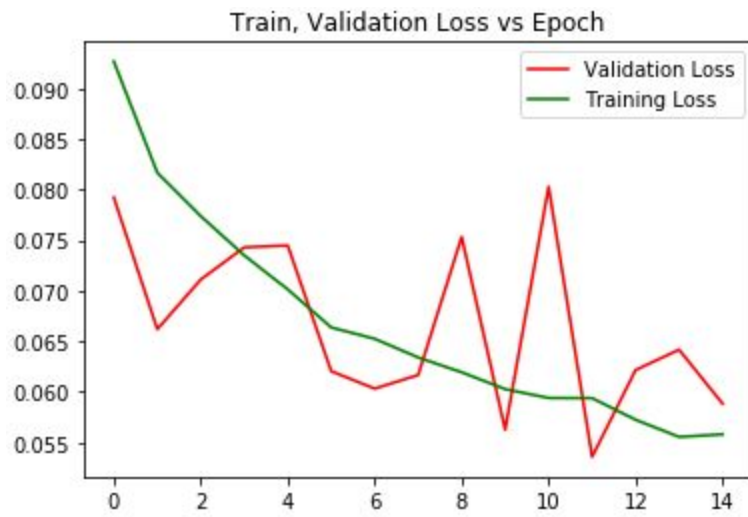


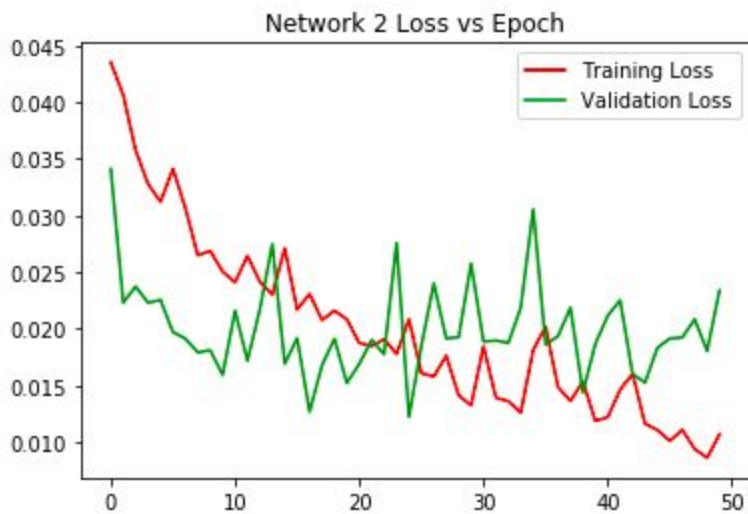
Network 1

Random initialization,Adam Optimization



Network 2

Random initialization,Adam Optimization



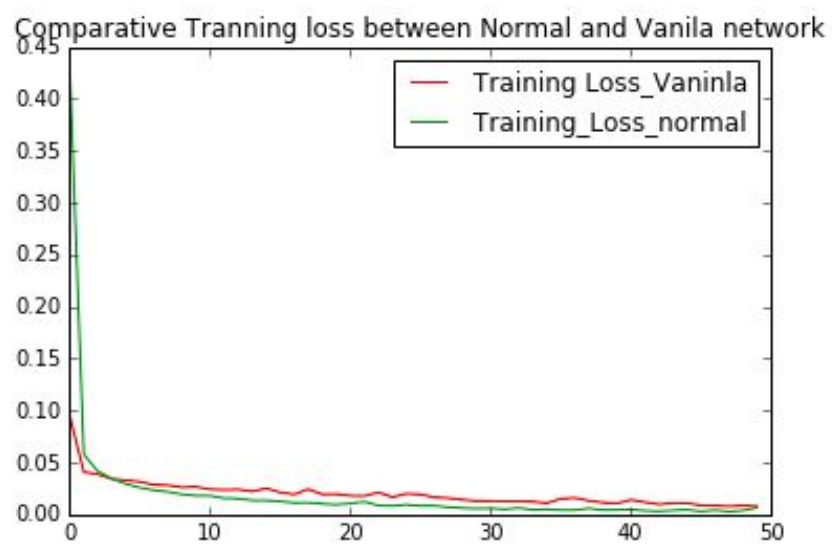
Training accuracy =0.764095238095

Total Accuracy: 0.696324936224

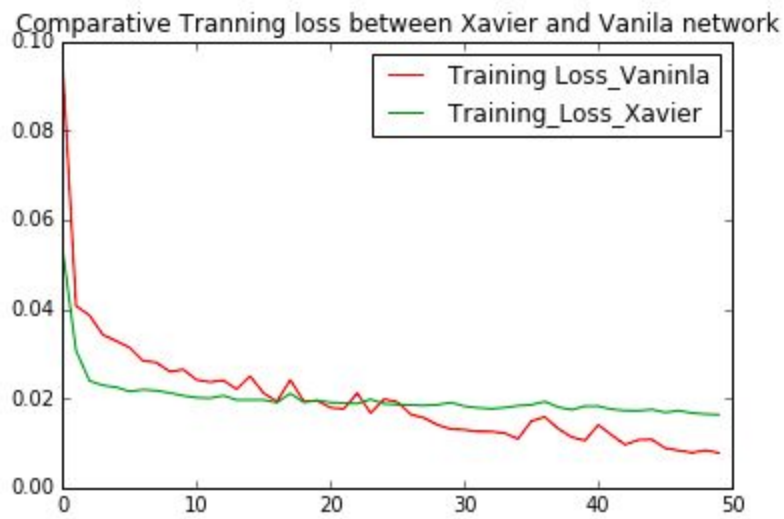
Task B

Experiment 1

Training Loss Comparison Between Normal initialization and Vanilla

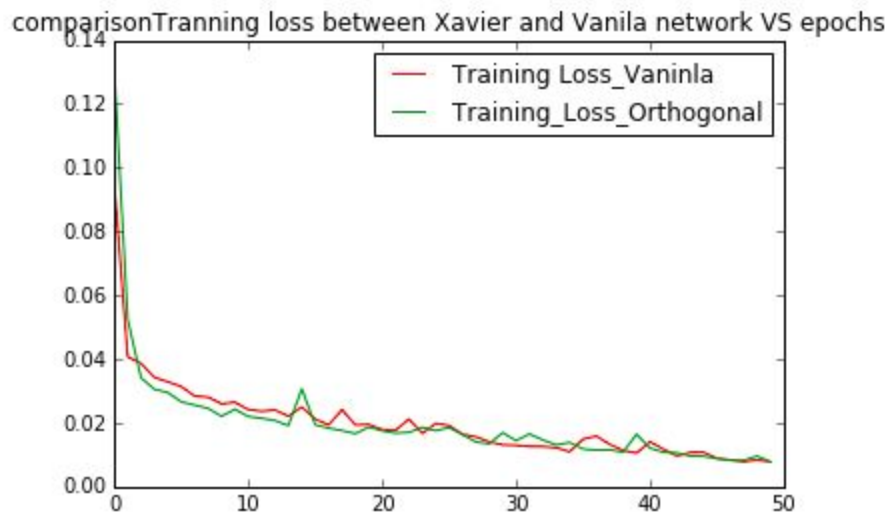


Training Loss Comparison Between Xavier initialization and Vanilla



EXP 4

Training Loss Comparison Between Orthogonal initialization and Vanilla



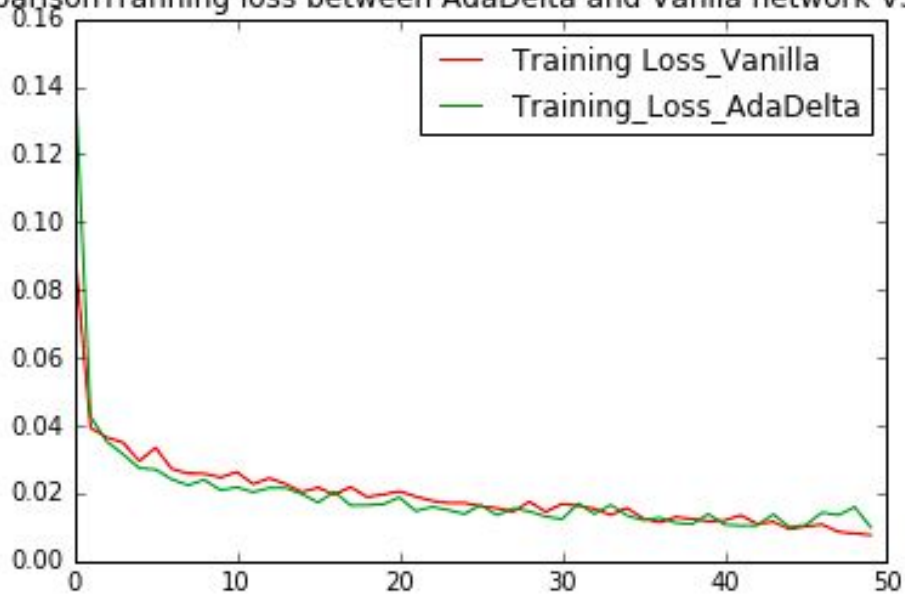
Experiment 4
SGD optimizer

comparisonTranning loss between SGD and Vanila network VS epochs

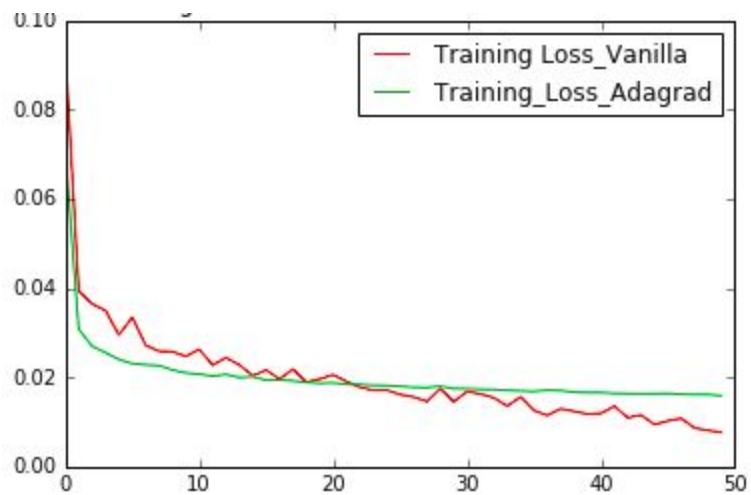


AdaDelta optimizer

comparisonTranning loss between AdaDelta and Vanila network VS epochs



Adagrad optimizer



Adagrad Total Accuracy: 0.864409279337

RmsProp optimizer

comparison Training loss between RmsProp and Vanila network VS epochs



RmsProp Total Accuracy: 0.89430205676

NAG optimizer

comparison Training loss between NAG and Vanila network VS epochs

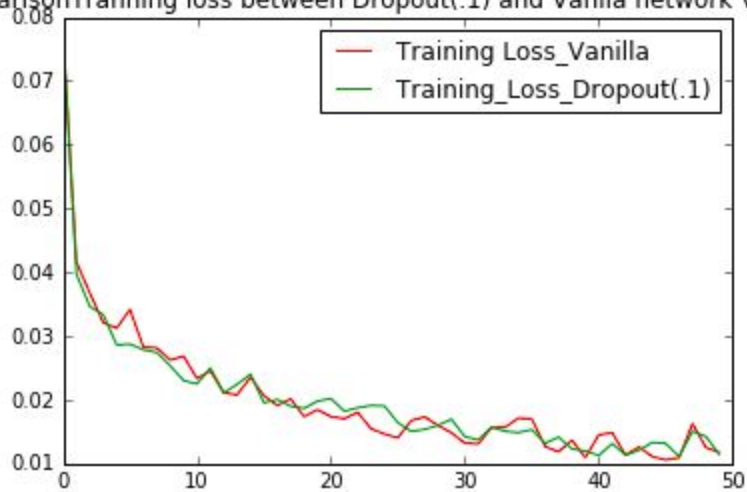


NAG Total Accuracy: 0.760265864158

Experiment 2

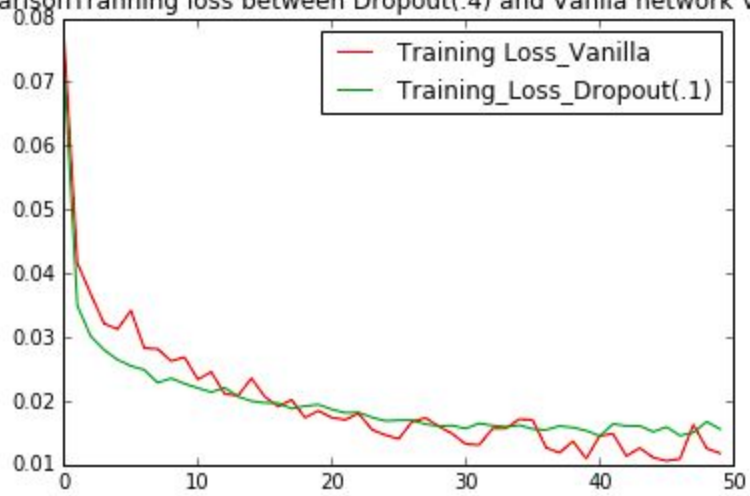
dropout(.1)

comparison Training loss between Dropout(.1) and Vanila network VS epochs

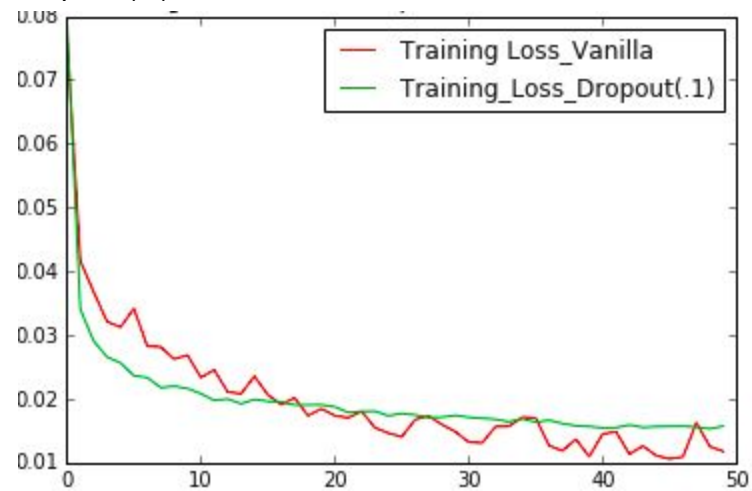


Dropout (.4)

comparison Training loss between Dropout(.4) and Vanilla network VS epochs

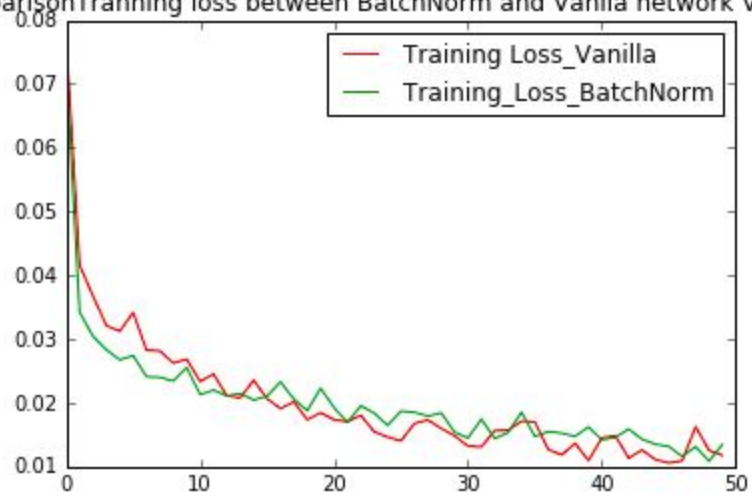


Dropout (.6)



BatchNorm

comparisonTranning loss between BatchNorm and Vanila network VS epochs



BatchNorm Total Accuracy: 0.868241788903