

Introduction

- PARKINSONISM = Parkinson's Disease (PD) and Atypical Parkinsonian Syndromes (APS)
- Progressive Supranuclear Palsy (PSP) and Multiple System Atrophy (MSA) belong to APS
- In early disease stage, most of PSP, MSA and PD patients develop similar symptoms
- → early differential diagnosis is a very challenging task and essential in assessing treatment/care, understanding the underlying pathophysiology and for the development of new therapies
- Speech disorder is frequently an early and prominent clinical feature of PD and APS

→ speech can be used as an objective marker in early differential diagnosis

Problem setting

- Most of (speech) research deal with discrimination between PD and healthy controls. The symptoms similarity with APS is not considered.
- → such research has a limited clinical impact. Differential diagnosis between PD and APS or within APS should be addressed first.
- We focus here discrimination between MSA and PSP (within APS)
- MSA and PSP are rare diseases → few patients can be studied
- → a small-data machine learning problem

How to build robust classifiers in differential diagnosis between MSA and PSP ?

Objective

- Dimension reduction is required to bypass the curse of dimensionality
- Typically only a 1d feature space may provide acceptable statistics
- → We need to design scalar variables for classification

We investigate standard linear and generalized linear classification models

Dataset

- From 2011 to 2014, 12 consecutive Czech patients with the clinical diagnosis of probable PSP (10 men, 2 women) and 13 patients with the diagnosis of probable MSA (6 men, 7 women) were recruited for the study.
- The diagnosis of PSP was established by the NINDS-PSP clinical diagnosis criteria
- The diagnosis of MSA was established according to consensus diagnostic criteria for MSA
- Speech recordings were performed in a quiet room with a low ambient noise level using a head-mounted condenser microphone
- We use speech data consisting in sustained phonations of the vowel /a/, fast /pa-/ta-/ka/ syllable repetitions and a monologue

Acoustic features

To allow easy future comparisons or reproduction, we consider a subset of 13 features (from the set used in a pioneer study [1]) that can be computed with existing and established scripts

- **Voicing features (Vf)**, on sustained vowels: Jitter, Shimmer, HNR, DUV, F0-SD, Vocal tremor
- **Articulation features (Af)**, on syllable repetition: Slow, Rapid and Irregular AMR (Alternating Motion Rates)
- **Prosodic features (Pf)**, on monologue: Monopitch, Number of pauses, Percentage of pause time (PPT), Intraword pause ratio

Classification models

- Principle Component Analysis (PCA) for unsupervised linear clustering
- Fisher Discriminant Analysis (FDA) for supervised linear classification
- Logistic Regression for (projected) feature normalization

Results

In all experiments, we use Leave-One-Speaker-Out (LOSO) training and a linear SVM with $C = 1$

PCA:

- Vf and Monopitch convey most of (first order) data variability
- PPT convey neutral information (can be discarded from Pf)
- Vf := Vf+Monopitch are orthogonal to the other features APf = Af+Pf
- → This suggests to apply FDA on Vf and APf separately

FDA:

- x-axis = FDA on Vf ; y-axis = FDA on APf (using all data)
- Black line = linear SVM frontier
- → a good separation between classes
- However, LOSO yields only 72% accuracy

FDA+LR:

- x(y)-axis = FDA followed by LR on Vf (APf) (using all data)
- Black line = linear SVM frontier
- → a much better separation between classes
- LOSO yields 80% accuracy

Variables factor map (PCA)

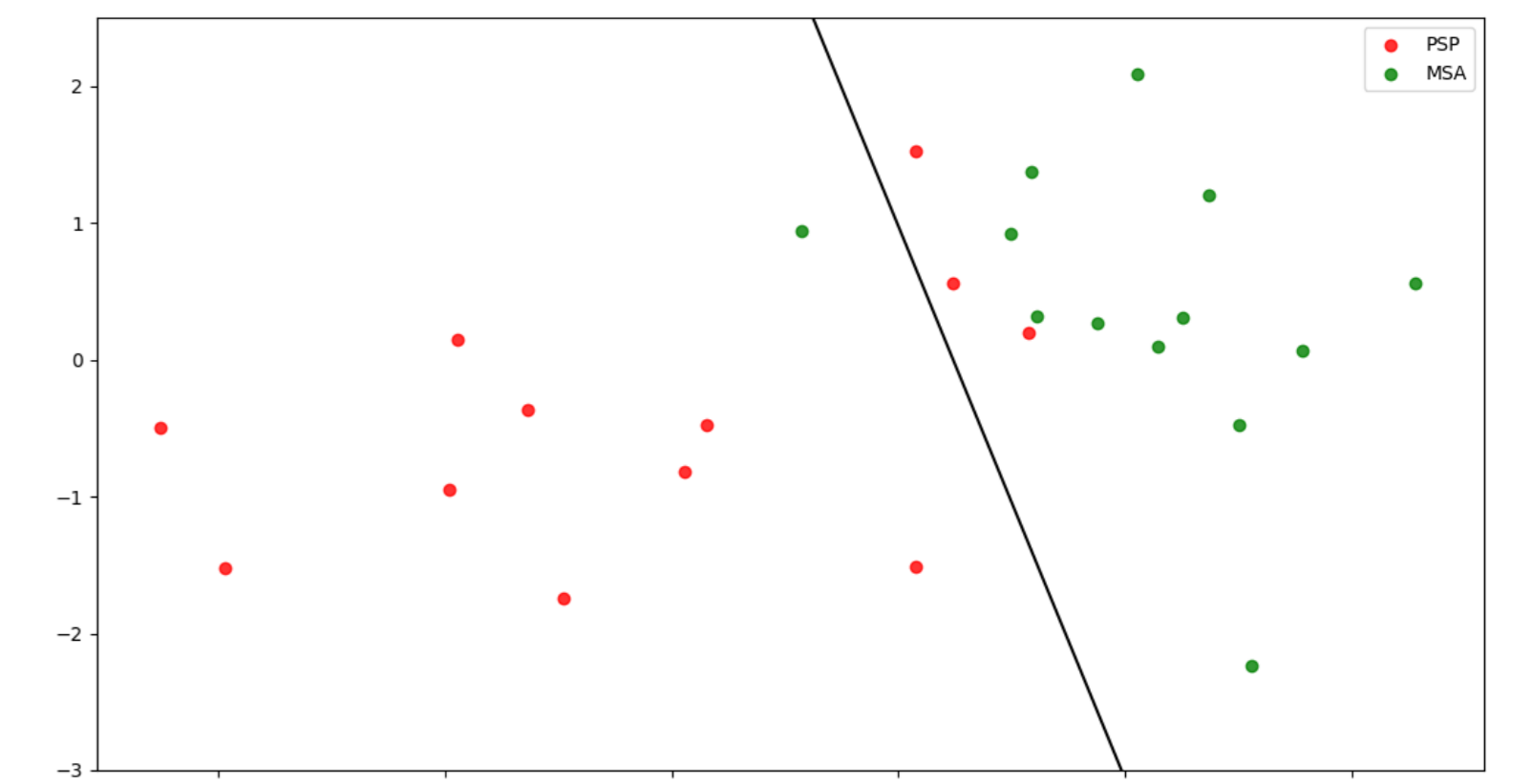
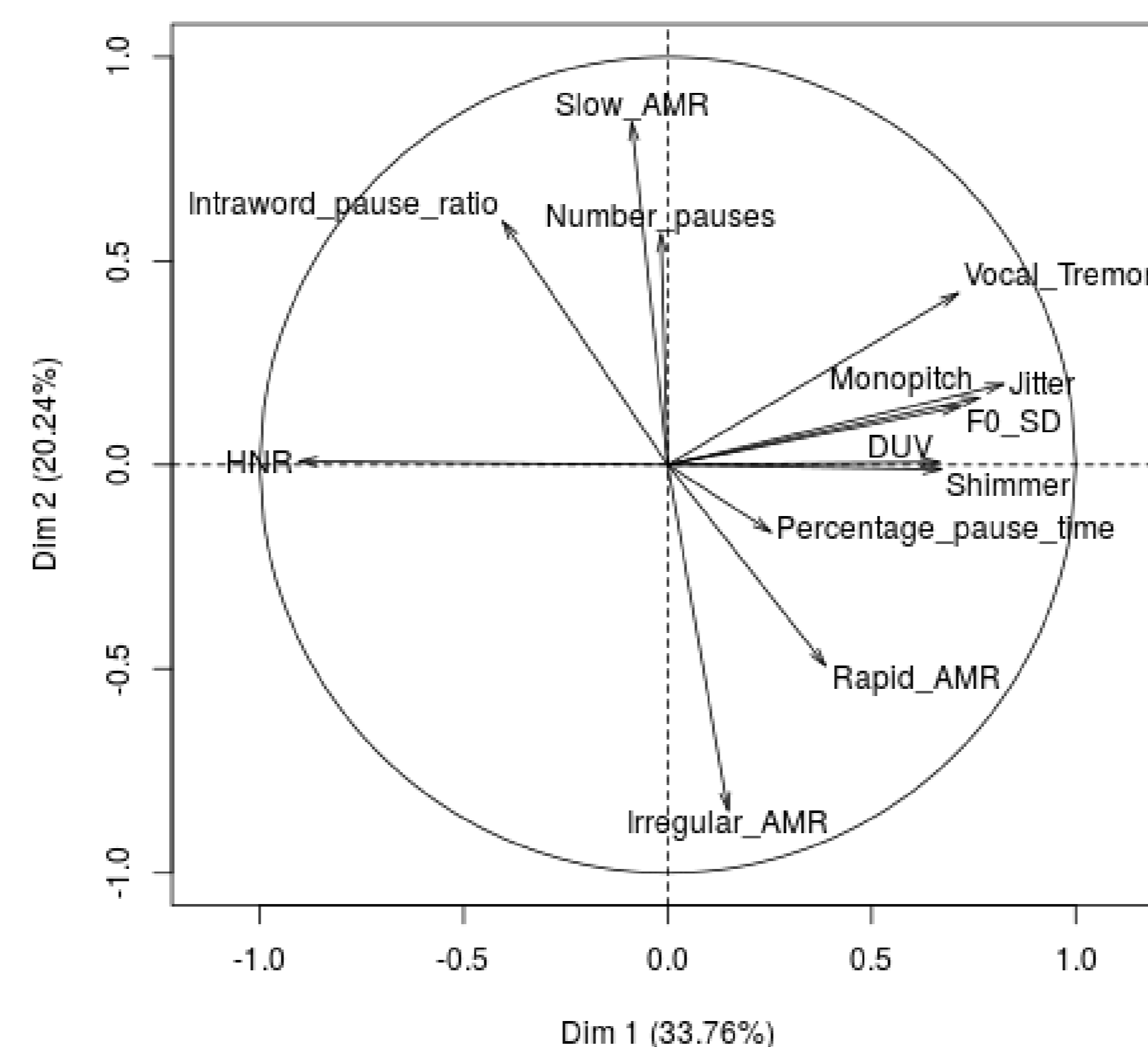


FIGURE 1: 2-dimensional FDA on Vf (x-axis) and APf (y-axis)

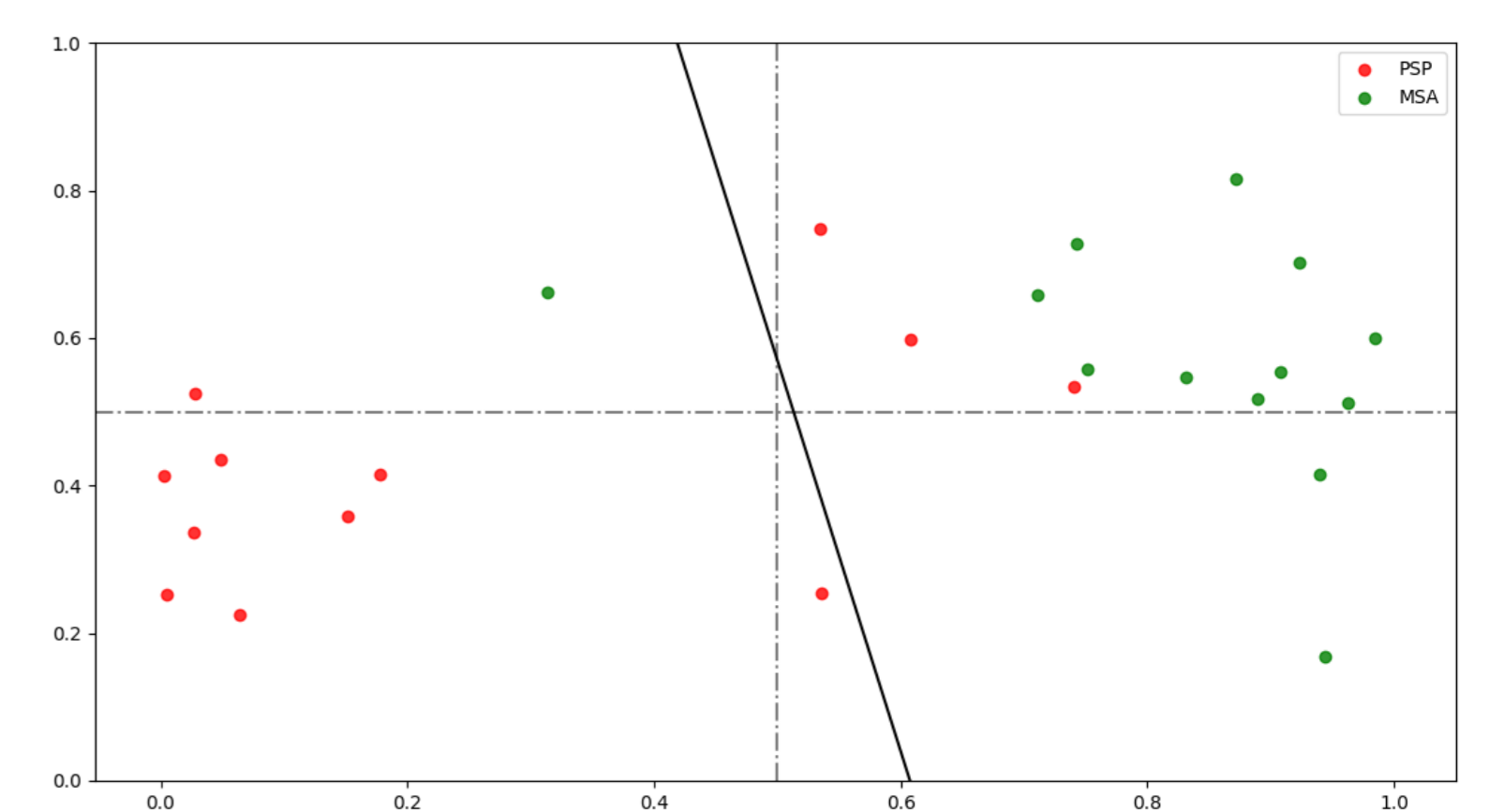


FIGURE 2: 2-dimensional FDA+LR on Vf (x-axis) and APf (y-axis)

Conclusion

- Differential diagnosis in Parkinsonism is a small-data learning problem
 - We investigated (log)linear dimension reduction models
 - FDA and LR can lead to robust classifiers
 - Important advantage: easy interpretation
 - However, no final conclusion can be drawn at this stage
 - (in)validation by additional data and studies is required
- Ongoing research : a large pilot study involving top French university hospitals in the field of Parkinsonism

[1] J. Ruzs, C. Bonnet, J. Klempir, T. Tykalova, E. Baborov, M. Novotny, A. Rulseh, and E. Ruzicka. "Speech disorders reflect differing pathophysiology in parkinsons disease, progressive supranuclear palsy and multiple system atrophy". *Journal of neurology*, vol. 262, pp. 992-1001, 2015.