

Divide-Conquer Transformer Learning for Predicting Electric Vehicle Charging Events Using Smart Meter Data

Fucai Ke¹, Hao Wang^{2,3,*}

¹Department of Human Centred Computing, Faculty of Information Technology, Monash University, Australia

²Department of Data Science and AI, Faculty of Information Technology, Monash University, Australia

³Monash Energy Institute, Monash University, Australia

Abstract—Predicting electric vehicle (EV) charging events is crucial for load scheduling and energy management, promoting seamless transportation electrification and decarbonization. While prior studies have focused on EV charging demand prediction, primarily for public charging stations using historical charging data, home charging prediction is equally essential. However, existing prediction methods may not be suitable due to the unavailability of or limited access to home charging data. To address this research gap, inspired by the concept of non-intrusive load monitoring (NILM), we develop a home charging prediction method using historical smart meter data. Different from NILM detecting EV charging that has already occurred, our method provides predictive information of future EV charging occurrences, thus enhancing its utility for charging management. Specifically, our method, leverages a self-attention mechanism-based transformer model, employing a “divide-conquer” strategy, to process historical meter data to effectively and learn EV charging representation for charging occurrence prediction. Our method enables prediction at one-minute interval hour-ahead. Experimental results demonstrate the effectiveness of our method, achieving consistently high accuracy of over 96.81% across different prediction time spans. Notably, our method achieves high prediction performance solely using smart meter data, making it a practical and suitable solution for grid operators.

Index Terms—Electric vehicle (EV), EV charging forecasting, nonintrusive load monitoring (NILM), smart meter, transformer

I. INTRODUCTION

As the popularity of electric vehicles (EVs) continues to surge, EV charging demand is increasing significantly. Un-coordinated EV charging poses substantial pressure on the power grid, primarily due to the considerable increase in peak load on distribution networks [1], [2]. The lifespan of network assets can be significantly impacted by overloading, as a result of increasing EV charging demand at homes [3], [4], [5]. The growing adoption of EVs presents new challenges for distribution system operators to plan and operate their systems. Therefore, it is essential to advance predictive models for EV charging events or called activities. Such predictive EV charging

This work was supported in part by the Australian Research Council (ARC) Discovery Early Career Researcher Award (DECRA) under Grant DE230100046.

*Corresponding author: Hao Wang (hao.wang2@monash.edu).

information is crucial for ensuring the stability and reliability of the network and holds substantial value for system operators [6], [7].

In recent years, there has been a large body of literature on various methods for predicting EV charging demand and enhancing EV charging management [8]. Such prediction studies often relied on a wide range of factors, including traffic data, ride-hailing trip data, arrival time, and charging rates [9], [10], [11], [12], [13], and utilized various optimization algorithms or machine learning models to predict different charging variables like charging rate [14], charging duration [15] and charging power [16]. All these methods provide useful insights into EV charging demand, primarily at charging stations [17]. However, a substantial number of charging events/activities occur at home rather than at charging stations, making the above existing studies less evident to system operators. More importantly, the above studies leveraged historical EV charging data to make predictions. However, behind-the-meter at-home charging is highly invisible to the operators and home charging data may not be available, making home charging event prediction even more challenging.

Several recent non-intrusive methods have gained prominence in various domains, for example, enhancing grid management through EV charging detection and disaggregation [18], [19], [20], [21], [22]. These innovative studies have demonstrated commendable performance in post-charging analysis or real-time detection, but a critical aspect remains unexplored: predicting future EV charging events, particularly in the context of at-home charging. In contrast to post-charging behavioral analysis, predictive information of future EV charging events plays a vital role in energy management. Therefore, *our work aims to fill this void by developing an effective algorithm for home EV charging event prediction using solely historical smart meter data*.

This paper introduces a novel method, namely Divide-Conquer Transformer EV charging event prediction (DCT-EV), for predicting EV charging events several minutes ahead with high accuracy based on smart meter data. We evaluate our proposed method using a real-world dataset and compare it with four representative machine learning and deep learning models

for different prediction requirements (minutes ahead). The main contributions of this paper are summarized as follow.

- **New practical task:** Unlike existing studies on predicting EV charging at charging stations based on historical charging data, our work focuses on a more challenging and practical task of predicting home charging events using historical residential meter data. Our work builds upon the NILM concept to provide predictive information about future EV charging events, deviating from post-charging analysis.
- **Non-intrusive, minute-interval, high-accuracy model:** DCT-EV, a self-attention-based model with a transformer backbone, divides long historical smart meter time-series into multiple sub-sequences and processes them in parallel. This approach enables effective handling of meter data to serve EV charging event prediction.
- **Extensive experiments and insights into EV charging event prediction using real-world datasets:** Our method, DCT-EV, outperforms benchmark models, establishing a new state-of-the-art in providing both 10-minute and 60-minute ahead EV charging event prediction. Extensive experiments provide valuable insights into 1) the impact of input history length on prediction performance, 2) the trade-off between model performance and the size of predictive time span, and 3) probability threshold selection.

The remainder of the paper is organized as follows. Section II presents the design of our DCT-EV model. Section III presents the experimental results, including comparisons between DCT-EV and benchmark models on a real-world dataset and the discussion of the evaluation results. Section IV concludes the paper and outlines some future work directions.

II. METHODOLOGY

The high-accuracy prediction of at-home EV charging events holds substantial value for energy management and grid operations. In this section, we present the structure of our proposed DCT-EV and its working principles in predicting future charging events using smart meter data, more specifically household load data. The model consists of three main components: the load embedding module, the EV representation learning module, and charging event prediction module, as illustrated in Figure 1.

In the electricity load embedding module, we first divide the lengthy smart meter data sequence into a few overlapping short sequences and then project the human-understandable and observable signals into high-dimension embedding vectors using convolutional layers. This approach can substantially reduce the computational resource and mitigate the inherent inductive bias for modeling load patterns. In the EV representation learning module, we boost the model's capacity by incorporating a self-attention-based transformer encoder. The encoder generates representations by taking into account the information from all sub-sequences. The charging event prediction module then predicts the near-future EV charging event and forecasts the probability of active EV charging over upcoming time steps.

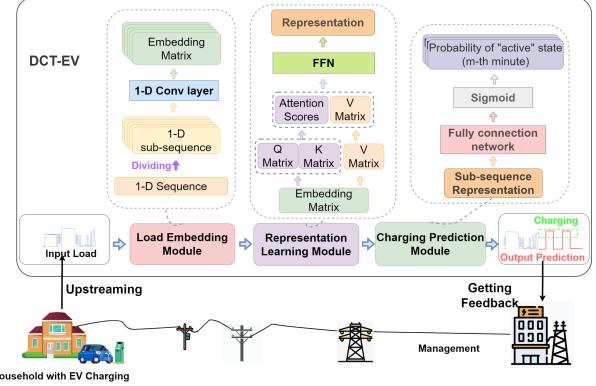


Fig. 1: DCT-EV structure and application.

A. Load Embedding Module

The objective of the load embedding module is to partition the long historical load sequence into several overlapping short sequences, drawing inspiration from [23], [24], [25]. This module can reduce computational resource consumption and enable the subsequent module to process the entire sequence in parallel.

The load embedding module takes each time series of electricity load from smart meter data as an input sequence, denoted as $\mathbf{x} = x_1, \dots, x_t, \dots, x_T \in \mathbb{R}^{1 \times T}$, where T is the length of the input sequence. The long input sequence \mathbf{x} is divided into $N = \frac{T-L}{\text{stride}} + 1$ number of overlapped sub-sequences, denoted as $\bar{\mathbf{x}} = \{\mathbf{x}^1, \dots, \mathbf{x}^n, \dots, \mathbf{x}^N\} \in \mathbb{R}^{N \times L}$, based on sub-sequence length L and a stride size of stride , where $\mathbf{x}^n = \{x_1^n, \dots, x_l^n, \dots, x_L^n\} \in \mathbb{R}^{1 \times L}$.

Each sub-sequence \mathbf{x}^n is then projected to a high-dimensional latent space $\chi^n = \{\chi_1^n, \dots, \chi_d^n, \dots, \chi_D^n\} \in \mathbb{R}^{1 \times D}$ using a convolutional layer as follows

$$\chi^n = \mathbf{x}^n \kappa, \quad (1)$$

where D represents the dimension of the latent space, and $\kappa \in \mathbb{R}^{L \times D}$ is the convolution layer kernal matrix.

The utilization of sub-sequence partitioning reduces information redundancy by treating each sub-sequence as a token instead of a single electricity load, thereby reducing the count of input tokens from T to N . The subsequent component, the EV representation learning module and charging event prediction module, constitute a self-attention backbone, demanding substantial memory resources, depending upon the number of the input tokens. When calculating the attention scores, the computational resources required for these two different sequence sizes exhibit a notable difference, with approximately $T^2 \gg N^2$, showing the advantage of our sub-sequence partitioning.

B. EV Representation Learning Module

The backbone of EV representation learning module is the self-attention mechanism, which is also a key element of the Transformer architecture [26]. We harness the attention mechanism to facilitate representation learning. Subsequently,

we employ a lightweight linear layer to interpret the essential information and make predictions.

The N embeddings of in-parallel sub-sequence are distinguished by their temporal order of interactions. To capture the intrinsic temporal dependencies within time series data, we implement **positional encoding**, a technique employing fixed sinusoids of varying frequencies [26], to model the relative ordering of time series data as

$$\chi_{\text{pos}} := \chi + \mathbf{b}^{\text{pos}}. \quad (2)$$

Multi-head attention layer serves as a representation learner for EV charging behavior by computing the attention scores for all contextualized embeddings of each sub-sequence. Each sub-sequence is mapped into the queries, keys, and values in each head h . More specifically, the mapping can be expressed as

$$\mathbf{Q}_h = \chi_{\text{pos}} \mathbf{W}_{h,Q}, \mathbf{K}_h = \chi_{\text{pos}} \mathbf{W}_{h,K}, \mathbf{V}_h = \chi_{\text{pos}} \mathbf{W}_{h,V}, \quad (3)$$

where $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V \in \mathbb{R}^{D \times D_m}$ are learnable mapping matrices, with D_m being the dimension after mapping, and \mathbf{K}' is the transpose of \mathbf{K} . Then, the scaled dot-product attention scores $\boldsymbol{\alpha}_h \in \mathbb{R}^{N \times N}$ and the hidden representation of load record $\boldsymbol{\psi}_h \in \mathbb{R}^{N \times D_m}$ are computed as

$$\boldsymbol{\alpha}_h = \text{Softmax}\left(\frac{\mathbf{Q}_h \mathbf{K}'_h}{\sqrt{d_{\mathbf{K}'_h}}}\right); \boldsymbol{\psi}_h = \boldsymbol{\alpha}_h \times \mathbf{V}_h, \quad (4)$$

where $d_{\mathbf{K}'_h} = D_m$ is the dimension of the key embedding matrix in head h . Thus the final attention is $\boldsymbol{\psi} \in \mathbb{R}^{N \times D}$ after summarizing all $\boldsymbol{\psi}_h$.

Feed-forward neural network (FFN) processes $\boldsymbol{\psi}$, which represents the output after layer normalization across all heads, and passes it through a two-layer multi-layer perceptron with ReLU activation to obtain the final output of the interaction encoder as

$$\mathbf{Z} := \text{ReLU}(\boldsymbol{\psi} \Phi_1 + \mathbf{b}_1) \Phi_2 + \mathbf{b}_2, \quad (5)$$

where $\Phi_1 \in \mathbb{R}^{D \times D'_m}, \Phi_2 \in \mathbb{R}^{D'_m \times D}$ and $\mathbf{b}_1 \in \mathbb{R}^{D'_m}, \mathbf{b}_2 \in \mathbb{R}^D$ are respectively the learnable weight matrices and bias vectors of the FFN, while $\text{ReLU}(\cdot)$ denotes the ReLU activation function.

As a result, the representation of each sub-sequence is denoted by $\mathbf{Z} \in \mathbb{R}^{N \times D}$, with each vector in the high-dimensional representation encapsulating the information of a sub-sequence.

C. Charging Event Prediction Module

Building upon the well-learned representation \mathbf{Z} , which is generated by the previous two modules processing the input load sequence, the charging event prediction module employs a trainable, lightweight fully connected network to obtain the final predicted probability of EV charging state in the near future.

The flattening technique is applied to reshape \mathbf{Z} , resulting in $\mathbf{Z} \in \mathbb{R}^{(N \times D) \times 1}$. For each future time step (m -th), the probability of EV charging event being “active” is computed by

$$\hat{y}_m = \sigma(W'_m \mathbf{Z}), \quad (6)$$

where $W'_m \in \mathbb{R}^{(N \times D) \times 1}$ is the weight matrix for m -th minute prediction, and $\sigma(\cdot)$ is the sigmoid activation function.

D. Loss Function

In the training of DCT-EV, the objective is to minimize the binary cross-entropy loss between the ground-truth labels and the corresponding prediction as follows

$$\mathcal{L} = - \sum_i \sum_m (y_{i,m} \log(\hat{y}_{i,m}) + ((1 - y_{i,m}) \log(1 - \hat{y}_{i,m}))), \quad (7)$$

where $y_{i,m}$ is the EV charging event state in the next m -th minute for user i , and $\hat{y}_{i,m}$ is the corresponding prediction.

III. NUMERICAL RESULTS AND ANALYSIS

In this section, we present the experimental settings, evaluation metrics, and corresponding results and discussions. The simulation experiment employs DCT-EV to predict future EV charging event in minute-interval based on historical meter data. As we will show in this section, DCT-EV can accurately predict near-future EV charging event.

A. Experimental Settings and Evaluation

In our experiment, we preprocess minute-interval residential electricity meter data with EV charging load records in 2018 from Pecan Street [27]. The dataset provides smart meter data aligned with EV charging profiles, which will be used to construct EV charging label for the experiment. After analyzing EV charging profiles, we set the EV charging event label as “charging” when the EV load is higher than 3 kW, and “not charging” otherwise. Note that our supervised learning method replies on the EV charging label to learn an effective prediction model.

To set up our experiment, we select homes with EVs and filter out homes without a ‘charging’ label in their records, resulting in 22 households with 8,063,175 electricity usage records for the experiment. For the selected households, we extract 402,296 EV charging records in total. In the real-world scenario, meter data is accessible to operators, but the EV charging records are not. Therefore, we use electricity usage as the model input, and EV charging records are merely used for validation. We divide the data into training and testing sets with an 80:20 ratio. The first 80% of each home’s records are for model training, the remaining 20% are reserved for testing.

The dataset exhibits an imbalance, primarily because EV charging events occur for only a few hours per week, resulting in limited EV charging records. As a classification problem, i.e., predicting future EV charging events, we utilize a set of performance metrics, such as recall, precision, accuracy (ACC), area under the ROC curve (AUC), average precision (AP), and F1 score (F1), along with mean squared error (MSE), to assess the model’s overall performance. All these metrics are classical indicators for assessing model performance in machine learning classification problems [28].

In our experiment, the model outputs a probability value representing the likelihood of an active charging event. To compute true positives, true negatives, false positives, and false negatives—necessary for subsequent F1 score, ACC ¹

¹@.5 means threshold is 0.5.

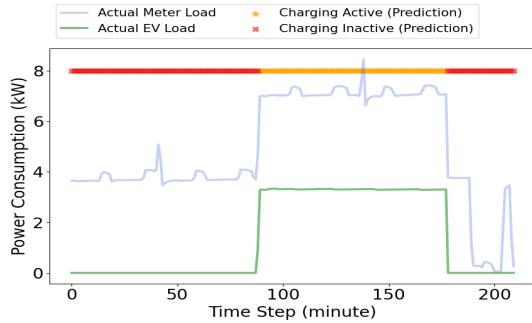


Fig. 2: EV charging event prediction for home #4526.

TABLE I: Performance comparisons between our DCT-EV method and baselines, including Random Forest, XGBoost, DNN, and LSTM. (Red: maximum and Blue: minimum.)

Prediction Span	Model Name	F1@.5	AUC	AP	ACC@.5	MSE
1-10 minutes	Random Forest	22.14%	56.50%	10.76%	98.55%	0.0144
	XGBoost	83.43%	88.34%	71.16%	98.63%	0.0137
	DNN	83.32%	97.55%	85.89%	98.45%	0.0122
	LSTM	83.91%	97.73%	86.89%	98.74%	0.0103
	DCT-EV	87.35%	97.90%	88.55%	98.80%	0.0102
1-60 minutes	DNN	49.74%	88.24%	55.20%	96.39%	0.0281
	LSTM	55.94%	88.96%	60.06%	96.68%	0.0260
	DCT-EV	63.78%	90.07%	61.59%	96.81%	0.0251

calculations, we adhere to a standard evaluation setup with a threshold set at 0.5. Accordingly, all predictions with a probability exceeding 0.5 are considered indicative of an “active” event, while those with a probability less than or equal to 0.5 are interpreted as “inactive” events.

B. Results and Discussions

DCT-EV exhibits outstanding performance, as illustrated in Figure 2, achieving high scores in AUC, AP, F1, ACC, surpassing all compared methods, and establishing a new state-of-the-art benchmark. These results highlight the robustness of the proposed method in predicting EV charging events based on historical smart meter records.

1) *Model Comparisons*: To validate the effectiveness of our method, we conduct comprehensive comparative studies with representative baseline models, including Random Forest [29], XGBoost [30], Deep Neural Networks (DNN) [31], and Long Short-Term Memory (LSTM) [32]. As depicted in Table I, it is evident that DCT-EV outperforms the other models, exhibiting higher F1 scores, AUC, AP, and ACC, as well as lower MSE.

2) *Impact of Input Length on Model Performance*: The input length can significantly impact model performance, as it influences the amount of information available to the model. As shown in Figure 3, the F1 score increases as the length of the size grows, but it plateaus when the size exceeds 180 minutes. While more information contributes to better prediction performance, it also increases computational resources needed. Hence, in this paper, we choose a length size of 180 minutes to tradeoff the performance and computational burden for our subsequent experiments.

3) *Trade-off Between Model Performance and the Size of Predictive Time Span*: Our experiments evaluate model per-

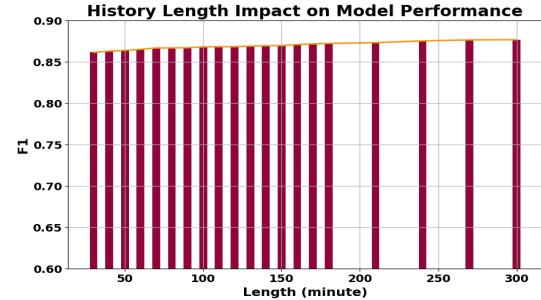


Fig. 3: The impact of input history length on the model performance for predicting EV charging event in the next 60 minutes.

TABLE II: The impact of the near-to-distant future prediction on model performance. Threshold @.5. (Red: maximum and Blue: minimum.)

m-th	AUC	AP	F1	Precision	Recall	ACC	MSE
1-min	99.63%	94.87%	90.88%	91.36%	90.40%	99.15%	0.0068
2-min	98.96%	92.67%	89.52%	90.39%	88.67%	99.03%	0.0079
3-min	98.50%	90.97%	88.70%	89.59%	87.83%	98.95%	0.0086
4-min	97.93%	89.20%	87.73%	88.93%	86.56%	98.86%	0.0094
5-min	97.30%	87.28%	86.70%	88.06%	85.38%	98.77%	0.0103
6-min	96.85%	85.55%	85.81%	87.08%	84.58%	98.68%	0.0110
7-min	96.41%	83.90%	84.79%	86.11%	83.51%	98.59%	0.0118
8-min	95.95%	82.41%	83.84%	85.25%	82.48%	98.51%	0.0125
9-min	95.46%	80.71%	82.75%	84.24%	81.31%	98.41%	0.0133
10-min	94.78%	78.75%	81.59%	83.31%	79.94%	98.31%	0.0142

formance over two different prediction time spans: next 10 minutes and next 60 minutes. Obviously, all models exhibit better performance in the next 1-10 minutes than 1-60 minutes as illustrated in table II. Due to the relatively lower performance of the two machine learning models—XGBoost and random forest in the 1-10 minutes task, we only conduct experiments for the 1-60 minutes ahead prediction for DNN, LSTM, and DCT-EV. All models’ performance drop in the 1-60 minutes task; therefore, predicting minute-interval events far into the future is indeed a more challenging task.

Furthermore, we conduct a series of experiments focused solely on predicting the state in the next m-th minute ($1 \leq m \leq 10$), as shown in Table II. The results reveal the same phenomena as discussed above.

4) *Probability Threshold Selection*: As discussed before, the model’s output is the probability of EV charging event. The selection of the probability threshold is crucial to determine whether the prediction is “active” or “inactive”. We calculate the models’ performance (e.g., F1 score) for two different prediction time spans. As depicted in Figure 4, the optimal threshold for DCT-EV falls between 0.6 and 0.8, whereas for DNN and LSTM, it lies in the range of 0.2 to 0.4. Furthermore, in comparison to DNN and LSTM, DCT-EV exhibits greater robustness to threshold selection effects.

IV. CONCLUSION

Accurate prediction of EV charging events is paramount for EV energy management and grid operations. Existing EV

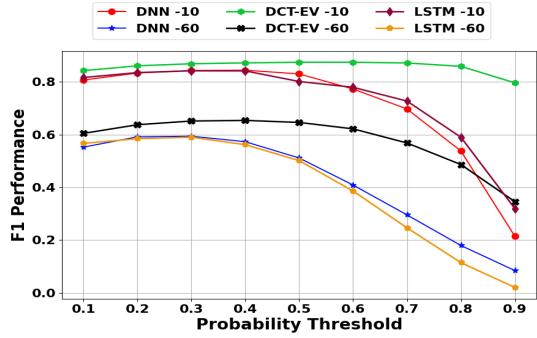


Fig. 4: Impact of probability threshold on the performance (F1 score) of DCT-EV, DNN, and LSTM for prediction tasks with 10/60 minutes ahead.

charging prediction methods, tailored to charging stations, overlooked the need for at-home EV charging predictions, especially when historical EV charging profile is not available. To address this problem, we introduced DCT-EV, a deep learning model leveraging parallel processing and self-attention mechanisms, to predict minute-by-minute future EV charging events with remarkable accuracy, exceeding 96%. Importantly, DCT-EV relies solely on historical smart meter data instead of EV charging profile, making it a valuable tool for grid management in the era of high EV adoption.

REFERENCES

- [1] Zhuowei Luo, Yonghua Song, Zechun Hu, Zhiwei Xu, Xia Yang, and Kaiqiao Zhan. Forecasting charging load of plug-in electric vehicles in china. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–8. IEEE, 2011.
- [2] Md Shariful Islam, N Mithulanthan, and Krischonme Bhumkittipich. Feasibility of pv and battery energy storage based ev charging in different charging stations. In *2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 1–6. IEEE, 2016.
- [3] Matteo Muratori. Impact of uncoordinated plug-in electric vehicle charging on residential power demand. *Nature Energy*, 3(3):193–201, 2018.
- [4] Md Murshadul Hoque, Mohsen Khorasany, Reza Razzaghi, Mahdi Jalili, and Hao Wang. Network-aware coordination of aggregated electric vehicles considering charge-discharge flexibility. *IEEE Transactions on Smart Grid*, 2022.
- [5] Jiayan Liu, Jan Peper, Gang Lin, Yang Zhou, Shrutarv Awasthi, Yong Li, and Christian Rehtanz. A planning strategy considering multiple factors for electric vehicle charging stations along german motorways. *International Journal of Electrical Power & Energy Systems*, 124:106379, 2021.
- [6] Oliver Frendo, Nadine Gaertner, and Heiner Stuckenschmidt. Improving smart charging prioritization by predicting electric vehicle departure time. *IEEE Transactions on Intelligent Transportation Systems*, 22(10):6646–6653, 2020.
- [7] Dingsong Cui, Zhenpo Wang, Peng Liu, Shuo Wang, Yiwen Zhao, and Weipeng Zhan. Stacking regression technology with event profile for electric vehicle fast charging behavior prediction. *Applied Energy*, 336:120798, 2023.
- [8] Kejun Qian, Chengke Zhou, Malcolm Allan, and Yue Yuan. Load model for prediction of electric vehicle charging demand. In *2010 International Conference on Power System Technology*, pages 1–6. IEEE, 2010.
- [9] Qiang Xing, Zhong Chen, Ziqi Zhang, Xueliang Huang, Zhaoying Leng, Kehui Sun, Yanxi Chen, and Haiwei Wang. Charging demand forecasting model for electric vehicles based on online ride-hailing trip data. *IEEE Access*, 7:137390–137409, 2019.
- [10] Mariz B Arias and Sungwoo Bae. Electric vehicle charging demand forecasting model based on big data technologies. *Applied energy*, 183:327–339, 2016.
- [11] Gan Li and Xiao-Ping Zhang. Modeling of plug-in hybrid electric vehicle charging demand in probabilistic power flow calculations. *IEEE Transactions on Smart Grid*, 3(1):492–499, 2012.
- [12] Ran Wang, Ping Wang, and Gaoxi Xiao. Two-stage mechanism for massive electric vehicle charging involving renewable energy. *IEEE Transactions on Vehicular Technology*, 65(6):4159–4171, 2016.
- [13] Lei Yu, Tianyang Zhao, Qifang Chen, and Jianhua Zhang. Centralized bi-level spatial-temporal coordination charging strategy for area electric vehicles. *CSEE Journal of Power and Energy Systems*, 1(4):74–83, 2015.
- [14] Yongquan Nie, CY Chung, and NZ Xu. System state estimation considering ev penetration with unknown behavior using quasi-newton method. *IEEE Transactions on Power Systems*, 31(6):4605–4615, 2016.
- [15] Olivier Beaude, Samson Lasaulce, Martin Hennebel, and Ibrahim Mohand-Kaci. Reducing the impact of ev charging operations on the distribution network. *IEEE Transactions on Smart Grid*, 7(6):2666–2679, 2016.
- [16] Yuchen Tang, Jin Zhong, and Math Bollen. Aggregated optimal charging and vehicle-to-grid control for electric vehicles under large electric vehicle population. *IET Generation, Transmission & Distribution*, 10(8):2012–2018, 2016.
- [17] Pengcheng You, Zaiyue Yang, Mo-Yuen Chow, and Youxian Sun. Optimal cooperative charging strategy for a smart charging station of electric vehicles. *IEEE Transactions on Power Systems*, 31(4):2946–2956, 2015.
- [18] Xudong Wang, Guoming Tang, Yi Wang, Srinivasan Keshav, and Yu Zhang. Evsense: A robust and scalable approach to non-intrusive ev charging detection. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, pages 307–319, 2022.
- [19] Shengyi Wang, Liang Du, Jin Ye, and Dongbo Zhao. A deep generative model for non-intrusive identification of ev charging profiles. *IEEE Transactions on Smart Grid*, 11(6):4916–4927, 2020.
- [20] ZhiLin Zhang, Jae Hyun Son, Ying Li, Mark Trayer, Zhouyue Pi, Dong Yoon Hwang, and Joong Ki Moon. Training-free non-intrusive load monitoring of electric vehicle charging with low sampling rate. *IECON Proceedings (Industrial Electronics Conference)*, 1:5419–5425, 2014.
- [21] Run Zhou, Yue Xiang, Yang Wang, Yuan Huang, and Shiwei Xia. Non-intrusive extraction and forecasting of residential electric vehicle charging load. In *2020 IEEE Sustainable Power and Energy Conference (iSPEC)*, pages 2141–2146. IEEE, 2020.
- [22] Cameron Martin, Fucai Ke, and Hao Wang. Non-intrusive load monitoring for feeder-level ev charging detection: Sliding window-based approaches to offline and online detection. In *2023 IEEE 7th Conference on Energy Internet and Energy System Integration (EI2)*. IEEE, 2023.
- [23] Haoyu He, Jianfei Cai, Zizheng Pan, Jing Liu, Jing Zhang, Dacheng Tao, and Bohan Zhuang. Dynamic focus-aware positional queries for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11299–11308, 2023.
- [24] Haoyu He, Jing Liu, Zizheng Pan, Jianfei Cai, Jing Zhang, Dacheng Tao, and Bohan Zhuang. Pruning self-attentions into convolutional layers in single path. *arXiv preprint arXiv:2111.11802*, 2021.
- [25] Yuqi Nie, Nam H Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. *arXiv preprint arXiv:2211.14730*, 2022.
- [26] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [27] The complete data available at. <https://www.pecanstreet.org/dataport/>. *Pecan Street: Austin, TX, USA*, Accessed in Sep. 2019.
- [28] Mohammad Hossin and Md Nasir Sulaiman. A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, 5(2):1, 2015.
- [29] Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.
- [30] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.
- [31] Fionn Murtagh. Multilayer perceptrons for classification and regression. *Neurocomputing*, 2(5–6):183–197, 1991.
- [32] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.