

CIFAR -10 Image Classification problem

The aim is to build a multilayer perceptron which will classify test images in 10 different categories(classes) using different hyperparameters like number of hidden layers its nodes, epochs, batch size, optimizer function, activation function n etc. So basically, comparison of different MLP's.

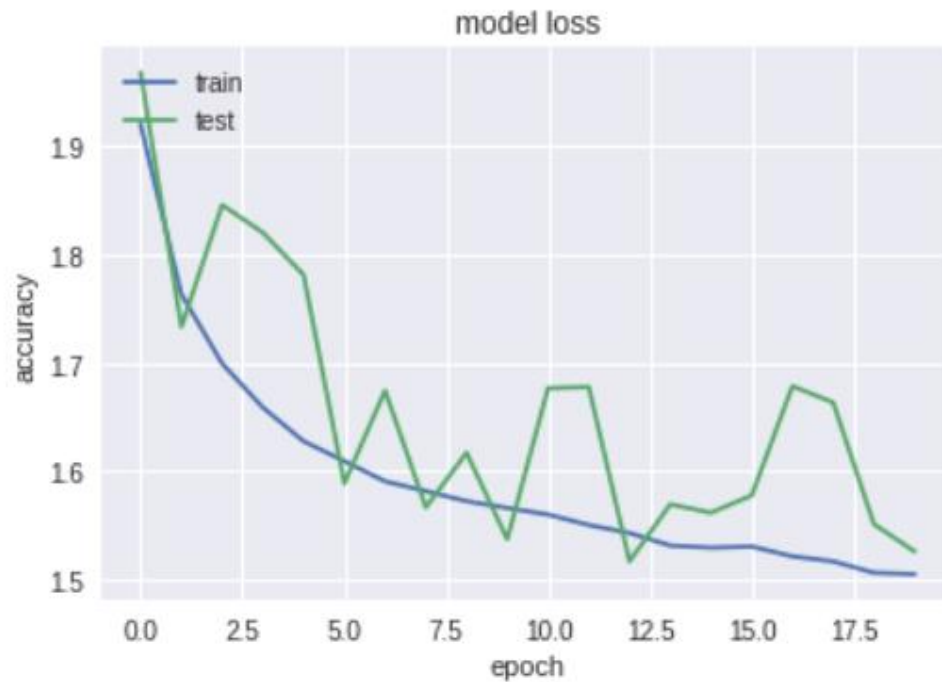
The best four models out of all the implemented 10 to 15 models are:

SR. NO	Epochs	Batch Size	Hidden Layers	Optimizer	Test Accuracy	Test Loss	Time Per Epoch GPU	Time per Epoch CPU
1	20	64	Sigmoid-512, Sigmoid-256	rmsprop	0.4565	1.52593429	6s	28s
2	200	128	Tanh-2048, tanh-1024, tanh-512	rmsprop	0.3442	1.887073356	10s	90s
3	40	64	Softmax-1024, softmax-512, softmax-256	SGD	0.2702	1.981236296	9s	93s
4	20	64	Relu - 2048 Relu-1024 Relu-512	SGD	0.5402	1.278560466	9s	168s

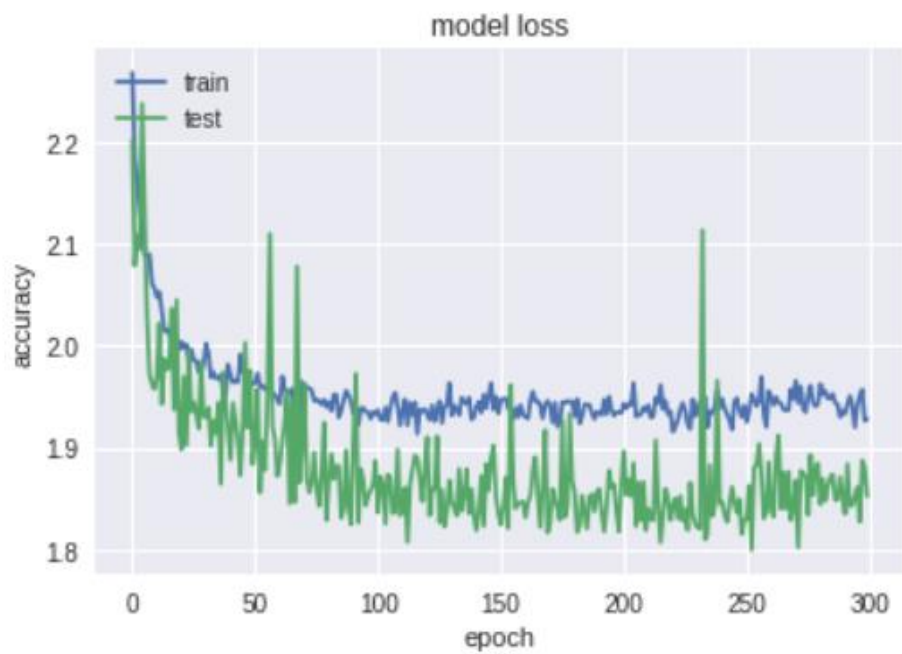
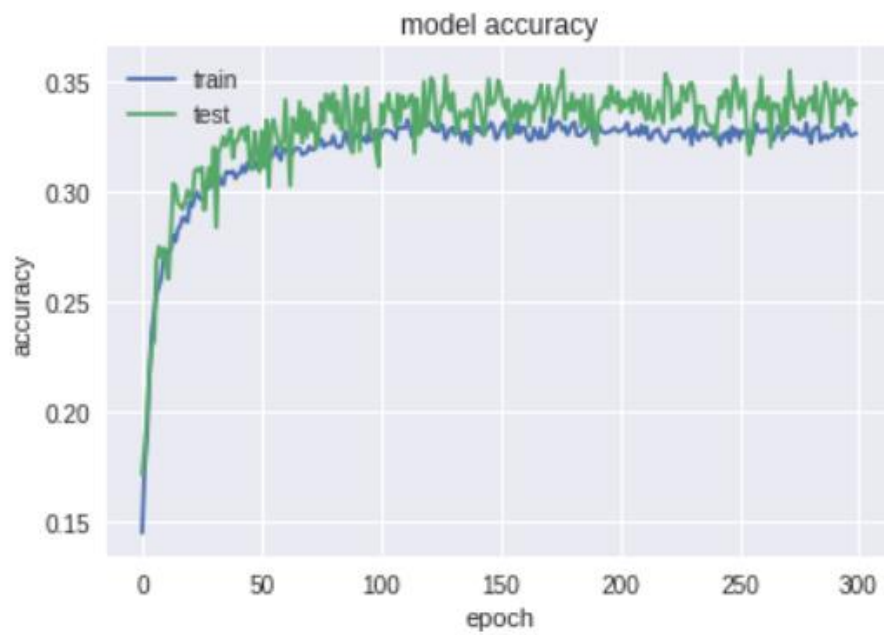
Clearly, GPU is much better than CPU saving a lot of computational time.

Sigmoid

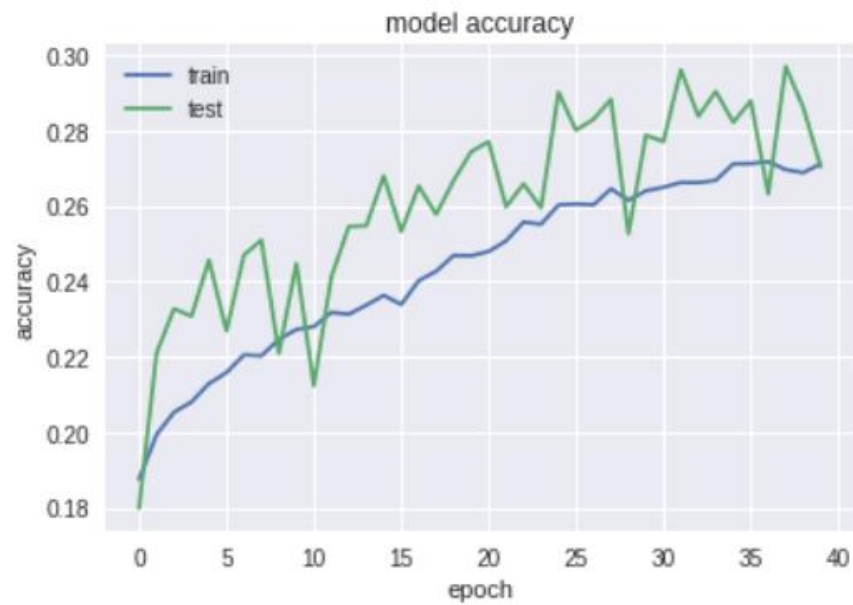
Training performing better than test problem of Overfitting might be an issue



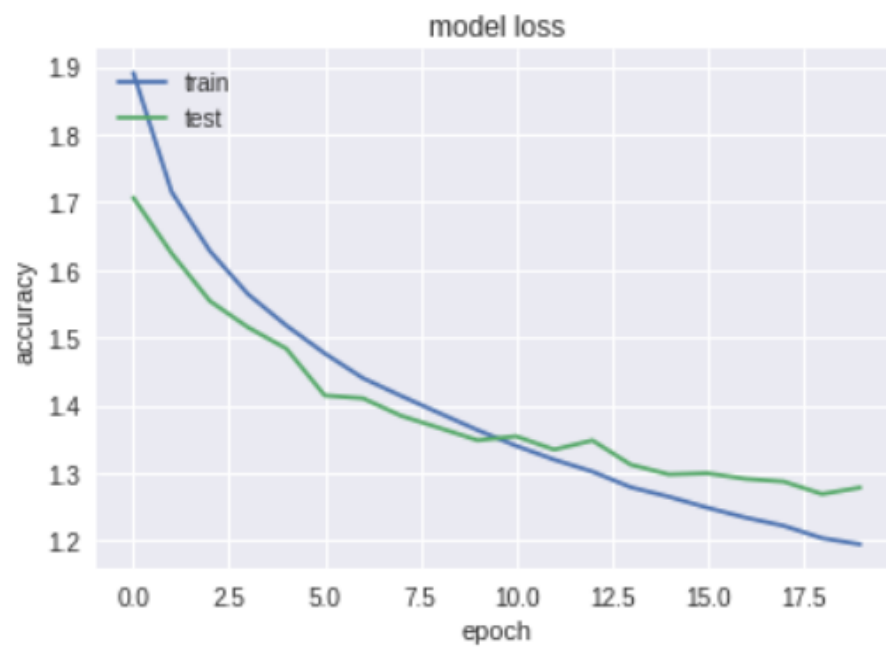
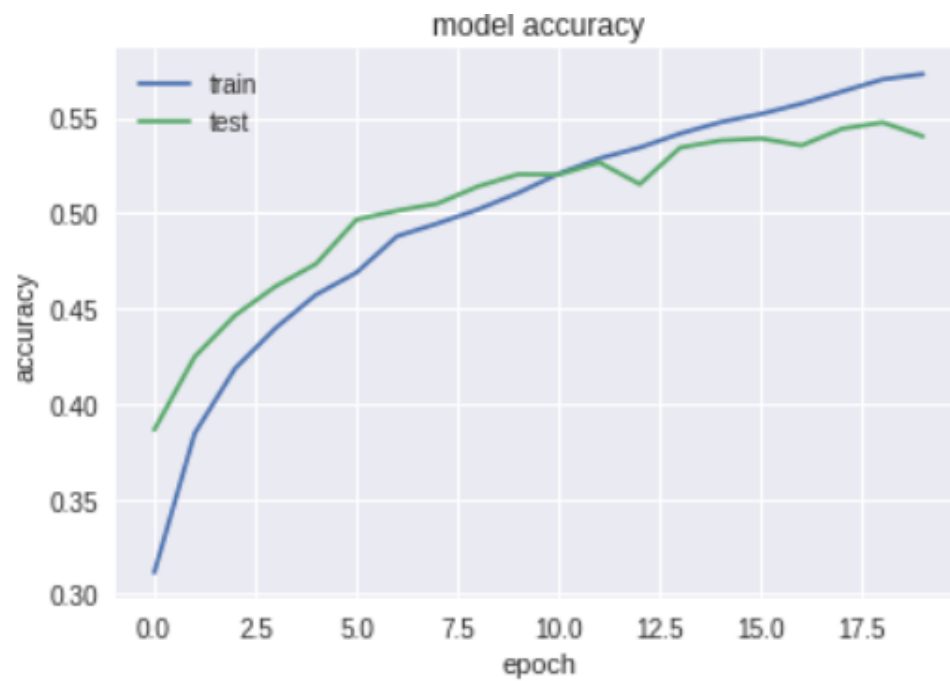
Tanh



Softmax



Relu



The best model according to me is the model with best accuracy which is 54% whose parameters are:

1. Hidden layers = 3
2. hidden_num_units = 2048, 1024, 512
3. output_num_units = 10
4. epochs = 20
5. batch_size = 64
6. output activation function = 'Softmax'
7. layer activation function = "Relu"
8. optimization function = "SGD"

The reason why I feel this is the best model is firstly RELU is used as an activation function for hidden layers which is widely used for image classification problems or CNN. Secondly, the output activation function is Softmax which will give probabilities for n number of classes (which is 10 in this Cifar-10 dataset) and the highest probability is the expected result. Thirdly, the learning rate is used to get an optimal accuracy using gradient descent to reduce the cost function. Finally, the batch-size is small which captures details as it learns slow and thus making the accuracy to increase.

My model is average it doesn't overfit or underfit as we can see that through accuracy plot above under RELU. To improve the model, we can try different combinations of epochs and batch sizes to achieve for a higher accuracy of the model. I tried increasing the epochs for the model, but overfitting was an issue, training was performing better than test. Another solution could be a Convolutional Neural Networks which is considered the best neural network for problems of image classification.