

Practica IoT

Group 9

June 2022

1 Data Generation and Inspection

First of all, we generated new data. We set up the Jupyter Notebook with Python 3.8 and Jupyter Notebook 6.4 on a Windows Computer and run the cells.

We decided to first let the parameters as they were and analyze the result. We can see in figure 1 that at certain times there are up to 25 students in the room, ca 5 times a week. So it makes sense to assume that from Monday to Friday are business days, whereas Saturday and Sunday less to no students are going to the seminar room.

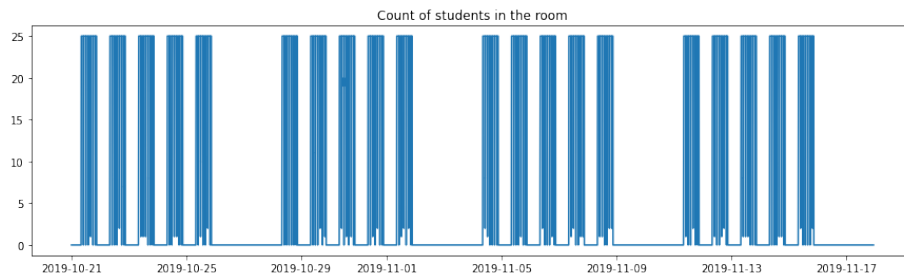


Figure 1: Count of students in the room in the original data.

Further, we plot the PACF plot in figure 2. As we can see, there are several significant spikes until 10. That is significant for finding the right coefficients (or order) since an ACF plot is a bar chart of the coefficients of correlation between a time series and its lags.

2 SARIMAX model

Next, we tried the SARIMAX model and adjust the parameters accordingly as described in the last section.

The predictions resulting are shown in figure 3.

Next, we use a simple function to compute the mean forecast error and we get the mean approx. $4.55e-06$.

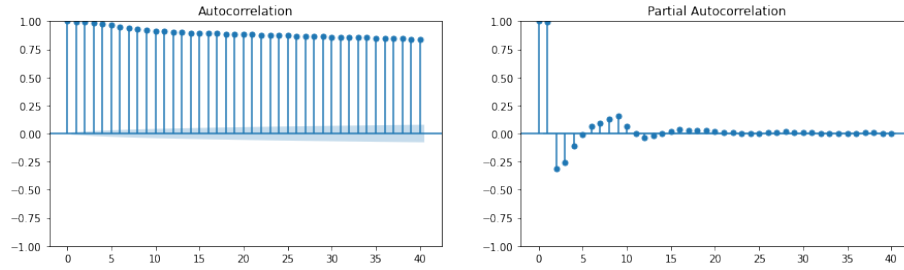


Figure 2: Plotting the pacf plot.

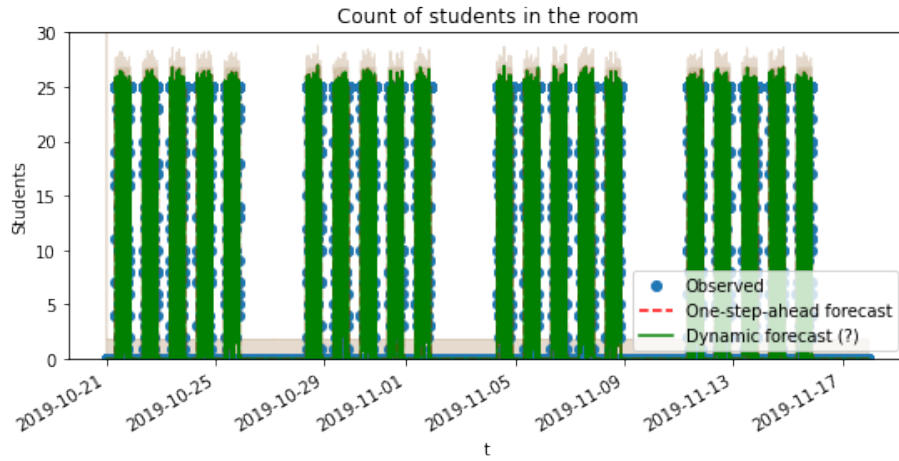


Figure 3: Predictions using SARIMAX.

3 ARIMA model

Next, we try the ARIMA model. We adjust the model accordingly. Also, we used `auto.arima()` to get the best parameters for the arima model. The predictions are shown in figure 4.

The mean forecast error is now at $7.53e-18$. So as you can see, the SARIMAX model was better fit to the data probabaly because we have this seasonal trend with having more students in the seminar room from Monday to Friday and less on weekends.

4 VARMAX model

Next, we also try the VARMAX model. The prediction is shown in figure 5 and the mean forecast error is now at $9.35e-18$.

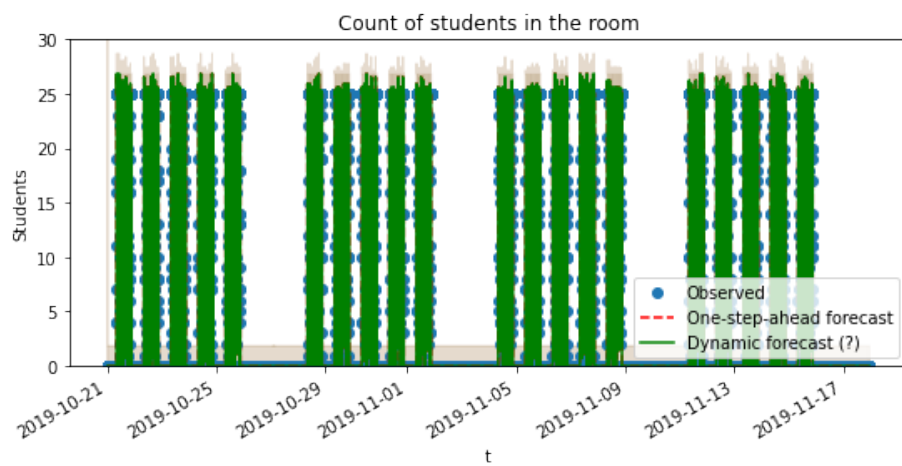


Figure 4: Predictions using ARIMA.

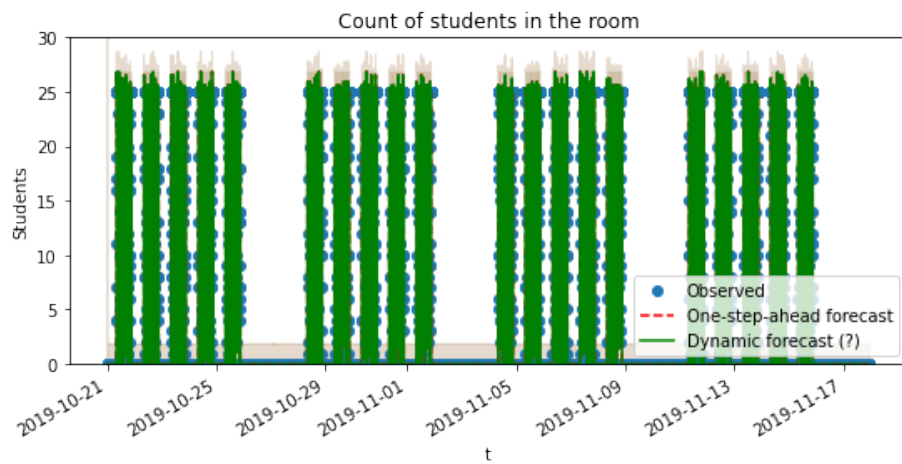


Figure 5: Predictions using VARMAX model.