Machine Learning Engineering & AI Bootcamp Capstone

Step 8: Scale Your Prototype

Summary

Time Estimate: 15 - 20 hours

In this step, your goal is to ensure that your ML/DL approach, which you've proved to be viable, can work with large volumes of data. You need to scale your prototype. Please work with your mentor to determine what that means for your problem.

Using scikit-learn, SparkML, Keras, TensorFlow, PyTorch, or some of the other technologies you have learned, implement your prototype at scale.

In case your earlier prototype was working with a subset, ensure that this scaled-up prototype can handle your complete dataset.

Think about what your capstone problem would look like in the real world:

- How much data would you need to handle?
- Can you scale your prototype to handle that volume of data using the approach and tools you have selected?

Project Submission Step

1. Implement the scaled version of your prototype and clearly document the trade-offs and implementation decisions you have to make to be able to scale your algorithm.

2. Submit the GitHub link to your notebook.