Clustering to Predict Electric Vehicle Behaviors using State of Charge data

Kunihiro Miyazaki
School of Engineering
The University of Tokyo
Tokyo, Japan
miyazaki@ioe.t.u-tokyo.ac.jp

Takayuki Uchiba
Sugakubunka Co., Ltd.
Tokyo, Japan
takayuki.uchiba@sugakubunka.com

Kenji Tanaka
School of Engineering
The University of Tokyo
Tokyo, Japan
tanaka@tmi.t.u-tokyo.ac.jp

***Abstract*—The concept of Vehicle to Grid (V2G) is attracting attention by its performance in improving the stability of power systems. V2G considers EVs as storage batteries as well as one option for power supply source. It is important to predict the behavior of the EV in advance in order to maintain the stability of the power supply from the EV. However, the behavior of EVs is highly random in individual vehicles, and the daily behavior patterns of individual EVs could differ from each other significantly. This makes it difficult to make a prediction model for multiple EV behaviors at once. Therefore, we propose a method to predict multiple EVs with similar behaviors. Collecting similar EV actions is expected to reduce the randomness and improve the accuracy of behavior prediction. In the experiment with a real-world State of Charge (SoC) record data, the behavior of each group is predicted with clustering techniques. We use distance-based clustering, and propose a method in which the number of clusters is interactively determined based on the analysis by using hierarchical clustering, as opposed to being automatically determined. As a result, we were able to demonstrate that the clustering method divides the whole set into several meaningful patterns. Moreover, it was confirmed that the prediction accuracy was higher when learning in groups divided by clustering compared to when using all vehicles.**

***Index Terms*—Electric Vehicle, Vehicle to Grid, State of Charge, Behavior Prediction, Clustering**

I. Introduction

Vehicle to Grid (V2G) is attracting attention by its performance in improving the stability of power systems. V2G considers EVs as storage batteries as well as one option for power supply source. It is important to predict the behavior of the EV in advance in order to maintain the stability of the power supply from the EV. If the behavior of EVs can be predicted in advance, an efficient distribution plan can be established.

The prediction of EV behavior has two major difficulties. First, there is significant randomness in individual vehicles. Human behavior always results in variance and outliers. The route and time used daily may be different, or the destination may change suddenly. If the variance is too large or involves outliers, learning of past actions will not be successful. Second, their behaviors often differ among individual cars. In particular, if the purpose is different, such as a private car or a company car, learning their behaviors with the same model cannot be expected to yield high accuracy.

Therefore, we propose a method to predict multiple EVs with similar behaviors. To identify individuals with similar patterns, we use distance-based clustering techniques. By conducting clustering as preprocessing, we expect two types of effect on the accuracy of prediction. First, by dividing the whole data into several groups with similar behaviors, we can avoid making prediction models based on EVs with significantly different behavior patterns which could lead to the confusion of prediction models. In clustering, it is said that if the whole data is well divided into several groups, the combination of clustering and supervised learning will improve the accuracy. A situation where the data is well separated can be described as when the variance within the group is greater than that of the individuals. In this study, since we can imagine several types of lifestyles related to the users of the vehicles, the data are also assumed to be divided into several patterns, and it is conceivable that the clustering associated with the supervised method will improve the accuracy. Second, by collecting the EVs of similar behaviors, we can get better accuracy of prediction compared to making models based on a single EV. Because the behavioral data of single EV are random and sparse, one can think of the difficulty in the prediction of such data with a large noise, and of addressing the problem by gathering the EVs with similar behavior to avoid the sparseness of data.

In the experiment with a real-world State of Charge (SoC) record data, the behavior of the group is predicted with clustering techniques. Prediction after clustering has already been studied in previous studies [1]–[4]. In those studies, distance-based clustering methods such as k-means are generally used, and the elbow method is used to determine the number of clusters [5], [6]. However, k-means is problematic in that the classification into clusters is sensitive to initial values. Also, elbow-method is overly subjective in addition to that the appropriate K value (number of clusters) is difficult to understand. Therefore, we propose a more reliable method based on hierarchical clustering which uses the Ward's method. The Ward's method uses the distance between clusters, which is more robust and more resistant to outliers than other methods which use the distance between representative points of

978-1-7281-7455-6/20/$31.00 ©2020 IEEE

clusters including a centroid as a criterion for connection [7]. Hierarchical clustering can determine better clusters by using an informative dendrogram. For the prediction after the cluster, we compared several popular Machine Learning methods and conducted comparative experiments while using the mean absolute error as a measure.

Given the above efforts, the result of the clustering in the end demonstrated that the whole dataset can be broken into several meaningful patterns. Moreover, it was confirmed that the prediction accuracy was higher when learning in groups divided by clustering compared to when using all vehicles. On the other hand, it was difficult to address the sparseness of the individual vehicle by clustering, and difficult to improve the accuracy of learning using only the individual vehicle over the prediction learned using the cluster, under the conditions of the study. The contribution of this research is as follows.

- We analyzed and predicted the real world EV behaviors with SoC data.
- We utilized clustering techniques for analysis of data, and demonstrated that the whole dataset can be broken into several meaningful patterns.
- We conducted one-day ahead prediction with a pre-processing of clustering, and showed the accuracy of the prediction was improved by using clustering in the EV behavior compared to when using all vehicles for prediction of individual EV behavior.

II. Related works

A. Analysis of SoC record data

As EV is a relatively new technology, the number of analysis of EVs and usage of SoC is limited, especially in terms of the analysis of EV behavior. [8] is an early one among the analyzes of EV behavior, where basic statistics are analyzed. [9] uses SoC to classify user types. Also, [10] and [11] uses various Machine Learning models to predict EV load.

B. Clustering

Clustering is one of the basic methods of unsupervised Machine Learning, and one of its purposes is to reduce preprocessing noise in supervised learning. It is a basic method to perform supervised learning such as linear regression for each of the divided models once by k-means [1]. In addition to prediction by regression, it is also used for recommender system [2]. Also, in the energy field, there are studies that predict energy consumption after clustering [3].

However, as far as we know, no research has yet been done using a combination of clustering and supervised learning on EV SoC data to predict EV behaviors. Some researches used clustering on SoC data, but these researches focus on SoC estimation, not on EV behavior [12]–[14].

III. Dataset

A. State of Charge data

In this research, we utilized a real-world SoC data. SoC is the charging rate in EV, which is represented by a numerical value from 0 to 100%, similarly to the charging rate of a mobile phone. In this study, the time scale of data is in one-minute units from 0:00 to 23:59. The data period is about one year from April 1, 2013, to March 31, 2014. The number of EVs is 446. We took great care in anonymity when treating data.

B. Extraction of actions from data

The kinds of EV actions are basically "Running," "Stop" and "Charging." "Charging" is performed by connecting to a charging station, and it takes a certain amount of time to refill, unlike gasoline refueling. It can be observed that the SoC rises during the charging time. The data currently acquired are raw data of SoC, and it is necessary to judge "Running", "Stop", and "Charging" before analysis and prediction of behaviors. In this research, from the result of observing the data, "Running" and "Charging" were defined as follows, and "Stop" is defined as other times from "Running" and "Charging."

- Running: The time where SoC keeps decreasing by more than 0.1% per minute, and the decrease continues for more than 5 minutes
- Charging: The time where SoC keeps increasing by more than 0.1% per minute, and the increase continues for more than 5 minutes

IV. Method

In this section, we introduce the methods used in our experiment and analysis.

A. Data Structure

In this research, the specific task is to predict when and how long the EV will run the next day, which is a realistic scenario even in practical cases. Regarding the data units, the following two factors are adopted in consideration of the actual driving pattern.

- Time range: every two hours from 6 am to 8 pm
- Day of the week: every date

The length of the run is measured by the time of running in the 2-hour section described above. The unit is minute.

B. Clustering

In this study, we perform clustering during preprocessing to improve prediction accuracy. The clustering method uses hierarchical clustering using the Ward's method. A cluster is selected based on having the lowest cost represented by the following equation when a certain cluster selects a cluster to be combined. When we consider combining two clusters C_1 and C_2, the cost of combination is

$$Cost(C_1, C_2) = \mathrm{E}(C_1 \cup C_2) - \mathrm{E}(C_1) - \mathrm{E}(C_2) \quad (1)$$

where $\mathrm{E}(C_1)$ is the distance from the centroid of cluster to each data points in the cluster. Supporsing $\mathbf{c}_i$ is a centroid of C_i, $\mathbf{x}$ is a data point of $\mathbf{C_i}$, and d is an euclidean distance, $\mathrm{E}(C_1)$ described as

$$\mathrm{E}(C_1) = \sum_{\mathbf{x} \in C_i} d(\mathbf{x}, \mathbf{c}_i)^2. \quad (2)$$

The Ward's method uses the distance between clusters as an index, which is more robust and more resistant to outliers than other methods which use the distance between representative points of clusters including a centroid as a criterion for connection. Since the cost is always non-negative, a monotone dendrogram can be created when used for agglomerative hierarchical clustering. This makes it difficult to produce a figure with low interpretability when visualized.

C. Visualization method

In the study, since it is hierarchical clustering, visualization by dendrogram is possible. The dendrogram is easy to visulaize as each data point and each class is combined into a new cluster. In particular, it is informative that the connection cost of a cluster and the number of data of each cluster are apparent. By selecting an appropriate threshold value for the dendrogram, the number of clusters can be selected intuitively and interactively. The easiness of selecting the appropriate number of clusters is important especially when one does not have a priori knowledge of clusters and wants to explore the clusters in various levels of granularity. K-means and elbow-method are often used for clustering and cluster number determination, but k-means is vulnerable to outliers. Also, the determination of the number of clusters in the elbow-method is unstable, and the number of data points in each cluster is difficult to intuitively understand.

D. Prediction

In this study, we predict the behavior of the next day using the methods of Simple Model and Machine Learning Models.

In the Simple Model, the average value in the past window is simply used as the prediction value. As will be described later, this time, two methods are used, one is to computate the average of individual cars and the other is to map the average of all data in all cars or clusters.

In Machine Learning models, the following three models are used. These are the basic methods of Machine Learning. We used scikit-learn, the python library, as the prediction toolkit.

- Linear Regression with L2 Regularization (LR)
- Random Forest Regrasssion (RFR)
- Support Vector Regression (SVR)

V. Cluster analysis

In the analysis by clustering, data accumulated over a period of 40 days starting from April 1, 2013 was used. As the feature value of the data, the mean value of the running time in each day-of-the-week x time range was used.

The Fig. 1 shows a dendrogram of hierarchical clustering by the Ward's method. The vertical axis represents the connection cost, and the horizontal axis represents the data points The dotted line is an example of a threshold when determining the number of clusters. In the figure, it is set to 300. Colors are displayed according to the clusters determined by the thresholds. As described above, by using the dendrogram, it is possible to intuitively determine a cluster. Currently, the threshold value is determined so as to yield four clusters,

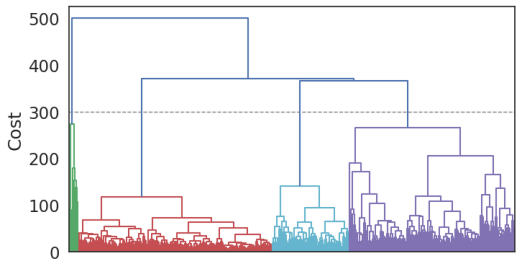


Fig. 1. Dendrogram of EV behaviors.

whereas it is also possible to set the threshold value deeper to increase the number of clusters.

We analyzed the four clusters divided by the method described above. In the following cluster numbers, the green, red, blue, and purple clusters from the left in the dendrogram are 1, 2, 3, and 4, respectively. The number of data points are 9, 194, 77, and 166, respectively.

The Fig. 2 shows the mean value of the running time in each day of the week x time range (data section) in each cluster with a heatmap. In the heatmap, the color depth is proportional to the mean of running time annotated in each box. Cluster 1 has a very large running time, cluster 2 has a small overall travel, cluster 3 has a greater amount of weekday travel than other time ranges, and cluster 4 has a moderate amount overall.

Further, the following Table I shows the characteristics of each cluster by comparing with the entire set of data. In the table, Z-score is calculated for each data section, and the top 10 of the absolute values are shown. When calculating the Z-score, the mean value (μ) and the standard deviation (σ) of each section in the entire data are calculated, then the observed value x is subtracted by the mean value and divided by the standard deviation as follows:

$$Z\text{-}score = \frac{x - \mu}{\sigma}. \tag{3}$$

The larger this value is, the larger the deviation of the data section from the mean of the whole, indicating that it is a more characteristic feature.

This table shows that cluster 1 has a particularly large running time around the weekend, such as Thursday night or Friday morning and Monday morning. It can be assumed by seeing overall running time that the car is in a company use. The most likely scenario is visiting business destinations before and after the weekend compared to the center of the week when work is generally heavy. In addition to that, because there is a lot of running time on the weekend, it can be assumed that the company usually has a heavy labor load. Looking at the cluster 2, it is assumed that the cluster is a privately-owned car that has a very small running time even during the daytime on weekdays, and is a cluster in which cars are not used very frequently even on weekends. In cluster 3, the running time is clearly increased in the morning and the afternoon hours on weekdays, and it is assumed that the vehicles are company cars which are used more moderately

compared to cluster 1. Cluster 4 has a large running time on the weekend and a large running time in the morning and evening on weekdays. It is assumed that cluster 4 is private cars which are also used for commuting.

VI. Experiment

A. Prediction Framework

In this study, we conduct one-day ahead prediction, where the running time in each time range of the next day is predicted for each car at the previous day. The term length used for input of forecast is 40 days. The prediction is repeated for 19 days, and the average value is used as the result of the prediction. The model is trained for 60 days up to the day before the prediction starts, and the input of training is also 40 days. This learning is performed for a total of four periods shown in Table II.

Number of clustering is performed on 4 and 7. Also, we try two types of input for the Machine Learning model. One is a method to pick up only the same data section as the prediction target in the past 40 days (SVR_1, RFR_1, LR_1). The other is all data points for the past 40 days including data sections with different time ranges and day of the weeks from targeted data sections (SVR_2, RFR_2, LR_2). The former has the same condition as the prediction using the historical mean.

B. Benchmark

As described in Method Section, we will use two benchmarks. Benchmark1 is a value representing the average value of historical values of each car (Ave_indiv). Benchmark2 is the average value of historical values of all vehicles or all EVs in the cluster mapped to individual vehicles (Ave_all).

C. Accuracy index

We used mean absolute error (MAE) as an accuracy index. MAE is calculated by averaging the absolute value of the difference between the ground-truth values and the predicted values. Smaller MAE indicates better accuracy.

D. Results

Table III, IV and V show the results of the experiment with no clustering, four clusters, and seven clusters, respectively. The row shows each prediction method, and the column shows the prediction period and their average. In each table, the column with the highest precision in each column is shown in bold. For SVR and LR, the output below 0 was fixed to 0, but the result was not largely affected.

Looking at the tables, the one with the highest accuracy in each table was Ave_indiv. The next most accurate is SVR_1. Accuracy generally increases as the number of clusters increases. Accuracy may vary significantly at different terms. In particular, since term3 extends over the year-end and new-year period, it seems that the learning has been greatly affected by its irregularity. LR2 gives a very large error in term2, probably because of the outlier.

Fig. 3 shows the accuracy of each cluster in SVR_1 in clusters 4 and 7. From this, it can be seen that the error also increases as the average running time increases. The reason why the accuracy increases as the number of clusters is increased is considered to be that the decomposition of these clusters having a large running time and to separate them during learning lead to an improvement in accuracy.

VII. Discussion

A. Effect of clustering

In the study, clustering has two meanings in terms of the impact for prediction accuracy. One is that it is more efficient to learn in several groups than to learn all data at once during learning. This was confirmed as the accuracy increased as the number of clusters increased.

The other is that it seems more efficient to learn together in groups rather than individual data. This was expected to confirm by observing that Machine Learning with clustering would have higher prediction accuracy than individual historical means, but this time we were not able to confirm it. There are three possible reasons. First, a behavior of an EV is different too much from other EVs, so the combining them did not lead to the improvement of accuracy. Second, because the mesh of data is extremely fine, the data are too sparse to be fulfilled by combining data with clustering. On the other hand, the reason why the SVR is coming close to the score of historical mean as the number of clusters increases is probably that EVs with less sparse data (having a large running time) could be learned together in some clusters. Third, terms might be too short to learn. In this case, 40 days of input were considered for practical use, but there are actually 5 or 6 of them whose day of the week and time range completely match the those of target data section. In this case, it is conceivable that the number of data is insufficient for Machine Learning and accuracy is reduced. A longer period of time could improve learning accuracy.

B. Comparison of models

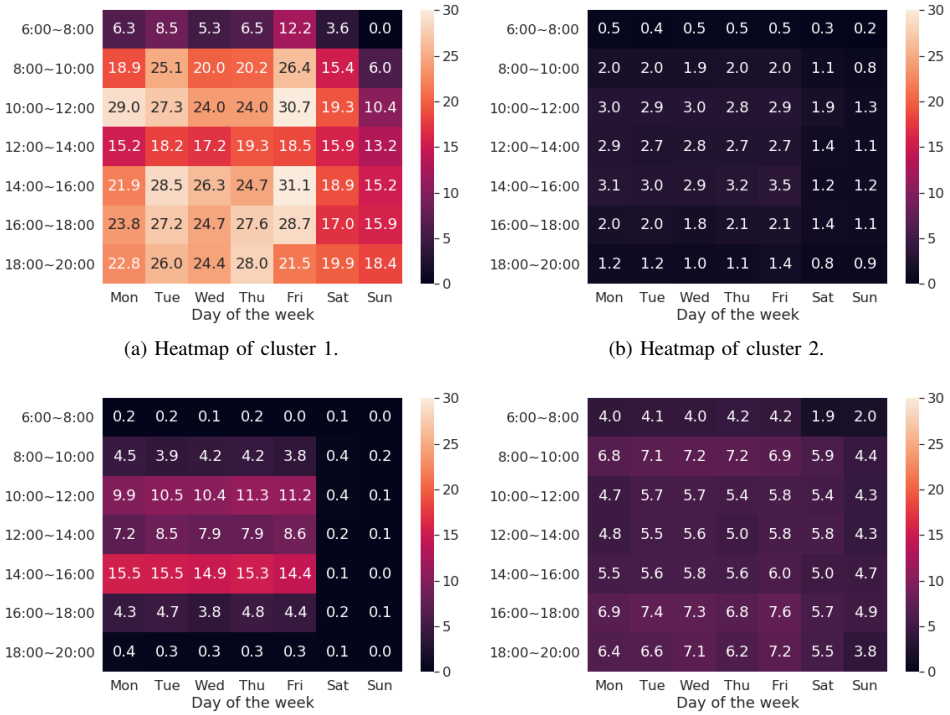
Among Machine Learning models, SVR was better this time. The reason of this result is considered to be attributed to the rubustness to outliers of each model. SVR is resistant to outliers. That is because, in SVR, predicted values will be calculated by only support vectors not by all data points unlike other model including LR and RFR. Moreover, from the same reason, SVR can learn well even when sample size is small. These features of SVR are considered to contribute to its better result. On the other hand, RFR and LR is vulnerable to outliers because they use squared error when calculating predicted values.

VIII. Conclusion

In this study, we analyzed the behavior of EV and performed one-day ahead prediction using clustering technology.

Given the above efforts, the result of the clustering demonstrated that the whole dataset can be broken into several meaningful patterns.

Moreover, it was confirmed that the prediction accuracy was higher when learning in groups divided by clustering compared to when using all vehicles.



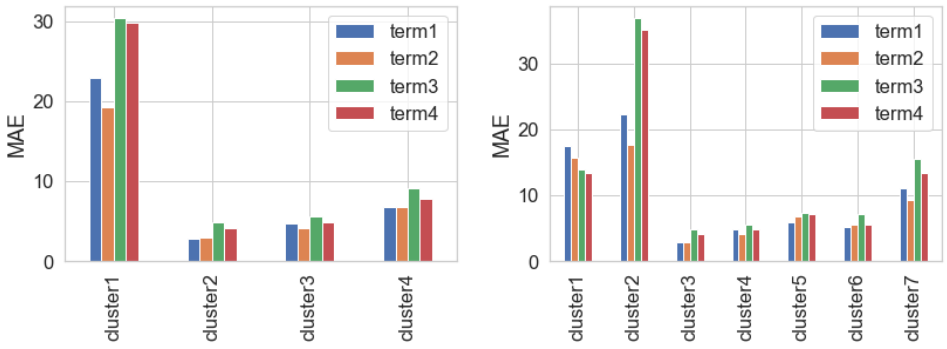
(a) Heatmap of cluster 1.

(b) Heatmap of cluster 2.

(c) Heatmap of cluster 3.

(d) Heatmap of cluster 4.

Fig. 2. Heatmap of the mean values of the running time in each data section in each cluster.



(a) Cluster num=4.

(b) Cluster num=7.

Fig. 3. Accuracy of each cluster.

REFERENCES

[1] S. M. Mostafa and H. Amano, "Effect of clustering data in improving machine learning model accuracy," *Journal of Theoretical and Applied Information Technology*, vol. 97, no. 21, pp. 2973–2981, 2019.

[2] T. K. Quan, I. Fuyuki, and H. Shinichi, "Improving accuracy of recommender system by clustering items based on stability of user similarity," in *2006 International Conference on Computational Inteligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06)*. IEEE, 2006, pp. 61–61.

[3] D. Hsu, "Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data," *Applied energy*, vol. 160, pp. 153–163, 2015.

[4] C. Yang, X. Shi, L. Jie, and J. Han, "I know you'll be back: Interpretable new user clustering and churn prediction on a mobile social application," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 914–922.

[5] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14. Oakland, CA, USA, 1967, pp. 281–297.

[6] T. M. Kodinariya and P. R. Makwana, "Review on determining number of cluster in k-means clustering," *International Journal*, vol. 1, no. 6, pp. 90–95, 2013.

[7] F. Murtagh and P. Legendre, "Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion?" *Journal of classification*, vol. 31, no. 3, pp. 274–295, 2014.

[8] J. Smart and S. Schey, "Battery electric vehicle driving and charging behavior observed early in the ev project," *SAE International Journal of Alternative Powertrains*, vol. 1, no. 1, pp. 27–33, 2012.

[9] C.-H. Lee and C.-H. Wu, "Learning to recognize driving patterns for

TABLE I
OUTSTANDING FEATURES OF EACH CLUSTER

dow_time	z_score	mean_cluster	mean_all
Thu_18:00	3.69	28.02	3.40
Fri_08:00	3.61	26.40	4.63
Mon_10:00	3.58	28.96	5.34
Fri_16:00	3.55	28.69	5.09
Fri_10:00	3.54	30.67	6.00
Thu_16:00	3.51	27.64	4.81
Tue_16:00	3.44	27.20	4.99
Tue_08:00	3.40	25.07	4.67
Tue_10:00	3.30	27.27	5.77
Sun_18:00	3.25	18.36	2.15

(a) Cluster 1.

dow_time	z_score	mean_cluster	mean_all
Wed_14:00	0.49	2.93	6.51
Thu_08:00	0.47	2.03	4.72
Tue_14:00	0.47	3.02	6.63
Wed_08:00	0.46	1.93	4.65
Tue_16:00	0.46	2.01	4.99
Mon_14:00	0.46	3.11	6.51
Thu_14:00	0.46	3.16	6.58
Mon_08:00	0.45	1.97	4.54
Tue_08:00	0.45	1.97	4.67
Wed_16:00	0.45	1.76	4.65

(b) Cluster 2.

dow_time	z_score	mean_cluster	mean_all
Mon_14:00	1.20	15.48	6.51
Thu_14:00	1.16	15.26	6.58
Tue_14:00	1.14	15.45	6.63
Wed_14:00	1.13	14.86	6.51
Fri_14:00	0.99	14.36	6.88
Thu_10:00	0.81	11.28	5.69
Fri_10:00	0.74	11.18	6.00
Tue_10:00	0.73	10.54	5.77
Wed_10:00	0.71	10.38	5.71
Mon_10:00	0.69	9.88	5.34

(c) Cluster 3.

dow_time	z_score	mean_cluster	mean_all
Sun_08:00	0.58	4.38	2.13
Sat_08:00	0.54	5.89	3.06
Wed_06:00	0.53	4.03	1.84
Sat_12:00	0.52	5.79	3.10
Sat_16:00	0.51	5.66	3.09
Sun_10:00	0.48	4.27	2.36
Wed_18:00	0.47	7.07	3.59
Fri_18:00	0.47	7.23	3.80
Mon_06:00	0.47	3.95	1.84
Sun_16:00	0.47	4.94	2.66

(d) Cluster 4.

TABLE II
TERMS FOR PREDICTION

	Train term	Test term
term1	2013-04-01 - 2013-06-04	2013-05-16 - 2013-07-14
term2	2013-06-25 - 2013-08-24	2013-08-05 - 2013-10-05
term3	2013-09-14 - 2013-11-17	2013-10-27 - 2013-12-29
term4	2013-12-09 - 2014-02-09	2014-01-21 - 2014-03-23

TABLE III
RESULTS FOR ALL DATA

	term1	term2	term3	term4	average
Ave_indiv	**4.96**	**4.88**	**6.29**	**5.57**	**5.43**
Ave_all	6.97	7.13	9.95	9.05	8.28
SVR_1	5.14	5.00	7.24	6.25	5.91
SVR_2	5.43	5.21	7.60	6.52	6.19
RFR_1	5.51	5.41	7.24	6.21	6.09
RFR_2	5.53	5.56	6.98	5.95	6.00
LR_1	5.47	5.23	6.99	6.03	5.93
LR_2	6.21	6.55	7.92	6.79	6.87

TABLE IV
RESULTS FOR 4 CLUSTERS

	term1	term2	term3	term4	average
Ave_indiv	**4.96**	**4.88**	**6.29**	**5.57**	**5.43**
Ave_all	6.26	6.49	9.18	8.34	7.56
SVR_1	5.03	4.91	7.12	6.17	5.81
SVR_2	5.47	5.27	7.20	6.08	6.01
RFR_1	5.47	5.39	7.20	6.14	6.05
RFR_2	5.63	5.59	6.96	5.97	6.04
LR_1	5.47	5.27	7.20	6.08	6.01
LR_2	12.93	1.09E+12	11.07	12.12	2.73E+11

TABLE V
RESULTS FOR 7 CLUSTERS

	term1	term2	term3	term4	average
Ave_indiv	**4.96**	4.88	**6.29**	**5.57**	**5.43**
Ave_all	6.03	6.32	8.96	8.15	7.36
SVR_1	**4.96**	**4.86**	7.08	6.11	5.75
SVR_2	5.16	5.02	7.31	6.31	5.95
RFR_1	5.47	5.44	7.27	6.17	6.09
RFR_2	5.66	5.58	7.02	5.99	6.06
LR_1	5.58	5.38	7.47	6.22	6.16
LR_2	12.94	1.09E+12	11.11	12.09	2.73E+11

collectively characterizing electric vehicle driving behaviors," *International Journal of Automotive Technology*, vol. 20, no. 6, pp. 1263–1276, 2019.

[10] S. Xydas, C. Marmaras, L. M. Cipcigan, A. Hassan, and N. Jenkins, "Electric vehicle load forecasting using data mining methods," 2013.

[11] Y.-W. Chung, B. Khaki, T. Li, C. Chu, and R. Gadh, "Ensemble machine learning-based algorithm for electric vehicle user behavior prediction," *Applied Energy*, vol. 254, p. 113732, 2019.

[12] X. Hu, S. E. Li, and Y. Yang, "Advanced machine learning approach for lithium-ion battery state estimation in electric vehicles," *IEEE Transactions on Transportation electrification*, vol. 2, no. 2, pp. 140–149, 2015.

[13] Y. Wei *et al.*, "A novel combined data mining algorithm for state-of-charge estimation in electric vehicle," *International Journal of Advancements in Computing Technology*, vol. 4, p. 662, 2012.

[14] T. Zahid, K. Xu, W. Li, C. Li, and H. Li, "State of charge estimation for electric vehicle power battery using advanced machine learning algorithm under diversified drive cycles," *Energy*, vol. 162, pp. 871–882, 2018.