Bi-weekly and Final Assignments

Overview

Bi-weekly assignments: You will be asked to complete 2 bi-weekly short assignments, one at the end of week 2 and the other at the end of week 4. You will be given a set of data and you will be asked to analyze this data set. We estimate that completing these assignments will require roughly 2 hours each.

How to submit your assignments: A dedicated space will be created (one in week 2 and one in week 4) and you will be required to upload your assignments there. We will provide more details on how to do this during week 2.

Bi-weekly assignment grading: Each assignment will be reviewed by a member of our Academic team. A Pass/Fail indication will be given, along with feedback remarks.

Final assignment: The final assignment will be to analyze a set of data. Ideally, the set of data should come from your own business. Alternatively, we could provide a set of data for anyone needing one. You will be asked to submit your final assignment at the end of week 6. We estimate that completing this assignments will require roughly 6 hours, depending on the dataset used.

Final assignment grading: Each final assignment will be reviewed by a member of our Academic team. A Pass/Fail indication will be given along with feedback remarks.

Eligibility for a certificate of accomplishment: In order to qualify to a Certificate of Accomplishment, you must submit the 2 bi-weekly and the final assignment, demonstrating engagement to the course. The certificate of accomplishment will be digitally verifiable via the Bitcoin Blockchain network.

Questions: If you still have any questions or problems please post it in the forum, or alternatively contact Demetrios Tseas at tseas.d@unic.ac.cy.

Bi-weekly Assignments

For these assignments you will be given two sets of time series that display different characteristics and properties. In the first bi-weekly assignment you will be asked to visualize, process, and analyze the series provided, while in the second to apply the forecasting methods presented in the Course and evaluate their performance.

1. First bi-weekly assignment

Exercise A. The first series of the dataset presents the daily sales of a retail store, expressed in Euros. You are asked:

- To visualize the original series.
- To identify the missing values and fill them using an appropriate method.
- To identify the outliers present and normalize their values.
- To create and visualize the respective weekly and monthly series.
- To decompose the monthly series and visualize its trend and seasonal components.
- To compute the average sales per weekday and month.

Exercise B. The second series of the dataset presents the daily sales of a product sold in the store, expressed in units. You are asked:

- To visualize the series.
- To compute the average daily demand of the product, the coefficient of variation of non-zero demands (CV^2), and the average number of time periods between two successive non-zero demands (ADI).
- To visualize the empirical distribution of the demand of the product and compute its 5%, 50% and 95% percentiles.
- To create and visualize the respective monthly series and comment on its seasonal pattern, if present.

2. Second bi-weekly assignment

Exercise A. The first series of the dataset presents the monthly totals of international airline passengers between 1949 and 1960, measured in thousands. You are asked:

- To seasonally adjust the series and compute its seasonal indexes. To do so, you should use both the additive and multiplicative expression of the decomposition method and select the most appropriate one.
- To generate forecasts for the following year using Simple, Holt and Damped Exponential Smoothing.
- To select the most appropriate forecasting method of the alternatives considered above using an in-sample accuracy measure of your choice.
- To provide the 80% and 95% prediction intervals for the following year using the forecasting method you have selected above.

Exercise B. The second and third series of the dataset present the power generated by a wind turbine at hourly level and the wind speed that has been forecasted for the same periods. You are asked:

- To visualize the power generation series.
- To plot power generation against wind speed and compute the correlation of the two series.
- To forecast the power that will be generated by the turbine in the following day using linear regression.
- To forecast the power that will be generated by the turbine in the following day using a neural network and a decision tree.

Exercise C. After completing exercises A and B, you will have to access another dataset, containing the actual future values of the series you have predicted in exercises A and B. You are asked:

- To measure the forecasting accuracy and bias of your point forecasts (series of exercises A and B).
- To comment on the correlation of in-sample and out-of-sample forecasting accuracy (series of exercise A).
- To measure the precision of the prediction intervals provided (series of exercise A).

Final Assignment

The aim of this assignment is to conduct a thorough analysis on a forecasting problem of your choice and report the steps you would have followed in order to provide an adequate solution based on what you have learned in the Course.

Ideally, the set of data should come from your own business. This will make sure that you have a deep understanding of the examined problem, being also a useful exercise for your company or organization. However, we will be able to provide a set of data for anyone needing one (more details about this dataset will be given at the end of week 4).

The report submitted for the final assignment should include:

- **A description of the problem being solved**. For example, you should state what does the data represent, what is the main objective of the application considered, why its results are important for your business, what is currently done in your company for dealing with this problem, and what are the main parameters that should be taken into consideration for solving it (e.g., input data, constrains, and forecasting horizon).
- **A presentation of the dataset used**. For example, you should present the series you want to predict and the external variables that will be considered (if any), commenting on their correlation. You should also process the data, if needed, and examined the seasonal and trend patterns present at various frequencies.
- **An overview of your "forecasting competition"**. Given that various forecasting methods may be appropriate for dealing with the problem being solved, you are expected to test various alternatives including both statistical and machine learning approaches (at least one of each kind). You should report the accuracy of the methods you have developed and comment on your results, thus justifying your final choice. Emphasis should be given to the metrics and processes used for performing the evaluation as objectively as possible. Note that the methods participating in the competition should be clearly described (inputs, outputs, and parameters considered).