TRADING & QUANTITATIVE RESEARCH REPORT

Intraday SEK Dynamics

Evidence from Cross-Asset Market Signals

In collaboration with:

Analysts: Malcolm Cederbaum, Ludvig Cederlund, and Sofie Melander
Supervisor: Simon Österberg

Introduction & Theory

Introduction

Intraday foreign exchange (FX) markets are among the most liquid and information dense financial markets, making short horizon exchange rate prediction particularly challenging. According to economic theory, Sweden is a small, open economy with strong trade and financial linkages. Consequently, the Swedish Krona (SEK) is often seen as a "cyclical" currency, meaning it is highly sensitive to global growth expectations and shifts in investor sentiment. It is well documented that the SEK typically strengthens when export demand increases or global risk appetite is high, while periods of global uncertainty lead to depreciation pressure as investors reduce their exposure to smaller currencies.

Since these broad macroeconomic factors are known to influence the SEK, the core objective of this project is to detect and apply these effects on an intraday timescale. The motivation is to investigate whether the factors describing, and ultimately predicting, the behaviour of the SEK can be identified before they are fully priced into the market.

To capture these dynamics, this study utilizes a diverse set of variables, including regional FX currency pairs, equity indices, cryptocurrencies and global risk indicators. Furthermore, to navigate the noisy intraday environment, this research leverages a unique informational edge: customer buy and sell volumes provided by SEB. Combining public cross-asset sentiment indicators with private institutional order flow creates a comprehensive framework for understanding intraday price discovery.

This study investigates whether it's possible to predict if EUR/SEK price will rise or fall at the 10-minute horizon. Rather than forecasting the exact return magnitude, the task is framed as a binary classification problem using the eXtreme Gradient Boosting (XGBoost) machine learning model.

Theory

Foreign exchange (FX) trading, also known as currency trading, is the conversion of one currency into another and is the largest financial market in the world by trading volume. Currencies are traded daily by banks, companies as well as individuals. Bank for International Settlements (BIS) estimated the average over-the counter turnover per day in April 2025 to be $9.6 trillion USD. This is an increase of 28% compared to 2022 [1]. The FX market is open 24-hours a day during the week but is closed on weekends.

Currencies are always traded in pairs, where the first currency is the base and the second is the quote currency. A currency pair that does not include the U.S. Dollar is referred to as a cross currency. Positions that are opened and closed within the same day are known as intraday trading.

The FX market has evolved significantly over the past decades. Initially dominated by telephone-based over-the-counter trading, the market has transformed into an electronic and global market with high-frequency and algorithmic trading. This evolution has increased market liquidity and speed of execution. The growth of electronic trading platforms and algorithmic strategies has also influenced intraday price patterns, introducing short-term correlations that can be analyzed for predictive modeling [4].

There are various factors that will affect the price movements. This could include actions or announcements from central banks, market news or shifts in investor sentiment. A common approach in intraday FX modeling is to incorporate proxies for global risk sentiment. Measures of market volatility, such as implied equity volatility indices, are widely used as indicators of risk-on and risk-off regimes. For risk-sensitive currencies, such as the Swedish krona (SEK) or the Norwegian krone (NOK), changes in global risk appetite can have a significant short-term impact on exchange rates [9].

Additionally, intraday FX data is characterized by a high level of noise and a low signal-to-noise ratio, making short-horizon forecasting particularly challenging. At such frequencies, price movements are strongly influenced by market microstructure elements such as order flow, liquidity provision, and bid–ask dynamics, while traditional macroeconomic relationships that are relevant at longer horizons tend to play a limited role. As a result, identifying predictive patterns in intraday exchange rate movements requires models that are flexible enough to capture complex, non-linear relationships [7].

When evaluating directional prediction models, it is essential to compare their performance against a simple benchmark. In a binary classification setting where the objective is to predict whether the exchange rate will move up or down, the natural baseline corresponds to a random classifier with equal probability for each outcome. Under this framework, a model that randomly predicts upward or downward movements with probability 0.5 achieves an expected accuracy of 50%.

This benchmark is particularly relevant in financial markets, where short-horizon price movements are often close to a random walk and therefore difficult to predict consistently. As a result, any predictive model must demonstrate performance above the 50% baseline to indicate the presence of exploitable information in the input features. The use of such naive benchmarks is standard practice in empirical asset pricing and financial machine learning studies, as it ensures that reported predictive performance reflects genuine signal rather than statistical noise [8, 6].

Given the non-linear and complex nature of intraday FX data, machine learning methods are often used for attempting to forecast exchange rate movements. One of these machine learning methods is Extreme Gradient Boosting, XGBoost. XGBoost can be implemented either as a regression or a classification model, depending on the definition of the target variable and the chosen loss function. In a regression setting, the model predicts a continuous outcome such as returns or price changes. In contrast, a classifier estimates the probability of discrete outcomes. In this study, XGBoost is implemented as a classifier, where the target variable is binary and represents the direction of the EUR/SEK in the upcoming 10 minutes.

XGBoost's algorithm is based on gradient-boosted decision trees and is designed to deliver high predictive performance while maintaining computational efficiency, particularly when working with large and complex datasets. A decision tree, which is the building block of XGBoost, is a model that makes predictions by repeatedly splitting the data into smaller homogeneous groups. Each split is chosen to reduce a loss function, meaning the tree tries to separate the data in a way that improves prediction accuracy. A single decision tree is easy to interpret but often too weak to achieve high performance. Boosting overcomes this limitation by creating many small, weak trees and combining their outputs to form a strong ensemble.

In gradient boosting, each decision tree is trained sequentially to correct the errors made by the previous ensemble. These errors are captured by the residuals, defined as the difference between the observed target values and the model's current predictions. At each boosting iteration, a new tree is fitted to these residuals, allowing the model to incrementally reduce the overall loss function. This residual-based learning process explains the structure illustrated in Figure 1. XGBoost includes penalties for the complexity of each tree, such as the number of leaves and the size of the leaf weights. By discouraging overly complex trees, the model becomes less likely to overfit and gains stronger generalization ability. This makes XGBoost well-suited for real-world problems where noise or limited data can otherwise lead to poor performance [5].

XGBoost has a number of parameters that can be tuned to improve the performance of the algorithm and to be able to control overfitting [10]. Some of the most important parameters are:

- *max_depth*: The maximum depth of one decision tree.
- *n_estimators*: The number of trees to be built.
- *learning_rate*: Controls the contribution of each new tree to the overall model by scaling the step size in the boosting process. A lower learning rate results in slower learning but can improve generalization when combined with a larger number of trees.

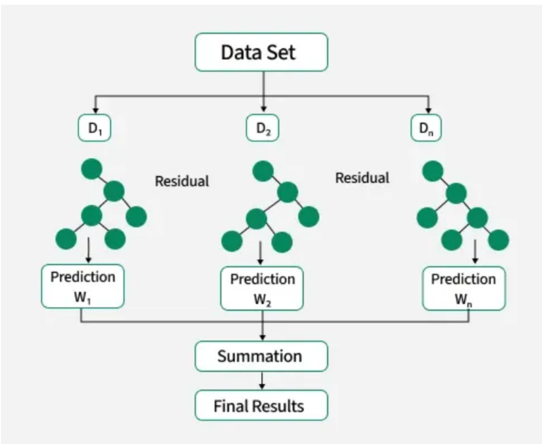


Figure 1: Visualization of XGBoost implementation.

Data & Method

Data

The analysis focuses on 10-minute EUR/SEK intraday directional movements over two periods: 2025-10-14 to 2025-11-25 and 2025-12-04 to 2026-01-30, restricted to the European trading session (09:00–17:30 CET). This window captures peak liquidity and aligns FX activity with other asset classes, ensuring the model captures the period of most active price discovery. Holiday-shortened trading days were excluded.

The decision to use two distinct non-overlapping time frames was due to both data availability constraints and methodological considerations. The initial period, combined with the first segment of the second period, formed the primary development dataset. This primary development dataset was used for the training (80%) and validation (20%) phases (hyperparameter tuning and feature selection). The remaining final portion of the second period was strictly reserved as an out-of-sample test set. By aggregating the periods into one large development set, the model gained access to a broader range of market behaviors, improving its ability to capture underlying intraday patterns and reducing the risk of overfitting. Simultaneously, holding out the final segment strictly for testing ensures that the finalized model is evaluated under unseen market conditions. Ultimately, this structure provides a robust estimation of the model's true predictive capabilities and significantly enhances the reliability of the final results.

To capture the complex dynamics driving the Swedish krona, explanatory variables were gathered from multiple asset classes and data providers, as summarized in Table 1. The selected asset classes have different impacts on the Swedish krona and can be categorized as follows:

Table 1: Overview of selected features and data sources.

Feature	Ticker	Source
EUR/SEK (Target)	EURSEK=X	Yahoo Finance
NOK/SEK	NOKSEK=X	Yahoo Finance
USD/JPY	USDJPY=X	Yahoo Finance
EUR/JPY	EURJPY=X	Yahoo Finance
USD/SEK	USDSEK=X	Yahoo Finance
DXY Index	DX-Y.NYB	Yahoo Finance
OMX Index	ˆOMX	Yahoo Finance
SSAB B	SSAB-B.ST	Yahoo Finance
SX5E Index	ˆSTOXX5E	Yahoo Finance
EUR/USD Volatility	EURUSDV1M	Bloomberg
BCOM Index	BCOM	Bloomberg
Bitcoin	BTCUSDT	Binance
Solana	SOLUSDT	Binance
SEB Buy Flow	-	SEB
SEB Sell Flow	-	SEB

- **Currency & FX Volatility:** The currencies we chose to look at included regional pairs, such as NOK/SEK, as well as global currency pairs like USD/JPY and EUR/JPY. We also included global indicators like the DXY (U.S. Dollar Index) and EURUSDV1M (Bloomberg 1-month volatility index). Since they both rely on the USD they add the perspective of market uncertainty and global risk sentiment.
- **Equity Indices & Commodities:** We utilize the OMX30 and SX5E (Euro Stoxx 50) to represent local Swedish and broader European equity markets. These indices serve as indicators of economic direction in the EU region. The overall commodity index (represented by BCOM, Bloomberg Commodity Index) is included to track broader commodity trends. Commodity prices often reflect shifts in global demand and inflation, which historically has had an impact on the valuation of the Swedish krona. Additionally, SSAB B is included to represent major Swedish export sectors, which are heavily influenced by international trade dynamics and currency fluctuations.
- **Cryptocurrencies:** While traditional assets are typically evaluated by price movements, this model specifically incorporates the trading volumes of Bitcoin and Solana. As high-beta speculative assets, cryptocurrencies respond rapidly to shifts in global risk sentiment. Due to continuous trading and the absence of regional market closures, cryptocurrency markets can incorporate global sentiment shocks almost instantaneously [3]. Consequently, cryptocurrency trading volumes can serve as real-time proxies for investor risk appetite, providing explanatory power for short-horizon movements in risk-sensitive exchange rates such as EUR/SEK.
- **Liquidity & Flow Indicators:** The only privileged data in our model consists of customer data from SEB regarding their buy and sell volumes in EUR/SEK. By including actual transaction volumes, the model gains insight into market positioning and underlying liquidity demand not visible in other data sources.

Method

Data from different assets was collected from Bloomberg, Yahoo Finance and Binance. To sort out less informative features, the features were divided up in their respective asset classes and evaluated by a linear regression model. Feature entries were lagged across multiple time scales, as certain variables exhibit limited predictive power at the 10-minute frequency but stronger signals at longer horizons. From this analysis, a final list of base features was selected, as presented in Table 1 in the Data section. Together with the buy and sell volumes provided by SEB, these acted as the input variables for the model.

Following the initial feature selection, the next crucial step was feature engineering. Standardized functions were developed to process the raw data, such as calculating lagged values, rolling means and spreads, to extract stronger predictive signals from the market. To prevent data leakage across trading days, since this report only covers the regional opening hours, all rolling features and lags were calculated intraday, ensuring that the previous trading session's closing data did not artificially bleed into the next day's opening signals.

XGBoost was implemented as a gradient-boosted decision tree classifier estimating:

$$P(y_t = 1 \mid X_t),$$

where X_t represents the vector of lagged market features at time t. Rather than relying strictly on the default 0.5 classification boundary, the decision threshold was treated as a tunable hyperparameter. A predicted probability above this optimized threshold is interpreted as an upward EUR/SEK move, while a value below it indicates a downward move.

XGBoost was used to build the model using the dataset of the "best" predicting features alongside the engineered features. To facilitate the model development, the primary development dataset (as introduced in the Data section) was partitioned into a training and a validation set. During the iterative tuning phase, the model was continuously trained on the training set and evaluated on the validation set. This back and forth with finetuning parameters of the XGBoost model, adding and deducting engineered features, alongside adding and deducting features, led to the finished model. Throughout this tuning phase, the model's performance was rigorously monitored using three key evaluation metrics:

1. **Accuracy Score** to measure overall correctness.
2. **Classification Reports** (Precision, Recall, and F1-score) to measure performance across both upward and downward predictions.
3. **Feature Importance** (Gain) to measure relative predictive contribution of each variable.

The evaluation process eventually led to the finished, optimized model, which was subsequently retrained on the combined training and validation set and then evaluated on the unseen test data.

Figure 2 illustrates the comprehensive workflow of the methodology, from initial data collection and feature selection to the final validation of the XGBoost model.

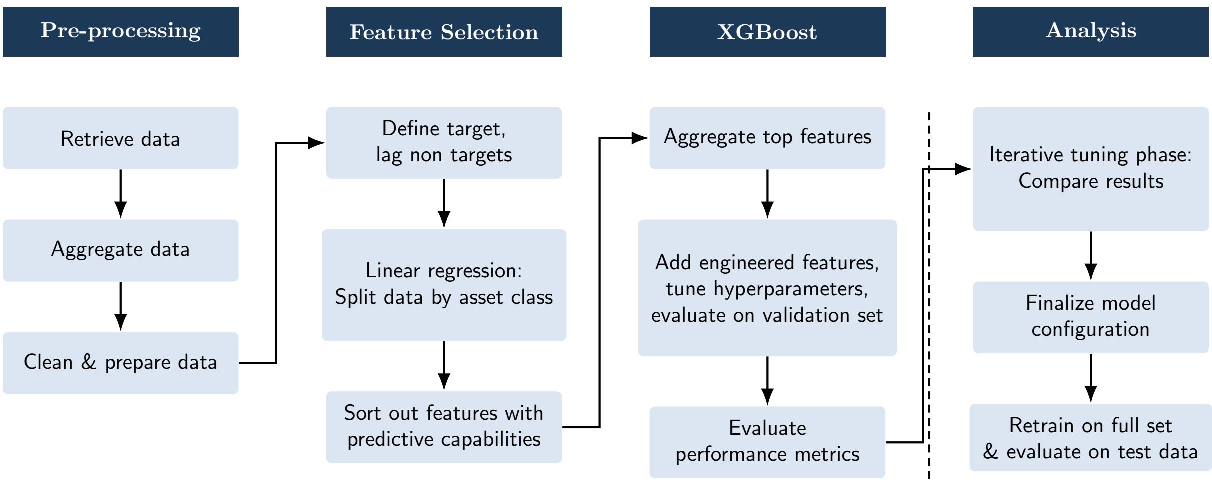


Figure 2: Flowchart of the report's method.

Results

The following section details the outcome of the XGBoost classification model. It is important to note that the results presented correspond only to the final, optimized model configuration derived from the iterative finetuning process described in the Method chapter. The final hyperparameter configuration for the model is presented in Table 2.

Table 2: Final hyperparameter configuration for the model.

Hyperparameter	Value
n_estimators	100
max_depth	3
learning_rate	0.05
threshold	0.53

In addition to the previously selected base features, the optimized model contained the following engineered features:

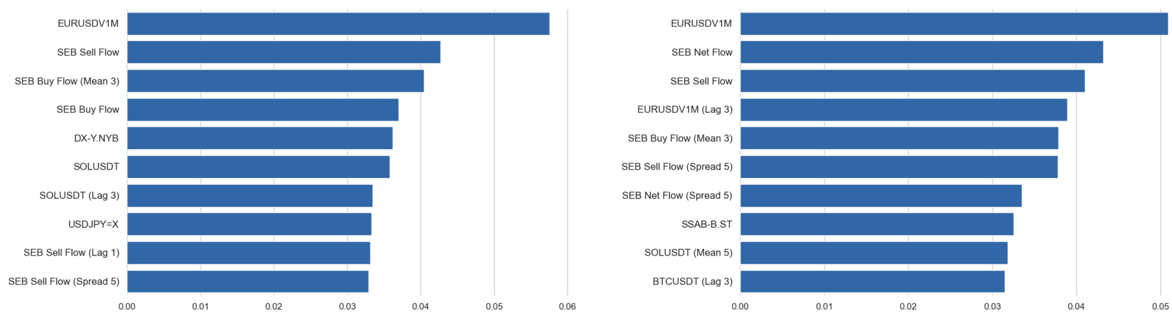
- **EUR/SEK (Target):** Lags of 1, 3 and 5 frames
- **EURUSDV1M, BTCUSDT, SOLUSDT:** Lags of 3 frames and rolling means of 5 frames
- **SEB Buy, Sell & Net Flow:** Lags of 1 and 3 frames, rolling means of 3 and 5 frames and rolling spreads of 5 frames

The presentation of the results is centered around two key dimensions: the predictive power of the input variables (Feature Importance) and the model's ability to correctly predict the direction on unseen data (Classification Performance). Within each of these dimensions, the model is evaluated on both the validation set and the test set, providing a clear comparison between the model's optimized performance and its true predictive power on unseen data, effectively assessing the risk of overfitting.

Feature Importance

This subsection outlines which features the model identified as most significant for predicting the 10-minute EUR/SEK directional changes. Feature importance is measured by the relative contribution (gain) of the corresponding feature to the model.

Figure 3 illustrates the 10 most important features derived from the model iterations. Figure 3a shows the importance derived from the model trained on the 80% training split, while Figure 3b shows the corresponding feature importance from the model after it was retrained on the combined training and validation set.



(a) Validation set

(b) Test set

Figure 3: Comparison of the Top 10 Feature Importances (gain) between the initial model and the finalized model.

Comparing the two distributions reveals both consistencies and notable shifts in feature dependence. The 1-month EUR/USD volatility index (EURUSDV1M) remains the single most critical predictor across both models. Furthermore, proprietary SEB flow data consistently dominates both top 10 lists. However, the specific engineered variants shift between the iterations; most notably, "SEB Net Flow" emerges as the second most important feature in the finalized model, despite not appearing in the initial model's top 10. Conversely, global macro indicators such as the U.S. Dollar Index (DX-Y.NYB) and USD/JPY, which were present in the initial model, fall out of the top 10 in the finalized model. Instead, they are replaced by features such as the SSAB B stock and lagged Bitcoin volume.

Classification Performance

To assess the model's generalization capabilities, its predictive performance is evaluated using accuracy scores, confusion matrices and classification reports. These metrics provide insight into the distribution of true positive and false positive predictions across both the validation and test set.

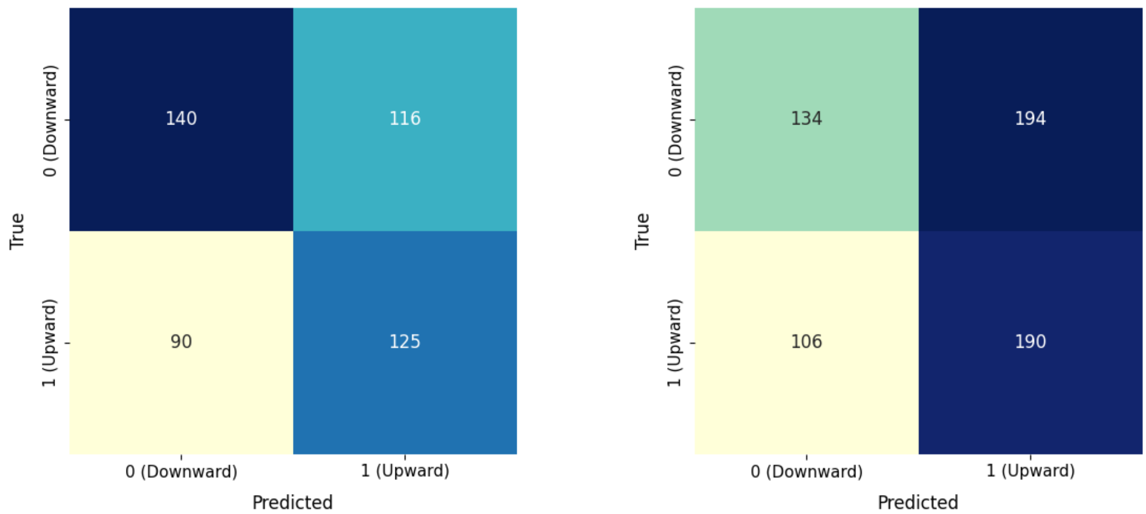
On the initial validation set (471 data points), the model achieved an overall accuracy of 56.26%. It demonstrated a relatively balanced performance, with a precision of 0.61 for downward movements (Class 0) and 0.52 for upward movements (Class 1). When subsequently evaluated on the test set (624 data points), the accuracy adjusted to 51.92%. The recall for upward movements (Class 1) increased to 0.64, but at the expense of recall of downward movements (Class 0), which dropped to 0.41. The detailed classification metrics for both datasets are outlined in Tables 3 and 4. Furthermore, the corresponding distributions of true and false predictions for both the validation and test set are illustrated side-by-side in Figure 4.

Table 3: Classification Report for the validation set.

Class	Precision	Recall	F1-score	Support
0 (Downward)	0.61	0.55	0.58	256
1 (Upward)	0.52	0.58	0.55	215

Table 4: Classification Report for the test set.

Class	Precision	Recall	F1-score	Support
0 (Downward)	0.56	0.41	0.47	328
1 (Upward)	0.49	0.64	0.56	296



(a) Validation set

(b) Test set

Figure 4: Confusion matrices for the validation (471 data points) and test (624 data points) set.

Analysis & Conclusion

Analysis
Our model did not prove to be the "money-machine" we obviously hoped for, but we would argue that the results found are significant and highly usable. In line with economic theory, modeling the rise of an asset's price should be harder than predicting its fall, due to the phenomenon of asymmetric volatility. The rise of an asset is often driven by fundamentals and is gradual, unlike decline which often happens fast (panic, actors cut their losses etc.) [2]. As aforementioned, the final model had a greater precision predicting downward motion on both the initial validation set and the test set.

However, it is important to interpret these results relative to the baseline benchmark for directional prediction. In a binary classification setting, a naive random classifier achieves an expected accuracy of 50%. The model's test accuracy of 51.92% therefore indicates a modest but non-trivial improvement over random guessing, suggesting that some predictive signal is present in the selected features. However, it is important to note that there are still notable limitations within our result. When evaluating the model on the test set, the recall for downward movements drops significantly. This resulted in an F1-score for downward movements that was lower than that for upward movements, signaling that we might have overfitted our model to our validation set. Before analyzing the impact of the selected and engineered features, it is worth highlighting an aspect of the predictions that could indicate the model's reliability despite this.

Through our collaboration with SEB, we were granted access to their proprietary EUR/SEK order flow data. As demonstrated in the Feature Importance section, both the initially trained model and the final retrained model assign significant predictive weight to this data source. As this data is exclusive to SEB, it likely captures a unique segment of market sentiment reflecting their specific client base. When developing quantitative prediction models, unique proprietary signals provide an informational edge against competitors in the market. With that being said, if the model heavily relies on one source it is going to be more prone to its limitations. While we utilize the exclusivity, we are also exposed to disturbances, such as variations in the trading frequency, and the specific exchange rates quoted by SEB. Furthermore, since not all firms trade through SEB, relying solely on this flow creates a blind spot. There are plenty of other market signals beyond buy and sell volumes, and if we do not capture on these, it could explain the drop in the test performance. However, we do still have a good precision when the model detects certain patterns (most probably tied to the SEB data). Thus we cannot reject the hypothesis that SEB's order flow data is being a valuable predictor of the EUR/SEK, making it highly interesting to further look into for a future project.

To be able to make rigid claims about the SEB order flow data being predictive, we would have to overcome one of our major obstacles: our lack of data. It is a well established principle within machine learning that more data equals better, or more interesting, results. It should be stated that in our case, the lack of data created difficulties in validating any results as well as evaluating if a result is overfitted or not. As it is not advised to continuously test back and forth between the validation and test set, we decided to make do with what we had. It is safe to say that more data would have aided this report tremendously, especially when utilizing XGBoost which is notoriously hungry for data. One way to solve this issue would be to either look at shorter time spans (this would gain more data points, but also redefine the scope of our work) or utilizing a bigger data source, which should be available in house at most trading firms.

While analyzing the chosen features for our model, some questions emerged. In our search for niche prediction signals, had we in some cases strayed from our initial path? The decision to include crypto assets was unconventional, and after the initial feature selection, only the largest crypto assets prevailed (Bitcoin and Solana). At first glance, it might seem strange that these would influence the price of EUR/SEK. However, while testing our model, dropping the crypto assets resulted in a notable loss of accuracy. Thus, despite an obvious causal link, these assets seemed to hold predictive value. One could argue that the appetite for cryptocurrencies in general says something about the current risk aversion at the market, which we know is something that affects the EUR/SEK since the SEK is a smaller, risk-sensitive currency.

Another unconventional feature is the Bloomberg Commodity Index (BCOM). Originating from the hypothesis that global copper prices could impact the exchange rates, since copper is a large Swedish export, the BCOM was included as a broader indicator of the commodity market. Sweden, being very restrictive regarding mining, might not affect this index directly, but it did seem to be useful when predicting the EUR/SEK. It is possible that the commodity trading influences the FX market since commodities are often more stable assets (which is also true for larger FX). Furthermore, it is arguable that commodities could also carry a sentiment of the market's current risk appetite since "hard assets", such as precious metals, are considered safe investments.

Including the 1-month EUR/USD implied volatility index (EURUSDV1M) adds the dimension of projected volatility for the currency pair. As the U.S. Dollar is widely considered a primary safe-haven currency, this metric is likely heavily tied to the current market risk sentiment.

As a final point of discussion regarding our feature selection, we must address the cross-currency pairs. While the EUR/SEK is undoubtedly influenced by most, if not all, active trading rates in the FX market, these relationships do not necessarily possess predictive capabilities. In fact, our final model does not include a single engineered feature based on cross-currencies. During the validation phase, features derived from cross-currencies initially showed promising results, however, incorporating too many of them led to overfitting. We still believe that some predictive behavior is present, since each feature was initially based on the lagged regression model, but we were unable to isolate these signals enough to avoid adding noise.

Following discussions with our supervisor, we chose a simple and foolproof approach to feature engineering to prevent data leakage. Feature engineering is crucial for highlighting underlying patterns for the XGBoost algorithm, balancing signal extraction with the risk of overfitting to the training data. Consequently, we only utilized lags, moving averages (mean), spreads and min/max values while testing our model. Each of them works on an adjustable lookback window, which decides the number of previous time steps considered. Ultimately, all engineered feature types except min/max values were retained in the final model. Exploring more complex feature engineering methods and the resulting impact on model performance would be an interesting approach to improve upon this work. Since feature engineering lacks a strict framework and requires continuous testing, it functions much like our proprietary SEB flow data: a key area where the model can achieve a unique quantitative edge.

While the methodology was constrained by data availability, we took strict precautions to not obtain false results. A common pitfall in machine learning is the iterative tuning of a model against the test set, overfitting the model rather than finding actual evidence of a signal. To avoid this, our test set was strictly reserved for final evaluation.

Our final results were similar to other top-performing iterations, suggesting that the predictive signal, while present, is sensitive to sample variance. This highlights the difficulty of isolating an 'optimal' configuration with a limited dataset. Future research should therefore prioritize a larger data volume or higher-frequency timeframes to better distinguish genuine market signals from statistical noise.

Conclusion

This study examined the short-term predictability of EUR/SEK intraday movements using a combination of public market indicators and proprietary SEB order flow data. The primary goal was to assess whether these features could provide an informational edge for directional forecasting at a 10-minute horizon.

Results indicate that SEB buy and sell flows contributed the most predictive value, suggesting that proprietary order data can capture unique aspects of market sentiment. Some public features, such as 1-month EUR/USD implied volatility, Bitcoin and Solana trading volumes and the Bloomberg Commodity Index, provided additional, though limited, predictive information. Conventional cross-currency pairs and equities showed minimal influence on the model at this intraday horizon, highlighting the challenge of extracting meaningful signals from traditional macroeconomic or fundamental indicators in noisy, short-term FX data.

Overall predictive performance remained modest, with test accuracy around 52%, and downward movements were slightly easier to forecast. These results reflect both the inherent difficulty of short-horizon FX prediction and the limitations of the dataset in terms of size, coverage, and noise. While SEB flow data appears informative, the model's reliance on a single proprietary source exposes it to potential concentration risk and limits generalization.

Future research could focus on expanding the dataset over longer periods or including additional participants to enhance robustness. Incorporating richer microstructure indicators, alternative sentiment proxies, and models designed to capture temporal dependencies, such as recurrent or attention-based neural networks—may improve predictive capability. Further exploration of regime-aware modeling and temporal stability under high volatility could also shed light on the reliability of short-term signals.

In conclusion, while high accuracy remains elusive in intraday FX forecasting, this study demonstrates that proprietary order flow data combined with selected market indicators can identify patterns beyond random chance. The findings provide a foundation for more data-intensive and methodologically advanced investigations into microstructure-informed, sentiment-driven forecasting.

References

References

[1] Bank for International Settlements. Otc foreign exchange turnover in april 2025. https://www.bis.org/statistics/rpfx25_fx.htm, 2025.

[2] Geert Bekaert and Guojun Wu. Asymmetric volatility and risk in equity markets. *The Review of Financial Studies*, 13(1), 2000. doi: 10.1093/rfs/13.1.1.

[3] Barbara Bedowska-Sójka, Joanna Górka, Danial Hemmings, and Adam Zaremba. Uncertainty and cryptocurrency returns: A lesson from turbulent times. *International Review of Financial Analysis*, 94, 2024. doi: 10.1016/j.irfa.2024.103330.

[4] Alain Chaboud, Dagfinn Rime, and Vladyslav Sushko. The foreign exchange market. In Refet S. Gürkaynak and Jonathan H. Wright, editors, *Research Handbook of Financial Markets*. Edward Elgar Publishing, 2023.

[5] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794, 2016. doi: 10.1145/2939672.2939785.

[6] Campbell R. Harvey, Yan Liu, and Heqing Zhu. ... and the cross-section of expected returns. *The Review of Financial Studies*, 29, 2016. doi: 10.1093/rfs/hhv059.

[7] Joel Hasbrouck. *Empirical Market Microstructure*. Oxford University Press, 2007.

[8] Marcos López de Prado. *Advances in Financial Machine Learning*. Wiley, 2018.

[9] Robert E. Whaley. Understanding the vix. *The Journal of Portfolio Management*, 35, 2009. doi: 10.3905/JPM.2009.35.3.098.

[10] XGBoost Developers. Notes on parameter tuning - xgboost documentation. https://xgboost.readthedocs.io/en/stable/tutorials/param_tuning.html, 2025.

Disclaimer

Disclaimer

These analyses, documents, and any other information originating from LINC Research & Analysis (henceforth "LINC R&A") are created for information purposes only, for general dissemination, and are not intended to be advisory. The information in the analysis is based on sources, data, and persons which LINC R&A believes to be reliable. LINC R&A can never guarantee the accuracy of the information. The forward-looking information found in this analysis is based on assumptions about the future, and is therefore uncertain by nature, and using information found in the analysis should therefore be done with care. Furthermore, LINC R&A can never guarantee that the projections and forward-looking statements will be fulfilled to any extent. This means that any investment decisions based on information from LINC R&A, any employee or person related to LINC R&A, are to be regarded to be made independently by the investor. These analyses, documents, and any other information derived from LINC R&A are intended to be one of several tools involved in investment decisions regarding all forms of investments regardless of the type of investment involved. Investors are urged to supplement with additional relevant data and information, as well as consult a financial adviser before any investment decision. LINC R&A disclaims all liability for any loss or damage of any kind that may be based on the use of analyses, documents, and any other information derived from LINC R&A.

Conflicts of Interest and Impartiality

To ensure LINC R&A's independence, LINC R&A has established compliance rules for analysts. In addition, all analysts have signed an agreement in which they are required to report any conflicts of interest. These terms have been designed to ensure that COMMISSION DELEGATED REGULATION (EU) 2016/958 of 9 March 2016, supplementing Regulation (EU) No 596/2014 of the European Parliament and of the Council concerning regulatory technical standards for the technical arrangements for objective presentation of investment recommendations or other information recommending or suggesting an investment strategy and for disclosure of particular interests or indications of conflicts of interest.

Other