Retrieval Augmented Time Series Forecasting

Sungwon Han, Seungeon Lee, Meeyoung Cha, Sercan O Arik, Jinsung Yoon

- ICML 2025

Introduction

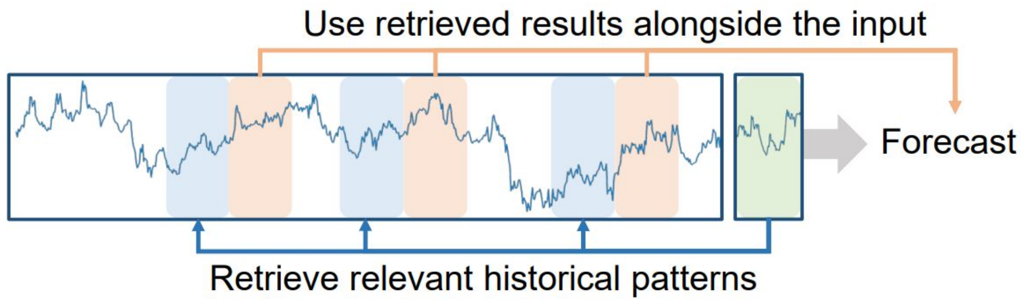
Limitation of Existing Time-Series Forecasting Models

- Deep-learning models have boosted time-series forecasting accuracy, yet they still struggle with **complex, non-stationary and rare** patterns.

- **Memorizing** every possible pattern in model weights is **inefficient** and **risks overfitting**.

Our Approach

- We introduce a **lightweight retrieval module** to externalize pattern knowledge, relieving the learning burden of the estimator.

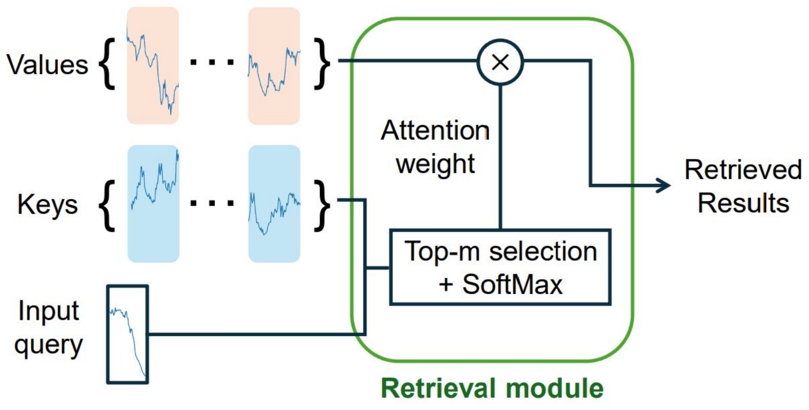


Retrieval Module Architecture

- With given input, slide a window over the full history (training dataset).
    - Store the **window as key** and its **immediate future as value**.

- Keep the top-m similar keys, convert similarity scores to weights.

- Return **weighted sum** of their value patches as retrieved results.

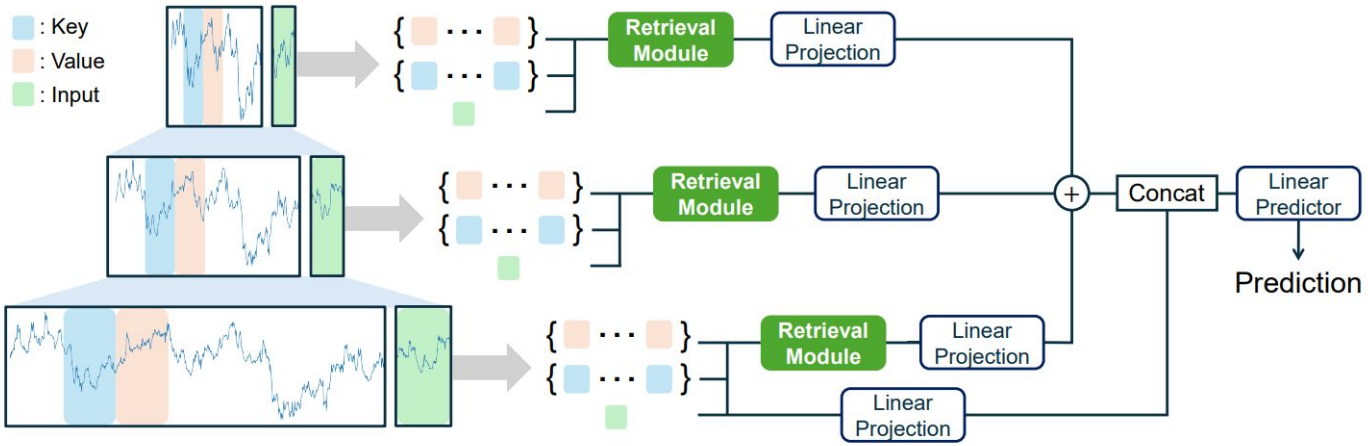


Main Model: RAFT

Extension with multi-periodicity

- Downsample input series with periods {1, 2, 4} to capture short- and long-term structure; independent retrieval path for each view.

- Project each period's retrieved vector into a common space, sum across periods, then concatenate with raw input features.



Highlighted Results

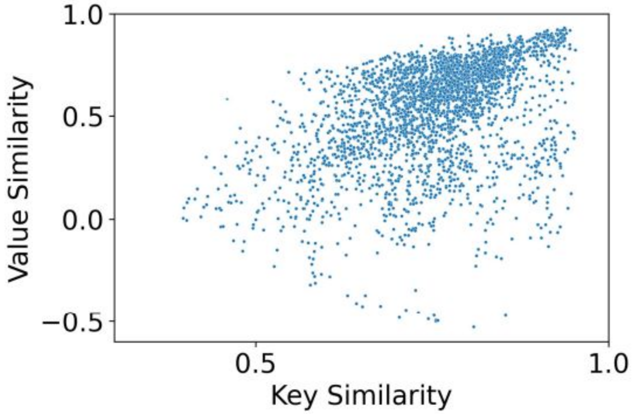
- Across 10 public benchmarks (ETT × 4, Electricity, Exchange, Illness, Solar, Traffic, Weather), RAFT attains an **86% average win ratio** over nine state-of-the-art baselines.

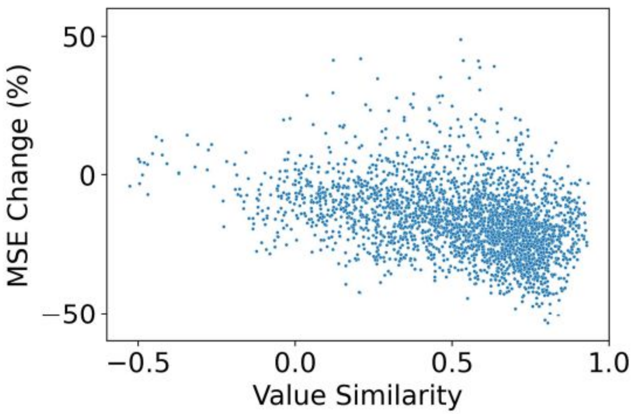
- Plugging the same retrieval module into **Transformer architecture** (e.g., AutoFormer) yields **consistent gains** (e.g., ETTh1 0.496 → 0.471 MSE).

Methods	RAFT	TimeMixer	PatchTST	TimesNet	MICN	DLinear	FEDformer	Stationary	Autoformer	Informer
ETTh1	**0.420**	0.447	0.516	0.495	0.475	0.461	0.498	0.570	0.496	1.040
ETTh2	**0.359**	0.364	0.391	0.414	0.574	0.563	0.437	0.526	0.450	4.431
ETTm1	**0.348**	0.381	0.406	0.400	0.423	0.404	0.448	0.481	0.588	0.961
ETTm2	**0.254**	0.275	0.290	0.291	0.353	0.354	0.305	0.306	0.327	1.410
Electricity	**0.160**	0.182	0.216	0.193	0.196	0.225	0.214	0.193	0.227	0.311
Exchange	0.441	**0.386**	0.564	0.416	0.315	0.643	1.195	0.461	1.447	2.478
Illness	2.097	2.024	**1.480**	2.139	2.664	2.169	2.847	2.077	3.006	5.137
Solar	0.231	**0.216**	0.287	0.403	0.283	0.330	0.328	0.350	0.586	0.331
Traffic	**0.434**	0.484	0.529	0.620	0.593	0.625	0.610	0.624	0.628	0.764
Weather	0.241	**0.240**	0.265	0.251	0.268	0.265	0.309	0.288	0.338	0.634

Discussion #1

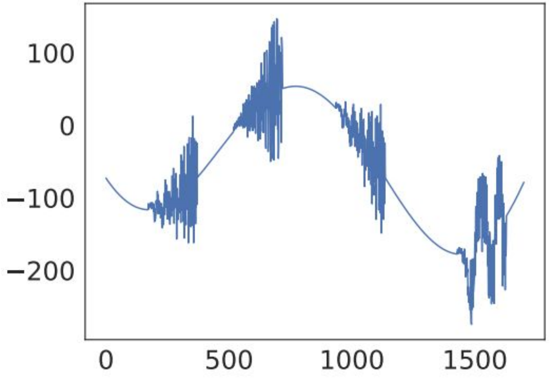
- Retrieval quality matters in forecasting performance.





Discussion #2

- Retrieval is more helpful when patterns are rare and temporally less correlated



Pattern occurrences	1	2	4
TimeMixer	0.2863	0.2305	0.2249
TimeNet	0.2448	0.1877	0.1938
MICN	0.2536	0.2445	0.2450
DLinear	0.3175	0.2059	0.2798
RAFT without retrieval	0.2694	0.2649	0.1894
RAFT with retrieval	0.1845	0.1818	0.1592
MSE decrease ratio	-31.5%	-31.4%	-16.0%

Summary

- **RAFT** introduces **retrieval** into time-series forecasting, externalizing pattern knowledge and reducing the model's learning burden.

- RAFT consistently **outperforms existing baselines** across 10 public benchmarks.

- Analysis shows that **retrieval quality correlates with performance gain**, and retrieval is especially **effective** when patterns are **rare and temporally uncorrelated**.

Takeaway: Adopting retrieval improves the performance of time-series forecasting.

Code:

Paper: