The model comparison with a single feature and features with an identifier on LSTM

Gonsoo Moon
Post Master Student of Computer Science
University of Texas at Dallas,
gonsoomoon@gmail.com

Abstract

This report is to judge whether a model can learn multiple household's electric consumption with its own identifier. Also, it compares an individual model per household to the model containing three households.

1. Data Set

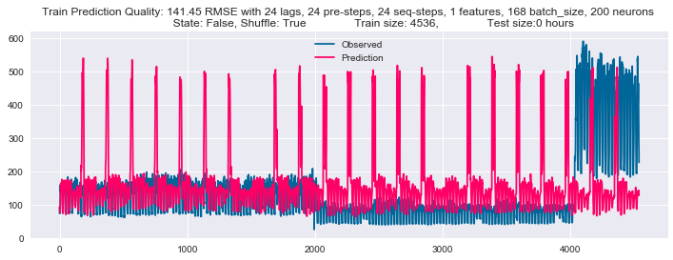
Using [2] the electricity consumption data set that is taken at 15 minute intervals over a four-year period from 2011 to 2014 is used. For experiment, I selected three households and a period of three months' data. The data is transformed for a supervised learning, consisting of features, 24 hour-series features of 0 hour to 23th hour for a day while labels are a 24-hour-series of 0 hour to 23th hour for the next day. Thus, the features and labels are corresponding to X and Y of the function, f: X -> Y, resulting in f mapping current day's electric consumption to the next day's one. For example, electric consumption of Sunday by hour is mapped to one of Monday by hour and likewise, is learned from Monday to Tuesday until Saturday to Sunday.

2. Two Stacked LSTM Model vs. MLP

The model is selected between MLP, Multilayer Perceptron, and LSTM (Long Short-Term Memory). Based on the hour-series data, I experimented the two models. two stacked LSTM is better than MLP while it is slightly better than one layer of LSTM. Thus, the two stacked LSTM was chosen and then it was optimized with the following hyper-parameters. a number of neurons is set to 200, the state parameter is set to being stateless, shuffle is set to being true, batch size is 168.

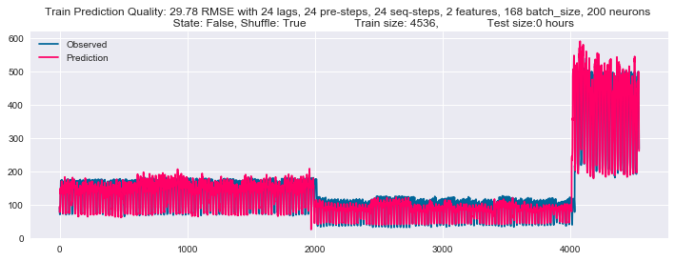
3. Two models: single feature model and two features' model

One model is trained for three households only using one feature, electric consumption. The other model is trained for two features such as electric consumption and households' identifier that is encoded as a series of a label such as 1, 2 and 3. The Fig.3-1 shows values of observation and prediction on the single feature model. The X axis means from 0 around 2000th hour, January to March, for the first household, from around 2000 to around 4000, January to March, for the second household, from around 4000 to the last hour, January, for the third household. The Y axis means electric consumption. The marks in blue are observations and ones in red are predictions. The predictions are not fitting to the observations for the three households.



<Fig. 3-1 observation and prediction on single feature model>

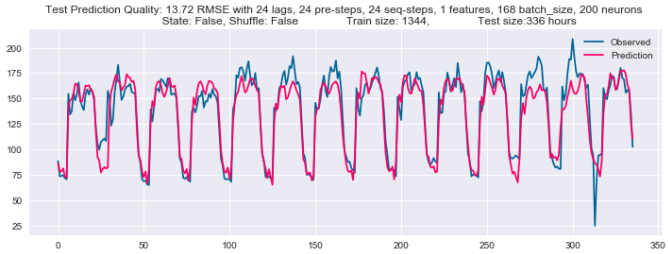
Conversely, the prediction on the Fig.3-2 shows fitting to the observations. each household's electric consumption shows a specific range. giving an example of from around 80 to around 180 for the first household. Thus, this model shows that if a specific identifier is fed into the model with other values, the model can be trained to recognize the identifier.



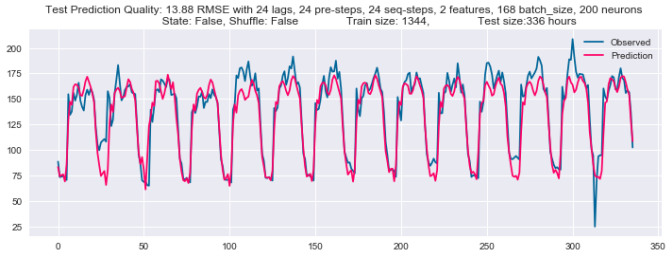
<Fig. 3-2 observation and prediction on two feature model >

4. Prediction per Household on Three Models

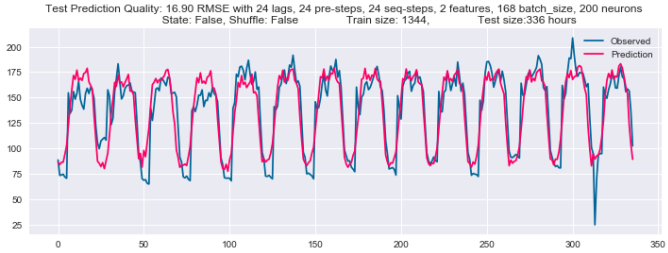
This experiment is conducted on three models for three households, e235, e250 and e252 respectively. The first model is trained on a single feature on a dedicated household. The second model is trained on two features including the household identifier on a dedicated household. The third household is trained on the two features on three households. The three models' structure and hyperparameters are equal but input dimensions are different. The first and second model shows 13.72 and 13.86 RMSE while the third model performs 16.90 RMSE for e235 household in <Fig. 4-1, 4-2, 4-3>. The first and second model are similar considering statistical significance. However, the third model on two features and three households is less accurate than them.



<Fig.4-1 observation and prediction on the first model for e235 >

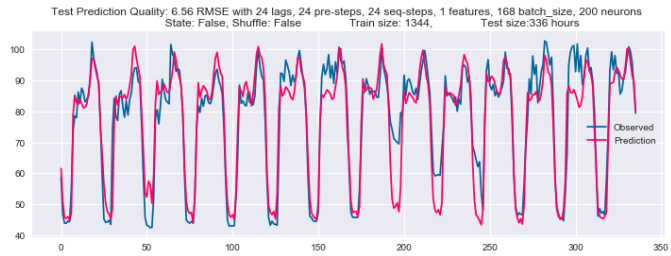


<Fig.4-2 observation and prediction on the second model for e235 >

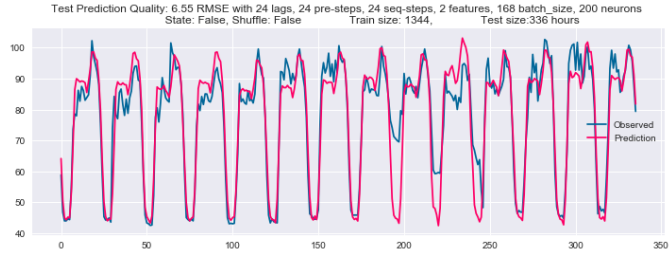


<Fig.4-3 observation and prediction on the third model for e235 >

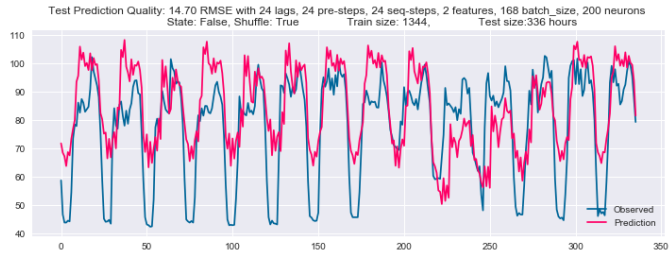
For e250 household, the three models perform 6.56, 6.55 and 14.70 RMSE in <Fig. 4-4, 4-5, 4-6> . The result of e250 is similar to e235.



<Fig.4-4 observation and prediction on the first model for e250 >

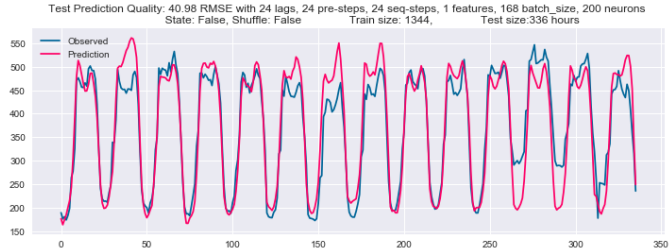


<Fig.4-5 observation and prediction on the second model for e250 >

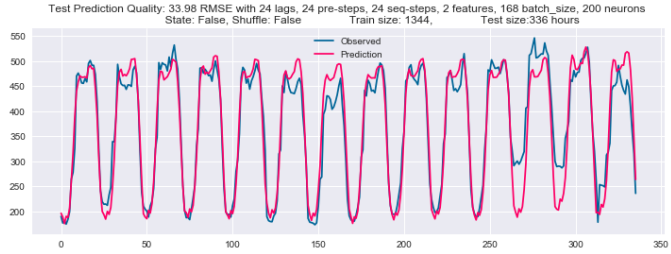


<Fig.4-6 observation and prediction on the third model for e250 >

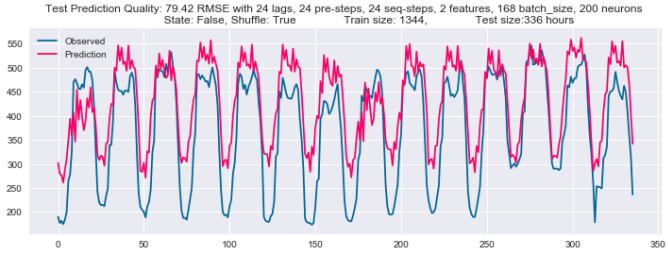
For e252, the three models perform 40.98, 33.98 and 79.42 RMSE in <Fig. 4-7, 4-8, 4-9>. The result of e252 is similar to e235 and e250.



<Fig.4-7 observation and prediction on the first model for e252 >



<Fig.4-8 observation and prediction on the second model for e252 >



<Fig.4-9 observation and prediction on the third model for e252 >

As a result, the models only for each household are better than the model for three households.

5. Conclusion

This report is for judging whether a model can distinguish each household with its own identifier. The model proves as experiment it can learn each household's pattern. Furthermore, each model only for a single household is better than the model learning three households.

6. References

[1] Antonio Gulli, Sujit Pal (2017). Deep Learning with Keras. Packt Publishing Ltd

[2] UCI Machine Learning Repository, https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014