SOLUTION ARCHITECTURE

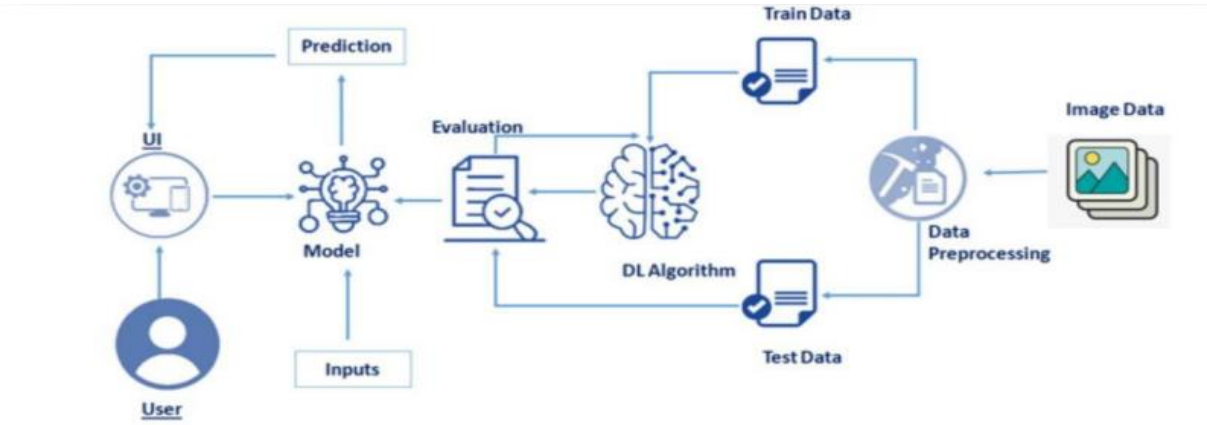
PROJECT: " A Novel Method for Handwritten Digit Recognition System"

Team ID: PNT2022TMID03792

PROJECT DESCRIPTION :

Everybody in the world has a distinctive writing style, and one of the exciting scientific initiatives now being performed is handwriting identification. It is the capacity of a computer to automatically detect and understand handwritten numbers or characters. As a consequence of advances in science and technology, every element of life is being digitalized to reduce the need for human labour. Therefore, a lot of real-time applications need handwritten digit recognition. For this recognition approach, the MNIST data collection, which comprises 70000 handwritten digits, is widely employed. We apply artificial neural networks to train these images and develop a deep-learning model. Users can upload images of handwritten numbers using a web application that has been created. The model looks at this image, and the UI receives the outcome of the detection.

TECHNICAL ARCHITECTURE :

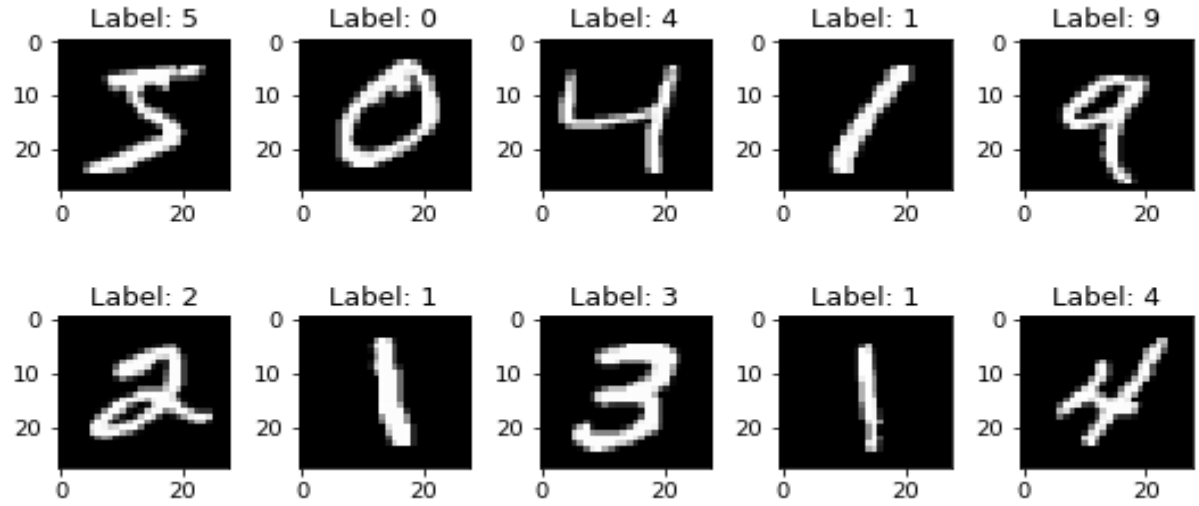


SOLUTION :

MNIST Dataset Description :

One of the intriguing research projects being pursued right now is handwriting identification because everyone in the globe has a distinct writing pattern. It is the potential of a computer to identify and interpret handwritten numerals or characters automatically. As a result of progress in science and technology, every aspect of life is becoming digitized to minimize the demand

for human manpower. Therefore, a lot of real-time applications need handwritten digit recognition. 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-labeled handwritten digit images. Each image has a total of 784 (2828) pixels or 28 pixels in height and 28 pixels in width. Each pixel has a single pixel value given to it. It shows the pixel's brightness or blackness (larger numbers indicate darker pixels). This pixel's integer value falls between 0 and 255.



PROCEDURE:

- Install the TensorFlow library.
- Create the dataset for the model.
- Create a single layer perceptron model to categocategorizeandwritten digits.
- Plot the accuracy change over time.
- Analyze the test data to evaluate the model.
- Speculate on the model summary.
- Include a hidden layer to transform the model into a multi-layer perceptron.
- Include Dropout in order to prevent overfitting and evaluate its effect on accuracy.
- Analyze the impact of adding more hidden layer neurons on accuracy.
- Test the impact of various optimizers on accuracy
- Increase the hidden layers and evaluate the influence on accuracy
- Assess the influence on accuracy after changing the batch size and epochs.

The MNIST dataset is commonly used for handwritten digit recognition. The dataset consists of 10,000 test photographs and 60,000 training shots. Artificial neural networks, which most closely mimic the human brain, are widely used in the field of image processing.

The recognition of handwritten digits using the MNIST dataset is an important effort that makes use of neural networks. It essentially detects handwritten and scanned numerals. With the help of an integrated GUI, our handwritten digit recognition technology goes a step further

and can now distinguish between handwritten digits put directly on the screen and those seen in scanned images.

APPROACH:

Three-layered neural networks will be used to approach this project:

INPUT LAYER: The input layer is responsible for sending data from our example systems to the layer above so that it can compute its activations.

HIDDEN LAYER: The network's nonlinear ties are given by hidden components known as activations, which together make up the hidden layer. There may be several covert levels, depending on our needs.

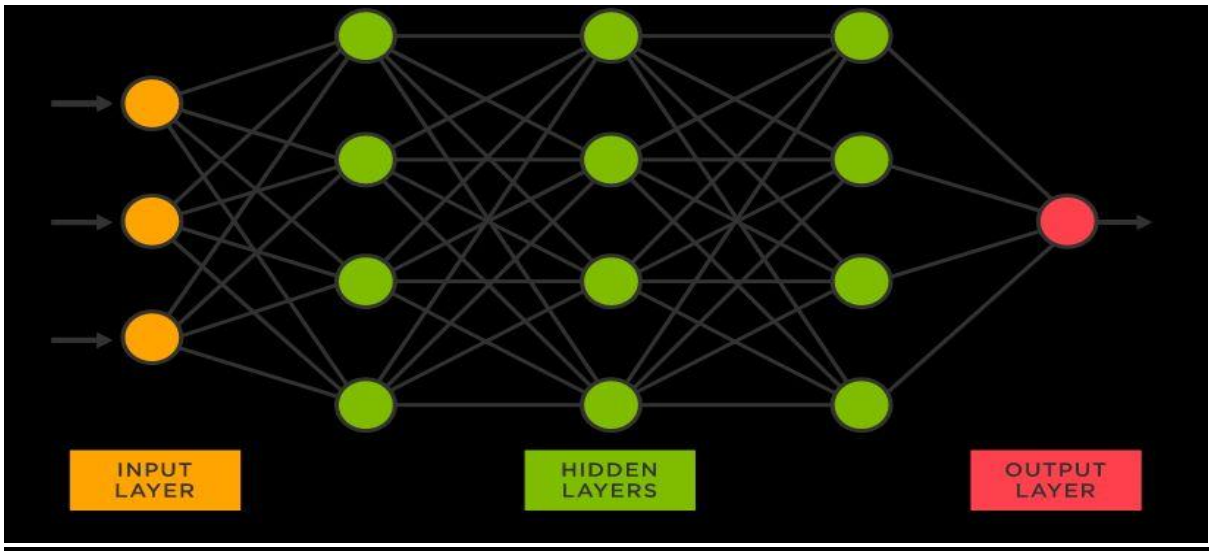
OUTPUT LAYER: The output layer's nodes are referred to as output units in this stratum. It gives us access to the final prediction made by the neural network which may be used to make final predictions.

A neural network is a simulation of how the brain processes. It consists of many layers that have different activations: these activations resemble the neurons in our brains. A neural network makes an effort to discover a collection of parameters from a set of data that could 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.

METHODOLOGY:

One hidden layer and one hundred activation units of a neural network have been used in practical solutions. The data was loaded from the a.mat file. Afterward, the features (X) and labels (Y) were extracted. Features are then divided by 255 to scale them into a range of [0,1] in order to prevent calculation overload. The data consists of 10,000 testing instances and 60,000 training examples. Feedforward is utilized to compute the hypothesis using the training

data, and backpropagation is then used to reduce the error between the layers. The regularisation parameter lambda is set at 0.1 to prevent overfitting. The optimizer runs 70 times to find the best model for the case.



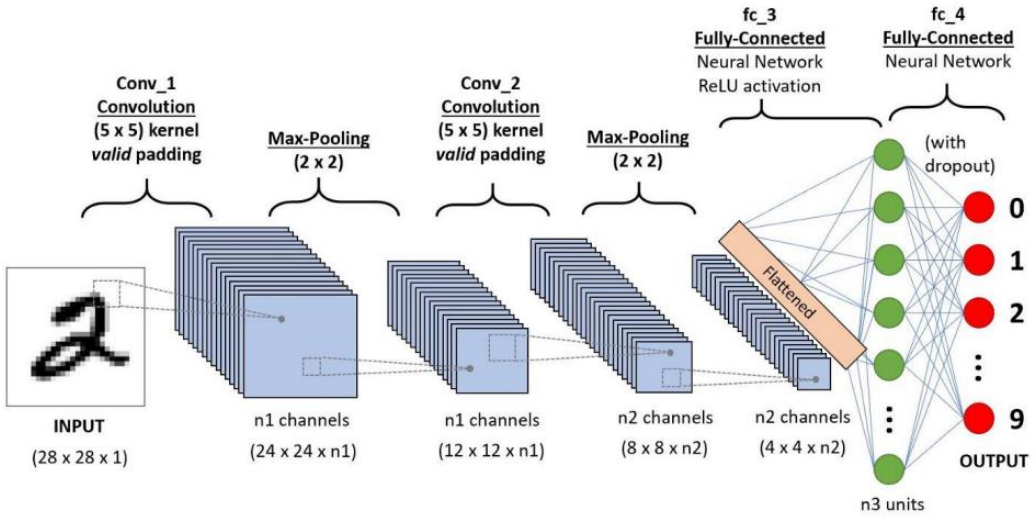
ALGORITHM :

Forward Propagation Architecture :

The CNN module's feature extraction and image categorization workflow is illustrated simply below. The architecture depicts the input layer, hidden layers, and output layer of the network. Convolution and subsampling are used during the network's feature extraction phase, which requires multiple layers.

EXPLANATION FOR THE PROPOSED SYSTEM:

- The architecture's top tier is the User layer. The user layer is made up of people who interact with the software and achieve the intended results.
- The following three tiers make up the frontend architecture of the application.
- The JavaScript, CSS, and HTML open-source platforms will be used to build the application.
- The software is deployed on the local host, which is visible in the browser. The user may upload pictures of the handwritten numerals to the app to have them converted to digital form.
- Between the database and display layers lies the business layer, which consists of logical computations driven by the client's request. There is also the service interface.
- The two datasets on the backend layer are Training Data and Test Data. The MNIST database already contains the training set, which consists of 60,000 instances, and the test set, which consists of 10,000 samples.
- The training technique makes use of a convolution neural network. The trained model will then be prepared to categorize the data.



WORKING:

Neural networks use a number of hidden layers to modify information after receiving it. Each cluster of neurons in a hidden layer is tightly connected to every other cluster of neurons in the layer above it. The neurons in one layer are completely independent of one another. The last layer to be fully connected is the "output layer."

CONVOLUTION LAYER :
The foundational component of a CNN is the convolutional layer. The parameters of the layer are a collection of learnable filters (or kernels) that have a small receptive field but cover the entire depth of the input volume. During the forward pass, each filter is convolved across the width and height of the input volume, and a two-dimensional activation map of the filter is produced by computing the dot product between each filter entry and the input. As a result, when a specific type of feature is found at a specific spatial location in the input, the network picks up filters that activate.

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 :
Subsampling, or down sampling, refers to reducing the overall size of a signal. Each feature map's spatial resolution is decreased by the subsampling layers. Reducing the impact of sounds and achieving shift or distortion invariance.

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:

TensorFlow is an open-source machine learning package that may be used for both research and production. TensorFlow offers APIs for desktop, mobile, online, and cloud applications to developers of different levels of experience. Refer to the sections below to get started. By using the python3 command at the terminal to scan the number digit and convert it to PNG format, we can produce text output and sound output.

RESULT:

Our techniques are always being improved because the field of machine learning is constantly evolving. There will always be novel solutions to issues. Three models were used to assess the application: Multi-Layer Perceptron (MLP), Convolution Neural Network, and (CNN). We obtain a different classifier accuracy for each model, indicating which is more accurate.