Binary classification by voice using machine learning in the Saarbrücken Voice Database

Introduction

Results of the application of **artificial intelligence** in medicine are very promising. Prediction-based diagnosis can help healthcare professionals to **support the diagnostic process** and the resolution of complex cases. The detection of both physiological and pathological variables **by voice** can be an **harmless, fast and low-cost alternative** for decision making.

Purpose

Classification of physiological and pathological variables by voice with artificial intelligence.

Methodology

Database:
1949 Recordings from Saarbrücker Voice Database pronouncing the vowel phrase "Guten Morgen, wie geht es Ihnen?" (Good mornig, how are you?)

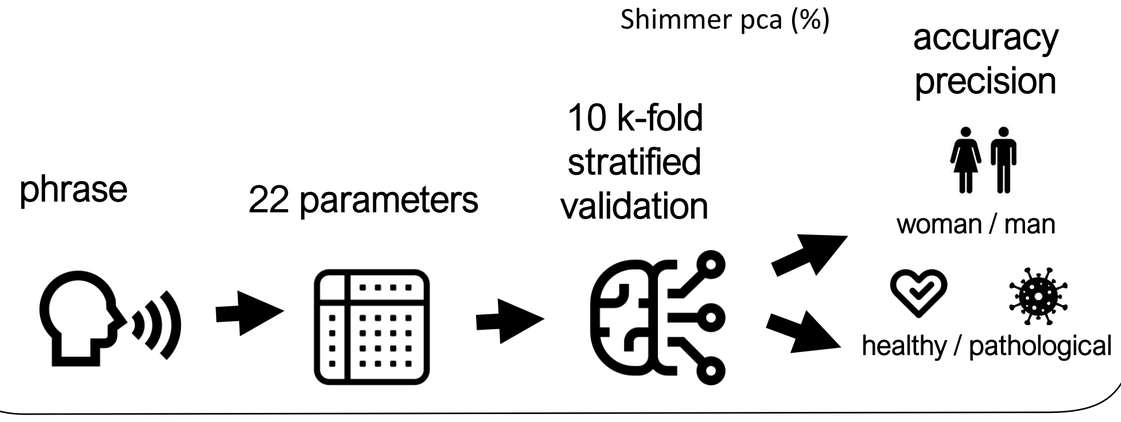
848 woman, 1101 man
619 healthy, 1330 pathological

Signal : 16-bit a 44,1kHz resolution

Feature extraction:

TEMPORAL AND ACOUSTIC ANALYSIS FEATURES
Duration (sec)
F0 mean (Hz)
F0 stdev (Hz)
F1 mean (Hz)
F2 mean (Hz)
F3 mean (Hz)
F4 mean (Hz)

PERTURBATION
HNR mean (dB)
HNR stdev (dB)
Jitter local (%)
Jitter local_abs (s)
Jitter rap (%)
Jitter ppq5 (%)
Jitter ddp (%)
Jitter pca (%)
Shimmer local (%)
Shimmer local dB (dB)
Shimmer apq3 (%)
Shimmer apq5 (%)
Shimmer apq11 (%)
Shimmer dda (%)
Shimmer pca (%)



Results

SUPPORT VECTOR MACHINE

woman/man

	W	M
W	767	81
M	83	1018

W M
Predicted

Accuracy: 92%
Precision: 92%

healthy/pathological

	H	P
H	350	269
P	276	1054

H P
Predicted

Accuracy: 72%
Precision: 79%

LOGISTIC REGRESSION

woman/man

	W	M
W	785	63
M	63	1038

W M
Predicted

Accuracy: 94%
Precision: 94%

healthy/pathological

	H	P
H	291	328
P	171	1159

H P
Predicted

Accuracy: 74%
Precision: 87%

NAIVE BAYES

woman/man

	W	M
W	727	121
M	115	986

W M
Predicted

Accuracy: 88%
Precision: 90%

healthy/pathological

	H	P
H	509	110
P	756	574

H P
Predicted

Accuracy: 55%
Precision: 43%

RANDOM FOREST

woman/man

	W	M
W	779	69
M	64	1037

W M
Predicted

Accuracy: 93%
Precision: 94%

healthy/pathological

	H	P
H	283	336
P	146	1184

H P
Predicted

Accuracy: 75%
Precision: 89%

ADA BOOST

woman/man

	W	M
W	769	79
M	70	1031

W M
Predicted

Accuracy: 92%
Precision: 94%

healthy/pathological

	H	P
H	296	323
P	216	1114

H P
Predicted

Accuracy: 72%
Precision: 84%

MULTILAYER PERCEPTRON

woman/man

	W	M
W	796	52
M	41	1060

W M
Predicted

Accuracy: 95%
Precision: 96%

healthy/pathological

	H	P
H	356	263
P	157	1173

H P
Predicted

Accuracy: 78%
Precision: 88%

Conclusions and further work

Whereas the discrimination between men and women by voice using artificial intelligence techniques obtains quite high results in terms of accuracy and precision. The accuracy and precision in discriminating between normal and pathological voices can be improved. In both cases, the artificial intelligence algorithm that obtains the best results is the **Multilayer Perceptron**.

Future work will consider other types of features as well as signal pre-processing techniques and other artificial intelligence models.

References

Saarbruecken Voice Database [Internet]. [cited 2023 Apr 26]. Available from: https://stimmdatenbank.coli.uni-saarland.de/help_en.php4

Parselmouth – Praat in Python, the Pythonic way — Parselmouth 0.4.3 documentation [Internet]. [cited 2023 Apr 27]. Available from: https://parselmouth.readthedocs.io/en/stable/

Tena A, Claria F, Solsona F, Meister E, Povedano M. Detection of Bulbar Involvement in Patients With Amyotrophic Lateral Sclerosis by Machine Learning Voice Analysis: Diagnostic Decision Support Development Study. JMIR Med informatics [Internet]. 2021 Mar 1 [cited 2023 Apr 26];9(3). Available from: https://pubmed.ncbi.nlm.nih.gov/33688838/

Yolanda Sabuco García

TRANSFORM4EUROPE DOCTORAL SYMPOSIUM 2023

Escuela Politécnica Superior
Universitat d'Alacant
Universidad de Alicante