Observations for prac 3 Optimizer

|  |  |  |
| --- | --- | --- |
| Model | Test Loss | Accuracy |
| Basic Model | 0.1023 | 0.9813 |
| SGD Optimizer | 0.0794 | 0.9855 |
| RMSProp Optimizer | 0.1331 | 0.9828 |
| Adagrad | 0.1156 | 0.986 |
| Adadelta | 0.1158 | 0.9859 |
| Adam | 0.1108 | 0.9819 |
| Adamax | 0.1049 | 0.9862 |
| Nadam | 0.1397 | 0.9785 |

Conclusion:

Least test loss: Stochastic Gradient Descent (0.0794049291110724).

Best test Accuracy: Adamax (0.9862)

|  |  |  |
| --- | --- | --- |
| Dataset | Loss | Accuracy |
| MNIST | 0.1023 | 0.9813 |
| FASHION\_MNIST | 0.3318 | 0.8768 |
| CIFAR\_10 | 1.3632 | 0.5149 |
| CIFAR\_100 | 3.5743 | 0.172 |
| IRIS | 0.1101 | 1.000 |