

Resnet-34

...

Model

1. Used resnet-34
2. Freezed all layers
3. Replaced the fc layer with our classification layer. We have two output classes.

Training

Epochs: 20

Optimizer: Adam

Lr Scheduling: cosine annealing

Optimizer

1. **Loss** : tells about how poorly the model is performing.
2. **Aim**: Minimize the loss.
3. **Optimization** : Process of minimizing the loss function.
4. **Learning rate**: How big each step we move in the direction of gradient.
5. Optimizers after every step it updates the parameters.

Adam

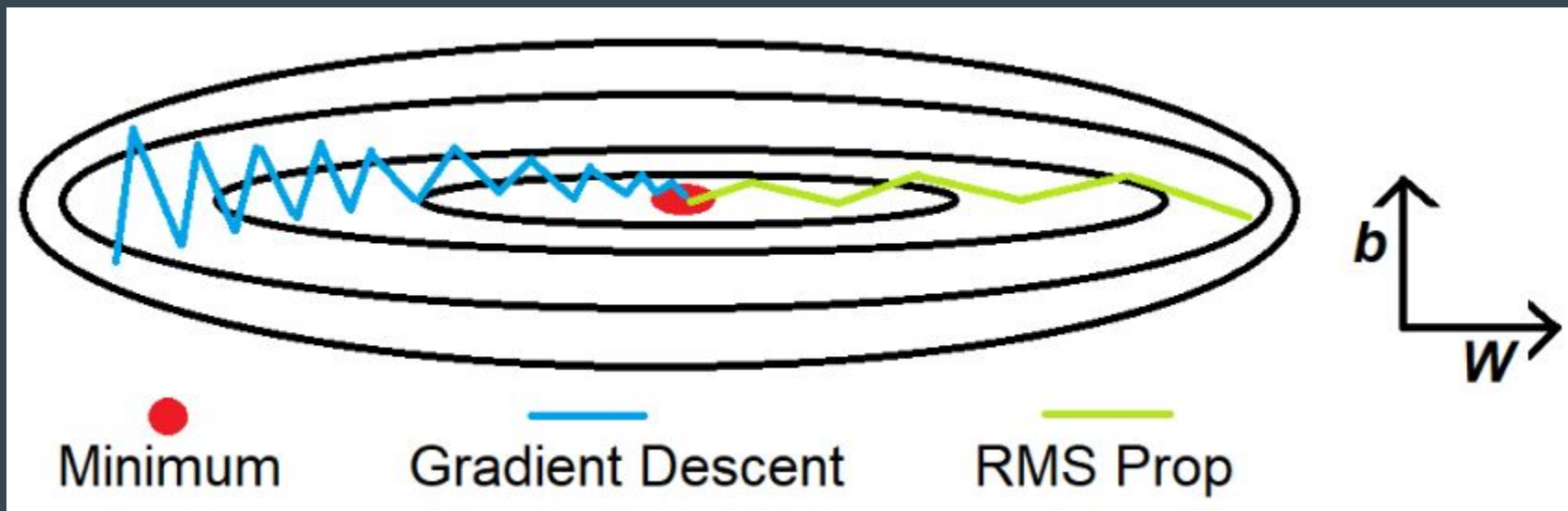
1. Adaptive moment estimation
2. Combination of momentum and rmsprop
3. Very effective
 - a. **Two ideas:**
 - i. If update params batch wise then noise comes into play.
 - ii. To denoise we use the idea of momentum.
 - iii. Adaptive learning rate - based on gradient it decided what lr needs to be set.
 1. If gradient is high then lr is low.
 2. If gradient is low then Lr is high.
4. Commonly used optimizer.

Momentum

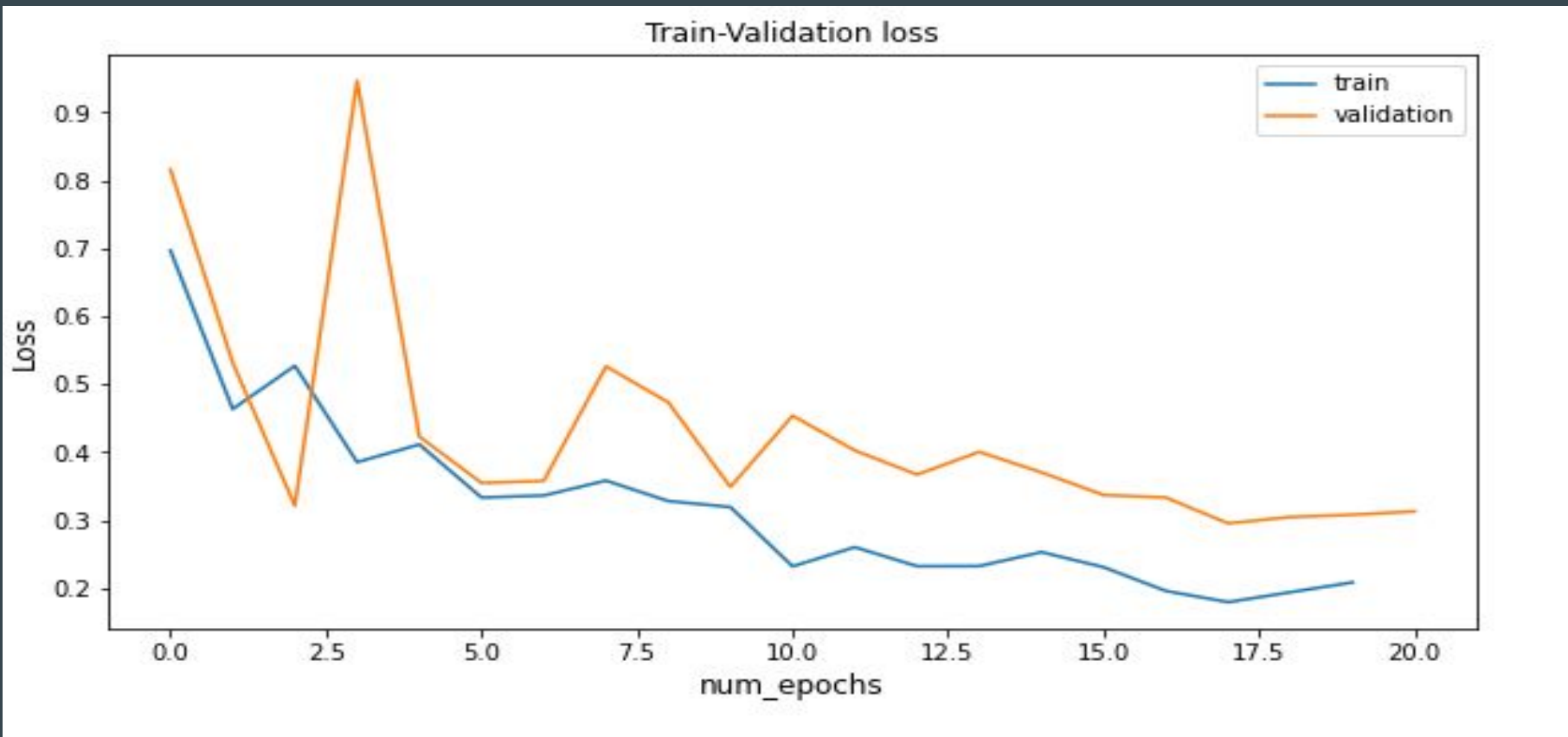
1. Overcomes the drawback of MB-SGD.
2. Previous update is added to the current update.
3. So repeated updates compounds in a particular direction. This avoids that oscillation.
4. Similar to momentum concept in physics.
5. This helps in convergence towards relevant direction and reduces the fluctuation to the irrelevant direction.

RMSprop

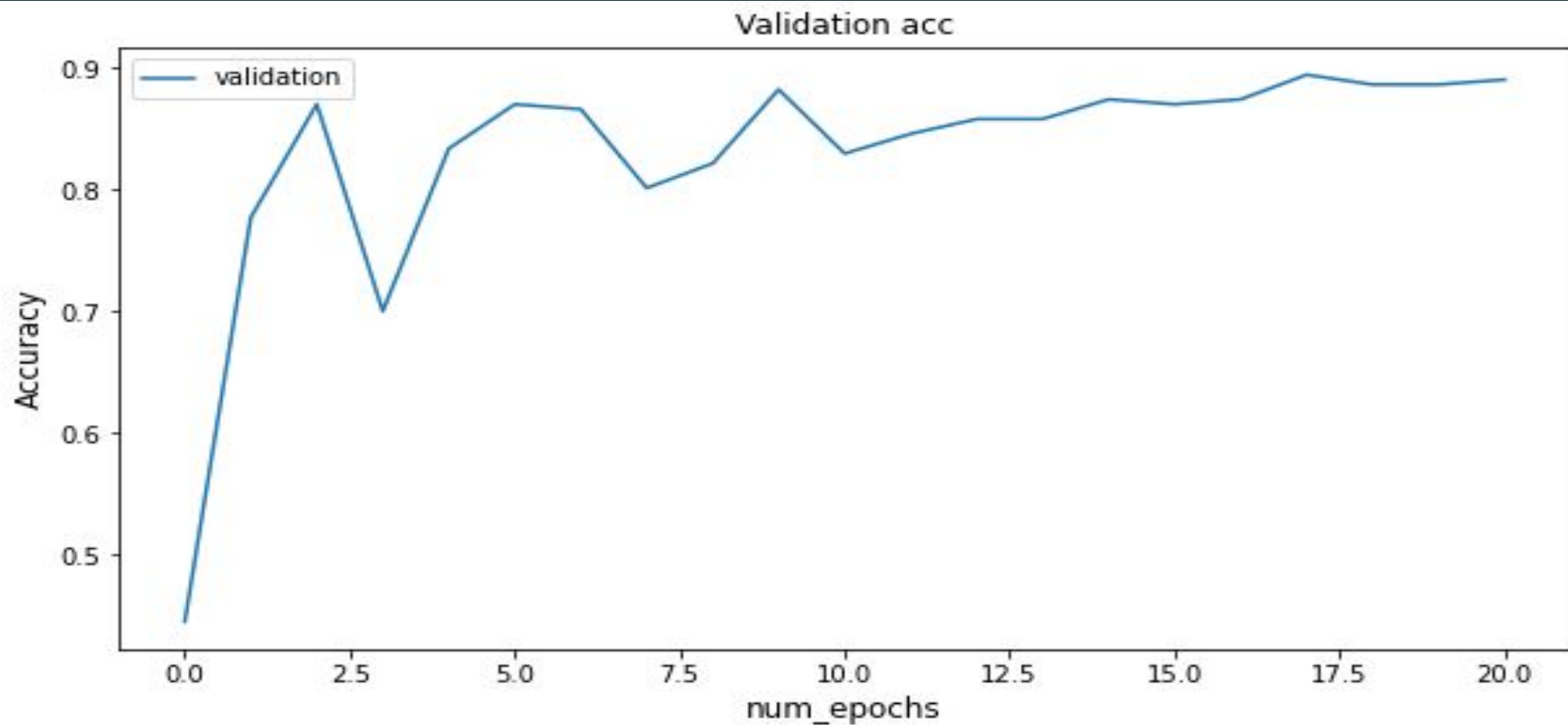
1. Gives less weightage to previous all derivatives and more weightage to recent moving average window.



Train Validation Loss



Validation Accuracy



Test Accuracy

Accuracy : 0.87999999952316284

Loss : 0.32042595744132996