

# Hisse Senedi Fiyatlarının Günlük Mum Yönü Tahmini:

YAP 470 Dönem Projesi Sunumu

# Proje Tanımı ve Amacı

- The primary objective of this project is to predict the direction (bullish/bearish) of the next daily candlestick in the stock market. The problem is framed as a binary classification task.
- The project aims to quantify the performance difference between these two approaches, evaluating the effectiveness of a dedicated time-series architecture (LSTM) against sophisticated feature engineering with traditional models.

## Traditional-ML Approach

- Xgboost
- RFtree
- MLP

## RNN

- LSTM

# Veri Seti ve Özellik Mühendisliği

- **Veri Kaynağı:** Kaggle – Stock Market Dataset (S&P 500 hisselerinin OHLCV verileri)
- **Ham Özellikler:** Open, High, Low, Close, Volume
- **Türetilmiş Özellikler (Feature Engineering):**
  - a. **Temel Göstergeler:** EMA (10, 20, 50), RSI, ATR
  - b. **Gelişmiş Göstergeler (LSTM için eklendi):** MACD, Bollinger Bantları, ROC.

# Akış Şeması:

Veri Yükleme ve Temizleme -> Özellik Mühendisliği -> Veriyi Kronolojik Ayırma (%80 Eğitim, %20 Test) -> Model Özellik Seçimi / Ölçekleme -> Model Eğitimi -> Test Verisi ile Değerlendirme.

## Testing Methods and Metrics:

Accuracy: Overall correctness

Precision & Recall: Especially important if one class is more prevalent

F1-score: Balances precision and recall

Confusion Matrix: Provides TP, FP, FN, TN breakdown

# XGBOOST

```
k_values = [5, 10, 15]
window_sizes = [3, 7]
split_ratio = 0.8
```

AAPL\_k5\_w3

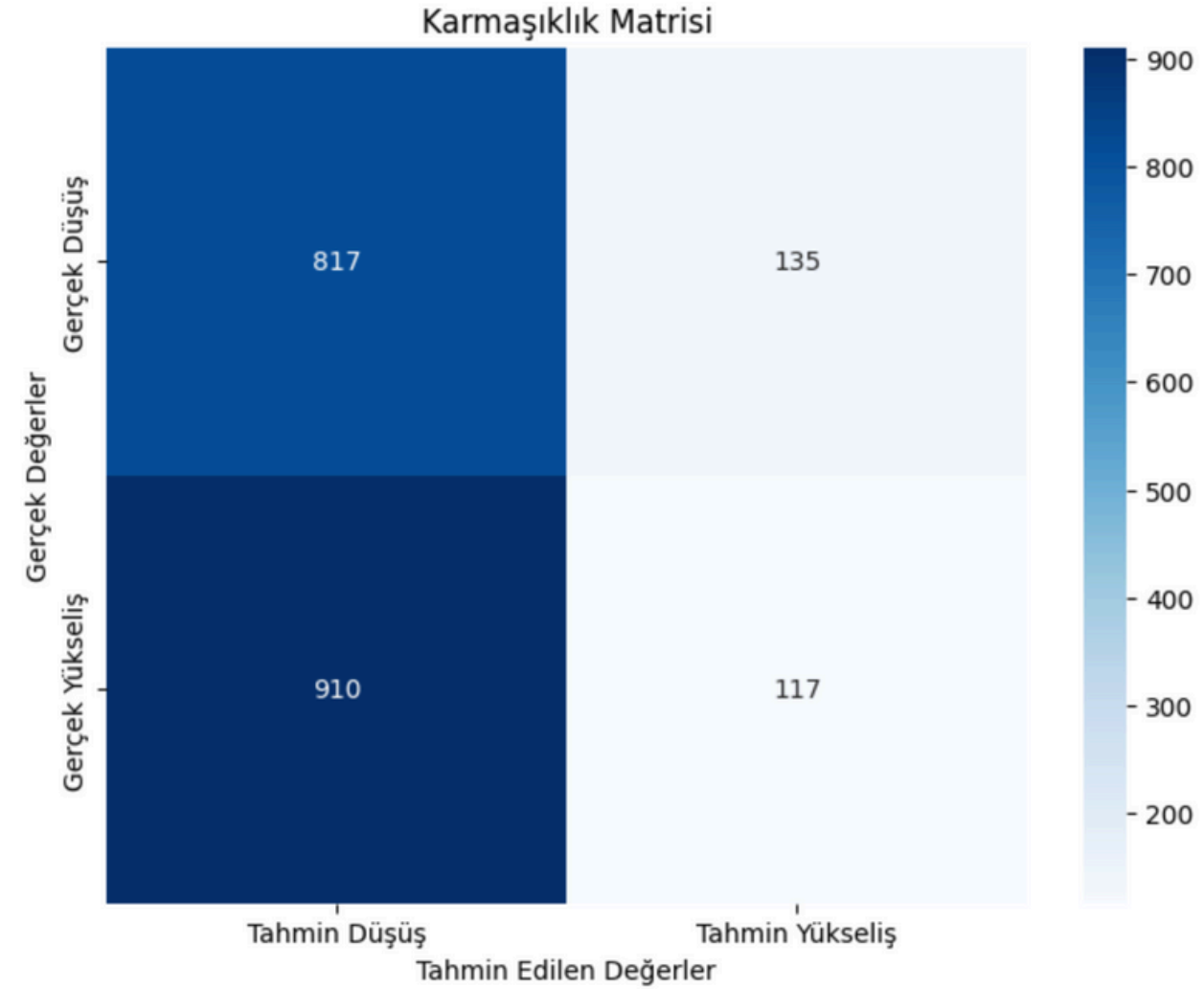
--- MODEL DEĞERLENDİRME SONUÇLARI ---

Modelin Test Verisi Üzerindeki Doğruluğu: 0.4720

Sınıflandırma Raporu:

	precision	recall	f1-score	support
Düşüş (0)	0.47	0.86	0.61	952
Yükseliş (1)	0.46	0.11	0.18	1027
accuracy			0.47	1979
macro avg	0.47	0.49	0.40	1979
weighted avg	0.47	0.47	0.39	1979

Karmaşıklık Matrisi:



# RForest

```
k_values = [5, 10, 15]
window_sizes = [3, 7]
max_depth_values = [10, 20, None]
split_ratio = 0.8
```

AAPL\_k5\_w3

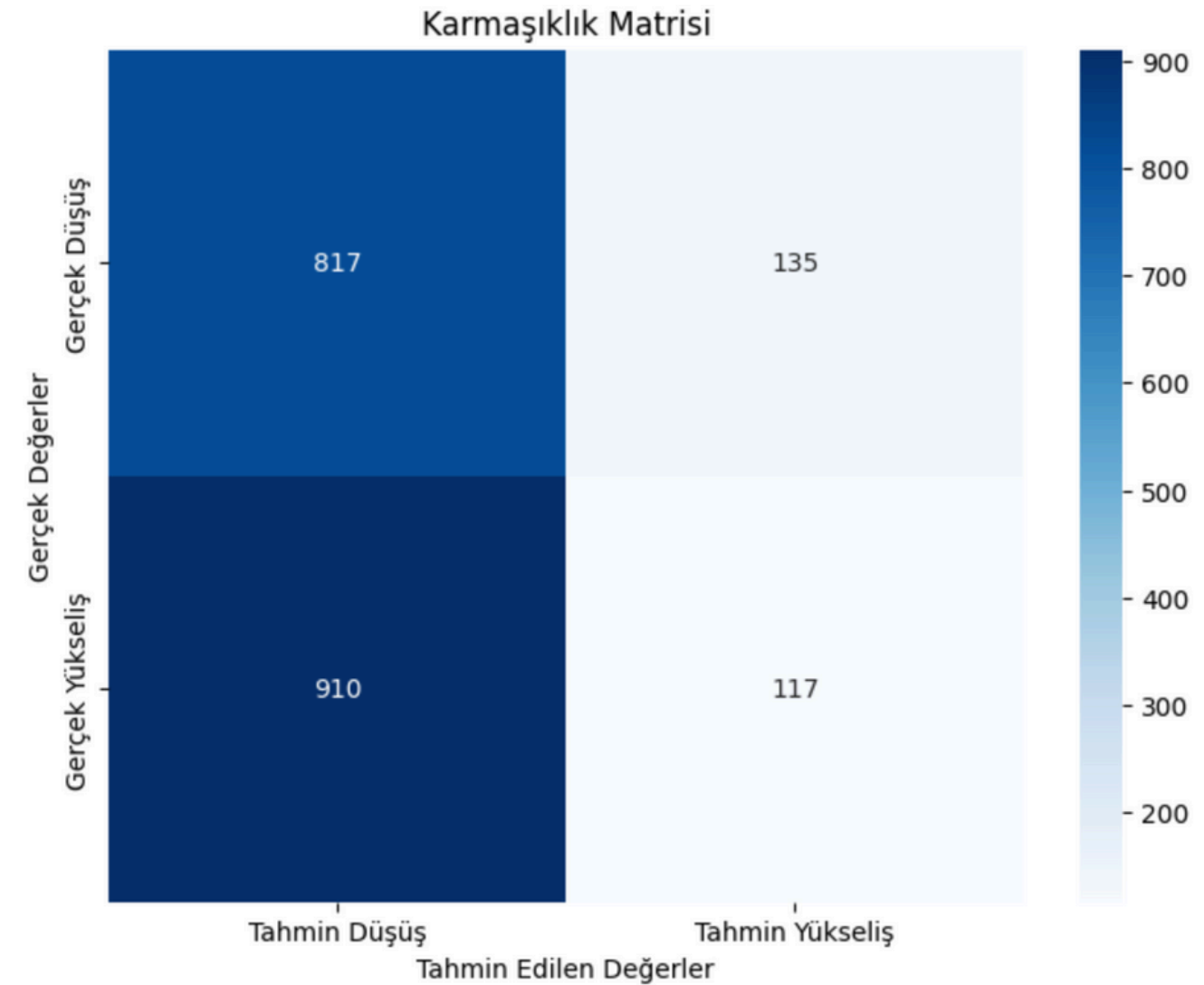
--- MODEL DEĞERLENDİRME SONUÇLARI ---

Modelin Test Verisi Üzerindeki Doğruluğu: 0.4720

Sınıflandırma Raporu:

	precision	recall	f1-score	support
Düşüş (0)	0.47	0.86	0.61	952
Yükseliş (1)	0.46	0.11	0.18	1027
accuracy			0.47	1979
macro avg	0.47	0.49	0.40	1979
weighted avg	0.47	0.47	0.39	1979

Karmaşıklık Matrisi:



# MLP

```
k_values = [5, 10, 15]
window_sizes = [3, 7]
hidden_layer_configs = [(50,), (100,), (50, 25), (100, 50)]
split_ratio = 0.8
```

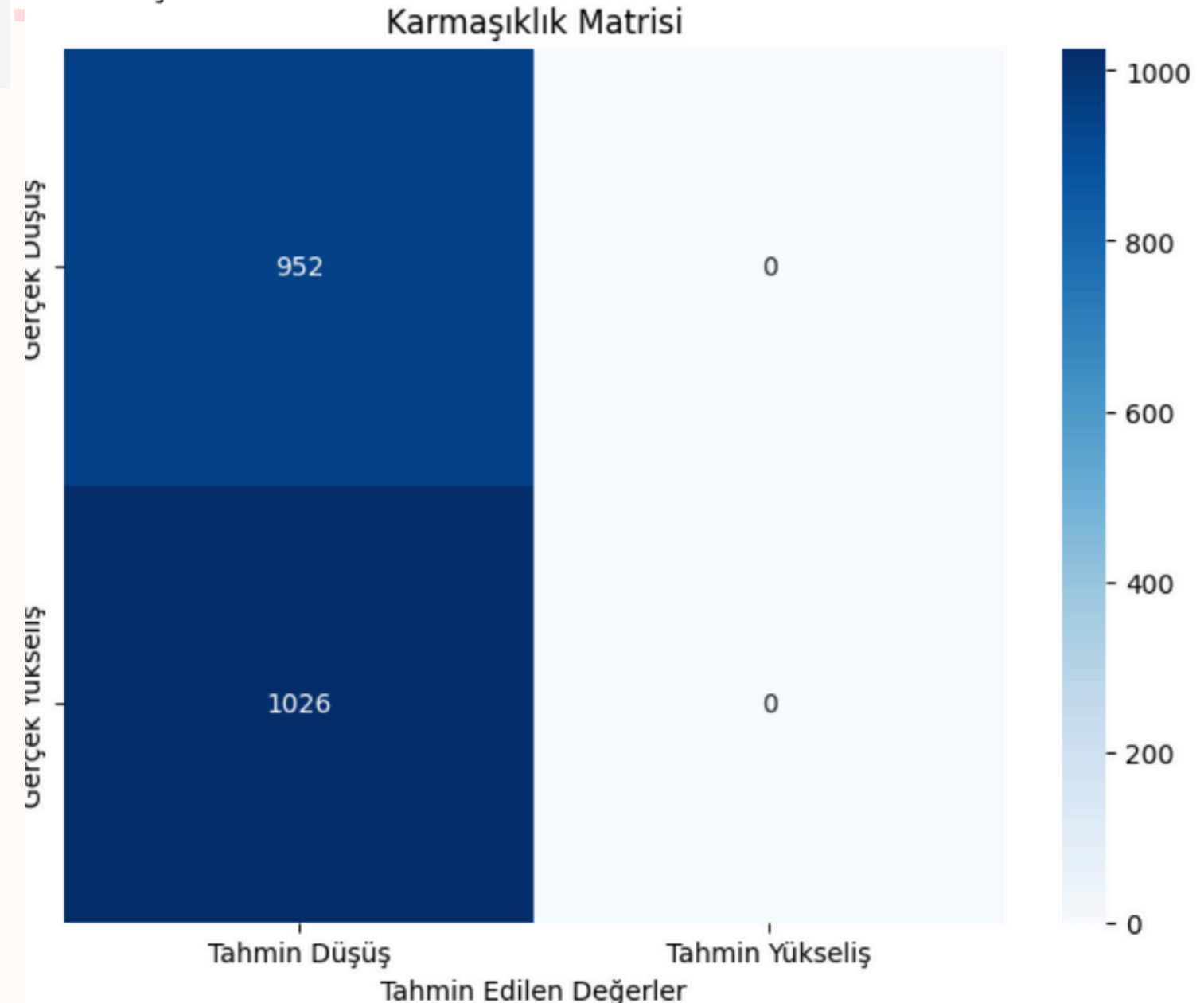
AAPL\_k5\_w7\_lyr100

Modelin Test Verisi Üzerindeki Doğruluğu: 0.4813

Sınıflandırma Raporu:

	precision	recall	f1-score	support
Düşüş (0)	0.48	1.00	0.65	952
Yükseliş (1)	0.00	0.00	0.00	1026
accuracy			0.48	1978
macro avg	0.24	0.50	0.32	1978
weighted avg	0.23	0.48	0.31	1978

Karmaşıklık Matrisi:



# LSTM

```
sequence_lengths_to_try = [3,10,20, 40,100]
lstm_units_to_try = [32, 50]
k_features_to_try = [5,8, 12]
```

AAPL\_s100\_lunit50\_k5

--- MODEL DEĞERLENDİRME SONUÇLARI ---

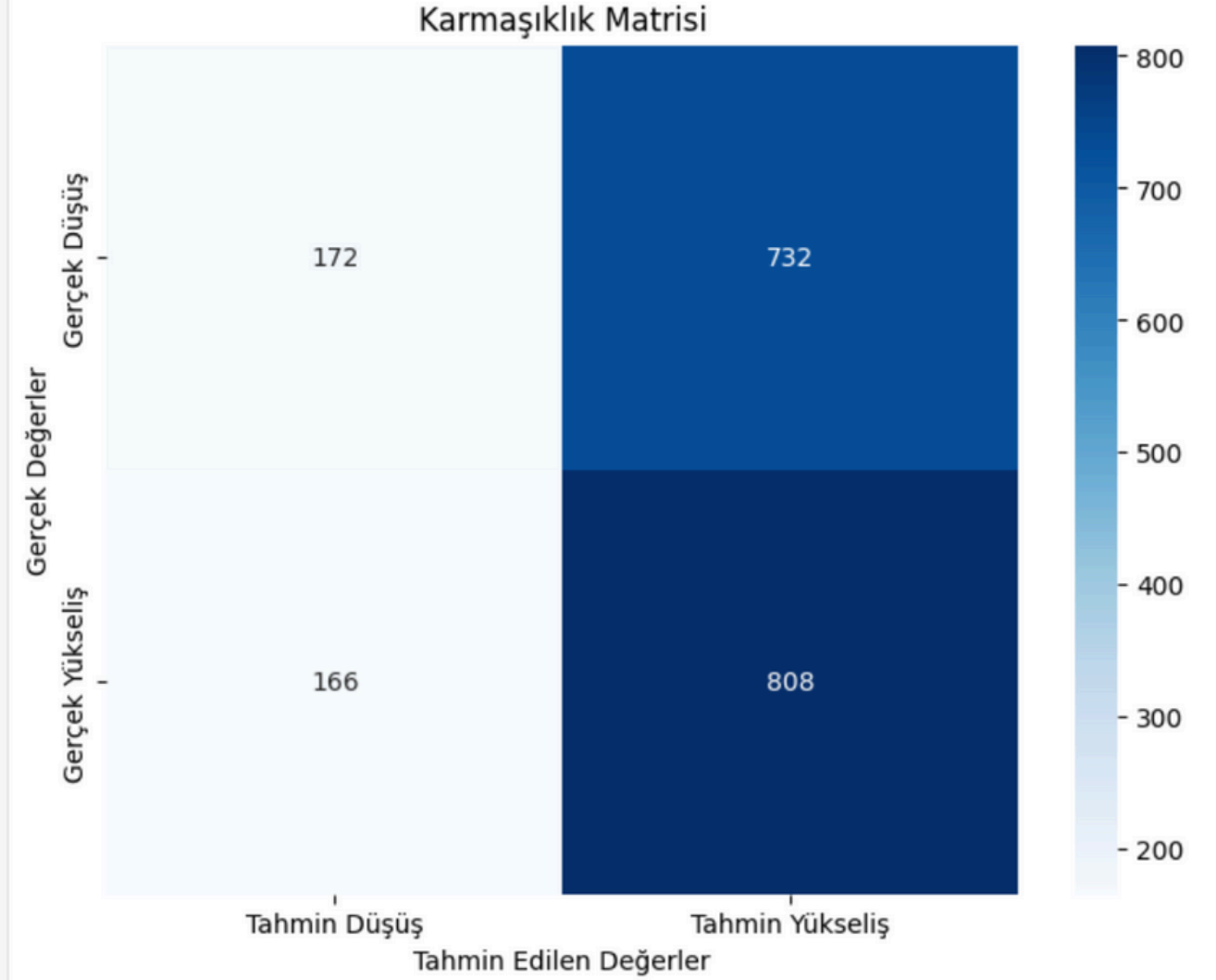
Modelin Test Verisi Üzerindeki Doğruluğu: 0.5218  
Modelin Test Verisi Üzerindeki Kaybı : 0.6929

59/59 ————— 1s 8ms/step

Sınıflandırma Raporu:

	precision	recall	f1-score	support
Düşüş (0)	0.51	0.19	0.28	904
Yükseliş (1)	0.52	0.83	0.64	974
accuracy			0.52	1878
macro avg	0.52	0.51	0.46	1878
weighted avg	0.52	0.52	0.47	1878

Karmaşıklık Matrisi:





# Modellerin Karşılaştırılması ve Genel Değerlendirme

Tartışmak istediklerinizi kısaca detaylandırın.

## Sonuç ve Değerlendirme

Bu çalışma kapsamında, Random Forest, MLP, XGBoost) ve derin öğrenme tabanlı LSTM modeli sistematik olarak karşılaştırılmıştır.

Geleneksel modeller arasında en tutarlı ve marjinal olarak en iyi performansı XGBoost sergilemiştir.

Nihayetinde, hem XGBoost hem de en iyi LSTM denemeleri, rastgele tahmin seviyesinin (%50) yalnızca bir miktar üzerine çıkabilmiş, ancak aralarında pratik anlamda kayda değer bir fark yaratamamıştır.

## Özet Bulgular (Maddesel Tasarım)

- Model Karşılaştırması: Geleneksel modeller arasında en stabil ve başarılı yaklaşım XGBoost olmuştur. MLP ise en kötü performansı sergilemiştir.
- Performans Sınırı: En iyi modeller dahi (XGBoost & LSTM), %51-55 bandını aşmakta zorlanmış ve yazı-tura atmaktan farksız sonuçlar üretmiştir.

Neler farklı yapılabilirdi?

- Veri Zenginleştirme: Modele dış veri kaynakları eklemek:
  - Piyasa duygu analizi (Haber başlıkları, sosyal medya).
  - Makroekonomik veriler (VIX, faiz oranları, enflasyon).

Ya da proje konusu daha farklı kurgulanabilirdi belirli bir dönemde fiyatlar artım eğilimi mi gösterecek düşüş eğilimi mi gibi