E3S Web of Conferences **351**, 01044 (2022) https://doi.org/10.1051/e3sconf/202235101044
ICIES'22

Feature Selection of Arabic Online Handwriting Using Recursive Feature Elimination for Parkinson's Disease Diagnosis

Meryem Amakrane[1], *Ghizlane Khaissidi*[1], *Mostafa Mrabti*[1], *Alae Ammour* [1], *Belahsen Faouzi*[2] *and Ghita Aboulem*[2]

[1] Laboratory LIPI ENS, USMBA Fez, Morocco
[2] Laboratory ERMSC, FMPF, CHU Hassan II Fez, Morocco

Abstract. Parkinson's disease (PD) is one of the most common neurodegenerative diseases affecting a large population worldwide. Parkinson's disease is characterized by rigidity, slowness of movement and tremors at rest, these syndromes are frequently manifested in the deterioration of handwriting. The aim of this article is to perform online Arabic handwriting analysis for two types of tasks, TASK 1: copying arabic imposed text and TASK 2: writing arabic desired text. A novel method of handwriting selection features is proposed to obtain the relevant features to efficiently identify subjects with PD, based on Recursive Feature Elimination with Cross-Validation (RFECV), three different RFE estimators were compared: Support Vector Machine, Decision Trees and Random Forest, the selected features have been fed to the same classifiers above to determine the best classifier for predicting Parkinson's disease. Result: An accuracy of 94.4% was obtained using SVM with Linear kernel, based on 55 features selected using RFE-SVM(Linear) for TASK 1, for TASK 2 an accuracy of 93.7% was obtained using SVM with RBF kernel, based only in 7 features selected using RFE-SVM(Linear). For all the classifiers used, this technique experimentally demonstrates an increase in performance metrics.

1 Introduction

Parkinson's disease (PD) is a neurodegenerative disease resulting primarily from the destruction of dopaminergic neurons in the substantia nigra. Clinical Characteristics of PD include resting tremor, rigidity, bradykinesia (slowness of movement), akinesia (absence of normal unconscious movements), micrographia (decreased size) and others [1]. PD is often not diagnosed until after the first motor symptoms appear when more than half of the DA nigrostriatal neurons are already lost [2]. However, early diagnosis of this pathology is crucial in order to limit crisis situations and ensure a better quality of life for patients over a longer period of time. Handwriting can be considered as one of the keys to the assessment of neurodegenerative diseases as it's a complex activity involving motor and cognitive components.
In this perspective, online handwriting analysis has become a new research approach in the healthy field and more particularly in the early diagnosis of neurodegenerative diseases. Which have led to numerous agreements between universities and hospitals to ensure the acquisition of online handwriting for neurodegenerative patients.
Related to our work, previous studies are done on Arabic language by Ammour et al, in [3] the study was carried out on the basis of 28 Parkinson's disease patients and 28 healthy controls, participants were of similar age and intellectual level. Focusing on the first task of copying an Arabic text, a new method was proposed based on a Supervised Learning Approach to characterize the online handwriting of PD and HC populations according to quantitative and qualitative parameters. The results obtained demonstrate that the fusion of these two aspects makes it possible to better discriminate parkinsonian patients from HCs.
In [4], the study confirms that fatigue occurs when writing in PD patients, a new proposal based on fragmentation of handwritten text into lines, comparing the full dynamics of new temporal and spectral features analyzed between patients with Parkinson's disease and healthy controls: Based on decision tree classifier an accuracy of 92.86% was achieved in the last line.
Indeed, our study is conducted using the same database of Ammour et al as part of the project called ENEMAR (Etude Neurologique de l'Ecriture des MARocains), the acquisition was carried out within the neurological department of the UHC Hassan II of Fez, under the ethical agreement (N° 03/15; July 10, 2015; Fez, Morocco) for the biomedical research of the FMPF.
In the upcoming section, the online handwriting database is introduced. Section 3 presents the methods of feature extraction and the classification used, the results are shown in section 4. The conclusion is provided in the final section.

2 Parkinson's disease handwriting dataset

Data acquisition was carried out in cooperation with neurological department of the Hospital Hassan II of Fez and the Laboratory of Computer Sciences and Interdisciplinary Physics (LIPI). The data collection was done using graphic tablet "WACOM Intuos Pro" with creative stylus "Inking". The tablet has a sample rate of 125 HZ, can sense 2048 pressure levels and record data when the stylus is in the air (up to 1.5cm above the tablet)[5]

2.1 Participants

The dataset used in this study is consists of 26 Parkinson's disease patients and 26 healthy controls. All the participants are Arabic mother tongue speakers, right-handed, and have achieved at a minimum six years of schooling. Neither one of them had a precedent or presence of any other neurodegenerative diseases (other than PD). All PD patients fulfilled the tasks after 1h to 30min from taking the L-DOPA medication. A neuropsychologist estimates the cognitive state of each patient via a complete neuropsychological test, the Unified Parkinson's Disease Rating Scale (UPDRS) [6] has been used to quantify the PD state, and the Mini-Mental State Examination (MMSE) [7] is applied for healthy control. Mean of age, UPDRS, and MMSE are represented in Table 1.

Table 1. Representative data of healthy controls (HC) and Parkinson's disease (PD).

	Mean age	Mean MMSE	Mean UPDRS
PD	54.23	27.11	10.65
HC	53.8	30	-

2.2 Handwriting task

Two types of tasks were analyzed in this work, the first and second exercise of the acquisition protocol [8]: the first task consist of copying a provided Arabic text contain six lines (TASK 1), the second task consist of writing a desired text (TASK 2) contain minimum four lines without being a text memorized a priori. An example of the first and second task for the same participant is presented in Fig. 1 and Fig. 2 respectively.

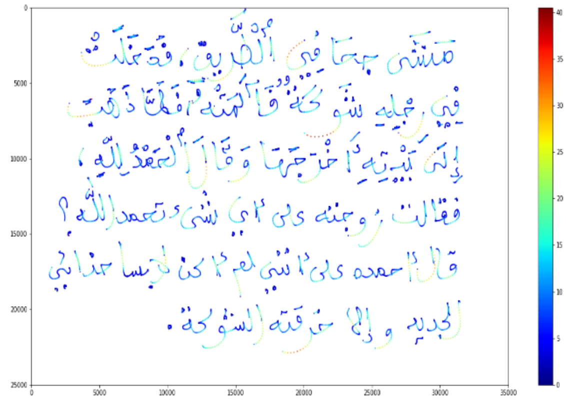


Fig. 1. Handwritten TASK 1 with on surface line recorded with tablet and related to velocity mm/s^2.

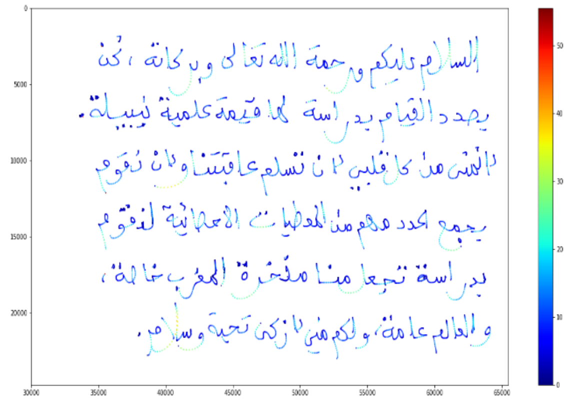


Fig. 2. Handwritten TASK 2 with on surface line recorded with tablet and related to velocity mm/s^2.

2.3 Feature extraction

For each point of the stylite trajectory, it was possible to extract the coordinates [X(n), Y(n)], the pressure (P) and the inclination [Az(n), Al(n)], this real time digital representation of handwriting is called the online Handwriting. These signals are used to calculate new features, the features are classified in the following category:

- Kinematics: velocity, acceleration, jerk, NCV [10], NCA [10], etc.
- Dynamics: pressure, pressure derivative, NCP [10] and relative NCP.
- Spatio-temporal: stroke size, stroke duration, on-surface and in-air time, etc.
- Inclination: azimuth\altitude angle, azimuth\altitude angle derivative.

Additionally, 11 basics statistical functionals (mean, median, mode, standard deviation, min, max, quartile 1, quartile 3, Interquartile Range, skewness, kurtosis) were computed. All features are scales between -1 and 1.

3 Feature selection and Classification

In order to remove the irrelevant features and achieve better accuracy, the following approach is adopted: the

https://doi.org/10.1051/e3sconf/202235101044

first step is to apply the Mann Whitney U statistical test, the second step is to select features based on the Recursive Feature Elimination with Cross-Validation (RFECV). The selected features were fed to three classifiers: Support Vector Machine (SVM), Decision Trees (DT) and Random Forest (RF).

3.1. Mann-Whitney Test

A statistical study of the previous calculated features was carried out in order to retain only the features which represent a statistical significance between Parkinson's patients and healthy controls.

Mann Withney test (MWW) [11] was used as a non-parametric test. All parameters were compared from the two populations and selected with a risk of error of less than 5%. Percentage of parameters retained following the application of the Mann Whitney test is presented in Table 2.

Table 2. Percentage of features retained from Mann Whitney test.

	Kinematic (%)	Dynamics (%)	Spatiotemp (%)	Inclination (%)
TASK 1	82.40	77.27	72.72	51.21
TASK 2	78.70	83.36	54.54	51.21

3.2 Recursive Feature Elimination

Recursive Feature Elimination is used in this study to select features that are more or most relevant in predicting the target variable. Selecting the right features is important to reduces the complexity of the model and improves their performances.

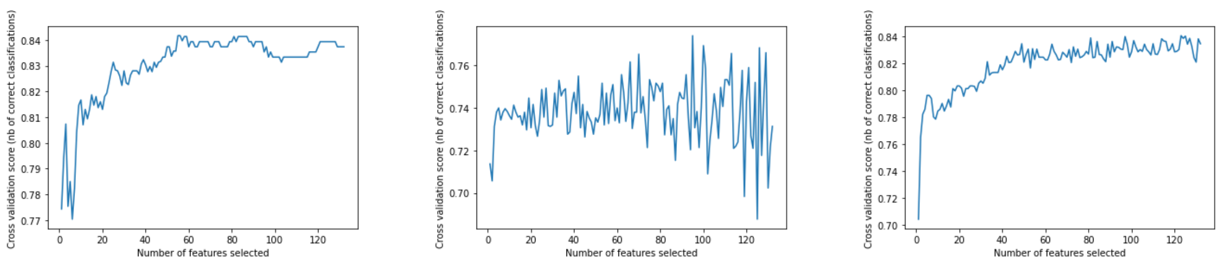
RFE is a wrapper-type feature selection algorithm. The Support Vector Machine with Linear kernel [12], Decision Trees and the Random Forest each one of those algorithms was used as the core of RFE to select the optimal number of features. Using RFE with cross-validation (RFECV), the fitting is accompanied by testing which make possible to automate the number of selected features. The cross-validation splitting strategy was conducted using stratified 10-fold.

The complete results of the selected features are presented in Table 3.

Table 3. Number of selected features with RFECV method.

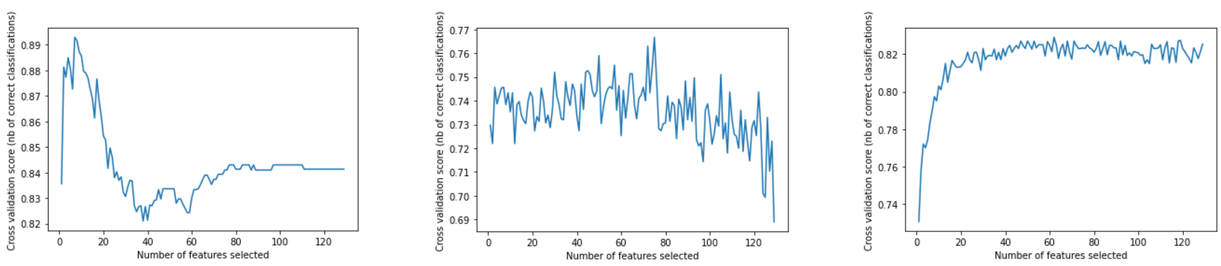
	RFECV-SVM	RFECV-DT	RFECV-RF
TASK1	55	95	123
TASK2	7	75	62

The optimal number of selected features using RFECV may vary from one exercise to another, the TASK 2 has less features selected compared to the TASK 1. This means that a good accuracy can be achieved with less features, this experiment also confirms that increasing features can have a negative impact on prediction as can be seen from Fig. 3 and Fig. 4.



(a) RFECV based on SVM (linear) (b) RFECV based on DT (c) RFECV based on RF

Fig. 3. Classification accuracy based on cross-validation for TASK 1 features.



(a) RFECV based on SVM (linear) (b) RFECV based on DT (c) RFECV based on RF

Fig. 4. Classification accuracy based on cross-validation for TASK 2 features.

3.3 Classification

In order to find the best model for classifying people with PD and HC, the following machine learning techniques were compared in this study: Support Vector Machine with Linear Kernel, Support Vector Machine with RBF Kernel, Decision Trees and Random Forest. The hyperparameters of all mentioned models are briefly described in Table 4. A 10-fold cross-validation was involved to evaluate each model. Which mean that our dataset is divided into 10 non-overlapping folds. Each fold is used once as validation while the remaining nine folds form the training set. The entire procedure has repeated ten times to ensure that each fold has been used as a test set. The performance of the classifiers was measured with the several statistics: F-score, accuracy, recall, and precision.

Table 4. Hyper-parameter specification of all prediction models.

Model	**Description**
SVM	SVM & Kernel = Linear C = 1.0 gamma = 'scale'
SVM	SVM & Kernel = RBF C = 1.0 gamma = 'scale'
DT	criterion = 'gini' max_depth = 4.0
RF	n_estimators = 100 criterion = 'gini'

4 Experimental results

To emphasize the importance of the proposed feature selection using Recursive Feature Elimination, the initial experiments were performed considering all features extracted from the statistical test. Subsequently the same experiments were carried out using the selected feature from RFE. Results demonstrate that the RFE feature selection technique improve the PD classification prediction scores for both writing exercises and best performed with SVM classifier (Linear and RBF kernel). Table 5 and Table 6 can be consulted for details of the results. According to the result found, the performance metrics scores represent a demonstrable improvement using features selected based on RFE-SVM(Linear) compared to RFE-RF and RFE-DT. For TASK 1, SVM (Linear kernel) remained best with prediction accuracy, recall, precision, and F1 scores of 0.94, 0.95, 0.851, and 0.94 respectively followed by SVM (RBF kernel), RF, and DT. For TASK 2, SVM (Linear kernel) and SVM (RBF kernel) represent similar high scores with a less number of features compared to TASK 1 followed by RF and DT. This study was carried out on the Python 3.8 platform based on "scikit-learn" machine learning library.

5 Conclusion

In this work, we proposed a novel handwriting selection feature method based on the elimination of recursive features with cross-validation (RFECV). This method was applied for 2 types of exercises in order to analyze and compare the tasks before and after the application of the RFE algorithm. The results obtained demonstrated the improvement of the classification using RFE; it has also been observed that the task 2 provide a similar level of discrimination compared to task 1 but with a much lower number of features.

Table 5. Classification results applied on TASK 1 using 10-fold cross-validation.

Feature	Classifier	Accuracy (%)	Recall (%)	Precision (%)	F1 (%)
MWW	SVM (RBF) SVM (Linear) RF DT	88.00 84.00 82.80 75.50	90.30 85.70 82.50 72.80	87.50 87.10 85.10 76.70	87.50 84.20 81.50 72.00
RFECV - SVM	SVM (RBF) SVM (Linear) RF DT	88.50 94.40 85.10 79.90	92.50 95.20 85.70 78.70	86.80 95.80 88.40 80.60	88.40 94.50 84.10 77.20
RFECV - DT	SVM (RBF) SVM (Linear) RF DT	88.40 81.80 82.50 73.10	91.80 81.10 81.30 72.70	87.00 85.10 84.10 80.30	88.10 81.10 80.60 73.80
RFECV - RF	SVM (RBF) SVM (Linear) RF DT	88.80 84.30 82.60 79.30	89.50 86.80 81.50 74.30	89.80 86.50 85.40 81.90	88.10 84.60 81.00 75.40

Table 6. Classification results applied on TASK 2 using 10-fold cross-validation.

Feature	Classifier	Accuracy (%)	Recall (%)	Precision (%)	F1 (%)
MWW	SVM (RBF) SVM (Linear) RF DT	86.20 80.10 82.40 68.50	87.70 81.30 81.20 70.80	86.10 83.00 85.30 70.80	85.40 79.40 80.60 67.20
RFECV - SVM	SVM (RBF) SVM (Linear) RF DT	93.70 92.50 87.40 82.00	96.00 96.80 87.50 78.30	92.80 91.00 88.40 87.20	93.70 92.80 86.40 79.80
RFECV - DT	SVM (RBF) SVM (Linear) RF DT	85.10 82.80 82.70 71.80	85.80 86.80 81.30 76.50	85.90 85.00 85.60 74.40	85.10 83.30 81.10 71.40
RFECV - RF	SVM (RBF) SVM (Linear) RF DT	86.20 82.60 82.80 71.00	88.20 85.30 81.20 74.80	86.10 84.00 85.90 73.60	85.80 82.20 81.10 70.50

References

1. Dauer, W., Przedborski, S.: Parkinson's Disease: Mechanisms and Models. Neuron (2003)
2. M. Man, J. H. K., Groenink, L., & Caiazzo, M.: Cell reprogramming approaches in gene-and cell-based therapies for Parkinson's disease. Journal of Controlled Release, 286,114–124 (2018)
3. Ammour, A., Aouraghe, I., Khaissidi, G., Mrabti, M., Aboulem, G., & Belhacen, F: A New Semi-Supervised Approach for Characterizing the Arabic on-line Handwriting of Parkinson's Disease Patients. Computer Methods and Programs in Biomedicine (2019)

4. Aouraghe, I., Alae, A., Ghizlane, K., Mrabti, M., Aboulem, G., & Faouzi, B.: A novel approach combining temporal and spectral features of Arabic online handwriting for Parkinson's disease prediction. Journal of Neuroscience Methods(2020)
5. Graphic tablet data sheet. https://www.wacom.com/fr-ca/products/pen-tablets/wacomintuos-pro
6. The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations. Movement Disorders (2003)
7. Folstein, M. F.: The Mini-Mental State Examination. Archives of General Psychiatry (1983).
8. I. Aouraghe, A. Ammour, G. Aboulem, G. Khaissidi, M. Mrabti, F. Belahsen, M.A. El-Yacoubi, S. Garcia-Salicetti: Automatic analysis of on-line handwriting for the early detection of neurodegenerative pathologies. Iste Openscience – Published by Iste Ltd. London, Uk – Openscience.Fr (2017).
9. Eichhorn, T. E., Gasser, T., Mai, N., Marquardt, C., Arnold, G., Schwarz, J., & Oertel, W. H.: Computational analysis of open loop handwriting movements in Parkinson's disease: A rapid method to detect dopamimetic effects. Movement Disorders (1996)
10. Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., & Faundez-Zanuy, M.: Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. Artificial Intelligence in Medicine (2016)
11. N. Nachar: The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. Tutorials in quantitative methods for psychology (2008)
12. Guyon, I., Weston, J., Barnhill, S. et al.: Gene Selection for Cancer Classification using Support Vector Machines. Machine Learning 46, 389–422 (2002)