

VERİ MADENCİLİĞİ YARDIMIYLA FUTBOL MAÇ TAHMİNLERİNİN GERÇEKLEŐTİRİLMESİ

ALİ AYDOĞAN

aliaydogan20@gmail.com

Özet

Bu çalışmamızın konusu son yıllarda toplum tarafından iyice benimsenen şans ve biraz da istatistiksel bilgi içeren bir oyun olan bahisin oranlarını kullanarak maç tahmini yapmaya çalışmaktır. Amacımız kişi veya kurumlara para kazandırmaktan uzak genel bir veri madenciliğı yöntemi kullanarak istatistiksel veriler üzerinden tahmin yapmaktır. Yaptığımız bu tahminlerin yapılan literatür çalışmalarına göre ve kullandığımız algoritmalarımıza göre başarı oranlarının belirlenmesidir.

1. GİRİŞ

Bilimsel bir alan olarak beden ve spor eğitimi, spor performanslarını geliřtirmek için bilimsel ilke ve teknikleri öngörmektedir. Mevcut veriler üzerinde ilişkiler kurularak maç tahminleri yapmak için yapılan sporun türü, çevre ve oyuncuların hedefleri öngörölerek bir deęerlendirme yapılmaktadır. Bu sonuçları üretmek için birçok yöntem geliştirilmiştir.

Bazı takımlar maçlara çıkarken herhangi bir deęerlendirme yapmazken, bazı takımlar maçı kazanmak için uzmanların yaptığı tarihsel, içgüdüsel ve deneyimsel verilere deęer vererek maçlara çıkmaktadır. Bu verilere güvenerek maça çıkan takımların daha çok başarı elde ettięi görölmektedir.

Son yıllarda veri madencilięinin deęerlendirme alanına giren bu sistem, takımların istatistiksel bilgilerine erişilerek bir deęerlendirme safhasından geçirilerek maç yapacak takımlar arasında maçı kimin kazanacağını da belirlemektedir. Bu sistemin karşılığı günümüzde iddaa olarak yerleşmiş ve insanlar bu sistem üzerinden para kazanmaya çalışmaya başlamışlardır.

2. PROJENİN TANIMI VE AMACI

Günümüzde bahis siteleri maçlar üzerinde yapmış olduğu derin araştırmalar üzerinde genel bir kanıya vararak birbirlerine yakın oranlar vermektedir. Bu oranlar maçı kazanacak takımı belirlemede büyük rol oynamaktadır. Biz de bu oranları ele alarak istatistiksel bu veriler üzerinden skor tahmini ve maçı kimin kazanacağı hususunda tahminler gerçekleştirmeye çalıştık.

Bahis siteleri bu oranları oluştururken birçok özelliği ön planda tutarak bir oranlama sistemi geliştirmiştir. Maç yapacak iki takımın oranlarının belirlenmesinde ev sahibi takımın evinde yaptığı maçlardaki başarısı, deplasman takımının deplasmanda yaptığı maçlardaki başarısı, iki takımın birbiriyle daha önce yaptıkları maçlardaki istatistikler, sakat ve cezalı oyuncular, o takımların o haftaya yakın haftadaki performansları, oynanacak stad vs. birçok özellik göz önünde bulundurularak oranlar oluşturulur.

Bu oranlar maçı kazanacak takımı belirlemede ya da maçın sonuçlarını kestirmede bahisseverlere ipucu olarak verilse de bazen tüm istatistikler doğru olarak sonuçlanmıyor ve sürpriz sonuçlar doğmasına sebep oluyor.

Çalışmamızı gerçekleştirmeden önce bu sürpriz sonuçların doğmasına sebep olan olayları ve daha önce bu konuda literatürde yapılan çalışmaları inceleyerek bir tahminleme gerçekleştirdik.

3. MAKİNE ÖĞRENME YÖNTEMİYLE MAÇ SONUÇLARINI TAHMİN ETME

[1] Albina Yezus, farklı futbol liglerindeki skorları analiz ederek maç sonuçları hakkında tahmin yapmak istemiş ve bunun için ilk önce maça etki eden özellikleri belirlemiş, daha sonra maça etki edecek bir model üretmeye çalışmış ve tahmin için kullanacağı makine yöntemlerini belirlemiştir.

3.1. Makine Öğrenme Algoritmaları

[1] Albina, makine öğrenme algoritmaları olarak SVM, Random Forest, KNN, Logistic Regression gibi algoritmaları belirlemiş, %70 başarıyla işe girişmiş ve amacı bahisten para kazanma gibi bir arzusu varmış.

3.2. Uygulama

Bu çalışmada analiz için veri kümesi oluşturulmuş, anahtar özelliklere karar verilmiş, veri çıkarımı yapılmış, makine öğrenme algortimaları devreye sokulmuş ve son olarak da algoritma geliştirip başarı oranının artırılması sağlanmıştır.

3.3. Analiz ve Veri Çıkarımı

Spor oyunlarında sonucu etkileyecek birçok faktörün ön planda olduğunu belirtmiştik. Ancak Albina, skora etki edecek maçın gidişatını değiştirecek hakem kararları, şike olayları, zor ve imkansız veri çıkarımı gibi faktörleri belirlemiştir.

Verilerini çıkarırken bu özellikleri dikkate alan Albina, takımların maç bilgisini, sezon bilgisini ve her an güncellenen puan tablolarını da hesaba katarak bir veri kümesi oluşturmuştur.

3.4. Anahtar Özellikler

Albina, maçın skoruna etki edecek özellikleri 2 sınıfa ayırmış ve onları dinamik ve statik özellikler olarak adlandırmıştır. Bu özelliklere göre bir çıkarım yaparak formüller üretmiştir.

Albina, statik özellikler olarak form, konsantrasyon, motivasyon gibi özellikleri ele almıştır. Hepsini için ayrı bir formülasyon geliştirmiştir. Bu formülleri geliştirirken, maçın derbi olup olmaması, tur sayıları, takımların her haftadaki maç sonuçları gibi özellikleri dikkate almıştır.

Form için;

$$\frac{1}{10} \sum_{k=1}^5 res_k$$

where res_k - result of the k^{th} match (value in $\{0, 1, 2\}$);

Konsantrasyon için;

$$1 - 2 * x$$

where x - the nearest match lost to the weak team (the difference between the current team and that team ≥ 7);

Motivasyon için;

$$\min\left(\max\left(1 - \frac{dist}{3 * left}, derby, \frac{tour + dist}{2}\right), 1\right)$$

where:

- derby – 1 if match is a derby and 0 otherwise;
- dist – distance to the nearest “key position”;
- left – tours left in the season;
- tour – 1 if left < 6 and 0 otherwise;
- key position – value in $\{1, 2, 3, 4, 5, 6, 17, 18\}$.



Dinamik özellikler olarak gol farkı, skor farkı, tarih gibi özellikler kullanılmıştır.

Gol farkı için;

$$\frac{1}{2} + \frac{dif}{2 * \max_dif}$$

where:

- dif – difference between goals;
- max_dif – maximal difference between goals;

Skor farkı için;

$$\frac{1}{2} + \frac{dif}{2 * \max_dif}$$

where:

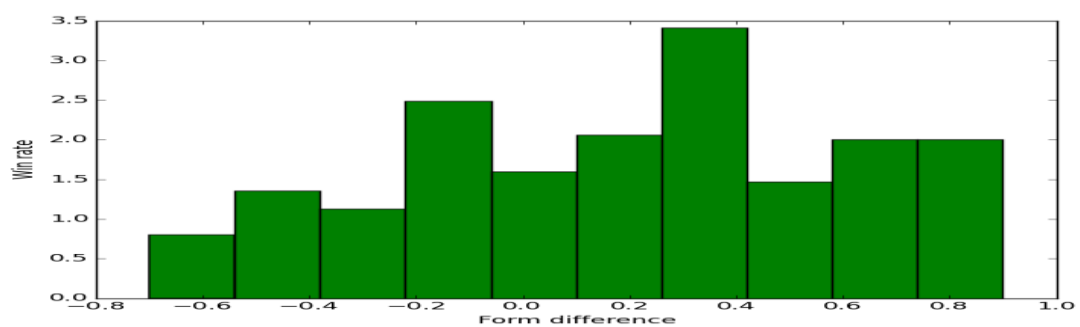
- dif – difference between scores;
- max_dif – maximal difference between goals;

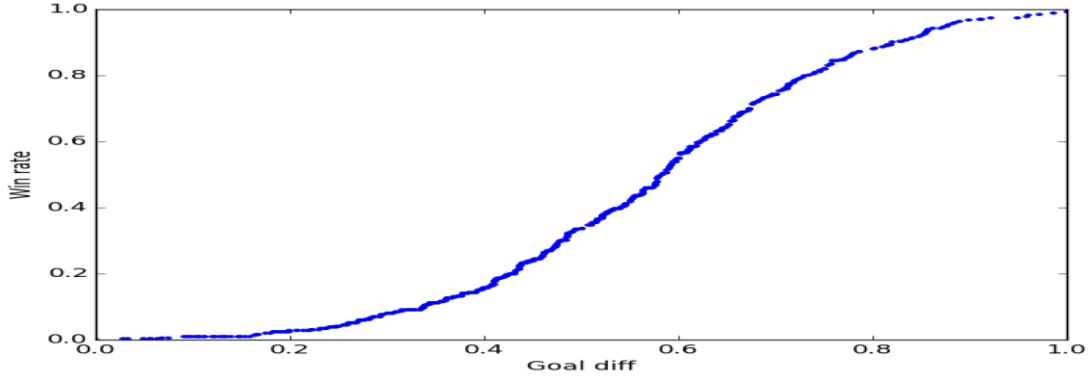
Tarih için;

$$\frac{p1 + p2}{2}$$

where p_i – result of i^{th} match (value in $\{0, 0.5, 1\}$).

Grafiksel olarak elde edilen formülle elde edilen başarılar;





3.5. Başarı Oranları, Materyaller ve Sonuçlar

Bu çalışmada [1]Albina, karar verdiği algoritmalarından 2 tanesini gerçekleştirmiş, KNN ile 0.55, Random Forest ile 0.634 başarı sağlamıştır. Kullandığı 9 özellik ve 640 obje ile İngiltere Prömiyer Ligi üzerinde %60 başarı sağlamıştır. Programlama dili olarak Phyton kütüphanesini kullanmıştır. Veri çıkarımı için Selenium, Urrlib, BeatifulSoup , makine öğrenmesi için SkLearn, veri analizi için Numpy,Pandas, MatPlotLib gibi kütüphaneleri kullanmıştır.

4. VERİ MADENCİLİĞİ YARDIMIYLA SKOR TAHMİNİ

[2] Islamic Azad Üniversitesi'nden birkaç kişinin yapmış olduğu bu çalışmada daha önce yapılan literatür çalışmaları incelenmiş ve bu yapılan çalışmalarda yakalanan en fazla başarı oranları incelenmiştir. Bu çalışma daha çok uygulanan makine öğrenme algoritmalarını tanıtmak olup bunların sonuçlarının karşılaştırılması yapılmıştır.

4.1. Sınıflandırma Teknikleri

Yapılan çalışmalarda genel olarak Yapay Sinir Ağları, SVM, Karar Ağaçları, Bayesian Metodu, Lojistik Regresyon ve Fuzzy Sistemleri kullanılmıştır.

4.1.1. YAPAY SİNİR AĞLARI

Yapay Sinir Ağları, karmaşık sistemleri makine öğrenme yoluyla verilen belirli girişlerden çıkışı tahmin etmek için kullanılan hesaplama yöntemleridir. Bu sistem, biyolojik olan bir sinir ağını taklit ettiği için bu isim verilmiştir. Taklit edilen bu sinir hücreleri nöronlar içerirler ve bu nöronlar çeşitli şekillerde birbirlerine bağlanarak ağı oluştururlar.

[2]Kahn, çok katmanlı algılayıcı bir nöral ağ yapısı kullanarak maç sonuçlarını tahmin etmeye çalıştı. Ağı Error Back-propagation algoritması ile eğitti. Aynı şekilde McCabe ve Travathan çok katmanlı bir yapı kullandı. Bunlar ise eğitim için Conjugate Gradient Metoduyla ağı eğittiler.

4.1.2. SVM

Sınıflandırma konusunda oldukça etkili ve basit olan bu yöntem, bir düzlemde bulunan 2 grup arasında bir sınır çizilerek iki grubu ayırmaya çalışır. Bu sınırın çizileceği yer iki grubun üyelerinin birbirlerine olan en uzak yer olarak seçilmelidir. Bu işlemin yapılması için iki gruba da yakın ve birbirine paralel iki sınır çizgisi çizilir ve bu sınır çizgileri birbirine yaklaştırılarak ortak sınır çizgisi üretilir.

Bu konuda yapılan çalışmada, basketbol maçlarının sonuçları tespit edilmiş, maçı kimin kazanacağı konusunda tahmin yapılmıştır.

4.1.3. KARAR AĞAÇLARI

Karar ağaçları öğrenmesinde, bir ağaç yapısı oluşturularak ağacın yaprakları seviyesinde sınıf etiketleri ve bu yapraklara giden ve başlangıçtan çıkan kollar ile de özellikler üzerindeki işlemler ifade edilmektedir.

Ağacın öğrenilmesi sırasında, üzerinde eğitim yapılan küme, çeşitli özelliklere göre alt kümelere bölünür, bu işlem, özyineli olarak tekrarlanır ve tekrarlama işleminin tahmin üzerinde etkisi kalmayana kadar sürer. Bu işleme özyineli parçalama adı verilir.

[2] Zdravevski ve Kulakov karar ağaçları için belirli özellikler seçerek WEKA üzerinde maç tahminleri üzerinde başarı oranlarını gözlemlemişlerdir.

4.1.4. BAYESİAN METODU

Bu sınıflandırıcıda, sisteme belirli bir oranda öğretilmiş veri sunulur. Öğretim için sunulan verilerin mutlaka bir sınıfı ve kategorisi bulunmalıdır. Öğretilmiş veriler üzerinden yapılan olasılık işlemleri ile, sisteme sunulan yeni test verileri, daha önce elde edilmiş olasılık değerlerine göre işletilir ve verilen test verisinin hangi kategoride olduğu tespit edilmeye çalışılır.

[2]Miljkovic, basketbol maç sonuçlarını tespit etmek için maç sonuçlarını kazanma ve kaybetme olarak kategorilendirerek belirli özelliklere göre tespit etmeye çalışmıştır.

4.1.5 FUZZY SYSTEMS

Bu yapıda, klasik kümelerin aksine bulanık kümelerde elemanların üyelik dereceleri $[0, 1]$ aralığında sonsuz sayıda değişebilir. Bunlar üyeliğin derecelerinin devamlı ve aralıksız bütünüyle bir kümedir. Keskin kümelerdeki soğuk-sıcak, hızlı-yavaş, aydınlık-karanlık gibi ikili değişkenler, bulanık mantıkta biraz soğuk, biraz sıcak, biraz karanlık gibi esnek niteleyicilerle yumuşatılarak gerçek dünyaya benzetilir. En önemli fark, böyle bir çatıda bilginin kaynağındaki küme üyeliğinin kesin tanımlanmış önkoşullarının olmayışı ve daha çok sorunlarla rastgele değişkenlerin hazır bulunmasındadır.

Trawinski ACB ligini tahmin etmek için bulanık bir model geliştirdi. Modeli oluştururken üç aşamalı bir model geliştirdi. Bu aşamalar, veri toplama, veri işleme ve Evrimsel Öğrenmeye dayalı Bilgi Çıkarımı için 10 farklı öğrenme algoritmasının gerçekleştirilmesinden oluşuyordu.

4.2. Sonuçları Değerlendirme

Kahn, Yapay Sinir Ağları ile yapmış olduğu çalışmada Uluslararası Futbol Ligi'nde 2003 yılında tüm sezonunda maç tahmin etmede %75 oranında başarı elde etmiştir. McCabe ve Travathan da bu konuda yaptığı çalışmada Yapay Sinir Ağları ile Amerika Futbol Ligi üzerinde % 65.1, Uluslararası Rugby Ligi'nde %63.2, İngiltere Prömiyer Ligi'nde %54.6, Super Rugby Ligi'nde % 67.5 başarı sağlamıştır.

Miljkovic, k-fold Cross Validation yöntemi ile yaptığı değerlendirmede başarıyı %67 olarak sağlamıştır.

Trawinski, 10-fold Cross Validation yöntemi ile 6 özellik kullandığı bu sistemde bulanık modelleme ile %82 başarı sağlamış ve spor oyunlarında görülen en yüksek başarıyı elde etmiştir.

5. NBA BASKETBOL LİĞİ TAHMİNİ

[3] Bu çalışmada Chenjie Cao, NBA verileri üzerinde inceleme yapmış, NBA'de oynayan 30 takımı iyice incelemiş, takımların 5 farklı bölgede oynayan tüm oyuncularını analiz ederek bir tahminleme gerçekleştirmiştir.

5.1. Daha Önceki Çalışmalar

Bernard Loeffelholz, Earl Bender yaptığı çalışmalarda Yapay Sinir Ağları kullanarak maçı hangi NBA takımının kazanacağı konusunda yaptığı çalışmada 2009 yılındaki çalışmalarıyla %74.33 başarı sağlamışlardır.

Bayesian Logic ve Markov Logic Networks yapılarını kullanan 2 araştırmacı bu algoritmalar ile 2006 senesinde bu lig üzerinde, BLOG ile %63, MLN ile %64 başarı sağlamışlardır.

5.2. Veri Toplama Aşaması

Cao, 2005-2006, 2008-2009 sezonundaki verileri eğitim, 2010-2011 sezonundaki verileri test verisi olarak kullanmış, tüm istatistikleri göz önünde bulundurmuştur.

Model oluşumunda hedef alınan bazı özellikler şunlardır:

- Son NBA sezonunda ev sahibi takımın kazanma oranları
- Son 5 günde ev sahibi takımın attığı sayı
- Son 2 maçta 2 takımın gösterdiği performanslar
- Ev sahibi ve deplasman takımlarının kaybettiği maçlardaki attıkları sayılar
- Ev sahibi ve deplasman takımlarının kaydettikleri 3'lük oranı ve sayısı

Cao, model oluşumunda bu özellikler gibi birçok özelliği incelemiştir.

5.3. Kullandığı Algoritmalar ve Başarı Oranları

Simple Logistics algoritmasını kullanan Cao, bu algoritma için maksimum iterasyon sayısı olarak 500, LogisBoost'un erken durdurulması için Heuristic parametresi 805, LogicBoost ağırlık düzeltme 0.02 olarak belirlemiştir ve başarı oranı %65.82 olarak çıkmıştır.

Cao, Naive Bayes üzerinde gerçekleştirdiği bu çalışmada %65.82 başarı elde etmiştir.

SVM' de ise gamma değerini 0.002, katsayıyı 0.5 olarak belirleyen Cao en yüksek başarısını %67.22 ile bu algoritmayla gerçekleştirmiştir.

Yapay Sinir Ağları'nı WEKA üzerinde deneyen Cao, Back-propagation algoritması için öğrenme oranı 0.1 , momentum oranı 0.05, epoch sayısı 700 olarak seçmiş ve %66.67 başarılı olmuştur.

6. MAÇ SONUÇLARINI TAHMİN ETMEDE YENİ BİR UYGULAMA

Bu çalışmamızda farklı maçların tahminlerini yaparken maç skoru yani ev sahibi takımın attığı gol sayısı ve deplasman takımının attığı gol sayısı ve maçı kimin kazanacağı ya da beraberlik durumunu tespit etmeye çalıştık.

6.1. Veri Kümesi Belirleme

Veri kümemizde maç tahminlerimizi tahmin etmekte kullanacağımız birçok özelliğimiz bulunmaktadır. Veri kümemizin açıklamaları <http://www.football-data.co.uk/notes.txt> belirtilmiştir. Veriler csv formatından farklı liglere ait verileri barındırmaktadır.

Veri kümemizde ligin ismi, maç tarihi, ev sahibi ve deplasman takımı bilgileri, önceki maçlardaki korner, kart sayısı, ofsayt, gol sayısı, köşe vuruşu gibi özellikler bulunmaktadır.

Ayrıca o maç için farklı bahis sitelerinin ev sahibi kazanır, deplasman kazanır, beraberlik gibi oranları verilmiştir.

6.2. Özellik Seçimi ve Veri Azaltma

Veri kümemiz sınıfımızın durumunu etkileyen en önemli unsurlardan oluşması gerektiğinden sınıfımıza etkisi olmayan gereksiz, gürültülü verileri veri setimizden çıkarmamız gerekmektedir. Bu nedenle maçın oynanma tarihi, ligin isimleri gibi veriler veri kümemizden çıkarılmıştır.

Ayrıca başarımızı etkileyen bazı değerler deneme-yanılma yöntemleriyle olup olmaması durumunda başarı oranları test edilip başarı oranımızı düşüren bu özellikler veri kümemizden çıkarılmıştır.

6.3. Sınıflandırma Teknikleri

Yapılan diğer literatür çalışmalarında Naive Bayes, Karar Ağaçları, Lojistik Regresyon, KNN, Random Forest, Bulanık Mantık gibi yapıları kullanmışlar ve test etmişlerdir. Biz de bu çalışmamızda KNN algoritmasını kullandık ve bu gerçekleştirdiğimiz algoritmayı diğer algoritmalarla karşılaştırdık.

6.3.1 KNN

Bu algoritmaya göre sınıflandırma sırasında çıkarılan özelliklerden, sınıflandırılmak istenen yeni bireyin daha önceki bireylerden k tanesine yakınlığına bakılır. Eğer k değeri çok küçük seçilirse oluşturacağımız model çok etkilenir. Çok büyük olursa da tek bir sınıf gibi olur. Bu nedenle k değerinin tam manasıyla uygun bir seçimi gerçekleştirilmelidir.

- İlk olarak K değerimizi belirledik.(farklı k değerleri için sonuçlar ürettik)
- Her veri için Öklid uzaklıkları hesaplandı.
- Bu değerlerden minimum uzaklıktaki en yakın k tane komşu bulundu.
- Bu uzaklıklara bağlı olarak en çok gelen kategori belirlendi.
- Ve o kategori sınıf etiketi olarak belirlendi.
- Bu algoritma her yeni ekleme için tekrar Öklid uzaklığı hesapladığı için dezavantajlıdır.

6.3.2 NAIVE BAYES

Daha önce Gaussian Naive Bayes nümerik değerler için yapmış olduğumuz çalışmanın sonuçları KNN algoritmamız ile karşılaştırılmıştır.

$$p(x_k|C_i) = g(x_k, \mu_{C_i}, \sigma_{C_i})$$
$$g(x_k, \mu_{C_i}, \sigma_{C_i}) = \frac{1}{\sigma_{C_i} \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x_k - \mu_{C_i}}{\sigma_{C_i}} \right)^2}$$

6.3.3 RANDOM FOREST

CSV dosyalarımızla WEKA üzerinde yapmış olduğumuz Random Forest ile incelemeyi KNN algoritmamız ile karşılaştırdık.

6.4. K Değerleri ve Farklı Liglerle Yapılan Çalışmalar

KNN algoritmamız için en uygun k değerini belirlemek için farklı k değerleriyle sonuçları gözlemledik. Ayrıca Naive Bayes ve Random Forest ile sonuçlarını karşılaştırdık.

İspanya Ligi'nde 20 takımın 4 sezonunu ele alarak gerçekleştirdiğimiz bu çalışmada $380 \times 4 = 1520$ satır ve 1 satırda 65 tane özelliği kullanarak yaptığımız bu çalışmada ilk 3 sezonu eğitim ve son sezonu test olarak kullandığımız bu çalışmada elde ettiğimiz sonuçlar tablodaki gibidir:

	Ev Sahibi Gol Sayısı Accuracy	Deplasman Takımı Gol Sayısı Accuracy	Maçı Kim Kazanır Ya da Beraber
K=8	%26.49	%28.25	%46.3
K=10	%27.33	%28.29	%47.5
K=12	%28.4	%29.12	%48.28
K=14	%27.25	%28.33	%47.48
Naive Bayes	?	?	%51.34
RandomForest(Weka)	?	?	%52.53

K=12 için farklı ligler üzerinde yaptığımız çalışmalarda en başarılı lig İspanya ligi olmuştur. İtalya, Fransa, Almanya, İngiltere liglerinde de yaptığımız bu çalışmada en az başarıyı maçı kimin kazanacağı tahminini yapmada %43 gibi bir oranla Fransa ligi olmuştur.

WEKA üzerinde yaptığımız çalışmada Random Forest için ağaç sayısını 100 olarak belirledik ve özellik seçimini de elimizdeki özellik sayısının karekökü yani elimizdeki 64 özelliğin karekökü 8 olarak seçtik

6.5. K-Cross Validation

Bu metrikte, yaptığımız çalışmada tüm sezonları test olarak kullanıp geri kalanları eğitim için kullanarak çalışmamızı doğrulamaya çalıştık.

K=12	Maçı Kim Kazanır Ya da Beraber
1. Sezon Test	%46.02
2. Sezon Test	%45.72
3. Sezon Test	%46.35
4. Sezon Test	%48.28

6.6 Bulgular ve Sonular

Yaptığımız bu alıřmada oynanan malarda ortaya ıkan hakem hataları, takımların oyuncularının oyundan atılmaları, sakatlıklar, ma esnasında hava kořulları, teknik direktörlerin verdikleri kararlar maın sonucunu doğrudan etkilediğinden ma skorunda sürpriz sonuçlar yaşanmasına sebep olmaktadır.

Analiz ve istatistikleri yanıltan bu sürpriz gelişmeler başarımızın düşmesine sebebiyet vermektedir. Yapılan literatür alıřmalarını da incelediğimizde %60 oranlarında başarı sağlanmıştır. Bu alıřmamızda uyguladığımız KNN algoritması %45 lerde başarı sağlamıştır. Ancak başka bir algoritma kullanarak, daha ok veri kullanarak, kilit özellikler eklenerek bu başarı oranı artırılabilir. alıřmalarımızı daha iyi bir algoritma ile ilerleyen zamanlarda deneyeceğiz.

KNN algoritmamız en yakın komřulara baktığından Real Madrid, Barcelona gibi takımların ok fazla gol attığı malarda sonuçlarımız az gol attığını öngörmektedir. ünkü Barcelona ve Real Madrid gibi takımlar her ma 4-5 gol atmıyorlar. Ayrıca bazı malarda bazı takımların hiç gol atmadığı malarda 3 -4 gol attığı görölmektedir. Bu tamamen verilen oranlarda bulunan en yakın komřulardaki bulunan kategorilerdeki gol sayılarıyla alakalı bir durumdur.

Veri kümemize ek olarak, hava kořulları, hakemin ismi, teknik direktörlerin ismi gibi birçok özellik eklenerek yeni ıkarımlar eklenebilir. Gerçekten bu alan ok analize ve istatistiksel bilgiye dayanan birçok bahis oynayanların ilgisini çekebilecek bir alandır.

KAYNAKLAR

- 1- Saint-Petersburg State University Mathematics and Mechanics Faculty, Albina Yezus
Scientific Adviser: Alexander Igoshkin, Yandex Mobile Department
“Predicting outcome of soccer matches using machine learning”
- 2- Maral Haghighat, Hamid Rastegari and Nasim Nourafza Faculty of Computer
Engineering, Najafabad Branch, Islamic, Azad University, Najafabad, Iran
“A Review of Data Mining Techniques for Result Prediction in Sports”
- 3- Chenjie Cao , Dublin Institute of Technology
“Sports Data Mining Technology Used in Basketball Outcome Prediction”