Time Series Analysis: an IoT-oriented overview

Part of the **Internet of Things** Course
Computer Engineering & AI Engineering Master's Degrees

Giacomo Salici
giacomo.salici@unimore.it

UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

Agenda

- ➔ Introduction on Time Series Analysis for IoT
- ➔ Time Series Forecasting
- ➔ Anomaly Detection on Time Series
- ➔ DarTS Library

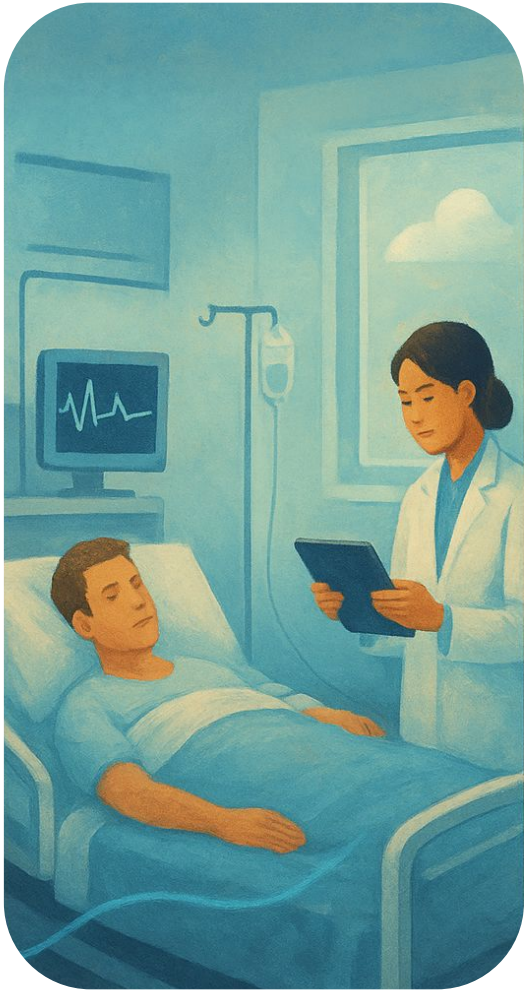
Why?

Because
in each **real-world**
scenario you can
think of, we have
time series.

Why in IoT?

IoT systems:

- ➔ Continuous sensor streams
- ➔ High volume, velocity, variety
- ➔ Unique challenges due to real time and embedded systems



Why in IoT?

Common IoT systems:

- ➔ Environmental sensors
- ➔ Smart meters
- ➔ Wearables
- ➔ Industrial machines

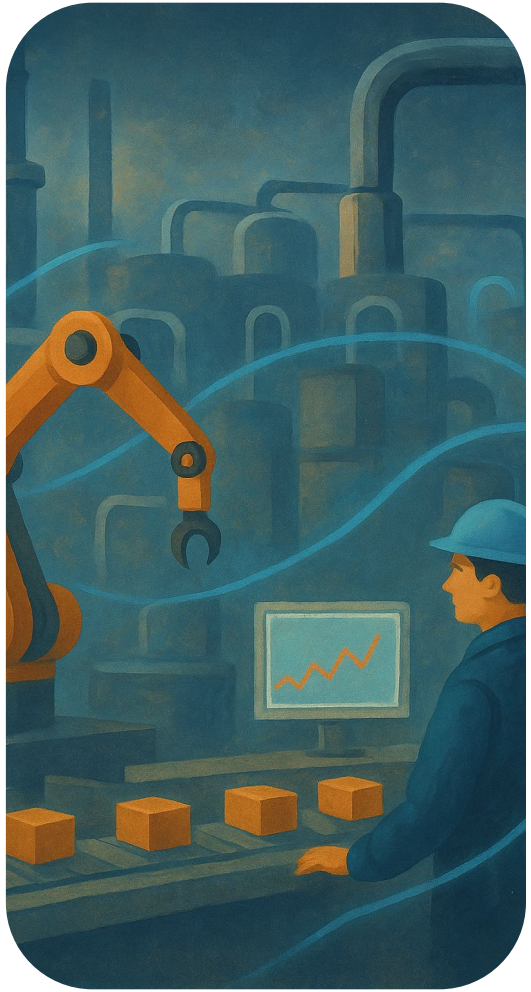
Why in IoT?
Industrial Examples

Predictive Maintenance:

➔ Detect equipment degradation
➔ Plan maintenance early

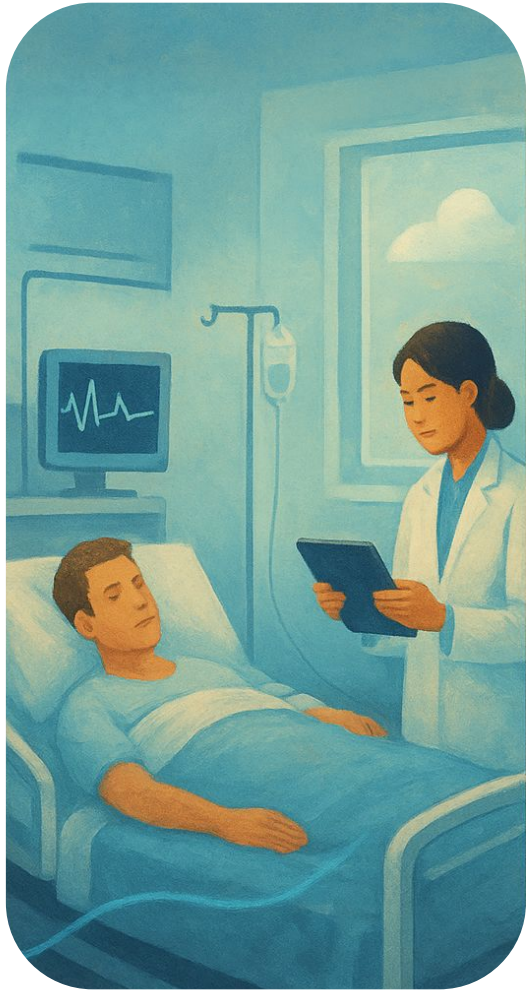
Sensor Fault Detection:

➔ Identify drift or failure
➔ Real-time alerts



Challenges

- ➔ Heterogeneous data
  - from industry to healthcare
- ➔ Missing / noisy values
- ➔ Real-time detection



What is a time series?

A time series is a **sequence** of points sampled with a certain frequency and indexed in **time order**.

$$y = \{y_1, y_2, \dots, y_t, \dots, y_T\} \quad , \quad y_t \in \mathbb{R}^N$$

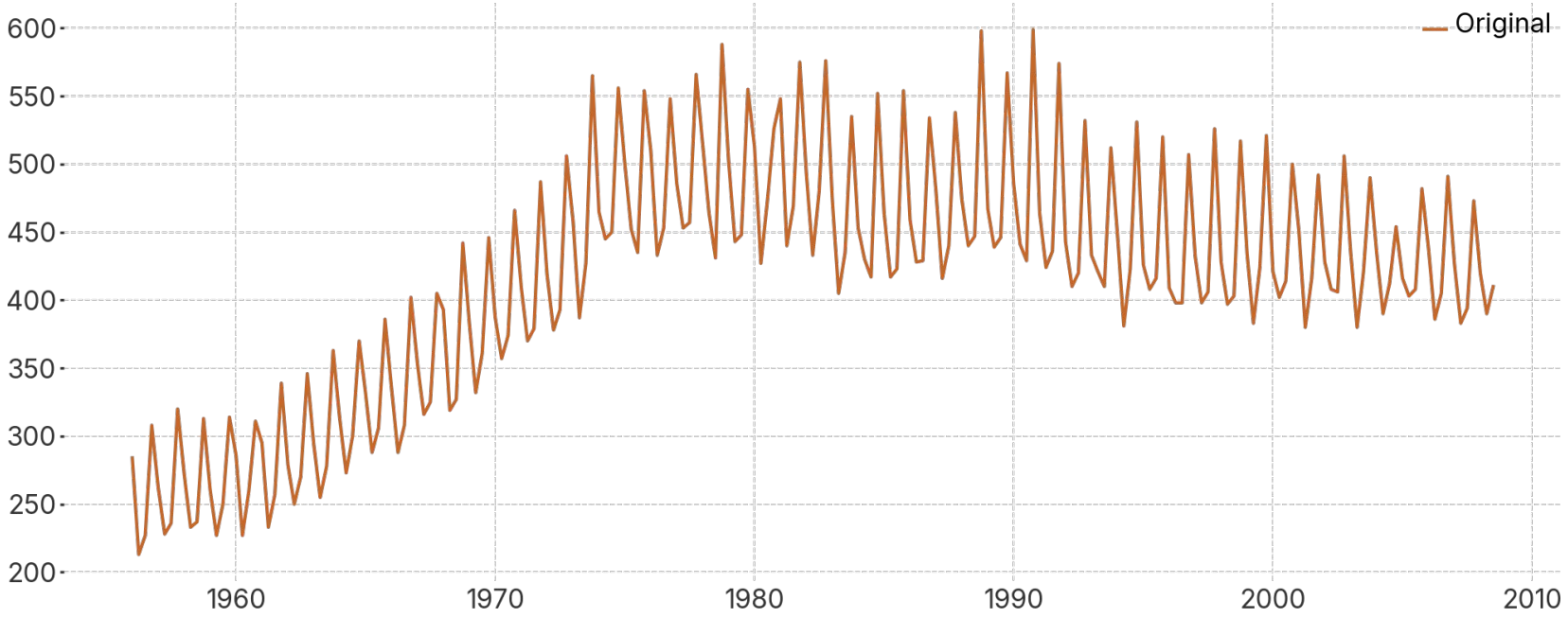
What is a time series?

We call it **univariate** if $n = 1$, **multivariate** if $n > 1$.

$$y_t = [y_1, y_2 \dots y_N]_t \in \mathbb{R}^N$$

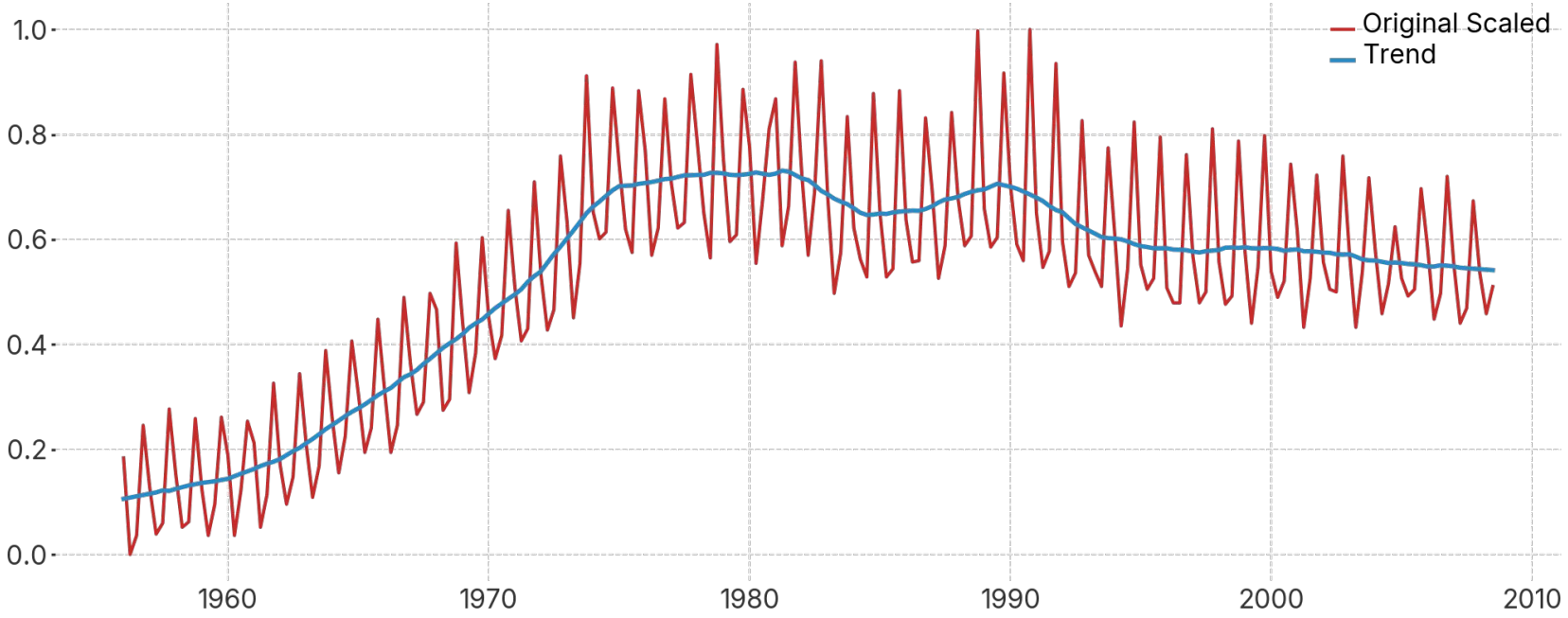
What is a time series?

An example about 🍺 production in 🇦🇺.



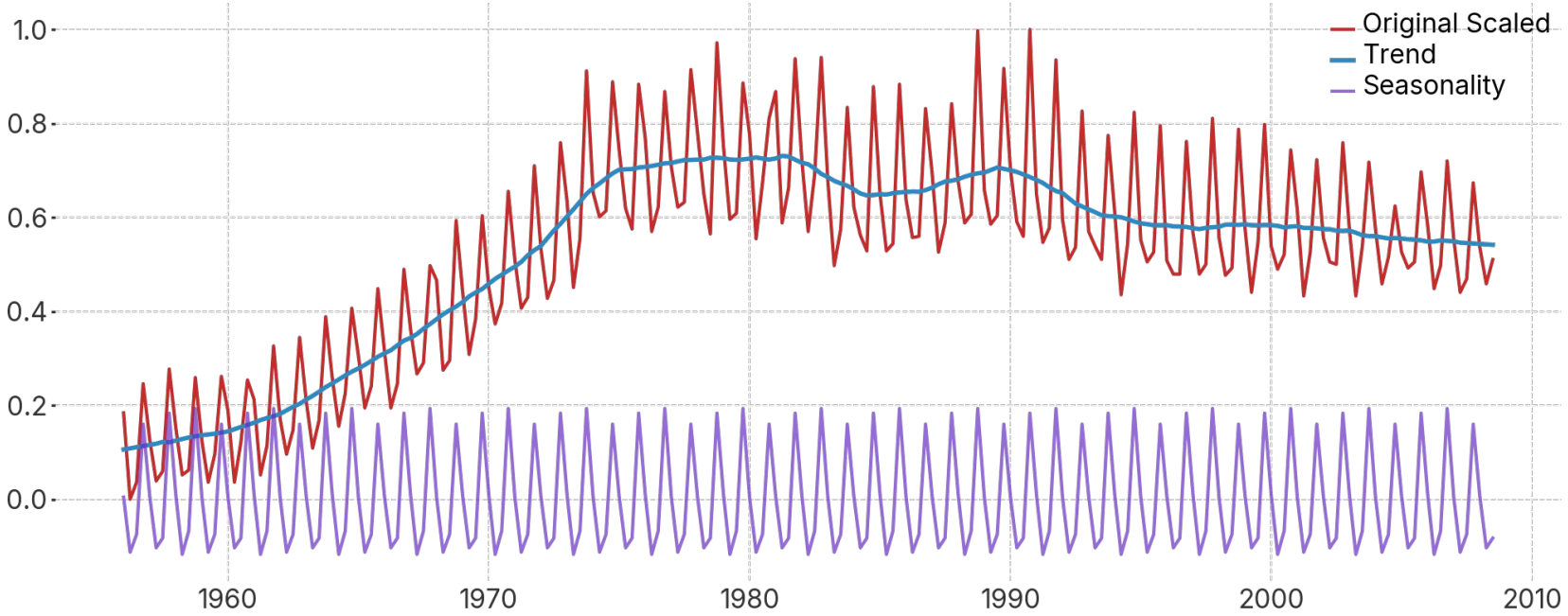
What is a time series?

An example about 🍺 production in 🇦🇺. With a **trend**.



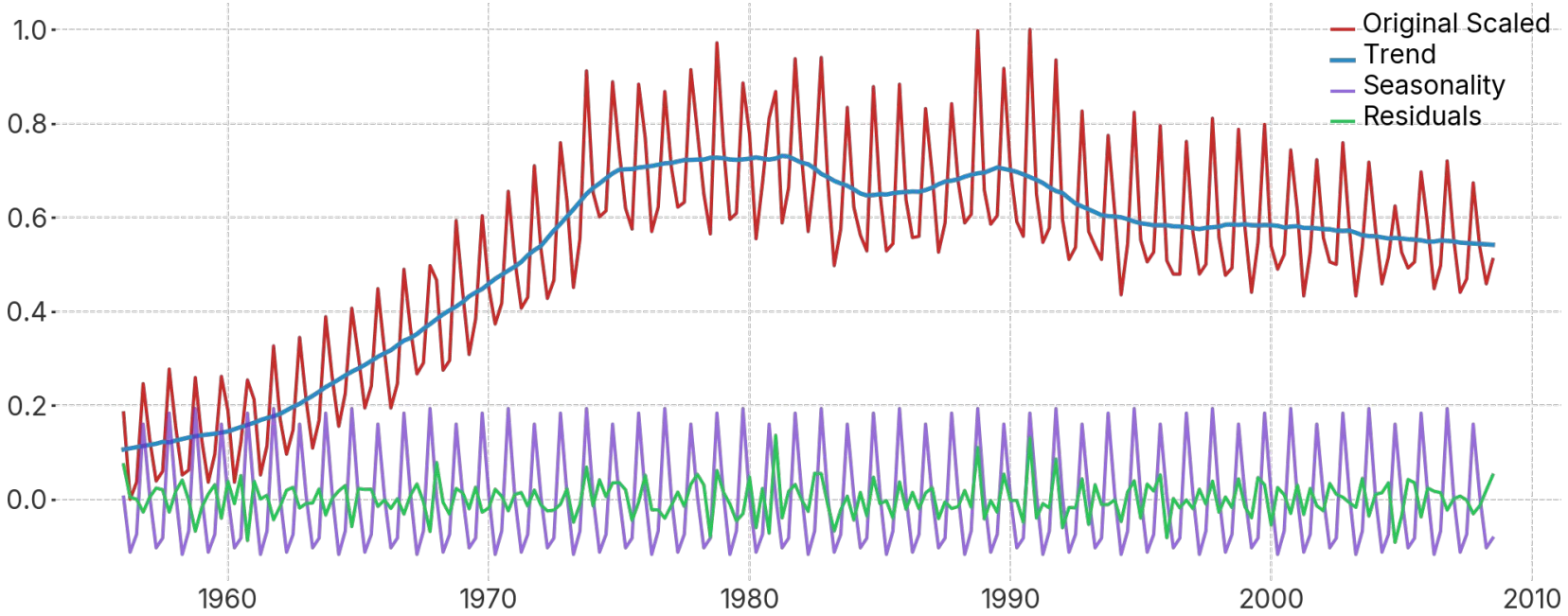
What is a time series?

An example about 🍺 production in 🇦🇺. With a **trend** and some **seasonality**.



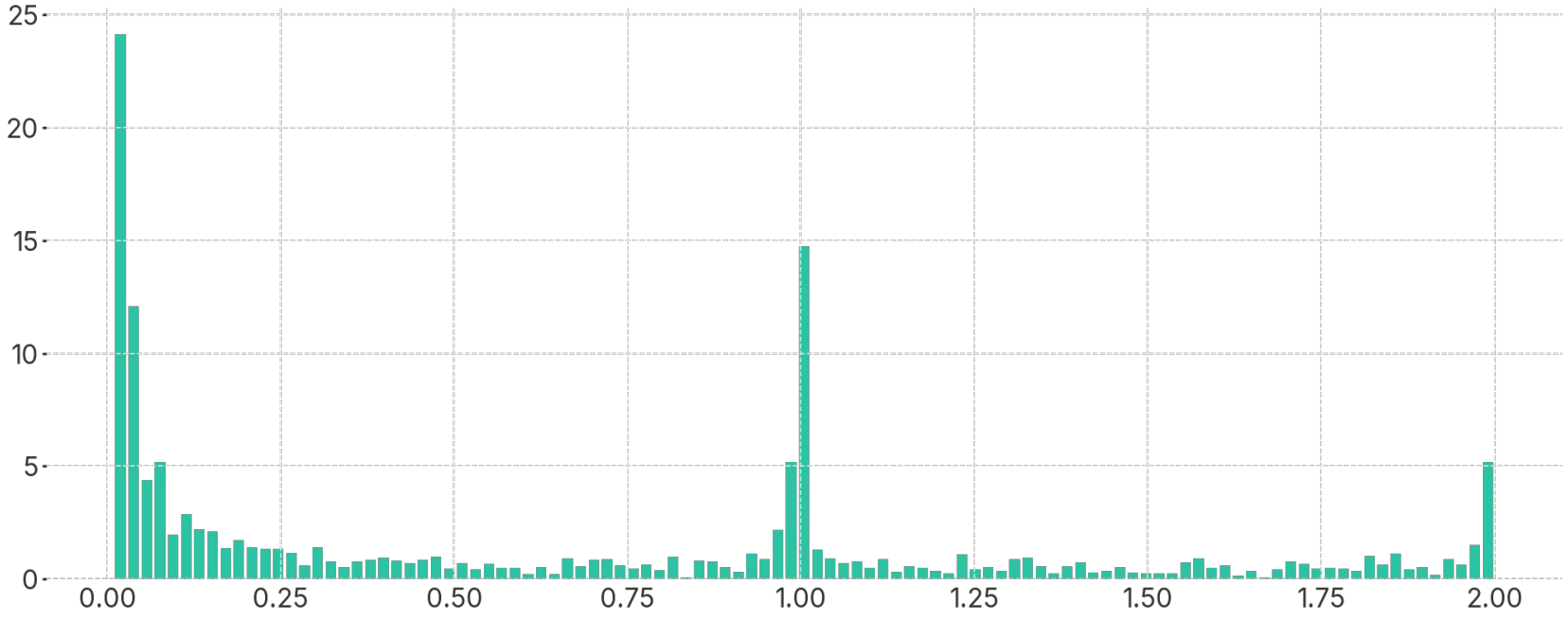
What is a time series?

An example about 🍺 production in 🇦🇺. With a **trend** and some **seasonality** and **residuals**.



What is a time series?

An example about 🍺 production in 🇦🇺. In the frequency domain.



What can we do with time series?

4 main tasks:

➔ classification
➔ imputation
➔ forecasting*
➔ anomaly detection*

*our focus for today

Part 1

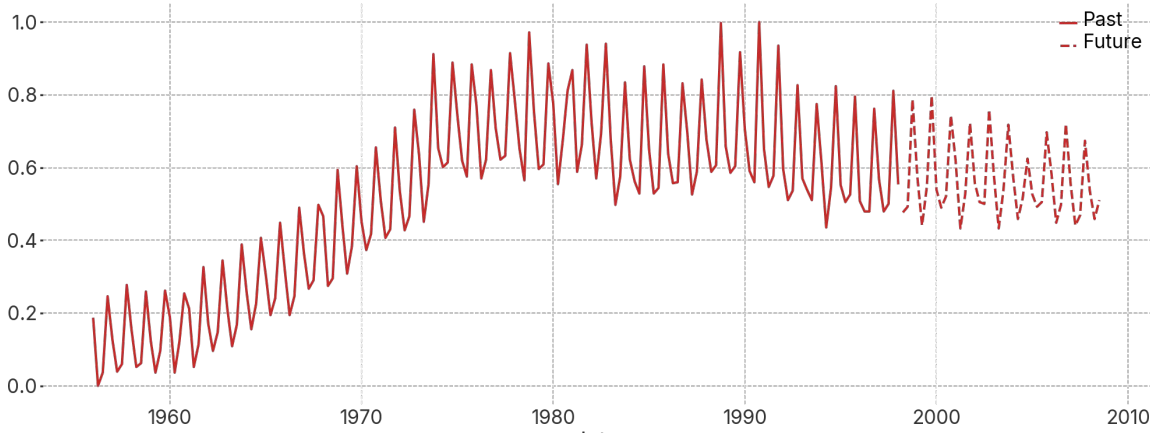
TS Forecasting

TS forecasting

Basic idea:

➔ analyze the past
➔ predict the future

Let’s do some tests.



TS forecasting

It’s a trade-off 🤝

TS forecasting

It’s a trade-off 🤝

TS forecasting:
Naives Methods

- ➔ Mean / Last Value
- ➔ Naive Shift

Not very useful 🫠

TS forecasting:
Naives Methods

➔ Moving Average
➔ Exponential Smoothing

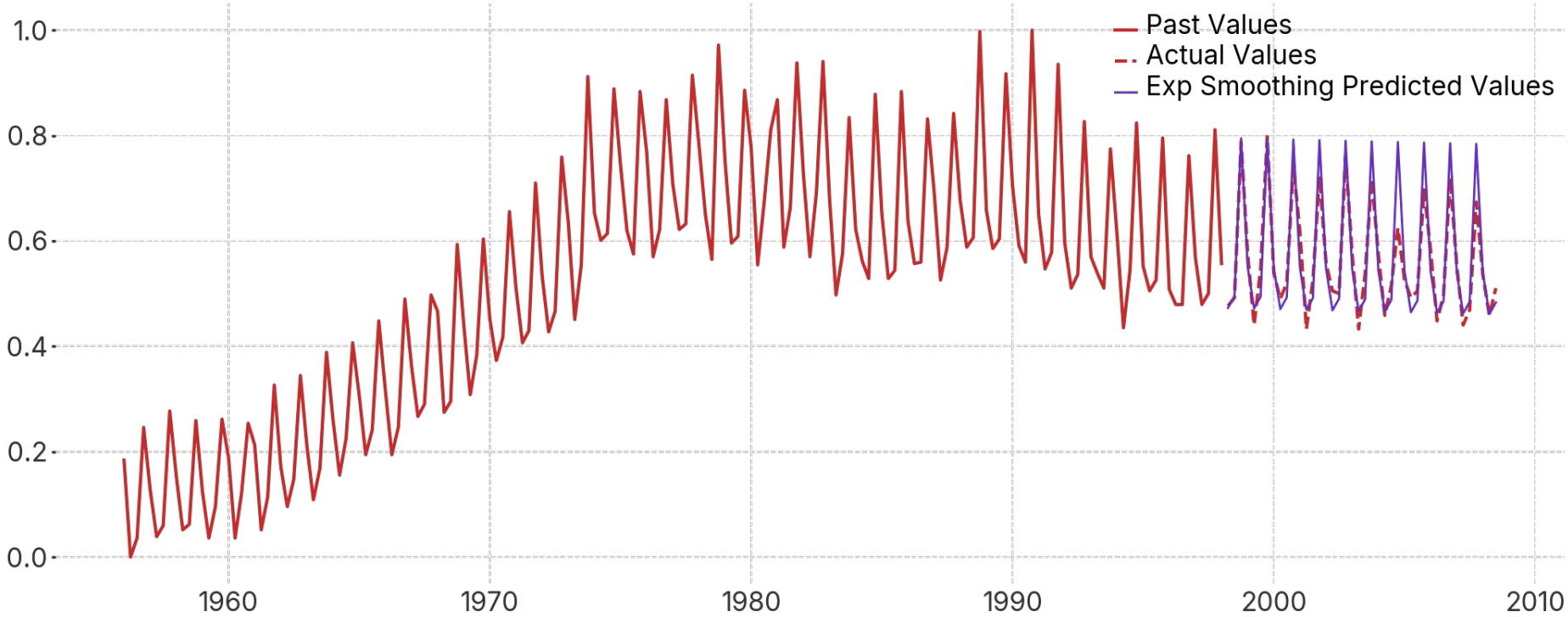
Very simple but sometimes very effective!

Exponential smoothing 📌
also called Exponential Moving Average

Weighted average of past observations with exponentially decreasing weights

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$$

TS forecasting:
Naives Methods



TS forecasting:
Statistical Based Methods

- ➔ ARMA
- ➔ ARIMA
- ➔ SARIMA
- ➔ VARIMA

ARIMA 📌
***A**uto**R**egressive **I**ntegrated*
***M**oving **A**verage*

Statistical model that combines autoregression, differencing, and moving averages to model temporal dependencies.

🔬 ARIMA

p = order of the autoregressive part (to use past steps)

d = degree of first differencing involved (to remove trends)

q = order of the moving average part (models errors)

B = Lag operator

$$\underbrace{(1-\phi_1 B-\cdots-\phi_p B^p)}\quad\underbrace{(1-B)^d y_t}\quad=\quad\underbrace{c+(1+\theta_1 B+\cdots+\theta_q B^q)\varepsilon_t}$$

We will not enter in the math rabbit hole 🐇
But for whose interested: [FPP book - Chap. 9]

TS forecasting:
Statistical Based Methods

- ➔ Prophet by Meta
  - Designed to handle seasonality, holidays, and trend changes effectively.
  - Combines decomposition and non-linearity

TS forecasting:
ML based methods

Especially useful with multivariate TS

➔ Regression
- SVR
- Random Forest

TS forecasting:
DL based methods

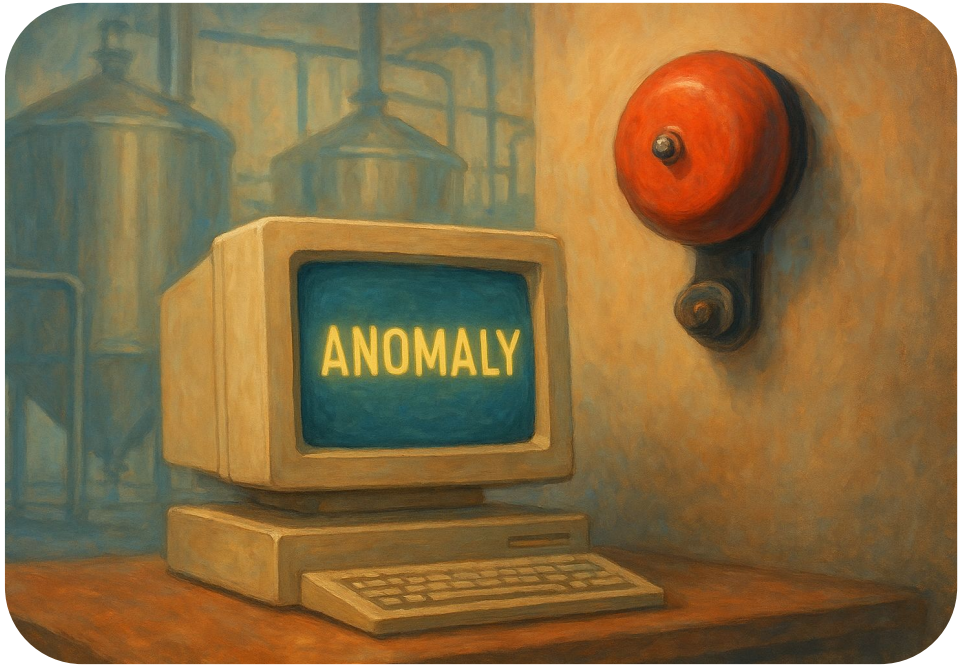
Often tailored on the specific dataset

Generally require more data and more compute power

➔ RNN/LSTM
➔ Transformers

Part 2

TS Anomaly Detection



TS anomaly detection

Naives Methods

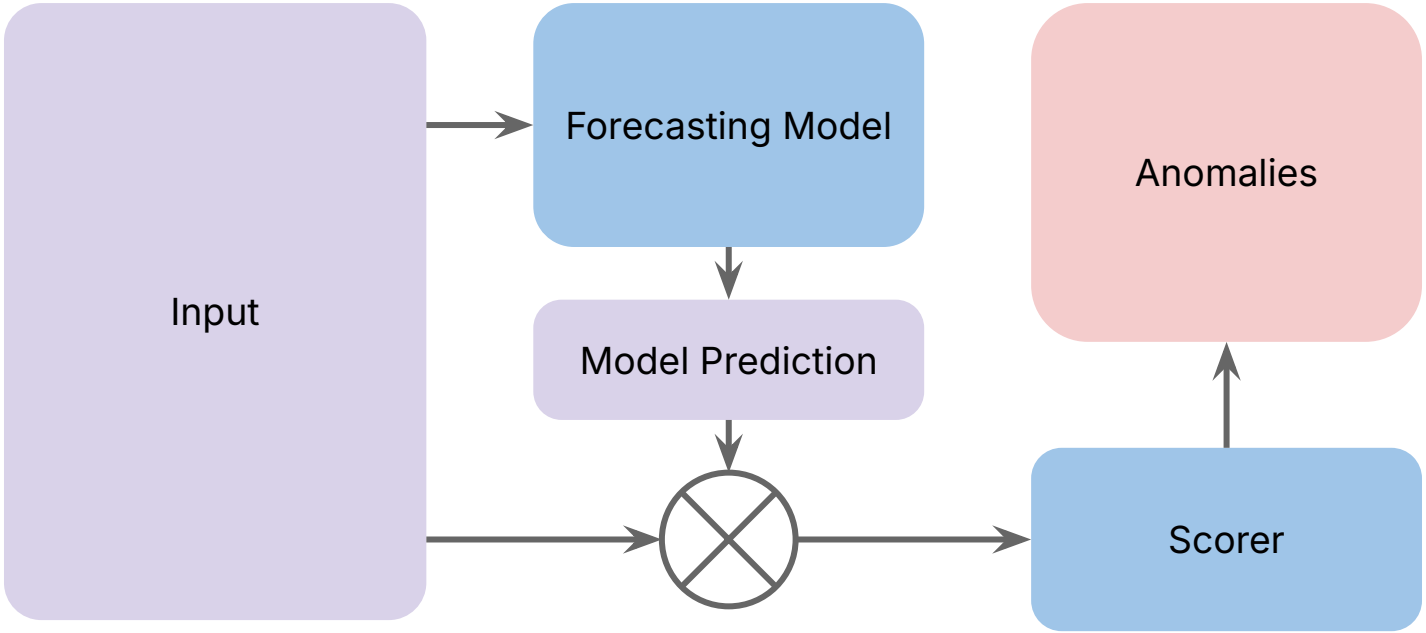
- ➔ Threshold-based
- ➔ Z-score / Deviation
- ➔ Rolling window check

Z-score 📌

$$Z = \frac{x - \mu}{\sigma}$$

TS anomaly detection

A less basic idea



TS anomaly detection
A less basic idea (main problem)

Problem: (often) we don’t have label*
about anomalies not even in the training set!

➔ We can estimate the outliers
➔ We can use unsupervised learning
➔ Anomalies depends on our environment

*a label is an information about the status (anomaly / normal) of the current time step

TS anomaly detection

A less basic idea (warnings)

- ➔ Anomalies are rare!
- ➔ Training set must be “normal” (= without anomalies)
- ➔ ⚠ Be aware of data drift

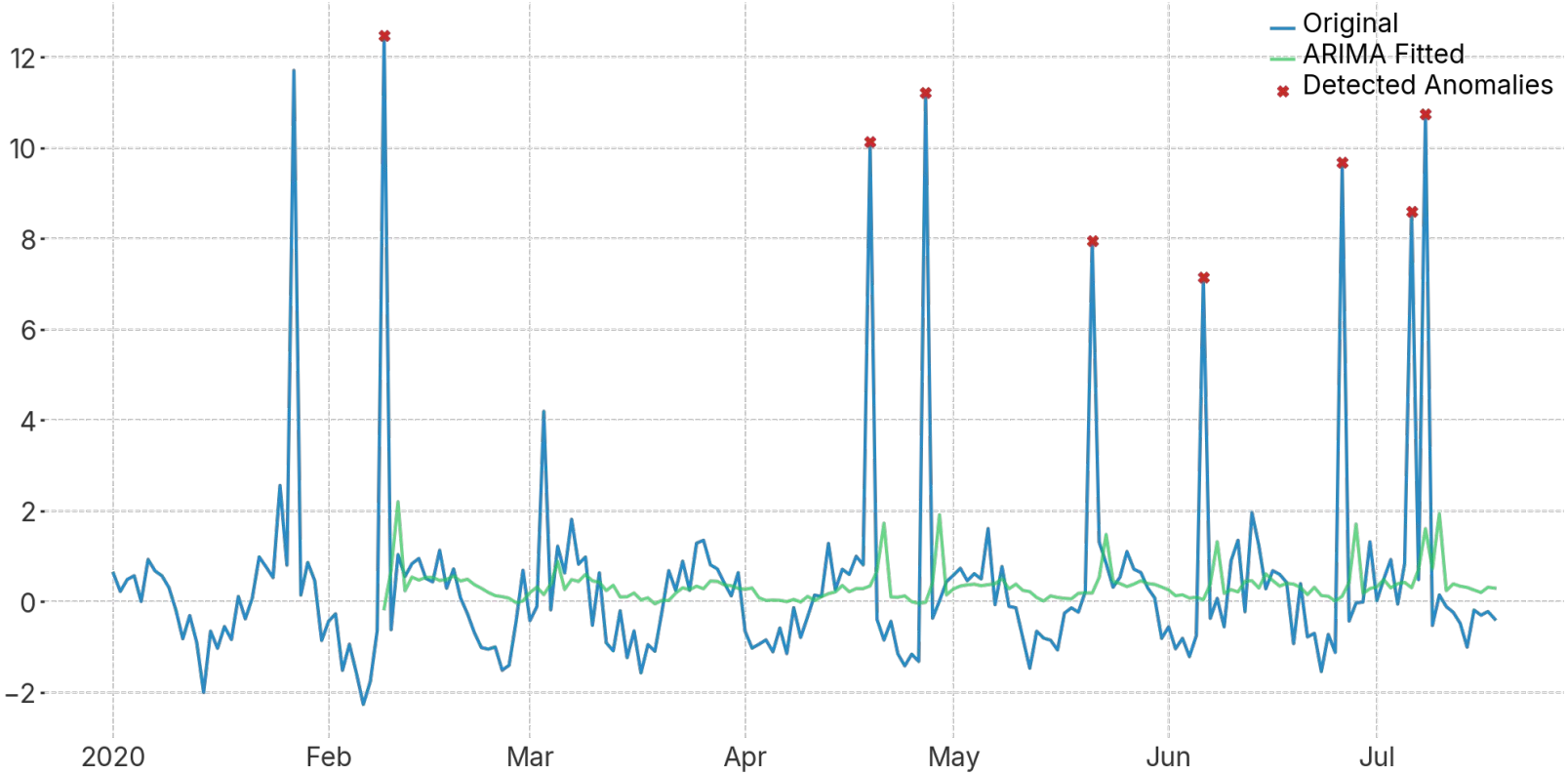
TS anomaly detection

Statistical Based Methods

- ➔ ARIMA residuals
- ➔ STL decomposition

TS anomaly detection

Statistical Based Methods



TS anomaly detection

ML/DL methods

➔ Isolation Forest / One-Class SVM

➔ Autoencoders

... and much more 😬

Part 3

A coding overview

Many Libraries...

Several libraries available for Python:

- ➔ Similar structure
- ➔ Similar dependencies
  - Matplotlib
  - Numpy
  - Pandas
  - SciKit Learn
- ➔ Perfect for quick tests
  - Not suitable for designing new SOTA models
- ➔ Wrappers of famous methods

Darts

SKTIME

NIXTLA

PROPHET

...similar codes

The core idea is the same for all the libraries.

Methods API borrowed from SKLearn.

```py
model.fit(train_data)
pred = model.predict(len(test_data))
error = mape(test_data, pred)
```

Hands-on moment!

The code (and also the plots) of today lecture has been made with Darts.

For the one interested:

👉 *github.com/jacksalici/tsa-lectures*

Hands-on moment!

In Darts, each model has several utility methods.

```py
model.fit(train_data) #train the model on a target series
model.predict(len(test_data)) #forecast future values
model.gridsearch(parameters, series, ...) #hyperparameter tuning
model.historical_forecasts(...) #simulate past predictions to evaluate
model.backtest(...) #compute error over historical forecasts
```

Darts

Hands-on moment!

In Darts, we have several models ready to use:

- ➔ **Baseline models:** Naive mean, seasonal, drift, moving average
- ➔ **Statistical models:** ARIMA, VARIMA, Exponential Smoothing, Theta, TBATS, etc...
- ➔ **SKLearn-like models:** regressors like RandomForest, LightGBM, Linear Regression
- ➔ **Deep learning** / PyTorch models: N-BEATS, RNN, Transformer, TCN, TFT, TiDE, TSMixer...
- ➔ Ensembles, Conformal models, Foundation models, etc.

👉 Read more here.

Darts

Do you want more 🍺 analysis?

Not only beer production datasets ofc...

Active research topics in the scientific community:

➔ Time Series foundation models
➔ Federated Learning
➔ Tiny ML

AImage Lab
University of Modena and Reggio Emilia

Thank you for your attention!

Feel free to reach me! 👋
giacomo.salici@unimore.it

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA