

RESEARCH TITLE

**DEVELOPING A BASIC NEURAL NETWORK TO CLASSIFY
IMAGES FROM THE MNIST DATASET**

Hamza BenTarif¹ Prof. Dr. Ali Okatan²

¹ Istanbul Aydin University, Institute of Graduate Study, Computer Engineering (With Thesis) Program, 34295, Küçükçekmece, Istanbul. Email: hbentareef@hotmail.com

² Istanbul Aydin University, Faculty of Engineering, Department of Computer Engineering, 34295, Küçükçekmece, Istanbul. Email: aliokatan@aydin.edu.tr

HNSJ, 2023, 4(7); <https://doi.org/10.53796/hnsj479>

Published at 01/07/2023

Accepted at 20/06/2023

Abstract

My thesis focuses on an important and currently popular topic: “Developing a basic neural network to classify images from the MNIST dataset”

This topic is very important for current and future research. The main goal of this project is to create a basic neural network capable of efficiently classifying images from the MNIST dataset, a critical benchmark for image classification tasks. The MNIST dataset is widely known and widely used in computer vision and machine learning. Designing an accurate neural network to classify images from this dataset is a crucial milestone in the development of more advanced computer vision systems. By exploring the basic architecture of a neural network, this study aims to provide an overview of the basic principles of image classification tasks. The PhD thesis deals with the complex process of training a neural network using the MNIST dataset, evaluating its performance and fine-tuning its classification accuracy. Through this research, researchers and professionals gain a deeper understanding of the basic concepts and techniques involved in developing neural networks for image classification. In addition, the results of this study contribute to the wider body of knowledge in computer vision and machine learning, enabling further advances and applications in areas such as object recognition, pattern analysis and visual perception. Finally, the development of a basic neural network to classify images from the MNIST dataset is a crucial step towards building more advanced and sophisticated computer vision systems. This doctoral thesis lays the foundation for further research and innovation in the field of image classification, providing valuable knowledge and methods for further research.

Key Words: neural network, image classification, MNIST dataset, computer vision, machine learning, classification accuracy.

1- INTRODUCTION

There are many libraries available for designing and training neural networks, making them fascinating tools for developing excellent AI systems that are also enjoyable to learn and use. In other words, creating your own neural network from start will enable you to comprehend neural networks' fundamental operation and inner workings much better.

70,000 grayscale photos of apparel and accessories are broken down into 10 separate classes in the Fashion MNIST dataset, which is well known in computer vision and machine learning. The square format of each image is 28x28 pixels, which is very modest compared to other huge image datasets, but the changes in the photos that distinguish them from one another make it difficult for machine learning models to correctly identify the images.

The dataset can be used as a starting point for learning and practicing how to create, assess, and apply convolutional deep learning neural networks to categorize images even though it is successfully resolved. This covers how to create a reliable testing tool to gauge model performance, how to debug model improvements, and how to store and load the model later to generate predictions based on new data.

The MNIST dataset, which consists of handwritten digits, is a well-known public dataset. It has a test set of 10,000 cases and a training set with 60,000 examples. It was made by recombining the sounds of American high school students and employees of the American Census Bureau. While the test set has 5,000 photographs from employees and 5,000 from students, the training set has 30,000 images from both employees and students.

A MNIST data set is a group of data that is typically shown as a table. The table's rows and columns each correspond to a different variable in the data collection. The values for each variable in this element are listed in this table. It may, for instance, provide an object's length and width. Depending on how many rows there are in the data set, there may be one or more elements.

A common dataset for computer vision and deep learning is the MNIST handwritten digit classification issue. Dataset for the MNIST Handwritten Digit Classification the Modified National Institute of Standards and Technology dataset is known by the abbreviation MNIST dataset.

It is a collection of 60,000 tiny square grayscale photographs, each measuring 28 by 28, comprising handwritten single digits between 0 and 9.

The assignment is to place a handwritten digit picture into one of ten classes that correspond to integer values from 0 to 9, inclusively.

The dataset can be used as a starting point for learning and practicing how to create, assess, and apply convolutional deep learning neural networks for image classification even if it has been successfully solved. This covers how to create a reliable test harness to gauge the model's performance, how to investigate model changes, and how to save the model and then load it to make predictions on fresh data.

Neural networks are used as a method for deep learning, as they are one of the many sub-fields of methods for applying artificial intelligence. It was proposed for the first time about 70 years ago in an attempt to simulate the way the human brain works, but it is much simpler than real neurons, as each artificial cell is linked to several layers, and each one has a certain weight that expresses the importance of this layer, in order to determine how the cell responds. nervousness when spreading data across it.

Numerous neural networks feature tens of thousands to millions of weights, each of which is

manually adjusted during training to reduce error,

Numerous open studies are being conducted on the architecture of neural networks. It is not unusual for a new architecture to perform better than existing ones at a particular task. Practitioners frequently choose an architecture based on a recent publication, reproduce it in large quantities for a new assignment, or make tiny changes for a gradual improvement.

The complexity of the learning process that we could achieve at the time was constrained by the number of neurons that neural networks could simulate; however, in recent years, thanks to significant hardware development progress, we have finally been able to build deep neural networks very, and trained on extremely large data sets as well. This has resulted in enormous advances in the field of artificial intelligence, specifically.

These qualitative leaps allowed machines to approach human capabilities, and even exceed them in performing some limited tasks. Among these tasks is its ability to recognize objects. Although machines have historically been unable to compete with the power of human vision, recent advances in deep learning have made it possible to build neural networks that can recognize objects, faces, texts, and even emotions!

2- BACKGROUND INFORMATION :

2.1: An Introduction to Neural Network:

A neural network is a type of computing model that is loosely based on the cortical structures of the brain. It is sometimes referred to as a parallel-distributed processing network. It is made up of ganglia or neurons, which are interconnected processing units that collaborate to provide the output function. The ability of each individual neuron in a neural network to act cooperatively determines the output of the network. In contrast to older binary computers or Von Neumann machines, which processed information sequentially (or sequentially), neural networks process information in parallel. The special characteristic of a neural network is that it can still carry out its overall function even if some of its individual neurons are not functioning since it depends on member neurons to carry out its function.

By training and then reversing an image recognition network, a neural network can produce an image. Additionally, it has the ability to take two photographs with different aesthetics and create a single output with the same attributes as both input images. A CNN with less complexity and high-quality output is the best option for image output.

Amazing results for image processing can be achieved with deep learning in methods that are still being studied. It's crucial to comprehend how to produce an image using a neural network, but doing so necessitates some understanding of how CNNs operate internally.

Sometimes the term "neural network theory" refers to a field of computational science that models complicated events using neural networks and/or conducts analytic research on the neural networks' underlying principles. Neural networks employ "agent networks" (entities of software or hardware linked to each other) as the computational architecture to solve issues, in contrast to artificial intelligence (AI), which uses conventional computational procedures to do so. Neural networks are trainable systems that can "learn" to solve complicated issues from a collection of models and generalize the "acquired knowledge" to handle unforeseen challenges, such as those in environmental and stock market forecasting. They are self-adapting systems, in other words.

An artificial neural network (ANN), also known as a simulated neural network (SNN), or simply a neural network (NN), is a network of connected artificial neurons that processes information using a mathematical or computational model and relies on a communication-

based approach to computing. An ANN is often an adaptive system that modifies its structure in response to information passing through the network, whether it be internal or external.

Neural networks can be used to model nonlinear statistical data in real-world applications. They can be applied to identify patterns in data or to model intricate relationships between inputs and outputs.

In other words, it is difficult to accept failure or error. Furthermore, compared to von Neumann machines, neural networks are better suited to fuzzy logic computing applications.

A network of biological neurons has historically been referred to as a neural network. Nowadays, the phrase is frequently used to describe artificial neural networks, which are made up of synthetic neurons or nodes. Consequently, the phrase "neural network" has various meanings:

Real biological neurons that are connected to or functionally coupled to the peripheral nervous system or the central nervous system make up biological neural networks. They are frequently recognized in the field of neuroscience as collections of neurons that, upon laboratory study, carry out a particular physiological function.

A subset of deep learning neural networks is the convolutional neural network (CNN). CNNs are a significant advancement in image recognition. They are typically employed in the background of picture categorization and are most usually utilized to assess visual imagery.

Deep neural networks—those with numerous intermediate (hidden) layers—are used to categorize images. As a result, the networks are able to extract intricate elements from the photos. The latter layers generate an output in the form of a vector (in this case, one with 10 components), where each element represents the likelihood that the image belongs to a particular class. The NN receives an image as input and extracts features that enable this to happen. There are numerous layers with various goals built between the input and the output. However, there are a few fundamental layers that every CNN must have. Modern CNNs are highly complicated and feature a variety of various layers or blocks of layers that enhance performance.

In the modern world, neural networks are widely used for image processing. As the output layer in a machine-learning model, it can generate images. All of the desired attributes for your image are present in the output layer. There are countless options for producing a picture using a convolutional neural network (CNN), if you learn how to do it.

A piece of software called Neural Network Upscaling Tool employs deep learning image processing to upscale photos beyond their initial resolution (number of pixels).

Computers can make intelligent decisions with minimal human intervention thanks to neural networks. This is why they can learn and model complex, nonlinear relationships between input and output data. They can perform the following duties, for example: make broad generalizations and deductions, which without explicit training; neural networks are capable of understanding unstructured material and drawing broad conclusions.

In case you are unfamiliar, the following terms related to machine learning are included:

- Name y stands for labels, while Name x stands for input data. The model's predictions are referred to by the term (pronounced \hat{y}).
- Our model learns using training data, which are the data sets.
- Even after the model has been trained, test data is kept private. We assess our model with test data.
- The loss function is a way to gauge how well a model predicts the future.

- The training process for the computational graph's weights is precisely under the control of the optimization algorithm.

However, because of the groundbreaking successes of generative pre-trained adapters (GPT) and other adapter-based architectures in natural language processing, vision (ViT) adapters have also lately grown in popularity.

In general, there are various methods that neural networks analyze and recognize images. The network design and the issue we are trying to solve will determine this. Neural networks frequently use images to solve the following issues:

- **Image classification:** Image categorization entails labeling or categorizing the image. For instance, if the picture features a dog or a cat.
- **Object detection:** Locate and recognize items in an image.
- **Image segmentation:** Image segmentation is the process of breaking up an image into groups of pixels that are each represented by a mask or a label.
- **Image generation:** Image generation is the process of creating new images depending on specific specifications.

We will not go through the additional image-related issues that neural networks may handle, such as picture annotation, image retrieval, feature identification, human posture assessment, and pattern transmission.

Perceptual networks, Hopfield networks, Boltzmann machines, fully connected neural networks, convolutional neural networks, recurrent neural networks, long-term memory neural networks, autoencoders, deep belief networks, and generative adversarial networks are just a few examples of the many machine learning algorithms referred to as neural networks. The majority of them are trained via the backpropagation algorithm.

2.2 Background information of the MNIST dataset

The MNIST dataset is a sizable collection of handwritten digits that is frequently used to hone different image processing algorithms. The database is frequently used for machine learning training and testing. It was produced by "remixing" samples taken from the initial datasets from NIST. The NIST training dataset came from US Census Bureau employees, and the testing dataset came from US high school students, hence the developers concluded that it was not suitable for machine learning research. Additionally, grayscale levels were rendered using NIST's black-and-white images that had been adjusted to fit a 28 x 28 pixel bounding box and anti-aliased.



Sample images from MNIST test dataset

10,000 test images and 60,000 training images are both present in the MNIST database. The NIST training data set was used for half of the training set and half of the test set, while the NIST test data set was used for the other half of the training set and half of the test set. The database's original designers maintain a record of some of the ways that have been tried out on it. They obtained an error rate of 0.8% in their original study using a direction support machine.

NIST created and made available Extended MNIST (EMNIST), a more recent data set, as the (ultimate) replacement for MNIST. Only handwritten number images were included by MNIST. All 19 photos from NIST's own database, a sizable collection of handwritten uppercase, lowercase, and digits, are included in EMNIST. The EMNIST photos underwent the same procedure as the MNIST images in order to be transformed to the same 28x28 pixel format. Therefore, it is more likely that tools that function with older and younger unmodified MNIST datasets will also function with EMNIST.

The MNIST handwritten digit collection is one of the most frequently used datasets for picture classification. With 60,000 photos at the time of its debut in the 1990s, it presented a significant challenge to the majority of machine learning algorithms.

Pixel resolution (with a 10,000-image test data collection). To put things in perspective, in 1995 at AT&T Bell Labs, a Sun SPARCStation 5 with a huge 64MB of RAM and 5 MFLOPs was regarded as state-of-the-art hardware for machine learning. The USPS automated letter sorting in the 1990s by focusing on achieving high accuracy in number recognition. Error rates under 1% have been attained with deep networks like LeNet-5, SVMs with constants (Schölkopf et al., 1996), and tangent distance classifiers.

MNIST has been used as a benchmark for contrasting machine learning methods for more than ten years. Even simple models by today's standards attain classification accuracy of more than 95%, making them inappropriate for differentiating between stronger and weaker models, despite the fact that it worked well as a normative data set. Furthermore, the dataset enables extremely high precision levels, which are uncommon in many classification situations. This pushed algorithmic research toward particular families of algorithms, such as active set methods and boundary-seeking active set algorithms that may benefit from clean data sets.

MNIST is now more of a safety measure than a benchmark. The problem ImageNet presents is more significant. Unfortunately, it will take a lot of experience to make the examples interactive because ImageNet is so big for many of the examples and images in this book. Instead, we will concentrate on the qualitatively comparable, but considerably smaller Fashion-MNIST data set released in 2017 in the following sections.

Thousands of samples of handwritten numerals 0 to 9 can be found in the MNIST collection. For optical character recognition (OCR) systems, which transform handwritten addresses into computer scripts, for example, and help postal agencies automate mail delivery systems, this data has been employed in machine learning research for decades.

Researchers used a novel classifier dubbed LIRA, a neural classifier with three neural layers based on Rosenblatt's perceptual principles, to obtain a best-case error rate of 0.42 percent in the database in 2004.

Some researchers have used a database subjected to random distortions to test AI systems. In these situations, the systems are frequently neural networks, and the distortions are frequently affine or elastic distortions. These systems can occasionally be quite effective; one such system had a 0.39 percent database error rate.

The MNIST set is a sizable set of manually typed digits. This dataset is particularly well-liked in the world of image processing. It is frequently employed to evaluate machine learning algorithms.

Data scientists that wish to test machine learning and pattern recognition methods on real-world data while putting little effort into preprocessing and formatting sometimes utilize the MNIST digits dataset.

The modified National Institute of Standards and Technology database is known as MNIST.

MNIST has 70,000 28×28 photographs of handwritten digits from 0 to 9 in its collection.

Training and testing sets have already been created for the dataset. Later on in the tutorial, we shall see this.

A sizable collection of handwritten numbers can be found in the MNIST collection. It is one of the most well-liked data sets in the image processing industry. It is frequently used to evaluate machine-learning techniques.

A sizable database of handwritten numbers called the MNIST database (Modified National Institute of Standards and Technology database) is frequently used to train different image processing systems. The database is frequently used for machine learning training and testing.

It was made by "re-mixing" samples from the initial datasets from NIST. Since the testing dataset was collected from American high school students, and the training dataset was collected from American Census Bureau employees, the designers of NIST concluded that it was unsuitable for machine learning research. In addition, the NIST black and white photos were anti-aliased, which added grayscale levels, and normalized to fit within a 28×28 pixel-bounding box.

10,000 test images and 60,000 training images are both present in the MNIST database. The NIST training data set was used for half of the training set and half of the test set, while the NIST test data set was used for the other half of the training set and half of the test set. The database's original designers maintain a record of some of the ways that have been tried out on it.

They obtained an error rate of 0.8% in their original study using a direction support machine. The EMNIST extended data set, which is comparable to MNIST and consists of 40,000 test images and 240,000 training images for MNIST's data set of handwritten numbers and letters, was released in 2017.

A foundation for testing image-processing systems is provided by MNIST. It is like the "hello world" of machine learning if you will. For the sole purpose of testing a new architecture or framework and making sure it functions, data scientists will train an algorithm using the MNIST dataset.

The binary pictures of handwritten digits from NIST Database 3 and NIST Special Database 1 were combined to create the MNIST database. SD-3 and SD-1 were initially identified by NIST as the training and test sets, respectively. However, SD-3 is far more recognizable and cleaner than SD-1. This can be explained by the fact that SD-3 was gathered from Census Bureau personnel while SD-1 was gathered from students in high school. It is necessary for the outcome of learning trials to be independent of the choice of the training and test set among all possible samples in order to draw valid conclusions. Thus, a new database that was created by rearranging the NIST datasets was required.

Computer vision and deep learning commonly employ the MNIST handwritten digit classification problem as a standard data set.

The dataset can be used as a starting point for learning and practicing how to create, assess, and apply convolutional deep learning neural networks to categorize images even though it is successfully resolved. This covers how to create a reliable testing tool to gauge model performance, how to debug model improvements, and how to store and load the model later to generate predictions based on new data.

Although the MNIST dataset has been successfully solved, it might serve as a valuable starting point for formulating and honing a method for resolving image classification problems with convolutional neural networks.

The aspect ratio was kept while the original MNIST images were size-normalized to fit a 2020 pixel box. As a result of the anti-aliasing technique utilized by the normalizing algorithm, the images have grey levels (i.e., pixels don't just have a value of black and white, but a level of greyness from 0 to 255).

By calculating the center of mass of each pixel and translating the image to place this point in the center of the 2828 field, the images were then centered in a 2828 pixel image.

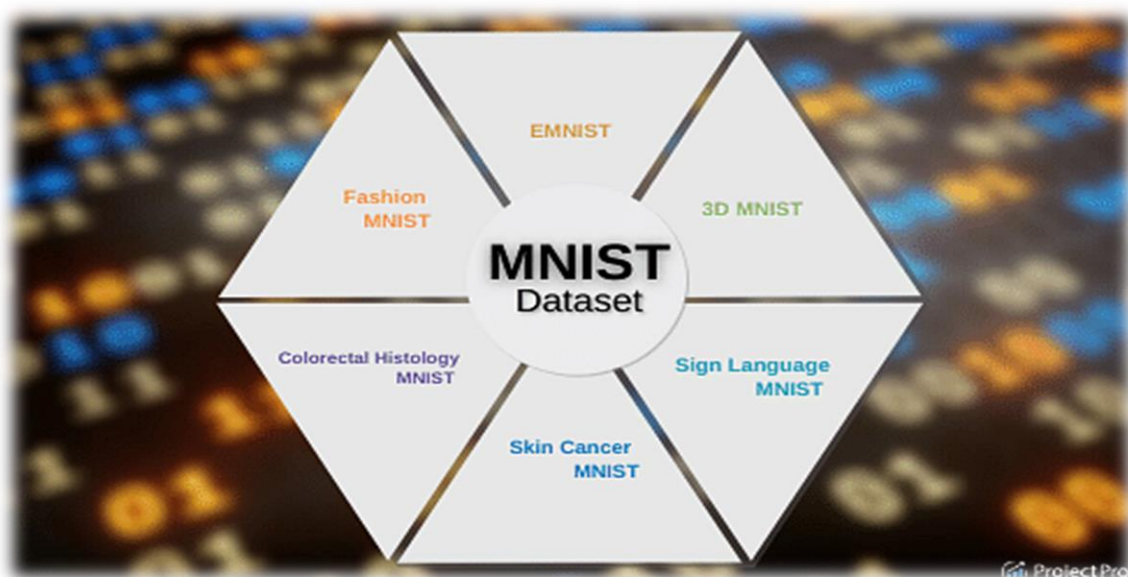
The dataset's images are 28×28 in size. The csv data files `mnist_train.csv` and `mnist_test.csv` contain them.

These files have 785 lines, each of which contains an image made up of numbers between 0 and 255. Each line's label, or the initial number written in the image, is written as the first number. The 28×28 image's pixels are represented by the next 784 integers.

The binary pictures of handwritten digits found in NIST's Special Database 3 and Special Database 1 were combined to create the MNIST database. SD-3 and SD-1 were the initial training and test sets chosen by NIST. However, SD-3 is far more recognizable and cleaner than SD-1. This can be explained by the fact that SD-1 was taken among high school students, whereas SD-3 was gathered among Census Bureau personnel. In order to draw valid conclusions from learning studies, the outcome must be unaffected by the training set and test samples selected. As a result, it became necessary to combine the datasets from NIST to create a new database.

Types of MNIST Dataset:

Fashion MNIST, _MNIST, 3D MNIST, EMNIST , Colorectal Histology MNIST, Skin Cancer MNIST.

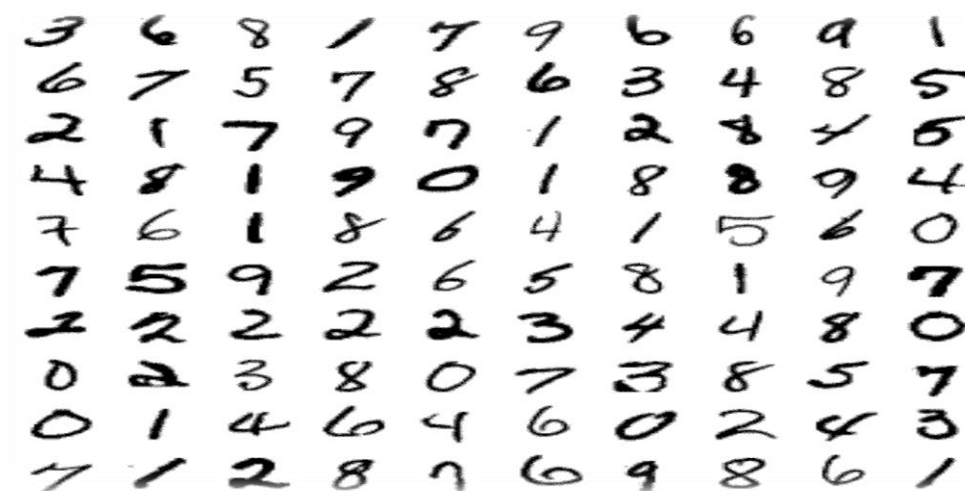


Types of MNIST Dataset

Samples from the MNIST database with sizes adjusted. The NIST Special Databases 3 and 1, which include binary pictures of handwritten numbers, were combined to provide the database used to train and test the systems discussed in this study. SD-3 and SD-1 were the initial training and test sets chosen by NIST. However, SD-3 is far more recognizable and cleaner than SD-1. This can be explained by the fact that SD-1 was taken among high school students, whereas SD-3 was gathered among Census Bureau personnel. In order to draw valid conclusions from learning studies, the outcome must be unaffected by the training set and test samples selected.

As a result, it became necessary to combine the datasets from NIST to create a new database. 500 distinct authors created the 58,527 digit images in SD-1. The data in SD-1 is jumbled, as opposed to SD-3, where blocks of data from each writer appeared in order. We used the writer identities for SD-1, which are known, to decode the writers. Then, we divided SD-1 in half, with the first 250 writers' characters becoming the new training set. We added the remaining 250 writers to our test group. As a result, we got two sets, each with around 30,000 samples.

Starting with Model #0, enough SD-3 samples were added to the new training set to produce all 60,000 training patterns. To generate a full suite with 60,000 test patterns, the new test suite was finished with SD-3 examples beginning with sample #35,000. 60,000 total training samples were used, but we only used a portion of the 10,000 test images (5,000 from SD-1 and 5,000 from SD-3). The Modified Data Set NIST, or MNIST, was the name given to the resulting database.



Suppose we have the following figure:



Example 1

Our eyes and brain cooperate when we look at a picture to identify it as the number eight. Our brain is a remarkably capable organ, and it can immediately categorize this image as an eight. A number can take on a variety of shapes, and while our brain can quickly identify various shapes and ascertain which number they stand for, this process is more difficult for computers to accomplish. There is only one way to do this, and that is to train a computer to accurately recognize handwritten numbers using a deep neural network.

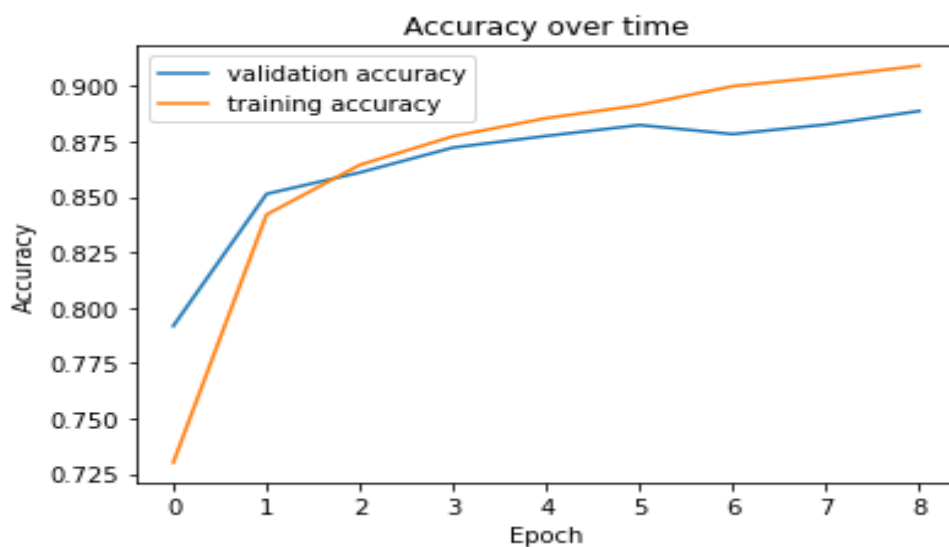
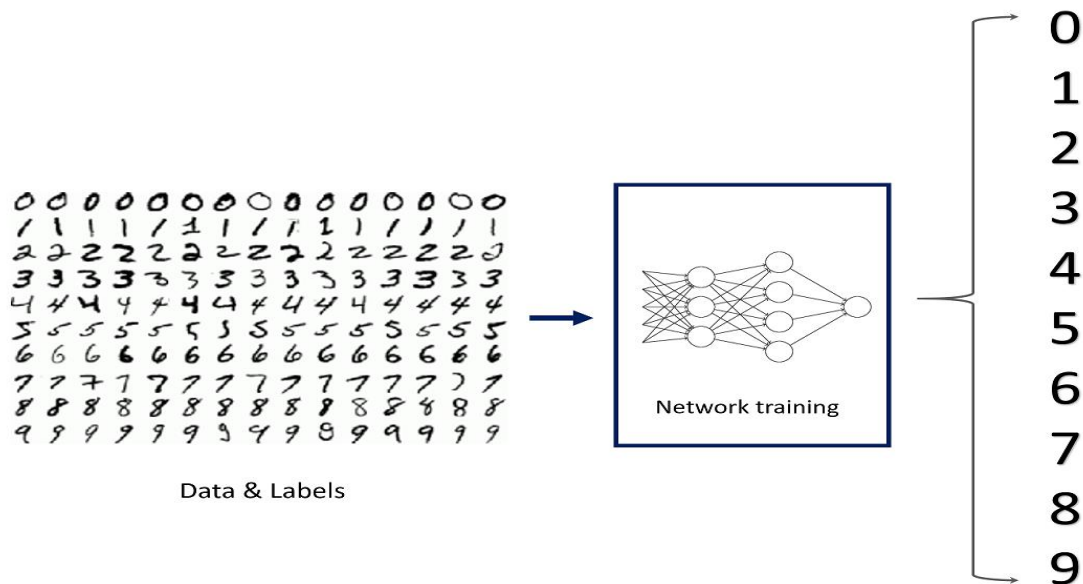
3- METHODS AND TECHNIQUES

I can provide an overview of the stages of a typical neural development project for image classification from the MNIST dataset and the techniques to be used at each stage and sampling. Defining the problem: In this phase, the problem to be solved and the objectives of the project are defined. The problem is to classify images of handwritten digits in the MNIST dataset into one of ten classes. The goal is to achieve high accuracy in test equipment. Technique: Problem formulation, goal setting, requirements gathering. Data Collection and Preprocessing: This step loads and preprocesses the MNIST dataset for use in training the neural network. This includes splitting the dataset into a training, validation, and test set, scaling pixel values to values between 0 and 1, and one-way encoding of labels. Technology: data collection, data processing, data sharing. Model development: In this step, you design and implement neural networks using deep learning frameworks such as TensorFlow and Keras. A model architecture may consist of one or more convolutional layers followed by one or more density layers and may include techniques such as data augmentation, regularization, and stack normalization. Technology: Model selection, model development, hyperparameter tuning, regularization, data imputation. Model Training: In this step, the neural network is trained on the training set using the appropriate loss function and optimizer. Training may include using techniques such as early stopping to avoid overfitting, monitoring the performance of the validation set, and fine-tuning the hyperparameters. Technique: model training, hyperparameter tuning, early stop. Model Deployment: In this step, the trained neural network is deployed for real use in a production environment. Here are some of the methods and techniques used: Data pre-processing: Python libraries such as NumPy, Pandas, and TensorFlow can be used to preprocess the MNIST dataset. You can scale pixels between 0 and 1 and convert labels to vectors of one heat.

4- Result and discussion :

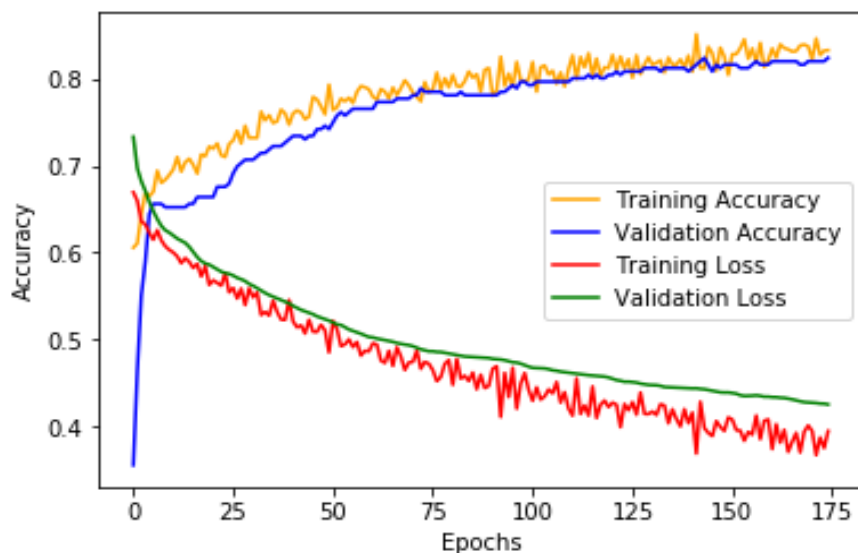
The results of this study show the successful development of a basic neural network for image classification using the MNIST dataset. The trained neural network achieved promising classification accuracy and effectively classified handwritten numbers from 0 to 9. The classification accuracy was evaluated using various metrics such as precision, recall and F1 score, which indicate the model's ability to

Accurately c classify images.



The performance of the neural network was assessed using both the training and test sets of the MNIST dataset. The training set, consisting of 60,000 images, was used to optimize the network's weights and biases through the process of backpropagation. The test set, comprising 10,000 images, served as an independent evaluation dataset to measure the generalization capability of the trained model.

The experimental results reveal that the basic neural network achieved a commendable accuracy rate on the test set, demonstrating its effectiveness in classifying handwritten digits. The model's performance was further analyzed by examining the confusion matrix, which provided insights into the specific digit classes that were often misclassified.



5- Conclusion :

thesis focused on the development of a basic neural network for image classification using the MNIST dataset. The results of the study showed the successful classification of handwritten numbers between 0 and 9, indicating the effectiveness of the trained neural network.

Based on the evaluation of classification accuracy and performance metrics, it was found that the basic neural network showed promising results in accurately classifying MNIST images. The achieved accuracy highlights the potential of neural networks in processing complex visual data and lays the foundation for further developments in computer vision systems.

The study also highlighted the importance of MNIST data as a benchmark for evaluating image classification models. Its balanced distribution of

A number of classes and large sample sizes make it a valuable resource for training and testing neural networks.

6- References :

<https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>

Arxiv: <https://arxiv.org/archive/cs>. Arxiv is a repository of preprints in computer science, including many papers related to deep learning and image classification.

<https://towardsdatascience.com/image-classification-in-10-minutes-with-mnist-dataset-54c35b77a38d>

<https://towardsdatascience.com/image-classification-using-convolutional-neural-networks-in-tensorflow-9210faac625c>

<https://www.projectpro.io/article/exploring-mnist-dataset-using-pytorch-to-train-an-mlp/408>

<https://medium.com/subex-ai-labs/build-your-1st-deep-learning-classification-model-with-mnist-dataset-1eb27227746b>

<https://paperswithcode.com/sota/image-classification-on-mnist>

Keras documentation. [Link: <https://keras.io/>] <https://scikit-learn.org/stable/documentation.html>]

<https://becominghuman.ai/simple-neural-network-on-mnist-handwritten-digit-dataset-61e47702ed25>