PROJECT REPORT

1.INTRODUCTION:

- The ever-growing number of digital sensors in the environment has led to an increase in the amount of digital data being generated. This includes data from satellites, weather stations, data from ‘‘internet of things’’ devices, and data collected by members of the public via smartphone applications, to name but a few.
- These new sources of data have contributed to the era of ‘‘Big Data’’ characterized by large volumes of data, of numerous types and quality, being generated at an increasing speed. Building efficient, scalable, and robust approaches to solve these problems is a difficult scientific challenge at the forefront of data science and machine learning specifically. Artificial intelligence (AI) techniques have profoundly transformed our ability to extract information from visual data. AI techniques have been applied for a long time in security and industrial domains, for example, in iris recognition or the detection of faulty objects in manufacturing.

1.1 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 for conservation 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 from experienced ornithologists. There should be a handy tool for them to capture, identify and share the beauty to the outside world.

Field naturalists can only use this web app from anywhere to identify the birds, flowers, mammals and other species they see on their hikes, canoe trips and other excursions.
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 the prediction of the bird when an image is been given.

1.2 Purpose:

To biological recording have to date typically focused on active sampling, that is, images collected specifically for the purpose of recording wildlife (e.g., wildlife recording apps or camera traps). However, this has neglected large amounts of image data that are not collected for the purposes of biological recording, but which nonetheless may contain useful information about biodiversity.

This includes social media imagery(e.g., Flickr and Instagram), CCTV, and imagery collected along linear infrastructure (e.g., Google StreetView). These unexploited image data could be rapidly analyzed using “AI naturalists” designed to locate potential images of biodiversity and classify what they see.

2. LITERATURE SURVEY:

"Bird classification using CNN" by Simon Haykin: This work presents a scenario with classification of birds using CNN technique based on color features. They used color images of birds with almost similar types. Image segmentation is carried in various stages. At first, the pixels are arranged and segmented on the basis of edges and spatial segmentation, where clustering is done. Next, the blocks are segmented using edge detection.

The computational efficiency increases for image and training becomes easier. This approach provides with better and robust results for different images. Here they took sparrow for the case study and evaluated the features of it using the steps up listed. The experimental results classify the effectiveness of proposed approach to improve the segmentation quality in aspects of precision and computational time .

2.1 Existing Problem:

By combining social media APIs with AI classifiers, we were able to build an AI naturalist capable of creating biodiversity datasets from previously unexploited data sources. However, we demonstrate that there are a number of biases in the data produced, some of which may be able to be mitigated against, that must be carefully considered before the data could be used in certain types of analyses.

2.2 References:

1.M. Schroeck, R. Shockley, J. Smart, D. Romero-Morales, P. Tufano Analytics: The Real-World Use of Big Data IBM Institute for Business Value (2012) Google Scholar

2.A.Y. Sun, B.R. Scanlon How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions Environ. Res. Lett., 14 (2019), p. 073001 View PDFCrossRefView Record in ScopusGoogle Scholar

3.M.A. Tabak, M.S. Norouzzadeh, D.W. Wolfson, S.J. Sweeney, K.C. Vercauteren, N.P. Snow, J.M. Halseth, P.A.D. Salvo, J.S. Lewis, M.D. White, et al. Machine learning to classify animal species in camera trap images: applications in ecology Methods Ecol. Evol., 10 (2019), pp. 585-590 View PDFCrossRefView Record in ScopusGoogle Scholar

4.R. Gibb, E. Browning, P. Glover-Kapfer, K.E. Jones Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring Methods Ecol. Evol., 10 (2019), pp. 169-185 View PDFCrossRefView Record in ScopusGoogle Scholar

5.B. Efron, T. Hastie Computer Age Statistical Inference Cambridge University Press (2016) Google Scholar

6.C.J. Lintott, K. Schawinski, A. Slosar, K. Land, S. Bamford, D. Thomas, M.J. Raddick, R.C. Nichol, A. Szalay, D. Andreescu, et al. Galaxy zoo: morphologies derived from visual inspection of galaxies from the Sloan digital sky survey Mon. Not. R. Astron. Soc., 389 (2008), pp. 1179-1189 View PDFCrossRefGoogle Scholar

7.K. Nguyen, C. Fookes, R. Jillela, S. Sridharan, A. Ross Long range iris recognition: a survey Pattern Recognit., 72 (2017), pp. 123-143

2.3 Problem Statement Definition:

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 for conservation groups like NCC.

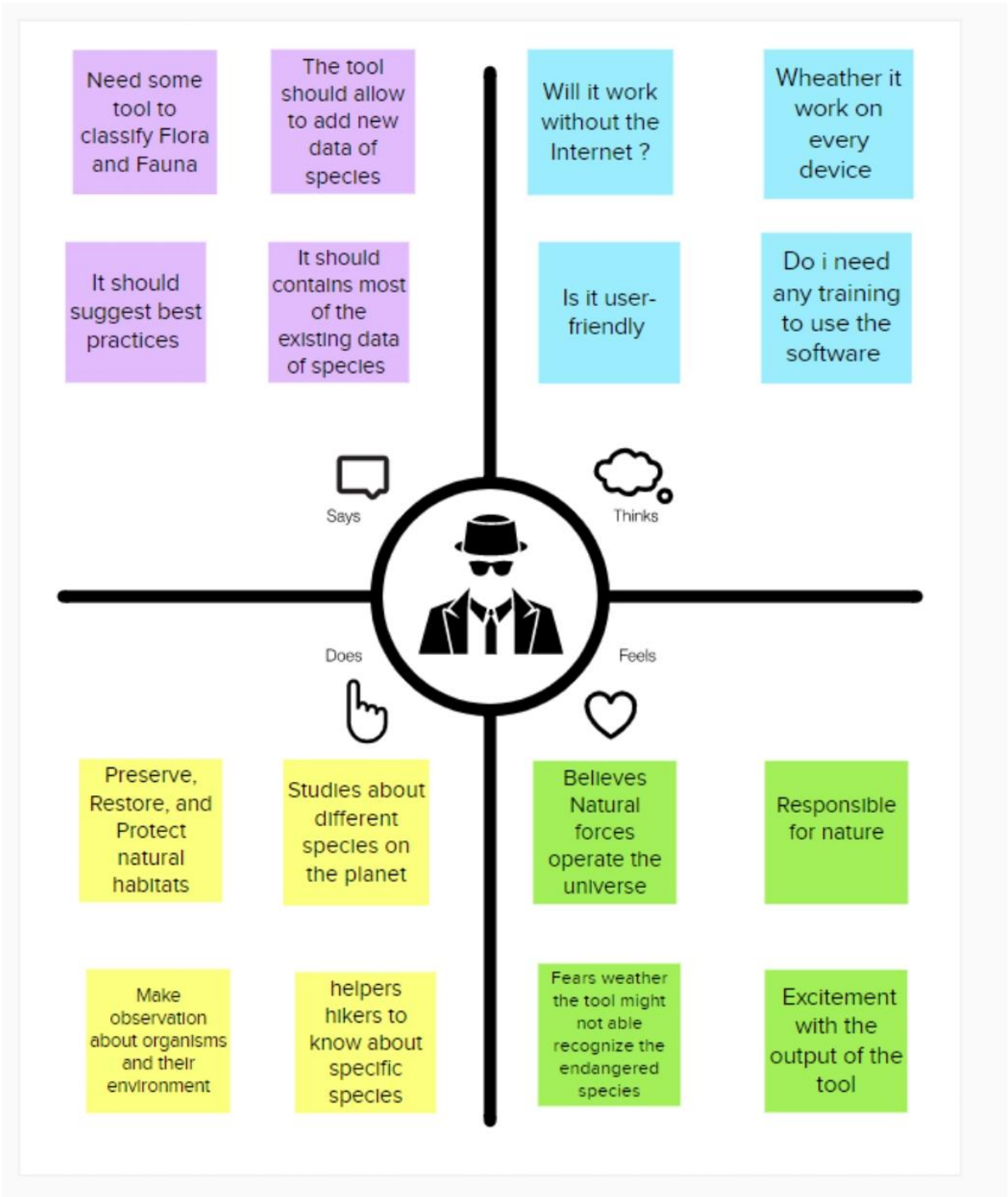
When venturing into the woods, field naturalists usually rely on common approaches like always carrying a guidebook around everywhere or seeking help from experienced ornithologists. There should be a handy tool for them to capture, identify and share the beauty to the outside world.

3. IDEATION & PROPOSED SOLUTION:

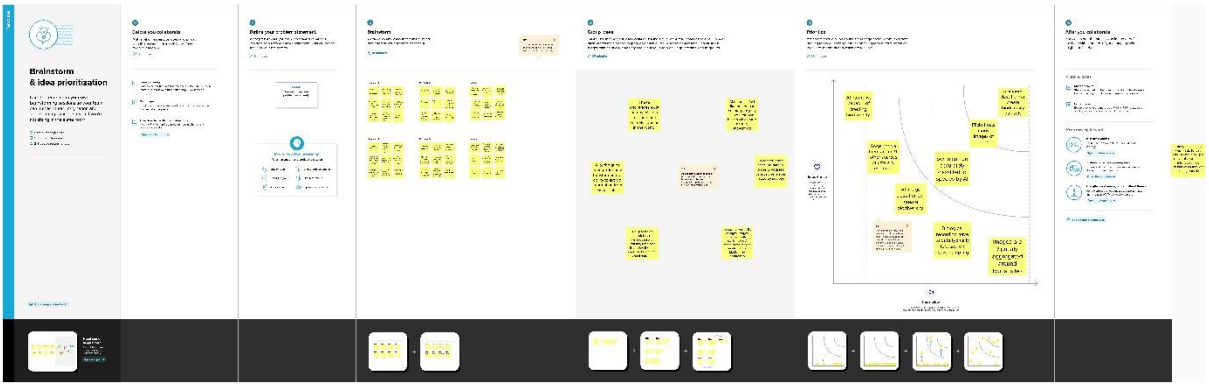
To identify a species in a forest or in any other place, we need to carry a heavy book or seek a professional like botanist or zoologist or an ornithologist, but there should be a handy tool for them to capture, identify and share the beauty to the outside world.

Idea / Solution description A system is built by using the Image/object recognition and classification using (CNN) Convolutional neural network which while using this system, we can capture the image of any animals and plants and can obtain the information about the flora and fauna at any time.

3.1 Empathy Map Canvas:



3.2 Ideation & Brainstorming:



3.3 Proposed Solution:

Parameter	Description
Problem Statement (Problem to be solved)	To identify a species in a forest or in any other place, we need to carry a heavy book or seek a professional like botanist or zoologist or an ornithologist, but there should be a handy tool for them to capture, identify and share the beauty to the outside world.
Idea / Solution description	A system is built by using the Image/object recognition and classification using (CNN) Convolutional neural network which while using this system, we can capture the image of any animals and plants and can obtain the information about the flora and fauna at any time.
Novelty / Uniqueness	Use of transfer learning in pre trained models to increase accuracy and training time along with data augmentation to increase the dataset size which will in turn yield more accuracy.
Social Impact / Customer Satisfaction	The user can identify the type of species faster and easier without searching in books page by page. It is a useful product for all the research analyst, Ornithologist, Biologist and Marine drivers who can instantly capture images of different species and are able to get all the relevant information about those breeds.

Business Model (Revenue Model)	The model could be open sourced, but we can get some revenue via ads. we will also add a few extra applications like storage and bookmarks permanently for a specific amount of payment.
Scalability of the Solution	The system apart from researchers can also be used by students, hikers or other people who are very much interested in the wildlife. This can also help children learn about different species and their sub species of different flora and fauna which can only be found in the other part of the world.

3.4 Problem Solution Fit:

4. REQUIREMENT ANALYSIS:

4.1 Functional Requirement:

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Google API
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Transactions	Through UPI, Credit/Debit cards and Net Banking.
FR-4	Authentication	• Through OTP sent to mobile. • User created secured passwords.
FR-5	Authorization	Basic Authorization.
FR-6	Administrative Function	Adding, Updating and Maintaining description data about various species.

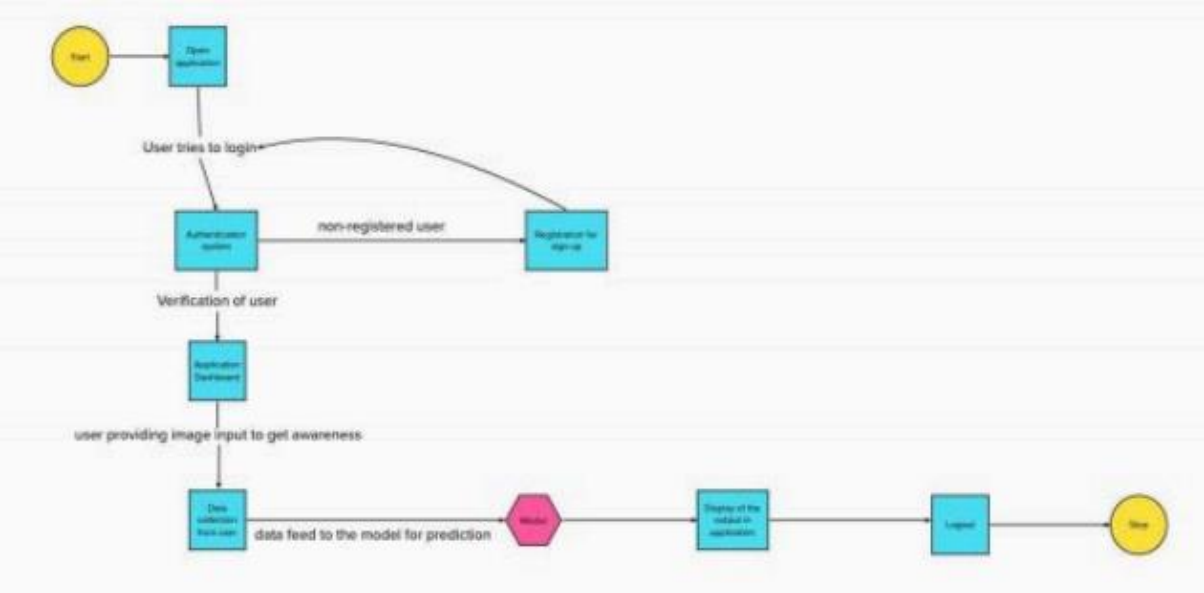
4.2 Non – Functional Requirement:

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	Our solution is demanded for scientific researchers Such as Ornithologists , Zoologists in order to predict and analyse about flora and fauna.
NFR-2	**Security**	Authentication process involves multilayer security to make user data and collected data more secured, also to avoid unknown authorization and data integrity issues. Most security methods include Encryption and Authorization
NFR-3	**Reliability**	Our framework should be reliable to cover wide range of species spanning across various habitats.
NFR-4	**Performance**	Data Augmentation to increase dataset size along with transfer learning to increase accuracy and performance for better working of application.
NFR-5	**Availability**	Our application possess full-time service (either offline or online) and dataset is constantly updated

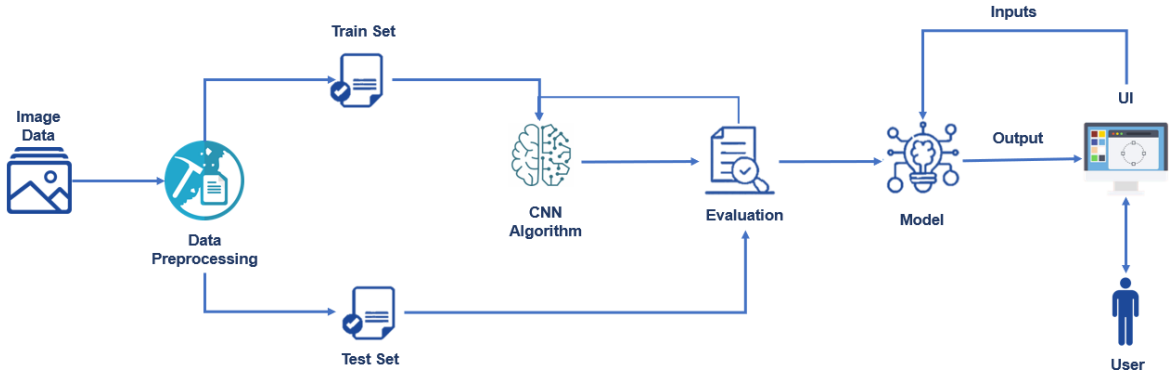
5. PROJECT DESIGN:

5.1 Data Flow Diagram:

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 the data is stored.



5.2 Solution and Technical Architecture:



5.3 User Stories:

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a biogeography, I can register for the application by entering my email, password, and confirming my password.	2	High	**J.Abinaya(TL)**
Sprint-1		USN-2	As a biogeography, I can register for the application through Gmail.	2	High	**B.Priyanka**
Sprint-1	User Confirmation	USN-3	As a biogeography, I will receive confirmation email once I have registered for the application.	2	High	**B.Langeshwari**
Sprint-1	Login	USN-4	As a biogeography, I can log into the application by entering email & password.	2	High	**S.Roja**

6. PROJECT PLANNING AND SCHEDULING:

6.1 Sprint planning and Estimation:

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned 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	8 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

6.2 Sprint Delivery Schedule:

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned 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	8 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

7. CODING AND SOLUTIONING:

```
 <!DOCTYPE html>
<html>
<style>

   /assign full width inputs/
   input[type=text],
   input[type=password] {
      width: 100%;
      padding: 12px 20px;
      margin: 8px 0;
      display: inline-block;
      border: 1px solid #ccc;
      box-sizing: border-box;
   }

   /set a style for the buttons/
   button {
      background-color: #4CAF50;
      color: white;
      padding: 14px 20px;
      margin: 8px 0;
      border: none;
      cursor: pointer;
      width: 100%;
   }

   /* set a hover effect for the button*/
   button:hover {
      opacity: 0.8;
```

```
}

/set extra style for the cancel button/
.cancelbtn {
    width: auto;
    padding: 10px 18px;
    background-color: #f44336;
}

/centre the display image inside the container/
.imgcontainer {
    text-align: center;
    margin: 24px 0 12px 0;
    position: relative;
}

/set image properties/
img.avatar {
    width: 40%;
    border-radius: 50%;
}

/set padding to the container/
.container {
    padding: 16px;
}

/set the forgot password text/
span.psw {
    float: right;
```

```
    padding-top: 16px;
}

/set the Modal background/
.modal {
    display: none;
    position: fixed;
    z-index: 1;
    left: 0;
    top: 0;
    width: 100%;
    height: 100%;
    overflow: auto;
    background-color: rgb(0, 0, 0);
    background-color: rgba(0, 0, 0, 0.4);
    padding-top: 60px;
}

/style the model content box/
.modal-content {
    background-color: #fefefe;
    margin: 5% auto 15% auto;
    border: 1px solid #888;
    width: 80%;
}

/style the close button/
.close {
    position: absolute;
    right: 25px;
```

```
    top: 0;
    color: #000;
    font-size: 35px;
    font-weight: bold;
}

.close:hover,
.close:focus {
    color: red;
    cursor: pointer;
}

/* add zoom animation*/
.animate {
    -webkit-animation: animatezoom 0.6s;
    animation: animatezoom 0.6s
}

@-webkit-keyframes animatezoom {
    from {
        -webkit-transform: scale(0)
    }
    to {
        -webkit-transform: scale(1)
    }
}

@keyframes animatezoom {
    from {
        transform: scale(0)
```

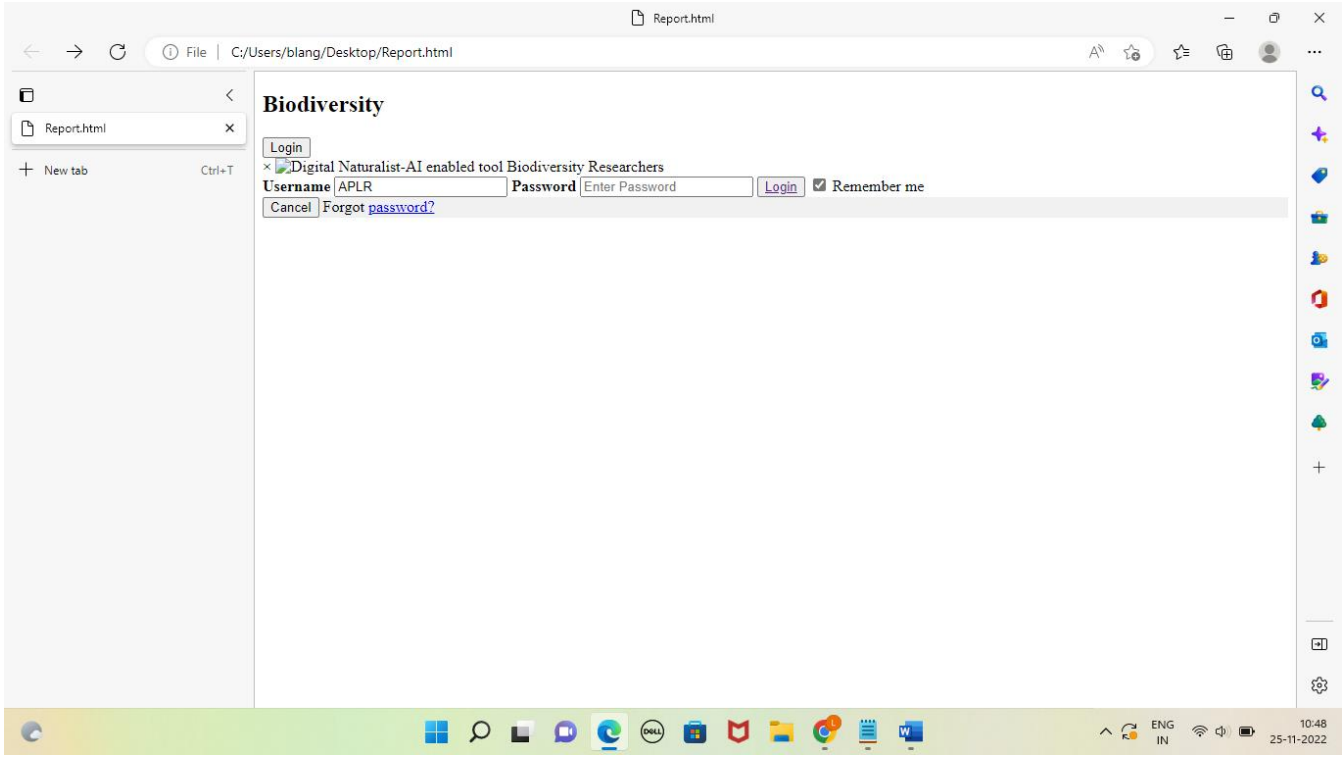
```
        }
        to {
            transform: scale(1)
        }
    }

    @media screen and (max-width: 300px) {
        span.psw {
            display: block;
            float: none;
        }
        .cancelbtn {
            width: 100%;
        }
    }
</style>

<body>

    <h2>Biodiversity</h2>
    <!--Step 1 : Adding HTML-->
    <button onclick="document.getElementById('id01').style.display='block'" style="width:auto;">Login</button>

    <div id="id01" class="modal">

        <form class="modal-content animate" action="/action_page.php">
            <div class="imgcontainer">
                <span onclick="document.getElementById('id01').style.display='none'" class="close" title="Close Modal">&times;</span>
```

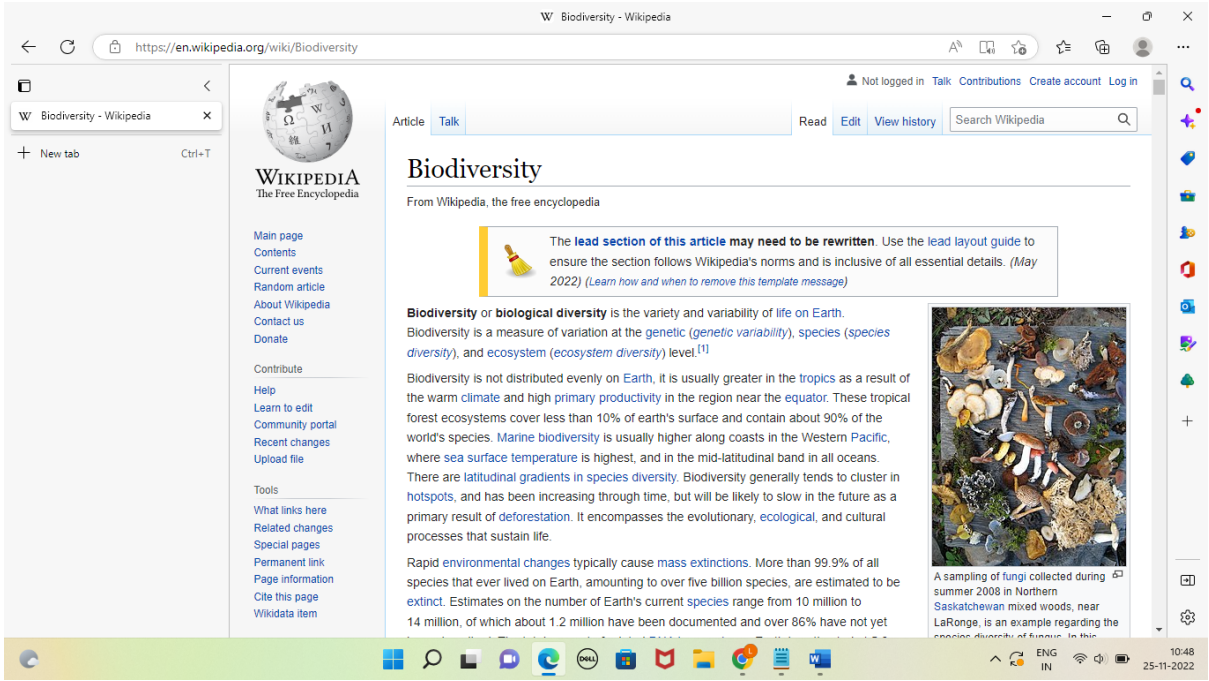
```
            <img src="C:\Users\KARTHI\Downloads_biodiversity2.jpg" alt="Digital
Naturalist-AI enabled tool Biodiversity Researchers" class="biodiversity2">
        </div>

        <div class="container">
            <label><b>Username</b></label>
            <input type="text" placeholder="Enter Username" name="uname" required>

            <label><b>Password</b></label>
            <input type="password" placeholder="Enter Password" name="psw" required>

            <button type="submit"><a href =
"https://en.wikipedia.org/wiki/Biodiversity" >Login </a></button>
            <input type="checkbox" checked="checked"> Remember me
        </div>

        <div class="container" style="background-color:#f1f1f1">
            <button type="button"
onclick="document.getElementById('id01').style.display='none'"
class="cancelbtn">Cancel</button>
            <span class="psw">Forgot <a href="#">password?</a></span>
        </div>
    </form>
  </div>

  <script>
    var modal = document.getElementById('id01');
    window.onclick = function(event) {
        if (event.target == modal) {
            modal.style.display = "none";
        }
    }
```

```
    </script>
</body>

</html>
```

8. TESTING:

8.1 Test Cases:





8.2 User Acceptance Testing:

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user.It also ensures that the system meets the functional requirements.

9. RESULTS:

Images Are Spatially Clustered Flickr searches returned a far greater number of images for central London (n = 55,176; 1,200 images/km2) than for the Peak District (n = 5,486; 46 images/km2). Images were taken between and April 26, 2003, and August 23, 2019 (Figure 1). By definition these are only the subset of images taken in these locations that had location data available. To obtain an indication of the proportion of images that did not have location data, we searched for all images that contained the word ‘‘flower’’ taken in the first week of July 2019, regardless of location information.

The ever-growing number of digital sensors in the environment has led to an increase in the amount of digital data being generated. This includes data from satellites, weather stations, data from ‘‘internet of things’’ devices, and data collected by members of the public via smartphone applications, to name but a few.

9.1 Performance Metrics:

1. Randomly Selected Example Images:

The top row (1–3) were all correctly identified to species by the AI classifier; 4 and 5 were classed as unidentifiable by our expert botanist, with 4 additionally classified as a representation; 6 was classed as identifiable, but as not being focused on a single species.

2. Spatial Distribution of Images:

The spatial distribution of Flickr images returned when searching with the term ‘‘flower’’ in (A) London (urban) and (B) the Peak District (rural). Gray/black dots show the location of individual images. Colored areas show regions of particularly high densities of images.

10. ADVANTAGES AND DISAVANTAGES:

10.1 Advantages:

- ✓ The project aims to create an application for the hikers to identify rare species of birds, flowers, mammals by giving a picture taken by them.
- ✓ To biological recording have to date typically focused on active sampling, that is, images collected specifically for the purpose of recording wildlife (e.g., wildlife recording apps or camera traps).
- ✓ However, this has neglected large amounts of image data that are not collected for the purposes of biological recording, but which nonetheless may contain useful information about biodiversity.

10.2Disadvantages:

- ✓ Naturalistic observation is a valuable tool because of its flexibility, external validity, and suitability for topics that can't be studied in a lab setting.
- ✓ They downsides of naturalistic observation include its **lack of scientific control, ethical considerations, and potential for bias from observers and subjects**.

11. CONCLUSION:

Field naturalists can only use theis web app from anywhere to identify the birds,flowers,animals,mammals and other species they see on the hikes,canoe trips and other excursions.

In this project ,we are creating a web application which uses a deep learning model,trained on different species of birds ,flowers and mammals.There is great diversity among naturalists,but some common ground too.

All naturalism begin with an admiring attitude towards science and its achievements.In many cases this admiring attitude is combine with a contempt or distrust for the way that philosophy has been or is conducted.For example,uses scepticism as a device to sweep away the hold Aristotelian foundations of knowledge.

12. FUTURE SCOPE:

Essentially , the proposed guidelines treat statistical comparison of ML based quality estimators as a multi dimensional problem.Accordingly,we seek to assess the predictors more holistically in teams of their local performance on specific test conditions ,their learning ability and the magnitude of treatment effect(to quantify the practical significant of the observed differences).

In contrast, the current approach tends to reduce this tasks to binary and global statistical decision making and does not reveal systematic weaknesses(or strengths) of the predictors.In order to provide a tool for practical use , software implementing the proposed guidelines is made publicly available.

Github link: https://github.com/IBM-EPBL/IBM-Project-30953-1660193242