	PES UNIVERSITY **(Established under Karnataka Act No. 16 of 2013)** **100 Ft. Road, BSK III Stage, Bengaluru – 560 085** **DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING** **SESSION: Jan-May 2021**

Course Title: Stochastic Models and Machine Learning	
Course code: UE20CS504	
Semester : I Sem	**Team Id:14**
SRN:PES1PG20CS021	**Name:Mehvish Hafiz**
SRN:PES1PG20CS023	**Name:Nallani Mounika**

ASSIGNMENT REPORT

1.Problem Statement:-Implement a Naive Bayes Classifier(NBC) , of English news paper Head Lines into Politics, Sports, Education, Healthcare, Finance (5- Class Labels)

Requirement Analysis:

The above problem aims at classifying News paper Head Lines by using naïve bayes classifier,and what the text will we give that can belongs to which classifier like politics,sports,education,healthcare and finance

Steps Involved:

1)Implement Naïve Bayes classifer.

2)Read a dataset which consists of a news paper .

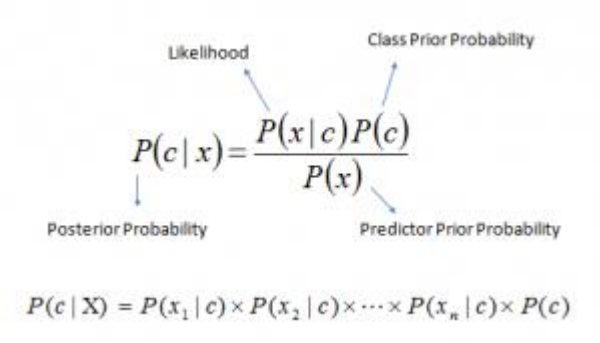
3) Metrics used to evaluate the classification report:

- Precision

- Recall
- F1 Score
- Description

Implementing Naïve Bayes Classifier:

Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other.



Before implementing the Gaussian Naive Bayes classifier we should write 2 simple assumptions:

Our data is normally distributed

We expect our data columns to be conditionally independent of each other

Let's finally implement this model in just 3 simple steps!

1.Preparing the dataset

2.Calculating train set statistics

3. Calculating labels for the validation set

Dataset Details:

1) The dataset training consists of 4 columns and about 10,000 data entries.

2) The columns include a unique ID, Text review, and a label indicating whether the review is positive or negative.

3) Split the data into training and testing datasets

4) Encoding of the tags needs to be done in order to convert the text tags into comparable tokens.

Output Screenshots

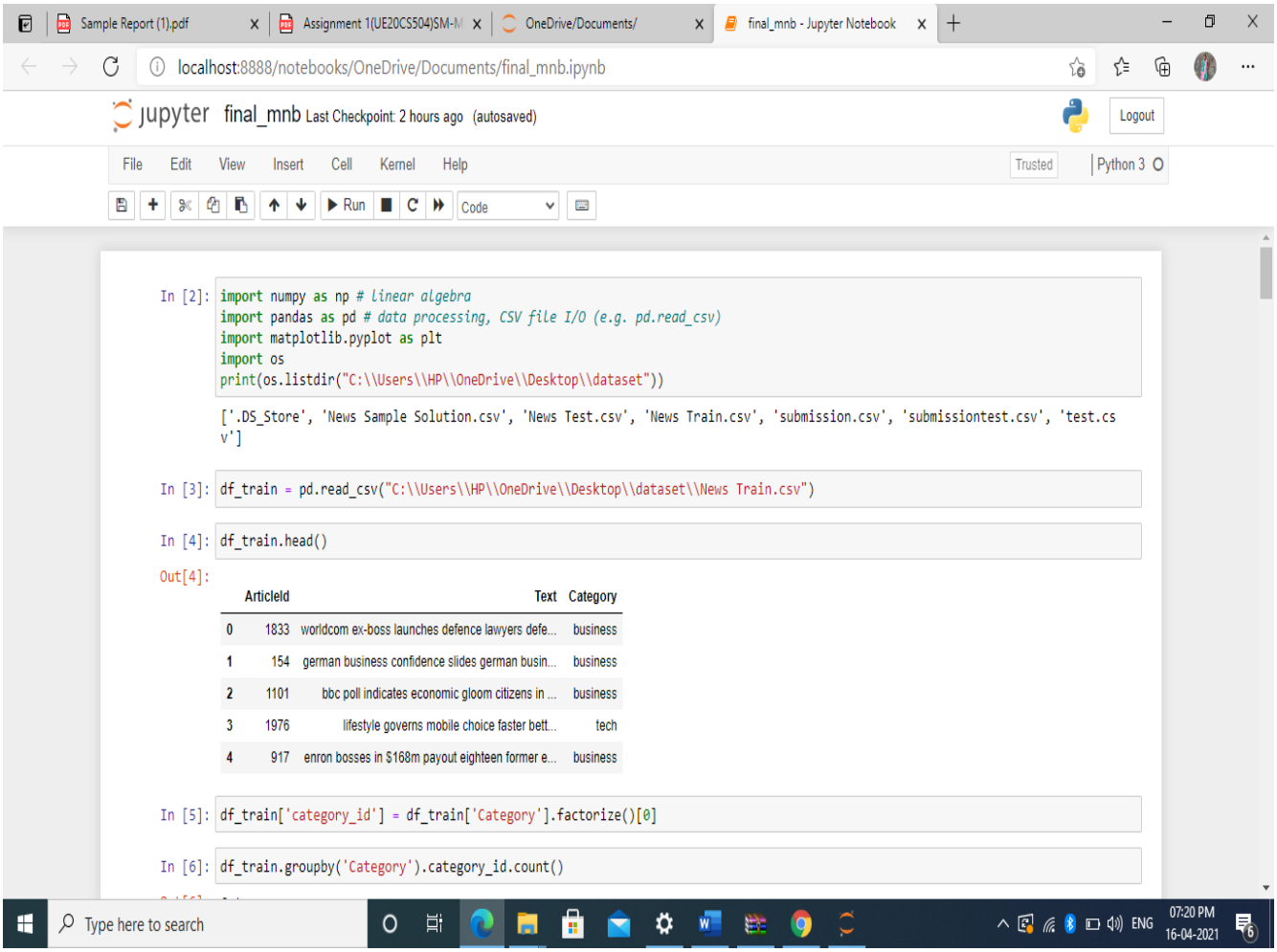
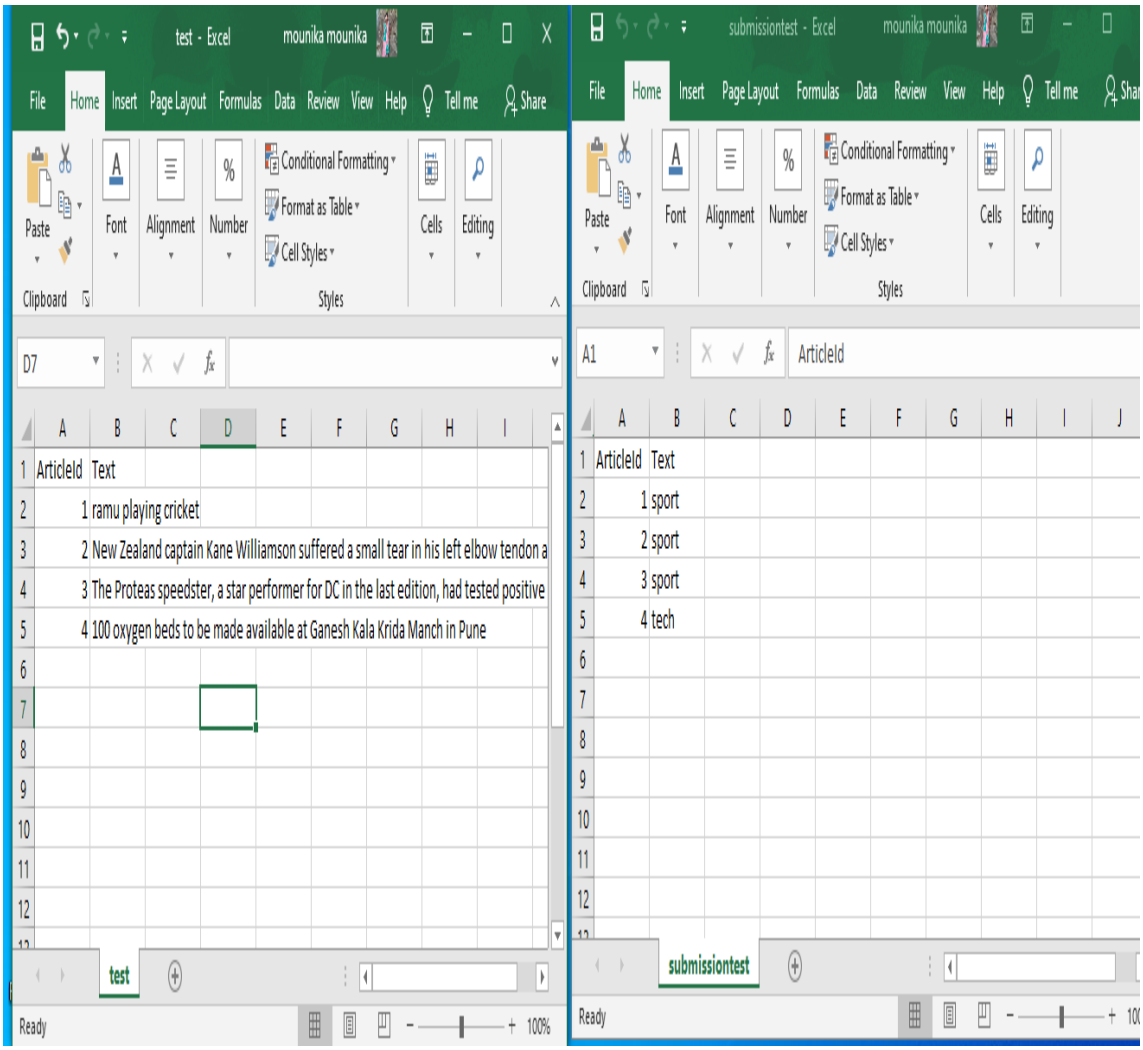


Fig 1: Training Dataset head



Learning Outcome

1)Understand one of the most popular and simple machine learning classification algorithms, the Naive Bayes algorithm

2)Learn how to implement the Naive Bayes Classifier in Python

3)We were able to learn to create and handle large datasets and efficiently use to to complete our assignment.

2.Problem Statement:- Apply One-Versus-Rest SVM and classify

SVM:-

It is a class of Machine Learning algorithms that are used quite frequently these days. Named after their method for learning a decision boundary, SVMs are binary classifiers – meaning that they only work with a 0/1 class scenario. In other words, it is not possible to create a multiclass classification scenario with an SVM natively

Fortunately, there are some methods for allowing SVMs to be used with multiclass classification. In this article, we focus on two similar but slightly different ones: one-vs-rest classification and one-vs-one classification

One-vs-Rest (OvR) Classification

The One-vs-Rest method can be used for creating a multiclass SVM classifier. Let's recall the multiclass assembly line that we discussed above. Here, the output is one out of three possible classes: {yellow, blue, red}.

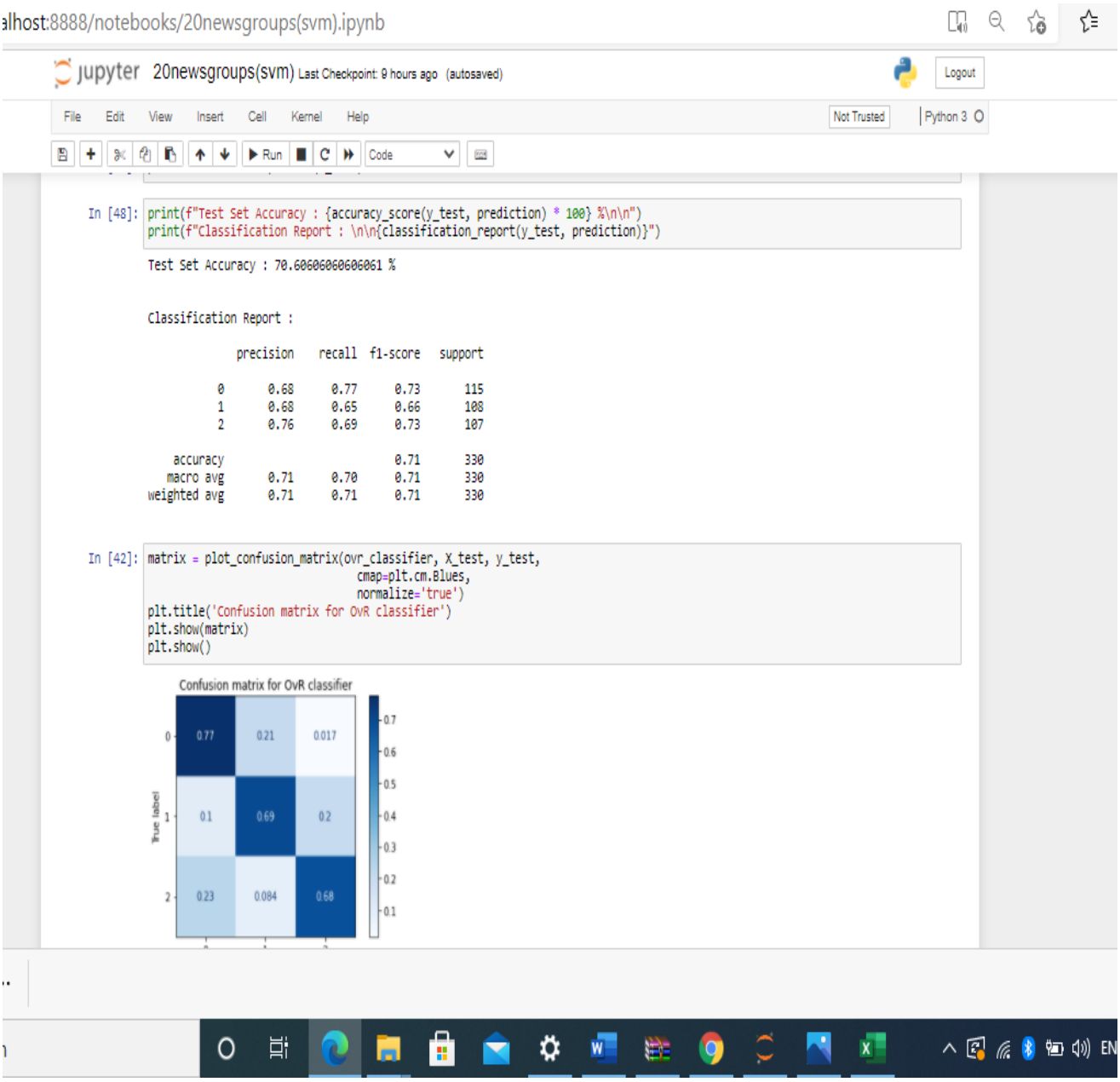
OvR binary classifier 1: yellow vs {blue, red}

OvR binary classifier 2: blue vs {yellow, red}

OvR binary classifier 3: red vs {blue, yellow}

Output Screenshots

.



Learning Outcome :-

1)How Naïve Bayes algorithm will be used.

2)How of SVM will work.

3) How to apply one vs rest SVM.

4) How to preprocess data for text classification

3.Problem Statement:- Implement a Multi Layer (One Input, One Output and One or more Hidden Layers) ANN for handwritten digit classification using MNIST dataset.

Requirement Analysis:

The above problem aims at being able to recognize the digit of a person in a picture when the system has been priorly trained by a set of pictures to allow for feature extraction. We are required to design and use a multilayer ANN or CNN to solve this problem.

Steps Involved:

1)Design and develop the neural network.

2) Provide the digit images as input for training the network

3) Using paint draw one digit and save .png/.jpeg formate and use that image and recogination.

4) Test the system with random test images to observe the accuracy score.

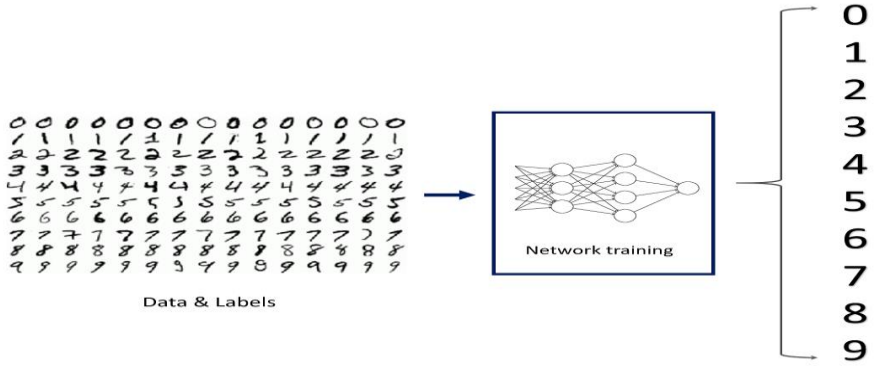
Description

Neural Networks:

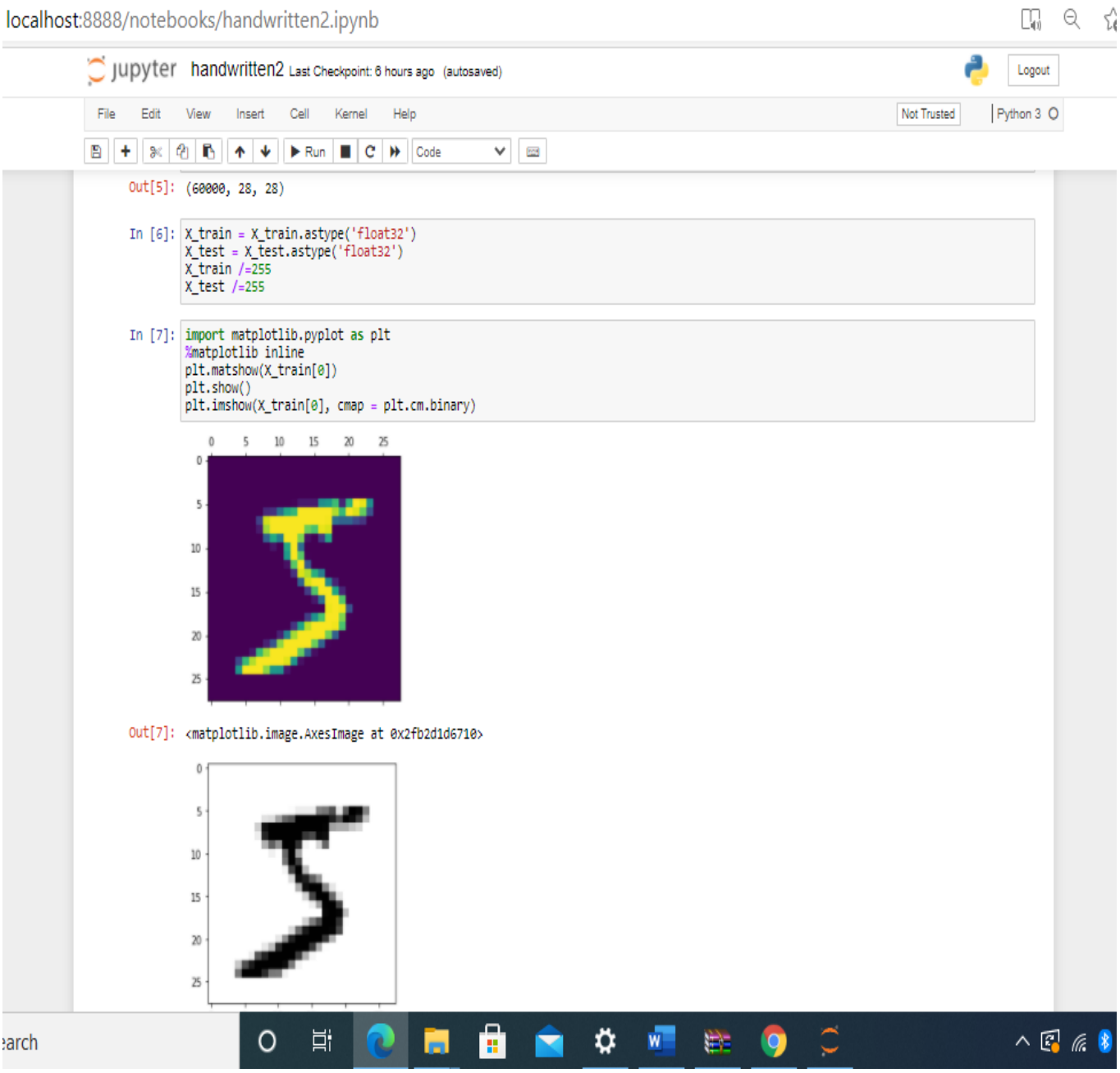
A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature.

MNIST handwritten digits dataset

In this article, we're going to work through a series of simple neural network architectures and compare their performance on the MNIST handwritten digits dataset. The goal for all the networks we examine is the same: take an input image (28x28 pixels) of a handwritten single digit (0–9) and classify the image as the appropriate digit.



Output Screenshots



localhost:8888/notebooks/handwritten2.ipynb

jupyter handwritten2 Last Checkpoint: 6 hours ago (autosaved)

Logout

File Edit View Insert Cell Kernel Help

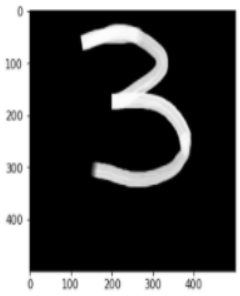
Not Trusted | Python 3

Run Code

In [108]:
```
path = r'C:\Users\HP\OneDrive\Desktop\Hand Writting\3.png'
img = cv2.imread(path)
```

In [109]:
```
plt.imshow(img)
```

Out[109]: <matplotlib.image.AxesImage at 0x2fb2d998cc0>



In [110]:
```
img.shape
```

Out[110]: (500, 500, 3)

In [111]:
```
gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
```

In [112]:
```
gray.shape
```

Out[112]: (500, 500)

In [113]:
```
resized = cv2.resize(gray, (28,28),interpolation = cv2.INTER_AREA)
```

In [114]:
```
resized.shape
```

Out[114]: (28, 28)

In [115]:
```
newimg = tf.keras.utils.normalize (resized, axis = 1)
```

re to search

Steps of Implementation:

1)The design and development of neural networks requires the use of Keras libraries.

2)The Neural network is designed with

3) input layer, several feature extraction middle layers and an output layer which condenses the learning of the model in a set of weights which are stored in a file

To acquire or collect the MNIST handwritten digit images.

4). To divide the input images into training and test images.

5). To apply the pre-processing technique to both the training dataset and the test dataset.

6.) To normalize the data so that it ranges from 0 to 1. 5. To divide the training dataset into batches of a suitable size.

7). To train the CNN model and its variants using the labelled data.

8). To use a trained model for the classification. 8. To analyze the recognition accuracy and processing time for all the variants.

Learning Outcome

1)We were able to learn Nerual network.

2)Know the main types of neural networks;

3)Know and apply the methods of training neural networks;

4)Know the application of artificial neural networks;

5)We were able to understand the hardware and software expenses involved in creating, processing and storing image repositories.