Evaluation of NN-based classification methods for Parkinson detection from small datasets

1 Description

Classification of neurological diseases and psychiatric conditions using data different types of tests is a relevant application area of machine learning algorithms [1, 2, 3]. Challenging questions is this type of problemas are: 1) How to find an appropriate representation of the data (feature extraction)?; and 2) How to improve classification for unbalanced datasets and few observations?

This project will use a database of collected handwriting samples, collected using a tablet, of Parkinson Disease (PD) patients and healthy subjects. The PD and control (healthy subjects) handwriting database, originally proposed in [1] consists of 62 PWP (People with Parkinson) and 15 healthy individuals. From all subjects, three types of handwriting recordings (Static Spiral Test (SST), Dynamic Spiral Test (DST) and Stability Test on Certain Point (STCP)) are taken. Also the drawings of spirals belongs to the PWP are included in the dataset as image. Therefore, this dataset could also be used for regression.

The dataset is available from `http://archive.ics.uci.edu/ml/datasets/` under the name of Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet. The original publication [1] provides details of the dataset.

2 Objectives

The goal of the project is the application Neural Networks for: I) Find suitable feature representations for this problem that are very usable for other ML classifiers, OR, II) Implement NN-based classifiers for this problem, OR III) The combination of I and II (e.g., using an RBM to find the features and a Multi-layer Perceptron to classify the problem using the extracted features). In case II), the students are free to decide which feature representation is more appropriate for the data. In case I), they can use any classifier with the NN-based features.

The student should: 1) Propose the representation and/or classifier. 2) Implement the solution of the classification problem. 3) Evaluate and discuss the results of the classifier. 4) Answer to the following questions in the report:

- What class of problems can be solved with the NN? (e.g., supervised vs unsupervised problems)
- What is the network architecture? (e.g., type and number of layers, parameters, connectivity, etc.).
- What is the rationale behind the conception of the NN?
- How is inference implemented? (e.g., How is the information extracted from the network?). Type of prediction or type of inference process.
- What are the learning methods used to learn the network ? Algorithms used for learning the network.

As in other projects, a report should describe the characteristics of the design, implementation, and results. A Jupyter notebook should include calls to the implemented function that illustrate the way it

works. The report should explain the feature engineering approach, describe the validation process, and show tables or figures with the accuracy of the classifiers.

3 Suggestions

- Classifiers implemented in the scikit-learn are recommended.
- Be aware that the number of samples is unbalaced since there are more PD samples than healthy ones. Stratified k-fold is suggested as a method to estimate the quality of the classifiers. Similarly, consider measures of the classifier performance that take into account the unbalancedness (e.g., accuracy by classes).
- Implementations can use any other Python library.

References

[1] M Isenkul, B Sakar, and O Kursun. Improved spiral test using digitized graphics tablet for monitoring parkinsons disease. In *Proceedings of the International Conference on e-Health and Telemedicine*, pages 171–175, 2014.

[2] Roberto Santana, Alexander Mendiburu, and Jose A Lozano. Multi-view classification of psychiatric conditions based on saccades. *Applied Soft Computing*, 31:308–316, 2015.

[3] Abdulwahab Sahyoun, Karim Chehab, Osama Al-Madani, Fadi Aloul, and Assim Sagahyroon. Parknosis: Diagnosing parkinson's disease using mobile phones. In *e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on*, pages 1–6. IEEE, 2016.