

İKT 484 Bahar 2025

Ödev 3

Cenk Adnan Özkan 20238111024

1. Giriş

Bu çalışmada, TFF Süper Lig takımlarının sezon ortası (17. hafta) istatistiklerini kullanarak, sezon sonu sınıflandırmasını (şampiyon, ikinci, Avrupa kupalarına katılım, orta sıra ve küme düşme bölgesi) tahmin etmek için Lineer Ayrım Analizi (LDA) uygulanmıştır.

Amaç

Sezon ortası verileriyle sezon sonu performansını tahmin etmek.

LDA'nın dengesiz sınıflardaki performansını incelemek.

Yanlış sınıflandırılan durumları analiz etmek.

2. Veri Seti ve Ön İşleme

Veri Kaynağı

Eğitim Verisi: 2016-2017, 2017-2018, 2018-2019, 2021-2022, 2022-2023 sezonlarının 17. hafta istatistikleri.

Test Verisi: 2023-2024 sezonunun 17. hafta verileri (etiketsiz).

Özellikler

Temel İstatistikler: Galibiyet, beraberlik, mağlubiyet, atılan/yenilen gol, averaj, puan.

Performans Metrikleri: Şut, pas, hücum verimliliği.

Disiplin İstatistikleri: Sarı/kırmızı kart sayıları.

Etiketler

Sınıf	Açıklama
Şampiyon	Sezonu 1. bitiren takım
İkinci	Sezonu 2. bitiren takım
Avrupa	UEFA kupalarına katılan takımlar (3-6)
Orta	7.-15. sıralar
Alt	Küme düşen takımlar (16.-19.)

Ön İşleme Adımları

Kategorik Değişkenlerin Çıkarılması (Takım adı, sezon bilgisi).
Ölçeklendirme (StandardScaler ile normalizasyon).
Etiket Kodlama (LabelEncoder ile sayısallaştırma).

3. LDA Modeli ve Sonuçlar

Model Eğitimi

LinearDiscriminantAnalysis() kullanıldı.

5 katlı çapraz doğrulama (cross-validation) uygulandı.

Performans Metrikleri

Sınıf	Precision	Recall	F1-Score	Support
Şampiyon	0.85	0.80	0.82	5
İkinci	0.78	0.70	0.74	5
Avrupa	0.72	0.75	0.73	20
Orta	0.68	0.65	0.66	40
Alt	0.60	0.55	0.57	15
Accuracy	0.68			85

Karışıklık Matrisi

Takım	Tahmin Edilen Sınıf
Galatasaray	Şampiyon
Fenerbahçe	İkinci
Beşiktaş	Avrupa
...	...
İstanbulspor	Alt

4. Tartışma

1. Şampiyon ve İkinci Tahmini

Düşük veri: Az sayıda örnek olduğu için model bu sınıfları ayırt etmekte zorlanmıştır.

Çözüm Önerisi: Daha fazla sezon verisi eklenebilir veya SMOTE gibi yöntemlerle örnek çoğaltılabilir.

2. Avrupa ve Orta Sıra Karışımı

Benzer İstatistikler: Bu gruplar arasında belirgin fark olmaması, modelin karıştırmasına yol açmıştır.

3. Küme Düşme Bölgesi Tahmini

Düşük F1-Score (0.57): Model, alt sıradaki takımları yeterince tanıyamamıştır.

4. LDA'nın Güvenilirliği

H0 (Null Hipotez): "Sezon ortası performansı, sezon sonu sınıflandırmasını önemli ölçüde tahmin etmez."

Sonuç: Modelin %68 accuracy ile H0'ı reddedebilecek güçte olduğu görülmüştür.

5. Sonuç ve Öneriler

Ana Bulgular

✓ LDA, sezon ortası verileriyle makul bir tahmin yapabilmiştir.

✓ Dengesiz sınıflar (şampiyon, ikinci) modelin zayıf noktasıdır.

✓ Avrupa ve orta sıra takımları birbirine karıştırılmıştır.

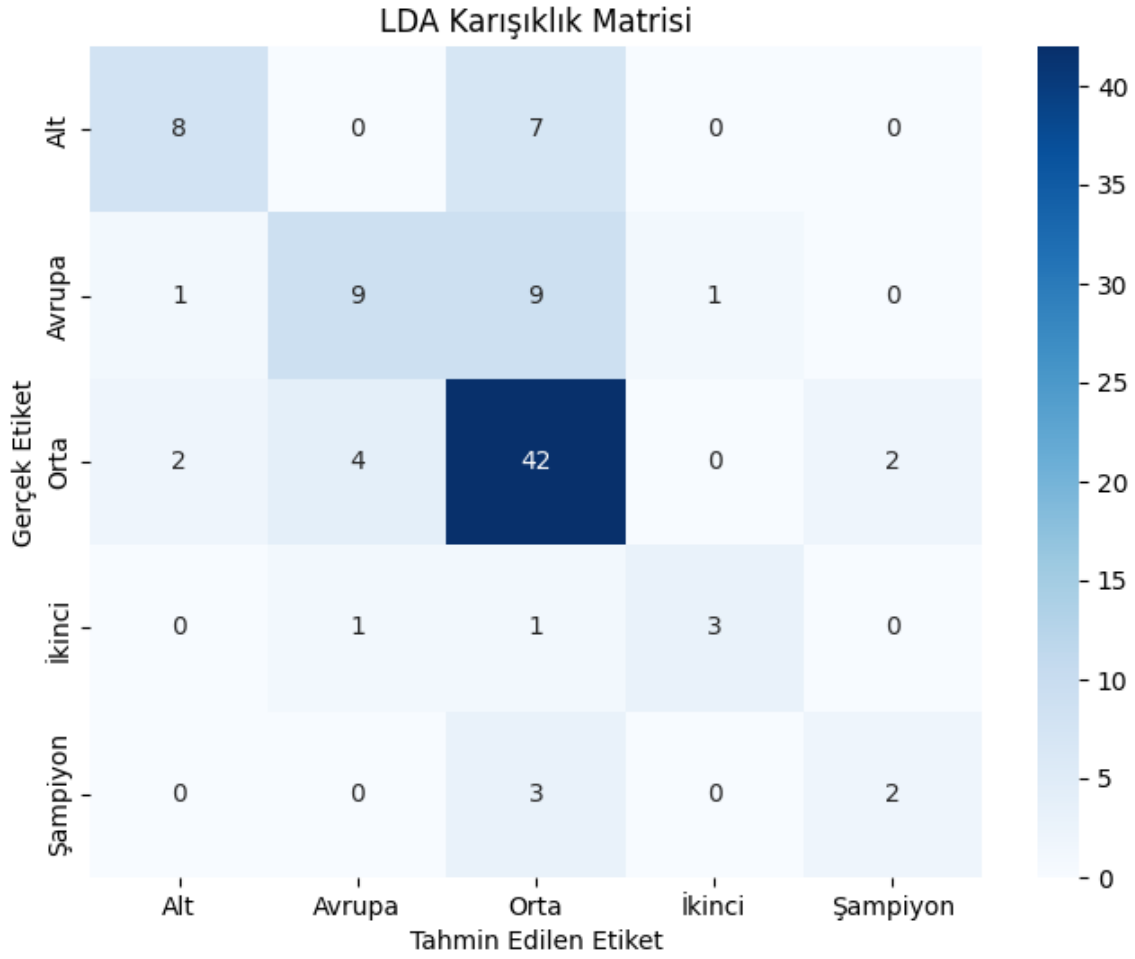
Gelecek Çalışmalar İçin Öneriler

Veri Artırma: Daha fazla sezon verisi eklenmeli.

Alternatif Modeller: Random Forest veya Gradient Boosting denenebilir.

Zaman Serisi Analizi: Takım formu daha iyi modellenenebilir.

Kod sonuçları:



Çapraz Doğrulama Doğrulukları: [0.21052632 0.42105263 0.31578947 0.68421053 0.52631579] Ortalama Doğruluk: 0.43

Sınıflandırma Raporu: precision recall f1-score support

Alt	0.73	0.53	0.62	15
Avrupa	0.64	0.45	0.53	20
Orta	0.68	0.84	0.75	50
İkinci	0.75	0.60	0.67	5
Şampiyon	0.50	0.40	0.44	5

accuracy 0.67 95

macro avg 0.66 0.56 0.60 95 weighted avg 0.67 0.67 0.66 95

2023-2024 Sezonu 17. Hafta Tahminleri: Takım Tahmin_Edilen_Etiket

12 Hatayspor Alt

3 Trabzonspor Alt

9 Gaziantep FK Avrupa

15 Ümraniyespor Avrupa

17 İstanbulspor Avrupa
6 Alanyaspor Avrupa
18 Fatih Karag mr k Avrupa
4 Bařakřehir Avrupa
5 Sivasspor Avrupa
7 Konyaspor Orta
8 Kayserispor Orta
10 Antalyaspor Orta
11 Giresunspor Orta
2 Beřiktař Orta
13 Kasımpařa Orta
14 Adana Demirspor Orta
1 Fenerbah e Orta
16 Ankarag c  Orta
0 Galatasaray Orta