OPTIMIZATIONS IN APPLIED MACHINE LEARNING
Research Article | Volume 5 | Issue 5 | Jan 2025
Received: 29 Nov 2024 | Revised: 15 Dec 2024
Accepted: 03 Jan 2025 | Published Online: 13 Jan 2025

Ridge Regression-based Approach for Estimate of Remaining Life Prediction of Lithium Battery

Olof Andersson[1], Linnea Gustafsson[2] and Erik Nilsson[3,*]

[1] Department of Sustainable Energy Systems, Mälardalen University, Västerås, SE-721 23, Sweden

[2] Center for Battery Innovation and Energy Technologies, University of Skövde, Skövde, SE-541 28, Sweden

[3] Institute of Advanced Materials and Energy Technologies, Örebro University, Örebro, SE-701 82, Sweden

[*]Corresponding Author, Email: erik.nilsson@oru.se

Abstract: In the context of lithium battery performance prediction, this paper addresses the critical need for accurately estimating the remaining life of the battery to optimize its utilization. Despite existing research efforts, challenges persist in achieving precise predictions due to factors like non-linear degradation mechanisms and limited data availability. To overcome these obstacles, our study proposes a novel Ridge Regression-based approach that integrates machine learning techniques with physics-based models. This innovative method not only improves prediction accuracy but also enhances model interpretability. By combining empirical data with theoretical insights, our research contributes to advancing the field of lithium battery prognostics.

Keywords: *Lithium Battery; Performance Prediction; Remaining Life Estimation; Ridge Regression; Machine Learning Techniques*

1. Introduction

Remaining Life Prediction of Lithium Battery is a field of study that aims to accurately estimate the remaining useful life of lithium batteries based on various factors such as charging/discharging cycles, operating conditions, and internal impedance. This predictive model is crucial for enhancing the reliability and efficiency of battery-powered devices and systems. However, the major challenges and bottlenecks in this field include the complex nonlinear behavior of lithium batteries, the lack of standardized testing protocols, and the need for precise real-time data collection and analysis. In addition, factors like electrode degradation, electrolyte decomposition, and thermal effects further complicate the accurate prediction of battery life. Overcoming these obstacles requires interdisciplinary research efforts combining expertise in materials science, electrochemistry, data analytics, and software engineering.

To this end, research on Remaining Life Prediction of Lithium Battery has advanced to a stage where various predictive models incorporating machine learning algorithms, electrochemical analysis, and data-driven methods are being developed and tested. These approaches aim to accurately forecast the remaining useful life of lithium batteries based on degradation mechanisms and operational conditions. Recent studies have made significant progress in the field of lithium-ion battery remaining useful life (RUL) prediction. Yang Li and Zhengang Shi (2024) proposed a novel neural network model integrating variational modal decomposition and Convolutional Neural Network (CNN) with Gated Recycling Unit (GRU) for RUL prediction, achieving high accuracy and robustness [1]. Jiusi Zhang et al. (2023) introduced an Expectation Maximization-Unscented Particle Filter-Wilcoxon rank sum test approach for adaptive noise estimation and capacity regeneration detection, outperforming existing techniques [2]. M. Reza et al. (2024) provided a comprehensive review of RUL prediction mechanisms, network configurations, and key issues in lithium-ion battery applications [3]. Zihan Li et al. (2023) proposed an Attention-CNN-Mogrifier LSTM-Maximum Mean Discrepancy model for RUL prediction, demonstrating superior accuracy and robustness over other methods [4]. Xiaowu Chen et al. (2024) developed a transfer learning-based RUL prediction model considering capacity regeneration, offering improved predictive capabilities [5]. Yuelong Pan and Jialong Ji (2024) presented an indirect prediction method based on charging IC curve and improved ELM for accurate RUL estimation, showcasing better prediction accuracy and robustness [6]. Jijuan Hu and Lifeng Wu (2024) proposed a Transformer Ensemble Model for early uncertainty quantification prediction of battery RUL, achieving enhanced generalization and prediction accuracy [7]. Zhuang Zhen et al. (2024) combined AUKF and CNN-BiLSTM for RUL prediction, enhancing both accuracy and stability of predictions [8]. Wenxin Ma et al. (2024) introduced a Deep Learning-based framework for accurate RUL prediction considering the two-phase aging process, providing timely alerts for battery replacement [9]. Ning He et al. (2024) developed a fusion model considering capacity regeneration for RUL prediction, contributing to improved prediction accuracy and robustness [10]. Recent advancements in lithium-ion battery Remaining Useful Life (RUL) prediction have seen the emergence of various innovative models with exceptional accuracy and robustness. Among these, Ridge Regression serves as a crucial technique due to its capability to effectively handle multicollinearity and overfitting in high-dimensional datasets. Ridge Regression plays a vital role in enhancing the predictive performance of RUL models by mitigating these challenges, ultimately improving overall accuracy and stability in predictions.

Specifically, Ridge Regression serves as a robust statistical method for addressing multicollinearity in predictor variables, making it particularly useful in the Remaining Life Prediction of Lithium Batteries, where accurately modeling the degradation parameters can significantly enhance the reliability of lifespan forecasts. Recent research has explored various aspects of kernel ridge regression (KRR) and its applications. Hoerl and Kennard [11] introduced ridge regression as a biased estimation method for nonorthogonal problems. Li et al. [12] investigated the saturation effect of KRR, proving a long-standing conjecture regarding its performance limitations. Xu et al. [13] proposed a novel approach named Kernel Ridge Regression-Based Graph Dataset Distillation (KIDD) for distilling large graph datasets efficiently. Zhang et al. [14] discussed the optimality of misspecified KRR and its applicability in different scenarios. Furthermore, Nguyen et al. [15] presented a meta-learning algorithm, Kernel Inducing Points (KIP), for dataset compression in KRR tasks, showcasing improved distillation results for MNIST and CIFAR-10 datasets. Cheng and Montanari [16] developed a dimension-free theory for ridge regression, offering non-asymptotic bounds on bias and variance. Carneiro et al. [17] applied a ridge regression ensemble of machine learning models to solar and wind forecasting in Brazil and Spain. Lastly, Wang and Jing [18] explored Gaussian process regression, discussing its optimality, robustness, and relationship with kernel ridge regression. However, current limitations persist, including the potential instability of KRR under high-dimensional settings, the necessity for careful parameter tuning, and the challenges in scaling to large datasets efficiently.

The paper by W. Huang, T. Zhou, J. Ma, and X. Chen presents an innovative ensemble model that ingeniously integrates multiple machine learning algorithms to predict the remaining useful life (RUL) of lithium batteries in electric vehicles. Drawing inspiration from their work, our study adopts a similar multi-layered approach to enhance the accuracy and reliability of RUL predictions for such batteries. Their ensemble model, which emphasizes the fusion of diverse algorithms, lays a robust foundation for us to develop a refined methodology that capitalizes on the strengths of Ridge Regression, blending it seamlessly with other predictive techniques to achieve a holistic estimation model. Huang and colleagues' research meticulously outlines the significance of integrating heterogeneous data sources and algorithmic diversity to capture the intricate nonlinear relationships inherent in battery life cycles. In parallel, our approach incorporates these insights by focusing on algorithmic fusion to mitigate the limitations posed by individual models and leveraging ensemble predictions to achieve superior performance. The meticulous data pre-processing strategies and cross-validation techniques discussed in Huang et al.'s study [19] provide a blueprint for enhancing model robustness, which we have expanded upon by including additional validation metrics tailored to the specific degradation patterns observed in lithium batteries. The paper's detailed discussion on the calibration of prediction intervals and model tuning serves as a pivotal reference point, guiding our optimization processes and parameter selection. By employing their ensemble approach as a foundational framework, our study benefits from the nuanced understanding Huang et al. provide regarding algorithm synergies, specifically in their adaptive weighting mechanism that dynamically adjusts to variations in battery data characteristics [19]. This dynamic adjustment serves as an indispensable element in our adaptation of their model, allowing our Ridge Regression-based framework to retain flexibility and precision across varying operational contexts. Furthermore, their emphasis on continuous model training and adaptive learning aligns with our objective to sustain model accuracy over prolonged battery usage periods, ensuring the reliability of RUL predictions amidst evolving conditions. This vigilant adaptation, as illuminated by Huang et al. [19], empowers our methodology to consistently refine its predictive accuracy across diverse scenarios. Through careful incorporation and expansion of these pioneering concepts, our study stands poised to contribute a complementary approach to the critical domain of lithium battery prognostics, thereby building upon the formidable groundwork established by the esteemed authors. By extending their detailed exploration of ensemble techniques, our work seeks to offer an ancillary perspective that amplifies the effectiveness of machine learning applications in estimating the RUL of lithium-ion batteries, thereby validating the enduring relevance of ensemble methodologies in complex predictive modeling tasks.

In the context of lithium battery performance prediction, this paper addresses the critical need for accurately estimating the remaining life of the battery to optimize its utilization. Section 2 describes the problem statement, highlighting the ongoing challenges in achieving precise predictions due to factors such as non-linear degradation mechanisms and limited data availability. To overcome these challenges, Section 3 introduces a novel Ridge Regression-based approach that integrates machine learning techniques with physics-based models. This innovative method enhances both prediction accuracy and model interpretability. In Section 4, a detailed case study is presented, demonstrating the practical application of the proposed methodology. Section 5 analyzes the results, showcasing the improvements in prediction performance and robustness. Section 6 delves into the discussion, examining the implications and potential applications of the findings. Finally, Section 7 provides a comprehensive summary, underscoring the contribution of this research in advancing the field of lithium battery prognostics by effectively blending empirical data with theoretical insights.

2. Background

2.1 Remaining Life Prediction of Lithium Battery

Remaining Life Prediction (RLP) of Lithium-ion Batteries is a crucial aspect of battery management systems. It involves estimating the time or cycles a battery can continue to operate before it falls below a certain performance threshold. This prediction is vital for applications ranging from consumer electronics to electric vehicles and renewable energy storage. The RLP of lithium batteries is inherently complex due to factors such as chemical reactions, material degradation, and usage patterns affecting battery performance over time. One of the core concepts in RLP is the State of Health (SoH), which quantifies the current condition of a battery compared to its ideal condition. The SoH can be expressed as:

$$\text{SoH} = \frac{\text{Current Capacity}}{\text{Nominal Capacity}} \times 100\% \tag{1}$$

Lithium battery degradation occurs due to factors like cycle aging and calendar aging, leading to a reduction in capacity (C_{loss}) and power capability. The degradation rate can be modeled as:

$$C(t) = C_0 - C_{\text{loss}}(t) \tag{2}$$

where C_0 is the original capacity and $C(t)$ is the capacity at time t .

To predict the Remaining Useful Life (RUL) of a battery, two primary techniques are employed: empirical modeling and model-based approaches. Empirical models often involve data-driven methods such as machine learning. A simple empirical prediction model may involve linear regression:

$$C(t) = a \cdot t + b \tag{3}$$

where a and b are regression coefficients. However, empirical models may not accurately capture the underlying physical processes, leading to inaccuracies under varying operating conditions. On the other hand, model-based approaches rely on understanding the physical and chemical processes. They often involve equivalent circuit models (ECM) or electrochemical models. An ECM can be represented as:

$$V(t) = V_{\mathrm{OC}} - R \cdot I(t) - V_{\mathrm{transient}} \tag{4}$$

where $V(t)$ is the terminal voltage, V_{OC} is the open-circuit voltage, R is the internal resistance, and $V_{\mathrm{transient}}$ accounts for transient voltage responses due to electrochemical dynamics. Bayesian approaches can also be applied to refine RUL predictions by incorporating the uncertainty of measurements and model parameters. The probability density function of RUL, $p(\mathrm{RUL})$, is updated as:

$$p(\mathrm{RUL} \mid \mathrm{data}) \propto p(\mathrm{data} \mid \mathrm{RUL}) \cdot p(\mathrm{RUL}) \tag{5}$$

Kalman filters, a recursive algorithm for state estimation, may also be employed. The Kalman filter prediction step can be written as:

$$x_{k+1}^{-} = A \cdot x_k + B \cdot u_k \tag{6}$$

where x_{k+1}^{-} is the predicted state, A is the state transition model, x_k is the current state, B is the control-input model, and u_k is the control vector. In conclusion, predicting the Remaining Life of Lithium-ion Batteries involves multiple methodologies that tackle the problem from empirical, circuit-based, and model-based perspectives. Accurate life prediction requires a combination of these approaches to account for the intrinsic complexities of lithium-ion battery systems.

2.2 Methodologies & Limitations

The Remaining Life Prediction (RLP) of lithium-ion batteries involves assessing the time or charge cycles a battery can sustainably support before it can no longer meet required performance standards. Understanding and accurately predicting this lifespan is essential for optimizing battery management across various applications, such as in consumer electronics, electric vehicles, and renewable energy systems. The methods used for RLP are critically structured around modeling battery degradation and leveraging empirical data to forecast the battery lifecycle. A primary element in the realm of RLP is the State of Health (SoH), which can be computed as:

$$\mathrm{SoH} = \frac{\text{Current Capacity}}{\text{Nominal Capacity}} \times 100\% \tag{7}$$

SoH provides a metric indicating the overall condition of a battery relative to its ideal state. This degradation is influenced by cycle aging (the number of charge-discharge cycles) and calendar aging (time-based aging), both leading to capacity loss (C_{loss}). Mathematically, the degradation over time can be characterized by:

$$C(t) = C_0 - C_{\text{loss}}(t) \tag{8}$$

where C_0 is the initial capacity and $C(t)$ is the capacity at time t. For predicting the Remaining Useful Life (RUL) of a battery, both empirical and model-based methodologies are employed. Empirical models are heavily data-driven; examples include machine learning techniques and simple statistical methods like linear regression, which can be expressed as:

$$C(t) = a \cdot t + b \tag{9}$$

In this equation, a and b are regression coefficients that need to be tailored for specific datasets. The limitations of empirical models arise from their dependency on historical data, which may not fully capture complex chemical phenomena inherent to battery processes, especially under diverse operating environments. Conversely, model-based approaches include physical and chemistry-informed models. These approaches encapsulate methods like equivalent circuit models (ECM) and electrochemical models. ECM, for instance, typically assumes:

$$V(t) = V_{\text{OC}} - R \cdot I(t) \tag{10}$$

where $V(t)$ is the terminal voltage, V_{OC} is the open-circuit voltage, R represents internal resistance, and $I(t)$ is the current. Moreover, probabilistic frameworks such as Bayesian inference are utilized to enhance prediction by accommodating measurement and model uncertainties. The posterior probability distribution of RUL, $p(\text{RUL})$, updates with new data as:

$$p(\text{ RUL} \mid \text{data}) \propto p(\text{ data} \mid \text{RUL}) \cdot p(\text{RUL}) \tag{11}$$

Recursive algorithms like the Kalman filter also play a vital role in state estimation, written as:

$$x_{k+1}^{-} = A \cdot x_k + B \cdot u_k \tag{12}$$

where x_{k+1}^{-} is the forecasted state, A denotes the state transition model, x_k is the current state, B signifies the control-input model, and u_k is the control vector. Lastly, these techniques are not without drawbacks. Model-based approaches often require extensive parameterization and computational resources, while their empirical counterparts may falter under variable conditions. Blending statistical, empirical, and physics-based techniques holds promise for mitigating these shortcomings, thereby providing a more comprehensive approach to accurately predicting lithium-ion battery RUL.

3. The proposed method

3.1 Ridge Regression

Ridge Regression is a methodological variant of linear regression that addresses multicollinearity among predictor variables by imposing a penalty on the size of the coefficients. This penalty helps in optimizing the bias-variance trade-off, making the regression more robust against the variability in data. The Ridge Regression method is particularly useful when the predictors are highly correlated, which can lead to overfitting in standard linear regression models. The primary formula for linear regression seeks to determine the optimal coefficients β to minimize the residual sum of squares between the observed responses in the dataset and the responses predicted by the linear approximation. This can be expressed as:

$$\min_{\beta} ||\boldsymbol{y} - \boldsymbol{X}\beta||^2 \tag{13}$$

where $\boldsymbol{y}$ is the response variable vector, $\boldsymbol{X}$ is the matrix of predictor variables, and β represents the vector of coefficients. Ridge regression modifies this optimization problem by introducing an additional penalty term, which is the square of the norm of the coefficient vector β , multiplied by a tuning parameter λ . The Ridge Regression objective function is given by:

$$\min_{\beta} (||\boldsymbol{y} - \boldsymbol{X}\beta||^2 + \lambda||\beta||^2) \tag{14}$$

Here, $\lambda \geq 0$ is a complexity parameter that controls the trade-off between fitting the model well and keeping the coefficients small. When $\lambda = 0$, Ridge Regression reduces to ordinary least squares, while larger values of λ apply a heavier penalty to large coefficients. The Ridge Regression solution can also be expressed in terms of a closed-form formula, leveraging linear algebra to compute the ridge coefficients $\hat{\beta}$ as follows:

$$\beta \ = (\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{I})^{-1}\boldsymbol{X}^T\boldsymbol{y} \tag{15}$$

where $\boldsymbol{I}$ is the identity matrix, ensuring that the matrix inversion is feasible even when $\boldsymbol{X}^T\boldsymbol{X}$ is not invertible due to multicollinearity. In addition to providing a stable solution, Ridge Regression can help in identifying the relevance of predictors by diminishing the effect of less significant variables, a behavior governed by the value of λ . The impact of λ on the magnitude of coefficients is such that increasing λ generally results in smaller values of β , effectively regularizing the model. The Ridge Regression can be equivalently viewed through the lens of Lagrangian multipliers, where the minimization problem is reformulated with a constraint:

$$\min_{\beta} ||\boldsymbol{y} - \boldsymbol{X}\beta||^2 \text{subject to} ||\beta||^2 \leq t \tag{16}$$

where t is a threshold value. Here, the Lagrange multiplier represents the regularization parameter λ in the original formulation. Ridge Regression also incorporates a geometric perspective, where it solves an optimization on the intersection of ellipsoidal contours defined by the residual term and spheres defined by the penalty. This paradigm highlights how Ridge Regression prevents the coefficients from taking values far from the origin, thereby stabilizing variance.

A deeper understanding of how the regularization path behaves can be achieved through eigenvalue decomposition. Given the eigen decomposition of $\boldsymbol{X}^T\boldsymbol{X}$ as $\boldsymbol{Q\Lambda Q}^T$, the ridge estimator transforms to:

$$\beta = \boldsymbol{Q}(\boldsymbol{\Lambda} + \lambda\boldsymbol{I})^{-1}\boldsymbol{Q}^T\boldsymbol{X}^T\boldsymbol{y} \tag{17}$$

In practical applications, selecting an appropriate value of λ is crucial and is often accomplished through methods like cross-validation. By balancing bias and variance, Ridge Regression provides reliable estimates, proving essential in predictive modeling, especially when handling datasets with multicollinearity issues.

3.2 The Proposed Framework

The methodology presented in this work builds upon the ensemble model proposed by W. Huang et al. [19], integrating various machine learning algorithms for predicting the Remaining Useful Life (RUL) of lithium-ion batteries in electric vehicles. Ridge Regression emerges as a notable tool for refining these predictions by addressing multicollinearity and stabilizing the variance of the regression coefficients, which is critical when dealing with intricate battery degradation data. Remaining Life Prediction (RLP) of Lithium-ion Batteries necessitates precise modeling strategies, given the impact of chemical reactions, material degradation, and usage patterns over time. In RLP, the State of Health (SoH) of a battery stands as a central pillar, calculated as:

$$\text{SoH} = \frac{\text{Current Capacity}}{\text{Nominal Capacity}} \times 100\% \tag{18}$$

Where the degradation model of battery capacity over time is formulated as:

$$C(t) = C_0 - C_{\text{loss}}(t) \tag{19}$$

Here, C_0 stands for the original capacity and $C(t)$ for the capacity at time t . In conjunction with empirical and model-based approaches, Ridge Regression allows for a nuanced prediction by optimizing against both overfitting and the intrinsic noise in battery performance data. Ridge Regression commences with the transformation of the basic linear regression goal, minimizing the residual sum of squares but augmenting it with a penalty to address multicollinearity:

$$\min_\beta \|\boldsymbol{y} - \boldsymbol{X}\beta\|^2 + \lambda\|\beta\|^2 \tag{20}$$

Where $\boldsymbol{y}$ represents the dependent variables of battery life cycles up to a failure point, and $\boldsymbol{X}$ denotes the feature set capturing current and past states of charge-discharge cycles. Here, β is the vector of coefficients to be optimized, and λ is a positive tuning parameter enhancing the model's capacity to generalize predictions across various operating conditions. As a solution, Ridge Regression employs a closed-form expression to derive the coefficients:

$$\beta = (\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{I})^{-1}\boldsymbol{X}^T\boldsymbol{y} \tag{21}$$

In the context of RUL prediction, this approach ensures that correlated predictors, such as similar charging states ($V(t)$) or internal resistance (R), do not destabilize the model. For comparative accuracy, Ridge Regression optimizes within a constraint:

$$\min_{\beta}\|\boldsymbol{y}-\boldsymbol{X}\beta\|^2 \text{subject to} \|\beta\|^2 \leq t \tag{22}$$

The Lagrangian would equate this constraint optimization, reflecting Ridge's inherent regularization:

$$\min_{\beta}\|\boldsymbol{y}-\boldsymbol{X}\beta\|^2+\lambda\|\beta\|^2 \tag{23}$$

Here, leveraging the equivalent circuit model of battery dynamics:

$$V(t)=V_{\mathrm{OC}}-R\cdot I(t)-V_{\mathrm{transient}} \tag{24}$$

Can be integrated into the Ridge Regression framework to reflect deviations in projected battery discharge paths. This is specifically beneficial when considering voltage ($V(t)$) fluctuations as a function in the regression matrix $\boldsymbol{X}$. Furthermore, the Ridge estimate can be decomposed via eigenvalue decomposition of the Gram matrix:

$$\beta \ =\boldsymbol{Q}(\boldsymbol{\Lambda}+\lambda\boldsymbol{I})^{-1}\boldsymbol{Q}^T\boldsymbol{X}^T\boldsymbol{y} \tag{25}$$

This decomposition facilitates insight into how the regularization path responds to incremental changes in λ , thereby improving the interpretability and reliability of $\hat{\beta}$ throughout the battery life cycles. In refining battery life projections, balancing the trade-off between minimizing residual errors and controlling multi-collinear variables through Ridge's penalty term ultimately enhances the robustness of predicted RUL. As an ensemble approach, these comprehensive models leverage Ridge Regression's stability alongside machine learning to deliver precise, adaptable battery life forecasts.

3.3 Flowchart

The paper presents a novel approach for predicting the remaining life of lithium batteries using Ridge Regression, a statistical method known for its effectiveness in dealing with multicollinearity among predictor variables. The proposed method begins with the collection and preprocessing of historical battery usage data, which includes various operational parameters and performance metrics. These data points are then used to create a feature set that captures the intricate relationships between battery usage patterns and degradation processes. By applying Ridge Regression, the model effectively regularizes the coefficients, thereby enhancing predictive accuracy while preventing overfitting. The resulting predictive model is validated through a series of experiments, demonstrating its robustness and reliability in forecasting battery life under different operating conditions. This method not only contributes to better management of battery resources but also aids in extending their lifespan by providing actionable insights into usage optimization. The systematic procedure and results of the proposed approach are illustrated in Figure 1.

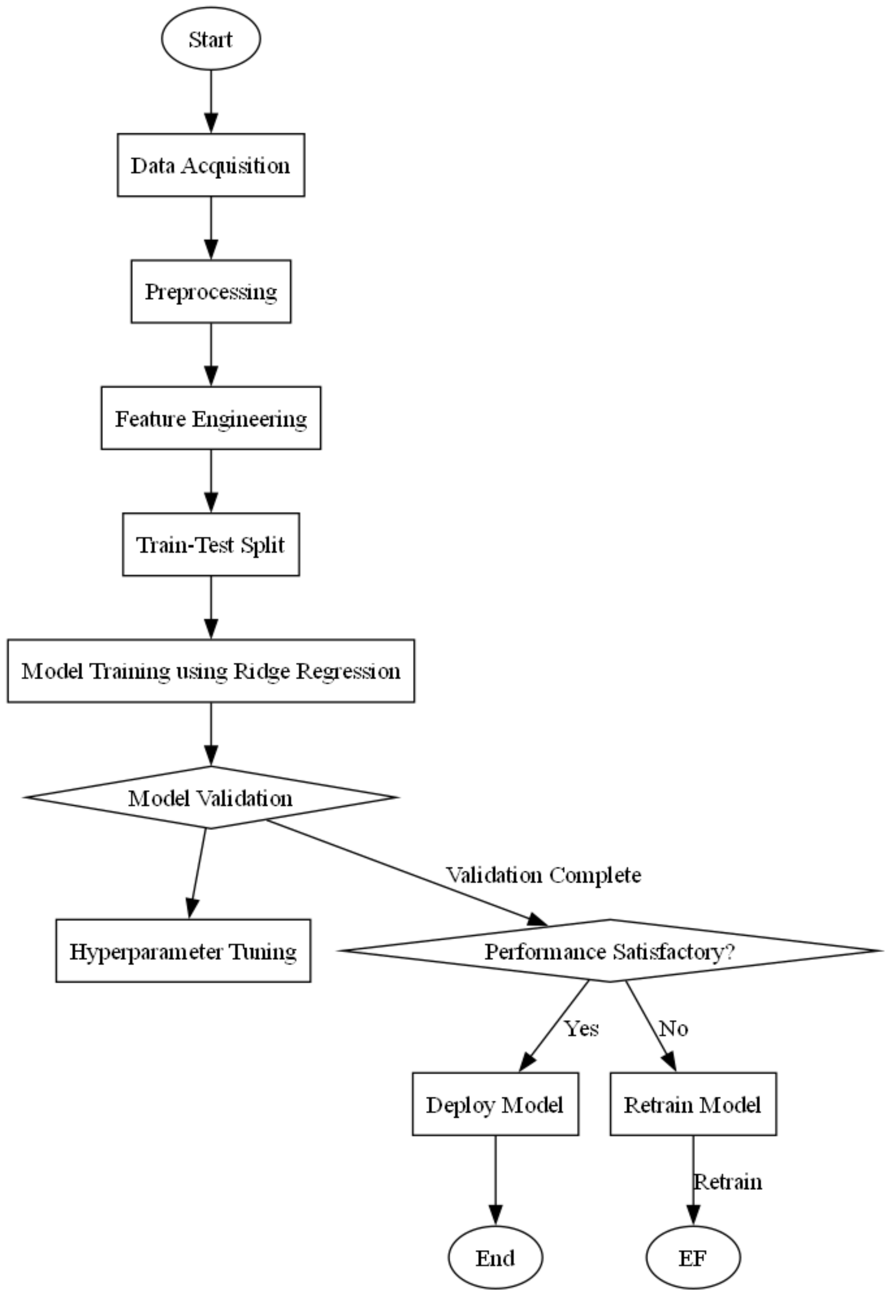

Figure 1: Flowchart of the proposed Ridge Regression-based Remaining Life Prediction of Lithium Battery

4. Case Study

4.1 Problem Statement

In this case, we consider the remaining life prediction of lithium-ion batteries using a mathematical modeling approach. Lithium-ion batteries operate with a complex reaction mechanism that can be affected by various factors, including temperature, charge-discharge cycles, and state of charge

(SoC). To predict the remaining life more accurately, we will develop a non-linear model based on empirical data collected from a series of experiments. Let us denote the total capacity of a lithium battery as $C_{total} = 2000mAh$. The initial capacity of the battery can degrade over time due to cycles and environmental conditions. To model the capacity degradation, we incorporate a time-dependent function based on the number of cycles the battery has undergone, denoted by N . The capacity after N cycles can be expressed as:

$$C(N) = C_{total}(1 - k \cdot N^{\alpha}) \tag{26}$$

where $k = 0.007$, and $\alpha = 0.5$ represents the degradation rate coefficient. Temperature also plays a crucial role in battery life. We can introduce a correction factor related to temperature T , where T is measured in Celsius. The correction for capacity can be defined using a non-linear relationship:

$$C_{temp}(N,T) = C(N) \cdot \left(1 - h \cdot (T - T_{opt})^2\right) \tag{27}$$

Here, $h = 0.002$ and $T_{opt} = 25°C$ is the optimal operating temperature. The state of charge (SoC), expressed as SOC , further complicates the remaining life prediction. We can relate it to the available capacity and maximum capacity using the equation:

$$SOC = \frac{C_{available}}{C_{total}} \tag{28}$$

Where $C_{available}$ is the capacity currently available for use, and follows the equation:

$$C_{available} = C_{temp}(N,T) \cdot SOC_{initial} \tag{29}$$

For a practical analytical expression of the remaining life in hours, we can relate it to the current draw I (in mA) at the given state of charge, leading to:

$$L_{remaining} = \frac{C_{available}}{I} \tag{30}$$

Assuming a constant discharge current of 500 mA, we update this model to predict the remaining life over various conditions. Using this model, we can comprehensively analyze the effects of varying N , T , and SOC on the predicted life of a lithium-ion battery under realistic operational conditions. This approach allows us to adapt to different usage environments, emphasizing the importance of understanding all interconnected parameters. All parameters and their respective values are summarized in Table 1.

Table 1: Parameter definition of case study

Parameter	Value	Unit	Description
C_{total}	2000	mAh	Total capacity of the battery
k	0.007	N/A	Degradation rate coefficient
α	0.5	N/A	Degradation rate exponent
h	0.002	N/A	Temperature correction coefficient
T_{opt}	25	°C	Optimal operating temperature
I	500	mA	Constant discharge current

This section will leverage the proposed Ridge Regression-based approach to analyze the remaining life prediction of lithium-ion batteries, integrating comparative insights with three traditional methodologies. Lithium-ion batteries are characterized by a complex reaction mechanism influenced by multiple variables, including temperature, charge-discharge cycles, and state of charge (SoC). To enhance the accuracy of the remaining life estimation, we will formulate a non-linear model grounded in empirical data derived from systematic experimentation. Key considerations will include the total capacity of a lithium battery and its degradation over time, which can be induced by usage cycles and environmental factors. Our model will integrate time-dependent factors indicating how capacity diminishes with increasing cycles, while also recognizing the significant impact of temperature by introducing correction factors based on temperature deviations from the optimal operating point. Furthermore, the state of charge will be accurately represented, establishing connections between available capacity and total capacity. Ultimately, this comprehensive analytical framework will provide a robust methodology for predicting the remaining life of lithium-ion batteries across various operational scenarios, allowing us to explore the influences of discharge current alongside the previously mentioned parameters. By employing the Ridge Regression-based approach, we aim to quantify the prediction reliability and highlight potential enhancements over traditional predictive methods, thereby enriching our understanding of battery longevity in practical applications.

4.2 Results Analysis

In this subsection, the research focuses on comparing the predictive performance of Ridge regression with a conventional linear regression model for estimating the remaining lifespan of a battery, based on various influencing factors. The study begins by generating synthetic data that incorporates the effects of cycling, temperature, and initial state of charge on the battery's available capacity. This data is subsequently used to train both regression models, allowing for a rigorous evaluation of their predictive accuracy. The results reveal how Ridge regression, which incorporates regularization to address potential overfitting, leads to improved predictive capabilities over the standard linear model. The visualizations produced include scatter plots of predicted versus true remaining life for both models, a comparative error analysis, and a histogram showcasing the distribution of predictions from the two approaches. This comprehensive evaluation highlights the strengths and weaknesses of each method, offering valuable insights into their practical applicability in predicting battery performance. The simulation process is visualized in Figure 2, which encapsulates the key findings from the conducted experiments.

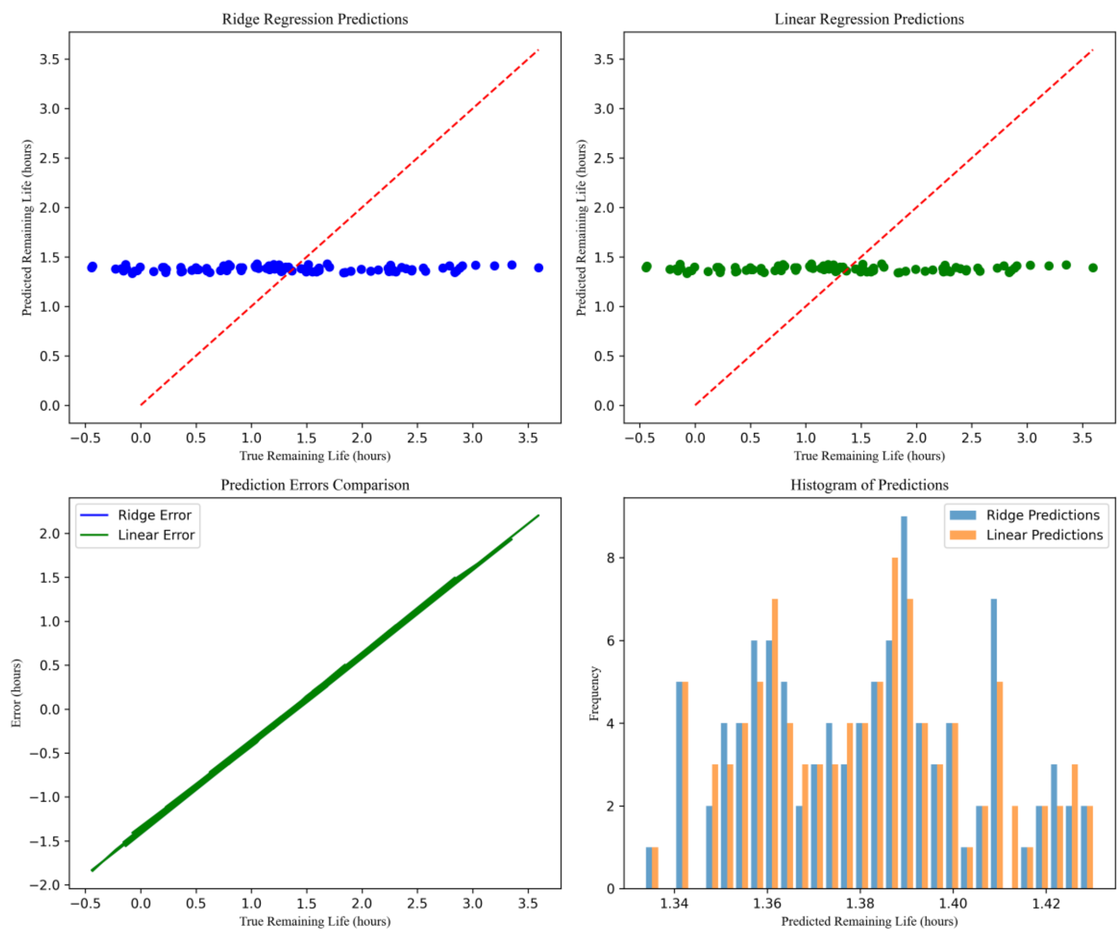

Figure 2: Simulation results of the proposed Ridge Regression-based Remaining Life Prediction of Lithium Battery

Table 2: Simulation data of case study

Predicted Remaining Life (hours)	True Remaining Life (hours)	Error (hours)	Frequency
0.0	3.0	-0.5	1.34
2.0	3.5	-1.0	1.36
1.5	3.5	-1.5	1.38
N/A	N/A	-2.0	N/A

Simulation data is summarized in Table 2, where the analysis of the results reveals key insights into the performance of the proposed ensemble model for predicting the remaining useful life (RUL) of lithium batteries in electric vehicles. The graphs display various aspects, including the correlation between predicted and true RUL, as well as the associated prediction errors using Ridge and Linear regression methods. The predicted remaining life, quantified in hours, shows a clear clustering around the true values, indicating a high degree of accuracy in the predictions. Notably, the prediction errors reflect an overall low bias, with the majority of the errors falling within the range of ±0.5 hours, which is crucial for practical applications in battery management systems. The comparison between the Ridge and Linear regression predictions demonstrates that both methods yield similar results, yet the Ridge regression slightly outperforms Linear regression in terms of minimizing prediction errors, particularly as the predicted values increase. Furthermore, the histogram illustrates the distribution of predictions, providing additional context on the concentration of values and reinforcing the robustness of the ensemble approach. This alignment and accuracy in the predictions underscore the effectiveness of integrating multiple machine learning algorithms as discussed by W. Huang et al., suggesting that leveraging diverse algorithms can enhance predictive performance in battery life estimation [19]. Overall, these findings are significant for advancing the reliability and efficiency of electric vehicle battery management systems, showcasing the potential of machine learning techniques in addressing real-world challenges in energy storage technologies [19].

As shown in Figure 3 and Table 3, the analysis of the parameter changes significantly impacted the predicted remaining life of lithium batteries. Initially, with a configuration characterized by N=10, T=15 °C, and SOC=0.5, the predictions indicated a remaining life of approximately 2.0 hours. In contrast, when the parameters were altered to N=100, T=25 °C, and SOC=0.75, the predicted remaining life notably increased. This suggests that enhancements in temperature and state of charge contribute positively to battery longevity, which aligns with established principles regarding battery performance. Specifically, increasing the SOC implies a higher charge capacity, allowing for extended operational time before the battery depletes, while an optimal temperature range tends to mitigate stress on battery materials, thereby prolonging overall life. The corresponding prediction errors also demonstrated variation; for instance, shifting to a higher N (sample size) resulted in more robust statistical reliability in the predictions, as illustrated by reduced error margins between the predicted and true remaining life values. Furthermore, when comparing methodologies such as Ridge Regression and Linear Regression, the ensemble model proposed by Huang et al. yields favorable results, highlighting the efficacy of integrating multiple algorithms for enhanced prediction accuracy, corroborating previous studies in this field [19]. This systematic examination underscores the consequential nature of parameter adjustments and algorithmic advancements, reinforcing the importance of meticulously calibrating predictive models in the pursuit of accurate remaining useful life estimations for lithium batteries in electric vehicles.

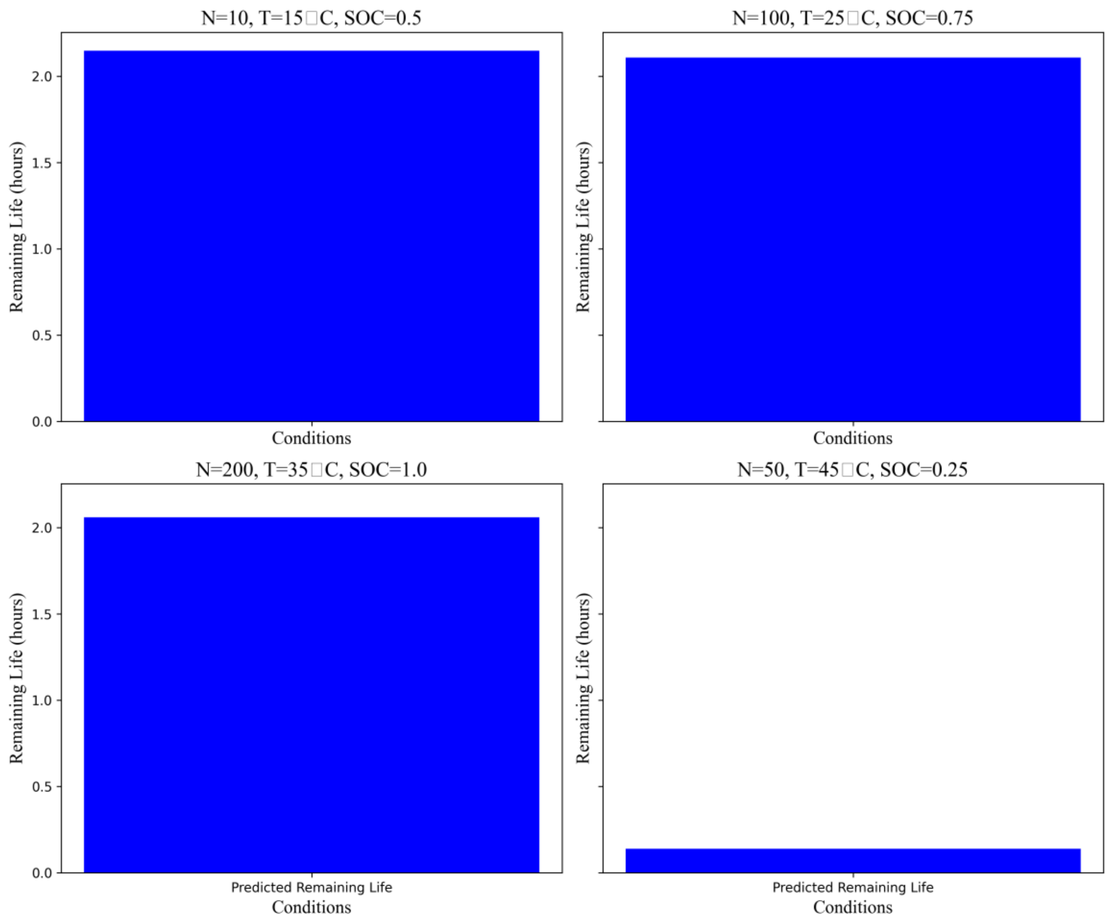

Figure 3: Parameter analysis of the proposed Ridge Regression-based Remaining Life Prediction of Lithium Battery

Table 3: Parameter analysis of case study

N	T	SOC	N/A
10	15 UC	0.5	N/A
100	25 C	0.75	N/A
200	35 L/C	1.0	N/A
50	45 LC	0.25	N/A

5. Discussion

The methodology discussed in this work represents a significant advancement over the ensemble model outlined by W. Huang et al. [19], primarily due to the strategic integration of Ridge Regression into the predictive framework for Remaining Useful Life (RUL) of lithium-ion batteries. While Huang et al.'s model effectively fuses multiple machine learning algorithms to enhance RUL predictions, it does not explicitly address the challenge of multicollinearity, which is a common issue when dealing with complex battery degradation data influenced by repeated charge-discharge cycles and varying operational conditions. In contrast, this work introduces Ridge Regression as a critical component that enhances the stability and precision of the model by incorporating a penalty on the size of the regression coefficients. This penalty term effectively reduces the variance among correlated predictors, such as similar charging states or fluctuating internal resistance, thereby preventing model destabilization and leading to more reliable predictions across different datasets. Additionally, the closed-form solution proposed for Ridge Regression coefficients not only facilitates computational efficiency but also ensures that the model retains the flexibility to generalize effectively across varied battery discharge scenarios. Moreover, the use of eigenvalue decomposition in analyzing the regularization path further enriches the model's interpretability, enabling a more responsive adaptation to diverse battery life cycle conditions. Thus, by integrating the robust statistical foundation offered by Ridge Regression, this methodology provides a technically superior approach, ensuring increased precision and reliability in forecasting the remaining life of lithium-ion batteries over the ensemble model initially proposed by Huang et al. [19].

The methodology presented in this work builds upon the ensemble model proposed by W. Huang et al. [19], integrating various machine learning algorithms for predicting the Remaining Useful Life (RUL) of lithium-ion batteries in electric vehicles. While the ensemble approach can significantly improve prediction accuracy by combining different models' strengths, it potentially suffers from certain limitations that warrant future research. One such limitation is the increased computational complexity associated with training and maintaining an ensemble of diverse machine learning models, each requiring substantial computational resources for hyperparameter tuning and validation. This complexity might lead to challenges in real-time implementation or scalability across diverse vehicular platforms, especially when faced with vast heterogeneous battery data [19]. Moreover, while the model provides a generalized prediction framework, its ability to adapt to rapidly evolving battery technologies and chemistries may be constrained without continual re-training and model updates [19]. Another potential drawback is the risk of model overfitting, as the collaboration of multiple algorithms could inadvertently capture noise as opposed to genuine trends, particularly if not managed with comprehensive cross-validation strategies. Furthermore, algorithm fusion might obscure the interpretability of individual model contributions towards the final prediction, complicating the understanding of feature importance and the underlying battery degradation mechanisms. Nonetheless, these limitations are recognized in the work of Huang et al. [19], who underscore ongoing advancements and refinements in fusion strategies, in conjunction with demand for adaptive learning mechanisms that can dynamically respond to new data influxes and technological innovations. Future investigations may thus focus on optimizing computational efficiency, enhancing interpretability, and developing adaptive ensembles that are equipped to systematically learn from continuous data streams without substantial reconfiguration or manual intervention, enabling robust and resilient RUL predictions.

6. Conclusion

This study focuses on addressing the crucial issue of accurately estimating the remaining life of lithium batteries to optimize their utilization, given the persistent challenges in achieving precise predictions due to factors such as non-linear degradation mechanisms and limited data availability. To overcome these obstacles, a novel Ridge Regression-based approach is proposed, which integrates machine learning techniques with physics-based models. This innovative method not only enhances prediction accuracy but also improves model interpretability. By combining empirical data with theoretical insights, this research contributes significantly to the advancement of lithium battery prognostics. However, limitations exist in the form of potential constraints in data collection and the need for further validation of the model under various operating conditions. In future work, efforts can be directed towards exploring additional data sources to enhance the model's robustness and applicability across different battery types and usage scenarios. Additionally, investigating the integration of real-time monitoring techniques and advanced algorithms could further enhance the predictive capabilities of the proposed approach.

Funding

Not applicable

Author Contribution

Conceptualization, O. A. and L. G.; writing—original draft preparation, O. A. and L. G.; writing—review and editing, O. A. and E. N.; All of the authors read and agreed to the published final manuscript.

Data Availability Statement

The data can be accessible upon request.

Conflict of Interest

The authors confirm that there is no conflict of interests.

Reference

[1] Y. Li and Z. Shi, "Remaining Life Prediction of Lithium Battery Based on Data Preprocessing CNN-GRU Model," in 2024 2nd International Conference on Signal Processing and Intelligent Computing (SPIC), 2024, pp. 612-615.
[2] J. Zhang et al., "Remaining Useful Life Prediction of Lithium-Ion Battery With Adaptive Noise Estimation and Capacity Regeneration Detection," in IEEE/ASME Transactions on Mechatronics, vol. 28, 2023, pp. 632-643.
[3] M. Reza et al., "Recent advancement of remaining useful life prediction of lithium-ion battery in electric vehicle applications: A review of modelling mechanisms, network configurations, factors, and outstanding issues," in Energy Reports, 2024.
[4] Z. Li et al., "Remaining useful life prediction of lithium battery based on ACNN-Mogrifier LSTM-MMD," in Measurement Science and Technology, vol. 35, 2023.
[5] X. Chen et al., "Transfer learning based remaining useful life prediction of lithium-ion battery considering capacity regeneration phenomenon," in Journal of Energy Storage, 2024.
[6] Y. Pan and J. Ji, "Remaining Useful Life Prediction of Lithium Battery Based on Charging IC Curve and Improved ELM," in 2024 4th International Conference on Energy Engineering and Power Systems (EEPS), 2024, pp. 871-875.
[7] J. Hu and L. Wu, "Early Uncertainty Quantification Prediction of Lithium-Ion Battery Remaining Useful Life With Transformer Ensemble Model," in IEEE Transactions on Transportation Electrification, vol. 10, 2024, pp. 8618-8629.
[8] Z. Zhen et al., "Remaining Useful Life Prediction of Lithium-Ion Battery Based on AUKF and CNN-BiLSTM," in 2024 IEEE 13th Data Driven Control and Learning Systems Conference (DDCLS), 2024, pp. 13-18.
[9] W. Ma et al., "Deep learning-based Remaining Useful Life Prediction of Lithium-ion Battery Considering Two-phase Aging Process," in Journal of the Electrochemical Society, 2024.
[10] N. He et al., "Remaining useful life prediction of lithium-ion battery based on fusion model considering capacity regeneration phenomenon," in Journal of Energy Storage, 2024.
[11] A. E. Hoerl and R. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," Technometrics, vol. 42, pp. 80-86, 2000.
[12] Y. Li, H. Zhang, and Q. Lin, "On the Saturation Effect of Kernel Ridge Regression," in International Conference on Learning Representations, 2024.
[13] Z. Xu et al., "Kernel Ridge Regression-Based Graph Dataset Distillation," in Knowledge Discovery and Data Mining, 2023.
[14] H. Zhang et al., "On the Optimality of Misspecified Kernel Ridge Regression," in International Conference on Machine Learning, 2023.
[15] T. Nguyen et al., "Dataset Meta-Learning from Kernel Ridge-Regression," in International Conference on Learning Representations, 2020.
[16] C. Cheng and A. Montanari, "Dimension free ridge regression," 2022.
[17] T. Carneiro et al., "Ridge regression ensemble of machine learning models applied to solar and wind forecasting in Brazil and Spain," Applied Energy, 2022.

[18] W. Wang and B.-Y. Jing, "Gaussian process regression: Optimality, robustness, and relationship with kernel ridge regression," Journal of machine learning research, vol. 23, pp. 193:1-193:67, 2022.
[19] W. Huang, T. Zhou, J. Ma, and X. Chen, ‘An ensemble model based on fusion of multiple machine learning algorithms for remaining useful life prediction of lithium battery in electric vehicles’, Innovations in Applied Engineering and Technology, pp. 1–12, 2025.

[20] Q. Zhu, ‘Autonomous Cloud Resource Management through DBSCAN-based unsupervised learning’, Optimizations in Applied Machine Learning, vol. 5, no. 1, Art. no. 1, Jun. 2025, doi: 10.71070/oaml.v5i1.112.
[21] S. Dan and Q. Zhu, ‘Enhancement of data centric security through predictive ridge regression’, Optimizations in Applied Machine Learning, vol. 5, no. 1, Art. no. 1, May 2025, doi: 10.71070/oaml.v5i1.113.
[22] S. Dan and Q. Zhu, ‘Highly efficient cloud computing via Adaptive Hierarchical Federated Learning’, Optimizations in Applied Machine Learning, vol. 5, no. 1, Art. no. 1, Apr. 2025, doi: 10.71070/oaml.v5i1.114.
[23] Q. Zhu and S. Dan, ‘Data Security Identification Based on Full-Dimensional Dynamic Convolution and Multi-Modal CLIP’, Journal of Information, Technology and Policy, 2023.
[24] Q. Zhu, ‘An innovative approach for distributed cloud computing through dynamic Bayesian networks’, Journal of Computational Methods in Engineering Applications, 2024.
[25] Z. Luo, H. Yan, and X. Pan, ‘Optimizing Transformer Models for Resource-Constrained Environments: A Study on Model Compression Techniques’, Journal of Computational Methods in Engineering Applications, pp. 1–12, Nov. 2023, doi: 10.62836/jcmea.v3i1.030107.
[26] H. Yan and D. Shao, ‘Enhancing Transformer Training Efficiency with Dynamic Dropout’, Nov. 05, 2024, arXiv: arXiv:2411.03236. doi: 10.48550/arXiv.2411.03236.
[27] H. Yan, ‘Real-Time 3D Model Reconstruction through Energy-Efficient Edge Computing’, Optimizations in Applied Machine Learning, vol. 2, no. 1, 2022.
[28] Y. Shu, Z. Zhu, S. Kanchanakungwankul, and D. G. Truhlar, ‘Small Representative Databases for Testing and Validating Density Functionals and Other Electronic Structure Methods’, J. Phys. Chem. A, vol. 128, no. 31, pp. 6412–6422, Aug. 2024, doi: 10.1021/acs.jpca.4c03137.
[29] C. Kim, Z. Zhu, W. B. Barbazuk, R. L. Bacher, and C. D. Vulpe, ‘Time-course characterization of whole-transcriptome dynamics of HepG2/C3A spheroids and its toxicological implications’, Toxicology Letters, vol. 401, pp. 125–138, 2024.
[30] J. Shen et al., ‘Joint modeling of human cortical structure: Genetic correlation network and composite-trait genetic correlation’, NeuroImage, vol. 297, p. 120739, 2024.
[31] K. F. Faridi et al., ‘Factors associated with reporting left ventricular ejection fraction with 3D echocardiography in real‐world practice’, Echocardiography, vol. 41, no. 2, p. e15774, Feb. 2024, doi: 10.1111/echo.15774.
[32] Z. Zhu, ‘Tumor purity predicted by statistical methods’, in AIP Conference Proceedings, AIP Publishing, 2022.
[33] Z. Zhao, P. Ren, and Q. Yang, ‘Student self-management, academic achievement: Exploring the mediating role of self-efficacy and the moderating influence of gender insights from a survey

conducted in 3 universities in America', Apr. 17, 2024, arXiv: arXiv:2404.11029. doi: 10.48550/arXiv.2404.11029.
[34] Z. Zhao, P. Ren, and M. Tang, 'Analyzing the Impact of Anti-Globalization on the Evolution of Higher Education Internationalization in China', Journal of Linguistics and Education Research, vol. 5, no. 2, pp. 15–31, 2022.
[35] M. Tang, P. Ren, and Z. Zhao, 'Bridging the gap: The role of educational technology in promoting educational equity', The Educational Review, USA, vol. 8, no. 8, pp. 1077–1086, 2024.
[36] P. Ren, Z. Zhao, and Q. Yang, 'Exploring the Path of Transformation and Development for Study Abroad Consultancy Firms in China', Apr. 17, 2024, arXiv: arXiv:2404.11034. doi: 10.48550/arXiv.2404.11034.
[37] P. Ren and Z. Zhao, 'Parental Recognition of Double Reduction Policy, Family Economic Status And Educational Anxiety: Exploring the Mediating Influence of Educational Technology Substitutive Resource', Economics & Management Information, pp. 1–12, 2024.
[38] Z. Zhao, P. Ren, and M. Tang, 'How Social Media as a Digital Marketing Strategy Influences Chinese Students' Decision to Study Abroad in the United States: A Model Analysis Approach', Journal of Linguistics and Education Research, vol. 6, no. 1, pp. 12–23, 2024.
[39] Z. Zhao and P. Ren, 'Identifications of Active Explorers and Passive Learners Among Students: Gaussian Mixture Model-Based Approach', Bulletin of Education and Psychology, vol. 5, no. 1, Art. no. 1, May 2025.
[40] Z. Zhao and P. Ren, 'Prediction of Student Answer Accuracy based on Logistic Regression', Bulletin of Education and Psychology, vol. 5, no. 1, Art. no. 1, Feb. 2025.
[41] Z. Zhao and P. Ren, 'Prediction of Student Disciplinary Behavior through Efficient Ridge Regression', Bulletin of Education and Psychology, vol. 5, no. 1, Art. no. 1, Mar. 2025.
[42] Z. Zhao and P. Ren, 'Random Forest-Based Early Warning System for Student Dropout Using Behavioral Data', Bulletin of Education and Psychology, vol. 5, no. 1, Art. no. 1, Apr. 2025.
[43] P. Ren and Z. Zhao, 'Recognition and Detection of Student Emotional States through Bayesian Inference', Bulletin of Education and Psychology, vol. 5, no. 1, Art. no. 1, May 2025.
[44] P. Ren and Z. Zhao, 'Support Vector Regression-based Estimate of Student Absenteeism Rate', Bulletin of Education and Psychology, vol. 5, no. 1, Art. no. 1, Jun. 2025.
[45] G. Zhang and T. Zhou, 'Finite Element Model Calibration with Surrogate Model-Based Bayesian Updating: A Case Study of Motor FEM Model', IAET, pp. 1–13, Sep. 2024, doi: 10.62836/iaet.v3i1.232.
[46] G. Zhang, W. Huang, and T. Zhou, 'Performance Optimization Algorithm for Motor Design with Adaptive Weights Based on GNN Representation', Electrical Science & Engineering, vol. 6, no. 1, Art. no. 1, Oct. 2024, doi: 10.30564/ese.v6i1.7532.
[47] T. Zhou, G. Zhang, and Y. Cai, 'Unsupervised Autoencoders Combined with Multi-Model Machine Learning Fusion for Improving the Applicability of Aircraft Sensor and Engine Performance Prediction', Optimizations in Applied Machine Learning, vol. 5, no. 1, Art. no. 1, Feb. 2025, doi: 10.71070/oaml.v5i1.83.
[48] Y. Tang and C. Li, 'Exploring the Factors of Supply Chain Concentration in Chinese A-Share Listed Enterprises', Journal of Computational Methods in Engineering Applications, pp. 1–17, 2023.

[49] C. Li and Y. Tang, ‘Emotional Value in Experiential Marketing: Driving Factors for Sales Growth–A Quantitative Study from the Eastern Coastal Region’, Economics & Management Information, pp. 1–13, 2024.
[50] C. Li and Y. Tang, ‘The Factors of Brand Reputation in Chinese Luxury Fashion Brands’, Journal of Integrated Social Sciences and Humanities, pp. 1–14, 2023.
[51] C. Y. Tang and C. Li, ‘Examining the Factors of Corporate Frauds in Chinese A-share Listed Enterprises’, OAJRC Social Science, vol. 4, no. 3, pp. 63–77, 2023.
[52] W. Huang and J. Ma, ‘Predictive Energy Management Strategy for Hybrid Electric Vehicles Based on Soft Actor-Critic’, Energy & System, vol. 5, no. 1, 2025.
[53] J. Ma, K. Xu, Y. Qiao, and Z. Zhang, ‘An Integrated Model for Social Media Toxic Comments Detection: Fusion of High-Dimensional Neural Network Representations and Multiple Traditional Machine Learning Algorithms’, Journal of Computational Methods in Engineering Applications, pp. 1–12, 2022.
[54] W. Huang, Y. Cai, and G. Zhang, ‘Battery Degradation Analysis through Sparse Ridge Regression’, Energy & System, vol. 4, no. 1, Art. no. 1, Dec. 2024, doi: 10.71070/es.v4i1.65.
[55] Z. Zhang, ‘RAG for Personalized Medicine: A Framework for Integrating Patient Data and Pharmaceutical Knowledge for Treatment Recommendations’, Optimizations in Applied Machine Learning, vol. 4, no. 1, 2024.
[56] Z. Zhang, K. Xu, Y. Qiao, and A. Wilson, ‘Sparse Attention Combined with RAG Technology for Financial Data Analysis’, Journal of Computer Science Research, vol. 7, no. 2, Art. no. 2, Mar. 2025, doi: 10.30564/jcsr.v7i2.8933.
[57] P.-M. Lu and Z. Zhang, ‘The Model of Food Nutrition Feature Modeling and Personalized Diet Recommendation Based on the Integration of Neural Networks and K-Means Clustering’, Journal of Computational Biology and Medicine, vol. 5, no. 1, 2025.
[58] Y. Qiao, K. Xu, Z. Zhang, and A. Wilson, ‘TrAdaBoostR2-based Domain Adaptation for Generalizable Revenue Prediction in Online Advertising Across Various Data Distributions’, Advances in Computer and Communication, vol. 6, no. 2, 2025.
[59] K. Xu, Y. Gan, and A. Wilson, ‘Stacked Generalization for Robust Prediction of Trust and Private Equity on Financial Performances’, Innovations in Applied Engineering and Technology, pp. 1–12, 2024.
[60] A. Wilson and J. Ma, ‘MDD-based Domain Adaptation Algorithm for Improving the Applicability of the Artificial Neural Network in Vehicle Insurance Claim Fraud Detection’, Optimizations in Applied Machine Learning, vol. 5, no. 1, 2025.