



RECOGNITION OF POWER QUALITY EVENTS USING WAVELET TRANSFORM, K-NEAREST NEIGHBOR ALGORITHM AND GAIN RATIO FEATURE SELECTION METHOD

Düzgün Akmaz^{*1} 

¹Munzur Üniversitesi, Elektrik-Elektronik Mühendisliği, 62000-Tunceli, Türkiye

Abstract

Original scientific paper

In this study, a method was applied to classify Power Quality (PQ) disturbance signals. In this method, a high accuracy classification success was aimed with a simple approach. In the method, Discrete Wavelet Transform (DWT) was applied to PQ signals containing noise and the signals were decomposed. Then, 80 different classification features were obtained by using different entropy values of the signals. The use of all these 80 features for classification leads to a longer training/testing time of the method. Therefore, it is necessary to determine the effective features between all classification features. In this study, 10 effective classification features were determined by the Gain Ratio (GR) feature selection method and K-Nearest Neighbor Algorithm (K-NN). K-NN was trained and tested with 10 determined classification features. The simulation results showed that the applied method was effective for classification. At the same time, the simulation results showed that 10 classification features determined by GR for the selected parameters provided a higher classification success in a shorter time than 80 classification features.

Keywords: Discrete wavelet transform, gain ratio feature selection method, K-nearest neighbour algorithm, power quality.

GÜÇ KALİTESİ OLAYLARININ DALGACIK DÖNÜŞÜMÜ, K-EN YAKIN KOMŞU ALGORİTMASI VE KAZANÇ ORANI ÖZELLİK SEÇME YÖNTEMİ KULLANILARAK TANINMASI

Özet

Orijinal bilimsel makale

Bu çalışmada Güç Kalitesi (GK) bozulumu sinyallerinin sınıflandırılması için bir yöntem uygulanmıştır. Bu yöntemde basit bir yaklaşımla yüksek doğrulukta bir sınıflandırma başarısı amaçlanmıştır. Yöntemde, Ayrık Dalgacık Dönüşümü (ADD), gürültü içeren GK sinyallerine uygulanarak sinyaller ayrıştırılmıştır. Daha sonra sinyallerin farklı entropi değerleri kullanılarak 80 farklı sınıflandırma özelliği elde edilmiştir. Bu 80 özelliğin hepsinin sınıflandırma için kullanılması yöntemin eğitim/test sürelerinin uzamasına yol açmaktadır. Bu yüzden bütün sınıflandırma özellikleri arasında etkili özelliklerin belirlenmesi gereklidir. Bu çalışmada Kazanç Oranı (KO) özellik seçme yöntemi ve K-En Yakın Komşu Algoritması (K-EYK) ile etkili 10 sınıflandırma özelliği belirlenmiştir. Belirlenen 10 sınıflandırma özelliği ile K-EYK eğitilmiş ve test edilmiştir. Benzetim sonuçları gerçekleştirilen yöntemin sınıflandırma için etkili olduğunu göstermiştir. Aynı zamanda benzetim sonuçları seçilen parametreler için KO ile belirlenen 10 sınıflandırma özelliğinin, 80 sınıflandırma özelliğine göre daha kısa sürede daha yüksek bir sınıflandırma başarısı sağladığını göstermiştir.

Anahtar Kelimeler: Ayrık dalgacık dönüşümü, kazanç oranı özellik seçme yöntemi, K-en yakın komşu algoritması, güç kalitesi.

1 Giriş

Günümüzde artan enerji ihtiyacı doğrultusunda birçok farklı enerji kaynağı güç sistemlerine dahil edilmektedir. Bu durum güç sistemlerindeki sinyallerin giderek bozulmasına ve düşük Güç Kalitesine (GK) neden olmaktadır. Düşük GK istenmeyen bir durum olup sistemdeki cihazlara zarar verebilmektedir. Bu doğrultuda birçok sinyal işleme yöntemi, özellik seçme metodu ve sınıflandırma algoritması GK sinyallerinin analizi için uygulanmıştır [1-4].

[1]'de Ayrık Dalgacık Dönüşümü (ADD) ve sinir ağları kullanılarak GK sinyalleri sınıflandırılmıştır. [2]'de Hiperbolik- Stocwell dönüşümü sinyal işleme aracı ve karar ağacı sınıflandırma algoritması kullanılmıştır. [3]'te Stocwell - dönüşümü/Varyasyon kip ayrıştırması sinyal işleme yöntemi ve destek vektör makineleri, GK bozulmalarının sınıflandırılması için kullanılmıştır. [4]'te Hilbert Huang dönüşümü ve ağırlıklı çift yönlü aşırı öğrenme makinesi, GK sinyallerinin sınıflandırılması için uygulanmıştır.

* Corresponding author.

E-mail address: dakmaz@munzur.edu.tr (D. Akmaz)

Received 04 November 2021; Received in revised form 15 December 2021; Accepted 21 January 2022

2587-1943 | © 2022 IJIEA. All rights reserved.

Doi: <https://doi.org/10.46460/ijiea.1019012>

Aynı zamanda GK sinyallerinin sınıflandırılması için geliştirilen bazı çalışmalarda farklı özellik seçme yöntemleri de kullanılmıştır [2,3]. Özellik seçme yöntemlerinde birçok farklı sınıflandırma özellikleri arasında sınıflandırma için etkili özellikler seçilmektedir. Böylece daha az sınıflandırma özelliği ile etkili bir sınıflandırma yapılabilmektedir. [2]'de genetik algoritma, [3]'te sıralı ileri seçim/sıralı geriye doğru seçim ve Gram-Schmidt ortogonalleşmesi, özellik seçme yöntemi olarak kullanılmıştır.

Bu çalışmada GK sinyallerinin sınıflandırılması için bir yöntem geliştirilmiştir. Yöntemde ilk olarak ADD, gürültü içeren GK sinyallerine uygulanarak sinyaller ayrıştırılmıştır. Daha sonra sınıflandırma özelliklerini elde etmek için sinyallerin enerji, shannon, log enerji, standart sapma, norm, ortalama, çarpıklık ve basıklık entropi değerleri alınmıştır. Bu entropi değerlerinin alınması ile 80 farklı sınıflandırma özelliği elde edilmiştir. 80 farklı sınıflandırma özelliği arasından etkili sınıflandırma özelliklerini belirlemek için Kazanç Oranı (KO) özellik seçme yöntemi ve k-En Yakın Komşu (K-EYK) algoritması kullanılmıştır. Bu algoritmaların kullanılması ile belirlenen 10 sınıflandırma özelliği ile en yüksek sınıflandırma başarısı elde edilmiştir. Benzetim sonuçları, seçilen parametreler için ADD sinyal işleme yöntemi, KO özellik seçme yöntemi ve K-EYK sınıflandırma algoritmasının başarılı olduğunu göstermiştir. Aynı zamanda benzetim sonuçları, 80 sınıflandırma özelliği yerine belirlenen 10 sınıflandırma özelliği kullanılarak daha yüksek bir sınıflandırma başarısı elde edildiğini göstermiştir.

Daha önce benzer parametrelerde ADD ve K-EYK yöntemi [5]'te GK bozulmuş sinyallerinin sınıflandırılması için uygulanmıştır. Fakat [5]'teki çalışmada yalnızca Shannon, Log enerji ve Norm entropi özelliklerinin uygulanan yöntemdeki başarısı incelenmiştir. Tamamlanan çalışmada herhangi bir özellik seçme yöntemi kullanılmamıştır. [5]'te 10 Shannon entropi değeri kullanılarak en yüksek sınıflandırma başarısı elde edilmiştir. Bu çalışmada ise 80 farklı sınıflandırma özelliği arasından KO yöntemi ile etkili olan 10 sınıflandırma özelliği belirlenmiştir. Elde edilen sonuçlar, bu çalışmadaki yöntemin sınıflandırma açısından [5]'teki yöntemden daha başarılı olduğu göstermiştir. Ayrıntılı karşılaştırma bölüm 5'te verilmiştir.

2 Ayrık Dalgacık Dönüşümü (ADD)

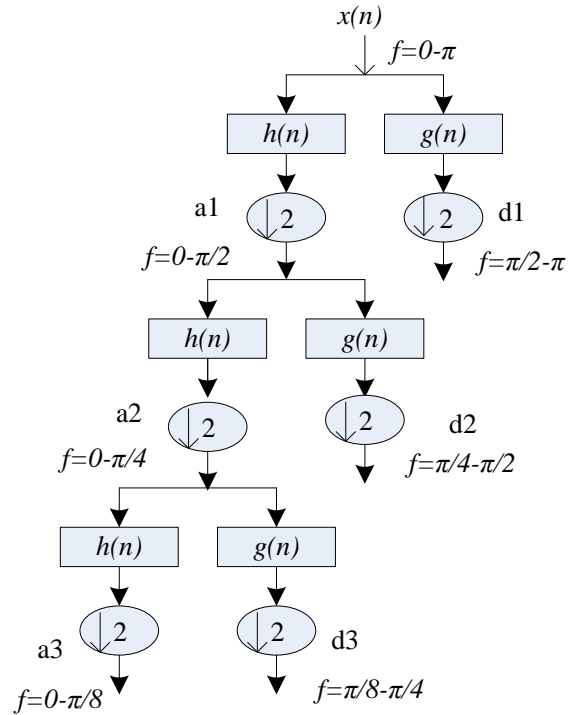
ADD, giriş sinyali $f(k)$ 'nin frekans bandını ilgili düşük ve yüksek frekans bileşenlerine ayırmak için alçak geçiren $h(k)$ ve yüksek geçiren $g(k)$ filtreleri kullanır [6]. Alçak geçiren filtre $h(k)$, ölçekleme fonksiyonundan belirlenir. Yüksek geçiren filtre $g(k)$, hem dalgacık hem de ölçekleme fonksiyonlarından belirlenir. Dalgacık ve ölçekleme fonksiyonları sırasıyla şu şekilde verilmiştir [6]:

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

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

burada n bir tam sayıdır. Alçak geçiren filtreleme A_j yaklaşık katsayılarını üretirken, yüksek geçiren filtreleme ayrıştırmanın D_j detay katsayılarını üretir [6].

ADD çoklu çözünürlük analizi, orijinal sinyalin çeşitli çözünürlük seviyelerinde farklı sinyallere ayrıştırılması esasına dayanır. İlk olarak orijinal sinyal, alçak geçiren ve yüksek geçiren filtreden geçirilir. Böylelikle sinyalin yaklaşık ve detay katsayıları elde edilir. Daha sonra elde edilen yaklaşık sinyale tekrardan alçak ve yüksek geçiren filtreler uygulanarak sinyaller ayrıştırılır. Bu işlem, belirlenen ayrıştırma seviyesine göre tekrarlanır. Böylece sinyaller istenilen frekans aralıklarında incelenebilmektedir. Üç seviyeli ADD uygulaması Şekil 1'de gösterilmiştir.



Şekil 1. Üç seviyeli ayrık dalgacık dönüşümü

Burada

$$x(n) = a_1 + d_1 \quad (3)$$

veya

$$x(n) = a_2 + d_1 + d_2 \quad (4)$$

veya

$$x(n) = a_3 + d_1 + d_2 + d_3 \quad (5)$$

Her ne kadar Şekil 1'de üç seviyeli ADD yöntemi gösterilse de uygulamalara bağlı olarak farklı seviyelerde ADD uygulanabilmektedir.

3 K-En Yakın Komşu Algoritması

Veri madenciliğinde kullanılan sınıflandırma algoritmalarından biri de K-EYK algoritmasıdır. Bu algoritmanın amacı, test veri setinde yer alan ve sınıfı

bilinmeyen bir örneğin eğitim veri setindeki örneklere olan uzaklığını ölçerek sınıflandırmaktır. Bu sınıflandırıcılar, eğitim örneklerini depolar ve sınıflandırma zamanı gelene kadar gerçek bir iş yapmaz. Yöntemin dezavantajı, tüm eğitim veri setinin depolanması için geniş bir alana ihtiyaç duymasındır. Eğitim aşamasında herhangi bir yöntem uygulanmadığı için çoğunlukla gürültülü eğitim verileri, veri tabanını gereksiz yere arttırır [7-8]. K-EYK, aynı uzaklık ölçüsünü kullanan k-en yakın komşu sınıflandırıcısıdır. En yakın komşuların sayısı nesne düzenleyicide açıkça belirtilebilir. En yakın komşuları bulma görevini hızlandırmak için farklı arama algoritmaları kullanılabilir. Kullanılan mesafe fonksiyonu, aramanın bir parametresidir. Öklit, Chebyshev, Manhattan ve Minkowski uzaklık ölçüleri olarak kullanılabilir [7].

4 Kazanç Oranı Özellik Seçme Yöntemi

S, m farklı sınıfa sahip s veri örneklerinden oluşan bir küme olsun. Belirli bir örneği sınıflandırmak için beklenen bilgi şu şekilde verilir [9]:

$$I(S) = -\sum_{i=1}^m p_i \log_2(p_i) \quad (6)$$

burada p_i rastgele bir örneğin C_i sınıfına ait olma olasılığıdır. Entropi veya A'nın alt kümelerine bölünmesine dayanan bilgi, şu şekilde verilir:

$$E(A) = -\sum_{i=1}^m I(S) \frac{S_{1i}+S_{2i}+\dots+S_{mi}}{s} \quad (7)$$

A üzerinde dallanma ile elde edilecek kodlama bilgisi,

$$Gain(A) = I(S) - E(A) \quad (8)$$

C4.5, aşağıda tanımlanan değeri kullanarak, bilgi kazancına normalleştirme uygulayan kazanç oranını kullanır:

$$SplitInfo_A(S) = -\sum_{i=1}^v (|S_i|/|S|) \log_2(|S_i|/|S|) \quad (9)$$

Yukarıdaki değer, eğitim veri seti S'nin, A özneliği üzerindeki bir testin v sonuçlarına karşılık gelen v bölümlerine bölünmesiyle üretilen bilgileri temsil eder [9]. Kazanç oranı şu şekilde tanımlanır [9]:

$$GainRatio(A) = Gain(A)/SplitInfo_{(S)} \quad (10)$$

En yüksek kazanç oranına sahip öznelik, bölme özneliği olarak seçilir.

5 Benzetim Sonuçları

Bu çalışmada sinüs, gerilim çukuru (sag), gerilim tepesi (swell), harmonik (harmonic), geçici rejim gerilim bileşenleri (transient), gerilim kesintisi (interruption), gerilim çukuru ile harmonik (sag with harmonic), gerilim tepesi ile harmonik (swell with harmonic) ve gerilim kırışması (flicker) sinyalleri incelenmiştir. Yukarıda belirtilen 9 farklı GK sinyallerinin matematiksel denklemleri [10]'da tamamlanan çalışmadan alınmıştır.

Sentetik olarak üretilmiş olan sinyallerin parametreleri [10]'da belirtilen aralıklar içinde olup rastgele seçilmiştir. Tablo 1' de bu sentetik sinyallerin matematiksel denklemleri bulunmaktadır. Her bir GK olayı için 200 farklı benzetim gerçekleştirilmiştir. Toplamda 9 farklı GK sinyali için 1800 farklı arıza meydana gelmiştir. Bu 1800 GK olayının yarısı eğitim için, diğer yarısı ise test için kullanılmıştır. GK sinyalleri gerçekte gürültü barındıran bir yapıya sahiptir. Bu yüzden tüm GK sinyallerine 50dB değerinde gürültü eklenmiştir.

Çalışmada GK sinyallerinin örnekleme frekansı 12.8 kHz olarak seçilmiştir. Her bir GK sinyali, çoklu çözünürlük analizi ile 9 seviyeye ayrıştırılmıştır. Dokuz seviyeli çoklu çözünürlüklü bir ayrıştırma işlemi gerçekleştirilerek frekans bantları; 1. seviye için 12,8–6,4 kHz, 2. seviye için 6,4–3,2 kHz, 3. seviye için 3,2–1,6 kHz, 4. seviye için 1,6–0,8 kHz, 5. seviye için 800–400 Hz, 6. seviye için 400–200 Hz, 7. seviye için 200–100 Hz, 8. seviye için 100–50 Hz ve 9. seviye için 50–25 Hz aralığında seçilerek detay katsayısı sinyalleri incelenebilmektedir.

Çoklu çözünürlük analizi için Daubechies 4 (db4), ana dalgacık olarak kullanılmıştır. Bu 9 seviyeli ayrıştırmanın her bir GK sinyaline uygulanması sonucunda, 9 detay ve 1 yaklaşık sinyal olmak üzere toplamda 10 farklı sinyal elde edilmiştir. Sınıflandırma özelliklerini elde etmek için bu sinyallerin 8 farklı entropi (enerji, shannon, log enerji, standart sapma, norm, ortalama, çarpıklık ve basıklık) değerleri alınmıştır. Bu entropi değerlerinin matematiksel denklemleri Tablo 2' de gösterilmiştir

Böylece her bir GK sinyalinde sınıflandırma için 80 (10 farklı sinyal*8 farklı entropi=80) farklı özellik elde edilmiştir. Sınıflandırma özellikleri elde edildikten sonra K-EYK algoritması ile GK olayları sınıflandırılmıştır. WEKA programı veri madenciliğinde ve farklı güç kalite çalışmalarında güncel olarak kullanılan bir programdır [11,12]. Bu yüzden, bu çalışmada K-EYK algoritması için WEKA programı kullanılmıştır. K-EYK algoritmasındaki en yakın komşu sayısı 1 olarak belirlenmiştir. Uzaklık ölçüsü olarak ise öklid formülü kullanılmıştır. Tablo 3, 80 farklı sınıflandırma özelliği ve K-EYK algoritması ile elde edilen benzetim sonuçlarını göstermektedir.

Bu tablodan görüldüğü gibi 80 farklı sınıflandırma özelliği kullanıldığında test verisindeki başarı %89 olmuştur. WEKA programı sonuçları, modeli eğitim verileri üzerinde test etmek için geçen sürenin 0.19 sn olduğunu gösterirken, verilen test setinde modeli test etmek için geçen sürenin ise 0.28 sn olduğunu göstermiştir. Tablo 3, test başarısının eğitim başarısına göre daha düşük olduğunu göstermiştir.

Çalışmada kullanılan 80 sınıflandırma özelliklerinden bazıları sınıflandırma için gereksizdir. Bu gereksiz verileri elemek ve sınıflandırmada etkili özellikleri belirlemek için KO yöntemi eğitim verilerine uygulanmıştır. K-EYK algoritmasına benzer şekilde KO yöntemi için de WEKA programı kullanılmıştır. KO yöntemin belirlediği özellikler etki değerlerine göre sıralanmıştır. Belirlenen bu özellikler ve K-EYK algoritması kullanılarak sınıflandırma başarısı incelenmiştir. Eğitim ve test verisi için en az veri ile en yüksek sınıflandırma başarısı elde edilmeye çalışılmıştır

Tablo 1. Güç kalitesi olaylarının matematiksel denklemleri.

Sınıf	Güç kalitesi	Güç kalitesi sinyallerinin matematiksel denklemleri	Parametreler
C ₁	Sinüs	$v(t)=\sin(\omega t)$	Frekans=50Hz
C ₂	Gerilim çukuru (Sag)	$v(t)=(1-\alpha(u(t-t_1)-u(t-t_2)))\sin(\omega t)$	$0.1 \leq \alpha \leq 0.9$ $T \leq t_2 - t_1 \leq 9T$
C ₃	Gerilim tepesi (Swell)	$v(t)=(1+\alpha(u(t-t_1)-u(t-t_2)))\sin(\omega t)$	$0.1 \leq \alpha \leq 0.8$ $T \leq t_2 - t_1 \leq 9T$
C ₄	Gerilim kesintisi (Interruption)	$v(t)=(1-\alpha(u(t-t_1)-u(t-t_2)))\sin(\omega t)$	$0.9 \leq \alpha \leq 1$ $T \leq t_2 - t_1 \leq 9T$
C ₅	Harmonik (Harmonic)	$v(t)=(\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t))$	$0.05 \leq \alpha_3 \leq 0.15$ $0.05 \leq \alpha_5 \leq 0.15$ $0.05 \leq \alpha_7 \leq 0.15$ $\sum \alpha_i^2 = 1$
C ₆	Geçici rejim (Transient)	$v(t)=(\sin(\omega t) + \alpha_{osc} \exp(-(t-t_1)/\tau_{osc}))\sin(\omega_{nosc}(t-t_1))$	$\tau_{osc}=0.008-0.04$ sn $\omega_{nosc}=100-400$ Hz
C ₇	Gerilim çukuru ile harmonik (Sag with harmonic)	$v(t)=(1-\alpha(u(t-t_1)-u(t-t_2))) (\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t))$	$0.1 \leq \alpha \leq 0.9$ $T \leq t_2 - t_1 \leq 9T$ $0.05 \leq \alpha_3 \leq 0.15$ $0.05 \leq \alpha_5 \leq 0.15$ $0.05 \leq \alpha_7 \leq 0.15$ $\sum \alpha_i^2 = 1$
C ₈	Gerilim yükselmesi ile harmonik (Swell with harmonic)	$v(t)=(1+\alpha(u(t-t_1)-u(t-t_2))) (\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t))$	$0.1 \leq \alpha \leq 0.8$ $T \leq t_2 - t_1 \leq 9T$ $0.05 \leq \alpha_3 \leq 0.15$ $0.05 \leq \alpha_5 \leq 0.15$ $0.05 \leq \alpha_7 \leq 0.15$ $\sum \alpha_i^2 = 1$
C ₉	Gerilim kırışması (Flicker)	$v(t)=(1+\alpha_f \sin(\beta_f \omega t))\sin(\omega t)$	$\alpha_f=0.1-0.2$ $\beta_f=5-10$ Hz

Tablo 2. Entropi formülleri.

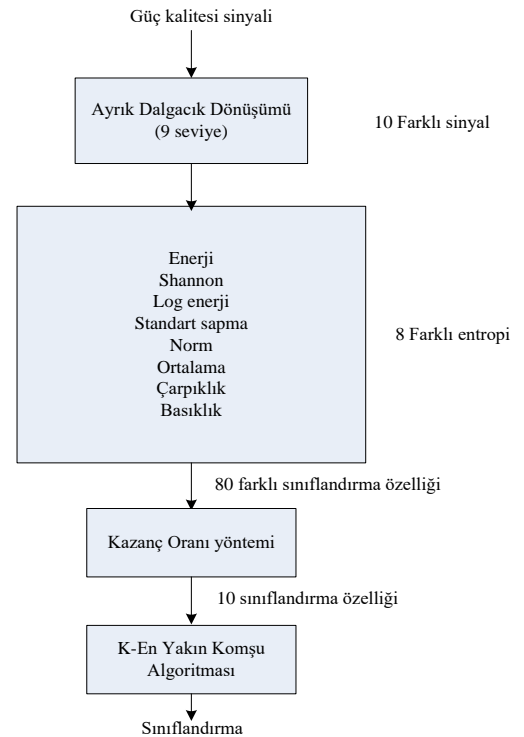
Enerji	$E_{ki} = \sum_{j=1}^N C_{ij} ^2$
Shannon entropi	$SE_{ki} = - \sum_{j=1}^N C_{ij}^2 \log(C_{ij}^2 \log(C_{ij}^2))$
Log	$LOE_{ki} = \sum_{j=1}^N \log(C_{ij}^2)$
Standart sapma	$\sigma_{ki} = \left(\frac{1}{N} \sum_{j=1}^N (C_{ij} - \mu_i)^2 \right)^{1/2}$
Norm	$NE_{ki} = \sum_{j=1}^N (C_{ij})^p$ $1 \leq p$
Ortalama	$\mu_{ki} = \frac{1}{N} \sum_{j=1}^N C_{ij}$
Çarpıklık	$SK_{ki} = \sqrt{\frac{1}{6N} \sum_{j=1}^N \left(\frac{C_{ij} - \mu_i}{\sigma_i} \right)^3}$
Basıklık	$KRT_{ki} = \sqrt{\frac{N}{24} \left(\frac{1}{N} \sum_{j=1}^N \left(\frac{C_{ij} - \mu_i}{\sigma_i} \right)^4 - 3 \right)}$

Tablo 3. 80 farklı sınıflandırma özelliği ve K-EYK yöntemi ile elde edilen benzetim sonuçları.

	Eğitim	Test
Başarı %	100	89
Süre (sn)	0.19	0.28

Tablo 4, KO yönteminin belirlediği farklı özellikler ve K-EYK algoritmasının sınıflandırma başarısını göstermektedir. Benzetimler en az 10 özellik kullanılması

ile test verisinde en yüksek sınıflandırma başarısı sağlandığını göstermiştir. Tablo 5' te ise belirlenen özelliklerin ayrıştırma seviyesi ve entropi tanımları gösterilmiştir. Bu tablodan görüldüğü gibi farklı seviyelerdeki sinyallerin farklı entropi değerleri kullanılmıştır. Bu çalışmadaki sınıflandırma için uygulanan yöntemin genel algoritması Şekil 2' de gösterilmiştir.

**Şekil 2.** Uygulanan sınıflandırma yöntemi adımları.

Tablo 4. KO yönteminin belirlediği özellikler ve K-EYK algoritmasının sınıflandırma başarısı.

Farklı veri sayısı	Eğitim sınıflandırma başarısı %	Test Sınıflandırma başarısı %	Belirlenen özellikler (KO)
1	100	53.44	18
2	100	68.88	18,38
3	100	69.11	18,38,8
4	100	72	18,38,8,19
5	100	72.55	18,38,8,19,9
6	100	75	18,38,8,19,9,49
7	100	75.33	18,38,8,19,9,49,39
8	100	76.55	18,38,8,19,9,49,39,48
9	100	90.88	18,38,8,19,9,49,39,48,44
10	100	98	18,38,8,19,9,49,39,48,44,26
11	100	97.77	18,38,8,19,9,49,39,48,44,26,29
12	100	97.11	18,38,8,19,9,49,39,48,44,26,29,59

Tablo 5. Seçilen sınıflandırma özellikleri.

	Özellik	Seviye ve katsayılar	Entropi
1	18	7. seviye detay	Shannon
2	38	7. seviye detay	Standart sapma
3	8	7. seviye detay	Enerji
4	19	8. seviye detay	Shannon
5	9	8. seviye detay	Enerji
6	49	8. seviye detay	Norm
7	39	8. seviye detay	Standart sapma
8	48	7. seviye detay	Norm
9	44	3. seviye detay	Norm
10	26	5. seviye detay	Log enerji

Tablo 6' da ise belirlenen 10 farklı sınıflandırma özelliği ile elde edilen benzetim sonuçlarını göstermektedir.

Tablo 6. KO ile belirlenen 10 farklı sınıflandırma özelliği ve K-EYK yönteminin benzetim sonuçları.

	Eğitim	Test
Başarı %	100	98
Süre (sn)	0.05	0.06

Tablo 3 ve Tablo 6 incelendiğinde belirlenen 10 özellik ile daha yüksek bir sınıflandırma başarısı sağlandığı belirlenmiştir. 80 özellik ile test verisi üzerinde %89' luk bir başarı sağlanırken, belirlenen 10 özellik ile test verisi üzerinde %98' lik bir sınıflandırma başarısı sağlanmıştır. Aynı zamanda eğitim ve test sürelerinin daha kısa sürede tamamlandığı gösterilmiştir. WEKA programı sonuçları belirlenen 10 sınıflandırma özelliği ile modeli eğitim verileri üzerinde test etmek için geçen sürenin 0.05 sn olduğunu gösterirken, verilen test setinde modeli test etmek için geçen sürenin ise 0.06 sn olduğunu göstermiştir.

Tablo 4' ten görüldüğü gibi KO seçme yöntemi ile en az 10 özellik ile en yüksek sınıflandırma başarısı elde edilmiştir. Fakat bu çalışmada yöntemin başarısını göstermek için aynı zamanda farklı özellik seçme yöntemlerinin de test edilmesi gereklidir. Bu yüzden bu çalışmada ayrıca Relieff, Bilgi kazancı (Information gain) ve ONE-R özellik seçme yöntemleri de kullanılmıştır. Bu özellik seçme yöntemleri ile ilgili ayrıntılı bilgi [13,14]' de bulunabilir. Tablo 7' de farklı özellik seçme yöntemleri ile belirlenen 10 sınıflandırma özelliğinin kullanılması ile elde edilen sınıflandırma sonuçları gösterilmiştir. Bu tablodan görüldüğü gibi en yüksek test sınıflandırma başarısı KO özellik seçme yöntemi ile elde edilmiştir. Bu yüzden çalışmada KO özellik seçme yöntemi kullanılmıştır.

Tablo 7. Farklı özellik seçme yöntemleri ile belirlenen 10 farklı sınıflandırma özelliği ve K-EYK yönteminin benzetim sonuçları.

Özellik seçme yöntemleri	Test başarısı %
KO	98
Relieff	84.66
ONE-R	77.88
Bilgi Kazancı	77.11

Tablo 8' de ise elde edilen sonuçlar [5]' teki çalışma ile karşılaştırılmıştır. Her iki çalışmada GK olayları için benzer veri tabanı kullanılmıştır. ADD ve K-EYK yönteminde aynı parametreler seçilmiştir. Bunun ile birlikte [5]' teki çalışmada yalnızca Shannon, Log energy ve Norm entropi değerlerinin yöntem başarıları incelenmiştir. [5]' te 9 seviyeli ADD uygulanarak 9 detay ve 1 yaklaşık olmak üzere 10 sinyal elde edilmiştir. Bu 10 sinyalin Shannon entropi değeri kullanılarak toplamda 10 sınıflandırma özelliği ile seçilen parametreler için en yüksek sınıflandırma başarısı elde edilmiştir.

Bu çalışmada farklı olarak sınıflandırma için özellik sayısı artırılmıştır. Dokuz seviyeli ADD' nin uygulanması sonrası elde edilen 10 sinyalin, 8 farklı entropi değeri alınarak 80 farklı sınıflandırma özelliği elde edilmiştir. Daha sonra 80 farklı sınıflandırma özelliği içinde KO özellik seçme yöntemi ile en etkili 10 sınıflandırma özelliği belirlenmiştir.

Tablo 8 incelendiğinde bu çalışmadaki test verisi üzerinde %98'lik bir başarı sağlanırken, [5]' teki yöntemde ise test başarısı %94.11 olmuştur. Bu değerler, seçilen parametreler için bu çalışmadaki yöntemin daha yüksek bir sınıflandırma başarısı sağladığını göstermiştir.

Aynı zamanda bu çalışmada elde edilen sonuçlar ile [1]' de tamamlanan çalışma sonuçları karşılaştırılmıştır. Tablo 8' den görüldüğü gibi kullanılan yöntemde daha az sınıflandırma özelliği kullanılarak daha yüksek sınıflandırma başarısı elde edilmiştir. Tablo 8' deki sonuçlar yöntemin GK bozulmalarının sınıflandırılması için kabul edilebilir bir başarı sağladığını göstermektedir.

Tablo 8. Önerilen yöntem karşılaştırılması.

Yöntem	Sinyal işleme yöntemi	Sınıflandırma algoritması	Özellik seçimi	Entropi	Özellik sayısı	Güç kalitesi sinyal sayısı	Test başarısı (50dB) %
[5]	ADD	K-EYK	-----	Shannon	10	9	94.11
[1]	ADD	Sinir Ağı	-----	Norm	13	7	95.11
Tamamlanan	ADD	K-EYK	KO	Tablo 4	10	9	98

6 Sonuçlar

Günümüzde farklı güç kalitesi bozulmaları şebekelerdeki cihazları olumsuz etkilemektedir. Bu yüzden farklı güç kalitesi bozulmalarının belirlenmesi ve sınıflandırılması gereklidir. Bu çalışmada ayrık dalgacık dönüşümü, kazanç oranı özellik seçme yöntemi ve K-En yakın komşu algoritması kullanılarak farklı güç kalitesi bozulmaları sınıflandırılmıştır. Benzetim sonuçları seçilen parametreler için uygulanan yöntemin başarılı olduğunu göstermiştir. Aynı zamanda benzetim sonuçları kazanç oranı özellik seçme yönteminin kullanılması ile daha yüksek bir sınıflandırma başarısının elde edildiğini göstermiştir.

Açıklamalar

Bu çalışmada Etik Kurul Onay belgesine gerek yoktur.

Kaynaklar

- [1] Uyar, M., Yildirim, S., & Gencoglu, M. T. (2008). An effective wavelet-based feature extraction method for classification of power quality disturbance signals. *Electric power systems Research*, 78(10), 1747-1755.
- [2] Ray, P. K., Mohanty, S. R., Kishor, N., & Catalão, J. P. (2013). Optimal feature and decision tree-based classification of power quality disturbances in distributed generation systems. *IEEE Transactions on Sustainable Energy*, 5(1), 200-208.
- [3] Abdoos, A. A., Mianaei, P. K., & Ghadikolaei, M. R. (2016). Combined VMD-SVM based feature selection method for classification of power quality events. *Applied Soft Computing*, 38, 637-646.
- [4] Sahani, M., & Dash, P. K. (2018). Automatic power quality events recognition based on Hilbert Huang transform and weighted bidirectional extreme learning machine. *IEEE Transactions on Industrial Informatics*, 14(9), 3849-3858.
- [5] Akmaz, D. (2021, June). The Effect of Different Entropy Values on the Success of the K Nearest Neighbors Algorithm for Classification of Power Quality Disturbances. *4th International Conference on Data Science and Applications*.

- [6] Erişti, H., Yıldırım, Ö., Erişti, B., & Demir, Y. (2013). Optimal feature selection for classification of the power quality events using wavelet transform and least squares support vector machines. *International Journal of Electrical Power & Energy Systems*, 49, 95-103.
- [7] Vijayarani, S., & Muthulakshmi, M. (2013). Comparative analysis of bayes and lazy classification algorithms. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(8), 3118-3124.
- [8] Choudhury, S., & Bhowal, A. (2015, May). Comparative analysis of machine learning algorithms along with classifiers for network intrusion detection. In *2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)* (pp. 89-95). IEEE.
- [9] Karegowda, A. G., Manjunath, A. S., & Jayaram, M. A. (2010). Comparative study of attribute selection using gain ratio and correlation based feature selection. *International Journal of Information Technology and Knowledge Management*, 2(2), 271-277.
- [10] Moravej, Z., Abdoos, A. A., & Pazoki, M. J. E. P. C. (2009). Detection and classification of power quality disturbances using wavelet transform and support vector machines. *Electric Power Components and Systems*, 38(2), 182-196.
- [11] Vinayagam, A., Veerasamy, V., Radhakrishnan, P., Sepperumal, M., & Ramaiyan, K. (2021). An ensemble approach of classification model for detection and classification of power quality disturbances in PV integrated microgrid network. *Applied Soft Computing*, 106, 107294.
- [12] Kiranmai, S. A., & Laxmi, A. J. (2018). Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy. *Protection and Control of Modern Power Systems*, 3(1), 1-12.
- [13] Budak, H. (2018). Özellik seçim yöntemleri ve yeni bir yaklaşım. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 22, 21-31.
- [14] Yildirim, P. (2015). Filter based feature selection methods for prediction of risks in hepatitis disease. *International Journal of Machine Learning and Computing*, 5(4), 258.