

## Digital Naturalist - AI Enabled tool for Biodiversity Researchers

### A PROJECT REPORT

***Submitted by***

|  |  |
| --- | --- |
| **Subasri N** | **822419106028** |
| **Aarthi J** | **822419106002** |
| **Janani R** | **822419106016** |
| **Abirami J** | **822419106004** |

### Team ID-PNT2022TMID47172

***for the course***

### HX8001 – Professional Readiness for Innovation, Employability and Entrepreneurship

***in***

**INFORMATION TECHNOLOGY**

## MRK Institute Of Technology-Cuddalore

**ANNA UNIVERSITY : CHENNAI 600 025**

## November 2022

ANNA UNIVERSITY : CHENNAI 600 025

# BONAFIDE CERTIFICATE

Certified that this project report **“Digital Naturalist - AI Enabled tool for Biodiversity Researchers”** is the bonafide work of **“Subasri N, Aarthi J, Abirami J ,Janani R”** who carried out the project work under my supervision. Our Team ID- **PNT2022TMID47172.**

|  |  |
| --- | --- |
| Mr.S.Sathyamoorthy | Mrs.T.Karthiga |
| **HEAD OF THE**  **DEPARTMENT** | **MENTOR** |
|  |  |
|  |  |
|  |  |

# TABLE OF CONTENTS

|  |  |  |
| --- | --- | --- |
| **CHAPTER NO** | **TITLE** | **PAGE NO** |
| **1** | **INTRODUCTION** |  |
|  | 1.1 Project Overview | 4 |
|  | 1.2 Purposeb | 4 |
| **2** | **LITERATURE SURVEY** |  |
|  | 2.1Existing problem | 5 |
|  | 2.2 References | 5 |
|  | 2.3 Problem Statement Definition | 6 |
| **3** | **IDEATION & PROPOSED SOLUTION** |  |
|  | 3.1Empathy Map Canvas | 7 |
|  | 3.2 Ideation & Brainstorming | 8 |
|  | 3.3 Proposed Solution | 11 |
|  | 3.4 Problem Solution fit | 12 |
| **4** | **REQUIREMENT ANALYSIS** |  |
|  | 4.1 Functional requirement | 13 |
|  | 4.2 Non-Functional requirements | 13 |
| **5** | **PROJECT DESIGN** |  |
|  | 5.1 Data Flow Diagrams | 14 |
|  | 5.2 Solution & Technical Architecture | 15 |
|  | 5.3 User Stories | 16 |
| **6** | **PROJECT PLANNING & SCHEDULING** |  |
|  | 6.1 Sprint Planning & Estimation | 17 |
|  | 6.2 Sprint Delivery Schedule | 19 |
|  | 6.3 Reports from JIRA | 20 |
| **7** | **CODING & SOLUTIONING** |  |
|  | 7.1 Feature 1 | 21 |

|  |  |  |
| --- | --- | --- |
|  | 7.2 Feature 2 | 21 |
| **8** | **RESULTS** |  |
|  | 8.1 Performance Metrics | 22 |
| **9** | **ADVANTAGES & DISADVANTAGES** | 25 |
| **10** | **CONCLUSION** | 26 |
| **11** | **FUTURE SCOPE** | 27 |
| **12** | **APPENDIX** |  |
|  | Source Code | 28 |
|  | GitHub & Project Demo Link | 33 |

1. **INTRODUCTION**

### Project Overview

A naturalist is someone who studies the patterns of nature, identifies a different kind of flora and fauna in nature. Being able to identify the flora and fauna around us often leads to an interest in protecting wild spaces, and collecting and sharing information about the species we see on our travels is very useful forconservation groups like NCC.

When venturing into the woods, field naturalists usually rely on common approaches like always carrying a guidebook around everywhere or seeking help g experienced ornithologists. There should be a handy toolfor them to capture, identify and share the beauty to the outside world.

In this project, we are creating a web application which uses a deep learning model, trained on different species of birds, flowers and mammals (2 subclasses in each for a quick understanding) and get theprediction of the bird when an image is been given.

### Purpose

* + - Augmenting a dataset to virtually increase the size of small datasets in order to make our machinelearning models work better.
    - Preprocess the images to a machine-readable format.
    - Applying CNN algorithm on the dataset.
    - How deep neural networks are predicting the class and subclass of a given image.
    - You will be able to know how to find the accuracy of the model.
    - You will be able to build web applications using the Flask framework.

# LITERATURE SURVEY

### Existing problem

Biases in our data arise in part from differences between the aims of the original data collectors (i.e., the photographers) and our aims as biodiversity researchers and ecologists. For example, the spatial distribution of our images was biased toward areas where extensive managed gardens or other displays exhibited large collections of flowering plants. These biases could be addressed by choosing alternative sources, changing the search terms used, or pre-filtering images. Images may also be biased taxonomically or in terms of certain traits, for example, toward species that are typically considered more photogenic due tolarge colorful flowers or leaves. Search terms could be modified to either focus on a specific sub-group, e.g., searching using scientific names, or to exclude non-target images, e.g., excluding images that include the words “show” or “garden” in their metadata. Finally, high-level image classifiers could be trained to remove images that are clearly not plants, e.g: removing images of animals, paintings. High-level classifiers developed to separate images that contain plants from those that do not, without looking to identify species, could be used to find images worthy of further examination in large datasets that do not have metadata (such as titles and descriptions), removing the need for keyword searches, such as that used in this study.

### References

1. **Aldhebiani AY (2018) Species concept and speciation. Saudi J Biol Sci 25:437–440.**
2. **AI naturalists might hold the key to unlocking biodiversity data in social media imagery, TA August, OL Pescott, A Joly, P Bonnet - Patterns, 2020 - Elsevier.**
3. **Digitalization to achieve sustainable development goals: Steps towards a Smart Green Planet, ME Mondejar, R Avtar, HLB Diaz, RK Dubey… - Science of the Total …, 2021 - Elsevier**
4. **The real-world use of big data, M Schroeck, R Shockley, J Smart, D Romero-Morales…**

**- IBM Global Business …, 2012**

* 1. **Problem Statement Definition**

# Main Problem statement (common):

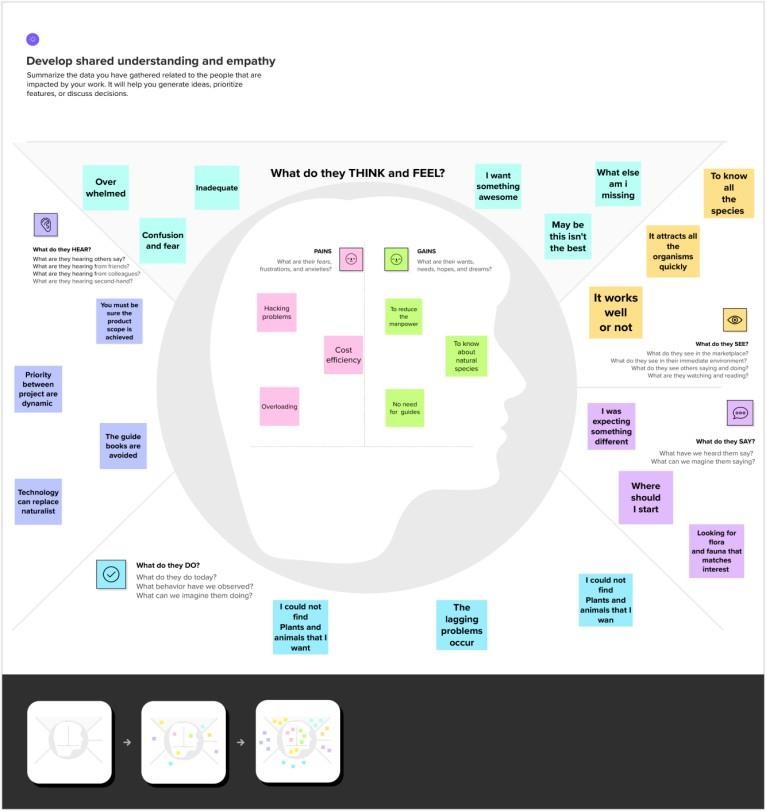
I. How might we help both experienced and inexperienced user to identify species of plants andanimals and their characteristics with related information?

### Specific problem statement:

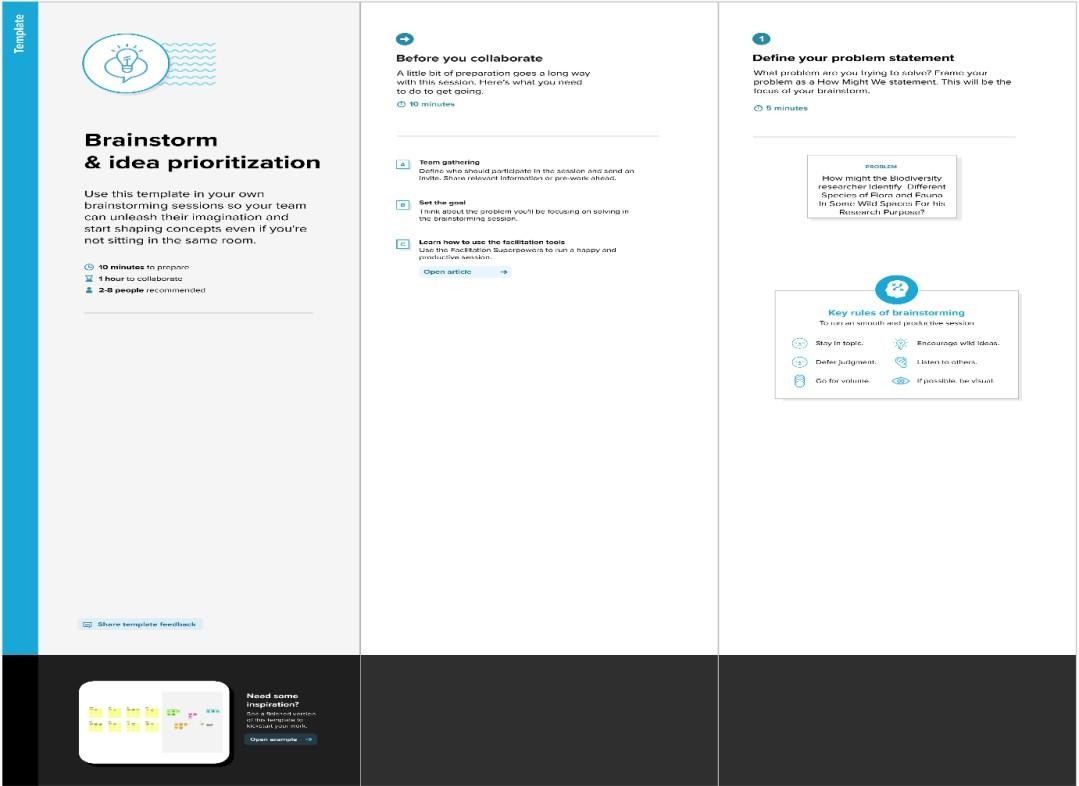
1. Inexperienced users need to know about poisonous plants and dangerous animals so that they canstay away from it.
2. Both experienced and inexperienced users need to know about the medicinal values of a plantbecause they need to use it in case of emergencies.
3. All the users need to know the types of species of birds, plants and animals so that they can learnabout it in more detailed manner.
4. All the users need to know about the rarity of the species of birds, animals or plants so that they canpreserve and save it.

# IDEATION & PROPOSED SOLUTION

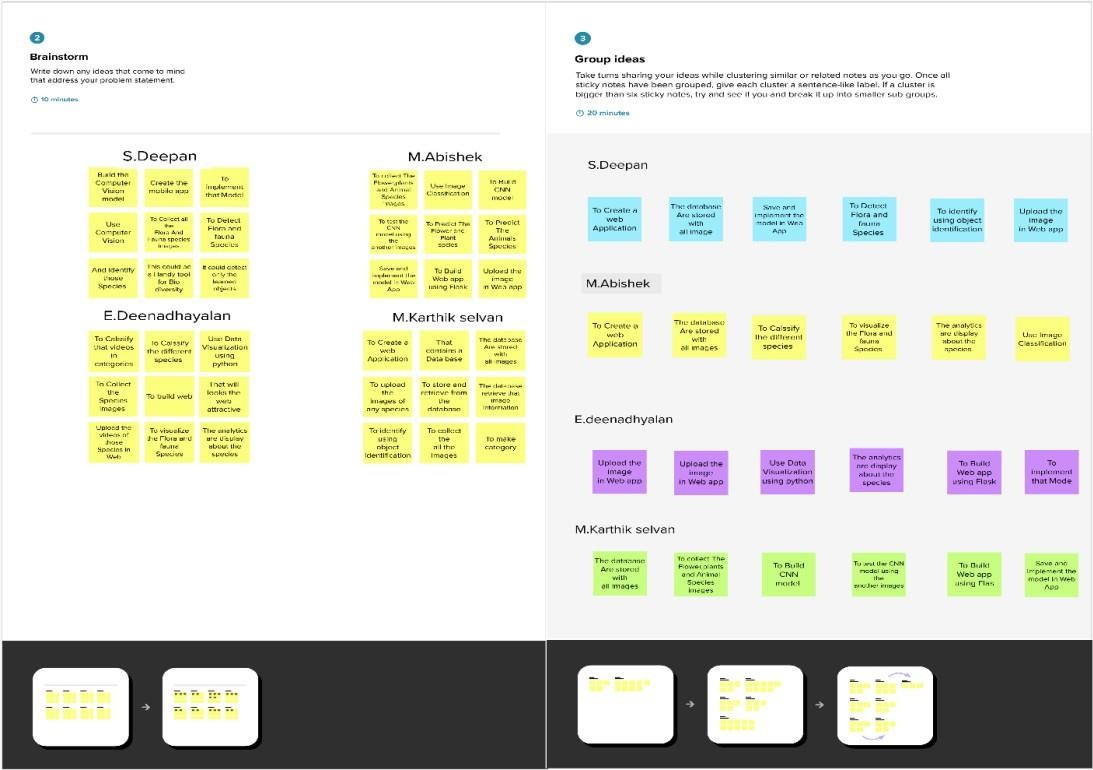
### Empathy Map Canvas



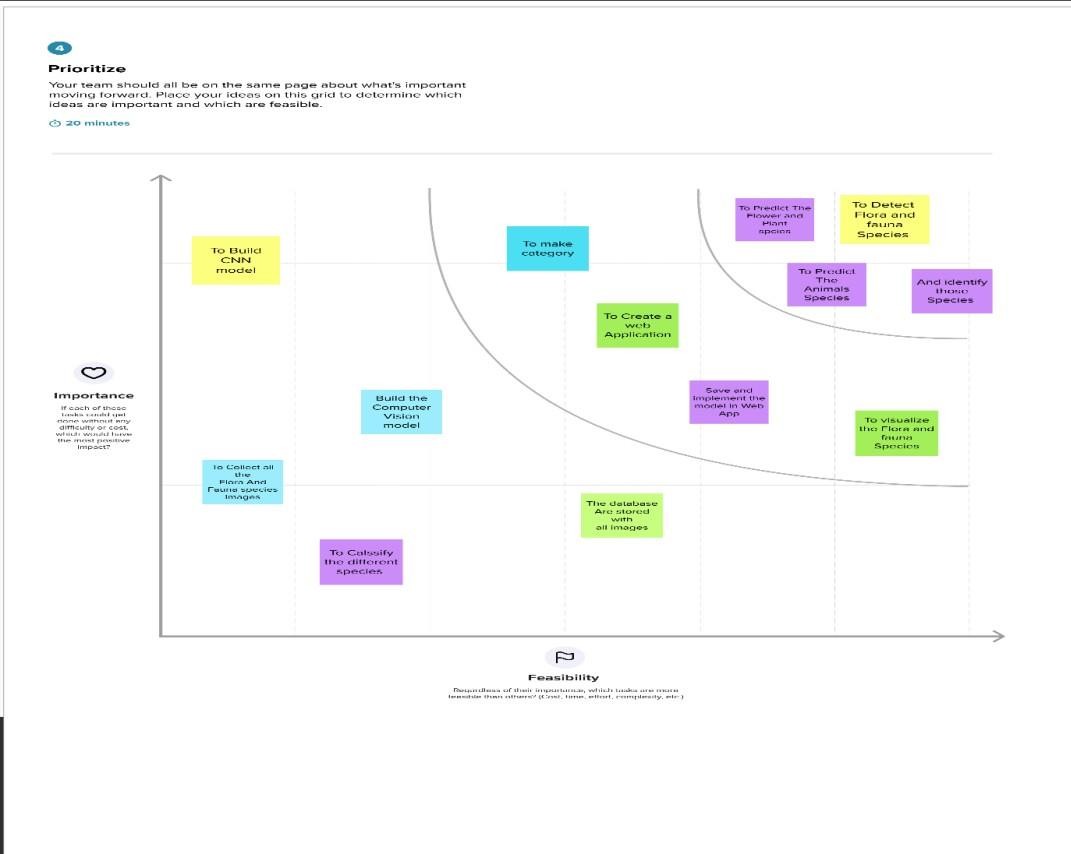
* 1. **Ideation & Brainstorming**



**Step-1: Team Gathering, Collaboration and Select the Problem Statement**



### Step-2: Brainstorm, Idea Listing and Grouping

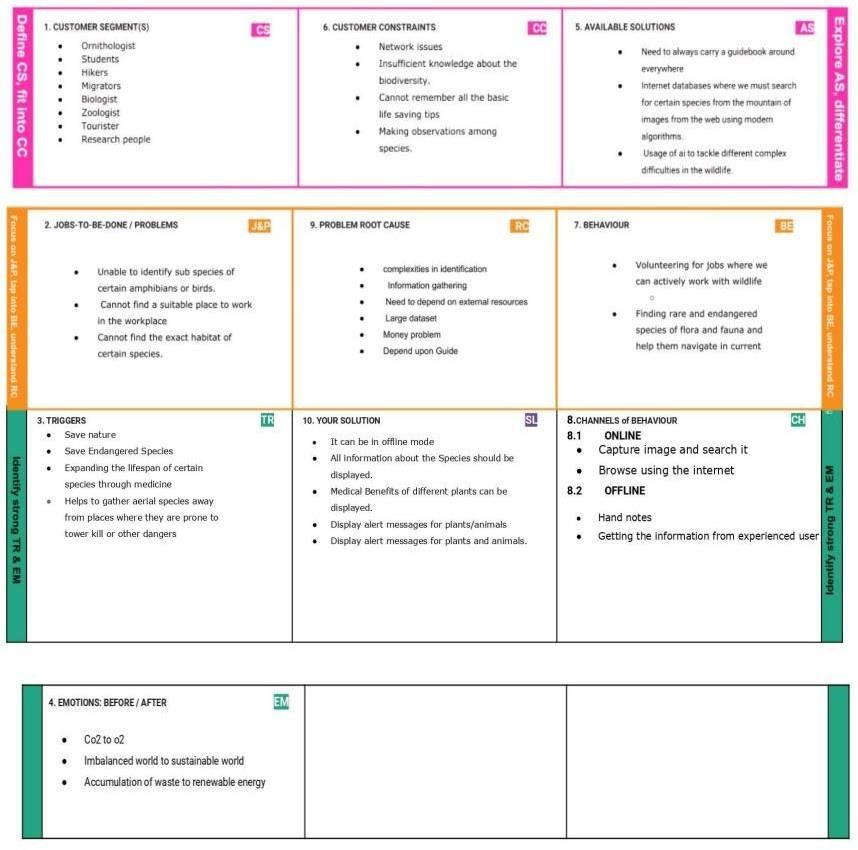


**Step-3: Idea Prioritization**

### Proposed Solution:

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Parameter** | **Description** |
| 1. | Problem Statement (Problem to be solved) | 1. How might we help both experienced and inexperienced user to identify species of plants   and animals and their characteristics with related information?   1. Inexperienced users need to know about poisonous plants and dangerous animals so that they can stay away from it. 2. Both experienced and inexperienced users need to know about the medicinal values of a plant because they need to use it in case of emergencies. 3. All the users need to know about the rarity of the species of birds, animals or plants so that they can preserve and save it. |
| 2. | Idea / Solution description | 1. Display Botanical names 2. Display alert messages for plants/animals using different colours 3. small description about them 4. Rarities of the species 5. What disease does the plant cure |
| 3. | Novelty / Uniqueness | 1. Providing alerts based on if a species is harmful or not 2. Alerting the user on the rarity of the species 3. Gives the complete description about the species being viewed 4. If the plant being viewed has a medicinal value , it gives a description about it. 5. Display the scientific name of the species. |
| 4. | Social Impact / Customer Satisfaction | Being able to identify the **flora** and **fauna** around us often leads to an interest in protecting wild spaces. |
| 5. | Business Model (Revenue Model) | 1. Can make money through subscription based. 2. Partnership with many laboratories and scientists around the world |
| 6. | Scalability of the Solution | 1. As the usage and user base of this application grows more feature can be added to the premium or subscription model. 2. We can introduce subscription models like free plan, business plan, educational plan and many more based on its usage 3. As the usage increase we can scale the application by releasing more languages based on the geographical usage. |

* 1. **Problem Solution fit**



# REQUIREMENT ANALYSIS

(Following are the functional & non-functional requirements of the proposed solution)

### Functional requirement.

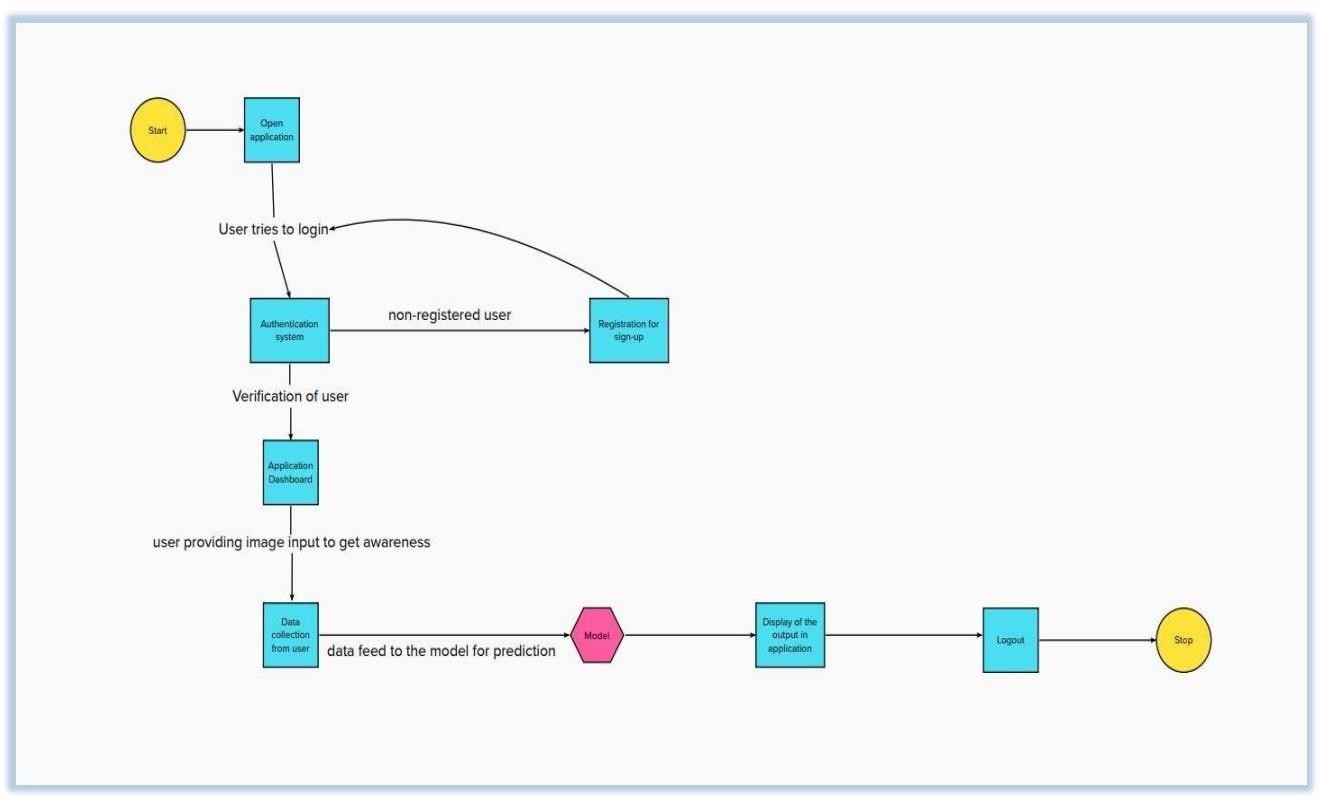
|  |  |  |
| --- | --- | --- |
| **FR No.** | **Functional Requirement (Epic)** | **Sub Requirement (Story / Sub-Task)** |
| FR-1 | User Registration | Registration through Form Registration through Gmail  Registration through LinkedIN |
| FR-2 | User Confirmation | Confirmation via Email Confirmation via OTP |
| FR-3 | Navigation Service | GPS |
| FR-4 | Database | My SQL, IBM Cloud |
| FR-5 | Premium features | Location sharing,  Adding information of new data by User |
| FR-6 | Updating and bug fixing | Updating the application based on user feedback |
| FR-7 | Final Output | Final description of the image (species) captured. |
| FR-8 | Alerts | System should alert about dangerous plants and animals |

* 1. **Non-functional Requirements**.

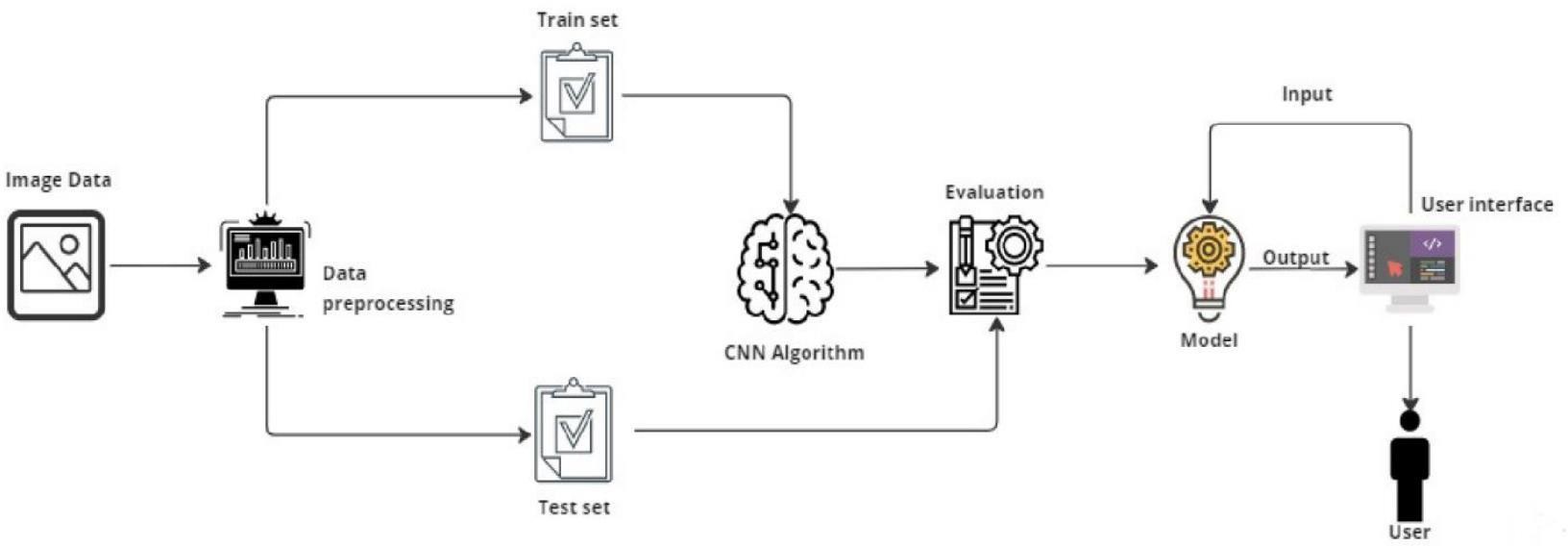
|  |  |  |
| --- | --- | --- |
| **FR No.** | **Non-Functional Requirement** | **Description** |
| NFR-1 | **Usability** | The users must be able to use the application without any problems and difficulties. The app is easy to view and does not strain the eyes. All information are in simple terms. The error rate of the final output must not be more than 20% |
| NFR-2 | **Security** | SHA-256, Encryptions, AES etc. |
| NFR-3 | **Reliability** | The system must perform without failure in 80 percent of the time. |
| NFR-4 | **Performance** | Under normal load, the system must show the results within 15 seconds, and under maximum it can take up to load 30 seconds |
| NFR-5 | **Availability** | The application will be available 99 % of the time in a month. |
| NFR-6 | **Scalability** | The system must be able to support 10,000 users while using it. As the usage and user base of this application grows,  more features can be added like languages based on the geographical usage, premium or subscription model, etc. |

# PROJECT DESIGN

### Data Flow Diagrams



**Solution & Technical Architecture**



### User Stories

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **User Type** | **Functional Requirement (Epic)** | **User Story Numbe r** | **User Story / Task** | **Acceptance criteria** | **Priority** |
| Customer (Mobile user) | Image capture | USN-1 | As a user, I can take photos of the plantlife, animals and birds | I can take photos whenrequired | High |
|  |  | USN-2 | As a user, I will receive processed information about the type of species | I can see the type of plant or animal or plant | High |
|  |  | USN-3 | As a user, I can share it with others | I can share using shareoption | Low |
|  | Data process | USN-4 | Data must be trained and tested and CNN algorithm must work properly. | I must see the correct processed information | High |
|  | Output | USN-5 | As a user, I can see the scientific nameof the species | I must see the correctdata | High |
|  |  | USN-6 | As a user, I can see the characteristicsand alert messages | I must see the correctdata | High |
| Administrator | Manage | USN-1 | As a admin I must add various data andedit information | I must edit the datapresent | High |

# PROJECT PLANNING & SCHEDULING

### Sprint Planning & Estimation

Use the below template to create product backlog and sprint schedule

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Spri nt** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** | **Stor y Poin**  **ts** | **Priority** |
| Sprint- | Registration | USN-1 | As a biogeography, I can register | 2 | High |
| 1 | for the application by entering my |
|  | Email, Password, and confirming |
|  | my password |
| Sprint- | User | USN-2 | As a biogeography, I will receive confirmation email once I have registered for the application | 1 | Medium |
| 1 | Confirmation |
|  |  |
|  |  |
|  |  |
|  |  |
| Sprint- | Login | USN-3 | As an biogeography, I can log into | 2 | High |
| 1 | the application by entering email & |
|  | password |
|  |  |
|  |  |
|  |  |
| Sprint- 2 | Data Collection | USN-1 | Download the dataset used in Digital Naturalist – AI Enabled tools for Biodiversity Researchers | 2 | Medium |
|
|
|
| Sprint- | Image | USN-1 | Improving the image data that | 1 | High |
| 2 | Preprocessin | suppresses unwilling distortions or |
|  | g | enhances some image features |
|  |  | important for further processing, |
|  |  | although performing some |
|  |  | geometric transformations of |
|  |  | images like rotation, scaling, etc. |
| Sprint- | Getting | USN-1 | Neural network are integral for |  | Medium |
| 3 | started with | teaching computers to think and |
|  | Convolution a | learn by classifying information, |
|  | l Neural | similar to how we as humans learn. |
|  | Network | With neural networks, the software |
|  |  | can learn to recognize images, for |
|  |  | example. Machines can also make |
|  |  | predictions and decisions with a |
|  |  | high level of accuracy based on |

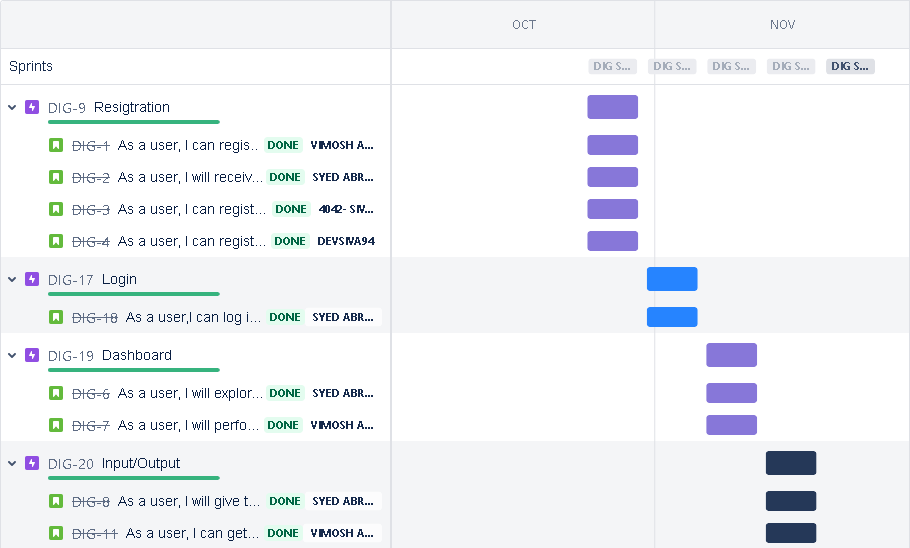
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | data inputs. |  |  |
| Sprint- | Evaluation and | USN-1 | well a model behaves after each | 1 | ⁷ |
| 3 | model saving | iteration of optimization. An |
|  |  | accuracy metric is used to measure |
|  |  | the algorithm's performance in an |
|  |  | interpretable way. The accuracy of |
|  |  | a model is usually determined after |
|  |  | the model parameters and is |
|  |  | calculated in the form of a |
|  |  | percentage. Saving The Model get |
|  |  | weights , set weights . |
| Sprint- | Application | USN-2 | After the model is built, we will be | 1 | High |
| 4 | Building | integrating it to a web application |
|  |  | so that normal users can also use it. |
|  |  | The users need to give the images |
|  |  | of species |
|  |  |  |
| Sprint- 4 | Train the Model on IBM | USN-2 | Build Deep learning model and computer vision Using the IBM cloud. | 2 |  |
|
|
|

### Sprint Delivery Schedule

Project Tracker, Velocity & Burndown Chart:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Total Story Points** | **Duration** | **Sprint Start Date** | **Sprint End Date (Planned)** | **Story Points Completed (as onPlanned End Date)** | **Sprint Release Date (Actual**  **)** |
| Sprint-1 | 20 | 4 Days | 24 Oct 2022 | 27 Oct 2022 | 20 | 29 Oct  2022 |
| Sprint-2 | 20 | 5 Days | 28 Oct 2022 | 01 Nov 2022 | 20 | 04 Nov  2022 |
| Sprint-3 | 20 | 7 Days | 02 Nov 2022 | 09 Nov 2022 | 20 | 11 Nov  2022 |
| Sprint-4 | 20 | 9 Days | 10 Nov 2022 | 18 Nov 2022 | 20 | 19 Nov  2022 |

* 1. **Reports from JIRA**



# CODING & SOLUTIONING

### Feature 1:

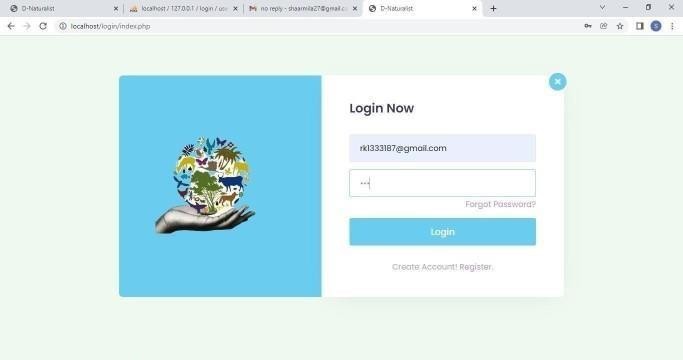
* + - Display Botanical names
    - Display alert messages for plants/animals using different colors
    - small description about them
    - Rarities of the species

### Feature 2:

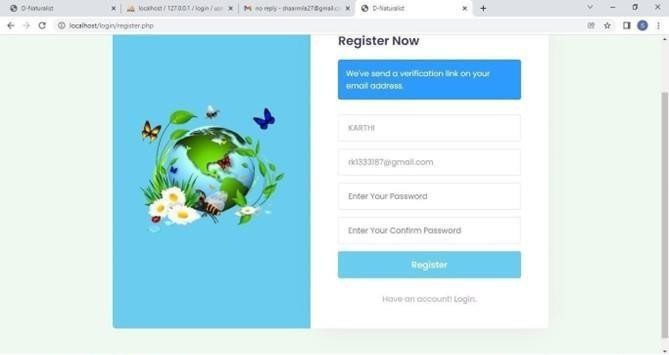
* + - What disease does the plant cure
    - Providing alerts based on if a species is harmful or not

# RESULTS

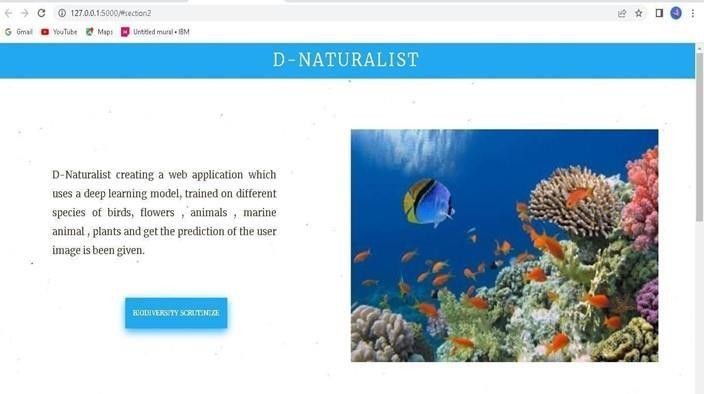
### Performance Metrics Registration Page:



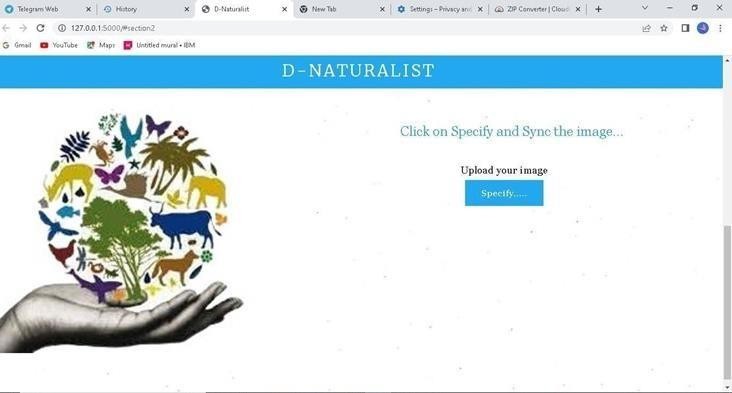
**Login Page :**



### Home Page(Dashboard):



**Input:**



**Output:**



# ADVANTAGES & DISADVANTAGES

## Advantages:

* + - It helps field biologists build their own experimental tools.
    - It helps designers explore new interactions with nature.
    - It paved the ways to discover new ways of experiencing the natural world.
    - It helps to create sustainable world by saving the endangered species.
    - Digital Naturalism unites biologists, designers, engineers, and artists to build and analyze new devices.
    - User can get detailed description of any kind of species.
    - It is a handy application for a person who is travel freak.

## Disadvantages:

* + - Proper network should be maintained to avoid hitches.
    - Difficult to classify sub-classes of same species.

# CONCLUSION

Assessment of regional biodiversity based on global scientific consensus is a scientific basis for the whole society and a tool for local to international discussion and decision making. In the new era of extinction, people would understand the value of (intrinsic or otherwise) of our state's threatened biodiversity. Fewer observations, data points and discoveries would be made of the natural world to help us measure our impacts on it. It provides the opportunity to build a positive identity with science or recognize the value of more holistic ways of thinking such as traditional ecological knowledge. It helps us to engage in environmental stewardship behaviours ranging from resource conservation to building resilience among vulnerable communities. By acknowledging our origins in evolution, the naturalist perspective also enhances our feeling of kinship with the other species with which we share this planet, and our desire to sustain and nurture the planet itself. All sentient beings, including humanity, owe their existence to conditions that extend far beyond us in space and time. The model which was used for the detection of digital naturalists using the species images from the wild life and the species with flora part and with fauna part will be displayed as well. From the resultant graphs, it is proven that the accuracy of the model has reached good level. If it is deployed in the real-time scenario then it will help many people in distinguishing between both without wasting the money on various machines. If the image is confirmed by the model, then the person can know the feature of the species. It can be the best way of practice for people to save money. As we know that the data plays a crucial role in every deep learning model, if the data is more specific and accurate about the species then that can help in reaching greater accuracy with better results in real- time applications.

# FUTURE SCOPE

### AI image classiﬁers can create biodiversity datasets from social media imagery

* + - **Flickr hosts many images of plants; some can be accurately classiﬁed to species by AI**

### Images are spatially aggregated around tourist sites and under-represent native species

* + - **Images focused on a single, non-horticultural, plant are most reliably**

# APPENDIX

### Source Code

**<--------------------------------------------- # index1.php >**

<?php

session\_start();

if (isset($\_SESSION['SESSION\_EMAIL'])) {

header("Location: welcome.php"); die();

}

include 'config.php';

$msg = "";

if (isset($\_GET['verification'])) {

if (mysqli\_num\_rows(mysqli\_query($conn, "SELECT \* FROM users WHERE code='{$\_GET['verification']}'")) > 0) {

$query = mysqli\_query($conn, "UPDATE users SET code='' WHERE code='{$\_GET['verification']}'");

if ($query) {

$msg = "<div class='alert alert-success'>Account verification has been successfully completed.</div>";

}

} else {

header("Location: index.php");

}

}

if (isset($\_POST['submit'])) {

$email = mysqli\_real\_escape\_string($conn, $\_POST['email']);

$password = mysqli\_real\_escape\_string($conn, md5($\_POST['password']));

$sql = "SELECT \* FROM users WHERE email='{$email}' AND password='{$password}'";

$result = mysqli\_query($conn, $sql);

if (mysqli\_num\_rows($result) === 1) {

$row = mysqli\_fetch\_assoc($result);

if (empty($row['code'])) {

$\_SESSION['SESSION\_EMAIL'] = $email;

header("Location: welcome.php");} else {

$msg = "<div class='alert alert-info'>First verify your account and try again.</div>";

}

} else {

$msg = "<div class='alert alert-danger'>Email or password do not match.</div>";

}

}

?>

<!DOCTYPE html>

<html lang="zxx">

<head>

<title>Digital Naturalist- </title>

<!-- Meta tag Keywords -->

<meta name="viewport" content="width=device-width, initial-scale=1">

<meta charset="UTF-8" />

<meta name="keywords" content="Login Form" />

<!-- //Meta tag Keywords -->

<link href="//fonts.googleapis.com/css2?family=Poppins:wght@300;400;500;600&display=swap" rel="stylesheet">

<!--/Style-CSS -->

<link rel="stylesheet" href="css/style.css" type="text/css" media="all" />

<!--//Style-CSS -->

<script src="https://kit.fontawesome.com/af562a2a63.js" crossorigin="anonymous"></script>

</head>

<body>

<!-- form section start -->

<section class="w3l-mockup-form">

<div class="container">

<!-- /form -->

<div class="workinghny-form-grid">

<div class="main-mockup">

<div class="alert-close">

<span class="fa fa-close"></span>

</div>

<div class="w3l\_form align-self">

<div class="left\_grid\_info"><img src="images/Loginpage.jpg" alt="">

required>

</div>

</div>

<div class="content-wthree">

<h2>Login Now</h2>

<p>Lorem ipsum dolor sit amet consectetur adipisicing elit. </p>

<?php echo $msg; ?>

<form action="" method="post">

<input type="email" class="email" name="email" placeholder="Enter Your Email"

<input type="password" class="password" name="password" placeholder="Enter Your

Password" style="margin-bottom: 2px;" required>

<p><a href="forgot-password.php" style="margin-bottom: 15px; display: block; text-align: right;">Forgot Password?</a></p>

<button name="submit" name="submit" class="btn" type="submit">Login</button>

</form>

<div class="social-icons">

<p>Create Account! <a href="register.php">Register</a>.</p>

</div>

</div>

</div>

</div>

<!-- //form -->

</div>

</section>

<!-- //form section start -->

<script src="js/jquery.min.js"></script>

<script>

$(document).ready(function (c) {

$('.alert-close').on('click', function (c) {

$('.main-mockup').fadeOut('slow', function (c) {

$('.main-mockup').remove();

});

});

});

</script>

</body>

</html>

### <--------------------------------------------- # app.py >

from future import division, print\_function import os

import numpy as np import tensorflow as tf

from tensorflow.keras.preprocessing import image from tensorflow.keras.models import load\_model from flask import Flask, request, render\_template from werkzeug.utils import secure\_filename global graph

#graph=tf.get\_default\_graph() # Define a flask app

app = Flask( name )

model = load\_model('nature1.h5')

print('Model loaded. Check http://127.0.0.1:5000/') @app.route('/', methods=['GET'])

def index():

# Main page

return render\_template('digital.html') @app.route('/predict', methods=['GET', 'POST']) def upload():

if request.method == 'POST':

# Get the file from post request f = request.files['image']

# Save the file to ./uploads

basepath = os.path.dirname( file )

file\_path = os.path.join(basepath, 'uploads', secure\_filename(f.filename)) f.save(file\_path)img = image.load\_img(file\_path, target\_size=(64,64))

x = image.img\_to\_array(img) x

= np.expand\_dims(x, axis=0)

#with graph.as\_default():

preds = np.argmax(model.predict(x)) found = ["animal- badger",

"animal- bat", "animal- bear", "animal- bee", "animal- dolphin", "animal- donkey", "animal- dragonfly", "animal- duck", "animal- eagle", "animal- elephant", "animal-flamingo", "animal-fly",

"animal-fox", "animal- gallina", "animal- gatto", "animal- hedgehog",

"animal- hippopotamus", "animal- hornbill", "animal- horse",

"animal- hummingbird"] print(preds)

text = found[preds]

if name == ' main ': app.run(threaded = False)

### GitHub & Project Demo Link

**GitHub Link :**

### REFERENCES

* 1. M.Mahesh, R. Rohan, V. Padmapriya, D. N. Sujatha (2022), ‘A Framework to Detect Hibiscus Flower Using YOLOV3 and SSD MobileNet’, ICDSMLA 2020, Vol.783, pp.699.
  2. Aparna, Saloni M P, Chandana M, Neha U K, Banushree D J, Prof. Naresh Patel K M (2021) ‘Digital Naturalist Using Deep Learning’, International Journal for Research in Applied Science & Engineering Technology (IJRASET), Vol.9 , issue 7, pp.2061-2065.
  3. Kamlesh Borana, Umesh More, Rajdeep Sodha, Prof. Vaishali Shirsath(2021) ‘Bird Species Identifier using Convolutional Neural Network’, International Journal of Engineering Research & Technology (IJERT), Vol.9 , issue 3, pp.340-344.
  4. Abdullah Albattal, Anjali Narayanan (2021) ‘Classifying Fish by Species Using Convolutional Neural Networks’. University of California San Diego.
  5. Li Liu and Paul W. Fieguth (2012) ‘Texture Classification from Random Features’, IEEE Transactions on Pattern Analysis and Machine Intelligence,Vol.34, No.11, pp.574-585.
  6. Mingyuan Xin and Yong Wang (2019) ‘Research on image classification model based on deep convolution neural network’, EURASIP Journal on Image and Video Processing, pp.01-11.

