Batching and Shuffling Data

What is A Batch?

- A batch is a portion of data that gets feed into the network
- For Example, let's say you are training a MNIST network (28 * 28)
  - Instead of training one image at a time, it's possible to train multiple flattened image at once
  - This would make training faster as it would allocate more resource to train
  - If we train 100 images at once then the batch size would be 100

Using TF Data

- If you need more control over the input data pipeline or needed to use data that doesn't easily fit into memory use: `tf.data`
- Single File:

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/7f85d846-b118-4a5a-bc8f-80a14d24a931/single_file.ipynb

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/a702620a-e056-4635-9c13-b7cb4b12fabc/single_file.py

- Multiple Files:

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/10215935-fa68-42d0-a62d-121ca8d4f1eb/multiple_files.ipynb

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/25ced4f0-b0c7-49a2-b530-d9d0a1008c75/multiple_files.py