

# QF634 Report : Beyond Conventional Charts: Signal Processing Indicators as Catalysts for Predicting Peaks/Troughs in Forex Trading

Andre LIM, LIU YiSi, Dylan LOO, WANG WenJie

December 2023

## 1 Introduction

### 1.1 Thesis Statement

The objective of our project is to answer a key question using ML methods: *Can non-traditional technical indicators (such as Fourier Transform, Kalman Filter, Hurst Exponent (also known as Signal Processing Indicators)) be used in conjunction with traditional technical indicators to boost the performance of ML models to predict buy / sell signal in forex markets?* Through this project we aim to answer this question, while experimenting with additional variables and features that were deemed to have potential to boost the performance of the final model.

In the final implementation, we used a trading time frame of 2013 to 2023. During this period, the value of JPY in USD terms dropped by about 41%, which made for an interesting challenge. Source code to our project can be found in our github repository: <https://github.com/andreignatius/quantresearch>.

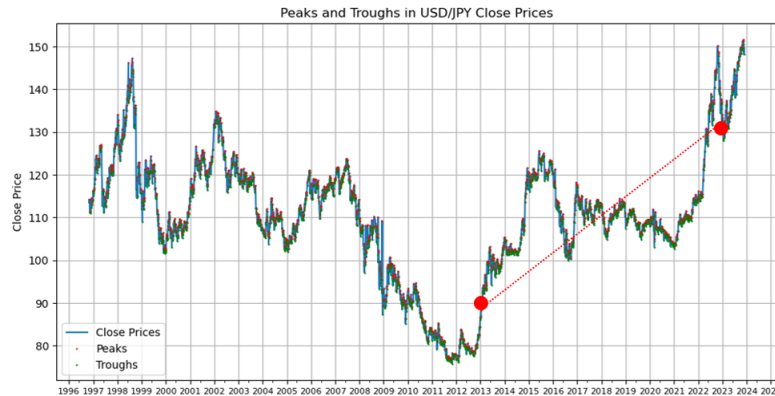


Figure 1: USD/JPY Exchange Rate From the mid 90s to the mid 20s

In order to generate our optimal peaks and troughs, we used the `scipy.signal.find_peaks` package in Python. At this juncture, it would be pertinent to highlight that in the actual implementation, we assess the performance of the model *not by its ability to perfectly predict* optimal peaks and troughs, but its ability to *consistently make good predictions* that are close enough to local market peaks and troughs such that each trade results in positive PnL, each holding period does not incur significant downside risk and that PnL is maximised.

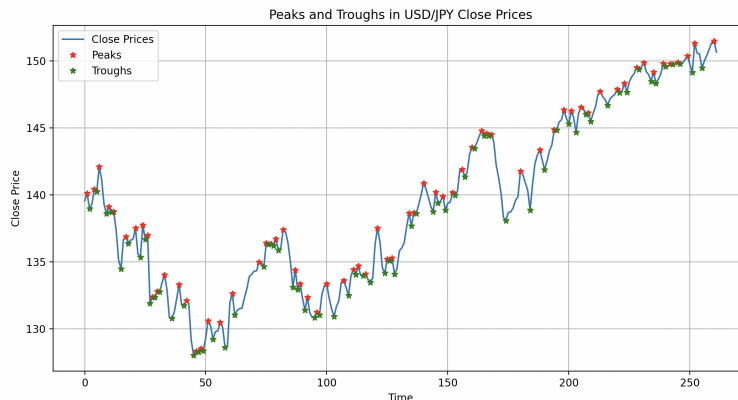


Figure 2: Peak/Trough Labeling using `scipy.signal.find_peaks` for USD/JPY Spot Rate from November 2022 to November 2023

Before proceeding, it also is worth disclaiming that the prices we used to determine optimal buy/hold/sell days were the 'Closing' price of the trading pair. Since the forex market is open 24 hours a day, precision of the definition of 'Close' price is important: The closing price in forex typically signifies the final price at the end of a trading session. For instance, if we were in the GMT time zone, the New York market close (around 5:00 PM EST or 10:00 PM GMT) is often considered the end of the trading day. The price at this point is recorded as the closing price for that particular day.

## 1.2 Theoretical Background

Hurst postulated that financial time series data could be modelled as a combination of discrete harmonic frequencies. This concept is part of the broader field of cyclical analysis, where financial market time series are considered to have a cyclical nature that can be decomposed into a series of sinusoidal components or waves, each with their own frequency, amplitude, and phase. These components correspond to different cycles that are superimposed upon each other, creating the observed market movements.

To help visualise the concept, let us take example of a musical chord, the A major chord with its triad A5, C#5, E5.

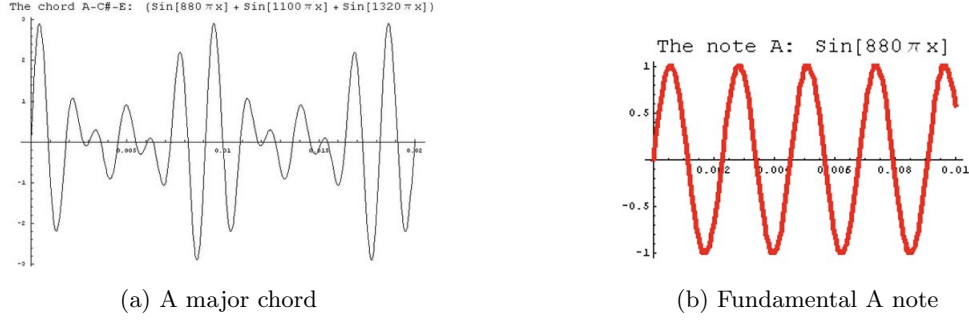


Figure 3: A major chord and fundamental A note

The A major chord wave is a combination of 3 other individual sinusoidal waves, note A at 880Hz, C $\sharp$  at 1100Hz, E at 1320Hz (values modified under different tuning system to reflect whole numbers). At the individual note level, the wave pattern appears clearly sinusoidal but when combined it looks noisier and more convoluted.

Let us examine with another visual example, how different notes add up to form a chord. This time with G major chord, with its constituent G3, B3, D3 notes (again the values are modified under different tuning system to reflect whole numbers).

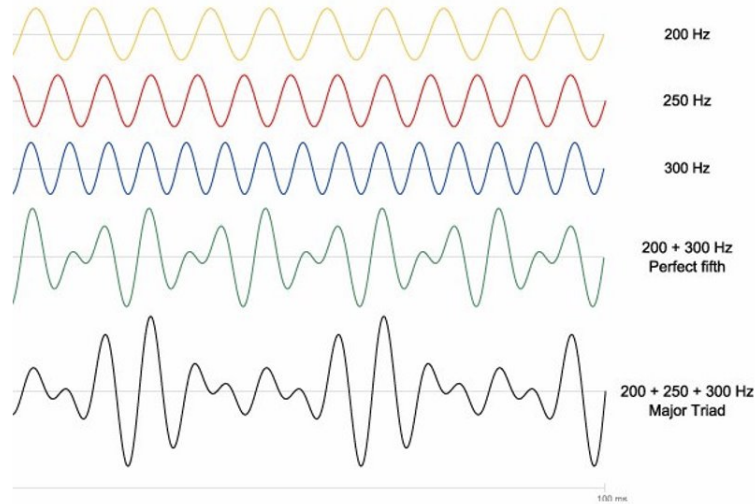


Figure 4: Visualization of fundamental notes combining together to form the G major musical chord

There are several observations and postulations in Hurst Cycle Theory

1. Cyclical Nature: Hurst observed that many systems exhibit cycles in a nested fashion, where smaller cycles exist within larger ones. This concept was originally applied to the analysis of river Nile floods but has since been extended to financial markets.

2. Predictability: The cycles are not perfectly periodic but have a degree of predictability. By identifying the cycle lengths, one can forecast the future movement of a time series to some extent.
3. Mean Reversion and Trend: The Hurst exponent ( $H$ ) is a key outcome of his work, which measures the tendency of a time series to revert to its mean ( $H < 0.5$ ), to trend ( $H > 0.5$ ), or to follow a random walk ( $H = 0.5$ ).
4. Market Analysis: In financial markets, Hurst Cycle Theory is used to identify potential market cycles for both short-term and long-term trading strategies. Traders and analysts look for patterns that repeat over time and use them to forecast market movements.
5. Fractal Nature: The theory also suggests that financial markets are fractal in nature. This means that market patterns can be self-similar at different scales, which is the basis for fractal market analysis.
6. Sentiment and Fundamentals: Hurst cycles can be influenced by market sentiment and fundamentals. For example, economic indicators, geopolitical events, and trader psychology can all influence the length and amplitude of cycles.

## 1.3 Motivations & Intuition

### 1.3.1 Why do forex movements matter (to an average person)?

In this globalized economy, many talented individuals are willing to leave their hometown to seek work opportunities abroad. However, a pressing concern remains for these migrant talent that their hard-earned income could be vulnerable to unfavourable forex movements, especially when they try to remit allowance back home or if they intend to return to their country of origin after a certain number of years.

To better illustrate this concern, let us take the example of a fictional character, Chad Nakamura. Chad Nakamura is an American citizen who found a job in Tokyo in 2013 and left the United States to start his career. As an Asian American with good Japanese language skills, it made sense to make the move.

Chad starts out in Tokyo with annual salary of 6,500,000 JPY which equates to 75,000 USD at the time of exchange rate on 1 Jan 2013 to work for Yoshi & Co. Over his time in the company, he got the following progression: he got annual increments of 2% per annum for good performance and 10% increment for promotions in 2016, 2018, 2021 and 2023. By 2023, his salary in JPY is now 65% higher than when he first start at Yoshi & Co. back in 2013.

He will learn however that the celebration is short-lived, as the fruits of his hard labour were been eroded by forex movements against his favour. Because in the same period, USD/JPY spot rate moved from 86.64 to 130.85.

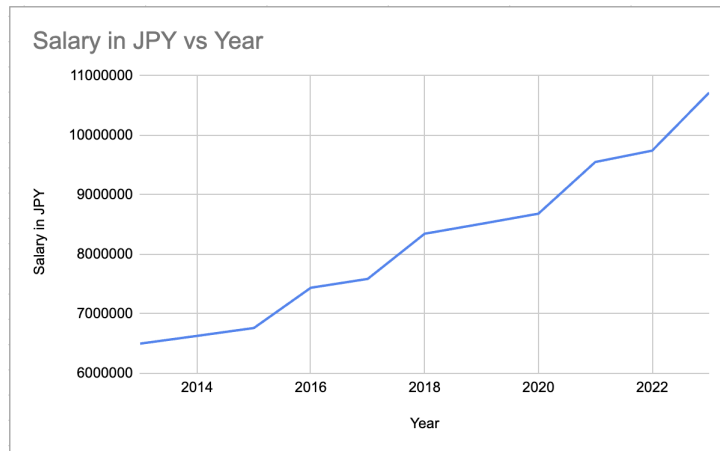


Figure 5: Chad's salary progression in JPY from 2013 to 2023

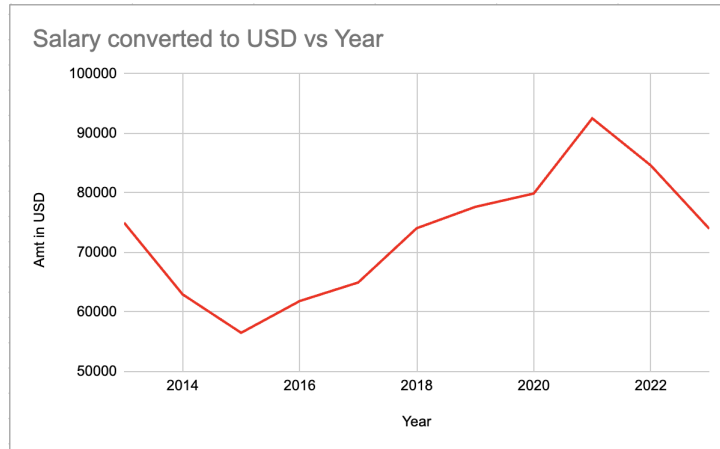


Figure 6: Chad's salary progression in USD from 2013 to 2023

In 2013 his salary of 6,500,000 JPY was worth 75,000 USD. However, his 2023 salary of 10,717,294 JPY is worth 73,984 USD. Throughout the sweat, tears and blood, Chad's salary when converted to his domestic currency in USD actually suffered -1.35% loss. This loss is extra devastating when one considers that it has not yet been adjusted for inflation.

Chad's story, along with many others like him, motivated us to create an automated trading strategy to hedge against forex movements and generate alpha.

### 1.3.2 Intuition from Theory (what can we do about it?)

Given what we have read from Hurst Cycle Theory and if we hold the postulate that financial time series data is as a combination of discrete harmonic frequencies, we can then align our research to explore various techniques in signal processing, such as Fourier transform, which breaks down time series into sinusoidal components. In the next chapter, we will explore the available signal processing techniques in greater depth.

## 2 Feature Generation

### 2.1 Signal Processing Indicators

#### 2.1.1 Fourier Transform

*What is it about?* Fourier transform is a mathematical technique that decomposes a complex signal into its constituent frequencies. The applied areas include signal processing, image analysis and sound decomposition etc. In financial industry, Fourier transform is utilised to convert the time series price data to dominant periodic trends.

$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt$ , where  $F(\omega)$  is the Fourier transformation of time series data  $f(t)$ . It tells us how much each frequency is presented in the time series data.

*Intuition as feature:* There are two domains in Fourier transformation: time-domain and frequency domain. In our case, Fourier transform takes time domain signal (Forex price) and converts it to frequency domain signal (the cyclical signal). The frequency domain signal shows the frequency components that dominate the price data.

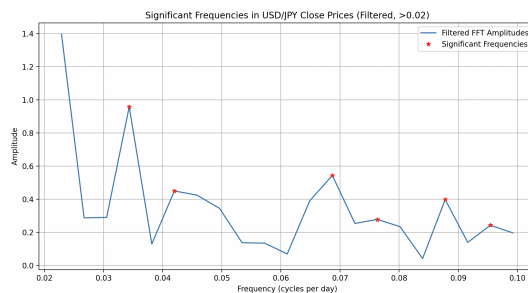


Figure 7: Dominant frequencies in USD/JPY Close Prices from Nov 2022 to Nov 2023

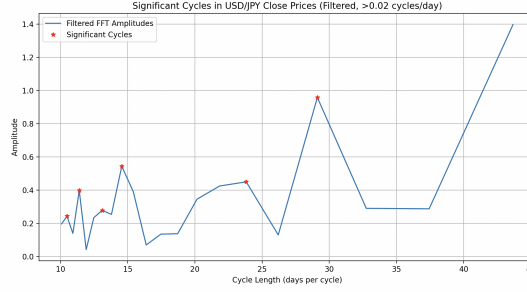


Figure 8: Dominant period cycles in USD/JPY Close Prices from Nov 2022 to Nov 2023

### 2.1.2 Kalman Filter

*What is it about?* Kalman Filter is a useful time series analysis algorithm in financial industry. Compared to Simple Moving Average, Kalman Filter puts more weight on more recent data on the calculation and takes statistical noises into consideration.

Kalman Filter works with two phases: the prediction phase and update phase. The prediction phase estimates the state at current timestamp. In the update phase, the innovation (difference between current prediction and current observation) is multiplied by Kalman gain (the weightage to most recent data) and combined with previous state estimation. This refined estimate is termed the posteriori state estimate.

The algorithm is a recursive process, thus can operate in real time. And it assumes the errors have a normal distribution.

*Intuition as feature:* Using Kalman Filter algorithm, we can derive the optimal state estimate of Forex spot rate combining prediction phase and update phase in a recursive process. This optimal state estimate gives us a smoothed and unbiased Forex spot rate which sifted out the unnecessary noise. The volatility (deviation from mean) can also be derived from Kalman Filter algorithm

### 2.1.3 Hurst Exponent

*What is it about?* Hurst Exponent is a mathematical measure of the long-term memory of a time series data that tells us the persistence of trend. First developed by Harold Edwin Hurst to solve the optimal dam size issue on Nile river, Hurst Exponent indicates us whether a time series is purely random, trending or mean reversion.

The equation for Hurst Exponential is:  $E \left[ \frac{R(n)}{S(n)} \right] = Cn^H$

where  $R(n)$  is the range of first  $n$  cumulative deviations from the mean,  $S(n)$  is the series of first  $n$  standard deviations,  $n$  is the span of the observation and  $C$  is a constant.

The most generic interpretation of Hurst Exponential is:

1. When  $H < 0.5$ , the time series is mean reverting, which indicates an anti-persistent behaviour
2. When  $H = 0.5$ , the time series exhibits Brownian motion with pure random walk
3. When  $H > 0.5$ , the time series is trending with a certain degree of predictability, which indicates a persistent behaviour

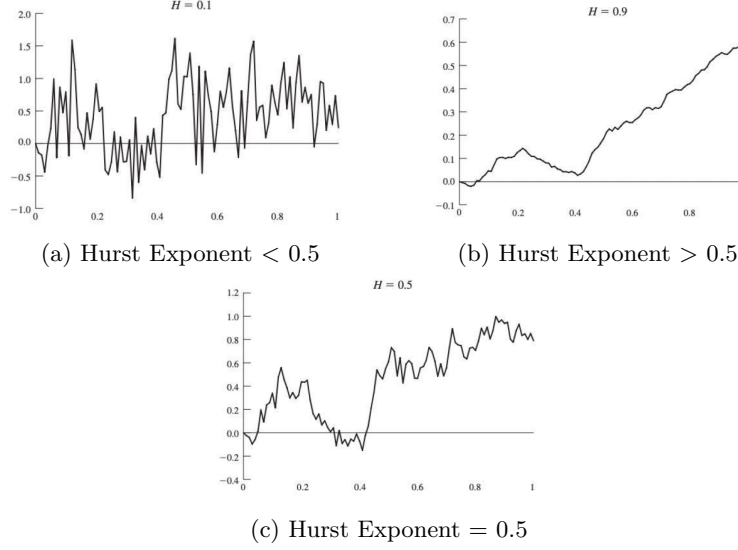


Figure 9: Simulation of fractional Brownian motion for different values of the Hurst exponent

*Intuition as feature:* Hurst Exponent tells us the domain of the current time series of Forex spot rate. We can use it to assess the essence of price movements. It gives us the information of whether the time series is currently mean-reverting or trending, and allows us to execute different strategies accordingly.

## 2.2 First and Second Order Derivatives of Indicators

*What is it about and why does it matter?* It should be noted that during this experiment, that identifying the best spot price to trade is not as important as **identifying the trend reversal turning point** of the spot price. The spot price data on its own has little predictive value without any reference to the date it belongs to but the date itself also has little value in predictive power as a date, once passed, will never repeat itself again.

*Intuition as feature:* Thus, the intuition came about that we were more interested in rate of change (first order derivative) or rate of rate of change (second order derivative) of the spot price in trying to time our trades in the spot market. For certain technical indicators like **Kalman Filter Estimate** and **Short / Long Moving Averages**, it made sense to compute the **first and second order derivative** values of those features and set those values as a new indicator.



## 2.3 Exogenous Economic Indicators

Spot forex exchange rate between JPY and USD is affected by several econometrics, such as the country's base rate, inflation, currency account balance, capital inflow and monetary policies in Japan and US. The correlation between exogenous variables and the USD/JPY forex rate is analysed in our model.

### 2.3.1 Uncovered Interest Rate Parity

Uncovered interest rate parity states that the expected change in spot exchange rate is equal to the interest rate differential between two countries. This theory does not involve the use of forward contracts and assumes the investor is risk-neutral.

The equation is:  $E[S_{t+1}] = S_t \times \frac{1+I_{Japan}}{1+I_{USA}}$

The expected spot USD/JPY forex exchange rate is determined by multiplying the current USD/JPY forex exchange rate with the ratio of the baseline interest rate in Japan to the baseline interest rate in the US.

### 2.3.2 Relative Purchasing Power Parity

The relative purchasing power parity (PPP) is based on the theory of law of one price, which states that without the transportation costs and other frictions, identical goods should have the same price in all countries. And relative PPP extends this concept and states that changes in exchange rates should exactly offset the price effects of any inflation differential between 2 countries.

The equation is:  $\frac{E[S_{t+1}]}{S_t} = \frac{1+InflationinJapan}{1+InflationinUSA}$

The expected change of spot USD/JPY forex exchange rate is equal to the ratio of inflation rate in Japan to inflation rate in US.

The relative PPP assumes long-term equilibrium, while in short run, other factors like speculative activities may affect the forex rate.

### 2.3.3 Current Account Balance

The current account provides a comprehensive overview of a country's trade, encompassing the exchange of goods, services, investment income, and unilateral transfers. It serves as a concise indicator of whether a nation is exporting more goods and services than it imports, thus signaling either a current account surplus or deficit.

Current account deficits lead to a depreciation of domestic currency because:

- Flow supply/demand mechanism: current account deficits in a country increase the supply of that currency in the currency markets, it puts downward pressure on the exchange value of the currency.
- Portfolio balance mechanism: countries with current account surpluses usually have capital account deficits, which typically take the form of investments in countries with current account

deficits. When investor countries decide to rebalance their investment portfolios, it can have a significant negative impact on the value of those investee country currency.

### 2.3.4 Capital Flows

Capital flows tend to be the dominant factor influencing exchange rates in the short term, as capital flows tend to be larger and more rapidly changing than goods flows.

The country with capital account surplus may be running a current account deficit by borrowing from abroad. There will be an issue for debt sustainability, as when the level of debt gets too high relative to GDP, investors may question the sustainability of this level of debt, leading to a rapid depreciation of the borrower's currency.

### 2.3.5 Mundell Fleming Model

The model evaluates the short-term impact of monetary and fiscal policies on exchange rates.

For our case, USD/JPY forex rate, both Japan and United States have high capital mobility (without capital mobility restriction). Therefore, expansionary monetary policy will decrease the interest rate and reduce the inflows of capital investment in physical and financial assets. The decrease in capital inflows reduces the demand for the domestic currency, resulting in depreciation of the domestic currency.

### 2.3.6 Proxies for Econometric Exogenous Variables

Econometric Exogenous Variable	Proxy
Base Interest Rate	1Y Sovereign Bond Yield
Inflation	Consumer Price Index (CPI)
Current Account Balance	Currency Account Balance
Capital Inflow	Foreign Direct Investment (FDI)
Monetary Policy	Money Supply (M2)

Table 1: Proxies for Econometric Exogenous Variables

## 2.4 Optimized Indicators

### 2.4.1 Traditional Technical Indicators

Simple Moving Average (SMA): is a trend indicator that smooths out the price data and identifies trends of property price. The calculation is simply by adding up all the data points over a specified period and dividing the sum by number of periods.

Exponential Moving Average (EMA): is also a trend indicator that gives us the smoothed out trend of property price. Different from SMA, it allocates more weight to short-term price change, thus can better reflect the latest market trend.

Average Directional Index (ADX): is derived from two rudimental indicators: Plus Directional Indicator (+DI) and Minus Directional Indicator (-DI). As a trend indicator, ADX tells us the strength of prevailing trend by calculating the division of Average True Range of DX by Smoothing Period.

Bollinger Bands: is a volatility indicator that consists of three lines: N-period SMA, an upper band with K times the N-period SMA deviation above the N-period SMA and a lower band with K times the N-period SMA deviation below the N-period SMA. It is dynamically adjusted to price volatility, and gives the trader signals of overbought and oversold when the middle band crosses the upper band and lower band.

Standard Deviation: is a statistical measurement of the variability and dispersion of property's price around its mean. It provides insights into the volatility of the property's price, which is associated with the risk and potential price swing.

Rate of Change (ROC): measures the percentage change in price between the current closing price and the closing price a certain number of periods ago. It helps the trader to identify the speed of price movement, which gives the insight of the strength of a trend.

Relative Strength Index (RSI): is a momentum indicator that measures the speed and change of price movements. RSI, which oscillates between 0 and 100, is calculated based on Relative Strength (RS) where RS is the average of x days' up close divided by the average of x days' down close. It helps the trader to identify the sell signal when RSI crossover upper band and buy signal when RSI crossover lower band.

Moving Average Convergence Divergence (MACD): is a momentum indicator which tells us the relation between two moving averages of a security's price. It helps the trader to identify the potential mean reversion (buy and sell signals) when crossover between MACD line and signal line happens. The MACD line is calculated by the difference between two periods' EMA and the Signal line is calculated by the EMA of the MACD line.

Stochastic Oscillator: is a momentum indicator that calculates the closing price of a property over a specific period (normally 14 periods). The stochastic oscillator helps to signal the potential overbought & oversold and mean reversion.

Momentum: is a momentum indicator which calculates the difference between today's closing price and N periods' ago closing price. It gives the trader the indication of momentum of the properties' price.

#### **2.4.2 Optimization of Technical Indicators**

On top of the traditional technical indicators, we aimed to test if there is opportunity to optimise parameters of certain technical indicators in order to generate stronger signal generating features. The technical indicators optimised include:

Simple Moving Average (SMA)	Stochastic Oscillator
Exponential Moving Average (EMA)	Bollinger Bands
Relative Strength Index (RSI)	Standard Deviation
Moving Average Convergence Divergence (MACD)	Average Directional Index (ADX)
Rate of Change (ROC)	Momentum

Table 2: Traditional Technical Indicators

The optimization is performed by systematically varying the above-mentioned parameters within specific ranges. You’ve utilized a parameter sweep approach where you’ve defined ranges for each parameter to explore multiple combinations. The backtesting library’s optimization function is used to test different combinations of these parameters.

After conducting the optimization process, the best combination of parameters obtained for maximizing the strategy’s performance is:

SMA n = 3	slowk period = 1
EMA n = 3	BB n = 9
RSI n = 7	STDDEV n = 14
short MA = 5	ADX n = 14
long MA = 15	ROC n = 7
signal MA = 5	MOM n = 7
fastk period = 5	

Table 3: Optimal Technical Indicators

These optimized parameter values are expected to enhance the strategy’s ability to generate buy/sell signals based on the technical indicators, potentially improving overall performance when applied to trading strategies. The optimization aims to strike a balance between responsiveness to market changes, reducing noise, and minimizing false signals, thereby increasing the strategy’s robustness and effectiveness.

These optimised technical indicators will eventually flow into our feature importance analysis to determine if they outperform the other features in generating buy and sell signal for the USD/JPY trading pair.

## 3 Model Selection

### 3.1 Selection Process & Methodology

For our forex trading strategy, where the primary objective is to classify potential buy or sell points, we need to concentrate on algorithms that excel in classification. Therefore, we conducted a model selection process specifically geared towards identifying the top three classification algorithms that are most effective and relevant for our trading strategy.

#### 3.1.1 Determine algorithms

The following classification Machine Learning algorithms were considered:

- **Logistic Regression:** A fundamental classification algorithm, typically used for binary classification problems but could be adapted for multi-class classification.
- **Support Vector Machine (SVM):** Versatile and powerful, SVM can be used for both classification and regression, but it's particularly well-known for its performance in classification.
- **K-Nearest Neighbors (KNN):** A simple, non-parametric, lazy learning algorithm used for classification. It classifies a data point based on how its neighbors are classified.
- **Decision Tree:** Primarily used for classification (but can also do regression). It's a tree-like model of decisions and their possible consequences.
- **Extra Trees (Extremely Randomized Trees):** Similar to Random Forest, this ensemble method is used for classification and regression, but it randomizes certain decisions and thresholds, hence 'extremely' randomized.
- **Random Forest:** An ensemble of Decision Trees, typically used for classification (and regression). It's particularly known for its high accuracy, robustness, and ease of use.
- **Gradient Boosted Trees (GBT):** An ensemble boosting method primarily used for classification and regression. It builds the model in a stage-wise fashion and generalizes them by allowing optimization of an arbitrary differentiable loss function.
- **Adaptive Boosting (AdaBoost):** A boosting algorithm that can be used with many other types of algorithms to improve performance. The output of the other learning algorithms ('weak learners') is combined into a weighted sum that represents the final output of the boosted classifier.

### 3.1.2 Model Comparison

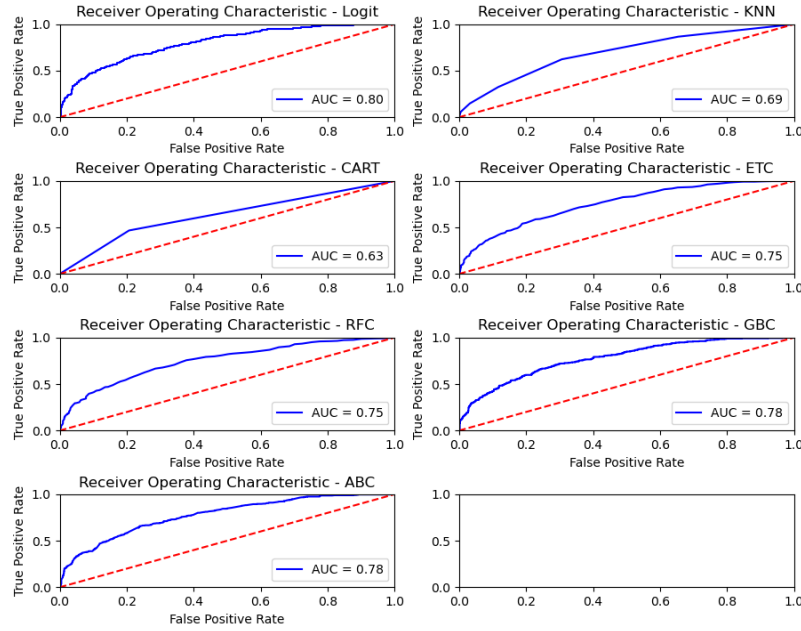


Figure 10: AUC Values of Machine Learning Models

Based on the train and test results of different algorithms to compare the performances. We used AUC (area under curve) of training and testing sets to evaluate the performances of different models.

### 3.1.3 Selected Models

After comparing the performance results of different models, we decided to proceed with Logistic Regression and Gradient Boosted Trees as they performed better than the others. For our third model, we went with Neural Nets.

## 3.2 Choice of ML Models

### 3.2.1 Logistic Regression

We eventually settled on using a logistic regression model with an elastic net penalty. The loss function for this setup was a combination of the logistic loss and the elastic net regularization.

Elastic net is a regularization technique that combines both L1 (Lasso) and L2 (Ridge) penalties.

Logistic Regression typically excels in binary classification tasks, modeling linear relationships and discerning trading signals based on input features. With that being said, LogReg can handle multi-class classification.

By default, Logistic Regression in scikit-learn handles multi-class classification using the one-vs-rest (OvR) scheme or the 'multinomial' scheme when the solver supports it. Since we selected to use the 'saga' solver, which supports 'multinomial', it can handle multi-class classification without explicitly setting the multi\_class parameter. The model will thus automatically choose the appropriate strategy based on the data provided. Since our target variables (y\_train and y\_test) contain three classes (**Buy Hold** and **Sell**), the model will perform multi-class classification using the 'multinomial' approach. This uses the softmax function for multi-class classification in a single model.

The softmax function is used in various multiclass classification methods, including logistic regression and neural networks, as it allows one to handle multiple classes in a probabilistic context. Each output from the softmax function is interpreted as the probability of the corresponding class given the input features.

During training, the softmax function is often paired with a cross-entropy loss (also known as log loss), which provides a measure of how different the predicted probabilities are from the actual class labels. This pairing is effective for training classifiers to predict the correct class probabilities.

### 3.2.2 Gradient Boosted Trees

Gradient Boosted Trees (GBT) is an ensemble machine learning technique that builds models in a stage-wise fashion. It involves the creation of a sequence of trees, where each subsequent tree aims to correct the errors made by the previous trees. The trees are added to the model sequentially until no further improvements can be made. It's a powerful technique used for both regression and classification problems.

For multi-class classification, Gradient Boosting can be used in two ways:

- One-vs-Rest (OvR) Scheme: Here, one decision tree is trained per class and is responsible for separating its class from the rest. This approach can be used for multi-class classification, but it can be computationally expensive as it requires training as many models as there are classes.
- Multiclass LogitBoost: This is a direct extension of boosting to the multiclass case. Instead of fitting a tree per class, it fits a tree per class per iteration but modifies the loss function

to be a multinomial deviance, which is a generalization of the binomial or binary deviance to the multiclass case.

Our GBT model uses the second approach. It inherently supports multi-class classification by optimizing the multinomial deviance loss function, which is a generalization of the logistic loss for multiple classes.

### 3.2.3 Neural Network (LSTM): (added)

Neural Network was not part of the earlier model selection process, but considering that Neural Network is a more advanced and sophisticated model compared to the other more basic ones, we wanted to see whether a more complex model would perform better.

Neural networks are a set of algorithms, modeled loosely after the human brain, designed to recognize patterns. They interpret sensory data through a kind of machine perception, labeling, or clustering raw input.

Within the context of multi-class classification, neural networks can effectively discriminate between more than two classes. This is typically done by including a final layer with a softmax activation function, which is what we did. The softmax function converts the output of the last neural network layer (which could be in any range depending on the previous activation functions) into a probability distribution over predicted output classes. This means the outputs of the softmax function are all non-negative and sum up to 1.

Our neural network model was designed for sequence prediction tasks (since we are working with time series data), making use of LSTM (Long Short-Term Memory) layers, which are a special kind of recurrent neural network (RNN) layer that can capture long-term dependencies and are often used in sequence prediction problems.

## 4 Feature Importance

Thus far we have discussed several features that could be implemented in the final model, including signal processing indicators, exogenous economic indicators, and optimised technical indicators. Now we will bring it all together by their importance in detecting peak and troughs through a feature importance analysis

We chose to use the Random Forest Classifier among other methodologies for several key reasons:

- **Handling of Non-Linearity:** Random Forest Classifier accommodates non-linear relationships between features and target signals, crucial in capturing complex patterns often present in Forex trading data.
- **Resilience to over-fitting:** Random Forest Classifier inherently mitigates over-fitting by averaging predictions across multiple decision trees, reducing the risk of learning noise or specific feature characteristics.



- **Manages Multicollinearity:** Random Forest Classifier is well-suited for Forex data due to its ability to handle correlated features effectively by employing a technique called feature randomization, which allows it to mitigate the impact of multicollinearity among indicators, ensuring robust predictions. Additionally, its ensemble nature and ability to handle large datasets make it particularly adept at capturing complex relationships within financial markets.
- **Outlier Resilience:** Random Forest Classifier can also handle outliers in the features without significantly impacting performance, allowing for robust feature importance determination despite potential data anomalies.

To ensure robustness in our feature importance analysis, we decided to test not just the indicators previously mentioned, but also a wide array of traditional technical indicators as well as higher order derivatives of some of the technical indicators and signal processing indicators. There are 2 main reasons why we decided that higher order derivatives of some of our features may provide additional performance to our PnL when implementing our models:

- **Refinement of Signal Accuracy:** Higher order derivatives can provide additional granularity and detail in understanding the rate of change or acceleration of price movements, potentially making signals more precise for predicting short-term fluctuations or changes in momentum.
- **Reducing Noise and False Signals:** Incorporating higher order derivatives might filter out some noise or false signals that could arise from using only first-order indicators, leading to a more accurate representation of market movements.

After running the features into the Random Forest Classifier, we plotted the feature importance of each feature tested into the following bar charts:

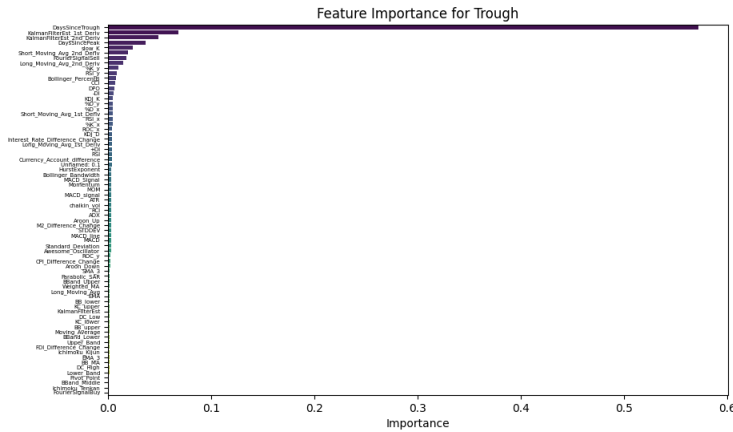


Figure 11: Feature Importance for Peaks

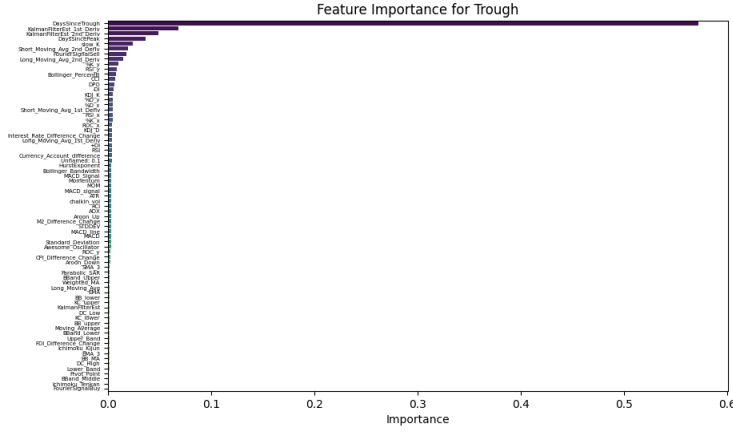


Figure 12: Feature Importance for Troughs

In both analyses of peak and trough importance, several similar features rank very highly, including DaysSincePeak/Trough, Kalman Filter, Fourier Signal, Higher order derivatives of LMA and SMA, and RSI. Noting these findings, we conducted rigorous testing on the different permutations and combinations of these important features through our models to determine which provided us with the most optimal PnL. Through this testing we landed on a final set of features, all of the features that ranked highly in feature importance ended up in our final feature-set:

Indicator	Category
Short_Moving_Avg_2nd_Deriv	Trend Technical Indicator + Higher Order Derivative
Long_Moving_Avg_2nd_Deriv	Trend Technical Indicator + Higher Order Derivative
DaysSincePeak	Signal Processing Indicator
DaysSinceTrough	Signal Processing Indicator
FourierSignalSell	Signal Processing Indicator
FourierSignalBuy	Signal Processing Indicator
KalmanFilterEst_1st_Deriv	Signal Processing Indicator + Higher order Derivative
KalmanFilterEst_2nd_Deriv	Signal Processing Indicator + Higher Order Derivative
RSI	Oscillator Technical Indicator (Optimized)
Slow %K	Oscillator Technical Indicator (Optimized)
Slow %D	Oscillator Technical Indicator (Optimized)
%K	Oscillator Technical Indicator
%D	Oscillator Technical Indicator
Interest_Rate_Difference_Change	Economic Exogenous Variable

Table 4: Description of Indicators

Among the final feature-set, there were several key observations that we made:

- **Balance of Traditional and Non-Traditional Indicators:** There is a mixture of signal processing indicators as well as traditional technical indicators. This shows that while traditionally accepted technical indicators are indeed valuable in assessing the USD/JPY trading pair, Signal Processing Indicators can indeed
- **Underperformance of Hurst Exponent,** Hurst Exponent did not appear in the final feature-set, despite initial literature suggesting its potential inclusion. This could indicate a lack of strong mean-reverting behavior in the spot price of USD/JPY, despite some periodic movement.
- **Optimized Technical Indicators:** Several optimized technical indicators performed strongly in our feature importance analysis, indicating that there is further opportunity to perform optimisation by adjusting the parameters of traditional technical indicators
- **Among the technical indicators that appeared,** all of them are oscillator-related technical indicators. This suggests frequent trend reversals, overbought or oversold conditions, and significant volatility or momentum shifts within the USD/JPY pair.
- **Interest Rate Difference Change:** Although not ranking highly in feature performance, we retained this strong performing exogenous econometric feature to mitigate overdependence on endogenous variables.

In the next section we will discuss the various parameters taken in to consideration when implementing our trading strategy.

## 5 Trading Strategy

### 5.1 Ground Rules / Scope

The rules of trading in our experiment are simple:

- avoid intraday trading ( we do not compete with other players in the HFT space ), focusing on inter-day trading
- Buy once, sell once ( commit to holding ) then repeat, avoiding repeat buys ( i.e. do not buy another position until you have sold the previous one )

### 5.2 Sliding-Window Methodology

Compared with normal data processing methods, sliding window allows us to perform: localized analysis, real-time processing, dynamic adaptability, temporal / spatial analysis, noise reduction and sequential data processing.

There are a multiple consideration points to take note of for managing time series data. Let us examine them.

### 5.2.1 Temporal Order of Data Points

Data points from training set needs to come before test set. The time order of data points have to be preserved, we cannot do random sampling, or rather sampling methods cannot randomize time sequence of data points.

### 5.2.2 Time Decay

For each data point, we assigned an integer age variable based on the number of days elapsed from data point record to "present day" (i.e. day of prediction). We then applied an exponential decay rate to derive its weight. All data points in the training set are then sampled according to their relative weights. If the training set is not large, no data points are discarded. As the training set grows, older data points are discarded based on sample selection based on their assigned relative weights.

### 5.2.3 Time Horizon

For each window, we used our ML model to train using 4 quarters of data before proceeding to test / predict for 2 quarters. As our window progressed, we retained data from the previous quarters for training purposes while applying time decay such that more recent data is prioritized over older data.

### 5.2.4 Class Weight Balancing

During our data preparation phase, we demarcated our data points into 3 main categories: **Buy**, **Hold**, **Sell**. During our inspection of training set, we noticed an over-representation of **Hold** labels, where the ratio of Buy : Sell : Hold ranged between 1 : 1 : 2 and 2 : 2 : 3. Thus we had to ensure class weight adjustment techniques to compensate for this.

### 5.2.5 Data Points

Our training sample set size was capped at 1500 data points to avoid over-fitting of training set. 1500 size was empirically selected after experimenting with different cap sizes.

## 6 Performance Evaluation

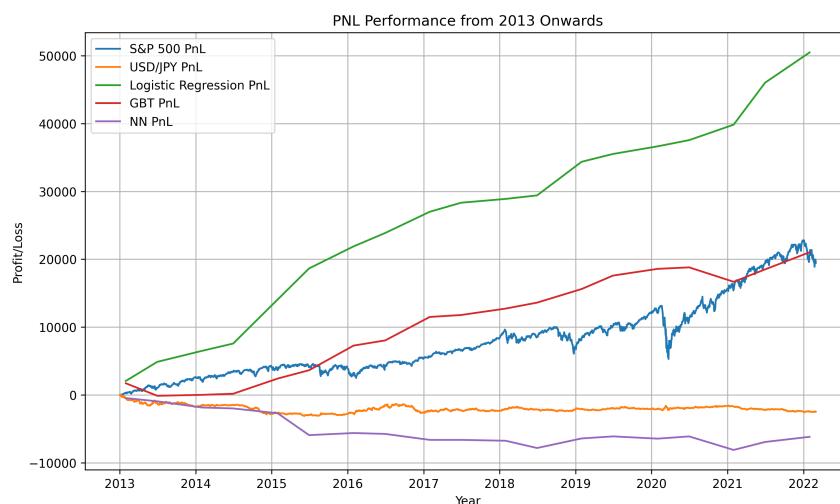


Figure 13: PnL Performance Chart Comparing Various Trading Strategies against S&P500 and Holding JPY Cash

When running all of our models, we generated the PnL distribution shown above. Overall, we can see that Logistic Regression outperformed all other models. It is also worth noting that Gradient Boosted Trees performed similarly to the S&P500 while Neural Networks gave the worst performance.

Metric	Benchmark JPY USD	S&P500	Logistic Regression	GBT	NN (LTSM)*
P&L (MTM)	-4,205 USD	21,428 USD	50,899 USD	22,740 USD	-6,177 USD
Excess Returns	NA	25,633 USD	55,104 USD	26,944 USD	-1,973 USD
Gain/Loss (%)	-41.05%	214%	508%	227%	-61%
Returns per 6 Months	-4.54%	11.22%	26.79%	11.97%	-3.25%
Returns per Annum	-9.08%	22.44%	53.59%	23.94%	-6.50%
Transaction Costs	NA	NA	422 USD	340 USD	292 USD
Interest Costs	NA	NA	1,113 USD	2,897 USD	1,853 USD
Std Dev of Returns	649	720	1911	1548	1038
Std Dev of Returns (%)	9.1%	17.6%	29.9%	26.9%	22.4%
Sharpe Ratio	-0.84	1.566	1.517	0.916	-0.100

Figure 14: Evaluation Table of Model Performance Benchmarked Against Forex Movement and S&P 500

Notes:

- Starting portfolio value is 10,000 USD

- Trading quantum remained fixed at 7,500 USD per lot
- Returns are displayed in Arithmetic form and not Geometric
- \*For NN the quantum was reset every period due to poor performance
- PnL evaluated on 4x leverage for consistency
- PnL Accounts for transaction costs
- Transaction cost is calculated as 1 USD per trade
- Aggregate interest costs are computed assuming borrowing rate of 5% per annum per borrowed quantum

We also conducted testing on different levels of leverage to assess its impact on PnL. We found that in general, there was a linear relationship between leverage and PnL depending on whether the direction was positive or negative. Furthermore we observed possible inflection points in the PnL distribution which could indicate some opportunity for further optimisation when implementing into our models.

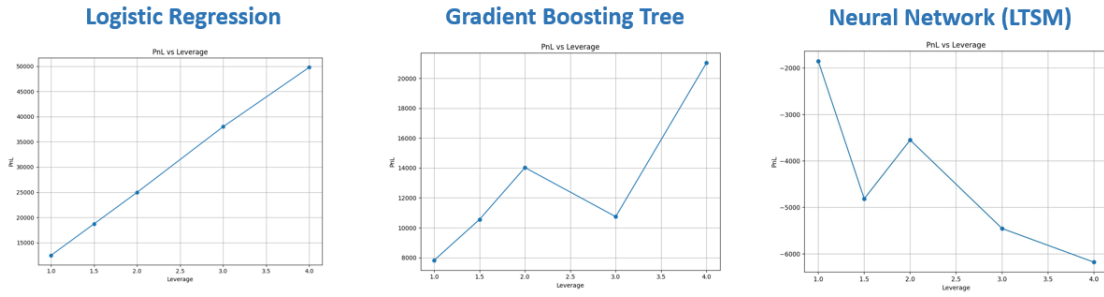


Figure 15: Visualisation of PnL against various levels of leverage for each model

## 7 Conclusion

Non-traditional Technical Indicators may indeed provide some edge in generating buy/sell signal. The ones that rank highly in importance may indicate some movement characteristics of the particular asset class.

Out of the signal processing indicators explored, **Fourier Transform** and **Kalman Filter** ranked highly, but *not Hurst Exponent*. This could be an affirmation that understanding the cyclical/periodic nature of the USD/JPY trading pair could yield good results. However, understanding the mean-reversion / trend nature of the time series based on recent historical data does not yield added forward predictive signal.

Furthermore, traditional Technical Indicators and Econometric Exogenous Variables may prove useful in generating signal when in silo, but may get pushed out when strong, additional indicators are introduced.

Though not the crux of our project, we also explored higher derivatives of many of our signal processing and traditional technical indicators. We found that though the indicator's raw values may not rank highly in feature importance, their 1st and/or 2nd derivatives may hold significant importance. In our case, we saw the significance of higher order derivatives of optimised SMA in determining feature importance.

## 8 Appendix

### 8.1 Features List

Category	Indicator	Description
Volatility	Average True Range	EMA of the true range values over a specified period, it indicates the volatility
	Bollinger PercentB	Bollinger band dynamically adjusted to price volatility, and shows signals of overbought and oversold
	Bollinger Bandwidth	Difference between upper BB band and lower BB band
	BBand Upper	Upper Bollinger band, SMA of N-periods + K times derivatives
	BBand Middle	SMA of N-periods
	BBand Lower	Lower Bollinger band, SMA of N-periods - K times derivatives
	Keltner Channel upper	Keltner Channel upper KC middle line plus multiple of Average True Range, it identifies the potential trend reversal points
	Keltner Channel lower	KC middle line minus multiple of Average True Range, it identifies the potential trend reversal points
	Standard Deviation	The standard deviation of asset price, it reflects the volatility
	Donchian Channels High	Higher Channel of Donchian Channels, identify the trend and potential reversal
	Donchian Channels Low	Lower Channel of Donchian Channels, identify the trend and potential reversal
	Chaikin vol	Identify the periods of increased or decreased volatility
Momentum	Awesome Oscillator	Difference between 34-period SMA and 5-period SMA, it identifies the bullish or bearish momentum
	Commodity Channel Index	Identify cyclical turns in the price of an asset. It is calculated by dividing Typical Price over Mean Deviation
	Momentum	Calculates the difference between today's closing price and N periods' ago closing price



Category	Indicator	Description
	Rate of Change x	percentage change in price between the current closing price and the closing price x periods ago
	Rate of Change y	percentage change in price between the current closing price and the closing price y periods ago
	Rank Correlation Index (RCI)	Combine price change data and time change data to identify potential changes in market sentiment
Trend	Aroon up	Measure strength of trend by calculating the number of periods since the highest closing price occurred
	Aroon down	Measure strength of trend by calculating the number of periods since the lowest closing price occurred
	ADX	Reveal the strength of prevailing trend by calculating the division of Average True Range of DX by Smoothing Period
	Moving Average	Smooths out the price data and identifies trends of property price
	EMA	Smooths out the price data and identifies trends of property price with more weightage allocation to short-term price change
	SMA 3	Moving average using 3 days look back period
	EMA 3	Exponential moving average using 3 days look back period
	Weighted MA	Moving average with weighted allocations
	Long Moving Avg	Moving average using long look back period
	Short Moving Avg 1st Derivative	Gradient of short moving average
	Short Moving Avg 2nd Derivative	Curvature of short moving average
	Long Moving Avg 1st Derivative	Gradient of long moving average
	Long Moving Avg 2nd Derivative	Curvature of long moving average
Oscillator	KDJ K	Difference between Signal line J and K line of Stochastic Oscillator
	KDJ D	Difference between Signal line J and D line of Stochastic Oscillator
	RSI x	Upper bond of RSI

Category	Indicator	Description
	RSI y	Lower bond of RSI
	RSI	Momentum Oscillator indicator to identify the overbought and oversold signals
	Percentage K x	This refers to the K line in the Stochastic Oscillator indicator applied to the x-axis. It measures the current closing price relative to the range of prices over a specific period, signaling potential overbought or oversold conditions in the x-axis market, helping traders identify potential reversal points
	Percentage D x	The D line in the Stochastic Oscillator applied to the x-axis, representing a moving average of the K x line. It smoothens the K x line's fluctuations, providing insights into potential trend changes or confirmations of market momentum shifts in the x-axis
	Percentage K y	Similar to K x, this indicates the K line in the Stochastic Oscillator applied to the y-axis. It measures the current closing price relative to the range of prices over a specific period, offering insights into potential overbought or oversold conditions in the y-axis market
Other Technical Indicators	Ichimoku Tenkan	Analyse short term market trends and potential reversal point
	Ichimoku Kijun	Analyse long term market trends and potential reversal point
	MACD	Identify the potential mean reversion
	MACD Signal	Signal component of MACD
	MACD Line	Line component of MACD
	Parabolic SAR	Identify the reversal points in the price direction
	Pivot Point	Identify potential support and resistance levels
Non Traditional Technical Indicators	Hurst Exponent	Persistence of current trend using Hurst Exponent
	Days Since Peak	Days elapsed since last peak
	Days Since Through	Days elapsed since last trough
	Price Change Since Peak	Price difference since last peak
	Price Change Since Through	Price difference since last through

Category	Indicator	Description
	Fourier Signal Sell	Sell Signal predicted by Fourier Transformed cycle
	Fourier Signal Buy	Buy Signal predicted by Fourier Transformed cycle
	Kalman Filter Est	Optimal state estimate of Forex spot rate combining prediction phase and update phase using Kalman Filter Algorithm
	Kalman Filter 1st Derivative	The gradient of Kalman Filter
	Kalman Filter 2nd Derivative	The curvature of Kalman Filter
Econometric Exogenous Variables	Change of Interest rate Difference	Measure the countries' Tbill yield
	Change of CPI Difference	Measure of a country's inflation
	Change of Current Account Difference	Country's balance of payments that records its economic transactions with the rest of the world
	FDI Difference	Measure of capital inflow to a country
	M2 Difference	Measure of money supply in an economy

Table 5: All Features Tested During Feature Importance Phase

Data	Source
Close Prices of USD JPY	Yahoo Finance (Ticker: USDJPY=X)
Close Prices of S&P500	Yahoo Finance Package (Ticker: GSPC)
1Y Sovereign Bond Yield	Bloomberg (Tickers: H15T1Y and GJTB12MO)
Consumer Price Index (CPI)	Bloomberg (Tickers: CPI YOY and JNCPIYOY)
Currency Account Balance	Bloomberg (Tickers: OEJPABDO and EHCAUS)
Foreign Direct Investment	Bloomberg (Tickers: FDINUSA and FDINJPN)
Money Supply (M2)	Bloomberg (Tickers: JMNSM2 and ECORUSN)

Table 6: Data Sources