EasyTS: The Express Lane to Long Time Series Forecasting

Tiancheng Zhang[1], Shaoyuan Huang[1], Cheng Zhang[2*], Xiaofei Wang[1], Wenyu Wang[3]

[1]Tianjin University,Tianjin, China
[2]Tianjin University of Finance and Economics, Tianjin, China
[3]Paiou Cloud Computing (Shanghai) Co., Ltd, Shanghai, China
zhangtiancheng@tju.edu.cn, hsy_23@tju.edu.cn, zhangcheng@tjufe.edu.cn, xiaofeiwang@tju.edu.cn, wayne@pplabs.org

Abstract

Responding to the escalating interest in long-term forecasting within the industry, we introduce EasyTS, a comprehensive toolkit engineered to streamline data collection, analysis, and model creation procedures. EasyTS acts as a unified solution, driving progress in long-term time series forecasting. The platform provides effortless access to various time series datasets, including a newly open-sourced multi-scenario dataset in the electricity domain. Integrated visualization and analysis tools help unveil inherent data features and relationships. EasyTS facilitates a user-friendly model validation approach with versatile evaluation criteria. This toolkit allows researchers to compare their models proficiently against renowned benchmarks. With our ongoing commitment to expanding our dataset collection and enhancing toolkit functionalities, we aspire to contribute significantly to the time series forecasting domain. Code is available at this repository: https://github.com/EdgeBigBang/EasyTS.git.

Introduction

In the epoch of the meteoric ascendancy of the Internet of Things (IoT), Long Time Series Forecasting (LTSF) has emerged as an imperative. It wields the potential to drive the ethos of low-carbon initiatives and frame future developmental trajectories. Emblematic models like Informer (Zhou et al. 2021), Pyraformer (Liu et al. 2021), TimesNet (Wu et al. 2022), DLiner (Zeng et al. 2023), DynEformer (Huang et al. 2023) bear testament to this growing momentum.

As LTSF models proliferate, numerous time series datasets are available but scattered across studies. This requires scholars to expend significant effort in the preliminary work of data collection, including writing specialized loading tools for these datasets. After the collection and preprocessing of datasets are completed, appropriate dataset analysis is also necessary. However, the diversity of time series analysis tools brings a learning curve, and these tasks often appear tedious and redundant. Additionally, after completing the time series data analysis, quickly building and evaluating models, as well as comparing new models with existing ones, also become time-consuming tasks.

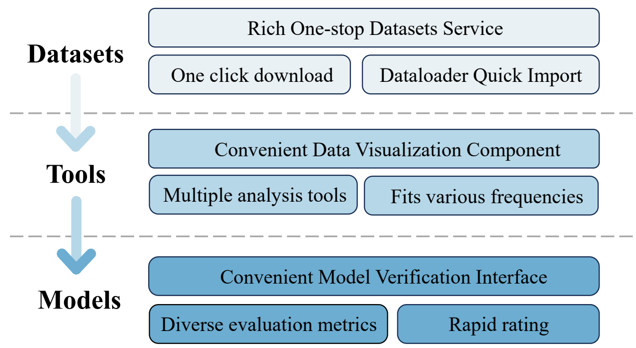

Figure 1: An overview of EasyTS

To simplify this process, this paper presents a tool, EasyTS, aiming to **lower the entry barrier** in the field of long time series forecasting and to accelerate the process of model construction and evaluation. We believe that, similar to AllenNLP (Gardner et al. 2018), EasyTS will help expedite research in long time series forecasting globally.

EasyTS Main Functionalities

An Overview

We introduce EasyTS with the aim to assist researchers in designing and evaluating Long Time Series Forecasting (LTSF) models in a convenient and efficient manner. The overall architecture is illustrated in Figure 1. EasyTS is structured into three progressively advanced levels. Initially, it offers a one-stop solution for datasets, allowing users to easily download and import **richly-scenarioed** time series datasets with a single click through dataloader. Subsequently, the toolkit embeds a variety of preprocessing and convenient **visualization analysis tools** to aid researchers in feature extraction and analysis. Building on this, an intuitive model building and validation interface is implemented for rapid model development and assessment. In this stage, EasyTS provides **diverse evaluation metrics** and benchmark models to ensure comprehensive model evaluation.

Rich One-stop Datasets Service

EasyTS currently encompasses multiple datasets from six different domains around the world, covering popular aca-

[*]Corresponding author.

Dataset	Timesteps	Fre	Scenario
FOOD	3*26205	5T	Electricity
MANU	26205	5T	Electricity
PHAR	2*26205	5T	Electricity
OFFICE	97803	5T	Electricity
ECL	321*26304	1H	Electricity
ETTm	2*69680	15T&1H	Electricity
Solar	105120	5T	Energy
Wind	29*262970	1H	Energy
APP_Flow	128*901	1H	Cloud
ECW	2*797*720	1H	Cloud
Weather	52696	10T	Weather
ILI	966	7D	Disease
Traffic	860*17544	1H	Transportation
METR-LA	207*34272	5T	Transportation
PEMS-BAY	325*52116	5T	Transportation

Table 1: Information of the datasets included in EasyTS

demic fields such as energy, transportation and cloud, with specific details provided in Table 1. In addition to the already open-sourced datasets, EasyTS introduces **four novel open-source datasets** related to electrical energy: FOOD, MANU, PHAR, and OFFICE (Bold indicates). These datasets are collected from real industrial scenarios and are cleaned for usability. The data service allows two methods to acquire datasets: specifying during model training which supports **automatic downloading** or downloading a specific dataset separately through an interface. Moreover, EasyTS implements dataset loading through Torch DataLoader and has crafted corresponding loading classes for each data scenario. Within the loading classes, parameters such as dataset division ratio and normalization methods can be customized. Also, to facilitate researchers using unreleased datasets, we have pre-configured loading classes amenable to third-party use, allowing the loading of private data and subsequent model training in a fixed data format using the toolkit.

Convenient Data Visualization Component

EasyTS expands a versatile set of analytical tools, as depicted in Figure 2. These tools can be used to analyze datasets with different sampling frequencies, using hourly

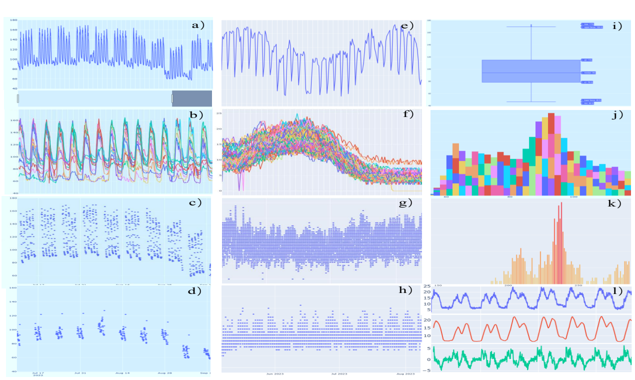

Figure 2: Various analytical tools provided by EasyTS

Model	Dataset	MSE	...	RMSE	CORR	DDTW
DLinear	FOOD	0.225	...	0.317	0.534	1.902

Table 2: Diverse time-series evaluation metrics

(dark) and minute (light) sampling as examples. Embarking on time series analysis necessitates an initial succinct overview of data to comprehend inherent characteristics promptly. EasyTS addresses this with an intuitive and streamlined overview tool, endowed with a slider for effortless perusal of the entire dataset, evidenced in Figure 2(a). Acknowledging the pivotal role of extreme values in numerous tasks, such as discerning maximum wind speed or peak human traffic, EasyTS is equipped with specialized tools, enabling users to extract these values at predefined frequencies, illustrated in Figure 2(e). To discern whether a dataset adheres to a specific distribution within a designated period, EasyTS permits the reconstruction of datasets by day or month, facilitating preliminary exploration of data periodicity as highlighted in Figure 2(b) and (f).

Time series often shows distinct patterns on weekdays and holidays. Additional tools, illustrated in Figure (c) and (d), allow in-depth investigation into these nuanced distinctions. For a more granulated insight, Figure (f) and (h) emphasizes disparities between working and non-working hours. Understanding the approximate value distribution within a dataset is paramount, and EasyTS, as shown in Figure 2(i), offers visual representations like box plots for different dataset quantiles and provides intuitive histogram statistics that amalgamate similar values on a singular axis, as seen in Figure j. To unearth potential cyclical patterns within time series data, tools based on the Fast Fourier Transform (FFT) are integrated into the toolkit, visualized in Figure 2(k),enabling the creation of statistical charts that portray the embedded cycles in time series segments. Additionally, EasyTS facilitates Seasonal-Trend decomposition using Moving Average (MA) (Wu et al. 2021), allowing users to deconstruct data into its integral trend and cyclical components.

Convenient Model Verification Interface

Based on the two aspects mentioned above, scholars can swiftly design their own models within the toolkit's specified modules. EasyTS provides diverse time-series evaluation metrics, such as MSE, MAE, MAPE, MSPE, RSE, RMSE, CORR, DDTW (Keogh and Pazzani 2012), to objectively assess the sophistication of the models, with scoring results in Table 2. Additionally, the toolkit also includes built-in baseline models like DLiner (Zeng et al. 2023) for users to make comparisons and improvements.

Conclusion

The EasyTS toolkit proposed in this paper, spanning datasets, tools, and models, effectively aids researchers, particularly novices, in the design of time series forecasting models. In the future, we will continue to develop this toolkit, encompassing a wider range of datasets, visualization tools, and widely recognized models.

Acknowledgments

This work is supported in part by China NSFC (Youth) through grant No. 62306208 and No. 62002260, in part by the Open Research Fund from Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ) under Grant No. GML-KF-22-03, in part by China NSFC through Grant No. 62072332, in part by the Tianjin Xinchuang Haihe Lab of ITAI under Grant No. 22HHXCJC00002.

References

Gardner, M.; Grus, J.; Neumann, M.; Tafjord, O.; Dasigi, P.; Liu, N. F.; Peters, M.; Schmitz, M.; and Zettlemoyer, L. 2018. AllenNLP: A Deep Semantic Natural Language Processing Platform. *ACL 2018*, 1.

Huang, S.; Wang, Z.; Zhang, H.; Wang, X.; Zhang, C.; and Wang, W. 2023. One for All: Unified Workload Prediction for Dynamic Multi-Tenant Edge Cloud Platforms. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 788–797.

Keogh, E. J.; and Pazzani, M. J. 2012. Derivative Dynamic Time Warping, 2001. *Irvine, California*.

Liu, S.; Yu, H.; Liao, C.; Li, J.; Lin, W.; Liu, A. X.; and Dustdar, S. 2021. Pyraformer: Low-complexity pyramidal attention for long-range time series modeling and forecasting. In *International conference on learning representations*.

Wu, H.; Hu, T.; Liu, Y.; Zhou, H.; Wang, J.; and Long, M. 2022. TimesNet: Temporal 2D-Variation Modeling for General Time Series Analysis. In *The Eleventh International Conference on Learning Representations*.

Wu, H.; Xu, J.; Wang, J.; and Long, M. 2021. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in Neural Information Processing Systems*, 34: 22419–22430.

Zeng, A.; Chen, M.; Zhang, L.; and Xu, Q. 2023. Are transformers effective for time series forecasting? In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, 11121–11128.

Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; and Zhang, W. 2021. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, 11106–11115.