

RMSprop

- Popular optimizer choice in neural networks.
- Uses an adaptive learning rate instead of treating η as a hyperparameter.
- Learning rate for a parameter is divided by an exponentially decaying average of the magnitude of gradients for that parameter.