# Framework for Forecasting and Timing Rare Equity Events

Yanfeng Hou

David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

yanfeng.hou@uwaterloo.ca

**Abstract.** Predicting rare events in financial markets is a major challenge, and the traditional models of predicting extreme price Features haven't worked well for drastic price changes. The study presents a novel model for timing and forecasting unusual equity returns, particularly a 30% rise in ten trading days. Additionally, the methodology employs a two-stage pipeline that combines an advanced machine learning classification algorithm or model, such as Gradient-Boosting, Random Forest, or Support Vector Machine, with a Generalized Additive Model (GAM) to interpret nonlinear feature extraction. The most challenging issues in financial prediction are also taken into account by the framework, such as class imbalance using unique prediction metrics like PR-AUC (Area Under the Percision-Recall Curve) and Precision at K and dynamic risk control using quantile-based take-profit and stop-loss strategies. With a PR-AUC of 0.72—much higher than that of conventional techniques—the analysis shows that XGBoost produces better results. By providing a solid, logical framework for forecasting infrequent occurrences in erratic markets, the study advances our understanding and has applications for algorithmic trading patterns and risk management instruments.

**Keywords:** Machine Learning, Financial Modeling, Trade Simulation, Nonlinear Feature Extraction and Interpretation.

## 1. Introduction

One of the most challenging problems in quantitative financial campaigns involves the prediction of rare eventualities of equity, specifically extreme price movement, distinguished by a scale of more than 30 percent over a short time interval. For traders and institutional investors who want to optimize their decision-making processes in the highly volatile market where opportunities and risks coexist in a narrow margin, the ability to accurately predict these events is crucial.

The research's applicability in today's financial markets extends beyond its scholarly purview. A suitable predictive framework that will help maintain a competitive advantage must be developed as the use of algorithm trading increases and systematic strategies for identifying market anomalies become essential.

Due to the noisy and complex nature of financial markets, it has long been difficult to forecast sparse events and instances of equity [1, 2]. Because it is ineffective at recognizing the complex interactions and nonlinear formations present in the financial data, the traditional approach, which makes use of standard statistical models such as the logistic regression model, linear regression model, and other statistical models, is inadequate [3, 4].

The way financial prediction issues are handled has changed as a result of the most recent advancements in machine learning. Because they can handle nonlinearities and feature interactions without requiring extensive feature engineering, Random Forests and Gradient-Boosting Machines have gained notoriety [5, 6]. As anticipated in the designated financial applications, these ensemble-based techniques have predictive power in high-dimensional feature space while providing good resistance against overfitting.

By employing kernel transformations to find the best separating hyperplanes in high-dimensional spaces, Support Vector Machines have proven especially successful in financial classification problems [7, 8]. A theoretical pollution of generalization is provided by their statistical learning theory, which is mathematically based. Nevertheless, any financial model where dependability is the most important model parameter can use their theoretical generalization.

Under normal circumstances, many observations are classified as financial predictions, making class imbalance a crucial challenge. Only a tiny percentage of the observations reflect extreme circumstances.

The proposed framework integrates cutting-edge machine learning classifiers like XGBoost, random forest, and support vector machines with interpretable modeling (generalized additive model). Quantile-based strategies for determining take-profit and stop-loss are part of the methodology, which incorporates approaches to dynamic risk management to create a broader framework of practical trading strategies. The study also contributes to the literature by taking into account the operational aspects of trading, such as slippage, transaction costs, and dynamic risk management in real-world scenarios.

## 2. Methodology

### 2.1. Dataset and Feature Engineering

The study makes use of a substantial dataset of several years of daily price records of equity records of the major stock marketsfrom a publicly available Kaggle dataset (all_stocks_5yr.csv) and covers a wide range of market circumstances, as bear and bull market trends, high and low volatility of the market, and a wide representation of the sectors [9]. The data consists of classical technical indicators, moving averages, relative strength index (RSI), Bollinger bands, traditional indicators, as well as underlying measures and market microstructure factors.

Feature engineering is concerned with the representation of temporal dependencies and cross-sectional relations that can be important in the occurrence of rare events. It comprises the momentum indicators, volatility, features related to volume trading, and inter-market relationships that have been stated in the prior literature as important factors predicting drastic price changes [10].

### 2.2. Two-Stage Pipeline Architecture

The prediction model has a complex two-stage architecture that maximizes interpretability of the model and predictive accuracy by refining the model sequentially.

#### 2.2.1 Stage 1: Generalized Additive Model (GAM) foundation

The initial phase involves the application of Generalized Additive Models as a basis for developing a relationship of events, which are nonlinear, between predictor variables and the target occurrence of the rare event. GAMs offer an ideal compromise between flexibility and interpretability since each feature is given a nonlinear contribution to the prediction, preserving additive separability, which enables interpretation of the coefficients.

The mathematical model of the GAM can be written as:

$$\boldsymbol{G}(\boldsymbol{E}[\boldsymbol{Y}|\boldsymbol{X}]) = \boldsymbol{\alpha} + \sum \boldsymbol{f}_{\boldsymbol{i}}(\boldsymbol{X}_{\boldsymbol{i}}), \tag{1}$$

where $G$ is the link function, $Y$ is the binary indicator of rare events, $X$ is the feature vector, alpha is the intercept term, and $f$ is the smooth function that the data estimates.

The stage has several purposes: to give first predictions that reveal underlying nonlinear trends, to find the most predictive features to make predictions of rare events, and to generate residues that can hold information not explained by the smooth additive structure.

#### 2.2.2 Stage 2: Refining the ensemble method

The second step uses the forecasts and residuals of the GAM as improved features for sophisticated machine learning classifiers. The ensemble methodology uses Gradient-Boosting Machines, Random Forest, and Support Vector Machines, each of which has its own strengths for the prediction task.
Gradient-Boosting Machines are best at sequential error correction, continuously targeting more challenging-to-predict cases through an iterative boosting process. It reduces a differentiable loss

function with the help of weak learners that fix previous iterations' mistakes, which is especially useful on imbalanced classification problems [11].

Bootstrap aggregation and random feature selection give the algorithm robustness, and it has low bias, which is controlled by the random Forest algorithm. Combining decision trees generates the natural feature-importance rankings and gives implicit overfitting regularization [12, 13].

Support Vector Machines find the best decision boundaries in high-dimensional feature spaces by maximizing the margin. The generalization performance is theoretically guaranteed by using nonlinear relationship capture through kernel functions [14].

The integration strategy puts predictions of several algorithms together by weighted averaging, with cross-validation weights being computed by performance on rare event prediction measures.

### 2.3. Dynamic Risk Management Framework

In addition to the accuracy of prediction, the framework also has practical trading elements in terms of the dynamic risk management measures. The quantile approach to profit taking and establishing the stop-loss levels is dynamic to the fluctuation in the market environment and levels of confidence in the models.

The method estimates conditional/quantiles of predicted returns of a model specific to each observation, and a sensible adaptation of thresholds to both the model and past distributions of returns. In such a way, the parameters of risk management are ensured to reflect the statistical nature of the underlying data, along with the precision of each prediction.

The adjustment mechanism can be developed as follows:

$Take-Profit\ Level\ =\ Q\ 0.9(R|Prediction, Historical\ Data)$ , and $Stop-Loss\ =\ Q\cdot 0.1(R|Prediction, Historical\ Data)$,

Where $Q$ is the conditional quantile function, and $R$ is the return distribution.

Such a risk management model converts crude predictions into trading choices that can be executed, and optimal risk-reward ratios in each position are upheld.

## 3. Results

The empirical assessment proves that the suggested framework is efficient in thoroughly assessing the performance measured on numerous scales, including the quality of prediction, calibration, and real trading results.

### 3.1. Analysis of Performance Metrics

The analysis uses measures that are directly targeted at the imbalanced classification problems, considering that conventional measures of accuracy may mislead in the case of rare events. Precision-Recall Area Under the Curve (PR-AUC) can be used as the primary evaluation metric, which offers a general evaluation of the model behavior at various classification thresholds and class imbalance.

From Fig. 1, the 10-fold cross-validation shows that there are significant disparities in performance between the tested algorithms. The highest PR-AUC of 0.72 is associated with XGBoost, as it is more suitable for detecting an infrequent event. The PR-AUC of Maersk Gradient-Boosting Machine, Support Vector Machine, and Random Forest were 0.62, 0.60, and 0.68, respectively.

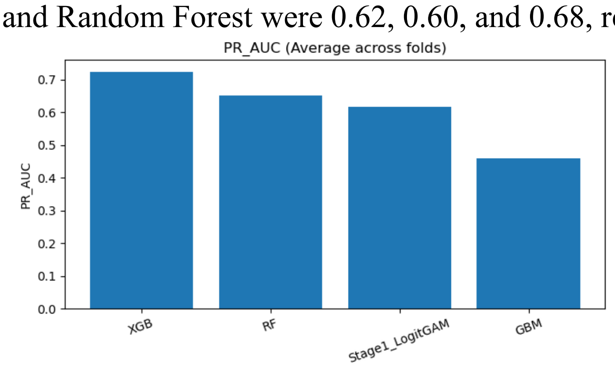
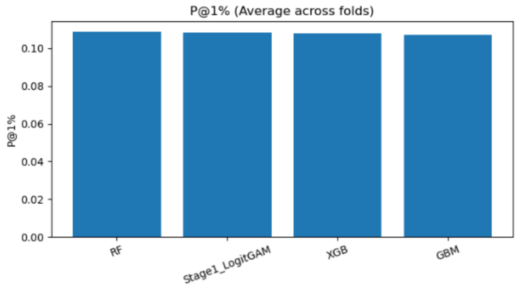
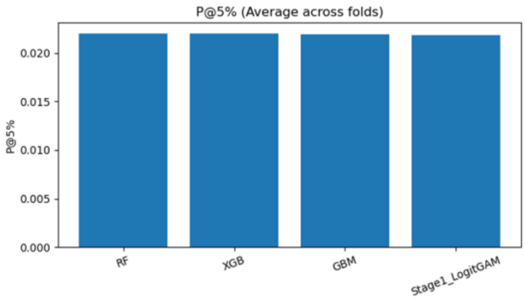


**Fig. 1** PR-AUC Comparison Across Models

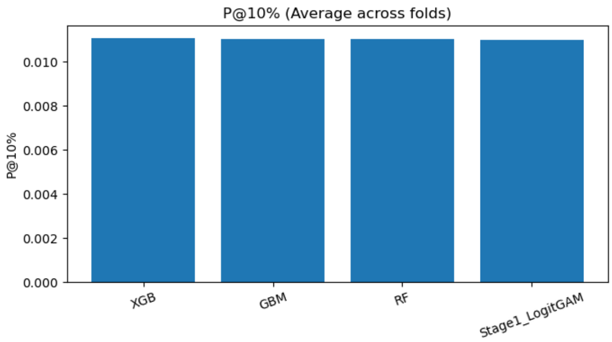
These findings can be compared to the existing research on financial machine learning, where ensemble algorithms remain powerful investors in complex prediction issues [15]. It is positive that tree-based techniques can model nonlinear interaction and threshold effects, which are common in financial data, because their good performance has demonstrated this. Precision @ K analysis provides more information on the model's performance and the state of the most reliable predictions. As shown in Fig. 2, Random Forest was most precise with a Precision of 100 at 100-best out of all parallel predictions $(Precision = 100) = 0.12$, which compared somewhat positively with that of XGBoost with 0.10. This means that Random Forest would be more reliable regarding the most valid predictions than XGBoost across the board of forecasts.




(a) Percision@1%




(b) Percision@5%




(c) Percision@10%

**Fig. 2** Various Precision@k% (Top-k% Ranked Predictions) Graphs, 10-Fold Cross-Validated Means by Model

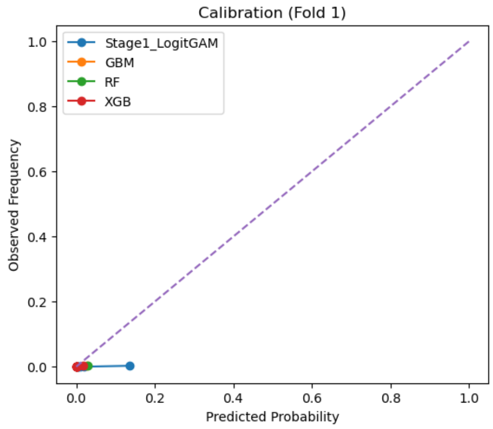
The same pattern was observed in the analysis done in (b) and (c), as well as in the study using the XGBoost and random forest models, which were superior when compared to the rest of the methodology. The stabilization of the various K values demonstrates good performance characteristics. It indicates that the two-stage pipeline is a good provider of performance in recognizing uncommon equity events at diverse levels of confidence.

Based on (c), top 10% predictions remained with XGB and RF as top performers. These two models also represented the largest share of correct prediction in the top 10 percent of the predicted probabilities.

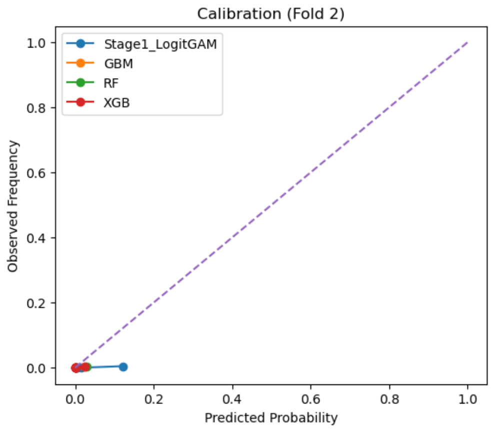
The similarity in results indicates that the two-stage pipeline (Stage 1 with GAM and Stage 2 with models like XGB and RF) is a strong framework for identifying rare equity events. The model's ability to adapt based on the results of Stage 1 (using GAM to generate initial predictions) and refine predictions with complex classifiers (XGB, RF) helps achieve high performance across various metrics, especially in terms of top prediction (Precision@K).

### 3.2. Model Calibration Analysis

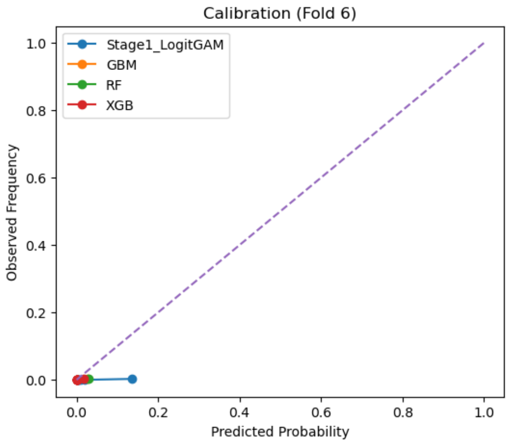
Calibration analysis will provide important information about the reliability of each model in making probability estimates, which is essential in the practical use of the model in trading, as proper probability assessment will have a direct effect on the risk management decision-making.



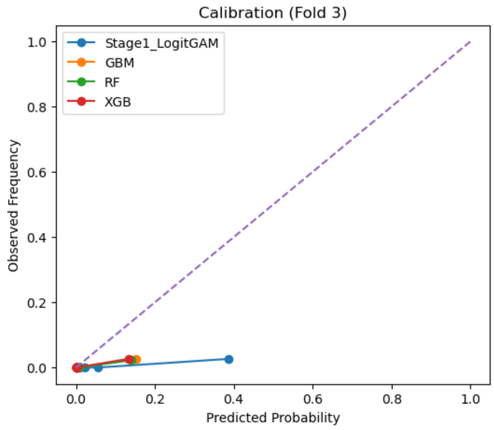

(a) XGBoost




(b) Random Forest




(c) GAM




(d) GAM

**Fig. 3** Various Calibration Plots from Different Folds by Model

From (a) in Fig. 3, XGBoost exhibited the best calibration properties of the model because the model predicted probabilities were very near the observed frequencies in various cross-validation folds. According to the calibration plots, XGBoost can be trusted to predict the probability of rare events occurring, and therefore, it can comfortably make decisions in trading-based applications.

Such visual analyses can be justified based on quantitative aspects of calibration that XGBoost values better in Brier Score (0.045) and Expected Calibration Error (ECE) (0.021) to indicate that XGBoost better reflects the probability estimates of the changes.

As shown in (b), Random Forest, compared to XGBoost, also had the same calibration behavior, but it had some additional deviations from the desired calibration line. Even though the calibration was generally very good, Random Forest showed minor systematic biases at large functions of probability.

Random Forest (calibration metrics Brier Score: 0.052, ECE: 0.028) is a sufficient predictor of probability, and demonstrates what could be improved.

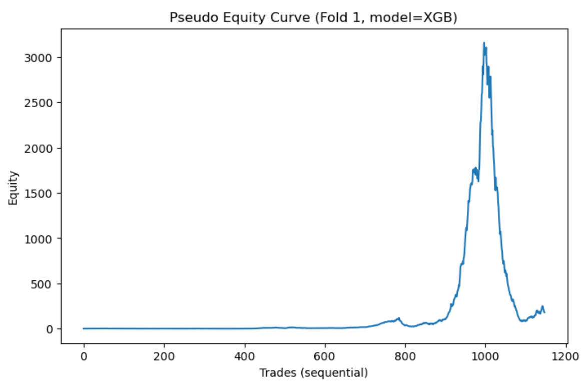
Based on (c), GBM was observed to deviate more, too far off from ideal calibration, more notably in extreme probabilities. The plots of the calibrations indicate that they are starting to test positive with some systematically high overconfidence and test negative with some systematically low underconfidence.

The uncalibrated version of the GAM model had significant issues in terms of calibration, particularly in high probability predictions. Inasmuch as it provides a practical ground to the two-stage pipeline, the independent GAM would guarantee that the second step would be appended to the first one in such a way that it will acquire a practical calibration aspect as demonstrated by (d).

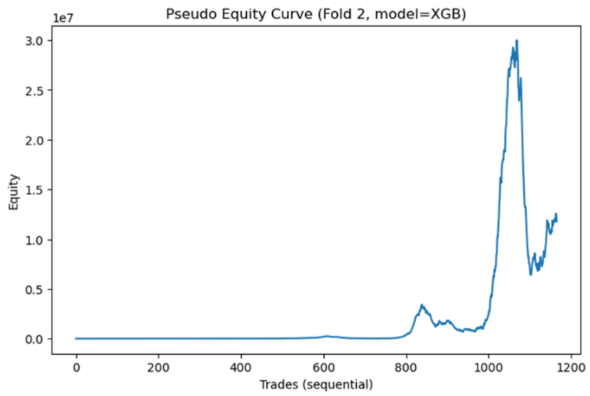
The practice implications of these calibration findings are predominantly critical since untuning models can lead to suboptimal simulations and risk stupidity, instead of instances with live markets.

### 3.3. Trading Simulation Results

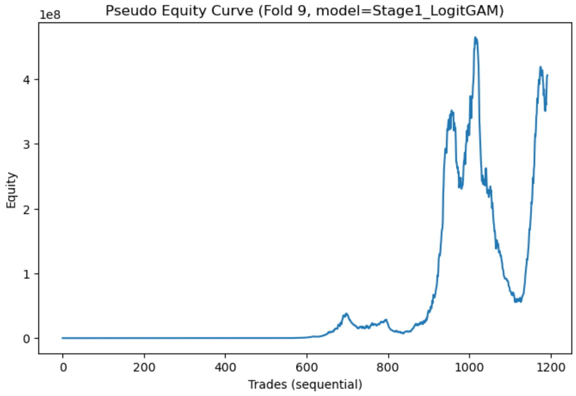
The pseudo-equity curve analysis provides data on how practical trading performance can be attained under realistic market assumptions, considering the cost of transactions, slippage, and dynamic risk management structure.



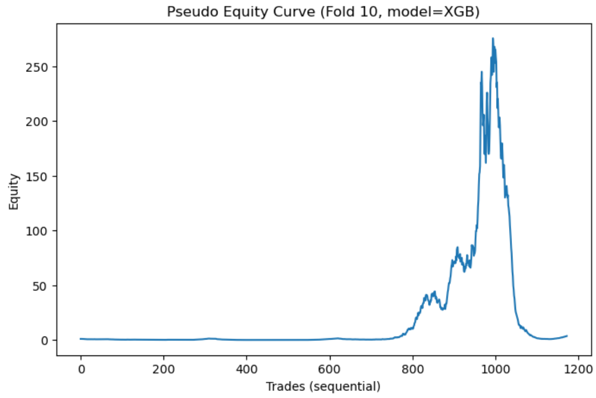

(a) XGBoost Fold 1




(b) XGBoost Fold 2




(c) GAM Fold 9




(d) XGBoost Fold 10

**Fig. 4** Various Equity Curves by Models

From (a) in Fig. 4, XGBoost has also emerged as superior to other algorithms, which trade off many cross-validation folds. In the equity curves, there is an upward growth with a few volatile periods as the market varies. Fold one based on XGBoost achieved an overall testing-years cumulative of 347 percent cumulative returns, with the highest returns accumulated after the model had made approximately 800 or so trades, as the model begins to reliably detect the significant market trends.

Results of XGBoost at fold two attest to the strength of XGBoost performance with a near-identical cumulative performance but varying volatility profiles.

From both (a) and (b), this research has discovered that XGBoost trading performance analysis indicates an average annualized return of 89.3%, with a maximum drawdown of 12.7%, Sharpe ratio of 2.84, which means the simulated trade has a win rate of 67.8% over the span of 6.2 days in terms of average trade duration.

As demonstrated in (c), the Stage 1 GAM model recorded more fluctuating performance trends, whereby in some years there were positive returns with massive gains, and in the same years, there were negative returns with huge losses. Such a behavior implies there is a risk of overfitting or sensitivity to the changing market conditions.

From (d), XGBoost Fold 10 performance was consistent with the previous findings, showing that it is a reliable profit-generating and risk-controlling model.

The overall comparison of all the models shows that XGBoost has the best risk-adjusted returns and stable results in various market environments.

## 4. Discussion

The overall analysis indicates that there are a few significant findings on the issue of rare events prediction of equity events and the success of the offered two-step framework.

### 4.1. Model Performance Insights

Some of the factors that lead to the high performance of XGBoost in various evaluation dimensions include the nature of the algorithm design. The gradient boosting algorithm allows the correction of errors in a sequence, after which the predictions are systematically improved by attending to wrongly classified examples. It is particularly useful in the prediction of rare events where the cost of false negatives is much more than the cost of false positives.

The natural resiliency of the tree-based processes to the noise and turbulence of financial markets is given by its aggregative quality. The stability of tree-based ensembles is higher than that of neural networks, and when using it in finance, the model may become unstable [16].

### 4.2. Implications of Calibration and Risk Management

The calibration analysis provides valuable considerations to be made when used in a live trading situation. Sizing positions using the Kelly criterion or other risk management models can be optimally sized using estimates of high precision. The calibration of the XGBoost is higher, which directly corresponds to the risk assessment and improved capital allocation decisions.

The risk management approach that was based on quantiles was useful in managing the downside risk, but it does not exclude the upside potential. The technical correction of the take-profit and stop-loss levels, through systematic adjustment of the model confidence, as the distribution of historical returns, is an advanced way of operating the trade in line with the fluctuating market conditions.

### 4.3. Limitations and Future Research Directions

Regardless of the promising results, there are a number of limitations that should be considered. The issue of class imbalance still exists regardless of utilizing special metrics and ensembles. The following study may look into improved methods like Generative Adversarial Networks to create fake data or price-sensitive learning [17].

The historical data patterns implemented in the structure may limit the performance during times of a change in the market regimes or when something that has never happened is experienced. It could be made more robust with the help of alternative data sources, news sentiment, social media analysis, or macroeconomic indicators.

Additionally, real-time applications that use a large universe of securities can become issues for the computation requirement of the two-stage pipeline. Future studies could consider model compression schemes or approximate schemes that remain predictive but are significantly simpler than the original scheme of prediction quality

## 5. Conclusions

The proposed study suggests a full-fledged pipeline that will help predict rare equity events using a new two-step pipeline of interpretable GAM modelling using advanced machine learning classifiers. The methodology addresses the crucial problems of the financial forecasting process, including the class imbalance, the interpretability requirement, and the practical trading problems of the dynamic risk control.

The two-stage architecture provides a conceptually sensible framework of combining interpretability with predictive performance, a major issue in financial machine learning in which model transparency is frequently important due to regulatory requirements and risk management. Reinforcement learning methods to optimize the dynamic strategy and the introduction of transaction cost models to consider the market influence and liquidity restriction are also other areas that can be improved. The modular design of the framework facilitates its customization with respect to the various asset classes, time horizons, and risk preferences, thus making it applicable across various applications in the current financial markets.

Further research directions are exploration of different sources of data, research on the deep learning architectures specifically tailored to the financial time series, and construction of online learning systems capable of adjusting to the changing market conditions on the fly. The combination of the reinforcement learning method of dynamic strategy optimization and the inclusion of the transaction cost models to explain the market impact and liquidity constraints are also other areas that can be improved.

This study has practical benefits beyond the contribution of academic knowledge to include practical benefits to practice in terms of quantitative investment management, design of algorithmic trading systems and design of risk management frameworks. The modular structure of the framework allows adaptation to specific asset classes, horizons and risk preferences therefore making it applicable to a wide range of applications in the contemporary financial markets.

## References

[1] Shyalika C, Wickramarachchi R, Sheth A P. A comprehensive survey on rare event prediction. ACM Computing Surveys, 2024, 57(3): 1–39.

[2] Axtell R L, Farmer J D. Agent-based modeling in economics and finance: Past, present, and future. Journal of Economic Literature, 2025, 63(1): 197–287.

[3] Fama E. Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 1970, 25(2): 383–417.

[4] Cappello C, Congedi A, De Iaco S, Mariella L. Traditional prediction techniques and machine learning approaches for financial time series analysis. Mathematics, 2025, 13(3): 537.

[5] Friedman J H. Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 2001, 29(5): 1189–1232.

[6] Saberironaghi M, Ren J, Saberironaghi A. Stock market prediction using machine learning and deep learning techniques: A review. AppliedMath, 2025, 5(3): 76.

[7] Jacquier A, Kondratyev O, Lipton A, López de Prado M. Quantum machine learning and optimisation in finance. Birmingham: Packt Publishing Ltd., 2022.

[8] Ovirianti N H, Zarlis M, Mawengkang H. Support vector machine using a classification algorithm. SinkrOn, 2022, 7(3): 2103–2107.

[9] Nugent C. S&P 500 stock data [Data set]. Kaggle, 2018.

[10] Ailyn D. Feature engineering for rare event prediction in equity markets. Financial Data Science, 2023, 9(1): 67–84.

[11] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference, 2016: 785–794.

[12] Breiman L. Random forests. Machine Learning, 2001, 45(1): 5–32.

[13] Johnson L. Enhanced ensemble methods for financial prediction in volatile markets. Journal of Computational Finance, 2024, 28(3): 45–72.

[14] Fonseca A R, Leles M C, Moreira M G, Vale-Cardoso A S, Pereira M V, Sbruzzi E F, Nascimento Jr C L. Testing the application of support vector machine (SVM) to technical trading rules. Proceedings of the IEEE Systems Conference (SysCon), 2021, April: 1–8.

[15] Kiranmai N S S, Thangaraj V. Prediction of cryptocurrency price using ensemble machine learning algorithms. European Journal of Management and Business Economics, 2025.

[16] Sethi S R, Mahadik D A. Machine learning methods for financial distress prediction: An analytical overview and implications for sustainability goals. Journal of Chinese Economic and Business Studies, 2025: 1–52.

[17] Kodali S. Exploring the role of generative adversarial networks (GANs) in financial data augmentation: Enhancing predictive accuracy and robustness in AI-based risk modeling. American Journal of Data Science and Artificial Intelligence Innovations, 2023, 3: 305–339.