**ICP4**

**BIG DATA ANALYTICS AND APPLICATIONS(CS5542)**

*Sravani Niharika Garapati (16328843)*

*Nikhila Chirumamilla (16323180)*

**QUESTION:**

Use the same data (that we used in ICP3 from keras.datasets import cifar10)) and use the model provided in ICP4 to perform image classification. You must change 4 hyper parameters in the source code. Report your findings in detail.

**What I have learned from this ICP:**

In this ICP, professor asked us to perform image classification on the CIFAR10 data set. In this ICP, we have created 4 deep learning which are interdependent on each other by changing 4 hyperparameters. The things which I have learned from this ICP are listed below

* I have learned many functions that were very useful in implementing the model and finding accuracy and many plotting functions.
* I have known which model and which layers to be used in the situation.
* I have known in detail about CIFAR 10 data set and to perform analysis on that data set in a thorough manner.
* Learnt in detail about normalizing and one hot encoding in a detailed manner.
* I have come to know that increasing number of epochs while fitting the model would give more accuracy.
* We have learnt more about accuracy and loss which are plotted as training and validation accuracy and training and validation loss
* In this ICP, we used SoftMax activation function and DropOut layers are hyperparameters. So, I have known much about them and learnt when and where to use them in the problem in future
* Learnt much about the layers of Conv2D, MaxPooling2D, Dense, Dropout which we have used in this ICP solution.
* At last, I have learned how to perform image classification using keras to a certain extent.

**ICP Description about the Task:**

In this ICP, we performed image classification on CIFAR 10 data set using Keras model. This is done in the following steps.

* As first step, I have imported all the necessary libraries
* Then, Loaded the CIFAR 10 data set for image classification
* Then performed the normalization on data to rescale them to range (0,1) by dividing with 255.0
* Performed One-Hot-Encoding on the data by using to\_categorical () function to convert the integer into 10 element binary vector
* Then built the Model1 for the data set

**Model-1**

* Used Sequential () Model, which is the best fit for the data
* Used 3 Conv2D layers each of 16, 32, 64 size
* Used 3 MaxPooling2D layers
* Used 2 Dense Layers for which I have passed number of classes in the dataset and for the other one the size is 128
* The activation functions used are “relu”
* Used Flatten layer
* ***Accuracy Observed***: 10%
* Compiled the Model-1 and has fit the model to the data set and has drawn plot for accuracies and loss
* Then built the Model2 for the data set

**Model-2(Hyperparameter 1)**

* Built Same model as Model-1
* Included “SoftMax” activation function in the Dense Layer
* We included “SoftMax” activation function because it is used when there are multiple layers for better performance
* ***Accuracy Observed:*** 72.70%
* Built the Model-3 for the data set as there is more training and validation loss incurred in the model 2

**Model-3(Hyperparameter 2)**

* Built Same model as Model-2
* Included an extra DropOut Layer with 0.2 score
* This DropOut layer resets the value to 0, which helps in preventing overfitting
* This reduced the training and validation loss to an extent which resulted in reduced accuracy
* ***Accuracy Observed:*** 70.31%
* Built the Model-4 for the data set to increase the accuracy of the model

**Model-4(Hyperparameter 3 & Hyperparameter 4)**

* Built the same model as Model-3
* Included multiple DropOut layers after each MaxPooling2D layer to reset values at each step
* Also increased number of epochs from 10 to 40 to improve accuracy which is considered as 4th hyperparameter
* ***Accuracy Observed:*** 74.61%
* Then downloaded some of the images from internet and predicted the result

**Challenges Faced in this ICP:**

The challenges which we faced while doing this ICP are listed below:

* It has become very difficult for us to identify the hyper parameters
* Also, it has been very difficult for us to understand and analyze the data

**Screenshots of Execution of Code:**

*Importing necessary libraries and CIFAR 10 dataset description*

**Graphical user interface, text, application, email

Description automatically generated**

*Loading the dataset*

**Graphical user interface, text, application, email

Description automatically generated**

*Printing the shapes and contents of the loaded data*

**Graphical user interface, text, application, email

Description automatically generated**

*Depicting number of classes and their names*

Graphical user interface, text

Description automatically generated with medium confidence

*Plotting first few images in the data set*

Graphical user interface, application

Description automatically generated

*Normalizing the data*

Graphical user interface, text, application

Description automatically generated

*Printing the normalized data*

*Text

Description automatically generated*

*Performing One Hot Encoding*

*Graphical user interface, text, application

Description automatically generated*

*Printing the data after performing one hot encoding*

*Table

Description automatically generated with low confidence*

*Graphical user interface, text, application

Description automatically generated*

*Building Model 1*

*Graphical user interface, text, application

Description automatically generated*

*Compiling the model and describing the summary of the model*

*Text

Description automatically generated with low confidence*

*Fitting the model by specifying the epochs and batch size*

*A picture containing text

Description automatically generated*

*Text

Description automatically generated*

*Plot for accuracies and loss for model 1*

*Chart

Description automatically generated with medium confidence*

*Calculating accuracy for Model 1*

*Text

Description automatically generated*

*Building Model 2*

*Text

Description automatically generated*

*Compiling the Model 2 and fitting the model to the dataset*

*Text

Description automatically generated with medium confidence*

*Accuracy and loss plot for model 2*

*Chart

Description automatically generated*

*Calculating accuracy for model 2*

*Graphical user interface, text

Description automatically generated*

*Building Model 3*

*Text

Description automatically generated*

*Compiling and fitting the model 3 to the dataset*

*Graphical user interface

Description automatically generated with medium confidence*

*Chart, line chart

Description automatically generated*

*Calculating accuracy for model 3*

*Text

Description automatically generated*

*Building Model 4*

*Graphical user interface, text, application

Description automatically generated*

*Compiling and fitting the model 4 to the data set*

*Table

Description automatically generated with medium confidence*

*Accuracy and Loss plot for Model 4*

*Histogram

Description automatically generated*

*Calculating the Accuracy of the whole finalized Model*

*Text

Description automatically generated*

*Importing the necessary libraries*

*Text

Description automatically generated*

*A picture containing text, bird, parrot, colorful

Description automatically generated*

*A cat sitting in a basket

Description automatically generated with low confidence*

*A picture containing text, fox, mammal

Description automatically generated*

*A picture containing text, frog

Description automatically generated*

*A horse running in a field

Description automatically generated with medium confidence*

*A screenshot of a plane

Description automatically generated with low confidence*

*A picture containing text, mammal

Description automatically generated*

*Graphical user interface, Word

Description automatically generated*

*A picture containing text

Description automatically generated*

*A picture containing graphical user interface

Description automatically generated*

**Video Link**