Due No Due Date

CODING DRILL DOWN

In the last session, we covered a lot of basics. Your target was to achieve 99.4% Test Accuracy within 20 Epochs while using less than 20k

In this session, we'll go through 10 Code Iterations to help us understand how do we target such a problem.

2. Set Transforms

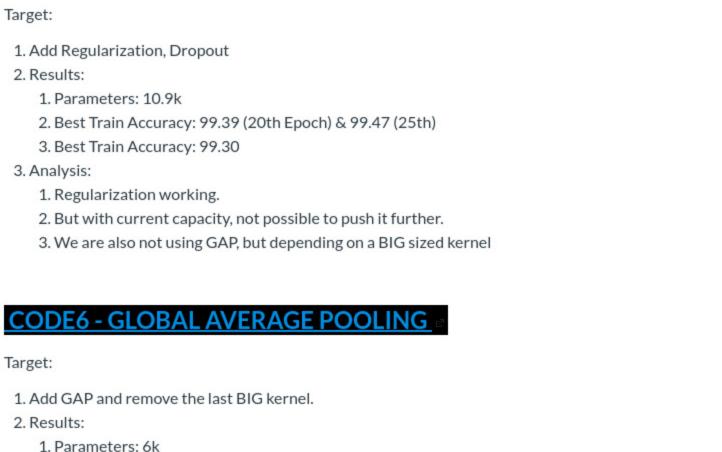
- 3. Set Data Loader
- 4. Set Basic Working Code Set Basic Training & Test Loop
- 6. Results: 1. Parameters: 6.3M 2. Best Training Accuracy: 99.99
- 3. Best Test Accuracy: 99.24 7. Analysis:
- Target:

3. Best Test Accuracy: 99.02

- 4. Analysis: 1. The model is still large, but working.
- Target:
- 2. Results: 1. Parameters: 10.7k

- CODE4 BATCHNORM
- 2. Results:
- 3. Analysis: 1. We have started to see over-fitting now.
- **CODE5 REGULARIZATION**

abhi kabhi lagta hai



- 2. Best Train Accuracy: 99.86
 - 2. We are comparing a 10.9k model with 6k model. Since we have reduced model capacity, reduction in performance is expected.

CODE7 - INCREASE CAPACIT

3. Best Test Accuracy: 98.13

- Increase model capacity. Add more layers at the end. 2. Result:
- 2. Best Train Accuracy: 99.33 3. Best Test Accuracy: 99.04 3. Analysis:
- 3. Closer analysis of MNIST can also reveal that just at RF of 5x5 we start to see patterns forming. 4. We can also increase the capacity of the model by adding a layer after GAP!
- CODE8 VOILA!

2. Perform MaxPooling at RF=5

Fix DropOut, add it to each layer

3. The model is not over-fitting at all.

3. Best Test Accuracy: 99.41 (9th Epoch) Analysis: 1. Works!

2. But we're not seeing 99.4 or more as often as we'd like. We can further improve it.

1. The model still showing over-fitting, possibly DropOut is not working as expected! Wait yes! We don't know which layer is causing

CODE9 - IMAGE AUGMENTATION

1. Parameters: 13.8k 2. Best Train Accuracy: 99.15 3. Best Test Accuracy: 99.5 (18th Epoch)

3. Analysis:

Target:

2. Results:

1. Add LR Scheduler

1. Parameters: 13.8k

- 3. Analysis: 1. Finding a good LR schedule is hard. We have tried to make it effective by reducing LR by 10th after the 6th epoch. It did help in getting to 99.4 or more faster, but final accuracy is not more than 99.5. Possibly a good scheduler can do wonders here! DISCIPLINE It's not that I'm so smart, it's just that I stay with problems longer.

~ Albert Einstein

further.net

Designing Models require discipline

Every step you take must have a purpose

Trying too many things without order or without any notes is useless

RECEPTIVE FIELD CALCULATIONS

 $n_{out} = \left| \frac{n_{in} + 2p - k}{\varsigma} \right| + 1$

convolution stride size $n_{out} = \left\lfloor \frac{n_{in} + 2p - k}{s} \right\rfloor + 1$

 $r_{out} = r_{in} + (k-1) * j_{in}$

- 3×1 3x3 conv, 512 3×1 Pool/2 2×2
- 3x3 conv, 512 3×1 3x3 conv, 512 3×1 Pool/2 2 x 2
- 3×1 3x3 conv, 256 Pool/2
- 1. 99.4% (this must be consistently shown in your last few epochs, and not a one-time achievement)
- 4. You must convince why have you decided that your target should be what you have decided it to be, and your analysis MUST be correct. 5. Evaluation is highly subjective, and if you target anything out of the order, marks will be deducted. 6. Explain your 5 steps using these target, results, and analysis with links to your GitHub files
- 7. Keep Receptive field calculations handy for each of your models. 8. When ready, attempt S5-Assignment Solution

- Parameters.

CODE 1 - SETUP Target: 1. Get the set-up right

1. Extremely Heavy Model for such a problem Model is over-fitting, but we are changing our model in the next step **CODE 2 - BASIC SKELETON**

1. Get the basic skeleton right. We will try and avoid changing this skeleton as much as possible. Results: 1. Parameters: 194k 2. Best Train Accuracy: 99.35

2. We see some over-fitting

- - <u> CODE3 LIGHTER MODEL</u>

2. Even if the model is pushed further, it won't be able to get to 99.4

- 1. Make the model lighter 2. Best Train Accuracy: 99.00
- 3. Best Test Accuracy: 98.98

3. Analysis: 1. Good model! 2. No over-fitting, model is capable if pushed further

1. Add Batch-norm to increase model efficiency.

Target:

- 1. Parameters: 10.9k 2. Best Train Accuracy: 99.9 3. Best Test Accuracy: 99.3

- 1. Adding Global Average Pooling reduces accuracy WRONG

3. Analysis:

- Target:
- 1. Parameters: 11.9k
- over-fitting. Adding it to a specific layer wasn't a great idea. 2. Quite Possibly we need to add more capacity, especially at the end.
- Target: Increase model capacity at the end (add layer after GAP)
- 1. Parameters: 13.8k 2. Best Train Accuracy: 99.39

4. Results:

- 4. Seeing image samples, we can see that we can add slight rotation.
- Target: 1. Add rotation, our guess is that 5-7 degrees should be sufficient. 2. Results:
 - 1. The model is under-fitting now. This is fine, as we know we have made our train data harder. 2. The test accuracy is also up, which means our test data had few images which had transformation difference w.r.t. train dataset

CODE 10 - OVERKILL

- - 2. Best Train Accuracy: 99.21 3. Best Test Accuracy: 99.45 (9th Epoch), 99.48 (20th Epoch)

Kabhi kabhi lagta hai apun hi bhagwan hai

 n_{in} : number of input features n_{out} : number of output features convolution kernel size convolution padding size

 $j_{out} = j_{in} * s$

output

FC8 4096

FC7 4096

FC6 4096

3x3 conv, 512

3x3 conv, 512

3x3 conv, 256

3x3 conv, 128

3x3 conv, 64

3x3 conv, 64

input

25088

7*7*512

14*14*512

14*14*512

28*28*512

28*28*256

56*56*256

56*56*128

112*112*128

112*112*64

224*224*64

224*224*3

Assignment:

1. Your new target is:

p:

s:

Pool/2 2×2 3x3 conv, 512 3 x 1

7 x 1

 3×1

 3×1

 3×1

 2×2

 3×1

 3×1

receptive

field

404 x 404

212 x 212

196 x 196

164 x 164

132 x 132

100 x 100

92 x 92

 76×76

60 x 60

44 x 44

 40×40

 32×32

 24×24

16 x 16

14 x 14

10 x 10

6 x 6

 5×5

 3×3

(k, s)'

 404×32

212 x 32

196 x 16

164 x 16

132 x 16

100 x 16

 82×8

76 x 8

60 x 8

44 x 8

40 x 4

 32×4

24 x 4

16 x 4

 14×2

10 x 2

6 x 2

5 x 1

 3×1

- 3x3 conv, 256 3×1
- 2×2 3x3 conv, 128 3×1

Pool/2

- 2. Less than or equal to 15 Epochs 3. Less than 10000 Parameters 2. Do this in minimum 5 steps 3. Each File must have "target, result, analysis" TEXT block (either at the start or the end)