



## TÜRKİYE ERKEKLER BASKETBOL LİĞİ MAÇ SONUÇLARININ MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE TAHMİNİ

Engin ÖZDEMİR, Serkan BALLI\*

Muğla Sıtkı Koçman Üniversitesi, Teknoloji Fakültesi, Bilişim Sistemleri Mühendisliği Bölümü, Muğla, Türkiye

Anahtar Kelimeler	Öz
<i>Türkiye Basketbol Ligi, Dört Faktör, Makine Öğrenmesi, Tahmin.</i>	Basketbol maçları dünyada en çok izlenen spor aktivitelerinden birisidir. Bu popülerlik sonucunda basketbol sporunda bilgi teknolojileri sık sık kullanılmaktadır. İstatistik ve veri depolama sistemlerinin gelişmesiyle maç istatistikleri, oyuncu özellikleri gibi parametreler artık çok kolay saklanabilmektedir. Müsabaka sonucu tahmini insanlara müsabaka öncesi bilgi vermesi açısından önemlidir. Müsabaka sonucu tahmini ile ilgili olarak Amerikan basketbol ligleri üzerine çalışmalar mevcut iken Avrupa basketbolu ligleri üzerine çalışmaların sayısı azdır. Literatürde bir ilk olarak bu çalışmada, müsabaka sonucu tahmini için Türkiye Erkekler Basketbol Süper Ligi maçları Dört Faktör ve DefansOfans modeli ile makine öğrenmesi yöntemleri beraber kullanılarak ele alınmıştır. Buna göre 2017-2018 tarihlerinde oynanan maçlar veri seti olarak kullanılmış ve Dört Faktör modeli ile birlikte incelenmiştir. Popüler makine öğrenmesi yöntemleri olan kNN, Lojistik Regresyon, Çok Katmanlı Algılayıcı, Naive Bayes, j48 ve Oylama ile kullanılmış ve çıkan sonuçlar değerlendirilmiştir. Sonuç olarak %96,55'lik tahmin başarıları yakalanmıştır.

## PREDICTION OF TURKISH MEN'S BASKETBALL SUPER LEAGUE GAME RESULTS WITH MACHINE LEARNING METHODS

Keywords	Abstract
<i>Turkish Basketball League, Four Factors, Machine Learning, Prediction.</i>	Basketball competitions are one of the most watched sports activities in the world. As a result of this popularity, information technologies are frequently used in basketball. With the development of statistics and data storage systems, parameters such as match statistics, player properties can now be stored very easily. The prediction of the competition result is important in that it provides information to people before the competition. While there are studies on American basketball leagues in this topic, there are few studies on European basketball leagues. As a first in the literature in this study, results of competitions for Turkish Men's Basketball Super League game prediction were examined by using Four Factor model, DefenseOfense model and machine learning methods together. Accordingly, the matches played between 2017-2018 were used as data set and examined together with the Four Factor model. Popular machine learning methods; kNN, Logistic Regression, Multilayer Perceptron, Naive Bayes, j48 and Voting were used and the results were evaluated. As a result, 96.55% predicted success was achieved.

### Alıntı / Cite

Özdemir, E., Ballı, S., (2020). Türkiye Erkekler Basketbol Ligi Maç Sonuçlarının Makine Öğrenmesi Yöntemleri ile Tahmini, Mühendislik Bilimleri ve Tasarım Dergisi, 8(3), 740-752.

Yazar Kimliği / Author ID (ORCID Number)	Makale Süreci / Article Process	
E. Özdemir, 0000-0002-6740-8444	Başvuru Tarihi / Submission Date	19.04.2020
S. Ballı, 0000-0002-4825-139X	Revizyon Tarihi / Revision Date	02.07.2020
	Kabul Tarihi / Accepted Date	14.07.2020
	Yayın Tarihi / Published Date	24.09.2020

\* İlgili yazar / Corresponding author: serkan@mu.edu.tr , +90-252-211-3106

## 1. Giriş (Introduction)

Spor müsabakaları günümüzde insanlar tarafından rağbet edilen en popüler hobi türlerinin başında yer almaktadır. Basketbol bu müsabakalardan birisidir. Spor yöneticileri makine öğrenmesi yöntemlerini oyuncu seçiminde, performans değerlendirmesi ve hatta sonuç tahmininde, spor meraklıları ise spor bahislerinde kullanmaktadır (Horvat vd., 2018).

İnternet üzerinde geçmiş maçlara dair sonuçlara ve istatistiklere ulaşmak diğer birçok alana göre çok daha kolaydır. Bu sebeple spor maçlarının sonucunun tahmin edilmesi geçmiş çalışmaların birçoğuna ilham vermiştir. Geçmiş çalışmaların çoğu futbol ve basketbol maçları üzerine olsa da, buz hokeyi, Amerikan futbolu, tenis gibi diğer spor dalları için de çalışmalar bulunmaktadır (Karaoğlu, 2015).

Peki, takımlar basketbol maçlarını nasıl kazanır, istatistiklerin kazanmaya etkisi nedir gibi sorulara cevap arayan Dean Oliver, Basketbol Başarısının Dört Faktörünü tanımlamıştır (Oliver, 2004). Bu tanıma göre şutların kazanmaya etkisi %40, top kayıplarının %25, ribauntların %20 ve serbest atışların ise %15'dir. Bu faktörler savunma ve hücum anındaki farkları, takımın kazanmasında belirleyici faktörleri oluşturmaktadır (Çiçekli vd., 2012). Kısacası sadece çok sayı atarak bir maçın kazanılamayacağını, diğer faktörlerin de maç kazanmaya etkisi olduğunu göstermektedir. Aynı zamanda bu faktörler hem hücum hem de defansif olarak ayrı ayrı hesaplanmaktadır.

Basketbol maç sonuçları tahmini ile ilgili olarak Amerikan basketbol ligleri üzerine çalışmalar mevcut iken Avrupa basketbolu ligleri üzerine yapılan çalışmaların sayısı azdır. Bu çalışmada bir ilk olarak Türkiye Erkekler Basketbol Süper Ligi maçları tahmin edilmiştir. Tahmin için DefansOfans modeli ve Dört Faktör modeli ile birlikte analiz edilmiştir. Amerikan basketbol ligleri üzerine Dört Faktör çalışmaları bulunurken ve hatta NBA resmi sitesinde bu konuda istatistiksel bilgiler yer alırken, Avrupa basketbolu ligleri üzerine Dört Faktör modeli kullanılarak bir tahmin çalışması daha önce yapılmamıştır. Horvat vd. (2018) tarafından yapılan çalışmada kullanılan DefansOfans modeli Euroleague veri seti üzerinde en çok başarı getiren model olmuştur. Bunun yanı sıra yeni melez modeller de eklenerek model sayısı arttırılmıştır ve maç istatistikleri üzerinden daha birçok model ile tahmin edilebileceği gösterilmiştir. Tahmin için kNN, Lojistik Regresyon, Çok Katmanlı Algılayıcı, Naive Bayes, j48 ve Oylama gibi makine öğrenmesi yöntemleri kullanılmıştır.

Çalışmanın ikinci bölümünde spor müsabakalarının sonucunu tahmin etmek için yürütülen geçmiş çalışmalara yer verilmiştir. Üçüncü bölümde bu çalışmada kullanılan makine öğrenmesi algoritmaları anlatılmıştır. Dördüncü bölümde elde edilen veri kümesi, veri kümesinin Dört Faktör ve DefansOfans modelleri üzerinde uygulanması sonucu elde edilen modeller ve algoritmalar ile tahmin edilmesi yer almaktadır. Beşinci bölümde tartışmalar sunulmuştur. Altıncı bölümde ise elde edilen sonuçlara değinilmiştir.

## 2. Kaynak Araştırması (Literature Survey)

Tablo 1'de basketbol tahminleri için geçmiş çalışmalara ait veri setleri, veri sayısı, en başarılı algoritma ve başarı oranı bilgileri gösterilmiştir.

Horvat vd. (2018) tarafından yapılan çalışmada, Euroleague 2012-2013 ve 2016-2017 sezonları arasındaki müsabaka sonuçları kNN metoduyla tahmin edilmeye çalışılmıştır. %84 oranında başarı yakalamışlardır. Çiçekli vd. (2012) tarafından yapılan çalışmada, Türkiye Erkekler Basketbol Süper Ligi 2010-2011 sezonu maçları Dört Faktör modeli ile incelenmiştir fakat tahmin yapılmamıştır. Zimmermann vd. (2013) yaptığı çalışmada, NCAA Ligi 2011-2012 sezonu maçlarını Ağırlıklı Dört Faktör metodu ile modelleyerek Yapay Sinir Ağları ile eğitmiş ve sonuçları %73 oranında doğru tahmin etmişlerdir. Zimmermann'ın (2016) yaptığı çalışmada, NBA Ligi 2014-2015 sezonunu Ağırlıklı Dört Faktör metodu ile modelleyerek birçok sınıflandırma metodu kullanarak sonuçları tahmin etmiştir. En etkili yöntemin Naive Bayes olduğu kararına varmıştır ve %65'lik başarı oranı yakalamıştır. Ozmen'in (2016) sunduğu çalışmada, Euroleague'in 14 sezonunu değerlendirerek her bir istatistiğin basketbol maçının kazanılmasındaki katkısını Probit Dağılımı yöntemi ile belirlemeye çalışmıştır. Erilli vd. (2013) yaptığı çalışmada, Euroleague 2011-2012 sezonu maçlarının skor tahminlerini Poisson Dağılımı ile bulmaya çalışmışlardır ve %73 oranında başarı sağlamışlardır. Cao (2012) sunduğu çalışmada, NBA Ligi 2005-2006 sezonu ile 2010-2011 sezonu arasındaki maçları sınıflandırma algoritmaları kullanarak tahmin etmeye çalışmıştır. En verimli yöntemin %69,67 başarı oranı ile Lojistik Dağılım olduğu sonucu çıkmıştır. Miljković vd. (2010) tarafından yapılan çalışmada, NBA Ligi 2009-2010 sezonu maçlarını Naive Bayes ve Lineer Dağılım ile sınıflandırmış ve sonuçları karşılaştırmışlardır. En etkili yöntemin Naive Bayes olduğu kararına varmıştır ve %67'lik başarı oranı yakalamıştır. Harmandeep ve Sushma (2017) tarafından yapılan çalışmada, NBA Ligi 2015-2016 sezonu maçlarını Destek Vektör Makinaları ve Bulanık Hibrit Destek Vektör Makinaları ile tahmin etmiştir. En verimli yöntemin %88,26 başarı oranı ile birlikte Destek Vektör Makinaları olduğu ortaya çıkmıştır. Ayyıldız (2018) sunduğu çalışmada, NBA Ligi 2015-2016

sezonu maçları Yapay Sinir Ağları ile eğitmiş ve sonuçları %90 oranında doğru tahmin etmiştir. Torres (2013) tarafından yapılan çalışmada, NBA Ligi 2005-2006 sezonu ile 2012-2013 sezonu arasındaki karşılaşmalar çeşitli makine öğrenmesi yöntemleriyle tahmin edilmeye çalışılmıştır. En iyi yöntemin %68,44 başarı oranı ile birlikte Çok Katmanlı Algılayıcı ile belirlendiği ortaya çıkmıştır. Thabtah vd. (2019) tarafından sunulan çalışmada, NBA Ligi Finallerini içeren 1980 yılından başlayan 2017'e kadar süren maçlar çeşitli modeller kullanılarak, özellik seçimi yardımıyla tahmin edilmeye çalışılmıştır. En başarılı yöntem Lojistik Model Ağacı algoritması ile %83 oranındadır. Caliwag vd. (2018) tarafından yapılan çalışmada, NBA Ligi 2015-2016 sezonu maçları Dört Faktör analizi ile birlikte çeşitli makine öğrenmesi yöntemleriyle tahmin edilmeye çalışılmıştır. En başarılı sonuç olan %70'e Kademeli Algoritma yardımıyla ulaşılmıştır.

Bu çalışmada ise Türkiye Erkekler Basketbol Süper Ligi 2017-2018 sezonu maçları çeşitli makine öğrenmesi yöntemleriyle tahmin edilmeye çalışılmış olup, en iyi sonucun %96,55 başarı oranı ile birlikte Lojistik Regresyon olduğu ortaya çıkmıştır.

**Tablo 1. Geçmiş Çalışmaların Listesi (List of Previous Studies)**

Yazar	Yıl	Veri Sayısı	Veri Kümesi	Algoritma	Başarı Oranı
Horvat vd.	2018	-	EuroLeague 2012-13 ile 2016-17 arası	kNN	%83,96
Zimmermann vd.	2013	21373 - Eğitim 5399 - Test	NCAA 2011-12	Yapay Sinir Ağları	%73,11
Zimmermann	2016	-	NBA 2014-15	Naive Bayes	%65,90
Ozmen	2016	3000'den fazla maç	EuroLeague 2000-01 ile 2013-14 arası	Probit Regresyon	-
Erilli vd.	2013	186 maç	EuroLeague 2011-12	Poisson Regresyon	%73,00
Cao	2012	-	NBA 2005-06 ile 2010-11 arası	Lojistik Dağılım	%69,67
Miljković vd.	2010	778 maç	NBA 2009-10	Naive Bayes	%67
Harmandeep	2017	-	NBA 2015-16	Destek Vektör Makineleri	%88,26
Ayyıldız	2018	396 - Eğitim 200 - Test	NBA 2015-16	Yapay Sinir Ağları	%90
Torres	2013	-	NBA 2005-06 ile 2012-13 arası	Çok Katmanlı Algılayıcı	%68,44
Thabtah vd.	2019	-	NBA Finalleri 1980 ile 2017 arası	Lojistik Model Ağacı	%83
Caliwag vd.	2018	-	NBA 2005-06	Kademeli Algoritma	%70
Bu Çalışma		260 maç	TBL 2017-2018	Lojistik Regresyon	%96,55

### 3. Makine Öğrenmesi Yöntemleri (Machine Learning Methods)

Spor maçları sonuçlarının tahmini için makine öğrenmesi yöntemlerine başvurulmaktadır. Bu çalışmada kullanılan makine öğrenmesi yöntemleri alt bölümlerde açıklanmıştır.

#### 3.1. k En Yakın Komşu (k Nearest Neighbor)

kNN algoritması gözetimli öğrenme algoritmalarındandır. Aynı zamanda hem sınıflandırma hem de regresyon problemlerinde kullanılmaktadır. Sınıflandırılacak olan yeni örnek için eğitim setinden en yakın mesafedeki k adet veriye bakılır ve bu k örnek çoğunluk olarak hangi sınıftan mevcut ise yeni örnek de o sınıfa dâhil edilir (Sağbaşı ve Ballı, 2016). Hesaplama yöntemleri arasında Öklid, Manhattan ve Minkowsky formülleri bulunur. En çok tercih edilen Eşitlik (1)'de verilen Öklid formülüdür. Bu çalışmada Öklid uzaklık ölçütü kullanılmıştır.

$$d = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (1)$$

d : İki nokta arasındaki doğrusal uzaklık

x : X noktası

y : Y Noktası

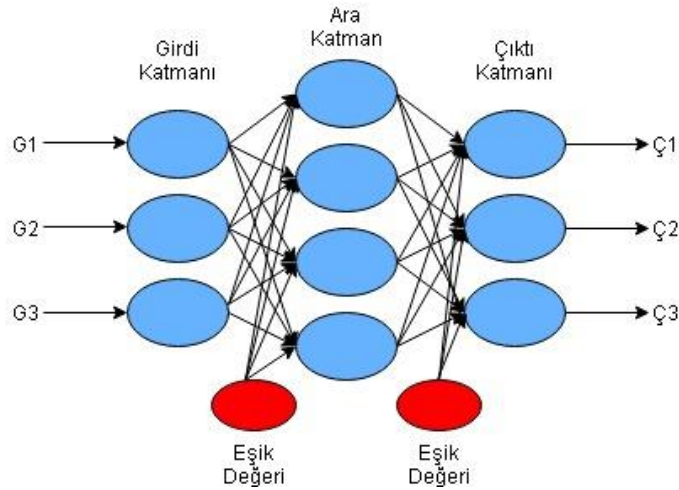
i : Boyut sayısı

### 3.2. Lojistik Regresyon (Logistic Regression)

Lojistik Regresyon, yalnızca ikili bir sınıflandırma problemini çözmeye yarayan bir regresyon analizidir. Lineer Regresyon modelinin gelişmiş hali olarak kategorik değişkenlerin sonuçlarını tahmin etmek için kullanılabilir. Yani bağımsız değişkenleri verilen, kategorik dağıtılan bağımlı değişkenlerin farklı olası sonuçlarını tahmin etmek için kullanılan modeldir (Sağbaş ve Ballı, 2016). Lineer Regresyon modelinde bağımlı ve bağımsız değişkenler sayısal olarak ifade edilmektedir. Örneğin bu çalışmada kullanılan modellerde ev sahibi takım maçı kazandı mı sorusuna cevap aranmaktadır. Sınıf ise ikili olarak 0 ve 1 olarak tutulmaktadır. Dolayısıyla test verilerinin sonuçları da iki sınıftan birisine dâhil olacaktır.

### 3.3. Çok Katmanlı Algılayıcı (Multilayer Perceptron)

Tek Katmanlı Algılayıcı, doğrusal olarak sınıflandırılmayan problemleri çözememektedir. Bu probleme çözüm üretmek için yapılan çalışmalar sonucunda Çok Katmanlı Algılayıcı modeli ortaya çıkmıştır. Geriye yayımlı çok katmanlı ağ yapısı içermektedir. Tek Katmanlı Algılayıcı modelinden farklı olarak ara katman içermektedir. Bu ara katman yardımıyla giriş katmanından gelen veriler çıktı katmanına gönderilmeden önce probleme göre uyarlanarak değiştirilebilir ve sistemin daha iyi karar vermesi sağlanır (Torres, 2013). Şekil 1'de örnek çok katmanlı algılayıcı modeli gösterilmiştir.



Şekil 1. Çok Katmanlı Algılayıcı Modelinin Yapısı (Structure of Multi Layer Perceptron Model)

### 3.4. Naive Bayes (Naive Bayes)

Adını İngiliz matematikçi Thomas Bayes'ten alan bu sınıflandırma algoritması temel olarak Bayes teoremini içermektedir ve olasılık hesabına dayanmaktadır. Test verisi üzerinden öğrenme işlemini gerçekleştirir ve en yüksek orandaki örneğini sınıfa dâhil eder (Sağbaş ve Ballı, 2017; Ballı ve Karasoy, 2019). Eşitlik (2)'de Bayes formülü verilmiştir.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (2)$$

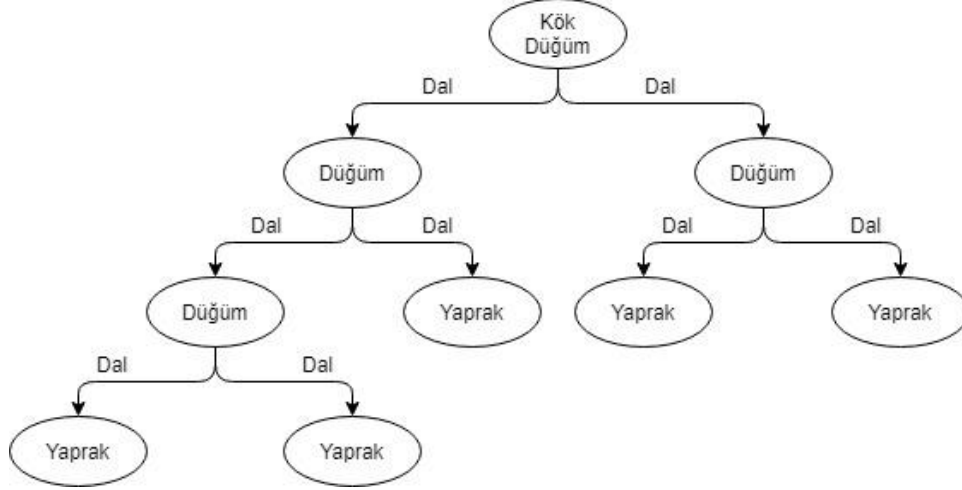
$P(A|B)$  : B olayı gerçekleştiği durumda A olayının meydana gelme olasılığı

$P(B|A)$  : A olayı gerçekleştiği durumda B olayının meydana gelme olasılığı

$P(A)$  ve  $P(B)$  : A ve B olaylarının önsel olasılıklarıdır.

### 3.5. j48 Yöntemi (j48 Method)

C4.5 olarak da bilinir. Sınıflandırma problemleri için en çok kullanılan karar ağacı algoritmalarından birisidir. ID3 algoritmasının eksikliklerini gidermek amacıyla geliştirilmiştir. ID3 algoritması sadece kategorik verilerle çalışırken, j48 algoritmasında sayısal verilerle çalışabilmektedir (Özkan ve Erol, 2015). Entropi hesabına dayanmaktadır (Sağbaşı vd., 2020). Şekil 2’de örnek bir karar ağacı yapısı gösterilmiştir.

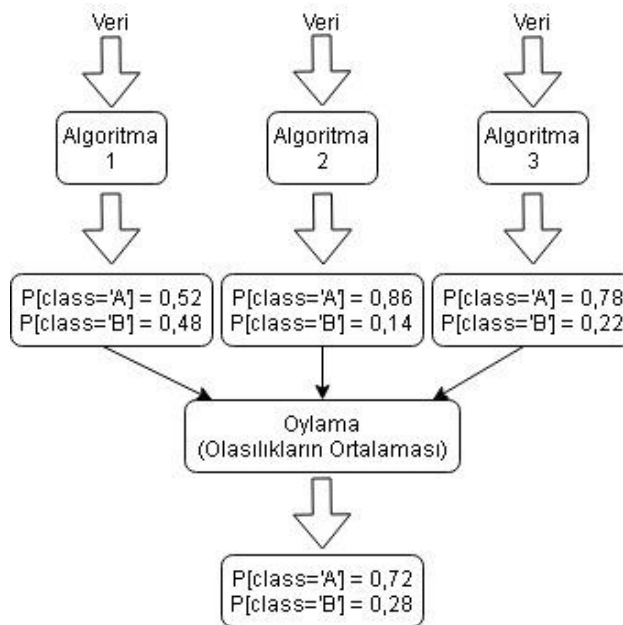


Şekil 2. Karar Ağacı Yapısı (Structure of Decision Tree)

### 3.6. Oylama (Voting)

Voting algoritması olarak da bilinir. Çoklu sınıflandırma yöntemlerinden birisidir. Bu çalışmada kullanılan diğer algoritmalar tekli sınıflandırma yöntemleri arasındadır. Birden çok tekli sınıflandırmayı tek bir karar doğrultusunda birleştirir ve tek bir sonuca bağlar. Altı adet karar bulunmaktadır (Nanğır, 2013). Bu çalışmada Olasılıkların Ortalaması kararı kullanılmıştır. Her bir tekli sınıflandırıcıya ait sonuçlar olumlu ve olumsuz olarak ikiye ayrılmaktadır. Her bir sonucun kendi arasında ortalaması alınarak sistemin genel kararını ortaya çıkarmaktadır. Şekil 3’de olasılıkların ortalaması kararına ait örnek bir çalışma sunulmuştur.

- Çoğunluk Oylaması (Majority Voting)
- Olasılıkların Ortalaması (Average of Probabilities)
- Olasılıkların Çarpımı (Product of Probabilities)
- En Yüksek Olasılık (Maximum Probability)
- En Düşük Olasılık (Minimum Probability)
- Ortanca (Median)



Şekil 3. Olasılıkların Ortalaması (Average of Probabilities)

### 3.7. Değerlendirme Ölçütleri (Evaluation Metrics)

Sınıflandırma algoritmaları ile yapılan tahminlerde doğruluk oranının yanında değerlendirmek için başka ölçütler de kullanılmaktadır. Şekil 4'de sınıflandırma işlemlerinde gerçekleşen ve tahmin edilen değerleri gösteren Karmaşıklık Matrisine yer verilmiştir (Celik vd., 2020). Doğru Pozitif (DP) ve Doğru Negatif (DN) alanları tahmini doğru gerçekleşen, Yanlış Pozitif (YP) ve Yanlış Negatif (YN) alanları ise tahmini yanlış gerçekleşen bölümlerdir (Işık ve Işık, 2020).

GERÇEKLEŞEN		TAHMİN EDİLEN	
Doğru	Yanlış	Doğru	Yanlış
Doğru Pozitif	Yanlış Pozitif	Doğru	Yanlış
Yanlış Negatif	Doğru Negatif	Doğru	Yanlış

Şekil 4. Karmaşıklık Matrisi (Confusion Matrix)

Doğruluk (Accuracy), tahmin işlemi doğru gerçekleşen değerlerin tüm değerlere bölünmesiyle hesaplanmaktadır. Eşitlik (3)'te formülüne yer verilmiştir.

$$Accuracy = \frac{DP + DN}{DP + DN + YP + YN} \quad (3)$$

Kesinlik (Precision), tahmin işlemi doğru olarak gerçekleşen değerlerden kaç tanesinin doğru olduğunu bize söylemektedir. Sınıflandırma işlemleri sonuçlarında bu değer yüksek çıkması beklenmektedir. Eşitlik (4)'te formülüne yer verilmiştir.

$$Precision = \frac{DP}{DP + YP} \quad (4)$$

Duyarlılık (Recall), doğru sınıflandırılmış olan pozitif değerlerin toplam pozitiflere bölünmesiyle hesaplanmaktadır. Eşitlik (5)'te formülüne yer verilmiştir.

$$Recall = \frac{DP}{DP + YN} \quad (5)$$

F ölçütü (F-measure), Kesinlik (Precision) ve Duyarlılık (Recall) değerlerinin harmonik ortalaması alınarak hesaplanmaktadır. Eşitlik (6)'da formülüne yer verilmiştir.

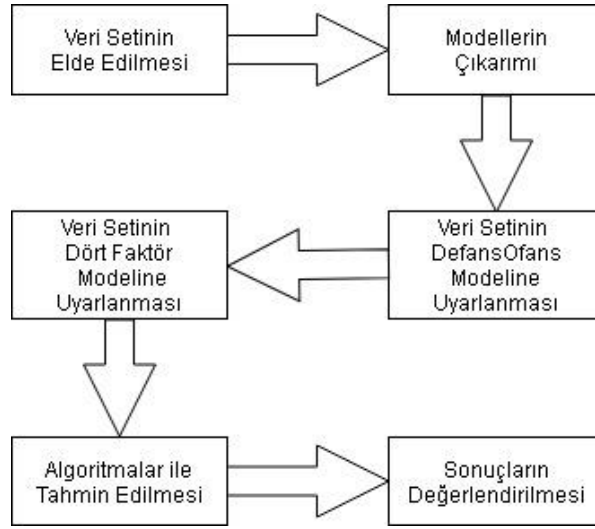
$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (6)$$

RMSE (Root Mean Square Error - Ortalama Hata Kareleri Kökü), tahmin verileri ile gerçek veriler arasındaki uzaklık hesabına dayanmaktadır. Eşitlik (7)'de formülüne yer verilmiştir.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} \quad (7)$$

### 4. Veri Kümesinin Modellenmesi ve Algoritmalar ile Tahmin Edilmesi (Modeling of the Data Set and Prediction with Algorithms)

Yapılan çalışmada maç sonuçlarının doğru tahmin edilmesi amaçlanmıştır. Bunun için yapılan işlemler Şekil 5'de gösterilmiştir. Çalışmada kullanılan veri kümesi, veri dönüşümleri, makine öğrenmesi yöntemleri alt bölümlerde detaylı olarak anlatılmıştır.



Şekil 5. Yapılan Çalışmanın Temel Akış Şeması (Basic Flow Chart of the Study)

#### 4.1. Veri Kümesi (Dataset)

Bu çalışmada veri kümesi olarak Türkiye Erkekler Basketbol Süper Ligi 2017-2018 sezonu maçları kullanılmıştır. [www.tblstat.net](http://www.tblstat.net) adlı web sitesinden maç istatistiklerini alabilmek için Android tabanlı bir uygulama yazılmıştır. Uygulama aracılığı ile alınan maç istatistikleri veritabanına kaydedilmiştir.

Ev sahibi ve deplasman takım için istatistikler ayrı ayrı tutulmuştur. Home, ev sahibini; Away, deplasman takımını ifade eder. Ev sahibinin attığı sayı ile deplasman takımının attığı sayı karşılaştırılır ve buna göre home\_win değişkeni 0 veya 1 olarak kaydedilir. Toplamda 260 adet maç değerlendirmeye alınmıştır ve sonuçları makine öğrenmesi teknikleriyle tahmin edilmeye çalışılmıştır. Elde edilen istatistiksel veriler ve açıklamaları Tablo 2’de verilmiştir:

Tablo 2. Veri Tabanında Bulunan İstatistikler ve Açıklamaları (Statistics and Descriptions in the Database)

Kısa Adı	Açıklama
2FGM (Field Goal Made)	2 sayılık isabet
2FGA (Field Goal Attempt)	2 sayılık deneme
3FGM (Field Goal Made)	3 sayılık isabet
3FGA (Field Goal Attempt)	3 sayılık deneme
FTM (Free Throw Goal Made)	Serbest atış isabeti
FTA (Free Throw Goal Attempt)	Serbest atış deneme
DReb	Defansif ribaunt
OReb	Ofansif ribaunt
Assists	Asist
Steals	Top çalma
Turnovers	Top kaybı

#### 4.2. Veri Dönüşümleri (Data Transformations)

Elde edilen veri seti üzerinden tahmin yapabilmek için model dönüşümlerine ihtiyaç duyulmuştur. Bunun için DefansOfans ve Dört Faktör modelleri baz alınmıştır. Bunlardan oluşan melez modellerle birlikte 8 adet model üretilmiştir ve maç sonuçları tahmin edilmeye çalışılmıştır.

##### 4.2.1. DefansOfans Modeli (DefenseOfense Model)

Horvat vd. tarafından (2018) yayınlanan makalede DefansOfans (DefenseOfense) olarak isimlendirdiği modelinde temel basketbol istatistiklerini kullanmıştır. Ev sahibi ve deplasman takımı istatistikleri üzerine yoğunlaşmıştır. Ancak bu çalışmada, veri setleri içerisinde blok ve faul verilerine ulaşamadığı için model üzerinde bu veriler kullanılmamıştır. Sonuç olarak, tüm istatistikler  $i\_offence$  ve  $i\_defence$  adında iki parametreye indirgenmiştir.

Orijinal Model için hesaplama işlemleri Tablo 3’de gösterilmiştir. Bu çalışmada kullanılan model parametreleri ve hesaplama işlemleri Tablo 4’de gösterilmiştir.

**Tablo 3.** DefansOfans Orijinal Modeli Hesaplama İşlemleri (Calculation Procedures of DefenseOfense Original Model)

İsim	Hesaplama Yöntemi
home_i_offence	home_2FGM+home_3FGM+home_FTM-(home_2FGA-home_2FGM+home_3FGA-home_3FGM+home_FTA-home_FTM)+ home_OReb+home_Assists+home_Fouls_Rv-home_Turnovers-home_Blocks_Ag
home_i_defence	(away_2FGM+away_3FGM+away_FTM)+(away_2FGA-away_2FGM+away_3FGA-away_3FGM+away_FTA-away_FTM)+home_DReb+home_Steals+ home_Blocks_Fv-away_OReb-home_Fouls_Com

**Tablo 4.** DefansOfans Modeli Yapısı (Structure of DefenseOfense Model)

İsim	Hesaplama Yöntemi	Tipi
home_i_offence	home_2FGM+home_3FGM+home_FTM- (home_2FGA-home_2FGM+home_3FGA-home_3FGM+home_FTA-home_FTM)+home_OReb+home_Assists-home_Turnovers	Özellik
home_i_defence	-(away_2FGM+away_3FGM+ away_FTM)+(away_2FGA-away_2FGM+away_3FGA-away_3FGM+away_FTA-away_FTM)+home_DReb+home_Steals-away_OReb	Özellik
home_win	Ev sahibi takım kazandı mı? (0,1)	Sınıf

#### 4.2.2. Dört Faktör Modeli (Four Factors Model)

Takımların basketbol maçlarını kazanmasının etkisini istatistikler üzerinden bulmaya çalışan Dean Oliver, Basketbol Başarısının Dört Faktörünü tanımlamıştır (Oliver, 2004). Bu tanıma göre şutların kazanmaya etkisi %40, top kayıplarının %25, ribauntların %20 ve serbest atışların %15’dir. Aynı zamanda bu temel değerler hem hücum hem de defansif olarak ayrı ayrı hesaplanmaktadır. Tüm istatistikler sonuç olarak offence\_rating ve defence\_rating adında iki parametreye indirgenmiştir. Dört Faktör hesaplama işlemleri Tablo 5 ve Tablo 6’da belirtilmiştir.

1. Effective Field Goal Percentage: Etkin Alan Sayısı yüzdesi olarak bilinir. Serbest atışlar dahil edilmez. Kazanmaya etkisi %40 olarak görülmektedir. Eşitlik (8)’de hesaplama işlemi belirtilmiştir.

$$efgp = \frac{FGM + 0.5 \times (3FGM)}{FGA} \quad (8)$$

2. Turnover Percentage: Top Kaybı yüzdesi olarak geçmektedir. Kazanmaya etkisi %25 olarak görülmektedir. Eşitlik (9)’da hesaplama işlemi belirtilmiştir.

$$tp = \frac{Turnovers}{0.96 \times (FGA - OReb - Turnovers + (0.475 \times FTA))} \quad (9)$$

3. Offensive Rebound Percentage: Ofansif Ribaunt yüzdesi olarak geçmektedir. Kazanmaya etkisi %20 olarak görülmektedir. Eşitlik (10)’da hesaplama işlemi belirtilmiştir.

$$orp = \frac{OReb}{OReb + DReb_{Rakip}} \quad (10)$$

4. Free Throw Rating: Serbest Atış yüzdesi olarak geçmektedir. Kazanmaya etkisi %15 olarak görülmektedir. Eşitlik (11)’de hesaplama işlemi belirtilmiştir.

$$ftr = \frac{FTA}{FGA} \quad (11)$$



**Tablo 5.** Dört Faktör Hesaplama İşlemleri (Calculation Operations of Four Factors)

İsim	Hesaplama Yöntemi
home_offence_efgp	$(\text{home\_2FGM} + \text{home\_3FGM} + (0.5 * \text{home\_3FGM})) / (\text{home\_2FGA} + \text{home\_3FGA})$
home_offence_tp	$\text{home\_Turnovers} / (\text{home\_2FGA} + \text{home\_3FGA} + (0.44 * \text{home\_FTA}) - \text{home\_OReb} + \text{home\_Turnovers})$
home_offence_orp	$(\text{home\_OReb}) / (\text{home\_OReb} + \text{away\_DReb})$
home_offence_ftr	$(\text{home\_FTA}) / (\text{home\_2FGA} + \text{home\_3FGA})$
home_defence_efgp	$(\text{away\_2FGM} + \text{away\_3FGM} + (0.5 * \text{away\_3FGM})) / (\text{away\_2FGA} + \text{away\_3FGA})$
home_defence_tp	$\text{away\_Turnovers} / (\text{away\_2FGA} + \text{away\_3FGA} + (0.44 * \text{away\_FTA}) - \text{away\_OReb} + \text{away\_Turnovers})$
home_defence_orp	$(\text{away\_OReb}) / (\text{away\_OReb} + \text{home\_DReb})$
home_defence_ftr	$(\text{away\_FTA}) / (\text{away\_2FGA} + \text{away\_3FGA})$

**Tablo 6.** Dört Faktör Modeli Yapısı (Structure of Four Factors Model)

İsim	Hesaplama Yöntemi	Tipi
home_offence_rating	$(0.4 * \text{home\_offence\_efgp}) + (0.25 * \text{home\_offence\_tp}) + (0.2 * \text{home\_offence\_orp}) + (0.15 * \text{home\_offence\_ftr})$	Özellik
home_defence_rating	$(0.4 * \text{home\_defence\_efgp}) + (0.25 * \text{home\_defence\_tp}) + (0.2 * \text{home\_defence\_orp}) + (0.15 * \text{home\_defence\_ftr})$	Özellik
home_win	Ev sahibi takım kazandı mı? (0,1)	Sınıf

#### 4.2.3. DefansOfans Detaylı Modeli (DefenseOfense Detailed Model)

DefansOfans modelini oluşturan 12 adet özellikten oluşmaktadır. Tablo 7’de bu özellikler ve açıklamaları gösterilmiştir.

**Tablo 7.** DefansOfans Detaylı Modeli Yapısı (Structure of DefenseOfense Detailed Model)

İsim	Açıklama	Tipi
home_2FGM	Ev sahibi takımın 2 sayılık isabeti	Özellik
home_2FGA	Ev sahibi takımın 2 sayılık denemesi	Özellik
home_3FGM	Ev sahibi takımın 3 sayılık isabeti	Özellik
home_3FGA	Ev sahibi takımın 3 sayılık denemesi	Özellik
home_FTM	Ev sahibi takımın serbest atış isabeti	Özellik
home_FTA	Ev sahibi takımın serbest atış denemesi	Özellik
home_DReb	Ev sahibi takımın defansif ribauntları	Özellik
home_OReb	Ev sahibi takımın ofansif ribauntları	Özellik
home_TReb	Ev sahibi takımın toplam ribauntları	Özellik
home_Assists	Ev sahibi takımın asistleri	Özellik
home_Steals	Ev sahibi takımın top çalmaları	Özellik
Home_Turnovers	Ev sahibi takımın top kayıpları	Özellik
home_win	Ev sahibi takım kazandı mı? (0,1)	Sınıf

#### 4.2.4. Dört Faktör Detaylı Modeli (Four Factors Detailed Model)

Dört Faktör modelini oluşturan 8 adet özellikten oluşmaktadır. Tablo 8’de bu özellikler ve açıklamaları gösterilmiştir.

**Tablo 8.** Dört Faktör Detaylı Modeli Yapısı (Structure of Four Factor Detailed Model)

İsim	Açıklama	Tipi
home_offence_efgp	Ev sahibi takımın ofansif etkin alan yüzdesi	Özellik
home_offence_tp	Ev sahibi takımın ofansif top kaybı yüzdesi	Özellik
home_offence_orp	Ev sahibi takımın ofansif ribaunt yüzdesi	Özellik
home_offence_ftr	Ev sahibi takımın ofansif serbest atış yüzdesi	Özellik
home_defence_efgp	Ev sahibi takımın defansif etkin alan yüzdesi	Özellik
home_defence_tp	Ev sahibi takımın defansif top kaybı yüzdesi	Özellik
home_defence_orp	Ev sahibi takımın defansif ribaunt yüzdesi	Özellik
home_defence_ftr	Ev sahibi takımın defansif serbest atış yüzdesi	Özellik
home_win	Ev sahibi takım kazandı mı? (0,1)	Sınıf

#### 4.2.5. Melez Modeller (Hybrid Models)

Dört Faktör ve DefansOfans modelleri literatürde kullanılan modellerdir. Yalnız her iki modelde de çok sayıda özellik 2 özelliğe indirgenmektedir. Dört Faktör modelinde toplam 8 özellik, DefansOfans modelinde ise 12 özellik ile hesaplama işlemlerinde kullanılmaktadır. İki modelin birleşimi ile farklı modeller elde edilmiş ve başarı oranını arttırıp arttırmadığı araştırılmıştır. Aşağıda modellerin yanında kaç adet özelliğe sahip olduğu bilgisi verilmiştir.

- Dört Faktör (2 özellik) ve DefansOfans (2 özellik) modeli birleştirilerek elde edilen yeni modelde 4 adet özellik bulunmaktadır.
- Dört Faktör Detaylı (8 özellik) ve DefansOfans (2 özellik) modeli birleştirilerek elde edilen yeni modelde 10 adet özellik bulunmaktadır.
- Dört Faktör (2 özellik) ve DefansOfans Detaylı (12 özellik) modeli birleştirilerek elde edilen yeni modelde 14 adet özellik bulunmaktadır.
- Dört Faktör Detaylı (8 özellik) ve DefansOfans Detaylı (12 özellik) modeli birleştirilerek elde edilen yeni modelde 20 adet özellik bulunmaktadır.

#### 4.3 Algoritmalar ile Tahmin Edilmesi (Prediction with Algorithms)

İlk olarak veri setleri ile oluşturulan 8 adet model Weka programı aracılığı ile tahmin edilmeye çalışılmıştır. Veri seti eğitim seti olarak yaygın kullanılan %66, %80 ve 10-katlı çapraz doğrulama şeklinde ayrılmıştır. kNN algoritmasında k değerleri için 1, 3, 5, 7, 9, 11, 13, 15 kullanılmıştır. Çıkan sonuçlar Tablo 9'da verilmiştir. Yüksek başarı oranına sahip sonuçlar kalın fontla belirtilmiş ve bu sonuçlarda kullanılan parametreler tablonun altında açıklanmıştır.

**Tablo 9.** Modellerin Algoritmalar ile Tahmin Edilmesi Sonucu Çıkan İlk Sonuçlar (The First Estimating Results of Models with Algorithms)

	kNN	Lojistik Regresyon	Çok Katmanlı Algılayıcı	Naive Bayes	j48
Model 1: DefansOfans	<b>%90,38</b>	%87,5	%86,36	%87,5	%84,2
Model 2: DefansOfans Detaylı	%82,95	%89,77	%88,46	%78,4	%84,6
Model 3: Dört Faktör	%73,07	%71,59	%72,72	%72,7	%71,5
Model 4: Dört Faktör Detaylı	%85,22	%78,07	%75	%82,9	%68,1
Model 5: Dört Faktör & DefansOfans	%88,46	%87,5	%87,5	%82,6	%82,9
Model 6: Dört Faktör Detaylı & DefansOfans	%87,5	<b>%92,04</b>	%86,36	%86,3	%80
Model 7: Dört Faktör & DefansOfans Detaylı	%82,95	%88,63	%80,77	%78,4	%78,8
Model 8: Dört Faktör Detaylı & DefansOfans Detaylı	%86,36	<b>%90,91</b>	<b>%90,91</b>	%81,8	%82,6

- Model 1 ve kNN algoritması için %90,38 çıkan sonuçta eğitim seti %80, algoritma parametresi k=15 olarak seçildi.
- Model 6 ve Lojistik Regresyon algoritması için %92,04 çıkan sonuçta eğitim seti %66 seçildi.
- Model 8 ve Lojistik Regresyon algoritması için %90,91 çıkan sonuçta eğitim seti %66 seçildi.
- Model 8 ve Çok Katmanlı Algılayıcı algoritması için %90,91 çıkan sonuçta eğitim seti %66 seçildi.

**Tablo 10.** Tablo 9'de Yer Alan 90+ Başarılı Olan Sonuçlar ve İyileştirmelerin Sonuçları (90+ Successful Results and Improvements Results in Table 9)

Sıra	Model	Eğitim Seti	Algoritma	Sonuç	F-Measure ROC RMSE	Yeni Eğitim Seti	Yeni Sonuç	F-Measure ROC RMSE
1	Model 6	%66	Lojistik Regresyon	%92,04	0,922 0,978 0,2371	%89	<b>%96,55</b>	0,965 0,994 0,178
2	Model 8	%66	Lojistik Regresyon	%90,91	0,907 0,967 0,2714	%89	<b>%93,10</b>	0,931 0,978 0,2246
3	Model 8	%66	Çok Katmanlı Algılayıcı	%90,91	0,907 0,928 0,299	%76	<b>%93,55</b>	0,934 0,953 0,26
4	Model 1	%80	kNN (k=15)	%90,38	0,905 0,938 0,3104	%88	<b>%93,55</b>	0,935 0,95 0,3969

Tablo 9'da çıkan sonuçlar arasında 90+ başarı oranı yakalanan sonuçların başarısını daha nasıl arttırılacağı üzerinde çalışılmış ve eğitim seti verilerini her biri için %50-%90 arasında değerlendirilmiş olup ve herhangi bir iyileştirme olup olmadığını anlamak için çıkan sonuçları bir öncekiyle karşılaştırılmıştır. Yeni çıkan sonuçlar Tablo

10'da verilmiştir. Aynı veri setleri, Oylama algoritması ile kullanılmıştır. Her bir çalışma için eğitim seti verilerini %50-%90 arasında değerlendirilerek en başarılı sonuca ulaşmaya çalışılmıştır. En başarılı sonuçlar Tablo 11'de belirtilmiştir. Tüm çalışmalar sonucu çıkan en başarılı algoritma ve modelleri Tablo 12'de gösterilmiştir.

**Tablo 11.** Tablo 10'de Yer Alan Sıralı Sonuçlar ile Oylama Algoritmaları ile Tahmin Edilmesi (Estimation with the Voting Algorithms with the Ordered Results in Table 10)

Sıra	Oylama (Lojistik Reg.& Çok Katmanlı Alg.)		Oylama (Lojistik Reg. & Çok Kat. Alg. & Naive Bayes)		Oylama (Lojistik Reg. & Çok Kat. Alg. & Random Forest)		Oylama (Lojistik Reg. & Çok Kat. Alg. & Naive Bayes & Random Forest)	
	Eğitim Seti	Sonuç	Eğitim Seti	Sonuç	Eğitim Seti	Sonuç	Eğitim Seti	Sonuç
1	%89	%93,10	%89	%93,10	%89	%93,10	%89	%93,10
2	%60	%93,27	%89	%93,10	%76	<b>%93,55</b>	%89	%93,10
3	%73	<b>%94,28</b>	%68	%92,77	%73	<b>%94,28</b>	%68	%92,77
4	%76	%88,71	%90	%88,46	%90	%92,30	%76	%88,71

**Tablo 12.** Tablo 11'de Yer Alan Sıralı Sonuçlara Göre En Başarılı Sonuçlar ve Algoritmalar (Most Successful Results and Algorithms According to the Ordered Results in Table 11)

Sıra	Model	Algoritma	Eğitim Seti	Sonuç	F-Measure ROC RMSE
1	Model 6	Lojistik Regresyon	%89	<b>%96,55</b>	0,965 0,994 0,178
2	Model 8	Oylama (Lojistik Reg. & Çok Kat. Alg. & Random Forest)	%76	<b>%93,55</b>	0,934 0,948 0,2585
3	Model 8	Oylama (Lojistik Reg.& Çok Kat. Alg.)	%73	<b>%94,28</b>	0,942 0,964 0,2442
4	Model 1	kNN (k=15)	%88	<b>%93,55</b>	0,935 0,95 0,3969

## 5. Tartışma (Discussion)

Yapılan çalışmada geçmiş çalışmalardan farklı olarak yeni bir veri seti kullanılmıştır. Elde edilen veri setleri iki adet ana model ve melez modeller kullanarak makine öğrenmesi yöntemleriyle tahmin edilmeye çalışılmıştır. Çalışma sonucunda en başarılı sonuç Lojistik Regresyon algoritması ile Model 6 kullanılarak elde edilmiştir.

Horvat vd. (2018) tarafından yapılan çalışmada kullanılan model ile beraber 6 adet yeni model oluşturulmuştur. DefansOfans modeline ek olarak Dört Faktör ve bunların birleşimiyle oluşan melez modeller ile yapılan çalışmadaki model sayısı arttırılmıştır. En başarılı 4 sonuç arasında yer alan 3 adet sonuçta melez modellerin kullanıldığı görülmüştür.

Çalışmanın amacı doğrultusunda Model 6'ın Lojistik Regresyon algoritması ile incelenmesi sonucunda %96,55 oranında en fazla başarıyı sağladığı görülmüştür. Elde edilen bu başarı, kullanılan model ve algoritmanın uygun olduğunu göstermektedir. Model 6 için RMSE değerleri diğerlerine göre en küçük, aynı zamanda ROC ve F-measure değerleri en yüksektir. Ayrıca Model 8'in Oylama (Lojistik Reg.& Çok Katmanlı Alg.) algoritması ile analizi neticesinde %94,28 oranında başarı sağladığı şeklinde yorumlanabilir. Model 8'in Oylama (Lojistik Reg. & Çok Kat. Alg. & Random Forest) algoritması ve Model 1'in kNN algoritması ile analizi neticesinde ise %93,55 oranında bir başarı yakaladığı sonucuna ulaşılmıştır. Ulaşılan bu sonuçlar kullanılan modellerin yapılan tahminlere uygunluğunu göstermektedir.

Ayyıldız tarafından 2018 yılında yapılan çalışma %90 başarı oranına ulaşarak geçmiş çalışmalar arasındaki en başarılı sonuçtur. Belirtilen çalışmada sadece Yapay Sinir Ağları kullanıldığı görülmüştür. Bu çalışmada ise kNN, Lojistik Regresyon, Çok Katmanlı Algılayıcı, Naive Bayes, j48 ve Oylama gibi makine öğrenmesi yöntemleri kullanılmıştır. En başarılı sonuca Lojistik Regresyon yardımıyla ulaşılmıştır. Ayrıca Oylama yöntemleri kullanılarak başarı oranlarının arttırıldığı görülmüştür.

## 6. Sonuçlar (Results)

Bu çalışmada literatürde bir ilk olarak, Türkiye Erkekler Basketbol Süper Ligi veri setinin Dört Faktör ve DefansOfans modeli ile birlikte incelenmesi amaçlanmıştır. Bu amaç doğrultusunda toplamda 260 adet maç değerlendirmeye alınmış ve sonuçları makine öğrenmesi yöntemleriyle tahmin edilmeye çalışılmıştır. En başarılı algoritmanın %96,55 başarı oranı ve %89 eğitim seti değerlendirmesiyle, Lojistik Regresyon olduğu görülmüştür.

Çalışmanın başında eğitim seti olarak yüzde 66, 80 ve 10-katlı çapraz doğrulama kullanılmıştır. Tablo 10'da görüldüğü üzere eğitim seti yüzdesini arttırmak başarı oranını arttırmıştır. Daha sonra oylama algoritmaları ile yapılan tahminlerde Model 8 için başarı oranını arttırdığı görülmüştür. Ayrıca Tablo 9'dan sonra yapılan çalışmalarda eğitim seti değerleri değiştirilerek daha iyi bir sonuca ulaşmak amaçlanmıştır. Eğitim seti değerlerini arttırmanın doğruluk oranını her zaman arttırmadığı gözlemlenmiştir.

Geçmiş çalışmaların çoğu Amerikan basketbol liglerini konu alırken, Avrupa basketbolu ligleri üzerine çalışmaların sayısı oldukça azdır. Avrupa basketbolu gelecek çalışmaların birçoğuna ilham olabilir. Ayrıca bu çalışma sonucunda en başarılı sonuçlar arasında melez modellere rastlanması sebebiyle, elde edilen istatistiksel veriler üzerinden birçok yeni model oluşturulabilir. Elde edilen geçmiş veriler yardımıyla gelecek maçların tahmini ilgili yeni çalışmalar yürütülebilir.

## Çıkar Çatışması (Conflict of Interest)

Yazarlar tarafından herhangi bir çıkar çatışması beyan edilmemiştir. No conflict of interest was declared by the authors.

## Kaynaklar (References)

- Ayyıldız, E., 2018. Amerika Basketbol Ligi (NBA) Maç Sonuçlarının Yapay Sinir Ağları ile Tahmini. Gaziantep Üniversitesi Spor Bilimleri Dergisi, 3(1), 40-53.
- Ballı, S., Karasoy, O., 2019. Development of content-based SMS classification application by using Word2Vec-based feature extraction. IET Software, 13(4), 295-304.
- Caliwag, J.A., Aragon, M.C.R., Castillo, R.E., Colantes, E.M.S., 2018. Prediction Basketball Results Using Cascading Algorithm. In Proceedings of the 2018 International Conference on Information Science and System, 64-68.
- Cao, C., 2012. Sports data mining technology used in basketball outcome prediction. M.Sc. Thesis, Dublin Institute of Technology, Dublin.
- Celik, O., Osmanoglu, U. O., Canakci, B., 2020. Sentiment Analysis From Social Media Comments. Journal of Engineering Sciences and Design, 8(2), 366-374.
- Çiçekli, U.G., Kocamaz, M., Soyuer, H., 2012. Basketbolda Top Hakimiyetinin Analizi ve Takım Performansı Üzerine Etkisi. Pamukkale Journal of Sport Sciences, 3(2), 38-58.
- Erilli, N.A., Ermis, E., Tasmektepligil, M.Y., 2013. Basketball "Turkish Airlines Euro League" 2011-12 season poisson regression simulation modeling. International Journal of Academic Research Part B, 5(5), 151-158.
- Harmandeep, K., Sushma, J., 2017. Machine Learning Approaches to Predict Basketball Game Outcome. 3rd International Conference on Advances in Computing, 1-7.
- Horvat, T., Job, J., Medved, V., 2018. Prediction of Euroleague Games based on Supervised Classification Algorithm k-Nearest Neighbours. 6th International Congress on Support Sciences Research and Technology Support, 20-21.
- Işık, N., Işık, A. H., 2020. Yapay Sinir Ağları İle Elektrostatik Lens Sistem Tasarımı, Mühendislik Bilimleri ve Tasarım Dergisi, 8(2), 388-396.
- Karaoğlu, B., 2015. Makine Öğrenmesi ile Spor Karşılaşmalarının Modellenmesi. Emo Bilimsel Dergi, 5(9), 1-6.
- Miljković, D., Gajić, L., Kovačević, A., Konjović, Z., 2010. The Use of Data Mining for Basketball Matches Outcomes Prediction. IEEE 8th International Symposium on Intelligent Systems and Informatics, 309-312.
- Nanğır, M., 2013. Türk Dili İçin Çoklu Sınıflandırıcı Yöntemler ile Duygu Sınıflandırma. Yüksek Lisans Tezi, İstanbul Kültür Üniversitesi, İstanbul.
- Oliver, D., 2004. Basketball on Paper: Rules and Tools for Performance Analysis. Potomac Books, Washington.
- Özkan, Y., Erol, Ç.S., 2015. Biyoformatik DNA Mikrodizi: Veri Madenciliği. Papatya Yayıncılık Eğitim, İstanbul.
- Özmen, M.U., 2016. Marginal contribution of game statistics to probability of winning at different levels of competition in basketball: Evidence from the Euroleague. International Journal of Sports Science & Coaching, 11(1), 98-107.
- Sağbaş, E.A., Ballı, S., 2016. Akıllı telefon algılayıcıları ve makine öğrenmesi kullanılarak ulaşım türü tespiti. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 22(5), 376-383.
- Sağbaş, E.A., Ballı, S., 2017. Akıllı saat algılayıcıları ile insan hareketlerinin sınıflandırılması. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 21(3), 980-990.
- Sağbaş, E.A., Korukoglu, S., Ballı, S. 2020. Stress detection via keyboard typing behaviors by using smartphone sensors and machine learning techniques. Journal of Medical Systems, 44(4), 1-12.
- Thabtah, F., Zhang, L., Abdelhamid, N., 2019. NBA Game Result Prediction Using Feature Analysis and Machine Learning. Annals of Data Science, 6(1), 103-116.
- Torres, R.A., 2013. Prediction of NBA games based on Machine Learning Methods. University of Wisconsin Madison.

- Zimmermann, A., 2016. Basketball predictions in the NCAA and NBA: Similarities and differences. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 9(5), 350-364.
- Zimmermann, A., Shi, Z., Moorthy, S., 2013. Predicting NCAA match outcomes using ML techniques – some results and lessons learned. In *ECML/PKDD 2013 Workshop on Machine Learning and Data Mining for Sports Analytics*.
- Türkiye Basketbol Süper Ligi İstatistikleri <http://www.tblstat.net/> (Erişim: 19.03.2020)