Neural Network from scratch with NumPy for MNIST

Akbar Tangirov

Department of Electrical and Computer Engineering

Ajou University in Tashkent

`202290398@ajou.uz`

March 1, 2025

Abstract

In this report, we describe the development of a neural network built from the ground up, using only NumPy to train and test on the MNIST dataset. This approach demonstrates a fundamental understanding of deep learning concepts without relying on high-level machine learning frameworks.

1 Introduction

The MNIST dataset is a benchmark dataset consisting of 28×28 pixel handwritten digits. Our project aims to implement and understand each component of a neural network, including forward propagation, loss calculation, backpropagation, and parameter updates.

2 Methodology

2.1 Data Preparation

We load the MNIST data, normalize it, and reshape images for input into the network.

2.2 Network Architecture

Our network includes an input layer, one or more hidden layers, and an output layer. We use common activation functions such as ReLU or Sigmoid, depending on requirements.

2.3 Training Process

We apply mean squared error or cross-entropy loss, then compute gradients during backpropagation. Model parameters are updated using gradient descent.

3 Results

After training, the model's accuracy on the MNIST test set is evaluated. We present a summary of key metrics, confusion matrices, and example predictions.

4 Conclusion

Building a neural network from scratch with NumPy deepens understanding of matrix operations, gradient-based optimizations, and the fundamentals of deep learning.