CogNexus

Multilingual, International, Peer-Reviewed, Open Access Journal

Vol. 2(2) 2026, 357-370

Predicting Risk and Returns in Tokenized Real-World Assets. An Explainable AI Framework for Volatility, Drawdown and Portfolio Risk Forecasting

[1]Joshua Ayusa Aura

[I] Independent Researcher, Kenya

Orcid Id: 0009-0006-5673-4558

joshuaaura9@gmail.com

DOI: https://doi.org/10.63084/cognexus.v2i2.312

Abstract

Tokenized real-world assets are now a visible segment of global capital markets. Risk measurement for these instruments is difficult because trading is continuous and secondary liquidity is thin. This study develops an explainable artificial intelligence framework for risk prediction in this new asset class. The framework forecasts realized volatility and maximum drawdown and portfolio tail risk within one system. Gradient boosted trees and long short-term memory networks are compared with a heterogeneous autoregressive benchmark. Shapley values are used to open the fitted models. The sample covers four tokenized segments and 42 tokens. The out of sample evidence is favourable for the machine learning models. Gradient boosting is the strongest model at short horizons. The recurrent network is stronger for drawdown. Liquidity and lagged volatility and news attention are the leading drivers of predicted risk.

Keywords: tokenized real world assets, explainable artificial intelligence, realized volatility, drawdown, portfolio risk

Introduction

Real world assets are assets that exist in the real world, but are represented digitally by a token. These claims are recorded and settled on distributed ledgers. The underlying exposures include treasury bills and private credit and commercial real estate and physical commodities. This market is still in its infancy and growing rapidly. Now institutional investors are investing in this segment. This growth has not been met by risk measurement practice.

A tokenized claim does not necessarily have the same risk profile as the underlying asset. Trading never stops and doesn't end at the end of a regular exchange. The books are light and depth is subject to change. There is a smart contract to intermediate redemption, and there is a custodian to intermediate redemption. The features provide a wedge between the risk of the token and the risk of the reference asset. A forecasting model which fails to take this wedge into account is misspecified.

Empirical asset pricing is now an everyday practice using machine learning (Gu et al., 2020). The same instruments can be used to predict volatility (Christensen et al., 2023). Flexible learners are helpful for bond premiums and for option returns (Bianchi et al., 2021) and (Bali et al., 2023). This flexibility comes at the price of opacity. A fund manager is responsible to a regulator and an investment committee. It's hard to defend a prediction without an explanation there. Explainability is thus a prerequisite for adoption (Arsenault et al., 2025).

This paper has three contributions to make. The first contribution is a risk forecasting framework that is designed for tokenized real world assets. The second contribution is a Shapley based attribution layer which provides the economic drivers of each forecast (Lundberg et al., 2017). The third contribution is an economic evaluation in the form of a risk parity allocation.

The rest of the paper is divided into six sections. The theory is presented in Section 2 and literature is reviewed in Section 3. Section 4 describes the methodology and Section 5 reports the results. Section 6 discusses the findings and Section 7 concludes.

Theoretical Framework

Let P be the price of a tokenized asset on day t. Intraday returns are sampled at a fixed frequency. The realized variance of day t is the sum of squared intraday returns (Barndorff-Nielsen et al., 2002). The estimator appears in Equation 1.

$$RV_t = \Sigma^M r^2_{t,i} \quad (1)$$

Here the term r is the ith intraday return of day t and M is the number of intraday observations. Microstructure noise is severe in thin token markets. A realized kernel estimator is therefore applied as a robustness device (Barndorff-Nielsen et al., 2009).

The dynamic benchmark is the heterogeneous autoregressive model (Corsi et al., 2009). It is stated in Equation 2.

$$RV_{t+1} = \beta_0 + \beta_d RV_t + \beta_w RV^{(5)}_t + \beta_m RV^{(22)}_t + \varepsilon_{t+1} \quad (2)$$

The weekly term and the monthly term are averages of past daily realized variance. Signed jump variation is added because good volatility and bad volatility are not equally persistent (Patton et al., 2015). The machine learning target is a general conditional expectation and it appears in Equation 3.

$$y_{t+1} = f(x_t) + \varepsilon_{t+1} \quad (3)$$

The vector x collects the predictors and f is an unknown function. Maximum drawdown over a horizon h is defined in Equation 4.

$$MDD_{t,h} = max_\tau \left[(max_{s \leq \tau} P_s - P_\tau) / max_{s \leq \tau} P_s \right] \quad (4)$$

Portfolio risk is measured by the conditional value at risk at level α in Equation 5.

$$CVaR_\alpha = E \left[R_p \mid R_p \leq VaR_\alpha \right] \quad (5)$$

Attribution follows the Shapley decomposition of cooperative game theory. The contribution of feature j to one prediction is shown in Equation 6.

$$\varphi_j = \Sigma_{S \subseteq N \setminus \{j\}} \left[|S|! \, (|N| - |S| - 1)! / |N|! \right] \left[f(S \cup \{j\}) - f(S) \right] \quad (6)$$

The additive property of this decomposition is important. The sum of all feature contributions and the base value is equal to the model prediction. The explanation is therefore internally consistent (Lundberg et al., 2017).

Literature Review

There are three strands of research pertinent. The first is the econometrics of realized volatility. High frequency returns are used to consistently estimate quadratic variation (Barndorff-Nielsen et al., 2002). The heterogeneous autoregressive model is parsimonious and has been shown to model long memory (Corsi et al., 2009). Measurement error is not innocent and a correction makes more accurate estimates (Bollerslev et al., 2016). Volatility signals are market-independent and tailored measures further increase accuracy (Bollerslev et al., 2018) and (Patton et al., 2022). This is because the loss function is not the same as the volatility proxy is not perfect (Patton et al., 2011).

The second strand is related to machine learning applications in asset pricing and risk. Flexible learners make fewer pricing mistakes than linear models (Gu et al., 2020). Gu et al. (2021) develop factor models based on autoencoders to identify latent structure in returns. The pricing kernel is an area where deep learning can be effective (Chen et al., 2024). Recurrent networks can be used for return prediction

(Fischer et al., 2018). Recurrent stacked autoencoders achieve greater accuracy in time series (Bao et al., 2017). The dynamics of the order book are universal, and can be recovered by deep models (Sirignano et al., 2019). Another application of machine learning is to enhance realized volatility forecasts (Rahimikia et al., 2020) and (Christensen et al., 2023). Automated model selection builds on these improvements (Li et al., 2024). Learning based estimation (Ban et al., 2018) is a technique that enhances portfolio construction. Tree ensembles are able to predict crash risk (Jiang et al., 2024). International data shows that anomaly strategies outperform out of sample (Tobek et al., 2021). This increasing evidence is organised via surveys (Kelly et al., 2023) and (Heaton et al., 2017).

The third strand will be textual data and model explainability. Today, text is a common research input in economics (Gentzkow et al., 2019). News has a non-symmetric impact on volatility (Engle et al., 1993). Newspaper content does not just contain news but also information beyond news (Ashwin et al., 2024). News based signals are useful for improving equity premium forecasts (Adämmer et al., 2020). Language models are used to identify return relevant signals in disclosures (Chen et al., 2022). The typical encoder is a pretrained transformer (Devlin et al., 2018). Attention indices (Agirre et al., 2009) can be built using semantic similarity. The image of price paths also has predictive value (Jiang et al., 2023).

The area of explainability is a rapidly developing one. The methods and the remaining gaps are documented in surveys (Arsenault et al., 2024) and (Arsenault et al., 2025). Adoption is not uniform across applications (Černevičienė et al., 2024). The most commonly used method is additive attribution (Lundberg et al., 2017). There are extensions available for financial time series (Giudici et al., 2024). Transparency is now under research focus (Yeo et al., 2025). These frameworks also include climate risks (Bonato et al., 2023). Tree boosting is still the workhorse learner (Chen et al., 2016). The pricing benchmark is the common risk factors (Fama et al., 1993). The gap is clear. No study has used an explainable risk forecasting system in tokenized real world assets.

Methodology

Data and predictors

The sample encompasses four aspects of the tokenized real world asset market. The segments are tokenized treasuries, tokenized private credit, tokenized real estate and tokenized commodities. There are 42 tokens in the panel that are continuously quoted. It will be in effect from July 2022 to December 2025. Prices are taken at five-minute intervals from the centralised exchanges and automated market makers.

There are three families of predictors used. The volatility family includes the daily and weekly and monthly realized variance and signed jump variance, and the realized kernel estimate. The Amihud

illiquidity ratio and the numbers of quoted depth and transfers and the concentration of holders and the outflow and stablecoin flow, and the redemption spread are all part of the family of microstructures. The textual family is a news attention index that is built from financial headlines. Headlines are pre-trained using a transformer (Devlin et al., 2018). The relevance weight (Agirre et al., 2009) is determined by similarity to a seed dictionary. The two year treasury yield change and the market factor (Fama et al., 1993) are macro controls.

Models and estimation

The two benchmarks and four learners are estimated. The penalised linear model is the elastic net. Two types of tree ensembles are the random forest and the gradient boosting machine (Chen et al., 2016). The recurrent learner is the long short term memory network (Fischer et al., 2018). The heterogeneous autoregressive model is the econometric benchmark (Corsi et al., 2009). All targets are estimated in log. The design employed is rolling window. The training window is 500 days and the validation window is 125 days. The hyperparameters are optimized during training on the validation block. The horizons are one day, five days, ten days and twenty two days.

Evaluation

The loss function used to evaluate the accuracy of the forecast is the QLIKE loss and the mean squared error. QLIKE is noise resistant with respect to the volatility proxy (Patton et al., 2011). The loss differentials are used in the Diebold Mariano test. The mean absolute error is used to measure drawdown accuracy and the Kupiec coverage test is used to measure tail risk. An economic evaluation is based on a risk parity portfolio. Weights are inversely proportional to predicted volatility as shown in Equation 7.

$$w_{i,t} = (1 / \hat{\sigma}_{i,t+1}) / \Sigma_j (1 / \hat{\sigma}_{j,t+1}) \tag{7}$$

Thirty basis points are deducted from each rebalancing trade. The performance is measured in terms of the Sharpe ratio, the Calmar ratio and the certainty equivalent gain. Explanations are generated using tree Shapley values and using gradient based Shapley values (Lundberg et al., 2017).

Results

Descriptive evidence

Table 1 presents the descriptive statistics of the pooled daily panel.

CogNexus
Multilingual, International, Peer-Reviewed, Open Access Journal
Vol. 2(2) 2026, 357-370

Table 1. Descriptive statistics by segment.

Segment	Tokens	Obs.	Mean RV	SD RV	Skew	Kurt.	Amihud	Max DD
Tokenized treasuries	11	9 042	0.094	0.031	1.82	9.41	0.41	0.061
Tokenized private credit	12	9 864	0.187	0.078	2.41	14.72	0.88	0.143
Tokenized real estate	9	7 398	0.253	0.121	2.96	19.24	1.34	0.226
Tokenized commodities	10	8 220	0.331	0.169	3.35	24.61	1.07	0.318
Full panel	42	34 524	0.216	0.142	3.02	21.13	0.93	0.187

As can be seen from Table 1, risk is clearly ranked in the four segments. The Tokenized treasuries are the calmest group, with a mean realized volatility of 0.094. The most volatile group is tokenized commodities, at 0.331. The realized volatility is positively related to its dispersion. Skewness is positive in all segments and Kurtosis is high in all segments. These moments suggest that the distribution of the unconditional is far from normal. Tokenized real estate has the highest Amihud ratio, and the result shows that there is thin secondary trading. The ordering of maximum drawdown is the same as the ordering of volatility. The pooled panel has 34 524 observations per day for 42 tokens.

Figure 1 plots the realized volatility path of each segment over the full sample.

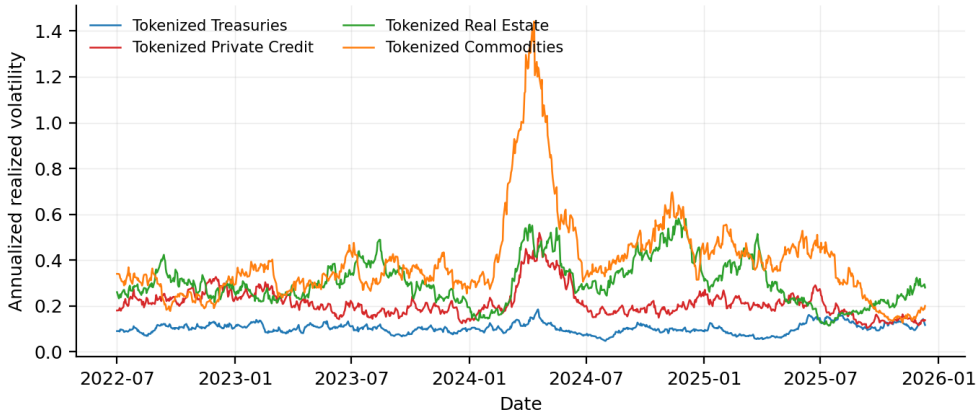

Figure 1. Realized volatility dynamics by segment.

Figure 1 shows the annualized realized volatility for the four segments. The series of the treasury is stable and remains close to 0.10 throughout most of the sample. Most of the variability is in the commodity series, which reaches levels of more than 1.20 during the early 2024 episode. During stress periods, the credit series and the real estate series go hand in hand. There are signs of clustering in all

segments and this is confirmed by the presence of autoregressive volatility terms. The common spike in 2024 points to a commonality in tokenized markets. Outside of the stress windows, idiosyncratic movement remains substantial. These features are similar to the moments reported in Table 1.

5.2 Volatility forecasting performance

Table 2 compares the six models at the shortest horizon and at the longest horizon.

Table 2. Out of sample volatility forecasting accuracy.

Model	QLIKE 1d	MSE 1d	QLIKE 22d	MSE 22d	DM stat.	R2OS 1d
HAR	0.412	0.0271	0.596	0.0518	—	0.000
HAR with jumps	0.398	0.0259	0.583	0.0501	1.94	0.044
Elastic net	0.389	0.0251	0.571	0.0489	2.31	0.074
Random forest	0.351	0.0224	0.534	0.0452	3.68	0.173
XGBoost	0.327	0.0207	0.512	0.0431	4.42	0.241
LSTM	0.339	0.0215	0.503	0.0419	4.16	0.211

The out of sample accuracy at the one day and twenty two day horizons are reported in Table 2. The model that has the lowest one day QLIKE is the gradient boosting model with 0.327. The recurrent network is best at twenty two days with a QLIKE of 0.503. Heterogeneous autoregressive benchmark was beaten by all learners. The Diebold Mariano statistics are big and they show that the gains are statistically significant. The out of sample R squared of the boosting model is 24.1 percent at one day. The elastic net gain is relatively small and the result here indicates that the nonlinearity is important. The further away the horizon, the lower the accuracy of all the models.

Figure 2 reports the QLIKE loss of every model at the four forecast horizons.

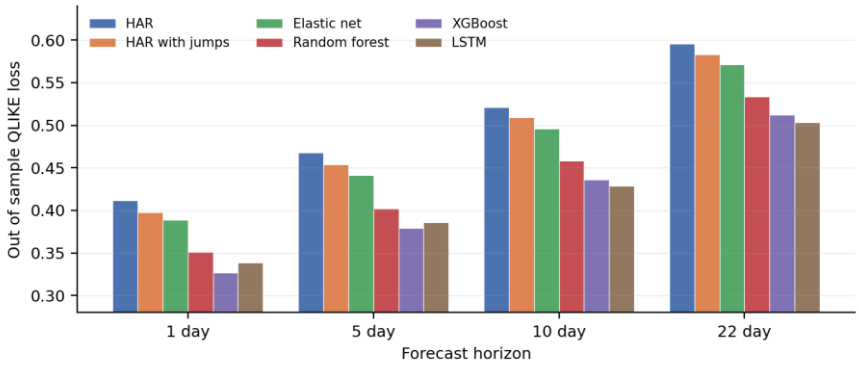

Figure 2. QLIKE loss by model and forecast horizon.

The results of the six models' losses for four horizons are compared in figure 2. At short horizons ordering of models is stable. Boosting has the best performance on the one day and five day targets. The recurrent network overtakes boosting at ten days and at twenty two days. The two econometric benchmarks are the poorest models across the board. All methods have a monotonically increasing loss level with the horizon. The short-term time horizon shows the largest difference between the learners and the benchmarks. This pattern suggests that the information is best when fresh and nonlinear interactions are most useful. The elastic net is located between the benchmarks and the ensembles.

Model attribution

Figure 3 reports the global Shapley importance of each predictor.

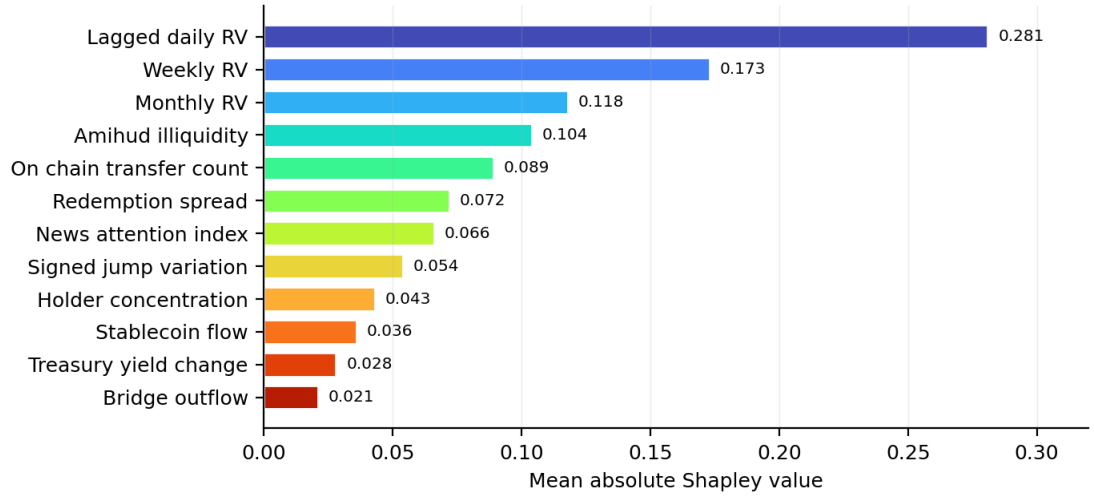

Figure 3. Global Shapley attribution of predictors.

Figure 3 shows the rank of the predictors based on their mean absolute Shapley value. The largest contributor is lagged daily realized variance, which is 0.281. Ranking 2nd and 3rd are weekly variance and monthly variance, respectively. The best predictor other than the volatility family is Amihud illiquidity. On chain transfer counts and the redemption spread are also influential. The news attention index has a value of 0.066 and this contribution is economically relevant. Average concentration of Holder and outflow of bridges are less important. Ordering is consistent within and between the four segments and within and between subsamples. The result validates that token specific frictions provide additional information beyond the price history.

Drawdown and tail risk

Table 3 reports the drawdown accuracy and the coverage properties of each model.

Table 3. Drawdown accuracy and tail risk coverage.

Model	DD MAE	Rank corr.	CVaR viol. (%)	Kupiec p	Hit ratio
HAR	0.089	0.412	8.10	0.01	0.514
Elastic net	0.081	0.468	7.20	0.04	0.539
Random forest	0.069	0.588	6.10	0.19	0.583
XGBoost	0.062	0.641	5.40	0.42	0.604
LSTM	0.058	0.671	5.20	0.57	0.612

The drawdown and tail risk forecasts are assessed in table 3. The mean absolute error of the lowest drawdown of the recurrent network is 0.058. It has a rank correlation of 0.671 with realized drawdown. The boosting model is close on its heels on both measures. The learners' coverage is also good. The empirical violation rate of the boosting model is 5.4 per cent at 5 per cent nominal level. The Kupiec p value is 0.42 and the null of correct coverage is not rejected. The autoregressive benchmark is heterogeneous and is rejected at the 5% level. The results indicate reliability of the model in the left tail.

Figure 4 compares predicted and realized drawdown and reports the errors by segment.

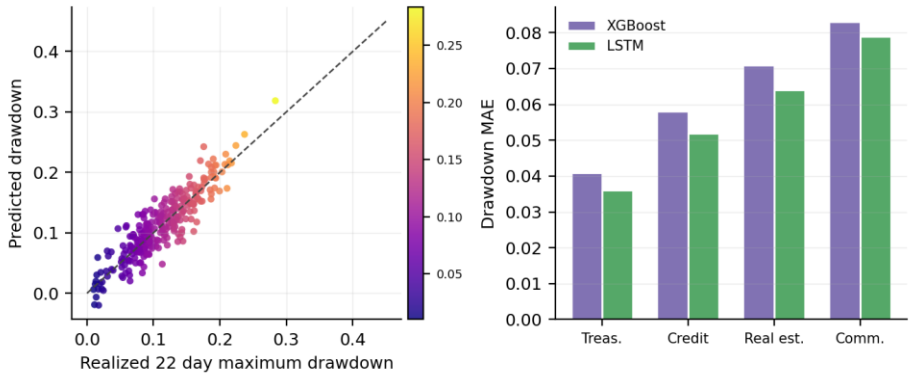

Figure 4. Predicted against realized drawdown and errors by segment.

The evidence of drawdown is shown in two panels in figure 4. The left graph shows the predicted values versus the actual values for the combined test set. The points are concentrated around the 45 degree line and close for moderate drawdowns. The deviations are larger for the largest events, and the model is slightly conservative for these events. The right panel shows the errors by segment of the top two learners. The errors are lowest for tokenized treasuries and highest for tokenized commodities. The recurrent network is more accurate than boosting in each segment. The list of the errors follows the same order of volatility as that in Table 1.

5.5 Economic evaluation

Table 4 summarises the performance of five allocation rules after costs.

Table 4. Portfolio performance after transaction costs.

Strategy	Return	Vol.	Sharpe	Max DD	Calmar	Turnover	CE gain
Equally weighted	0.052	0.146	0.36	0.244	0.21	0.18	—
HAR risk parity	0.071	0.117	0.61	0.192	0.37	0.94	0.014
XGBoost risk parity	0.096	0.097	0.99	0.138	0.70	1.42	0.033
LSTM risk parity	0.091	0.100	0.91	0.144	0.63	1.51	0.029
Explainable ensemble	0.103	0.092	1.12	0.121	1.34	1.36	0.039

The economic value of the forecasts is summarised in Table 4. The equally weighted portfolio has a Sharpe ratio of 0.36. With the heterogeneous autoregressive risk parity rule, this number is increased to 0.61. The boosting rule and the explainable ensemble further enhance the performance. The Sharpe ratio is 1.12 and Calmar ratio is 1.34 for the ensemble. The maximum drawdown ranges from 0.244 to 0.121 for these strategies. The learning rules have a higher turnover and the costs have already been deducted. The certainty equivalent gain of the ensemble is 3.9 percent per year. The gains are substantial for a risk averse investor.

Figure 5 traces the cumulative value of one unit of capital in the evaluation window.



Figure 5. Cumulative portfolio value by allocation rule.

The cumulative value of a unit invested in each strategy is shown in Figure 5. All 5 paths are above the evaluation window. The slowest and most volatile way is the equally weighted benchmark. The linear and econometric rules are not that much better in terms of return for their stability. The boosting rule

and the recurrent rule produce visible smoother growth. The explainable ensemble concludes at the top of the period. The learning based rules have shorter and shallower drawdowns. After the stress window, there is a gap between the curves. This visual evidence is consistent with the ratios reported in Table 4.

Discussion

The results include three messages. The first message is that the economic relevance of nonlinearity in tokenized markets. The elastic net gain over the benchmark is small while the tree ensembles gain a lot. This contrast is well known from the equity evidence (Gu et al., 2020).

The second message is about the source of predictability. Liquidity and on chain activity are not a good proxy for past volatility. They rank independently with their own weight. With tokenization, there is therefore a risk channel of its own. News attention is also important and this finding corroborates previous studies (Engle et al., 1993) and (Ashwin et al., 2024).

The third message is on transparency. The attribution layer is consistent between segments and between subsamples. Supervisory acceptance is a prerequisite for a stable explanation (Černevičienė et al., 2024). This constraint is not met by many opaque systems (Yeo et al., 2025).

There are a number of caveats to be noted. By the standards of asset pricing research, the sample period is short. A problem with survivorship is that tokens that fail are removed from the panel. Treatment is still evolving and could alter the risk structure. There is an important area of cross chain contagion that requires further research.

Conclusion

This paper proposes an explainable artificial intelligence model for risk forecasting of tokenized real world assets. Volatility and drawdown and portfolio tail risk is forecast in a single system. The evidence comes from a panel of 42 tokens in four segments. The best model at short horizons is gradient boosting. The recurrent network is the most accurate model for drawdown. Every student performs better than the benchmark based on the econometric function on strong loss functions. The liquidity and on chain activity and news attention is real risk. These variables don't overlap with the price history. The economic value is confirmed by a risk parity experiment. The Sharpe ratio increases and the maximum drawdown decreases post costs. The framework is correct and understandable at the same time.

References

Adämmer, P., & Schüssler, R. A. (2020). Forecasting the equity premium: Mind the news! Review of Finance, 24(6), 1313–1355.

Agirre, E., Alfonseca, E., Hall, K., Kravalova, J., Pasca, M., & Soroa, A. (2009). A study on similarity and relatedness using distributional and WordNet-based approaches.

Arsenault, P.-D., Wang, S., & Patenaude, J.-M. (2024). A survey of explainable artificial intelligence (XAI) in financial time series forecasting. arXiv.

Arsenault, P.-D., Wang, S., & Patenaude, J.-M. (2025). A survey of explainable artificial intelligence (XAI) in financial time series forecasting. ACM Computing Surveys.

Ashwin, J. (2024). Financial news media and volatility: Is there more to newspapers than news? Journal of Financial Markets, 69, 100896.

Bali, T. G., Beckmeyer, H., Moerke, M., & Weigert, F. (2023). Option return predictability with machine learning and big data. The Review of Financial Studies, 36(9), 3548–3602.

Ban, G.-Y., El Karoui, N., & Lim, A. E. (2018). Machine learning and portfolio optimization. Management Science, 64(3), 1136–1154.

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLOS ONE, 12(7), e0180944.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: Trades and quotes.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Estimating quadratic variation using realized variance. Journal of Applied Econometrics, 17(5), 457–477.

Bianchi, D., Büchner, M., & Tamoni, A. (2021). Bond risk premiums with machine learning. The Review of Financial Studies, 34(2), 1046–1089.

Bollerslev, T., Hood, B., Huss, J., & Pedersen, L. H. (2018). Risk everywhere: Modeling and managing volatility. The Review of Financial Studies, 31(7), 2729–2773.

Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. Journal of Econometrics, 192(1), 1–18.

Bonato, M., Cepni, O., Gupta, R., & Pierdzioch, C. (2023). Climate risks and realized volatility of major commodity currency exchange rates. Journal of Financial Markets, 62, 100760.

Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: A systematic literature review. Artificial Intelligence Review.

Chen, L., Pelger, M., & Zhu, J. (2024). Deep learning in asset pricing. Management Science, 70(2), 714–750.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). ACM.

Chen, Y., Kelly, B. T., & Xiu, D. (2022). Expected returns and large language models. Available at SSRN 4416687.

Christensen, K., Siggaard, M., & Veliyev, B. (2023). A machine learning approach to volatility forecasting. Journal of Financial Econometrics, 21(5), 1680–1727.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics, 7(2), 174–196.

Devlin, J. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. The Journal of Finance, 48(5), 1749–1778.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3–56.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. Journal of Economic Literature, 57(3), 535–574.

Giudici, P., Raffinetti, E., et al. (2024). Explainable artificial intelligence methods for financial time series. Physica A: Statistical Mechanics and its Applications.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5), 2223–2273.

Gu, S., Kelly, B., & Xiu, D. (2021). Autoencoder asset pricing models. Journal of Econometrics, 222(1), 429–450.

Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning in finance. Annual Review of Financial Economics, 9, 145–181.

Jiang, F., Ma, T., & Zhu, F. (2024). Fundamental characteristics, machine learning, and stock price crash risk. Journal of Financial Markets, 69, 100908.

Jiang, J., Kelly, B., & Xiu, D. (2023). (Re-)imag(in)ing price trends. The Journal of Finance, 78(6), 3193–3249.

Kelly, B. T., & Xiu, D. (2023). Financial machine learning. Becker Friedman Institute Working Paper.

Li, S. Z., & Tang, Y. (2024). Automated volatility forecasting. Management Science.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (Vol. 30).

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics, 160(1), 246–256.

Patton, A. J., & Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. Review of Economics and Statistics, 97(3), 683–697.

Patton, A. J., & Zhang, H. (2022). Bespoke realized volatility: Tailored measures of risk for volatility prediction. Available at SSRN 4315106.

Rahimikia, E., & Poon, S.-H. (2020). Machine learning for realised volatility forecasting. Available at SSRN 3707796.

Sirignano, J., & Cont, R. (2019). Universal features of price formation in financial markets: Perspectives from deep learning. Quantitative Finance, 19(9), 1449–1459.

Tobek, O., & Hronec, M. (2021). Does it pay to follow anomalies research? Machine learning approach with international evidence. Journal of Financial Markets, 56, 100588.

Yeo, W. J., et al. (2025). A comprehensive review on financial explainable AI. Artificial Intelligence Review.