Modeling Stock Prices Based on Risk Indicators

A Comparative Study of Machine Learning Models Using Time Series Analysis

Authors: Shiyue (Cynthia) Zhou, Kaisen Yao, Arko Bhattacharya, Skye Augsorn, Mona Saeed

Instructor: Prof. David Ye

Institution: Duke University, IDS 789

Author Note

[The technical component of this project is presented in a GitHub repository[1].]

[1] GitHub repository: https://github.com/CynthiaShiyue/IDS789-Financial-Model-Project

Table of Content

Abstract

This study investigates the relationship between UBS stock returns and various risk factors, utilizing daily data from January 2021 to January 2024. Five machine learning models—ARIMAX, Bayesian Model, Decision Trees, Gradient Boosting Machines, and Neural Networks—were employed to assess the predictive capabilities of these risk factors in forecasting future stock returns. The models were evaluated based on their Mean Squared Forecast Error (MSFE), with all achieving values below 0.001, demonstrating high predictive accuracy. The Neural Network model exhibited the smallest MSFE, while the other models showed minimal differences, with MSFE values all under 0.0002. By comparing the predictive performance of these machine learning models, this study highlights the effectiveness of incorporating risk factors into stock return predictions. The findings provide empirical evidence supporting the utility of machine learning in forecasting stock returns under liquid market conditions. For investors, understanding the influence of risk factors on future stock returns can help optimize investment strategies. This study offers insights into risk-aware investment decision-making, emphasizing the importance of evaluating risk indicators to enhance forecasting accuracy.

1. Introduction

In financial markets, investors and financial institutions often quantify risk to help explain the economics of a company or an asset. Risk factors such as illiquidity or market volatility are at the core of economic instability. In this way, it will be more accurate to reflect the true market conditions by incorporating these factors into the analysis. For example, the VIX index, known as the "fear index," measures the implied volatility of S&P 500 index options (Kuepper, 2024). The VIX is usually used by investors and financial institutions as a risk management indicator to evaluate potential portfolio volatility and develop trading strategies. By analyzing different categories of risk indicators, investors can predict asset prices and market trends based on the changing trends of these indicators, thereby optimizing investment decisions, adjusting portfolios, reducing losses, and increasing returns.

Although risk indicators are widely used in financial markets, how to effectively quantify these risk factors and integrate them into forecasting models remains a major challenge. In the process of optimizing forecasting models by incorporating risk factors, it is particularly important to select appropriate factors. Because different risk factors contribute differently to forecast accuracy, how to filter and evaluate these factors scientifically has become an important topic.

To address these challenges, this study focuses on analyzing the role of key risk factors in predicting UBS stock returns. By comparing the prediction accuracy of multiple machine learning models, this study analyzes different risk factors as predictors and stock returns as the outcome variable. Using mean squared prediction error (MSFE) as a measure, this study evaluates the prediction performance of each model and further emphasizes the effectiveness of integrating risk factors into stock return prediction. In addition, this study explores the relationship between stock returns and multiple risk indicators, analyzes the overall situation of the company and its risk management practices through stock price models, and evaluates the predictive power of these indicators. Ultimately, this study aims to achieve the broader goal of improving risk-aware investment strategies and promoting the application and development of machine learning technology in financial markets.

2. Data Source

2.1. Data Overview

The dataset for this study was sourced using the Yahoo Finance package in Python[2], focusing on daily financial and market data for UBS Group AG (UBS) from January 4, 2021, to January 5, 2024. This period was chosen to capture a wide range of market dynamics and fluctuations, providing a comprehensive basis for analyzing UBS's stock performance. The dataset includes UBS stock prices as the target variable and a carefully selected set of explanatory variables representing financial, macroeconomic, and market risk factors. These variables collectively aim to capture the multifaceted influences on UBS stock returns.

Data quality was ensured through a rigorous cleaning process. Missing values, duplicate records, and outliers were identified and removed to improve the dataset's reliability. All explanatory variables were standardized and normalized to account for differences in scale, ensuring consistency across the dataset and facilitating seamless integration into the modeling workflow. The target variable, UBS's daily log return, was calculated to address non-stationarity issues associated with raw stock prices. This transformation ensures compatibility with statistical and machine learning models, making the data suitable for robust risk and return forecasting. The dataset was divided into two subsets: a training set (January 4, 2021–January 2, 2023) for model

[2] The dataset was retrieved using the Yahoo Finance package, a Python library for accessing financial data directly from Yahoo Finance. For implementation details, refer to the extract.py script.

development and a testing set (January 3, 2023–January 5, 2024) for evaluation, stored as training_dataset.csv and testing_dataset.csv.

2.2. Data Description

The dataset comprises two main types of variables: the target variable, which represents UBS's daily return, and eight explanatory variables, each carefully chosen to capture critical drivers of stock performance.

2.2.1. Target Variables

The target variable was transformed into the natural logarithmic return of UBS stock prices, ensuring stationarity and compatibility with modeling techniques. This approach enhances the analysis by providing a reliable measure of percentage changes in stock prices while mitigating non-stationary trends.

2.2.2. Explanatory Variables

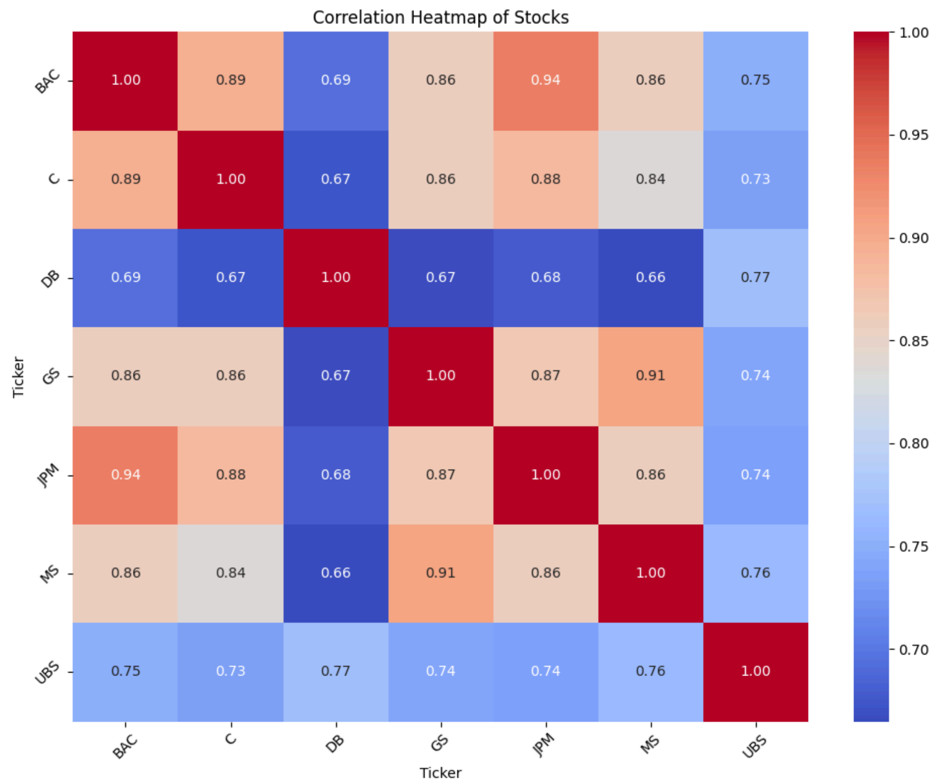
The explanatory variables encompass a diverse range of factors to provide a comprehensive understanding of UBS stock performance. Liquidity risk indicators include the bid-ask spread, which measures liquidity by reflecting the difference between buying and selling prices, and trading volume, which indicates market activity and investor sentiment by capturing the total shares traded daily. These indicators offer insights into how readily UBS stock can be traded without significant price impact.

Macroeconomic indicators such as the SPY 500 and FTSE 100 indices are included to account for broader economic conditions. The SPY 500 tracks the performance of 500 leading U.S. companies, serving as a proxy for the U.S. market, while the FTSE 100 provides insights into European market trends by reflecting the performance of the largest companies listed on the London Stock Exchange.

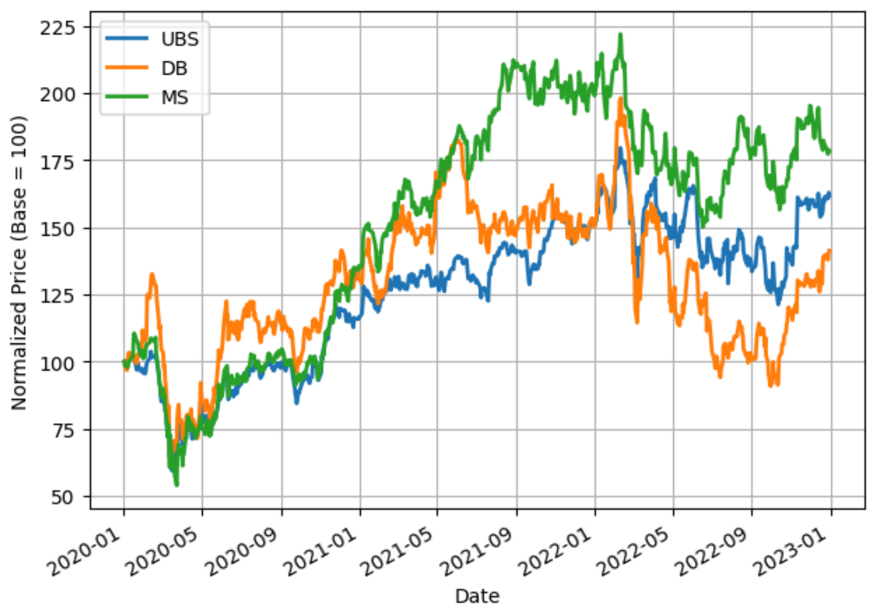
Market risk is assessed using the CBOE Volatility Index (VIX), commonly referred to as the “fear index.” This index measures expected market volatility, offering a gauge of investor sentiment and risk perception. Currency pair volatility is captured through the EUR/CHF exchange rate, reflecting fluctuations between the Euro and Swiss Franc. This is particularly relevant to UBS given its Swiss operations and regional exposure.

Commodity prices, specifically oil and gold, were also included. Oil prices, measured through WTI Crude Futures, reflect economic growth, inflationary pressures, and energy costs, while gold prices, measured through Gold Futures, act as a hedge against inflation and a measure of investor confidence during periods of uncertainty. These variables offer critical insights into the external factors influencing global financial stability.

To analyze UBS’s performance in comparison to its competitors, a Python-based methodology was employed to process and visualize financial data . The adjusted closing prices of UBS and its peers, including Morgan Stanley (MS) and Deutsche Bank (DB), were retrieved using the Yahoo Finance API. Daily returns were calculated to assess short-term performance variability, followed by computing a correlation matrix to identify relationships between stock movements. The two competitors most closely correlated with UBS were determined through this analysis, providing valuable benchmarks for evaluating performance trends. A heatmap [figure 2.2.2(a)] was generated to visualize the correlations across the financial sector, highlighting UBS’s alignment with other institutions. Additionally, normalized stock prices were plotted [figure 2.2.2(b)] to compare UBS’s performance trajectory with those of its top correlated competitors, offering an intuitive view of relative market trends over time.



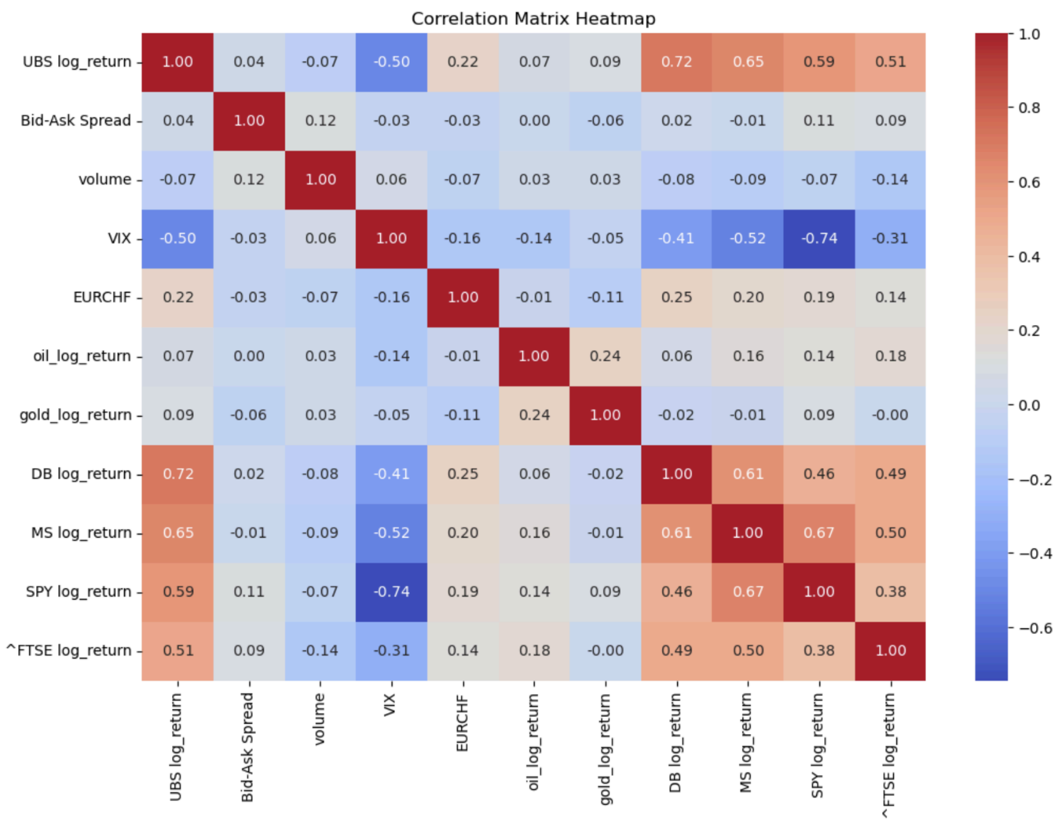
[Figure 2.2.2(a)] Correlation for UBS and its competitors' stock movements



[Figure 2.2.2(b)] Normalized Stock Price Movements for UBS and Top 2 Competitors

2.2.3. Correlation Analysis

Together, these explanatory variables provide a balanced and comprehensive view of the factors driving UBS's stock returns, forming a robust foundation for the modeling process.



[Figure 2.2.3] Correlation Matrix among all variables

Figure 2.2.3 illustrates the correlation between the target variable (UBS log_return) and the predictors (bid-ask spreads, trading volume, SPY 500 and FTSE 100 indexes, VIX, EUR/CHF, and oil and gold prices). By observing the first column in the matrix of Figure 1, we can see that volume and the VIX variables have a negative index, which implies a negative relationship between them: when volume and VIX decrease, the target variable increases. In contrast, all other variables have a positive relationship with the target variable. In addition, when we compare the absolute values in the first column of the matrix, we can find that VIX, SPY log_return, ^FTSE log_return, DB log_return, and MS log_return have a correlation index exceeding 0.5 with the target variable. This indicates that these variables are likely to contribute significantly to our prediction results. On the contrary, for those variables with a small correlation index, Bid-Ask Spread (0.04), volume (-0.07), and gold_log_return (0.09), they may not significantly contribute to predicting UBS stock returns. These risk factors with a high correlation index can help us explain the target variable and improve the accuracy of predicting UBS stock returns. However, we cannot directly exclude variables with a low correlation index based on the correlation matrix alone, because of the properties of the correlation matrix, which only evaluates linear relationships and may ignore nonlinear or complex interactions. It is important to do further analysis and research to analyze these low-correlation variables.

3. Methodology

This study aims to model and predict the daily log returns of UBS's stock using a comprehensive set of explanatory variables and a diverse range of modeling techniques. First, we transform raw closing prices into daily log returns to ensure stationarity and facilitate reliable statistical inference. We then carefully select a broad array of explanatory variables — encompassing liquidity measures (e.g., bid-ask spread, trading volume), macroeconomic and market indicators (e.g., S&P 500, FTSE 100, VIX), currency pairs (EUR/CHF), commodity prices (oil, gold), and peer institutions' stock returns (Deutsche Bank, Morgan Stanley) — to capture the multifaceted environment influencing UBS's returns. The dataset undergoes rigorous preprocessing, including handling missing values, outlier adjustments, and appropriate standardization, ensuring a robust foundation for subsequent modeling efforts.

Following data preparation, we implement an array of modeling approaches to forecast UBS's stock returns. We begin with traditional time series models, including ARIMA and ARIMAX, to establish a benchmark and incorporate exogenous variables. We then explore a Bayesian framework to incorporate prior knowledge and manage uncertainty systematically. To capture non-linear relationships and interaction effects, we employ advanced machine learning methods: decision trees, random forests, and gradient boosting machines (GBM). Finally, we assess the performance of a Long Short-Term Memory (LSTM) network, leveraging its capacity to model complex temporal dynamics. Forecast accuracy is consistently evaluated using the Mean Squared Forecast Error (MSFE), providing a standardized measure of prediction quality. This multifaceted approach allows us to compare a range of models, identify the factors driving UBS's returns, and select the most effective forecasting methodology.

3.1. Variables Calculation

The study involves two types of variables: the target variable, representing UBS's daily log stock return, and explanatory variables, which include financial and macroeconomic indicators influencing stock performance. These variables collectively aim to uncover patterns and drivers of risk and return.

3.1.1. Target Variable

The target variable is the daily log return of UBS's stock price. Let P_t denote the closing price of the UBS stock on day t. The daily log return Y_t is defined as: $Y_t = log(\frac{P_t}{P_{t-1}})$. This transformation enhances stationarity and normalizes the distribution of returns, allowing for more reliable estimation and inference in subsequent modeling steps.

3.1.2. Explanatory Variables

To comprehensively capture the multifaceted nature of UBS's stock return movements, we incorporate a diverse set of explanatory variables spanning liquidity measures, macroeconomic and market indices, currency and commodity prices, and peer institution performance. With the exception of certain liquidity indicators, all variables are expressed in terms of daily log returns to ensure consistency and comparability across different measures.

3.1.2.1 Liquidity Measures

Liquidity is a critical factor influencing stock returns, as it reflects the ease with which assets can be bought or sold in the market without affecting their price. We employ two primary liquidity indicators: the bid-ask spread and trading volume.

Bid-Ask Spread:

To capture market microstructure effects, we utilize the Corwin and Schultz (2012) estimator for the bid-ask spread. This estimator leverages high and low prices over consecutive days to derive a robust proxy for transaction costs and market liquidity. Let H_t and L_t denote the high and low prices on day t, respectively. The estimator is defined as: $\alpha = \sqrt{\frac{\beta}{\gamma}}$, where $\beta = (ln\,\frac{H_t}{L_t})^2 + (ln\,\frac{H_{t-1}}{L_{t-1}})^2$, $\gamma = (ln\,\frac{H_t}{L_{t-1}})^2$. The final bid-ask spread estimate S_t is derived from α through the Corwin-Schultz formulation: $S_t = \alpha \cdot \sigma$, where σ represents the standard deviation of the bid and ask prices. This measure is retained in its raw form to directly reflect implicit transaction costs and market liquidity without the need for logarithmic transformations.

Trading Volume:

Trading volume serves as a direct proxy for market activity and liquidity conditions. We incorporate raw daily trading volume V_t without any transformation, as it inherently represents the level of market participation and does not require normalization for interpretability.

3.1.2.2 Market and Macroeconomic Indices

Macroeconomic factors and broad market trends significantly influence individual stock returns. We include major stock market indices to capture these overarching effects.

S&P 500 (SPY) and FTSE 100 (FTSE):

Let P_t^{SPY} and P_t^{FTSE} represent the closing levels of the S&P 500 and FTSE 100 indices on day t, respectively. The daily log returns for these indices are calculated as: $R_t^{SPY} = ln\left(\frac{P_t^{SPY}}{P_{t-1}^{SPY}}\right)$, $R_t^{FTSE} = ln\left(\frac{P_t^{FTSE}}{P_{t-1}^{FTSE}}\right)$. These log returns encapsulate the broad market sentiment and serve as benchmarks for assessing UBS's performance relative to major economic indicators.

3.1.2.3 Market Risk Indicator

CBOE Volatility Index (VIX):

The VIX is a widely recognized measure of market risk and investor sentiment regarding volatility. Let P_t^{VIX} denote the VIX level on day t. The daily log return is defined as:

$R_t^{VIX} = ln\,(\frac{P_t^{VIX}}{P_{t-1}^{VIX}})$. This variable captures the market's expectation of near-term volatility, providing insights into risk perceptions that may influence UBS's stock returns.

3.1.2.4 Currency and Commodity Prices

Fluctuations in currency exchange rates and commodity prices can have profound effects on financial institutions, especially those with significant international exposure.

EUR/CHF Exchange Rate:

Let $P_t^{EUR/CHF}$ represent the EUR/CHF exchange rate on day t. The daily log return is calculated as: $R_t^{EUR/CHF} = ln\,(\frac{P_t^{EUR/CHF}}{P_{t-1}^{EUR/CHF}})$. This log return reflects currency-related risks that could impact UBS's international operations and financial performance.

Commodity Prices (Oil and Gold):

For oil and gold prices, denoted by P_t^{Oil} and P_t^{Gold}respectively, the daily log returns are defined as: $R_t^{Oil} = ln\,(\frac{P_t^{Oil}}{P_{t-1}^{Oil}})$, $R_t^{Gold} = ln\,(\frac{P_t^{Gold}}{P_{t-1}^{Gold}})$. These log returns capture the influence of essential input costs and reserve asset prices on UBS's financial stability and risk profile.

3.1.2.5 Peer Institution Stock Prices

Benchmarking against peer institutions provides contextual insights into UBS's performance within the industry.

Deutsche Bank (DB) and Morgan Stanley (MS):

Let P_t^{DB} and P_t^{MS} denote the closing stock prices of Deutsche Bank and Morgan Stanley on day tt, respectively. The daily log returns are calculated as: $R_t^{DB} = ln\,(\frac{P_t^{DB}}{P_{t-1}^{DB}})$, $R_t^{MS} = ln\,(\frac{P_t^{MS}}{P_{t-1}^{MS}})$. Including these variables allows for the assessment of industry-specific shocks and competitive dynamics that may affect UBS's stock returns.

3.2 Data Cleaning and Preprocessing

Prior to model estimation, all data series underwent rigorous cleaning and preprocessing to ensure accuracy and suitability for analysis. Missing observations were addressed by either imputing them using appropriate statistical techniques or removing them entirely, thereby preventing biased parameter estimates and distortions within the dataset. Extreme values, or outliers, were identified and treated through methods such as winsorization, which mitigates their undue influence on model stability and performance. Additionally, explanatory variables were standardized to have zero mean and unit variance where applicable. This standardization step enhances comparability across variables and stabilizes numerical optimization processes during model fitting.

3.3 Modeling Approach

In this section, we apply five different machine learning models to do the stock's return forecasting. To quantify our prediction accuracy and forecasting error, we use the Mean Squared Forecast Error (MSFE) method, which evaluates the average squared difference between predicted and actual values. This error evaluation method provides a way to estimate how well the model performs. In general, smaller MSFE values indicate higher predictive accuracy.

3.3.1 ARIMA (AutoRegressive Integrated Moving Average)

The AutoRegressive Integrated Moving Average (ARIMA) model is a cornerstone in time series forecasting, known for its ability to model linear dependencies in sequential data. In this study, ARIMA was employed as an initial approach to predict UBS stock returns, serving as a benchmark model due to its simplicity and interpretability. By decomposing a time series into three components—autoregression (AR), differencing (I), and moving averages (MA)—ARIMA captures temporal dependencies while ensuring stationarity.

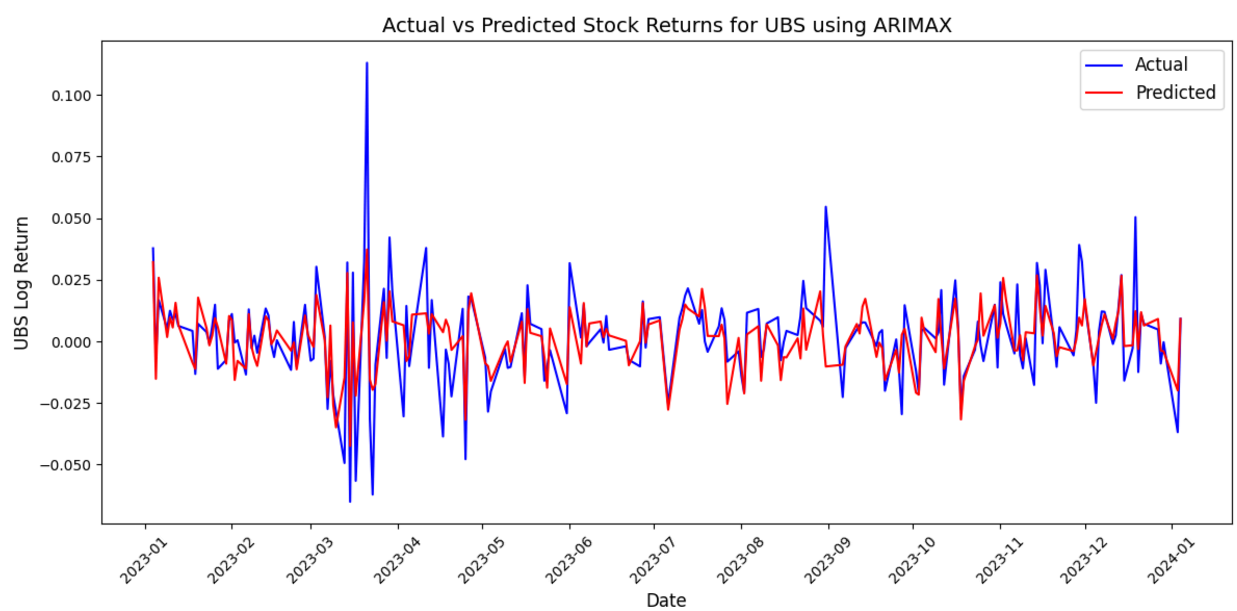
The autoregressive (p) component captures the influence of past observations, the differencing (d) component stabilizes the series by removing trends, and the moving average (q) component accounts for dependencies on past forecast errors. These elements are parameterized as (p, d, q). The general mathematical form of ARIMA is:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where y_t represents the series at time t, ϕ are the autoregressive coefficients, θ are the moving average coefficients, and ε_t is the error term.

For ARIMA to yield accurate predictions, the time series must be stationary. This assumption was confirmed in this study using the Augmented Dickey-Fuller (ADF) test, which produced an ADF statistic of -16.02 and a p-value of 6.17e-29, indicating stationarity without requiring differencing ($d = 0$). Despite this, the best-fitting ARIMA model identified using the Akaike Information Criterion (AIC) was ARIMA(0,0,0), effectively reducing the model to a constant mean. This result underscored ARIMA's limitations in capturing the underlying patterns of UBS stock returns, necessitating the inclusion of exogenous predictors.

To address the limitations of ARIMA, the ARIMAX model was employed, extending ARIMA by incorporating external risk factors such as the bid-ask spread, trading volume, VIX, and commodity prices. These exogenous variables allowed ARIMAX to account for external drivers of stock returns, enhancing its explanatory power. The optimal ARIMAX model identified was ARIMA(0,0,1), selected based on AIC minimization, which produced a significantly lower Mean Squared Forecast Error (MSFE) of 0.000181, demonstrating superior performance compared to ARIMA.



[Figure 3.3.1] ARIMAX Analysis: Predicted UBS Stock's Return VS Actual UBS Stock's Return

While ARIMAX improved forecasting accuracy by leveraging external predictors, it encountered convergence issues during parameter estimation, highlighting potential multicollinearity among exogenous variables or data limitations. Additionally, like ARIMA, ARIMAX assumes linear relationships, which restricts its ability to model complex nonlinear dynamics in financial data.

In conclusion, ARIMA provided a useful starting point for time series modeling, but its limitations necessitated a shift to ARIMAX to incorporate exogenous variables. ARIMAX significantly enhanced the model's predictive accuracy, showcasing the importance of integrating external factors in financial forecasting. However, both models remain constrained by their linear assumptions, pointing to the need for more advanced techniques, such as machine learning models, to capture the complexities of financial data.

3.3.2 Bayesian Model

According to the characteristics of the Bayesian model, “The Bayesian statistical paradigm uses the rules and language of probability to quantify uncertainty about all unknown aspects of phenomena that generate observed data. This core characteristic of the paradigm makes it particularly suitable for forecasting, with uncertainty about the unknown values of future observations automatically expressed in terms of a probability distribution” (Martin et al., 2024). As a statistical method that captures uncertainty and incorporates prior knowledge, Bayesian approaches excel in using new data to estimate parameters. Due to the inherent randomness of time series data, our target variable—being part of a time series—is influenced by past values and carries prior information. When we expand the testing dataset, the new data provides additional information that may influence the probabilities and the way the data supports the target variable. Based on these features, we applied Bayesian methods in time series models to construct our prediction model.

In order to improve the accuracy of our prediction model, we will create a weekday variable extracted from the date of UBS stock returns and use it as a hierarchical prior.

According to the definition of hierarchical priors, “Hierarchical priors, which allow parameters to share information through higher-level distributions, enable partial pooling, reduce overfitting, and improve predictions, especially in cases with structured or sparse data” (Gelman et al., Bayesian Data Analysis, 2013). Due to these advantages, we will use hierarchical priors to estimate parameters and construct our Bayesian model. In general, the Bayesian approach starts with an original set of prior probabilities. After obtaining data from an experiment, we find the posterior probabilities. Those posterior probabilities become the new prior for the next experiment, which we combine with new data to get updated prior probabilities (Buyske, 2021). Based on these definitions, we construct the fitted model as follows.

$$Hyperpriors: \mu_{weekday} \sim N(0, 10), \quad \sigma_{weekday} \sim HalfNormal(10)$$

$$Priors: \alpha_{weekday}[j] \sim N(\mu_{weekday}, \sigma_{weekday}), \quad j = 0, 1, ..., 6 \ (Weekdays)$$

$$\beta_k \sim N(0, 10), \quad k = 1, ..., p \ (Predictors)$$

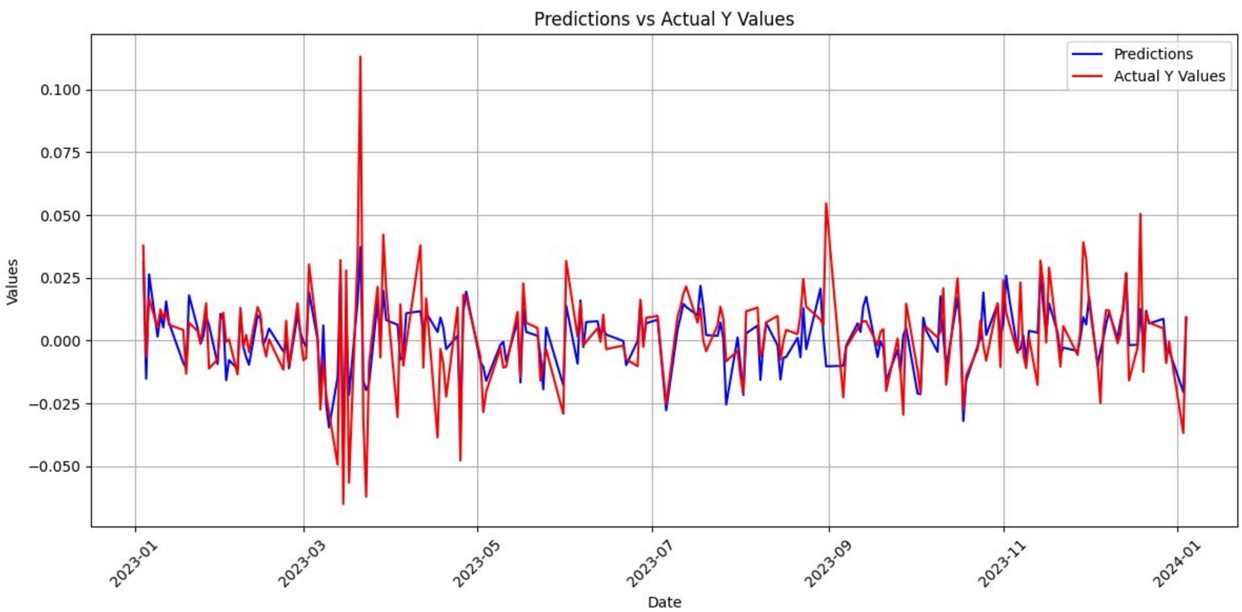
$$\sigma \sim HalfNormal(10)$$

$$Linear\ Model: \mu_i = \alpha_{weekday}[weekday(i)] + \sum_{k=1}^{p} \beta_k X_{ik}, \quad i = 1, ..., n$$

$$Likelihood: y_i \sim N(\mu_i, \sigma), \quad i = i, ..., n$$

$$Prediction\ for\ Test\ Data: \hat{y}_j = \alpha_{weekday}[weekday(j)] + \sum_{k=1}^{p} \beta_k X_{jk}, \quad j = 1, ..., m$$

After applying our training dataset (2021–2023), which includes risk factor predictors and actual UBS stock returns, to the Bayesian prediction model, we obtain the predicted stock returns. To evaluate the prediction results, we will compare them with the actual UBS stock returns from our testing dataset (2023–2024).



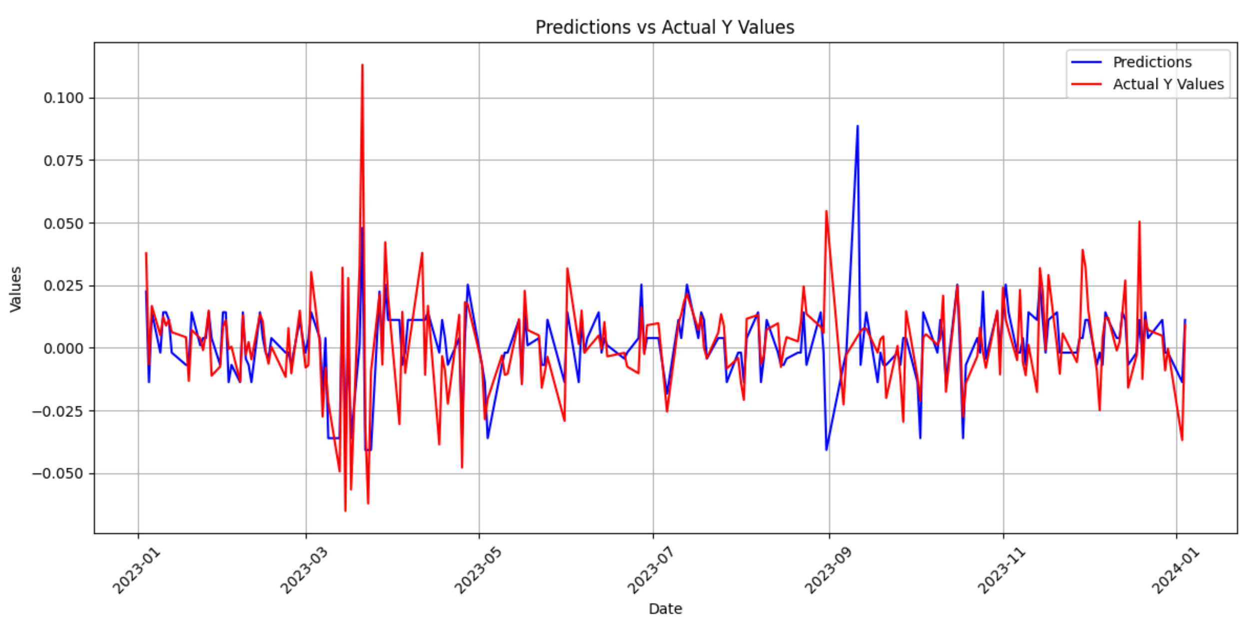
[Figure 3.3.2] Bayesian Data Analysis: Predicted UBS Stock's Return VS Actual UBS Stock's Return

Figure 3.3.2 illustrates the comparison between the predicted UBS stock returns and the actual UBS stock returns. Overall, the plot shows that the actual and predicted returns follow a similar trend. The differences between the predicted and actual values (red y-value minus blue y-value for the same x) are generally small, indicating that the model performs well in leveraging prior data for prediction. By applying the MSFE method to our prediction, we obtain an MSFE of 0.00018 for the Bayesian model. This small value demonstrates the model's high forecasting accuracy.

In conclusion, applying a Bayesian model with risk factors enables accurate predictions of stock returns.

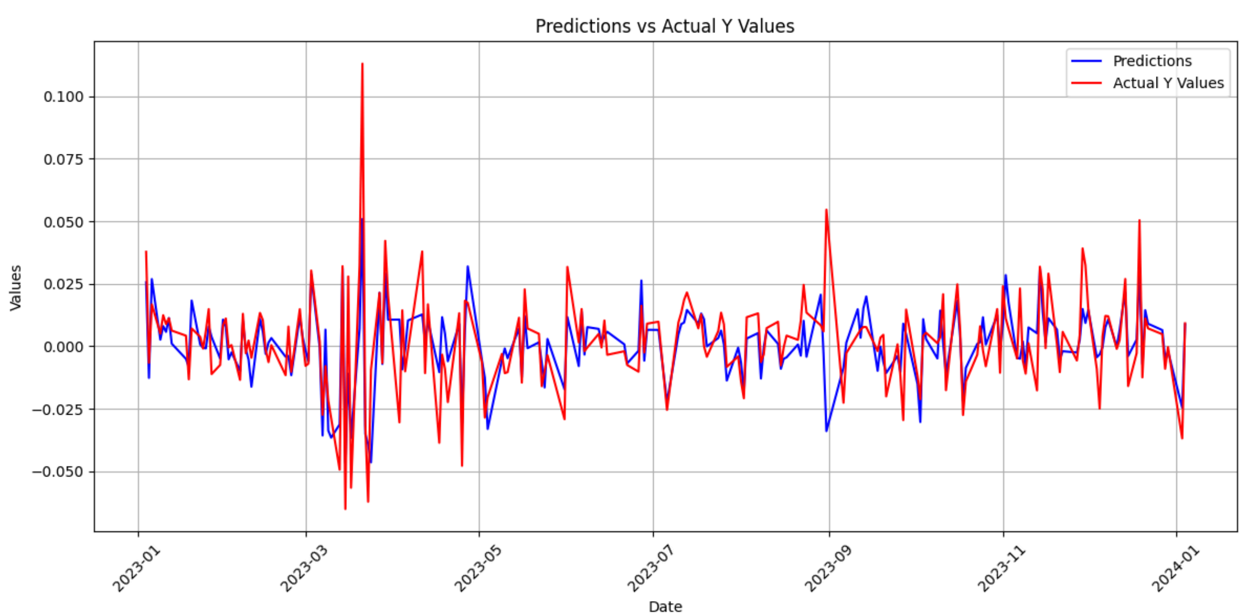
3.3.3 Decision Trees

The decision tree model was utilized as a non-parametric supervised learning method to predict UBS log returns. This model works by recursively splitting the dataset into subsets based on feature values to minimize prediction error at each step, making it suitable for capturing non-linear relationships and complex patterns in data (Breiman et al., 1984). To enhance the model's predictive capability, various features were engineered. Lagged values of UBS log returns were introduced to account for temporal dependencies, and additional variables like day, month, and year were extracted from the date column to capture seasonal effects. Interaction terms were also created, such as the product of the bid-ask spread and volume (spread volume) and the product of oil log returns, and the VIX index (vix_oil), to account for economic interactions. Hyperparameter optimization was performed using GridSearchCV to fine-tune parameters like maximum depth, minimum samples per split, and split evaluation criteria. Despite these efforts, the decision tree model achieved a modest R^2 score of 0.386, indicating it could explain only a limited portion of the variance in UBS log returns. While the Mean Squared Forecast Error (MSFE) and Mean Absolute Error (MAE) were low; 0.01 and 0.0002 respectively, suggesting the model was not overfitting, the R^2 score highlighted the need for further improvement.



[Figure 3.3.3.1] Decision Tree: Predicted UBS Stock's Return VS Actual UBS Stock's Return

To address the limitations observed with the decision tree model, a random forest model, introduced by Breiman (2001), was implemented. This ensemble learning method builds multiple decision trees during training and averages their predictions, reducing the likelihood of overfittings and improving generalization. The random forest model also utilized the engineered features and interaction terms from the decision tree approach. GridSearchCV was again employed to optimize hyperparameters such as the number of estimators, maximum depth, and minimum samples per split. Although hyperparameter tuning did not result in substantial gains, the random forest model demonstrated improved performance, achieving an R^2 score of 0.54. This indicates that 54% of the variance in UBS log returns was explained, a significant improvement over the decision tree model. Furthermore, the low MSFE and MAE; 0.008 and 0.0002, respectively, reinforced the model's accuracy. These results suggest that the random forest's ability to combine multiple trees effectively captured more nuanced patterns within the data.



[Figure 3.3.3.2] Random Forest: Predicted UBS Stock's Return VS Actual UBS Stock's Return

3.3.4 Gradient Boosting Machines (GBM)

Gradient Boosting Machines (GBM) were employed as an advanced ensemble learning method to predict UBS stock returns. GBM builds upon the foundation of decision trees but addresses their limitations through an iterative process of error correction. Unlike single decision trees or random forests that build trees independently, GBM constructs trees sequentially, with each new tree specifically trained to correct the errors of its predecessors (Friedman, 2001).

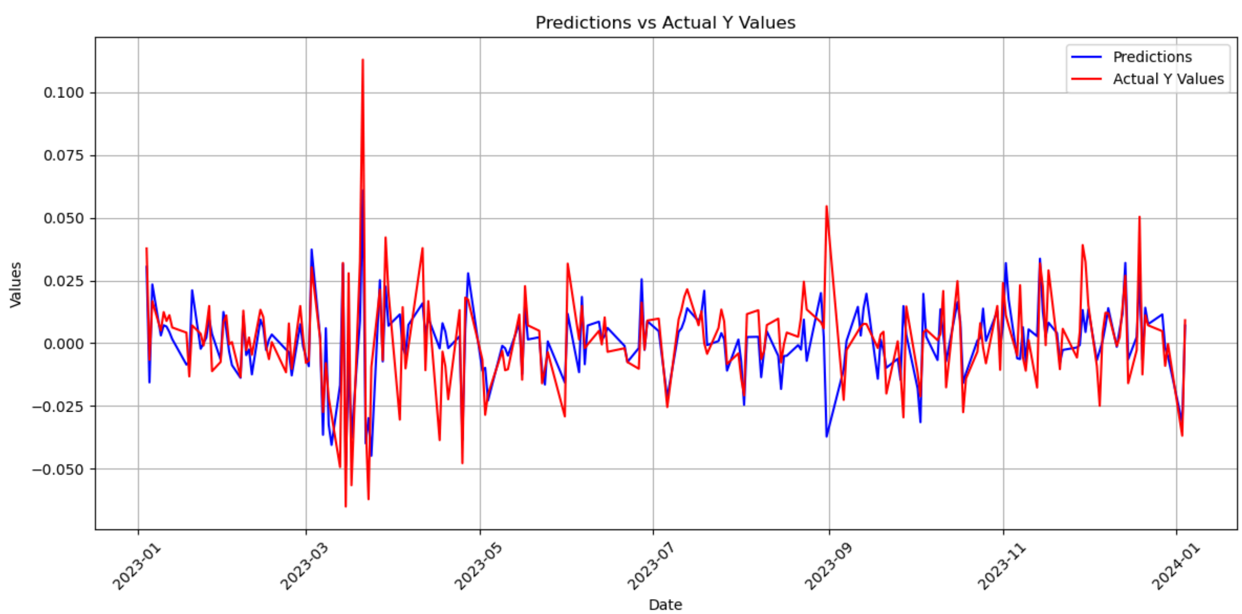
The model integrates multiple predictors including market liquidity indicators (bid-ask spread, trading volume), volatility measures (VIX), currency risk (EUR/CHF), commodity market influences (oil and gold returns), peer performance (DB and MS returns), and broader market indices (SPY and FTSE returns). These predictors were carefully selected to capture different aspects of market dynamics that could influence UBS stock returns. The mathematical framework of GBM can be expressed as: $F(x) = \sum_{i=1}^{n} \gamma_i h_i(x)$, where F(x) is the final prediction, $h_i(x)$ represents individual decision trees, and γ_i are the weights learned during the boosting process.

The implementation utilized key hyperparameters including 100 estimators (trees), a learning rate of 0.1, and a maximum tree depth of 3. These parameters were chosen to balance model complexity with generalization ability. The relatively shallow tree depth helps prevent overfitting, while the moderate learning rate allows for gradual model improvement. The subsample ratio of 0.8 introduces randomness into the training process, further enhancing the model's robustness.

Feature importance analysis revealed significant insights into the drivers of UBS stock returns. Deutsche Bank (DB) returns emerged as the most influential predictor, accounting for

45.5% of the model's predictive power, followed by trading volume (17.3%) and Morgan Stanley (MS) returns (11.5%). This hierarchy suggests that peer bank performance, particularly European banks, strongly influences UBS stock movements. Market indices like SPY (7.3%) showed moderate importance, while macroeconomic factors such as EUR/CHF exchange rates (2.2%) and oil returns (2.3%) had relatively minor impacts.

The model demonstrated strong predictive performance with an R-squared value of 0.506, indicating it explains approximately 50.6% of the variance in UBS stock returns. The Mean Squared Forecast Error (MSFE) of 0.000192 and Root Mean Squared Error (RMSE) of 0.013846 suggest high prediction accuracy, comparing favorably with other modeling approaches in this study. These metrics demonstrate the model's ability to capture both linear and non-linear relationships in the data while maintaining generalization capability.



[Figure 3.3.4] GBM: Predicted UBS Stock’s Return VS Actual UBS Stock’s Return

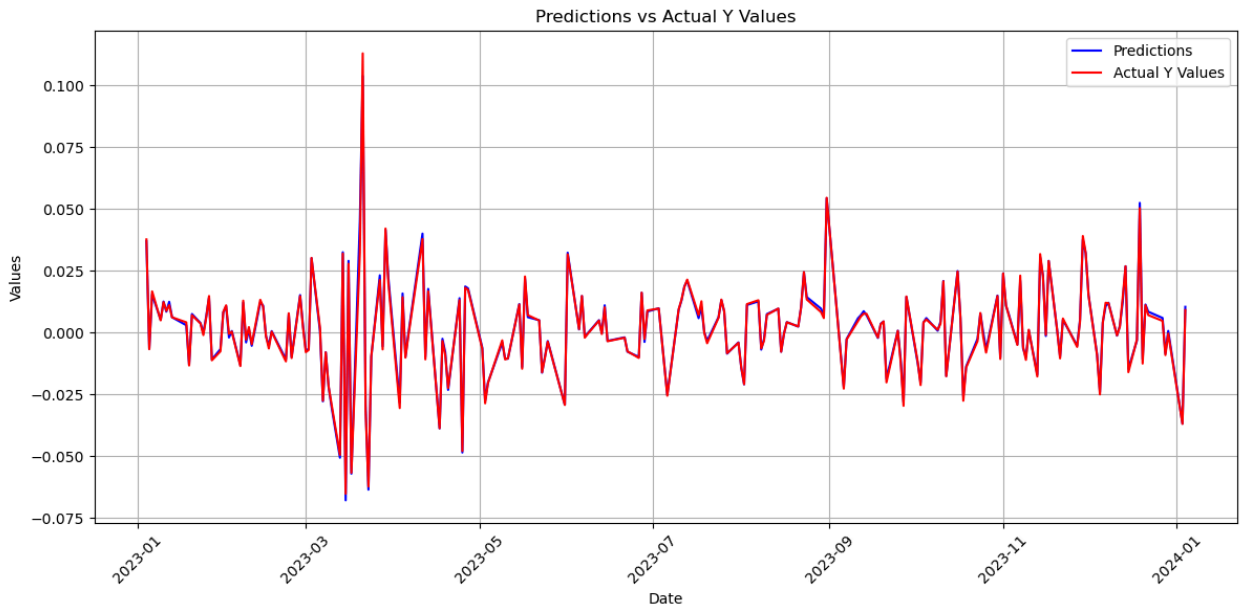
However, GBM's implementation faced certain challenges. The model's computational intensity required careful parameter tuning to maintain efficiency. Additionally, while GBM excels at capturing non-linear relationships, its black-box nature makes it more difficult to interpret compared to simpler models like ARIMA. The risk of overfitting was mitigated through cross-validation and careful hyperparameter selection, but this required significant computational resources.

In conclusion, GBM proved to be a powerful tool for predicting UBS stock returns, effectively capturing complex market dynamics through its iterative boosting process. The model's ability to identify Deutsche Bank returns as the primary predictor aligns with economic intuition about the interconnectedness of European banking stocks. While computational overhead and interpretation challenges exist, the strong predictive performance and meaningful feature importance rankings demonstrate GBM's value for financial forecasting tasks.

3.3.5 Long Short-Term Memory Networks (LSTM)

The Long Short-Term Memory (LSTM) network, a type of Recurrent Neural Network (RNN), was tested as part of the model comparison for predicting UBS log returns. LSTMs are particularly suited for time-series data due to their ability to capture long-term dependencies and temporal relationships, addressing issues like the vanishing gradient problem (Hochreiter & Schmidhuber, 1997). This makes LSTMs a powerful choice for financial data, where patterns evolve over time. The data preparation steps for the LSTM model included standardizing features such as UBS log returns, bid-ask spread, volume, VIX, and other indicators using StandardScaler to stabilize and improve performance. The standardized data was reshaped into a three-dimensional format to match the sequential input requirements of the model. The dataset was transformed into sequences of lagged UBS log returns, with each sequence serving as the input and the subsequent return as the target. Additional features, such as spread_volume, vix_oil, and temporal information like day, month, and year, were included to capture market interactions and seasonality.

The LSTM architecture comprised multiple layers tailored for sequential data. The first LSTM layer, with 128 units, was configured to return sequences, allowing further processing by a second LSTM layer with 64 units. Dropout regularization was applied after each LSTM layer, randomly setting 20% of weights to zero during training to mitigate overfitting. A dense layer with 32 units captured additional representations, while the output layer, a single dense unit, predicted the next UBS log return. The model was optimized using the Adam optimizer, known for its efficiency in handling sparse gradients and non-stationary objectives, and used Mean Squared Error (MSE) as the loss function. It was trained over 50 epochs with mini-batches of size 32, effectively learning patterns while preventing overfitting. Testing data underwent the same preprocessing steps, including standardization and reshaping, ensuring consistency between training and testing phases.



[Figure 3.3.5] LSTM Prediction Analysis: Predicted UBS Stock's Return VS Actual UBS Stock's Return

The LSTM model achieved the lowest Mean Squared Forecast Error (MSFE) of 3.9229e-6, underscoring its accuracy in minimizing forecasting errors and outperforming other models in terms of forecast precision. Despite its strengths, the LSTM model's "black-box" nature remains a limitation, making it difficult to interpret how specific features influence predictions. This lack of transparency can be a drawback in financial applications that demand clear model explanations. Nonetheless, the LSTM's ability to capture complex temporal relationships and long-term dependencies highlights its potential as a forecasting tool. Future work could focus on improving model interpretability or exploring hybrid approaches that balance predictive performance with transparency, particularly in scenarios where understanding the influence of features is critical.

4. Conclusion

This study examined the relationship between UBS Group AG's stock returns and a range of financial, macroeconomic, and risk-related indicators using various machine learning and statistical models. The evaluation of predictive performance was conducted using the Mean Squared Forecast Error (MSFE), with lower values indicating superior accuracy.

Among the models, the Long Short-Term Memory (LSTM) neural network emerged as the most accurate, achieving an MSFE of 0.00000392. This performance underscores its ability to capture nonlinear relationships and complex temporal dependencies in financial data. However, the LSTM model's high computational demands and lack of interpretability may limit its applicability in resource-constrained or transparency-critical settings.

ARIMAX and Bayesian models, both with MSFE values of 0.000181, delivered comparable performance, each with unique strengths. ARIMAX leveraged exogenous predictors such as VIX and trading volume, demonstrating its ability to incorporate external influences into forecasts. Meanwhile, the Bayesian model excelled in providing probabilistic forecasts, offering valuable insights for assessing uncertainty—an important consideration when understanding risk in dynamic systems like stock markets.

Gradient Boosting Machines (GBM) and Decision Trees trailed slightly behind, with MSFE values of 0.000192 and 0.0002, respectively. Although GBM is generally robust, its performance in this study may have been constrained by limited feature engineering or hyperparameter tuning. Decision Trees, while interpretable, provided a baseline for comparison but struggled with generalization.

This study highlights a clear trade-off between accuracy and interpretability in forecasting models. Advanced models such as LSTM and Random Forest deliver high accuracy, making them suitable for tasks requiring precise predictions. However, simpler models like ARIMAX and Bayesian methods offer transparency, computational efficiency, and utility in understanding uncertainty. The consistent importance of variables such as trading volume, the VIX index, and peer bank performance reinforces their role as key drivers of UBS stock returns.

5. Discussion

5.1 Shortcomings, and Future Enhancements

While this study showcased the predictive capabilities of a diverse set of models, several overarching limitations need to be addressed. One major challenge was the relatively short time horizon of the dataset, covering only three years. This limited scope may have constrained the models' ability to generalize across varying market conditions, such as major financial shocks or long-term economic cycles. Expanding the dataset to include a longer historical period or additional stock datasets could provide broader insights and enhance model reliability.

The interpretability of the models also varied widely. Simpler models like ARIMAX provided transparency and computational efficiency, making them more suitable for tasks requiring explainability. However, they struggled to capture the complex nonlinear relationships inherent in financial data. Advanced models such as Gradient Boosting Machines (GBM) and Long Short-Term Memory (LSTM) networks excelled in predictive accuracy, leveraging their ability to capture intricate patterns and temporal dependencies. Nonetheless, their "black-box" nature and computational intensity presented challenges in understanding how specific features influenced predictions.

Another limitation was the selection and representation of explanatory variables. While the inclusion of variables like the VIX, trading volume, and peer bank performance (e.g., Deutsche Bank and Morgan Stanley returns) improved predictions, potential multicollinearity and limited feature engineering might have constrained the models' performance. Future studies could explore advanced techniques like interaction terms, non-linear transformations, or feature selection algorithms to enhance the representation of these predictors.

From a methodological perspective, the models exhibited distinct challenges. While simpler approaches like Decision Trees and Bayesian models offered flexibility, they were limited in their ability to capture the volatility and complexity of UBS stock returns. Meanwhile, advanced methods like GBM and LSTM required substantial computational resources and were sensitive to hyperparameter tuning, which may limit their scalability in real-world applications.

To address these shortcomings, future research should consider hybrid approaches that combine the strengths of multiple models. For example, linear models like ARIMAX could provide a foundation, with machine learning models like GBM or LSTM fine-tuning predictions in non-linear regions. Incorporating external data sources, such as sentiment analysis, macroeconomic indicators, or intraday stock movements, could also enhance predictive power. Additionally, ensemble methods that aggregate predictions across multiple models may help strike a balance between accuracy, interpretability, and robustness.

The insights from this study also offer significant practical applications, particularly for investors and portfolio managers. Understanding the influence of risk factors—such as VIX, trading volume, and macroeconomic indicators—on future stock returns can help optimize investment strategies. For example, ARIMAX's ability to incorporate external predictors makes it suitable for evaluating the impact of specific market events on stock performance, while LSTM’s superior accuracy can support high-frequency trading strategies that rely on capturing nonlinear patterns. These findings emphasize the importance of evaluating risk indicators in financial forecasting, providing a framework for more informed, risk-aware decision-making.

In summary, this study highlights the importance of tailoring models to specific forecasting objectives. While advanced models demonstrate superior accuracy, simpler models offer significant value when transparency and computational efficiency are prioritized. Bridging this trade-off will be critical for advancing financial forecasting methodologies and improving investment strategies through risk-aware optimization.

6. Acknowledgement

We would like to express our sincere gratitude to Professor David Ye for his insightful guidance, constructive feedback, and continuous support, which were instrumental in shaping the direction and execution of this research. We would like to acknowledge the Duke MIDS QFC track for providing the infrastructure and resources necessary to undertake this study. We would also like to thank our peers for their valuable critiques and discussions, which enhanced the rigor and depth of our analysis. Lastly, we would like to recognize the support and encouragement of our families and colleagues, whose contributions made the completion of this work possible.

References

1. Bildirici, M. E., & Ozaksoy Sonustun, F. (2018). The effects of oil and gold prices on oil-exporting countries. Energy Strategy Reviews, 22, 290–302. https://doi.org/10.1016/j.esr.2018.10.004
2. Buyske, S. (2021). Bayesian data analysis, Statistics 365 course materials. Rutgers University, Department of Statistics.
3. Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and Regression Trees. Wadsworth International Group.
4. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.
5. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian data analysis (3rd ed.). CRC Press.
6. Kenton, W. (n.d.). Liquidity risk. Investopedia. Retrieved December 7, 2024, from https://www.investopedia.com/terms/l/liquidityrisk.asp
7. Kuepper, J. (2024, August 7). VIX. Investopedia. Retrieved December 7, 2024, from https://www.investopedia.com/terms/v/vix.asp
8. Martin, G. M., Frazier, D. T., Maneesoonthorn, W., Loaiza-Maya, R., Huber, F., Koop, G., inance: A modern review. International Journal of Forecasting, 40(2), 811–839. https://doi.org/10.1016/j.ijforecast.2023.05.002
9. Corwin, S. A., & Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance, 67*(2), 719-760. https://doi.org/10.1111/j.1540-6261.2012.01729.x

Non-technical Research - Credit Suisse: Unveiling the Risks and Collapse Behind a Global Financial Giant

Abstract

The unprecedented acquisition of Credit Suisse by UBS in 2023 marked a pivotal moment in banking history, fundamentally reshaping Switzerland's financial landscape and sending ripples through the global banking sector. This research examines the complex chain of events and systemic issues that led to this historic merger, with particular emphasis on risk management failures, regulatory oversight gaps, and broader implications for financial institution resilience. Through detailed analysis of the Greensill Capital and Archegos Capital crises, along with other contributing factors, this study reveals how the combination of inadequate risk management, reputational damage, and liquidity issues can precipitate the downfall of even the most established financial institutions. The findings underscore the critical importance of maintaining robust risk management frameworks and proactive governance structures in preserving financial institution stability.

1. Background

Credit Suisse's journey from its founding in 1856 to its eventual acquisition by UBS in 2023 spans over 165 years of banking history. Originally established to support Swiss infrastructure development, the bank gradually expanded its operations to become a global financial powerhouse. A significant milestone in this expansion was the acquisition of First Boston in 1988, which solidified Credit Suisse's position in global investment banking and demonstrated its ambitions for international growth. The bank's operations encompassed various business segments, including investment banking, wealth management, asset management, and private banking, serving a diverse international clientele.

However, the bank's latter years were marked by a series of challenges and controversies. The 2021 crises involving Greensill Capital and Archegos Capital Management proved particularly damaging, exposing significant weaknesses in the bank's risk management practices and oversight mechanisms. These incidents, combined with other reputational issues and operational challenges, began to erode market confidence in the institution. The situation deteriorated further as the bank struggled with liquidity concerns and mounting losses, ultimately leading to its acquisition by UBS in a historic deal orchestrated by Swiss regulators to prevent a broader financial crisis.

2. Statement of Problems

The research addresses several interconnected critical problems that contributed to Credit Suisse's downfall. First, the study examines how Credit Suisse's risk management practices failed to prevent or adequately respond to the Greensill and Archegos crises, despite the bank's sophisticated risk management infrastructure. This includes analyzing the specific breakdowns in risk assessment, monitoring, and mitigation that allowed these situations to develop.

Second, the research investigates the role of liquidity risk in the bank's ultimate collapse. This encompasses both market liquidity risk, affecting the bank's ability to sell assets without significant price impact, and funding liquidity risk, relating to the bank's capacity to meet its financial obligations. The study examines how these risks interacted and amplified each other during the crisis period.

Third, the research explores how governance and oversight inadequacies contributed to the crisis. This includes examining internal control mechanisms, decision-making processes, and the effectiveness of board oversight in managing the bank's risk exposure and strategic direction.

Finally, the study analyzes the systemic implications that necessitated UBS's intervention, including the potential consequences for the Swiss banking sector and the global financial system had the acquisition not occurred.

3. Key Findings

3.1 Risk Management Failures

The investigation revealed significant failures in Credit Suisse's risk management practices that ultimately contributed to its downfall. A primary issue was the bank's over-reliance on third-party risk mitigation, particularly evident in the Greensill Capital case. The bank placed excessive trust in Greensill's business model and insurance coverage without conducting adequate independent risk assessment. This was compounded by inadequate monitoring of leverage, most notably demonstrated in the Archegos Capital crisis, where the bank failed to properly assess and limit its exposure to highly leveraged positions.

Poor due diligence practices were consistently evident across multiple cases. In the Greensill situation, Credit Suisse failed to thoroughly investigate the underlying risks of supply chain finance arrangements and the concentration of exposure to single counterparties. The bank's high concentration of risky clients further exacerbated these issues, showing a systematic failure to diversify risk exposure appropriately.

3.2 Liquidity Crisis Development

The study identified a complex interplay of liquidity risks that ultimately led to Credit Suisse's crisis. Market liquidity risk manifested through the rapid devaluation of assets, particularly during the Archegos crisis when the bank needed to liquidate positions quickly, leading to substantial losses. This was exacerbated by funding liquidity risk as market confidence eroded, making it increasingly difficult for the bank to secure stable funding sources.

The situation was further complicated by significant cash flow mismatches and operational pressures. As market confidence declined, the bank faced increasing difficulty in managing its day-to-day operations and maintaining adequate liquidity buffers. This created a downward spiral effect on the bank's reputation, where each new revelation of problems led to further erosion of market confidence, making it increasingly difficult to maintain stable funding sources.

3.3 Governance and Oversight Issues

The analysis exposed critical weaknesses in Credit Suisse's governance structure and oversight mechanisms. Weak internal controls and communication channels prevented early warning signs from reaching appropriate decision-makers in time for effective intervention. The bank's risk escalation procedures proved inadequate, with many warning signals either being ignored or not properly communicated to senior management and the board.

A persistent focus on short-term gains over long-term stability emerged as a recurring theme in decision-making processes. This was evident in the bank's aggressive pursuit of high-risk clients and complex financial products without adequate consideration of long-term risks. The bank's response to regulatory concerns was often insufficient, suggesting a broader governance failure in managing relationships with supervisory authorities.

4. Shortcomings and Future Research Directions

This research faced several limitations that present opportunities for future investigation. Access to internal documentation and decision-making processes was necessarily limited, restricting our ability to fully understand the internal dynamics that contributed to the crisis. Future research with greater access to internal records could provide deeper insights into the organizational factors that led to risk management failures.

The study also identified a need for deeper analysis of cross-border regulatory coordination in managing international banking crises. The Credit Suisse case highlighted challenges in coordinating regulatory responses across multiple jurisdictions, suggesting an important area for future research. Additionally, there is a significant opportunity for comparative analysis with other banking crises to identify common patterns and potential preventive measures.

5. Conclusions

The collapse of Credit Suisse and its subsequent acquisition by UBS provides crucial lessons for the financial sector. First and foremost, it demonstrates that even the most established financial institutions can fail when risk management, governance, and transparency are compromised. The case underscores the critical importance of maintaining robust risk management frameworks that can adapt to evolving market conditions and new types of financial risks.

The bank's downfall serves as a powerful reminder of the interconnected nature of different types of risks in modern banking. Market risk, credit risk, and reputational risk can quickly transform into liquidity risk, creating a devastating feedback loop that can bring down even the largest institutions. This highlights the need for holistic risk management approaches that consider the interplay between different risk categories.

For the broader financial sector, the case offers important guidance on building resilience and maintaining stability. It emphasizes the importance of balancing profit-seeking activities with prudent risk management and the need to maintain stakeholder trust through transparent operations and strong governance. The Credit Suisse case will likely influence regulatory approaches and risk management practices in the banking sector for years to come.