**RESULTS**

* 2-layer CNN architecture with stride=1, padding=0
* Adam Optimizer with weight decay = 0.0001
* Trained for 15 epochs
* No. of configurations tested = 200
* Runtime= 8 hours

Default configuration:

Mini Batch size: 128

Learning Rate: 0.01

* Optimized loss Value: 0. 05043112114071846 (CrossEntropyLoss)

Incumbent configuration:

Mini Batch size: 64

Learning Rate: 0.0064434254058602985

* Accuracy: 92%
* Optimized loss Value: 0.009865626692771912(CrossEntropyLoss)

