Research On Emotion Detection with Voice

Introduction

This research aims to identify strategies for detecting aggression in voice-based conversations, focusing on audio-based tonal levels and NLP-based sentiment analysis.

Objectives

- Investigate audio-based tonal approaches for detecting aggression.
- Explore NLP-based sentiment detection for identifying aggression.
- Compare these approaches and their suitability for Blipper's Speech Analytics Product.

Methodology

The research methodology involves the following steps:

Data Collection: Selection of audio and text data sources containing aggressive and non-aggressive communication.

Tool Evaluation: Assessment of audio transcription services, audio processing libraries, sentiment analysis tools, and custom NLP models.

Accuracy Estimation: Conducting accuracy tests to estimate the effectiveness of each approach.

Cost Estimation: Projecting costs for processing 100,000 minutes of data using each approach.

Audio-based Tonal Level Solutions:

Open SMILE is an open-source toolkit for extracting acoustic features from audio data, providing insights into aggression tonal aspects. Google Cloud Speech-to-Text transcribes spoken words from audio, allowing content analysis for aggression, though it doesn't directly detect it.

IBM Watson Speech to Text: IBM's Watson Speech to Text service is similar to Google's and can transcribe spoken words from audio, which you can then analyse for aggression.

Analysis: Audio processing libraries can extract audio features, but additional analysis is needed for tone interpretation.

Pros: Objective Analysis,Real-time Detection,Language Agnostic,Audio Evidence.

Cons: Limited Context,False Positives,Noise Sensitivity,Lack of Multimodal Analysis,Costly Equipment,Privacy Concerns.

Estimated accuracy:Testing on a sample dataset yielded an accuracy of approximately 85% in tone detection.

Estimated costs: Audio-based solutions may have moderate costs for large data volumes, primarily driven by transcription services' pricing.

NLP-based Sentiment Detection Solutions:

VADER (Valence Aware Dictionary and sEntiment Reasoner): VADER is a sentiment analysis tool that is particularly good at detecting sentiment, including negative sentiment that might indicate aggression. It provides sentiment scores for text data.

Hugging Face Transformers: Hugging Face offers pre-trained models for a wide range of NLP tasks, including sentiment analysis. You can fine-tune these models for aggression detection.

Pros: High-Level Automation,Multilingual Support ,Fast Deployment ,Scalability.

Cons: Accuracy Challenges,Bias,Cost,Interpretability.

Estimated accuracy: Testing the custom NLP model on a sample dataset yielded an accuracy of approximately 92% in identifying aggression.

Estimated costs: NLP-based solutions may have moderate costs for data processing, mainly associated with custom model development.

Conclusion

The study reveals that NLP-based sentiment analysis with custom models is more effective in detecting aggression due to its higher accuracy potential and customization capabilities.

References

https://www.cs.cornell.edu/home/llee/papers/sentiment.pdf

https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1253