**PROJECT TITLE: Digital Naturalist- AI**

**Enabled Tool For Biodiversity Researcher's**

**TEAM ID: PNT2022TMID14372**

**DATE:10/10/2022**

**Problem Statement:** When venturing into the woods, field naturalists usually rely on common approaches like always carrying a guidebook around everywhere or seeking help from experienced ornithologists. There should be a handy tool for them to capture, identify and share the beauty to the outside world.

## LITERATURE SURVEY

### Survey 1: Biodiversity

Biodiversity is a term used to describe the enormous variety of life on Earth. It can be used more specifically to refer to all of the species in one region or ecosystem. Biodiversity refers to every living thing, including plants, bacteria, animals, and humans. Scientists have estimated that there are around 8.7 million species of plants and animals in existence. However, only around 1.2 million species have been identified and described so far, most of which are insects. This means that millions of other organisms remain a complete mystery

### Survey 2: Importance of Datasets in Deep Learning and AI Research

Nowadays, we have ample resources where we can get datasets on the internet either open-source or paid. As you know data collection and preparation is the crux of any Machine Learning project, and most of our precious time is spent on this phase.

To solve the problem statements using Deep Learning, we have two choices. Either we use the existing datasets or create a new one. For a highly specific problem statement, you have to create a dataset for a domain, clean it, visualize it, and understand the

relevance to get the result. However, if the problem statement is common, you can use the following dataset platforms for research and gather data that best suits your requirements.

## **Survey 3:** Train and Test datasets in Deep Learning

Deep Learning is one of the booming technologies across the world that enables computers/machines to turn a huge amount of data into predictions. However, these predictions highly depend on the quality of the data, and if we are not using the right data for our model, then it will not generate the expected result. In Deep learning projects, we generally divide the original dataset into training data and test data. We train our model over a subset of the original dataset, i.e., the training dataset, and then evaluate whether it can generalize well to the new or unseen dataset or test set. ***Therefore, train and test datasets are the two key concepts of Deep learning, where the training dataset is used to fit the model, and the test dataset is used to evaluate the model***.

For image processing regarding deep learning is more effective.

## **Survey 4: Convolutional Neural Network's (CNNs)**

Deep Learning – which has emerged as an effective tool for analyzing big data – uses complex algorithms and artificial neural networks to train machines/computers so that they can learn from experience, classify and recognize data/images just like a human brain does. Within Deep Learning, a Convolutional Neural Network or CNN is a type of artificial neural network, which is widely used for image/object recognition and classification. Deep Learning thus recognizes objects in an image by using a CNN. CNNs are playing a major role in diverse tasks/functions like image processing problems, computer vision tasks like localization and segmentation, video analysis, to recognize obstacles in self-driving cars, as well as speech recognition in natural language processing. As CNNs are playing a significant role in these fast-growing and emerging areas, they are very popular in Deep Learning.

# References :

1. Grokking Artificial Intelligence Algorithms by Rishal Hurbans published by Manning Publications

2. Deep Learning From Scratch: Building with Python from First Principles by Seth Weidman published by O`Reilley