Intelligent Reasoning Systems Project Proposal

Ashiwin Rajendran (A0291191E)
Jithin Krishnan (A0249481W)
Prem Varijakzhan (A0291913B)
Shankar Sai Ganesh (A0168149U)

17th March 2024

Introduction

Overview

- Introducing an advanced mobile app for revolutionizing birdwatching in Singapore.
- Utilizes machine learning for easy bird species recognition and sighting insights.
- Accessible bird identification and bird sighing predictions for enthusiasts and experts alike.

Importance

- Addresses the knowledge gap in birdwatching, enriching the experience for all.
- Harnesses Singapore's rich biodiversity for educational and conservation efforts.

Goals

- Utilize historical data, weather records and environmental data to predict bird sightings, aiding enthusiasts in optimizing their birdwatching plans.
- Enable precise, instant bird identification through image recognition technology and bird description provided by the users
- Map out prime birdwatching locations for optimal birdwatching experiences.
- Embed an LLM-powered chat feature to facilitate seamless access of the features that the App offers

Project Background / Market Context

Background

- Singapore's geographical location and lush environment make it a prime location for bird watching.
- The city attracts enthusiasts armed with binoculars and field guides.
- Over 400 bird species inhabit Singapore's urban and natural landscapes.
- Bird watchers have the opportunity to observe a diverse array of resident and migratory bird species.

Challenges

- Bird watchers in Singapore face difficulties in locating, identifying, and understanding bird species.
- Despite the abundance of birding opportunities, the vast number of bird species present poses challenges.
- Understanding the behaviours of these species can be particularly challenging.

Market Landscape

Landscape

- Singapore is a hotspot for over many bird species and offers a vibrant mix of urban and natural birding spaces.
- These spaces are enhanced by technology.
- Singapore's birding community includes strong enthusiasts and conservationists.
- Singapore presents ample market potential for innovative birdwatching solutions.
- This market potential arises from the city's status as a birding hotspot.
- There is a demand for solutions catering to both enthusiasts and conservationists.

Literature Review / Market Research

Key Players

- BirdCast (Predict only migration patterns, Dees not predict for specific locations in Singapore)
- Ebird (doesn't predict but users able to update their sightings)
- Singapore Bird Database -> No prediction but historical data of sightings in Singapore

Trends

- Increased interest in birdwatching in Singapore [1]
- Initiatives to promote conservation and data collection on birds in Singapore [2]
- Citizen Science – Bird Photographers helping in improving the database of bird information [3]

Market Demands

- No app specifically tailored to Singapore's habitats that predicts bird sightings
- Challenges highlighted by enthusiasts in locating and identifying birds in Singapore [4][5]

[1] https://www.channelnewsasia.com/commentary/singapore-bird-watching-spots-nature-park-race-sungei-buloh-2490576

[2] https://birdsociety.sg/2023/08/29/the-bird-society-of-singapore-is-officially-launched/

[3] https://lkcnhm.nus.edu.sg/wp-content/uploads/sites/10/app/uploads/2017/06/2009nis27-30.pdf

[4] https://www.birdforum.net/threads/was-looking-for-birding-in-singapore-landed-here.442034/

[5] https://www.birdforum.net/threads/guide-for-singapore-stop-over-needed.440297/

Project Scope

Sighting Prediction

Utilize machine learning models to analyze historical sighting data, offering users predictions on when and where they might see specific bird species.

Display bird-watching hotspots on embedded maps.

Bird Information

Provide detailed information about birds, including habitat, behavior, and identification tips, supplemented with maps highlighting birdwatching hotspots.

Species Identification

Implement text, and image recognition technologies to identify bird species from user submissions.

Chatbot Interface

Use the LLM API (GPT API) to facilitate natural language interactions, guiding users through information queries and data submissions.

Application Features

- Utilizes historical sighting data to generate accurate predictions on when and where specific bird species can be observed.
- Offers comprehensive information on bird species, including details on their habitats and behaviors, to educate users on where to find certain species and how to identify them.
- Incorporates maps highlighting birdwatching hotspots and habitats, aiding users in planning their birdwatching excursions more effectively.
- Utilizes Large Language Models for species identification through user-submitted text descriptions, and photos
- Integrates a chatbot interface for communicating with users in natural language, efficiently processing queries and facilitating interactions with the app.
- Ensures seamless interaction between the frontend mobile application, the Large Language Models (LLMs), and the backend application for API and trained machine learning prediction models

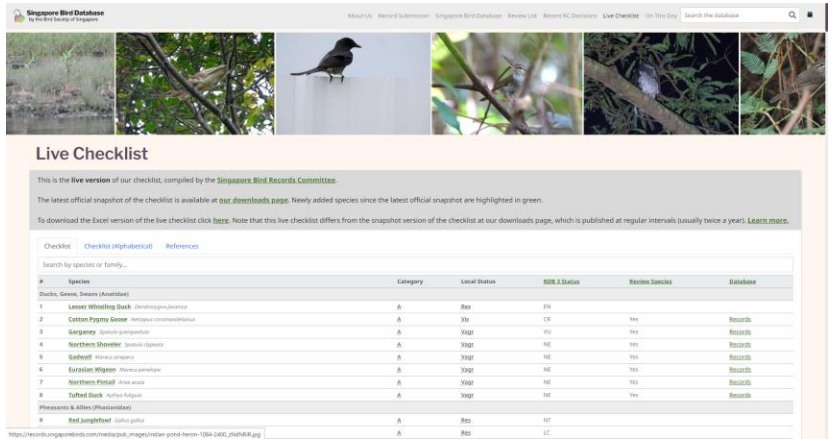
Data Collection

Bird Sightings Information

- Utilise data from the Singapore Birds Database[1], encompassing detailed text descriptions of various bird species observed in Singapore.
- Conduct web scraping to gather historical bird sightings information, extracting relevant data such as date, location, and species descriptions.

Historical Weather and Environmental Records

- Acquire historical weather and environmental records from the Meteorological Service Singapore[2].
- Employ web scraping techniques to collect data on weather conditions, including temperature, precipitation, wind speed, and other relevant environmental factors·



[1] https://records.singaporebirds.com/

[2] http://www.weather.gov.sg/home/

Data Processing

Processing Data

- Location and dates birds were spotted will be categorized according to each species of birds. Only bird species with location & date information will be used in training of the model.
- Historical weather and environmental variables will be added for bird species prediction.
- Description of species (text) and photos submitted by the users will be used for identification of birds. Details of all common birds in Singapore are to be gathered.

System Design

Android Chatbot Application

- This is the frontend that the user interacts with. It's a chatbot within an Android application where users can input their queries or data.

Django Backend Server (API)

- The Android application communicates with a Django backend server. Django is a high-level Python web framework that encourages rapid development and clean, pragmatic design. The server likely handles requests from the mobile application and communicates with the database and other services.

LLM CHATGPT API

- This is the interface for interacting with a Large Language Model, specifically ChatGPT. This API provides functionalities related to bird identification through natural language processing, where users can ask questions or submit descriptions.

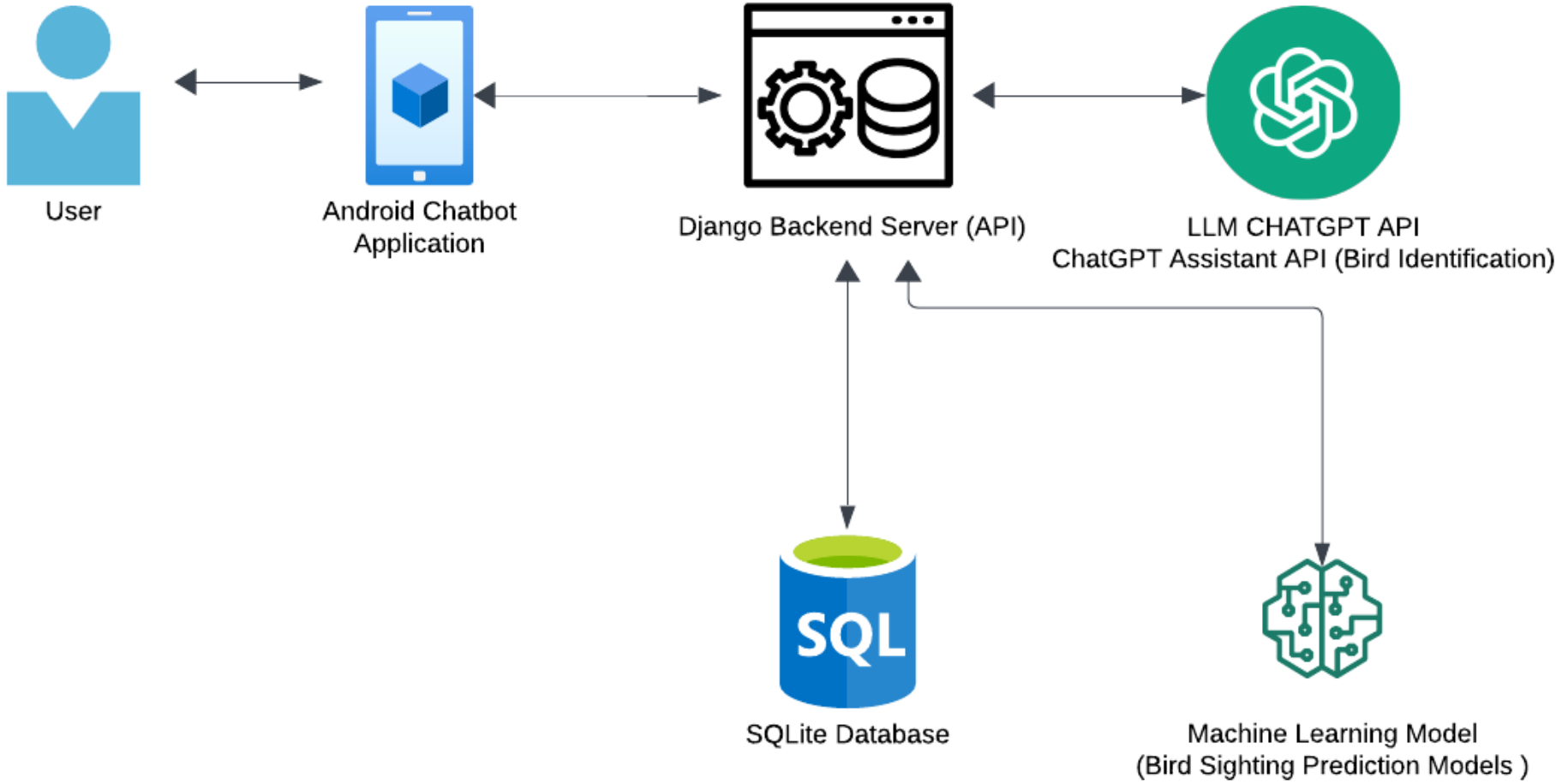
SQLite Database

- This is the data storage component of the system. SQLite is a C-language library that implements a small, fast, self-contained, high-reliability, full-featured SQL database engine. The database stores user data, bird sighting data, and other relevant information that the system uses to function.

Machine Learning Model (Bird Sighting Prediction Models)

- These are the AI models that predict bird sightings. They likely use historical data from the SQLite database to make predictions about when and where specific bird species might be observed.

Architectural Diagram



Reasoning techniques/Algorithms

Random Forest

- For predicting the occurrence of specific bird species at various locations and times.
- Capable of handling imbalanced datasets, which is beneficial when some species have fewer sightings than others.
- Its versatility allows for addressing both classification (categorizing into different species) and regression (estimating the number of sightings) problems. This makes it adept for predicting categorical variables such as species type and continuous outcomes like the frequency of sightings at specific locations.

Regression Models

- For predicting bird abundance in specific locations in Singapore for the hotspot map, Poisson Regression can be used to model the relationship between environmental factors and the total number of birds sighted. These models can provide insights into how various environmental factors and past data influence overall bird abundance.

Large Language Model (Chat GPT API)

- Bird Identification : ChatGPT Assistant API that's trained with the prepared data for text-based bird identification and ChatGPT API is to be used for image-based bird identification

Conclusion

- In conclusion, our proposal utilizes a robust combination of predictive models and advanced processing techniques to address the challenge of bird sighting prediction and identification.