# Ev Fiyat Tahmin Modelleri Karşılaştırma Raporu

## Performans Metrikleri Açıklamaları

- **MAE (Mean Absolute Error - Ortalama Mutlak Hata):** Modelin tahmin ettiği değerler ile gerçek değerler arasındaki farkların mutlak değerlerinin ortalamasıdır. Daha düşük MAE, daha iyi bir model performansını gösterir.  
  
- **MSE (Mean Squared Error - Ortalama Kare Hata):** Tahmin hatalarının karelerinin ortalamasıdır. Büyük hataları daha fazla cezalandırdığı için büyük hata yapan modelleri daha belirgin hale getirir. Küçük değerler daha iyi model performansını gösterir.  
  
- **RMSE (Root Mean Squared Error - Karekök Ortalama Kare Hata):** MSE’nin karekökü alınarak elde edilir. Ölçü birimi verinin birimiyle aynıdır ve hataların büyüklüğünü daha anlaşılır hale getirir. Daha düşük RMSE, daha iyi bir model anlamına gelir.  
  
- **R² (R-Karesi - Determinasyon Katsayısı):** Modelin bağımsız değişkenlerin bağımlı değişkeni açıklama oranını gösterir. 1’e ne kadar yakınsa model o kadar iyi tahmin yapıyor demektir. Negatif veya sıfıra yakın değerler modelin kötü performans gösterdiğini belirtir.

## Genel Analiz ve Sonuçlar

1. En Düşük Hata (MAE, MSE, RMSE): LightGBM ve CatBoost modelleri en düşük hata oranlarını vererek en iyi sonuçları sundu.  
2. En Yüksek R² Skoru: CatBoost (0.1967) ve LightGBM (0.1957) en yüksek R² skorlarına sahip.  
3. Karar Ağaçları ve Random Forest: Tek başına karar ağaçları (Decision Tree) daha yüksek hata oranına sahipken, Random Forest modeli hataları biraz düşürdü.  
4. Gradient Boosting Modelleri: Gradient Boosting, XGBoost ve LightGBM benzer sonuçlar vermesine rağmen LightGBM en iyi performansı gösterdi.  
5. MLP (Yapay Sinir Ağı): Nispeten iyi bir performans gösterse de Boosting algoritmaları kadar başarılı olamadı.  
6. AdaBoost Kötü Performans Gösterdi: AdaBoost en kötü sonuçları verdi. Bu modelin yüksek bias ve düşük kapasiteye sahip olması bu durumu açıklayabilir.

## Modellerin Performans Karşılaştırması

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | MAE | MSE | RMSE | R² Skoru |
| Linear Regression | 0.6275 | 0.7877 | 0.8875 | 0.1611 |
| Decision Tree Regressor | 0.6221 | 0.8135 | 0.9019 | 0.1336 |
| Random Forest Regressor | 0.6143 | 0.7834 | 0.8851 | 0.1657 |
| Gradient Boosting Regressor | 0.6104 | 0.7585 | 0.8709 | 0.1922 |
| XGBoost Regressor | 0.609 | 0.7597 | 0.8716 | 0.1909 |
| LightGBM Regressor | 0.6065 | 0.7552 | 0.869 | 0.1957 |
| MLP Regressor | 0.6018 | 0.771 | 0.8781 | 0.1789 |
| AdaBoost Regressor | 0.8282 | 1.0099 | 1.0049 | -0.0755 |
| TabNet Regressor | 0.6188 | 0.7661 | 0.8753 | 0.1841 |
| CatBoost Regressor | 0.6059 | 0.7543 | 0.8685 | 0.1967 |

## ****🔹 En düşük MAE: CatBoost (0.6059)****

## ****🔹 En düşük MSE: CatBoost (0.7543)****

## ****🔹 En düşük RMSE: CatBoost (0.8685)****

## ****🔹 En yüksek R²: CatBoost (0.1967)****

## ****3. Hangi Model Hangi Konuda Daha İyi?****

### ****Genel Olarak En İyi Performansı Gösteren Model: CatBoost****

* Tüm metriklerde en düşük hata ve en yüksek R² değerine sahip.
* LightGBM ve XGBoost da benzer performans gösteriyor, ancak CatBoost biraz daha iyi.

### ****Düşük Hesaplama Maliyeti ile İyi Performans: LightGBM****

* Büyük veri setlerinde hızlı çalışıyor.
* CatBoost’a yakın sonuçlar veriyor, ancak eğitim süresi daha kısa.

### ****Derin Öğrenme Alternatifleri: MLP ve TabNet****

* MLP ve TabNet, non-lineer ilişkileri yakalamada başarılı.
* Ancak, performansları CatBoost ve LightGBM’den biraz daha düşük.
* Eğitim süreleri uzun ve hiperparametre ayarına daha duyarlılar.

### ****Daha Basit ve Açıklanabilir Model: Linear Regression****

* Açıklanabilirlik açısından güçlü, ancak doğruluk düşük.
* Verilerde güçlü lineer bağıntılar yoksa zayıf performans gösteriyor.

### ****Zayıf Performans Gösteren Model: AdaBoost****

* MAE, MSE ve RMSE değerleri en yüksek.
* R² değeri negatif, yani rastgele tahmin yapmaktan bile kötü.
* Küçük veri setlerinde iyi çalışabilir ama burada başarısız oldu.