CUNEF UNIVERSIDAD

http://computational-intelligence.com/

group

FLINS-ISKE 2024

Voice-based mood recognition: An application to Mental Health

R. Hidalgo, V. López and D. Urgelés

Quantitative methods, CUNEF University, Madrid, 28040, Spain
ricardo.hidalgo@cunef.edu

Table of contents

Introduction

- The research focuses on using voice data to predict episodes of emotional disturbance and enable early interventions[1].
- Today's technology allows real-time monitoring of individual behavior through activity sensors[2], and this is especially useful in the study of emotional disorders[3] where clinical data may be limited.[4]
- Bip4Cast [5] project (ACERTA) to improve the prediction and treatment of emotional disorders [6]. Mood states vs. Emotional states

Data Collection

Language	Samples
German voices	500
Spanish voices	1500

Public anonymized audio databases

Both datasets contain the voices of women and men reading a series of sentences with different emotional states (happy, anxious, fearful, disgusted, angry, and different kinds of neutral, such as fast or slow neutral)

Voice signal processing

1.- The audio must be in WAV format, and in this work at 256kb/s, 16kHz, 1 channel, PCM (Little/Signed)

2.- → Parselmouth [7] metadata extraction →

3.- Generation of metadata file

Voice metadata

- Acoustic amplitude - Perceived intensity or volume.
- Pitch - Refers to the perception of the frequency of a sound.
- Harmonicity - Presence of harmonics. It allows one to distinguish between different instruments and voices.
- Intensity - Perceived loudness
- Speech rate - Speed at which someone speaks
- Articulation rate - Clarity and precision of speech
- Energy - Strength or amplitude of a voice signal
- MFCC 1:12 - The spectral envelope of a sound, capturing the unique characteristics of how energy is distributed across different frequencies.
- Formants 1:5 - Resonant frequencies of the vocal tract that shape the sound

Finally we will have → 119 variables ← including the class

Classification system

- Why SVM?
  - Already used in the classification of emotional states.
  - Higher performance than CNN [8]
- SVM + RBF (Radial Basis Function) kernel
  - Objective function

$$Minimize \frac{||\bar{w}||^2}{2} + C\sum_i \xi_i \qquad \gamma = \sum_i \xi_i$$

being: $\bar{w}$, control variables vector
ξ_i, slack variable is introduced for each data point.
C, regularization parameter.
γ, defines how much influence a single training example has

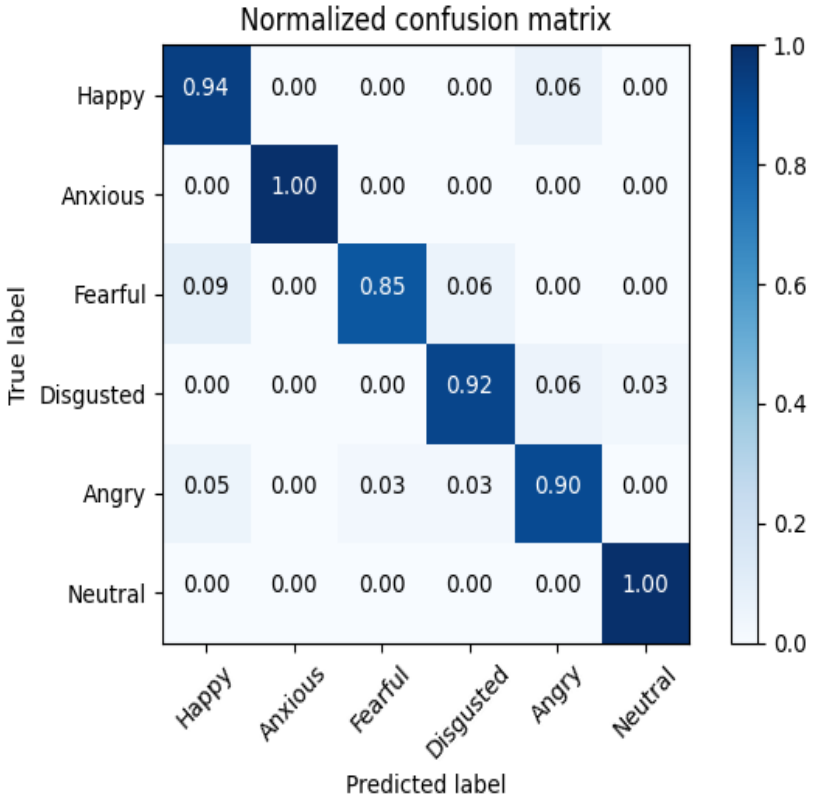
The proper choice of C and γ is critical for SVM performance.
Automatic process.

- Extra Trees classifier to determine the main sound features for the SVM model.

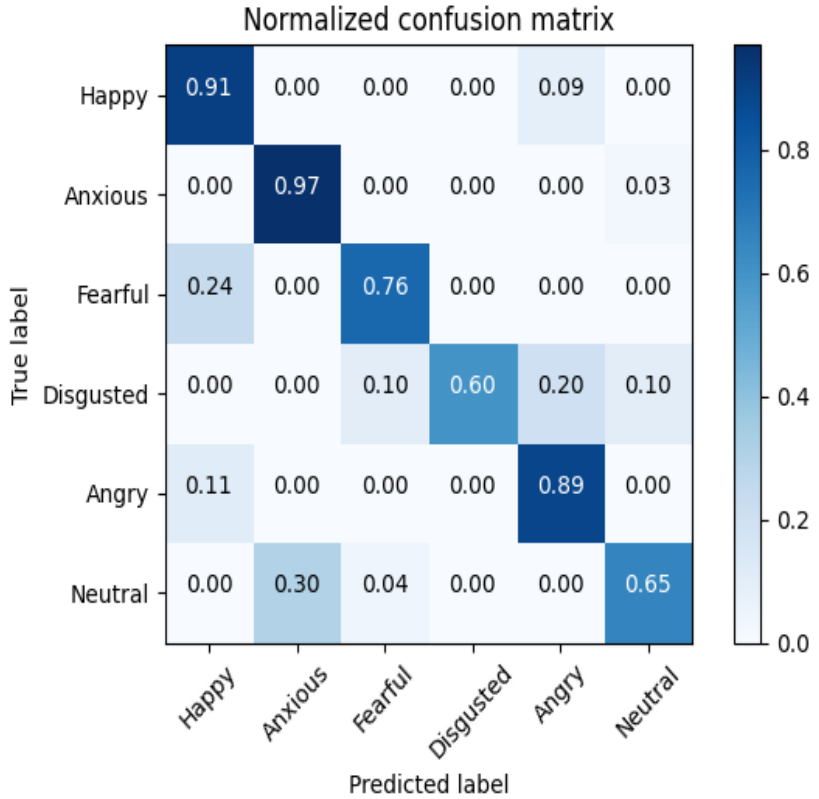
Bip4Cast Data integration

- Bip4Cast project aims to predict mood states: depression, mania, mixed, or euthymia
- This work analyzes emotions detected through voice recordings: happiness, sadness, rage, anger, irritability, panic, anxiety, exaltation, and fear
- Emotions and mood states are related [9]
- Integrating the classifications[10] into the Bip4Cast project will be done by assigning a dimension 3 soft vector indicating Depressive, Manic, or Euthymic degrees in the fuzzy logic sense.
  I.e.: (.2, .7,.1) $\Rightarrow$ state closer to mania than to euthymia

Classification result



(a) Spanish audios



(b) Mixed Audios

Classification result details

Mapping emotions into mood states

Degrees of membership[9] normalized

Emotion	depressive	maniac	mixed
happiness	.138	.552	.309
sadness	.345	.309	.346
irritability	.183	.440	.377
panic	.270	.309	.421
anxiety	.288	.328	.383
anger	.175	.474	.351

Conclusions

- Audio records are useful in determining the emotion expressed by individuals since these records can be successfully classified using machine learning techniques.
- Voice records are useful to enrich the information collected about an individual with other techniques in previous projects, such as the Bip4cast project where motion sensor data and clinical data are used.

Future work

- Integrate the resulting information into previous work.
- Fuzzy modeling of the voice variables.

References I

[1] Pavél Llamocca, Axel Junestrand, Milena Čukić, Diego Urgelés Puértolas, and Victoria López. Data source analysis in mood disorder research. 2018.

[2] Pavel Llamocca, Victoria López, and Milena Čukić. The proposition for bipolar depression forecasting based on wearable data collection. *Frontiers in Physiology*, 29(12):doi: 10.3389/fphys.2021.777137., 2022.

[3] Milena Čukić and Victoria López. Progress in objective detection of depression and online monitoring of patients based on physiological complexity. *Frontiers in Psychiatry*, 13:doi: 10.3389/fpsyt.2022.828773, 2022.

[4] Victoria López, Gabriel Valverde, Julio Anchiraico, and Diego Urgelés. Specification of a cad prediction system for bipolar disorder. In *Uncertainty Modelling in Knowledge Engineering and Decision Making: Proceedings of the 12th International FLINS Conference*, pages 162–167. World Scientific, 2016.

[5] Pavel Llamocca, Diego Urgelés, Milena Čukić, and Victoria Lopez. Bip4cast: Some advances in mood disorders data analysis. In *Proceedings of the 1st International Alan Turing Conference on Decision Support and Recommender Systems, London*, 2019.

[6] Pavel Llamocca, Victoria López, Matilde Santos, and Milena Čukić. Personalized characterization of emotional states in patients with bipolar disorder. *Mathematics*, 9(11):1174, 2021.

[7] Yannick Jadoul, Bill Thompson, and Bart de Boer. Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, 71:1–15, 2018.

[8] Sharifalillah Nordin Mohammad Shahrul Izham Sharifuddin and Azliza Mohd Ali. Comparison of cnns and svm for voice control wheelchair. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 9:387, 2020.

[9] Chantal Henry, Katia M'Baïlara, Alain Desage, Sébastien Gard, David Misdrahi, and Eduard Vieta. Towards a reconceptualization of mixed states, based on an emotional-reactivity dimensional model. *Journal of Affective Disorders*, 101(1-3):35–41, 2007.

References II

[10] Milena Čukić, Victoria López, and Juan Pavón. Classification of depression through resting-state electroencephalogram as a novel practice in psychiatry. *Journal of Medical Internet Research*, 22(11):e19548, 2020.