Prediction of Crowdedness in Public Transportation

Kenneth Scheel, Laurits Fromberg and Nikolaj Espinoza

s174488, s174512, s163848

Technical University of Denmark

Introduction

Covid-19 restrictions have limited the number of people allowed in public transportation. The public transportation authority `Movia` would therefore like a model that can predict the crowdedness of a bus at a given time point, and has supplied the necessary data. This is done in order to inform people and thus enable them to either find an alternative way of transportation or reschedule their journey. Likewise, it facilitates the appropriate authorities and companies to pursue the necessary precautions, such as adjustment of the capacities, to ensure social distancing and hereby the decrease of widespread infection of Covid-19. The task is the supervised learning problem of predicting the future crowdedness in public transportation using *Rejsekort* data for a given bus route.

Data and Assumptions

The provided data consists of *Rejsekort* data for 60 days made with one hour resolutions from 2/9-2020 at midnight to 01/11-2020 at 11 PM. However, this data does not account for approximately 67% of travellers due to other purchase options. This percentile will depend upon the time of day with for instance more commuter travelling during rush hours, making it reasonable to assume that the model will under-perform during rush hours. We will now state the assumptions for the data.

- For the data 33% uses *Rejsekort*, scaling by 3 will give representative results for all travellers.
- All travelers check-in at the beginning of their journey and check-out at their final destination.

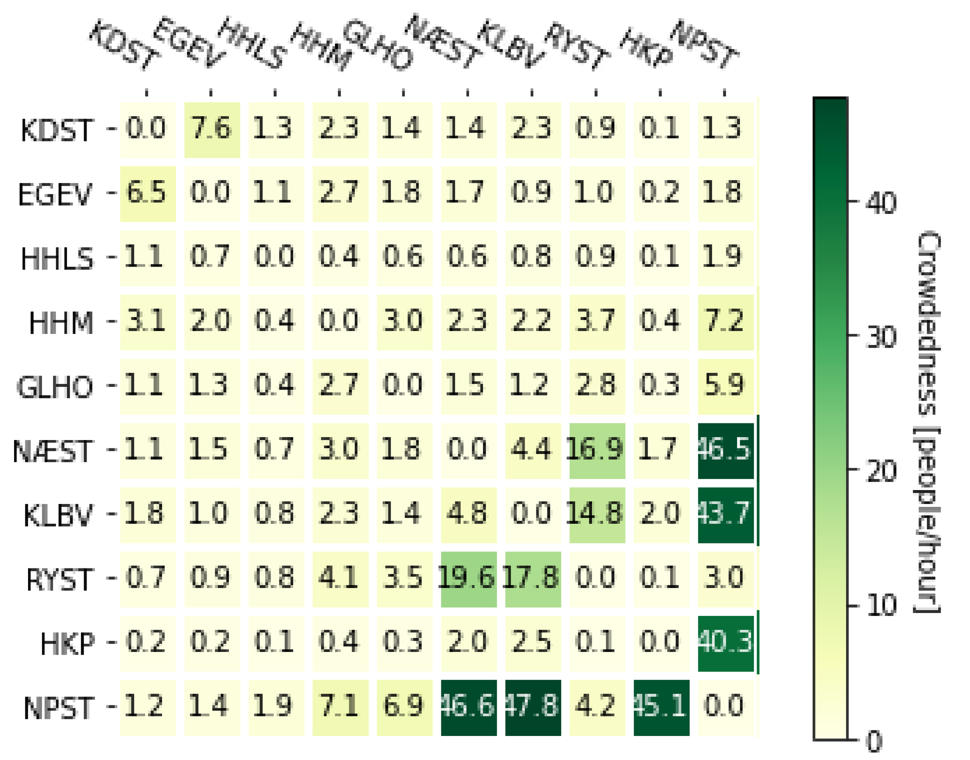

Figure 1: Average crowdedness on bus route 150S during the 60 days.

We have as an extension also considered weather data from the Danish Meteorological Institute `DMI` as the weather is expected to influence whether travellers use public transportation. The data consists of variables for the wind speed, the wind direction, the temperature and the amount of precipitation.

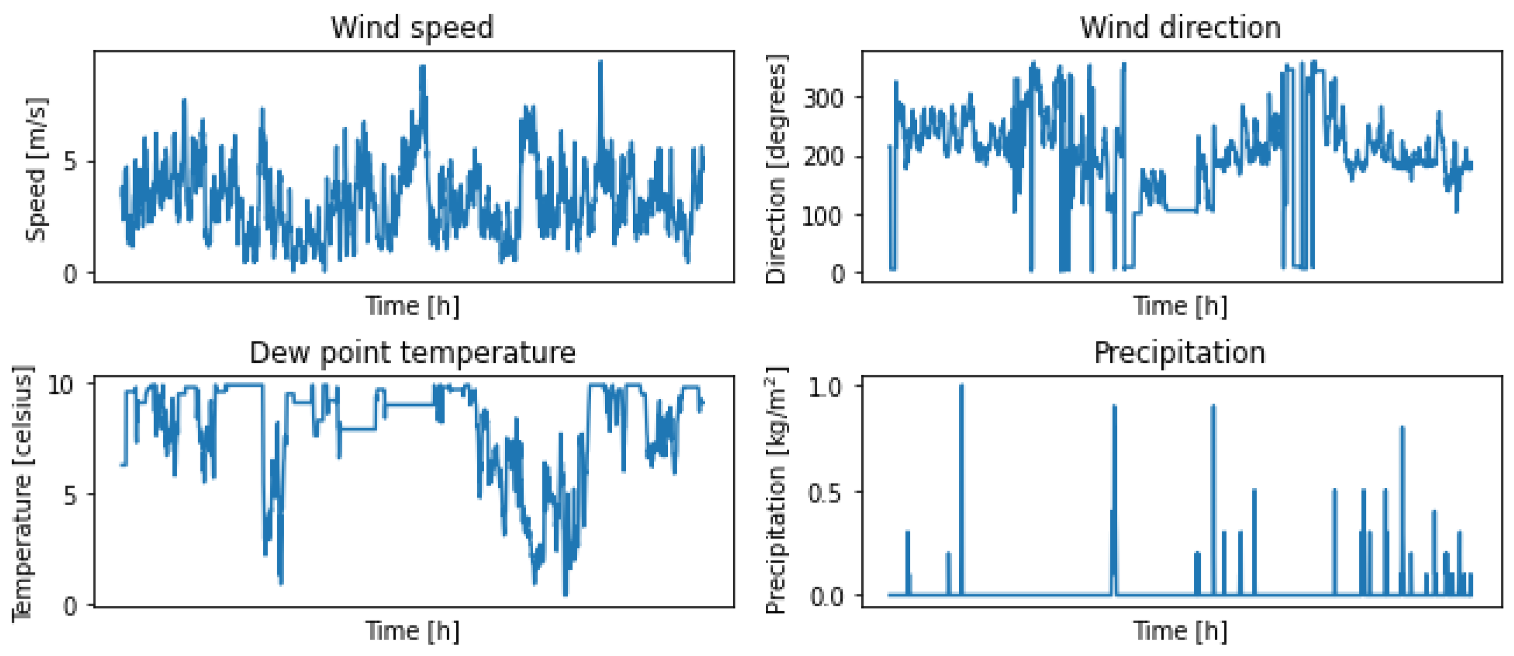

Figure 2: Weather variables during the 60 day period.

The target data is $k \times k$ matrices (where k is the number of stops on a given bus route), with the number of people who have taken a given journey, as alluded to in figure 1. These matrices are timeslices, and we can therefore think of the data as a collection of k^2 time series. The measurements from `DMI` were taken from a weather station in Copenhagen. Missing observations in the data have been artificially filled by previous values, such that we have weather measurements for each time slice.

Models and Results

The model is an artificial neural network, which combines the encoder part of the autoencoder, as well as a long short-term memory recurrent neural network due to the temporal behaviour of the data. Furthermore, different baseline models have been created in order to compare and examine the relative performance of the network and whether any insight is gained by the complex model.

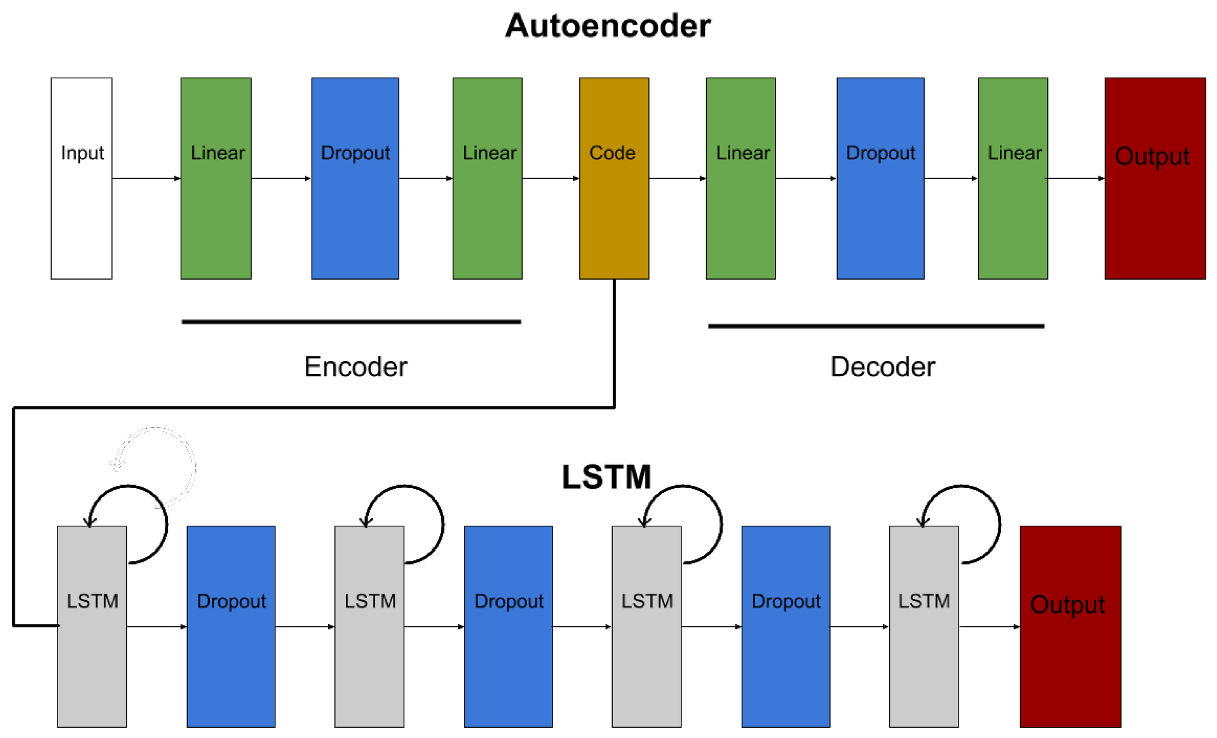

Figure 3: The architecture of the models.

The model architecture for the AE part of the models consist of Linear layers, dropout layers and uses the ReLU activation function. The recurrent part of the model consists of LSTM layers, dropout layers and also uses the ReLU activation function. Furthermore, the learning rate is reduced when not having achieved any significant improvement in the minimization within a given time frame.

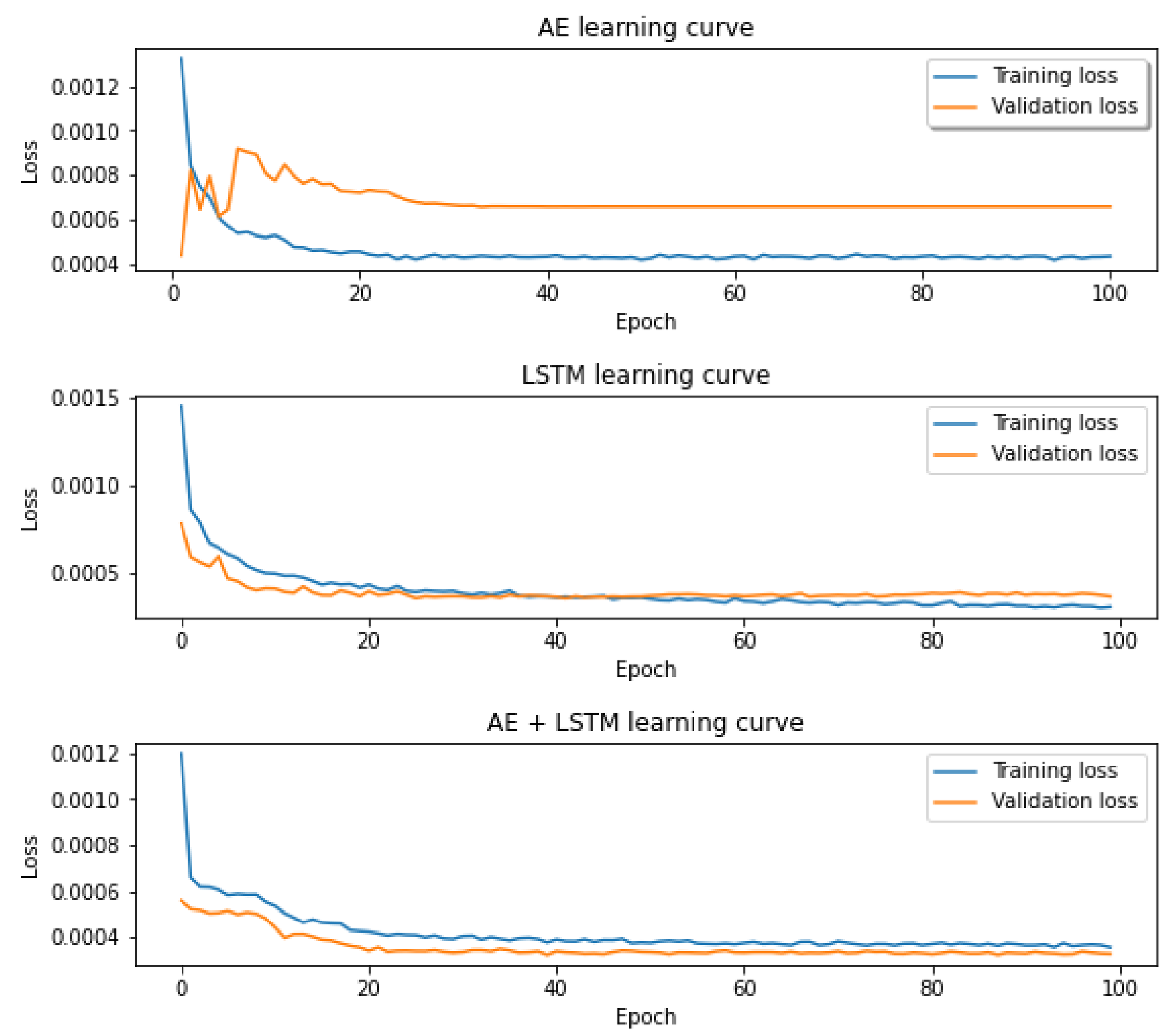

Figure 4: Learning curves for the models.

The training of the artificial neural networks has been achieved with the use of the *Adam* optimization algorithm with the mean squared error as its loss function in order to find the most suitable weights.

Model\Metric	MAE	MSE	RMSE
Baseline 1	2373.9	66761.3	258.4
Baseline 2	2379.3	108720.6	329.7
Baseline 3	2120.8	81986.1	286.3
AE	1762.2	107276.8	327.5
LSTM	1123.2	19643.2	140.2
AE + LSTM	1105.1	17321.9	131.6

Table 1: The error measures for each model applied to the test set.

Predictions for a Single Journey

We will now use the baselines and trained networks to make predictions for the test set. We have chosen a journey between two of the stops on route 150S, such that we can visualize the time series.

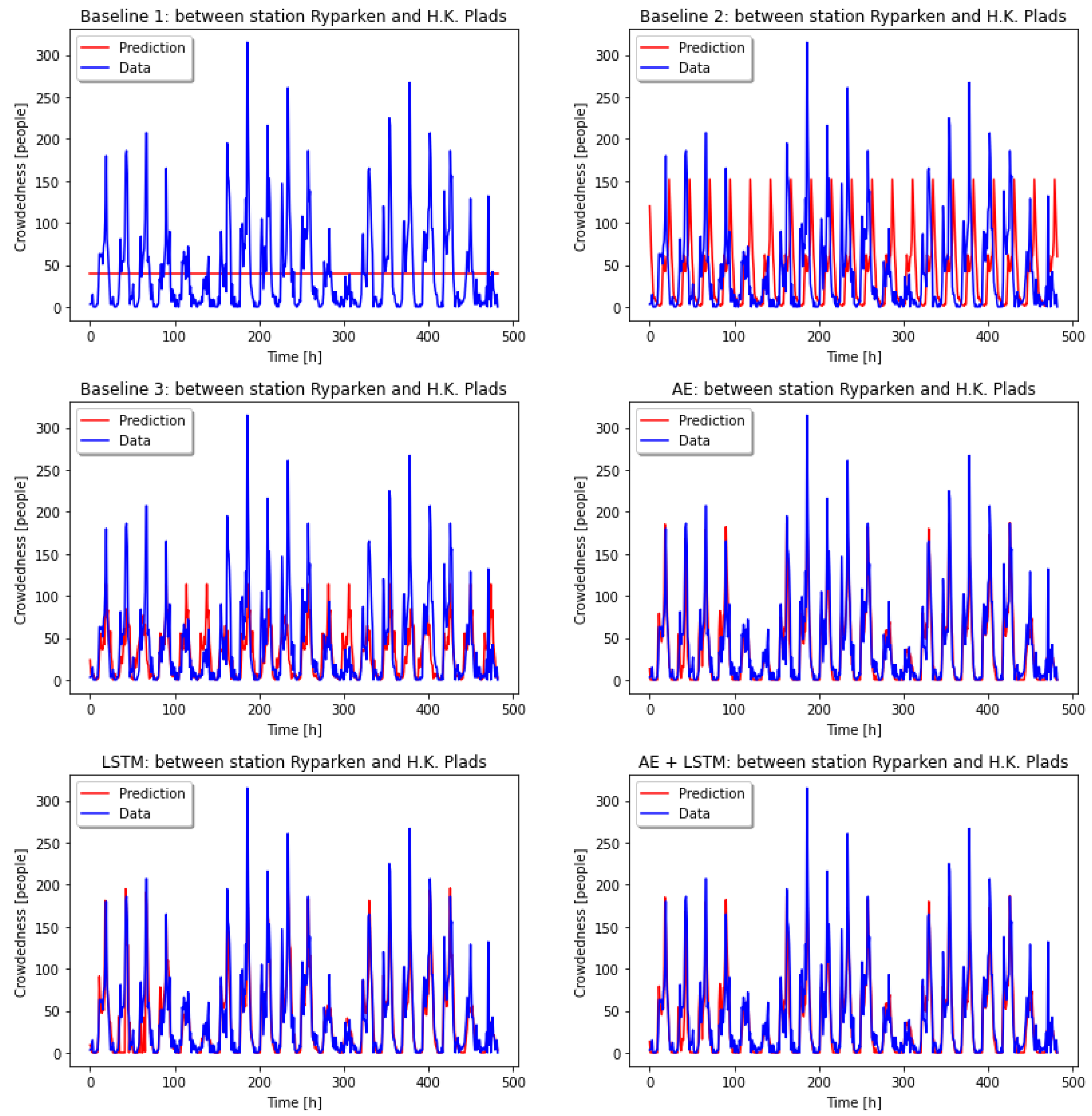

Figure 5: Model crowdedness predictions for the journey: Ryparken St. to Hans Knudsens Plads.

Conclusion and Future work

It is evident from the obtained results that the more complex artificial neural networks are able to describe the stochastic dynamics of the systems better than the simpler baseline models. This is seen from the different metrics where the most complex model provides the smallest values, as expected. It is also seen that the learning curves for both training loss and validation loss converge. The difference between the two (the generalization gap) is small, indicating a good fit. For future work it would be interesting to examine further the influence of other means of public transportation for instance delay on connecting trains and metros.

Figure 6: QR-code for the travel plan of bus route 150S.