

CSE 363 / ECE 343: Machine Learning Project Proposal

Title: Gender and Emotion Classification by Voice

Anikait Agrawal
anikait22072@iiitd.ac.in

Anant Kaushal
anant22067@iiitd.ac.in

Ansh Varshney
ansh22083@iiitd.ac.in

Abdullah Shujat
abdullah22013@iiitd.ac.in

1. Motivation

Recognizing emotion and gender from speech have become critical areas of research as it improves human-computer interaction, privacy and security, and speaker identification. Emotion recognition makes our systems more intuitive and empathetic, enhancing user experiences in virtual assistants, telemedicine and customer service. Similarly, gender recognition from speech is vital for smart devices and surveillance systems.

2. Related Work

Various works that are still going on in the field of Gender and Emotion Recognition are as follows:

2.1. A New Gender Detection Algorithm Considering the Non-Stationarity of Speech Signal :

A Pitch based algorithm has been developed for gender detection using analysis of the non-stationary behavior of the voice signal. The Concept of peak detection is used to get the dominant pitch, the frequencies, and the fundamental frequency in the sound signal. [\[Link\]](#)

2.2. Speech Emotion Recognition using CNN and LSTM models:

This study explores the use of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) models which demonstrated significant improvements in accuracy by combining the strengths of CNN and LSTM models. [\[Link\]](#)

2.3. Multi-Type Features Separating Fusion Learning For Speech Emotion Recognition:

This research focuses on the fusion of different types of features (e.g., acoustic and linguistic) for emotion recognition from speech shows that that separating and then fusing these features leads to better performance compared to using them together from the start. [\[Link\]](#)

3. Timeline (Tentative 12 Weeks)

Week 1-2 : Data Collection, Pre-processing and Data Visualization for Gender Recognition.

Week 3-4 : Logistic Regression, Naïve Bayes, Support Vector Machines (SVM).

Week 5-6 : Analysis and Performance of models for Gender, Hyperparameter Tuning, Check For Model Overfitting and Underfitting.

Week 7-8 : Data Collection, Pre-processing and Data Visualization for Emotion Recognition.

Week 9-10 : Random Forest, Decision Trees for Emotion Recognition, Support Vector Machines(SVM), Integration of Both the Models.

Week 11 : Analysis and Performance of models for Emotion ,Hyperparameter Tuning, Check For Model Overfitting and Underfitting.

Week 12 : Report Writing.

4. Individual Task

Tasks	Team Member /s
Data Collection	Anikait and Anant
Pre-processing and Data Visualization	Abdullah And Ansh
Feature Extraction	Ansh
Logistic Regression, Naïve Bayes,	Anikait And Abdullah
Decision Trees, Random Forest, Support Vector Machine (SVM)	Anant And Ansh
Analysis and performance of models	Anikait And Ansh
Hyperparameter Tuning, Check for model Overfitting and Underfitting	Abdullah And Anant
Report Writing	Ansh, Anikait, Abdullah, Anant

5. Final Outcome

The objective and ultimate goal of the project is to develop and design a highly accurate Machine Learning model capable of classifying and identifying both gender and emotional states from speech, enhancing the way machines understand and interact with human emotions and identities. Correctly identifying gender helps tailor user interactions to be more relevant, whereas detecting emotions is essential for building systems that respond with empathy and understanding. The project focuses on extracting key audio features such as pitch tone and rhythm, which are critical for distinguishing gender and emotions. These models will undergo thorough assessment using metrics such as Accuracy, Precision, Recall, and F1-Score.