Below table (table\_1) shows batch\_risk, test\_risk, accuracy and test accuracy for different subset, batch\_size and learning\_rate.

For table 1: subset training size affects accuracy and risk. Computation time is fast, risk is high. Batch size and learning rate affects accuracy and risk.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Batch\_size | Learning\_rate | Subset\_size | Epoch | Batch\_risk | Risk\_item | Accuracy | Test\_accuracy |
| 64 | 0.01 | 2000 | 9999 | 2.1708999887157675e-10 | 2.5456395185265817 | 1.0 | 0.8779 |
| 64 | 0.0001 | 100 | 9999 | 0.0004941953145229556 | 7.794979266348429 | 1.0 | 0.1488 |
| 64 | 0.1 | 2000 | 9999 | 3.28188908750791e-09 | 16.102805327839153 | 1.0 | 0.1122 |
| 64 | 0.01 | 500 | 9999 | 3.2480659105142848e-09 | 26.220342353488878 | 1.0 | 0.12 |
| 256 | 0. 001 | 500 | 9999 | 6.21043010252034e-05 | 8.861352639176138 | 1.0 | 0.1157 |
| 2000 | 0.001 | 500 | 9999 | 8.953223977384239e-05 | 0.9341344672731162 | 1.0 | 0.8365 |
| 64 | 0.001 | 500 | 9999 | 2.288315093265438e-09 | 1.7512380598875879 | 1.0 | 0.8509 |
| 64 | 0.1 | 500 | 9999 |  |  |  |  |
| 2000 | 0.01 | 500 | 9999 | 3.1984725447719887e-06 | 1.1927371542978025 | 1.0 | 0.8395 |
| 256 | 0.01 | 500 | 9999 | 3.2216497328659006e-08 | 1.5524908426773814 | 1.0 | 0.8421 |
| 256 | 0.01 | 2000 | 9999 | 2.02093579534852e-06 | 10.840338921992531 | 1.0 | 0.0963 |

Below table (table\_2) shows batch\_risk,test\_risk,accuracy and test\_accuracy for subset size of 50000 and different batch\_size and learning\_rate.

For table 2: almost for 500 epochs accuracy and test\_accuracy is 90 percent and more but after that it starts decreasing gradually.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Batch\_size | Learning\_rate | Epoch | Batch\_risk | Risk\_item | Accuracy | Test\_accuracy |
| 64 | 0.01 | 9999 | nan | nan | 0.09904 | 0.098 |
| 20 | 0.001 | 9999 | nan | nan | 0.09864 | 0.098 |
| 256 | 0.1 | 9999 | nan | nan | 0.09904 | 0.098 |

The size of the training data set is a major determinant of classification accuracy. with larger training sets typically resulting in superior performance compared to smaller training sets. Larger the batch size: The slower the training loss decreases, The higher the minimum validation loss, The less time it takes to train per epoch, The more epochs it takes to converge to the minimum validation loss.