Project Design Phase-I

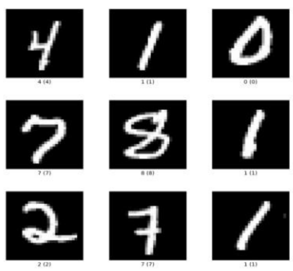
Solution Architecture

Date	5 October 2022
Team ID	PNT2022TMID41168
Project Name	A Novel Method for Handwritten Digit Recognition System
Maximum Marks	4 Marks

MNIST Dataset:

One of the interesting research projects is the recognition of handwriting. It is the ability of a computer to automatically recognise and comprehend handwritten numbers or letters. Every aspect of life is being digitalized to lessen the need for human labour as a result of advancements in science and technology. Thus, handwritten digit recognition is required in many real-time applications. The MNIST data collection, which contains 70000 handwritten digits, is frequently utilised for this recognition method. In order to train these photos and create a deep learning model, we use artificial neural networks. A web application is developed that allows users to upload pictures of handwritten numbers. This image is examined by the model, which then sends the results back to the user interface.

There are 60,000 training and 10,000 testing labelled handwritten digit images in the MNIST Handwritten Digit Recognition Dataset. There are 28 pixels in height and 28 pixels in width in each image, for a total of 784 (2828) pixels. A single pixel value corresponds to each pixel. It tells whether a pixel is bright or dark (larger numbers indicates darker pixel). An integer from 0 to 255 makes up this pixel value.



PROCEDURE:

• Set up the most recent TensorFlow library. •Configure the model's dataset.

•Construct a single layer perceptron model to categorise the handwritten digits.

•Plot the accuracy change over time.

•Assess the model based on the test data.

• Examine the summary of the model.

•To construct a multi-layer perceptron, add a hidden layer to the model.

•Incorporate Dropout to avoid overfitting and evaluate its impact on accuracy.

• Adding more Hidden Layer neurons and evaluating the impact on accuracy.

• Test the impact of various optimizers on accuracy.

• Boost the number of hidden layers and assess the impact on accuracy.

•Test the impact of changing the batch size and epochs on accuracy.

This project will be approached utilising a three-layered neural network.

The input layer:

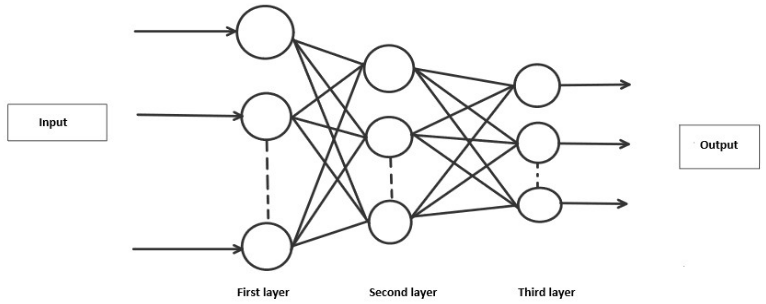
▪ It transfers the characteristics from our example layers to the following layer so that the subsequent layer's activations can be calculated.

The hidden layer:

▪ These ties for the network are built up of hidden units known as activations. Depending on our needs, there can be a variety of concealed layers.

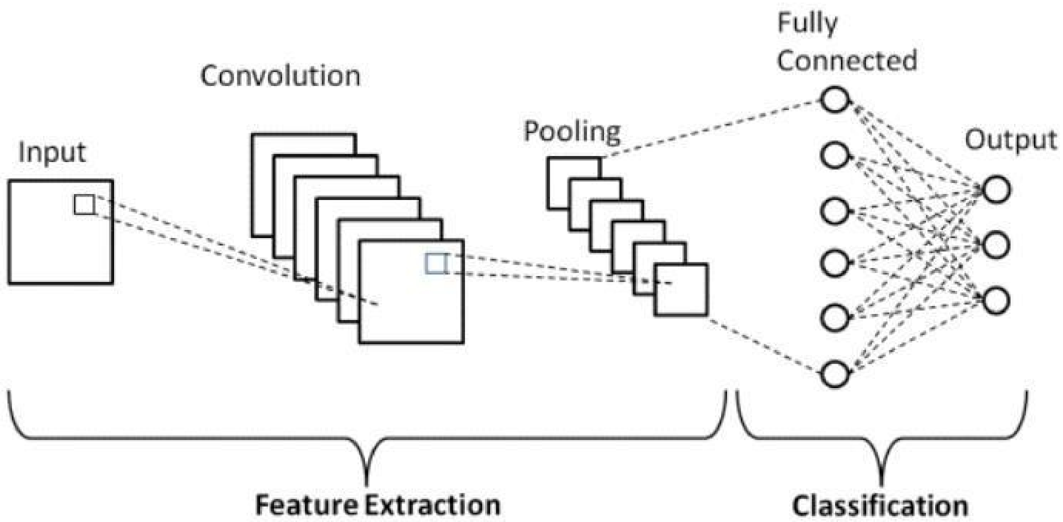
The output layer:

▪ The nodes in this layer are referred to as output units. It gives us access to the neural network's final prediction, which may be used to make final predictions.



Forward Propagation Process:

It's a simple process through which the CNN module will extract the features and categorise the image using them. The architecture displays the network's input layer, hidden layers, and output layer. The feature extraction phase of the network involves multiple layers and uses convolution and subsampling.



Working:

Neural networks take in data and process it through a number of secret layers. Each hidden layer is composed of a group of neurons, each of which is completely linked to every neuron in the layer above. A single layer of neurons has totally independent functioning. "Output layer" refers to the final layer that is entirely connected.

Convolution Layer:

The fundamental component of a CNN is the convolutional layer. The parameters of the layer are a set of learnable filters that cover the entire depth of the input volume but have a narrow receptive field. Each filter is convolved across the width and height of the input volume during the forward pass, computing the dot product between each filter entry and the input to create a two-dimensional activation map of the filter. As a result, the network picks up filters that turn on when they spot a certain kind of feature at a particular location in the input.

Feature Extraction:

The weights of each neuron in a feature are the same. In this manner, the same feature is recognised by all neurons at various locations in the input image. Limit the number of unrestricted parameters.

Subsampling Layer: Reducing the overall size of a signal is referred to as subsampling, sometimes known as down sampling. Each feature map's spatial resolution is decreased by the subsampling layers. Shift or distortion invariance is attained, and the impact of sounds is lessened.

Pooling layer:

In a Convent architecture, it is typical to sporadically introduce a Pooling layer between succeeding Conv layers. In order to decrease the number of parameters and computation in the network and, as a result, control overfitting, it gradually shrinks the spatial size of the representation. Every depth slice of the input is independently processed by the Pooling Layer, which then applies the MAX operation to resize each slice spatially.

TensorFlow:

An open-source machine learning library for both research and production is called TensorFlow. TensorFlow provides developers of all skill levels with APIs for desktop, mobile, web, and cloud applications. To get started, refer to the sections below. We can achieve text output and sound output by scanning the number digit and converting it to png format using the python3 command in terminal.

RESULT:

We are not claiming that our results are infallible, as with any endeavour in the fields of machine learning and image processing. There is always opportunity for methodological development in the field of machine learning; there will always be a fresh new idea that solves the same problem more effectively. Three models were used to test the application: Convolution Neural Network, Multi-Layer Perceptron (MLP), and (CNN). The classifier accuracy varies with each model, allowing us to determine which is more accurate.