

## Dalgacık Paket Dönüşümü, ReliefF Özellik Seçimi ve Topluluk Öğrenme Algoritması Tabanlı Bir Kısmi Deşarj Arızası Tespit Yöntemi

Belkıs ERİŞTİ\*

<sup>1</sup> Elektrik ve Enerji Bölümü, Teknik Bilimler MYO, Mersin Üniversitesi, Mersin, Türkiye  
\*<sup>1</sup> beristi@mersin.edu.tr

(Geliş/Received: 17/04/2023;

Kabul/Accepted: 29/05/2023)

**Öz:** Enerji nakil hatlarında birçok arıza olayı meydana gelebilmektedir. Özellikle hatlarda faz iletkenlerinin çevresel bitki örtüleriyle ve birbirleriyle temas etmeleri sonucunda oluşan arızalar sıklıkla meydana gelir. Bu şekilde oluşabilecek arızaların önüne geçebilmek için özellikle enerji nakil hatlarında izolasyonlu iletkenler yaygın olarak kullanılmaktadır. Ancak izolasyonlu iletkenlerin yalıtım malzemesinde meydana gelebilecek deformasyonlar bu iletkenlerde kısmi deşarj (KD) adı verilen olaylara sebep olabilirler. Oluşabilecek çok daha büyük arızaların önüne geçebilmek için KD'lerin hızlı bir şekilde tespit edilmesi gerekir. Bu çalışmada, iletim hatlarında meydana gelen KD'lerin tespiti için dalgacık paket dönüşümü (DPD), ReliefF özellik seçim yaklaşımı ve topluluk öğrenme algoritma sınıflandırıcı tabanlı etkili bir tespit yaklaşımı önerilmiştir. Bu yaklaşımın en önemli özelliği, KD verilerinin DPD kullanarak etkili frekans bantlarına dayanan özellikler elde edilmesi ve ReliefF yaklaşımı kullanılarak bu özellikler içerisinde tespit performansı yüksek özelliklerin seçilmesidir. Önerilen tespit sistemi VSB gerçek veri seti kullanılarak test edilmiş ve 89,22% doğruluk oranı elde edilmiştir. Literatürde VSB veri seti kullanan benzer çalışmalarla karşılaştırıldığında başarımın oldukça yüksek olduğu ve önerilen yaklaşımın KD tespiti için etkili bir performans sergilediği görülmüştür.

**Anahtar kelimeler:** Kısmi deşarj, arıza tespiti, dalgacık paket dönüşümü, ReliefF, topluluk öğrenme algoritması.

### A Partial Discharge Fault Detection Method Based on Wavelet Packet Transform, ReliefF Feature Selection and Ensemble Learning Algorithm

**Abstract:** Many faults can occur in power transmission lines. Especially in power transmission lines, faults occur frequently as a result of phase conductors coming into contact with environmental vegetation and each other. Insulated conductors are widely used, especially in power transmission lines, in order to prevent malfunctions that may occur in this way. However, deformations that may occur in the insulating material of insulated conductors may cause events called partial discharge (PD) in these conductors. PD need to be detected quickly in order to prevent much larger failures that may occur. In this paper, an effective detection approach based on wavelet packet transform (WPT), ReliefF feature selection approach and ensemble learning algorithm classifier is proposed for the detection of PD in the transmission line. The most important advantage of this approach is to obtain features based on effective frequency bands by using WPT of PD data and to select features with high detection performance among these features by using ReliefF approach. The proposed detection system is tested using the VSB real dataset and an accuracy rate of 89,22% is obtained. When compared with similar studies using VSB dataset in the literature, it has been seen that the performance is quite high and the proposed approach has an effective performance for PD detection.

**Key words:** Partial discharge, fault detection, wavelet packet transform, ReliefF, ensemble learning algorithm.

#### 1. Giriş

Teknolojinin gelişmesiyle beraber günümüzde elektrik enerjisine olan talebin artması, elektrik güç şebekelerini sadece boyut olarak büyütmeyle kalmamış, aynı zamanda daha karmaşık bir hal almasına sebep olmuştur. Bu büyüme, elektrik enerjisi şebekelerinin üretimden başlayarak iletim ve dağıtımına kadar olan tüm bölümlerini kapsamaktadır [1]. Elektrik şebekelerinin uzun süreli olarak kararlı bir şekilde çalışması, ekonomik kayıpların önlenmesi, bakım maliyetlerinin düşürülmesi, ekipmanların zarar görmemesi gibi durumlar için sürekli araştırma ve geliştirme çalışmaları yapılmaktadır [2]. Ancak iletim hatlarında meydana gelen beklenmedik arızalar elektrik şebekelerinin kararsız çalışmasına sebep olur. Bitki örtüsünün enerji nakil hatlarıyla temas etmesi ve iletkenlerin birbirine çarpması gibi olaylardan kaynaklı çeşitli arıza ve yangınlar meydana gelebilmektedir. Bu olayların önüne geçebilmek için enerji hatlarında yaygın olarak izolasyonlu iletkenler kullanılır. Ancak bahsedilen bu çevresel olaylar, izolasyonlu iletkenlerin yalıtım sistemini bozarak KD'lere sebep olabilir. KD'ler iletim hatlarının izolasyon yapısında oldukça önemli tahribatlar oluşturabilir ve onarılmadan bırakılırsa çok büyük arızalara yol açabileceği gibi yangınlara da sebep olabilirler. KD, elektrik alan şiddetinin yalıtım malzemesinin

\* Sorumlu yazar: [beristi@mersin.edu.tr](mailto:beristi@mersin.edu.tr). Yazarın ORCID Numarası: <sup>1</sup> 0000-0003-1276-2347

kırılma gücünü aştığı durumlarda, yalıtım malzemesinin yüzeyinde meydana gelen küçük bir elektrik kıvılcımıdır. Dolayısıyla, iletim hatlarının oluşturabileceği elektrik kesintileri, orman yangınları gibi önemli problemlerin önlenmesi için KD'lerin tespit edilmesi ve KD'lerin oluşturduğu izolasyon arızalarını ortadan kaldırılması oldukça önemli bir iletim hattı koruma araştırma konusu olmuştur [3].

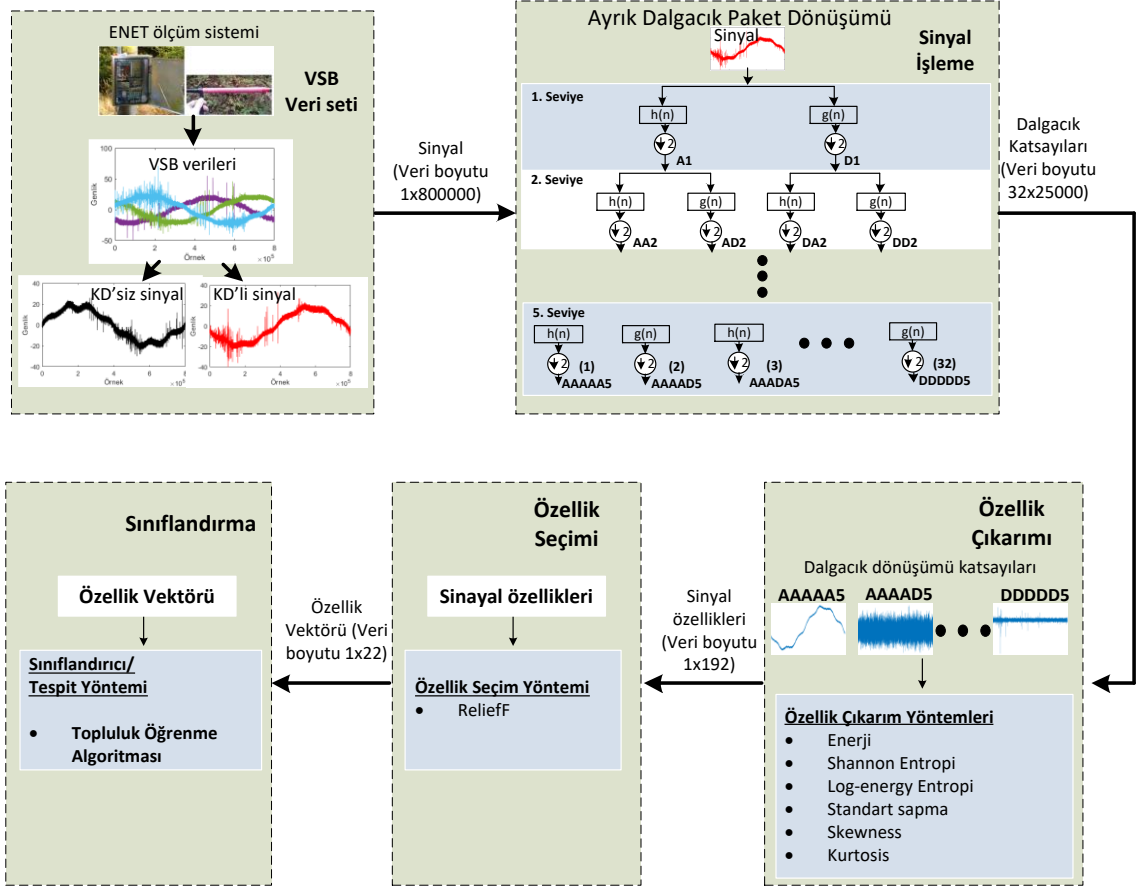
Literatürde enerji hatlarında KD tespiti için birçok çalışma sunulmuştur. [4]'de ölçüm gürültüsüyle bozulan büyük ölçekli verilerde KD tespiti için gürültüsü azaltılmış uzun kısa süreli bellek (long-short term memory-LSTM) tabanlı bir döngü optimizasyonu sınıflandırıcısı önerilmektedir. KD tespit sistemi gürültü azaltma, özellik çıkarma ve LSTM sinir ağı bir bütün olarak birleştirilmiştir. Sistem, havai elektrik hatlarından ölçülen büyük gerçek veriler kullanılarak doğrulanmıştır. [5]'de yapılan çalışmada yalıtılmış havai hat iletkenlerinde arıza tespiti ve sınıflandırması için üç faz sinyallerindeki gizli özellik bilgisinden daha iyi faydalanmak amacıyla çok kanallı evrimsel sinir ağı (convolutional neural network-CNN)-LSTM yapısını kullanmıştır. Üç faz gerilim sinyallerine hızlı Fourier dönüşümü (fast Fourier transform-FFT) uygulanmış, düşük ve yüksek frekans bileşenleri elde edilerek farklı filtre uzunluklarına sahip üç paralel CNN yapısı kullanılmıştır. Ayrıca farklı ölçekli özellikleri sırayla oluşturmak için LSTM yapısı kullanılarak önerilen bu yöntem test edilmiştir. [6]'da üç fazlı bir güç sisteminde KD tespiti için öğrenmeye dayalı çok ölçekli bir özellik mühendisliği yapısı geliştirilmiştir. [7]'de makine öğrenimi algoritmalarını kullanarak KD sinyali modellerini tespit etmeye ve bu potansiyel hataları azaltmaya yönelik bir yaklaşım sunulmuştur. [8]'de herhangi bir özellik çıkarımı yapılmadan güçlü KD tespiti sağlayan CNN'ye dayalı uçtan uca bir sistem önerilmiştir. [9]'da, KD tespiti için derin öğrenme tabanlı bir arıza tespiti yaklaşımı sunulmuştur. [3]'te, spektrogram özellik çıkarma ve CNN'nin avantajlarını birleştiren makine öğrenimi algoritmaları kullanılarak KD tespiti için yüksek frekanslı gerilim sensörlerinden alınan ölçümler kullanılmıştır. [10]'da, yalıtılmış enerji nakil hatlarında arıza tespiti için derin öğrenme tabanlı bir KD model analizi sunulmuştur. Önerilen yöntemin LSTM ve destek vektör makine (DVM) tabanlı yöntem kullanılarak karşılaştırılmış ve karşılaştırma sonrasında LSTM yönteminin daha iyi performans gösterdiği görülmüştür. [11]'de, yalıtılmış havai iletim hatlarında KD'lerin tespiti için CNN tabanlı arıza sınıflandırma algoritması önerilmiş olup, veri setindeki gerilim sinyallerine AlexNet, VGG16 ve ResNet gibi önceden eğitilmiş popüler CNN mimarileri uygulanmıştır. [12]'de, dalgacık dönüşümü (DD) kullanılarak gürültü filtreleme sonrasında KD sinyal özellikleri elde edilmiş ve LSTM kullanılarak KD hatası tespit edilmiştir. [13]'de klasik arıza tespiti yaklaşımlarının eksiklikleriyle başa çıkmak için ikili sınıflandırmaya dayalı otomatik bir KD arıza tespiti yaklaşımı önerilmiştir.

Literatürde gerçekleştirilen makaleler incelendiğinde KD arıza sinyallerinin tespit edilmesinin oldukça zorlu olduğu görülmektedir. Çalışmalarda çoğunlukla VSB veri seti kullanılarak tespit sistemlerinin performansı belirlenmektedir [14]. Bu literatür çalışmalarında genel olarak KD sinyallerine DD, FFT gibi sinyal işleme yaklaşımları uygulandıktan sonra DVM gibi geleneksel sınıflandırıcılar veya yenilikçi CNN, LSTM gibi derin öğrenme yaklaşımları uygulanarak tespit işlemi gerçekleştirilmektedir. Geleneksel sınıflandırıcılar kullanılarak gerçekleştirilen tespit işlemlerinde tespit doğruluğunun 70%-80% civarlarında olduğu görülmektedir. Yenilikçi derin öğrenme yaklaşımları kullanan tespit sistemlerinde ise başarımın 80%-90% aralığında olduğu hatta bazı çalışmalarda 90%'ın üzerinde başarımlar elde edilebildiği gözlemlenmiştir. Bununla beraber yenilikçi tespit yaklaşımları, başarımlarını arttırılabilseler bile özellikle karmaşık bir yapıya sahip olması, oldukça yüksek hesaplama yükü getirmesi ve yüksek grafik işlemcilerle ihtiyaç duyması gibi olumsuz yönleri de bulunmaktadır. Bu makale çalışmasında ise KD sinyallerine 5-seviyeli DPD uygulanarak elde edilen 32 farklı frekans bandında bulunan DPD katsayılarına 6 farklı özellik çıkarımı uygulanarak ayırt edici özellikler çıkarılmıştır. Elde edilen  $1 \times 192$  boyutundaki özellikler içerisinden tespit performansı yüksek özellikler, Relief özellik seçimi yaklaşımı ile belirlenmiştir. Böylece özellik seçimi işlemi sonucunda  $1 \times 22$  düşük veri boyutuna sahip ve tespit performansı yüksek bir özellik vektörü elde edilmiştir. Son aşamada ise bu özellik vektörü, topluluk öğrenme algoritmasına (TÖA'na) giriş olarak sunulmuştur. Önerilen DPD, Relief ve TÖA tabanlı KD tespit sistemini test etmek için VSB gerçek veri seti kullanılmıştır. Elde edilen sonuçlar incelendiğinde literatürde birçok derin öğrenme tabanlı tespit sistemlerinin doğruluk oranının bile üzerinde olan 89.22% tespit doğruluğuna sahip sonuç elde edilmiştir. Sonuç olarak bu makale çalışmasında doğruluk oranı yüksek, hesaplama yükü düşük ve yüksek tespit performansına sahip bir KD tespit sistemi sunulmuştur. Bu çalışmanın diğer bölümler ise şu şekildedir; önerilen çalışmanın algoritması ve VSB veri kümesi Bölüm 2'de verilmiştir. Bölüm 3'te ise önerilen tespit sistemine ilişkin testler, performans analizleri ve sonuçları verilmiştir. Son bölüm olan 4'de ise tartışma ve sonuçlar sunulmuştur.

## 2. Metodoloji

Bu çalışmada önerilen KD tespit sisteminin asıl görevi, enerji nakil hatlarında özellikle izoleli iletkenlerde meydana gelen KD arıza türünü tespit etmektir. Bu tespit sistemi, KD ile ilişkili anormal gerilim sinyalleri üzerine

gerçekleştirilen bir dizi analizlere dayanır. Ayrıca bu tespit sisteminde, normal gerilim sinyalleri üzerine de aynı analizler yapılarak ikili sınıf tabanlı bir tespit yaklaşımı oluşturulur. Şekil.1’de önerilen KD tabanlı sistemin genel yapısı gösterilmiştir. Önerilen KD tespit sistemi için gerçekleştirilen analizler genel olarak dört aşamadan oluşmaktadır. KD tespit sisteminde ilk olarak ham gerilim sinyallerine ayrık DPD tabanlı sinyal işleme yaklaşımı uygulanır. Daha sonra, ilk aşamada elde edilen DPD katsayılarına özellik çıkarımı işlemi gerçekleştirilir. Bir sonraki aşamada ise ReliefF özellik seçim algoritması kullanılarak elde edilen özelliklerin içerisinde ayırt edici özellikler seçilir. Son aşamada ise sonuç özellik vektörü TÖA sınıflandırıcıya uygulanarak KD tespit işlemi gerçekleştirilir.



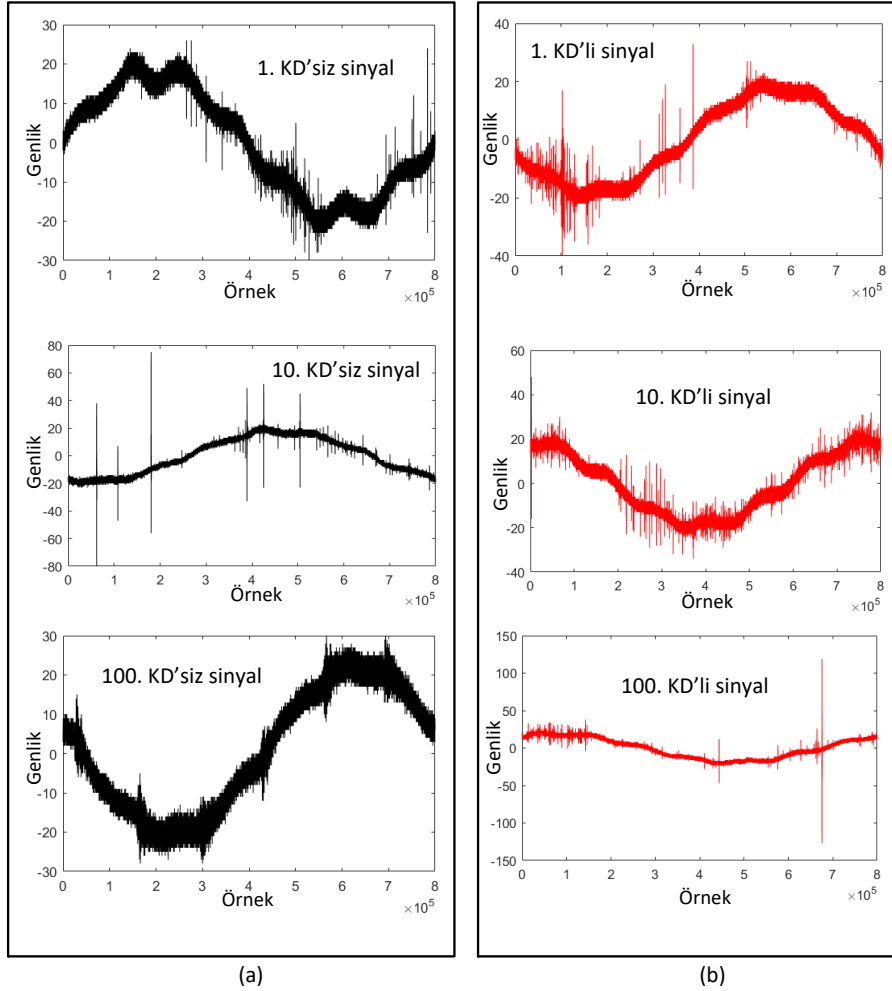
Şekil 1. Önerilen KD tespit algoritmasının genel yapısı.

Önerilen KD tespit sisteminin performansını belirlemek için Ostrava Teknik Üniversitesi ENET araştırma merkezi tarafından oluşturulan VSB veri seti kullanılmıştır. VSB veri seti, gerçek güç hatlarının gerilim sinyallerini yakalamak için yeni bir cihaz kullanarak Ostrava Teknik Üniversitesindeki ENET araştırma merkezi tarafından oluşturulduğu için gerçek zamanlı KD verilerinden oluşmaktadır [14, 15]. ENET araştırma merkezi KD’leri ölçmek için izoleli iletkenlerin üzerine takılabilen basit bir ölçüm cihazı geliştirmiştir. Şekil 2’de gösterildiği gibi yalıtılmış iletken boyunca KD voltaj sinyali elde etmek için izoleli iletken etrafına tek katmanlı bir bobin sarılmıştır. ENET’in ölçüm sistemi, KD ölçümü için geliştirilen diğer sistemlere kıyasla düşük maliyet ve yüksek güvenilirliğe sahiptir.

VSB veri seti, iki ayrı alt veri setinden oluşmaktadır. İlk veri setinde, 0 olarak etiketlenmiş KD’siz gerilim sinyali (8.186) ve 1 olarak etiketlenmiş KD’li gerilim sinyalinden (525) oluşan 8.711 veri bulunmaktadır. Her bir gerilim verisi, tek periyottan oluşan 40 MHz örnekleme frekansına sahip 800000 örnekten oluşmaktadır. Şekil 3’te, ilk veri seti içerisinde üç KD’li sinyal ve KD olmayan sinyal örnekleri (1., 10. ve 100. veri) gösterilmiştir. Genel olarak KD’li sinyal, KD’li olmayan sinyale göre daha büyük gürültü pik değerlerine sahiptir. Her iki sinyal arasında belirleyici bir eşik değeri bulunmaması ve gürültü etkisinin oldukça yüksek olması nedeniyle bu iki sinyal türünü herhangi bir analiz işlemi gerçekleştirilmeden birbirlerinden ayırt etmek oldukça zordur. İkinci veri setinde ise KD’nin varlığına ilişkin herhangi bir durum belirtmeyen 20037 veri bulunmaktadır.



Şekil 2. ENET ölçüm sistemi ve izoleli iletken yüzeyindeki tek katlı bobin.

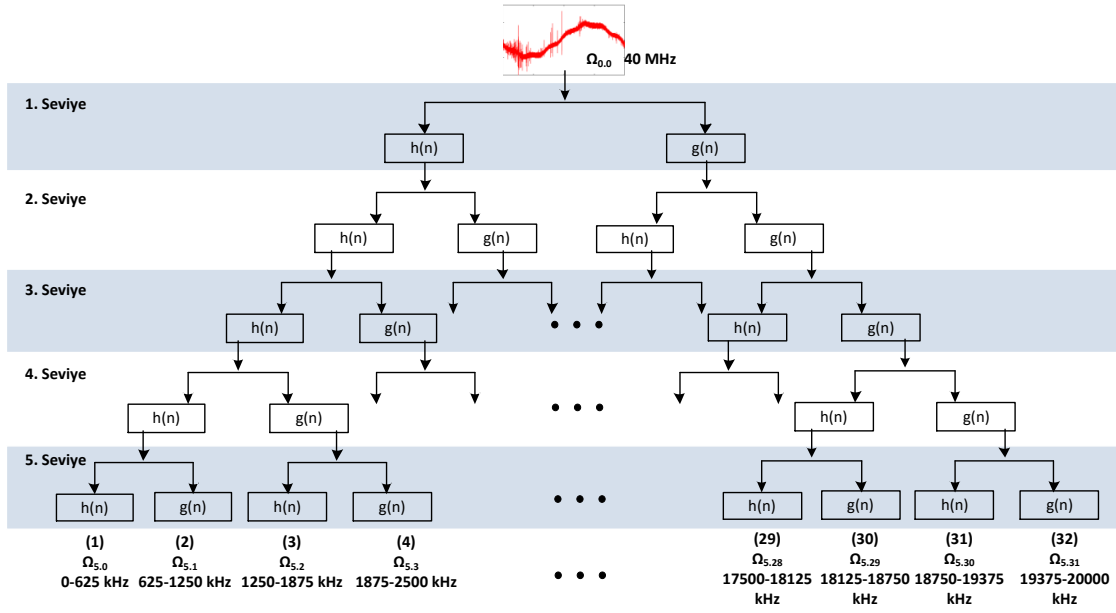


Şekil 3. VSB veri seti örnekleri, a) KD'siz normal sinyal ve b) KD'li sinyal.

## 2.1. Sinyal İşleme

İzoleli enerji hatlarında meydana gelen ve zaman-genlik bilgilerinden oluşan bir KD'yi tespit etmek oldukça zordur. VSB veri setinden de görülebileceği gibi her iki sinyal sınıfına ait ayırt edici bir genlik bilgisi yoktur. Ayrıca, sinyal verileri çok yüksek veri boyutuna sahiptir. Bu nedenle frekans bilgilerini analiz etmek oldukça etkin bir yaklaşım olacaktır. Bu çalışmada, DPD kullanılarak her iki sınıfın sinyalleri için farklı frekans bantlarında frekans bilgileri elde edilmiştir. Dönüşüm sonrasında her iki sınıf verileri için ayırt edici bilgiler elde edilmiş olup veri boyutu değişmemiştir. Veri boyutu indirgeme işlemi ise DPD katsayıları

kullanılarak bir sonraki özellik çıkarımı aşamasında gerçekleştirilmiştir. Gerçekleştirilen DPD işleminde literatürde elektromanyetik geçici analizleri ve arıza analizleri için yaygın olarak tercih edilen Daubechies 4 (db4) ana dalgacığı kullanılmıştır. Bu çalışmada KD verilerine 5-seviyeli DPD analiz işleminin en uygun olduğu sonucuna varılmıştır. Analiz seviyesi azaltılırsa işlem yükü azalırken hassas frekans bant aralıkları elde edilmesinde zorluklarla karşılaşılabilir. Analiz seviyesi azaltılırsa hassas frekans bant aralığı elde edilirken işlem yükü problemi ve büyük veri boyutuna sahip özellik vektörü oluşturma problemi ortaya çıkacaktır. Özellik çıkarımı için kullanılan sinyale 5-seviyeli DPD analizi uygulanarak toplam 32 frekans alt bandı içerisinde DPD katsayıları elde edilmiştir. Böylece, bu frekans bantlarının her birinde sinyalin KD'li veya normal bir sinyal olduğunu ayırt edebilecek önemli frekans bilgileri bulunmaktadır. Şekil 4'de bu çalışmada gerçekleştirilen 5-seviyeli DPD dönüşümü sonucunda elde edilen alt uzay ağacı ve sonuç frekans bant aralıkları verilmiştir.



Şekil 4. 5-seviyeli DPD dönüşümü ile elde edilen 32 alt frekans bandı.

DD sinyal işleme tekniği, dalgacıklar kullanılarak özellikle durağan ve periyodik olmayan sinyallerin keskin yükselmelerinin ve süreksizliklerinin belirlenmesinde çok etkili bir yöntemdir [16]. Dalgacıklar kelimesinin tam anlamıyla sınırlı süreye ve sıfır ortalama değere sahip küçük dalgalar olarak tanımlanabilir. Bir işlevi veya bir veri kümesini hem zaman hem de frekansta yerleştirebilen matematiksel işlevlerdir. Dalgacıkların temeli, Haar'ın 1909'daki tezine dayanır [17]. Fourier dönüşümü ile sadece sinyalin frekans bilgisi elde edilir. DD, Fourier dönüşümüne göre sinyalin hem frekans bilgisini hem de konum bilgisini yakalar. Daha sonra sinyal, farklı frekans bantlarında sinyali analiz etmek için kullanılan yaklaşık ve detay bilgilerine ayrıştırılır. DPD, sinyalin DD kıyasla daha fazla filtreden geçirildiği bir tür DD'dir. DD'de her seviyede sadece yaklaşık katsayısı ayrıştırılır. DPD'de ise hem yaklaşık hem de detay katsayısı kendi detay ve yaklaşık katsayısını elde etmek için ayrıştırılır. DPD, DD ile karşılaştırıldığında sinyalin ayrıştırılması için frekans çözünürlüğü üzerinde daha iyi kontrol sağlar [18]. Şekil 4'de görüldüğü gibi DPD yöntemi, kök düğümü  $\Omega_{0,0}$  (orijinal sinyal) olan bir alt uzay ağacı olarak düşünülür. Genel olarak, DPD ağaç alt uzayının kökü  $\Omega_{j,k}$  ile temsil edilir, burada  $j$  ölçeği,  $k$  ise mevcut ölçek içindeki alt bant indeksini gösterir. Ayrıca, bu kök düğüm  $\Omega_{0,0}$ , yaklaşık uzayına ( $\Omega_{j,k} \rightarrow \Omega_{j+1,2k}$ ) ve detay uzayına ( $\Omega_{j,k} \rightarrow \Omega_{j+1,2k+1}$ ) ayrıştırılır. Bu,  $\Omega_{j,k}$ 'nin  $\{\phi_j(t - 2^j)\}_{k \in \mathbb{Z}}$  ortogonal tabanlarını,  $\Omega_{j+1,2k}$ 'nin  $\{\phi_{j+1}(t - 2^{j+1})\}_{k \in \mathbb{Z}}$  ve  $\{\psi_{j+1}(t - 2^{j+1})\}_{k \in \mathbb{Z}}$  iki ortogonal tabanına bölerek gerçekleştirilir. Aşağıdaki denklemlerde gösterildiği gibi  $\phi$  ve  $\psi$  sırasıyla ölçekleme ve dalgacık fonksiyonlarıdır [19]:

$$\phi(k) = \sqrt{2} \sum_n h(n) \phi(2k - n) \quad (1)$$

$$\psi(k) = \sqrt{2} \sum_n g(n) \phi(2k - n) \quad (2)$$

Burada  $n$  örnek sayısını,  $h$  düşük geçiren filtreyi ve  $g$  yüksek geçiren filtreyi temsil eder. Alçak geçiren filtreler yaklaşık katsayılarını ( $A$ ) oluştururken, yüksek geçiren filtre çıkışında ise detay katsayıları ( $D$ ) elde edilir. Bu durum aşağıdaki denklemler ile tanımlanır.

$$A_{j+1}(k) = \sum_n h(n - 2k) A_j(n) \quad (3)$$

$$D_{j+1}(k) = \sum_n g(n - 2k) A_j(n) \quad (4)$$

## 2.2. Özellik Çıkarımı

Bu aşamada, sinyal işleme sonucunda elde edilen ve 32 frekans bant aralığında bulunan DPD katsayılarına özellik çıkarım yaklaşımları uygulanarak KD verilerini temsil eden ayırt edici özellikler çıkarılır. Bu amaç için literatürde yaygın olarak kullanılan 6 farklı özellik çıkarım yaklaşımı DPD katsayılarına uygulanmıştır. Bu özellik çıkarım yaklaşımları: Enerji (Öz1), Shannon entropi (Öz2), Log-enerji entropi (Öz3), Standart sapma (Öz4), Skewness (Öz5) ve Kurtosis (Öz6)'dir. Aşağıda verilen bu tekniklere ait eşitlikler kullanılarak DPD katsayılarından her bir frekans bandını temsil eden bir özellik vektörü elde edilir.

$$\text{Öz1 } E = \sum_{j=1}^N |C_{ij}|^2 \quad (5)$$

$$\text{Öz2 } SE = - \sum_{j=1}^N C_{ij}^2 \log(C_{ij}^2) \quad (6)$$

$$\text{Öz3 } LEE = \sum_{j=1}^N \log(C_{ij}^2) \quad (7)$$

$$\text{Öz4 } \sigma = \left( \frac{1}{N} \sum_{j=1}^N (C_{ij} - \mu_i)^2 \right)^{\frac{1}{2}} \quad (8)$$

$$\text{Öz5 } Sk = \sqrt{\frac{1}{6N} \sum_{j=1}^N \left( \frac{C_{ij} - \mu_i}{\sigma_i} \right)^3} \quad (9)$$

$$\text{Öz6 } Kr = \sqrt{\frac{N}{24} \left( \frac{1}{N} \sum_{j=1}^N \left( \frac{C_{ij} - \mu_i}{\sigma_i} \right)^4 - 3 \right)} \quad (10)$$

Bu denklemlerde  $C_{ij}$  dalgacık katsayılarıdır ve  $\mu$  ise dalgacık katsayılarının ortalama değeridir. Bu aşama sonucunda her bir özellik çıkarım yaklaşımı için 32 ayrı özellik elde edilir. 6 özellik çıkarımı yaklaşımı sonucundan elde edilen özellikler sırasıyla eklenerek  $1 \times 192$  boyutunda sonuç özellik vektörü elde edilir. Böylece özellik çıkarma aşaması sayesinde hem ayırt edici özellikler elde edilir hem de özellik uzayının boyutsallığı azaltılır.

## 2.3. Özellik Seçimi

Özellik seçimi, sınıflandırma problemlerinde giriş veri kümesi içerisinde sonucu olumsuz yönde etkilemeyen verilerin çıkarılması olarak tanımlanabilir. Bunun yanı sıra özellik seçimi işlemi kullanılarak sınıflandırıcı başarımının artırılması beklenir. Böylece, sınıflandırıcı başarımı artırılırken sınıflandırıcının işlem yükü azaltılır ve sınıflandırma hızı da artırılmış olunur. Bu çalışmada, özellik çıkarımı aşamasında elde edilen  $1 \times 192$  boyutunda sonuç özellik vektörünün veri boyutunu azaltarak sınıflandırıcı performansını arttırmak için ReliefF özellik seçimi yöntemi kullanılmıştır. Özellik seçimi algoritmalarının en başarılılarından biri olan ReliefF algoritması, 1992

yılında Kira ve Rendell tarafından tanıtılmıştır [20]. 1994 yılında bu algoritma, Kononenko tarafından çok sınıflı problemler için yeniden geliştirilmiştir [21]. ReliefF algoritması, farklı kategorilerdeki her örnekten en yakın komşu örnekleri seçerek birden çok problem sınıfını çözmeyi amaçlar. Algoritma bu işi yaparken, aynı sınıfın komşularına farklı değerler veren öngörücüleri cezalandırırken farklı sınıfların komşularına farklı değerler verenleri ise ödüllendirir [22]. Özellik seçimini başarılı bir şekilde tamamlayabilen bu algoritma aynı zamanda veri türlerinin özelliklerini sınırlamaksızın yüksek verimlilikler sunar. ReliefF aşağıda verilen eşitlik ile tanımlanır.

$$W_f^{i+1} = W_f^i + \sum_{c \neq \text{class}(x)} \frac{\frac{p(x)}{1 - p(\text{class}(x))} \sum_{j=1}^k \text{diff}_f(x, M_j(x))}{m * k} - \sum_{j=1}^k \text{diff}_f(x, H_j(x)) / (m * k) \quad (11)$$

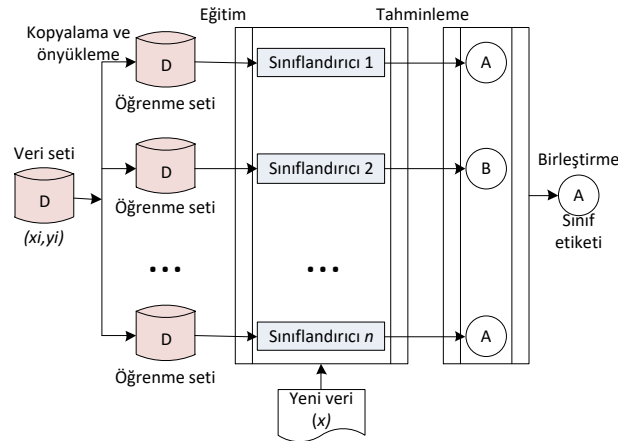
Burada;  $\text{diff}$  iki örnek arasındaki özellik mesafesini,  $H_jx$  örneğinden komşu örnekleri,  $M_j$  farklı sınıfların komşularından komşu örnekleri ve  $p$  ise sınıfın olasılığını temsil eder [23].

## 2.4. Sınıflandırma ve Değerlendirme Metrikleri

Önerilen KD tespit sisteminin son aşaması ikili sınıflandırmaya dayalı tespit aşamasıdır. Bu aşamada sınıflandırıcı ve parametrelerinin seçimi doğrudan tespit sisteminin başarımını belirleyecektir. Bu nedenle özellikle son yıllarda ikili sınıflandırma yaklaşımı işlemlerinde oldukça yaygın olarak kullanılan TÖA sınıflandırıcı olarak belirlenmiştir.

Tek bir öğrenme algoritmasına dayalı karar verme süreçleri, kullanılan öğrenme modelinin sınırlamaları nedeniyle hatalı çıktılar üretebilir. Örneğin yüksek boyutlu verilerle uğraşırken, fazla/yetersiz uydurma sorunları ortaya çıkabileceği gibi özellikle birçok girdinin etkili olduğu saha uygulamalarında girdi değişkenlerinde meydana gelen küçük değişimlerin çıktıyı önemli ölçüde etkilemesi ile yanlış tahminlere neden olabilir [24]. Bu gibi sınırlamaların üstesinden gelebilmek için Dietterich, tek bir öğrenme algoritmasının kararlılığını ve tahmin doğruluğunu artırmak amacıyla sınıflandırma, kümeleme ve regresyon problemlerinde topluluk öğrenme kavramını tanıtmıştır [25].

Topluluk öğrenimi, daha doğru ve istikrarlı sonuçlar için birleşik bir bütünsel öğrenme modeli elde etmek üzere birden fazla tekli öğrenme modelini birleştirebilir. Bu tek öğrenme modellerinin her birine "zayıf öğrenen", en sonunda tek tip olarak oluşturulan entegre modele de "güçlü öğrenen" denir [26]. Literatürdeki mevcut birçok çalışmada [27-29], TÖA yaklaşımının, klasik tek sınıflandırıcının kullanıldığı makine öğrenmesi yöntemlerine göre daha yüksek doğruluk oranına sahip sınıflandırma performansı gösterdiği görülmüştür. Şekil 5'de TÖA sınıflandırıcının genel yapısı gösterilmiştir.



Şekil 5. TÖA sınıflandırıcının genel yapısı.

TÖA sınıflandırıcı kendi içerisinde torbalama (bagging), yükseltme (boosting), yığılma (stacking) ve oylama (voting) gruplarına ayrılmaktadır [30]. Torbalama yöntemi en çok bilinen TÖA yaklaşımlarından biridir. Bu yöntemde, önyükleme (bootstrap) denilen yöntem ile tüm veri seti içerisinde rastgele ve tekrarlayabilen örnekler seçilerek farklı eğitim setleri elde edilir. Her bir sınıflandırıcı farklı bir eğitim verisi ile eğitilir. Yeni bir veri sınıflandırılırken, her sınıflandırıcıdan ayrı ayrı sonuçlar elde edilir. Elde edilen sonuçlar arasından oylama yoluyla nihai sınıf etiketi belirlenir. Yükseltme yönteminde öncelikle temel bir öğrenme algoritması seçilir. Seçilen

algoritma, başlangıçta her biri eşit ağırlığa sahip bir örnek alt kümesi üzerinde çalıştırılır ve ortaya çıkan model, örnekleri sınıflandırır. Algoritma, bir önceki sınıflandırma işleminde hatalı sınıflandırılan örneklerin ağırlıkları artırılarak güncellenen eğitim veri setinden yeniden çekilerek oluşturulan yeni bir örnek alt kümesi üzerinde tekrar çalıştırılır. Bu şekilde, aynı temel algoritma tekrar tekrar çalıştırılır. Yinelemeli bir süreç olarak, model sayısında veya doğrulukta bir sınıra ulaşılan kadar öğrenici eklemeye devam edilir. Sonuç olarak, nispeten zayıf bir algoritma güçlendirilmiş olur.

$$H_m(x) = \sum_{m=1}^M y_m h_m(x) \quad [m = 1, \dots, M] \quad (12)$$

Burada  $y_m$ , m. hedef değerini,  $h_m(x)$ , her aşamada kayıp fonksiyonunu en aza indirmek için seçilen temel öğreniciyi temsil eder [31]. Yığılma yönteminde birden fazla sınıflandırma algoritması aynı eğitim seti ile eğitilerek birden fazla sınıflandırma modeli elde edilmiş olur. Her bir modelden elde edilen çıktılar, ara bir katmandaki meta sınıflandırıcıya girdi olarak gönderilir ve elde edilen çıktı sınıf etiketi olarak belirlenir. Oylama yönteminde, birden çok model oluşturulur. Daha sonra bu modeller daha iyi sonuçlar elde etmek için birleştirilir. Bu yöntemde, tek bir veri seti üzerinde birden fazla öğrenme algoritması uygulanabildiği gibi tek bir algoritma yine aynı veri seti üzerinde farklı parametreler ile de kullanılabilir [30].

TÖA sınıflandırıcı çıkışında elde edilen başarımların değerlendirme metrik değerleri elde edilerek önerilen KD tespit yaklaşımının performansı belirlenmiştir. Bu çalışmada literatürde yaygın olarak kullanılan Doğruluk (Accuracy, ACC), Duyarlılık (Sensitivity, SEN), Özgülük (Specificity, SPE), Kesinlik (Precision, PRE) ve F1-skor (F1-score) metrik değerleri elde edilmiştir. Bu değerlendirme metrikleri, bir sınıflandırma sürecinden elde edilen dört ana sonuç kullanılarak belirlenir: Doğru Pozitif (True Positive, TP), Doğru Negatif (True Negative, TN), Yanlış Pozitif (False Positive, FP) ve Yanlış Negatif (False Negative, FN)'dir. [32]'de belirtildiği gibi bu beş metrik aşağıdaki eşitlikler kullanılarak elde edilir.

$$ACC = \frac{TP+TN}{TP+FN+TN+FP} \quad (13)$$

$$SEN = \frac{TP}{TP+FN} \quad (14)$$

$$SPE = \frac{TN}{TN+FP} \quad (15)$$

$$PRE = \frac{TP}{TP+FP} \quad (16)$$

$$F1-score = \frac{2 \times TP}{2 \times TP + FN + FP} \quad (17)$$

### 3. Deneysel Sonuçlar ve Tartışma

Önerilen yaklaşımda KD tespit sistemi olarak tanımlansa da test aşamasının sonucu ikili sınıflandırma şeklindedir. Bu sınıflandırma işleminde '0' normal sinyal ve '1' KD'li sinyal olarak etiketlenmiştir. Tespit sisteminin eğitim ve test işlemlerinde VSB veri seti kullanılmıştır. VSB test veri seti herhangi bir temel sınıf bilgisi içermediğinden daha detaylı bir analiz gerçekleştirilememektedir. Bu nedenle, 8711 veri boyutuna sahip eğitim veri seti, önerilen algoritmanın eğitim ve test aşamalarında kullanılmıştır. VSB eğitim veri seti içerisinde de normal sinyal verileri ve KD'li sinyal veri sayıları arasında ciddi oranda bir fark bulunup KD'li veri sayısı, normal veri sayısından oldukça azdır. Yapılan bu çalışmada ise dengesiz veri seti durumu oluşturmamak için KD'li verilerin tamamı (525) ve normal verilerin ilk 1050 tanesi (KD'li verilerin 2 katı) alınarak sınıflandırma işleminde kullanılmıştır. Alınan bu veri kümesinin % 70'i eğitim, % 30'u test aşamasında kullanılmıştır. Literatürde sınıflandırma işlemlerinde en uygun veri setini oluşturabilmek için KD'li verilere veri parçalama ve yapay veri üretme gibi farklı bazı yöntemler uygulanmaktadır [33,34].

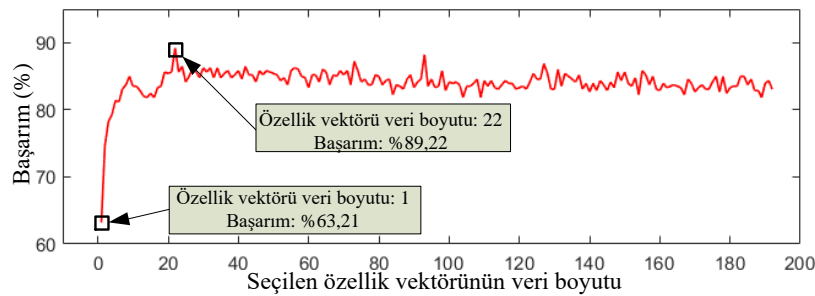


Çalışma kapsamında kullanılan TÖA mimarisi; DVM, C4.5, k-en yakın komşu ve karar ağacı yöntemlerine dayalı dört temel makine öğrenmesi sınıflandırıcısından oluşmaktadır. TÖA yöntemi olarak analizlerde en iyi performans gösteren yükseltme topluluk öğrenmesi yöntemi kullanılmıştır. Önerilen tespit sistemi için gerçekleştirilen ilk analizde TÖA sınıflandırıcı girişine uygun etkili özellik vektörü elde edilmiştir. Bunun için her bir VSB veri seti örneği için DPD katsayıları elde edilerek bu katsayılarla özellik çıkarımı işlemi uygulanmıştır. 6 farklı özellik çıkarım tekniği sonucunda her bir veri için toplam 192 örneğe sahip özellik vektörü elde edilmiştir. Bu özellik vektörüne ReliefF özellik seçimi işlemi uygulanarak ayırt etme performansı yüksek özellikler sıralanmıştır. Böylece ilk analiz sonucunda yüksek sınıflandırma performansına sahip özellikler sıralanmıştır. Tablo 1’de en etkili ilk 22 özelliğe ait bilgiler verilmiştir. Tablo 1’den de görülebileceği gibi ilk sıralardaki etkili özellikler Log-enerji entropi ve standart sapma özellikleridir. Gerçekleştirilen DPD analiz işleminde temel frekans içerden 0-625 kHz bant aralığına ait özellikler 1., 33., 65., 97., 129. ve 161. özelliklerdir. Temel frekans bandının bir üzerindeki bant aralığı ise 625-1250 kHz olup bu bant aralığını temsil eden özellikler ise 2., 34., 66., 98., 130. ve 162. özelliklerdir. Tablo 1’den de görülebileceği gibi seçilen 22 ayırt edici özellik içerisinde bu iki bant aralığındaki özellikler bulunmamaktadır. Seçilen 22 özellik 1,25 MHz üzerindeki yüksek frekanslara ait özelliklerdir.

**Tablo 1.** ReliefF özellik seçimi sonrasında elde edilen ilk 20 özelliğin sınıf ve frekans bant bilgileri

Sırası	No’su	Sınıfı	Sırası	No’su	Sınıfı
1	81	Log-enerji entropi	12	99	Standart sapma
2	103	Standart sapma	13	3	Enerji
3	92	Log-enerji entropi	14	121	Standart sapma
4	71	Log-enerji entropi	15	4	Enerji
5	90	Log-enerji entropi	16	39	Shannon entropi
6	76	Log-enerji entropi	17	181	Kurtosis
7	104	Standart sapma	18	100	Standart sapma
8	68	Log-enerji entropi	19	82	Log-enerji entropi
9	35	Shannon entropi	20	96	Log-enerji entropi
10	36	Shannon entropi	21	67	Log-enerji entropi
11	7	Enerji	22	105	Standart sapma

İkinci analizde ise ilk analizde sıralama yapılan özelliklerin sınıflandırma performansı tespit edilir. Böylece bu analiz sonucunda önerilen tespit sisteminin TÖA sınıflandırıcı girişindeki özellik kümesi belirlenir. Bu analiz işleminde ilk olarak TÖA sınıflandırıcı girişine sadece birinci özellik giriş olarak uygulanmış sınıflandırma performansı elde edilmiştir. Daha sonrasında bu özelliğe ikinci sıradaki özellik eklenerek sınıflandırma sonucu elde edilmiştir. Böylece ardışık olarak son sıradaki özelliğe kadar sırasıyla özellikler eklenerek özellik vektörü oluşturulur ve oluşturulan özellik vektörlerinin sınıflandırma performansları belirlenir. Şekil 6’da, ikinci analiz sonucunda elde edilen sınıflandırma sonuçlarının değişimi gösterilmiştir. Bu sonuçlardan da görülebileceği gibi en iyi başarımla, ReliefF özellik seçimi sonucunda belirlenen ilk 22 özellikten oluşturulan özellik vektörü ile elde edilmiştir. Ayrıca sadece ilk sıradaki Log-enerji entropi özelliği ile sınıflandırma işlemi yapıldığında bile oldukça iyi bir başarımla elde edilmiştir.



**Şekil 6.** ReliefF özellik seçiminden elde edilen özelliklerin TÖA ile sınıflandırma sonuçları.

Üçüncü analiz çalışmasında ise önerilen KD tespit sisteminin başarımına yer verilmiştir. Ayrıca özellik seçimi olmadan bütün özellik çıkarım yöntemleri için başarımlar elde edilmiştir. Şekil 7’de önerilen KD tespit yaklaşımının başarımına ait karar matrisi gösterilmiştir. Bu karar matrisinden görülebileceği gibi önerilen 22 özellik kullanılarak elde edilen doğruluk değeri %89,22’dir. Tablo 2’de ise bu analiz aşamasında elde edilen tüm sonuçlara yer verilmiştir. Değerlendirme metrik değerlerine göre doğruluk değeri en yüksek özellik çıkarım yaklaşımı enerji yöntemidir. Duyarlılık değeri en yüksek olan yöntem ise Log-enerji entropi yöntemidir. İkinci analiz çalışmasında da görülebileceği gibi en etkili özellikler duyarlılık performansı yüksek olan Log-enerji entropi yöntemidir. Ayrıca, bütün özellik çıkarım yöntemlerinden elde edilen özellik vektörünün doğruluk değeri, her bir yöntemin doğruluk değerine göre oldukça yüksektir. Bu analiz sonuçlarından da görülebileceği gibi önerilen özellik seçimi tabanlı KD tespit sistemi, sınıflandırıcı girişinde daha düşük veri boyutuna sahipken başarımların değeri ise bütün yöntemlere göre daha etkilidir.

**Tablo 2.** Önerilen KD tespit sistemi değerlendirme metrik sonuçları

Model	Veri boyutu	Değerlendirme metrikleri				
		ACC	SEN	SPE	PRE	F1-score
Öz1+TÖA	1×32	0,8487	0,9024	0,7241	0,8810	0,8916
Öz2+TÖA	1×32	0,7294	0,9665	0,1931	0,7304	0,8320
Öz3+TÖA	1×32	0,7082	0,8384	0,4138	0,7639	0,7994
Öz4+TÖA	1×32	0,6786	0,7805	0,4483	0,7619	0,7711
Öz5+TÖA	1×32	0,6829	0,9512	0,0759	0,6996	0,8062
Öz6+TÖA	1×32	0,6490	0,7744	0,3655	0,7341	0,7537
Öz1+ Öz2+ Öz3+ Öz4+ Öz5+ Öz6+TÖA	1×192	0,8647	0,9543	0,6621	0,8646	0,9072
<b>ReliefF+TÖA (Önerilen tespit sistemi)</b>	<b>1×22</b>	<b>0,8922</b>	<b>0,9390</b>	<b>0,7862</b>	<b>0,9086</b>	<b>0,9235</b>

	Normal	KD’li sinyal	Doğruluk oranı	Hata oranı
Normal	308	20	%93,9	%6,1
KD’li sinyal	31	114	%78,62	%21,38
Doğruluk			%89,22	

**Şekil 7.** ReliefF özellik seçiminden elde edilen 22 özellik için elde edilen başarımların sonuçları**Tablo 3.** Literatürde önerilen KD tespit sistemleri ile karşılaştırma

Makale	Yöntem	Veri sayısı	Veri oranı (%) (Eğitim/Test)	ACC
[6]	Tekrarlayan sinir ağı	8711	60/20 (20% doğrulama)	0.8820
[34]	Uzun kısa süreli bellek	8711	80/20	0.7876
[35]	Anomali tabanlı teknik + Destek vektör makine	8711	82/18	0.7984
[35]	Anomali tabanlı teknik + Temel bileşenler analizi	8711	82/18	0.7928
[36]	Rastgele orman	8711	75/25	0.8920
[37]	Çift yönlü kısa süreli bellek	8500	76/24	0.8880
[37]	Konvolüsyonel sinir ağı + Çift yönlü uzun kısa süreli bellek	8500	76/24	0.9389
<b>Bu çalışma</b>	<b>DPD + ReliefF + TÖA</b>	<b>1575</b>	<b>70/30</b>	<b>0.8922</b>

Son analiz çalışmasında ise bu makalede elde edilen KD tespit doğruluk sonuçları ile literatürde VSB veri seti kullanılarak elde edilen benzer çalışmalarda elde edilen doğruluk sonuçları karşılaştırılmıştır. Tablo 3’te, gerçekleştirilen karşılaştırma işlemi sonuçlarına yer verilmiştir. Karşılaştırma sonuçlarından da görülebileceği gibi bu çalışmada sunulan KD tespit sistemi, son yıllarda yapılan birçok çalışmaya göre oldukça yüksek bir başarımlara sahiptir. Bu çalışmalar içerisinde işlem yükü ve karmaşıklığı yüksek olan derin öğrenme yaklaşımları da

bulunmaktadır. Önerilen KD tespit sistemi, derin öğrenme yaklaşımlarına göre işlem karmaşıklığı açısından oldukça sade bir yapıya sahip olup doğruluk değeri de yüksektir. Bunun yanı sıra son yıllarda literatürde doğruluk oranı 0,90'ın üzerinde sonuçlara sahip bazı çalışmaların yer aldığı görülmektedir. Bu çalışmalar incelendiğinde özellikle işlem yükü ve karmaşıklığı olan ileri sinyal işleme ve derin öğrenme yaklaşımlarının kullanıldığı gözlemlenmiştir. İlerleyen çalışmalarda ise DD tabanlı derin öğrenme yaklaşımları kullanılarak KD tespit sisteminin geliştirilmesi de planlanmaktadır.

#### 4. Sonuçlar

Bu makale, enerji iletim hatlarında meydana gelen KD arıza tespiti alanındaki araştırma boşlukları konusunda araştırmacılara ufuk açıcı bilgiler sunmaktadır. Elektrik enerjisi endüstrisinde hızlı ve etkili arıza tespiti, ekonomik kayıpların en aza indirilmesinin yanı sıra hizmet kalitesinin en üst düzeye çıkarılmasına yol açacaktır. Böylece bu makalede tespit performansı yüksek ve birçok makalede sunulan sonuçlardan daha etkili bir yapıya sahip DPD, ReliefF ve TÖA yöntemlerine dayalı bir KD tespit sistemi önerilmiştir. Önerilen KD tespit yaklaşımı, sinyal işleme, özellik çıkarımı, özellik seçimi ve sınıflandırma olmak üzere dört aşamadan oluşur. Sunulan yaklaşımda sistemin tanıma performansının yüksek olması için her bir aşamada oldukça önemli yöntemler ve çalışmalar yürütülmüştür. DPD yöntemi ile ham KD verileri içerisinde eşit ve 625'er Hz artan 32 bant aralığında normal ve KD'li verileri birbirinden ayırt edebilecek frekans bileşenlerine ait DPD katsayıları elde edilmiştir. Elde edilen bu katsayılar ise literatürde yaygın kullanılan 6 farklı özellik çıkarım yöntemi uygulanarak veri boyutu 1/4167 oranında indirgenmiş ve KD verilerini tespit eden  $1 \times 192$  veri boyutuna sahip etkili bir özellik vektörü elde edilmiştir. ReliefF yöntemi kullanılarak gerçekleştirilen özellik seçimi aşamasında ise önerilen sistemin başarımını daha da arttıracak işlemler gerçekleştirilmiştir. Veri boyutu  $1 \times 22$  ye düşürülerek bir özellik vektörü oluşturulmuştur. Bu özellik vektörü ise önerilen sistemin işlem yükünün azaltılmasında ve tespit performansının artırılmasında önemli bir rol oynamaktadır. Son aşamada ise TÖA dayalı sınıflandırma işlemi gerçekleştirilerek tespit sonuçları elde edilmiştir. Önerilen KD tespit sisteminin performansını test etmek için VSB veri seti kullanılmış ve 89,22% doğruluk oranı ile yüksek bir tespit başarımına ulaşılmıştır. Elde edilen sonuçlara göre önerilen yaklaşım etkili bir yapıya sahip olduğu görülmüştür.

#### Kaynaklar

- [1] Elmasry, W., Wadi, M. EDLA-EFDS: A novel ensemble deep learning approach for electrical fault detection systems. *Electric Power Systems Research* 2022; 207: 107834.
- [2] Elmasry, W., Wadi, M. Detection of faults in electrical power grids using an enhanced anomaly-based method. *Arabian Journal for Science and Engineering* 2022; 47(11): 14899–14914.
- [3] Wang, W., Yu, N. Partial discharge detection with convolutional neural networks. In 2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS) August 2020; IEEE. pp. 1–6.
- [4] Xu, N., Gooi, H. B., Wang, L., Zheng, Y., Wang, W., Yang, J. Loop Optimization Noise-Reduced LSTM Based Classifier for PD Detection, *IEEE Transactions on Industry Applications* 2023; 59(1): 392–402.
- [5] Xi, Y., Tang, X., Li, Z., Shen, Y., Zeng, X. Fault detection and classification on insulated overhead conductors based on MCNN-LSTM. *IET Renewable Power Generation* 2022; 16(7): 1425–1433.
- [6] Huang, C., Ding, S., Li, S., Liu, R. LMFE: Learning-Based Multiscale Feature Engineering in Partial Discharge Detection. *IEEE Transactions on Neural Networks and Learning Systems* 2022; doi: 10.1109/TNNLS.2022.3222671.
- [7] Bajwa, B., Butani, C., Patel, C. A novel approach towards predicting faults in power systems using machine learning. *Electrical Engineering* 2022; 104: 363–368.
- [8] Michau, G., Hsu, C. C., Fink, O. Interpretable detection of partial discharge in power lines with deep learning. *Sensors* 2021; 21(6): 2154.
- [9] Tehrani, P., Levorato, M. Frequency-based multi task learning with attention mechanism for fault detection in power systems. In 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm) November, 2020; IEEE. pp. 1–6.
- [10] Ahmad, D., Wang, S., Alam, M. Long short term memory based deep learning method for fault power line detection in a MV overhead lines with covered conductors. In 2020 21st National Power Systems Conference (NPSC) December 2020; IEEE. pp. 1–4.
- [11] Kalanidhi, K., Baskar, D., Kumar, V. Transmission Power Line Fault Detection using Convolutional Neural Networks. In Proceedings of the First International Conference on Computing, Communication and Control System, I3CAC 2021, 7-8 June 2021; Bharath University, Chennai, India.
- [12] Qu, N., Li, Z., Zuo, J., Chen, J. Fault detection on insulated overhead conductors based on DWT-LSTM and partial discharge. *IEEE Access* 2020; 8: 87060–87070.

- [13] Wadi, M. Fault detection in power grids based on improved supervised machine learning binary classification. *Journal of Electrical Engineering* 2021;72(5): 315–322.
- [14] VSB Power Line Fault Detection, Kaggle, 2018, (<https://www.kaggle.com/c/vsb-power-line-fault-detection/data>).
- [15] ENET Centre, VSB, 2020, (<https://cenet.vsb.cz/en/>).
- [16] Chui, C. K.. An introduction to wavelets (Vol. 1). Academic press, 1992.
- [17] Daubechies, I. Ten lectures on wavelets. Society for industrial and applied mathematics, 1992.
- [18] Desai, R., Porob, P., Rebelo, P., Edla, D. R., Bablani, A. EEG data classification for mental state analysis using wavelet packet transform and Gaussian process classifier. *Wireless Personal Communications* 2020;115(3): 2149–2169.
- [19] Khushaba, R. N., Kodagoda, S., Lal, S., Dissanayake, G. Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. *IEEE transactions on biomedical engineering* 2010; 58(1): 121–131.
- [20] Kira, K., Rendell, L. A. (1992, July). The feature selection problem: Traditional methods and a new algorithm. In *Aaii* 1992: 2(1992a); 129–134.
- [21] Kononenko, I. Estimating attributes: Analysis and extensions of RELIEF. In *ECML*, April 1994: 94, pp. 171–182.
- [22] Tuncer, T., Ertam, F. Neighborhood component analysis and reliefF based survival recognition methods for Hepatocellular carcinoma. *Physica A: Statistical Mechanics and its Applications* 2020; 540: 123143.
- [23] Özyurt, F. Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures. *The Journal of Supercomputing* 2020; 76(11) : 8413–8431.
- [24] Kapucu, C., Cubukcu, M. A supervised ensemble learning method for fault diagnosis in photovoltaic strings. *Energy* 2021; 227: 120463.
- [25] Dietterich, T. G. Ensemble methods in machine learning. In *Multiple Classifier Systems: First International Workshop 2000, MCS 2000 Cagliari, Italy, June 21–23*, pp. 1–15.
- [26] He, L., Cheng, Y., Li, Y., Li, F., Fan, K., Li, Y. An improved method for soil moisture monitoring with ensemble learning methods over the Tibetan plateau. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 2021; 14: 2833–2844.
- [27] Yu, L., Wang, S., Lai, K. K. Credit risk assessment with a multistage neural network ensemble learning approach. *Expert systems with applications* 2008; 34(2) : 1434–1444.
- [28] Yu, Z., Chen, H., You, J., Wong, H. S., Liu, J., Li, L., Han, G. Double selection based semi-supervised clustering ensemble for tumor clustering from gene expression profiles. *IEEE/ACM transactions on computational biology and bioinformatics* 2014; 11(4) : 727–740.
- [29] Wang, H., Ma, J. Wang G., Hao J., Ma J., Jiang H. A comparative assessment of ensemble learning for credit scoring, *Expert Systems with Applications* 2011; 38(1) : 223–230.
- [30] Akram, V. K., Taşer, P. Y. Telsiz Duyurga Ağlarda Bizans Saldırılarının Topluluk Öğrenme-tabanlı Tespiti. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi* 2020; 22(6) : 905–918.
- [31] Akcan, F., Sertbaş, A. Topluluk Öğrenmesi Yöntemleri ile Göğüs Kanseri Teşhisi. *Electronic Turkish Studies*( 2021; 16(2): 511–528.
- [32] Hossin, M., Sulaiman, M. N. A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process* 2015; 5(2) : 1–11.
- [33] N. V. Chawla, N. Japkowicz and A. Kotcz. Special issue on learning from imbalanced data sets. *ACM SIGKDD Explorations Newsletter* 2004; 6(1): 1–6.
- [34] Dong, M., Sun, J. Partial discharge detection on aerial covered conductors using time-series decomposition and long short-term memory network. *Electric Power Systems Research* 2020; 184: 106318.
- [35] Wadi, M., Elmasry, W. An anomaly-based technique for fault detection in power system networks. In *2021 International Conference on Electric Power Engineering–Palestine (ICEPE–P) 2021*; IEEE. pp. 1–6.
- [36] Vantuch, T., Prilepok, M., Fulneček, J., Hrbáč, R., Mišák, S. Towards the text compression based feature extraction in high impedance fault detection. *Energies* 2019;12(11): 2148.
- [37] Li, Z., Qu, N., Li, X., Zuo, J., Yin, Y. Partial discharge detection of insulated conductors based on CNN-LSTM of attention mechanisms. *Journal of Power Electronics* 2021; 21: 1030–1040.