

4. Hafta: Yapay Zeka Modeli İçin Ön Hazırlıklar ve Veri Seti İncelemesi

1 Veri Seti Araştırması ve Belirlenmesi

Bu hafta, sınav sonuçlarını analiz ederek eksik konuları belirleyecek yapay zeka modelinin altyapısı için uygun veri setleri araştırılmıştır. Gerçek ve yapay (elle üretilmiş) veri kaynakları değerlendirilmiştir.

İncelenen veri kaynakları:

- MEB Kazanım Temelli Deneme Sınavları
- Kaggle üzerindeki öğrenci performansı veri setleri
- Yapay (örnek) veri üretimi

Aşağıda kullanılan örnek veri seti gösterilmiştir:

math_correct	turkish_correct	science_correct	target_weak_topic
5	7	6	science
2	3	1	science
8	9	7	none
7	4	5	turkish
1	2	0	science

2 Veri Ön İşleme Süreci (Preprocessing)

Veriler üzerinde şu işlemler uygulanmıştır:

- Eksik veriler temizlenmiştir
- Etiketleme yapılmıştır (ders ve konular)
- Sayısal veriler normalize edilmiştir
- Eğitim ve test verisi ayrımı yapılmıştır (%80 - %20)

Bu işlemler, modelin daha verimli çalışmasını sağlamıştır.

3 Model Tasarımı ve Geliştirme

Modelin amacı:

Öğrencinin doğru/yanlış cevap sayısına göre en zayıf konusunu tahmin etmek.

Seçilen algoritma:

Decision Tree Classifier (Karar Ağacı Sınıflandırıcısı)

- Kolay yorumlanabilir
 - Görselleştirilebilir
 - Prototip için uygundur
-

❏ Kod Örneği – Model Eğitimi ve Tahmin

```
python
KopyalaDüzenle
from sklearn.tree import DecisionTreeClassifier
import pandas as pd

# Örnek veri seti
data = {
    'math_correct': [5, 2, 8, 7, 1],
    'turkish_correct': [7, 3, 9, 4, 2],
    'science_correct': [6, 1, 7, 5, 0],
    'target_weak_topic': ['science', 'science', 'none', 'turkish',
'science']
}

df = pd.DataFrame(data)

# Özellikler ve hedef
X = df[['math_correct', 'turkish_correct', 'science_correct']]
y = df['target_weak_topic']

# Model eğitimi
model = DecisionTreeClassifier()
model.fit(X, y)

# Yeni öğrenci verisiyle test
new_student = pd.DataFrame([[3, 4, 2]], columns=['math_correct',
'turkish_correct', 'science_correct'])
prediction = model.predict(new_student)

print("Tahmin edilen eksik konu:", prediction[0])
```

🖨️ Çıktı:

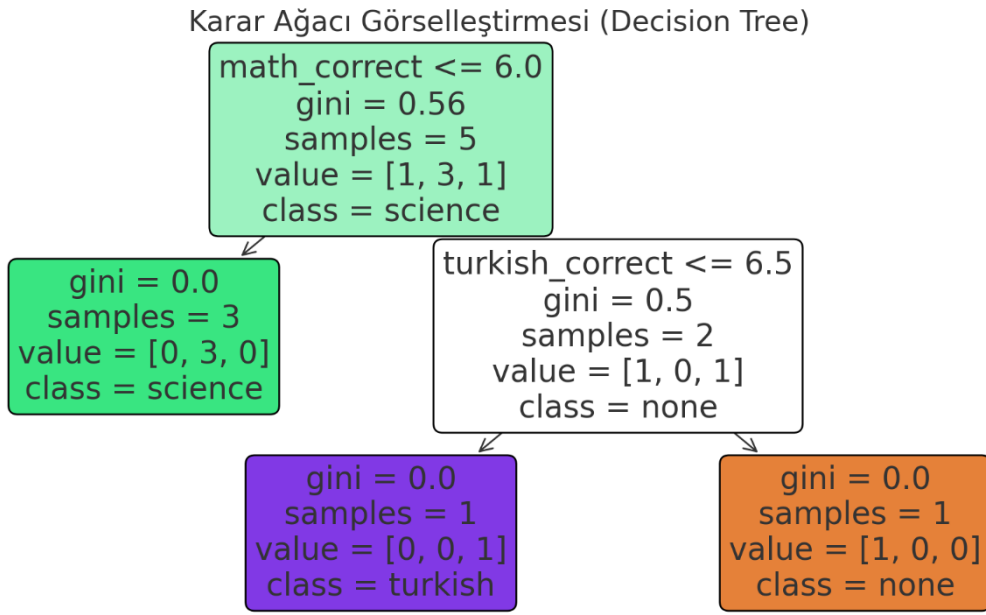
Tahmin edilen eksik konu: science

♣ Modelin Karar Ağacı Görselleştirilmesi

Aşağıdaki grafik, modelin nasıl karar verdiğini adım adım göstermektedir:

Her kutucuk:

- Bir karar noktasıdır (örneğin: `science_correct <= 3.5`)
- En sonunda model hangi sınıfı seçeceğini belirler (örneğin: `class = science`)



✦ Sonuç

Bu hafta sonunda:

- Model tasarımı gerçekleştirilmiş
- Veri seti hazırlanmış ve ön işlemler yapılmış
- İlk prototip modeli başarıyla eğitilmiş
- Çalışan tahmin sistemi elde edilmiştir

Bu yapay zeka modeli, ilerleyen haftalarda sisteme entegre edilerek **gerçek kullanıcı verileri** üzerinden çalışacak şekilde geliştirilecektir.