**Assignment 2**

**Convolution**

**Report**

**Introduction:** **In this study, we will assess the performance of developing a convolution neural network using the Cats and Dogs example, adjusting and determining which sample size and approach is most beneficial throughout the model-building stage. .**

**Methodology:**

**We created six Scratch Models and three Pre-Trained Models in a variety of setups. These configurations differ in terms of the number of layers, nodes, optimizers, dropout rates, and other characteristics.**

**Scratch Models**

**Validation Accuracy, Test Accuracy, and Test loss**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model no** | **Training sample size** | **Validation and Test sample size** | **Validation  Accuracy** | **Test Accuracy** | **Test Loss** |
| 1 | 1000 | 500,500 | 0.74 | 0.73 | 0.55 |
| 1a | 1000 | 500,500 | 0.79 | 0.81 | 0.42 |
| 1b | 1000 | 500,500 | 0.77 | 0.75 | 0.50 |
| 1c | 1000 | 500,500 | 0.80 | 0.83 | 0.37 |
| 2 | 5000 | 500,500 | 0.99 | 0.87 | 0.42 |
| 3 | 10000 | 500,500 | 0.91 | 0.88 | 0.30 |

**Pre-Trained Models**

**Validation Accuracy, Test Accuracy, and Test Loss**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model no** | **Training Sample Size** | **Validation and test sample size** | **Validation  Accuracy** | **Test Accuracy** | **Test Loss** |
| Model 4 | 1000 | 500,500 | 0.97 | 0.979 | 0.20 |
| Model 5 | 5000 | 500,500 | 0.975 | 0.976 | 0.12 |
| Model 6 | 10000 | 500,500 | 0.98 | 0.986 | 0.044 |

**Observations:**

* An unregularized Model of Cats and Dogs example with 1000 training samples, 500 validation samples, and 500 test samples yielded a relatively low accuracy of 73%. Because the training sample is tiny, this is an example of overfitting.
* By combining multiple strategies, we can improve the model's performance while keeping the sample size constant at 1000. In the model, I utilized three approaches to do this.

1)Dropout Method

2)Data Augmentation

3)Data Augmentation and dropout method.

• It was discovered that the model trained using data augmentation and the dropout strategy performed better.

• Increase accuracy by training with additional data. We tried increasing the training samples to 5000 and 10000, which resulted in better accuracy.

• When the training sample size was raised to 10000, the validation accuracy rose to 98%. I've used Max pooling, Data Augmentation, a dropout approach with a dropout rate of 0.5, and Early halting.

• The validation accuracy rose to 97% when we raised the training sample size to 50000.

• Regularized models appeared to be more accurate than unregularized ones.

• We can see that when models with training sizes of 1000, 5000, and 10000 were pre-trained, both validation and test accuracy increased.

• In both scratch and pre-trained models, the training sample size of 10000 provided the maximum accuracy.

**Conclusion:**

To summarize, the size of the training sample is significant in enhancing model accuracy since it prevents overfitting. Furthermore, hyper tuning factors like as max-pooling, data augmentation, and the dropout approach aid in further boosting the model's performance.

We see a considerable increase in accuracy when the models are pre-trained, thus we infer that pretraining the model, together with the training sample size, has a major effect.

1.The Model 1 unregularized Model of Cats and Dogs example with 1000 training samples, 500 validation samples, and 500 test samples has a relatively poor accuracy of 73%. Because the training sample is tiny, this is an example of overfitting.

2.By combining several strategies, we may improve the model's performance while keeping the sample size constant at 1000. For this, I utilized three strategies on the model: a) Drop out Technique

b) Data Enrichment

c) The data augmentation and drop out approach.

3.It was discovered that the model trained utilizing data augmentation and the dropout strategy performed better.

4.Train with additional data: Training with more data improves accuracy. We tried increasing the training samples to 5000 and 10000, and the accuracy increased.

5.When the models were pretrained, the accuracy increased to about 99%.