



# FET445 Veri Madenciliği

**Bitcoin (BTC-USD) Fiyat Tahmini: Gelişmiş Ensemble ve Deep Learning Modeller**

**Fantastik Dörtlü**

Youtube Link: <https://www.youtube.com/watch?v=0pJi8GshvBE>

Github Repo: <https://github.com/ilkayGkbdk/Advanced-BTC-Prediction>

# Proje Ekibi

## Fantastik Dörtlü

İsim Soyisim	Öğrenci Numarası	Email
Barchinoy KODIROVA	22040101112	<a href="mailto:barchinoykodirova@stu.topkap.edu.tr">barchinoykodirova@stu.topkap.edu.tr</a>
Berkant ŞİMŞEK	22040101038	<a href="mailto:berkantsimsek@stu.topkap.edu.tr">berkantsimsek@stu.topkap.edu.tr</a>
İlkay GÖKBUDAK	22040101047	<a href="mailto:ilkaygokbudak@stu.topkap.edu.tr">ilkaygokbudak@stu.topkap.edu.tr</a>
Ömer AVCI	22040101024	<a href="mailto:omeravci@stu.topkap.edu.tr">omeravci@stu.topkap.edu.tr</a>

# Problemin Tanımı ve Amaç

- **Problem:** Cripto para piyasalarındaki yüksek volatilite ve belirsizlik nedeniyle Bitcoin fiyatlarını tahmin etme zorluğu.
- **Amaç:** Geçmiş 3 yıllık (1095 gün) verileri kullanarak, makine öğrenmesi ve derin öğrenme yöntemleriyle gelecek 30 günün fiyat haraketlerini tahmin etmek.
- **Kapsam:** Regresyon problemi (Sayısal fiyat tahmini).

# Veri Seti ve Ön İşleme

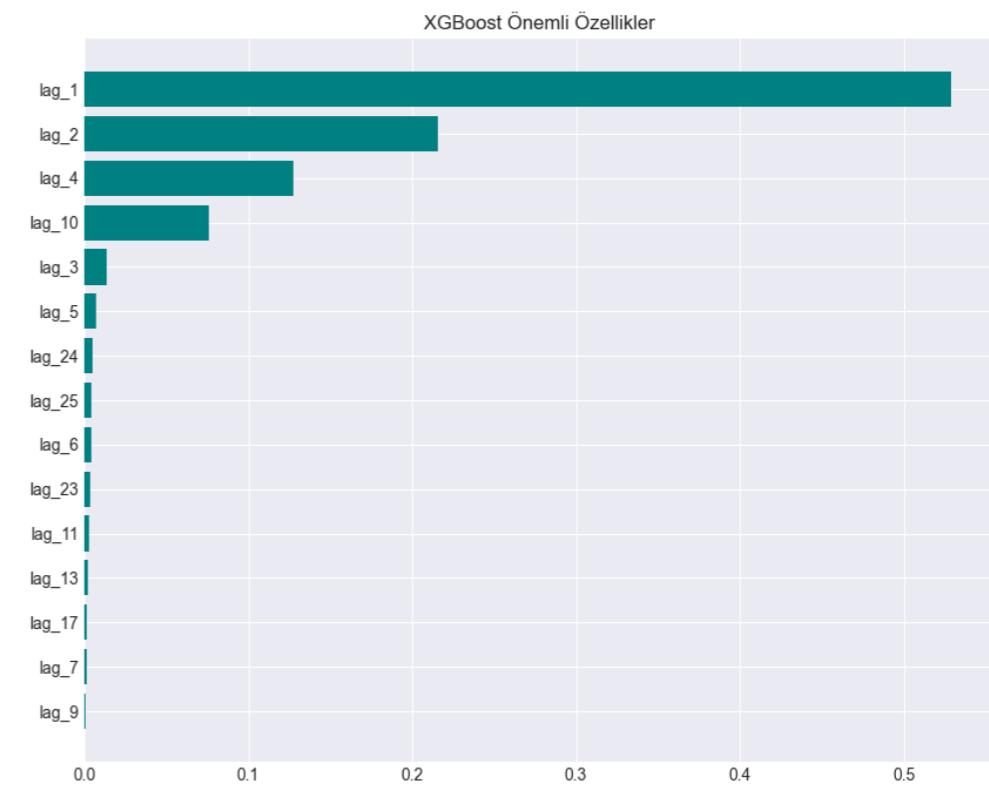
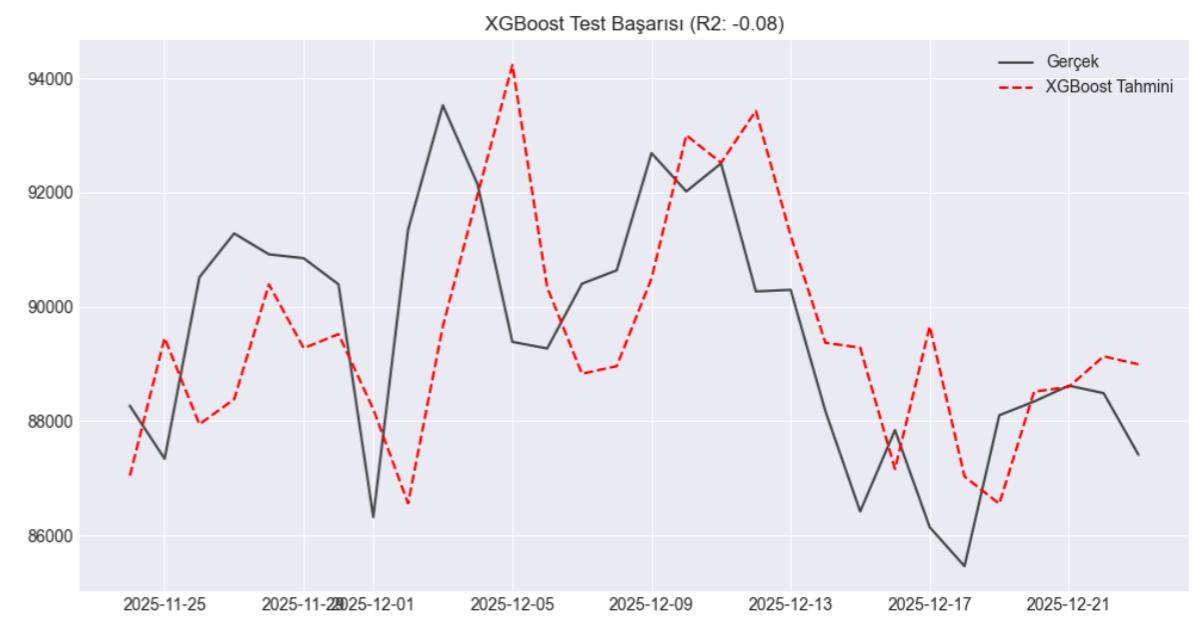
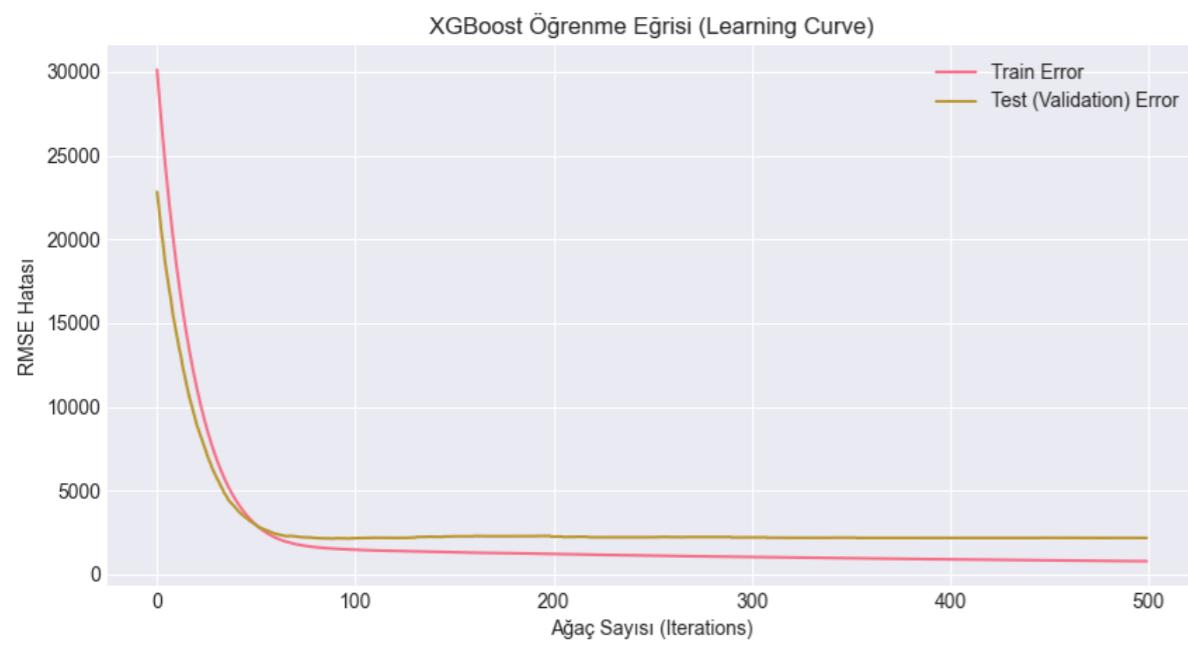
- **Veri Kaynağı:** Yahoo Finance API (yFinance).
- **Veri Aralığı:** 3 Yıl (Günlük Veri - Daily Close).
- **Kullanılan Özellikler (Features):**
  - **Lag Features:** Geçmiş 30 günün fiyatları ( $t-1, t-2, \dots, t-30$ ) model girdisi olarak kullanıldı.
  - **Scaling:** Derin öğrenme modelleri (GRU/LSTM) için *MinMaxScaler* ile 0-1 noramlizasyonu yapıldı.
- **Eğitim-Test Ayrımı:**
  - Zaman serisi yapısını bozmamak için *Shuffle=False* kullanıldı.
  - Verinin son bölümü (Test Seti) modelin görmediği gelecek senaryosu olarak ayrıldı.

# Metodoloji ve Performans Metrikleri

- **Kullanılan 4 Gelişmiş Model:**
  - **XGBoost** (Gradient Boosting - Best Model)
  - **LightGBM** (Hızlı ve Hafif Boosting)
  - **PyTorch GRU** (Gated Recurrent Unit)
  - **PyTorch LSTM** (Long Short-Term Memory)
- **Başarı Kriterleri (Metrikler):**
  - **MAPE (Mean Absolute Percentage Error):** Yüzdesel hata payı (En kritik metrik).
  - **RMSE (Root Mean Squared Error):** Standart sapma cinsinden hata büyüklüğü.
  - **R<sup>2</sup> (R-Squared):** Modelin varyansı açıklama gücü.

**En İyi Model  
(Best Model)**

**XGBoost**



# En İyi Model (Best Model)

## XGBoost

- **Neden En İyisi?**
  - Test verisi üzerinde **%1.96** gibi çok düşük bir **MAPE** oranı yakaladı.
  - RMSE değeri **2182 USD** ile en düşük sapmaya sahip model oldu.
- **Optimizasyon (Hyperparameter Tuning):**
  - **Yöntem:** *RandomizedSearchCV + TimeSeriesSplit*.
  - **Optimize Edilenler:** Ağaç sayısı (`n_estimators`), Derinlik (`max_depth`), Öğrenme hızı (`learning_rate`).
- **Feature Importance:** Model, fiyatı tahmin ederken en çok **son 1-3 günün (Lag\_1, Lag\_2)** fiyatlarına ağırlık verdi.

# Modellerin Karşılaştırma Tablosu

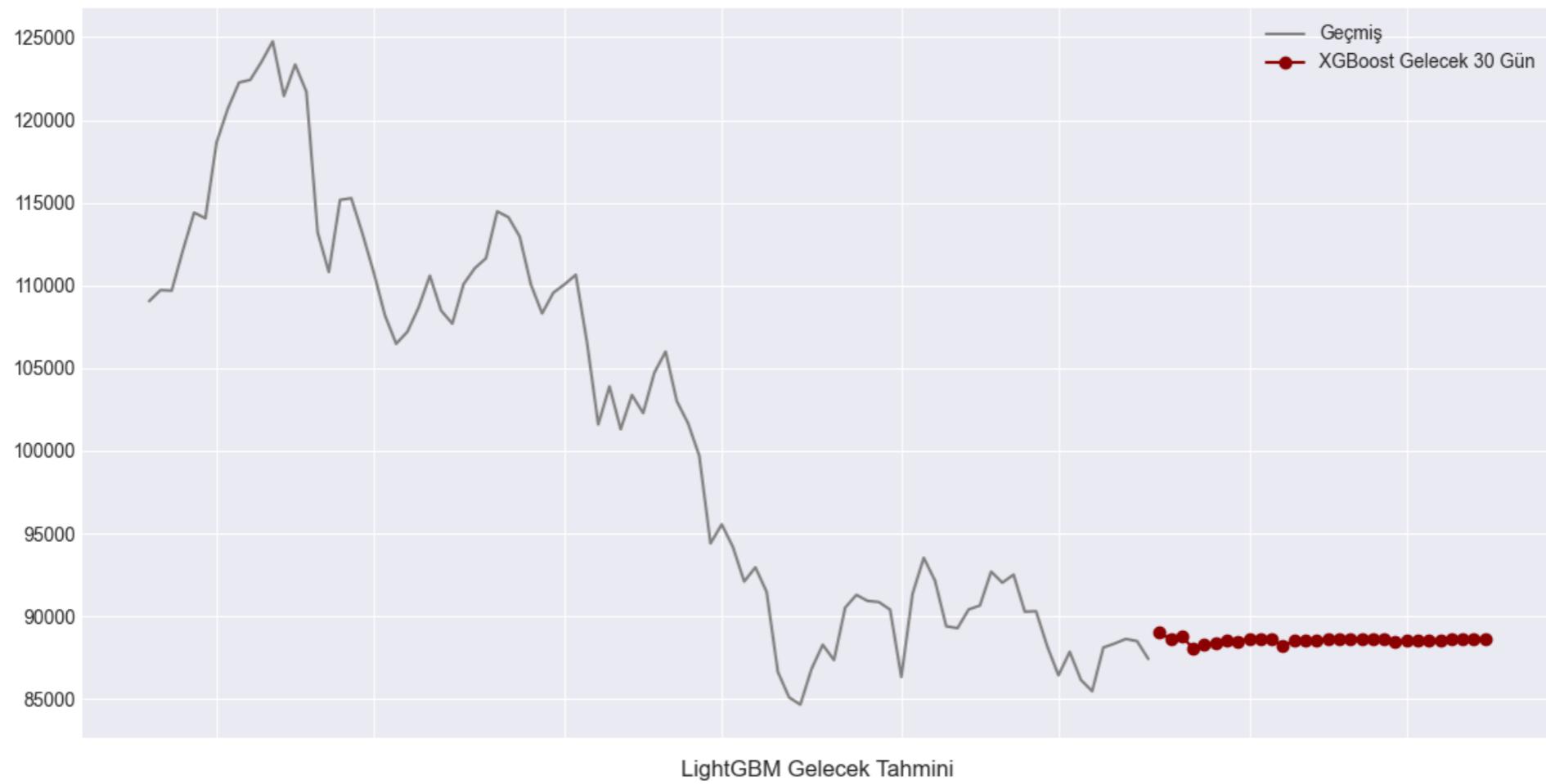
Model	Model Türü	MAPE (Hata%)	RMSE (USD)	R^2 Score	Durum
XGBoost	Ensemble	%1.96	2,182	-0.08*	En İyi Sonuç
LightGBM	Ensemble	%2.10	2,306	-0.21	Başarılı
GRU	Deep Learning	%2.12	2,806	0.94	Yüksek Korelasyon
LSTM	Deep Learning	%21.47	25,558	-3.57	Geliştirilmeli

\*Not: XGBoost ve LightGBM kısa vadeli test setinde trend değişimine denk geldiği için R2 negatif görünse de, MAPE değerinin düşüklüğü (yaklaşık %2 hata) tahminlerin fiyata çok yakın olduğunu kanıtlamaktadır.

# **Gelecek 30 Günlük Tahminler (Future Forecast)**

**XGBoost vs LightGBM**

XGBoost Gelecek Tahmini



LightGBM Gelecek Tahmini



# Gelecek 30 Günlük Tahminler

## XGBoost vs LightGBM

- **Simülasyon Yöntemi:** Recursive Forecasting (Özyinelemeli Tahmin).
- **Beklenti:** XGBoost modeline göre Bitcoin, önümüzdeki 30 gün boyunca mevcut seviyelerini koruyarak **yatay/hafif yukarı** yönlü bir trend izleyebilir.
- **Tahmin Verileri (Örnek):**
  - XGBoost Tahmini (30. Gün): ~88,600 USD seviyeleri.
  - LightGBM Tamil (30. Gün): ~92,100 USD Seviyeleri.

# Sonuç ve Değerlendirme

- **Çıkarımlar:**
  - **Ensemble Modeller (XGBoost/LightGBM)**, finansal zaman serilerinde daha düşük hata oranı ile çalışarak Derin Öğrenme modellerini (LSTM) geride bıraktı.
  - **GRU Modeli**, LSTM'e göre çok daha başarılı sonuç verdi ve %2.12 hata payı ile XGBoost'a güçlü bir rakip oldu.
  - **LSTM**, veri seti boyutu veya mimari karmaşıklığı nedeniyle bu problemlerde “underfit” (yetersiz öğrenme) kaldı.
- **Gelecek Çalışmalar:** Modele “Haber Analizi” (Sentiment Analysis) eklenerek ani düşüş/yükselişler daha iyi yakalanabilir.

**Bizi Dinledığınız  
İçin Teşekkür  
Ederiz**

**İyi Çalışmalar.**