

PAPER DETAILS

TITLE: Destek Vektör Makinelerinin Wi-Fi Tabanlı İç Mekan Lokalizasyon Tespitinde Kullanımı ve Çekirdek Fonksiyon Seçiminin Sınıflandırma Performansına Etkisi

AUTHORS: Ebru EFEOGLU

PAGES: 1370-1382

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/2191975>

Destek Vektör Makinelerinin Wi-Fi Tabanlı İç Mekan Lokalizasyon Tespitinde Kullanımı ve Çekirdek Fonksiyon Seçiminin Sınıflandırma Performansına Etkisi: Anaokulu Öğrencilerinin Konumlarını Bulma Örneği

Ebru EFEÖĞLU^{1*}

¹Kütahya Dumlupınar Üniversitesi, Mühendislik Fakültesi, Yazılım Mühendisliği Bölümü, Kütahya

¹<https://orcid.org/0000-0001-5444-6647>

*Sorumlu yazar: ebru.efeoglu@dpu.edu.tr

Araştırma Makalesi

ÖZ

Makale Tarihçesi:

Geliş tarihi: 14.01.2022

Kabul tarihi: 07.05.2022

Online Yayınlanma: 12.12.2022

Anahtar Kelimeler:

Kablosuz sinyal gücü

Destek vektör makineleri

Çekirdek fonksiyon

İç mekan konum belirleme

Son yıllarda çocuk güvenliğindeki sorunlar nedeniyle iç mekanlarda çocukların yerini tespit etme çalışmalarının önemi artmıştır. Bu çalışmada kablosuz sinyal gücü ve Destek Vektör Makineleri sınıflandırma algoritması kullanılarak farklı odalardaki çocukların konumları tespit edilmiştir. Ayrıca algoritmanın sınıflandırma performansı üzerindeki etkileri farklı çekirdek fonksiyonları kullanılarak incelenmiş ve performans analizi için çapraz doğrulama tekniği kullanılmıştır. Performans analizinde, Karışıklık matrisi, Kesinlik, Duyarlılık, F-Skoru, AUC, Doğruluk, Kappa istatistik değeri ve Kök Ortalama Kare Hata değeri incelenmiştir. Analizi sonucunda doğrusal çekirdek fonksiyonları tercih edildiğinde en yüksek performans metrikleri elde edilmiş ve %97,9 başarı elde edildiği gözlemlenmiştir.

The Use of Support Vector Machines in Wi-Fi Based Indoor Localization Detection and the Effect of Kernel Function Selection on Classification Performance: An Example of Finding the Location of Kindergarten Students

Research Article

ABSTRACT

Article History:

Received: 14.01.2022

Accepted: 07.05.2022

Published online: 12.12.2022

Keywords:

WiFi

Support vector machines

Kernel

Indoor localization

In recent years, the importance of locating children indoors has increased due to the problems in child safety. In this study, the locations of children in different rooms were determined by using wireless signal strength and Support Vector Machines classification algorithm. In addition, the effects of the algorithm on the classification performance were examined using different kernel functions and cross-validation technique was used for performance analysis. In the performance analysis, Confusion matrix, Precision, Sensitivity, F-Score, AUC, Accuracy, Kappa statistical value and Root Mean Square Error value were examined. As a result of the analysis, when linear kernel functions are preferred, the highest performance metrics were obtained and 97.9% success was observed

To Cite: Efeoglu E. Destek Vektör Makinelerinin Wi-Fi Tabanlı İç Mekan Lokalizasyon Tespitinde Kullanımı ve Çekirdek Fonksiyon Seçiminin Sınıflandırma Performansına Etkisi: Anaokulu Öğrencilerinin Konumlarını Bulma Örneği. Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi 2022; 5(3): 1370-1382.

1. Giriş

Anaokulları çocukların sosyalleştiği, bilgi ve beceri öğrendiği, deneyimler elde ettiği yerlerdir (Yıldız ve Bektaş, 2021). Çocuklar okullarda çevresini tanıma merakı içerisindeyler ve onlar için okul yapısı birer öğretim aracıdır (Sıvacılar, 2021). Bu nedenle çocuk psikolojisinin gelişimini, güvenliğini, işlevselliğini geliştirmek için anaokulu iç mekân oyun alanlarına yönelik kavramsal yaklaşımlar incelenmiş ve uygun alanlar tasarlanmıştır (Öztürk ve Tayılga, 2021). Son zamanlarda ebeveynlerin

anaokulların güvenliğine ilişkin endişelerinin artması okullarda güvenlik önlemlerinin arttırılmasına yönelik talepleri de arttırmıştır (Karadaş, 2021). Okullarda güvenliğin sağlanması öğrencilerin sağlığı ve ebeveynlerin huzuru için gereklidir (Karadaş, 2021). Okullardaki güvenlik probleminin aşılmasında yeni teknolojilerden faydalanılmaktadır. Oldukça karmaşık ve zor olan konum belirleme uygulamaları, sayısı ve kullanım oranı her geçen gün artan akıllı telefonlar (Caneren ve ark., 2021) sayesinde yapılabilmektedir. Örneğin, android tabanlı bir akıllı telefon ile Web ve GPS tabanlı hizmetler için yerel uygulamalar tasarlanmıştır (Shu ve ark., 2009). Java tabanlı akıllı telefonlarda GPS sistemlerinden yararlanılarak kullanıcının konumunu gösteren sistemler (Chandra ve ark., 2011), acil durum senaryoları (Liao ve ark., 2011) oluşturulmuştur. Yine Android işletim sistemine sahip bir telefona web tabanlı çocuk takip sistemi oluşturulmuştur (Bayıroğlu ve Ayan, 2014). Çocukların dış ortamda konumunun belirlenmesinde GPS yönteminin kullanılmasına karşın iç ortamlarda bu yöntem iyi sonuçlar vermemektedir (Bejuri ve Mohamad, 2014). Nedeni uydu sinyallerinin iç ortamlarda bulunan duvarlar nedeniyle zayıflamasıdır. Anaokulu çocuklarının güvenliğinin tam olarak sağlanabilmesi için bakım personeline yardımcı olabilecek, çocukların okul içindeki konumlarını belirleyebilecek sistemlere ihtiyaç duyulmaktadır. Bina içinde konum belirleme çalışmalarında Bluetooth sinyal gücüne dayalı sistemler (Rida ve ark., 2015), Radyo Frekans sinyal gücü ölçümlerine dayalı sistemler (Seco ve ark., 2013) ve çoklu optik alıcılar (Yasir ve ark., 2015) kullanılmıştır. Ayrıca (Chen ve Huang, 2009) tarafından yeni bir konum belirleme yöntemi önerilmiştir. Bu çalışmada kablosuz sinyal gücü kullanılarak bina içi konum tespiti yapılmaktadır.

Çalışmanın amacı anaokullarındaki çocukların bina içindeki konumlarını yüksek doğruluk oranı ile belirleyebilmektir. Bina içinde çocukların yeri tespit edilebildiğinde Anaokulunda çalışan bakım personelinin iş yükü azalacak ve çocukların güvenliği sağlanmış olacaktır. Güvenlik sorunu yaşamayan ebeveynler çocuklarını gönül rahatlığı ile okula gönderebilecek ve daha verimli çalışabileceklerdir.

2. Materyal ve Metot

Bu bölüm, çalışmada kullanılan veri seti, sınıflandırma algoritmaları ve performans metrikleri hakkında bilgiler içermektedir.

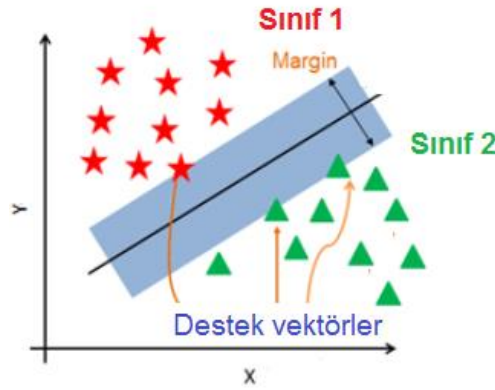
2.1. Veri Seti

Çalışmada kullanılan veri seti UCI (Frank, 2010) kütüphanesinden alınmıştır. Veri setinde farklı odalarda bulunan insanların cep telefonları ile 7 farklı kablosuz sinyal gücü ölçümünden oluşan 2000 adet ölçüm bulunmaktadır. Daha önce bu veriler (Rohra ve ark., 2017) tarafından kullanılarak bulanık parçalı sürü optimizasyonu yöntemi ile konum belirleme çalışması yapılmıştır. Ayrıca bu veriler konum belirleme çalışmasında farklı sınıflandırma algoritmalarının karşılaştırılmasında kullanılmıştır (Sabancı ve ark., 2018). Bu çalışmada ise Destek Vektör Makineleri (DVM) algoritması ile konum

belirleme çalışması yapılmış ve farklı çekirdek fonksiyon seçiminin DVM algoritmasının konum bulma performansına etkisi de incelenmiştir.

2.2. Destek Vektör Makinesi (DVM)

İstatistiksel öğrenme teorisini temel alan parametrik olmayan bir makine öğrenmesi yöntemidir (Cortes ve Vapnik, 1995) Algoritma, sınıfları birbirinden ayırabilecek en uygun hiper düzlemi bulmaya çalışır. Hiper düzlem ve Destek Vektörler Şekil 1’de gösterilmiştir. Burada hiper düzlem, farklı sınıflara ait elemanlar arasında ayrım yapan bir karar düzlemidir. Bu düzleme en yakın olan veri noktaları da destek vektörleri oluşturmaktadır. En yakın sınıf noktalarındaki iki çizgi arasındaki boşluk Marj olarak adlandırılır. Bu değerin büyük olması marjın iyi olduğunu ifade eder (Cortes ve Vapnik, 1995).



Şekil 1. Hiper düzlem ve destek vektörler

Doğrusal olmayan verilerin sınıflandırılmasında DVM kullanıldığında doğrusal olarak ayrılmayan girdi uzayının bir üst boyuta değişik yapıdaki çekirdek fonksiyonlarla taşınır. Bu çalışmada değişik DVM modellerini oluşturmak amacıyla Doğrusal, Gaussian, Polinomial, Radyal tabanlı fonksiyon ve Sigmoid çekirdek fonksiyonları kullanılmıştır.

Doğrusal çekirdek fonksiyonu: Doğrusal çekirdek, en basit çekirdek fonksiyondur. Doğrusal bir çekirdek, verilen herhangi iki gözlemde normal nokta çarpımı olarak kullanılabilir. İki vektör arasındaki çarpım, her bir girdi değeri çiftinin çarpımının toplamıdır ve aşağıdaki gibi ifade edilir.

$$K(x_i, x_j) = x_i^T x_j \quad (1)$$

Burada $K(x_i, x_j)$ kernel fonksiyonunu temsil eder.

Polinomial çekirdek fonksiyonu: Bir polinom çekirdeği, doğrusal çekirdeğin daha genelleştirilmiş bir şeklidir. Polinom çekirdeği, eğri veya doğrusal olmayan girdi uzayını ayırt edebilir. Polinomial çekirdek,

$$K(x_i, x_j) = (\alpha x_i^T x_j + p)^d \quad (2)$$

şeklinde ifade edilir. Buradaki x_i ve x_j girdi uzayındaki vektörlerdir. Denklemden alfa (α) eğimi, p sabit terimi ve d polinom derecesini göstermektedir. Polinomal çekirdeği, tüm eğitim verilerinin normalleştirildiği problemler için uygundur.

Gaussian radyal tabanlı fonksiyon: Bazı girdi uzaylarında öznelik vektörleri olarak temsil edilen iki örnek x_i ve x_j üzerindeki Radyal tabanlı fonksiyon çekirdeğinin bir örneği olan Gauss çekirdeği ise

$$K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2) \quad (3)$$

şeklinde tanımlanır. Burada, $\|x_i - x_j\|^2$ iki özellik vektörü arasındaki kare Öklid mesafesi olarak kabul edilebilir. σ , Gauss çekirdeğinin genişliğini belirleyen ve çekirdeğin performansını önemli ölçüde etkileyen bir parametredir.

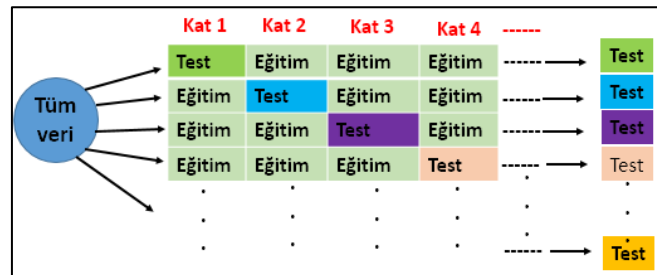
Sigmoid çekirdek fonksiyonu: Bu fonksiyon çekirdeği Hiperbolik Tanjant Çekirdeği ve Çok Katmanlı Algılayıcı çekirdeği olarak da adlandırılır. Eğimi α ve kesme sabitini δ ile gösterilirse Sigmoid fonksiyonu,

$$K(X_i, X_j) = \tanh(\alpha x_i^T x_j + \delta) \quad (4)$$

gibi ifade edilebilir. α için ortak bir değer ($1 / N$) kullanılabilir. Burada N değeri veri boyutunu ifade etmektedir.

2.3. Performans Değerlendirme Yöntemi ve Performans Metrikleri

Model istatistiksel performansını aynı veriler üzerinde eğitmek ve değerlendirmek, aşırı iyimser bir sonuç verir. Bu problemin üstesinden gelmek için algoritmanın çıktısının yeni veriler üzerinde test edilmesi gerekir (Arlot ve Celisse, 2010). Bunun için veri setinin parçalara ayrılması ve bir kısmının eğitim için bir kısmının ise test için kullanılması gerekir. Fakat veri seti parçalara ayrıldığında verinin dağılımına göre modelin eğitim ve testinde bazı sapmalar oluşabilir. Bu sapmaları en aza indirebilmek için k -kat çapraz doğrulama işlemi uygulanır (Şekil 2). k -kat çapraz doğrulama yönteminde veri seti k eşit parçaya bölünür ve her parçanın eğitim ve test verisi olarak kullanımı sağlanır. k genellikle bu çalışmada olduğu gibi 10 seçilir.



Şekil 2. k -kat çapraz doğrulama yöntemine ilişkin grafiksel gösterim

2.3.1. Performans Metrikleri

Sınıflandırma yapılırken genellikle pozitif ve negatif olmak üzere iki sınıf tanımlanır (Orozco-Arias ve ark., 2019). Bu nedenle, beklenen sonuçlar pozitif olarak sınıflandırıldıysa ve pozitif sınıfta yer alıyorsa doğru pozitif (DP), reddedilmiş ancak negatif sınıfa ait değilse yanlış negatif (YN) olarak sınıflandırılabilir. Öte yandan, negatif sınıfta yer alan ve pozitif olduğu tahmin edilen örnekler yanlış pozitif (YP), değilse doğru negatif (DN) oluşturur (Orozco-Arias ve ark., 2019). Bu sonuçlardan yararlanılarak performans metrikleri hesaplanır. Kullanılan veri seti sınıflandırılırken performans ölçümünde kullanılacak metriklerin seçiminin iyi yapılması gerekir. Her metriğin değerlendirilen algoritmaların çeşitli yönlerini ölçen belirli özellikleri vardır (Li ve ark., 2009). Sınıflandırma performansı hakkında en fazla bilgi içeren metrik karışıklık matrisidir ve matris yapısı Şekil 3'te verilmiştir. Matriste mavi renkle gösterilenler DP ve DN değerleridir. Bu değerler doğru tahmin edilen verilerin sayısını gösterir. Matriste bulunan diğer değerler YN ve YP değerleridir. Bu değerlerde yanlış tahmin edilen verilerin sayısını göstermektedir.

		Tahmin edilen Sınıf			
		A	B	C	D
Gerçek Sınıf	A	Kişinin A odasında olduğu tahmin edilmiş ve kişi gerçekten de A odasında	Kişinin B odasında olduğu tahmin edilmiş ve kişi gerçekte A odasında	Kişinin C odasında olduğu tahmin edilmiş ve kişi gerçekte A odasında	Kişinin D odasında olduğu tahmin edilmiş ve kişi gerçekte A odasında
	B	Kişinin A odasında olduğu tahmin edilmiş ve kişi gerçekte B odasında	Kişinin B odasında olduğu tahmin edilmiş ve kişi gerçekten de B odasında	Kişinin C odasında olduğu tahmin edilmiş ve kişi gerçekte B odasında	Kişinin D odasında olduğu tahmin edilmiş ve kişi gerçekte B odasında
	C	Kişinin A odasında olduğu tahmin edilmiş ve kişi gerçekte C odasında	Kişinin B odasında olduğu tahmin edilmiş ve kişi gerçekte C odasında	Kişinin C odasında olduğu tahmin edilmiş ve kişi gerçekten de C odasında	Kişinin D odasında olduğu tahmin edilmiş ve kişi gerçekte C odasında
	D	Kişinin A odasında olduğu tahmin edilmiş ve kişi gerçekte D odasında	Kişinin B odasında olduğu tahmin edilmiş ve kişi gerçekte D odasında	Kişinin C odasında olduğu tahmin edilmiş ve kişi gerçekte D odasında	Kişinin D odasında olduğu tahmin edilmiş ve kişi gerçekten de D odasında

Şekil 3. Karışıklık matrisi yapısı

Bu matris kullanılarak diğer metrikler hesaplanır. Kesinlik doğru tahminlerin yüzdesi (Schietgat ve ark., 2018) Duyarlılık doğru tespit edilen gerçek örneklerin yüzdesidir (Rawal ve Ramaswamy, 2011), Kesinlik ve Duyarlılık değerleri,

$$\text{Kesinlik} = \frac{DP}{DP+YP} \quad (5)$$

$$\text{Duyarlılık} = \frac{DP}{YN + DP} \quad (6)$$

biçimindedir. Bu metriklerin birlikte kullanılması tavsiye edilmektedir fakat sadece bu metrikler algoritmanın performansını tam olarak ifade etmediğinden diğer metriklere de ihtiyaç duyulmaktadır.

Örneğin Düşük duyarlılık ve yüksek kesinliğe sahip iki modeli ve bunun tersini karşılaştırmak zordur. Bu durumda F-ölçü değeri kullanılır. F-ölçütü aşağıdaki bağıntı ile hesaplanabilir.

$$F_Skoru = \frac{2 * Duyarlilik * Kesinlik}{Duyarlilik + Kesinlik} \quad (7)$$

Bir başka kullanılan performans ölçme metriği de Kappa istatistik değeridir. Bu değer, sınıflandırma sonucunda yapılan sınıflandırma ile bir veri kümesindeki gerçek sınıflandırmalar arasında ne kadar uyum olduğunu ölçmek için kullanılır. Sınıflandırma algoritmasının olasılıksal doğruluğu $P(x)$, Aynı veri kümesinde yapılan sınıflandırma olasılığının ağırlıklı ortalaması $P(y)$ ile gösterilirse Kappa değeri şu şekilde ifade edilir.

$$Kappa = \frac{P(x) - P(y)}{1 - P(y)} \quad (8)$$

ROC eğrisinin altında kalan alanın değeri de bize sınıflandırma performansı hakkında bilgi verir. Fakat pozitif ve negatif veri kümeleri dengesiz olması durumunda ROC eğrileri kümelerinde (Kamath, De Jong, ve Shehu, 2014) anlamsız performans ölçümleri yapabilir. Bir ROC eğrisinde, yatay eksen yanlış pozitif oranı (YPO), dikey eksen doğru pozitif oranı (DPO) gösterir. Bu eğrinin altındaki alan (AUC) sınıflandırma ölçütü olarak kullanılır. YPO ve DPO değerleri sırasıyla aşağıda belirtildiği gibi hesaplanır:

$$YPO = \frac{YP}{YP + DN} \quad (9)$$

$$DPO = \frac{DP}{DP + YN} \quad (10)$$

Yine dengesiz veri setlerinde yanıltıcı sonuçlar veren doğruluk metriği şu şekilde hesaplanır:

$$Doğruluk = \frac{DP + DN}{DP + DN + YP + YN} \quad (11)$$

RMS, gerçek değerler ile model tarafından tahmin edilen değerler arasındaki farkları ölçeklendirmek için kullanılan hata değeri ölçütlerinden en çok tercih edilen hata ölçütüdür.

$$RMS = \sqrt{\frac{1}{n} \sum_{k=1}^n (T_{ik} - A_k)^2} \quad (12)$$

Şeklinde hesaplanır. Burada T_{ik} tahmin edilen değerdir ve A_k objektif değerdir. Hata değerinin sıfıra yaklaşması, sınıflandırma algoritmasının doğru tahmininin arttığı anlamına gelir.

3. Bulgular ve Tartışma

Farklı amaçlar için iç mekanda konum belirleme çalışmaları yapılmıştır. Yapılan çalışmalarda makine öğrenimi tekniklerinden sıklıkla yararlanılmıştır. İnsanların konumlarının otomatik olarak, kısa bir süre içinde belirlenebilmesi özellikle çocuk güvenliği açısından önemlidir. Makine öğrenmesi tekniklerinden yararlanılarak Wi-fi tabanlı konum belirleme çalışmalarında genellikle algoritma performans değerlendirmesi yapılırken sadece doğruluk değeri ve karışıklık matrisi göz önüne alınarak yapılmıştır (Sabancı ve ark., 2018). Oysaki bir algoritmanın sınıflandırma performansının değerlendirmesi için tek başına bu değerler yeterli değildir. Bununla birlikte literatürde bulunan çalışmalar incelendiğinde Yapay sinir ağları, Aşırı öğrenme algoritması, Karar ağaçları algoritmalarının kullanıldığı görülmüştür. Bu algoritmalar ile yapılan sınıflandırmalarda elde edilen doğruluk değerleri sırasıyla 96,6, 96,6 ve 95,8'dir (Sabancı ve ark., 2018). DVM algoritması ise doğrusal fonksiyon kullanıldığında %97,9 doğruluk oranı elde edilmiştir. Bu nedenle DVM algoritmasının Wi-fi tabanlı iç mekan konum belirlemede bu algoritmalarından daha başarılı olduğu söylenebilir. Bu çalışmada DVM algoritmasının performans değerlendirmesi yapılırken farklı metrikler göz önüne alınarak yapılmıştır. Kullanılan metriklerden ilki karışıklık matrisidir. Bu çalışmayı diğer çalışmalardan ayıran özelliklerden biri de performans değerlendirmesi yapılırken hem çapraz doğrulama öncesi hem de çapraz doğrulama sonrası performans analizinin yapılmasıdır. Çapraz doğrulama öncesi ve sonrası yapılan sınıflandırma sonucu elde edilen karışıklık matrisi Şekil 4'te verilmiştir. Şekil incelendiğinde Doğrusal fonksiyon ve polinomal fonksiyonun birbirine yakın sonuçlar verdiği ama doğrusal fonksiyonun daha fazla doğru tahmin sayısına sahip olduğu görülmektedir. Radyal temel fonksiyonu çapraz doğrulama öncesinde bütün odaları doğru tahmin etmesine karşın çapraz doğrulama sonrası yapılan sınıflandırmada doğru tahmin sayısında epeyce bir düşüş gözlenmiştir. Sigmoid fonksiyon kullanıldığında ise çapraz doğrulama öncesinde sadece D odasında bulunan kişileri doğru sınıflandırdığı görülmüştür. Algoritmanın çapraz doğrulama sonrası D odası doğru tahmin sayısının da düşüş gözlenmiştir.

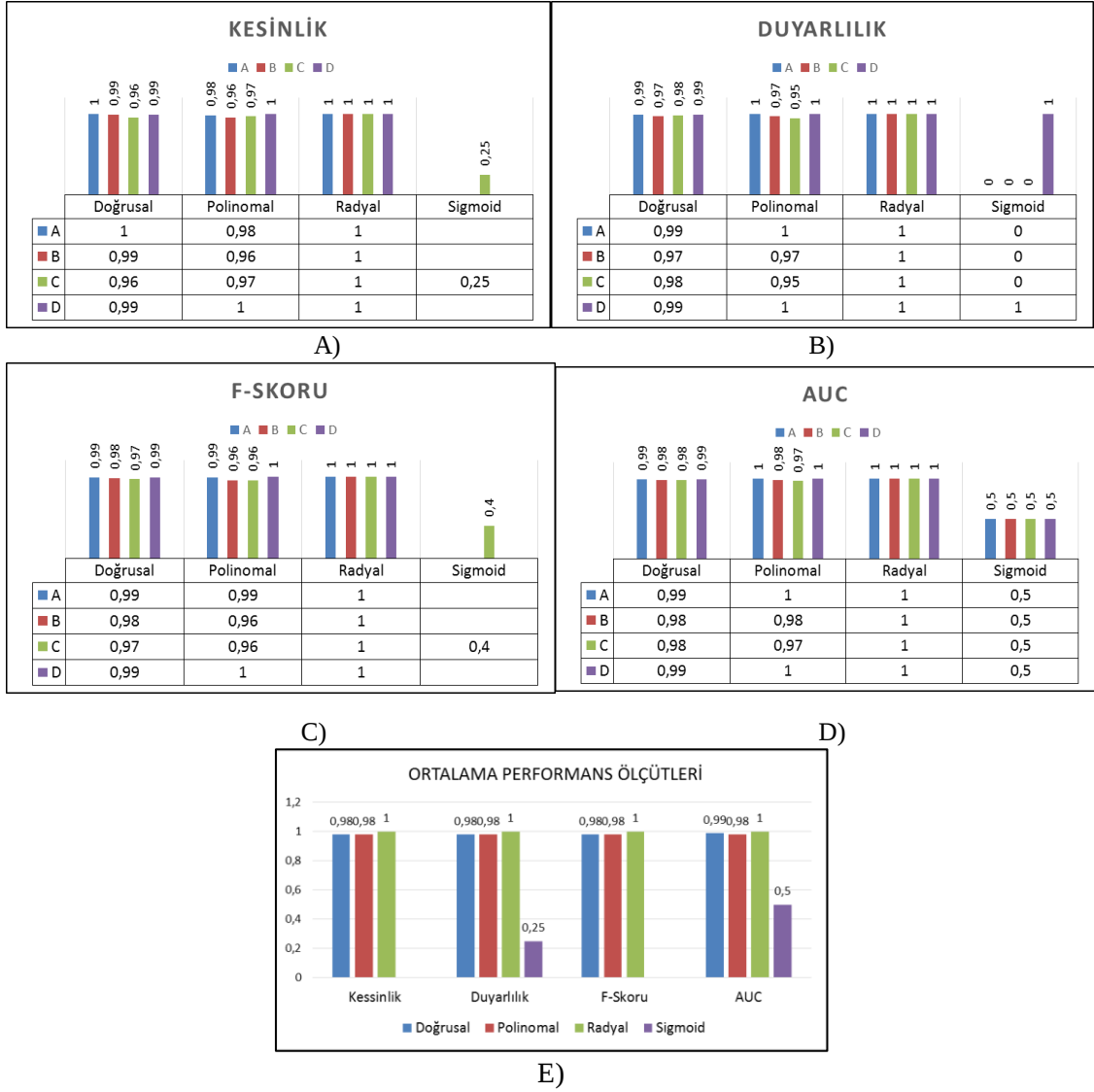
Diğer performans metriklerinin sonuçları çapraz doğrulama öncesi ve sonrası ayrı şekil ve tablolarda verilmiştir. Çapraz doğrulama öncesi yapılan sınıflandırma sonucu elde edilen performans metriklerinden Kappa, doğruluk ve hata değeri Tablo 1'de diğer performans metrikleri grafikler halinde Şekil 5'te verilmiştir. Tablo 1'e bakıldığında en düşük Kappa, doğruluk değeri ve en yüksek hata değeri sigmoid fonksiyonu kullanımında elde edilmiştir. Kappa değerinin düşük olması tahmin edilen sınıf ile gerçek sınıflar arasında düşük bir uyum olduğunu gösterir (Cohen, 1960). Doğruluk değerinin düşük olması da bu fonksiyon kullanılarak yapılan sınıflandırmada doğru tahmin edilen veri sayısının az olduğunu ifade etmektedir. Bu sonuç Sigmoid fonksiyon kullanılması durumunda algoritmanın başarısız sonuçlar verdiğini göstermektedir. En başarılı sonuç doğrusal fonksiyon

kullanımında elde edilmiştir. Doğrusal fonksiyon kullanımında Kappa değeri 0,98 doğruluk değeri de %98,8’dir.

DOĞRUSAL													
		Tahmin edilen Sınıf							Tahmin edilen Sınıf				
Çapraz doğrulama öncesi	Gerçek Sınıf		A	B	C	D	Çapraz Doğrulama sonrası	Gerçek Sınıf		A	B	C	D
		A	499	0	1	0			A	495	0	3	2
		B	0	485	15	0			B	0	482	18	0
		C	0	5	493	2			C	3	6	488	3
		D	0	0	1	499			D	2	0	4	494
POLİNOMAL													
		Tahmin edilen Sınıf							Tahmin edilen Sınıf				
Çapraz doğrulama öncesi	Gerçek Sınıf		A	B	C	D	Çapraz Doğrulama sonrası	Gerçek Sınıf		A	B	C	D
		A	500	0	0	0			A	496	0	1	3
		B	0	487	13	0			B	0	476	24	0
		C	1	20	479	0			C	3	28	466	3
		D	0	0	0	500			D	2	0	3	495
RADYAL													
		Tahmin edilen Sınıf							Tahmin edilen Sınıf				
Çapraz doğrulama öncesi	Gerçek Sınıf		A	B	C	D	Çapraz Doğrulama sonrası	Gerçek Sınıf		A	B	C	D
		A	500	0	0	0			A	376	124	0	0
		B	0	500	0	0			B	0	499	1	0
		C	0	0	500	0			C	0	149	351	0
		D	0	0	0	500			D	0	145	0	355
SİGMÖİD													
		Tahmin edilen Sınıf							Tahmin edilen Sınıf				
Çapraz doğrulama öncesi	Gerçek Sınıf		A	B	C	D	Çapraz Doğrulama sonrası	Gerçek Sınıf		A	B	C	D
		A	0	0	0	500			A	200	150	50	100
		B	0	0	0	500			B	200	150	50	100
		C	0	0	0	500			C	200	150	50	100
		D	0	0	0	500			D	200	150	50	100

Şekil 4. Çapraz doğrulama öncesi ve sonrası elde edilen karışıklık matrisi

Şekil 5 incelendiğinde Sigmoid fonksiyonu kullanımında A,B ve D odası için kesinlik ve F-skoru değeri, A,B ve C odası için duyarlılık değerinin hesaplanamadığı görülmektedir. AUC değeri ise tüm odalar için 0,5 olarak hesaplanmıştır. Ortalama metrik değerleri incelendiğinde en yüksek değerlerin doğrusal fonksiyon kullanımında en düşük değerlerin ise sigmoid fonksiyon kullanımında gerçekleştiği görülmektedir.

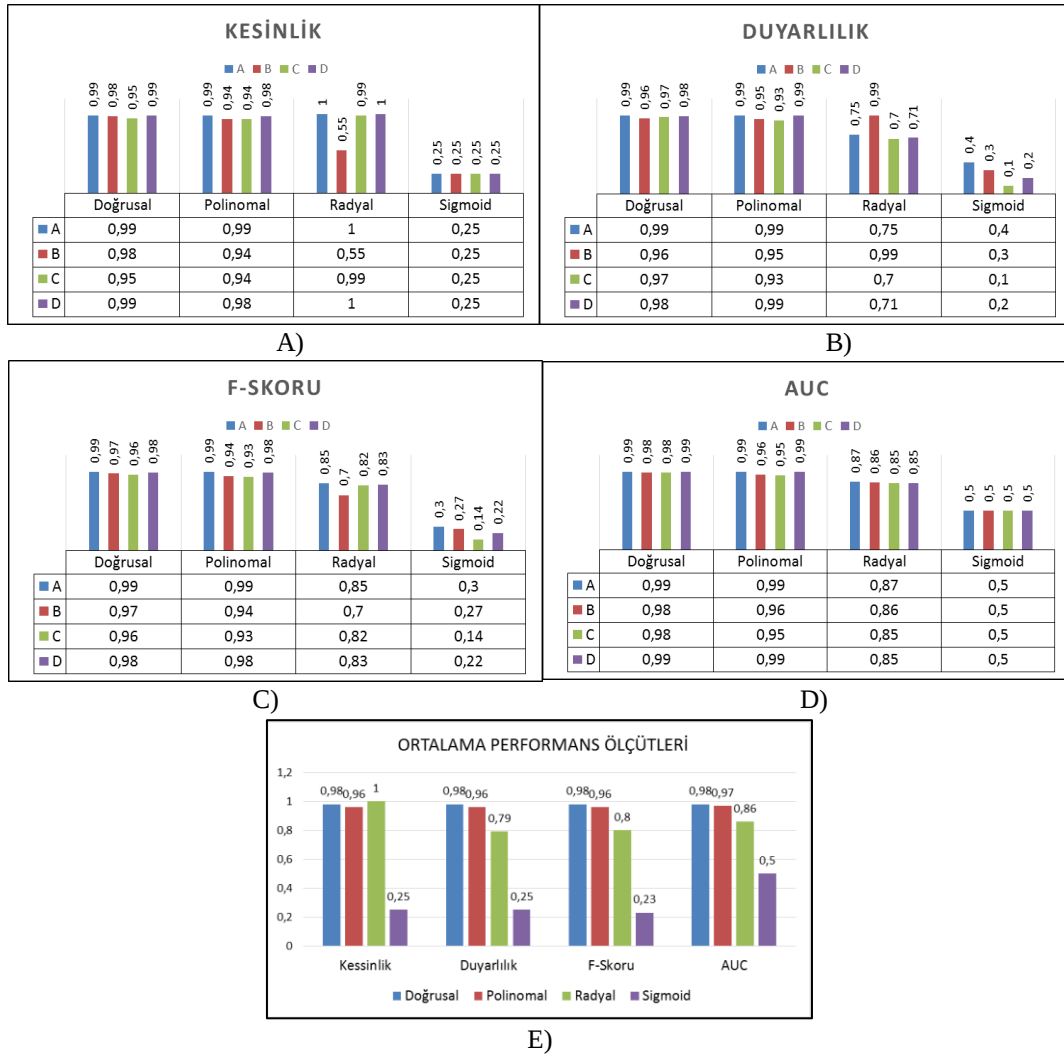


Şekil 5. Çapraz doğrulama öncesi performans metrikleri a)Kesinlik b)Duyarlılık c)F-skoru d)AUC e)Metriklerin ortalama değerleri

Tablo 1. Çapraz doğrulama yapılmadan elde edilen performans metrikleri

Algoritma	Kappa	Doğruluk (%)	Rms
Doğrusal	0,98	98,8	0,07
Polinomal	0,97	98,3	0,09
Radyal	1	100	0
Sigmoid	0	25	0,62

Çapraz doğrulama uygulaması sonrası sınıflandırma performansının incelenmesi için hesaplanan performans metriklerini gösteren grafikler Şekil 6’da verilmiştir.



Şekil 6. Çapraz doğrulama sonrası performans metrikleri a)Kesinlik b)Duyarlılık c)F-skoru d)AUC e)Metriklerin ortalama değerleri

Doğrusal çekirdek kullanılmasıyla en yüksek metrik değerleri A odası için hesaplanmıştır. A odasının Kesinlik, Duyarlılık, AUC ve F-skoru değerlerinin 0,99 olarak hesaplandığı görülmektedir. Tablo 2’de Doğruluk Kappa ve RMS değerleri verilmiştir.

Tablo 2. Çapraz doğrulama sonrası elde edilen performans metrikleri

Algoritma	Kappa	Doğruluk (%)	Rms
Doğrusal	0,97	97,9	0,10
Polinomal	0,95	96,6	0,12
Radyal	0,72	79,05	0,32
Sigmoid	0	25	0,61

Çapraz doğrulama öncesi ortalama performans metrik değerlerine bakıldığında Doğrusal ve Polinomal fonksiyon kullanıldığında Kesinlik, Duyarlılık, F-skoru, AUC ve doğruluk değerlerinin 0,98 olduğu görülmüştür. Radyal fonksiyon kullanıldığında tüm metrik değerleri 1 değerini almıştır. En düşük metrik değerleri Sigmoid fonksiyonu kullanıldığında elde edilmiştir. Algoritmaların performans

değerlendirmesinde önemli bir yer tutan ve veri tabanında bulunmayan verinin sınıflandırılma başarısının test edildiği çapraz doğrulama yöntemi uygulandıktan sonra yapılan performans analizinde Radyal fonksiyonun performans metrik değerlerinde bir düşüş gözlenmiştir. Bunun sebebi bu fonksiyonun daha önce karşılaşmadığı veriyi sınıflandırmada başarısız olduğunu göstermektedir.

Sonuçlar

Bu çalışmada kullanımı oldukça yaygınlaşmış, kullanım yaşı anaokulu çocuklarına kadar inmiş olan akıllı telefonlar güvenlik amaçlı kullanılmıştır. Akıllı telefonlar tarafından kayıt edilen kablosuz sinyal gücü kullanılarak anaokulu öğrencilerinin konumları belirlenmiştir. Bunun için sınıflandırma problemlerinde oldukça başarılı olan ve sıklıkla tercih edilen DVM algoritması seçilmiştir. Algoritma kullanılırken farklı çekirdek fonksiyonlarının kullanımı ile sınıflandırma başarısındaki değişkenlik incelenmiştir. Doğrusal fonksiyon kullanımı ile %100 'e yakın bir başarı sağlanırken Sigmoid fonksiyon kullanıldığında kötü sonuçlar verdiği çeşitli performans metrikleri kullanılarak gösterilmiştir. Doğrusal fonksiyon kullanıldığında çapraz doğrulama sonrası yapılan sınıflandırmada Kappa değeri 0,97 doğruluk değeri 97,9'dur. AUC ve F-skoru değeri ise 0,98 olarak hesaplanmıştır. Ayrıca doğrusal fonksiyon kullanıldığında elde edilen hata değeri diğer fonksiyonlara oranla daha düşüktür. En kötü sınıflandırma ise Sigmoid fonksiyonu kullanıldığında gerçekleşmiştir. Bu fonksiyonun doğruluk oranı %25'tir. Sonuç olarak bina içlerinde, kapalı alanlarda Sinyal gücü ve DVM algoritması kullanıldığında doğrusal fonksiyon seçiminin daha başarılı sonuçlar vereceği anlaşılmıştır.

Çıkar Çatışması Beyanı

Makale yazarı herhangi bir çıkar çatışması olmadığını beyan eder.

Araştırmacıların Katkı Oranı Beyan Özeti

Yazar makaleye %100 oranında katkı sağlamış olduğunu beyan eder.

Kaynaklar

- Arlot S., Celisse A. A survey of cross-validation procedures for model selection. *Statistics Surveys* 2010; 4: 40-79.
- Bayıroğlu H., Ayan K. Android üzerinde web tabanlı çocuk takip sistemi. *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi* 2014; 18(2): 87-91.
- Bejuri W., Mohamad MM. Wireless LAN/FM radio-based robust mobile indoor positioning: an initial outcome. *Int J Softw Eng Appl* 2014; 8(2): 313-324.

- Caneren G., Doğan AH., Öcalan T. Akıllı mobil telefonlardan elde edilen çift frekanslı ham GNSS gözlemleri ile PPP konum belirleme performansının araştırılması. *Jeodezi ve Jeoinformasyon Dergisi* 2021; 8(2): 120-130.
- Chandra A., Jain S., Qadeer MA. Implementation of location awareness and sharing system based on GPS and GPRS using J2ME, PHP and MYSQL. Paper presented at the 2011 3rd International Conference on Computer Research and Development, 2011, sayfa no: 216-220.
- Chen RC., Huang SL. A new method for indoor location base on radio frequency identification. Paper presented at the WSEAS International Conference. *Proceedings. Mathematics and Computers in Science and Engineering* 2009; 618-627.
- Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960; 20(1): 37-46.
- Cortes C., Vapnik V. Support-vector networks. *Machine learning* 1995; 20(3): 273-297.
- Frank A. UCI machine learning repository 2010. <http://archive.ics.uci.edu/ml>.
- Kamath U., De Jong K., Shehu A. Effective automated feature construction and selection for classification of biological sequences. *PloS one*, 2014; 9(7): e99982.
- Karadaş H. Okul öncesi eğitim kurumlarında yaşanan güvenlik ve sağlık problemlerinin paydaş görüşlerine göre incelenmesi. *Anatolian Journal of Health Research* 2021; 2(1): 7-14
- Li X., Nsofor GC., Song L. A comparative analysis of predictive data mining techniques. *International Journal of Rapid Manufacturing* 2009; 1(2): 150-172.
- Liao YC., Jeng JT., Chuang CC., Chen JC. Systematic design for the global positional systems with application in intelligent google android phone. Paper presented at the *Proceedings 2011 International Conference on System Science and Engineering* 2011, sayfa no: 182-186.
- Orozco-Arias S., Isaza G., Guyot R., Tabares-Soto R. A systematic review of the application of machine learning in the detection and classification of transposable elements. *PeerJ*, 2019; 7, e8311.
- Öztürk B., Tayılga G. Anaokulu iç mekân oyun alanlarına yönelik kavramsal yaklaşımlar: iç mimarlık tasarım stüdyosu deneyimi. *Turkish Online Journal of Design Art and Communication* 2021; 11(4): 1226-1238.
- Rawal K., Ramaswamy R. Genome-wide analysis of mobile genetic element insertion sites. *Nucleic Acids Research*, 2011; 39(16): 6864-6878.
- Rida ME., Liu F., Jadi Y., Algawhari AA., Askourih A. Indoor location position based on bluetooth signal strength. Paper presented at the 2015 2nd International Conference on Information Science and Control Engineering.
- Rohra JG., Perumal B., Narayanan SJ., Thakur P., Bhatt RB. User localization in an indoor environment using fuzzy hybrid of particle swarm optimization & gravitational search algorithm with neural networks. Paper presented at the *Proceedings of Sixth International Conference on Soft Computing for Problem Solving* 2017; 769-773.

- Sabancı K., Yigit E., Ustun D., Toktas A., Aslan MF. Wifi based indoor localization: application and comparison of machine learning algorithms. Paper presented at the 2018 XXIIIrd International Seminar/Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory (DIPED), sayfa no: 246-251
- Schietgat L., Vens C., Cerri R., Fischer CN., Costa E., Ramon, J., Blockeel H. A machine learning based framework to identify and classify long terminal repeat retrotransposons. PLoS Computational Biology, 2018; 14(4): e1006097.
- Seco F., Jiménez, AR., Zampella F. Joint estimation of indoor position and orientation from RF signal strength measurements. Paper presented at the International Conference on Indoor Positioning and Indoor Navigation, 28-31 th October 2013.
- Shu X., Du Z., Chen R. Research on mobile location service design based on Android. Paper presented at the 2009 5th International Conference on Wireless Communications, Networking and Mobile Computing, 2009.
- Sıvacılar S. Sürdürülebilirlik ilkesinde dayalı okul öncesi eğitim yapısı tasarlanması ve fiziksel etkenlerin iş sağlığı güvenliği açısından incelenmesi. Kırklareli Üniversitesi 2021; 89.
- Yasir M., Ho SW., Vellambi BN. Indoor position tracking using multiple optical receivers. Journal of Lightwave Technology, 2015; 34(4): 1166-1176.
- Yıldız S., Bektaş F. İlkokul öğrencilerinin okul yolunda karşılaşılabileceği risklerin belirlenmesi; Fındıklı örneği. Artvin Çoruh Üniversitesi Uluslararası Sosyal Bilimler Dergisi 2021; 7(2): 82-104.