PROJECT DEVELOPMENT PHASE
SPRINT-II

Date	05 November 2022
TeamID	PNT2022TMID49483
Project Name	Digital Naturalist – AI Enabled Tool for Biodiversity Researchers
MaximumMarks	8Marks

Image Preprocessing

Click Here To view The Project (Hyperlink)

#Import The ImageDataGenerator Library:

```
import numpy as np
import tensorflow as tf
import keras
import keras.backend as K
from keras.optimizers import SGD, Adam, Adagrad, RMSprop
from keras.applications import *
from keras.preprocessing import *
from keras.preprocessing.image import ImageDataGenerator
from keras.callbacks import EarlyStopping, ModelCheckpoint
from keras.models import Sequential
from keras.layers import Dense, Conv2D, MaxPool2D, Flatten, Activation, BatchNormalization, Dropout
from keras.utils.np_utils import to_categorical
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt
import glob
from PIL import Image
import os
from os import listdir
```

#Make A List of Paths To All Folders Where You Have Data :

```
train_datagen = ImageDataGenerator(rescale=1./255,
                                   zoom_range=0.2,
                                   horizontal_flip=True)

test_datagen = ImageDataGenerator(rescale=1./255)
```

#Loading Images Into Machine Understandable Data :

Animal Dataset :

```
# Passing training data to train variable for animals
xtrain = train_datagen.flow_from_directory('/content/datasetbd/train/animals',
                                           target_size=(64,64),
                                           class_mode='categorical',
```

```
                                batch_size=100)

# Passing testing data to test variable for animals
xtest = test_datagen.flow_from_directory('/content/datasetbd/test/animals',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)
```

Birds Dataset:

```
# Passing training data to train variable for birds
xtrain1 = train_datagen.flow_from_directory('/content/datasetbd/train/birds',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)

xtest1 = test_datagen.flow_from_directory('/content/datasetbd/test/birds',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)
```

Flowers Dataset:

```
# Passing training data to train variable for flowers
xtrain2 = train_datagen.flow_from_directory('/content/datasetbd/train/flowers',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)

# Passing testing data to test variable for flowers
xtest 2= test_datagen.flow_from_directory('/content/datasetbd/test/flowers',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)
```

Marine Animal Dataset:

```
# Passing training data to train variable for marine animals
xtrain3 = train_datagen.flow_from_directory('/content/datasetbd/train/marine animals',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)


# Passing testing data to test variable for marine animals
xtest3= test_datagen.flow_from_directory('/content/datasetbd/test/marine animals',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)
```

Plants Dataset:

```
# Passing training data to train variable for plants
xtrain4 = train_datagen.flow_from_directory('/content/datasetbd/train/plants',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)

# Passing testing data to test variable for plants
xtest4 = test_datagen.flow_from_directory('/content/datasetbd/test/plants',
                                target_size=(64,64),
                                class_mode='categorical',
                                batch_size=100)
```