HAND-DRAWN DIGITS RECOGNITION

##### A PROJECT REPORT

###### ***Submitted by***

###### 

**YUVARAJ.M**

**THANIGAHAI VEL.M**

**SURIYA NARAYANAN.A**

**UMMANENI HEMANTH**

**BACHELOR OF ENGINEERING**

*in*

# ELECTRONICS AND COMMUNICATION ENGINEERING

 

VELAMMAL ENGINEERING COLLEGE, SURAPET, CHENNAI-66



ANNA UNIVERSITY:: CHENNAI 600 025

**ODD SEM 2018-19**

APPENDIX 2

VELAMMAL ENGINEERING COLLEGE:CHENNAI 600 066

**BONAFIDE CERTIFICATE**

Certified that this project report **“**HANDWRITTEN DIGITS RECOGNITION**”** is the bonafide work of “**THANIGAHAI VEL.M , YUVARAJ.M , SURIYA NARAYANAN.A , UMMANENI HEMANTH** who carried out the project work under my supervision

**Mr.V.Magesh Dr.S.MARYJOANS**

**Value Added Course Incharge**

**Dept. of Electronics and Professor & Head**

**Communication Engineering Dept. of Electronics and**

**Velammal Engineering College Communication Engineering**

**Chennai-66 Velammal Engineering College**

**Chennai-66**

**External Supervisor**

**Relearn Pro**

**TABLE OF CONTENT**

|  |  |  |
| --- | --- | --- |
| **S.NO** | **CONTENT** | **PAGE NO** |
| 1 | ABSTRACT | 1 |
| 2 | OBJECTIVE | 2 |
| 3 | SOFTWARE DETAILS | 3 |
| 4 | DESCRIPTION | 4 |
| 5 | APPLICATIONS | 5 |
| 6 | OUTCOME OF THE PROJECT | 6 |

**ABSTRACT**

The objective of this project is to identify the number in the handwritten format that was given to us in csv format. Each image is 28 pixels in height and 28 pixels in width, for a total of 784 pixels in total. The training dataset’s first column consists of label which denotes the number being represented by that particular row.Each pixel’s value were given and there was not much pre-processing involved in the dataset as each pixel’s value has to be utilised and they were already in int64 type. Larger the number in a particular pixel darker the pixel. We have used Random Forests to meet the requirements. With Random Forests we were able to attain a pretty good accuracy of 96% for 30% data from the train dataset as test data. We have added a separate file(python notebook) along this mail containing our code and a another file in csv format containing the predictions of our model for the given test dataset.

**OBJECTIVE**

**Aim of the project**

The ultimate aim of the project is to predict the Hand-drawn numbers as accurately as possible

**SOFTWARE DETAILS**

We have used PYTHON 3.6.5 and Anaconda,Inc.product version:5.2.0(JUPYTER NOTEBOOK) to implement our model. We have used the libraries Pandas and Numpy to read the dataset file and to implement it in the model. We used the scikit-learn library to implement the train\_test\_split , for Random Forest and to calculate the mean squared error .

**DESCRIPTION**

The train data set was used to train the model. We used the Pandas in particular to read the csv file. All columns ‘except the label’ column was saved to the variable ‘x’ and the ‘label’ column was saved in variable ‘y’. From scikit-learn the Random Forests was imported and implemented on the training data. The model is thus obtained. The model is then used to predict the label for the test dataset. The result of the model is then stored in a separate file in csv format.

**APPLICATION**

1. The model can be trained to recognize handwritten alphabets to understand the notes of students and convert them to digital data
2. Model can also be used to recognize the handwritten mathematical problems and solve them
3. This allows easy conversion of handwritten data to digital data

**OUTCOME OF THE PROJECT**

1. We got a clear picture over the concept of Random Forests and basic algorithms
2. We will find ease in implementing these algorithms in the near future
3. We will also be able to implement these algorithms in our future projects