# Assignment 1

**12/29/2024**

10 Points Possible

Attempt 1

In Progress

**NEXT UP: Submit Assignment**

Add Comment

**Unlimited Attempts Allowed**

12/1/2024 to 12/29/2024

**Details**

This assignment is about feature extraction via dimensionality reduction using variants of autoencoders. Use the CIFAR10 dataset provided in Keras, after conversion to gray-level images! Use randomly selected 70% of the dataset as training set and remaining 30% as the test set.

Task 1: Perform standard PCA with 70% of the training dataset and identify the eigenvectors associated with top eigenvalues with 95% total energy. With these, train a logistic regression classifier to classify the images into 10 classes. Draw the ROC curve for the test dataset. Repeat the same with randomized PCA and compare.

Task 2: Train a single layer autoencoder with linear activation function and appropriately mean and variance normalized input with constraint that encoder weight matrix and decoder weight matrix are transpose w,r,t, each other. Compare the eigenvectors obtained in step 1 with those obtained using the autoencoders. Explain your observations.

Task 3: Train an appropriate deep convolutional autoencoder with same dimension of latent space. Calculate the reconstruction error fand compare that with a single hidden layer autoencoder (with sigmoid activation at the autoencoder and linear at the decoder) for the test dataset. What will be the reconstruction error if the hidden nodes are distributed equally (approximately) among 3 hidden layers in a new 3 hidden layer autoencoder with sigmoid activation at the autoencoder and linear at the decoder final layer?

Task 4. Train a deep convolutional autoencoder with MNIST dataset and using extracted features train a MLP classifier with 7 outputs (7 segment LED display) that are representative of 10 digits. For example images of "0" will be classified as

1

1 1

1

1 1

1

7 will be "classified" as

  1

0  1

  0

0  1

  0

Generate the confusion matrix for the corresponding test dataset.

Task 5: Upload both *.ipynb with all outputs embedded and corresponding *.html files.

Marks will be deducted for inadequate training resulting in higher errors in all tasks!