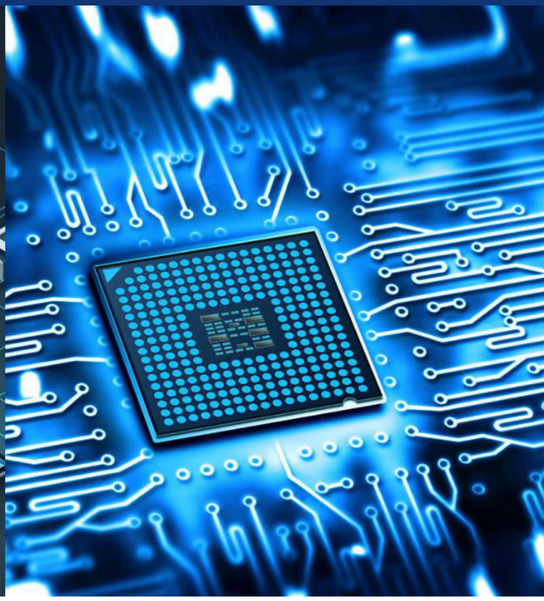
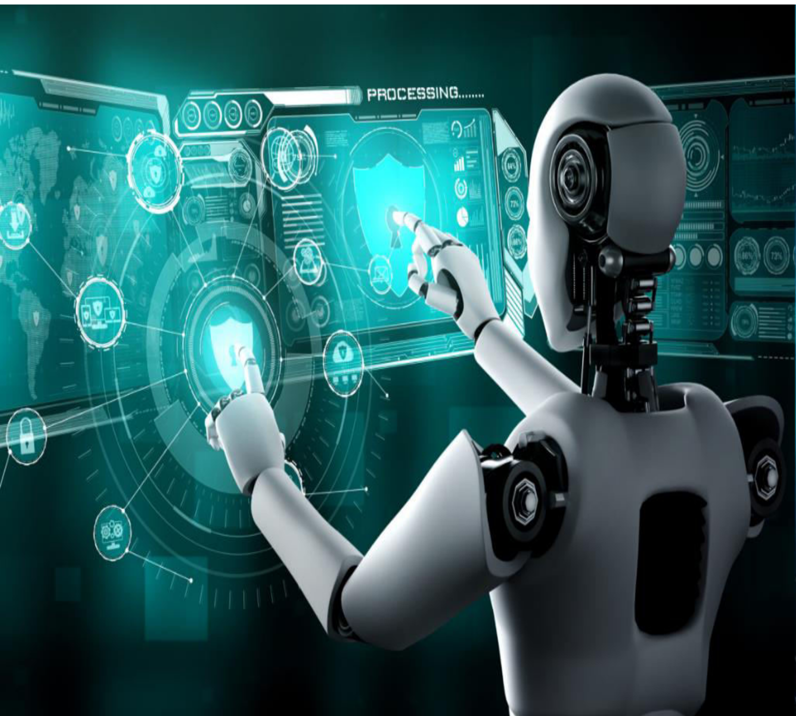
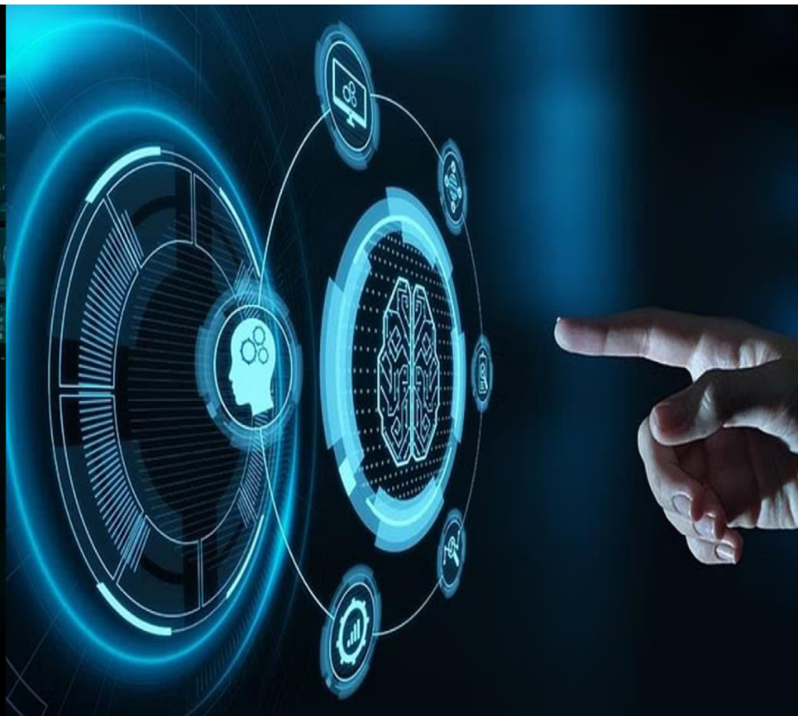
ISSN(O): 2320-9801
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 2, February 2025

© 2025 IJIRCCE | Volume 13, Issue 2, February 2025| DOI: 10.15680/IJIRCCE.2025.1302017

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Machine Learning-Driven Portfolio Optimization in Long/Short Equity Strategies: A Sector-Based Approach with Macroeconomic Integration and Stress Testing

Rugved Shitole[1], Sunny Patil[2], Mrs. Varsha Pandagre[3], Mrs. Deepali Hajare[4]

Student, Department of Artificial Intelligence and Data Science, D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune, India[1,2]

Assistant Professor, Dr. D. Y. Patil International University, Akurdi, Pune, India [3,4]

ABSTRACT: This paper describes a hedge fund styled portfolio optimization through the lens of a multi-sector long/short equity trading strategy driven by machine learning. This strategy methodically accounts for growth sectors such as Telecom, Tech, Shipping and Energy whilst using defensive sectors like Utilities, Healthcare and bond-based ETFs to hedge. Using historical financial data and technical indicators to create sector-specific machine learning models—Support Vector Machines (SVM), Random Forest (RF) and Deep Neural Networks (DNN)—this methodology determines long and short positions at the sector level, with evaluations at both the sector level and portfolio level over a specified holding period.

Dynamic threshold based on key American economic indicators (CPI, Federal Funds Rate and GDP growth) executes the trades which allows for quick adaptation of an algorithm to changing conditions. The Sharpe ratio, alpha, volatility, max drawdown and risk-reward ratio performance metrics measure this strategy against benchmark indices such as the S&P 500. The strategy is also stress tested to help gauge how the strategy would perform during recessionary periods and under different economic scenarios. The article highlights how machine learning can help achieve risk-adjusted returns in an evolving capital market environment by improving long/short equity strategies at the sector level within a transaction-cost-aware framework.

KEYWORDS: Portfolio Optimization; Hedge Fund Strategy; Long/Short Equity; Machine Learning; Support Vector Machines (SVM); Random Forest (RF); Deep Neural Networks (DNN)

I. INTRODUCTION

As traditional hedge fund strategies have always used advanced techniques for maximizing the alpha of the portfolio, there are methods that combine growth and defensive exposures in a way that maximizes both return and risk. Of these, none are more precious than long/short equity strategies due to their versatility of going long on securities anticipated to appreciate and short on securities predicted to depreciate. The aim is to maintain an equal combination of these positions in order to profit from expected price movements and lower total market exposure. However, long/short approaches based primarily on rules take inspiration from the fundamentals of human judgement and hence lack scalability and consistency across different market regimes.

As financial markets grow in complexity and become more sensitive to macroeconomic signals, long/short equity strategies must evolve accordingly. Static thresholds and limited data sources in conventional approaches present significant challenges for achieving effective portfolio optimization, as they may fail to capture the dynamic conditions needed for generating alpha in volatile markets. Machine learning-driven models offer an innovative solution by uncovering hidden patterns in data, enabling more nuanced, sector-specific long and short positions to improve portfolio performance.

This study introduces a long-short equity portfolio optimization strategy driven by machine learning techniques that predicts returns at the sector level and updates allocations over time through macroeconomic signals. In addition, the

© 2025 IJIRCCE | Volume 13, Issue 2, February 2025| DOI: 10.15680/IJIRCCE.2025.1302017

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

long and short decision thresholds are variable, adjusted in real time depending on inflation, interest rates, and GDP growth. The model adjusts these boundaries by utilizing economic fundamentals, making it flexible to the market environment and allowing for more accurate and robust programming of the portfolio.

This boosted approach includes measurements in High Watermark Return and mean variance such as Sharpe Ratio, alpha, volatility and maximum drawdown. In addition, putting the model through a recession stress test by modifying returns at the sector level confirms it continues to hold under recessionary conditions. In this way, we present a signal-driven long / short equity trading strategy approach that adds macroeconomic information-focused sector level and optimizing sector exposures while as well, owning predictions powered by machine learning. This is a very nimble and economically-tuned method of long/short equity trading — providing economic data at the very heart of portfolio decision making.

II. RELATED WORK

The long/short equity strategy, the most popular hedge fund strategy as of today, is widely considered to be pioneered by Alfred Winslow Jones in the 1940s — "the father of the hedge fund." The paradigm was Jones', a former Fortune magazine beat writer turned sociologist, who built leveraged long positions with a short program to limit market exposure. By being long on undervalued stocks and short on overvalued stocks, he tried to achieve the double objective of lowering market risk in his portfolio and profiting from stock-specific moves. This caught Jones to hedge declines in emerging stocks with rises among the growing, therefore increasing risk-adjusted total returns in a well-balanced offer.

In the subsequent years throughout the 1980s and 1990s, this approach propagated itself further due to rising hedge funds with newly accessible environmental their products as well as data. In the years that followed, quantitative finance began to shape the strategy with initial quant adopters modelling long and short positions using past earnings, valuation, momentum etc as inputs. This was an equity release theme that attracted investors, for the first time in a while, who sought limited market exposure whilst the realized returns were still lofty. For example, given the rudimentary investment palette of institutional investors long/short funds became a popular option -- combining both bullish and defensive elements in a single strategy.

Now move into the 21st century and machine learning and data science have completely redefined long/short equity strategies, enabling complex algorithms utilizing large datasets including alternative data as well as real-time economic indicators. Modern approaches frequently rely on cloud computing to apply complex machine learning models that forecast stock movements and adjust long and short positions in the sectors dynamically. Modern long/short strategies seek alpha via data-driven systems that incorporate macroeconomic fundamentals/indicators and sentiment analysis while adapting to changing market conditions.

III. METHODOLOGY

1. The long/short equity trading strategy is designed to take advantage of rise/fall in price across multiple sectors but based on machine learning model and extensive data sources. This approach consists of steps from data pre-processing to model building, predictive analytics and performance measurement.
2. Data Collection and Pre-processing:
   Historical stock prices are retrieved for instance from Yahoo Finance and Alpha Vantage and technical indicators (moving averages, RSI, Bollinger Bands) computed using TA-Lib. Data from the FRED API
3. There are a few macroeconomic indicators through which trends are very well prone. This step is typically normalization, and functions for machine learning model preparation

4. Method 2: Model Selection and Training

a) Random Forest — For modelling sector-wise returns by being able to capture non-linearity and interaction among predictors of stock level data along with the ability to generalize across the different sectors.

b) Support Vector machine (SVM): Learn historical price action and classify security as Long/Short, Helps in sector wise weightage & good at support plot.

© 2025 IJIRCCE | Volume 13, Issue 2, February 2025| DOI: 10.15680/IJIRCCE.2025.1302017

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

c) Deep Neural Network (DNN): DNN were used for learning complex, multi-dimensional structure in the stock data at sector levels (example: Tech, Telecom & Energy)

5. The models are trained using sector-divided datasets, so each model is tailored to a given sector thus providing long/short opportunities even from split industries.

6. Signal Generation and Portfolio Construction for Optimization Signals are generated by the model for each sector to determine how the portfolio is built. Long- and short-designated stocks are apportioned in the portfolio's long and short positions, respectively, while preselected hedges' securities enter the mélange to buffer downturns. We optimize the portfolio in a way that balances the sector-specific signals with hedging to provide stronger performance across multiple market environments.

7. Back-testing & Performance Evaluation: This strategy starts in Jan 2020 and its back-tests going up to the same month in 2024. Performance metrics as Total Returns, Sharpe ratio, Alpha, Risk Reward Ratio, Volatility and Maximum drawdown shall be calculated by Back-testing the algorithm so that you can evaluate how profitable a strategy will be for you and we will have some insight about the risk on your portfolio. Strategist chi on strategy v various large indices (returns are relative to the S&P 500).

8. Sector-Wise Performance Analysis Similar to average historical modeling performance, we perform quant and stats in terms of sectoral results where models are compared by cumulative as well as annualized returns for every individual sector. The return by sector pie charts show which sectors helped or hurt the performance of a strategy, and to what extent.

IV. FORMULAS USED

Portfolio work in a dynamic market-driven strategy allocations are adjusted according to optimization, to inflationary, interest-driven economic thresholds rates, and GDP growth. The strategy can be organized like this

1. Data Collection and Preprocessing:
- Sector-Wise Historical Stock Data: The sector-wise historical collecting the stocks pricing from Yahoo, Alpha Vantage and Finance individually. It powers the training the model and calculating returns
- Macroeconomic Variables: Various macroeconomic data (FRED API containing inflation rate, interest rate GDP growth, the 10-Year Treasury yield & (Fed fund) (risk-free rate).

2. Risk-Free Rate Calculation
The third element is the risk-free rate which you obtain through average of 10-year Treasury yields that are related to the comparative standard of outcome metrics.
Formula for Average Risk-Free Return:
- Average Risk-Free Return = Rationality Test/Market Cost of Equity

i.e, Average Risk-Free Return = $(1 / n) * \Sigma\, r\_i$

3. Dynamic Calculation for Threshold.
Long / short positions are dynamic in nature, and its threshold varies with macro-economic environment. The base threshold is modified according to the following:
- If inflation is higher than 2.5%, also, subtract 2 from the threshold.
- If the interest rate is higher than 3.0%, subtract 3 from the threshold.
- If GDP growth is more than 3.0%, increase threshold by 2.
- Else, leave the threshold as-is.

4. Sector Return
For each of the individual sector return in which different machine learning models (Random Forest, SVM, DNN) gives prediction of returns. Percentage Return of Each Sector

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The Formula:
• Return_sector = (Sector Return / Total Model Return) * 100%

5. Threshold Based Long/Short Decision
At the threshold system, each sector goes to a either long (bottom line) or short (top line) according to predicted returns and macroeconomic indicators.

6. Total Return Calculation
Long returns are added, and short returns are subtracted to compute total return.
Return Formula:
Total Return = Initial Cash Balance + Σ (Initial Cash Balance *(Sector Return / 100) if action 'Long') – Π ((1- (Sector Return / 100)) if action 'Short')

7. Method to calculate performance metrics
• Sharpe Ratio = Mean of Cumulative Returns / Standard
Deviation of Cumulative Returns
• VOLATILITY = STDEV of TOTAL RETURNS
• Max Drawdown = max(Peak - Value) / Peak)
• Alpha = Mean of Cumulative Returns - Benchmark Cumulative Return
• Risk-Reward Ratio = Average Return / Max Drawdown (if Max Drawdown ≠ 0)

8. Benchmarking Against S&P 500
It returns the cumulative return of whose S&P
performance is measured against to the Reliance is not that the strategy will work for sure.

V. RESULTS AND COMPARISON

Metrics	**Random Forest (RF)**	**Support Vector Machine (SVM)**	**Deep Neural Network (DNN)**
Initial Investment	$100,000,000	$100,000,000	$100,000,000
Total Return (%)	57.63%	61.80%	50.00%
Final Investment Value	$143,177,352.53	$132,198,495.76	$132,325,035.51
Final Investment Value (%)	43.18%	32.20%	32.33%

Metrics	**Random Forest (RF)**	**Support Vector Machine (SVM)**	**Deep Neural Network (DNN)**
Sharpe Ratio	1.7372	1.3886	1.9379
Volatility	0.8092	1.0856	0.6293
Maximum Drawdown	0.99	0.89	0.92
Alpha	1.13	1.23	0.94
Risk-Reward Ratio	1.42	1.69	1.33

© 2025 IJIRCCE | Volume 13, Issue 2, February 2025| DOI: 10.15680/IJIRCCE.2025.1302017

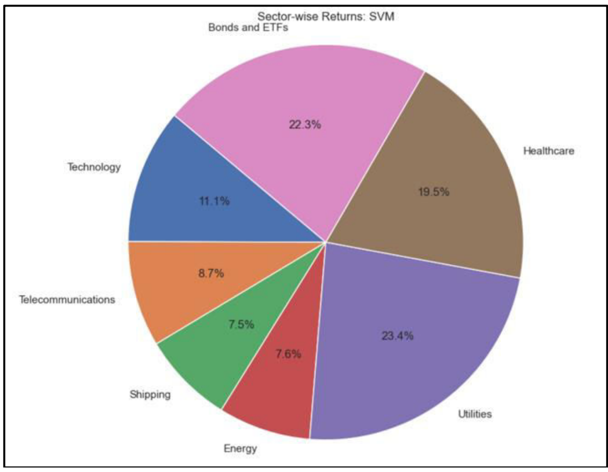
www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sector	Random Forest (RF)	Support Vector Machine (SVM)	Deep Neural Network (DNN)
Technology	11.49% (Long)	11.06% (Long)	13.23% (Long)
Telecommunications	3.49% (Short)	8.65% (Short)	10.93% (Long)
Shipping	8.74% (Short)	7.53% (Short)	8.00% (Short)
Energy	10.64% (Long)	7.56% (Short)	8.85% (Short)
Utilities	21.60% (Long)	23.40% (Long)	11.39% (Long)
Healthcare	22.80% (Long)	19.54% (Long)	23.45% (Long)
Bonds and ETFs	21.23% (Long)	22.25% (Long)	24.15% (Long)
Stress Test Results (Recession Scenario)			
Technology	-3.38%	-3.16%	-3.39%
Telecommunications	-2.99%	0.35%	0.46%
Shipping	-9.96%	-10.34%	-11.00%
Energy	-13.87%	-15.33%	-15.57%
Utilities	17.45%	19.46%	10.69%
Healthcare	23.14%	22.08%	21.73%
Bonds and ETFs	14.23%	15.75%	14.08%
Total Return Under Recession Scenario	24.63%	28.80%	17.00%

Note: The cumulative return for the S&P 500 for the same period is approximately 26.75%.



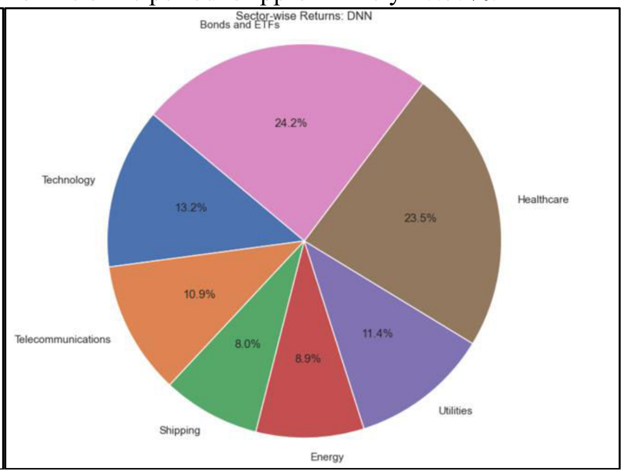


Fig.1. Sector Wise Returns of

Fig. 2. Sector Wise Returns of

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Support Vector Machine (SVM) Deep Neural Network (DNN)

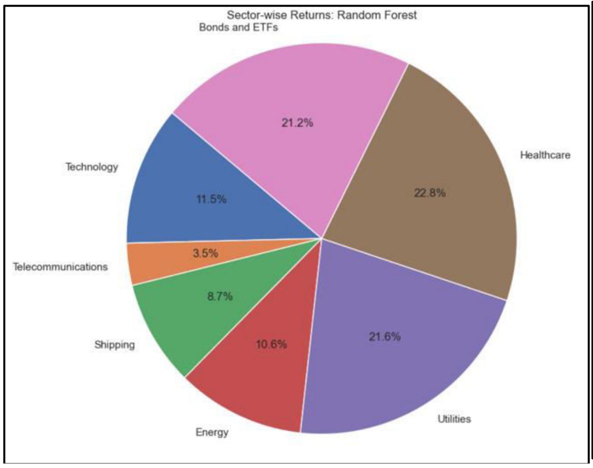


Fig. 3. Sector Wise Returns of Random Forest (RF)

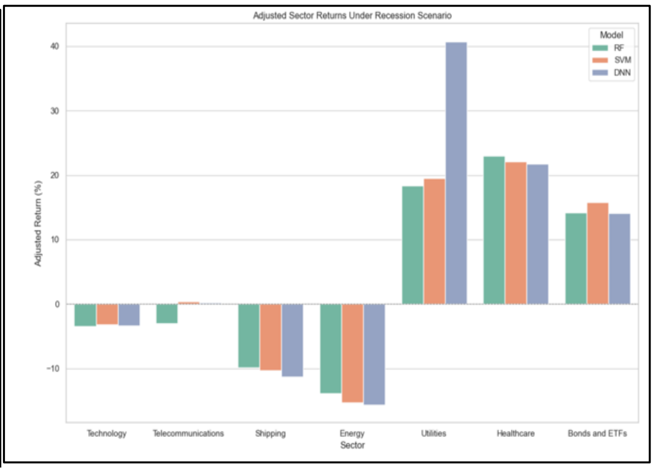


Fig 4. Stress Test Result

VI. SCOPE

This paper describes how to design, implement and perform a long/short equity trading strategy based on machine learning techniques that locates and isolates marketplace opportunities within the U.S. stock market which possess high profit potential. This study has three-fold scope: it traces the historical background of long/short strategies from their birth to history of high dimensionality approaches. It goes on to detail how the architecture and pipeline prepare our data (data preparation), what are the correct things to extract (feature engineering), how well we can expect our models to do on unseen examples (evaluation metrics) so that we can test alternative methods for predicting stock price movement, as well as sector returns.

The analysis combines three major machine learning models — Random Forest, Support Vector Machine (SVM), and Deep Neural Network (DNN) that are specially tuned for market prediction. We meet sector-specific predictions as well as market-wide predictions using a largescale dataset consisting of historical stock prices, technical indicators and macroeconomic factors. All 4 offers performance (including Total Returns, Final Investment Value (%), Sharpe ratio, Alpha, Volatility and Risk reward ratio and Benchmark Comparison with S&P 500) on a basis of Metric
That also extends to the risk management techniques of the strategies, such as position sizing and stop-loss mechanisms which are designed to mitigate losses but at the same time maximize returns. This paper aims to shed light on model performance with regard of capturing market structures, robustness of the strategy framework and other aspects related to ML infusion into long/short equity trading. With this framework we aspire to give some context regarding advanced quantitative trading strategies in modern algorithmic data driven financial markets with each having their own advantages and disadvantages.

REFERENCES

[1] Ang, A., Goetzmann, W. N., & Schaefer, S. M. (2009). Evaluation of active management of the Norwegian Government Pension Fund – Global. Norwegian Ministry of Finance.
[2] Avellaneda, M., & Lee, J. H. (2010). Statistical arbitrage in the U.S. equities market. Quantitative Finance, 10(7), 761-782.
[3] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794.
[4] Huang, Y., & Zhou, G. (2017). Asymmetry in long-short portfolios: New insights from high-dimensional asset pricing models. The Journal of Financial and Quantitative Analysis, 52(6), 2575-2600.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[5] Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Stock price prediction using support vector machines: A comprehensive review. Expert Systems with Applications, 124, 226-251.
[6] Huang, C., et al. (2020). Deep learning for finance: Deep portfolios. Applied Stochastic Models in Business and Industry, 36(1), 102-120.
[7] Jiang, Z., Xu, D., & Liang, J. (2020). A machine learning approach to predict stock returns with financial ratios and macroeconomic variables. Finance Research Letters, 32, 101183.
[8] Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. European Journal of Operational Research, 259(2), 689-702.
[9] Lopez de Prado, M. (2018). Advances in financial machine learning. Hoboken, NJ: John Wiley & Sons.
[10] Tsantekidis, A., Passalis, N., et al. (2017). Using deep learning to detect price change indications in financial markets. Proceedings of the 39th International Conference on Financial Engineering and Applications (FEM), 1-8.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING