



JOTHIKA K

Final Project

PROJECT TITLE

Adversarial Training for robustness
in speech emotion recognition with CNN

AGENDA

- 1.Introduction to speech recognition
- 2.Project overview
- 3.End users for this project
- 4.Solution and its value proposition
- 5.The wow in my solution
- 6.Modelling
- 7.Results



PROBLEM STATEMENT



Speech is a fundamental mode of human communication, carrying not only semantic content but also rich emotional cues. In this project I took Speech emotion recognition using CNN(convolutional Neural Network). Understanding and recognizing emotions from speech can greatly enhance human-computer interaction, ranging from customer service applications to mental health monitoring systems . However, accurately discerning emotions solely from speech signals poses significant challenges due to the complexity and variability of human emotions, linguistic expressions, and environmental factors.



PROJECT OVERVIEW

It Introduce the project's aim to develop a system capable of recognizing emotions from speech signals. Highlight the significance of speech emotion recognition in improving human-computer interaction, mental health assessment, and other applications. This project begins with the selection and acquisition of a suitable dataset containing audio recordings of speech with labeled emotional states. Preprocessing steps are applied to clean the data, remove noise, and extract relevant features from the audio signals. The developed speech emotion recognition system has potential applications in various domains, including virtual assistants, sentiment analysis in customer service, and mental health monitoring.



WHO ARE THE END USERS?

End users for a speech emotion recognition project could vary depending on the intended application and context.

Education Sector: Teachers and educators could use speech emotion recognition to assess student engagement and emotional state during lectures or online learning sessions, adapting their teaching methods accordingly.

Mental Health Professionals: Speech emotion recognition systems could assist mental health professionals in diagnosing and monitoring conditions such as depression, anxiety, or PTSD by analyzing the emotional content of speech.

Entertainment Industry: Gaming companies and virtual reality developers could integrate speech emotion recognition to create more immersive experiences by adjusting gameplay based on the player's emotions and reactions.

YOUR SOLUTION AND ITS VALUE PROPOSITION

The solution for a speech emotion recognition project entails the development of a sophisticated system capable of accurately detecting and interpreting emotions conveyed through speech. In healthcare, therapists and mental health professionals can utilize it to monitor patients' emotional states, facilitating more effective assessment and treatment. Market researchers gain deeper insights into consumer sentiment, guiding product development and marketing strategies. In education, adaptive learning environments are created, where educators can tailor their approaches based on students' emotional engagement. Security measures are bolstered with the ability to detect emotional distress, enhancing safety in various settings.



THE WOW IN YOUR SOLUTION

High Accuracy: Achieving exceptionally high accuracy in detecting and interpreting emotions from speech signals can be a significant "wow" factor. This could involve surpassing industry benchmarks or even human-level performance in emotion recognition tasks.

Real-time Processing: Implementing real-time emotion recognition capabilities can be impressive, especially in applications like customer service or gaming where immediate responses are crucial for user engagement and satisfaction.

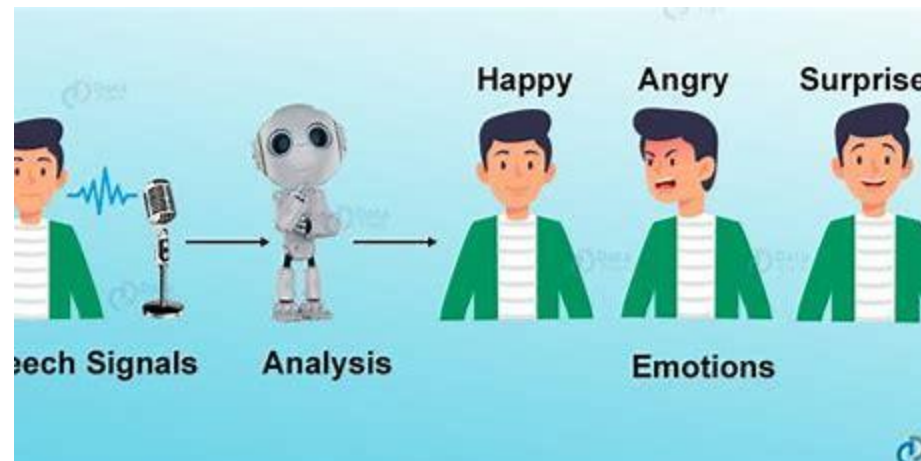
User Experience Enhancement: Implementing features that significantly enhance user experience, such as providing real-time feedback on emotional states during interactions with virtual assistants or educational tools, can be a noteworthy "wow" factor.



MODELLING

Data Collection and Preprocessing: Gather a dataset of speech samples labeled with corresponding emotion categories (e.g., happy, sad, angry). Preprocess the audio data, which may include tasks such as noise removal, normalization, and feature extraction.

Model Selection: Particularly, convolutional neural networks (CNNs) are widely used for learning representations from audio data.



Model Training: Train the selected model on the labeled dataset using appropriate training algorithms and optimization techniques.

Model Evaluation: Evaluate the trained model's performance on a held-out test set to assess its generalization ability.

Monitoring and Maintenance: Continuously monitor the model's performance in production to detect drift or degradation over time. Retrain or update the model as necessary to maintain optimal performance and adapt to evolving data distributions or user preferences.

RESULTS

In this study, we developed a speech emotion recognition system using methodology or model used, e.g., deep learning techniques such as recurrent neural networks (CNNs). Our primary goal was to accurately classify emotions from speech data into predefined categories such as happiness, sadness, anger, and neutral.

In conclusion, our study contributes to the growing body of research in speech emotion recognition and provides insights into the development of effective models for automatic emotion detection from speech. While there is still room for improvement, the advancements made in this project pave the way for applications in areas such as human-computer interaction, affective computing, and mental health monitoring.