**REPORT FILE**

**TITLE:**

Hand-written Digit Recognition

**ABSTRACT**

The reliance of humans over machines has never been so high such that from object classification in photographs to adding sound to silent movies everything can be performed with the help of deep learning and machine learning algorithms. Likewise, Handwritten text recognition is one of the significant areas of research and development with a streaming number of possibilities that could be attained. Handwriting recognition (HWR), also known as Handwritten Text Recognition (HTR), is the ability of a computer to receive and interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices . Apparently, in this paper, we have performed handwritten digit recognition with the help of MNIST datasets using Support Vector Machines (SVM), Multi-Layer Perceptron (MLP) and Convolution Neural Network (CNN) models. Our main objective is to compare the accuracy of the models stated above along with their execution time to get the best possible model for digit recognition.

**Keywords:** Deep Learning, Machine Learning, Handwritten

Digit Recognition, MNIST datasets, Multi-Layered Perceptron (MLP), and Convolution Neural

Network (CNN).

TABLE OF CONTENTS

1. INTRODUCTION....................................................................................5

2. METHODOLOGY…………………………………………………………….6

2.1 Data set

2.2 Convolutional neural network

2.3 Visualization

3. IMPEMENTATION……………………………………………...…………...8

3.1 Pre-processing

3.2 Convolutional Neural Network

4. CONCLUSION ……………………………………………………………10

5. RESULT …………………………………………………………………...11

6. FUTURE ENHANCEMENT ……………………………………………... 12

## **INTRODUCTION**

Handwritten digit recognition is the ability of a computer to recognize the human handwritten digits from different sources like images, papers, touch screens, etc, and classify them into 10 predefined classes (0-9). This has been a topic of boundless-research in the field of deep learning. Digit recognition has many applications like number plate recognition, postal mail sorting, bank check processing, etc . In Handwritten digit recognition, we face many challenges because of different styles of writing of different peoples as it is not an Optical character recognition. This research provides a comprehensive comparison between different machine learning and deep learning algorithms for the purpose of handwritten digit recognition. For this, we have used Support Vector Machine, Multilayer Perceptron, and Convolutional

Neural Network. The comparison between these algorithms is carried out on the basis of their accuracy, errors, and testing-training time corroborated by plots and charts that have been constructed using matplotlib for visualization.

The accuracy of any model is paramount as more accurate models make better decisions. The models with low accuracy are not suitable for real-world applications.

Ex- For an automated bank cheque processing system where the system recognizes the amount and date on the check, high accuracy is very critical. If the system incorrectly recognizes a digit, it can lead to major damage which is not desirable. That’s why

an algorithm with high accuracy is required in these real world applications. Hence, we are providing a comparison of different algorithms based on their accuracy so that the most accurate algorithm with the least chances of errors can

be employed in various applications of handwritten digit recognition.

This project report provides a reasonable understanding of machine learning and deep learning algorithm CNN for handwritten digit recognition. It furthermore gives the information about which algorithm is efficient in performing the task of digit recognition. In further sections of this paper, we will be discussing the methodology and implementation of the algorithm for the fairer understanding of it. Next, it presents the conclusion and result. Moreover, it will also give you some potential future enhancements that can be done in this field.

## **METHODOLOGY**

The comparison of the algorithms (Support vector

machines, Multi-layered perceptron and Convolutional neural

network) is based on the characteristic chart of each algorithm

on common grounds like dataset, the number of epochs,

complexity of the algorithm, accuracy of each algorithm,

specification of the device (Ubuntu 20.04 LTS, i5 7th gen

processor) used to execute the program and runtime of the

algorithm, under ideal condition.

## DATASET

Handwritten character recognition is an expansive research

area that already contains detailed ways of implementation

which include major learning datasets, popular algorithms,

features scaling and feature extraction methods. MNIST

dataset (Modified National Institute of Standards and

Technology database) is the subset of the NIST dataset

which is a combination of two of NIST’s databases: Special

Database 1 and Special Database 3. Special Database 1 and

Special Database 3 consist of digits written by high school

students and employees of the United States Census Bureau,

respectively. MNIST contains a total of 70,000 handwritten

digit images (60,000 - training set and 10,000 - test set) in

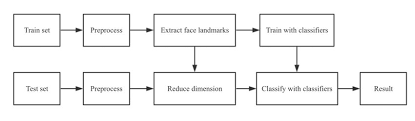
28x28 pixel bounding box and anti-aliased. All these images

have corresponding Y values which apprises what the digit

is.

## CONVOLUTIONAL NEURAL NETWORK

CNN is a deep learning algorithm that is widely used for image recognition and classification. It is a class of deep neural networks that require minimum pre-processing. It inputs the image in the form of small chunks rather than inputting a single pixel at a time, so the network can detect uncertain patterns (edges) in the image more efficiently. CNN contains 3 layers namely, an input layer, an output layer, and multiple hidden layers which include Convolutional layers, Pooling layers (Max and Average pooling), Fully connected layers (FC), and normalization layers [12]. CNN uses a filter (kernel) which is an array of weights to extract features from the input image. CNN employs different activation functions at each layer to add some non-linearity [13]. As we move into the CNN, we observe the height and width decrease while the number of channels increases. Finally, the generated column matrix is used to predict the output .



## VISUALIZATION

In this project, we have used the MNIST dataset (i.e. handwritten digit dataset) to compare different level algorithm of deep and machine learning (i.e. CNN) on the basis of execution time, complexity, accuracy rate, number of epochs and number of hidden layers (in the case of deep learning algorithms). To visualize the information obtained by the detailed analysis of algorithms we have used bar graphs and tabular format charts using module matplotlib, which gives us the most precise visuals of the step by step advances of the algorithms in recognizing the digit. The graphs are given at each vital part of the programs to give visuals of each part to bolster the outcome.

## **IMPLEMENTATION**

To compare the algorithms based on working accuracy, execution time, complexity, and the number of epochs (in deep learning algorithms) we have used two different classifiers:

ANN - Multilayer Perceptron Classifier

Convolutional Neural Network Classifier

We have discussed in detail about the implementation of each algorithm explicitly below to create a flow of this analysis to create a fluent and accurate comparison.

## PRE-PROCESSING

Pre-processing is an initial step in the machine and deep learning which focuses on improving the input data by reducing unwanted impurities and redundancy. To simplify and break down the input data we reshaped all the images present in the dataset in 2-dimensional images i.e (28,28,1).Each pixel value of the images lies between 0 to 255 so, we Normalized these pixel values by converting the dataset into ’float32’ and then

dividing by 255.0 so that the input features will range between 0.0 to 1.0. Next, we performed one-hot encoding to convert the y values into zeros and ones, making each number categorical, for example, an output value 4 will be converted into an array of zero and one i.e [0,0,0,0,1,0,0,0,0,0].

## CONVOLUTIONAL NEURAL NETWORK

The implementation of handwritten digit recognition by Convolutional

Neural Network [15] is done using Keras. It is an open-source neural

network library that is used to design and implement deep learning models.

From Keras, we have used a Sequential class which allowed us to create model layer-by-layer. The dimension of the input image is set to 28(Height), 28(Width), 1(Number of channels). Next, we created the model whose first layer is a Conv layer [20]. This layer uses a matrix to convolve around the input data across its height and width and extract features from it. This matrix is called a Filter or Kernel. The values in the filter matrix are weights. We have used 32 filters each of the dimensions (3,3) with a stride of 1. Stride determines the number of pixels shifts. Convolution of filter over the input data gives us activation maps whose dimension is given by the formula: ((N + 2P - F)/S) + 1 where N= dimension of input image, P= padding, F= filter dimension and S=stride. In this layer, Depth (number of channels) of the output image is equal to the number of filters used. To increase the non-linearity, we have used an activation function that is Relu. Next, another convolutional layer is used in which we have applied 64 filters of the same dimensions (3,3) with a stride of 1 and the Relu function. Next, to these layers, the pooling layer [22] is used which reduces the dimensionality of the image and computation in the network. We have employed MAX-pooling which keeps only the maximum value from a pool. The depth of the network remains unchanged in this layer. We have kept the pool-size (2,2) with a stride of 2, so every 4 pixels will become a single pixel. To avoid overfitting in the model, Dropout layer is used which drops some neurons which are chosen randomly so that the model can be simplified. We have set the probability of a node getting dropped out to 0.25 or 25%. Following it, Flatten Layer is used which involves flattening i.e. generating a column matrix (vector) from the 2-dimensional matrix. This column vector will be fed into the fully connected layer [24]. This layer consists of 128 neurons with a dropout probability of 0.5 or 50%. After applying the Relu activation function, the output is fed into the last layer of the model that is the output layer. This layer has 10 neurons that represent classes (numbers from 0 to 9) and the SoftMax function [25] is employed to perform the classification. This function returns probability distribution over all the 10 classes. The class with the maximum probability is the output. The explosive growth of the Service based websites over the last two decades, has captured the attention of many researchers and academics from various scientific fields. Recognizing the diversity of the nature of this website, which complies with the diverse nature of the Internet, this report tries to examine the topic from a variety of viewing angles. Once commercial exploitation of these technologies becomes widespread, a host of social, cultural and political issues arise.” Although the main goal of the project is to create a service based for house hold employment, the background report’s goal is to examine the whole idea of different small labours works in depth and look at different issues that arise. It does not focus on examining the technical and development side of building a website. Furthermore it does not focus on literature work that deals solely with employment. The reason that this approach was followed, is that the general rules and the theoretical background that define the framework upon which, a service website is built, are the same, despite the fact that the idea behind each service website may differ. The report tries to analyse the methods and the strategies that lead to a successful project and the reasons behind this success. Examining the topic of India’s unemployment as a whole, could lead to more secure conclusions, regarding the characteristics of a successful and prosperous Nation.



## **RESULT**

After implementing all the three algorithms that are SVM, MLP and CNN we have compared their accuracies and execution time with the help of experimental graphs for perspicuous understanding. We have taken into account the Training and Testing Accuracy of all the models stated above. After executing all the models, we

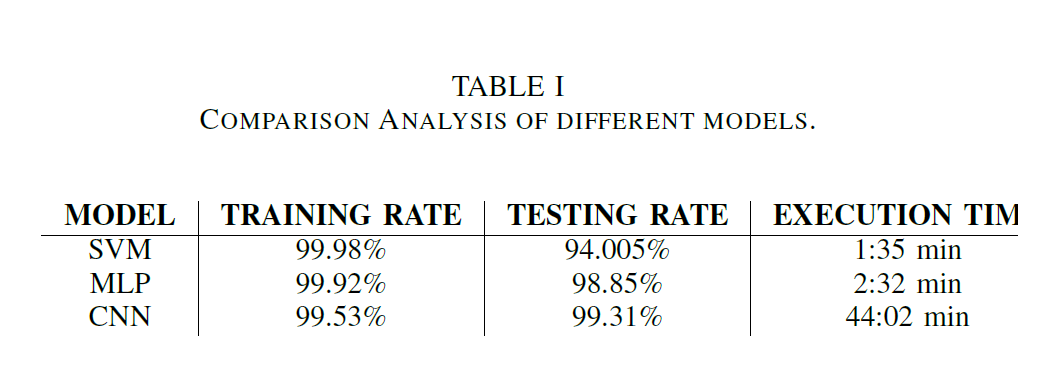
found that SVM has the highest accuracy on training data while on testing dataset CNN accomplishes the utmost accuracy. Additionally, we have compared the execution time to gain more insight into the working of the algorithms. Generally, the running time of an algorithm depends on the number of operations it has performed. So, we have trained our deep learning model up to 30 epochs and SVM

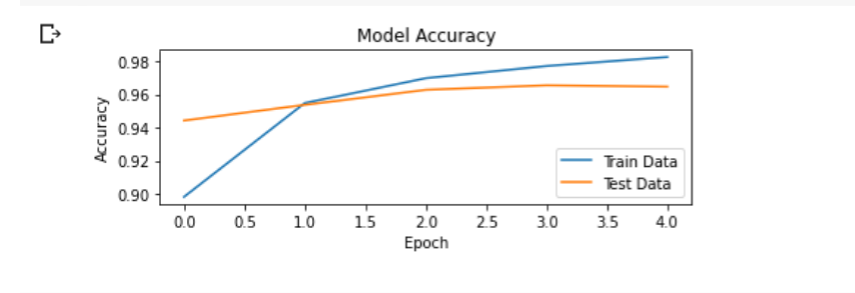
models according to norms to get the apt outcome. SVM took the minimum time for execution while CNN accounts for the maximum running time. This table represents the overall performance for each model.

The table contains 5 columns, the 2nd column represents model name, 3rd

and 4th column represents the training and testing accuracy of models, and

5th column represents execution time of models.





## **CONCLUSION**

In this research, we have implemented three models for handwritten digit recognition using MNIST datasets, based on deep and machine learning algorithms. We compared them based on their characteristics to appraise the most accurate model among them. Support vector machines are one of the basic classifiers that’s

why it’s faster than most algorithms and in this case, gives the maximum training accuracy rate but due to its simplicity, it’s not possible to classify complex and ambiguous images as accurately as achieved with MLP and CNN algorithms. We have found that CNN gave the most accurate results for handwritten digit recognition.

So, this makes us conclude that CNN is best suitable for any type of prediction problem including image data as an input. Next, by comparing execution time of the algorithms we have concluded that increasing the number of epochs without changing the configuration of the algorithm is useless because of the limitation of a certain

model and we have noticed that after a certain number of epochs the model starts overfitting the dataset and give us the biased prediction.

## **FUTURE ENHANCEMENT**

The future development of the applications based on algorithms of deep and machine learning is practically boundless. In the future, we can work on a denser or hybrid algorithm than the current set of algorithms with more manifold data to achieve the solutions to many problems. In future, the application of these algorithms lies from the public to high-level authorities, as from the differentiation of the algorithms above and with future development we can attain high-level functioning applications which can be used in the classified or government agencies as well as for the common people, we can use these algorithms in hospitals application for detailed medical diagnosis, treatment and monitoring the patients, we can use it

in surveillances system to keep tracks of the suspicious activity under the system, in fingerprint and retinal scanners, database filtering applications, Equipment checking for national forces and many more problems of both major and minor category. The

advancement in this field can help us create an environment of safety, awareness and comfort by using these algorithms in day to day application and high-level application (i.e. Corporate level or Government level). Application-based on artificial intelligence and deep learning is the future of the technological world because

of their absolute accuracy and advantages over many major problems.