TASK 3.1

As a part of the wildlife and environmental monitoring project, you need to identify the species of animals and birds. Consider yourself part of a wildlife sanctuary/refuge team with access to video devices. There are designated camera spots within the territory. How would you employ Machine Learning and its aiding technology for this purpose? Explain how you can preprocess the sources, and classify the animals or birds in the video clips. You have to mention the algorithm or method you will be using and the reasoning. Mathematical explanations would be beneficial.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Step 1: The first step would collect the data from the photos clicked by the cameras and preprocess it. In preprocessing, we have to identify the object and outline it so as to remove its background and reduce impurity in data (noise reduction). We divide the picture frame by frame and go through feature extraction.

Step 2: We have to train the model to identify the features on its own and decide, for we have to create a CNN (Convolutional Neural Networks) model. The model will learn from the image frames to identify the features. We can use the **Decision tree method** to train the CNN model, for simplicity, we can use binary values such as ‘Yes’ and ‘No’. The attributes can be wings, patterns, limbs, claws, etc. with root attributes such as ‘is animal?’. The positive value in the root attribute will separate the species as an animal and the negative as a bird.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Frame | Wings | Patterns | Limbs | Claws | Is animal? |
| Frame1 | Yes | Yes | No | No | No |
| Frame2 | No | Yes | Yes | No | Yes |
| Frame3 | No | No | Yes | Yes | No |
| Frame4 | No | Yes | Yes | No | Yes |
| Frame5 | Yes | Yes | No | Yes | No |
| Frame6 | No | No | Yes | No | Yes |
| Frame7 | No | No | Yes | No | Yes |

Step 3: We need to implement the mathematical logic behind the training model, we can use the Information gain method.

Information Gain = (Entropy before split) – (weighted entropy after split)

=

Entropy = - p(A).log2(p(A)) – p(B).log2(p(B))

Here, p(A) is the positive outcome ratio, and p(B) is the negative outcome ratio.

The higher the information gain, the higher the certainty of true value.

Step 4: Just like step 2, we should add more branches to the decision tree to separate the data into species-specific classes. This is just the simplest of the possible models. The complexity of the model rises with movement recognition, light adjustment, environment noise, etc. Also, we have taken binary values as Yes and No but the complex model will recognize the values differently and cluster them into groups, i.e. using clustering and decision making model.