Project Initiation Document

Project Details:

Project Title	Using neural networks to classify bird species from songs and calls
Student	Bradley Smith
Course	Data Science and Analytics
Project code	PJE40
Client	N/A
Supervisor	Farzad Arabikhan
Date	September 2025

1. Client / Target Audience

There is no one specific client for this project, however the target audience for the development of this project would be people who have a mild interest in birding and/or those interested in using neural networks for classification of audio files.
A potential further target audience would be software developer(s) who would be developing an application to help identify birds based on a specific call.

2. Project background and problem to be solved

In the modern world, many people are disconnected from nature, with the fast growth of technology and rise in convenience. When people take a step back and try to re-establish their roots with nature, one common way is to take a walk in nature. The most common sound people hear are the sounds of many birds calling. However, these birds are impossible to notice, be it due to their size, speed, or even their timid nature, leading people to never actually see these birds. As a result, one way people are able to properly identify these birds is through “banding” these birds, by catching them in nets, potentially causing unnecessary stress to the birds, especially if just for the purpose of identification of a species. My goal is to reduce this approach to identifying birds, saving birding enthusiasts/organisations lots of time, effort and resources, allowing them to identify birds in a way that causes minimal stress to these birds.

3. Project aim and objectives

The final goal of this project is to have a machine learning model that will accurately classify a testing set of bird calls and songs.

The aim of this project and the necessary objectives are:

- Create a machine learning model that can classify over 100 species of bird with an F1 score of at least 0.7. This is because of the Kaggle dataset leaderboards showing performance metrics in the ranges of 0.65-0.8 for F1 scores.
- If the above is not possible, create a machine learning model that can classify up to 100 species of bird with an F1 score of at least 0.9. This is due to having a smaller number of species to classify.
- To do these, I will:
    - Create multiple types of machine learning models and compare the performance metrics of each.
    - Combine data from existing BirdCLEF and Cornell Lab of Ornithology datasets and collate the 5 datasets from 2020-2024 to investigate a larger range of data.
        - I will not include any duplicate audio files to skew learning capabilities and reduce redundancy.

4. Project constraints

One constraint for the project will be the spectrum of bird species that the model is able to classify. If all the datasets that my model is trained on do not include calls from a specific species of bird, then using any audio recordings of said species of bird will be pointless.
One limiting factor for this project is also that there is no budget. If there are any audio recordings or databases locked behind a one-time payment, or a subscription, that data will not be available for analysis, even if it could have been vital for this project's potential.
One potential constraint would be the combination of technology that the machine learning model is developed on, as well as the deadline of the project. If the model is trained on a less powerful machine, then the time taken to fully train the model over one epoch could take considerably longer. The fastest method of training a complex ML model - GPU acceleration - can only be as fast as the most powerful workstation available to me. Given the deadline of the project, any time I need to fine-tune the model to produce a single output could change from minutes per epoch, up to hours (if the model is complex enough.)

5. Project management

For my project, the Agile methodology seems to best apply to the development of my model. However, I will be also planning stages of development which will, within themselves, have Agile development aspects. I will be splitting the project into phases of development, reiterating each phase until I am satisfied, and only once I am satisfied with the end result, I move on to the next phase. This will likely be split up into sections such as: data collection, data cleaning, data preprocessing, machine learning model development, and finally model testing. With the final two phases, I will go back and forth between the two to ensure my model accurately meets testing criteria, fine-tuning the model each time testing results are not accurate enough. I aim to dedicate a minimum of 2 weeks for each phase, with the final two phases requiring most of my time.

6. Tasks and timescales

No	Stage	Dates	Main Tasks
1	Project Initiation	22nd Sept 2025 - 1st Oct 2025	Establishing supervisor and submitting PID.
2	Literature Review	22nd Sept 2025 - 19th Dec 2025	Reading and critically reviewing all relevant literature to the subject.
3	Ethics Form Application	8th Oct 2025 - 5th Nov 2025	Writing and submitting ethics form.
4	Satisfactory Progress Presentation	17th Nov 2025 - 8th Dec 2025	Establishing development progress for the model.
5	Field Data Collection	10th Nov 2025 - 25th Feb 2026	Collecting audio recordings of wild birds to test any/all models.
6	Dataset Handling	6th Oct 2025 - 24th Oct 2025	Procuring, cleaning and exploring any relevant datasets.
7	Model Prototype	24th Oct 2025 - 20th Nov 2025	Developing a small-scale model prototype as proof of concept.
8	First Full Model Iteration	24th Nov 2025 - 15th Dec 2025	Develop a full-scale model to test and assess capabilities.
9	Critical Analysis #1	15th Dec 2025 - 19th Dec 2025	Thorough analysis of all previous work mentioned above, detailing any necessary changes to be made.
10	Second Model Iteration	22nd Dec 2025 - 12th Jan 2026	Redevelopment of the model following critical analysis #1. Training and re-testing.
11	Critical Analysis #2	12th Jan 2026 - 16th Jan 2026	Analysis of differences between first and second iteration of models, potential investigation into other machine learning models.
12	Further Developments	19th Jan 2026 - 20th Mar 2026	Repeats of steps 10-11 or complete redevelopment of the model if applicable until a final model is established and tested.
13	Report Submission	19th Jan 2026 - 6th May 2026	Writing of dissertation up to and including submission.

7. Facilities and resources

For this project, there are various pieces of hardware and software that I will need to complete this to the best of my ability.
To facilitate the flexibility of my working locations, any machines should have Git, a version control architecture that syncs workstations to allow for up-to-date progress on any machine that is connected to a repository.
All the software/libraries I plan to use are available on all major operating systems (Windows, Mac, Linux), so operating system differences are not an issue.
In terms of hardware requirements, I require a computer that has capability of developing machine learning models, ideally with the use of TensorFlow, which is a library which can utilise a machine's GPU to greatly accelerate learning/training of a machine learning model.

- I have constant access to a PC at my home that satisfies these requirements. The hardware specification is as follows:
    - Memory: 32GB DDR4-2133
    - CPU: Intel Core i9-9900k (8 cores, 16 threads)
    - GPU: NVidia GeForce RTX 3060
    - Storage: 2TB NVME M.2 drive, 1TB NVME M.2 drive, 1TB SATA HDD, totalling 4TB of total storage.
- When I am not at home, development can be resumed via PCs at the University, where the hardware specifications of workstations are unknown, due to the variation among locations containing PCs. These PCs are not available 24/7, due to the buildings closing overnight.

In regards to datasets, I will be mainly training/testing my model on the Cornell Lab of Ornithology competition datasets on Kaggle. The below table describes the datasets.

https://www.kaggle.com/organizations/cornelllabofornithology/competitions

Dataset/Competition	Dataset Size (GB)	Total Audio File Count	Total Unique Classes
Cornell Birdcall Identification: https://www.kaggle.com/competitions/birdsong-recognition	25.35	21,375	264
BirdCLEF 2021 - Birdcall Identification: https://www.kaggle.com/competitions/birdclef-2021	42.18	62,874	397
BirdCLEF 2022: https://www.kaggle.com/competitions/birdclef-2022	6.61	14,852	152
BirdCLEF 2023: https://www.kaggle.com/competitions/birdclef-2023	5.29	16,941	264
BirdCLEF 2024: https://www.kaggle.com/competitions/birdclef-2024	23.43	24,459	182
Combined (Removing duplicates, best approximation)	75.77	112,891	992

8. Project risks

No	Description	Likelihood	Impact	Mitigation/Avoidance
1	Storage failure	Low	Loss of data for the project on a given system, resulting in hours/days worth of time loss.	Use GitHub to store progress in the Cloud, and keep physical copies on multiple systems.
2	Unavailability of participants to procure testing data.	Medium	Lack of fresh testing data reduces real world applicability of the model. Potentially affecting final testing accuracy.	Personally seek testing data where possible, or find files online of already-classified birds to test the model on.

9. Project deliverables

Outside of my final report document, other files that will be delivered are any procured audio files by myself or participants, Python notebook(s) containing the process of developing this model. There will also be a file for each phase's model iteration, stored in a file format native to TensorFlow. There will also be anonymised documents from each participant stating where and when they recorded the audio files, for how long, and any additional notes to add to these recordings.

10. Research

I will do secondary research on many topics, including the ethics/safety of bird catching, whether any psychological stress is put on these birds who get caught in such nets, as well as other methodologies to identify birds without physically catching and temporarily restraining them.
For my model development, I will also do secondary research on data analysis on audio files, neural network optimization techniques with big datasets, as well as appropriate methodologies for classification analysis with hundreds of output nodes.
I do not plan to do any primary research.

11. Legal, ethical, professional, social issues

Legal:

- When myself or participants are recording any audio files for classification analysis, we will ensure we adhere to all laws regarding trespassing, nature reserves, as well as ensuring all audio files are recorded for public use, and do not contain any third party human voices, as there could be a potential violation of audio file usage without any third party's permission.

Ethical:

- One potential ethical issue with this project is that when recording any audio files of birds, we must take care not to cause any psychological stress to the birds or any other wildlife in the area, as the aim of this project is to identify birds without causing any stress at all.
- When storing information on the audio files, any and all information must be handled such that any sensitive information cannot be obtained, such as extremely detailed location information, and that all participant information is anonymised such that they cannot be identified from the files.

Professional:

- Any data or information submitted in my deliverables must be fully secure and not exploitable or contain malicious code. I aim to source all of my data from verifiable sources, and thoroughly check for any abnormalities.
- One problem is the accuracy of individually sourced data of bird songs. Unless we already know which bird is making the call, it would be difficult/impossible to determine if the model can accurately predict which bird is making the call. If this is the case, professional birders will be consulted and will help classify the bird call. If this is not possible, the data will not be used for testing.

12. Supervisor meetings

Upon discussion with my supervisor, we plan to have weekly meetings at 12:30pm on Fridays, in person. If it is not possible to have the meetings in person, they could be rescheduled or they will be on a video call instead.

13. Declarations

Please tick the following declarations:

Usage (optional)	
☑	I give permission for this document to be made available to other students as examples of previous work.
Authenticity of work	
☑	I confirm that I have read the University rules in respect of plagiarism and student misconduct.
☑	I understand that if I use work from an external source, I must reference and cite the source in any work that I produce.
Ethics	
☑	I understand that ethical approval will be needed for this project regardless of whether I carry out primary research or not.
☑	I understand that if I do conduct primary research, I must include all raw data and research documentation in my final project report.
☑	I understand that not including the raw data can mean potentially failing the project.
☑	I understand that not obtaining ethical approval is grounds for academic misconduct and possible project failure.
Use of AI Artificial Intelligence (AI), such as ChatGPT, can be used as a tool to help assist and inform the initial development of your work but it is not a replacement for your own critical thinking and analysis. We require that your work is your own original content, demonstrating your knowledge, skills, and critical thinking abilities. To that end, AI-generated content must not be included in any work you submit for assessment unless suitably referenced. Not doing so is a breach of the University's Academic Regulations as outlined in the Student Conduct Policy and constitutes an assessment offence.	
☑	I have read the University AI policy and if relevant, I will acknowledge the use of AI in any work I produce for this project with appropriate referencing.
☑	I understand that not appropriately referencing the use of AI is an assessment offence equal to plagiarism and may result in a penalty which remains on my student record.

Name: Bradley Smith
Date: 23rd September 2025

Appendix A: Gantt chart

The following is my project development timeline. The full image can be found in my GitHub repository. There is also a ".gantt" file which can be opened online in the Online Gantt website.