ZKG INTERNATIONAL ISSN: 2366-1313

Sorcery of Gaussian Naïve Bayes: Identifying Handwritten Digits with Machine Learning

Dr.Pon Partheeban ponparthee@stellamaryscoe.edu.in

Mr.C.Bastin Rogers bastinrogers@stellamaryscoe.edu.in

Dr. F. R. Shiny Malar SHINYMALAR@stellamaryscoe.edu.in

Mrs.R.Shiny shiny@stellamaryscoe.edu.in

Mr.R.Dinesh dinesh@stellamaryscoe.edu.in

Department Of Computer Science Engineering
Stella Mary's College Of Engineering, Tamilnadu, India

Abstract- Nowadays, the scope of machine learning and deep learning studies is increasing day by day. Handwriting recognition is one of the examples in daily life of this field of work. Data storage in digital media is a method that almost everyone is using nowadays. At the same time, it has become a necessity for people to store their notes in digital media and even take notes directly in the digital environment. As a solution to this need, applications have been developed that can recognize numbers, characters, and even text from handwriting using machine learning and deep learning algorithms. Moreover, these applications can recognize numbers, characters, and text from handwriting and convert them into visual characters. This project investigated the performance comparison of machine learning algorithms commonly used in handwriting recognition applications and which of them is more efficient. As a result of the study, the accuracy was 98.66% with an artificial neural network, 99.45% with a convolutional neural network, 97.05% with K-NN, 83.57% with Naive Bayes, 97.71% with a support vector machine, and 88.34% with a decision tree. This study also developed a handwriting recognition system for numbers similar to those mentioned in the applications. A desktop application interface was developed for end users to show the instant performance of some of these algorithms and allow them to experience the handwriting recognition system.

Keywords: Machine Learning, Handwriting Recognition, Deep Learning, MNIST

1. INTRODUCTION

Today, the applications of artificial intelligence and data science are advancing very rapidly and are now used in almost all fields. Most of these applications include machine learning and deep learning. One of the application areas is handwriting

ISSN: 2366-1313

recognition, which is an evolving field. Handwriting recognition enables communication between machines and humans, and it is a field that aims to facilitate this communication. A lot of information is now stored in digital media. In daily life, almost everyone has started storing their data and notes in digital media and using electronic diaries. They tend to take notes in the digital environment using the keyboard, touch screen, and smart pens of smartphones and tablets that they always have with them. Handwriting recognition systems are needed more and more every day for the following reasons: For data to be stored in digital media with recognition once handwritten, or for data previously written to be transferred to digital media as optical characters. With handwriting recognition systems beginning to evolve out of this need, today's technology has moved away from keyboards to touch systems, and writing in the digital environment has begun to be done with handwriting instead of the small keyboards of touch screens. In addition, handwritten text documents, lettering on signs, etc. have begun to be transferred to digital in sectoral areas. Handwriting recognition systems are constantly evolving and being integrated into intelligent systems. Many *corresponding Author electronic tablets and digital agendas are now offered to users with handwriting recognition system software and smart pens. At the same time, in addition to everyday use, these systems are also being used in areas such as education, health, and security, which are evolving every day. In banks, handwriting recognition systems are used in areas such as reading check amounts and forms and reading and sorting incoming mail addresses [1]. In addition, studies are being conducted daily for the systems to be used by children on smart boards and tablets, mainly to be used in education and to support teaching. Handwriting recognition is the computer recognition of handwritten letters, numbers, and characters. This process, which is very simple for a human, is difficult for computers. In other words, making sense of lines, symbols, and their combined shapes at the word level is difficult for computers. Handwriting features such as the presence of characters that are different in many languages, the fact that each person has different handwriting, and the presence of combined handwriting make it difficult for computer systems to recognize handwriting. This topic is not yet fully developed, and it is an area of limited efficiency [2]. When this technology is developed, which is mainly used in tablet computers and for which there are already examples, it will be possible to store and organize any handwritten information in a digital environment without using a keyboard.

Handwriting recognition technologies can be studied in two different groups. One is interactive (online), and the other is non-interactive (offline). Interactive systems are

ISSN: 2366-1313

systems that recognize and classify by following the movements of the writing tool as it writes the handwriting. They are usually used in devices with a touch surface, such as tablets, phones, smartboards, etc. Every movement during writing is controlled by the device. This is an important factor in increasing the accuracy rate. The disadvantage of these systems is that they have to give instant results, and they have to be fast-running systems as they have to keep up with the writing speed [2]. Noninteractive systems are an attempt to recognize information previously written on paper by digitizing it using methods such as photography and scanning. These systems are not expected to produce instantaneous output, so there are ways to work more slowly. However, since handwriting may not be very smooth, especially in old handwritten documents, the accuracy rates in the recognition process are low. To increase this rate, extensive pre-processing of the photo is required first. However, as a first step, the text should be divided into sections. Lines, sentences, words, and characters should be divided according to their size, and then preprocessing steps should be applied to each divided part before classification [1]. The advantage of this method is that it can be used for all kinds of documents that have existed for years and should be transferred to electronic media.

2. RELATED STUDIES

In the work of Mohd Razif Shamsuddin and his colleagues who analyzed machine learning models for the MNIST dataset, they obtained two different versions of the MNIST dataset, grayscale and binary (black and white), with data preprocessing. They also performed the model training with these two different datasets. In the grayscale dataset, the accuracy rate was 99.4% with the CNN model, 94% with the random forest model, and 94% with the Extremely Randomized Trees model. In the binary data set, the accuracy rate was 90.1% with the CNN model, 91% with the Random Forest model, and 92% with the Extremely Randomized Trees model. As a result of these experiments, they showed the effect and importance of proper selection of data preprocessing steps and methods in machine learning on success rate [3]. Abien Fred M. Agarap created an architecture for machine learning by combining convolutional neural networks and support vector machines for image classification problems. In addition to the model where he used the ReLu function as the activation function in the CNN structure and then added a support vector machine, he created another model where the same CNN structure determined the activation function of the output layer as a softmax function. 2 models named "CNNSoftmax" and "CNN-SVM" were trained with MNIST and

Fashion-MNIST datasets and compared the results. In the study, 60000 training data and 10000 test data from both datasets were used. With the structure of "CNN-Softmax," it achieved a 99.23% success rate in the MNIST dataset and a 91.86% success rate in the FashionMNIST dataset. With the structure of "CNNSVM," it achieved a 99.04% success rate in the MNIST dataset and a 90.72% success rate in the FashionMNIST dataset [4].

Mine Altınay Günler Pirim argued that the output weights in the hidden layers in the structure of the neural network used in handwriting recognition can be used to extract the feature vectors of the image. For this purpose, it used MNIST and USPS datasets in its experiments with MATLAB; using the artificial neural network, the support vector machine, and the Euclidean distance classifier algorithm as the classifier, the success rates were measured. Using the structure developed in the study, it achieved success rates of 99.64% in the support vector machine model with the MNIST dataset, 98.01% in the Euclidean distance classifier model, and 98.56% in the artificial neural network model. With the USPS dataset, it achieved 97.47% success rates in the support vector machine model, 94.52% in the Euclidean distance classifier model, and 94.18% in the artificial neural network model [5]. Aoudou Salouhou tested and compared deep learning algorithms for image classification and handwriting recognition using Fashion-MNIST, MNIST, CIFAR-10, and Arabic datasets in his study. To classify Arabic characters, the Arabic dataset created by Ahmed and his friends was used [6]. He used deep neural networks, convolutional neural networks, and recursive neural network structures from deep learning algorithms. In his experiments with the MNIST dataset, the deep neural network model provided a 99.53% accuracy rate, the convolutional neural network model provided a 99.88% accuracy rate, and the iterative neural network model provided a 99.05% accuracy rate. In his experiments with the Arabic dataset, the deep neural network model provided 96.48% accuracy, the convolutional network model provided 99.00% accuracy, and the recursive neural network model provided 96.94% accuracy [7].In her studies on handwriting recognition with machine learning algorithms, Rabia Karakaya conducted tests with support vector machines, decision trees, random forests, artificial neural networks, the K-nearest neighbor algorithm, and the K-mean algorithm using the MNIST dataset. She conducted her work using the Scikit-Learn library and tools. In the test results of the models trained with the MNIST dataset using all 60000 data points, she achieved a 90% accuracy rate by using the polynomial kernel function in the support vector machine model. It also achieved 87% accuracy in the decision tree model, 97% in the random forest model,

ISSN: 2366-1313

and 97% in the artificial neural network model. In the K-Nearest Neighbor algorithm model, it achieved 96% accuracy in 865.932 seconds of test time and 98% accuracy in the K-Mean algorithm model [8]. In their work, Shubham Sanjay Mor and his colleagues developed a system and an Android application to recognize handwritten characters and numbers using the EMNIST dataset and the CNN structure. At the end of the experiments they conducted in the CNN model they created, they used the "Adamax" function as an optimization parameter, with which they achieved the highest performance. The model they created recognizes 62 handwritten characters and has an accuracy rate of 87.1% [9].

3. MATERIAL AND METHOD

MNIST Dataset: The dataset used for training and testing this system is the MNIST (Modified National Institute of Standards and Technology) dataset. It is a dataset created to process and make sense of the image. It consists of images of handwritten numbers, each of which is 28x28 pixels in size. It contains 60,000 training (train) and 10,000 verification (test) images as classification [10]. It is a commonly preferred dataset in studies in similar domains.

Libraries: The software libraries used in this study are TensorFlow, Scikit Learn, Keras, and Numpy. In the following sections of this study, some mathematical values are used to make comparisons in performance evaluations and analyses. These mathematical values are defined in the mentioned libraries, and their definitions are as follows:

Confusion Matrix: A confusion matrix, or error matrix, is a performance measurement method for machine learning classification algorithms. It gives information about the accuracy of the predictions. It is a table that contains four combinations of predictions and actual values [11]. With the help of this matrix, we can calculate values like precision, recall, support, accuracy, specificity, and F1 score.

Accuracy: It is a value that indicates the accuracy rate. It is the ratio of predictions classified as correct to all predictions.

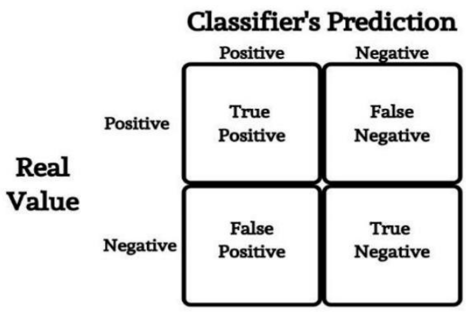

Figure 1. Confusion Matrix

Precision: It is the value that indicates how much of the positively predicted data was predicted correctly. The higher this ratio is, the more accurate the predictions were [12]. It is calculated as;

$$TP/(TP + FP) \quad (1)$$

ISSN: 2366-1313

Recall: It is the ability of the classifier to find all positive samples. It is a metric that shows how much of the data that should be positively predicted is positively predicted [12]. It is calculated as;

$$TP/(TP + FN) \qquad (2)$$

F1 Score: The F1 score can be interpreted as a weighted harmonic average of the Precision and Recall scores. It is a measure of the level of performance exhibited by the model. It is often used to compare models.

Data Processing :Data sets are processed through multiple phases, as with any MA study courses. Data pre processing, cutting, feature extraction and classification are the key procedures employed in the image classification and handwriting recognition applications.

Preprocessing data: Data preparation is performed ready for analysis and subsequent phase; such operations as correction, conversion, cleaning, decrease of size, standardization and noise reduction. In handwriting recognition data the data preparation processes required might be stated in the following way:

Thresholding: This is the way the image is converted into a binary picture, namely a black and white one. The basic objective is to highlight the image and identify the item with this technique. Pixels are calculated as black or White according to the supplied threshold value, depending on the type of threshold approach used [13].

Noise reduction: Refers to processes carried out to further clear the picture. The highlights can be sharpened in texts by a technique that is needed. It aims to produce a clear picture through this approach. Methods such as a medium filter, median filter, Gaussian smoothing filter and masking approach can minimize noise.

Normalization: Slope correction is an additional name for standardization, one of the image processing procedures. This procedure corrects the curvature of the text in handwritten text photographs and eliminates the skew. Histograms are used to identify and correct curvature and path, as with many other image processing procedures. In the early parts of the normalization operations, Bosinosvic and Srihari Method (BSM) is commonly used [14]. The angle between the horizontal axis in the text correction and the horizontal axis of the written text is known as the slope, and the angle between the vertical axis in the text correction and the vertical axis of the written text is known as the slant. Normalization is another term for handwriting correction using slant and slope [15].

ISSN: 2366-1313

Feature extraction: Each distinguishing feature in the photos may be designated as an attribute. Attribute information is composed of numerical data derived from the separation of features from the picture [16]. Approaches like histograms, projection-based methods, Fourier and Wavelet transforms, or defining letters as a set of basic shapes like curves and lines are utilized at this step [2]. The information collected from this stage has a direct impact on the recognition stage, and the qualities of this information have a direct impact on the recognition stage's efficiency.

Classification: During the classification phase, the attributes of the data in the picture are compared to the classes in the database to determine which class the picture belongs to. Many various approaches are utilized at this step, including template matching, neural networks, classification algorithms, statistical learning, and structural learning. The data sets employed are critical for this stage's high performance and accuracy, and they should be prepared to contain as many samples and kinds as feasible.

4. Proposed Model

The implementation steps for the project are as follows:

1) The project contains necessary libraries such as Tensorflow, Keras, Scikit Learn, and Numpy.

2) The project includes the MNIST dataset from the Tensorflow-Keras library.

The implementation steps for the project are as follows:

3) The dataset was labeled and split into training and test data.

4) The data was subjected to preprocessing processes.

5) Following the preparation processes, the data was normalized and characteristics were retrieved. 6) Algorithms to be utilized with libraries have been defined.

7) For each method, the parameters to be utilized during training were determined by the most appropriate method for the dataset. It was intended to get the best possible results.

8) The dataset was trained using the "Fit" function, and model training was completed. The Scikit Learn library was widely utilized during learning in KNN, SVM, Decision Tree, and Naive Bayes algorithms, while Tensorflow and Keras libraries were employed in neural networks.

9) The model was evaluated using the "Predict" function on a portion of the dataset that had been put aside for testing.

10) Finally, to assess the performance of the trained and tested model, the classification report and confusion matrix were generated and shown as a report using the "classification report" and "confusion matrix" functions of the Scikit Learn library's "metrics" module. Precision, recall, f1 score, support, and accuracy scores generated from the confusion matrix findings are shown in tabular form

ISSN: 2366-1313

in the classification report. This table contains enough information to allow you to examine and evaluate the performance results.

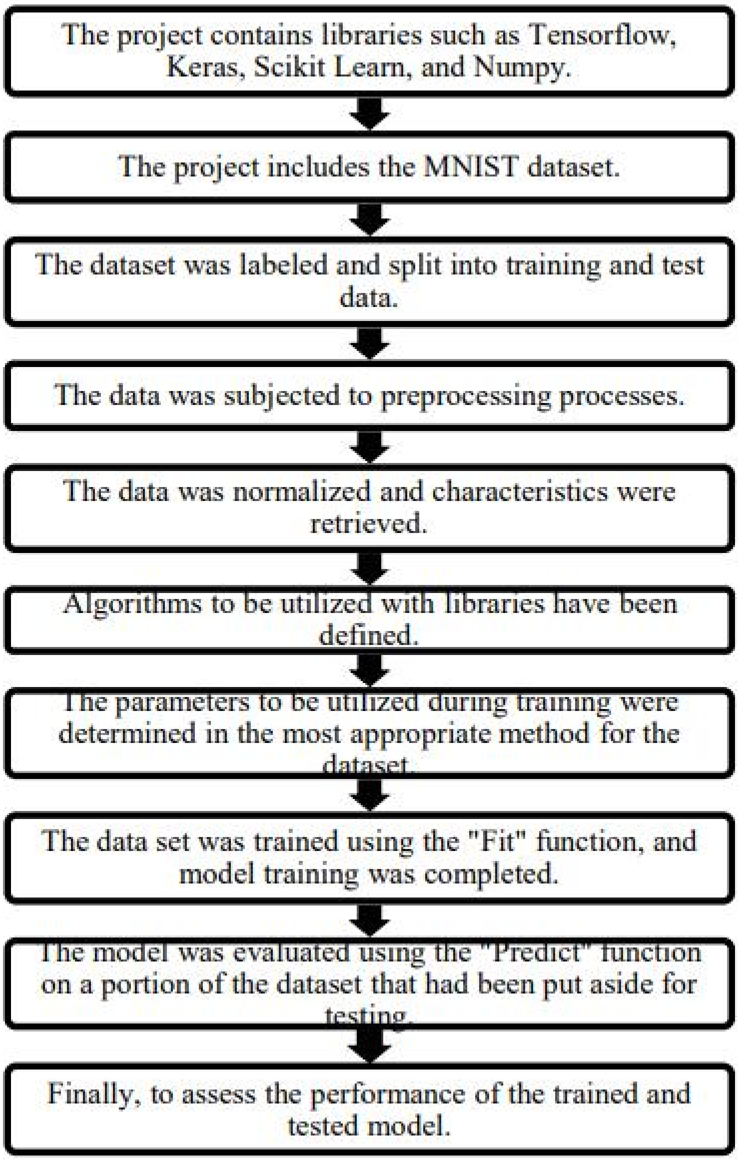

Figure 2. The Implementation Steps

5. RESEARCH AND PERFORMANCE RESULTS

The tests carried out, the models constructed using the methods utilized, the hyper-parameter modifications, and the performance results are all detailed in depth in this section. All 70000 bits of data from the data set were used, with 60000 reserved for training and 10000 designated for testing, and the ratios of these portions were kept consistent throughout all ss

Artificial Neural Network (ANN)

Deep learning is based on neural networks, which were formed by modeling the neuron structure of humans and adapting this neuron structure to machines. The human neuron structure was used to generate features such as brain structure, learning, and information use. To detect real-world relationships, artificial neural networks have been built. They can also conduct classification, pattern recognition, grouping, and estimation. They may process many inputs and generate results [17]. The network structure is formed by the combination of artificial neural network cells within the framework of particular rules. Layers are generated when these brain networks, or neurons, join together. Receiving inputs and sending outputs are handled by cells and layers. As a result, they are linked to the outside world [18]. This structure's layers are divided into three sections: the input layer, the concealed layer, and the output layer. "Multiple-layer perceptron neural network (MLPNN)" refers to artificial neural networks that have one or more hidden layers in addition to the input and output layers. MLPNN structures are artificial neural network structures that are employed and represented in much artificial neural network research nowadays [19]. The study's artificial neural network is a three-layer, fully linked network. In order to build the neural network, the "Sequential" model from the Keras library's "model" module was used. According to the “sequential” paradigm, a neural network will be formed. The "Dense" class was then added to the "layers" module to generate neural network layers. The number of neurons in the layers was calculated using these modules, and the neural network was built. Neurons in the neural network's hidden layers have activation functions [19]. A neuron's activity is determined by its activation functions. As a result, they are extremely important in neural networks. Different functions are selected based on where they will be utilized and how they will affect performance. The most preferred and recommended activation functions in neural network layers are the “relu”, "sigmoid,” and “softmax” functions. In the neural networks in this study, while the "relu" activation function is used in the input layers, the "softmax" function, which is the most preferred in multiple classification problems, is used in the output layers. The “Softmax” function generates values between 0 and 1, and these values show the probability that each input belongs to a class. After creating the neural network structure, it was compiled with the necessary parameters.

Figure 5. Interface implementation classification process result

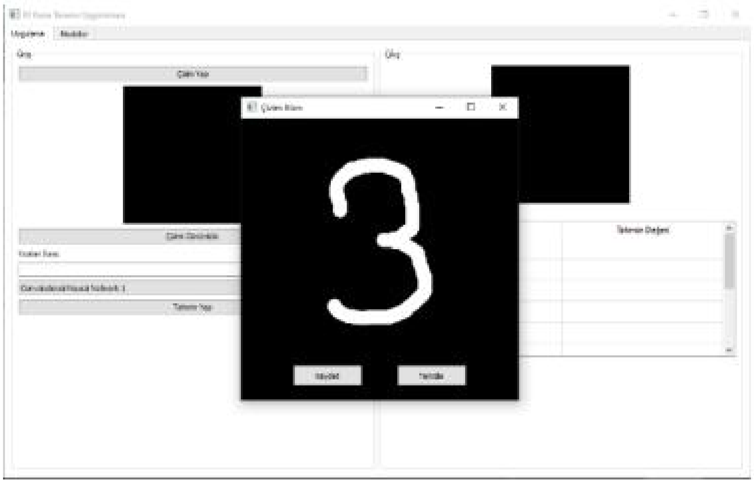

Figure 6. Interface application drawing window

Among the algorithms previously described in this study, details of models trained using convolutional neural networks, multiple-layer perceptron neural networks, naive bayes, support vector machines, K-nearest neighbors, and decision tree algorithms can be accessed in the "Modeler" tab. Then, on the main screen, they select the model they want to use in the classification process. The classification process starts with the “Tahmin Yap (Predict)” button. After the classification process is finished, an “Sınıflandırma tamamlandı (Classification is complete).” information message is received.

6. CONCLUSION AND EVALUATION

In this study, the achievements and performances of machine learning algorithms in handwriting recognition processes were examined. An interface has been developed for test studies by sampling recognition models with different algorithms. These algorithms were selected as ANN, CNN, K-NN, Naive Bayes algorithm, SVM, and decision trees; all of them were discussed in detail, and experiments were carried out on the MNIST dataset for each of them. During these experiments, Python programming language and, accordingly, Keras and Tensorflow libraries for the MNIST dataset, CNN and ANN structures, and the Scikit Learn library were used for test result reports of K-NN, Naive Bayes algorithm, SVM, decision trees, and models. The result reports and accuracy value were obtained using the Scikit Learn library.

In the study, a total of six different algorithms were used. ANN and CNN models were the most successful when the models were compared in terms of accuracy. A success rate of 98.66% was achieved with the ANN-Adam model and 99.45% with the CNN-Adam model.

ISSN: 2366-1313

Experimental results have shown that neural networks are more successful than other algorithms studied. However, neural network architecture and selected activation functions significantly affect performance. In traditional classification algorithms, it is observed that the K-NN and SVM models can achieve a success rate of over 97%. As a result, almost all classification algorithms were examined with the MNIST data set, unlike similar studies in the literature. Experiments were carried out with many different parameters in all of them, and it was aimed to create the most efficient combination.

REFERENCES

[1] I. S. MacKenzie and K. Tanaka-Ishii, Text entry systems: mobility, accessibility, universality. San Francisco, Calif: Morgan Kaufmann, 2007. doi: 10.1016/B978-0-12-373591-1.X5000-1.

[2] P. Duygulu, “El Yazısı Tanıma,” in Bilişim Ansiklopedisi, Papatya Yayıncılık, 2006.

[3] M. R. Shamsuddin, S. Abdul-Rahman, and A. Mohamed, “Exploratory Analysis of MNIST Handwritten Digit for Machine Learning Modelling,” Communications in Computer and Information Science, vol. 937, pp. 134–145, 2019, doi: 10.1007/978-981-13-3441-2_11.

[4] A. F. M. Agarap, “An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification,” arXiv, pp. 5–8, 2019.

[5] M. A. Günler Pirim, “Neural Network Based Feature Extraction for Handwriting Digit Recognition,” Ankara, 2017.

[6] A. El-Sawy, M. Loey, and H. El-Bakry, “Arabic Handwritten Characters Recognition using Convolutional Neural Network,” WSEAS Transactions on Computer Research, vol. 5, pp. 11–19, 2017.

[7]. G. Jagga Rao, Y. Chalapathi Rao, Dr. Anupama Desh Pande "Detection For 6G-NOMA Based Machine Learning Optimization for Successive Adaptive Matching Pursuit Analysis", Q3, pp. 1803-1812, Jan 2020.

[8]. Sudha, Y. Chalapathi Rao, G. Jagga Rao " Machine Learning based Copy-Move Forgery Detection with Forensic Psychology Ultra-Hd images "in Volume 81, Nov-Dec-2019.

[9]. Dr. B Sankara Babu, Srikanth Bethu, K. Saikumar, G. Jagga Rao, "Multispectral Satellite Image Compression Using Random Forest Optimization Techniques" Journal of Xidian University, in Volume 14, Issue 5-2020.

ISSN: 2366-1313

[10]. G. Jagga Rao, Y. Chalapathi Rao, "Human Body Parts Extraction in Images Using JAG-Human Body Detection (JAG-HBD) Algorithm Through MATLAB" Alochana Chakra Journal, Volume IX, Issue V, May/2020.

[11]. Dr. k. Raju, A. Sampath Dakshina Murthy, Dr. B. Chinna Rao,G. Jagga Rao "A Robust and Accurate Video Watermarking System Based On SVD Hybridation For Performance Assessment" International Journal of Engineering Trends and Technology (IJETT) – Volume 68 Issue 7 - July 2020.

[12]. G. Jagga Rao, Y. Chalapathi Rao, Dr. Anupama Desh Pande "A Study of Future Wireless Communication: 6G Technology Era " volume 14, issue 11,2020.

[13]. G. Jagga Rao, Y. Chalapathi Rao, Dr. Anupama Desh Pande "Deep Learning and AI-Based millimeter Wave Beamforming Selection for 6G With Sub-6 GHz Channel Information" Volume 21 : Issue 11 – 2020.

[14] A. Vinciarelli and J. Luettin, "A New Normalization Technique for Cursive Handwritten Words," Pattern Recognition Letters, vol. 22, no. 9, pp. 1043–1050, 2001, doi: 10.1016/S0167-8655(01)00042-3.

[15] B. Yılmaz, "Design of A Mobile Device Application with Handwriting Recognition to Make Learning Easy For Students Who Have Learning Disabilities," Istanbul, 2014.

[16] H. H. Çelik, "Recognition of Handwritten Numerals by Using Neural Network," Istanbul, 1999. [17] O. A. Erdem and E. Uzun, "Turkish Times New Roman, Arial, And Handwriting Characters Recognition by Neural Network," journal of the Faculty of Engineering and Architecture of Gazi University, vol. 20, no. 1, pp. 13–19, 2005.

[18] H. A. Şahin, "Comparison of Artificial Neural Networks and Different Optimization Methods," Samsun, 2020.

[19] E. Öztemel, Yapay Sınır Ağlari. İstanbul: Papatya Yayıncılık, 2012. [Online]. Available: http://papatyabilim.com.tr/PDF/yapay_sinir_aglari.pdf

[20] B. Ma, X. Li, Y. Xia, and Y. Zhang, "Autonomous deep learning: A genetic DCNN designer for image classification," Neurocomputing, vol. 379, pp. 152– 161, 2020, doi: 10.1016/j.neucom.2019.10.007.

[21] E. Yalçın, "Binary-Data Multi-Criteria Recommender Systems Based on Naive Bayes Classifier," Eskişehir, 2016