**CNN architecture to classify the MNIST handwritten dataset**

**using Machine Learning.**

Joy, Ovizit Das  
18-37945-2Computer Science & Engineering  
American International University-Bangladesh  
Dhaka,Bangladesh  
ovidasjoy@gmail.com

Mahjabin Rahman  
19-40464-1Computer Science & Engineering  
American International University-Bangladesh  
Dhaka,Bangladesh  
[mahzabin1699@gmail.com](mailto:mahzabin1699@gmail.com)

Asif, Abu Junaed Mohd  
19-40464-1Computer Science & Engineering  
American International University-Bangladesh  
Dhaka,Bangladesh  
junaedasif@gmail.com

Fahim, Abu Talha  
18-38125-2Computer Science & Engineering  
American International University-Bangladesh  
Dhaka,Bangladesh  
[emonfahim789@gmail.com](mailto:emonfahim789@gmail.com)

***Abstract***

MNIST contains a huge number of handwritten text data sets, which are frequently used for CNN deep model training, testing, and validation.In this tutorial, we created an efficient model with several relus and pooling layers.On the MNIST data set, which of the following is rated with 98.45% accuracy?This model is also tested on a similar type of random picture data set, which provides excellent accuracy results.Even a simple three-layer neural network can easily achieve accuracies of more than 98 percent (i.e. an error of less than 2%).To play with that dataset, no preprocessing is required because all digits are size-normalized and centered in the image.

# **Introduction**

Recognizing handwritten numbers is critical, and it can be used in a variety of ways in online handwriting recognition. There are numerous challenges to face when attempting to solve this problem. Handwritten digits contain a wide range of strokes, widths, thicknesses, orientations, and margin lengths, making identification more difficult. The goal was to use CNN to recognize a handwritten digit with a similar pattern and develop a model that could categorize a digit based on its pattern. A Convolutional Neural Network (CNN) is a form of artificial neural network that is used to analyse visual data in deep learning. The greatest performance on the handwritten digit recognition challenge from random photos and the MNIST dataset was demonstrated in this paper using an unique CNN model. The images are of handwritten numeral digits retrieved from a source. Modifying the model architecture was attempted to increase the accuracy to over 98 percent. To determine actual accuracy, it is also tested with various optimizers (Adam, SGD, RSMProp)to identify the actual accuracy.

**Methodology**

Recognizing handwritten numbers is critical, and it can be used in a variety of ways in online handwriting recognition.

There are numerous challenges to face when attempting to solve this problem. Handwritten digits contain a wide range of strokes, widths, thicknesses, orientations, and margin lengths, making identification more difficult. The goal was to use CNN to recognize a handwritten digit with a similar pattern and develop a model that could categorize a digit based on its pattern. A Convolutional Neural Network (CNN) is a form of artificial neural network that is used to analyse visual data in deep learning. The greatest performance on the handwritten digit recognition challenge from random photos and the MNIST dataset was demonstrated in this paper using an unique CNN model. The images are of handwritten numeral digits retrieved from a source. Modifying the model architecture was attempted to increase the accuracy to over 98 percent. To determine actual accuracy, it is also tested with various optimizers (Adam, SGD, RSMProp)to identify the actual accuracy.

## **Dataset**

Modified The National Institute of Standards and Technology (MNIST) is a free database for handwritten digits that serves as a standard for machine learning algorithms.It is similar to TIDigit, a speech database created by Texas Instruments that performs speech recognition tasks. The MNIST dataset is being used in our project.The images of digits in this dataset were taken from a variety of scanned documents, and each image is Greyscaled and 28\*28 pixels in size.It employs 60,000 images to train the network and 10,000 images to assess how well the network learned to classify the images.Some of the MNIST dataset's sample images are shown below.

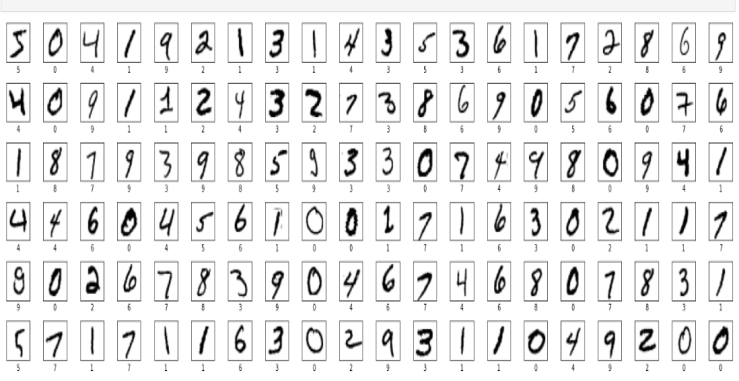


Fig: Images of our proposed MNIST dataset

An API is provided to automatically download and

extractimages and labels from the MNIST dataset in Keras.

The task is to place a given input image of a handwritten digit into one of the ten classes representing integer values ranging from 0 to 9 inclusively.The distribution of training data in MNIST is depicted below.

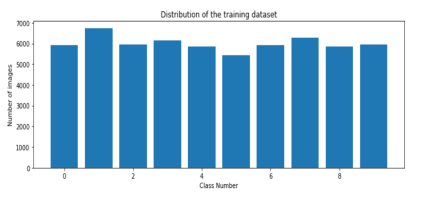
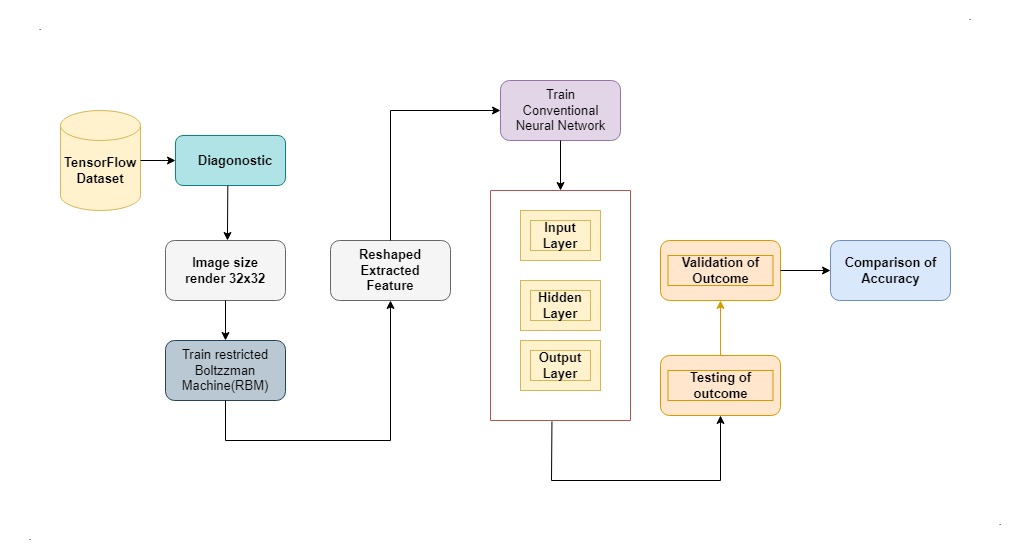


Fig : Bar diagram showing the distribution of training Dataset.

**Used Dataset**

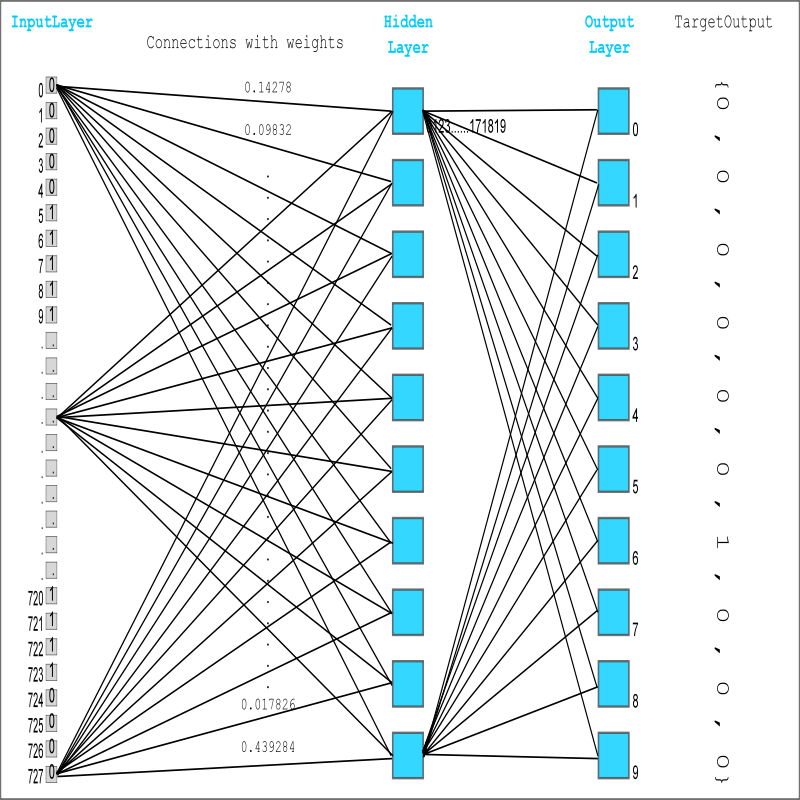
Here we used TensorFlow Datasets: A collection of ready-to-use datasets.TensorFlow Datasets is a collection of datasets ready to use, with TensorFlow or other Python ML frameworks, such as Jax.All datasets are exposed as tf.data.Datasets , enabling easy-to-use and high-performance input pipelines.To get started see the guide and our list of datasets.Keras includes the MNIST dataset as a built-in dataset.It is made up of 70,000 28x28 grayscale images, each displaying a single handwritten digit from 0 to 9. The training set is made up of 60,000 images, while the test set is made up of 10,000 images.The handwritten digits have been normalized and focused on consistency. The dataset is loaded from Keras using the load data() function.TensorFlow is a fantastic machine learning library created by Google's Brain Team and released as open source in 2015. It's designed to be simple to use and applicable to both numeric and neural system challenges, as well as varied spaces. TensorFlow is a low-level tool for doing entangled math that is aimed at specialists who know what they're doing when it comes to building exploratory learning structures, playing with them, and turning them into working programs. For the most part, it may be seen of as a programming framework in which calculations can be represented as graphs. The graph's nodes represent math activities, while the edges represent multi-dimensional information clusters (tensors) that link them together.

**Diagram**



**Fig: Whole Process**

**Network layer**

****

From one to three layers,

The addition of a hidden layer between input and output is the first major alteration in the architecture. So, how did a one-layer network become into a three-layer network? It's just a different name scheme. The output layer of the 1-layer network had only one layer of perceptrons. In our earlier approach, the network's input was not part of the network; instead, we referred to an MNIST picture, a variable outside of the network, anytime this input was required (when calculating a node's output and updating a node's weights).We found it advantageous to add the input feed inside the network when we revamped it. It enables the network to be treated as a standalone object (data structure) without the need for external references. As a result, the 1-layer network becomes a 3-layer network by adding a hidden layer and treating the input as a network layer as well.

##### **Acknowledgment**

Convolutional neural networks (CNNs) are extremely successful in perceiving the structure of handwritten characters/words in ways that allow for the automatic extraction of different characteristics, making CNN the best solution for solving handwriting recognition challenges. The goal of the proposed work is to investigate several design alternatives for CNN-based handwritten digit recognition, such as the number of layers, stride size, receptive field, kernel size, padding, and dilution. Our goal is to attain equivalent accuracy utilizing a pure CNN design rather than an ensemble architecture, because ensemble architectures have a higher computational cost and testing complexity.As a result, a CNN design is proposed to attain accuracy that is even greater than ensemble systems while also reducing operational complexity and expense.Furthermore, we demonstrate an appropriate combination of learning parameters in CNN design that leads to a new absolute record in categorizing MNIST handwritten digits.

**Discussion**

Deep learning methods for handwritten digit recognition have been implemented. The most commonly used machine learning algorithms, KNN, SVM, RFC, and CNN, were trained and tested on the same data in order to compare classifiers. A high level of accuracy can be obtained by utilizing these deep learning techniques. In comparison to other research methods, this method focuses on which classifier works best by improving classification model accuracy by more than 99 percent. A CNN model with Keras as the backend and Tensorflow as the software can achieve an accuracy of about 98.72 percent. In this preliminary experiment, CNN achieves an accuracy of 98.72 percent, KNN achieves an accuracy of 96.67 percent, and RFC and SVM perform poorly. RMSprop optimizer is preferred because to its accuracy, according to the report's findings. This optimizer has a 0.79 percent loss ratio. It has pseudo-curvature information and is a fairly robust optimizer. It is well-suited to stochastic objectives, making it suitable for micro batch learning. It is more convergent than momentum.

##### **References**

1. <https://mmlind.github.io/posts/simple_3-layer_neural_network_for_mnist_handwriting_recognition/>
2. <https://arxiv.org/pdf/2106.12614.pdf?fbclid=IwAR1qxpJgrxiO82hFspUQs65H0e_REQMBjx1i89uMDH29QNpd6nbFujC4sP0>
3. <https://ijcrt.org/papers/IJCRT2106267.pdf?fbclid=IwAR16ljadAxdOv_J_O_mG2RON7Z8KP8UTbzEPl1JCv0U2qGUcwWbnVME5Cbs>
4. <https://ijisrt.com/wp-content/uploads/2019/06/IJISRT19JU358.pdf?fbclid=IwAR3DZN-pMF00CZPb97rOaRVdgr6HvfkbW-nDlB5TAiYXat4t9wwQCKYfzaY>