MI1 - Project 2: Sea Monsters Checkpoint
Group: The Sea Monsters
          Group Leader: Sharaf Tariq   Group Members: Alexa Owen, Audrey Himes
DS 4002
March 17, 2023

Narrative: Since the start of the COVID-19 pandemic, the number of people spending time outdoors has dramatically increased [1]. Whether they are visiting national parks or taking hikes through a neighborhood, people are enjoying nature now more than ever. In order to be respectful towards the land, it is important to have some understanding of the flora and fauna with which you interact. So, what do you do if you encounter some type of unknown plant while camping or on a hike? At one point in time it may have been sufficient to search for characteristics of this plant and take a half-hearted stab at identifying it. However, it would be much more efficient (and safe) to simply snap a photo of the plant and be given an identification on the spot. Turning plant identification for the average non-botanist into a definitive science rather than a guessing game can make nature more educational, safe, and enjoyable. We are seeking to create a model that can do just this. We want to contribute in a positive way to America's newfound love of the outdoors and to help people safely learn and understand the nature within which we coexist.

Hypothesis: The model will correctly be able to identify the type of plant based on an image of its leaf with an accuracy of 0.75 or higher.

Research Question: Can different types of plants be classified solely based on images of their leaves?

Modeling Approach: We will be using a Keras model to conduct image classification. Keras is an API for Python on TenserFlow [2]. The model will take JPEG image files from the disk and assign them to different classifications based on features in the images. We will have a training dataset with input images with known classifications, then test the precision of the model on another subset of the data that has not yet been used [3]. This algorithm will help us develop and train the machine learning model with very few lines of code [4]. We will be able to create a simple algorithm that is highly efficient and easy to use for image classification of plants.

References:

[1] R. Kestenbaum, "The Pandemic Has Forever Changed The Active Outdoor Business," *Forbes,* para. 1 & 2, Aug. 18, 2021. [Online]. Available: https://www.forbes.com. [Accessed Mar. 16, 2023].

[2] F. Chollet, "Image classification from scratch," *keras.io*, para. 1, Nov. 10, 2022. [Online]. Available: https://keras.io/examples/vision/image_classification_from_scratch. [Accessed Mar. 14, 2023].

[3] TenserFlow, "Image classification," *tensorflow.org*, para. 1, Dec, 15, 2022. [Online]. Available: https://www.tensorflow.org/tutorials/images/classification. [Accessed Mar. 14, 2023].

[4] B. G. Regmi, "Image Classification with tf.keras (Introductory Tutorial)," *Medium*, para. 1, Jan. 11, 2021. [Online]. Available: https://medium.com/analytics-vidhya/image-classification-with-tf-keras-introductory-tutorial-7e0ebb73d044. [Accessed Mar. 15, 2023].

https://www.forbes.com/sites/richardkestenbaum/2021/08/18/the-pandemic-has-forever-changed-the-active-outdoor-business/?sh=1d896176674f

https://www.kaggle.com/datasets/kritikseth/fruit-and-vegetable-image-recognition
https://www.tensorflow.org/tutorials/images/classification
https://keras.io/examples/vision/image_classification_from_scratch/
https://keras.io/about/

Resources:
https://keras.io/examples/vision/image_classification_from_scratch/
https://www.tensorflow.org/tutorials/images/classification
https://tensorflow.rstudio.com/tutorials/keras/classification

Data:
https://www.kaggle.com/datasets/mdwaquarazam/microorganism-image-classification

https://www.kaggle.com/datasets/aseemdandgaval/23-pet-breeds-image-classification

https://www.kaggle.com/datasets/alxmamaev/flowers-recognition

https://www.kaggle.com/datasets/shaunthesheep/microsoft-catsvsdogs-dataset

https://www.kaggle.com/datasets/amandam1/120-dog-breeds-breed-classification

https://www.kaggle.com/datasets/miljan/stanford-dogs-dataset-traintest

https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri

https://www.kaggle.com/datasets/ihelon/lego-minifigures-classification

https://www.kaggle.com/datasets/moltean/fruits