Annotated Bibliography
Utilizing Machine Learning for High Frequency Algorithmic Trading

Masixole Boya (Student number: 1869204)

March 24, 2024

[Peng and Lee, 2021] Peng, Y.L. and Lee, W.P., 2021. Data selection to avoid overfitting for foreign exchange intraday trading with machine learning. Applied Soft Computing, 108, p.107461.

Aim

To propose the 'path loss', a new metric that aims to resolve the issues of overfitting in algorithmic trading in the foreign exchange (FOREX) market, and the issue of deciding which currency pair to pick and with what frequency.

Style/Type:

Journal article

Cross references

The aspect of algorithmic trading focused on here is the same as in **Ilić and Brtka, 2018**, albeit in different contexts (foreign exchange market in Peng and Lee, and broader trading strategies in Ilić and Brtka). They both emphasize the importance of developing effective strategies while mitigating risks such as overfitting and randomness. Just as **Ayitey Junior *et al.*, 2023**, both papers utilize machine learning techniques for trading purposes. While **Peng and Lee** propose a metric to address overfitting, **Ayitey Junior *et al.*** conduct a systematic literature review on machine learning models for forex market forecasting, including the methods proposed by **Peng and Lee** in their analysis.

Summary

The paper proposes a novel metric called "path loss" to tackle overfitting issues in algorithmic trading within the foreign exchange (FOREX) market, addressing the challenge of selecting currency pairs and frequencies. It introduces key performance metrics including accuracy, in-sample return, and path loss to evaluate trading models' effectiveness while minimizing overfitting. The selection process for currency pairs and frequencies aims at maximizing accuracy and in-sample return while minimizing path loss. Machine learning models are trained and tested using minute-level data from Dukascopy, incorporating features like spreads, price changes, and volume differentials.

Four popular machine learning models are evaluated, with hyperparameter optimization through iterative training and testing using the moving window technique. Strategies treating FX trading as a Markov Decision Process outperform others. The paper also compares various metrics for selecting currency pairs and frequencies, highlighting the superiority of the path loss metric. Experimental results demonstrate the effectiveness of the proposed approach, showing superior accuracy, in-sample return, and F1-score. Additionally, the paper discusses the profitability ratio, considering factors like slippage, commissions, and spreads. Overall, utilizing the path loss metric improves trading performance and aids in selecting exchange pairs and frequencies effectively, mitigating the challenges of overfitting.

[Ilić and Brtka, 2018] Ilić, V. and Brtka, V., 2018. Evaluation of algorithmic strategies for trading on foreign exchange market.

Aim

The aim is to firstly take a brief look at the Foreign Exchange market and secondly, to look at algorithmic strategies for automated trading while advising on strategy development, and performance evaluation.

Style/Type

Journal article

Cross references

Both this paper and **Peng and Lee, 2021** discuss algorithmic trading strategies, with **Peng and Lee** proposing a metric to address overfitting specifically in the foreign exchange market, which can be an addition to the strategies discussed in here.
As with **Ayitey Junior et al., 2023**, both papers contribute to the broader understanding of algorithmic trading strategies. **Ilić and Brtka**'s discussion of various strategies provides context for the machine learning models examined in Ayitey Junior *et al.*'s systematic review.

Summary

The paper extensively discusses algorithmic strategies tailored for trading in the foreign exchange market (Forex or FX), stressing the vital role of these strategies in mitigating randomness and emotional influences. These strategies encompass a series of instructions governing position openings and closures grounded in technical analysis principles. They typically entail entry and exit rules, risk management protocols, and position sizing strategies. Simulations on historical data serve to furnish preliminary insights into strategy efficacy, aiding in assessing expected performance and risk levels.
Subsequent optimizations seek to fine-tune parameter values, optimizing strategy performance based on historical data by running multiple iterations with varied parameter sets to maximize profits, minimize losses, and reduce risks. The paper also highlights the utilization of strategy testers, such as MetaTrader's module, for optimization and financial data analysis. Furthermore, it suggests leveraging neural networks, data mining, and evolutionary techniques for identifying optimal trading parameters.

Additionally, the paper provides an overview of major currencies, currency pairs, and two primary approaches to market analysis: fundamental and technical. It underscores the structured formulation and evaluation of trading strategies, emphasizing a systematic approach for development and assessment. This comprehensive synthesis encompasses strategy formulation, simulation, optimization, and market analysis, offering a robust framework for algorithmic trading in the Forex market.

[Moosa, 2015] Moosa, I., 2015. The regulation of high-frequency trading: A pragmatic view. Journal of Banking Regulation, 16, pp.72-88.

Aim

To explore arguments that support and those against the regulation of High Frequency Trading (HFT), while categorizing HFT strategies.

Style/Type

Journal article

Cross references

Although from different perspectives, **Ayitey Junior *et al.*, 2023**, as in this paper, address the topic of high-frequency trading. In particular, **Moosa** explores arguments for and against regulation, while Ayitey Junior et al. conduct a systematic review of machine learning models for forex market forecasting, which seemingly includes strategies relevant to Moosa's discussion.
Vo and Yost-Bremm, 2018 act on the discussion of Moosa - the impact of high-frequency trading on market dynamics. Moosa discusses the regulatory aspects, while Vo and Yost-Bremm propose a high-frequency trading strategy for cryptocurrency, that contributes to the controversy introduced by Moosa subject.

Summary

The paper explores arguments both for and against the regulation of High-Frequency Trading (HFT) while categorizing HFT strategies. It addresses concerns regarding market stability, price discovery, volatility, liquidity, and fairness, challenging assertions that HFT worsens volatility or liquidity. Instead, it suggests that HFT can positively impact market linkages, liquidity provision, and price efficiency. While acknowledging risks such as abusive practices and technological malfunctions, it emphasizes the benefits of HFT to market functioning. The paper advocates for targeted regulatory measures to address specific risks, preserving innovation and competition.

HFT is characterized by the use of advanced technology and algorithms for strategies like market making and arbitrage, involving high portfolio turnover and frequent order cancellations. Despite some market participants viewing HFT as monopolizing superior

technology, it's argued that HFT contributes to market efficiency by quickly processing new information and aiding in price discovery. While intra-day price moves may not always correlate with fundamentals, HFT's role in market volatility, as seen in events like the 2010 flash crash, is noted. By presenting a nuanced understanding of HFT's impact and advocating for targeted regulation, the paper aims to balance concerns regarding market abuse while fostering innovation and competition.

[Vo and Yost-Bremm, 2018] Vo, A. and Yost-Bremm, C., 2018. A high-frequency algorithmic trading strategy for cryptocurrency. Journal of Computer Information Systems.

Aim

To develop a HFT strategy for Bitcoin by utilizing machine learning algorithms and financial indicators derived from minute-level price data. The objectives are to pre-process data from cryptocurrency exchanges, transform the data into financial indicators, and train a trading model using machine learning techniques.

Style/Type

Journal article

Cross references

The work of **Moosa, 2015** extensively examines regulatory frameworks and their impact on high-frequency trading practices, evaluating arguments for and against regulation, and discussing potential implications for market stability and fairness. On the other hand, "**Vo and Yost-Bremm** go on to make this practical by developing a high-frequency trading strategy specifically tailored for cryptocurrency markets. Their research explores the same, although not all, technical aspects such as algorithm design and order execution in the context of cryptocurrency trading, shedding light on unique challenges and opportunities in this rapidly evolving landscape. Each study contributes distinct insights into the complex dynamics of high-frequency trading, enriching our understanding of its implications and strategies.

Summary

The paper aims to develop a high-frequency trading (HFT) strategy tailored for Bitcoin, leveraging machine learning algorithms and financial indicators derived from minute-level price data. It acknowledges existing literature on cryptocurrency markets, machine learning in HFT, and investor biases, but presents a novel approach by combining machine learning with financial indicators for Bitcoin trading. Utilizing the Design Science Research (DSR) methodology, it pre-processes data from exchanges and applies Random Forest (RF) algorithms to train the trading model, comparing its performance with a Deep Learning (DL) model. The strategy incorporates technical indicators like Relative Strength Index (RSI), Stochastic Oscillator, Williams $\%R$, Moving Average Convergence Divergence (MACD), and On-Balance Volume (OBV). The paper evaluates the strategy against foreign currency exchanges, particularly the Japanese Yen (JPY) to U.S. Dollar (USD) exchange rate.

Additionally, the paper provides descriptive and prescriptive knowledge to Design Science Research (DSR), focusing on creating and evaluating an HFT strategy for Bitcoin using the random forest (RF) machine learning technique. It emphasizes the scarcity of studies on HFT trading algorithms for Bitcoin, aiming to establish a benchmark strategy for the research community. HFT involves rapid asset buying and selling with short holding periods, exploiting order imbalances and news trading. The developed trading strategy demonstrates robust performance even with default settings, outperforming a long-only strategy. The paper encourages fintech research engagement and highlights the role of behavioral finance theories in research ideation and execution.

[Marudulu, 2020] Marudulu, L., 2020. Portfolio optimisation approaches towards investment in the forex market (Doctoral dissertation, North-West University (South Africa)).

Aim

To propose and evaluate a portfolio optimization and risk management approach for the Forex market. The primary objectives of the paper are to go through the following steps to achieve the aim: Fitting Forecasting Models, Constructing Efficient Frontiers, Risk Reduction, Comparing Optimization Models, Backtesting Analysis.

Style/Type

Doctoral (Academic) dissertation

Cross References

This paper goes further than **Ayitey Junior *et al.*, 2023**, although they share an objective - researching optimization approaches for trading in the forex market. **Marudulu** focuses on portfolio optimization and risk management, which is only an aspect considered in the machine learning models reviewed by Ayitey Junior et al.
While Marudulu's focus lies mainly on portfolio optimization, **Peng and Lee, 2021** propose a metric to address overfitting in algorithmic trading, and this complements Marudulu's optimization approach where the 'effectiveness' of trading models is the main goal.

Summary

The paper aims to propose and evaluate a portfolio optimization and risk management approach specifically tailored for the Forex market. It delineates steps including Fitting Forecasting Models, Constructing Efficient Frontiers, Risk Reduction, Comparing Optimization Models, and Backtesting Analysis to achieve this aim. The study focuses on portfolio optimization models customized for currency trading, exploring Markowitz (M-V), semi-mean-absolute deviation (SMAD), and conditional value-at-risk (CVaR) models. Assumptions include trader risk aversion, standard Forex trading account operation, and 1:1 leverage usage, with USD serving as the quote currency for simplified analysis. Methodology involves defining terms like capital amount invested (C), proportion of capital invested in each security (w_i), and unit price of securities (P_{it}), with key concepts including returns (R_{it}), mean (μ_i), variance (σ_i^2), and covariance (σ_{ij}). The study introduces the performance function $F_\alpha(w, \gamma)$ and employs convex programming formulations for optimization.

Additionally, it discusses mean and variance of security returns, standard deviation, and fundamental analysis in currency trading. The paper presents a loss function associated with the decision vector and permissible portfolios, along with a lemma on convexity and degree of being differentiable. Results and discussion encompass trading strategy assumptions, efficient portfolio evaluation, and profitability assessment. Comparison of deterministic FS and SES models for forecasting indicates SES model superiority. Optimal capital allocation for different return rates using M-V, SMAD,

and CVaR approaches is presented. Overall, the paper offers a comprehensive examination of portfolio optimization strategies in the Forex market, demonstrating forecast-based optimization models' efficacy, particularly the SMAD portfolio, in risk mitigation and profitability enhancement.

[Ayitey Junior *et al.*, 2023] Ayitey Junior, M., Appiahene, P., Appiah, O. and Bombie, C.N., 2023. Forex market forecasting using machine learning: Systematic Literature Review and meta-analysis. Journal of Big Data, 10(1), p.9.

Aim

The aim is to, through a systematic literature review (SLR), understand and summarize the current algorithms and models used for forex market forecasting with machine learning techniques.

Type/Style

Journal article

Cross References

The theme is this paper is the same as given by **Marudulu, 2020** but different aspects of forex trading are explored and thus a more complete picture is painted. While **Ayitey Junior *et al.*** conduct a systematic review of machine learning models for forex market forecasting, **Marudulu** focuses on portfolio optimization and risk management. The findings of Ayitey Junior et al. could inform the optimization strategies discussed by Marudulu.

This paper, together with **Peng and Lee, 2021** examine how machine learning models repsond to some algorithmic trading strategies. **Peng and Lee** end by proposing a metric to address overfitting, which they present as a problem experienced by some of the machine learning models examined by **Ayitey Junior et al**. in their systematic review. Overall, both papers aim to enhance trading effectiveness by addressing specific challenges in algorithmic trading.

Summary

The paper titled "Forex market forecasting using machine learning: Systematic Literature Review and meta-analysis" aims to offer an overview of machine learning models applied in the FX market through a Systematic Literature Review (SLR) of 60 papers published from 2010 to 2021. The evaluation techniques encompass validation techniques and assessment measures for prediction models. Employing a snowballing strategy, the authors identify additional sources by scrutinizing reference lists and citations of initially selected papers. The final list of articles undergoes multiple screening stages, including filtering based on inclusion and exclusion criteria, quality assessment, and data extraction.

The reviewed papers commonly utilize assessment metrics such as MAE, RMSE, MAPE, and MSE, with EURUSD being the most traded currency pair. LSTM and Artificial Neural Network emerge as the most frequently used machine learning algorithms for FX market prediction. However, the findings suggest unresolved concerns and difficulties in currency prediction using machine learning, indicating room for further development and advanced strategies.

Through the SLR, the research delineates a diverse landscape of machine learning methodologies employed in forecasting, with LSTM and ANN as prevalent algorithms. Evaluation techniques predominantly rely on percentage split validation, with metrics like MAE and RMSE commonly used to assess model performance. This synthesis of existing literature provides insights into current trends and challenges in machine learning-based FX market forecasting, paving the way for future advancements in the field.