 doi:10.20944/preprints202106.0137.v1

Machine Learning Methods for Classification Multiple Sclerosis with different stages

Atrin Barzegar
[1] Computer Engineering Department, Sapienza University, Rome, Italy
Email: barzegar.1839406@studenti.uniroma1.it

Yas Barzegar
[2] Mechanical Engineering Department, Sapienza University, Rome, Italy
Email: barzegar.1799951@studenti.uniroma1.it

Abstract: Multiple sclerosis (MS) is a debilitating disease of the brain and spinal cord (central nervous system). In MS, the immune system attacks the protective sheath (myelin) that covers the nerve fibers, causing communication problems between the brain and the rest of the body. Eventually the disease can cause permanent damage or nerve damage. The signs and symptoms of MS are very different and depend on the extent of the nerve damage and which nerves are affected. Some people with severe MS may lose the ability to walk independently or completely, while others may experience a long recovery period without any new symptoms. Most people with MS have a relapsing-remitting illness. They experience periods of new symptoms or recurrences that occur over days or weeks and usually improve somewhat or completely. Following these recurrences, there are periods of recovery that can last for months or even years. In this Project, we used some methods of machine learning in order to evaluate the precision and accuracy of Methods to Predict and classification of Multiple Sclerosis with different stages. In order to calculate accuracy, precision, recall Fscore we used some different method such as Art Fuzzy, SVM, Decision tree to compare the classes two by two. To improve the results we used the method of Adaptive fuzzy optimization. we used two options Genetic algorithm and particle swarm optimization.

Keywords: Multiple sclerosis, Machine Learning, precision, Decision tree, Art Fuzzy, SVM

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 June 2021 doi:10.20944/preprints202106.0137.v1

1-Multiple sclerosis:

Multiple sclerosis (MS) is a chronic complex neurodegenerative disease, targeting the central nervous system (CNS) and widely believed to be autoimmune in nature (WEN-JUAN HUANG et al, 2017). Multiple sclerosis (MS) is the most frequent disabling neurological disease in young adults. Whileits etiology remains unknown, MS is a chronic disease of the central nervous system, characterized by inflammation, demyelination, and neurodegenerative pathological processes (Polman et al., 2011). In 85% of the patients, disease onset is characterized by a first acute clinical episode [called clinically isolated syndrome (CIS)], including optic neuritis, paresthesia, paresis, and fatigue (McDonald et al., 2001), evolving into a relapsing-remitting (RRMS) course, and after a delay varying between 15 and 20 years, into a secondary-progressive (SPMS) course, leading to long- term disability. The remaining 15% of MS patients starts with the primary-progressive course (PPMS) (Miller et al., 2005a,b). A revised version of this classification of the MS clinical courseshas been proposed in 2014 by Lublin et al. (2014). In this revision, two main forms of the diseaseare considered: the relapsing remitting MS (CIS and RRMS patients) and the progressive MS (SPMS and PPMS patients), each one being wether active or not active. The course of the diseaseand the risk for developing Permanent disability are very different from one patient to another andthe prediction of long-term disability is still not possible in a new MS patient. Today's neurologist challenge is to predict the individual patient evolution and response to therapy based on the clinical, biological, and imaging markers available from disease onset. Long-term clinical studieshave been conducted to determine the clinical predictors of disability accumulation in MS (Degenhardt et al., 2009; Soldán et al., 2015). In RRMS and SPMS, several negative prognostic factors were identified such as the onset of progression, higher early relapse rate, greater disabilityin the first 5 years, and shorter interval to the second relapse. However, none of these predictors are available at the beginning of the disease (Confavreux et al., 2003; Confavreux and Vukusic, 2006; Scalfari et al., 2010).

2-Assessment of ground reaction forces

Assessment of ground reaction forces and moments (GRF&M) is an important stage in the biomechanical analysis procedure. Conventionally, these measures are recorded using force plate (FP) systems, which, despite their high accuracy, have several significant limitations (Angelos Karatsidis et al, 2017) . Force plates are mechanical sensing systems designed to measure the ground reaction forces and moments involved in human movements. A force plate relies on the use of load cells to determine forces. Force platform, whichis also known as force plate, is used to measure the GRF exerted by the ground on a body in contactwith it. Force plates provide 3D GRF components, the center of pressure components defining the location of the force vector, and orthogonal moment components. Firstly, the fixed position of the plates on the ground together with the requirement to step with the whole foot on the plate for a successful measurement may cause subjects to alter their natural gait pattern. Moreover, due to their high cost, most laboratories are equipped with one or a couple of FPs, which makes trackingmany successive steps during over ground walking impossible.

3-Methodology and Procedure of Project:

In this project We have 2 types of data from foot : 1- normal person 2-MS patients with 3 different stages: 1-MS Stage1 2-MS stage2 3-MS stage3. We have 50 normal person and 50 MS patients with sample frequency of foot 100 HZ. We import data of all normal and patients in Load Data function. The load target determine and show the normal and MS patients and the stages of MS. The command num2string convert the number in to string which has 2 digit. The data obtained by Force plate from the foot of people. The Picture 1 shows the sample data of MS Patient.

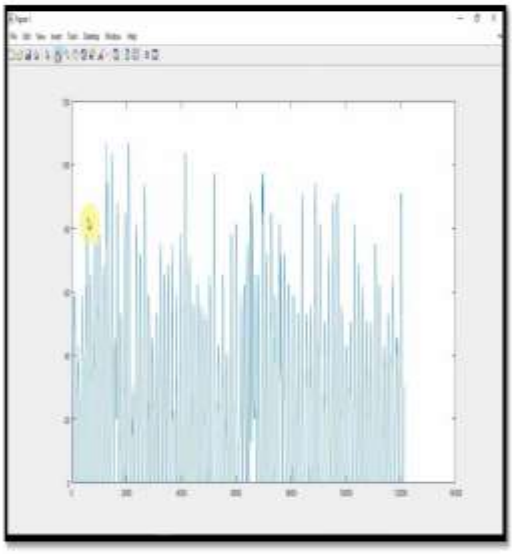

Figure 1:Plot a sample of MS data

3-1 De-noising wavelet:

In the next step we used Denoising is the process with which we reconstruct a signalfrom a noisy one. The Wavelet transform performs a correlation analysis, therefore the output is

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 June 2021 doi:10.20944/preprints202106.0137.v1

expected to be maximal when the input signal most resembles the mother wavelet. Donoho's has initially proposed the fixed thresholding based denoising of signals and images. Here, the value of threshold (t) is computed as:

$$t = \sigma\sqrt{2\,log(n)/n}$$

$$\sigma = \frac{\text{MAD}}{0.6745}$$

MAD is the median of wavelet coefficients and n is the total number of wavelet coefficients. There are four types of Thresholding rules mostly used by different researchers on Denoising applications. In this project we used Denoising with threshold Method which divided the median in to 0.67 with 3 degree wavelet with coif 4.we do this for both group MS and Normal. filtering data and smoothing.

```
%% Denoising Wavelet code
for i=1:50
        for j=1:19
                Data_temp=Data_N{i}(:,j);
                Data_N_D{i}(:,j)    =
                wden(Data_temp,'rigrsure','s','one',3,'coif4');
                Data_temp=Data_M{i}(:,j);
                Data_M_D{i}(:,j)    =
                wden(Data_temp,'rigrsure','s','one',3,'coif4');

        end
end
```

In Picture 2 shows the MS Denois vs MS noise.

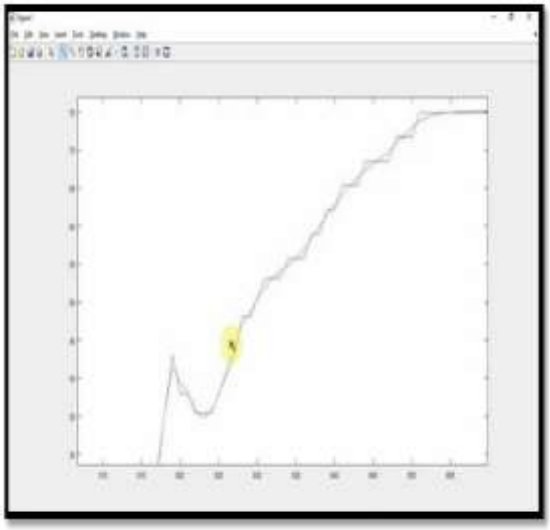

Figure 2: Plot data of MS Denois vs MS noise

3-2 Windowing and Feature Selection:

Data is windowed by selecting a window size. window size determines the number of data points that constitute successive inputs to the model, by which the subsequent output value will be predicted. A small window size might result in insufficient amount of input data to capture a pattern in the data while a large window size might yield a too sophisticated pattern. A large window size also leads to a longer trainingprocess. The process of windowing a signal involves multiplying the time record by a smoothing window of finite length whose amplitude varies smoothly and graduallytowards zero at the edges. The length, or time interval, of a smoothing window is defined in terms of number of samples. Multiplication in the time domain isequivalent to convolution in the frequency domain. Therefore, the spectrum of the windowed signal is a convolution of the spectrum of the original signal with the spectrum of the smoothing window. Windowing changes the shape of the signal in the time domain, as well as affecting the spectrum.

```
%% Windowing
Data_temp=Data(1:60*fs,:);
TotalTime=size(Data_temp,1);
WindowTime=20 *fs; % 20 second
StartWin=(1:WindowTime:TotalTime-
WindowTime+1)';
StopWin=(WindowTime:WindowTime:TotalTi
me)';
Win_Ind=[StartWin StopWin];
```

3-3 Feature Selection:

After that we extracted features of each window for both normal and MS Patients and sent them into the output. The features of all windows shows the features of the signal. Features include Statistics parameters, Frequency parameters, Band Power, Mean Frequency,Median, kurtosis and standard deviation. In the next step we used dimension reduction technique. In statistics, machine learning, and information theory, dimensionality reduction or dimension reduction isthe process of reducing the number of random variables under consideration by obtaining a set of principal variables. Approaches can be divided into feature selection and feature extraction. We used 3 methods of dimension reduction such as LLE, Isomap and Laplacian.

3-4 Dimension reduction:

In the next step we used dimension reduction technique. In statistics, machine learning, and information theory, dimensionality reduction or

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 June 2021 doi:10.20944/preprints202106.0137.v1

dimension reduction is the process of reducing the number of random variables under consideration by obtaining a set of principal variables. Approaches can be divided into feature selection and feature extraction. We used 3 methods of dimension reduction such as LLE, Isomap and Laplacian.

```
%% Dimension Reduction code:
% Features_new=
real(compute_mapping(FeaturesAll, 'LLE',7));
% Features_new=
real(compute_mapping(FeaturesAll
, 'Isomap',7)); Features_new=
real(compute_mapping(FeaturesAll
, 'Laplacian',7));
TargetAll=TargetAll(1:size(Features_new,1));
```

3-5 Test-Train:

In this step we converted data to train and test in 4 classes: normal, MS stage 1, MS stage 2, MS stage 3. We selected 70% of data into train and 30% into test. The function find the target of stage1, stage 2 and stage 3 and put them into index.

3-6 Shuffle

In the next step we shuffle the data. "Shuffle" refers to an improvement of the system. Up till now, and as it was explained before, to train the ANN obtained from the chromosome, its total patterns set is generated and split into two different subsets, train and test patterns subset. Data Shuffling. Simply put, shuffling techniques aim to mix up data and can optionally retain logical relationships between columns. ... Then, it allows production data to be safely used for purposes such as testing and training since all the statistics distribution stays valid. Shuffling is a process of redistributing data across partitions (aka repartitioning) that may or may not cause moving data across JVM processes or even over the wire (between executors on separate machines). Shuffling is the process of data transfer between stages. If test data reach to target test the system is good.

```
%% Shuffle code:
IndRand=randperm(numel(TargetTrain));
TrainData=TrainData(IndRand,:);
TargetTrain=TargetTrain(IndRand);
IndRand=randperm(numel(TargetTest));
TestData=TestData(IndRand,:);
TargetTest=TargetTest(IndRand);
```

3-7 Decision tree

In the next step we used decision tree in Matlab software. We used the function of fitctree. We imported train data and target train to the system and the system gave us out put the out train and also out test. We used the command multiclass param to compare the accuracy, precision, score, recall of the classes two by two. multiclassparam function received target train and target test. Multiclassparam used the function of conv2vec. This function convert the gold standard into binary and also we used in this step another function which is calculate_info_theory calculate the parameters t_p,t_n.

3-8 ART Fuzzy:

A Fuzzy Adaptive Resonance Theory (ART) model capable of rapid stable learning of recognition categories in response to arbitrary sequences of analog or binary input patterns is described. Fuzzy ART incorporates computations from fuzzy set theory into the ART 1 neural network, which learns to categorize only binary input patterns. The generalization to learning both analog and binary input patterns is achieved by replacing appearances of the intersection operator (∩) in ART 1 by the MIN operator (Λ) of fuzzy set theory.

3-9 SVM:

SVM is a supervised machine learning algorithm which can be used for classification or regression problems. It uses a technique called the kernel trick to transform your data and then based on these transformations it finds an optimal boundary between the possible outputs.

3-10 Adaptive fuzzy optimization

To improve the results we used the method of Adaptive fuzzy optimization. we used two options Genetic algorithm and particle swarm optimization. In PSO method we used Train FIS cost function. This function received the parameters of fuzzy and multiply them into a coefficient and set the parameters. With this set parameters and used evalfis command. The out put of this function is not integer so we round the output. And after that we used the command multiclassparam.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 June 2021 doi:10.20944/preprints202106.0137.v1

4-Result and Discussion:

In the figure 3 shows the precision with art fuzzy method. the value of precision percentage is 0.5.

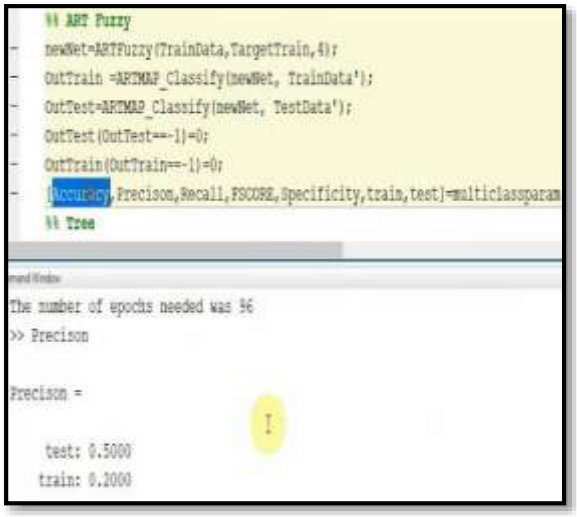

Figure 3: precision with art fuzzy method

In the figure 4 shows the accuracy with art fuzzy method. the value of accuracypercentage is 0.75.

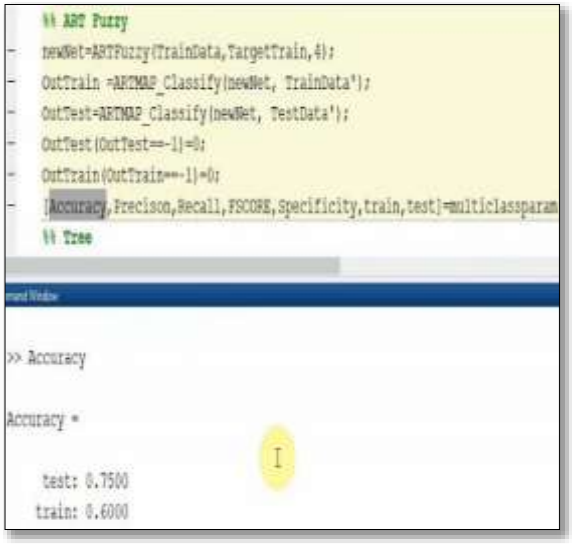

Figure 4: accuracy with art fuzzy method

In the figure 5 show the accuracy with Decision tree method. the value of accuracy Percentage is 0.73.

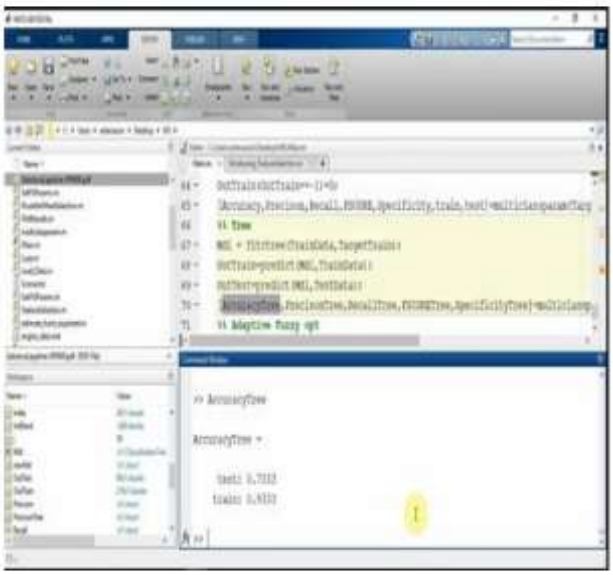

Figure 5: accuracy with Decision tree method

In the figure 6 shows the precision with Decision tree method. the value of precision percentage is 0.46.

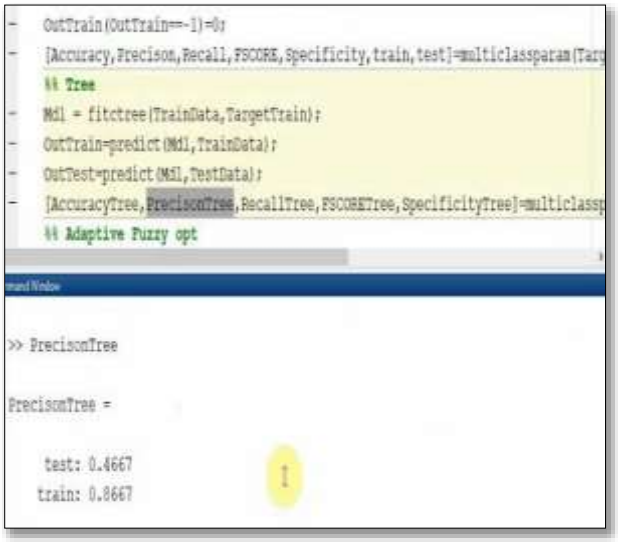

Figure 6: precision with Decision tree method

In the figure 7 shows the precision with SVM method. the value of precisionpercentage is 0.5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 June 2021 doi:10.20944/preprints202106.0137.v1

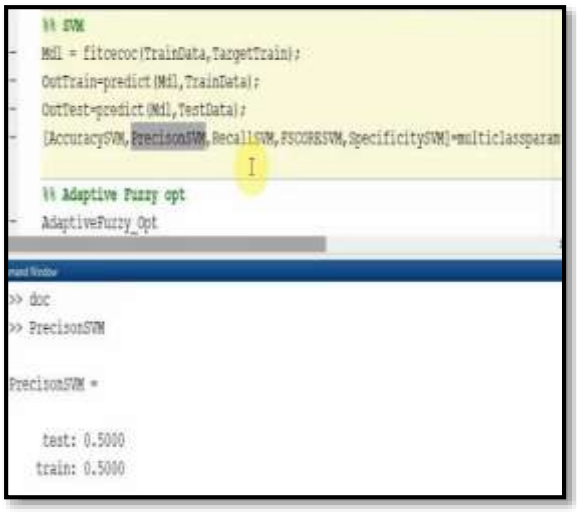

Figure 7: precision with SVM method

In the figure 8 shows the specificity with SVM method. the value of specificitypercentage is 0.83.

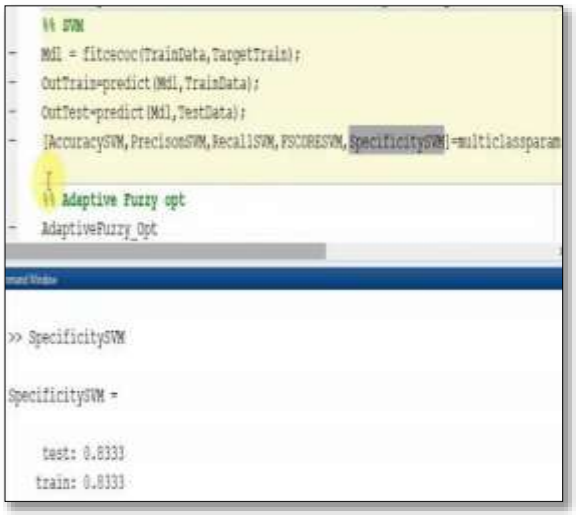

Figure 8: specificity with SVM method

In Figure 9 shows the percentage of specificity with Adaptive Fuzzy optimization .the value of Percentage is 0.82.

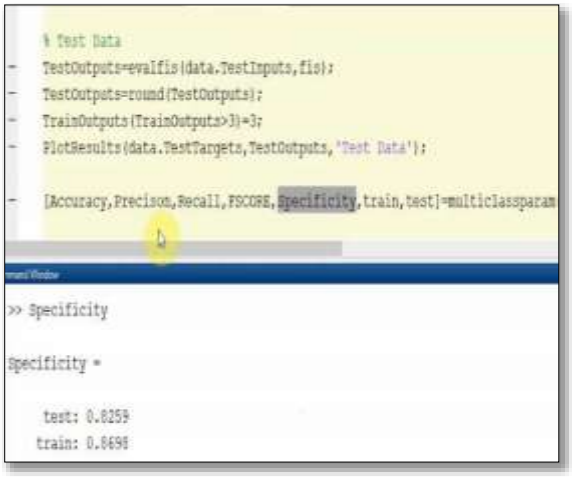

Figure 9:specificity with adaptive fuzzy opt

5-Conclusion:

In this Project, we used some methods of machine learning in order to evaluate the precision and accuracy of Methods to Predict and classification of Multiple Sclerosis with different stages. In order to calculate accuracy, precision, recall Fscore we used some different method such as Art Fuzzy, SVM, Decision tree. The figures 10 and 11 show the result of both train data and test data with target and output data and error value. In the figure 10 shows the root mean square error is 1.286 for test data and Mean error is 0.78 for test data.

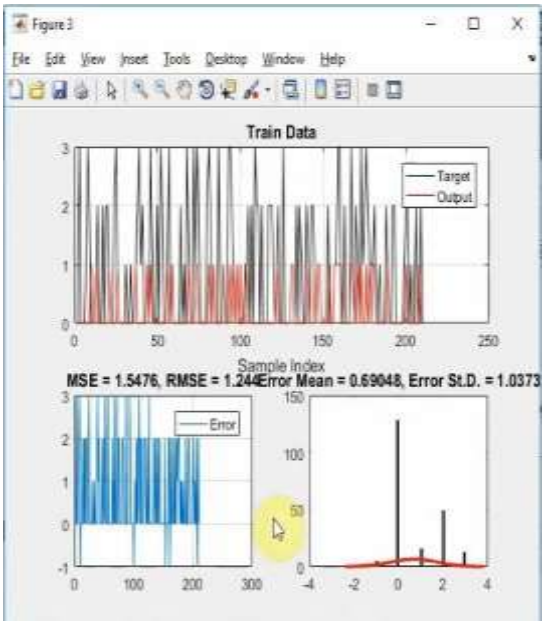

Figure 10: Result of Train data

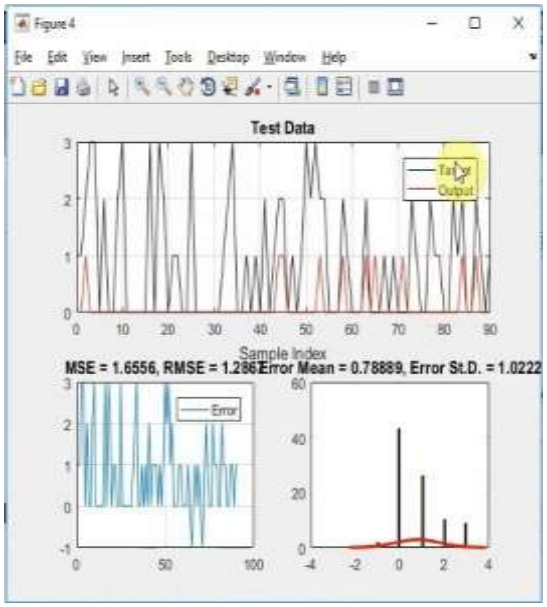

Figure 11: Result of Test data

References:

[1] WEN-JUAN HUANG, WEI-WEI CHEN and XIA ZHANG,(2017), Multiple sclerosis: Pathology, diagnosis and treatments, Department of Neurology, Xuzhou Central Hospital, Xuzhou, Jiangsu 221009, P.R. China, DOI: 10.3892/etm.2017.4410

[2] Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., et al. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the mcdonald criteria. Ann. Neurol. 69, 292–302. doi: 10.1002/ana. 22366

[3] McDonald, W. I., Compston, A., Edan, G., Goodkin, D., Hartung, H.- P., Lublin, F. D., et al. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. Ann. Neurol. 50, 121–127. doi: 10.1002/ ana.1032

[4] Miller, D., Barkhof, F., Montalban, X., Thompson, A., and Filippi, M. (2005a). Clinically isolated syndromes suggestive of multiple sclerosis, part 1: natural history, pathogenesis, diagnosis, and prognosis. Lancet Neurol. 4, 281–288. doi: 10.1016/S1474-4422(05) 70071-5

[5] Degenhardt, A., Ramagopalan, S. V., Scalfari, A., and Ebers, G. C. (2009). Clinical prognostic factors in multiple sclerosis: a natural history review. Nat. Rev. Neurol. 5, 672. doi: 10.1038/nrneurol.20 09.178

[6] Soldán, M. M. P., Novotna, M., Zeid, N. A., Kale, N., Tutuncu, M., Crusan, D. J., et al. (2015). Relapses and disability accumulation in progressive multiple sclerosis. Neurology 84, 81–88. doi: 10.1212/WNL.0000000000 001094

[7] Confavreux, C., and Vukusic, S. (2006). Natural history of multiple sclerosis: a unifying concept. Brain 129, 606–616. doi: 10.1093/brain/ awl007

[8] Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sørensen, P. S., Thompson, A. J., et al. (2014). Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology 83, 278–286. doi: 10.1212/WNL.0000000000 000560

[9] Angelos Karatsidis, Giovanni Bellusci , H. Martin Schepers (2017), Estimation of Ground Reaction Forces and Moments During Gait Using Only Inertial Motion Capture, doi:10.3390/s17010075 w

Authors:

[1] Atrin Barzegar, Computer Engineering Department, Sapienza University, Rome, Italy.
Email: barzegar.1839406@studenti.uniroma1.it

[2] Yas Barzegar, Mechanical Engineering Department, Sapienza University, Rome, Italy
Email: barzegar.1799951@studenti.uniroma1.it