**Experiment 3**

Kaashan Panjwani

J035 (J2)

B.Tech DS 3rd Year DL

**1. rmsprop**

acc: 0.9966

val\_loss: 0.1012

val\_acc: 0.9804

**2. Stochastic gradient descent**

acc: 0.9999

val\_loss: 0.0786

val\_acc: 0.9854

**3. RMSProp**

acc: 0.9986

val\_loss: 0.1459

val\_acc: 0.9807

**4. Adagrad**

acc: 1.0000

val\_loss: 0.1091

val\_acc: 0.9839

**5. Adadelta**

acc: 1.0000

val\_loss: 0.1116

val\_acc: 0.9842

**6. Adam**

acc: 0.9979

val\_loss: 0.0977

val\_acc: 0.9834

**7. Adamax**

acc: 1.0000

val\_loss: 0.1092

val\_acc: 0.98618

**8. Nesterov Adam**

acc: 0.9945

val\_loss: 0.1665

val\_acc: 0.9726

Conclusion :

After doing this experiment we observe that the best optimiser for this dataset is **Nesterov Adam** and the least optimal one being **rmsprop**.

Using the best optimiser i.e **Nesterov Adam**, we try and improve accuracies of other datasets :

**1. fashion\_mnist**

acc: 0.9214

val\_loss: 0.3430

val\_acc: 0.8857

**2. Cifar10**

acc: 0.1000

val\_loss: 14.5063

val\_acc: 0.1000

**3. Cifar100**

acc: 0.2686

val\_loss: 3.3171

val\_acc: 0.2208

**4. Iris**

loss: 0.0536

acc: 0.9797