11th MALAYSIA STATISTICS CONFERENCE 2024
Data and Artificial Intelligence: Empowering the Future
Sasana Kijang, Bank Negara
19th September 2024

Time Series Modeling of Start-Stop Battery Electric Vehicle Charging

Syahrizal Salleh[1]; Roslinazairimah Zakaria[1]; Siti Roslindar Yaziz[1]

[1] Universiti Malaysia Pahang Al-Sultan Abdullah, Center for Mathematical Sciences, Kuantan, Pahang, Malaysia

Abstract: Relying solely on stored energy from electric charges in their battery packs, Battery Electric Vehicles (BEVs) propel their electric motors without the need for traditional combustion engines. Meeting this growing demand requires electricity utility providers to enhance electricity generation capabilities and upgrade distribution grids for BEVs charging stations. In the raw form, the start and stop charging data from BEVs charging stations is not capable to demonstrate the electricity load demand from BEVs charging activities. To address this limitation, transforming the data into a continuous time series format is essential for effective modeling of charging behavior, enabling trend and seasonality forecasting. Secondary data for this study was sourced from the My Electric Avenue project in the UK where 209 Nissan Leaf BEVs were leased to participants and its usage behavior is recorded. Using data from January 1, 2014 to December 15, 2014, the transformation involves counting simultaneous charging by augmenting data between start and stop times. A significant correlation exists between electricity demand from BEV charging and concurrent sessions. The corresponding Mean Absolute Percentage Error (MAPE) using the LSTM model on observed data recorded at 1.38% and its Root Mean Squared Error (RMSE) at 0.51. The LSTM model serves as a validation tool for the transformed time series suitability for forecasting. Electric utility companies can utilize a similar LSTM model, leveraging localized data, for forecasting long-term electricity demand from local BEV charging. This model can integrate into planning tools focused on upgrading electric generation and transmission facilities.

Keywords:
Battery Electric Vehicle; Electricity Demand; Time Series; Forecasting; Box-Jenkins; SARIMA

1. Introduction:

It is known that road transport is the second largest contributor to CO_2 emissions in the European Union (EU). If no drastic steps are taken to reduce the growth, it could be the largest contributor of CO_2 emissions in the EU by 2050 as highlighted by Krause et al. (2020). About 25% of global emission comes from vehicle tailpipe according to Teixeira and Sodre (2018). The transportation sector contributed 16% to the global CO_2 emission (Abdul Latif, 2021).

The adoption of battery electric vehicles has reached such a high popularity that most European countries committed to shift away from internal combustion engines (ICE) by 2030 and 2040 time frame (Krause et. al, 2020). Mass adoption of BEV will put stress on the country's national electric grid and electric generation since BEV need to be plugged into the electric grid for charging. Electric utility providers will need to plan for this new electricity demand created by BEV.

To anticipate and accommodate the burgeoning demand for electricity spurred by Battery Electric Vehicles (BEVs), electricity utility entities must proactively enhance their electricity generation and distribution grid network capacities. A vital aspect of this preparation involves adopting a robust forecasting methodology to rationalize substantial investments in infrastructure. In this investigation, the Long Short-Term Memory (LSTM) model is selected to project the electricity demand arising from BEVs during charging, when they are connected to the electric grid.

The choice of the LSTM model is grounded in its practicality for analyzing time series data, ensuring a systematic approach to forecasting in the context of BEV charging activities. Employed as a validation tool, the LSTM model serves to affirm the appropriateness of the transformed time series data for accurate time series forecasting. This methodological approach not only aids in forecasting the demand trajectory but also provides a rational basis for strategic decisions related to infrastructure development and capacity augmentation by electricity utility companies.

2. Methodology:

The raw form of start-stop battery electric vehicle charging data need to be transformed from the raw multivariate discrete time series data into univariate continuous time series data. Then, the final transformed data will be suitable to be used in the LSTM model for modeling and forecasting electric load demand from BEV during active charging session.

The transformed data is capable to address the simultaneous BEV charging. Refer to Figure 1 for the transformation step flow chart built using 5 data frames. The first data frame, Dataframe 1 is the raw data that contains start and stop date of BEV charging activities. The start and stop date is converted from string data type into datetime data type. The start date is set as the index of Dataframe 1. Dataframe 2 is additional transformation steps to the Dataframe 1. Desired date range can be selected easily while the start date of BEV charging is still the index of Dataframe 2. After selecting new date range, the index is reset and new dummy key column is added to Dataframe 2.

The third data frame, Dataframe 3 is built from scratch. Date range of this new data frame need to be identical to the Dataframe 2. The frequency interval of this data frame also need to match Dataframe 2. New column time is added to Dataframe 3 to track time from beginning to the end of date range in same frequency as Dataframe 2. New dummy key column is added to Dataframe 3. Dataframe 4 is result of left merge Dataframe 2 with Dataframe 3 on key column. The resulting outcome is data on column key in Dataframe 3 will fill data from start date to end date of BEV charging activities. New column *cnt_charge_events* is created to keep track of charging duration for each charging event.

The final data frame, Dataframe 5 is result of grouping all values in column time in a time series sequence. For time value that overlaps between multiple charging events, the number of overlaps is summed up and stored in column *cnt_charge_events*. Finally, the single time flow in minute-frequency interval is set as new index of Dataframe 5.

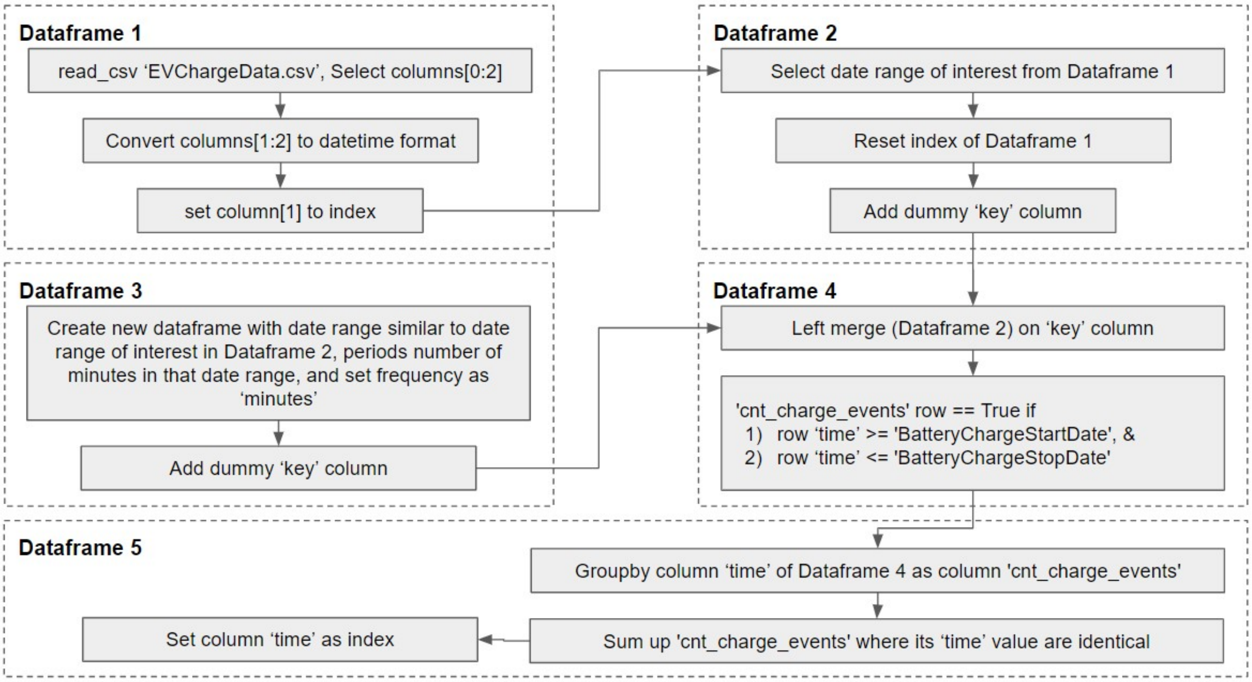

Figure 1 Flowchart of proposed procedure for start-stop data transformation process.

The Recurrent Neural Network (RNN) is a specialized neural network architecture tailored for processing sequential data. LSTM network successfully address the shortcomings of traditional RNNs by incorporating gate control mechanisms that integrate short-term memory with long-term memory (Hochreiter, 1997). The LSTM model is chosen to model the feature engineered time series data. The model will then be used to perform 1-step ahead forecasting.

Figure 5 visualizes a single LSTM cell, its input layer, output layer, and internal hidden layers. The LSTM cell take input from previous cell state memory (c_{t-1}), previous cell hidden state (h_{t-1}), and current input data (x_t). Internally, the intermediate state of forget gate (f_t), input gate (i_t), and prior cell state ($\acute{c}_t$) is calculated. Output from the LSTM cell are current cell state (c_t), current hidden cell state (h_t), and output gate state (o_t).

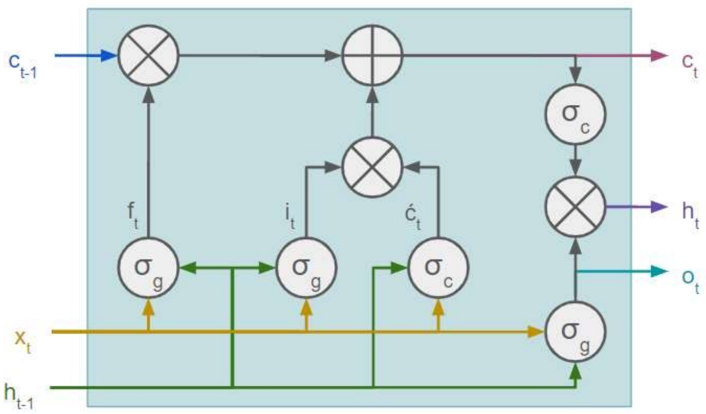

Figure 5: Visualization of single LSTM cell

The LSTM cell is defined by Eq. 1-6.

$$f_t = \sigma_g\left(W_f * x_t + U_f * h_{t-1} + b_f\right) \tag{1}$$

$$i_t = \sigma_g(W_i * x_t + U_i * h_{t-1} + b_i) \tag{2}$$

$$o_t = \sigma_g(W_o * x_t + U_o * h_{t-1} + b_o) \tag{3}$$

$$\acute{c}_t = \sigma_c(W_c * x_t + U_c * h_{t-1} + b_c) \tag{4}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \acute{c}_t \tag{5}$$

$$h_t = o_t \cdot \sigma_c(c_t) \tag{6}$$

The constants W_f, W_i, W_o, W_c, U_f, U_i, U_o, and U_c are weight matrices. The constants b_f, b_i, b_o, and b_c are biases. Both weight matrices and biases are not time-dependant. The σ_g is a sigmoid function defined by Eq. 7, and the σ_c is tanh hyperbolic tangent function defined by Eq. 8.

$$\sigma_g = (1 + e^{-x})^{-1} \tag{7}$$

$$\sigma_c = \frac{e^{2x}-1}{e^{2x}+1} \tag{8}$$

The hidden LSTM layer is structured as a sequential single layer, encompassing 125 units of LSTM cells. In the training process, the data was divided into sequences of 60 data points, each associated with a single expected output. The expected output corresponds to the subsequent data point in the training dataset. Example of a sequential single layer LSTM network is shown in Figure 6.

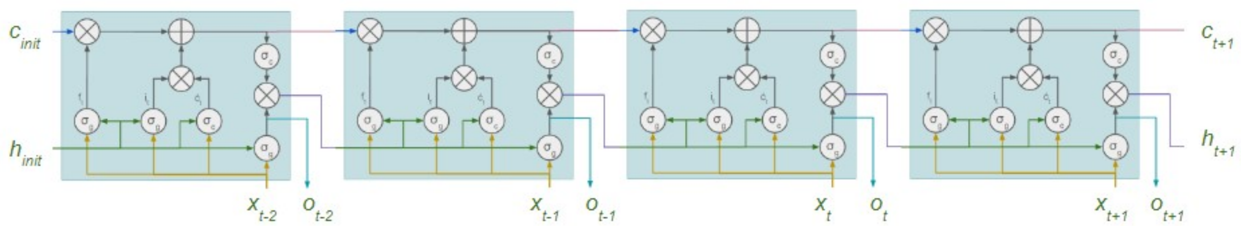

Figure 6: Visualization of four units of LSTM cells in sequential single layer configuration

3. Result and Discussion:

Figure 2 and 3 present the raw multivariate discrete data of start and stop charge date from *EVChargeData.csv*, sourced from the My Electric Avenue project. The plot in the main window of Figure 2 shows BEV start charge date for the whole selected date range. Resolution of the *x*-axis is in minutes, while the *y*-axis is count number of BEV that start charging. The small window on Figure 2 is a 24-hours snippet of the BEV start charge date. Each spike on the plot represents BEV start charge date (and time) that happened on each minute interval. It can be observed that multiple BEV start charge event has count of one with exception of one event where two BEV started charging on the same minute, thus giving count of two.

The plot in main window of Figure 3 shows BEV stop charge date for the whole selected date range. Characteristic of *x*-axis and *y*-axis on this plot is the same with Figure 2. The small window in Figure 3 is a 24-hours snippet of the BEV stop charge date. The date and time range of the snippet window in Figure 3 is identical to the snippet window in Figure 2. It can be observed also that all BEV stop charge event in the plot has count of one.

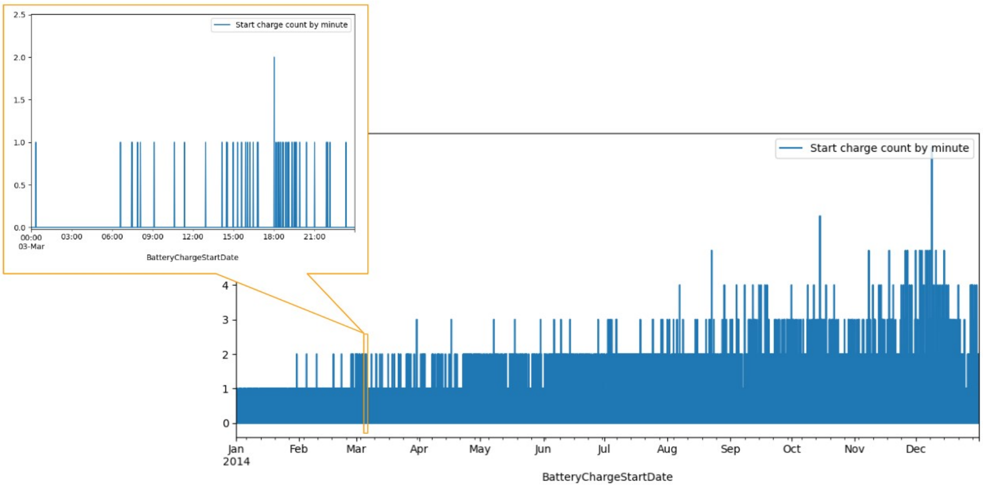

Figure 2: Start charge data sourced from the My Electric Avenue project

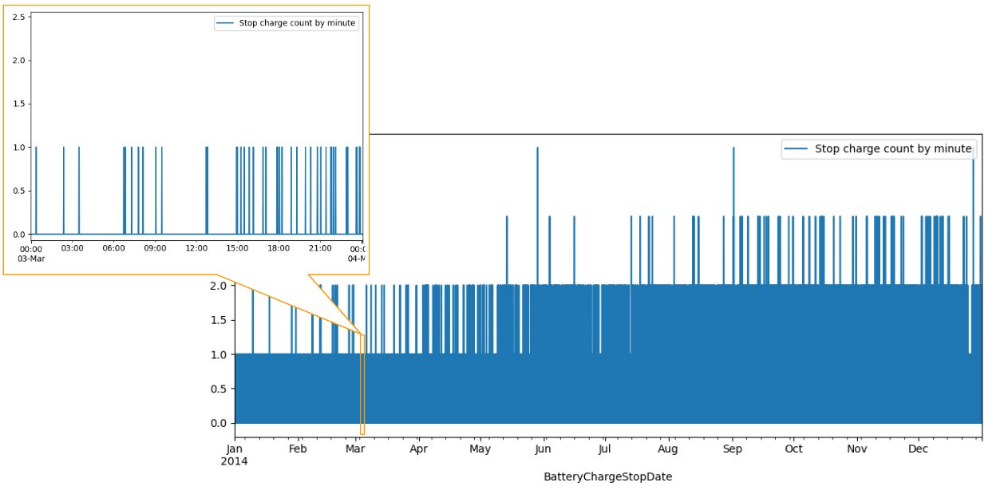

Figure 3: Stop charge data sourced from the My Electric Avenue project

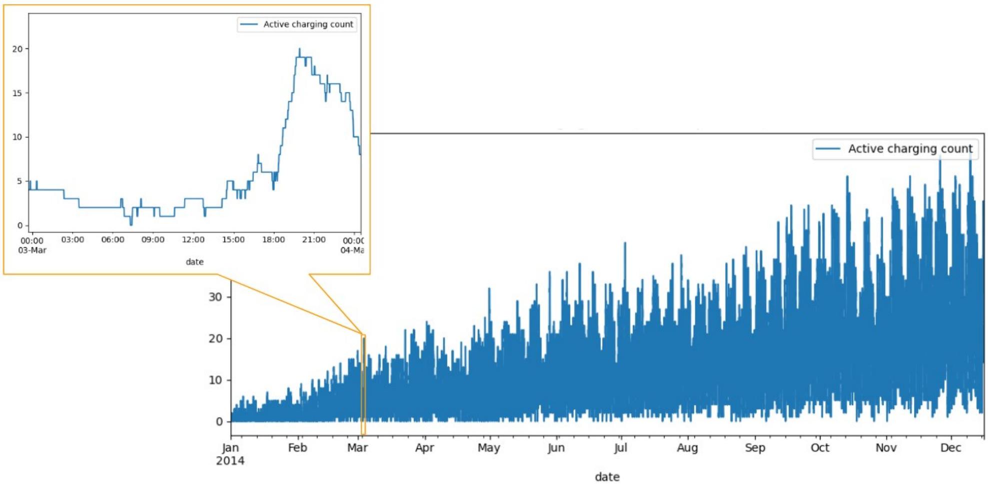

Figure 4: Feature engineered continuous time series data generated from the start and stop charge data sourced from the My Electric Avenue Project

It is known that charging BEV can take hours to fully charge its batteries. Each BEV that is actively charging will continuously draw power from power source. As another BEV starts charging while the first BEV is still charging, the two BEV will draw twice the power from power source if both charging stations use identical charging equipment. Taking the BEV start and stop charging data in its raw form does not explain the scenario of concurrent charging events. Using steps visualized in Figure 1, the outcome is continuous time series data as plotted in Figure 4. Characteristics of *x*-axis and *y*-axis of the plot in Figure 4 is the same as the plots in Figure 2 and 3. The small snippet window in Figure 4 has the same date range as the 24-hours snippet as in Figure 2 and Figure 3. Augmenting data between start and stop date of active BEV charging activities gives us insight of active concurrent charging activities.

The LSTM model employed in this study is configured for supervised learning. To impart the behavioral patterns of the transformed data to the LSTM network, each of the STL decomposed components of the observed data is partitioned into training and testing sets. Specifically, the training data comprises 80% of the dataset, starting from the initial entry, while the testing data consists of the remaining subsequent 20% of the dataset, similar to Koohfar (2023). Based on the observed data, the training dataset contains 452,304 data points, while the testing dataset consists of the remaining 50,256 data points.

4. Conclusion:

The BEV transformed data is found suitable to be used in time series modeling using LSTM method. The LSTM method proposed in this study has achieved MAPE value of 1.38% and RMSE value of 0.51, at epoch 20. However, there is room for improvement in the MAPE result from the residual components, suggesting potential enhancements for better overall accuracy in predicting BEV charging behavior. Load forecasting spans various temporal scales, including long-term, mediumterm, short-term, and ultra-short-term forecasting, each tailored to specific prediction periods. Ultra-short-term load forecasting (USTLF) is a specialized focus that involves estimating power consumption within a timeframe ranging from a few minutes to hours ahead. This type of forecasting is integral for facilitating rational tariff adjustments, ensuring smooth system operation, and enhancing economic efficiency as highlighted by Tong (2023).

5. Acknowledgement:

This research was funded by Universiti Malaysia Pahang Al-Sultan Abdullah (grant number RDU233008) under a special publication grant from the Research and Innovation Department, and Doctorate Research Scheme (Salleh, S.).

References:

1. Krause, J., Thiel, C., Tsokolis, D., Samaras, Z., Rota, C., Ward, A., Prenninger, P., Coosemans, T., Neugebauer, S. and Verhoeve, W., 2020. EU road vehicle energy consumption and CO2 emissions by 2050–Expert-based scenarios. Energy Policy, 138, p.111224.

2. Teixeira, Ana Carolina Rodrigues, and José Ricardo Sodré. "Impacts of replacement of engine powered vehicles by electric vehicles on energy consumption and CO2 emissions." Transportation Research Part D: Transport and Environment 59 (2018): 375-384.

3. Abdul Latif, S.N., Chiong, M.S., Rajoo, S., Takada, A., Chun, Y.Y., Tahara, K. and Ikegami, Y., 2021. The trend and status of energy resources and greenhouse gas emissions in the Malaysia power generation mix. Energies, 14(8), p.2200.

4. My Electric Avenue. EA Technology. [Online]. Available: https://eatechnology.com/media/iwjjxzeg/1myelectricavenue-i2evprojectsummaryreport.pdf.

5. Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9, no. 8 (1997): 1735-1780.

6. Koohfar, S., Woldemariam, W., & Kumar, A. Prediction of Electric Vehicles Charging Demand: A Transformer-Based Deep Learning Approach. Sustainability, 15(3), 2105 (2023).

7. Tong, C., Zhang, L., Li, H., & Ding, Y. Attention-based temporal–spatial convolutional network for ultra-short-term load forecasting. Electric Power Systems Research, 220, 109329 (2023).