Final Project Report

Running the code:

1. Unzip the folder after downloading, then upload all .ipynb files in google collab, or open them with jupyter notebook.

Introduction:

In this project, we utilize the CIFAR dataset, a well-known benchmark in image classification, to evaluate and compare the performance of four distinct machine learning models: a Multilayer Perceptron (MLP), a Convolutional Neural Network (CNN), an Encoder model, and a ResNet (Residual Network). Each of these models represents a unique approach to handling image data, ranging from basic neural network architectures like MLP to more advanced deep learning techniques in CNNs and ResNet. Our goal is to explore how effectively each model classifies images from the CIFAR dataset, thereby gaining insights into their strengths, limitations, and potential applications in the field of image classification.

Conclusion of each model:

1. **MLP:**

   In evaluating the Multilayer Perceptron (MLP) model with the CIFAR dataset, I achieved an accuracy of 51%, indicating a moderate level of effectiveness in image classification. This somewhat limited performance can be attributed to the MLP's nature as a fully connected network, which does not specialize in capturing the spatial and local pixel relationships crucial in image data. Consequently, while MLPs offer a fundamental understanding of neural networks, their structure inherently lacks the efficiency needed for complex image datasets. In contrast, a Convolutional Neural Network (CNN), with its convolutional layers designed to process pixel data and capture spatial hierarchies, would likely be a more ideal choice for enhancing accuracy and performance in such tasks.

2. **CNN:**

   Upon implementing a Convolutional Neural Network (CNN) for the CIFAR dataset, the model demonstrated an improved accuracy of 60%, a notable enhancement over the MLP model. This increase can be attributed to CNN's architecture, which is specifically tailored for image processing. CNNs utilize convolutional layers to efficiently identify and learn spatial hierarchies and patterns in image data, such as edges, textures, and shapes. This ability to focus on local receptive fields and apply shared weights makes CNNs adept at handling the intricacies of image datasets like CIFAR. However, the 60% accuracy also suggests room for further optimization in the network's architecture, such as adjusting filter sizes or layer configurations, to better capture the complexities of the dataset and further improve the classification performance.

3. **Autoencoder:**

In analyzing the performance of an autoencoder on the CIFAR dataset, which resulted in an overall mean squared error (MSE) of 0.1439 and class-specific MSEs ranging from 0.1300 to 0.1532, a key challenge emerged in determining the model's comparative effectiveness. Unlike traditional classification accuracy metrics, MSE focuses on the model's ability to reconstruct images, making direct comparisons with other models like MLP or CNN, which are evaluated based on classification accuracy, less straightforward. This discrepancy in evaluation metrics stems from the inherent difference in the primary objectives of the models: while MLP and CNN aim to categorize images into distinct classes, the autoencoder's goal is to capture and reproduce the underlying features of the images. Therefore, while the MSE provides valuable insights into the autoencoder's reconstruction capabilities, it doesn't directly translate to a measure of classification accuracy, complicating the assessment of whether it is a "better" model for the CIFAR dataset in comparison to MLP or CNN.