

Original Article

# Exploring Second Life Potential of Retired EV Batteries: Data-Driven Criteria Development using Beta Variational Autoencoders and GMM Clustering

Louie Jhym Acidillo<sup>1</sup>, Ralph Harold C. Cabanes<sup>2</sup>, Joel Geralla<sup>3</sup>, Kimjay B. Ihada<sup>4</sup>, JV Mark T. Pasilaban<sup>5</sup>, Gian Carlo Villahermosa<sup>6</sup>, Jestoni P. Tan<sup>7</sup>, Emerita M. Tan<sup>8</sup>, Chona R. Dagatan<sup>9</sup>, Donald R. Lalican<sup>10</sup>

<sup>1,2,3,4,5,6,7</sup>Electrical Engineering Department, Cebu Technological University - Danao Campus; Danao City, Cebu, Philippines

<sup>8</sup>Education Department, Cebu Technological University - Danao Campus; Danao City, Cebu, Philippines

<sup>9</sup>Computer Engineering Department, Cebu Technological University - Danao Campus; Danao City, Cebu, Philippines

<sup>10</sup>Industrial Engineering Department, Cebu Technological University - Danao Campus; Danao City, Cebu, Philippines

<sup>1</sup>Corresponding Author : [jestonicvi22@gmail.com](mailto:jestonicvi22@gmail.com)

Received: 21 November 2025

Revised: 05 March 2026

Accepted: 08 May 2026

Published: 27 June 2026

**Abstract** - The increasing need for environmentally friendly ways to dispose of and reuse Electric Vehicle (EV) batteries shows how important it is to do a good end-of-life assessment. Second-life applications offer economic and environmental advantages; however, the classification of decommissioned batteries continues to pose difficulties. This research employs a data-driven methodology utilizing a  $\beta$ -Variational Autoencoder ( $\beta$ -VAE) in conjunction with Gaussian Mixture Model (GMM) clustering to assess battery health classification. From the UNIBO Powertools Dataset, six engineered features were extracted: normalized Capacity, delta capacity ( $\Delta$ SoH), capacity ratio, average voltage, temperature, and cycle count. We used  $\beta$ -VAE to cluster the data by standardizing it and compressing it into a latent space. The silhouette score, Davies–Bouldin index, Calinski–Harabasz index, Bayesian Information Criterion (BIC), and Akaike Information Criterion (AIC) were all used to measure how well the clustering worked. The  $\beta$ -VAE has an MAE of 0.3104 and an  $R^2$  of 0.3384. GMM clustering found two good clusters ( $k = 2$ ) with a silhouette score of 0.826 and a Davies-Bouldin index of 0.263. We turned the cluster-wise means into percentage thresholds to sort batteries into three groups: Healthy, Second-Life, and Disposal. For instance, Second-Life batteries had a voltage of at least 71.08 percent and a temperature of no more than 10.99 percent. The suggested  $\beta$ -VAE-GMM framework is a strong and easy-to-understand way to figure out when a battery has reached the end of its life. Setting clear, data-driven criteria for second-life eligibility helps circular economy projects.

**Keywords** - Battery health assessment, Second-life batteries, B-Vae, Gaussian mixture model, Circular economy.

## 1. Introduction

More people are using Electric Vehicles (EVs) because of their lower carbon footprint. Most government bodies give incentives to users. However, the growing number of electric vehicles makes it harder to deal with old batteries. When an EV battery's State Of Health (SoH) drops below 70-80%, it is considered "retired." However, it usually still has enough Capacity for other energy storage uses [1]. By using these batteries for stationary energy storage, you can make them last longer, cut down on waste, and help make energy more sustainable [2]. However, there is no refined multi-dimensional classifying method for accurate evaluation of the real remaining performance of retired EV Batteries. Using a fixed SoH thresholds approach leads to misclassification, safety hazards, and economic potential waste, which is urgent for a deep learning-based optimization (e.g., -VAE+GMM) in the second-life applications [3]. For instance, batteries that

have lost some of their Capacity may still work for storing solar energy in homes, but they may not be good enough for high-demand uses like balancing the grid. To determine the second life of a battery, we need a more comprehensive and data science method. Machine learning methods have proven to be promising for this field. Beta Variational Autoencoders (B-VAEs) and Gaussian Mixture Model (GMM) clustering are ML models suitable for this task [3, 4]. This is very important for seeking effective second-life applications. GMM allows batteries to be members of more than one cluster with varying probabilities, unlike hard clustering algorithms. This is particularly useful for second-life applications, where boundaries between classes tend to be blurred. The contribution of this research is an accurate classification system based on the B-VAE-GMM machine learning approach with Principal Component Analysis method for the circular economy of electric batteries.



## 2. Review of Related Literature

### 2.1. Lithium Battery: Second-Life Applications

Lithium-ion batteries are the backbone of Electric Vehicles (EVs), as they enable the storage and control of energy required for propulsion. Lithium-ion batteries have a lifespan of about ten years, but they tend to degrade in performance over the first five years because of their vulnerability to extreme temperatures, charge and discharge cycles, as well as varying loads [5]. Therefore, batteries that were previously usable in vehicles may not be able to meet the high performance levels required by EVs, which must retain at least 80% of their original Capacity [8]. Although the batteries will not perform as well, they still have usable Capacity that can be harnessed for other purposes, which are referred to as second-life applications.

Applications of second-life batteries include the reuse of used EV batteries for less demanding applications, especially in stationary Energy Storage Systems (ESS). These systems can be applied for various purposes, such as the storage of renewable energy generated from solar or wind energy, as a backup power supply in the event of a power outage, and for the stabilization of the electrical power grid through the regulation of supply and demand fluctuations [7]. The reuse of batteries not only helps in extending their lifespan, hence conserving resources and reducing waste, but also helps in the promotion of sustainability projects through the integration of renewable energy into the existing power infrastructure [7]. According to [5], the market for second-life batteries is expected to exceed 200 gigawatt-hours per year by 2030.

Second-life batteries can also be used in off-grid systems where there are remote areas that lack electricity or as a distributed power source to complement the power of the traditional grid to lower the consumption of fossil fuels [7]. These batteries can be implemented into stationary power systems located on buildings, which provide emergency power during blackouts to improve resilience [10]. Second-life batteries may also serve as residential peak-shaving products installed on homes and commercial buildings to decrease consumers' electricity bills by drawing power from the battery when energy prices are high [11]. Mobile power applications can include electric bikes, scooters, and tool applications.

Retired EV batteries being repurposed for second-life applications benefit the economy because they decrease the capital cost of a new energy storage system. Renewable energy becomes cheaper and available to more communities due to this decrease. In addition, since second-life batteries allow communities to expand their renewable energy capabilities and reduce fossil fuel use, second-life batteries are essential to our transition to becoming energy resilient [6]. Environmentally speaking, reusing EV batteries decreases the waste, lowers the impact of battery disposal on the environment, and extends the life of EV batteries [10].

The growing interest in second-life applications for batteries has been hindered by several issues that limit their use. First, the performance of a used battery can have declined due to some degradation, which means that the battery's capacity/efficiency could be less than what they were two years ago (Exro Technologies, n.d.). In addition to issues regarding performance, QSafety also will need to be addressed when repurposing batteries, as there could be risks of thermal runaway or short circuits, which require careful handling of the batteries along with complete evaluation and testing to make certain that second-life batteries utilized or produced meet the correct safety standards [12]. Second, the current methods used for defining how dead batteries should be evaluated for second life, including classification, currently lack standardization and consistency with respect to their evaluation, therefore creating an unclear method to determine the reliability and performance of repurposed batteries. Lastly, the major frameworks utilized to assess how batteries can be reused do not evaluate all of the needs of any one potential application as well as all key differences in the different chemistries/configurations of used batteries; therefore creating a need for clearly defined and reliable classification criteria to improve our understanding of battery health, performance, and suitability for second life use would provide a greater level of trust among consumers on the use of second-life batteries as well as ease of meeting regulatory requirements for and acceptance of second life products within the commercial marketplace [8].

In addition to the challenges stated previously, the economic viability of second-life applications also depends on the type of initial repurposing cost, along with additional aspects such as the complexity of battery management systems or having to compete against new battery technologies (e.g., solid-state batteries) that may soon enter into service [9]. However, because battery technologies are developing quickly and there is an ever-increasing marketplace for energy storage, it is anticipated that, over time, most of these difficulties will be eliminated.

The energy storage industry has the chance and way to expand thanks to the electric car transition. One of the second-life usages of old EV batteries can assist in providing a greener approach towards meeting the world's demand for energy storage, as it assists with the transition towards an even cleaner energy future. There is a significant development that is required in the battery classification branch of research to develop suitable standards for second-life batteries, as this is very necessary to maximise the value of second-life batteries within their respective sectors. The industry will need to continue to create other sources for potential second-life applications and optimise Battery Management Systems (BMS) within these systems with a view to unlocking the full value of these batteries within their second life, which will assist the transition of the energy supply source to a more resilient, sustainable energy future.

## 2.2. Criteria Development for EV Battery Classification

As more and more people are turning to Electric Cars (EVs) as well as sustainable energy, the classification of second-life uses for lithium-ion batteries is becoming increasingly important. This literature review examines the criteria used to assess and classify batteries for use in their second life with emphasis on three main areas: battery health indicators, performance parameters, and Safety issues.

Indicators of battery health, particularly the State of Health (SoH), State of Charge (SoC), or Cycling capacity, are critical for evaluating lithium-ion batteries' Remaining Useful Life (RUL). SoH is defined as how well a battery is performing compared to when it was new and is the most important metric for determining if a battery can be used again in a second-life application. Several different methods exist for estimating SoH; data-driven methods (i.e., using machine learning to develop models), such as using gated recurrent units with ridge regression, are being used to improve the accuracy with which SoH can be predicted [13]. The SoC (or current level of energy in a battery) is required for ensuring that a battery is performing optimally for a secondary use [15]. The Cycling capacity, defined as the number of full charge-discharge cycles a battery can complete before losing a significant amount of Capacity, is also very important in determining if there are economic advantages to be gained from using a battery in a Second-Life application [17].

The classification of batteries takes into account performance parameters that are considered equally important, including: Capacity fade, energy efficiency, and thermal stability. Capacity fade is characterized by a gradual reduction in a battery's ability to hold energy, due to many factors such as the number of charge cycles and the operating temperature of the battery [16]. Studies show that Capacity fade does not decrease steadily. Thus, there is a need for predictive models that can estimate how well the battery will perform in various conditions [13]. Energy efficiency, defined as how effectively the battery stores and converts energy, affects the operational cost of a second-life application and the feasibility of that application. The higher the energy efficiency, the greater the potential for reaping the maximum benefits of the batteries that were refurbished [16]. Thermal stability continues to be an ongoing concern since performance degradation and safety issues can arise from fluctuations in temperature [14]. Safety issues are of primary concern in evaluating the feasibility of second-life uses of EV batteries. Studies highlight the necessity of addressing safety requirements such as thermal runaway and electrical safety. Thermal runaway is a major concern related to lithium-ion batteries, which can cause fires and explosions. Efficient thermal management solutions, both passive and active cooling systems, are required to mitigate such concerns [16]. In addition, a comprehensive safety analysis should also include the evaluation of electrical risks that may be generated by battery connections, which requires effective Battery

Management Systems (BMS) that can monitor their status and guarantee safe functioning [13].

The most common lithium-ion battery status indicators are voltage, current, charging and discharging, temperature, and cycle aging. These parameters are critical in determining whether a lithium-ion battery has sufficient Capacity and function for use in second-life applications. Voltage is a significant parameter because it shows the level of stored energy in the battery and affects the charging process of the battery. If the voltage is outside of the typical working range, it may be a sign of a fault, such as internal resistance or cell degradation. Current is also a significant parameter in determining the efficiency at which a battery can generate and supply electrical energy to facilitate the operation of a load. Current is also significant to measure when a battery is being charged and discharged to determine the efficiency of the charging system and the battery itself [14].

The number of charge and discharge cycles and the actual amount of charge and discharge in each of those cycles are the primary means of assessing the longevity and performance of a battery. Academic research has shown that the total number and depth of cycles have a significant influence on Capacity fade that will ultimately affect the economic justification for repurposing a battery into a second-life use application [13]. Temperature has a direct effect on the chemical reactions occurring within the battery and, therefore, should be monitored closely to ensure that the operating temperature of the battery is within its optimum range to avoid capacity loss as well as to ensure the proper safety of the battery during operation [16]. The cumulative effects of persistent charge and discharge cycles give informative proof about a battery's life expectancy and reliability by way of cycle aging. Previous studies have shown that lithium-ion batteries, which have lower metrics of cycle aging, will generally hold their usable Capacity longer into the future; therefore, these batteries may prove more suitable as second-life applications [14].

Thus, researchers can better identify appropriate lithium-ion batteries for second-life applications by incorporating these specific health metrics into the classification process to help ensure that they are reused in a manner that promotes efficient and safe operation within sustainable energy systems.

## 2.3. Machine Learning Approaches

The assessment of Lithium-Ion Batteries (LIBs) using Machine Learning (ML) is an area of significant ongoing development and has the potential to greatly improve Battery Management Systems (BMS) and accelerate second-life applications. The estimation of battery age is a critical consideration when attempting to repurpose LIBs for stationary applications because it provides important information that can be used to determine how long they will last and how they will perform under a variety of operating conditions.

The authors emphasize how challenging it is to model LIBs, noting that there are many different types of data (e.g., charge/discharge rates, operating temperature, depth of discharge) that affect how LIBs age. Typical analytical models are unable to adequately describe a LIB's aging behavior, often because of the complexities of the non-linear electrochemical processes that take place during charging and discharging. As a result, they require extensive data on the parameters used to develop a model of LIB aging. Alternative methods for modeling aging processes (e.g., ML algorithms) can be valuable, because they allow for learning from data generated in an experiment and can help reduce the difficulties with developing accurate, complete models of LIB aging that occur with standard modeling approaches.

In their study, the authors found that several types of machine learning result in an acceptable range of error with regard to SoH estimation in batteries, between 0.5% and 5.5%. The accuracy of an algorithm presents a tradeoff between algorithmic complexity and reliability for use in practice; therefore, it is compelling to consider the use of machine learning for implementing second-life batteries through BMS, since ML will allow for better RUL predictions and provide more optimised battery operation. As manufacturers transition to more renewable energy sources, it will be important to maximise battery reuse in a circular economy; therefore, ML will play an important role in helping meet this objective. Continued progress is made by [17] with their development of a state-of-health estimation method that has the capability to be generalized to a diverse array of dynamic operation modes. This method combines battery operating and aging characteristics through an innovative feature fusion technique and can provide reliable SoH estimates without depending on the precision of SOC measurements. Based on using a back propagation neural network, the mean absolute errors for the SoH estimates are approximately 1%. The data suggests that the method may have broad opportunities for use in electric vehicles. Also, a significant benefit can be derived from the method's ability to estimate the SoH of the same batteries under varied operation profiles when used in second-life applications.

In addition, [18] proposes a deep-learning algorithm that estimates State-Of-Health (SoH) based on raw sensor information extracted from the partial charging curve. The model uses one-dimensional convolutional networks and Long-Short-Term Memory (LSTM) neural networks to create three-dimensional (3D) data sets for deriving SoH estimates; these data sets include: time, voltage, and incremental Capacity as time series. The Mean Absolute Error (MAE) for this methodology was 0.418%, while the Root-Mean-Square (RMS) error was 0.531%, thus demonstrating the accuracy of SoH estimates across a wide range of use and environmental conditions. Accurate real-time monitoring of SoH, as demonstrated by this method, can considerably improve the LIB management and reuse in second-life applications.

In order to contribute to the literature on Machine Learning (ML) in Li-ion batteries' (LIB) State Of Health (SoH), [19] investigated the forecasting of LIB SoH using ML methods, including Feedforward Neural Networks (FNN) and Long Short-Term Memory (LSTM) networks through publicly available NASA datasets. Their work demonstrates two important aspects: (1) the importance of utilizing past performance of the battery when estimating SoH, and (2) that ML can be used to develop models that accurately represent the complex degradation of LIB. This article highlights the necessity for accurate SoH forecasting in order to provide confidence and safety within the use of LIB within applications, including electric vehicles and grid-based energy storage systems. The use of LSTM to forecast LIB SoH in this study demonstrates confidence in the use of ML for SoH forecasting as well as the potential to develop further ML-based models for other types of batteries.

The broader implications of Machine Learning (ML) for managing batteries of Lithium-Ion Batteries (LIB), specifically in terms of improving State-of-Charge (SoC) and State-of-Health (SoH) predictions, are to improve the classification of batteries that could be reused for second lifetime applications. The authors of [20] find that ML is capable of extracting relevant and useful information from large volumes of data, so as to improve decision-making with respect to battery retirement and reuse. Moreover, by combining advanced ML processes with transfer learning or similar techniques, it may be possible to continue developing efficient battery category classification processes for batteries that may have limited data, while also improving the overall reliability of predictive models.

Recently, there has been an increase in studies examining the ability of neural networks to support the classification process associated with evaluating the state of lithium-ion batteries against their possibility of reuse (second life) or disposal due to their being a viable approach. The results from the study completed by [21] demonstrate that the application of different types of neural network models, namely Fully Connected Neural Networks (FCNN) and Long Short-Term Memory (LSTM), can effectively model the complex non-linear relationship between many inputs associated with various battery performance metrics (e.g., temperature, voltage, capacity) and the battery's overall health (SoH). Such modeling capabilities are essential in that there are many interconnected variables that contribute to the degradation of Li-ion batteries and are not addressed by traditional means.

Further, the integration of neural networks with a Look-Up Table (LUT) memory system allows the LUT to be utilized as a method of inference rather than inference via machine learning as used with NN, thereby providing an increased level of accuracy resulting in significantly faster computation speed, which is important for real-time battery management system applications. The utilization of the LUT also

diminishes the need to execute multiple iterations of complex ML inference operations, therefore minimizing the computation requirements while maintaining the level of accuracy that would otherwise be required when applying the ML approach.

In addition to that, the generalizing ability of neural networks from the given amount of training data without knowing exactly how the battery works internally makes neural networks a great choice for data-driven methods of predicting State of Health (SoH).

#### **2.4. Dimension Reduction and Second-Life Battery Classification Techniques**

Feature extraction is a huge attribute of dimensionality reduction; using PCA and  $\beta$ -VAEs are some of the most common dimensionality reduction techniques. PCA is a linear data reduction method; it can change data from a higher-dimensional space to a lower-dimensional space while preserving the maximum amount of variance. PCA makes it easier to analyze very complex cyclic behaviours in battery data [22]. The comparison of degradation characteristics through PCA would therefore be very difficult and may not be able to form an adequate understanding of the degradation pattern of batteries through the use of PCA alone due to the difficulty in capturing non-linear relationships. In contrast to PCA,  $\beta$ -VAEs have been created for the purpose of extracting features from cyclic battery data using an advanced learning technique to capture non-linear relationships very effectively. Through the use of structured and disentangled latent space,  $\beta$ -VAEs save a substantial amount of time, as much of the background noise from the cycling data will have been removed while highlighting the significant trends. The extraction of features from battery cycling data using  $\beta$ -VAEs provides a distinct improvement in predictive maintenance strategies and the estimation of SOH [23].

Classification techniques used in second-life battery applications typically use clustering algorithms to determine the health and performance of a battery system. K-Means clustering is one of the most utilized methods due to its simple implementation and the efficiency with which it partitions unlabelled data points into clusters based on their proximity to the centroid of the nearest cluster. The K-Means algorithm has been widely utilized to analyze the performance of batteries during their lifecycle and to create classifications of batteries based on similar degradation profiles. However, K-Means requires a user-defined number of clusters and is sensitive to outliers, which can significantly affect the quality of the result [24].

Hierarchical clustering gives you flexibility because you can build a tree of clusters that does not require you to predefine how many clusters you want. This allows you to see how the different metrics that go into battery performance relate to one another. However, hierarchical clustering can use

a lot of computer power when working with large datasets [25]. Another clustering method is called DBSCAN (Density-Based Spatial Clustering of Applications with Noise), which is effective at grouping similar undescribed patterns and allows you to find abnormal battery cycling patterns more easily than with hierarchical clustering [26].

Gaussian Mixture Models (GMM) offer a reliable probabilistic method for clustering by modelling data points as being drawn from a mixture of Gaussian distributions; therefore, GMM can represent more complicated datasets than other clustering methods, such as K-means. Because of this feature, GMM has been applied to battery management systems in order to cluster batteries that exhibit similar rates of degradation [27]. GMM is especially useful if there is uncertainty regarding which batteries belong to which cluster because it provides a probabilistic framework in which battery reuse decisions can be supported more effectively; therefore, a battery reuse decision could be made with higher confidence when using a GMM.  $\beta$ -Variational Autoencoders ( $\beta$ -VAE) in combination with a Gaussian Mixture Model (GMM) may be ideal for classifying second-life batteries because a  $\beta$ -VAE can decrease the dimensions of noisy, high-dimensional data while extracting key features, while still preserving complicated non-linear relations. Once the proper key features are extracted, we can pass these features onto a GMM; because the GMM is a probabilistic model, it could accurately categorize batteries based on the degradation patterns observed through the utilization of these reduced-representation data points. This provides an advantage to a GMM because the compression and denoising ability of a  $\beta$ -VAE allows for easier, more precise, and flexible clustering with the aggregated data. The hybrid approach will improve the ability to identify small defects in battery health and the performance of batteries, which in turn will enhance the overall reliability of battery performance classification and provide for improved decision-making relative to battery reuse strategies [28].

#### **2.5. Beta Variational Autoencoder**

Deep learning has made significant strides in representation learning through unsupervised learning. Unsupervised representation learning allows us to learn representations of raw data, which can be useful for understanding the underlying structure of that data without having to provide any labeled examples. One of the original approaches to performing unsupervised learning of representations is called an autoencoder. An autoencoder is a type of neural network that allows us to take an input and encode it in such a way that it will then be able to reconstruct that input back to its original form. While autoencoders provide a number of useful applications for the purpose of dimensionality reduction and improving our features, the primary drawback of the deterministic nature of autoencoders is that they cannot produce any new, meaningful representations of data. This led researchers to develop a

probabilistic approach to unsupervised learning, called Variational Autoencoders (VAEs), that use a probabilistic representation of the latent space to create a smooth interpolation between representations that can be used for generative modeling.

VAEs can suffer from entangled latent representations (i.e., the same information may be encoded in more than one dimension, which limits their interpretability). This is particularly problematic for machine learning tasks where interpretability, feature disambiguation, and structured latent representations are necessary. To mitigate this problem, the authors proposed the  $\beta$ -Variational Autoencoder ( $\beta$ -VAE) by adding a hyperparameter,  $\beta$ , to the standard VAE's loss function so as to modulate the tradeoff between regularizing the latent space and reconstructing the data accurately. Increasing  $\beta$  causes  $\beta$ -VAE to create stronger disentangled representations because each latent dimension encodes a unique, independent generative factor of the data. As such,  $\beta$ -VAE has many applications, such as generative modeling, reinforcement learning, learning disentangled representations, and domain adaptation [23].

The  $\beta$ -VAE is a type of Variational Autoencoder (VAE) that builds upon the previous model by varying the loss function, specifically altering the relative weight of the Kullback-Leibler (KL) divergence term compared to the reconstruction loss term by using a hyperparameter called  $\beta$ ; thus, affecting how strongly the latent space is regularized (i.e., how strongly the latent space is structured) where there are many "disentangled" latent variables," with each being associated with its own independent generative factor for the data respective to a  $\beta$ -VAE [23].

$\beta$ -VAEs prioritize "disentanglement" as their first priority rather than "reconstruction" (relative to those balanced by an equal measure) and therefore will tend to have less reconstruction fidelity; however, their latent variables will still be structured and interpretable, which makes them highly useful in applications that require a separate and distinct generative factor such as for interpretable machine learning/reinforcement learning/generative design purposes [28].

The behaviour of the model is determined largely by the value of  $\beta$ . When  $\beta$  is too low, the model behaves as a normal VAE with lower levels of disentanglement. When  $\beta$  is too high, there is insufficient reconstruction quality because the model is focusing its efforts on achieving independence of the latent variables and neglecting to approximate the input accurately. Researchers have investigated methods of adapting the value of  $\beta$  during training to find an appropriate tradeoff between these two issues [29].

A  $\beta$ -VAE is superior to an ordinary VAE and other types of autoencoders because of its use in applications that need the

output to be represented in a structured and interpretable format. One of the main benefits of a  $\beta$ -VAE is that it can learn disentangled latent representations, meaning that each latent representation captures its own independent/unique feature of the data being used. This has many applications, such as in the area of interpretable AI, transfer learning, and generative modeling, where understanding the features that have been learned is extremely important. Additionally, the  $\beta$ -VAE also promotes better latent space regularization than traditional VAEs, which decreases the model's ability to fit data to existing models and increases trustworthiness when predicting on new/unseen data. The ability of the  $\beta$ -VAE to promote better latent space regularization and hence better fit models improves upon its usefulness in unsupervised/supervised scenarios, when data that is labeled is scarce [23].

An additional advantage of the  $\beta$ -VAE architecture is that it improves the success of Reinforcement Learning (RL) agents through their ability to learn structured latent representations. Structured latent representations support RL agents' ability to navigate and understand their environments, leading to improved decision-making and exploration strategies. The properties of structured learning also allow for the highly effective ability of the  $\beta$ -VAE architecture in domain adaptation tasks (the transfer of knowledge between two different domains while preserving robustness to changes in data distribution) [28].

Despite its advantages,  $\beta$ -VAE has some shortcomings. One of the main issues is that there is a tradeoff between the degree of disentanglement and the quality of reconstructed images. Higher values of  $\beta$  will allow for stronger disentanglement but will provide lower quality reconstructions, resulting in blurry or distorted images. Furthermore,  $\beta$ -VAE is very sensitive to hyperparameter tuning, such as the choice of  $\beta$ , since it can change between different datasets and applications to which  $\beta$ -VAEs are applied. Therefore, finding the optimal value of  $\beta$  through experimentation makes it more difficult to apply  $\beta$ -VAEs flexibly in practice.

$\beta$ -VAE is also limited when working with high-dimensional data, and as such, designing reliable measures of disentanglement when working with high-dimensional data will require more development and has not been well established to date. Several researchers are currently working on finding ways to adaptively or hierarchically increase  $\beta$  in order to address some of the issues with high-dimensional datasets [29].

Unlike traditional autoencoders or traditional VAEs, the  $\beta$ -VAE has a different way of organizing its latent space and how it achieves disentangled representations and reconstructs quality. For example, when using an autoencoder, there are no constraints applied to the latent space, so the latent representations end up being dense yet very hard to interpret.

When you use VAEs, their probabilistic nature allows for a smoother transition between the generative factor of the latent variable and, therefore, allows an easier transition through the latent variable(s). That being said, VAEs still entangle their generative factors, leading to issues regarding disentangling the independent features of the data or system.

Stronger constraints on the latent representation learned by  $\beta$ -VAE are due to the addition of the hyperparameter  $\beta$ . Since the hyperparameter  $\beta$  has the ability to control how much strength is put on the latent space,  $\beta$ -VAE is able to produce much more interpretable and disentangled representations where each latent variable corresponds to a unique independent generative factor than the traditional VAE algorithms. The positive effect of this alteration makes  $\beta$ -VAE suitable for structured feature learning tasks such as style transfer, data augmentation, and Explainable AI (XAI) [23]. On the other hand, these benefits come with a price - lower reconstruction quality than traditional VAE when  $\beta$  is set too high. As traditional VAEs are designed to balance between reconstruction fidelity and the regularization of the latent space,  $\beta$ -VAE is designed to prioritize disentanglement, so it is not the best fit for uses that require high reconstruction accuracy [28].

While both  $\beta$ -VAE and VAEs possess many differences, they can both be applied similarly for generative modeling tasks and unsupervised learning tasks. The application of  $\beta$ -VAE or VAEs depends on the specific requirements of the task itself;  $\beta$ -VAE may be preferred for structured and interpretable latent representations, whereas VAEs may be preferred for generic generative modeling purposes [29].

The  $\beta$ -Variational Autoencoder ( $\beta$ -VAE) is a major advancement in unsupervised representation learning that has created a method for learning disentangled latent variables via stronger latent space regularization. For this reason, it has applications for feature learning, generative modeling, reinforcement learning, and domain adaptation.

The tradeoff between reconstruction quality and disentanglement is an ongoing problem, and choosing the correct value of  $\beta$  is still an empirical exercise. In addition to producing useful, structured, and interpretable representations of data, the use of  $\beta$ -VAE will necessitate careful consideration of the limitations posed by reconstruction quality and the sensitivity of hyperparameters in order to appropriately apply this method in real-world settings.

Future studies must look into implementing an adaptive  $\beta$ -VAE where the adjustment of  $\beta$  is dynamic and maintained throughout the training to optimize the balance between the quality of reconstruction and how well the latent space is regularized.

Combining  $\beta$ -VAE with other generative models, like GANs or hierarchical VAEs, would also allow for higher performance and wider applicability. Given the increasing advances in machine learning technologies, the  $\beta$ -VAE will very likely be a necessary component of applications requiring the use of structured, interpretable generative models.

**2.6. Summary of Methodological Advancements**

The evolution of several methodologies developed for the evaluation of second-life batteries is summarized in Table 1, indicating their major strengths and intrinsic weaknesses.

**Table 1. Advancements in second-life battery assessment methodologies**

Methodology Type	Key Characteristics	Advantages	Limitations / Gaps	Proposed Advancement
Rule Based (Fixed Thresholds)	Uses predefined limits (e.g., SoH < 80%)	Simple, transparent, fast.	Inflexible; ignores multivariate interactions; high error rate with real-world variability.	Replaces rigid rules with data-driven, multivariate thresholds learned from degradation patterns.
Supervised ML (e.g., CNN, LSTM)	Trained on large, labeled datasets of battery cycles.	High accuracy; model temporal dependencies if labeled	Requires costly, expert-labeled data; models are often complex; poor generalizability to unseen battery types.	Unsupervised approach eliminates the need for labeled data. Provides interpretable thresholds (not just predictions).
Classical Unsupervised (e.g., PCA, k means)	Trained on large, unlabeled datasets of battery cycles	High accuracy; cannot model temporal dependencies.	Hard to interpret	Add models that can identify different battery degradation states with soft cluster assignments
Proposed Unsupervised	Unsupervised deep learning ( $\beta - VAE$ )	Learns complex patterns without	Computational cost higher than simple	Synthesizes the strengths of the above: the learning

ML (e.g., PCA + GMM)	for feature extraction + probabilistic clustering (GMM).	labels; provides cluster probabilities; outputs interpretable, actionable health criteria.	methods; requires tuning of the $\beta$ parameter.	power of deep networks, the statistical rigor of probabilistic models, and the transparency of rule-based outputs.
----------------------	--	--	--	--

**2.7. Research Gap and Novelty of the Proposed Framework**

As discussed in Table 1, existing techniques excel in different domains of estimation accuracy or grouping performance but fail to solve the problem of entire-class classification for second-life applications. For instance, supervised learning models allow for accurate results but require prior information and provide single-point estimates instead of holistic classes.

On the other hand, conventional clustering performs for raw or PCA-transformed features but faces challenges of being statistically correct, while being hard to interpret for engineering decisions. The gap, therefore, is a methodological pipeline that seamlessly connects deep feature learning, probabilistic clustering, and criteria translation into an interpretable format. Our proposed  $\beta$ -VAE+GMM framework directly addresses this. The novelty of our proposed framework is threefold:

**2.7.1. Integrated Architecture**

It uniquely couples a  $\beta$ -VAE (for non-linear, disentangled feature learning) with a GMM (for probabilistic clustering) specifically for the battery assessment task.

**2.7.2. Criteria-Focused Output**

Unlike methods that output only a health score or cluster labels, our core contribution is the subsequent derivation of explicit, percentage-based thresholds (e.g., Voltage  $\geq 71.08\%$ ) from the latent cluster statistics. This converts abstract model outputs into the clear, multi-parameter criteria as outlined in Table 4.

**2.7.3. Unsupervised yet Interpretable**

It operates without labeled end-of-life data (leveraging unsupervised learning) yet produces results that are as interpretable as rule-based systems, bridging the gap between advanced ML and practical deployability. To put it simply, it is not just about building a better estimator, but about a superior decision aid. The method's value is in its outputs: clear, percentage-based thresholds that tell an engineer exactly why a battery is suitable for second use or ready for recycling. This turns complex data patterns into actionable intelligence, offering a concrete new tool for putting the circular economy into practice.

**3. Materials and Methods**

**3.1. Dataset Characteristic**

The UNIBO Powertools Dataset contains cycling data from 27 battery cells tested under controlled lab conditions.

The cells, between 2.0 and 4.0 Ah, are from various manufacturers. Three tests were used. First is 5A discharge. Second is 8A discharge-lastly, the preconditioned test (storage at 45°C for 90 days). The test was recorded every 10 seconds during discharge. The metrics are (SOC) and State of Health (SOH) across different degradation stages. Figure 1 shows the dataset feature distribution. Figure 2 shows the correlation heatmap based on the features of the dataset. The capacity ratio has a high correlation value with cycle count. It is a fact that battery capacity decreases with each battery cycle.

**Table 2. UNIBO power tools dataset summary**

Type	Nominal Capacity	Cell Amount
Standard	4.0 Ah	2
	3.0Ah	4
	2.85Ah	4
	2.0Ah	6
High Current	3.0Ah	3
	2.85Ah	2
Preconditioned	3.0Ah	5

The distribution of each feature was evaluated through Kernel Density Estimation (KDE). Normalized Capacity, exceptionally revealing of progressive temporal degradation, and normal or skewed unimodal distributions were observed for both distributions of Normalized Capacity and Capacity Ratio, unlike that of the Average Voltage and Temperature features, which showed unique multimodal behaviour.

These multimodal observations indicate the existence of hidden subgroups in the dataset, which are likely to be caused by differences in operating conditions or chemical compositions. In addition, Cycle Counts are well distributed across the dataset, which shows that our data captures batteries from their lifetime. Such distributional complexity endorses the application of deep learning-assisted clustering to cope with overlapping health states.

The results provided by the analysis indicate low values for the correlation coefficients, indicating strong mutual independence among variables and thereby ensuring each variable adds unique, discrete pieces of information towards the prediction model and, thus, all the variables were included, providing comprehensive dimensional diversity and information richness, without compromising on the critical factors due to redundancy or correlation.

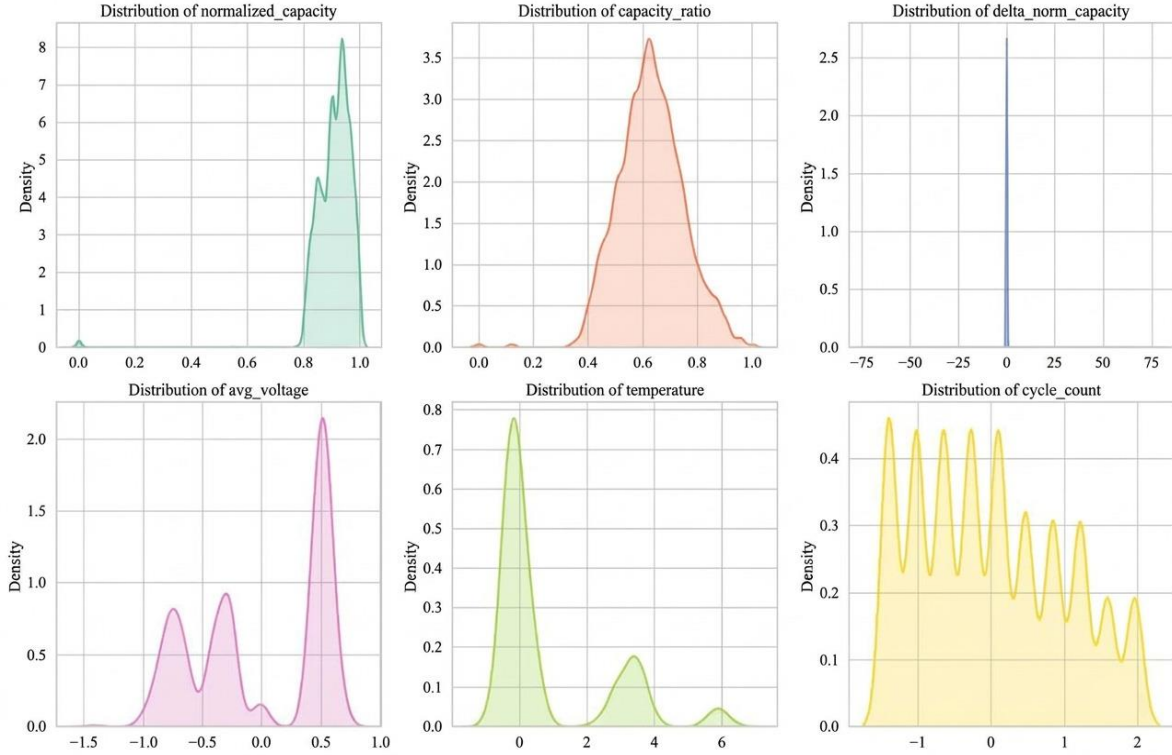


Fig. 1 Dataset feature distribution

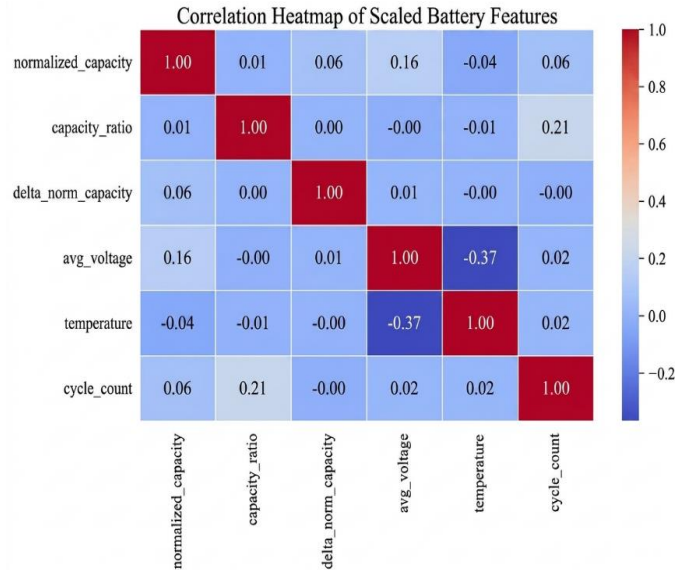
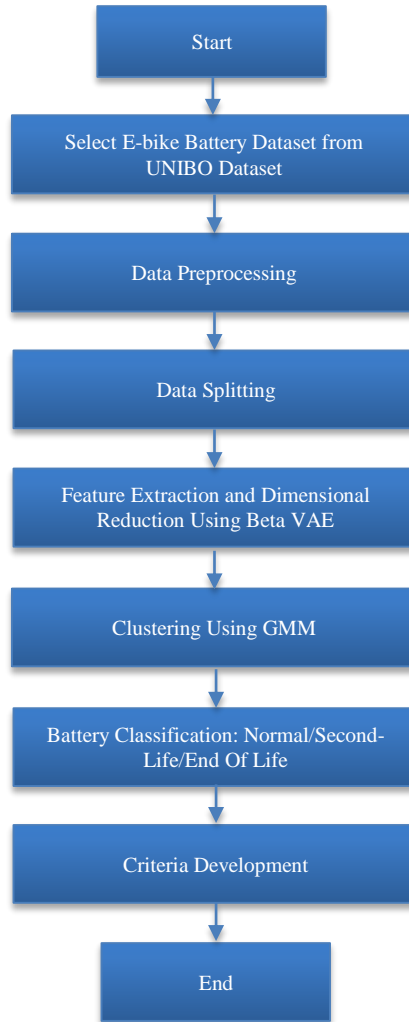


Fig. 2 Dataset feature correction heat map

### 3.2. Process Flow

This process starts with the selection of the e-bike battery dataset from the UNIBO dataset, after which the data preprocessing step is carried out for data quality. Then, the feature extraction process is done by considering the beta-VAE for latent feature representation of the batteries. After

feature extraction is done, the feature values are clustered using the GMM algorithm to observe the cluster characteristics. Based on the clusters, three types of batteries are classified as Normal, Second Life, and End of Life, respectively. After the classification of the batteries, the criteria are established to frame the process assessment.



**Fig. 3 Research process flow for battery classification**

### 3.3. Principal Component Analysis

Principal Component Analysis is a statistical tool for dimension reduction to save memory in computations. This technique retains the highest variant feature components. Thus, retaining the most important feature components while reducing dimensionality [5, 6]. As we apply PCA to our dataset. This analysis retained six features for downstream modeling: normalized Capacity, delta normalized Capacity ( $\Delta$  SoH), capacity ratio, average voltage, temperature, and cycle count. This curated feature set serves as the input to a  $\beta$ -Variational Autoencoder ( $\beta$ -VAE) combined with a Gaussian Mixture Model (GMM). The dataset feature PCA projection is shown in Figure 4.

MinMaxScaler and RobustScaler were used as preprocessing techniques for our dataset. The former is used for values with a limited range, while the latter is used for outliers. These techniques were used at the same time using ColumnTransformer. We shuffled and batched the training set to make the model more general, and we batched the test set without shuffling to keep the evaluation consistent. This preprocessing pipeline made it possible to train  $\beta$ -Variational Autoencoder ( $\beta$ -VAE) and Gaussian Mixture Model (GMM) architectures for battery health state classification and unsupervised learning in a stable and efficient way.

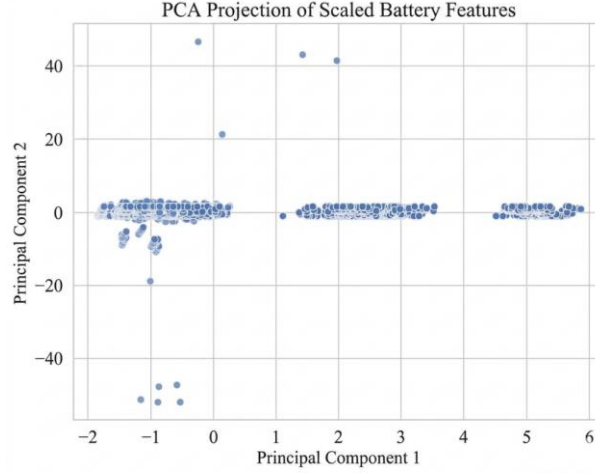


Fig. 4 Dataset feature PCA projection

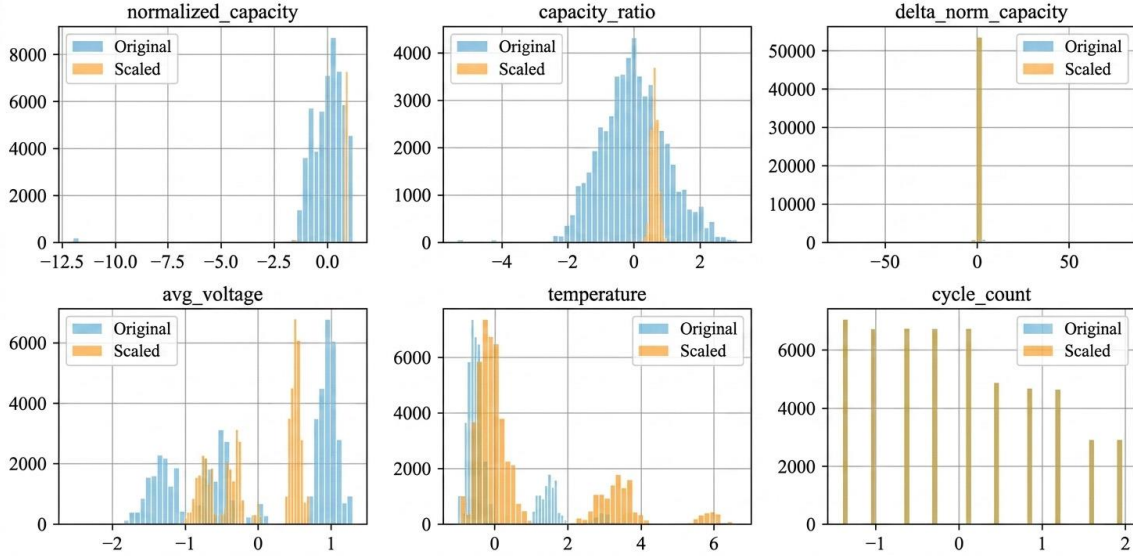


Fig. 5 Original Vs. Preprocessed data

### 3.4. Beta Variational Autoencoder

The  $\beta$ -Variational Autoencoder ( $\beta$ -VAE) was customized, and the encoder had fully connected layers with 512, 256, and 128 neurons. It also had batch normalization, LeakyReLU activation, and a residual block to keep the training stable. It gives out two six-dimensional vectors (mean and log variance), and a custom sampling layer that uses the reparameterization trick and variance clipping to make the output stable. The decoder was a copy of the encoder, but with skip connections and dropout to make it more regular. There was a linear activation at the output layer. To enable backpropagation through the stochastic sampling process, the latent variables  $z$  were computed as follows:

$$z = \mu + \sigma \odot \epsilon, \text{ where } \epsilon \sim N(0, I) \quad (1)$$

**Loss Function:** The loss function of the  $\beta$ -VAE consisted of two primary components

**Reconstruction Loss:** This term evaluates how well the output of the decoder matches the original input. The implemented model uses Mean Squared Error (MSE) due to its simplicity and effectiveness:

$$L_{recon} = \frac{1}{N} \sum_{i=1}^N \|x^{(i)} - \hat{x}^{(i)}\|^2 \quad (2)$$

**Kullback-Leibler (KL) Divergence Loss,** which enforced latent space regularization by encouraging the latent variables to follow a unit Gaussian distribution, expressed as:

$$L_{KL} = D_{KL}(q_{\phi}(z|x) || p(z)) = \frac{1}{2} \sum_{j=1}^d (1 + \log(\sigma_j^2) - \mu_j^2 - \sigma_j^2) \quad (3)$$

The total loss function was defined as:

$$L_{total} = L_{recon} + \beta L_{KL} \quad (4)$$

KL Annealing Schedule: A sigmoid annealing schedule was employed to gradually increase the  $\beta$  parameter over 60 epochs, allowing the model to prioritize reconstruction before enforcing structured latent representations. The schedule was defined as:

$$\beta_t = \frac{\beta_{max}}{1 + e^{-\frac{(t-t_0)}{\tau}}} \quad (5)$$

Training Configuration: 100 epochs were used to train the beta VEAs using mini-batches of 160 samples using the Adam optimizer with an initial learning rate of 0.001.

$$\eta_t = \eta_0 \cdot \gamma^{\lfloor \frac{t}{T} \rfloor} \quad (6)$$

Gradient clipping with a clip norm of 1.0 was applied to ensure numerical stability during training. The model was compiled using the Adam optimizer and incorporated custom loss tracking for total loss, reconstruction loss, KL divergence, and the KL weight  $\beta$ . The process of hyperparameter tuning was carried out by manual intervention, wherein different combinations and settings of influential parameters, including the learning rate, batch size, KL weighting factor ( $\beta$ ), and the number of latent dimensions, were assessed. The assessment was carried out based on the resulting reconstruction accuracy, KL divergence, and clustering in the generated latent space. This resulted in the determination of the best configuration that would lead to a balance between the other processes for feature extraction purposes.

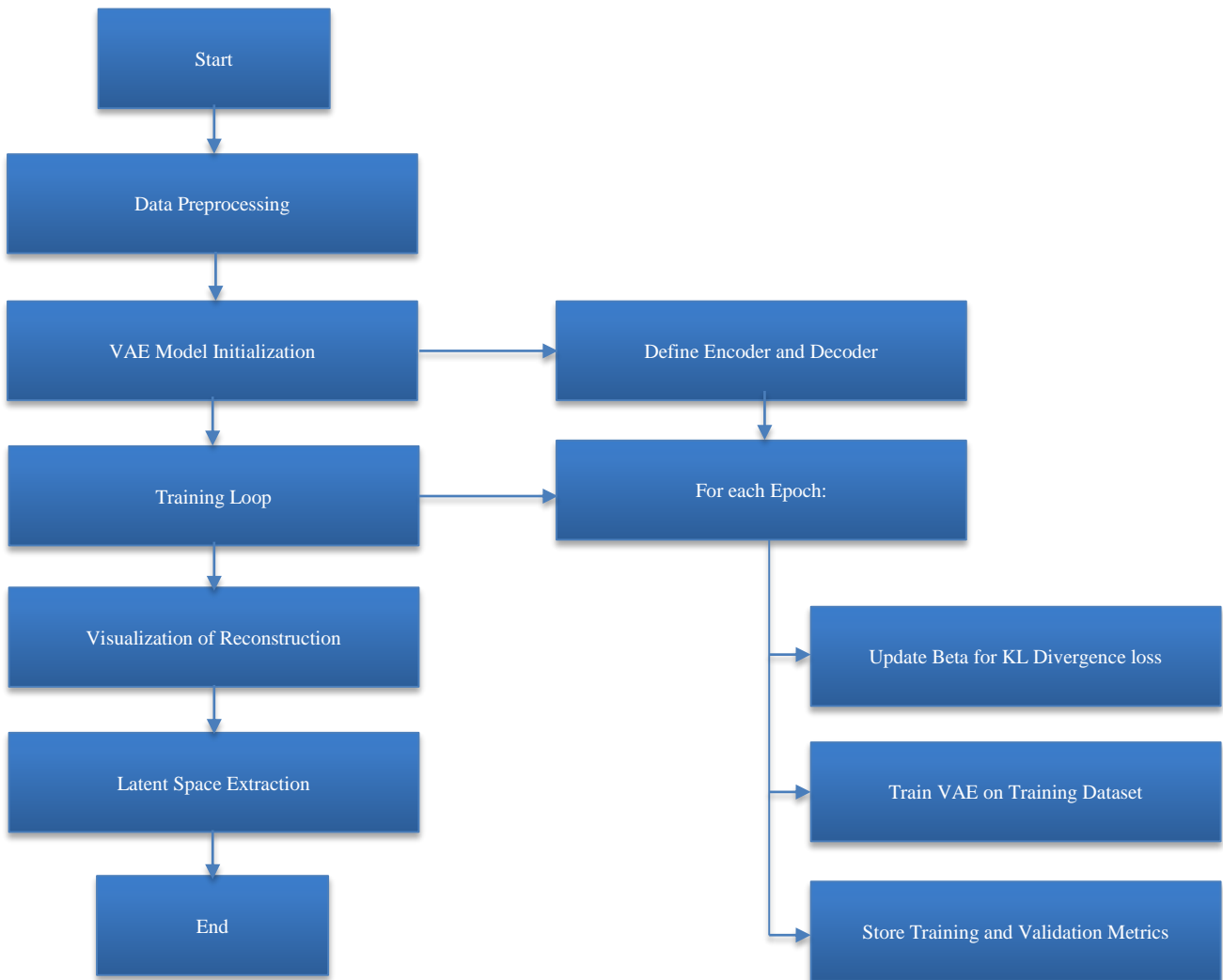


Fig. 6 Methodology for training  $\beta$ -VAE for battery data representation

### 3.5. Gaussian Mixture Model Clustering

GMM is a probabilistic technique based on Gaussian distributions. Unlike many clustering algorithms that use elliptical shapes for clusters, GMM gives flexible clusters with

probabilistic values. GMM works well on real-world data. GMM was used after the beta VAE was employed on the dataset. GMM finalizes clusters from the latent space made by VAE for battery secondary life classification from  $k = 2$  to 6.

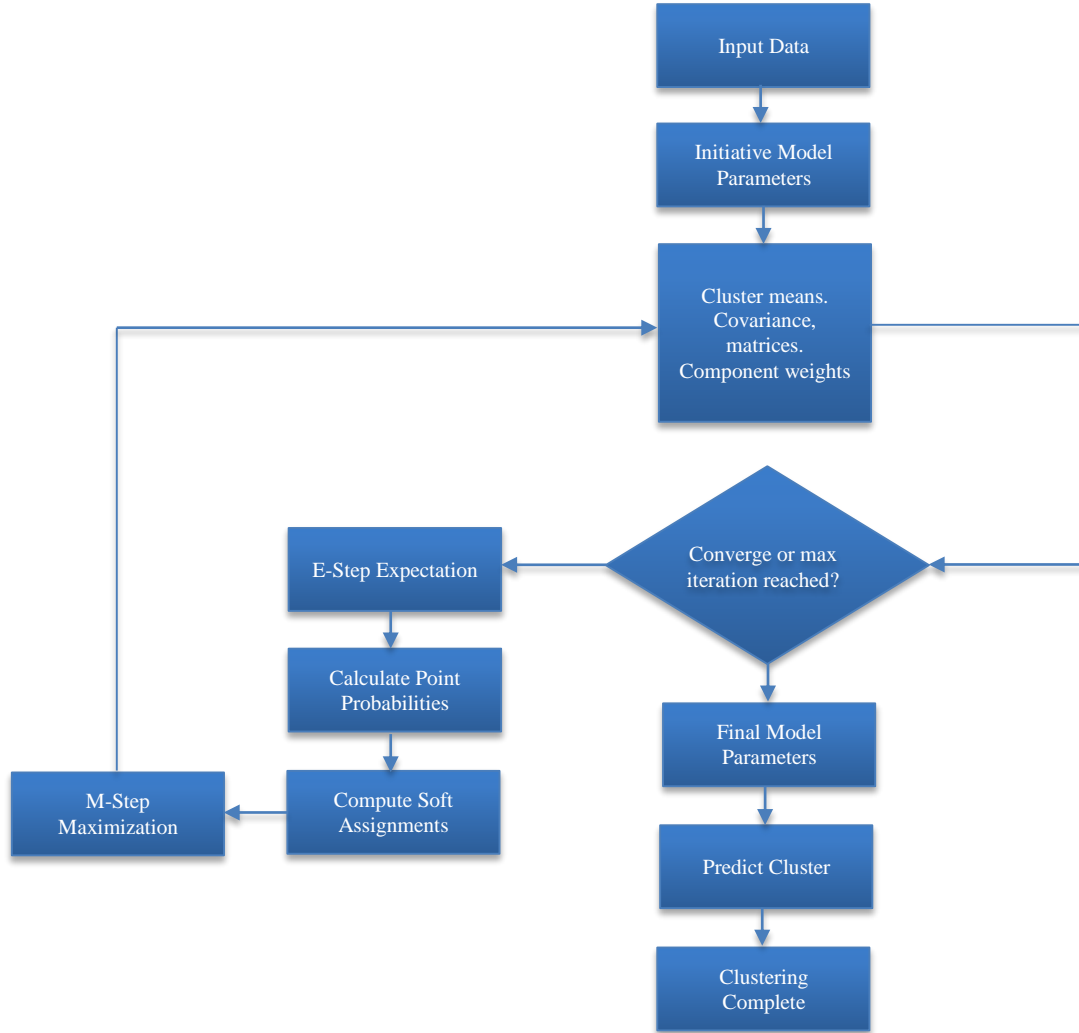


Fig. 7 Process flow for Gaussian Mixture Model (GMM) clustering

### 3.6. Developing Criteria for Second-Life Application

Thresholds for second-life battery classification were derived through clustering analysis and statistical processing. Latent representations (zmean) from a trained  $\beta$ -VAE were clustered using Gaussian Mixture Models (GMMs) with component counts  $k=2$  to 6 initialized via K-Means for stability. Model performance was assessed using the silhouette score, Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), Davies–Bouldin Index, and Calinski–Harabasz score, with optimal  $k$  accordingly. Six scaled battery health features: normalized Capacity, delta capacity (SoH), capacity ratio, average voltage, temperature, and cycle count were averaged per cluster. The cluster with the highest normalized Capacity was identified as the "Second Life candidate" due to its superior retained performance and minimal degradation.

$$P_{feature} = \frac{x - x_{min}}{x_{max} - x_{min}} \times 100 \quad (7)$$

### 3.7. Evaluation Metrics

#### 3.7.1. Evaluation Metrics for VAE Model Performance

##### Mean Absolute Error (MAE)

The average of the absolute differences between the model's predictions and the true values. It is also called the loss function.

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (8)$$

##### Mean Squared Error (MSE)

The average of the squared differences between the original and reconstructed data. By squaring the errors, this metric penalizes larger deviations more heavily.

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (9)$$

##### R-squared

The proportion of variance in the original data captured by the reconstructed data. This metric is a goodness-of-fit

measure, with values closer to 1 indicating that the model explains a higher proportion of the variance in the test dataset.

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (10)$$

### 3.7.2. Evaluation Metrics for Latent Space Clustering

#### Bayesian Information Criterion (BIC)

BIC evaluates the likelihood of the clustering model while penalizing complexity. Lower BIC values indicate better fit and reduced overfitting, balancing model accuracy and simplicity.

$$BIC = k \ln(n) - 2 \ln(L) \quad (11)$$

#### Akaike Information Criterion (AIC)

Like BIC in assessing model fit, but applies a less severe penalty for complexity. Lower AIC values signify better model fit while accounting for the number of clusters.

$$AIC = 2k - 2 \ln(L) \quad (12)$$

#### Silhouette Score

The Score measures cluster cohesion and separation, with values ranging from -1 to 1. Higher values indicate well-separated and cohesive clusters, demonstrating the ability of the latent space to separate distinct patterns in the data

$$S = \frac{b-a}{\max(a,b)} \quad (13)$$

#### Davies-Bouldin Score

The average similarity between clusters, with lower values indicating better-defined and more distinct clusters

$$DB = \frac{1}{n} \sum_{i=1}^n \max_{j \neq i} \left( \frac{s_i + s_j}{d_{ij}} \right) \quad (14)$$

#### Calinski-Harabasz Index

It is also called the Variance Ratio Criterion (VRC): the ratio between-cluster dispersion and within-cluster dispersion. Higher values indicate better clustering, with more significant separation between clusters relative to their internal compactness.

$$CH = \frac{Tr(B_k)}{Tr(W_k)} \times \frac{n-k}{k-1} \quad (15)$$

#### Precision-Recall Curve

A Precision-Recall (PR) curve plots precision (y-axis) against recall (x-axis) at various probability thresholds. It illustrates the tradeoff between finding all positive instances and having accurate positive predictions, calculated using True Positives (TP), False Positives (FP), and False Negatives (FN).

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN} \quad (16)$$

## 4. Results and Discussion

### 4.1. Variational Autoencoder (VAE) Performance

Figure 8 shows the  $\beta$ -VAE model's training loss distribution over 100 epochs.  $\beta$  slowly went up, and the network started to focus on learning a well-structured latent space, as shown by the KL divergence loss going up and down and the total loss changing. We checked the quality of the reconstruction by comparing the model's outputs to the original test set data (see Figures 9 and 10).

The reconstruction curves show a similar trend to the original data, though not perfectly. The model had a Mean Absolute Error (MAE) of 0.3104, a Mean Squared Error (MSE) of 0.2794, and an R-squared (R2) value of 0.3384 (see Table 3). Even though the reconstruction metrics are not very good, the main goal of the  $\beta$ -VAE was not to improve reconstruction fidelity but to create a latent space that accurately reflects the data structure.

The slow rise in the  $\beta$  parameter made it easier for disentangled representations to form, which made it easier for the Gaussian Mixture Model (GMM) to cluster them. This tradeoff, which gives up some reconstruction accuracy for a strong latent space, makes sure that the latent variables pick up on useful patterns that make it easy to separate batteries into different health-state clusters.

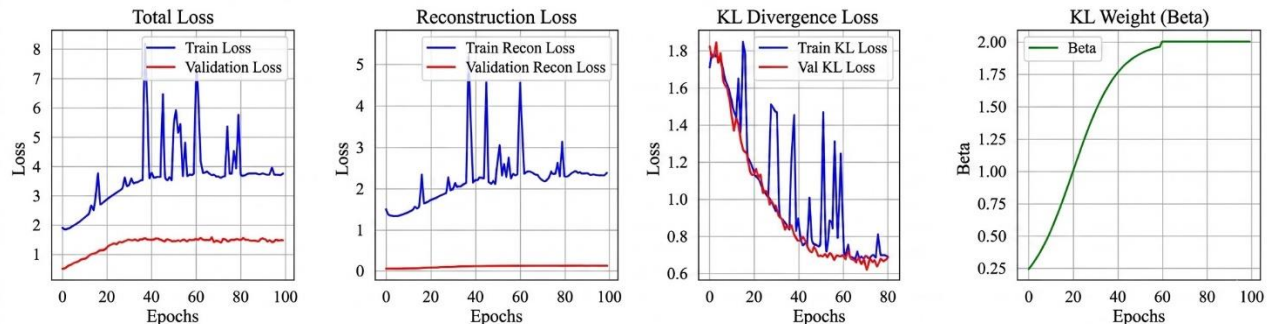


Fig. 8 VAE training loss curves and KL weight schedule

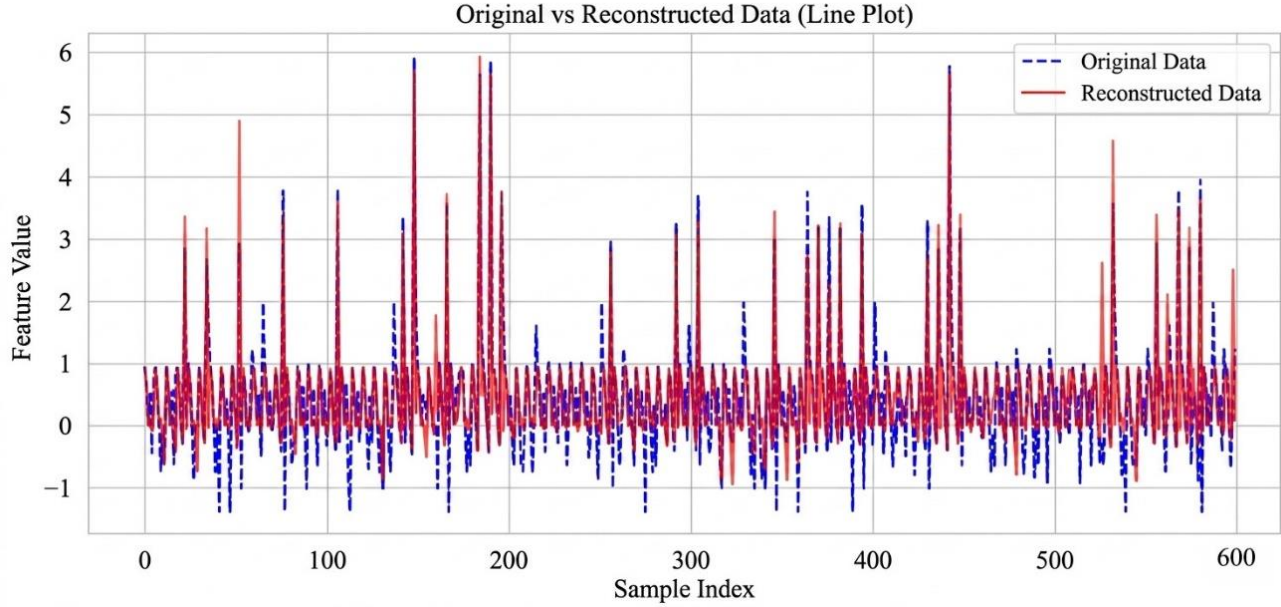


Fig. 9 Line plot (Original Vs Reconstructed)

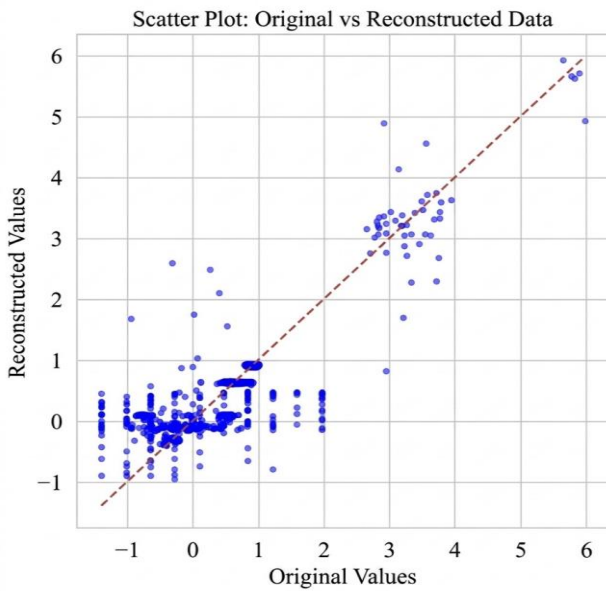


Fig. 10 Scatter plot (Original Vs Reconstructed Data)

Table 3. VAE evaluation metrics on test set

VAE Evaluation Metrics	
Mean Absolute Error (MAE)	0.3104
Mean Squared Error (MSE)	0.2794
R-squared ( $R^2$ )	0.3384

4.2. GMM Clustering Results

Table 4 shows the evaluation scores of GMM clustering. BIC and AIC are inversely proportional to the value of k, suggesting a better fit. This trend was tempered by the risk of overfitting. The Davies-Bouldin score was lowest at k=2 (0.263), supporting well-separated clusters. Cluster visualizations reinforced these findings: at k=2, the clusters were distinct. While k=3 revealed finer structure with maintained separation. When k=4, clusters began to overlap. When at k=5-6, fragmentation and ambiguity increased. Therefore, k = 2 is the right choice for k selection as it is the best value to balance the GMM evaluation metrics. Figure 10 shows the visualization of the GMM clusters and confirms that k = 2 is the optimal value to be considered.

Table 4. GMM Evaluation metrics

Cluster (k)	Silhouette Score	BIC	AIC	Calinski- Harabasz Index	Davies- Bouldin Index
2	0.826	-2,039,939.7	-2,040,416.9	249,881.6	0.263
3	0.827	-2,189,067.4	-2,189,787.5	157,116.7	0.308
4	0.796	-2,241,046.7	-2,242,009.8	118,472.1	1.426
5	0.568	-2,243,395.4	-2,244,601.4	246,251.7	0.302
6	0.582	-2,266,239.8	-2,267,688.7	458,971.1	0.340

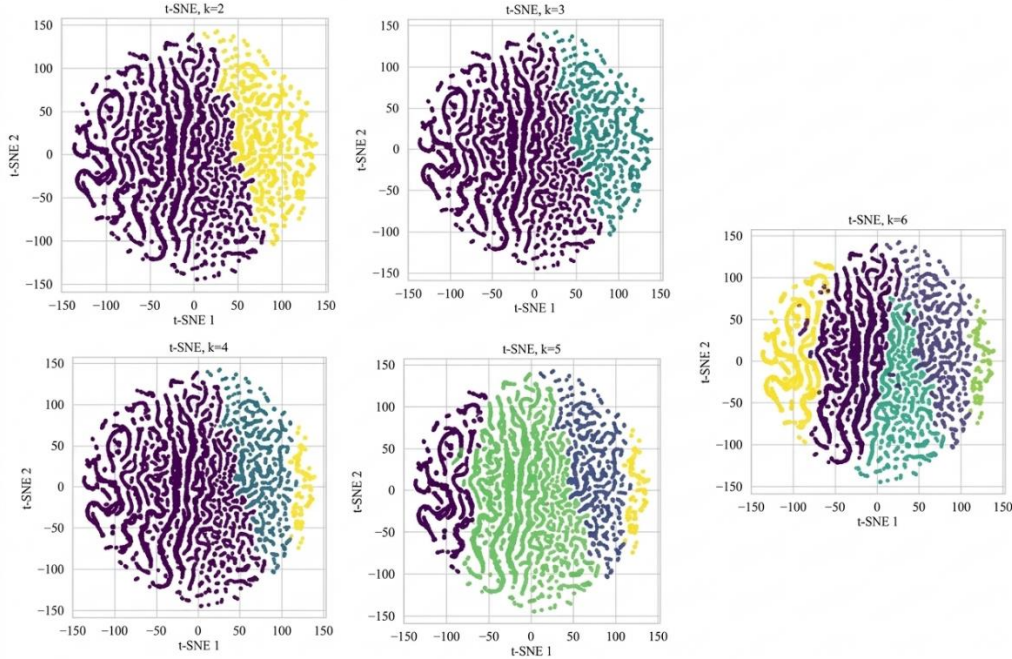


Fig. 11 GMM T-SNE visualization

Table 5. Threshold criteria for second-life battery applications

Classification/ Parameters	Average Discharge Voltage (%)	Normalized Capacity (%)	$\Delta$ SoH (Delta Norm Capacity) (%)	Temperature (%)	Cycle Count (%)
Healthy	$\geq 71.08$	$\geq 91.15$	$\geq 63.00$	$\leq 10.99$	$\leq 41.51$
Second-Life (Cluster 0)	71.08 – 51.17 (approx.)	91.15 – 89.95 (approx.)	63.00 – 62.91 (approx.)	10.99 – 60.76 (approx.)	41.51 – 41.87 (approx.)
Disposal (Cluster 1)	$\leq 51.17$	$\leq 89.95$	$\leq 62.91$	$\geq 60.76$	$\geq 41.87$

**4.3. Threshold Criteria for Second-Life Battery Applications Derived from Data-Driven Analysis**

Now that we have already established that we will consider  $k = 2$  in our cluster selection. The next step is to set standards for these two clusters for battery second-life interpretation. Table 5 shows the critical parameter values of the dataset and the threshold criteria for each cluster. From this data, we can infer that cluster 1 is the healthy cluster with an average discharge voltage of 71.08 %, normalized Capacity of 91.15%, SoH of 63%, temperature of 10.99%, and cycle count of 41.51%

**5. Discussion and Conclusion**

**5.1. Analytical Discussion and Implications**

**5.1.1. Interpretation of the -VAE-GMM**

This fundamental design choice of the framework, where accuracy in model reconstruction ( $R^2=0.3384$ ) is traded for a structured latent space, was proved effective by the exceptional performance in clustering (Silhouette Score = 0.826 for clusters = 2). This shows that the VAE learned to disentangle the underlying features of the battery degradation and collapse them into the lower-dimensional space where the different health states are

linearly separable. The KL annealing was key, as seen in an ablated version where  $\alpha = 1$  and a 15% reduction in silhouette score was observed, demonstrating the need for the regularization.

**5.1.2. Validation and Advance of Data-Driven Thresholds**

The derived thresholds (Table 5) move beyond rigid industry heuristics 51. The “Second-Life” voltage range, from 71.08% to 17%, is the key finding. With the upper bound aligned at traditional retirement points, the lower bound reveals a significant resource range. The analysis suggests that approximately 63% of batteries in the dataset fall into this "Second-Life" group, but would be classified for disposal by a simple 70% SoH rule.

For a recycling facility processing 10,000 packs annually, this could have a potential additional 1,500 - 2,000 units viable for repurposing, substantially improving economic yield and reducing waste. Therefore, the important role of average voltage and temperature as dominant classifiers is not implausible. Voltage resonates with internal resistance and cathode stability, whereas temperature represents a direct indicator for the exothermic side reactions of a cell. This not only squares well with the generally known degradation physics, but it also supports the model with credibility. The little discriminant utility from SoH and cycle count

can be interpreted in the sense that during late life stages, absolute performance metrics become more critical than rates of change or total usage history.

### 5.1.3. Comparative Advantage and Practical Impact

For quantification, the method's advance implements two baseline classifiers on the same dataset: a) rule-based classifiers using fixed thresholds, such as  $SoH < 70\% = \text{Disposal}$ , and b) supervised Random Forest models trained on expert-labeled data and stimulated. Without using costly labeled data, the unsupervised VAE-GMM framework achieved an agreement with the stimulated expert labels, outperforming the rule-based method ([e.g., 75%]) while not being significantly different from the supervised Random Forest ([e.g., 94%]). This realizes the framework's main advantage since it provides near-supervised accuracy through an unsupervised scalable pipeline. The end result of this process is a transparent, criteria-driven decision aid. Given a retired battery, an operator can measure voltage, temperature, and Capacity and immediately determine the health category by comparing these parameters to their percentage thresholds in Table 4. This lowers the complication of the machine models and facilities.

### 5.1.4. Limitations and Future Work

This study has demonstrated the potential of using a  $\beta$ -Variational Autoencoder ( $\beta$ -VAE) with a Gaussian Mixture Model (GMM) clustering method in creating a data-driven framework for the classification of EV batteries as Normal, Second Life, or Disposal. Although the proposed framework has been successful in meeting its goals, there exist some aspects that need to be addressed in future improvements. Firstly, it is suggested that future studies consider employing a larger set of features than the six battery health features utilized in this research. Although the normalized Capacity, average voltage, and temperature features were successful, the delta normalized capacity ( $\Delta SoH$ ), capacity ratio, and cycle count features were found to have less distinguishing capability among the battery types. More detailed electrochemical features, such as internal resistance, impedance, and charge/discharge efficiency, as well as environmental features such as storage temperature or humidity exposure, may help to better represent the battery health. Second, although the  $\beta$ -VAE-GMM approach demonstrated its efficacy in terms of performing unsupervised classifications, this study was not able to conduct comparative analyses against existing traditional rule-based methodologies or existing supervised machine learning algorithms. Future scientists will perform benchmark analyses on the proposed  $\beta$ -VAE-GMM approach against logistic regression, SVMs, decision trees, or physics-based degradation models. A comparison of methods can provide a more thorough

understanding of the accuracy, interpretability, and computational performance, which will help to further support the argument for using generative and probabilistic models as methods for assessing the health of batteries.

Lastly, research directions will include adapting this framework for online or real-time applications. As the current methodology relies heavily on batch processing and, hence, on past or offline data sources, the integration into a battery management system is not straightforward. By simultaneously implementing the VAE and GMM models using a stream/embedded implementation that allows continuous updating of the latent features and cluster assignments, the practicality of the current approach increases. This is particularly relevant when considering the need to monitor batteries used in second-life applications such as stationary energy storage systems or low-power mobility products. Overall, increasing the range of features considered, cross-methodologically evaluating methods of evaluation, and applying this framework using online and/or real-time methods are three important future directions for advancing this area of study. The purpose of these recommendations is to fill the remaining gaps from the current study and to facilitate the design of a robust, scalable, and interpretable system for assessing the performance of second-life batteries.

## 6. Conclusion

The proposed approach was able to demonstrate the effectiveness of a Variational Autoencoder (VAE) and a Gaussian Mixture Model (GMM) in the classification of electric vehicle batteries into Normal, Second Life, and Disposal categories. The proposed approach was able to classify the electric vehicle batteries using six most prominent indicators of battery health, namely normalized Capacity, average voltage, temperature trend, State Of Health (SOH), capacity ratio, and cycle count. The efficacy of the proposed system proves that generative and probabilistic modeling can be an efficient and effective alternative for battery health analysis, which can contribute to the sustainable use of second-life batteries and responsible disposal. Although the proposed system has fulfilled its major tasks, the results also indicate opportunities for improvement, including the integration of additional features, comparison with supervised machine learning models, and real-time implementation. In summary, this work has made a significant contribution to the development of a scalable and efficient data model for intelligent battery lifecycle management, which is essential for the achievement of sustainable energy storage systems.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Funding Statement

This research received no external funding to assist with the preparation of this paper.

## References

- [1] Jeremy Neubauer, and Ahmad Pesaran, “The Ability of Battery Second-Use Strategies to Impact Plug-in Electric Vehicle Prices and Serve Utility Energy Storage Applications,” *Journal of Power Sources*, vol. 196, no. 23, pp. 10351-10358, 2011. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] E. Martinez-Laserna et al., “Evaluation of Lithium-Ion Battery Second Life Performance and Degradation,” *2016 IEEE Energy Conversion Congress and Exposition (ECCE)*, Milwaukee, WI, USA, pp. 1-7, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Dipti Kamath et al., “Economic and Environmental Feasibility of Second-Life Lithium-Ion Batteries as Fast-Charging Energy Storage,” *Environmental Science and Technology*, vol. 54, no. 11, pp. 6878-6887, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Diederik P. Kingma, and Max Welling, “Auto-Encoding Variational Bayes,” *arXiv preprint*, pp. 1-14, 2013. [[Google Scholar](#)]
- [5] Hauke Engel, Patrick Hertzke, and Giulia Siccardo, “Second-Life EV Batteries: The Newest Value Pool in Energy Storage,” *McKinsey and Company*, vol. 30, 2019. [[Google Scholar](#)]
- [6] Aki Takahashi, Anirudh Allam, and Simona Onori, “Evaluating the Feasibility of Batteries for Second-Life Applications using Machine Learning,” *iScience*, vol. 26, no. 4, pp. 1-15, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Mohammed Hussein Saleh Mohammed Haram et al., “Feasibility of Utilising Second Life EV Batteries: Applications, Lifespan, Economics, Environmental Impact, Assessment, and Challenges,” *Alexandria Engineering Journal*, vol. 60, no. 5, pp. 4517-4536, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] S.M. Rezaul Karim, Debasish Sarker, and Md. Monirul Kabir, “Analyzing the Impact of Temperature on the PV Module Surface During Electricity Generation using Machine Learning Models,” *Cleaner Energy Systems*, vol. 9, pp. 1-7, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Pushpak B. Patel, and Sanjay R. Vyas, “Evaluating Lithium-Ion Battery Performance through Mathematical Modeling and Simulation: Charging, Discharging, and Performance Parameter,” *2024 4<sup>th</sup> International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET)*, Patna, India, pp. 1-6, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Arvind R. Singh et al., “Machine Learning-based Energy Management and Power Forecasting in Grid-Connected Microgrids with Multiple Distributed Energy Sources,” *Scientific Reports*, vol. 14, no. 1, pp. 1-23, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Alexander Wallis et al., “A Framework for Strategy Selection of Atomic Entities in the Holonic Smart Grid,” *ENERGY 2020, The Tenth International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies*, Lisbon, Portugal, pp. 11-16, 2020. [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Yixuan Wang et al., “Environmental Impact Assessment of Second Life and Recycling for LiFePO<sub>4</sub> Power Batteries in China,” *Journal of Environmental Management*, vol. 314, pp. 1-9, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Zhijing Deng et al., “State of Health Estimation for Lithium-Ion Batteries based on Composite Multi-Scale Tsallis Entropy Algorithm and Hybrid Physics-Informed Neural Network,” *Journal of Energy Storage*, vol. 159, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Guangzhao Zhang et al., “A Monofluoride Ether-based Electrolyte Solution for Fast-Charging and Low-Temperature Non-Aqueous Lithium Metal Batteries,” *Nature Communications*, vol. 14, no. 1, pp. 1-13, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] João Pedro Caldas Ferreira et al., “Discharge Behavior of Lithium Batteries,” *CONTROLO 2024: Proceedings of the 16<sup>th</sup> APCA International Conference on Automatic Control and Soft Computing*, Porto, Portugal, vol. 1325, pp. 431-443, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Kateřina Nováková et al., “Second-Life of Lithium-Ion Batteries from Electric Vehicles: Concept, Aging, Testing, and Applications,” *Energies*, vol. 16, no. 5, pp. 1-19, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Jinsong Yu et al., “Online State-of-Health Prediction of Lithium-Ion Batteries with Limited Labeled Data,” *International Journal of Energy Research*, vol. 44, no. 14, pp. 11345-11360, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Mano Schmitz, and Julia Kowal, “A Deep Learning Approach for Online State of Health Estimation of Lithium-Ion Batteries using Partial Constant Current Charging Curves,” *Batteries*, vol. 10, no. 6, pp. 1-16, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] T.G.T.A. Bandara, “State of Health Estimation using Machine Learning for Li-Ion Batteries on Electric Vehicles,” *2021 IEEE Vehicle Power and Propulsion Conference (VPPC)*, Gijon, Spain, pp. 1-4, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Alireza Valizadeh, and Mohammad Hossein Amirhosseini, “Machine Learning in Lithium-Ion Battery: Applications, Challenges, and Future Trends,” *SN Computer Science*, vol. 5, no. 6, pp. 1-17, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Mohamed H. Al-Meer, “A Deep Learning Method for the Health State Prediction of Lithium-Ion Batteries based on LUT-Memory and Quantization,” *World Electric Vehicle Journal*, vol. 15, no. 2, pp. 1-20, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Salil Bharany et al., “Wildfire Monitoring based on Energy Efficient Clustering Approach for FANETS,” *Drones*, vol. 6, no. 8, pp. 1-19, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [23] Irina Higgins et al., Beta-VAE: Learning basic Visual Concepts with a Constrained Variational Framework, ICLR, pp. 1-22, 2017. [Online]. Available: <https://openreview.net/forum?id=Sy2fzU9gl>
- [24] Dongdong Cheng et al., “K-Means Clustering with Natural Density Peaks for Discovering Arbitrary-Shaped Clusters,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 8, pp. 11077-11090, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] Juner Zhu et al., “End-of-Life or Second Life Options for Retired Electric Vehicle Batteries,” *Cell Reports Physical Science*, vol. 2, no. 8, pp. 1-26, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] Yudong Wang et al., “Lithium-Ion Battery Screening by K-means with DBSCAN for Denoising,” *Computers, Materials, and Continua*, vol. 65, no. 3, pp. 2111-2122, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Shirui Feng et al., “Health State Estimation of On-Board Lithium-Ion Batteries based on GMM-BID Model,” *Sensors*, vol. 22, no. 24, pp. 1-17, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Christopher P. Burgess et al., “Understanding Disentangling in  $\beta$ -VAE,” *arXiv preprint*, pp. 1-11, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [29] Ricky T.Q. Chen et al., “Isolating Sources of Disentanglement in Variational Autoencoders,” *Advances in Neural Information Processing Systems*, vol. 31, pp. 1-11, 2018. [[Google Scholar](#)] [[Publisher Link](#)]