Q10

Due Jun 7 at 6:30am **Points** 200 **Questions** 11 **Available** May 31 at 9am - Jun 7 at 6:30am 7 days
Time Limit 30 Minutes

Instructions

You have 30 minutes to solve the quiz.

This quiz was locked Jun 7 at 6:30am.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	9 minutes	160 out of 200

Score for this quiz: **160** out of 200
Submitted Jun 6 at 9:08pm
This attempt took 9 minutes.

Question 1 — 10 / 20 pts

Which of the following are true?

Correct Answer — ☐ Comparatively, at the midele of the training, learning rate can be much higher than at later stages
Correct! — ☑ While updating the weights of one kernel, we must assume other kernel is constant
Correct! — ☑ For gradient ascent we take positive value of the gradients
Correct Answer — ☐ A high dropout value can lead to lot of fluctuations at the later stages of training in VA

Question 2 — 20 / 20 pts

Which of the following are true?

Correct! — ☑ As the VA increases, we should increase the batch size keeping the learning rate constant
Correct! — ☑ As the VA increases, we should reduce the learning rate, keeping batch size constant
Correct! — ☑ Even if we use momentum with SGD, the learning rate remains constant.
Correct! — ☑ In SDG (academically), batch size is 1

Question 3 — 10 / 20 pts

It is proven that if we add gradient perturbation (small noise in gradients), we can avoid hitting the problem of weights getting stuck in plateaus.

Which of the following can have a similar effect:

You Answered — ☑ L1/L2 regularization
☐ ReLU
Correct! — ☑ Patch Gaussian
Correct! — ☑ Dropout

Question 4 — 20 / 20 pts

Match the following: Left is the problem, right is the solution

Correct! — **Weight Plateaus** — Gaussian Noise
Correct! — **UnderFitting** — Remove Dropout
Correct! — **OverFitting** — Image Augmentation
Correct! — **Slow Convergence** — Increase LR

Other Incorrect Match Options:
- CutOut
- SGD
- Reduce LR

Question 5 — 20 / 20 pts

Assume that the value of a specific weight was 4.

The derivative of the Loss Function w.r.t. this weight is 100

If we used a learning rate of 0.01, after the backprop step, what would be the value of the new weight?

Correct! — 3

Correct Answers — 3.0
3

Question 6 — 10 / 10 pts

We are working on a custom dataset, where we have 10 classes, but only 100 images for each class. Which Optimization Algorithm should provide us better results?

☐ SGD
☐ SGD with reducing learning rates
Correct! — ☑ Adam
☐ SGD with Momentum

Question 7 — 10 / 10 pts

In the momentum algorithm, what would be the value of the $v^{(0)}$ for the very first time?

☐ Would be required to be calculated
Correct! — ☑ 0
☐ would be set to a random value

Question 8 — 20 / 20 pts

What all would be the benefits of adding momentum term to SGD?

Correct! — ☑ Faster Convergence
Correct! — ☑ Solving weight saddling problem
Correct! — ☑ Solving weight plateauing problem

Question 9 — 10 / 20 pts

We "know" for sure that we are stuck in local minima. What all could we try?

Correct! — ☑ Add momentum to our optimizer if we haven't done so yet
Correct Answer — ☐ Change the optimizer for sometime
Correct! — ☑ Increase the learning rate
Correct Answer — ☐ Add image augmentation

Question 10 — 10 / 20 pts

We defined a network and then ran an LR finder on it.

After a few tests on learning rates, we do not see any change in the loss function. What all could be wrong?

Correct Answer — ☐ We haven't yet tried varied range of possible LRs
Correct! — ☑ Network is incapable of learning

Question 11 — 20 / 20 pts

Match the following

Correct! — **SGD with Momentum** — Most DNNs
Correct! — **Adaptive Optimizers** — GANs & RL

Other Incorrect Match Options:
- NLP
- Object Detection Networks

Quiz Score: **160** out of 200