**Handwritten digit classification**

**using Machine Leaning (REPORT)**

**Hardware Required**

∙ Standard computer with at least i3 processor Standard computer with 6GB of RAM

∙ Standard computer with 20GB of free space

∙ Active Internet Connectivity with good bandwidth

. Battery output (keep the laptop charged)

**Software Required**

∙ Jupyter Notebook / Anaconda Navigator

∙ Chrome / Mozilla Browser

∙ Python 3.8 or above

**Libraries Required**

∙ Numpy & Pandas

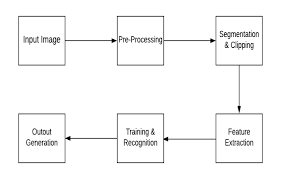
∙ Matplotlib/Seaborn

∙ Scikit Learn

∙ Python 3.8 or above

**Product design & Modelling**

In this document is to look into the design possibilities of the proposed system, such as architecture design, block diagram, sequence diagram, data flow diagram and user interface design of the system in order to define the steps such as pre-processing, feature extraction, segmentation, classification and recognition of digits

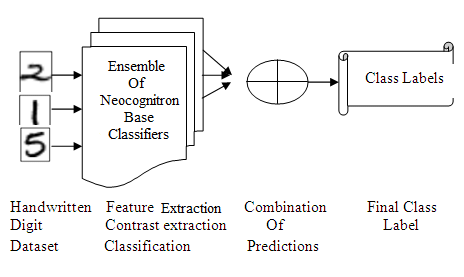


The above Figure 1 illustrates the architecture diagram   
of the proposed system. The proposed model contains the four   
stages in order to classify and detect the digits:   
A. Pre-processing   
B. Segmentation   
C. Feature Extraction   
D. Classification and Recognition

**A. Pre-Processing:**   
The role of the pre-processing step is it performs various tasks on the input image. It basically upgrades the image by making it reasonable for segmentation. The fundamental motivation behind pre-processing is to take off a fascinating example from the background. For the most part, noise filtering, smoothing and standardization are to be done in this stage. The pre-processing additionally characterizes a smaller portrayal of the example. Binarization changes over a gray scale image into a binary image.

**B. Segmentation:**   
Once the pre-processing of the input images is completed, sub-images of individual digits are formed from the sequence of images. Pre-processed digit images are segmented into a sub-image of individual digits, which are assigned a number to each digit. Each individual digit is   
resized into pixels. In this step an edge detection technique is being used for segmentation of dataset images.   
  
**C. Feature Extraction:**   
After the completion of pre-processing stage and segmentation stage, the pre-processed images are represented in the form of a matrix which contains pixels of the images that are of very large size. In this way it will be valuable to represent the digits in the images which contain the necessary information. This activity is called feature extraction. In the feature extraction stage redundancy from the data is removed.   
 **D. Classification and Recognition:**   
In the classification and recognition step the extracted feature vectors are taken as an individual input to each of the following classifiers. In order to showcase the working system   
model extracted features are combined and defined using following three classifiers:   
 K-Nearest Neighbor   
 Random Forest Classifier   
 Support Vector Machine

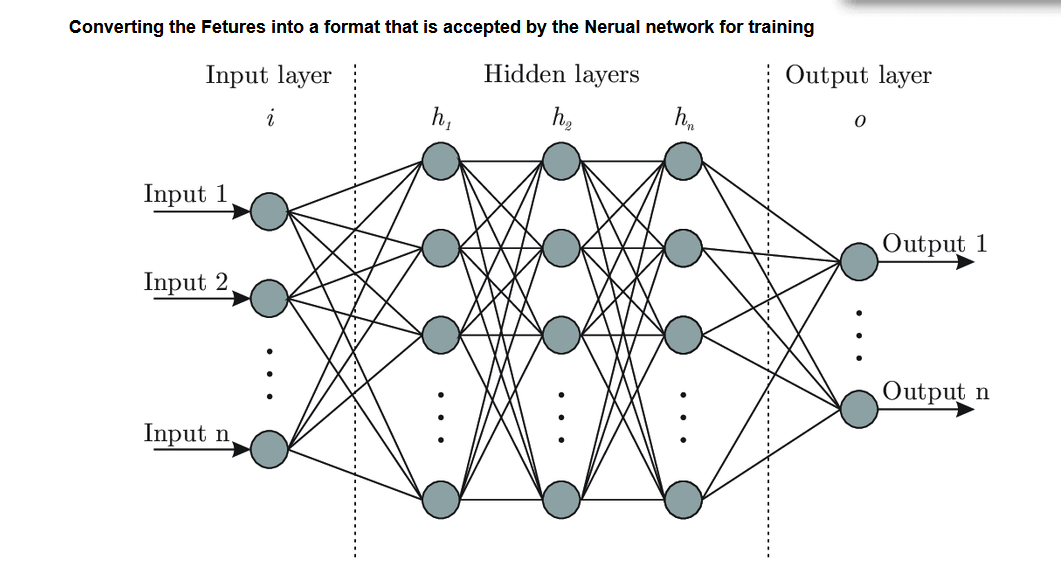
**Methodology**



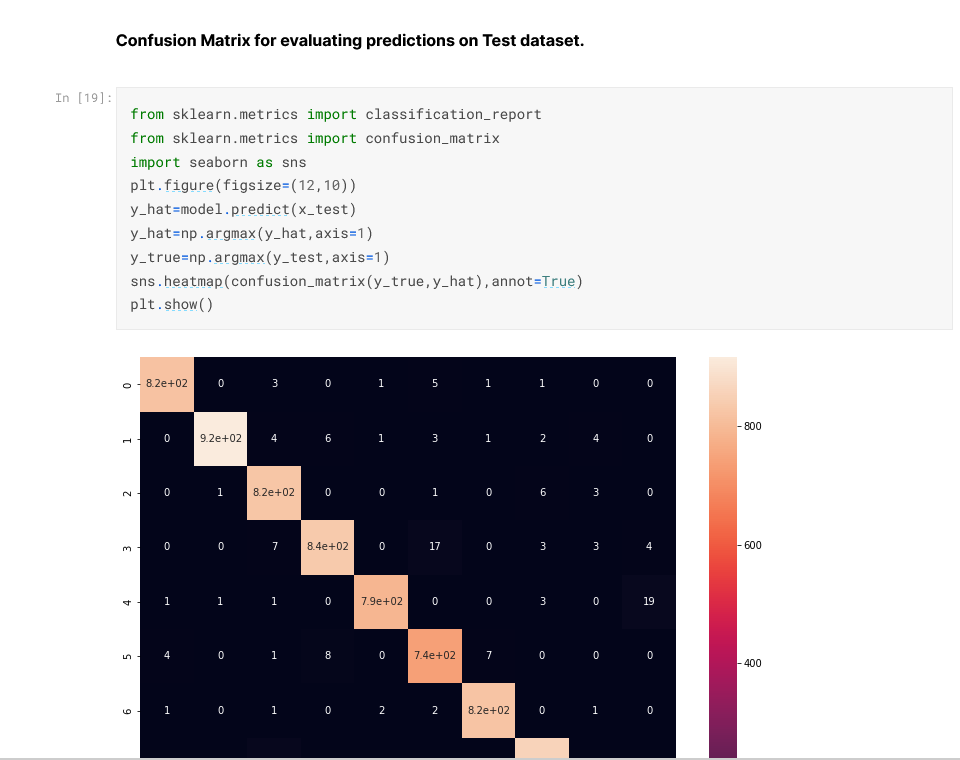
Classically one can guess intelligently for the appropriate size of base neural network, in this case digitized handwritten digits to convergence. In principle, it is quite possible for a classifier

learn input patterns of any complexity provided that training database is quite adequately large

enough. Where as in case in which network size is less than optimum, it is poorly trained. Several techniques have been to improvise the generalization ability of artificial neural networks but for the time being we’ll use classic ML algorithms only.



**Outputs :**

****

