Sameer Vivek Pande
2017CS10371

Report Part-c

Best Performing Design :

Batch Size = 200
Learning Rate = 1 (adaptive)
Activation = Softplus (last layer softmax)
number of Hidden Layers = 1
Nodes in Hidden Layer = 120
Regularziation Type = L2
Regularization Parameter = 0.1

Other observations

Number of Iterations = 15000

Here are some the results of multiple tests done

Batch Size = 100

Learning Type	Learning Rate (base rate)	Hidden Layers	Regularization Parameter	Activation	Best Validation Accuracy % (Approx.)	Training : Validation Split
1	0.5	[200,200]	0.1	SoftPlus	38.1	9:1
1	0.5	[100,100]	0.1	SoftPlus	33.7	9:1
2	1	[120]	0.1	SoftPlus	35	9:1
1	0.5	[100,100]	0.1	Relu	15	9:1

1	0.5	[100,100]	0.1	Leaky_Relu	31	9:1
1	0.5	[100,50,10]	0.01	Sigmoid	16	9:1
1	0.5	[100,100,10]	0.1	Sigmoid	15	9:1
2	10	[50]	0.01	Sigmoid	28	9:1
2	10	[50]	0.01	Tanh	26	9:1
1	0.5	[100,100]	0.1	Tanh	33	9:1

Many were observations were recorded:

1. In general regularization parameter of 0.1 performed better than 0.01 . Using regularization parameters in range like 1 , 2 etc. gave poor results with all activations

2. Three hidden layer models overfitted extremely with all activation functions and gave best val accuracy of arround 16 % for sigmoid and softplus activations

3. Softplus performs bettern than tanh. Tanh has been seen to perform better than sigmoid functions
4. Leaky Relu has performed better but unable to been SoftPlus
5. Using two layers led to overfitting without improving validation accuracy by a lot. Instead adding more neurons in a single layered model increased validation accuracy

Batch Size = 200

Learning Type	Learning Rate (base rate)	Hidden Layers	Regularization Parameter	Activation	Best Validation Accuracy % (Approx.)	Training : Validation Split
2	1	[120]	0.1	SoftPlus	39	9:1
2	1	[120]	0.1	Sigmoid	32	9:1
2	1	[120]	0.1	Tanh	34	9:1

1. For batch size 200, adaptive learning rates worked performed by keeping max iterations same
2. Using base learning rate 1 gives better results for max_iterations 15000
3. Comparing three activation functions for best result obtained (used in final submission)

Report Part -d

Best Working Design

Learning Type = Adaptive
Base Learning Rate = 1
Number of iterations = 15000
Batch size = 120
Activation Type = SoftPlus
Filter Used = Gaber (Theta = 45 degrees, Frequency = 0.25)
Feature Added = Apply Gaber filter on each image and using the image as a feature

(E.g. 32 X 32 image gave 1024 sized feature vector. Now It will give 2048 sized vector. Just by appending the vector obtained by applying filter on the image)

1. Tried different filters – Gaber Filters, Sobel Filter, Canny Filter (edge detection)
2. For the best working model parameters

Gaber Filter -> Increases Accuracy by (1-1.5)%

Sobel Filter -> Accuracy almost same as without filter

Canny Filter -> Accuracy of model reduced (perhaps required more training or different model to give better accuracy)

Batch Size = 200

Learning Type	Learning Rate (base rate)	Hidden Layers	Regularization Parameter	Activation	Best Validation Accuracy % (Approx.)	Filter
2	1	[120]	0.1	SoftPlus	40	Gaber
2	1	[120]	0.1	SoftPlus	39	Sobel
2	1	[120]	0.1	SoftPlus	34	Canny