

# 3 Ingredients to cook MOMENT

## The Time-series Pile

Public time-series datasets are small, task-specific and scattered

**The Time-series Pile:** Large repository of public time-series spanning diverse domains, frequencies & modeling tasks

**Collection of 4 task-specific public benchmarks** for long & short-horizon forecasting, anomaly detection, and classification

## Multi-dataset pre-training

Diverse time-series characteristics (lengths, amplitudes, frequencies) make multi-dataset pre-training hard

**Systematically solutions** to time-series challenges: Channel independence, fixed-length inputs, reversible instance normalization

## Evaluation

Holistic benchmarks to evaluate models on diverse datasets and tasks in limited supervision settings in early stages

**Carefully designed benchmark** on 5 time-series modeling tasks: short- and long-horizon forecasting, anomaly detection, classification, and imputation

Focus on evaluating MOMENT against both SoTA deep learning and statistical baselines, on many task-specific datasets, using multiple evaluation metrics, exclusively in limited supervision settings (e.g., zero-shot imputation)

# MOMENT: A Family of Open Time-series Foundation Models

Mononito Goswami<sup>1</sup>, Konrad Szafer<sup>\*1</sup>, Arjun Choudhry<sup>\*1</sup>, Yifu Cai<sup>1</sup>, Shuo Li<sup>2</sup>, Artur Dubrawski<sup>1</sup>

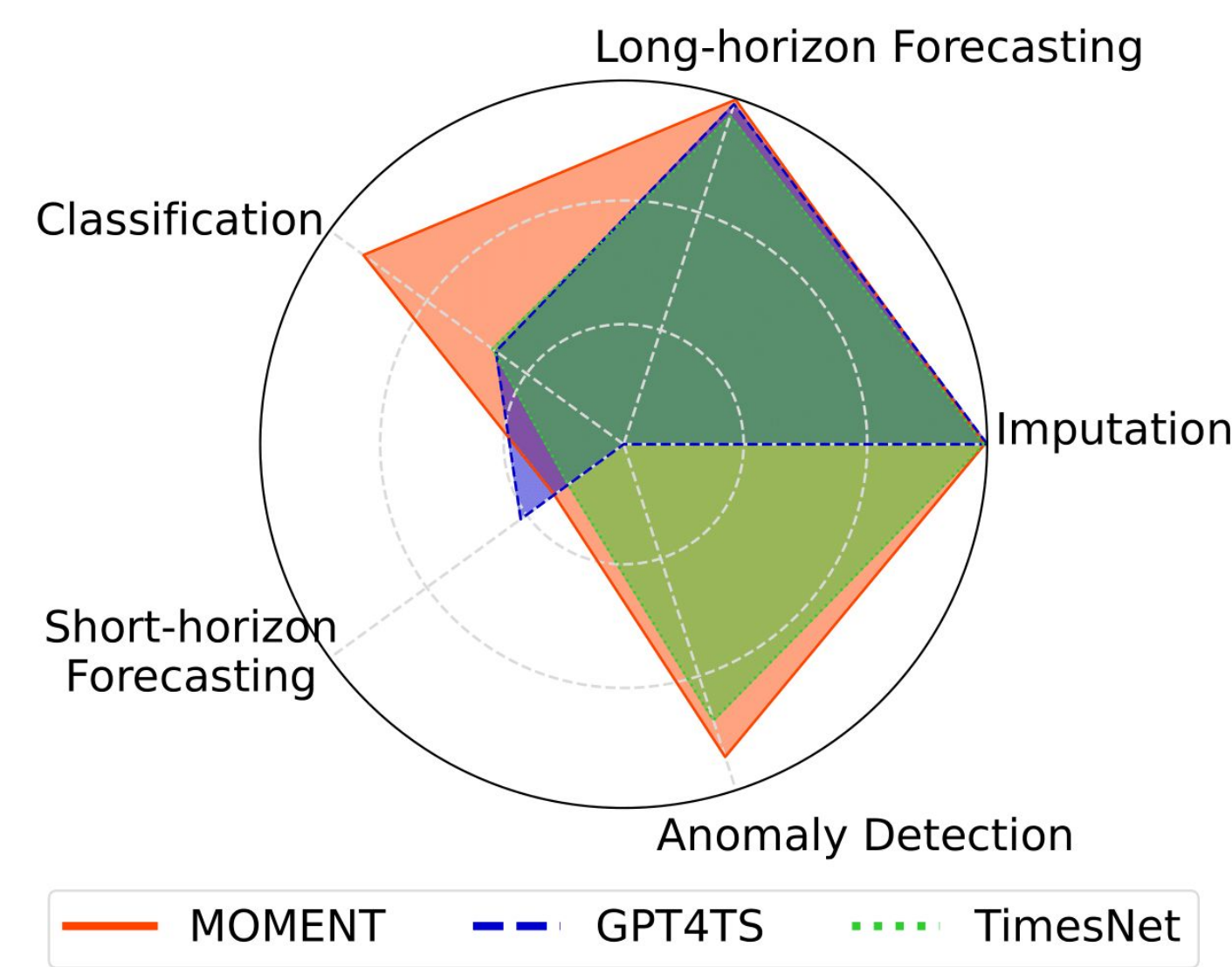
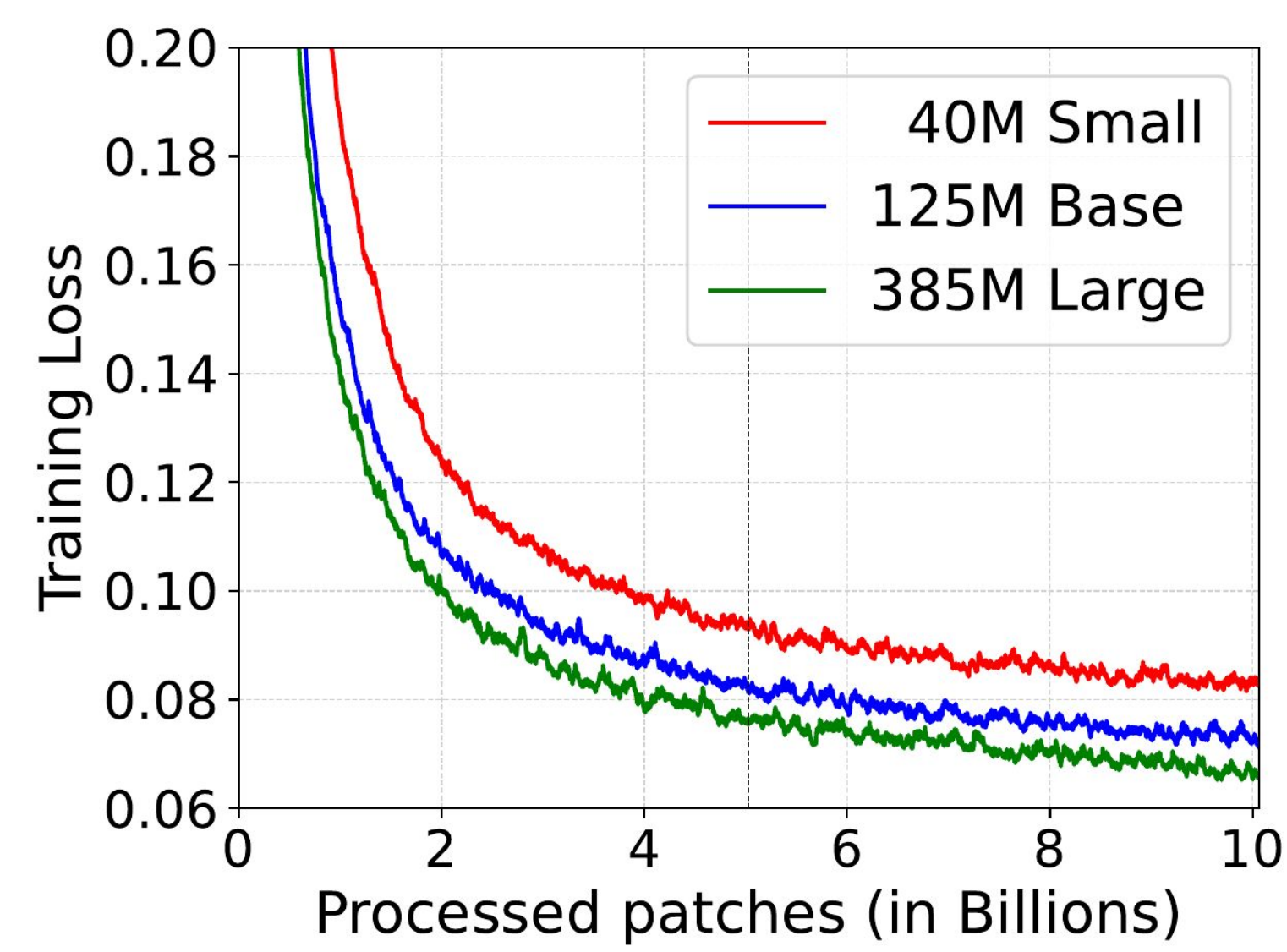
<sup>1</sup>Auton Lab, Robotics Institute, Carnegie Mellon University  
<sup>2</sup>University of Pennsylvania

## TL;DR

We introduce the Time-series Pile, a large repository of public time-series, use it to pre-train a family of large models. We test these models on a benchmark with diverse tasks and datasets in limited supervision settings.



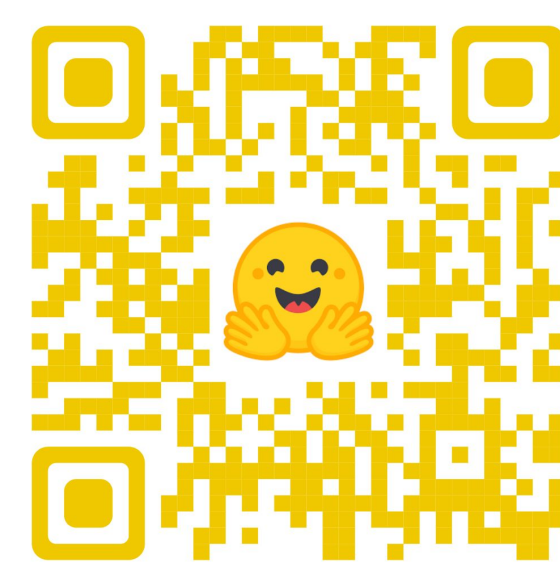
Modeling time-series data, now just a MOMENT away!



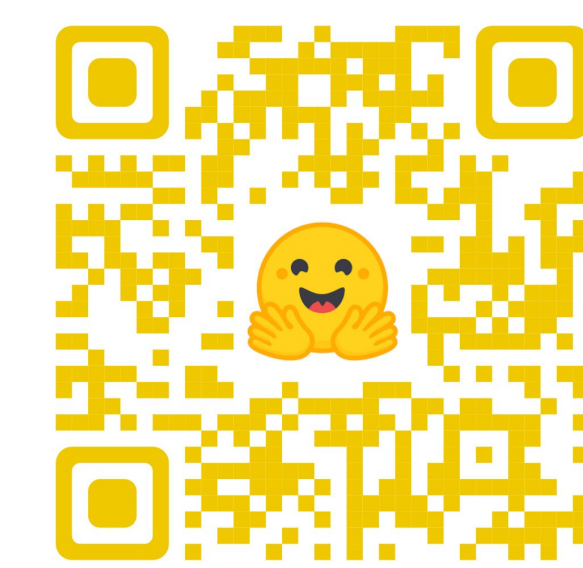
Paper



Research code



Model Weights



Time-series Pile

# Findings



## MOMENT does well on multiple tasks

Long-horizon forecasting

Methods/Metric	MOMENT <sub>L</sub>		Time-LLM		GPT4TS		PatchTST		DLinar		TimesNet		FEDFormer		Stationary		LightTS		N-BEATS		
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	
Weather	96	0.154	0.209	-	-	0.162	0.212	0.149	0.198	0.176	0.237	0.172	0.220	0.217	0.296	0.173	0.223	0.182	0.242	0.152	0.210
	192	0.197	0.248	-	-	0.204	0.248	0.194	0.241	0.220	0.282	0.219	0.261	0.276	0.336	0.245	0.285	0.227	0.287	0.199	0.260
	336	0.246	0.385	-	-	0.254	0.386	0.245	0.382	0.265	0.399	0.280	0.352	0.399	0.500	0.321	0.388	0.282	0.334	0.258	0.311

## Anomaly detection

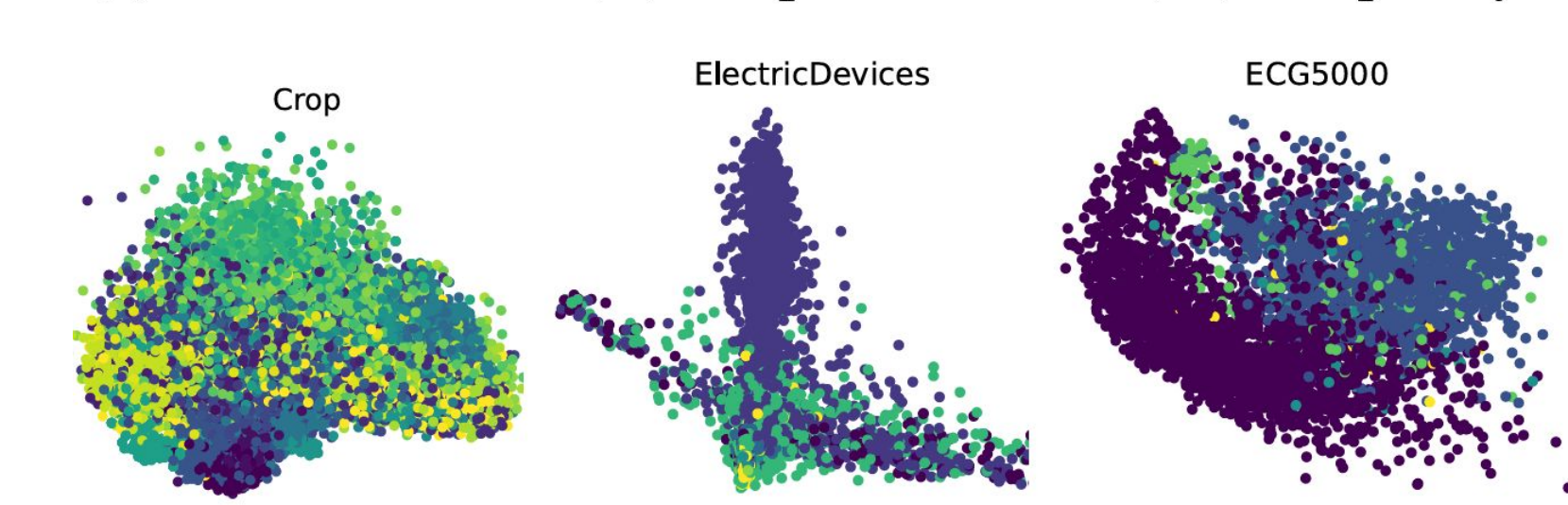
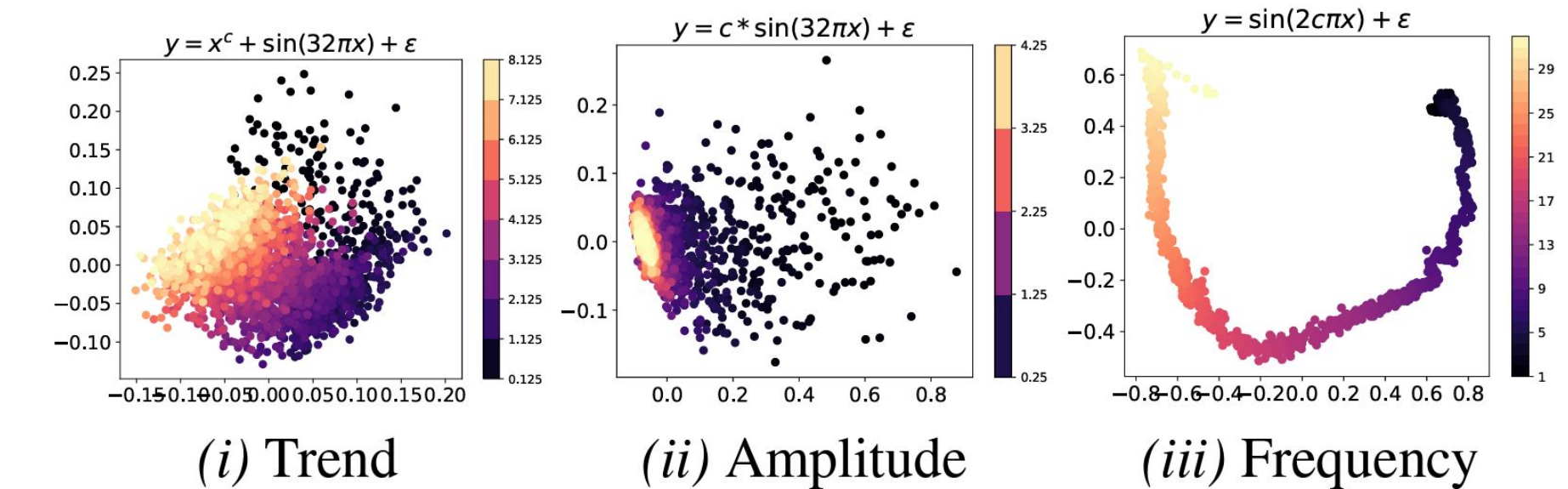
Metric	MOMENT <sub>0</sub>		MOMENT <sub>L</sub>		GPT4TS		TimesNet		Anomaly Transformer		DGHL		k-NN	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Adj. $F_1$	0.585	0.628	0.24	0.537	0.492	0.425	0.554	0.585	0.628	0.24	0.537	0.492	0.425	0.554
	0.683	0.778	0.31	0.541	0.432	0.331	0.595	0.683	0.778	0.31	0.541	0.432	0.331	0.595
	0.377	0.373	0.366	0.389	0.401	0.365	0.393	0.377	0.373	0.366	0.389	0.401	0.365	0.393
	3.410	3.005	4.862	3.642	4.326	5.071	3.681	3.410	3.005	4.862	3.642	4.326	5.071	3.681
	3.00	3.00	5.00	5.00	4.00	5.25	3.75	3.00	3.00	5.00	5.00	4.00	5.25	3.75

Statistical or non-transformer based methods are great!

Zero-shot or linear probed MOMENT often beats models tailored for specific tasks  
Check out paper for complete results ...

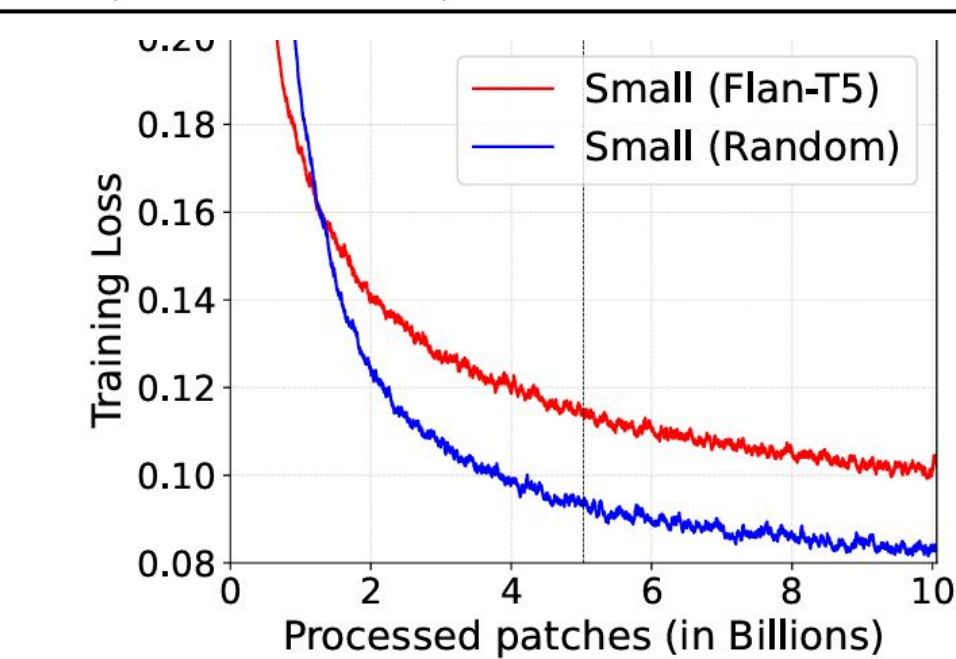


## MOMENT has interesting properties



What is MOMENT learning?

Model	Bit Memory	MNIST	CIFAR-10	IMDb
GPT-2	1,000	0.975	0.711	0.867
Flan-T5	1,000	0.987	0.672	0.861
MOMENT	1,000	0.982	0.620	0.872



MOMENT & Large Language Models

