



Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi

<https://dergipark.org.tr/tr/pub/yyufbed>



Araştırma Makalesi

Dengesiz Veri Kümelerinin Sınıflandırılmasında Poligon Alan Metriğinin Sınıflandırıcı Performans Değerlendirilmesi için Kullanılması

Önder AYDEMİR

Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 61080,
Trabzon, Türkiye

Önder AYDEMİR, ORCID No: 0000-0002-1177-8518

Sorumlu yazar e-posta: onderaydemir@ktu.edu.tr

Makale Bilgileri

Geliş: 01.02.2022

Kabul: 26.04.2022

Online Ağustos 2022

DOI: 10.53433/yyufbed.1066340

Anahtar Kelimeler

Dengesiz veri,

Poligon alan metriği,

Sınıflandırıcı performansı

Öz: Son yıllarda makine öğrenmesi yöntemleri birçok disiplinde kullanılmaktadır. Araştırmacılar en uygun sınıflandırıcıyı belirlemek için birçok yöntemi yine birçok metrik ile birlikte test etmektedir. Özellikle düzensiz veri setleri için sınıflandırıcı performanslarını karşılaştırmak için kullanılan klasik metrikler (sınıflandırma doğruluğu, özgüllük, duyarlılık, eğri altında kalan alan, Jaccard indeksi ve F metriği) büyük tabloların oluşması ile birlikte takip edilebilirliği zorlaştırmaktadır. Diğer taraftan ise bir sınıflandırıcı bir metrik açısından yüksek performans sağlarken başka bir metrik açısından düşük performans sağlayabilmektedir. Tüm bunlar en uygun sınıflandırıcının belirlenmesini zorlaştırmaktadır. Bu çalışmada düzensiz veri setlerinde sınıflandırıcı performanslarını karşılaştırmak için poligon alan metriğinin (PAM) kullanılabileceği gösterilmiştir. Bu metrik sınıflandırma doğruluğu, özgüllük, duyarlılık, eğri altında kalan alan, Jaccard indeksi ve F metriğinin düzgün bir altıgen üzerinde oluşturdukları poligonun alanı üzerinden hesaplanırken, klasik metriklerin değerlerini de bu düzgün altıgen üzerinde görselleştirme esasına dayanmaktadır. Bu yöntem ile sınıflandırıcı performanslarının etkin bir biçimde karşılaştırılabileceği sonucuna varılmıştır.

Using the Polygon Area Metric for Evaluation of Classifier Performance in the Classification of Unbalanced Datasets

Article Info

Received: 01.02.2022

Accepted: 26.04.2022

Online August 2022

DOI: 10.53433/yyufbed.1066340

Keywords

Classifier performance,

Polygon area metric,

Unbalanced data

Abstract: In recent years, machine learning methods have been used in many disciplines. Researchers test many methods together with many metrics to determine the most suitable classifier. Classical metrics (classification accuracy, specificity, sensitivity, area under the curve, Jaccard index and F metric) used to compare classifier performances, especially for irregular data sets, make traceability difficult with the formation of large tables. On the other hand, while a classifier provides high performance in terms of one metric, it may provide low performance in terms of another metric. All this complicates the determination of the most suitable classifier. In this study, it has been shown that polygon area metric (PAM) can be used to compare classifier performances, in irregular data sets. While this metric is calculated over the classification accuracy, specificity, sensitivity, area under the curve, Jaccard index and the area of the polygon formed by the F metric on a regular hexagon, it is based on visualizing the values of classical metrics on this regular hexagon. It has been concluded that this method can perform classifier performances effectively.

1. Giriş

Makine öğrenmesi algoritmaları günümüzün hemen her disiplninde kullanılan, zaman kazandıran, performans arttıran ve çeşitli açılardan kazanç sağlayan yöntemlerdir (Alsheikh ve ark., 2014; Al-Garadi ve ark., 2019). Araştırmacılar ele aldıkları probleme en uygun ve en yüksek performansı sağlayan makine öğrenmesi (sınıflandırıcı) yöntemini belirlemek için birçok deney gerçekleştirirler. Bu seçim, büyük oranda hem dikkate alınan öznelik uzayına hem de sınıflandırıcı performansını belirlemede kullanılan metriğe bağlıdır. En yaygın kullanılan metrik sınıflandırma doğruluğu (classification accuracy) olmakla birlikte duyarlılık (sensitivity), özgüllük (specificity), eğri altında kalan alan (area under curve), Jaccard indeksi, F-ölçütü (F-measure) ve Kappa istatistiği de sınıflandırıcı performanslarını değerlendirmede kullanılmaktadır (Liu ve ark., 2022; Rodríguez-de-Vera ve ark., 2022). Sınıflandırıcı modelinin oluşturulmasında dikkate alınan performans metriği çok önemlidir.

Her biri farklı bir anlam ifade etse de çok sayıda metrik kullanmak en uygun sınıflandırıcıyı belirlemede zorluk çıkarmakta ve aynı zamanda büyük tabloların ortaya çıkmasına sebebiyet vermektedir. Ancak yine de bazı durumlarda, özellikle dengesiz dağılıma (sınıfların içerdiği örnek sayılarının birbirine eşit olmadığı veri setleri) sahip sınıflandırma problemlerinde (Liu ve ark., 2014; Hossin ve ark., 2015; Huang ve ark., 2021; Vuttipittayamongkol ve ark., 2021), tek başına bir metrik kullanmakta yeterli olmayabileceği gibi anlamsız da kalabilmektedir. Örneğin, bir grubunda (*sınıf 1*) 900 ve diğer grubunda (*sınıf 2*) 100 örnek olan iki sınıflı bir test kümesi ele alınsın. Modellenen sınıflandırıcının, tüm örnekleri (900+100=1000) *sınıf 1* olarak tahmin ettiği varsayılınsın. Bu durumda kullanılan sınıflandırıcıya ait sınıflandırma doğruluğu (SD) $900/1000=0.9$ ya da yüzde cinsinden verilecek olursa %90 olarak hesaplanır ki bu değer, hiçbir şekilde bu sınıflandırıcının gerçek performansını yansıtmayacaktır. Bu sorunu ortadan kaldırmak için en az duyarlılık ve özgüllük metriklerini de hesaplamak gerekir. Verilen bu örnek için bu metriklerden biri %0 ve diğeri %100 olarak hesaplanacaktır ve böylece sınıflandırıcının aslında hiçbir başarısının olmadığı anlaşılacaktır. Araştırmacılar buna benzer problemlerin önüne geçebilmek ve sınıflandırıcı performanslarını doğru analiz edebilmek için birden fazla metrik kullanma yoluna giderler. Ancak bu da beraberinde iki sorunu getirmektedir. Bunlardan biri büyük ve takip edilebilirliği zor kıyaslama tablolarının oluşması, ikincisi ise bir sınıflandırıcının bir metrik için yüksek diğeri bir metrik için düşük performans sergileyebilmesidir (Liu ve ark., 2014; Hossin ve ark., 2015)

Özellikle dengesiz dağılıma sahip verilerin sınıflandırılmasında sınıflandırıcı performanslarının değerlendirilmesi için Fawcett (2006) duyarlılık (D) ve özgüllük (Ö) metriklerini birlikte kullanarak eğri altında kalan alanı (EAKA) metriğini türetmiştir (Fawcett, 2006). Bu metrik medikal veri setlerinin sınıflandırılmasında performans değerlendirilmesi için sıklıkla kullanılır (Shia & Chen, 2021; Yin ve ark., 2021). İstatistiksel bir yaklaşım olan Jaccard indeksi (Jİ) ise tahmin edilen etiketler ile gerçek etiketlerin benzerlik oranını bulurken, en önemli dezavantajı *doğru negatif* oranını dikkate almamasıdır (Mun ve ark., 2017). Bu metrik daha çok görüntü segmentasyon performans değerlendirmelerinde kullanılır (Roy ve ark., 2022; Militello ve ark., 2022; Yeung ve ark., 2022). Literatür yakından incelendiğinde aynı ya da farklı tip sınıflandırma problemlerinde belirlenmiş bir metrik kullanımı mevcut değildir. Araştırmacılar sınıflandırıcı performansını ölçmede ya birden fazla metrik kullanımına ya da yeni bir metrik önermeye yönelmektedir. Önerilecek yeni bir metriğin şu özelliklere sahip olması beklenmektedir: 1) İki sınıflılarla birlikte çok sınıflı problemlere de uygulanabilme, 2) Hesaplama karmaşıklığı ve süresinin düşük olması, 3) Performans karşılaştırması için ayırt edici olma, 4) Sınıflandırıcı performansı hakkında olabildiğince bilgilendirici olmalı, 5) Azınlık sınıfının performans içerisindeki yerini anlamlı olmalı, 6) Olası tüm senaryolar (en kötü, en iyi veya tam tersi sınıflandırma gibi) için hesaplanabilmeli.

Bu çalışmada dengesiz veri kümelerinin sınıflandırılmasında poligon alan metriğinin (PAM) (MATLAB, 2018; Aydemir, 2021) sınıflandırıcı performans değerlendirilmesi için kullanılmasının hem büyük tabloların oluşmasının önüne geçeceği hem de sınıflandırıcıların performans karşılaştırılması için kullanılarak en uygun sınıflandırıcıyı belirlemek için kullanılabilmesi önerilmektedir. PAM'ın etkinliğini ortaya koymak için gerçek ve yapay olarak üretilmiş dengeli ve dengesiz dağılıma sahip toplam dört adet veri seti kullanılmıştır. Bunlara ek olarak sınıflandırmada en iyi, en kötü ve tüm örnekleri diğeri sınıf olarak etiketleme senaryoları da değerlendirilmiştir. Elde edilen sonuçlar PAM'ın

sınıflandırıcı performanslarını ortaya koymada çok sayıda metrik kullanılmadan tek başına yeterli olabileceğini göstermiştir.

2. Materyal ve Yöntem

Bu bölümde kullanılan veri setleri tanıtılmış, ardından sınıflandırıcı metrikleri ve PAM matematiksel olarak açıklanmıştır.

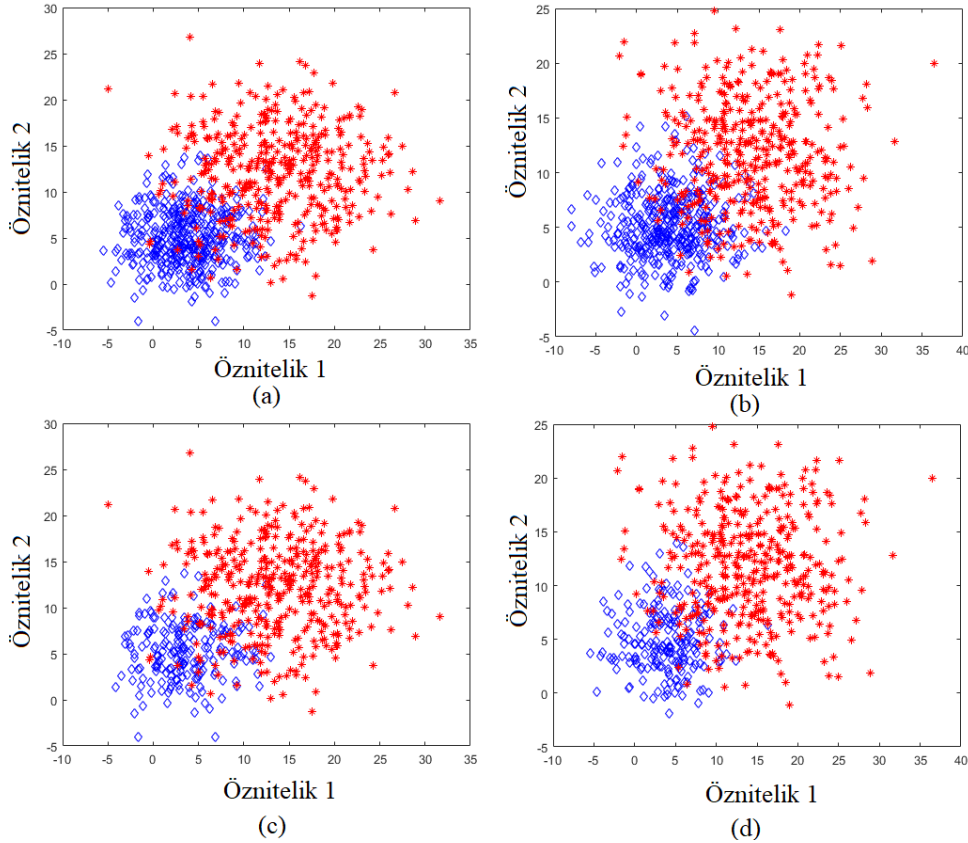
2.1. Veri setleri

2.1.1. Yapay veri setleri

Yapay veri setleri Matlab ortamında *randn* fonksiyonu kullanılarak türetilmiştir. PAM'ın düzensiz veri dağılımlarında etkinliğinin anlaşılabilmesi için oluşturulan düzenli ve düzensiz veri setlerinin özellikleri Çizelge 1'de ve dağılımları Şekil 1'de verilmiştir. Bu istatistiksel özellikler sınıflandırma aşamasında %50 ile %100 arasında sınıflandırma doğruluğu hesaplayabilecek şekilde seçilmiştir. Veri seti 1'de her bir sınıftan 800 adet örnek sayısı (ÖS) üretilirken, düzensiz olan veri seti 2'de *sınıf 1* ve *sınıf 2* için sırasıyla 400 ve 800 ÖS üretilmiştir. Bu sayıların yarısı eğitim yarısı da test kümesi için rastgele olarak seçilmiştir. Şekil 1'de mavi noktalar *sınıf 1* ve kırmızı noktalar ise *sınıf 2* örneklerini göstermektedir.

Çizelge 1. Yapay veri seti özellikleri

Veri Seti	Sınıf 1				ÖS	Veri Seti	Sınıf 2				ÖS
	x-ekseni		y-ekseni				x-ekseni		y-ekseni		
	Ortalama	Varyans	Ortalama	Varyans			Ortalama	Varyans	Ortalama	Varyans	
Veri Seti 1	4	14	5	9.6	800	Veri Seti 1	14	37.9	12	24.6	800
Veri Seti 2	4	14	5	9.6	400	Veri Seti 2	14	37.9	12	24.6	800



Şekil 1. Yapay veri seti dağılımı, (a): Yapay veri seti 1 eğitim verisi, (b): Yapay veri seti 1 test verisi, (c) Yapay veri seti 2 eğitim verisi, (d) Yapay veri seti 2 test verisi.

2.1.2. Gerçek veri setleri

Bu çalışmada gerçek veri seti olarak sağlıklı bir katılımcıya gözleri kapalı olduğu sırada 4 farklı kokunun koklatılmasıyla elde edilen EEG işaretleri kullanılmıştır (Kroupi ve ark., 2014). Bu deney için etik izin ilgili araştırma grubu tarafından alınmıştır. Kullanılan kokular ile bunlardan kaydedilen örnek sayıları şöyledir: 1- 17 örnek kediotu (KO), 2- 21 örnek nilüfer çiçeği (NÇ), 3- 19 örnek peynir kokusu (PK), 4- 23 örnek gül suyu (GS). Veriler rahat bir koltukta oturmakta olan katılımcının deney esnasında yaklaşık 1 metre uzaklıkta bulunan bir ekrana baktığı sırada ekrandan takip ettiği komutlara uyması istenmiştir. İçini göstermeyen şişelerin içerisinde bulunan kokular kişilere 1-2 cm uzakta tutulurken ekranda “Kokla” komutunu görmesiyle otomatik ve rastgele kapağı açılan şişedeki kokuyu 2 saniye boyunca katılımcı koklamıştır. EEG sinyalleri 250 Hz örnekleme frekansı ve toplam 216 elektrot ile kaydedilmiştir. Katılımcının 2 saniye kokladığı verinin son 1 saniyelik kısmı tüm veriden çıkartılmıştır ve işaretler bir koku kaydı için 250 örneğe düşürülmüştür. Dengeli veri seti için KO ile GS örnekleri birleştirilerek bir sınıf ve NÇ ile PK birleştirilerek diğer sınıf oluşturulmuştur. Dengesiz veri seti için ise GS bir sınıfı oluştururken diğer üçüne ait örnekler diğer sınıfı oluşturmuştur. Bu durum Çizelge 2’de gösterilmiştir.

Çizelge 2. Gerçek veri seti özellikleri

Veri Seti	Sınıf 1	Sınıf 2
Gerçek Veri Seti 1	KO+GS (ÖS=40)	NÇ+PK (ÖS=40)
Gerçek Veri Seti 2	GS (ÖS=23)	KO+NÇ+PK (ÖS=57)

2.2. Sınıflandırıcı metrikleri

Sınıflandırıcı metrikleri bir makine öğrenmesi uygulamasının performansını değerlendirmek için kullanılan sayısal değerlerdir. Bu metriklerin birçoğu 0 ile 1 arasında bir değere sahiptir ve yüzde cinsinden de %0 ile %100 arasında hesaplanabilirler. Genel olarak bu değerler büyüdükçe sınıflandırıcı performansı da artmış olur. İki sınıflı bir problemde birçok metrik Çizelge 3’te verilen hata matrisinden hesaplanır. Bu tabloda ele alınan sınıflardan biri “pozitif” ve diğeri “negatif” olarak isimlendirilmek üzere; *doğru pozitif (DP)*: doğru tahmin edilen pozitif örneklerin sayısını, *yanlış pozitif (YP)*: yanlış tahmin edilen (hatalı olarak pozitif sınıfa atanan) negatif örneklerin sayısını, *doğru negatif (DN)*; doğru tahmin edilen negatif örnek sayısını ve *yanlış negatif (YN)*: yanlış tahmin edilen (hatalı olarak negatif sınıfa atanan) pozitif örneklerin sayısını ifade etmektedir.

Çizelge 3. Hata matrisi

		Tahmin Edilen Etiket	
		Pozitif	Negatif
Gerçek Etiket	Pozitif	Doğru Pozitif	Yanlış Negatif
	Negatif	Yanlış Pozitif	Doğru Negatif

Bu değişkenler kullanılarak SD, D, Ö, Jİ ve F-ölçütü (F) sırasıyla aşağıdaki eşitliklerle hesaplanmaktadır (Militello ve ark., 2022).

$$\text{Sınıflandırma Doğruluğu} = \frac{DP + DN}{DP + DN + YP + YN} \quad (1)$$

$$\text{Duyarlılık} = \frac{DP}{DP + YN} \quad (2)$$

$$\text{Özgüllük} = \frac{DN}{DN + YP} \quad (3)$$

$$Jaccard\ indeks = \frac{DP}{DP + YP + YN} \quad (4)$$

$$F\ ölçütü = \frac{2DP}{2DP + YP + YN} \quad (5)$$

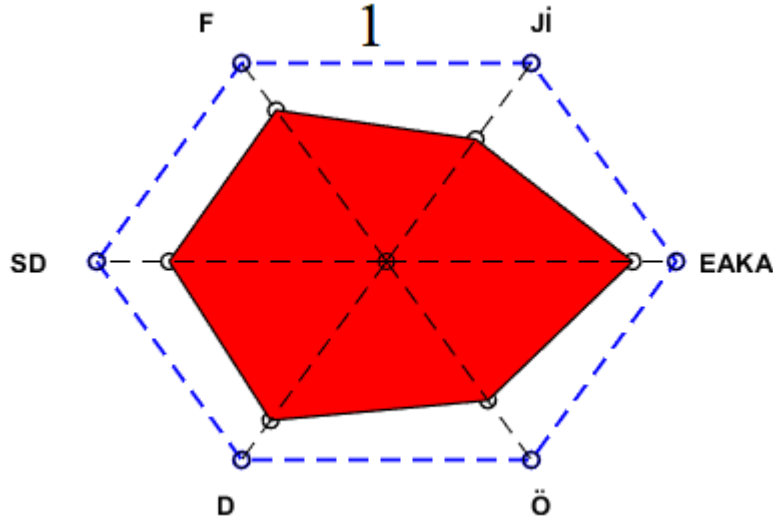
Bu metriklerden başka EAKA metriği Eşitlik 6’da verildiği gibi hesaplanır.

$$Eğri\ altında\ kalan\ alan = \int_0^1 f(x) dx \quad (6)$$

Bu eşitlikte $f(x)$, farklı kesme noktaları için yanlış pozitif oranın ($I-\bar{O}$) fonksiyonunda gerçek pozitif oranın (D) çizildiği bir alıcı işletim karakteristik eğrisidir. PAM ise Şekil 2’de gösterildiği gibi ayırıt uzunluğu 1 olan düzgün bir altıgenin merkezinden köşelerine çizilen doğrular üzerine düşen SD, D, Ö, Jİ, F ve EAKA metriklerinin oluşturduğu kırmızı alandan (KA) hesaplanır. KA bir açısı 60 derece olan 6 adet üçgenden oluşmaktadır. Her bir üçgenin alanı iki kenar uzunluğu ile bu kenarlar arasındaki açının (60 derece) sinüsünün çarpımının yarısına eşittir. Buradan hareketle KA, Eşitlik 7’de verildiği gibi hesaplanabilir. Bu eşitlikte a_i ve b_i ; i . üçgenin 60 derecelik açısına komşu olan kenarları ifade etmektedir. Ayırıt uzunluğu 1 olan düzgün altıgenin alanı 2.59807 olduğundan ve PAM değerinin 0-1 arasına normalize edilmesinden dolayı şekildeki kırmızı alan Eşitlik 8’de verildiği gibi 2.59807 değerine bölünmüştür.

$$KA = \frac{\sqrt{3}}{4} \sum_{i=1}^6 a_i \cdot b_i \quad (7)$$

$$PAM = \frac{KA}{2.59807} \quad (8)$$



Şekil 2. PAM’in hesaplanması.

3. Bulgular

Bu çalışmada PAM’in sık kullanılan altı metriği birlikte görselleştirip bir de bunların oluşturduğu poligon alanını metrik olarak hesaplaması özelliği bu metriğin düzensiz veri setlerinde sınıflandırıcı performanslarının değerlendirilmesinde etkin olarak kullanılabileceği gösterilmiştir.

Bunun için yapay ve gerçek düzenli ve düzensiz veri setleri kullanılmıştır. Yapay veri setleri için öznelikler Matlab ortamında üretilirken, gerçek veri seti için öznelikler (Aydemir, 2017) çalışmasında önerilen gama bandı güçleri kullanılmıştır. Bu öznelikler literatürde sıklıkla kullanılan *k*-en yakın komşuluk (*k*-EYK) (Mahami ve ark., 2021), destek vektör makineleri (DVM) (Sadiq ve ark., 2020), Bayes (Al-Salman ve ark., 2022) ve karar ağacı (KA) (Yuvaraj ve ark., 2021) algoritmaları ile sınıflandırılmıştır. Yapay ve gerçek veri setlerine ait klasik metrik sonuçları Çizelge 4, Çizelge 5, Çizelge 6 ve Çizelge 7’de verilmiştir.

Çizelge 4. Yapay veri seti 1’e ait klasik metrik sonuçları

Sınıflandırıcı	SD	D	Ö	EAKA	Jİ	F
<i>k</i> -EYK	0.89	0.89	0.91	0.89	0.80	0.89
DVM	0.90	0.89	0.91	0.90	0.81	0.90
Bayes	0.90	0.90	0.90	0.90	0.82	0.90
KA	0.85	0.87	0.84	0.85	0.75	0.86

Klasik metrik sonuçlarına bakıldığına hangi sınıflandırıcının diğerine göre üstün olduğuna karar verebilmek için tüm metrikler birbiri ile kıyaslanmalıdır. Ancak, görüldüğü gibi bir metrik bir sınıflandırıcı da daha yüksek olabilirken, diğer metrik bakımından düşük kalabilmektedir. Örneğin Çizelge 4’te *SD* metriği açısından DVM ve Bayes en yüksek sonuca sahipken, *Ö* metriği bakımından *k*-EYK ve DVM en yüksek sonuca sahiptir.

Çizelge 5. Yapay veri seti 2’ye ait klasik metrik sonuçları

Sınıflandırıcı	SD	D	Ö	EAKA	Jİ	F
<i>k</i> -EYK	0.91	0.90	0.92	0.91	0.87	0.93
DVM	0.90	0.92	0.87	0.90	0.86	0.93
Bayes	0.9	0.91	0.91	0.91	0.87	0.93
KA	0.85	0.89	0.78	0.83	0.80	0.89

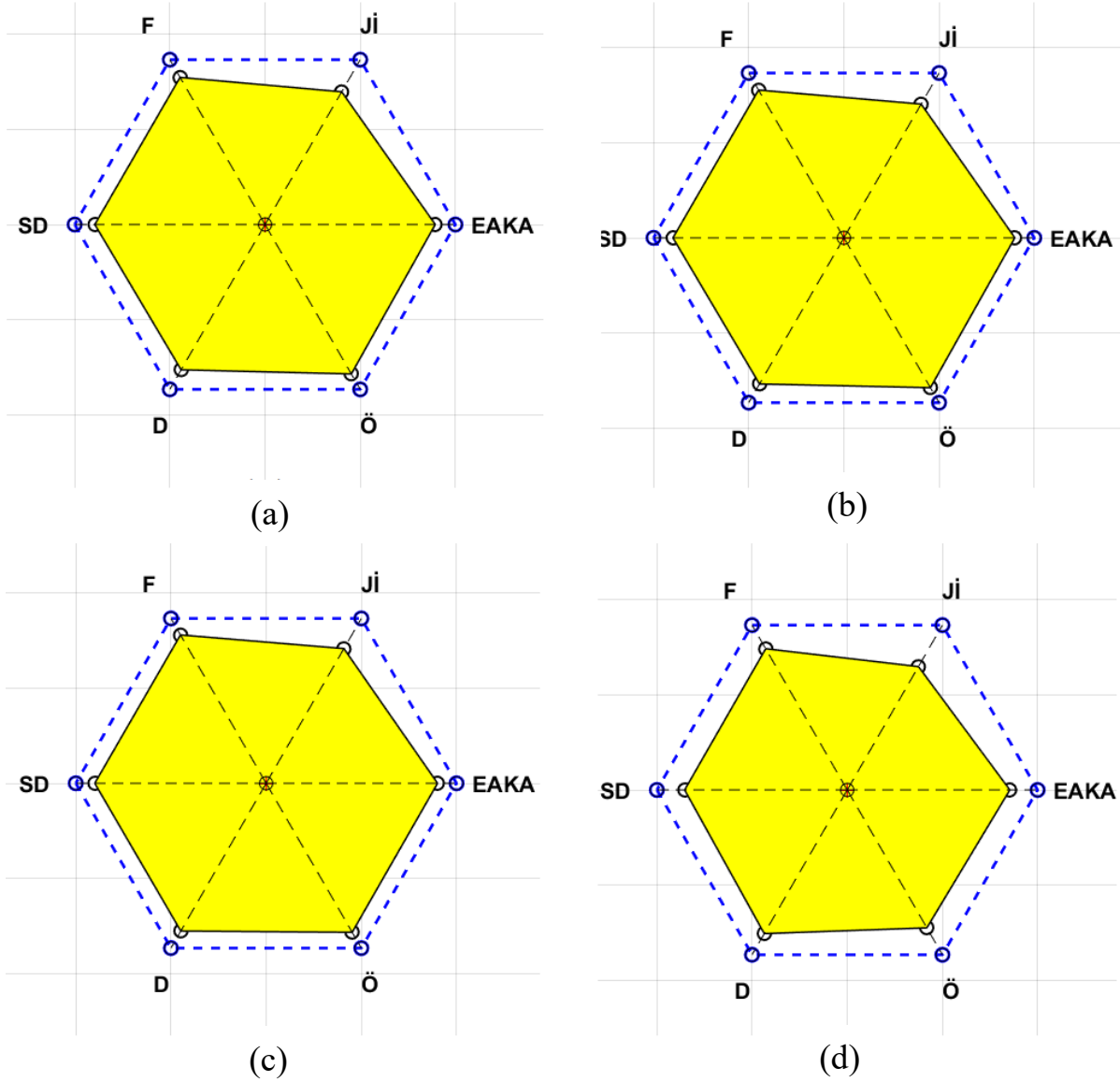
Benzer biçimde Çizelge 5’te *SD* bakımından en yüksek performans 0.91 ile *k*-EYK için elde edilmişken, *D* bakımından en yüksek performans 0.92 ile DVM için, *Jİ* bakımından 0.87 ile *k*-EYK ve Bayes en yüksek performansı sağlamıştır. Benzer durumlar diğer klasik metrik sonuçlarına ait diğer tablolarda da görülmektedir. Daha önce de ifade edildiği gibi, her bir klasik metrik farklı bilgiler sağlıyor olsa da sınıflandırıcı karşılaştırmasında ve en başarılı sınıflandırıcıyı belirlemede zorluklara neden olmaktadır.

Çizelge 6. Gerçek veri seti 1’e ait klasik metrik sonuçları

Sınıflandırıcı	SD	D	Ö	EAKA	Jİ	F
<i>k</i> -EYK	0.93	0.95	0.90	0.93	0.86	0.93
DVM	1.00	1.00	1.00	1.00	1.00	1.00
Bayes	0.70	0.40	1.00	0.70	0.40	0.57
KA	0.63	0.50	0.75	0.63	0.40	0.57

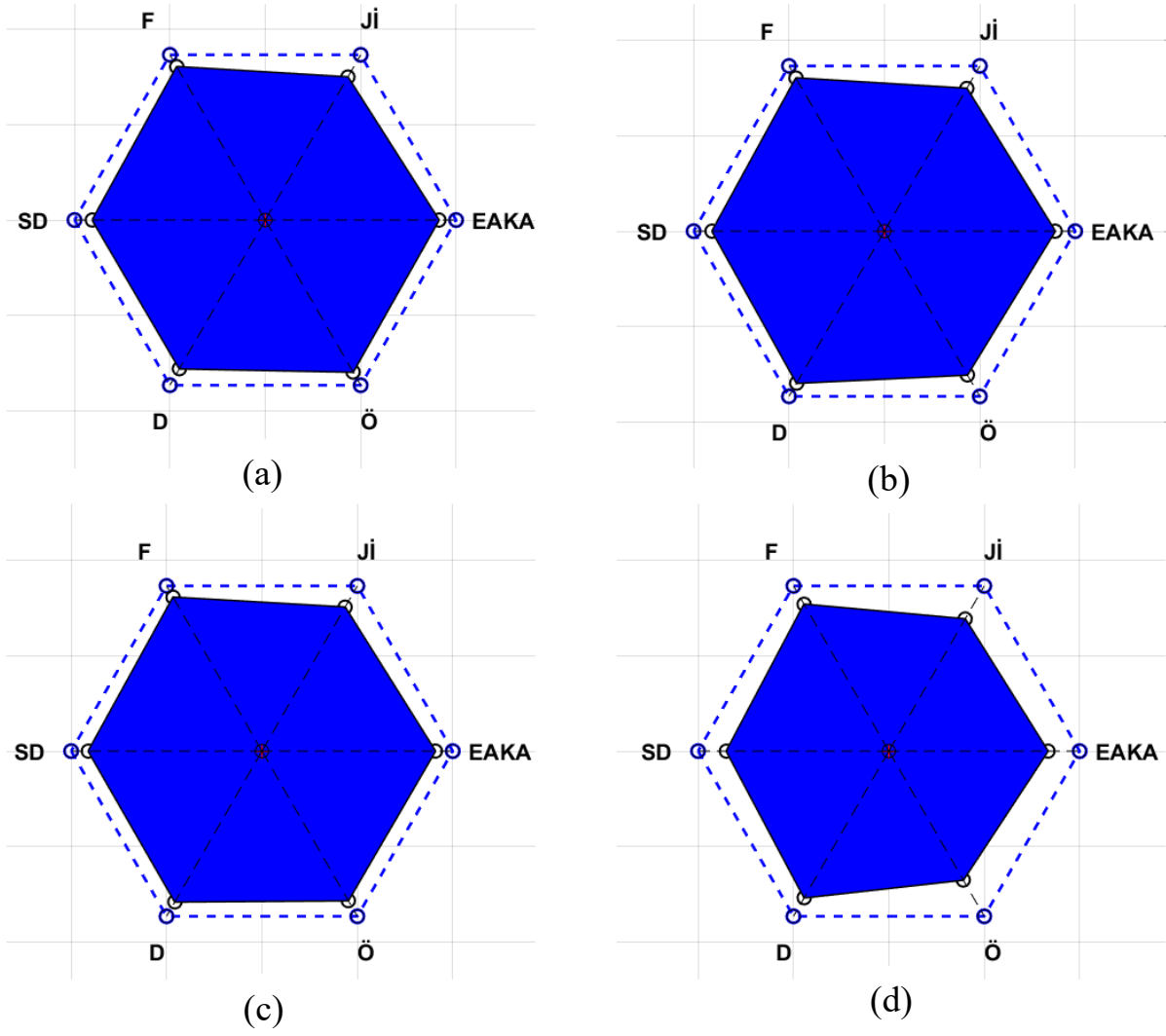
Çizelge 7. Gerçek veri seti 2’ye ait klasik metrik sonuçları

Sınıflandırıcı	SD	D	Ö	EAKA	Jİ	F
<i>k</i> -EYK	0.85	0.93	0.67	0.80	0.82	0.90
DVM	0.81	0.90	0.58	0.74	0.77	0.87
Bayes	0.59	0.52	0.75	0.63	0.47	0.64
KA	0.63	0.66	0.58	0.62	0.56	0.72



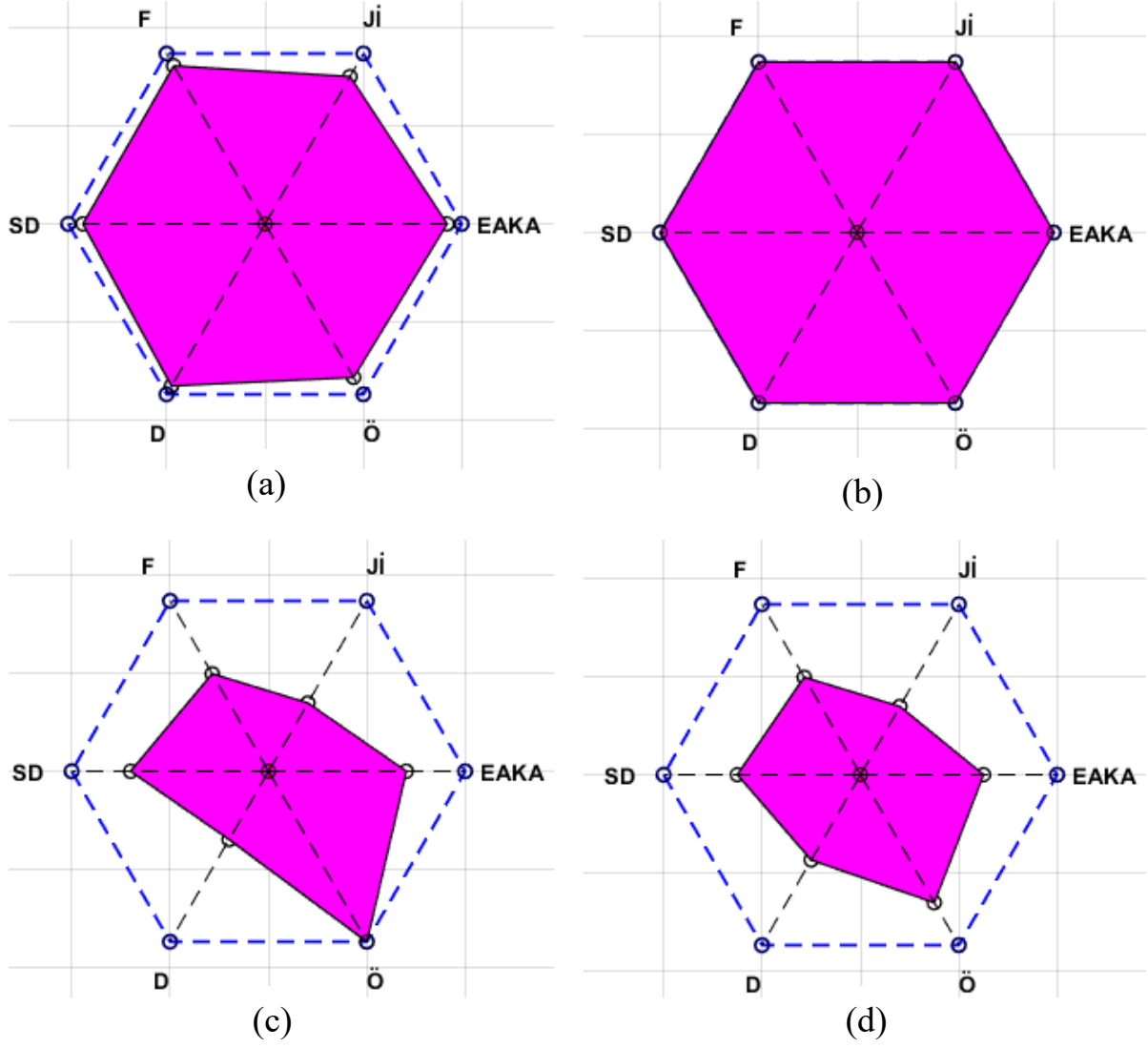
Şekil 3. Yapay veri seti 1'e ait PAM grafiği sonuçları, a) k-EYK, b) DVM, c) Bayes, d) KA.

Bunlara karşılık düzensiz veri setleri için sınıflandırıcı karşılaştırması için önerilen PAM'a ait sonuçlar ise Çizelge 8'de ve Şekil 3, Şekil 4, Şekil 5 ve Şekil 6'da verilmiştir. Özellikle bu şekiller bir sınıflandırıcının hangi klasik metrik açısından güçlü ve zayıf olduğunu hızlı bir şekilde görselleştirirken, bu metriklerin oluşturduğu poligonun alanı da sınıflandırıcının genel başarısını değerlendirmek için kullanılmaktadır. Örneğin Şekil 3'te Jİ'nin diğer metriklerle göre daha düşük başarıya olduğu, Şekil 5-c'de ise Ö bakımından en yüksek başarıya sahip olduğu hızlıca görülebilmektedir.



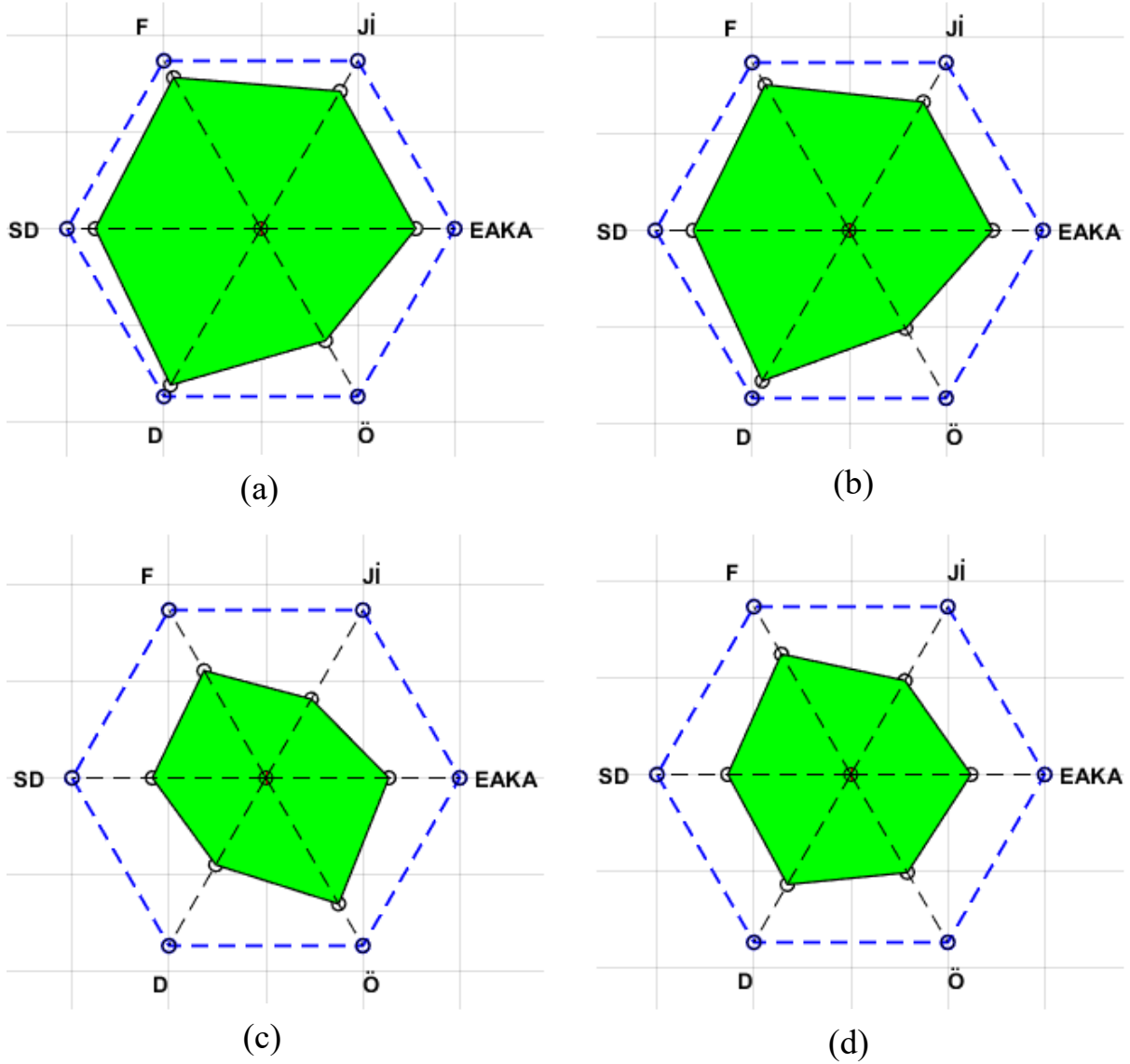
Şekil 4. Yapay veri seti 2'ye ait PAM grafiği sonuçları, a) k-EYK, b) DVM, c) Bayes, d) KA.

Düzensiz veri setlerinin sınıflandırıcı değerlendirilmesinde özellikle SD metriğinin tek başına yeterli anlam taşımadığı daha önce ifade edilmişti. Bu durum Şekil 4-d'de ve Şekil 6-a'da net olarak görülmektedir. Bu şekillerde SD yüksek performans sağlamış gibi görünse de aslında bu ilgili sınıflandırıcının Ö bakımından düşük performansta kaldığı görülmektedir.



Şekil 5. Gerçek veri seti 1'e ait PAM grafiği sonuçları, a) k -EYK, b) DVM, c) Bayes, d) KA.

Sınıflandırıcı karşılaştırmasında ise Çizelge 8'e bakıldığında yapay veri seti 1 için DVM ve Bayes'in, yapay veri seti 2 için k -EYK ve Bayes'in, gerçek veri seti 1 için k -EYK ve DVM'nin ve gerçek veri seti 2 için ise sadece k -EYK'nın en iyi performansa sahip olduğu görülmektedir.



Şekil 6. Gerçek veri seti 2'ye ait PAM grafiği sonuçları, a) k -EYK, b) DVM, c) Bayes, d) KA.

Çizelge 8. Tüm veri setleri için PAM sonuçları

Sınıflandırıcı	Yapay veri seti 1	Yapay veri seti 2	Gerçek veri seti 1	Gerçek veri seti 2
k -EYK	0.77	0.82	0.89	0.75
DVM	0.78	0.80	1.00	0.74
Bayes	0.78	0.82	0.62	0.50
KA	0.70	0.71	0.64	0.69

4. Tartışma ve Sonuç

Bu çalışmada sınıflandırıcıların birbirleriyle ve kendi içinde klasik metrikler bakımından kolayca karşılaştırılmasına imkân veren PAM'ın düzensiz veri setlerinde kullanılabileceği gösterilmiştir. PAM sunduğu görsel ile hem SD, Ö, D, EAKA, Jİ ve F klasik metriklerinin değerlerini sunmakta hem de bu metriklerin oluşturduğu poligonun alanı ile sınıflandırıcının genel performansını ortaya koymaktadır. Böylece klasik metriklerin oluşturduğu büyük tablolar ve sınıflandırıcı

karşılaştırmasındaki zorluklar da ortada kalkmaktadır. PAM'a alternatif olarak aritmetik ortalama, geometrik ortalama veya harmonik ortalama gibi yöntemlerde kullanılabilir ancak bu yöntemler bazı özel durumlar için yanıltıcı sonuçlar üretebilir. Örneğin, geometrik ortalama da herhangi bir çarpan sıfır değerine sahip olursa sonuç da sıfır olur ki bu istenilen bir performans değerlendirme sonucu olmayacaktır. Benzer durum harmonik ortalama da ise tanımsız durum ortaya çıkarır ve ortalama hesaplanamaz. Aritmetik ortalama ise bir veri dizi içerisinde aşırı küçük ya da aşırı büyük değerler var ise aritmetik ortalama o aşırı değere yaklaşma gösterir ki bu da sınıflandırıcı performansını ölçmede yanıltıcı bir durum ortaya koyar. Bunlar yerine klasik metriklerin durumunu bir altıgen üzerinde radar çizimi gibi görselleştirerek özel durumlar (modelin tüm örnekleri tek bir sınıfa atması, hiç doğru sınıflandırma yapılamaması vb.) da dahil olmak üzere tüm durumlar hakkında bilgi sağlanması bakımından PAM önemli bir avantaj sağlamaktadır.

Kaynakça

- Al-Garadi, M. A., Hussain, M. R., Khan, N., Murtaza, G., Nweke, H. F., Ali, I., Mujtaba, G., Chiroma, H., Khattak, H. A., & Gani, A. (2019). Predicting cyberbullying on social media in the big data era using machine learning algorithms: Review of literature and open challenges. *IEEE Access*, 7, 70701-70718. doi: 10.1109/ACCESS.2019.2918354
- Al-Salman, W., Li, Y., Wen, P., Miften, F. S., Oudah, A. Y., & Al Ghayab, H. R. (2022). Extracting epileptic features in EEGs using a dual-tree complex wavelet transform coupled with a classification algorithm. *Brain Research*, 147777. doi: 10.1016/j.brainres.2022.147777
- Alsheikh, M. A., Lin, S., Niyato, D., & Tan, H. P. (2014). Machine learning in wireless sensor networks: Algorithms, strategies, and applications. *IEEE Communications Surveys & Tutorials*, 16(4), 1996-2018. doi: 10.1109/COMST.2014.2320099
- Aydemir, O. (2021). A new performance evaluation metric for classifiers: polygon area metric. *Journal of Classification*, 38(1), 16-26. doi: 10.1007/s00357-020-09362-5
- Aydemir, O. (2017). Olfactory recognition based on EEG gamma-band activity. *Neural Computation*, 29(6), 1667-1680. doi: 10.1162/NECO_a_00966
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874. doi: 10.1016/j.patrec.2005.10.010
- Huang, B., Zhu, Y., Wang, Z., & Fang, Z. (2021). Imbalanced data classification algorithm based on clustering and SVM. *Journal of Circuits, Systems and Computers*, 30(2), 2150036. doi: 10.1142/S0218126621500365
- Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 1. doi: 10.5121/ijdkp.2015.5201
- Kroupi, E., Yazdani, A., Vesin, J. M., & Ebrahimi, T. (2014). EEG correlates of pleasant and unpleasant odor perception. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 11(1), 1-17. doi: 10.1145/2637287
- Liu, H., Li, J., Cao, H., Xie, X., & Wang, Y. (2022). Prediction modeling of geogenic iodine contaminated groundwater throughout China. *Journal of Environmental Management*, 303, 114249. doi: 10.1016/j.jenvman.2021.114249
- Liu, Y., Zhou, Y., Wen, S., & Tang, C. (2014). A strategy on selecting performance metrics for classifier evaluation. *International Journal of Mobile Computing and Multimedia Communications*, 6(4), 20-35. doi: 10.4018/IJMCMC.2014100102
- Mahami, A., Rahmoune, C., Bettahar, T., & Benazzouz, D. (2021). Induction motor condition monitoring using infrared thermography imaging and ensemble learning techniques. *Advances in Mechanical Engineering*, 13(11). doi: 10.1177/16878140211060956
- MATLAB. (2018). MathWorks web sayfası: www.mathworks.com/matlabcentral/fileexchange/74136-polygon-area-metric-for-classifier-evaluation
- Militello, C., Ranieri, A., Rundo, L., D'Angelo, I., Marinozzi, F., Bartolotta, T. V., ... & Russo, G. (2022). On unsupervised methods for medical image segmentation: investigating classic approaches in breast cancer DCE-MRI. *Applied Sciences*, 12(1), 162. doi:10.3390/app12010162

- Mun, J., Jang, W. D., Sung, D. J., & Kim, C. S. (2017, September). *Comparison of objective functions in CNN-based prostate magnetic resonance image segmentation*. IEEE International Conference on Image Processing (ICIP), 3859-3863. doi: 10.1109/ICIP.2017.8297005
- Rodríguez-de-Vera, J. M., Bernabé, G., García, J. M., Saura, D., & González-Carrillo, J. (2022). Left ventricular non-compaction cardiomyopathy automatic diagnosis using a deep learning approach. *Computer Methods and Programs in Biomedicine*, 214. doi:10.1016/j.cmpb.2021.106548
- Roy, A., Singh, B. K., Banchhor, S. K., & Verma, K. (2022). Segmentation of malignant tumours in mammogram images: A hybrid approach using convolutional neural networks and connected component analysis. *Expert Systems*, 39(1). doi: 10.1111/exsy.12826
- Sadiq, M. T., Yu, X., Yuan, Z., & Aziz, M. Z. (2020). Identification of motor and mental imagery EEG in two and multiclass subject-dependent tasks using successive decomposition index. *Sensors*, 20(18). doi: 10.3390/s20185283
- Shia, W. C., & Chen, D. R. (2021). Classification of malignant tumors in breast ultrasound using a pretrained deep residual network model and support vector machine. *Computerized Medical Imaging and Graphics*, 87. doi: 10.1016/j.compmedimag.2020.101829
- Vuttipittayamongkol, P., Elyan, E., & Petrovski, A. (2021). On the class overlap problem in imbalanced data classification. *Knowledge-Based Systems*, 212. doi: /10.1016/j.knosys.2020.106631
- Yeung, H. W. F., Zhou, M., Chung, Y. Y., Moule, G., Thompson, W., Ouyang, W., Cai, W., & Bennamoun, M. (2022). Deep-learning-based solution for data deficient satellite image segmentation. *Expert Systems with Applications*, 191. doi: 10.1016/j.eswa.2021.116210
- Yin, L., Lin, X., Liu, J., Li, N., He, X., Zhang, M., 2021. Investigation on Nutrition Status and Clinical Outcome of Common Cancers (INSCOC) Group. Classification tree-based machine learning to visualize and validate a decision tool for identifying malnutrition in cancer patients. *Journal of Parenteral and Enteral Nutrition*, 45(8), 1736-1748. doi: 10.1002/jpen.2070
- Yuvaraj, N., Chang, V., Gobinathan, B., Pinagapani, A., Kannan, S., Dhiman, G., & Rajan, A. R. (2021). Automatic detection of cyberbullying using multi-feature based artificial intelligence with deep decision tree classification. *Computers & Electrical Engineering*, 92, 107186. doi: 10.1016/j.compeleceng.2021.107186