

Time Series Forecasting: Machine Learning and Deep Learning with R and Python

Marco Zanotti

2021-2022

In the last 15 years, business requests related to time series forecasting changed dramatically. Business needs evolved from predicting at most 100, low frequency (mainly monthly or quarterly) data, to forecasting 10.000, high frequency (mainly daily data) time series. Unfortunately, the classical tools that were highly used in the past years may not be the best anymore, both in terms of accuracy and computationally. Moreover, the review of M-Series Forecasting Competitions suggests that newer ML and DL models, or stacking ensemble techniques may be very accurate compared to easier solutions.

For these reasons, nowadays the “time series forecasting” data scientist is required to be capable of providing business forecasting solutions tackling both scalability and accuracy, constantly keeping up-to-date with new methods.

Course Structure

The aim of the course is to teach the students how time series forecasting problems can be solved in practice. The state-of-the-art techniques are presented from a very practical point of view, throughout R tutorials on each main topic. Python algorithms are also presented and used within R by means of the *reticulate* package. Theoretical concepts are left to those who are interested in and bibliographic references are listed at the end of the course.

- 10-14 hours: practical lectures on the main contents
- 1 hour: virtual meeting with Matt Dancho
- 4-8 hours: group projects
- 2 hours: results' presentations and discussion

Contents

- Time Series Manipulations, Transformations & Visualizations
- Time Series Features Engineering
- Time Series Models
- Machine Learning Models
- Deep Learning Models
- Automatic Machine Learning
- Hyperparameter Tuning
- Ensemble Learning
- Recursivity of ML Models

Evaluation

Students will be divided into 5 groups (3 students each) and will be asked to compete on a Time Series Forecasting challenge. They are expected to provide a fully reproducible project showing the statistical methods used to solve the forecasting problem at hand.

The evaluation will take into consideration:

- the correctness of the methods used (10pt)
- the final average forecasting performance (5pt)
- the total time spent on developing the project (5pt)
- the total computation time (1pt)
- the originality of the approach (5pt)
- the presentation of results (4pt)

Duration & Calendar

The course earns 3 ECTS for a total of 20 hours, divided into 5 lectures (4 hours each).

1. Saturday 2022-01-29, 9.00 - 13.00
2. Saturday 2022-02-05, 9.00 - 13.00
3. Saturday 2022-02-12, 9.00 - 13.00
4. Saturday 2022-02-19, 9.00 - 13.00
5. Saturday 2022-03-05, 9.00 - 13.00

Lectures take place at University of Milan, Via Conservatorio 7.

The 1 hour virtual meeting with Matt Dancho will take place during the week after 18pm.

Eligible Students & Requirements

15 students

Requirements:

- Knowledge of R / Python
- Basic knowledge of RMarkdown / Jupiter Notebook
- GitHub Account
- Knowledge of the theoretical framework related to ML/DL models
- Strong interest in time series domain analysis
- Competitiveness

Application

Motivation letter has to be sent to dse@unimi.it

Application deadline: 2022-01-09

For more information about the course contact zanottimarco17@gmail.com or look at the Course Syllabus