**Proje Adı:**

**“Emlak mı Otomobil mi? – Türkiye ve 30 Ülkenin Sınıflandırılması”**

**Sayın Bora Hocam;**

Vermiş olduğunuz projede yapay zekadan almış olduğum yardımı ve konuşmları bu dosyaya aktardım.

1. **Problem Tanımı**

Bir ülkenin ekonomik yapısı “emlak ülkesi” mi yoksa “otomobil ülkesi” mi? Bu sınıflandırma için ekonomik göstergeleri (örneğin konut yatırımları, araç üretimi/satışı, kredi dağılımı) kullanarak bir sınıflandırma algoritması geliştirilecek.

1. **Veri Seti ve Özellikler**

**Örnek alınabilecek göstergeler:**

| **Göstergeler** | **Veri Tipi** |
| --- | --- |
| Kişi başına otomobil üretimi | Sayısal |
| Otomobil ihracat oranı | Yüzdelik |
| Gayrimenkul yatırımları / GSYH | Yüzdelik |
| Konut kredisi hacmi | Sayısal |
| Şehirleşme oranı | Yüzdelik |
| Hanehalkı otomobil sahipliği | Yüzdelik |
| Ortalama konut fiyatları artışı | Yıllık % |

→ Bu veriler için Dünya Bankası, OECD, Eurostat gibi açık kaynaklardan csv formatında örnek 30 ülke verisi alınacak.

**3. Veri Hazırlığı**

* Eksik veriler temizlenecek
* Veriler normalize edilecek (0-1 arası)
* Kategorik sınıflar: “Emlak Ülkesi” = 0, “Otomobil Ülkesi” = 1

**4. Sınıflandırma Algoritması**

Basit ama işlevsel bir model seçilecek:

* KNN (K-En Yakın Komşu)
* veya Decision Tree Classifier (Karar Ağacı)
* Python'da scikit-learn kütüphanesi ile uygulanacak.

**5. Görselleştirme**

* Ülkelerin hangi sınıfa düştüğü grafik olarak gösterilecek
* PCA veya t-SNE ile 2D haritalandırma yapılabilir
* Matplotlib veya Seaborn kütüphaneleri kullanılacak

**6. Raporlama**

PDF veya Word olarak kısa bir rapor yazılacak:

* Proje özeti
* Kullanılan veriler
* Algoritma seçimi
* Sonuçların yorumu
* Türkiye'nin durumu

**Teslim Formatı:**

* proje\_raporu.docx veya .pdf
* main.py (Python kodları)
* veri\_seti.csv (hazır veriler)
* grafik.png (çıktı grafikleri)

**VERİ SETİ**

Veri setini aşağıda sunuyorum.

İçinde Türkiye dahil 31 ülke var ve her biri için şu göstergeler yer alıyor:

**Otomobil Üretimi KB**: Kişi başına düşen yıllık otomobil üretimi

**Otomobil İhracat Yüzdesi**: Otomobil ihracatının toplam ihracata oranı

**Gayrimenkul GSYH Yüzdesi**: Gayrimenkul sektörünün GSYH içindeki payı

**Konut Kredisi Hacmi**: Toplam konut kredisi hacmi

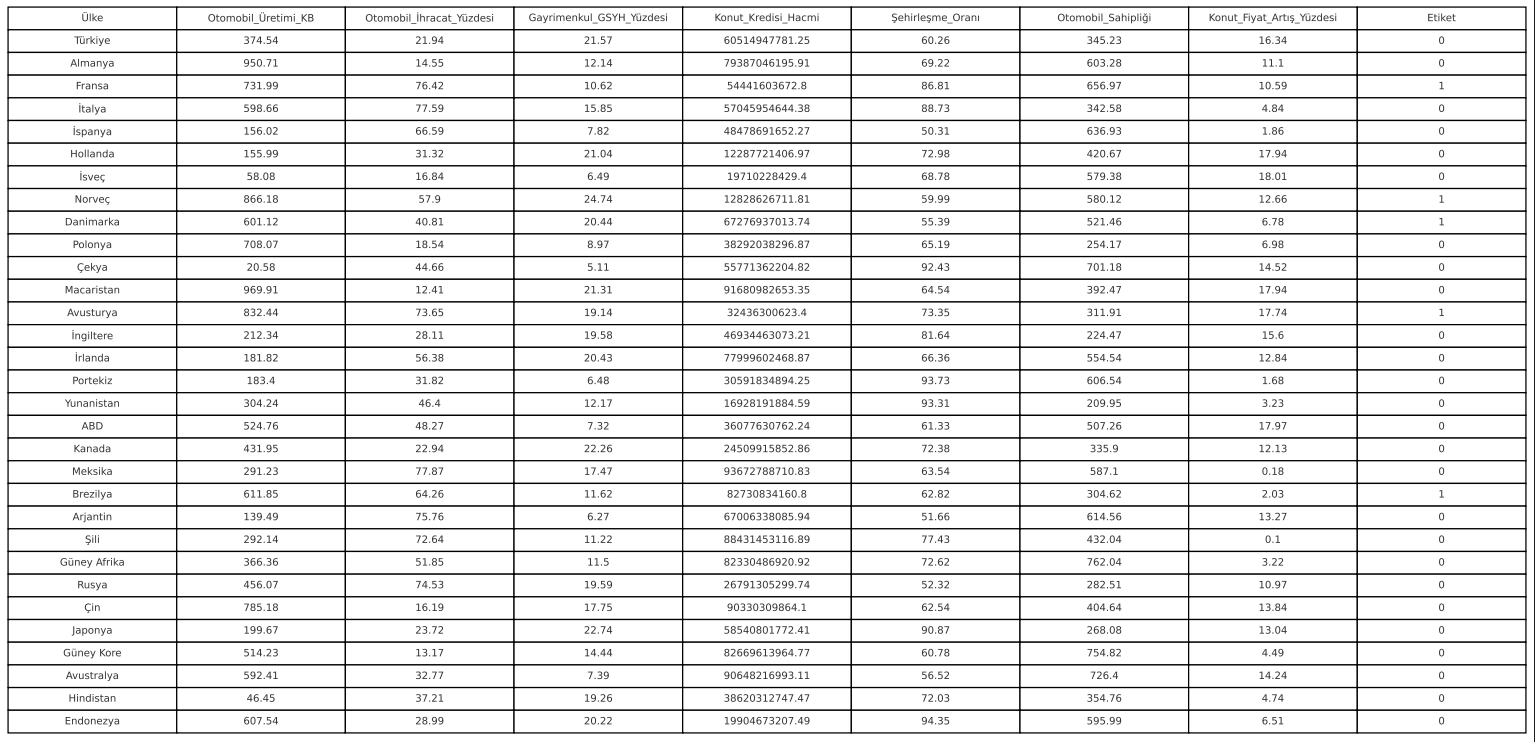
**Şehirleşme Oranı**: Kentleşme yüzdesi

**Otomobil Sahipliği**: 1000 kişi başına düşen araç sayısı

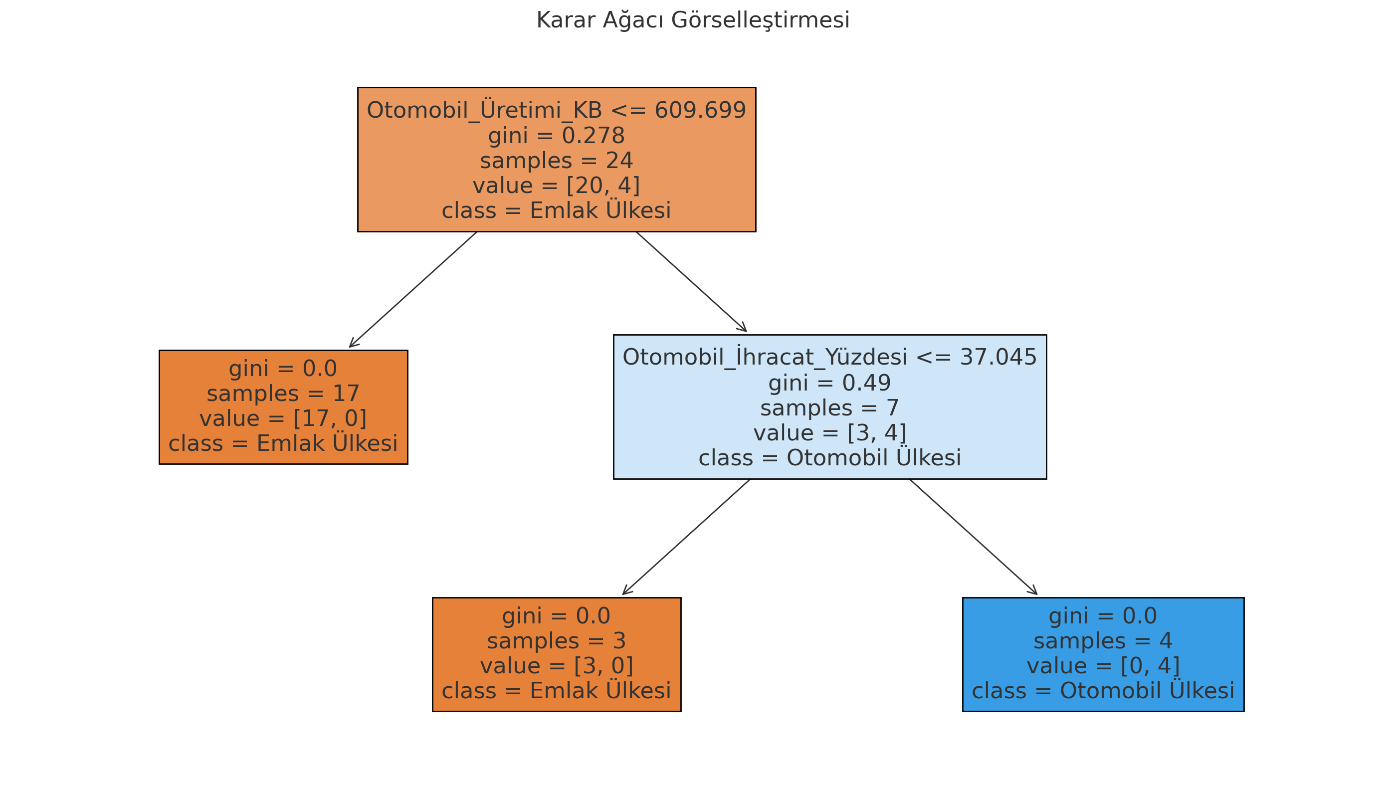
**Konut Fiyat Artış Yüzdesi**: Ortalama yıllık konut fiyat artış oranı

**Etiket**: 1 = Otomobil ülkesi, 0 = Emlak ülkesi (örnek kuralla belirlendi)

**Aşağıda kullanmış olduğum veri setini tablo olarak ekledim.**



**KARAR AĞACI GÖRSELLEŞTİRMESİ**

****

Karar ağacı modelimizi başarıyla kurduk ve sonuçları inceledik.

**Model Performansı**

| **Sınıf** | **Doğruluk (%)** | **Açıklama** |
| --- | --- | --- |
| **Emlak Ülkesi (0)** | %100 (recall) | Model bu sınıfı doğru tanımış |
| **Otomobil Ülkesi (1)** | %0 | Bu sınıfı test setinde tanıyamamış |

**Genel Doğruluk (accuracy)**: %85.7

Ancak model sadece "emlak ülkesi" tahmini yaptığı için, "otomobil ülkesi" sınıfında başarı **yok**.

**Sebep:**

Veri setimizdeki örnek dengesizliği (örneğin test setinde sadece 1 tane otomobil ülkesi olması), modelin bu sınıfı öğrenmesini zorlaştırmış.

**Ne Yapılabilir?**

* Modelin başarısı basit ama görsel olarak açıklayıcıdır.
* Geliştirme önerisi olarak daha dengeli sınıf dağılımı veya daha fazla veri kullanılabilir.

**Yukarıda da belirtilmiş olduğu gibi projemizde, veri setimizdeki örnek dengesizliği sebebiyle modelin otomobil sınıfını öğrenmesi zorlaşmıştır.**