# Chirp: Birdsong Recognition

Google Developer Student Clubs

Indian Institute of Technology Kanpur

## Overview of the Project:

Birdsong recognition plays a crucial role in various fields, from ecological research to wildlife monitoring. The 'Chirp' project aims to develop an efficient and accurate bird call recognition system using machine learning techniques. The primary goal is to enable the identification of bird species based on their unique vocalizations.

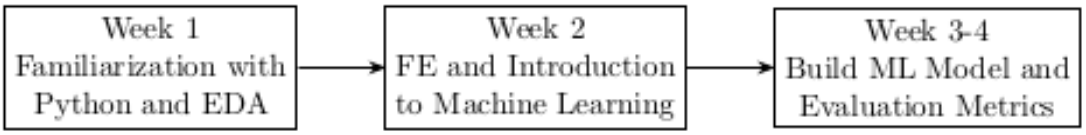
## Tech Stacks:

- **Programming Languages:** Python
- **Libraries and Frameworks:** NumPy, Pandas, Matplotlib, Seaborn, Librosa, Sklearn, Py-Torch
- **Tools:** Jupyter Notebook, Google Colab

## Project Timeline:

- **Week 1:**
  - Day 1: Introduction to NumPy, the scientific computing library of Python.
  - Day 2: Introduction to Pandas, the data analysis library of Python.
  - Day 3: Reading on various different types of possible plots and introduction to Matlotlib and Seaborn, visualization libraries of Python.
  - Days 4 and 5: Cleaning, preprocessing and EDA of a practice CSV file.
  - Days 6 and 7: Addressing doubts encountered so far + brush-up of concepts.
- **Week 2:**
  - Days 1 and 2: Learning about audio and its various features.
  - Days 3 and 4: Extracting learned features from a practice dataset.
  - Day 5: Basic introduction to Machine Learning.
  - Day 6: Application of ML models to audio data.
  - Day 7: Addressing doubts encountered so far + brush-up of concepts.
- **Week 3:**
  - Days 1 and 2: Getting familiar with the dataset and EDA + FE.
  - Days 3 and 4: Learning about Neural Networks, Pytorch and ResNet.
  - Days 5 and 6: Understanding ML models that can be used for our dataset.
  - Day 7: Addressing doubts encountered so far + brush-up of concepts.

- **Week 4:**
    - Day 1 and 2: Building our models from scratch.
    - Day 3 and 4: Fine-tuning and evaluation of performance of the models.
    - Day 5 and 6: Improvement of the models.
    - Day 7: Addressing doubts encountered so far + brush-up of concepts.




## All Resources to be Used and Delivered:

1. **Learning Materials:** Curated tutorials, articles, and documentation.
2. **Hands-On Exercises:** Practical coding assignments to reinforce learning.
3. **Projects:** Real-world projects to apply acquired skills.
4. **Interactive Sessions:** Live sessions for doubt resolution and discussions.
5. **Feedback Mechanism:** Regular feedback and assessment to track progress.
6. **Final Project:** A capstone project to demonstrate proficiency in machine learning concepts.

**Exploratory Data Analysis (EDA)**

Initiate a thorough EDA of CSV files for data cleaning, preprocessing, and relationship exploration. Identify and address outliers to ensure data integrity.

**Feature Extraction (FE) from Audio Files**

Conduct extensive Feature Extraction from audio files, capturing beats per minute, spectrograms, zero-crossing rates, band energy ratios and other important features with the aim to help distinguish unique acoustic signatures for different bird species.

**Machine Learning Model Development**

Utilize various machine learning models for bird species classification based on extracted audio features. Evaluate model performance metrics to determine the most effective algorithm.

**Performance Evaluation**

Assess model performance across different algorithms, emphasizing accuracy, precision, recall, and F1 score. Identify optimal models and analyze their robustness in handling diverse birdcall patterns.

## Prerequisite:

- Basic understanding of Python programming.
- Familiarity with machine learning concepts is beneficial and highly recommended.

## Expected Number of Intakes: 35

## Evaluation Metrics:

- **Attendance:** Minimum **80%** attendance in sessions is mandatory. 12 sessions expected.
- **Week 1 and 2: Practical assignments** related to EDA and FE are to be completed.
- **Week 3: Performance evaluation** of the practice model.
- **Week 4: Performance evaluation** of the final fine-tuned model.

## Project Mentors:

Bhavna Jayswal (211213)
Github

## YouTube Playlist:

https://www.youtube.com/playlist?list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0

## Attendance Sheet:

https://docs.google.com/spreadsheets/d/1jyp1enzChoPlkczVaJGKm_kBhjseelNeEf5yUCx-LDU/edit#gid=0

# Session Details:

## Session 1: 13-12-23

https://www.youtube.com/watch?v=gn0rk8q414I&list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0&index=1&t=501s

**Agenda:**

1. Detailed explanation of the problem statement and week-wise work distribution.
2. Introduction to NumPy and Pandas

**Resources Provided:**

1. **NumPy:** (Updated 13-12-23)
   NumPy Theory
   NumPy Documentation (You can easily access and read documentations for any library.)
   Practice for NumPy

2. **Pandas:** (Updated 13-12-23)
   Pandas Course

Pandas Exercises
Theory + Questions
Pandas Aggregation
Pandas Sorting

**Practice Assignment (Optional but highly recommended):**
Based on your learning from the above resources, try to answer the questions given in the following Colab Notebook (The last question is optional since we have not studied matplotlib yet):
https://colab.research.google.com/drive/1MGkWmXrKp2wXCl_q1SXlL7-pT10Hx3ds

3. **Matplotlib:** (Updated 14-12-23)
   Types of Plots
   Matplotlib Tutorial
   3D Plots
   Matplotlib Exercises

4. **Seaborn:** (Updated 15-12-23)
   Seaborn Tutorial (Codes)
   Seaborn Tutorial (YT)

**Quiz 1** (Taken on 15-12-23): Quiz 1: Pandas & NumPy
**Make-up Quiz 1** (Taken on 16-12-23): Make-up Quiz 1: Pandas & NumPy

## Session 2: 16-12-23

https://www.youtube.com/watch?v=VTjQWxOb-8k&list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0&index=2

**Agenda:** Demonstration of exploratory data analysis (EDA) of a mock dataset using NumPy, Pandas, Matplotlib, Seaborn and Sklearn

**Resources Provided:**

1. **EDA:** (Updated 16-12-23)
   EDA Guide
   EDA Basic Tutorial (Codes)
   EDA Tutorial (YT)
   EDA (Kaggle)

2. **Sklearn:** (Updated 16-12-23)
   Sklearn Tutorial (YT)

3. **Audio Features:** (Updated 18-12-23)

Features of Audio
More Features of Audio
MFCC

**The Mock EDA Notebook used in the session:**
https://colab.research.google.com/drive/156aGz9QytU-hTB4mNpIXumvaBUMA8kd0?usp=sharing

**The dataset used in the session:**
https://drive.google.com/file/d/1dhieKI016737rCuNeWEvRDtj-eDF8XAj/view?usp=sharing

**Assignment 1:**
Choose any one of the following unclean datasets (or a different one based on your own preference) and perform EDA on them. Make your data as clean as possible, and make necessary inferences from the graphs you plot and/or the data you see in various columns. Write all your inferences in the notebook as you go.

**Datasets to be used for the assignment:**
https://drive.google.com/file/d/1rJZP6eRBR7ypnvCG1d1mr_9d0BjcZoo7/view?usp=sharing
https://drive.google.com/file/d/1dhieKI016737rCuNeWEvRDtj-eDF8XAj/view?usp=sharing
https://drive.google.com/file/d/1P3XAwPP01Y4-wdFkLJafC5o1y1u6B_W-/view?usp=sharing
https://drive.google.com/file/d/1bl9bRz_h5hhu1biq-M4_tS5cKyeGvvdQ/view?usp=sharing

**Submission deadline:** 21-12-23 11:59 PM

**Submission Link:** https://forms.gle/TiWG46fqDUHCf16e9

**Quiz 2** (Taken on 19-12-23): https://forms.gle/NRe5jk3iGQCFRjkM6

## Session 2.5: 20-12-23

https://www.youtube.com/watch?v=q1C-6oJj2g8&list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0&index=3

**Agenda:** Addressing the dos and don'ts of EDA.

## Session 3: 24-12-23

https://www.youtube.com/watch?v=bQSnS4M4hQs&list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0&index=4

**Agenda:** Demonstration of extraction of all the features of an audio using the Librosa library.

**Resources Provided:**

1. **Feature Engineering:** (Updated 20-12-23)

ML for Signal Processing (Videos 1, 2, 3 & 12 for now)
FE Kaggle Course (VERY important)
Audio Signal Processing (YT Playlist) (Important)

**The notebook used in the session:**
https://colab.research.google.com/drive/1aT4bNme5MLxYotnCAg0sbIG64WjPP-yC?usp=sharing
**The audios used in the session:**
https://drive.google.com/file/d/11VOovo7YJ-F8IsrAAT10_G7FMk4bDjJH/view?usp=sharing
https://drive.google.com/file/d/1UzwKp1k6RT8gIHdRKOmELcxxeLW--XyJ/view?usp=sharing

**Assignment 2:**
As demonstrated in the session, you need to find the similarities in the features of various audios of the same number, and differences in the audios of different numbers. Use the values you find/plots you create to discover which of the features can best distinguish the audios and be used to recognize similar ones. Write your analysis in a structured way within your notebook with a conclusion summarizing everything you infer from the extracted features.

**Audios to be used for the assignment:**
https://drive.google.com/drive/folders/1M9K__JCpAn7ancHOG_Xp7bcCgNSM-wtN?usp=sharing
**Audios taken from:**
https://www.kaggle.com/datasets/alanchn31/free-spoken-digits

**Submission Deadline:** 26.12.23 11:59 PM

**Submission Link:** https://forms.gle/q7wq8XmfGFt2NXGj6

## Session 4: 26-12-23

https://www.youtube.com/watch?v=3wo_TJL1U0k&list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0&index=5

**Agenda:**

1. Explanation of differences between traditional machine learning models and neural networks.
2. Demonstration of working and building of traditional machine learning methods using the best features found in the assignment that can be used to differentiate between various audios in order to classify future audios.

**Resources Provided:**

1. **Neural Networks** (Updated 26-12-23)
   What are Neural Networks (YT)
   Working of a Neural Network (YT Videos 1 to 7)
   Activation Functions (YT)

Activation Functions (Blog)
**Optional:** Machine Learning Algorithms (YT playlist for anyone who is interested in Traditional ML Algorithms – ONLY for additional knowledge; does not have a relation with our project.)

2. **PyTorch** (Updated 30-12-23)
   Introduction to PyTorch (YT Playlist)

3. **Building a Neural Network** (Updated 30-12-23)
   Sample Model 1 (Blog)
   Sample Model 2 (YT Playlist – 4 videos – Preferred)

**Assignment 3:**
The two sample models given to you are functional. For your next assignment, what you need to do is create a similar model to classify the provided audio data for 10 different categories. You can use any number of audios from every category. You can also decide how many categories you want to fit in your model. It is not necessary to use all of them, but it is necessary to use a minimum of 4 categories with at least 25 audios for each category. Building the actual model here should not take up much of your time, so what I need you to do is try to understand each and every line of code there is.
**Note:** The provided models are built using Tensorflow, and not PyTorch. (Both of them have similar functions so if you know one library, the other should not be as difficult to understand.)

**Dataset to be used for the assignment:**
https://www.kaggle.com/datasets/chrisfilo/urbansound8k?select=UrbanSound8K.csv

**Submission Deadline:** 03-01-24 11:59 PM

**Submission Link:** https://forms.gle/M1uSBEp6fbHnJx8b6

## Session 5: 03-01-24

https://www.youtube.com/watch?v=yg5w4cCe5Fs&list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0&index=6

**Agenda:** Line-by-line explanation of a sample neural network model built using Tensorflow.

**Quiz 3** (Taken on 04-01-23): https://forms.gle/b7W7q27qwN6bW5Qc6
**Re-Quiz 3** (Taken on 09-01-24): https://forms.gle/5b6D6oedbgJJKBEv6

## Session 6: 05-01-24

https://www.youtube.com/watch?v=9IvuN4aw0Pk&list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0&index=7

**Agenda**: Step-by-step demonstration of EDA of the final 'Chirp' dataset to be used.

**Resources Provided:**

1. **Batch Normalisation** (Updated 08-01-24)
   Batch Normalisation (YT)

2. **ResNet** (Updated 08-01-24)
   ResNet Model from Scratch (YT)

3. **Evaluation Metrics** (Updated 09-01-24)
   Defining Different Metrics (Blog)
   Implementation of the Metrics (YT)

**The Final Dataset to be used:**
https://www.kaggle.com/competitions/birdsong-recognition/data (*train.csv* is the file to be used for EDA. All our work from this point onwards will be done on this dataset.)

**Final Project Assignment I:**
As demonstrated in the session, do the EDA on your own for the *train.csv* file. In addition to that, pick the birdsong audios of any 5 birds, two audios *each*. So in total you'll have 10 birdsongs. Similar to what you have done in your FE assignment, find the features that show similarities between two audios and differences between varied audios. In this case, try not to get only numerical values for features. Plot graphs instead to make your analysis more readable and explanatory at a glance to other users.

Since the data is very large, go to the 'Code' tab on the Kaggle page and create a New Notebook. All the files you need will already be there. If you are working on Colab with the one CSV file, it's fine for EDA, but from the FE part to the final model, make a Notebook on Kaggle and work there.

## Session 7: 10-01-24

https://www.youtube.com/watch?v=g7IhpOZiZBQ&list=PL_fkUfeFmd2vYQTvMOTYHXHeVFuahkQq0&index=8

**Agenda:** Demonstration and explanation of each line of code used in building the final model for birdsong recognition

**Final Project Assignment II:**
As shown in the session, follow along and create your own neural network model using the ResNet50 architecture. The ‘submission.csv’ file you create would have 75 entries for only 15 audios since we are predicting a label for every 5 seconds, so at the end of your submission, you will need to find the mode among the predictions for every audio and submit the csv file with only 15 unique audio IDs and their predicted label.

**Submission Deadline:** 11-01-24 11:59 PM

## Final Submission Link: https://forms.gle/mNVBaopRNsir62e4A

**Quiz 4** (Taken on 11-1-24): https://forms.gle/UxYWPQazvY3E5Rnf8
**Re-quiz 4** (Taken on 14-1-24): https://forms.gle/YqDo2z5bHrpaHiJ88

## Resources:

Link for Combined Resources