ICT303 – Advanced Machine Learning and Artificial Intelligence

Topic 7: Implementation Details and Discussions

Hamid Laga
H.Laga@murdoch.edu.au
Office: 245.1.020

How to Get in Touch with the Teaching Team

- Internal and External Students
  - Email: H.Laga@murdoch.edu.au.
- Important
  - In any communication, please make sure that you
    - Start the subject of your email with ICT303
    - Include your student ID, name, and the lab slot in which you are enrolled.
  - We will do all our best to answer your queries within 24 hrs.

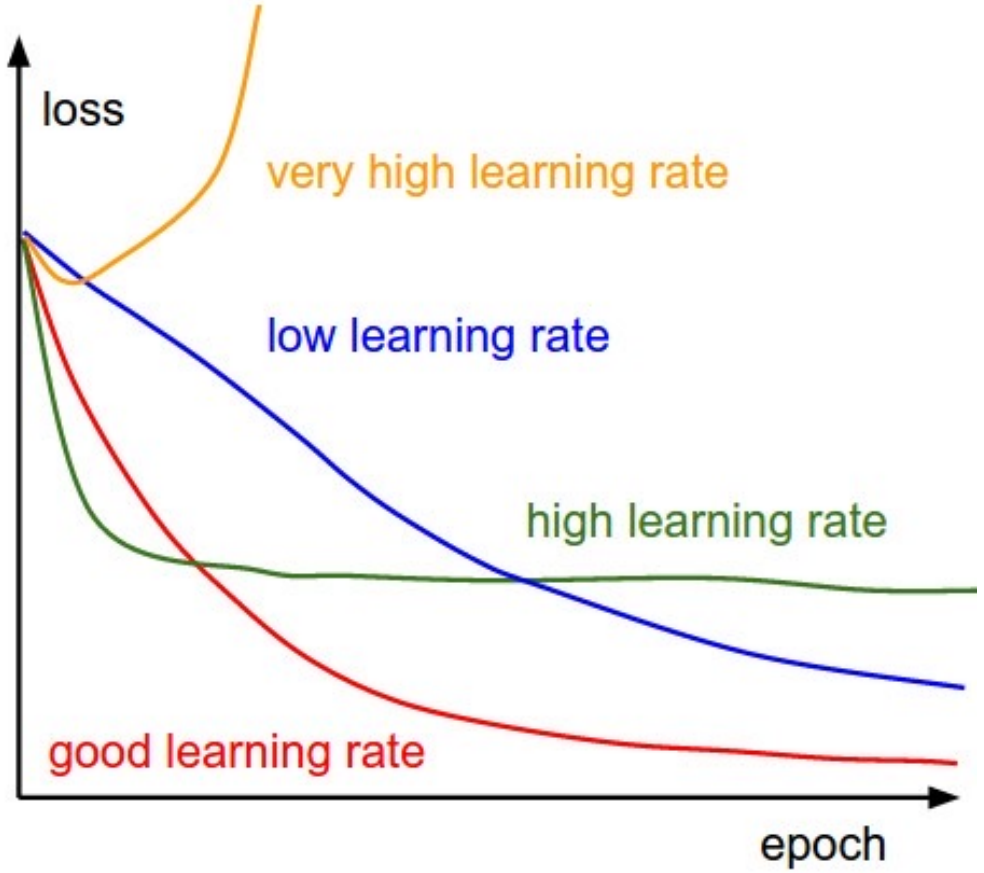
In this Lecture

- Hyper parameters
  - Learning rate
  - Batch size
- Data normalization
- ResNet18 implementation
- Summary

- Learning objectives
  - Understand the effect of some hyper parameters
  - Tune hyperparameters
  - Importance of data normalization
  - Implement ResNet18

Hyperparameters – Learning Rate

- Effect of the learning rate on training and loss function



https://cs231n.github.io/neural-networks-3/#loss-function

Hyperparameters – Learning Rate

- Use a schedule
  - Start with a large learning rate (e.g., 0.1)
  - Keep decreasing it after a few epochs until epoch == 30 or 50
  - Then keep it constant until convergence (at this stage the learning rate is low)
- Rate of decrease
  - Linear
  - Exponential, i.e., every time multiply it by a value that is lower than 1, e.g., 0.98
  - Use a decay factor, e.g., Decay = 0.01
    - New LR = Initial LR divided by (1 + Decay X Steps)

https://cs231n.github.io/neural-networks-3/#loss-function

Hyperparameters – Learning Rate

- Additional reading
  - Importance of the learning rate
    - https://cs231n.github.io/neural-networks-3/#loss-function
  - Using learning rate scheduler in PyTorch
    - https://machinelearningmastery.com/using-learning-rate-schedule-in-pytorch-training/
- What I didn't cover
  - Momentum update
    - https://cs231n.github.io/neural-networks-3/#loss-function

Hyperparameters – Batch Size

- Affects both the training time and generalization of the model
  - Smaller batch size
    - The network will learn from each individual but it takes long time to converge
  - Larger batch size
    - Fast training
    - Risk of not being able to generalize well – it will not capture nuances/variabilities in the data
  - Acceptable sizes
    - 16, 32, 64, 128, 256, 512, 1024
    - Make sure it fits within your GPU memory

Data Normalization

- Normalization is the process of bringing the data into a given range
- Normalizing the input
  - It is important especially of the input are of different types
  - This is usually done by
    - finding the mean and standard deviation
    - Subtract the mean
    - Divide by the standard deviation
  - This transforms the data so that its mean is 0 and its standard deviation is 1.

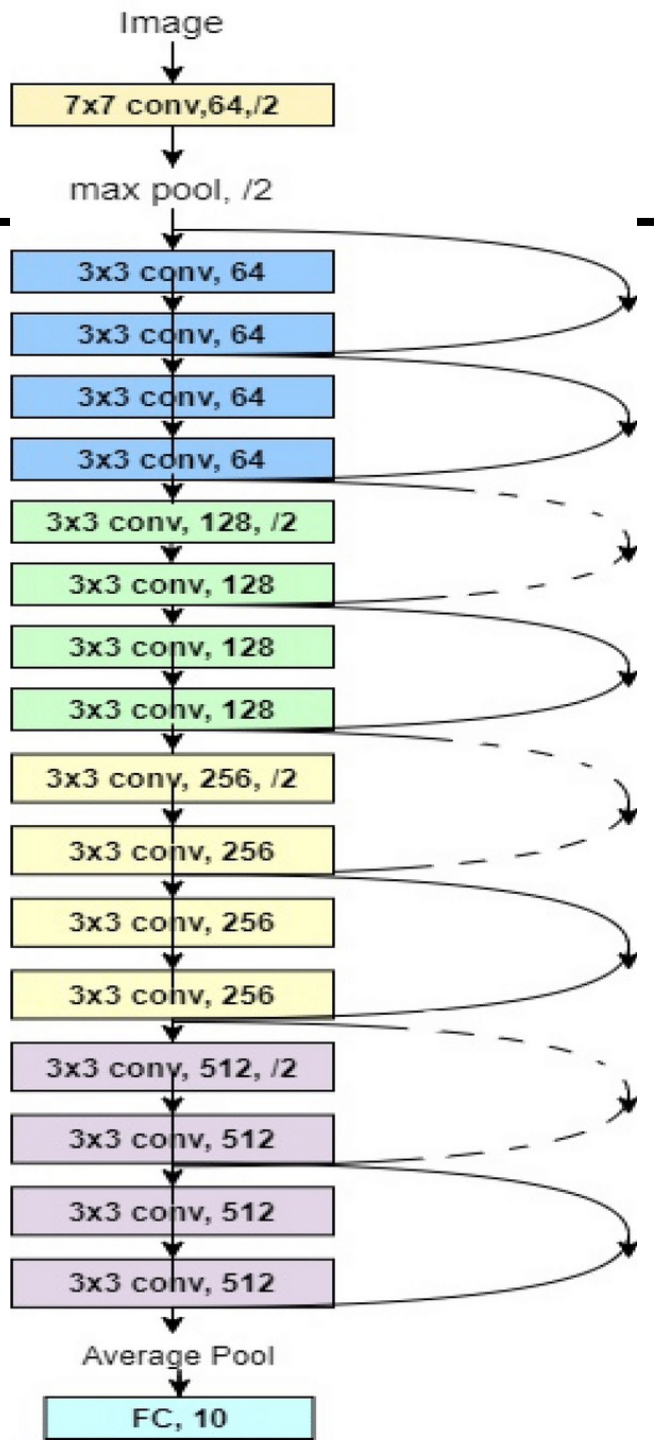
Data Normalization – Important Note

- Normalizing the input
  - It is important especially of the input are of different types
  - This is usually done by
    - finding the mean and standard deviation of the training data
  - Normalize the training data
    - Subtract the mean
    - Divide by the standard deviation
  - Normalize the validation data
    - Subtract the mean of the training data
    - Divide by the standard deviation of the training data
  - Normalize the test data
    - Subtract the mean of the training data
    - Divide by the standard deviation of the training data

Data Normalization – Batch Normalization

- Batch normalization standardizes the inputs to a layer for each mini-batch
  - Important for training very deep neural networks
  - Stabilizing the learning process and dramatically reducing the number of training epochs required to train deep networks
- At each hidden layer
  - Computes the mean and variance of the activation values across the batch
  - Normalizes the activation vector using the computed mean and variance
- PyTorch provides nn.BatchNorm2d or nn.BatchNorm1d
  - Can be applied before or after activation
  - Research shows it is best when applied before

ResNet18 Implementation



Next Week

- More about training (practical issues to take into account)

Questions