Zero-shot Time Series Forecasting in Financial Data Analysis: Prospects and Challenges

Kostromina Alina
Sber AI Lab, MLTools
October 2024

СБЕР

Data Characteristics and Problem Statements

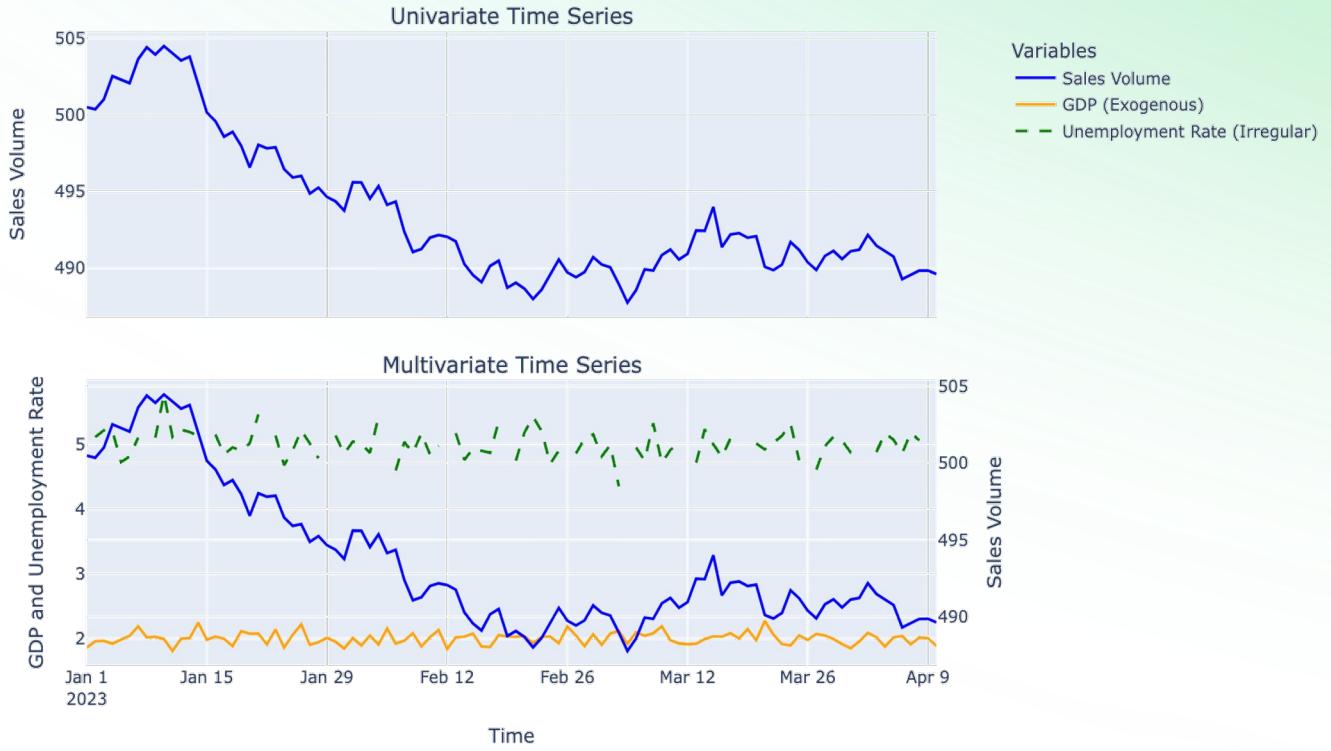
Forecasting instruments

Zero-shot models

Time Series Forecasting

Let’s define the terminology

- **A time series** is a sequence of values ordered by time.
- **A multivariate time series** is a structure where multiple individual time series are considered simultaneously**.**
- **A regular time series** is a time series with evenly spaced time intervals between data points.
- **Additional features (exogenous variables, covariates)** are external factors not generated by the system but influencing the target variables.

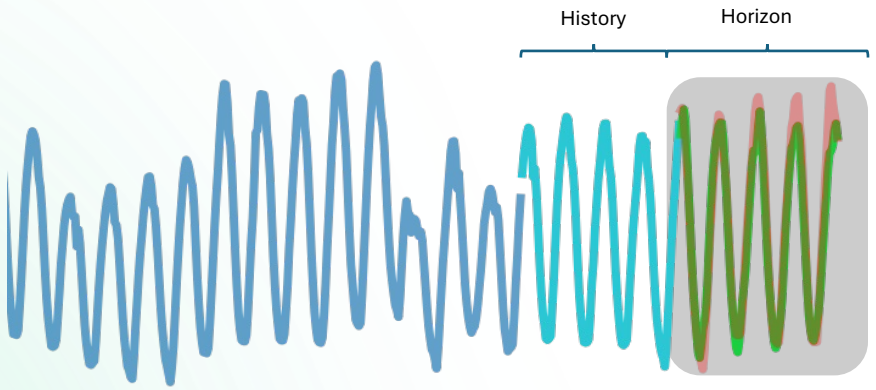


*The top subplot represents the **univariate** TS, while the bottom subplot — the **multivariate** TS. Both `Sales Volume` and `GDP` are **regular**, whereas the `Unemployment Rate` is **irregular**. `GDP` and `Unemployment Rate` are **exogenous features** relative to `Sales Volume`.*

Time Series Forecasting

Let's define the terminology

- **Forecasting Task:** given a history, the goal is to forecast over a horizon.



The history is not all the available data points, but rather the context length currently considered for generating predictions.

- **Forecasting Metrics** (between the prediction and the actual series):
  - Mean Squared Error (MSE)

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$$

  - Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$$

  - Mean Absolute Percentage Error (MAPE)

$$MAPE = 100 \times \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|$$

  - Etc.

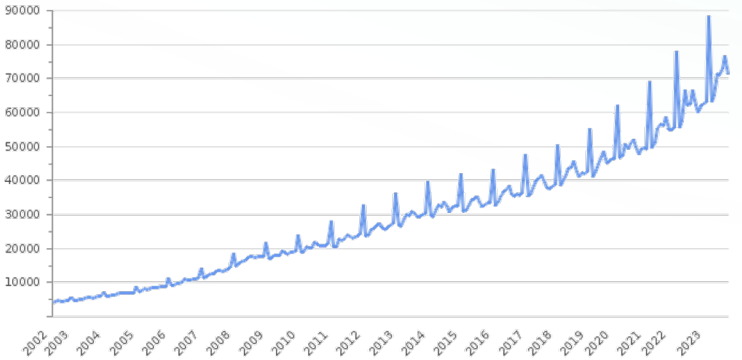
Characteristics of Financial Data and Tasks

which affect how they need to be analyzed

- We are not considering stock market data (**asset prices, market indices**, etc.).
- We are also not considering **irregular** time series.

Interest rate in Russia (%)

- We are focusing on macro- and microeconomic indicators, as well as financial data that reflect the activities of agents (such as **data on loans and deposits, labor market indicators**, and so on).



Average wage in Russia (Rub / Month)

Characteristics of Financial Data and Tasks

which affect how they need to be analyzed

What properties of the data do we face?

1. Short time series
2. Data Instabilities and External Disruptions
3. Numerous exogenous variables and contextual information

What do we expect from the model?

1. The model should work efficiently under data scarcity.
2. The model should be flexible and adapt to changing macro-level distributions.
3. The model should be able to select relevant factors and account for their dynamics.

СБЕР

Data Characteristics and Problem Statements

Forecasting instruments

Zero-shot models

How to generate a forecast?

Groups of methods commonly used in forecasting tasks

Naive methods

- (Seasonal) Naive
- Mean, Median

Statistical methods

- ETS
- Theta
- ARIMA

ML methods

- ElasticNet
- GBMs

DL methods

- DLinear
- NBEATS
- PatchTST
- GPT2

How these methods meet our requirements?

Methods Group	Data scarcity	Flexibility and adaptability	Exogenous variables
Naive methods	+	-	-
Statistical methods	+	-	+-
ML methods	+-	+-	+
DL methods	-	+	+

СБЕР

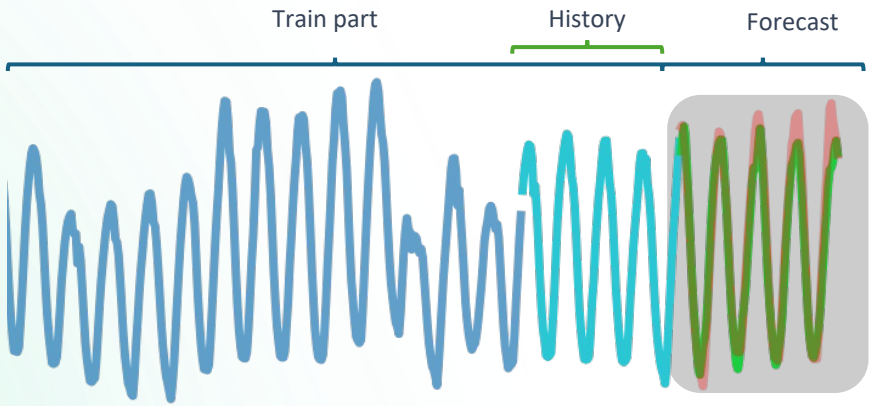
Data Characteristics and Problem Statements

Forecasting instruments

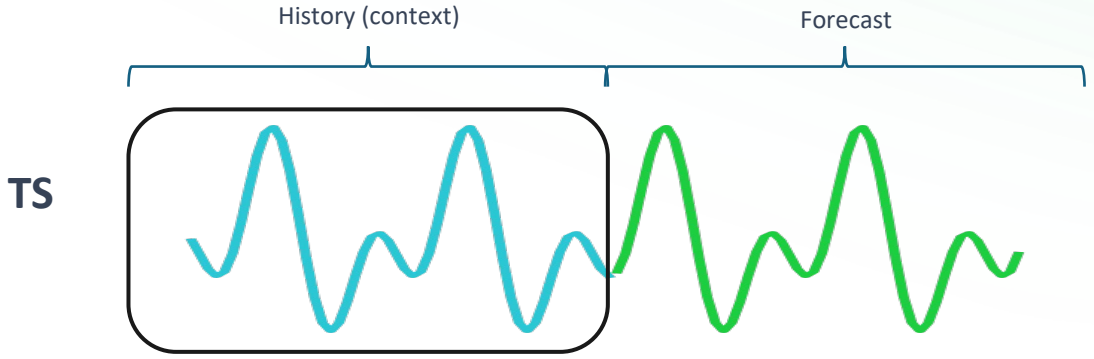
Zero-shot models

Supervised vs Zero-shot models

Supervised: Train the model to predict labels for new data based on patterns identified in the **training data**.



Zero-shot: Train the model to predict labels for new data **without training on the target dataset**, based on patterns identified in the unrelated data.



The variety of Zero-shot models

Zero-shot models for time series are an actively developing area

LLM

Non-adapted LLM

- LSTPrompt
- PromptCast
- LLMTime

Adapted LLM

- Time-LLM
- FPT
- Chronos
- UniTS
- DAM

Specialized

Using synthetic data

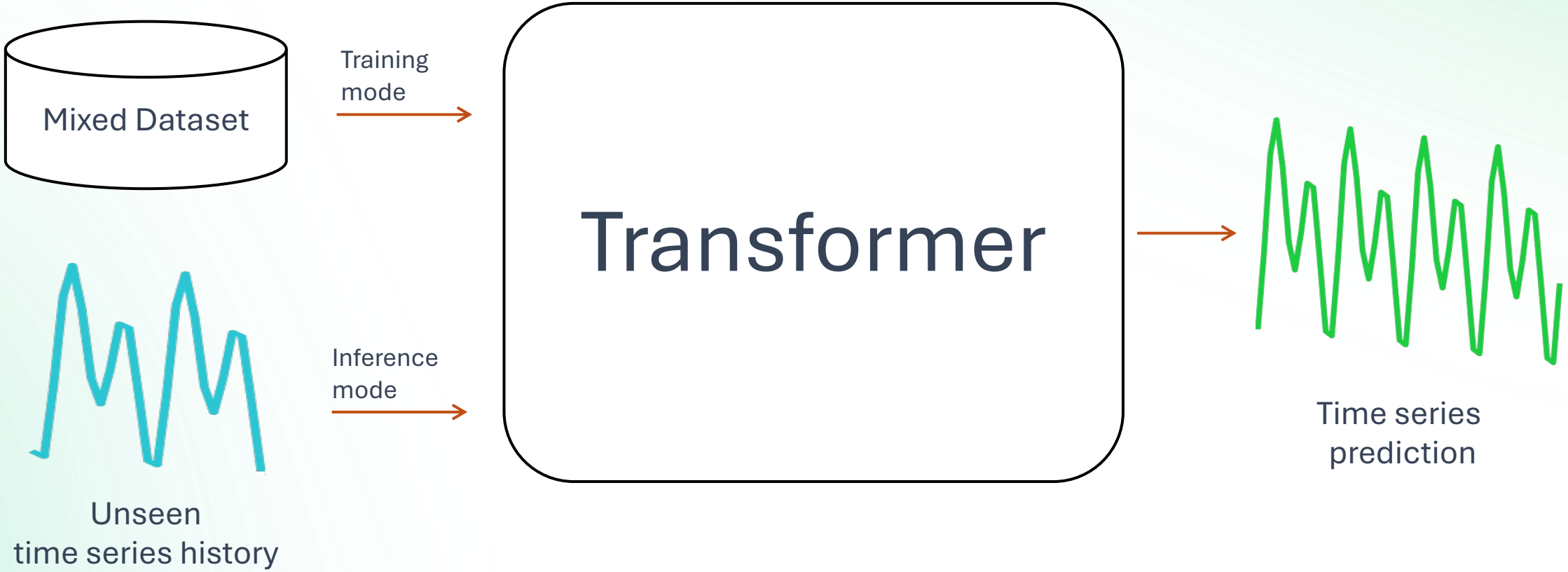
- ForecastPFN

Using real data

- GPHT
- MOIRAI
- Moment

Zero-shot models

- Zero-shot models for time series are mainly Transformers
- They require a large and diverse dataset for pretraining



Kuvshinova, K., Tsymboi, O., Kostromina, A., Simakov, D., & Kovtun, E. (2024). Towards Foundation Time Series Model: To Synthesize Or Not To Synthesize?. *arXiv preprint arXiv:2403.02534*.

Zero-shot for business case

We can adapt the training data properties to a specific task.

- In this case, we are forecasting **macroeconomic** time series
- Dozens to hundreds of **short** time series, which are mostly **independent**.

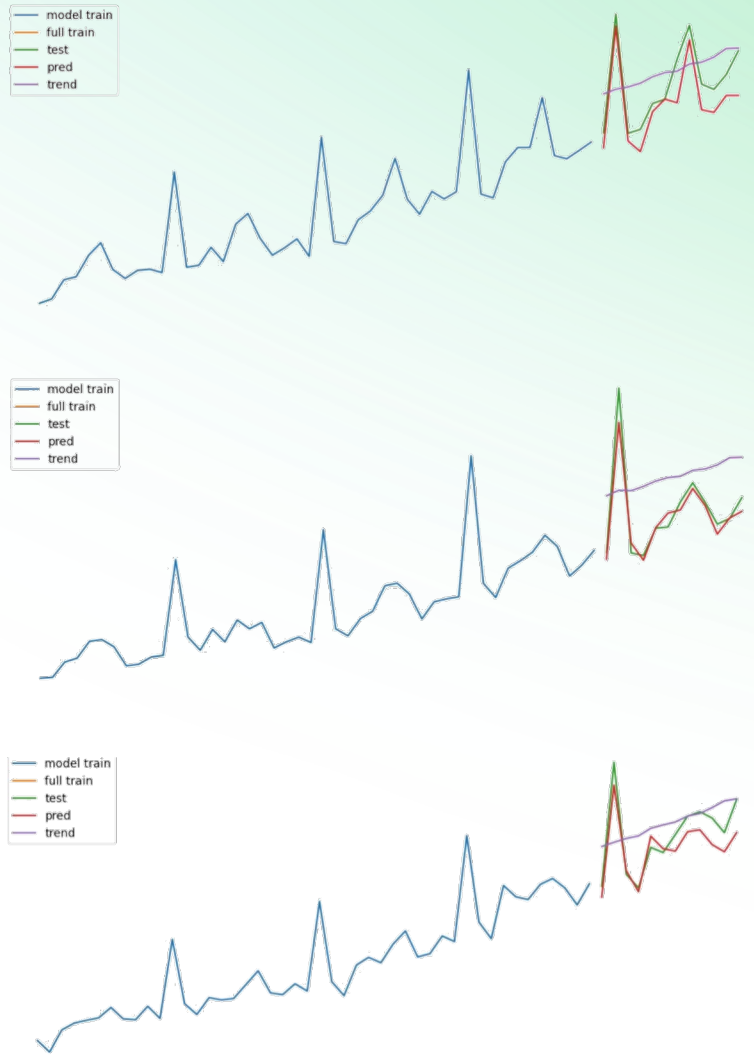
MAPE:

- Zero-shot – x%
- GBDT – **↑0.7%**
- Prophet x~100 – **↓0.2%**

Time (train + inference):

- Zero-shot – **< 1sec**
- GBDT – **30 sec**
- Prophet x~100 – **5 min**

Great quality with a speed advantage!



Predictions of a Zero-shot model trained not only to forecast the continuation of the time series but also to decompose it into trend and seasonality.

Zero-shot with exogenous data

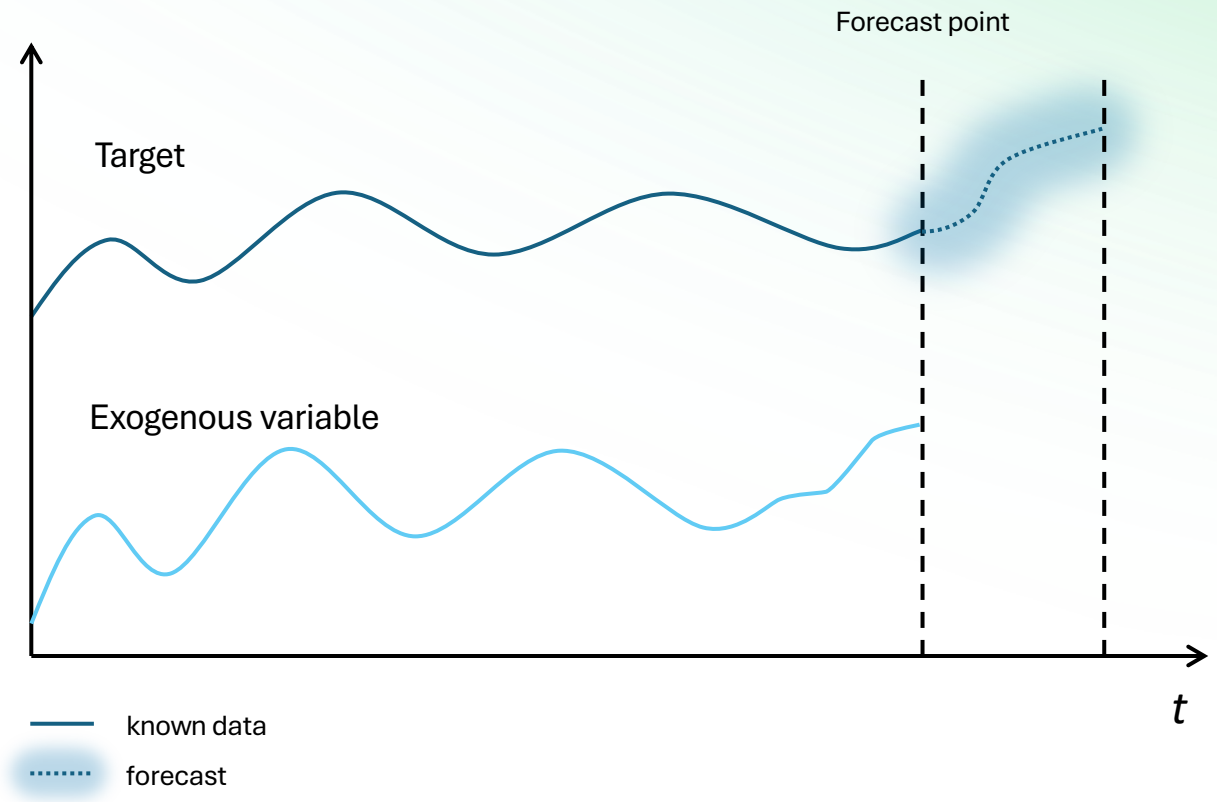
The functionality of Zero-shot model can be extended to handle additional information.

Zero-shot with additional features is an unexplored scientific topic.

Possible solution:

- There are additional time series. The task is to transform these series into a forecast.
- Use TabPFN (Tabular Prior-Data Fitted Network).

- There is no open-source TabPFN for regression tasks yet, but there is a paper from the NeurIPS 2024 Workshop*.



* Hoo, S. B., Müller, S., Salinas, D., & Hutter, F. The Tabular Foundation Model TabPFN Outperforms Specialized Time Series Forecasting Models Based on Simple Features. In *NeurIPS 2024 Third Table Representation Learning Workshop*.

How does Zero-shot fit the requirements from data?

Let’s return to our table

Methods Group	Data scarcity	Flexibility and adaptability	Exogenous variables
Naive methods	+	-	-
Statistical methods	+	-	+-
ML methods	+-	+-	+
DL methods	-	+	+
Zero-shot	+	+	+

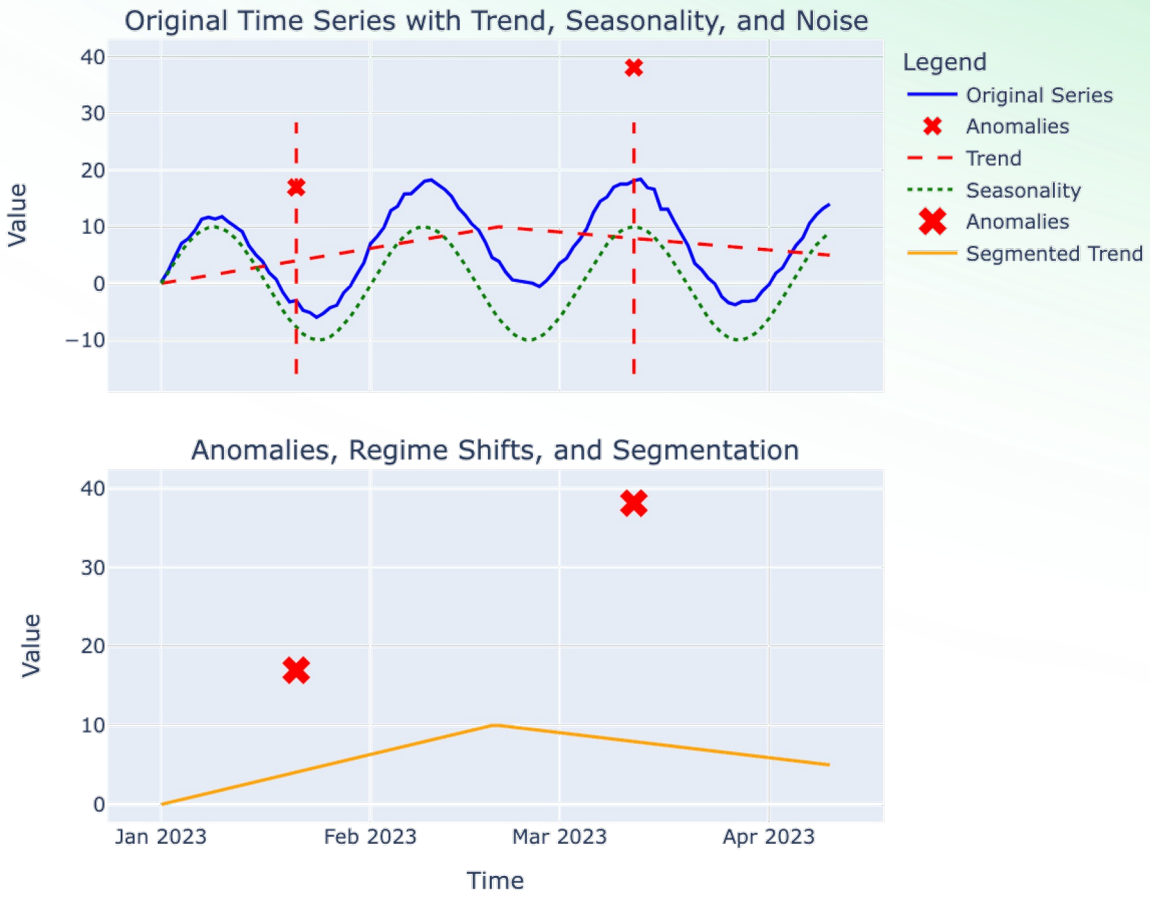
Possible directions for the development and challenges to overcome

1. Development of Multitasking:

- **Decomposition** into trend, seasonality, and noise, with tasks such as:
  - removal of trend, seasonality, or noise
  - **segmentation** based on trend
  - predicting the **trend type**
  - forecasting Fourier coefficients; **seasonality** period and **type**
- Detection of **anomalies** and **regime shifts**
- Generative tasks such as predicting stochastic differential equation **(SDE) parameters, interval forecasting, and filling in missing values**

2. Challenges:

- Integrating forecasting and regression tasks in one model
- Developing effective methods for generating synthetic data
- Finding efficient approaches to utilize exogenous features



Representation of different tasks: decomposition into trend, seasonality, and noise, along with detection of anomalies and regime shifts.

Thank you for your attention!

Kostromina Alina (tg: @elineii)

Sber AI Lab, MLTools