#### A close-up of a logo Description automatically generatedA logo of a company Description automatically generated

**Joint Tech Internship Community Program**

# **DEEP LEARNING ASSIGNMENT 1**

**Name** : Rashmika K R

**Degree**: B.Tech AIML

#### **Problem Description**

#### **Kaggle Link for the Dataset:** [**Bird Species Classification**](https://www.kaggle.com/datasets/gpiosenka/100-bird-species?select=EfficientNetB0-525-%28224+X+224%29-+98.97.h5)

#### **Content**: This dataset comprises 84,635 training images of 525 bird species, plus 2,625 images each for testing and validation. Its clean, high-resolution images make it suitable for developing precise bird identification models, which can be used in wildlife research, ecological studies, and conservation projects.

#### **Images**:

#### **Training Set**: 84,635 images

#### **Test Set**: 2,625 images (5 images per species)

#### **Validation Set**: 2,625 images (5 images per species)

#### **Image Specifications**:

#### **Size**: 224 x 224 pixels

#### **Channels**: 3 (RGB)

#### **Format**: JPEG

#### **Structure**:

#### Each dataset (train, test, validation) contains 525 subdirectories, one for each bird species.

#### **File Naming**: Training images are numbered sequentially with zero padding (e.g., 001.jpg, 002. jpg). Test and validation images are named 1.jpg to 5.jpg per species.

### **Basic Terminologies**

1. **Neural Network**: A neural network is a machine learning program, or model, that makes decisions like the human brain, by using processes that mimic how biological neurons work together to identify phenomena, weigh options and arrive at conclusions. For this dataset, a neural network will be used to classify images of birds based on their features.
2. **Neuron**: A neuron is a unit in a neural network that processes input and outputs a signal. An artificial neuron (also referred to as a perceptron) is a mathematical function. It takes one or more inputs that are multiplied by values called “weights” and added together. This value is then passed to a non-linear function, known as an ***activation function***, to become the neuron’s output. In the context of this dataset, neurons will help identify and classify features in bird images.
3. **Layer**: A layer is a collection of neurons in a neural network. For the bird dataset, multiple layers will be used to process the image data, including convolutional layers for feature extraction and dense layers for classification.
4. **Input Layer**: The input layer is where the image data enters the neural network. This layer will accept the data and pass it to the rest of the network. For this dataset, each image is a 224 x 224 x 3 array (RGB), which will be fed into the input layer.
5. **Hidden Layer**: Hidden layers are either one or more in number for a neural network.Hidden layers process input from previous layers to learn complex patterns. They perform multiple functions at the same time such as data transformation, automatic feature creation etc. In this dataset, hidden layers will help in learning features specific to different bird species.
6. **Output Layer**: The output layer produces the final classification of each image. For this dataset, it will have 525 neurons, each representing one of the bird species.
7. **Convolutional Layer**: A convolutional layer applies filters to the image to extract features such as edges and textures. This is crucial for processing the bird images to recognize various features.
8. **Convolutional Neural Network (CNN)**: A Convolutional Neural Network (CNN) is a type of deep learning algorithm that is particularly well-suited for image recognition and processing tasks. It is made up of multiple layers, including convolutional layers, pooling layers, and fully connected layers. Visual processing inspires the architecture of CNNs in the human brain, and they are well-suited for capturing hierarchical patterns and spatial dependencies within images.CNNs are designed for image processing and classification tasks. They are ideal for this dataset as they can automatically learn and extract features from bird images to classify them into 525 species.
9. **Recurrent Neural Network (RNN)**: Recurrent Neural Network works better than a simple neural network when data is sequential like Time-Series data and text data. RNNs are typically used for sequential data, not image data. They are less relevant for this dataset, which is focused on static images.
10. **Activation Function**: An activation function in the context of neural networks is a mathematical function applied to the output of a neuron. The purpose of an activation function is to introduce non-linearity into the model, allowing the network to learn and represent complex patterns in the data. Without non-linearity, a neural network would essentially behave like a linear regression model, regardless of the number of layers it has. An activation function determines the output of a neuron. In this dataset, common activation functions include:
    * **ReLU (Rectified Linear Unit)**: ReLU is less computationally expensive than tanh and sigmoid because it involves simpler mathematical operations. At a time only a few neurons are activated making the network sparse and making it efficient and easy for computation. Used in convolutional layers to introduce non-linearity and handle the features from the bird images.
    * **Sigmoid**: Usually used in the output layer of binary classification, where the result is either 0 or 1, as the value for the sigmoid function lies between 0 and 1 only so, the result can be predicted easily to be ***1*** if the value is greater than **0.5** and ***0*** otherwise. It is less commonly used for multi-class image classification.
    * **Tanh**: The activation that works almost always better than the sigmoid function is the Tanh function also known as the **Tangent Hyperbolic function**. It’s a mathematically shifted version of the sigmoid function. Provides a range of -1 to 1, less common in CNNs.
    * **Softmax**: If the output is for multi-class classification then, Softmax is very useful to predict the probabilities of each class.Used in the output layer to convert the final layer’s outputs into probabilities for each bird species.
11. **Forward Propagation**: Passing the input data through the network to obtain predictions. For this dataset, forward propagation will classify bird images into one of 525 species.
12. **Backpropagation**: A method for updating the model weights based on the error between predictions and actual labels. It helps improve the model’s accuracy on the bird images.
13. **Loss Function**: Measures the difference between predicted and actual labels. For this dataset, categorical cross-entropy is commonly used for multi-class classification.
14. **Cost Function**: Also known as the loss function, it quantifies the overall error of the model. It helps guide the optimization process to improve model performance.
15. **Gradient Descent**: An optimization algorithm used to minimize the loss function by adjusting model parameters. It helps in fine-tuning the model to achieve better classification accuracy.
16. **Learning Rate**: Determines how quickly the model learns during training. A well-chosen learning rate can accelerate convergence and improve the model’s performance on the bird dataset.
17. **Batch Size**: The number of images processed in one iteration of model training i.e The batch size is a hyperparameter that defines the number of samples to work through before updating the internal model parameters. For this dataset, the batch size affects training efficiency and memory usage.
18. **Epoch**: One complete pass through the entire training set. The number of epochs is a hyperparameter that defines the number of times that the learning algorithm will work through the entire training dataset. For this dataset, multiple epochs are necessary to ensure the model learns effectively from the bird images.
19. **Overfitting**: When the model performs well on training data but poorly on test or validation data. This can occur if the model learns specific details from the training images that do not generalize well.
20. **Underfitting**: When the model is too simple to capture the patterns in the data, leading to poor performance on both training and test data. This can be mitigated by increasing model complexity or training time.
21. **Training Set**: The subset of the dataset used to train the model. For this dataset, it consists of 84,635 images of bird species.
22. **Validation Set**: Used to tune and evaluate the model during training. This dataset includes 2,625 images and helps in assessing model performance.
23. **Test Set**: Used to evaluate the final performance of the model. This dataset also consists of 2,625 images, providing an unbiased measure of model accuracy.
24. **Cross-Validation**: Technique for evaluating model performance by dividing the data into folds. It ensures that the model generalizes well and avoids overfitting.
25. **Hyperparameters**: Parameters set before training that influence model performance (e.g., learning rate, batch size). Tuning these hyperparameters can optimize the model for the bird dataset.
26. **Model Parameters**: The weights and biases learned during training. These parameters are adjusted to minimize the loss function and improve classification accuracy.
27. **Regularization**: Techniques to prevent overfitting by adding constraints or penalties. Methods like dropout or L2 regularization can be applied to improve model generalization.
28. **Dropout**: A regularization method where randomly selected neurons are ignored during training to prevent overfitting. It helps the model generalize better on unseen images.
29. **Weight Initialization**: Setting the initial values of model weights before training. Proper initialization helps in achieving faster convergence and better model performance.
30. **Normalization**: Scaling input data to a standard range (e.g., 0 to 1). For the bird dataset, normalizing image pixel values can improve model training and performance.
31. **Standardization**: Scaling data to have a mean of 0 and a standard deviation of 1. This can help with model stability and performance, especially if the input data has varying scales.

#### 

#### 

#### 

#### 