



# Kahramanmaraş Sütçü İmam University

## Journal of Engineering Sciences



Geliş Tarihi : 24.10.2024  
Kabul Tarihi : 10.12.2024

Received Date : 24.10.2024  
Accepted Date : 10.12.2024

### DALGACIK DÖNÜŞÜMÜ TABANLI NAİVE BAYES SINIFLANDIRICI İLE TÜRKİYE’NİN TAMAMINI İÇEREN BÖLGELERDEKİ LİNYİT KÖMÜRÜNÜN KALİTE TESPİTİ

#### QUALITY DETERMINATION OF LIGNITE COAL IN THE REGIONS INCLUDING THE WHOLE OF TURKEY USING WAVELET TRANSFORM BASED NAIVE BAYES CLASSIFIER

Sevcan AYTAÇ<sup>1\*</sup>(ORCID: 0000-0001-6689-2337)

<sup>1</sup> Fırat Üniversitesi, Elektronik ve Otomasyon, Elazığ, Türkiye

\*Sorumlu Yazar / Corresponding Author: Sevcan AYTAÇ, sevcanaytackorkmaz@gmail.com

#### ÖZET

Günümüzde Türkiye sondaj sahalarından elde edilen karmaşık linyit kömürünün kabul edilebilir kalite ve nispeten zayıf kalitede olduğunun tespiti işlemi enerji santrallerinde, diğer alanlarda ve enerji üretiminde hayati öneme sahiptir. Özellikle, birden fazla karmaşık parametrelerinin doğru bir şekilde kalitelerine göre belirlenmesi yatırım kararlarında büyük ölçüde önem kazanmıştır. Bu makalede, Türkiye Kömür İşletmeleri Genel Müdürlüğünden alınan, MTA tarafından yayınlanan linyit envanterinde Türkiye'nin tamamını içeren 96 linyit kömürünün 96 sondaj bölgesine ait nem, kül, kükürt ve kalori içerikleri ele alınmıştır. Belirtilen nem, kül ve kükürt gibi 3 parametre, bağımlı değişken olan kalori değeri üzerinde etkili olmaktadır. Kalori ne kadar yüksekse kömür kalitesi o kadar iyidir. İlk olarak verilere K-Ortalamalar kümeleme algoritması uygulanarak kömürünün kalori değerleri 2 gruba ayrılmıştır. 2 gruba ayrılan bu değerler doğrulama verisi olarak kullanılmıştır. Dalgacık katsayılarından elde edilen özellik değerleri Naive Bayes sınıflandırıcısı ile sınıflandırılmıştır. Sınıflandırma performansları kıyaslandığında Daubechies dalgacık ailesinden olan db4'ün diğer dalgacık ailelerine ve CA dalgacık katsayılarının CH, CV ve CD katsayılarına göre daha yüksek performans sergilediği tespit edilmiştir. Nem, kül ve kükürt içeriklerinin Dalgacık Dönüşümü katsayıları hesaplanarak, Naive Bayes yöntemi ile sınıflandırma performansları kıyaslanmıştır. Daubechies ailesinden olan db4'ün yaklaşık katsayıları ile elde edilen kömürün kalite tespit oranının en yüksek olduğu ve %100 olarak bulunduğu tespit edilmiştir.

**Anahtar Kelimeler:** Linyit kömür, k-ortalamalar, dalgacık dönüşümü, naive bayes sınıflandırıcı

#### ABSTRACT

Today, the process of determining whether the complex lignite coal obtained from Turkey's drilling fields is of acceptable quality or relatively poor quality is of vital importance in power plants, other areas and energy production. In particular, the accurate determination of multiple complex parameters according to their quality has gained great importance in investment decisions. In this article, the moisture, ash, sulfur and calorie contents of 96 drilling regions of 96 lignite coals covering the whole of Turkey in the lignite inventory published by MTA, received from the General Directorate of Turkish Coal Enterprises, are discussed. The 3 parameters mentioned, such as moisture, ash and sulfur, affect the caloric value, which is the dependent variable. The higher the calories, the better the coal quality. First, by applying the K-Means clustering algorithm to the data, the caloric values of coal were divided into 2 groups. These values, divided into 2 groups, were used as validation data. Feature values obtained from wavelet coefficients were classified with the Naive Bayes classifier. When the classification performances were compared, it was determined that db4, which is from the Daubechies wavelet family, showed higher performance than other wavelet families and CA wavelet coefficients compared to CH, CV and CD coefficients. Wavelet Transform coefficients of moisture, ash and sulfur contents were calculated and classification performances were compared with the Naive Bayes method.

ToCite: AYTAÇ, S., (2025). DALGACIK DÖNÜŞÜMÜ TABANLI NAİVE BAYES SINIFLANDIRICI İLE TÜRKİYE’NİN TAMAMINI İÇEREN BÖLGELERDEKİ LİNYİT KÖMÜRÜNÜN KALİTE TESPİTİ. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 28(1), 403-413.

It has been determined that the quality detection rate of coal obtained with the approximation coefficients of db4, which is from the Daubechies family, is the highest and is 100%.

**Keywords:** Lignite coal, k-means, wavelet transform, naive bayes classifier

## GİRİŞ

Nem, kül ve kükürt gibi linyit kalitesini etkileyen değişkenler enerji santrallerinde, diğer alanlarda ve enerji üretiminde hayati öneme sahiptir. Bu nedenle bu değişkenlerin modellenmesi yatırım kararlarında yararlıdır (Moon vd., 2006; Tutmez vd., 2013; Korkmaz ve Esmeray, 2018). Dünya projeksiyonuna benzer şekilde, Türkiye'de bilinen linyit yataklarının çoğu düşük kalori değerine, yüksek oranda kül, uçucu madde, nem ve kükürt içeriklerine sahiptir. Bu nedenle, teknik ve ekonomik olarak işletmeler için uygulanabilir olan bu türdeki linyitlerin çoğunun, pazarlamadan önce yıkanarak zenginleştirilmesi gerekir (Korkmaz ve Esmeray, 2018; Senguler, 2010). Linyit, kömürün alt kademesi olmasına rağmen dünyanın birçok yerinde geniş rezervler edinilebilir. Türkiye dünyanın en önemli linyit üreticilerinden biridir (Korkmaz ve Esmeray, 2018; IEA, 2000). Linyit kömürü, Türkiye ulusal ürününün büyümesine önemli ölçüde katkıda bulunduğu için, linyit alanlarında genel bir değerlendirme yaparak, genel bir bakış açısı sağlamakta ve nispeten teknik ve ekonomik değerlendirmeler yapma olanağı sağlamaktadır. Linyit kalitesine ilişkin parametreler arasındaki ilişkiyi sınıflandırmak ve modellemek için, literatürde birçok deneysel ve matematiksel / istatistiksel teknikler kullanılmıştır (Korkmaz ve Esmeray, 2018; Galetakis vd., 2002).

Bir çalışmada, Hindistan kömürlerini sınıflandırmak için, ilk kez "Kendini Düzenleyen Harita (Self Organization Map-SOM)" olarak bilinen klasik denetimsiz bir kümeleme tekniği olan "K-Ortalamalar Kümeleme" ve Yapay Zeka Tabanlı (Artificial Intelligence-AI) doğrusal olmayan kümeleme yöntemi kullanılmıştır. Bu çalışma ile K-Ortalamaları ve SOM tabanlı sınıflandırma sonuçlarının %75 kömür numunelerinde benzerlik sergilediği bulunmuştur (Korkmaz ve Esmeray, 2018; Pandit vd. 2011).

Başka bir araştırma çalışmasında, Hindistan'ın tüm büyük kömür alanlarından spontan yanmaya karşı yüksek ve düşük duyarlılığa sahip farklı safhalardan 51 kömür numunesi toplanmıştır. Toplanan bu kömürlerin damarlarının sınıflandırılması için bulanık K-Ortalama yönteminin kullanıldığı bir yaklaşım sunulmaktadır. Bu yaklaşımla, kömür numunelerinin nemi, uçucu maddesi, kül içeriği ve geçiş noktası sıcaklığını parametreler olarak kullanarak önerilen algoritma kömür damarlarını üç farklı kategoriye ayırmak için uygulanmıştır. Bu sınıflandırma, saha mühendislerine maden yangınlarının oluşmasını önlemek için önceden iyileştirici tedbirler almak için yararlı olacağı bilgisini vermektedir (Korkmaz ve Esmeray, 2018; Sahu vd., 2012).

Başka bir makalede, Hindistan'ın tüm büyük kömür alanlarından kendiliğinden yanmaya karşı yüksek ve düşük duyarlılık derecelerine sahip olan çeşitli derecelerde 29 kömür numunesi toplanmıştır. Bu makale, kümeleme probleminin doğrusal bir tayin modeli haline getirilmesini ve spontan yanmaya yol açan kömür damarlarının sınıflandırılması için ayırık parçacık sürüsü optimizasyon yaklaşımının uygulanmasını sunmaktadır (Korkmaz ve Esmeray, 2018; Sahu vd., 2011).

Ayrıca, literatürde kömürün nem, kül, kükürt ve kalori değerlerinin sınıflandırılması için çok farklı yaklaşımlar bulunmaktadır. Bu çalışmalardan biri olan makalede Yapay Sinir Ağları, Bulanık Çıkarım Sistemleri, Genetik Algoritmalar ve bunların melez biçimleri gerekli kömür parametrelerin tahmin edilmesi ve öngörülen modellerin geliştirilmesi için kullanılmıştır (Yılmaz vd., 2010).

Kömürlerin Brüt Kalori Değerinin (Gross Calorific Value-GCV) belirlenmesi, kömür karakterizasyonu için çok önemlidir; zor, pahalı, zaman alıcı ve tahrip edicidir. Bu çalışmada GCV'nin öngörülmesi için Yapay Sinir Ağlarının farklı öğrenme algoritmaları olan Çok Katmanlı Algılayıcı (Multi-Layered Perceptron-MLP), Radyal Tabanlı Fonksiyon (Radial Basis Function-RBF)-(tam), RBF (K-Ortalamalar) ve RBF (Kendini Düzenleyen Harita (Self Organization Map-SOM) gibi modelleri kullanılmıştır. Bu makalenin bir sonucu olarak, tüm modeller GCV'yi tahmin etmek için yüksek performans sergilemiştir. Yapay Sinir Ağları (Artificial Neural Network-ANN)'nin dört farklı algoritması neredeyse aynı tahmin etme kabiliyetine sahip olmuş olsa da, MLP'nin doğruluğunun diğer modellerden nispeten daha yüksek olduğu belirtilmiştir (Korkmaz ve Esmeray, 2018; Yılmaz vd., 2010).

Bir araştırmada linyit kömür numunelerinde, Yakın Kızıl Ötesi (Near-infrared-NIR) spektrumlarını kullanarak nem tespiti yapılmıştır. Çalışmadaki model, kısmi En Küçük Kareler Regresyonu (Least Squares Regression-PLS) ve Ayırık Dalgacık Dönüşümü (Discrete Wavelet Transform-DWT) uygulanarak geliştirilmiştir. Öncelikle, NIR

spektrumu, montaj ve sıkıştırma için DWT ile önceden işlendi. Daha sonra, sıkıştırılmış veriler, kömür örneklerinde nem belirlenmesi için PLS ile regresyon modeli oluşturmak için kullanılmıştır. Farklı çözünürlük ölçeklerinde sıkıştırma performansı araştırılmıştır. Sıkıştırılmış veriyi kullanarak, PLS ham spektrumlardan daha doğru sonuç elde edildiği belirtilmiştir. Ayrıca, PLS modelinde de ana bileşen sayısı araştırılmıştır. Sonuçlar, DWT-PLS'nin bitümlü kömür ve linyit kömürün nem analizi için tatmin edici bir tespit performansı elde edebildiğini göstermektedir (Korkmaz ve Esmeray, 2018; Yang vd., 2014).

Başka bir çalışmada, çevrimiçi kömür kalori değer tahmin yöntemine dayalı bir Destek Vektör Regresyonu (Support Vector Regression-SVR) modellenmesi yapılmıştır. Dalgacık Dönüşümü ve PCS işlemi ile kömür alevi radyasyon özellikleri çıkarılmıştır. SVR modellenmesi ile radyasyon karakteristik değişkenleri ile kömür kalori değeri arasındaki ilişki modeli oluşturulmuştur. Doğru SVR yapım parametreleri, Grid hesaplaması ile seçilmiştir. Deney sonuçlarının, SVR modeli tahmini ile kömür kalori değerinin laboratuvar sonuçları arasında iyi bir tutarlılık gösterdiği belirtilmiştir (Korkmaz ve Esmeray, 2018; Cheng vd., 2012).

Bununla birlikte, genel olarak jeosistemlerde kullanılan problem yapılarından bazıları, lineer regresyonda aşına olunan sürekli cevaplar yerine sonuç değişkenleri farklılıklar içerebilir. Bu nedenle, bazı durumlarda, konvansiyonel yaklaşımlar yerine, sonuç çıktılarını için bazı analizler yapılmalıdır. Literatürde şimdiye kadar Linyit Kömürünün kalite tespiti için K-Ortalamlar, Dalgacık Dönüşümü ve Naive Bayes sınıflandırma yöntemleri birlikte kullanılmamıştır. Bu makalede, verilerin arasındaki ilişki ve güçlü yönleri hakkında sınıflandırma yapmak için K-Ortalamlar, Dalgacık Dönüşümü ve Naive Bayes sınıflandırma yöntemleri birlikte kullanılmıştır. Ayrıca, bu makalede Naive Bayes sınıflandırıcısının kullanılmasının nedenlerinden birincisi bu yöntemin tahmine dayalı modelleme yapmak için basit ve bunun yanında güçlü bir algoritma olmasıdır. İkinci ise az sayıdaki verilerle güçlü tahmin kabiliyetine sahip kullanışlı ve avantajlı bir yöntem olmasıdır. Üçüncüsü ise, diğer sınıflandırıcılara göre daha iyi performans sağladığı için bu makalede kullanılmıştır.

Bu çalışmada, linyitin nem, kül ve kükürt içeriği ve tepki değişken kalori değeri gibi belirleyici kalite parametreleri arasındaki ilişkinin modellenmesine odaklanılmıştır. İlk olarak Türkiye'de linyit bölgelerinden elde edilen tüm ortalama kalori değerleri K-Ortalamlar yöntemi ile iki kategoriye ayrılmıştır. İkinci olarak kalori değerlerine göre iki kategoriye ayrılan nem, kül ve kükürt gibi özelliklere Dalgacık Dönüşümü yöntemi uygulanmıştır. Son olarak ise Naive Bayes sınıflandırıcısı yöntemiyle iki kategoriye ayrılan nem, kül ve kükürt değerlerine göre sınıflandırılma başarımları karşılaştırılmıştır. Bu makalenin literatüre katkıları aşağıdaki gibi sıralanmaktadır:

Türkiye Kömür İşletmeleri Genel Müdürlüğünden alınan, MTA tarafından yayınlanan lintit envanterinde Türkiye'nin tamamını içeren 116 bölgede linyit değerleri elde edilmiş olup bu çalışmada eksik veriler çıkarılarak 96 linyit kömürünün 96 sondaj bölgesine ait nem, kül, kükürt ve kalori içerikleri ele alınmış olup daha önceden K-ortalamlar kümeleme algoritması kullanılmamış olup bu makalede, K-ortalamlar kümeleme algoritması ile kalori değerleri 2 gruba ayrılarak bu değerler sınıflandırmada doğrulama verisi olarak kullanılmıştır.

1- Bugüne kadar, Türkiye'nin tamamını içeren bölgelerde bulunan karmaşık nem, kül, kükürt ve kalori gibi linyit kömürü verilerin kabul edilebilir kalite ve nispeten zayıf kalite diye iki sınıfa ayrılması için K-Ortalamlar, Dalgacık Dönüşümü ve Naive Bayes sınıflandırıcı yöntemleri birlikte kullanılmamıştır. Bu yöntemler birlikte kullanılarak bugüne kadar literatürde yapılmayan çok modaliteli farklı bir yöntemle sınıflandırma yapılmıştır.

2- Belirtilen nem, kül ve kükürt gibi 3 parametre, bağımlı değişken olan kalori değeri üzerindeki etkisine bakılarak kömürün bu kalori değerlerine göre çok modaliteli bir yöntemle “kabul edilebilir kalite” ve “nispeten zayıf kalite” olmak üzere 2 kategoriye ayrılmıştır. Kalori ne kadar yüksekse kömür kalitesinin o kadar iyi olduğu tespit edilmiştir.

### **Problem Belirleme**

Linyit kalitesinin değerlendirilmesinde önemli parametrelerden biri olan kalori değeri, santrallerde yakılan kaliteli kömürün önemli bir ölçüsüdür. Kalori değeri genellikle yüksek ısıtma değeri veya brüt kalori değeri olarak belirtilir. Genel olarak, linyit belirsizdir ve bazı bölgelerde büyük rezervler olmasına rağmen, kalori değerleri düşük olarak kaydedilebilir (Korkmaz ve Esmeray, 2018; Chelgani vd. 2010). Nem, kül ve kükürt gibi linyit kalitesini etkileyen değişkenler enerji santrallerinde, diğer alanlarda ve enerji üretiminde hayati öneme sahiptir. Bu nedenle adı geçen değişkenlerin modellenmesi yatırım kararları için yararlıdır.

Türkiye Kömür İşletmeleri Genel Müdürlüğünden alınan, MTA tarafından yayınlanan lintit envanterinde Türkiye'nin tamamını içeren 116 bölgede linyit değerleri elde edilmiş olup bu çalışmada eksik veriler çıkarılarak 96 linyit kömürünün 96 sondaj bölgesi için linyit kömürü kalite sınıflandırma yapılması amaçlanmıştır. Bu bölgeler için elde edilen kalori değerleri, K-Ortalamalar kümeleme yöntemiyle, kabul edilebilir kalite ve nispeten zayıf kalite olmak üzere 2 kategoriye ayrılmıştır. Bu kategorilere 2200 kcal / kg ile 5574 kcal / kg arasındaki kalori değerlerinin kabul edilebilir kalite seviyesinde olduğu ve 868 kcal / kg ile 2196 kcal / kg arasındaki kalori değerlerinin nispeten zayıf

**Tablo 1.** Türkiye Kömür İşletmeleri Genel Müdürlüğünden Alınan, 96 Bölgenin Nem, Kül, Kükürt ve Kalori Değerleri (MTA Tarafından Yayınlanan Linyit Envanteri Verileri)

No	Nem	Kül	Kükürt	AID (kcal/kg)	No	Nem	Kül	Kükürt	AID (kcal/kg)
1	41,00	28,00	2,10	1298	49	25,73	30,90	2,51	2539
2	49,72	21,34	1,16	1385	50	19,67	35,74	3,29	2832
3	38,66	14,79	3,39	2820	51	25,00	22,61	2,41	2800
4	10,25	40,52	1,25	3500	52	27,50	36,57	2,02	1908
5	14,53	10,31	1,81	5574	53	32,50	22,50	1,50	2300
6	21,71	34,35	4,04	2557	54	48,00	25,00	1,10	1150
7	14,83	48,70	3,39	1988	55	25,00	25,00	5,40	2250
8	26,44	25,36	2,79	2839	56	50,00	11,00	1,10	2180
9	10,00	36,11	5,20	3000	57	41,70	20,25	1,54	1344
10	35,36	28,69	1,46	2004	58	45,57	25,00	1,10	1155
11	36,83	22,63	1,64	2232	59	48,88	27,00	1,66	1430
12	41,06	23,34	1,19	1470	60	22,90	15,30	4,42	4063
13	15,25	20,16	2,75	4020	61	4,41	25,54	6,51	5200
14	20,46	27,24	0,98	3120	62	32,00	43,00	1,20	2080
15	16,00	22,00	3,16	3800	63	15,00	41,00	1,60	2560
16	20,00	8,00	2,00	3260	64	13,00	33,00	1,30	3150
17	16,70	39,80	1,67	2700	65	15,00	28,00	0,70	2070
18	47,00	24,00	0,60	1460	66	18,00	45,00	0,70	2310
19	10,09	20,98	9,00	4755	67	15,00	36,00	1,20	2940
20	9,00	28,00	6,91	4100	68	18,00	40,00	1,20	2080
21	23,00	31,00	1,80	2340	69	30,90	17,43	5,80	3600
22	53,00	10,73	2,50	1754	70	52,00	19,00	2,00	1143
23	34,00	26,00	1,50	1900	71	48,07	24,08	0,92	1063
24	31,00	26,00	4,20	2100	72	29,00	25,00	3,20	2196
25	11,67	27,05	3,71	3840	73	31,00	27,00	3,30	1642
26	24,00	24,00	2,00	2500	74	32,00	29,00	2,20	1775
27	37,00	11,00	1,00	3190	75	30,00	31,00	1,30	1863
28	23,00	25,00	4,20	3000	76	30,00	22,00	4,50	2229
29	26,00	16,00	4,00	3678	77	27,00	25,00	4,30	4200
30	28,40	21,30	4,85	2960	78	34,00	22,00	3,00	2185
31	21,61	52,62	1,09	1278	79	29,00	28,00	3,10	2356
32	48,47	28,59	0,71	868	80	26,00	24,00	2,80	2670
33	23,00	23,00	1,60	3150	81	25,58	24,33	2,72	2671
34	23,00	46,00	1,40	1470	82	33,00	27,00	2,30	2168
35	31,30	42,55	0,76	976	83	42,00	22,00	1,30	1903
36	20,42	43,11	4,54	1878	84	42,24	29,69	2,51	1231
37	40,00	34,00	2,60	3000	85	41,20	32,80	1,01	1244
38	38,00	20,00	2,00	2500	86	50,00	21,00	2,02	1300
39	19,20	23,29	0,71	3500	87	41,06	19,11	2,07	2120
40	31,61	26,97	1,33	2200	88	42,40	16,56	1,70	2194
41	17,07	35,73	4,07	2754	89	42,06	20,30	1,12	1699
42	34,53	34,12	1,30	1484	90	22,13	37,38	1,56	2266
43	5,66	36,44	3,74	4500	91	33,28	27,33	1,41	2359
44	11,76	40,89	1,94	2500	92	28,28	31,14	1,57	2277
45	28,59	30,39	1,60	2304	93	32,46	28,90	1,90	2075
46	35,00	19,00	0,20	2570	94	26,96	36,37	0,50	2089
47	21,00	45,00	0,50	971	95	5,32	31,46	4,10	4926
48	8,00	44,00	1,10	3100	96	13,76	43,80	4,09	2502

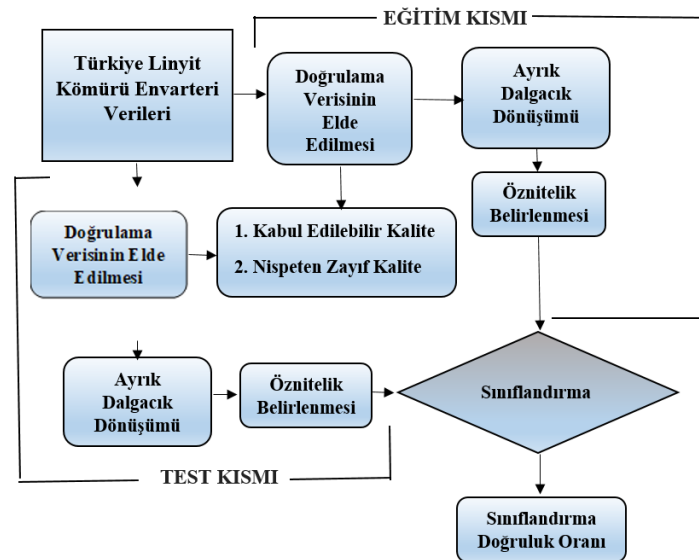
kalitede olduğu kabul edilmiştir. Bulunan bu kritik değerlere dayanarak, bu çalışmada linyit kalite parametreleri olan nem, kül ve kükürt içeriklerine göre sınıflandırma tahmini yapılması amaçlanmıştır. Bu çalışmada ifade edilen problemde, yanıt değişkeni olan linyit kalitesi, kabul edilebilir (1) ve nispeten zayıf (0) gibi sadece iki değer alabilir. İki hedef değişken kategorisinin üyeliğini tahmin etmek için çoklu öngörücü değişkenlerin etkisini tanımlayan bir sınıflandırma düşünülmüştür.

### Veri Seti

Tablo 1’de, Türkiye Kömür İşletmeleri Genel Müdürlüğünden alınan, 96 bölgenin nem, kül, kükürt ve kalori değerlerini yani MTA tarafından yayınlanan linyit envanteri verilerini göstermektedir. 96 bölgenin verileri vardır. Sınıflandırmada, 48 bölgenin verileri eğitim amaçlı, 48 bölgenin verileri test amaçlı kullanılmıştır.

### YÖNTEM

Bu makalede kullanılan linyit kömürü verileri, Türkiye Maden Tetkik ve Arama Genel Müdürlüğü'nün yetkili kamu kuruluşu tarafından yayınlanan Türkiye Linyit Envanterinden elde edilmiştir (Tutmez vd., 2013; MTA, 2010). Bu orijinal veriler, Türkiye'nin tamamını içeren 116 bölge için linyit değerleri elde edilmiş olup bu çalışmada eksik veriler çıkarılarak 96 linyit kömürünün 96 sondaj bölgesine ait nem, kül, kükürt ve kalori içerikleri ele alınmıştır. Bu makalede linyitin kabul edilebilir kalite ve nispeten zayıf kalite olmak üzere kalite sınıfının belirlenmesi için önerilmiş sistem Şekil 1’de görülmektedir.



Şekil 1. Önerilen Sistemin Genel Yapısı

### K-Ortalamalar Yöntemi

Çok değişkenli istatistiksel bir teknik olan K-Ortalamalar yöntemi, homojen alt gruplara ayrılması ve benzerliklerine göre sınıflandırılması için kullanılmaktadır. En çok bilinen kümeleme yöntemlerinden biri K-Ortalamalar yöntemidir. Hiyerarşik olmayan bir yapıya sahip olan K-Ortalamalar yöntemi, bir T veri setine ait n adet değişkeni ve B adet özellik vektörünü A adet kümeye ayırma ve sınıflandırma özelliğine sahiptir (Korkmaz ve Esmeray, 2018; Muda vd., 2011).

Bu yöntem, daha önceden seçilmiş A adet kümeye ait merkezlerin tespit edilmesi ile başlar. Ve daha sonra her bir değişken benzerlik ölçütü yardımıyla en yakın küme merkezine atanır (Korkmaz ve Esmeray, 2018; Ozcift vd., 2008). Giriş veri setinde bulunan her bir değişkenin 2 kümeye atanmasından sonra, her bir küme için küme merkezi yeniden hesaplanır. Küme merkezi yeniden hesaplandıktan sonra ise bu değişkenler elde edilen yeni küme merkezlerinin yerleşimine bağlı olarak yeni farklı kümelere atanabilir. Bu işlem küme üyeliklerinde herhangi bir değişim olmayana kadar tekrarlanır. İncelenen bir problemde, K adet özellik vektörüne ve n adet değişkene sahip bir T veri seti  $T = \{t_k | k = 1, 2, \dots, A\}$  şeklinde tanımlanabilir. Bu veri setinde k. özellik vektörü  $X_k = [x_{k1}, x_{k2}, \dots, x_{kn}]$ ,  $x_k \in R^d$  şeklinde yazılabilir (Korkmaz ve Esmeray, 2018; Muda vd., 2011; Fırat vd., 2012).

Denklem (1)'de, verilen amaç fonksiyonunun veri setinin alt kümelerine ayrıldığında en küçük kümeye ayrılması ana hedefdir. Bu denklemde verilen amaç fonksiyonu ile en küçük küme elde edilmesi esası, her bir özellik vektörü ile en yakın küme merkezi arasındaki mesafe ölçütünün ve dolayısıyla benzer yapıya sahip verilerin aynı kümede toplanmasını sağlamaya dayanmaktadır. Mesafe ölçütünün hesaplanması için genellikle Denklem (2)'de verilen öklit mesafe ölçütü kullanılmaktadır (Korkmaz ve Esmeray, 2018; Fırat vd., 2012; Leśniak ve Isakow, 2009).

$$J(S:T) = \sum_{i=1}^A \sum_{k=1}^B n_{ik}^2(t, S_i) \quad (1)$$

Denklem (1)'de verilen  $n_{ik}^2$  aşağıdaki gibi tanımlanmıştır.

$$n_{ik}^2 = \left\| t_k^{(i)} - S_i \right\|^2 \quad (2)$$

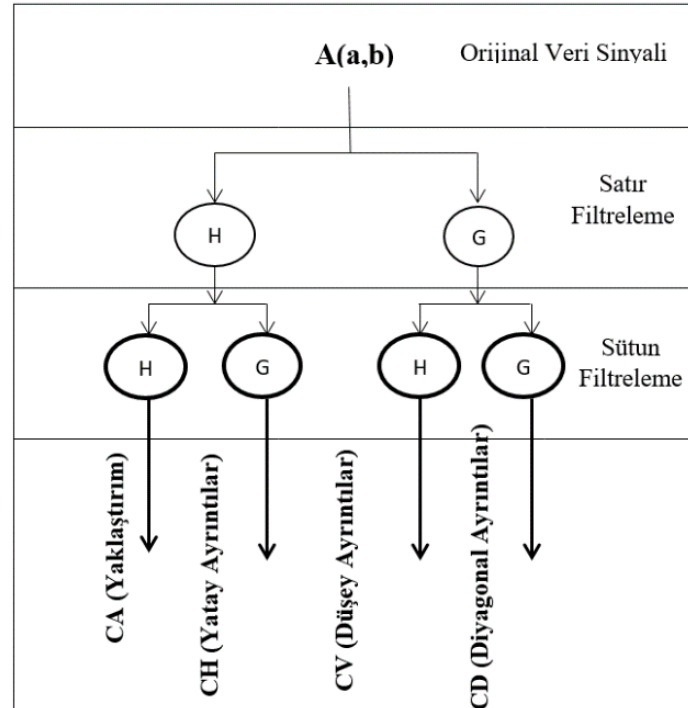
$$S_i = \frac{1}{B} \sum_{k=1}^B t_k^{(i)} \quad (3)$$

### Dalgacık Dönüşümü Özellikleri

Ayrık Dalgacık Dönüşümü öznelik çıkartımı için son yıllarda sinyal işleme, biyomedikal işaret işleme, haberleşme gibi birçok değişik alanda yaygın olarak kullanılan bir yöntemdir (Korkmaz ve Esmeray, 2018; Leśniak ve Isakow, 2009). Ayrık Dalgacık Dönüşümü (ADD), tüm zaman boyunca dalgacığın zaman düzleminde ölçeklenmiş ve ötelenmiş değerleriyle çarpılan sinyalin toplamına denir. ADD fonksiyonu, imge ya da işaret için istenilen seviyede kullanılabilir (Korkmaz ve Esmeray, 2018; Leśniak ve Isakow, 2009). ADD alt bantlara ayırma yöntemi verilerin, yüksek frekansa ait bileşenini dar bir aralıkta, düşük frekansa ait bileşenini de daha geniş bir bant aralığında analiz ederek, sinyale ait ayırt edici özelliklerin elde edilmesine olanak sağlar. ADD kullanılarak dördüncü seviyeye kadar alt bant ayrışımının matematiksel ifadesi Denklem (4)'de gösterilmektedir (Korkmaz ve Esmeray, 2018; Leśniak ve Isakow, 2009).

$$I_{BI}(u, v) = ADD\{I(x, y), B \in \{AA, YA, AY, YY\}, 1 < I < 4\} \quad (4)$$

Şekil 2'de işaretin bir seviye alt bantlara ayrılma işlemi gösterilmektedir.



Şekil 2. Orijinal Veri Sinyalinin Bir Seviye ADD Alt Bant Ayrışımı (Korkmaz ve Esmeray, 2018)

Şekil 2'deki  $G[b]$  yüksek geçiren filtre ve  $H[b]$  alçak geçiren filtreyi temsil etmektedir. Bu alt bantlara ayrıştırma işleminde  $H[b]$  işarete ait yaklaşık katsayıları ifade ederken  $G[b]$  ise işarete ait detay katsayılarını ifade etmektedir.

(Korkmaz ve Esmeray, 2018; Karhan ve Ergen, 2016; Sengur vd., 2008). İşaret öncelikle satır esasına göre filtrelemeye tabi tutulur. Bu filtreleme sonucunda, alçak ve yüksek geçiren filtrelemeden iki grup veri elde edilir. Elde edilen verilerde örnek azaltması yapılır. Ardından bu iki alt grup veriler sütun esasına göre yine alçak ve yüksek geçiren filtrelerden geçirilir. Bu alt iki grup veriden de yine iki alt veri grubu elde edilir. Sonuçta yaklaşım, yatay, dikey ve diyagonal olmak üzere dört alt veri grubu elde edilmiş olur. Yaklaşım verileri yine benzer şekilde ayrıştırılarak ADD devam ettirilebilir.

Çalışmalarda seçilmesi gereken dalgacık ailesi çeşidinin belirlenmesinde herhangi bir algoritma bulunmamaktadır. Zaman düzleminde ana dalgacık, üzerinde çalışılacak sinyale görsel olarak ne kadar çok benzerse, seçilen Dalgacık Dönüşümü yönteminden o kadar çok verim elde edilir. Fakat bu şekilde bir seçim yapılması hemen hemen olanaksız görünmektedir. Bu sebeple tüm dalgacık çeşitleri denenmelidir. Diğer filtreler kullanıldığında sınıflandırma doğruluk oranları düşük çıkmıştır. Bu yüzden dalgacık dönüşümünde kullanılan wavelet filtreleri diğer filtrelere göre avantajlı olduğu için kullanılmıştır. Genel amaçlı olarak kullanılan dalgacık Daubechies'tir. Diferansiyel bir operatör olan Haar Dalgacığı, Daubechies1, Symlet1, Coiflet1 ve BiorSpline1.1 dalgacık aileleri ile aynıdır (Korkmaz ve Esmeray, 2018; Sengur vd., 2008).

### **Naive Bayes Sınıflandırıcısı**

Naive Bayes sınıflandırıcısı makine öğrenmesi ve veri madenciliği alanında yaygın kullanıma sahip denetimli öğrenme sınıfında bir algoritmadır. Belirli bir veri kümesini belirli bir sınıfa ait olma olasılığını tahmin etmede kullanılır. Sınıflandırmada, Bayes Karar teoremini esas alarak ve olasılık hesaplamalarını kullanarak en yüksek olasılığa sahip kararı seçmeyi hedefler. Bu yöntemde, her bir özelliğin sınıf içerisindeki diğer özelliklerden bağımsız olduğu kabul edilmektedir. Bu sınıflandırıcı basit hesaplamalar yaparak olasılık tabanlı olarak bir olayın gerçekleşme ihtimalini hesaplamaktadır. Naive Bayes sınıflandırıcısı, tahmine dayalı modelleme yapmak için basit ve bunun yanında güçlü bir algoritma olduğu için ve az sayıdaki verilerle güçlü tahmin kabiliyetine sahip kullanışlı ve avantajlı bir yöntem olduğu için tercih edilir. Naive Bayes sınıflandırıcı, Bayes Teoremini temel alan ve bağımsız varsayımlara dayanan olasılıklı bir sınıflayıcıdır. Bayes teoremi koşullu olasılıklar ile marjinal olasılıklar arasındaki ilişkiyi gösteren ve bir rassal değişkeni referans alan sınıflandırıcı yöntemidir (McCallum ve Nigam, 1998).

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \quad (5)$$

Denklem (5)'de  $P(X|Y)$  Y olayının gerçekleştiği durumda X olayının meydana gelme olasılığıdır.  $P(Y|X)$  X olayı gerçekleştiği durumda Y olayının meydana gelme olasılığıdır.  $P(X)$  ve  $P(Y)$  ise X ve Y olaylarının önsel olasılıklarıdır (Yongkui vd., 2014). Bir sınıflandırma işlemi, birçok özellikten ve bir hedef değişkeninden oluşur.

$$p(T|K_1 \dots K_k) = \frac{p(T)p(K_1 \dots K_n|T)}{p(K_1 \dots K_n)} \quad (6)$$

T verilen hedef ve K özellikleri temsil eder. Naive Bayes sınıflandırıcı genel olarak bütün koşullu olasılıkların çarpımıdır (McCallum ve Nigam, 1998; Yongkui vd., 2014).

## **SONUÇLAR VE TARTIŞMA**

Bugüne kadar makine öğrenme alanında birçok çalışma bulunmaktadır (Tuncer ve Ertam, 2020; Tuncer vd., 2019; Korkmaz, 2021; Korkmaz ve Binol, 2018; Aytaç Korkmaz, 2020; Doğan ve Ergen, 2022; Özdemir ve Türkoğlu, 2022). Bu makalede, makine öğrenme yöntemlerinden K-Ortalamalar, Dalgacık Dönüşümü ve Naive Bayes yöntemleri, Türkiye Maden Tetkik ve Arama Genel Müdürlüğü'nün yetkili kamu kuruluşu tarafından yayımlanan Türkiye Linyit Envanteri'nden elde edilen linyit kömürü verilerine uygulanmıştır. Türkiye Maden Tetkik ve Arama Genel Müdürlüğü'nün yetkili kamu kuruluşu tarafından yayımlanan Türkiye Linyit Envanteri'nden elde edilen linyit kömürü verilerinden olan nem, kül, kükürt ve kalori değerlerinden öncelikle kalori değerleri doğrulama verisi olarak kullanılmıştır. Bu kalori değerlerine göre K-Ortalamalar kümeleme yöntemiyle veriler "kabul edilebilir kalite" ve "nispeten zayıf kalite" olmak üzere 2 kategoriye ayrılmıştır. Bu kategorileşme yöntemi ile 2200 kcal / kg ile 5574 kcal / kg arasındaki kalori değerlerinin kabul edilebilir kalite seviyesinde olduğu ve 868 kcal / kg ile 2196 kcal / kg arasındaki kalori değerlerinin nispeten zayıf kalitede olduğu kabul edilmiştir. K-Ortalamalar yöntemiyle 2 kategoriye ayrılan 96 sondaj lokasyonu bölgesindeki linyit kömürlerinde doğrulama verisi olarak, 54 verinin kabul edilebilir kaliteli, 42'sinin ise nispeten zayıf kaliteli linyit kömürünün olduğu saptanmıştır. Daha sonra bu çalışmada linyit kömürünün kalite parametreleri olan nem, kül ve kükürt içerikleri esas alınarak Dalgacık Dönüşümü yöntemi ile

yaklaşım (CA), yatay (CH), düşey (CV), diyagonal (CD) katsayı değerleri bulunmuştur. Bu katsayılar bulunurken Daubechies, Symlet, Coiflet, Reversebior, Dmey, Biorpline ve Haar dalgacık aile gruplarından faydalanılmıştır. Bu katsayıların değerleri ayrı ayrı Naive Bayes sınıflandırıcısına verilerek “kabul edilebilir kalite” ve “nispeten zayıf kalite” olmak üzere iki sınıfa ayrılarak sınıflandırma doğruluk oranları bulunmuştur. Sınıflandırma yapılırken toplam verilerin 48’i test amaçlı 48’i ise eğitim amaçlı kullanılmıştır. Daha sonra bulunan bu sınıflandırma doğruluk oranı sonuçları karşılaştırılmıştır. Bu sonuçların karşılaştırılması Tablo 2, Tablo 3 ve Tablo 4’de gösterilmektedir.

**Tablo 2.** Reversebior Dalgacık Türü İçin Dalgacık Katsayılarına Göre Naive Bayes Sınıflandırıcı Performansları

	<b>Dalgacık Katsayılarına göre Naive Bayes Sınıflandırıcı Performansı</b>			
	<b>CA</b>	<b>CH</b>	<b>CV</b>	<b>CD</b>
rbio1.1	86,95	73,91	43,47	34,78
rbio1.3	86,95	60,86	47,82	56,52
rbio1.5	86,95	60,86	47,82	34,78
rbio2.2	91,30	78,26	56,52	47,82
rbio2.4	91,30	78,26	52,17	52,17
rbio2.6	91,30	78,26	52,17	56,52
rbio2.8	91,30	78,26	47,82	47,82
rbio3.1	86,95	69,56	43,47	47,82
rbio3.3	86,95	69,56	43,47	47,82
rbio3.5	86,95	69,56	43,47	47,82
rbio3.7	86,95	69,56	43,47	52,17
rbio3.9	86,95	69,56	47,82	52,17
rbio4.4	86,95	78,26	56,52	47,82
rbio5.5	86,95	73,91	56,52	56,52
rbio6.8	86,95	78,26	52,17	43,47

K-Ortalama kümeleme yöntemiyle kalori değerlerine göre kümelenen linyit kömürünün nem, kül ve kükürt değerleri, Tablo 2’de reversebior dalgacık ailesi türleri için CA, CH, CV ve CD katsayı değerleri kullanılarak Naive Bayes sınıflandırıcısı ile sınıflandırılmıştır. Tablo 2’deki sonuçlar incelendiğinde en yüksek sınıflandırma performansının reverseior dalgacık türlerinden rbio2.2, rbio2.4, rbio2.6 ve rbio2.8’in CA katsayıları ile %91.30 olarak bulunduğu görülmektedir.

K-Ortalama kümeleme yöntemiyle kalori değerlerine göre kümelenen linyit kömürünün nem, kül ve kükürt değerleri, Tablo 3’de Daubechies dalgacık ailesi türleri için CA, CH, CV ve CD katsayı değerleri kullanılarak Naive Bayes sınıflandırıcısı ile sınıflandırılmıştır. Tablo 3’deki sonuçlar incelendiğinde en yüksek sınıflandırma performansının Daubechies dalgacık türlerinden db4’ün CA katsayıları ile % 100 olarak bulunduğu görülmektedir.

**Tablo 3.** Daubechies Dalgacık Türü İçin Dalgacık Katsayılarına Göre Naive Bayes Sınıflandırıcı Performansları

	<b>Dalgacık Katsayılarına göre Naive Bayes Sınıflandırıcı Performansı</b>			
	<b>CA</b>	<b>CH</b>	<b>CV</b>	<b>CD</b>
db1	86,95	73,91	43,47	34,78
db2	82,60	60,86	56,52	56,52
db3	82,60	73,91	39,13	43,47
db4	100	91,30	39,13	43,47
db5	86,95	86,95	52,17	39,13
db6	78,26	78,26	34,78	34,78
db7	91,30	82,60	47,82	43,47
db8	73,91	78,26	43,47	43,47
db9	82,60	86,95	65,21	65,21
db10	86,95	82,60	47,82	47,82
db11	73,91	78,26	52,17	47,82
db12	91,30	86,95	65,21	60,86
db13	65,21	73,91	47,82	65,21
db14	82,60	78,26	52,17	56,52

K-Ortalama kümeleme yöntemiyle kalori değerlerine göre kümelenen linyit kömürünün nem, kül ve kükürt değerleri, Tablo 4’de Dmey, Haar, Coif, Sym ve Bior dalgacık ailesi türleri için CA, CH, CV ve CD katsayı değerleri kullanılarak Naive Bayes sınıflandırıcısı ile sınıflandırılmıştır. Tablo 4’deki sonuçlar incelendiğinde en yüksek



sınıflandırma performansının Coif3, Bior1.3 ve Bior1.5'in CA katsayıları ile % 91.30 olarak bulunduğu görülmektedir.

**Tablo 4.** Diğer Dalgacık Türleri İçin Dalgacık Katsayılarına Göre Naive Bayes Sınıflandırıcı Performansları

	Dalgacık Katsayılarına göre Naive Bayes Sınıflandırıcı Performansı			
	CA	CH	CV	CD
dmey	86,95	82,60	56,52	56,52
haar	86,95	73,91	43,47	34,78
coif1	86,95	78,26	52,17	56,52
coif2	69,56	73,91	43,47	39,13
coif3	91,30	91,30	52,17	43,47
coif4	78,26	78,26	39,13	34,78
coif5	86,95	82,60	47,82	43,47
sym1	86,95	73,91	43,47	34,78
sym2	82,60	60,86	56,52	56,52
sym3	82,60	73,91	39,13	43,47
sym4	69,56	65,21	43,47	43,47
sym5	86,95	86,95	56,52	47,82
sym6	69,56	69,56	43,47	39,13
sym7	69,56	69,56	43,47	39,13
sym8	69,56	69,56	43,47	34,78
bior1.1	86,95	73,91	43,47	34,78
bior1.3	91,30	65,21	43,47	34,78
bior1.5	91,30	69,56	43,47	34,78
bior2.2	78,26	69,56	56,52	60,86
bior2.4	82,60	69,56	56,52	60,86
bior2.6	82,60	69,56	56,52	60,86
bior2.8	82,60	69,56	56,52	60,86
bior3.1	86,95	56,52	43,47	60,86
bior3.3	78,26	69,56	43,47	60,86
bior3.5	78,26	69,56	43,47	60,86
bior3.7	78,26	69,56	43,47	60,86
bior3.9	78,26	65,21	43,47	60,86
bior4.4	86,95	69,56	52,17	56,52
bior5.5	86,95	86,95	52,17	52,17
bior6.8	86,95	69,56	52,17	56,52

Bu çalışmada elde edilen sonuçlar ile literatürde önerilen benzer çalışmaların sonuçları arasında aşağıda belirtilen üç nedenden dolayı doğrudan ve ayrıntılı karşılaştırma yapmak mümkün olamamıştır. İlk olarak, her çalışmada farklı veri doğrulama yöntemleri ve farklı makine öğrenmesi yöntemleri ve kullanılmaktadır. İkincisi, her çalışma, analiz türüne göre farklı tahmin edici değişkenleri olan ve farklı sayıda örnek içeren farklı veri kümeleri kullanmaktadır. Son olarak, oluşturulan her veri seti, örneğin kömürün çıkarıldığı bölge ve kömürün kalitesi gibi farklı veri özelliklerine sahip olabilmektedir.

Ancak, literatür çalışmalarından anlaşılacağı üzere günümüzde Türkiye'de bulunan karmaşık nem, kül, kükürt ve kalori gibi linyit kömürü verilerinin kabul edilebilir kalite ve nispeten zayıf kalite diye iki sınıfa ayrılması için K-Ortalamlar, Dalgacık Dönüşümü ve Naive Bayes sınıflandırıcı yöntemleri birlikte kullanılmamıştır. İlk olarak karmaşık olarak Türkiye Kömür İşletmeleri Genel Müdürlüğü tarafından verilen verilere K-Ortalamlar kümeleme algoritması uygulanmıştır. Bu K-Ortalamlar algoritmasının kümelediği kalori değerleri, nem, kül ve kükürt gibi linyit kömürü parametrelerinin sınıflandırılmasında doğrulama verisi olarak kullanılmıştır. Daha sonra nem, kül ve kükürt gibi linyit kömürü parametrelerinin Daubechies, Symlet, Coiflet, Reversebior, Dmey, Biorspline ve Haar gibi dalgacık aileleri yardımıyla CA, CH, CV ve CD dalgacık katsayıları bulunmuştur. Bu dalgacık katsayılarından elde edilen özellik değerleri Naive Bayes sınıflandırıcısı ile sınıflandırılmıştır. Sınıflandırma performansları kıyaslandığında nem, kül ve kükürt linyit kömürü parametrelerine uygulanan Daubechies dalgacık ailesinden olan db4'ün diğer dalgacık ailelerine ve CA dalgacık katsayılarının ise CH, CV ve CD katsayılarına göre daha yüksek performans sergilediği tespit edilmiştir. Kaliteli ve nispeten zayıf kaliteli linyit kömürünün yüksek performansla ayırt edilmesi yatırımcılara yatırım amaçlı fayda sağlarken, linyit kömürü kullanıcılarına da sağlık açısından fayda sağlayacağı düşünülmektedir.

Gelecek çalışmalarda, linyit verilerinin bilinen x, y koordinat sistemleri ve kalori değerleri referans alındığında iki boyutlu olduğu için bazı sınırlamalar olduğu için, 3 boyutlu olarak sondaj vurulmayan farklı x ve y noktalarındaki kalori sonuçları regresyon yöntemleriyle tahmin edilebilir. Bu çalışmada veri sayısı az olduğu için LSTM Derin öğrenme yöntemleri kullanılmamıştır. Bu yüzden, gelecek çalışmalarda, çok fazla sayıda bölgenin nem, kül, kükürt ve kalori değerleri temin edilebilirse, LSTM derin öğrenme tabanlı yöntemlerde kullanılarak analiz yapılabilir.

## TEŞEKKÜR

Yazar İnönü üniversitesinde öğretim üyesi olan Prof. Dr. Bülent TÜTMEZ'e teşekkür etmektedir.

## KAYNAKLAR

- Aytac Korkmaz, S. (2020). Grade level of lignite coal datas in the different areas with decision tree, random forest, and discriminant analysis methods. *Applied Artificial Intelligence*, **34(11)**, 755-776.
- Cheng, Y., Xu, L., Li, X., & Guo, Z. (2012, July). Online coal calorific value prediction from mutiband coal/air combustion radiation characteristics. In *Instrumentation and Control Technology (ISICT), 2012 8th IEEE International Symposium on*, 309-313, IEEE.
- Chelgani S C, Mesroghli S H, Hower J C. (2010). Simultaneous prediction of coal rank parameters based on ultimate analysis using regression and artificial neural network. *International Journal of Coal Geology*, **83(1)**, 31-34.
- Doğan, G. & Ergen, B. (2022). Karayollarındaki Asfalt Çatlaklarının Tespiti İçin Yeni Bir Konvolüsyonel Sinir Ağı Tabanlı Yöntem . *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, **34 (2)** , 485-494 . DOI: 10.35234/fumbd.1014951.
- Fırat M., Dikbaş F., Koç AC., ve Güngör M. (2012). K-Ortalamalar Yöntemi ile Yıllık Yağışların Sınıflandırılması ve Homojen Bölgelerin Belirlenmesi. *İMO Teknik Dergi*, **383**, 6037-6050.
- Galetakis M J, Theodoridis K, Kouridou O. (2002). Lignite quality estimation using ANN and adaptive neuro-fuzzy inference systems (ANFIS). *APPCOM*: 425-431.
- IEA, (2000). International energy annual. France. *International Journal of Coal Science & Engineering (China) Energy Agency*.
- Karhan Z., Ergen B. (2016). Content based medical image classification using discrete wavelet and cosine transforms. (2015). *23rd Signal Processing and Communications Applications Conference (SIU)*, (pp:1445-1448).
- Korkmaz, S. A. (2021). Classification of histopathological gastric images using a new method. *Neural Computing and Applications*, **33(18)**, 12007-12022.
- Korkmaz, S. A., & Binol, H. (2018). Classification of molecular structure images by using ANN, RF, LBP, HOG, and size reduction methods for early stomach cancer detection. *Journal of Molecular Structure*, **1156**, 255-263.
- Korkmaz, S. A., & Esmeray, F. (2018, March). Quality lignite coal detection with discrete wavelet transform, discrete fourier transform, and ANN based on k-means clustering method. In *Digital Forensic and Security (ISDFS), 2018 6th International Symposium on* (pp. 1-6). IEEE.
- Korkmaz, S. A., & Poyraz, M. (2014). A New Method Based for Diagnosis of Breast Cancer Cells from Microscopic Images: DWEE—JHT. *Journal of medical systems*, **38(9)**, 92.
- Leśniak, A., & Isakow, Z. (2009). Space-time clustering of seismic events and hazard assessment in the Zabrze-Bielszowice coal mine. Poland. *International Journal of Rock Mechanics and Mining Sciences*, **46(5)**, 918-928.
- McCallum, A., & Nigam, K. (1998, July). A comparison of event models for Naive Bayes text classification. In *AAAI-98 workshop on learning for text categorization*, **752**, pp. 41-48.
- Moon C J, Whateley M K G, Evans A M. (2006). "Introduction to mineral exploration" India: Blackwell Publishing.
- MTA, (2010). Lignite inventory of Turkey, general directorate of mineral research and exploration (MTA) in Turkey. Ankara (in Turkish).
- Muda, Z., Yassin, W., Sulaiman, M. N., & Udzir, N. I. (2011, July). Intrusion detection based on K-Means clustering and Naive Bayes classification. In *Information Technology in Asia (CITA 11), 2011 7th International Conference on* (pp. 1-6). IEEE.

- Ozcift, A., & Gulten, A. (2008). Assessing effects of pre-processing mass spectrometry data on classification performance. *European Journal of Mass Spectrometry*, **14(5)**, 267-273.
- Özdemir, E. & Türkoğlu, İ. (2022). Yazılım Güvenlik Açıklarının Evrimsel Sinir Ağları (CNN) ile Sınıflandırılması, *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, **34(2)**, 517-529. DOI: 10.35234/fumbd.1076870.
- Pandit, Y. P., Badhe, Y. P., Sharma, B. K., Tambe, S. S., & Kulkarni, B. D. (2011). Classification of Indian power coals using K-means clustering and Self Organizing Map neural network. *Fuel*, **90(1)**, 339-347.
- Sahu, H. B., Mahapatra, S. S., & Panigrahi, D. C. (2012). Fuzzy c-means clustering approach for classification of Indian coal seams with respect to their spontaneous combustion susceptibility. *Fuel processing technology*, **104**, 115-120.
- Sahu, H. B., Mahapatra, S. S., Sirikasemsuk, K., & Panigrahi, D. C. (2011). A discrete particle swarm optimization approach for classification of Indian coal seams with respect to their spontaneous combustion susceptibility. *Fuel processing technology*, **92(3)**, 479-485.
- Senguler I. (2010). Lignite explorations in Turkey: new projects and new reserves. //17th Annual International Pittsburgh Coal Conference, İstanbul, Turkey.
- Sengur, A., Turkoglu, I., & Ince, M. C. (2007). Wavelet packet neural networks for texture classification. *Expert systems with applications*, **32(2)**, 527-533.
- Sengur, A., Turkoglu, I., & Ince, M. C. (2008). Wavelet oscillator neural networks for texture segmentation. *Neural Network World*, **18(4)**, 275.
- Tuncer, T., & Ertam, F. (2020). Neighborhood component analysis and reliefF based survival recognition methods for Hepatocellular carcinoma. *Physica A: Statistical Mechanics and its Applications*, **540**, 123143.
- Tuncer, T., Dogan, S., & Akbal, E. (2019). A novel local senary pattern based epilepsy diagnosis system using EEG signals. *Australasian physical & engineering sciences in medicine*, **42(4)**, 939-948.
- Tutmez B., HOZATLI B., CENGİZ A.K. (2013). An overview of Turkish lignite qualities by logistic analysis, *Journal of Coal Science & Engineering China*, **19:2**, 113-118.
- Yang, X. L., Wang, F., Wang, W. C., Chen, Y. X., & Chen, J. S. (2014). DWT-PLS Regression on Near-Infrared Spectra for Moisture Determination of Coal. In *Advanced Materials Research*, **827**, 209-212. Trans Tech Publications.
- Yilmaz, I., Erik, N. Y., & Kaynar, O. (2010). Different types of learning algorithms of artificial neural network (ANN) models for prediction of gross calorific value (GCV) of coals. *Scientific Research and Essays*, **5(16)**, 2242-2249.
- Yongkui, S., Pengrui, L., Ying, W., Jingyu, Z., & Meijie, L. (2014). The Prediction of the Caving Degree of Coal Seam.