Short-term Trading strategy on G10 Currencies

Dhruv Agrawal*, Man Sing Ho**, Sergey Chigrinov***, Ritabrata Bhattacharyya****

* WorldQuant University, dhruva1@stanfordalumni.org (D. Agrawal)

** WorldQuant University, mshoalbert@gmail.com (M.S. Ho),

*** WorldQuant University, chigrinov.s.88@gmail.com (S. Chigrinov)

**** WorldQuant University, ritabrata.bhattacharyya@wqu.org (R. Bhattacharyya)

Abstract

This paper presents research on a profitable trading strategy for G10 currencies.

We will devise trading strategies by considering realistic trading scenarios, analyze the performance of such strategies on out of sample data, identify the risks of these trading strategies, explain why the trading strategy works, and summarize and draw conclusions.

We will use technical indicators like Moving average (MA), Exponentially-weighted moving average, Ichimoku, Relative strength index, stochastic oscillators, Williams %R, Commodity Channel Index, and Bollinger bands. We will also use fundamental indicators like interest rate differentials from one currency to another.

We will build a machine learning based model as it would be better equipped to build dynamic trading rules to capture profitable trading opportunities. We used an ensemble of machine learning algorithms - logistic regression with four principal components, and a voter classifier with random forest, extremely randomized trees, logistic regression with 4 principal components from the features, and support vector machine with 4 principal components from the features, giving 27% and 27% annualized returns, and 1.47 and 1.60 Sharpe ratio respectively.

Keywords: *Moving average (MA), exponentially-weighted moving average, Ichimoku, relative strength index, stochastic oscillators, Williams %R, Money Flow Index, Bollinger bands, support vector machine, Naïve Bayes, extreme Gradient Boosting, Light Gradient Boosted Machine, Adaptive Boosting, random forest, extremely randomized trees, logistic regression, principal component analysis, voting classifier.*

1 Introduction

The G10 currencies are a group of currencies that are among the most used and traded currencies in the world. The G10 currencies list is as follows:

- United States dollar (USD)
- Canadian dollar (CAD)
- Japanese yen (JPY)
- Australian dollar (AUD)
- New Zealand dollar (NZD)
- Euro (EUR)
- Pound sterling (GBP)
- Swiss franc (CHF)
- Norwegian krone (NOK)
- Swedish krona (SEK)

The benefits of trading G10 currencies compared with counterparts in emerging markets (EM currencies) is that:

1. G10 currencies have freely floating exchange rates and the prices are determined by market forces of demand and supply.

2. G10 currencies account for the majority of daily turnover in the foreign exchange (FX) market, which suggests trading liquidity.

3. G10 currencies are the currencies used by the largest industrialized countries. These currencies' volatility should be relatively lower than most currency pairs outside of G10 which could be subject to larger political and policy risk.

A major drawback of trading G10 currencies is that every "evident" market inefficiency gets exploited very quickly due to the high number of market participants. Therefore, it is very hard to create a long-lasting strategy relying on a fixed set of rules. Machine learning offers a dynamic way of managing trading rules, therefore, it may be used for this purpose.

So in this paper, we are going to use machine learning techniques on technical and fundamental indicators to come up with a profitable trading strategy for G10 currencies.

2 Theoretical Framework

We are going to rely on the following framework:

1. Technical analysis

2. Fundamental analysis

3. Overall research on G10 currencies

We would start with technical analysis, which should be the most relevant to short-term trading.

2.1 Technical analysis

In finance, technical analysis is a technique for forecasting the direction of prices through the study of past market data, primarily price and volume, assuming that the market is weak-form efficient under efficient market hypothesis. Technical analysis has long been a part of the finance practice and has been studied in academic finance literature too. Technical analysis is based on the following philosophy, as Neely and Weller (2011)[1] suggested that:

1. Market actions discount everything, such that price history incorporates all relevant information in the current price.
2. Assets prices move in a trend where people buy (sell) assets when the price is rising (falling), in anticipation of higher (lower) prices in the future.
3. History repeats itself.

Technical indicators can be categorized into various types accordingly to Technical Analysis by Hobson in 2011[2].

1. Trend indicators

Moving averages (MA)

The simple moving average is called a rolling average to smooth out short-term fluctuations and show longer-term trends and cycles.

N-period movement average is suggested as follows:

$$\mathrm{MA_t} = \frac{\sum_{i=0}^{n} \text{Closing price}_{t-i}}{n}$$

Some of the popular selections for period n would be 5, 20, and 50, 200 as the gold cross and death cross.

Pukthuanthong-Le, Levich, and Thomas III[3] in 2006 employed three types of moving average comparisons to generate a trading signal (comparing 5-day moving average with the 20-day moving average, 1-day moving average with the 5-day moving average, and 20-day moving average to the 200-day moving average). They bought when the short term average is greater than the long term average and hold the position until the signal changes.They suggested that there were opportunities to use such technical trend-following rules but profits were substantially reduced in recent days.

Exponentially-weighted moving average

Another option is to allocate specific weights to historical prices. A common example is the exponentially-weighted moving average.

---

1 Neely, Christopher J. & Weller, Paul A. (2011). Technical Analysis in the Foreign Exchange Market. Working Paper 2011-001B, Research Division, Federal Reserve Bank of St. Louis. http://research.stlouisfed.org/wp/2011/2011-001.pdf

2 Hobson, M.P. (2011). Technical Analysis. Cavendish Astrophysics, University of Cambridge. http://www.mrao.cam.ac.uk/~mph/Technical_Analysis.pdf

3 Pukthuanthong-Le, K., Levich, R.M. & Thomas III, Lee.R. (2006). Do Foreign Exchange Markets Still Trend? The Journal of Portfolio Management 34(1). http://people.stern.nyu.edu/rlevich/wp/PLT1.pdf

The change of the trend is assumed to happen when a "faster" MA (the one with a smaller number of days included) intersects with a "slower" MA. This is to add weight for each of the data points that decreases exponentially.

$$\mathrm{EWMA_t} = \mathrm{weight}\ (\mathrm{Closing\ price_t}) + (1 - \ \mathrm{weight})\ (\mathrm{EWMA_{t-1}})$$

Ichimoku

Ichimoku, developed by Goichi Hosoda, measures moving averages for present and future market conditions. Ichimoku has been used widely after Nicole Elliott further discussed in her 2007 book Ichimoku Charts (Twomey, 2012)[4]. The methodologies of Ichimoku is further discussed in the Summer 2008 Journal of Technical Analysis[5] with an application focus in Japanese equity markets.

There are five key components of the Ichimoku indicator:

I. **Tenkan-sen -** Known as Conversion Line. It is the average of the highest and lowest prices of an asset over the last nine periods:

$$\mathrm{TS} = \frac{\max(\mathrm{s9})\ + \min(\mathrm{s9})}{2}$$

II. **Kijun-sen -** Known as Base Line. It is the average of the highest and lowest prices of an asset over the last 26 periods:

$$\mathrm{KS} = \frac{\max(\mathrm{s26}) + \min(\mathrm{s26})}{2}$$

III. **Senkou Span A** - Known as Leading span A. It is one of the two Cloud boundaries and it's the midpoint between **Tenkan-sen** as the Conversion Line and **Kijun-sen** as the Base Line with the formula below:

$$\mathrm{SSA} = \frac{\mathrm{TS} + \mathrm{KS}}{2}$$

IV. **Senkou Span B** - Known as Leading Span B. It is given by half of the difference between the last 52 periods high and the last 52 periods low using the formula below:

$$\mathrm{SSB} = \frac{\max(\mathrm{s52}) - \min(\mathrm{s52})}{2}$$

V. **Chikou Span -** known as the "lagging span". It is created by using closing prices 26 periods behind the latest closing price of an asset.

---

4 Towmey, B. (2012). Inside the Currency Market – Mechanics, Valuation, and Strategies, P.289-290. John Wiley & Sons, Inc, ISBN 978-0-470-95275-7 (hardback); ISBN 978-1-118-14933-1 (ebk); ISBN 978-1-118-14934-8 (ebk); ISBN 978-1-118-14935-5 (ebk).

5 Lashinski, V. (2008). Ichimoku Kinko Hyo, P.42 to 49. Journal of Technical Analysis 2008 Summer / Fall, Issue 65. Market Technicians Association, Inc.. https://www.yumpu.com/en/document/read/11279567/journal-of-technical-analysis-market-technicians-association

To add to the 2008 paper in the Journal of Technical Analysis which has been focusing on the Japanese equity market, Deng and Sakurai in 2014[6] designed two trading strategies based on the support/resistance level of Ichimoku and conducted simulated trading on short-term foreign exchange rates. Their trading strategies are applied to five currency pairs: USDJPY, EURUSD, GBPUSD, USDCHF, and AUDUSD which are highly relevant to our G10 currencies topic. Their research results showed that the average return of one trading strategy that is based on Ichimoku could be better than other baseline strategies, which builds the ground of our short-term G10 currencies trading strategies.

2. Momentum Oscillators

Momentum is a trend-following strategy, where the strategy buys the assets which have performed well in the past and sells the assets which have performed badly. Okunev and White (2001)[7] used moving average rule and concluded that there was potential to generate excess returns in foreign exchange markets by adopting a momentum strategy and suggested not all foreign exchange markets operate efficiently. However, Pukthuanthong-Le and Thomas (2008) found that the profitability of trend following eroded for major currencies and their associated cross exchange rates around the mid-1990s[8]. Rohrbach, Suremann, and Osterrieder (2017)[9] also suggested momentum trading strategies for G10 worked well until the 2008 global financial crisis and are no longer profitable.

The **Relative strength index** is a momentum oscillator created by J. Welles Wilder which measures the price movements' speed and change. RSI with n-day lookback period is calculated as follows:

$$\text{RSI} = 100 - \frac{100}{\left(1 + \frac{\text{Average of Upward Price Change in n} - \text{day period}}{\text{Average of Downward Price Change in n} - \text{day period}}\right)}$$

RSI oscillates between zero and 100 and traditionally the RSI suggests an overbought signal when above 70 and oversold signal when below 30.

---

6 Deng, S. & Sakurai, A. (2014). Short-term foreign exchange rate trading based on the support/resistance level of Ichimoku Kinkohyo. Institute of Electrical and Electronics Engineers. https://ieeexplore.ieee.org/abstract/document/6948127

7 Okunev, J. & White, Derek (2001). Do Momentum Based Strategies Still Work In Foreign Currency Markets?. http://wwwdocs.fce.unsw.edu.au/banking/workpap/currency.pdf

8 Pukthuanthong, K. and Thomas III (2008), Lee R. Weak-Form Efficiency in Currency Markets., Financial Analysts Journal 64(3). https://www.researchgate.net/publication/228258482_Weak-Form_Efficiency_in_Currency_Markets

9 Rohrbach, J., Suremann, S. & Osterrieder, J. (2017). Momentum and Trend Following Trading Strategies for Currencies Revisited - Combining Academia and Industry. Zurich University of Applied Sciences. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2949379

Anderson and Li (Anderson and Li, 2015)[10] reviewed whether the standard configuration of RSI below 30 and RSI above 70 as buy or sell threshold would generate any trading profits, and also any recalibration of the threshold would lead to trading profits by trading USDCHF. They found that changing the buying and selling threshold could still give profits and suggested the market is neither in strong-form efficiency (all information including public information and private information, is accounted for in current asset prices, and no information can give an investor an advantage on the market) nor semi-strong form efficiency (public information is part of a assets current price, investors cannot utilize either technical or fundamental analysis, though information not available to the public can help investors).

Stochastic (%K %D) is also another well-known indicator popularized by George Lane in the 1950s, which is a method based on the observation that as price decreases, the daily close prices tend to accumulate ever closer to their extreme lows of the daily range. Conversely, as price increases, the daily close prices tend to accumulate ever closer to the extreme highs of the daily range[11]. Stochastic is widely used in technical analysis and included in some of the key technical analysis textbooks including Murphy (1999)[12].

There are three components for %K%D:

1. **Fast Stochastic oscillating %K with n day lookback period**

   A stochastic oscillator is a momentum indicator comparing a particular closing price of a security to a range of its prices over n days as follows:

$$\%\mathrm{K} = \frac{\text{Closing price}_{\mathrm{T}} - \text{Lowest price within n day period}}{\text{Highest price within n day period} - \text{Lowest price within n day period}}$$

   This measures on a percentage basis of 0 to 100 where the closing price is in relation to the total price range for the n-day period.

2. **3-period moving average of %K (%D)** as a 3-period moving average of %K with formula as follows:

$$\%\mathrm{D}_{\mathrm{T}} = \frac{\%\mathrm{K}_{\mathrm{T}} + \%\mathrm{K}_{\mathrm{T-1}} + \%\mathrm{K}_{\mathrm{T-2}}}{3}$$

10 Anderson, B. & Li, S. (2015). An investigation of the relative strength index. Banks and Bank Systems, Volume 10, Issue 1, 2015. https://www.researchgate.net/publication/283691105_An_investigation_of_the_relative_strength_index

11 Lane, George C. (1990). Lane's Stochastics. Stocks & Commodities V. 2:3 (87-90). https://www.forexfactory.com/attachment.php/1410270?attachmentid=1410270&d=1397915452

12 Murphy, John J. (1999). Technical analysis of the financial markets. Larry Williams % R, Point and Figure Charting, ISBN 0-7352-0066-1.

3. **3-period slow stochastic oscillating %D** as 3-period of moving average of %K with formula as follows:

$$\%\text{Slow } D_T = \frac{\%D_T + \%D_{T-1} + \%D_{T-2}}{3}$$

These formulas produce two lines oscillating between 0 and 100. The D line is a slower line and the K line is a faster line and the signal to watch is a divergence between the price of the underlying market and the D line when the D line is in an oversold or overbought area.

There is a prior study (Bhavani and Pichai, 2016)[13] of trading EUR and USD as one of the G10 currency pairs in the foreign exchange market using technical analysis tools, including Stochastic Oscillator, Relative Strength Index (RSI), Bollinger Bands, and Parabolic Stop and Reversal (PSAR) and suggests the benefits of algorithmic trading to overcome mental concentration issues.

When talking about oscillators, one of the well-known ones is called **William percentage range** (William %R). William % R was developed by Larry William to see where today's close was in Relationship to the Range of the last "X" time period. This was first mentioned in Larry's book, "How to Select Stocks for Immediate & Substantial Gains" in 1967[14]. By then the William % R had been used widely in technical analysis and included in some of the key technical analysis textbooks including Murphy (1999)[15].

Williams %R is calculated as follows:

$$WR = \frac{\text{HighestHigh} - \text{CurrentClose}}{\text{HighestHigh} - \text{LowestLow}} * 100$$

The indicator oscillates between 0 and 100. This is used as a buying sign when %R is closer to 100 or conversely to sell when %R is close to 0.

There is another indicator called the **Commodity Channel Index (CCI)**. It is an oscillator created by Donald Lambert in 1980 for the commodities market to identify the oversold and overbought position, similar to what RSI intends to provide. Commodity Channel Index with 20-day lookback period definition is as follows

1. First, the day's Typical Price is calculated:

---

13 Bhavani, N. Ganga & Pichai, A. (2016). A Study of the Euro and U.S Dollar in the foreign exchange market using technical analysis tools. Innovative Journal of Business and Management 5:5, November –December (2016) 116 – 101., https://www.researchgate.net/publication/312623592_A_STUDY_OF_THE_EURO_AND_US_DOLLAR_IN_THE_FOREIGN_EXCHANGE_MARKET_USING_TECHNICAL_ANALYSIS_TOOLS

14 William, L. (2011). Williams Percent R Indicator (%R). Larry Williams CTI Publishing, 2011. http://williamspercentr.com/

15 "Technical analysis of the financial markets", Larry Williams % R, Point and Figure Charting, ISBN 0-7352-0066-1 John J. Murphy, 1999.

$$\text{Typical Price} = \frac{\text{High price} + \text{Low price} + \text{Close price}}{3}$$

2. Calculate mean absolute deviation of price

$$\text{Mean Deivation} = \frac{\sum_{i=1}^{20} |\text{Typical Price} - 20 - \text{period moving average of Typical Price}|}{20}$$

3. Calculate Commodity Channel Index

$$\text{Commodity Channel Index} = \frac{\text{Typical Price} - 20 - \text{period moving average of Typical Price}}{0.015 \text{ x Mean Deviation}}$$

We typically use the 20-day period to calculate the typical price and mean absolute deviation.

With the constant of 0.015, the majority of the Commodity Channel Index would fall between -100 and 100 and work similarly to the RSI. CCI over 100 suggests a strong uptrend and trade should be closed when reverting below 100. Similarly, CCI below -100 suggests a strong downward trend and short sell should be covered when this above -100 level.

Maitah, Prochazka, Cermak, and Šrédl[16] researched on CCI and concluded that this rule is profitable in capturing volatility using mean deviation during a volatile period. Roudgar[17] also used CCI as one of the technical indicators among other technical indicators, including simple moving average, moving average convergence divergence, stochastic oscillator, and RSI. His results showed over 60% of the trades as profitable.

Volume Indicators

The **Money Flow Index (MFI)** is a momentum indicator measuring the money flow into and out of a security over a specified period of time. The MFI is calculated by accumulating positive and negative Money Flow values and normalized it into the MFI oscillator form. Since this momentum indicator is adding up the trading volume to the RSI (Relative Strength Index), it is also known as the volume-weighted RSI.

Money Flow Index definition is as follows:

1. First, the day's Typical Price is calculated:

$$\text{Typical Price} = \frac{\text{High price} + \text{Low price} + \text{Close price}}{3}$$

---

[16] Maitah, M., Prochazka, P., Cermak, M. & Šrédl, K. (2016). Commodity Channel Index: Evaluation of Trading Rule of Agricultural Commodities. International Journal of Economics and Financial Issues, 2016, 6(1), 176-178. http://www.econjournals.com/index.php/ijefi/article/download/1648/pdf

[17] Roudgar, E. (2012). Forecasting Foreign Exchange Market Trends: Is Technical Analysis Perspective Successful?. Institute of Graduate Studies and Research, Eastern Mediterranean University. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.938.1143&rep=rep1&type=pdf

2. Next, Money Flow is calculated by multiplying the period's Typical Price by the volume.

$$\text{Money Flow} = \text{Typical Price} \times \text{Volume}$$

3. If today's Typical Price is greater than yesterday's Typical Price, it is considered Positive Money Flow. If today's price is less, it is considered Negative Money Flow.

4. Calculate the Money Flow Ratio by adding up all the positive money flows over the last 14 periods and dividing it by the negative money flows for the last 14 periods.

5. Money Flow Index is calculated as follows:

$$\text{Money Flow Index} = 100 - \frac{100}{1 + \text{Money Flow Ratio}}$$

Marek and Markova (2020)[18] discussed using the S&P 500 that MFI may be more profitable than a buy-and-hold strategy with the caveat that parameters of MFI need to be optimized. MFI is also referred to in various papers researching the foreign exchange market, such as Bartkus (2018)[19] and Ilic and Brtka (2011)[20]. However, we noted that the foreign exchange market is made on an over-the-counter (OTC) basis where volume data is usually only available daily from reports produced by market infrastructure providers such as Depository Trust & Clearing Corporation. Intraday volume data is usually not available.

3. Volatility Indicators

Bollinger bands

Bollinger bands were created by John Bollinger in the 1980s. The logic for this indicator is simple yet efficient. If the price is 2-standard deviations away from its mean, most likely it will revert. This is an indicator designed for trading in a volatile market when the direction of whether the asset's price moves up or down is uncertain. However, it fails when there is a strong trend. Bollinger bands with n-period and 2 standard deviations are as follows:

$$\text{Upper band: UB} = \text{n} - \text{period MA} + 2 * \text{Std}(\text{price}_{\text{n}})$$

---

18 Marek, P. & Markova, V. (2020). Optimization and Testing of Money Flow Index. 19th Conference of Applied Mathematics. https://www.researchgate.net/publication/339029303_Optimization_and_Testing_of_Money_Flow_Index

19 Bartkus, C. (2018). Using the relative strength index for active investments in the foreign exchange market. http://doi.prz.edu.pl/pl/pdf/zim/358

20 Ilic, V. & Brtka, V. (2011). Evaluation of algorithmic strategies for trading on foreign exchange market. Information and Communication Technologies for Small and Medium Enterprises (ICT-SME's2011). https://www.researchgate.net/publication/299287406_Evaluation_of_algorithmic_strategies_for_trading_on_foreign_exchange_market

$$\text{Lower band: LB} = \text{n} - \text{period MA} - 2 * \text{Std}(\text{price}_\text{n})$$

Lento, Gradojevic, and Wright (2017) performed an analysis using Bollinger Bands. Apart from the traditional setup of a 20-day moving average with 2 standard derivations, they also introduced variants of the 30-day moving average with 2 standard derivations, 20-day moving average with 1 standard derivation. One of the interesting observations from their analysis is that they could generate better returns than the buy-and-hold strategy in the CADUSD pair which is highly relevant to our study in G10 currency training.

2.2 Fundamental analysis

Other than technical indicators, a currency's value would be determined by the demand and supply in the foreign exchange market. This could be one of the modeling factors that explain the trend of the prices, rather than only historical prices. Per Reserve Bank of Australia[21], with an example of the Australian dollar, the fundamental factors affecting currencies spot price include:

1. **International trade in goods and services -** Demand for a country's currency will increase if its' exports increase and it will decrease if its' imports increase.

2. **Capital flows -** Interest rate differentials can affect capital flows and influence the exchange rate in the medium term. One of the well-known strategies related to interest rate differentials is called the carry trade, which is to invest and fund a higher yield currency with a lower yield currency to capture the profits arising from interest rate differentials.

3. **Terms of trade –** This is the ratio of an index of a country's export prices to an index of its import prices**.** Deteriorating (improving) terms of trade translate into a weaker (stronger) currency since the country has to spend more to import the same amount of products.

4. **Purchasing power parity and relative inflation rates -** The purchasing power parity theory suggests the exchange rate is affected by relative rates of inflation between countries in the long run. Inflation would be included in the nominal interest rates and thus the exchange rate. Countries with higher inflation would see their currencies fall in value.

The major drawback of fundamental data is that it is too low frequency for short term trading. However, if the release results are not as expected by investors, it may lead to instantaneous shocks at the moment when the data gets published.

2.3 General G10 currency research

21 Hamilton, A. (2018). Understanding Exchange Rates and Why They Are Important. Bulletin -December 2018 | Finance. https://www.rba.gov.au/publications/bulletin/2018/dec/understanding-exchange-rates-and-why-they-are-important.html

It is interesting to note that G10 currencies also showed some connectedness among them (Betendorf and Heinlein, 2019)[22]. They can be classified as commodity currencies (the Australian dollar, the Canadian dollar, and the New Zealand dollar), European currencies (the Euro,the Norwegian krone, the Swedish krona), and safe haven/carry trade financing currencies (the Swiss franc, the US dollar, and the Japanese Yen). Rohrbach, Suremann, and Osterrieder (2017)[23] also showed similar return clusters for using momentum and trend following strategies and suggested momentum trading strategies for G10 worked well until the 2008 global financial crisis but are no longer profitable.

The following research gives useful context to our work:

- In 2000 Nasution and Agah[24] explored the use of neural networks to forecast currency exchange rates by experimenting with USDJPY in the early days of 2000 when neural network was not popular due to limited computing power. They compared their results with linear prediction using mean and median of the past five previous days as the forecasts. The neural network approach gave a smaller tracking error.
- In the same year, Yao and Tan[25] studied the use of moving average feeding into neural networks with USDPY, USDEUR, USDGBP, USDCHF, and AUDUSD and suggested profitability with simple technical indicators in the out sample and paper portfolio.
- In 2004, Hryshko and Downs[26] presented trading strategies based on the machine learning methods of genetic algorithms and reinforcement learning.
- In 2017, Carapuco[27] explored reinforcement learning in the EURUSD market and generated a yearly profit of 16.3%.

---

22 Bettendorf, T. & Heinlein, R. (2019). Connectedness between G10 Currencies: Searching for the Causal Structure. Discussion Paper Deutsche Bundesbank No 06/2019. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3337431

23 Rohrbach, J., Suremann, S. & Osterrieder, J. (2017). Momentum and Trend Following Trading Strategies for Currencies Revisited - Combining Academia and Industry. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2949379

24 Nasution, Bona P., & Agah, A. (2000). Currency Exchange Rate Forecasting with Neural Networks. Journal of Intelligent Systems 10.3. https://www.degruyter.com/view/journals/jisys/10/3/article-p219.xml

25 Yao, J. & Tan, C.L. (2000). A case study on using neural networks to perform technical forecasting of forex. Neurocomputing 34(1-4):79-98. https://www.researchgate.net/publication/222517139_A_case_study_on_using_neural_networks_to_perform_technical_forecasting_of_forex

26 Hryshko, A. & Downs, T. (2004). System for foreign exchange trading using genetic algorithms and reinforcement learning. International Journal of Systems Science 35(13). https://www.researchgate.net/publication/43447963_System_for_foreign_exchange_trading_using_genetic_algorithms_and_reinforcement_learning

- Song (2017)[28] using evolutionary reinforcement learning fitting with various neurons and suggested using Genetic Algorithms in the optimization under Recurrent Reinforcement Learning suggested potential profitability on EURUSD.
- In the same year, Baasher and Fakhr[29] used classification and machine learning techniques to predict the foreign exchange market. Techniques used included support vector machines, bagging trees, maximally collapsing metric learning, neighborhood component analysis, class-based principal component analysis, and cluster-class-based and cluster-cluster based linear discriminant analysis.
- In 2019, Tsai and Wang[30] used deep reinforcement learning for trading foreign currencies and suggested profitability with the right choice of reward selection.
- In the same year Chihab, Bousbaa, Chihab and Bencharef[31] reviewed various algorithmic trading strategies including artificial neural networks, genetic algorithms, support vector machine, random forest. They focused their research on random forest and probit regression and proved that such machine learning techniques could be effective in improving prediction accuracy.

3 Methodology

3.1 Approach

We take a currency pair, sample the data based on ticks, use the triple barrier labeling method for the returns data, get technical indicators from the tick price and economic indicators and use their fractional differentiation for input into machine learning algorithms such as random forest and SVM.

We analyze the results, use an ensemble of the machine learning algorithms for prediction, and reiterate the process for the next currency pair.

3.2 Methodology

27 Carapuco, J.M.B (2017). Reinforcement Learning Applied to Forex Trading. Técnico Lisboa. https://www.sciencedirect.com/science/article/abs/pii/S1568494618305349

28 Song, Y. (2017). A Forex Trading System Using Evolutionary Reinforcement Learning. Worcester Polytechnic Institute. https://web.wpi.edu/Pubs/ETD/Available/etd-050117-071501/unrestricted/ysong.pdf

29 Baasher, A.A. and Fakhr, M.W. (2017). FOREX Trend Classification using Machine Learning Techniques. Computer Science Department, Arab Academy for Science and Technology Cairo. https://www.researchgate.net/publication/303080330_Forex_trend_classification_using_machine_learning_techniques

30 Tsai, Y.C & Wang, C.C. (2018). Deep Reinforcement Learning for Foreign Exchange Trading. https://arxiv.org/abs/1908.08036

31 Chihab, Y., Bousbaa, Z., Chihab, M., Bencharef, O. & Ziti, S. (2019). Algo-Trading Strategy for Intraweek Foreign Exchange Speculation Based on Random Forest and Probit Regression. https://www.hindawi.com/journals/acisc/2019/8342461/

Our overall methodology is defined by the underlying methodologies in each step. This includes:

- Overall machine learning flow – as the overall program on training machine learning models.
- Data preparation methodology – Where do we get the data, how our data is prepared, and what kinds of data transformation techniques are adopted during the data preparation stage.
- Machine learning methodology – What kinds of machine learning methodologies are adopted.

3.2.1 Overall machine learning flow

To build our machine learning algorithm, we would like to go through the following steps according to Google Machine Learning[32] This includes:

Step 1: Gather Data

Step 2: Explore Your Data

Step 2.5: Choose a Model

Step 3: Prepare Your Data

Step 4: Build, Train, and Evaluate Your Model. We are going to split our data set into 50%, 25%, and the remaining 25% for training, validation, and testing purposes respectively. Then we will perform cross-validation for our training set data.

Step 5: Tune Hyperparameters

Step 6: Deploy Your Model

3.2.2 Data preparation methodology

3.2.2.1 Source of data and data transformation

There are two types of data included in this research. These are foreign exchange spot data and the interest rate data.

For foreign exchange, we obtain the spot data from Forex Capital Markets (FXCM), between 22nd June 2014 to 22nd June 2020 with an interval of every two hours by Coordinated Universal Time (UTC). The following are the foreign currency pairs downloaded:

- AUDUSD
- AUDCAD

32 Machine Learning Guide (2018). Google. https://developers.google.com/machine-learning/guides/text-classification

- AUDJPY
- EURUSD
- GBPUSD
- NZDUSD
- USDCAD
- USDJPY

For interest-rate data, considering that the cross-currency swap market is priced based on a 3-month floating rate index as the market convention, we obtained the following interbank 3-month interest rate below.

Table 1 Currency Fixing Time

Currency	Interbank 3-month rate selected	Daily Fixing Time	Data Source
AUD	Bank Bill Swap Rate 3-month (BBSW)	Australian Eastern Standard Time 10:30am	Reserve Bank of Australia Daily Interest Rate data[33]
NZD	Bank Bill yield 3-month (BKBM)	New Zealand Standard Time 10:30am	Reserve Bank of New Zealand Wholesale interest rates - B2[34]
USD	London Interbank Offer Rate (LIBOR) 3-month	London Time 11:55am	Economic Research, Federal Reserve Bank of St. Louis [35]
JPY	London Interbank Offer Rate (LIBOR) 3-month	London Time 11:55am	Economic Research, Federal Reserve Bank of St. Louis [36]
GBP	London Interbank Offer Rate (LIBOR) 3-month	London Time 11:55am	Economic Research, Federal Reserve Bank of St. Louis [37]
EUR	London Interbank Offer Rate (LIBOR) 3-month	London Time 11:55am	Economic Research, Federal Reserve Bank of St. Louis [38]
CAD	Canadian Dollar Offered Rate (CDOR) 3-month, known as Banker Acceptance 3-month rate	Eastern Time 10:15am	2014 to 2018: Bank of Canada[39] 2018 and beyond: Investment Industry Regulatory Organization of Canada[40]

33 Statistical Tables - Interest Rates - Interest Rates and Yields – Money Market – Daily – F1. Reserve Bank of Australia. https://rba.gov.au/statistics/tables/#interest-rates . .

34 Wholesale interest rates - B2. Reserve Bank of New Zealand. https://www.rbnz.govt.nz/statistics/b2

35 3-Month London Interbank Offered Rate (LIBOR), based on U.S. Dollar (USD3MTD156N). Economic Research, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/USD3MTD156N

36 3-Month London Interbank Offered Rate (LIBOR), based on Japanese Yen (JPY3MTD156N). Economic Research, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/JPY3MTD156N

37 3-Month London Interbank Offered Rate (LIBOR), based on British Pound (GBP3MTD156N). Economic Research, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/GBP3MTD156N

38 3-Month London Interbank Offered Rate (LIBOR), based on Euro (EUR3MTD156N)Economic Research, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/EUR3MTD156N

39 Canadian Interest Rates and Monetary Policy Variables: 10-Year Lookup. Bank of Canada. https://www.bankofcanada.ca/rates/interest-rates/canadian-interest-rates/

40 Canadian Bankers' Acceptance (BA) Rates. Investment Industry Regulatory Organization of Canada. https://www.iiroc.ca/industry/marketmonitoringanalysis/Pages/BA-Rate.aspx

Given the spot data is available every two hours while the interbank interest rate data is available daily, we need to join the data by standardizing everything into UTC first. The following is the hours between the local time and UTC with the consideration of daily light saving in individual regions

Table 2 Hours ahead / behind UTC with and without daylight saving

Time zone	Hours ahead / behind UTC at start of daylight saving	Hours ahead / behind UTC at end of daylight saving
Australian Eastern Standard Time	+11	+10
New Zealand Standard Time	+13	+12
London Time	+1	0
Eastern Time	-4	-5

Then the interest rate r_t would be equal to the fixing rate where previous fixing time $\leq t <$ next fixing time. If the fixing rate is not available on a particular day, it is assumed that there is no change from the previous fixing rate.

3.2.2.2 Data differential methodology

In differentiation methodology, we try to implement fractional differentiation rather than the usual integer differentiation which allows us to make the data stationary while retaining some memory. Fractional differentiation is discussed by Prado [41] in 2018.

The following is the methodology for the fractional differential:

Let the backshift operator B applied to an observation X_t, where $B^k X_t = X_{t-k}$ for any integer $k \geq 0$.

For a backshift model, we are modelling the $(1-B)^d = \sum_{k=0}^{\infty} \frac{\prod_0^{k-1}(d-\mathrm{i})}{k!}(-\mathrm{B})^k$

$$= \sum_{k=0}^{\infty} (-\mathrm{B})^k \prod_{i=0}^{k-1} \frac{d-i}{k-i} = 1 - dB + \frac{d(d-1)}{2!}\mathrm{B}^2 - \frac{d(d-1)(d-2)}{3!}\mathrm{B}^3 + \cdots$$

3.2.2.3 Data labeling methodology

In data labeling methodology, we implement the triple-barrier method discussed by Prado [42] in 2018. It labels an observation using the first barrier touched out of three barriers. The following are the steps for labeling :

1. Set two horizontal barriers and one vertical barrier. Horizontal barriers are defined by profit-taking and stop-loss limits as a function of the estimated volatility.

---

41 Prado, M. L. D. (2018). Section 5 Fractionally Differentiated Features. Advances in Financial Machine Learning. ISBN 978-1-119-48211-6

42 Prado, M. L. D. (2018). Section 3.4 The Triple-Barrier Method. Advances in Financial Machine Learning. ISBN 978-1-119-48211-6

2. The third barrier is the number of bars elapsed since the position was taken as the limit for expiration.

3. Then:

   a). If the upper barrier is hit first, the observation is labeled as 1.
   b). If the vertical barrier is hit first the observation is labeled as the sign of the return, or 0.
   c). If the lower barrier is hit first, the observation is labeled as -1.

The feature of the triple-barrier method is that it is path-dependent. To label an observation, we need to look back at the previous observations in the vertical barrier. The following is the graphical illustration by Prado[43].

Figure 1 The Triple-barrier Method Illustration by Prado.

We are implementing the three barriers labeling as follows:

1. We first set the horizonal barrier as 70% of the mean of the estimated daily volatility.

   We would calculate the one-period return using the previous bid against the current ask as the short position return, and using the previous ask against the current bid as the long position return. Once the cumulative long position or short position goes above or below the mean daily volatility, we would label this as an event.

[43] Prado, M.L.d (2018). The 10 Reasons Most Machine Learning Funds Fail. http://www.smallake.kr/wp-content/uploads/2018/07/SSRN-id3104816.pdf

2. Once we have an event, we take the third barrier as the 3 days after the time of the event.

3. Then:

   a). If the upper barrier at 70% of the daily volatility is hit first, the observation is labeled as 1.
   b). If the vertical barrier as the 3-day period is hit first the observation is labeled as the sign of the return, or 0.
   c). If the lower barrier at -1 x 70% of the daily volatility is hit first, the observation is labeled as -1.

3.2.2.4 Technical indicators

For technical indicators, we select the following technical indicators:

1. **The Exponential moving average (EMA) -** We will be taking the 3-period exponential moving average as fast MA and 7-period exponential moving average as slow MA.

2. **Bollinger bands -** We will be using the 20-period interval with 2 standard deviations as the standard configuration of Bollinger bands.

3. **Commodity Channel Index -** Here we perform a modification on the typical price where we replace this with the close price.

4. **Stochastics** –We will be using the 14-period interval for %K and %D as the standard configuration of stochastics.

5. **Williams %R -** We will be using 14-period interval as the lookback period.

6. **Ichimoku -** We will be using the standard configurations mentioned above.

7. **Relative Strength Index (RSI) -** As the well-known technical indicators and prior research suggested RSI could still be a profitable indicator with some twists in parameters. We will be using the 14-day period, a value of 70 as the overbought signal, and a value of 30 as the oversold signal.

3.2.2.5 Feature engineering

From the above, we are going to derive four sets of features as our trading signals to apply machine learning on.

Features where fractional differencing is applied: These include

- **Price:** price_frdif
- **Exponential moving average:** slow_frdif, fast_frdif,
- **Bollinger bands:** average_frdif, lower_band_frdif, upper_band_frdif,
- **Ichimoku:** kijun_sen_frdif, tenka_sen_frdif, senkou_span_b_frdif, senkou_span_a_frdif

1. **Features where fractional differencing is not applied:** These include

- **Price:** price
- **Exponential moving average:** fast, slow, ema_side,
- **Bollinger bands**: average, upper_band, lower_band, standard_deviation, bb_side
- **Stochastics:** %K, %D, so_side
- **Commodity Channel Index:** CCI
- **Williams %R :** wr, wr_side
- **Ichimoku:** tenka_sen, kijun_sen, senkou_span_a, senkou_span_b, chikou_span. ic_side
- **Relative Strength Index:** RSI, rsi_side
- **Lagged price:** T-1, T-2, T-12
- **Auto correlation:** Autocor_1_lag, Autocor_2_lag, Autocor_4_lag, Autocor_6_lag
- **Lagged 1-period return:** T-1_1per_rtn, T-2_1per_rtn, T-12_1per_rtn
- **Lagged period returns**: T-1_rtn, T-2_rtn, T-12_rtn,
- **Interest rate differentials:** between currencies: ir_d1
- Day of week indicators: Monday, Tuesday, Wednesday Thursday, Friday, Sunday,
- **Month indicators:** January, February, March, April, May, June, July, August, September, October, November, December
- **End of month indicator:** EOM

2. **Full set of features** as the union set of (1.) and (2.) above.

3. **Features pre-selected by a random forest with 5-split cross-validation in the training set.** The pre-selection will be based on the full set of features.

3.3 Dimension reduction with Principal Component Analysis (PCA)

PCA is a dimensionality-reduction method that is used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one while preserving as much information as possible[44].

Let $\mathrm{X} = (\mathrm{X}_1, \mathrm{X}_2, \dots \mathrm{X}_n)^T$ be an n-dimensional random vector.
Let $\Sigma = \mathrm{E}[\mathrm{XX}^T]$ is the variance-covariance matrix of X.
Let $\mathrm{Y} = \mathrm{AX}$ where $\mathrm{A} \epsilon \mathbb{R}^{n \times n}$ is the matrix to be determined where $\mathrm{Cov(Y)} = \mathrm{A}\Sigma\mathrm{A}^T$.

To find Y, we will use the spectral decomposition technique where we let $\Sigma = \mathrm{V}\Lambda\mathrm{V}^T$ where Λ is a diagonal matrix diag $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$are the eigenvalues of Σ. V is an orthogonal matrix with the i-th column $\mathrm{V}_i \ \epsilon \ \mathbb{R}^n$ being the i-th standardized eigenvector of Σ.

Let $\Sigma = \mathrm{V}\Lambda\mathrm{V}^T$ be the spectral decomposition of Σ and also let $\mathrm{Y} = \mathrm{V}^T\mathrm{X}$. The covariance matrix of Y is Cov(Y) = $\mathrm{V}^T\Sigma\mathrm{V} = \mathrm{V}^T\mathrm{V}\Lambda\mathrm{V}^T\mathrm{V} = \Lambda$. Given Λ is a diagonal matrix, the

[44] Jaadi, Z. (2020). A Step by Step Explanation of Principal Component Analysis. https://builtin.com/data-science/step-step-explanation-principal-component-analysis

components of Y are uncorrelated and Y is then called the principal component matrix of X. We could see $\mathrm{Var}(\mathrm{Y}_i) = \lambda_i$.

Then we could consider how principal component Y vectors could capture the variance of X. $\sum_{i=1}^{n} Var(\mathrm{Y}_i) = \sum_{i=1}^{n} \lambda_i = trace(\Lambda) = trace(\mathrm{V}^T\Sigma\mathrm{V}) = trace(\mathrm{V}\mathrm{V}^T\Sigma) = trace(\Sigma) = \sum_{i=1}^{n} Var(\mathrm{X}_i)$ to restore the variance of X.

Given we are using the intraday data with a 5-year observation period, we try to adopt principal component analysis to reduce the data dimension where we train models with the following principal component analysis application:

- Random Forest with 4 principal components of the features,
- Support Vector Machine with 4 principal components of the features,
- Support Vector Machine with 20 principal components of the features
- Naive Bayes classifier with 6 principal components of the features.

3.3.1 Machine learning methodology for classification.

We look at the below supervised machine learning classification models.

3.3.1.1 Logistic Regression

It is used to classify the dependent variable based on a sigmoid function applied to a linear combination of independent variables and a decision boundary to the output of this function. Logistic regression looks similar to the multivariable linear regression, with dependent variable Y replaced by the log-odds. The following is the logistic regression formulations for n independent variables:

$$log\frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n, \text{ where } p = P(Y = k)$$

The log-odds could be converted into odds by taking exponential on both sides.

$$\frac{p}{1-p} = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n}$$

By putting the term into the same side, we could get

$p = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n}}{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + 1}} = \frac{1}{1+e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}}$, which is the sigmoid function giving $0 \leq p \leq 1$. Given the probability is found in the form of dependent variable Xs, the beta coefficients can be estimated by using various techniques, such as ordinary least squares and maximum likelihood estimation.

3.3.1.2 Naive Bayes

It is used to classify the output variable based on the input variables using Bayes theorem and the assumption that given the value of the output variable, the input variables are independent of each other[45]. This is based on the conditional probability of the probability of a particular category depending on its independent variables $\boldsymbol{x}$, i.e. $p(C_k|\,\boldsymbol{x} = x_1, x_2, \ldots, x_n)$.

Using Bayes' theorem the conditional probability is $p(C_k|\,\boldsymbol{x}) = \frac{p(C_k)p(x|C_k)}{p(x)}$. Here we could find the joint distribution of the category and independent variables $\boldsymbol{x}$ using the chain rule of conditional probability.

$$p(C_k, x_1, x_2, \ldots, x_n) =$$
$$p(\,x_1, x_2, \ldots, x_n, C_k) = p(\,x_1|x_2, \ldots, x_n, C_k)\,p(x_2, \ldots, x_n, C_k) =$$
$$p(\,x_1|x_2, \ldots, x_n, C_k)\,p(x_2|\,x_3, \ldots, x_n, C_k)p(x_3, \ldots, x_n, C_k) =$$
$$p(\,x_1|x_2, \ldots, x_n, C_k)\,p(x_2|\,x_3, \ldots, x_n, C_k)\ \ldots.\,p(x_n|\,C_k)p(C_k)$$

Assuming conditional independence where all features in $\boldsymbol{x}$ are mutually independent, conditional on the category C_k as $\boldsymbol{x}$ are the independent variables, we could see $p(\,x_i|x_{i+1}, \ldots, x_n, C_k) = p(\,x_i|C_k)$.

Therefore we could see $p(C_k|\,\boldsymbol{x}) \propto\ p(C_k)p(x_1|\,C_k)p(x_2|\,C_k) \ldots p(x_n|\,C_k) = p(C_k)\prod_{i=1}^{n} p(x_i|\,C_k)$

Back to our conditional probability $p(C_k|\,\boldsymbol{x}) = \frac{p(C_k)p(x|C_k)}{p(x)}$, we could see $p(C_k|\,\boldsymbol{x}) = \frac{p(C_k)\prod_{i=1}^{n} p(x_i|\,C_k)}{p(x)} = \frac{p(C_k)\prod_{i=1}^{n} p(x_i|\,C_k)}{\sum_k\, p(C_k)p(x|C_k)}$

The corresponding Bayes classifier would be $\hat{y} = \operatorname{argmax} p(C_k)\prod_{i=1}^{n} p(x_i|\,C_k)$.

3.3.1.3 Overview of Decision tree

We would like to first discuss decision trees below following Wahlstorm lecture on Tree-based methods, Bagging and Random Forest[46].

Decision trees can be applied to both regression and classification.

Regression trees look at the mean of training data within a given region after they partition the data into L regions, $\mathrm{R}_1, \mathrm{R}_2, \mathrm{R}_3, \ldots, \mathrm{R}_L$.

The prediction model is $\widehat{y_*} = \sum_{l=1}^{L} \widehat{y_l}\, \mathbb{I}\{x_* \in R_l\}$ where $\mathbb{I}\{x_* \in R_l\}$ is the indicator function

[45] Gandhi, Rohith (2018). Naive Bayes Classifier. https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c

46 Wahlstrom, N. (2020). Lecture 6 – Tree-based methods, Bagging and Random Forests, Statistical Machine Learning, Division of Systems and Control, Department of Information Technology, Uppsala University. http://www.it.uu.se/edu/course/homepage/sml/lectures/Lecture6_handout.pdf

$$\mathbb{I}\{x_* \in R_l\} = \begin{cases} 1, \text{if } x \in R_l \\ 0, \text{if } x \notin R_l \end{cases}$$

and $\hat{y}_l$ is a constant prediction within each region. For regression tree $\hat{y}_l = average\,\{y_i : x_* \in R_l\}$.

For each decision tree, we select a random subset of features at each node to decide the optimal split.

Classification trees are similar to regression trees with two differences:

First, the class prediction for each region is based on the proportion of data points for each class in the region where.

$$\widehat{z_{lm}} = \frac{1}{nl} \sum_{i:x_i \in R_l} \mathbb{I}\{y_i \in m\}$$

The above defines the proportion of training observations in the **i**-th region that belong to the **m**-th class. Then the probability can be appropriated by

$$p(y = m \mid x) \approx L \sum_{l=1} \widehat{z_{lm}} \mathbb{I}\{x \in R_l\}$$

Here we would be using the following classifier:

3.3.1.3.1 Extreme Gradient Boosting

Extreme Gradient Boosting, known as XGBoost, as a training algorithm implementing gradient boosting decision tree algorithm[47]. Gradient boosting produces a prediction model in the form of an ensemble of decision trees where data points for which model is weaker are more likely to be included in building subsequent decision trees. XGBoost is an implementation of gradient boosting that uses a regularized model formalization to control over-fitting and utilizes parallel computation on a single machine[48].

3.3.1.3.2 Light Gradient Boosted Machine

Light Gradient Boosted Machine, known as LightGBM is a training algorithm on decision trees that increases the efficiency of the model and reduces memory usage[49]. It splits the tree leaf wise unlike other algorithms which split the tree level wise. The leaf-wise

---

[47] Khandelwal, P.(2017). Which algorithm takes the crown: Light GBM vs XGBOOST? https://www.analyticsvidhya.com/blog/2017/06/which-algorithm-takes-the-crown-light-gbm-vs-xgboost/

[48] Nishida, K. (2017). Introduction to Extreme Gradient Boosting in Exploratory. https://blog.exploratory.io/introduction-to-extreme-gradient-boosting-in-exploratory-7bbec554ac7

[49] Khandelwal, P.(2017). Which algorithm takes the crown: Light GBM vs XGBOOST? https://www.analyticsvidhya.com/blog/2017/06/which-algorithm-takes-the-crown-light-gbm-vs-xgboost/

algorithm (growing on the same leaf) can reduce more loss than the level-wise algorithm and results in much better accuracy. It is very fast so it is called "Light".

3.3.1.3.3 Adaptive Boosting

Adaptive Boosting (AdaBoost) first builds a model on a subset of the data with all observations being given equal weights[50]. Errors are calculated based on the predictions made on the whole dataset (from the actual values). Higher weights are given to the datapoints predicted incorrectly and a new model is calculated. This process is repeated until the error function converges or the maximum limit for the number of estimators is reached.

3.3.1.3.4 Random Forest

Random forest fits decision trees on different bootstrap random samples.

One of the decision tree problems would be when to stop the split. The problem of too many steps of splitting would lead to a model with too many leaves but with large variance and overfitting to a training set which results in less predictive power with the test or production data. Therefore random forests correct for overfitting of decision trees.

Random forests use the bagging technique where we train a separate deep tree $\widehat{y^b}(x)$ for each independent data 1, 2, 3, …, B

$$y_{\mathrm{bag}}\widehat{(x)} = \frac{1}{B}\sum_{b=1}^{B}\widehat{y^b}(x)$$

We could see the variance is reduced by the factor B by averaging.

Random forest is constructed by bagging and for each split, in each tree only a random subset $q \leq p$ inputs are considered as splitting variables.

3.3.1.3.5 Extremely Randomized Trees

Extremely Randomized Trees, known as Extra trees, work similar to random forests with two key differences[51]. It consists of randomizing strongly both attribute and cut-point choice while splitting a tree node[52].

3.3.1.4 **Support Vector Machine**

---

[50] Albert, S. (2018). Boosting with AdaBoost and Gradient Boosting. https://medium.com/diogo-menezes-borges/boosting-with-adaboost-and-gradient-boosting-9cbab2a1af81#:~:text=Boosting%20with%20AdaBoost%20and%20Gradient%20Boosting%201%20Adaptive,is%20an%20advanced%20implementation%20of%20the%20Gradient%20Boosting.

[51] Bhandari (2018), ExtraTreesClassifier. Medium.com. https://medium.com/@namanbhandari/extratreesclassifier-8e7fc0502c7

[52] Ernst, D. Geurts, P. & Wehenkel, L. (2006). Extremely randomized trees. https://link.springer.com/content/pdf/10.1007/s10994-006-6226-1.pdf

Following Berwick Lectures in Artificial Intelligence in 2011[53], Support Vector Machine (SVM) is a supervised learning algorithm that is used to learn a hyperplane that can solve the binary classification problem. Support vectors are the data points that lie closest to the decision surface as the data points most difficult to classify. SVM maximizes the margin between the classes by defining the decision surface in the form of:

$$\mathbf{w}^T\mathbf{x} + b = 0, \text{where}$$

$\mathbf{w}$ is a weight vector, $\mathbf{x}$ is input vectors, $\mathbf{b}$ is bias.

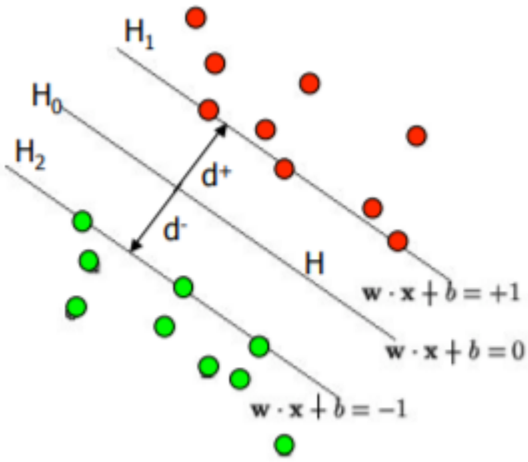
Therefore the margin of separation d as the separation between the hyperplane and the closest data point for a given weight vector $\mathbf{w}$ and bias $\mathbf{b}$ where

$$\mathbf{w}^T\mathbf{x} + b \geq 0 \; for \; d_i = 1$$

$$\mathbf{w}^T\mathbf{x} + b \leq 0 \; for \; d_i = -1$$

The following is an illustration of how the support vectors look like in a 2-dimensional space.

Figure 2 Support Vector Illustration in 2-dimensional space



In order to maximize the margin d we could alternatively minimize $\|\mathbf{w}\|$ with the condition that there are no data points between hyperplane 1 and hyperplane 2. This becomes a quadratic programming problem to minimize $\frac{1}{2}||\mathbf{w}||$ for $\mathbf{y}_i(\mathbf{w}^T\mathbf{x}_i) - b - 1 = 0$ as a constrained optimization problem.

3.3.1.5 **Voting**

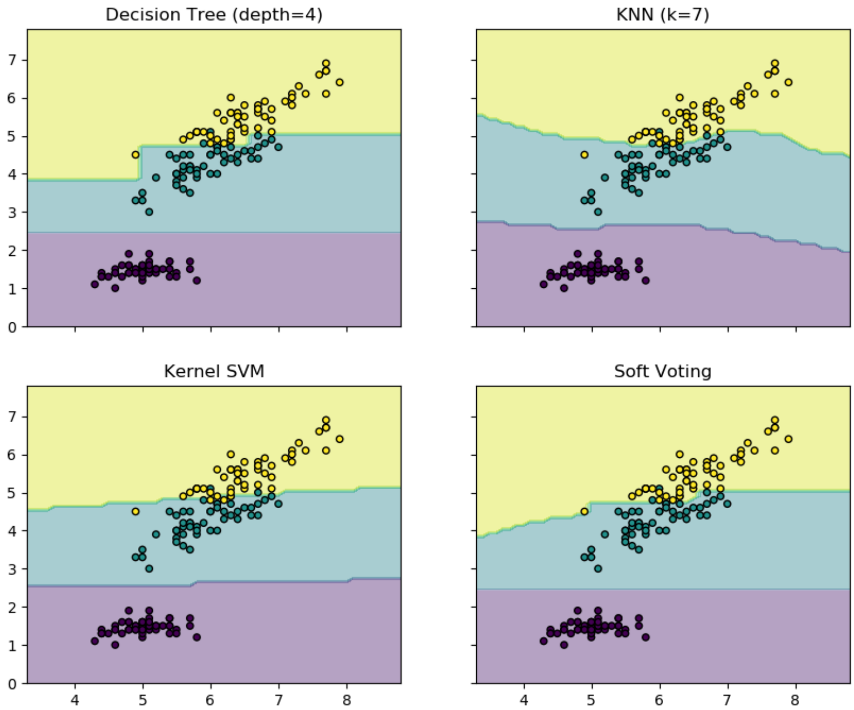
---

53 Berwick, B. (2011), An Idiot's guide to Support vector machines (SVMs), 6.034 Artificial Intelligence. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology. http://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf

Voting classifier takes the input predicted by various classifiers and assigns a class probability for each of the class region. Then the voting classifier would take the average of the class probabilities as the final class probability.

The following is a graphical illustration on scikit-learn website[54] for using three classifiers - decision tree with depth 4, k-nearest neighbors with number of neighbors being 7, kernel support vector machine, and how a soft voting result looks like with the three classifiers as the input.

Figure 3 Voting Classifier Illustration from scikit-learn



3.3.1.6 List of Machine Learning used

We are going to apply the following machine learning models:

1. Light gradient boosted machine
2. Extreme gradient boosting
3. Random forest

---

[54] Scikit learn. Plot the decision boundaries of a VotingClassifier. https://scikit-learn.org/stable/auto_examples/ensemble/plot_voting_decision_regions.html#sphx-glr-auto-examples-ensemble-plot-voting-decision-regions-py

4. Random forest with features pre-selected by another Random Forest
5. Random forest with 4 principal components from the features
6. Extremely randomized trees
7. Adaptive boosting
8. Logistic regression with 4 principal components from the features
9. Support vector machine
10. Support vector machine with 4 principal components from the features
11. Support vector machine with 20 principal components from the features
12. Naïve Bayes
13. Naïve Bayes with 6 principal components from the features
14. Soft voting with (3), (6), (8), and (10) above.

3.4 Desired outcomes

The desired outcome for this paper is to identify a profitable trading strategy in G10 currencies. The following is the way in which we select our machine learning model for training phase, validation phase, and testing phase.

3.4.1 Training phase

In the training set we are going to identify the models with relatively high accuracy under cross-validation, together with higher returns, lower volatilities, and higher Sharpe ratio.

3.4.2 Testing phase

In the testing set we are going to again look at trained model performance using validation set data with the model identified in the validation stage and examined by testing set data by using returns, volatilities, and Sharpe ratio to see if the performance persists in out-sample testing.

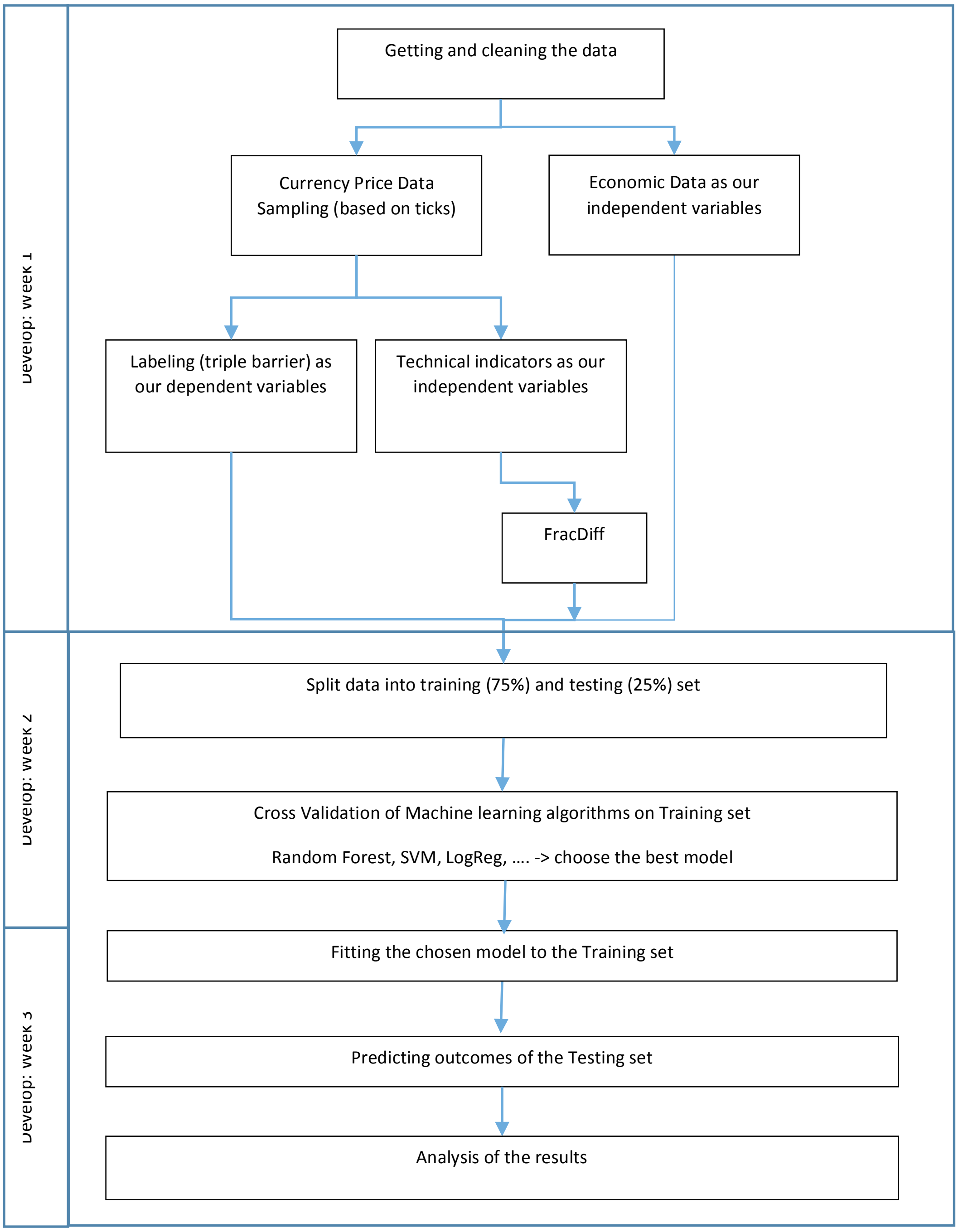
3.5 Intended working plan

We are going to split our work within the three weeks of developing our methodology as follows:

Step 1: Gather Data – Week 1
Step 2: Explore Your Data – Week 1
Step 2.5: Choose a Model – Week 1
Step 3: Prepare Your Data – Week 1
Step 4: Build, Train and Evaluate Your Model – Week 2/3
Step 5: Tune Hyperparameters – Week 2/3
Step 6: Deploy Your Model – Week 3

The below diagram outlines how we conducted our research study:

Figure 4 Development Plan



4 Results

The following are our results:

4.1.1 Training phase

The following are the testing phase results with 75% of the data with 5-split cross-validation.

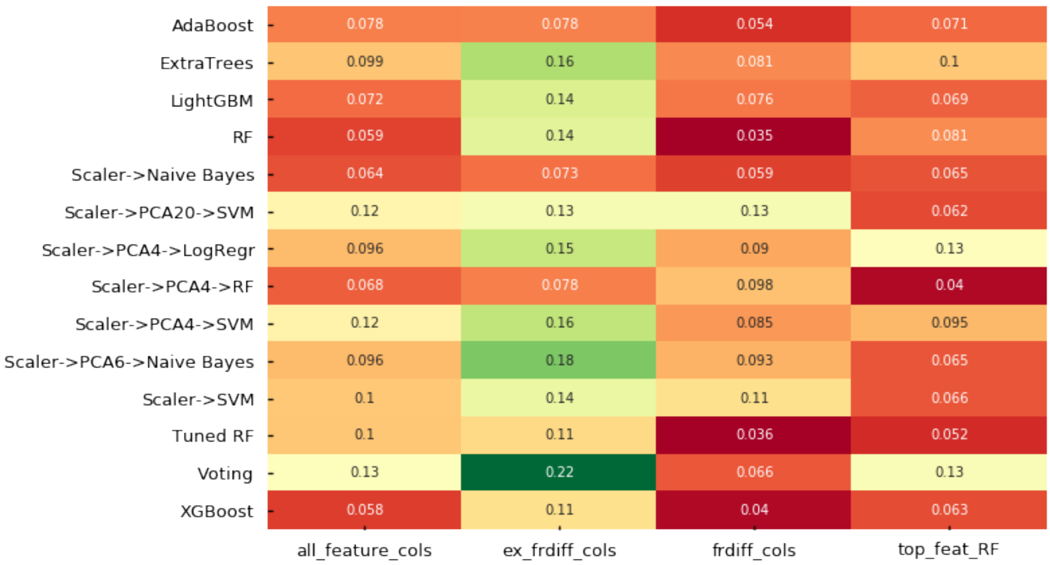
Figure 5 Average Accuracy under 5-split cross-validation for the training set

For accuracy:

- From a model perspective, it is observed that most models have a similar accuracy score with the exception of Naive Bayes classification.

- From a features perspective, we do not see a significant difference for one feature set against the others.

Figure 6 Average returns across currency pairs for the training set



For return across currency pairs:

- From a model perspective, it is observed that extra trees, logistic regression with 4 principal components, support vector machine with 4 principal components, support vector machine with 20 principal components, and voting classifier performs relatively better than other algorithms.

- From a feature perspective, it is observed that models with features without undergoing fractional differencing performs better than other feature sets.

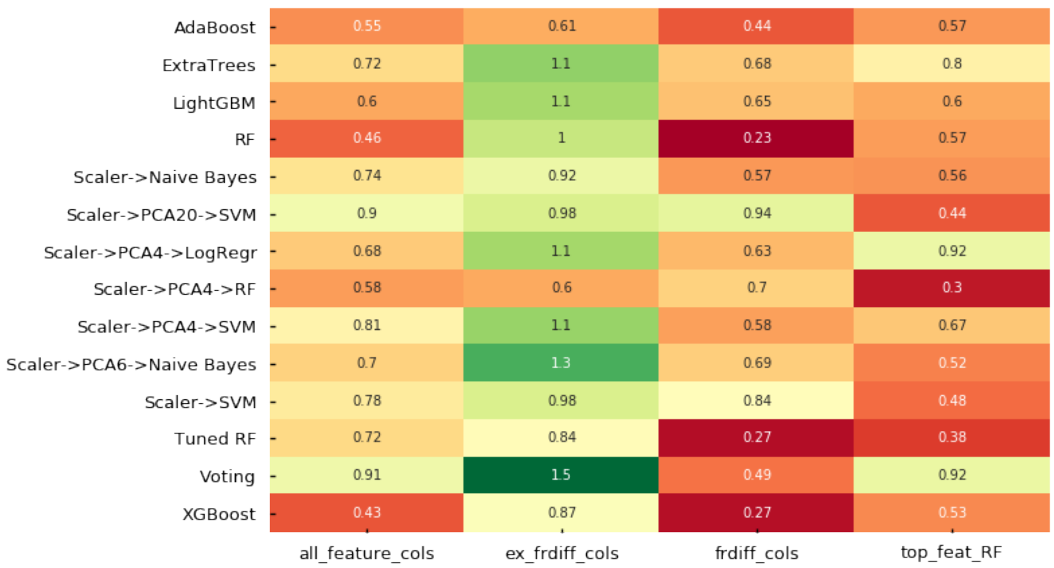
Figure 7 Average volatilities across currency pairs for the training set

For standard deviations across currency pairs:

- From a model perspective, it is observed that the Naive Bayes model gives minimal standard deviations while other models give similar level of volatility.

- From a features perspective, we do not observe significant differences across features with the exception of the feature set with Naïve Bayes model.

Figure 8 Sharpe ratios across currency pair for training set



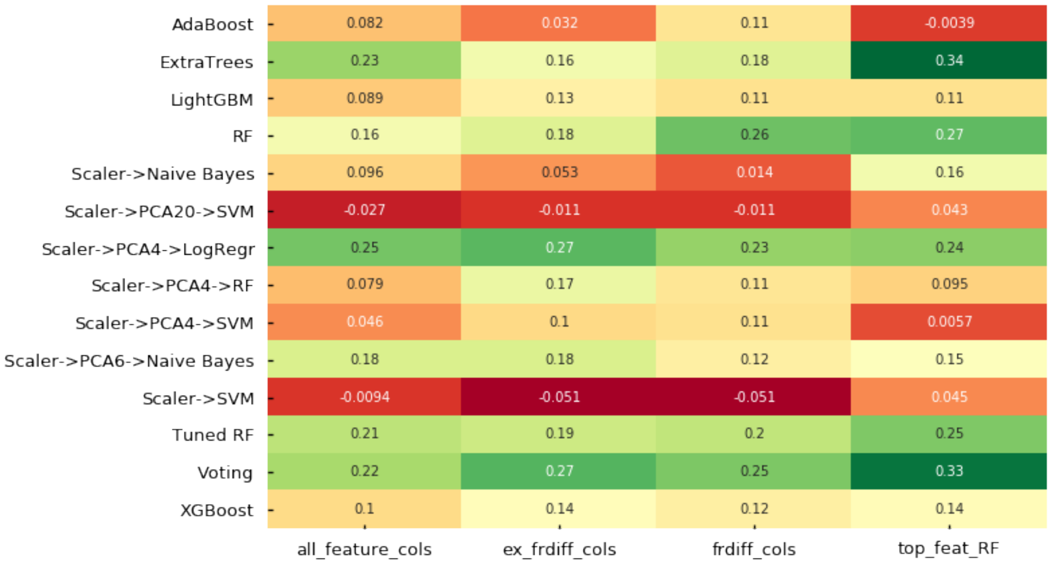
For Sharpe Ratio across currency pairs:

- From a model perspective, given the volatility is largely consistent across features, similar to the observations we see from the return, extra trees, logistic regression with 4 principal components, and voting classifier perform relatively better, followed by support vector machine with 20 principal components, and support vector machine with 4 principal components than other algorithms.

- From a features perspective, given the volatility is largely consistent across features, it is observed that models with features without fractional differencing perform better than other feature sets.

4.1.2 Testing set phase

Figure 9 Average returns across currency pairs for the testing set

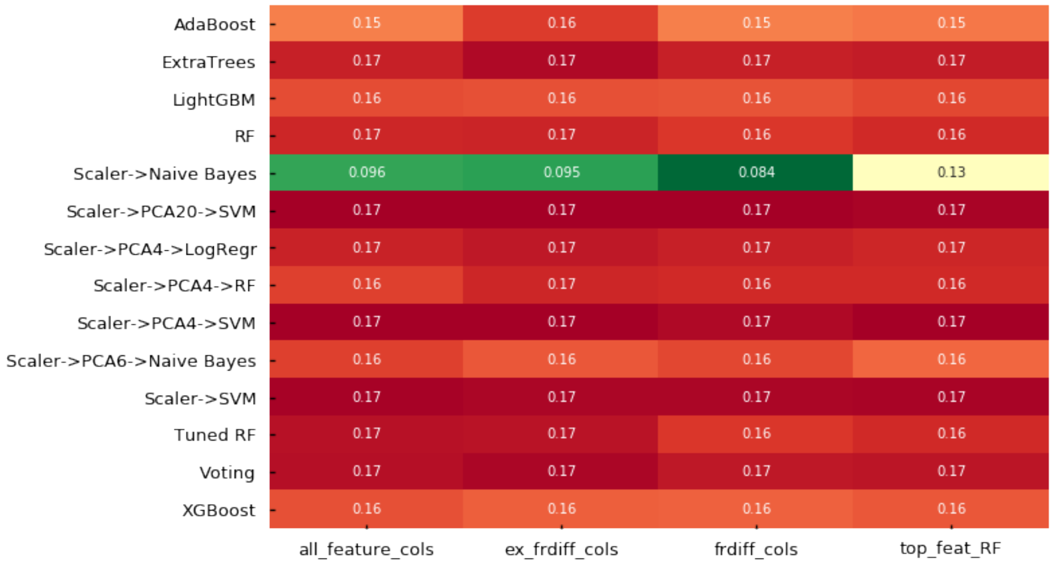


For returns across currency pairs:

- From a model perspective, it is observed that voting classifier, logistic regression with 4 principal components, and extra trees as our candidate model perform well, with Voting classifier working the best, followed by logistic regression with 4 principal components, and extra trees. Support vector machine with 4 principal components still gives fair prediction results while support vector machine with 20 principal components performs poorly, suggesting potential overfitting and data mining bias with 20 principal components.

- From a features perspective, we could see models with features not undergoing fractional differencing perform fairly with other feature sets.

Figure 10 Average volatilities across currency pairs for the testing set



For standard deviations across currency pairs:

- From a model perspective, similar to cross-validation results, it is observed that the Naive Bayes model gives minimal standard deviations while other models give similar levels of volatility.

- From a features perspective, similar to cross-validation results, we do not observe significant differences across features with the exception of the feature set with the Naïve Bayes model.

Figure 11 Average Sharpe ratio across currency pairs for the testing set

For Sharpe Ratio across currency pairs:

- From a model perspective, given that the volatility is largely consistent across features, it is observed that voting classifier, logistic regression with 4 principal components and extra trees as our candidate models perform well. Here voting classifier works best, followed by logistic regression with 4 principal components, and then extra trees. Support vector machine with 4 principal components gives below-average prediction results while support vector machine with 20 principal components performs poorly, suggesting potential overfitting and data mining bias with 20 principal components.

- From a feature perspective, we could see models with features not undergoing fractional differencing perform fairly with other feature sets.

The above reinforced our observations in the cross-validation during this out-of-sample testing. It is concluded the voting classifier, logistic regression with 4 principal components as our candidate models performs best while extra trees performs well in the out-sample testing, among the same set of features without undergoing fractional differencing.

Considering the consistency of model performance between the validation and out-sample testing phase, we will be using the voting classifier, logistic regression with 4 principal components with features without undergoing fractional differencing.

5 Discussion

Here we are going to drill into how the model performs for individual currency pairs form our backtesting results. We are leveraging the existing pyfolio package to produce the backtesting results for us, which may have deviation between our implementation for annualized return, annualized volatility, and Sharpe ratio presented above. It is noted that our backtesting period covers the COVID-19 market crash in March, which could be one way to look at the model robustness in handling crisis scenarios which the model may not be trained specifically for.

5.1.1 Logistic regressions with 4 principal components with features not applied for fractional differencing

The following is the backtesting results across currencies from logistic regressions with 4 principal components with top features pre-selected by random forest.

	AUD/USD	AUD/CAD	AUD/JPY	EUR/USD	GBP/USD	NZD/USD	USD/CAD	USD/JPY	Average
Annual return	40%	1%	-4%	7%	45%	40%	23%	61%	27%
Cumulative returns	33%	1%	-3%	5%	40%	33%	19%	44%	22%
Annual volatility	23%	17%	21%	12%	16%	20%	11%	14%	17%
Sharpe ratio	1.59	0.17	-0.07	0.66	2.37	1.78	1.88	3.37	1.47
Calmar ratio	1.35	0.06	-0.09	0.78	3.54	2.12	1.85	6.93	2.07
Stability	0.52	0.03	0.14	0.47	0.89	0.50	0.48	0.78	0.48
Max drawdown	-0.30	-0.26	-0.40	-0.09	-0.13	-0.19	-0.13	-0.09	-0.20
Omega ratio	1.26	1.03	0.99	1.10	1.43	1.30	1.31	1.66	1.26
Sortino ratio	2.49	0.24	-0.10	1.00	3.80	2.80	2.95	5.79	2.37
Skew	0.09	-0.09	-0.05	0.17	-0.05	0.00	0.02	0.06	0.02
Kurtosis	-0.66	-0.23	-0.68	-1.06	-0.26	-0.61	-1.05	-0.57	-0.64
Tail ratio	1.24	0.98	0.93	1.09	1.22	1.19	1.26	1.32	1.15
Daily value at risk	-3%	-2%	-3%	-1%	-2%	-2%	-1%	-2%	-2%

From the above summary, we could observe that all USD cross pair performs quite well while EUR/USD would be the worst among them. Non-USD cross pair performs badly.

When looking into the cumulative return time series graphs in the Reference section, it is observed that:

1. A consistent pattern during the COVID market crash where all the currency pair returns shedded during March. This is suggested the model does not perform during the market crash, which is expected with our validation data for training is more on business-as-usual trading days.

2. A strong return for most currency pairs in the post-COVID period. This is because the economic conditions and the foreseeable economic recovery gives the investor directions to invest in the foreign market, driving clear trends in the spot price movement where our model can detect the technical signals from the model training.

3. JPY pairs (USDJPY and AUDJPY) perform consistently in the pre-COVID crisis and generates positive returns. However, AUDJPY performs badly in the COVID market crash and was not able to recover to positive cumulative returns in the backtesting period.

From the above, we could further analyze and test on trading the USDJPY pair given it consistency performance between pre-COVID and post-COVID period.

5.1.2 Voting Classifier

	AUD/USD	AUD/CAD	AUD/JPY	EUR/USD	GBP/USD	NZD/USD	USD/CAD	USD/JPY	Average
Annual return	38%	7%	-5%	19%	45%	29%	42%	42%	27%
Cumulative returns	32%	6%	-4%	14%	40%	24%	34%	31%	22%
Annual volatility	23%	17%	23%	11%	17%	21%	11%	14%	17%
Sharpe ratio	1.53	0.46	-0.09	1.61	2.31	1.33	3.12	2.50	1.60
Calmar ratio	1.29	0.26	-0.11	2.77	3.66	1.64	6.97	4.21	2.59
Stability	0.52	0.00	0.15	0.66	0.69	0.50	0.84	0.51	0.48
Max drawdown	-0.30	-0.26	-0.40	-0.07	-0.12	-0.18	-0.06	-0.10	-0.19
Omega ratio	1.25	1.07	0.99	1.25	1.41	1.22	1.56	1.46	1.28
Sortino ratio	2.40	0.67	-0.13	2.60	3.80	2.03	5.02	4.29	2.59
Skew	0.10	-0.05	-0.05	0.20	0.08	0.02	-0.15	0.25	0.05
Kurtosis	-0.66	-0.23	-0.88	-1.08	-0.46	-0.51	-0.97	-0.62	-0.68
Tail ratio	1.24	0.99	1.06	1.15	1.34	1.14	1.26	1.35	1.19
Daily value at risk	-3%	-2%	-3%	-1%	-2%	-2%	-1%	-2%	-2%

From the above summary, we could observe that there is an improvement for annualized return for AUD/CAD, EUR/USD and USD/CAD while deterioration in other currency pairs.

1. Similar to what was observed in the logistic regression model, a consistent pattern during the COVID market crash where all the currency pair returns shedded during March. This is suggested the model does not perform during the market crash, which is expected with our validation data for training is more on business-as-usual trading days.

2. Similar to what was observed in the logistic regression model, a strong return for most currency pairs in the post-COVID period. This is because the economic conditions and the foreseeable economic recovery gives the investor directions to invest in the foreign market, driving clear trends in the spot price movement where our model can detect the technical signals from the model training.

3. Comparing between the logistic regression model and voting classifier, it is observed the cumulative return has a similar pattern within each currency. Recalling the voting classifier is getting the average class probability from the four selected classified of which logistic regression model is one of them, this is suggested that the logistic regression model has a core effect on the voting classifier while the model improves the prediction results by three other models reinforcing the signals.

4. USDJPY pair also consistently performs in the pre-COVID crisis under the voting classifier.

6 Conclusion

We trained various machine learning models with technical indicators and also interest rate differentials as our fundamental data as our independent variables, to classify trading signals which help generate positive returns. For our independent variables, we also made use of various approaches, including fractional differencing, its counterpart, and also features pre-selected by a random forest to see if such an ensemble of machine learning models would improve the prediction results.

From our machine learning model flow, we successfully identified models with relatively good, and also most importantly consistent performance in the training phase with cross-validation and out-sample testing. Models selected after the out-sample testing phase, which are logistic regression with 4 principal components, and voting classifier with selected 4 models have 27% annualized average return each across currency pairs, and average Sharpe ratio of 1.47 and 1.6 across currency pairs respectively.

We also performed model backtesting for each of the currency pairs to look at the model performance with explicit consideration of the model performance for the pre-COVID market crash as business-as-usual trading days and the COVID market-crash as our out-sample testing period to examine for model robustness. We observed that both models selected did not perform during the COVID market-crash as expected where our training data for the model covers business-as-usual trading days. Both models could be performing by examining the pre-COVID and post-COVID market crashes with USDJPY currency pairs to enhance the average returns and Sharpe ratio across selected currency pairs as future works.

Disclaimer

7 References

3-Month London Interbank Offered Rate (LIBOR), based on British Pound (GBP3MTD156N). Economic Research, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/GBP3MTD156N

3-Month London Interbank Offered Rate (LIBOR), based on Euro (EUR3MTD156N). Economic Research, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/EUR3MTD156N

3-Month London Interbank Offered Rate (LIBOR), based on Japanese Yen (JPY3MTD156N). Economic Research, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/JPY3MTD156N

3-Month London Interbank Offered Rate (LIBOR), based on U.S. Dollar (USD3MTD156N). Economic Research, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/USD3MTD156N

Albert, S. (2018). Boosting with AdaBoost and Gradient Boosting. https://medium.com/diogo-menezes-borges/boosting-with-adaboost-and-gradient-boosting-9cbab2a1af81#:~:text=Boosting%20with%20AdaBoost%20and%20Gradient%20Boosting%201%20Adaptive,is%20an%20advanced%20implementation%20of%20the%20Gradient%20Boosting.

Anderson, B. & Li, S. (2015). An investigation of the relative strength index. Banks and Bank Systems, Volume 10, Issue 1, 2015. https://www.researchgate.net/publication/283691105_An_investigation_of_the_relative_strength_index

Baasher, A.A. and Fakhr, M.W. (2017). FOREX Trend Classification using Machine Learning Techniques. Computer Science Department, Arab Academy for Science and Technology Cairo. https://www.researchgate.net/publication/303080330_Forex_trend_classification_using_machine_learning_techniques

Bartkus, C. (2018). Using the relative strength index for active investments in the foreign exchange market. http://doi.prz.edu.pl/pl/pdf/zim/358

Berwick, B. (2011), An Idiot's guide to Support vector machines (SVMs), 6.034 Artificial Intelligence. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology. http://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf

Bettendorf, T. & Heinlein, R. (2019). Connectedness between G10 Currencies: Searching for the Causal Structure. Discussion Paper Deutsche Bundesbank No 06/2019. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3337431

Bhandari (2018), ExtraTreesClassifier. Medium.com. https://medium.com/@namanbhandari/extratreesclassifier-8e7fc0502c7

Bhavani, N. Ganga & Pichai, A. (2016). A Study of the Euro and U.S Dollar in the foreign exchange market using technical analysis tools. Innovative Journal of Business and Management 5:5, November –December (2016) 116 – 101., https://www.researchgate.net/publication/312623592_A_STUDY_OF_THE_EURO_AND_US_DOLLAR_IN_THE_FOREIGN_EXCHANGE_MARKET_USING_TECHNICAL_ANALYSIS_TOOLS

Canadian Bankers' Acceptance (BA) Rates. Investment Industry Regulatory Organization of Canada. https://www.iiroc.ca/industry/marketmonitoringanalysis/Pages/BA-Rate.aspx

Canadian Interest Rates and Monetary Policy Variables: 10-Year Lookup. Bank of Canada. https://www.bankofcanada.ca/rates/interest-rates/canadian-interest-rates/

Carapuco, J.M.B (2017). Reinforcement Learning Applied to Forex Trading. Técnico Lisboa. https://www.sciencedirect.com/science/article/abs/pii/S1568494618305349

Chihab, Y., Bousbaa, Z., Chihab, M., Bencharef, O. & Ziti, S. (2019). Algo-Trading Strategy for Intraweek Foreign Exchange Speculation Based on Random Forest and Probit Regression. https://www.hindawi.com/journals/acisc/2019/8342461/

Deng, S. & Sakurai, A. (2014). Short-term foreign exchange rate trading based on the support/resistance level of Ichimoku Kinkohyo. Institute of Electrical and Electronics Engineers. https://ieeexplore.ieee.org/abstract/document/6948127

Ernst, D. Geurts, P. & Wehenkel, L. (2006). Extremely randomized trees. https://link.springer.com/content/pdf/10.1007/s10994-006-6226-1.pdf

Gandhi, Rohith (2018). Naive Bayes Classifier. https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c

Hamilton, A. (2018). Understanding Exchange Rates and Why They Are Important. Bulletin -December 2018 | Finance. https://www.rba.gov.au/publications/bulletin/2018/dec/understanding-exchange-rates-and-why-they-are-important.html

Hobson, M.P. (2011). Technical Analysis. Cavendish Astrophysics, University of Cambridge. http://www.mrao.cam.ac.uk/~mph/Technical_Analysis.pdf

Hryshko, A. & Downs, T. (2004). System for foreign exchange trading using genetic algorithms and reinforcement learning. International Journal of Systems Science 35(13). https://www.researchgate.net/publication/43447963_System_for_foreign_exchange_trading_using_genetic_algorithms_and_reinforcement_learning

Jaadi, Z. (2020). A Step by Step Explanation of Principal Component Analysis. https://builtin.com/data-science/step-step-explanation-principal-component-analysis

Ilic, V. & Brtka, V. (2011). Evaluation of algorithmic strategies for trading on foreign exchange market. Information and Communication Technologies for Small and Medium Enterprises (ICT-SME's2011). https://www.researchgate.net/publication/299287406_Evaluation_of_algorithmic_strategies_for_trading_on_foreign_exchange_market

Khandelwal, P.(2017). Which algorithm takes the crown: Light GBM vs XGBOOST? https://www.analyticsvidhya.com/blog/2017/06/which-algorithm-takes-the-crown-light-gbm-vs-xgboost/

Lane, George C. (1990). Lane's Stochastics. Stocks & Commodities V. 2:3 (87-90). https://www.forexfactory.com/attachment.php/1410270?attachmentid=1410270&d=1397915452

Lashinski, V. (2008). Ichimoku Kinko Hyo, P.42 to 49. Journal of Technical Analysis 2008 Summer / Fall, Issue 65. Market Technicians Association, Inc.. https://www.yumpu.com/en/document/read/11279567/journal-of-technical-analysis-market-technicians-association

Machine Learning Guide (2018). Google. https://developers.google.com/machine-learning/guides/text-classification

Maitah, M., Prochazka, P., Cermak, M. & Šrédl, K. (2016). Commodity Channel Index: Evaluation of Trading Rule of Agricultural Commodities. International Journal of Economics and Financial Issues, 2016, 6(1), 176-178. http://www.econjournals.com/index.php/ijefi/article/download/1648/pdf

Marek, P. & Markova, V. (2020). Optimization and Testing of Money Flow Index. 19th Conference of Applied Mathematics. https://www.researchgate.net/publication/339029303_Optimization_and_Testing_of_Money_Flow_Index

Murphy, J. (1999). Technical analysis of the financial markets. Larry Williams % R, Point and Figure Charting, ISBN 0-7352-0066-1.

Nasution, Bona P., & Agah, A. (2000). Currency Exchange Rate Forecasting with Neural Networks. Journal of Intelligent Systems 10.3. https://www.degruyter.com/view/journals/jisys/10/3/article-p219.xml

Neely, Christopher J. & Weller, Paul A. (2011). Technical Analysis in the Foreign Exchange Market. Working Paper 2011-001B, Research Division, Federal Reserve Bank of St. Louis. http://research.stlouisfed.org/wp/2011/2011-001.pdf

Nishida, K. (2017). Introduction to Extreme Gradient Boosting in Exploratory. https://blog.exploratory.io/introduction-to-extreme-gradient-boosting-in-exploratory-7bbec554ac7

Okunev, J. & White, Derek (2001). Do Momentum Based Strategies Still Work In Foreign Currency Markets?. http://wwwdocs.fce.unsw.edu.au/banking/workpap/currency.pdf

Rohrbach, J., Suremann, S. & Osterrieder, J. (2017). Momentum and Trend Following Trading Strategies for Currencies Revisited - Combining Academia and Industry. Zurich University of Applied Sciences. Zurich University of Applied Sciences. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2949379

Prado, M. L. D. (2018). Section 5 Fractionally Differentiated Features. Advances in Financial Machine Learning. ISBN 978-1-119-48211-6

Prado, M. L. D. (2018). Section 3.4 The Triple-Barrier Method. Advances in Financial Machine Learning. ISBN 978-1-119-48211-6

Prado, M.L.D. (2018). The 10 Reasons Most Machine Learning Funds Fail. http://www.smallake.kr/wp-content/uploads/2018/07/SSRN-id3104816.pdf

Pukthuanthong-Le, K. and Thomas III (2008), Lee R. Weak-Form Efficiency in Currency Markets., Financial Analysts Journal 64(3). https://www.researchgate.net/publication/228258482_Weak-Form_Efficiency_in_Currency_Markets

Pukthuanthong-Le, K., Levich, R.M. & Thomas III, Lee.R. (2006). Do Foreign Exchange Markets Still Trend? The Journal of Portfolio Management 34(1). http://people.stern.nyu.edu/rlevich/wp/PLT1.pdf

Roudgar, E. (2012). Forecasting Foreign Exchange Market Trends: Is Technical Analysis Perspective Successful?. Institute of Graduate Studies and Research, Eastern Mediterranean University. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.938.1143&rep=rep1&type=pdf

Scikit learn. Plot the decision boundaries of a VotingClassifier. https://scikit-learn.org/stable/auto_examples/ensemble/plot_voting_decision_regions.html#sphx-glr-auto-examples-ensemble-plot-voting-decision-regions-py

Song, Y. (2017). A Forex Trading System Using Evolutionary Reinforcement Learning. Worcester Polytechnic Institute. https://web.wpi.edu/Pubs/ETD/Available/etd-050117-071501/unrestricted/ysong.pdf

Statistical Tables - Interest Rates - Interest Rates and Yields – Money Market – Daily – F1. Reserve Bank of Australia. https://rba.gov.au/statistics/tables/#interest-rates.

Swaminathan, S. (2018). Logistic Regression — Detailed Overview. https://towardsdatascience.com/logistic-regression-detailed-overview-46c4da4303bc

Towmey, B. (2012). Inside the Currency Market – Mechanics, Valuation, and Strategies, P.289-290. John Wiley & Sons, Inc, ISBN 978-0-470-95275-7 (hardback); ISBN 978-1-118-14933-1 (ebk); ISBN 978-1-118-14934-8 (ebk); ISBN 978-1-118-14935-5 (ebk).

Tsai, Y.C & Wang, C.C. (2018). Deep Reinforcement Learning for Foreign Exchange Trading. https://arxiv.org/abs/1908.08036

Wahlstrom, N. (2020). Lecture 6 – Tree-based methods, Bagging and Random Forests, Statistical Machine Learning, Division of Systems and Control, Department of Information Technology, Uppsala University. http://www.it.uu.se/edu/course/homepage/sml/lectures/Lecture6_handout.pdf

Wholesale interest rates - B2. Reserve Bank of New Zealand. https://www.rbnz.govt.nz/statistics/b2

Willis, O. (2020). Australia is flattening the coronavirus curve. But should we try to squash it like New Zealand? https://www.abc.net.au/news/health/2020-04-15/coronavirus-elimination-australia-new-zealand/12150302

William, L. (2011). Williams Percent R Indicator (%R). Larry Williams CTI Publishing, 2011. http://williamspercentr.com/

Yao, J. & Tan, C.L. (2000). A case study on using neural networks to perform technical forecasting of forex. Neurocomputing 34(1-4):79-98. https://www.researchgate.net/publication/222517139_A_case_study_on_using_neural_networks_to_perform_technical_forecasting_of_forex

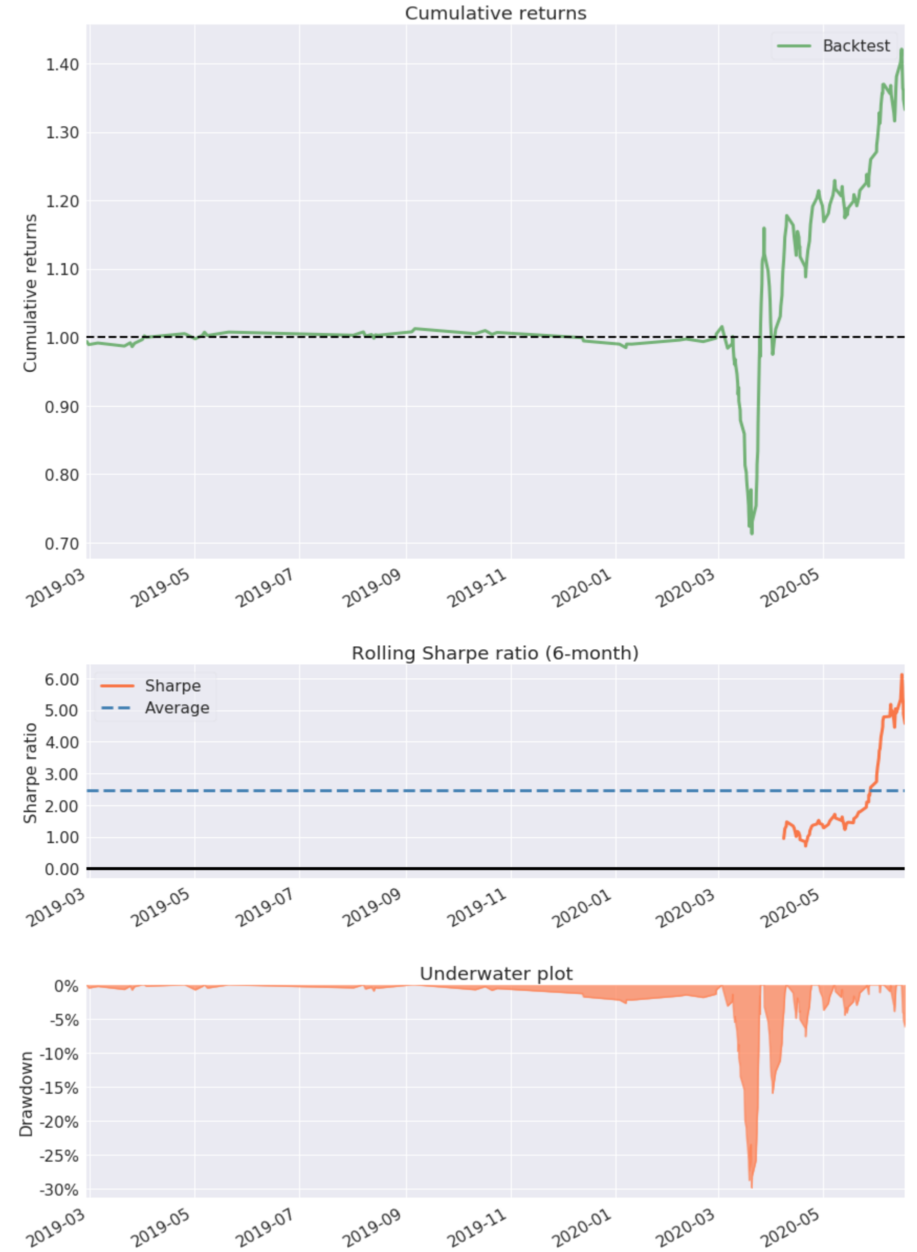
8 Appendix

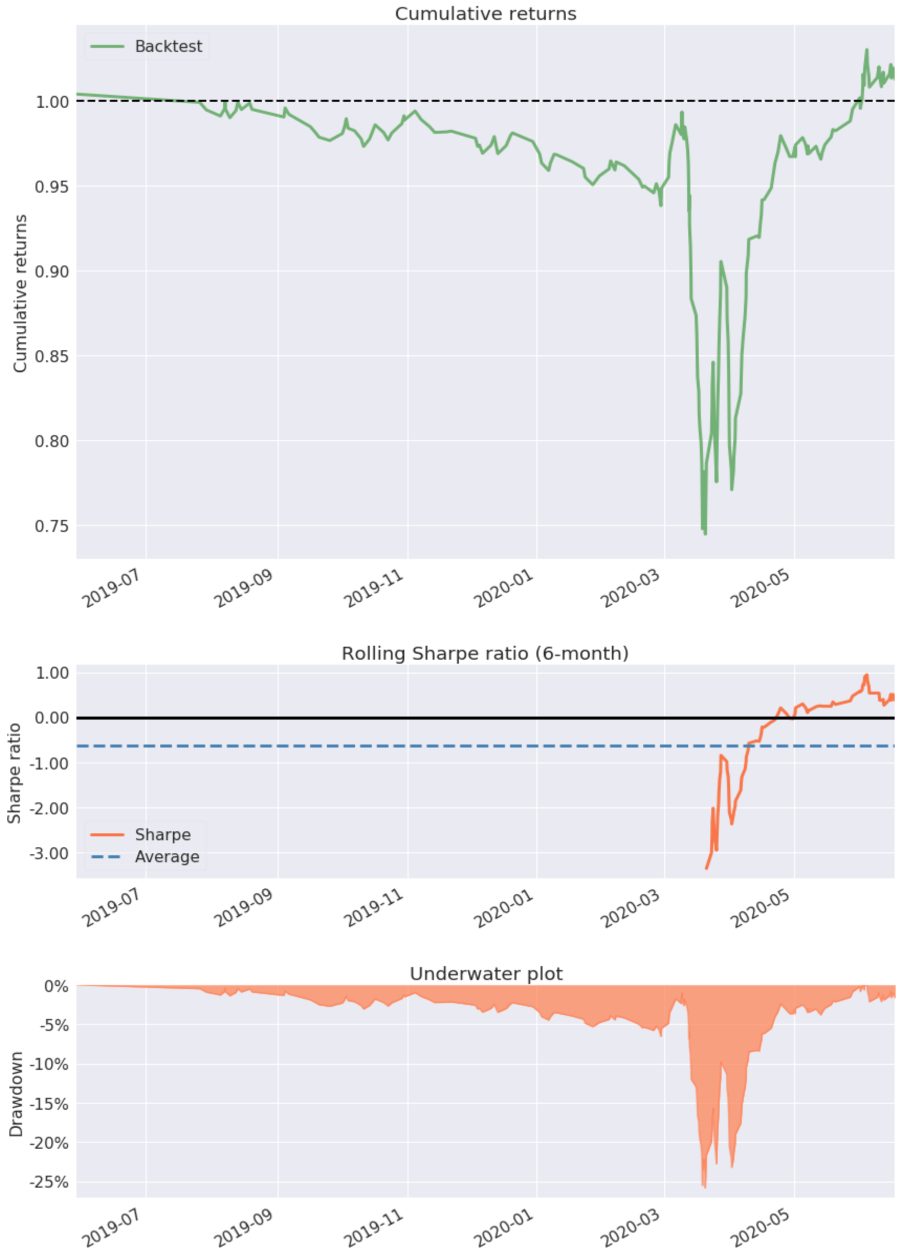
8.1 Code Reference

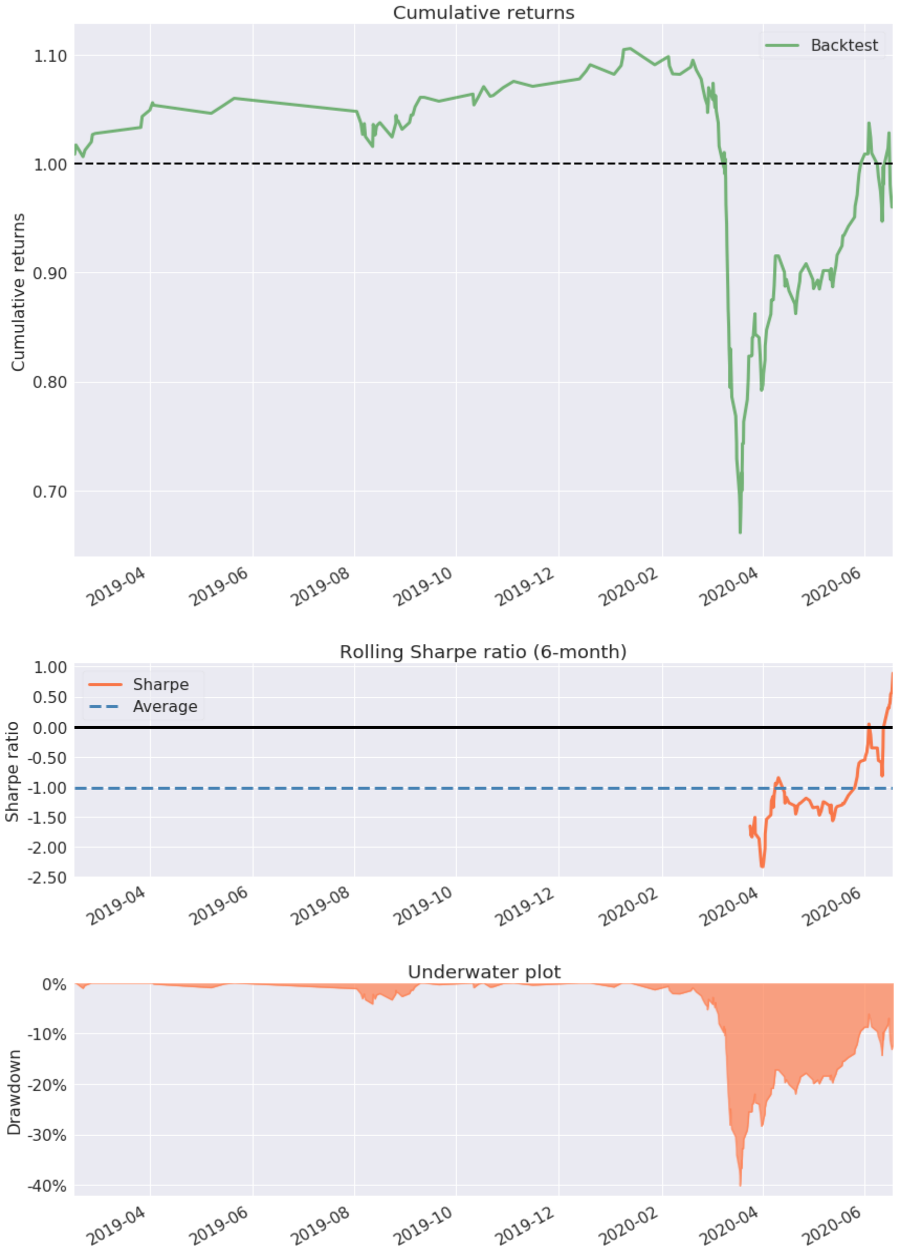
This project adopts a collaborative GitHub environment and the source code of Jupyter Notebook and Python code for the project is available at https://github.com/schigrinov/capstone

8.2 Backtesting results

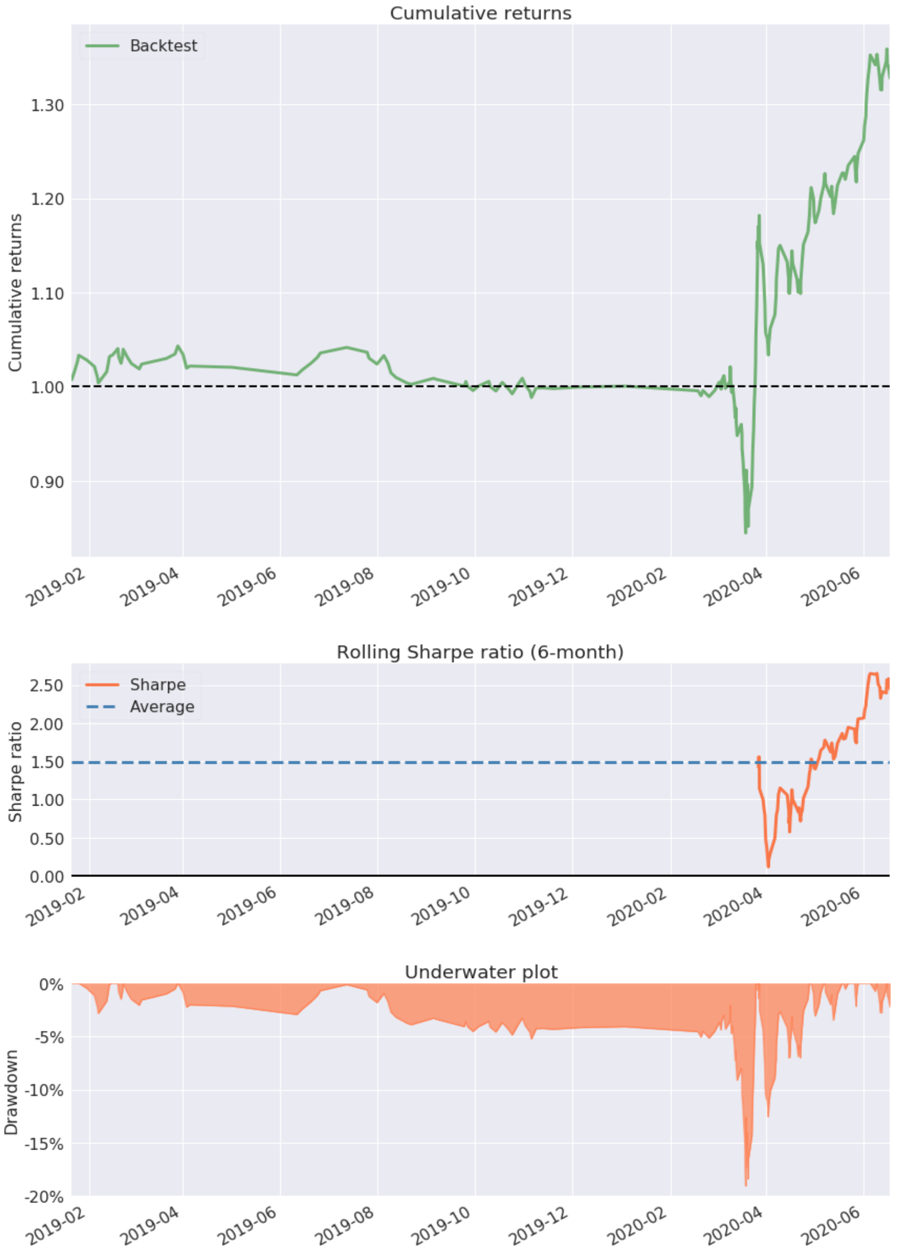
8.2.1 Logistic regressions with 4 principal components with features not applied for fractional differencing

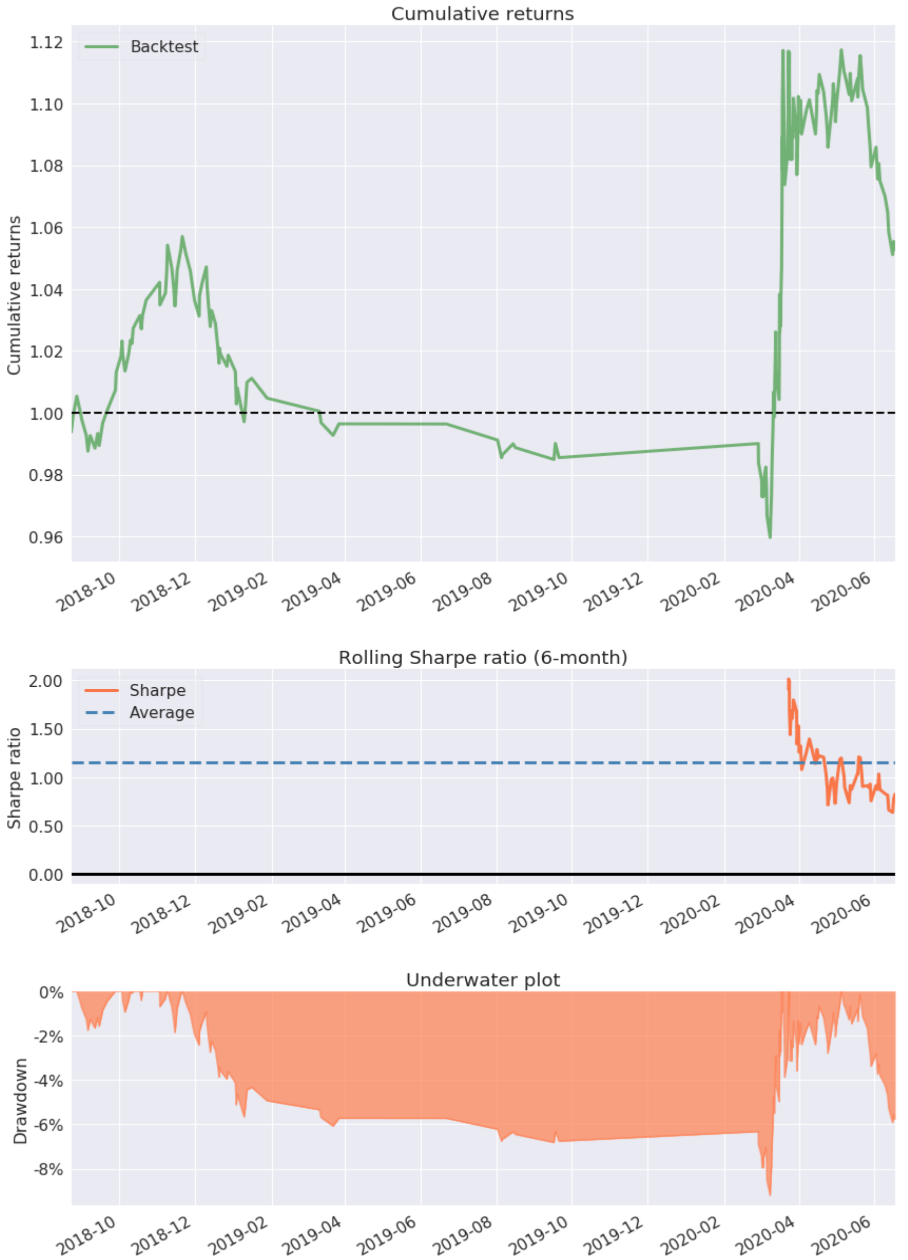
Currency	Backtesting results
AUDUSD	

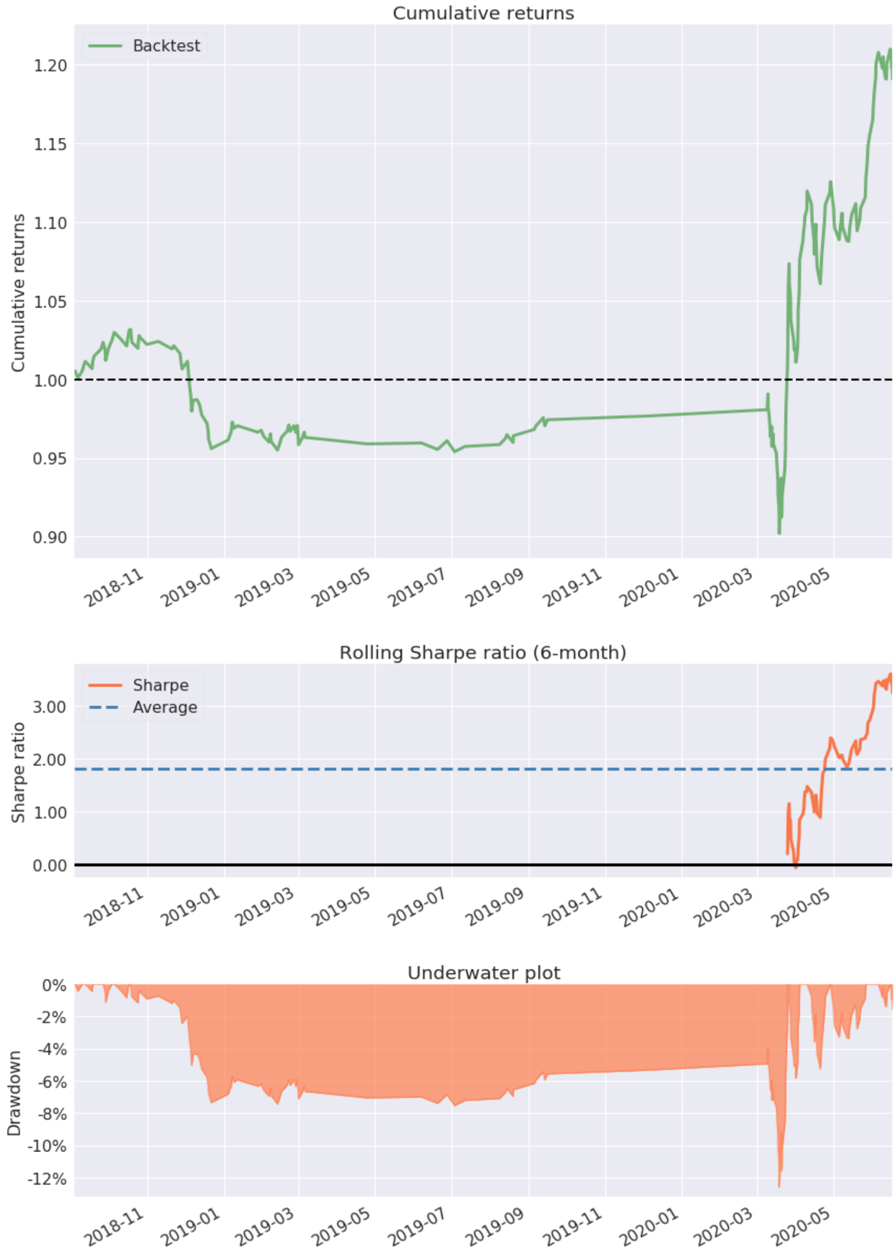
Currency	Backtesting results
AUDCAD	

Currency	Backtesting results
AUDJPY	

Currency	Backtesting results
GBPUSD	

<table>
<tr><th>Currency</th><th>Backtesting results</th></tr>
<tr><td>NZDUSD</td><td>
</td></tr>
</table>

Currency	Backtesting results
EURUSD	

Currency	Backtesting results
USDCAD	

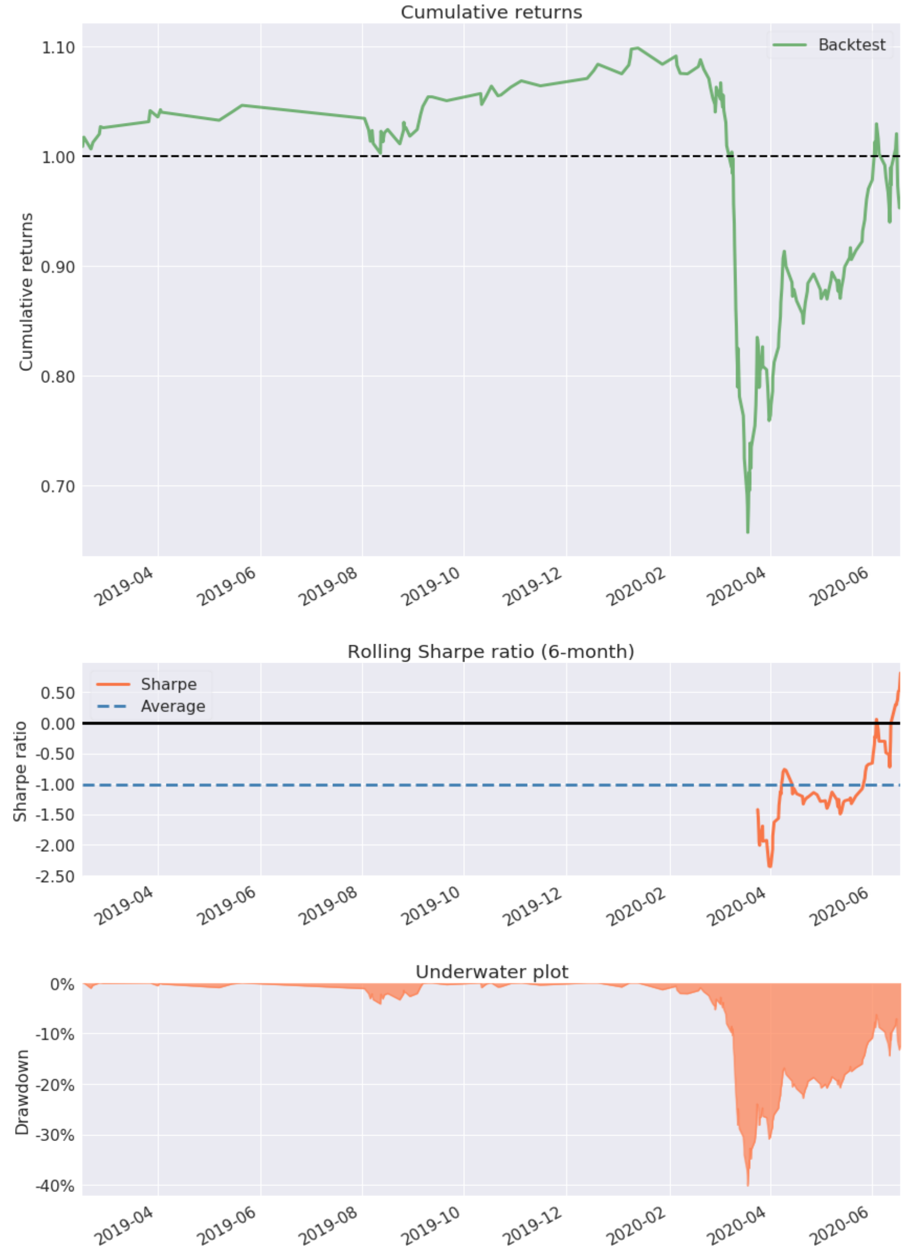
Currency	Backtesting results
USDJPY	

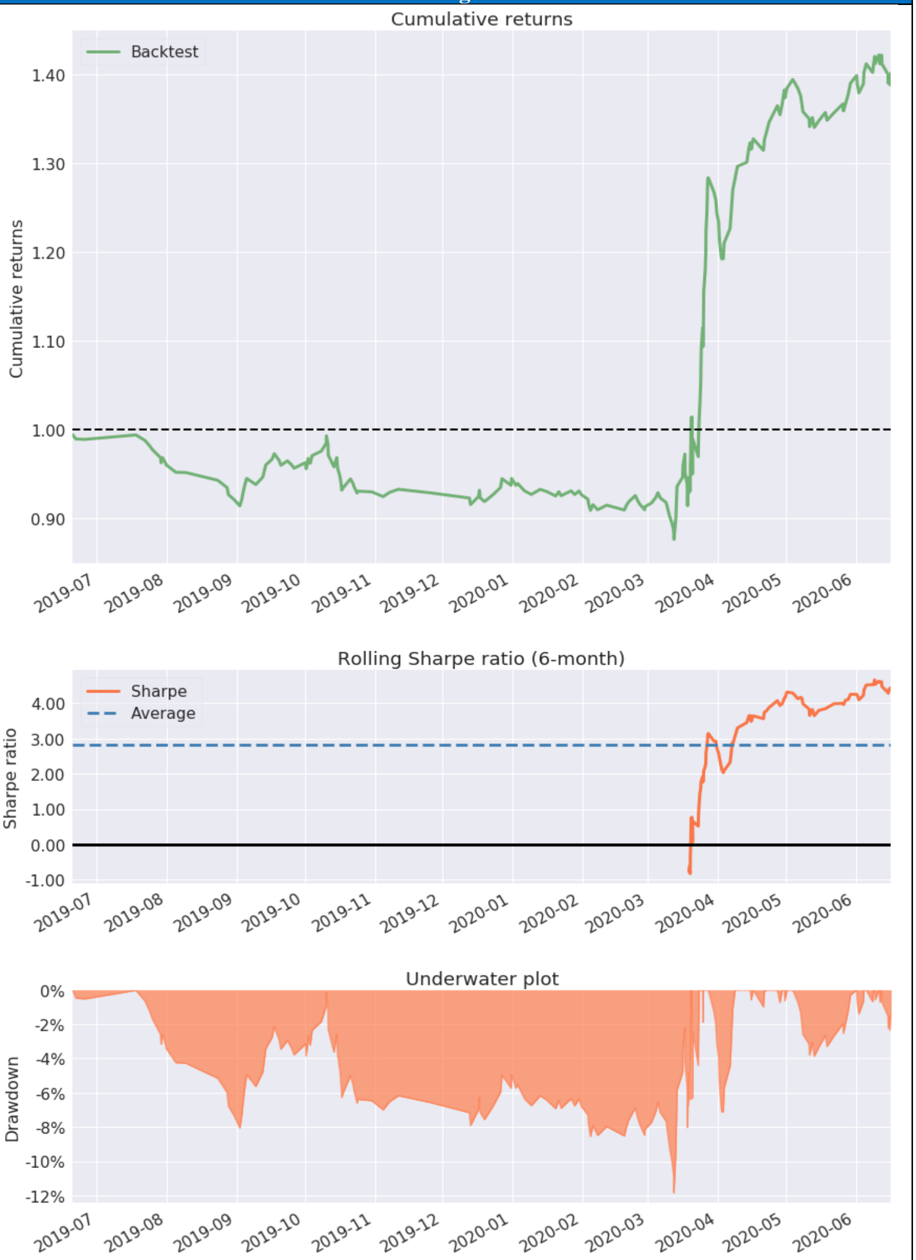
8.2.2 Voting Classifier

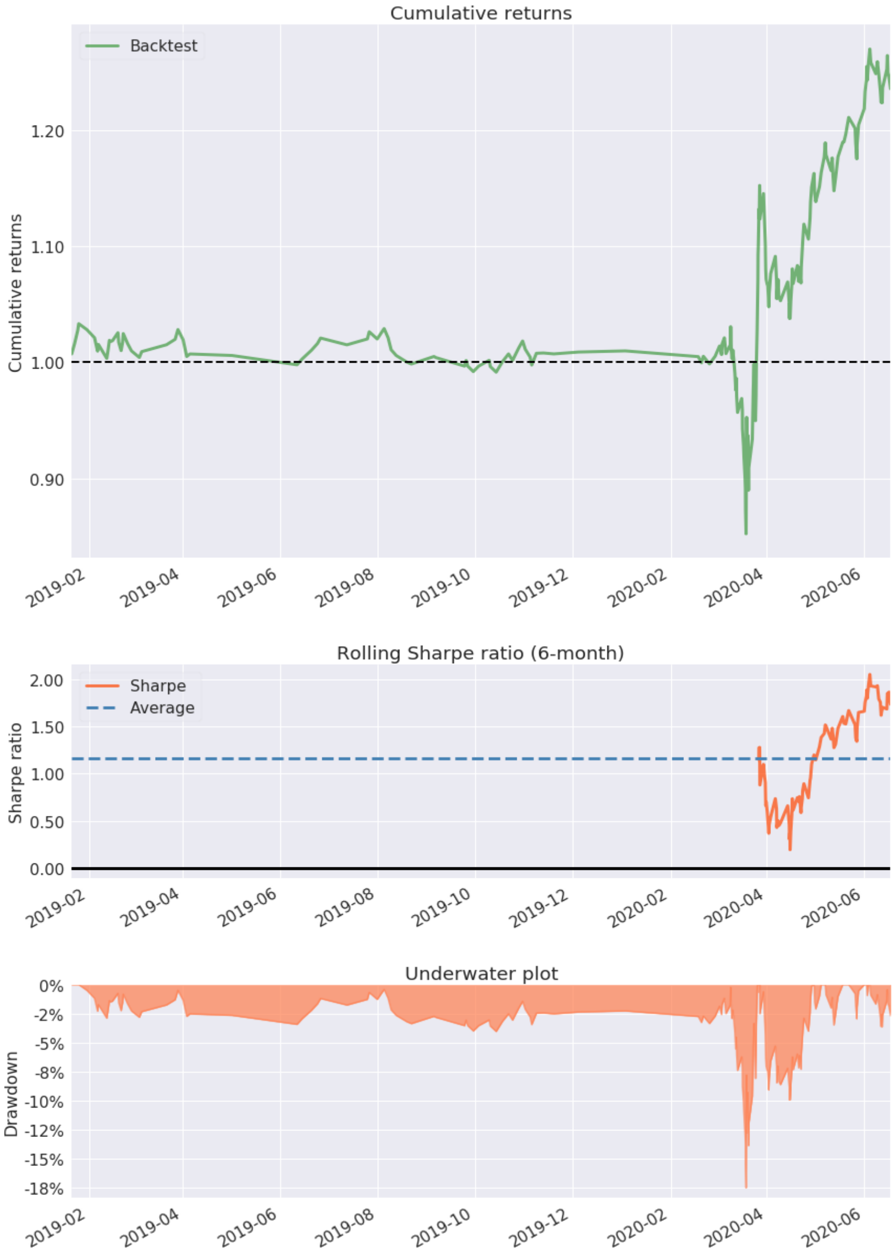
Currency	Backtesting results

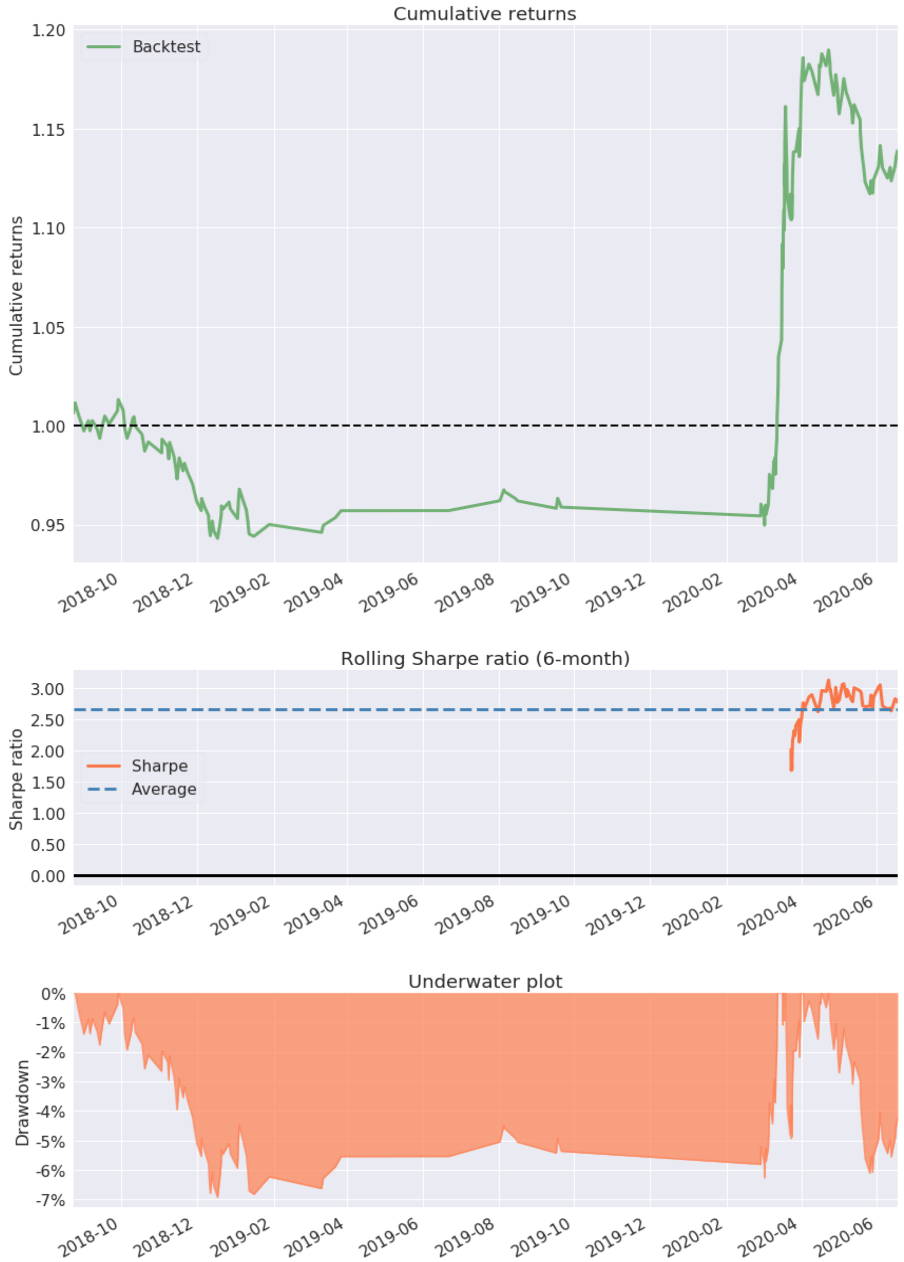
Currency	Backtesting results
AUDUSD	

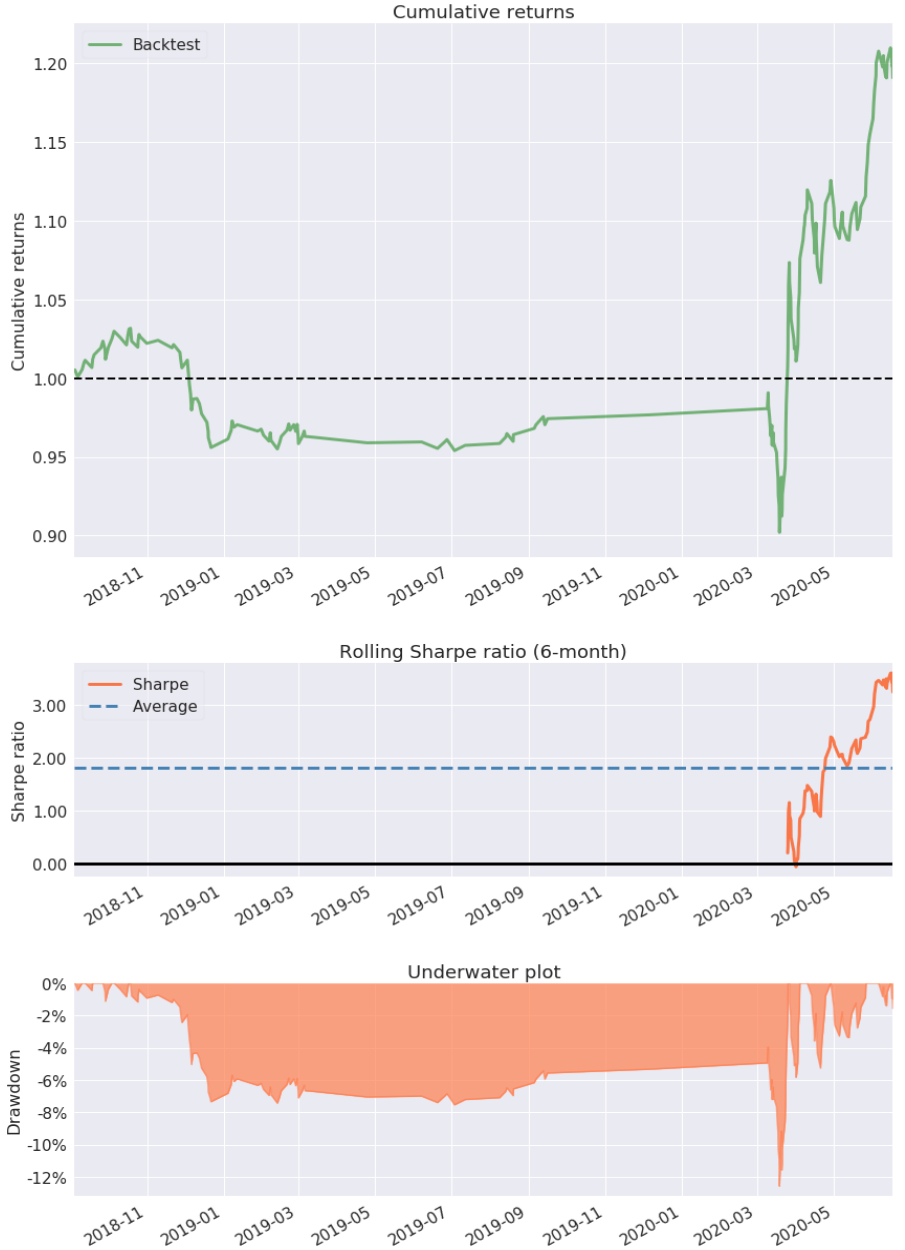
Currency	Backtesting results
AUDCAD	

Currency	Backtesting results
AUDJPY	

Currency	Backtesting results
GBPUSD	

Currency	Backtesting results
EURUSD	

Currency	Backtesting results
NZDUSD	

Currency	Backtesting results
USDCAD	

Currency	Backtesting results
USDJPY