Image Data - Recognising Handwritten Alphabets

Dataset Link - mnist.zip

SPRINT 1 - Create DataFrame from raw Image Files

Description

MNIST ("Modified National Institute of Standards and Technology") is the de facto "hello world" dataset of computer vision. Since its release in 1999, this classic dataset of handwritten images has served as the basis for benchmarking classification algorithms. As new machine learning techniques emerge, MNIST remains a reliable resource for researchers and learners alike.

In this SPRINT, your goal is to preprocess the image files given to you and create a pandas dataframe.

Task A - Download the *'mnist_data.zip'* and read the data in a pandas dataframe.
Task B - Use your Data Engineering skills to create a dataframe which can annotate each image into one of the 26 classes.

SPRINT 2 - Build a model

Task A - Perform data preprocessing on the given image data and convert it into numerical vectors.

Task B - Given an image, build a model that can recognise the character.

Client Expectations

1. Show me some nice analysis on the given data.
2. Show me the comparison of various ML models.
3. Model should be light for deployment.
4. Model should have very less latency.
5. Create a REST API to interact with the model.