Finvisor

# AI-DRIVEN STOCK MARKET FORECASTING

## A Hybrid Approach of News Analytics & Time-Series Data

*Insightful Data, Smarter Predictions.*

**ISTANBUL BILGI UNIVERSITY**

2024-2025, Spring

Aslı Gizem Ulusoy,

Ramiz Kadayıfçı,

Umutcan Adıgüzel

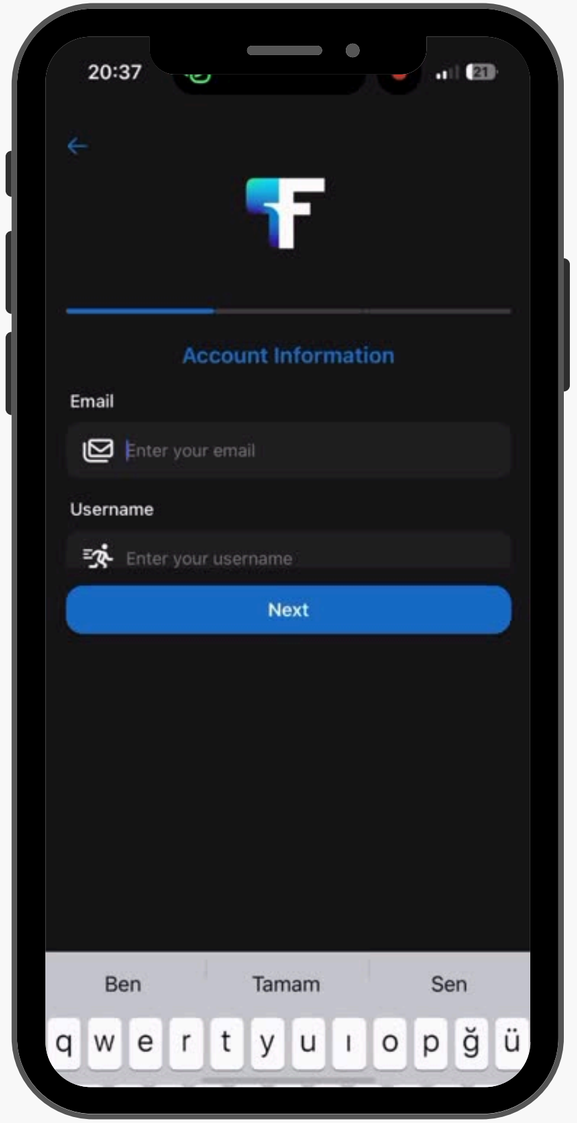
Advisor: Dr. Emel Küpçü

Finvisor

# Problem

“Stock price prediction is highly **complex**. It demands **expertise and accuracy**.”

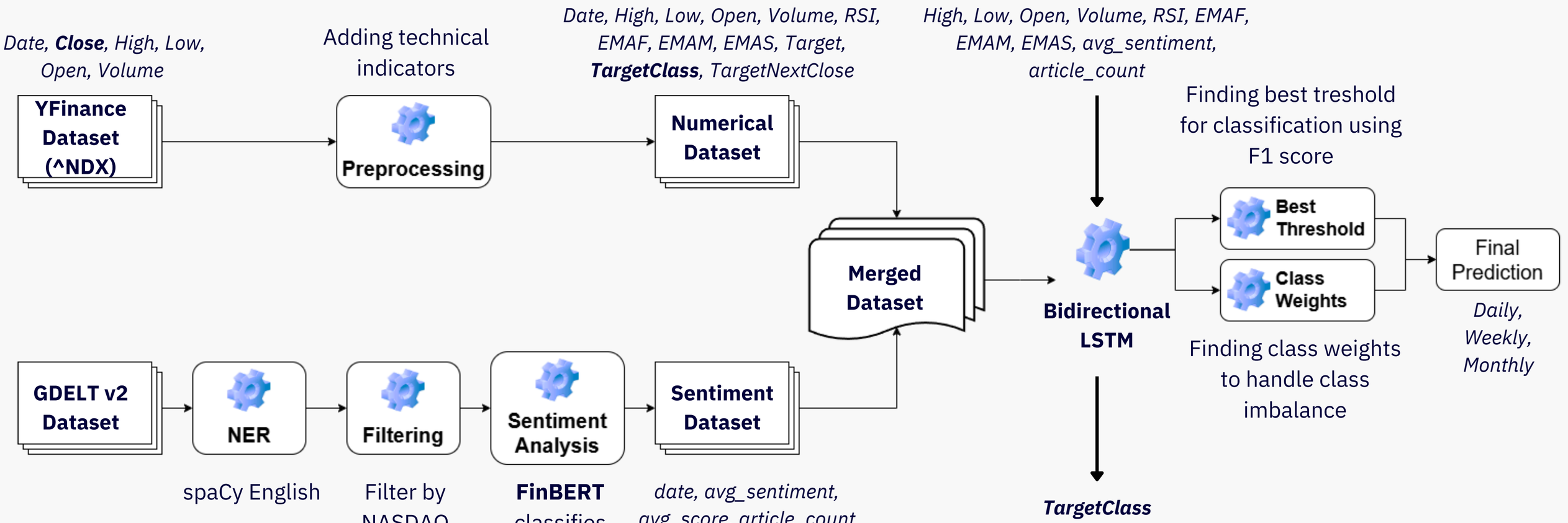
- **Global news** significantly **impacts** market movements.
- Analyzing news impact **requires expertise**.
- Professionals struggle to keep up with dynamic global news.
- Need for **integration** of **historical data** and real-time **news**.




Methodology

# Methodology

System Design




Finvisor

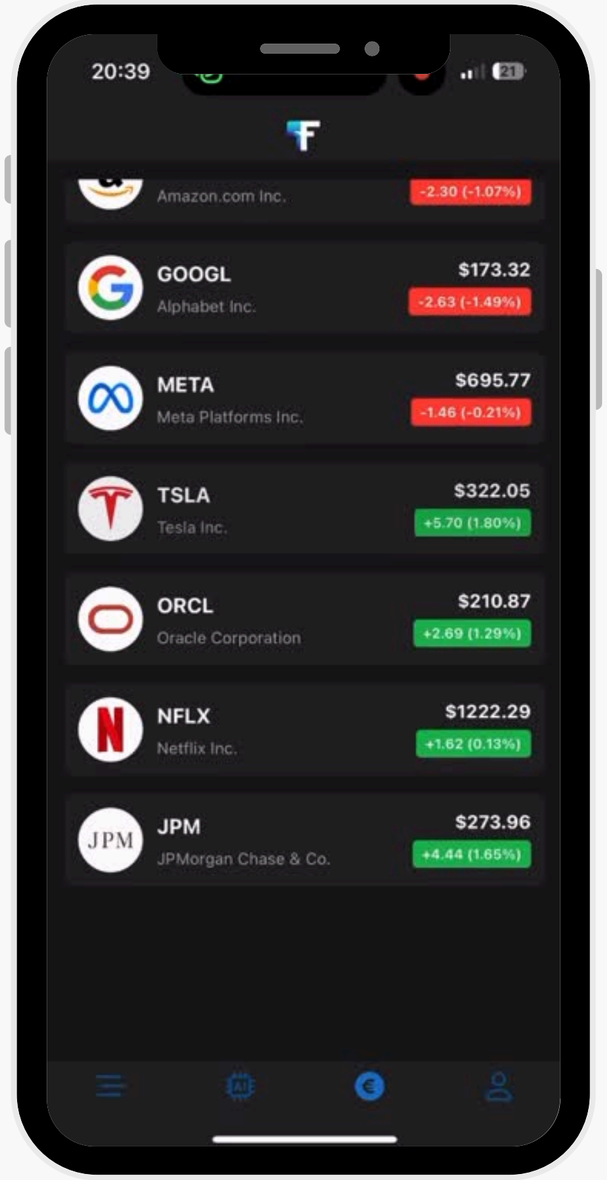
# System Design

**Mobile App (React Native JS)**

- Social Media
- Stock Tracking
- AI Analysis
- Redux State Management

**API Server (Laravel PHP)**

- User API
- Stock API
- Analysis Integration
- Authentication (JWT Token Based)



# Implementation & Challenges

## Progress

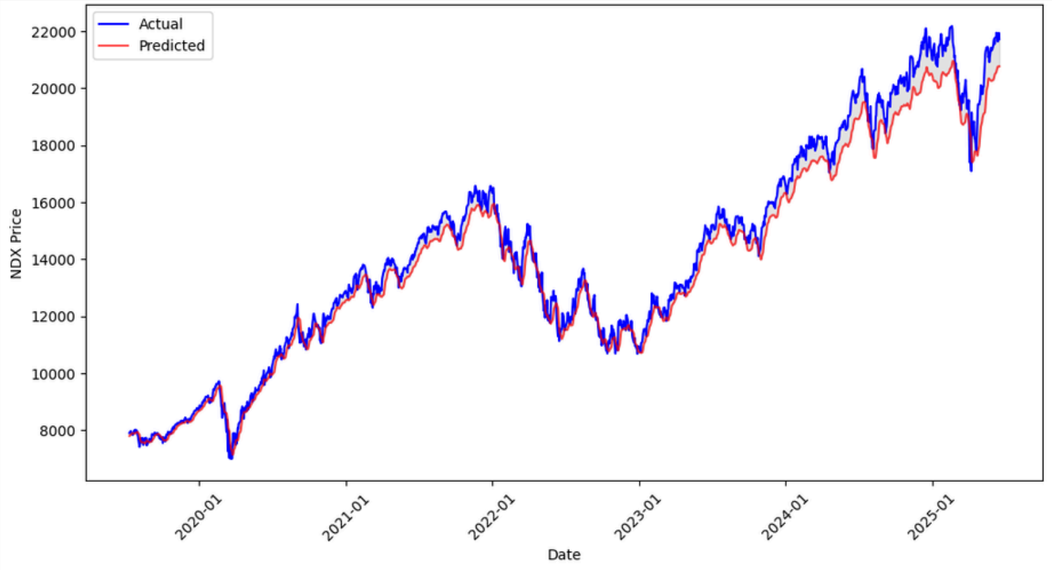
- Numerical analysis module redesigned using target class prediction.
- Previous tests used a small dataset (Pulse); the final version uses and analyzes GDELT v2.
- Previous version lacked sentiment analysis; the final version integrated a FinBERT-based pipeline.
- Integrated social media functions; login/signup , share post with/without media , comment , follow.
- BunnyCDN image upload system.
- Alpha Vintage stock data integration

## Challenges

- Aligning timeframes between news and stock price data.
  - Developed a pipeline.
- Finding relevant datasets with desired inputs
  - GDELT: a large-scale, real-time global news database.
- Strong performance on training data but poor generalization to unseen data.
  - Dropout and early stopping techniques are used.
- Limited free Stock API Token
  - Used 10 Trend Stocks to visualize
- Cross-platform compatibility issues
  - Used native packages

# Results & Testing

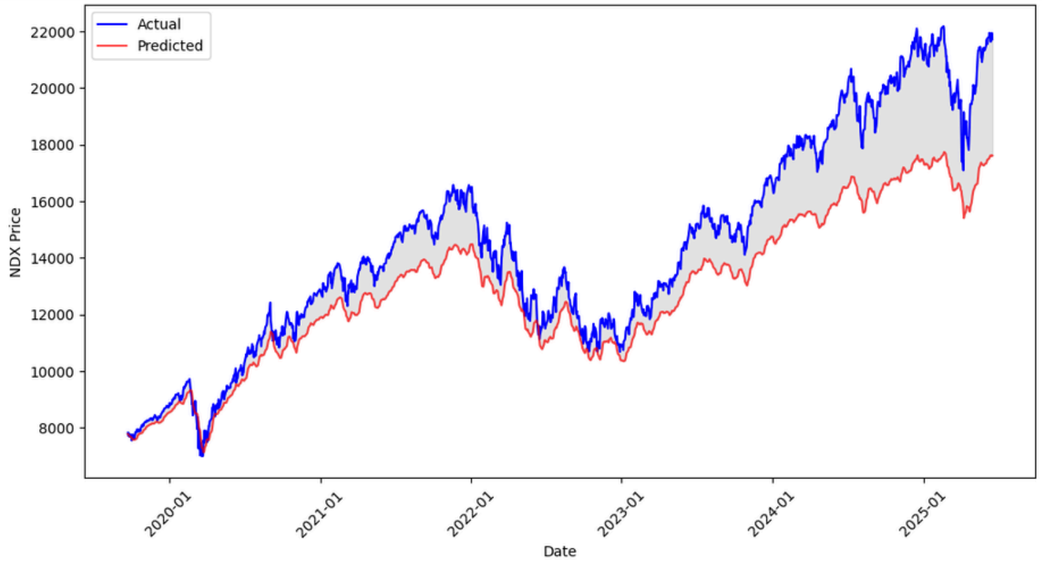
## Full Scaled Data, Close Price Prediction




1985/10/01 - 2025/06/17
10006 rows,
Close Price

R-squared Score: 0.9771
Explained Variance Score: 0.9880
Directional Accuracy: 51.47%
Mape Percentage: **97.00%**

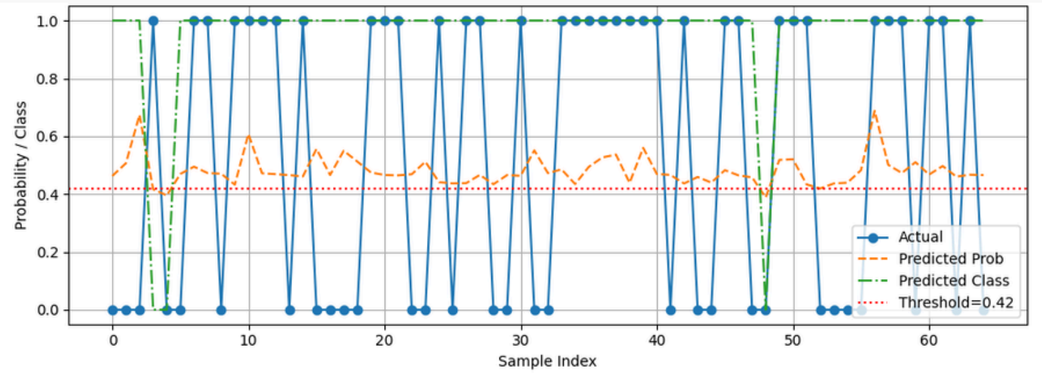
## Seperate Scaled, Close Price Prediction




1985/10/01 - 2025/06/17
10006 rows,
Close Price

R-squared Score: 0.7353
Explained Variance Score: 0.9048
Directional Accuracy: 50.66%
Mape Percentage: **90.46%**

## Seperate Scaled, Target Class Prediction




2025/01/01 - 2025/06/01
101 rows,
TargetClass

**Accuracy: 0.72**
**F1- 0: 0.62**
**F1- 1: 0.78**

Finvisor

# Results & Testing

## TargetClass, Without Sentiment

| Model Name | Recall | F1-score | Accuracy |
|---|---|---|---|
| Logistic Regression | 0: 0.09<br>1: 0.87 | 0: 0.15<br>1: 0.68 | 0.53 |
| Random Forest | 0: 0.23<br>1: 0.78 | 0: 0.30<br>1: 0.66 | 0.54 |
| XGBoost | - | - | 0.47 |
| MLP Neural Net | 0: 0.51<br>1: 0.52 | 0: 0.47<br>1: 0.54 | 0.51 |
| LSTM | 0: 0.26<br>1: 0.71 | 0: 0.31<br>1: 0.62 | 0.51 |

## TargetClass, With Sentiment

| Model Name | Recall | F1-score | Accuracy |
|---|---|---|---|
| Logistic Regression | 0: 0.75<br>1: 0.40 | 0: 0.60<br>1: 0.68 | 0.56 |
| Random Forest | 0: 0.50<br>1: 0.80 | 0: 0.57<br>1: 0.73 | 0.67 |
| LSTM | 0: 0.50<br>1: 0.80 | 0: 0.57<br>1: 0.73 | 0.67 |
| **Bidirectional LSTM** | 0: 0.50<br>1: 0.90 | 0: 0.62<br>1: 0.78 | **0.78** |

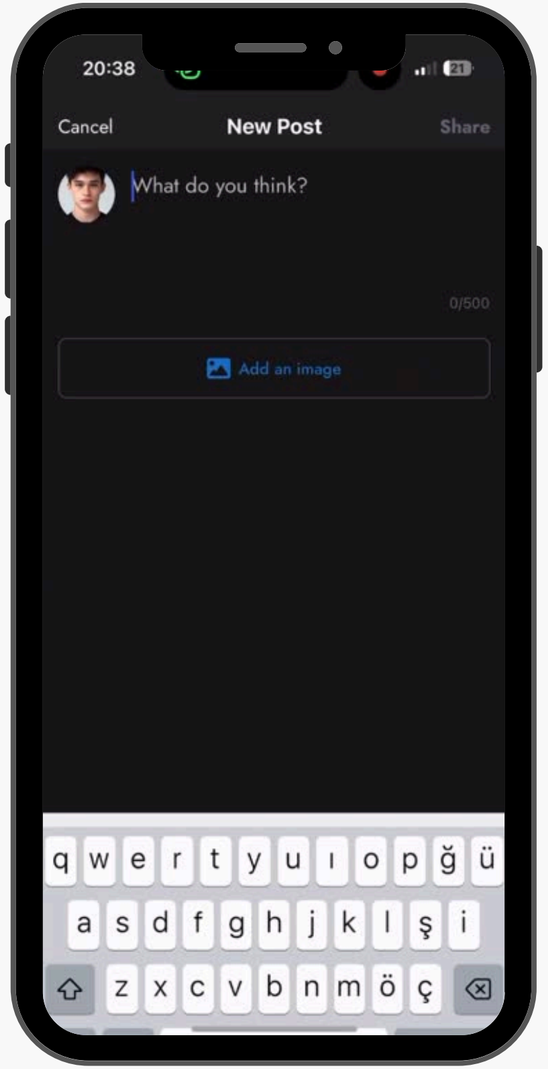
Sentiment Analysis

F1 Score Support

Class Weights

# Conclusion & Acknowledgments

- Predicting **target classes** rather than relying solely on closing prices provides deeper insight into **stock movements.**
- Integrating **sentiment analysis** significantly enhances the model's **predictive accuracy.**
- Stocks listed on indices like NASDAQ generally exhibit an **upward trend** over time, resulting in **class imbalance**. To address this;
  - **F1 score threshold**,
  - **class weighting** is applied to ensure balanced learning.
- Incorporating **technical indicators** (RSI, EMAF, EMAM, EMAS) enhances the model’s ability to capture **stock movement patterns.**
- The **volume** of relevant **news articles** analyzed improves models accuracy.



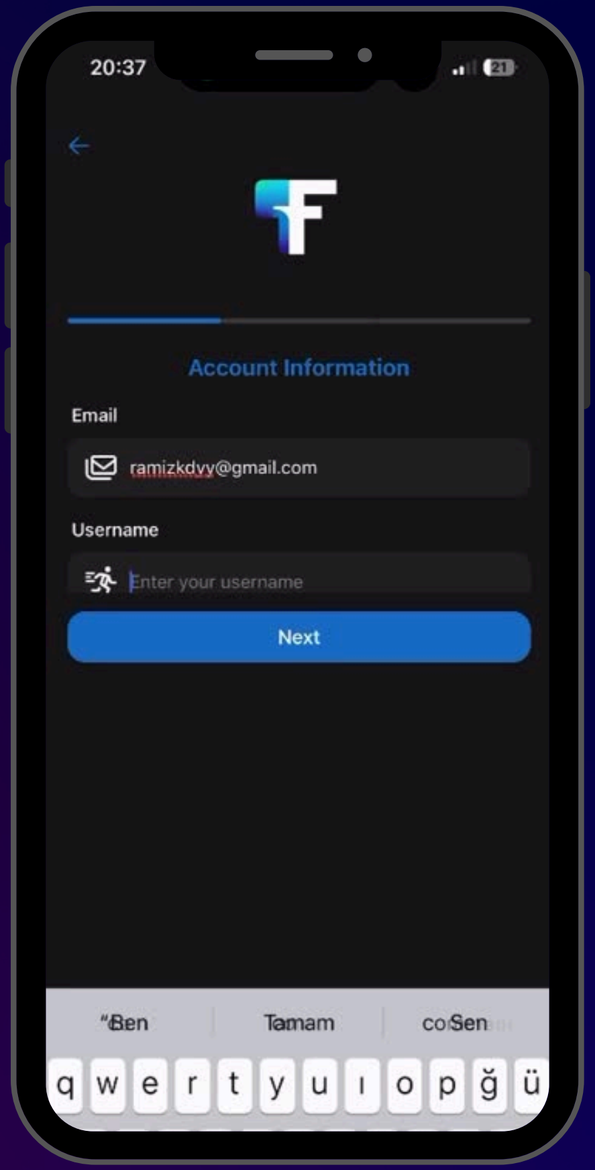
Finvisor

# Future Work & Recommendations

- Instead of binary sentiment (positive/negative), use **granular categories** like *strong positive, neutral, or weak negative* for better accuracy.
- **News impact** can persist for days or months and should be temporally accounted for.
- Ensuring **continuous availability** by fetching data from **multiple sources** and using alternatives when one is unreachable.
- Instead of limited stock APIs, microservices can **fetch full data** from YFinance.
- While broad filters work for **index-level analysis** (e.g., NDX), single-stock analysis (e.g., AAPL) requires **stricter news filtering** for relevance.
- Given the high volatility of the stock market, **using tick data (high-frequency trade data)** instead of daily prices can provide richer insights.
- **Economic statistics** can be added to adapt to global trends.

Finvisor

# DEMO




- Social Media, Post and Comment
- NDX Stock Movement Prediction
- Stock Tracking System
- Create and edit your own profile
- Follow other users

Finvisor

# Thank You! Questions?

**ASLI GIZEM ULUSOY**
121200107

**UMUTCAN ADIGÜZEL**
120200087

**RAMİZ KADAYIFÇI**
120200027

## Refrences

Hochreiter S, Schmidhuber J (1997) Long Short-Term Memory. MIT Press 9(8):1735-1780