

Classification of Neurodegenerative Diseases using Machine Learning Methods

Fatih Aydin^{*1}, Zafer Aslan²

Accepted : 11/03/2017 Published: 30/03/2017

Abstract: Abstract: In this study, neurodegenerative diseases (Amyotrophic Lateral Sclerosis, Huntington's disease, and Parkinson's disease) were diagnosed and classified using force signals. In the classification, five machine learning algorithms Averaged 2-Dependence Estimators (A2DE), K star (K*), Multilayer Perceptron (MLP), Diverse Ensemble Creation by Oppositional Relabeling of Artificial Training Examples (DECORATE), Random Forest) were compared by the 10-fold Cross Validation method. K* classifier gave the best outcome among these algorithms. As a result of quad classification of the K* classifier, the best classification accuracy was 99.17%. According to the first three and five principal component qualifications which are created from these 19 features, the best classification accuracies of K* classifier were 95.44% and 96.68% respectively

Keywords: Neurodegenerative diseases, Machine Learning, K* classifier, Dimension Reduction, Principal Component Analysis.

1. Introduction

Neurodegeneration is a general term which expresses structure and function loss of neurons including their death. Many neurodegenerative diseases like Amyotrophic Lateral Sclerosis (ALS), Parkinson Disease (PD) and Huntington Disease (HD) occur as a consequence of the neurodegenerative processes. Some of the symptoms of the neurodegenerative diseases as ALS, HD and PD are mutual. Some of these symptoms are physical [1]. One of the activities where physical disorders appear is 'walking' [2]. Early detection of neurodegenerative disease is crucial to prevent the progression of the disease. Furthermore, in certain neurodegenerative diseases, early diagnosis of the disease changes the course of the treatment [3-6]. Many published articles give an idea about the diagnosis and classification of neurodegenerative diseases. Here is brief information about these articles.

In the study carried out by Banaie et al. (2011), a new automatic approach was developed in order to classify the patients who can move by using the features that have come out from the gait signals. In the study, four groups were chosen. These were: people suffering from ALS, HD, PD and Control Object (CO). The classification algorithm used in this study was a quadratic Bayesian classifier [7]. In the study carried out by Daliri (2012), an approach was offered for the diagnosis of neurodegenerative diseases (ALS, HD and PD) based on gait dynamics. Support Vector Machine (SVM) algorithm, was used for diagnosis, by employing different kernels. The best performance was obtained from radial basis kernel function [8]. In the study realized by Lee and Lim (2012), Wavelet based feature extraction was performed by using gait characteristics, in order to classify the Parkinson Disease. Classification was performed by using extracted features and weighted fuzzy membership function neural network (NEWFM) [9]. Drotár et al. (2014) argue that dysgraphia, one of the first motor symptoms observed in PD, can help tell such patients from healthy control objects, using kinematic changes in handwriting to diagnose PD. They first extract the features of the handwriting and form data sets, after which they use SVM to distinguish the classes

[10]. Akdemir et al. (2014) rely on Brain 18-Fluorodeoxyglucose Positron Emission Tomography (18F-FDG PET) to differentially diagnose some Parkinsonisms (PD, MSA, PSP, KBD, dementia with Lewy bodies (DLB) and control objects). The images were analyzed visually and using NeuroQ software [11]. Lee et al. (2014) analyze videos showing gait and posture characteristics of patients with PD to differentiate tem from Progressive Supranuclear Palsy (PSP) and Multiple System Atrophy (MSA) [12]. Pilleri et al. (2014) use heart rate and circadian rhythm of patients with PD and MSA for differential diagnosis. They then put the data to t-test and ANCOVA test to compare the two groups [13]. Navarro-Otano et al. (2014) study 123I-meta-iodobenzylguanidine (123I-MIBG), heart rate, odor recognition, and [123I]FP-CIT (DaTscan) SPECT to differentiate PD from Vascular Parkinsonism (VP). The resulting data is then evaluated statistically [14]. Salvatore et al. (2014) rely on T1 weighted MRI to differentiate PD from PSP, using Principle Component Analysis (PCA) for feature extraction and SVM for classification of diseases [15]. Feng et al. (2014) use 3.0T MRI data for a differential diagnosis of PD and MSA, and statistical tests for comparison [16]. Baudrexel et al. (2014) compare 18F-FDG PET and Diffusion Weighted Imaging (DWI) for a differential diagnosis of PD, MSA and PSP, and then use ANOVA and ROC analysis for classification of the diseases [17]. Huertas-Fernández et al. (2015) use [123I]FP-CIT (DaTscan) SPECT method for differential diagnosis between PD and VP, and then use Logistic Regression, Linear Differential Analysis and SVM for classification of the diseases [18]. Zanigni et al. (2015) rely on MR Proton Spectroscopy (1H MRS) method for differential diagnosis between PD and other parkinsonian syndromes (MSA-C, MSA-P and Richardson syndrome (PSP-RS)). They focus on cerebellum differently from other studies, and analyze statistically data obtained in study [19]. Bradvica et al. (2015) rely on Transcranial Sonography and odor test for differential diagnosis between PD and Essential Tremor (ET) but also compare results of Transcranial Sonography and Dopamin Transporter Scan (DaTSCAN), and find out high degree of compatibility between them [20]. In the study

performed by Xia et al. (2015), a system based on machine learning was developed to classify the neurodegenerative diseases (ALS, HD, PD and normal subjects (healthy subjects)). The algorithms used in the classification were SVM, Random Forest, Multiple Layer Perceptron (MLP) and k-Nearest Neighbors (kNN) [21]. Vranová et al. (2016) rely on Clusterin protein level in cerebrospinal fluid for differential diagnosis between neurodegenerative diseases (PD, DLB, AD, MSA, PSP). They use statistical analysis in comparing the data [22]. Drotár et al. (2016) rely on kinematic and pressure characteristics of handwriting for differential diagnosis between patients with PD and control objects, and then use kNN, AdaBoost and SVM in classifying the two groups [23].

In the studies on the diagnosis of neurodegenerative diseases, diagnosis is done either by focusing on particular diseases (especially PD and CO group) or frequent diseases as ALS, HD and PD are analyzed one by one with the control groups. However, in this study, quad classification was established by using all disease groups and control groups data together. Moreover, the accuracy of this study is higher than the other studies.

2. Material and Methods

In this study, statistical features were derived from gait force signals. After this step, the dataset was formed. Then, among the statistical features that were produced, the features that give more information were chosen using various feature selection methods. Those feature selection methods are: InfoGainAttributeEval, ChiSquaredAttributeEval, ConsistencySubsetEval, GainRatioAttributeEval, ReliefFAttributeEval, SVMAttributeEval and SymmetricalUncertAttributeEval. And the process of feature selection is performed by taking the average of the methods. Chosen features were examined experimentally on whether they increase the performance of Machine Learning algorithms. Thus, feature vector that gives the most information was formed. Since the dimension of this vector was big, its dimension was reduced by using Principal Component Analysis. Feature extraction was completed with this dimension reduction. Finally, the best classifier was chosen using the obtained features. Flow chart of this process is shown in Figure 1.

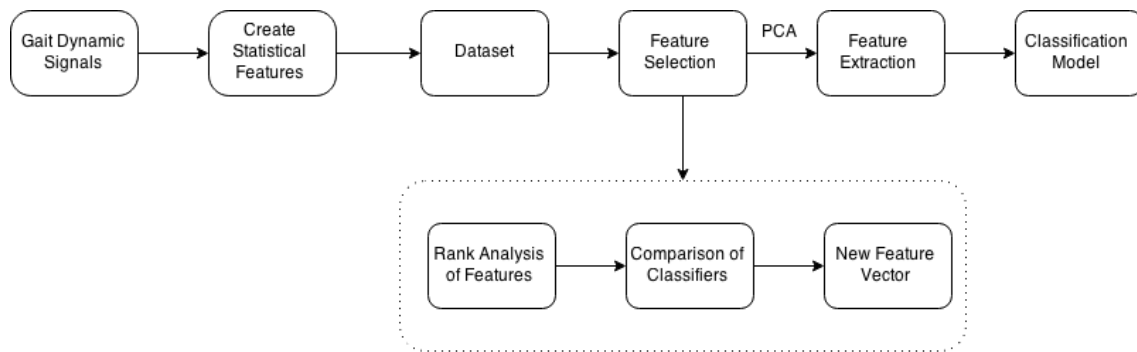


Figure 1. Flowchart of the proposed system for diagnosis and classifying of Neurodegenerative diseases (ALS, PD, HD).

2.1. Gait Dynamics Signals

Gait analysis (GA) used in this study is a kinetic analysis which is one of the measures of 3D GA. The only data which can be measured in kinetic analysis is Ground reaction force vector (GRFV). GRFV is measured with plates that gauge the total impulse of foot on the ground. Data used in the study were collected from the right and left feet of Hausdorff's 64 subjects (15 PD patients, 20 HD patients, 13 ALS patients and 16 healthy people). These data were taken from PhysioNet database [24]. These gait signals from the subjects were sampled with 300 Hz frequency and there are five minutes of data for each subject. The

mechanism which is used to obtain those data was developed by Hausdorff et al [25]. Hausdorff examined ALS and its changes on the gait rhythm. That study is based on the measurements of duration and length of strides on normal individuals and ALS patients [26]. In another study of Hausdorff et al., duration of double step was examined in relation to age and Huntington disease [27]. In that study, 5-minute force signals that were taken from each patient (right and left foot) were split up for 1 minute. So, 640 new data (320 for right foot and 320 for left foot) were created. Neurodegenerative patients' force signals on right and left foot were shown in Figure 2 and Figure 3 respectively.

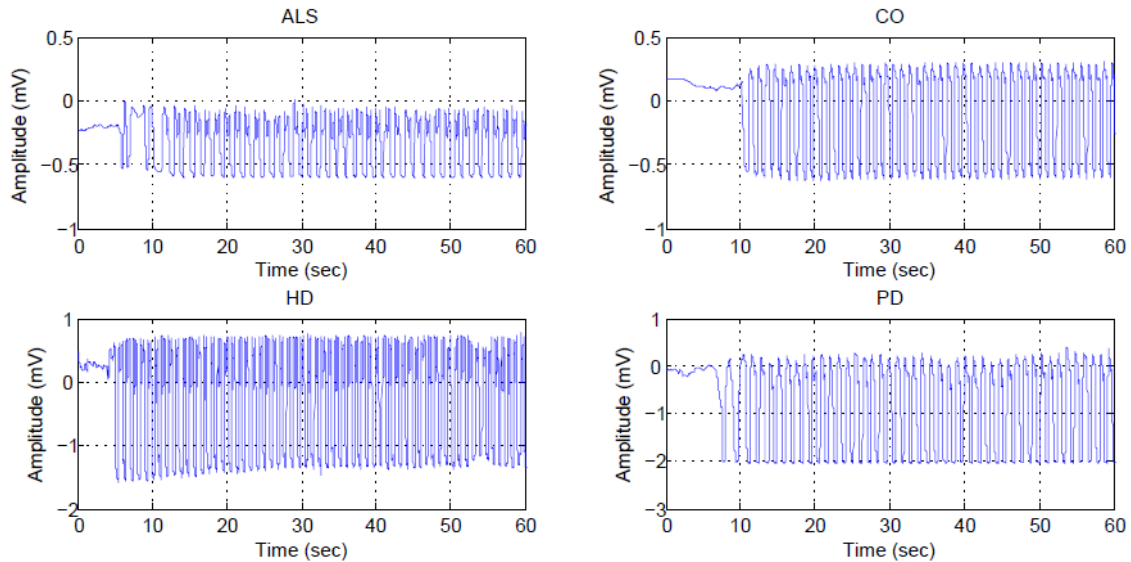


Figure 2. The left foot force-sensitive resistor data for one ALS patient (age: 68, sex: male), one PD patient (age: 77, sex: male), one HD patient (age: 42, sex: male) and one CO subject (age: 57, sex: female). The x-axis label marks the sampling timestamp of the gait signal, and the y-axis label is its amplitude.

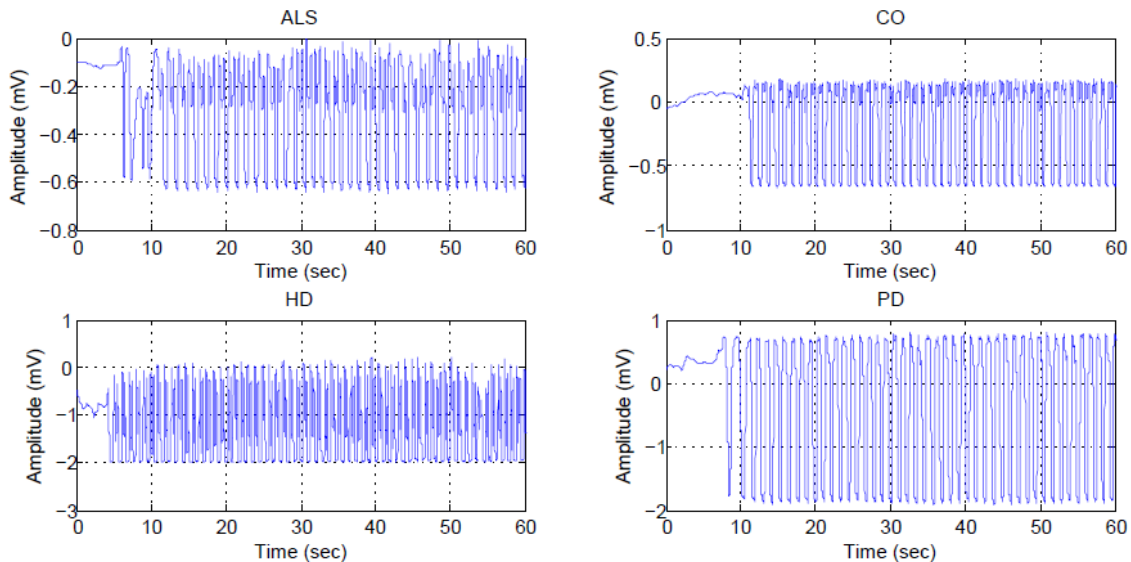


Figure 3. The right foot force-sensitive resistor data for one ALS patient (age: 68, sex: male), one PD patient (age: 77, sex: male), one HD patient (age: 42, sex: male) and one CO subject (age: 57, sex: female). The x-axis label marks the sampling timestamp of the gait signal, and the y-axis label is its amplitude.

2.2. Statistical Features and Preparation of the Datasets

Statistical parameters were formed by using 320 gait signals (force signal) and each parameter was matched to one feature. 26 statistical features (for left foot and right foot) and their abbreviations are as follows: Maximum (F1-F14), minimum (F2-F15), mean (F3-F16), standard deviation (F4-F17), skewness (F5-F18), kurtosis (F6-F19), first quartile (F7-F20), median (F8-F21), third quartile (F9-F22), interquartile range (F10-F23), range (F11-F24), trimmed mean (F12-F25), mean absolute deviation (F13-F26).

The number of statistical features obtained from gait signals was 13. These features were created separately for right and left foot. Later, the dataset was formed. There were 4 datasets: Dataset 1 containing 13 features and only the data of the left foot; Dataset 2 containing 13 features and only the data of the right foot; Dataset 3 containing 13 features and the data of both right and left feet; Dataset 4 containing 26 features and the data of both right and left feet. The classification of all these datasets was done in order to

interpret the causes of the classification results better. Therefore, the causes of classification results could be interpreted better.

2.3. Kstar (K*) Classifier

The algorithm that gave the best results in classification problems in this study is K* algorithm. K* algorithm is also a simple instance-based classification and resembles kNN algorithm [28]. K* algorithm uses entropic measurement based on the probability of transforming an instance within another one through random selection among all possible transforms. In fact, the transform complexity of an instance within another one is the distance between the instances. There are many benefits of using entropic distance. These are; using missing values and real valued attributes.

Let I be a possibly infinite set of instances and T a finite set of transformations on I . Each $t \in T$ maps instances to instances: $t: I \rightarrow I$. K* function is calculated as shown in the statement (Eq. 3).

$$\bar{t}(x) = t_n(t_{n-1}(\dots t_1(x)\dots)) \text{ where } \bar{t} = t_1, \dots, t_n \quad (1)$$

$$\sum_u p(\bar{t}_u) = p(\bar{t}) \quad (2)$$

$$K^*(y_i, x) = -\log_2 P^*(y_i, x) \quad (3)$$

In this statement, P* is the probability of all transformational paths from x to y. It is shown in statement (Eq. 4).

$$P^*(y_i, x) = \sum_{\bar{t} \in P_i(\bar{t}(a)=b)} P(\bar{t}) \quad (4)$$

3. Methodology Statistical Criteria used for the Analysis of Experiments

Some criteria were required in order to decide how successful a classifier is at the end of the learning process. The criteria are as follows:

Kappa Statistics was used to measure consistency between the predicted and the observed classifications in a dataset [29]. Kappa value is given in (Eq. 5). P(a) shows the accuracy of the classifier and P(e) is the expected accuracy gained by the classifier that makes a random prediction on the same data set. In other words, it is the probability of occurrence of the prediction.

$$K = \frac{P(a) - P(e)}{1 - P(e)} \quad (5)$$

F-Measure value (Eq. 6) is stated as the harmonic average of precision and recall values. In terms of the examples classified as positive and negative, precision gives the rate of accurately classified positive ones among the examples predicted as positive.

$$F - Measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (6)$$

F-Measure is particularly an important criterion in the preparation process of training data in order to increase the performance of the classifier. Accordingly, in this study, it is aimed to get a happy graph [30] from learning curve drawn between F-Measure and data size.

Bias-Variance tradeoff has a key role to understand Machine learning algorithms. Use of Bias-Variance tradeoff in experimental studies has gradually increased in recent years [29]. Terms of Bias and Variance help explain how superior the simple predictors can be to the complicated ones and how superior model groups can be to the simple models [31].

Some methodologies were applied within the experiments in this study in order to measure the performance of the learning system correctly, in other words, in order to guarantee the accuracy of experiment results. This method is known as k-fold cross validation (k-fold CV). K-fold CV method was used to determine if training data is enough for a classifier to learn. K-fold CV offers

a method of dividing data into about k equal parts in order to predict a classifier's accuracy. The answer of the question: "why the k value was 10?" in this study is that it was observed that the correct number was about 10 in order to make the most accurate prediction of the rate of error in the common tests done by using different techniques on various training and test sets. Additionally, theoretical proofs support the statement [29]. Leave-one-out cross-validation (LOOCV) which is the specialized phase of k-fold CV method was used sometimes while testing performances of Machine Learning classifiers. This method was used particularly to reveal how different from one another the models created by the classifiers were.

4. Results and Discussion

The experiments were started by choosing the optimum of four datasets gathered from the raw data. Once this dataset was chosen, all of the experiments were done using that dataset. The classifier results for election process of the dataset are shown in Table 1. In order to evaluate the efficiency of the classifiers used in our experiments, we have calculated two baselines: a majority based random predictor; a class distribution based random predictor and a most recent candidate matching predictor. Majority based classification (ZeroR classifier) was done by assigning the most frequent class to each instance, which was the HD class in our dataset. Class distribution based random classification was carried out by randomly assigning classes to the instances on the basis of their distributions. Baseline classifiers were used to compare the performances of other classifiers. Other classifiers were expected to perform better than baseline classifiers. The five chosen classifiers performed better when compared to baseline classifiers. Except the baseline classifiers, the classifiers at the Dataset 4 performed better at classifying in comparison to others according to the results shown on Table 1.

Table 1. The classifying results of classifiers on datasets

Classifier	Dataset-1	Dataset-2	Dataset-3	Dataset-4
Baseline-1 (ZeroR)	31.25	31.25	31.25	31.25
Baseline-2 (Rand)	25.94	26.88	23.91	24.38
A2DE	70.63	75.00	67.34	90.31
MLP	73.75	70.94	66.56	91.25
DECORATE	88.44	87.19	84.38	93.75
K*	88.75	91.25	87.34	96.25
Random Forest	88.75	89.06	86.41	94.69
<i>Average (%)</i>	<i>66.79</i>	<i>67.37</i>	<i>63.88</i>	<i>74.55</i>

Table 2. The results of the classifiers after parameter change

Classifier	Parameter	Accuracy	Precision	Recall	F-measure	Kappa
K*	globalBlend=14	96.88	0.97	0.97	0.97	0.96
DECORATE	desiredSize=45	94.69	0.95	0.95	0.95	0.93
Random Forest	default	94.69	0.95	0.95	0.95	0.93

After Dataset 4 was chosen as training data, the next step was the

selection of the classifier which achieved the highest performance. The parameters of the highest performance classifiers K*, DECORATE and Random Forest were changed in order to inspect the accuracy, precision, recall, f-measure and kappa values. In the light of these results, it was observed that classifier K* had better results compared to other classifiers. The results are shown on Table 2.

After Dataset 4 was chosen as the training data, K* classifier was chosen, the following step was the feature selection. The features that were not required for feature selection were eliminated to increase the accuracy of the classifier, and to decrease the required time. The results were calculated by averaging the seven methods used in feature selection. These results express the effect of each feature in terms of classifying accuracy. According to the data, the F11 feature, which was also called the maximum and minimum range of the left, provided the highest amount of information. The lowest information provided was in the F5 feature which was also called the skewness value.

After the ranks of the features were acquired, starting with the primary information providing feature, new Feature datasets were created by adding secondary and tertiary and lesser information providing features to the primary one. In total, 26 feature datasets were created. The first set contains only the {F11} feature, the second set contains the {F11, F4} features and the final set contains all the {F11, F4, F15, ..., F18, F5} features. According to the accuracy and Root Relative Squared Error (RRSE) values of the model created by Classifier K* with default parameters, the feature dataset which consisted of the first 19 features achieved the highest Accuracy = 98.75% and the lowest Root Mean Squared Error (RRSE) = 15.56%. According to this statement, the first 19 features had the highest information value, and these features were chosen to create the Feature dataset. This dataset consisted of: {F11, F4, F15, F10, F17, F14, F23, F24, F2, F20, F1, F7, F9, F13, F21, F26, F3, F8, F16}.

The Accuracy and RRSE value of Classifier K* using the parameter of Globalblend values are shown in Figure 4. According to all these results, the most suitable Globalblend value was 19.

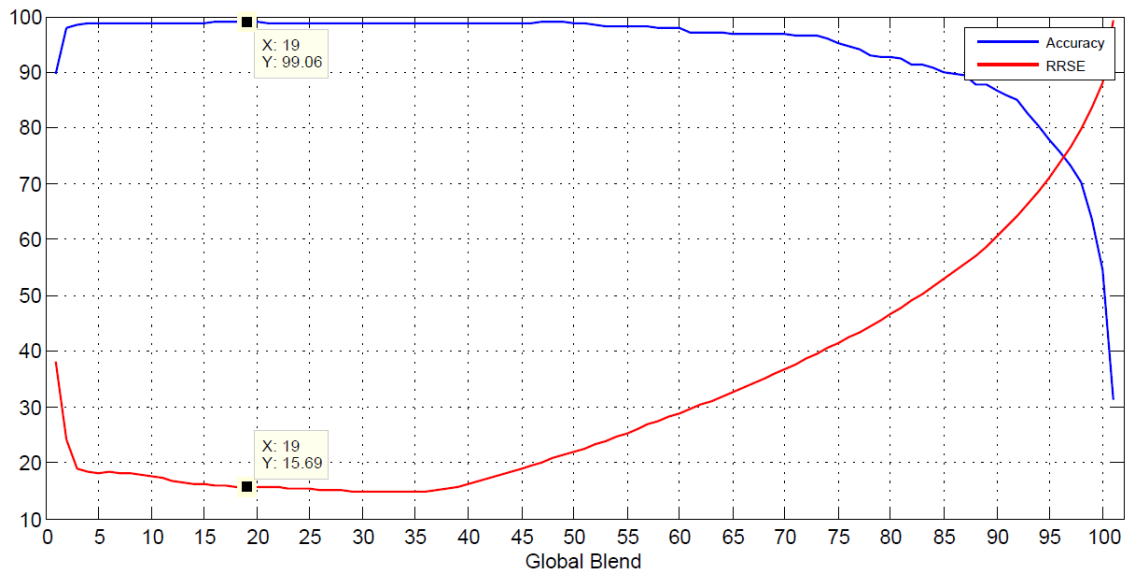


Figure 4. Accuracy and RRSE value of classifier K* according to Global blend parameters.

The Graph of bias-variance is shown in Figure 5. This is a graph that demonstrates the error origination point of a classifier; whether it is bias based or variance based. If the error origination point is based on bias, the convergence of the values of $RMSE_{cv}$ and $RMSE_{training}$ stops at a certain value when the amount of data is increasing. In other words, the error could not be corrected because of bias even if data numbers are the highest. Additionally, there was a significant gap between the values of $RMSE_{cv}$ and $RMSE_{training}$. If the error origination point is based on variance, the situation is the opposite of bias based error. When the amount of

data increases, the values of $RMSE_{cv}$ and $RMSE_{training}$ start to get closer, and the gap between those two the lines closes. On the other hand, the increase in the amount of data does not always decrease the errors caused by variance. When Figure 5 is inspected, it is seen that when the amount of data increases, the gap between two values progressively decreases. In addition, regardless of the amount of data presented, $RMSE_{training}$ always had the value of zero, but the value $RMSE_{cv}$ progressively decreases and finally had the value of 0.07. In conclusion, the complexity of local model created by classifier K* and the complexity of data fit each other.

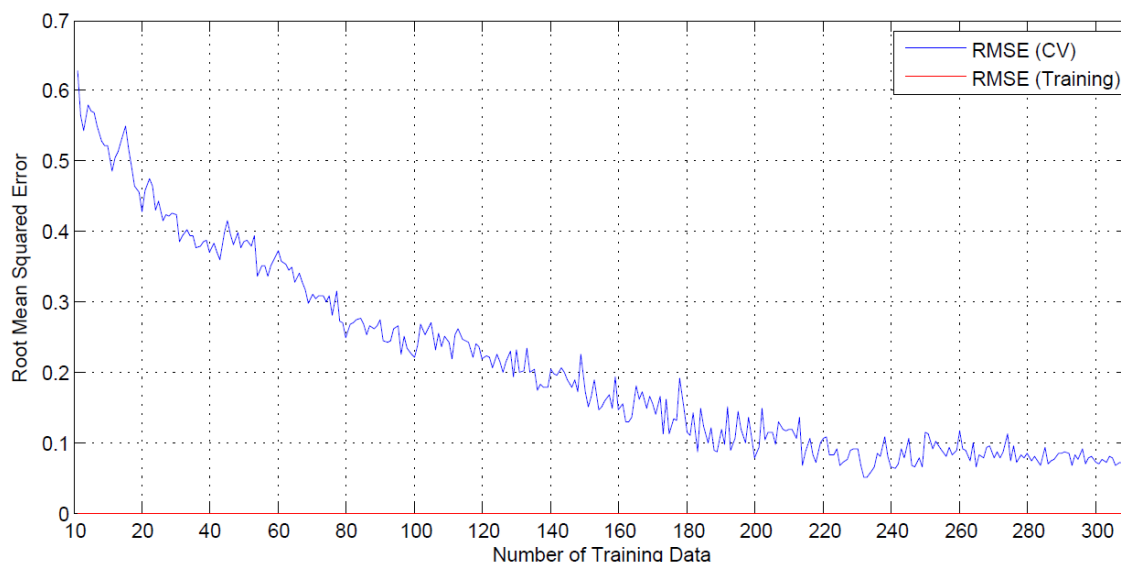


Figure 5. Bias-variance graph of classifier K*

Happy Graph for classifier K* is shown on Figure 6. When Figure 6 is observed, it is seen that when the amount of data is 241, 242 and 243, the value of F-measure is 0.992. In addition, when the amount of data is 320, it is observed that the value of F-measure is 0.991. According to this information, the amount of data needed

for classifier training can be limited to 241. When the amount of data was 241, it was observed that classifier K* achieved Accuracy = 99.17% and RMSE = 0.05 values. According to these results, the amount of data should be limited to 241.

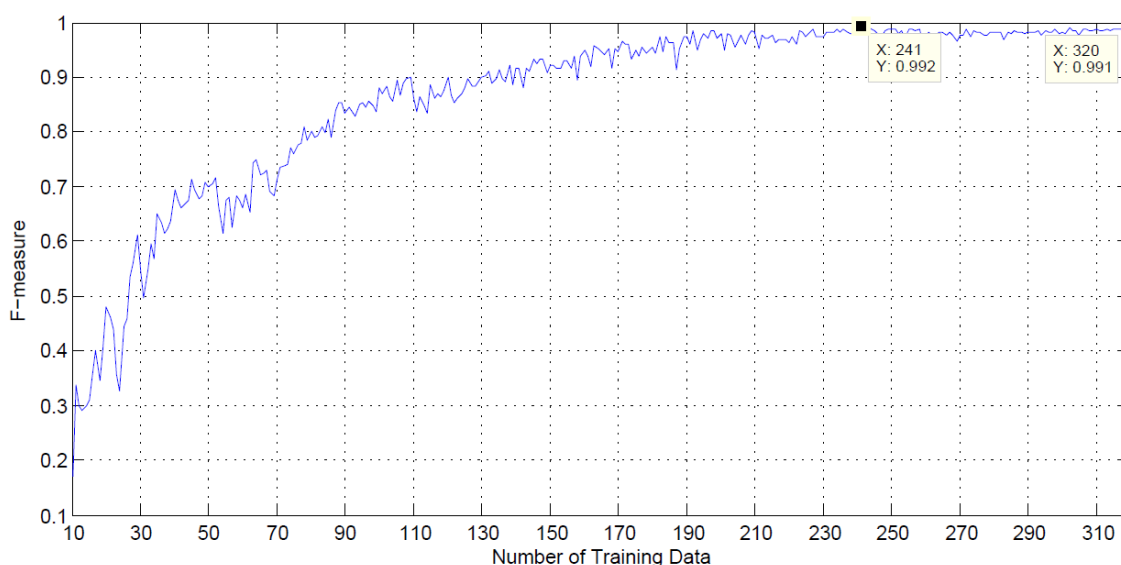


Figure 6. Happy Graph for classifier K*

Table 3. Summary table of classifier K* after classifying process

Statistical summary	Result
Correctly Classified Instances	99.1701 %
Incorrectly Classified Instances	0.8299 %
Kappa statistic	0.9888
Mean absolute error	0.004
Root mean squared error	0.0507
Relative absolute error	1.0817 %
Root relative squared error	11.7975 %
Total Number of Instances	241

The classifying results of classifier K* are shown on Table 3. According to the results, the correct classifying percentage of samples is 99.1701, as 239 samples were classified correctly while only 2 samples were classified incorrectly. It is also possible that classifier achieved perfect learning as the Kappa value was 0.9888. In addition to this, four error measurement statistics proved really low error values.

Detailed accuracy table of classifier K* is shown in Table 4. According to this table, classifier K* created a local model which learns all the classes (ALS, CO, HD, PD) excellently because the precision, recall, F-measure and Receiver Operating Characteristic (ROC) Area values higher than 0.5 were the desired results. All of the values are either 1 or very close to 1 in Table 4. Thus, it is possible to claim that classifier K* created a model that could make generalization on all classes.

Table 4. Detailed accuracy table of K* classifier by class

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	1	0	1	1	1	1	ALS
	0.983	0	1	0.983	0.992	1	CO
	1	0.006	0.988	1	0.994	1	HD
	0.982	0.005	0.982	0.982	0.982	1	PD
Weighted Avg.	0.992	0.003	0.992	0.992	0.992	1	

The percent variability explained by each principal component is shown in Figure 7. The total amount of principal components was 19. Only the first three of the components explain 95.967% of the variability. Hence, the others are not shown in Figure 7. The first

three principal components, feature vectors and data are shown in Figure 8. Additionally, how 19 features contributed to the three principal components was shown in Figure 8 with a vector consisting of length and direction.

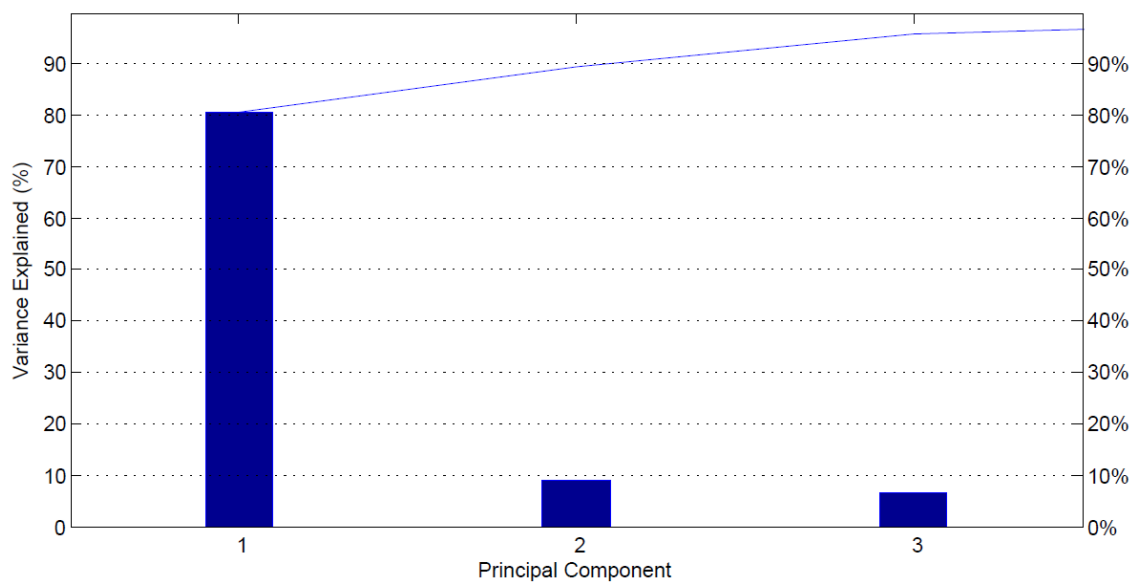


Figure 7. The percent variability explained by each principal component

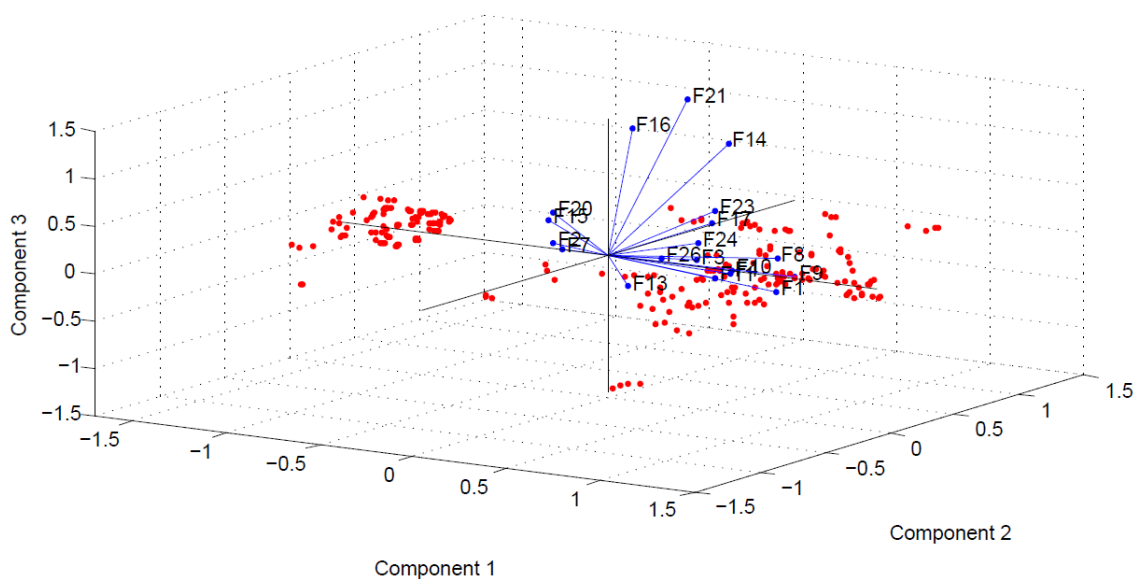


Figure 8. The representation of the first three principal components

Moreover, new features created by the principal components are shown in Figure 8. Classifier K* had the accuracy value of 95.4357% and RMSE value of 0.1555 (GlobalBlend = 14) in classifying when the first three features were chosen. Therefore, 230 out of 241 data were classified correctly with the use of only three features. Additionally, the total data explanation percentage of three principal components was 95.976%. Classifier K* had the accuracy value of 96.6805% and RMSE value of 0.1229 (GlobalBlend = 20) in classifying when the first five features were chosen. Therefore, 233 out of 241 data were classified correctly via using only five features. In addition to these, the total data explanation percentage of five principal components was 98.3201%. It was observed that when other features were added, there was no increase in the results of classifying.

5. Conclusion

In this study, gait signals were used to classify neurodegenerative diseases and control objects. Thirteen statistical features for each left and right foot were created from these signals. Four datasets were created from the constructed features. Compared to others, the dataset which consisted of both left and right data was learned better by all of the classifiers. After dataset was chosen, the most successful classifier was defined as classifier K*. Classifier K* achieved better performance results in comparison to the classifiers of A2DE, MLP, DECORATE and Random Forest. After that step, the election of the features with the highest informative value was carried out. At this step, 19 out of 26 features were elected. (F1, F2, F3, F4, F7, F8, F9, F10, F11, F13, F14, F15, F16, F17, F20, F21, F23, F24, F26) After the election of features, the parameters that had the best performance for classifier K* were chosen. Therefore, GlobalBlend parameter value was defined to be 19. The gap between the classifying results of Classifier K* were observed when the value of k was changed during CV process. It was observed that the performance of classifier K* didn't change when the value of k was changed during CV process. According to that information, it is possible to state that classifier K* had a stable condition. After that step, number adjustment of training data used for the training process of classifier K* which consisted of 320 numbers of data was carried out. A happy graph was created and the number of data required for the training process was defined to be 241. Therefore, during 10-fold CV process, classifier K* achieved the results of Accuracy = 99.1701% and RMSE = 0.0507 when GlobalBlend parameter was 19 and training dataset consisting of 241 number of data with 19 features was used.

The 19 features which were obtained by feature selection process were used to create new features by PCA method. These features were created by first three and first five principal components. During 10-fold CV process for the first three principal components, classifier K* achieved the results of Accuracy = 95.4357% and RMSE = 0.1555 when GlobalBlend parameter was 14 and a training dataset consisting of 241 number of data was used. During 10-fold CV process for the first five principal components, Classifier K* achieved the results of Accuracy = 96.6805% and RMSE = 0.1229 when GlobalBlend parameter was 20 and a training dataset consisting of 241 number of data was used.

Acknowledgements

We would like to thank Istanbul Aydin University and Kırklareli University for supporting us to realize our studies.

References

- [1] Kaynak Selekler (2010). Alois alzheimer ve alzheimer hastalığı. *Türk Geriatri Dergisi*. Vol. 13. Pages. 9-14.
- [2] Nadire Özaras and Selim Yalçın (2002). Normal yürüme ve yürüme analizi. *Turkish Journal of Physical Medicine and Rehabilitation*. Vol. 48(3).
- [3] National Collaborating Centre for Chronic Conditions (2006). *Parkinson's Disease: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care*. London, UK: Royal College of Physicians.
- [4] Jeffrey M. Bronstein, Michele Tagliati, Ron L. Alterman, Andres M. Lozano, Jens Volkmann, Alessandro Stefani, Fay B. Horak, Michael S. Okun, Kelly D. Foote, Paul Krack, Rajesh Pahwa, Jaimie M. Henderson, Marwan I. Hariz, Roy A. Bakay, Ali Rezai, William J. Marks, Elena Moro, Jerrold L. Vitek, Frances M. Weaver, Robert E. Gross and Mahlon R. DeLong (2011). Deep brain stimulation for Parkinson disease: an expert consensus and review of key issues. *Arch. Neurol.* Vol. 68 (2). Pages. 165.
- [5] Samuel Frank and Joseph Jankovic (2010). Advances in the pharmacological management of huntington's disease. *Drugs*. Vol. 70(5). Pages. 561-571.
- [6] Patrick Russell and Roger Harrison (2014). What is amyotrophic lateral sclerosis?. *Clinical Pharmacist*. Vol. 6(7).
- [7] Masood Banaie, Mohammad Pooyan and Mohammad Mikaili (2011). Introduction and application of an automatic gait recognition method to diagnose movement disorders that arose of similar causes. *Expert Systems with Applications*. Vol. 38(6). Pages. 7359-7363.
- [8] Mohammad R. Daliri (2012). Automatic diagnosis of neurodegenerative diseases using gait dynamics. *Measurement*. Vol. 45(7). Pages. 1729-1734.
- [9] Sang-Hong Lee and Joon S. Lim (2012). Parkinson's disease classification using gait characteristics and wavelet-based feature extraction. *Expert Systems with Applications*. Vol. 39(8). Pages. 7338-7344.
- [10] Drotár, P., Mekyskaa, J., Rektorová, I., Masarováb, L., Smékala, Z. ve Faundez-Zanuy, M. 2014. "Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease", *Computer Methods and Programs in Biomedicine*, 117(3), 405-411.
- [11] Akdemir, Ü.Ö., Tokçaa, B., Karakuş, A. ve Kapucu, L.Ö. 2014. "Brain 18F-FDG PET Imaging in the Differential Diagnosis of Parkinsonism", *Clinical Nuclear Medicine*, 93(3), 220-226.
- [12] Lee, S., Kim, M., Lee, H.M., Kwona, K., Kim, H. ve Koh, S. 2014. "Differential diagnosis of parkinsonism with visual inspection of posture and gait in the early stage", *Gait & Posture*, 39(4), 1138-1141.
- [13] Pilleri, M., Levedianos, G., Weis, L., Gasparoli, E., Facchini, S., Biundo, R., Formento-Dojot, P. ve Antonini, A. 2014. "Heart rate circadian profile in the differential diagnosis between Parkinson disease and multiple system atrophy", *Parkinsonism and Related Disorders*, 20(2), 217-221.
- [14] Navarro-Otano, J., Gaig, C., Muxi, A., Lomeña, F., Compta, Y., Buongiorno, M.T., Martí, M.J., Tolosa, E. ve Valldeoriola, F. 2014. "123I-MIBG cardiac uptake, smell identification and 123I-FP-CIT SPECT in the differential diagnosis between vascular parkinsonism and Parkinson's disease", *Parkinsonism and Related Disorders*, 20(2), 192-197.
- [15] Salvatorea, C., Cerasab, A., Castiglioni, I., Gallivanonec, F., Augimerib, A., Lopezd, M., Arabiae, G., Morellie, M.,

- Gilardi, M.C. ve Quattrone, A. 2014. "Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and Progressive Supranuclear Palsy", *Journal of Neuroscience Methods*, 222, 230-237.
- [16] Yi Xia, Qingwei Gao and Qiang Ye (2015). Classification of gait rhythm signals between patients with neuro-degenerative diseases and normal subjects: Experiments with statistical features and different classification models. *Biomedical Signal Processing and Control*. Vol. 18. Pages. 254-262.
- [17] Feng, J., Huang, B., Yang, W., Zhang, Y., Wang, L., Wang, L. ve Zhong, X. 2014. "The putaminal abnormalities on 3.0T magnetic resonance imaging: can they separate parkinsonism-predominant multiple system atrophy from Parkinson's disease?", *Acta radiologica*, 56(3), 322-328.
- [18] Baudrexel, S., Seifried, C., Penndorf, B., Klein, J.C., Middendorp, M., Steinmetz, H., Grünwald, F. Ve Hilker, R. 2014. "The Value of Putaminal Diffusion Imaging Versus 18-Fluorodeoxyglucose Positron Emission Tomography for the Differential Diagnosis of the Parkinson Variant of Multiple System Atrophy", *Movement Disorders*, 29(3), 380-387.
- [19] Huertas-Fernández, I., García-Gómez, F.J., García-Solís, D., Benítez-Rivero, S., Marín-Oyaga, V.A., Jesús, S., Cáceres-Redondo, M.T., Lojo, J.A., Martín-Rodríguez, J.F., Carrillo, F. ve Mir, P. 2015. "Machine learning models for the differential diagnosis of vascular parkinsonism and Parkinson's disease using [123I]FP-CIT SPECT", *European journal of nuclear medicine and molecular imaging*, 42(1), 112-119.
- [20] Zanigni, S., Testa, C., Calandra-Buonaura, G., Sambati, L., Guarino, M., Gabellini, A., Evangelisti, S., Cortelli, P., Lodi, R. ve Tonon, C., 2015. "The contribution of cerebellar proton magnetic resonance spectroscopy in the differential diagnosis among parkinsonian syndromes", *Parkinsonism and Related Disorders*, 21(8), 929-937.
- [21] Bradvica, I.S., Mihaljevic', I., Butkovic'-Soldo, S., Kadojic', D., Titlic', M., Bradvica, M. ve Kralik, K. 2015. "Transcranial sonography and the pocket smell test in the differential diagnosis between parkinson's disease and essential tremor", *Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 36(8), 1403-1410.
- [22] Vranová, H.P., Hényková, E., Mareš, J., Kaiserová, M., Menšíková, K., Vašík, M., Hlušík, P., Zapletalová, J., Strnad, M., Stejskal, D. ve Kaňovský, P. 2016. "Clusterin CSF levels in differential diagnosis of neurodegenerative disorders", *Journal of the Neurological Sciences*, 361, 117-121.
- [23] Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z. ve Faundez-Zanuy, M. 2016. "Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease", *Artificial Intelligence in Medicine*, 67, 39-46.
- [24] PhysioNET. Gait Dynamics in Neuro-Degenerative Disease Data Base. <http://www.physionet.org/physiobank/database/gaitnidd/>. [Accessed: 13.04.2015].
- [25] Jeffrey M. Hausdorff, Zvi Ladin and Jeanne Y. Wei (1995). Footswitch system for measurement of the temporal parameters of gait. *J Biomech*. Vol. 28(3). Pages. 347-351.
- [26] Jeffrey M. Hausdorff, Susan L. Mitchell, Renee Firtion, Chung-Kang Peng, Merit E. Cudkovicz, Jeanne Y. Wei and Ary L. Goldberger (1997). Altered fractal dynamics of gait: reduced stride-interval correlations with aging and huntington's disease. *J Applied Physiology*. Vol. 82(1). Pages. 262-269.
- [27] Jeffrey M. Hausdorff, Apinya Lertratanakul, Merit E. Cudkovicz, Amie L. Peterson, David Kaliton and Ary L. Goldberger (2000). Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis. *J Applied Physiology*. Vol. 88. Pages. 2045-2053.
- [28] Pamela K. Douglas, Sam Harris, Alan L. Yuille and Mark S. Cohen (2011). Performance comparison of machine learning algorithms and number of independent components used in fMRI decoding of belief vs. disbelief. *NeuroImage*. Vol. 56(2). Pages. 544-553.
- [29] Ian H. Witten and Eibe Frank (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd ed. San Francisco, CA, USA: Morgan Kaufmann.
- [30] Stuart J. Russell and Peter Norvig (2003). *Artificial Intelligence: A Modern Approach*. 2nd ed. Upper Saddle River, New Jersey, USA: Prentice Hall.
- [31] Pedro Domingos (2000). A unified bias-variance decomposition and its applications, In: *Proceedings of the Seventeenth International Conference on Machine Learning*, 29 June-2 July 2000; Stanford, CA, USA. San Francisco, CA, USA: Morgan Kaufmann. pp. 231-238.