



Yapay öğrenme yöntemleri ve dalgacık dönüşümü kullanılarak nöro dejeneratif hastalıkların teşhisi

Fatih Aydın^{1*}, Zafer Aslan²

¹Kırklareli Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Bilgisayar Programcılığı Programı, 39020, Kırklareli, Türkiye

²İstanbul Aydın Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği, 34295, İstanbul, Türkiye

Ö N E Ç İ K A N L A R

- Ayrık dalgacık dönüşümü ile istatistiksel öznitelik oluşturma
- Tek kural algoritması ile yüksek enformasyon veren özniteliklerin seçilmesi
- Yapay öğrenme yöntemleri ile nöro-dejeneratif hastalıkların sınıflandırılması

Makale Bilgileri

Geliş: 30.03.2016

Kabul: 22.02.2017

DOI:

10.17341/gazimmfd.337621

Anahtar Kelimeler:

Nöro-dejeneratif hastalıklar,
yapay öğrenme,
dalgacık dönüşümü,
radyal tabanlı fonksiyon ağı,
boosting

ÖZET

Bu çalışma da yere uygulanan kuvvet sinyalleri kullanılarak Amyotrofik lateral skleroz (ALS), Huntington hastalığı (HD) ve Parkinson hastalığı (PD) gibi nöro dejeneratif hastalıkların (NDD) sınıflandırılmasında kullanılabileceği önerilmektedir. Deneyler 16 kontrol bireyi (CO), 13 ALS, 20 HD ve 15 PD'ye ait veriler kullanılarak gerçekleştirildi. İlk olarak kuvvet sinyalleri, Discrete Meyer (dmey) dalgacığı kullanılarak yedinci seviyeye kadar ayrıştırıldı. Yeni oluşan sinyallerden yedinci seviyedeki yaklaşım sinyali seçildi. Bu sinyal üzerinde tepe (peak) analizi gerçekleştirilerek sinyalin lokal maksimumları, tepe'nin x eksenindeki değerleri, tepe genişliği ve tepe çıkıntıları elde edildi. Daha sonra bu dört tepe özelliğinin her birinden 15 adet temel istatistiksel öznitelik elde edildi. Böylelikle sol ayak için 60 ve sağ ayak için 60 olmak üzere toplamda 120 öznitelik elde edildi. Daha sonra OneRules sınıflandırıcı kullanılarak bu öznitelikler içerisinde en çok enformasyon veren öznitelikler seçildi. Bir sonraki aşamada ise Radyal Tabanlı Fonksiyon Ağı (RBFNetwork), Adaptif Yükseltme (Adaboost) ve Eklemeli Lojistik Regresyon (LogitBoost) algoritmaları kullanılarak ALS-CO için %93,1 doğruluk, HD-CO için %97,22 doğruluk, PD-CO için %83,87 doğruluk ve NDD-CO için %92,18 doğruluk elde edildi.

Diagnosis of neuro degenerative diseases using machine learning methods and wavelet transform

H I G H L I G H T S

- Generating statistical features with discrete wavelet transform
- Selecting features with high information using one rule algorithm
- Classification of neuro-degenerative diseases using machine learning methods

Article Info

Received: 30.03.2016

Accepted: 22.02.2017

DOI:

10.17341/gazimmfd.337621

Keywords:

Neuro-degenerative diseases,
machine learning,
wavelet transform,
radial basis function
network,
boosting

ABSTRACT

This study suggests that the force signals applied to the ground may be used to classify neuro-degenerative diseases (NDD) such as Amyotrophic lateral sclerosis (ALS), Huntington's disease (HD) and Parkinson's disease (PD). The experiments were performed using data with 16 control subjects (CO), 13 ALS, 20 HD and 15 PD. Firstly, the force signals were separated up to level-7 using Discrete Meyer (dmey) wavelet. Among the new signals, the approach signal at the seventh level was selected. The local maximums of the peaks, peak locations, peak widths and peak prominences were obtained by performing peak analysis on this signal. Then, 15 basic statistical features from each of these four peak features were obtained. Thus, 60 for each of left and right foot, 120 features were obtained. Among these 120 features, the ones giving the highest information were selected using OneRules classifier. Respectively, 93.1%, 97.22%, 83.87% and 92.18% accuracy was obtained on ALS-CO, HD-CO, PD-CO and NDD-CO datasets using Radial Basis Function Network (RBFNetwork), Adaptive Boosting (Adaboost) and Additive Logistic Regression (LogitBoost) algorithms.

1. GİRİŞ (INTRODUCTION)

Nörodejenerasyon nöronların ölümü de dahil, nöronların yapı ve işlev kaybını ifade eden genel bir ifadedir. Amyotrofik lateral skleroz (ALS), Parkinson hastalığı (PD), Alzheimer (AD) ve Huntington hastalığı (HD) gibi birçok nöro-dejeneratif hastalık nöro-dejeneratif süreçlerin bir sonucu olarak ortaya çıkar. Nöro-dejeneratif hastalıklar tedavi edilemez ve bu ilerleyen bozulma sinir hücrelerinin ölümüne neden olmaktadır [1]. ALS, HD ve PD'de ortak belirti hastalığın kendisini ilk aşamada fiziksel olarak göstermesidir. Fiziksel bozuklukların kendisini gösterdiği süreçlerden biri de yürüme eylemidir. Modern yürüme analizi; normal ve patolojik yürüyüşü inceleyebilmemiz için eklem ve kasların fonksiyonlarını sayısal veriye dönüştürür [2]. Yürüme, kendi kendini tekrar eden hareketlerden oluşur. Bu hareketler, bir yürüyüş döngüsü olarak tanımlanır [3]. Aslında yürüme, beyinde başlar. Bu nedenle yürüme sorunlarında değerlendirme, beyinden medulla spinalise, oradan da kas ve eklemlere doğru olmalıdır [4].

Alt ekstremité hareketinin her sekansı (yürüme siklusu) bir ağırlık taşıyıcı periyot ve bir self-ilerleme intervali içerir. Normal yürüme siklusu süresinin %60'ı duruşta, %40'ı sallanmada harcanır. Bu intervallerde harcanan süre yürüme hızına bağlı olarak değişir. Ayrıca bireyler arası farklılıklar da vardır [5]. Newton'un üçüncü kanununa (etkiye karşı zıt yönde tepki) göre ayaktaki bir insanın yerde oluşturduğu ağırlık kuvvet vektörüne; büyüklüğü aynı, yönü ters bir kuvvet ile karşılık verir. Buna Yer Tepkime Kuvvet Vektörü (YTKV) denir. Yürürken YTKV, vücut ağırlığı ve hareketi sağlayan kas kuvvetlerinin bileşkesine karşı oluşur ve yürüme eylemi boyunca yönü ve büyüklüğü sürekli değişir [6].

Bu makale aşağıda sıralanan bölümlerden oluşmaktadır: 2. Bölümde nöro-dejeneratif hastalıkların önceden tanısı ile ilgili yapılmış çalışmalara yer verildi. 3. Bölümde sınıflandırma aşamasında kullanılan verilerin nasıl elde edildiği belirtildi. 4. Bölümde yürüme kuvvet sinyallerinin ayrıştırılmasında kullanılan sinyal işleme yöntemi olan dalgacık dönüşümü açıklandı. 5. Bölümde eğitim verilerinin sınıflandırılmasında kullanılan yapay öğrenme algoritmaları açıklandı. 6. Bölümde nöro-dejeneratif hastalıkların sınıflandırılması sonucunda ortaya çıkan sonuçların değerlendirilmesi için kullanılan istatistikî kriterler ve deney yöntemleri tanımlandı. 7. Bölümde eğitim kümesinin hazırlanma süreci açıklandı. 8. Bölümde deney sonuçlarına ve değerlendirmelere yer verildi. Son olarak 9. Bölümde de sonuçlara yer verildi.

2. İLİŞKİLİ ÇALIŞMALAR (RELATED WORKS)

Nöro-dejeneratif hastalıkların (ALS, HD, PD) türünün erken belirlenmesi hastalığın ilerlemesini geciktirmek açısından önemlidir. Bunun yanı sıra bazı nöro-dejeneratif hastalıklarda teşhisin erken konması tedavinin seyrini değiştirmektedir [7]. Bu durum Bronstein vd. [8] tarafından

yapılan PD çalışmalarında, Frank ve Jankovic [9] tarafından yapılan HD çalışmalarında ve son olarak Russell ve Harrison [10] tarafından yapılan ALS hastalığı çalışmalarında belirtilmektedir. Bu nedenle erken teşhisin ve yüksek oranda doğruluk veren bir sınıflandırma yönteminin geliştirilmesi önemli olmaktadır. Nöro-dejeneratif hastalıkların teşhisi ve sınıflandırılmasına dair fikir veren birçok makale yayınlanmıştır. Han ve grubu tarafından yapılan çalışmada Parkinson hastalarının yürüyüşünü güvenli bir şekilde analiz etmek için yeni bir yürüme faz ayırıcı metot, yürüme boyunca açının üç eksenli ivmesini analiz etmek üzere geliştirilmiştir. Bu çalışmada yürüme fazlarını ayıran karakteristik tepeler yani hızlanma ve yavaşlama esnasındaki maksimum noktalar kullanılarak sınıflandırma doğruluğu %97,6 ve %99,4 olarak elde edilmiştir [11]. Dutta ve gurubu tarafından yapılan çalışmada Elman'ın yinelemeli sinir ağı (ERNN) sağlıklı ve patolojik yürüyüşün (ALS, PD, HD) otomatik saptanması için kullanılmış ve ilgili yürüme örüntülerinden patolojik yürümelerdeki nörolojik bozuklukların teşhisi gerçekleştirilmiştir.

Sağlıklı ve patolojik yürüyüşün sınıflandırma doğruluğu %90,6 olarak elde edilmiştir. ALS, PD ve HD gruplarının sınıflandırma doğruluğu ise %89,8 olarak elde edilmiştir [12]. Wu ve Krishnan [13] tarafından yapılan çalışmada 16 sağlıklı birey ve 13 ALS'li bireyden kaydedilen salınım-aralığı verileri kullanılarak zaman serisi dalgalanmalarını ölçmek için sinyal dönüş sayısı metodu kullanılmıştır. Bu çalışmadaki sınıflandırma doğruluğu %89,66 olarak elde edilmiştir. Wu ve Krishnan [14] tarafından yapılan çalışmada Parkinson hastalarının yürüme değişimlerini değerlendirmek için yürümenin iki alt fazı (salınım ve durma) ve adım aralığı parametre olarak seçilmiş ve bu parametrelerin olasılık yoğunluk fonksiyonunu kestirmek için parametrik olmayan Parzen-window metodu kullanılmıştır. Bu çalışmadaki sınıflandırma doğruluğu %90,32 olarak elde edilmiştir. Banaie ve grubu [15] tarafından yapılan çalışmada, yürüme sinyallerinden çıkarılan öznitelikler kullanılarak hareket edebilen hastaları (ALS, HD, PD) sınıflandırmak için kuadratik Bayes normal sınıflandırıcı kullanılarak yeni otomatik bir yaklaşım geliştirilmiştir. Bu çalışmadaki sınıflandırma doğruluğunun ortalama sonucu %86,96 olarak elde edilmiştir. Manap ve grubu [16] tarafından yapılan çalışmanın amacı normal yürüme boyunca Parkinson hastalarındaki anormal yürüme örüntülerini tanımlamakta kullanılan parametreleri incelemektir. Bu çalışmadaki sınıflandırma doğruluğunun sonucu en yüksek %95,63 olarak elde edilmiştir. Daliri [17] tarafından yapılan çalışmada, yürüme dinamiklerine dayalı nöro-dejeneratif hastalıkların (ALS, HD ve PD) teşhisi için bir yaklaşım önerilmiştir. Bu çalışmada ALS-CO için %96,79 doğruluk, PD-CO için %89,33 doğruluk ve HD-CO için %90,28 sınıflandırma doğruluğu elde edilmiştir. Lee ve Lim [18] tarafından yapılan çalışmada Parkinson hastalığını sınıflandırmak için yürüme karakteristikleri kullanılarak Wavelet tabanlı öznitelik çıkarma işlemi gerçekleştirilmiştir. Bu çalışmadaki sınıflandırma doğruluğu en yüksek %77,33 olarak elde edilmiştir. Xia ve grubu [19] tarafından yapılan

çalışmada nöro-dejeneratif hastalıkların (ALS, Huntington ve Parkinson) ve normal deneklerin sınıflandırılması için Yapay Öğrenmeye dayalı bir sistem geliştirilmiştir. Bu çalışmada NDD-CO için %96,83 doğruluk, ALS-CO için %96,55 doğruluk, PD-CO için %100 doğruluk ve HD-CO için %100 sınıflandırma doğruluğu elde edilmiştir. Nair vd. [20] tarafından yapılan çalışmada PH'yi MSA'dan ayıran ayırıcı tanı için 3-T MR görüntüleme tekniği kullanılmıştır. Görüntülerin değerlendirilmesi ve bu verilerin etiketlenmesi iki bağımsız nöroradyolojist tarafından gerçekleştirilmiştir. Son olarak bir karar ağacı modeli oluşturularak PH'nin ve MSA'dan ayrılması sağlanmıştır. Ota vd. [21] tarafından yapılan çalışmada PH'nin, MSA-P ve MSA-C'den ayırıcı tanısı üzerine bir çalışma gerçekleştirilmiştir. Çalışmada bireylerden T1 ağırlıklı veriler ve Yayınımsal MRI verileri elde edilmiştir.

Daha sonra sırasıyla Voxel-based morphometry (VBM) ve Tract-based spatial yöntemler kullanılarak veriler analiz edilmiştir. Drotár vd. tarafından yapılan çalışmada PD'li hastalarda ilk görülen özelliklerden olan el yazısının bozulmasının analiz edilmesi ile PD'li hastaların sağlıklı kontrol bireylerinden ayrılabilceği düşünülmüştür. PD'nin tespit edilmesinde el yazısının kinematikindeki değişim kullanılmıştır. Bu açıdan öncelikle el yazısı özellikleri çıkarılmıştır. Özelliklerin çıkarılmasından sonra veri kümesi oluşturulmuş ve Destek Yöney Makinası (DYM) ile sınıfların birbirlerinden ayrılması sağlanmıştır [22]. Akdemir vd. [23] tarafından yapılan çalışmada bazı parkinsonizmler'in (PH, Multi Sistem Atrofi (MSA), Progresif Supranükleer Paralizi (PSP), Kortikobazal Dejenerasyon (KBD), Lewy cisimcikli demans (LCD) ve kontrol bireyleri) ayırıcı tanısı için Brain 18F-FDG PET görüntüleme kullanılmıştır. Görüntülerin analizi görsel olarak ve NeuroQ yazılımı kullanılarak gerçekleştirilmiştir. Lee vd. [24] tarafından yapılan çalışmada PD'nin PSP ve MSA'dan ayrılması için yürüme ve duruş özellikleri kullanılmıştır. Bunun için PD, PSP ve MSA hastalarının videoları analiz edilmiştir. Pilleri vd. [25] tarafından yapılan çalışmada PD ve MSA'nın ayırıcı bir tanısı için kalp atım hızı ve kan basıncının sirkadiyen ritim özellikleri kullanılmıştır. İki hastalık grubunun karşılaştırılmasında t-testi ve ANCOVA testi yapılmıştır. Navarro-Otano vd. [26] tarafından yapılan çalışmada PD ve Vasküler Parkinsonizm (VP) arasındaki ayırıcı tanı için 123I-meta-iodobenzylguanidine (123I-MIBG) kalp alımı, koku tanıma ve 123I-FP-CIT SPECT analizleri yapılmıştır. Bu analizler sonucunda elde edilen verilerin değerlendirilmesinde istatistiksel testler kullanılmıştır. Salvatore vd. [27] tarafından yapılan çalışmada PD ve PSP arasındaki ayırıcı tanı için T1 ağırlıklı MRI verileri kullanılmıştır. Özellik seçiminde Temel Bileşen Analizi kullanılmıştır. Hastalık gruplarının sınıflandırılmasında ise DYM kullanılmıştır. Feng vd. [28] tarafından yapılan çalışmada PD ve MSA arasındaki ayırıcı tanı için 3.0T MRI verileri kullanılmıştır. Bu hastalık gruplarının karşılaştırılması için istatistiksel testler kullanılmıştır. Baudrexel vd. tarafından yapılan çalışmada PD, MSA ve PSP arasındaki ayırıcı tanı için 18-Fluorodeoxyglucose Positron Emission Tomography (FDG-

PET) ve Difüzyon Ağırlıklı Görüntüleme (DAG) yöntemleri karşılaştırılmış ve kullanılmıştır. Bu hastalık gruplarının sınıflandırılmasında ANOVA ve ROC analizi kullanılmıştır [29]. Huertas-Fernández vd. tarafından yapılan çalışmada PH ve VP arasındaki ayırıcı tanı için [123I]FP-CIT SPECT yöntemi kullanılmıştır. Verilerin sınıflandırılmasında Lojistik Regresyon, Doğrusal Ayırıcı Analiz ve DYM kullanılmıştır [30]. Zanigni vd. [31] tarafından yapılan çalışmada PH ve diğer parkinsonyen sendromlar (MSA-C, MSA-P ve Richardson sendromu (PSP-RS)) arasındaki ayırıcı tanı için MR Proton Spektroskopisi (1H-MRS) yöntemi kullanılmıştır. Diğer çalışmalardan farklı olarak bu çalışmada beyincik üzerine odaklanılmıştır. Çalışmada elde edilen verilerin analizinde istatistiksel analiz kullanılmıştır.

Bradvice vd. [32] tarafından yapılan çalışmada PD ve Esansiyel Tremor (ET) arasındaki ayırıcı tanı için Transrakniyal Sonografi ve koku testi kullanılmıştır. Ayrıca Transrakniyal Sonografi sonuçları Dopamin Transporter Scan (DaTSCAN) sonuçları ile de karşılaştırılmış ve yüksek derecede uyumluluk bulunmuştur. Verilerin analizi istatistiksel analiz kullanılarak yapılmıştır. Vranová vd. [33] tarafından yapılan çalışmada nörodejeneratif hastalıkların (PD, LCD, Alzheimer hastalığı, MSA, PSP) ayırıcı tanısı için Clusterin proteininin beyin omirilik sıvısındaki seviyesi kullanılmıştır. Verilerin karşılaştırılmasında istatistiksel analiz kullanılmıştır. Drotár vd. [34] tarafından yapılan çalışmada PD'yi normal bireylerden ayıran ayırıcı bir tanı sistemi için el yazısının kinematik ve basınç özellikleri kullanılmıştır. Bu özelliklerden elde edilen verilerin sınıflandırılmasında k-en yakın komşuluk algoritması, AdaBoost ve DYM kullanılmıştır. Aydın ve Aslan tarafından yapılan çalışmada yere uygulanan kuvvet sinyalleri kullanılarak nöro-dejeneratif hastalıkların (ALS, HD, PD) teşhisi ve sınıflandırılması gerçekleştirilmiştir. Sınıflandırmada beş Yapay Öğrenme algoritması (A2DE, K*, Multilayer Perceptron, DECORATE, Random Forest) 10-katlı çapraz doğrulama yöntemi kullanılarak karşılaştırılmıştır. Öznitelik çıkarımı için ise Temel Bileşen Analizi yöntemi kullanılmıştır [35]. Bu makalede diğer çalışmalardan farklı olarak özellik oluşturmak için tepe analizine dayalı tanımlayıcı veri özetleme yöntemi kullanıldı. Ayrıca bu çalışmada öznitelik seçimi için ilk defa One Rules (OneR) algoritması kullanıldı. Böylelikle sınıflandırma sırasında en fazla bilgi veren özniteliklerin seçilmesi gerçekleştirildi. Sınıflama işlemi boyunca en iyi doğruluk oranını veren algoritmalar ise RBFNetwork, AdaBoost ve LogitBoost olarak belirlendi.

3. VERİ TOPLAMA (DATA COLLECTION)

Bu araştırma çalışmasında kullanılan veriler Hausdorff ve diğerlerinin, 64 deneğin (15 denek PD, 20 denek HD, 13 denek ALS ve 16 denek de sağlıklı birey) sağ ve sol ayaklarından elde ettikleri işlenmemiş kuvvet verileri olup PhysioNET veritabanından [36] alınmışlardır. Bu yürüme işaretleri deneklerden 300 Hz de örneklenmiş olup her denek için 5'er dakikalık verilerdir. Bu verilerin elde edilmesinde kullanılan düzenek Hausdorff vd. [37] tarafından

geliştirilmiştir . Bu araştırmada kullanılan yürüme analizi 3D Yürüme Analizi (3D Gait Analysis)'nin ölçümlerinden biri olan kinetik analizdir. Kinetik analizde ölçülebilen veri YTKV'dir. YTKV, vücudun yere uyguladığı bileşke kuvveti ölçen ve basınca duyarlı plakalardan oluşan kuvvet platformu ile ölçülür.

4. DALGACIK DÖNÜŞÜMÜ (WAVELET TRANSFORM)

Dalgacık dönüşümü, gürültü temizleme ve sıkıştırma gibi sinyal işleme uygulamalarında başarılı sonuçlar veren bir sinyal dönüşüm tekniğidir [38]. Wavelet dönüşümünün taban fonksiyonlarına dalgacık denilmekte olup, dalgacıklar bir ölçek ve dalgacık işlevi ile tanımlanmaktadır. Dalgacık dönüşümü, Fourier dönüşümünün durağan olmayan sinyaller üzerindeki eksikliklerini gidermek için geliştirilmiş bir dönüşüm yöntemidir [39]. Bu yöntem; düşük frekanslardaki sinyaller için geniş, yüksek frekanslardaki sinyaller için dar olacak şekilde değişebilen pencere boyutlarına sahip olduğundan bütün frekans aralıklarında uygun zaman-frekans çözünürlüğü sağlamaktadır.

Bu durum Procházka [40] ve Morlet'in [41] çalışmalarında gösterilmiştir. Dalgacık dönüşümünde, Fourier dönüşümünde kullanılan sinüzoidal fonksiyonlar değil de dalgacık fonksiyonları kullanıldığından keskin süreksizlikleri içeren sinyaller için uygun olmaktadır. Bu durum Mallat [42] ve Miner'in [43] çalışmalarında gösterilmiştir. Dalgacık dönüşümü sinyalin tüm frekans-zaman bilgisini korur. Bu sebeptendir ki durağan olmayan sinyallerin işlenmesi geleneksel metotlardan daha iyi sonuç sağlar. Dalgacığın en önemli avantajlarından biri, büyük sinyallerin yerel analizine imkan sağlamasıdır. Ayrıca süreksiz sinüzoidal bir sinyalin Fourier dönüşümünde, sinyalin uç noktalarında tepe değerleri oluşurken Dalgacık dönüşümü ile süreksizlik civarındaki değişim de gerçek zamanlı olarak gözlemlenebilmektedir [44]. Sürekli dalgacık dönüşümünün formülü Eş. 1'de gösterilmektedir. Bu eşitlikte $f(t)$ sinyal ya da fonksiyon, a ölçekleme parametresi, $\Psi(t)$ ifadesi temel dalgacık ve τ ise zaman eksenini boyunca dalgacık fonksiyonunun dönüşümüdür [45].

$$T(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \Psi\left(\frac{t-\tau}{a}\right) dt \quad (1)$$

5. YAPAY ÖĞRENME (MACHINE LEARNING)

Yapay Öğrenme bir dizi örnekten bir görev öğrenen mantıksal ya da ikili operasyonlara dayalı otomatik hesaplama yordamlarını kapsayan bir çalışma alanıdır [46]. Yapay Öğrenme bu açıdan yapay zeka ile ilgili görevleri yerine getiren sistemlerdeki değişimle ilgilenir [47]. Bu nedenle yapay zekanın bir dalıdır. Bilgi keşfi esnasında Yapay Öğrenme, bilgi keşfi sürecindeki adımlardan biri olan çıkarımsal algoritmaların uygulanmasını gerçekleştirmek için çok yaygın bir şekilde kullanılır [48]. Çıkarımsal Yapay Öğrenme algoritmaları etiketli veriden yani bilinen bir çıkışa sahip veriden örüntüleri öğrenebilir [49]. Yapay öğrenme alanında, tüm gerçek dünya problemlerine ait eğitim verileri

üzerinde iyi bir öğrenme sağlayan tek bir öğrenme algoritması yoktur. Bundan dolayı algoritma seçimi deneysel yolla yapılır [50]. Aynı zamanda sınıflandırıcıların eğitim verisi üzerinde oluşturdukları model eğitim verisine göre değişmektedir. Bu durum en iyi modelin seçimini gündeme getirmektedir. Ayrıca en iyi öğrenme algoritmasının olmadığını, her algoritmanın çok başarılı olduğu veriler olduğu gibi çok başarısız olduğu veriler de vardır. Bu durum "no free lunch" teoremi [51] olarak bilinmektedir. "no free lunch" teoremi Wolpert tarafından diğer optimizasyon algoritmalarına da uyarlanmıştır [52]. Bu çalışmadaki sınıflandırma problemi WEKA içerisinde yer alan algoritmalar kullanılarak çözülmüştür. WEKA ile verileri işleyebilir, işlenmiş verilerin sonuçlarını analiz edebilir ve hiç kod yazmadan WEKA'nın işlevselliğinden yararlanabiliriz [53]. Bu çalışmadaki sınıflandırma probleminde kullanılan algoritmalar aşağıda tanıtılmaktadır.

5.1. Radyal Tabanlı Fonksiyon Ağı (Radial Basis Function Network)

Radyal tabanlı fonksiyon ağı (RBFNetwork), aktivasyon fonksiyonu olarak radyal tabanlı fonksiyonlar kullanan bir yapay sinir ağıdır. RBFNetwork kavramsal açıdan kNN (k-en yakın komşuluk algoritması) ve k-Means algoritmalarına benzemektedir. RBFNetwork sınıflandırıcısında temel düşünce, bir test verisinin eğitim kümesi içerisindeki diğer verilere olan uzaklığını hesaplamaktır. Bu yönü ile kNN sınıflandırıcıya benzemektedir. RBFNetwork tipik olarak üç katmandan oluşmaktadır. Bunlar: giriş katmanı, gizli katman ve çıkış katmanıdır. Gizli katman non-linear aktivasyon fonksiyonunu içermektedir. Çıkış katmanı ise doğrusal bir yapıdadır. RBFNetwork sınıflandırıcı ileri beslemeli bir yapıya sahiptir. RBFNetwork kaotik zaman serileri modellemede Moody [54] ve Casdagli [55] tarafından kullanılmış ve başarılı sonuçlar vermiştir.

Ayrıca RBFNetwork algoritması Broomhead [56] ve Matej [57] tarafından interpolasyon uygulamaları için kullanılmış ve başarılı sonuçlar vermiştir. RBFNetwork sınıflandırıcı da eğitim ve öğrenme aşaması çok hızlıdır. Ayrıca RBFNetwork'ün performansını radyal tabanlı fonksiyonun merkez lokasyonu etkilemektedir [58]. Weka'da yer alan RBFNetwork, aktivasyon fonksiyonu olarak normalize edilmiş Gauss radyal tabanlı fonksiyonu kullanır. Gauss radyal tabanlı aktivasyon fonksiyonu Eş. 2 de gösterilmektedir.

$$\phi(X) = \exp\left(-\frac{X^2}{2\sigma^2}\right) \quad (2)$$

Bu eşitlikte; X ifadesi giriş yöneyini tanımlarken, σ ifadesi giriş yöneylerinin standart sapmasını belirtmektedir. Ayrıca diğer temel fonksiyonlar için k-Means kümeleme algoritmasını kullanmaktadır. Öğrenme esnasında, ayrık sınıf problemleri için lojistik regresyon'u kullanırken, nümerik sınıf problemleri için doğrusal regresyon'u kullanır.

5.2. Adaptif Yükseltme (Adaptive Boosting)

Adaptif Yükseltme (AdaBoost) algoritması bir meta öğrenme algoritması ve aynı zamanda bir boosting algoritmasıdır. Bu tür algoritmalar bir yapay öğrenme yöntemi olan topluluk öğrenme yöntemini (ensemble method) kullanırlar. Bu yöntemdeki amaç, güçsüz olan her bir öğrencinin bir araya getirilerek daha iyi bir öğrencinin oluşturulmasıdır. AdaBoost algoritması şu şekilde çalışmaktadır: Algoritma, X uzayından çekilen x_i 'leri ve bu x_i 'lerle ilişkilendirilen Y etiket uzayının bir elemanı olan y_i ikililerinin kümesi olan ve m örnekten oluşan bir $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ eğitim kümesini giriş olarak alır. Bunun yanısıra AdaBoost algoritması zayıf öğrenci olarak kendisine bir öğrenci seçer. Bu öğrenci t defa çalıştırılır. Zayıf öğrenci ilk defa çalışırken S eğitim kümesinin tüm örneklerinin ağırlığını eşit olarak alır. Örneklerin ağırlıkları, zayıf öğrencinin her çalışmasında güncellenir. Bu güncelleme işlemi zayıf öğrencinin oluşturmuş olduğu hipotezlerin yapmış olduğu yanlış sınıflandırmalara göre yapılır. Yani bir önceki hipotezde yanlış olarak sınıflandırılan örneklerin ağırlıkları artırılır ve bir sonraki hipotezin oluşturulması için gerekli olan eğitim kümesinde bu yanlış örneklerin yer alma olasılıkları artırılır. Böylelikle bir eğitim kümesi içerisinde aynı örnekten birden fazla yer alabilmektedir. Sonuç olarak bu işlemlerin sonucunda t adet hipotez oluşturulmuş olur. Bu hipotezler daha sonrasında birleştirilerek tek bir hipotez elde edilir [59].

5.3. Eklemeli Lojistik Regresyon (Additive Logistic Regression)

Eklemeli Lojistik Regresyon (LogitBoost) algoritması bir başka boosting algoritmasıdır. Bu algoritma AdaBoost algoritmasının sorun yaşadığı çok gürültülü verilerden kaynaklı aşırı uygunluk (overfitting) problemini çözmek için geliştirilmiştir [60]. LogitBoost bu sorunu çözmek için eğitim hatasını doğrusal bir biçimde azaltmaktadır [61]. Bu yüzden daha iyi bir genelleme sağlamaktadır. Kısaca AdaBoost, Eş. 3 te görüldüğü gibi exponansiyel kaybı azaltırken, LogitBoost Eş. 4 te belirtilen lojistik kaybı azaltmaktadır.

$$J(F) = E[e^{-yF(x)}] \quad (3)$$

$$J(F) = E[-\log(1 + e^{yF(x)})] \quad (4)$$

Sonuç olarak LogitBoost lojistik kayıp fonksiyonunu kullanarak aşırı uygunluk sorununa neden olan verilerin ağırlığını daha da arttırmaktadır [62]. Böylelikle AdaBoost algoritmasına göre aşırı uygunluk sorununun üstesinden daha iyi gelmektedir.

5.4. Tek Kural (One Rule)

Yapay Öğrenme sistemleri tarafından üretilen sınıflandırma kuralları iki kriter bakımından değerlendirilir. Bu kriterler: Bağımsız bir test verisi üzerinde onların sınıflandırma doğruluğu ve onların komplekslikleridir [63]. Bazı sınıflandırıcıların birçok veri kümesi üzerinde şartıcı bir

biçimde yüksek doğruluk verdiği görülmektedir. Rendell ve Seshu [64] tarafından yapılan çalışmada zor konseptlerin içsel bir doğruluğa sahip olduğu ancak geleneksel yöntemlerle öğrenilemeyeceği gösterilmiştir. Shavlik ve Mooney [65] tarafından yapılan çalışmada geri beslemeli bir sinir ağı, ID3 ve algılayıcı öğrenme algoritmalarının çeşitli veri kümeleri üzerinde başarılı sonuçlar verdiği gösterilmiştir. Buntine ve Niblett [66] tarafından yapılan çalışmada karar ağaçlarındaki ayırma kuralları değiştirilerek veri kümeleri üzerindeki başarımlar incelenmiştir. Clark ve Niblett [67] tarafından yapılan çalışmada yeni bir kural öğrenme algoritması geliştirilmiş ve çeşitli veri kümeleri üzerinde başarılı sonuçlar vermiştir. Ayrıca Mingers [68] karar ağaçları için beş budama yöntemi kullanarak çeşitli veri kümeleri üzerinde bunları karşılaştırmıştır. Sonuçta yeni geliştirilen ya da düzenlenen tüm bu öğrenme algoritmaları çeşitli veri kümeleri üzerinde başarımlar göstermektedir. Weka'da One Rules (OneR) sınıflandırıcı olarak isimlendirilen "1-rules" sınıflandırıcı veri kümeleri üzerinde çok basit kurallar bulmayı amaçlayan bir sınıflandırıcıdır. OneR sınıflandırıcı tek bir niteliğe dayalı bir nesneyi sınıflandıran kurallardan oluşur. Bu açıdan tek seviyeli bir karar ağacına benzemektedir. Bunun dışında OneR sınıflandırıcı çok kompleks hipotezler üreten sınıflandırıcılarla rekabet edebilecek tahminler yapabilmektedir [63]. Weka'da ise OneR sınıflandırıcı kuralları en düşük hata oranına göre değil en yüksek doğru örneğe göre oluşturur. Aynı zamanda hata oranı aynı olduğunda bir kuralı rastgele seçmez [69].

5.5 Sıfır Kural (Zero Rules)

"0-rules" olarak bilinen sınıflandırıcı Weka'da ZeroR sınıflandırıcı olarak isimlendirilmektedir. Aynı zamanda çoğunluk sınıflandırıcı olarak da bilinmektedir. Temel sınıflandırıcı olarak kullanılır. Daha çok diğer sınıflandırıcıların performansını ölçmek için kullanılır. Yani diğer tüm sınıflandırıcıların doğruluk oranının ZeroR sınıflandırıcısının üzerinde olması sınıflandırıcıların performans kriteri açısından çok önemlidir. ZeroR sınıflandırıcı şu şekilde çalışmaktadır: Eğitim verileri içerisinde frekansı en çok olan sınıf tüm veriler için çıkış değeri olarak kabul edilir. Böylelikle frekansı yüksek olan sınıf'ın oranı sınıflandırıcının doğruluk oranını verecektir.

6. DENEYLERİN ANALİZİNDE KULLANILAN İSTATİSTİKSEL KRİTERLER (STATISTICAL CRITERIA USED FOR THE ANALYSIS OF EXPERIMENTS)

Bir sınıflandırıcının, sınıflandırma işlemi sonunda ne kadar performanslı olduğunu değerlendirmek için çeşitli ölçütler kullanılmalıdır. Bu çalışmada da sınıflandırma işleminin başarımını değerlendirmek için kullanılacak ölçütler şunlardır: Kappa İstatistiği, F-Ölçütü değeri, Ortalama Karesel Hatanın Karekökü ve bias-varyans ikilemi'dir. Kappa değeri bir grup veri üzerinde tahmin edilen ve gözlenen sınıflandırmalar arasındaki uyumu ölçmek için kullanılan bir araçtır [53]. Kappa değeri ilk olarak Cohen [70] tarafından ortaya konuldu [53]. Cohen'e göre

Kappa değeri, iki gözlemci arasındaki uyum derecesini ölçmek için kullanılan bir ölçüttür [70]. Kappa değeri Eş. 5'te gösterilmektedir. P(a) sınıflandırıcının doğruluğunu olasılıksal olarak göstermektedir ve P(e) aynı veri kümesi üzerinde rastgele tahmin yapan sınıflandırıcının beklenen doğruluğu yani yapılan tahminlerin olasılığının ağırlıklı ortalamasıdır. Cohen'in geliştirdiği Kappa değerine yapılan en önemli eleştiri bazen sezi ile beklenenden farklı sonuçlar üretmesidir [71]. Kappa değeri -1 ve 1 arasında değer almaktadır. -1 değeri tam bir uyumsuzluğu yani ters yönlü bir ilişki olduğunu gösterirken, 1 değeri ise mükemmel uyumu belirtmektedir. Kappa değeri 0,4 ve üzeri bir değerde ise şansın ötesinde bir uyumdan bahsedilir [72].

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

Eş. 9'da gösterilen F-Ölçütü değeri, Eş. 6 daki Kesinlik (precision) ve Eş. 7'deki Hassasiyet (recall) değerlerinin harmonik ortalaması alınarak hesaplanır. Kesinlik, pozitif olarak tahmin edilen örnekler içerisinde doğru olarak sınıflandırılan pozitiflerin oranı olarak ifade edilir. Hassasiyet ise gerçek sınıfı pozitif olan örnekler içerisinde doğru olarak sınıflandırılan pozitif örneklerin oranı olarak ifade edilir [73]. Eş. 8'de gösterilen Doğruluk Oranı (accuracy) ise tüm örnekler içerisinde doğru olarak sınıflandırılmış örnek sayısını vermektedir. İyi öğrenilen bir sınıf için F-ölçütü değeri genellikle 0,5 ve yukarısı olarak alınır.

$$\text{Kesinlik} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (6)$$

$$\text{Hassasiyet} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (7)$$

$$\text{Doğruluk Oranı} = \frac{\text{True Positive} + \text{True Negative}}{\text{Number of Instances}} \quad (8)$$

$$F\text{-Ölçütü} = \frac{2 * \text{Kesinlik} * \text{Hassasiyet}}{\text{Kesinlik} + \text{Hassasiyet}} \quad (9)$$

Ortalama Karesel Hatanın Karekökü (Root Mean Squared Error (RMSE)) Ortalama Karesel Hata (Mean Squared Error (MSE)) değerinin karekökünün alınmasıyla hesaplanır. Bir tahmin edicinin hata oranı rastgele bir tahmin yapmasından ya da doğru tahmin yapmasını sağlayacak bilgiye sahip olmamasındandır [74]. MSE ve RMSE değeri sıfıra yaklaştıkça sınıflandırıcının doğru tahmini de artmaktadır. Ayrıca her gerçek dünya problemi için MSE / RMSE'nin kabul edilebilir değeri aynı değildir. MSE ve RMSE'nin denklemleri sırasıyla Eş. 10 ve Eş. 11'de yer almaktadır.

Buna göre p, yapılan tahmini değerleri; a ise gerçek değerleri ifade etmektedir.

$$MSE = \frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n} \quad (10)$$

$$RMSE = \sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}} \quad (11)$$

Bias-Varyans ikilemi Yapay Öğrenme algoritmalarını anlamada önemli bir yöntemdir. Son dönemde yapılan çalışmalarda Bias-Varyans ikileminin kullanımı önemli bir biçimde artmaktadır [53]. Bias / Varyans ikilemi çok naif sınıflandırıcıların çok karmaşık sınıflandırıcılar karşısında ne kadar performanslı olabildiklerini ve model toplulukların da naif modeller karşısında ne kadar performanslı olabildiklerini açıklar. Bias-Varyans ikileminin hata fonksiyonu içerisindeki konumlanışı, 0-1 hata fonksiyonunun yanı sıra karesel kayıplar için de gösterilmiştir [75]. Bu gösterim Geman, Bienenstock ve Doursat [76] tarafından yapılmıştır. Her bir örneğin gerçek değeri a ve tahmini değeri p olmak üzere; ortalama karesel hata'nın bias-varyans ayrışımı olarak gösterimi Eş. 12'de gösterilmektedir. Gürültünün ihmal edildiği durumlarda ortalama karesel hata Eş. 13'teki eşitliğe dönüşür.

$$E[(p_i - a_i)^2] = \text{Varyans}(\text{Gürültü}) + \text{Bias}^2 \text{Varyans}(p_i) \quad (12)$$

$$MSE = \text{Varyans}(p) + \text{Bias}^2 \quad (13)$$

Yapılan deneylerin sonuçlarının doğruluğunu garanti altına almak için deneylerde bazı metodolojiler uygulanmıştır. Bu metodoloji k-katlı çapraz doğrulama (k-fold cross validation) metodu olarak bilinmektedir. Eğitim kümesinin bir sınıflandırıcı tarafından öğrenilmesi için yeterli olup olmadığını belirlemek için de k-katlı çapraz doğrulama (k-fold CV) yöntemi kullanılmaktadır. k-fold CV, verinin hemen hemen k eşit parçaya bölünmesi esasına dayanır. Böylelikle sınıflandırıcı k adet eğitilir ve test edilir. Sınıflandırıcı, her zaman veri üzerinde bir eksik katla eğitilir [48]. k-fold CV sınırlı veriye sahip iki sınıflandırıcıyı kıyaslamak için de tercih edilir [77]. Ayrıca WEKA aracı, k-fold CV yöntemi sırasında eğitim verisi ile test verisine dağılmış örneklerin sınıfları arasındaki dengesizliği önlemek için "stratification" süreci uygular. Böylelikle istatistiksel açıdan güvenli bir sonuç elde edilmiş olur [78]. k değerinin bu çalışmada 10 alınmasının nedeni ise, farklı sınıflandırıcılar kullanılarak yapılan testlerde beklenen doğrulukta bir tahminin yapmak için doğru sayının yaklaşık olarak 10 olduğu görülmektedir. Aynı zamanda yapılan bazı teorik çalışmalarda k değerinin 10 olması gerektiğini belirtmektedir [53].

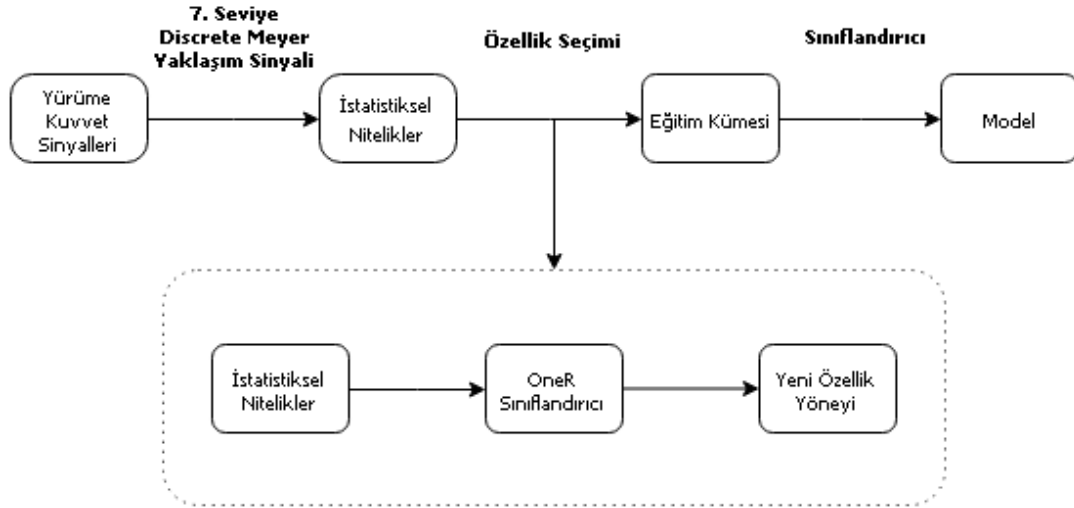
7. EĞİTİM VERİSİNİN HAZIRLANMASI (PREPARATION OF TRAINING DATA)

Bu çalışma da öncelikle yürüme kuvvet sinyalleri, Discrete Meyer (dmey) dalgacığı kullanılarak yedinci seviyeye kadar ayrıştırıldı. Yeni oluşan sinyallerden yedinci seviyedeki yaklaşım sinyali seçildi. Bu sinyal üzerinde peak analizi gerçekleştirilerek sinyalin lokal maksimumları, peak'in x değerleri, peak genişliği ve peak çıkıntıları elde edildi. Daha sonra bu dört peak özelliğinin her birinden 15 adet istatistiksel öznitelik elde edildi. Böylelikle sol ayak için 60 ve sağ ayak için 60 olmak üzere toplamda 120 öznitelik elde edildi.

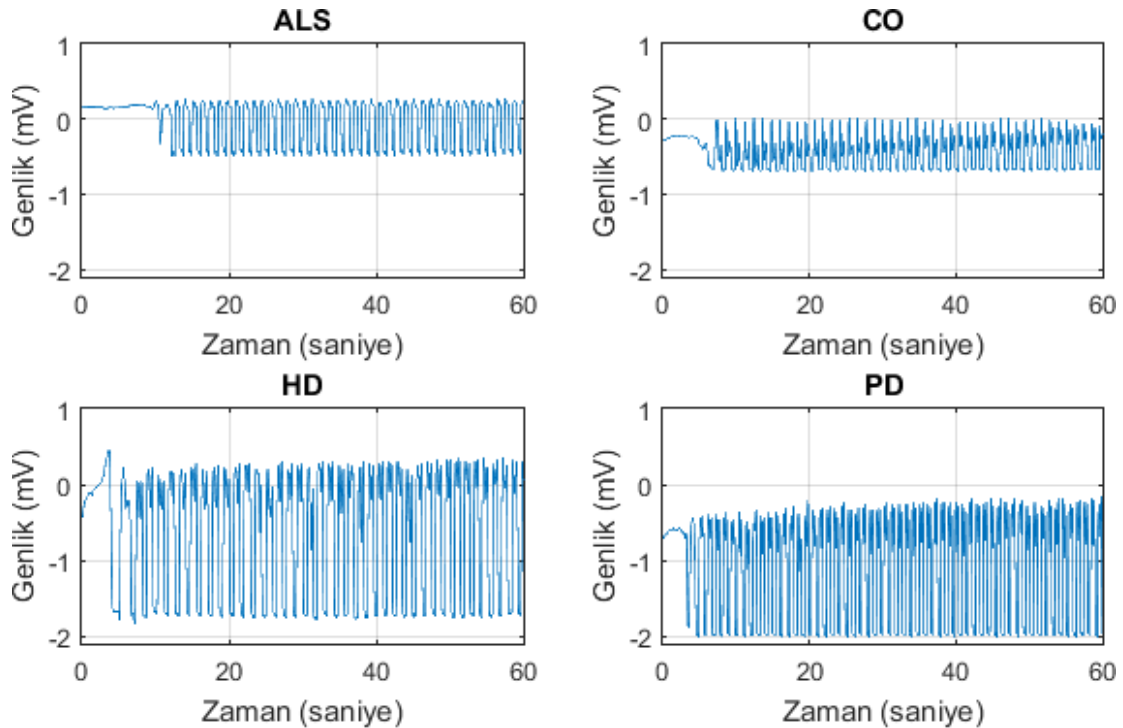
Bu istatistiksel öznitelikler: maksimum, minimum, ortalama, standart sapma, basıklık, çarpıklık, ortanca değer, çeyreklikler arası aralık, ortalamadan mutlak sapma, ortancadan mutlak sapma, toplam, shannon entropisi, logaritmik entropi, vektörün birinci normu ve vektörün ikinci normudur. Daha sonra OneRules sınıflandırıcı kullanılarak bu öznitelikler içerisinde en çok enformasyon veren öznitelikler seçildi. Daha sonra seçilen özniteliklere göre eğitim kümesi oluşturuldu. Bu süreç Şekil 1'de gösterilmektedir.

8. SONUÇLAR VE TARTIŞMALAR (RESULTS AND DISCUSSIONS)

Deneyler yapılırken bazı düzenekler sabit seçildi. Bu deney düzenekleri şunlardır: Deneyler aksi belirtilmediği sürece 10-fold CV tekniği kullanılarak gerçekleştirildi. Ayrıca aksi belirtilmediği sürece tüm sınıflandırıcılar varsayılan parametrelerle çalıştırıldı. Nöro-dejeneratif hastalar ve kontrol bireyine ait sol ve sağ ayak kuvvet sinyalleri sırasıyla Şekil 2 ve Şekil 3'te gösterilmektedir.



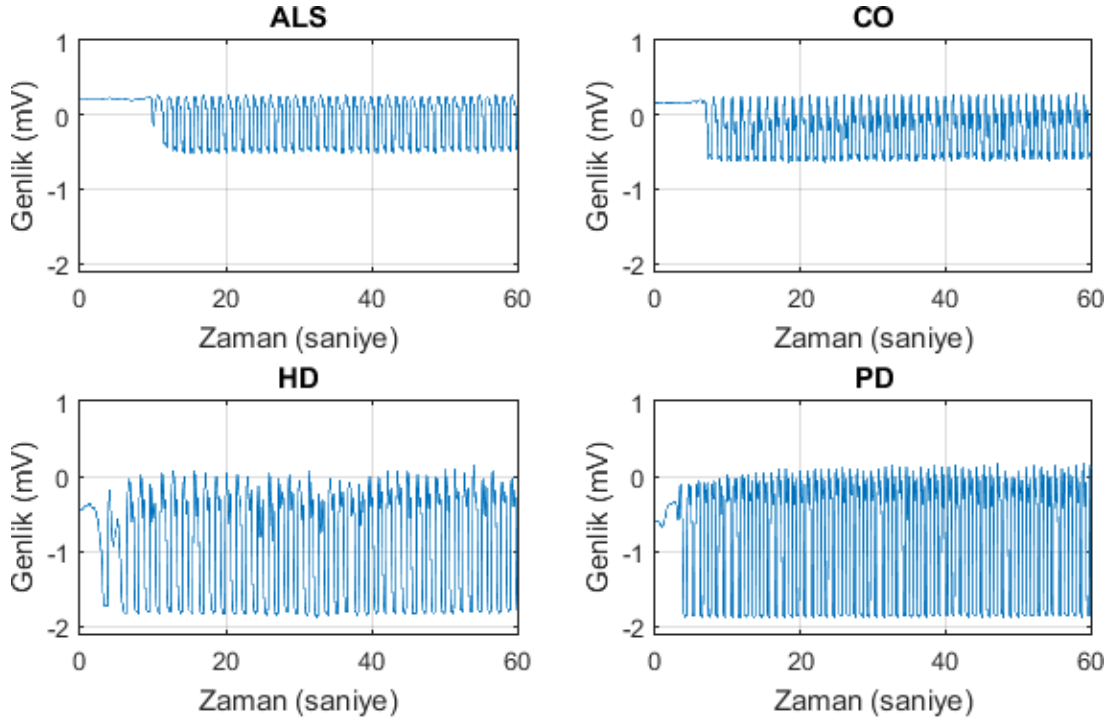
Şekil 1. Yürüme sinyallerinin işlenmesinden modelin oluşturulmasına kadar geçen süreç
(Elapsed process from creation of the model to processing of gait signal)



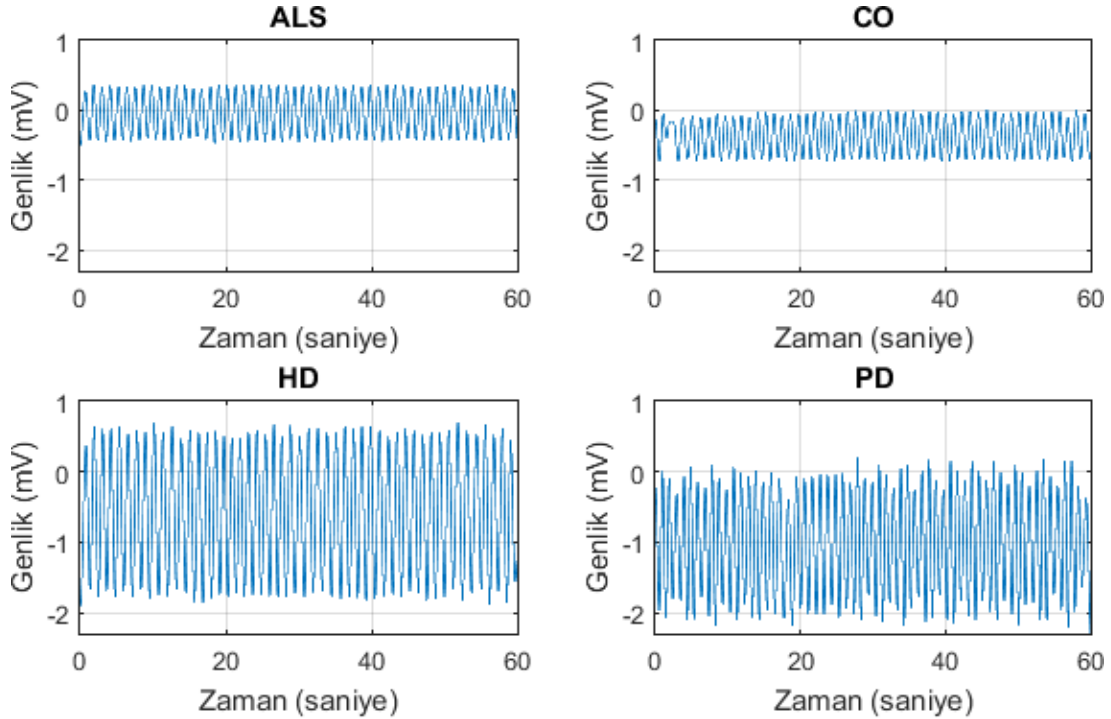
Şekil 2. ALS, PD, HD hastası ve CO bireyi için sol ayak kuvvet-duyarlı direnç verisi
(The left foot force-sensitive resistor data for one ALS patient, PD patient, HD patient and CO subject)

Her hastadan elde edilen 5 dakikalık sinyalin sadece ilk 1 dakikalık kısmı Şekil 2 ve Şekil 3'te gösterilmektedir. Bu sinyallerin herbiri için Discrete Meyer dalgacı uygulanarak 7. seviyeye kadar dalgacık dönüşümü gerçekleştirildi. Tüm bu süreçte ilk olarak sinyal

gürültülerden temizlendi. Sonrasında Discrete Meyer dalgacı kullanılarak 7. seviyeye kadar sinyal ayrıştırıldı. Son olarakta 7. seviye yaklaşım sinyali kullanılarak sinyal tekrar oluşturuldu. Tekrar oluşturulan bu sinyaller sol ve sağ ayak için sırasıyla Şekil 4 ve Şekil 5'te gösterilmektedir. Sol



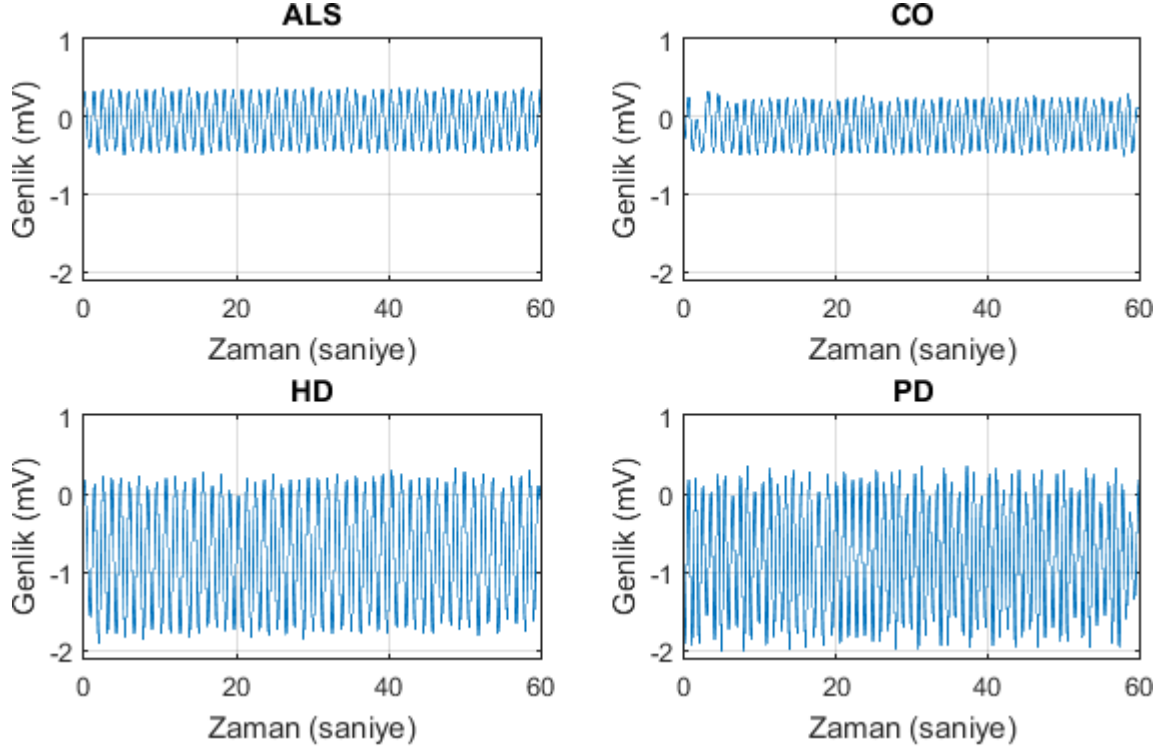
Şekil 3. ALS, PD, HD hastası ve CO bireyi için sağ ayak kuvvet-duyarlı direnç verisi
(The right foot force-sensitive resistor data for ALS patient, PD patient, HD patient and CO subject)



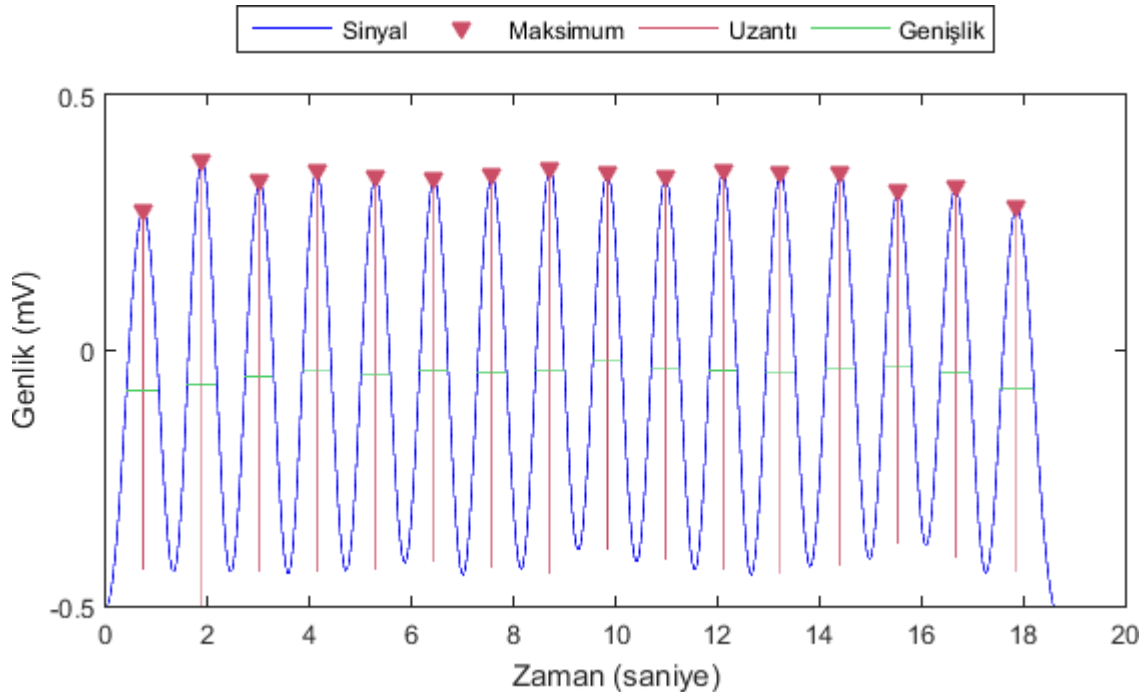
Şekil 4. Sol ayak için Discrete Meyer dalgacı kullanılarak 7. seviye yaklaşım sinyali ile oluşturulan yeni sinyal
(Reconstructed approximation signal at level-7 using Discrete Meyer wavelet for the left foot)

ve sağ ayak için oluşturulan bu yeni sinyaller üzerinde tepe analizi (peak analysis) yapıldı. Tepe analizi sonucunda 4 yöney elde edildi. Bu yöneyler sırasıyla yerel maksimumlar (local maxima), tepe konumları (peak locations), tepe genişlikleri (peak widths) ve tepe uzantıları (peak prominences)'dir. (Şekil 6.) Bu 4 yöneyin her birine 15 istatistiksel fonksiyon uygulandı. Böylelikle sol ayak için 60

ve sağ ayak için 60 olmak üzere toplamda 120 istatistiksel öznelik elde edildi. Bu 4 yöney üzerine uygulanan istatistiksel fonksiyonlar Tablo 1'de gösterilmektedir. Bu tabloda sol ve sağ ayaklar için kullanılan öznelik kodları aynıdır. Bu çalışmada herbir tepe yöneyi için 15 adet istatistiksel özellik kullanılmıştır. Bu özelliklerin oluşturulma nedeni ise verinin tümünü ifade edebilecek



Şekil 5. Sağ ayak için Discrete Meyer dalgacığı kullanılarak 7. seviye yaklaşım sinyali ile oluşturulan yeni sinyal (Reconstructed approximation signal at level-7 using Discrete Meyer wavelet for the right foot)



Şekil 6. Yeni sinyalin tepe analizi (Peak analysis of reconstructed approximation signal)

Tablo 1. Yaklaşım sinyallerinden çıkarılan istatistiksel öznitelikler ve bu özniteliklerin MATLAB formülleri/fonksiyonları
(The statistical features created from the gait signals and MATLAB formulas/functions of these features)

İstatistiksel Fonksiyon	Sol/Sağ Ayak Yöneyleleri				Formül
	Maks.	Konum	Genişlik	Uzantı	
Maksimum	F1	F16	F31	F46	max(x)
Minimum	F2	F17	F32	F47	min(x)
Ortalama	F3	F18	F33	F48	$\frac{1}{N} \sum_{i=1}^N x_i$
Standart Sapma	F4	F19	F34	F49	$\sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2}$
Çarpıklık	F5	F20	F35	F50	$\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}\right)^3}$
Basıklık	F6	F21	F36	F51	$\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2\right)^2}$
Ortanca	F7	F22	F37	F52	median(x)
Aralık	F8	F23	F38	F53	max(x) - min(x)
Ortalamadan Mutlak Sapma	F9	F24	F39	F54	mad(x, 0);
Ortancadan Mutlak Sapma	F10	F25	F40	F55	mad(x, 1)
Toplam	F11	F26	F41	F56	sum(x)
Shannon entropisi	F12	F27	F42	F57	wentropy(x,'shannon')
Logaritmik entropi	F13	F28	F43	F58	wentropy(x,'log energy')
Yöney 1. normu	F14	F29	F44	F59	norm(x, 1)
Yöney 2. normu	F15	F30	F45	F60	norm(x)

özetler elde edilebilir. Bu özetleme (Tanımlayıcı Veri Özetleme) teknikleri, verinin tipik özelliklerini belirlemek için kullanılır. Bu tekniklerin yapmış oldukları ölçümler üçe ayrılır. Bunlar: Dağıtım ölçümü (distributive measure), cebirsel ölçüm (algebraic measure) ve bütünsel ölçüm (holistic measure)'dir. Dağıtım ölçümü, verilen bir verikümesini alt parçalara bölerek herbir parçanın ölçümünün ayrı ayrı hesaplanmasını ve daha sonra bu sonuçların birleştirilmesi ilkesine dayanan bir ölçüm tekniğidir. Örneğin toplam (sum), sayma (count), maksimum (max) ve minimum (min) fonksiyonları dağıtım ölçümü göre çalışmaktadır. Cebirsel ölçüm, bir ya da birden daha çok dağıtım ölçümü bir cebirsel fonksiyonun uygulanmasıyla hesaplanan bir ölçümdür.

Örneğin ortalama (average), toplam/sayma şeklinde hesaplandığından dolayı cebirsel bir ölçümdür. Bütünsel ölçüm, bir bütün olarak tüm veri üzerinde yapılan ölçümdür. Yani verinin parçalara ayrılması söz konusu değildir. Bütünsel ölçümü örnek olarak ortanca (median) fonksiyonu verilebilir. İstatistiksel özniteliklerin belirlenmesinden sonra, bu özniteliklere ait değerler

kullanılarak dört adet eğitim kümesi (ALS-CO, HD-CO, PD-CO ve NDD-CO) oluşturuldu. Bu işlemde sonra, bu eğitim kümeleri üzerinde OneR sınıflandırıcı kullanılarak öznitelik seçimi işlemi gerçekleştirildi. Öznitelik seçimi ile birlikte herbir eğitim kümesinin öznitelik yöneyi birbirlerinden farklı oldu. Tablo 2'de eğitim kümelerine göre seçilen öznitelik yöneyleri gösterilmektedir. Öznitelik seçiminde, OneR sınıflandırıcının hem sol hem de sağ ayak için kullanılan öznitelik yöneyleri üzerinde başarılı olması sonucu öznitelikler seçilmektedir. Eğer bir öznitelik sol ayak için yüksek enformasyon verirken sağ ayak için vermiyorsa bu durumun rastlantısal olabileceği ihtimalinden dolayı bu öznitelik seçilmemiştir. Bu nedenle Tablo 2'de yer alan öznitelikler hem sol hem de sağ ayak özniteliklerini temsil etmektedir. Öznitelik seçiminden sonra dalgacık ailesinin seçimi aşamasına geçildi. Tablo 3'te dalgacık ailesine göre eğitim kümeleri üzerinde varsayılan parametreye göre en yüksek doğruluk oranını veren sınıflandırıcılar gösterilmektedir. Bu sonuçlara göre Dmeyer (dmey) dalgacık ailesi eğitim kümeleri üzerinde en yüksek doğruluğu vermiştir. Bu nedenle dalgacık ailesi olarak bu çalışmada dmey dalgacığı kullanılmıştır.

Tablo 2. Eğitim kümelerine göre yeni öznelik yöneyleri (New feature vectors according to training data)

Eğitim Kümesi	Öznelik Yöneyi (Sol/Sağ Ayak için)	Öznelik Sayısı
ALS-CO	F1, F2, F3, F7, F10, F11, F14, F15, F16, F18, F21, F25, F26, F27, F28, F29, F30, F31, F32, F33, F34, F36, F37, F38, F39, F40, F41, F42, F43, F44, F45, F46, F47, F48, F52, F55, F56, F57, F58, F59, F60, F61, F62, F63, F67, F70, F71, F74, F75, F76, F78, F81, F85, F86, F87, F88, F89, F90, F91, F92, F93, F94, F96, F97, F98, F99, F100, F101, F102, F103, F104, F105, F106, F107, F108, F112, F115, F116, F117, F118, F119, F120	82
HD-CO	F4, F6, F8, F9, F10, F17, F21, F26, F28, F29, F31, F33, F34, F37, F38, F39, F40, F45, F46, F47, F49, F50, F52, F53, F54, F55, F56, F57, F58, F59, F60, F64, F66, F68, F69, F70, F77, F81, F86, F88, F89, F91, F93, F94, F97, F98, F99, F100, F105, F106, F107, F109, F110, F112, F113, F114, F115, F116, F117, F118, F119, F120	62
PD-CO	F6, F8, F9, F21, F22, F23, F24, F26, F28, F29, F30, F32, F34, F37, F38, F39, F40, F41, F42, F43, F44, F45, F46, F48, F49, F50, F52, F53, F54, F55, F57, F58, F60, F66, F68, F69, F81, F82, F83, F84, F86, F88, F89, F90, F94, F97, F98, F99, F100, F101, F102, F103, F104, F105, F106, F108, F109, F110, F112, F113, F114, F115, F117, F118, F120	65
NDD-CO	F32, F34, F39, F40, F92, F94, F99, F100	8

Tablo 3. Dalgacık ailesi ve eğitim kümesine göre sınıflandırıcı sonuçları (Classifier results according to wavelet family and training set)

Dalgacık	Eğitim Kümesi	Sınıflandırıcı	Doğruluk Oranı	RMSE	Kappa
DMeyer	ALS-CO	RBFNetwork	%93,1034	0,2486	0,8585
	HD-CO	AdaBoostM1	%91,6667	0,2584	0,8302
	PD-CO	AdaBoostM1	%83,871	0,4124	0,6791
	NDD-CO	LogitBoost	%87,5	0,3135	0,6364
Daubechies	ALS-CO	RandomForest	%93,1034	0,2893	0,8606
	HD-CO	RBFNetwork	%77,7778	0,4549	0,55
	PD-CO	ClassificationViaClustering	%67,7419	0,5477	0,4105
Haar	NDD-CO	RBFNetwork	%78,125	0,4236	0,2632
	ALS-CO	NNge	%93,1034	0,2626	0,8626
	HD-CO	AttributeSelectedClassifier	%83,3333	0,4138	0,6582
Symlets	PD-CO	RandomCommitte	%77,4194	0,4162	0,5507
	NDD-CO	RotationForest	%76,5625	0,4151	0,3023
	ALS-CO	IBk	%93,1034	0,2556	0,8606
Battle-Lemarie	HD-CO	Logistic	%83,3333	0,4156	0,6667
	PD-CO	LWL	%77,4194	0,4377	0,5526
	NDD-CO	MultiBoostAB	%84,375	0,3909	0,6154
Battle-Lemarie	ALS-CO	ADTree	%93,1034	0,2593	0,8606
	HD-CO	LMT	%86,1111	0,3388	0,7205
	PD-CO	ClassificationViaClustering	%67,7419	0,5477	0,4053
	NDD-CO	Bagging	%76,5625	0,4277	0,1892

Tablo 3'teki sonuçlara göre dalgacık ailesi olarak en iyi sonucu dmey dalgacığı verdikten sonra yapılan tüm deneylerde dalgacık ailesi olarak dmey kullanılmıştır. ALS-CO eğitim kümesi üzerinde en iyi sonuçları veren ve birbirlerinden farklı varsayımlarla öğrenme sürecini gerçekleştiren sınıflandırıcılar Tablo 4'te gösterilmektedir. Bu sınıflandırıcılardan ilk ikisi temel sınıflandırıcıdır. Randomize isimli sınıflandırıcı eğitim kümesindeki verileri rastgele bir biçimde sınıflandırarak öğrenme işlemi gerçekleştirmektedir. Randomize ve ZeroR sınıflandırıcıları diğer sınıflandırıcıların performanslarını kıyaslamak için kullanılır. Yani diğer sınıflandırıcıların doğruluk oranlarının temel sınıflandırıcıların doğruluk oranlarından her zaman daha büyük olması istenmektedir. Böylelikle diğer

sınıflandırıcıların performanslı bir öğrenme gerçekleştirdikleri söylenebilir. Tablo 4'teki sonuçlara göre RBFNetwork sınıflandırıcı diğer sınıflandırıcılara göre daha performanslı olduğu görüldü. RBFNetwork; Decorate ve NNge sınıflandırıcıları ile aynı doğruluk oranına sahiptir. Ancak RMSE değerlerine bakıldığında en az hatayı RBFNetwork sınıflandırıcının yaptığı görülmektedir. Eğer RMSE değerleri eşit olsaydı bu durumda ise Kappa değerine bakılacaktı. Kappa değeri en büyük olan sınıflandırıcı en iyi sınıflandırıcı yani performanslı sınıflandırıcı olarak seçilecekti. Tablo 4'teki sonuçlara göre RBFNetwork sınıflandırıcının doğruluk oranı %93,1034'dir. Yani sınıflandırma işlemi sonunda 27 örnek doğru olarak sınıflandırılırken sadece 2 örnek yanlış olarak

sınıflandırılmıştır. Ayrıca RMSE değeri 0,2452 ve Kappa değeri 0,8585 olduğu için sınıflandırıcının mükemmel bir öğrenme gerçekleştirdiğini söyleyebiliriz. RBFNetwork sınıflandırıcı için en yüksek doğruluk oranı minStdDev parametresinin değiştirilmesi ile elde edilmiştir. Bu parametre Gauss dağılımının standart sapmasını belirlemektedir. Yani bu standart sapma Gauss gürültüsünü değiştirmektedir. Bu değer belirlenmesi gerçek dünya problemlerine göre değişmekte olup deneylerle bu değere karar verilir. Bu değer varsayılan değeri 0,1'dir. Bu deney için bu değer 0,03 olarak belirlenmiştir. Tablo 5'te RBFNetwork sınıflandırıcının detaylı doğruluk tablosu gösterilmektedir. Buna göre RBFNetwork sınıflandırıcı tüm sınıfları (ALS, CO) çok iyi derecede öğrenen bir model üretmiştir. Çünkü Kesinlik, Anma, F-ölçütü ve ROC Alanı değerlerinin 0,5'ten büyük olması bir sınıflandırıcı için arzu edilen bir durumdur. Tablo 5'te ise tüm değerler ya 1 ya da 1'e çok yakındır. Bu nedenle RBFNetwork sınıflandırıcının tüm sınıflar üzerinde genelleme yapabilen bir model oluşturduğunu söyleyebiliriz.

HD-CO eğitim kümesi üzerinde en iyi sonuçları veren ve birbirlerinden farklı varsayımlarla öğrenme sürecini gerçekleştiren sınıflandırıcılar Tablo 6'da gösterilmektedir. Tablo 6'daki sonuçlara göre AdaBoostM1 sınıflandırıcı diğer sınıflandırıcılara göre daha performanslıdır. AdaBoostM1 sınıflandırıcı için en yüksek doğruluk oranı Classifier, numIterations ve useResampling parametrelerinin değiştirilmesi ile elde edilmiştir. Classifier parametresi REPTree olarak belirlenmiştir. Bu parametre temel sınıflandırıcıyı belirleyen bir parametredir. numIterations parametresi ise 30 olarak belirlenmiştir. Bu parametre REPTree sınıflandırıcının 30 defa çalıştırılarak 30 adet model oluşturulmasını sağlamak için kullanılır. Ayrıca bu modeller daha sonra birleştirilerek tek bir modelin oluşturulması sağlanır. useResampling parametresi ise bu

çalışmada True değerini almıştır. Bu parametrenin True değerini alması sonucunda sınıflandırıcının eğitiminde kullanılan örneklerin tekrar ağırlıklandırılması ve modelin iyi öğrenemediği örneklerin tekrar öğrenme kümesinde yer alması sağlanmış olur. Tablo 7'de AdaBoostM1 sınıflandırıcının detaylı doğruluk tablosu gösterilmektedir. Buna göre AdaBoostM1 sınıflandırıcı tüm sınıfları (HD, CO) çok iyi derecede öğrenen bir model üretmiştir. Bu nedenle AdaBoostM1 sınıflandırıcının tüm sınıflar üzerinde genelleme yapabilen bir model oluşturduğunu söyleyebiliriz.

PD-CO eğitim kümesi üzerinde en iyi sonuçları veren ve birbirlerinden farklı varsayımlarla öğrenme sürecini gerçekleştiren sınıflandırıcılar Tablo 8'de gösterilmektedir. Tablo 8'deki sonuçlara göre AdaBoostM1 sınıflandırıcı diğer sınıflandırıcılara göre daha performanslıdır. numIterations parametresi ise 30 olarak belirlenmiştir. Bu parametre AdaBoostM1 sınıflandırıcının 30 defa çalıştırılarak 30 adet model oluşturulmasını sağlamak için kullanılır. Ayrıca bu modeller daha sonra birleştirilerek tek bir modelin oluşturulması sağlanır. Tablo 9'da AdaBoostM1 sınıflandırıcının detaylı doğruluk tablosu gösterilmektedir. Buna göre AdaBoostM1 sınıflandırıcı tüm sınıfları (PD, CO) iyi derecede öğrenen bir model üretmiştir. Bu nedenle AdaBoostM1 sınıflandırıcının tüm sınıflar üzerinde genelleme yapabilen bir model oluşturduğunu söyleyebiliriz.

NDD-CO eğitim kümesi üzerinde en iyi sonuçları veren ve birbirlerinden farklı varsayımlarla öğrenme sürecini gerçekleştiren sınıflandırıcılar Tablo 10'da gösterilmektedir. Tablo 10'daki sonuçlara göre LogitBoost sınıflandırıcı diğer sınıflandırıcılara göre daha performanslıdır. numIterations parametresi ise 100 olarak belirlenmiştir. Bu parametre LogitBoost sınıflandırıcının 100 defa çalıştırılarak 100 adet model oluşturulmasını sağlamak için kullanılır. Ayrıca bu

Tablo 4. ALS-CO eğitim kümesi üzerinde sınıflandırıcıların özet tablosu (Summary table of the classifiers on ALS-CO training data)

Sınıflandırıcı	Parametre	Doğruluk (%)	RMSE	Kappa
Randomize	-	55,1724	-	-
ZeroR	-	55,1724	0,5026	0
RBFNetwork	<i>minStdDev=0,03</i> <i>artificialSize=2,4</i>	93,1034	0,2452	0,8585
Decorate	<i>unpruned=True</i> <i>desiredSize=25</i>	93,1034	0,2478	0,8606
NNge	<i>numFoldersMIOption=1</i>	93,1034	0,2626	0,8606
BayesNet	<i>alpha=3,0</i>	89,6552	0,3214	0,7924
NBTree	-	89,6552	0,3216	0,7924
LWL	<i>KNN=3</i>	86,2069	0,3714	0,7129
HyperPipes	-	82,7586	0,4382	0,6588

Tablo 5. Sınıflara göre RBFNetwork sınıflandırıcının detaylı doğruluk tablosu
(Detailed accuracy table of RBFNetwork classifier by the classes)

	TP	FP	Kesinlik	Anma	F-ölçütü	ROC Alanı	Sınıf
	0,846	0	1	0,846	0,917	0,933	ALS
	1	0,154	0,889	1	0,941	0,933	CO
Ağırlıklı Ortalama	0,931	0,085	0,939	0,931	0,93	0,933	

modeller daha sonra birleştirilerek tek bir modelin oluşturulması sağlanır. Tablo 11'de LogitBoost sınıflandırıcının detaylı doğruluk tablosu gösterilmektedir.

Buna göre LogitBoost sınıflandırıcı tüm sınıfları (NDD, CO) iyi derecede öğrenen bir model üretmiştir. Bu nedenle LogitBoost sınıflandırıcının tüm sınıflar üzerinde genelleme

Tablo 6. HD-CO eğitim kümesi üzerinde sınıflandırıcıların özet tablosu (Summary table of the classifiers on HD-CO training data)

Sınıflandırıcı	Parametre	Doğruluk (%)	RMSE	Kappa
Randomize	-	36,1111	-	-
ZeroR	-	55,5556	0,4982	0
AdaBoostM1	<i>Classifier=REPTree</i> <i>numIterations=30</i> <i>useResampling=True</i>	97,2222	0,1841	0,9441
LWL	<i>Classifier=ADTree</i> <i>numOfBoostingIterations=40</i> <i>searchPath='Expand a random path'</i>	91,6667	0,2677	0,8323
RandomForest	-	91,6667	0,3428	0,8302
DTNB	-	88,8889	0,3411	0,6625
SimpleLogistic	<i>errorOnProbabilities=True</i> <i>numBoostingIterations=10</i> <i>useAIC=True</i> <i>weightTrimBeta=0,1</i>	86,1111	0,3747	0,7205
BayesNet	<i>alpha=0,01</i>	77,7778	0,4584	0,55
HyperPipes	-	77,7778	0,4722	0,55

Tablo 7. Sınıflara göre AdaBoostM1 sınıflandırıcının detaylı doğruluk tablosu
(Detailed accuracy table of AdaBoostM1 classifier by the classes)

	TP	FP	Kesinlik	Anma	F-ölçütü	ROC Alanı	Sınıf
Ağırlıklı Ortalama	0,95	0,125	0,905	0,95	0,927	0,956	HD
	0,875	0,05	0,933	0,875	0,903	0,956	CO
Ağırlıklı Ortalama	0,917	0,092	0,917	0,917	0,916	0,956	

Tablo 8. PD-CO eğitim kümesi üzerinde sınıflandırıcıların özet tablosu (Summary table of the classifiers on PD-CO training data)

Sınıflandırıcı	Parametre	Doğruluk (%)	RMSE	Kappa
Randomize	-	45,1612	-	-
ZeroR	-	35,4839	0,5047	-0,2917
AdaBoostM1	<i>numIterations=30</i> <i>numOfBoostingIterations=13</i>	83,8710	0,4099	0,6778
ADTree	<i>searchPath='Expand a random path'</i>	80,6452	0,4021	0,6125
IBk	<i>KNN=2</i> <i>crossValidate=True</i> <i>hiddenLayers=1</i>	77,4194	0,5154	0,5489
MultilayerPerceptron	<i>momentum=0,3</i> <i>trainingTime=250</i>	70,9677	0,5106	0,4200
BayesNet	<i>alpha=0,01</i>	64,5161	0,5477	0,2940
NNge	-	64,5161	0,5957	0,2969
VFI	<i>weightByConfidence=False</i>	61,2903	0,4981	0,2185

Tablo 9. Sınıflara göre AdaBoostM1 sınıflandırıcının detaylı doğruluk tablosu
(Detailed accuracy table of AdaBoostM1 classifier by the classes)

	TP	FP	Kesinlik	Anma	F-ölçütü	ROC Alanı	Sınıf
Ağırlıklı Ortalama	0,867	0,188	0,813	0,867	0,839	0,788	PD
	0,813	0,133	0,867	0,813	0,839	0,788	CO
Ağırlıklı Ortalama	0,839	0,160	0,840	0,839	0,839	0,788	

yapabilen bir model oluşturduğunu söyleyebiliriz. olduğumuz çalışmamız karşılaştırılmıştır. Bu karşılaştırma Tablo 12’de konu ile ilişkili çalışmalarla birlikte yapmış tablosunda kullanılan eğitim kümesi, veri elde etme yöntemi,

Tablo 10. NDD-CO eğitim kümesi üzerinde sınıflandırıcıların özet tablosu
(Summary table of the classifiers on NDD-CO training data)

Sınıflandırıcı	Parametre	Doğruluk (%)	RMSE	Kappa
Randomize	-	35,9375	-	-
ZeroR	-	75	0,4342	0
LogitBoost	<i>numIterations=100</i>	92,1875	0,2869	0,7872
ADTree	<i>searchPath='Expand the heaviest path' / 'Expand the best z-pure path'</i>	90,625	0,2955	0,75
RBFNetwork	<i>minStdDev=0</i>	82,8125	0,3615	0,5111
DecisionTable	<i>ridge=1.0E-9</i>	81,25	0,381	0,4545
KStar	<i>crossVal=4</i>	81,25	0,4077	0,5714
BayesNet	-	79,6875	0,395	0,5273
HyperPipes	<i>estimator=BMAEstimator</i>	75	0,4231	0
	-			

Tablo 11. Sınıflara göre LogitBoost sınıflandırıcının detaylı doğruluk tablosu
(Detailed accuracy table of LogitBoost classifier by the classes)

	TP	FP	Kesinlik	Anma	F-ölçütü	ROC Alanı	Sınıf
	0,958	0,188	0,939	0,958	0,948	0,906	NDD
	0,813	0,042	0,867	0,813	0,839	0,906	CO
Ağırlıklı Ortalama	0,922	0,151	0,921	0,922	0,921	0,906	

Tablo 12. İlişkili çalışmaların karşılaştırma sonuçları (Comparison results of related works)

Çalışma	Eğitim Kümesi	Veri Elde Etme Yöntemi	Sınıflandırıcı	Doğruluk Oranı
Önerilen Yöntem	ALS-CO	Yere uygulanan dikey kuvvet	Radyal Tabanlı Fonksiyon Ağı	%93,1034
	HD-CO		AdaBoostM1	%97,2222
	PD-CO		AdaBoostM1	%83,871
	NDD-CO		LogitBoost	%92,1875
[19]	ALS-CO	Yürüme ritim sinyali	Destek Yöney Makinaları, Rastgele Ormanlar, Çok Katmanlı Algılayıcı, k En Yakın Komşuluk Algoritması	%96,55
	HD-CO			%100
	PD-CO			%100
	NDD-CO			%96,83
[18]	PD-CO	Yere uygulanan dikey kuvvet	Ağırlıklandırılmış Bulanık Üyeli Sinir Ağı	%77,33
[17]	ALS-CO	Yürüme ritim sinyali	Destek Yöney Makinaları	%96,79
	HD-CO			%90,28
	PD-CO			%89,33
[16]	PD-CO	Yürüme ritim sinyali	Çok Katmanlı Algılayıcı	%95,63
[15]	ALS, HD, PD, CO	Yere uygulanan dikey kuvvet	İkinci Dereceden Bayes Normal Sınıflandırıcı	%86,957 (Ort.)
[14]	PD-CO	Yürüme ritim sinyali	Parzen Window Metodu	%90,32
[13]	ALS-CO	Yürüme ritim sinyali	En Küçük Kareler Destek Yöney Makinaları	%89,66
[12]	NDD-CO	Yürüme ritim sinyali	Elman's Yinelenen Sinir Ağı	%90,6
	ALS, HD, PD, CO			%89,8
				(Ort.)
[11]	PD-CO	Yürüme analizi (initial contact)	Giyilebilir Etkinlik İzleme Sistemi	%97,6
		Yürüme analizi (end contact)		%99,4

kullanılan sınıflandırıcılar ve bu sınıflandırıcıların doğruluk sonuçları gösterilmektedir. Bu sonuçlara göre yere uygulanan dikey kuvvet verileri kullanılarak elde edilen sonuçlara göre çalışmamız diğer çalışmalara göre en yüksek doğruluk oranına sahiptir. Çalışmamızın ortalama doğruluk oranı %91,596'dır. Yürüme ritim sinyalleri kullanılarak gerçekleştirilen karşılaştırma sonuçlarına göre [19] çalışması diğer çalışmalara göre en yüksek doğruluk oranına sahiptir. [19] çalışmasının ortalama doğruluk oranı %98,345'dir.

9. SONUÇLAR (CONCLUSIONS)

Bu çalışmada nöro-dejeneratif hastalıkların sınıflandırılması için kuvvete duyarlı alıcılardan elde edilen yürüme kuvvet sinyalleri kullanıldı. Bu sinyaller, ilk olarak Discrete Meyer dalgacığı kullanılarak yedinci seviyeye kadar ayrıştırıldı. Ayrıştırma işleminden sonra yüksek ve düşük frekanslı sinyaller elde edildi. Bu sinyallerden yüksek frekanslı olanlar detay sinyali olarak bilinirken, düşük frekans olanlar yaklaşık sinyali olarak bilinmektedir. Bu ayrıştırma işleminden sonra yedinci seviyedeki yaklaşım sinyali seçildi. Bu seçim işlemine deneysel sonuçlar ışığında karar verilmiştir. Daha sonrasında bu 7. seviye yaklaşım sinyali üzerinde tepe analizi gerçekleştirilerek sinyalin yerel maksimumları, bu yerel maksimumların tepe konumları, tepe genişlikleri ve tepe çıkıntılarını bulundu. Sonrasında bu dört tepe özelliğinin her birinden 15 adet temel istatistiksel öznitelik elde edildi. Böylelikle sol ayak için 60 ve sağ ayak için 60 olmak üzere toplamda 120 öznitelik elde edildi. Bu aşamadan sonra ALS-CO, HD-CO, PD-CO ve NDD-CO veri kümeleri oluşturuldu. Bu veri kümeleri için 120 öznitelik yerine bu veri kümelerindeki örüntüyü ortaya çıkarmak için en fazla enformasyonu veren özniteliklerin seçimi işlemi gerçekleştirildi. Bunun için OneRules sınıflandırıcı kullanıldı. Bir sonraki aşamada ise RBFNetwork, Adaboost ve LogitBoost algoritmaları kullanılarak ALS-CO için %93,1 doğruluk, HD-CO için %97,22 doğruluk, PD-CO için %83,87 doğruluk ve son olarak NDD-CO veri kümesi için %92,18 doğruluk elde edildi. Böylelikle ALS-CO, HD-CO ve NDD-CO veri kümelerini temsil eden çok iyi modeller oluşturulurken, PD-CO veri kümesini temsil eden iyi bir model oluşturulmuştur. Sonuç olarak nöro-dejeneratif hastalıkların erken ve doğru teşhisinde yapay öğrenme yöntemleri ve dalgacık dönüşümü yönteminin etkin bir biçimde kullanılabileceği görülmüştür.

TEŞEKKÜR (ACKNOWLEDGEMENT)

Yazarlar bu araştırma çalışması sırasındaki desteklerinden dolayı Kırklareli Üniversitesi ve İstanbul Aydın Üniversitesine teşekkür ederler.

KAYNAKLAR (REFERENCES)

1. JPND Research. What is Neurodegenerative Disease?. <http://www.neurodegenerationresearch.eu/about/what/>. Yayın tarihi Şubat 7, 2012. Erişim tarihi Haziran 18, 2017.
2. Özaras N., Yalçın S., Normal Yürüme ve Yürüme Analizi, Turkish Journal of Physical Medicine and Rehabilitation, 48 (3), 2002.
3. Barr A.E., Biomechanics and Gait, Orthopaedic Knowledge Update 7: Home Study Syllabus, Editör: Koval K. J., American Academy of Orthopaedic Surgeons, Rosemont, Illinois, A.B.D., 31-38, 2002.
4. Kanatlı U., Yetkin H., Songür M., Öztürk A., Bölükbaşı S., Yürüme Analizinin Ortopedik Uygulamaları, Türk Ortopedi ve Travmatoloji Birliği Derneği Dergisi, 5 (1-2), 53-59, 2006.
5. Zengin S., İhmal Edilmiş Aşıl Tendon Rüptürlerinde Lindholm Yöntemi ve Sonuçlarının Yürüme Analizi ile Değerlendirilmesi, Doktora Tezi, T.C. Sağlık Bakanlığı Okmeydanı Eğitim ve Araştırma Hastanesi Ortopedi ve Travmatoloji Kliniği, 2007.
6. Özaras N., Yalçın S., Yürüme Analizi, (2. Baskı), Avrupa Tıp Kitapçılık, İstanbul, 2002.
7. NCC-CC (Collaborating Centre for Chronic Conditions), Parkinson's Disease: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care, Royal College of Physicians of London, London, 2006.
8. Bronstein J.M., Tagliati M., Alterman R. L. and et al., Deep brain stimulation for Parkinson disease: an expert consensus and review of key issues, Archives of neurology (Arch. Neurol), 68 (2), 2011.
9. Frank S., Jankovic J., Advances in the Pharmacological Management of Huntington's Disease, Drugs, 70 (5), 561-71, 2010.
10. Russell P., Harrison R., What is amyotrophic lateral sclerosis, Clinical Pharmacist, 6 (7), 2014.
11. Han J., Jeon H.S., Yi W.J., Jeon B.S., Park K.S., Adaptive windowing for gait phase discrimination in Parkinsonian gait using 3-axis acceleration signals, Medical & biological engineering & computing (MBEC), 47, 1155-1164, 2009.
12. Dutta S., Chatterjee A., Munshi S., An automated hierarchical gait pattern identification tool employing cross-correlation-based feature extraction and recurrent neural network based classification, Expert Systems, 26 (2), 202-217, 2009.
13. Wu Y., Krishnan S., Computer-aided analysis of gait rhythm fluctuations in amyotrophic lateral sclerosis, Medical & biological engineering & computing (MBEC), 47, 1165-1171, 2009.
14. Wu Y., Krishnan S., Statistical Analysis of Gait Rhythm in Patients With Parkinson's Disease, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 18 (2), 150-158, 2010.
15. Banaie M., Pooyan M., Mikaili M., Introduction and application of an automatic gait recognition method to diagnose movement disorders that arose of similar causes, Expert Systems with Applications, 38, 7359-7363, 2011.
16. Manap H.H., Tahir N.M., Yassin A.I.M., Statistical Analysis of Parkinson Disease Gait Classification using Artificial Neural Network, IEEE Signal Processing and Information Technology (ISSPIT), 60-65, 2011.

17. Daliri M.R., Automatic diagnosis of neuro-degenerative diseases using gait dynamics, *Measurement*, 45, 1729-1734, 2012.
18. Lee S., Lim J.S., Parkinson's disease classification using gait characteristics and wavelet-based feature extraction, *Expert Systems with Applications*, 39, 7338-7344, 2012.
19. Xia Y., Gao Q., Ye Q., 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*, 18, 254-262, 2015.
20. Nair S.R., Tan L.K., Ramli N.M., Lim S.Y., Rahmat K., Nor H.M. A decision tree for differentiating multiple system atrophy from Parkinson's disease using 3-T MR imaging, *European Society of Radiology*, 23 (6), 1459-1466, 2013.
21. Ota M., Nakata Y., Ito K., Kamiya K., Ogawa M., Murata M., Obu S., Kunugi H., Sato N. Differential Diagnosis Tool for Parkinsonian Syndrome Using Multiple Structural Brain Measures, *Computational and Mathematical Methods in Medicine*, 2013 (6), 1-10, 2013.
22. Drotár P., Mekyskaa J., Rektorová I., Masarováb L., Smékala Z., Faundez-Zanuy M., Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease, *Computer Methods and Programs in Biomedicine*, 117 (3), 405-411, 2014.
23. Akdemir Ü.Ö., Tokçaeer B., Karakuş A., Kapucu L.Ö., Brain 18F-FDG PET Imaging in the Differential Diagnosis of Parkinsonism, *Clinical Nuclear Medicine*, 93 (3), 220-226, 2014.
24. Lee S., Kim M., Lee H.M., Kwona K., Kim H., Koh S., Differential diagnosis of parkinsonism with visual inspection of posture and gait in the early stage, *Gait & Posture*, 39 (4), 1138-1141, 2014.
25. Pilleri M., Levedianos G., Weis L., Gasparoli E., Facchini S., Biundo R., Formento-Dojot P., Antonini A., Heart rate circadian profile in the differential diagnosis between Parkinson disease and multiple system atrophy, *Parkinsonism and Related Disorders*, 20 (2), 217-221, 2014.
26. Navarro-Otano J., Gaig C., Muxi A., Lomeña F., Compta Y., Buongiorno M.T., Martí M.J., Tolosa E., Valldeoriola F., 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, 2014.
27. Salvatorea C., Cerasab A., Castiglioni I., Gallivanonec F., Augimerib A., Lopezd M., Arabiae G., Morellie M., Gilardi M.C., Quattrone A., Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and Progressive Supranuclear Palsy, *Journal of Neuroscience Methods*, 222, 230-237, 2014.
28. Feng J., Huang B., Yang W., Zhang Y., Wang L., Wang L., Zhong X., 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, 2014.
29. Baudrexel S., Seifried C., Penndorf B., Klein J.C., Middendorp M., Steinmetz H., Grünwald F., Hilker R., 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, 2014.
30. 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., Mir P., 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, 2015.
31. Zanigni S., Testa C., Calandra-Buonaura G., Sambati L., Guarino M., Gabellini A., Evangelisti S., Cortelli P., Lodi R., Tonon C., The contribution of cerebellar proton magnetic resonance spectroscopy in the differential diagnosis among parkinsonian syndromes, *Parkinsonism and Related Disorders*, 21 (8), 929-937, 2015.
32. Bradvica I.S., Mihaljević I., Butković-Soldo S., Kadojić D., Titlić M., Bradvica M., Kralik K., 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, 2015.
33. Vranová H.P., Hényková E., Mareš J., Kaiserová M., Menšíková K., Vašítk M., Hlušítk P., Zapletalová J., Strnadb M., Stejskal D., Kaňovský P., Clusterin CSF levels in differential diagnosis of neurodegenerative disorders, *Journal of the Neurological Sciences*, 361, 117-121, 2016.
34. Drotár P., Mekyska J., Rektorová I., Masarová L., Smékál Z., Faundez-Zanuy M., Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease, *Artificial Intelligence in Medicine*, 67, 39-46, 2016.
35. Aydın F., Aslan Z. Classification of Neurodegenerative Diseases using Machine Learning Methods, *International Journal of Intelligent Systems and Applications in Engineering*, 5 (1), 1-9, 2017.
36. PhysioNET. Gait Dynamics in Neuro-Degenerative Disease Data Base. <http://www.physionet.org/physiobank/database/gaitnnd/>. Güncelleme tarihi Ekim 28, 2016. Erişim tarihi Haziran 18, 2017.
37. Hausdorff J. M., Ladin Z., Wei J. Y., Footswitch system for measurement of the temporal parameters of gait, *Journal of Biomechanics (J. Biomech.)*, 28, 347-351, 1995.
38. Haşiloğlu A., Dalgacık dönüşümü ve yapay sınırlar ile döndürmeye duyarsız doku analizi ve sınıflandırma,

- Turkish Journal of Engineering and Environmental Sciences (Turk J Engin Environ Sci.), 25, 405-413, 2001.
39. Shaker M.M., EEG waves classifier using Wavelet transform and Fourier transform, International Journal of Biological and Life Sciences (Int. J. Biol. Sci.), 1 (2), 85-88, 2005.
 40. Procházka A., Jech J., Smith J., Wavelet transform use in signal processing, 31st International Conference in Acoustics., Prague, Czech Technical University, 209-213, 1994.
 41. Morlet J., Arens G., Fourgeau E., Giard D., Wave propagation and sampling theory, Part1: Complex signal land scattering in multilayer media, Journal of Geophysics, 47, 203-221, 1982.
 42. Mallat S. G., Multiresolution approximations and Wavelet orthonormal bases of $L_2(\mathbb{R})$, Transactions of the American Mathematical Society (Journal of the AMS), 315 (1), 69-87, 1989.
 43. Miner N.E., An introduction to Wavelet theory and analysis, Sandia Raporu, Sandia National Laboratories, Albuquerque, New Mexico, California, A.B.D., 1998.
 44. Huang N.E., Introduction to the Hilbert-Huang Transform and Its Related Mathematical Problems, Hilbert Huang transform and its applications, Vol. 5, Editör: Huang N. E., Shen S. S. P., World Scientific Publishing, Hackensack, A.B.D., 1-24, 2005.
 45. Drozdov A., Pomortsev I., Tyutyukin K., Baloshin Y., Comparison of Wavelet Transform and Fourier Transform Applied to Analysis of Non-Stationary Processes, Nanosystems: Physics, Chemistry, Mathematics (Nanosist.: fiz. him. mat.), 5 (3), 363-373, 2014.
 46. Donald M., Spiegelhalter D.J., Taylor C.C., Machine Learning: Neural and Statistical Classification, Overseas Press, 2009.
 47. Nilsson N.J., Introduction to Machine Learning: An Early Draft of a Proposed Textbook, Robotics Laboratory, Department of Computer Science, Stanford University. [http://ai.stanford.edu/people/nilsson/MLBO OK.pdf](http://ai.stanford.edu/people/nilsson/MLBO_OK.pdf). Yayın tarihi Kasım 3, 1998. Erişim tarihi Haziran 18, 2017.
 48. Kohavi R., Provost F., Glossary of Terms, Machine Learning, 30, (2-3), 271-274, 1998.
 49. Duda R.O., Hart P.E., Stork D.G., Pattern classification, 2nd ed., Wiley-Interscience, New York, 2001.
 50. Alpaydm E., Introduction to Machine Learning, 3rd ed., The MIT Press, 2014.
 51. Wolpert D., The Lack of A Priori Distinctions between Learning Algorithms, Neural Computation, 1341-1390, 1996.
 52. Wolpert D.H., Macready W.G., No Free Lunch Theorems for Optimization. IEEE Transactions on Evolutionary Computation (TEVC), 1 (1), 67-82, 1997.
 53. Witten I.H., Frank E., Data mining : practical machine learning tools and techniques, 3rd ed., Morgan Kaufmann, 2011.
 54. Moody J., Fast learning in networks of locally-tuned processing units, Neural Computation, 1, 281-294, 1989.
 55. Casdagli M., Nonlinear prediction of chaotic time series, Physica D Nonlinear Phenomena (Physica D), 35, 335-356, 1989.
 56. Broomhead D.S., Lowe D., Multivariable functional interpolation and adaptive networks, Complex Systems, 2, 321-355, 1988.
 57. Matej S., Lewitt R.M., Practical considerations for 3 D image reconstruction using spherically volume elements, IEEE Transactions on Medical Imaging (IEEE-TMI), 15, 68-78, 1996.
 58. Mashor M.Y., Hybrid Training Algorithm for RBF Network, In International Journal of the Computer, The Internet and Management (IJCIM), 8 (2), 50-65, 2000.
 59. Freund Y., Schapire R. E., Experiments with a new boosting algorithm, In: Thirteenth International Conference on Machine Learning, San Francisco, A.B.D., 148-156, 1996.
 60. Ratsch G., Onoda T., Muller K.R., Soft margins for AdaBoost, Machine Learning, 42, 287-320, 2001.
 61. Friedman J., Hastie T., Tibshirani R., Additive logistic regression: a statistical view of boosting, The Annals of Statistics, 28 (2), 337-407, 2000.
 62. Cai Y., Feng K., Lu W., Chou K., Using LogitBoost classifier to predict protein structural classes, Journal of Theoretical Biology, 238 (1), 172-176, 2006.
 63. Holte R.C., Very simple classification rules perform well on most commonly used datasets, Machine Learning, 11, 63-91, 1993.
 64. Rendell L., Seshu R. Learning Hard Concepts Through Constructive Induction, Computational Intelligence, 6, 247-270, 1990.
 65. Shavlik J., Mooney R.J., Towell G., Symbolic and Neural Learning Algorithms: An Experimental Comparison, Machine Learning, 6, 111-143, 1991.
 66. Buntine W., Niblett T., A Further Comparison of Splitting Rules for Decision-Tree Induction, Machine Learning, 8, 75-86, 1992.
 67. Clark P., Niblett T., The CN2 Induction Algorithm, Machine Learning, 3, 261-283, 1989.
 68. Mingers J., An Empirical Comparison of Pruning Methods for Decision Tree Induction, Machine Learning, 4 (2), 227-243, 1989.
 69. Buddhinath G., Derry D., A Simple Enhancement to One Rule Classification, Department of Computer Science & Software Engineering, University of Melbourne, Australia, 2007.
 70. Cohen J., A coefficient of agreement for nominal scales, Educational and Psychological Measurement (EPM), 20 (1), 37-46, 1960.
 71. Gwet K., Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters, (4th ed.), Advanced Analytics, 2014.
 72. Landis J. R., Koch G. G., The measurement of observer agreement for categorical data, Biometrics, 33, 159-174, 1977.

73. Haltaş A., Alkan A., Karabulut M., Performance Analysis of Heuristic Search Algorithms in Text Classification, Journal of the Faculty of Engineering and Architecture of Gazi University, 30 (3), 417-427, 2015.
74. Lehmann E.L., Casella George, Theory of Point Estimation, (2nd ed.), Springer, New York, A.B.D., 2003.
75. Domingos P., A Unified Bias-Variance Decomposition and its Applications, In Proc. 17th International Conf. on Machine Learning, San Francisco, CA, A.B.D., 231-238, 2000.
76. Geman S., Bienenstock E., Doursat R., Neural networks and the bias/variance dilemma, Neural Computation, 4 (1), 1-58, 1992.
77. Mitchell T. M., Machine Learning, McGraw-Hill Science/Engineering/Math, 1997.
78. Bouckaert R.R., Frank E., Hall M., Kirkby R., Reutemann P., Seewald A., Seuse D., WEKA Manual for 3-6-15. <https://sourceforge.net/projects/weka/files/documentation/3.6.x/>. Yayın tarihi Aralık 19, 2016. Erişim tarihi Haziran 18, 2017.