Residential EV Charging Characterization via Forecasting, Vehicle Classification, and Behavior Identification

Antonio Aguilar, Justin Luke, Robert Spragg | {antonioa, jthluke, spragg}@stanford.edu

CS 229: Machine Learning, Autumn 2018

INTRODUCTION & MOTIVATION

Electric Vehicles (EV) can serve as mobile energy storage assets for the electricity grid by performing demand response, participating in ancillary services, and supporting renewable energy integration [1]. This can be achieved by optimally controlled "smart charging" of the EVs. While the flexibility for EVs to provide grid services has been studied in the public and commercial setting, representation of EV charging in residential environments is less understood [2].

Our project aims to characterize residential EVs by forecasting charging load, classifying car models, and identifying patterns in user behavior. These findings can provide insight for utilities and policymakers when selecting distribution infrastructure investments and implementing customer demand response programs.

DATA & FEATURES

The dataset used is from the Pecan Street PECAN STREET **research project in Austin, TX. The data consist of one-minute resolution, sub-metered power load data from 67 homes with EVs. For some homes, EV car models were labeled. We first developed a heuristic to identify EV charging sessions from raw data. For each session, we derived: session energy [kWh], time variables (e.g. day of week, # hour of day), session length, heuristically identified tails, quantiles of tail power and duration. For each home, we heuristically estimated their car's max charging rate and max battery capacity. Lastly, power data was aggregated across all users, producing daily power curves, averaged across one week. These derived features are utilized as inputs to our models.**

DISCUSSION & RESULTS

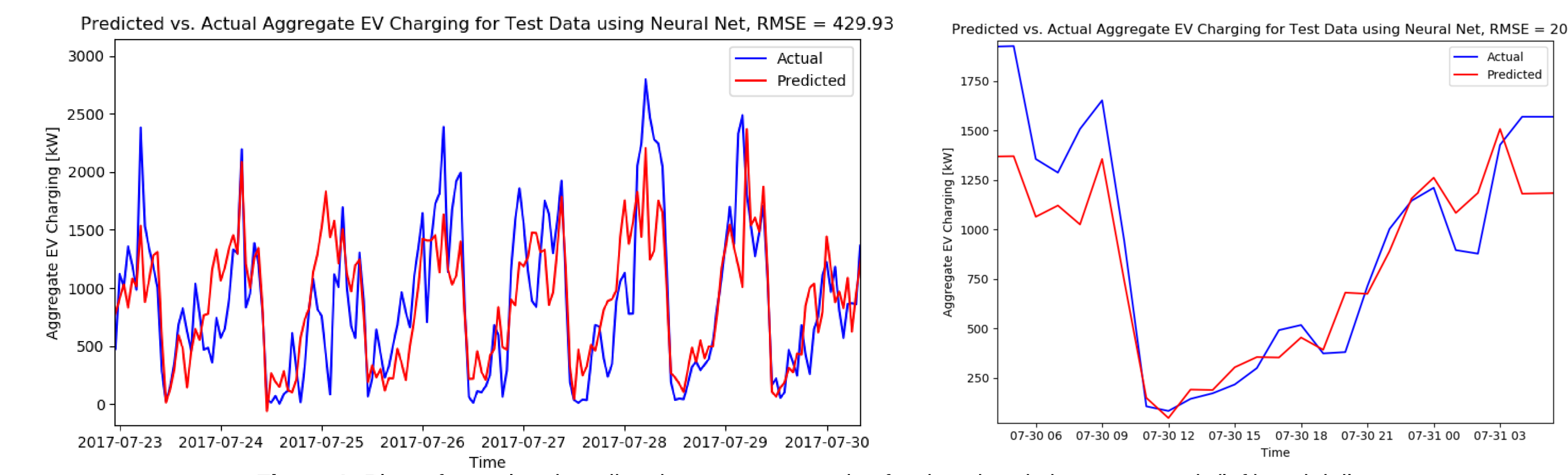

Figure 1: Plots of actual and predicted power consumption for day ahead given past week (left) and daily average of week ahead given past four weeks' daily averages (right). The model adapts well to variations on subsequent days and weeks.

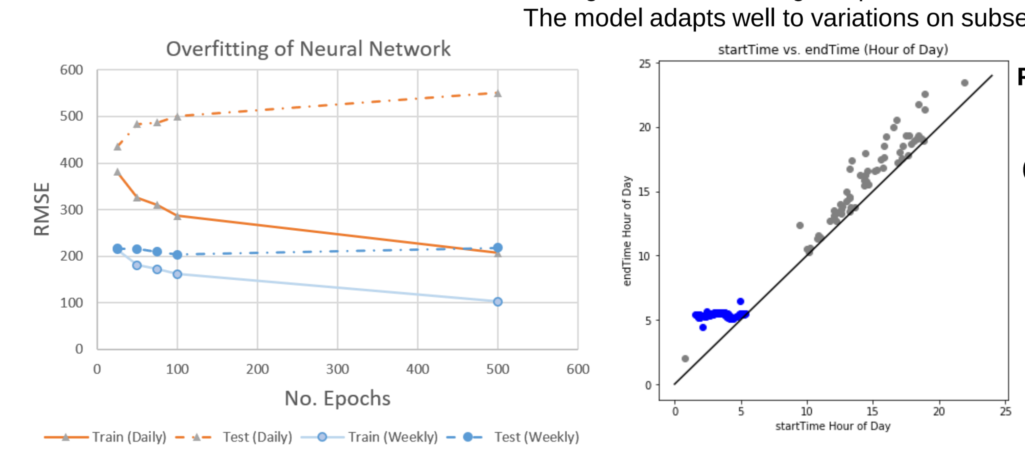

Figure 2: (left) – Demonstration of overfitting on train and test sets. The daily predictions were subject to higher variance than weekly predictions due to higher uncertainty. (right) – Demonstration of identification of programmable charger using density-based spatial clustering of applications with noise (DBSCAN)

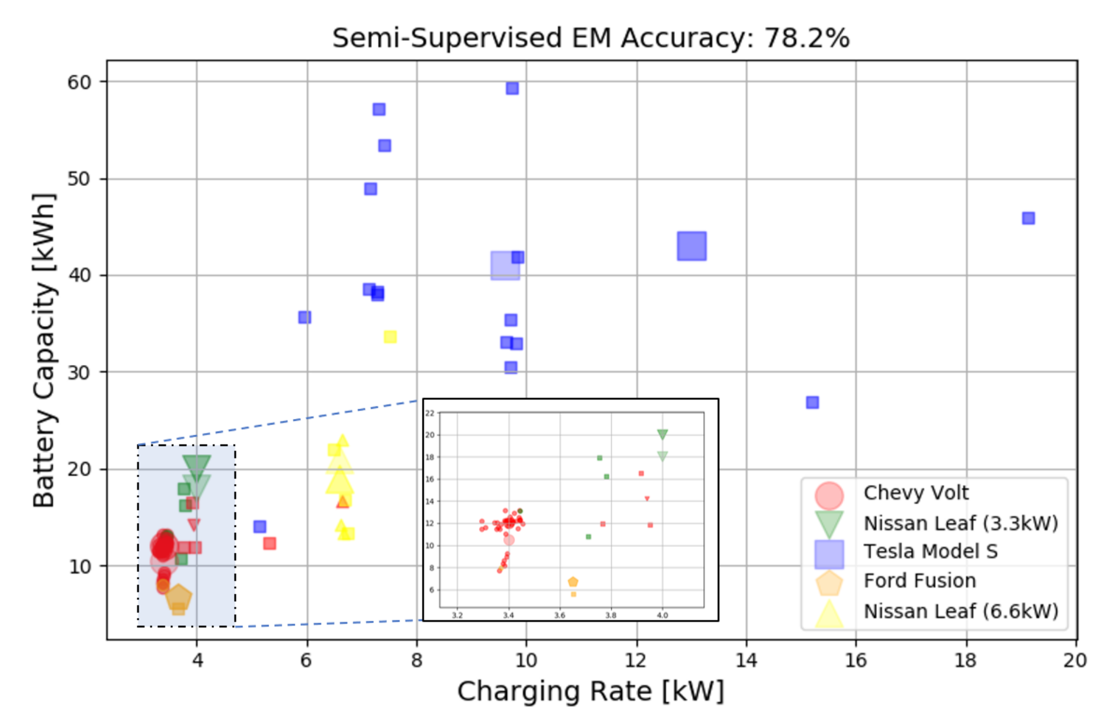

Figure 3: Plot of vehicle model classification. Semi-supervised GMM EM algorithm utilized charging rate and estimated battery capacity features extracted from original time series.

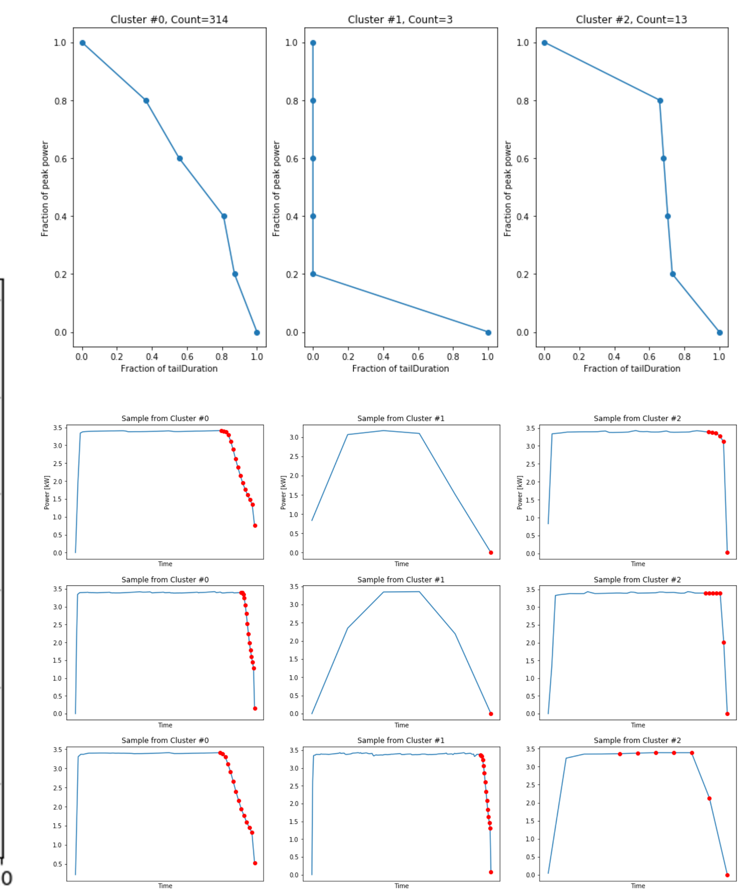

Figure 4: Demonstration of charging tail identification using K-means clustering. The fraction of sessions that end abruptly (#1, 2) provides insight into an EV's load flexibility.

METHOD	TRAIN RMSE (Daily, Weekly) m=(476, 65)	TEST RMSE (Daily, Weekly) m=(244, 34)
Linear Regression	(304.31, 7.15)	(496.44, 437.43)
SVM Regression	(355.82, 109.19)	(429.88, 235.67)
Elastic Net	(322.89, 4.82)	(464.15, 441.32)
Kernelized Ridge Regression	(320.74, 0.12)	(460.51, 370.88)
Feedforward NN	(302.20, 166.42)	(435.06, 212.53)

A feedforward fully-connected neural network resulted in the lowest prediction error. The neural network is able to uncover charging patterns that occur during weekdays, weekends, and holidays. Hyperparameter tuning and regularization was able to limit overfitting of the irregular load data. While our model performed well on aggregate power data, we did not accurately predict the charging schedule at the individual EV level. This is a challenging problem due to the high variance of charging behavior for individual users, as well as the relative sparseness of charging sessions.

MODELS

A feedforward neural network was the top performing estimator among several regression estimators used for forecasting EV charging. Neural network structure was identified using grid search across layers, units, L2-regularization term, and learning rate. ReLU activation was used in each layer. A semi-supervised Gaussian Mixture Model (GMM) was used to classify EV car models. k-Means clustering was used to identify typical behavior patterns among users.

Semi-supervised GMM: For samples $x^{(i)} \in \mathbb{R}^2\ \forall i = 1, \dots, m$, labeled samples $\tilde{x}^{(i)} \in \mathbb{R}^2\ \forall i = 1, \dots, \tilde{m}$, and latent labels $\tilde{z}^{(i)} = j\ \forall j = 1, \dots, k$:

Semi-supervised GMM E-step:

$$w_j^{(i)} = p\left(z^{(i)} = j \middle| x^{(i)}; \theta\right) = \frac{|\Sigma_j|^{-0.5} \exp\left(-\frac{1}{2}\left(x^{(i)}-\mu_j\right)^T \Sigma_j^{-1}\left(x^{(i)}-\mu_j\right)\right)\phi_j}{\sum_{l=1}^{k} |\Sigma_l|^{-0.5} \exp\left(-\frac{1}{2}\left(x^{(i)}-\mu_l\right)^T \Sigma_l^{-1}\left(x^{(i)}-\mu_l\right)\right)\phi_l}$$

Semi-supervised GMM M-step:

$$\mu_j \leftarrow \frac{\sum_{i=1}^{m} w_j^{(i)} x^{(i)} + \sum_{i=1}^{\tilde{m}} \alpha \tilde{x}^{(i)} 1\{\tilde{z}^{(i)} = j\}}{\sum_{i=1}^{m} w_j^{(i)} + \sum_{i=1}^{\tilde{m}} \alpha 1\{\tilde{z}^{(i)} = j\}}, \quad \phi_j = \frac{\left(\sum_{i=1}^{\tilde{m}} \alpha 1\{\tilde{z}^{(i)} = j\} + \sum_{i=1}^{m} w_j^{(i)}\right)}{\alpha \tilde{m} + m}$$

$$\Sigma_j \leftarrow \frac{\sum_{i=1}^{m} w_j^{(i)}\left(\left(\tilde{x}^{(i)}-\mu_j\right)\left(\tilde{x}^{(i)}-\mu_j\right)^T\right) + \alpha \sum_{i=1}^{\tilde{m}}\left(\left(\tilde{x}^{(i)}-\mu_j\right)\left(\tilde{x}^{(i)}-\mu_j\right)^T\right) 1\{\tilde{z}^{(i)} = j\}}{\sum_{i=1}^{m} w_j^{(i)} + \sum_{i=1}^{\tilde{m}} \alpha 1\{\tilde{z}^{(i)} = j\}}$$

k-Means: For cluster labels $c^{(i)} \in \mathbb{R}\ \forall i = 1, \dots, m$, data $x^{(i)} \in \mathbb{R}^n\ \forall i = 1, \dots, m$, and cluster centroids $\mu_j \in \mathbb{R}^n\ \forall j = 1, \dots, k$.

1) Randomly initialize cluster centroids:

$$\mu_j, \forall j = 1, \dots, k$$

2) Update and repeat until convergence:

$$c^{(i)} \leftarrow \arg\min_j \left\|x^{(i)} - \mu_j\right\|^2\ \forall i = 1, \dots, m,$$

$$\mu_j \leftarrow \frac{\sum_{i=1}^{m} 1\{c^{(i)} = j\} x^{(i)}}{\sum_{i=1}^{m} 1\{c^{(i)} = j\}}\ \forall j = 1, \dots, k$$

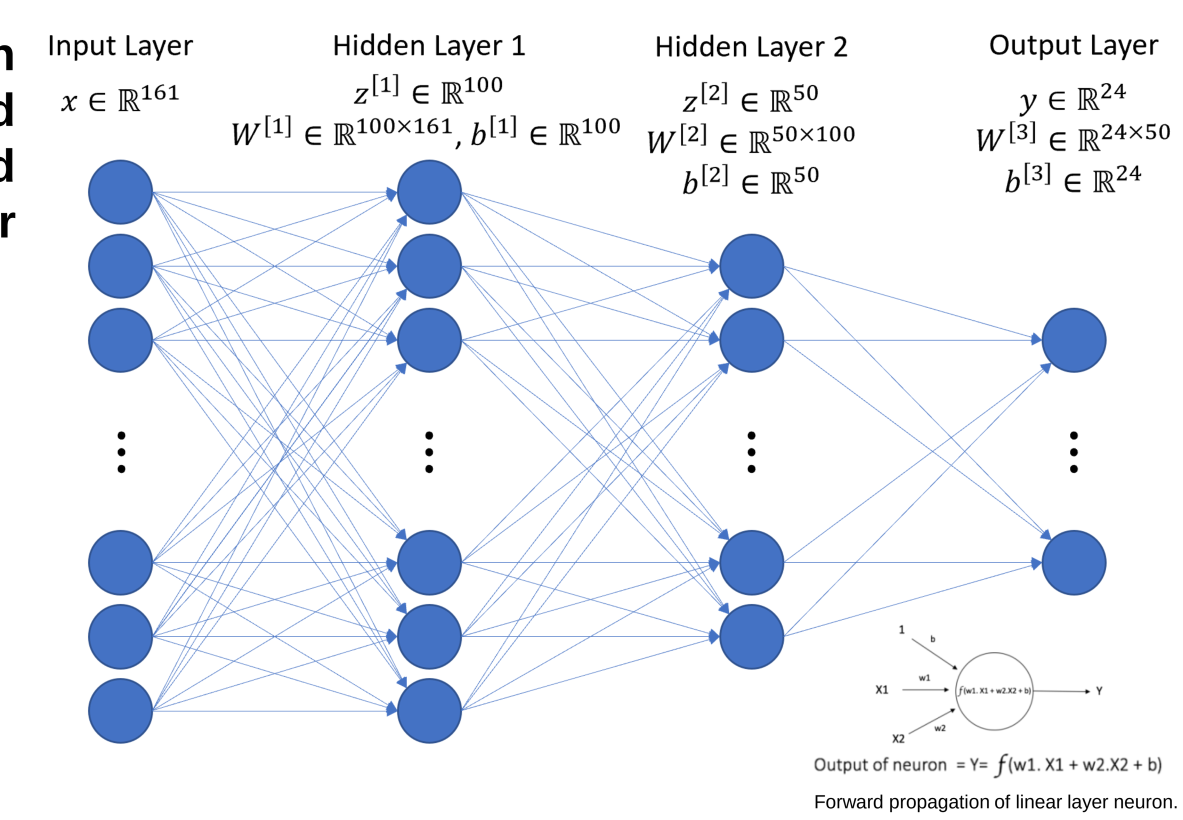

Figure 5: Fully connected neural network with two hidden layers used to predict daily EV charging demand given day of week, month, and past charging demand. ReLU activation was used in each layer

CONCLUSION & FUTURE WORK

We have developed a neural network that can forecast total load 51% more accurately than baseline linear regression on daily EV charging load averaged over one week. Furthermore, we are able to classify an EV model from its charging history with 78% accuracy. The range in typical EV charging load throughout the day is approximately 95% of the peak load. Upon inspection of utility costs and typical transformer layout, we estimate that smart charging of EV loads could save utilities $800 per EV in grid improvement costs through distribution upgrade deferral.

Future work will involve estimating power demand for individual EVs via utilization of an LSTM model. We hope to then further quantify load flexibility of each vehicle, and expand our scope to include ChargePoint data from the entire Bay Area.

REFERENCES

[1] Sadeghianpourhamami, N. et al. (2018). Quantitative analysis of electric vehicle flexibility: A data-driven approach. *Electrical Power and Energy Systems, 95, 451-462*

[2] Schuller, A., Flath, C., Gottwalt, S. (2015). Quantifying load flexibility of electric vehicles for renewable energy integration. *Applied Energy, 151, 335-344*

ACKNOWLEDGMENTS

We would like to acknowledge Dr. Yuting Ji (CEE), Professor Ng, Professor Dror and the CS229 teaching staff for providing mentorship and support.