Gait Assessment of Patients with Parkinson's Disease using Inertial Sensors and Non-Linear Dynamics Features

Paula Andrea Pérez Toro
BSc. student in Electronics Engineering

Advisor: **Prof. Juan Rafael Orozco Arroyave Ph.D.**

Co-Advisor: **MSc. Juan Camilo Vazquez Correa**

GITA research group, University of Antioquia.

paula.perezt@udea.edu.co

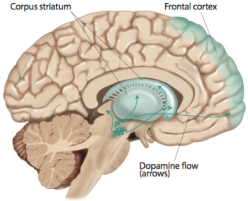
October 29, 2018

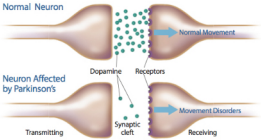
Outline

Introduction

Context: Parkinson's Disease

- Second neuro-degenerative disorder worldwide.
- 6.000.000 Parkinson's patients around the world. 220.000 are from Colombia.
- Neurologists evaluated PD according to MDS-UPDRS-III scale (Goetz et al. 2008).

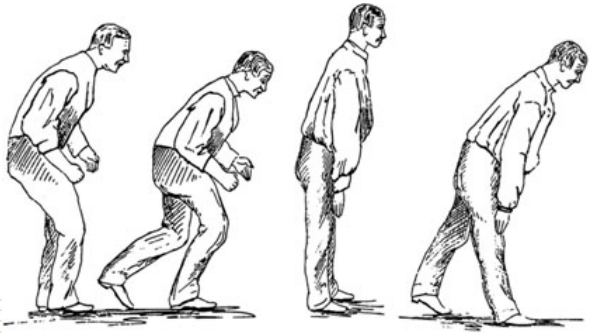




Context: Parkinson's Disease

Motor symptoms

- Resting tremor.
- Rigidity.
- Postural instability.
- Bradykinesia.
- Freezing gait.



Hypothesis and objectives

Hypothesis

Gait signals collected with inertial sensors help in the assessment of the neurological state of patients with PD in different stages of the disease (low, intermediate, and severe).

Hypothesis and objectives

Objectives

General Objective To develop a methodology based on gait analysis and pattern recognition techniques, to perform the automatic classification and evaluation of the neurological state of PD patients according to the MDS-UPDRS-III scale **Goetz2008**

Hypothesis and objectives

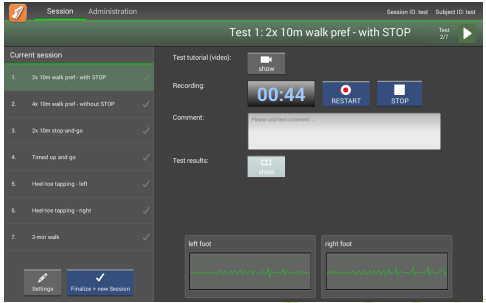
Objectives

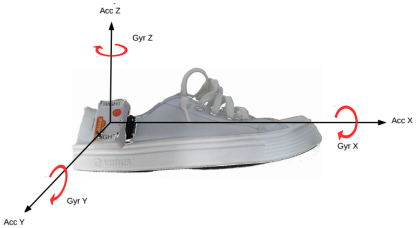
Specific Objective

1. To model several gait tasks performed by PD and HC subjects using different non-linear dynamics features and probabilistic representations of Poincaré maps.
2. To analyze the suitability of different classification and regression methods to model the neurological state of Parkinson's disease patients.
3. To evaluate the developed methodology with several performance metrics.

Gait adquisition and database

Gait Acquisition





Gait signals were captured with the eGaIT system[1]

[1] Embedded Gait analysis using Intelligent Technology, `http://www.egait.de/`

General information about the gait data.

Table: General information of the subjects. **PD** patients: Parkinson's disease patients. **HC**: healthy controls. μ: mean. σ: standard deviation. **T**: disease duration.

	PD patients		YHC subjects		EHC subjects	
	male	female	male	female	male	female
Number of subjects	17	28	26	18	23	22
Age ($\mu \pm \sigma$)	65 $\pm$ 10.3	58.9 $\pm$ 11.0	25.3 $\pm$ 4.8	22.8 $\pm$ 3.0	66.3 $\pm$ 11.5	59.0 $\pm$ 9.8
Range of age	41-82	29-75	21-42	19-32	49-84	50-74
T ($\mu \pm \sigma$)	9 $\pm$ 4.6	12.6 $\pm$ 12.2				
Range of duration of the disease	2-15	0-44				
MDS-UPDRS-III ($\mu \pm \sigma$)	37.6 $\pm$ 21.0	33 $\pm$ 20.3				
Range of MDS-UPDRS-III	8-82	9-106				

PD patients: Parkinson's disease patients.
HC: healthy controls (Elderly and Young)

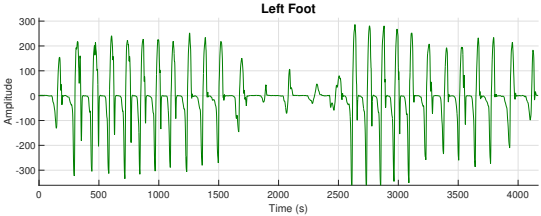
Database and tasks

We considered two gait tasks :

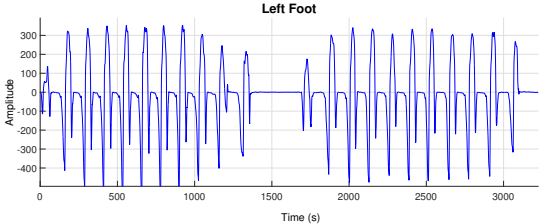
- 4x10m: this consist of walk in a straight line 10 meters and turned around the right side returning back twice.
- 2x10m: this consist of walk in a straight line 10 meters and turned around the right side returning back with a short pause.

Time Series

Female PD patient.
Age:*52.*
MDS-UPDRS=*49*



Female Healthy Young Control.
Age:*23*



Gyroscope Z

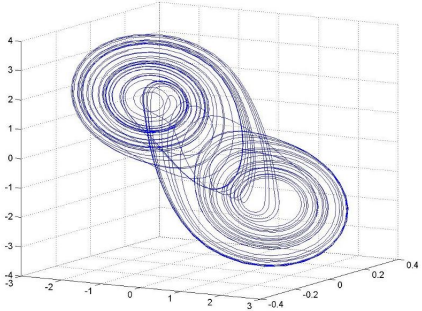
Feature Extraction

Non-linear Dynamics

Gait signals are not linear. This kind of signal shows a non-stationary behaviour.

We focus on non-linear Dynamics systems to describe patterns of gait complexity in patients with Parkinson's disease.

Non-linear Dynamics: Attractors (Phase Space)



Chua's Attractor

- In order to analyze the non-linear properties of the gait signals, the time series has to be projected into a high dimensional space, known as attractor (Taylor 2005).

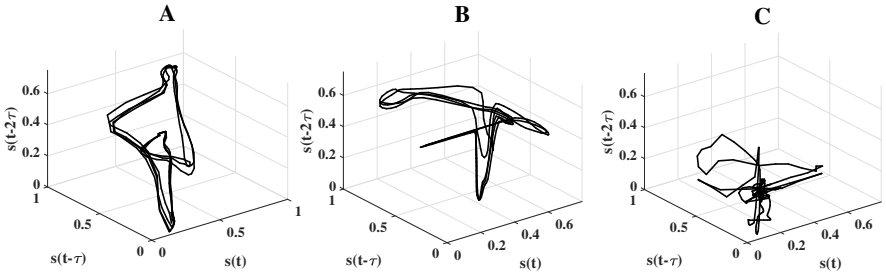
Non-linear Dynamics: Attractors (Phase Space)

- In order to analyze the non-linear properties of the gait signals, the time series has to be projected into a high dimensional space, known as attractor (Taylor 2005).
- From a single time series $\boldsymbol{S_t}$, a phase space can be constructed as follows:

$$\boldsymbol{S_t} = \left\{ s_t, s_{t+\tau}, \ldots s_{t+(m-1)\tau} \right\} \tag{1}$$

τ:delay-time.
m:embedding dimension, a point in the reconstructed phase space.

Non-linear Dynamics: Attractors (Phase Space)



(A) Female YHC, age=23. (B) Female EHC, age=52. (C) Female PD patient, age=52, MDS-UPDRS=49.

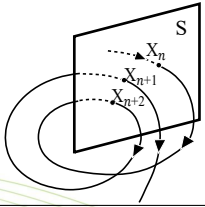
Non-linear Dynamics: Measures

Ten measures were performed. These measures are related with:

- Entropy.
- Space occupied by the attractor.
- Stability.
- Periodicity.
- Large-range dependency and trends.
- Repetitiveness patterns.

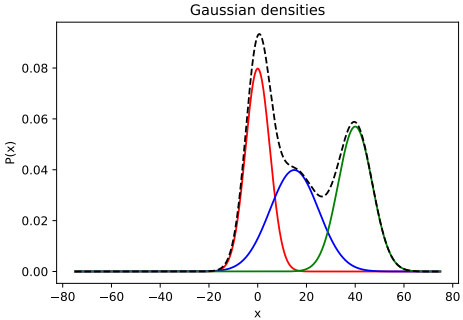
Poincaré Section

- The Poincaré sections also can be used to assess the NLD properties of the signals
- This application takes each point of this section at the first point at which the orbit containing it returns to it.



A clustering algorithm is performed to model the Poincaré Section in a probabilistic way.

Gaussian Mixture Model (GMM)



- Soft version of K-Means: EM algorithm for GMM.
- GMM searchs a mixed of gaussian probability distributions that best model any dataset.

Gaussian Mixture Model (GMM)

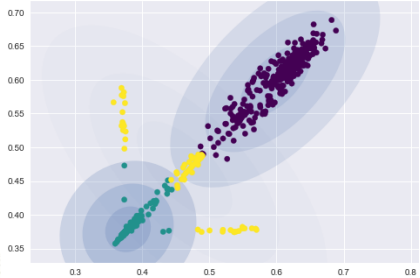
The goal is to estimate μ_k (means), Σ_k (co-variances) and ω_k (weight) to the likelihood L maximization:

$$L(\boldsymbol{X}|\mu_k, \Sigma_k) = \prod_{t=1}^{n} \sum_{k=1}^{K} \omega_k P_k(\boldsymbol{x_t}|\mu_k, \Sigma_k) \tag{2}$$

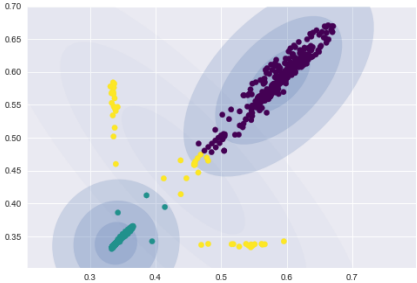
where K is the clusters number, n is the Poincaré dimensions number in $\mathbf{X}$ and P_k the probability density.

Gaussian Mixture Model (GMM)

Female PD patient.
Age:*52.*
MDS-UPDRS=*49*



Female Healthy Young Control.
Age:*23*



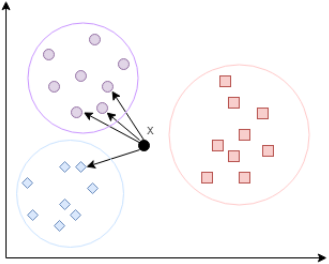
Gyroscope Z

Classification

Classification: K-Nearest-Neighbors (KNN)

- KNN (Bishop 2006) uses a majority vote among the **k**, defining competencies as a distance measure d

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2} \tag{3}$$

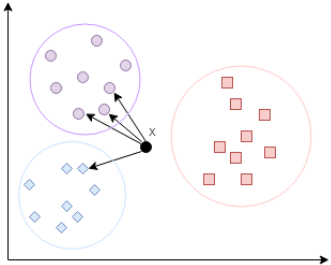


New input data in accordance with their distances

Classification: K-Nearest-Neighbors (KNN)

- For the input x, the class with the highest probability is assigned.

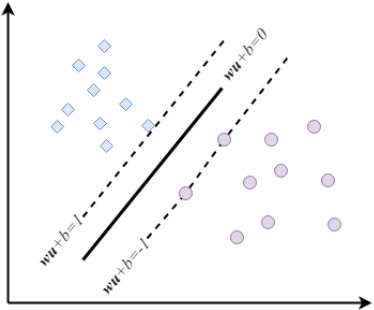
$$P(Y = j|X = x) = \frac{1}{k}\sum I(Y^{(i)} = j) \tag{4}$$



New input data in accordance with their distances

Classification: Support Vector Machine (SVM)

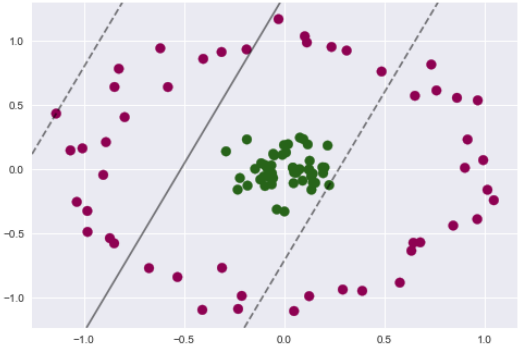
- SVM (Bishop 2006) outputs a class identity for every new vector **u**, by modeling best fitting hyperplane.



SVM Best fitting hyperplane

Classification: Support Vector Machine (SVM)

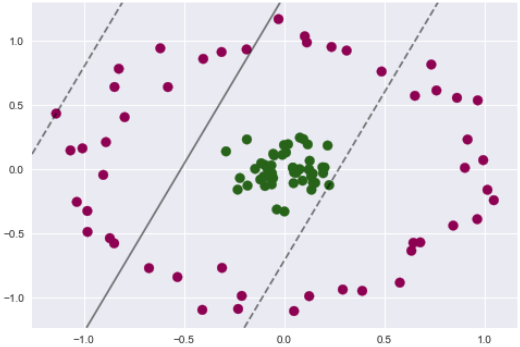
Linear Kernel



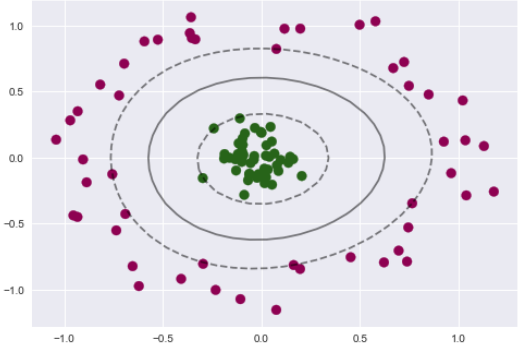
Classification: Support Vector Machine (SVM)

- A Gaussian kernel transforms the feature space into one linearly separable.

Lineal Kernel

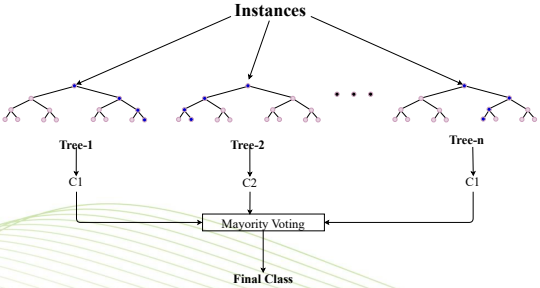


Gaussian Kernel



Classification: Random Forest (RF)

- Random Forest (RF) consists of a classification tree set.
- Each one contributes with one vote to assign a class.



Architecture of the random forest model

Regression

Support Vector Regression (SVR)

- Let us to predict the value of the scale ($\widehat{y}$) using a function of losses $L(y, \widehat{y})$.This function is calculated with the follow equation:

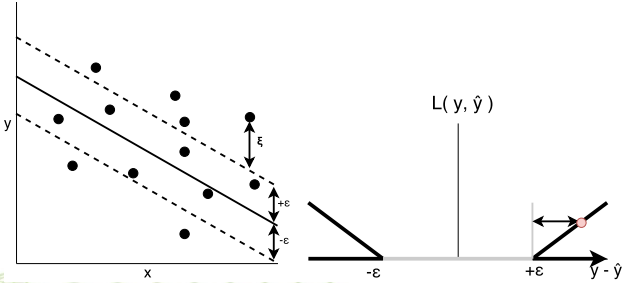
$$L(y, \widehat{y})) = \begin{cases} 0 & \textit{if } |y - \widehat{y}| \leq \varepsilon \\ |y - \widehat{y}| - \varepsilon & \textit{otherwise.} \end{cases} \tag{5}$$

- The predicted values $\widehat{y}$ are estimated using the equation 6, where ω_j sets the weight of each support vector, and b is the independent term.

$$\widehat{y} = \sum_{j=1}^{m} \omega_j g_j(\boldsymbol{x}) + b \tag{6}$$

Neurological State Prediction: SVR

Support Vector Regression (SVR)



Loss Function

- A linear kernel transforms the feature space into one linearly separable.

Experiments and Results

Results: Biclass Classification

Five folds are chosen to perform the classification. These folds were balanced by gender and shoe type.

Table: Classification Results: Fusion Left

Features/ Classificator	NLD Accuracy ($\mu \pm \sigma$)	Poincaré-GMM Accuracy ($\mu \pm \sigma$)	NLD+Poincaré-GMM Accuracy ($\mu \pm \sigma$)
KNN	80.0%±8.4	57.8%±9.0	83.3%±6.0
SVM	83.3%±6.8	57.8%±4.0	86.8%±8.3
RF	83.3%±8.8	83.7%±2.7	87.7%±6.4

Table: Confusion Matrix: Fusion Left
Random Forest
NLD+Poincaré-GMM

Class	EHC	PD
EHC	40	5
PD	7	38

Results: Biclass Classification

Five folds are chosen to perform the classification. These folds were balanced by gender and shoe type.

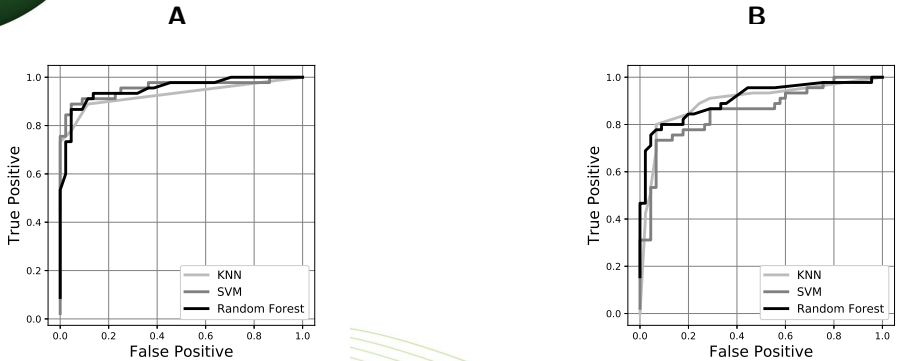
Table: Classification Results: Fusion Left

Features/ Classificator	NLD Accuracy ($\mu \pm \sigma$)	Poincaré-GMM Accuracy ($\mu \pm \sigma$)	NLD+Poincaré-GMM Accuracy ($\mu \pm \sigma$)
KNN	80.0%±8.4	57.8%±9.0	83.3%±6.0
SVM	83.3%±6.8	57.8%±4.0	86.8%±8.3
RF	**83.3%±8.8**	**83.7%±2.7**	**87.8%±6.4**

Table: Confusion Matrix: Fusion Left Random Forest NLD+Poincaré-GMM

Class	EHC	PD
EHC	40	5
PD	7	38

Results: Biclass Classification



ROC curve graphics of the best NLD Features results. A) PD vs YHC. B) PD vs EHC. In both cases the fusion of features from both feet and both tasks are considered.

Results: Neurological state prediction

Five folds are chosen to perform the regression. These folds were balanced by MDS-UPDRS-III scale.

Table: Regression. **MAE:** mean absolute error. ρ: Spearman's Correlation.

Features /UPDRS	**NLD (ρ/MAE)**	**Poincaré-GMM (ρ/MAE)**	**NLD+Poincaré-GMM (ρ/MAE)**
General (Left Fusion)	0.65/12.95	0.26/18.00	0.05/15.46
Lower Limbs (Both 4x10)	0.31/8.26	-0.09/8.46	0.07/7.93

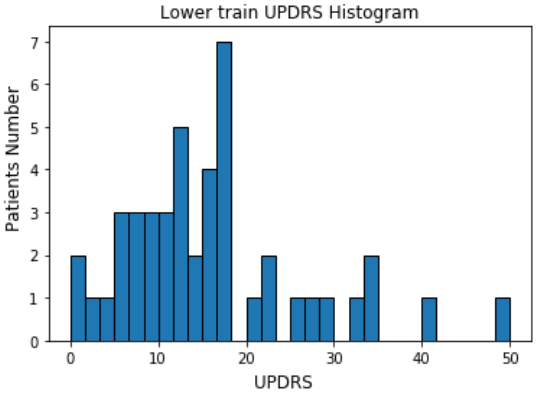
Results: Neurological state prediction

Five folds are chosen to perform the regression. These folds were balanced by MDS-UPDRS-III scale.

Table: Regression. **MAE:** mean absolute error. ρ: Spearman's Correlation.

Features /UPDRS	NLD (ρ/MAE)	Poincaré-GMM (ρ/MAE)	NLD+Poincaré-GMM (ρ/MAE)
General (Left Fusion)	**0.65/12.95**	0.26/18.00	0.05/15.46
Lower Limbs (Both 4x10)	**0.31/8.26**	-0.09/8.46	0.07/7.93

Results: Multiclass Classification



EHC and three UPDRS categories were chosen:

- Elderly Healthy Controls (EHC): elderly control (45+ years old).
- Category 1: UPDRS for lower limbs among 0-10.
- Category 2: UPDRS for lower limbs among 11-17.
- Category 3: UPDRS for lower limbs among 18+.

Results: Multiclass Classification

Five folds are chosen to perform the classification. These folds were balanced by gender and UPDRS category.

Table: Classification Results: Fusion Left

Features/ Classificator	NLD Accuracy ($\mu \pm \sigma$)	Poincaré-GMM Accuracy ($\mu \pm \sigma$)	NLD+Poincaré-GMM Accuracy ($\mu \pm \sigma$)
KNN	62.2%±9.5	52.9%±7.1	57.3%±4.1
SVM	62.2%±13.3	51.0%±3.7	65.2%±8.1
RF	61.1%±12.1	56.0%±3.6	61.8%±5.1

Table: Confusion Matrix: Fusion Left Support Vector Machine NLD+Poincaré-GMM

Class	EHC	UPDRS Category 1	UPDRS Category 2	UPDRS Category 2
EHC	41	1	3	0
UPDRS Category 1	5	4	3	2
UPDRS Category 2	5	0	11	0
UPDRS Category 3	6	4	3	2

Results: Multiclass Classification

Five folds are chosen to perform the classification. These folds were balanced by gender and UPDRS category.

Table: Classification Results: Fusion Left

Features/ Classificator	NLD Accuracy ($\mu \pm \sigma$)	Poincaré-GMM Accuracy ($\mu \pm \sigma$)	NLD+Poincaré-GMM Accuracy ($\mu \pm \sigma$)
KNN	62.2%±9.5	52.9%±7.1	57.3%±4.1
SVM	62.2%±13.3	51.0%±3.7	65.2%±8.1
RF	61.1%±12.1	56.0%±3.6	61.8%±5.1

Table: Confusion Matrix: Fusion Left Support Vector Machine NLD+Poincaré-GMM

Class	EHC	UPDRS Category 1	UPDRS Category 2	UPDRS Category 2
EHC	41	1	3	0
UPDRS Category 1	5	4	3	2
UPDRS Category 2	5	0	11	0
UPDRS Category 3	6	4	3	2

Conclusions

Conclusions

- The fusion of several features and tasks is more effective in the classification process, i.e., both tasks provide complementary information to discriminate between PD patients and EHC subjects.

Conclusions

- The fusion of several features and tasks is more effective in the classification process, i.e., both tasks provide complementary information to discriminate between PD patients and EHC subjects.
- Results indicate the presence of the cross laterality effect(Sadeghi et al. 2000).

Conclusions

- The fusion of several features and tasks is more effective in the classification process, i.e., both tasks provide complementary information to discriminate between PD patients and EHC subjects.
- Results indicate the presence of the cross laterality effect(Sadeghi et al. 2000).
- The fusion among Poincaré-GMM and NLD features, shows us a reduction of standard deviation and increment the accuracy, indicating more stability and efficiency.

Conclusions

- The fusion of several features and tasks is more effective in the classification process, i.e., both tasks provide complementary information to discriminate between PD patients and EHC subjects.
- Results indicate the presence of the cross laterality effect(Sadeghi et al. 2000).
- The fusion among Poincaré-GMM and NLD features, shows us a reduction of standard deviation and increment the accuracy, indicating more stability and efficiency.
- The reason because MDS-UPDRS-III has higher results than with the subscore of lower limbs is the range of the total UPDRS is larger and some parameters were a little bit affected by this.

Future Work

- Analyze movement signals captured from smarthphones.
- Combine kinematics features from gait signals.
- The proposed approach can be extended to other applications. For instance the discrimination between PD and other neurological disorders with similar symptoms, such as Huntington's disease, amyotrophic lateral sclerosis, or essential tremor.

References

Bishop, Christopher M. *Pattern recognition and machine learning*. springer, 2006.

Goetz, Christopher G et al. "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results". In: *Movement disorders* 23.15 (2008), pp. 2129–2170.

Sadeghi, Heydar et al. "Symmetry and limb dominance in able-bodied gait: a review". In: *Gait & posture* 12.1 (2000), pp. 34–45.

Taylor, Robert LV. "Attractors: nonstrange to chaotic". In: *Society for Industrial and Applied Mathematics, Undergraduate Research Online* (2005), pp. 72–80.