**Yaşam Tarzı Verileriyle Work-Life Balance Tahmini – Ensemble ve Model Karşılaştırma Yaklaşımı**

**Proje Amacı**

Bu projede, bireylerin yaşam tarzı alışkanlıklarına göre **iş-yaşam dengesi skorlarını** tahmin etmeyi hedefledim.  
Kullanılan veri seti, 15.000'den fazla kişinin uyku süresi, stres seviyesi, sosyal çevresi gibi özelliklerini içeriyor.  
Amaç: Bu verilerden yola çıkarak kişinin **Work-Life Balance Score**'unu tahmin etmek.

**Veri Seti Özeti**

* Kaynak: Kaggle – *Lifestyle and Wellbeing Dataset*
* Örnek sayısı: 15.977
* Hedef değişken: WORK\_LIFE\_BALANCE\_SCORE
* Kullanılan özellikler:
  + DAILY\_STRESS, SLEEP\_HOURS, DAILY\_STEPS
  + SOCIAL\_NETWORK, ACHIEVEMENT, TIME\_FOR\_PASSION
  + ve diğer yaşam tarzı göstergeleri

**Yöntem**

Projede **ensemble model yaklaşımı** ve **model karşılaştırması** kullandım.

* İlk olarak, farklı algoritmaların performansını görmek için **cross-validation** ile model karşılaştırması yaptım.
* Kullanılan modeller:
  1. **Linear Regression**
  2. **Random Forest Regressor**
  3. **XGBoost Regressor**

Bu karşılaştırma sayesinde en düşük hata oranına sahip modeli belirledim.

* Ardından en başarılı iki modelin tahminlerini birleştirerek **ensemble (ortalama tahmin)** oluşturdum.

**Veri Ön İşleme**

* Eksik değerler sayısal sütunlarda ortalama ile, kategorik sütunlarda mod ile dolduruldu.
* Kategorik değişkenler one-hot encoding ile sayısala çevrildi.
* Eğitim ve test verisi %80–%20 oranında ayrıldı.

**Sonuçlar**

| **Model** | **Mean Squared Error (MSE)** |
| --- | --- |
| Linear Regression | *A değeri* |
| Random Forest | *B değeri* |
| XGBoost | *C değeri* |
| Ensemble (Ortalama) | *D değeri* |

Cross-validation ile yapılan karşılaştırma, XGBoost’un en düşük hata ile en başarılı model olduğunu gösterdi.  
Ensemble model ise iki güçlü algoritmanın avantajlarını birleştirerek daha dengeli tahminler sundu.

**Görselleştirme**

Tahmin edilen skorlar ile gerçek skorlar karşılaştırıldı.  
Grafikte, modelin bireylerin iş-yaşam dengesi skorlarını ne kadar başarılı tahmin ettiğini görebiliyoruz.

**Sonuç ve Öneriler**

* Model, yaşam tarzı verileriyle iş-yaşam dengesi skorunu başarılı şekilde tahmin etti.
* Cross-validation sayesinde en uygun model seçildi.
* Daha fazla veri ve hiperparametre optimizasyonuyla doğruluk artırılabilir.
* Aynı yaklaşım farklı sağlık veya psikolojik skorlar için de uygulanabilir.