Abstract:

Using the MNIST dataset and CIFAR – 10 dataset three different models were built. They are: SoftMax regression, Multi-Layer Perceptron (MLP) and CNN where I designed and trained image classifiers for these two datasets. Also, prettytable( ) is used to load the outputs into tabular values for ease of comparison of the results.

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 soft-max regression, Multilayer perceptron (MLP), Convolutional Neural network (CNN) models on the given datasets and attain high performance along with a balance of bias-variance tradeoff.

Exploratory Data Analysis:

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

Since, both our datasets are multi-class classification, *categorical cross entropy* was used as the loss measure.

Model and Performance:

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

MNIST:

SoftMax model:

Initially, a model was created with SGD optimizer and a SoftMax optimizer, later the optimizer was from SGD to ADAM to attain an accuracy value of 0.91.

Multi-layer Perceptron:

In MLP model, various activation units like sigmoid, ReLu, and optimizers which includes SGD and Adam were used. With the introduction of Batch Normalization, the model started to overfit very early (2 epochs) with 0.97 accuracy score. So, dropouts were introduced, and best accuracy value attained is 0.983

Convolution Neural Network:

In this model, a simple CNN network with 2 layers was used where the activation function used was ReLu and an Adam optimizer. This model attained an accuracy of 0.99 with a very low training error and validation error.

Model	#Layers	Activation	optimizer	Epochs	Accuracy	Training Error	Validation error
Softmax Regression	0	Softmax	SGD	50	0.917	0.309	0.298
Softmax Regression	0	Softmax	ADAM	4	0.92	0.27	0.27
MLP	2	Sigmoid	SGD	80	0.917	0.288	0.277
MLP	3	Sigmoid	SGD	80	0.9	0.34	0.33
MLP	3	Sigmoid	Adam	3	0.95	0.14	0.14
MLP	3	ReLu	SGD	20	0.97	0.09	0.1
MLP	3	ReLu	ADAM	3	0.983	0.05	0.06
MLP-Batch Normalization	3	ReLu	ADAM	2	0.978	0.07	0.08
MLP-BN-Dropout	3	ReLu	ADAM	25	0.983	0.052	0.051
CNN	3	ReLu	ADAM	9	0.99	0.03	0.02

Figure 1 (MNIST): Models with number of layers, activation, optimizer, Epochs, Accuracy, Train error and Validation error values for each model.

CIFAR-10:

SoftMax model:

In this model, SGD optimizer and SoftMax activation was used to attain an accuracy of 0.36. Later the model by changing the optimizer to Adam and an accuracy of 0.387 was attained.

Multi-Layer Perceptron:

Like MNIST dataset, in CIFAR-10 as well multiple models were trained with varying activation functions (Sigmoid, ReLu), optimizers (SGD, Adam), kernels (He_normal), batch normalization and dropouts. Among the MLP models, the best accuracy score of 0.471 was attained for the MLP network with ReLu activation and Adam optimizer model.

Convolution Neural Network:

Different CNN models were trained by using ‘same’ padding, kernel of sizes (3,3),(5,5), combination of ((2,2),(4,4)), batch normalization and the dropouts. The best among them was CNN with multiple kernels, ReLu activation function, Adam optimizer with max-pooling and dropouts in it which attained an accuracy of 0.76.

Model	# Layers	Activation	optimizer	Epochs	Accuracy	Training Error	Validation error
Softmax	0	softmax	SGD	4	0.36	1.814	1.813
Softmax	0	softmax	ADAM	4	0.387	1.78	1.77
MLP	4	Sigmoid	SGD	20	0.14	2.289	2.288
MLP	4	ReLu	ADAM	5	0.471	1.48	1.48
MLP - Batch Normalization	4	ReLu	ADAM	1	0.38	1.7	1.8
MLP-BN-Dropout	4	ReLu	ADAM	40	0.47	1.488	1.379
CNN with (3,3) kernels	2	ReLu	ADAM	2	0.626	1.048	0.979
CNN with (5,5) kernels	2	ReLu	ADAM	3	0.657	0.975	0.972
CNN with multiple kernels	5	ReLu	ADAM	6	0.758	0.684	0.675

Figure 2 (CIFAR-10): Models with number of layers, activation, optimizer, Epochs, Accuracy, Training error and Validation error values for each model.

Conclusions:

Presented three different models (SoftMax regression, MLP and CNN) on MNIST and CIFAR-10 datasets for measuring the performance using categorical loss entropy as a loss measure. Among these three models, we can observe that the performance of CNNs outweighed MLP and SoftMax regression by a large margin. However, the accuracy score obtained on the CIFAR-10 dataset is lower than MNIST dataset.

References:

[1] The MNIST DATABASE http://yann.lecun.com/exdb/mnist/

[2] The CIFAR-10 dataset https://www.cs.toronto.edu/~kriz/cifar.html