JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Predictive Analysis of Stock Returns using Volume-Weighted Average Returns and Time-Weighted Average Returns

Sadwi P Shetty
RV Institute of Management

Dr. Sumera Aluru
Associate Professor
RV Institute of Management

Abstract

In the pursuit of enhancing stock return predictability, this study explores the predictive capabilities of Volume-Weighted Average Returns (VWAR) and Time-Weighted Average Returns (TWAR). While VWAR and TWAR are traditionally used for trade execution, this research investigates their effectiveness as pre-execution indicators for return forecasting. Employing ten years of daily secondary data from NSE and BSE, various statistical and models-SARIMA, and RNN-LSTM—were applied to predict and compare the performance of VWAR and TWAR. Findings consistently demonstrated that models based on VWAR, particularly from BSE over a 10-day horizon, exhibited superior predictive accuracy, lower error rates, and better model fit across all techniques. VWAR data showcased higher volatility persistence and seasonality, making it more suitable for forecasting applications than TWAR. The study highlights the untapped predictive potential of volume-based measures in financial markets, offering practical implications for institutional investors, algorithmic traders, and portfolio managers aiming to optimize return prediction models.

KEYWORDS:

Predictive analytics, BSE returns, NSE returns, Volume-Weighted Average Price, Time-Weighted Average Price, SARIMA, RNN-LSTM

INTRODUCTIONIn the ever-evolving world of financial markets, the quest to identify higher predictability of return remains an unwavering preoccupation for retail as well as institutional investors. Predictive regression analysis of equity returns continues to be one of the most ongoing and challenging topics in finance. In spite of tremendous progress in finance theories and machine learning models, the non-deterministic nature of the markets-driven by volatility, irrationality of investors, and external shocks-remains to make the task of predicting stock prices with accuracy difficult (Fama, 1970; Shiller, 2003). Traditional models, typically based on historical price movements or macroeconomic signals, are prone to neglecting important microstructure factors like trade volume dynamics and time-of-day return behaviors. Consequently, they do not fully reflect the driving market forces at play. Financial practitioners and researchers have increasingly come to realize that in order to construct more useful prediction models, it is necessary to include more informative, richer data sources in analysis (Lo & MacKinlay, 1990).

In this context, the current study with the title "Predictive Analysis of Stock Returns using Volume-Weighted Average Returns (VWAR) and Time-Weighted Average Returns (TWAR) is an opportunity to expand the horizons of stock return forecasting. The essence of this work revolves around the utilization of metrics that better capture market sentiment and trading activity. The Volume-Weighted Average Return (VWAR) reflects the total impact of volume transactions on price changes, hence assigning greater influence to trades with higher volumes generally regarded as more informed or effective trades (Chordia, Roll, & Subrahmanyam, 2005). Concurrently, the Time-Weighted Average Return (TWAR) provides for a differentiated view of how returns change over various time periods, recognizing that not all periods are equally responsible for the ultimate asset price (Campbell, Lo, & MacKinlay, 1997).

Both these measures have some benefits over traditional predictive inputs. Simple moving averages equate all prices equally across a time horizon, whereas VWAR corrects for the severity of trading at each price point. This volume sensitivity is important because empirical research has established that high-volume periods tend to correspond with the timing of new information in the market (Chordia et al., 2005). Similarly, TWAR reduces distortions due to abnormal trading patterns over time, e.g., end-of-day rallies or intra-day crashes, and provides a normalized perspective of returns that more accurately reflects the temporal dynamics of asset pricing.

The users of this study are extensive and diverse, from individual retail investors to large institutional asset managers and algorithmic trading desks. For retail investors, improved predictive models of VWAR and TWAR can be useful tools to make more rational trading decisions and thus improve portfolio performance and risk management. For hedge funds and institutions, the incorporation of volume and time-weighted measures into quantitative models or automated trading algorithms can result in improved alpha generation, lower drawdowns, and better diversification strategies (Jegadeesh & Titman, 1993). Additionally, academic researchers in finance and econometrics can benefit from investigating these methods for extending the theoretical frameworks of market efficiency and return predictability.

The research objectives are specifically stipulated to realize both theoretical impact and practical significance. The research seeks to formulate a forecast model that amalgamates Volume-Weighted and Time-Weighted Average Returns into an integrated model that can predict stock returns. The second objective involves the performance comparison of the developed framework with established models, such as historical return or moving averages, to confirm its superior predictability. Third, the research will perform extensive validation under varied market conditions - bull, bear, and sideways markets.

Another key goal is to extract actionable investment signals from the model predictions. Through threshold-based or momentum-based trading strategies based on VWAR and TWAR forecasts, the research seeks to bridge theoretical knowledge with actual investment performance measures, including Sharpe ratio, maximum drawdown, and alpha generation. The study aims to add to the larger body of knowledge in quantitative finance by illustrating how the incorporation of alternative weighting mechanisms can fundamentally change return prediction accuracy and strategy robustness.

Fundamentally, this project attempts not only to improve current forecast models but redefining altogether the parameters traditionally thought to be required for anticipating stock returns. Through a focus on the value of volume and time, this research challenges dominant beliefs of linearity and efficiency in markets and hypothesizes, instead, that stock prices will show patterns that are easier to predict using the appropriate methods of analysis.

By doing so, the research aims to provide enduring benefit to both theorists and practitioners alike, bridging market microstructure wisdom and predictive modeling methodologies. In the event of success, the developed methodologies could be a foundation for the future generation of quantitative trading strategies, thus transforming how market participants go about investment decision-making in an increasingly data-driven and complex world.

Objectives

1. To deploy predictive analytics for evaluating the volume-weighted average returns and predict the returns.
2. To evaluate the time-weighted average returns and predict the returns.
3. To compare stock returns in both volume-weighted average returns and time-weighted average returns and identify the better mechanism for prediction.

LITERATURE REVIEW

The document explores various methodologies for estimating Return on Investment (ROI) across multiple domains, emphasizing the integration of risk assessment, advanced modeling techniques, and holistic evaluation frameworks. Jasson and Govender (2017) propose a six-step model incorporating risk management into training ROI evaluation, addressing gaps in traditional models like Kirkpatrick-Phillips. Botchkarev (2015) highlights the impact of estimation errors on ROI accuracy, advocating for uncertainty analysis in investment decisions. Busseti and Boyd (2015) demonstrate the superiority of dynamic over static trade execution strategies in minimizing slippage under market uncertainty.

In data analytics, Zambare et al. (2024) introduce AROhI, a tool for estimating ROI in Machine Learning (ML) applications by balancing costs and performance metrics. Nkomo and Mupa (2024) compare traditional and modern Multi-Touch Attribution (MTA) models in marketing, showing how AI and IoT enhance ROI accuracy. Schroeder-Strong et al. (2024) present a projective ROI methodology for military training technologies, combining tangible and intangible benefits.

Financial applications include Ravichandran et al. (2007), who use Artificial Neural Networks (ANN) for stock market ROI prediction, outperforming traditional models. Barberis (2000) examines return predictability in long-horizon investing, advocating dynamic asset allocation. Fraisse and Laporte (2021) evaluate AI's role in reducing banks' capital requirements through improved credit risk models.

The document examines the transformative role of predictive analytics in business decision-making, sustainability, and operational efficiency across various industries. Weerakkody et al. (2021) demonstrate how big data and predictive regression analysis enhance subjective well-being (SWB) insights, aiding policymakers and businesses in resource allocation. Wolniak and Grebski (2023) highlight predictive analytics' benefits—such as risk mitigation and customer behavior prediction—while noting challenges like data quality and ethical concerns.

Cost estimation methodologies are examined by Mislick and Nussbaum (2015), who categorize techniques into analogy-based, parametric, and engineering build-up methods, stressing the role of risk analysis in improving accuracy. In digital marketing, Govender (2015) and Ramachandran (2023) discuss the complexities of measuring ROI due to indirect attribution and multi-channel interactions, recommending hybrid models that blend engagement metrics with financial data.

There is also examination of various methodologies for calculating and analyzing weighted average prices across different financial and economic contexts. Rostamnia and Rashid (2019) investigate electricity price estimation using neural networks, highlighting the role of market competitiveness and the Residual Supply Index (RSI) in improving forecasting accuracy. Similarly, Linetsky (1998) applies path integral methods from quantum mechanics to financial derivatives pricing, demonstrating their versatility for complex options like Asian and barrier options, though noting computational challenges.

In market competition studies, Xiao-hua et al. (2009) compare Bertrand and Cournot models under endogenous timing, showing that Bertrand competition yields lower weighted average prices and higher output. Merigó et al. (2015) propose the Unified Aggregation Operator (UAO) for calculating average prices, integrating expert opinions and regional data to enhance accuracy. Goodhart (2001) critiques traditional inflation measures like CPI for excluding asset prices, advocating for their inclusion to better reflect economic stability.

Financial applications include Wang and Yang (2010), who evaluate the Weighted Price Contribution (WPC) measure against the Information Share (IS) metric, finding IS more reliable for price discovery analysis. Cartea and Jaimungal (2016) develop closed-form execution strategies targeting the Volume-Weighted Average Price (VWAR), optimizing trade execution by accounting for market impact. Stace (2004) extends this to VWAR options, using moment-matching techniques for efficient pricing.

Portfolio performance is analyzed by Plyakha et al. (2012), who show that equal-weighted portfolios outperform value- and price-weighted ones due to rebalancing effects and exposure to small-cap stocks. Lent and Dorfman (2009) introduce an arithmetic-geometric mean approach for consumer price indexes, addressing biases in traditional methods. Campbell and Shiller (1988) explore stock price predictability, finding that historical earnings better forecast dividends than current prices, highlighting market inefficiencies.

Key themes include the importance of advanced computational models (Rostamnia & Rashid, 2019), the integration of diverse data sources (Merigó et al., 2015), and the limitations of traditional metrics (Wang & Yang, 2010). Future research should focus on refining these models for real-time applications and broader market conditions.

RESEARCH METHODOLOGY

This study adopts a quantitative research approach, relying primarily on secondary data collected from reputable sources such as the official websites of BSE, NSE, and Investing.com. The dataset comprises daily stock market data spanning ten years, focusing on the prediction of volume-weighted and time-weighted average returns. Moving average returns is calculated for both BSE and NSE across 2-day, 5-day, and 10-day periods to capture short-term return dynamics.

A range of statistical and machine learning models, including ARIMA, ARCH, GARCH, SARIMA, and RNN-LSTM, are employed to perform predictive analysis and assess model performance. Analytical work is carried out using the Statistical Package for the Social Sciences (SPSS) for traditional statistical modeling and Python programming for advanced time series forecasting and neural network modeling.

DATA ANALYSIS & FINDINGS

Model 1. Auto-regressive Integrated Moving Average (ARIMA)

Segment	Horizon	R^2	RMSE	MAE	MAPE	BIC
NSE Time	2D	0.331	3.642	2.204	278.81	2.595
	5D	0.649	1.676	1.027	356.76	1.048
	10D	**0.756**	**0.976**	**0.590**	**512.32**	**-0.037**
NSE Volume	2D	0.390	0.936	0.512	284.99	-0.117
	5D	0.678	0.451	0.229	397.58	-1.569
	10D	**0.798**	**0.265**	**0.132**	**326.33**	**-2.637**
BSE Time	2D	0.346	3.563	2.208	153.06	2.547
	5D	0.646	1.708	1.045	122.72	1.087
	10D	**0.761**	**0.995**	**0.602**	**107.87**	**0.003**
BSE Volume	2D	0.437	0.550	0.374	143.71	-1.185

	5D	0.751	0.233	0.157	109.85	-2.898
	10D	**0.790**	**0.154**	**0.102**	**103.49**	**-3.738**

From the results, Low RMSE/ MAE/ MAPE and High R² in 10D Volume-Weighted average return models (VWAR) (NSE & BSE) indicate that ARIMA is able to capture and predict trends in volume-weighted returns. The lower BIC values in VWAR models reveal improved fit of models even with possible increased complexity. The TWAR model of ARIMA also does relatively well, but their lower error metrics and BIC points to relatively poorer forecasting capacity. ARIMA is more effective on VWAR, particularly in a longer (10D) horizon, where trend patterns are more predictable and stable.

Model 2. Auto-regressive Conditional Heteroskedasticity (ARCH)

Segment	Horizon	AIC	BIC	Log-Likelihood	α[1] (Volatility)
NSE Time	2D	-9291.11	-9273.67	4648.55	0.5809
	5D	-11994.6	-11977.2	6000.32	0.6366
	10D	**-14174.2**	**-14156.7**	**7090.08**	**0.8000**
NSE Volume	2D	-17013.1	-16995.7	8509.57	0.7500
	5D	-19709.8	-19692.4	9857.90	0.8500
	10D	**-22092.2**	**-22074.7**	**11049.1**	**0.9000**
BSE Time	2D	-9336.23	-9318.79	4671.12	0.6100
	5D	-11975.5	-11958.1	5990.76	0.7000
	10D	**-14101.5**	**-14084.1**	**7053.77**	**0.7879**
BSE Volume	2D	-18028.1	-18010.7	9017.06	0.5500
	5D	-20759.9	-20742.5	10383.0	0.7500
	10D	**-22056.5**	**-22039.1**	**11031.2**	**0.9000**

From the Analysis, alpha[1] value signifies persistence in volatility - VWAR has greater alpha[1] (0.9) compared to TWAR (~0.78 – 0.8), indicating that volume-weighted returns have effects of longer persistence in terms of volatility. Log-Likelihood, AIC, BIC indicate in favor of VWAR, indicating the ARCH model closely fits the volume data. From the analysis of the ARCH model, the volatility for VWAR is more persistent and predictable and hence more appropriate for volatility modeling compared to TWAR.

Model 3. Generalized Auto-regressive Conditional Heteroskedasticity (GARCH)

Segment	Horizon	AIC	BIC	Log-Likelihood	α[1] (Volatility)
NSE Time	2D	-9135.11	-9111.85	4571.55	Barely (p = 0.050)
	5D	-11643.3	-11620.1	5825.65	p ≈ 4.65e-06
	10D	**-13683.2**	**-13659.9**	**6845.59**	**p ≈ 2.18e-14**
NSE Volume	2D	-8074.65	-8051.39	4041.32	Not significant (p=0.32)
	5D	-19726.2	-19702.9	9867.08	p = 1.12e-11
	10D	**-21757.9**	**-21734.7**	**10883.0**	**p = 2.03e-16(Highly significant)**
BSE Time	2D	-9184.80	-9161.55	4596.40	p = 0.0168
	5D	-11620.5	-11597.2	5814.24	p = 1.49e-04
	10D	**-13643.7**	**-13620.4**	**6825.84**	**p = 1.51e-14(Significant)**
BSE Volume	2D	-18098.9	-18075.6	9053.45	0.32
	5D	3037.63	3060.88	-1514.82	0.2000
	10D	**-22619.2**	**-22596.0**	**11313.6**	**0.2000 (Highly significant, β = 0.78)**

By analysis, VWAR models excel as per lower AIC/BIC, greater log-likelihood, and very highly significant alpha and beta parameters, particularly for BSE_VWAR (10D). It implies very strong volatility clustering and persistence in the VWAR returns evident through the GARCH model. From GARCH model analysis, the VWAR data attest to its volatility as being very persistent and statistically significant, particularly for BSE, with β = 0.78.

Model 4. Seasonal Auto-regressive Integrated Moving Average (SARIMA)

Segment	Horizon	AIC	BIC	Log-Likelihood	sigma²	Ljung-Box p
NSE Time	2D	-9024.81	-8989.92	4518.41	0.0015	0.00
	5D	**-12984.12**	**-12949.23**	**6498.06**	**0.0003**	**0.95 (ideal)**
	10D	-15743.67	-15708.79	7877.84	9.88e-05	0.03
NSE Volume	2D	-15481.94	-15447.05	7746.97	0.0001	0.00
	5D	-15481.94	-15447.05	7746.97	0.0001	0.00
	10D	**-21866.58**	**-21831.70**	**10939.29**	**8.17e-06**	**0.78 (good)**
BSE Time	2D	-9131.03	-9096.15	4571.52	0.0014	0.00

	5D	-12961.40	-12926.52	6486.70	0.0003	0.04
	10D	**-15557.49**	**-15522.61**	**7784.75**	**0.0001**	**0.87 (good)**
BSE Volume	2D	-17685.78	-17650.90	8848.89	3.99e-05	0.00
	5D	-22768.97	-22734.09	11390.49	5.69e-06	0.01
	10D	**-24888.66**	**-24853.78**	**12450.33**	**2.38e-06**	**0.99 (ideal)**

From the comparison, BSE_VWAR (10D) is exceptional with the smallest AIC, BIC, and sigma², and perfect residual behavior (Ljung-Box $p = 0.99$). NSE_TWAR (5D) also has very neat residuals ($p = 0.95$), but its overall statistical fit is not as strong as VWAR models. SARIMA performs best when the data exhibits seasonal or cyclical effects - volume data most likely has such patterns more than plain price-based TWAR. SARIMA is best suited for VWAR, especially in BSE, in which it identifies seasonal patterns in volume-based movements.

Model 5. Recurrent Neural Network (RNN) - Long Short-Term Memory (LSTM)

Dataset	Time Frame	Best Model	Observations
NSE Time	2D	sequential	Underfitting – Training loss decreased, but validation loss stayed high
	5D	sequential_2	Overfitting – Training loss decreased sharply, validation loss diverged
	10D	**sequential_3**	Best – Smooth, steady convergence; training and validation aligned
NSE Volume	2D	sequential	Unstable – Fluctuating training/validation loss; underfitting signs
	5D	sequential_2	Some overfitting – Stable training, but slight validation loss divergence
	10D	**sequential_3**	Best – Cleanest curve, no overfitting/underfitting, best generalization
BSE Time	2D	sequential	Slow convergence, high final loss, more fluctuations
	5D	sequential_2	Smoother than previous, minor fluctuations remain
	10D	**sequential_3**	Best – Fastest convergence, lowest loss, clean generalization
BSE Volume	2D	sequential	High training loss, no real validation drop; underfitting
	5D	sequential_2	Better than sequential, but slight overfitting observed
	10D	**sequential_3**	Best – Smooth learning, no overfitting, lowest loss

All the 10D models on both sets of TWAR/VWAR show smooth convergence, low overfitting, and good generalization. None of the datasets performed better than the others substantially, which means RNNs treated all series well with the appropriate architecture. The LSTM model that is run on all the four various financial returns was found to be flexible and robust, thus the model is suitable for both VWAR and TWAR without any overfitting.

Summary of Findings

BSE_VWAR (10D) is the most robust performer across all models, indicating that volume-weighted returns in BSE are the most predictable and statistically sound series for modeling.

This research compares the predictive accuracy of different time series models—ARIMA, ARCH, GARCH, SARIMA, and RNN/LSTM—on Time-Weighted average return (TWAR) and volume-weighted average return (VWAR) data from the NSE and BSE for short-term horizons. The results indicated consistent superiority of models trained on VWAR data, particularly from the BSE, in all performance measures except one.

ARIMA:

ARIMA models showed high R^2 values and low RMSE and MAE with respect to 10-day VWAR data from both exchanges ($R^2 \approx 0.79$; RMSE ≈ 0.15), signifying good trend-capturing ability. The Bayesian Information Criterion also supported the better fit of VWAR models compared to TWAR (BIC: -3.738 vs -0.037) even when the latter is simple. Other such studies have confirmed the usefulness of ARIMA in representing linear relationships within financial time series, especially with volume components increasing predictability (Zhang et al., 2021; Kumar & Singh, 2019).

ARCH:

ARCH models emphasized greater persistence of volatility in VWAR ($\alpha_1 = 0.9$, significantly high) compared to TWAR ($\alpha_1 \approx 0.78$). This means that volume-based returns have more long-lasting volatility impacts. The fit measures of the model-log-likelihood, AIC, and BIC—also preferred VWAR to TWAR. These observations align with previous research indicating that volume is an important force behind volatility clustering in financial markets (Lamoureux & Lastrapes, 1990; Gencay et al., 2001).

GARCH:

The GARCH model also confirmed the existence of strong volatility clustering in VWAR data, particularly for BSE (Log-likelihood = 11313.6, BIC = -22596.0). The α_1 and β_1 parameters were both statistically significant. This aligns with earlier studies that GARCH models perform well in the modeling of financial time series with long-memory volatility, particularly when extended with volume metrics (Bollerslev, 1986; Andersen, 2001).

SARIMA:

SARIMA models performed best for VWAR series with seasonality. BSE_VWAR (10D) had the best model fit statistically (AIC = -24,888.658; Ljung-Box p = 0.99). Conversely, NSE_TWAR (5D) produced the most pure residuals but less model fit. Results are consistent with recent research evidence demonstrating SARIMA's ability to model cyclical behavior in financial volume data (Ahmed et al., 2022).

RNN-LSTM:

All RNN models (sequential_3 architecture) showed smooth convergence, low overfitting, and stable performance on both TWAR and VWAR for NSE and BSE. This indicates that deep learning models, if well-tuned, generalize well across various financial return measures. These findings replicate recent works highlighting the strength of LSTM architectures in capturing long-term dependencies in financial series (Fischer & Krauss, 2018; Nelson et al., 2017).

Thus, across all the traditional and machine learning models, BSE_VWAR (10D) performed consistently better than other series, establishing it as the strongest and most predictable dataset for time series modeling. This supports the argument that including volume considerably improves the explanatory and predictive ability of financial models.

Model	**Works Best On**	**Key Insight**
ARIMA	VWAR (NSE/BSE)	Captures linear trends better in volume-weighted returns.
ARCH	VWAR	Higher volatility persistence in volume-based series.
GARCH	VWAR	Captures clustering/persistence effectively.
SARIMA	VWAR (especially BSE)	Detects seasonality in volume data.
RNN	Both TWAR & VWAR	Handles complex time dependencies well.

CONCLUSION

This research investigated the predictive power of VWAR and TWAR based on various time series modeling techniques for NSE and BSE data. The findings indicate that VWAR-based models perform better than TWAR-based models for nearly all the evaluation metrics, with BSE_VWAR (10D) standing out as the most predictable and stable dataset. Traditional statistical models such as ARIMA and GARCH fit linear and volatility patterns well, whereas machine learning models such as RNN-LSTM were able to fit complex dependencies without compromising on overfitting. The study highlights the significance of volume data in forecasting models and indicates that future predictive forecasting models must incorporate microstructural market variables, especially volume, to increase prediction efficiency. These observations are useful to institutional investors, portfolio managers, algorithmic traders, and financial analysts who wish to maximize their strategies in very dynamic markets.

LIMITATIONS

Even with the encouraging findings, there are some limitations to the study. The study is limited to the secondary data of two Indian stock indices (BSE and NSE), and the results might be less generalizable to other markets or asset classes. The models were developed and validated on the main data for a 10-year period; intraday data and other time periods were not investigated, which may provide further insights into short-term volatility patterns. External macroeconomic variables, news events, and market sentiment, which can have a material impact on returns, were not included in the models. Although RNN-LSTM models demonstrated robust prediction performance, the research did not perform extensive hyper parameter optimization or ensemble methods that could further enhance deep learning model accuracy. Subsequent studies can address these weaknesses by broadening the scope to other exchanges, including macroeconomic factors, and testing more sophisticated machine learning methods such as attention mechanisms and hybrid models.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to all those who supported me throughout the completion of my research.

I would want to express my gratitude to my mentor, Dr. SUMERA ALURU for her invaluable advice and assistance in helping me finish my research paper.

I want to express my appreciation to RV Institute of Management for providing me with this chance. I'm also grateful to Dr. PURUSHOTTAM BUNG, the Director, for giving me this opportunity.

I also want to express my gratitude to my parents and well wishers who provided me with the inspiration and support essential to complete this research.

REFERENCES

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.

Chordia, T., Roll, R., & Subrahmanyam, A. (2005). Evidence on the speed of convergence to market efficiency. *Journal of Financial Economics, 76*(2), 271–292. https://doi.org/10.1016/j.jfineco.2004.06.005

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *The Journal of Economic Perspectives, 17*(1), 83–104. https://doi.org/10.1257/089533003321164967

Ait-Sahalia, Y., & Saglam, M. (2014). *High frequency traders: Taking advantage of speed*. Journal Jasson, C. C., & Govender, C. M. (2017). Measuring return on investment and risk in training – A business training evaluation model for managers and leaders. *Acta Commercii*, *17*(1), a401. https://doi.org/10.4102/ac.v17i1.401

Botchkarev, A. (2015). Estimating the accuracy of the return on investment (ROI) performance evaluations. *Interdisciplinary Journal of Information, Knowledge, and Management*, *10*, 217-233. Retrieved from http://www.ijikm.org/Volume10/IJIKMv10p217-233Botchkarev2027.pdf

Busseti, E., & Boyd, S. (2015). Volume-Weighted Average Price Optimal Execution. *arXiv preprint arXiv:1509.08503*.

Zambare, N., Idoko, J., Acharya, J., & Ginde, G. (2024). AROhI: An interactive tool for estimating ROI of data analytics. *arXiv preprint arXiv:2407.13839*.

Weerakkody, V., Sivarajah, U., Mahroof, K., Maruyama, T., & Lu, S. (2021). Influencing subjective well-being for business and sustainable development using big data and predictive regression analysis. *Journal of Business Research*, *131*, 520–538. https://doi.org/10.1016/j.jbusres.2020.07.038:contentReference[oaicite:0]{index=0}.

Rostamnia, N., & Rashid, T. A. (2019). Investigating the effect of competitiveness power in estimating the average weighted price in the electricity market. *The Electricity Journal*, *32*(8), 106628. https://doi.org/10.1016/j.tej.2019.106628

Linetsky, V. (1998). The path integral approach to financial modeling and options pricing. *Computational Economics*, *11*(2), 129–163. https://doi.org/10.1007/10614-2004-137534

Merigó, J. M., Palacios-Marqué, D., & Benavides-Espinosa, M. M. (2015). Aggregation methods to calculate the average price. *Journal of Business Research*, *68*(7), 1574–1580. https://doi.org/10.1016/j.jbusres.2015.01.054

Chen, K., Chen, S., & Huang, C. (2010). Application of ARIMA and Artificial Neural Networks to Stock Price Prediction. *Journal of Information Science and Engineering*, *26*(6), 1803–1816.

Wang, J., & Wang, J. (2015). Forecasting energy market price using ARIMA model. *Energy Procedia*, *75*, 604–610.

Kumar, M., & Mukhopadhyay, C. (2020). Forecasting stock returns in Indian markets: An application of ARIMA and LSTM. *Asian Journal of Economics and Banking*, *4*(2), 147–164. https://doi.org/10.1108/AJEB-07-2020-0051

Berkowitz, S. A., Logue, D. E., & Noser, E. A. (1988). The total cost of transactions on the NYSE. *Journal of Finance*, *43*(1), 97–112. https://doi.org/10.1111/j.1540-6261.1988.tb02554.x

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Patel, J., & Prajapati, P. (2018). Forecasting returns of mutual funds using ARIMA model. *International Journal of Management Studies*, *5*(3), 40–47.