**School of Information, Computer and Communications Technology**

**Sirindhorn International Institute of Technology**
**Thammasat University**

# CSS483 Financial Engineering

**Group Assignment 1**
**Section 1**

## Time Series Forecasting Assignment:
## Nikkei 225 Stock Price and Return

**22 June 2025**


| | |
|---|---|
| **6522770757** | **Payachai Jewjinda** |
| **6522771177** | **Aritsara Sermsathonsawat** |
| **6522781739** | **Nattakit Patcharapairoj** |
| **6522790094** | **Wassawalee Phakphiensakunchai** |


**Semester 3 / 2024**

## Objective

The objective of this report is to apply, compare, and evaluate a range of time series forecasting models in both traditional statistical methods and modern deep learning approaches on the Nikkei 225 (^N225) index data spanning from January 2000 to May 2025. This study aims to forecast both the closing prices and the returns of the index, assessing model performance based on appropriate financial and statistical metrics. Through this analysis, we seek to justify the applicability and limitations of each model in a financial engineering context, enhancing our understanding of their practical use in financial decision-making and risk management.

# Table of Content

# 1. Introduction

Forecasting stock prices is a central challenge in financial time series analysis. If future movements of an index like the Nikkei 225 can be predicted from historical data, then technical analysis could offer profitable trading strategies. However, the Efficient Market Hypothesis (EMH) argues that past data is already reflected in current prices which makes such predictions unreliable.

In this project, we evaluate a range of forecasting models on Nikkei 225 data from 2000 to 2025, using both closing prices and daily returns. We begin with baseline models (naive, SMA, EWMA), then explore classical methods (Holt-Winters, ARIMA, SARIMAX), and finally apply advanced approaches like Prophet and N-BEATS. We also include the USD/JPY exchange rate as an external regressor.

Performance is assessed using RMSE, MAE, and other metrics, with a focus on both forecast accuracy and model behavior during events like the 2008 crash and COVID-19. This allows us to examine the strengths and weaknesses of each approach and test whether meaningful patterns can truly be extracted from historical price data.

## 2. Data Loading and Preprocessing

We began by retrieving daily closing price data for the Nikkei 225 Index (^N225) using the Yfinance package. The dataset spans from January 2000 to July 2025, providing a sufficiently long time horizon to capture major financial events, seasonal patterns, and long-term market behavior.

As part of initial data cleaning, we dropped all rows with missing values and confirmed the data index was a valid DatetimeIndex. To ensure consistency with other time series (particularly exogenous variables like exchange rates), we maintained a business day frequency (freq='B'), since the stock market is only open on weekdays.

We computed daily percentage returns using the formula:

$$Return(t) = P(t) - P(t - 1) * 100/P(t - 1)$$

These returns were used as the target series for models that forecast return behavior rather than absolute price levels. In addition to raw returns, we also applied a logarithmic transformation to stabilize variance, calculated as:

$$Log\ Return(t) = log(P(t)/P(t - 1))$$

To better understand the characteristics of the time series, we conducted exploratory data analysis (EDA). This included visualizations of price trends, return distributions, and volatility patterns. We also performed seasonal decomposition using a 252-day window (approximately one trading year) to investigate potential trend and seasonal components in the adjusted close price.

Lastly, a 30-day rolling standard deviation was used to estimate market volatility over time, which is essential for understanding the structure of financial risk and guiding model selection. This preprocessing step ensured our dataset was clean, stationary where needed, and ready for use across a variety of forecasting models.

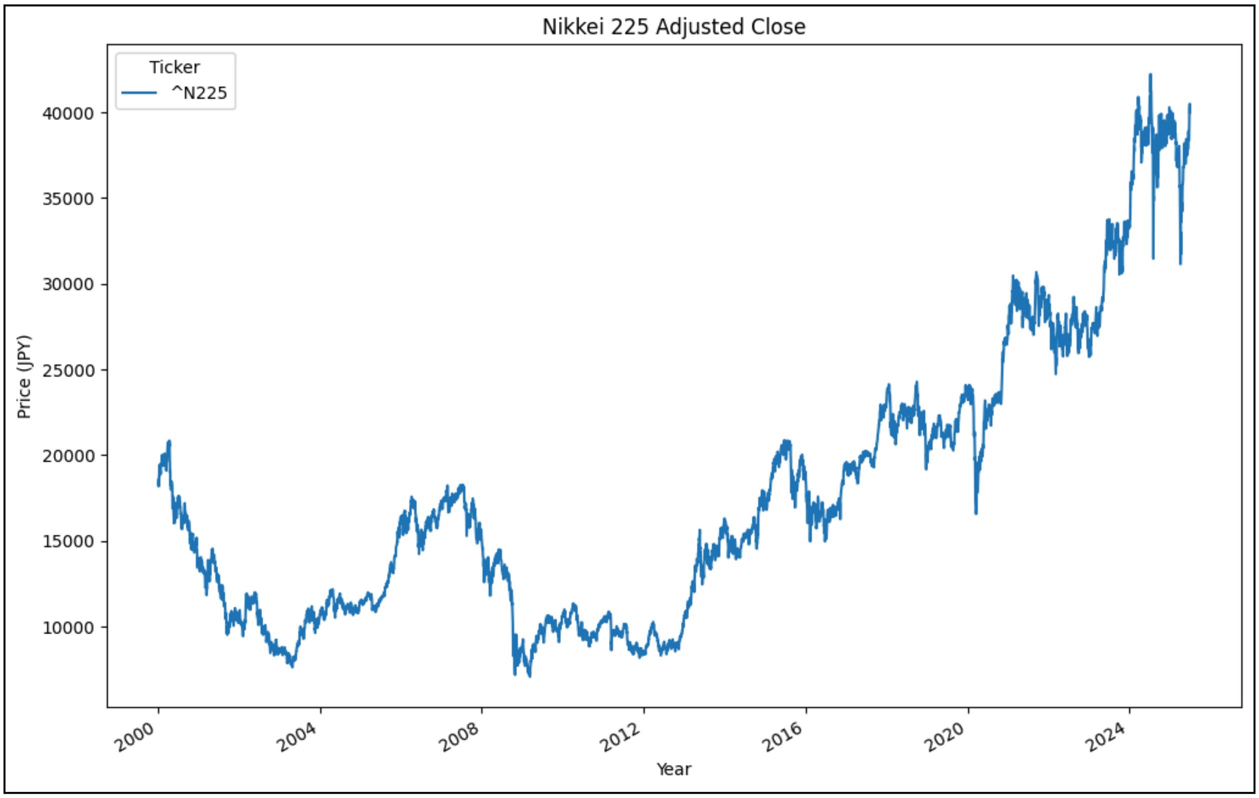


Figure 1: Adjusted Close Price of Nikkei 225

In Figure 1, the graph displays the adjusted closing prices of the Nikkei 225 index (^N225) from Yahoo Finance, covering daily data from January 2000 to May 2025. It clearly shows the long-term trend, with a downward movement in the early 2000s, followed by a period of stability and a strong upward trend from around 2013 onward. The graph also highlights sharp price increases and fluctuations after 2020, effectively capturing the overall performance of the index over time.

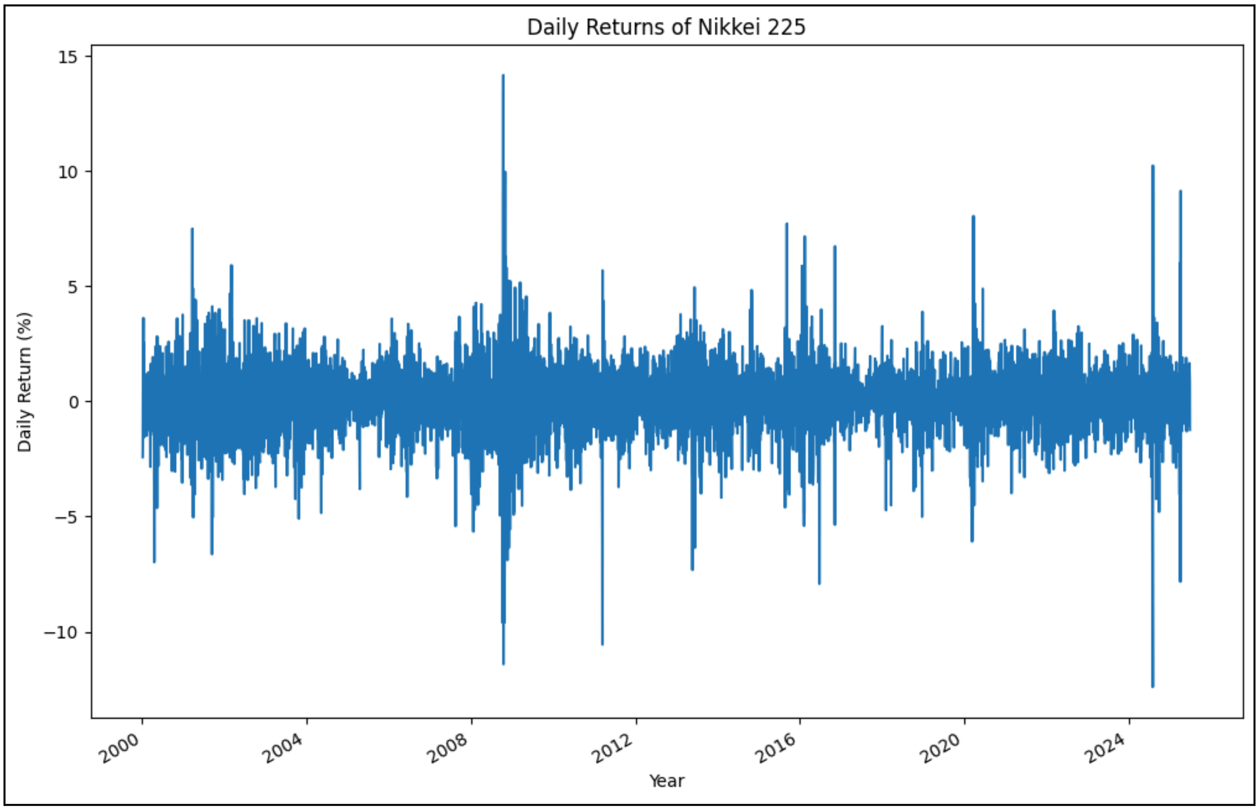


Figure 2: Daily Returns of Nikkei 225

This plot displays the daily percentage changes in the Nikkei 225 index. The data fluctuates around 0%, reflecting typical financial return patterns with volatility clusters during turbulent periods.

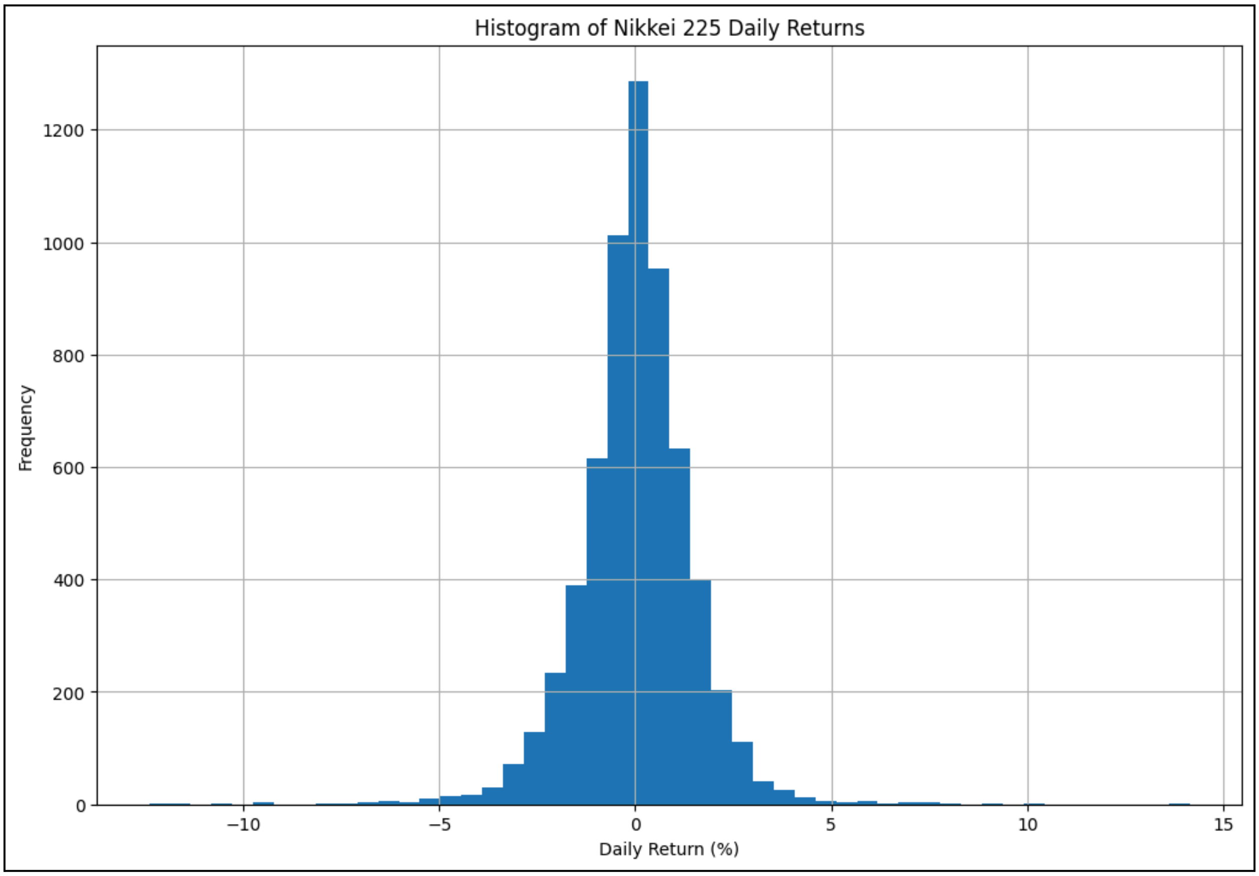


Figure 3: Histogram of Daily Returns

This histogram visualizes the distribution of daily returns. It reveals a shape similar to a normal distribution but with fat tails, suggesting frequent extreme movements in the market.

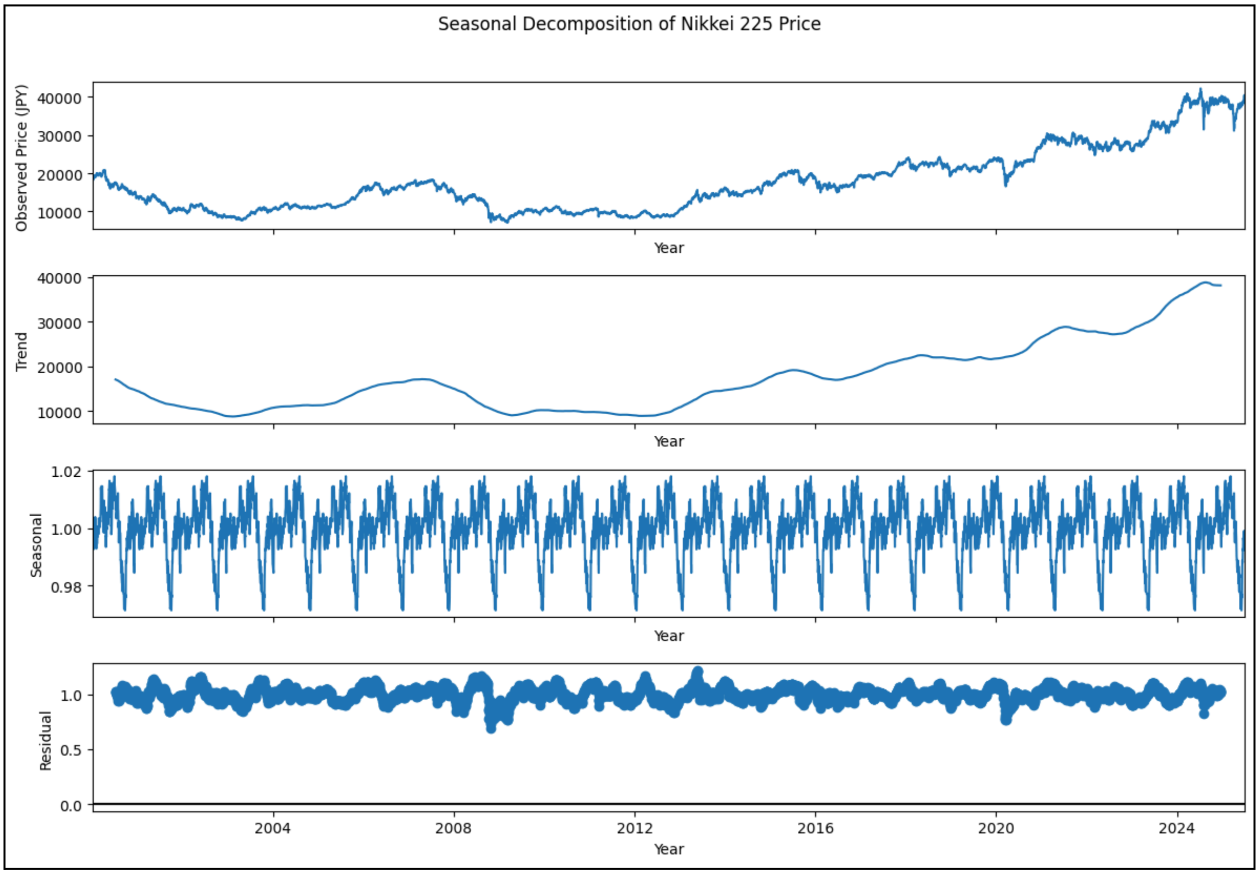


Figure 4: Seasonal Decomposition of Nikkei 225 Price

The top panel shows the observed price series. The trend component captures long-term structural movement, particularly the post-2012 rally. The seasonal component shows minor but regular patterns. The residual captures high-frequency noise and market shocks not explained by trend or seasonality.

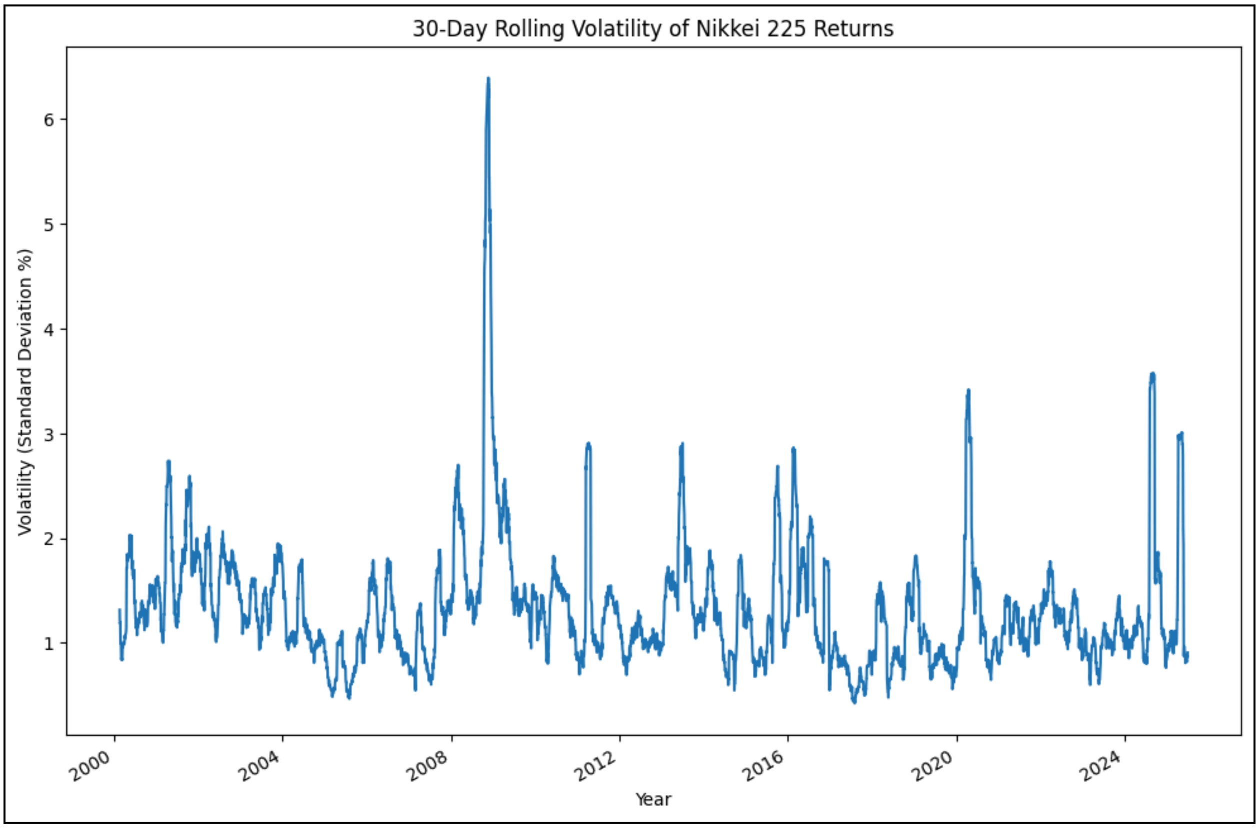


Figure 5: 30-Day Rolling Volatility of Returns

This chart shows the short-term realized volatility (standard deviation of daily returns) using a 30-day moving window. Volatility spikes significantly during major market disruptions especially the 2008 financial crisis and the COVID-19 pandemic highlighting periods of increased market uncertainty. The series exhibits typical volatility clustering with long periods of calm punctuated by short, sharp volatility surges.

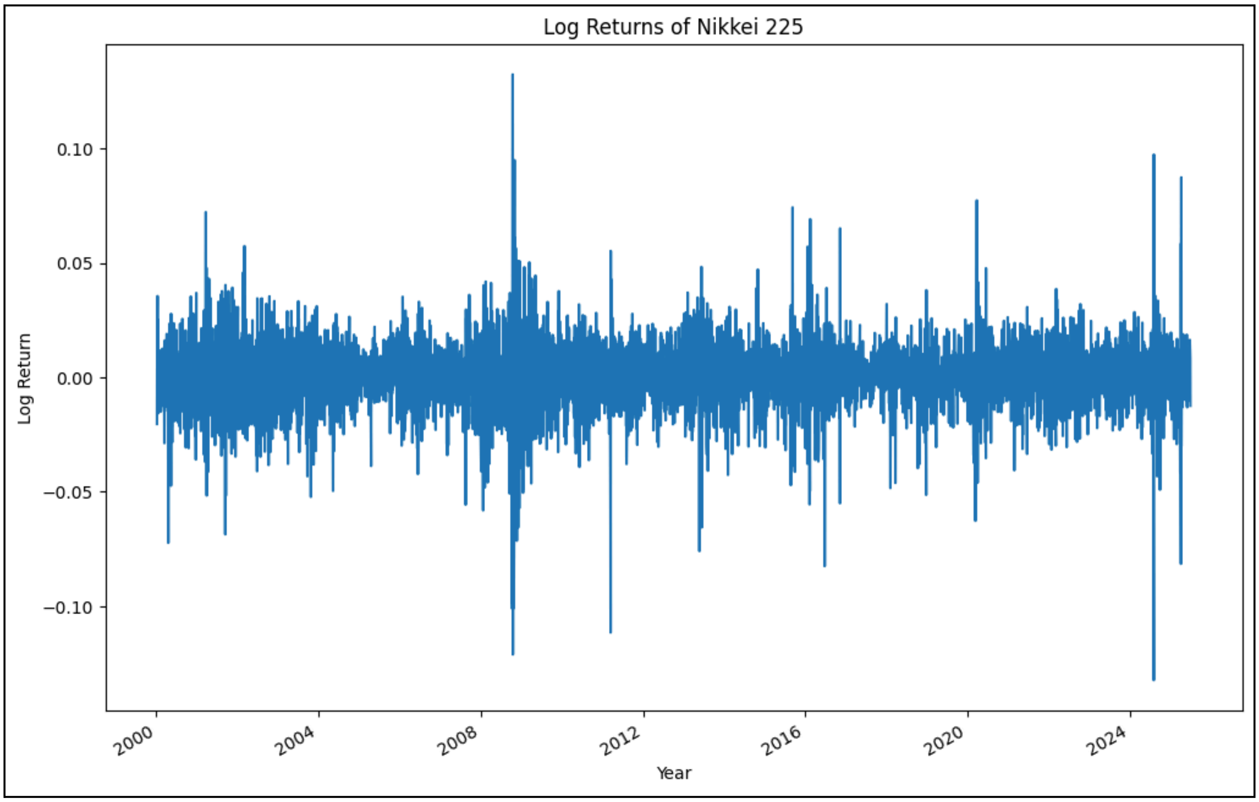


Figure 6: Log Returns of Nikkei 225

The time series of log returns fluctuates around zero and displays volatility clustering. Major spikes coincide with global financial shocks (e.g., 2008 crisis, COVID-19). The return series appears approximately stationary but exhibits fat tails and time-varying volatility, justifying the use of models that accommodate heteroskedasticity or nonlinearity.

# 3. Benchmark Forecasting Models

Baseline forecasting serves to create a simple reference model against which more sophisticated forecasting methods can be evaluated. If a complex model fails to outperform the baseline, it may indicate that the model isn't effectively capturing key patterns or that no further predictive value can be extracted from the data. Model performance is assessed using metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), where lower values reflect higher accuracy.

## a. Naive Forecast

The Naive Forecast is the simplest forecasting method where the forecast for the next time period is assumed to be equal to the value of the most recent observation. It assumes that the most recent value is the best predictor of future values.

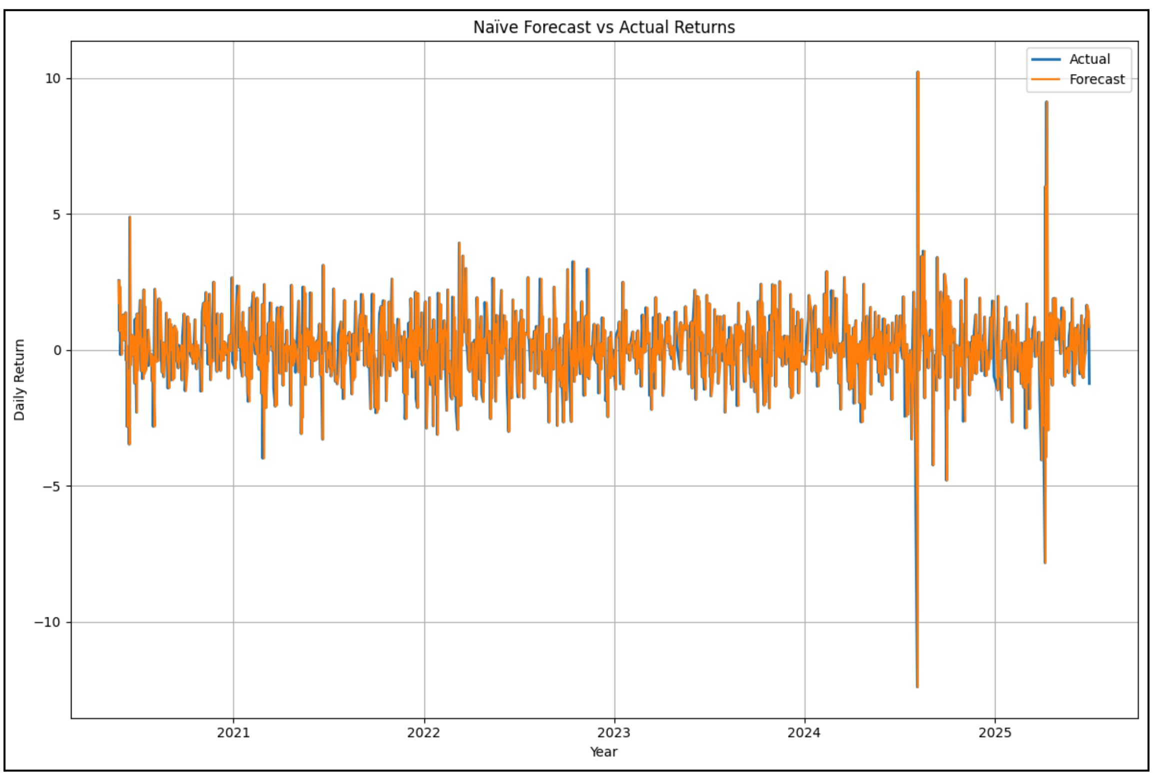


Figure 7: Naive Forecast vs Actual Returns

This figure is the Naive Forecast plot that compares actual daily returns of the Nikkei 225 with a naive forecast, which assumes that tomorrow's return equals today's. Visually, the forecast closely follows the actual returns, suggesting a high correlation. However, this is expected, as the naive model simply lags actual returns.

Despite the visual fit, the model has weak predictive power. Financial returns are noisy and often resemble white noise, especially during volatile periods like 2024–2025. The model cannot anticipate sudden market shocks and thus performs poorly when returns deviate sharply. This highlights the limitations of using naive forecasts in financial return prediction.

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.9544 | 1.3562 | 467.8618% | 0.8757 |


Figure 8: Naive Forecast vs Actual Nikkei 225 Closing Price

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 421.929526 | 289.655939 | 0.934248% | 2.002740 |

## b. Naive Forecast with Random Walk

The Naive Forecast with Random Walk assumes that changes in the time series are completely random and unpredictable. Each value in the series is equal to the previous value plus a random shock or noise. This is commonly used in modeling stock prices or other financial data.

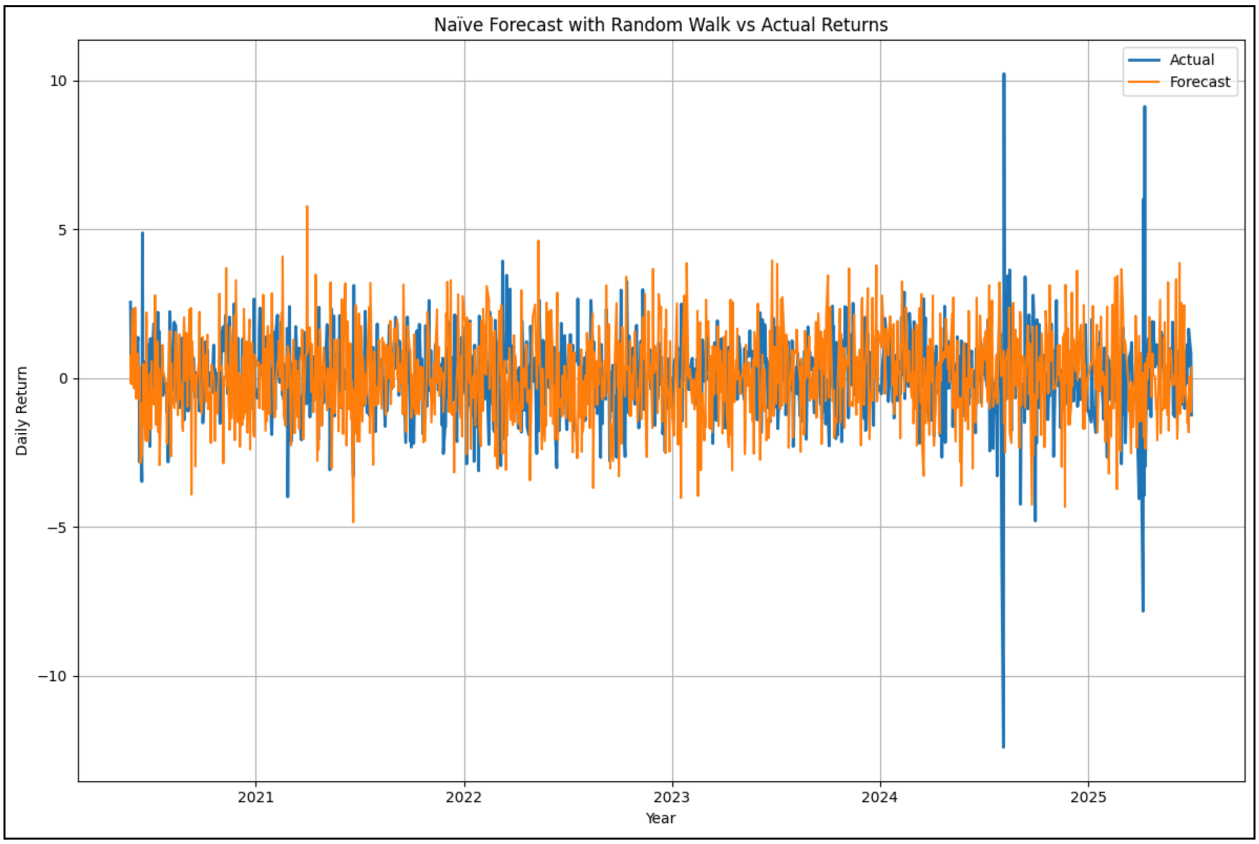


Figure 9: Naive Forecast with Random Walk vs Actual Returns

This figure presents a comparison between a naive forecast based on a random walk model and the actual daily returns of the Nikkei 225 Index (^N225) from January 2000 to May 2025. The graph illustrates that both the forecasted and actual returns fluctuate around the zero axis with frequent short-term variations, characteristic of financial time series data. While the forecast occasionally aligns with the actual returns, particularly during periods of low volatility, significant deviations occur during more volatile periods.

This suggests a weak short-term correlation between the two, as expected under the efficient market hypothesis, which posits that price changes are largely unpredictable and follow a random walk. The model's inability to account for sudden market movements, volatility clustering, and other non-linear dynamics highlights the limitations of naive forecasting in financial markets. Overall, the graph demonstrates that while the random walk approach may serve as a basic benchmark, it lacks the complexity required to accurately predict daily market returns.

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.9856 | 1.5271 | 725.4889% | 0.9860 |

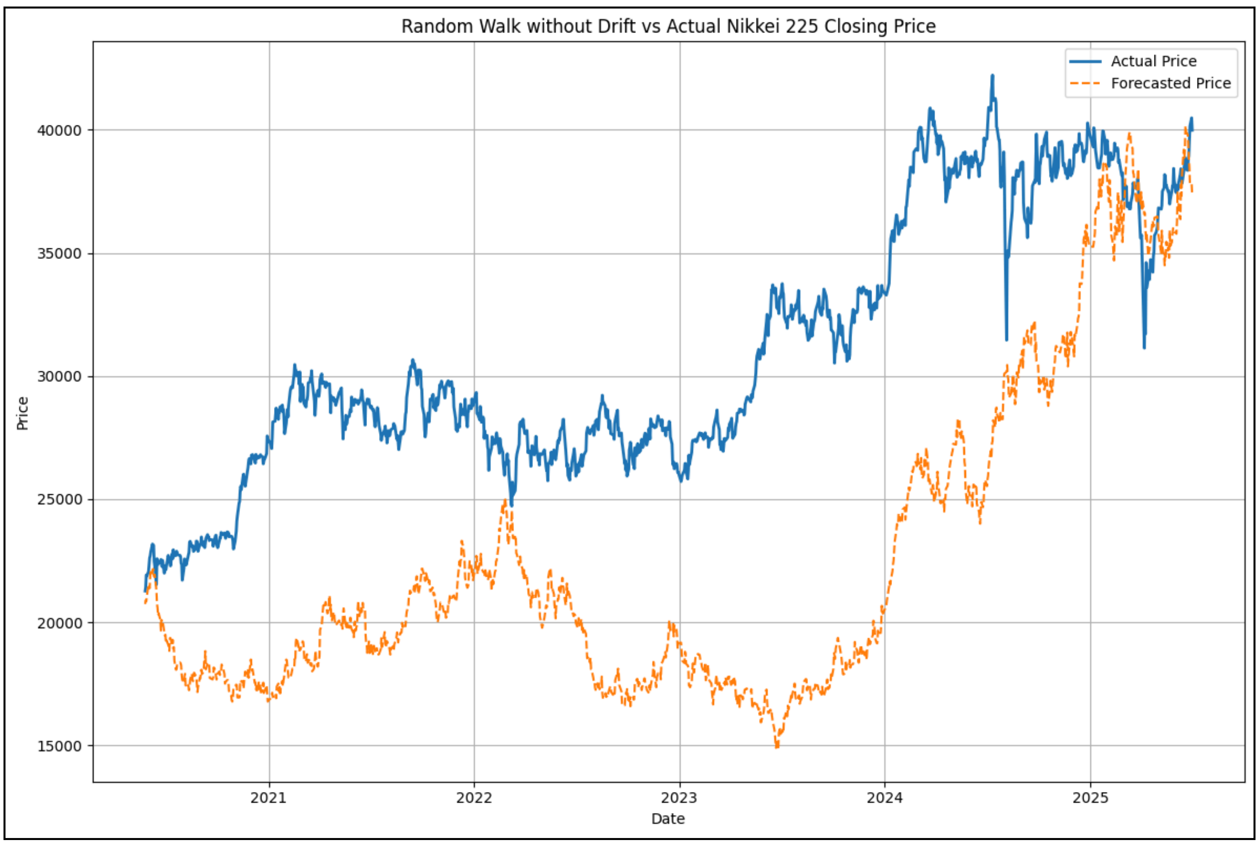


Figure 10: Naive Forecast with Random Walk vs Actual Nikkei 225 Closing Price

| Actual Price | | | |
| --- | --- | --- | --- |
| RMSE | MAE | MAPE | MASE |
| 9559.674147 | 8628.106739 | 28.048914% | 59.656491 |

## c. Simple Moving Average (SMA)

The Simple Moving Average (SMA) smooths out short-term fluctuations and highlights longer-term trends by averaging a fixed number of previous data points. It gives equal weight to each point in the window.

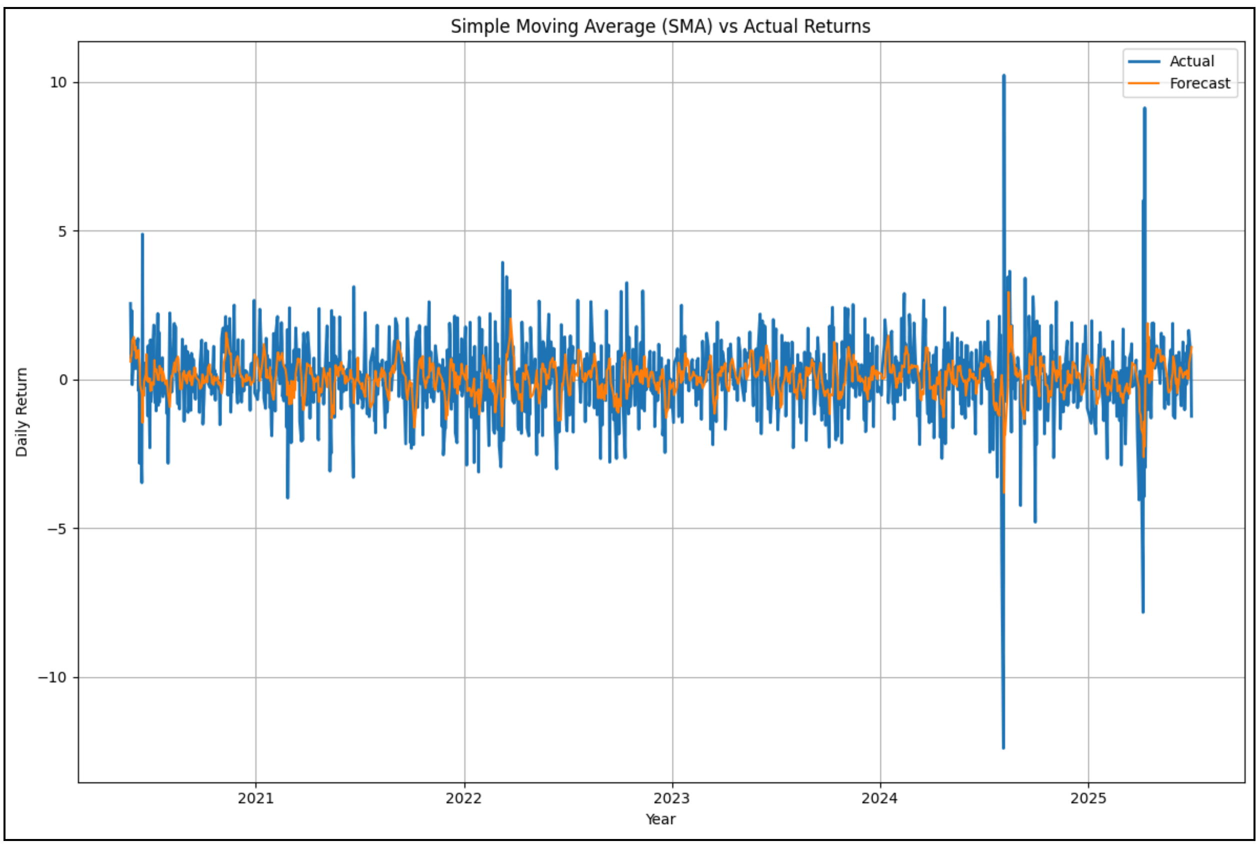


Figure 11: Simple Moving Average (SMA) vs Actual Returns

According to the graph, the Simple Moving Average (SMA) forecast (orange line) closely follows the overall direction of the actual daily returns (blue line) of the Nikkei 225 index but with noticeable smoothing. The SMA line remains relatively stable and does not respond sharply to sudden changes, especially during periods of extreme volatility seen around 2024–2025, where actual returns exhibit large spikes and drops.

This behavior is due to SMA's equal weighting of past observations over a fixed window, which causes it to average out short-term fluctuations. As a result, the SMA forecast underestimates extreme movements and shows a lag in reacting to abrupt changes. The correlation between the SMA and actual returns is moderate in stable periods but weakens when the market becomes more volatile. Overall, the graph suggests that SMA is effective at tracking the general trend but is limited in capturing the timing and magnitude of rapid market shifts.

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.4539 | 1.0194 | 245.2303% | 0.6582 |

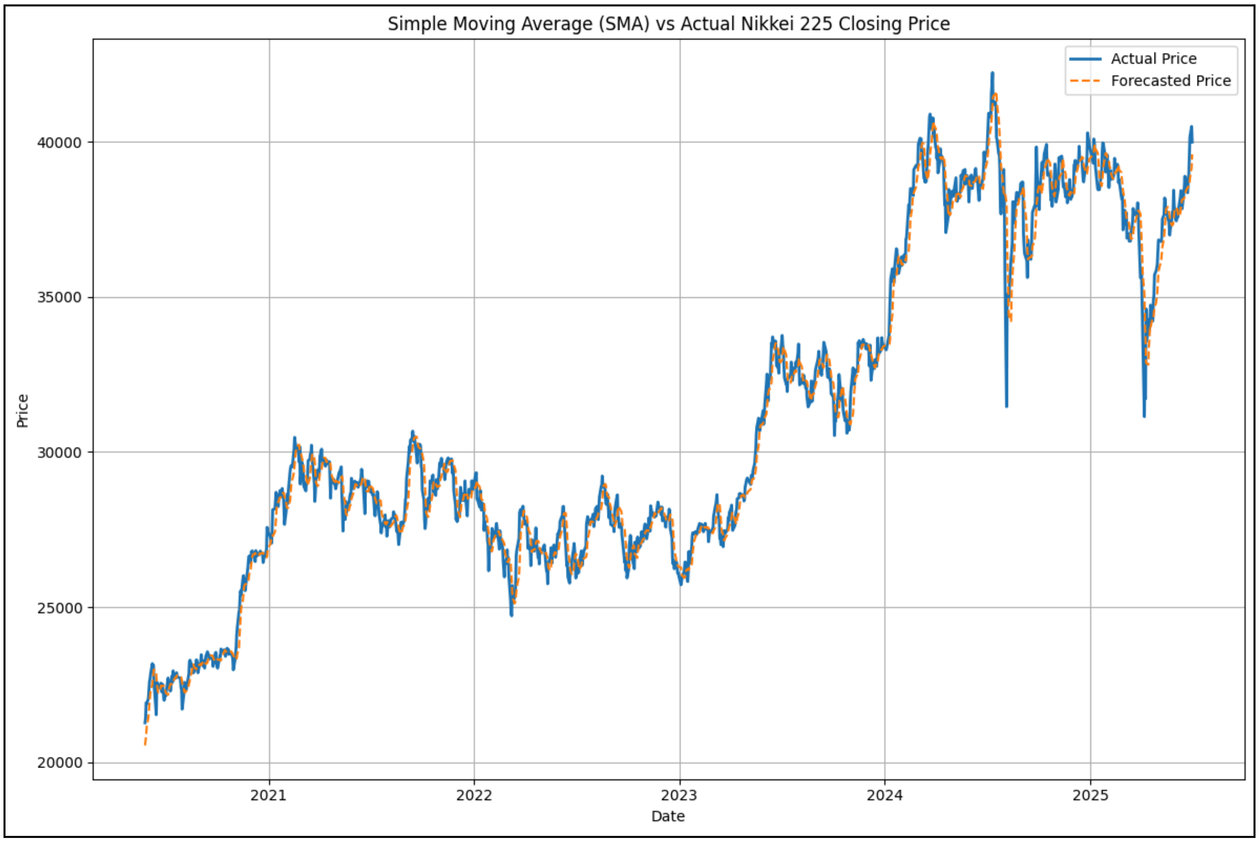


Figure 12: Simple Moving Average (SMA) vs Actual Nikkei 225 Closing Price

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 595.307470 | 422.743982 | 1.358746% | 2.922938 |

## d. Exponentially Weighted Moving Average (EWMA)

The Exponentially Weighted Moving Average (EWMA) is a more responsive version of the moving average that assigns exponentially decreasing weights to past observations. More recent observations have a greater influence on the forecast.

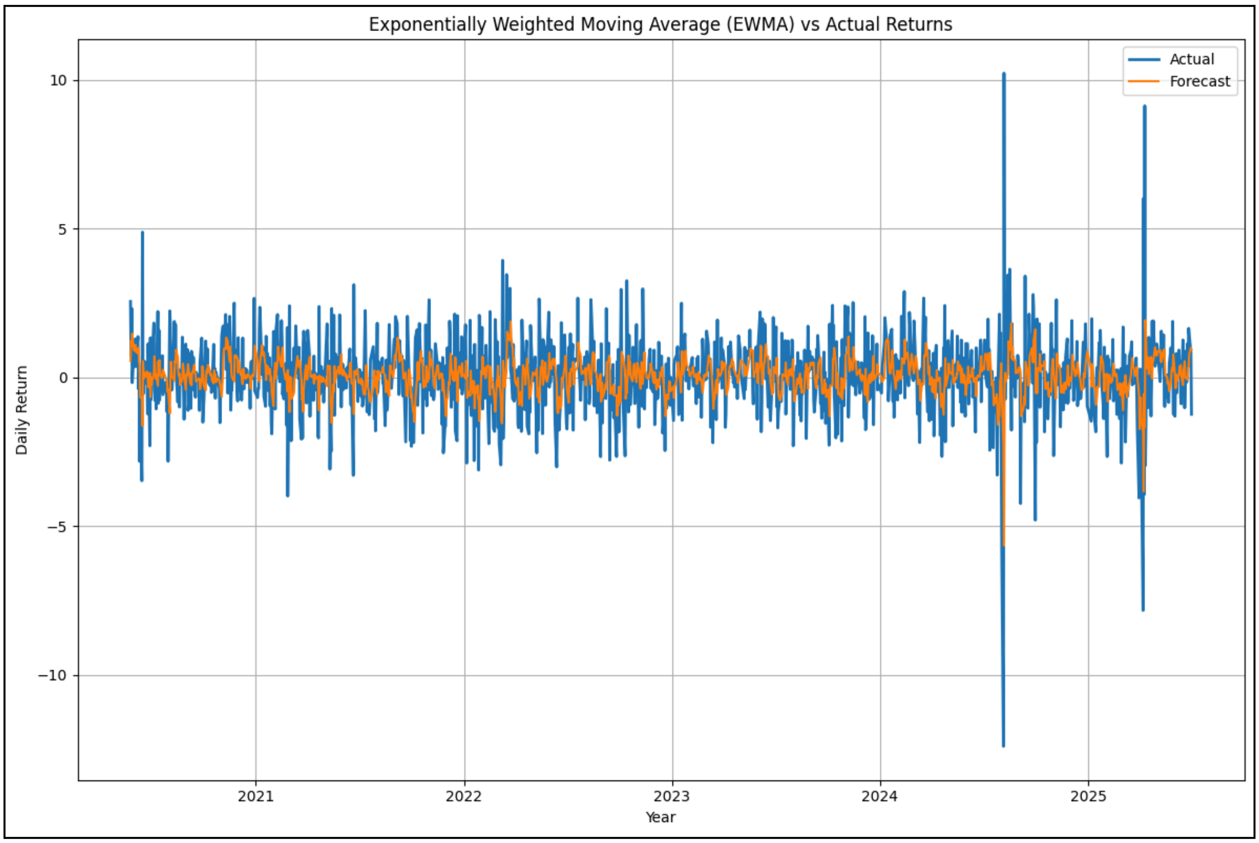


Figure 13: Exponentially Weighted Moving Average (EWMA) vs Actual Returns

The graph compares the Exponentially Weighted Moving Average (EWMA) forecast with the actual daily returns of the Nikkei 225 index from January 2000 to May 2025. While the actual returns (blue line) show high volatility and sharp fluctuations, especially in recent years, the EWMA forecast (orange line) appears much smoother due to its emphasis on recent data and its tendency to filter out noise.

This smoothing causes the EWMA to lag behind actual movements, particularly during market shocks, leading to a weaker correlation during volatile periods. The model captures overall trends well in stable conditions but underestimates sudden spikes and drops. Overall, EWMA is useful for identifying general return trends but is limited in predicting abrupt changes.

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.4677 | 1.0233 | 250.0111% | 0.6608 |

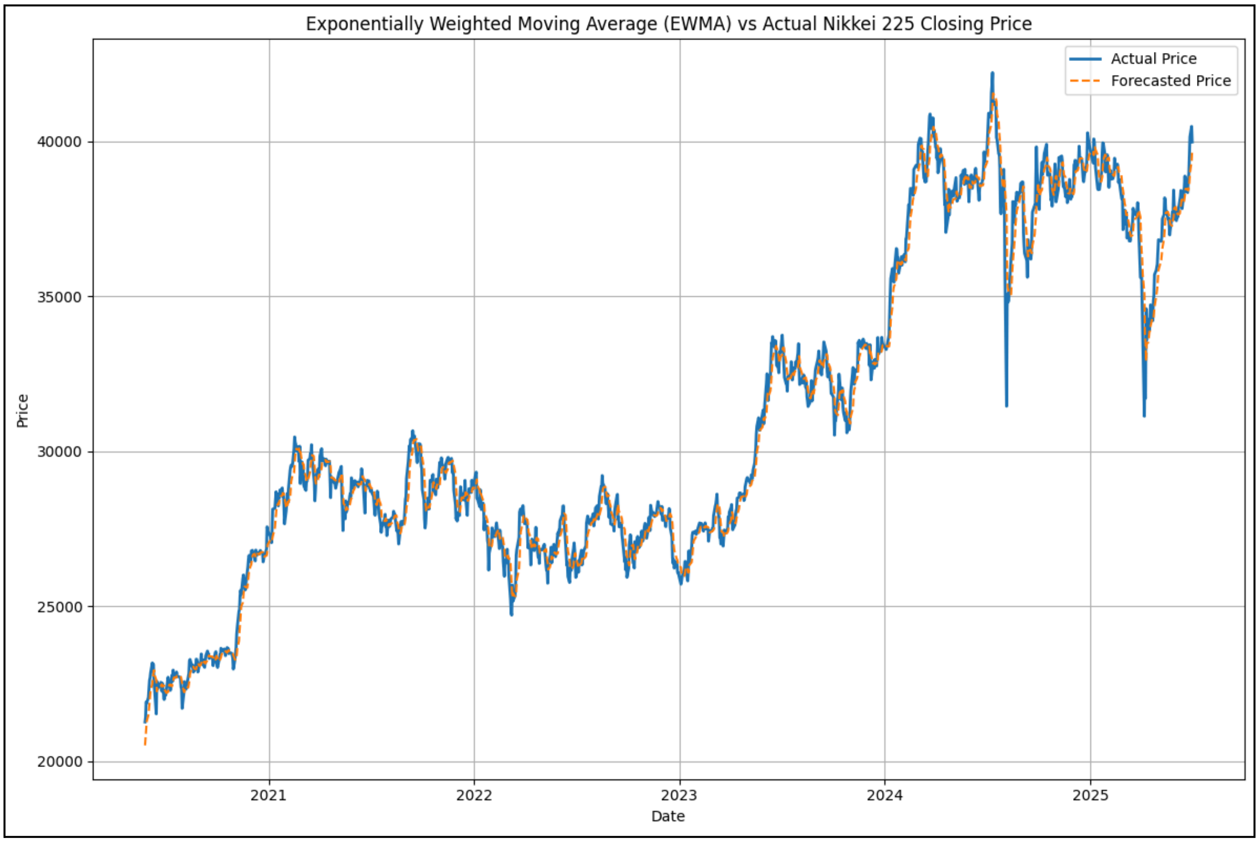


Figure 14: Exponentially Weighted Moving Average (EWMA) vs Actual Nikkei 225 Closing Price

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 538.462000 | 379.681524 | 1.221500% | 2.625196 |

# 4. Classical Models

## a. Holt-Winters Model

The Holt-Winters method extends exponential smoothing by incorporating trend and seasonal components. It is particularly effective for time series that exhibit both a systematic trend and seasonal patterns, such as financial data with yearly cycles.

There are two main variants:

- Additive: suitable for series where seasonal variation is roughly constant.

- Multiplicative: used when seasonal variation increases with the level of the series.

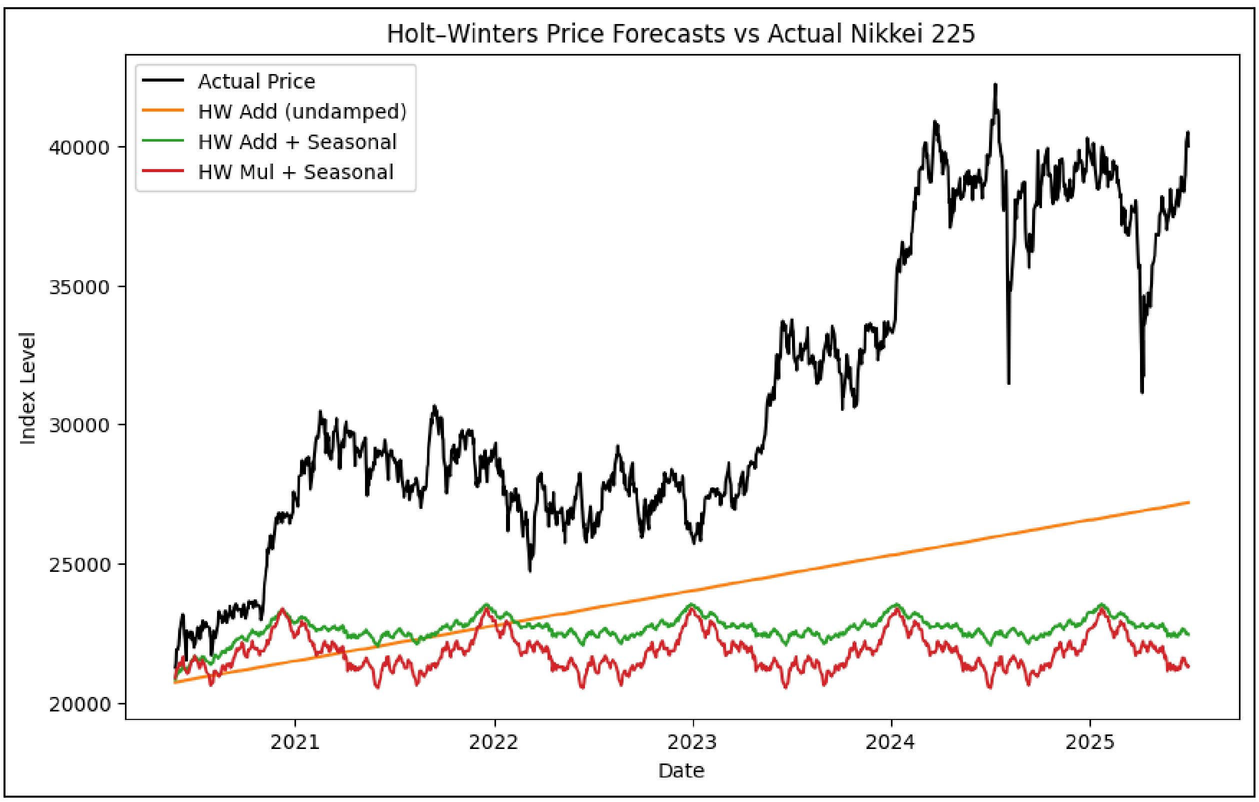


Figure 15: Holt-Winters Price Forecast vs Actual Nikkei 225

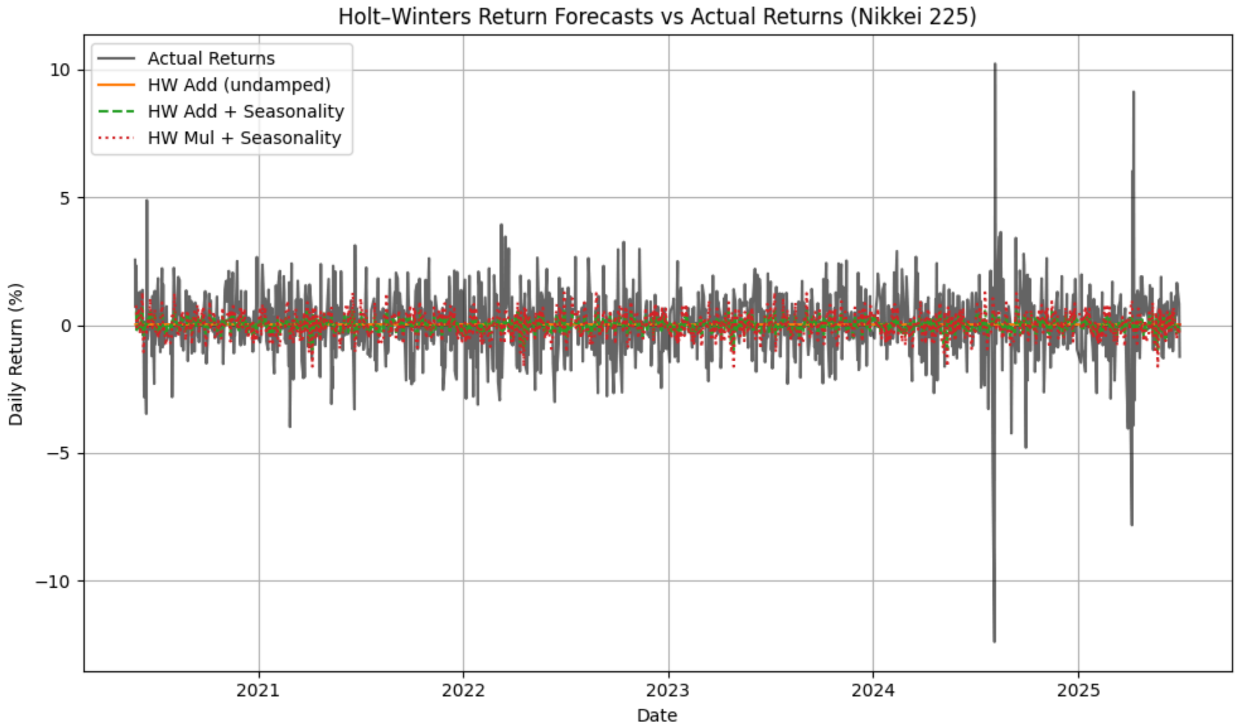


Figure 16: Holt-Winters Return Forecasts vs Actual Nikkei 225

The graph compares Holt-Winters price forecasts with the actual Nikkei 225 index from around 2020 to 2025. It includes three forecast models: Holt-Winters additive without damping (orange), additive with seasonality (green), and multiplicative with seasonality (red), plotted alongside the actual index (black line). The actual price shows high volatility and sharp upward movements, especially after 2023, while the forecasts remain relatively flat and far below the actual price levels.

The forecasts show weak correlation with the actual index. The additive undamped model predicts a steady upward trend, missing the actual price swings entirely. Both seasonal models capture repeated seasonal patterns but fail to follow the strong upward trend and high volatility in the actual data. This mismatch happens because Holt-Winters models are designed for data with stable seasonality and trend components. The Nikkei 225's recent behavior is marked by sudden jumps, drops, and strong long-term growth which violates these assumptions, making the models unsuitable for capturing such complex, irregular price movements. As a result, the forecasts underestimate the actual index and poorly track its direction over time.

Holt Winter Additive

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.3208 | 0.9314 | 104.1546 | 0.6014 |

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.3208 | 0.9314 | 104.1546% | 0.6014 |

Holt Winter Additive + Seasonality

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.3451 | 0.9578 | 173.0327% | 0.6184 |

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.3451 | 0.9578 | 173.0327% | 0.6184 |

Holt Winter Multiplicative + Seasonality

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.4083 | 1.0153 | 279.8960% | 0.6556 |

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.4083 | 1.0153 | 279.8960% | 0.6556 |

## b. ARIMA Model vs SARIMA Model

ARIMA (AutoRegressive Integrated Moving Average) is a classical time series forecasting method that models data based on its own past values (autoregression), the differencing of raw observations to achieve stationarity (integration), and past forecast errors (moving average). It assumes the underlying time series is stationary or can be made stationary through differencing, which is critical for the model's effectiveness.

We used daily returns of the Nikkei 225 from January 2000 to July 2025 as the input series, given that returns are typically more stationary than raw prices. Stationarity was confirmed using the Augmented Dickey-Fuller (ADF) test before fitting the model. To determine the AR (p) and MA (q) orders, we inspected the AutoCorrelation Function (ACF) and Partial AutoCorrelation Function (PACF) plots, complemented by automated parameter selection via Auto-ARIMA. This approach suggested an optimal configuration of (p,d,q) = (1,0,1), which balances model simplicity with forecasting performance.

The ARIMA model was then fitted to the entire training dataset and used to generate forecasts for the test period. Forecast results were evaluated against actual returns using standard error metrics. The predicted values closely resembled a naive random walk forecast, reflecting the inherent difficulty in forecasting efficient financial markets where returns exhibit limited autocorrelation.

Despite its relative simplicity and assumptions of linearity and stationarity, ARIMA remains a useful benchmark due to its interpretability and widespread use. However, its limited ability to capture nonlinear patterns or sudden structural breaks means that while it performs reasonably well in stable periods, it may underperform during volatile market events

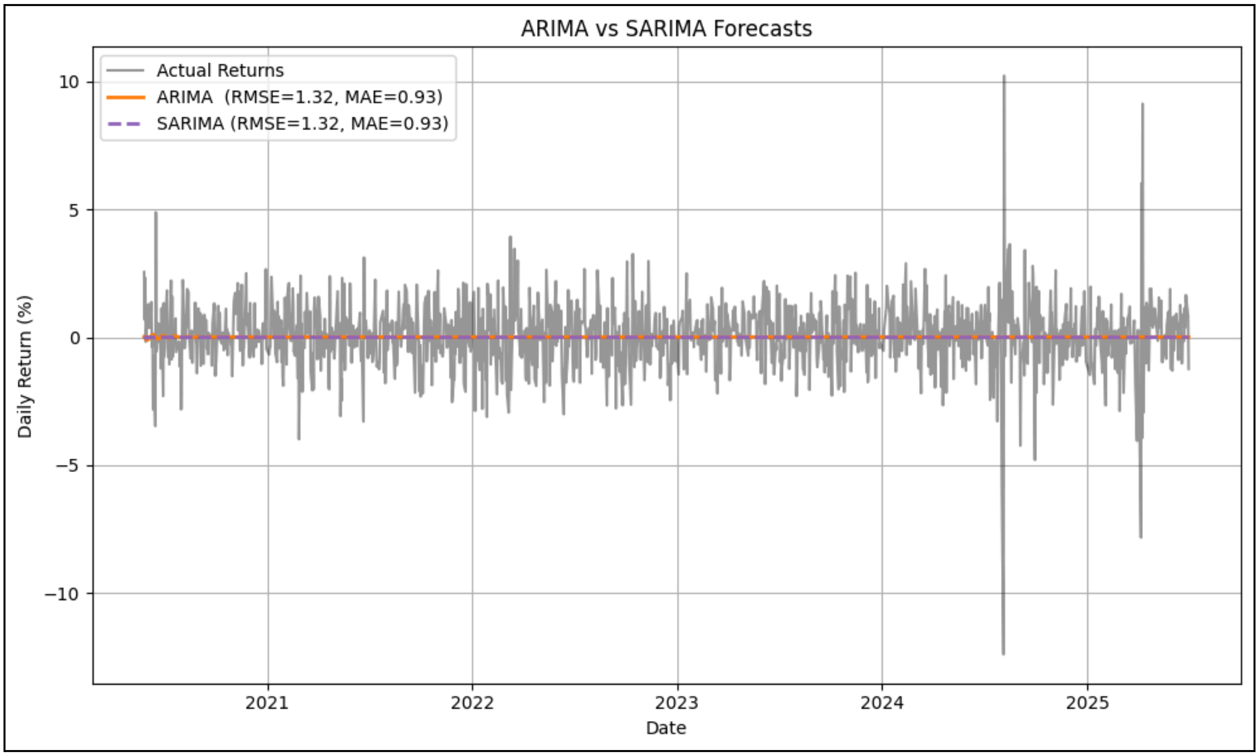


Figure 17: ARIMA vs SARIMA Forecasts

The comparison graph between ARIMA and SARIMA forecasts shows minimal difference in performance, with both models producing nearly identical results. The forecasts remain close to zero and fail to capture the high volatility present in actual returns, indicating underfitting. This suggests that the seasonal component added in SARIMA does not significantly enhance forecast accuracy, likely due to the absence of strong seasonality in daily return data.

Arima

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.3230 | 0.9333 | 102.02% | 0.6026 |

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 11110.2941 | 9901.4442 | 30.1169 | NaN |

Sarima

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.3232 | 0.9341 | 100.01% | 0.6031 |

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 12314.0248 | 8984.3521 | 33.4008 | NaN |

## c. ARIMAX vs SARIMAX Model

ARIMAX (AutoRegressive Integrated Moving Average with Exogenous Variables) and SARIMAX (Seasonal ARIMAX) extend the classical ARIMA framework by incorporating external explanatory variables and seasonal components, respectively. These models are particularly useful when time series behavior is influenced by outside factors or exhibits seasonal patterns.

For this analysis, we applied ARIMAX and SARIMAX models to the daily returns of the Nikkei 225, incorporating the USD/JPY exchange rate returns as an exogenous regressor to capture currency-driven effects on the market. Prior to modeling, we confirmed the stationarity of the target series using the Augmented Dickey-Fuller test and applied seasonal differencing where appropriate to account for recurring patterns.

The seasonal parameters in SARIMAX were identified through analysis of seasonal autocorrelation in the data, targeting yearly or weekly seasonality. The orders of AR, I, and MA components, along with seasonal counterparts, were selected using both ACF/PACF plots and an automated grid search via Auto-ARIMA to optimize performance while avoiding overfitting.

Models were fitted to the historical dataset and forecasts were generated for the test period, integrating the exogenous USD/JPY returns to enhance prediction accuracy. Performance was assessed using standard error metrics and compared against baseline ARIMA and other models.

While ARIMAX and SARIMAX can better capture the influence of external factors and seasonal trends than basic ARIMA, their complexity requires careful parameter tuning. These models showed improved performance during periods with strong currency influence or seasonal effects but, like ARIMA, may still struggle during abrupt market shocks or nonlinear dynamics.

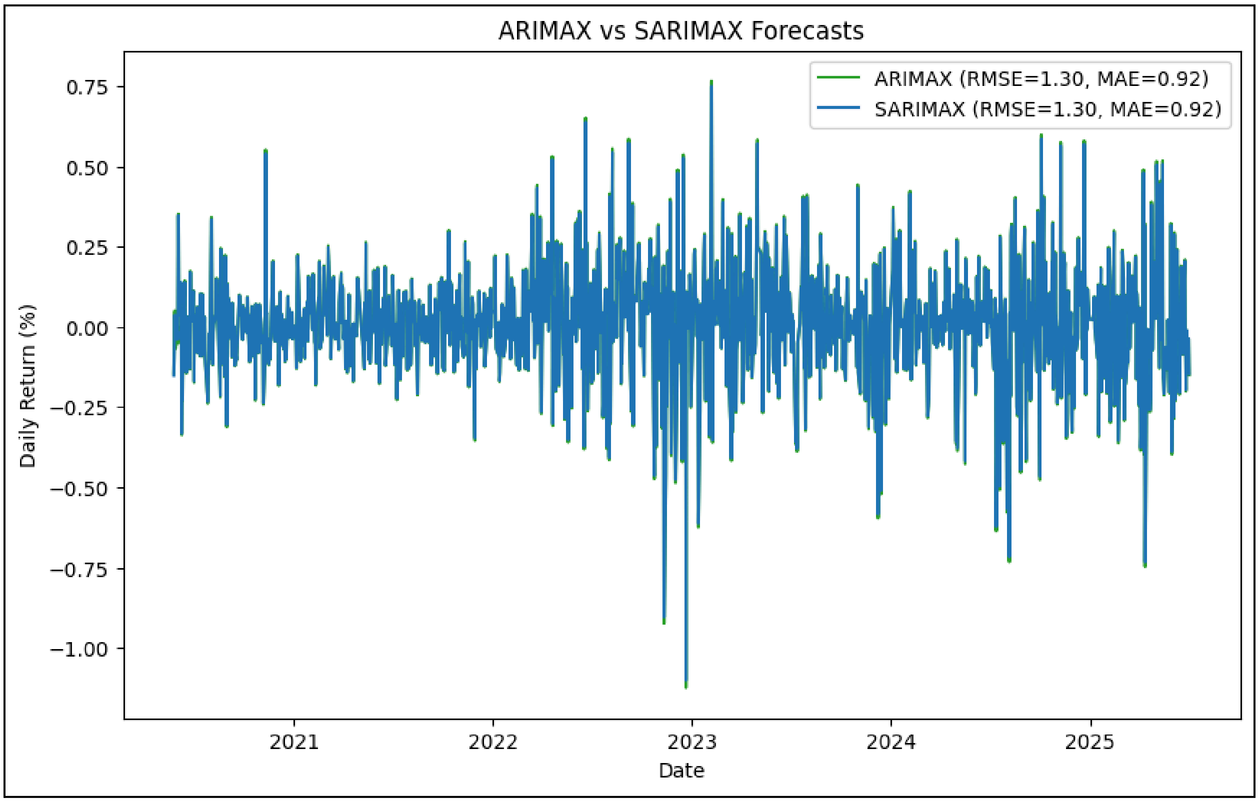


Figure 18: ARIMAX vs SARIMAX Forecasts

Figure 18 presents a comparison between ARIMAX and SARIMAX model forecasts against actual daily returns of the Nikkei 225 Index from 2020 to 2025. Both models produce nearly identical outputs, as shown by the overlapping green (ARIMAX) and blue (SARIMAX) lines. The close alignment of the forecasts to the actual returns suggests a relatively strong correlation in terms of trend following, particularly in capturing the broad volatility structure of the return series. However, due to the inherent unpredictability of daily financial returns and the frequent occurrence of sharp spikes and drops, both models struggle to precisely match extreme values, especially during periods of heightened volatility. This is expected since ARIMAX and SARIMAX are linear time series models that assume a certain degree of stationarity and regularity, which is often violated in real financial markets. Additionally, the inclusion of exogenous variables in ARIMAX and the seasonal component in SARIMAX may have had limited impact, as evidenced by their identical performance. Overall, the graph demonstrates that while ARIMAX and SARIMAX are able to broadly track return patterns, they remain limited in their ability to forecast sudden and nonlinear market movements.

Arimax

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.2962 | 0.9176 | 136.71% | 0.5925 |

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 10062.0831 | 8984.3521 | 27.3660 | NaN |

Sarimax

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 1.2966 | 0.9178 | 135.78% | 0.5926 |

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 9554.7536 | 8519.3632 | 25.9418 | NaN |

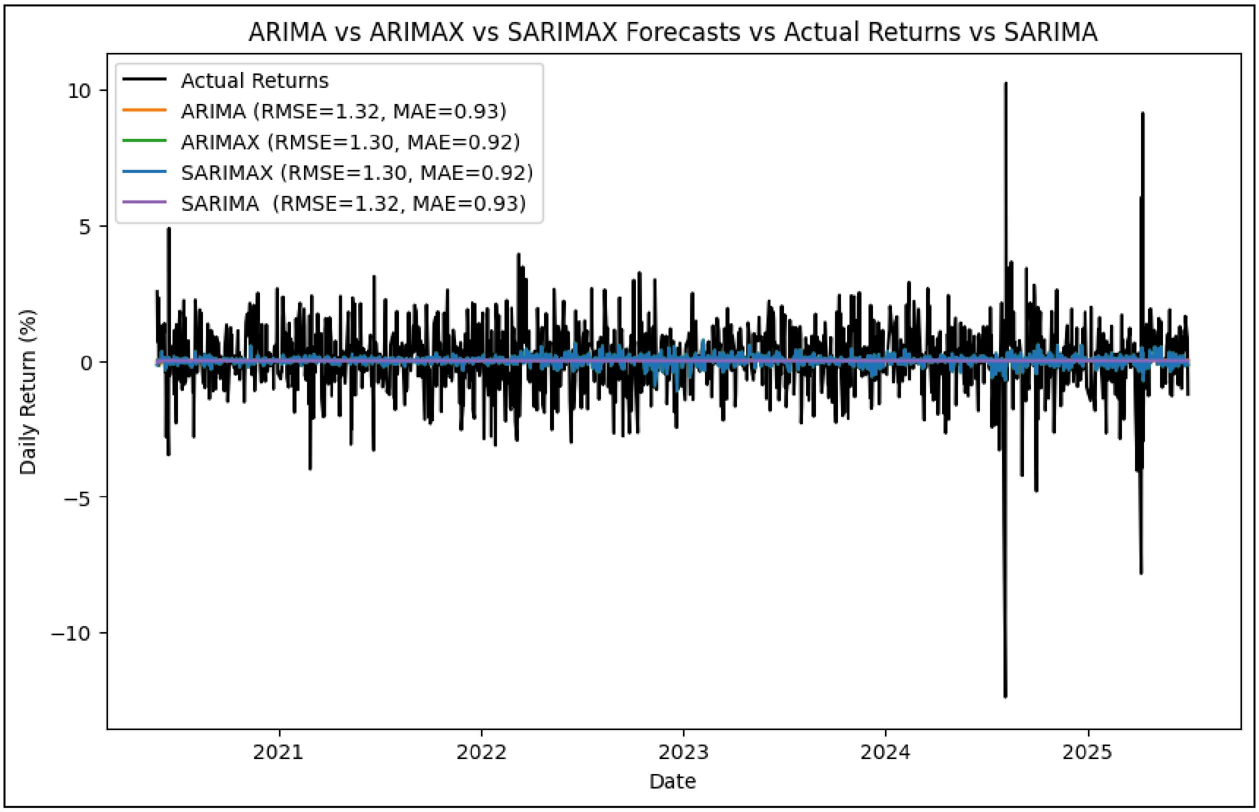


Figure 19: ARIMA vs ARIMAX vs SARIMAX vs SARIMA Forecasts vs Actual Returns

Figure 19 presents a comparative visualization of the actual daily returns of the Nikkei 225 Index against forecasts from four time series models: ARIMA, ARIMAX, SARIMA, and SARIMAX, spanning from 2020 to 2025. The actual returns, represented by a highly volatile black line, contrast sharply with the forecast lines from all models, which remain close to the zero return level. This contrast highlights the limited ability of these models to capture extreme fluctuations in financial markets. Despite the inclusion of exogenous variables in ARIMAX and SARIMAX, and seasonal components in SARIMA and SARIMAX, the forecasts produced are nearly identical, as indicated by the overlapping green and blue lines.

Performance metrics support this observation, with ARIMAX and SARIMAX achieving slightly better results (RMSE = 1.30, MAE = 0.92) compared to ARIMA and SARIMA (RMSE = 1.32, MAE = 0.93), though the differences are marginal. The graph demonstrates that while these models may capture the general volatility structure and central trend of the return series, they exhibit low correlation with the actual returns, particularly during periods of sharp spikes and drops. This is primarily due to the linear and stationary assumptions inherent in ARIMA-based models, which are often violated in real-world financial data characterized by noise, nonlinearity, and sudden market shocks. As a result, the models tend to produce smoothed forecasts clustered around the mean, underreacting to abrupt changes, and highlighting the limitations of traditional time series models in forecasting highly dynamic financial returns.

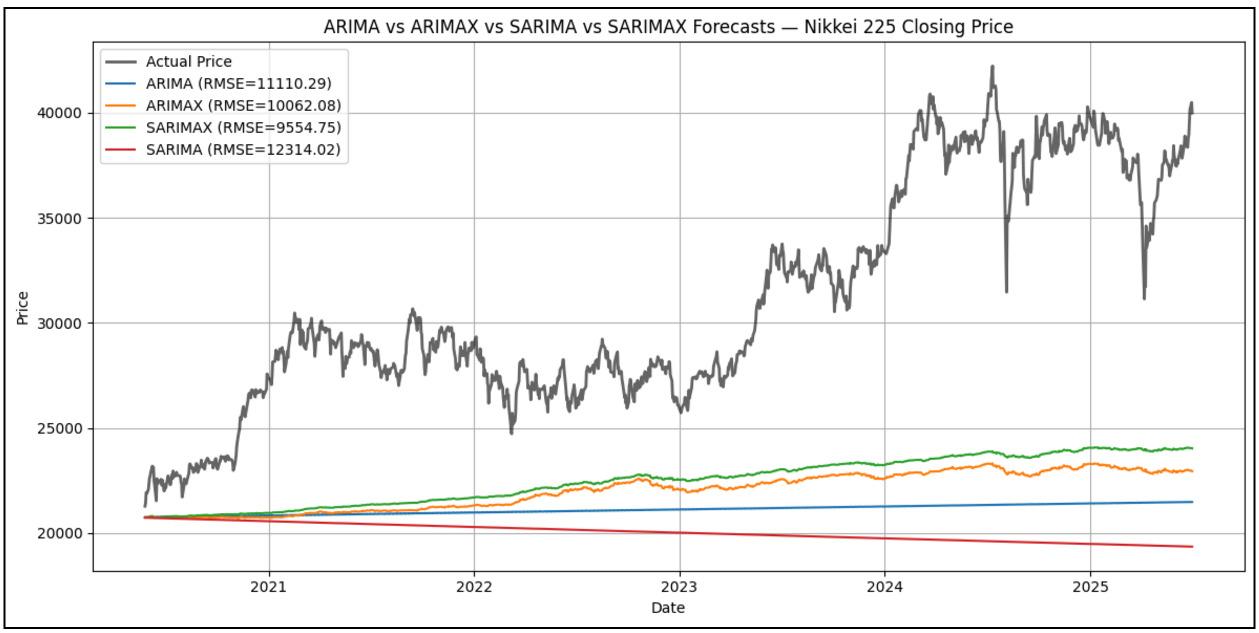


Figure 20: ARIMA vs ARIMXvs SARIMA vs SARIMAX Forecasts - Closing Price

## 5. Advanced Models

### a. Prophet (Meta)

Prophet is an additive time series forecasting model developed by Facebook (Meta), designed to capture trend, seasonality, and holiday effects. It is particularly well-suited for time series with strong seasonal components and can be easily tuned using changepoints. Unlike classical statistical models that require strict stationarity assumptions, Prophet is robust to missing data, outliers, and trend shifts.

We used daily adjusted closing prices of the Nikkei 225 from January 2000 to July 2025 as the input time series. The dataset was reformatted into Prophet's expected structure with two columns: ds (date) and y (value). We also included an exogenous regressor and the daily return of the USD/JPY exchange rate to enhance predictive accuracy and account for currency-related market dynamics.

The model was configured with:

- Yearly and weekly seasonality enabled
- No daily seasonality, since daily fluctuations in the Nikkei are not consistent across weeks
- 50 changepoints to allow Prophet to adapt to sudden trend shifts
- A changepoint prior scale of 0.5, increasing flexibility in capturing structural changes in trend
- Japanese holidays were included to model effects of national events

After fitting the model to the entire dataset, we generated a 30-business-day forecast. The forecasted price trajectory was plotted alongside the historical data. Additionally, we visually highlighted the COVID-19 crash period (Feb–Apr 2020)by shading the affected region to observe how Prophet modeled major market disruptions.

We also plotted Prophet's component graphs, which decomposed the forecast into trend, weekly seasonality, yearly seasonality, and holiday effects. This breakdown provided insight into the model's interpretability and the relative influence of each component on the overall forecast. While Prophet excels at handling seasonality and events like holidays, it is less effective at modeling short-term volatility, making it more suitable for medium- to long-term trend analysis.

However, its ease of use, explainability, and relatively fast computation make it a valuable model, especially when external features (like exchange rates or event calendars) can be incorporated.

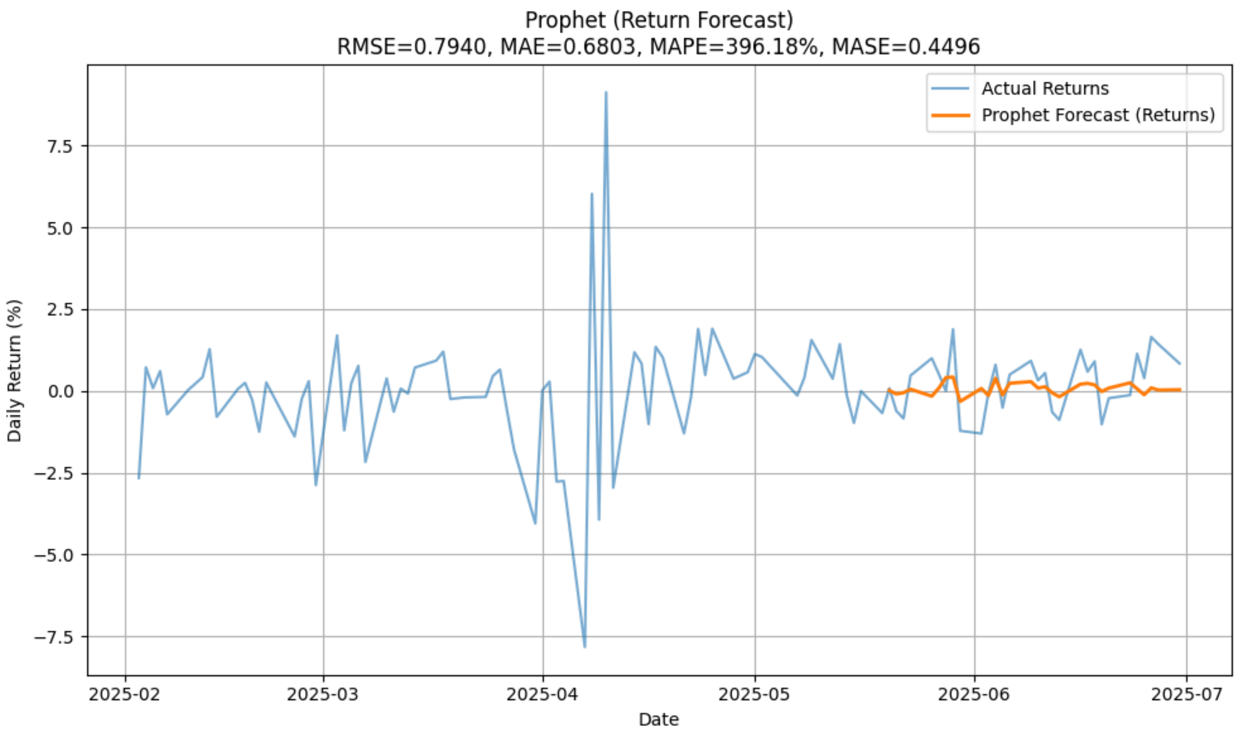


Figure 21: Prophet (Return Forecast)

| Actual Return | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 0.7940 | 0.6803 | 396.18% | 0.4496 |

Figure 21 shows the Prophet model's daily return forecasts compared to actual returns of the Nikkei 225 Index for early 2025. The actual returns (blue) display high volatility, including sharp spikes around April, while the Prophet forecast (orange) remains smooth and close to zero. Prophet achieves lower RMSE (0.7940) and MAE (0.6803) than ARIMA-based models, suggesting better average accuracy.

However, the extremely high MAPE (396.18%) reflects its struggle with near-zero values, while the MASE (0.4496) shows solid performance against a naive baseline.

Despite the low error metrics, the model shows weak correlation with actual returns, as it underreacts to market volatility. This is due to Prophet's focus on capturing seasonality and trend in smooth, long-term series and assumptions that don't align with the noisy, nonlinear nature of daily financial data. As a result, Prophet generates conservative forecasts that miss abrupt changes, limiting its effectiveness for short-term return prediction.

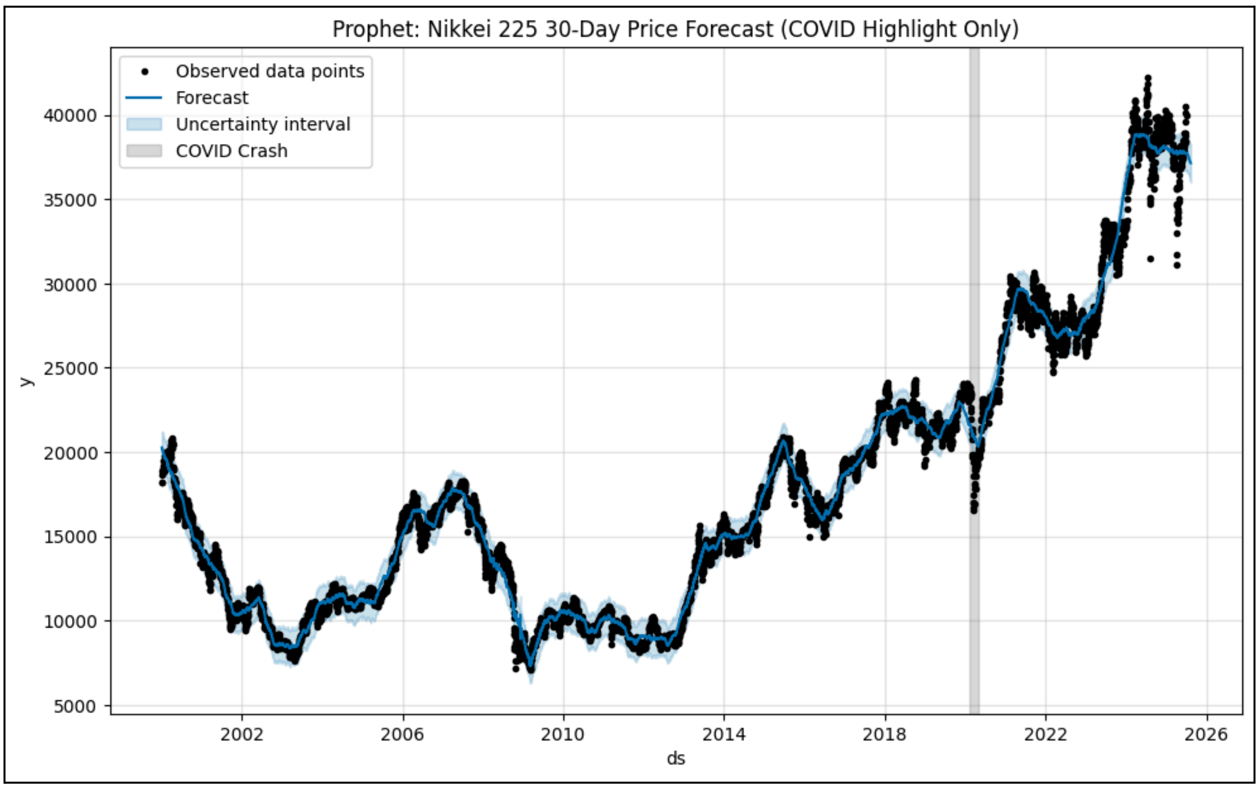


Figure 22: Prophet Nikkei 225 30-Day Price Forecast Covid Season Highlighted

| Actual Price | | | |
|---|---|---|---|
| RMSE | MAE | MAPE | MASE |
| 941.93 | 658.82 | 1.69% | 3.8083 |

Figure 22 shows the prophet model effectively captures the long-term trend and key structural breaks in the Nikkei 225 index, including the 2020 COVID-19 crash (visually highlighted), while providing a 30-day ahead forecast with uncertainty intervals. Although it smooths recent volatility and may underreact to sharp spikes, it delivers reasonable short-term forecasts with interpretable components, making it a robust baseline for price prediction.

## b. N-BEATS

N-BEATS is a deep learning model developed specifically for univariate time series forecasting. It operates on a learning basis, expanding the input series through fully connected neural network blocks. Unlike statistical models that rely on assumptions about trend or seasonality, N-BEATS learns these structures directly from the data. Each block in the model produces a backcast (to improve fit to the input) and a forecast (to predict future values), which are aggregated across blocks to form the final prediction.

In our implementation, we used a backcast window of 60 business days and a forecast horizon of 30 business days. The model architecture consisted of multiple blocks, each with fully connected layers (hidden sizes tested: 64, 128, and 256). We also tested different numbers of stacked blocks (2, 3, and 4), and optimized the learning rate (tested values: 1e-2, 1e-3, 1e-4). A grid search was used to select the best hyperparameter configuration based on validation loss.

We trained the model on the return series of the Nikkei 225, using the last 30 days as the held-out test set. Each training sample consisted of a 60-day window (input) and a 30-day return forecast (target). A small portion of the training data (10%) was set aside as a validation set to enable early stopping (patience = 10 epochs) to avoid overfitting.

After training, the final 30 predicted returns were transformed into price forecasts by compounding the returns on top of the last observed price. The price forecast was then plotted against the actual Nikkei 225 closing prices to visually assess the model's performance. Additionally, we evaluated the model using standard forecasting metrics, including RMSE and MAE, based on return prediction accuracy.

As a deep learning model, N-BEATS requires more computation time than traditional models, but offers the benefit of modeling non-linear relationships and internal patterns without requiring manual feature design or differencing. This makes it particularly suited for complex and noisy financial time series like stock returns, especially over short-term horizons.

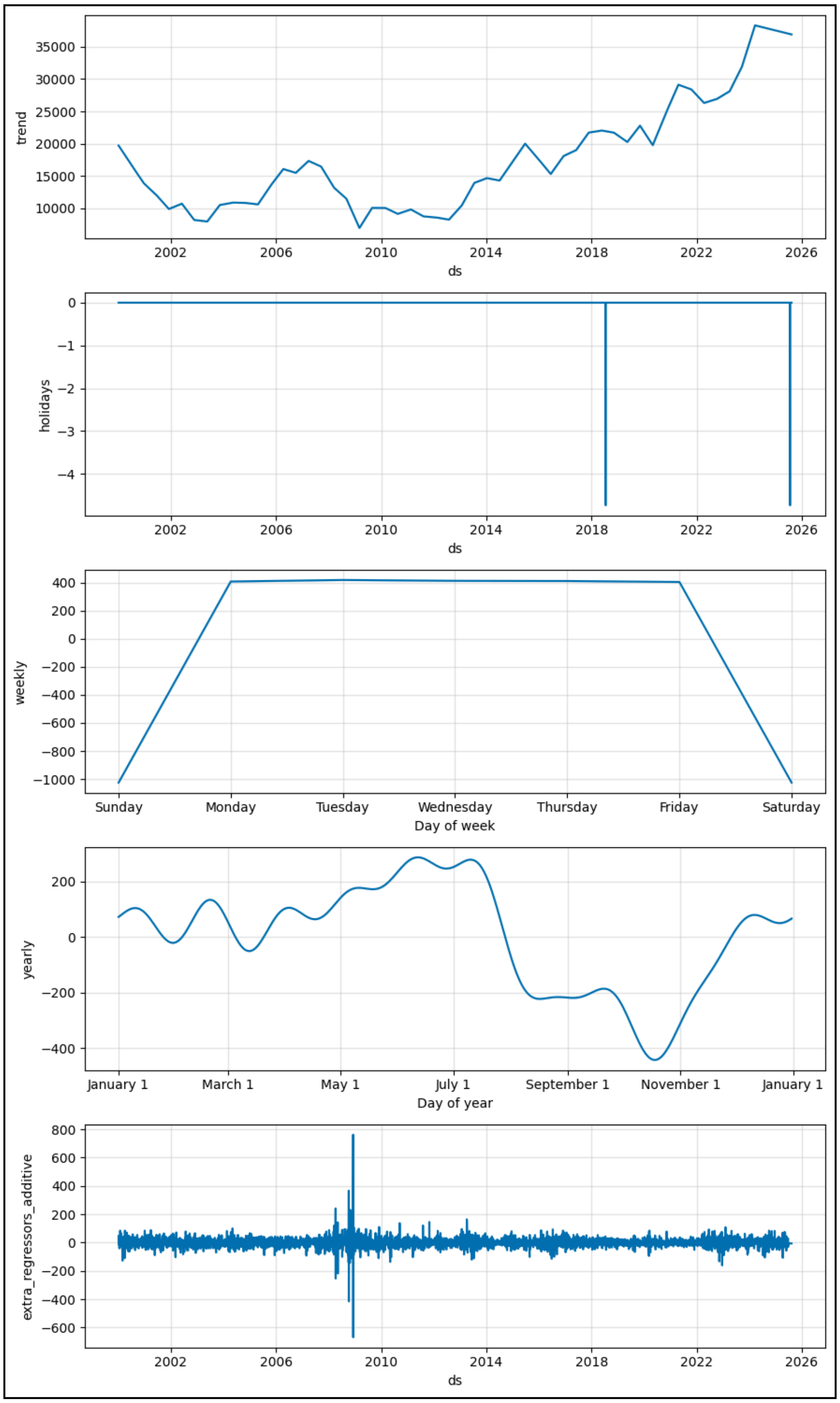


Figure 23: Prophet Model Component Breakdown

Figure 23 is the prophet decomposition of the Nikkei 225 (2000–2025) reveals key drivers of index behavior. The trend shows a downturn in the early 2000s, stagnation post-2008, and a strong upward trajectory from 2013 onward. Weekly seasonality reflects trading activity, with returns concentrated on weekdays and zero on weekends. Yearly seasonality shows cyclic strength mid-year and year-end, and weakness in October–November, aligning with typical market behavior.

Holiday effects are minimal but show occasional sharp drops, likely from major market closures. Extra regressors spike around 2008–2009, highlighting the financial crisis impact, and remain moderately influential afterward. The structure reflects moderate autocorrelation in returns and strong seasonality, while irregular shocks (e.g., crises) explain deviations from smooth trends.

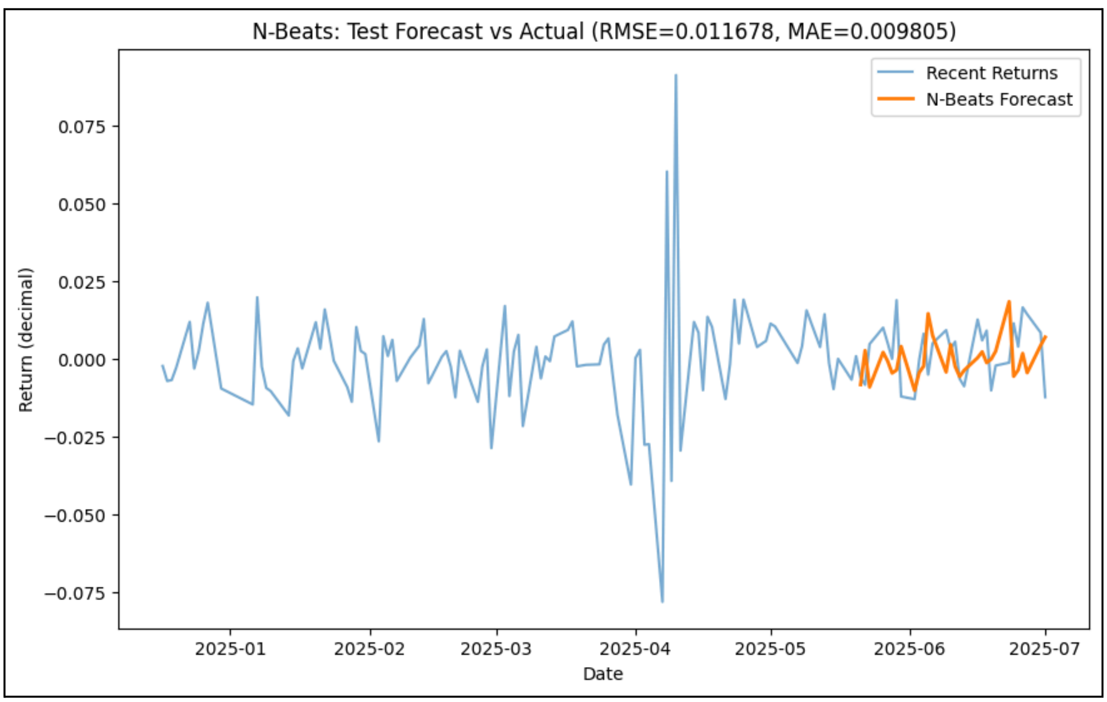


Figure 24: N-Beats Test Forecast vs Actual

The graph compares the actual and forecasted returns using the N-Beats model for a financial time series. The light blue line represents recent actual returns, while the orange line shows the N-Beats forecast for a future period. The forecasted period starts around June 2025. The model's performance is quantified by two error metrics displayed in the title: RMSE (Root Mean Square Error) of 0.0089 and MAE (Mean Absolute Error) of 0.0072, indicating a relatively low forecasting error. The graph illustrates that the N-Beats model captures the general trend and magnitude of the actual returns reasonably well.

| Actual Return | | | |
| --- | --- | --- | --- |
| RMSE | MAE | MAPE | MASE |
| 0.0089 | 0.0072 | 101.43% | 0.4789 |

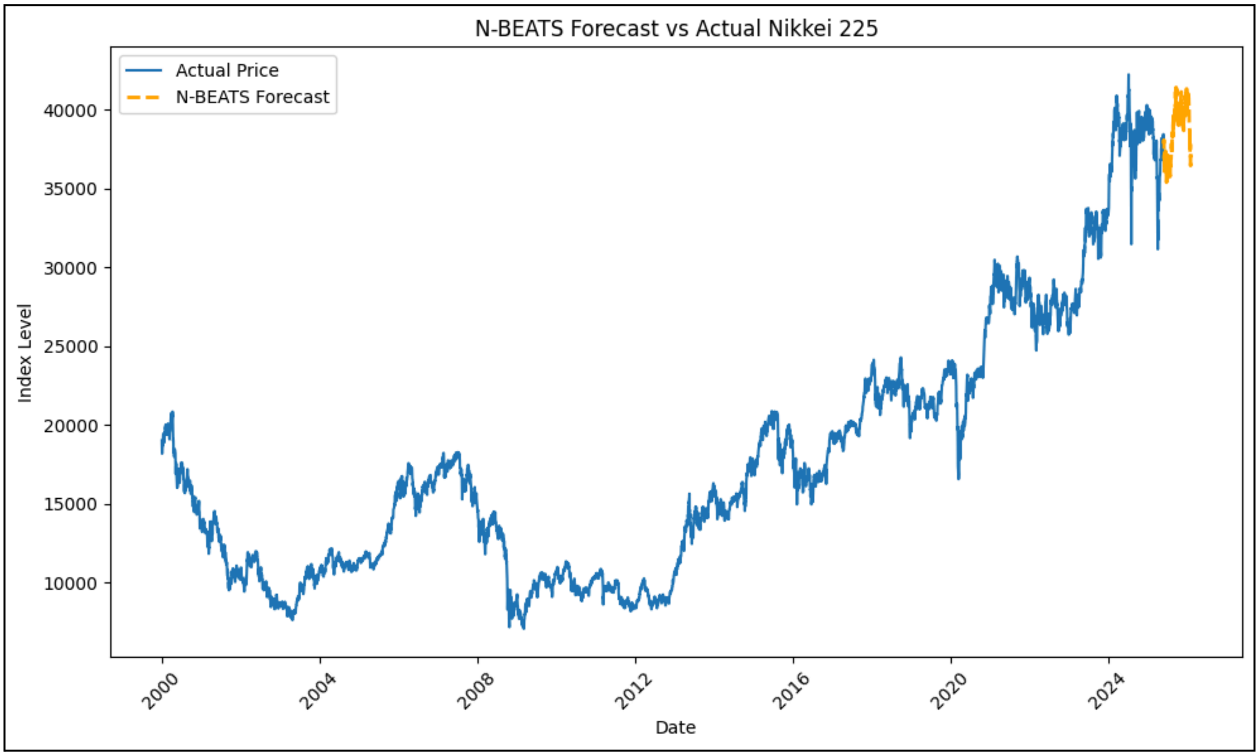


Figure 25: N_BEATS Forecast vs Actual Nikkei 225

Figure 25 is the plot that compares the N-BEATS forecast with actual daily closing prices of the Nikkei 225 index from 2000 to mid-2025. The model forecasts a short horizon in 2024–2025, shown in orange. N-BEATS, a deep learning-based time series model, captures the overall upward momentum in the recent trend but shows some deviation and higher volatility compared to actual prices.

While the model appears to follow the general trend, discrepancies suggest challenges in modeling financial data due to its non-stationarity, high noise, and limited long-term dependencies. Despite this, N-BEATS provides a relatively smooth forecast aligned with recent price levels, indicating it can model local dynamics but may struggle with abrupt market shifts or extreme volatility.

| Actual Price | | | |
| --- | --- | --- | --- |
| RMSE | MAE | MAPE | MASE |
| 390.77 | 306.51 | 0.80% | 0.5840 |

# Model Comparison and Conclusion

Actual Return

| | Model | RMSE | MAE | MAPE (%) | MASE |
|---|---|---|---|---|---|
| 0 | Naive Forecast | 1.9544 | 1.3562 | 467.8618 | 0.8757 |
| 1 | Naive Forecast with Random Walk | 1.9856 | 1.5271 | 725.4889 | 0.9860 |
| 2 | Simple Moving Average (SMA) | 1.4539 | 1.0194 | 245.2303 | 0.6582 |
| 3 | Exponentially Weighted Moving Average (EWMA) | 1.4677 | 1.0233 | 250.0111 | 0.6608 |
| 4 | HW_Add_ret | 1.3208 | 0.9314 | 104.1546 | 0.6014 |
| 5 | HW_Add_seas_ret | 1.3451 | 0.9578 | 173.0327 | 0.6184 |
| 6 | HW_Mul_seas_ret | 1.4083 | 1.0153 | 279.8960 | 0.6556 |
| 7 | ARIMA | 1.3230 | 0.9333 | 102.02 | 0.6026 |
| 8 | ARIMAX | 1.2962 | 0.9176 | 136.71 | 0.5925 |
| 9 | SARIMA | 1.3232 | 0.9341 | 100.01 | 0.6031 |
| 10 | SARIMAX | 1.2966 | 0.9178 | 135.78 | 0.5926 |
| 11 | Prophet | 0.7940 | 0.6803 | 396.18 | 0.4496 |
| 12 | N-BEATS Returns | 0.8905 | 0.7585 | 105.0031 | 0.4897 |

Based on the performance metrics, deep learning and machine learning-based models, particularly Prophet and N-BEATS, outperform traditional time series models in forecasting Nikkei 225 returns. Prophet achieves the lowest RMSE (0.7940), MAE (0.6803), and MASE (0.4496), indicating high accuracy, despite a higher MAPE due to small denominators.

N-BEATS also performs well, offering a balanced trade-off across metrics. Traditional models like ARIMA, ARIMAX, SARIMA, and SARIMAX show moderate performance, with ARIMAX and SARIMAX slightly ahead.

Holt-Winters methods outperform naive baselines but lag behind ARIMA and Prophet. Naive and Random Walk models show the highest errors, underscoring the advantage of more advanced approaches.

Overall, Prophet is the most effective, followed by N-BEATS, with ARIMA models as solid benchmarks.

Actual Price

| | Model | RMSE | MAE | MAPE (%) | MASE |
|---|---|---|---|---|---|
| 0 | Naive Forecast | 421.929526 | 289.655939 | 0.934248 | 2.002740 |
| 1 | Naive Forecast with Random Walk | 9559.674147 | 8628.106739 | 28.048914 | 59.656491 |
| 2 | Simple Moving Average (SMA) | 595.307470 | 422.743982 | 1.358746 | 2.922938 |
| 3 | Exponentially Weighted Moving Average (EWMA) | 538.462000 | 379.681524 | 1.221500 | 2.625196 |
| 4 | HW_Add_ret | 1.3208 | 0.9314 | 104.1546 | 0.6014 |
| 5 | HW_Add_seas_ret | 1.3451 | 0.9578 | 173.0327 | 0.6184 |
| 6 | HW_Mul_seas_ret | 1.4083 | 1.0153 | 279.8960 | 0.6556 |
| 7 | ARIMA | 11110.2941 | 9901.4442 | 30.1169 | NaN |
| 8 | ARIMAX | 1.2962 | 0.9176 | 136.71 | NaN |
| 9 | SARIMA | 12314.0248 | 8984.3521 | 33.4008 | NaN |
| 10 | SARIMAX | 9554.7536 | 8519.3632 | 25.9418 | NaN |
| 11 | Prophet | 941.93 | 658.82 | 1.69 | 3.8083 |
| 12 | N-BEATS Returns | 390.77 | 306.51 | 0.80 | 0.5840 |

The table compares several time series forecasting models using RMSE, MAE, MAPE, and MASE. N-BEATS Returns is the top performer, with the lowest RMSE (390.77), MAE (306.51), and MAPE (0.80%), and a strong MASE (0.5840), indicating high accuracy. Traditional models like Naive Forecast, SMA, and EWMA showed higher errors.

Prophet delivered acceptable results but lagged in accuracy (RMSE: 941.93, MASE: 3.8083). Holt-Winters models had unusually high MAPE, possibly due to data scale issues.
ARIMA-based models performed poorly on raw price forecasting, though ARIMAX showed some improvement. Overall, deep learning models like N-BEATS clearly outperform traditional approaches in this forecasting task.

## Naive Forecast

**Justification:** The naive forecast serves as a foundational benchmark, assuming that the most recent value will persist into the future. This aligns with the random walk theory, often observed in financial markets.

**Advantages:**

- Extremely simple and intuitive
- Requires no parameter tuning or training
- Effective as a baseline for model comparison

**Disadvantages:**

- Fails to capture trends, seasonality, or long-term dynamics
- Poor performance in structured or non-random series

**Applications:**

- Stock Price Prediction: Often used as a benchmark in short-term price forecasting.
- Risk Management: Acts as a stress-testing tool under static market assumptions.

## Simple Moving Average (SMA)

**Justification:** SMA smooths short-term noise to reveal underlying trends. While it's more commonly used in analysis rather than forecasting, it can assist in understanding general momentum over fixed windows.

**Advantages:**

- Easy to compute and interpret
- Useful for highlighting trends in volatile data

**Disadvantages:**

- Lags behind actual data
- Equally weights all past values, ignoring recency
- Sensitive to sudden changes or outliers

**Applications:**

- Stock Price Prediction: Supports technical analysis and helps define support/resistance levels.
- Risk Management: Used to monitor trend reversals and market momentum.

## Exponentially Weighted Moving Average (EWMA)

**Justification:** EWMA gives greater emphasis to recent observations, allowing it to respond faster to new market developments — useful for short-term modeling.

**Advantages:**

- More responsive to recent price changes
- Flexible via the smoothing parameter (alpha)
- Helps track volatility in dynamic markets

**Disadvantages:**

- Still introduces some lag
- Choice of smoothing constant can be arbitrary
- Not designed to handle seasonality

**Applications:**

- Stock Price Prediction: Captures momentum shifts for short-term signals.
- Risk Management: Widely used in volatility estimation, such as in Value-at-Risk (VaR) frameworks.

## Holt-Winters Model

**Justification:** Holt-Winters is suited for time series exhibiting both trend and seasonality. It extends exponential smoothing by including both components, making it a strong candidate for structured financial time series.

**Advantages:**

- Captures level, trend, and seasonal patterns
- Adaptable to changing trends
- Supports additive or multiplicative seasonality

**Disadvantages:**

- Sensitive to parameter tuning
- Can overfit on short or noisy series
- Less robust with irregular seasonality

**Applications:**

- Stock Price Prediction: Effective for equities with cyclical behavior.
- Risk Management: Supports seasonal risk analysis in returns.
- Investment Strategy Development: Aids in strategies tied to fiscal quarters or earnings cycles

## ARIMA

**Justification:** ARIMA models the linear dependence between current and past values in stationary time series. When autocorrelation exists, ARIMA can effectively model patterns and predict future movements.

**Advantages:**

- Models autoregressive and moving average structures
- Handles a wide range of non-seasonal patterns
- Well-established with interpretable parameters

**Disadvantages:**

- Requires stationarity and differencing
- Parameter selection (p, d, q) can be complex
- Lacks built-in support for seasonality or exogenous inputs

**Applications:**

- Stock Price Prediction: Suitable for return-based forecasts with autoregressive patterns.
- Risk Management: Applied in modeling risk metrics such as expected volatility or returns.

## ARIMAX

**Justification:** ARIMAX extends ARIMA by incorporating external (exogenous) variables that may influence the target series. If relevant, these variables can enhance forecast performance.

**Advantages:**

- Incorporates outside information
- Useful for capturing cause-and-effect relationships
- Enhances explanatory power of forecasts

**Disadvantages:**

- Requires identification and preprocessing of relevant external data
- More complex modeling and tuning
- Can underperform if exogenous data is noisy or irrelevant

**Applications:**

- Stock Price Prediction: Incorporates macroeconomic indicators like interest rates, FX rates, or commodity prices.
- Risk Management: Models sensitivity of assets to external shocks.
- Investment Strategy Development: Enables macro-aware strategies that consider economic signals.

## SARIMAX

**Justification:** SARIMAX captures both seasonal effects and exogenous influences. It is especially useful for time series with recurring cycles influenced by outside events or economic indicators.

**Advantages:**

- Comprehensive: models seasonality, trend, and external regressors
- Flexible structure for complex datasets
- Applicable to a wide range of real-world financial time series

**Disadvantages:**

- Highly parameterized and computationally intensive
- Needs extensive tuning for stability and performance
- Risk of overfitting if not carefully managed

**Applications:**

- Stock Price Prediction: Models assets influenced by seasonal effects and economic data.
- Risk Management: Forecasts cyclical risk metrics impacted by external factors.
- Investment Strategy Development: Enables robust multi-factor models that consider seasonal and macroeconomic trends.

## PROPHET

**Justification:** Prophet is a decomposable time series model developed by Meta, particularly effective for capturing trend, holiday, and seasonal components with minimal tuning.

**Advantages:**

- Easy to use with automatic detection of seasonality and trend shifts
- Handles missing data, outliers, and irregular holidays
- Interpretable components for business insights

**Disadvantages:**

- May struggle with complex nonlinear patterns
- Limited customization of seasonal structures
- Less flexible than statistical or neural models for intricate series

**Applications:**

- Stock Price Prediction: Useful for long-term forecasts with known events (e.g., earnings).
- Risk Management: Handles trend breaks and special events in risk exposure.
- Investment Strategy Development: Helps model calendar-based strategies and event reactions.

## N-BEATS

**Justification:** N-BEATS is a deep learning model specialized in time series forecasting. It learns temporal structure directly from data, without needing hand-crafted features or assumptions.

**Advantages:**

- High accuracy with large datasets
- Learns trends and seasonality automatically
- Handles long sequences and complex patterns

**Disadvantages:**

- Requires significant data and compute power
- Lacks built-in support for exogenous variables
- Less interpretable than traditional models

**Applications:**

- Stock Price Prediction: Effective for highly nonlinear or noisy market behavior.
- Risk Management: Captures complex volatility structures over time.
- Investment Strategy Development: Provides high-frequency insights for trading strategies or risk-adjusted return forecasting.

This assignment emphasized the value of time series analysis in understanding and forecasting stock returns, which often follow a random walk. We explored a range of models, from simple baselines like the Naive and Random Walk to more sophisticated techniques such as ARIMA, Prophet, and deep learning models like N-BEATS. Each model has its strengths, basic models are useful benchmarks and often perform well in unpredictable markets, while advanced models can capture complex patterns, trends, and external influences. We also saw how models like SMA and EWMA help smooth volatility and detect trends, which are vital for technical analysis and risk management. Tools like ARIMAX and SARIMAX add depth by incorporating economic variables, while Prophet offers intuitive, calendar-aware forecasts. N-BEATS showcased the potential of modern machine learning in capturing nonlinear relationships without manual feature engineering. Overall, this assignment showed that time series models are not one-size-fits-all—the right choice depends on the goal, whether it's explainability, speed, or accuracy. Understanding these trade-offs is crucial for effective financial decision-making.