**Project Topic: Speech Emotion Recognition**

**Abstract**

As human beings speech is amongst the most natural way to express ourselves. We depend so much on it that we recognize its importance when resorting to other communication forms like emails and text messages where we often use emojis to express the emotions associated with the messages. As emotions play a vital role in communication, the detection and analysis of the same is of vital importance in today’s digital world of remote communication. Emotion detection is a challenging task, because emotions are subjective. There is no common consensus on how to measure or categorize them. We define a SER system as a collection of methodologies that process and classify speech signals to detect emotions embedded in them. Such a system can find use in a wide variety of application areas like interactive voice based-assistant or caller-agent conversation analysis. In this study we attempt to detect underlying emotions in recorded speech by analysing the acoustic features of the audio data of recordings.

**Introduction**

Emotion recognition in spoken dialogues has been gaining increasing interest all through current years. Speech Emotion Recognition (SER) is a hot research topic in the field of Human Computer Interaction (HCI). It has a potentially wide applications, such as the interface with robots, banking, call centers, car board systems, computer games etc. For classroom orchestration or E-learning, information about the emotional state of students can provide focus on enhancement of teaching quality. Speech emotion recognition is essentially a sequence classification problem, where the input is a variable-length sequence and the output is one single label. That is why we have chosen recurrent neural networks in our work. In this experimental work, we have used Multivariate Linear Resgression (MLR), Support Vector Machine (SVM) and Recurrent Neural Networks (RNN) classifiers to identify the emotional state of spoken utterances. In order to demonstrate the high effectivennes of the MFCC and MS features extraction for emotion classification in speech, we provide results on two open emotional databases (Berlin-DB and Spanish-DB). The performance and robustness of the recognition systems will be easily affected if it is not well-trained with suitable database. Therefore, it is essential to have sufficient and suitable phrases in the database to train the emotion recognition system and subsequently evaluate its performance. In this section, we detail the two emotional speech databases used in our experiments: Berlin Database and Spanish Database.

**Aim And Objectives**

Communication is the key to express oneself. Humans use most part of their body and voice to effectively communicate. Hand gestures, body language, and the tone and temperament are all collectively used to express one’s feeling. Though the verbal part of the communication varies by languages practiced across the globe, the non-verbal part of communication is the expression of feeling which is most likely common among all. Therefore, any advanced technology developed to produce a social environment experience also covers understanding emotional context in speech. Improvements in the field of emotion detection positively impact a multitude of applications. Some of the research areas that benefit from automating the emotion detection technique include psychology, psychiatry, and neuroscience. These departments of cognitive sciences rely on human interaction, where the subject of study is put through a series of questions and situations, and based on their reactions and responses, several inferences are made. A potential drawback occurs as few people are classified introverts and hesitate to communicate. Therefore, replacing the traditional procedures with a computer-based detection system can benefit the study. Similarly, the practical applications of the speech-based emotion detection are many. Smart home appliances and assistants (Examples: Amazon Alexa [2] and Google Home [3]) are ubiquitous these days. Additionally, customer care-based call centers often have an automated voice control which might not please most of their angry customers. Redirecting such calls to a human attendant will improve the service. Other applications include eLearning, online tutoring, investigation, personal assistant (Example: Apple Siri [4] and Samsung S Voice [5]) etc. A very recent application could be seen in self-driving cars. These vehicles heavily depend on voice-based controlling. An unlikely situation, such as anxiety, can cause the passenger to utter unclear sentences. In these situations, understanding the emotional content expressed becomes of prime importance.

**Existing System-Disadvantages**

The speech emotion detection system is implemented as a Machine Learning (ML) model. The steps of implementation are comparable to any other ML project, with additional fine-tuning procedures to make the model function better. The flowchart represents a pictorial overview of the process (see Figure 1). The first step is data collection, which is of prime importance. The model being developed will learn from the data provided to it and all the decisions and results that a developed model will produce is guided by the data. The second step, called feature engineering, is a collection of several machine learning tasks that are executed over the collected data. These procedures addressthe several data representation and data quality issues. The third step is often considered the core of an ML project where an algorithmic based model is developed. This model uses an ML algorithm to learn about the data and train itself to respond to any new data it is exposed to. The final step is to evaluate the functioning of the built model. Very often, developers repeat the steps of developing a model and evaluating it to compare the performance of different algorithms. Comparison results help to choose the appropriate ML algorithm most relevant to the problem.

**Proposed System-Advantages**

The problem of speech emotion recognition can be solved by analysing one or more of these features. Choosing to follow the lexical features would require a transcript of the speech which would further require an additional step of text extraction from speech if one wants to predict emotions from real-time audio. Similarly, going forward with analysing visual features would require the excess to the video of the conversations which might not be feasible in every case while the analysis on the acoustic features can be done in real-time while the conversation is taking place as we’d just need the audio data for accomplishing our task. Hence, we choose to analyse the acoustic features in this work.

**FEASIBILITY ANALYSIS**

During system analysis, a feasibility study of the proposed system is carried out to see whether it is beneficial to the users. It is a test of the proposed system, regarding its workability, impact on organization, the ability to meet the user requirements. It’s performed to choose the system that meets the performance requirements at the least cost. Primarily, whether a project that was initiated by one has to be considered for further development or not will be decided through preliminary investigation. As part of this investigation, feasibility analysis of the System was conducted. Based on the outcome of this feasibility analysis, the request for the project is approved. The study was done in four areas:

* Technical Feasibility
* Behavioral Feasibility
* Economic feasibility
* Operational Feasibility
* Legal Feasibility
* Time Feasibility
* **Technical Feasibility**

This study is carried out to check the technical feasibility of the system, that is, the technical requirement of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands being placed on the client. Technical Feasibility centers on the existing computer system hardware, software, etc. and to some extent how it can support the proposed addition. This involves financial considerations to accommodate technical enhancements. Technical support is also a reason for the success of the project.

The techniques needed for the system should be available and it must be reasonable to use. Technical Feasibility is mainly concerned with the study of function, performance, and constraints that may affect the ability to achieve the system. By conducting an efficient technical feasibility we need to ensure that the project works to solve the existing problem area.

**Table 2.1: Technical feasibility Report**

|  |  |  |
| --- | --- | --- |
| **Serial No** | **Technical Feasibility** | **Result** |
| 1 | Does the infrastructure design meet the need specified during the Identification Phase? | Yes. |
| 2 | Do we currently possess the necessary technology? | Yes. |
| 3 | Do we possess the necessary technical expertise and is the schedule reasonable? | Yes. |
| 4 | Is the relevant technology mature enough to be applied to our problem? | Yes. |
| 5 | Can the main technological changes in the service delivery be satisfactorily estimated? | Yes |

* **Behavioral Feasibility**

Proposed projects are beneficial only if they can be changed in to information system that will meet operation requirement of the organization. People are inherently resistant to change and computers have been known to facilitate changes. An estimate should be made of how strong reaction the user staff is likely to have towards the development of a computerized system. The new proposed system is very much useful to users and there for it will accept broad audience from around the world**.Table 2.2: Behavioral feasibility Report**

|  |  |  |
| --- | --- | --- |
| **Serial No** | **Behavioural Feasibility** | **Result** |
| 1 | Was there any negative reaction when the new system was implemented? | No. |
| 2 | Is the system helpful for all kinds of users? | Yes. |
| 3 | Are there a still employee who doesn’t understand the new system? | No. |

* **Economic feasibility**

Economic feasibility study present tangible and intangible benefits from the prefect by comparing the development and operational cost. The technique of cost benefit analysis is often used as a basis for assessing economic feasibility. This system needs some more initial investment than the existing system, but it can be justifiable that it will improve quality of service.

The user should be able to distinguish the existing one and proposed one, so that one must be able to appreciate the characteristics of the proposed system, the manual one is not highly reliable and also is considerably fast. The proposed system is efficient, reliable and also quickly responding. This project is economically feasible. It does not require much cost to be involved in the overall process. The overall objectives are in easing out the requirement processes.

**Table 2.3: Economic feasibility Report**

|  |  |  |
| --- | --- | --- |
| **Serial No** | **Economic Feasibility** | **Result** |
| 1 | Can the project be done within given cost constraints? | Yes. |
| 2 | Can this system be implemented with a minimal cost? | No. |
| 3 | When the new workers were selected, did it exceed the targeted budget? | No. |
| 4 | Was there any need for new hardware’s and software’s? | Yes. |
| 5 | Was there any alternative financial arrangements? | No |

* **Operational Feasibility**

Suppose for moment that technical and economic resources are both judged adequate.  The systems analyst must still consider the operational feasibility of the requested project.  Operational feasibility is dependent on human resources available for the project and involves projecting whether the system will operate and be used once it is installed. If users are virtually led to the present system, see no problems with it, and generally are not involved in requesting a new system, resistance to implementing the new system will be strong.  Chances for it ever becoming operational are low.

**Table 2.4: Operational feasibility Report**

|  |  |  |
| --- | --- | --- |
| **Serial No** | **Operational Feasibility** | **Result** |
| 1 | Does the management support the project? | Yes. |
| 2 | Can or will the end users and management adapt to the new system? | Yes. |
| 3 | Will there be a change in the working environment of the end users? | Yes. |
| 4 | Was there any negative comment from end users? | No |

* **Legal Feasibility**

It determines whether the proposed system conflicts with legal requirements, e.g., a data processing system must comply with the local data protection regulations and if the proposed venture is acceptable in accordance to the laws of the land. This assessment investigates whether any aspect of the proposed project conflicts with legal requirements like zoning laws, data protection acts or social media laws.

* **Time Feasibility**

Time feasibility is a measure of how reasonable the project timetable is. Given our technical expertise, are the project deadlines reasonable? Some projects are initiated with specific deadlines. It is necessary to determine whether the deadlines are mandatory or desirable.

**SYSTEM REQUIREMENT**

Front End : HTML5, CSS3, JavaScript, Bootstrap

Back End : Mysql 5.1 or above

Technology : Python –Flask Framework

Platform : Windows 10

Web Server : Mysql Server2010

Tools : Sublime, Sqlyog, Anaconda Navigator

Web browser : Google Chrome

**Module Description**

**1.Admin**

* Login
* Send Response to Feedback
* User Details

**2. User**

* **Register**
* **Login**
* **Send Feedback**
* **Edit/View/Delete My Profile**
* **Speech Emotion Recognition**

**Tables:**

**1.Login-login\_id,username,password,Usertype**

**2.user-user\_id,fname,lname,place,email,phone**

**3.feedback=feed\_id,user\_id,feedback,reply,date**

**Future Scope**

1. Data Processing. So to start, we need audio data of human voices with labeled emotions. ...
2. Acoustic Feature Extraction. Now we are all set with the audio files and labels. ...
3. Filtering and Splitting the Dataset. ...
4. Acoustic Model Building and Scoring Using Deep Learning.

**Conclusion**

Speech Emotion Recognition (SER) is the **task of recognizing the emotion from speech** irrespective of the semantic contents. However, emotions are subjective and even for humans it is hard to notate them in natural speech communication regardless of the meaning.Speech emotion analysis refers **to the use of various methods to analyze vocal behavior as a marker of affect** (e.g., emotions, moods, and stress), focusing on the nonverbal aspects of speech.