**SYNOPSIS**

**MINOR PROJECT REPORT**

**ON**

**“**Speech Emotion Recognition”

Logo

Description automatically generated

Submitted By:

Nitika Devgan

Submitted To:

Sushma Mam

**Table of Contents**

|  |
| --- |
| Contents |
| 1. Project Objective |
| 1. Project Scope |
| 1. Methodology |
| 1. Tools / Platform, Hardware and Software Requirement specifications |
| 1. Pictorial Representation |
| 1. Conclusion |
| 1. References |

1. Title of the Project:

**“Speech Emotion Recognition”**

2. Project Objective:

The main objective of this project is to build a model that recognize emotion from speech using the librosa and sklearn libraries and the RAVDESS dataset.  Speech emotion recognition is a technology that extracts emotion features from computer speech signals, compares them, and analyzes the feature parameters and the obtained emotion changes.Recongnizing emotions from audio signals requires feature extraction and classifier training.

The feature vector is composed of audio signal elements that characterize the specific characteristics of the speaker (such as pitch, volume, energy), which is essential for training the classifier model to accurately recognize specific emotions.

3. Project Scope:

Emotion is an integral part of human behavior and inherited property in all mode of communication. We, human is well trained thought your experience reading recognition of various emotions which make us more sensible and understandable. But in case of machine, however, it can easily understand content-based information such as information in text, audio or video but still far behind to access the depth behind the content. It is need of the era that machine should also be trained to understand emotions correctly for better understanding and to avoid any miscommunication. Present study comes into domain of emotion recognition from audio conversation. Moreover, Audio emotion analysis has many applications in various sectors like healthcare, banking, defense, and IT. On the other part, text emotions are easy to decode as there is no role of factors like tone and pitch, but in case of audio emotion analysis both the factors need attention for better accuracy. Also there are several factors like noise, disturbance, and various pauses in communication which results in degrading the accuracy. It is challenging task to make machine to understand the emotion of the respondents.

From the preceding flowchart, some of the factors that influence the accuracy of speech emotion recognition are the size and quality of emotion corpus, the extraction of features, and the selection of classifiers. Hence, three aspects comprise the primary focus of this research: 3 emotion corpus, feature extraction and classifier design. Among the three, only emotion corpus and feature extraction are generally involved in this paper, which will be the basis for this section.

**Proposed System and Advantages**

The proposed system involves using machine learning algorithms to analyze voice signals and extract features related to the emotional state of the speaker. The system can be trained on a dataset of voice recordings with labeled emotions and can then be used to classify the emotions of new voice inputs. The advantages of the proposed system include:

* Real-time emotion detection
* Non-intrusive and non-contact method
* Ability to analyze emotions from remote locations.
* Higher accuracy and consistency than subjective interpretation
* Use Cases
* The Voice Emotion Detector can be used in various fields, including:
* Mental health: to assist therapists and counselors in monitoring the emotional states of their clients.
* Customer service: to provide better and personalized support to customers based on their emotional needs.
* Marketing: to analyze the emotional responses of users to advertisements and products
* Education: to monitor the emotional states of students and provide personalized support

3. Methodology:

Stage 1: Ideation

The idea hit when I realized that there are so many things around you which you don’t want and also want to sell those things and to buy anything in reasonable price so from there it hit there can be a platform where buying and selling can be done but in addition there will be bidding on products which we help to sell a product at high price or at reasonable price. And to buy anything from this platform.

Stage 2: Choosing the tech stack.

The Voice Emotion Detector project is developed using Python programming language and various libraries, such as:

* librosa: for audio signal processing and feature extraction

**Librosa** is powerful Python library built to work with audio and perform analysis on it. It is the starting point towards working with audio data at scale for a wide range of applications such as detecting voice from a person to finding personal characteristics from an audio.

It helps us to implement audio signal analysis for music, reference implementation of common methods and buildings blocks for Music information retrieval(MIR).

* numpy: for numerical computing and array operations
* matplotlib: for data visualization
* scikit-learn for machine learning algorithms and models

Stage 3: Wireframes and prototype creation

A wireframe is just an outline or schema for a potential web page. It offers the appearance of a page's grid and contains all the blocks that need to be positioned on the page. It is possible to give a more accurate estimate and sweat the details of the functionality when there are wireframes present.

Stage 4: UI design

 Here creation of user interface and incorporate brand colours and aspects after the wireframes are finished. Then design assets will be the focus of work.

Stage 5: Building the website.

 The majority of website development happens at this phase. The coding part begins here. The frontend entails creating all design elements that visitors will see, whereas the backend entails any scripting, coding, and backend operations. Once the details are worked out, it's time to start getting ready for launch and beyond.

Stage 6: Testing and Deployment

In this part testing of all the features and functionality that have been planned works out here. Before deployment, we'll try to find and fix any minor errors there may be. After all of the tests are finished, the website will be hosted on any appropriate platform.

Stage 7: Maintenance

After the launch of website, there is always work to be done. It's crucial to consider potential areas for development as your website develops.

4. Facilities required for proposed work: sw /hw required for development .

HARDWARE:

Processor : CORE13

Hard Disk : 250 GB

RAM :8GB

SOFTWARE:

Operating System : WINDOWS

Programming Language : PYTHON

System Architecture

Diagram

Description automatically generated

6. Pictorial representation through Diagrams

Use Case Diagram

Diagram

Description automatically generated

Activity Diagram

Input Audio

Load Trained Model

Audio Recognition

Saving Audio File

Audio Extraction

Feature Extraction

Predicting Emotion

Display Emotion

7.CONCLUSIONS

The Voice Emotion Detector project is a promising application of machine learning and signal processing in the field of emotion detection. The proposed system has the potential to revolutionize the way emotions are analyzed and understood and can be applied in various fields for better human interaction and support.

A multi-feature fusion and Multi-lingual fusion speech emotion recognition algorithm is proposed based on the RNN with improved local attention mechanism. Four Chinese and English speech emotion datasets are used in the process. Three sets of contrast experiments are then employed to choose the best feature combination, verify its effect on Multi-lingual fusion and prove the effectiveness of the improved model. The said experiments reveal that the proposed algorithm can effectively enhance speech recognition accuracy. Such algorithm denotes a useful reference in resolving cases in which insufficient speech emotion data are experienced.

8. REFERENCES

* <https://deliverypdf.ssrn.com/delivery.php?ID=007123072013099125001127097119127069029073072014083066027080065092082012092095104024043042013058038026045074016100001105098110006016023001060066090067109019088026113023059078090108076119003077024116084071066028025002108110085084027065081006023027081090&EXT=pdf&INDEX=TRUE>
* <https://arxiv.org/ftp/arxiv/papers/2001/2001.05908.pdf#:~:text=The%20present%20acoustic%20features%20typically,depends%20on%20the%20frame%20involved>.