Sentiment Analysis Using Speech Recognition

PROJECT ID (PCSE-08)

PROJECT SYNOPSIS
OF MAJOR PROJECT

BACHELOR OF TECHNOLOGY
COMPUTER SCIENCE AND ENGINEERING

SUBMITTED BY

Team leader-Alisha Raghav(2100290100019)
Team member-Dhruvi Gupta(2100290100056)
Team member-Aakansha Mittal(2100290100001)

Project Guide
Prof. Hriday Kumar Gupta

KIET Group of Institutions, Delhi-NCR, Ghaziabad(UP)
Department of Computer Science and Engineering
October 202

Table of contents

Introduction

Sentiment Analysis refers to a group of procedures used to analyse textual data as a component of NLP in order to identify the sentiment included in the input text.

Even if it's not common, this feeling can typically be categorised as good, negative, or neutral. Instead, you might wish to construct a model that finally produces a polarity with a wider range, such as extremely positive, positive, neutral, negative, and very negative. In addition, you may attempt to forecast an emotion (such as happiness, sadness, anger, etc.) rather than a polarity.

Sentiment analysis has advanced significantly over the past few years as a result of the use of machine learning and deep learning techniques, which can complete this task quickly and accurately.

By highlighting the improvement in valence prediction using the concordance correlation coefficient loss function, it aims to incorporate sentiment into speech emotion identification. It emphasises on how non-linguistic information must be understood by machines in order for there to be natural human-machine interactions.

It explores how sentiment analysis (SA) and emotion recognition (ER) have converged, and it makes the case for comparing the two tasks using the same model. Speech analysis benefits of direct sentiment extraction from acoustic variables are discussed, along with the integration of multimodal data for enhanced performance.

With the advent of deep learning algorithms and the availability of big, labelled datasets, speech recognition with NLP continues to advance. It has developed into a crucial component of contemporary technology, enabling voice-activated gadgets and enhancing human-computer connection.

Rationale

- **Tech Landscape Significance**: "Sentiment Analysis using Speech Recognition" project addresses a critical need in the tech landscape.
- **Emotion Understanding**: Focuses on understanding emotions conveyed through speech.
- **Healthcare Applications**: Essential for remote patient monitoring, enabling timely interventions during telemedicine.
- **Enhanced Patient Care**: Utilizes emotional cues for improved patient care and outcomes.
- **Educational Impact**: Enhances education by increasing student engagement through the analysis of spoken responses.
- **Teaching Adaptation**: Allows educators to adjust teaching strategies based on student emotions.
- **Content Platform Enhancement**: Tailors recommendations to user sentiment, boosting satisfaction and engagement.
- **Improved Human-Machine Interactions**: Ultimately elevates the quality of interactions between humans and machines.

Objectives

The project aims to achieve the following objectives:

- The idea behind creating this project is to build a machine learning model that could detect emotions from the speech we have with each other all the time.
- Develop a robust speech recognition system to convert spoken language into text.
- Implement a sentiment analysis algorithm to determine the sentiment (positive, negative, neutral) of the transcribed text.
- Create a user-friendly interface to input spoken content and visualise sentiment results.

Literature Review

1.Speech Emotion analysis using machine learning for depression recognition.

In order to identify depression using speech analysis, the paper investigates a range of methodology and datasets, from acoustic characteristics and classification algorithms to deep learning methods and multimodal approaches. The purpose of the study is to improve the precision and efficacy of speech data-based depression diagnosis.

2.Sentiment Analysis on Speaker Specific Speech Data

This paper proposes a generalised model that analyses the content and speakers' identities in an audio file containing a discussion between two persons. It does this by automatically turning the audio file into text and conducting speaker recognition. In this study, we suggested a straightforward system to carry out the aforementioned work. The system performs well with the artificially generated dataset, and we are currently working to expand the dataset and improve the system's scalability.

The system currently only has the ability to handle conversations between two speakers, where only one speaker should speak at a time. Although the system is accurate at understanding the sentiment of the speakers in conversational dialogue, it has some limitations.

3.Sentiment Analysis and Emotion Recognition from Speech Using Universal Speech Representations

Also, a study presented a deep neural network-based method for comprehending spoken language that combines sentiment analysis & speech recognition. The system employed recurrent neural networks (RNNs) for sentiment analysis after using convolutional neural networks (CNNs) to extract acoustic information from speech inputs. The ability to identify emotions from spoken language using this deep learning model has shown encouraging results.

4. Sentiment extraction from natural audio streams.

Overall, Speech recognition systems now include sentiment analysis. The classification of spoken utterances into subjective and objective categories was done by the authors using traditional machine learning approaches. By removing meaningless or deceptive phrases from the speech data, our method enhanced the polarity classification's accuracy

Feasibility Study

Making a sentiment analysis project using speech recognition feasible involves careful planning, resource allocation, and addressing potential challenges. Here are some steps to ensure the feasibility of your project:

Project Scope Definition: Clearly define the scope of your project. Decide whether you'll focus on a specific domain (e.g., customer reviews, social media comments) and the languages you'll support. This helps manage expectations and prevents scope creep.
Technical Resources: Evaluate the availability of the required hardware and software. Ensure that you have access to computers with sufficient processing power and memory for training machine learning models.
Data: Ensure you have access to a substantial amount of labelled audio data for training your models. This data should cover a diverse range of sentiments and speakers.
Technical Skills: Assess your team's technical skills in programming, machine learning, speech recognition, and sentiment analysis. If necessary, plan for skill development or consider collaborating with experts.
Learning Curve: Consider the time needed to learn new tools, libraries, or techniques.
Break down your project into smaller milestones with specific tasks and deadlines. This helps in tracking progress and managing time effectively.
Noise Handling: Real-world audio data often contains background noise. Research and implement noise reduction techniques to improve the accuracy of speech recognition.
Accurate Transcription: Speech recognition may not be 100% accurate. Plan for post-processing or additional data cleaning steps to enhance the quality of transcribed text.
User Interface Design:
Design an intuitive user interface that's easy to use and visually appealing. User feedback during development can help refine the interface.
Hardware and Software Requirements:
Ensure the hardware and software you plan to use are available and compatible. This includes the necessary speech recognition libraries, machine learning frameworks, and development environments.
Testing and Validation:
Allocate time for rigorous testing and validation. Test your system with a diverse set of audio samples to ensure accurate speech recognition and sentiment analysis results.
Backup Plans:Have contingency plans in case certain resources become unavailable or if you encounter unexpected challenges. This might involve finding alternative datasets, switching to different models, or adapting your project's scope.
Be mindful of ethical considerations related to data privacy and usage, especially if you're collecting or utilising user-generated content.
Depending on your project's scale, consider any costs associated with purchasing specialised hardware, cloud services for training models, or acquiring quality labelled datasets.

By addressing these aspects, you can enhance the feasibility of your sentiment analysis project using speech recognition and increase the chances of successful completion within the defined constraints.

Methodology

Building a sentiment analysis project using speech recognition in Python involves several steps. Here's a high-level overview of the process:

Step 1: Data Collection
Collect audio data containing spoken content with labelled sentiment. You can use publicly available datasets or record your own audio samples. Make sure to label each audio sample with its corresponding sentiment label (positive, negative, neutral).

Step 2: Audio Preprocessing
Preprocess the audio data to convert it into a format suitable for speech recognition. You can use libraries like librosa to extract features from the audio, such as Mel-frequency cepstral coefficients (MFCCs), which are commonly used for speech analysis.

Step 3: Speech Recognition
Use a speech recognition library like SpeechRecognition or pyAudio to convert the preprocessed audio into text. These libraries usually provide interfaces to popular speech recognition engines like Google Web Speech API or CMU Sphinx.

Step 4: Sentiment Analysis Model
Train a sentiment analysis model using the transcribed text from the speech recognition step. You can use machine learning frameworks like scikit-learn or deep learning frameworks like TensorFlow or PyTorch for this purpose. Your model should be trained on the labeled audio data to predict sentiment labels.

Step 5: User Interface
Create a user interface that allows users to input audio files or record their own speech. The interface should utilise the speech recognition module to convert the spoken content into text and then pass the text through the sentiment analysis model for sentiment classification. You can use libraries like tkinter for creating graphical user interfaces.

Step 6: Integration and Testing
Integrate the speech recognition and sentiment analysis components within the user interface. Test the system using various audio samples with different sentiments to ensure accurate sentiment classification.

Step 7: Deployment

Once you are satisfied with the functionality and accuracy of your project, you can consider deploying it. You can deploy it as a standalone application, a web application, or even integrate it into existing platforms.

Facilities required for proposed work

1. Audio input: A device that captures audio, such as a microphone, to record spoken language.
2. Audio Preprocessing:There are technologies for enhancing and reducing noise in audio that can help speech recognition be more accurate.
3. Natural Language Processing (NLP) Tools:Using NLP packages or APIs, the transcribed speech can be converted to text and the text's useful linguistic properties can be extracted. The Natural Language Toolkit, spaCy, and NLTK are three well-known NLP libraries.
4. Datasets: A set of speech data with labels for use in training and perfecting your sentiment analysis model. Transcripts of speeches that have related sentiment labels should be included in this collection.
5. Computing Resources: A sufficient amount of computational capacity, like a powerful computer or cloud-based servers, to process and analyze big audio files.
6. Data Retention: A database or file storage system for sentiment analysis findings, transcriptions, and audio recordings.
7. User Interface:If the sentiment analysis system is designed to be user-facing, a user interface (UI). Users of this UI can receive the findings of sentiment analysis in the form of sentiment scores or illustrative displays of sentiment.
8. Testing and Evaluation: A method for measuring the precision and potency of your sentiment analysis model. This includes assessing performance using test and validation datasets.
9. Updates and Continuous Monitoring: A strategy for constant evaluation and revision of the sentiment analysis model to guarantee its accuracy throughout time.
10. Ethics and Legal Aspects:When using speech data, be mindful of the legal and moral ramifications, especially with regard to privacy, permission, and data protection laws.Keep in mind that sentiment analysis is a challenging task, and your system's precision will

Expected outcomes

The expected outcomes of the "Sentiment Analysis using Speech Recognition" project encompass several key deliverables. Firstly, we anticipate the development of a robust and accurate sentiment analysis model that can effectively detect and interpret emotions in spoken language. This model is expected to exhibit high precision and recall rates. Additionally, we aim to create a user-friendly interface or application that integrates this technology, making it accessible and practical for end-users. Furthermore, we anticipate the generation of valuable insights into various fields, such as customer service, healthcare, education, and content recommendation, where sentiment analysis can be applied to enhance user experiences, decision-making, and engagement. Ultimately, the project's success is expected to contribute to the advancement of emotionally-aware AI technologies, improving human-machine interactions across diverse sectors.

References[IEEE format]

1.Chaudhari, Shivaji J., and Ramesh M. Kagalkar. "Automatic speaker age estimation and gender dependent emotion recognition." International Journal of Computer Applications 117, no. 17 (2015).

2. Paula Lopez-Otero, Laura Docio-Fernandez, “Analysis of gender and identity issues in depression detection on de-identification speech”, Computer Speech & Language, E.E Telecommunication campus – 2021 dio.org/10.1016/j.csl.2020.101118

3. https://getthematic.com/sentiment-analysis/ (2023)

4. gan, Chenquan, Lu Wang, Zufan Zhang, and Zhangyi Wang. "Sparse attention based separable dilated convolutional neural network for targeted sentiment analysis." Knowledge-Based Systems 188 (2020).