# Valuation and Profit Allocation for Electric Vehicle Battery Data in a Data Market

**Junkang Chen, Guannan He**

**Department of Industrial Engineering and Management, College of Engineering, Peking University**

## Introduction

The global shift towards sustainable transportation has catalyzed a paradigmatic transformation in the automotive industry, with Electric Vehicles (EVs) at its forefront. In the face of fast-developing electric mobility, the safety of EV batteries becomes more and more important, driving the need for robust anomaly detection models, further driving the demand for existing EV battery data. Especially, for newly founded EV companies which lack extensive data, data markets offer a solution.

In such a data market, data valuation is of great importance in data transaction, leading to efficiency and fairness in data market. Although there are plenty of works about data valuation in machine learning tasks, to the best of our knowledge, few works have focused on the specific scenario of EV battery, especially on the task of anomaly detection. To fill this gap, we adapt EV battery anomaly detection methods to existing data valuation methods under a data trading market. Further, to stimulate data trading and incentivize data sellers to provide higher quality data, we propose revenue allocation schemes to allocate part of data buyer's revenue.

## Methods

In our study, we consider a data trading market including three kinds of agents: data buyers, data sellers and data trading platforms. A data buyer purchase data from data sellers through the data platform according to their specific needs. Each data transaction involves a single buyer and multiple sellers, and the transaction process is shown in Figure 1.


Figure 1: The framework of the EV battery data trading and profit allocating.

As for profit allocation, we propose 3 groups of profit allocation schemes, each corresponding to a data valuation method, as listed in Table 1. Data valuation is based on DyAD, an anomaly detection method based on Dynamical Variational Autoencoder. The data value for a battery is given by the expected direct cost saved for the EV company after training the model with the data, tuning the threshold to achieve the least sum of inspection and fault cost.

Table 1: The introduction to proposed profit allocation schemes.

| Scheme | | Properties | Data Valuation | Revenue Allocation |
|---|---|---|---|---|
| Average Scheme (AS) | | The value for each seller is the same, thus each seller gets the same payoff. | $\phi_i = C$ | $\xi_i = \frac{\phi_i}{\sum_{i=1}^{N}\phi_i}\cdot \eta V(D)$, which, for LCS, can be simplified to: $\xi_i = \eta\phi_i$ |
| Schemes based on properties of single data | Quantity Based Scheme (QTBS) | The data value of each seller is based on the quantity, which is the number of charging segments in our case. | $\phi_i = N_i^{\text{snippet}}$ | |
| | Quality Based Scheme (QLBS) | The data value of each seller is based on the initial improvement it brings to model performance. | $\phi_i = \max\{V(\{i\}) - V(\phi), 0\}$ | |
| Schemes based on the cooperative game of all sellers | Leave One Out Scheme (LOOS) | The data value of each seller is based on the decrease of model performance when leaving it out alone. | $\phi_i = \max\{V(D) - V(D\setminus\{i\}), 0\}$ | |
| | Shapley Value Scheme (SVS) | The data value of each seller is computed by averaging the marginal improvement of model performance over subsets of data of other data sellers. | $\phi_i = \frac{1}{N}\sum_{S\subseteq D\setminus\{i\}} \frac{\max\{V(S\cup\{i\}) - V(S), 0\}}{\binom{N-1}{\lvert S\rvert}}$ | |
| | Least Core Scheme (LCS) | The data value of each seller is evaluated by solving the Least Core problem, which aims to find the payoff allocation within the core of the game that minimizes the payoff of the worst-off coalition | $\min e$ s.t. $\sum_{i\in D}\phi_i = V(D) - V(0)$, $\sum_{i\in S}\phi_i + e \geq V(S) - V(0)$, $\phi_i \geq 0$ | |

## Results

In our experiments, we choose 23 data sellers, of which 15 have data of a normal EV battery and 8 have the data of an abnormal EV battery. The data comes from real-world EV lithium batteries, where time series data are segmented into several charging snippets of equal length. The allocated profit for each seller under different allocation schemes are listed in Table 2. The heatmap of the correlation of payoffs under different schemes is presented in Figure 2.

Table 2: Payoffs for sellers under different schemes.

| | | | Revenue Allocation Schemes | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | AS | QTBS | QLBS | LOOS | SVS | LCS |
| Seller Index | Normal | 1 | 1.11 | 1.02 | 0.00 | 1.30 | 0.79 | 0.63 |
| | | 2 | 1.11 | 0.55 | 2.96 | 0.78 | 1.51 | 0.65 |
| | | 3 | 1.11 | 1.86 | 0.39 | 1.81 | 0.54 | 1.29 |
| | | 4 | 1.11 | 0.49 | 5.61 | 1.55 | 1.19 | 0.81 |
| | | 5 | 1.11 | 1.07 | 0.00 | 1.04 | 0.78 | 1.50 |
| | | 6 | 1.11 | 3.92 | 1.68 | 1.55 | 1.47 | 1.45 |
| | | 7 | 1.11 | 0.25 | 1.70 | 1.29 | 1.80 | 1.38 |
| | | 8 | 1.11 | 0.93 | 2.41 | 1.55 | 2.43 | 1.35 |
| | | 9 | 1.11 | 0.28 | 0.22 | 1.29 | 0.94 | 0.89 |
| | | 10 | 1.11 | 0.45 | 1.46 | 1.04 | 0.24 | 0.16 |
| | | 11 | 1.11 | 2.91 | 0.00 | 1.29 | 2.31 | 2.43 |
| | | 12 | 1.11 | 2.54 | 1.37 | 1.40 | 0.96 | 1.12 |
| | | 13 | 1.11 | 1.98 | 0.06 | 1.29 | 1.34 | 1.65 |
| | | 14 | 1.11 | 4.01 | 1.76 | 1.29 | 0.79 | 0.98 |
| | | 15 | 1.11 | 0.03 | 0.00 | 1.04 | 0.98 | 0.88 |
| | | Average | 1.11 | 1.49 | 1.31 | 1.30 | 1.21 | 1.14 |
| | Abnormal | 16 | 1.11 | 0.22 | 1.62 | 0.98 | 0.66 | 1.56 |
| | | 17 | 1.11 | 0.42 | 3.17 | 1.33 | 0.65 | 0.85 |
| | | 18 | 1.11 | 0.03 | 1.05 | 0.66 | 1.02 | 0.42 |
| | | 19 | 1.11 | 0.86 | 0.00 | 0.92 | 0.68 | 1.52 |
| | | 20 | 1.11 | 0.20 | 0.06 | 1.02 | 0.75 | 1.25 |
| | | 21 | 1.11 | 0.41 | 0.00 | 0.82 | 0.76 | 1.18 |
| | | 22 | 1.11 | 0.52 | 0.00 | 0.00 | 1.52 | 0.94 |
| | | 23 | 1.11 | 0.57 | 0.00 | 0.26 | 1.40 | 0.63 |
| | | Average | 1.11 | 0.40 | 0.74 | 0.75 | 0.93 | 1.04 |

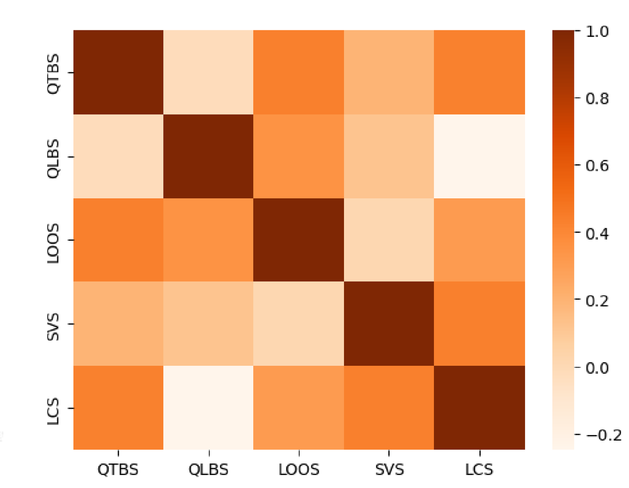


Figure 2: Correlation of payoffs under different schemes

To compare the efficacy of different schemes, we throw data away in order of value from the highest to lowest and from the lowest to highest, respectively. The results are plotted in Figure 3. It can be seen that under the cooperative game based schemes, the market is relatively fair and efficient, because the data buyer loses less without the lowly-paid sellers and loses more with the highly-paid sellers.

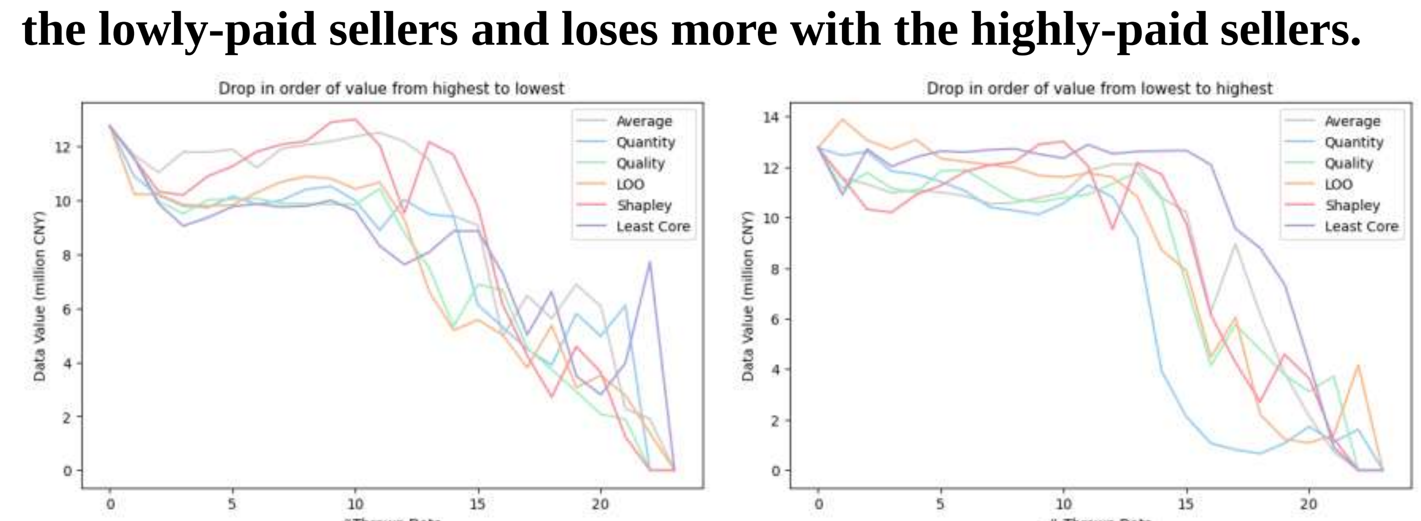


Figure 3: The total data value when throwing data away in order of value from the highest to lowest (left) and from the lowest to highest (right).

## Conclusions

In conclusion, our research delves into the realm of electric vehicle (EV) battery data trading markets, focusing on data valuation and revenue allocation. Our exploration extends to data valuation methodologies, encompassing leave-one-out, Shapley value and the least core algorithm. To explain data value from an economic perspective, we utilize a utility function considering the direct economic costs saved for the EV company to refine the evaluation process.

Based on data value, we further propose revenue allocation schemes to allocate part of EV company's revenue to data sellers. A case study is conducted based on real world EV battery dataset to illustrate how the different revenue allocation schemes allocate payoffs to data sellers.

## References

Jingzhao Zhang, Yanan Wang, Benben Jiang, Haowei He, Shaobo Huang, Chen Wang, Yang Zhang, Xuebing Han, Dongxu Guo, Guannan He, et al. Realistic fault detection of li-ion battery via dynamical deep learning. Nature Communications, 14(1):5940, 2023.

Amirata Ghorbani and James Zou. Data shapley: Equitable valuation of data for machine learning. In International conference on machine learning, pages 2242–2251. PMLR, 2019.

Tom Yan and Ariel D Procaccia. If you like shapley then you'll love the core. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 35, pages 5751–5759, 2021.