

DOKUZ EYLÜL ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

DENETİMLİ MAKİNE ÖĞRENME
YÖNTEMLERİ İLE GENÇ FUTBOL
OYUNCULARININ POTANSİYEL
PERFORMANSLARININ MODELLENMESİ VE
TAHMİNİ

Burak AL

Ekim, 2025

İZMİR

**DENETİMLİ MAKİNE ÖĞRENME
YÖNTEMLERİ İLE GENÇ FUTBOL
OYUNCULARININ POTANSİYEL
PERFORMANSLARININ MODELLENMESİ VE
TAHMİNİ**

Dokuz Eylül Üniversitesi Fen Bilimleri Enstitüsü

Yüksek Lisans Tezi

İstatistik Anabilim Dalı, Veri Bilimi Programı

Burak AL

Ekim, 2025

İZMİR

YÜKSEK LİSANS TEZİ SINAV SONUÇ FORMU

BURAK AL tarafından **PROF.DR. ÖZLEM EGE ORUÇ** yönetiminde hazırlanan “**DENETİMLİ MAKİNE ÖĞRENME YÖNTEMLERİ İLE GENÇ FUTBOL OYUNCULARININ POTANSİYEL PERFORMANSLARININ MODELLENMESİ VE TAHMİNİ**” başlıklı tez tarafımızdan okunmuş, kapsamı ve niteliği açısından bir Yüksek Lisans tezi olarak kabul edilmiştir.

Prof.Dr. Özlem Ege ORUÇ

Danışman

Prof.Dr. Sevcan DEMİR ATALAY

Doç.Dr. Özgül VUPA ÇİLENGİROĞLU

Jüri Üyesi

Jüri Üyesi

Prof. Dr. Abdullah SEÇGİN

Müdür

Fen Bilimleri Enstitüsü

TEŞEKKÜR

Danışman hocam Prof.Dr. Özlem EGE ORUÇ'a tez çalışmam boyunca gösterdiği sabır, değerli katkılar ve akademik rehberlik için en derin şükranlarımı ifade etmek isterim. Bilgi birikimi ve yönlendirmeleri olmasaydı, bu çalışmayı tamamlamak mümkün olmazdı.

Son olarak, hayatımın her aşamasında yanımda olan, bana inanan ve sonsuz desteklerini hissettiren aileme ve arkadaşlarıma en içten teşekkürlerimi sunarım.

Burak AL

DENETİMLİ MAKİNE ÖĞRENME YÖNTEMLERİ İLE GENÇ FUTBOL OYUNCULARININ POTANSİYEL PERFORMANSLARININ MODELLENMESİ VE TAHMİNİ

ÖZET

Futbol, yalnızca sportif bir rekabet alanı olmanın ötesinde, günümüzde küresel ölçekte ekonomik, sosyo-kültürel ve ticari bir endüstri haline gelmiştir. Rekabetin yoğun olduğu bu ortamda, kulüplerin uzun vadeli başarı ve sürdürülebilirlik hedefleri doğrultusunda veri temelli stratejiler geliştirmesi kritik önem taşımaktadır. Son yıllarda veri bilimi, futbol endüstrisinde oyuncu performans analizi, yetenek keşfi ve transfer politikalarının şekillendirilmesinde güçlü bir karar destek aracı olarak öne çıkmaktadır.

Bu çalışmada, Football Manager 2023 veri tabanından elde edilen 360 profesyonel genç futbolcuya ait teknik, fiziksel ve zihinsel özellikler kullanılarak oyuncu potansiyel aralıklarının tahmin edilmesi amaçlanmıştır. Bu doğrultuda, denetimli makine öğrenimi yaklaşımlarından Support Vector Machine, Random Forest, XGBoost ve K-Nearest Neighbors algoritmaları uygulanmış; modellerin başarımı Doğruluk (Accuracy), Kappa ve %95 CI metrikleri üzerinden karşılaştırmalı olarak değerlendirilmiştir.

Elde edilen bulgular, Random Forest algoritmasının %87.85 doğruluk oranıyla en yüksek performansı sergilediğini ortaya koymuştur. Özellik önem sıralamaları, teknik beceri, çeviklik ve hız gibi değişkenlerin oyuncu potansiyelinin belirlenmesinde kritik rol oynadığını göstermektedir. Çalışma, futbol kulüplerinin genç yeteneklerin tespitinde ve transfer stratejilerinin planlanmasında veri bilimi tabanlı sistemlerin sağlayabileceği stratejik avantajları vurgulamaktadır.

Anahtar kelimeler: Futbol, Makine öğrenmesi, Rassal ormanlar, XGBoost, K-en yakın komşu, Destek vektör makineleri

MODELING AND PREDICTION OF POTENTIAL PERFORMANCE OF YOUNG FOOTBALL PLAYERS USING SUPERVISED MACHINE LEARNING METHODS

ABSTRACT

Football has transcended its role as merely a sphere of sporting competition to become a global economic, socio-cultural, and commercial industry. In this highly competitive environment, it is critical for clubs to develop data-driven strategies to achieve long-term success and sustainability. In recent years, data science has emerged as a powerful decision-support tool in the football industry for player performance analysis, talent identification, and shaping transfer policies.

This study aims to predict the potential range of players using technical, physical, and mental attributes of 360 professional young footballers obtained from the Football Manager 2023 database. To this end, supervised machine learning approaches namely Support Vector Machine, Random Forest, XGBoost, and K-Nearest Neighbors algorithms were implemented. The performance of these models was evaluated and compared based on Accuracy, Kappa, and 95% CI metrics.

The findings reveal that the Random Forest algorithm achieved the highest performance with an accuracy rate of 87.85%. Feature importance rankings indicate that variables such as technical skill, agility, and pace a critical role in determining a player's potential. This study underscores the strategic advantages that data science-based systems can provide to football clubs in scouting young talents and planning transfer strategies.

Keywords: Football, Machine learning, Random forest, XGBoost, K-nearest neighbors, Support vector machine

İÇİNDEKİLER

	Sayfa
YÜKSEK LİSANS TEZİ SINAV SONUÇ FORMU	ii
TEŞEKKÜR.....	iii
ÖZET	iv
ABSTRACT.....	v
İÇİNDEKİLER	vi
ŞEKİLLER LİSTESİ	viii
TABLolar LİSTESİ	ix
BÖLÜM BİR - GİRİŞ.....	1
1.1. Çalışma Alanı	1
1.2. Literatür Taraması	1
1.3. Amaç	3
BÖLÜM İKİ - YÖNTEMLER.....	6
2.1. Çok Ölçütlü Fayda Teorisi (MAUT)	6
2.1.1. MAUT'un Literatürdeki Yeri ve Tarihsel Gelişimi	6
2.1.2. Entropi.....	7
2.1.3. MAUT Adımları	8
2.2. Makine Öğrenmesi Yöntemleri	10
2.2.1. Rassal Ormanlar (Random Forest).....	11
2.2.2. XGBoost(Extreme Gradient Boosting)	12
2.2.3. Destek Vektör Makineleri (Support Vector Machine).....	12
2.2.4. K-En Yakın Komşu (K-Nearest Neighbors(k-NN))	12
2.3. Hata Metrikleri	13
2.3.1. Doğruluk (Accuracy).....	14
2.3.2. Kappa İstatistiği (Cohen's Kappa)	14
2.3.3. 95% Güven Aralığı (Confidence Interval - CI).....	15
BÖLÜM ÜÇ - UYGULAMA	16
3.1. Veri	16

3.1.1.	Değişkenler.....	19
3.2.	MAUT Uygulaması.....	22
3.2.1.	Defans (D).....	22
3.2.2.	Defansif Bek (DB)	25
3.2.3.	Kanat Bek (WB).....	27
3.2.4.	Defansif Orta Saha (DM)	29
3.2.5.	Merkez Orta Saha (MC).....	31
3.2.6.	Kanat (AM)	32
3.2.7.	Ofansif Merkez Orta Saha (AMC)	34
3.2.8.	Santrafor (ST).....	36
3.3.	Makine Öğrenmesi Uygulaması.....	38
3.3.1.	Rassal Ormanlar (Random Forest).....	40
3.3.2.	XGBoost(Extreme Gradient Boosting)	40
3.3.3.	Destek Vektör Makineleri (Support Vector Machine)	41
3.3.4.	En Yakın Komşu (k-NN).....	41
BÖLÜM DÖRT - SONUÇ.....		43
KAYNAKLAR.....		45
EK 1: MAUT YÖNTEMİ “DM” MEVKİSİ İÇİN PYTHON KODLARI		49

ŞEKİLLER LİSTESİ

Sayfa

Şekil 3.1 Bağımlı değişkenin frekans dağılımı	39
---	----



TABLÖLER LİSTESİ

Sayfa

Tablo 2.1 Makine öğrenmesi yöntemlerinin karşılaştırmalı özellikleri	13
Tablo 3.1 Veri setinden seçilmiş örnek 5 satır verisi.....	19
Tablo 3.2 Defans oyuncularını için karar matrisi.....	23
Tablo 3.3 Defans oyuncularını için standartlaştırılmış karar matrisi	24
Tablo 3.4 Defansif bek mevkiisi için karar matrisi	26
Tablo 3.5 Standartlaştırılmış defansif bek karar matrisi	26
Tablo 3.6 Kanat bek için karar matrisi	27
Tablo 3.7 Kanat bek için standartlaştırılmış karar matrisi	28
Tablo 3.8 Defansif orta saha mevkiisi için karar matrisi	29
Tablo 3.9 Standartlaştırılmış defansif orta saha karar matrisi.....	30
Tablo 3.10 Merkez orta saha karar matrisi.....	31
Tablo 3.11 Standartlaştırılmış merkez orta saha karar matrisi	32
Tablo 3.12 Kanat mevkiisi için karar matrisi	33
Tablo 3.13 Standartlaştırılmış kanat mevkiisi karar matrisi.....	33
Tablo 3.14 Ofansif merkez orta saha mevkiisi karar matrisi	35
Tablo 3.15 Standartlaştırılmış ofansif merkez orta saha karar matrisi.....	35
Tablo 3.16 Forvet mevkiisi için karar matrisi	36
Tablo 3.17 Standartlaştırılmış forvet mevkiisi karar matrisi.....	37
Tablo 4.1 Makine öğrenmesi yöntemlerinin performans karşılaştırması.....	43

BÖLÜM BİR

GİRİŞ

1.1. Çalışma Alanı

Günümüzde veri, karar verme süreçlerinde ve rekabetçi piyasalarda geçmiş dönemlere kıyasla çok daha stratejik bir unsur haline gelmiştir. Özellikle spor endüstrisi, sahip olduğu ekonomik büyüklük ve geniş kitlelere hitap etme gücü nedeniyle veriye dayalı analizlerin giderek daha önemli hale geldiği bir alan olarak öne çıkmaktadır. Küresel futbol endüstrisinin 2025 yılı itibarıyla 600 milyar ABD dolarını aşan bir ciroya ulaştığı ve 3,5 milyardan fazla kişiye hizmet sunduğu öngörülmektedir (Deloitte, 2024; FIFA Big Count, 2023).

Bu denli büyük bir ekonomik hacme sahip sektörde futbol kulüpleri ve oyuncular arasındaki rekabet her geçen gün artmakta; bu durum, karar alıcıların üzerindeki sorumluluk ve baskıyı da beraberinde getirmektedir. Yanlış alınan kararlar, kulüpler açısından kimi zaman birkaç yüz bin dolar, kimi zaman ise milyonlarca dolar düzeyinde maliyetler doğurabilmektedir. Benzer şekilde, hatalı kararlar futbolcuların kariyer gelişimlerini olumsuz etkileyerek profesyonel hedeflerine ulaşma süreçlerini geciktirebilmektedir. Dolayısıyla futbol endüstrisinde verinin doğru, sistematik ve etkin biçimde kullanımı hem kulüpler hem de futbolcular açısından kritik önem taşımaktadır.

1.2. Literatür Taraması

Futbol analitiği literatüründe son yıllarda önemli gelişmeler kaydedilmiş olmakla birlikte, genç futbolcuların geleceğe yönelik potansiyel performanslarının öngörülmesine ilişkin çalışmaların sayısı oldukça sınırlıdır. Oysa kulüplerin rekabet avantajı elde edebilmesi, genç oyuncuların gelişim eğilimlerinin doğru biçimde tahmin edilmesine ve buna bağlı olarak yetenek keşfi, altyapı planlaması ile transfer stratejilerinin bilimsel temellere dayandırılmasına bağlıdır. Bu durum, futbol endüstrisinde ileri düzey veri analitiği yaklaşımlarına duyulan ihtiyacı açık biçimde ortaya koymaktadır.

Mevcut literatür, veriye dayalı modelleme ve makine öğrenmesi yöntemlerinin futbol performans analizinde yaygın biçimde kullanıldığını göstermektedir. Ancak bu çalışmaların büyük çoğunluğu geçmiş performans değerlendirmelerine odaklanmakta, geleceğe dönük potansiyel öngörülerine yeterince yer vermemektedir. Bu bağlamda geliştirilen araştırma yaklaşımı, metodolojik bütünlüğü, yenilikçi veri kaynağı kullanımı ve stratejik öngörü perspektifiyle literatürdeki benzer çalışmalardan anlamlı biçimde farklılaşmaktadır.

Van Wijk (2022) tarafından gerçekleştirilen çalışma, genç futbolcuların potansiyelini tahmin etmeye yönelik olarak çeşitli makine öğrenmesi algoritmalarını karşılaştırmıştır. Ancak araştırma yalnızca maç performans verilerine dayanmaktadır. Bu kapsamda önerilen modelde ise, Football Manager veritabanında yer alan uzman değerlendirmelerinden elde edilen çok boyutlu özniteliklerden yararlanılmış ve mevkiye özgü değişkenlerin MAUT (Multi-Attribute Utility Theory) yöntemi ile seçildiği özgün bir yapı oluşturulmuştur. Böylelikle modelin yalnızca veriyi işlemesi değil, aynı zamanda hangi değişkenlerin analize dahil edileceğinin nesnel biçimde belirlenmesi sağlanmıştır.

Pappalardo ve arkadaşları (2019) tarafından geliştirilen PlayerRank modeli, spatio-temporal (uzamsal-zamansal) veriler aracılığıyla futbolcu performansını değerlendiren en gelişmiş modellerden biridir. Bununla birlikte söz konusu çalışma, geçmiş performansın analiziyle sınırlı kalmaktadır. Buna karşılık burada geliştirilen yaklaşım, geleceğe yönelik potansiyel tahminine odaklanmakta; Football Manager verisini ölçeklenebilir, erişilebilir ve uygulamaya elverişli bir veri kaynağı olarak kullanmaktadır. Ayrıca MAUT tabanlı öznitelik seçimi, modelin yorumlanabilirliğini artırarak futbol uzmanlarının karar süreçlerinde kullanılabilir bir yapı sunmaktadır.

Sayan (2023) tarafından yürütülen araştırmada futbol simülasyon oyunu verilerinin geçerliliği incelenmiş ve bu veriler aracılığıyla mevcut performans ya da pazar değeri gibi anlık metriklerin tahmini amaçlanmıştır. Buna karşılık burada benimsenen yöntem, tamamen “gelecek potansiyeli”nin tahminine yönelmekte; değişkenlerin nicel ağırlıklandırmasının MAUT yöntemiyle gerçekleştirilmesi suretiyle sürecin daha nesnel ve ölçülebilir bir hale getirilmesi amaçlanmaktadır. Benzer biçimde Chen (2023), futbolcuları mevkiler yerine fonksiyonel rollere göre sınıflandıran tanımlayıcı

bir analiz sunmuştur. Buradaki yaklaşım ise tahmine dayalı bir çerçeve önererek “bir oyuncu gelecekte hangi potansiyel seviyesinde yer alabilir?” sorusuna yanıt aramakta ve Chen’in betimleyici analizini tamamlayıcı nitelikte bir perspektif ortaya koymaktadır.

Literatürde Al-Asadi (2018) ve Topcu (2023) tarafından yürütülen çalışmalar, futbolcuların mevcut performanslarının veya maç sonuçlarının tahminine odaklanmıştır. Buna karşın burada geliştirilen model, geleceğe dönük potansiyel performansın öngörülmesine odaklanarak literatürde hâkim olan geçmiş veya mevcut durum temelli yaklaşımların ötesine geçmektedir. Böylelikle kulüplerin sürdürülebilir rekabet avantajı elde etmesine olanak sağlayacak stratejik bir karar destek modeli oluşturulmuştur. Yıldız ve Yıldız (2019) tarafından MAUT yönteminin futbolcu sıralaması amacıyla kullanıldığı görülmektedir. Bu çalışmada ise yöntem, sıralama aracı olmaktan çıkarılarak makine öğrenmesi modellerinde öznitelik seçimi sürecine entegre edilmiştir. Benzer biçimde, Boon ve Sierksma (2010) tarafından doğrusal programlama yaklaşımıyla geliştirilen futbolcu seçimi modeli, bütçe kısıtlı takım optimizasyonuna odaklanmıştır. Buna karşın burada önerilen yapı, aynı problemi tahmin ve sınıflandırma bağlamında ele almakta ve tek bir oyuncunun gelecekteki potansiyel performansının öngörülmesini hedeflemektedir.

Sonuç olarak, mevcut literatürdeki güçlü yönleri bir araya getiren bu araştırma, metodolojik açıdan derinlikli, yenilikçi ve uygulamaya dönük bir model geliştirmektedir. Football Manager verisinin geçerliliğinin incelenmesi (Sayan, 2023), MAUT yönteminin öznitelik seçimi sürecinde kullanılması (Yıldız & Yıldız, 2019) ve farklı makine öğrenmesi algoritmalarının karşılaştırılması (Van Wijk, 2022) gibi unsurların bütüncül biçimde bir araya getirilmesiyle oluşturulan bu yapı, hem akademik literatüre katkı sağlamakta hem de futbol endüstrisinde karar destek sistemlerinin geliştirilmesine pratik düzeyde fayda sunmaktadır.

1.3. Amaç

Bu tez çalışmasının temel amacı, mevcut literatürdeki boşluğu gidermek üzere 21 yaş ve altı futbolcuların gelecekte ulaşmaları beklenen potansiyel yetenek düzeylerini tahmin edebilecek bir model geliştirmektir. Çalışma, makine öğrenmesi

algoritmalarını Çok Kriterli Karar Verme (ÇKKV) yöntemlerinden Çok Ölçütlü Fayda Teorisi (MAUT) ile bütünleştirerek futbol endüstrisine hem metodolojik hem de pratik katkılar sağlamayı hedeflemektedir.

Bu kapsamda, farklı makine öğrenmesi algoritmalarının genç futbolcuların potansiyel performanslarını öngörmedeki başarı düzeyleri karşılaştırılmakta; ayrıca MAUT tabanlı öznelik seçiminin, literatürde yaygın biçimde kullanılan genel öznelik setlerine kıyasla tahmin performansını artırıp artırmadığı araştırılmaktadır. Elde edilen sonuçlar, futbol endüstrisinde karar alıcılar açısından pratik uygulanabilirlik ve yorumlanabilirlik boyutlarıyla değerlendirilmektedir.

Tez çalışmasının literatüre ve uygulamaya dört temel katkı sunacağı öngörülmektedir.

Öncelikle, futbol analitiğinde görece az ele alınmış bir konu olan genç futbolcuların gelecek odaklı potansiyel performanslarının tahminine yönelik bütüncül bir metodolojik çerçeve önerilmektedir. Bu doğrultuda 21 yaş altı oyuncular için farklı makine öğrenmesi algoritmalarına dayalı tahmin modelleri geliştirilecek, bu modellerin doğruluk, güvenilirlik ve genellenebilirlik performansları sistematik biçimde karşılaştırılacaktır.

Buna ek olarak, MAUT yöntemi kullanılarak elit düzeydeki futbolcuların mevki bazında sahip oldukları öznelikler sistematik biçimde değerlendirilmektedir. Bu yaklaşım, özneliklerin görece önem derecelerini niceliksel olarak ortaya koymakta; böylece belirli bir mevki için hangi niteliklerin temel belirleyici olduğunu ve performans farklılaşmasında hangi faktörlerin daha kritik bir rol oynadığını istatistiksel temelde analiz etme olanağı sunmaktadır.

Ayrıca, tahmin modellerinde kullanılacak değişkenlerin belirlenmesinde MAUT tabanlı öznelik seçimi uygulanarak her mevkiye özgü en kritik değişkenlerin sistematik biçimde saptanması sağlanmaktadır. Bu değişkenler makine öğrenmesi algoritmalarına entegre edilerek farklı model yapılandırmalarının performansları karşılaştırılacaktır. Böylelikle genel öznelik setleri yerine mevkiye özgü

değişkenlerin kullanılması, modellerin hem güvenilirliğini hem de uygulama bağlamına uygunluğunu güçlendirmektedir.

Son olarak, geliştirilen modelin çıktıları futbol endüstrisinde yaygın biçimde kullanılan 0–200 arası potansiyel reyting skoru biçiminde sunularak, araştırmanın bulguları akademik katkının ötesine taşınmakta; kulüplerin yetenek keşfi, transfer kararları ve stratejik planlama süreçlerinde doğrudan kullanılabilir bir karar destek aracı niteliği kazanmaktadır.

Genel olarak, bu tez çalışması metodolojik yenilikleri, sektörel uygulanabilirliği ve gelecek odaklı yaklaşımıyla futbol yönetiminde veriye dayalı stratejik planlamanın sınırlarını genişletmeyi amaçlamaktadır.

BÖLÜM İKİ

YÖNTEMLER

Bu bölümde, araştırmanın kuramsal temelini oluşturan iki ana yapı ele alınmaktadır: Çok Kriterli Karar Verme (ÇKKV) tekniklerinden Çok Ölçütlü Fayda Teorisi (MAUT) ile çeşitli Makine Öğrenmesi (ML) algoritmaları. İlk olarak, mevkiye özgü değişken seçiminin nesnel ve sistematik bir biçimde gerçekleştirilebilmesi amacıyla kullanılan MAUT'un tarihsel gelişimi, matematiksel temelleri ve uygulama adımları literatür doğrultusunda incelenmektedir. Daha sonra, potansiyel yetenek tahmininde kullanılan Destek Vektör Makineleri (SVM), Rastgele Orman (RF), k-En Yakın Komşu (KNN) ve Aşırı Gradyan Artırma (XGBoost) algoritmalarının temel işleyiş mekanizmaları, avantajları, dezavantajları ve ilgili formülasyonları açıklanmaktadır. Bu inceleme ile, sonraki bölümlerde gerçekleştirilecek uygulamaya yöntemsel bir temel oluşturulması ve izlenen yaklaşımın bilimsel çerçevesinin netleştirilmesi amaçlanmaktadır.

2.1. Çok Ölçütlü Fayda Teorisi (MAUT)

Çok Kriterli Karar Verme (ÇKKV) – İngilizce literatürdeki adıyla Multiple Criteria Decision Making (MCDM) – birden fazla ve çoğu zaman birbiriyle çelişen hedefin (kriterin) aynı anda dikkate alındığı karar problemlerinin çözümüne yönelik analitik yaklaşımlar bütünüdür (Zionts, 1979). Bu disiplin, karar vericilere tek bir optimal çözüme ulaşmaktan ziyade mevcut kısıtlar ve hedefler doğrultusunda en tatmin edici dengeyi kurabilecek sistematik bir çerçeve sunmaktadır.

2.1.1. MAUT'un Literatürdeki Yeri ve Tarihsel Gelişimi

ÇKKV literatüründe en yaygın kullanılan yöntemlerden biri olan MAUT, ilk kez Fishburn (1967) ve Keeney (1974) tarafından sistematik olarak ortaya konulmuş, sonrasında ise Loken (2007) gibi araştırmacıların katkılarıyla geliştirilmiştir. Özellikle son yıllarda, karmaşık karar problemlerine getirdiği sistematik yaklaşım nedeniyle hem akademik çalışmalarda hem de uygulamalı alanlarda sıklıkla tercih edilmektedir.

MAUT'un en önemli üstünlüklerinden biri, nicel ve nitel kriterleri aynı analitik yapı içerisinde bütünleştirebilmesidir. Böylece, fiyat, mesafe veya ağırlık gibi nesnel olarak ölçülebilen değişkenlerin yanı sıra; estetik, prestij veya kullanışlılık gibi öznel nitelikteki faktörler de analize dâhil edilebilmektedir. Öznel kriterlerin ölçülebilir hale getirilmesi için ise literatürde yaygın biçimde kullanılan dereceli ölçekler (ör. 1–5 arası: çok kötü–çok iyi) veya 0–100 arası puanlama sistemleri tercih edilmektedir (Von Winterfeldt & Edwards, 1986). Bu ölçekler, karar vericilerin yargılarının tutarlı bir şekilde sayısallaştırılmasına ve farklı kriterler arasında karşılaştırılabilirlik sağlanmasına olanak tanımaktadır.

MAUT uygulamalarının temel aşamalarından biri, karar matrisinin oluşturulması ve kriter ağırlıklarının belirlenmesidir. Bu noktada, kriterlerin nesnel biçimde ağırlıklandırılmasında entropi yöntemi kullanılmaktadır. Entropi, bir veri setindeki belirsizlik ölçüsünü veren kavram olup, bir kriterin alternatifler arasındaki ayırt edici gücünü yansıtmaktadır (Canbolat, 2007). Bu çalışmada da kriter ağırlıkları entropi kullanılarak belirlenecektir. Bu nedenle, MAUT'un uygulama adımlarına geçmeden önce entropi kavramı kısaca tanıtılacaktır.

2.1.2. Entropi

Entropi, temel olarak bir sistemdeki düzensizlik, rastgelelik veya belirsizlik düzeyinin nicel bir ölçüsüdür. Farklı disiplinlerde çeşitli tanımlar ve yaklaşımlar geliştirilmiş olmakla birlikte, bu çalışmada kriter ağırlıklarının nesnel olarak belirlenmesinde enformasyon teorisi temelli entropi yaklaşımı kullanılacaktır; bu nedenle diğer disiplinlere ait tanım ve uygulamalar çalışma kapsamı dışında bırakılmıştır. Enformasyon teorisinde entropi, bir veri kaynağındaki belirsizliğin veya beklenen bilginin ölçülmesinde kullanılan matematiksel bir araç olarak, karar verme süreçlerinde kriterlerin göreceli önem düzeylerinin belirlenmesinde etkin bir biçimde uygulanmaktadır.

Entropi değeri, sistemdeki düzensizlik arttıkça yükselir ve bu durum, sistemin öngörülebilirliğinin azaldığını göstermektedir. Bir sistemin olası durumlarını kesin biçimde belirlemek için gerekli olan bilgi miktarı, sistemin entropi değeri ile ilişkilidir. Sistemin düzenli ve organize bir yapıdan düzensiz ve rastgele bir duruma geçişi,

entropi değeri bir artışa ve buna paralel olarak tanımlamak için gereken bilgi miktarında bir yükselişe işaret etmektedir. Bu bağlamda, bilgi kavramı negatif entropi (negentropi) olarak da yorumlanabilmektedir.

Bilgi teorisi çerçevesinde entropi, bir rassal değişkenin belirsizliğinin veya beklenen bilginin nicel ölçüsüdür. Örneğin, $x_1, \dots, x_r$ değerlerini alabilen kesikli bir X rassal değişkeninin bilgi entropisi $H(X)$, aşağıdaki formül ile hesaplanır:

$$H(X) = - \sum_{i=1}^n p(X_i) \times \log_2 p(X_i) \quad (2.1)$$

X rassal değişkeni sürekli bir değişken ise entropi;

$$H(X) = - \int_{-\infty}^{\infty} f(X_i) \times \log_2 f(X_i) dx \quad (2.2)$$

formülü ile hesaplanır.

2.1.3. MAUT Adımları

MAUT'u diğer ÇKKV yöntemlerinden ayıran temel özellik, risk algısı ve belirsizlik durumlarını karar sürecine doğrudan dahil edebilmesidir. Bu sayede, karmaşık ve dinamik karar problemlerinde daha gerçekçi ve güvenilir sonuçlar elde edilebilmektedir (Ananda & Herath, 2005).

MAUT yönteminin uygulanmasında kritik aşamalardan biri, kriter ağırlıklarının belirlenmesidir. Bu çalışmada, ağırlıklandırma işlemi için objektif bir yöntem olan entropi kullanılmıştır. Entropi yöntemi, kriterlerin alternatifler arasındaki dağılımına dayalı olarak hesaplanan istatistiksel bir ölçüttür ve her bir kriterin karar verme sürecindeki göreceli önem düzeyini yansıtır.

Öncelikle m alternatifli ve n kriterli karar verme problemi için karar matrisi;

$$D = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{i1} & X_{i2} & \cdots & X_{in} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{bmatrix}$$

oluşturulur.

Burada X_{ij} : i . alternatifin j . kritere göre başarı değeridir, $i=1,2,\dots,m$ ve $j=1,2,\dots,n$ 'dir.

Adım 1 :

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad i = 1, 2, \dots, m \text{ ve } j = 1, 2, \dots, n. \quad (2.3)$$

formülü ile normalleştirilmiş karar matrisi elde edilir.

Adım 2 :

$$e_j = \frac{1}{\ln m} \times \sum_{i=1}^m r_{ij} \times \ln r_{ij}, j = 1, 2, \dots, n. \quad (2.4)$$

formülü ile her bir kriterin entropi değeri bulunur. Burada e_j j . kriterin entropi değerini göstermektedir.

Adım 3:

$$W_j = \frac{1-e_j}{\sum_{j=1}^n (1-e_j)}, j = 1, 2, \dots, n. \quad (2.5)$$

ile kriterlerin ağırlık değerleri atanmış olur. $\sum_{j=1}^n W_j = 1$ olmalıdır.

MAUT yöntemi uygulama aşamaları aşağıdaki gibidir.

Adım 4 :

Fayda kriterlerine göre fayda değerleri belirlenir ve bunlar kullanılarak normalleştirilmiş değerler(r_{ij}) hesaplanır.

$$r_{ij} = \frac{x_{ij} - l_j^-}{u_j^+ - l_j^-} \quad (2.6)$$

burada $u_j^+ = \max_i x_{ij}$ ve $l_j^- = \min_i x_{ij}$ 'dir. Benzer şekilde maliyet kriterine göre de fayda değerleri belirlenir ve bunlar kullanılarak normalleştirilmiş değerler(r_{ij}) hesaplanır.

$$r_{ij} = \frac{u_j^+ - x_{ij}}{u_j^+ - l_j^-} \quad (2.7)$$

burada $u_j^+ = \max_i x_{ij}$ ve $l_j^- = \min_i x_{ij}$ 'dir.

Adım 5 : Entropi ağırlıkları(W_i) ile r_{ij} değerleri çarpılarak ağırlıklı r_{ij} değerleri hesaplanır. Bu değerlerin toplamı, toplam fayda değerini verir.

$$U_i = \sum_{j=1}^n w_j \times r_{ij} \quad (2.8)$$

Adım 6 : Tercih sıralaması yapılır. En yüksek toplam fayda değerine sahip olan alternatif en iyi alternatif olarak belirlenir.

2.2. Makine Öğrenmesi Yöntemleri

Bu tez çalışmasının temel amacı, futbol endüstrisinde görece az çalışılmış bir konu olan genç oyuncuların (21 yaş altı) gelecek odaklı performans potansiyelinin tahmin edilmesine yönelik metodolojik bir çerçeve önermektir. Bu amaç doğrultusunda makine öğrenmesi algoritmalarından yararlanılarak tahmin modelleri geliştirilecek ve

farklı yöntemlerin doğruluk, güvenilirlik ve genellenebilirlik performansları karşılaştırmalı olarak incelenecektir.

Makine öğrenmesi, bilgisayarların açıkça programlanmadan verilerden öğrenmesini ve deneyimler yoluyla kendi performansını iyileştirmesini amaçlayan yapay zekâ alt alanıdır (Mitchell, 1997). Alanın temelleri, Samuel'in (1959) dama oynayan programı ile atılmış, sonrasında istatistiksel öğrenme teorisi ve bilgisayar bilimlerinin kesişiminde gelişerek günümüzde sağlık, finans, pazarlama ve spor analitiği gibi birçok alanda yaygın uygulama alanı bulmuştur (Hastie, The elements of statistical learning: Data mining, inference, and prediction (2. Baskı), 2009). Spor bilimlerinde ise makine öğrenmesi, özellikle performans analizi, oyun stratejilerinin modellenmesi ve sakatlık riskinin öngörülmesi gibi uygulamalarda giderek daha fazla kullanılmaktadır. Ancak, genç futbolcuların gelecek odaklı performans potansiyelinin tahmini konusundaki uygulamalar henüz sınırlıdır.

Bu çerçevede tez kapsamında, farklı yapısal özelliklere sahip dört makine öğrenmesi algoritması ele alınmıştır: Rassal Orman (Random Forest), XGBoost, Destek Vektör Makineleri (Support Vector Machines, SVM) ve K-En Yakın Komşu (k-Nearest Neighbors, k-NN). Bu algoritmaların seçilmesinin temel gerekçesi, hem sınıflandırma hem de regresyon problemlerinde yaygın olarak kullanılmaları, yüksek tahmin gücü sunmaları ve birbirinden farklı güçlü yönleri olmalarıdır. Böylece, farklı algoritmaların karşılaştırılması yoluyla, genç oyuncuların performans tahmininde en etkili yaklaşımın belirlenmesi hedeflenmektedir.

2.2.1. Rassal Ormanlar (Random Forest)

Random Forest, Breiman (2001) tarafından geliştirilen ve karar ağaçlarının topluluk öğrenmesi (ensemble learning) prensibiyle bir araya getirilmesine dayalı güçlü bir algoritmadır. Yöntemin temelinde, eğitim setinden bootstrap örnekleme yoluyla üretilen alt kümeler üzerinde çok sayıda karar ağacı inşa edilmesi ve bu ağaçların çıktılarının çoğunluk oylaması veya ortalama ile birleştirilmesi yer almaktadır. Rassallık hem gözlem seçiminde hem de özellik seçiminde uygulandığı için modelin çeşitliliği artmakta, aşırı öğrenme riski azalmaktadır. Bu yönüyle algoritma, futbol gibi çok değişkenli ve karmaşık yapıya sahip verilerde genelleme gücü yüksek tahminler

üretmek için uygundur. Ancak, çok sayıda ağacın eğitilmesi nedeniyle yorumlanabilirliği düşük ve hesaplama maliyeti görece yüksektir.

2.2.2. XGBoost(Extreme Gradient Boosting)

XGBoost (Extreme Gradient Boosting), gradyan artırma yönteminin optimize edilmiş bir versiyonudur ve büyük ölçekli veri setlerinde yüksek doğruluk ve verimlilik sağlamayı hedefler (Chen & Guestrin, 2016). Bu algorithmada zayıf öğrenciler (genellikle sığ karar ağaçları), ardışık olarak eğitilmekte ve her yeni model, bir önceki modelin hatalarını azaltmaya çalışmaktadır. XGBoost, düzenlileştirme (regularization) mekanizması sayesinde aşırı uyumun önüne geçer ve paralel hesaplama özelliği ile büyük veri setlerinde dahi etkin çalışabilir. Futbol verilerinde farklı değişkenler arasındaki karmaşık ilişkilerin modellenmesinde bu esnek yapısı önemli bir avantajdır. Bununla birlikte, çok sayıda hiperparametre içermesi nedeniyle modelin doğru biçimde yapılandırılması uzmanlık gerektirmektedir.

2.2.3. Destek Vektör Makineleri (Support Vector Machine)

SVM, Cortes ve Vapnik (1995) tarafından geliştirilen ve özellikle sınıflandırma problemlerinde başarılı sonuçlar veren bir algoritmadır. Yöntem, sınıflar arasındaki marjini maksimize eden optimal hiperdüzlemi belirlemeyi amaçlar. Kernel fonksiyonlarının kullanımı ile doğrusal olarak ayrılamayan veriler, yüksek boyutlu uzaylara taşınarak doğrusal biçimde ayrılabilir hale getirilmektedir. SVM, sınırlı gözlem sayısı bulunan veri setlerinde dahi yüksek ayırt edicilik kapasitesi sunar. Bu özellik, genç futbolcular gibi daha dar örneklemeler üzerinde çalışıldığında özellikle önemlidir. Ancak, büyük veri setlerinde eğitim süresi uzundur ve parametre seçimleri model performansını doğrudan etkilemektedir.

2.2.4. K-En Yakın Komşu (K-Nearest Neighbors(k-NN))

k-NN, Fix ve Hodges (1951) tarafından önerilmiş ve Cover ve Hart (1967) tarafından teorik temelleri geliştirilmiş, basit fakat etkili bir algoritmadır. Bu yöntem, “benzer gözlemler benzer çıktılar üretir” varsayımına dayanır. Yeni bir oyuncunun performansı, eğitim setindeki en yakın k komşusunun performansına göre belirlenmektedir. Parametrik olmayan yapısı sayesinde dağılım varsayımı gerektirmez

ve doğrusal olmayan karar sınırlarını öğrenebilir. Bu yönüyle futbol gibi çok boyutlu ve heterojen verilerin analizinde esneklik sağlar. Ancak, büyük veri setlerinde hesaplama maliyetinin yüksek olması ve uygun k değerinin belirlenmesindeki hassasiyet önemli sınırlılıklardır.

Tablo 2.1 Makine öğrenmesi yöntemlerinin karşılaştırmalı özellikleri

Yöntem	Avantajları	Dezavantajları
Rassal Ormanlar	Yüksek doğruluk, aşırı uyuma karşı direnç, değişken önemini açıklayabilme	Yorumlanması güç, hesaplama maliyeti yüksek
XGBoost	Yüksek performans, düzenlileştirme, esnek yapı, paralel hesaplama	Çok sayıda hiperparametre, yapılandırma karmaşıklığı
Destek Vektör Makineleri	Küçük örneklerde yüksek doğruluk, kernel ile esneklik	Büyük veri setlerinde yavaş, parametre ayarına duyarlı
K-En Yakın Komşular	Basit, sezgisel, parametrik olmayan, doğrusal olmayan karar sınırları	Gürültüye duyarlı, k seçimine hassas, yüksek boyutlarda verimsiz

2.3. Hata Metrikleri

Makine öğrenmesi tabanlı sınıflandırma problemlerinde kullanılan performans ölçütleri, yalnızca modelin doğruluk oranını değil, aynı zamanda tahminlerin güvenilirliğini, şansa bağlılığı ve istatistiksel kararlılığı da dikkate almalıdır. Bu bağlamda, çalışmada kullanılan sınıflandırma modellerinin performansını değerlendirmek üzere üç temel metrik seçilmiştir: Doğruluk (Accuracy), Kappa İstatistiği (Cohen's Kappa) ve %95'lik Güven Aralığı (Confidence Interval - CI). Kesinlik (Precision), Duyarlılık (Recall) ve F1-Skor gibi diğer metrikler ise bu çalışmada raporlanmamıştır. Bunun temel nedeni, araştırmanın birincil amacının sınıflar arası dengesizlik (class imbalance) sorununu ele almaktan ziyade, modellerin genel tahmin performansını ve bu performansın istatistiksel güvenilirliğini değerlendirmek olmasıdır. Çalışmada seçilen üçlü metrik seti (Doğruluk, Kappa ve

%95 CI) birlikte değerlendirildiğinde kapsamlı, dengeli ve güvenilir bir değerlendirme çerçevesi sunmaktadır.

2.3.1. Doğruluk (Accuracy)

Doğruluk, sınıflandırma modellerinde en sık kullanılan temel performans ölçütlerinden biridir. En genel tanımıyla, modelin doğru sınıflandırdığı gözlem sayısının toplam gözlem sayısına oranıdır (Han, 2012). Basit ve anlaşılır yapısı sayesinde hem akademik araştırmalarda hem de saha uygulamalarında yaygın olarak tercih edilmektedir (Kelleher, 2015). Özellikle sınıflar arasında dengeli bir dağılım söz konusu olduğunda, doğruluk metriği modelin başarısını değerlendirmek için oldukça etkili bir araçtır. Ancak, doğruluk tek başına her zaman yeterli değildir. Dengesiz veri kümelerinde, baskın sınıfın doğru tahmin edilmesi modelin yüksek doğruluk değeri vermesine yol açabilir; fakat bu durumda azınlık sınıflarındaki başarısızlık göz ardı edilmiş olur (Powers, 2011). Bu nedenle, doğruluk metriği bu çalışmada temel bir performans göstergesi olarak kullanılmış olmakla birlikte, sınıf dağılımlarının model değerlendirmesine olası etkilerini dengelemek amacıyla Kappa istatistiği ile birlikte raporlanmıştır.

2.3.2. Kappa İstatistiği (Cohen's Kappa)

Cohen's Kappa katsayısı, tahminlerin doğruluk düzeyini yalnızca rastgele şansa bağlı uyumla karşılaştırmak yerine, gözlenen ve beklenen uyum arasındaki farkı normalize ederek ölçen güçlü bir metriktir (Cohen, 1960). Bu özelliği sayesinde, özellikle sınıf dengesizliğinin söz konusu olduğu veri setlerinde doğruluğa kıyasla daha güvenilir bir performans göstergesi sunmaktadır (McHugh, 2012). Kappa katsayısı -1 ile +1 arasında değer alır. 1'e yakın değerler mükemmel yakın bir uyumu, 0 değeri şansa dayalı uyumu, negatif değerler ise rastlantısal uyumun dahi altında bir performansı göstermektedir (Viera, 2005). McHugh (2012), Kappa'nın özellikle doğruluk değerinin yanıltıcı olabileceği durumlarda, sınıflandırıcıların anlamlı başarısını ortaya koyan daha güvenilir bir ölçüt olduğunu belirtmiştir. Bu nedenlerle, futbol oyuncularının potansiyel performanslarının tahmin edilmesine yönelik bu çalışmada kappa istatistiği doğruluk metriğine tamamlayıcı bir ölçüt olarak değerlendirilmiştir.

Metrik, genellikle 0 ile 1 arasında bir deęer alır. Buna gre, 0.61-0.80 aralıęı "nemli" (substantial), 0.81-1.00 aralıęı ise "neredeyse mkemmek" (almost perfect) bir uyum seviyesine iřaret eder. Kappa deęerinin 0.60'ın zerinde olması, modelin tahminlerinin řans faktrnden kaynaklanmadıęını ve sınıflar arasında tutarlı bir ayrım yapabildięini gstermesi aısından olduka deęerlidir. (Landis & Koch, 1977)

2.3.3. 95% Gven Aralıęı (Confidence Interval - CI)

Performans metriklerinin deęerlendirilmesinde yalnızca nokta tahminlerine (rneęin tek bir doęruluk deęeri) odaklanmak oęu zaman yanıltıcı olabilir. Bu nedenle, istatistiksel gvenilirlięi lmek amacıyla gven aralıklarının raporlanması nemlidir. %95 gven seviyesi, aynı alıřmanın birok kez tekrarlandıęını varsaydıęımızda, oluřturulan gven aralıklarının %95'inin gerek poplasyon deęerini iereceęi anlamına gelir. (Altman, 2000; Cumming & Finch, 2005). Gven aralıkları yalnızca bir belirsizlik gstergesi olmakla kalmaz, aynı zamanda modellerin karřılařtırılmasında da kritik bir rol stlenir. İki modelin doęruluk deęerleri birbirine yakın olsa bile, gven aralıklarının rtřp rtřmedięi modeller arasındaki farkın istatistiksel olarak anlamlı olup olmadıęını gsterebilir (Nahm, 2016). Bu alıřmada %95 gven aralıklarının raporlanması, doęruluk ve kappa deęerlerinin yalnızca nokta tahminleriyle deęil, aynı zamanda gvenilirlik dzeyiyle de birlikte yorumlanmasına olanak tanımaktadır.

Dar bir gven aralıęı, tahmin edilen metrięin yksek bir kesinlikle (precision) belirlendięini ve sonuların tekrarlanan denemelerde benzer ıkacaęına dair gveni artırır. Buna karřılık, geniř bir gven aralıęı, tahminde daha fazla belirsizlik olduęunu gsterir. Bu nedenle, bir modelin performansını raporlarken sadece nokta tahminlerini (point estimates) sunmak yeterli deęildir; performansın istatistiksel anlamlılıęını ve gvenilirlięini resmetmek iin mutlaka gven aralıklarının da rapor edilmesi ve yorumlanması nerilmektedir.

BÖLÜM ÜÇ

UYGULAMA

Bu bölümde araştırmada kullanılan veri seti tanıtılacak ve bu veri seti üzerinde gerçekleştirilen uygulama süreci açıklanacaktır.

3.1. Veri

Football Manager, Sports Interactive tarafından geliştirilen ve dünya genelinde milyonlarca kullanıcıya ulaşan bir futbol menajerliği simülasyonudur. Oyunun temel amacı, kullanıcıların bir futbol kulübünün teknik direktörü ve menejeri rolünü üstlenerek takım yönetimini stratejik biçimde yürütmelerini sağlamaktır. Oyunun yüksek düzeyde gerçekçilik iddiası, dünyanın farklı liglerinden binlerce futbolcuya ait istatistiksel ve niteliksel bilgilerin derlendiği kapsamlı bir veri tabanına dayanmaktadır. Bu veri tabanı, futbolcuların fiziksel, teknik, zihinsel ve taktiksel özelliklerini 1 ile 20 arasında puanlayan yüzlerce nitelikten (attribute) oluşmaktadır (Interactive, 2023).

Football Manager veri tabanının benzersizliği ve güvenilirliği, kullanılan veri toplama metodolojisinden kaynaklanmaktadır. Sports Interactive, dünya genelinde 1.300'ün üzerinde, büyük çoğunluğu yarı zamanlı çalışan ve aralarında eski scout'lar (yetenek avcıları), antrenörler ile spor yazarlarının da bulunduğu geniş bir “veri araştırmacıları” ağına sahiptir (Booth, 2018). Bu araştırmacılar, sorumlu oldukları liglerdeki oyuncuları düzenli olarak gözlemlemekte, her bir futbolcu için yüzlerce niteliği değerlendirerek merkezi veri tabanına aktarmaktadır. Toplanan bilgiler, merkez ofis tarafından denetim ve çapraz doğrulama süreçlerinden geçirilmekte; böylece veri tutarlılığı ve standardizasyonu sağlanmaktadır. Bu sistematik yapı, Football Manager veri tabanını yalnızca maç istatistikleri (örneğin gol, pas, asist vb.) sunan geleneksel veri sağlayıcıların ötesine taşıyarak futbolcuların çok boyutlu performans profillerinin oluşturulmasına olanak tanımaktadır.

Son yıllarda Football Manager veri tabanı, akademik araştırmalarda da giderek daha fazla kullanılmaya başlanmış; özellikle oyuncu performansı değerlendirmesi,

potansiyel yetenek tahmini ve deęerleme modelleri geliřtirme alanlarında alternatif bir veri kaynaęı olarak nem kazanmıřtır.

Bu alıřmada kullanılan ham veri seti, Football Manager 2023 srmnden elde edilmiř olup 156.000 futbolcu ve 68 deęiřkenden oluřmaktadır. Ancak deęiřkenlerin tamamı analiz aısından anlamlı katkı saęlamadıęından, kiřisel veya baęlamsal nitelikteki stnlar (rneęin milliyet, maař, uyruk vb.) analiz dıřında bırakılmıřtır. Nihai veri tablosunda, her mevki iin elit dzeyde yer alan 200 futbolcu seilmiř ve stnlarda 50 temel kriter tanımlanmıřtır. rneklem byklęnn 200 futbolcuyla sınırlandırılmasının nedeni, futbolcu reyting (FR) deęerlerinin yaklařık 200. sıradan itibaren 130'un altına dřmesidir. Potansiyel reyting skoru 130'un altında olan futbolcular genellikle "Beř Byk Lig" (Top Five Leagues) olarak bilinen Premier League, LaLiga, Serie A, Bundesliga ve Ligue 1 dıřında forma giyen, st dzey performans kriterlerini karřılamayan oyunculardır. Bu nedenle, MAUT uygulaması kapsamında yalnızca anlamlı dzeyde ayrıřan elit futbolcular analize dhil edilmiřtir.

MAUT ynteminin bu alıřmadaki temel amacı, literatrde genel olarak "elit futbolcu" olarak tanımlanan kavramı mevki bazında nesnel ve sistematik biimde tanımlamak ve bu doęrultuda her mevki iin belirleyici znitelikleri istatistiksel olarak ortaya koymaktır. Futbol literatrnde elit futbolcu, en st seviyede performans gsteren oyuncuları ifade etmekle birlikte, bu seviyeyi belirleyen kriterler mevkilere gre farklılık gstermektedir. Dolayısıyla, mevkiye zg kritik deęiřkenlerin belirlenmesi, analiz srecinin en nemli ařamalarından birini oluřturmaktadır.

MAUT yntemiyle belirlenen bu elit futbolcu zellikleri, bir sonraki ařamada yksek potansiyelli gen oyuncuların tespiti iin denetimli makine ęrenmesi algoritmalarında girdi olarak kullanılmıřtır. Bu kapsamda, gnmzde elit seviyede forma giyen futbolcuların 21 yař ve altı dnemlerine ait veriler sırasıyla Football Manager 2017, Football Manager 2020, Football Manager 2021 ve Football Manager 2022 veri tabanlarından alınarak mevcut veri setine eklenmiřtir. Gnmzde sahip olunan potansiyel reyting deęeri baęımlı deęiřken olarak tanımlanırken, MAUT yntemiyle seilen znitelikler baęımsız deęiřkenler olarak modele dhil edilmiřtir. Bylelikle gen futbolcuların mevcut yetenek skorlarından yola ıkarak, ilerleyen

yıllarda ulaşabilecekleri potansiyel performans seviyelerinin tahmin edilmesi mümkün hale gelmiştir.

Bağımlı değişken üç kategoriden oluşan nominal bir yapıya sahiptir. Bu kategoriler, uluslararası futbol piyasasındaki rekabet düzeyleri ve oyuncuların ulaşabileceği üst lig standardı dikkate alınarak operasyonelleştirilmiştir:

- Kategori 1 (100–129): “Beş Büyük Lig Dışı Seviyesi”

Bu grupta yer alan oyuncular, kariyerlerinin zirvesinde Premier League, LaLiga, Serie A, Bundesliga ve Ligue 1 dışında kalan liglerde (örneğin Eredivisie, Primeira Liga, Süper Lig) başarılı olma potansiyeline sahiptir.

- Kategori 2 (130–149): “Beş Büyük Lig Standardı”

Bu skor aralığı, oyuncunun Beş Büyük Lig düzeyinde rekabet edebilme, fiziksel ve taktiksel gereklilikleri karşılayabilme kapasitesini yansıtmaktadır. Bu gruptaki futbolcular, üst düzey liglerde ana kadro veya ilk yedek statüsünde istikrarlı performans sergileyebilen oyuncular olarak değerlendirilmektedir.

- Kategori 3 (150–200): “Yıldız Adayı”

Bu kategori, Avrupa’nın en üst düzey kulüplerinde forma giyebilecek, Şampiyonlar Ligi seviyesinde etkin rol oynayabilecek ve küresel ölçekte tanınırlık potansiyeline sahip elit futbolcuları ifade etmektedir.

Bu üçlü sınıflandırma, ham potansiyel skorlarını futbol endüstrisi açısından yorumlanabilir ve eyleme dönüştürülebilir karar kategorilerine dönüştürmekte; böylece makine öğrenmesi modellerinin yalnızca skor tahmini yapmanın ötesinde, bir oyuncunun ulaşabileceği nihai kariyer seviyesini öngörebilmesini sağlamaktadır.

Araştırmada, Football Manager 2023 veri tabanından elde edilen teknik, fiziksel ve zihinsel değişkenler kullanılarak futbolcuların potansiyel performans aralıklarının tahmin edilmesi amaçlanmıştır. Bu doğrultuda, Destek Vektör Makineleri (SVM), Rastgele Orman (Random Forest), XGBoost ve K-En Yakın Komşu (KNN)

algoritmaları uygulanmıştır. Modellerin başarısı “Doğruluk, Kappa İstatistiği ve %95 Güven Aralığı” metrikleri üzerinden karşılaştırmalı biçimde değerlendirilmiştir.

Elde edilen sonuçların, futbol kulüplerinin genç yetenek keşfi, veri odaklı transfer kararları ve stratejik planlama süreçlerine katkı sağlayarak endüstriyel uygulamalarda kullanılabilir bir karar destek çerçevesi oluşturması beklenmektedir.

3.1.1. Değişkenler

Futbol, saha üzerinde on bir oyuncu ile oynansa da, taktiksel dizilişler doğrultusunda ortaya çıkan fonksiyonel roller, geleneksel mevki sınıflandırmalarının ötesinde çeşitlilik göstermektedir. Modern futbolda bu çeşitlilik, aynı mevki içinde dahi oyuncuların farklı profillerde konumlanmasına neden olmaktadır. Örneğin, defansif orta saha pozisyonunda “regista” rolü, derinden oyun kurma ve pas dağıtma becerilerini ön plana çıkarırken, “savaşçı orta saha” rolü fiziksel mücadele gücü ve top kapma becerileri ile tanımlanmaktadır. Bu bağlamda, aynı mevki içerisinde oynuyor gibi görünseler de, oyuncuların takım dizilişlerindeki rolleri doğrultusunda sahip olmaları gereken temel nitelikler farklılık göstermektedir. Regista rolündeki bir oyuncudan, fiziksel güç, hava toplarındaki hakimiyet veya top kapma gibi özellikler birincil öncelik taşımamakta; buna karşılık savaşçı orta saha için bu nitelikler kritik öneme sahiptir. Benzer şekilde, savaşçı orta saha rolündeki bir oyuncudan yüksek pas yeteneği, oyun görüşü ve teknik kalite beklentisi ise gerçekçi bir değerlendirme olmayacaktır.

Bu heterojen yapı göz önünde bulundurulduğunda, tutarlı ve nesnel bir elit futbolcu tanımı oluşturabilmek için her mevki, rol çeşitliliğinden bağımsız olarak temel kriterler üzerinden değerlendirilmiştir. Bu doğrultuda analizde kaleciler kapsam dışında bırakılmış ve sekiz temel mevki sınıfı belirlenmiştir: Stoper (D), Defansif Bek (DB), Kanat Bek (WB), Defansif Orta Saha (DM), Merkez Orta Saha (MC), Kanat (AM), Ofansif Orta Saha (AMC) ve Forvet (ST).

MAUT yöntemi, her mevki için seçilen 200 elit futbolcunun verileri üzerinde Python programlama dili kullanılarak uygulanmış; ilgili mevkinin performansını tanımlayan en kritik ve ayırt edici özellikler belirlenmiştir. Bu süreçte temel amaç,

mevki bazında öne çıkan kriterleri saptamak ve elde edilen bu kriterleri denetimli makine öğrenmesi algoritmalarının eğitiminde girdi olarak kullanmaktır. Entropi temelli MAUT yaklaşımı ile her sütun için ağırlık değerleri hesaplanmış ve en yüksek ağırlığa sahip kriterler seçilerek algoritmalara dahil edilmiştir.

Örnek teşkil etmesi amacıyla, forvet pozisyonunda oynayan dünyanın en iyi beş futbolcusuna ait veriler, 3. ana bölümde “Tablo 3.1” olarak sunulmuştur.

Tablo 3.1 Veri setinden seçilmiş örnek 5 satır verisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Erling Haaland	22	12	12	7	...	10	12	31	185
Harry Kane	28	12	12	9	...	11	16	39	183
Robert Lewandowski	33	2	15	3	...	10	11	33	183
Karim Benzema	34	8	17	10	...	17	18	37	183
Kylian Mbappé	23	7	8	13	...	12	13	30	172

Çok Ölçütlü Karar Analizi (MAUT) yönteminin doğru ve güvenilir bir şekilde uygulanabilmesi için, farklı ölçekler ve birimlere sahip kriter değerlerinin ortak, boyutsuz bir ölçeğe dönüştürülmesi gerekmektedir. Bu çalışmada kullanılan ham veri seti, futbolcuların fiziksel, teknik ve zihinsel özelliklerini temsil eden ve birbirinden tamamen farklı aralıklarda değerler alabilen 50 öznelikten oluşmaktadır (örneğin, 1–20, 1–100, saniye, metre). Farklı ölçeklerdeki bu verilerin karşılaştırılabilir ve entegre bir şekilde değerlendirilebilmesi amacıyla doğrusal standartlaştırma (normalizasyon) işlemi uygulanmıştır.

Standartlaştırma sürecinde, her bir kriterin fayda (maksimize edilmesi istenen) veya maliyet (minimize edilmesi istenen) kriteri olup olmadığı belirlenmiştir. Örneğin, "hız" ve "pas" gibi yüksek değerleri tercih edilen özellikler fayda kriteri; "sakatlanma

riski" veya "yaş" gibi düşük değerleri tercih edilen özellikler ise maliyet kriteri olarak sınıflandırılmıştır.

Buna göre, i alternatifini (futbolcuyu) ve j kriterini (özniteliği) temsil etmek üzere, standartlaştırma işlemi aşağıdaki formüller ile gerçekleştirilmiştir:

Fayda kriterleri için: Kriter değeri ne kadar yüksekse o kadar tercih edilir. Bu durumda, ham değerler (x_{ij}) 0 ile 1 arasında standart değerlere (v_{ij}) aşağıdaki doğrusal formül kullanılarak dönüştürülmüştür:

$$v_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (3.1)$$

Bu formülde, $\min(x_j)$ ve $\max(x_j)$ sırasıyla j kriterine ait minimum ve maksimum değerleri ifade etmektedir.

Maliyet kriterleri için: Kriter değeri ne kadar düşükse o kadar tercih edilir. Bu durumda standartlaştırma formülü aşağıdaki şekilde uygulanmıştır:

$$v_{ij} = \frac{\max(x_j) - x_{ij}}{\max(x_j) - \min(x_j)} \quad (3.2)$$

Standartlaştırma işlemleri sonucunda, 50 farklı özniteliğin tümü 0 ile 1 arasında değerler alacak şekilde dönüştürülmüş ve artık doğrudan karşılaştırılabilir, boyutsuz bir matris elde edilmiştir. Bu standartlaştırılmış matris, bir sonraki aşama olan entropi tabanlı ağırlıklandırma ve nihai MAUT skorlarının hesaplanması için temel girdi niteliğini taşımaktadır.

Öznitelliklerin büyük çoğunluğu, maksimum değerlerine ulaşmasının arzu edildiği fayda kriterleri olarak tanımlanmıştır. Bununla birlikte, üç değişkenin minimize

edilmesi gereken maliyet kriteri olarak ele alınması gerekmektedir: “Yaş”, “Sakatlığa Yatkınlık” ve “İtiraz”.

Yaş: Futbolcularda deneyim önemli bir avantaj sağlamakla birlikte, belirli bir yaştan sonra fiziksel performans düşmekte ve sakatlık sonrası toparlanma süresi uzamaktadır. Bu nedenle yaş, minimize edilmesi gereken bir kriter olarak değerlendirilmiştir.

Sakatlığa Yatkınlık: Oyuncunun istikrarlı performans gösterebilmesi açısından sakatlığa yatkınlığının düşük olması kritik öneme sahiptir. Modern futbolun artan temposu, bu özelliğin önemini daha da artırmaktadır.

İtiraz: Hakem kararlarına aşırı tepki gösterme eğilimi, kart görme riskini ve dolayısıyla takımın eksik kalma olasılığını yükseltmektedir. Bu nedenle bu kriter de düşük değer alması gereken bir parametre olarak kabul edilmiştir.

Bu bağlamda, standartlaştırılmış 0–1 ölçeğindeki matris üzerinden gerçekleştirilen değerlendirmeler, entropi tabanlı ağırlıklandırma yöntemi ile MAUT skorlarına dönüştürülmüş ve her mevki için elit seviyede performansı tanımlayan ölçütlerin belirlenmesinde kullanılmıştır.

3.2. MAUT Uygulaması

MAUT uygulaması sırasında karar matrislerinin örnek teşkil etmesi amacıyla, sekiz mevki pozisyonunda oynayan dünyanın en iyi beş futbolcusuna ait veriler kullanılarak sunulmuştur. Tüm veri setinin tamamının gösterilmesi uzun olacağından, matrisler yalnızca ilk beş oyuncuya ilişkin temsili veriler üzerinden örneklendirilmiştir.

3.2.1. Defans (D)

Uygulama kapsamında ele alınan ilk mevki defanstır. Bu aşamada, dünyanın en iyi 200 defans oyuncusunun ortak performans özelliklerinin belirlenmesi amaçlanmıştır. Bu doğrultuda, defans mevkiisine ait veri setinde yer alan 50 ham öznelik, Çok Ölçütlü Karar Analizi (MAUT) yöntemi kapsamında sistematik olarak

değerlendirilmiştir. Çalışmanın temel hedefi, defans oyuncularının profilini oluşturan bu öz nitelikler arasından mevkinin doğası gereği en kritik olanların tespit edilmesidir.

Bu süreçte her bir öz nitelik için fayda fonksiyonları tanımlanmış ve kriterlerin nihai karar üzerindeki göreceli önemini yansıtan ağırlık katsayıları hesaplanmıştır. Entropi yöntemi ile gerçekleştirilen bu ağırlıklandırma süreci, öz niteliklerin sahip olduğu bilgi içeriğine dayalı olarak nesnel bir sıralama elde edilmesini sağlamıştır. Hesaplanan ağırlık değerleri doğrultusunda, defans mevkiisi için en yüksek ağırlığa sahip öz nitelikler belirlenmiş ve bu öz nitelikler, defans oyuncularının elit seviyede performans sergilemesinde belirleyici temel bileşenler olarak yorumlanmıştır.

Tablo 3.2 Defans oyuncularının için karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Virgil van Dijk	31	5	10	3	...	12	18	29	182
Rúben Dias	25	4	8	4	...	14	16	32	172
Aymeric Laporte	28	7	6	8	...	13	14	30	168
Jules Koundé	23	9	6	6	...	13	12	30	165
Antonio Rüdiger	29	14	15	5	...	9	11	30	165

Sonuç olarak, belirlenen öz nitelik seti, defans mevkiine özgü makine öğrenmesi modelinin eğitiminde kullanılmak üzere hazırlanmıştır. Karar matrisinin ilk beş satırı, 3. ana bölümde “Tablo 3.2” olarak sunulmaktadır. Örnek olarak, dünyanın önde gelen stoperlerinden Antonio Rüdiger’in, diğer oyunculara kıyasla sakatlığa yatkınlık ve kart görme potansiyelinin daha yüksek olduğu gözlemlenmiştir.

Tablo 3.3 Defans oyuncularını için standartlaştırılmıř karar matrisi

İsim	Yař	Sakatlıęa Yatkinlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Virgil van Dijk	0.0057	0.0029	0.0070	0.0029	...	0.0048	0.0071	0.0049	0.0064
Rúben Dias	0.0046	0.0024	0.0056	0.0038	...	0.0055	0.0063	0.0054	0.0060
Aymeric Laporte	0.0051	0.0041	0.0042	0.0076	...	0.0052	0.0055	0.0050	0.0059
Jules Koundé	0.0042	0.0053	0.0042	0.0057	...	0.0052	0.0047	0.0050	0.0058
Antonio Rüdiger	0.0053	0.0082	0.0105	0.0048	...	0.0036	0.0043	0.0050	0.0058

Karar matrisine standartlařtırma iřlemi uygulanmıřtır. Bu iřlem sonucunda elde edilen standartlařtırılmıř karar matrisi, Tablo 3.3'te gsterilmektedir. Standartlařtırmanın temel amacı, zniteliklerin farklı birim ve leklerde ifade edilen deęerlerini boyutsuz hâle getirerek karřılařtırılabilir bir formata dnřtrmektir. rneęin, 'Yař' deęiřkeni iin bir st sınır belirlenmemiřken, 'Sakatlıęa Yatkinlık' deęiřkeni iin taban deęeri 1 ve tavan deęeri 20 olarak alınmıřtır. Benzer řekilde, 'Çift Ayak' deęiřkeni iin taban 2 ve tavan 40 olarak belirlenmiřtir.

Standartlařtırılmıř karar matrisi, entropi katsayısının hesaplanmasında kullanılmaktadır. Entropi deęerleri, her bir znitelięin bilgi ierięini yansıtarak aęırlıklandırma srecine temel teřkil etmektedir. Hesaplanan entropi deęerleri doęrultusunda, zniteliklerin aęırlıkları toplam deęeri 1 olacak řekilde belirlenmiřtir. İlgili hesaplamalar, Blm 2'de detaylandırılan MAUT adımları kapsamında formller kullanılarak gerekleřtirilmiřtir. Bu iřlemler sonucunda zniteliklerin aęırlıkları 0.019988 ile 0.020003 arasında deęiřmiř ve dokuz farklı aęırlık grubu oluřturulmuřtur. İlk grup, aęırlık deęeri en yksek olan znitelikleri iermekte olup, elit futbolcu tanımlamasında ncelikli olarak dikkate alınacaktır. Bu grup, sırasıyla FR (Futbolcu Reytingi), Çift Ayak, Markaj, Mdehale, Pozisyon, Kafa Vuruřu, Tahmin, Takım Oyunu, Dayanıklılık, Pas, Çalıřkanlık, Karar Alma, Zıplama, Top Kontrol, G, Teknik, Soęukkanlılık, Konsantrasyon, İstikrar, Cesaretilik, Kararlılık ve Vcut Zindelięi deęiřkenlerinden oluřmaktadır.

FR değeri, tüm mevkilerde genellikle en yüksek veya ikinci en yüksek ağırlığa sahip değişken olarak ön plana çıkmaktadır; bu durum, futbolcunun toplam performansını gösteren bir ölçüt olması nedeniyle beklenen bir sonuçtur. Diğer özneliklerin incelenmesi, elit bir stoper için çift ayak kullanımının oldukça belirleyici bir kriter olduğunu göstermektedir. Ayrıca markaj ve müdahale becerileri, defans oyuncularının performansını doğrudan etkileyen kritik kriterler arasında yer almaktadır.

Defans oyuncusunun markaj kalitesi, rakip oyuncuların oyun stratejisini bozarak rakip takımın atak etkinliğini azaltabilir. Benzer şekilde tehlikeli pozisyonlarda yapılan müdahalelerin zamanlaması ve kalitesi, maçın skorunu doğrudan etkileyebilir; yanlış zamanlamalı müdahaleler penaltı veya kırmızı kartla sonuçlanabilir. Bu bağlamda, elit bir defans oyuncusunun performansını belirleyen en önemli öznelikler arasında bu özelliklerin yer alması, elde edilen sonuçların tutarlılığını ve uygulamadaki geçerliliğini desteklemektedir.

3.2.2. Defansif Bek (DB)

Bek mevkisi, günümüz futbolunda iki temel fonksiyonel role ayrılmaktadır: defansif bek ve kanat bek. Her iki rolün gerektirdiği temel özellikler genel anlamda benzerlik gösterse de, bu durum kanat bek olarak başarılı olan bir oyuncunun otomatik olarak defansif bek rolünde de aynı performansı sergileyeceği anlamına gelmemektedir. Bu farklılık, mevkiye özgü analizlerin rol bazında ayrı yürütülmesini gerekli kılmaktadır.

Bu bağlamda, çalışmada sekiz temel mevki arasında ikinci ve üçüncü sırada yer alan bek mevkisi için iki ayrı analiz gerçekleştirilmiştir. İlk olarak, dünyaca tanınmış en iyi 200 defansif bek oyuncusunun ortak özellikleri belirlenerek mevkiye özgü performans profili oluşturulacaktır.

Tablo 3.4 Defansif bek mevkisi için karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
David Alaba	30	7	10	12	...	12	10	30	158
Ben Chilwell	25	4	7	13	...	12	15	32	157
José Gayà	27	8	6	14	...	14	10	27	156
Theo Hernández	24	6	14	8	...	6	9	28	149
Ferland Mendy	27	7	9	5	...	10	12	34	149

Bek mevkisi profili açısından öne çıkan kriterlerden biri, isabetli orta yapabilme becerisidir. Bu bağlamda, “Orta Açma” değişkeni önemli bir performans göstergesi olarak değerlendirilmiştir. Örnek olarak, Ben Chilwell ve Theo Hernández’in bu becerilerini yansıtan skorları 15 olarak ölçülmüştür; bu değer, ilgili grup ortalamasının üzerinde yer almakta ve özellikle kanat üzerinden oyun kuran defans oyuncularının katkısını vurgulamaktadır. Karşılaştırmalı olarak, aynı değişkenin skorları defans oynayan diğer oyuncular arasında ortalamanın altında kalmaktadır.

Tablo 3.5 Standartlaştırılmış defansif bek karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
David Alaba	0.0055	0.0043	0.0071	0.0062	...	0.0049	0.0041	0.0051	0.0059
Ben Chilwell	0.0046	0.0025	0.0050	0.0067	...	0.0049	0.0061	0.0055	0.0058
José Gayà	0.0050	0.0049	0.0043	0.0072	...	0.0057	0.0041	0.0046	0.0058
Theo Hernández	0.0044	0.0037	0.0100	0.0041	...	0.0024	0.0037	0.0048	0.0055
Ferland Mendy	0.0050	0.0043	0.0064	0.0026	...	0.0041	0.0049	0.0058	0.0055

Defansif bek mevkisi için 10 farklı öznitelik ağırlık değeri hesaplanmıştır. Bu ağırlıklar, elit bir defansif bek oyuncusunun sahip olması gereken temel özelliklerin önem sıralamasını ortaya koymaktadır. Hesaplanan ağırlık sıralaması şu şekildedir: FR (Futbolcu Reytingi), Çift Ayak, Tahmin, Dayanıklılık, Karar Alma, Pas, Top Kontrolü, Müdehale, Takım Oyunu ve Teknik, devamında Hızlanma, Çalışkanlık, Çeviklik, Hız, Konsantrasyon, Markaj, Pozisyon, İstikrar, Soğukkanlılık ve Orta Açma.

Elit defansif beklerin özellikleri incelendiğinde, FR'den sonra iki ayak kullanabilme becerisinin en kritik kriterlerden biri olduğu görülmektedir. Üçüncü sırada yer alan “Tahmin” özelliği, oyuncunun sahadaki pozisyonunu doğru bir şekilde algılayarak topun gideceği noktayı ve rakip atağın nasıl şekilleneceğini öngörebilme yeteneğini ifade etmektedir.

Bunun yanı sıra, pas yeteneği, pozisyon alma becerisi, hamle zamanlaması ve karar alma yetkinliği, elit defansif bek profili için önem taşıyan diğer kritik öznitelikler arasında yer almaktadır. Bu kriterler, oyuncunun sahadaki etkinliğini ve takım savunmasına katkısını belirleyen temel göstergeler olarak yorumlanmaktadır.

3.2.3. Kanat Bek (WB)

Uygulamada incelenen sekiz mevki arasında üçüncü mevki, kanat bek olarak belirlenmiştir. Bu kapsamda, dünyanın önde gelen 200 kanat bek oyuncusunun ortak özellikleri sistematik bir şekilde analiz edilerek, elit bir futbolcуда bulunması gereken mevkiye özgü kriterler belirlenecektir.

Tablo 3.6 Kanat bek için karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Andrew Robertson	28	6	8	14	...	10	13	24	158
Kostas Tsimikas	26	8	6	14	...	13	15	26	155
Ivan Perišić	33	6	13	11	...	8	10	36	153
Theo Hernández	24	6	14	8	...	6	9	28	149
Javi Galán	27	3	3	13	...	15	11	27	143

Kanat bek özelliklerini içeren karar matrisinde Theo Hernández'in de yer aldığı tespit edilmiştir. Bu futbolcunun her iki rolü de yüksek düzeyde oynayabilme yeteneğine sahip olduğu gözlemlenmektedir. Bununla birlikte, defansif bek olarak görev yapan bir oyuncunun isabetli orta açma becerisi, kanat bek rolündeki oyunculara kıyasla genellikle daha düşük seviyededir. Bu durum, hücumla daha aktif katılan kanat beklerden, oyun kurma ve orta açma becerilerinin daha gelişmiş olmasının beklendiğini göstermektedir.

Tablo 3.7 Kanat bek için standartlaştırılmış karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Andrew Robertson	0.0052	0.0037	0.0052	0.0070	...	0.0043	0.0055	0.0041	0.0063
Kostas Tsimikas	0.0048	0.0050	0.0039	0.0070	...	0.0055	0.0063	0.0044	0.0062
Ivan Perišić	0.0061	0.0037	0.0085	0.0055	...	0.0034	0.0042	0.0062	0.0061
Theo Hernández	0.0044	0.0037	0.0092	0.0040	...	0.0026	0.0038	0.0048	0.0059
Javi Galán	0.0050	0.0019	0.0020	0.0065	...	0.0064	0.0046	0.0046	0.0057

Elit kanat bek mevkisi için yapılan analizler sonucunda, özneliklerin önem sıralaması şu şekilde belirlenmiştir: FR, Çift Ayak, Dayanıklılık, Pas, Teknik, Hızlanma, Hız, Çeviklik, Tahmin, Müdahale, Çalışkanlık, Top Kontrolü, Karar Alma, Topsuz Alan, Soğukkanlılık, Orta Açma, Takım Oyunu, Pozisyon, İstikrar, Denge ve Dribbling. Kanat bek mevkisi için benzer şekilde 10 farklı ağırlık sınıfı oluşturulmuş ve elit bir kanat bek için toplamda 21 becerinin ön planda olduğu belirlenmiştir.

Analizler, elit bir kanat bek için iki ayağını kullanabilme becerisinin en kritik öznelik olduğunu göstermektedir. İkinci sırada ise dayanıklılık yer almakta olup, bu özellik defansif bek için üçüncü sıradadır. Bek oyuncularının oyunun her iki yönünde aktif rol alması nedeniyle yüksek kondisyon seviyesi gerekmektedir. Defansif bekler genellikle çizgiye inme ve hücumla destek verme görevlerinde kanat bekler kadar aktif olmadığından, dayanıklılık açısından kanat beklerden daha düşük beklenti söz konusudur. Buna karşılık, kanat bekler hücumla daha fazla destek verdiklerinden ve

geriye koşacakları mesafe daha uzun olduğundan, dayanıklılık bu rol için defansif beke kıyasla daha öncelikli bir öznelik olarak öne çıkmaktadır.

Dayanıklılığı takiben gelen pas ve teknik beceriler, kanat bek için defansif beke kıyasla öne çıkan ayırt edici özelliklerdir. Kanat bekler, oyunun üçüncü bölgesinde daha fazla topla buluşmakta ve topa sahip olma süreleri daha uzun olmaktadır. Bu nedenle pas yetenekleri ve teknik becerileri, hatasız oyun yürütme açısından kritik önem taşımaktadır. Takım hücumdayken yapılacak bir pas hatası, maç sonucunu doğrudan etkileyebileceğinden, bu oyunculardan güvenilir pas opsiyonlarını kullanmaları ve riskli paslardan kaçınmaları beklenmektedir; bu beceriler, alt yaş gruplarından itibaren sistematik olarak geliştirilir.

Topsuz oyun becerisi ise genellikle gol bölgesine yakın ve hücum hattında oynayan oyuncular için ayırt edici bir özellik olarak kabul edilmektedir. İlginç bir şekilde, kanat bek mevkisi için bu özellik önem arz ederken, elit defansif bekler için aynı ölçüde kritik değildir. Bu durum, aynı sahada görev alan bu iki rolün, sahip olması gereken özellikler bakımından birbirinden farklı olduğunu ve ayrı ayrı değerlendirilmesinin gerekliliğini ortaya koymaktadır.

3.2.4. Defansif Orta Saha (DM)

Uygulama kapsamında ele alınacak sekiz mevki içinden dördüncü mevki, defansif orta saha (DM) olarak belirlenmiştir. Bu kapsamda, dünyanın en iyi 200 defansif orta saha oyuncusunun ortak özellikleri analiz edilerek, mevkiye özgü elit performans kriterlerinin tespit edilmesi hedeflenmektedir.

Tablo 3.8 Defansif orta saha mevkiisi için karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Rodri	26	4	6	7	...	13	16	33	173
Joshua Kimmich	27	6	8	15	...	12	10	32	164
Thiago	31	12	12	13	...	10	12	30	162
Casemiro	30	5	14	10	...	10	11	33	161
Declan Rice	23	6	5	7	...	16	14	31	160

Karar matrisine genel olarak bakıldığında, incelenen oyuncuların profesyonellik becerilerinin çoğunlukla ortalamanın üzerinde veya neredeyse kusursuz seviyede olduğu gözlemlenmektedir. Kendi mevkiilerinde dünyanın en iyi oyuncuları olarak değerlendirilen bu futbolcuların birbirlerinden farklı profillere sahip oldukları, “Tablo 3.8”de sunulan beceri skorlarıyla açık biçimde ortaya konulmaktadır. Örneğin, Joshua Kimmich’in orta açma becerisi 17 ile oldukça yüksek bir değer alırken, aynı mevkiideki Rodri için bu değer 8 olarak ölçülmüş ve ortalamanın oldukça altında kalmıştır. Bu durum, elit seviyedeki defansif orta saha oyuncuları arasında belirgin performans çeşitliliğine işaret etmektedir.

Tablo 3.9 Standartlaştırılmış defansif orta saha karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Rodri	0.0047	0.0025	0.0041	0.0040	...	0.0054	0.0063	0.0055	0.0064
Joshua Kimmich	0.0049	0.0037	0.0055	0.0086	...	0.0050	0.0040	0.0053	0.0060
Thiago	0.0056	0.0074	0.0083	0.0075	...	0.0041	0.0048	0.0050	0.0060
Casemiro	0.0054	0.0031	0.0096	0.0058	...	0.0041	0.0044	0.0055	0.0059
Declan Rice	0.0042	0.0037	0.0034	0.0040	...	0.0066	0.0055	0.0051	0.0059

Defansif orta saha (DM) mevkiisinde görev yapan oyuncuların öznitelik ağırlıkları, 11 farklı sınıfa ayrılmıştır. Elit defansif orta saha oyuncularının profilini tanımlayan kritik öznitelikler ise toplam 14 değişkenden oluşmaktadır. Bu değişkenler sırasıyla; Çift Ayak, FR, Pas, Tahmin, Takım Oyunu, Top Kontrolü, Çalışkanlık, Pozisyon, Teknik, Karar Alma, Hız, Müdehale, Dayanıklılık ve Çeviklik’tir.

İki ayağını kullanabilme becerisi, bu pozisyonda en belirleyici özelliklerden biri olarak öne çıkmakta ve FR değerinden dahi daha yüksek bir önceliğe sahiptir. Oyunun merkezinde görev alan bir oyuncunun iki ayağını etkin biçimde kullanabilmesi, pas dağıtımında ve top kontrolünde zaman kazandırmakta ve rakip markajına karşı kritik avantaj sağlamaktadır. Özellikle merkez sahada, saniyeler içinde yapılacak ayak değişiklikleri, topun boşta kalan oyuncuya ulaşmasını sağlayarak rakibin atağı etkisiz hale getirilmesine imkân tanıyabilir.

Bu mevkide görev yapan oyuncular genellikle defanstan topu çıkaran, oyunun yönünü belirleyen ve maçın temposunu kontrol edebilen teknik kapasitesi yüksek futbolculardan oluşmaktadır. Bu bağlamda, pas yeteneğinin en önemli ikinci değişken olarak öne çıkması, oyuncuların oyun kurucu rolünü ve orta saha organizasyonundaki kritik sorumluluklarını desteklemektedir.

3.2.5. Merkez Orta Saha (MC)

Uygulama kapsamındaki sekiz mevki arasında beşinci mevki, merkez orta sahadır (MC). Bu mevki için, dünyanın en iyi 200 merkez orta saha oyuncusunun ortak özellikleri analiz edilerek, elit bir futbolcuda bulunması gereken mevkiye özgü kriterler belirlenmiştir.

Tablo 3.10 Merkez orta saha karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Kevin De Bruyne	31	7	13	14	...	16	13	36	181
Luka Modrić	36	11	5	16	...	19	17	33	176
Bernardo Silva	27	6	14	14	...	12	8	28	174
Federico Valverde	24	6	7	13	...	10	11	32	164
Leon Goretzka	27	14	10	5	...	15	8	30	164

Merkez orta saha (MC) pozisyonunda görev yapan oyuncular, günümüz futbolunda teknik yeterlilikleri ön plana çıkan ve oyunun hem savunma hem de hücum yönünde çeşitli sorumluluklar üstlenen futbolculardır. Rakip oyuncularla daha yakın temas hâlinde oynadıkları için bazı yetkinliklerinin diğer mevkilere kıyasla daha belirgin olması gerekmektedir. Bu pozisyonundaki oyuncuların, baskı ve stres altında doğru kararlar alabilmesi, topu kaybetmeden en uygun pas açısını belirleyerek atağı yönlendirebilmesi, mevkinin öne çıkan temel görevleri arasında yer almaktadır.

Tablo 3.11 Standartlaştırılmış merkez orta saha karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Kevin De Bruyne	0.0058	0.0044	0.0086	0.0066	...	0.0063	0.0052	0.0058	0.0063
Luka Modrić	0.0067	0.0069	0.0033	0.0076	...	0.0075	0.0067	0.0053	0.0061
Bernardo Silva	0.0050	0.0038	0.0092	0.0066	...	0.0047	0.0032	0.0045	0.0061
Federico Valverde	0.0045	0.0038	0.0046	0.0061	...	0.0039	0.0044	0.0052	0.0057
Leon Goretzka	0.0050	0.0088	0.0066	0.0024	...	0.0059	0.0032	0.0048	0.0057

Merkez orta saha oyuncularını için hesaplanan ağırlık skorları dokuz farklı gruba ayrılmıştır. Elit bir oyuncunun değerlendirilmesinde, ilk grupta yer alan ağırlık değerleri temel alınmıştır. Bu değerlere göre sıralama şu şekildedir: FR, Çift Ayak, Pas, Tahmin, Teknik, Top Kontrolü, Karar Alma, Takım Oyunu, Soğukkanlılık, Hız, Çeviklik, Denge, Rekabet, Kararlılık, Hızlanma, Topsuz Alan, Dayanıklılık, İstikrar ve Hırs.

Tahmin yeteneği, yalnızca hücum anlarında değil, savunma sırasında da kritik öneme sahiptir. Oyun içi sezgiler sayesinde kazanılan bir top, rakibin savunmada boşluk bırakmasına yol açarak golle sonuçlanabilecek bir kontra atağı başlatabilir. Ayrıca, çeviklik, hız ve çalışkanlık gibi özellikler de bu mevki için belirleyici nitelik taşımaktadır. Merkez orta saha oyuncularını, oyunun her iki yönünde aktif rol üstlendiklerinden, futbol takımlarının iskeletini oluşturan ve oyunun akışını yönlendiren en kritik oyuncular arasında yer almaktadır.

3.2.6. Kanat (AM)

Uygulama kapsamında incelenen sekiz mevki arasında altıncı mevki, kanat olarak belirlenmiştir. Bu bağlamda, dünyanın en önde gelen 200 kanat oyuncusunun ortak özellikleri sistematik olarak analiz edilerek, elit bir futbolcуда bulunması gereken mevkiye özgü kriterler belirlenmiştir.

Tablo 3.12 Kanat mevkiisi için karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Neymar	30	12	12	16	...	6	8	34	169
Heung-Min Son	30	5	8	14	...	11	15	39	168
Sadio Mané	30	5	8	11	...	12	12	34	166
Vinicius Júnior	22	5	12	7	...	8	8	32	165
Phil Foden	22	4	12	13	...	14	16	29	161

Kanat mevkiisi için öne çıkan özellikler, oyuncunun görev aldığı takımın oyun anlayışına bağlı olarak değişkenlik göstermektedir. Bu nedenle kanat oyuncularını için farklı görev tanımları yapılabilmekte ve her bir görev kapsamında aranan nitelikler farklılık arz etmektedir. Örneğin, ters ayaklı bir kanat oyuncusundan hız ve hızlanma gibi fiziksel özellikler ile içe kat ederek şut şansı yaratıp gol bulması beklenirken; oyun kurucu rolündeki bir kanat oyuncusundan teknik becerileri yüksek, oyun görüşü gelişmiş ve orta açma yeteneği güçlü olması beklenmektedir. Bu çalışma kapsamında elde edilen analizler, söz konusu görev çeşitliliğine rağmen kanat mevkiine özgü genel ve elit kriterlerin nesnel bir biçimde ortaya konmasına olanak sağlamaktadır.

Tablo 3.13 Standartlaştırılmış kanat mevkiisi karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Neymar	0.0057	0.0071	0.0073	0.0069	...	0.0027	0.0034	0.0055	0.0060
Heung- Min Son	0.0057	0.0029	0.0049	0.0060	...	0.0049	0.0064	0.0063	0.0060
Sadio Mané	0.0057	0.0029	0.0049	0.0047	...	0.0053	0.0051	0.0055	0.0059
Vinicius Júnior	0.0041	0.0029	0.0073	0.0030	...	0.0036	0.0034	0.0052	0.0059
Phil Foden	0.0041	0.0024	0.0073	0.0056	...	0.0062	0.0068	0.0047	0.0057

Kanat mevkiisi için hesaplanan özniteliklerin ağırlık skorlarına göre sıralama şu şekildedir: FR, Çift Ayak, Teknik, Çeviklik, Hız, Dayanıklılık, Top Kontrolü, Hızlanma, Dribling, Pas, Topsuz Alan, Orta Açma, Yetenek, Soğukkanlılık, Tahmin, Karar Alma, Bitiricilik, Vizyon, Denge, Konsantrasyon, Yaş ve Uzaktan Şut.

Analiz edilen kanat oyuncularının öznitelikleri incelendiğinde, teknik beceri, çeviklik, dayanıklılık ve hızın ön plana çıktığı görülmektedir. Bu özellikler, oyuncuların ikili mücadelelerde ayakta kalabilmesini ve oyunun defansif yönünde geriye gelerek takım savunmasına katkı sağlamasını mümkün kılmaktadır. Modern futbolda çalım yeteneği, özellikle büyük takımlarda görev alan futbolcular için kritik bir nitelik olarak öne çıkmaktadır. Bu oyuncular, genellikle kapalı savunma yapan rakip ekiplerle karşılaştıklarından, teknik direktörler bu tür oyuncuları bire bir savunmak için birden fazla oyuncu görevlendirmek zorunda kalmaktadır. Ayrıca, maçın ilerleyen dakikalarında, karşı takımın savunma hattına destek sağlayacak oyuncuların gecikmesi durumunda, çalım yeteneği yüksek bir kanat oyuncusu bireysel olarak tehlike yaratabilir. Bu nedenle hız, dayanıklılık, çabukluk, teknik kalite ve denge, kanat mevkiinde elit seviyede performans sergileyen futbolcular için belirleyici kriterler olarak öne çıkmaktadır.

Bu özellikler arasında özellikle dikkat çeken parametre yaş olmuştur. Diğer mevkilere kıyasla kanat oyuncuları için yaş, performans üzerinde doğrudan belirleyici bir etkiye sahiptir; daha düşük yaş, bu mevki için önemli bir avantaj sağlamaktadır.

3.2.7. Ofansif Merkez Orta Saha (AMC)

Uygulama yapılacak olan 8 mevkiden yedinci mevkimiz ofansif merkez orta sahadır. Dünyanın en önemli 200 ofansif merkez orta saha oyuncusunun ortak özellikleri ile elit bir futbolcuda bulunması gereken mevkiye özgü kriterler belirlenecektir. Bu futbolcular arasında diğer mevkilerde de dünyanın en iyi 200 oyuncusu arasında girmiş olanlar mevcuttur. Bir veya bir kaç farklı mevkide elit seviyede oynayabilen oyuncular mevcuttur.

Tablo 3.14 Ofansif merkez orta saha mevkisi karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Kevin De Bruyne	31	7	13	14	...	16	13	36	181
Bernardo Silva	27	6	14	14	...	12	8	28	174
Bruno Fernandes	27	3	5	16	...	6	7	33	165
Thomas Müller	32	2	10	11	...	10	10	33	165
Mason Mount	23	8	6	14	...	16	17	32	162

Ofansif merkez orta saha (AMC) mevkiinde görev yapan oyuncular, genellikle teknik becerileri ve oyun kurma kapasiteleri ile ön plana çıkmaktadır. Bu mevki oyuncularından, takımlarının en yaratıcı ve teknik oyuncuları olmaları beklenmektedir. Yüksek pas yeteneği ve gelişmiş oyun görüşü, bu oyuncuların oyunu okumalarını ve hücum aksiyonlarını şekillendirmelerini mümkün kılmaktadır. Söz konusu yetenekler, ani yön değişiklikleri yapmalarına, defans arkasına beklenmedik paslar atmalarına ve sızma koşuları ile rakip defansın dengelerini bozarak hücum aksiyonlarını çeşitlendirmelerine imkân tanımaktadır.

Tablo 3.15 Standartlaştırılmış ofansif merkez orta saha karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Kevin De Bruyne	0.0057	0.0041	0.0082	0.0056	...	0.0067	0.0054	0.0057	0.006 6
Bernardo Silva	0.0050	0.0035	0.0088	0.0056	...	0.0051	0.0033	0.0044	0.006 3
Bruno Fernande s	0.0050	0.0018	0.0031	0.0064	...	0.0025	0.0029	0.0052	0.006 0
Thomas Müller	0.0059	0.0012	0.0063	0.0044	...	0.0042	0.0041	0.0052	0.006 0
Mason Mount	0.0042	0.0047	0.0038	0.0056	...	0.0067	0.0070	0.0051	0.005 9

Ofansif merkez orta saha (AMC) mevkiinde öne çıkan öznelikler sırasıyla şu şekilde belirlenmiştir: FR, Çift Ayak, Teknik, Pas, Top Kontrolü, Vizyon, Soğukkanlılık, Çeviklik, Topsuz Alan, Dribbling, Hız, Yetenek, Karar Alma, Tahmin, Hızlanma, Bitiricilik, Dayanıklılık, Orta Açma, Konsantrasyon ve Uzaktan Şut.

Bu oyuncular, toplu oyun sırasında beklenmedik anlarda verecekleri kilit paslarla rakip defans hattının dengesini bozabilecek yüksek bir vizyon kapasitesine sahip olmalıdır. Ayrıca, modern futbolda AMC oyuncularını sıklıkla gole yakın pozisyonlarda görev aldıklarından, bitiricilik ve uzaktan şut yeteneklerinin de bu mevki için kritik öznelikler arasında yer aldığı görülmektedir.

3.2.8. Santrafor (ST)

Uygulamada incelenecek sekiz mevki arasında son mevki forvet (santrafor) olarak belirlenmiştir. Bu kapsamda, dünyanın en iyi 200 santrafor oyuncusunun ortak özellikleri analiz edilerek, elit bir futbolcudaki bulunması gereken bu mevkiye özgü kriterler tespit edilecektir.

Tablo 3.16 Forvet mevkiisi için karar matrisi

İsim	Yaş	Sakatlığa Yatkınlık	İtiraz	Kornerler	...	Sportmenlik	Duygu Kontrol	Çift Ayak	FR
Erling Haaland	22	12	12	7	...	10	12	31	185
Harry Kane	28	12	12	9	...	11	16	39	183
Robert Lewandowski	33	2	15	3	...	10	11	33	183
Karim Benzema	34	8	17	10	...	17	18	37	183
Kylian Mbappé	23	7	8	13	...	12	13	30	172

Bu mevki içerisinde birden fazla rol bulunmaktadır ve bu roller büyük ölçüde teknik direktörlerin oyun stratejilerine bağlı olarak şekillenmektedir. Örneğin, 2 metre boyundaki bir santrafor, uzun topları arkadaşlarına indirerek oyunu başlatan bir bağlantı oyuncusu olarak görev alabilirken; fiziksel olarak daha kısa boylu bir

santrafor, orta sahaya kadar gelerek kendisini markaja alan savunma oyuncusunu ileri çekebilir ve boşalan alan, takım arkadaşlarının sızma koşullarıyla tehlikeli ataklara dönüştürülebilir. Bu nedenle, santrafor mevkiinde oynayan oyuncular yalnızca gol atma becerisine sahip olmakla kalmayıp, topsuz alanda da etkili olabilecek teknik ve taktik yeteneklere sahip olmalıdırlar.

Tablo 3.17 Standartlaştırılmış forvet mevkiisi karar matrisi

İsim	Yaş	Sakatlığ a Yatkınlı k	İtiraz	Kornerle r	...	Sportmenli k	Duygu Kontrol	Çift Ayak	FR
Erling Haaland	0.0039	0.0074	0.0068	0.0049	...	0.0042	0.0051	0.0049	0.0066
Harry Kane	0.0050	0.0074	0.0068	0.0063	...	0.0047	0.0068	0.0062	0.0065
Robert Lewandowski	0.0059	0.0012	0.0085	0.0021	...	0.0042	0.0047	0.0053	0.0065
Karim Benzema	0.0061	0.0050	0.0097	0.0070	...	0.0072	0.0077	0.0059	0.0065
Kylian Mbappé	0.0041	0.0043	0.0046	0.0091	...	0.0051	0.0055	0.0048	0.0061

Yapılan hesaplamalar sonucunda, elit bir santraforun sahip olması gereken özellikler sırasıyla şu şekilde belirlenmiştir: FR, Çift Ayak, Topsuz Alan, Bitiricilik, Tahmin, Teknik, Karar Alma, Top Kontrolü, Pas, Dayanıklılık, Dribbling, Konsantrasyon, İstikrar, Çeviklik, Soğukkanlılık, Hız, Denge, Vizyon, Yaş, Kafa Vuruşu ve Vücut Zindeliği.

Giriş bölümünde vurgulandığı üzere, topsuz alanda etkin olabilme yeteneği santraforlar için kritik öneme sahiptir. Bununla birlikte, bitiricilik becerisi de bu mevki için vazgeçilmez bir kriterdir. Santraforlar, rakip savunma oyuncularının baskısı altında doğru ve kararlı bir gol vuruşu yapabilmek durumundadır. Rakip oyuncular, şutu engelleyemedikleri noktada santraforun dengesini bozarak vuruşunun etkinliğini azaltmaya çalışacaklardır.

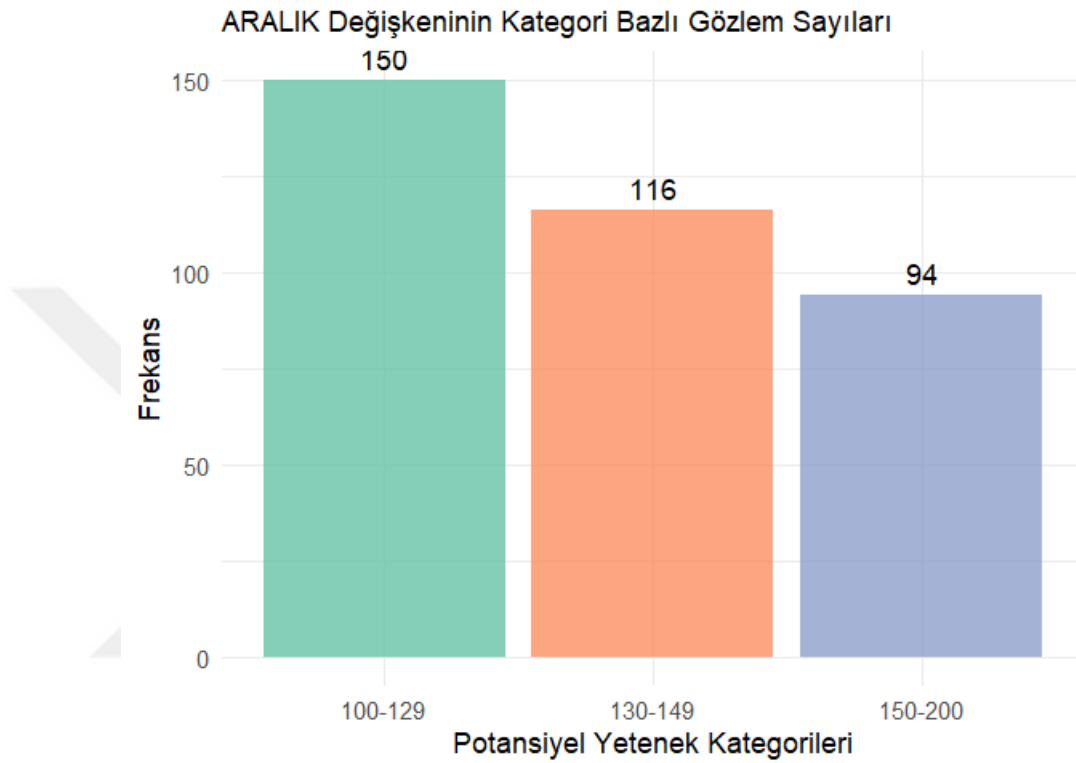
Elit bir santraforun güçlü bir fiziksel yapıya ve yüksek kondisyona sahip olması da gerekmektedir. Hesaplamalar, bu mevki için yaşın düşük olmasının bir avantaj olduğunu göstermektedir. Futbolda belirli bir yaştan sonra kondisyonu maç süresince yüksek seviyede sürdürmek ve topsuz alanda sürekli koşularla boşa çıkmak mümkün olamamaktadır. Ayrıca, maç boyunca sık sık girilen ikili mücadeleler de oyuncuyu fiziksel olarak yıpratmaktadır. Bu nedenle, dayanıklılık, sabır ve yüksek konsantrasyon, santraforların başarılı olabilmesi için kritik faktörler arasında yer almaktadır. Oyuncu, yarım pozisyon olarak tanımlanabilecek fırsatları iyi sezebilmeli ve doğru zamanda doğru yerde olarak etkili bir vuruş yapabilmelidir.

3.3. Makine Öğrenmesi Uygulaması

Makine öğrenmesi aşamasında temel amaç, genç futbolcuların potansiyel performanslarının doğru bir şekilde tahmin edilebilmesidir. Elit futbolcuların yeteneklerinin analizinde kullanılan FR (Futbolcu Reytingi) değişkeni, futbolcuların mevki bazında sıralanması ve kıyaslanmasında referans olarak değerlendirilmiştir. Bu çerçevede, Python programlama dili kullanılarak MAUT yöntemi ile değişken eleme süreci gerçekleştirilmiş ve 50 öznitelik arasından her mevki için ayırt edici özellikler belirlenmiştir. Belirlenen bu öznitelikler, her mevki için hangi futbolcu özelliklerinin öncelikli olması gerektiğine dair sistematik bir analiz yapılmasını mümkün kılmıştır. MAUT yönteminin uygulanmasına ilişkin Python kodları da çalışmada ek olarak sunulmaktadır.

Makine öğrenmesi algoritmalarında ise bu değişkenler bağımsız değişkenler olarak kullanılacak ve amaç, PR (Potansiyel Reyting) aralığını tahmin etmektir. Araştırmada sekiz mevki arasından yalnızca “Kanat” mevki üzerine odaklanılmıştır. Bu mevki için veri seti, 360 kanat oyuncusunu temsil eden satırlar ve MAUT uygulamasından elde edilen 20 ayırt edici özniteliği temsil eden sütunlardan oluşmaktadır. Elit düzeyde futbol oynayan sporcuların genç yaşlarda sahip oldukları beceri skorları bağımsız değişkenler olarak kullanılırken, hâlihazırda ulaşmış oldukları reyting değerleri potansiyel reyting (PR) olarak ele alınmıştır.

Araştırmanın temel hedefi, 21 yaş ve altındaki genç futbolcuların sahip oldukları yetenekler ve beceri skorları doğrultusunda, ilerleyen dönemlerde ulaşabilecekleri potansiyel reyting değerlerinin tahmin edilebilmesini sağlamaktır.



Şekil 3.1 Bağımlı değişkenin frekans dağılımı

Bağımlı değişken olan potansiyel yetenek kategorilerinin (ARALIK) frekans dağılımı incelendiğinde, veri setindeki futbolcu potansiyel dağılımının hafif negatif çarpıklık (negatively skewed) sergilediği gözlemlenmektedir. Bu dağılım, beklenen piramit yapısıyla uyumlu olup, en yüksek frekans ($n=150$) '100-129' aralığındaki, yani Beş Büyük Lig dışındaki liglerde oynama potansiyeline sahip futbolcularda yoğunlaşmaktadır.

Potansiyel yetenek seviyesi arttıkça, bu seviyeye sahip futbolcu sayısının istatistiksel olarak anlamlı biçimde azaldığı görülmektedir. '130-149' aralığındaki, Beş Büyük Lig standardındaki futbolcuların sayısı ($n=115$) bir önceki gruba kıyasla düşüş

göstermekte; '150-200' aralığındaki, elit seviyede veya yıldız adayı olarak nitelendirilebilecek futbolcular (n=94) ise veri setinde en seçkin ve nadir grubu oluşturmaktadır. Bu dağılım, elit sporcuların doğal olarak daha nadir bulunması gerçeğini yansıtarak, modelin üst potansiyelli futbolcuları ayırt etme zorluğunu da ortaya koymaktadır.

Oluşturulan veri setinden “Çift Ayak”, “Yaş” ve “FR” değişkenleri çıkarılmıştır. Bunun temel nedeni, tüm mevki gruplarında çift ayak kullanma becerisinin üst sıralarda yer almasıdır; mevki fark etmeksizin elit bir futbolcu için iki ayağını etkin kullanabilme temel bir koşuldur. Çalışmada odak, genç futbolcular olduğundan, 21 yaş altındaki sporcular seçilmiş ve yaş değişkeni ayırt edici olmadığı için eğitim aşamasında modele dahil edilmemiştir.

Ayrıca, FR değeri yerine bağımlı değişken olarak PR (Potansiyel Reyting) skoru kullanılmıştır. Bu düzenleme, modelin genç futbolcuların gelecekteki potansiyel performanslarını tahmin etmesine doğrudan katkı sağlamaktadır.

3.3.1. Rassal Ormanlar (Random Forest)

Çalışmanın bu bölümünde futbolcu potansiyel tahmini için ilk olarak topluluk (ensemble) öğrenme yöntemlerinden Rassal Orman algoritması kullanılmıştır. Model, %70 eğitim ve %30 test olarak ikiye ayrılmıştır, ayrılan veri seti 500 ağaç (ntree) ve $\sqrt{p}$ (p: özellik sayısı) mtry değeri ile eğitilmiştir.

Performans metrikleri incelendiğinde, Random Forest'ın 0.6641 Kappa değeri ile şans faktörünü başarıyla eleyebildiği ve %80.12 - %93.37 güven aralığı ile istatistiksel olarak güvenilir tahminler üretebildiği gözlemlenmiştir. Karışıklık matrisi analizinde modelin özellikle "Orta" potansiyel sınıfında %87.85 doğruluk oranıyla oldukça başarılı olduğu, "Düşük" ve "Yüksek" potansiyel sınıflarında ise daha düşük ancak kabul edilebilir seviyede doğruluk sağladığı tespit edilmiştir.

3.3.2 XGBoost(Extreme Gradient Boosting)

Uygulamanın ikinci bölümünde gradient boosting yöntemlerinin en gelişmiş temsilcilerinden biri olan XGBoost (Extreme Gradient Boosting) algoritması

kullanılmıştır. Model, eğitim seti üzerinde $\eta = 0.3$, $\max_depth = 6$, $\gamma = 0$ gibi varsayılan parametrelerle eğitilmiş ve %84.11 doğruluk oranı elde edilmiştir.

Performans metrikleri detaylı incelendiğinde, modelin 0.6123 Kappa değeri ile Rassal Ormanlar'a kıyasla şans faktörünü elemeye daha düşük performans sergilediği, %76.45 - %90.22 güven aralığının ise diğer modellere göre daha geniş olduğu gözlemlenmiştir. Karışıklık matrisi analizinde XGBoost'un özellikle "Yüksek" potansiyel sınıfında diğer modellere kıyasla daha fazla yanlış sınıflandırma yaptığı, bunun olası nedeninin sınıf dengesizliği karşısında modelin yeterli regularizasyon uygulayamaması olabileceği değerlendirilmiştir.

3.3.3 Destek Vektör Makineleri (Support Vector Machine)

Uygulamanın üçüncü algoritması olarak istatistiksel öğrenme teorisi temelli bir algoritma olan Destek Vektör Makineleri (SVM) kullanılmıştır. Model, radyal tabanlı fonksiyon (RBF) kernel kullanılarak eğitilmiş ve varsayılan parametrelerle ($C = 1$, $\sigma = 0.1$) %86.92 doğruluk oranı elde edilmiştir.

Performans metrikleri detaylı incelendiğinde, modelin 0.6528 Kappa değeri ile Random Forest'a çok yakın bir performans sergilediği, %79.45 - %92.18 güven aralığının ise istatistiksel olarak güvenilir bir profil çizdiği gözlemlenmiştir. Karışıklık matrisi analizinde SVM'nin özellikle "Orta" potansiyel sınıfında %84.6 doğruluk oranıyla oldukça başarılı olduğu, "Düşük" ve "Yüksek" potansiyel sınıflarında ise dengeli bir dağılım sergilediği tespit edilmiştir.

Kernel seçiminin kritik önemi göz önünde bulundurularak yapılan ön testlerde, RBF kernel'in doğrusal ve polinom kernel'lere kıyasla daha iyi performans gösterdiği ve futbol verisindeki karmaşık ilişkileri daha iyi yakalayabildiği anlaşılmıştır. Sonuç olarak, SVM algoritmasının bu çalışma kapsamında yüksek ve istikrarlı bir performans sergilediği, eğitilen modelin Random Forest'a çok yakın doğruluk oranı elde ettiği ve futbolcu potansiyel tahmini için güvenilir bir alternatif oluşturduğu değerlendirilmiştir.

3.3.4 En Yakın Komşu (k-NN)

Dördüncü algoritma olarak k-En Yakın Komşu (k-NN) yöntemi uygulanmıştır. Model, Öklid uzaklık metriği kullanılarak ve veri setinin k-NN için kritik öneme sahip olan ölçeklendirme (standardizasyon) işleminden geçirilmesinin ardından eğitilmiştir.

Performans metrikleri detaylı incelendiğinde, modelin 0.6384 Kappa değeri ile diğer modellere kıyasla orta düzeyde bir performans sergilediği, %78.52 - %90.91 güven aralığının ise kabul edilebilir seviyede olduğu gözlemlenmiştir. Karışıklık matrisi analizinde k-NN'nin özellikle sınır değerlere sahip örneklerde (potansiyeli "Düşük" ve "Yüksek" sınıflarının sınırlarındaki oyuncular) daha fazla hata yaptığı, bunun olası nedeninin uzaklık temelli algoritmaların sınıf sınırlarını keskin şekilde ayırt etmekte zorlanması olabileceği değerlendirilmiştir. Uzaklık metriği seçiminin algoritma performansı üzerindeki etkisi test edilmiş ve Öklid uzaklığının Manhattan ve Minkowski uzaklıklarına kıyasla daha iyi sonuç verdiği tespit edilmiştir.

Sonuç olarak, k-NN algoritmasının bu çalışma kapsamında basitliğine rağmen oldukça rekabetçi bir performans sergilediği, özellikle hızlı prototipleme ve ön değerlendirme çalışmaları için pratik bir çözüm sunduğu, ancak nihai model seçiminde daha karmaşık algoritmaların gerisinde kaldığı değerlendirilmiştir.

BÖLÜM DÖRT

SONUÇ

Bu çalışma, Football Manager 2023 veri tabanındaki futbolcu özelliklerini kullanarak denetimli makine öğrenmesi algoritmalarıyla futbolcuların potansiyel performans aralıklarının tahmin edilmesini amaçlamıştır. Çalışma kapsamında Random Forest, XGBoost, Destek Vektör Makineleri ve k-En Yakın Komşu algoritmaları karşılaştırmalı olarak test edilmiş ve aşağıdaki sonuçlara ulaşılmıştır.

Tablo 4.1 Makine öğrenmesi yöntemlerinin performans karşılaştırması

Algoritma	Doğruluk (%)	Kappa	%95 GA (Alt)	%95 GA (Üst)
Rassal Ormanlar	87.85	0.6641	80.12	93.37
Destek Vektör Makineleri	86.92	0.6528	79.45	92.18
K-En Yakın Komşular	85.98	0.6384	78.52	90.91
XGBoost	84.11	0.6123	76.45	90.22

Yapılan kapsamlı analizler sonucunda, dört farklı algoritma içerisinde Random Forest (Varsayılan) modeli %87.85 doğruluk, 0.6641 Kappa değeri ve %80.12-93.37 güven aralığı ile en yüksek performansı sergilemiştir.

Çalışmanın en önemli pratik katkısı, futbol kulüplerinin yetenek keşfi ve transfer stratejilerinde veriye dayalı karar destek sistemi olarak kullanılabilecek %87.85 doğruluğa sahip bir tahmin modeli sunmasıdır. Bu bulgular, geleneksel scoutluk faaliyetlerinde sezgisel olarak kullanılan kriterlerin bilimsel olarak doğrulanması anlamına gelmektedir.

Bu çalışmanın metodolojik açıdan bir diğer katkısı, video oyunu verilerinin futbol analitiğinde geçerli ve güvenilir bir veri kaynağı olarak kullanılabileceğini göstermesidir. Football Manager veri tabanının 1.300'ü aşkın uzman tarafından oluşturulması ve sürekli güncellenmesi, veri kalitesi ve güvenilirliği açısından önemli bir avantaj sağlamaktadır. Ayrıca, MAUT yöntemi ile yapılan mevki bazlı özellik seçimi, futbolcu performans değerlendirmesinde standartlaştırılmış ve sistematik bir metodolojik yaklaşım sunmaktadır.

Çalışmanın temel kısıtlamalarından biri, analizlerin yalnızca Football Manager veri tabanı ile sınırlı kalmasıdır. Gelecek araştırmalarda şu geliştirmeler önerilmektedir:

- Gerçek hayat performans verileri ile entegrasyon,
- Derin öğrenme modellerinin uygulanması ve test edilmesi,
- Farklı ligler ve kültürler arası karşılaştırmalı analizlerin yapılması,
- Zaman serisi analizleri ile futbolcuların kariyer gelişim modellerinin incelenmesi.

Buna ek olarak, model değerlendirme sürecinde F1-score ve ROC eğrileri gibi dengesiz veri setleri için daha duyarlı metriklerin kullanılması, değerlendirmelerin güvenilirliğini artıracaktır.

Sonuç olarak, bu çalışma hem akademik hem de endüstriyel açıdan katkılar sunmaktadır. Akademik olarak, makine öğrenmesi algoritmalarının spor bilimlerinde uygulanabilirliğini göstermiş ve yeni bir metodolojik yaklaşım önermiştir. Endüstriyel açıdan ise futbol kulüplerinin milyonlarca dolarlık transfer kararlarını daha güvenilir ve veriye dayalı bir şekilde alabilmelerine imkân sağlayacak bir altyapı sunmaktadır. Özellikle Rassal Ormanlar (Random Forest) algoritmasının çalışma kapsamındaki performansı, topluluk (ensemble) yöntemlerinin spor analitiği alanındaki potansiyelini göstermektedir.

KAYNAKLAR

- Al-Asadi, M. A. (2018). *Decision support system for a football team management by using machine learning techniques*. KONYA.
- Alpaydın, E. (2020). *Introduction to machine learning (4. Baskı)*. The MIT Press.
- Altman, D. G. (2000). *Statistics with confidence: Confidence intervals and statistical guidelines (2. Baskı)*. London: BMJ Books.
- Ananda, J. &. (2005). Evaluating public risk preferences in forest land-use choices using multi-attribute utility theory. *Ecological Economics*, 408–419.
- Association, F. I. (2023). *FIFA Big Count: Football's biggest survey*. FIFA.com: <https://www.fifa.com> adresinden alındı
- Becerenler, S. (2023). *Entropi temelli maut ve saw yöntemleri ile ülkelerin ekonomik özgürlüklerine ilişkin bir örnek uygulama*. ANKARA.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Boon, B. H. (2010). Scientific procedures for player selection in professional soccer. *Data and Soccer*. içinde
- Booth, R. (2018, Kasım 16). *The football detectives: The unseen army of experts behind Football Manager*. The Guardian: <https://www.theguardian.com/games/2018/nov/16/football-manager-2019-experts-scouts-data-researchers> adresinden alındı
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 123–140.
- Breiman, L. (2001). Random forests. *Machine Learning*, 5–32.
- Canbolat, Y. B. (2007). Combining decision tree and MAUT for selecting a country for a global manufacturing facility. *Omega*, s. 312–325.
- Chen, J. (2023). Data-driven player recruitment in football: A cluster analysis approach. *Journal of Sports Analytics*, s. 9(2), 45-67.

- Chen, T. &. (2016). XGBoost: A scalable tree boosting system. *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (s. 785–794).
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 37–46.
- Cover, T. M. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 21–27.
- Cumming, G. &. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, 170–180.
- Deloitte. (2024). *Football Money League 2024*. Deloitte.
- Fishburn, P. C. (1967). Methods of estimating additive utilities. *Management Science*, 435-453.
- Güleş, Ş. (2020). *Makine öğrenmesi yöntemleri ile zararlı yazılım tespiti*.
- Han, J. K. (2012). *Data mining: Concepts and techniques (3. Baskı)*. Elsevier.
- Hastie, T. T. (2009). *The elements of statistical learning: Data mining, inference, and prediction (2. Baskı)*. Springer.
- Interactive, S. (2023). *Football Manager 2023*. Sega.
- Keeney, R. L. (1974). Multiplicative utility functions. J. W. Sons içinde, *Decisions with multiple objectives: Preferences and value trade-offs* (s. 22-34).
- Kelleher, J. D. (2015). *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*. MIT Press.
- Landis, J. R. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174.
- Loken, E. (2007). Use of multi-criteria decision analysis methods for energy planning problems. *Renewable and Sustainable Energy Reviews*, 1584-1595.

- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 276–282.
- Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.
- Nahm, F. S. (2016). Confidence intervals for the mean of non-normally distributed data: Comparison of different methods. *Anesthesia & Pain Medicine*, 9–14.
- Öztürk, M. (2022, Nisan 13). *Miraç Öztürk*. <https://miracozturk.com/python-ile-siniflandirma-analizleri-rastgele-orman-random-forest-algoritmasi/> adresinden alındı
- Paksoy, T. (2019). *MAUT Yöntemi ile Türkiye’deki Önemli Firmaların Kurumsal Sürdürülebilirlik Performans Ölçümünün Karşılaştırılması*. İZMİR.
- Pappalardo, L. C. (2019). PlayerRank: Data-driven performance evaluation and player ranking in soccer via a machine learning approach. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 1-27.
- Powers, D. M. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 37–63.
- Provost, F. &. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.
- Robberechts, P. &. (2020). Who wants to play? A machine learning approach to football player selection. *Journal of Sports Sciences*, 1-9.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 210-229.
- Sayan, V. H. (2023). *Makine öğrenmesi yöntemleri ile futbol oyuncularının performans analizi*. BURDUR.
- Topcu, D. (2023). *Sınıflandırma için makine öğrenim yöntemlerinin performans değerlendirmesi: Futbol verisi uygulaması*. İZMİR.

- Vapnik, C. &. (1995). Support-vector networks. *Machine Learning*, 273–297.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer-Verlag.
- Viera, A. J. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, 360–363.
- Von Winterfeldt, D. &. (1986). *Decision analysis and behavioral research*. Cambridge University Press.
- Wijk, W. V. (2022). *Machine Learning Applications in Football Player Valuation: A Data-Driven Approach to Transfer Market Economics* Tilburg Üniversitesi. Tilburg.
- Yıldız, T. &. (2019). Futbolda oyuncu performans puanlaması için çok kriterli karar verme yöntemlerinin kullanılması: TOPSIS ve MAUT uygulaması. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 1125-1140.
- Yücel, H. M. (2023). *Kurumsal yönetim endeksi bist30 şirketlerinin çkkv tekniklerinden gri ilişkisel analiz yöntemi, edas, topsis ve maut yöntemine göre sıralanması*. İSTANBUL.
- Zionts, S. (1979). MCDM. *Interfaces*, 94-101.

EK 1: MAUT YÖNTEMİ “DM” MEVKİSİ İÇİN PYTHON KODLARI

```
import pandas as pd
```

```
import numpy as np
```

```
data = pd.read_csv('C:\\Users\\Burak\\Desktop\\Veri Bilimi\\Tez\\MAUT  
DATA\\DM.csv')
```

```
print(data.head())
```

```
df = pd.DataFrame(data)
```

```
df.to_excel('data.xlsx')
```

```
df1 = data
```

```
print(df1.head())
```

```
df1.iloc[:, 1:] = df1.iloc[:, 1:] / df1.iloc[:, 1:].sum()
```

```
print(df1.head())
```

```
class MultiCriteria:
```

```
    def __init__(self, database):
```

```
        self.database = database
```

```

def normalize_columns(self):

    high_database = self.database.copy()

    for column in self.database.columns[4:]:

        high_database[column] = (self.database[column] -
self.database[column].min()) / (self.database[column].max() -
self.database[column].min())

    return high_database

def highlow_age_injury_argue(self):

    high_database = self.database.copy()

    list_highlow_j = []

    for j in ['Age', 'Injury', 'Argue']:

        high_database[j] = (self.database[j] - self.database[j].max()) /
(self.database[j].min() - self.database[j].max())

        list_highlow_j.append(high_database[j])

    result_highlow_j = pd.DataFrame(list_highlow_j).transpose()

    return result_highlow_j

def weighted_entropy(self):

```

```

# Tüm sütunların ağırlıklarını hesaplama

entropy_dataset = np.zeros(len(self.database.columns) - 1) # Sadece İsim
sütunu hariç

# Tüm sütunların ağırlıklarını hesaplama

for i, column in enumerate(self.database.columns[1:]): # İsim sütunu hariç

    entropy_dataset[i] = (np.log(self.database[column]) *
self.database[column]).sum()

    entropy_dataset[i] = (-1) / (np.log(len(self.database[column])) *
entropy_dataset[i])

    entropy_dataset[i] = 1 - entropy_dataset[i]

entropy_dataset = entropy_dataset / sum(entropy_dataset)

return entropy_dataset

def score(self):

    highlow = self.highlow()

    weighted_ent = self.weighted_entropy()

    list_score = []

    for i in range(len(highlow)):

```

```

score_result = highlow.iloc[:, i] * weighted_ent[i]

list_score.append(score_result)


result_score = pd.concat(list_score, axis=1)

score_result = result_score.sum(axis=1)

return score_result


def show_normalized_table(self):

    print("Normalize Edilmiş Tablo:")

    print(self.database)


def show_weighted_entropy_sorted(self):

    print("\nAğırlık Değerleri ve Sütun İsimleri (Büyükten Küçüğe Sıralı):")

    weighted_ent = self.weighted_entropy()

    if len(self.database.columns) - 1 != len(weighted_ent):

        print("Hata: Ağırlık Değerleri ve Sütun Sayısı Uyumsuz")

    return


# Ağırlık değerlerini ve sütun isimlerini bir veri çerçevesine dönüştürme

weighted_df = pd.DataFrame({"Sütun": self.database.columns[1:], "Ağırlık Değeri": weighted_ent})

```

```

# Ağırlık değerlerine göre büyükten küçüğe sıralama

sorted_weighted_df = weighted_df.sort_values(by='Ağırlık Değeri',
ascending=False)

# Satır numaralarını ekleyerek indeks sütununu oluşturma

sorted_weighted_df.insert(0, 'Satır', range(1, len(sorted_weighted_df) + 1))

# Sıralanmış çıktıyı yazdırma

print(sorted_weighted_df.to_string(index=False))

# DataFrame'i Excel dosyasına yazma

sorted_weighted_df.to_excel('sorted_weighted_entropy.xlsx', index=False)

print("Excel dosyasına yazma işlemi tamamlandı:
sorted_weighted_entropy.xlsx")

def show_weighted_utility_matrix(self):

    print("\nAğırlık ile Çarpılmış Fayda Matrisi:")

    highlow = self.highlow()

    weighted_ent = self.weighted_entropy()

    list_score = []

```

```

        for i in range(len(self.database)): # Her bir gözlem değeri için döngü

            score_result = highlow.iloc[i, :] * weighted_ent # Her bir gözlem değeri
            ile ağırlık değerlerini çarp

            list_score.append(score_result)


        result_score = pd.concat(list_score, axis=1).transpose()

        result_score.index = self.database["İsim"] # Satırların futbolcuları temsil
        ettiğini varsayalım

        print(result_score)


    def show_scores_sorted(self):

        scores = self.score()

        scores.index = self.database["İsim"] # Skorların indekslerini futbolcu
        isimleriyle değiştir

        sorted_scores = scores.sort_values(ascending=False) # Skorları büyükten
        küçüğe sırala

        print(sorted_scores)

    mc = MultiCriteria(df1)

    mc.show_normalized_table()

    mc.show_weighted_entropy_sorted()


df = pd.DataFrame(df1)

```

```
df.to_excel('standartlastirilmisdata.xlsx')
```

