

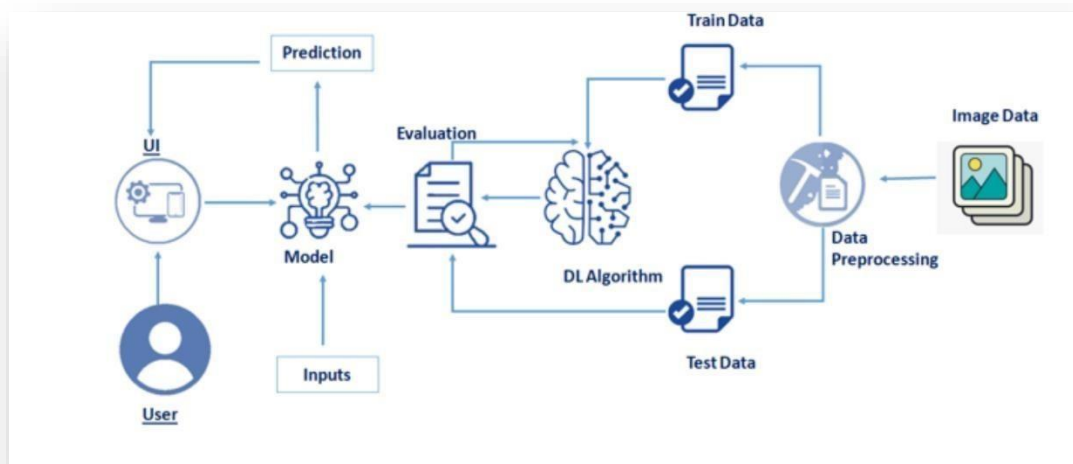
SOLUTION ARCHITECTURE

Date	19 September 2022
Team ID	PNT2022TMID01546
Project Name	A Novel Method for Handwritten Digit Recognition System
Maximum Marks	4 Marks

PROJECT DESCRIPTION:

Everyone in the world has a distinctive writing style, and one of the exciting scientific initiatives now being performed is handwriting identification. It refers to a computer's capacity to automatically discern and understand handwritten numbers or letters. As a result of advances in science and technology, every element of life is being digitized to reduce the need for human labor. In many real-time applications, handwritten digit recognition is therefore necessary. This recognition method typically uses the MNIST data collection, which has 70000 handwritten digits. We employ artificial neural networks to train these images and build a deep learning model. The creation of a web application that enables users to upload images of handwritten digits. This image is examined by the model, and the detected result is returned to the UI.

TECHNICAL ARCHITECTURE:

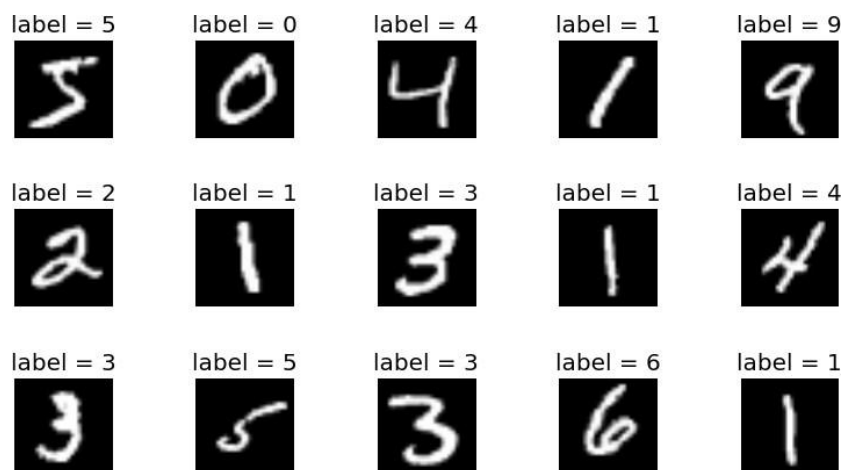


SOLUTION:

MNIST Dataset Description:

One of the exciting research projects being undertaken right now is handwriting identification because everyone in the globe has a different writing style. It refers to a computer's capacity to automatically discern and understand handwritten numbers or letters. As a result of advances in science and technology, every element of life is being digitized to reduce the need for human labor. In many real-time applications, handwritten digit recognition is therefore necessary. This recognition method typically uses the MNIST data collection, which has 70000 handwritten digits. We employ artificial neural networks to train these images and build a deep learning model. The creation of a web application that enables users to upload images of handwritten digits. This image is examined by the model.

The MNIST Handwritten Digit Recognition Dataset contains 60,000 training and 10,000 test labelled handwritten digit images. 784 (28x28) pixels, or 28 pixels in height and 28 pixels in breadth, make up each image. Each pixel has a single pixel value given to it. It shows the pixel's brightness or blackness (larger numbers indicate darker pixel). This pixel's integer value falls between 0 and 255.



PROCEDURE:

- Install the TensorFlow library.
- Prepare the model's dataset.
- For the purpose of classifying the handwritten digits, create a single layer perceptron model.
- Plot the accuracy change over time.
- Analyze the test data to evaluate the model.
- Speculate on the model summary.
- To make the model a multi-layer perceptron, add a hidden layer.
- To avoid overfitting and assess its impact on accuracy, include Dropout.
- Increase the number of hidden layer neurons and assess how accuracy is affected.
- Test the impact of various optimizers on accuracy.
- Increase the hidden layers and assess the accuracy impact.
- Change the batch size and epochs, then assess the impact on accuracy.

The MNIST collection is frequently used for the recognition of handwritten digits. The dataset consists of 60,000 training images and 10,000 test shots. Artificial neural networks are heavily used in the field of image processing because they most closely mimic the human brain.

The MNIST dataset was used in a significant effort that used neural networks to recognize handwritten numbers. In essence, it can read handwritten and scanned digits.

With the help of an integrated GUI, our handwritten digit recognition technology goes a step further and can now distinguish between handwritten numbers put directly on the screen and those found in scanned images.



APPROACH:

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

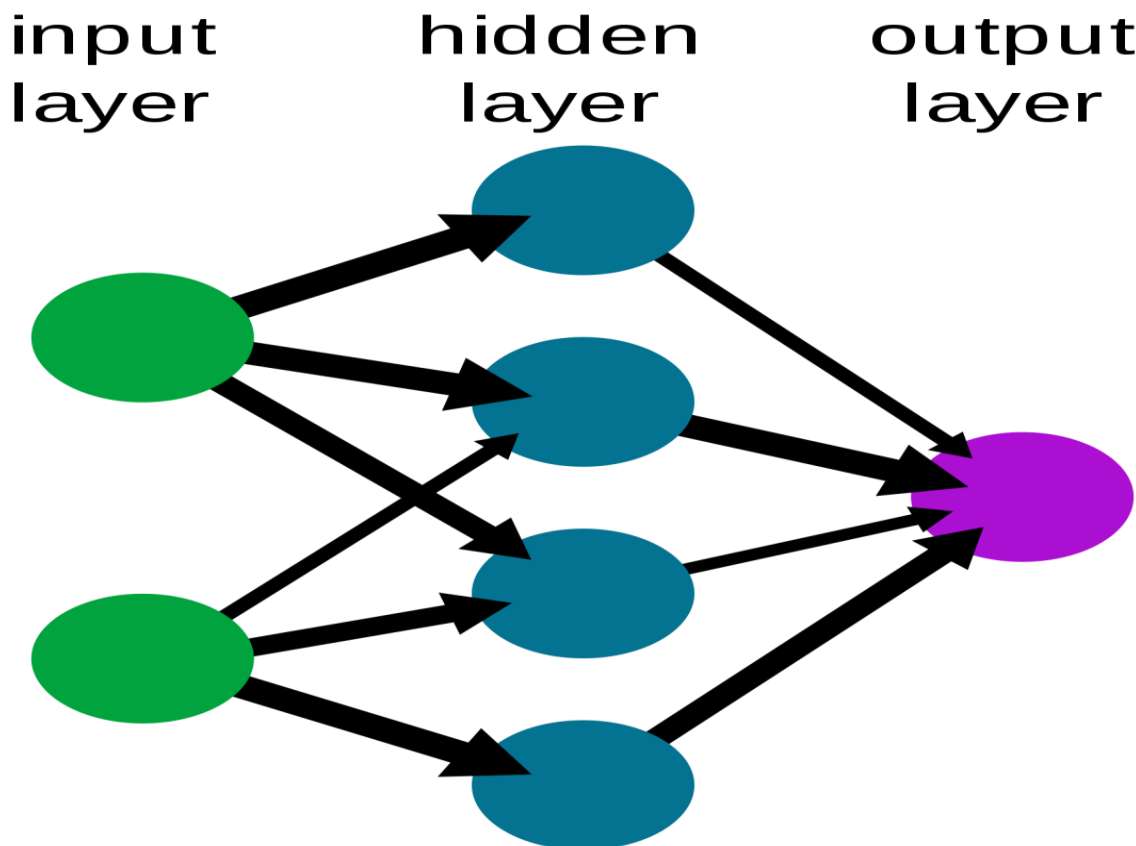
- **The input layer:** The input layer transfers the information from our example systems to the following layer so that the latter can compute its activations.
- **The hidden layer:** The network's nonlinear ties are provided by hidden units termed activations that make up the hidden layer. Depending on our needs, there can be a variety of concealed layers.
- **The output layer:** The nodes in this stratum 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.

A neural network is a simulation of how the brain works. It consists of many layers that have different activations; these activations resemble the neurons in our brain. A neural network makes an effort to discover parameters from a set of data that may help in comprehending the underlying linkages. The network can produce the optimal result without modifying the output criterion since neural networks can adapt to changing input it will happen in the program.

METHODOLOGY:

A neural network with one hidden layer and 100 activation units has been put into practice (excluding bias units). The features (X) and labels (Y) were retrieved after the data was loaded from a.mat file. To prevent overflow during computation, features are then scaled into a range of $[0,1]$ by dividing by 255. 10,000 testing cases and 60,000 training examples make up the data. With the training data, feedforward is used to calculate the hypothesis, and backpropagation is then used to lower the error between the layers. To combat overfitting, the regularization parameter λ is set to 0.1. To identify the model that fits the situation the optimizer runs for 70 times.

A simple neural network



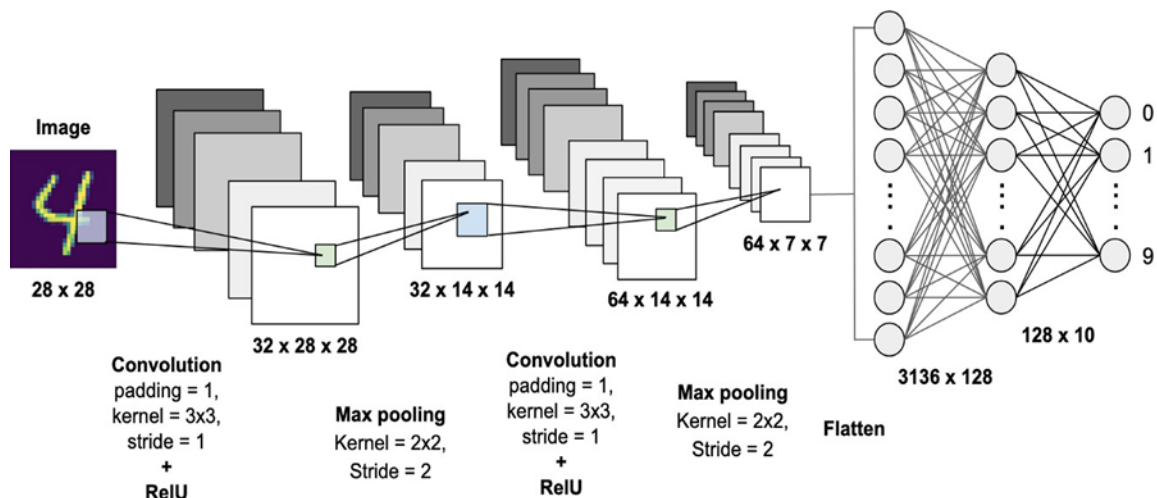
ALGORITHM:

Forward Propagation Architecture:

A simple process for the CNN module's feature extraction and picture classification is shown below. The network's input layer, hidden layers, and output layer are all displayed in the architecture. The feature extraction phase of the network involves multiple layers and uses convolution and subsampling.

EXPLANATION FOR THE PROPOSED SYSTEM

- The User layer is the top layer of the architecture. The users who engage with the programme and get the desired outcomes make up the user layer.
- The frontend architecture of the application is made up of the following three levels.
- The application will be created on the open-source JavaScript, CSS, and HTML platform.



- ❖ The localhost, which is displayed in the browser, is where the programme is deployed. The user will be able to upload images of the handwritten numbers to the app to have them digitized.
- ❖ The business layer, which consists of logical calculations based on the client's request, sits between the database and view layers. The service interface is also included.
- ❖ Training Data and Test Data make up the backend layer's two datasets. The training set, which consists of 60,000 cases, and the test set, which consists of 10,000 examples, have already been separated into the MNIST database.
- ❖ A convolution neural network is utilized as the training algorithm. This will get the trained model ready to classify the data.

WORKING:

- ❖ After receiving an input, neural networks change it using a number of hidden layers.
- ❖ Each group of neurons in a hidden layer is completely linked to every other neuron in the layer above it.
- ❖ One layer of neurons have perfect independence from one another.
- ❖ The "output layer" is the final layer to be fully connected.

CONVOLUTION LAYER:

The foundational component of a CNN is the convolutional layer. The parameters of the layer are a set of learnable filters (or kernels) 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 it detects a certain kind of feature at a particular spatial location in the input.

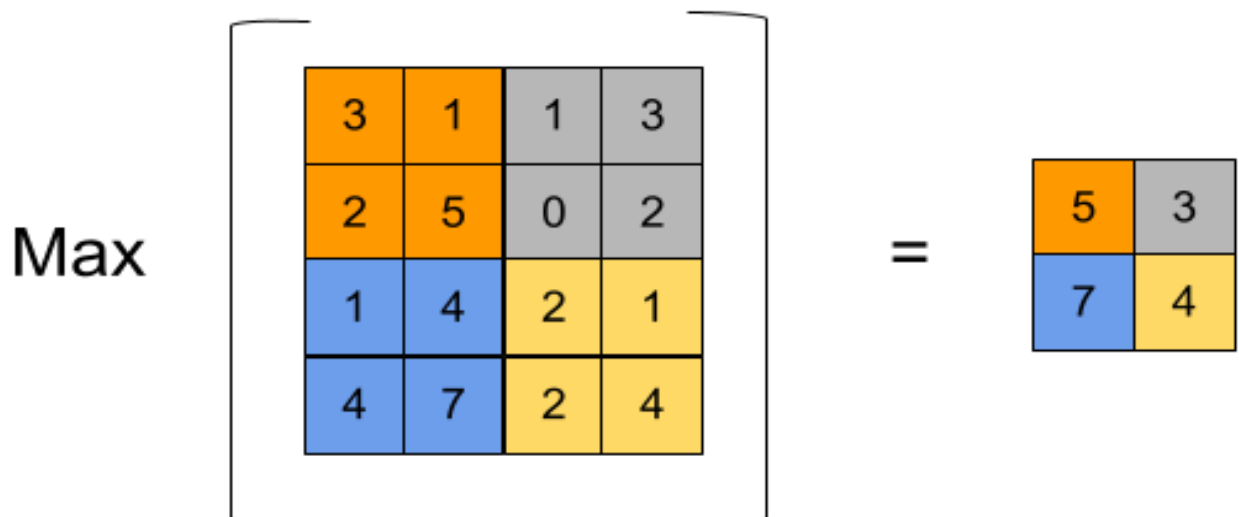
FEATURE EXTRACTION:

All neurons in a feature share the same weights. In this way all neurons detect the same feature at different positions in the input image. Reduce the number of free parameters.

SUBSAMPLING LAYER:

Subsampling, or down sampling, refers to reducing the overall size of a signal. The subsampling layers reduce the spatial resolution of each feature map. Reduce the effect of noises and shift or distortion invariant is achieved.

POOLING LAYER:



It is common to periodically insert a Pooling layer in-between successive Conv layer in a Convnet architecture. Its function is to progressively reduce the spatial size of the representation to reduce the number of parameters and computation in the network, and hence to also control overfitting. The Pooling Layer operates independently on every depth slice of the input and resizes it spatially, using the MAX operation.

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 the terminal.

RESULT:

As with any study or project conducted in the fields of machine learning and image processing, we do not consider our results to be perfect.

Because machine learning is a field that is constantly evolving, there is always room for improvement in your approaches. There will always be a brand-new concept that more successfully addresses a certain issue. The application was evaluated using three models: Multi-Layer Perceptron (MLP), Convolution Neural Network, and (CNN). With each model, we get a different classifier accuracy, showing which is better.