Fachbereich AIHI
Universität Salzburg
Univ.-Prof. Dr. Wolfgang Trutschnig

Salzburg, June 2026
SS 25/26

Exercise Sheet 13 @'Applied AI Using R'

This is (really) the last exercise sheet of the semester. We work with timeseries and again very hands-on.

Exercise 65.
We again work with the ATM dataset loaded in `http://www.trutschnig.net/07_Sprout.R` Use `https:/trutschnig.net/08_Trelawny.R` (in particular the last block) to fit a linear trend, additive seasonality and ARIMA to the ATM dataset and test its performance in the same manner we checked the performance for the RF and the NN. How does this simple model perform in comparison to the other ones?

The following two exercises are a first glance to your final project. Here is s quick description of the data: A company has 1,000 employees with laptops (lap0001 to lap1000), each of which has up to 100 different services (i.e., apps or operating systems) installed (ser001 to ser100). A central database logs errors for each laptop and each installed service — the dataset in the next two exercises contains exactly this information over a one-year period. The column special.day is set to 1 for public holidays and bridge days, and 0 otherwise. If no error is recorded for a service on a laptop, this automatically means that the service is not installed on that laptop.

Exercise 66.
Download `http://www.trutschnig.net/Errors.rds` and import the dataset in R. Produce 5 nice and informative ggplot2 graphics summarizing the data; work, e.g., with heatmaps counting the daily/weekly total number of errors per service[i], work with timeseries, etc.

Exercise 67.
Use the timeseries of the daily total of errors per service from the previous exercise. Forecast the daily total number of errors per services for the first 14 days in 2020 using at least two different models (RF, NN, ARIMA, etc.).

[i] i.e., consider the sum of the errors per service and laptop and aggregate over laptops