ECE 8870 Project 1

Qazi Zarif Ul Islam

Pawprint: qzic2d

Neural Networks

University of Missouri

University of Missouri-columbia

02/29/2024

1 Technical Description

The project described henceforth is one that aims to understand the working 'principle' of a neural network. In principle, a neural network has several core components; the inputs and outputs, the computational units known as 'neurons', the weights, the activation function, an optimizer and an error function. A neural network's components work together to form composite functions that ultimately combine to generate a final, function. Thus objective of 'understanding' the working principle of a neural network implies the attempt to understand how the core components behave in different configurations. With this aim in mind, this project contains two parts; Part 1- 'Sanity Check': In this part, 4 multi-layered perceptrons of different configurations were employed to classify a synthetic, linearly separable classes of 3 classes. Part 2- 'MLP vs CNN': In this part, the prediction performance of a convolutional neural net (CNN) and an MLP (Multi-layered perceptrons) were compared.

In the following sections will start with an explanation on the operation of a neural network. Then, activation functions and their derivatives shall be mathematically described and illustrated. Then, the algorithms used in this project for optimizing the neural networks shall be described, which are **Stochastic Gradient Descent (SGD)** and **Adam**. Finally the so called 'learning task' with respect to the datasets in both parts will be defined.

1.1 Operation of Neural Networks

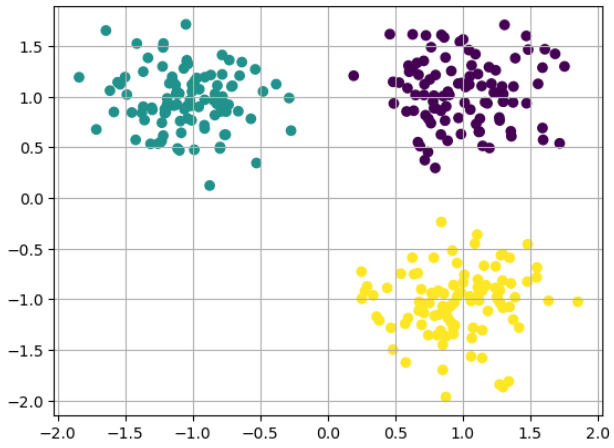

Figure 1: Synthetic data

2 Experiments & Results

The experiments were run in two Parts; Part 1- A classification task to classify a dataset containing 3 classes and Part 2- A classification task to classify the 'MNIST' dataset.

2.1 Part 1

The aim of this part was to verify that the program worked. The dataset used in this section is shown in Fig 1. The dataset was generated such that it would be linearly separable. This would facilitate the verification process since generating a discriminant function for a linearly separable dataset is a trivial task for a Neural Network using non-linear activation functions.

This simple classification task was also used as a means to study the activation functions; **Rectilinear Unit**, **Sigmoid** and **Hyperbolic tan**. Thus, several architectures were employed, each being composed of one of the 3 activation functions for all its hidden layers.

3 Conclusions and future work

The aim was not to attain high classification performance but rather to see the system in action.

Disregarded output classes sometimes, which benefitted performance. The merit of a classifier based on fuzzy sets lies within the interplay of the rules, membership functions and domain knowledge. The domain knowledge

The systems 1 and 2 had 3 fuzzy subsets and with the the better rule bases in system 3, a performance.

Thus, partitioning the overlapped region by having more subsets number and shape of fuzzy subsets over output domain could've been changed