

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/366588770>

# Bölüm 4 MAÇ SONUÇLARININ SINIFLANDIRILMASINDA ETKİN ROL OYNAYAN DEĞİŞKENLERİN BELİRLENMESİ: ANADOLU EFES BASKETBOL TAKIMI ÖRNEĞİ

Chapter · December 2022

CITATIONS

0

READS

26

1 author:



Enes Filiz

Balıkesir University

15 PUBLICATIONS 62 CITATIONS

SEE PROFILE

## **BÖLÜM 4**

### **MAÇ SONUÇLARININ SINIFLANDIRILMASINDA ETKİN ROL OYNAYAN DEĞİŞKENLERİN BELİRLENMESİ: ANADOLU EFES BASKETBOL TAKIMI ÖRNEĞİ**

*Enes FİLİZ<sup>1</sup>*

<sup>1</sup> Dr. Öğr. Üyesi Enes FİLİZ, Balıkesir Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Balıkesir (ORCID: 0000-0002-8006-9467)

## 1. Giriş

Basketbol günden güne popülerliğini arttıran bir spor dalıdır. NBA (Amerikan Basketbol Ligi) öncülüğünde insanların bu spora olan bakış açısı değişmiş ve taraftar sayısı gittikçe artmaya başlamıştır. Avrupa’da ise özellikle Euroleague gibi bir organizasyon olması, Avrupa’daki taraftarların büyük eğlenceden uzak kalmamasını sağlamaktadır. Avrupa’nın önemli takımlarının ve basketbolcularının bu büyük platformda kendilerini göstermeleri onlara olan ilginin ve taraftar sayılarının artmasına neden olmaktadır. Basketbol maç sonuçlarına olan ilginin yanı sıra büyük boyutlu maç veri setlerinin ortaya çıktığı görülmektedir. Maçların kazananları hakkında mümkün olan en iyi sonuçları sağlamak amacıyla çeşitli sistemler geliştirmiştir. Basketbol ve diğer spor dallarında yapılan mücadeleler tamamen bağımsız olaylar olmamakla birlikte insan faktörü de olaya dahil edildiğinden, analiz sürecinin uygun seçimi çok önemlidir. Takım sporlarında maç verileri aracılığıyla hücum ve savunma stratejilerini belirlemek ve sınıflandırma yapmak amacıyla makine öğrenmesi algoritmalarının kullanımı, son zamanlarda literatürde ve global spor dünyasında önemli bir ilgi görmektedir.

Maç sonuçlarının sınıflandırılmasına etki eden değişkenlerin belirlenmesi problemine bir yaklaşım olarak, bu çalışma da gösterilecek olan makine öğrenmesi algoritması ve öznelik seçim algoritmalarının kullanılmasıdır. Makine öğrenmesi algoritmalarından olan lojistik regresyon algoritması sınıflandırmada pek çok çalışmada kullanılmıştır (Dreiseitl & Ohno-Machado, 2002; Schein & Ungar, 2007; Cheng & Hüllermeier, 2009; Filiz & Öz, 2019; Itoo & Singh, 2021). Bu bağlamda literatür incelendiğinde makine öğrenmesi algoritmalarının farklı spor alanlarında (Arabzad vd., 2014; Igiri, 2015; Berrar vd., 2019; Filiz, 2022) olduğu gibi basketbol alanında da önemli çalışmalara konu olduğu belirlenmiştir. Jain & Kaur (2017), basketbol maç sonuçlarını tahmin etmek için destek vektör makineleri ve hibrit bulanık destek vektör makinelerini kullanmıştır. Çene (2018), Euroleague maçlarında oyunun sonucuna etki eden faktörleri belirlemeyi amaçlamıştır. Cai, vd. (2019), basketbol takımlarının son durumunu ortaya çıkararak maç sonuçları tahmini için karma bir topluluk öğrenme yöntemi ortaya koymuşlardır. Tian, vd. (2019), basketbolcuların farklı savunma stratejilerini

sınıflandırmak için makine öğrenmesini kullanmışlardır. Sarlis, vd. (2021), NBA takımlarının verilerini kullanarak sakatlıklara ilişkin etkin özellikleri belirlemek ve bunların oyuncu ve takım performansı üzerindeki etkisini belirlemek için makine öğrenmesi algoritmalarından yararlanmışır. Ulaş (2021), çalışmasında NBA takımlarının değerlerini etkileyen faktörler üzerine araştırma yapmıştır. NBA takımlarının özellikleri ve benzerliklerini belirlemek için makine öğrenmesi tekniklerinden yararlanmışır. Mahmood, vd.(2021), basketbolun yükselen yıldızlarını tahmin etmek için makine öğrenmesinden yararlanmışlardır.

Literatürde basketbol maçının sonuçlarına etki eden faktörleri belirlemek için öznelik seçim algoritmaları kullanıldığı belirlenmiştir. Basketbol veya farklı maçların sonuçlarını tahmin etmek ve maç için stratejiler belirlemede farklı makine öğrenmesi algoritmaları kullanılmış ve maç için önemli değişkenleri belirlemek için öznelik seçiminden yararlanmışır (Miljković, vd.. 2010; Trawiński, 2010; Lin, vd. 2014; Pai vd. 2017; Leicht, vd., 2017; Giasemidis, 2020; Chen vd., 2021). Ji, (2020), basketbolcuların spor duruşunu geliřtirmek amacıyla öznelik seçimi yöntemi kullanmışır. Nataliani (2021), basketbolcuların pozisyonunu belirlemek için boy, kilo, yaş ve vücut kitle indeksinden oluşan ilgili fiziksel özellik ya da özellikleri seçmek için öznelik seçiminden yararlanmışır.

Literatür çerçevesindeki incelemeler doğrultusunda çalışma, maç sonuçlarının sınıflandırmasında etkin rol oynayan değişkenlere odaklanacaktır. Bu bağlamda Türkiye Basketbol Ligi'nin en başarılı takımlarından biri olan Anadolu Efes Basketbol Takımının maç sonuçları incelenecektir.

Çalışmanın amacı Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılmasında etkin rol oynayan değişkenleri belirlemektir. Ayrıca farklı öznelik seçim algoritmalarından yararlanılacaktır. Farklı öznelik seçim algoritmaları ile belirlenen etkin değişkenler yardımıyla en başarılı sınıflandırma sonuçlarını veren öznelik seçim algoritması/algoritmaları ortaya çıkarılacaktır.

## 2. Materyal ve metod

### 2.1. Veri seti

Çalışmanın amacı doğrultusunda Anadolu Efes Basketbol Takımı'nın 2015/2016 - 2020/2021 sezonları arasındaki toplam 422 maçı (Türkiye ligi, Türkiye ligi playoff, Türkiye kupası, Euroleague ve Euroleague playoff) çalışmaya dahil edilmiştir. Maçkolik internet sayfasından alınan Anadolu Efes Basketbol Takımı'nın maç verileri doğrultusunda analizler yapılmıştır (Maçkolik internet sayfası). Çalışmada 32 bağımsız, 1 bağımlı değişken (maç sonucu) olmak üzere toplam 33 değişken kullanılmıştır. Bu değişkenler Tablo 1'de verilmiştir.

Çalışmada Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılmasında etkin rol oynayan değişkenlerin belirlenmesi için Info gain öznitelik seçim algoritması (IG), Correlation attribute öznitelik seçim algoritması (CA), One-R öznitelik seçim algoritması, ReliefF öznitelik seçim algoritması (RF) kullanılmıştır. Ayrıca belirtilen öznitelik seçim algoritmaları arasındaki sınıflandırma sonuçlarını belirlemek ve karşılaştırma yapabilmek için lojistik regresyon algoritmasından (LR) yararlanılmıştır.

**Tablo 1.** Çalışmada kullanılan değişkenler

Değişkenler	Tanımı
<i>Tür</i>	1:Türkiye ligi, 2: Lig Playoff, 3: Türkiye Kupası, 4: Euroleague, 5: Euroleague playoff
<i>Taraf</i>	1: Ev sahibi, 2: Deplasman
<i>Birinci periyod sonucu</i>	0: Beraberlik, 1: Ev sahibi, 2: Deplasman
<i>İkinci periyod sonucu</i>	0: Beraberlik, 1: Ev sahibi, 2: Deplasman
<i>Üçüncü periyod sonucu</i>	0: Beraberlik, 1: Ev sahibi, 2: Deplasman
<i>Dördüncü periyod sonucu</i>	0: Beraberlik, 1: Ev sahibi, 2: Deplasman
<i>Devre arası fark</i>	
<i>İlgili takımın iki sayı isabet oranı</i>	

---

*İlgili takımın üç sayı isabet oranı*

---

*İlgili takımın serbest atıř isabet oranı*

---

*Rakip takımın iki sayı isabet oranı*

---

*Rakip takımın üç sayı isabet oranı*

---

*Rakip takımın serbest atıř isabet oranı*

---

*İlgili takımın savunma ribaundu sayısı*

---

*İlgili takımın hücum ribaundu sayısı*

---

*İlgili takımın asist sayısı*

---

*İlgili takımın top kaybı sayısı*

---

*İlgili takımın top çalma sayısı*

---

*İlgili takımın blok sayısı*

---

*İlgili takımın faul sayısı*

---

*Rakip takımın savunma ribaundu sayısı*

---

*Rakip takımın hücum ribaundu sayısı*

---

*Rakip takımın asist sayısı*

---

*Rakip takımın top kaybı sayısı*

---

*Rakip takımın top çalma sayısı*

---

*Rakip takımın blok sayısı*

---

*Rakip takımın faul sayısı*

---

*İlgili takımda çift haneli sayı atan oyuncu sayısı*

---

*İlgili takımın rotasyonda kullandıęı oyuncu sayısı*

---

*İlgili takımda sayı üreten oyuncu sayısı*

---

*İlgili takımda double double yapan oyuncu sayısı*

---

---

*İlgili takımda 20 dk. Üzerinde oynayan oyuncu sayısı*

---

*Maç sonucu*

0: Yenilgi, 1: Galibiyet

---

## 2.2. LR algoritması

Veri setindeki bağımlı ve bağımsız değişkenler arasındaki ilişkiyi incelemek için kullanılan bir analizdir. Temel amacı, değişkenler arasındaki ilişkiyi ortaya çıkarmaktır. Bağımlı değişkenin ikili olduğu durumlarda kullanıldığı gibi; çoklu bağımlı değişken olduğu durumlarda da uygulanabilmektedir. Sınıflandırma, kümeleme veya modelleme gibi analizlerde kullanımı bulunmaktadır. Lojistik regresyon fonksiyonu;

$$\pi(x) = [1 + \exp(-\beta_0 - \beta_1 X)]^{-1}$$

şeklinde gösterilebilir (Agresti, 1996)

## 2.3. Öznitelik Seçim Algoritmaları

Makine öğrenmesi algoritmalarının uygulanmasında önemli noktalardan birisi de sınıflandırmada etkin rol oynayan değişkenlerin belirlenmesidir. Sınıflandırmada etkin rol oynayan değişkenlerin belirlenmesi hem zaman kazanımını hem de işlem kolaylığını sağlar. Etkin rol oynayan değişkenlerin belirlenmesiyle birlikte sınıflandırma da belirgin bir düşüş olmaması hatta sınıflandırma sonuçlarına yakın sonuçlar vermesi beklenir. Etkin rol oynayan değişkenlerin belirlenmesi için öznitelik seçim algoritmalarından yararlanılır. IG öznitelik seçim algoritması, CA öznitelik seçim algoritması, One-R öznitelik seçim algoritması, RF öznitelik seçim algoritması kullanılmıştır.

### 2.3.1. IG öznitelik seçim algoritması

IG öznitelik seçim algoritması entropi kavramını temel alarak işlem yapmaktadır. Sistemdeki düzensizliği belirten entropi, ne kadar yüksek olursa o veri setinde daha fazla bilgiden söz edilebilir. Öznitelik seçimi yapmak için her bir değişken için entropi hesaplaması yapmaktadır. Entropi için belirlenen eşik değerinden düşük olan değişkenler sınıflandırma sürecine dahil edilmemesi gerekmektedir. IG öznitelik seçim algoritmasının formülü aşağıdaki gibidir;

$$\begin{aligned}
 IG(t) = & - \sum_{i=1}^m P(C_i) \log P(C_i) \\
 & + P(t) \sum_{i=1}^m P(C_i | t) \log P(C_i | t) \\
 & + P(\bar{t}) \sum_{i=1}^m P(C_i | \bar{t}) \log P(C_i | \bar{t})
 \end{aligned}$$

m toplam sınıf sayısını;  $P(C_i)$   $C_i$  sınıfının olasılığını;  $P(t)$  ve  $P(\bar{t})$ ,  $t$ 'nin kümede olup olmamasının olasılığını;  $P(C_i | t)$  ve  $P(C_i | \bar{t})$  ise  $t$ 'nin kümeye ait olduğu ya da olmadığı zamanda  $C_i$ 'nin koşullu olasılığını ifade etmektedir. Bu tanımlama, ikili sınıflandırma modellerinde kullanılanlardan durumdan daha genel bir tanımdır (Lewis & Ringuette, 1994).

### 2.3.2. CA öznitelik seçim algoritması

CA öznitelik seçim algoritması Pearson korelasyon katsayısı ile ikili değişkenler arasında ilişkiyi inceleyerek etkin öznitelikleri belirlemektedir. Her bir değişken için ayrı ayrı işlem yapmaktadır. Eğer değişkenler nominal ölçek ile hesaplanırsa, genel korelasyon değeri için ağırlıklı ortalama değeri belirlenerek sonuçlar ortaya çıkar (Gnanambal, et al., 2018).

### 2.3.3. One-R öznitelik seçim algoritması

One-R öznitelik seçim algoritması en düşük hata oranına sahip özniteliği belirlemeye odaklanmaktadır. Eğitim veri setindeki değişkenleri hata oranına göre sıralama yapmaktadır. Bunu yapmak için tek düzeyli bir karar ağacı oluşturur. Sayısal değer bulunan değişkenler üzerinde çalışmaktadır. Yüksek doğrulukla başarılı sonuçlar veren, basit bir uygulamadır (Kabakchieva, 2013).

### 2.3.4. RF öznitelik seçim algoritması

RF öznitelik seçim algoritması değişkenler arasındaki ilişkileri incelemektedir. Veri setinden seçilen bir örneğin, kendi sınıfındaki örnekler ile olan yakınlığı ve diğer sınıftaki örneklerle olan uzaklığı ile

model oluşturmaktadır. Öncelikle ikili sınıflandırma problemlerinde kullanılmış ardından çoklu sınıflandırma problemlerinde de kullanılmaya başlanmıştır. RF öznelik seçim algoritmasının formülü aşağıda verilmiştir;

$$diff(A, I_1, I_2) = \begin{cases} 0 & value(A, I_1) = value(A, I_2) \\ 1 & diğer durumlarda \end{cases}$$

$$diff(A, I_1, I_2) = \frac{|value(A, I_1) - value(A, I_2)|}{\max(A) - \min(A)}$$

$I_1$  ve  $I_2$  örneği için A özneliğinin değerleri arasındaki farkı hesaplamaktadır. A'nın maksimum ve minimum değerleri, tüm örnek kümesi üzerinden belirlenir. Bu normalleştirme, hem ayrık hem de sürekli özellikler için ağırlık güncellemelerinin 0 ile 1 arasında olmasını sağlar (Robnik-Šikonja & Kononenko 2001)

#### 2.4. Sınıflandırma Başarısı Performans Ölçütleri

Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılmasında etkin değişkenlerin belirlenmesi ve bu etkin değişkenler yardımıyla sınıflandırma başarılarını incelemek amacıyla farklı performans ölçütleri kullanılması gerekmektedir. Öncelikli olarak Dp oranı ile sonuçlar incelenmiş ardından Kappa istatistiği ( $\kappa$ ), ortalama mutlak hata (MAE: Mean Absolute Error) ve alıcı işlem karakteristiği eğrisi (ROC: Receiver Operating Characteristic) değerlerinden yararlanılmıştır. Bazı değerlerin hesaplanması için karşılaştırma matrisine ihtiyaç duyulmaktadır. Karşılaştırma matrisi tablo 2' de verilmiştir:

**Tablo 2.** Karşılaştırma matrisi

		Tahmin	
		a	b
Gerçek Değer	a	DP (Doğru Pozitif)	YN (Yanlış Negatif)
	b	YP (Yanlış Pozitif)	DN (Doğru Negatif)

*Dp oranı*, doğru sınıflandırılmış pozitif örneklerin, tüm pozitif örneklere oranı ile hesaplanmaktadır.

$$Dp\ oranı = \frac{DP}{DP + YN}$$

$\kappa$  istatistiği, sınıflandırma algoritmasının performansını gösteren önemli ölçütlerden biridir. 0 ile 1 arasında değer alan  $\kappa$  istatistiği, 1'e ne kadar yakın olursa o sınıflandırma algoritmasının o kadar başarılı sınıflandırma yaptığını göstermektedir. *Mae*, hata değerini ortaya koyan bir göstergedir. Tahmin değerleri ile gerçek değerler arasındaki mutlak farkın ortalamasını hesaplamaktadır. *Roc alanı*, algoritmaların sınıflandırma performans ölçümünde başarılı sonuçlar veren bir eğri olarak tanımlanmaktadır. Eğrinin altında kalan alan sınıflandırma başarısını göstermektedir. 1'e ne kadar yakın değerler elde edilirse sınıflandırma algoritmasının da o denli başarılı olduğunu göstermektedir. Ortaya çıkarılan eğri sayesinde sınıflandırma algoritmaları görsel olarak kıyaslama imkânına erişmektedir (Bradley, 1997).

### 3. Deneysel Kurulum

Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılmasında etkin rol oynayan değişkenleri belirlemek amacıyla

uygulama kurgulanmıştır. Çalışmanın uygulama prensibi aşağıda verilmiştir.

- Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılması tüm değişkenler kullanılarak LR ile yapılmıştır.

- Farklı öznitelik seçim algoritmaları ile her bir öznitelik seçim algoritması için etkin öznitelikler ve önem düzeyleri belirlenmiştir.

-Her bir öznitelik seçim algoritmasının belirlediği etkin değişkenler yardımıyla Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırma başarısı ortaya çıkarılmıştır. Ayrıca öznitelik seçimindeki benzerlik ve farklılıklar incelenmiştir.

Çalışmada makine öğrenmesi algoritmalarında eğitim ve test veri seti seçimi için k-katlı çapraz doğrulama işlemi uygulanmış ve k=10 seçilmiştir. Analizler Java kaynak kodlu sınıflandırma ve kümeleme gibi alanlarda kullanılan weka programında yapılmıştır.

#### 4. Bulgular

Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılmasında etkin rol oynayan değişkenlerin belirlenmesinde farklı öznitelik seçim algoritmalarından yararlanılmıştır. Bu öznitelik seçim algoritmaları arasında karşılaştırma yapabilmek amacıyla LR analizi kullanılmıştır. Tüm değişkenler kullanılarak elde edilen sınıflandırma başarısı Tablo 3' de; doğru sınıflandırılan maç sayıları Tablo 4' te verilmiştir. Şekil 1' de ise Roc alanı grafiği gösterilmiştir.

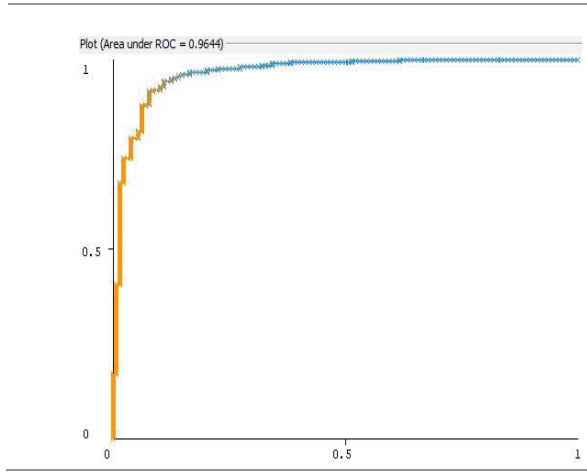
**Tablo 3.** Tüm değişkenler kullanılarak elde edilen sınıflandırma sonucu

	<i>Dp Oranı</i>	$\kappa$	<i>Mae</i>	<i>Roc Değeri</i>
<i>LR</i>	%92,7	0,826	0,096	0,964

**Tablo 4.** Tüm değişkenler kullanılarak elde edilen başarılı sınıflandırma sayıları

	0	1
0	112	16
1	15	279

\*0:yenilgi, 1:galibiyet

**Şekil 1.** Tüm değişkenler kullanılarak elde edilen Roc alanı

Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılması için LR algoritması kullanılmıştır. Tablo 3' e göre Dp oranı %92,7 ile doğru sınıflandırma başarısı elde etmiştir.  $\kappa$  istatistiği (0.826), Mae (0.096) ve Roc değeri (0.964) bu sonucu desteklemektedir. Tablo 4' e göre Anadolu Efes Basketbol Takımı'nın incelenen 422 maçının 391(112 yenilgi, 279 galibiyet) tanesini doğru sınıflandırdığı belirlenmiştir.

Farklı öznelik seçim algoritmaları ile Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılmasında etkin rol oynayan değişkenler ve önem düzeyleri tablo 5' te; etkin değişkenler yardımıyla elde edilen sınıflandırma sonuçları tablo 6' da; etkin değişkenler yardımıyla elde edilen başarılı sınıflandırılmış maç sayıları tablo 7' de

verilmiştir. Şekil 2’de ise farklı öznitelik seçim algoritmaları ile elde edilen Roc alanı grafikleri gösterilmiştir.

**Tablo 5.** Farklı öznitelik seçim yöntemlerinin belirlediği etkin değişkenler

<i>IG</i>		<i>CA</i>	
<i>Önem Düzeyi</i>	<i>Değişkenler</i>	<i>Önem Düzeyi</i>	<i>Değişkenler</i>
0.2410	<i>Devre arası fark</i>	0.5110	<i>Devre arası fark</i>
0.1498	<i>İlgili takımın asist sayısı</i>	0.4273	<i>İlgili takımın asist sayısı</i>
0.1243	<i>Rakip takımın savunma ribaundu sayısı</i>	0.4062	<i>İlgili takımın iki sayı isabet oranı</i>
0.1134	<i>İlgili takımın iki sayı isabet oranı</i>	0.3752	<i>Rakip takımın savunma ribaundu sayısı</i>
0.0841	<i>İlgili takımın üç sayı isabet oranı</i>	0.3237	<i>İlgili takımın üç sayı isabet oranı</i>
0.0611	<i>İlgili takımın savunma ribaundu sayısı</i>	0.3234	<i>İlgili takımın savunma ribaundu sayısı</i>
0.0528	<i>Rakip takımın üç sayı isabet oranı</i>	0.3097	<i>İlgili takımında çift haneli sayı atan oyuncu sayısı</i>
0.0505	<i>Rakip takımın iki sayı isabet oranı</i>	0.2833	<i>Rakip takımın iki sayı isabet oranı</i>
0.0493	<i>İlgili takımında çift haneli sayı atan oyuncu sayısı</i>	0.2749	<i>Rakip takımın üç sayı isabet oranı</i>
0.0477	<i>Tür</i>	0.2499	<i>Tür</i>
0.0401	<i>İlgili takımın faul sayısı</i>	0.2447	<i>İlgili takımın faul sayısı</i>
0.0301	<i>Rakip takımın faul sayısı</i>	0.1857	<i>İlgili takımında sayı üreten oyuncu sayısı</i>
0.0287	<i>Rakip takımın serbest atış isabet oranı</i>	0.1713	<i>Rakip takımın serbest atış isabet oranı</i>
0.0272	<i>İlgili takımında sayı üreten oyuncu sayısı</i>	0.1609	<i>Taraf</i>

0.0255	Rakip takımın top kaybı sayısı	0.1589	İlgili takımın top çalma sayısı
<b>One-R</b>		<b>RF</b>	
<b>Önem Düzeyi</b>	<b>Deęişkenler</b>	<b>Önem Düzeyi</b>	<b>Deęişkenler</b>
77.48815	Devre arası fark	0.10024	Taraf
72.74882	Rakip takımın savunma ribaundu sayısı	0.05604	Üçüncü periyod sonucu
70.85308	İlgili takımda çift haneli sayı atan oyuncu sayısı	0.05052	Devre arası fark
70.61611	Rakip takımın üç sayı isabet oranı	0.05024	Dördüncü periyod sonucu
70.61611	İlgili takımın iki sayı isabet oranı	0.04348	Tür
70.61611	Rakip takımın iki sayı isabet oranı	0.04292	Rakip takımın savunma ribaundu sayısı
70.14218	İlgili takımın top çalma sayısı	0.03791	İkinci periyod sonucu
69.90521	Rakip takımın asist sayısı	0.03700	İlgili takımın asist sayısı
69.66825	İlgili takımda 20 dk. Üzerinde oynayan oyuncu sayısı	0.03649	Birinci periyod sonucu
69.66825	Dördüncü periyod sonucu	0.03378	İlgili takımın üç sayı isabet oranı
69.66825	Taraf	0.03159	İlgili takımın iki sayı isabet oranı
69.66825	Birinci periyod sonucu	0.02226	İlgili takımın faul sayısı
		0.02157	Rakip takımın iki sayı isabet oranı

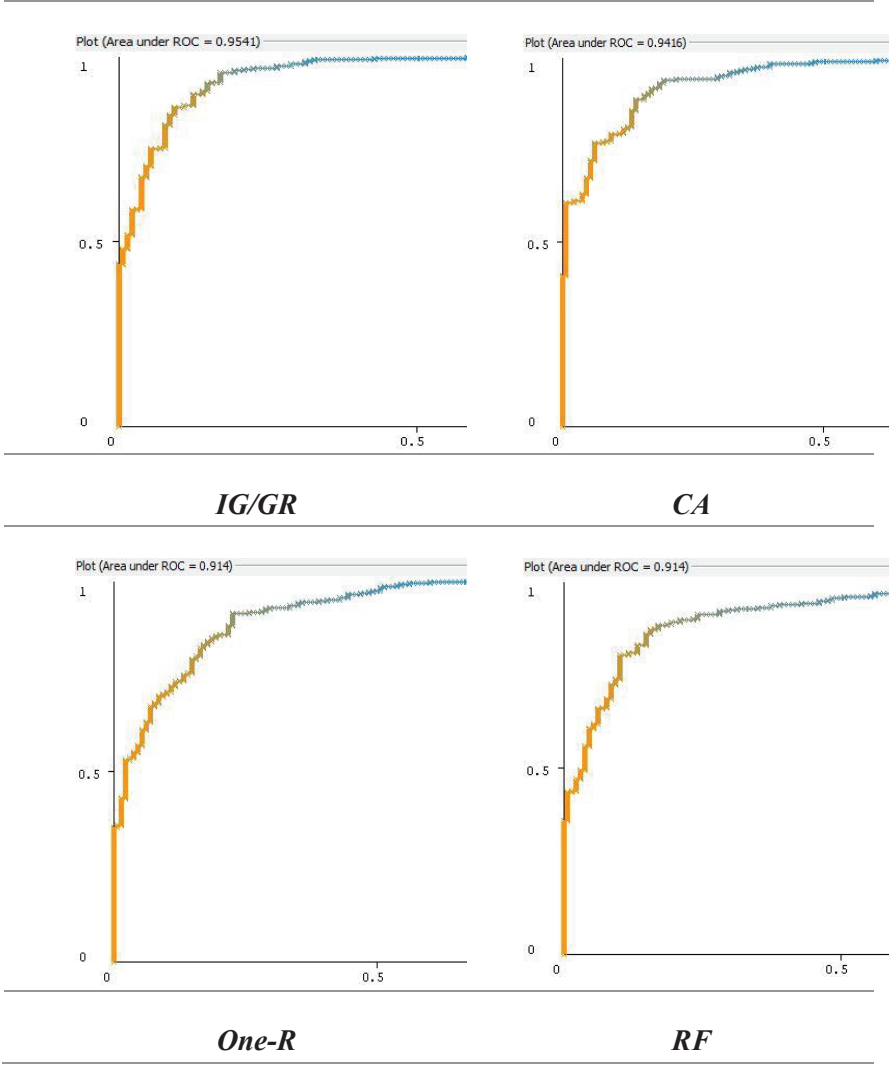
**Tablo 6.** Etkin deęişkenler kullanılarak elde edilen sınıflandırma sonuçları

	<i>IG</i>	<i>CA</i>	<i>One-R</i>	<i>RF</i>
	(15 deęişken)	(15 deęişken)	(12 deęişken)	(13 deęişken)
<i>Dp Oranı</i>	<b>%91,0</b>	%89,6	%86,3	%86,3
<i>κ</i>	0,7841	0,7511	0,6675	0,6719
<i>Mae</i>	0,1322	0,1565	0,1977	0,1967
<i>Roc Deęeri</i>	0,9541	0,9416	0,9140	0,9140

**Tablo 7.** Etkin deęişkenler kullanılarak elde edilen başarılı sınıflandırma sayıları

	<i>IG</i>		<i>CA</i>		<i>One-R</i>		<i>RF</i>	
	<i>0</i>	<i>1</i>	<i>0</i>	<i>1</i>	<i>0</i>	<i>1</i>	<i>0</i>	<i>1</i>
<i>0</i>	106	22	104	24	94	34	97	31
<i>1</i>	16	278	20	274	24	270	27	267

\*0:yenilgi, 1:galibiyet



**Őekil 2.** Farklı znitelik seim algoritmaları ile elde edilen sınıflandırma başarılarının Roc alanları

Tablo 5’ te Anadolu Efes Basketbol Takımı’nın ma sonuçlarının sınıflandırılmasında etkin rol oynayan deęişkenler ve nem dzeyleri gsterilmiřtir. Kullanılan znitelik seim algoritmalarının farklı sonuçlar verdięi gzlemlenmiřtir. Drt farklı znitelik seim algoritmalarında ortak olarak belirlenen etkin znitelikler ‘Devre arası fark’, ‘İlgili takımın iki sayı isabet oranı’, ‘Rakip takımın iki sayı isabet oranı’, ‘Rakip takımın savunma ribaundu sayısı’ olarak belirlenmiřtir. Ü znitelik seim

algoritmasının ortak olarak belirlediği fakat One-R öznitelik seçim algoritması tarafından etkin olarak belirlenmeyen öznitelikler ise *'Tür'*, *'İlgili takımın üç sayı isabet oranı'*, *'İlgili takımın asist sayısı'*, *'İlgili takımın faul sayısı'* olarak ortaya konmuştur. Yine üç öznitelik seçim algoritması tarafından etkin bulunup RF tarafından etkin bulunmayan değişkenler *'Rakip takımın üç sayı isabet oranı'*, *'İlgili takımda çift haneli sayı atan oyuncu sayısı'* olarak belirlenmiştir. Genel olarak bakıldığında önem düzeyi olarak *'Devre arası fark'* değişkeninin maç sonuçlarının sınıflandırılmasında en etkin rol oynadığı söylenebilir.

Tablo 6' da dört farklı öznitelik seçim yöntemiyle elde edilen etkin değişkenler kullanılarak elde edilen sınıflandırma başarı sonuçları verilmiştir. IG öznitelik seçim algoritması ile belirlenen 15 etkin değişken ile yapılan sınıflandırmada Dp oranı %91.0 ile en başarılı sınıflandırma olarak ortaya çıkmıştır.  $\kappa$  istatistiği (0,7841), Mae (0,1322), Roc değeri (0,9541) bu sonucu desteklediği görülmüştür. CA öznitelik seçim algoritmasının belirlediği 15 etkin değişken ile yapılan sınıflandırmanın başarı oranı ise %89.6; One-R öznitelik seçim algoritmasının belirlediği 12 etkin değişken ile yapılan sınıflandırma başarı oranı %86.3; RF öznitelik seçim algoritmasının belirlediği 13 etkin değişken ile yapılan sınıflandırma başarı oranı %86.3 olarak belirlenmiştir.

Tablo 7' de IG öznitelik seçim algoritması ile belirlenen etkin değişkenler kullanılarak Anadolu Efes Basketbol Takımı'nın incelenen 422 maçının 384 (106 yenilgi, 278 galibiyet) tanesini doğru sınıflandırdığı belirlenmiştir. CA öznitelik seçim algoritması ile belirlenen etkin değişkenler ile 422 maçının 378 (104 yenilgi, 274 galibiyet) tanesini; One-R öznitelik seçim algoritması ile belirlenen etkin değişkenler ile 422 maçının 364 (94 yenilgi, 270 galibiyet) tanesini; RF öznitelik seçim algoritması ile belirlenen etkin değişkenler ile 422 maçının 364 (97 yenilgi, 267 galibiyet) tanesini doğru sınıflandırdığı ortaya konmuştur.

## 5. Değerlendirme ve Tartışma

Çalışmanın amacı doğrultusunda, Türkiye Basketbol Ligi'nin en başarılı takımlarından biri olan Anadolu Efes Basketbol Takımının maç sonuçları incelenmiştir. Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılmasında etkin rol oynayan değişkenler öznitelik seçim algoritmaları ile belirlenmiştir. Farklı öznitelik seçim algoritmaları

ile belirlenen etkin değişkenler yardımıyla en başarılı sınıflandırma sonuçlarını veren öznitelik seçim algoritması/algoritmaları ortaya çıkarılmıştır. Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılmasında tüm değişkenler kullanılarak yapılan analiz sonucunda %92.7'lik bir başarı elde edilmiştir. Bunun anlamı çalışmadaki bağımsız değişkenler elde olduğunda Anadolu Efes Basketbol Takımı'nın maç sonuçlarının sınıflandırılması ~%93 oranında doğru yapılabileceği ortaya konmuştur.

Çalışmanın odak noktası olan etkin değişkenlerin belirlenmesinde farklı öznitelik seçim algoritmaları kullanılmış ve farklı sonuçlar elde edilmiştir. Kullanılan öznitelik seçim algoritmalarının ortak olarak belirlediği değişkenler '*Devre arası fark*', '*İlgili takımın iki sayı isabet oranı*', '*Rakip takımın iki sayı isabet oranı*', '*Rakip takımın savunma ribaundu sayısı*' dir. Literatürde benzer şekilde, isabet oranı, ribaunt gibi değişkenlerin maç sonucunda etkili olduğu incelenmiştir (Leicht, vd., 2017; Jain & Kaur, 2017; Çene, 2018; Chen vd., 2021). Ayrıca takımın oynadığı lig (tür), asist, faul sayıları, rakip takımın isabet oranı ve rakip takımın çift haneli sayı üreten oyuncu sayısı gibi değişkenler de etkin rol oynayan değişkenler olarak belirlenmiştir. Etkin değişkenlerin yanı sıra '*İlgili takımın serbest atış isabet oranı*', '*İlgili takımın ve rakip takımın hücum ribaunt sayıları*', '*İlgili takımın ve rakip takımın blok sayıları*', '*İlgili takımın top kaybı sayısı*', '*İlgili takımın rotasyonda kullandığı oyuncu sayısı*' ve '*İlgili takımda double double yapan oyuncu sayısı*' değişkenleri ise hiçbir öznitelik seçim algoritması tarafından etkin değişken olarak bulunmamıştır. Anadolu Efes Basketbol Takımı özelinde bu etkin ve etkin olmayan değişkenler maç sonuçları için teknik ekibe ve oyun yöneticilerine yol gösterici olabileceği düşünülmektedir.

Çalışmadaki bir diğer önemli nokta etkin değişkenlerin kullanımının ne derece faydalı olduğudur. Bu doğrultuda öznitelik seçim algoritmaları ile elde edilen değişkenler yardımıyla yapılan sınıflandırmada ~%86-91 oranında başarı elde ettiği görülmüştür. Sonuç olarak daha az değişken ile (15 değişken) sınıflandırma da çok fazla düşünüş yaşamadan başarılı sınıflandırma oranına ulaşılabileceği ortaya konmuştur. Öznitelik seçim algoritmaları arasından da IG öznitelik seçim algoritması ile elde edilen etkin değişkenlerin en başarılı sınıflandırma oranını verdiği belirlenmiştir. Bu çalışma özelinden yola çıkarak benzer çalışmalarda IG öznitelik seçim algoritmasının kullanılabilirliği önerilebilir.

Çalışmada bazı kısıtlamalar mevcuttur. Yalnızca Anadolu Efes Basketbol Takımı'nın 2015-2016/2020-2021 arasında oynadığı 422 maç (Türkiye ligi, Türkiye ligi playoff, Türkiye kupası, Euroleague ve Euroleague playoff) incelenmiştir. Çalışmada makine öğrenmesi algoritması olarak yalnızca LR algoritması; öznitelik seçim algoritması olarak IG, CA, One-R, RF öznitelik seçim algoritması kullanılmıştır. Maç sonuçlarına etki eden değişkenler sadece Anadolu Efes Basketbol Takımı için belirlenmiştir.

## Kaynaklar

- Agresti, A. (1996). *An Introduction to Categorical Data Analysis*, John Wiley and Sons. Inc.
- Berrar, D., Lopes, P., & Dubitzky, W. (2019). Incorporating domain knowledge in machine learning for soccer outcome prediction. *Machine Learning*, 108(1), 97–126. <https://doi.org/10.1007/s10994-018-5747-8>
- Bradley, A. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7), 1145–1159. [http://dx.doi.org/10.1016/S0031-3203\(96\)00142-2](http://dx.doi.org/10.1016/S0031-3203(96)00142-2).
- Cai, W., Yu, D., Wu, Z., Du, X., & Zhou, T. (2019). A hybrid ensemble learning framework for basketball outcomes prediction. *Physica A: Statistical Mechanics and Its Applications*, 528(16), Article 121461. <https://doi.org/10.1016/j.physa.2019.121461>
- Chen, W. J., Jhou, M. J., Lee, T. S., & Lu, C. J. (2021). Hybrid basketball game outcome prediction model by integrating data mining methods for the national basketball association. *Entropy*, 23(4), 477.
- Cheng, W., & Hüllermeier, E. (2009). Combining instance-based learning and logistic regression for multilabel classification. *Machine Learning*, 76(2), 211-225.
- Çene, E. (2018). What is the difference between a winning and a losing team: insights from Euroleague basketball. *International Journal of Performance Analysis in Sport*, 18(1), 55-68.
- Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, 35(5-6), 352-359.
- Filiz, E. (2022). Evaluation of Match Results of Five Successful Football Clubs With Ensemble Learning Algorithms. *Research Quarterly for Exercise and Sport*, 1-10. <https://doi.org/10.1080/02701367.2022.2053647>
- Filiz, E., & Öz, E. (2019). Finding the Best Algorithms and Effective Factors in Classification of Turkish Science Student Success. *Journal of Baltic Science Education*, 18(2), 239-253.

- Giasemidis, G. (2020). Descriptive and Predictive Analysis of Euroleague Basketball Games and the Wisdom of Basketball Crowds. arXiv preprint arXiv:2002.08465.
- Gnanambal, S., Thangaraj, M., Meenatchi, V. T., & Gayathri, V. (2018). Classification algorithms with attribute selection: an evaluation study using WEKA. *International Journal of Advanced Networking and Applications*, 9(6), 3640-3644.
- Igiri, C. P. (2015). Support vector machine–based prediction system for a football match result. *IOSR Journal of Computer Engineering (IOSRJCE)*, 17(3), 21–26. <https://doi.org/10.9790/0661-17332126>
- Ito, F., & Singh, S. (2021). Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. *International Journal of Information Technology*, 13(4), 1503-1511.
- Jain, S., & Kaur, H. (2017, September). Machine learning approaches to predict basketball game outcome. In *2017 3rd International Conference on Advances in Computing, Communication & Automation (ICACCA)(Fall)* (pp. 1-7). IEEE.
- Ji, R. (2020). Research on basketball shooting action based on image feature extraction and machine learning. *IEEE Access*, 8, 138743-138751.
- Kabakchieva, D. (2013). Predicting student performance by using data mining methods for classification. *Cybernetics and information technologies*, 13(1), 61-72.
- Leicht, A. S., Gómez, M. A., & Woods, C. T. (2017). Explaining match outcome during the men’s basketball tournament at the Olympic Games. *Journal of sports science & medicine*, 16(4), 468.
- Lewis, D. D., & Ringuette, M. (1994, April). A comparison of two learning algorithms for text categorization. In *Third annual symposium on document analysis and information retrieval* (Vol. 33, pp. 81-93).
- Lin, J., Short, L., & Sundaresan, V. (2014). Predicting national basketball association winners. *CS 229 FINAL PROJECT*, 1-5.
- Mackolik Internet Sayfası, [www.mackolik.com](http://www.mackolik.com), Erişim Tarihi: 15.04.2022
- Mahmood, Z., Daud, A., & Abbasi, R. A. (2021). Using machine learning techniques for rising star prediction in basketball. *Knowledge-Based Systems*, 211, 106506.

- Miljković, D., Gajić, L., Kovačević, A., & Konjović, Z. (2010, September). The use of data mining for basketball matches outcomes prediction. In *IEEE 8th international symposium on intelligent systems and informatics* (pp. 309-312). IEEE.
- Nataliani, Y. (2021). Feature-reduction Fuzzy c-means Clustering for Basketball Players Positioning. *JOIV: International Journal on Informatics Visualization*, 5(4), 415-421.
- Pai, P. F., ChangLiao, L. H., & Lin, K. P. (2017). Analyzing basketball games by a support vector machines with decision tree model. *Neural Computing and Applications*, 28(12), 4159-4167.
- Robnik-Šikonja, M., & Kononenko, I. (2001). Comprehensible interpretation of relief's estimates. In *Machine Learning: Proceedings of the Eighteenth International Conference on Machine Learning (ICML2001)*, Williamstown, MA, USA. San Francisco: Morgan Kaufmann (pp. 433-40).
- Sarlis, V., Chatziilias, V., Tjortjis, C., & Mandalidis, D. (2021). A data science approach analysing the impact of injuries on basketball player and team performance. *Information Systems*, 99, 101750.
- Schein, A. I., & Ungar, L. H. (2007). Active learning for logistic regression: an evaluation. *Machine Learning*, 68(3), 235-265.
- Trawiński, K. (2010, July). A fuzzy classification system for prediction of the results of the basketball games. In *International conference on fuzzy systems* (pp. 1-7). IEEE.
- Tian, C., De Silva, V., Caine, M., & Swanson, S. (2019). Use of machine learning to automate the identification of basketball strategies using whole team player tracking data. *Applied Sciences*, 10(1), 24.
- Ulas, E. (2021). Examination of National Basketball Association (NBA) team values based on dynamic linear mixed models. *Plos one*, 16(6), e0253179.