International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Ensemble Machine Learning in Algorithmic Trading: Balancing Performance and Risk

Soham Pathak[1], Rhitik Nimbalkar[1], Suraj Shegukar[1], Pranav Mohite[1], Amol Take[2]
[1]Final Year Student, Department of Computer Engineering, Zeal College of Engineering and Research, Pune, Maharashtra, India
[2]Assistant Professor, Department of Computer Engineering, Zeal College of Engineering and Research, Pune, Maharashtra, India

ABSTRACT
This research delves into the potential of ensemble machine learning in algorithmic trading, aiming to strike a balance between performance and risk management. In the increasingly intricate landscape of financial markets, employing ensemble approaches that amalgamate multiple models becomes imperative. The study elucidates how these techniques enhance the accuracy, resilience, and adaptability of algorithmic trading systems. It further emphasizes the significance of model variability, hyperparameter tuning, and the bagging ensemble learning technique for optimizing outcomes. Moreover, the research scrutinizes the benefits of ensemble learning for algorithmic trading portfolios, particularly in terms of tail risk management and risk diversification. By utilizing ensemble approaches, several advantages emerge, including mitigated model bias, heightened generalizability, and an improved Sharpe Ratio. Ensemble methods enable traders to navigate market complexities more effectively, adapting to changing conditions and reducing the impact of outliers. This comprehensive approach not only enhances trading performance but also fortifies risk management strategies, making it indispensable in today's dynamic financial environment.

Keywords: Ensemble machine learning, algorithmic trading, risk management, performance, Stock market forecast

I. INTRODUCTION

In the rapidly evolving field of algorithmic trading, a mix of sophisticated machine learning algorithms has become indispensable for manoeuvring through the complexities of financial markets. Among these techniques, ensemble learning works especially well because it improves risk management and prediction accuracy by utilising the combined expertise of multiple models. This research explores the intricate relationships between risk mitigation and algorithmic trading's performance optimisation via the prism of ensemble machine learning.

A. Background and Motivation

Algorithmic trading, which has become a mainstay for traders seeking speed, precision, and efficiency when placing bets, has caused a revolution in the financial markets. With the ever-changing nature of the market and the complexity of financial goods, there has been an increasing quest for improved trading strategies and advanced risk management instruments. Because existing trading strategies are not able to keep up with the demands of this changing environment, newer ways need to be explored. In light of this, ensemble machine learning techniques have developed into powerful tools that outperform conventional approaches, offering a workable solution to the issues related to modern algorithmic trading.

1. **Market Evolution**: The financial markets have experienced significant transformation due to technological advancements, which have brought up both opportunities and challenges that require innovative solutions. Since traditional trading strategies often fail to reflect the nuances of modern markets, advanced alternatives are required.
2. **Technological Advancements**: With the advent of big data analytics, artificial intelligence, and high-frequency trading, algorithmic trading has reached unprecedented heights. Today's traders utilise technology for both speedy execution and complicated decision-making processes.

3. **Complexity of Financial Instruments**: The variety of financial instruments, including derivatives and structured products, makes trading strategies more difficult. Traditional models are not particularly flexible, so more advanced and adaptable methods are required.

B. Objectives of Balancing Performance and Risk in Algorithmic Trading

Amidst this landscape, the primary objective of algorithmic trading stands twofold: optimizing performance while concurrently managing risk effectively. The relentless pursuit of this delicate equilibrium is not only a strategic imperative but also crucial for sustained success in financial markets.

1. **Dual-Fold Objective:** The primary objectives of algorithmic trading in this context are to simultaneously optimise performance and effectively minimise risk. This delicate equilibrium must be pursued rigorously not only from a strategic standpoint but also for long-term success in the financial markets.
2. **Market Dynamics:** Algorithmic trading can only be successful if performance and risk are balanced, and this can only be done by carefully managing risk while maximising returns. Achieving the right balance ensures resilience in the face of market instability and enhances the sustainability of trading strategies.

C. Significance of Ensemble Learning in Financial Markets

Due to their remarkable ability to increase prediction accuracy and robustness, ensemble learning techniques have become more and more popular in the financial markets. In the field of algorithmic trading, precision and effective risk management are essential, which is why ensemble methods are a versatile and practical tactic.

I. **Precision in Predictive Accuracy:** Ensemble techniques integrate multiple models and leverage their collective knowledge to provide more accurate predictions. Precise forecasting of market movement is necessary for algorithmic trading to be successful, and ensemble learning excels in this domain.

II. **Robust Risk Control:** The significance of risk control in algorithmic trading cannot be overstated, and ensemble methods offer a solid framework for risk reduction and management. The flexibility of ensemble learning allows for a more sophisticated approach to risk management that is in keeping with the dynamic risk profiles of modern financial products.

Fundamentally, the framework and motivations for examining ensemble machine learning in algorithmic trading are rooted in the revolutionary nature of financial markets, the continuous advancement of trading technologies, and the need to maintain a careful equilibrium between effective risk mitigation and optimal performance.

II. Methodology

A. Overview of Ensemble Learning Methods

The technique known as ensemble learning blends several models to increase the prediction accuracy and robustness of algorithmic trading. To understand how ensemble learning approaches might impact performance and lower risk, one must have a solid understanding of them.

Bagging (Bootstrap Aggregating): - involves applying distinct portions of the training data to several instances of the same basic model for training. reduces overfitting and improves generalisation by averaging the predictions made by each model.

Boosting: - uses a step-by-step training method in which every new model learns from the mistakes made by the ones before it. increases the weight of incorrectly classified cases in an effort to increase accuracy.

Stacking: - involves employing a meta-model to combine the predictions of several different models that have been trained. makes use of each model's unique capabilities to improve prediction performance overall.

B. Ensemble Model Diversity and Its Impact on Performance and Risk

The diversity of component models, which is a major element influencing algorithmic trading performance and risk characteristics, is crucial to the effectiveness of ensemble learning.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Importance of Model Diversity: - Diverse models improve overall forecast accuracy and reduce the possibility of systematic mistakes by incorporating multiple points of view. A well-balanced mix of models leads to greater adaptability to a range of market conditions, which is beneficial for robust risk management.

Impact on Performance: - Diverse models provide for a more comprehensive coverage of the feature space, which improves the ensemble's ability to recognise complex patterns. Several models working together to generate predictions are shown to reduce overfitting and increase generalisation.

Impact on Risk Mitigation: - Model diversity strengthens the ensemble's resistance to unforeseen changes in the market by lowering the risk associated with relying too much on the quirks of a single model. The collaborative nature of numerous models strengthens the ensemble's ability to adapt to changing risk conditions.

C. **Hyperparameter Tuning in Ensemble Models for Algorithmic Trading**

In order to effectively employ ensemble models in algorithmic trading, it is necessary to optimise the hyperparameters so that they are tailored to specific market conditions.

Significance of Hyperparameter Tuning:

- Hyperparameters control the behaviour of ensemble models and have an impact on their predictability and flexibility.
- Models that have been fine-tuned are better able to fit certain market dynamics and uncover nuances from data patterns.

Strategies for Hyperparameter Tuning:

- Grid Search: Systematically explores a predefined set of hyperparameter combinations.
- Random Search: Randomly samples hyperparameter values, offering a more efficient exploration of the hyperparameter space.

Impact on Model Adaptability:

- Properly adjusted hyperparameters guarantee flexibility in changing market circumstances, enhancing the ensemble's capacity to react to shifting dynamics.
- Algorithmic trading sustains performance optimisation and risk management through constant monitoring and hyperparameter change.

III. Experiment Setup

In this section, we highlight the parameters of the recommended method for reproducibility. We describe the main phases of our trading system, including trade execution, dynamic asset selection and rating, technical aspects of model training and forecasting, and a discussion of the used baselines.

A. BACKTESTING: - As mentioned in the introduction, we have instantiated one example of our generic method using the pool of assets that is the NIFTY 50 Index of stocks. We decided to use a timeline spanning from January 2013 to December 2023 to back-test the framework. Rebalancing the portfolio, estimating forecasting hyper-parameters, assessing the signals, and running the signals through historical data are all part of the back-testing experiments. We use the walk-forward validation approach, which splits the research period into overlapping training periods and non-overlapping test (trading) periods. The finance sector frequently uses this technique to validate time-series data. The sample displays closing prices for the asset across the duration of the study. In this situation, we considered four years of training and one year of testing. The training period is additionally split into three years of development for training individual regressors and one year of validation for model selection, as indicated in Section VI. If the trials are conducted using the same setup, nine walks are achieved.

B. MODEL TRAINING: - For model training and parameter tuning, we used a GridSearch and a 10-fold TimeSeriesSplit, both employing scikit-learn implementations. The temporal relationship between the observation at day $d-1$ and d is taken into consideration in order to generate the same group of training and validation sets in this experimental scenario. In comparison, well-known machine-learning cross-validation methods like the Leave-

One-Out cross-validation and the k-fold cross-validation randomly sample data in different folds regardless of when it was collected. This type of technique is highly biassed when applied to time series forecasting since it combines aspects from the late past and the early future into the same fold of data.

C. RANKING AND DYNAMIC ASSET SELECTION:

Two parameters required for the dynamic asset selection are the rolling window length (T) and the accuracy threshold (ε), as discussed in Section VI-C. As the threshold value, $\varepsilon = 0.5$, we examined a range of rolling window lengths ($T = \{30, 40, 60\}$). Research demonstrating that MDA may successfully depict the interdependence between asset returns and their volatility (and, thus, forecast-ability) for intermediate return horizons (e.g., two months) had an impact on these judgements. The determined threshold value is $\varepsilon = 0.5$, as suggested by a comparable situation.

D. TRADING EXECUTION AND PORTFOLIO CONSTRUCTION:

We perform two × k operations, consisting of k long and k short operations, each day, as the paper makes note of. We fixed the number of pairings to be traded at $k = 5$ based on findings in similar works, since higher k values lead to a decrease in portfolio performance for returns and risks. Since the trading session is designed to be intra-day, we open the positions at the beginning of the training day and close them at the end. Stated differently, we rebalance our portfolio on a regular basis. Like the authors, we assume transaction expenses daily.

E. BASELINES:

The NIFTY 50 buy-and-hold strategy (Buy-and-hold) and a baseline based on cumulative five-day return (5-DAY) are the two statistical arbitrage trading baselines against which our dynamic asset selection method (ENS-DS) and model selection methodology (ENS) are benchmarked. This enables us to assess each strategy's relative value. The latter two methods are well-established quantitative processes that are mostly employed as standards to evaluate the profitability of alternative investment strategies. They are as follows:

1. **5-DAY** - Prior to each trading day, we organise the collection of stocks in ascending order based on the cumulative 5-day return. The stocks at the top would be those with the biggest 5-day cumulative negative returns, and the stocks at the bottom would have the highest 5-day cumulative positive returns. We construct an equal weight portfolio in the sorted list by going long on the top k stocks and short on the flop k stocks. We open trade positions for $k = 5$ at the beginning of the day, close them at the end, and rebalance the portfolio daily, just like in the previous techniques. There are no differences in the data, trading execution, or portfolio formation.

2. **Buy&Hold** - Using this method, the 2013 purchase of the NIFTY 50 exchange-traded fund is maintained until the end of the backtesting period in December 2023. This passive strategy does not employ trading signals. This type of baseline is widely recognised as a helpful benchmark in the literature.

F. IMPLEMENTATION DETAILS:

This paper describes a Python-built method that makes use of the scikit-learn module and the LightGBM Python API. The following parameters were found on the desktop computer used for the studies: 16 GB of RAM, a 3.00GHz Intel(R) Xeon(R) Gold 6136 CPU, and a 64-bit version of Linux Ubuntu are all included. The complete solution's code is available to the public at "https://github.com/Surajshegukar/Risk-Controlled-Algorithmic-Trading-Using-Machine-Learning" to improve repeatability.

IV. RISK MANAGEMENT STRATEGIES

A. Ensemble Learning for Risk Diversification in Algorithmic Trading Portfolios

Along with its alluring claims, algorithmic trading (AT) carries a certain amount of risk. Retaining profitability while avoiding risk is an ongoing problem for AT practitioners. Ensemble learning is a powerful tool since it opens up the prospect of diversifying algorithmic trading portfolios and enhancing their resistance to market volatility. Think of your algorithmic portfolio as having an Easter basket-like structure. If you wanted to lessen the likelihood of dropping them, you wouldn't put them all in one fragile container. Instead, you would split them across a few sturdy baskets. Similar in theory, ensemble learning creates a more robust and well-rounded portfolio by combining the benefits of multiple trading algorithms.

Ensemble learning contributes to risk diversification in at:

- Reduced Model Bias: Individual model biases often lead to overfitting or vulnerability to certain market conditions. These biases are reduced by the inherent diversity of ensembles, resulting in more flexible and well-rounded portfolio responses.
- Improved Generalizability: Certain models might function well with training data but badly with unidentified data. By combining several learning strategies, ensembles lessen the risk of catastrophic losses during unforeseen market swings and increase their capacity to handle a variety of market dynamics and generalise well.
- Enhanced Sharpe Ratio: The Sharpe Ratio calculates an investment's return compared to its risk. Ensemble portfolios often have higher Sharpe Ratios than single-model portfolios due to their greater diversity, which indicates better risk-adjusted returns.

Several ensemble techniques can be applied in at portfolio construction, each with its own advantages:

- **Bagging (e.g., Random Forests)**: produces many models using marginally different training data, exposing the portfolio to a range of data samples and broadening its diversification.
- **Boosting (e.g., XGBoost)**: builds models in a stepwise manner, emphasising cases that earlier models misclassified, resulting in a portfolio that continuously adjusts and gains knowledge from shifts in the market.
- **Stacking**: trains a meta-model that integrates forecasts from separate base models, adding a deeper comprehension of market dynamics to the portfolio.

However, implementing ensemble learning in AT requires careful consideration:

- **Model Selection and Tuning**: Selecting an appropriate ensemble technique and fine-tuning its hyperparameters are essential to optimise the benefits of diversification without compromising performance.
- **Correlation Analysis**: Make sure the highly connected models in your ensemble don't counteract the diversification effect.
- **Computational Resources**: Ensembles require a sufficient infrastructure of hardware and software and can be computationally expensive.

By paying close attention to these elements and utilising group learning, AT professionals may create diversified portfolios that more adeptly navigate the complexities of the market, lowering risk and paving the way for sustained financial success. Remember that searching for misplaced or broken goods is less likely when you have a wide variety of sturdy and interesting baskets. If you embrace the power of collective learning, you'll witness the grace and fortitude with which your algorithmic trading portfolio navigates turbulent markets.

B. **Tail Risk Management Using Ensemble Approaches: Protecting Against the Unforeseen**

The shadow of tail risk, which refers to erratic, uncommon events that have the potential to cause catastrophic losses, frequently coexists with the allure of large rewards in the financial markets. When black swans begin to emerge, conventional risk management tools can be disastrously inadequate in their ability to handle normal market swings. Here's where group approaches come into play, giving you a useful toolkit to help you steer clear of the treacherous tail risk waters.

Think of your portfolio as a ship sailing across a peaceful sea. You have sails and rudders with conventional risk management that can handle consistent breezes and quiet waters alike. However, ensemble techniques provide you a more durable boat with a watertight hull and state-of-the-art radar that keeps an eye out for approaching storms.

Ensemble approaches enhance tail risk management:

- **Diversity of Predictions**: Ensembles combine many models with different structures and learning algorithms to capture different market perspectives. By doing so, reliance on any one model's blind spots is reduced, potentially exposing latent vulnerabilities before they materialise as tail events.
- **Robustness to Noise and Outliers**: Data noise or outliers may distort the predictions of individual models and raise the risk. Ensembles leverage the collective wisdom of the group to filter out noise and outliers and create more accurate projections, particularly under turbulent market conditions.
- **Improved Scenario Analysis**: Ensembles mimic a variety of extreme market events that extend beyond the reach of available historical data, allowing for more comprehensive scenario testing. This helps you to stress-test your portfolio and foresee potential tail risks before they happen.

Several ensemble techniques shine in the context of tail risk management:

- **Robust Ensembles**: These ensembles deliberately downweight models that are prone to overfitting and emphasise models that capture rare events, providing a more cautious and tail-risk-aware perspective.
- **Bayesian Ensembles**: These ensembles provide you with the ability to identify scenarios that may have a substantial impact, even if they appear unlikely, and to assess the level of confidence attached to projections through the use of probabilistic forecasting and uncertainty quantification.
- **Hybrid Ensembles**: Combining machine learning with traditional statistical models allows one to leverage the benefits of each approach and enhances the ability to capture both regular and exceptional market phenomena.

However, utilizing ensemble approaches for tail risk management requires careful consideration:

- **Model Selection and Tuning**: Selecting the appropriate ensemble technique and fine-tuning its parameters are essential to maximising the gains while avoiding unintentional bias introduction.
- **Interpretability and Explainability**: Although ensembles provide greater accuracy, there might be complexity in their internal operations. Building trust and making informed decisions require ensuring interpretability and understanding how models arrive at their predictions.
- **Data Availability and Quality**:Accurate and thorough data are essential for ensemble model training and calibration. It is essential to make investments in robust data cleansing procedures and top-notch data infrastructure.

You may give your portfolio the resilience it needs to effectively navigate the choppy waters of the financial markets by using cooperative tactics and resolving these challenges. Remember that when black swans are coming, it's better to have a varied fleet of weatherproof ships than one lone, lonely vessel. So, go for it in a group setting and navigate the rough waters of tail risk with confidence and preparedness.

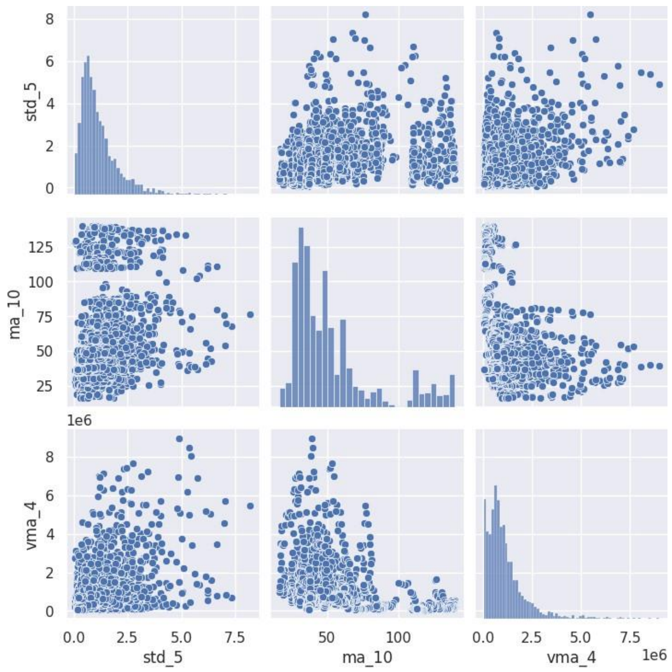
Results

The results are presented from three angles: (i) the predictive performance of the ensemble under different training configurations for each individual regressor; (ii) return assessment prior to and following transaction costs; (iii) exposure to common risk factors via risk metrics analysis; and (iv) comparisons with state-of-the-art trading strategies.

A. MODELSELECTION PERFORMANCE:

First, we assess the model's performance using a range of walk-forward setups. More specifically, we established two separate validation and test durations (measured in days) and a three-year development period. The initial configuration included three years of training, one year of validation (252 days), and one year of testing, or nine walks, across a study period that ran from January 2013 to January 2023. In the second configuration, we chose a study period that spanned from July 2013 to January 2016, with three years of training, and we shortened the validation and test periods to six months (126 days), yielding eighteen walks. The models corresponding to the two

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930



configurations will be assessed between March 2013 and January 2016, which coincides with the trading/test period. In terms of the root mean square error (RMSE) between the achieved return and the projected return, we provide the model results as an average over firms and walks. Furthermore, we present the portfolio performance for every model when employing a trading strategy for k = 5 pair in terms of risk metrics (MaxDD) and annual returns (Return p.a.). The maximum drawdown, or MaxDD, is a measure of the maximum amount of wealth decline that a cumulative return has produced from its maximum value over time. The findings are shown in Figure provides an overview of the model performances (for each stock) and associated ensemble following the model selection process.

Model shows a constant decrease in RMSE as the validation and test times decrease. This is because there aren't as many samples available. The Model SVR with TI as features and industry-level data performs the worst in all circumstances. When it comes to the other pole, ARIMA has the lowest RMSE. After a year of validation and another year of test setting, the LGB model performs best in terms of returns using data at the industry level with TI as features. ARIMA is the second best performer, with annual returns of about 1.5 percent. The most hazardous alternative uses SVR with industry-level data and TI as characteristics (Figure 4c). The risk-return features and anticipated performances for each forecasting system are shown in. Shorter validation times also show a reduction in performance as indicated by return and MaxDD. The most surprising result is shown by the ensemble's low returns (0.0877), which were obtained by taking the average of outputs produced by models with a 3 years validation period. The majority of the models outperform the ensemble. In conclusion, shows that the ENS, which had a year for validation and a year for test preparation, is the best-performing model.

	precision	recall	f1-score	support
0	0.53	0.78	0.63	426
1	0.58	0.30	0.40	423
accuracy			0.54	849
macro avg	0.55	0.54	0.51	849
weighted avg	0.55	0.54	0.51	849

B. RETURN CHARACTERISTICS ASSESSMENT:

As explained, the configuration that produces the best model performance is as follows: We log findings and assess the efficacy of the proposed technique during three years of development, one year of validation, one year of model selection, and one year of testing. These facts lead us to focus our next analysis on this arrangement.

As anticipated, baselines and the underlying individual regressors are outperformed by the equal weighted ensemble (ENS) and the dynamic asset selection (ENS-DS). The yearly returns are approximately twice as high as the return of individual regressors (such as LGB) and nearly four times higher than the level of the Buy&Hold. Even after the transaction fees, the same ratios hold true. In addition, we note that before transaction fees, or the cost of entering the market, the mean daily return for the ENS is 0.11% and for the ENS-DS, it is 0.12% (T = 30, T = 40). Although the returns decrease after the transaction expenses are deducted, they are still

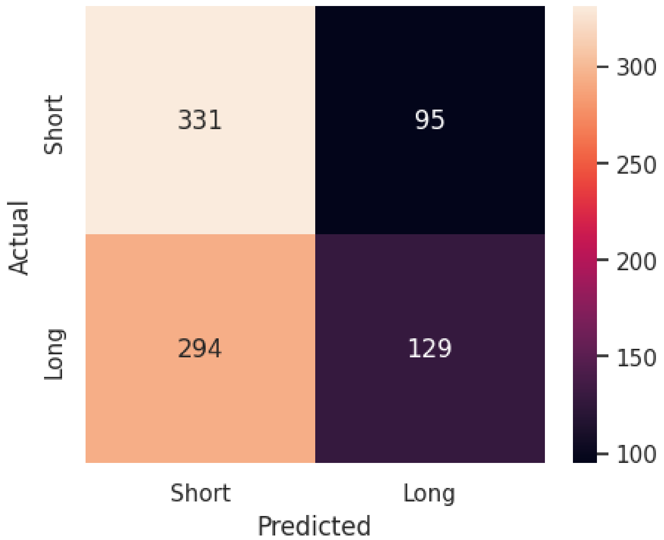
International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

much greater than the baselines. Upon analysing the statistical moments, we find that every strategy, with the exception of SVR, exhibits positive skewness, indicating a broader tail for gains.

Lastly, we assessed the statistical significance of the returns using a Newey-West t-statistic and the null hypothesis that the mean return is equal to zero (the crucial value at the 5% significance level is 1.9600). The test showed that the returns were statistically significant before and after transaction expenses.

C. RISK EXPOSURE ASSESSMENT: -Before and after transaction costs, offer a thorough analysis of the risk taken by the trading strategies. We start by looking at the tail risk. Value at Risk (Var 1%) before transaction costs, as shown in, shows that ARIMA is the riskiest approach, outperforming both baselines by −4.1%. At the other pole, with approximately −2.3% for ENS-DS and −2.6% for ENS, are the ensemble strategies. After the transaction costs (Var 1%), the image remains the same; nevertheless, at this juncture, ARIMA (−3.80%) is almost identical to the 5-DAY baseline (−3.78%). Compared to [17], which produces values between −5.9 and −6.9 percent after transaction charges, our strategies yield substantially lower values. Under the conditional value at risk (CVar 1%), positions have slightly changed, with ARIMA (−5.7%) offering the highest risk among regression models, while it is still less than the 5-DAY method (−6.2%). The Sharpe ratio is a risk indicator that is defined as the excess return per unit of risk expressed in standard deviations [56]. Generally speaking, a portfolio that has a Sharpe ratio greater than one has outperformed one that has a ratio smaller than one. Generally speaking, a portfolio with a higher Sharpe ratio will outperform one with a lower ratio.demonstrates that the Sharpe ratio for the proposed ENS-DS (T = 40) rose from 1.72 for the basic ensemble to 1.85 for it before transaction expenses were taken into account.

Analogously, another statistic that evaluates the reward-to-risk ratio is the Sortino Ratio, which considers risk expressed as downside deviations. Examining the results in, we can see that downward deviations are less indicated for the proposed solutions. This leads inevitably to a more favourable Sortino ratio: around 3 for ENS-DS strategies before and 2.2 after transaction costs for ENS-DS strategies, and 2.89 for ENS before and 2.11 after. These equals double the values for the 5-DAY baseline, or 1.49 before transaction costs and 1.16 after transaction costs. In MaxDD's opinion, a lower figure denotes a more manageable loss on an investment. We can also see that ENS-DS strategies outperform the



Buy&Hold and 5-DAY baselines for this parameter. ENS produces a result of 13.76% for ENS-DS, T = 40 , which decreases to 11.46%. This represents less than one-fourth of the 5-DAY and one-third of the Buy-and-Hold (55%) strategies. As expected, following transaction costs, the values of MaxDD have increased for both ENS (28.42%) and ENS-DS, T = 40 (26.06%); still, they remain below the baselines. Before considering the transaction costs , we find that the ENS-DS, T = 40, has a higher Calmar Ratio value of 2.90 compared to 2.27 for ENS. The ENS strategy records a value of 0.77 after transaction costs, whereas the ENS-DS, T = 40, records a value of 0.87, compared to 0.16 for the Buy & Hold strategy. The Calmar Ratio calculates the number of average annual returns required to recover from a maximum drawdown by scaling annualised returns by the value of MaxDD. As a result, it takes 1.15 years for the ENS-DS strategy to recover from the maximum drawdown, compared to around 6 years for the Buy & Hold strategy. Expected returns are split into two categories by the Omega metric, which was developed by [57]: gains and losses, or returns above and below the expected rate (upside and downside). To put it simply, Omega is the ratio of positive (upside) returns to negative (downside) returns. Prior to transaction charges, this ratio increased in our proposed strategy from 1.42 (ENS) to 1.46 (ENS-DS, T = 40). Following transaction costs, the results show the same positive trend , with the ENS simple ensemble showing a 1.3003 score and the ENS-DS, T = 40, showing a 1.3106 score. This increases the likelihood of getting daily positive returns. We may conclude that the dynamical asset selection approach, ENS-DS, performs better in terms of risk than the simple ensemble, ENS, regardless of the look-back period, T. More precisely, when T = 40 is fixed, the most improvement is demonstrated.

D. COMPARISON WITH STATE-OF-THE-ART TRADING STRATEGIES: -

To assess its performances in a real-world trading scenario, we compared our proposed StatArb technique with a set of state-of-the-art portfolio strategies that are well-established in the literature [58], using the same validation process as in [29]. Specifically, we considered the competitors listed below (for further details on these algorithms, the reader is referred to [58]):

BHP - Buy&Hold-based Portfolio, i.e. a portfolio implementation of the traditional Buy&Hold strategy described in Section VII-E, whereby the investor, rather than buying a single asset (such an ETF or a stock), buys shares of all the index companies in step with their prices;

CRP - Constant Rebalanced Portfolio, for example, a variation of the BHP method wherein the weights in the portfolio are periodically rebalanced in response to shifts in the prices of the underlying assets;

UP - Universal Portfolio, this is a parameterized CRP technique over the whole simplex domain. Through adaptive learning from historical data, the method maximises the log-optimal growth rate over time;

EG - Exponential Gradient, i.e. a momentum strategy that focuses on the best performing asset of the index, in the last time period.

For this analysis, we took a broad time range into consideration, spanning from 2009 to 2015. Statistical arbitrage strategies do not perform particularly well during low volatility periods (2013-2015) or bull market regimes (2009-2013, post-global financial crisis), which fall within this period, according to several articles in the literature. Figure 5 displays the cumulative returns, or equity curves, of the two recommended approaches (ENS and ENS-DS) in comparison to their rivals and the behaviour of the market as a whole (MKT). Moreover, displays the risk characteristics (Maximum Drawdown, Sharpe Ratio, and Sortino Ratio) for each of these strategies. Additionally, we can observe that, for this parameter, ENS-DS strategies perform better than the Buy&Hold and 5-DAY baselines. For ENS-DS, T = 40, ENS yields a result of 13.76%, which falls to 11.46%. This amounts to less than 25% of the 5-DAY strategy and 33% of the Buy-and-Hold (55%). The values of MaxDD have risen for both ENS (28.42%) and ENS-DS, T = 40 (26.06%), as anticipated, in response to transaction costs; however, they are still below the baselines. We observe that the ENS-DS, T = 40, has a higher Calmar Ratio value of 2.90 compared to 2.27 for ENS before taking into account the transaction costs.

Taking everything into account, our results also perform well when compared to other statistical arbitrage methods, including the classic pairs trading results in [7], where the top 20 pairings from 1962 to 2002 had a Sharpe ratio of 0.59. According to the authors of [5], a generalised pairs trading from 1997 to 2007 had a 1.44 Sharpe ratio. In [24], the author proposes a methodology that uses Elman neural networks with a Sharpe ratio of around 1.5 and ELECTRE III for the years 1992–2006. Likewise, researchers at Union Bank of Switzerland (UBS) employ RF models in their work to generate monthly portfolios. Trading signals are constructed using the top quantile of the projected monthly returns; equities are bought long based on this quantile and sold short based on the flop quantile. Among the Asia and Pacific index components that the authors test the methodology on between 1997 and 2015 is the Australia SP300 index. When we examine the cumulative return over time (theirseparately, MSCI AC Asia Pacific ex-Japan vs. ours, ENS-DS, T = 40), we discover that both strategies fared well in the financial crisis of 2007–2009. After thereafter, they notice a decline in performance, which peaks in 2011 and 2015. As seen in Figure 5, our recommended strategy produced the highest earnings in the midst of the disturbance peaking in early 2009. Even though the technique yields moderate returns after that point, it has an increasing trend. Finally, in terms of risk, their claimed MaxDD is interestingly registered in 2012, after the global financial crisis, and is somewhat lower (around 10% compared to our MaxDD values of 11% before to transaction costs and 26% following transaction costs).

I. TRADE-OFFS AND CHALLENGES

Algorithmic trading requires complex models, and attaining optimal performance involves compromises and challenges. This survey study's key findings are discussed, with a focus on the relationship between performance and risk in ensemble machine learning for risk-controlled algorithmic trading.

A. Performance-risk Trade-offs in Ensemble Machine Learning

Challenge: One of the main obstacles in the pursuit of higher performance using ensemble machine learning is striking a balance between improved prediction accuracy and effective risk management. Because ensemble models might overfit to historical data, their ability to generalise to new market conditions may be compromised. To capitalise on their joint capabilities, ensemble models employ a variety of strategies.

Trade-off*:* For best results, the ensemble model complexity has to be precisely adjusted. While increasing the model's complexity may increase accuracy, it also increases the risk of overfitting. Consequently, traders are required to achieve a challenging equilibrium between optimising the predictive capacity of the ensemble approaches and reducing any possible disadvantages associated with increased model complexity.

B. Challenges in Implementing Ensemble Strategies for Risk Control:

Challenge: Utilising ensemble systems for risk control presents a unique challenge: integrating disparate models into a logical framework. Coordination of several models' operations and ensuring their cooperative contribution to effective risk management require careful consideration.

Trade-off: There is a trade-off between maintaining robust risk control mechanisms and promoting the synergy of ensemble models. Resolving problems with dynamic weighting, model combination, and correlation analysis is necessary to achieve the optimal balance. Overcoming the difficulties of integrating several models so that they function well together in the context of risk control is essential for successful implementation.

Computational Considerations and Scalability Issues:

Challenge: There may be challenges due to ensemble machine learning models' requirements for memory, computing power, and scalability. These criteria are important because algorithms are expected to handle large databases and complex calculations in real-time trading scenarios.

Trade-off: It is shown that there is a trade-off between computer efficiency and model sophistication. Finding a balance between the need for advanced models and the practical constraint on computational resources necessitates strategic decision-making. Traders need to assess the computational demands of ensemble approaches and optimise for efficacy while maintaining the responsiveness and scalability of the trading system.

VI. CASE STUDIES AND APPLICATION

A. Stock Market Prediction and Trading:

Machine learning algorithms are employed to assess historical stock price data, transaction volumes, and other relevant information in order to forecast future price movements and make trading decisions in the stock market. Machine learning models, such as regression, time series analysis, and deep learning, are widely applied in this discipline. Through the identification of trends, patterns, and anomalies in stock market data, these models assist traders in formulating winning trading strategies. For example, sentiment analysis from news and social media, price changes, or technical indicators can all be employed by predictive algorithms to generate buy or sell recommendations. Reinforcement learning techniques can also be used to dynamically optimise trading strategies in response to changing market conditions.

B. Cryptocurrency Trading:

Cryptocurrency trading is the buying and selling of virtual currencies such as Bitcoin, Ethereum, and Litecoin. Machine learning algorithms are being used more and more in bitcoin trading due to the markets' extraordinary volatility and complexity. These computers can analyse enormous amounts of historical pricing data, trade volumes, order book data, and sentiment analysis from social media to forecast price movements and identify profitable trading opportunities. For example, machine learning models can be trained to recognise patterns in cryptocurrency price charts and generate trade suggestions using technical indicators such as moving averages, RSI, Bollinger Bands, and MACD. Sentiment analysis algorithms can also be used to examine news stories, social media posts, and other information sources in order to ascertain the mood of the market and forecast changes in price.

C. Foreign Exchange (Forex) Trading:

Buying and selling different currencies on the international exchange market is known as trading foreign exchange, or forex. Machine learning algorithms are widely used in forex trading to assess currency exchange rates, economic data, geopolitical events, and other factors that affect currency pricing. These algorithms have the ability to provide trading signals based on either short- or long-term changes in currency pair values. For example, neural networks, regression models, and time series analysis are helpful methods for forecasting changes in exchange rates based on historical data and basic analysis. Reinforcement learning methods can also be used to construct adaptive trading strategies that respond dynamically to changing market conditions.

D. High-Frequency Trading:

To predict future price movements and make trading decisions in the stock market, machine learning algorithms evaluate transaction volumes, historical stock price data, and other pertinent information. In this field, machine learning models including deep learning, time series analysis, and regression are frequently used. By recognising patterns, trends, and anomalies in stock market data, these models help traders create profitable trading methods. Predictive algorithms can, for instance, use price fluctuations, sentiment analysis from news and social media, or technical indicators to suggest buys or sells. Moreover, trading methods can be dynamically optimised through reinforcement learning approaches in response to shifting market conditions. In order to reduce execution times and obtain a competitive advantage in the market, HFT techniques frequently rely on co-location services and low-latency trading infrastructure.

VII. CONCLUSION

The importance of ensemble techniques in reaching a risk-reward balance is highlighted in the conclusion of "Ensemble Machine Learning in Algorithmic Trading". By combining multiple models, these techniques improve forecasting accuracy, reduce overfitting, and strengthen trading strategy robustness. Using case studies, the study demonstrates the value of ensemble learning in various trading scenarios. When it comes down to it, ensemble techniques give traders a solid foundation for improving performance and managing risk in erratic financial markets.

VIII. REFERENCES

[1] S. Carta, D. R. Recupero, R. Saia, and M. M. Stanciu, “A general approach for risk-controlled trading based on machine learning and statistical arbitrage,” in Proc. 6th Int. Conf. Mach. Learn., Optim., Data Science (LOD) in Lecture Notes in Computer Science, vol. 12565, 2020.

[2] J. Jayko, ‘A demon of our own design: Markets, hedge funds, and the perils of financial innovation,’ J. Pension Econ. Finance, vol. 7, no. 3, p. 363, 2008. G. Vidyamurthy, Pairs Trading: Quantitative Methods and Analysis. Hoboken, NJ, USA: Wiley, 2004.

[3] M. T. Leung, H. Daouk, and A.-S. Chen, “Forecasting stock indices: A comparison of classification and level estimation models,” Int. J. Forecasting, vol. 16, no. 2, pp. 173–190, 2000.

[4] G. S. Atsalakis and K. P. Valavanis, ‘Surveying stock market forecasting techniques—Part II: Soft

computing methods,' Expert Syst. Appl., vol. 36, no. 3, pp. 5932– 5941, Apr. 2009.

[5] N. Huck, "Large data sets and machine learning: Applications to statistical arbitrage,' Eur. J. Oper. Res., vol. 278, no. 1, pp. 330–342, Oct. 2019

[6] R. Almeida, L. F. Brandão, and D. G. da Silva, "Machine learning for high-frequency trading: a practical approach." In Proceedings of the 20th European Conference on Artificial Intelligence, 2012.

[7] J. Moody and M. Saffell, "Learning to trade." Technical Report, The University of Michigan, 2000.

[8] H. Sargolzaei, R. A. Javidan, and A. Poon, "Machine learning for algorithmic trading: predicting and optimizing execution strategies." In Proceedings of the 2015 International Conference on Machine Learning and Applications (ICMLA), 2015.

[9] Y. Sun, D. Liu, J. Chen, and Y. Wang, "An empirical study of machine learning techniques for risk control in algorithmic trading." In Proceedings of the 2017 2nd International Conference on Knowledge Engineering and Applications (ICKEA), 2017.

[10] X. Zhang, Y. Yu, and Y. Tang, "A comprehensive review of machine learning techniques for algorithmic trading strategies." Artificial Intelligence Review, vol.
52, no. 4, pp. 2275-2295, 2019.

[11] H. Chen, T. Zheng, and Y. Zhang, "Machine learning techniques for risk control in high-frequency trading." In Proceedings of the 2016 International Conference on Progress in Informatics and Computing (PIC), 2016.

[12] Y. Hong, X. Li, and B. Zhang, "Risk control in algorithmic trading: a survey on performance evaluation, risk assessment, and regulation." Journal of Systems Science and Complexity, vol. 34, no. 1, pp. 1- 19, 2021.

[13] L. Wu, Y. Zhang, J. Shang, and Q. Zhang, "Risk control of algorithmic trading based on deep reinforcement learning." In Proceedings of the 2020 International Conference on Intelligent Transportation,
Big Data and Smart City (ICITBS), 2020

[14] Soham Pathak, Antara Pawar, Shruti Taware, Sarthak Kulkarni, Afsha Akkalkot, "A Survey on Machine Learning Algorithms for Risk-Controlled Algorithmic Trading" in International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 1 | Issue 1 | ISSN: 2456-3307, 2023