

[Home](#)[Service](#)[About Us](#)[Contact](#)

NETFLIX



Stock Price Forecasting System



A Scalable Apache Spark Solution

June 14th, 2025

[Explore Now](#)

01



Team members

Vo Hoai Bao – ITITU21038

Thai Thanh Phat – ITITU21274

Nguyen Phuc Vinh – ITITU21350

Pham Trung Dung – ITITU21007

Nguyen Duc Duy – ITITU21008





TABLE OF CONTENTS

1

Introduction

2

Data Investigation
and Preprocessing

3

Feature Engineering

4

Model Development

5

Model Evaluation

6

Real-Time
Forecasting Pipeline



TABLE OF CONTENTS

7

Performance
Optimization

8

Visualization &
Dashboards

9

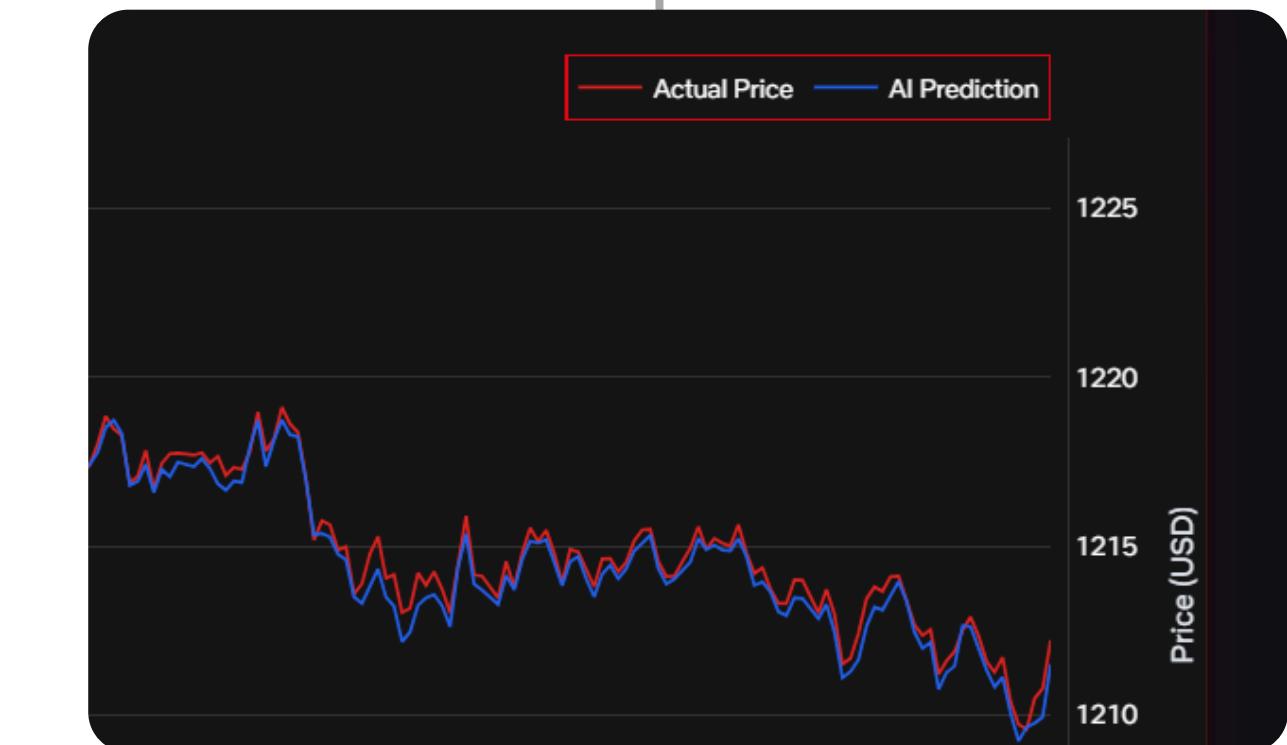
Conclusion



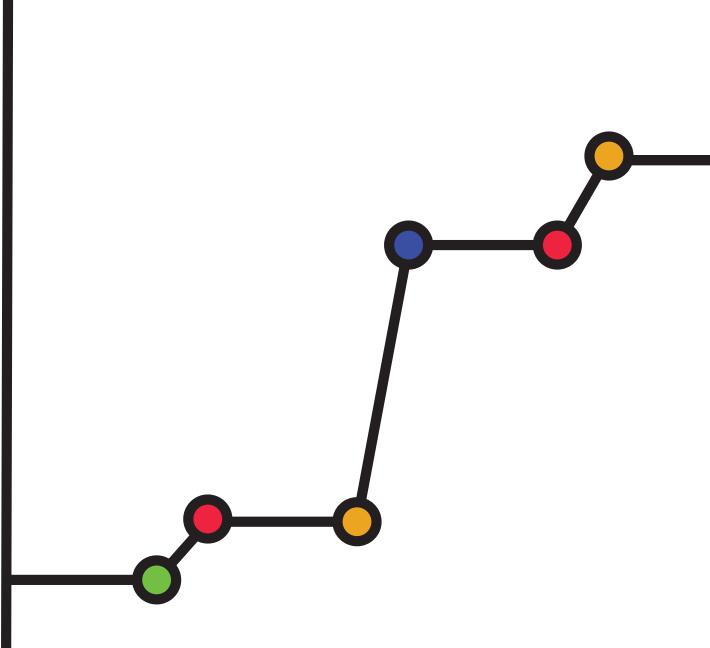
- Supporting Netflix's investment strategies
- Build accurate, real-time stock price forecasting system

Introduction

- Data preprocessing, feature engineering, modeling, deployment
- Limitations: Market volatility, data availability



Data Investigation and Preprocessing



Data Source:
Historical Netflix stock data
(Kaggle dataset)

Handle missing values
Adjust for stock splits and dividends
Align timestamps

Key trends and anomalies identified



Key Features

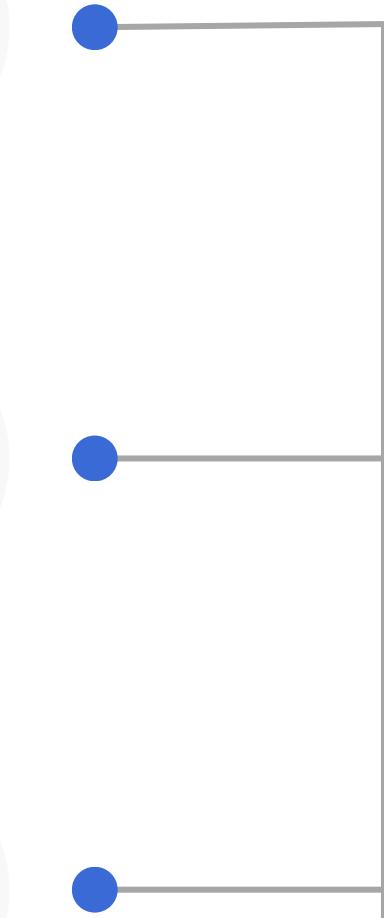
Price data



Volume data

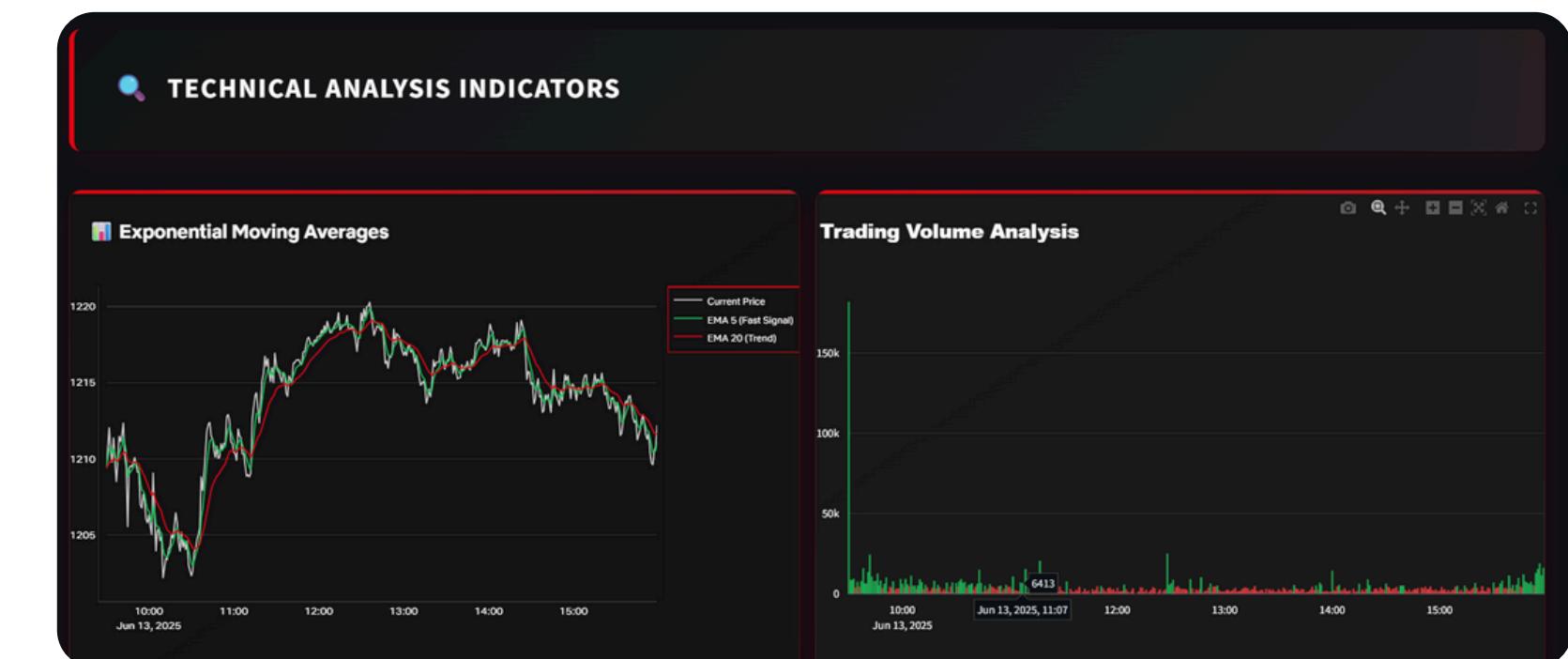


Technical indicators



Rationale

Features selected for predictive power



NETFLIX

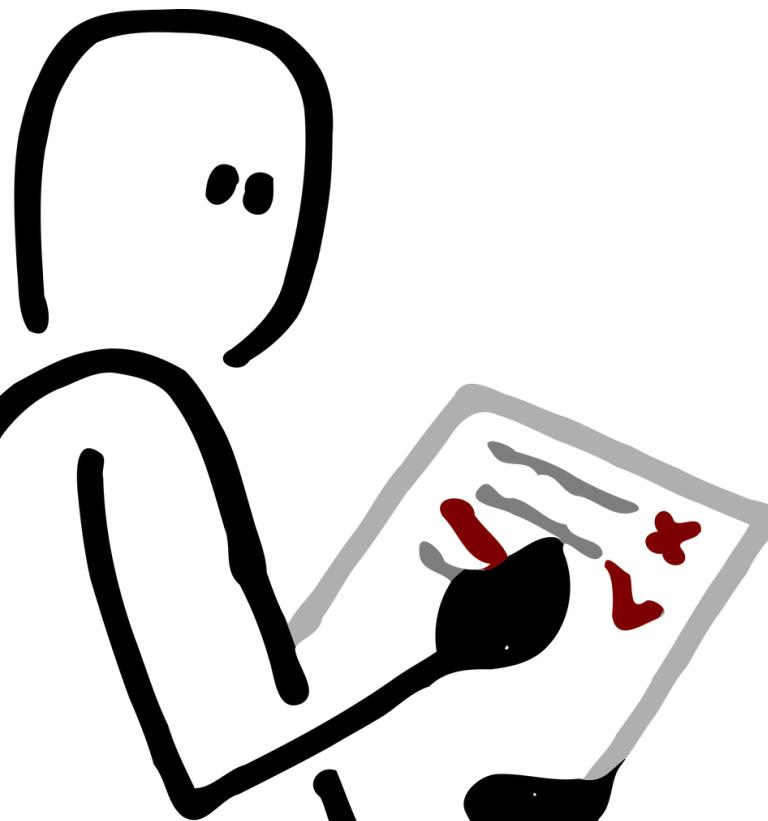
- Approach: Temporal train-test split
- Implementation: Apache Spark for distributed computing

Model:
• Linear Regression (baseline)
• ARIMA/SARIMA (time series)
• LSTM/GRU (deep learning)
• Ensemble and hybrid models



Model Development

Model Evaluation



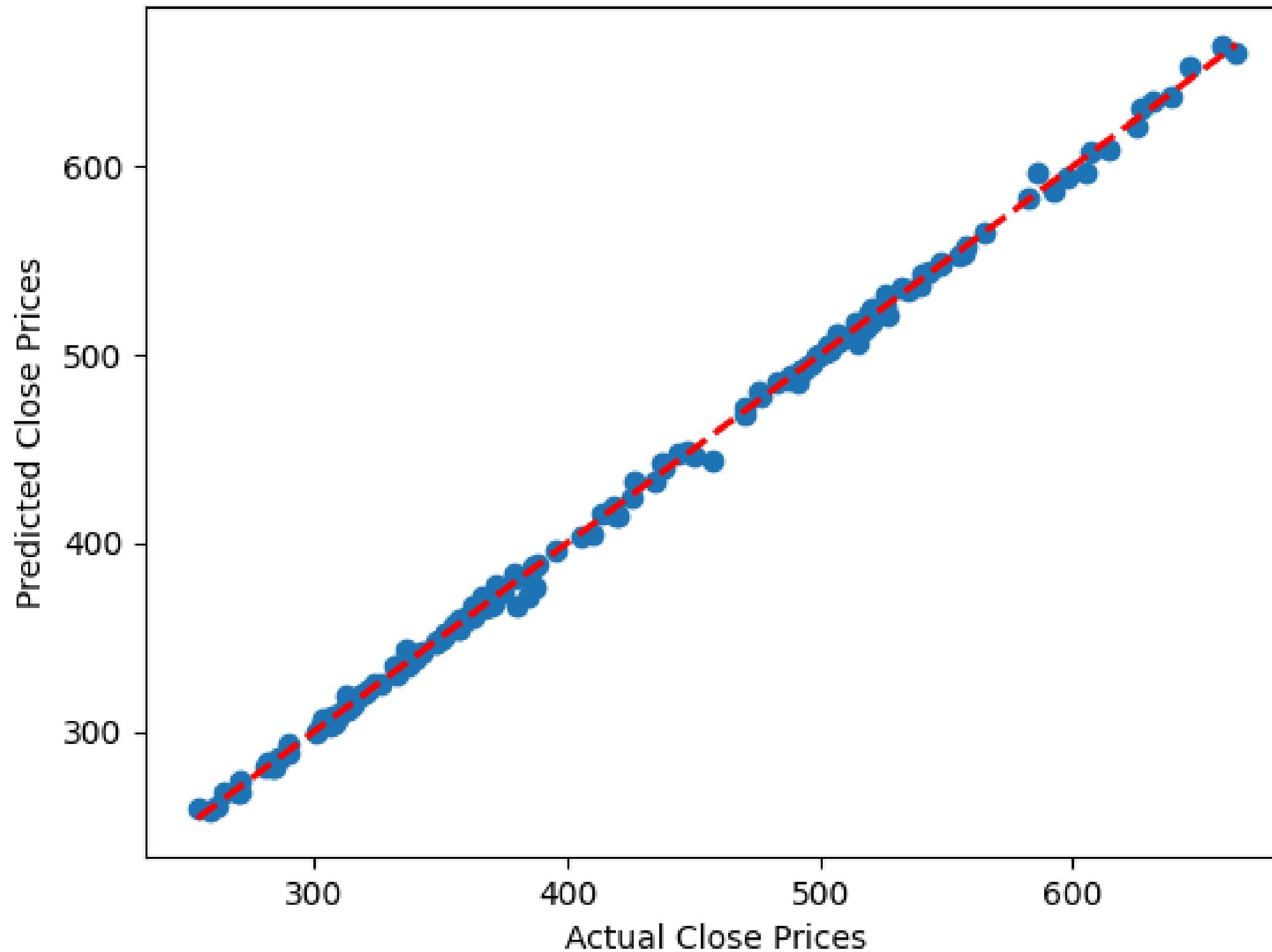
Metrics: MAE, MSE, RMSE, directional accuracy

Results: Performance comparison across models

Backtesting: Historical validation outcomes

Best Model: Linear Regression

Actual vs Predicted Close Prices



Accuracy of Linear Regression model

Real-Time Forecasting Pipeline

1

Architecture:
Spark streaming for
real-time data

3

Output:
Example forecasts

2

Components:
+ Data ingestion
+ Model integration for live predictions
+ Alerting for price movements/anomalies

[Home](#)[Service](#)[About Us](#)[Contact](#)

Stock Price vs AI Predictions

— Actual Price — AI Prediction



Performance Optimization

Techniques

Spark job configurations
(e.g., executor settings)

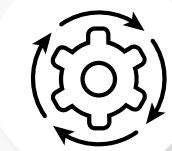


Distributed training and inference



Outcomes

Reduced latency,
improved scalability



Visualization & Dashboards



**Forecasted vs
actual stock
prices**



**Real-time
monitoring for
traders**



**Model
performance
metrics**

CONCLUSION



**Accurate,
scalable
forecasting
system**



**Enhanced
investment
decision-
making**



**Additional data
sources,
advanced models**



Thank You

