

CONTENT

* Problem Statement.
* Dataset.
* Proposed Methodology.
* Results.

Problem Statement:

Audio data is used by the vocal feeling recognition system. It takes a segment of speech as input and identifies the emotions the orator is expressing. You can recognize many emotions such as happiness, sadness, surprise, anger, and so on.

Dataset:

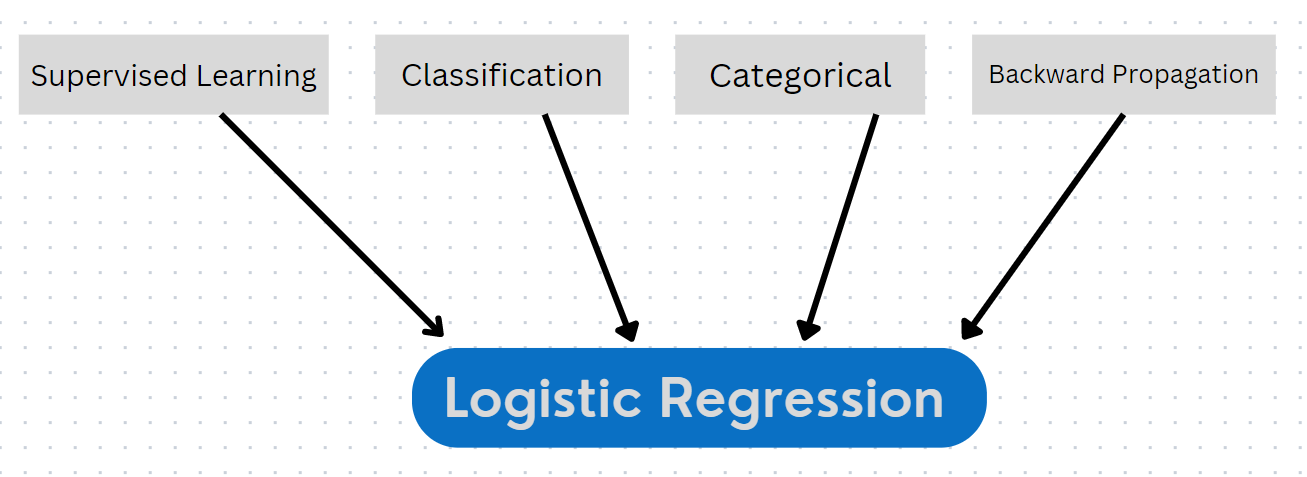
The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS).

[The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) | Zenodo](https://zenodo.org/record/1188976#.YFZuJ0j7SL8)

Proposed Methodology:

Logistic Regression:

Logistic regression is a supervised machine learning algorithm that accomplishes binary classification tasks by predicting the probability of an outcome, event, or observation. The model delivers a binary or dichotomous outcome limited to two possible outcomes: yes/no, 0/1, or true/false.



Classification:

The **Classification** algorithm is a **Supervised Learning** technique that is used to identify the **category** of new observations based on training data. In Classification, a program learns from the given dataset and then classifies new observation into several classes or groups. Audio Emotion Recognizer is a **Multi-class Classification problem as it results in** more than two outcomes (fear, happy, sad, angry, calm, neutral).

Backward Propagation:

* **Stochastic Gradient Descent-** **Stochastic Gradient Descent (SGD)** is a simple yet very efficient approach to fitting linear classifiers and regressors under convex loss functions such as (linear) support Vector Machine and logistic Regression.
* **Implementation:**

Let (x(i),y(i)) be the training example

Cost(θ, (x(i),y(i))) = (1/2) Σ( hθ(x(i)) - y(i))2

Jtrain(θ) = (1/m) Σ Cost(θ, (x(i),y(i)))

Repeat {

For i=1 to m{

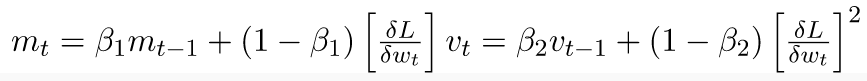
θj = θj – (learning rate) \* Σ( hθ(x(i)) - y(i))xj(i)

For every j =0 …n

}

}

* **Adam Solver -** Adaptive Moment Estimation is an algorithm for optimization technique for gradient descent. The method is efficient when working with large problem involving a lot of data or parameters. It requires less memory and is efficient. Intuitively, it is a combination of the ‘gradient descent with momentum’ algorithm and the ‘RMSP’ algorithm.
* **Mathematical Aspect of Adam Optimizer: -**



**Parameters Used:**

**1. ϵ =** a small +ve constant to avoid 'division by 0' error when (vt -> 0). (10-8)

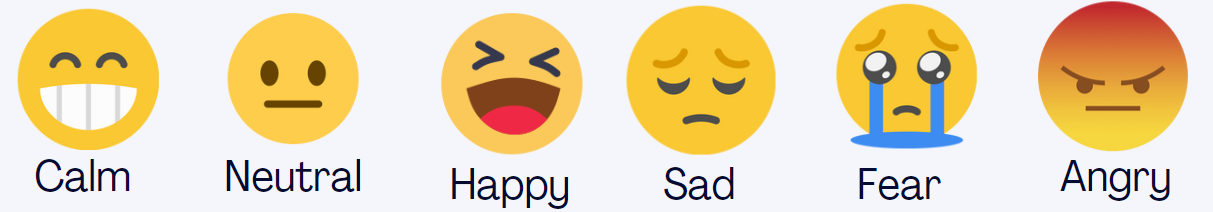
**2. β1 & β2 =** decay rates of average of gradients in the above two methods. (β1 = 0.9 & β2 = 0.999)

**3. α —** Step size parameter / learning rate (0.001)

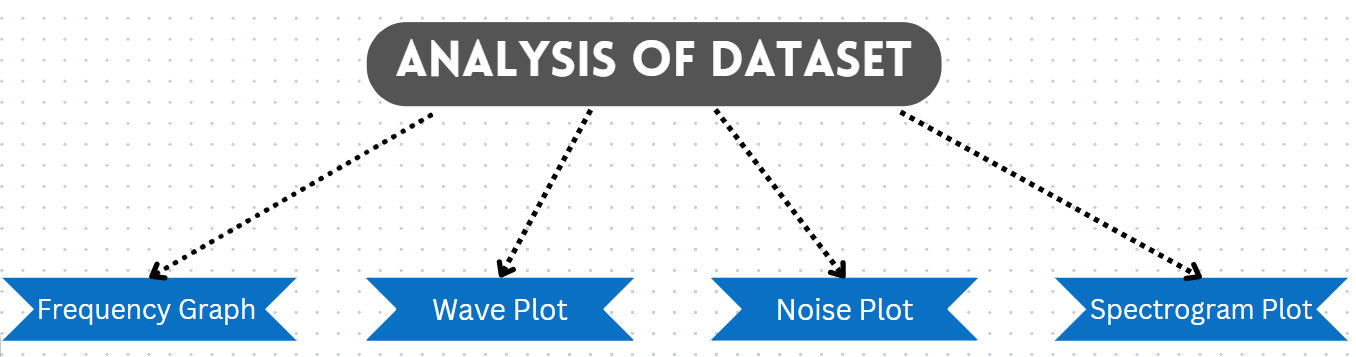
Code:

[GitHub - Anunaya07/Speech-Emotion-Recognizer](https://github.com/Anunaya07/Speech-Emotion-Recognizer)

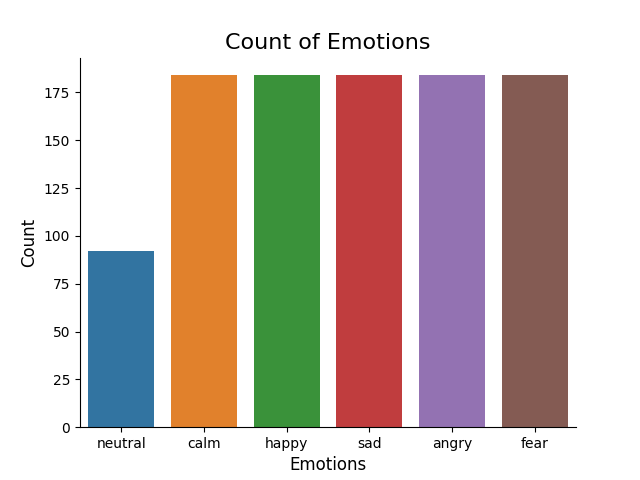
Emotions recognized:



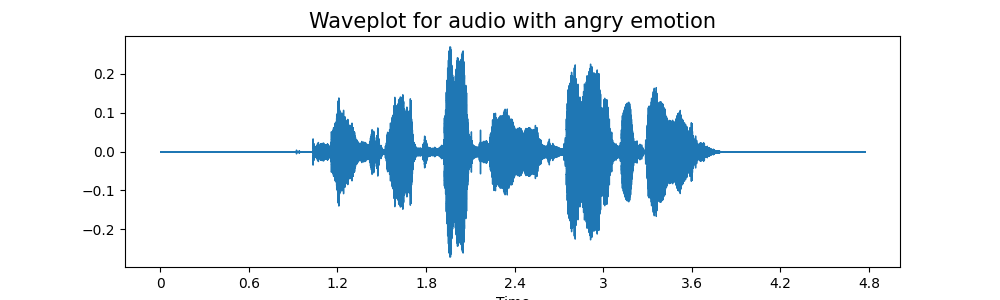
Analysis of Dataset:

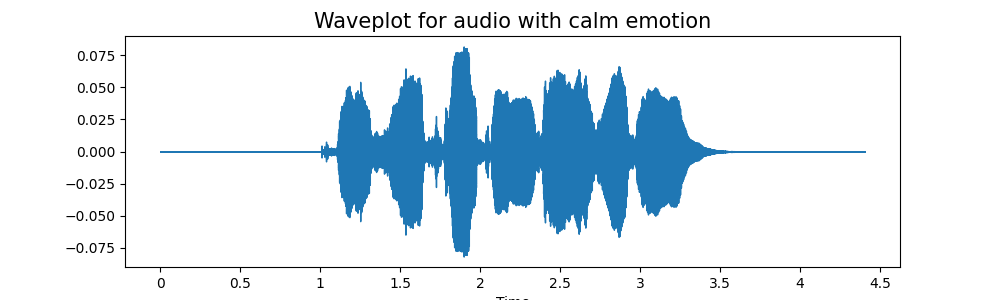
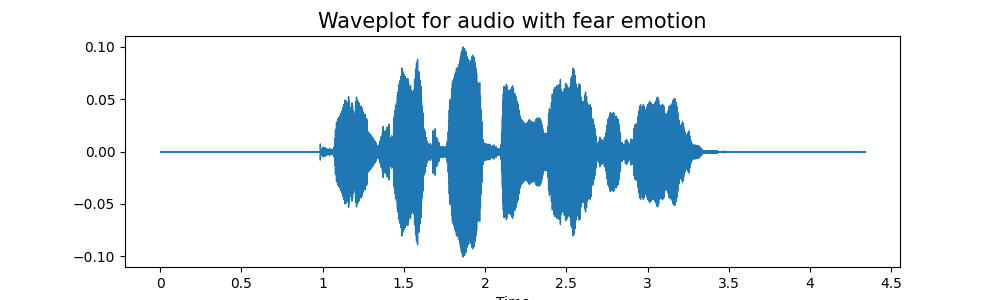
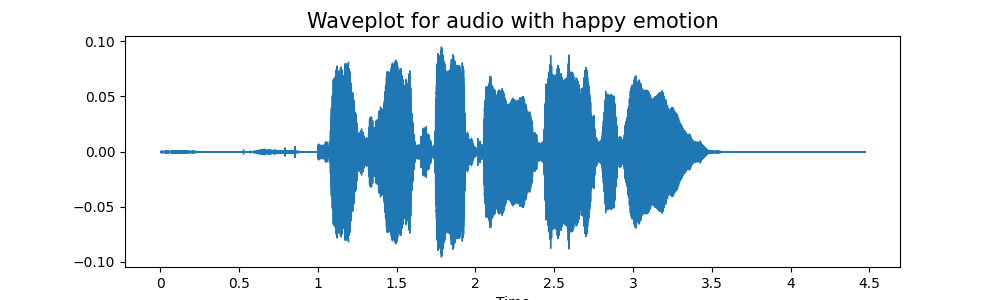
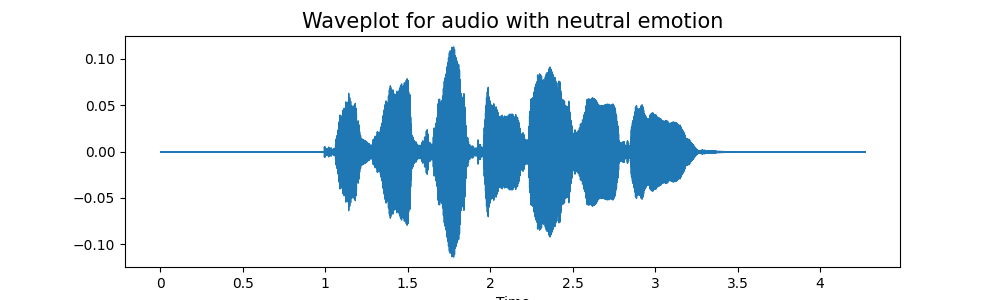


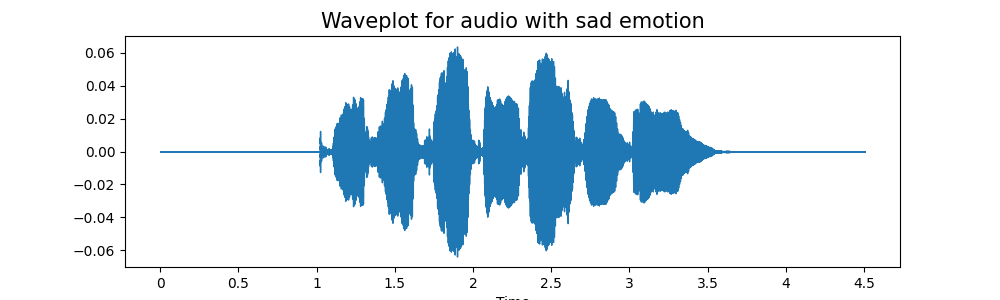
Frequency Graph:



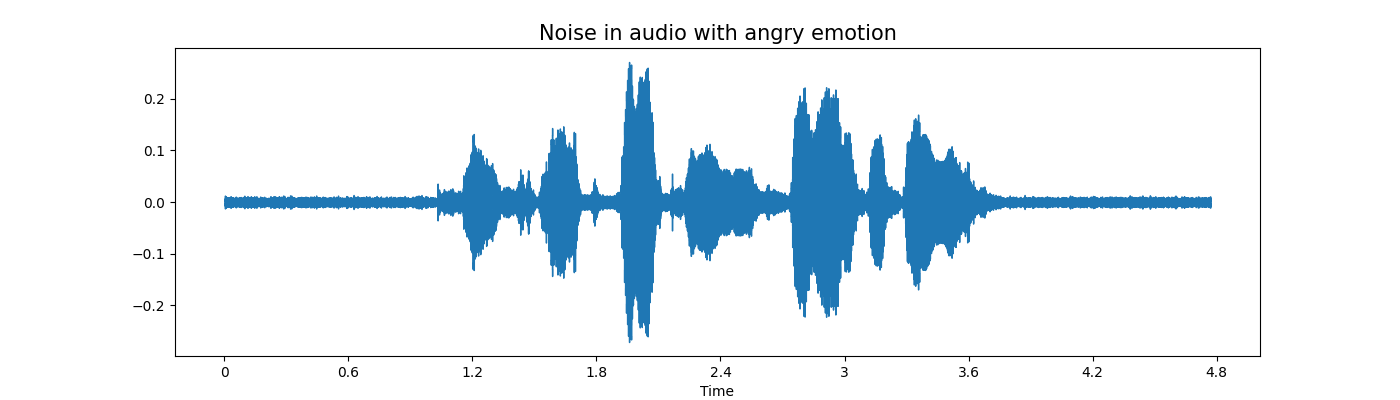
**Wave plot analysis:**

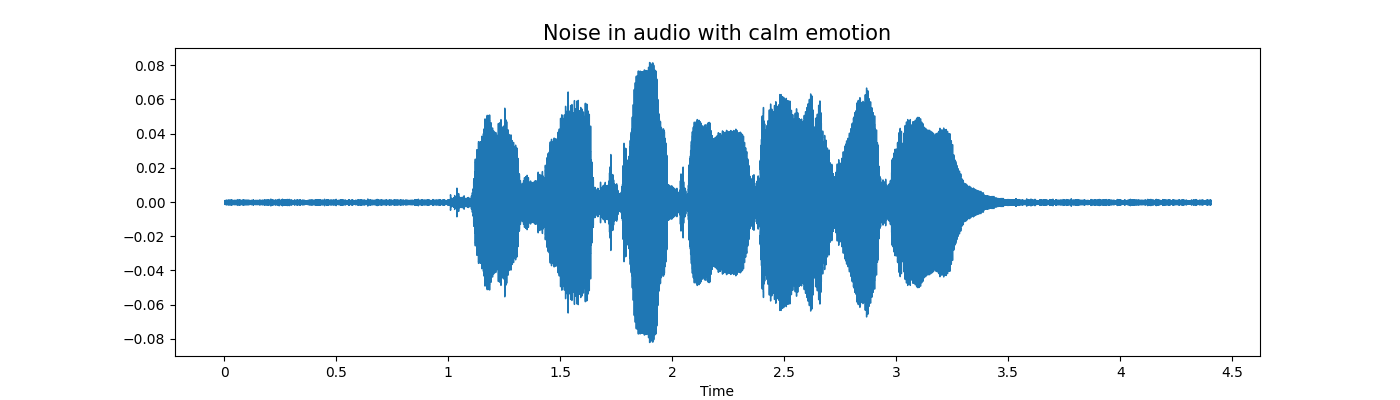


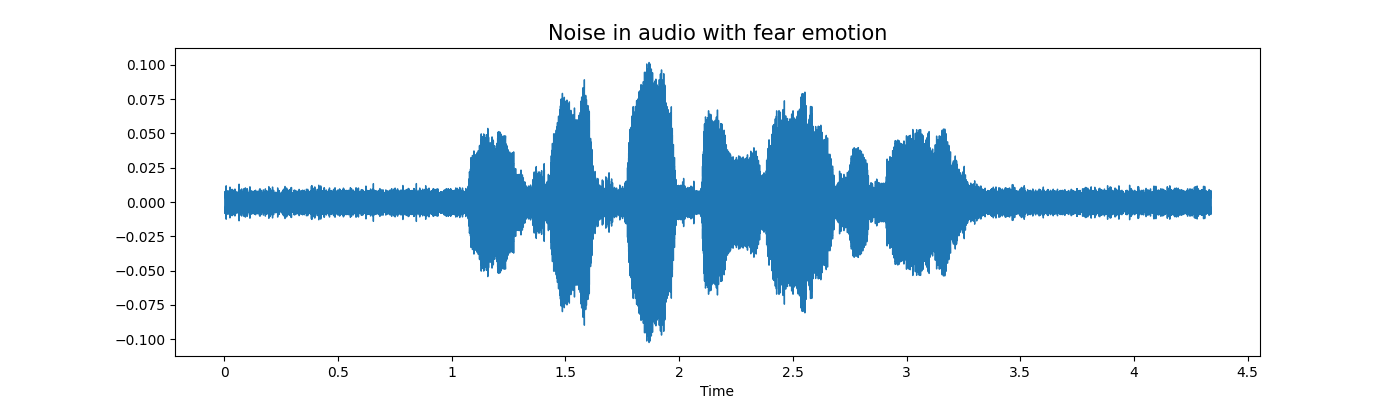
   

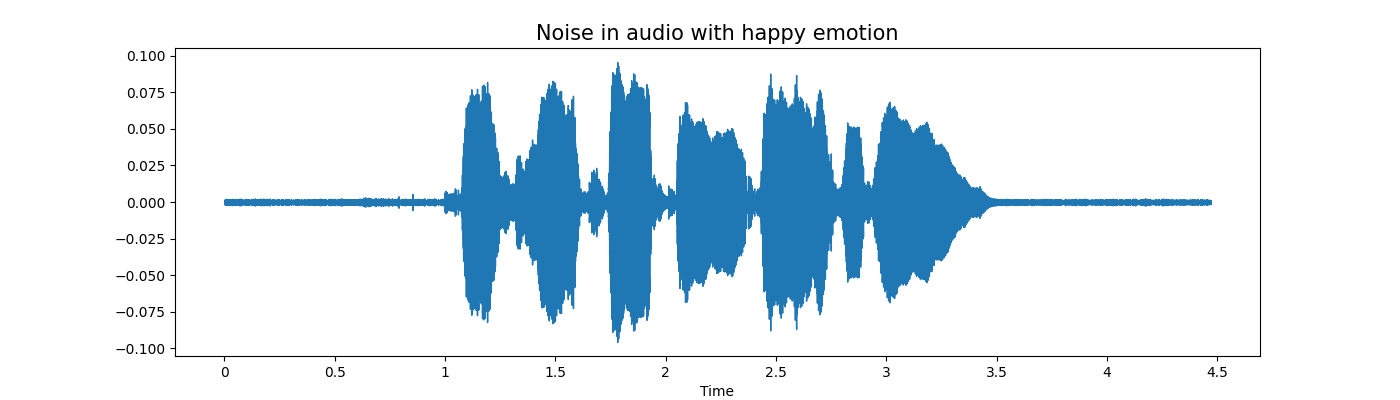


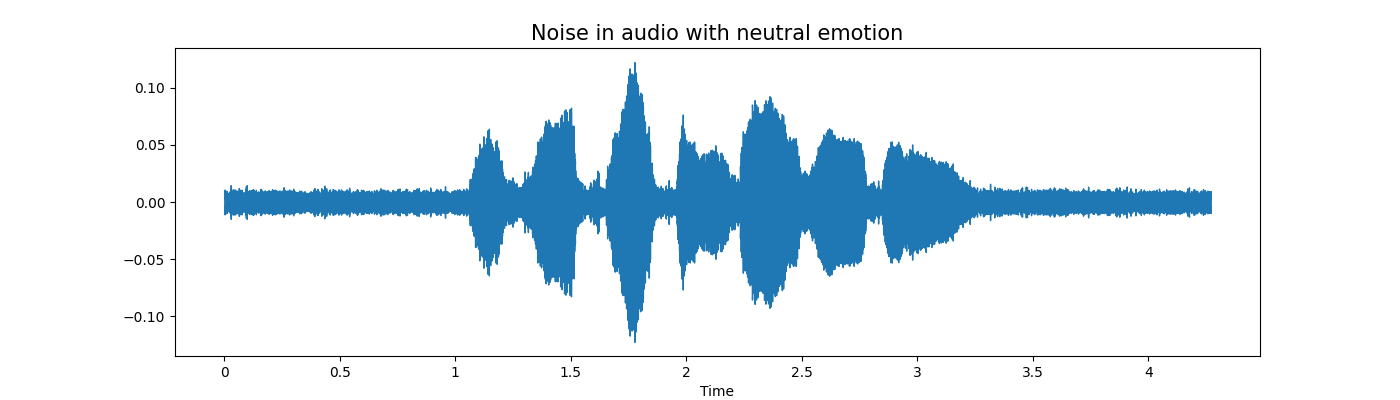
**Noise Analysis:**

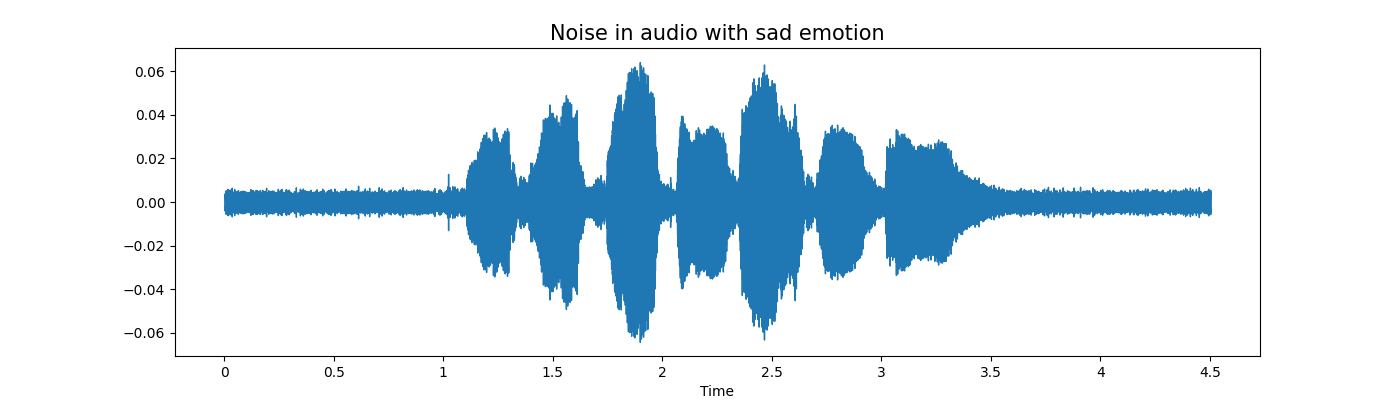




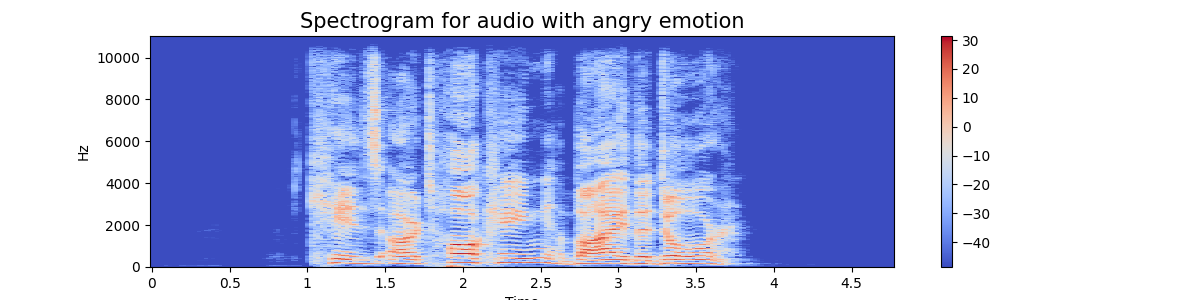


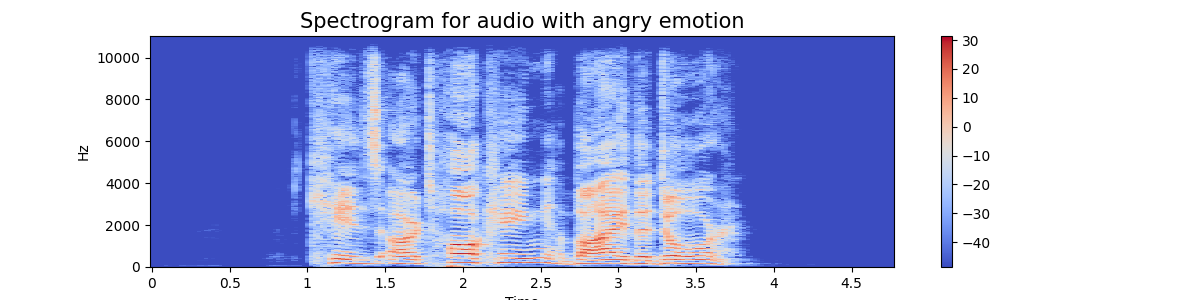
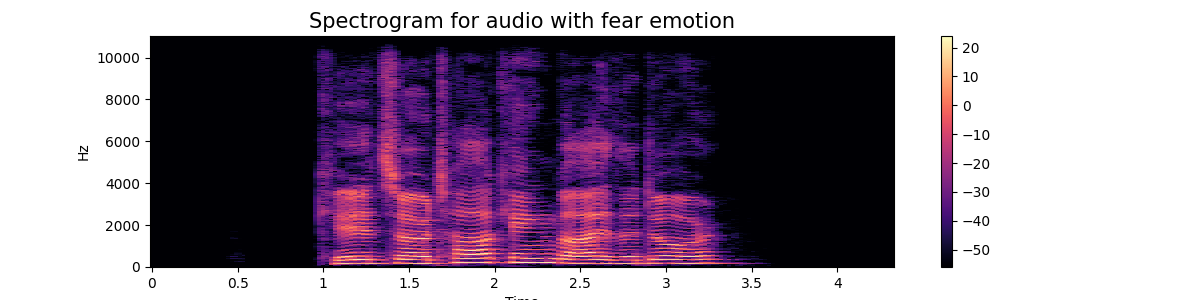
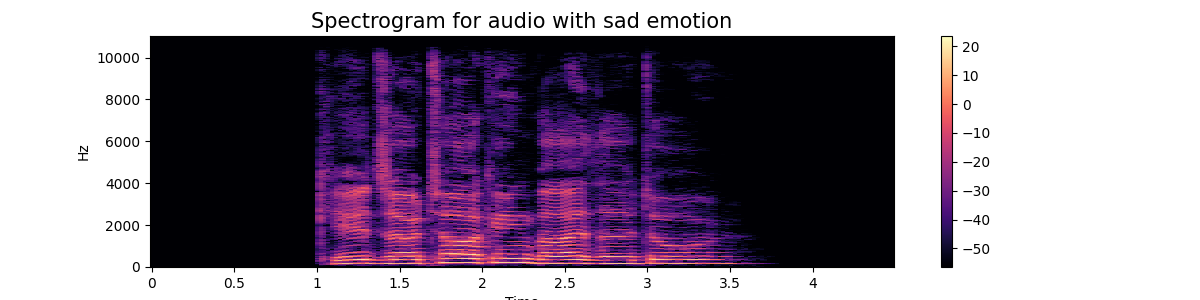






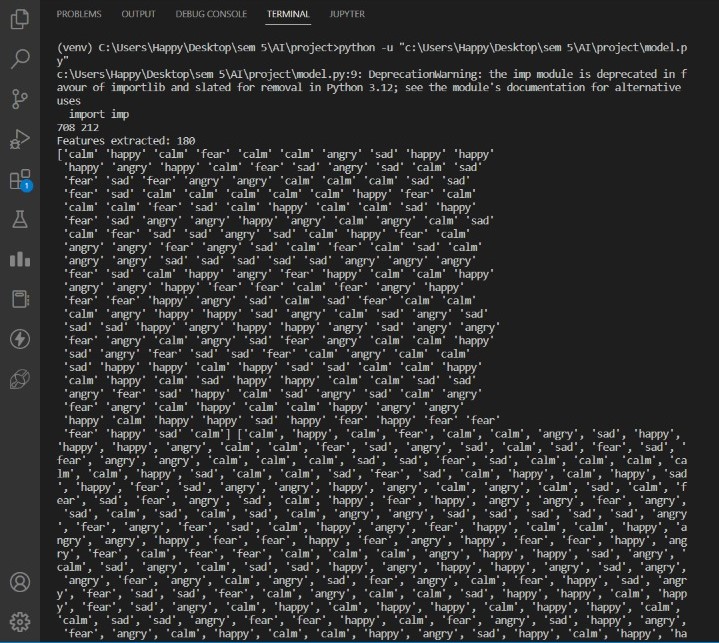
**Spectrogram analysis:**

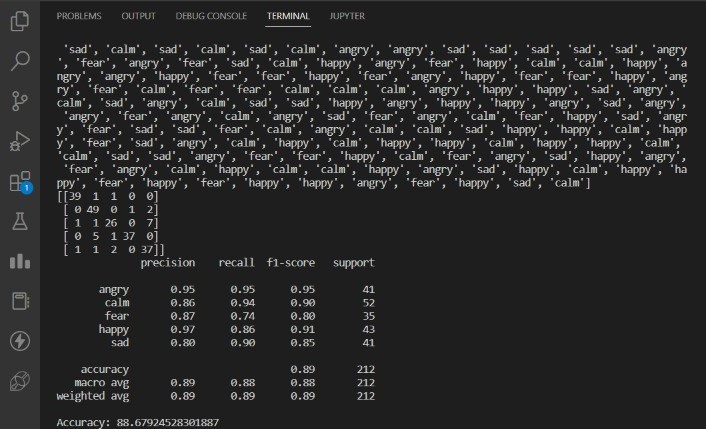


Results: -

Dataset –





Screenshot of Website UI:

