CSA 250: Deep Learning Project II Report

Modukuri Sumanth Kumar

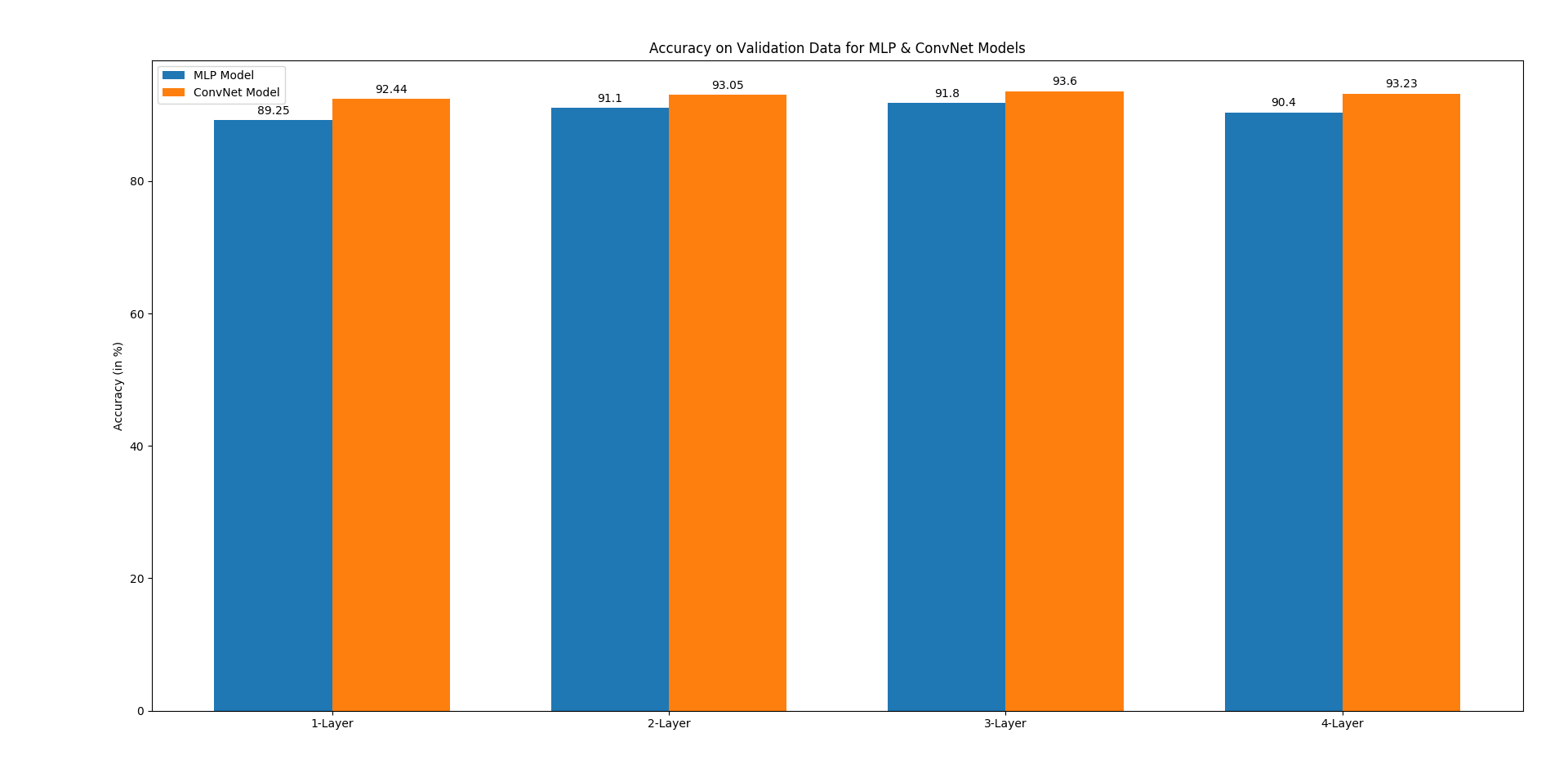
M.Tech AI [sumanthkumar@iisc.ac.in](mailto:sumanthkumar@iisc.ac.in) S/RNo.: 17025

Tools Used:

* Python 3
* Tensorflow 2.x
* Keras

Model Comparison:

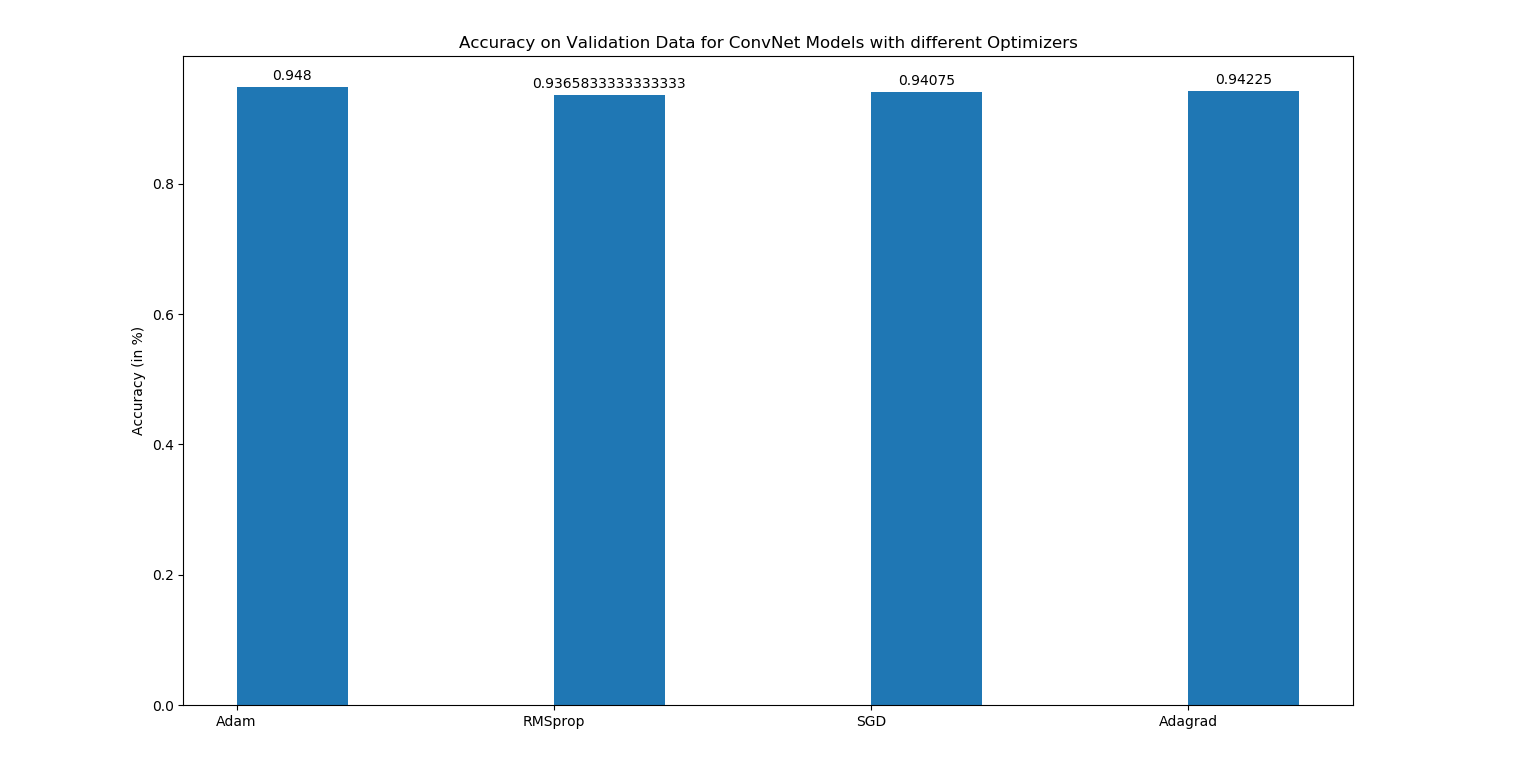
* What made you choose your current architecture? and How many different architectures have you considered?

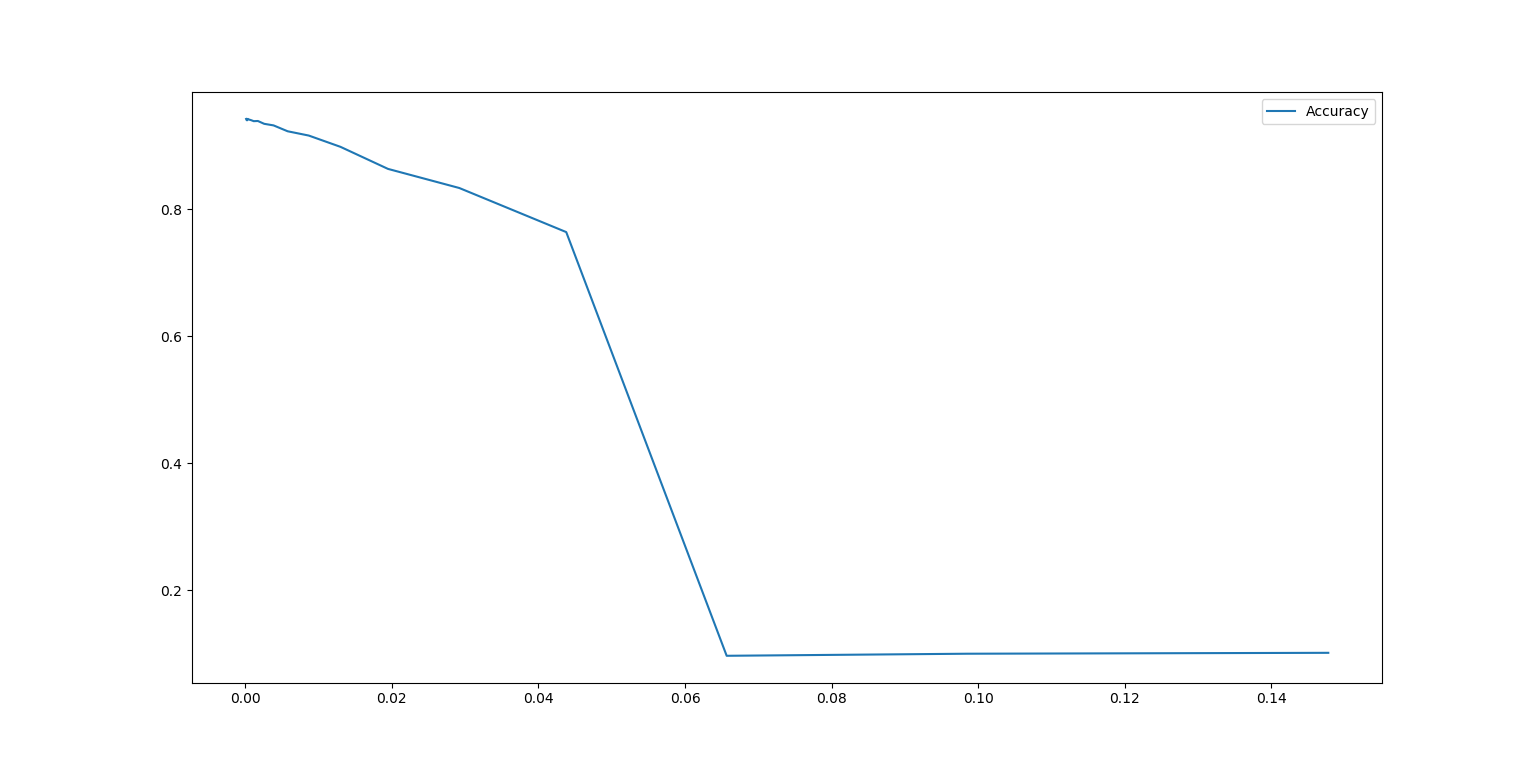


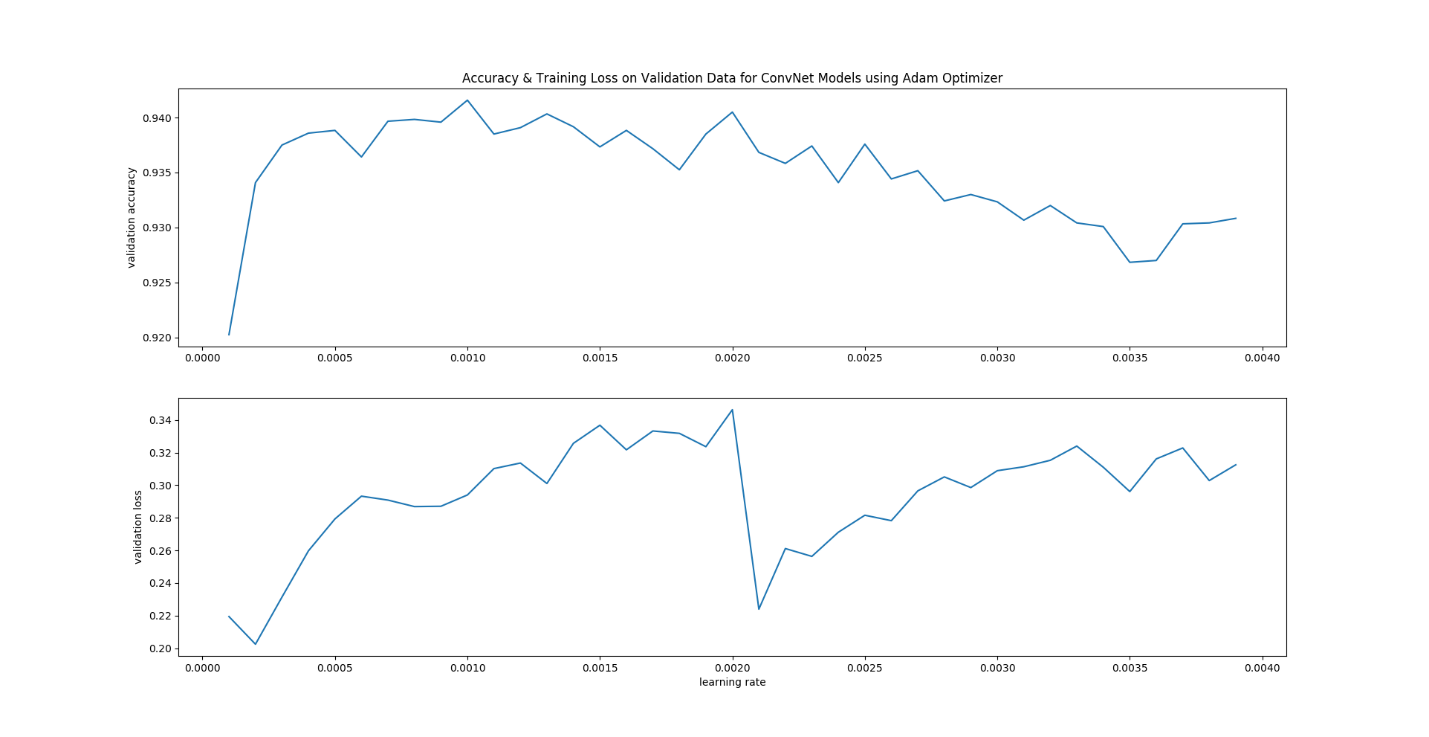
Result:

ConvNet: 3-Layer

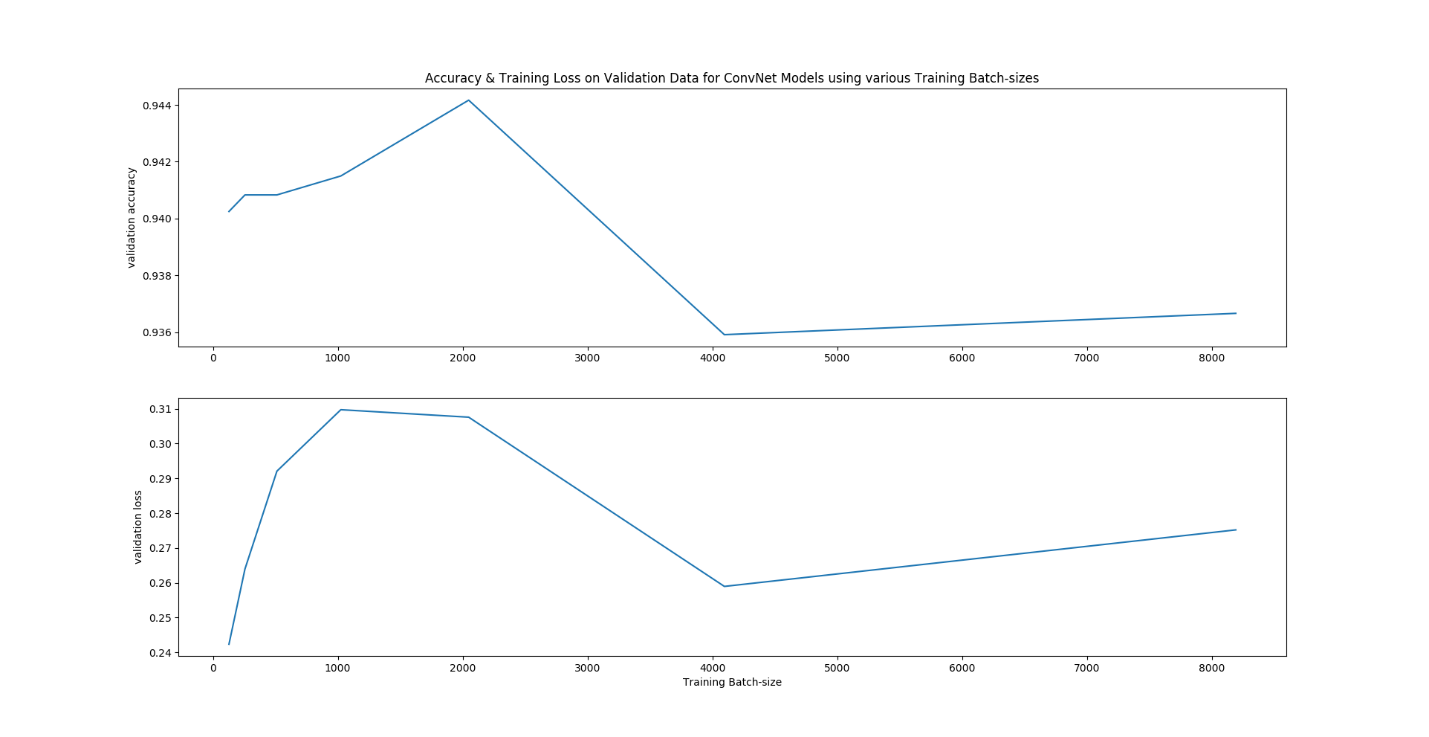
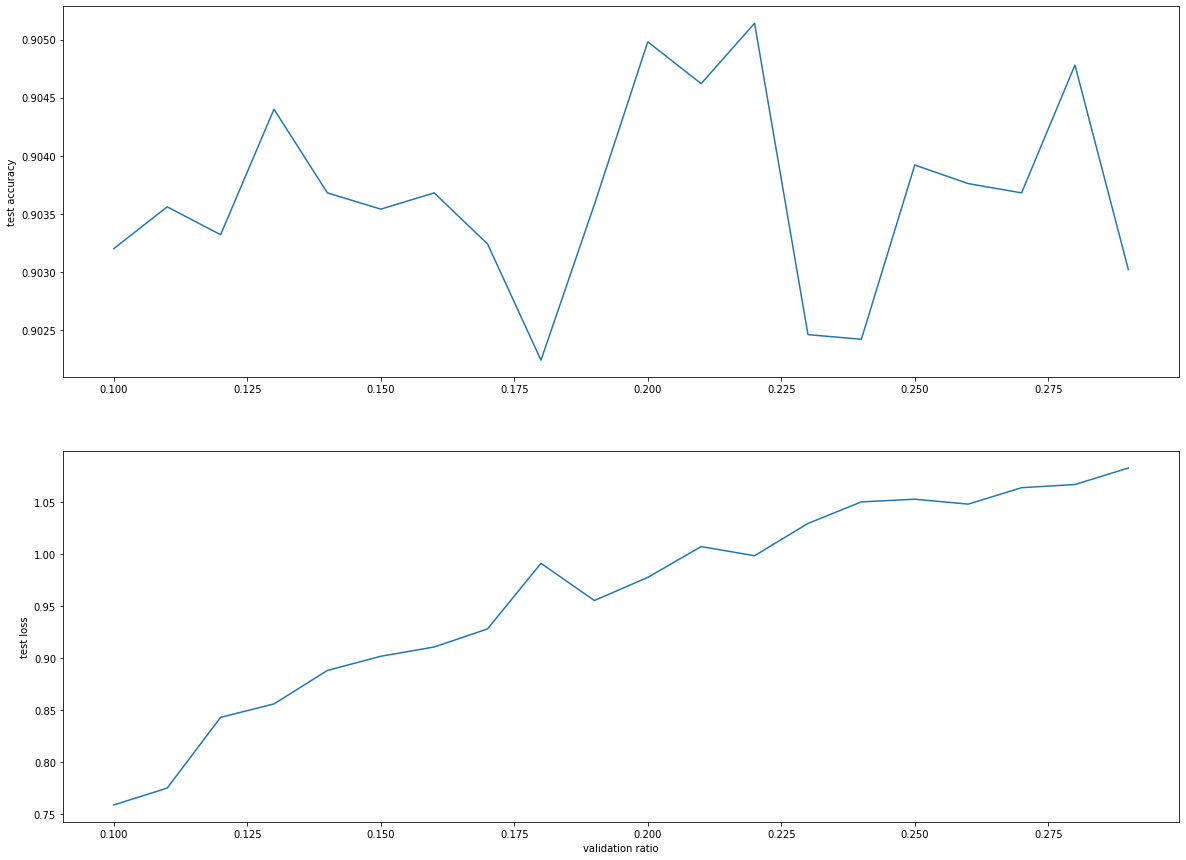
MLP model: 3-Layer

* How did you arrive at your choice of hyperparameters - number of neurons/layers, activation function, learning rate etc.?
  + Number of layers:  
    Result: **3-Layer Models** (above bar graph)
  + Activation function:  
      
    Result: **Adam**
  + Learning rate:  
      
      
      
      
      
      
      
    Learning rate vs Accuracy

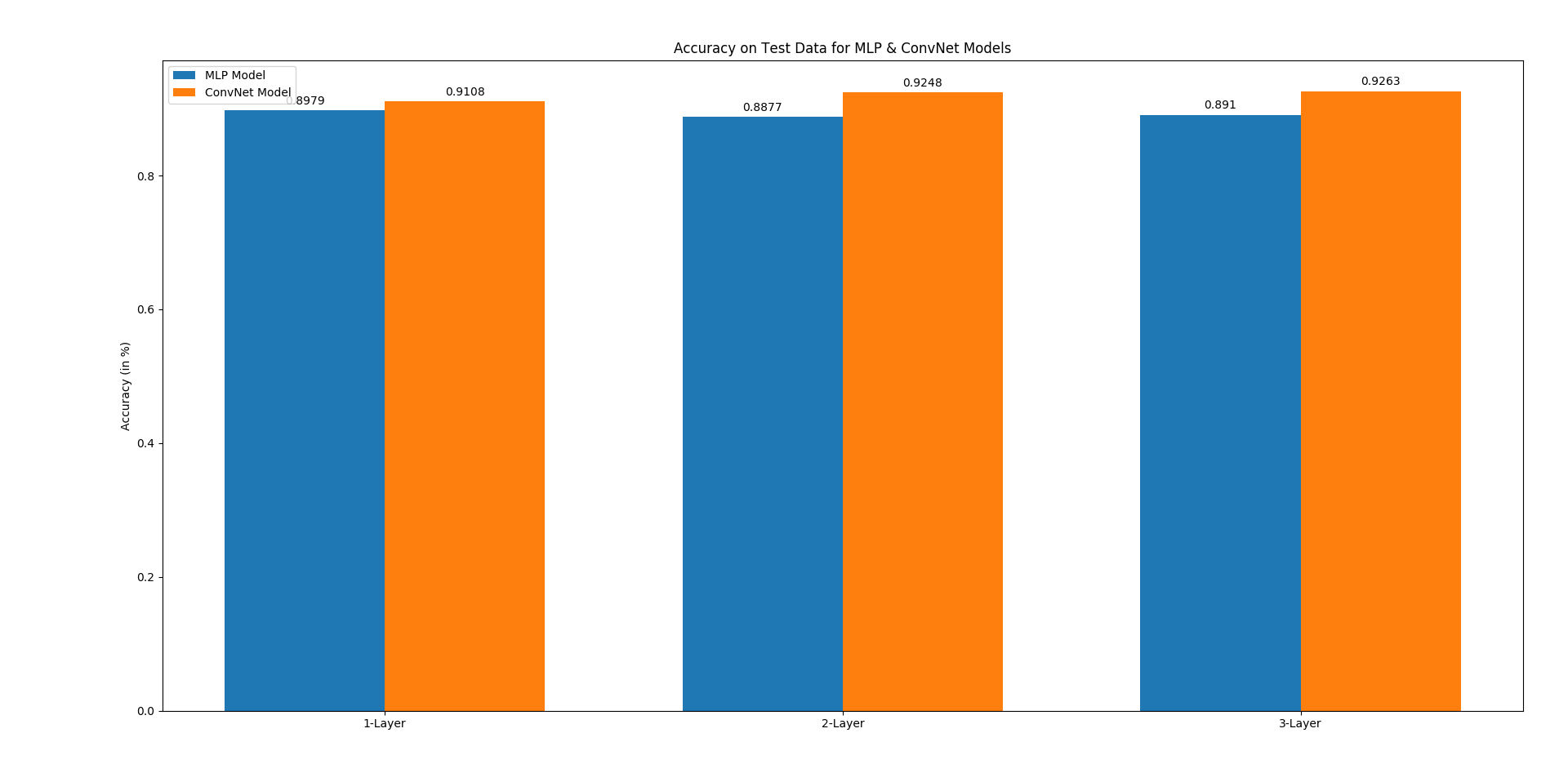




Result: **0.001** gave best result

* + Batch-size:  
      
      
    Result: **2048** gave best result
* How did you validate your model?  
    
    
    
  Result: **0.2** gave best results on test data, I used 48000 samples for training & 12000 for validating where the division is done randomly

Models Results:



Observation:

ConvNet with 3 layers performed the best, its performance was increasing from 1 to 3, then was invariant of increasing the layers.

MLP models were giving similar performance for any architectures with reasonable nodes. It could not get better results than ConvNets, & MLP’s model file size is much higher than ConvNets. ConvNets were time taking to train but delivered the best results with 3-Layer architecture.

References:

https://github.com/markjay4k/Fashion-MNIST-with-Keras