Advanced Production and Industrial Engineering
R.M. Singari and P.K. Kankar (Eds.)

doi:10.3233/ATDE220716

Machine Learning Algorithm for Sorting of Battery Packs at Smart Manufacturing Industries

ASHRAF MOHAMMAD[a], JIEH-REN CHANG[a,1] , and TIEN -TAI CHANG[c]
[a] *Department of Electronic Engineering, National Ilan University, Yilan City, Taiwan (R.O.C)*
[c] *Plant Manager, JS Power co. Ltd, Zhunan City, Taiwan (R.O.C)*

Abstract. Energy conversion, high efficiency, and longer product life batteries are required to allow the strong enabling and efficient integration of battery technology into our society. Lithium-ion cells are used for the formation of battery packs due to their good characteristics provided for a long time. In manufacturing industries, battery packs are produced at a massive rate. Pouch cell capacity inside the pack has to be balanced to make the output of the battery stable and efficient. Sorting pouch cells became a critical task to build a battery pack. So, the pouch cells have to be sorted into groups based on their capacity. A methodology was proposed in this paper for sorting pouch cells using the hybrid clustering algorithm. Results were compared on various clustering algorithms. Optimal clustering was selected by analyzing the results of the machine learning model.

Keywords. Battery packs, Lithium-ion, Pouch cells, Sorting, Clustering algorithm, Machine learning, Optimal clustering.

1. Introduction

Nowadays, electric vehicles are gradually being accepted by consumers. Lithium-ion batteries with high energy density, long cycle life, and low self-discharge rate are commonly used as energy storage devices for electric vehicles [1][2]. As the usage increases, the manufacturing rate should also need to be increased to meet the market demands.

Lithium-ion batteries have lately been improved for a variety of new uses, including power tools, UPS power backup, and also for electric cars. While lithium-ion batteries offer several advantages in terms of performance, they do need electrical monitoring and control systems. The minimal quantity of energy delivered by any battery may be used to gauge its performance. The phrases amp-hours and watt-hours are commonly used to characterize a battery's capacity. The energy density of a perfect battery would be extremely high or infinite. A battery pack is a collection of any number of identical batteries or individual battery cells. To achieve the necessary voltage, capacity, or power density, they can be set in series, parallel, or a combination of the two. Lithium-ion pouch cells as shown in figure 1 are grouped together to build lithium-ion battery packs.

[1] Jieh Ren Chang, Ashraf Mohammad, Department of Electronic Engineering, National Ilan University, Yilan City, Taiwan (R.O.C); E-mail: jrchang@niu.edu.tw.

Figure 1. Lithium-ion pouch cell.

In this project, new methodology was proposed to perform grouping of cells based on full charge capacity collected from the learning stage. This paper contributes several clustering algorithms from machine learning techniques. The structure of this paper as follow background work was discussed in section II. Section III provides the proposed methodology of the project. Experimentation and Results are shown in section IV and Section V concludes the paper.

2. Literature Review

In the literature, a physical model was developed to simulate the operation of the parallel-connected packs [3], a group of different degraded Lithium-ion battery cells was selected to build various battery packs and test them using a battery test bench. Various battery cell balancing methodologies have been developed in past study and evaluated their relationship with battery performance. In this paper [4] the cell inconsistency problem was investigated and its influence over parallel-connected packs for electric vehicles and plug-in hybrid electric vehicles applications. Batteries are prone to failure caused by charge imbalance in the batteries connected in either series or parallel. The balancing of the batteries at various temperatures, they are cycled, as well as the load vibration frequencies. An inconsistency model has been proposed based on the relationship between voltage and capacity inconsistency [5]. The statistical characters of discharge voltage are represented by a three-parameter Weibull Model than Normal Model. Seventeen battery pack models, based on four-series-connected battery pack structure models and three battery cell models, are introduced in [6]. Several experiments were carried out to collect realistic battery test data for parameter identification and model comparisons. Most popular clustering techniques such as K-Means [7], Hierarchical Clustering [8], Louvain Clustering [9] were experimented throughout this research work.

In our work, from the data collected, the average full charge capacity of each pouch cell was 8700mAh [10]. Every cell doesn't have the same discharging rate due to either technical manufacturing issues or the internal composition of chemicals inside cells, which affects full charge capacity. As a result, there will be a huge effect on the whole battery pack. Due to a shortage of data, the reliability of data-driven solutions is severely hampered. Researchers have devised a number of estimation/prediction approaches. Model-based, differential analysis-based, and machine learning-based are the three categories that they fall under. Because of extrinsic elements like temperature, mathematical models utilized by model-driven approaches to depict degradation processes are unstable. To achieve this sorting, every black box has to undergo the learning process to observe each pack's discharging rate. Capacity-based Black box sorting for energy storage system automation line still under research, which requires battery pack holders escalating the equipment size.

3. Methodology

This section gives a brief description of the clustering techniques in machine learning. Before implementing the machine learning techniques, need to consider whether to follow supervised or unsupervised algorithms.

Unsupervised learning models, operate independently to expose the internal representation of large datasets of unlabeled data. In this paper all the clustering techniques are unsupervised, As the battery packs capacity was unpredictable before the learning cycle stage at the manufacturing process. Full charge capacity was collected from the learning stage of manufacturing process. Charge capacity was numerical value and it cannot be labelled but it can be represented in the form of distributions. Unsupervised learning technique was adopted to make the adaptive grouping of battery packs irrespective of their capacity.

Clustering is a machine learning approach that divides data points into clusters so that data points in the same cluster are more similar to data points in other clusters than data points in other areas. It is the process of grouping related data points together. There are different types of clustering approaches. Based on previous research studies widely used clustering techniques such as K-Means Clustering, Hierarchical Clustering & Louvain Clustering are experimented in this research to group nearest capacity cells together.

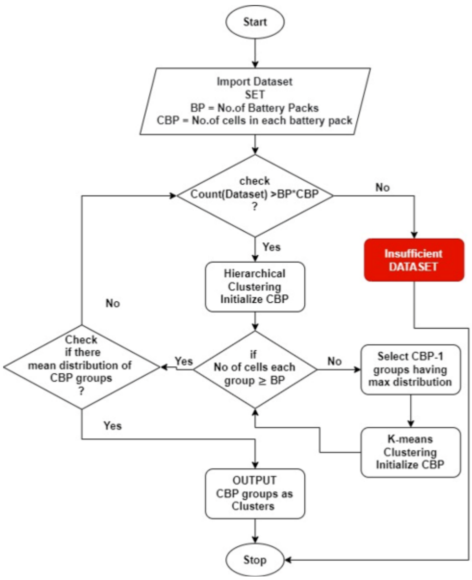

Figure 2. Flow chart of battery pack clustering algorithm.

Each clustering approach was applied to the pouch cell dataset in this study. Based on the findings, a hybrid clustering approach combining k-means and hierarchical clustering as shown in figure 2 was used to the dataset in order to automate the process of sorting pouch cells into battery packs. This algorithm can able to choose the maximum distribution groups on each iteration. As a process it removes the battery with having full charge capacity (FCC) is out of range based on distribution. Louvain clustering was not used in the creation of the hybrid clustering method because, as previously stated, the Louvain clustering algorithm finds the number of clusters and cannot regroup the clusters depending as required. However, an experiment on Louvain clustering was conducted, and the results were examined.

4. Experimentation and Results

Data was gathered from the battery pack manufacturing industry's pouch cell production process. 38621 lithium-ion pouch cells are represented in this dataset. Because grouping of cells would have been performed on the basis of cell capacity, only full charge capacity was considered as an input parameter. The number clusters will be considered based on the cells required to create the battery pack. As one battery pack, four pouch cells are linked in series. As a result, the dataset must be divided into four groups. Table 1 shows the statistics of the dataset.

Table 1. Dataset Exploration

Parameters	Total No of Records	Capacity Range (mAh)	Mean capacity (mAh)
Cell barcode, Full Charge Capacity (FCC)	38621	8500 - 9974	8837

The dataset was initially subjected to K-means clustering. k was set to 4 by default. K-means clustered all of the data points into four clusters. A pouch cell from each cluster will be chosen to balance the battery pack. As an outcome, for every range of pouch cells, there should be a mean distribution. Each cluster was represented with different color, In the distribution plots as shown in following figures, x-axis represents number of cluster and y-axis represents number of pouch cells. The distribution of pouch cells in each cluster is shown in table 2. Through this analysis we can observe that no mean distribution amongst clusters using k-means.

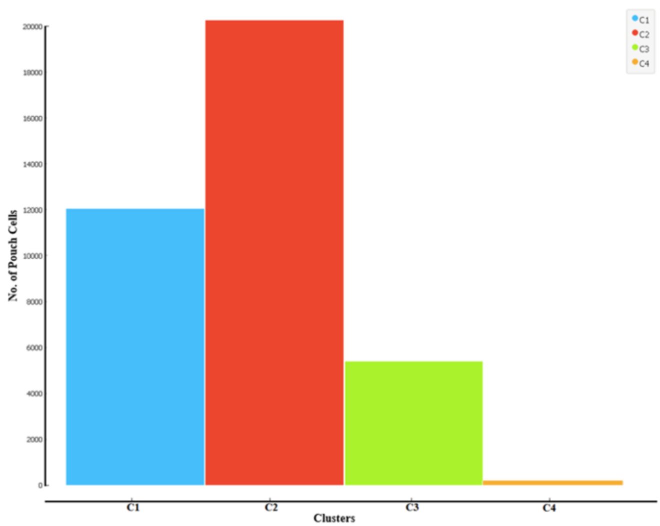

Figure 3. Distribution of K-Means Clustering.

On the same dataset, hierarchical clustering was conducted. Figures 4 i.e., hierarchy plot represents the hierarchy based on charge capacity of pouch cells. The clusters do not have a mean distribution. These clusters have the same maximum distributions as K-means, with just two clusters having the maximum distributions. The distribution of pouch cells in each cluster is shown in table 2.

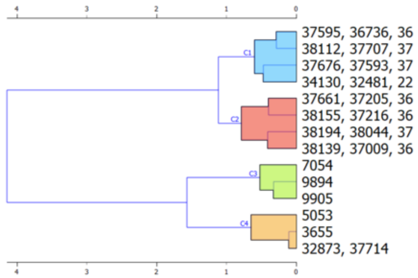

Figure 4. Hierarchy plot of Hierarchical Clustering.

On the same dataset, Louvain clustering was also performed. All datapoints were categorized into 516 clusters with 1000 k nearest neighbors. Required number of clusters cannot be initialized using this algorithm. Which means that we cannot able to set number of clusters. Based on k nearest neighbors, small change in FCC classified as new group. Only four clusters with their mean distribution are required for the manufacturing line.

In this algorithm hierarchical clustering is implemented first so that there is a possibility to choose groups on hierarchy as shown in figure 4. This helps to choose the groups with maximum distributions. All the nearest neighbors are grouped together. The distribution of pouch cells in each cluster using hybrid clustering is shown in figure 5. From table 2, It can be observed that there is mean distribution in between cluster 1 (C1) and cluster 2 (C2) using K-means. Similarly for hierarchical clustering. Cluster 3 (C3) and Cluster 4 (C4) doesn't have at least half of cells compared to remaining two clusters. But using Hybrid clustering. Cells are distributed among 4 clusters. The final output has 38253 pouch cells distributed among 4 clusters.

Table 2. Comparison of clustering results with different algorithms

Clusters	K-Means	Hierarchical Clustering	Hybrid Clustering
C1	11929	15872	6937
C2	21009	22382	11411
C3	5277	3	7749
C4	46	4	12156

5. Conclusion

Based on Machine learning techniques, a new hybrid clustering algorithm approach was developed for sorting of pouch cells. Experiments were made on most popular clustering techniques and compared with developed hybrid clustering algorithm. By delving into the previous section's results. Out of 38621 pouch cells, the combined clustering algorithm has distributed 38523 pouch cells into four groups. 368 cells are seen to be out of range, contributing for less than 1% of the overall dataset. This method can be used in the battery pack manufacturing process to create massive cell clusters. As soon as each pouch cell has completed the learning cycle, grouping may ensue. This algorithm can automate the production line by saving time and quick sorting. In the future, this algorithm might be used with combination of various algorithms for complete automation of the battery manufacturing process.

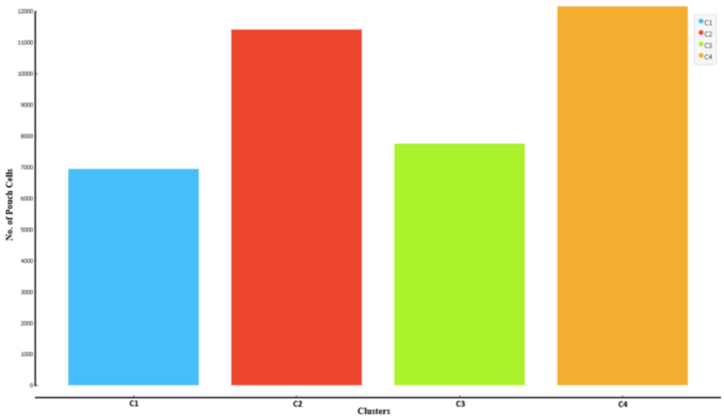

Figure 5. Distribution of Hybrid Clustering.

References

[1] Li, X., Xu, J., Hong, J., Tian, J., & Tian, Y. (2021). State of energy estimation for a series-connected lithium-ion battery pack based on an adaptive weighted strategy. Energy, 214, 118858. doi: https://doi.org/10.1016/j.energy.2020.118858.

[2] Arora, S., Abkenar, A. T., Jayasinghe, S. G., & Tammi, K. (2021). Materials and Manufacturing Methods for Advanced Li-ion Batteries. Heavy-Duty Electric Vehicles, 69–104. doi: https://doi.org/10.1016/b978-0-12-818126-3.00007-5.

[3] F. Si, Z. Li, X. Long, F. Jiang and W. Hua, "Research on Estimation of Battery State of Electric Vehicle Battery Management System," 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), 2020, pp. 465-469, doi: 10.1109/ITOEC49072.2020.9141648.

[4] Z. B. Omariba, L. Zhang and D. Sun, "Review of Battery Cell Balancing Methodologies for Optimizing Battery Pack Performance in Electric Vehicles," in IEEE Access, vol. 7, pp. 129335-129352, 2019, doi: 10.1109/ACCESS.2019.2940090.

[5] L. Wang, Y. Fang, L. Wang, F. Yun, J. Wang and S. Lu, "Understanding Discharge Voltage Inconsistency in Lithium-Ion Cells via Statistical Characteristics and Numerical Analysis," in IEEE Access, vol. 8, pp. 84821-84836, 2020, doi: 10.1109/ACCESS.2020.2992206.

[6] F. Feng et al., "A Practical and Comprehensive Evaluation Method for Series-Connected Battery Pack Models," in IEEE Transactions on Transportation Electrification, vol. 6, no. 2, pp. 391-416, June 2020, doi: 10.1109/TTE.2020.2983846.

[7] K. P. Sinaga and M. Yang, "Unsupervised K-Means Clustering Algorithm," in IEEE Access, vol. 8, pp. 80716-80727, 2020, doi: 10.1109/ACCESS.2020.2988796.

[8] W. Yang, X. Wang, J. Lu, W. Dou and S. Liu, "Interactive Steering of Hierarchical Clustering," in IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 10, pp. 3953-3967, 1 Oct. 2021, doi: 10.1109/TVCG.2020.2995100.

[9] T. M. Low, D. G. Spampinato, S. McMillan and M. Pelletier, "Linear Algebraic Louvain Method in Python," 2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 2020, pp. 223-226, doi: 10.1109/IPDPSW50202.2020.00050.

[10] Tian, H., Li, A. & Li, X. SOC estimation of lithium-ion batteries for electric vehicles based on multimode ensemble SVR. J. Power Electron. 21, 1365–1373 (2021). https://doi.org/10.1007/s43236-021-00279-9