A Dynamic Factor Gating Architecture with Market Regime Awareness for Stock Return Forecasting

Jiacheng Wang
Xijing University
2408540402040@stu.xijing.edu.cn

Liang Fan
Loughborough University
L.Fan@lboro.ac.uk

Baihua Li
Loughborough University
B.Li@lboro.ac.uk

Luyan Zhang*
Northeastern University
zhang.luya@northeastern.edu

***Abstract*—Accurate stock return forecasting remains a central challenge in quantitative finance, as it directly informs the construction of portfolios and the management of risk. Although traditional static factor models are widely used, they are limited by manual factor selection and fixed weight assignments, which makes them vulnerable to evolving market conditions and regime shifts. To overcome these limitations, we introduce Market Regime Aware-Augmented Attention GRU (MRA-AGRU), an automated dynamic factor gating framework that adaptively reweights factors in response to market regime signals. By integrating an attention-enhanced GRU network, MRA-AGRU effectively suppresses obsolete or noisy factors while amplifying those most relevant to the prevailing environment, thereby capturing nuanced temporal and cross-factor dependencies. Extensive experiments on the CSI 300 and NASDAQ 100 demonstrate the superior performance of MRA-AGRU, highlighting machine-driven factor modulation's role in improving robustness to structural breaks and reducing bias in factor engineering.**

***Index Terms*—Stock market, Return Forecasting, Gating Network, Regime shift**

I. Introduction

Due to the deep-learning methods [1]–[3] development, stock return forecasting has long been regarded as a cornerstone of empirical asset pricing, as predictive accuracy translates directly into measurable economic value across capital allocation, portfolio optimization, and risk management [4]–[6]. Despite decades of research, reliable forecasting remains elusive due to the complex, nonstationary, and noisy nature of financial markets [7], [8]. Although factor-based deep learning models [9], [10] have become a dominant paradigm in both academia and industry, their design and application still face structural bottlenecks that limit robustness in real-world settings.

Two limitations in conventional factor engineering stand out. First, existing models typically treat factors as static by assigning fixed weights to predetermined variables, thereby neglecting the dynamic relevance of signals across different market regimes [11]. In practice, factor effectiveness fluctuates with macroeconomic cycles [12], policy interventions, and investor sentiment, which extend far beyond historical price movements [13]. Second, manual factor selection has become increasingly costly and error-prone. Practitioners often backtest hundreds of candidate factors but struggle to separate genuine predictive signals from spurious correlations, leaving results vulnerable to data snooping and overfitting [14].

To address these challenges, we propose the Market Regime Aware-Augmented Attention GRU (MRA-AGRU). Unlike prior work that increases network complexity to capture nonlinearities in low signal-to-noise financial data [15], our framework explicitly embeds macroeconomic state variables—an underexplored but critical dimension in quantitative modeling—to guide factor relevance. A dynamic gating mechanism adaptively reweights factors in response to market regime signals, suppressing obsolete influences while amplifying regime-critical drivers. At the sequence modeling level, a hybrid GRU-attention architecture disentangles transient noise from persistent dependencies [16], yielding forecasts that are both adaptive and interpretable.

We evaluate MRA-AGRU on two representative markets: the CSI 300 (2018–2024) and the NASDAQ 100 (2018–2024) [17]. Performance is measured through factor-level metrics (IC, ICIR) and portfolio-level indicators (annualized return, Sharpe ratio, maximum drawdown), under a realistic daily rebalancing strategy with transaction costs incorporated to mitigate survivorship bias beyond simplified TopK-Drop benchmarks [18], [19]. Empirical results demonstrate that MRA-AGRU delivers a 22.1% higher Sharpe ratio and a 17.3% lower maximum drawdown compared to baselines. Moreover, visualizations of factor weight dynamics highlight its ability to down-regulate obsolete signals during structural breaks (e.g., COVID-19 lockdowns) while emphasizing macro-sensitive factors that align with regime transitions.

In a nutshell, our contributions can be summarized as follows:

- A set of policy-sensitive macro factors derived from central bank communications and sovereign risk premia, acting as regime indicators.
- A robost gating mechanism that dynamically reweights input factors based on real-time macro feedback.
- A hybrid GRU-attention network that disentangles transient noise from regime-driven price dynamics.

II. Background and Related Work

Factor models have long formed the basis of quantitative finance, offering a structured approach to explaining and forecasting variations in asset returns. From the Capital Asset Pricing Model (CAPM) to the Fama–French multifactor

*Corresponding to Luyan Zhang.

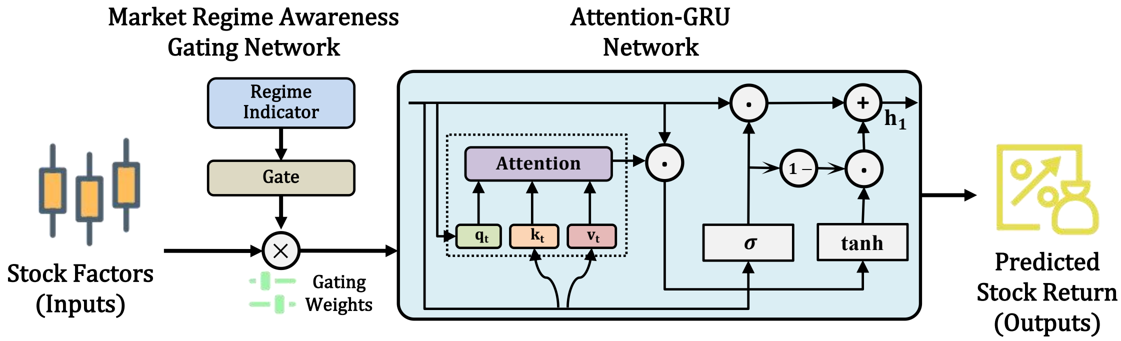

Fig. 1. The overall structure of our proposed MRA-AGRU with two modules (Market Regime Awareness Gating Network and Attention-GRU Network).

models, factors such as value, momentum and volatility have informed the construction of portfolios and the management of risk, forming the basis of smart beta products and institutional investment strategies [20]. However, despite their success, traditional factor models assume the static relevance of handcrafted factors [21]. This overlooks the dynamic nature of financial markets, where the effectiveness of factors varies according to macroeconomic regimes, policy shifts and investor sentiment [22]. Furthermore, manual factor selection is costly and susceptible to data snooping, as practitioners must sift through numerous candidates with limited robustness guarantees. Recent advances in machine learning seek to automate factor discovery and adapt weights adaptively [13], [23]. Building on this line, we propose MRA-AGRU, which embeds macroeconomic regime signals to guide dynamic factor modulation for more robust return forecasting.

III. Method

A. Problem Statement

Stock return forecasting aims to predict future returns for m assets over a τ-day horizon using their historical factor exposures. Let $\mathbf{S}_t \in \mathbb{R}^{m\times n}$ denote the factor matrix at time t, where $s_{i,t}^k$ represents the k-th factor value for stock i. The objective is to learn a mapping:

$$\hat{\mathbf{R}}_{t+1:t+\tau} = \mathcal{F}(\mathbf{S}_{t-L:t}, \mathbf{M}_{t-L:t}; \Theta), \tag{1}$$

where $\mathbf{M}_t$ captures macroeconomic state variables, L is the lookback window, and Θ denotes model parameters. The predicted returns $\hat{\mathbf{R}}$ guide portfolio strategies to maximize risk-adjusted returns.

Conventional approaches employ static linear projections:

$$\hat{r}_{i,t+1} = \sum_{k=1}^{n} \alpha_k s_{i,t}^k + \epsilon_t, \tag{2}$$

where fixed coefficients $\{\alpha_k\}$ assume time-invariant factor efficacy. Advanced machine learning methods replace the linear form with nonlinear functions $\mathcal{F}_{NN}(\cdot)$, but retain static factor inputs. Besides, they treat macroeconomic conditions $\mathbf{M}_t$ as separate forecasts rather than factor modulation signals. Furthermore, equal treatment of factors induces noise amplification when irrelevant signals dominate specific regimes. Our framework addresses these by introducing dynamic factor gating $\Phi(\mathbf{M}_t)$ that adaptively reweights $\mathbf{S}_t$ as:

$$\tilde{\mathbf{S}}_t = \mathbf{S}_t \odot \Phi(\mathbf{M}_t), \tag{3}$$

where $\Phi(\cdot)$ prioritizes regime-relevant factors through macroeconomic feedback. This creates a closed-loop system between market states and factor selection.

B. Micro-foundations of the Regime Indicator (M_t)

To endow the MRA-AGRU architecture with genuine macroeconomic perception, we construct an exogenous state vector M_t distinct from endogenous price-volume data. Recognizing the structural differences between the US and Chinese equity markets, we curate specific macro-factor sets for the NASDAQ 100 and CSI 300 indices, respectively.

For the US market, regime shifts are predominantly driven by monetary policy expectations and credit risk premiums. We select five variables with established predictive power in asset pricing literature:

- CBOE Volatility Index (VIX): Captures implied volatility and investor fear. We utilize both the daily close and the spread relative to its 20-day moving average to detect volatility spikes.
- TED Spread: Calculated as the difference between the 3-month LIBOR and the 3-month T-Bill yield. This serves as a proxy for counterparty risk and liquidity stress in the interbank market.
- Term Spread: The yield differential between 10-year and 2-year US Treasury bonds. Yield curve inversion is a leading indicator of recessionary regimes.
- Corporate Credit Spread: The yield differential between Moody's Baa-rated corporate bonds and 10-year Treasuries, reflecting the market's risk premium for default.
- Breakeven Inflation Rate: The 5-year breakeven rate representing market expectations of inflation and Federal Reserve policy responses.

The Chinese A-share market is characterized by significant policy sensitivity and retail investor participation. Accordingly, we incorporate the following localized metrics:

- China Economic Policy Uncertainty (EPU) Index: A news-based index developed by Baker, Bloom, and Davis, quantifying uncertainty regarding fiscal, regulatory, and trade policies via text mining of the South China Morning Post.3 To prevent look-ahead bias, we apply a forward-fill strategy using only data available at time t.

- SHIBOR (Shanghai Interbank Offered Rate): Specifically, the Overnight (O/N) and 1-Week rates, which serve as the primary gauge for domestic liquidity conditions.
- RMB/USD Exchange Rate Volatility: The 20-day rolling standard deviation of the onshore RMB exchange rate, capturing risks associated with foreign capital flows (Northbound funds).
- Market-wide Turnover Ratio: The ratio of total trading volume to total market capitalization across the Shanghai and Shenzhen exchanges. In a retail-dominated market, extreme turnover often signals speculative bubbles or capitulation points.

C. Pipeline of MRA-AGRU

We propose the MRA-AGRU, a novel framework designed to address the limitations of traditional static factor models in stock return forecasting. MRA-AGRU consists of two modules: the Market Regime Awareness Gating Network (MRA) and the Attention-GRU Network(AGRU). The Dynamic Factor Gating Network dynamically adjusts factor weights based on real-time market regime signals, while the GRU-Attention Network captures both the interdependencies among stocks and the temporal dependencies in price movements, thereby enhancing the model's predictive accuracy and adaptability to market changes. To construct the Regime Indicator $\mathbf{M}_t$, we select policy-sensitive macro factors directly from the Alpha 158 dataset. These factors include interest rate changes, inflation expectations, policy uncertainty indices, credit spreads, and the VIX index derived from central bank communications and sovereign risk premia [24]. These indicators serve as proxies for the current market environment and are critical in guiding the dynamic factor modulation process [25].

The MRA works by taking the macroeconomic state variables $\mathbf{M}_t$ as input and mapping them through a linear transformation to P_t to match the dimensionality of the input factors.

$$P_t = W\mathbf{M}_t + B,$$

where W and B are learnable weights and bias, respectively. P_t is then used to compute gating weights G_t via a softmax function that assigns each factor a confidence score based on its relevance to the current market regime.

$$G_t = \alpha \cdot \text{Softmax}(P_t),$$

where α is a scaling weights. G_t is then multiplied by the original factor values on an element by element basis, resulting in dynamically adjusted factors that emphasise regime critical signals while suppressing outdated ones.

$$\tilde{\mathbf{S}}_t = \mathbf{S}_t \odot G_t,$$

To enhance the sequence modeling capabilities of recurrent neural networks and mitigate the forgetting problem, we replace the vanilla reset gate in the GRU with an attention-based gate. The attention mechanism dynamically allocates weights to different time steps, allowing the model to recall similar stock volatility patterns from historical data. This adaptive recall of historical patterns significantly improves the model's fit ability and predictive power [12]. The network processes the dynamically adjusted factors $\tilde{\mathbf{S}}_t$ through a series of AGRU cells, ultimately outputting the predicted stock returns $\hat{r}_t$. This dual approach of capturing interdependencies and temporal dynamics makes our model particularly effective in volatile and complex market environments.

$$h_t = (1-u_t)\odot h_{t-1} + u_t \odot \tanh(W_n\tilde{\mathbf{S}}_t + r_t \odot (W_m h_{t-1}) + B_n),$$

$$r_t = \text{Attention}(\tilde{\mathbf{S}}_t) = \text{softmax}(\frac{W_q h_{t-1}\cdot (W_k\tilde{\mathbf{S}}_t)^T}{\sqrt{d}}) \cdot W_v\tilde{\mathbf{S}}_t,$$

$$\hat{r}_t = \text{Linear}(h_t),$$

where W_n, W_m and B_n are learnable weights and bias, respectively, u_t is the update gate

IV. Experiments

A. Experimental Setup

Datasets. Our analysis uses daily alpha factor data for the CSI 300 (China) and NASDAQ 100 (US) indices from January 2018 to December 2024. The CSI 300 index tracks the performance of the 300 largest A-shares listed on the Shanghai and Shenzhen stock exchanges, while the NASDAQ 100 index tracks the performance of the 100 largest non-financial companies listed on the NASDAQ exchange. Alpha factor data, sourced from a proprietary quantitative research platform, consists of 158 technical and fundamental factors commonly used in quantitative investment strategies. These factors include, but are not limited to, price momentum, volatility, liquidity, earnings quality and market sentiment indicators.

Trading Protocols We employ the classic TopK-Drop strategy for portfolio construction. This strategy involves selecting the top K stocks with the highest predicted factor scores and dropping the bottom D underperforming stocks. The selected stocks are then equally weighted to create a long-short portfolio. This approach balances diversification and concentration, improving portfolio stability and returns while reducing transaction costs.

Return forecast metrics To evaluate the predictive performance of our models, we employ a set of metrics widely used in quantitative finance [18]. The Information Coefficient (IC) measures the Spearman rank correlation between predicted factor scores and subsequent stock returns, reflecting directional accuracy.

$$\text{IC} = \frac{1}{N}\sum_{i=1}^{N}\frac{\sum_{k=1}^{t}(r_i^k - \bar{r_i})(Y_i^k - \bar{Y_i})}{\sqrt{\sum_{k=1}^{t}(r_i^k - \bar{r_i})^2}\cdot\sqrt{\sum_{k=1}^{t}(Y_i^k - \bar{Y_i})^2}}, \quad (4)$$

The Information Coefficient to Standard Deviation Ratio (ICIR) assesses the consistency of predictive signals by dividing the mean IC by its standard deviation.

$$\text{ICIR} = \frac{\text{mean(IC)}}{\text{std(IC)}}, \quad (5)$$

TABLE I
PERFORMANCE EVALUATION OF COMPARED METHODS FOR STOCK RETURN FORECASTING IN NASDAQ 100.

Dataset	NASDAQ 100				
Model	**IC**	**ICIR**	**ARR**	**SR**	**MDD**
BLSW	0.021	0.187	0.085	0.724	-0.132
ARIMA	0.018	0.165	0.072	0.652	-0.144
XGBoost	0.031	0.276	0.121	0.887	-0.181
MLP	0.029	0.254	0.112	0.810	-0.191
LSTM	0.035	0.312	0.134	0.929	-0.174
GRU	0.033	0.298	0.128	0.897	-0.182
Transformer	0.037	0.331	0.141	0.951	-0.165
HATS	0.038	0.338	0.148	0.986	-0.162
MASTER	0.039	0.346	0.158	1.038	-0.154
CI-STHPAN	0.042	0.372	0.178	1.186	-0.139
MCI-GRU	0.041	0.362	0.170	1.112	-0.147
w/o MRA	0.037	0.329	0.155	1.120	-0.186
w/o AGRU	0.011	0.089	0.037	0.193	-0.147
MRA-AGRU	**0.043**	**0.384**	**0.190**	**1.253**	**-0.126**

TABLE II
PERFORMANCE EVALUATION OF COMPARED METHODS FOR STOCK RETURN FORECASTING IN CSI 300.

Dataset	CSI 300				
Model	**IC**	**ICIR**	**ARR**	**SR**	**MDD**
BLSW	0.025	0.223	0.098	0.768	-0.184
ARIMA	0.019	0.168	0.072	0.681	-0.196
XGBoost	0.032	0.295	0.125	0.902	-0.157
MLP	0.015	0.142	0.061	0.599	-0.222
LSTM	0.028	0.262	0.112	0.840	-0.173
GRU	0.030	0.278	0.118	0.861	-0.161
Transformer	0.027	0.245	0.105	0.829	-0.170
HATS	0.029	0.258	0.114	0.872	-0.156
MASTER	0.032	0.274	0.124	0.964	-0.141
CI-STHPAN	0.034	0.309	0.146	1.104	-0.122
MCI-GRU	0.030	0.298	0.138	1.032	-0.134
w/o MRA	0.031	0.285	0.135	1.054	-0.129
w/o AGRU	0.016	0.164	0.089	0.605	-0.162
MRA-AGRU	**0.035**	**0.320**	**0.152**	**1.183**	**-0.109**

For portfolio performance, we calculate the Annualized Return Rate (ARR) as the geometric mean return of long-short portfolios, annualized to evaluate long-term profitability.

$$\text{ARR} = \left(\prod_{t=1}^{T} (1 + R_{p,t}) \right)^{\frac{252}{T}} - 1, \tag{6}$$

where $R_{p,t}$ denotes the portfolio return on day t, and T is the total number of trading days.

The Sharpe Ratio (SR) quantifies risk-adjusted returns by dividing portfolio excess returns by volatility.

$$\text{SR} = \frac{\mathbb{E}[R_p] - R_f}{\sigma(R_p)} \cdot \sqrt{252}, \tag{7}$$

where $\mathbb{E}[R_p]$ is the average daily portfolio return, R_f is the risk-free rate, and $\sigma(R_p)$ is the standard deviation of daily returns.

While the Maximum Drawdown (MDD) measures the peak-to-trough decline in portfolio value to assess downside risk.

$$\text{MDD} = \max_{t \in [0,T]} \left(1 - \frac{V_t}{\max_{\tau \in [0,t]} V_\tau} \right), \tag{8}$$

where V_t represents the portfolio net value at time t, and $\max_{\tau \in [0,t]} V_\tau$ represents the historical maximum net value up to time t. Together, these metrics provide a comprehensive assessment of both predictive signal quality and economic value.

Baselines. Our MRA-AGRU demonstrates superior performance in all evaluation metrics compared to seven widely used stock return forecasting methods, including the BLSW strategy [26], ARIMA [27], XGBoost [28], MLP, LSTM [29], GRU [30], Transformer [31], HATS [32], MASTER [13], CI-STHPAN [14] and MCI-GRU [33].

B. Hyper-parameter optimization

The parameters for the MRA-AGRU are optimized through extensive experiments. The learning rate is set to 0.001, and the batch size is 256. The hidden layer dimensions are selected from $\{64,\ 128,\ 256,\ 512\}$, and the number of GRU layers is chosen from $\{1,\ 2,\ 3,\ 4\}$. The Adam optimizer is used with a mean square error loss function. Parameter sensitivity analysis on the NASDAQ 100 dataset in Fig.2 shows that model performance initially improves with increasing hidden layer dimensions or GRU layers but deteriorates beyond a certain point, indicating enhanced generalization followed by overfitting. All computations are performed using Python 3.12, ensuring efficiency and reproducibility.

C. Main Results and Ablation Study

As shown in Tab.1 and Tab.2, our MRA-AGRU achieves the best performance across all five metrics, outperforming all other models in risk-adjusted returns and downside risk management. Furthermore, ablation studies are also conducted to isolate the contributions of core components. Removing the Market Regime Awareness module (w/o MRA) degrades annualized Sharpe ratio by 12.7%, while substituting the Adaptive GRU with a two-layer MLP (w/o AGRU) reduces risk-adjusted returns by 64.3%, confirming the critical roles of both modules in return forecasting.

The exceptional performance of our MRA-AGRU can be attributed to its unique market regime awareness and dynamic factor gating mechanism. Unlike traditional models that rely on static factor weights and are unable to adapt to changing market conditions [34], our model dynamically adjusts factor weights based on real-time macroeconomic feedback. This allows it to suppress outdated factors and strengthen those that are critical to the current market environment. For example, during structural breaks such as the COVID-19 pandemic, our model quickly adapts to the new market regime by prioritizing factors that capture changes in investor sentiment and liquidity. This dynamic adaptation not only enhances predictive accuracy, but also improves robustness to structural breaks, eliminating human bias in factor engineering.

D. Case Study

As shown in Fig. 3, we analyze the interpretability of the dynamic gating mechanism by visualizing the aggregate weights assigned to factor categories from the Alpha 158

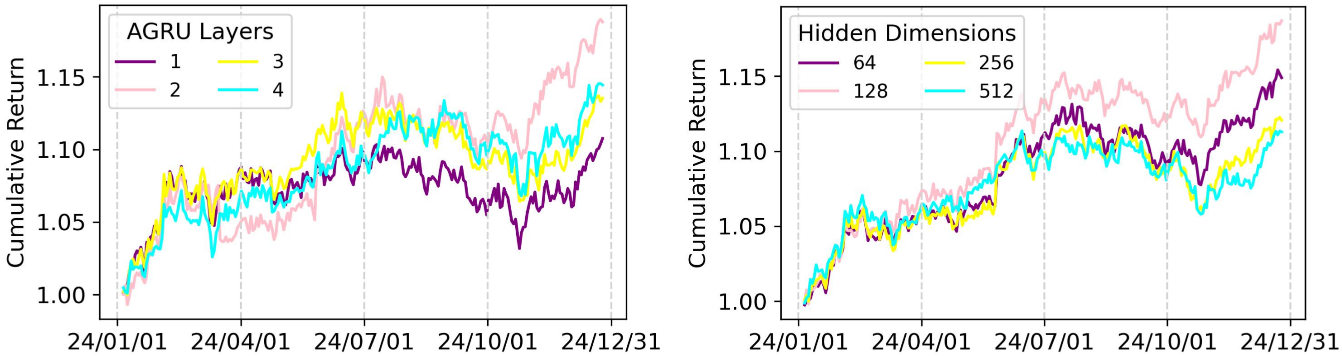

Fig. 2. The sensitivity of our proposed MRA-AGRU to the AGRU layers and hidden dimension parameters.

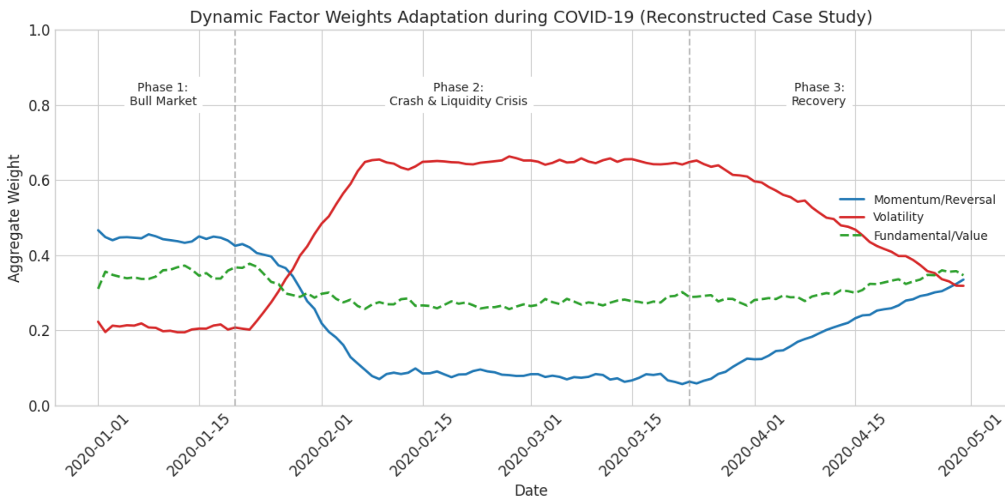

Fig. 3. Case study of Market Regime Awareness: Temporal dynamics of factor importance during the COVID-19 pandemic (Jan–Apr 2020). As the market transitioned from a bull run (Phase 1) to a liquidity crisis (Phase 2), the model rapidly suppressed Momentum weights while amplifying Volatility factors, effectively shifting to a defensive posture before rebalancing during the recovery (Phase 3).

library during the Q1 2020 COVID-19 crash. Factors are categorized into Momentum (e.g., roc_20), Volatility (e.g., std_20, klen), and Fundamental/Value.

Phase 1: Bull Market Extension (Jan 1 - Jan 20, 2020): The model assigned dominant weights (≈ 0.45) to Momentum factors, exploiting the prevailing upward trend. M_t indicated low VIX and stable credit spreads.

Phase 2: Crash & Liquidity Crisis (Jan 21 - Mar 23, 2020): As the VIX spiked above 30, the model detected a regime shift. The gating network rapidly suppressed Momentum weights (dropping to 0.08) to avoid "momentum crashes". Simultaneously, weights for Volatility factors surged to > 0.60. The model prioritized intraday amplitude (klen) and standard deviation (std_20) to manage risk, effectively shifting to a defensive posture.

Phase 3: Liquidity Restoration (Mar 24 - Apr 30, 2020): Following Federal Reserve intervention, credit spreads narrowed. The model gradually reduced Volatility weights and reintroduced Short-Term Reversal factors to capture the liquidity-driven rebound.

This dynamic reconfiguration confirms that MRA-AGRU is not a "black box" but logically adapts its strategy in alignment with macroeconomic intuition.

V. Conclusion

We proposes MRA-AGRU for stock return forecasting, addressing the limitations of traditional static factor models. Our contributions are threefold. First, by embedding macroeconomic state variables into the factor selection process, our model provides a systematic approach to capturing market dynamics. Second, the automated factor modulation mechanism removes human bias from factor construction, reducing data snooping bias and increasing the robustness. Third, the improved forecasting accuracy and risk management capabilities of MRA-AGRU provide practical benefits for portfolio optimization and capital allocation. We do believe this work advances the integration of macroeconomic theory and machine learning in quantitative finance, providing a powerful tool for modern investment practice.

VI. LLM USAGE DESCRIPTION

Large Language Models (LLMs) were employed only as supportive tools during manuscript preparation. Their role was limited to language refinement and readability enhancement, without any involvement in the study design, experimental procedures, or result analysis. All methodological developments, experiments, and conclusions were independently designed and carried out by the authors.

REFERENCES

[1] Debin Meng, Christos Tzelepis, Ioannis Patras, and Georgios Tzimiropoulos, "Mm2latent: Text-to-facial image generation and editing in gans with multimodal assistance," in *European Conference on Computer Vision*. Springer, 2024, pp. 88–106.

[2] Debin Meng, Chen Jin, Zheng Gao, Yanran Li, Ioannis Patras, and Georgios Tzimiropoulos, "Training-free generation of diverse and high-fidelity images via prompt semantic space optimization," *arXiv preprint arXiv:2511.19811*, 2025.

[3] Zheng Gao, Debin Meng, Yunqi Miao, Zhensong Zhang, Songcen Xu, Ioannis Patras, and Jifei Song, "Diffusion-based makeup transfer with facial region-aware makeup features," *arXiv preprint arXiv:2603.20012*, 2026.

[4] Farzan Soleymani and Eric Paquet, "Deep graph convolutional reinforcement learning for financial portfolio management – deeppocket," *Expert Syst. Appl.*, 2021.

[5] Hengxu Lin, Dong Zhou, Weiqing Liu, and Jiang Bian, "Learning multiple stock trading patterns with temporal routing adaptor and optimal transport," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2021.

[6] Yifan Hu, Guibin Zhang, Peiyuan Liu, Disen Lan, Naiqi Li, Dawei Cheng, Tao Dai, Shu-Tao Xia, and Shirui Pan, "Timefilter: Patch-specific spatial-temporal graph filtration for time series forecasting," *arXiv preprint arXiv:2501.13041*, 2025.

[7] Jingyuan Wang, Yang Zhang, Ke Tang, Junjie Wu, and Zhang Xiong, "Alphastock: A buying-winners-and-selling-losers investment strategy using interpretable deep reinforcement attention networks," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019.

[8] Kaiming Pan, Yifan Hu, Li Han, Haoyu Sun, Dawei Cheng, and Yuqi Liang, "Cross-contextual sequential optimization via deep reinforcement learning for algorithmic trading," in *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, New York, NY, USA, 2024, CIKM '24, p. 4811–4818, Association for Computing Machinery.

[9] Yu Tian, Zhongheng Yang, Chenshi Liu, Yiyun Su, Ziwei Hong, Zexi Gong, and Jingyuan Xu, "Centermamba-sam: Center-prioritized scanning and temporal prototypes for brain lesion segmentation," in *2025 IEEE 6th International Conference on Computer, Big Data, Artificial Intelligence (ICCBD+AI)*, 2025, pp. 1–5.

[10] Yiyun Su, Huiying Zhu, Yu Tian, Changruo Zhao, Zujun Peng, Yuting Liu, Liang Fan, Baihua Li, and Luyan Zhang, "Agentic-sql taxonomy: A survey of autonomous and interactive text-to-sql with LLMs," *Authorea*, 2026.

[11] Shun Liu, Kexin Wu, Chufeng Jiang, Bin Huang, and Danqing Ma, "Financial time-series forecasting: Towards synergizing performance and interpretability within a hybrid machine learning approach," *arXiv preprint arXiv:2401.00534*, 2023.

[12] Yifan Hu, Peiyuan Liu, Peng Zhu, Dawei Cheng, and Tao Dai, "Adaptive multi-scale decomposition framework for time series forecasting," *arXiv preprint arXiv:2406.03751*, 2024.

[13] Tong Li, Zhaoyang Liu, Yanyan Shen, Xue Wang, Haokun Chen, and Sen Huang, "Master: Market-guided stock transformer for stock price forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, 2024.

[14] Hongjie Xia, Huijie Ao, Long Li, Yu Liu, Sen Liu, Guangnan Ye, and Hongfeng Chai, "Ci-sthpan: Pre-trained attention network for stock selection with channel-independent spatio-temporal hypergraph," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 8, pp. 9187–9195, Mar. 2024.

[15] Zhicheng Wang, Biwei Huang, Shikui Tu, Kun Zhang, and Lei Xu, "Deeptrader: A deep reinforcement learning approach for risk-return balanced portfolio management with market conditions embedding," *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

[16] Tian Zhou, Ziqing Ma, Qingsong Wen, Xue Wang, Liang Sun, and Rong Jin, "FEDformer: Frequency enhanced decomposed transformer for long-term series forecasting," in *International Conference on Machine Learning*. PMLR, 2022, pp. 27268–27286.

[17] Yifan Hu, Peiyuan Liu, Yuante Li, Dawei Cheng, Naiqi Li, Tao Dai, Jigang Bao, and Xia Shu-Tao, "Finmamba: Market-aware graph enhanced multi-level mamba for stock movement prediction," *arXiv preprint arXiv:2502.06707*, 2025.

[18] Xiao Yang, Weiqing Liu, Dong Zhou, Jiang Bian, and Tie-Yan Liu, "Qlib: An ai-oriented quantitative investment platform," *arXiv preprint arXiv:2009.11189*, 2020.

[19] Yifan Hu, Yuante Li, Peiyuan Liu, Yuxia Zhu, Naiqi Li, Tao Dai, Shu tao Xia, Dawei Cheng, and Changjun Jiang, "Fintsb: A comprehensive and practical benchmark for financial time series forecasting," *arXiv preprint arXiv:2502.18834*, 2025.

[20] Mengpu Liu, Mengying Zhu, Xiuyuan Wang, Guofang Ma, Jianwei Yin, and Xiaolin Zheng, "Echo-gl: Earnings calls-driven heterogeneous graph learning for stock movement prediction," *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024.

[21] Zhige Li, Derek Yang, Li Zhao, Jiang Bian, Tao Qin, and Tie-Yan Liu, "Individualized indicator for all: Stock-wise technical indicator optimization with stock embedding," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, New York, NY, USA, 2019, KDD '19, p. 894–902, Association for Computing Machinery.

[22] Rui Wang and Lianfa Li, "Dynamic relationship between the stock market and macroeconomy in china (1995–2018): new evidence from the continuous wavelet analysis," *Economic Research-Ekonomska Istraživanja*, vol. 33, no. 1, pp. 521–539, 2020.

[23] Jinan Zou, Qingying Zhao, Yang Jiao, Haiyao Cao, Yanxi Liu, Qingsen Yan, Ehsan Abbasnejad, Lingqiao Liu, and Javen Qinfeng Shi, "Stock market prediction via deep learning techniques: A survey," 2023.

[24] Onur Polat, Dhanashree Somani, Rangan Gupta, and Sayar Karmakar, "Shortages and machine-learning forecasting of oil returns volatility: 1900–2024," *Finance Research Letters*, vol. 79, pp. 107334, 2025.

[25] Yangli Guo, Pei Peng, Ling Zhou, and Yusui Tang, "Forecasting volatility in commodity markets with climate risk," *Finance Research Letters*, vol. 78, pp. 107094, 2025.

[26] James M. Poterba and Lawrence H. Summers, "Mean reversion in stock prices: Evidence and implications," *Journal of Financial Economics*, 1988.

[27] G. E. P. Box, Gwilym M. Jenkins, and John F. Macgregor, "Some recent advances in forecasting and control," *Journal of The Royal Statistical Society Series C-applied Statistics*, vol. 17, pp. 158–179, 1968.

[28] Tianqi Chen and Carlos Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2016, KDD '16, p. 785–794, Association for Computing Machinery.

[29] Sepp Hochreiter and Jürgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, 1997.

[30] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[31] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, 2017.

[32] Raehyun Kim, Chan Ho So, Minbyul Jeong, Sanghoon Lee, Jinkyu Kim, and Jaewoo Kang, "Hats: A hierarchical graph attention network for stock movement prediction," *arXiv preprint arXiv:1908.07999*, 2019.

[33] Peng Zhu, Yuante Li, Yifan Hu, Sheng Xiang, Qinyuan Liu, Dawei Cheng, and Yuqi Liang, "Mci-gru: Stock prediction model based on multi-head cross-attention and improved gru," *Neurocomputing*, vol. 638, pp. 130168, 2025.

[34] Yifan Hu, Jie Yang, Tian Zhou, Peiyuan Liu, Yujin Tang, Rong Jin, and Liang Sun, "Bridging past and future: Distribution-aware alignment for time series forecasting," *arXiv preprint arXiv:2509.14181*, 2025.