Title: Project Presentation – Wildlife Vocalization Classification for Conservation Monitoring – Project based on Audio Analysis and Birds detection in audio clips.

Subtitle: Purnima Rangavajjula - Roll Number: AM.EN.U4CSE21046.

**About the Dat﻿aset:** The dataset consists of 10 second audio clips.

A portion of the audio clips consists of bird sounds, and the and the rest consists of other miscellanious sounds.

Features are extracted from the Audio clips and processed to make it fit for the models presented

The features consist of:

* Spectral\_Centroid
* Spectral\_Bandwidth
* Spectral\_Contrast
* Spectral\_Rolloff
* Chroma
* Spectral\_Flatness
* Zero\_Crossing\_Rate
* HNR

The dataset has 4730 values and 11 columns

Logistic Regression:

Logistic regression is a type of algorithm used for such binary classification problems.

features are your input variables. you assign a weight that represents its importance in predicting. You multiply each feature by its weight and add them up. This gives you a weighted sum. ou pass the weighted sum through a special function called the logistic function (or sigmoid function). This function squashes the weighted sum into a value between 0 and 1. The output from the logistic function can be interpreted as the probability. You then choose a threshold (like 0.5) – if the probability is above this threshold, you predict according to that. This threshold is your decision boundary. During training, the algorithm adjusts the weights based on the difference between its prediction and the actual outcome (whether the email is really spam or not). It aims to minimize this difference, improving its ability to make accurate predictions.

Logistic regression takes in input features, assigns weights to them, combines them, applies a logistic function, and then uses the result to make a binary prediction.

Naïve Bayes:   
Naive Bayes is a probabilistic machine learning algorithm that is based on Bayes' theorem. It is particularly popular for classification tasks, such as spam detection, sentiment analysis.

The "naive" in Naive Bayes comes from the assumption that features used to describe an observation are conditionally independent, given the class label.

Bayes' theorem describes the probability of an event based on prior knowledge of conditions that might be related to the event.

The formula is as follows:

P(A∣B)=P(B)P(B∣A)×P(A)​

In the context of machine learning:

P(A∣B) is the probability of class A given the observed features B.

P(B∣A) is the probability of observing the features B given the class A.

P(A) is the prior probability of class A.

P(B) is the probability of observing the features B

Types of Naive Bayes:

Gaussian Naive Bayes: Assumes that the features follow a Gaussian distribution.

raining and Classification:

During training, the algorithm calculates the probabilities needed for Bayes' theorem using a labeled dataset.

During classification, the algorithm predicts the class with the highest posterior probability based on the observed features.

Advantages and Limitations:

Advantages: Simple, fast, and performs well in many real-world situations, especially when the naive assumption holds.

Limitations: The naive assumption may not always reflect the true relationships between features. It might not perform well on highly complex tasks.