

## PART TWO (IMAGE CLASSIFICATION)

### MOTIVE

- Creating a dataset to be used in a machine that would help us determine population patterns and behavioral attributes of different insects.

### THE TASK

- To create and prepare a dataset that can be used to train a model for image classification.

### STEPS INVOLVED

#### I. **Data collection:**

- We used the *bing-image-downloader* module in python to collect images of various known insects.
  - We extracted insect images using six different names to ensure diversity.

#### II. **Preprocessing and feature engineering:**

- a. we resized all the images to one scale to ensure uniformity.
- b. we performed *min\_max* normalization since the color patterns were very spread.

#### III. **Labeling:**

We labeled with focus on two things:

- a. Class
  - We put different insects into different folders which would be their classes whereby the names of the insects are set as class names.
- b. Object
  - We created object labels by annotating each image using bounding boxes and the names of the insects but as in this case as the object names.

### CHALLENGES AND SOLUTIONS

- We had challenges using python code separate the images into various class folders. We instead used manual moving of files without code.
- We faced challenges where a much more sophisticated polygon was needed. We resorted to using the rectangular bounding box instead.

