



Makine Öğrenmesi ile Yetenek
Avcılığı Sınıflandırma

miuul

Scout'lar tarafından izlenen futbolcuların özelliklerine verilen puanlara göre, oyuncuların hangi sınıf (average, highlighted) oyuncu olduğunu tahminleme.

Scout: Gözlemci ya da yetenek avcısı, gelecek vadettiği düşünülen sporcuları gözlemleyerek mevcut yeteneklerini ve potansiyellerini tespit eden uzman kişi

Veri Seti Hikayesi

Veri seti Scoutium'dan maçlarda gözlemlenen futbolcuların özelliklerine göre scoutların değerlendirdikleri futbolcuların, maç içerisinde puanlanan özellikleri ve puanlarını içeren bilgilerden oluşmaktadır.

scoutium_attributes.csv

8 Değişken

10.730 Gözlem

527 KB

task_response_id	Bir scoutun bir maçta bir takımın kadrosundaki tüm oyunculara dair değerlendirmelerinin kümesi
match_id	İlgili maçın id'si
evaluator_id	Değerlendiricinin(scout'un) id'si
player_id	İlgili oyuncunun id'si
position_id	İlgili oyuncunun o maçta oynadığı pozisyonun id'si 1: Kaleci 2: Stoper 3: Sağ bek 4: Sol bek 5: Defansif orta saha 6: Merkez orta saha 7: Sağ kanat 8: Sol kanat 9: Ofansif orta saha 10: Forvet
analysis_id	Bir scoutun bir maçta bir oyuncuya dair özellik değerlendirmelerini içeren küme
attribute_id	Oyuncuların değerlendirildiği her bir özelliğin id'si
attribute_value	Bir scoutun bir oyuncunun bir özelliğine verdiği değer(puan)

scoutium_potential_labels.csv

5 Değişken

322 Gözlem

12 KB

task_response_id	Bir scoutun bir maçta bir takımın kadrosundaki tüm oyunculara dair değerlendirmelerinin kümesi
match_id	İlgili maçın id'si
evaluator_id	Değerlendiricinin(scout'un) id'si
player_id	İlgili oyuncunun id'si
potential_label	Bir scoutun bir maçta bir oyuncuyla ilgili nihai kararını belirten etiket. (hedef değişken)



Proje Görevleri

Görevler

Adım 1: scoutium_attributes.csv ve scoutium_potential_labels.csv dosyalarını okutunuz.

Adım 2: Okutmuş olduğumuz csv dosyalarını merge fonksiyonunu kullanarak birleştiriniz.

("task_response_id", 'match_id', 'evaluator_id' "player_id" 4 adet değişken üzerinden birleştirme işlemini gerçekleştiriniz.)

Adım 3: position_id içerisindeki Kaleci (1) sınıfını veri setinden kaldırınız.

Adım 4: potential_label içerisindeki below_average sınıfını veri setinden kaldırınız.(below_average sınıfı tüm verisetinin %1'ini oluşturur)

Adım 5: Oluşturduğunuz veri setinden “pivot_table” fonksiyonunu kullanarak bir tablo oluşturunuz. Bu pivot table'da her satırda bir oyuncu olacak şekilde manipülasyon yapınız.

Adım 1: İndekste “player_id”,“position_id” ve “potential_label”, sütunlarda “attribute_id” ve değerlerde scout’ların oyunculara verdiği puan “attribute_value” olacak şekilde pivot table’ı oluşturunuz.

			4322	4323	4324	4325	4326	4327	..
player_id	position_id	potential_label							..
1355710	7	average	50.500	50.500	34.000	50.500	45.000	45.000	..
1356362	9	average	67.000	67.000	67.000	67.000	67.000	67.000	..
1356375	3	average	67.000	67.000	67.000	67.000	67.000	67.000	..
	4	average	67.000	78.000	67.000	67.000	67.000	78.000	..
1356411	9	average	67.000	67.000	78.000	78.000	67.000	67.000	..

Adım 2: “reset_index” fonksiyonunu kullanarak indeksleri değişken olarak atayınız ve “attribute_id” sütunlarının isimlerini stringe çeviriniz.

Adım 6: Label Encoder fonksiyonunu kullanarak “potential_label” kategorilerini (average, highlighted) sayısal olarak ifade ediniz.

Adım 7: Sayısal değişken kolonlarını “num_cols” adıyla bir listeye atayınız.

Adım 8: Kaydettiğiniz bütün “num_cols” değişkenlerindeki veriyi ölçeklendirmek için StandardScaler uygulayınız.

Adım 9: Elimizdeki veri seti üzerinden minimum hata ile futbolcuların potansiyel etiketlerini tahmin eden bir makine öğrenmesi modeli geliştiriniz. (Roc_auc, f1, precision, recall, accuracy metriklerini yazdırınız.)

Adım 10: Değişkenlerin önem düzeyini belirten feature_importance fonksiyonunu kullanarak özelliklerin sıralamasını çizdiriniz.

miuul

miuul.com