Abstract:

Using the MNIST dataset and CIFAR – 10 dataset three different models were built. They are: VAE, GAN and WGAN where I designed and trained the elements of these models for the two datasets.

Introduction:

MNIST dataset [1] has 60,000 training sample and 10,000 test samples where the input samples are images of handwritten digits in the range [0,1,2,3,4,5,6,7,8,9] with 10 classes in total. So, it's a multiclassification problem where we input an image and see the output being one among the 10 classes.

CIFAR-10 dataset [2] has 50,000 training samples and 10,000 test samples where the input samples are RGB images with different types of vehicles and animals with class labels specified as below:

-Airplane - class '0'

- automobile - class '1'

- bird - class '2'

- cat - class '3'

- deer - class '4'

- dog - class '5'

- frog - class '6'

- horse - class '7'

- ship - class '8'

- truck - class '9'

This dataset is also a multiclass classification problem with class labels in the range [0,9], so, for a given input sample i.e. an image we must predict the class of it.

Objective:

The goal of the assignment is to use VAE, GAN, WGAN models on the given datasets and build a image that is present in the dataset by passing a random noise vector / latent dimension.

Exploratory Data Analysis:

The MNIST dataset and CIFAR-10 datasets used were downloaded from TensorFlow keras datasets.

MNIST dataset:

There are 10 classes in the dataset and each class has approximately 6000 samples (training data) and 1000 samples (testing data) making it balanced.

Training Samples: 60000

Testing Samples: 10000

CIFAR – 10:

In cifar-10 as well we have 10 classes with the classes having an exact value of 5000 samples each (training data) and 1000 samples each (testing data).

Training Samples: 50000

Testing Samples: 10000

Model:

We train the data on both the datasets using the three different models.

1. Variational Autoencoder.
2. Generative adversarial networks.
3. Wasserstein Generative Adversarail Networks.

Performance and Conclusion:

All the three models, VAE, GAN and WGAN performed well on MNIST dataset as all the them were able to build images similar to the train images dataset, where as all the models didn't do well CIFAR10 dataset. However comparatively GAN and WGAN proved to be more effective than VAE model.

References:

[1] The MNIST DATABASE

[2] The CIFAR-10 dataset

[3]VAE, GAN, WGAN Code is highly inspired from these links.