E3S Web of Conferences **545**, 01004 (2024) https://doi.org/10.1051/e3sconf/202454501004
ICSREE 2024

A Comparative Analysis of Machine Learning Algorithms for Aggregated Electric Chargepoint Load Forecasting

Chang Li[1*], *Miao* Zhang[1], *Kevin* Förderer[1], *Jörg* Matthes[1] and *Veit* Hagenmeyer[1]

[1]Karlsruhe Institute of Technology, Institute for Automation and Applied Informatics, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

Abstract. With the development of electric vehicles in the last years, the number of electric chargepoints are expanding rapidly. Accordingly, the aggregated load demand from different electric chargepoints is increasing significantly. Due to the unpredictability of charging behaviour, it is difficult to build white-box models to analyse the patterns and to predict the load profiles, which is essential for other tasks such as demand side management. Thus, in this work, four different models based on machine learning and deep learning algorithms namely Random Forest (RF), Support Vector Regression (SVR), Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) are applied to a massive real-world open dataset from the UK, published in 2018, to compare the forecast performance of each algorithm with the modified persistence model as the baseline. The raw data are first pre-processed to generate the aggregated load demand by hour and then used for training and forecasting with a predictive horizon of 72 hours. The results are compared by using two common descriptive statistics, i.e., normalized Root-Mean-Square Error (nRMSE) and Mean Absolute Percentage Error (MAPE). In comparison we find that the GRU generates the lowest prediction error with 5.12% MAPE and 8.24% nRMSE in January 2017 and the modified persistence model generates the overall lowest prediction error with 2.88% MAPE and 3.76% nRMSE in July 2017.

1 Introduction

The rapid development of electric vehicle and the corresponding deployment of chargers will have an increasing impact on energy demand and interaction with existing grids [1,2]. However, due to the uncoordinated deployment of charging stations [3] and the system fluctuations regarding charging behaviors [4], the charging environment is dynamic [5]. This makes the use of traditional modeling methods such as white-box models difficult for an accurate analysis and forecast [6], which is essential for operational decision making such as demand side management. For instance, the results of short-term load forecasting can help utilities to optimize generation and to ensure grid stability in the short term.

* Corresponding author: chang.li@kit.edu

Recently, many researchers have investigated in the machine learning- and deep learning-based models for time series data analysis and forecasting [7-10] and highlighted the advantages of some algorithms such as LSTM [11,12]. The aggregated load demand can also be pre-processed as time series data. Therefore, it's possible to build and train different models based on these machine learning and deep learning algorithms.

The present work contributes to comparing four different models based on machine learning (RF, SVR) and deep learning (GRU, LSTM) algorithms in aggregated load demand analysis and forecast with the modified persistence model as the baseline. Another important goal is to test the capabilities of these algorithms by utilizing the selected dataset. These models are built and trained on the massive real-world open dataset from the UK, published in 2018 [13], which contains more than 3 million raw data points covering the whole country. This dataset is found to be useful in other research topics such as the performance analysis of battery-assisted charging [14] and demand response [15]. In the present work, these raw data are pre-processed into time series data by hour at first. Then the first 744 hours in January, April, July and October respectively are used as training set with cross-validation for model training and the next 72 hours of data are reserved for test. By utilizing the descriptive statistics, i.e., nRMSE and MAPE, the results of different models are compared, and the performances are evaluated.

The remainder of the paper is divided into four parts: Section 2 presents related work in load demand research of electric chargepoints with machine learning algorithms. The raw data and the results of the preprocessing are described in Section 3. In Section 4, a brief description of each algorithm used for training and forecasting is given. Section 5 presents, analyses and discusses the predicted aggregated load demand with different models. Finally, the main conclusions of this work are highlighted in Section 6.

2 Related Work

Based on the type of approach, the related work can be generally divided into two categories. The first category focuses on one specific algorithm and then optimizes the algorithm or compares with other approaches. For instance, a fast-charging power demand analysis and forecast based on LSTM is introduced in [3] and briefly compared with other RNN to highlight the slight advantages of LSTM in the fast-charging scenario. In [11], the authors use genetic algorithms (GA) to optimize a LSTM model for short to medium term aggregate load forecasting and briefly compare with other machine learning approaches to show the better performance of the proposed approach. In [16], an optimized AC (Ant Colony) algorithm is proposed and has shown outstanding performance in accuracy and computational efficiency when simulating EV charging load profiles. In [17] three different Transformer training strategies are compared with other models such as LSTM, MLP (Multi-layer Perceptron) to discuss the advantages and disadvantages of different models with different training strategies for load forecasting.

Another category focuses on the differences between various datasets by utilizing several algorithms. In [18], four different machine learning approaches are applied to two datasets to explore the differences regarding customer privacy when forecasting charging load. Similarly, in [19] two different machine learning algorithms are applied to a total preprocessed dataset which consists of multiple sources of data for predicting charging station utilization. In [20] the authors propose a deep transfer learning method named as DTr-CNN to tackle the problem with missing labeled training data for time series prediction in some actual situations. Through the experiments across different datasets, the effectiveness of the method is proven compared to other algorithms.

Based on the related work, we summarize that the comparative analysis results of different machine learning and deep learning approaches is greatly impacted by different datasets. Thus, it's necessary to conduct a comparative methodological analysis of a particular dataset before utilizing that dataset for further research steps such as demand side management. To the best knowledge of the authors, there is no straightforward comparative analysis based on the dataset [13]. Hence, the utilization and comparative analysis of dataset [13] can provide new and useful results for a further operational decision-making when considering the aggregated load of electric chargepoints as part of demand side management.

3 Data preprocessing

The raw data for the experimental statistics on the usage of domestic electric vehicle chargepoints are released by Department for Transport in the UK in December 2018. The raw data contain 3.2 million charging events recorded across approximately 25,000 funded domestic chargepoints through the whole year of 2017 [13]. Table 1 shows a part structure of the raw data.

Table 1. Part structure of the raw data.

Start Date	Start Time	End Date	End Time	Energy [kWh]	Plugin Duration [h]
2017-12-31	23:59:23	2018-01-01	18:20:23	8.8	18.35
2017-12-31	23:59:00	2018-01-01	00:03:00	10.2	0.07
…	…	…	…	…	…

Based on the date and time, the raw data are pre-processed to generate aggregated load demand by hour. In the preprocessing, any plug-in events that were less than 3 minutes in length are treated as anomalies and therefore excluded. Besides, it's assumed that the charging power is constant throughout the plugin duration based on statistical observation and the simplified piecewise-linear charging profile model in [21]. All pre-processed data for the year 2017 are presented in Fig. 1. For a rapid training, forecast and analysis, the data of January, April, July and October as representative months in each season are extracted separately as inputs, which are presented in Fig. 2 and Fig. 3.

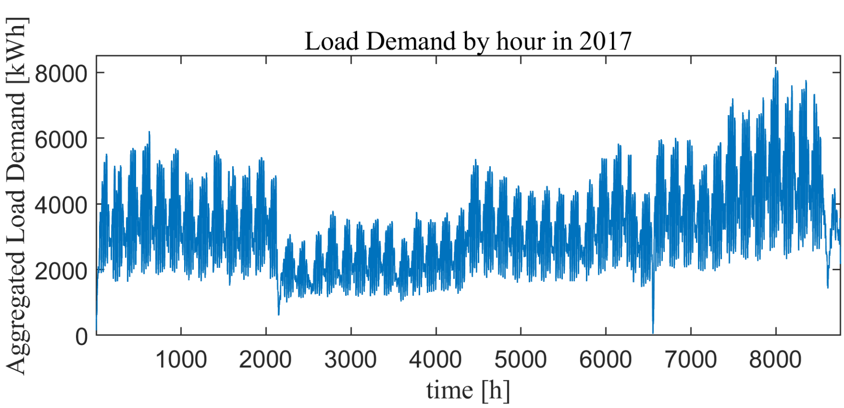

Fig. 1. Aggregated Load Demand by hour in the year 2017.

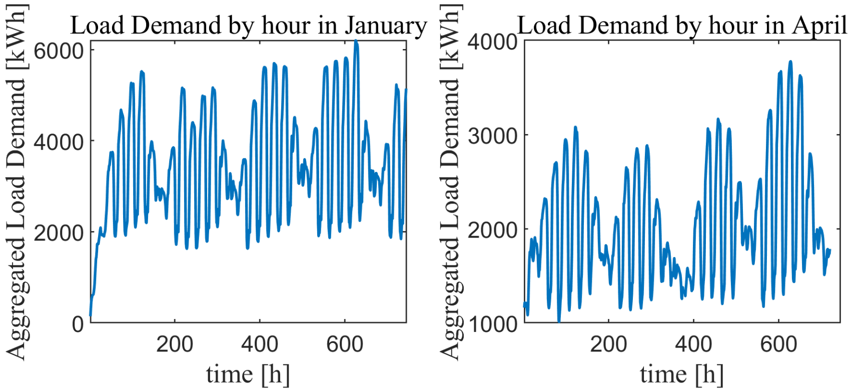

Fig. 2. Aggregated Load Demand by hour in January and April 2017.

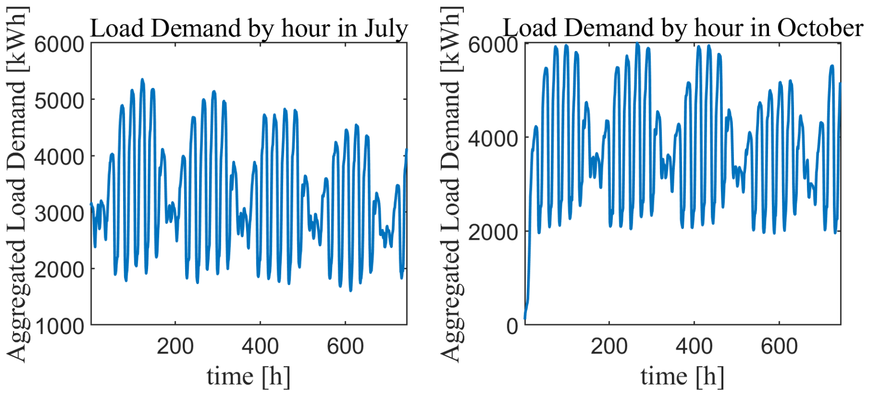

Fig. 3. Aggregated Load Demand by hour in July and October 2017.

4 Model training algorithms and modified persistence model

In the present work, four different algorithms are chosen for model training based on [22]. They have been proven to be powerful in time series data training in different scenarios as mentioned in Section 2. In this section, each of them is briefly described. Besides, the definition of our modified persistence model as the baseline is also included in this section.

4.1 Random Forest (RF)

As an ensemble learning method for classification and regression problems [23], RF has been widely used in many classification and regression problems. For time series forecasting, it requires that the time series dataset be transformed into a supervised learning problem first. Fig. 4 shows this transformation process, i.e., sliding window, with an input size of one as an example, where Y is the value at each time step. However, there is a limitation of this method that cannot be ignored, i.e., random forest cannot extrapolate. It means that predicted values are always within the range of the training set. In this work, different input sizes are tested to find an ideal parameter. Finally, we create a bagged regression ensemble object with an input size of 5 together with the temporal features of days such as Monday, Tuesday etc. as the 6th input, to use bootstrap aggregation method for model training, since there are no significant improvements with further increased input sizes.

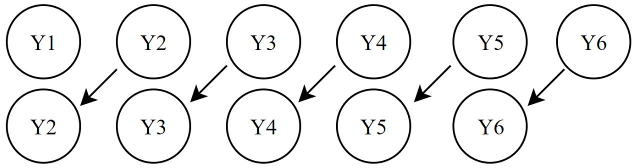

Fig. 4. Transformation of time series data into a supervised learning problem with input size of one.

4.2 Support Vector Regression (SVR)

As a variant of Support Vector Machine (SVM) for regression tasks, SVR has a great advantage to learn the nonlinear relationship between input data and a target output value with a margin around the target values by introducing a kernel function such as gaussian, linear and polynomial. Since it's designed to build regression models for nonstationary numeric values, it's selected with a linear kernel function for tasks like aggregated load demand forecasting in our work. To make meaningful comparisons, the setup for input size is kept consistent with RF.

4.3 Long Short-Term Memory (LSTM)

For predicting data based on time series while avoiding the vanishing gradient problem, LSTM has been developed as a modified version of traditional RNN. By introducing the so-called gates, LSTM can regulate the flow of information and maintain valuable information. In comparison to other RNN, LSTM can deal with large amounts of data and time steps more easily [24]. Based on these advantages, it's been chosen as one of the algorithms in the paper. The implementation is based on the library PyTorch.

4.4 Gated Recurrent Unit (GRU)

Gated recurrent units (GRU), introduced in 2014 by Kyunghyun Cho *et al.* [25], are also a gating mechanism in recurrent neural networks (RNN), which corresponds to a simplified version of LSTM. On consequence of its concise topology, GRU shows good performance with limited computational resources as an advantage. For the sake of comparison with LSTM, GRU is also included in the present work. The implementation is based on the library PyTorch.

4.5 Modified Persistence Model

The persistence model [26] is often used as a trivial reference model when different forecast models are compared. In this work, a modified version of the persistence model is defined by considering the temporal impacts. Instead of calculating the future value by assuming that no changes happen between the current time step and next time step, we use the values a week ago of the same time period, i.e., same days in the week as presented in Fig. 5.

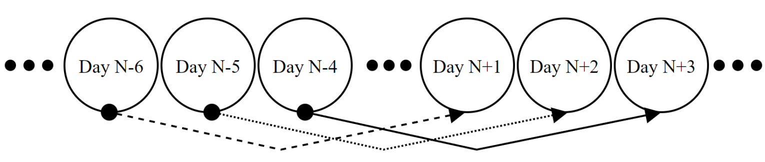

Fig. 5. Modified persistence model.

5 Results and discussions

As mentioned in Section 1, the training set with cross-validation contains the first 744 hours data in January, April, July and October and the test set is the subsequent 72 hours. The hyperparameters for RF and SVR are automatically optimized in MATLAB and the hyperparameter setting for LSTM and GRU in PyTorch is shown in Table 2. Table 3 summarizes the two descriptive statistics, i.e. nRMSE and MAPE, for each algorithm and the modified persistence model. The detailed plots are presented from Fig. 6 to Fig. 9.

The best results in each month are bolded. For instance, the best forecast results in January are given by GRU with a nRMSE of 8.24% and MAPE of 5.12%. However, GRU does not always give the best results in other months. For example, the results of LSTM are the best in April with a nRMSE of 7.57% and MAPE of 5.81%. And in the same month, even the baseline has a lower MAPE than the results with GRU. Therefore, the evaluation and discussion should be carried out based on each month.

Table 2. Hyperparameter setting for LSTM and GRU.

	LSTM	**GRU**
Input Size	1	1
Hidden Size	150	150
Hidden Layer	1	1
Epoch	200	200
Learning Rate	0.01	0.01
Criterion	MSEloss	MSEloss
Optimizer	Adam	Adam

Table 3. Summary of descriptive statistics for each algorithm.

		Persistence Model	**RF**	**SVR**	**LSTM**	**GRU**
nRMSE	January	16.35%	17.04%	32.27%	9.84%	**8.24%**
	April	11.80%	41.62%	35.46%	**7.57%**	27.82%
	July	**3.76%**	30.98%	33.40%	25.84%	10.69%
	October	11.72%	19.44%	34.99%	9.71%	**8.49%**
MAPE	January	12.18%	11.95%	31.23%	6.27%	**5.12%**
	April	9.15%	28.63%	31.29%	**5.81%**	18.42%
	July	**2.88%**	25.02%	33.72%	19.89%	8.94%
	October	7.68%	11.87%	34.23%	**6.08%**	6.48%

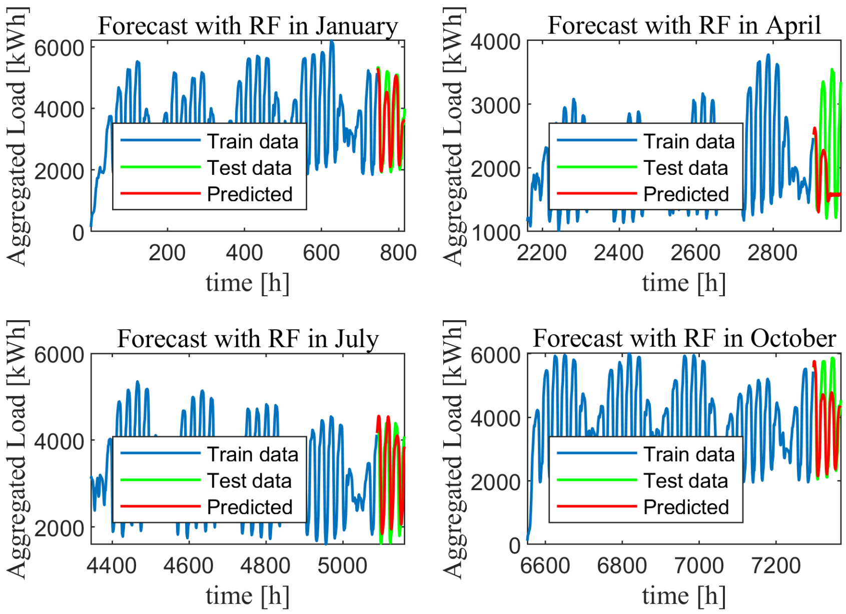

Fig. 6. Load Demand Forecast with RF.

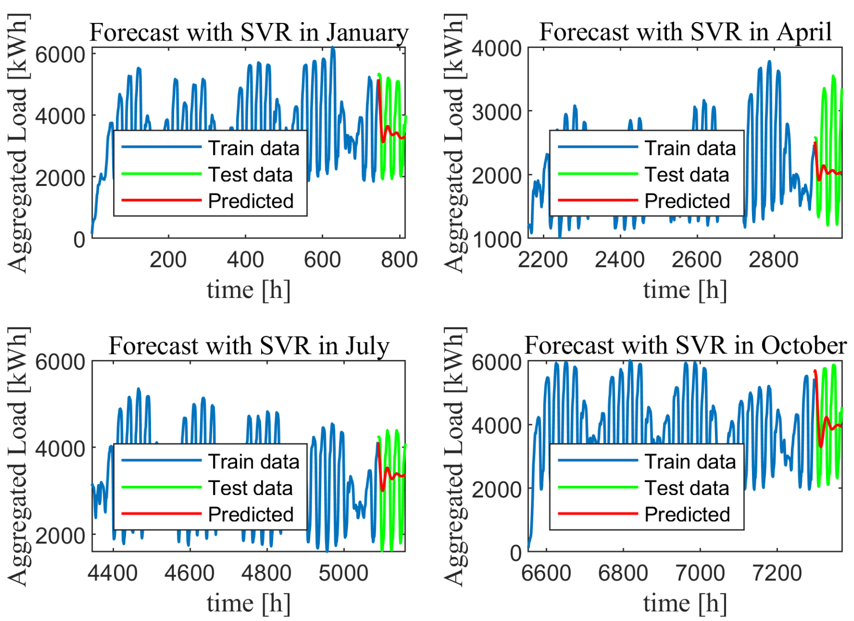

Fig. 7. Load Demand Forecast with SVR.

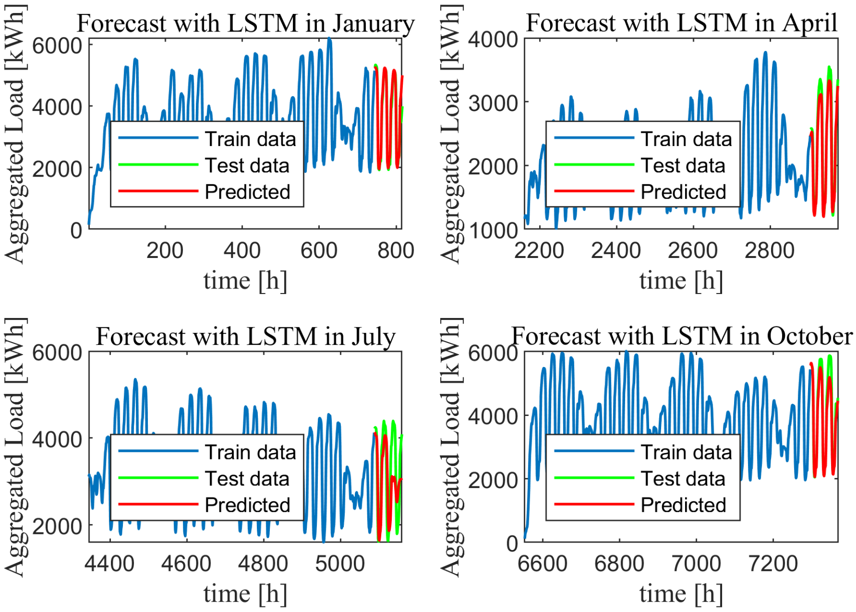

Fig. 8. Load Demand Forecast with LSTM.

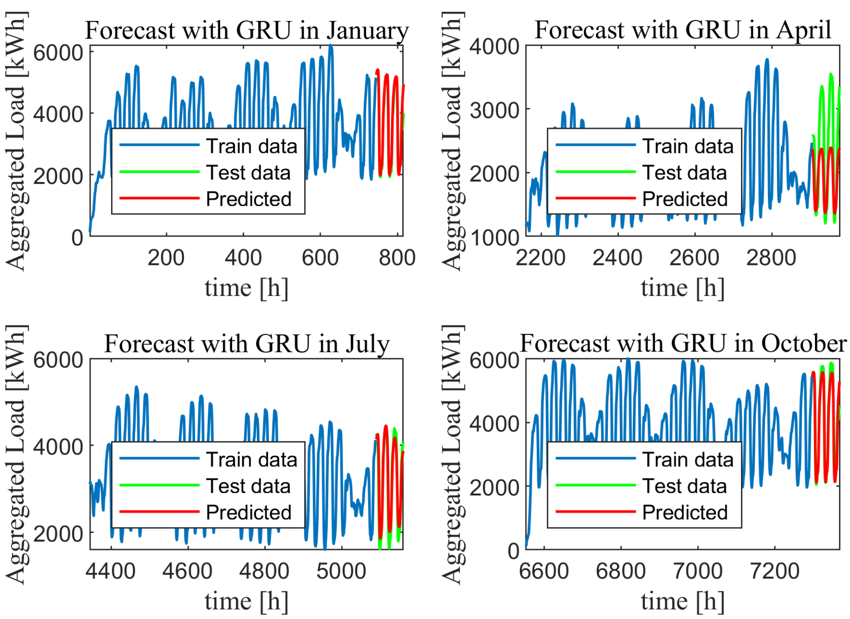

Fig. 9. Load Demand Forecast with GRU.

In April, it's worth noting that the prediction errors of almost all machine learning algorithms, except LSTM, are larger than the results of the modified persistence model. In July, the baseline results give even the overall lowest error with a nRMSE of only 3.76% and a MAPE of only 2.88%. There are two possible reasons for the relatively inaccurate forecast results in these two months, especially when utilizing RF and SVR. First, we have only used a relatively small training set i.e. one month for each season for a rapid training which could limit the quality of the trained model of all methods.

Besides, the number of features to consider when looking for the best split in RF impacts the quality of the results. If too many features are considered at each split, the model may overfit. If too few are considered, the model may underfit. Similarly, the features that capture patterns in time series data such as temporal trends have a huge impact when using RF and SVR for forecast. Therefore, more training data and different input sizes in different months together with new features such as different hour in the day should be considered as further steps to improve the quality of the trained model by utilizing RF [27,28] and SVR.

On the other hand, with a smaller input size of one, LSTM achieves the smallest prediction error in April which reflects the superiority of this deep learning-based algorithm with the hyperparameter setting in Table 2 in this scenario. However, the results in other months reveal the limitations of the current setup. The results in July show that when the value range fluctuates widely, in our case, the range is decreasing gradually, the trained model cannot predict the load demand with a good accuracy. On the contrary, the results in October are more accurate when the value range is more stable. One possible reason is that the method of closed loop forecasting within LSTM and GRU is implemented in the current work, which can be less accurate when compared to open loop forecasting because they don't have access to the true values during the prediction process.

Furthermore, these facts could also be caused due to lack of other features such as different hours in the day when training the model. The selection of relevant features such as temporal features is crucial for time series forecasting in this scenario. The aggregated load demand has shown different patterns on each hour according to the figures in Section 3. Therefore, to further improve the accuracy of the prediction, more new features such as pattern of different hours in the day should be included in the training process.

Moreover, a third possible reason for the relatively poor results in April with GRU and in July with LSTM would be that the model architecture of both deep learning-based methods contains only one hidden layer. Generally, adding more hidden layers can help improve the accuracy of the model by allowing it to capture more patterns in the training set. However, since the focus of this paper is not on optimizing a particular algorithm, different hidden layers in a LSTM or GRU model are not compared in the current work. Similarly, the hyperparameters in both algorithms are not fine-tuned for a better performance which could limit the ability of the model to capture more complex patterns in April and July. Thus, it's necessary to optimize the hyperparameters in the next step to tackle the inaccuracies in specific months.

6 Conclusions

The present paper investigates the accuracy of different machine learning and deep learning algorithms for aggregated load demand analysis and forecast with a modified persistence model as the baseline. The aggregated load demand is pre-processed with an hourly interval based on a massive real-world dataset consisting of the uncoordinated deployment of electric chargepoints over the UK. With a predictive horizon of 72 hours, we used the first 744 hours data in January, April, July and October 2017 for training and the next 72 hours data for prediction and as test set respectively. The forecasting results indicate that the modified persistence model achieved the overall best accuracy in July 2017 with a nRMSE of 3.76% and a MAPE of 2.88% and GRU presented the best accuracy in January 2017 with a nRMSE of 8.24% and a MAPE of 5.12%. In April, the smallest prediction error is produced by using LSTM. Moreover, the results with RF and SVR are worse than the baseline in almost all months, except that in January RF has a slightly smaller MAPE than the baseline, which shows the limitations of these two algorithms in our scenario. These results and conclusions could be used as a basis for a more comprehensive comparison when considering other important features in a longer time series data such as electricity price over time.

As further steps, larger training set for RF should be considered to improve the accuracy. The model architecture together with the hyperparameters for LSTM and GRU could be further tuned, compared and optimized for a better performance, especially in April and July. Moreover, new features such as electricity price over time and specific hours in the day could be included to improve the accuracy of results and further to help operational decision making such as flexibility optimization and demand side management.

Acknowledgements

This work is supported by the Energy System Design (ESD) Program of the Helmholtz Association (HGF) within the structure 37.12.02.

References

1. Savari G F, Sathik M J, Raman L A, et al. Assessment of charging technologies, infrastructure and charging station recommendation schemes of electric vehicles: A review. Ain Shams Engineering Journal, 2023, 14(4): 101938.
2. Das H S, Rahman M M, Li S, et al. Electric vehicles standards, charging infrastructure, and impact on grid integration: A technological review. Renewable and Sustainable Energy Reviews, 2020, 120: 109618.
3. Chang M, Bae S, Cha G, et al. Aggregated electric vehicle fast-charging power demand analysis and forecast based on LSTM neural network. Sustainability, 2021, 13(24): 13783.
4. Gong L, Cao W, Liu K, et al. Demand responsive charging strategy of electric vehicles to mitigate the volatility of renewable energy sources. Renewable Energy, 2020, 156: 665-676.
5. Alsabbagh A, Wu B, Ma C. Distributed electric vehicles charging management considering time anxiety and customer behaviors. IEEE Transactions on Industrial Informatics, 2020, 17(4): 2422-2431.
6. Chen Y, Guo M, Chen Z, et al. Physical energy and data-driven models in building energy prediction: A review. Energy Reports, 2022, 8: 2656-2671.
7. Lim, Bryan, and Stefan Zohren. "Time-series forecasting with deep learning: a survey." Philosophical Transactions of the Royal Society A 379.2194 (2021): 20200209.
8. Torres J F, Hadjout D, Sebaa A, et al. Deep learning for time series forecasting: a survey. Big Data, 2021, 9(1): 3-21.
9. Masini R P, Medeiros M C, Mendes E F. Machine learning advances for time series forecasting. Journal of economic surveys, 2023, 37(1): 76-111.
10. Shen Z, Zhang Y, Lu J, et al. A novel time series forecasting model with deep learning. Neurocomputing, 2020, 396: 302-313.
11. Bouktif, S.; Ouni, A.; Serhani, M.A. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. Energies 2018, 11, 1636.
12. Song X, Liu Y, Xue L, et al. Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. Journal of Petroleum Science and Engineering, 2020, 186: 106682.

13. Data.gov.uk. "Electric Chargepoint Analysis 2017: Domestics", 2018. Department for Transport. https://www.data.gov.uk/dataset/
14. Ali J, Dyo V, Zhang S. Battery-assisted electric vehicle charging: Data driven performance analysis. 2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe). IEEE, 2020: 429-433
15. Fang C, Zhao X, Xu Q, et al. Aggregator-based demand response mechanism for electric vehicles participating in peak regulation in valley time of receiving-end power grid. Global Energy Interconnection, 2020, 3(5): 453-463.
16. Yang S, Wu M, Yao X, et al. Load modeling and identification based on ant colony algorithms for EV charging stations. IEEE Transactions on Power Systems, 2014, 30(4): 1997-2003.
17. Hertel, M., Beichter, M., Heidrich, B. et al. Transformer training strategies for forecasting multiple load time series. Energy Inform 6 (Suppl 1), 20 (2023).
18. Majidpour M, Qiu C, Chu P, et al. Forecasting the EV charging load based on customer profile or station measurement? Applied energy, 2016, 163: 134-141.
19. Pevec D, Babic J, Kayser M A, et al. A data-driven statistical approach for extending electric vehicle charging infrastructure. International journal of energy research, 2018, 42(9): 3102-3120.
20. Ye R, Dai Q. Implementing transfer learning across different datasets for time series forecasting. Pattern Recognition, 2021, 109: 107617.
21. Zhang P, Qian K, Zhou C, et al. A methodology for optimization of power systems demand due to electric vehicle charging load. IEEE Transactions on Power Systems, 2012, 27(3): 1628-1636.
22. Mahesh B. Machine learning algorithms-a review. International Journal of Science and Research (IJSR), 2020, 9(1): 381-386.
23. Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001).
24. Zhu J, Yang Z, Guo Y, Zhang J, Yang H. Short-Term Load Forecasting for Electric Vehicle Charging Stations Based on Deep Learning Approaches. Applied Sciences. 2019; 9(9):1723.
25. Cho, Kyunghyun, et al. "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation". Association for Computational Linguistics. 2014
26. Notton, Gilles, and Cyril Voyant. "Forecasting of intermittent solar energy resource." Advances in renewable energies and power technologies. Elsevier, 2018. 77-114.
27. Schwenk K, Phipps K, Briegel B, et al. A Benchmark for Parking Duration Prediction of Electric Vehicles for Smart Charging Applications. 2021 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2021: 1-8.
28. Lahouar A, Slama J B H. Random forests model for one day ahead load forecasting. Irec2015 the sixth international renewable energy congress. IEEE, 2015: 1-6.