Machine Learning for the Estimation of UPDRS score

Krisztian Buza, Noémi Ágnes Varga
BioIntelligence Lab, Institute of Genomic Medicine and Rare Disorders, Semmelweis University
Budapest, Hungary, buza@biointelligence.hu, noemiagnesvarga@biointelligence.hu

Background

Parkinson's disease (PD) chronic progressive neurodegenerative condition affecting cca. 7 million in the U.S., primarly elderly people. One of the key components of successful treatment is regular monitoring of the patient's status. Unified Parkinson's Disease Rating Scale (UPDRS) it's worldwide used scoring system to clinical evaluation and follow up the progression of the disease. 90% of PD patients exhibit speech impairment [1].

UPDRS Parts	Time consumption	Usage frequency
Mentation, Behavior and Mood	10 min	60%
Activities of daily living	15 min	80%
Motor Examination	15 min	98%
Complications of therapy	5 min	65%

Aims

In order to continuously measure the patient's UPDRS score, tele-monitoring was proposed. The basic idea is to estimate UPDRS score based on previous UPDRS scores and biomedical voice recordings that can potentially be captured while the patient makes telephone or Skype calls on his/her smartphone or tablet.

Materials and Methods

Data: Parkinsons Telemonitoring Dataset from the UCI data repository [2].

Algorithms and Software: we used the following regression methods to estimate UPDRS (motor) score:

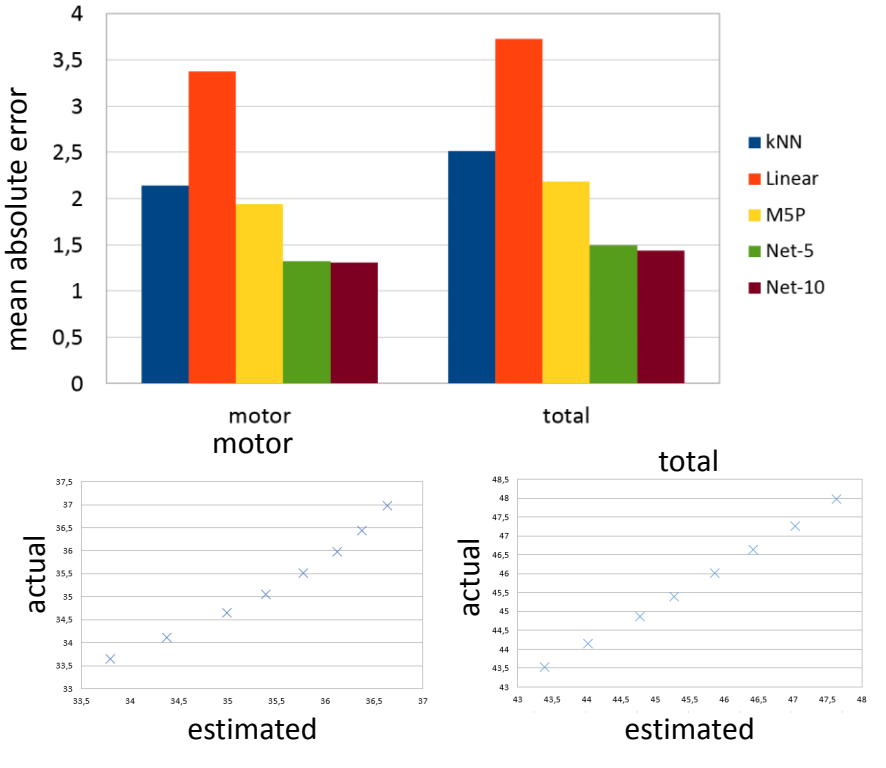
- k-Nearest Neighbor regression with k=10 (kNN)
- Linear Regression (Linear),
- M5P regression tree (M5P)
- Artificial Neural Network: multilayer perceptron with one hidden layer containing 5 and 10 nodes respectively (Net-5, Net-10)

We used the Weka-implementation [3] of Linear, M5P and MLP.

Experimental Settings and Evaluation Metrics:

- As predictors we used the time since recruitment and the four most relevant biomedical voice features according to [2]. These voice features are HNR, RPDE, DFA and PPE.
- *Personalized Telemonitoring* scenario: adaptive system for the estimation of UPDRS score. Temporal train-test splits, first 2/3 of the data is used as training data, the last 1/3 of the data is used as test data. A separate model is build for each patient using only the instances corresponding to that patient resulting in personalized models.
- We averaged the estimations belonging to the same day → final estimation of UPDRS score for a particular day
- We estimated both motor and total UPDRS scores.
- Primary evaluation metric: mean absolute error (MAE)
  We also used normalized mean absolute error NMAE and root mean squared error (RMSE) and observed similar trends.

Results



motor and total UPDRS scores estimated by Net-10 (horizontal axis) vs. actual UPRDS scores (vertical axis) for a selected patient

Discussion

- Out of the examined models, artificial neural networks seem to perform best.
- While automated estimation of UPDRS score with personalised apps running on smartphones or tablets is a promising research direction, however, further research is required in order to achieve accuracy that is high enough in case of real-world applications, esp. if only few training data is available for each patient or the voice data is captured under daily-life conditions (as opposed to well-controlled laboratory settings).
- Additional data sources (based on the user's interaction with the smartphone or tablet, such as her typing patterns or preformance in logical games, etc.) could improve the estimation of UPDRS score.

References

[1] Christopher G. Goetz, Barbara C. Tilley et al. (2008), Movemenz Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale Presentation and Clinimetric Testung Results
[2] Max A. Little, Patrick E. McSharry, Eric J. Hunter, Lorraine O. Ramig (2009), *Suitability of dysphonia measurements for telemonitoring of Parkinson's disease*, IEEE Transactions on Biomedical Engineering, 56(4):1015-1022
[3] Witten, Ian H., and Eibe Frank (2011), *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, third edition.

Acknowledgements - This research was performed within the framework of the grant of the Hungarian Scientific Research Fund - OTKA 111710 PD. This poster was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.