GUJARAT TECHNOLOGICAL UNIVERSITY

CHANDKHEDA, AHMEDABAD

AFFILIATED

SARVAJANIK COLLEGE OF ENGINEERING AND TECHNOLOGY

A

PROJECT REPORT

ON

"AI-BIRDie"

PROJECT TYPE : UDP

B.E. IV, SEMESTER – VII, COMPUTER – (Shift-2)

SUBMITTED BY :

GROUP NO : 12_S2

SR. NO	NAME	ENROLLMENT
1	SMIT JIVANI	160420107524
2	VISHAL JOBANPUTRA	160420107525
3	NEHA DADARWALA	160420107540
4	JAIMI SHETA	160420107558
5	JAY SONANI	160420107559

Prof. (Dr.) Nirali Nanavati

(FACULTY GUIDE)

Prof. (Dr.) Pariza Kambhoj

(HEAD OF THE DEPARTMENT)

ACADEMIC YEAR : (2019 – 2020)

"AI-BIRDie"

A PROJECT REPORT

Submitted by

SMIT JIVANI (160420107524)

VISHAL JOBANPUTRA (160420107525)

NEHA DADARWALA (160420107540)

JAIMI SHETA (160420107558)

JAY SONANI (160420107559)

In fulfillment for the award of the

degree Of

BACHELOR OF ENGINEERING

In

COMPUTER ENGINEERING

Sarvajanik College of Engineering and Technology,
Surat.

Gujarat Technological University, Ahmedabad.
November, 2019

SARVAJANIK COLLEGE OF ENGINEERING AND TECHNOLOGY
Dr. R.K.DESAI MARG, ATHWALINES,
SURAT-395001

DEPARTMENT OF COMPUTER ENGINEERING

CERTIFICATE

This is to certify that the project entitled *AI-BIRDie* has been carried out by Smit Jivani (160420107524), Vishal Jobanputra (160420107525), Neha Dadarwala (160420107540), Jaimi Sheta (160420107558), Jay Sonani (160420107559), students of B.E.IV (CO-Shift2), Semester-VII, under my guidance in fulfillment of the degree of Bachelor of Engineering in Computer Engineering of Gujarat Technological University, Ahmedabad for the academic year Nov-2019.

Signature of
Guide

---------- Prof. (Dr.) Nirali Nanavati

Signature of
Head of the Department

---------- Prof. (Dr.) Pariza Kambhoj

Signature of Jury Members

---------- , ---------- , ---------- , ----------

ACKNOWLEDGEMENTS :

We would like to acknowledge the valuable guidance and constant motivation which we have received throughout the course of our project from our guide Prof. (Dr.) Nirali Nanavati. We are also grateful to have Prof (Dr.) Arnav Bhavsar, Professor at Indian Institute of Technology (IIT), Mandi as our External guide who have guided and encouraged us in our project work, we also thank sir for giving us his valuable time and insightful advice.

We thank all our faculty members of Computer Department, SCET who have supported and imparted their knowledge which have helped us in our project work.

ABSTRACT:

Birds are important indicators of ecosystem health. They play vital roles in pollination, keeping the population of insects under control, and in-plant dispersal .Birds are vulnerable to human-induced change in their surroundings and several species all over the world, are under the threat of population decline. The use of technology supplements many traditional field studies, which are usually human as well as cost-intensive and are limited in their scalability. Birds can be detected by both sight and sound. Hence, sounds, images and videos are useful in detecting and identifying bird species. The information provided by these modalities can be combined to provide reliable mechanisms to detect and identify birds. Our goal is to create a one-stop application for all the bird watchers and wildlife enthusiasts out there in India. Leveraging the power of AI, the application will contain features like bird detection and recognition from images and audio, trivia and a personal dashboard service.

TABLE OF CONTENTS:

LIST OF FIGURES:

CHAPTER 1: INTRODUCTION:

1.1 PROBLEM SUMMARY:

- Birds are important indicators of ecosystem health
- They play vital roles in pollination, keeping the population of insects under control, and in plant dispersal.
- Birds are vulnerable to human-induced change in their surroundings and several species all over the world, are under the threat of population decline.
- For ecologists, the use of technology supplements many traditional field studies, which are usually human as well as cost-intensive and are limited in their scalability.
- Technology can be used to create systems which cut down human effort and can scale up to collect and analyse vast amount of data.
- According to our prior art search, there is no single application in the market right now which provides the service of birds species identification through image and audio.
- There is no application right now in the Indian subcontinent which provides the service of acoustic detection for Indian bird species.

- Sounds are captured using microphones, and images and videos using cameras.
- Of course, the problem is much more complex than what is described above. The microphone records not just bird sounds, but all sounds near it. This includes humans talking, passing vehicles, the wind, other animals etc. Similarly, the camera captures birds only when they come into the field of view.
- Sometimes when captured through the camera, birds could be occluded by leaves, may be out of focus or be only partly visible. Thus, even detecting whether a bird is present, partly present, or not present is a challenge in itself.
- Varying degrees of noise in the signals, including background noise for sounds, varying degrees of pose, illumination for images and videos increase the challenges

1.2 AIM AND OBJECTIVES OF THE PROJECT:

AIM:

Our goal is to create a one-stop application for all the bird watchers and wildlife enthusiasts out there in India. Leveraging the power of AI, the application will contain features like bird detection and recognition from images and audio, trivia and a personal dashboard service

OBJECTIVES:

1. For visual identification, our objective is to develop algorithms for :
    a. Bird Detection in images
    b. Bird Identification from images
2. For acoustic identification, our objective is to develop algorithms for:
    a. Bird Detection
    b. Species Identification

1.3 PROBLEM SPECIFICATION:

1.3.1 Functional Requirements :

- **Bird classification based on Audio:** AI-Birdie provides bird classification based on audio. The application takes bird audio as input and identifies the bird, gives bird name, its image and trivia as output.

  Input: Bird audio
  Output: Bird name, trivia

- **Bird classification based on Image:** AI-Birdie provides bird classification based on image. The application takes bird image as input and identifies the bird, gives bird name, its audio and trivia as output.

  Input: Bird image
  Output: Bird name, trivia
- **Bird suggestion based on text input:** Birds are suggested based on user input on some specific questions about bird such as size, color, location.

Input: user input to system defined questions
Output: Birds Suggestion list

- **Bird Hotspots:** AI-Birdie provides the facility of hotspots. Based on user location, the birds which are frequently found in their location are suggested.

  Input: User location
  Output: Birds suggestion list

- **Bird Injury Helpline:** AI-Birdie provides nearby veterinarian hospital helpline number and address in case a user finds a bird injured.

  Input: User location
  Output: Veterinarian hospital address and number

- **Personalized User Dashboard:** AI-Birdie provides a personalized dashboard to user. It consists of history of all the bird citing of the user.

1.3.2 Non - Functional Requirements :

- **Usability:** Anyone familiar in using simple application can operate the system since it have user friendly user interface. The system user interface is simple to understand.

- **Reliability:** The AI-Birdie can do transactions efficiently. For invalid and malfunctioned operation the system will restart in order to prevent data loss as well as safe operation within 5 seconds.

- **Performance:** AI-Birdie offers quick response.

- **User interface:** The user interface is friendly which is easy to use. And having attractive frame structure which is prepared in assumption with other related systems.

1.3.3 Materials and Methods used in Project :

- It is the user friendly application for bird watchers which helps them to identify different bird species based on audio or image clipping. It provides bird suggestion in their location which are frequently cited there. It also gives bird suggestion based on text input - various features such as size, color, area.

- The main perspective of the AI-Birdie is to contribute the birds population count. This application also provides a helpline numbers and nearby veterinary hospital address in case a bird is found injured.

  - **Software Requirement**
    → Front-End :- Flutter (iOS /Android)
    → Back-End :- CNN, ML Framework, Firebase(BaaS)

  - **Hardware Requirement**
    → The system is implemented in Intel(R) Core(TM) i7 processor with 8GB RAM,
    64 bit computer with GPU
    → AWS (IaaS)

1.4 BRIEF LITERATURE REVIEW:

1.4.1 Visual Detection:

1. **Inception v3:**

Convolutional networks are at the core of most state of-the-art computer vision solutions for a wide variety of tasks. Since 2014 very deep convolutional networks started to become mainstream, yielding substantial gains in various benchmarks. Although increased model size and computational cost tend to translate to immediate quality gains for most tasks (as long as enough labeled data is provided for training), computational efficiency and low parameter count are still enabling factors for various use cases such as mobile vision and big-data scenarios. Here we are exploring ways to scale up networks in ways that aim at utilizing the added computation as efficiently as possible by suitably factorized convolutions and aggressive regularization. We benchmark our methods on the ILSVRC 2012 classification challenge validation set demonstrate substantial gains over the state of the art: 21.2% top-1 and 5.6% top-5 error for single frame evaluation using a network with a computational cost of 5 billion multiply-adds per inference and with using less than 25 million parameters. With an ensemble of 4 models and multi-crop evaluation, we report 3.5% top-5 error and 17.3% top-1 error.

2. **Mobile net:**

We present a class of efficient models called MobileNets for mobile and embedded vision applications. MobileNets are based on a streamlined architecture that uses depthwise separable convolutions to build light weight deep neural networks. We introduce two simple global hyperparameters that efficiently trade off between latency and accuracy. These hyper-parameters allow the model builder to choose the right sized model for their application based on the constraints of the problem. We present extensive experiments on resource and accuracy tradeoffs and show strong performance compared to other popular models on ImageNet classification. We then demonstrate the effectiveness of MobileNets across a wide range of applications and use cases including object detection, finegrain classification, face attributes and large scale geo-localization

3. **Dense net:**

Recent work has shown that convolutional networks can be substantially deeper, more accurate, and efficient to train if they contain shorter connections between layers close to the input and those close to the output. In this paper, we embrace this observation and introduce the Dense Convolutional Network (DenseNet), which connects each layer to every other layer in a feed-forward fashion. Whereas traditional convolutional networks with L layers have L connections—one between each layer and its subsequent layer—our network has L(L+1) 2 direct connections. For each layer, the feature-maps of all preceding layers are used as inputs, and its own feature-maps are used as inputs into all subsequent layers. DenseNets have several compelling advantages: they alleviate the vanishing-gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters. We evaluate our proposed architecture on four highly competitive object recognition benchmark tasks (CIFAR-10, CIFAR-100, SVHN, and ImageNet). DenseNets obtain significant improvements over the state-of-the-art on most of them, whilst requiring less computation to achieve high performance.

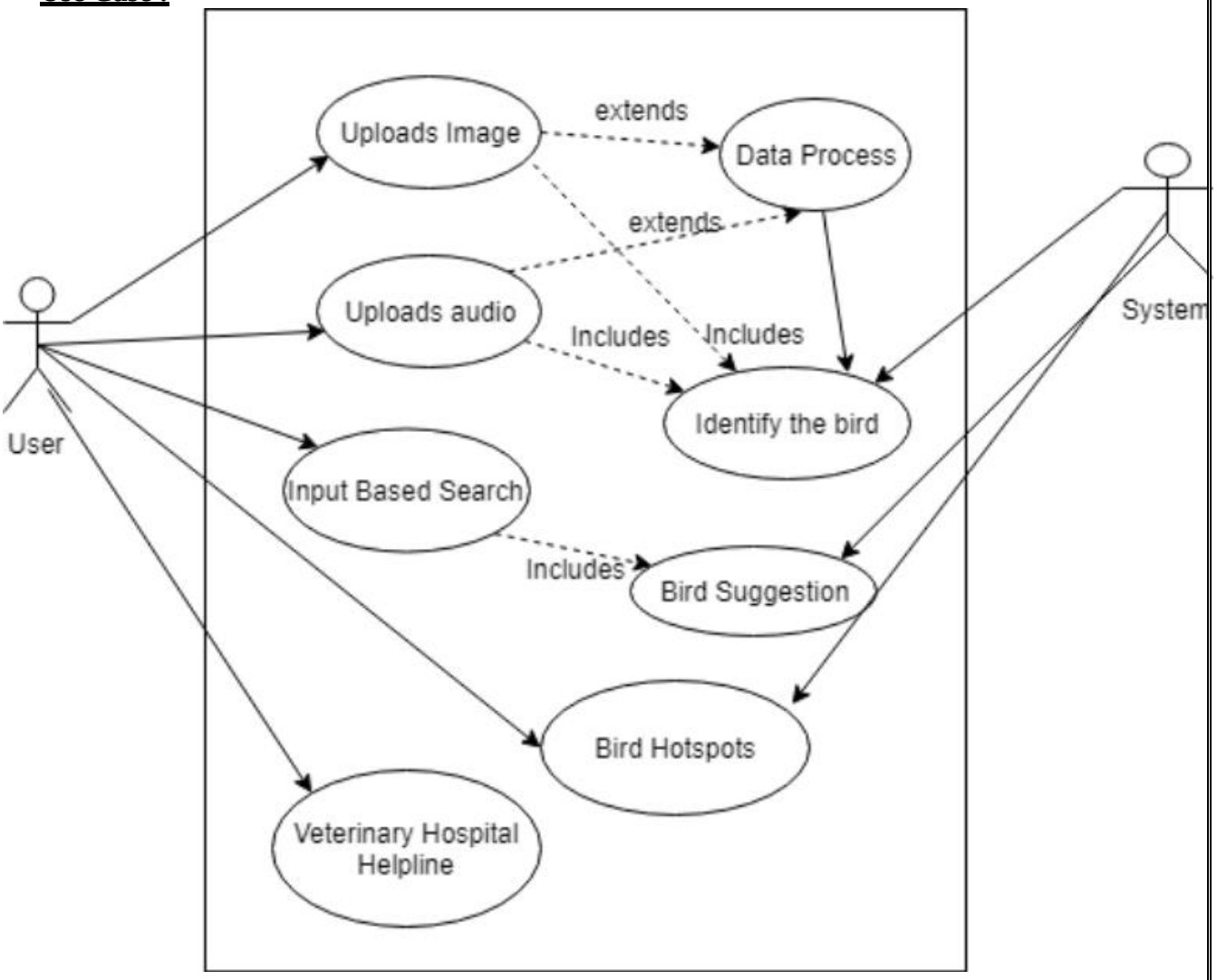
1.4.2 **Acoustic Detection** :

1. **Wave Gan:**

Audio signals are sampled at high temporal resolutions, and learning to synthesize audio requires capturing structure across a range of timescales. Generative adversarial networks (GANs) have seen wide success at generating images that are both locally and globally coherent, but they have seen little application to audio generation. In this paper we introduce WaveGAN, a first attempt at applying GANs to unsupervised synthesis of raw-waveform audio. WaveGAN is capable of synthesizing one second slices of audio waveforms with global coherence, suitable for sound effect generation. Our experiments demonstrate that—without labels—WaveGAN learns to produce intelligible words when trained on a smallvocabulary speech dataset, and can also synthesize audio from other domains such as drums, bird vocalizations, and piano.

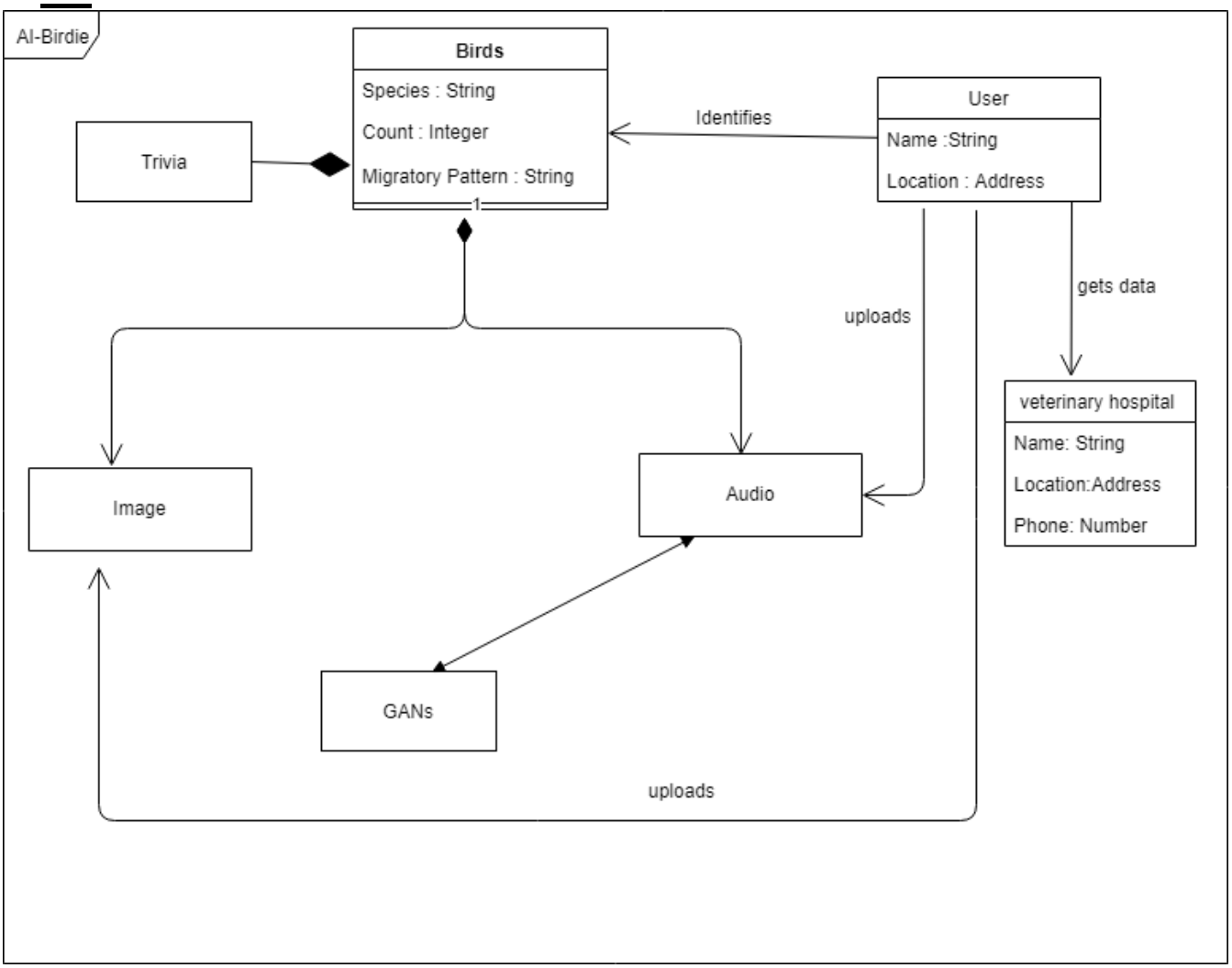
Requirement Analysis and Design

Use Case :



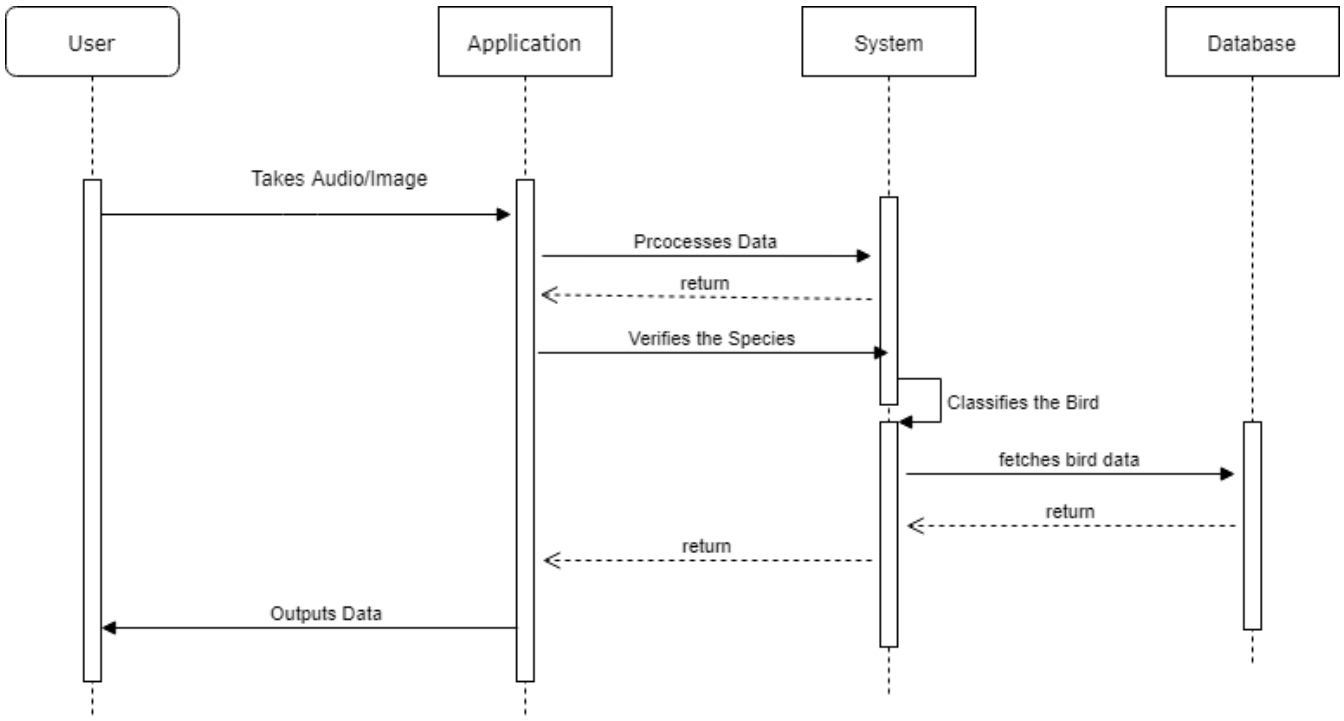
[Figure No: 1 – Use Case]

Class Diagram :



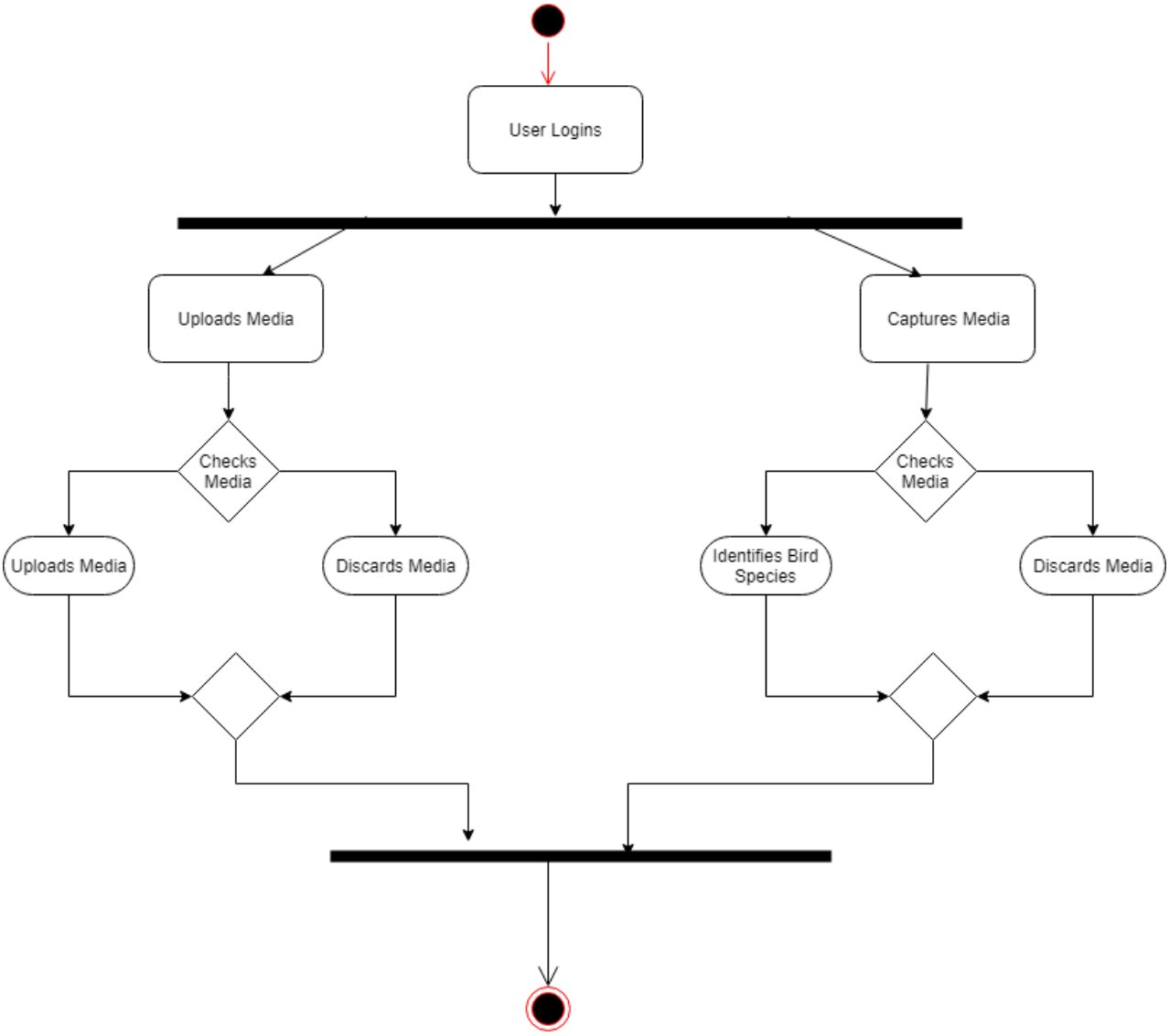
[Figure No : 2 – Class Diagram]

SEQUENCE DIAGRAM :



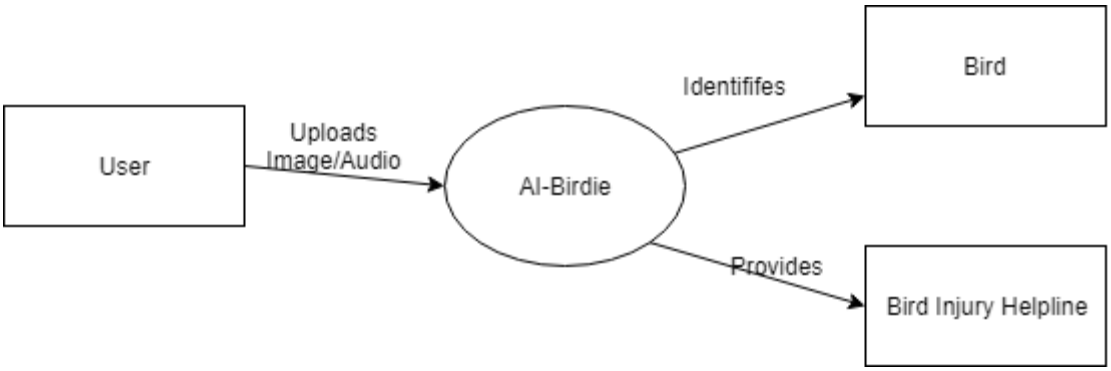
[Figure No : 3 – Sequence Diagram]

ACTIVITY DIAGRAM :



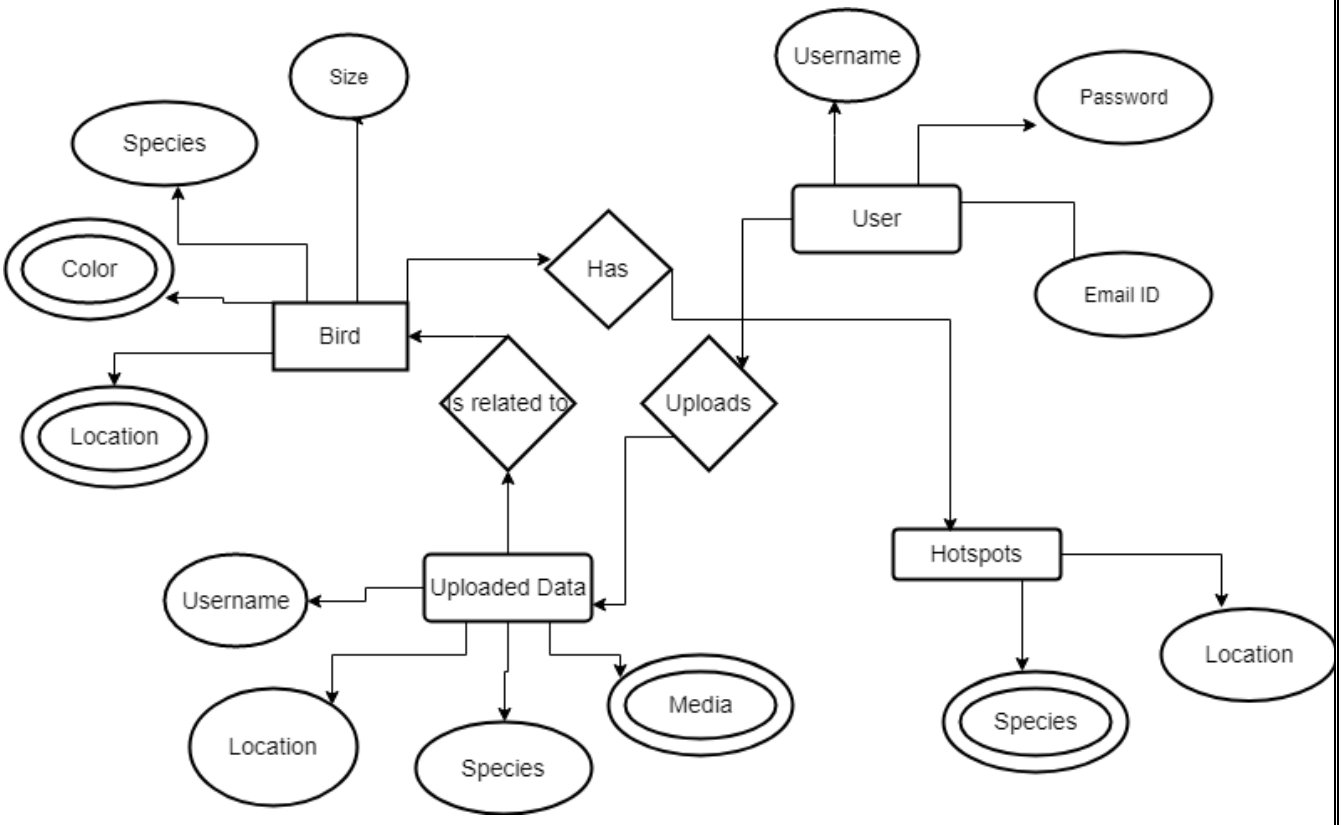
[Figure No : 4 –Activity Diagram]

Data Flow Diagram:



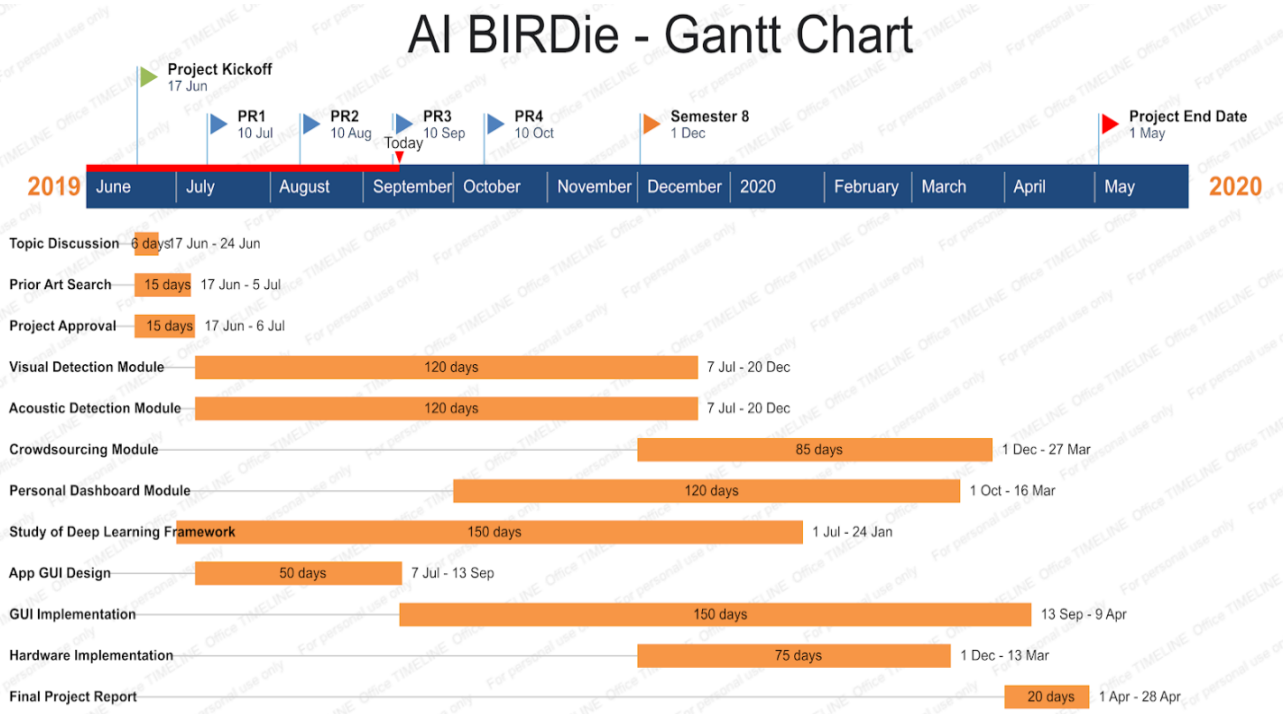
[Figure No : 5 –Data Flow Diagram]

ER Diagram:



[Figure No : 6 – ER Diagram]

Gantt Chart :

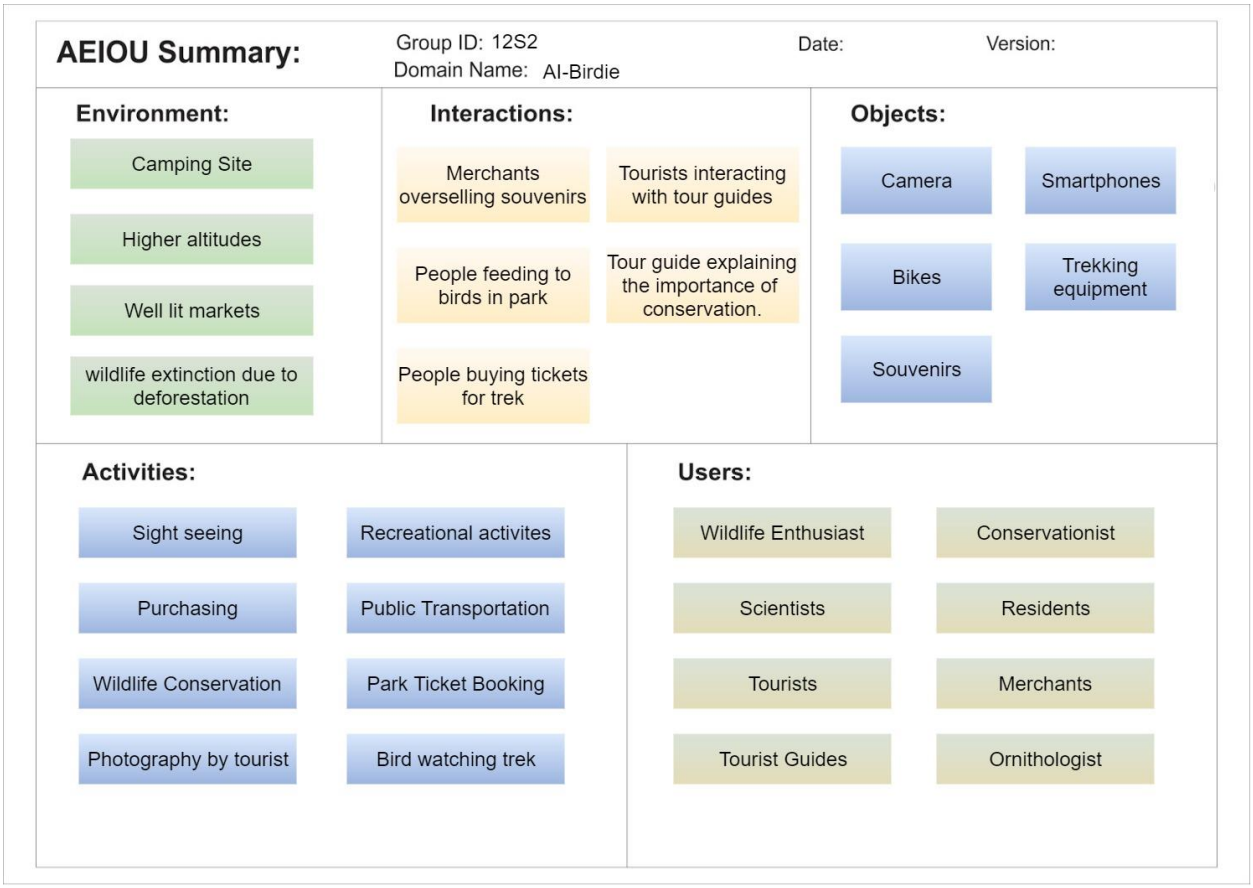

AI BIRDie - Gantt Chart
Project Kickoff
17 Jun
PR1
10 Jul
PR2
10 Aug
PR3
10 Sep
Today
PR4
10 Oct
Semester 8
1 Dec
Project End Date
1 May
2019
June
July
August
September
October
November
December
2020
February
March
April
May
2020
Topic Discussion
6 days
17 Jun - 24 Jun
Prior Art Search
15 days
17 Jun - 5 Jul
Project Approval
15 days
17 Jun - 6 Jul
Visual Detection Module
120 days
7 Jul - 20 Dec
Acoustic Detection Module
120 days
7 Jul - 20 Dec
Crowdsourcing Module
85 days
1 Dec - 27 Mar
Personal Dashboard Module
120 days
1 Oct - 16 Mar
Study of Deep Learning Framework
150 days
1 Jul - 24 Jan
App GUI Design
50 days
7 Jul - 13 Sep
GUI Implementation
150 days
13 Sep - 9 Apr
Hardware Implementation
75 days
1 Dec - 13 Mar
Final Project Report
20 days
1 Apr - 28 Apr

CHAPTER 2: DESIGN : ANALYSIS, DESIGN METHODOLOGY AND IMPLEMENTATION STRATEGY

For making our Canvases we have observed people in general from all walks of life. Few Locations we have observed are our local area, tourists point etc.

2.1 AEIOU CANVAS

It Stands for: Activity, Environment, Interaction, Object and User. This Canvas is the detailed observation of the location selected.

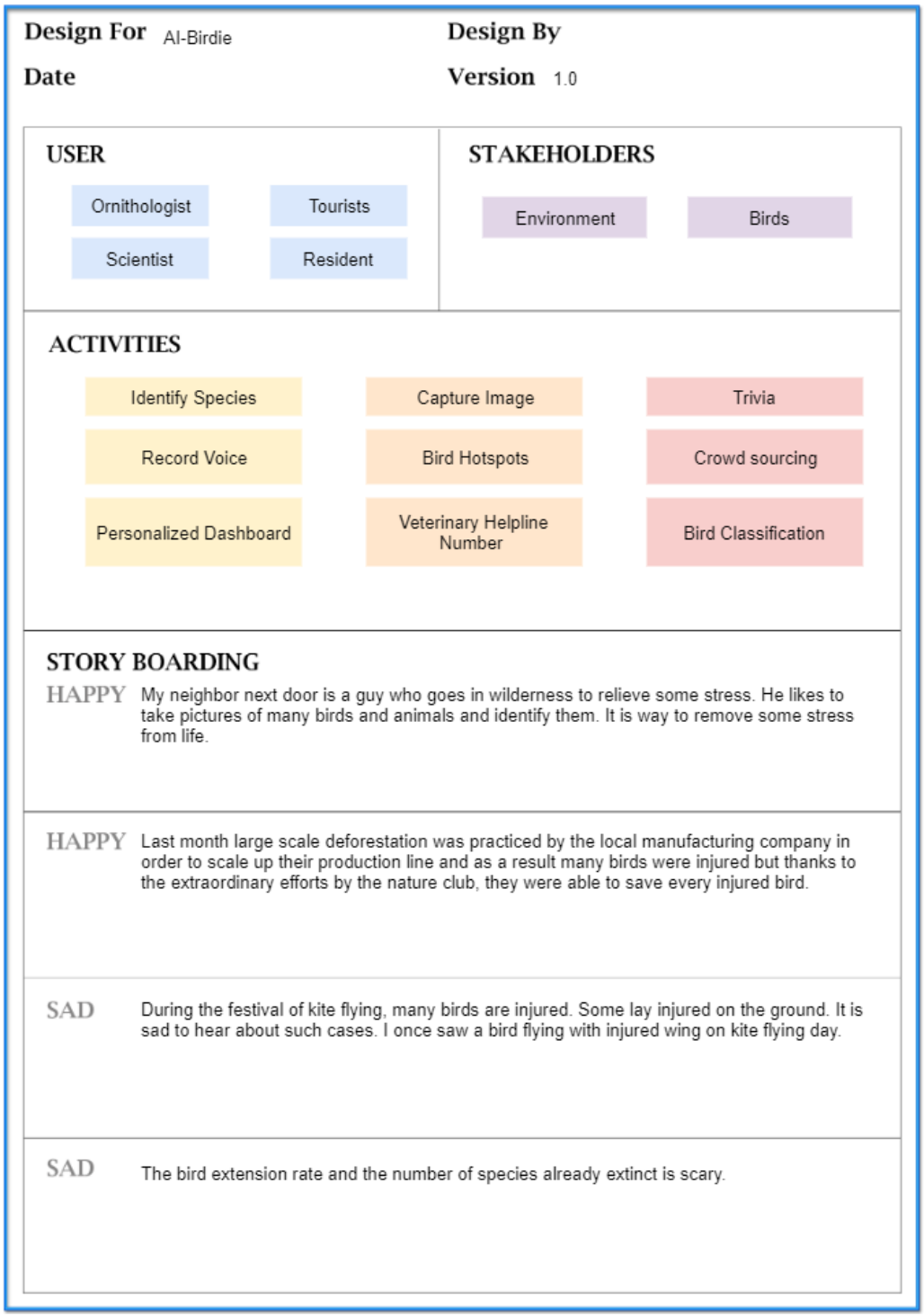


[Figure No : 7 – AEIOU Summary Canvas]

In Activities we observed that at tourist places, tourists were going on trekking trials, were snapping the nature, getting to now their environment by wildlife enthusiast, people purchasing tickets to various parks and local shops, people using the buses. The Environment was quite silent during night time and there was hustle during morning. Object involved are binoculars, smartphones, hiking equipment, camera etc. Users were tourist, tour guides, merchants, ornithologist etc.

2.2 Empathy Mapping Canvas

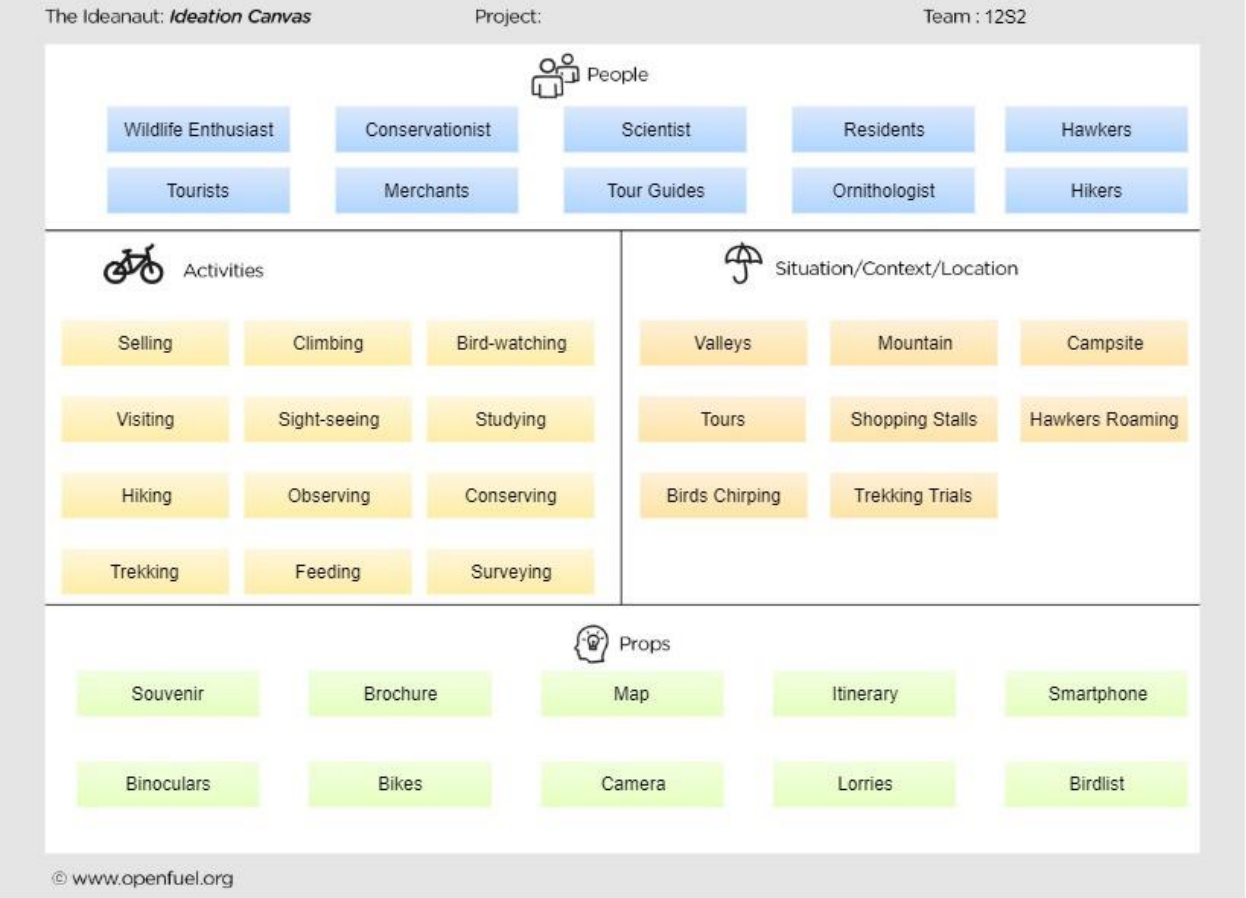
Empathy mapping helped us to look from the perspective of the prospective users which were involved in different activities and interactions.



[Figure No : 8 – Empathy Mapping Canvas]

2.3 Ideation Canvas

Ideation canvas helped us to think about different potential ideas for our project which involved multiple users. It helped us to understand different activities in a better way according to its context.

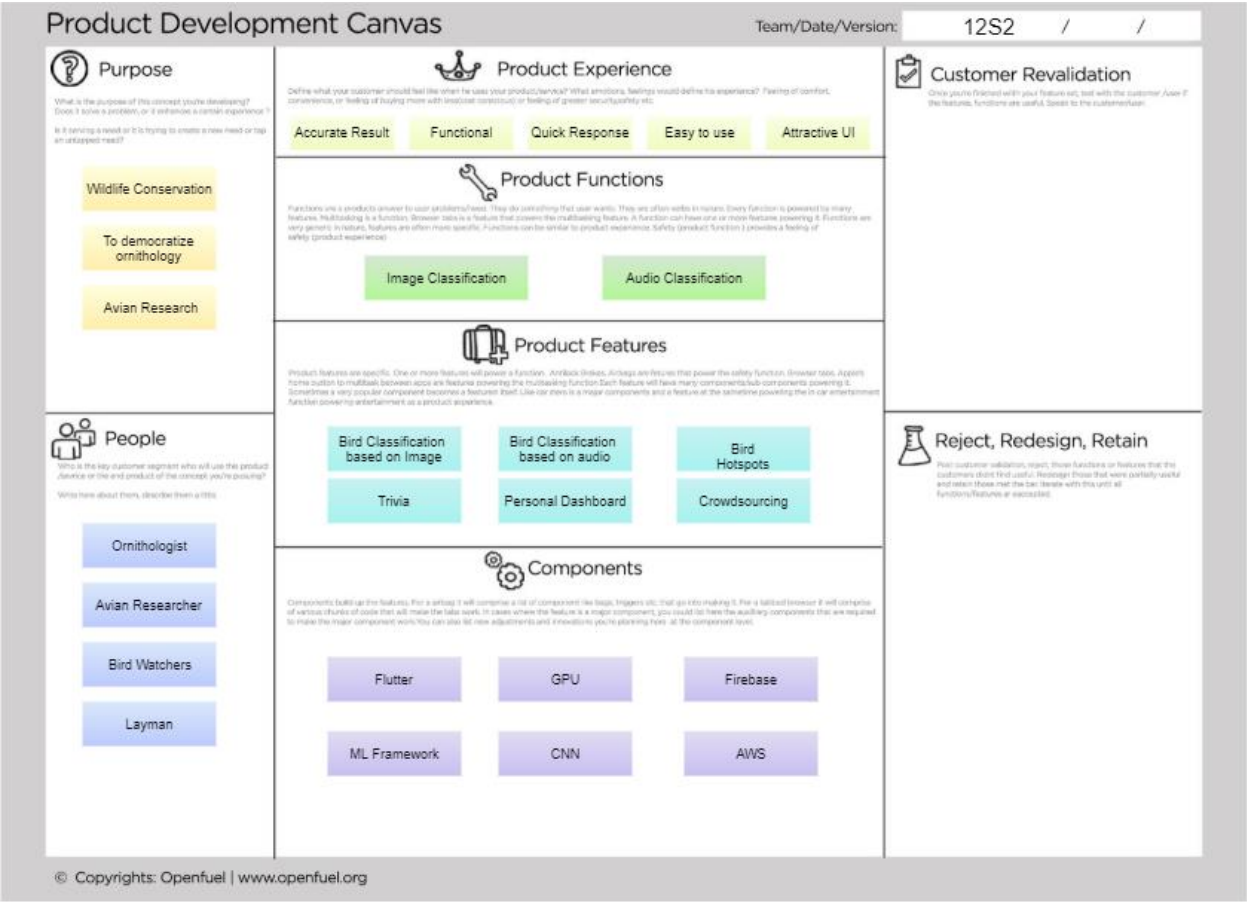


[Figure No : 9 – Ideation Canvas]

Ideation canvas made us find a common activity done by Locals, Photographers, Tourists, wildlife enthusiast etc. and it was that all these people were interested in the wildlife and plantations around them and were interested in knowing their surroundings..

2.4 Product Development Canvas

PDC Canvas is a canvas where we finally pitch our idea of bird Classification using image or audio which we got after performing all activities up till Ideation Canvas. It contains purpose, features and functions of our project along with the components involved.



[Figure No : 10 – Product Development Canvas]

The purpose of our project is to help users identify the birds they see in their surroundings. Main feature of our project is to classify the birds based on image input and audio input. We also provide a function to user to contribute image or audio to the database.

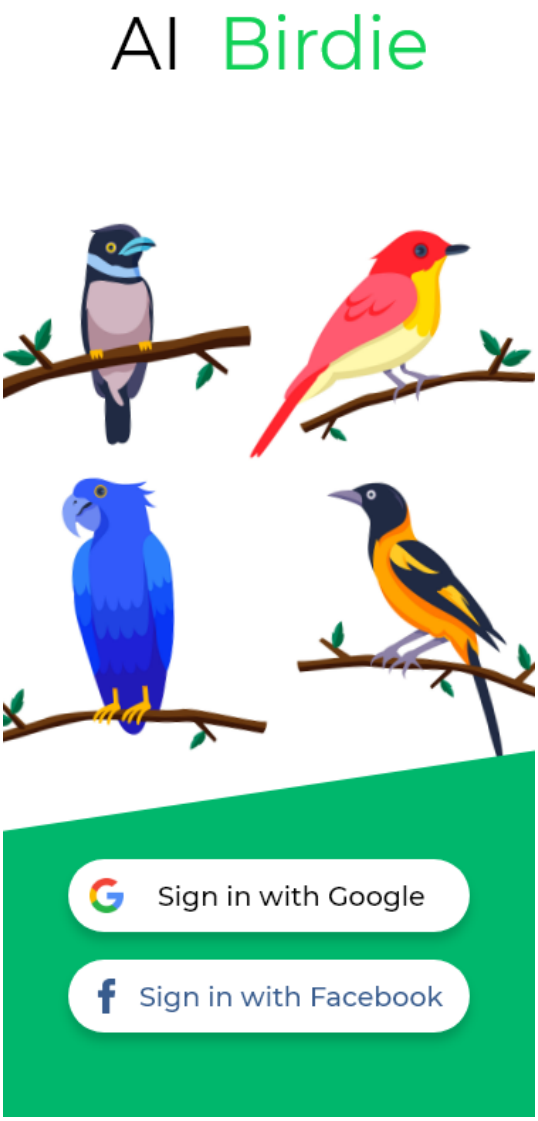
CHAPTER 3: IMPLEMENTATION

- Our Project consists of three main modules :
  - Flutter Application
  - Visual Detection
  - Acoustic Detection

Following are the screen shots of the application UI/UX developed by us till now.

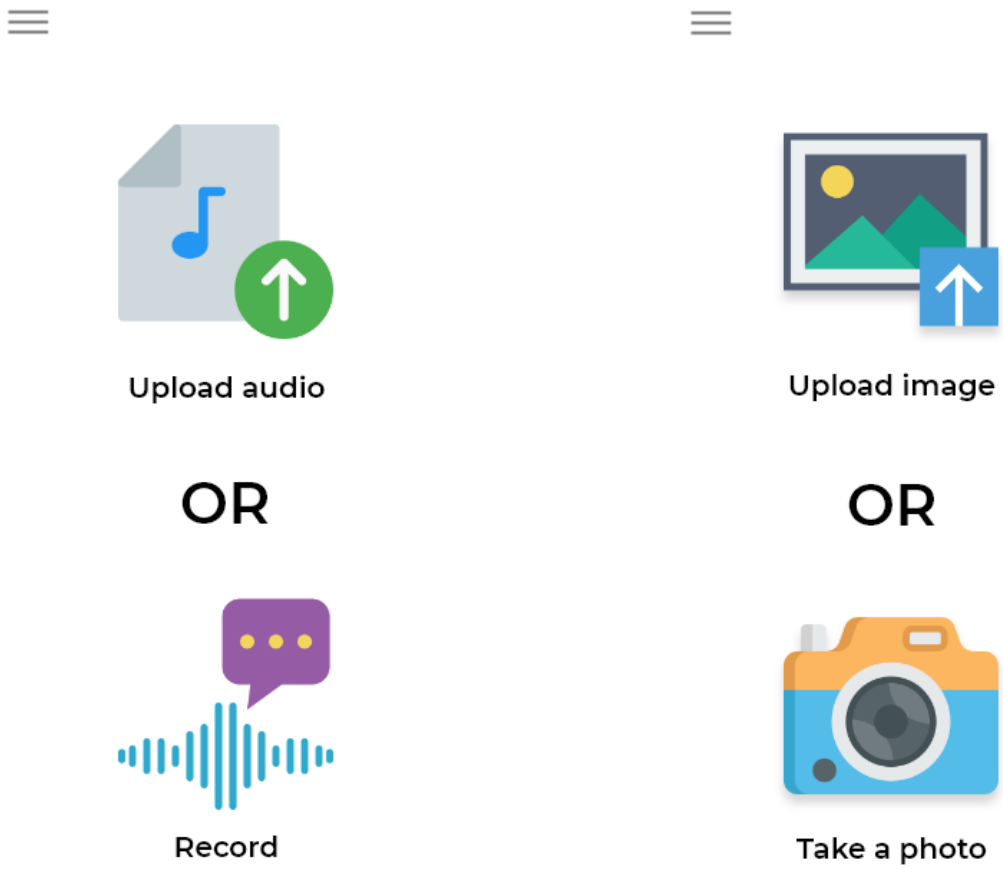
Welcome Screen

Login Screen



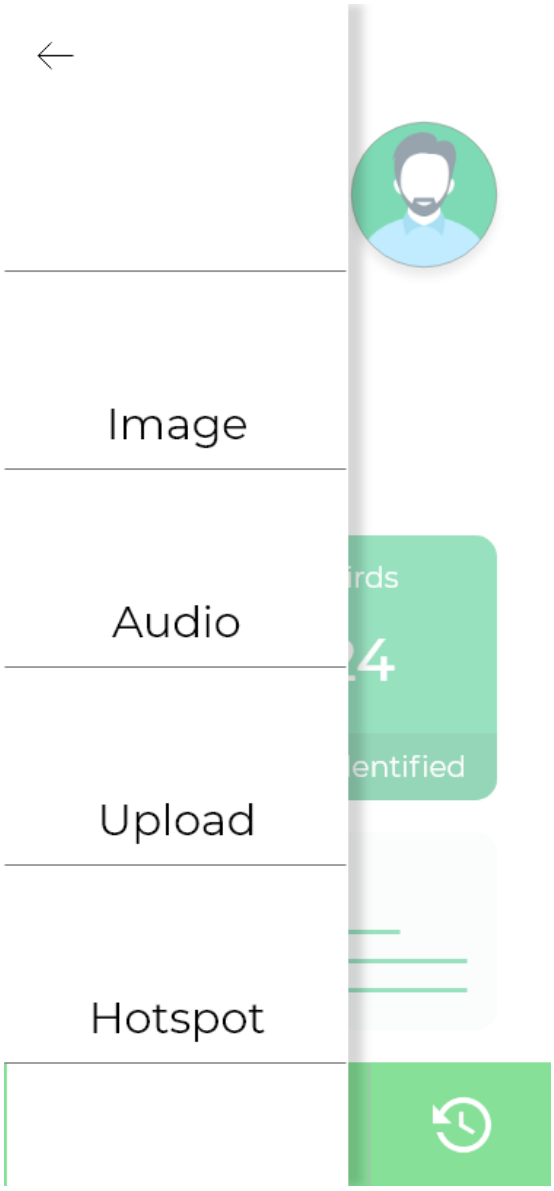
Acoustic Detection

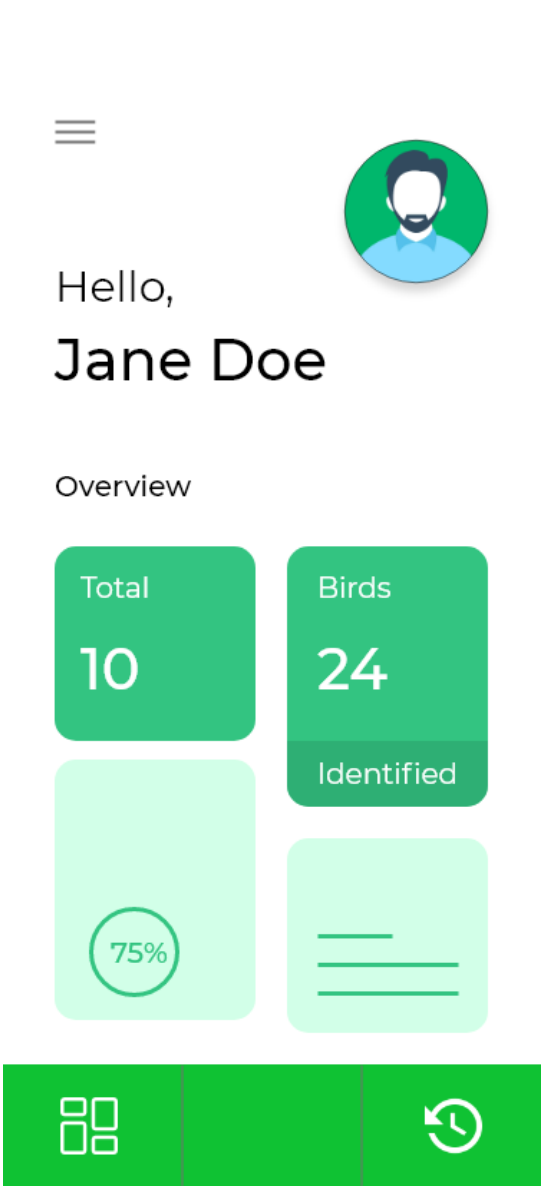
Visual Detection



- After the user logins into the application, the user is redirected to the dashboard.
- There are different functionalities available on the navigation bar, from where he/she can choose to seamlessly use that particular service.
- There are following options available in the navigation bar:
    - Visual Detection
    - Acoustic Detection
    - Upload option for crowdsourcing
    - Bird Hotspot

Navigation Bar **Dashboard**

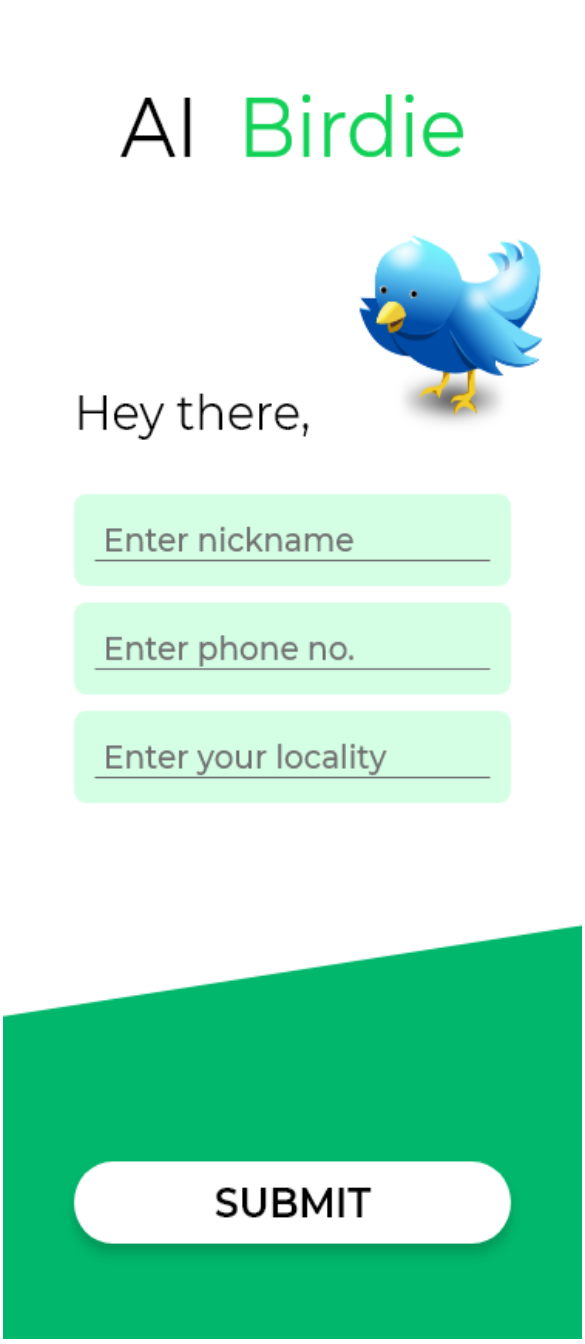




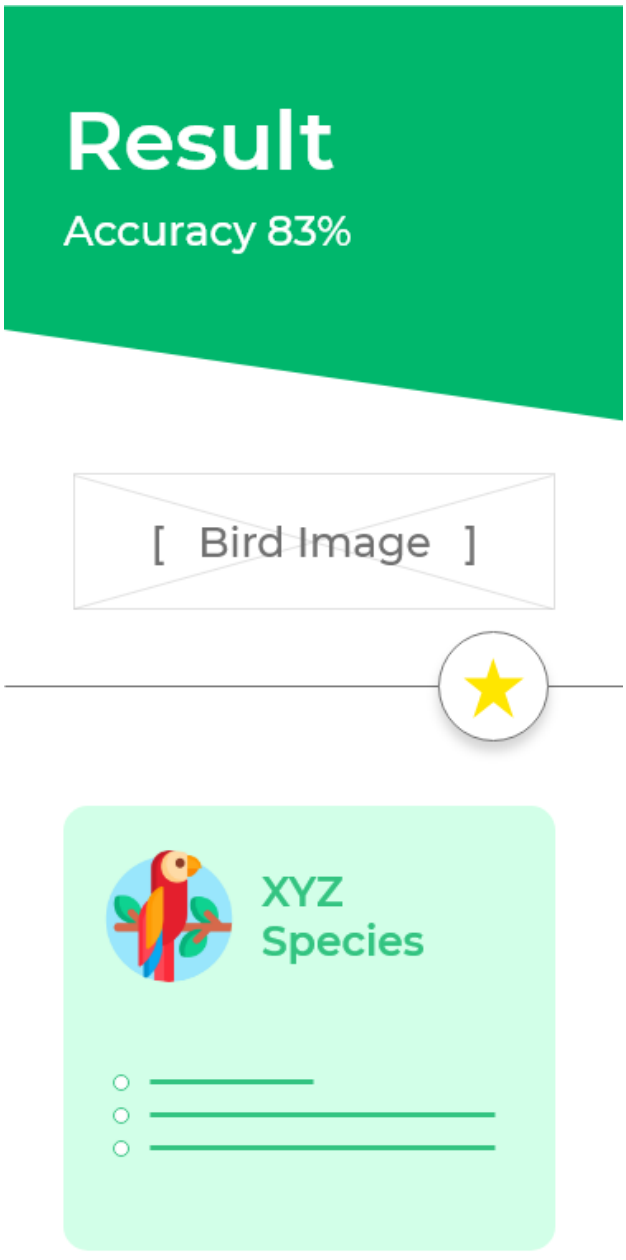
After the user has selected the particular option, he/she will get the results along with the trivia in the result page

The same will be updated in the user dashboard.

Details



Result Page



CHAPTER 4: CONCLUSION

- Mobile net has shown efficient response time as compared to dense net. How it performs on large dataset is subject further review
- WaveGan is able to generate hours of audio in just seconds and will be a perfect addition to augment the existing dataset.
- Our goal is to create a seamless User Experience for the customer and will be working on the same goal.
- The PSAR activities broaden our vision to explore new innovative ways.
- The Design Engineering Canvases helped us define our use cases better.

REFERNCES

https://birds.iitmandi.ac.in/index.php

https://birdcount.in/

http://dcase.community/documents/challenge2018/technical_reports/DCASE2018_Thakur_25.pdf

https://arxiv.org/pdf/1807.05812.pdf

Cornell lab of ornithology - https://www.birds.cornell.edu/home/

https://www.audubon.org/news/testing-out-song-sleuth-new-app-identifies-birds-their-calls

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6366782/

https://www.birdgenie.com/our-story/

https://www.nature.com/articles/d41586-019-00746-1

https://www.sciencemag.org/news/2018/07/watch-out-birders-artificial-intelligence-has-learned-spot-birds-their-songs

https://lifehacker.com/shazam-for-birds-three-apps-that-recognize-bird-calls-1797955537

https://birdwatchinghq.com/birdingapps/