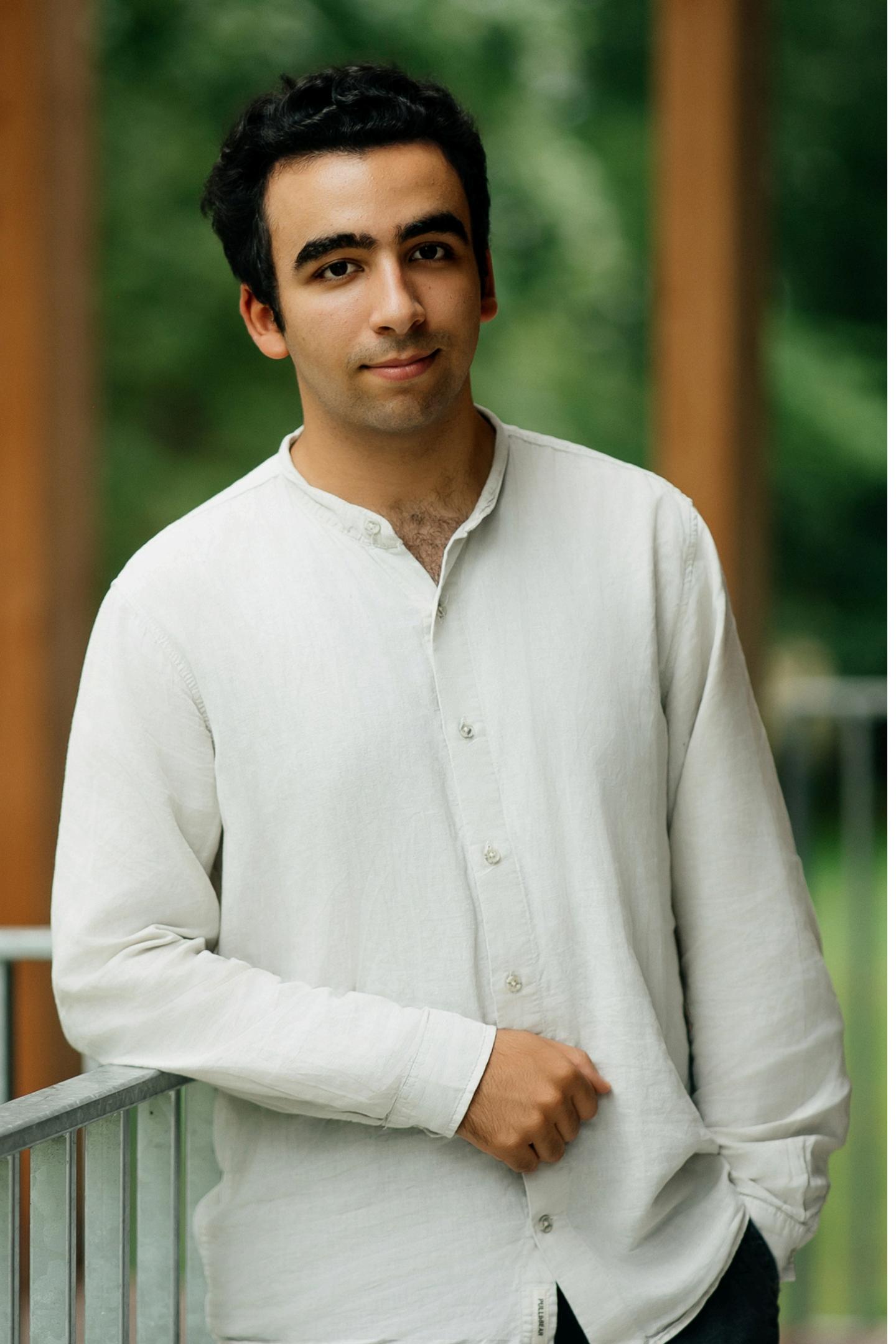


Development of a Prediction Model for NTT Stock Prices

NTT株価予測モデルの開発

Ilias Kalalou – NTT AI Report

A portrait photograph of a young man with dark hair and a slight smile. He is wearing a light-colored, long-sleeved button-down shirt. His left hand is resting on a metal railing. The background is blurred green foliage.

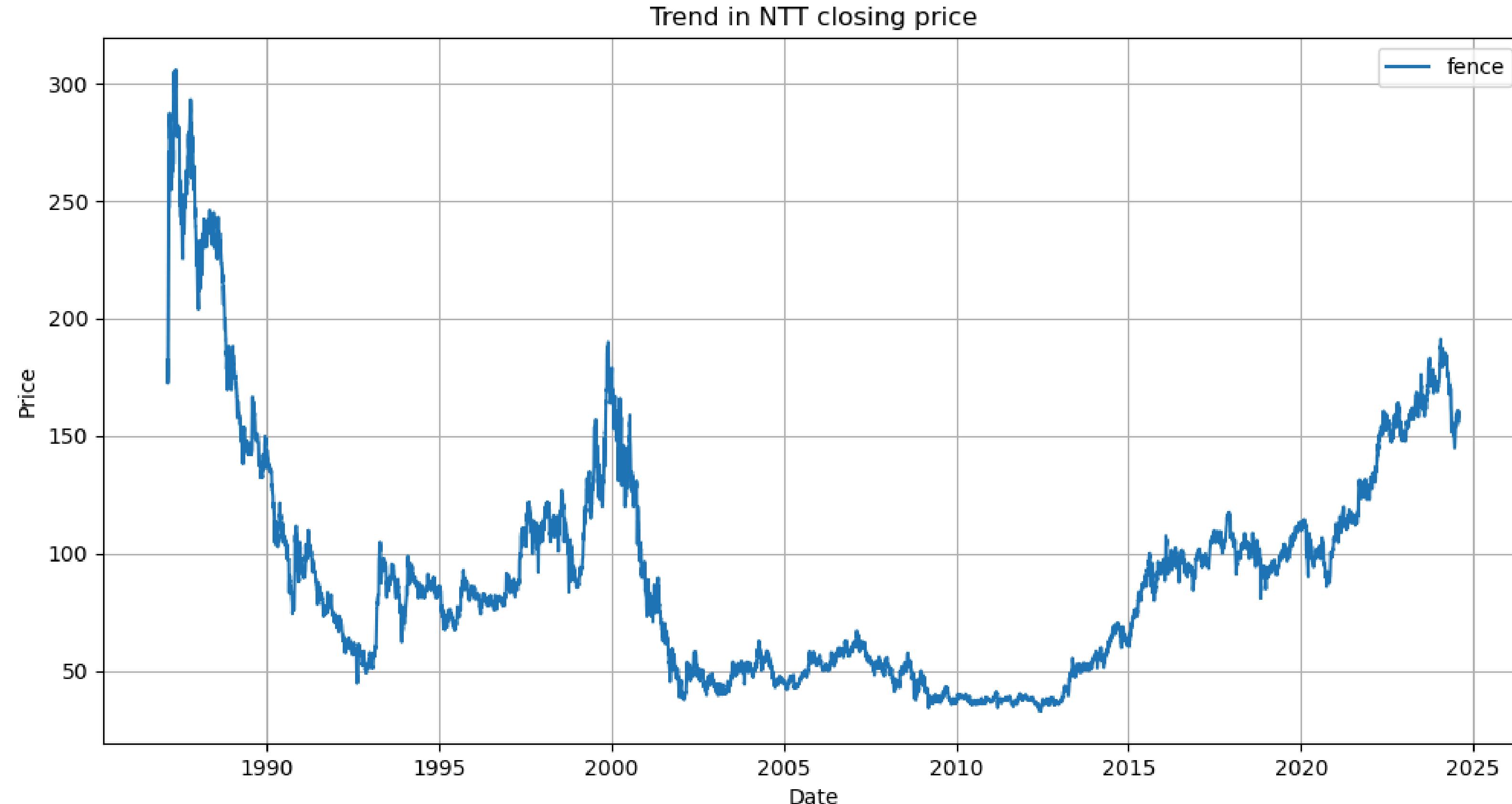
About Me - 自己紹介

- ILIAS KALALOU
- Born in France, 2004 - 2004年にフランスで生まれました
- Currently in the 3rd year at EPITA (equivalent to Bac+3) - 現在、EPITAの3年生で、Bac+3に相当します
- Internship as Lead Developer at a law firm - 弁護士事務所でリードデベロッパーとしてインターンシップ
- My hobbies : Japanese culture, going to the cinema - 日本文化、映画鑑賞

Background and Objectives - 背景と目的

- Importance of stock price prediction and its challenges - 株価予測の重要性とその課題
- Objectives: Understand data, preprocess, develop multiple prediction models, evaluate performance, and refine models. - 目的: データの理解、前処理、複数予測モデルの開発、パフォーマンス評価、モデルの改善

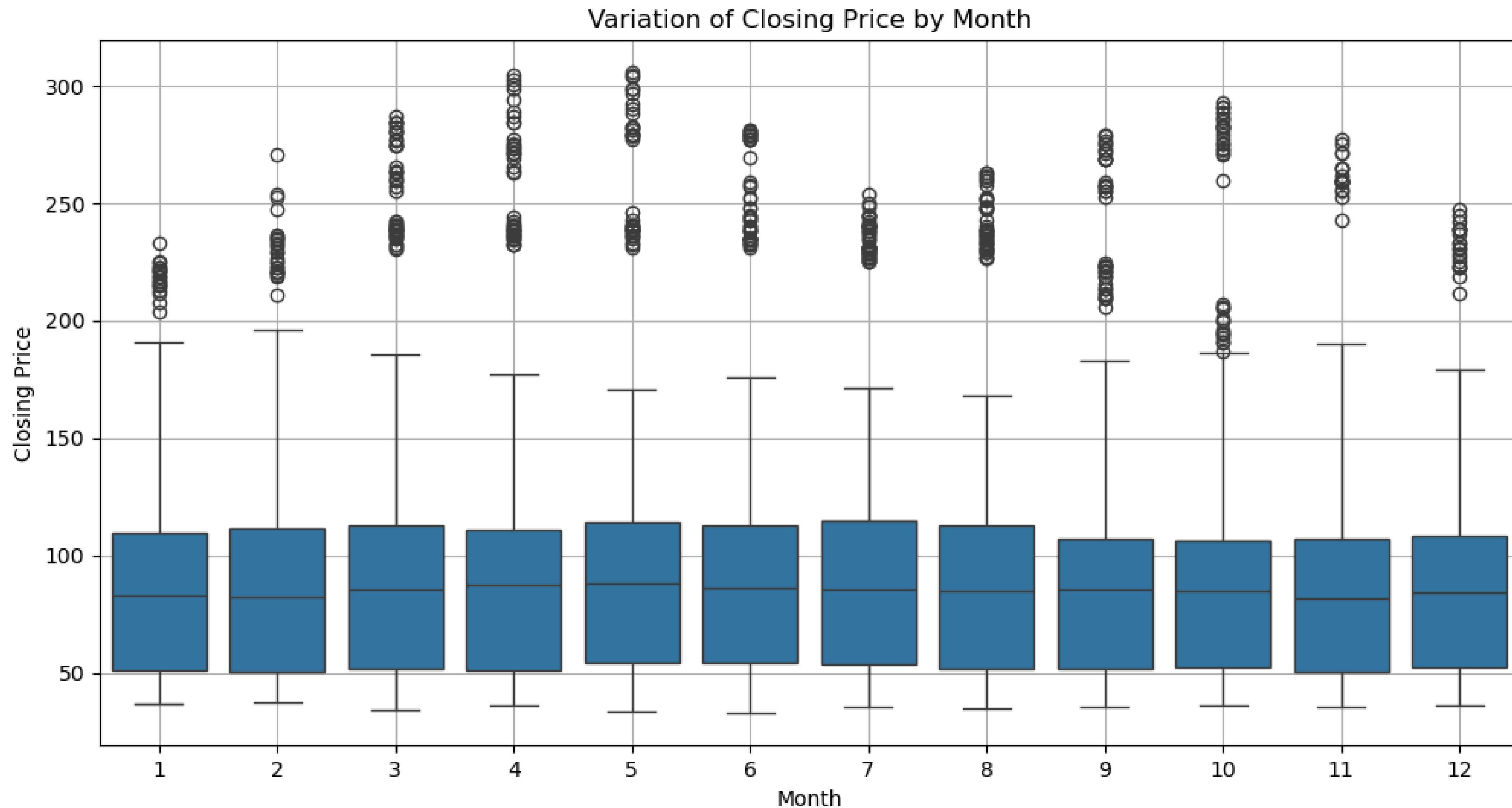
Trend in Closing Price - 終値のトレンド



Distribution of Closing Price - 終値の分布



Monthly Variation of Closing Price - 月ごとの終値の変動



Introduction to AI Models

AIモデルの紹介

Overview of the various AI approaches used in this project (ARIMA, LSTM, GRU, TCN)

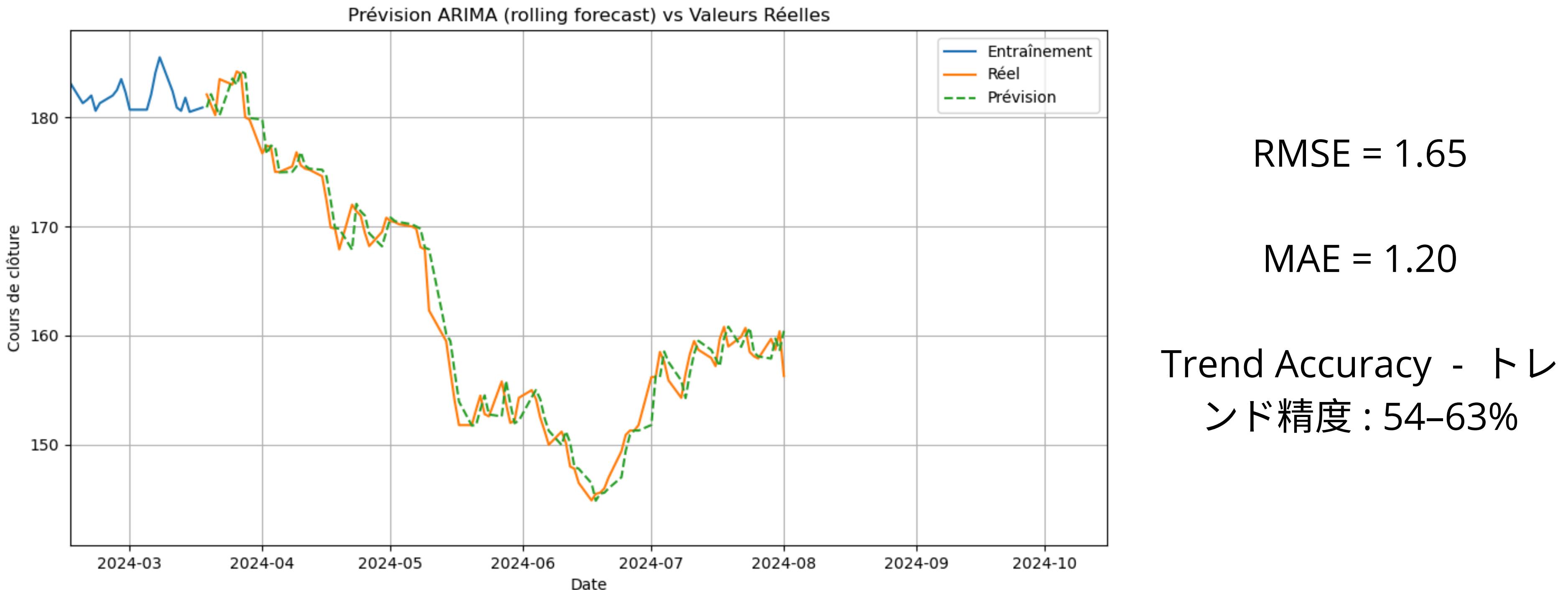
本プロジェクトで使用した各種AI手法 (ARIMA、LSTM、GRU、TCN) の概要

What is ARIMA? - ARIMAとは?

- Explanation: ARIMA stands for AutoRegressive Integrated Moving Average.
 - ARIMAは自己回帰和分移動平均モデルの略です。
- Advantages: Simple, interpretable, and effective for one-step forecasting.
 - シンプルで解釈しやすく、一步先の予測に有効です。
- Disadvantages: Assumes linearity; multi-step forecasts may accumulate errors.
線形性の仮定があり、複数ステップ先の予測では誤差が蓄積する可能性があります。

ARIMA Model – Results and Graph

ARIMAモデルの結果とグラフ



Discussion and Improvement Opportunities

モデルの考察と改善の

Discussion : The ARIMA model achieved a trend accuracy between 54% and 63%, which is robust for financial time series given their inherent volatility and unpredictability.

考察: ARIMAモデルは、54%～63%のトレンド精度を達成しており、これは金融時系列の内在する変動性と不確実性を考慮すると堅牢な性能です。

Future Improvements: Incorporate exogenous variables using an ARIMAX model to include external factors like macroeconomic indicators.

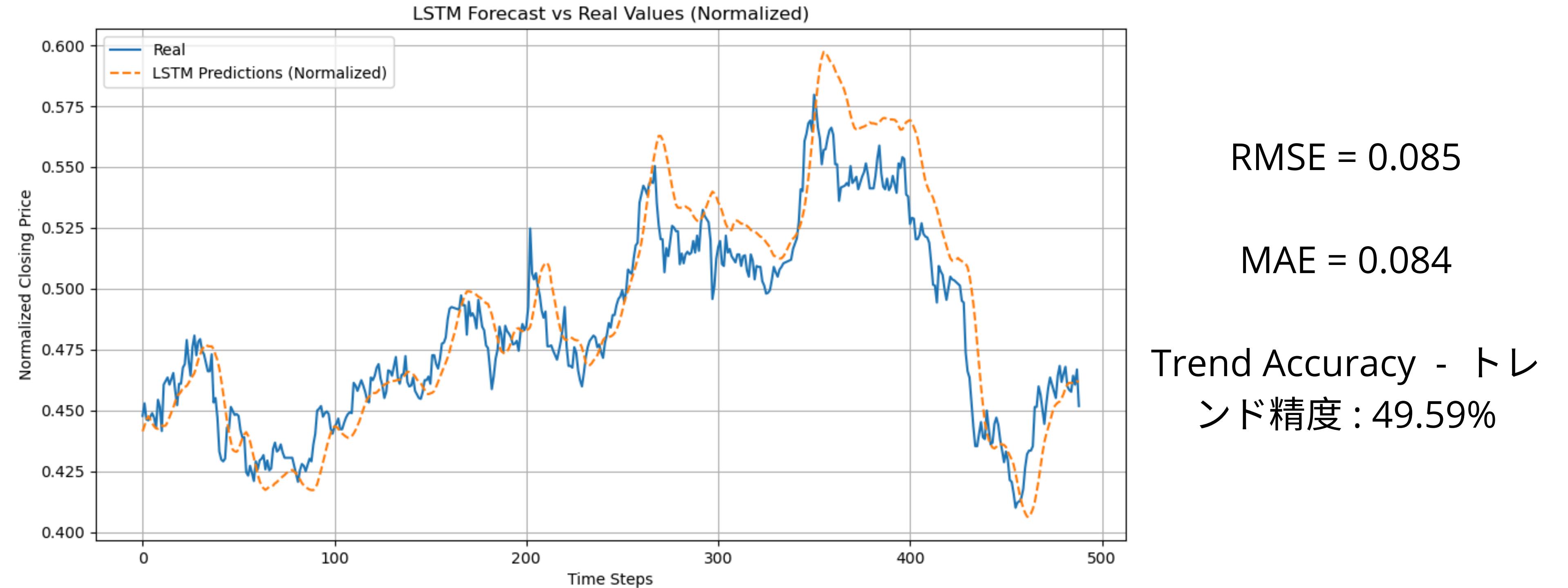
今後の改善点: 外部要因（例: マクロ経済指標）を組み込むARIMAXモデルの採用。

What is LSTM? - LSTMとは?

- Explanation : LSTM is a type of recurrent neural network designed to capture long-term dependencies in time series data by mitigating the vanishing gradient problem.
- LSTM は、勾配消失問題を緩和することで時系列データの長期依存性を捉えるリカレントニューラルネットワークの一種です。
- Advantages : Captures long-term dependencies, Models non-linear relationships.
長期依存性の捕捉が可能 , 非線形な関係をモデル化
- Disadvantages: Computationally intensive, Highly sensitive to data scaling and requires large datasets - 計算コストが高い , データのスケールに敏感で、大量のデータが必要

LSTM Model - Results & Graphs

LSTM モデルの結果とグラフ



Discussion and Improvement Opportunities

モデルの考察と改善の

Discussion : LSTM shows potential for capturing non-linearities, but raw data errors are high while normalized data errors are low due to scaling effects.

考察: LSTMは非線形性を捉える潜在能力がありますが、生データでは誤差が大きく、正規化によりスケールが縮小されています。

Future Improvements: Explore deeper or bidirectional architectures and optimize hyperparameters.

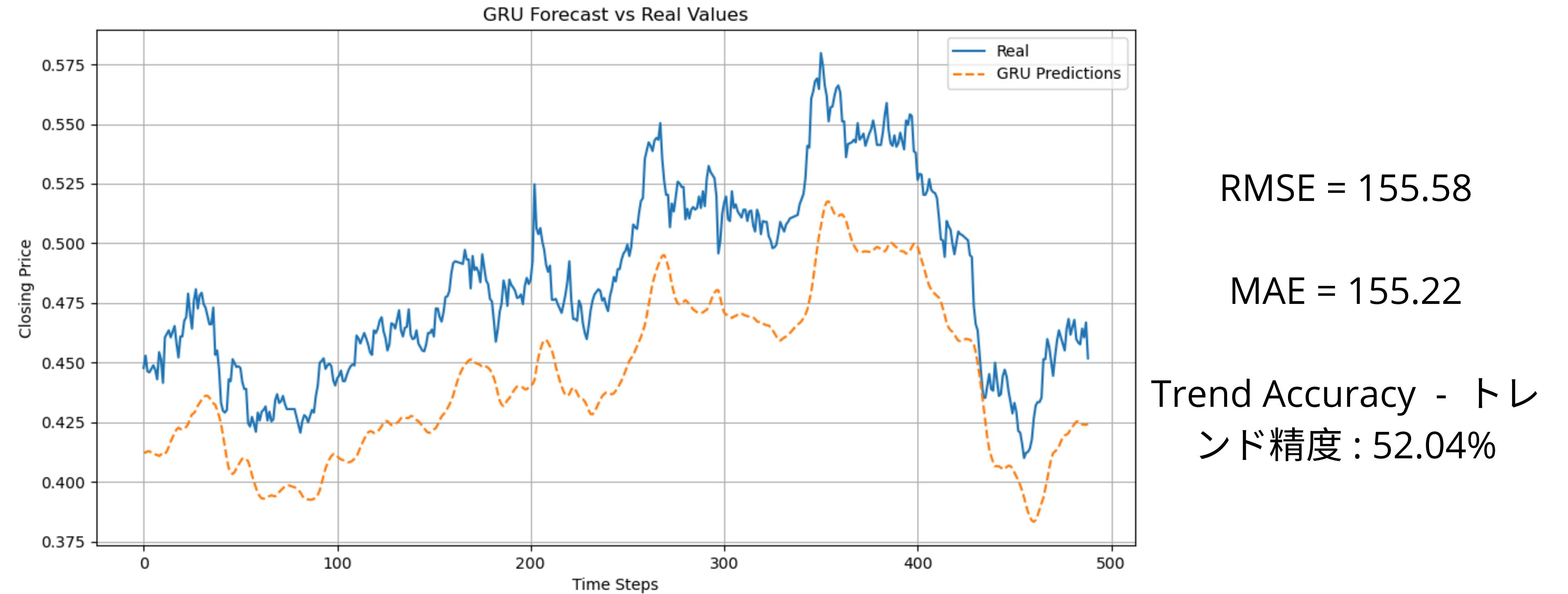
今後の改善点: より深いまたは双方向のアーキテクチャの検討とハイパラメータの最適化。

What is GRU? - GRUとは?

- Explanation : The Multivariate GRU model integrates multiple features to predict the logarithmic closing price, aiming to capture complex interactions in financial data.
多変量GRUモデルは、始値、最高値、最低値、出来高等の複数の特徴量を統合して、対数変換された終値を予測し、金融データの複雑な相互作用を捉えようとするモデルです。
- Advantages : Utilizes multivariate information , Parameter-efficient compared to LSTM , Stable training via gradient clipping.
多変量情報の活用、LSTMに比べパラメータ効率が良い、勾配クリッピングによる安定した学習
- Disadvantages: Highly sensitive to hyperparameters , Requires rigorous data normalization - ハイパーパラメータに非常に敏感、厳格な正規化が必要

GRU Model – Results & Graphs

GRU モデルの結果とグラフ



Discussion and Improvement Opportunities

モデルの考察と改善の

Discussion : The multivariate GRU reduces prediction errors slightly, yet its trend accuracy around 52.04% shows room for improvement.

考察: 多変量GRUは誤差を若干削減しますが、トレンド精度は約48～49%で改善の余地があります。

Future Enhance feature engineering and fine-tune hyperparameters.

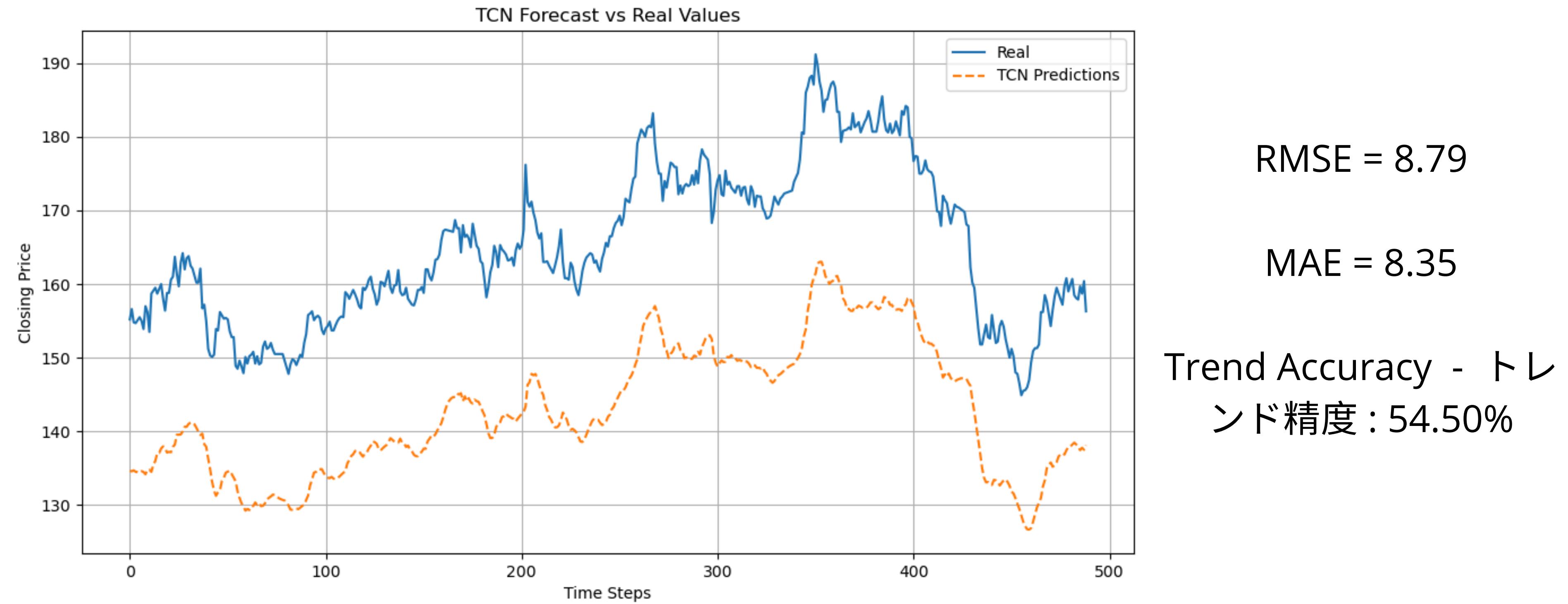
今後の改善点: 特徴量エンジニアリングの強化とハイパラメータの微調整。

What is TCN? - TCNとは?

- Explanation : TCN (Temporal Convolutional Network) uses dilated convolution blocks to capture time dependencies at multiple scales without recurrent connections.. -
TCN（時系列畳み込みネットワーク）は、拡張畳み込みブロックを使用して再帰的な接続を用いずに複数のスケールで時系列依存性を捉えます
- Advantages : Captures multi-scale dependencies, Allows parallel processing, Stabilizes gradients - マルチスケールの依存性を捕捉 , 並列処理が可能, 勾配が安定
- Disadvantages: Sensitive to kernel size, number of layers, dilation, and dropout settings. Less interpretable compared to classical linear models - カーネルサイズ、層数、拡張率、ドロップアウト設定に敏感 従来の線形モデルに比べ解釈が難しい

TCN Model – Results & Graphs

TCN モデルの結果とグラフ



Discussion and Improvement Opportunities

モデルの考察と改善の

Discussion : TCN effectively captures multi-scale dependencies, but its sensitivity to scaling requires careful parameter tuning.

考察: TCNはマルチスケールの依存性を効果的に捉えますが、スケーリングに敏感なためパラメータ調整が必要です。

Future : Optimize dilation rates, dropout, and residual connections.
今後の改善点： 拡張率、ドロップアウト、残差接続の最適化。

Conclusion - 結論

Thanks.