs.