**Team 1 Proposal: Project 3 – Image Classification Birds**

Team Members:

* Danielle Dejean
* Baruch Gottesman
* Kaidon Kennedy
* Carolyn Scheese
* Harpreet (Monty) Singh

Data Set:

* As a team we discussed various options, finally deciding on an image classification problem. We decided to see if we could do a prediction model for bird species.
* As a team we settled on 2 different data sets. We plan to use the first data set to do our initial project. The data sets are very similar, so if we have sufficient time, we may use the second data set to repeat our work and compare the results.
* Winged Wonders: The Birds-200. Kaggle.
* <https://www.vision.caltech.edu/datasets/cub_200_2011/>
* Bird Species Classification 200 Categories. Kaggle. <https://www.kaggle.com/datasets/kedarsai/bird-species-classification-220-categories>
* We aim to use data scraping the web or certain websites, for the “fun facts” about the various species.

Communication:

* We created a Slack Channel named *project-3-group-1* to help us communicate effectively.
* We set up a GitHub project (project3\_imageclassification) with branches.

Plans for Working with the Data

* Create an .ipynb file to clean the data.
* Use a “Supervised ML” model because there is a column that identifies which bird species using classification methods.
* Thus far we have been familiarizing ourselves with various datasets; settling on 2 from Kaggle that have 11,000+ images and 200 species.
* We plan to use and apply the following libraries: tensorflow, keras, tensorflow.keras.optimizers import Adam, sklearn, OneHotEncoder, EarlyStopping, matplotlib.pyplot, seaborn, pandas, numpy, gradio.

Project Output

* We plan to create an interactive Gradio app which users can use to predict the species and include a “fun fact”.
* Time permitting, this project may be expanded to encompass other ideas that may come to mind.

Presentation Planning

* We will use PPT for our final presentation.