

PROJECT REPORT

**A NOVEL METHOD FOR HANDWRITTEN
DIGIT RECOGNITION SYSTEM**

Submitted By

PNT2022TMID09973

HARIKRISHNA S	720719104067
JICKSON SAM PAUL J	720719104072
MOHANAKRISHNAN S	720719104102
MUTHA J SUBASH SAI	720719104108
SURYA NAGA SRINIVAS	

TABLE OF CONTENTS

INDEX NO	TITLE
1	INTRODUCTION
2	OBJECTIVE
3	IDEATION PHASE 3.1 Literature Survey 3.2 Empathy Map 3.3 Ideation 3.4 Problem Statement
4	PROJECT DESIGN PHASE 1 4.1 Proposed Solution 4.2 Problem Solution Fit 4.3 Solution Architecture
5	PROJECT DESIGN PHASE 2 5.1 Customer Journey Map 5.2 Solution Requirements 5.3 Data Flow Diagrams 5.4 Technology Stack
6	PROJECT PLANNING PHASE 6.1 Prepare Milestone and Activity List 6.2 Sprint Delivery Plan
7	PROJECT DEVELOPMENT
8	CONCLUSION
9	FUTURE SCOPE
10	REFERENCES

1. INTRODUCTION

Handwriting recognition is one of the compelling research works going on because every individual in this world has their own style of writing. It is the capability of the computer to identify and understand handwritten digits or characters automatically. Because of the progress in the field of science and technology, everything is being digitized to reduce human effort. Hence, there comes a need for handwritten digit recognition in many real-time applications. The MNIST data set is widely used for this recognition process and it has 70000 handwritten digits. We use Artificial neural networks to train these images and build a deep learning model. Web application is created where the user can upload an image of a handwritten digit. This image is analyzed by the model and the detected result is returned to the UI.

2. OBJECTIVES

- To utilize the concepts and techniques of the Artificial Neural Network and Convolutional Neural Networks and provide a web application and a novel method for recognizing handwritten digits.
- To solve more complex problems and make humans' jobs easier by implementing handwritten digit recognition systems.
- This type of system can be widely used in the world to recognize zip code or postal code for mail sorting and in the Banking Sector too where more handwritten numbers are involved like account number, figure of cash and checks.
- To make it easy for the Postal department and courier services while finding the digits written.
- To help old people who have eye sight issues with handwritten digits.

3. IDEATION PHASE

3.1 LITERATURE SURVEY

[1]. This project dives into the fundamentals of machine learning victimization associate degree approachable and well-known artificial language, Python. And here we'll be reviewing 2 main components: 1st, we'll be learning regarding the aim of Machine Learning and wherever it applies to the important world. Second, we'll get a general summary of Machine learning topics like supervised learning, model analysis, and Machine Learning algorithms.

[2] Handwritten number recognition is a challenging problem. It is the ability of a computer to receive and interpret intelligible handwriting input from sources such as documents, photographs, touch-screens and other devices. Researchers have been researching in this area in recent years. The main objective of our system is to recognize isolated handwritten digits. Different users have their own handwriting styles. Here the main challenge is to let computer systems understand these different handwriting styles and recognize them. The proposed system deals with such a problem. The system starts by acquiring an image containing handwritten digits, this image is digitized by using techniques like KNN. Numerous recognition techniques are available currently. The idea is to take a large number of handwritten digits, known as training examples and then develop a system which can learn from those training examples. Furthermore, by increasing the number of training examples, the system can learn more on handwritten patterns and improve its accuracy.

[3] In the pattern recognition field, growing interest has been shown in recent years for multiple classifier systems and particularly for bagging, boosting and random sub-spaces. Those methods aim at inducing an ensemble of classifiers by producing diversity at different levels. Following this principle, Breiman introduced in 2001 another family of methods called random forest. Our work aims at studying those methods in a strictly pragmatic approach, in order to provide rules on parameter settings for practitioners. For that purpose we have experimented with the forest-RI algorithm, considered as the random forest reference method, on the MNIST handwritten digits database. In this paper, we describe random forest principles and review some methods proposed in the literature. We present our experimental protocol and results. We finally draw some conclusions on random forest global behavior according to their parameter tuning.

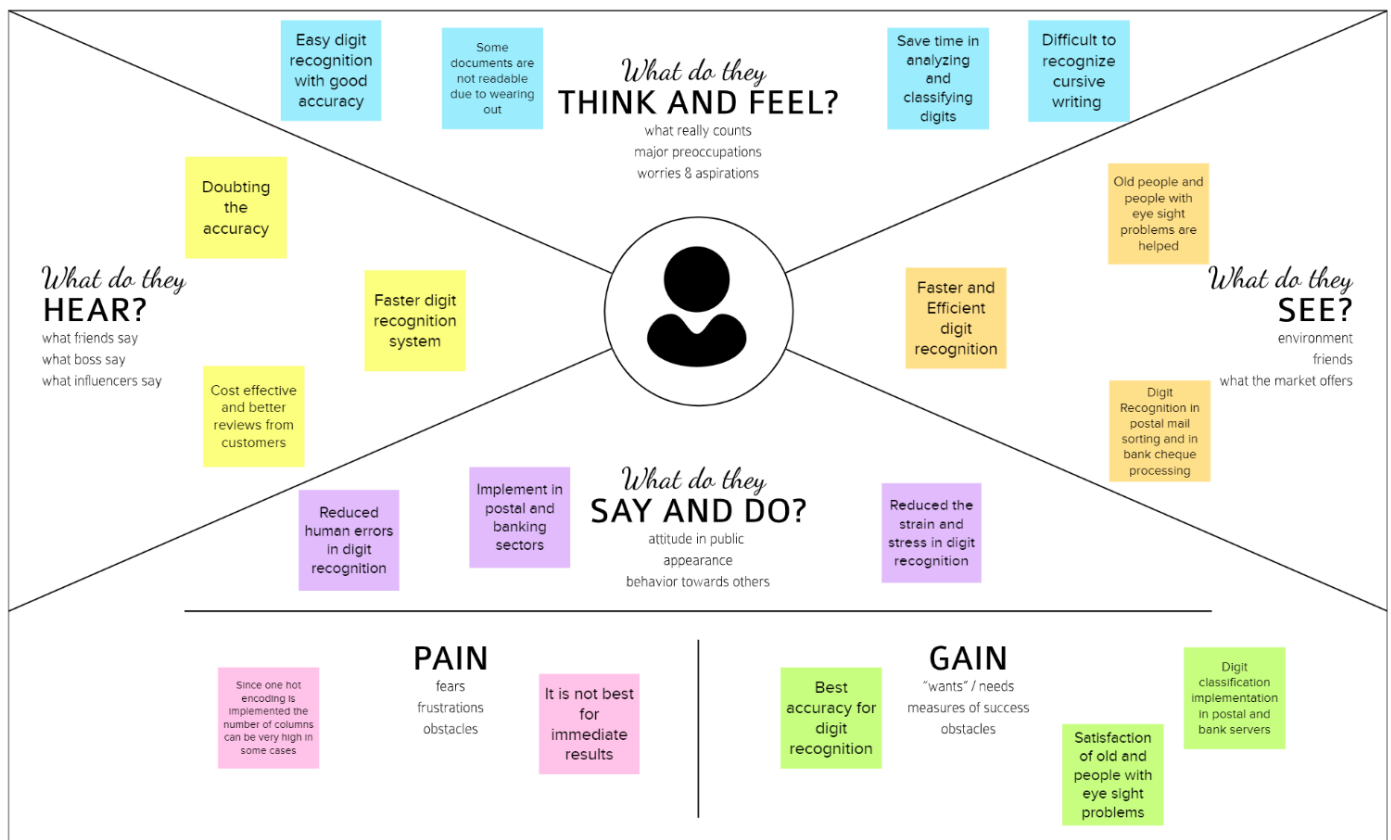
[4] T. Som have discussed fuzzy membership function based approach for HCR. Character images are normalized to 20 X 10 pixels. Average image (fused image) is formed from 10 images of each character. Bounding box around character is determined by using vertical and horizontal projection of character. After cropping image to bounding box, it is resized to 10 X 10 pixels size. After that, thinning is performed and thinned image is placed in one by one row of 100 X 100 canvas. Similarity score of the test image is matched with fusion image and characters are classified.

[5] Renata F. P. Neves has proposed SVM based offline handwritten digit

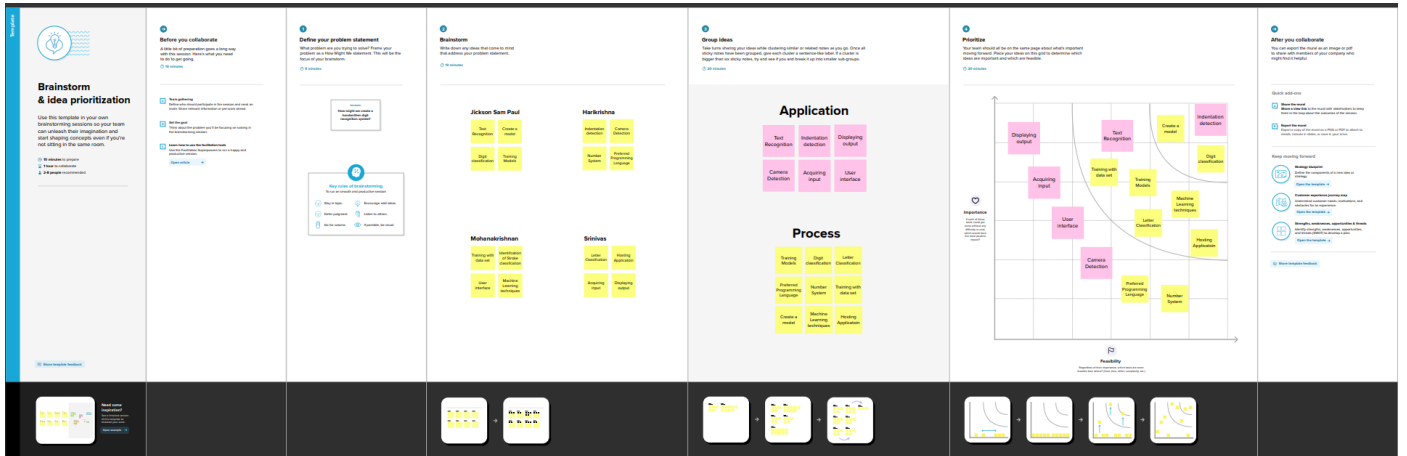
recognition. Authors claim that SVM outperforms the Multilayer perceptron classifier. Experiment is carried out on NIST SD19 standard dataset. Advantage of MLP is that it is able to segment non-linearly separable classes. However, MLP can easily fall into a region of local minimum, where the training will stop assuming it has achieved an optimal point in the error surface. Another hindrance is defining the best network architecture to solve the problem, considering the number

of layers and the number of perceptrons in each hidden layer. Because of these disadvantages, a digit recognizer using the MLP structure may not produce the desired low error rate.

3.2 EMPATHY MAP



3.3 IDEATION



3.4 PROBLEM STATEMENT

Problem Statement of the project : A Novel Method For Handwritten Digit Recognition System

Problem Statement (1) :

I am Mrs.Indira and I am 48 years old. I am working in a bank at the cheque counter. I have difficulty in reading numbers because of my poor eyesight caused by aging. Whenever I am trying to read the digits are too small for me to recognize which makes me feel useless even though I am one of the most appreciated employees in the branch I am working.

Problem Statement (2) :

I am Mr.Indian working in a post office. I have difficulty in

reading pincode because of the different writing styles of people. Every time I try to read, the pincode are unclear for me to recognize which makes it hard to classify the area it needs to be categorized into.

Who does the problem affect?	People who have to Work with handwritten digits
What methodology is used to solve the issue?	Artificial intelligence and neural networks are used to identify and name the digits that can be taken for note.
What solution to solve this issue?	An AI system is introduced to identify different styles of handwritten digits. By using tensor flow, neural networks we create an AI system for recognizing handwritten digits.
Where does the issue occur?	In areas like post offices and business organizations which are in need to recognize handwritten digits in their daily work basis, especially for aged employers it's hard to recognize different digits.

4. PROJECT DESIGN PHASE 1

4.1 PROPOSED SOLUTION

S.NO	PARAMETER	DESCRIPTION
1.	Problem Statement	Recognizing handwritten digits which are very hard to classify due to the difference in people's handwriting and for people with poor eyesight.
2.	Solution	Handwritten digit recognition by using tensorflow and neural networks will solve the stated problem by recognizing handwritten digits with great accuracy.
3.	Novelty	With the help of tensorflow and neural networks this new system is better than the other existing systems which are for recognizing alphabets. This new system can work with better accuracy for recognizing digits.
4.	Social Impact & Customer Benefit	People with eyesight issues find it hard to recognize handwritten digits and can use this system to recognize the handwritten digits correctly. Bank employees can use this system for faster accurate digit recognition. Postal department and courier services can easily find the digits written.
5.	Scalability of the Solution	This system can be developed into a quick and easy to use application. Has the potential for becoming a widely used phone application for digit recognition.

4.2 PROBLEM SOLUTION FIT

Project Title: A Novel Method For Handwritten Digit Recognition System

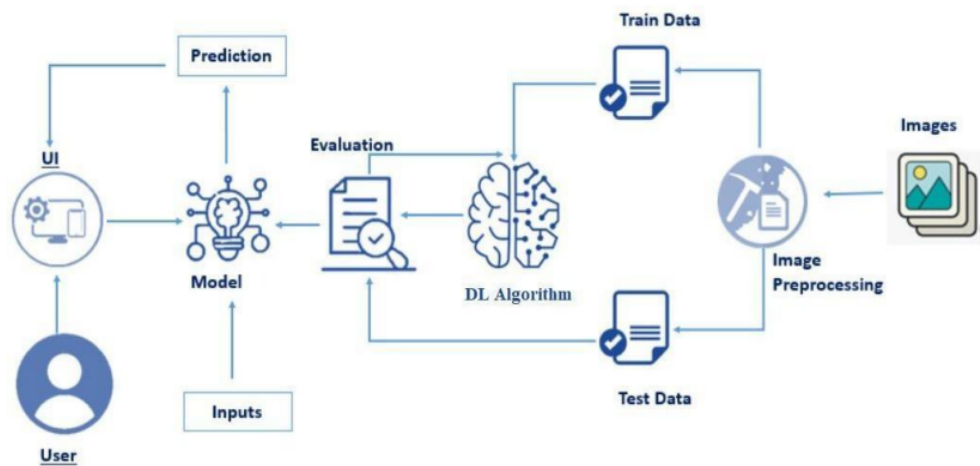
Project Design Phase-I - Solution Fit Template

Team ID: PNT2022TMID09973

Define CS, fit into CC	<div>1. CUSTOMER SEGMENT(S)<div>CS</div></div> <div>2. Used in Banking and Post Office</div> <div>3. Register Office</div> <div>4. Passport and visa Office</div> <div>5. Government Document Verification Office</div> <div>6. Aadhar Sector</div> <div>7. Medical Department</div> <div>8. Old Age People</div>	<div>6. CUSTOMER CONSTRAINTS<div></div></div> <div>1. Time</div> <div>2. Accuracy</div> <div>3. Ease to access</div> <div>4. Imperfect findings</div>	<div>5. AVAILABLE SOLUTIONS<div>AS</div></div> <div>1.In past they get trouble in finding handwritten digits</div> <div>2. Using this system, they can resolve this type of problems</div> <div>3. Pros of this system is quick recognition and Accurate prediction</div> <div>4. Cons are using this system Knowledge about the system is required</div>	Explore AS, differentiate
	<div>2. JOBS-TO BE DONE / PROBLEMS<div>J&P</div></div> <div>Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one;</div> <div>explore different sides. Jobs to be done: To identify the digits in the manually written forms, Cheques filled by people in banks, Phone numbers written manually in register notebook of hospitals. Problems: Dim lighting and weak eyesight</div>	<div>9. PROBLEM ROOT CASUE<div>RC</div></div> <div>What is the real reason that this problem exists? What is the backstory behind the need to do this job?</div> <div>i.e. customers have to do it because of the change in regulations. Handwritten digits are in different fonts and sizes, hard to recognize the digits due to various factors such as dim lighting, weakening eyesight</div>	<div>7. BEHAVIOUR</div> <div>1. Designing the best software that more quickly and accurately identifies the handwritten digits.</div> <div>2. To address the problem, they can take scan copy of the handwritten digit and upload it in the software</div>	

4. SOLUTION ARCHITECTURE

Solution Architecture Diagram for A Novel Method For Handwritten Digit Recognition System :



5. PROJECT DESIGN PHASE 2

5.1 CUSTOMER JOURNEY MAP

<div>1 Phases</div> <div>High-level steps your user needs to accomplish from start to finish</div>	MOTIVATION	GIVE INPUT IMAGES	PROCESSING	OUTPUT
<div>2 Steps</div> <div>Detailed actions your user has to perform</div>	<div>To recognize handwritten digits</div> <div>For recording handwritten digits</div>	<div>Place the handwritten digits near the perceptor</div> <div>Make sure they are placed well</div>	<div>Proceed with recognition option</div> <div>Wait until the process is complete</div>	<div>See the recognized digit</div>
<div>3 Feelings</div> <div>What your user might be thinking and feeling at the moment</div>	<div>Eager</div> <div>Frustration</div>	<div>Happy to place it correctly</div> <div>Confusion</div>	<div>Excited to see result</div> <div>Worry</div>	<div>Happy and satisfied to see the result</div> <div>Not happy if the result is not expected</div>
<div>4 Pain points</div> <div>Problems your user runs into</div>	<div>Being stressed because of his/her inability to recognize digits</div>	<div>Worried if they placed it well</div>	<div>Thinking that the process might fail</div>	<div>If the digits are recognized correctly or not</div>
<div>5 Opportunities</div> <div>Potential improvements or enhancements to the experience</div>	<div>Making it available for many</div>	<div>Making it more easy to use</div>	<div>Recognize the digits much faster</div>	<div>Making it more accurate and precise</div>

5.2 SOLUTION REQUIREMENTS

Project Design Phase-II Solution Requirements (Functional & Non-functional)

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Gmail/Messengers.
FR-2	User Login	Login through registered Username and Password.
FR-3	Uploading Files	Input the handwritten images into the application.
FR-4	Input Perception	Recognize characters from images , collect data and prepare it for training.
FR-5	Feature extraction	Analyzing the images and deriving characteristics from these images that identify each specific element.
FR-6	Recognizing digits	Display the recognized digits to the user.

Non-functional Requirements:

Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	The software must be smooth and user friendly so that the users can use it efficiently with less learning to do. To recognize the digits from bank cheque,papers,entry in forms etc.
NFR-2	Security	Provide security by authentication system in which the users must login with their username and password.
NFR-3	Reliability	Application will be more reliable in use than the previous methods.
NFR-4	Performance	Has to recognize messy writings and provide output efficiently. The Application can provide almost precise accuracy.

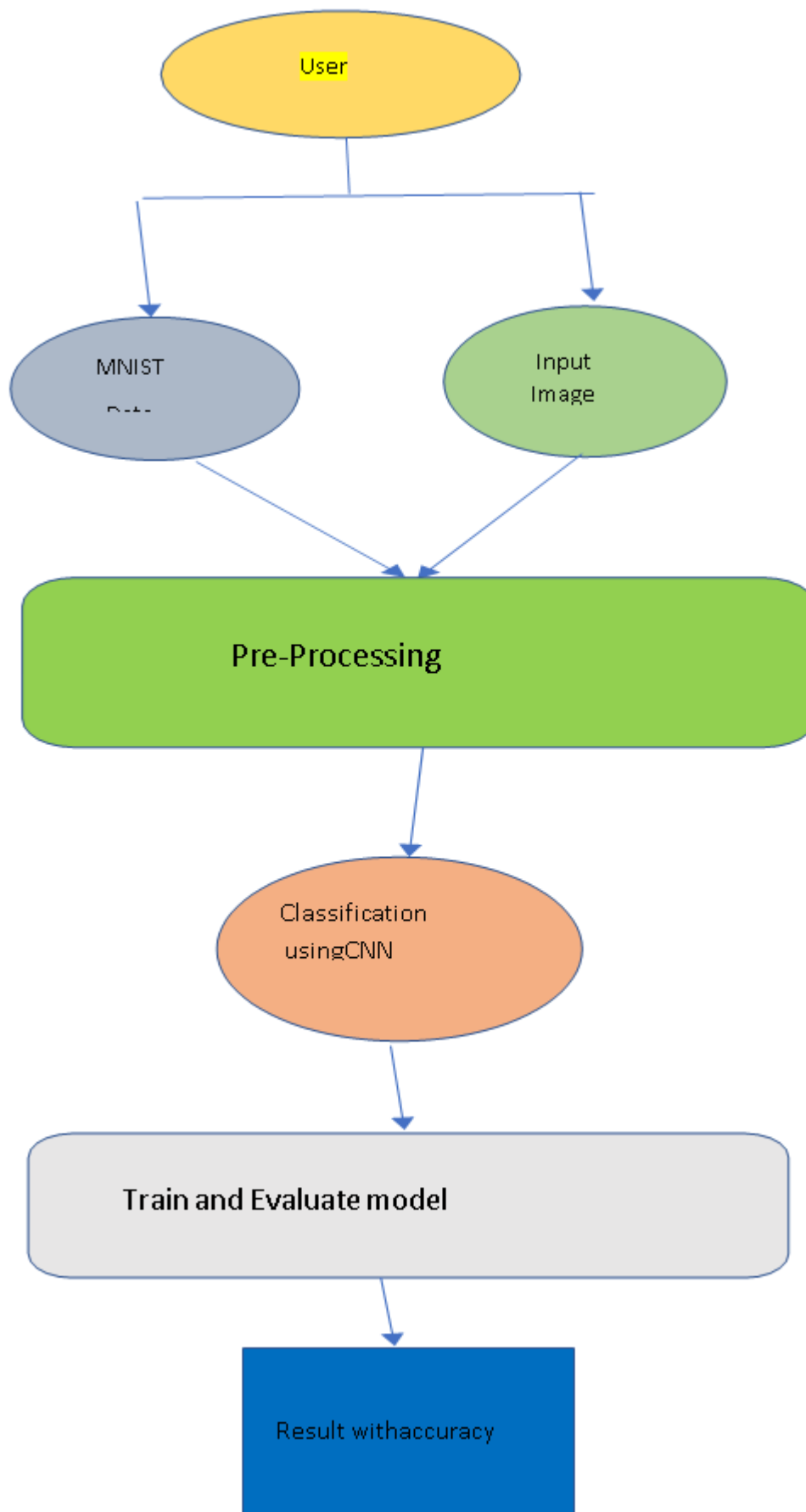
NFR-5	Availability	Has to be available to the users from anywhere to give the handwritten images as input.
NFR-6	Scalability	It can handle many users at the same time providing faster response and better recognition.

5.3 DATA FLOW DIAGRAMS

Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

Example:



User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer	Dashboard	USN-1	As a user, they can see the information regarding the prediction of handwritten digit recognition.	I can see the information regarding digit recognition.	High	Sprint 1
	Launch	USN-2	On clicking the launch button, it will redirect the user to a page where the images to be predicted can be uploaded.	I can see the launch button.	High	Sprint 1
	Upload	USN-3	Users can select the image from the local storage.	I can upload the image.	High	Sprint 2
	Predict	USN-4	Once the image is uploaded, it will predict the respective		High	Sprint 3

			image.			
	Display	USN-5	The predicted image will be displayed with the accuracy chart.	I can see the result with accuracy.	High	Sprint 4

5.4 TECHNOLOGY STACK

Technical Architecture:

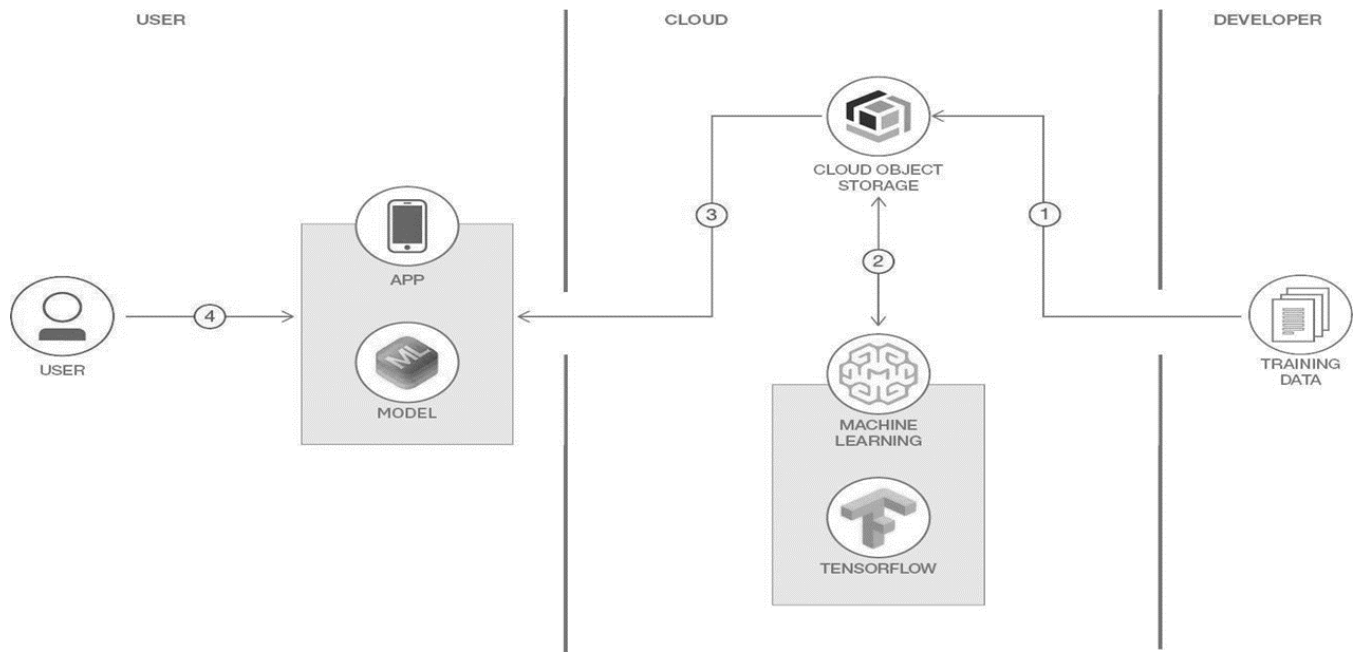


Table-1 : Components & Technologies:

S.No	Component	Technology
1	User Interface	HTML, CSS, JavaScript / React Js etc.
2	Application Logic-1	Python (ML & DL Algorithms)
3	Database	MySQL, NoSQL, etc.
4	Cloud Database	IBM DB2, IBM Cloudant etc.
5	File Storage	IBM Block Storage or Other StorageService or Local File System

6 PROJECT PLANNING PHASE

6.1 MILESTONE & ACTIVITY LIST

Milestone:

Modern Technology is increasing and optimizing the Performance of the Artificial Intelligences (AI) Model Based A Novel Method for Handwritten Digit Recognition System, is helpful for post office workers ,data digitization officers etc..to prevent the typing of data manually which can identify the Digits within a process of capturing the Image on the papers and Machine Learning Algorithm will give digit value. In this Project Milestone will be given the Best Solution for the data digitizers using the complete friendly and simple user interface web application to fetch the solution on its own. In addition, we are planning to add alphabets and also special characters in future. It can completely convert handwritten data into digital data format.

Activity List:

In Project Management Planning is an Important task to scheduling the phrase of the project to the Team Member. In this Activity can shows the various activity are allocated and Done by the Team Members! In Project we can Split into the Four Step of Phrases are

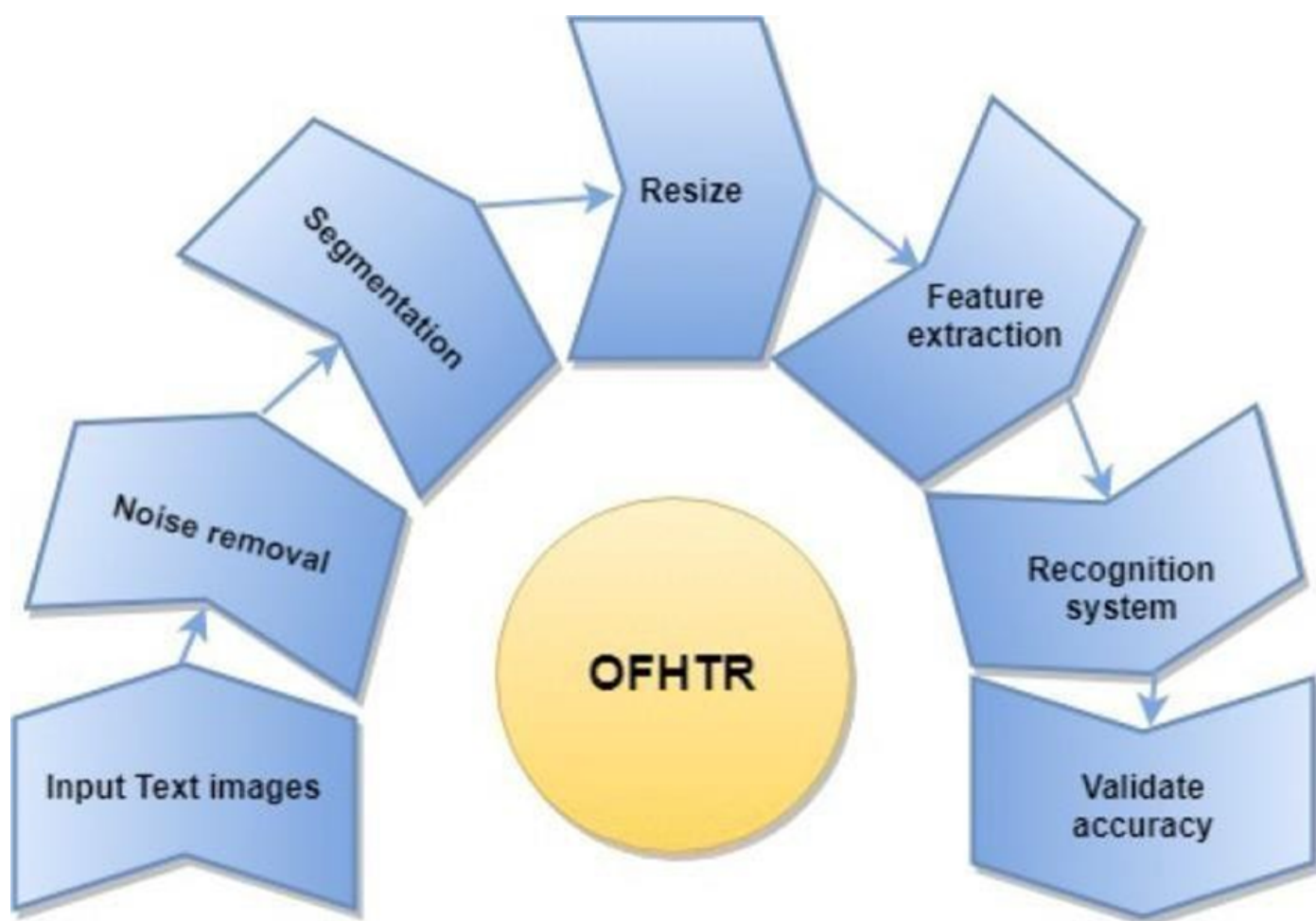
Phase 1: Information Collection and Requirement Analysis.

Phase 2 : Project Planning and Developing Modules.

Phase 3 : Implementing the High Accuracy Deep Learning Algorithm to Perform.

Phase 4 : Deploying the Model on Cloud and Testing the Model and UI Performance





6.2 SPRINT DELIVERY PLAN

PROJECT TRACKER:

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (On Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	5 Days	1 Nov 2022	05 Oct 2022	20	05 Oct 2022
Sprint-2	20	5 Days	06 Nov 2022	10 Nov 2022	20	10 Nov2022
Sprint-3	20	5 Days	11 Nov 2022	15 Nov 2022	20	15 Nov 2022
Sprint-4	20	5 Days	16 Nov 2022	20 Nov 2022	20	20 Nov 2022

VELOCITY: SPRINT - 1

Sprint duration = 5 days

Velocity of team = 20 points

Average Velocity (AV) =
$$\frac{\text{Velocity}}{\text{Sprint duration}}$$

$$AV = 20/5 = 4$$

Average Velocity = 4

VELOCITY: Sprint 1 - 4

Sprint duration = 20 days

Velocity of team = 80 points

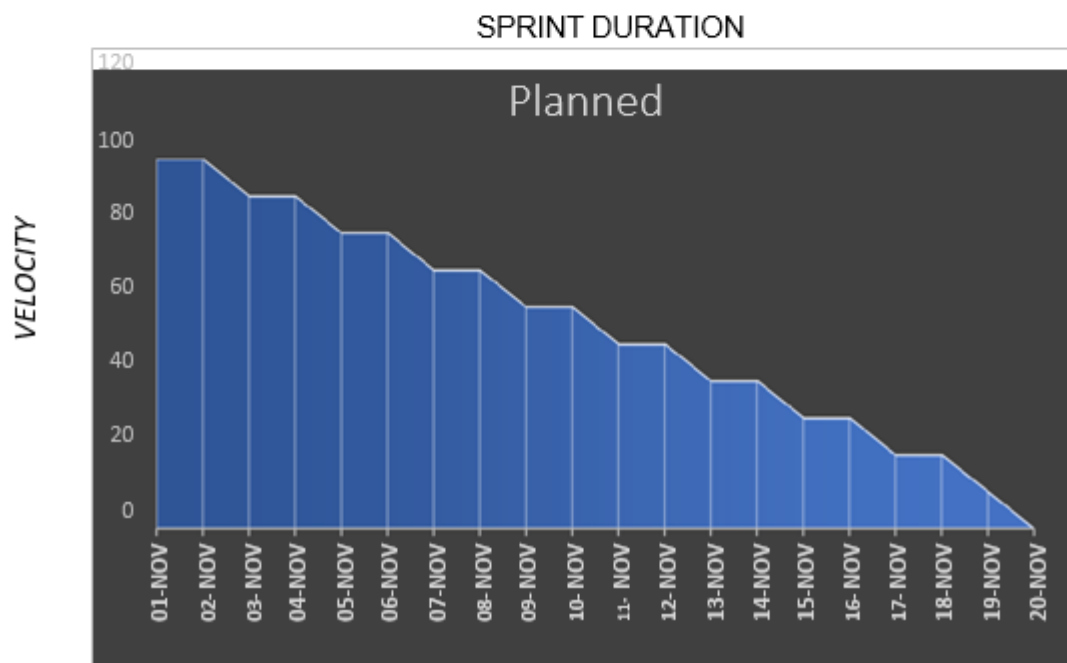
$$\text{Average Velocity (AV)} = \frac{\text{Velocity}}{\text{Sprint duration}}$$

$$AV = 80/20 = 4$$

Total Average Velocity = 4

BURNDOWN CHART:

- A *Burn Down Chart* plots the amount of work remaining to perform against the amount of time. In agile software development approaches like Scrum, it is frequently employed.
- However, burn down charts can be applied to any project containing measurable progress over time.



7. PROJECT DEVELOPMENT

```
import numpy as np
import os
from PIL import Image
from flask import Flask, request, render_template, url_for
from werkzeug.utils import secure_filename, redirect
from gevent.pywsgi import WSGIServer
from keras.models import load_model
from keras.preprocessing import image
from flask import send_from_directory
UPLOAD_FOLDER = 'D:/ibm/data'
app = Flask(__name__)
app.config['UPLOAD_FOLDER'] = UPLOAD_FOLDER
model = load_model("./models/mnistCNN.h5")
@app.route('/')
def index():
    return render_template('index.html')
@app.route('/predict', methods=['GET', 'POST'])
def upload():
    if request.method == "POST":
        f = request.files["image"]
        filepath = secure_filename(f.filename)
        f.save(os.path.join(app.config['UPLOAD_FOLDER'], filepath))

        upload_img = os.path.join(UPLOAD_FOLDER, filepath)
        img = Image.open(upload_img).convert("L")
        img = img.resize((28, 28))

        im2arr = np.array(img)
        im2arr = im2arr.reshape(1, 28, 28, 1)

        pred = model.predict(im2arr)

        num = np.argmax(pred, axis=1)

        return render_template('predict.html', num=str(num[0]))
if __name__ == '__main__':
    app.run(debug=True, threaded=False)
```

8. CONCLUSION

This project demonstrated a web application that uses machine learning to recognize handwritten numbers. Flask, HTML, CSS, JavaScript, and a few other technologies were used to create this project. The model predicts the handwritten digit using a CNN network. During testing, the model achieved a 99.61% recognition rate. The proposed project is scalable and can easily handle a huge number of users. Since it is a web application, it is compatible with any device that can run a browser. This project is extremely useful in real- world scenarios such as recognizing number plates of vehicles, processing bank cheque amounts, numeric entries in forms filled up by hand (tax forms) and so on. There is so much room for improvement, which can be implemented in subsequent versions.

9. FUTURE SCOPE

This project is far from complete and there is a lot of room for improvement. Some of the improvements that can be made to this project are as follows:

- ✓ Add support to detect from digits multiple images and save the results
- ✓
- ✓ Add support to detect multiple digits
- ✓
- ✓ Improve model to detect digits from complex images
- ✓ Add support to different languages to help users from all over the world This project has endless potential and can always be enhanced to become better.

Implementing this concept in the real world will benefit several industries and reduce the workload on many workers, enhancing overall work efficiency.

10. REFERENCES

- [1] Saiteja, Jeisetti & Rao, Podila & Bharadwaj(2022). HANDWRITTEN DIGIT RECOGNITION. International Journal of Computer Science and Mobile Computing. 11. 45-54. 10.47760/ijcsmc.2022.v11i01.007
- [2] Kumar, A & Kumar, S & Kumar, J & Shri, S & Narayanan, Geetha. (2021). HANDWRITTEN DIGIT RECOGNITION. Xi'an Jianzhu Keji Daxue Xuebao/Journal of Xi'an University of Architecture & Technology. XII. 3136-3143.
- [3] S. Bernard Laboratoire LITIS, EA 4108 UFR des Sciences, Université de Rouen, France S. Adam Laboratoire LITIS, EA 4108 UFR des Sciences, Université de Rouen, France L. Heutte Laboratoire LITIS, EA 4108 UFR des Sciences, Université de Rouen, France
- [4] T.Som, Sumit Saha, "Handwritten Character Recognition Using Fuzzy Membership Function", International Journal of Emerging Technologies in Sciences and Engineering, Vol.5, No.2, pp. 11-15, Dec 2011.
- [5] Renata F. P. Neves, Alberto N. G. Lopes Filho, Carlos A.B.Mello, CleberZanchettin, "A SVM Based Off-Line Handwritten Digit Recognizer", International conference on Systems, Man and Cybernetics, IEEE Xplore, pp. 510-515, 9-12 Oct, 2011, Brazil