Full Stack Deep Learning Spring 2021 Labs

Welcome!

As part of Full Stack Deep Learning 2021, we will incrementally develop a complete deep learning codebase to understand the content of handwritten paragraphs.

We will use the modern stack of PyTorch and PyTorch-Ligtning

We will use the main workhorses of DL today: CNNs, RNNs, and Transformers

We will manage our experiments using what we believe to be the best tool for the job: Weights & Biases

We will set up continuous integration system for our codebase using CircleCI

We will package up the prediction system as a REST API using FastAPI, and deploy it as a Docker container on AWS Lambda.

We will set up monitoring that alerts us when the incoming data distribution changes.

Sequence:

- Lab Setup: Set up our computing environment.
- Lab 1: Intro: Formulate problem, structure codebase, train an MLP for MNIST.
- Lab 2: CNNs: Introduce EMNIST, generate synthetic handwritten lines, and train CNNs.
- Lab 3: RNNs: Using CNN + LSTM with CTC loss for line text recognition.
- Lab 4: Transformers: Using Transformers for line text recognition.
- Lab 5: Experiment Management: Real handwriting data, Weights & Biases, and hyperparameter sweeps.
- Lab 6: Data Labeling: Label our own handwriting data and properly store it.
- Lab 7: Paragraph Recognition: Train and evaluate whole-paragraph recognition.
- Lab 8: Continuous Integration: Add continuous linting and testing of our code.
- Lab 9: Deployment: Run as a REST API locally, then in Docker, then put in production using AWS Lambda.
- Lab 10: Monitoring: Set up monitoring that alerts us when the incoming data distribution changes.