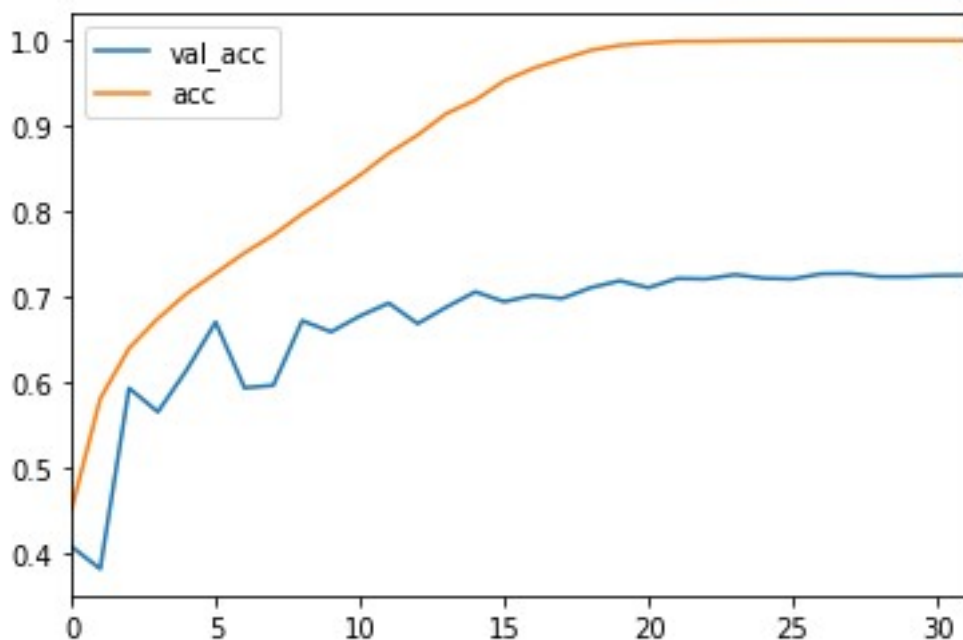


Name – Sameer Vivek Pande
Entry No – 2017CS10371

1.) Part a:- Model Submitted

Optimizer = 'sgd' (default keras optimizer for sgd)
epochs = 20,
batch size = 100
validation_split = 0.1
loss = categorical_crossentropy

validation accuracy = 72.58%
train accuracy = 100%



2.) Part b:

Model Submitted – VGG Net

reference:- <https://github.com/geifmany/cifar-vgg>

Here is the summary of the model used

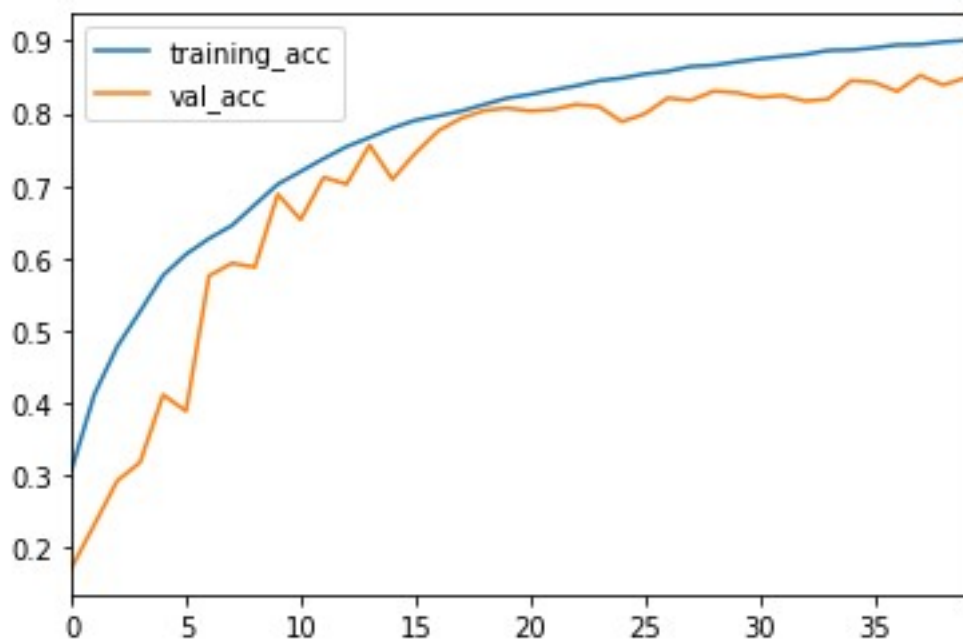
Layer (type)	Output Shape	Param #
=====		
conv2d_1 (Conv2D)	(None, 32, 32, 64)	1792
batch_normalization_1 (Batch Normalization)	(None, 32, 32, 64)	256
dropout_1 (Dropout)	(None, 32, 32, 64)	0
conv2d_2 (Conv2D)	(None, 32, 32, 64)	36928
batch_normalization_2 (Batch Normalization)	(None, 32, 32, 64)	256
max_pooling2d_1 (MaxPooling2D)	(None, 16, 16, 64)	0
conv2d_3 (Conv2D)	(None, 16, 16, 128)	73856
batch_normalization_3 (Batch Normalization)	(None, 16, 16, 128)	512
dropout_2 (Dropout)	(None, 16, 16, 128)	0
conv2d_4 (Conv2D)	(None, 16, 16, 128)	147584
batch_normalization_4 (Batch Normalization)	(None, 16, 16, 128)	512
max_pooling2d_2 (MaxPooling2D)	(None, 8, 8, 128)	0
conv2d_5 (Conv2D)	(None, 8, 8, 256)	295168
batch_normalization_5 (Batch Normalization)	(None, 8, 8, 256)	1024
dropout_3 (Dropout)	(None, 8, 8, 256)	0
conv2d_6 (Conv2D)	(None, 8, 8, 256)	590080
batch_normalization_6 (Batch Normalization)	(None, 8, 8, 256)	1024
dropout_4 (Dropout)	(None, 8, 8, 256)	0
conv2d_7 (Conv2D)	(None, 8, 8, 256)	590080
batch_normalization_7 (Batch Normalization)	(None, 8, 8, 256)	1024
max_pooling2d_3 (MaxPooling2D)	(None, 4, 4, 256)	0

conv2d_8 (Conv2D)	(None, 4, 4, 512)	1180160
batch_normalization_8 (Batch Normalization)	(None, 4, 4, 512)	2048
dropout_5 (Dropout)	(None, 4, 4, 512)	0
conv2d_9 (Conv2D)	(None, 4, 4, 512)	2359808
batch_normalization_9 (Batch Normalization)	(None, 4, 4, 512)	2048
dropout_6 (Dropout)	(None, 4, 4, 512)	0
conv2d_10 (Conv2D)	(None, 4, 4, 512)	2359808
batch_normalization_10 (Batch Normalization)	(None, 4, 4, 512)	2048
max_pooling2d_4 (MaxPooling2D)	(None, 2, 2, 512)	0
flatten_1 (Flatten)	(None, 2048)	0
dense_1 (Dense)	(None, 512)	1049088
batch_normalization_11 (Batch Normalization)	(None, 512)	2048
dropout_7 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 10)	5130
=====		
Total params: 8,702,282		
Trainable params: 8,695,882		
Non-trainable params: 6,400		

The regularization parameter(weight decay) for all the layers is 0.00025
The dropout value is set to 0.5 for all dropout layers

batch_size = 100
epochs = 40
optimizer = SGD(learning rate = 0.008, decay = 1e-6, momentum=0.9,nesterov=True)
loss = 'categorical_crossentropy'

Validation Accuracy :- 84.8 %
Training Accuracy :- 90%



Other models tried:-

1.) VGG Net with regularization set to 0.0008 and dropout to 0.5, rest parameters similar to original model

after 40 epochs:-

validation accuracy = 81 %, Training Accuracy = 82.5%

2.) VGG Net with regularization set to 0.0005 and dropout 0.4 for all layers, rest similar to original model

after 40 epochs:

validation accuracy = 84 %, Training Accuracy = 93 %

(Though accuracy of this model is similar to our model but training accuracy for our model is slightly lower than 93 %. Hence our model is better since less overfit)

3.) VGG Net with regularization set to 0.001 and rest same

validation accuracy = 81.8 % , Training Accuracy = 92.8 %

4.) AlexNet for 15 epochs with l2 regularization parameter 0.01

validation accuracy = 67 % , training accuracy = 97%

5.) AlexNet for 15 epochs with no regularization

validation accuracy = 73.44 % , Training Accuracy = 97.60 %