Progress Report Savaiz

Training

Completed Deep Learning Videos

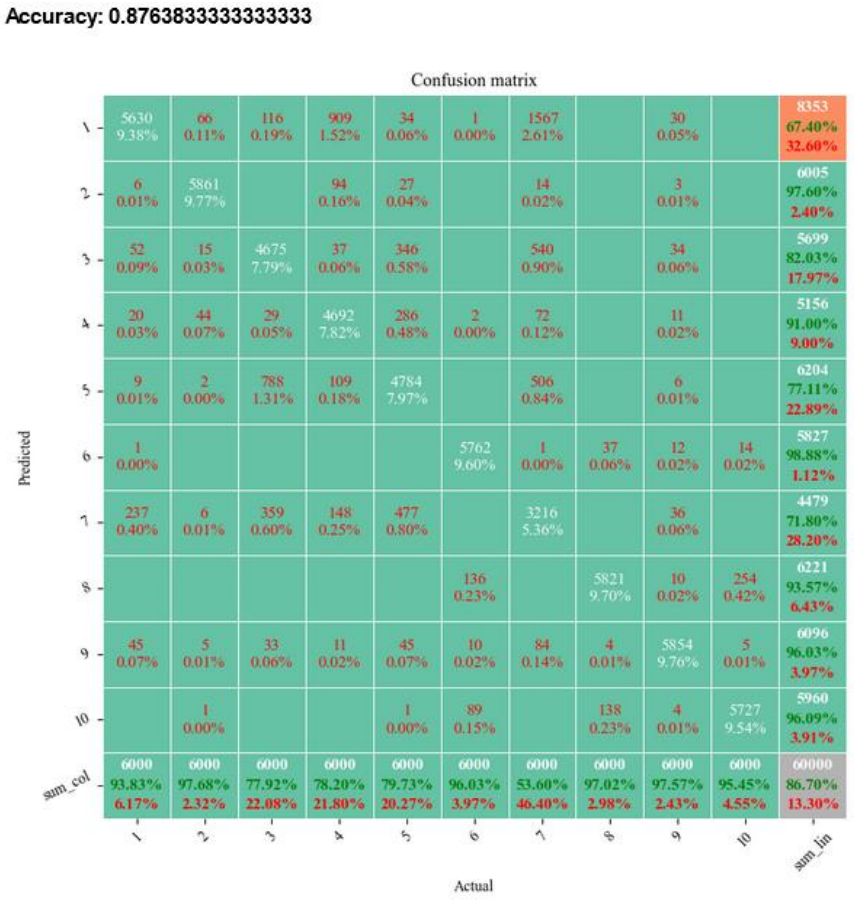
Completed PyTorch Tutorials

Attempted all tasks

BadNets (Week 1)

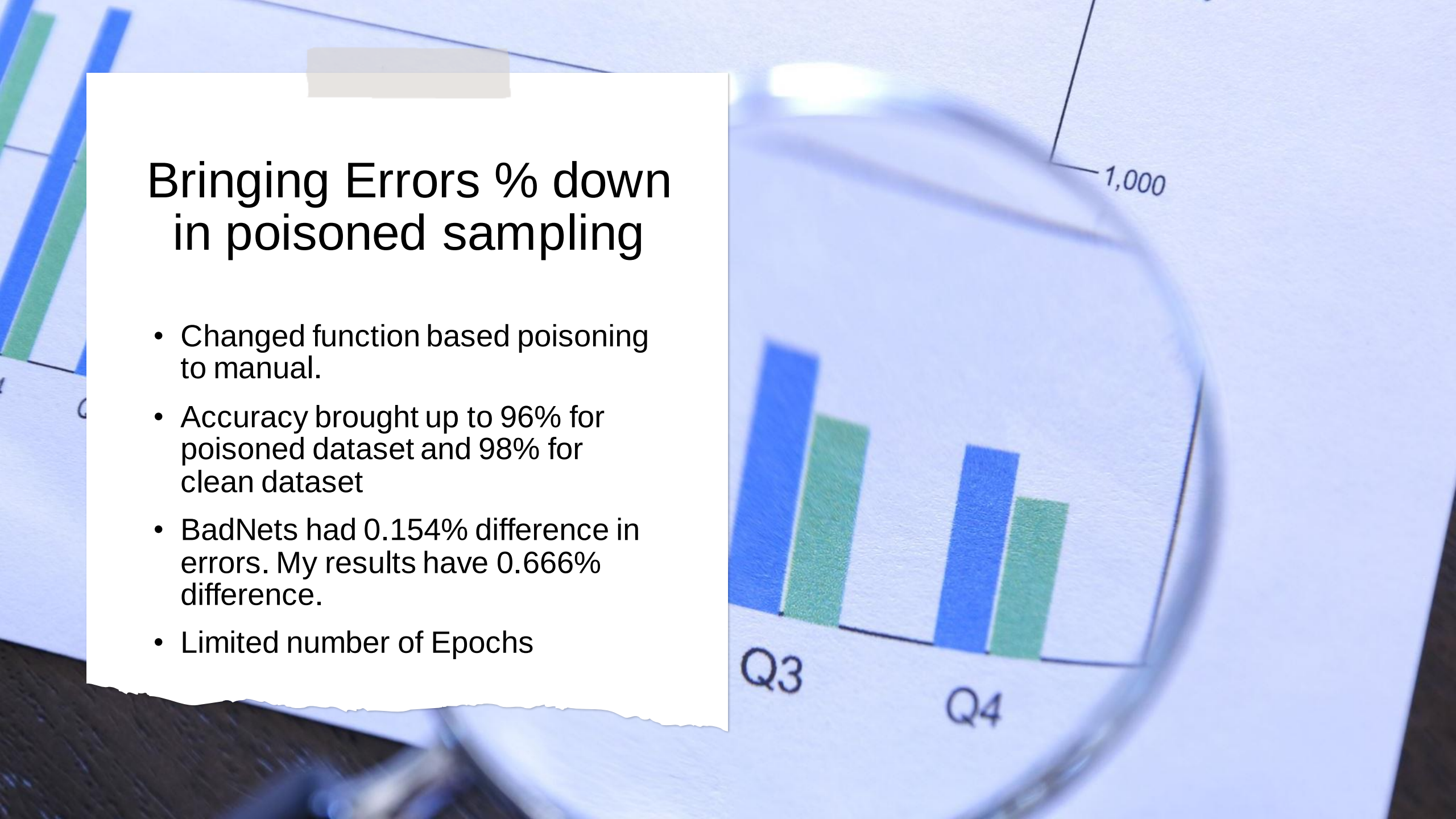
Clean

Poisoned



Bringing Errors % down in poisoned sampling

- Changed function based poisoning to manual.
- Accuracy brought up to 96% for poisoned dataset and 98% for clean dataset
- BadNets had 0.154% difference in errors. My results have 0.666% difference.
- Limited number of Epochs



Other Observations

- 50% backdoored samples: 97% accuracy
- single target (1,9): 98% accuracy
- 2 targets: 96% later 98%

Current hurdles

LIRA

- Understanding the algorithm
- Slow performance

Narcissus

- CIFAR
- Remote access