Modelling Forex Market Reflexivity using Self-Exciting Point Process and Ensemble Learning

Master's Thesis

Devendra Shintre

dshintre@student.ethz.ch

Department of Mathematics - Seminar for Statistics
Chair of Entrepreneurial Risks
ETH Zürich

Supervisors:
Prof. Dr. Didier Sornette
Sumit Kumar Ram
Dr. Markus Kalisch

September 26, 2019

Acknowledgements

I would first like to thank my thesis supervisor Prof. Didier Sornette of the Chair of Entrepreneurial Risks at ETH Zurich for agreeing to give me the opportunity to work on this topic and for the various breakfast meetings providing meaningful insights pertaining to the implementation of our work in the real world.

I would also like to thank my co-supervisor Sumit Kumar Ram for providing much needed guidance. Sumit not only steered me in the right direction, but also allowed the thesis to be my own work. His guidance and help was essential to the research.

I would also like to acknowledge Dr. Markus Kalisch of Seminar for Statistics at ETH Zurich for providing me support as the second supervisor in this interdisciplinary thesis.

I would also like to thank Jan-Christian Gerlach for insightful discussions. I would like to extend my thanks to Adriana Schellenbaum for administrative help and the chair of Entrepreneurial Risks for providing high performance computing environment and making the computationally intensive experiments feasible.

Abstract

Efficient market hypothesis (Bachelier, 1900) has continued to dominate the discourse of finance and advocates the absence of arbitrage opportunities based on technical analysis. Its applicability and limitations have been pointed out as stylized facts in the economics literature. The thesis tries to address the inefficiencies or reflexivity in forex market.

We assume that the forex market is reflexive (Soros, 2015) and model the endogenous component of the market as a multi-variate self exciting conditional point process. We use power law memory kernels for modeling the endogenous correlations (Bouchaud, Kockelkoren, & Potters, 2006). Based on this market model we design a set of features using the past log returns, for a fixed time window, which take advantage of an ensemble of predefined power law memory kernels to classify the future returns, with the help of a random forest (RF) classifier.

An algorithmic trader is designed to exploit the predictions from our model, which takes long and short positions, uses trailing stop loss to cut the excess loss and makes trades within a fixed holding period. We train RF model on 23-10-07 14:17 to 15-11-15 17:18 ($\sim$ 8 years, 1 min sampling forex data for "AUDCAD" pair, 3,004,315 datapoints), train the trader on 15-11-15 17:18 to 12-04-16 00:32 (5 months, 150,000 datapoints). We test RF on 15-11-15 17:18 to 05-09-16 04:34 (10 months, 300,000 datapoints) and trader on 12-04-16 00:32 to 05-09-16 04:34 (5 months, 150,000 datapoints).

With our trained model we predict the drift of the price time series during holding period with an increment of 10% precision when compared with the random predictions. Our trader achieved Sharpe ratio of 6.56 on test set and outperforms the Buy and Hold strategy and noise trader with sufficient margin.

Contents

Chapter 1

Introduction

Efficient market hypothesis has been historically dominant school of thought in the field of finance. The thesis intends to challenge this hypothesis for the case of forex market by capturing the inefficiencies by modelling the reflexivity. The reflexivity is modelled as an ensemble of self exiting point processes or multivariate self exciting conditional point process, with memory kernel as power law and subsequently used to predict the direction for a fixed holding period using machine learning. Same model is later extended to an algorithmic trader.

Since determining the exact form of the power law memory kernel is non-trivial, a predefined ensemble of power law exponents is used along with the random forest algorithm to determine the functional relationship linking these exponents with the memory kernels and the future direction for the fixed holding period.

In this chapter, I will first introduce the terminology, then I will discuss the dataset used for the thesis, and then I will end the introduction stating our goals.

1.1 Terminology

In this section, the terminology used in the thesis would be explained briefly.

1.1.1 Exchange Market

The foreign exchange market (Forex, FX, or currency market) is a global decentralized or over-the-counter (OTC) market for the trading of currencies (Wikipedia contributors, 2019d). This market determines the foreign exchange rate. It includes all aspects of buying, selling, and exchanging currencies at current or determined prices. In terms of trading volume, it is by far the largest market in the world, followed by the Credit market.

1.1.2 Efficient Market Hypothesis

Efficient market hypothesis (Wikipedia contributors, 2019b) assumes that all the prices we observe are fair. The degree of "Fairness" depends on the form of the hypothesis namely - Strong, Semi-strong, and weak form. Even in the weakest form, the hypothesis reject the possibility of arbitrage opportunities based on the historical data. It assumes that the market follows geometric Brownian motion.

Weak Form

Weak form of efficient market hypothesis claims that historical price data holds no predictive power.

Semi-strong Form

Semi-strong form of efficient market hypothesis rejects the arbitrage opportunities based on technical and fundamental analysis.

Strong Form

According to the Strong form of efficient market hypothesis, even with the information, which is not public, there are no opportunities to make excess profit without taking excessive risks.

1.1.3 Asymmetric information

Efficient market hypothesis assumes that all the players in the financial world have direct or indirect access to the same information. However, it has been observed (Hasbrouck, 1988) that this is typically not the case in reality. The information asymmetry present amongst the market participants leads to decision making under uncertainty which systematically departs from the rational economic decision. This gives rise to "Reflexivity".

1.1.4 Reflexivity

Reflexivity (Wikipedia contributors, 2019g) is social phenomenon modelled as positive feedback cycles. It is widely observed in animals, birds, and human behavior. Herd of sheep following a certain and not pre-planned direction is a good example of reflexivity. Notably herd behavior is not random in this case. Act of a single individual has capacity to influence the future dynamics. Riots and strikes can also be good example. Similar behaviour is observed in stocks, primarily during bubbles and crashes.

1.1.5 Trading

A trade (Wikipedia contributors, 2019i) is an exchange of a security (stocks, bonds, commodities, currencies, derivatives or any valuable financial instrument) for money. A trade typically involves a promise to pay in the currency of the country where the "exchange" is located. The trading price or the price at which a financial instrument is traded, is determined by the supply and demand for that financial instrument.

1.1.6 Algorithmic Trading

When trading signal is generated based a pre-programmed algorithm, the trading strategy is known as algorithmic trading (Wikipedia contributors, 2019a). Popular "algos" include Percentage of Volume, Pegged, VWAP, TWAP, Implementation shortfall, Target close. Although the basic algorithms are well known, their modifications are usually proprietary. It is used by investment banks, pension funds, mutual funds, and hedge funds on a regular basis. The reason being that, these institutional traders need to execute large orders in markets, and they naturally cannot support all of them at once. However recently it has gained popularity with the retail investors as well. According to even the weakest form of efficient market hypothesis, Algorithmic trading is impossible for making profits, as any intelligent trading-algo is equivalent to noise trading.

Momentum Trading

When trading signal is generated based on past momentum trends only, the trading strategy is known as momentum trading (Lee & Swaminathan, 2000). Price momentum is similar to momentum in physics, where mass multiplied by velocity determines the likelihood that an object will continue on its path. Momentum investing seeks to take advantage of market volatility by taking short-term positions in stocks going up and selling them as soon as they show signs of going down. The investor then moves the capital to new positions. A momentum investor looks to take advantage of investor herding by leading the pack in and being the first one to take the money and run. Weak form of Efficient market hypothesis rejects the possibility of momentum trading.

1.1.7 Random Forest

Random forest(Breiman, 2001) is ensemble supervised learning algorithm. Statistical methods can be broadly classified into two categories, Bayesian and frequentist approach. Each can be then subdivided into supervised and unsupervised learning methods. Supervised learning can in turn be further subdivided into classification and regression. Both these problem can be solved by tree based and non-tree based approaches. Random forest is tree based ensemble approach, which can be used for both classification and regression purpose. A random forest is an ensemble of decision trees. A decision tree is formed by recursively partitioning the dataset, in a way which encourages the homogeneity in the partition. Because of the ensemble technique, Random forest is not only capable of taking advantage of the reduced bias from a fully partitioned tree, but also reduced variance provided by averaging.

1.2 Data

Analysis is done on "AUDCAD" pair, one min tick data starting from "2007-10-23 14:17" to "05-09-16 04:34". There are 3,304,315 data points in total.

1.3 Goal

1.3.1 Primary

There are two objectives of this study:

1. **The goal is to prove that $\exists$ intermittent reflexivity patterns in high frequency forex data, which hold potential to predict the future drift. Further, the patterns are not spurious and rather statistically robust and can be tested against all forms of efficient market hypothesis.**
2. **The reflexivity patterns are not only predictable but also exploitable with a very simple trading algorithm.**

Primary goal is to substantiate the existence of intermittent reflexivity in the currency market. The idea is to design features based on historical data and to test their efficacy against the null model created on synthetic price series following geometric Brownian motion and shuffled return time series of the real data. In order to utilize these features, Random forest would be used to predict the future drift.

Once the effectiveness of the reflexivity based features is verified, subsequent goal is to design automated algorithmic trading agent to replace the human decision making involved.

Chapter 2

Market Efficiency

Efficient Market Hypothesis forms the cornerstone of the today's world of finance. It provides a strong conceptual framework to understand the financial world.

2.1 Efficient Market Hypothesis

Complexity of the financial world we live in, exceeds our capacity to comprehend it (Soros, 2015). We try to comprehend it with simplistic theories like Efficient market hypothesis (EMH henceforth) which are comprehensible. EMH assumes that markets are efficient. That is, all the prices that are visible are fair and there are no statistical or technical arbitrage opportunities (Bachelier, 1900). Theoretically it claims that neither fundamental nor technical analysis (and even inside information in for strong form of EMH) can consistently generate positive alpha.

The EMH was publicised and brought to the public attention by Fama and Malkiel (Malkiel & Fama, 1970). They argued that stocks always trade at their fair value. There are neither undervalued stocks nor inflated prices. It is impossible to outperform the market through expert stock selection or market timing. Only pure luck or increased risk exposure can explain the out-performance of an inverstor. The view has been supported by many scholars. Based on the distribution of abnormal returns of US mutual funds, which was very similar to what one would be expected if no fund managers had any skill, the views were further substantiated by (Fama & French, 2010).

According to EMH stocks are neither undervalued nor overvalues, they are exactly priced. Only way an investor can earn excess returns is by increasing his risk exposure. There are no free lunches and hence no means to consistently generate returns without taking risk. It is impossible to outperform the market consistently.

In the world of finance, there are two schools of thoughts. One which says market is efficient (Fama, 1965) and one which criticises this hypothesis (Malkiel, 2003).

Statistical testing methodology (Null hypothesis significance testing), which is usually followed, has its own limitations (Wikipedia contributors, 2019h). Failure to reject, is not an evidence for for presence. This argument is typically offered by the critics of EMH. As there are no quantitative measures to ascertain the efficiency of given market, Statistical significance testing becomes impossible. According to critics (Campbell et al., 1997), EMH can only be accepted until one manages find a real life example against the hypothesis. However proponents of EMH have continued to argue their case with equally strong evidences (Jensen, 1978).

Occurrence of crashes and bubbles did draw a lot of criticism for EMH. (Campbell, 2014) and (Shiller, 2014) have discussed in great depth about the future of EMH. (Thaler & Ganser, 2015) have suggested alternatives like behavioral economics.

2.2 Foundations

EMH assumes that all investors have rational expectation and all are utility maximizing agents. Please note that it does not require all the agents to be rational. Ensemble of agent is correct irrespective of accuracy of individual agents. Market on the whole is always fair. When new information is presented, agents change their expectations in accordance. Some will over-react and some will under-react. As a group the decisions are random and hence there are no arbitrage opportunities, especially when transaction costs are considered.

EMH is typically presented in three different forms, depending on the strength of assumptions - Weak form, Semi- strong form, and Strong form.

2.2.1 Weak Form

Weak form of EMH assumes that current market price only reflects the past prices, and subsequently there are no arbitrage opportunities based on the past prices. Technical analysis alone will never be sufficient to provide excess returns. There are no patterns in the price time series which can be exploited. Correlation between future prices and current price series is insignificant. However, fundamental analysis, may provide arbitrage opportunities in this case.

2.2.2 Semi-Strong Form

Markets are assumed to adjust in an unbiased fashion to not only historical data, but also the publicly available data. Therefore neither technical nor fundamental analysis can provide excess returns without significant risk exposure. Not only the patterns are absent in the prices series, but also reading the business reports of the company or following the relevant geopolitical news will not provide any hints about the future price trends.

2.2.3 Strong Form

Strong form of EMH claims that share prices reflects all information - Historic, public and private and no one can get excess returns. Typically governing bodies impose restrictions on legally sharing the private information. Strong form goes on to make the bold claim, that in spite of those restrictions, market and thereby investor are directly or indirectly privy to the inside information.

2.3 Mathematical formulations of EMH

EMH claims that market returns follow random walk or geometric Brownian motion i.e. $X \sim \mathcal{N}(\mu, \sigma^2)$. In simple terms, markets returns are white noise and can be given by Gaussian distribution as in equation (2.1) (Wikipedia contributors, 2019e).

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.1}$$

In figure 2.1, the plots corresponding to Gaussian log returns are shown along with the corresponding price time series. We would like to bring the absence of clustering in log returns (volatility clustering to be specific) to the reader's attention. Lack of pattern in white noise makes it unpredictable except for mean and variance.

If $X \sim N(\mu, \sigma^2)$, then permutations of realization of X will also follow Gaussian distribution. In simple words, Normal distribution are invariant under permutations. Typically Gaussian distributions are not fat-tailed - i.e. outliers are rare events. However, in real life (Figure 2.1) outliers and subsequently fat tailed distributions do occur.

2.3.1 Weak Form

(Samuelson, 2016) has mathematically proved the weak form of EMH (Equation 2.2). Here V^* is the fundamental value and P_t is the price at time t.

$$P_t = E[V^* | I_t]$$
$$P_{t+1} = E[V^* | I_{t+1}]$$

where $I_t = (P_{t-1}, P_{t-2}, P_{t-3}, ...)$

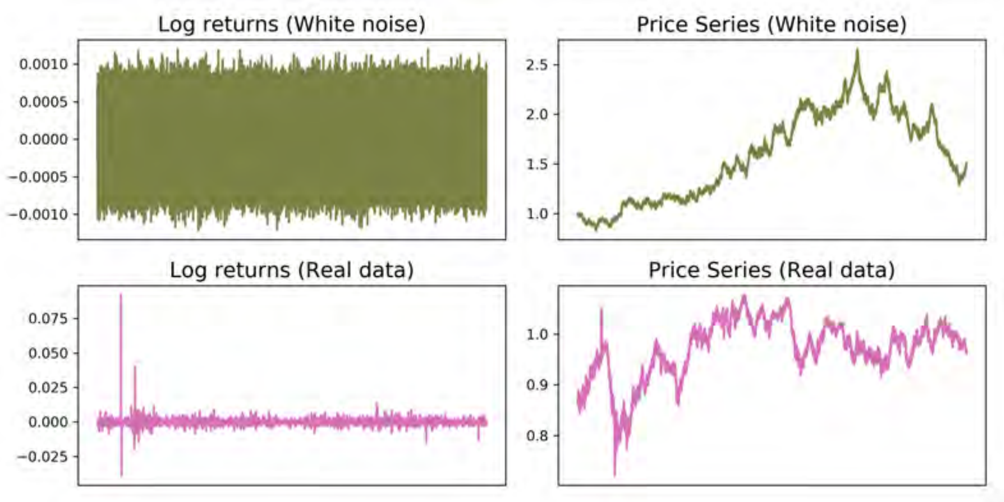

Figure 2.1: White noise log returns were synthesised and subsequently converted to price time series and plotted. Real FX time series and the corresponding log returns were also calculated and plotted for comparison. Volatility clusters are present in real time series in contrast with the synthetic one.

$$E[P_{t+1} - P_t|I_t] = E[E[V^*|I_{t+1}] - E[V^*|I_t]|I_t]$$
$$E[P_{t+1} - P_t|I_t] = E[V^*|I_t] - E[V^*|I_t] = 0$$

$\implies$ Price behaviour follows random walk.

$$\implies P_t = P_{t-1} + \epsilon_t$$

where $\epsilon_t \sim \ i.i.d.(0, \sigma^2)$

$$\implies E(P_t|I_t) = P_{t-1} \tag{2.2}$$

2.3.2 Semi-strong Form

Although it is hard to mathematically prove the semi strong form of EMH, it tends to occur quite frequently in real life. Jump in stock prices on the day of merger announcement and subsequent stability (relative) in the stock price is perfect example of semi strong form of EMH.

2.3.3 Strong Form

(Jensen, 1954) suggested a mathematical measure, Jensen's alpha, to assess the performance of a mutual fund, and subsequently to substantiate the strong form of EMH. It works in the context of CAPM (Capital Asset Pricing Model). It is defined in (Wikipedia contributors, 2018) as follow.

$$\alpha_J = R_i - [R_f + \beta_{iM} \cdot (R_M - R_f)]$$

Where R_i is the realized return (on the portfolio), R_M is the market return, R_f is the risk-free rate of return, and β_{iM} is the beta of the portfolio.

If Strong form of EMH is invalid for a particular fund manager, then the corresponding Sharpe ratio i.e. $\frac{R_i - R_f}{\beta_{iM} R_M}$ will be better than benchmark portfolio. This enables us to do the following statistical test

$$H_0 : \alpha_J > 0$$

Which is equivalent to testing

$$H_0 : \frac{R_i - R_f}{\beta_{iM} R_M} > \frac{R_M - R_f}{\sigma_M}$$

2.4 Contradicting Views

Critics argue that in real life market returns do not follow random walk and there are volatility clusters (Figure 2.1). Even the proponents of EMH (Sewell, 2012) have agreed that EMH, by itself, is not a well-defined and empirically refutable hypothesis. (Campbell et al., 1997) and (Cuthbertson, Hayes, & Nitzsche, 1997) have argued that any test of EMH is a joint test of an equilibrium returns model and rational expectations and hence by design an irrefutable hypothesis.

(Filimonov, Bicchetti, Maystre, & Sornette, 2014) and (Sornette, Woodard, & Zhou, 2009) have discussed bubble formation in commodities market and oil market respectively as an endogenous process. Markets are subjected to internal feedback loops (Filimonov et al., 2014), which are created by the collective behavior. e.g. herding or informational cascades. Prices do influence the fundamentals and these influenced set of fundamentals then proceed to change expectations, thus influencing prices.

EMH by definition, cannot explain reflexivity. EMH reached the height of its dominance in academic circles around the 1970s (Shiller, 2003). Later occurrence of reflexive/herding behavior was observed more frequently during bubbles and crashes. (Mandelbrot, 1997) observed volatility clustering in financial markets as a phenomenon where "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes". Such observations along with application of ARCH processes to financial markets (Engle, 1982) bolstered the critical view towards EMH.

2.5 Conclusion

EMH in its weakest form claims that there is no extra information in the historical time series data. Volatility clusters do not exists. Log returns follow geometric Brownian motion. Consequently, there are no arbitrage opportunities.

However, as apparent in Figure 2.1, volatility clusters do exists in the FX market. Thus the corresponding return time series, and consequently technical analysis does hold predictive power concerning future returns.

Chapter 3

Reflexivity and its Quantification

Since the introduction of EMH (Chapter 2), there has been endless debate going on between the proponents and critics of EMH. (Soros, 2013) defined two concepts, fallibility and reflexivity. There are facts in the world, which corresponds to actual reality. However, our view of the world, as a thinking participant never corresponds perfectly to the real state of affairs. (Soros, 2013) recognised this as fallibility. These imperfect views yield the capacity to influence the real world, of which they are participant, as the participants start to act. If an investor believes in EMH, this belief may influence his investment behaviour, which in turn may change the market in which they are investing (it need not make it more efficient). This is principle of reflexivity (Soros, 2013).

Our stand is that currency markets are not efficient even in the weak form. FX Markets are intermittently reflexive. In other words, there exists positive and negative feedback cycles, which can make the market movements predictable in short term horizon. The mechanism of this reflexivity is discussed in (Koutmos & Saidi, 2001) and (De Long, Shleifer, Summers, & Waldmann, 1990). The theory and its quantification methods are discussed in great depths in (Filimonov & Sornette, 2012). (Filimonov et al., 2014) have discussed the reflexivity as an endogenous process in the case of commodities market.

Our hypothesis is that Reflexivity follows the dynamics of multi-variate self exciting conditional point process following power law as memory kernel and the corresponding response function for the feedback loops can be estimated using the same.

3.1 Reflexivity

Reflexivity is defined as positive feedback effect of market sentiment, where rising prices attract more buyers, and thereby creating more demand. Consequently, prices are driven further up till the point, where they become unsustainable. Self reinforcing effect of market sentiments causes such positive feedback loops, or bubbles. Similar effect can be observed in crashes, where sinking prices promotes selling, and thereby creating the scarcity of buyers. The prices fall rapidly till a certain point (Bankruptcy in the worst case scenario).

There are two ways to model finance, one from game theoretical perspective and other from behavioral finance. (Sen, 1969) has proposed a mathematical model for quasi-transitivity in reflexivity and collective decision. (Soros, 2015) on the other hand advocates against mathematical models. From game theoretical perspective all players are rational, and therefore outcome is deterministic. From the lens of behavioral finance, participants are emotional human beings, and hence irrational. Idea of reflexivity falls somewhere in between. Market participants are neither completely emotional nor rationally stubborn.

3.1.1 Herding behavior

Even without deciding explicitly, a group or a herd will follow the certain direction. This introduces a form of predictability while forecasting. Such behavior not only occur in case of animals and birds, but also for humans during riots, strikes, and demonstrations. There is also benign herding involved in everyday decision making process. Just like in real life, social factors like herding effect are typically observed in finance (Caparrelli, D'Arcangelis, & Cassuto, 2004) as well.

3.2 Feedback

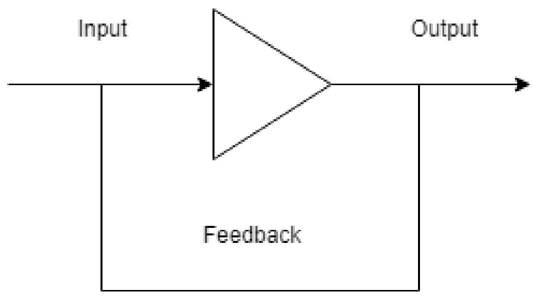

Figure 3.1: Circuit representation of a feedback loop

Feedback (Figure 3.1) occurs when output of a system is rerouted as input. This is typically a part of cause and effect chain and affects the subsequent outputs (Wikipedia contributors, 2019c). The notion of causality plays pivotal role while understanding a feedback system. There are two types of feedback, positive and negative.

3.2.1 Positive feedback

Positive feedback occurs when small disturbances caused by the perturbation of a system tend to amplify the perturbation, i.e. when output and input are in phase. There is typically exponential or super exponential growth accompanied by chaotic behavior, which is naturally divergent from equilibrium.

3.2.2 Negative feedback

Negative feedback occurs when output is fed back to a system in a way such that it dampens the oscillations. Negative feedback tends to promote stability and typically leads the system back to the equilibrium state.

3.3 Self exiting point process

A point process or point field is a collection of mathematical points randomly located on some underlying mathematical space such as the real line, the Cartesian plane, or more abstract spaces (Wikipedia contributors, 2019f). In the case of finance, if returns are modelled as a point process, they would be normally distributed for efficient markets. However, this under-predicts the occurrence of the outliers or dragon kings (Sornette, 2009). (Bouchaud, Farmer, & Lillo, 2009) provides a brief summary of occurrence of deviation from efficiency. Therefore, modelling the returns as a Gaussian should be avoided. As an alternative we aim to model the returns as self-exciting Hawke's process (Figure 3.2), which can produce the feedback effects and can explain the dragon kings.

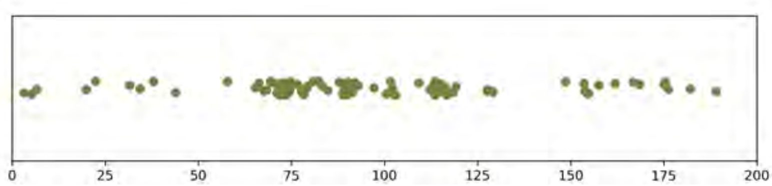

Figure 3.2: Simulation of Hawke's process

Intuitively, a process is self-exciting if the occurrence of past points makes the occurrence of future points more probable. A point process is called self exiting if $cov(N(s,t), N(t,u)) > 0$ for $s < t < u$ (where cov means covariance). The idea was initially proposed in (Hawkes, 1971). Hawke's Process N_t is a simple point process, whose conditional intensity is defined as (Wikipedia contributors, 2019f)

$$\begin{aligned} \lambda(t) &= \mu(t) + \int_{-\infty}^{t} \nu(t-s) dN_s \\ &= \mu(t) + \sum_{T_k<t} \nu(t-T_k) \end{aligned} \tag{3.1}$$

Where $\nu : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a kernel function quantifying the positive feedback of past events T_i on current intensity $\lambda(t)$. $\mu(t)$ is non-stationary function representing the expected, predictable, or deterministic part of the intensity.
$\{T_i : T_i < T_{i+1}\} \in \mathbb{R}$ is the time of occurrence of the i^{th} event of the process.

3.3.1 Multi-variate self exciting conditional point process

Multi-variate self exciting conditional point process (Saichev, Maillart, & Sornette, 2013) is extension of unidirectional point process to multivariate space conditional on the recent events in the fixed time horizon. The conditionality is introduced for computational reasons.

$$\lambda_j(t|H_t) = \lambda_j^0(t) + \sum_{k=1}^{m} \Lambda_{kj} \int_{(-\infty,t)\times R} f_{k,j}(t-s) g_k(x) N_k(ds \times dx) \tag{3.2}$$

Here H_t denotes the whole past history up to time t, λ_j^0 is the rate of spontaneous (exogenous) events of type j. Λ_{kj} is the $(k,j)^{th}$ element of the matrix of coupling between the different types which quantifies the ability of a type k-event to trigger a type j-event. The memory kernel $f_{k,j}(t-s)$ gives the probability that an event of type k that occurred at time s < t will trigger an event of type j at time t. The function $f_{k,j}(t-s)$ is the distribution of waiting times $t-s$ between the impulse of event k which impacted the system at some time s and the occurrence of an event of type j at time t. The fertility (or productivity) law $g_k(x)$ of events of type k with mark x quantifies the total average number of first-generation events of any type triggered by an event of type k (Saichev et al., 2013).

3.3.2 Power law

When one dependent variable varies with some power of independent variable, the relationship generated is known as power law. We believe the distribution of waiting time between the action leading to strikes or herding behaviour is distributed as a power law. The assumptions, implementation, and reasoning behind the usage of power law as a memory kernel for Hawke's process is discussed in Section 5.2.

3.3.3 Relation with EMH (Chapter 2)

When memory kernel (ν) is zero, Hawke's process boils down to simple Poisson point process. This would follow up from EMH. In this case, the past time series holds zero predictive power, as $cov(N(s,t), N(t,u)) = 0$. Consequently, no predictions about the future can be made.

ν is a very important parameter of the process. Its value determines the existence of temporal cluster and their size. For non-zero ν, there exists a pattern which can be exploited to identify the clusters and consequently, make intelligent predictions about the future.

Chapter 4

Machine Learning Model

The direction prediction is a classification problem. The models which are used to typically solve a classification problem are logistic regressions, Support vector machines, Linear Discriminant Analysis, Random forest, Neural Network etc. Idea is to find the true distribution. All models will lead to the same solution under certain conditions. Since Random Forest (Breiman, 2001) offers fair accuracy in ML models, and it is not as complex as Neural network, it is usually a safe choice. Random Forest deals with Bias Variance trade-off in very elegant manner. We are using Random forest classifier with 500 trees. In order to tackle class imbalance, we will use balanced sub-samping to re-weight the costs.

4.1 Classification problem

A learning problem is classified into two categories; Supervised and Unsupervised learning problem. Supervised learning problem can be further bifurcated into two categories; Regression and Classification.

4.1.1 Regression

In regression problems continuous value of output signal is estimated. Estimating the exam score can be an example of a regression problem.

4.1.2 Classification

When output variable is a factor instead of continuous signal, the problem is typically referred as classification. Output is typically the group in which the observation belongs.

4.2 Learning Problem

What exactly we want our model to learn? This is the statement of a learning problem. We want predict the direction or polarity for a fixed future holding period.

Let us get introduced to model as a whole.

4.2.1 Features

Feature construction is explained in detail in Chapter 5. Briefly, for given time, we take the weighted sum of historical returns for a fixed look-back time horizon (weighted with given γ). We then separate the positive and negative part, and take their ratio which would then serve as the feature value corresponding to the given γ and t.

4.2.2 Target Variable

For target, we are looking at next n steps in the future. We will predict whether it will stay up, go down or keep fluctuating in the next n steps.

4.3 ML Model

There are various models which can be used to solve a classification problem (logistic regression, SVM, Random forest, Neural network etc.) There are typically two ways to chose a particular model - cross validation or using domain knowledge. We are using the domain knowledge. Logistic regressions are comparatively the least accurate, but are highly explainable. Neural network on the other other hand, can be very accurate, but unfortunately offers zero explainability. We chose random forest because it makes a nice compromise between accuracy and explainability.

Random forest was first proposed as an ensemble method by (Ho, 1995). He advocated "Divide and conquer" ideology and proved that one can gain accuracy with intelligent feature selection for each tree. (Barandiaran, 1998) built on this and reiterated the importance of selecting relevant feature on the accuracy. (Amit & Geman, 1997) introduced the idea of searching over random subset. This lead to (Breiman, 2001), where he introduced method of building trees using CART (Breiman, 2017) like procedure and bagging for random forest. This procedure to build a forest is widely accepted in the industry and we would be using the same.

4.3.1 Decision Tree

Decision tree is simply recursive partitioning of the data, done in a way which optimally segregates the classes. It is a greedy algorithm, which finds the optimum cut based on loss function (like gini score).

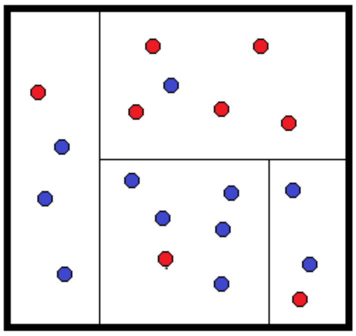

Figure 4.1: Visualization of a decision tree as method which recursively partitions the dataset.

4.3.2 Random Forest

Decision tree typically has a tendency to overfit. When we let the the decision tree grow to its full height (or partitioning depth), it overfits the data. On the other hand, with just one branch it typically underfits. It is very difficult to judge the optimal height. One solution is cross validation, other easy solution is random forest (Breiman, 2001). In random forest, we first bootstrap the data (For (Breiman, 2001) version, we randomly select some features as well to make the bootstrapped data more diverse). We then train a tree on each bootstrapped sample to the full height. This makes each tree less biased, but largely increases the variance. We aggregate the results from all the trees to get the final result. This reduces the variance. This gives us a very nice trade-off between bias and variance.

Chapter 5

Modelling Reflexivity - Feature Construction

5.1 Previous work

There has been extensive academic work studying the causality behind reflexivity. (Sornette, 2006) explored reflexivity as an endogenous process. Existence of reflexivity was not only proved, but a measure analogous to "criticality" safety measure in nuclear plant was also devised to quantify reflexivity (Filimonov & Sornette, 2012). Other notable work includes (Soros, 2015), who has provided philosophical background to understand the behaviour. (Soros, 2015) tries to delineate the differences in natural sciences and finance. In the world of finance, discovery of a theory affects the postulate of the of the theory. This is not typically the case in natural sciences (Soros, 2015). This was attributed to reflexivity. In simpler words, the behavior of agent may have a significant impact on the ensemble behavior. This is in direct contradiction against EMH (Chapter 2), where ensemble behavior is assumed to follow random walk.

Following the philosophy of self-exciting processes, we have developed a measure which can be useful in predicting the intermittent predictable patterns.

To put everything into perspective, we start with the assumptions that markets are reflexive in nature. Cycles of positive feedback exists along with the volatility clusters. Positive feedback follows the dynamics of multi-variate self exciting conditional point process.

5.2 Power law assumption

Next comes a very critical assumption, which is central to our theory. We assume that $\nu(t-s)$ term in Equation 3.1 follows decays with power law given by

$$y = ax^{-\gamma} \tag{5.1}$$

We believe the distribution of waiting time between the action leading to crowd gathering or herding behaviour (Sjöberg, Albrectsen, & Hjältén, 2000) is distributed as a power law. This assumption is borrowed from physics, finance (Bouchaud, 2001) and collective animal behaviour literature (Garcimartín et al., 2015). There is abundant finance literature observing and suggesting power law dependence. (Bouchaud, Gefen, Potters, & Wyart, 2004) analysed in Paris Bourse and observed the non-normality. (Bouchaud et al., 2006) observed and suggested power law dependence from empirical evidence based on a set of stocks from the Paris Bourse. (Lillo & Farmer, 2004) observed similar behavior at London Stock Exchange.

Estimating shape of the memory kernel or γ is not easy in this case. Here is where recent advances in computing techniques and machine learning would come to our rescue. We would synthesize an ensemble of γ and perform the corresponding transformations (Section 5.3). We would allow our machine learning function to learn the relationship between the future returns and the ensemble of memory kernels generated using the γ. This also offers us more degrees of freedom and thereby flexibility, as we are using ensemble of γ instead of committing to single value.

Since we are modelling as ensemble of γ for a fixed time horizon, we are making an implicit assumption that volatility clusters are generated by multi-variate self exciting conditional point process.

The proposed structure of power law decay is visualized in figure 5.1.

5.3 Bidirectional Reflexivity

We introduce a method to calculate multi-scale reflexivity for the financial time series. This measure takes into account the past price movements and encodes the information for different scales of activity.

In essence, for given time t, we take the weighted sum of historical returns for a fixed look-back time horizon (weighted with given γ). We then separate the positive and negative part, and take their ratio and call it $\mathcal{P}_\gamma(t)$, which is the feature value corresponding to γ and t.

The detailed procedure is described below. We call $S(t)$ the price of asset at time t, and $R(t)$ as return on the asset at time t, defined as $Log(\frac{S(t)}{S(t-1)})$. We define multi-scale response functions as

$$\phi_\gamma(\tilde{t}, t) = |t - \tilde{t}|^{-\gamma} \tag{5.2}$$

For a given γ and t, the response function can give a scaling factor at time $\tilde{t}$ according to $\phi_\gamma(\tilde{t}, t)$.

We construct the multi-scale kernel tensor with elements $M^{+}_{\gamma k}(t) = \phi_\gamma(k,t) * R^{+}(t-k)$ and $M^{-}_{\gamma k}(t) = \phi_\gamma(k,t) * R^{-}(t-k)$.

Here $R^{+}(t) = max(R(t), 0)$, $R^{-}(t) = min(R(t), 0)$, and

$k \in [0, \tau]$, τ being the look-back time horizon and $\gamma \in [0, \gamma_m]$.

Finally, we introduce the multi-scale reflexivity parameter for a given value of γ at time t,

$$\mathcal{P}_\gamma(t) = \frac{\sum_{k=0}^{\tau} M^{+}_{\gamma k}(t)}{\sum_{k=0}^{\tau} M^{-}_{\gamma k}(t)} \tag{5.3}$$

Another core assumption is that the reflexivity exist in bearish as well bullish market. Based on these two assumptions, our machine learning model will chose appropriate kernel and estimate the polarity for the currency pair.

5.3.1 Visualization

The proposed structure of feature weight decay ($\phi_\gamma(t)$ in equation 5.2) is visualized in figure 5.1.

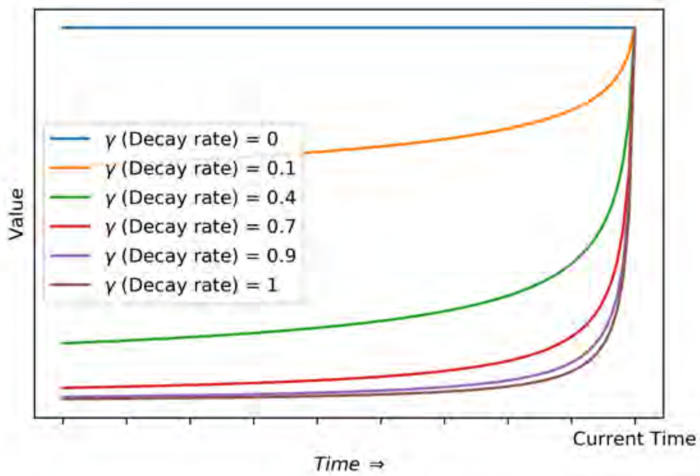

Figure 5.1: Feature weight decay for the past time horizon - exhaustive γ_i were considered for power law decay rate (while looking back in time).

The behaviour of real data is studied in the figure: 5.2. It is high frequency (1 min) tick data corresponding to AUDCAD pair from 22-10-07 13:18 to 09-

11-07 10:38. (The figure is currently shown for representation purpose only, our analysis was carried out on much bigger dataset)

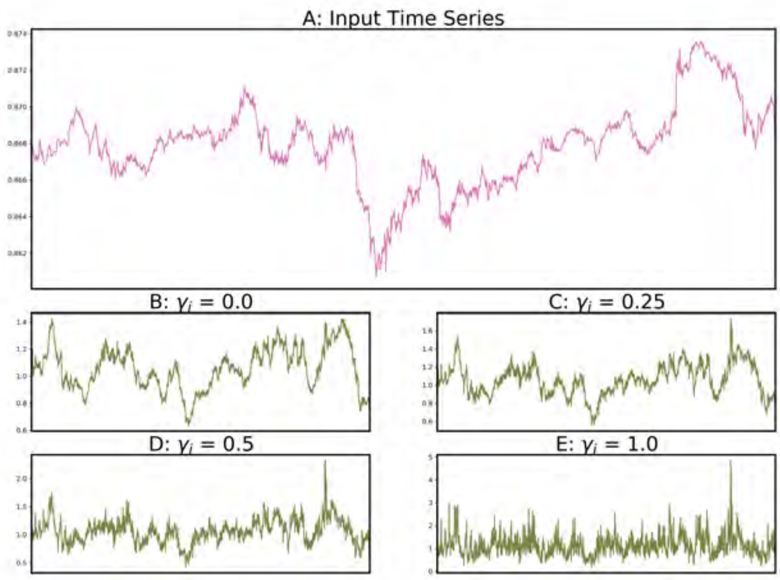

Figure 5.2: Feature construction on real data (AUDCAD pair from 22-10-07 13:18 to 09-11-07 10:38) - Wide range of γ_i in power law decay was applied on the given series to separate different decay rates and identify different scales of changes. Features were constructed based on Equation 5.3 and plotted as the subplots. Different γ_i are supposed capture different information about the history. When $\gamma_i = 1$, the feature constructed would capture the impact from the recent events more effectively. Whereas, $\gamma_i = 0$ would give equal weightage to all the events in given time window. All other γ_i would fall somewhere in between.

Kindly note how different features of the time series are captured with different γ in Figure 5.2. $\gamma = 1$ would give very high weight to the most recent event and all others weights would be negligible. $\gamma = 0$ would allot equal weights to all historical weights (Figure 5.1). For all other γs the weights would follow the distribution somewhere in between. That is exactly what is visible in Figure 5.2. For $\gamma = 1$, the feature peaks when there is sudden change in price time series and consequently also in the log returns. For $\gamma = 0$, as all the previous log returns are equally weighted, it would just act as average of all log returns in the given memory window.

All the possible cases were visualized on real world data and synthetic data exhaustively to study the effect of feature transformation.

Possible cases include monotonous decay, monotonous growth, Recent changes in trend, Change in behavior long ago (halfway in the time horizon), white noise, and sudden change.

First examples would be gradual change (Figure 5.3). Smaller values of γ_i capture this gradual trend change exactly, whereas larger values, Which are responsible for short term horizon, fluctuate more frequently as expected.

We would like to bring to reader's attentions, the elegant way in which different γ manage to capture a specific behaviour and thus specific γ corresponds to specific behavior.

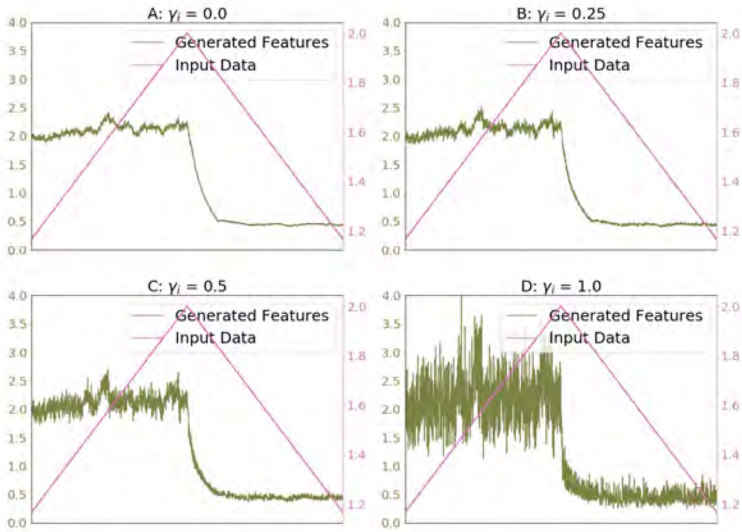

Figure 5.3: Growth for the first half and then decay for the rest, captured exactly by small γ_i

Second example would be sudden change (Figure 5.4). Here again short term γ (γ close to 1) are more effective in capturing the short term change.

The exhaustive analysis of all possible input time scenario for synthetic data and corresponding feature behaviour is presented in Appendix section A.1.

5.4 Target definition

Colloquially, we want to define target $\mathcal{B}(t)$, as the direction in next n time steps (Equation 5.4).

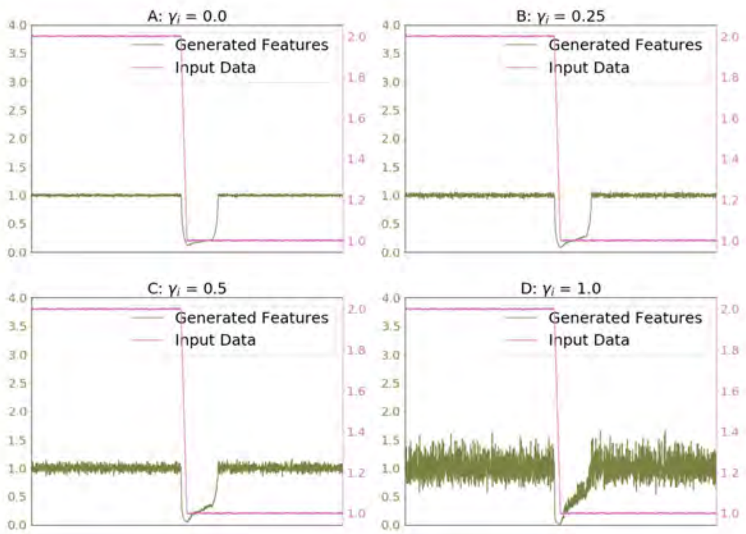

Figure 5.4: Sudden drop - captured precisely by higher γ, which corresponds to smaller time horizon

$$\mathcal{B}(t) = \begin{cases} 1, & \text{If prices strictly goes up in next n steps} \\ -1, & \text{If prices strictly goes down in next n steps} \\ 0, & \text{Otherwise} \end{cases} \tag{5.4}$$

We define $T_k(t)$ as $\sum_{j=1}^{k} R(t+j)$, $k \in [1, \tau_f]$ and introduce our target variable $\mathcal{B}(t)$ as Equation 5.5.

$$\mathcal{B}(t) = \begin{cases} 1, & \text{If } \forall\ k \in [1, \tau_f],\ T_k(t) > 0 \\ -1, & \text{If } \forall\ k \in [1, \tau_f],\ T_k(t) < 0 \\ 0, & \text{If } \exists\ k_1, k_2\ \in [1, \tau_f]\ \ni\ T_{k_1}(t) * T_{k_2}(t) < 0 \end{cases} \tag{5.5}$$

In order to get a meaningful prediction about $\mathcal{B}(t)$, we are interested in a supervised learning algorithm, which can learn the function $\mathcal{Q}$ from the past values of $\mathcal{P}_j(t)$ so that for the future time stamps t, we can calculate $\mathcal{B}(t)$ with $\mathcal{Q}(P_j(t)) \rightarrow \mathcal{B}(t)$. We will use random forest as per the discussion in Section 4.3, however any other machine learning should also work.

5.5 Data Preparation

Forex data was obtained from (*www.histdata.com*, 2018). The data is structured as exchange rate pairs and values are recorded at every minute. Our current Analysis is done on "AUDCAD" pair, one min tick data starting from "23-10-07 14:17" to "05-09-16 04:34". There are 3,304,315 data points in total.

5.5.1 Missing Value Treatment

There were two types of temporal discontinuity observed in the data, typical and atypical.

Typical

Typical temporal discontinuity resulted from holidays. Forex market is closed on weekend and standard holidays. Sometimes the temporal discontinuity was accompanied by large fluctuation (greater than 2 standard deviations). The fluctuation was accepted as the structure of the data, as it might have been caused because of sudden geopolitical shift. No treatment has been carried out for typical temporal discontinuities.

Atypical

Atypical discontinuity consisted of temporal discontinuity spanning less than a day. 1.5% of overall datapoints fall under such category. Linear interpolation was carried out for such atypical missing values.

Chapter 6

Trader design

Predicting direction or polarity of a stock is by design, more difficult problem than designing a trading algorithm. Minuscule amount of gain in overall accuracy can be easily magnified using leveraging techniques.

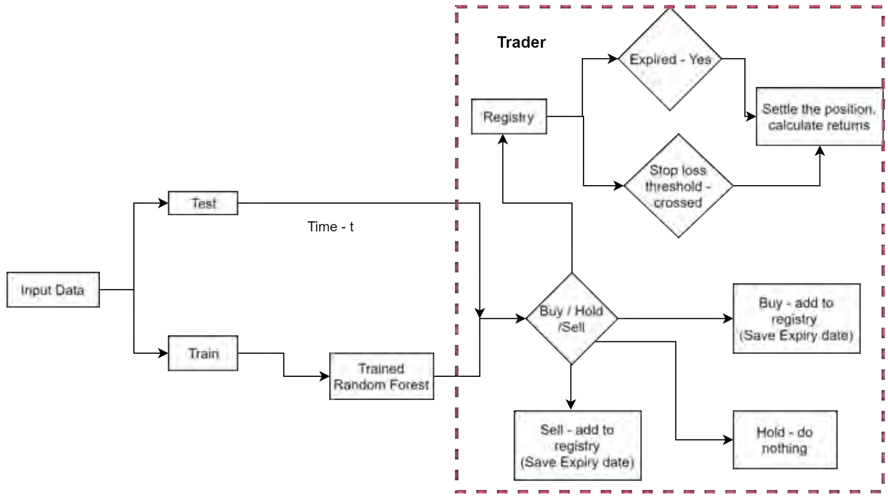

Figure 6.1: Trader Algorithm

6.1 Defining Direction

Predicting the direction for the next step accurately is next to impossible. Therefore we decided to predict the direction in next n-steps. There are two ways to define directions. First, would be defining direction based on changes from opening value to closing value within the n-step interval. However, this would not include information about highs and lows in this period. Therefore we opted for a different way. We defined direction as positive when both highs and lows for the given n-step period are above the current value. Similarly direction was defined as negative when both highs and lows were below the current value. Direction was assumed to be zero in rest of the cases. The same is described in mathematical notations in Equation 5.4. This provided us with a better sense of direction from trading point of view, as this was much more structured definition.

6.2 Trading Strategy

The suggested trading strategy (Figure 6.1) is quite simple. Reflexivity Trader (Trader based on our strategy) will be first trained training set. A time stamp would then be provided serially as the input for back testing and predictions would be noted. Reflexivity Trader would go long when the predicted direction is positive, and would go short when it is negative. Reflexivity trader will then make a note of transaction and transaction expiry time, which is determined by Maximum holding period. Maximum holding period is determined by n (defined in Section 6.3). As time progresses, transaction would be settled prematurely if stop loss threshold is breached. Otherwise, it would be settled at the expiry time.

6.2.1 Stop Loss

We implemented trailing stop loss strategy. Trader would settle the position earlier, if current value crosses a certain threshold. The threshold is determined by pre-determined allowable risk exposure and back testing. Since our gains in predictive power were minuscule, implementing stop loss was very important in our strategy in order to keep our risk exposure in check.

6.3 Holding Period

Holding period (n) was defined as maximum allowable time up to which the position will be held. The position can be settled prematurely if current loss crosses the threshold (It is τ_f in Equation 5.5).

6.3.1 Class Imbalance

Our way of defining holding period has imposed some restrictions on the choice of optimal n in n-steps. As holding period increases, current values becomes less and less likely to be either highest or lowest value. This is because, as we increase the time window in future, we are more likely to start including entire cycles and not just the path upwards or downwards in our window. Mathematically, by construction of $B(t)$ (Section 5.4), as we increase n, we introduce class imbalance between 0 and 1/-1 (Figure 6.2).

This would introduce an artifact of class imbalance, which should be avoided if possible. We ran simulation of class frequency count for different values of n (holding period). The results are shown in Figure 6.2 and 7.1.

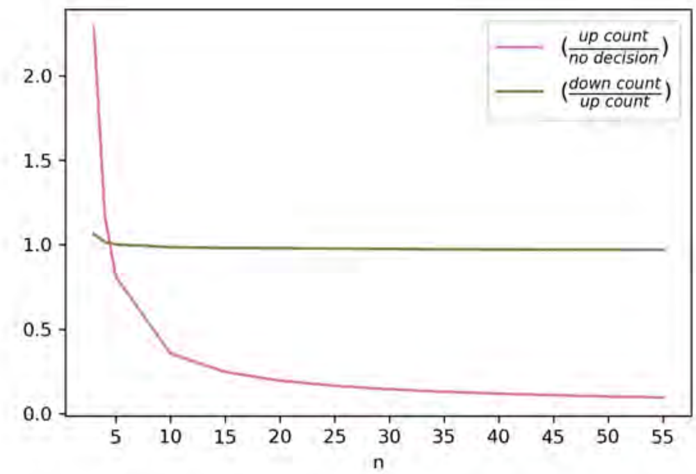

Figure 6.2: Here we are visualizing class frequency distribution against n (holding period). holding period is plotted on X axis. We are plotting two ratios on Y axis. Since we have a three class (-1,0,1) problem, we have plotted (Number of instances of class -1 / Number of instances class of 0) as one of the ratios and (Number of instances of class -1/ Number of instances of class 1) as other one to get a holistic picture. As n increases, the price tends to fluctuate more and hence, no decision count in training data increases rapidly. In other words, as n increases, number of instances of class 0 increase in proportion with either of the other two. The number of instances of class 1 remains same as that of class -1.

6.3.2 Precision gain

Another important factor, which we considered before deciding the holding period was gain in predictive precision. The results are discussed in Section 7.1.

Chapter 7

Results

7.1 Holding Period Analysis

Holding period is a hyper-parameter, that should be set according to the user requirements. We have designed an algorithm that would serve as guiding framework while deciding the holding period. First, n or holding period to be considered is narrowed down. Then target variable distribution for each of these n is calculated. Next random forest is trained on the train dataset and the results were tested on the test data for each n. The results are tabulated in Table 7.1.

7.1.1 Metric Selection

Since we are dealing with a three class problem with class imbalance, selection of metric is pivotal. Precision can be quite useful in our case. F1 score is also important but less useful than precision, rather the combination of precision score for all the scores is useful. By this we eliminate the biased measurements. Another useful criterion is Direction prediction percent.

Table 7.1: Holding period (n) analysis - Simulations were carried out for different n, while observing precision gain and class imbalance. Metrics observed were F1 score, percent of down prediction (Direct. Prediction), Precision and its improvements.

Holding Period	F1 Score	Direct. Pred %	Precision			Shuffled Precision			Improvement in Precision		
			-1	0	1	-1	0	1	-1	0	1
4	0.33	41%	0.34	0.35	0.33	0.32	0.35	0.32	3.23%	1.10%	3.47%
5	0.36	27%	0.29	0.43	0.29	0.28	0.43	0.29	3.97%	0.14%	3.15%
8	0.44	6%	0.23	0.57	0.23	0.21	0.57	0.22	6.14%	0.05%	4.86%
10	0.49	2.9%	0.21	0.62	0.21	0.19	0.62	0.19	10.10%	0.13%	10.45%
12	0.53	1.5%	0.19	0.66	0.20	0.17	0.66	0.17	10.21%	0.04%	13.97%
15	0.57	0.7%	0.17	0.70	0.17	0.15	0.70	0.15	14.83%	0.14%	12.56%
18	0.61	0.4%	0.16	0.73	0.16	0.14	0.73	0.14	15.58%	0.01%	18.86%
20	0.63	0.3%	0.15	0.74	0.16	0.13	0.74	0.13	19.11%	0.06%	26.29%

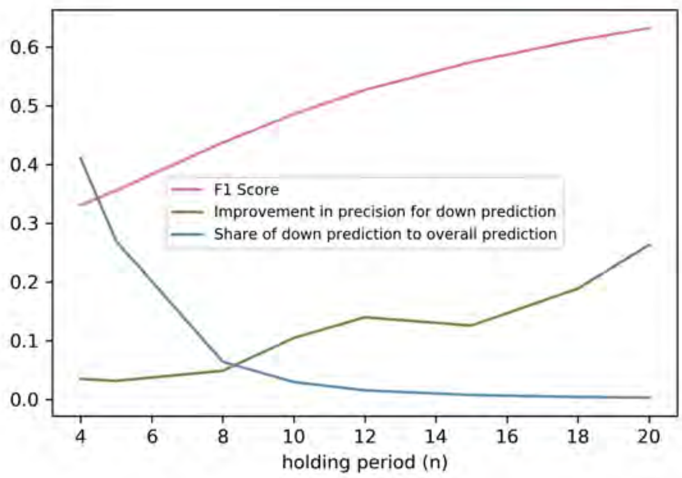

Figure 7.1: Visualization of Table 7.1 - We are comparing class imbalance introduced verses precision gained in order to decide optimal holding period (n). Here holding period (n) is plotted on X axis. On Y axis, F1 score is plotted in its absolute value along with two ratios. For fractional improvement as one of these ratios, two values of precision were calculated. First is the precision of predicted data with respect to test data and second is the precision of the randomly shuffled test data with test data. Fractional improvement was calculated as fractional change of the first precision from the second and plotted as Improvement in precision for down prediction. Class imbalance ratio (number of predictions of class -1 / total instances) is the other ratio which is plotted as share of down prediction to overall prediction. From the figure, it is apparent that F1 score is misleading metric in our case as it does not take the class imbalance ratio into account.

Direction Prediction Percent

Direction prediction percent is fraction of down predictions in overall predictions. As holding period is increased, entire cycles of fluctuations start getting included in the holding period. This increases the occurrences of zeros in target variables and reduces ones and minus ones. Thus direction prediction percent starts reducing with increasing holding period (n), and causes class imbalance, which in turn is an important parameter that is needed to be considered before applying a machine learning technique.

Precision

Precision is defined as the ratio of true positives to total number of positives predicted. It intuitively makes sense in our case.

F1 Score

Precision would produce different number for each class. Our analysis might become far easier, if we get a single number across all classes instead. Therefore we considered F1 score. It is harmonic mean of precision and recall. However, if we calculate F1 score over entire data, we are still missing the important artifact introduced by class imbalance. There are several reasons (Powers, 2011) why F1 score is criticized and has been observed to be biased as an evaluation metric. F1 Score values are misleading in our case as well (Table 7.1). Therefore we need to consider each class separately and hence we decided on class wise precision.

7.1.2 Shuffling

In order to calculate the incremental jump we are getting, we compare the prediction from our model to a noise predictor. We shuffle the actual test values randomly to get noise predictions. This approach maintains the class imbalance, hence it is not a confounding variable anymore.

We then calculate the precision of this shuffling against the test data. This gives a baseline, above which we can quantify our gains in precision.

7.1.3 Data Sufficiency

To check if the test size is large enough, we carried out data sufficiency analysis and visualized the growth in precision with time or sample size in the test data set (in serial manner). Form Figure 7.2, it is apparent that precision values start stabilizing after 100000 data points. Temporal evolution of precision is tested separately for possibility of discrepancies in temporal values. The results are shown in Appendix (Figure A.14).

7.1.4 Conclusion

Figure 7.1 and Table 7.1 provides an algorithm to select future time horizon based on the client requirements. For now, we have selected our future time horizon as ten minutes. It offers optimum trade-off between class-imbalance and precision gain. We not only want to gain precision, but we also wish to make enough number of predictions to make profit. In Figure 7.1, the trade-off is visible.

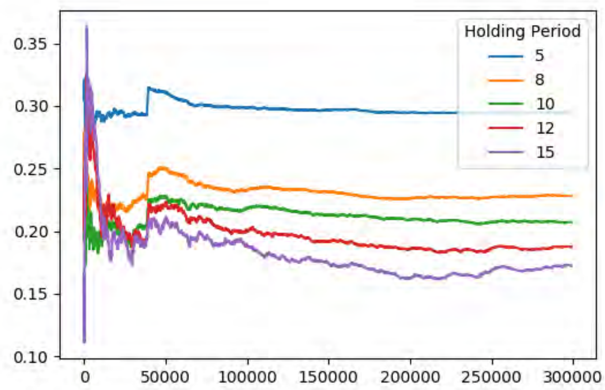

Figure 7.2: Precision is plotted against increasing sample size. Precision values start stabilizing after 100000 data points.

When n is ten, we get around 10% down prediction and around 10% jump in precision.

7.2 Testing for Reflexivity

Our first set of tests includes the test for reflexivity. With this test we want to check that our predictions are because of reflexivity and not because of the machine learning model. For this we test the results from random forest with hand designed input as features against null model. We generate the null model as geometric Brownian motion. With this procedure, we get rid of volatility clusters, which are quintessential to financial time series. This null model checks for the possibility of forward bias as well. It ensures that we are not unknowingly providing future information. In presence of a forward bias, the model should be capable of predicting white noise with the same accuracy. We would use the series shown in Figure 2.1 as null model (We would like to remind the reader that for this plot, we have generated white noise log returns). We would then generate features in the same way as described in Chapter 5. We would then test both the models.

7.2.1 Cross Validation

In order to get unbiased estimate of the results, we are testing them on a separate test dataset. As written in Section 5.5, there are approximately 3.3 million data points in total. We will train the model on the first 3,004,315 data points (23-10-07 14:17 to 15-11-15 17:18 for "AUDCAD" pair) and test on the last 300,000 data points (15-11-15 17:18 to 05-09-16 04:34, "AUDCAD" pair). For feature construction, we will select a look back horizon (τ) of 1500 datapoints (Which approximates to one day).

7.2.2 Results from Random Forest

We are visualizing two aspects of the results, confusion matrix for precision and Variable importance factor (VIF).

VIF

VIF is a very useful tool available with Random forest. Unlike Neural network, with VIF random forests gain explainability. It is a visual depiction of variable importance in classification. Figure 7.3 shows the importance of features in deciding the direction.

Confusion Matrix

We have also plotted confusion matrix (Figure 7.4) to better visualize the incremental prediction gain. We are comparing Precision for each class. Figure 7.4 is another visualization of the part of the information in Table 7.1.

7.2.3 Null Model Comparison

As discussed in previous sections, we generated synthetic data using geometric Brownian motion to break the inherent volatility clusters present in the real data. We then applied same procedure, i.e. created the features to capture reflexivity, trained Random forest and cross validated it. In Figure 7.5 the results from null model variable importance factor are plotted. There is a clear distinction between the VIF plots, as different features are influential for both model.

Figure 7.6 shows the confusion matrix for precision calculated on the null model. When compared with Figure 7.4, precision gain is clearly visible.

7.2.4 Conclusion

From Figure 7.6 and Figure 7.4, the gain in precision is very clear. Therefore we can reject the null hypothesis that the forex data in not reflexive and can also validate that the history of log returns exhibit a degree of predictive power. Since number of features is a hyper-parameter, a number of tests were carried out by varying number for features. The results are discussed in great depth in Section 7.3.6.

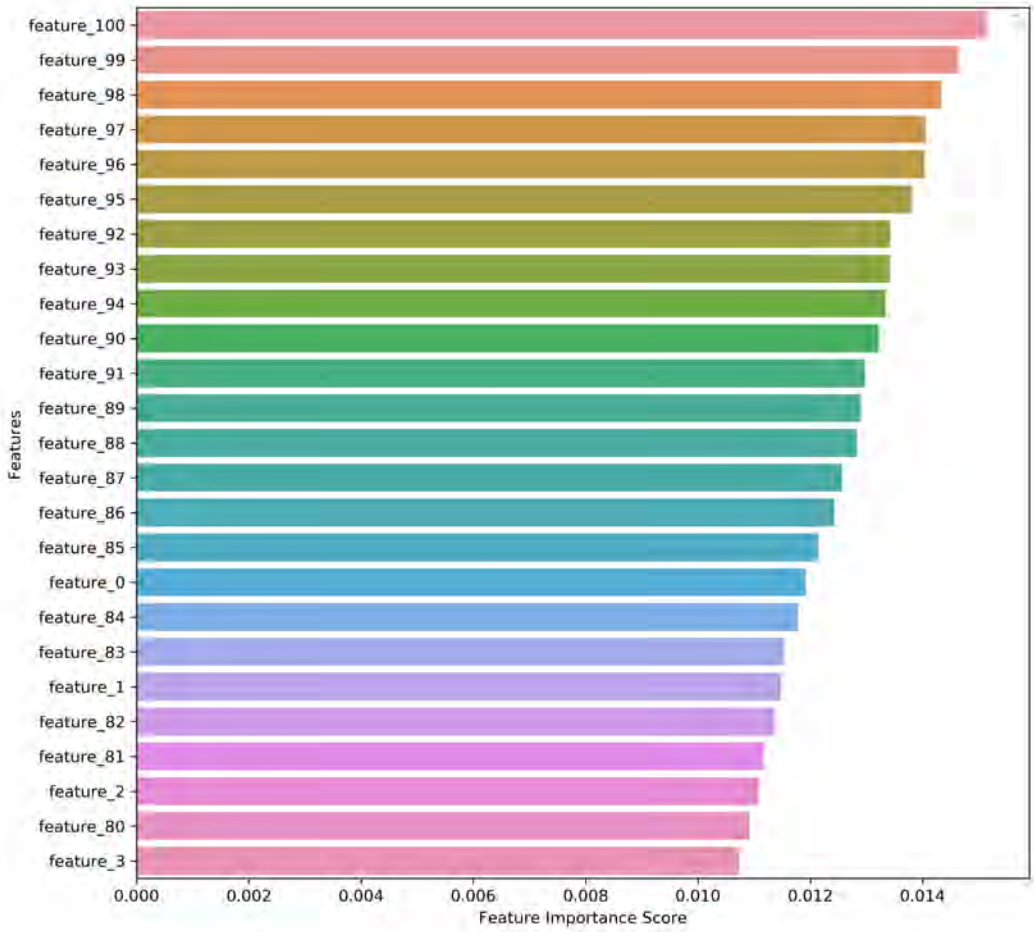

Figure 7.3: Variable importance factor on real data - feature_100 corresponds to $\gamma = 1$ and feature_0 corresponds to $\gamma = 0$ in equation 5.2, all other features fall in between.

7.3 Trader Testing

After we have tested and validated the existence of reflexivity, A trader which will utilize arbitrage opportunities produced from intermittent reflexivity was designed. The design of the trader is discussed in detail in Chapter 6. The trader was also used to test the effectiveness of number of features which were

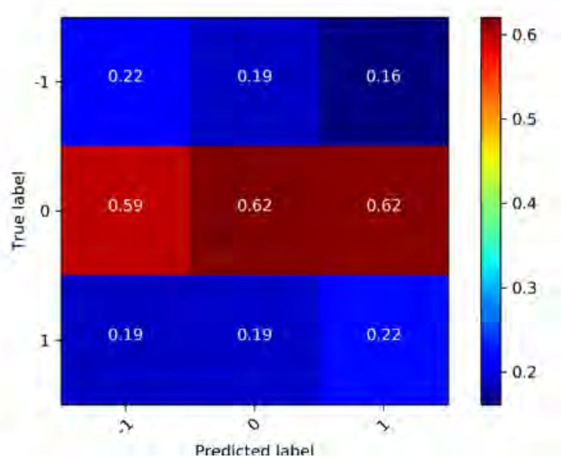

Figure 7.4: Normalized Confusion matrix for Random forest trained on the real data - The entries correspond to precision of each variable.

considered.

7.3.1 Cross Validation

We always kept the training and testing set separate to get reliable estimates. Analysis is done on "AUDCAD" pair, one min tick data starting from "23-10-07 14:17" to "05-09-16 04:34". There are 3,304,315 data points in total. (Section 5.5). We trained our random forest model on approximately 3 million datapoints from 23-10-07 14:17 to 15-11-15 17:18. We later fine tuned the threshold on the first 150,000 datapoints (from 15-11-15 17:18 to 12-04-16 00:32) after it. Finally, we validated on the next 150,000 datapoints (from 12-04-16 00:32 to 05-09-16 04:34) to get an estimate of real performance.

7.3.2 Trading Fees

Trading fee is not easy to estimate in the forex market. Typically trading fee is determined by the spread. Two different prices are typically quoted for currency pairs: the bid and ask price. The "bid" is the price at which you can sell the currency. The "ask" is the price at which you can buy the currency. The difference between bid and ask is known as the spread. The spread is typically the trading fee. The spread is measured in pips, which is the smallest unit of price movement of a currency pair. for our dataset of AUDCAD pair 1 pip = 10^{-5}. We also ran the simulations to study the performance of the trader with varying trading fees (Figure 7.7 and 7.8). As expected, trader performance increases with reduced trading fees.

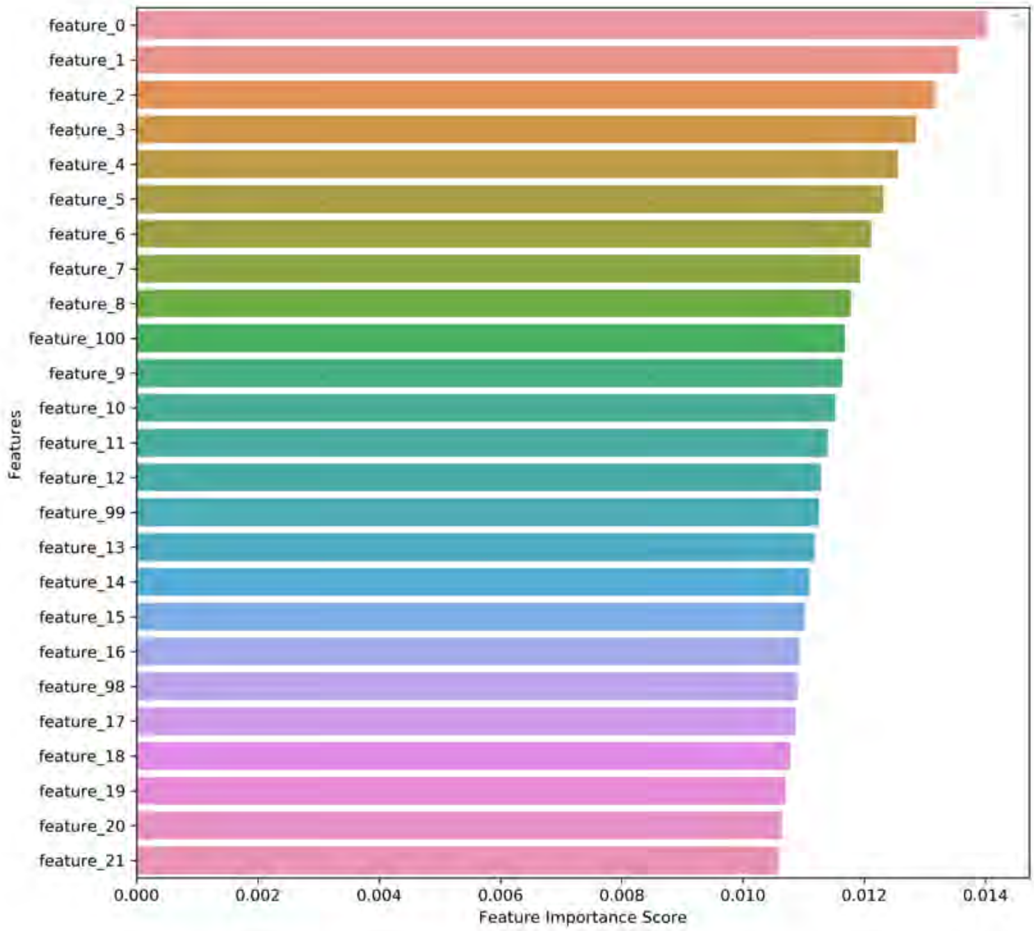

Figure 7.5: Variable importance factor on synthetic data - feature_100 corresponds to $\gamma = 1$ and feature_0 corresponds to $\gamma = 0$ in equation 5.2, all other features fall in between.

7.3.3 Threshold Determination

While designing the trader, we implemented trailing stop loss strategy (Subsection 6.2.1). This strategy requires fixing a threshold. As domain knowledge is insufficient to decide this, we decided to leverage the data to design such threshold. We ran simulations for different thresholds and studied it. Since trading fees are not fixed, we also observed the change in behavior for different trading

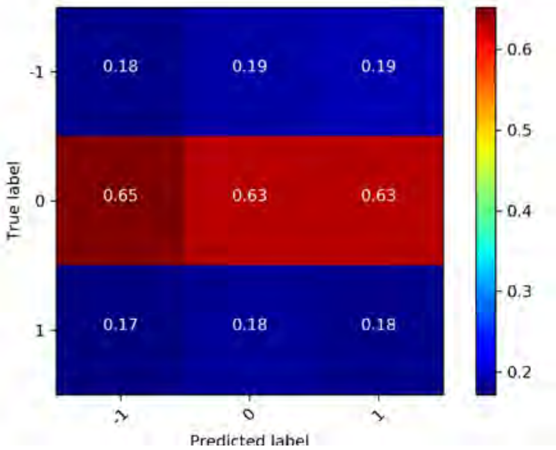

Figure 7.6: Normalized Confusion matrix for Random forest trained on the null data - The entries correspond to precision of each class. When compared with Figure 7.4, incremental gain is clearly visible.

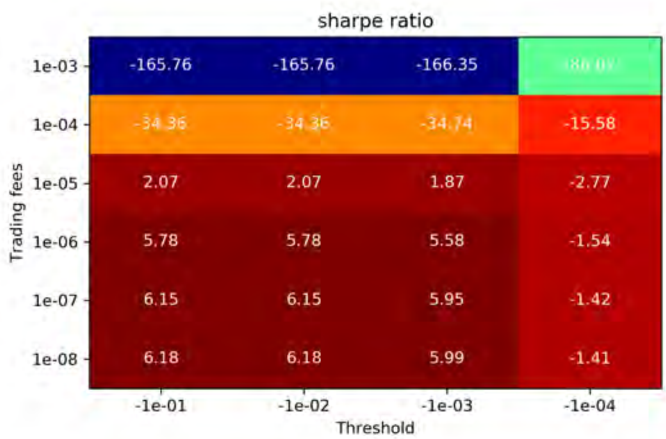

Figure 7.7: Sharpe ratio measures excess return per unit deviation. It is defined as $S_a = \frac{E[R_a - R_b]}{\sigma_a} = \frac{E[R_a - R_b]}{\sqrt{\mathrm{var}[R_a - R_b]}}$. Simulations of Reflexivity Trader (Our suggested algorithm) were carried out for different trading fees and thresholds and corresponding Sharpe ratios were observed.

fees. As the threshold is moved away from the zero, it rarely activates. Thus, the performance starts attending a limiting value as one pushes the threshold to limiting value. On the other hand, when threshold is closer to zero, it gets activated quite often, and one looses money on trading fees.

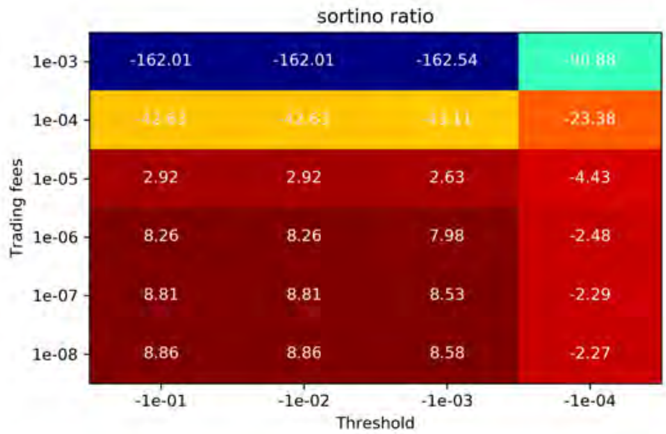

Figure 7.8: Sortino ratio is the variation of Sharpe ratio that factors only the downside risk instead of total standard deviation. $S_a = \frac{E[R_a - R_b]}{\sigma_d}$, where σ_d corresponds to downside deviation. Simulations of Reflexivity Trader (Our suggested algorithm) were carried out for different trading fees and thresholds and corresponding Sortino ratios were observed.

Metrics

We decided to measure efficacy of our algorithm with following metrics.

1. **Sharpe Ratio -** Sharpe ratio is universally accepted metric used to measure risk adjusted performance of an investment. It measures excess return per unit deviation. It is defined as $S_a = \frac{E[R_a - R_b]}{\sigma_a} = \frac{E[R_a - R_b]}{\sqrt{\mathrm{var}[R_a - R_b]}}$. Results are shown in Figure 7.7.

2. **Sortino Ratio -** Sortino ratio is the variation of Sharpe ratio that factors only the downside risk instead of total standard deviation. $S_a = \frac{E[R_a - R_b]}{\sigma_d}$, where σ_d corresponds to downside deviation. Results are shown in Figure 7.8.

For all these ratios, bigger value implies higher profitability for the trader.

7.3.4 Results for different Thresholds

Threshold

Based on Figure 7.7 and 7.8 threshold away from zero, like -0.01 seemed like good choice. For threshold close to zero, trader will have no patience and will settle almost instantly. Threshold of -0.01 means that, If overall change drops below -0.01, we would say that our prediction was incorrect and we will settle the position in order to manage the losses. It was also observed that if the threshold becomes too low (very close to zero), it always activates and minuscule gains achieved are eaten up by the trading fees. If our predictions are good enough, we do not want threshold to be activated often. The results validate the efficacy of our predictions, as the metric starts attaining a limiting value below some threshold.

Trading Fees

Another important behaviour observed is the behavior of the trader for different trading fees. If trading fees are below 10 pips our trader is profitable for a weak threshold away from zero. If trading fees are more than or equal to 10 pips, then irrespective of the threshold, our trader can never be profitable. Thankfully, spread below 10 pips is frequently observed in the forex market.

7.3.5 Testing against other Strategies

In order to do holistic testing, we ran the test against the traders utilizing other strategies: namely buy and hold, and noise trader. As holding period is a parameter decided by the user, we ran the tests for different holding periods as well. We fixed the the threshold to -0.01 and trading fees to 10^{-5} or 1 pip.

Buy and Hold

Buy and hold is a passive investment strategy in which assets are bought and held for a long period regardless of fluctuations in the market.

Noise Trader

When the decision are made randomly, we call the strategy as noise trading. Since we wanted to test the fact the gains are due to the predictions and not because of the strategy, we kept the fraction of decision constant, and randomly shuffled the predictions and provided it as an input to the trader.

The result from the analysis are shown in Table 7.2. The algorithm suggested, reflexivity trader, is profitable than other two trading strategies. Same results are visualized in Figure 7.10,7.9, and 7.11.

In Figure 7.10, the trader does not have enough patience, and hence performs poorly. In Figure 7.11, trader has too much patience, and which affects its performance. From Figure 7.9, it is apparent that our trading strategy beats both noise trader, buy and hold by significant margin for appropriate hyper

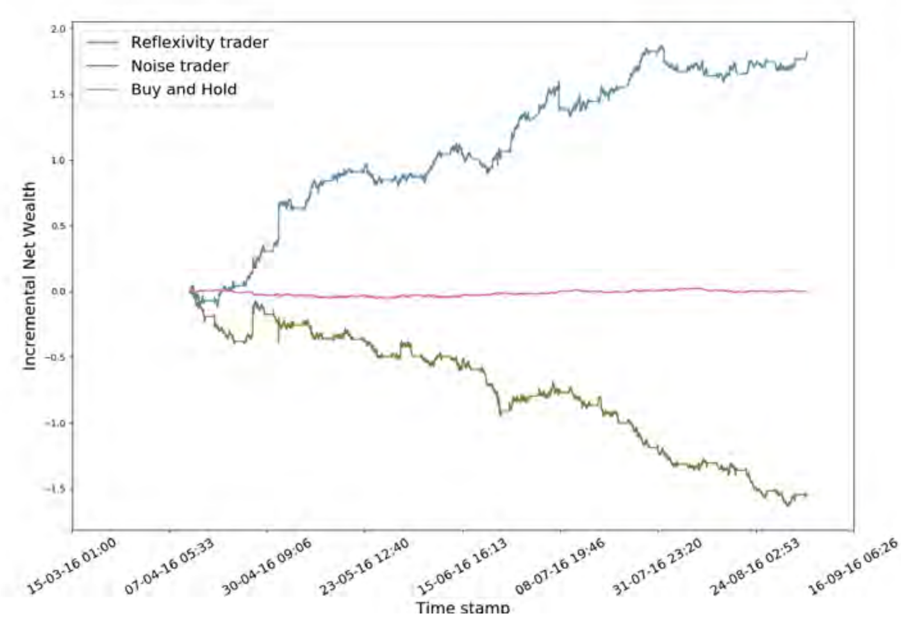

Figure 7.9: Comparing absolute returns of the trader on the test set for holding period of 10 steps - Reflexivity Trader(our algorithm) did significantly better than other strategies.

parameters. Holding period of 10 appears to be the the sweetest spot for this case. Exhaustive analysis for different holding periods is visualized in appendix section A.2.

7.3.6 Input Features Analysis

For the analysis presented above, we have fixed the number of of features. The trader is working and profitable. For further improvement, one natural question that would arise is, are considering too many features? Are we learning some amount of noise instead of trend? For this we conducted a small analysis by reducing the number of input features. Since the features on extreme appeared to have more impact in Figure 7.3 and 7.5, while reducing the number features, we considered the features only from either of the extremes. The results are in Table 7.3. Further analysis is needed to explain the jump in the trader performance, when only 20 features are considered. One possible explanation could be that we are unnecessarily dealing with 100 features, when only 20 would be sufficient. However, extensive simulations should be carried out substantiate this claim.

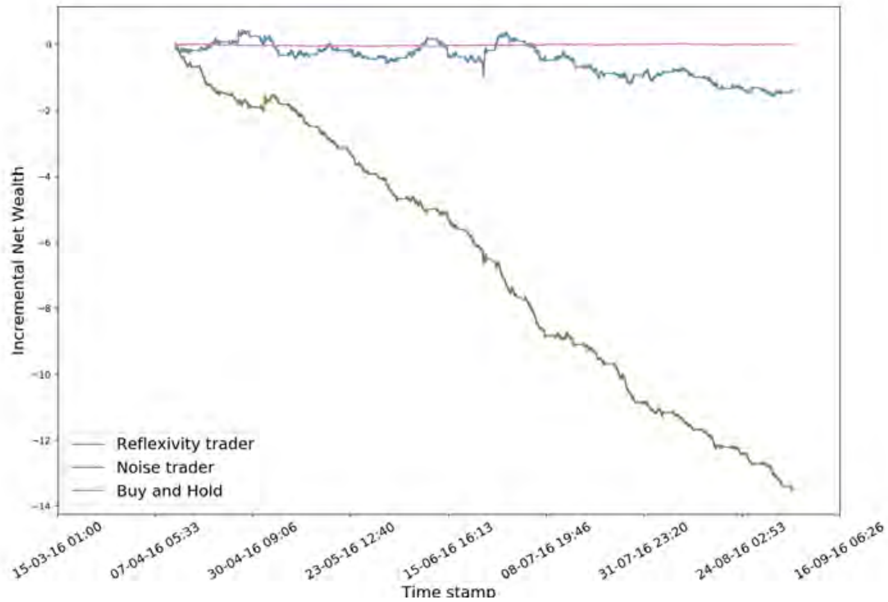

Figure 7.10: Comparing absolute returns of the trader on the test set for holding period of 4 steps - Reflexivity Trader(our algorithm) did worse than buy and hold in absolute terms.

7.3.7 Conclusion

In order to get a quantitative measure of the performance we compared the standard ratios (Sharpe Ratio and Sortino Ratio). The ratios are discussed in great depth in section 7.3.3. The results are tabulated in Table 7.2. Noise trader appears to perform the worst, even worse than buy and hold. It is to be expected because of the losses incurred by trading fees. Reflexivity trader appears to beat Buy and hold strategy by margin, when Sharpe and Sortino ratios are considered.

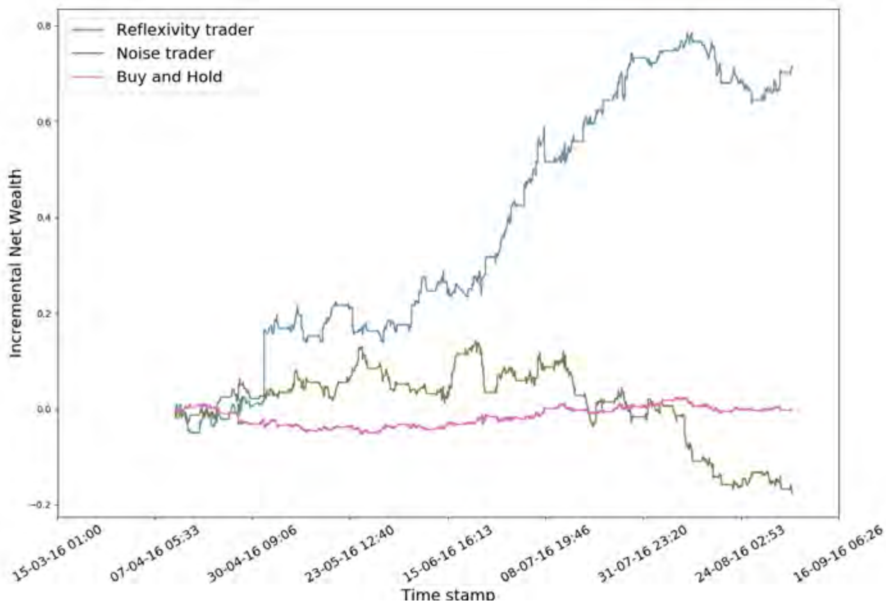

Figure 7.11: Comparing absolute returns of the trader on the test set for holding period of 20 steps - although Reflexivity Trader (our algorithm) did better than other strategies, performance gain is negligible.

Holding	Sharpe Ratio			Sortino Ratio			No. of
Period	Reflexivity	Noise	Buy & Hold	Reflexivity	Noise	Buy & Hold	Trades
4	-2.35	-21.98	-0.07	-3.30	-29.88	-0.10	122228
5	0.54	-16.86	-0.05	0.77	-23.24	-0.07	84765
8	3.91	-9.06	-0.05	5.97	-12.45	-0.07	21712
10	6.56	-5.56	-0.05	10.52	-8.00	-0.06	10058
12	5.70	-4.84	-0.05	9.63	-6.92	-0.06	5416
15	5.58	-2.61	-0.04	10.28	-4.10	-0.06	2582
18	5.55	-2.76	-0.05	10.14	-3.36	-0.07	1409
20	4.94	-1.72	-0.06	9.47	-2.36	-0.08	1049

Table 7.2: Reflexivity Trader (Strategy proposed by us) performs consistently better than buy & hold and noise trader. As expected, as holding period increases number of trades reduce.

No. of Features	Sharpe ratio	Sortino ratio	No. of Trades
10	1.43	2.04	14156
20	4.37	6.93	12327
30	1.94	2.86	11363
50	3.16	4.80	10019

Table 7.3: Trader performance for different input size - There is an improvement in trader performance when only 20 features are considered.

Chapter 8

Conclusion

8.1 Primary Objective

The primary goal of the thesis was to model the reflexivity and substantiate its presence in the forex market. In Section 7.2.3, we rejected the null model, which was developed assuming that the currency markets are efficient. We have created features which are capable of capturing intermittent reflexive behavior of the market. These features were later used to predict the future drift of the price time series using Random forest classifier. Predictive success validates our hypothesis that forex markets are intermittently reflexive.

8.2 Secondary Objective

The secondary objective was to utilize the reflexive behaviour and design a trader capable of providing better results with lesser risk as compared to standard market practices. In Section 7.3.5, we compared our trading strategy, which makes the decision based on reflexivity, to other trading practices, such as noise trader and Buy and Hold. Our trader performed better than others in the high frequency regime.

8.3 Future Works

We have already established the presence of reflexivity. We have also designed a primary trader capable of providing higher risk with lower volatility. The next step would be to improve this trader. One way to improve the trader would be to use better machine learning technique. Section 7.3.6 hinted than better features selection may hold potential to improve the trader performance.

Another approach, would be to use more sophisticated algorithm. In (Moody, Wu, Liao, & Saffell, 1998), implementation of reinforcement learning for algorithmic trading is discussed in great details. In the current version, we are solving two problems separately - namely direction prediction and trading decision. Reinforcement learning would combine both problems and has potential to outperform the current trader.

Appendix A

Appendix

A.1 Reflexivity Features for synthetic data

In this section we will exhaustively discuss the behavior of the feature construction for all possible data structures. Each decay rate (γ_i) is supposed to capture different aspect.

Intuitively, feature value for given time horizon will be greater than one if there is overall growth in the given horizon. In the opposite scenario, when there is overall decay then numerical value will be less than one.

A.1.1 Continuous decay

In this case, we are considering continuous decay. Here, we expect feature value to be less than one for the entire range of γ (Figure A.1).

A.1.2 Long term decay - Short term uncertainty

Here the price is following decay in the long term horizon. However in short term the price follows no particular direction. Higher values of γ_i will grow quickly towards the end in this case (Figure A.2).

A.1.3 Long term decay - short term growth

Here long downward drift of price movement is interrupted by short term growth. The growth in higher value of γ_i is more apparent than the previous case (Figure A.3).

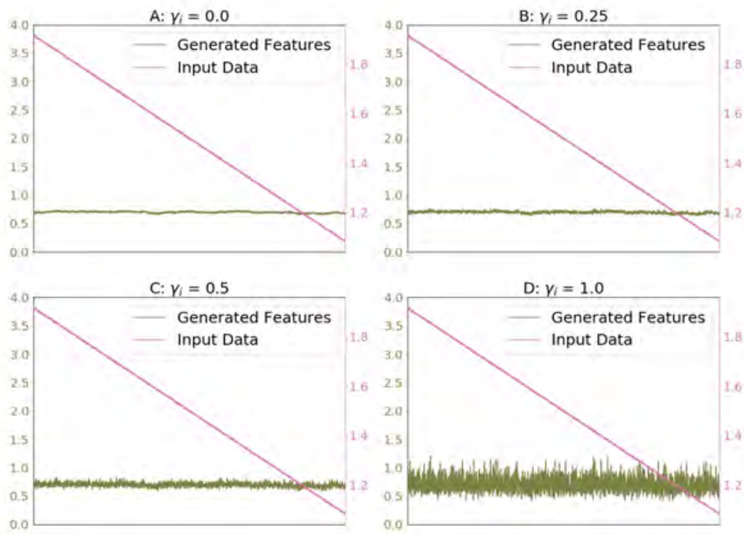

Figure A.1: We are considering overall decay (Decay in both long term and short term horizon)

A.1.4 Long term constant - short term decay

Here, overall change in the complete time horizon is zero. Price grows for the first half, and then decays equally for the rest. Smaller values of γ_i capture this trend change exactly, whereas larger values fluctuate more frequently (Figure 5.3).

A.1.5 Only fluctuations

Here the series follows random walk and fluctuates completely. All γ_i's fail to capture the trend (Figure A.4).

A.1.6 First decay and then growth

This is exactly opposite of the case discussed in Section A.1.4. Here price decays for the first half and then grows equally for the second. As expected the changes are best captured by the smaller values of γ_i's (Figure A.5).

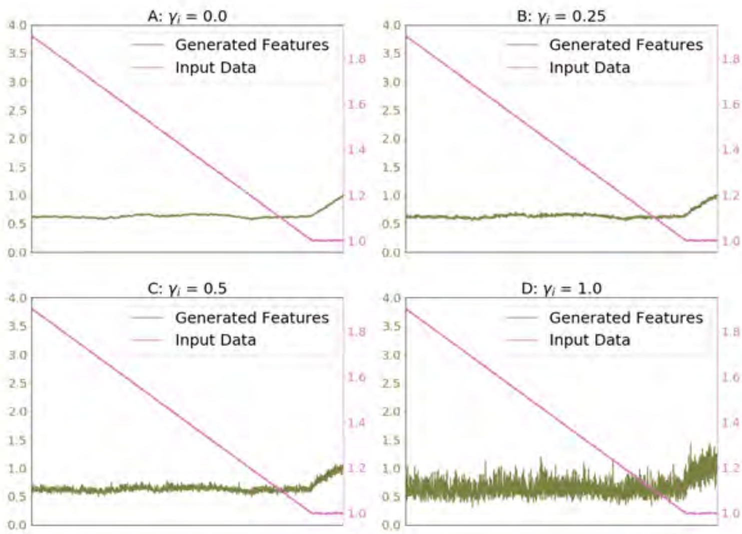

Figure A.2: We are considering long term decay but short term fluctuations

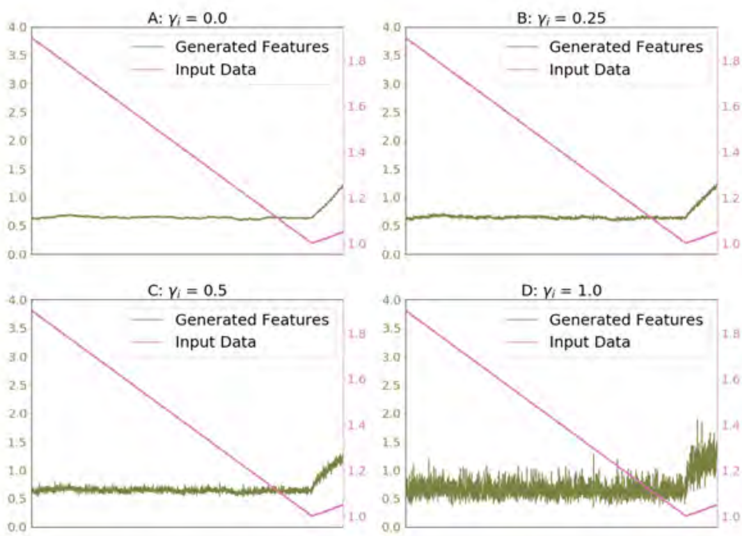

Figure A.3: We are considering long term decay but short term growth

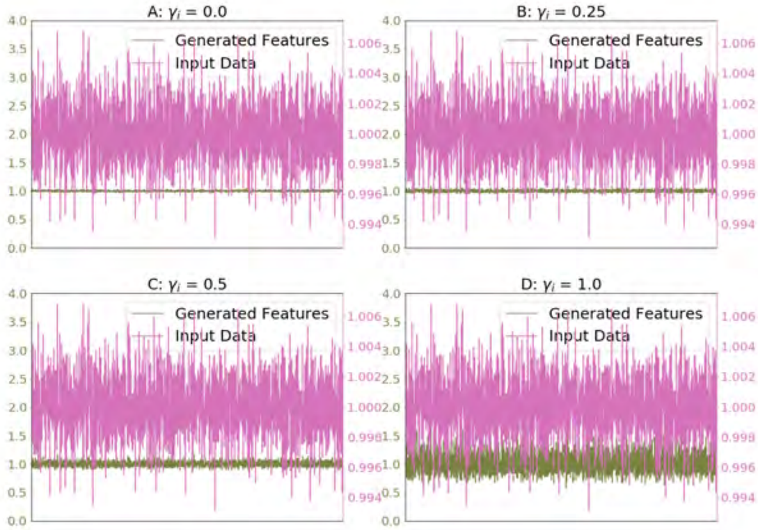

Figure A.4: Random walk, All γ_i's fail to capture any trends

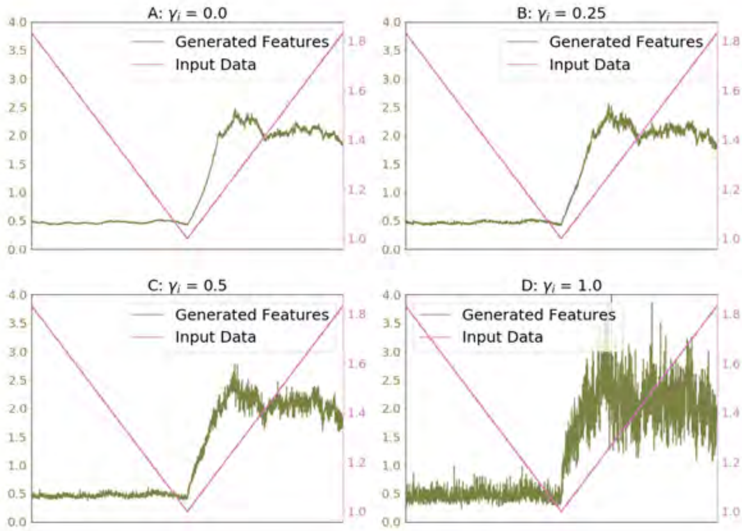

Figure A.5: Decay for the first half and then equal growth for the rest, captures exactly by small γ_i corresponding to longer time horizons

A.1.7 Long term growth and short term decay

We are considering the series which has grown significantly in the longer time horizon, but has started to decay recently. Again the changes are captured quite nicely by the range of γ_i offered (Figure A.6).

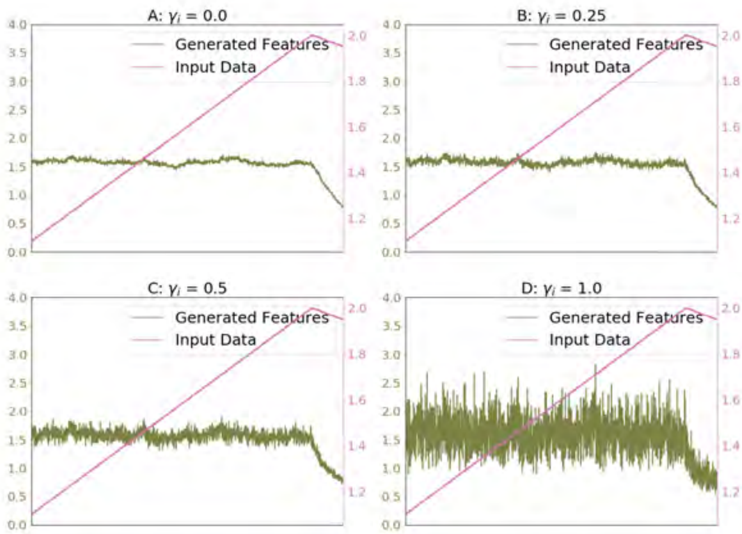

Figure A.6: We are considering long term growth but short term decay

A.1.8 Long term growth and recent uncertainty

The series has been growing constantly till some point, and then it started fluctuating around a mean. The changes are captured by smaller γ_i, however the change is not as drastic as the previous case (Figure A.7).

A.1.9 Continuous growth

This is exact opposite of the first case discussed, as we have monotonous growth. Small γ_i's have captured this trend quite effectively (Figure A.8).

A.1.10 Special case - sudden drop

Here we are considering the special case of sudden drop in price. Higher values of γ_i are effectively designed to capture this scenario (Figure 5.4).

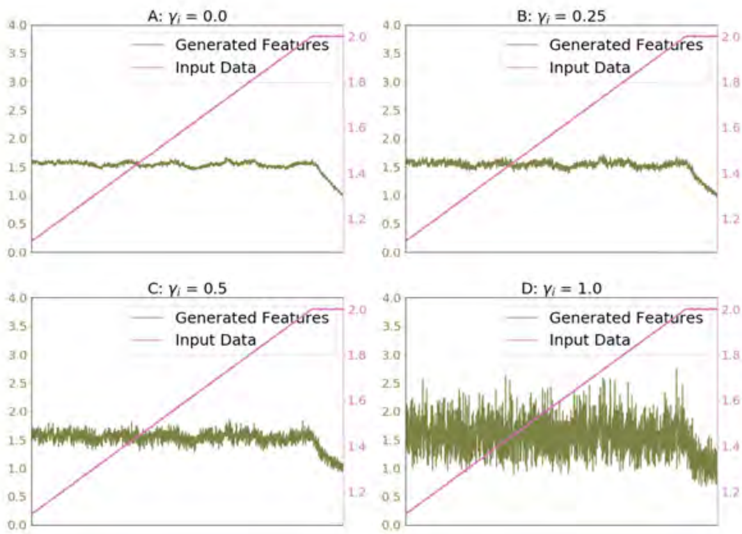

Figure A.7: Price has historically grown, but recently became uncertain

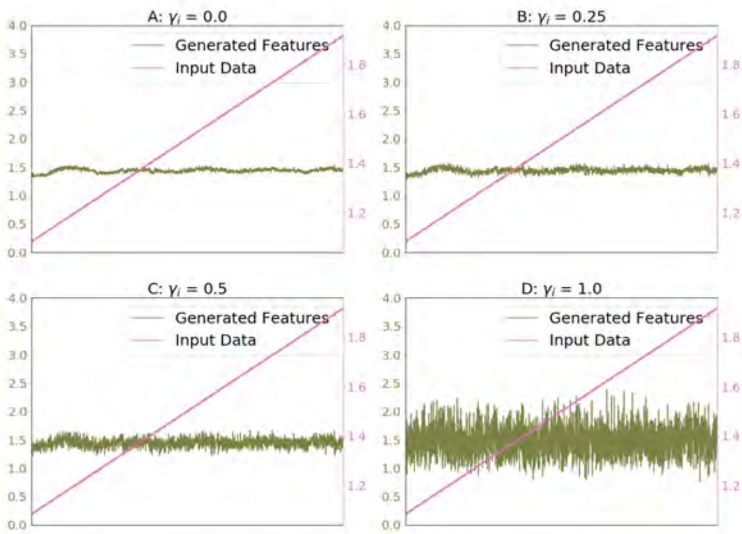

Figure A.8: Monotonous growth

A.2 Trader Visualization

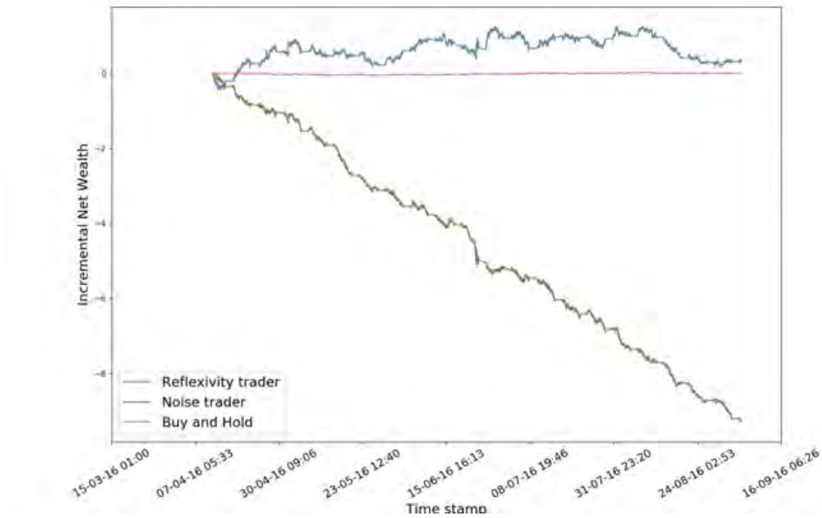

Figure A.9: Comparing absolute returns of the trader on the test set for holding period of 5 steps - Reflexivity Trader(our algorithm) started doing a bit better

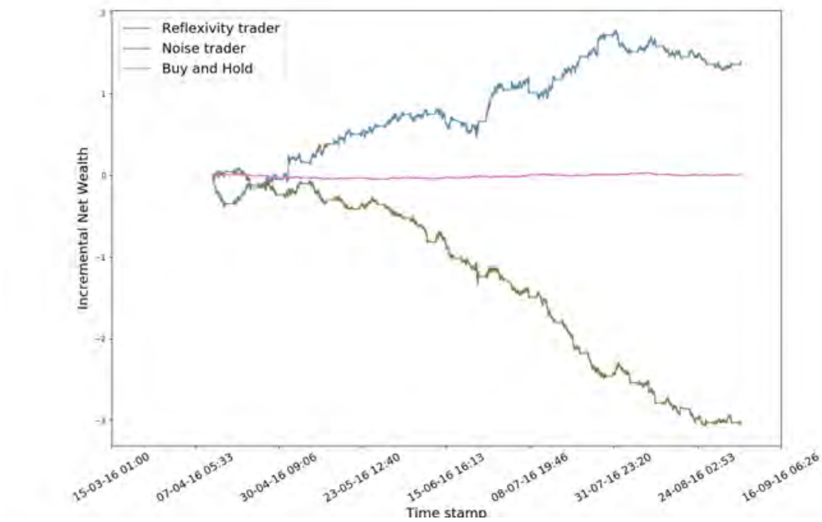

Figure A.10: Comparing absolute returns of the trader on the test set for holding period of 8 steps - Reflexivity Trader(our algorithm) has enough patience and does significantly better.

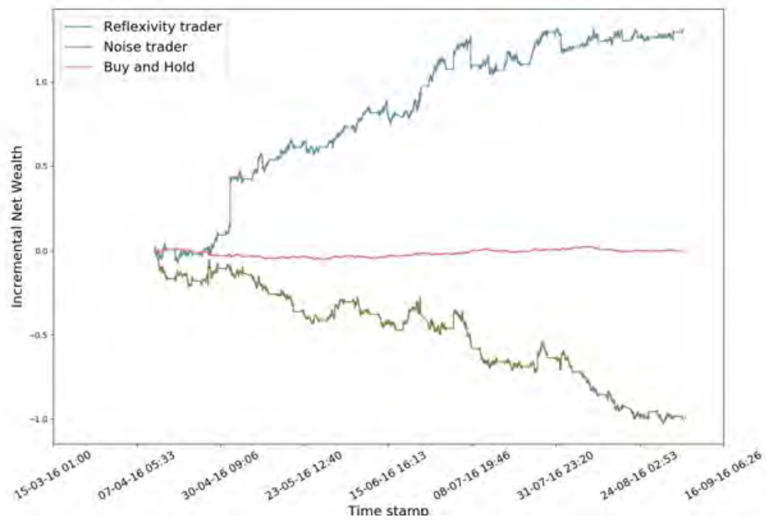

Figure A.11: Comparing absolute returns of the trader on the test set for holding period of 12 steps - Reflexivity Trader's (our algorithm) performance starts reducing.

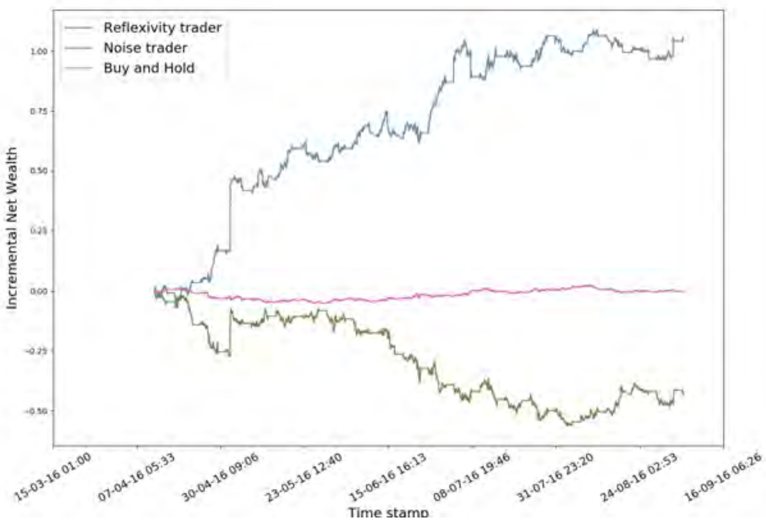

Figure A.12: Comparing absolute returns of the trader on the test set for holding period of 15 steps - Reflexivity Trader's (our algorithm) performance further reduces.

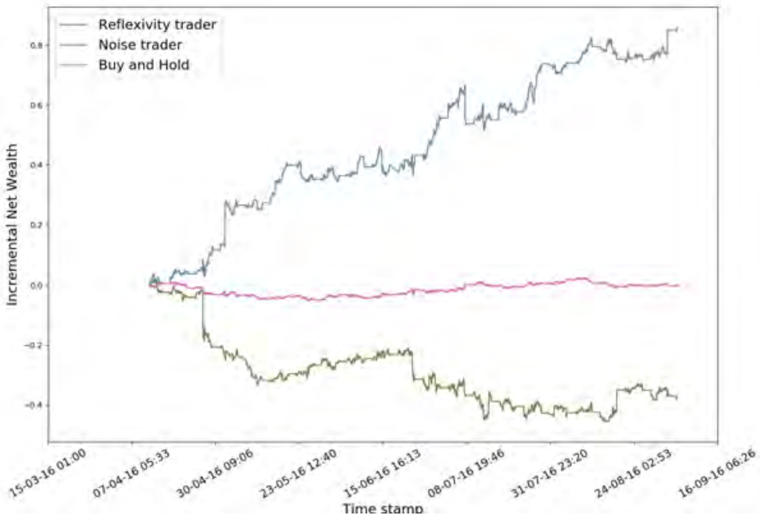

Figure A.13: Comparing absolute returns of the trader on the test set for holding period of 18 steps - Reflexivity Trader's (our algorithm) performance further reduces.

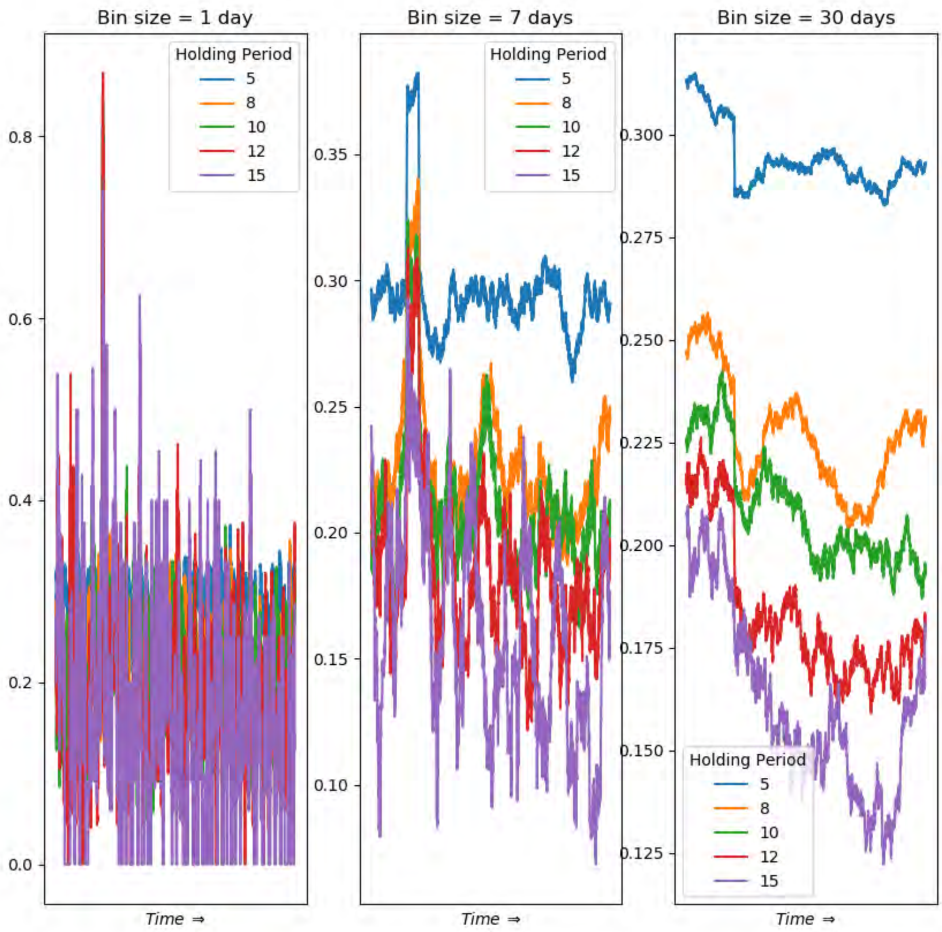

Figure A.14: Evolution of precision over time. We present the evolution of test set precision (Precision score in y and time index in x) for class -1 over time for different holding periods (figure legend) and for different rolling windows (subfigure panels) over which the precision is calculated. We don't have huge discrepancy between the temporal values, which tells us that the overall performance improvement is because of the increment of precision at each time step.

References

Amit, Y., & Geman, D. (1997). Shape quantization and recognition with randomized trees. *Neural computation*, *9*(7), 1545–1588.

Bachelier, L. (1900). Théorie de la spéculation. In *Annales scientifiques de l'école normale supérieure* (Vol. 17, pp. 21–86).

Barandiaran, I. (1998). The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell*, *20*(8), 1–22.

Bouchaud, J.-P. (2001). Power laws in economics and finance: some ideas from physics.

Bouchaud, J.-P., Farmer, J. D., & Lillo, F. (2009). How markets slowly digest changes in supply and demand. In *Handbook of financial markets: dynamics and evolution* (pp. 57–160). Elsevier.

Bouchaud, J.-P., Gefen, Y., Potters, M., & Wyart, M. (2004). Fluctuations and response in financial markets: the subtle nature of 'random'price changes. *Quantitative finance*, *4*(2), 176–190.

Bouchaud, J.-P., Kockelkoren, J., & Potters, M. (2006). Random walks, liquidity molasses and critical response in financial markets. *Quantitative finance*, *6*(02), 115–123.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Breiman, L. (2017). *Classification and regression trees.* Routledge.

Campbell, J. Y. (2014). Empirical asset pricing: Eugene fama, lars peter hansen, and robert shiller. *The Scandinavian Journal of Economics*, *116*(3), 593–634.

Campbell, J. Y., Champbell, J. J., Campbell, J. W., Lo, A. W., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets.* princeton University press.

Caparrelli, F., D'Arcangelis, A. M., & Cassuto, A. (2004). Herding in the italian stock market: a case of behavioral finance. *The Journal of Behavioral Finance*, *5*(4), 222–230.

Cuthbertson, K., Hayes, S., & Nitzsche, D. (1997). The behaviour of uk stock prices and returns: is the market efficient? *The Economic Journal*, *107*(443), 986–1008.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *the Journal of Finance*, *45*(2), 379–395.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, *38*(1), 34–105.

Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *The journal of finance*, *65*(5), 1915–1947.

Filimonov, V., Bicchetti, D., Maystre, N., & Sornette, D. (2014). Quantification of the high level of endogeneity and of structural regime shifts in commodity markets. *Journal of international Money and finance*, *42*, 174–192.

Filimonov, V., & Sornette, D. (2012). Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. *Physical Review E*, *85*(5), 056108.

Garcimartín, A., Pastor, J., Ferrer, L., Ramos, J., Martín-Gómez, C., & Zuriguel, I. (2015). Flow and clogging of a sheep herd passing through a bottleneck. *Physical Review E*, *91*(2), 022808.

Hasbrouck, J. (1988). Trades, quotes, inventories, and information. *Journal of Financial Economics*, *22*(2), 229–252.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, *58*(1), 83–90.

Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (Vol. 1, pp. 278–282).

Jensen, M. C. (1954). The performance of mutual funds in the period 1954-64. *Journal of Finance (May 1968)*.

Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of financial economics*, *6*(2/3), 95–101.

Koutmos, G., & Saidi, R. (2001). Positive feedback trading in emerging capital markets. *Applied Financial Economics*, *11*(3), 291–297.

Lee, C. M., & Swaminathan, B. (2000). Price momentum and trading volume. *the Journal of Finance*, *55*(5), 2017–2069.

Lillo, F., & Farmer, J. D. (2004). The long memory of the efficient market. *Studies in nonlinear dynamics & econometrics*, *8*(3).

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of economic perspectives*, *17*(1), 59–82.

Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, *25*(2), 383–417.

Mandelbrot, B. B. (1997). The variation of certain speculative prices. In *Fractals and scaling in finance* (pp. 371–418). Springer.

Moody, J., Wu, L., Liao, Y., & Saffell, M. (1998). Performance functions and reinforcement learning for trading systems and portfolios. *Journal of Forecasting*, *17*(5-6), 441–470.

Powers, D. M. (2011). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation.

Saichev, A., Maillart, T., & Sornette, D. (2013). Hierarchy of temporal responses of multivariate self-excited epidemic processes. *The European Physical Journal B*, *86*(4), 124.

Samuelson, P. A. (2016). Proof that properly anticipated prices fluctuate randomly. In *The world scientific handbook of futures markets* (pp. 25–38).

World Scientific.

Sen, A. (1969). Quasi-transitivity, rational choice and collective decisions. *The Review of Economic Studies*, *36*(3), 381–393.

Sewell, M. (2012). The efficient market hypothesis: Empirical evidence. *International Journal of Statistics and Probability*, *1*(2), 164.

Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *Journal of economic perspectives*, *17*(1), 83–104.

Shiller, R. J. (2014). Speculative asset prices. *American Economic Review*, *104*(6), 1486–1517.

Sjöberg, M., Albrectsen, B., & Hjältén, J. (2000). Truncated power laws: a tool for understanding aggregation patterns in animals? *Ecology letters*, *3*(2), 90–94.

Sornette, D. (2006). Endogenous versus exogenous origins of crises. In *Extreme events in nature and society* (pp. 95–119). Springer.

Sornette, D. (2009). Dragon-kings, black swans and the prediction of crises. *International Journal of Terraspace Science and Engineering 2*(1), 1-18.

Sornette, D., Woodard, R., & Zhou, W.-X. (2009). The 2006–2008 oil bubble: Evidence of speculation, and prediction. *Physica A: Statistical Mechanics and its Applications*, *388*(8), 1571–1576.

Soros, G. (2013). Fallibility, reflexivity, and the human uncertainty principle. *Journal of Economic Methodology*, *20*(4), 309–329.

Soros, G. (2015). *The alchemy of finance*. John Wiley & Sons.

Thaler, R. H., & Ganser, L. (2015). *Misbehaving: The making of behavioral economics*. WW Norton New York.

Wikipedia contributors. (2018). *Jensen's alpha — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Jensen%27s_alpha&oldid=870607127. ([Online; accessed 21-August-2019])

Wikipedia contributors. (2019a). *Algorithmic trading — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Algorithmic_trading&oldid=910630682. ([Online; accessed 20-August-2019])

Wikipedia contributors. (2019b). *Efficient-market hypothesis — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Efficient-market_hypothesis&oldid=910425958. ([Online; accessed 20-August-2019])

Wikipedia contributors. (2019c). *Feedback — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Feedback&oldid=904025503. ([Online; accessed 12-August-2019])

Wikipedia contributors. (2019d). *Foreign exchange market — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Foreign_exchange_market&oldid=911001563. ([Online; accessed 20-August-2019])

Wikipedia contributors. (2019e). *Normal distribution — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Normal_distribution&oldid=908137280. ([Online; accessed 21-August-2019])

Wikipedia contributors. (2019f). *Point process — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Point_process&oldid=883904758. ([Online; accessed 22-August-2019])

Wikipedia contributors. (2019g). *Reflexivity (social theory) — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Reflexivity_(social_theory)&oldid=906528273. ([Online; accessed 20-August-2019])

Wikipedia contributors. (2019h). *Statistical hypothesis testing — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Statistical_hypothesis_testing&oldid=909688699. ([Online; accessed 20-August-2019])

Wikipedia contributors. (2019i). *Trade (financial instrument) — Wikipedia, the free encyclopedia.* https://en.wikipedia.org/w/index.php?title=Trade_(financial_instrument)&oldid=887670581. ([Online; accessed 20-August-2019])

www.histdata.com. (2018). Retrieved from https://www.histdata.com/download-free-forex-data/

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Modelling Forex Market Reflexivity using Self-Exciting Point Process and Ensemble Learning

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	First name(s):
Shintre	Devendra

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date

Zurich,30-08-2019

Signature(s)

DSShintre

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.