

Recent Advancements in AI for Finance

A Computer Science Perspective

Structure

Predictive Tasks

- Continuous Price Prediction
- Binary Trend / Movement Classification
- Ranking-based Stock Selection

Decision Making Tasks

- Portfolio Optimisation
- Quantitative Trading

Others

- LLMs for Finance

Full list of papers

<https://github.com/junhua/awesome-finance-ai-papers>

Predictive Task Formulations

Continuous Price Prediction

- Predicting the closing prices
- Change in closing prices (i.e. return)
- Volatility (i.e. dispersion of returns / log return)

Binary Trend/Movement Classification

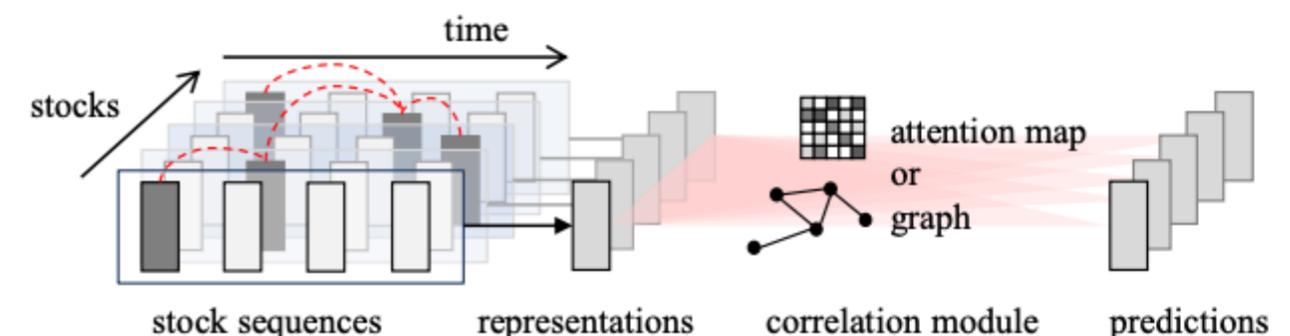
- Binary classification on price movement from date t to future $t+w$

Ranking-based Stock Selection

- Stock ranking by return ratio

Feature representation (Encoder)

- Pre-trained embedding
- Graph modelling



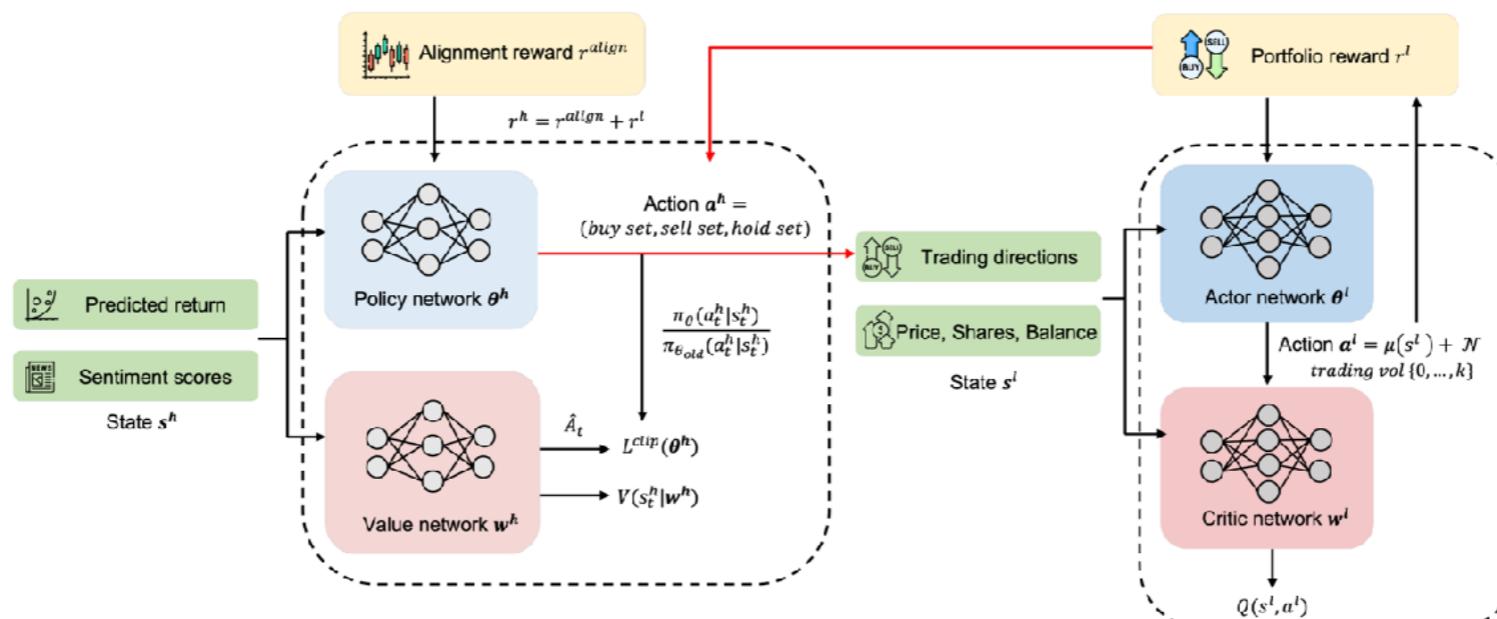
Decision-Making Formulations

Portfolio Optimization

- Formulation: MDP, POMDP, HMDP, (S)HMM
- Goal: assets allocation (optimise portfolio weights)

Quant Trading

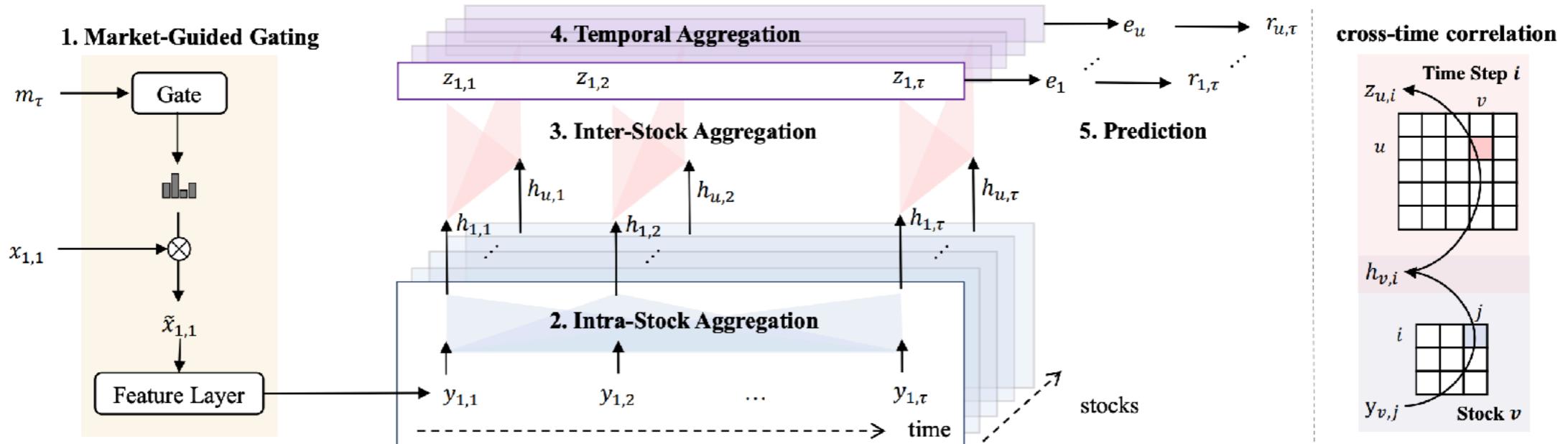
- Goal: trade execution (stock, action, price, amount...)



Continuous Price Prediction

MASTER: Market-Guided Stock Transformer for Stock Price Forecasting, AAAI'24

Goal: Improve stock price forecasting by modeling momentary cross-time stock correlations and dynamic feature selection.



Contributions:

- (1) Propose Market-Guided Stock Transformer (MASTER) that captures realistic stock correlations through alternating intra-stock and inter-stock aggregation
- (2) Leverage market-guided gating mechanism to automatically select relevant features based on market conditions
- (3) Show superior performance through experiments in both ranking and portfolio-based metrics

Dataset	Model	IC	ICIR	RankIC	RankICIR	AR	IR
CSI300	XGBoost	0.051 ± 0.001	0.37 ± 0.01	0.050 ± 0.001	0.36 ± 0.01	0.23 ± 0.03	1.9 ± 0.3
	LSTM	0.049 ± 0.001	0.41 ± 0.01	0.051 ± 0.002	0.41 ± 0.03	0.20 ± 0.04	2.0 ± 0.4
	GRU	0.052 ± 0.004	0.35 ± 0.04	0.052 ± 0.005	0.34 ± 0.04	0.19 ± 0.04	1.5 ± 0.3
	TCN	0.050 ± 0.002	0.33 ± 0.04	0.049 ± 0.002	0.31 ± 0.04	0.18 ± 0.05	1.4 ± 0.5
	Transformer	0.047 ± 0.007	0.39 ± 0.04	0.051 ± 0.002	0.42 ± 0.04	0.22 ± 0.06	2.0 ± 0.4
	GAT	0.054 ± 0.002	0.36 ± 0.02	0.041 ± 0.002	0.25 ± 0.02	0.19 ± 0.03	1.3 ± 0.3
	DTML	0.049 ± 0.006	0.33 ± 0.04	0.052 ± 0.005	0.33 ± 0.04	0.21 ± 0.03	1.7 ± 0.3
CSI800	MASTER	$0.064^* \pm 0.006$	0.42 ± 0.04	$0.076^* \pm 0.005$	0.49 ± 0.04	0.27 ± 0.05	2.4 ± 0.4
	XGBoost	0.040 ± 0.000	0.37 ± 0.01	0.047 ± 0.000	0.42 ± 0.01	0.08 ± 0.02	0.6 ± 0.2
	LSTM	0.028 ± 0.002	0.32 ± 0.02	0.039 ± 0.002	0.41 ± 0.03	0.09 ± 0.02	0.9 ± 0.2
	GRU	0.039 ± 0.002	0.36 ± 0.05	0.044 ± 0.003	0.39 ± 0.07	0.07 ± 0.04	0.6 ± 0.3
	TCN	0.038 ± 0.002	0.33 ± 0.04	0.045 ± 0.002	0.38 ± 0.05	0.05 ± 0.04	0.4 ± 0.3
	Transformer	0.040 ± 0.003	0.43 ± 0.03	0.048 ± 0.003	0.51 ± 0.05	0.13 ± 0.04	1.1 ± 0.3
	GAT	0.043 ± 0.002	0.39 ± 0.02	0.042 ± 0.002	0.35 ± 0.02	0.10 ± 0.04	0.7 ± 0.3
CSI300	DTML	0.039 ± 0.004	0.29 ± 0.03	0.053 ± 0.008	0.37 ± 0.06	0.16 ± 0.03	1.3 ± 0.2
	MASTER	$0.052^* \pm 0.006$	0.40 ± 0.06	0.066 ± 0.007	0.48 ± 0.06	$0.28^* \pm 0.02$	$2.3^* \pm 0.3$

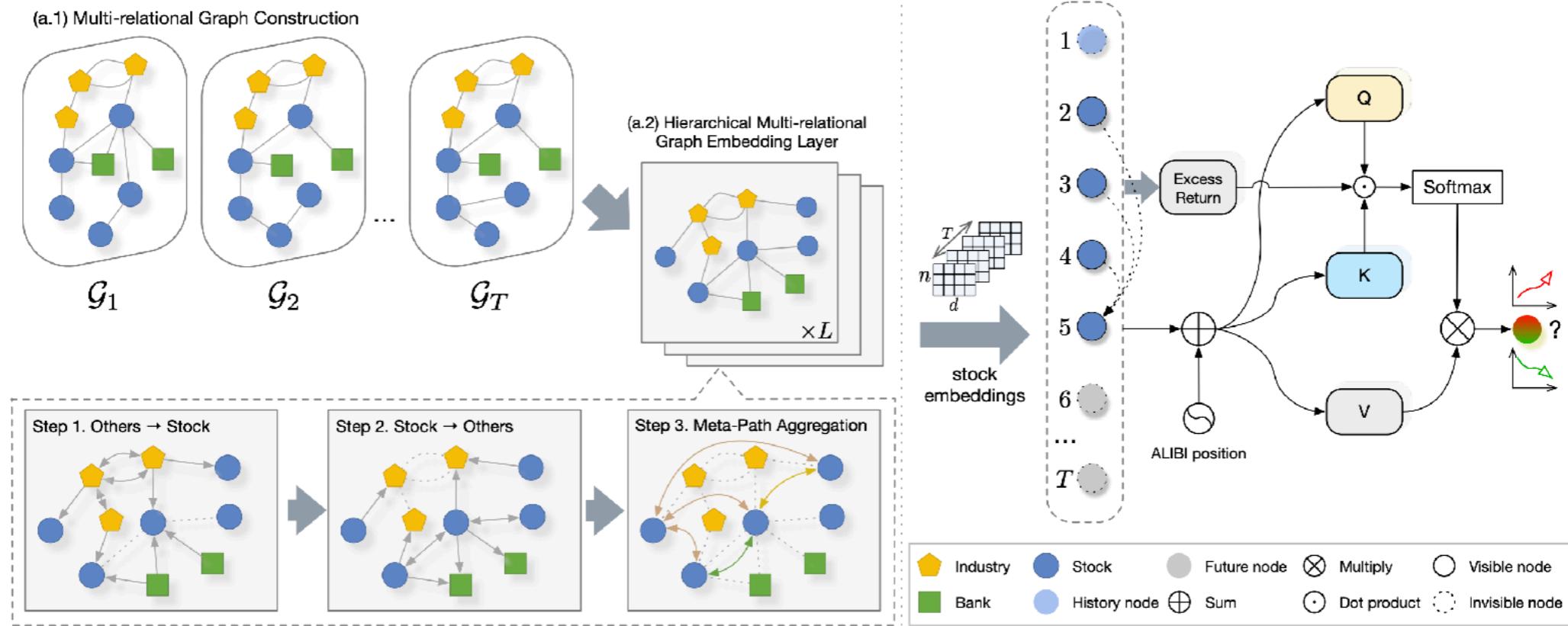
Table 1: Overall performance comparison. The best results are in bold and the second-best results are underlined. And * denotes statistically significant improvement (measured by t-test with p-value < 0.01) over all baselines.

Model	IC	ICIR	RankIC	RankICIR	AR	IR
(MA)STER	0.064 ± 0.003	0.43 ± 0.02	0.074 ± 0.004	0.48 ± 0.04	0.25 ± 0.03	2.1 ± 0.3
(MA)STER-Bi	<u>0.058 ± 0.005</u>	<u>0.38 ± 0.04</u>	<u>0.066 ± 0.008</u>	<u>0.41 ± 0.05</u>	0.19 ± 0.03	1.6 ± 0.2
Naive	0.041 ± 0.008	0.30 ± 0.05	0.046 ± 0.007	0.32 ± 0.04	0.18 ± 0.05	1.6 ± 0.6
Clustering	0.044 ± 0.003	0.36 ± 0.02	0.049 ± 0.005	0.39 ± 0.04	0.18 ± 0.04	1.7 ± 0.3

Table 2: Experiments on CSI300 to validate the effectiveness of proposed stock transformer architecture. The best results are in bold and the second-best results are underlined.

MDGNN: Multi-Relational Dynamic Graph Neural Network for Comprehensive and Dynamic Stock Investment Prediction, AAAI'24

Goal: Captures both multifaceted relationships between stocks and their temporal evolution for more accurate stock investment prediction.



Contributions:

- (1) Formalise multifaceted stock relationships through empirical analysis of market connections
- (2) Introduce hierarchical graph embedding with temporal pattern extraction architecture to represent stock markets
- (3) Effective information propagation leads to strong performance surpassing SOTA on real-world datasets

Methods	CSI 100				CSI 300			
	IC	IR	CR	Prec@30	IC	IR	CR	Prec@30
MLP	0.0027 (2.25e-03)	0.0282 (2.26e-02)	0.1166 (8.10e-03)	0.4751 (8.17e-04)	0.0039 (9.42e-04)	0.0314 (1.53e-02)	0.1721 (1.08e-02)	0.4958 (1.01e-03)
LSTM	0.0040 (1.27e-03)	0.0335 (1.31e-02)	0.1289 (1.90e-03)	0.4808 (2.70e-04)	0.0049 (6.84e-04)	0.0345 (1.09e-02)	0.1859 (1.29e-02)	0.4958 (1.99e-03)
Transformer	0.0058 (2.50e-03)	0.0422 (1.51e-02)	0.1383 (7.47e-02)	0.4987 (3.22e-03)	0.0063 (1.95e-03)	0.0442 (1.28e-02)	0.2122 (1.14e-01)	0.5065 (6.03e-03)
GAT	0.0031 (9.08e-04)	0.0274 (7.63e-03)	0.1534 (2.45e-02)	0.4812 (2.31e-03)	0.0066 (1.50e-03)	0.0454 (2.46e-02)	0.2653 (2.42e-02)	0.4991 (3.00e-04)
GCN	0.0038 (1.34e-03)	0.0305 (9.36e-03)	0.1616 (8.64e-03)	0.4927 (2.33e-03)	0.0075 (9.85e-04)	0.0674 (3.80e-02)	0.2816 (2.88e-02)	0.5055 (2.04e-03)
RGCN	0.0104 (1.29e-03)	0.0578 (7.47e-03)	0.1912 (2.84e-02)	0.4985 (2.59e-03)	0.0090 (1.69e-03)	0.0845 (1.42e-02)	0.5159 (5.32e-02)	0.5104 (1.85e-03)
HAN	0.0108 (4.08e-04)	0.0525 (2.69e-03)	0.2267 (2.48e-02)	0.4997 (3.25e-03)	0.0086 (4.68e-03)	0.0848 (4.53e-02)	0.3511 (5.72e-02)	0.5112 (4.53e-03)
HGT	0.0112 (1.35e-03)	0.0657 (7.46e-03)	0.2384 (1.98e-02)	0.5036 (4.72e-03)	0.0115 (2.05e-03)	0.0874 (1.17e-02)	0.4108 (5.65e-02)	0.4923 (6.93e-03)
EvolveGCN	0.0065 (3.54e-04)	0.0538 (3.18e-03)	0.1815 (2.81e-02)	0.4961 (2.26e-03)	0.0080 (3.46e-04)	0.5012 (4.69e-03)	0.4989 (6.09e-02)	0.4830 (3.11e-03)
HTGNN	0.0118 (3.76e-03)	0.0724 (2.45e-02)	0.2643 (8.23e-02)	0.5039 (3.54e-03)	0.0192 (7.59e-04)	0.1773 (9.94e-03)	0.4653 (7.03e-02)	0.5126 (1.12e-03)
MDGNN	0.0123 (2.75e-03)	0.0746 (1.59e-02)	0.2741 (8.11e-02)	0.5081 (3.22e-03)	0.0322 (2.43e-03)	0.2488 (4.19e-03)	0.9828 (1.13e-02)	0.5232 (3.01e-03)

Table 2: Results of methods on public datasets. The last row in each dataset indicates the percentage of improvements gained by the proposed method w.r.t the best-performed baseline. Prec@k is a shortened form of Precision@k.

DoubleAdapt: A Meta-learning Approach to Incremental Learning for Stock Trend Forecasting, KDD'23

Goal: Use incremental learning to mitigate distribution shifts of market evolution

Contributions:

- (1) Incremental learning framework to cope with distribution shifts
- (2) formulate bi-level optimisation with data and model adaptations
- (3) superior performance on real-world datasets against various distribution shifts

Model	Method	CSI 300					CSI 500						
		IC	ICIR	RankIC	RankICIR	Return	IR	IC	ICIR	RankIC	IR		
Transformer	RR	0.0449	0.3410	0.0462	0.3670	0.0881	1.0428	0.0452	0.4276	0.0469	0.4732	0.0639	0.9879
	DDG-DA	0.0420	0.3121	0.0441	0.3420	0.0823	1.0018	0.0450	0.4223	0.0465	0.4634	0.0681	1.0353
	IL	0.0431	0.3108	0.0411	0.2944	0.0854	0.9215	0.0428	0.3943	0.0453	0.4475	0.1014	1.5108
	MetaCoG	0.0463	0.3493	0.0434	0.3133	0.0952	0.9921	0.0449	0.4643	0.0469	0.4629	0.1053	0.8945
	C-MAML	0.0479	0.3560	0.0448	0.3405	0.0986	1.0537	0.0477	0.4620	0.0468	0.4861	0.0930	1.4923
	DoubleAdapt	0.0516	0.3889	0.0475	0.3585	0.1041	1.1035	0.0492	0.4653	0.0490	0.4970	0.1330	1.9761
LSTM	RR	0.0592	0.4809	0.0536	0.4526	0.0805	0.9578	0.0642	0.6187	0.0543	0.5742	0.0980	1.5220
	DDG-DA	0.0572	0.4622	0.0528	0.4415	0.0887	1.0583	0.0636	0.6181	0.0540	0.5783	0.1061	1.6673
	IL	0.0594	0.4661	0.0546	0.4362	0.1089	1.2553	0.0576	0.5550	0.0553	0.5660	0.1249	1.8461
	MetaCoG	0.0515	0.4131	0.0505	0.4197	0.1013	1.1133	0.0573	0.5673	0.0549	0.5908	0.1384	2.0546
	C-MAML	0.0568	0.4601	0.0517	0.4381	0.0963	1.1145	0.0582	0.5863	0.0550	0.5898	0.1315	1.9770
	DoubleAdapt	0.0632	0.5126	0.0567	0.4669	0.1117	1.3029	0.0648	0.6331	0.0594	0.6087	0.1496	2.2220
ALSTM	RR	0.0630	0.5084	0.0589	0.4892	0.0947	1.1785	0.0649	0.6331	0.0575	0.6030	0.1211	1.8726
	DDG-DA	0.0609	0.4915	0.0581	0.4823	0.0966	1.2227	0.0645	0.6298	0.0573	0.6029	0.1042	1.6091
	IL	0.0626	0.4762	0.0585	0.4489	0.1171	1.3349	0.0596	0.5705	0.0579	0.5712	0.1501	2.1468
	MetaCoG	0.0581	0.4676	0.0570	0.4695	0.1140	1.3228	0.0576	0.5874	0.0571	0.6086	0.1403	2.0857
	C-MAML	0.0636	0.5064	0.0588	0.4765	0.1085	1.2432	0.0647	0.6490	0.0598	0.6330	0.1644	2.4636
	DoubleAdapt	0.0679	0.5480	0.0594	0.4882	0.1225	1.4717	0.0653	0.6404	0.0607	0.6170	0.1738	2.5192
GRU	RR	0.0629	0.5105	0.0581	0.4856	0.0933	1.1428	0.0669	0.6588	0.0586	0.6232	0.1200	1.8629
	DDG-DA	0.0623	0.5045	0.0589	0.4898	0.0967	1.1606	0.0666	0.6575	0.0582	0.6234	0.1264	1.9963
	IL	0.0633	0.4818	0.0596	0.4609	0.1166	1.3196	0.0637	0.6093	0.0617	0.6291	0.1626	2.3352
	MetaCoG	0.0560	0.4443	0.0545	0.4503	0.0992	1.1014	0.0603	0.5741	0.0585	0.5720	0.1587	2.2635
	C-MAML	0.0638	0.5085	0.0595	0.4865	0.1121	1.3210	0.0646	0.6498	0.0600	0.6494	0.1693	2.5064
	DoubleAdapt	0.0687	0.5497	0.0621	0.5110	0.1296	1.5123	0.0686	0.6652	0.0632	0.6445	0.1748	2.4578

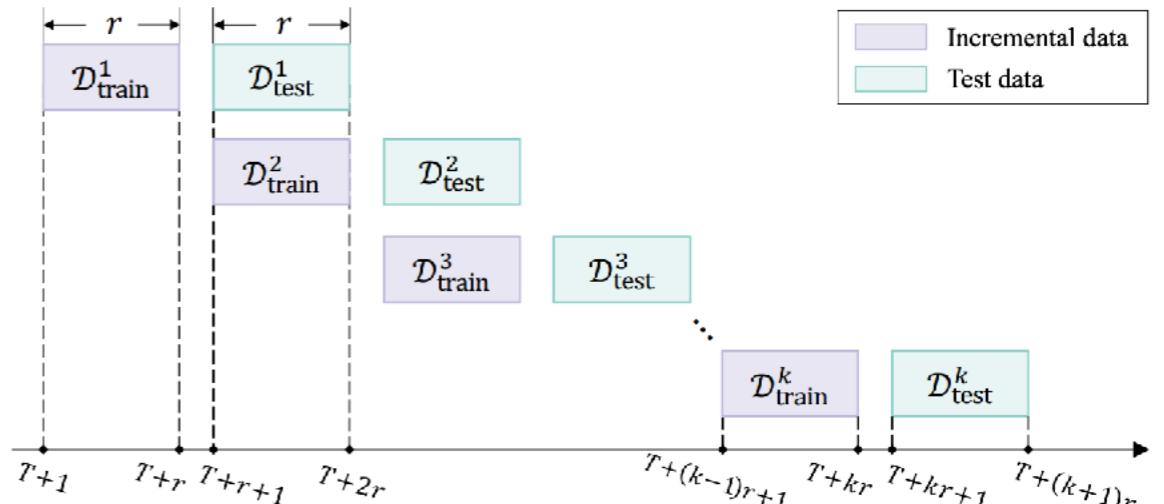


Figure 2: Illustration of IL for stock trend forecasting. r is the timespan of incremental data or test data.

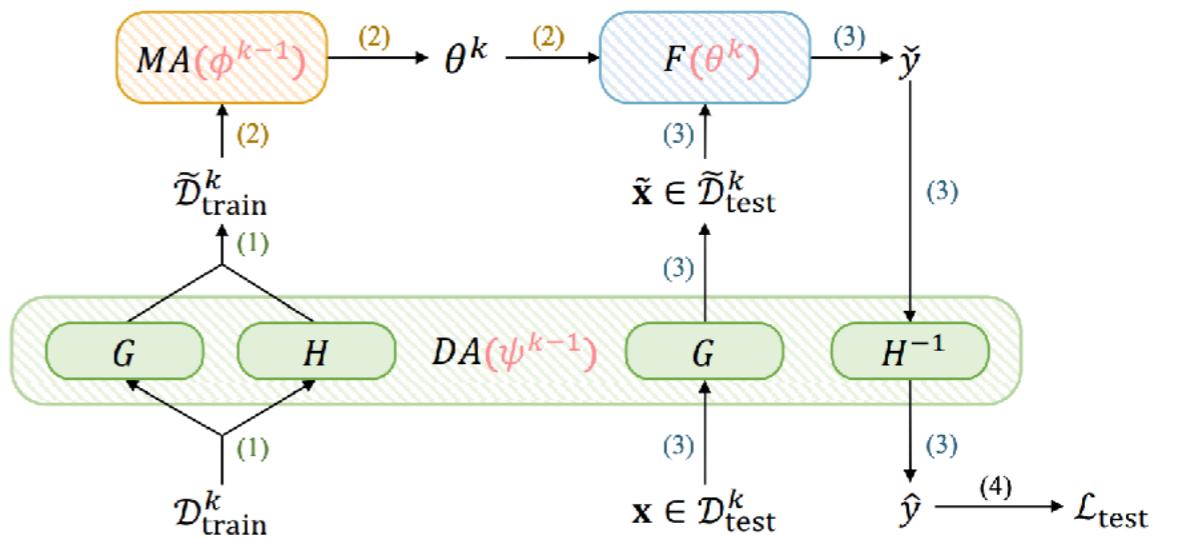


Figure 4: Overview of DoubleAdapt with a data adapter DA and a model adapter MA . The parameters are shown in red.

From News to Forecast: Iterative Event Reasoning in LLM-Based Time Series Forecasting, NeurIPS'24

Goal: Enhance time series forecasting by integrating news data through large language models and generative agents to improve prediction accuracy.

Contributions:

- (1) Developed a framework that combines news with numerical time series data through language models for enriched contextual analysis
- (2) Created a multi-domain dataset pairing time series data with curated news articles for research advancement
- (3) Implemented LLM-based agents for dynamic news filtering and continuous refinement of forecasting accuracy

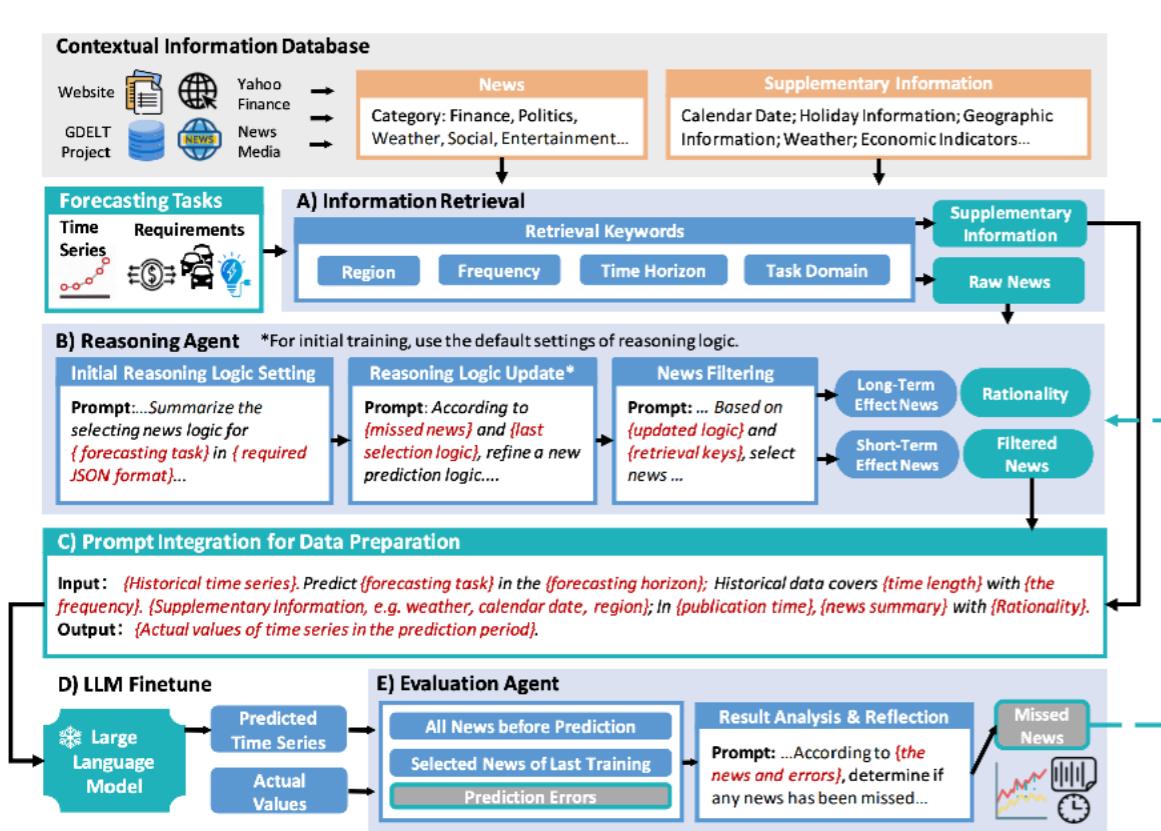


Figure 1: Integrating textual information in time series forecasting. (A) We retrieve relevant original news and supplementary information from our comprehensive database based on information such as the geographic location and time frame of the prediction task. (B) LLM-based agents analyze and select relevant news for different forecasting horizons. (C & D) The selected news and contextual information are combined with time series data for fine-tuning the LLM forecasting model. (E) Discrepancies between predictions and ground truth trigger a review of historical news and data to reprocess missed information and refine reasoning logic.

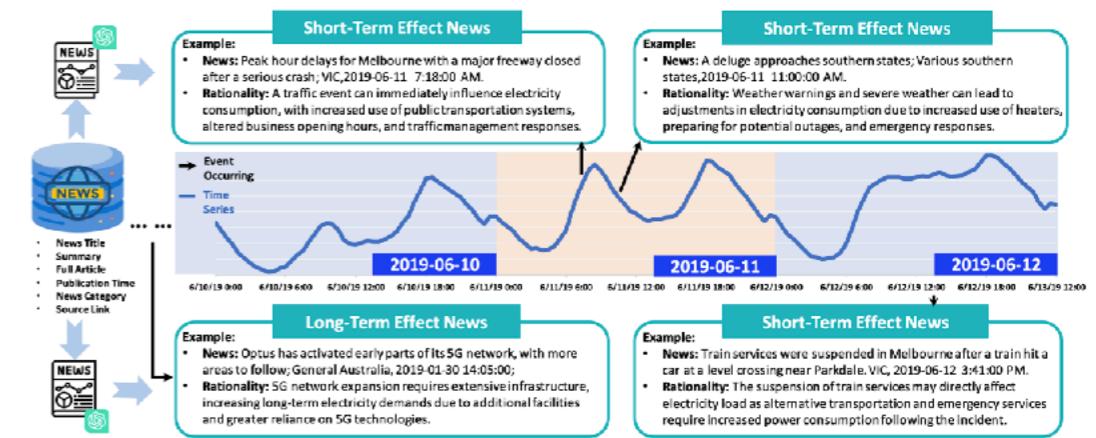


Figure 2: Relationship between news and time series. This figure illustrates the news filtered by the reasoning agents, using the example of Australia's state-level electricity demand. It features load data in Victoria state and selected news from June 10 to 12, 2019. The black arrow indicates time-specific events, the blue curve shows load fluctuations. The x-axis represents time, and the y-axis displays load values in kilowatts. The blue box displays the short-term impact news and long-term impact news selected by the reasoning agent (e.g., traffic incidents or new construction projects).

Domains	Metrics	Ours	Auto. [59]	In. [67]	DIn. [64]	iTrans. [38]	FILM [68]	Times. [58]	Pyra. [36]	PatchTST [40]	FED. [69]	GPT4TS [78]
Electricity	MAE	180.96	349.43	282.56	255.7	233.58	254.05	237.49	220.32	234.46	238.77	236.91
	MSE _{x10^-3}	78.62	251.79	166.07	161.59	135.27	153.90	134.42	97.61	133.53	133.96	142.60
	RMSE	280.39	501.78	407.52	401.98	367.79	392.3	366.64	312.42	305.41	366	377.62
	MAPE	5.15%	10.63%	8.94%	7.29%	6.86%	7.36%	6.81%	6.87%	6.56%	6.75%	6.61%
Exchange	MAE _{x10^3}	4.83	9.27	1.75	6.96	5.12	6.44	5.24	14.6	6.73	8.98	15.05
	MSE _{x10^4}	0.42	1.36	4.76	0.91	0.45	0.77	0.45	3.55	0.77	1.28	4.01
	RMSE _{x10^3}	0.65	1.17	2.18	9.52	0.671	0.875	0.673	1.88	0.875	1.13	2.00
	MAPE	0.65%	1.23%	2.32%	0.92%	0.68%	0.85%	0.70%	1.94%	0.90%	1.21%	1.34%
Traffic	MAE _{x10^2}	1.43	2.49	4.44	1.70	1.56	1.65	1.61	1.51	1.84	1.74	1.64
	MSE _{x10^3}	1.78	2.19	5.27	1.67	1.54	1.71	1.49	0.98	1.54	1.43	1.45
	RMSE _{x10^2}	4.22	4.68	7.26	4.09	3.93	4.14	3.86	3.13	3.92	3.79	3.81
	MAPE	2.68	4.28	12.27	5.74	3.20	3.28	3.17	9.22	2.85	3.96	2.84
Bitcoin	MAE _{x10^-3}	13.41	27.64	162.47	50.90	16.21	17.65	16.38	123.71	13.52	24.60	13.66
	MSE _{x10^-6}	3.67	5.26	12.75	7.13	4.03	4.20	4.05	11.12	3.68	4.96	3.70
	RMSE _{x10^-3}	4.95%	7.61%	21.28%	10.39%	5.70%	5.84%	5.64%	16.16%	5.13%	6.97%	5.08%

DiffSTOCK: Probabilistic relational Stock Market Predictions using Diffusion Models, ICASSP'24

Goal: de-noise financial data to achieve better market prediction

Contributions:

1. Showcase effective use of DDPM
2. Propose MaTCHS Denoising Model to exploit inter-stock relations with historical features
3. Achieve SOTA for movement prediction and portfolio management

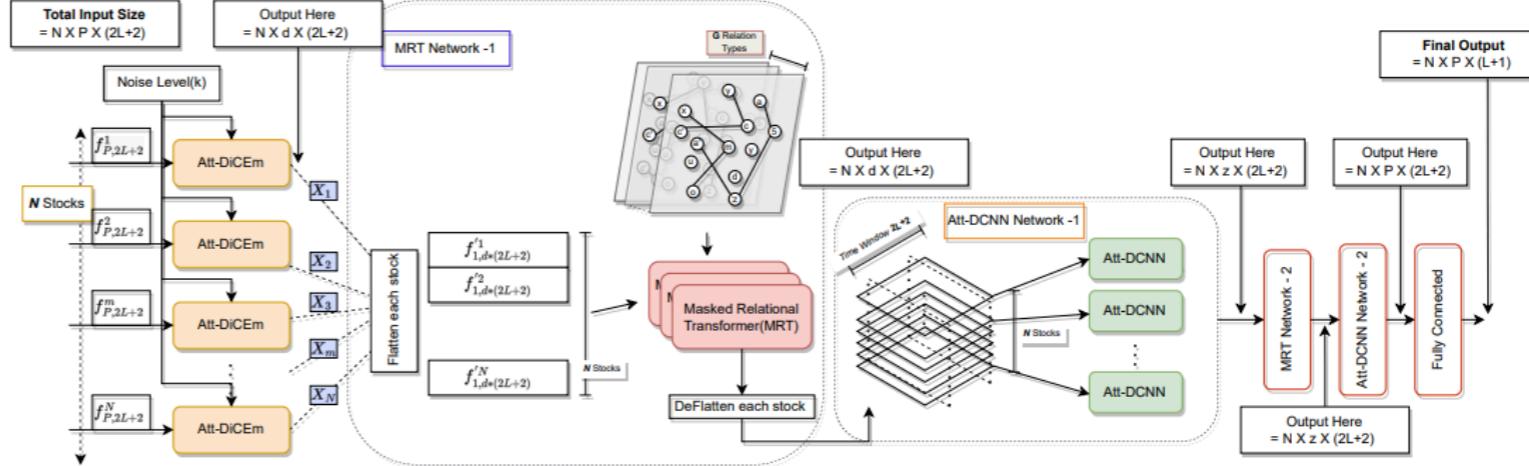


Fig. 1. *MaTCHS* Denoising Model: Masked Transformer and Convolutional network for Hypergraph relation based Stock time-series generation.

Table 1. Evaluation Results over StockNet Dataset[23]

Model	$F1 \uparrow$	Accuracy \uparrow	MCC \uparrow
RAND	$0.502 \pm 8e-4$	$0.509 \pm 8e-4$	$-0.002 \pm 1e-3$
TA	$0.513 \pm 1e-3$	$0.514 \pm 1e-3$	$-0.021 \pm 2e-3$
ARIMA - [27]	$0.529 \pm 5e-2$	$0.530 \pm 5e-2$	$-0.004 \pm 7e-2$
- [28]			
RandomForest - [29]	$0.527 \pm 2e-3$	$0.531 \pm 2e-3$	$0.013 \pm 4e-3$
TSLDA - [30]	$0.539 \pm 6e-3$	$0.541 \pm 6e-3$	$0.065 \pm 7e-3$
HAN - [31]	$0.572 \pm 4e-3$	$0.576 \pm 4e-3$	$0.052 \pm 5e-3$
StockNet - TechnicalAnalyst - [23]	$0.546 \pm -$	$0.550 \pm -$	$0.017 \pm -$
StockNet - FundamentalAnalyst - [23]	$0.572 \pm -$	$0.582 \pm -$	$0.072 \pm -$
StockNet - IndependentAnalyst - [23]	$0.573 \pm -$	$0.575 \pm -$	$0.037 \pm -$
FA	$0.559 \pm -$	$0.562 \pm -$	$0.056 \pm -$
StockNet - DiscriminativeAnalyst - [23]	$0.575 \pm -$	$0.582 \pm -$	$0.081 \pm -$
StockNet - HedgeFundAnalyst - [23]	$0.530 \pm 7e-3$	$0.532 \pm 7e-3$	$0.093 \pm 9e-3$
GCN[25]	$0.560 \pm 2e-3$	$0.562 \pm 2e-3$	$0.117 \pm 6e-3$
HATS - [26]	$0.570 \pm -$	$0.572 \pm -$	$0.148 \pm -$
Adversarial LSTM - [6]	$0.605 \pm 2e-4$	$0.608 \pm 2e-4$	$0.195 \pm 6e-4$
MAN-SF - [32]	$0.609 \pm 2e-4$	$0.613 \pm 2e-4$	$0.198 \pm 6e-4$
STHGCN - [5]	$0.568 \pm 2e-3$	$0.572 \pm 2e-3$	$0.168 \pm 6e-3$
MaTCHS (This work) - AttDiCEm i.e. without relations	$0.585 \pm 2e-3$	$0.587 \pm 2e-3$	$0.175 \pm 6e-3$
MaTCHS (This work) - Aggregated relations	$0.611 \pm 2e-3$	$0.612 \pm 2e-3$	$0.206 \pm 6e-3$
MaTCHS (This work)	$0.623 \pm 2e-3$	$0.621 \pm 2e-3$	$0.214 \pm 6e-3$
MaTCHS (with Diffusion w/o Adap. Noise)	$0.631 \pm 2e-3$	$0.634 \pm 2e-3$	$0.225 \pm 6e-3$
MaTCHS (with Diffusion)			

Table 2. Evaluation Results over NASDAQ and NYSE Dataset(2 decimal places disp.) [6]

Model	NYSE		NASDAQ	
	SR@5	IRR@5	SR@5	IRR@5
ARIMA [27]	$0.33 \pm 3e-3$	$0.10 \pm 5e-3$	$0.55 \pm 1e-3$	$0.10 \pm 6e-3$
A-LSTM [33]	$0.81 \pm 4e-3$	$0.14 \pm 7e-3$	$0.97 \pm 5e-3$	$0.23 \pm 3e-3$
GCN [25]	$0.70 \pm 3e-3$	$0.10 \pm 6e-3$	$0.75 \pm 4e-3$	$0.13 \pm 1e-3$
HATS [26]	$0.73 \pm 5e-3$	$0.12 \pm 2e-3$	$0.80 \pm 6e-3$	$0.15 \pm 7e-3$
DQN [34]	$0.72 \pm 5e-3$	$0.12 \pm 4e-3$	$0.93 \pm 5e-3$	$0.20 \pm 6e-3$
iRDPG [35]	$0.85 \pm 7e-3$	$0.18 \pm 3e-3$	$1.32 \pm 5e-3$	$0.28 \pm 4e-3$
Rank LSTM [36]	$0.79 \pm 1e-3$	$0.12 \pm 6e-3$	$0.95 \pm 4e-3$	$0.22 \pm 2e-3$
GCN [25]	$0.72 \pm 7e-3$	$0.16 \pm 3e-3$	$0.46 \pm 4e-3$	$0.13 \pm 5e-3$
RSR-E [6]	$0.88 \pm 6e-3$	$0.20 \pm 3e-3$	$1.12 \pm 5e-3$	$0.26 \pm 4e-3$
RSR-I [6]	$0.95 \pm 1e-3$	$0.21 \pm 3e-3$	$1.34 \pm 6e-3$	$0.39 \pm 5e-3$
STHAN-SR [5]	$1.10 \pm 1e-3$	$0.255 \pm e-3$	$1.40 \pm 7e-3$	$0.44 \pm 1e-2$
MaTCHS	$1.13 \pm \pm e-3$	$0.267 \pm e-3$	$1.45 \pm 7e-3$	$0.45 \pm 1e-2$
MaTCHS(Agg)	$0.97 \pm \pm e-3$	$0.221 \pm e-3$	$1.34 \pm 7e-3$	$0.40 \pm 1e-2$
MaTCHS(16)	$1.14 \pm \pm e-3$	$0.270 \pm e-3$	$1.46 \pm 7e-3$	$0.46 \pm 1e-2$
MaTCHS with Diffusion w/o Adap. Noise	$1.15 \pm \pm e-3$	$0.274 \pm e-3$	$1.48 \pm 7e-3$	$0.46 \pm 1e-2$
MaTCHS with Diffusion	$1.18 \pm \pm e-3$	$0.285 \pm e-3$	$1.52 \pm 7e-3$	$0.48 \pm 1e-2$
% Improv. (SOTA w.r.t. STHAN-SR)	7.92	9.81	6.18	8.07

Double-Path Adaptive-correlation Spatial-Temporal Inverted Transformer for Stock Time Series Forecasting, KDD'25

Goal: De-noise and model financial data in spatial-temporal graph to perform price movement prediction

Contributions:

- (1) Two-step Locality-Aware-Attention method to identify profitable samples
- (2) Refinement-labelling noise in the selected profitable samples
- (3) Outperform baseline models in 3 real-world financial datasets

Notes / Innovations:

Temporal: inverted features as tokens $Y_1(a_1, b_1..) \rightarrow A(y_1, y_2 ..)$

Spatial (dual paths): feature-wise and price-wise temporal relations

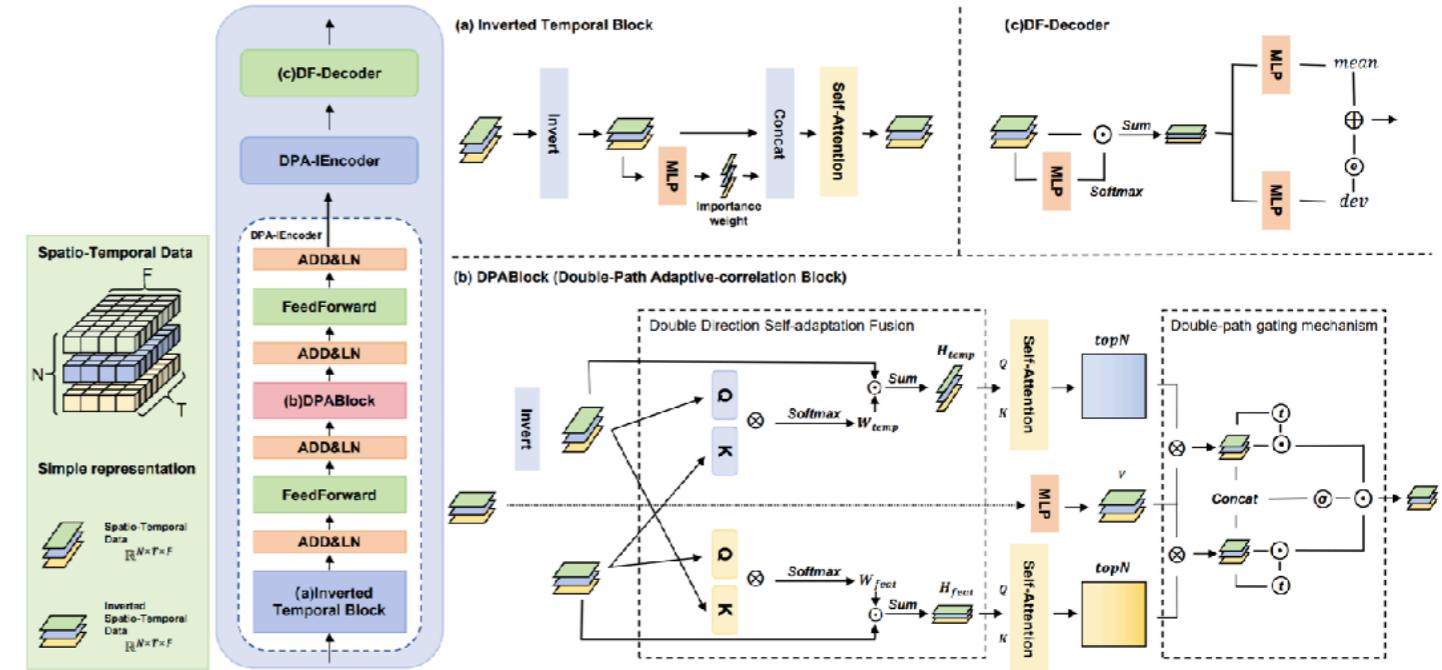


Figure 1: The overview of the proposed Double-Path Adaptive-correlation Spatial-Temporal Inverted Transformer.

	CSI500										CSI1000									
	IC	PNL	A_RET	A_VOL	MAXD	SHARPE	WINR	PL	IC	PNL	A_RET	A_VOL	MAXD	SHARPE	WINR	PL				
STGNN	ASTGCN	0.1022	0.3758	0.3727	0.0872	0.1813	4.2740	0.6281	2.0069	0.0840	0.7716	0.7652	0.1019	0.1285	6.8903	0.6694	3.0834			
	STGCN	0.0349	0.0466	0.0462	0.0676	0.1693	0.6841	0.5537	1.1236	0.0404	0.0388	0.0385	0.0865	0.2850	0.4449	0.5000	0.9310			
	MSTGCN	0.0732	0.0328	0.0325	0.0694	0.2245	0.4692	0.5793	1.9280	0.0919	0.0804	0.0732	0.1180	0.1333	7.3996	0.5901	3.2232			
	MTGNN	0.1176	0.4652	0.4614	0.1005	0.1652	4.5845	0.6116	2.1265	0.0941	1.1608	0.0756	0.0324	15.2259	0.7759	23.4614				
	STEMGNN	0.0845	0.2124	0.2107	0.0842	0.1223	2.5016	0.5702	1.5455	0.0176	0.0154	0.0153	0.0205	0.0141	0.7448	0.0455	1.6529			
RNN	DCRNN	0.0609	0.0899	0.0892	0.0839	0.2084	1.0630	0.5289	1.1794	0.0420	0.1080	0.1071	0.0850	0.2552	1.2591	0.4793	0.8216			
	TPA_LSTM	0.0052	-0.0959	-0.0951	0.0717	0.2364	-1.3272	0.4959	0.7985	0.0344	0.0413	0.0409	0.0811	0.2304	0.5048	0.4959	0.9201			
FORMER	FC_LSTM	0.1193	0.5402	0.5357	0.1056	0.1428	5.0750	0.6033	2.3893	0.0482	1.0291	1.0206	0.1041	0.1053	9.8080	0.7190	5.4323			
	CROSSFORMER	0.0134	0.1447	0.1435	0.0779	0.2596	1.8423	0.4463	0.7295	0.0365	0.0357	0.03508	0.0964	0.2412	3.6383	0.5909	1.7650			
	ITTRANSFORMER	0.0225	-0.0517	-0.0513	0.1031	0.2853	-0.4976	0.5331	0.9226	0.0391	0.1947	0.1931	0.0945	0.2717	2.0267	0.5614	1.3757			
	DPA-STIFormer	0.1312	0.8794	0.8721	0.1033	0.1164	8.4395	0.7273	4.2453	0.1404	1.2809	1.2703	0.1111	0.0791	12.4611	0.8058	9.2447			
	NYSE										NASDAQ									
	IC	PNL	A_RET	A_VOL	MAXD	SHARPE	WINR	PL	IC	PNL	A_RET	A_VOL	MAXD	SHARPE	WINR	PL				
STGNN	ASTGCN	0.0400	0.0314	0.0322	0.3251	0.3936	0.9893	0.5171	0.8252	0.0323	0.2578	0.2644	0.1879	0.1369	1.4070	0.5128	1.2638			
	STGCN	0.0150	0.1009	0.1034	0.2707	0.1926	0.3821	0.5128	1.0667	0.0383	0.3145	0.3225	0.2408	0.1766	1.3396	0.5085	1.2675			
	MSTGCN	0.0351	0.0767	0.0786	0.2887	0.3035	0.2723	0.5085	1.0189	0.0227	0.2588	0.2655	0.1776	0.0973	1.1524	0.4915	1.2845			
	MTGNN	-0.0035	-0.0291	-0.0299	0.1209	0.0604	-0.1000	0.4957	0.9831	0.0163	0.1375	0.1410	0.1537	0.1076	0.9179	0.5385	1.1803			
	STEMGNN	0.0254	-0.1557	-0.1597	0.2335	0.2691	-0.5839	0.4915	0.8772	0.0171	0.0273	0.0280	0.2240	0.2597	0.1249	0.4786	1.0224			
RNN	DCRNN	0.0016	0.0122	0.0125	0.1642	0.2281	0.0763	0.4701	1.0128	-0.0056	0.0253	0.0259	0.1394	0.1273	0.1858	0.5284	1.0325			
	TPA_LSTM	0.0260	0.1729	0.1774	0.2557	0.3244	0.6937	0.5385	1.2037	0.0270	-0.0899	-0.0922	0.2065	0.3087	-0.4465	0.4786	0.9253			
FORMER	FC_LSTM	-0.0012	0.1255	0.1287	0.3391	0.3694	0.3795	0.4872	1.0676	0.0083	0.2479	0.2542	0.2206	0.2393	1.2508	0.4829	1.2476			
	CROSSFORMER	0.0097	0.0204	0.0209	0.2767	0.2766	0.0755	0.5128	1.0128	0.0231	0.0611	0.0627	0.2081	0.1456	0.3010	0.4786	1.0544			
	ITTRANSFORMER	0.0011	-0.5296	-0.5431	0.1141	0.5510	-4.7592	0.4316	0.3288	-0.0097	-0.0931	-0.0954	0.1158	0.1185	-0.8241	0.5369	0.5229			
	DPA-STIFormer	0.0453	0.2135	0.2190	0.2987	0.1999	1.8115	0.5427	1.1256	0.0416	0.3902	0.4002	0.3199	0.1462	1.4945	0.5470	1.2320			

Table 2: Comparison results on four datasets. ↓ indicates that the smaller the metric is better. The best result is in bold.

GARCH-Informed Neural Networks for Volatility Prediction in Financial Markets, ArXiv'24

Goal: De-noise and model financial data in spatial-temporal graph to predict volatility of financial markets

Contributions:

- (1) Propose GINN (inspired by PINN) adding GARCH predictions to regularise an LSTM's loss
- (2) Empirically discover GINN-0's ability to predict without ground truth data
- (3) Superior performance of GINN and GINN-0 across multiple global stock market indices

Notes:

- GARCH: Generalised Autoregressive Conditional Heteroscedasticity model, a standard model for forecasting stock volatility

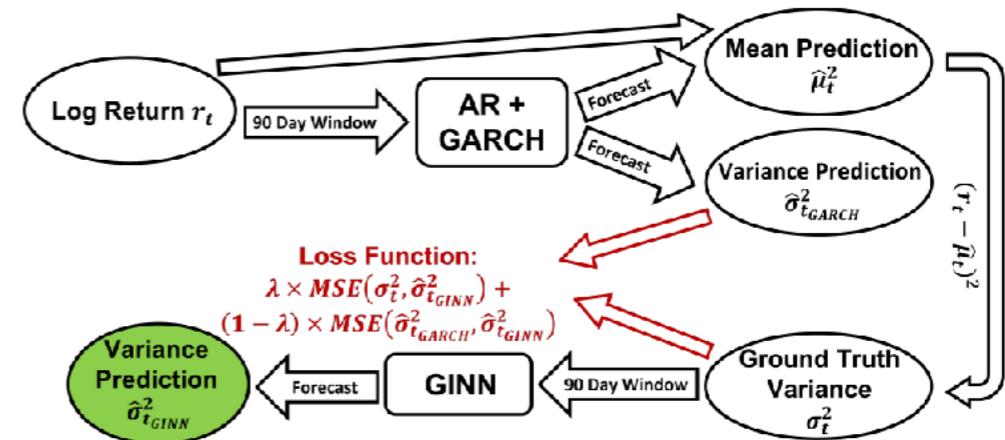


Fig. 4. The variance prediction workflow diagram for the GINN model

TABLE I
PERFORMANCE RESULTS OF MODELS IN THE OUT-OF-SAMPLE TESTING DATASET

Data	Perform. Metric	GARCH	GJR-GARCH	TGARCH	LSTM	GINN	GINN-0
GSPC	R2	0.299	-0.053	-0.092	0.011	0.329	0.319
	MSE	16.471	24.473	25.263	23.241	15.774	16.004
	MAE	1.094	1.159	1.162	1.181	1.042	1.041
DJIA	R2	0.266	-0.061	0.017	0.011	0.339	0.295
	MSE	22.049	31.477	29.112	29.726	19.847	21.173
	MAE	1.122	1.192	1.169	1.185	1.068	1.053
NYA	R2	0.266	-0.159	-0.131	0.031	0.323	0.321
	MSE	19.654	30.593	29.783	25.950	18.114	18.170
	MAE	1.098	1.192	1.168	1.139	1.045	1.051
RUT	R2	0.267	-0.094	-0.005	-0.016	0.323	0.313
	MSE	13.221	19.486	17.817	18.318	12.214	12.388
	MAE	1.068	1.129	1.105	1.145	0.976	0.974
HSI	R2	0.035	-0.003	0.079	-0.154	0.052	0.058
	MSE	5.823	6.096	5.601	6.969	5.721	5.685
	MAE	1.116	1.136	1.124	1.217	1.154	1.145
NIK	R2	0.004	0.073	0.100	-0.158	0.059	0.066
	MSE	6.153	5.763	5.595	7.154	5.812	5.770
	MAE	1.110	1.086	1.075	1.127	1.036	1.047
FTSE	R2	-0.004	-0.034	0.072	-0.036	0.154	0.144
	MSE	16.391	16.851	15.100	16.918	13.819	13.979
	MAE	1.177	1.197	1.155	1.263	1.081	1.088
Color Coding Scheme							
Best Performance		Second Best		Third Best		Worst Performance	

Stock Movement Prediction Based on Bi-typed Hybrid-relational Market Knowledge Graph via Dual Attention Networks, TKDE'23

Goal: Use graphs to represent time-evolving dynamics between and within stocks

Contributions:

1. Model stock momentum spillover by bi-typed hybrid-relational knowledge graph
2. Dual Attention Networks to for stock prediction
3. Propose two benchmark datasets and show promising performance of proposed approach

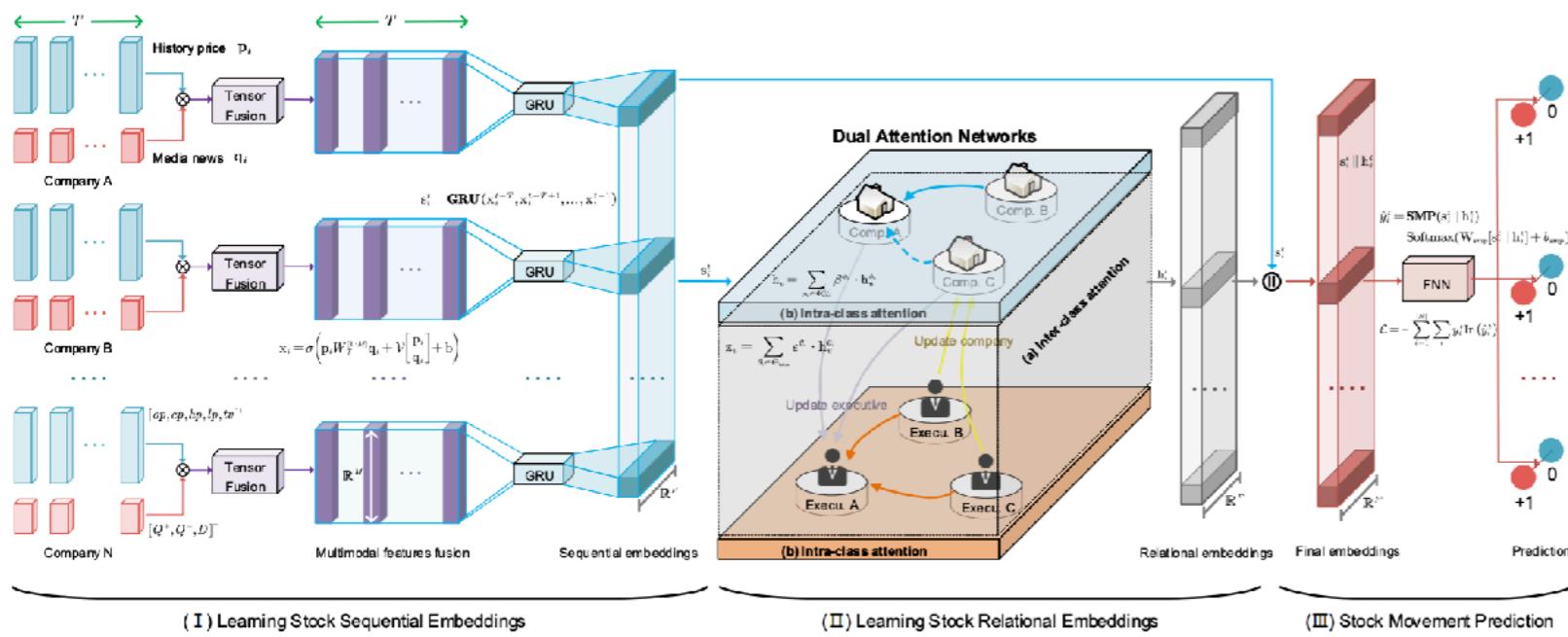


Fig. 2. The overall framework of the proposed method. **(I) Learning Stock Sequential Embeddings** based on Tensor Fusion and GRU. Tensor Fusion is the Neural Tensor Network (NTN) to learn the fusion of the technical indicators vector p_i and media news sentiment features q_i . The GRU is designed to learn the sequential embedding s_i^t . **(II) Learning Stock Relational Embeddings** by a dual mechanism to model the mutual affects and inner interaction among the bi-typed entities (i.e. companies and executives) alternately, including: **(a)** inter-class attention, and **(b)** intra-class attention. The former aims to deal with the interaction between listed companies and their associated executives and the latter aims to learn the interaction among the same type of entities. **(III) Stock Movement Prediction** via Feed-forward Neural Network (FNN) with the learned firm embeddings.

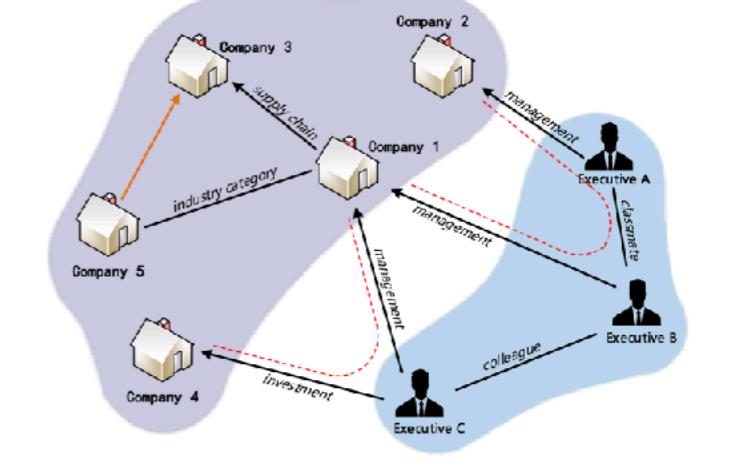


Fig. 1. Example of a bi-typed (i.e. listed companies, and executives) hybrid-relational (i.e. explicit relations, and implicit relation) market knowledge graph (MKG). The relational information in MKG is essential for stock prediction but has not been well utilized in previous works.

Methods	CSI100E		CSI300E	
	Accuracy	AUC	Accuracy	AUC
LSTM [Hochreiter and Schmidhuber 1997]	51.14	51.33	51.78	52.24
GRU [Cho et al. 2014]	51.66	51.46	51.11	52.30
GCN [Kipf and Welling 2017]	51.58	52.18	51.68	51.81
GAT [Velickovic et al. 2018]	52.17	52.78	51.40	52.24
RGCN [Schlichtkrull et al. 2018]	52.33	52.69	51.79	52.59
HGT [Hu et al. 2020]	53.01	52.51	51.70	52.19
MAN-SF [Sawhney et al. 2020a]	52.86	52.23	51.91	52.48
STHAN-SR [Sawhney et al. 2021]	52.78	53.05	52.89	53.48
AD-GAT [Cheng and Li 2021]	54.56	55.46	52.63	54.29
DANSMP (ours)	57.75	60.78	55.79	59.36

Table 5. Stock prediction results of different models.

Methods	CSI100E		CSI300E	
	IRR	SR	IRR	SR
LSTM [Hochreiter and Schmidhuber 1997]	-4.57%	-2.1713	-0.38%	-0.326
GRU[Cho et al. 2014]	-2.55%	-1.053	-3.73%	-1.197
GCN [Kipf and Welling 2017]	1.59%	0.719	3.55%	1.873
GAT [Velickovic et al. 2018]	0.3%	0.050	-1.82%	-1.121
RGCN [Schlichtkrull et al. 2018]	6.41%	3.789	-3.64%	-1.905
HGT[Hu et al. 2020]	2.54%	1.716	0.36%	0.076
MAN-SF [Sawhney et al. 2020a]	-2.91%	-1.590	1.38%	0.604
STHAN-SR[Sawhney et al. 2021]	-0.12%	-0.092	5.41%	1.565
AD-GAT [Cheng and Li 2021]	2.34%	1.190	15.12%	4.081
DANSMP (ours)	10.18%	4.112	16.97%	4.628

Table 6. Profitability of all methods in back-testing.

Binary Trend Classification

ECHO-GL: Earnings Calls-Driven Heterogeneous Graph Learning for Stock Movement Prediction, AAAI'24

- Goal: extract from earnings calls and learn complex dynamics of stock relations
- Contributions: (1) **first** to model stock relations derived from earnings calls (2) capture the stocks' spatial-temporal relations with multimodal heterogeneous graph (3) superior movement prediction performance on real-world datasets

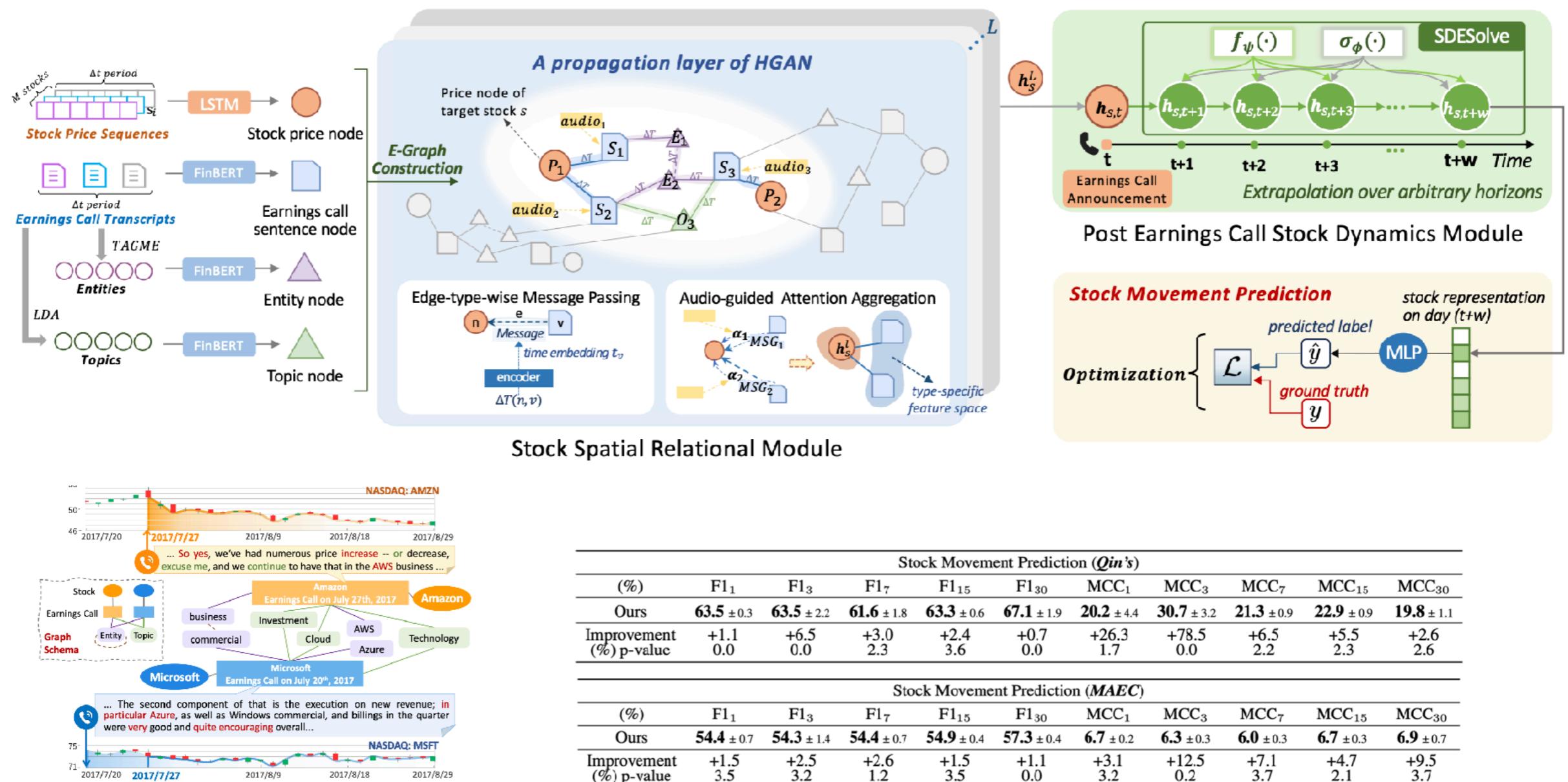


Figure 1: An example of earnings calls-driven stock relations' effect on stock movements.

Trade When Opportunity Comes: Price Movement Forecasting via Locality-Aware Attention and Iterative Refinement Labelling, IJCAI'24

Goal: De-noise financial data to improve price movement prediction

Contributions:

- (1) Two-step Locality-Aware-Attention method to identify profitable samples
- (2) Refinement-labelling noise in the selected profitable samples
- (3) Outperform baseline models in 3 real-world financial datasets

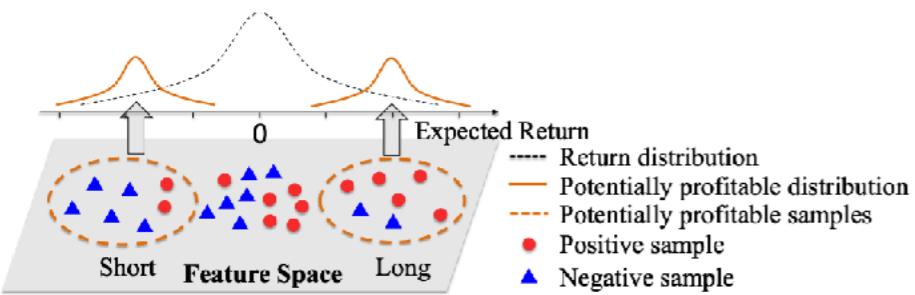


Figure 1: Illustration of the potential profitable samples. The half top figure represents the probability density function (PDF) of expected return over corresponding samples (best viewed in color).

Methods	China's A-share stocks			Cryptocurrency			Ranking Count	
	PR(%)	WLR	AR	PR(%)	WLR	AR	1 st	2 nd
Quantitative Investment Methods	ALSTM	51.9	1.133	3.24E-3	47.9	0.933	5.9E-5	0 0
	TabNet	51.8	1.299	5.14E-3	51.0	0.890	1.4E-5	0 0
	Transformer	53.2	1.230	5.73E-3	38.7	0.888	-5.0E-5	0 0
	Adamct	52.7	1.309	5.73E-3	49.3	1.177	1.3E-4	0 2
	LightGBM	55.0	1.331	7.26E-3	51.0	0.890	1.4E-5	1 1
	DoubleEnsemble	54.0	1.225	5.75E-3	-	-	-	0 0
Time-series Methods	TCTS	55.6	0.913	2.09E-3	-	-	-	0 0
	iTransformer	53.8	1.095	4.01E-3	-	-	-	0 0
	PatchTST	53.0	1.274	5.17E-3	-	-	-	0 0
Noisy Labels Methods	TimesNet	55.5	1.131	6.16E-3	-	-	-	0 0
	CNLCU	52.6	1.233	5.03E-3	52.9	1.220	1.2E-4	1 0
	FINE	55.3	1.070	4.35E-3	56.3	0.863	9.7E-5	0 1
Ours	SEAL	56.6	1.200	5.95E-3	53.0	0.969	7.2E-5	0 1
	LA-Attention	56.6	1.142	5.27E-3	51.2	0.826	9.5E-5	0 1
	RA-Labeling	55.2	1.038	3.53E-3	56.2	1.034	1.4E-4	0 1
	LARA	59.1	1.274	7.79E-3	57.8	1.059	1.5E-4	4 0

Table 1: Quantitative comparisons among different methods on the China's A-share stocks and the cryptocurrency (BTC/USDT). - means that the corresponding method is either not implemented or unsuitable for corresponding setting. We retrieve the top 1000 signals with the highest probability for each experiment. The best performance is highlighted in **bold**. The second best is highlighted with underline.

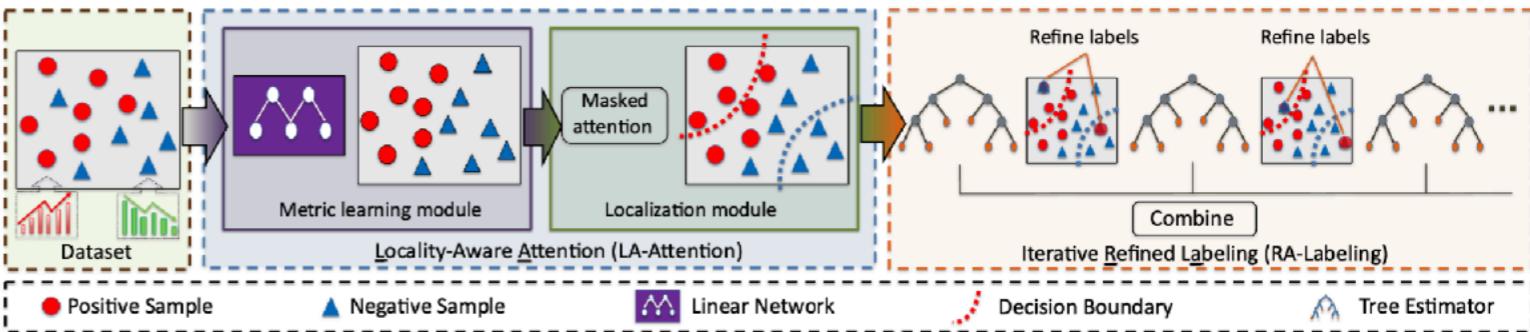


Figure 2: The workflow of our proposed LARA framework. LARA first extracts potentially profitable samples from the noisy market and then refines their labels. LARA consists of two sequential components: LA-Attention and RA-Labeling.

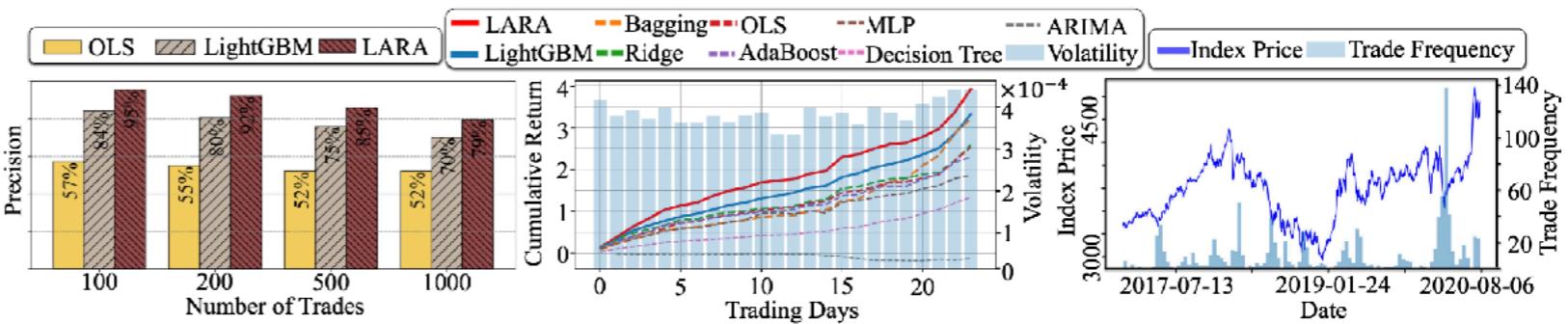


Figure 4: **Left:** Precision with the different number of trades (#Transactions) among three methods on 512480.SH. **Middle:** Quantitative comparisons over cumulative return between LARA and a set of baselines on 512480.SH. The right y-axis illustrates the volatility [Campbell et al., 1997] on each trading day. **Right:** Plot of the trade frequency in the China's A-share market and the corresponding stock index price. The left y-axis denotes the stock index price while the right y-axis denotes the trade frequency.

Learning to Generate Explainable Stock Predictions using Self-Reflective Large Language Models, WWW'24

Goal: enable LLMs to generate explainable stock predictions by self-reflection and reinforcement learning

Contributions:

- (1) Identifies limitations of LLMs in weighing market factors and lack of expert-annotated samples for stock predictions.
- (2) Proposes SEP framework using self-reflection and PPO techniques for autonomous generation of stock predictions.
- (3) Validates SEP's effectiveness through experiments and demonstrates generalizability to portfolio optimization tasks.

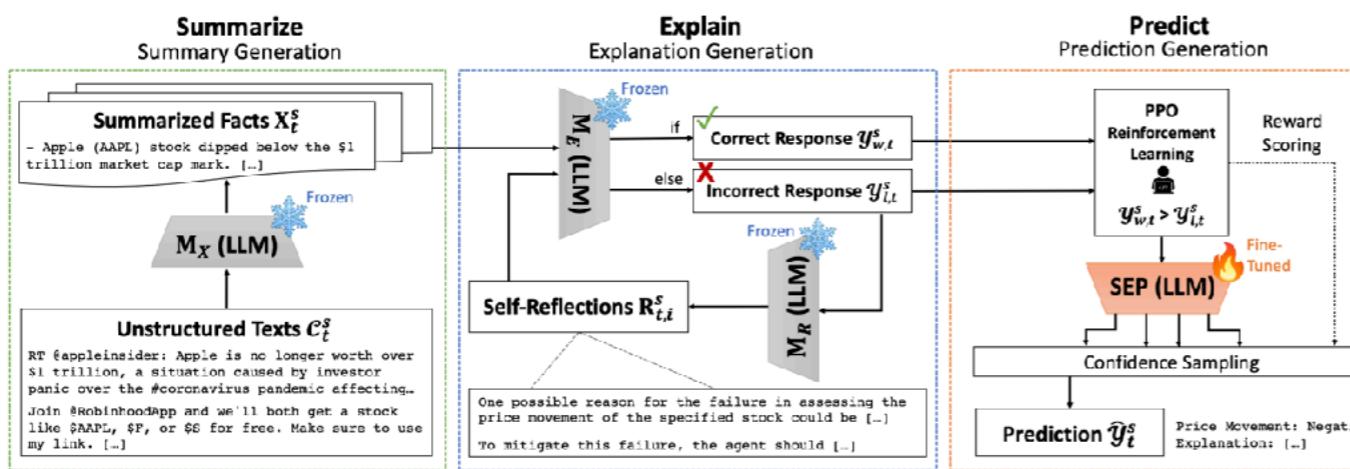


Figure 2: Overall framework of our proposed SEP method, which consists of three components: Summarize, Explain and Predict.

Table 1: Performance comparisons in accuracy and MCC of our SEP model against baselines. The best results are boldfaced.

Models	Top 1 Stock, GPT-3.5				Remaining Stocks, Vicuna				
	All Texts		Informative Texts		All Texts		Informative Texts		
	Accuracy	MCC	Accuracy	MCC	Accuracy	MCC	Accuracy	MCC	
Deep-Learning Models	VAE+Att	49.96	0.0046	-	-	49.83	0.0070	-	-
	GRU+Att	50.15	0.0125	-	-	50.77	0.0189	-	-
	Transformer	50.06	0.0089	-	-	50.17	0.0135	-	-
Large Language Models	GPT-3.5	20.80	0.0094	29.35	0.0298	17.57	0.0027	22.99	0.0052
	Vicuna	40.85	0.0114	45.29	0.0368	39.66	0.0115	43.30	0.0301
	FinGPT	47.61	0.0158	51.56	0.0384	45.76	0.0161	46.12	0.0379
	SEP (Ours)	51.38	0.0302	54.35	0.0993	47.59	0.0203	50.57	0.0508

Table 2: Comparison of explanation quality from our SEP model against baselines. The best results are boldfaced.

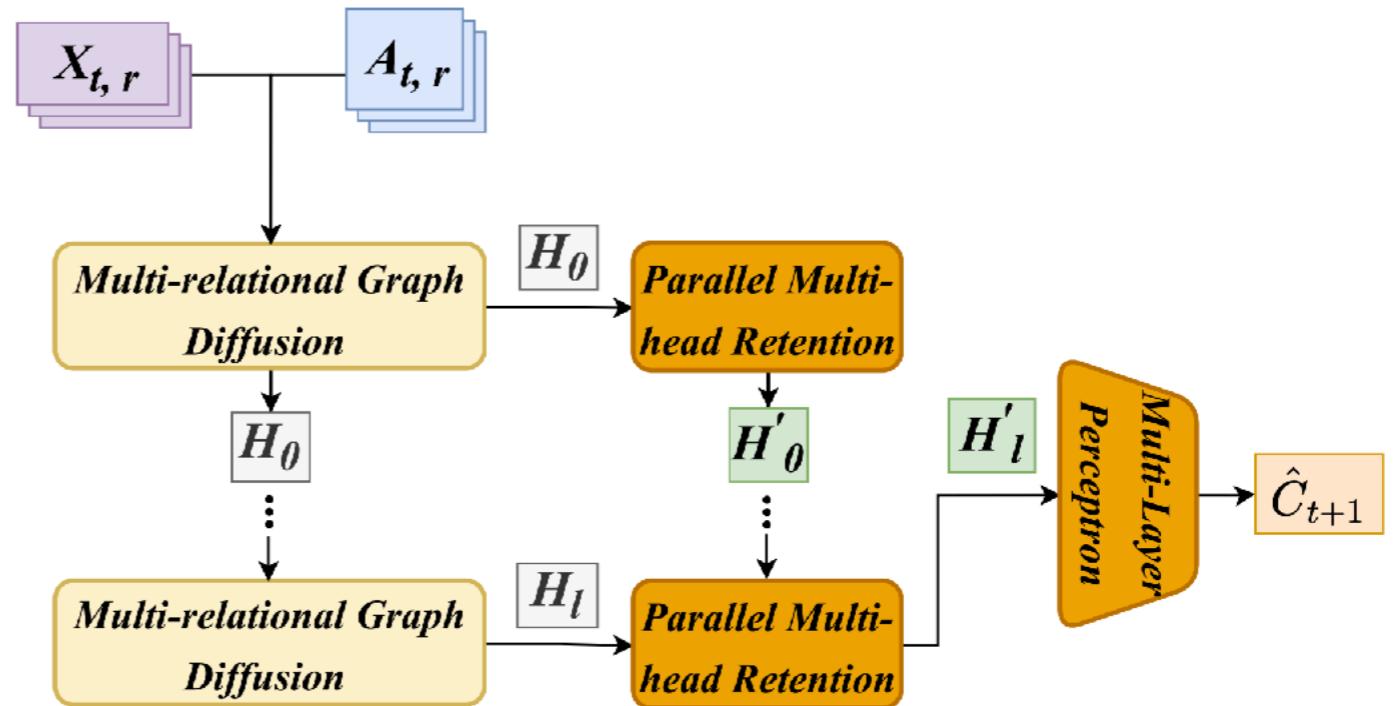
Metric	GPT-3.5	Vicuna	SEP (Ours)
Relevance to Stock Movement	5.407	5.396	5.449
Financial Metrics	2.957	3.146	3.334
Global & Industry Factors	3.180	3.576	3.700
Company Developments	3.905	4.066	4.224
Temporal Awareness	3.951	4.066	4.170
Balance of Positive & Negative	4.030	4.084	4.224
Contextual Understanding	4.012	4.098	4.193
Clarity & Coherence	6.271	6.325	6.439
Consistency with Information	5.575	5.652	6.006
Sensitivity to Updates	4.112	4.172	4.362

Multi-relational Graph Diffusion Neural Network with Parallel Retention for Stock Trends Classification, ICASSP'24

Goal: Use graphs to represent time-evolving dynamics between and within stocks

Contributions:

- (1) quantifies inter-stock dynamics using information entropy and signal energy
- (2) dual channel graph representation of hierarchical intra-stock features
- (3) Effectively forecast trends with 3 real-world datasets over seven years



Method	NASDAQ			NYSE			SSE		
	Acc(%)	Mcc	F1	Acc(%)	Mcc	F1	Acc(%)	Mcc	F1
DA-RNN [15]	56.33±1.15	0.04±4.06×10 ⁻³	0.54±0.02	57.28±0.76	0.05±2.23×10 ⁻³	0.56±0.01	57.03±0.42	0.04±2.41×10 ⁻³	0.55±0.02
HMG-TF [8]	58.32±0.41	0.10±1.79×10 ⁻³	0.57±0.01	59.11±0.35	0.09±3.81×10 ⁻³	0.59±0.01	58.90±0.36	0.11±3.01×10 ⁻³	0.59±0.01
DTML [5]	57.56±0.67	0.06±1.98×10 ⁻³	0.58±0.01	58.78±0.45	0.08±2.00×10 ⁻³	0.60±0.01	59.63±0.21	0.09±5.42×10 ⁻³	0.59±0.01
HATS [6]	50.37±1.80	0.01±4.79×10 ⁻³	0.48±0.02	51.93±0.76	0.02±6.55×10 ⁻³	0.50±0.03	53.13±0.47	0.02±5.07×10 ⁻³	0.50±0.01
GraphWaveNet [17]	59.19±0.55	0.06±6.83×10 ⁻³	0.60±0.02	62.14±1.08	0.07±3.20×10 ⁻³	0.59±0.02	60.78±0.23	0.06±2.93×10 ⁻³	0.57±0.01
HyperStockGAT [13]	57.23±0.71	0.06±5.36×10 ⁻³	0.59±0.02	59.34±0.46	0.08±5.73×10 ⁻³	0.61±0.02	58.36±0.22	0.09±4.10×10 ⁻³	0.58±0.02
TPGNN [18]	60.42±0.49	0.10±3.45×10 ⁻³	0.61±0.02	61.81±0.19	0.11±4.45×10 ⁻³	0.60±0.02	62.69±0.10	0.12±1.66×10 ⁻³	0.60±0.02
MGDPR	62.77±0.65	0.13±4.49×10⁻³	0.62±0.01	64.54±0.20	0.13±1.88×10⁻³	0.63±0.01	63.90±0.32	0.14±2.01×10⁻³	0.62±0.02

MANA-Net: Mitigating Aggregated Sentiment Homogenization with News Weighting for Enhanced Market Prediction, CIKM'24

Goal: improve market prediction by developing a novel method that better handles sentiment aggregation from large financial news datasets

Contributions:

- (1) Identified "aggregated sentiment homogenization" issue in existing financial news aggregation methods.
- (2) Utilising market-news attention mechanism for weighted sentiment aggregation.
- (3) Superior prediction performance compared to benchmark.

Figure 3: An overview of MANA-Net. The market prices and news sentiments data are as defined in Section 3.1. Blue arrows show the feed-forward process and grey arrows show the back-propagation process.

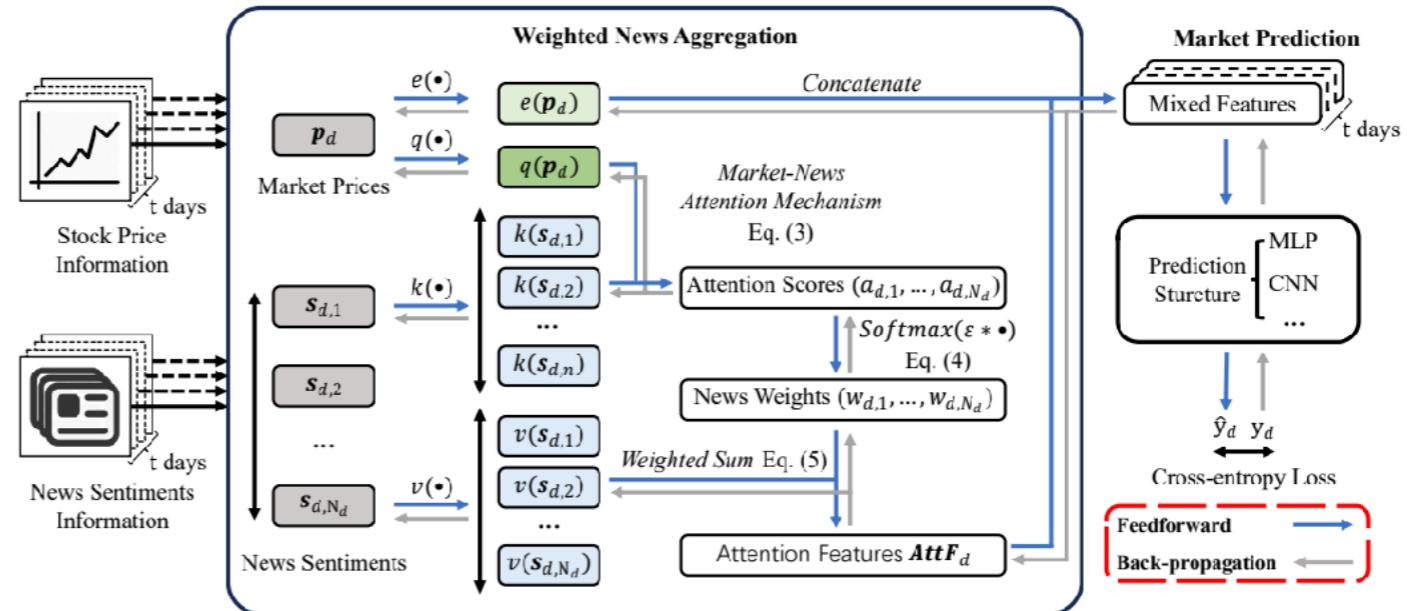


Table 4: Market prediction results of prevalent aggregation methods and advanced prediction methods. Abbreviations for these methods are detailed in Section 6.2. The "Price Only" lines serve as the control group, relying solely on prices for predictions. Results for the GPT-4 on the NASDAQ 100 dataset are exclusive due to its less competitive results on S&P 500 dataset and its limited availability compared to GPT-3.5. Underlined values highlight the highest PnL and SR achieved across all settings.

Method	1-day			3-day			5-day			10-day			20-day		
	Acc	PnL	SR												
Results of S&P 500 Market															
Price Only	0.535	0.026	0.870	0.512	0.030	0.886	0.525	0.024	0.592	0.531	0.029	0.820	0.546	0.026	0.686
CF	0.498	0.024	0.619	0.533	0.008	0.309	0.536	0.015	-0.051	0.530	0.021	0.621	0.526	0.024	0.915
SenF	0.539	0.021	0.767	0.496	0.010	-0.164	0.554	0.030	1.083	0.525	0.038	1.013	0.554	0.045	1.085
SumF	0.512	-0.022	-0.637	0.514	0.019	0.276	0.503	-0.019	-0.109	0.501	0.024	0.814	0.513	0.036	0.853
AF	0.535	0.017	0.515	0.541	0.037	0.943	0.513	-0.004	-0.098	0.509	0.026	0.732	0.515	0.023	0.712
FAF	0.518	0.018	0.389	0.531	0.031	0.845	0.515	-0.016	-0.716	0.511	0.012	0.339	0.503	0.029	0.597
MANA-Net	0.537	0.031	1.110	0.518	0.029	0.978	0.542	0.041	1.331	0.552	0.056	1.624	0.528	0.045	1.533
Ensemble	0.522	0.018	0.886	0.514	0.031	0.672	0.510	0.018	0.824	0.534	0.031	1.343	0.545	0.028	1.028
FPT	0.508	0.002	0.323	0.524	0.033	0.917	0.520	0.042	1.169	0.528	0.025	0.969	0.547	0.035	1.273
GPT-3.5	0.522	0.022	0.914	0.530	0.030	1.133	0.528	0.022	0.946	0.522	0.014	0.765	0.524	0.014	0.769
GPT-4	0.504	0.032	0.696	0.520	0.008	0.194	0.516	0.008	-0.104	0.484	0.003	-0.244	0.494	0.005	-0.109
Results of NASDAQ 100 Market															
Price Only	0.512	-0.021	-0.834	0.498	-0.057	-1.537	0.517	-0.010	-0.176	0.557	0.024	0.950	0.517	-0.008	0.016
CF	0.548	0.010	0.711	0.482	-0.045	-1.227	0.527	-0.003	0.136	0.533	-0.005	0.231	0.507	-0.023	-0.475
SenF	0.528	0.024	0.150	0.510	-0.020	-0.257	0.569	0.023	0.937	0.505	-0.035	-0.594	0.517	-0.016	-0.381
SumF	0.536	0.010	0.570	0.514	-0.019	-0.217	0.545	0.006	0.588	0.527	0.009	0.457	0.540	0.014	0.883
AF	0.504	-0.035	-0.984	0.536	-0.013	0.059	0.521	0.008	0.306	0.513	-0.026	-0.516	0.491	-0.024	-0.884
FAF	0.552	-0.002	0.401	0.536	-0.034	-0.425	0.539	-0.016	-0.101	0.525	-0.016	-0.120	0.527	-0.006	0.059
MANA-Net	0.528	0.004	0.592	0.550	0.023	0.891	0.549	0.020	0.901	0.561	0.031	1.017	0.538	0.041	1.347
Ensemble	0.540	-0.006	0.107	0.534	0.007	0.628	0.546	0.010	0.627	0.532	0.020	0.985	0.542	0.024	1.095
FPT	0.536	0.009	0.523	0.537	0.003	0.457	0.549	0.009	0.766	0.540	0.017	0.837	0.506	0.003	0.322
GPT-3.5	0.538	0.010	0.795	0.522	0.014	0.882	0.526	0.015	0.692	0.530	0.017	0.896	0.522	0.009	0.639

Saliency-Aware Interpolative Augmentation for Multimodal Financial Prediction, COLING'24

Goal: propose a saliency-guided hierarchical Mixup technique to improve multimodal financial prediction using conference call data.

Contributions:

- (1) Introduce SH-Mix, combining modality and fusion-level saliency for data augmentation.
- (2) Demonstrate 3-7% improvement over existing methods through comprehensive experiments.
- (3) Prove SH-Mix's versatility across different tasks, modalities, and model architectures.

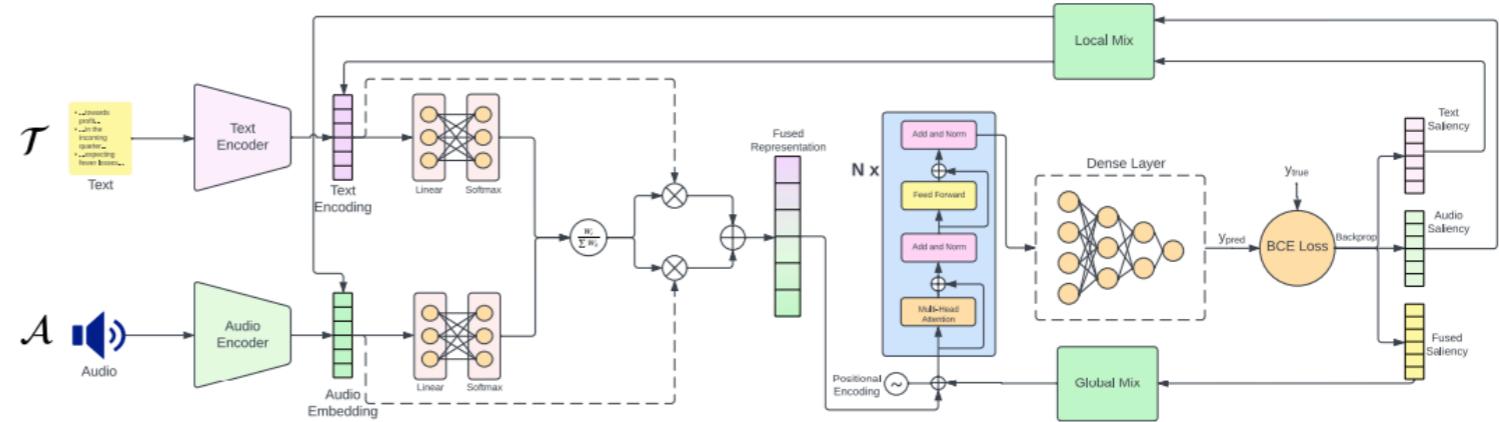


Figure 2: **SH-Mix Overview:** Input text \mathcal{T} and audio \mathcal{A} are encoded to yield respective embeddings. Attention weights for each modality are extracted, followed by fusion through weighted summation. The fused representation is fed to transformer block and a dense layer, to obtain the loss. Saliencies are computed for text, audio, and the fused representation via backpropagation, as detailed in section 3.1. These saliencies drive Local-Mix and Global-Mix, resulting in two sets of mixed inputs.

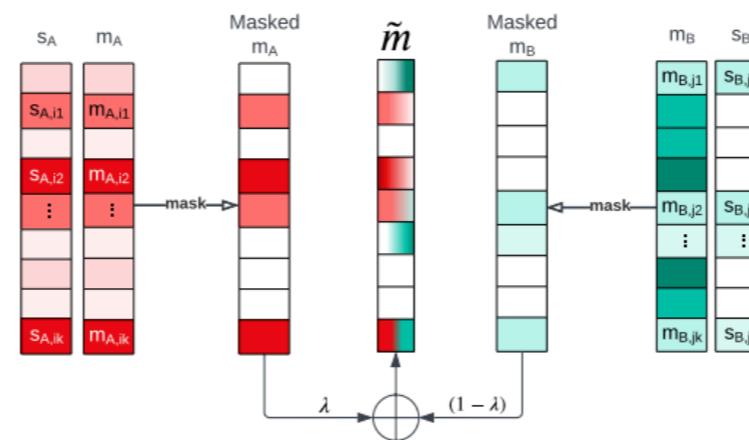


Figure 3: **Local-Mix:** Given any audio/text utterance m_A and m_B for input samples A and B, we mix the most salient portions of m_A with the least salient portions of m_B while zeroing out the remaining features to obtain the mixed utterance \tilde{m} . More salient features are darker.

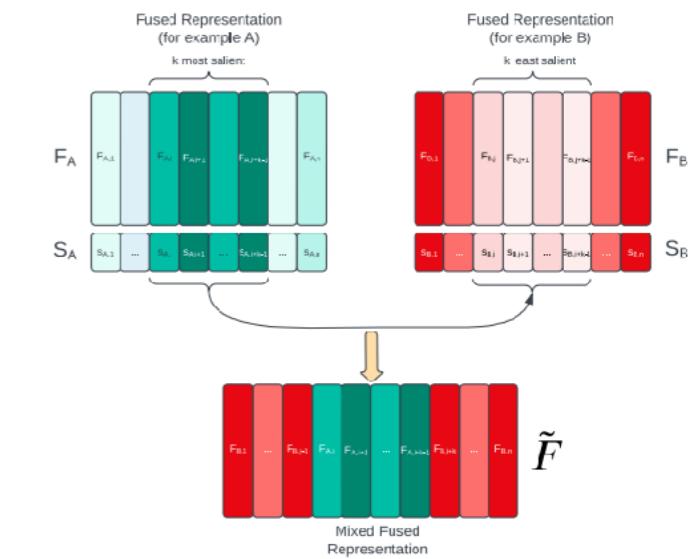


Figure 4: **Global-Mix:** Given fused representations F_A and F_B of input samples A and B, we replace the most salient span in F_A with the least salient span in F_B to obtain the mixed fused representation \tilde{F} . More salient utterances are colored darker.

Ranking-based Selection

CI-STHPAN: Pre-trained Attention Network for Stock Selection with Channel-Independent Spatio-Temporal Hypergraph, AAAI'24

Goal: to develop a pre-trained attention network for stock selection that addresses non-stationary dynamics and complex market dependencies.

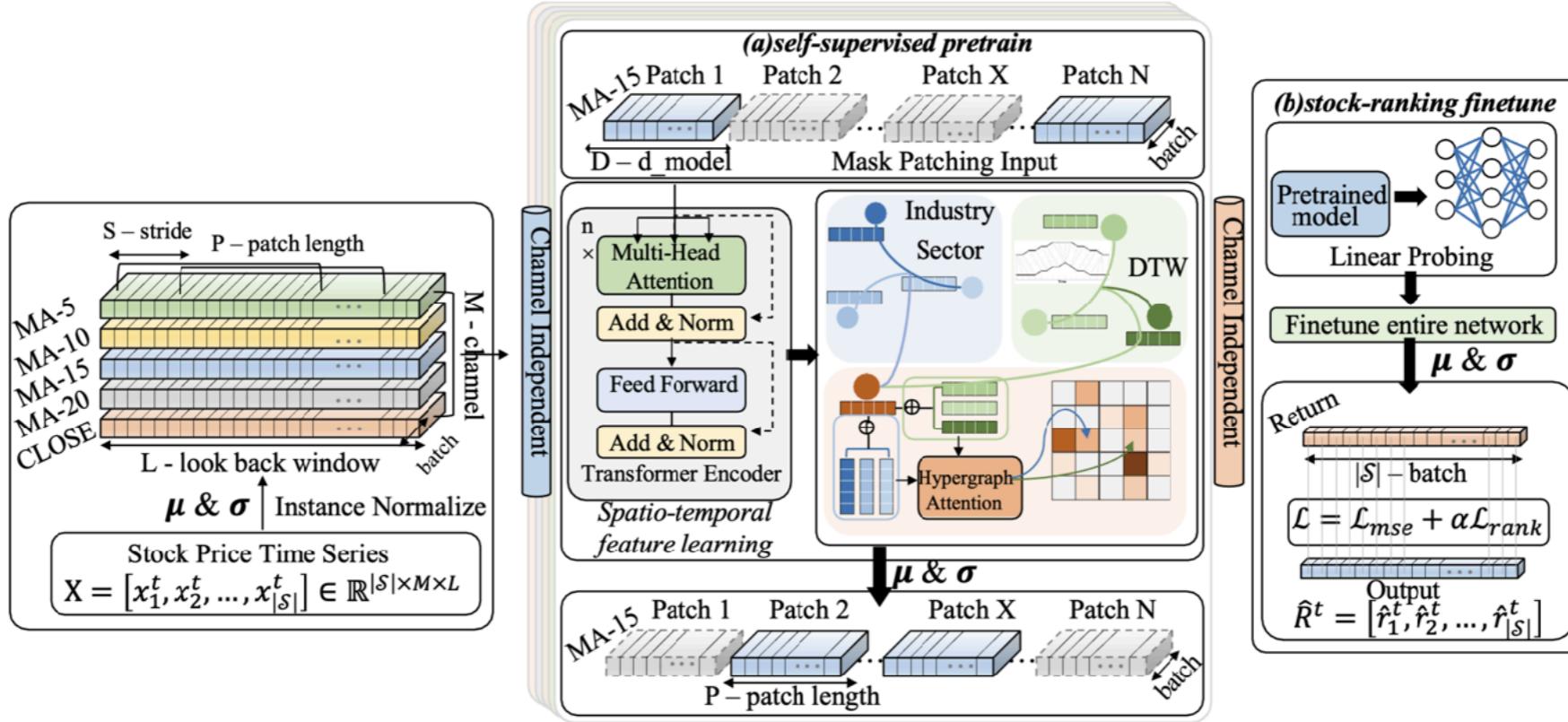


Figure 3: Overview of the proposed CI-STHPAN. A two-stage framework for stock selection, involving Transformer and HGAT based stock time series self-supervised pre-training and stock-ranking based downstream task fine-tuning.

Contributions:

- (1) Introduces a two-stage framework combining self-supervised pre-training with stock-ranking downstream fine-tuning
- (2) Implements channel independence and reversible normalisation to handle financial time series distribution shifts
- (3) Creates dynamic stock hypergraphs based on time series similarity rather than predefined relationships

	Model	Methods	NASDAQ		NYSE	
			IRR	SR	IRR	SR
CLF	ARIMA (Wang and Leu 1996)	RNN with ARIMA Features	0.10	0.55	0.10	0.33
	HGCluster (Luo et al. 2014)	Stock trend to hypergraph clustering	0.10	0.06	0.11	0.10
	Adv-LSTM (Feng et al. 2018)	Adversarial training	0.23	0.97	0.14	0.81
	HATS (Kim et al. 2019)	Diverse stock relationships	0.15	0.80	0.12	0.73
	HMG-TF (Ding et al. 2020)	Multiscale Gaussian prior	0.19	0.83	0.13	0.75
	LSTM-RCGN (Li et al. 2021)	Using news to predict overnight stock	0.13	0.75	0.10	0.70
REG	HATR (Wang et al. 2021)	Multi-scale local combinations	0.31	0.92	0.14	0.76
	SFM (Zhang, Aggarwal, and Qi 2017)	State-frequency memory	0.09	0.16	0.11	0.19
	DA-RNN (Qin et al. 2017)	Two-stage Attention-RNN	0.14	0.71	0.13	0.66
RL	DQN (Carta et al. 2021)	Maximise the gain function	0.20	0.93	0.12	0.72
	iRDPG (Liu et al. 2020)	Intelligent trading agents	0.28	1.32	0.18	0.85
	RAT (Xu et al. 2021a)	Relation-aware transformers	0.40	1.37	0.22	1.03
RAN	SAE-LSTM (Bao, Yue, and Rao 2017)	Wavelet Transform & Stacked Autoencoder	0.22	0.95	0.12	0.79
	RSR-I (Feng et al. 2019)	Temporal graph convolution	0.39	1.34	0.21	0.95
	STHAN-SR (Sawhney et al. 2021)	Spatio-temporal attention hypergraph	0.44	1.42	0.33	1.12
	ALSP-TF (Wang et al. 2022)	Adaptive long-short pattern Transformer	0.53	1.55	0.41	1.24
	CI-STHPAN (Ours)	Spatial-temporal pre-training	0.66	2.01	0.79	2.14

Table 1: Profitability comparison with Classification (CLF), Regression (REG), Reinforcement Learning (RL), and Ranking (RAN) baselines. Bold & underline depict the best & second-best results($p < 0.01$).

Automatic De-Biased Temporal-Relational Modeling for Stock Investment Recommendation, IJCAI'24

Goal: An automatic de-biased model to improve stock investment recommendations by addressing temporal-relational biases and distribution shifts

Contributions:

- (1) Introduces ADB-TRM framework that mitigates both stock-level biases and market distribution shifts
- (2) Enhances model adaptability through adversarial training to improve resilience against market volatility
- (3) Demonstrates superior performance across multiple stock exchanges with significantly improved investment returns

Notes

- Meta: stable correlations (sector and industry)
- Micro-level bias: inter-stocks influence
- Macro-level bias: market's non-stationary properties and economic cycles
- De-biasing: learn the bias and adversarially remove it

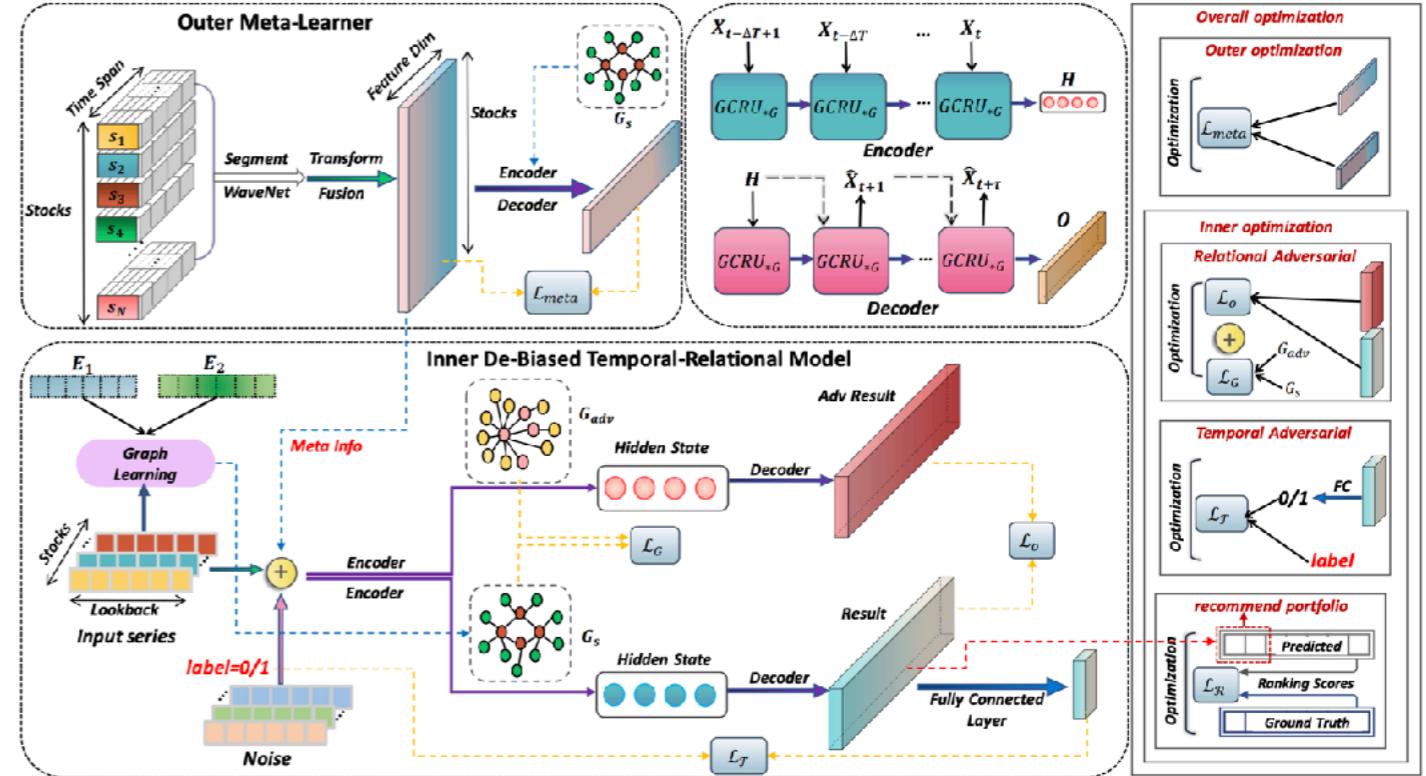


Figure 2: The framework of ADB-TRM. ADB-TRM is mainly divided into two parts: the outer meta-learner and the inner de-biased TRM. The inner part combines GRU with the graph convolutional operation to jointly learn the temporal-relational patterns and applies temporal-relational adversarial training to enhance the model's robustness against temporal-relational biases. The outer meta-learner is responsible for interacting with the inner part and inferring relative invariant information from the global perspective to alleviate the distribution shifts.

	Methods	NASDAQ		NYSE		TSE		
		SR	IRR	SR	IRR	SR	IRR	
CLF	ARIMA [Wang and Leu, 1996]	Auto Regressive Integrated Moving Average to fit non-stationary stock price data	0.55	0.10	0.33	0.10	0.47	0.13
	ALSTM [Feng et al., 2019a]	Adversarial LSTM simulates stock stochasticity during training	0.97	0.23	0.81	0.14	1.10	0.43
	HATS [Kim et al., 2019]	Hierarchical graph attention model to aggregate information from multi-graph	0.80	0.15	0.73	0.12	0.96	0.31
	HMG-TF [Ding et al., 2020]	Enhanced Transformer for learning multi-scale features of Finance data	0.83	0.19	0.75	0.13	1.05	0.33
RL	DQN [Carta et al., 2021]	Annotation-free ensembled RL method for maximizing return function	0.93	0.20	0.72	0.12	1.08	0.31
	iRDPG [Liu et al., 2020]	An enhanced model combines deep reinforcement learning and imitation learning	1.32	0.28	0.85	0.18	1.10	0.55
	RAT [Xu et al., 2021]	Relation-aware Transformer for portfolio selection with reinforcement learning	1.37	0.40	1.03	0.22	1.20	0.64
REG	SFM [Zhang et al., 2017]	State Frequency Memory recurrent network for modeling multi-level time frequency	0.16	0.09	0.19	0.11	0.08	0.07
	MTGNN [Wu et al., 2020]	Adaptive GNN framework with dilated inception module for time-series forecasting	0.82	0.29	0.94	0.17	1.01	0.33
	THGNN [Xiang et al., 2022]	A temporal and heterogeneous GNN based on learning dynamic relations	0.88	0.31	0.78	0.13	1.15	0.41
RAN	RSR-E [Feng et al., 2019b]	Temporal GCN based on similarity measure as relation weight	1.12	0.26	0.88	0.20	1.07	0.50
	RSR-I [Feng et al., 2019b]	Temporal GCN based on neural network to calculate relation weight	1.34	0.39	0.95	0.21	1.08	0.53
	STHAN-SR [Sawhney et al., 2021a]	A temporal-relational hypergraph attentive architecture for stock selection	1.42	0.44	1.12	0.33	1.19	0.62
	HyperStockGAT [Sawhney et al., 2021b]	Hyperbolic graph attention network on the Riemannian Manifolds for stock selection	1.40	0.44	1.10	0.25	1.20	0.75
	ALSP-TF [Wang et al., 2022c]	A temporal-relation adaptive transformer architecture for stock selection	1.55	0.53	1.24	0.41	1.27	0.71
	RT-GCN [Zheng et al., 2023]	A relational-temporal GCN based on three relation-aware strategies	1.49	0.48	1.22	0.37	1.29	0.78
	ADB-TRM (Ours)	Automatic de-biased temporal-relational model for stock selection	1.66	0.66	1.42	0.58	1.38	0.93
	Improve	Improvements over state-of-the-art	7.10%	24.53%	14.52%	41.46%	6.98%	19.23%

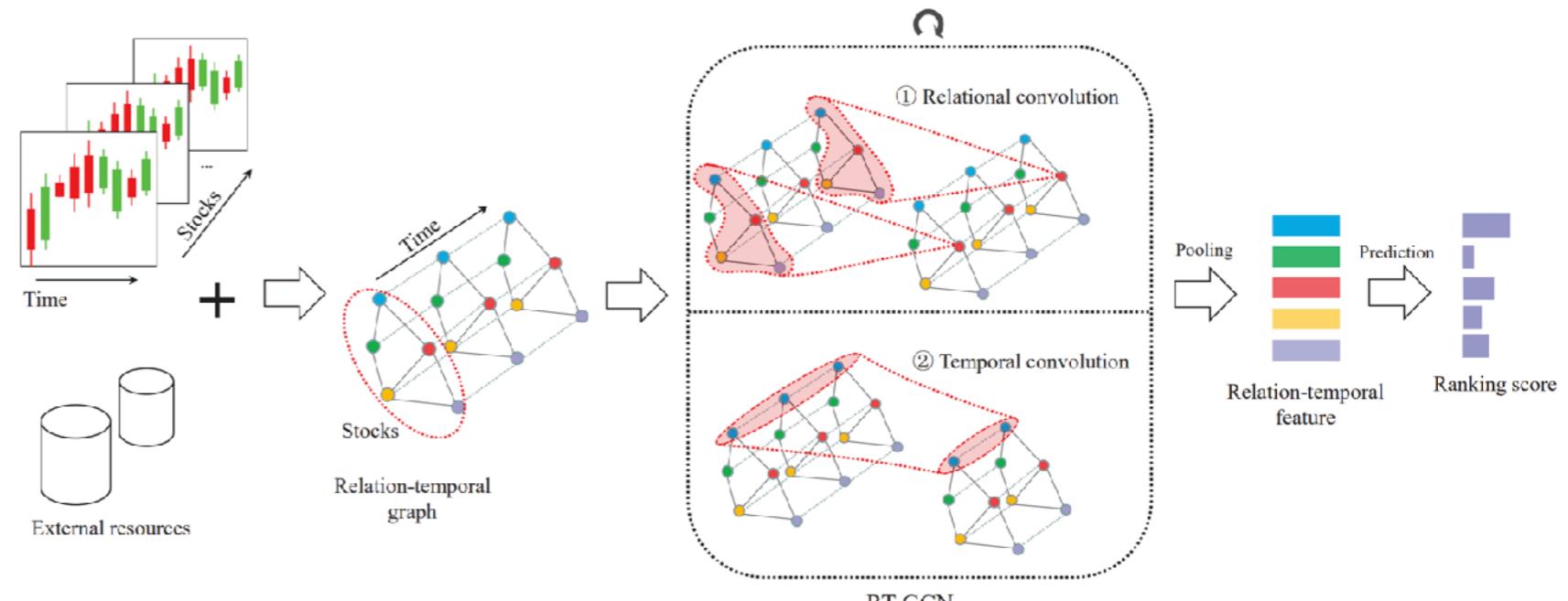
Table 1: Comparison of profitability with Classification (CLF), Reinforcement Learning (RL), Regression (REG), and Ranking (RAN) baselines. The improvement is statistically significant ($p < 0.01$) under Wilcoxon's signed rank test.

Relational Temporal Graph Convolutional Networks for Ranking-Based Stock Prediction, ICDE'23

Goal: Ranking-based stock prediction to maximise investment returns with RT-GCN

Contributions:

- (1) Models stock relations and temporal features as a relation-temporal graph for efficient processing
- (2) Develops a pure convolutional network structure for faster training in stock prediction
- (3) Implements three relation-aware strategies to optimize information propagation between connected stocks



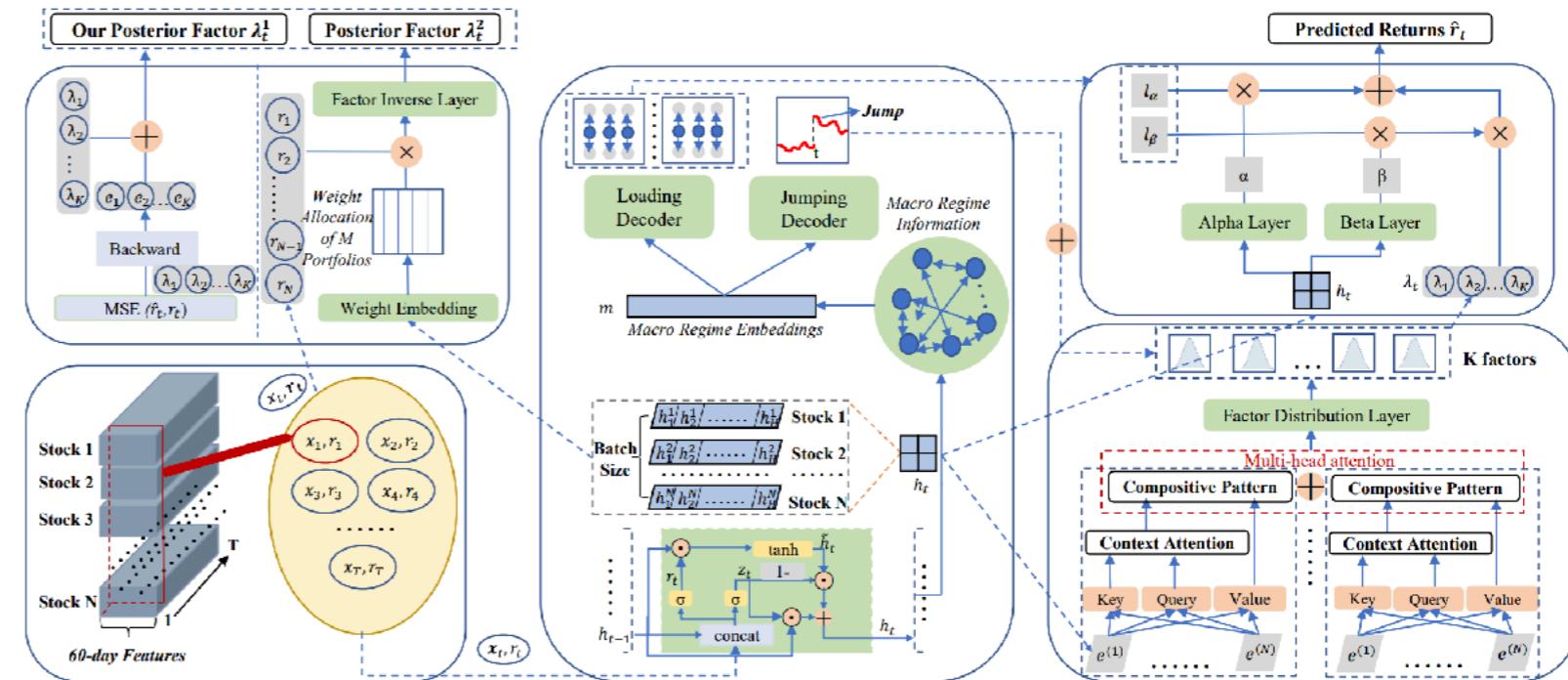
Models	NASDAQ				NYSE				CSI				
	MRR	IRR-1	IRR-5	IRR-10	MRR	IRR-1	IRR-5	IRR-10	MRR	IRR-1	IRR-5	IRR-10	
CLF	ARIMA [14]	-	0.12	0.20	0.23	-	0.10	0.17	0.28	-	-0.13	-0.09	0.07
	A-LSTM [41]	-	0.23	0.40	0.59	-	0.43	0.51	0.69	-	-0.07	0.15	0.11
REG	SFM [1]	0.032	0.11	0.28	0.36	0.029	0.13	0.23	0.31	0.007	-0.21	0.10	0.17
	LSTM [16]	0.027	0.13	0.22	0.28	0.013	0.09	0.15	0.19	0.011	-0.33	0.07	0.12
RL	DQN [18]	0.041	0.20	0.34	0.58	0.033	0.12	0.33	0.60	0.009	0.11	0.13	0.15
	iRDPG [19]	0.049	0.29	0.37	0.46	0.040	0.20	0.28	0.57	0.010	0.12	0.14	0.15
RAN	Rank_LSTM [16]	0.041	0.35	0.42	0.82	0.029	0.30	0.64	0.63	0.016	-0.24	0.14	0.19
	RSR_I [9]	0.043	0.81	0.93	0.88	0.032	0.66	0.98	1.01	0.015	-0.04	0.29	0.33
	RSR_E [9]	0.055	0.89	0.83	0.84	0.048	0.88	0.92	0.95	0.021	0.29	0.26	0.28
	RT-GAT [31]	0.049	0.11	0.67	0.87	0.044	0.48	0.87	0.97	0.027	0.27	0.30	0.25
Ours	RT-GCN (U)	0.055	0.45	0.71	0.82	0.040	0.65	0.88	0.96	0.017	0.13	0.15	0.19
	RT-GCN (W)	0.059	0.63	0.86	0.98	0.047	0.80	0.95	0.99	0.022	0.29	0.26	0.35
	RT-GCN (T)	0.061	1.25	0.97	1.03	0.056	0.92	1.10	1.13	0.031	0.35	0.35	0.38
Improvement		10.9%	40.4%	4.3%	17.0%	16.7%	4.5%	12.2%	11.9%	14.8%	20.7%	16.7%	15.2%
p-value		3.05e-4	9.16e-5	6.10e-5	3.05e-4	1.52e-4	4.27e-4	7.63e-4	3.05e-5	3.05e-5	1.53e-4	6.10e-5	4.82e-4

RSAP-DFM: Regime-Shifting Adaptive Posterior Dynamic Factor Model for Stock Returns Prediction, IJCAI'24

Goal: proposes a dynamic factor model that predicts stock returns by adapting to macroeconomic regimes through dual shifting mechanisms

Contributions:

- (1) Introduces gradient-based posterior factor construction and bilevel optimization for more robust factor modelling
- (2) Develops first dual regime shifting structure to explicitly model macroeconomic impacts continuously
- (3) Achieves state-of-the-art performance on A-share market datasets through comprehensive empirical validation



Model	CSI100			CSI300			CSI500		
	IC	ICIR	RANK IC	IC	ICIR	RANK IC	IC	ICIR	RANK IC
XGBoost	0.0517	0.3096	0.0547	0.3744	0.0517	0.4642	0.0596	0.5696	0.0517
LightGBM	0.0502	0.3211	0.0595	0.3858	0.0560	0.4703	0.0654	0.5608	0.0555
MLP	0.0413	0.2466	0.0427	0.2676	0.0358	0.2587	0.0427	0.3246	0.0353
Transformer	0.0475	0.2668	0.0578	0.3435	0.0415	0.3094	0.0564	0.428	0.0396
ALSTM	0.0574	0.3340	0.0655	0.3866	0.0658	0.5394	0.0763	0.6231	0.0586
GATS	0.0583	0.3526	0.0678	0.4208	0.0717	0.5655	0.0804	0.6520	0.0680
AdaRNN	0.0657	0.3865	0.0697	0.4222	0.0640	0.5077	0.0769	0.6148	0.0705
IGMTF	0.0622	0.3678	0.0708	0.4352	0.0669	0.5250	0.0765	0.6196	0.0653
TRA	0.0674	0.4157	0.0755	0.4784	0.0638	0.5207	0.0731	0.6084	0.0688
TCTS	0.0593	0.3429	0.0632	0.3738	0.0763	0.6633	0.0806	0.6828	0.0718
FactorVAE	0.0487	0.3084	0.0518	0.3321	0.0528	0.4417	0.0627	0.5350	0.0606
RSAP-DFM	0.0768	0.4260	0.0870	0.4852	0.0855	0.6194	0.0961	0.6849	0.0854

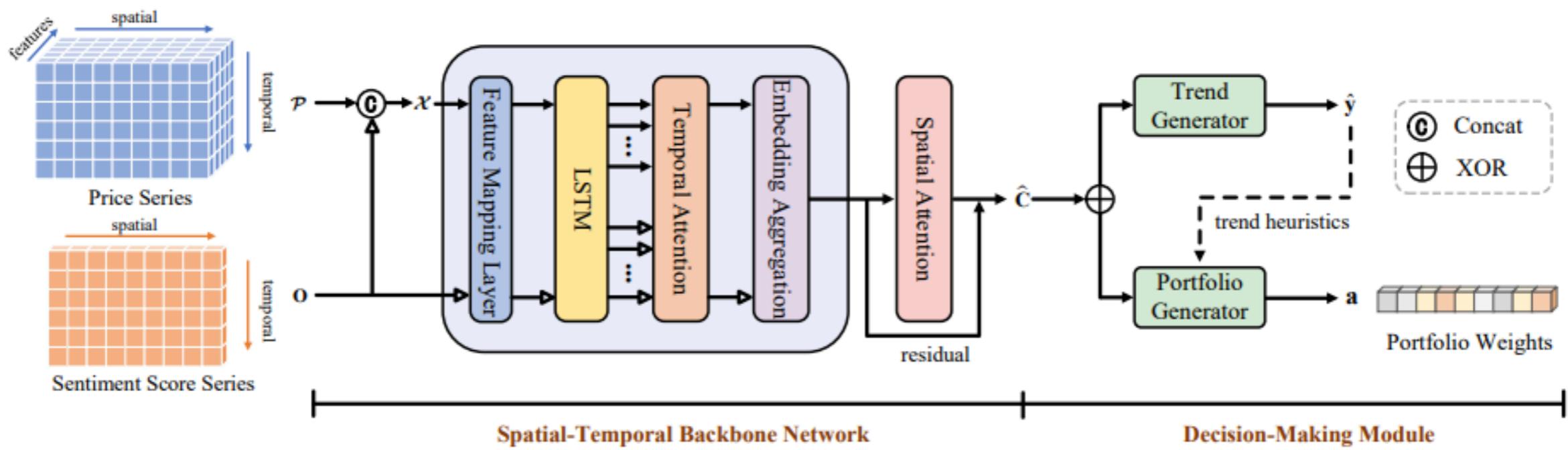
Portfolio Optimisation

Developing A Multi-Agent and Self-Adaptive Framework with Deep Reinforcement Learning for Dynamic Portfolio Risk Management, AAMAS'24

- Goal: dynamically balance the trade-off between portfolio returns and risk management
 - Contributions: (1) Novel Multi-Agent Framework (MASA) (2) Reward-Based Guiding Mechanism (3) Superior experimental results
 - Cooperative agents:
 1. Trend observing agent
 2. Return optimising agent
 3. Risk managing agent
 - Reward-based guiding mechanism with return reward (J_r) and Action reward (J_{JS}), helps maintain diversity in generated action sets while adapting to volatile market conditions
 - Variants of MASA outperform SOTA in CSI300, DJIA and S&P500
-
- The diagram illustrates the MASA framework. It features three main agents: a Market Observer (yellow box), an RL-based Agent (blue box), and a Solver-based Agent (orange box). The Market Observer receives Market States from the Trading Environment and provides signals σ_s and v_m to both the RL-based Agent and the Solver-based Agent. The RL-based Agent uses a Softmax function to generate an action a^{RL} . The Solver-based Agent performs optimization based on constraints like $\frac{1}{2}x^T Px + q^T x$ and $Gx + s = h$, producing a control action a^{ctrl} . These actions are combined at a central point to produce the final action a^{Final} . The Trading Environment provides Market States to all agents. Rewards are used to guide the agents: a Return Reward J_r is added to the RL-based Agent's output, and an Action Reward $J_{JS(a^{RL}||a^{Final})}$ is used to refine the final action.
- | Market | CSI 300 | | | | DJIA | | | | S&P 500 | | | |
|------------------|-------------|--------------|-------------|---------------|--------------|--------------|-------------|---------------|--------------|--------------|-------------|---------------|
| Models | AR(%) ↑ | MDD(%) ↓ | SR↑ | Risk↓ | AR(%) ↑ | MDD(%) ↓ | SR↑ | Risk↓ | AR(%) ↑ | MDD(%) ↓ | SR↑ | Risk↓ |
| CRP | 7.19 | 33.96 | 0.19 | 0.0131 | 11.44 | 19.66 | 0.58 | 0.0095 | 18.09 | 37.12 | 0.65 | 0.0143 |
| EG | 7.19 | 33.94 | 0.19 | 0.0131 | 11.39 | 19.66 | 0.57 | 0.0095 | 18.03 | 36.88 | 0.65 | 0.0142 |
| OLMAR | -3.50 | 55.67 | -0.17 | 0.0217 | -13.89 | 59.89 | -0.54 | 0.0179 | 29.34 | 68.52 | 0.55 | 0.0275 |
| PAMR | -14.24 | 49.32 | -0.47 | 0.0223 | -37.72 | 81.72 | -1.35 | 0.0174 | 3.63 | 59.08 | 0.05 | 0.0263 |
| CORN | -2.06 | 59.28 | -0.13 | 0.0219 | 1.62 | 41.76 | 0.00 | 0.0122 | -6.90 | 62.97 | -0.21 | 0.0200 |
| RMR | 5.77 | 41.96 | 0.07 | 0.0215 | -12.98 | 61.23 | -0.51 | 0.0180 | -4.12 | 85.48 | -0.09 | 0.0280 |
| EIII | 6.84 | 31.77 | 0.18 | 0.0122 | 10.81 | 18.24 | 0.58 | 0.0089 | 16.50 | 35.80 | 0.63 | 0.0135 |
| PPN | 6.74 | 31.19 | 0.18 | 0.0119 | 10.50 | 17.95 | 0.57 | 0.0087 | 16.65 | 34.17 | 0.65 | 0.0130 |
| RAT | 6.78 | 31.48 | 0.18 | 0.0121 | 10.62 | 18.19 | 0.57 | 0.0088 | 17.03 | 34.92 | 0.65 | 0.0133 |
| TD3-Profit | 7.18 | 33.97 | 0.19 | 0.0128 | 11.45 | 19.65 | 0.58 | 0.0095 | 18.09 | 37.12 | 0.65 | 0.0143 |
| MASA-MLP | 8.87 | 31.78 | 0.27 | 0.0119 | 13.17 | 19.89 | 0.69 | 0.0088 | 22.49 | 26.50 | 0.92 | 0.0116 |
| MASA-LSTM | 8.72 | 31.83 | 0.26 | 0.0121 | 13.50 | 19.58 | 0.71 | 0.0087 | 22.12 | 26.61 | 0.90 | 0.0117 |
| MASA-DC | 8.70 | 31.77 | 0.25 | 0.0120 | 15.52 | 16.21 | 0.80 | 0.0086 | 14.88 | 24.29 | 0.60 | 0.0112 |

Trend-Heuristic Reinforcement Learning Framework for News-Oriented Stock Portfolio Management, ICASSP'24

- Goal: fuse stock trend and financial news information in portfolio management
- Contributions: (1) multimodal information fusion (2) spatial-temporal backbone + RL decision making module (3) significant improvement against SOTA in ARR, ASR & MDD for DJIA & SSE-50



- Sentiment: a measure of investors' collective view (market outlook, company news, etc.), by FinGPT
- **Spatial**: different stocks; **Temporal**: different time steps/periods
- Temporal model: LSTM; Spatial model: self-attention
- Decision-making: (1) $\hat{C} \rightarrow \hat{y}$; (2) $(\hat{C}, \hat{y}) \rightarrow a$

Dataset	DJIA				SSE 50				
	Methods	ARR(%)	AVol	ASR	MDD(%)	ARR(%)	AVol	ASR	MDD(%)
Market		14.35	0.188	0.764	33.13	10.71	0.217	0.490	34.36
CRP		14.69	0.187	0.794	33.09	12.07	0.195	0.626	26.25
ONS		14.05	0.230	0.611	34.59	16.96	0.209	0.813	31.34
RMR		16.64	0.356	0.474	75.52	12.50	0.454	0.286	66.07
EIIIE		26.15	0.215	1.099	26.97	23.40	0.216	1.039	23.77
SARL		24.80	0.284	1.003	30.01	20.60	0.230	0.915	14.28
DeepTrader		28.59	0.250	1.117	28.63	31.55	0.307	1.089	17.09
Ours		33.45	0.234	1.206	25.33	35.34	0.286	1.105	19.75

Reinforcement Learning with Maskable Stock Representation for Portfolio Management in Customizable Stock Pools, WWW'24

- Goal: adapting RL to handle portfolio management with customised stock pools
- Contributions: (1) Learnable masks for unfavourable stocks (2) Meaningful self-supervised stock embeddings (3) Re-weighting mechanism on favourable stocks

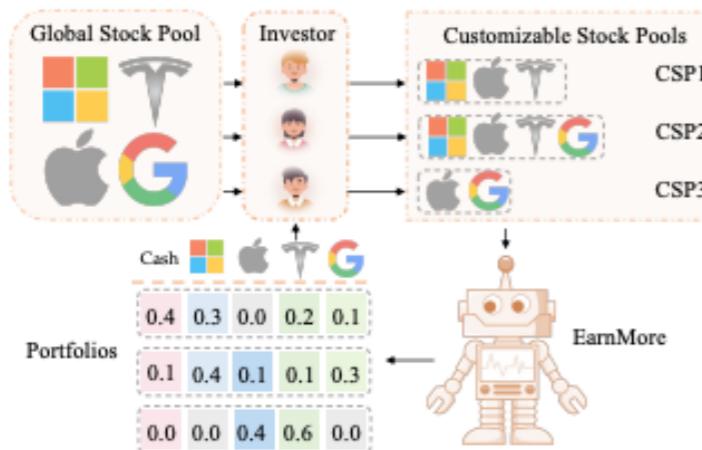


Figure 1: Overview of portfolio management by EarnMore in customizable stock pools (CSPs).

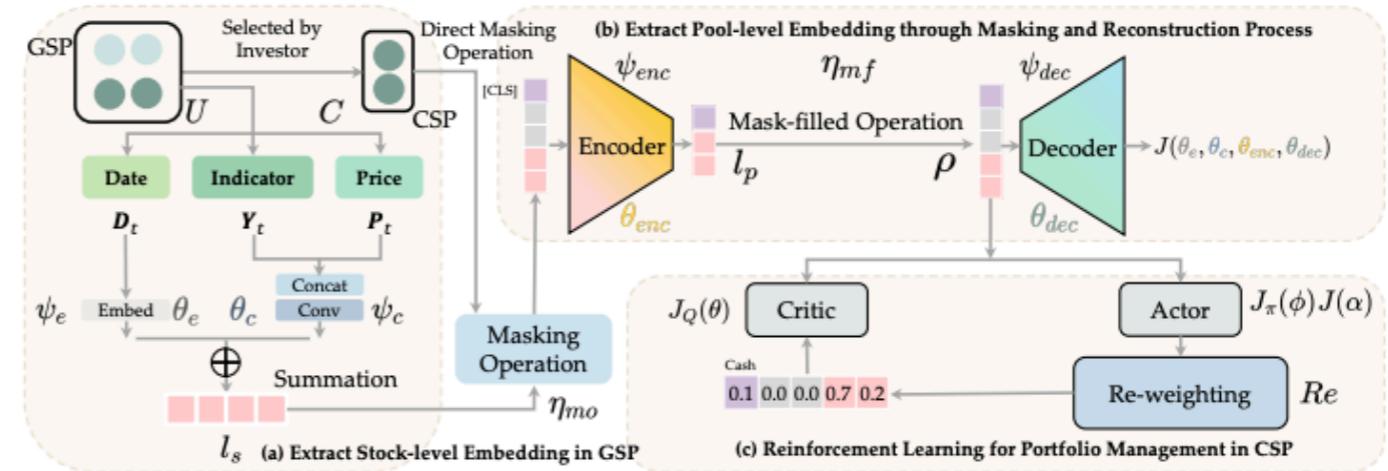


Figure 2: The overall architecture of EarnMore. Module (a) is used to extract stock-level embeddings from GSP. Module (b) is the masking and reconstruction process to learn pool-level embeddings. Module (c) is an agent with masked token awareness.

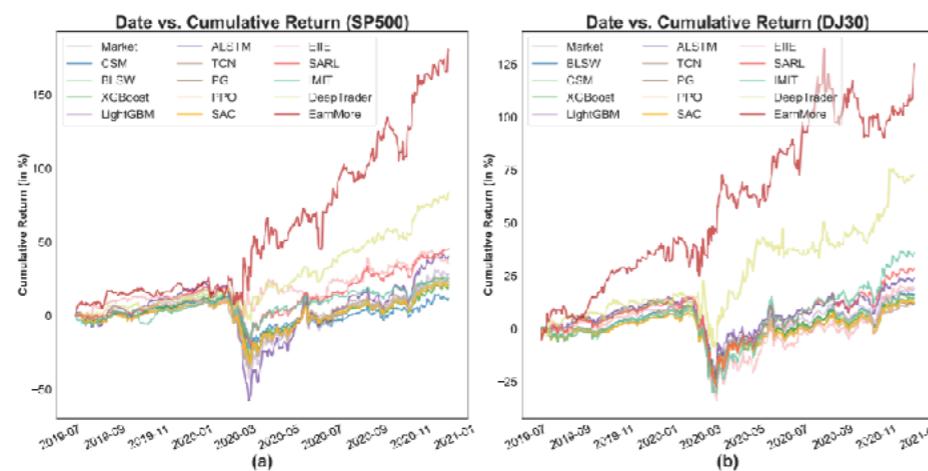


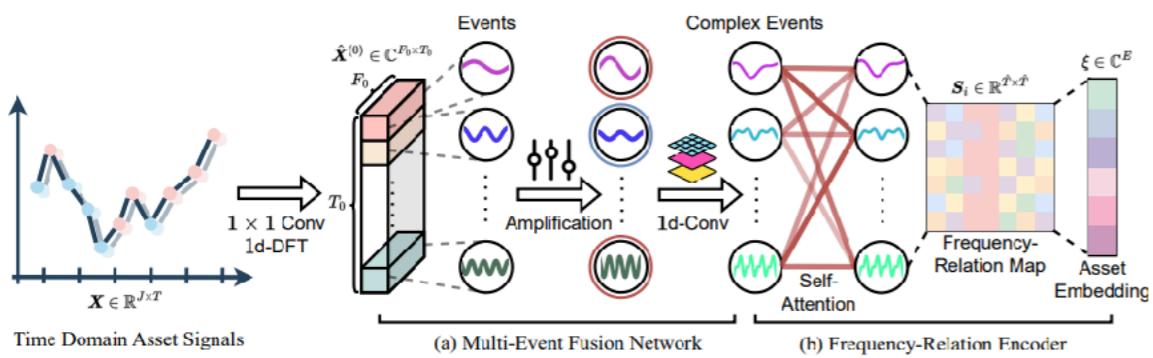
Figure 3: Performance on GSP for SP500 and DJ30

Pool	Strategies	Stock Index		SP500		DJ30	
		ARR%↑	SR↑	ARR%↑	SR↑	ARR%↑	SR↑
CSP1	SARL	34.330	0.820	24.140	0.638		
	IMIT	20.973	0.860	20.071	0.920		
	DeepTrader	34.030	0.793	27.740	0.757		
	EarnMore	122.610	2.278	53.990	1.810		
CSP2	SARL	17.000	0.570	20.020	0.820		
	IMIT	7.971	0.486	11.841	0.751		
	DeepTrader	34.030	0.793	38.470	0.955		
	EarnMore	110.110	2.279	43.400	1.549		
CSP3	SARL	18.090	0.760	10.910	0.480		
	IMIT	21.193	1.220	6.851	0.496		
	DeepTrader	61.320	1.489	16.840	0.601		
	EarnMore	93.670	2.120	43.460	1.572		

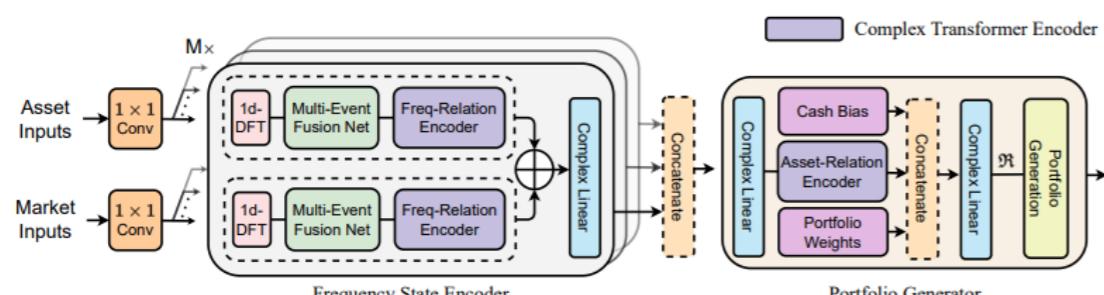
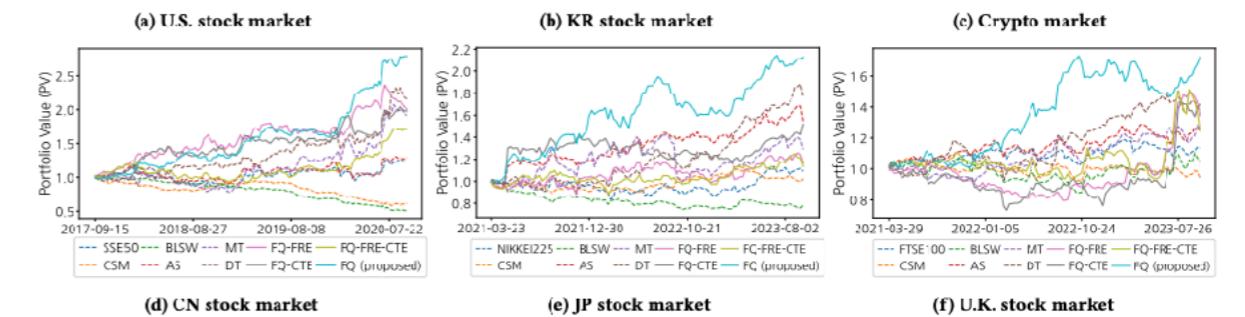
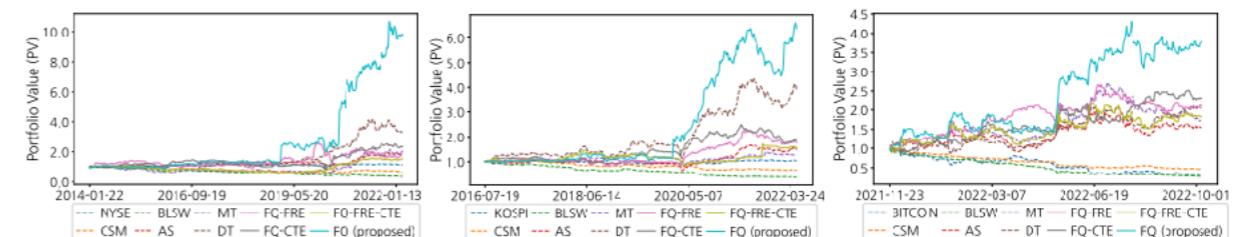
Table 3: Performance comparison on SP500 and DJ30 with Customizable Stock Pools. Underlined metrics indicate the best-performing results.

FreQuant: A Reinforcement-Learning based Adaptive Portfolio Optimization with Multi-frequency Decomposition, KDD'24

- Goal: Adaptive Portfolio Optimization via Multi-Frequency Quantitative Analysis
- Contributions: (1) DRL by frequency (2) Realistic formulation with transaction fees (3) Superior performance on diverse market datasets with up to 2x higher ARR



Dataset	U.S. stock market	KR stock market	Crypto market	CN stock market	JP stock market	U.K. stock market				
Methodology	ARR(%)↑	ASR↑	ARR(%)↑	ASR↑	ARR(%)↑	ASR↑	ARR(%)↑	ASR↑	ARR(%)↑	ASR↑
CSM [19]	-3.774	-0.238	-6.279	-0.545	-84.36	-3.122	-15.86	-1.358	1.621	0.139
BLSW [28]	-11.51	-0.706	-16.64	-1.400	-145.0	-4.397	-21.79	-1.722	-9.612	-0.817
AlphaStock [43]	9.141	0.495	6.922	0.363	70.56	0.966	11.63	0.486	20.06	1.089
DeepTrader [45]	16.80	0.810	27.84	0.994	102.6	1.598	29.40	1.154	26.22	1.190
MetaTrader [26]	13.69	0.429	5.509	0.368	82.17	1.208	27.31	0.940	15.05	0.526
FQ-FRE-CTE	7.213	0.360	10.17	0.427	93.32	1.318	20.25	1.015	6.697	0.399
FQ-FRE	12.69	0.434	12.06	0.570	99.28	1.905	25.94	1.219	8.769	0.404
FQ-CTE	13.21	0.542	13.42	0.663	116.7	1.834	25.16	1.284	20.29	0.835
FQ (proposed)	36.25	0.770	38.57	1.008	179.4	2.370	36.45	2.109	41.93	1.263



Automate Strategy Finding with LLM in Quant investment, ArXiv'24

Goal: Propose an LLM-based framework for automated alpha mining and portfolio management in quantitative investing.

Contributions:

- (1) Develops novel LLM framework to extract alpha factors from multimodal financial information
- (2) Introduces multi-agent system to evaluate market conditions and alpha factor relationships
- (3) Creates dynamic weight optimization to adapt strategies across varying market environments

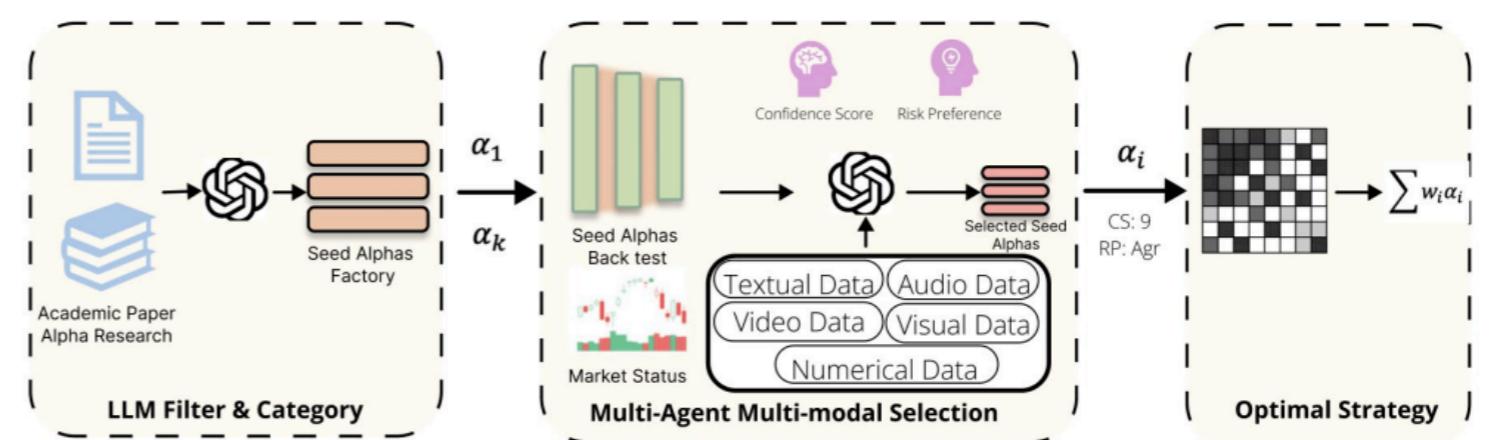
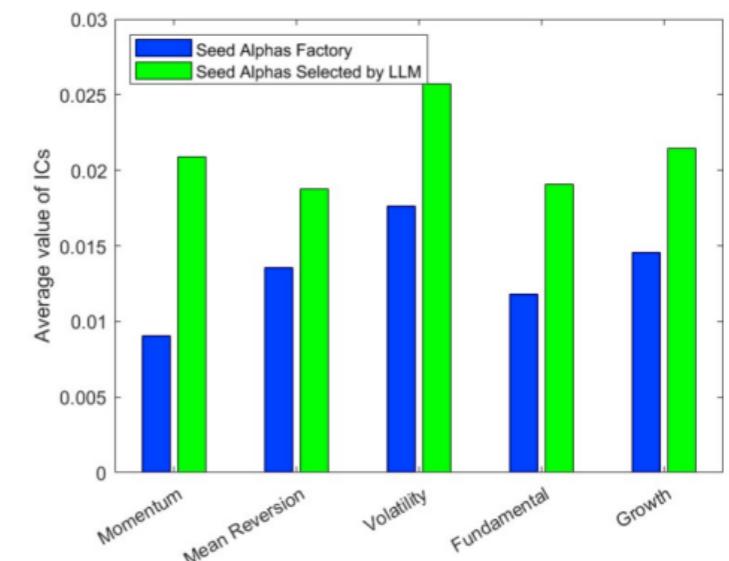


Figure 3: Overview of the strategy generate process (CS stands for confidence score; RP stands for risk preference).

Data Type	Description	Examples
Textual Data	Financial reports, academic papers, news articles, and other textual documents.	Trading forums' sentiment analysis and stock predictions, company disclosures, financial statements, Sina Finance
Numerical Data	Historical stock market data, financial metrics, and performance indicators.	Returns, log returns, annualized returns, volatility
Visual Data	Charts, graphs, and other visual representations of financial data.	Kline charts, trading charts
Audio Data	Financial news broadcasts.	Financial morning news radio, stock review radio, market discussion radio
Video Data	Financial news channels.	CCTV Securities Information Channel, CCTV News Broadcast (news affecting China's stock market)

#	Alpha	Weight	IC(SSE50)
1	(CLOSE - DELAY(CLOSE, 14))	-0.1459	0.0209
2	(RSI - DELAY(RSI, 14))	-1.0265	-0.0225
3	(CLOSE - DELAY(SMA(CLOSE, 14), 7))	-0.1978	0.0193
4	(MA(CLOSE, 20) - CLOSE)	0.0556	-0.0186
5	(SMA(CLOSE, 20) - CLOSE)	-0.945	-0.0186
6	(MAX(HIGH, 20) - CLOSE)	-0.4053	-0.0185
7	(100-RSI)	-0.3199	0.0194
8	(BOLL_UP - BOLL_DOWN) / SMA(CLOSE, 20)	3.6186	0.0278
9	STD(CLOSE, 10) / STD(CLOSE, 50)	-0.183	0.0236
10	VOLUME / MARKET_CAP	-3.2145	-0.0194
11	VOLUME * CLOSE	-0.0058	0.0187
12	(EPS / DELAY(EPS, 1) - 1)	-1.8351	-0.0215
Weighted Combination			-0.0587



Mitigating Extremal Risks: A Network-Based Portfolio Strategy, ArXiv'24

Goal: Propose a portfolio strategy using network theory and extremal dependence to mitigate extreme market risks.

Contributions:

- (1) Develops network model reflecting extremal dependencies between stocks using extreme value theory
- (2) Utilizes maximum independent sets and community detection for optimal portfolio diversification
- (3) Demonstrates effectiveness of community-based portfolio strategy versus sector-based and market approaches

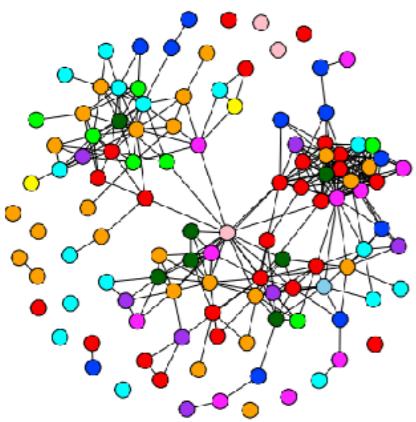


Figure 3: A stock network is classified into 11 industry sectors, with each sector represented by a different color.

Algorithm 1 Portfolio construction using EDM.

Input: Adjusted price of each stock $P_i(t)$, $i = 1, \dots, n$, at time t .

- Step 1: Compute the log return $r_i(t)$ for stock i , and then calculate the pairwise EDM based on Eq.(7);
- Step 2: Denote each stock as a vertex, and use a threshold-based approach to construct networks;
- Step 3: Divide networks into clusters using a proper criterion such as by sector or by community;
- Step 4: Solve for the maximum independent set of each cluster;
- Step 5: Use risk measurement indicators such as VaR or ES for each maximum independent set, and construct a portfolio optimization model by minimizing the overall risk.

Output: Optimal portfolio with minimum risk.

Table 1: Comparison of network parameters of different thresholds.

Threshold	Isolated vertex	Average degree	Network diameter	Graph density	Average clustering coefficient	Average path length
0.05	0	79.30973	0.15605	0.70812	0.80294	0.09323
0.1	2	26.63717	0.44465	0.23783	0.53524	0.21363
0.15	13	5.71681	1.61079	0.05104	0.53920	0.61485
0.2	61	1.16814	1.05244	0.01043	0.48649	0.40297
0.25	98	0.19469	0.53741	0.00174	0.75000	0.31666

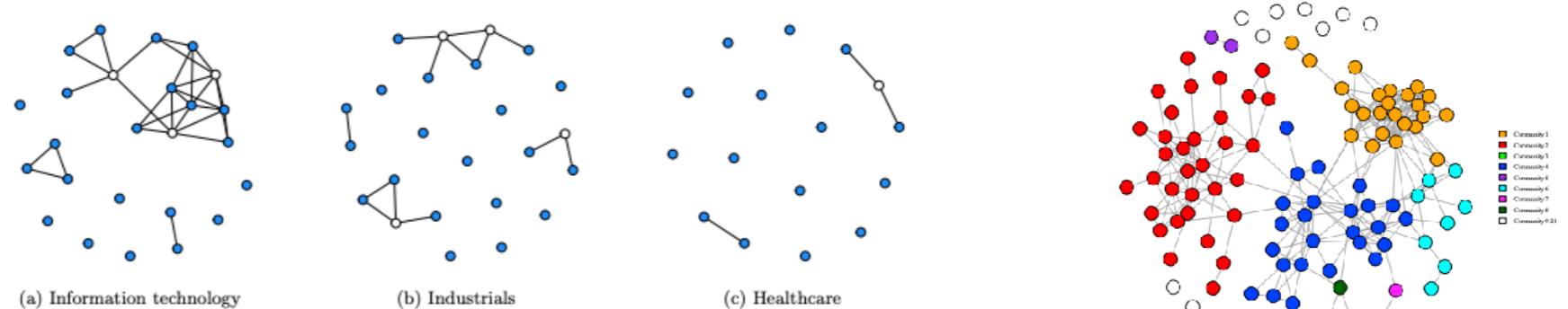


Figure 4: Sub-networks extracted from the original network; the blue nodes represent the maximum independent set of results of the sub-network by sector.

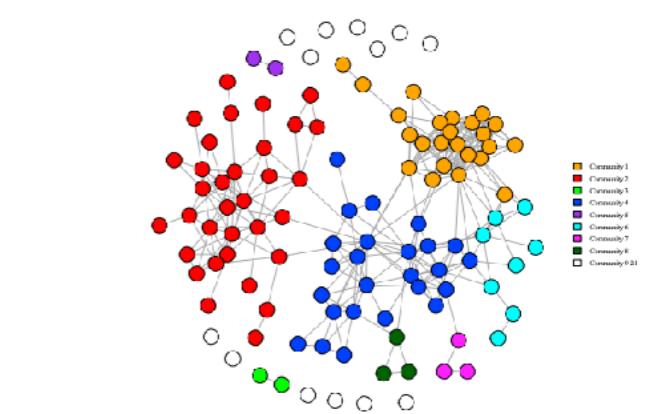


Figure 5: The stock network consists of 13 communities, with each community represented by a different color.

Risk-Managed Sparse Index Tracking Via Market Graph Clustering, ICASSP'24

Goal: Propose a market graph-based portfolio optimization method for better risk-managed index tracking.

Contributions:

- (1) Created a risk-managed tracking method using market-graph neutrality and turnover sparsity.
- (2) Developed algorithm enforcing strict investment diversity across market-graph derived clusters.
- (3) Demonstrated improved performance compared to sector-based neutrality through empirical validation.

$$\begin{aligned} \min_{\mathbf{w}} \frac{1}{T} \|\mathbf{r}^b - \mathbf{X}\mathbf{w}\|_2^2 + \lambda \|\mathbf{w} - \mathbf{w}_{pre}\|_1, \\ \text{s.t. } \begin{cases} \|\mathbf{w}_{C_i}\|_0 \leq k_i, & \forall i \in \{1, \dots, K\}, \\ \mathbf{1}^\top \mathbf{w}_C = b_i, & \forall i \in \{1, \dots, K\}, \\ \mathbf{w} \in S_{l,u}, \end{cases} \end{aligned} \quad (5)$$

Algorithm 1 PDS-based algorithm for solving (5)

Input: \mathbf{r}^b, \mathbf{X}
Output: Output signal \mathbf{w}
Initialization: $\mathbf{w} = \mathbf{0}$

```

1: while A stopping criterion is not satisfied do
2:    $\mathbf{w}^{(n+1)} = \text{prox}_{\iota_{S_k}}(\mathbf{w}^{(n)} - \gamma_1(-\frac{2}{T}\mathbf{X}^\top(\mathbf{r}^b - \mathbf{X}\mathbf{w}^{(n)})) +$ 
       $\mathbf{v}_1^{(n)} + \mathbf{v}_2^{(n)} + \mathbf{v}_3^{(n)})$ 
3:    $\mathbf{v}_1^{(n)} \leftarrow \mathbf{v}_1^{(n)} + \gamma_2(2\mathbf{w}^{(n+1)} - \mathbf{w}^{(n)})$ 
4:    $\mathbf{v}_2^{(n)} \leftarrow \mathbf{v}_2^{(n)} + \gamma_2(2\mathbf{w}^{(n+1)} - \mathbf{w}^{(n)})$ 
5:    $\mathbf{v}_3^{(n)} \leftarrow \mathbf{v}_3^{(n)} + \gamma_2(2\mathbf{w}^{(n+1)} - \mathbf{w}^{(n)})$ 
6:    $\mathbf{v}_1^{(n+1)} = \mathbf{v}_1^{(n)} - \gamma_2 \text{prox}_{\frac{\lambda}{\gamma_2} \|\cdot - \mathbf{w}_{pre}\|_1}(\frac{1}{\gamma_2} \mathbf{v}_1^{(n)})$ 
7:    $\mathbf{v}_2^{(n+1)} = \mathbf{v}_2^{(n)} - \gamma_2 P_{S_b}(\frac{1}{\gamma_2} \mathbf{v}_2^{(n)})$ 
8:    $\mathbf{v}_3^{(n+1)} = \mathbf{v}_3^{(n)} - \gamma_2 P_{S_{l,u}}(\frac{1}{\gamma_2} \mathbf{v}_3^{(n)})$ 
9:    $n \leftarrow n + 1$ 
10: end while
11: return  $\mathbf{w}^{(n)}$ 

```

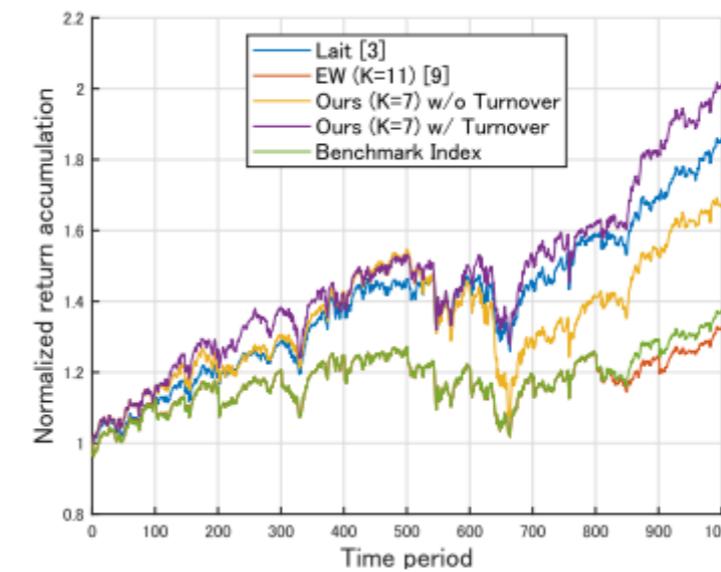


Table 1. The tracking performance measured in MDTE[bps] and Sharpe ratio. Dataset: S&P500, initial capital: \$10000.

Time period	Metric	Turnover	Sector neutrality		Market-graph neutrality					
			Lait [3]	Ours	EW [9] K = 11	Ours K = 7	Ours K = 9	Ours K = 11	Ours K = 13	Ours K = 15
2012-2017	MDTE	w/o	0.60	1.07	0.33	1.23	0.95	1.08	1.07	1.23
		w/	-	0.86	-	0.77	0.75	0.75	0.72	0.70
	Sharpe ratio	w/o	1.21	0.73	0.60	0.65	0.95	0.87	0.78	0.88
		w/	-	0.99	-	1.30	1.16	1.20	1.18	0.98
2017-2022	MDTE	w/o	0.78	1.68	0.68	1.74	1.55	1.53	1.78	1.80
		w/	-	1.78	-	1.77	1.55	1.37	1.75	1.60
	Sharpe ratio	w/o	0.28	0.31	0.14	0.37	0.29	0.28	0.29	0.31
		w/	-	0.30	-	0.30	0.35	0.34	0.34	0.37

Quantitative Trading

StockFormer: Learning Hybrid Trading Machines with Predictive Coding, IJCAI'23

Goal: Fuse market data with future trends and assets correlations in trading policies

Contributions:

- (1) Introduces diversified multi-head attention to better capture diverse temporal patterns in concurrent financial time series
- (2) Develops three specialized Transformer branches to learn long-term, short-term and relational market representations
- (3) Creates a unified decision framework combining predictive states with RL for flexible trading policy optimization

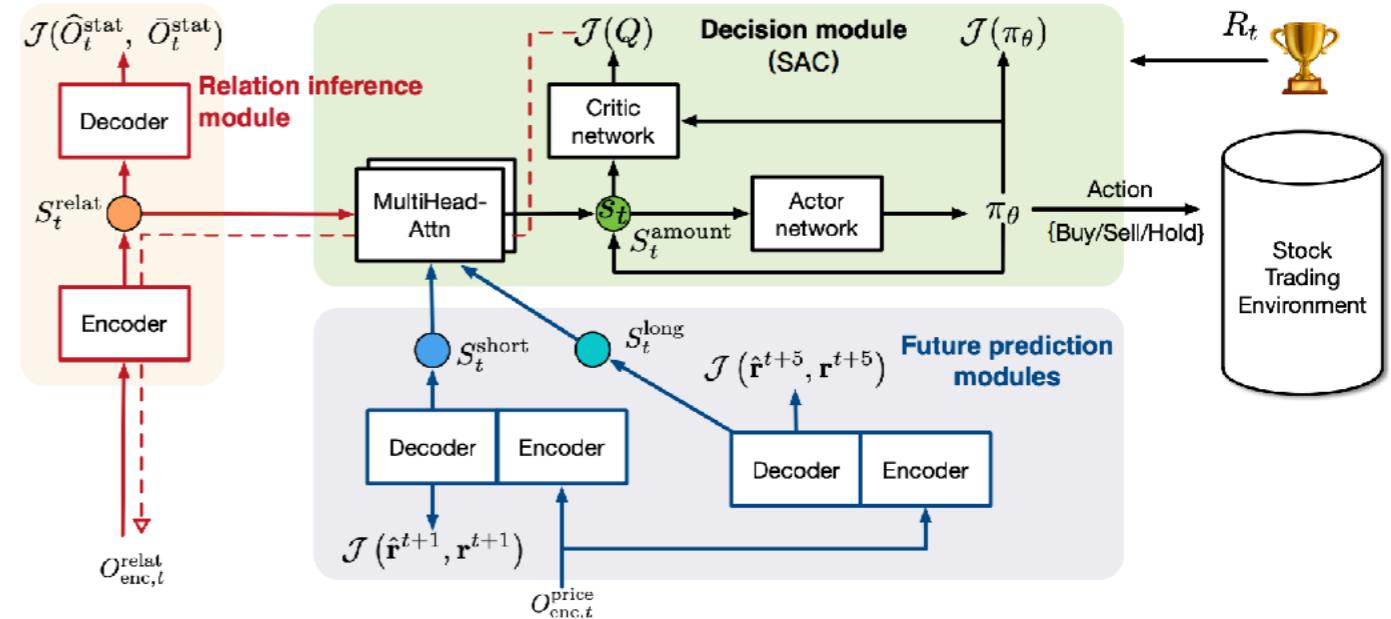


Figure 2: Unlike previous RL-for-finance methods, StockFormer builds the decision module upon learned representations provided by a relation inference module and two future prediction modules. The decision module contains a couple of multi-head attention layers that integrate the compositional representations, an actor network, and a critic network. In particular, the critic propagates its gradients of state values back into the relation inference module (Solid arrow: the data flow; Dashed arrow: the gradient flow of the critic loss).

Method	CSI-300				NASDAQ-100				Cryptocurrency	
	PR [↑]	AR [↑]	SR [↑]	MDD [↓]	PR [↑]	AR [↑]	SR [↑]	MDD [↓]	PR [↑]	SR [↑]
Market benchmark	0.24	0.08	0.51	0.18	1.05	0.30	1.16	0.30	-	-
Min-Var [Basak and Chabakauri, 2010]	0.11	0.04	0.38	0.13	0.59	0.18	0.99	0.26	-0.09	0.02
HATR [Wang et al., 2021]	-0.02	-0.01	0.14	0.45	0.08	0.28	0.25	0.41	-0.65	-0.66
Relational Ranking [Feng et al., 2019]	-0.01	-0.01	0.17	0.45	0.96	0.28	0.95	0.30	-	-
FactorVAE [Duan et al., 2022]	1.37	0.38	1.27	0.17	1.20	0.33	0.99	0.24	-	-
AutoFormer [Wu et al., 2021]	0.08	0.03	0.25	0.40	-0.25	-0.10	-0.21	0.42	-0.27	-0.16
Our future prediction module ($t + 5$)	0.89	0.27	0.73	0.49	0.66	0.20	0.64	0.44	-0.40	-0.71
SARL [Ye et al., 2020]	1.59	0.39	1.38	0.31	1.22	0.30	0.99	0.34	0.06	0.43
FinRL-SAC [Liu et al., 2021]	1.76	0.42	1.41	0.34	1.38	0.34	1.24	0.33	0.10	0.55
FinRL-DDPG [Liu et al., 2021]	1.43	0.36	1.23	0.39	0.83	0.22	0.87	0.33	0.15	0.60
StockFormer	2.47	0.54	1.73	<u>0.31</u>	1.71	0.40	1.39	<u>0.31</u>	0.24	0.75

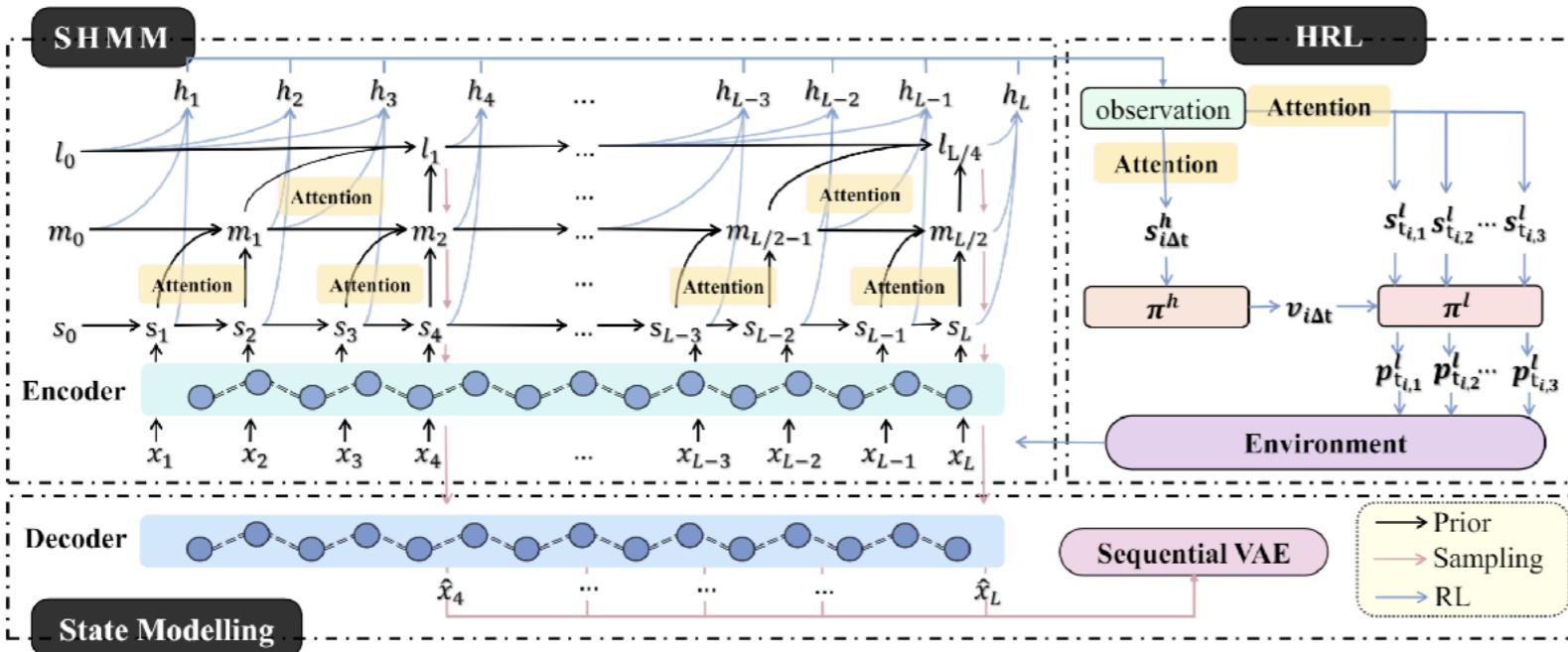
Table 3: Quantitative results in portfolio return (PR), annual return (AR), Sharpe ratio (SR), and maximum drawdown (MDD) on the test splits of the stock market datasets. Transaction costs are included in buying and selling actions. For stock prediction methods (Rows 3-7), we use the “buy-and-hold” strategy. In the last two columns, we show the results on the cryptocurrency dataset.

MacMic: Executing Iceberg Orders via Hierarchical Reinforcement Learning, IJCAI'24

Goal: Fuse market data with future trends and assets correlations in trading policies

Contributions:

- (1) Formulates order execution as hierarchical MDP, separating volume scheduling from order placement decisions
- (2) Introduces stacking Hidden Markov Model for extracting multi-granularity market state representations effectively
- (3) Develops hybrid learning approach combining hierarchical RL with imitation learning techniques



Methods	PA (bps)↑	PA-std (bps)↓	WR↑	GLR↑	AFI↓
TWAP	-0.12	3.12	0.49	0.97	0
VWAP	-3.29	4.87	0.39	0.78	0.04
AC	-1.24	4.21	0.45	1.02	0
DDQN	-1.21	7.09	0.47	0.99	0.05
PPO	-0.98	7.01	0.52	0.92	0.03
ODP	0.32	6.78	0.53	1.07	0.10
HALOP	2.89	6.12	0.65	1.12	0.07
MacMic	3.31	6.23	0.72	1.37	0.02

Table 2: The comparison results of the proposed method and the benchmarks on CSI100.

Methods	PA (bps)↑	PA-std (bps)↓	WR↑	GLR↑	AFI↓
TWAP	-0.19	4.31	0.49	1.01	0
VWAP	-2.89	4.72	0.35	0.79	0.02
AC	-2.31	4.39	0.39	0.96	0
DDQN	-1.98	6.89	0.45	0.97	0.03
PPO	-1.20	6.79	0.46	0.96	0.04
ODP	1.03	6.82	0.54	1.05	0.12
HALOP	2.75	5.87	0.63	1.13	0.05
MacMic	3.14	5.69	0.74	1.28	0.01

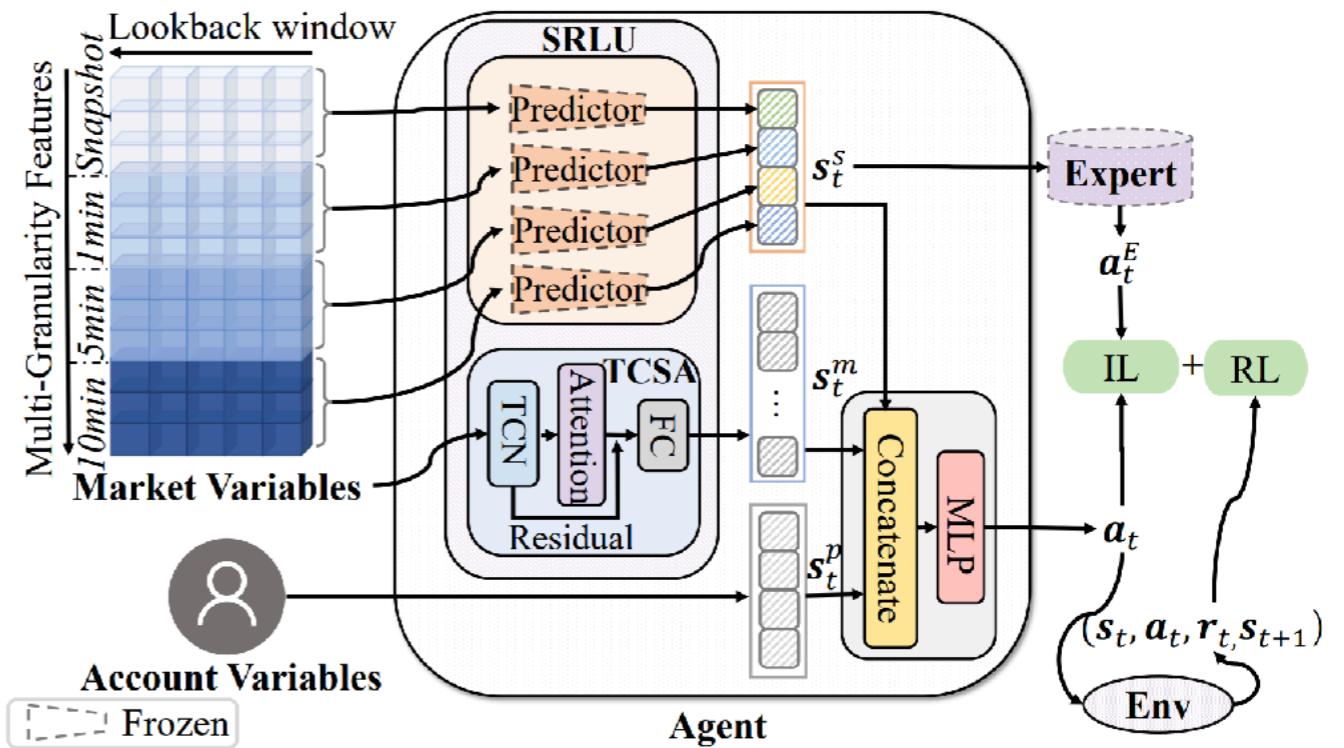
Table 3: The comparison results of the proposed method and the benchmarks on NASDAQ100.

IMM: An Imitative Reinforcement Learning Approach with Predictive Representation Learning for Automatic Market Making, IJCAI'24

Goal: Imitative reinforcement learning approach for automated market making with multi-price level order strategies

Contributions:

- (1) Formulates market making as MDP with effective state/action definitions for multi-price strategies.
- (2) Develops state representation learning unit incorporating multi-granularity signals and temporal-spatial attention networks.
- (3) Creates imitative reinforcement learning framework to extract knowledge from expert trading strategies.



	RB			FU			CU			AG		
	EPnL[10 ³]↑	MAP[unit]↓	PnLMAP↑	EPnL[10 ³]↑	MAP[unit]↓	PnLMAP↑	EPnL[10 ³]↑	MAP[unit]↓	PnLMAP↑	EPnL[10 ³]↑	MAP[unit]↓	PnLMAP↑
FOIC	3.23 ± 4.35	255 ± 111	14 ± 22	-7.79 ± 9.25	238 ± 135	-43 ± 56	-33.05 ± 27.63	206 ± 141	-161 ± 224	-48.39 ± 28.83	189 ± 154	-250 ± 335
LIIC	2.26 ± 3.32	123±32	20 ± 29	-6.89 ± 6.66	115± 30	-66 ± 69	-24.19 ± 14.83	150 ± 20	-164 ± 513	-38.9 ± 26.2	142 ± 45	-302 ± 243
LTIIC	9.16 ± 4.87	65 ± 6	139 ± 68	8.26 ± 2.64	52 ± 3	160 ± 50	-16.74 ± 15.81	112 ± 109	-190 ± 203	-32.57 ± 22.8	128 ± 22	-264 ± 166
<i>RL_{SD}</i>	4.36 ± 1.64	38 ± 4	114 ± 38	7.31 ± 5.38	76 ± 29	90 ± 46	-19.7 ± 17	214 ± 109	-92 ± 298	-25.43 ± 23.83	107 ± 37	-237 ± 235
<i>DRL_{OS}</i>	8.22 ± 3.70	51 ± 4	156 ± 61	11.03 ± 13.87	37 ± 3	30 ± 36	-18.9 ± 18.02	647 ± 2367	-99 ± 147	-28.39 ± 27.92	169 ± 154	-167 ± 135
IMM	16.46 ± 9.10	96± 13	165 ± 74	28.10 ± 10.27	102 ± 14	274 ± 89	-4.86 ± 10.17	111 ± 28	-43 ± 87	-14.5 ± 20.2	102 ± 14	-274 ± 89

EarnHFT: Efficient Hierarchical Reinforcement Learning for High Frequency Trading, AAAI'24

Goal: efficient high-frequency cryptocurrency trading that maintains performance across different market conditions

Contributions:

- (1) Introduces Q-teacher with optimal action values to enhance training efficiency of RL agents.
- (2) Constructs diverse agent pool through preference-based training for different market trend conditions.
- (3) Develops minute-level router to dynamically select appropriate trading strategies from agent pool.

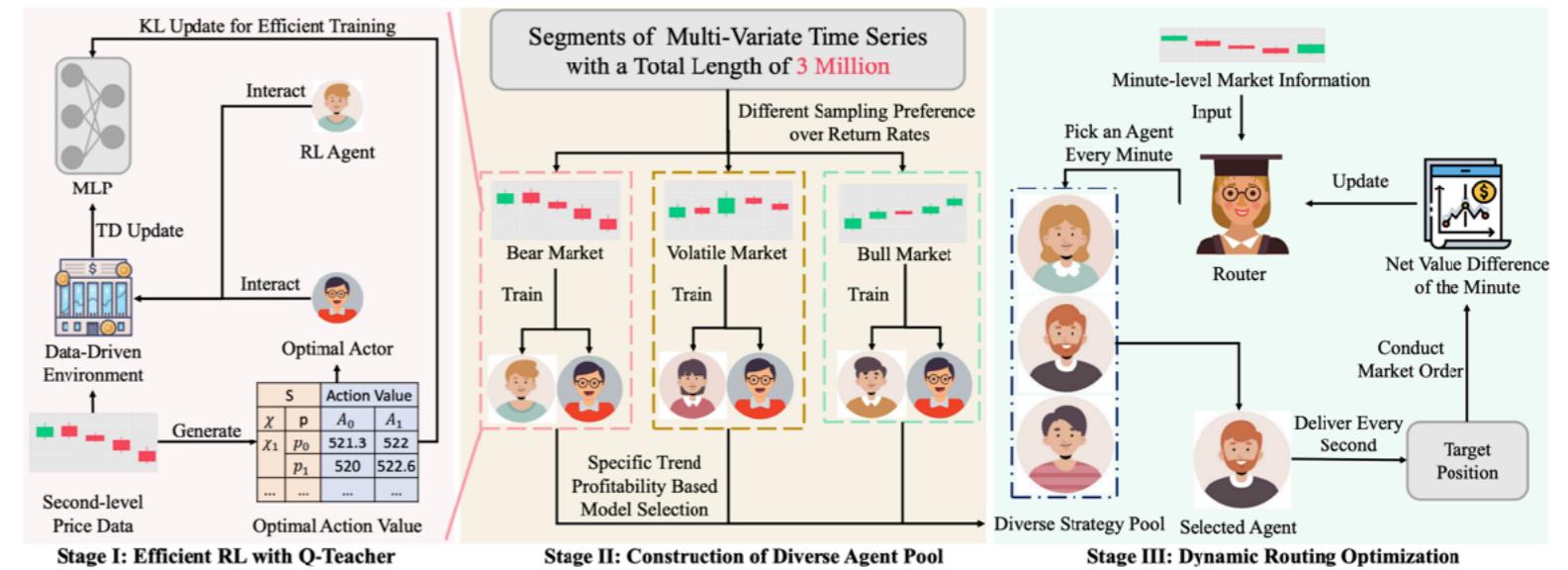


Figure 1: The overview of EarnHFT. First, we compute a Q-teacher for enhancing the performance and training efficiency of second-level RL agents. Then, we efficiently train diverse RL agents under various market trends and select a small fraction of them to form an agent pool based on profitability. Finally, we train a minute-level router which dynamically picks a second-level agent from the pool to achieve stable performance across different markets.

		Profit				Risk-Adjusted Profit		Risk Metrics				Profit				Risk-Adjusted Profit				Risk Metrics	
Market	Model	TR(%)↑	ASR↑	ACR↑	ASoR↑	AVOL(%↓)	MDD(%↓)	Market	Model	TR(%)↑	ASR↑	ACR↑	ASoR↑	AVOL(%↓)	MDD(%↓)						
BTCU	DRA	-4.56	-4.28	-19.57	-4.65	42.25	9.24	BTCT	DRA	-2.65	-4.82	-17.48	-4.77	21.18	5.84						
	PPO	-3.61	-5.25	-22.74	-5.71	27.76	6.41		PPO	-0.60	-14.74	-35.80	-0.14	1.59	0.65						
	CDQNRP	-2.83	-2.91	-14.85	-3.31	37.61	7.38		CDQNRP	-0.60	-19.52	-37.88	-0.74	1.20	0.61						
	DQN	-3.48	-12.37	-35.01	-11.86	11.57	4.09		DQN	0.47	4.21	27.94	1.14	4.38	0.66						
	MACD	-6.07	-10.11	-25.17	-8.05	24.84	9.98		MACD	-4.02	-5.80	-24.21	-4.32	26.87	6.44						
	IV	-2.99	-3.78	-14.24	-3.27	31.35	8.32		IV	-12.01	-17.83	-38.99	-13.90	27.68	12.66						
ETH	EarnHFT	0.72	1.22	10.77	0.93	27.08	3.07	GALA	EarnHFT	0.99	1.34	7.76	1.01	32.40	5.61						
	DRA	-33.37	-9.06	-32.23	-9.20	163.25	45.88		DRA	10.56	4.77	41.63	4.44	92.43	10.60						
	PPO	-22.61	-10.11	-31.17	-10.39	96.12	31.17		PPO	10.56	4.77	41.63	4.44	92.43	10.60						
	CDQNRP	-6.82	-24.41	-40.19	-3.11	11.46	6.96		CDQNRP	5.22	4.51	39.42	4.16	47.27	5.41						
	DQN	-11.02	-9.47	-32.81	-8.43	47.76	13.79		DQN	2.94	3.55	32.02	2.66	34.08	3.78						
	MACD	-4.29	-1.78	-8.71	-1.19	79.64	16.35		MACD	2.37	1.79	11.45	1.22	62.89	9.84						
	IV	-27.42	-12.27	-36.01	-9.00	99.67	33.96		IV	13.95	6.74	55.67	5.41	81.79	9.91						
	EarnHFT	4.52	2.92	14.30	1.78	67.92	13.89		EarnHFT	19.41	9.77	79.08	7.79	74.94	9.26						

Table 2: Performance comparison on 4 Crypto markets with 6 baselines including 2 policy-based RL algorithms, 2 value-based RL algorithms, and 2 rule-based methods. Results in pink, green, and blue show the best, second-best, and third-best results.

A Multimodal Foundation Agent for Financial Trading: Tool-Augmented, Diversified, and Generalist, KDD'24

Goal: leverages Multimodal LLMs for comprehensive, tool-augmented financial trading decisions

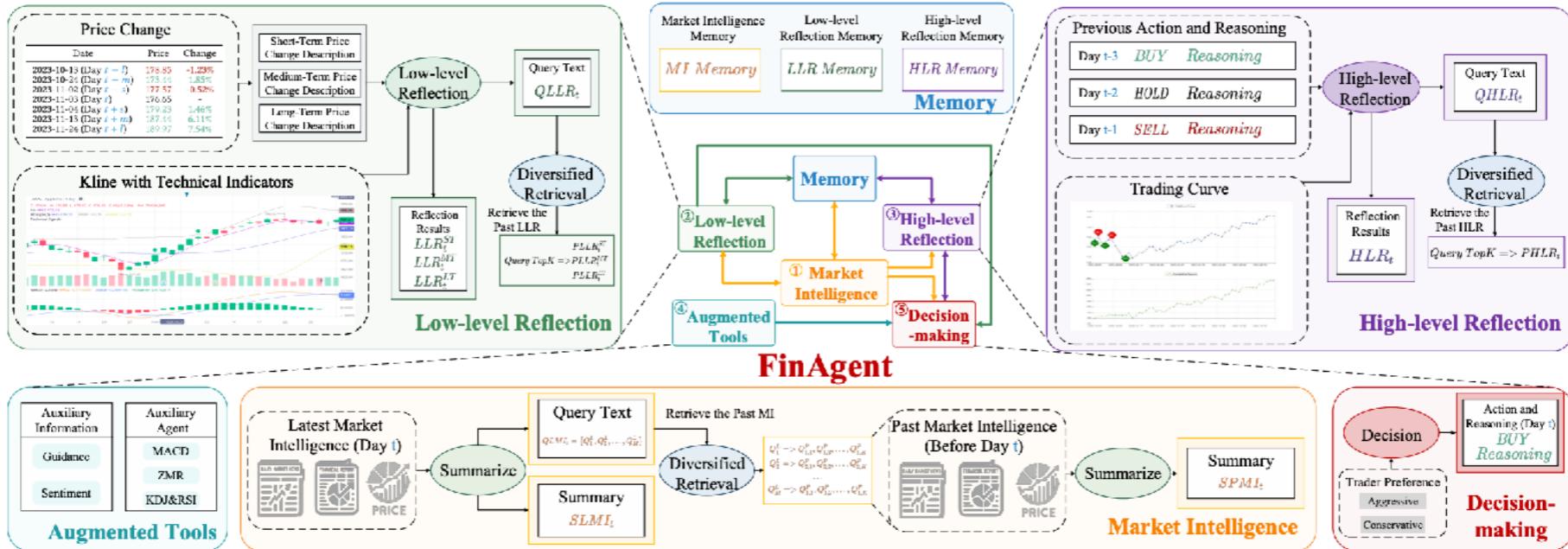


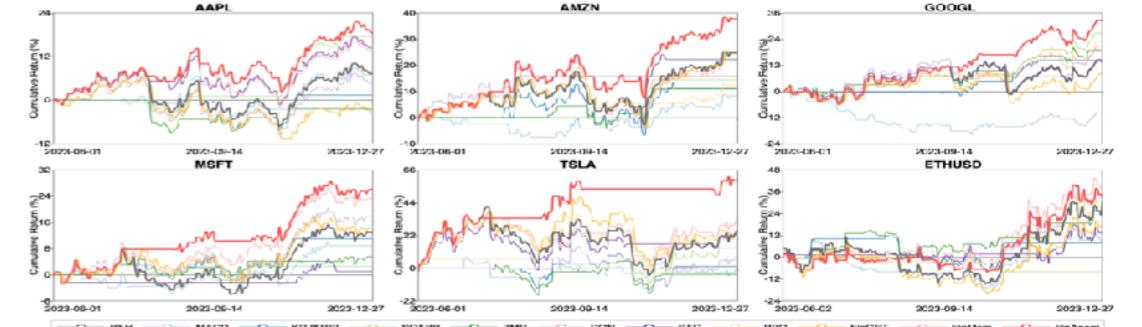
Figure 2: The overall architecture of FinAgent. The ordinal numbers in the figure represent the order of execution, where augmented tools are implemented with the decision-making module.

Contributions:

- (1) Developed a market intelligence module extracting insights from multimodal data including prices, visuals, and news
- (2) Created a dual-level reflection system analyzing market movements and evaluating past trading decisions
- (3) Integrated expert trading strategies and technical indicators as auxiliary tools to enhance decision-making

Table 4: Performance comparison of all methods on six profitable metrics. Results in red, yellow and green show the best, second best and third best results on each dataset. The improvement row is the FinAgent over the best-performing baselines.

Categories	Models	AAPL			AMZN			GOOGL			MSFT			TSLA			ETHUSD		
		ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓												
Market	B&H	13.0	0.6	14.78	42.35	1.68	17.38	22.47	0.71	12.97	22.49	0.64	12.52	37.4	0.72	32.65	29.26	0.87	23.21
Rule-based	MACD	11.84	0.72	10.38	14.27	0.71	7.84	-18.0	-0.89	20.07	15.23	0.77	8.34	-4.9	-0.02	14.15	10.24	0.47	24.32
	ZMR	2.17	0.17	11.88	19.36	0.65	17.27	24.39	2.13	2.03	18.64	1.06	7.78	2.14	0.17	24.73	8.87	0.51	16.95
	ZMR	3.91	0.22	8.88	18.73	0.84	7.89	32.51	1.65	5.38	9.85	0.71	6.22	-7.28	0.09	19.5	29.35	1.23	13.11
ML&DL-based	LGBM	16.93	1.47	2.52	29.34	0.72	17.41	24.77	0.7	12.98	19.28	0.67	12.56	15.57	0.84	3.88	24.91	0.72	22.95
	LSTM	16.97	0.54	11.93	15.91	0.46	17.41	24.86	0.7	12.98	18.66	0.68	11.75	17.35	0.73	4.44	36.09	1.03	21.5
	Transformer	17.11	0.96	7.53	52.66	1.11	4.96	13.69	0.6	12.93	17.44	1.16	2.59	39.7	1.01	8.17	31.0	1.02	12.95
RL-based	DQN	7.92	0.4	14.88	27.45	1.17	5.27	34.4	1.59	7.15	50.44	1.18	10.56	15.07	0.44	28.12	29.81	1.18	9.53
	SAC	24.84	1.12	11.98	38.33	1.07	13.84	23.8	0.75	13.07	22.02	0.82	12.52	42.23	0.87	26.19	8.76	0.76	10.05
	PPO	13.28	0.51	14.78	21.17	0.7	13.84	38.29	1.3	8.45	11.32	0.68	17.51	33.61	0.75	28.35	34.75	1.31	11.12
LLM-based	FinGPT	-5.46	-0.17	16.23	49.95	1.1	18.94	12.28	0.64	13.0	25.1	0.97	9.84	33.43	0.75	31.47	21.57	0.68	25.55
	FinMem	23.76	1.11	10.39	40.07	1.03	18.53	31.27	1.11	8.97	40.58	1.5	50.01	25.77	1.27	14.72	12.57	13.59	
Ours	FinAgent	31.9	1.43	10.4	65.1	1.61	13.2	56.15	1.78	8.45	44.74	1.79	5.57	92.27	2.01	12.14	43.08	1.18	12.72
	Improvement(%)	28.39	-	-	51.64	37.61	-	46.64	-	-	10.25	19.33	-	84.39	93.27	-	-	-	-

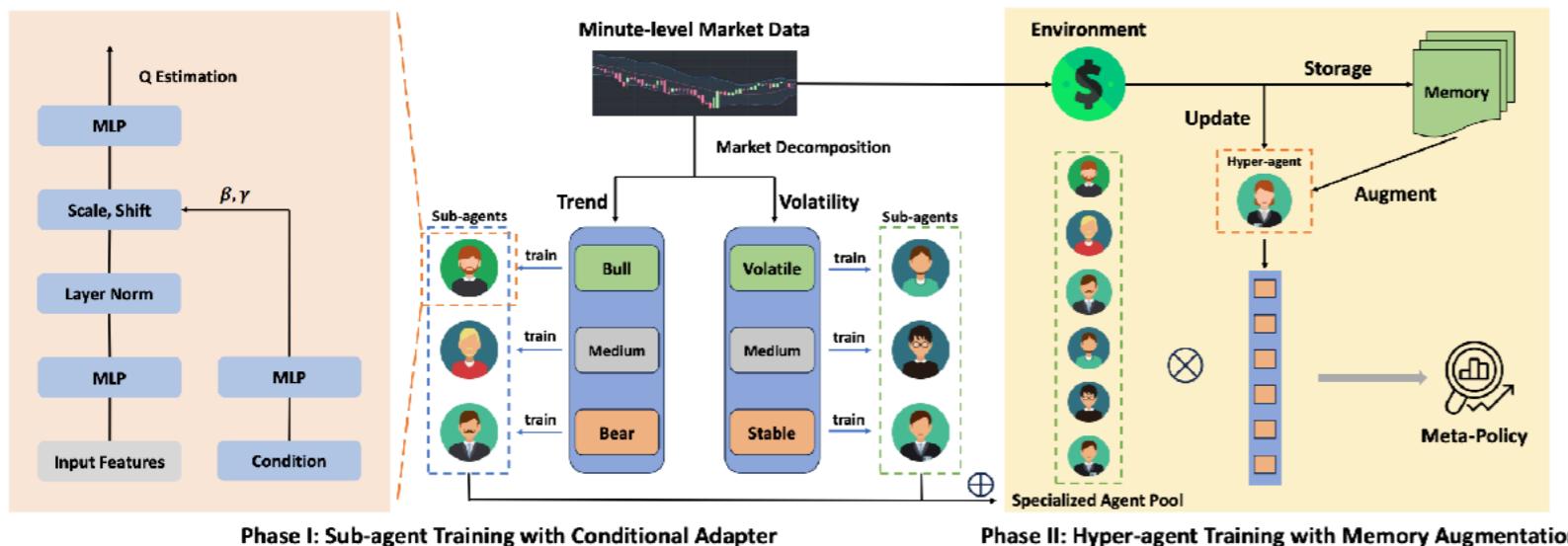


MacroHFT: Memory Augmented Context-aware Reinforcement Learning On High Frequency Trading, KDD'24

Goal: Multiple condition-adaptive agents with a memory-enhanced hyper-agent for profitable meta-policy for cryptocurrency market

Contributions:

- (1) Introduces market decomposition using trend and volatility indicators to train specialized trading agents.
- (2) Develops conditional adaptation for sub-agents to adjust trading policies based on market context.
- (3) Creates a memory-enhanced hyper-agent that effectively combines diverse trading strategies from sub-agents.



Model	BTCUSDT			ETHUSDT		
	TR(%)↑	ASR↑	MDD(%)↓	TR(%)↑	ASR↑	MDD(%)↓
w/o-CA	1.69	0.36	7.24	14.27	2.42	<u>7.57</u>
w/o-MEM	2.03	0.49	6.49	12.73	1.26	20.87
MacroHFT	<u>3.03</u>	<u>0.61</u>	<u>5.41</u>	<u>39.28</u>	<u>3.89</u>	9.67

Model	DOTUSDT			LTCUSDT		
	TR(%)↑	ASR↑	MDD(%)↓	TR(%)↑	ASR↑	MDD(%)↓
w/o-CA	-16.79	-1.18	31.66	-6.71	-0.36	18.83
w/o-MEM	2.41	0.34	27.23	-8.66	-0.58	22.03
MacroHFT	<u>13.79</u>	<u>0.97</u>	<u>15.89</u>	<u>18.16</u>	<u>1.50</u>	14.24

Table 3: Performance comparison of models across four datasets. Underlined results represent the best performance

Efficient Continuous Space Policy Optimization for High-frequency Trading, KDD'23

Goal: DRL for continuous-space portfolio optimization in high-frequency trading

Contributions:

- (1) **First** to address high-frequency portfolio optimization using continuous-space policy optimization with DRL framework
- (2) Proposed efficient learning by separating static market states from dynamic portfolio states and using probabilistic programming
- (3) Demonstrated superior performance through extensive experiments on three real-world datasets and successful deployment

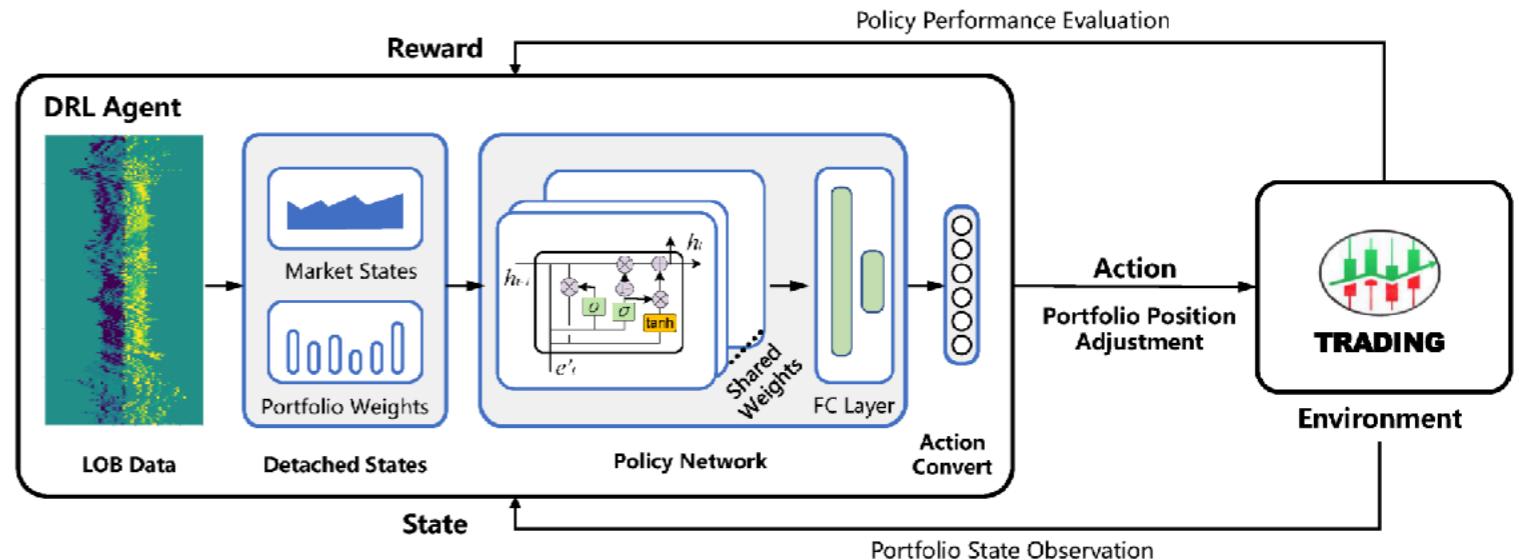


Figure 2: The model framework of our proposed efficient continuous space policy optimization for high-frequency trading. (a) Left: The illustration of the DRL agent, which includes the state detach processing, the policy network with shared-weights recurrent neural network and the action convertor that produce the adjusted portfolio position; (b) Right: displays the environment of our proposed framework, which generates rewards for different DRL actions and simulates the financial status on the stock market.

Table 2: The results of models' Alpha performance metrics on DOW30, COIN and SSE50 datasets. CAR denotes the cumulative adjusted (excess) return, MDD is the maximum drawdown, SR means Sharpe ratio and Calmar is the Calmar ratio.

Dataset	DOW30				COIN				SSE50			
	CAR↑	MDD↓	SR↑	Calmar↑	CAR↑	MDD↓	SR↑	Calmar↑	CAR↑	MDD↓	SR↑	Calmar↑
LightGBM	11.63	0.1761	0.1305	0.3204	0.01	0.0480	0.0039	0.0117	2.01	0.0648	0.0231	0.0309
ALSTM	11.02	0.1889	0.1237	0.2836	0.06	0.0487	0.0273	0.0748	3.81	0.0649	0.5358	0.6086
DDPG	25.75	0.0989	1.0669	1.2242	0.06	0.0408	0.0331	0.0852	6.45	0.1031	0.2625	0.6495
A3C	58.56	0.1351	1.7203	1.9143	0.09	0.0408	0.0357	0.0924	6.56	0.1294	0.2204	0.5259
PPO	31.31	0.0667	1.3203	2.1800	0.98	0.0738	0.3290	0.7142	8.85	0.1278	0.2962	0.7191
SAC	74.91	0.1817	1.5289	1.7710	0.18	0.0291	0.1067	0.3358	10.16	0.1167	0.3469	0.9043
ADARNN	15.62	0.1852	0.1754	0.4057	0.06	0.0447	0.0262	0.0763	3.38	0.0661	0.4469	0.5316
TRA	20.99	0.1757	0.2333	0.5679	0.17	0.0483	0.0603	0.1824	4.41	0.0572	0.5632	0.7999
DRPO	88.74	0.1653	1.7252	2.2560	1.02	0.0072	2.5007	7.5420	20.02	0.1045	1.1417	1.9935

Optimizing Trading Strategies in Quantitative Markets using Multi-Agent Reinforcement Learning, ICASSP'24

Goal: Optimize trading strategies by combining portfolio insurance methods with multi-agent reinforcement learning.

Contributions:

- (1) Integration of CPPI strategy with MADDPG framework for enhanced portfolio management
- (2) Development of TIPP-MADDPG approach for dynamic portfolio protection and risk management
- (3) Empirical validation showing improved performance over traditional quantitative trading methods

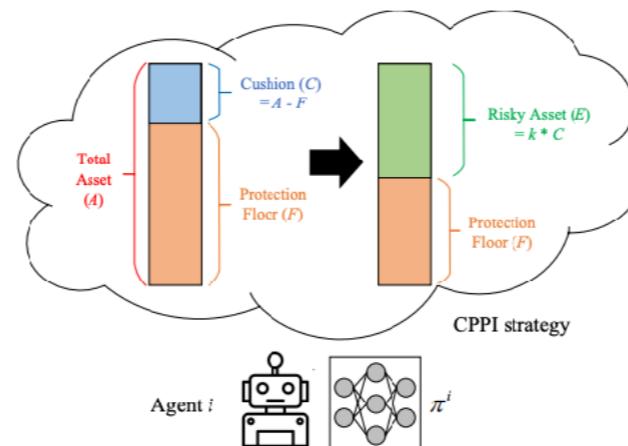


Fig. 1. The principle of CPPI strategy for agent i .

Algorithm 1 MADDPG with insurance strategy

```

1: Initialize the rule-based policy  $\Phi$  according to the insurance strategy.
2: Initialize  $Q_{\phi_i}, \pi_{\theta_i}, Q_{\phi'_i}, \pi_{\theta'_i}$  for all  $i \in \{1, \dots, N\}$ .
3: while training not finished do
4:   Initialize initial state  $s$  and a random process  $\mathcal{N}$  for action exploration.
5:   for each episode do
6:     For each agent select action  $a_i = \pi_{\theta_i}(s) + \mathcal{N}_i$ .
7:     if  $a_i \notin \Phi(s)$  then
8:       adjust  $a_i$  with  $\Phi(s)$ .
9:     end if
10:    Execute joint action  $a$  and observe reward  $r$  and next state  $s'$ .
11:    Store experience  $\langle s, a, r, s' \rangle$  in replay buffer  $\mathcal{D}$ .
12:    Sample a minibatch of  $K$  experiences from  $\mathcal{D}$ .
13:    for each agent do
14:      Update the critic  $Q_{\phi_i}$  by minimizing Eq.(4).
15:      Update the actor  $\pi_{\theta_i}$  using Eq.(3).
16:    end for
17:    Update target networks for each agent:

$$\phi'_i \leftarrow \tau \phi_i + (1 - \tau) \phi'_i,$$


$$\theta'_i \leftarrow \tau \theta_i + (1 - \tau) \theta'_i.$$

18:  end for
19: end while

```

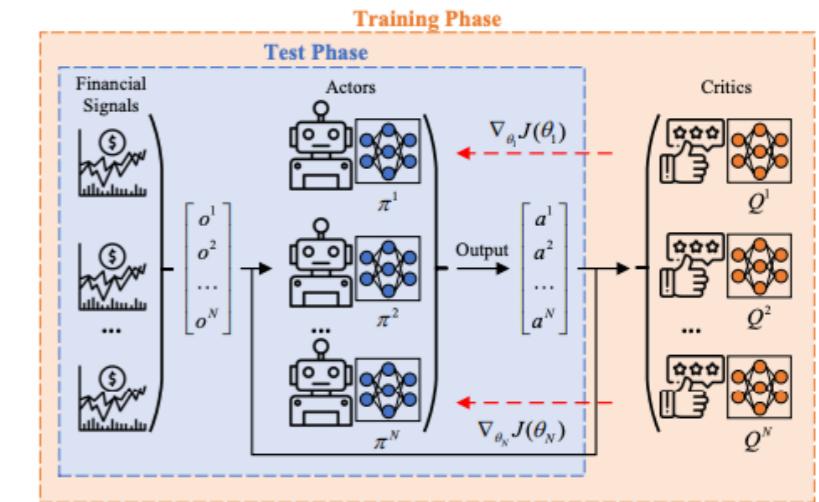


Fig. 2. A schematic of MADDPG in the quantitative market environment.

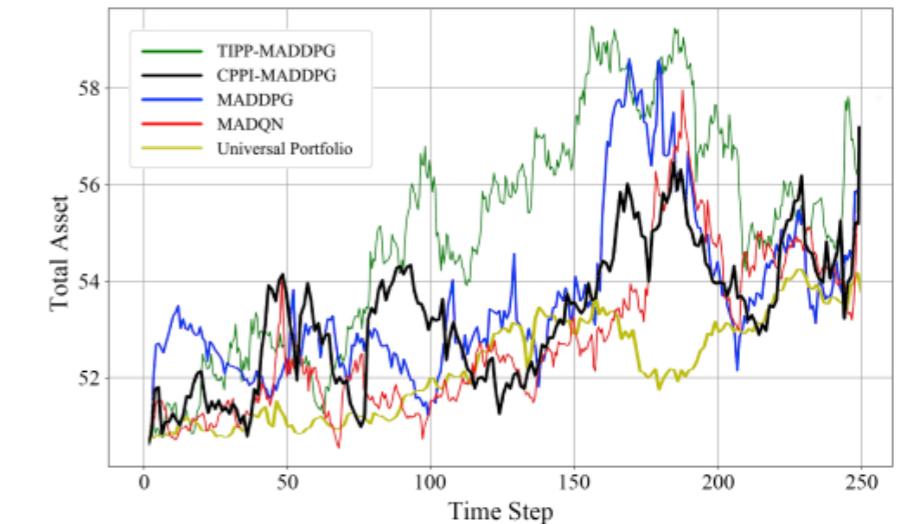


Fig. 4. The performances of portfolios with different strategies (The metrics of Total Asset and Time Step: 10^3 RMB and Day).

Hierarchical Reinforced Trader(HRT): A Bi-Level Approach for Optimizing Stock Selection and Execution, ArXiv'24

Goal: Introduces a hierarchical reinforcement learning approach to optimize stock selection and trade execution.

Contributions:

- (1) Proposes HRT agent with phased alternating training for coordinating high and low-level trading decisions
- (2) Demonstrates superior Sharpe ratio on S&P 500 stocks across bullish/bearish market conditions
- (3) Shows how hierarchical framework reduces dimensionality issues and improves trading diversification performance

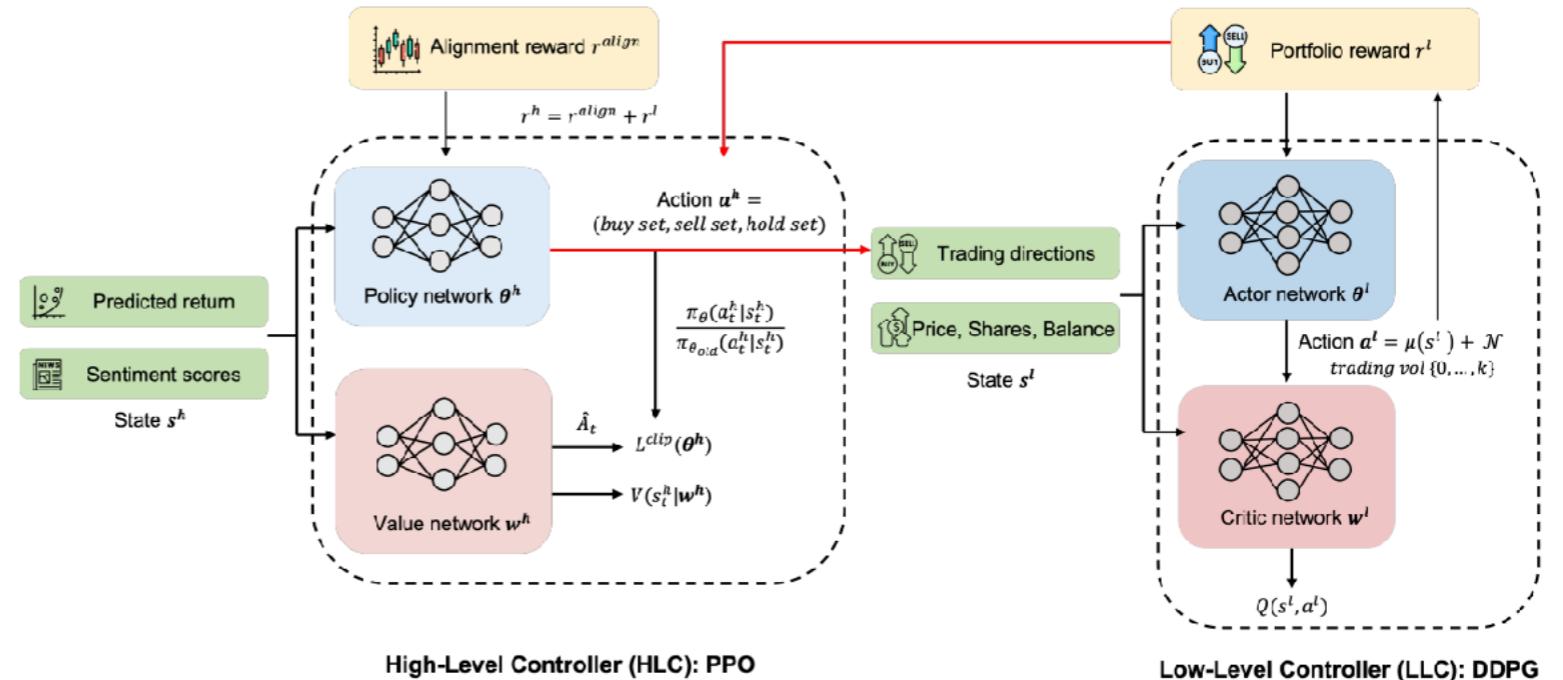
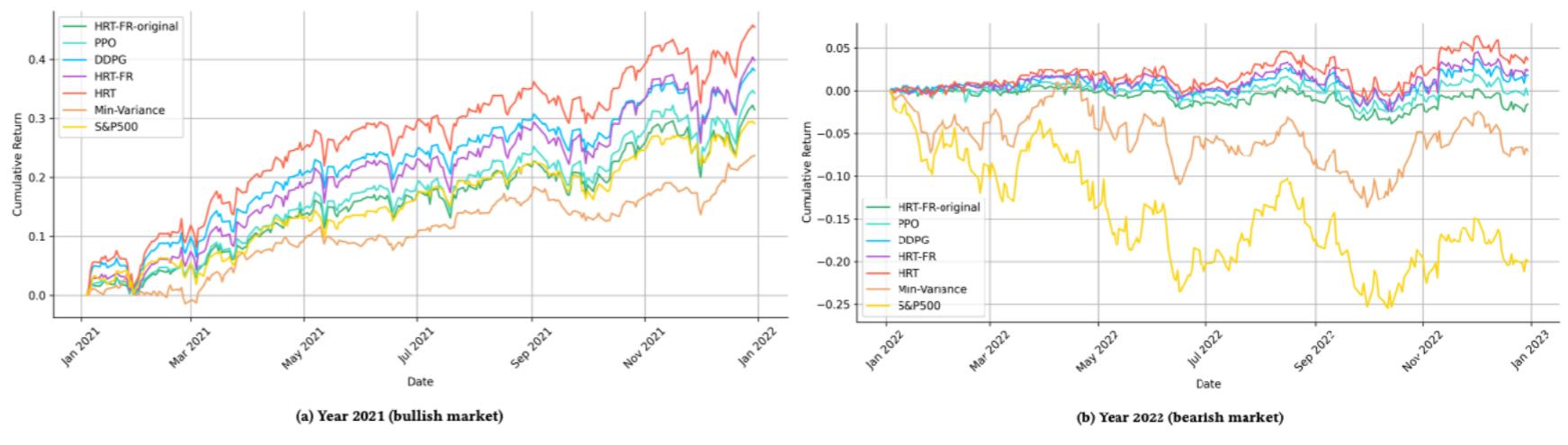


Figure 2: Overview of the Hierarchical Reinforced Trader (HRT) architecture. Interactions between the HLC and LLC are indicated by the red arrows.



Knowledge Retrieval and Augmentation

Extracting Financial Events from Raw Texts via Matrix Chunking, COLING'24

Goal: Propose a method to extract financial events from raw Chinese texts using matrix chunking.

Contributions:

- (1) Introduce CFTE task for extracting Chinese financial events directly from unstructured text
- (2) Create FINEED dataset with rich linguistic features and two-dimensional annotation method
- (3) Develop MACK method using matrix chunking for fault-tolerant financial event extraction



Figure 1: Example of an *Equity Pledge* event triggered by *pledged* with three corresponding event arguments.

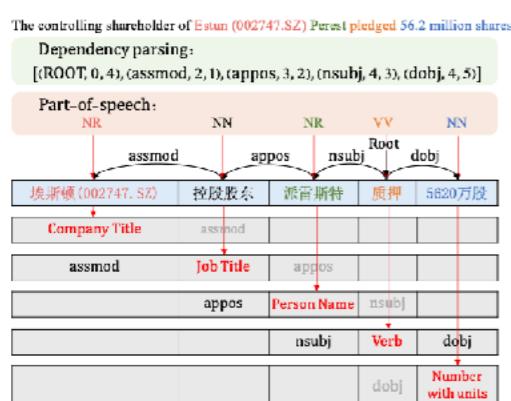


Figure 2: An example of the 2-D matrix. The red tags correspond to each attentive position. The black tags are directly assigned according to dependency parsing tags. The grey tags are determined based on symmetry. The empty places are *O* tags, indicating an absence of any tag.

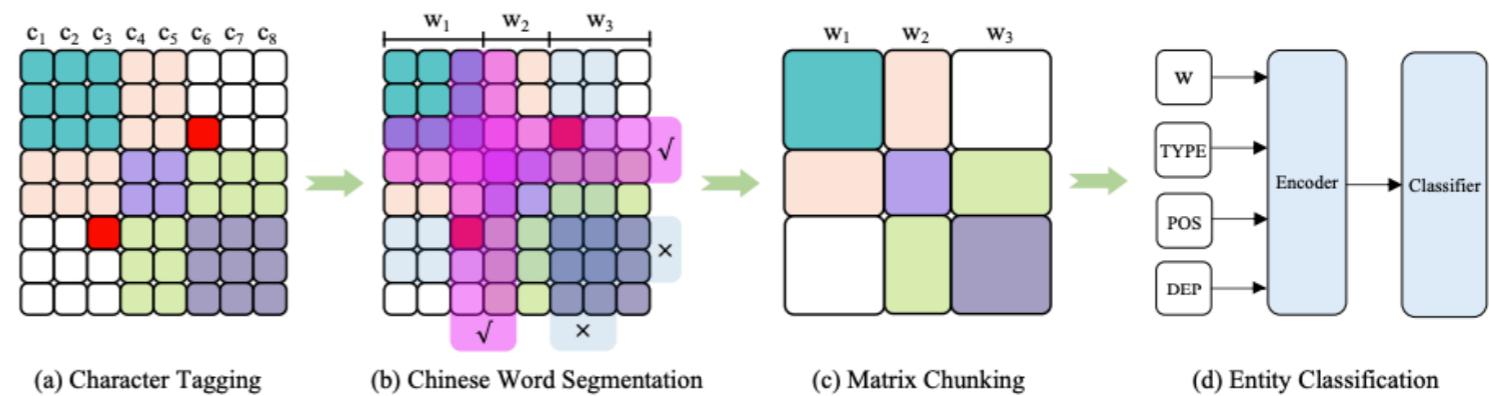


Figure 3: An illustration of the MACK method, including four key tiers. (a) Implement character-level tagging using distinct colors for different tags. Tags in red denotes incorrect predictions. (b) Segment the text based on the statistics of adjacent rows and columns at the corresponding position. (c) Partition the matrix into submatrices using word coordinates, and assign the mode tag of each submatrix as the designated tag. (d) Detect event triggers and classify entities into the event roles by leveraging multiple features.

Dataset	Train	Dev	Test	Total
ACE 2005	529	30	40	599
FINEED	3000	1000	1000	5000

Table 1: Comparison between the FINEED dataset and ACE 2005 event extraction dataset.

Dataset	Entity Type	Total
FEED (Li et al., 2021)	person, organization, date, number	4
FINEED	company title, branch office, national institution, organization name, department, person name, job position, job title, location name, number with units, time point, time interval	12

Table 2: Comparison of entity types between different event extraction datasets.

FinReport: Explainable Stock Earnings Forecasting via News Factor Analyzing Model, WWW'24

Goal: automatically generates comprehensive stock analysis reports for ordinary investors using news and market data

Contributions:

- (1) **First** work to create an automated financial report system that collects, analyzes, and summarizes market information
- (2) Proposes three modules (news factorization, return forecasting, risk assessment) to generate reliable and explainable reports
- (3) Demonstrates superior performance through extensive real-world experiments comparing against baseline methods and benchmarks

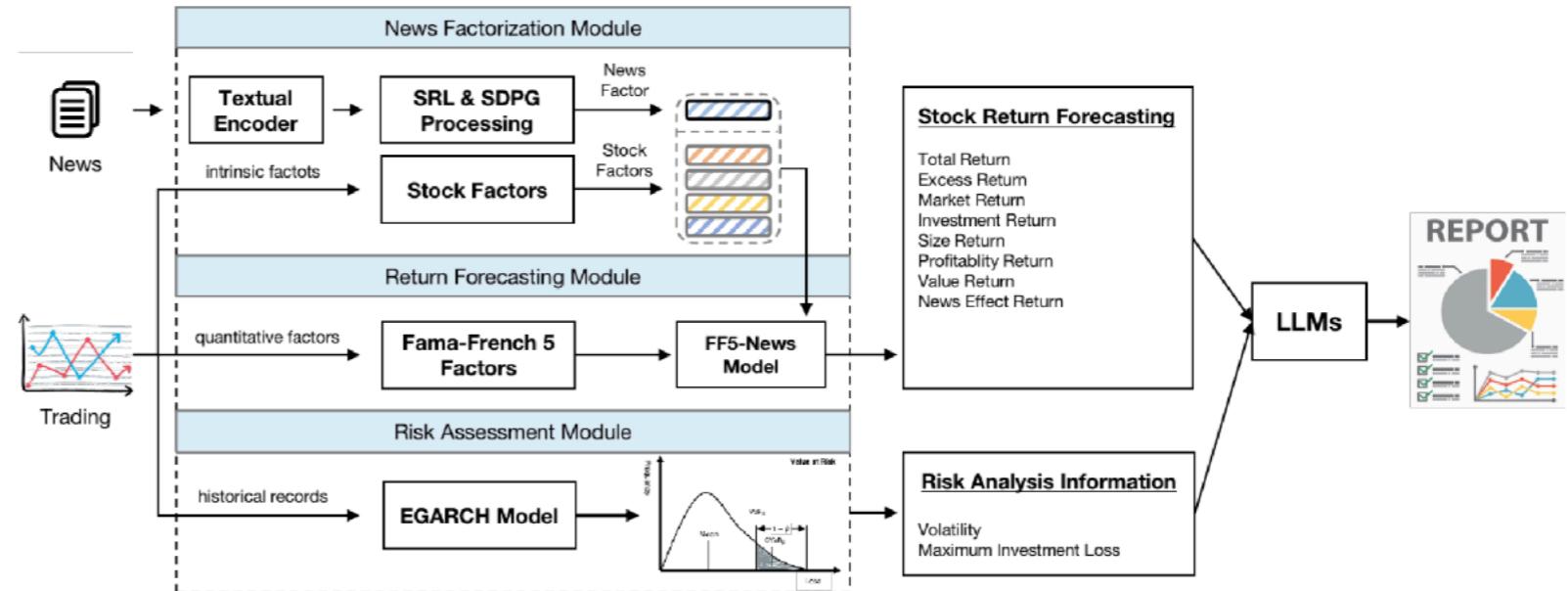


Figure 2: The overall architecture of our FinReport for explainable stock returns forecasting.

Model	Reference	Resource	Accuray	F1 Score	Recall	Precision
StockNet [28]	ACL 2018	News	46.72	44.44	46.68	47.65
HAN Stock [13]	WSDM 2018	News	57.35	56.61	57.20	58.41
Bert Chinese [9]	ACL 2019	News	59.11	58.99	59.20	59.07
ERNIE-SKEP [25]	ACL 2020	News	60.66	60.66	60.59	61.85
XLNET Chinese [6]	EMNLP 2020	News	61.14	61.19	61.09	61.60
RoBERTa WWM Ext [6]	EMNLP 2020	News+Factors	61.34	61.48	61.32	61.97
Chinese Lert Large [7]	Arxiv 2022	News	64.37	64.30	64.31	64.34
Chinese Pert Large [8]	Arxiv 2022	News+Factors	66.36	66.16	66.69	66.4
Self-supervised SRLP [34]	FinNLP 2022	News	65.09	65.03	65.07	65.02
		News+Factors	67.37	67.27	67.73	67.28
		Factors	59.76	59.71	59.75	59.72
Factors only		Factors	63.74	63.66	63.71	63.67
SRL&SDPG		News	66.10	66.01	66.09	66.04
SRL & Factors	Ours	News+Factors	69.48	69.28	69.41	69.54
SDPG & Factors		News+Factors	73.12	72.97	72.96	73.04
SRL & SDPG & Factors		News+Factors	75.40	75.12	75.23	75.42

Analyze stock 002069.SZ. Today, the following news has emerged:
On April 14, 2021, Zoneco Group reported an estimated Q1 net loss of 30.50 million yuan, compared to a 3.7139 million yuan profit last year. The company aims to enhance core business focus, streamline operations, and manage risks.

FinReport
Based on the analysis from our model, the following impacts on the stock's returns are expected from this news.

Return Forecast
Market Factor: Announcement may lead to 3% stock decline, impacting factor returns.
Size Factor: Reduced asset size suggests limited gains, causing a 1% decrease.
Valuation (BP) Factor: Net loss may lower investor valuation, causing about 1.5% decrease.
Profitability Factor: Net loss might affect future earnings perception, causing a projected 3% decrease.
Investment Factor: Anticipated loss may foster cautious investment, raising returns by 0.6%.
News Effect Factor: Expected loss may cause negativity, but ongoing efforts could maintain attention, resulting in 2% decrease.

Risk Assessment: Historical data reveals heightened stock volatility in 30 days, indicating potential for abnormal fluctuations. Expected max decline: 10%, indicating substantial risk.

Overall trend prognosis: Negative

Summary: After considering the revenue forecast and risk assessment information, we believe that the stock carries an extremely high level of risk, and the anticipated rate of return decline is expected to be above 8%.

This report serves as a reference; it is not investment advice. Kindly exercise caution and acknowledge personal risks.

Large Language Models as Financial Data Annotators: A Study on Effectiveness and Efficiency, COLING'24

Goal: Evaluate large language models as financial data annotators compared to experts and crowd-workers.

Contributions:

- (1) First study comparing LLMs against domain experts/crowdworkers for financial data annotation tasks
- (2) Comprehensive analysis of 3 LLMs with different parameters and prompting approaches
- (3) Introduces reliability index to identify trustworthy samples needing human review

Text: The predecessor **Mississippi Power Company** was incorporated under the laws of the State of Maine on November 24, 1924 and was admitted to do business in Mississippi on **December 23, 1924** and in Alabama on December 7, 1962.

Relation type: Organization–Date

Expert Label: NO/OTHER RELATION

Crowdworker Label: FORMED ON

Figure 1: Example of relation extraction task from REFinD dataset.

Select date of formation relationship described in one sentence. Given a single sentence: The predecessor **Mississippi Power Company** was incorporated under the laws of the State of Maine on November 24, 1924 and was admitted to do business in Mississippi on **December 23, 1924** and in Alabama on December 7, 1962. With 2 highlighted phrases: Mississippi Power Company and December 23, 1924. Select a multiple choice answer from options below, which best describes the relation between Mississippi Power Company and December 23, 1924.

Please choose the MOST appropriate relation from the following options:

1. Mississippi Power Company is/was formed on December 23, 1924
2. Mississippi Power Company is/was acquired on December 23, 1924
3. no/other relation between Mississippi Power Company and December 23, 1924

Annotator	Type	Temperature Setting	Micro-Averaged F1 Score/ Accuracy(%)					
			Zero-Shot Prompt		Few-Shot Prompt		Few-Shot CoT Prompt	
			simple prompt	full instruction	1-shot	5-shot	1-shot CoT	5-shot CoT
LLM	GPT-4	0.2	67.4/63.4	68.5/64.6	65.0/60.1	67.6/63.8	64.5/58.4	68.4/ 65.4
	GPT-4	0.7	67.6/63.6	68.4/64.6	65.0/60.0	67.7/63.9	64.6/58.4	68.4/ 65.4
	PaLM 2	0.2	62.3/53.9	62.2/53.8	66.4/60.1	66.0/59.2	64.7/55.9	65.6/57.2
	PaLM 2	0.7	64.5/56.0	64.4/56.0	67.3/60.9	68.7/63.8	64.9/57.4	65.9/59.2
	MPT Instruct	0.2	20.0/21.9	31.1/27.6	18.6/18.0	42.5/36.7	20.1/18.5	45.2/36.1
	MPT Instruct	0.7	20.8/24.7	24.8/27.3	22.7/24.2	30.5/31.1	22.2/23.2	33.9/30.8
	Ensemble (All LLMs)	0.2	65.2/60.1	66.0/60.7	63.9/58.1	68.1/63.3	63.3/56.4	68.8/63.8
	Ensemble (GPT-4 w Palm 2)	0.2	67.2/63.2	68.6/64.7	65.0/60.1	67.8/64.0	64.3/58.1	68.2/65.2
	Ensemble (GPT-4 w MPT Instruct)	0.2	67.2/63.2	68.6/64.7	65.0/60.1	67.8/64.0	64.3/58.1	68.2/65.2
	Ensemble (Palm 2 w MPT Instruct)	0.2	62.6/54.3	61.9/53.6	66.7/60.5	66.1/59.4	64.5/55.7	65.4/56.9
Human	Mturk Annotators	-	-	38.6/40.7	-	-	-	-

Table 1: Annotator performance in terms of micro-averaged F1-Score and accuracy against expert assigned labels.

Figure 2: Full instruction prompt example.

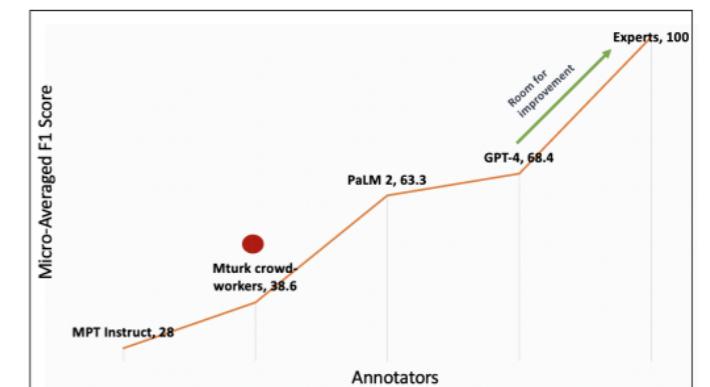


Figure 3: Annotator performance in terms of micro-averaged F1-Score under full instruction prompt.

Prompting for Numerical Sequences: A Case Study on Market Comment Generation, COLING'24

Goal: Investigates optimal input representations for language models to generate market commentary from numerical data.

Contributions:

- (1) Presents performance analysis of zero-shot and few-shot approaches for market commentary generation
- (2) Compares different prompting strategies for converting numerical sequences into effective prompts
- (3) Provides insights into which prompt designs work best for numerical data-to-text tasks

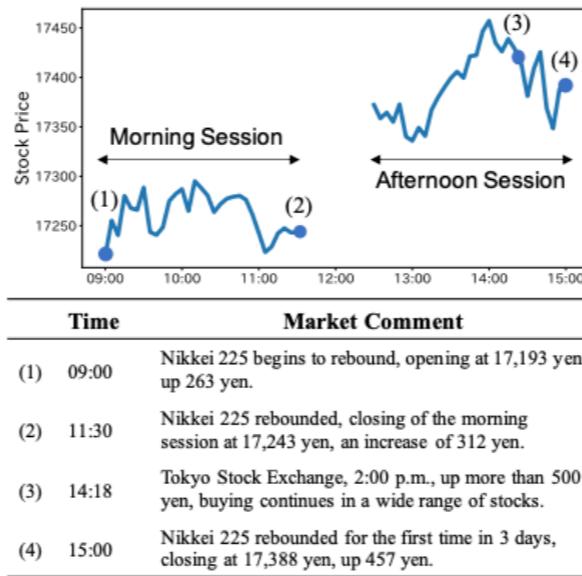


Figure 1: Example of market comment generation. We represent the time-series numerical data using a line graph. The objective of this task is to generate comments on the stock price movement at a specific time.

Output the market comment at the current time in the form of a <comment>market comment</comment>.

###

Input:

[INPUT FORMAT (short-term)]
[INPUT FORMAT (long-term)]

Output:

Nikkei225 closes at large, rebounding yen strength pushes mainstay stocks higher

###

Input:

[INPUT FORMAT (short-term)]
[INPUT FORMAT (long-term)]

Output:

Table 3: The template we use for the few-shot setting.

	0-shot			5-shot			10-shot		
	BLEU	METEOR	BERTScore	BLEU	METEOR	BERTScore	BLEU	METEOR	BERTScore
Direct	0.01	0.48	60.30	8.26	25.22	73.50	9.39	26.55	73.96
Column	0.38	14.06	65.33	8.30	24.99	73.35	9.49	26.00	73.65
Row	0.42	8.86	64.83	9.16	26.33	73.76	10.49	27.88	74.31
Python List	0.36	16.16	65.01	8.32	25.32	73.54	9.59	26.51	73.87
Python List (nested)	0.40	8.94	65.77	9.15	26.77	74.01	9.86	27.42	74.15
Python Dictionary	0.44	9.60	65.40	9.17	26.42	73.96	10.41	28.25	74.56
HTML Table	0.35	12.26	63.93	8.30	26.10	73.92	8.45	26.44	74.08
LaTeX Table	0.44	15.11	67.59	8.36	26.10	73.76	9.53	27.56	74.02
Text (English)	0.30	15.66	60.13	8.49	25.92	73.97	9.10	26.95	74.52
Text (Japanese)	0.03	0.95	55.35	8.51	26.55	74.00	9.26	27.60	74.21

Table 5: Comparison of methods in terms of BLEU, METEOR, and BERTScore.

Temporal Relational Reasoning of Large Language Models for Detecting Stock Portfolio Crashes, ArXiv'24

Goal: Proposes a framework using LLMs to detect stock portfolio crashes by analyzing news through temporal-relational reasoning.

Contributions:

- (1) Identifies limitations of LLMs in complex financial reasoning requiring temporal-relational network analysis
- (2) Develops framework simulating human cognitive capabilities for processing financial information across time
- (3) Demonstrates framework's effectiveness through experiments and extends application to crisis detection

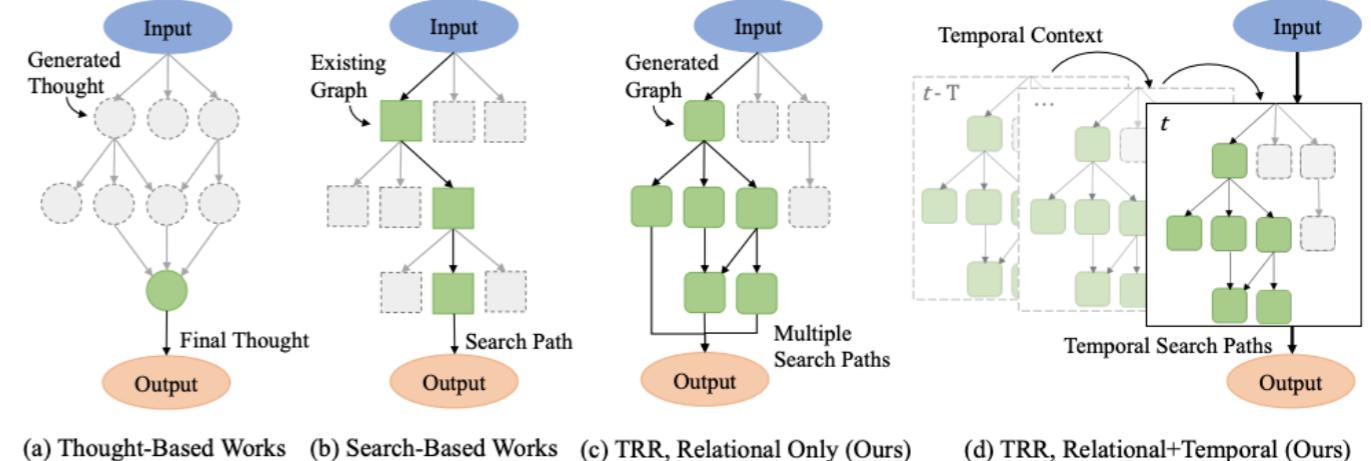


Figure 1: Illustration of thought-based frameworks (e.g., TOT [43], GoT [6]), search-based frameworks (e.g., ToG [43]), and our TRR framework, with Relational Only and Relational+Temporal variants.

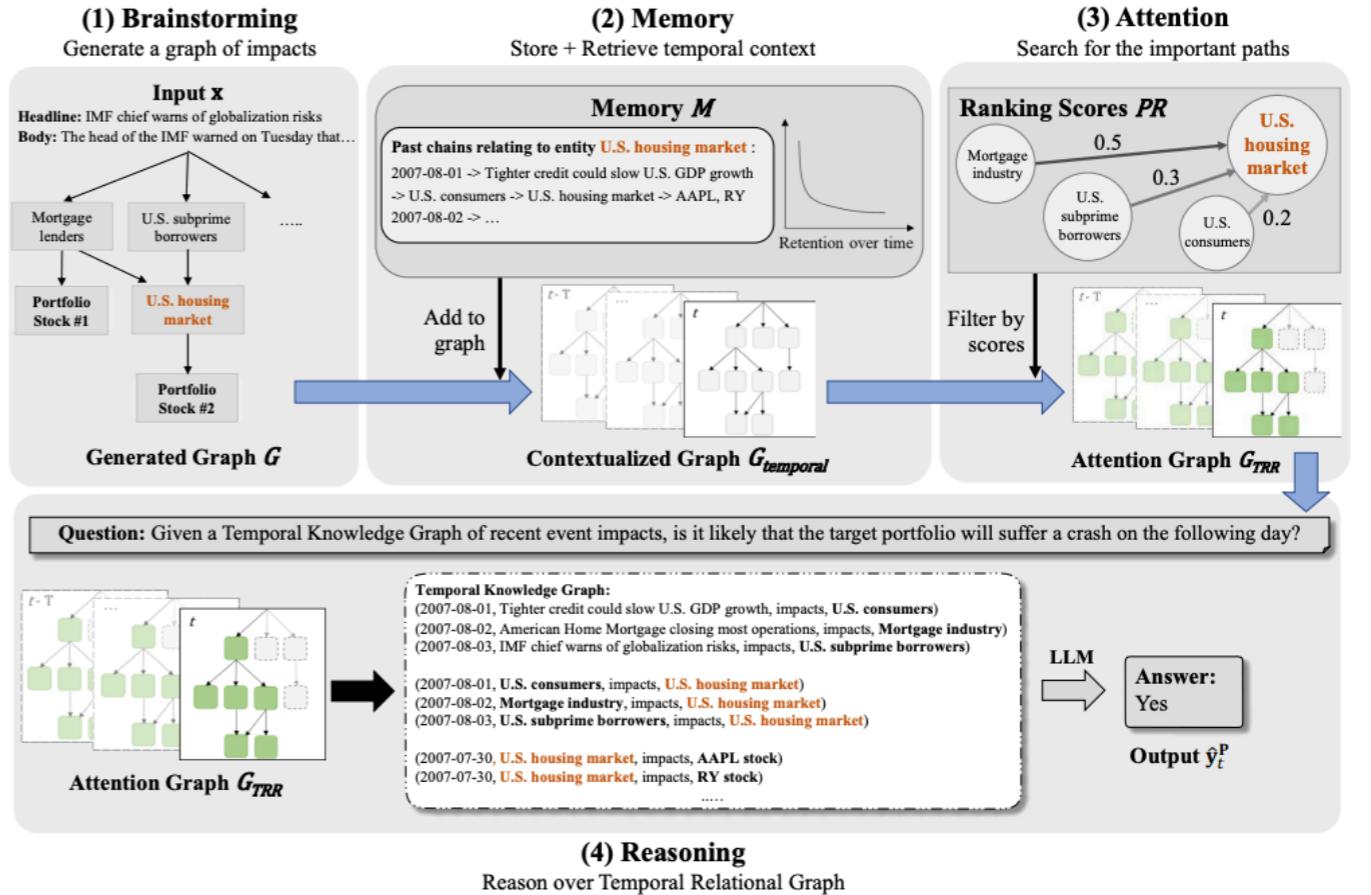


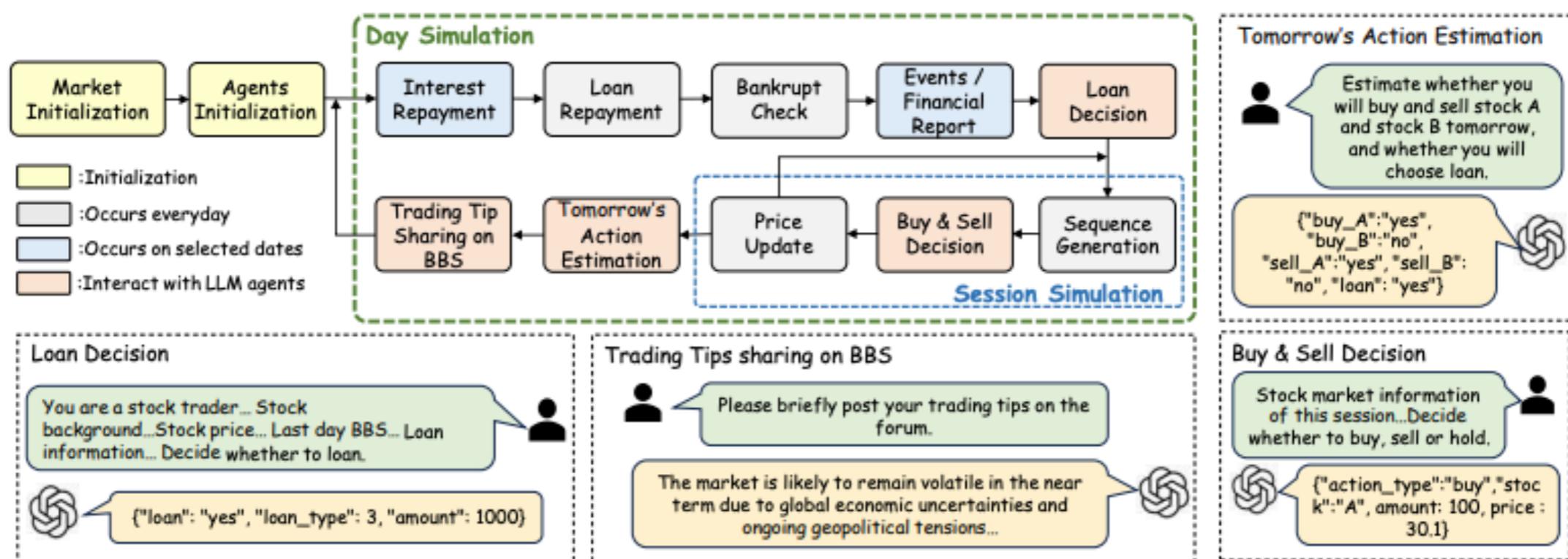
Figure 2: The components of TRR which emulates the human cognitive capabilities used for solving complex problems, that include (1) brainstorming, (2) memory, (3) attention and (4) reasoning skills.

When AI Meets Finance (StockAgent): Large Language Model-based Stock Trading in Simulated Real-world Environments, ArXiv'24

Goal: Develop an multi-LLM-agent system to simulate real-world stock trading and analyze external factors' impact.

Contributions:

- (1) Created StockAgent framework simulating investor behavior through LLM-driven agents in realistic markets
- (2) Demonstrated differing trading patterns between GPT and Gemini models through empirical analysis
- (3) Analyzed how external factors like financial reports affect AI agents' trading decisions



EconAgent: Large Language Model-Empowered Agents for Simulating Macroeconomic Activities, ACL'24

Goal: Develop LLM-powered economic agents for macroeconomic simulation that exhibit human-like decision-making capabilities.

Contributions:

- (1) **First** integration of LLMs into macroeconomic simulations, bridging AI and economics domains.
- (2) Created simulation environment with LLM agents producing realistic macroeconomic phenomena.
- (3) Demonstrated enhanced realism in economic simulations through LLM-based agent decision making.

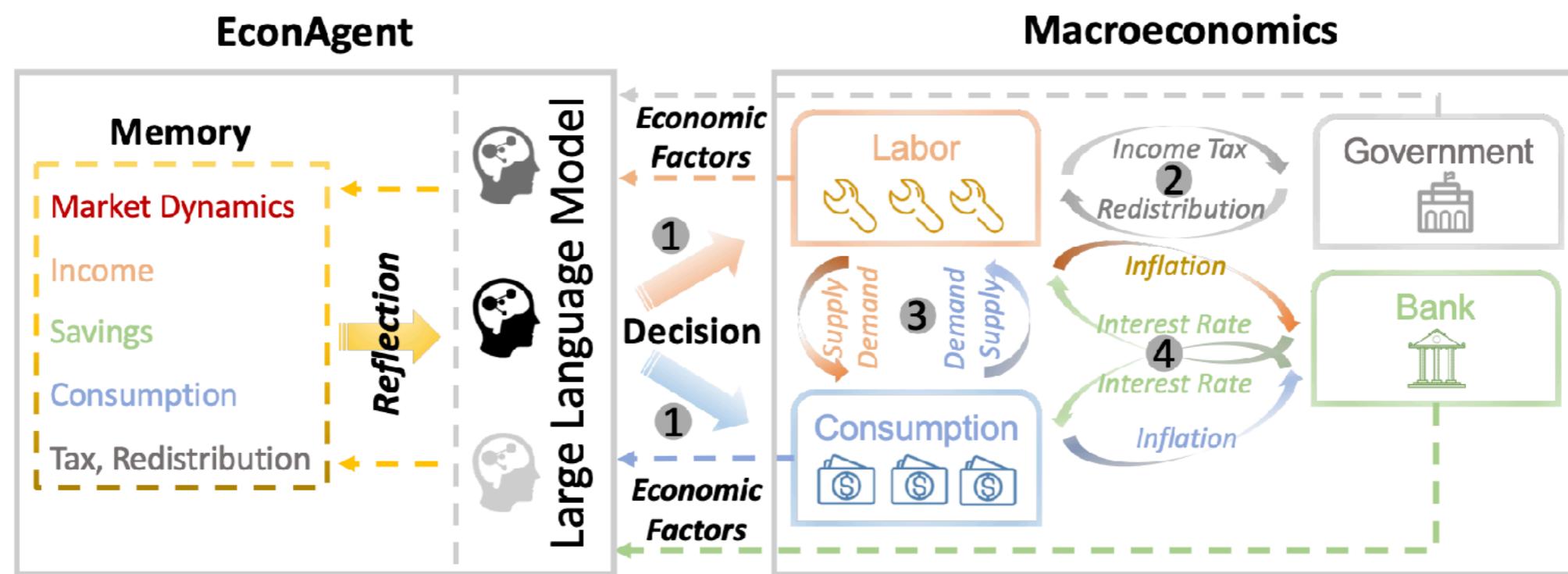


Figure 1: The illustration of our EconAgent (left) and simulation environment (right).

Datasets

Market-GAN: Adding Control to Financial Market Data Generation with Semantic Context, AAAI'24

Goal: Develop a GAN-based financial market simulator that generates data with controllable semantic context.

Contributions:

- (1) Created Contextual Market Dataset combining market dynamics, stock history, and ticker data
- (2) Designed Market-GAN architecture integrating GAN, autoencoder, and knowledge transfer supervisors
- (3) Introduced two-stage training scheme ensuring intrinsic market distribution with multiple objectives

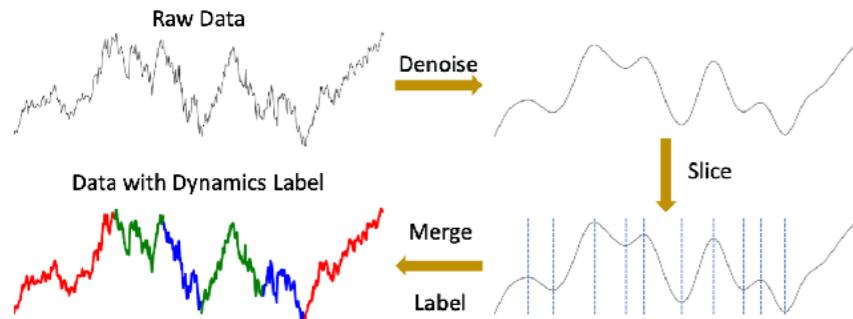


Figure 1: Overview of the market dynamics modeling

Algorithm 1: Market Dynamic Modeling

Input: Time-series x
Parameter: Number of Dynamics $|D|$, Minimum Length l_{min} , Threshold θ , Distance Constraint c
Output: Dynamics Label d

- 1: Denoise x with a low-pass filter into x' .
- 2: Slice x' into batches set B by extremums.
- 3: For each batch b_i in B , merge b_i with b_{i+1} length of $b_i < l_{min}$
- 4: **while** Clustering not converge **do**
- 5: Calculate the slope of each batch b_i with linear regression models and label their dynamics d into D categories by the percentile of their slope. Merge b_i with b_{i+1} if $\text{distance}(b_i, b_{i+1}) < \theta$ and label distance $|d_i - d_{i+1}| \leq c$
- 6: **end while**

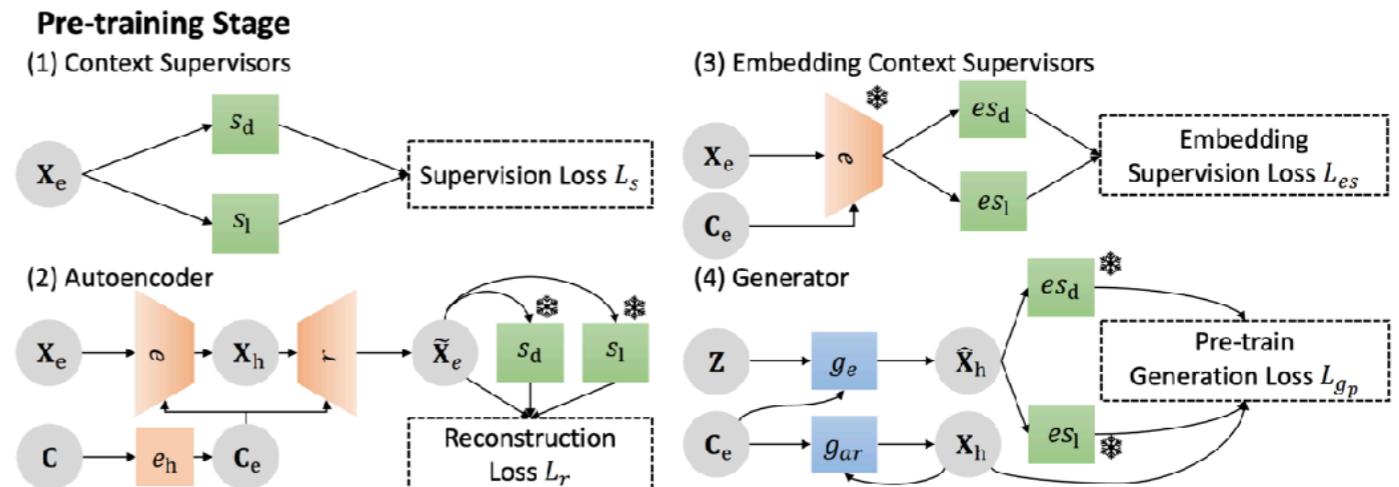


Figure 3: Training scheme of Market-GAN in the pre-training stage. The snowflake indicates the network parameters are frozen in the stage.

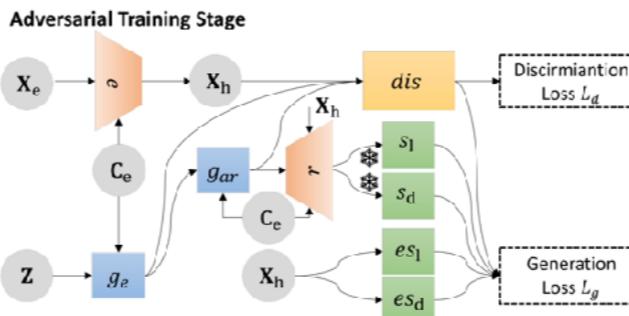


Figure 4: Training scheme of Market-GAN in the adversarial training stage. The snowflake indicates the network parameters of s_l , s_d are frozen in the stage.

Training Set Data	TimesNet	TCN	LSTM	GRU
Only Real Data	1.76 ± 0.71	6.63 ± 0.76	20.9 ± 20.9	17.2 ± 17.9
SigCWGAN w/o TF	2.71 ± 8.07	15.0 ± 26.3	19.1 ± 34.8	17.6 ± 33.7
SigCWGAN w TF	5.64 ± 12.7	18.1 ± 22.8	24.0 ± 25.0	20.7 ± 23.7
TimeGAN w/o TF	1.73 ± 0.72	6.59 ± 0.70	19.5 ± 20.1	16.6 ± 17.5
TimeGAN w TF	1.36 ± 0.48	6.52 ± 0.72	19.1 ± 19.8	14.8 ± 16.3
CGMMN w/o TF	1.74 ± 4.98	8.92 ± 12.3	11.1 ± 19.4	8.74 ± 16.5
CGMMN w TF	1.38 ± 0.50	6.53 ± 0.73	19.1 ± 19.8	14.8 ± 16.3
RCGAN w/o TF	1.68 ± 0.67	6.64 ± 0.83	16.6 ± 18.1	13.9 ± 15.5
RCGAN w TF	1.40 ± 0.55	6.70 ± 1.91	19.2 ± 19.7	15.0 ± 16.2
Market-GAN(CTB)	1.26 ± 0.75	6.48 ± 0.73	12.9 ± 16.7	11.1 ± 14.2
Market-GAN(RNN)	1.11 ± 0.66	6.39 ± 0.92	9.89 ± 14.9	8.41 ± 12.4

Table 4: Prediction SMAPE loss on the test set using generated data of different models to augment the training set. TF is the abbreviation of the data transformation layer.

FNSPID: A Comprehensive Financial News Dataset in Time Series, KDD'24

Goal: A comprehensive financial dataset integrating stock prices and news sentiment for market prediction.

Contributions:

- (1) Built large-scale dataset combining 29.7M stock prices with 15.7M news records
- (2) Demonstrated dataset size improves prediction accuracy and sentiment enhances transformer models
- (3) Developed reproducible procedure for continuous dataset updates and maintenance

	FNSPID (ours)	Reuters	Benzinga	Bloomberg	Lenta	Lutz's	Farimani's	SemEval*	SEntFiN 1.0
Time Stamp	Yes	Yes	Yes	Yes	Yes	No	No	No	No
Text Type	Article	Article	Article	Article	Article	Sentence	Sentence	Headline	Headline
Number of News	15698563	8556324	3252885	447341	800974	1000	21867	1142	10753
Symbol	Yes	No	Yes	No	No	No	No	No	No
Summarization	Yes	No	No	No	No	No	Yes	No	No
Sentiment Score	Integer	-	-	-	-	Integer	-	Real	Integer
URL	Yes	No	Yes	No	No	No	No	No	No
Language	Many	Eng	Eng	Eng	Ru	Eng	Eng	Eng	Eng
Stock Price	Yes	No	No	No	No	No	Yes	No	No

Table 1: Comparison of existing datasets for Time Series Financial Analysis. FNSPID stands out with the highest volume of news data and includes unique features not found in other benchmark datasets. In the label, SemEval* stands for SemEval-2017 Task5 dataset.

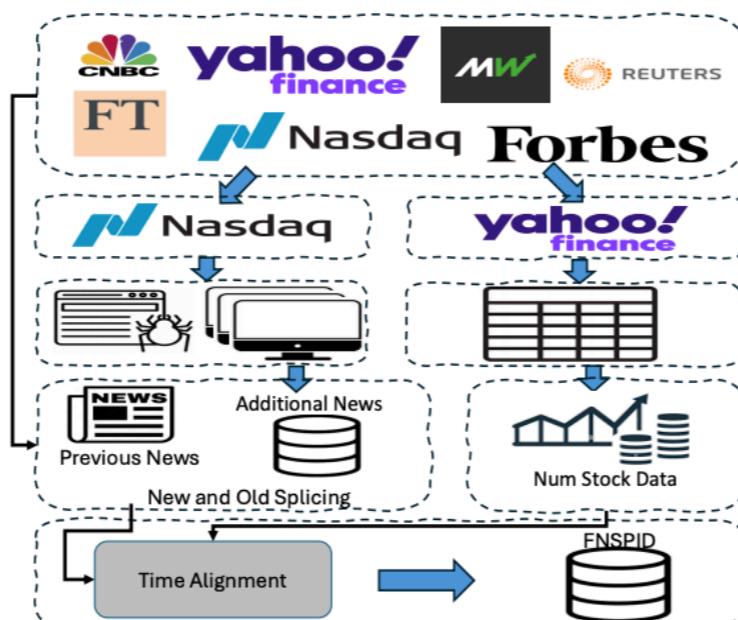


Figure 1: Data Collection Process from website selection in the first level box; data segmentation in second level boxes; data collection for web scraping on left and numerical data collection on right; data organization on fourth level boxes and final FNSPID build-up on the last level box.

System: Forget all your previous instructions. You are a financial expert with stock recommendation experience. Based on a specific stock, score for range from 1 to 5, where 1 is negative, 2 is somewhat negative, 3 is neutral, 4 is somewhat positive, 5 is positive. 10 summarized news will be passed in each time, you will give score in format as shown below in the response from assistant.

User: "News to Stock Symbol -- AAPL: Apple (AAPL) increase 22% ## News to Stock Symbol -- AAPL: Apple (AAPL) price decreased 30% ## News to Stock Symbol -- MSFT: Microsoft (MSFT) price has no change"
ASSISTANT: "5, 1, 3"
User: "News to Stock Symbol -- AAPL: Apple (AAPL) announced iPhone 15 ## News to Stock Symbol -- AAPL: Apple (AAPL) will release VisonPro on Feb 2, 2024"
Assistant: "4, 4"
USER: ## News to Stock Symbol -- {symbol}: {text}

Date	2022-06-03 00:00:00
Symbol	AAPL
Headline	Consider Alphabet Stock Even in a Recession
Text	After six straight red weeks, the bulls may rejoice with two consecutive green days. This is where the fear of missing out kicks in for most investors and they blindly jump back in. Today, we will contemplate the prospects of doing so with Alphabet (NASDAQ:GOOG,NASDAQ:GOOGL) stock. But first, we should discuss the bigger...
URL	https://www.nasdaq.com/articles/consider-alphabet-stock-even-in-a-recession
LSA Sum	But investors will be shy about risking money if they think a big recession is coming. 7 Overlooked Value Stocks to Buy Before Wall Street Catches On ...
Luhn Sum	Today, we will contemplate the prospects of doing so with Alphabet (NASDAQ:GOOG,NASDAQ:GOOGL) stock. Judging by their statements, they...
TexRank Sum	The reason why experts are now calling for disaster is the rhetoric from the Fed. Ticker Company Price GOOG Alphabet Inc. \$2,202.40 GOOG Stock....
LexRank Sum	These are conditions that Wall Street deems as recessionary. Current investors of GOOG stock have realistic expectation..

AlphaFin: Benchmarking Financial Analysis with Retrieval Augmented Stock-Chain Framework, COLING'24

Goal: create a framework for financial analysis combining stock prediction and question-answering using LLMs

Contributions:

- (1) Introduces AlphaFin dataset combining financial reports, news, and chain-of-thought data for LLM training
- (2) Develops Stock-Chain framework integrating retrieval-augmented generation for real-time financial analysis
- (3) Demonstrates superior performance with 30.8% annualized return and enhanced question-answering capabilities

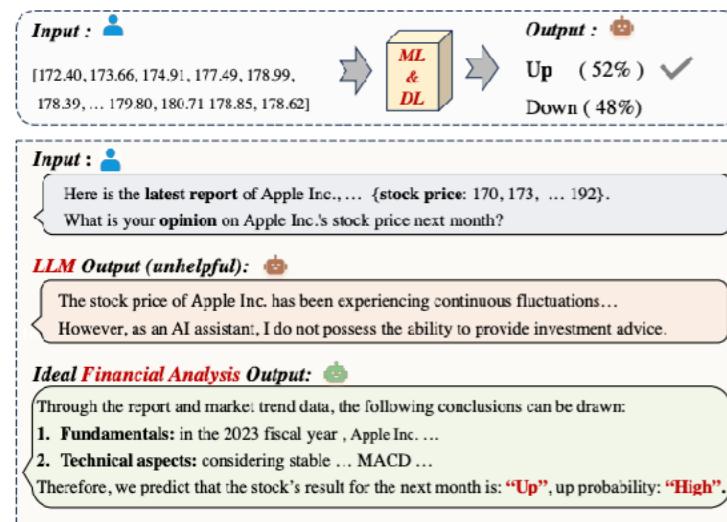


Figure 1: An example of the financial analysis task, including stock trend prediction and financial Q&A. Traditional ML&DL methods merely provide uncertain forecasts (Up/Down) without any justification, while original LLMs could offer analysis of the prediction but unhelpful.

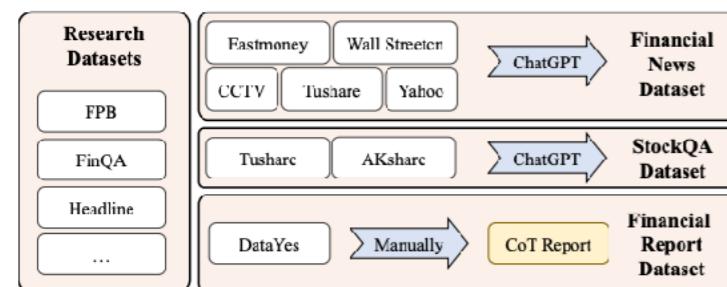


Figure 2: The data source and preprocessing of the proposed AlphaFin datasets.

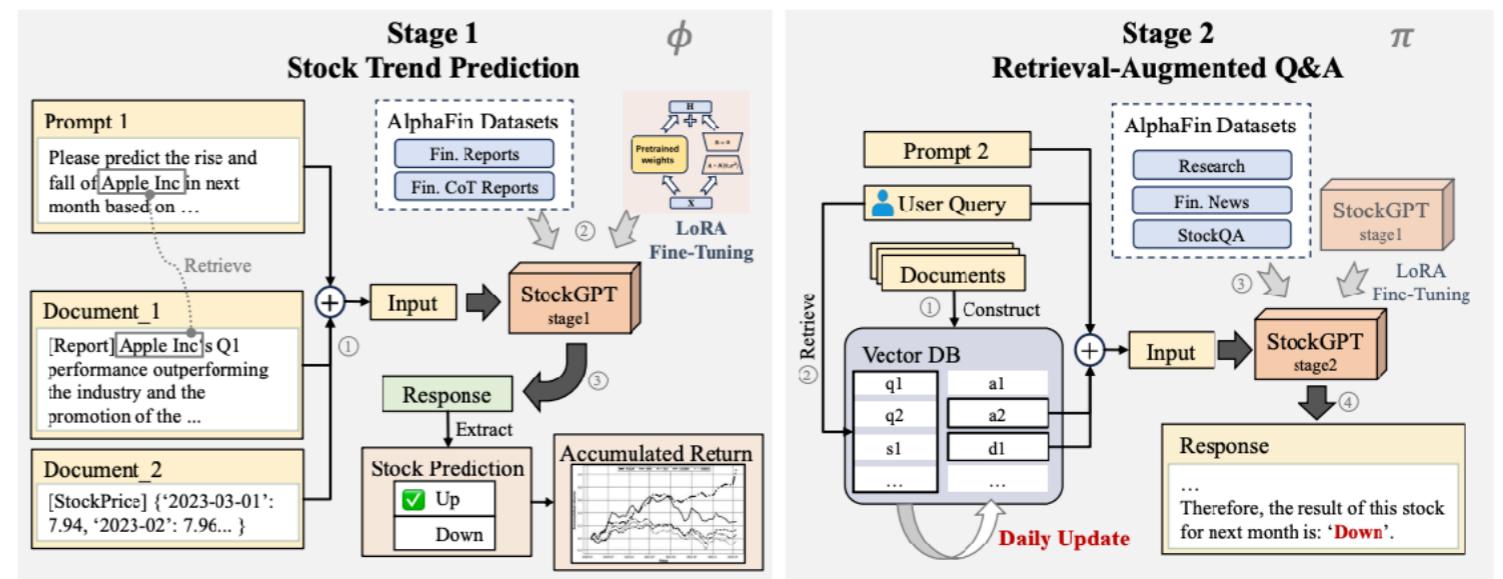
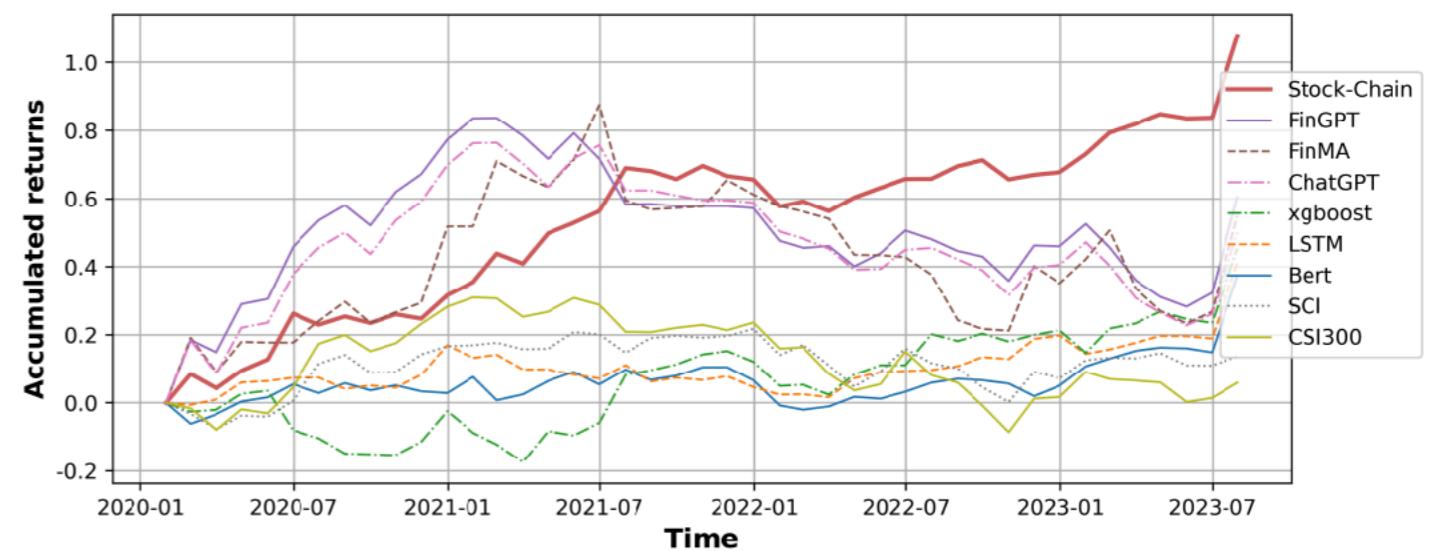


Figure 3: An illustration of the Stock-Chain framework of the two stages in financial analysis.



StockEmotions: Discover Investor Emotions for Financial Sentiment Analysis and Multivariate Time Series, AAAI'23

Goal: Create and analyze a dataset of stock market emotions for financial sentiment analysis and forecasting.

Contributions:

- (1) Created StockEmotions dataset with 10k annotated comments capturing investor psychology in markets
- (2) Developed a hybrid human-AI annotation pipeline combining language models and expert input
- (3) Demonstrated dataset's value through sentiment classification and time series prediction experiments

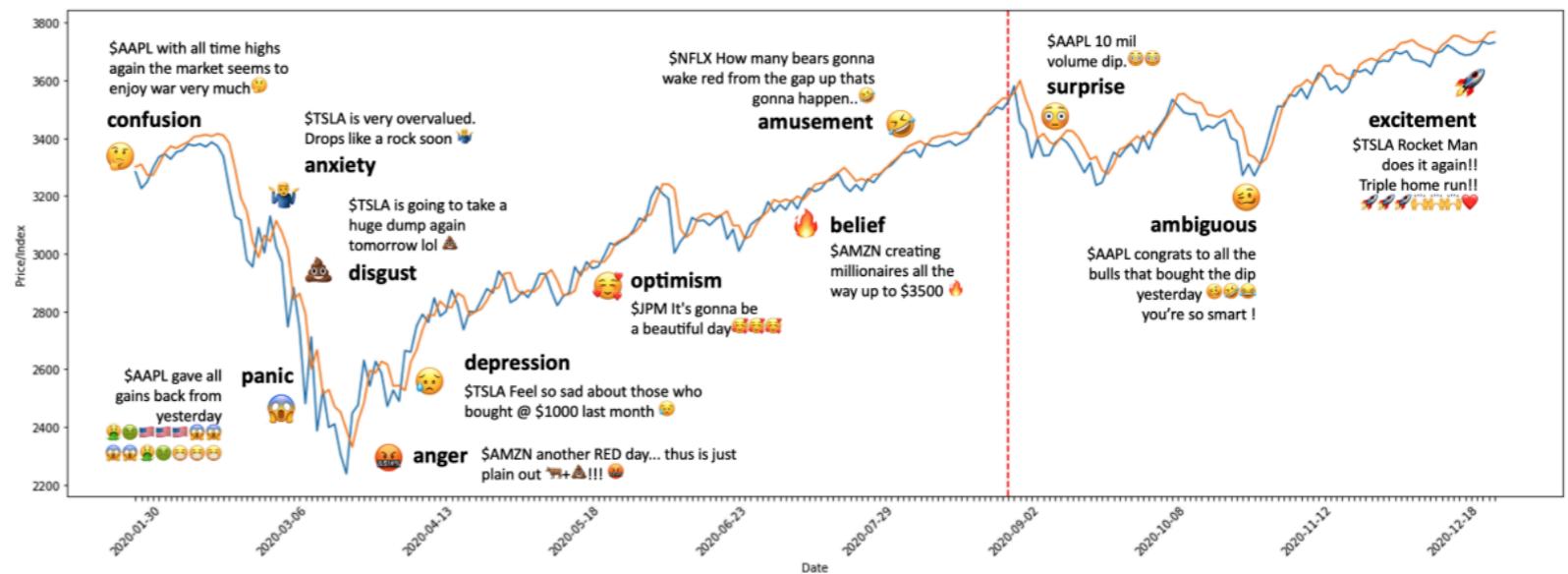


Figure 1: Example from StockEmotions dataset showing investor psychology on the stock market. A combination of input data (stock price index, text and emoji, and emotion label) is used on a Temporal Attention LSTM for multivariate time series forecasting. (blue line = the actual S&P index, orange line = the prediction S&P index with a rolling window size 5, red vertical line = the point of data splitting; Further details are in the experiments section.)

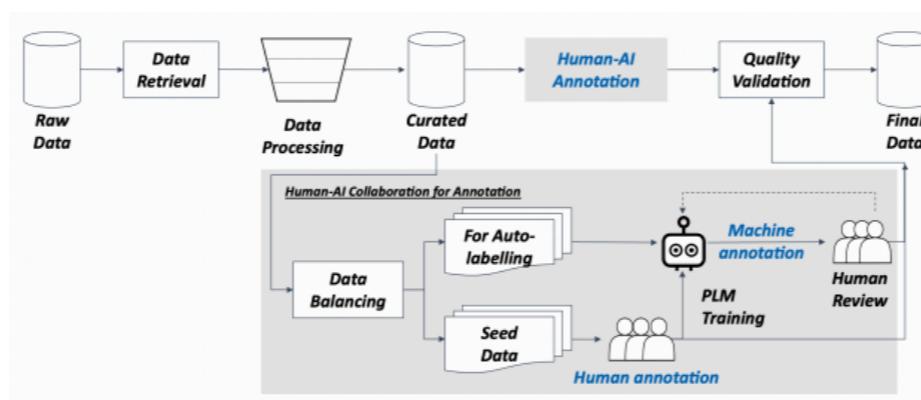


Figure 2: An overview of dataset creation pipeline.

Number of Utterance	10,000
Number of Sentiment	2 - bullish (55%), bearish (45%)
	12 - ambiguous(9%), amusement(8%), anger(4%), anxiety(14%), belief(9%), confusion(6%), depression(2%), disgust(13%), excitement(14%), optimism (16%), panic(3%), surprise(3%)
Number of Emotion	
Avg. Length	19.2 tokens per utterance
Unique Emoji	761
Time Period	01 Jan 2020 - 31 Dec 2020

Table 1: Key statistics of StockEmotions. Each label shows the proportion in the total dataset.

Time Series Models

Timer: Generative Pre-trained Transformers Are Large Time Series Models, ICML'24

Goal: Develop a large-scale pre-trained transformer model for time series analysis in data-scarce scenarios.

Contributions:

- (1) Created UTSD dataset with 1B time points and unified format for heterogeneous time series
- (2) Developed Timer, a generative pre-trained transformer showing strong few-shot generalization capabilities
- (3) Demonstrated versatility across forecasting, imputation, and anomaly detection with state-of-the-art results

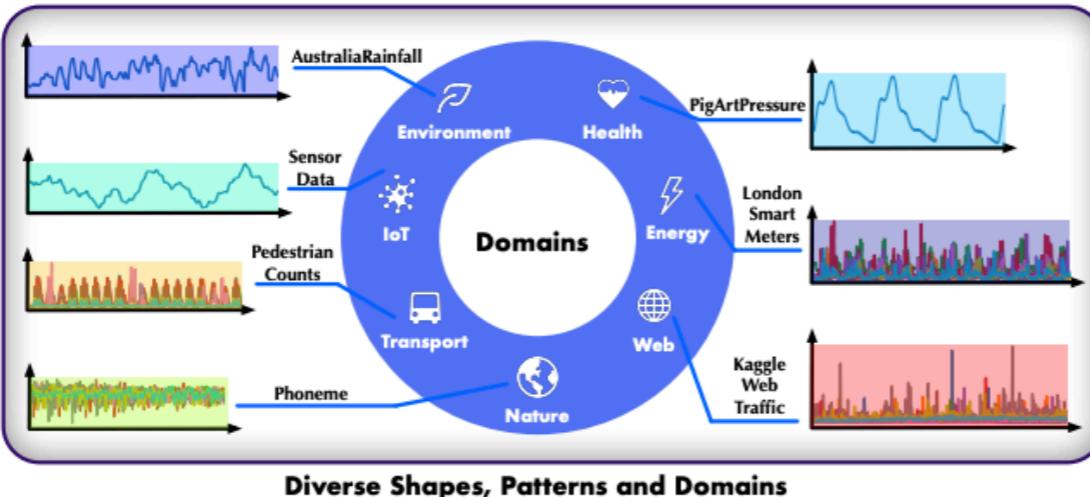
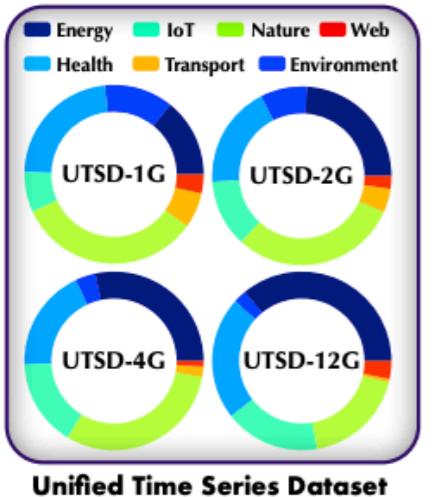


Figure 2. Illustration of Unified Time Series Dataset (UTSD) that is composed of various time series domains with hierarchical capacities.

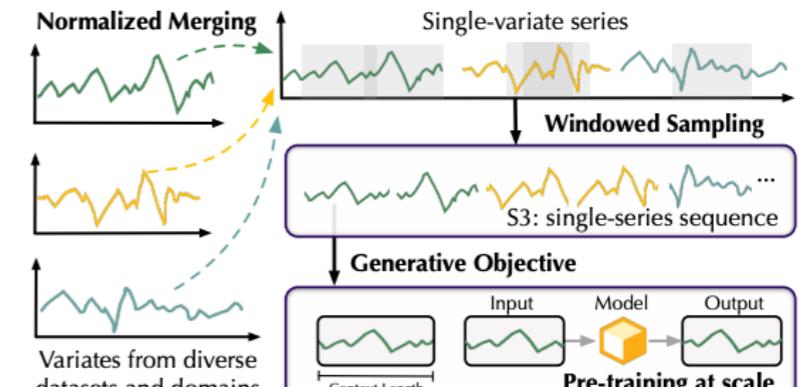


Figure 3. Pre-training strategy for heterogeneous time series.

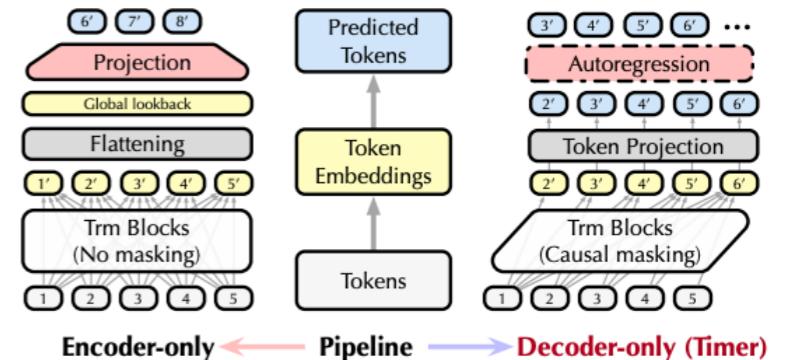


Figure 4. Architectures of typical Transformer-based forecasters.

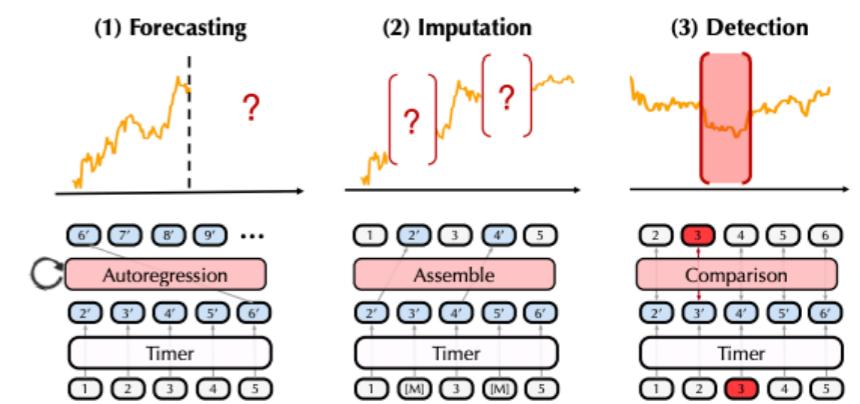


Figure 5. Illustration of our generative task unification: (1) Generative pre-trained Timer can naturally predict the next series by the iterative autoregression; (2) By introducing masked tokens during adaptation, Timer generates imputations with the previous context and assemble them with the observed part; (3) We propose predictive anomaly detection by predicting normal series in advance.

MOMENT: A Family of Open Time-series Foundation Models, ICML'24

Goal: Introduces MOMENT, an open-source foundation model family for general-purpose time series analysis.

Contributions:

- (1) Creates Time Series Pile, a large diverse public dataset for pre-training models
- (2) Develops strategies to handle multi-dataset training challenges in time series modeling
- (3) Establishes comprehensive benchmark framework for evaluating limited-supervision time series models

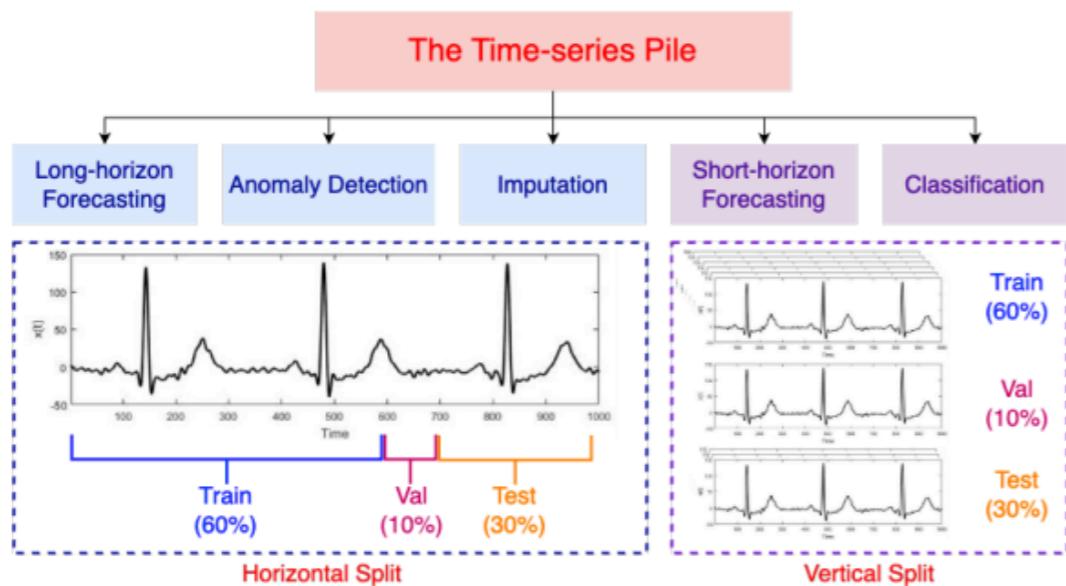


Figure 2. Time Series Pile data splits. To avoid data contamination, we carefully partition all datasets into disjoint train, validation, and test splits. We adhere to the predefined splits provided by the creators of each dataset. In cases where such splits are unavailable, we randomly sample 60% of the data for training, 10% for validation, and 30% for testing. We only use the training splits of all datasets for pre-training.

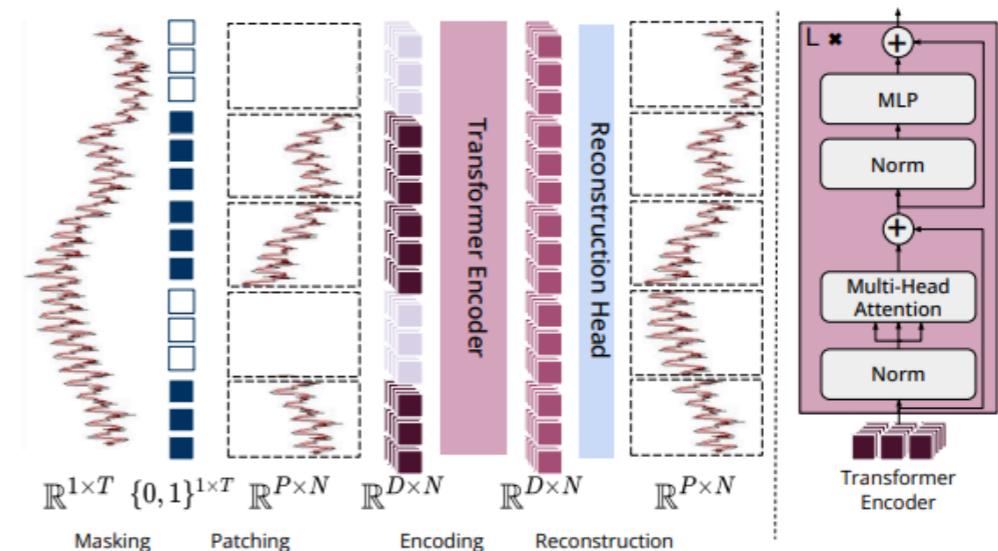


Figure 3. Overview of MOMENT. A time series is broken into disjoint fixed-length sub-sequences called patches, and each patch is mapped into a D -dimensional patch embedding. During pre-training, we mask patches uniformly at random by replacing their patch embeddings using a special mask embedding [MASK]. The goal of pre-training is to learn patch embeddings which can be used to reconstruct the input time series using a light-weight reconstruction head.

TimeMixer: Decomposable Multiscale Mixing for Time Series Forecasting, ICLR'24

Goal: Proposes TimeMixer, a novel MLP-based model for time series forecasting using multiscale mixing.

Contributions:

- (1) Introduces multiscale mixing approach that disentangles temporal variations at different sampling scales
- (2) Develops Past-Decomposable-Mixing block to handle seasonal and trend components separately efficiently
- (3) Creates Future-Multipredictor-Mixing block that combines predictions from multiple scales effectively

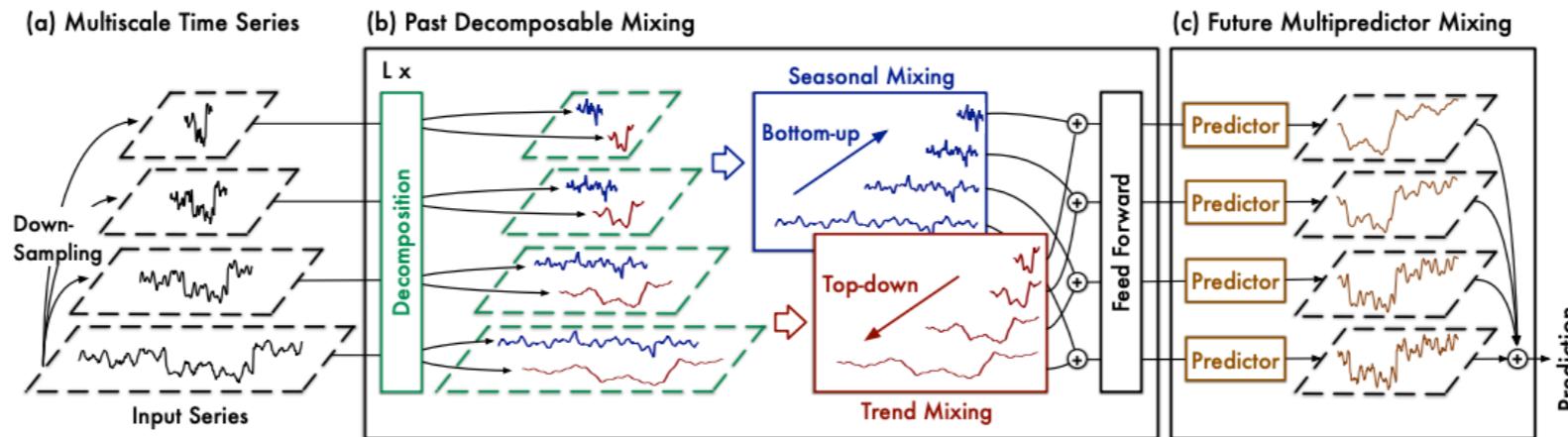


Figure 1: Overall architecture of TimeMixer, which consists of Past-Decomposable Mixing and Future-Multipredictor-Mixing for past observations and future predictions respectively.

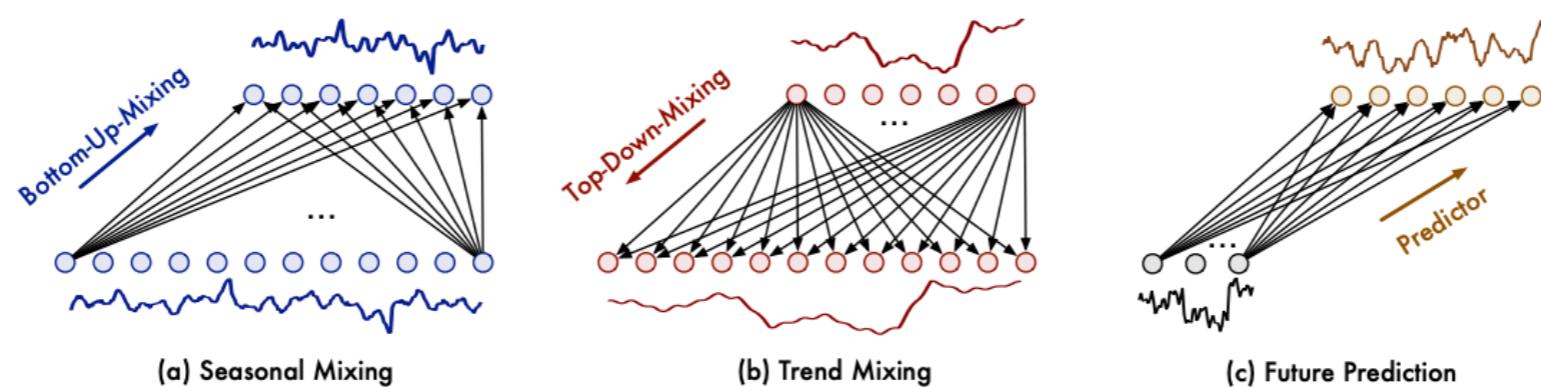


Figure 2: The temporal linear layer in seasonal mixing (a), trend mixing (b) and future prediction (c).

TimesNet: Temporal 2d-variation modeling for general time series analysis, ICLR'23

Goal: Propose TimesNet, a general-purpose neural network for diverse time series analysis tasks.

Contributions:

- (1) Introduces multi-periodicity analysis to disentangle complex temporal patterns into simpler variations
- (2) Transforms 1D time series into 2D space enabling better modeling of temporal relationships
- (3) Achieves state-of-the-art performance across five major time series tasks with unified architecture

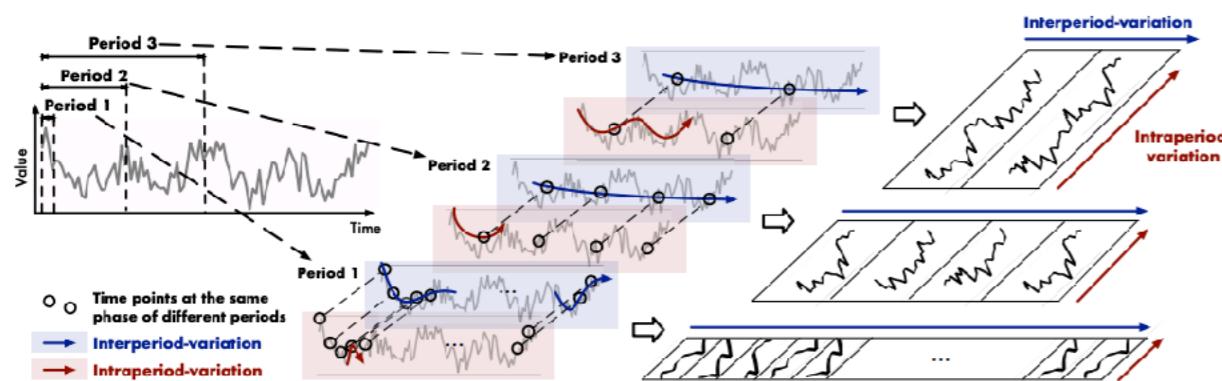


Figure 1: Multi-periodicity and temporal 2D-variation of time series. Each period involves the **intraperiod-variation** and **interperiod-variation**. We transform the original 1D time series into a set of 2D tensors based on multiple periods, which can unify the intraperiod- and interperiod-variations.

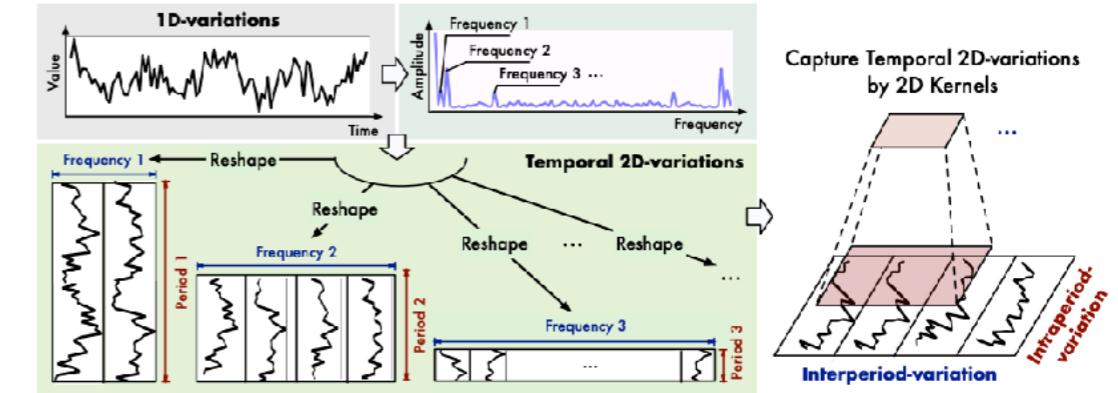


Figure 2: A univariate example to illustrate 2D structure in time series. By discovering the periodicity, we can transform the original 1D time series into structured 2D tensors, which can be processed by 2D kernels conveniently. By conducting the same reshape operation to all variates of time series, we can extend the above process to multivariate time series.

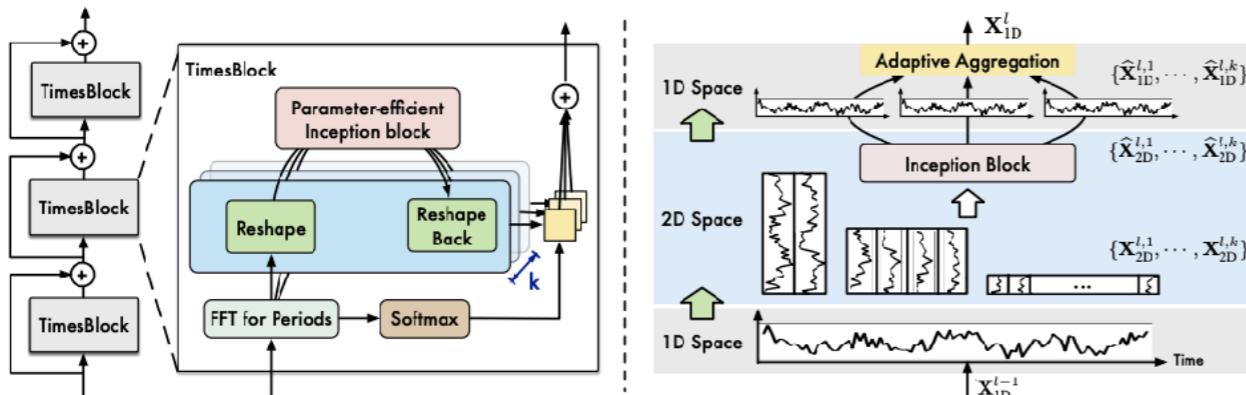


Figure 3: Overall architecture of TimesNet. TimesNet is stacked by TimesBlocks in a residual way. TimesBlocks can capture various temporal 2D-variations from k different reshaped tensors by a parameter-efficient inception block in 2D space and fuse them based on normalized amplitude values.

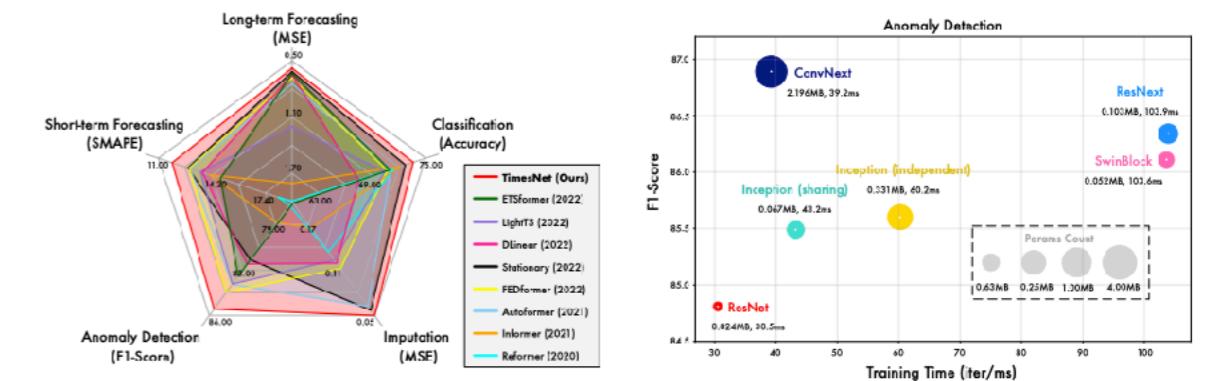


Figure 4: Model performance comparison (left) and generality in different vision backbones (right).

A Time Series is Worth 64 Words: Long-term Forecasting with Transformers, ICLR'23

Goal: Proposes an efficient Transformer model for time series forecasting using patches and channel-independence.

Contributions:

- (1) Introduces patch-based time series segmentation to capture local patterns and reduce computation.
- (2) Implements channel-independence architecture allowing separate processing of individual time series features.
- (3) Demonstrates superior performance in both supervised learning and self-supervised representation transfer.

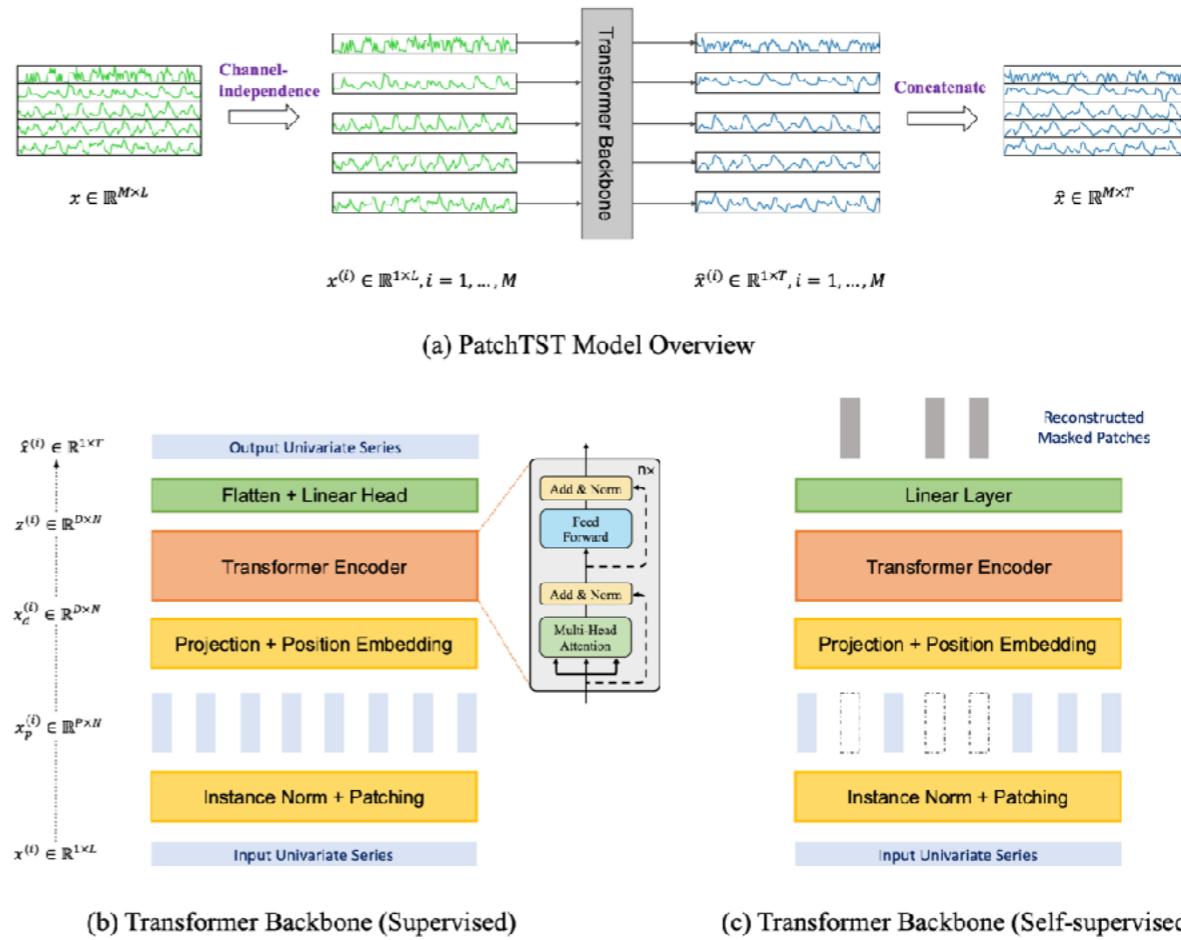


Figure 1: PatchTST architecture. (a) Multivariate time series data is divided into different channels. They share the same Transformer backbone, but the forward processes are independent. (b) Each channel univariate series is passed through instance normalization operator and segmented into patches. These patches are used as Transformer input tokens. (c) Masked self-supervised representation learning with PatchTST where patches are randomly selected and set to zero. The model will reconstruct the masked patches.

Thank you

Full list of papers: <https://github.com/junhua/awesome-finance-ai-papers>