A

Major Project

On

HANDWRITTEN DIGITS RECOGNITION USING MACHINE LEARNING

(Submitted in partial fulfillment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

BY

Gampala Ramlikhith(177R1A0514)

Akshay Shivanagaram(177R1A0551)

Venkata Phani Krishna Puppala(167R1A05C0)

Under the Guidance of

N.SRAVANTHI

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
CMR TECHNICAL CAMPUS
UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi)
Recognized Under Section 2(f) & 12(B) of the UGC Act.1956,
Kandlakoya (V), Medchal Road, Hyderabad-501401.
2017-2021

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR
GROUP OF INSTITUTIONS
EXPLORE TO INVENT

CERTIFICATE

This is to certify that the project entitled "**HANDWRITTEN DIGITS RECOGNITION USING MACHINE LEARNING**" being submitted by **Gampala Ramlikhith(177R1A0514), Akshay Shivanagaram(177R1A0551), Venkata Phani Krishna Puppala(167R1A05C0)** in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, is a record of bonafide work carried out by him/her under our guidance and supervision during the year 2020-21.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Ms.N.Sravanthi
Assistant Professor
INTERNAL GUIDE

Dr. A. Raji Reddy
DIRECTOR

Dr. K. Srujan Raju
HoD

EXTERNAL EXAMINER

Submitted for viva voice Examination held on ____________________

ACKNOWLEDGEMENT

Apart from the efforts of us, the success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project. We take this opportunity to express our profound gratitude and deep regard to our guide

Ms. N. Sravanthi , Assistant Professor for her exemplary guidance, monitoring and constant encouragement throughout the project work. The blessing, help and guidance given by her shall carry us a long way in the journey of life on which we are about to embark.

We also take this opportunity to express a deep sense of gratitude to Project Review Committee (PRC) Coordinators : : **Mr. J. Narasimha Rao, Mr. B. P. Deepak Kumar, Mr. K. Murali, Dr. Suwarna Gothane and Mr. B. Ramji** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to the Head of the Department **Dr. K. Srujan Raju** for providing excellent infrastructure and a nice atmosphere for completing this project successfully.

We are obliged to our Director **Dr. A. Raji Reddy** for being cooperative throughout the course of this project. We would like to express our sincere gratitude to our Chairman Sri. **Ch. Gopal Reddy** for his encouragement throughout the course of this project

The guidance and support received from all the members of **CMR TECHNICAL CAMPUS** who contributed and who are contributing to this project, was vital for the success of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement without which this assignment would not be possible. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project.

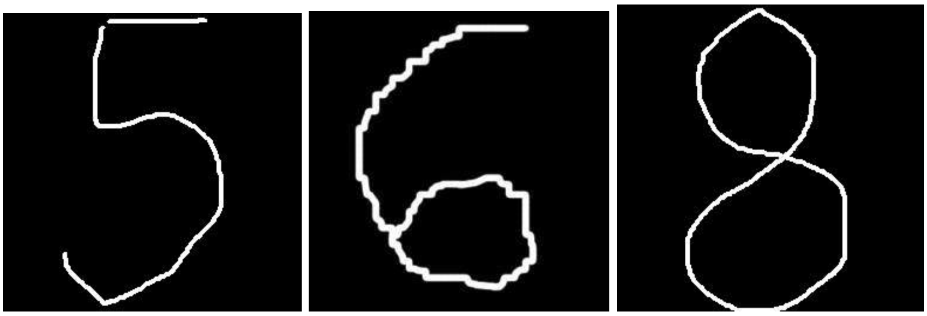
Gampala Ramlikhith(177R1A0514)

Akshay Shivanagaram(177R1A0551)

Venkata Phani KrishnaPuppala(167R1A05C0)

ABSTRACT

The human visual system is one of the wonders of the world. Consider the following sequence of handwritten digits:



Most people effortlessly recognize those digits as 5,6,8. That ease is deceptive. We carry in our heads a supercomputer, tuned by evolution over hundreds of millions of years, and superbly adapted to understand the visual world. Recognizing handwritten digits isn't easy. Rather, we humans are stupendously, astoundingly good at making sense of what our eyes show us. But nearly all that work is done unconsciously. And so we don't usually appreciate how tough a problem our visual systems solve.

The difficulty of visual pattern recognition becomes apparent if you attempt to write a computer program to recognize digits like those above. What seems easy when we do it ourselves suddenly becomes extremely difficult. Simple intuitions about how we recognize shapes - "a 9 has a loop at the top, and a vertical stroke in the bottom right" - turn out to be not so simple to express algorithmically. When you try to make such rules precise, you quickly get lost in a morass of exceptions and caveats and special cases. It seems hopeless.

LIST OF FIGURES

FIGURE NO	FIGURE NAME	PAGE NO
Figure 3.1	Project Architecture	6
Figure 3.3	Use case diagram	7
Figure 3.4	Class diagram	8
Figure 3.5	Activity diagram	9
Figure 3.6	Sequence diagram	10
Figure 3.7	Component diagram	11

LIST OF SCREENSHOTS

SCREENSHOT NO.	SCREENSHOT NAME	PAGE NO.
Screenshot 5.1	Module installation error	13
Screenshot 5.2	Installation of scipy module	13
Screenshot 5.3	Input	14
Screenshot 5.4	Output	14

TABLE OF CONTENTS

ABSTRACT i

LIST OF FIGURES ii

LIST OF SCREENSHOTS iii

1. INTRODUCTION 1

1.1 PROJECT PURPOSE 2

1.2 PROJECT SCOPE 2

1.3 PROJECT OVERVIEW 2

2. SYSTEM ANALYSIS 3

2.1 PROBLEM DEFINITION 3

2.2 EXISTING SYSTEM 3

2.2.1 LIMITATIONS OF THE EXISTING SYSTEM 3

2.3 PROPOSED SYSTEM 4

2.3.1 ADVANTAGES OF PROPOSED SYSTEM 4

2.4 SYSTEM REQUIREMENTS SPECIFICATION 4

2.4.1 FUNCTIONAL REQUIREMENTS 4

2.4.2 NON-FUNCTIONAL REQUIREMENTS 4

2.5 HARDWARE & SOFTWARE REQUIREMENTS 5

2.5.1 HARDWARE REQUIREMENTS 5

2.5.2 SOFTWARE REQUIREMENTS 5

3. ARCHITECTURE 6

3.1 PROJECT ARCHITECTURE 6

3.2 DESCRIPTION 6

3.3 USECASE DIAGRAM 7

3.4 CLASS DIAGRAM 8

3.5 ACTIVITY DIAGRAM 9

3.6 SEQUENCE DIAGRAM 10

3.7 COMPONENT DIAGRAM 11

4. IMPLEMENTATION 12

4.1 SAMPLE CODE 12

5. SCREENSHOTS 13

6. TESTING 15

6.1 INTRODUCTION TO TESTING 15

6.2 TYPES OF TESTING 15

6.2.1 UNIT TESTING 15

6.2.2 INTEGRATION TESTING 15

6.2.3 FUNCTIONAL TESTING 17

6.3 TEST CASES 20

6.3.1 LOADING PATH 20

6.3.2 DIGIT RECOGNITION FROM IMAGE 21

6.4 DISCUSSION OF RESULTS 21

7. CONCLUSION & FUTURE SCOPE 27

7.1 PROJECT CONCLUSION 27

7.2 FUTURE SCOPE 27

8. BIBILOGRAPHY 28

8.1 REFERENCES 28

8.2 WEBSITES 28

1. INTRODUCTION

1. INTRODUCTION

Digit recognition system is the working of a machine to train itself or recognizing the digits from different sources like emails, bank cheque, papers, images, etc. and in different real-world scenarios for online handwriting recognition on computer tablets or system, recognize number plates of vehicles, processing bank cheque amounts, numeric entries in forms filled up by hand and so on.

Handwriting recognition (HWR) is the ability of a computer to receive and handwritten input from sources such as paper documents, photographs, touch screens and other devices. Data science practitioners apply machine learning algorithms to numbers, text, images, video, audio, and more to produce artificial intelligence systems that perform tasks which ordinarily require human intelligence. Scikit learn (formerly scikits learn) is a free software machine learning library for the programming language. It features various classification, regression, and clustering algorithms.

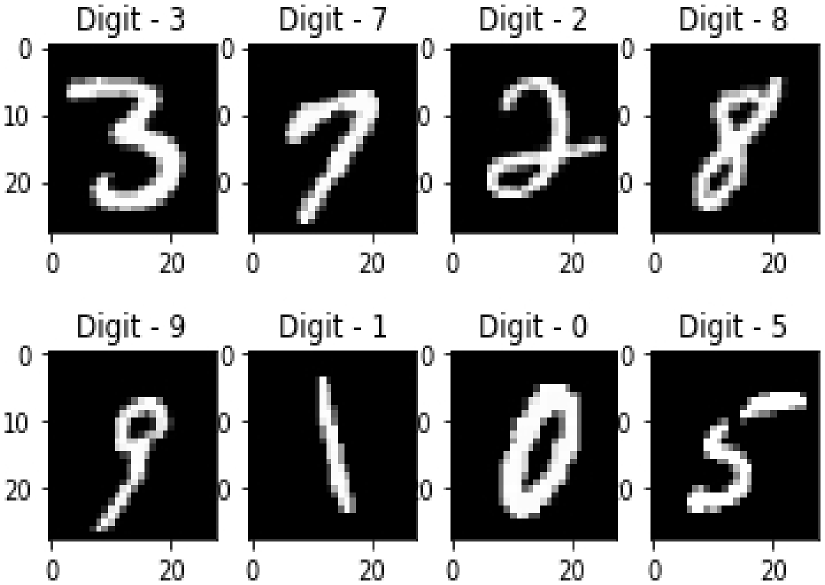


Fig 1.1 Hand Written Digits

Handwritten digit recognition has gained so much popularity from the aspiring beginner of machine learning and deep learning to an expert who has been practicing for years. Developing such a system includes a machine to understand and classify the images of handwritten digits as

10 digits (0–9). Handwritten digits from the MNIST database are already famous among the community for many recent decades. The scope of this article is to compare the different classifiers with different parameters and try to achieve near-human performance.

1.1 PROJECT PURPOSE

Here the purpose is to train the system to learn and identify hand written digits which is given in image format by using image libraries (like pillow, sklearn.digits).

1.2 PROJECT SCOPE

The scope is to make machine to learn and identify the handwritten digits that is 0-9 only that we give in image format by using image libraries like pillow,Sklearn.digit which are used to work with digit recognition.

1.3 PROJECT OVERVIEW

It consists of digit images recognition in which machine is trained itself for recognizing the digit from image. Handwritten recognition (HWR) is ability of a computer to receive and intelligible handwritten input from sources such as paper documents photographs and other devices. Sckit learn is a free software machine learning library for programming language. Developing such a system includes a machine to understand and classify the images of handwritten digits as 10 digits (0–9).

2. SYSTEM ANALYSIS

2. SYSTEM ANALYSIS

SYSTEM ANALYSIS

System Analysis is the important phase in the system development process. The System is studied to the minute details and analyzed. The system analyst plays an important role of an interrogator and dwells deep into the working of the present system. In analysis, a detailed study of these operations performed by the system and their relationships within and outside the system is done. A key question considered here is, “what must be done to solve the problem?” The system is viewed as a whole and the inputs to the system are identified. Once analysis is completed the analyst has a firm understanding of what is to be done.

2.1 PROBLEM DEFINITION

It takes difficult for the system to recognize, identify and classify the handwritten digits.

2.2 EXISTING SYSTEM

In existing system it is easy to recognize the system numericals than the handwritten digits. Often it is difficult to identify the handwritten digits as they differ from writing of person to person.

2.2.1 LIMITATIONS OF EXISTING SYSTEM

- The system uses images which breaks into pixels in identifying the numbers and their sequences of pixels. This makes identification very difficult and the results always vary in real time.
- The handwritten digits are not always of the same size, width, and justified to margins as they differ from person to person.
- The uniqueness and variety in the handwriting of different individuals also influence the formation and appearance of the digits.

- The problem is faced more when many people write a single digit with a variety of handwritings.

2.3 PROPOSESD SYSTEM

The proposed system uses image libraries to identify the handwritten digits. Here the image libraries like scipy, pillow (it is used to read or scan image), sklearn.dataset (which is used to work with only digits) used in python. By using these libraries the machine will be trained and can easily classify and identify the handwritten digits.

2.3.1 ADVANTAGES OF THE PROPOSED SYSTEM

- It recognizes the digits in different styles.
- It can easily scan and read the image using image libraries.
- The handwritten digit can be of any size or shape can easily be identified by the system.
- Python has good graphical libraries. The output is more effective using graphical libraries in python.

2.4. SYSTEM REQUIREMENTS SPECIFICATION

2.4.1 FUNCTIONAL REQUIREMENTS

- Loading image.
- Image visualization.
- Recoginsing hand written digit from image.
- Predict in real-time the value of number
- System will process the input given by the user only when it is given in the image format only.
- System will display the error message to the user, when the given input by the user was not in the required format.
- System will detect the number present in the image.
- And also system identify the given number and display the result (i.e.given number).

2.4.2 NON-FUNCTIONAL REQUIREMENTS

- As the name suggests these are the requirements that are not directly interacted with the specific functions delivered by the system.

- The system will retrieve the handwritten text regions only if the image contains written text in it.
- To work with handwritten digits recognition by using machine learning mainly it should have machine learning library that is scikit-learn.
- It provides the users to load the image easily if there is a Scikit-image library only.

2.5 HARDWARE & SOFTWARE REQUIREMENTS

2.5.1 SOFTWARE REQUIREMENTS.

Table 2.1: Showing Software Requirements

Operating System	Windows 7/8
Language	Python 3.5
Browser	Mozilla firefox(or any browser)

2.5.2 HARDWARE REQUIREMENTS

Table 2.2: Showing Hardware Requirements

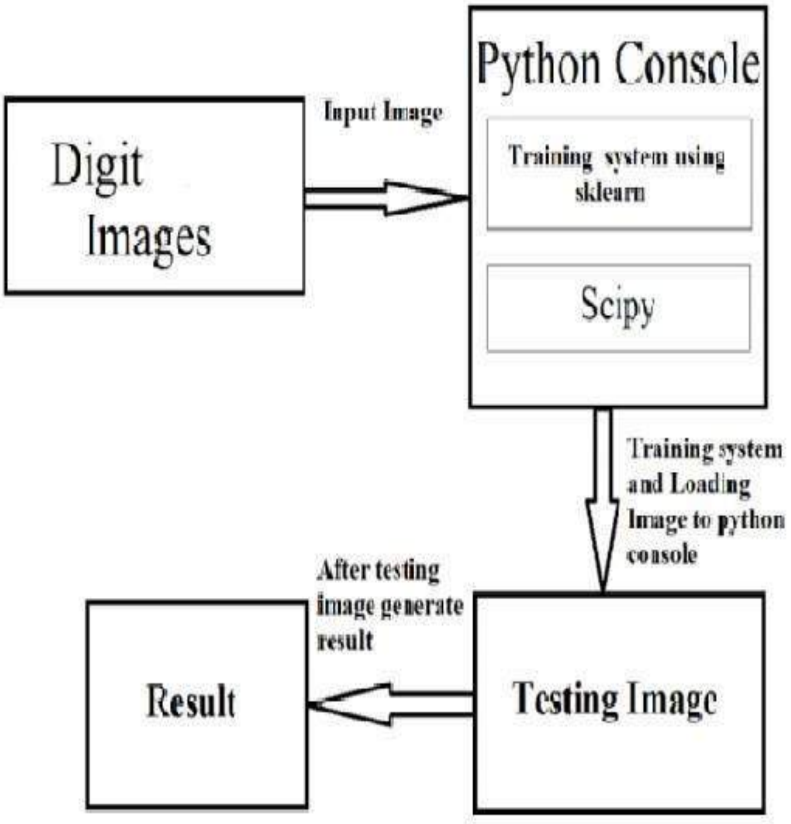
Processor	Pentium 4
RAM	2GB or 4GB RAM
HARD DISK	40GB

3. ARCHITECTURE

3. ARCHITECTURE

3.1 PROJECT ARCITECTURE

This project architecture shows the procedure followed for breed detectionusing machine learning, starting from input to final prediction.


Digit Images
Input Image
Python Console
Training system using sklearn
Scipy
Training system and Loading Image to python console
After testing image generate result
Result
Testing Image

3.2 DESCRIPTION

The below figure describes the process of hand written digit recognition. Initially digits of image are taken as input and given that image to the python console. Load the data sets using installed sklearn, scipy module in these modules data sets are already available we are just required to import the module and use those modules. Later converting the image into data sets and testing the image. If the generated result is correct considered it as correct. If the system unable to identify then the system 1 as output.

3.3 USE CASE DIAGRAM

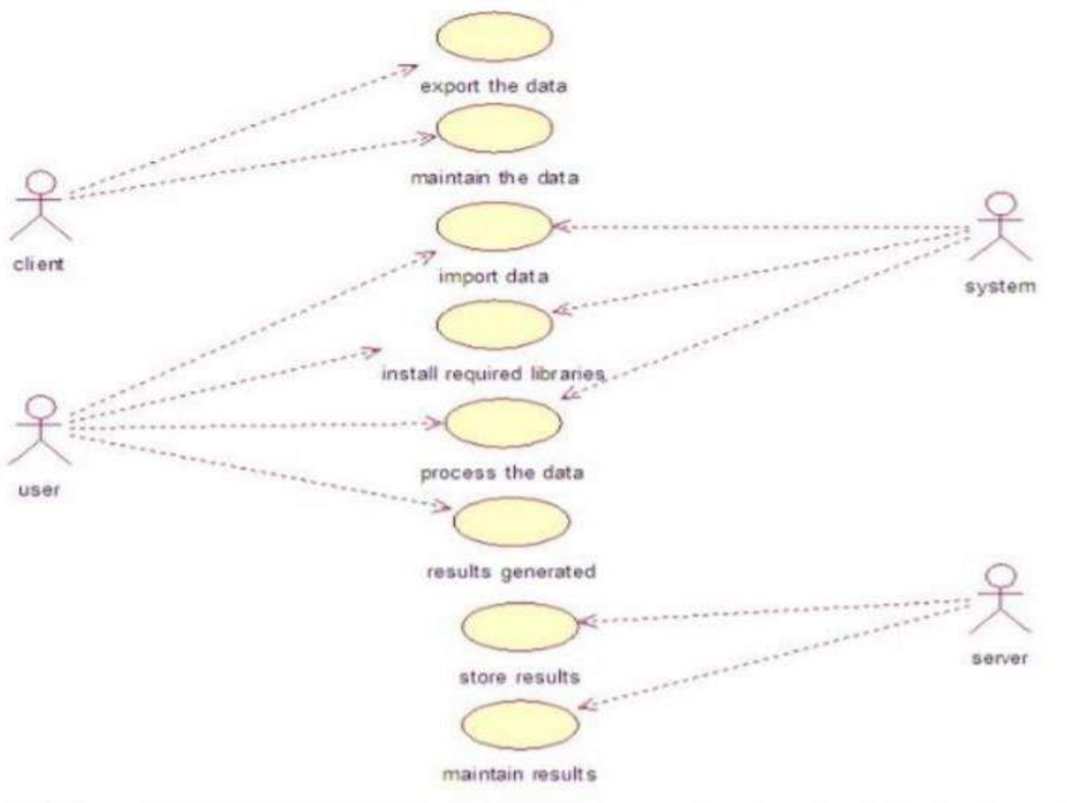


Fig: 3.1 Usecase diagram for hand wirtten recognition

3.4 CLASS DIAGRAM

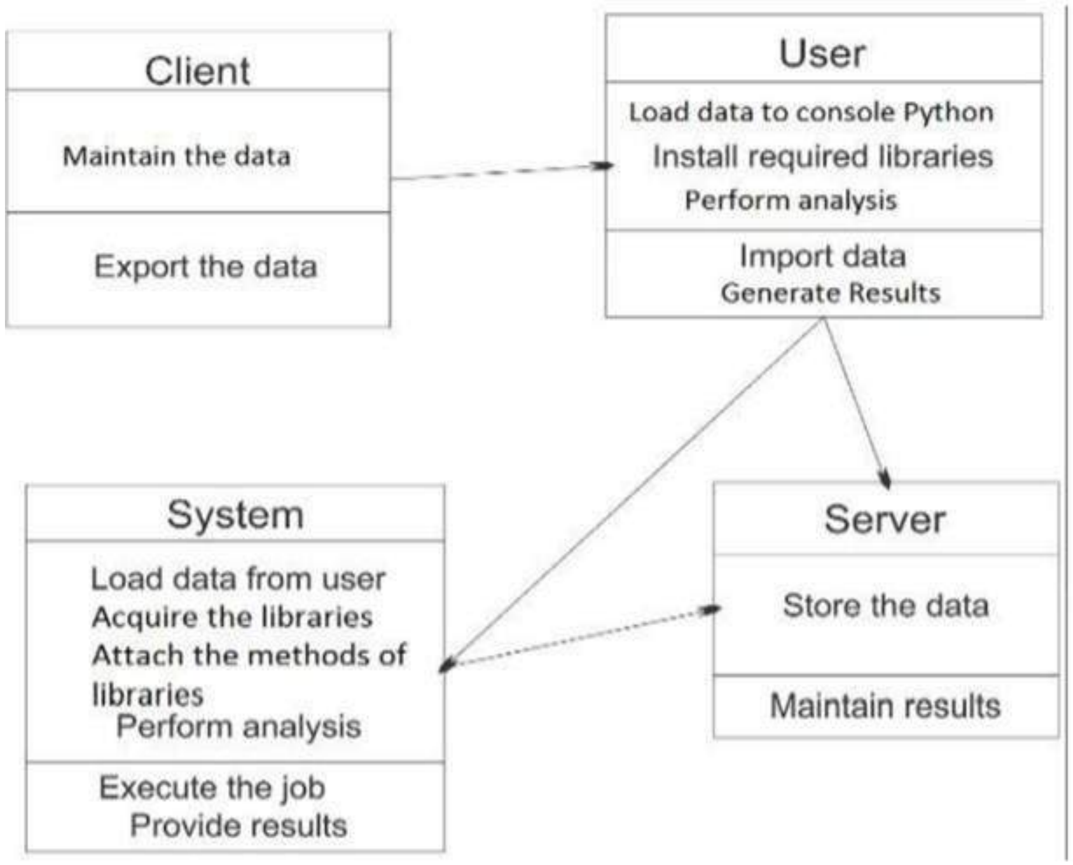


Fig:3.2 Class diagram for hand wirtten recognition

3.5 ACTIVITY DIAGRAM

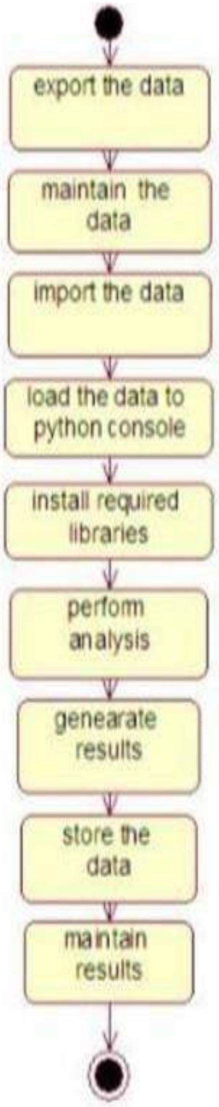


Fig: 3.3 Activity diagram for hand wirtten recognition

3.6 SEQUENCE DIAGRAM

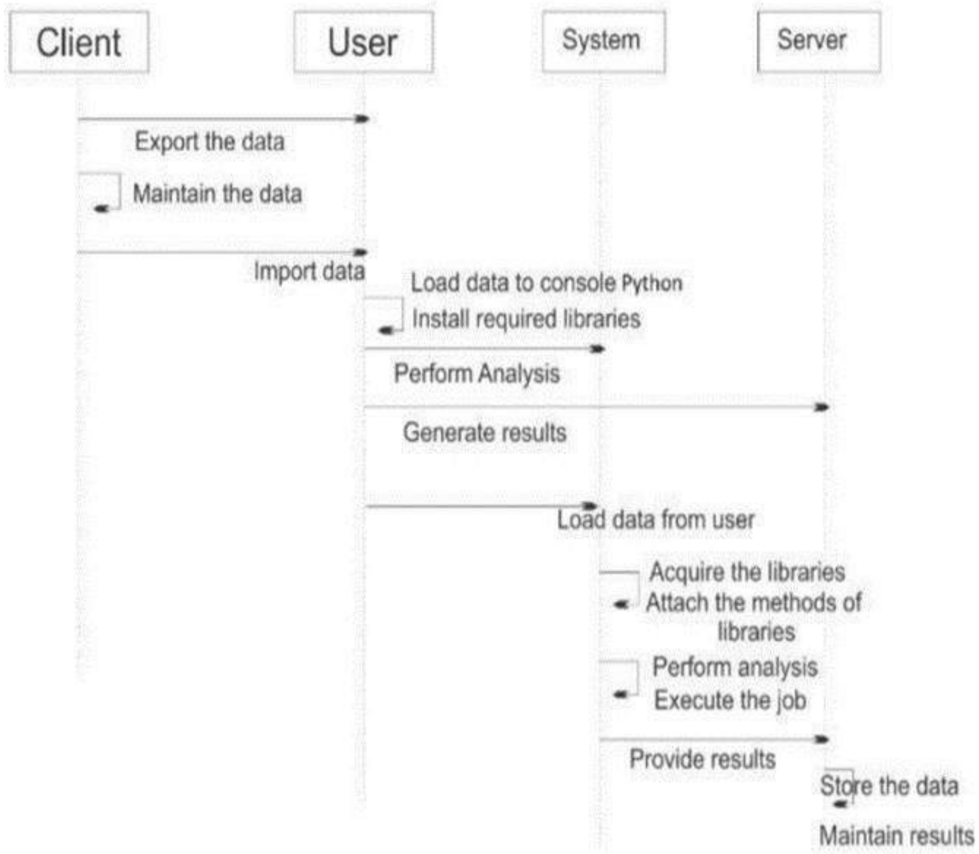


Fig:3.4 Sequence diagram for hand wirtten recognition

3.7 COMPONENT DIAGRAM

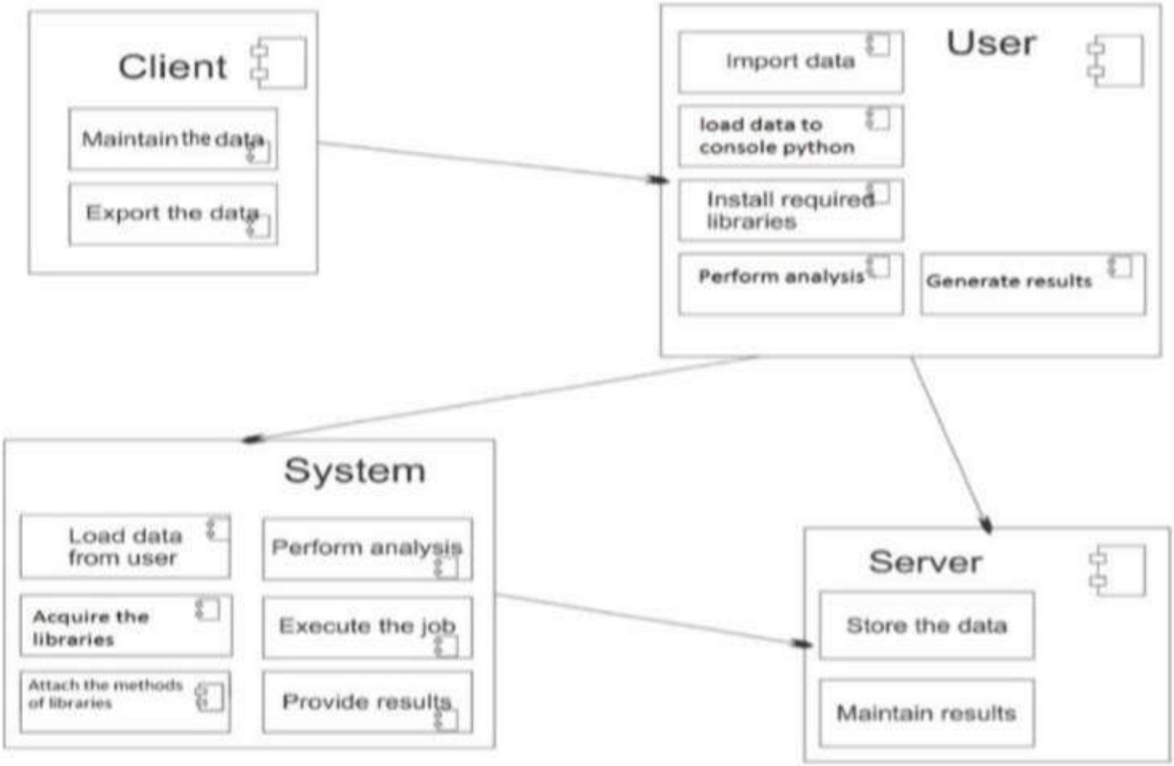


Fig: 3.5 Component diagram for hand wirtten recognition

4. IMPLEMENTATION

4. IMPLEMENTATION

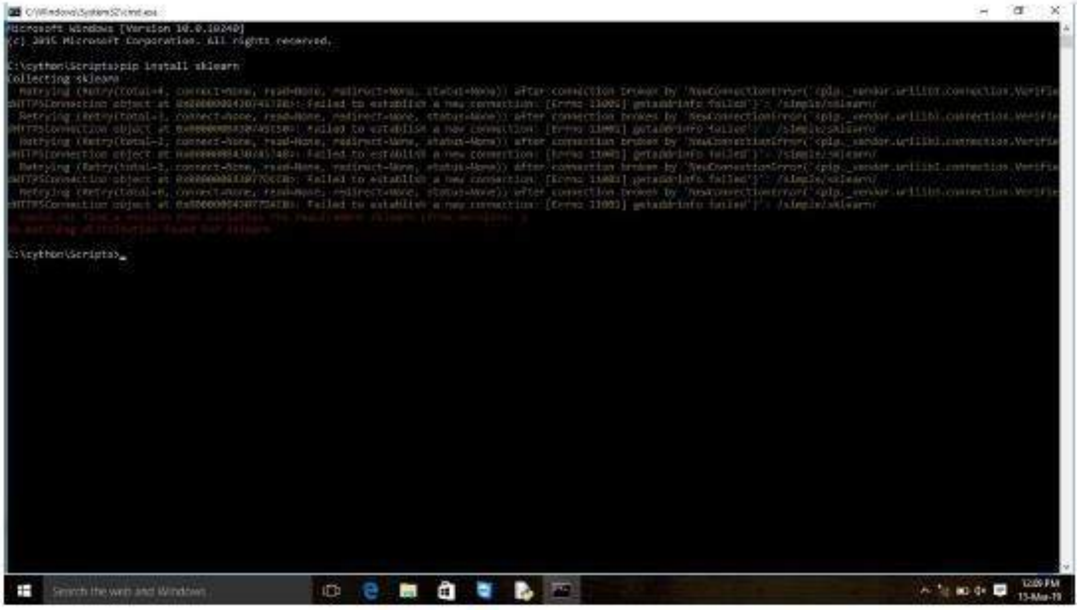
4.1 SAMPLE CODE

```
from sklearn import datasets
from sklearn.svm import SVC
from scipy import misc
        digits = datasets.load_digits()
features = digits.data
labels = digits.target
         clf = SVC(gamma = 0.001)
clf.fit(features, labels)
img = misc.imread("C:/Users/NAVYA/Desktop/project19/123.jpg")
img = misc.imresize(img, (8,8))
img = img.astype(digits.images.dtype)
img = misc.bytescale(img, high=16, low=0)
         x_test = [ ]
for eachRow in img:
for eachPixel in eachRow:
x_test.append(sum(eachPixel)/3.0)
         print(clf.predict([x_test]))
```

5. SCREENSHOTS

5.SCREENSHOTS

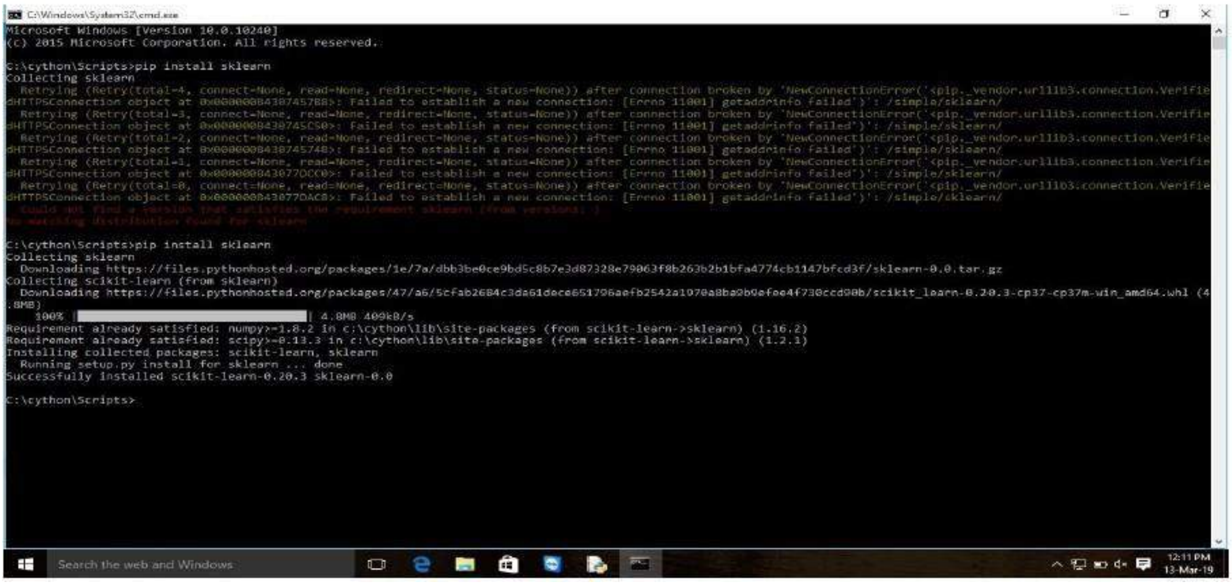
5.1 MODULE INSTALLATION ERROR



```
Microsoft Windows [Version 10.0.10240]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Retrying (Retry(total=4, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x00000084307457B8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=3, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843074SC50>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=2, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745748>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=1, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DCC0>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=0, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DAC8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Could not find a version that satisfies the requirement sklearn (from versions: )
No matching distribution found for sklearn

C:\cython\Scripts>_
```

Fig5.1 :MODULE INSTALLATION ERROR

5.2 INSTALLATION OF SCIPY MODULE



```
Microsoft Windows [Version 10.0.10240]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Retrying (Retry(total=4, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x00000084307457B8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=3, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843074SC50>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=2, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745748>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=1, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DCC0>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=0, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DAC8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Could not find a version that satisfies the requirement sklearn (from versions: )
No matching distribution found for sklearn

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Downloading https://files.pythonhosted.org/packages/1e/7a/dbb3be0ce9bd5c8b7e3d87328e79063f8b263b2b1bfa4774cb1147bfcd3f/sklearn-0.0.tar.gz
Collecting scikit-learn (from sklearn)
  Downloading https://files.pythonhosted.org/packages/47/a6/5cfab2684c3da61dece651796aefb2542a1970a8ba9b9efee4f730ccd90b/scikit_learn-0.20.3-cp37-cp37m-win_amd64.whl (4
.8MB)
    100% |████████████████████████████████| 4.8MB 409kB/s
Requirement already satisfied: numpy>=1.8.2 in c:\cython\lib\site-packages (from scikit-learn->sklearn) (1.16.2)
Requirement already satisfied: scipy>=0.13.3 in c:\cython\lib\site-packages (from scikit-learn->sklearn) (1.2.1)
Installing collected packages: scikit-learn, sklearn
  Running setup.py install for sklearn ... done
Successfully installed scikit-learn-0.20.3 sklearn-0.0

C:\cython\Scripts>
```

FIG 5.2: INSTALLATION OF SCIPY MODULE

5.3 INPUT

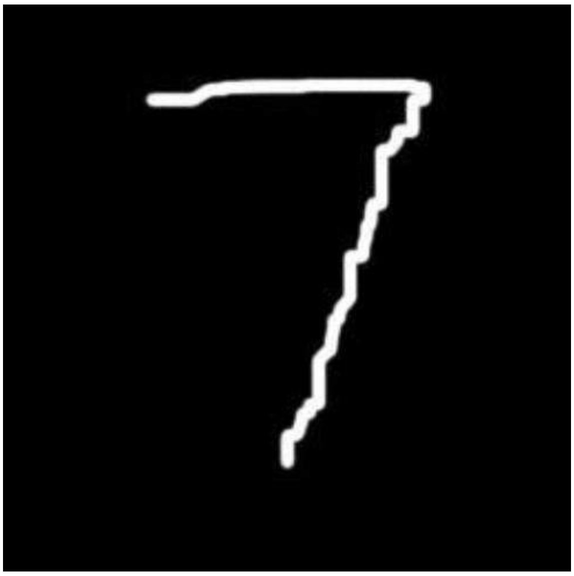
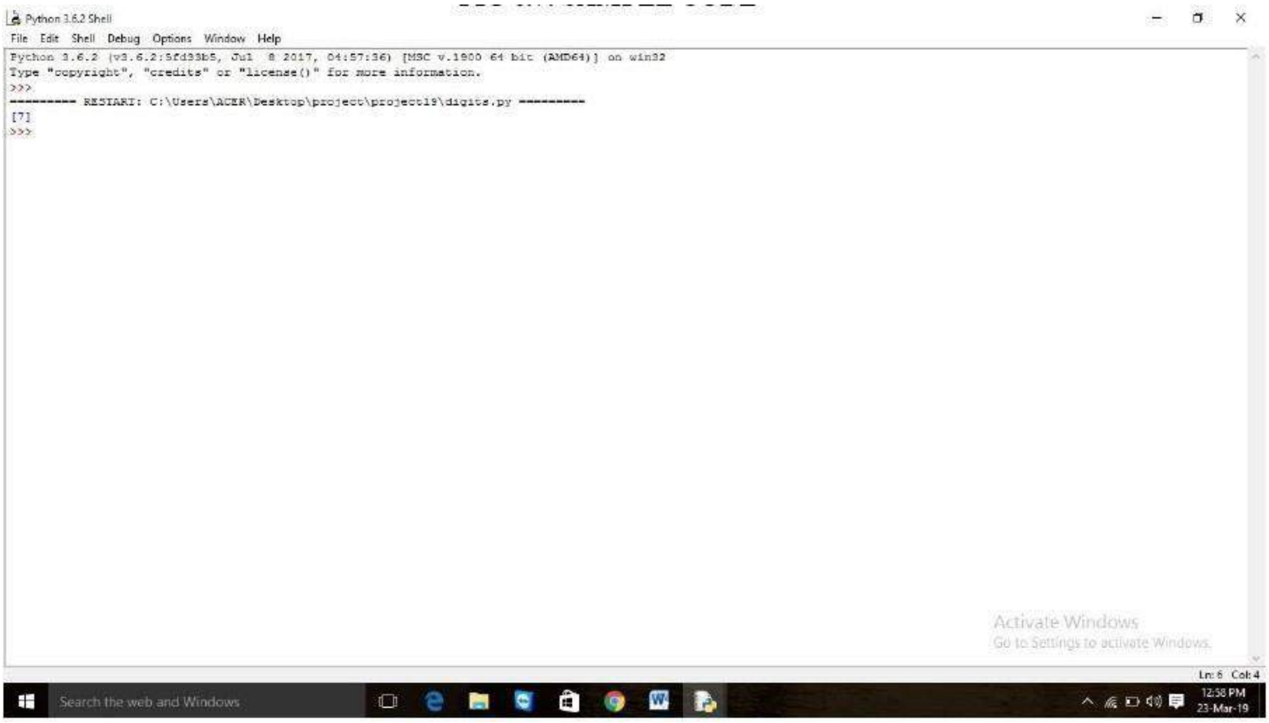


FIG 5.3:INPUT

5.4 OUTPUT


Python 3.6.2 Shell
File Edit Shell Debug Options Window Help
Activate Windows
Go to Settings to activate Windows.
Ln: 6 Col: 4
Search the web and Windows
12:58 PM
23-Mar-19

```
Python 3.6.2 (v3.6.2:5fd33b5, Jul  8 2017, 04:57:36) [MSC v.1900 64 bit (AMD64)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
========= RESTART: C:\Users\ACER\Desktop\project\project19\digits.py =========
[7]
>>>
```

FIG 5.4:OUTPUT

6. TESTING

6. TESTING

6.1 INTRODUCTION TO TESTING

Designers compose unit tests to check their own particular code. Unit testing varies from coordination testing, which affirms that parts function admirably together, and acknowledgment testing, which affirms that an application does what the client anticipates that it will do. Unit tests are so named in light of the fact that they test a solitary\unit of code. A unit test is completely computerized, non-intelligent, and double—that is, it either succeeds or fizzles. So running your code and looking at its yield to check whether it works is not a test . Nor is composing a bit "test pilot" that drives your code or enables you to check logs to check whether it's working accurately.

When you compose unit tests, at times you feel constrained to change your code just to encourage the test, more often than not when you have to test a private strategy or trait. Doing as such is a terrible thought. In the event that you ever feel enticed to make a private technique open only to test purposes, don't do it. Testing is intended to enhance the nature of your code, not diminish it.

6.2 TYPES OF TESTING

6.2.1 Unit Testing

Unit testing is a level of software testing where individual units/ components of software are tested. The purpose is to validate that each unit of the software performs as designed. A unit is the smallest testable part of any software. It usually has one or a few inputs and usually a single output.

6.2.2 Integration Testing

Integration Testing is a level of software testing where individual units are combined and tested as a group. The purpose of this level of testing is to expose faults in the interaction between integrated units. Test drivers and test stubs are used to assist in Integration Testing.

6.2.2.1 Integration Test Approaches

There are four types of integration testing approaches. Those approaches are the following

6.2.2.2 Big-Bang Integration Testing

It is the simplest integration testing approach, where all the modules are combining and verifying the functionality after the completion of individual module testing. In simple words, all the modules of the system are simply put together and tested. This approach is practicable only for very small systems. If once an error is found during the integration testing, it is very difficult to localize the error as the error may potentially belong to any of the modules being integrated. So, debugging errors reported during big bang integration testing are very expensive to fix.

6.2.2.3 Bottom-Up Integration Testing

In bottom-up testing, each module at lower levels is tested with higher modules until all modules are tested. The primary purpose of this integration testing is, each subsystem is to test the interfaces among various modules making up the subsystem. This integration testing uses test drivers to drive and pass appropriate data to the lower level modules.

6.2.2.4 Top-Down Integration Testing

Top-down integration testing technique used in order to simulate the behavior of the lower-level modules that are not yet integrated. In this integration testing, testing takes place from top to bottom. First high-level modules are tested and then low-level modules and finally integrating the low-level modules to a high level to ensure the system is working as intended.

6.2.2.5 Test Drivers

Test Drivers are used during Bottom-up integration testing in order to simulate the behaviour of the upper level modules that are not yet integrated. Test Drivers are the modules that act as temporary replacement for a calling module and give the same output as that of the actual product. Drivers are also used when the software needs to interact with an external system and are usually complex than stubs.

6.2.2.6 Test Stubs

Stubs are used during Top-down integration testing, in order to simulate the behavior of the lower-level modules that are not yet integrated. Stubs are the modules that act as temporary replacement for a called module and give the same output as that of the actual product. Stubs are also used when the software needs to interact with an external system.

6.2.2.3 Acceptance Testing

Acceptance testing, a testing technique performed to determine whether or not the software system has met the requirement specifications. The main purpose of this test is to evaluate the system's compliance with the business requirements and verify if it is has met the required criteria for delivery to end users.

6.2.3 Functional Testing

Functional testing is a type of software testing whereby the system is tested against the functional requirements/specifications.

Functions (or features) are tested by feeding them input and examining the output. Functional testing ensures that the requirements are properly satisfied by the application. This type of testing is not concerned with how processing occurs, but rather, with the results of processing. It simulates actual system usage but does not make any system structure assumptions. During functional testing, Black Box Testing technique is used in which the internal logic of the system being tested is not known to the tester.

6.2.4 Black Boxtesting

Black box testing is a type of software testing in which the functionality of the software is not known. The testing is done without the internal knowledge of the products.

Black box testing can be done in following ways:

6.2.4.1 Syntax Driven Testing

This type of testing is applied to systems that can be syntactically represented by some language. For example- compilers, language that can be represented by context free grammar. In this, the test cases are generated so that each grammar rule is used at least once.

6.2.4.2 Equivalence Partitioning

It is often seen that many type of inputs work similarly so instead of giving all of them separately we can group them together and test only one input of each group. The idea is to partition the input domain of the system into a number of equivalence classes such that each member of class works in a similar way, i.e., if a test case in one class results in some error, other members of class would also result into same error.

The technique involves two steps:

6.2.4.3 Identification Of Equivalence Class

Partition any input domain into minimum two sets: valid values and invalid values. For example, if the valid range is 0 to 100 then select one valid input like 49 and one invalid like 104.

6.2.4.4 Generating Test Cases

(i) To each valid and invalid class of input assign unique identification number.

(ii) Write test case covering all valid and invalid test case considering that no two
invalid inputs mask each other.

6.2.5 White Box Testing

White box testing techniques analyze the internal structures the used data structures, internal design, code structure and the working of the software rather than just the functionality as in black box testing. It is also called glass box testing or clear box testing or structural testing.

6.2.5.1 Working Process of White Box Testing

Input

Requirements, Functional specifications, design documents, source code.

Processing

Performing risk analysis for guiding through the entire process.

Output

Preparing final report of the entire testing process.

Testing Techniques Statement Coverage

In this technique, the aim is to traverse all statement at least once. Hence, each line of code is tested. In case of a flowchart, every node must be traversed at least once. Since all lines of code are covered, helps in pointing out faulty code.

Branch Coverge

In this technique, test cases are designed so that each branch from all decision points is traversed at least once. In a flowchart, all edges must be traversed at least once.

Advantages

- White box testing is very thorough as the entire code and structures are tested.
- It results in the optimization of code removing error and helps in removing extra lines of code.
- It can start at an earlier stage as it doesn’t require any interface as in case of black box testing.
- Easy to automate.

Disadvantages

- Main disadvantage is that it is very expensive.
- Redesign of code and rewriting code needs test cases to be written again.
- Testers are required to have in-depth knowledge of the code and programming language as opposed to black box testing.
- Missing functionalities cannot be detected as the code that exists is tested.
- Very complex and at times not realistic.

6.2.3 Validation Testing

- The process of evaluating software during the development process or at the end of the development process to determine whether it satisfies specified business requirements.
- Validation Testing ensures that the product actually meets the client's needs. It can also be defined as to demonstrate that the product fulfills its intended use when deployed on appropriate environment.
- It answers to the question, Are we building the right product?

6.3 TESTCASES

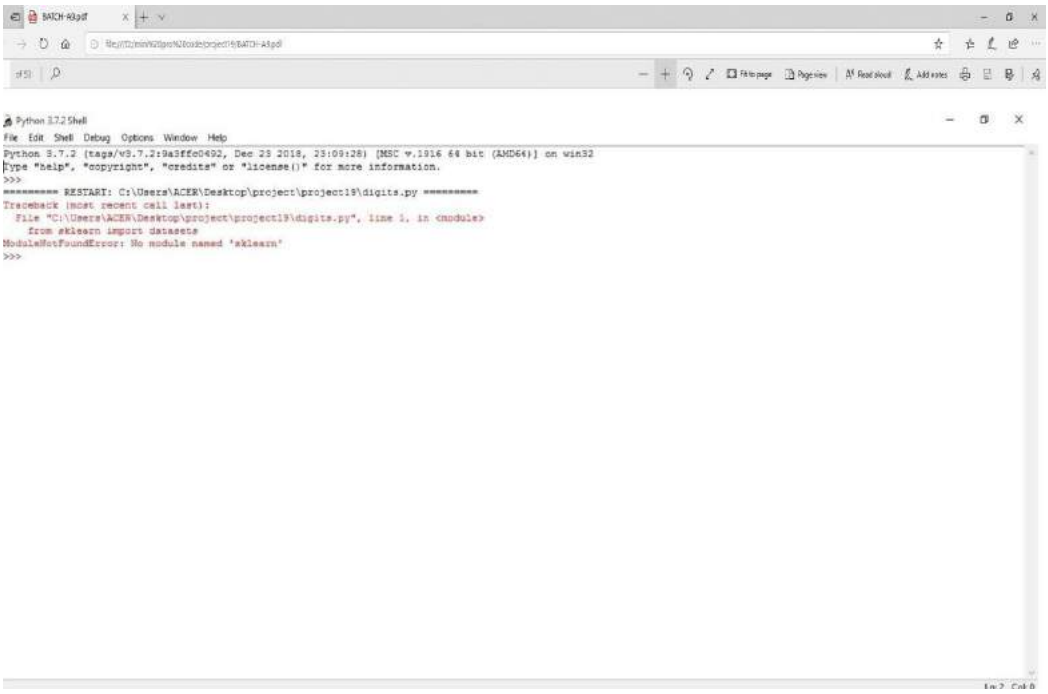
Table 6.3.1: Loading Path

Test case#:1	Priority(h,1): high	
Test objective	Correct Path details	
Test description	Path need be checked	
Requirements verified	Path checked in user system	
Test environments	(IDLE)	
Test setup	User initiates any control mechanism	
Actions	Expected results	Actual results
Incorrect path	Allows user to enter correct stem path	Incorrect system path
Correct path	Console	Console
Pass: yes	Conditional pass	Fail

Table 6.3.2: digit recognition from image

Test case#:1	Image identification	
Test objective	Identify the image	
Test description	Machine should identify the written digit from the image	
Requirements verified	Verified	
Test environments	(IDLE)	
Test setup	Windows	
Actions	Expected results	Actual results
Incorrect image	Correct recognition of writt git from image	Correct recognition ritten digit from image
Correct image	Console	Console
Pass: yes	Conditional pass	Pass

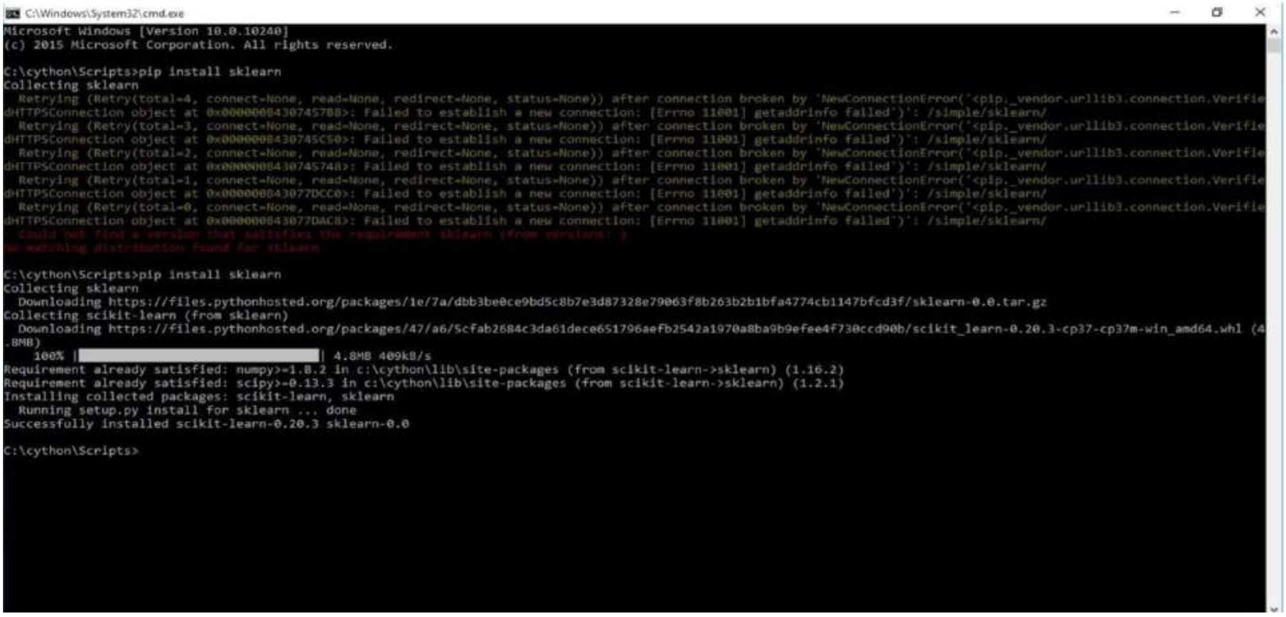
6.4 DISCUSSION OF RESULTS



```
Python 3.7.2 (tags/v3.7.2:9a3ffc0492, Dec 23 2018, 23:09:28) [MSC v.1916 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
========= RESTART: C:\Users\ACER\Desktop\project\project19\digits.py =========
Traceback (most recent call last):
  File "C:\Users\ACER\Desktop\project\project19\digits.py", line 1, in <module>
    from sklearn import datasets
ModuleNotFoundError: No module named 'sklearn'
>>>
```

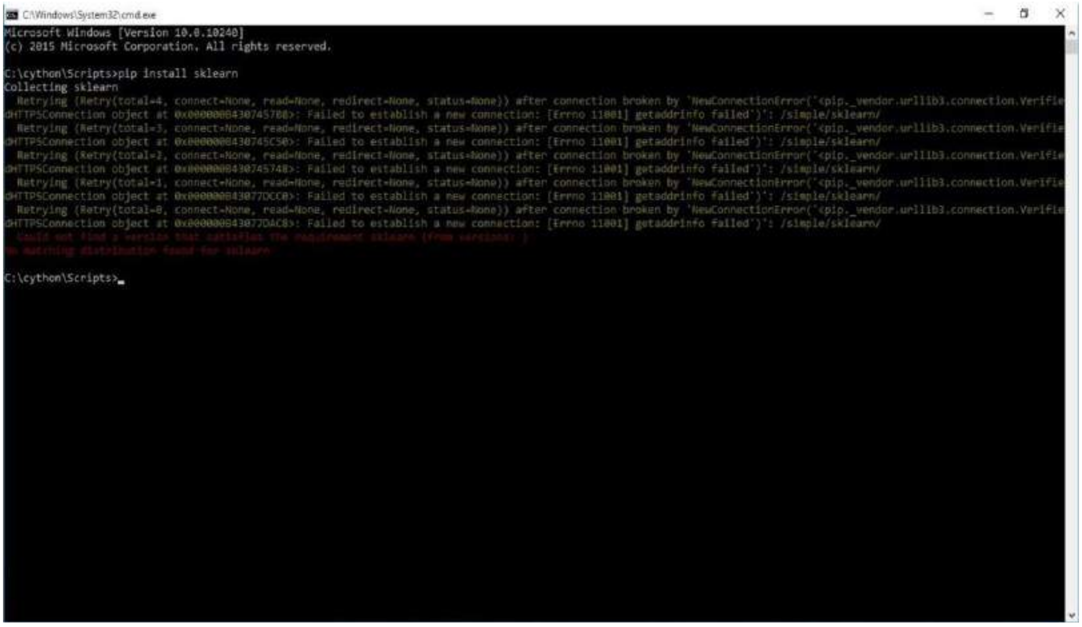
Fig 6.3.1 Module failed error

- With out installing the sklearn library wie will not get the out put because it is a machine library .
- After running the program without installing the modules then we will get error.
- The module not found error as sklearn module not found error.



```
Microsoft Windows [Version 10.0.10240]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Retrying (Retry(total=4, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x00000084307457B8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=3, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745C50>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=2, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745748>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=1, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DCC0>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=0, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DAC8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Could not find a version that satisfies the requirement sklearn (from versions: )
No matching distribution found for sklearn

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Downloading https://files.pythonhosted.org/packages/1e/7a/dbb3be0ce9bd5c8b7e3d87328e79063f8b263b2b1bfa4774cb1147bfcd3f/sklearn-0.0.tar.gz
Collecting scikit-learn (from sklearn)
  Downloading https://files.pythonhosted.org/packages/47/a6/5cfab2684c3da61dece651796aefb2542a1970a8ba9b9efee4f730ccd90b/scikit_learn-0.20.3-cp37-cp37m-win_amd64.whl (4
.8MB)
    100% |████████████████████████████████| 4.8MB 409kB/s
Requirement already satisfied: numpy>=1.8.2 in c:\cython\lib\site-packages (from scikit-learn->sklearn) (1.16.2)
Requirement already satisfied: scipy>=0.13.3 in c:\cython\lib\site-packages (from scikit-learn->sklearn) (1.2.1)
Installing collected packages: scikit-learn, sklearn
  Running setup.py install for sklearn ... done
Successfully installed scikit-learn-0.20.3 sklearn-0.0

C:\cython\Scripts>
```

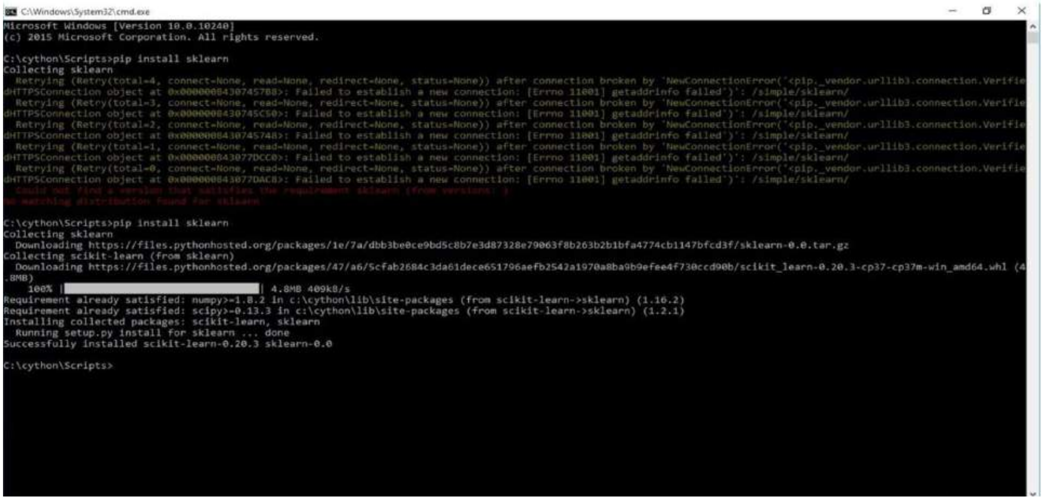
6.3.1 Installation error



```
Microsoft Windows [Version 10.0.10240]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Retrying (Retry(total=4, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x00000084307457B8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=3, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745C50>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=2, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745748>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=1, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DCC0>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=0, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DAC8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Could not find a version that satisfies the requirement sklearn (from versions: )
No matching distribution found for sklearn

C:\cython\Scripts>_
```

Fig 6.3.2 Module installation error

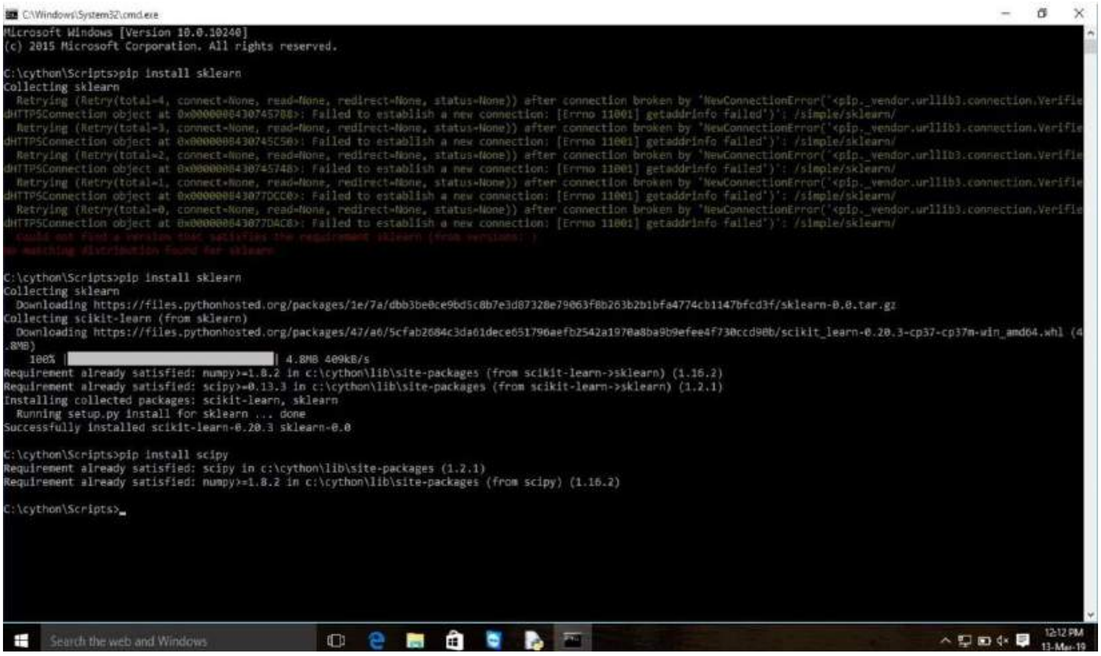
- For installing the modules we need internet connectivity.
- If we are not connecting to the internet to install modules then we will get network failure error.
- We need to install modules in command prompt.



```
Microsoft Windows [Version 10.0.10240]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Retrying (Retry(total=4, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x00000084307457B8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=3, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745C50>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=2, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745748>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=1, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DCC0>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=0, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DAC8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Could not find a version that satisfies the requirement sklearn (from versions: )
No matching distribution found for sklearn

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Downloading https://files.pythonhosted.org/packages/1e/7a/dbb3be0ce9bd5c8b7e3d87328e79063f8b263b2b1bfa4774cb1147bfcd3f/sklearn-0.0.tar.gz
Collecting scikit-learn (from sklearn)
  Downloading https://files.pythonhosted.org/packages/47/a6/5cfab2684c3da61dece651796aefb2542a1970a8ba9b9efee4f730ccd90b/scikit_learn-0.20.3-cp37-cp37m-win_amd64.whl (4
.8MB)
    100% |████████████████████████████████| 4.8MB 409kB/s
Requirement already satisfied: numpy>=1.8.2 in c:\cython\lib\site-packages (from scikit-learn->sklearn) (1.16.2)
Requirement already satisfied: scipy>=0.13.3 in c:\cython\lib\site-packages (from scikit-learn->sklearn) (1.2.1)
Installing collected packages: scikit-learn, sklearn
  Running setup.py install for sklearn ... done
Successfully installed scikit-learn-0.20.3 sklearn-0.0

C:\cython\Scripts>
```

Fig 6.3.3 Installation of sklearn module

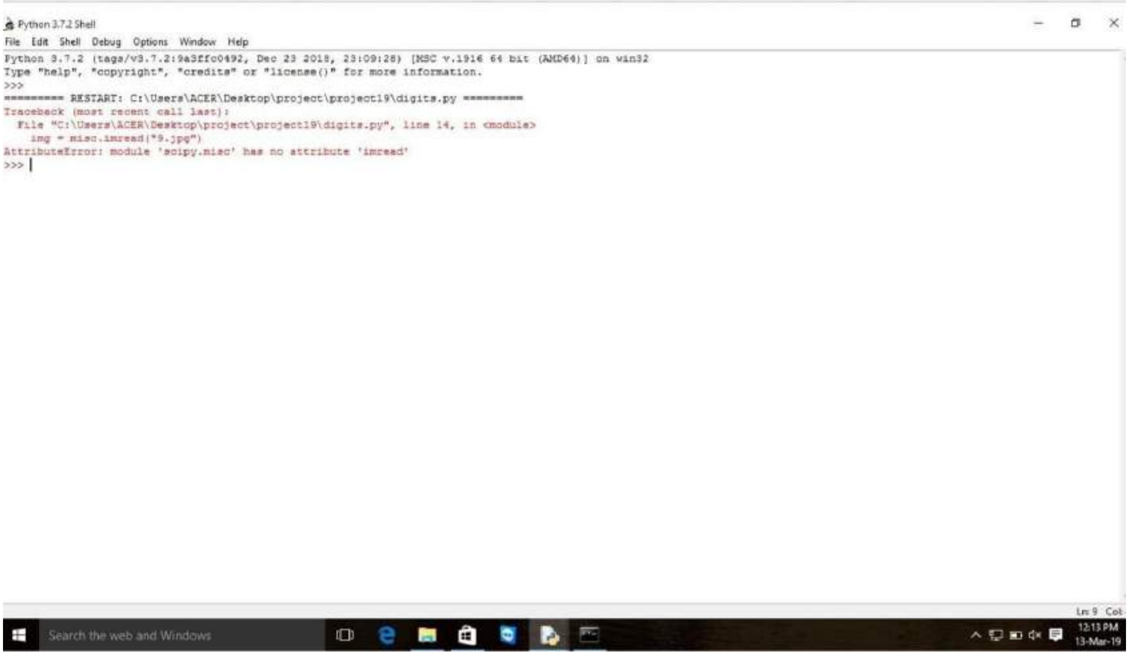
- We install sklearn module to import the data sets to our program.
- The sklearn is a machine library for hand written digit recognition from image.
- It include algorithms for filtering,feature detection, and more.



```
Microsoft Windows [Version 10.0.10240]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Retrying (Retry(total=4, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x00000084307457B8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=3, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745C50>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=2, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x0000008430745748>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=1, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DCC0>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Retrying (Retry(total=0, connect=None, read=None, redirect=None, status=None)) after connection broken by 'NewConnectionError('<pip._vendor.urllib3.connection.Verifie
dHTTPSConnection object at 0x000000843077DAC8>: Failed to establish a new connection: [Errno 11001] getaddrinfo failed')': /simple/sklearn/
  Could not find a version that satisfies the requirement sklearn (from versions: )
No matching distribution found for sklearn

C:\cython\Scripts>pip install sklearn
Collecting sklearn
  Downloading https://files.pythonhosted.org/packages/1e/7a/dbb3be0ce9bd5c8b7e3d87328e79063f8b263b2b1bfa4774cb1147bfcd3f/sklearn-0.0.tar.gz
Collecting scikit-learn (from sklearn)
  Downloading https://files.pythonhosted.org/packages/47/a6/5cfab2684c3da61dece651796aefb2542a1970a8ba9b9efee4f730ccd90b/scikit_learn-0.20.3-cp37-cp37m-win_amd64.whl (4
.8MB)
    100% |████████████████████████████████| 4.8MB 409kB/s
Requirement already satisfied: numpy>=1.8.2 in c:\cython\lib\site-packages (from scikit-learn->sklearn) (1.16.2)
Requirement already satisfied: scipy>=0.13.3 in c:\cython\lib\site-packages (from scikit-learn->sklearn) (1.2.1)
Installing collected packages: scikit-learn, sklearn
  Running setup.py install for sklearn ... done
Successfully installed scikit-learn-0.20.3 sklearn-0.0

C:\cython\Scripts>pip install scipy
Requirement already satisfied: scipy in c:\cython\lib\site-packages (1.2.1)
Requirement already satisfied: numpy>=1.8.2 in c:\cython\lib\site-packages (from scipy) (1.16.2)

C:\cython\Scripts>_
```

Fig 6.3.4 Installation of scipy module

- Installation of Scipy module is which provides convenient and fast N-dimensional array manipulation.
- The Scipy library is built to work with Numpy arrays and provides many user-friendly and efficient numerical practices such as routines for numerical integration and optimization.
- Together, they run on all popular operating systems, are quick to install and are free of charge.
- Numpy and Scipy are easy to use, but powerful enough to depend on by some of the world's leading scientists and engineers.


Python 3.7.2 Shell
File Edit Shell Debug Options Window Help
Ln: 9 Col:
Search the web and Windows
12:13 PM
13-Mar-19

```
Python 3.7.2 (tags/v3.7.2:9a3ffc0492, Dec 23 2018, 23:09:28) [MSC v.1916 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
========== RESTART: C:\Users\ACER\Desktop\project\project19\digits.py ==========
Traceback (most recent call last):
  File "C:\Users\ACER\Desktop\project\project19\digits.py", line 14, in <module>
    img = misc.imread("9.jpg")
AttributeError: module 'scipy.misc' has no attribute 'imread'
>>> 
```

Fig 6.3.5 Attribute error after running program

- When you get an attribute error in Python.
- Numpy and Scipy are easy to use, but powerful enough to depend on by some of the world's leading scientists and engineers.
- Itt means you tried to access the attribute value of, or assign an attribute value to, a Python object or class instance in which that attribute simply does not exist.

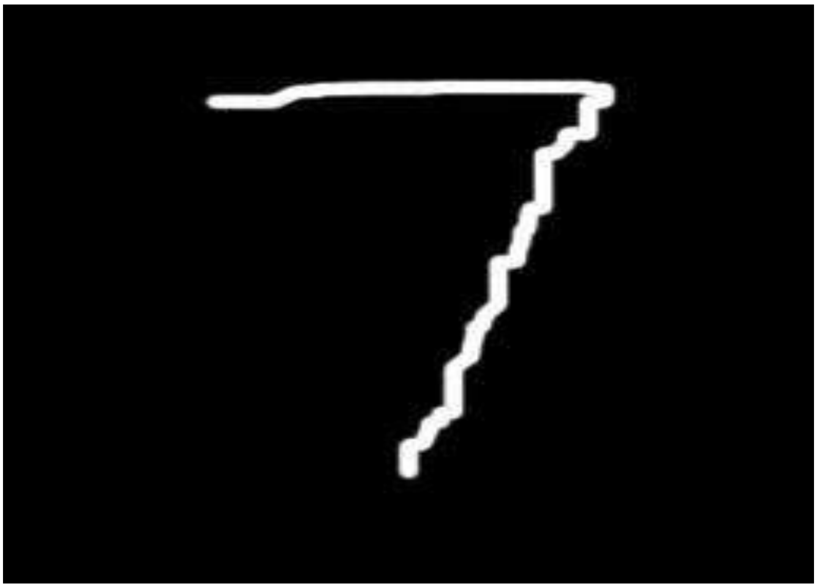
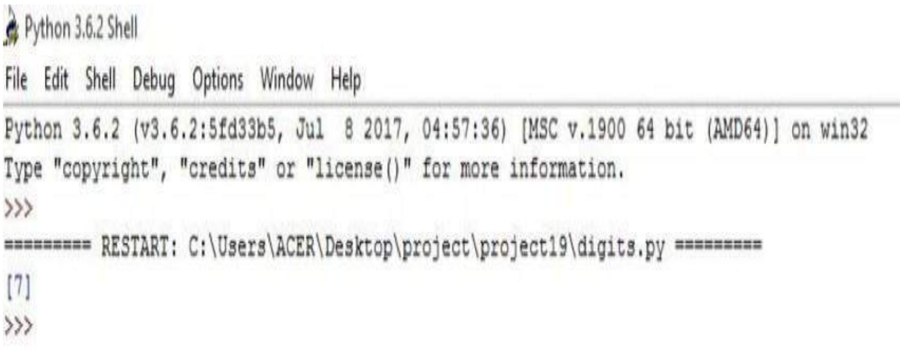


Fig 6.3.6 Input

- Here we give our input image where the hand written digit is written in that image

- That hand written digit 7 is need to be recognized by the machine .
- Then the machine should generate output as digit 7.



```
Python 3.6.2 (v3.6.2:5fd33b5, Jul  8 2017, 04:57:36) [MSC v.1900 64 bit (AMD64)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
========= RESTART: C:\Users\ACER\Desktop\project\project19\digits.py =========
[7]
>>>
```

Fig 6.3.7 Output

- Here in previous step we give image as input.
- In that image there is a hand written digit.
- Finally by using image libraries like scikit-learn,numpy and all it generate result to us on python console screen.

7. CONCLUSION

7. CONCLUSION AND FUTURE SCOPE

7.1 CONCLUSION

However now we are implementing the hand written digit recognition using Sklearn machine learning now we can easily implement such intelligence to identify hand written digits to machines and computers in order to cater our needs in the industrial applications. Previously mat lab was used for such simulations but in such implementations one does not have full control, or the ability to understand that what is happening behind the application.

7.2 FUTURE SCOPE

We have a chance to implement this project to character recognition and we can work on remaining numbers also moreover. We plan to use the same technique to identify signatures for processing check in banking Industry and secondly to develop a face recognition system for HRM Department for student attendance system based on Computer Vision.

8. BIBILOGRAPHY

8. REFERENCES

8.1 REFERENCES

[1] Software engineering Ian Somerville,seventh edition,Pearson education.

[2] The elements of Uml 2.0 martin fowler,Pearson education.

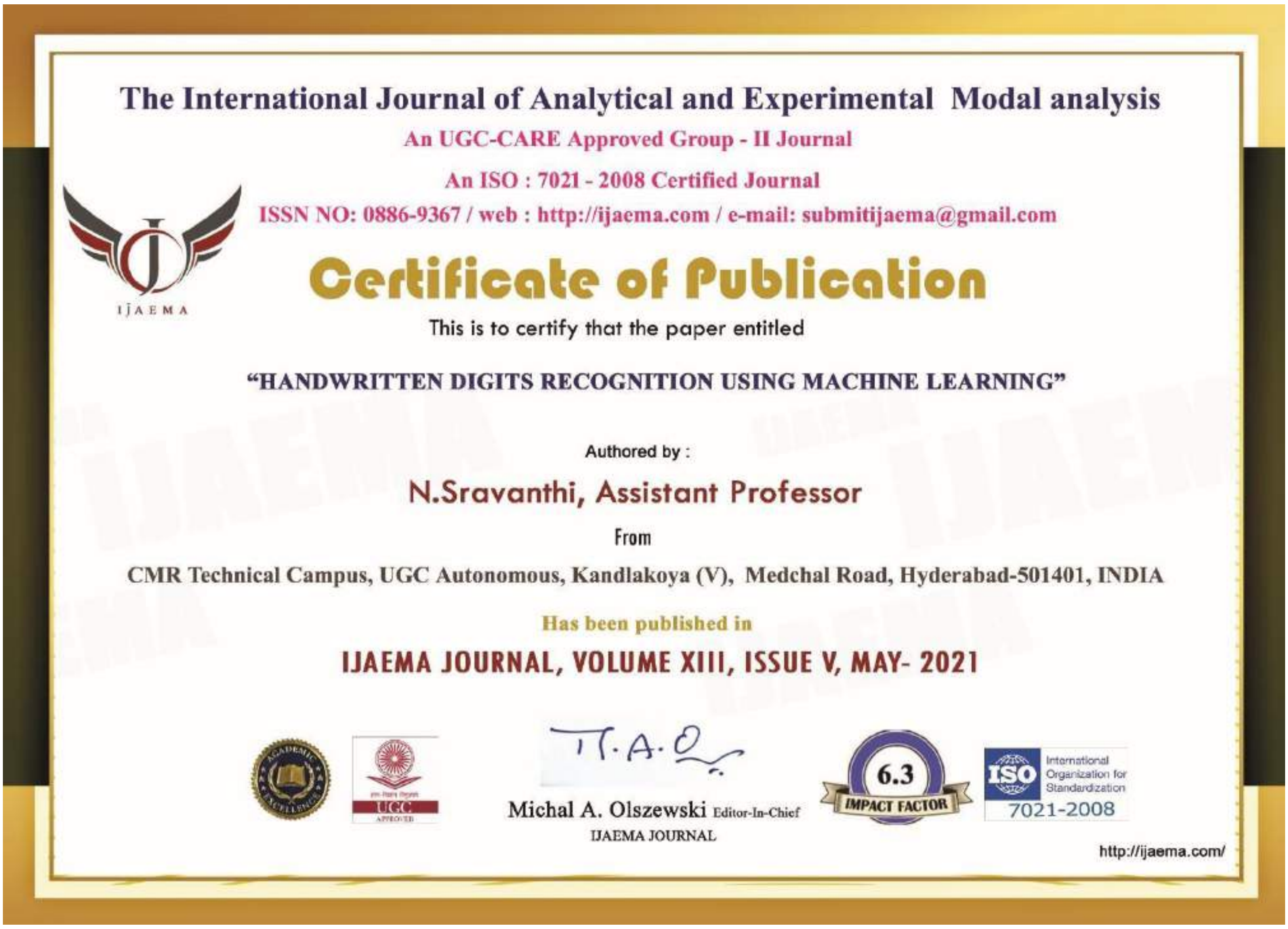
[3] Think Python First Edition, by Allen B.Downey,Orielly publishing.

8.2 WEBSITES

[1] https://en.wikipedia.org/wiki/Handwriting_recognition

[2] https://www.datarobot.com/wiki/data-science/

[3] https://github.com/akshay3134/hand-written-digit-recognition-CMRTC-CSE



The International Journal of Analytical and Experimental Modal analysis

An UGC-CARE Approved Group - II Journal

An ISO : 7021 - 2008 Certified Journal

ISSN NO: 0886-9367 / web : http://ijaema.com / e-mail: submitijaema@gmail.com

IJAEMA

Certificate of Publication

This is to certify that the paper entitled

"HANDWRITTEN DIGITS RECOGNITION USING MACHINE LEARNING"

Authored by :

N.Sravanthi, Assistant Professor

From

CMR Technical Campus, UGC Autonomous, Kandlakoya (V), Medchal Road, Hyderabad-501401, INDIA

Has been published in

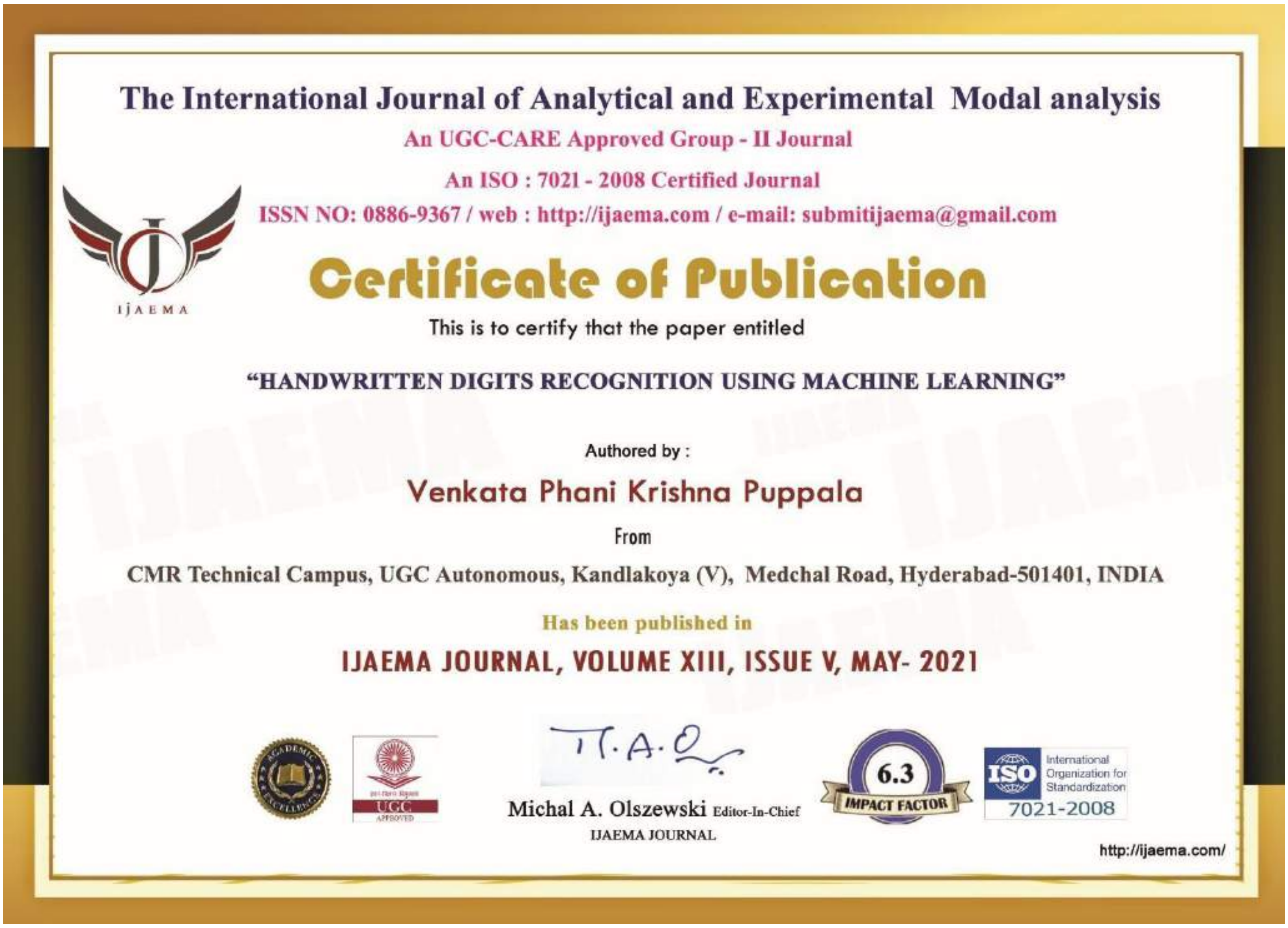
IJAEMA JOURNAL, VOLUME XIII, ISSUE V, MAY- 2021

Michal A. Olszewski Editor-In-Chief
IJAEMA JOURNAL

6.3 IMPACT FACTOR

ISO International Organization for Standardization 7021-2008

http://ijaema.com/



The International Journal of Analytical and Experimental Modal analysis

An UGC-CARE Approved Group - II Journal

An ISO : 7021 - 2008 Certified Journal

ISSN NO: 0886-9367 / web : http://ijaema.com / e-mail: submitijaema@gmail.com

IJAEMA

Certificate of Publication

This is to certify that the paper entitled

"HANDWRITTEN DIGITS RECOGNITION USING MACHINE LEARNING"

Authored by :

Venkata Phani Krishna Puppala

From

CMR Technical Campus, UGC Autonomous, Kandlakoya (V), Medchal Road, Hyderabad-501401, INDIA

Has been published in

IJAEMA JOURNAL, VOLUME XIII, ISSUE V, MAY- 2021

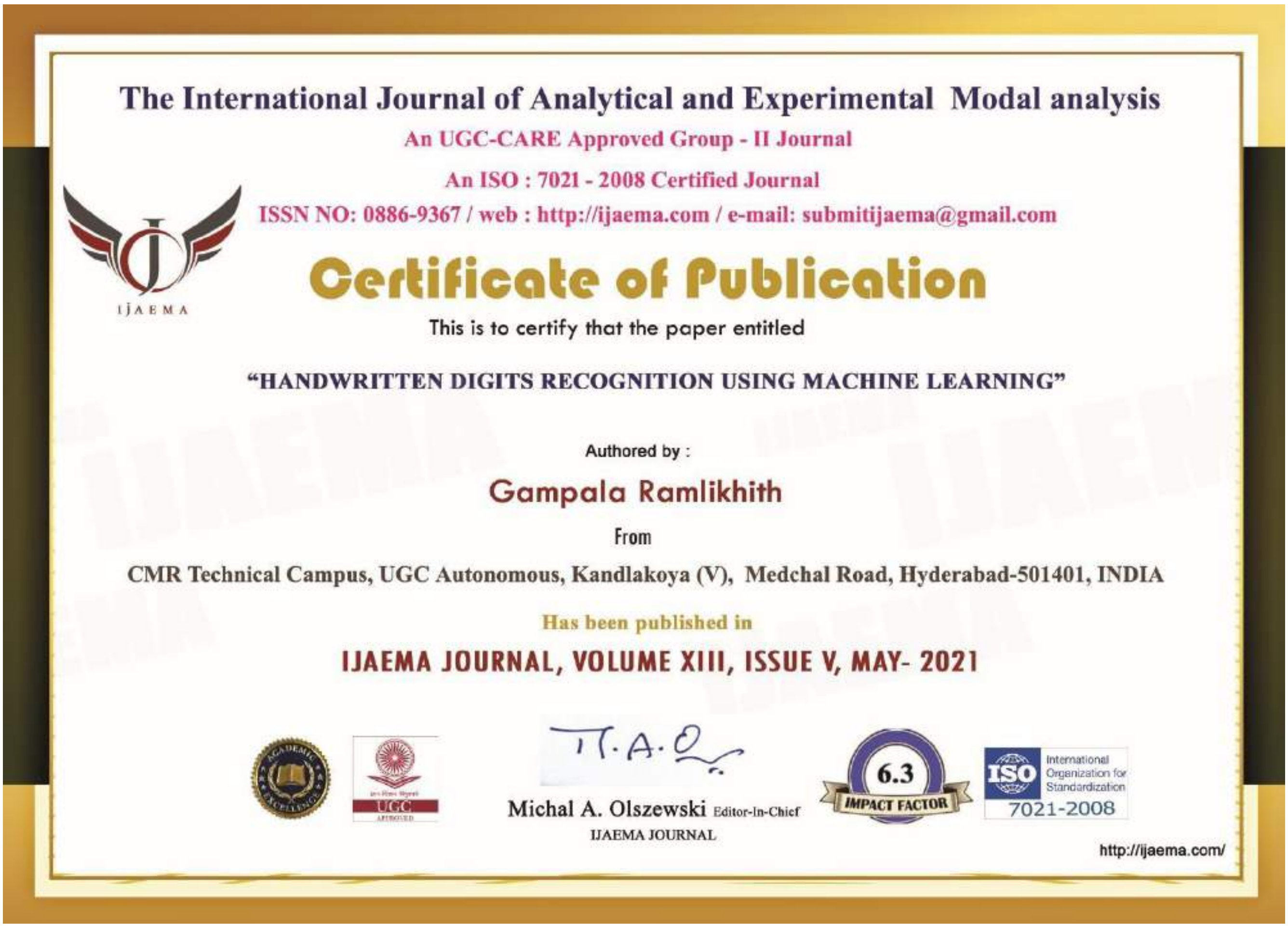
UGC APPROVED

Michal A. Olszewski Editor-In-Chief
IJAEMA JOURNAL

6.3 IMPACT FACTOR

ISO International Organization for Standardization 7021-2008

http://ijaema.com/



The International Journal of Analytical and Experimental Modal analysis

An UGC-CARE Approved Group - II Journal

An ISO : 7021 - 2008 Certified Journal

ISSN NO: 0886-9367 / web : http://ijaema.com / e-mail: submitijaema@gmail.com

IJAEMA

Certificate of Publication

This is to certify that the paper entitled

"HANDWRITTEN DIGITS RECOGNITION USING MACHINE LEARNING"

Authored by :

Gampala Ramlikhith

From

CMR Technical Campus, UGC Autonomous, Kandlakoya (V), Medchal Road, Hyderabad-501401, INDIA

Has been published in

IJAEMA JOURNAL, VOLUME XIII, ISSUE V, MAY- 2021

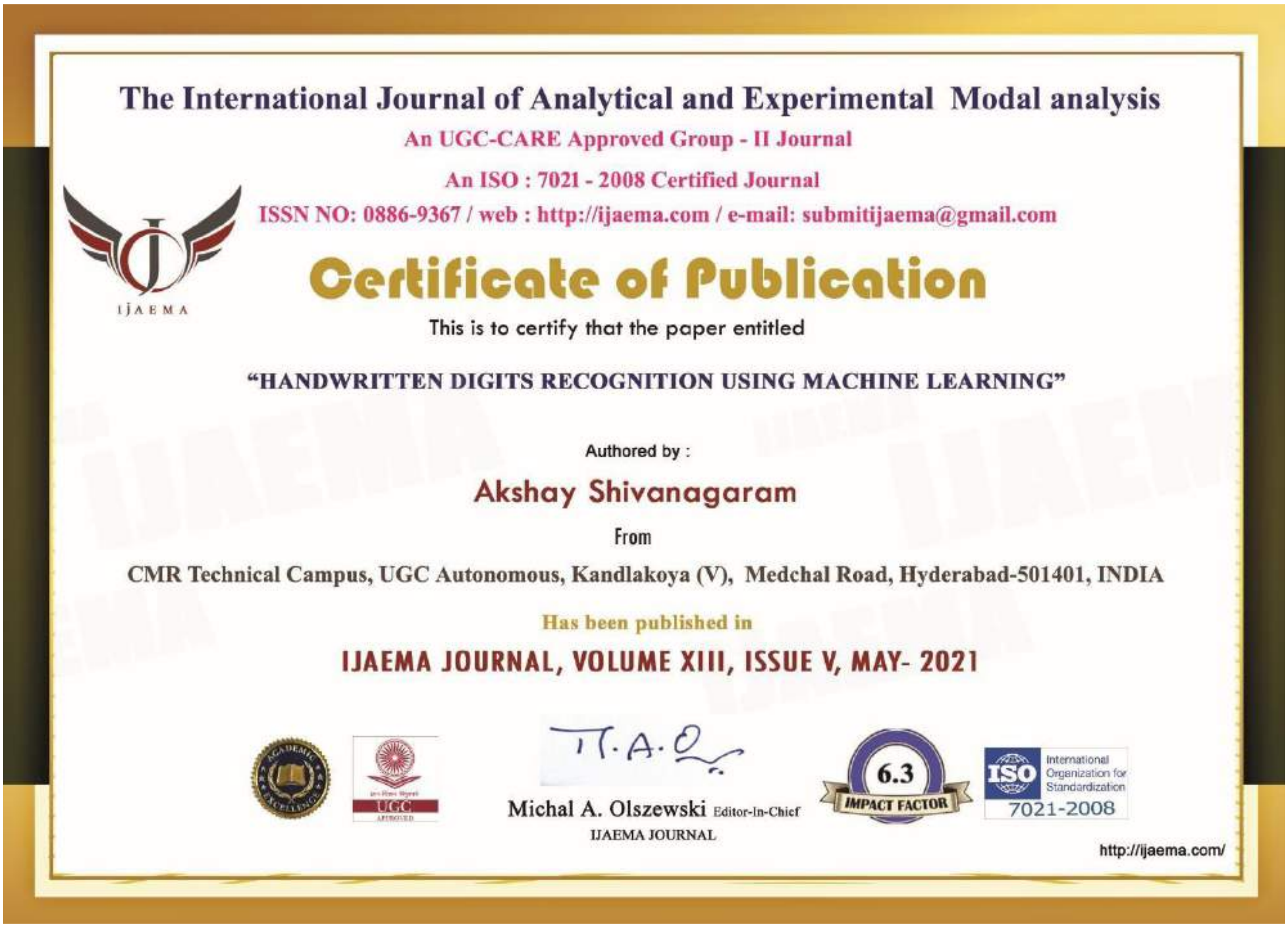
UGC APPROVED

Michal A. Olszewski Editor-In-Chief
IJAEMA JOURNAL

6.3 IMPACT FACTOR

ISO International Organization for Standardization 7021-2008

http://ijaema.com/



The International Journal of Analytical and Experimental Modal analysis

An UGC-CARE Approved Group - II Journal

An ISO : 7021 - 2008 Certified Journal

ISSN NO: 0886-9367 / web : http://ijaema.com / e-mail: submitijaema@gmail.com

IJAEMA

Certificate of Publication

This is to certify that the paper entitled

"HANDWRITTEN DIGITS RECOGNITION USING MACHINE LEARNING"

Authored by :

Akshay Shivanagaram

From

CMR Technical Campus, UGC Autonomous, Kandlakoya (V), Medchal Road, Hyderabad-501401, INDIA

Has been published in

IJAEMA JOURNAL, VOLUME XIII, ISSUE V, MAY- 2021

ACADEMIC EXCELLENCE

UGC APPROVED

Michal A. Olszewski Editor-In-Chief
IJAEMA JOURNAL

6.3 IMPACT FACTOR

ISO International Organization for Standardization 7021-2008

http://ijaema.com/

HANDWRITTEN DIGITS RECOGNITION USING MACHINE LEARNING

Gampala Ramlikhith
UG Scholar, Computer Engineering
Department, CMR Technical Campus,
UGC Autonomous, Kandlakoya (V),
Medchal Road, Hyderabad-501401,
INDIA

Venkata Phani Krishna Puppala,
UG Scholar, Computer Engineering
Department, CMR Technical Campus,
UGC Autonomous, Kandlakoya (V),
Medchal Road, Hyderabad-501401,
INDIA

Akshay Shivanagaram,
UG Scholar, Computer Engineering
Department, CMR Technical Campus,
UGC Autonomous, Kandlakoya (V),
Medchal Road, Hyderabad-501401,
INDIA

N.Sravanthi
Assistant Professor, CMR Technical
Campus, UGC Autonomous,
Kandlakoya (V), Medchal Road,
Hyderabad-501401, INDIA

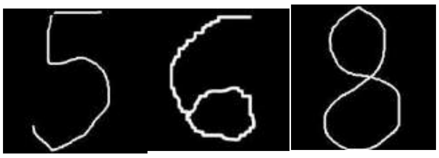
Abstract:

The human visual system is one of the wonders of the world. Consider the following sequence of handwritten digits. Most people effortlessly recognize those digits as 5,6,8. That ease is deceptive. We carry in our heads a supercomputer, tuned by evolution over hundreds of millions of years, and superbly adapted to understand the visual world. Recognizing handwritten digits isn't easy. Rather, we humans are stupendously, astoundingly good at making sense of what our eyes show us. But nearly all that work is done unconsciously. And so we don't usually appreciate how tough a problem our visual systems solve. The difficulty of visual pattern recognition becomes apparent if you attempt to write a computer program to recognize digits like those above. What seems easy when we do it ourselves suddenly becomes extremely difficult. Simple intuitions about how we recognize shapes - "a 9 has a loop at the top, and a vertical stroke in the bottom right" - turn out to be not so simple to express algorithmically. When you try to make such rules precise, you quickly get lost in a morass of exceptions and caveats and special cases. It seems hopeless.

. ***Keywords***—*Hand Written Digits, Machine Learning, Logistic Regression.*

I Introduction

Digit recognition system is the working of a machine to train itself or recognizing the digits from different sources like emails, bank cheque, papers, images, etc. and in different real-world scenarios for online handwriting recognition on computer tablets or system, recognize number plates of vehicles, processing bank cheque amounts, numeric entries in forms filled up by hand and so on.



Handwriting recognition (HWR) is the ability of a computer to receive and handwritten input from sources such as paper documents, photographs, touch screens and other devices. Data science practitioners apply machine learning algorithms to numbers, text, images, video, audio, and more to produce artificial intelligence systems that perform tasks which ordinarily require human intelligence. Scikit learn (formerly scikits learn) is a free software machine learning library for the programming language. It features various classification, regression, and clustering algorithms. Handwritten digit recognition has gained so much popularity from the aspiring beginner of machine learning and deep learning to an expert who has been practicing for years. Developing such a system includes a machine to understand and classify the images of handwritten digits as 10 digits (0–9). Handwritten digits from the MNIST database are already famous among the community for many recent decades. The scope of this article is to compare the different classifiers with different parameters and try to achieve near-human performance.

II. LITERATURE SURVEY

Nowadays, more and more people use images to represent and transmit information. It is also popular to extract important information from images. Image recognition is an important research area for its widely applications[1, 2]. In the relatively young field of computer pattern recognition, one of the challenging tasks is the accurate automated recognition of human handwriting. Indeed, this is precisely a challenging problem because there is a considerable variety in handwriting from person to person. Although, this variance does not cause any problems to humans, yet, however it is more difficult to teach computers to recognize general handwriting [3]. For the image recognition problem such as handwritten classification, it is very important to make out how data are represented in images[1].

The data here is not the row pixels, but should be the features of images which has high level representation[2, 4]. For the problem of handwritten digit recognition, the digit’s structure features should be first extracted from the strokes. Then the extracted features can be used to recognize the handwritten digit. The high performance of large-scale data processing ability is the core technology in the era of big data. Most current classification and regression machine learning methods are shallow learning algorithms [4]. It is difficult to represent complex function effectively, and its generalization ability is limited for complex classification problems[5, 6]. Deep learning is a multilayer neural network learning algorithm which emerged in recent years. Applications of deep learning to various problems have been the subject of a number of recent studies ranging from image classification and speech recognition to audio classification [5, 7-9]. It has brought a new wave to machine learning, and making artificial intelligence and human-computer interaction advance with big strides.

III. PROPOSED METHODOLOGY

At this stage, because all images in the database are clean and without noise, no noise reduction technique is required here. But in a real system we need to remove noise from the images. In any document there could be optical noises present along with the documents. Especially in the handwritten documents the character shapes may not be always unique. Hence the preprocessing is mandatory. We will first apply an

Erosion with 3 X 3 structuring elements which will eliminate the one bit errors and give a smooth edge. Then the characters are dilated with 2 X 2 elements.

Segmentation :

After the preprocessing step, an image of sequence of digit is decomposed into sub-images of individual digit. Preprocessed input image is segmented into isolated digit by assigning a number to each digit using a labeling process. This labeling provides information about number of digits in the image. Each individual digit is uniformly resized into 100 X 70 pixels for classification and recognition stage [16].

Feature Extraction:

After the segmentation step, the Segmented Image is given as input to feature extraction module. The statistical features of the histogram; mean and standard deviation, will be extracted **from** the images.

Training:

After the Feature Extraction step, each of the proposed algorithms (CNN, DBN, DNN) is trained separately with the training images.

Classification & Recognition:

After the training step, “the classification & Recognition stage is the decision making part of a recognition system and it uses the features extracted in the previous stage. A feed forward back propagation neural network having two hidden layers with architecture of 54-100-100-38 is used to perform the classification. The hidden layers use log sigmoid activation function, and the output layer is a competitive layer, as one of the digits is to be identified. The feature vector is denoted as X where X = (f1, f2,…,fd) where f denotes features and d is the number of zones into which each digit is divided. The number of input neurons is determined by length of the feature vector d. The total numbers of digits' n determine the number of neurons in the output layer. The number of neurons in the hidden layers is obtained by trial and error [16]. The most Compact network is chosen and presented as shown in Fig. It is to recognize handwritten digits using the three algorithms in which each algorithm recognizes the image in its own way process. After the training process, the Digits are compared by an expert to assess the accuracy of the tip. Also, the precision, the expense of performance and execution time are compared..

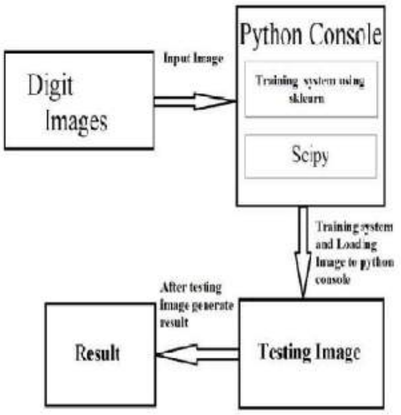

Digit Images
Input Image
Python Console
Training system using sklearn
Scipy
Training system and Loading Image to python console
After testing Image generate result
Result
Testing Image

Fig 1: Project Architecture

The below figure describes the process of hand written digit recognition. Initially digits of image are taken as input and given that image to the python console. Load the data sets using installed sklearn, scipy module in these modules data sets are already available we are just required to import the module and use those modules. Later converting the image into data sets and testing the image. If the generated result is correct considered it as correct. If the system unable to identify then the system 1 as output

IV RESULT ANALYSIS

After preprocessing and feature extraction of our dataset, 80% of the dataset was selected for training and 20% of the dataset was selected for testing. For error calculation, we are using scikit-learn metrics.

Fig 2: Input image to the model

V CONCLUSION

However now we are implementing the hand written digit recognition using Sklearn machine learning now we can easily implement such intelligence to identify hand written digits to machines and computers in order to cater our needs in the industrial applications. Previously mat lab was used for such simulations but in such implementations one does not have full control, or the ability to understand that what is happening behind the application.

VI REFERENCES

[1]. Xiaofeng Han and Yan Li (2015), "The Application of Convolution Neural Networks in Handwritten Numeral Recognition" in International Journal of Database Theory and Application, Vol. 8, No. 3, pp. 367-376.

[2]. Yoshihiro Shima, Meisei, Yumi Nakashima, Michio Yasuda, Meisei (2017), "Pattern Augmentation for Handwritten Digit Classification based on Combination of Pre-trained CNN and SVM", 6th international Conference on informatics, Electronics and vision (ICIEV) and 7th International Symposium n Computational medical and health technology (ISCMHT).

[3]. Xiao-Xiao Niu n , Ching Y. Suen (April 2012), "A novel hybrid CNN–SVM classifier for recognizing handwritten digits", Elsevier, Vol. 45, Issue 4, Pages 1318-1325.

[4]. Caiyun Ma, Hong Zhang (2015), "Effective Handwritten Digit Recognition Based on Multi-feature Extraction and Deep Analysis", 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pp. 297-301.

[5]. Matthew Y.W. Teow Artificial Intelligence Lab (21 October 2017), "Understanding Convolutional Neural Networks Using A Minimal Model for Handwritten Digit Recognition", 2017 IEEE 2nd International Conference on Automatic Control and Intelligent Systems (I2CACIS 2017), Kota Kinabalu, Sabah, Malaysia, pp. 167-172.

[6]. Dan ClaudiuCires¸an, Ueli Meier, Luca Maria Gambardella, Jurgen Schmidhuber (March 2010), ¨Deep, Big, Simple Neural Nets for Handwritten Digit Recognition", arXiv, pp. 1-14.

[7]. Y. Le Cun, L. D. Jackel, B. Boser, J. S. Denker, H. P. Graf, I. Guyon, D. Henderson, R. E. Howard, W. Hub, "Handwritten Digit Recognition : Applications of Neural Network Chips and Automatic Learning" NATO ASI series F: Computer and system sciences, Vol. 68, pp. 41-46.

[8]. T Siva Ajay (July 2017), "Handwritten Digit Recognition Using Convolutional Neural Networks" International Research Journal of Engineering and Technology (IRJET), Vol. 04, Issue 07, pp. 2971-2976.

[9]. Li Deng (November 2012), "The MNIST Database of Handwritten Digit Images for Machine Learning Research", Best of the web series, IEEE signal processing magazine, pp. 141-142.

[10]. Rafael M. O. Cruz, George D. C. Cavalcanti and Tsang Ing Ren (2010), "Handwritten Digit Recognition Using Multiple Feature Extraction Techniques and Classifier Ensemble." 17th Internatonal conference on systems, signals and image processing (IWSSIP), pp. 215-218.

[11] Prasadu Peddi (2019), "AN EFFICIENT ANALYSIS OF STOCKS DATA USING MapReduce", ISSN: 1320-0682, Vol 6, issue 1, pp:22-34.

[12] Prasadu Peddi (2018), Data sharing Privacy in Mobile cloud using AES, ISSN 2319-1953, volume 7, issue 4.