Project Title:
Anomaly Detection in Digits using Variational Autoencoders (VAE)

Project Participants:

Anson Antony (Responsible for data preparation and preprocessing, In charge of VAE model implementation, Handles testing)
Aditya Singh (In charge of VAE model implementation, training, anomaly detection, results interpretation)
Note: Our team is composed of two members, each having a distinct role, ensuring a streamlined workflow and efficient collaboration.

Problem Description:
Our goal is to train a model on the digits 0 through 8 and leverage it to detect the digit 9 as an anomaly.

Inputs: Images of handwritten digits (0-8).
Outputs: Anomaly score for each input image when testing with the digit 9.
Why is it interesting? Demonstrating that a model can be trained on a subset of data and can detect unseen data (in this case, the digit 9) as anomalies has significant implications for fraud detection, quality control, and other anomaly detection applications.

Algorithms:

Expected to use: Variational Autoencoders (VAE).
Why appropriate: VAEs can reconstruct input data and, in the process, determine how well the data conforms to the training data distribution. Anomalies will have higher reconstruction errors.
Typical usage: VAEs are generally used for generative tasks and data reconstruction.
Prior usage: VAEs have been used previously in literature for anomaly detection tasks.

Data Sets:

Source: Presumably, the dataset would be the MNIST dataset or a subset of it.
Existing/New: It's an existing dataset.
Preprocessing Steps: The dataset would require normalization, and only digits 0-8 would be used for training.

Libraries and Tools:

Libraries: The typical libraries for this task would include TensorFlow or PyTorch for building and training the VAE, and possibly scikit-learn for auxiliary tasks.

References:
TensorFlow Documentation
PyTorch Documentation

Results:

Ideal Outcome: The VAE, when trained on digits 0-8, should be able to detect digit 9 images as anomalies with high confidence.
Expected Results: We anticipate higher reconstruction errors for digit 9 as compared to digits 0-8.
Comparisons: We might compare the average reconstruction error of digits 0-8 with that of digit 9.
Risks: If the VAE doesn't differentiate the digit 9 effectively, we might need to revisit our model architecture or training strategy.

Note: Might scale it up to new related task eg image anomaly detection depending on complexity and scope of the current problem statement