batteries

Article

The Role of Machine Learning in Enhancing Battery Management for Drone Operations: A Focus on SoH Prediction Using Ensemble Learning Techniques

Büşra Çetinus [1], Saadin Oyucu [2], Ahmet Aksöz [3] and Emre Biçer [1,*]

[1] Battery Research Laboratory, Faculty of Engineering and Natural Sciences, Sivas University of Science and Technology, 58010 Sivas, Türkiye; 200102036@sivas.edu.tr
[2] Department of Computer Engineering, Faculty of Engineering, Adıyaman University, 02040 Adıyaman, Türkiye; saadinoyucu@adiyaman.edu.tr
[3] Mobilers Team, Sivas Cumhuriyet University, 58050 Sivas, Türkiye; aaksoz@cumhuriyet.edu.tr
* Correspondence: emre.bicer@sivas.edu.tr

Abstract: This study considers the significance of drones in various civilian applications, emphasizing battery-operated drones and their advantages and limitations, and highlights the importance of energy consumption, battery capacity, and the state of health of batteries in ensuring efficient drone operation and endurance. It also describes a robust testing methodology used to determine battery SoH accurately, considering discharge rates and using machine learning algorithms for analysis. Machine learning techniques, including classical regression models and Ensemble Learning methods, were developed and calibrated using experimental UAV data to predict SoH accurately. Evaluation metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) assess model performance, highlighting the balance between model complexity and generalization. The results demonstrated improved SoH predictions with machine learning models, though complexities may lead to overfitting challenges. The transition from simpler regression models to intricate Ensemble Learning methods is meticulously described, including an assessment of each model's strengths and limitations. Among the Ensemble Learning methods, Bagging, GBR, XGBoost, LightGBM, and stacking were studied. The stacking technique demonstrated promising results: for Flight 92 an RMSE of 0.03% and an MAE of 1.64% were observed, while for Flight 129 the RMSE was 0.66% and the MAE stood at 1.46%.

Keywords: UAV data analysis; machine learning; regression models; Ensemble Learning; Li-ion

Citation: Çetinus, B.; Oyucu, S.; Aksöz, A.; Biçer, E. The Role of Machine Learning in Enhancing Battery Management for Drone Operations: A Focus on SoH Prediction Using Ensemble Learning Techniques. *Batteries* **2024**, *10*, 371. https://doi.org/10.3390/batteries10100371

Academic Editors: Rodolfo Dufo-López and Benben Jiang

Received: 5 July 2024
Revised: 2 October 2024
Accepted: 14 October 2024
Published: 18 October 2024

1. Introduction

"Drone" is a concept that refers to aircraft of various sizes and features that are controlled by remote control, also known as unmanned aerial vehicles (UAV) which are increasingly preferred in various civilian fields; these areas of use include activities such as live monitoring, expansion of wireless internet access areas, remote sensing, search-and-rescue activities, cargo distribution, security and surveillance, precision agriculture, and control of civil structures. Smart UAV technologies are among the important innovations in UAV technology, with advantages such as reducing risks and providing cost savings, especially in areas such as civil construction management [1].

UAVs' ease of deployment, low maintenance costs, superior mobility, and ability to stay airborne for long periods of time make them a viable option for a range of civilian and military applications. These features make UAVs ideal for these tasks [2]. UAVs primarily rely on either internal combustion engines or batteries for propulsion power sources. In this context, we specifically concentrate on battery-operated drones. Battery-operated drones offer several notable advantages, including zero emissions, simplified maintenance, and straightforward controllability. However, they also face limitations, particularly concerning flight range and duration [3].

Energy consumption is one of the key factors of UAVs affecting efficient use, endurance, cost, and emissions. Energy consumption is also related to drone design, environment, drone dynamics, and operations [4]. In addition, the battery pack must be capable of providing sufficient electrical power to match the power consumed by aerodynamic drag, considering losses from the propeller, motor, and motor controller, and additionally the avionic system and other electronics [5]. The capacity of a battery is determined based on its application, and this capacity decreases as the battery ages. Additionally, as the battery ages, its impedance increases, which in turn reduces its ability to provide energy effectively.

The state of health (SoH) of a battery reflects its ability to store and supply energy based on its fundamental conditions, considering the energy and power requirements of the application. SoH serves as a crucial indicator for assessing the level of degradation experienced by a battery over time. Monitoring SoH in the operational processes of UAVs is critical for the successful continuation of missions. The condition of Li-ion batteries requires particular attention in order to ensure that UAVs operate uninterruptedly throughout their mission period and do not pose any safety risk in the air. Therefore, battery SoH estimation is a great challenge for operation endurance due to the aging mechanism of the battery. To tackle this situation, various methods have been developed as experimental techniques and adaptive battery models [6]. The determination of SoH in battery packs directly impacts the endurance and operational limits of UAVs. As batteries age, the operational time of drones decreases, affecting their overall performance. Additionally, because State of Charge (SoC) data are linked to SoH, inaccuracies in SoC measurements might mislead the estimation of remaining capacity, potentially leading to premature shutdown or loss of the drone during operation [7]. Andrioaia et al. discussed the necessity of estimating the Remaining Useful Life (RUL) and predicting the capacity of Li-ion batteries to prevent UAVs from losing autonomy, comparing the performance of three machine learning models—Support Vector Machine for Regression (SVMR), Multiple Linear Regression (MLR), and Random Forest (RF)—to estimate the RUL. This method might be implemented in UAVs' Predictive Maintenance (PdM) systems [8]. In addition, Zhang et al. addressed the issue of limited endurance in UAVs for the estimation of the state of charge (SOC) of Li-Po batteries, maximizing energy utilization and enhancing UAV endurance. They used the extended Kalman filter algorithm, based on an equivalent circuit model, to estimate SOC [9]. Accurate prediction of SoH can ensure UAVs battery lifespan and efficiency. Shibl et al. presented a UAV BMS based on machine learning (ML) models predicting SoC and estimating SoH using battery voltage, current, and ambient temperature data. Deep Neural Networks (DNN) and Long Short-Term Memory (LSTM) were used for SoC prediction, while Random Forest (RF) was used for SoH estimation [10].

In various applications, such as steady-level flight for fixed-wing aircraft or hovering conditions for multi-rotor vehicles, a constant power discharge process represented the actual battery loading; however, in the present study, a robust and diverse testing methodology was employed, involving the testing of the battery under various C rates at ten-second intervals to mimic different operational states of the drone [11]. This involved altering the position of an actual drone every 10 s while logging current and voltage data. By analyzing the discharge current associated with position changes, we determined the battery's behavior across different operational scenarios. Subsequently, these data were utilized to calculate SoH information of the battery. A single cell underwent rigorous testing for 285 cycles using a specialized battery test system, during which capacity data were recorded. The effect of discharge rate on available capacity was considered for variable current discharges. Afterwards, having the capacity values in response to discharge current for each cycle, machine learning algorithms were then applied to the data to calculate and assess the SoH information of the battery accurately to determine an efficient source of SoH information (Tables S1 and S2).

Our research group has recently published studies utilizing machine learning for estimating the SoH in electric vehicles (EVs). Specifically, it introduced a methodology for comparative analysis, focusing on classical and deep learning approaches, and assessed improvements made to LSTM and Bi-LSTM methods using evaluation metrics such as MSE, MAE, RMSE, and R-squared. The objective of the study was to drive advancements in EV technology by predicting Li-ion Battery (LIB) performance [12]. Additionally, another study was conducted as an assessment of traditional ML and advanced deep learning models for predictive modeling, encompassing techniques such as Random Forests, XGBoost, Elastic Net, MLP, CNNs, and RNNs, as well as hybrid models such as RNN-LSTM and CNN-LSTM. The findings underscored the potential of deep learning-based hybrid models in improving battery SoH assessments [7]. Accurately managing the capacity of L-ion batteries is crucial for enhancing the cost-effectiveness of large-scale energy storage systems. Therefore, in a study by Oyucu et al., various machine learning models—AdaBoost, gradient boosting, XGBoost, LightGBM, CatBoost, and Ensemble Learning—were used to predict the discharge capacity of LIBs. The study also utilized SHAP (Shapley additive explanations) values within the explainable artificial intelligence (XAI) framework to analyze key features affecting predictions [13].

Apart from the previous studies, in the present study, classical regression models such as Linear Regression (LR), Lasso, ElasticNet, and Ridge Regression (RR) were developed to predict the SoH of UAVs [14,15]. These models were calibrated using comprehensive experimental data obtained from UAV operations to interpret battery aging and operational performance. Following these initial models, Ensemble Learning techniques including Bagging, Boosting, and Stacking were employed to enhance the accuracy of SoH predictions [16]. These methods combined predictions from various regression models to achieve more robust and accurate predictions, thereby playing a critical role in the operational continuity and reliability of UAV batteries.

Two different evaluation metrics were used to measure the accuracy of SoH predictions for UAV batteries using the developed regression and Ensemble Learning models. The first evaluation metric is Root Mean Squared Error (RMSE), indicating how much deviation the model's predictions showed from the actual values by calculating the square root of the mean of squared errors, providing a broad perspective on the magnitude of the model's errors. The second metric was Mean Absolute Error (MAE), which was calculated by taking the mean of the absolute values of errors, presenting the average of the model's errors more directly. By evaluating the model's performance from different perspectives using these two metrics, it was demonstrated how reliable the developed prediction models were.

The results obtained indicated improvements among different models; however, they also demonstrated a decrease in success rates in cases where model complexity increased. This phenomenon was particularly pronounced in models using more complex Ensemble Learning techniques. As model complexity increases, theoretically, the model is expected to learn the data better and generalize; however, in some cases, overfitting issues have arisen, negatively impacting the model's performance on independent datasets. The results have shown that controlling and balancing complexity during the model development process is a significant factor directly affecting success rates. Therefore, while increasing model complexity, the importance of integrating techniques that maintain generalization ability has been emphasized.

This article evaluates the feasibility of machine learning techniques for developing and validating SoH predictions for UAV batteries. It extensively discusses the data collection and processing methods used for estimating battery SoH values. Factors such as temperature and cycle that affect battery performance are examined, and how these factors might be integrated into SoH prediction models is emphasized. The transition from simple regression models to more complex Ensemble Learning methods is described. Each model's implementation and the comparative analysis of their results identifies the advantages and limitations of the models.

Section 3 evaluates the prediction results obtained through different machine learning models using RMSE and MAE metrics to measure model performance. Section 5 focuses on the integration potential of the results into practical scenarios and discusses the limitations encountered in using machine learning methods for UAV battery SoH prediction. The article concludes with suggestions for future research, including the development of new algorithms to improve prediction model accuracy, experiments on more diverse datasets, and enhancing the applicability of SoH prediction models in real-world conditions.

The novel contributions of this study can be listed as below:

- Integration of diverse machine learning techniques: The combination of classical regression models and advanced Ensemble Learning techniques (Bagging, Boosting, and Stacking) is utilized to obtain accurate battery SoH estimation. In particular, the stacking of models is deployed for improving prediction accuracy and robustness by leveraging the strengths of several models.
- Robust testing approach: Within this study, a robust test method is designed for battery SoH estimation on various discharge rates and real operational scenarios. These scenarios mainly simulate the actual usage conditions of drones more accurately than standard tests by changing the drone's operational conditions every 10 s to measure battery behavior in dynamic scenarios. Consequently, this testing process crucially improves the feasibility of the SoH predictions to real-world drone environments.
- Evaluation with real UAV data: For the evaluation process, this study presents a real-world dataset providing calibration and experimental validation of the machine learning models under realistic operational factors, which is critical for practical applications in drone battery management.

2. Materials and Methods

This study aims to estimate the SoH for a UAV by using a literature dataset [13]. Rodrigues and co-workers conducted a study using DJI Matrice 100 Quadcopter (DJI, Shenzhen, China) to work with various parameters such as the battery voltage and current, altitude, wind conditions, wind speed, angular velocity, acceleration, and load within a specific area. In the context of this article, we used some of the datasets.

2.1. Battery Tests and Data Collection

The dataset in this article included not only voltage and current logs but also GPS data, range, altitude, acceleration, aerial time, wind speed, etc. [17]. A total of 209 flights were conducted in the study, and characteristics such as range and altitude were analyzed for these flights. However, the SoH was not shared after each flight. In this article, flights 92 and 129 were considered as having the longest range declared, and current and voltage information was utilized. Based on these data, each current value discharged from the battery was calculated to create a discharge algorithm. Using this algorithm, a test methodology with 285 cycles was conducted on high-energy 18650 batteries. The discharge algorithm, detailed in the supporting information, involved changing the current every ten seconds, repeated approximately 167 times throughout the test. The dataset from these 285 cycles, comprising discharge currents, voltage readings, and capacity values, was then used as input data for estimating the SoH by using machine learning techniques. In feature engineering, firstly data samples are extracted with respect to cycle feature values from 1 to 285. Consequently, a histogram based on the current (A) feature is extracted to analyze current values data distribution. Afterwards, a time-series analysis is conducted between the relative time feature and the capacity (Ah) target feature to analyze how the capacity feature values change over relative time feature values. Ultimately, a correlation matrix is generated between the current and capacity features to better understand the relationship information between capacity feature values and current feature values. Additionally, some 'null' valued samples are separated from the data to provide cleaner data as an input to the models.

2.2. Machine Learning Techniques

This study explores the application of machine learning techniques, specifically classical regression models such as LR, Lasso, ElasticNet, and RR, to predict the SoH of UAV batteries. These models were initially calibrated using comprehensive UAV operational data to interpret battery aging and performance. Subsequently, Ensemble Learning methods such as Bagging, Boosting, and Stacking were employed to enhance SoH prediction accuracy by combining predictions from multiple regression models. Evaluation metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) were used to measure model accuracy, highlighting the reliability of the developed prediction models [7].

This study found that while model complexity might improve predictions, overly complex models might suffer from decreased success rates, especially with complex Ensemble Learning techniques, due to potential overfitting issues. The finding emphasizes the importance of controlling complexity to maintain generalization ability and improve success rates. The article discusses data collection, processing methods, and the integration of factors such as temperature and cycle into SoH prediction models. It also evaluates model performance using RMSE and MAE metrics and discusses limitations and future research directions, including the development of new algorithms, diverse dataset experiments, and improving the applicability of SoH prediction models in real-world conditions.

Finally, estimation of the battery capacity is the direct measure for determining the SoH state of the battery. Consequently, as capacity decreases, SoH also decreases, which shows the battery aging and degradation. Therefore, SoH estimation and evaluation are intrinsically accomplished by performing the battery capacity estimation and comparing the estimated battery capacity values with ground truth capacity values in the dataset with the machine learning techniques and evaluation metrics mentioned above.

2.2.1. Linear Regression

LR in the context of cycle, current, and capacity for batteries is a statistical modeling technique used to understand and predict how the number of cycles, the magnitude of discharge current during cycles, and the resulting capacity of the battery are related [18]. LR with cycle, current, and capacity for batteries involves modeling the relationship between these variables using a linear equation. The independent variables are cycle and discharge current, representing the number of cycles and the magnitude of discharge current, respectively. The dependent variable is capacity, indicating the amount of charge the battery may hold. By collecting data on cycle currents and corresponding capacities across multiple charge–discharge cycles, LR allowed us to determine the linear relationship between these variables. The LR model provided insights into how changes in cycle and current impacted the capacity of the battery. The resulting regression Equation (1) is given as

$$Capacity = \beta_0 + \beta_1 \times Cycle + \beta_2 \times Current + \epsilon \tag{1}$$

where

- *Capacity* is the battery's capacity;
- *Cycle* the number of cycles;
- *Current* is the discharge current during the cycles;
- β_0 is the intercept, representing the capacity when the cycle current is zero;
- β_1 and β_2 are the coefficients for cycle and current, respectively, indicating how they influence capacity;
- ϵ is the error term.

This equation helps in predicting the capacity of the battery for different cycle currents based on the observed data. LR analysis is valuable for optimizing battery performance, understanding degradation patterns, and designing efficient charging and discharging strategies.

2.2.2. Lasso Regression

Lasso regression is a method that applies L1 norm regularization to independent variables in regression models to approach their coefficients towards zero. This method reduces the complexity of variables in the model and sets the coefficients of insignificant variables to zero, thereby creating simpler and more generalizable models. Lasso stands for "Least Absolute Shrinkage and Selection Operator", referring to its ability to shrink coefficients while performing variable selection [19].

The main difference between LR (Linear Regression) and Lasso lies in Lasso's inclusion of a regularization term during model creation. LR aims to determine coefficients that minimize the relationship between independent and dependent variables, whereas Lasso additionally minimizes the sum of absolute values of coefficients. This additional term makes Lasso more resistant to overfitting and enables it to build models with fewer variables. These characteristics make Lasso more advantageous, especially when the number of variables exceeds the number of observations. Therefore, Lasso regression can be used to select the most effective features from battery data and accurately predict SoH.

2.2.3. ElasticNet Regression

ElasticNet regression performs variable selection and regularization in regression models by using a combination of both L1 and L2 norm regularization. This method combines the variable selection advantages of Lasso with the performance in grouped variables of Ridge regression. ElasticNet provides an effective solution, especially in cases where features exhibit high correlation with each other. The model is optimized with a penalty term that includes both the sum of squares of coefficients (L2) and the sum of absolute values (L1), making the model resistant to overfitting and improving its generalization ability [20].

Unlike Linear and Lasso regression, ElasticNet regression has a more robust modeling structure against the multitude of variables and correlation problems among them, thanks to its regularization terms. LR does not include regularization, so the problem of multicollinearity among variables may negatively impact the model's accuracy. ElasticNet minimizes these issues with regulatory terms, providing more stable and reliable predictions.

2.2.4. Ridge Regression

Ridge Regression (RR), applied to the context of cycle, current, and capacity for batteries, involves using a regularization technique to enhance the stability and accuracy of the LR model [14]. RR is a variant of LR that includes a regularization term to the ordinary least squares (OLS) regression equation. In the case of cycle, current, and capacity analysis for batteries, RR aims to model the relationship between these variables while addressing potential issues such as multicollinearity and overfitting.

The RR Equation (2) with cycle current and capacity as variables is represented as

$$Capacity = \beta_0 + \beta_1 \times Cycle + \beta_2 \times Current + \epsilon + \lambda \sum_{i=1}^{p} {\beta_i}^2 \tag{2}$$

where

- *Capacity* is the number of battery's capacity;
- *Cycle* is the number of cycles;
- *Current* is the discharge current during cycles;
- β_0 is the intercept;
- β_1 and β_2 are the coefficients for cycle and current, respectively;
- ϵ is the error term;
- λ is the regularization parameter that controls the strength of regularization;
- $\sum_{i=1}^{p} {\beta_i}^2$ is the sum of squared coefficients, excluding the intercept.

In our use of the RR method, the regularization term $\lambda \sum_{i=1}^{p} {\beta_i}^2$ was added to the least squares equation. This term penalized large coefficients, preventing overfitting and reducing the impact of multicollinearity in the model. The parameter λ determines the extent of regularization, with higher values leading to greater shrinkage of coefficients. By incorporating RR into the analysis of cycle, current, and capacity data for batteries, we improved the model's robustness, handled correlated predictors more effectively, and produced more reliable predictions for battery performance.

2.3. Ensemble Modeling in Machine Learning

Ensemble methods in machine learning were designed to integrate predictions from multiple base models to achieve an overall performance improvement. This approach was used to enhance the model's predictive ability for unknown data and to better generalize learning on training data. The underlying motivation for using ensemble methods was to optimize the balance between bias and variance by combining a set of weak prediction models strongly, thus creating more accurate and reliable models. In this context, methods such as the Random Forest algorithm excel by combining multiple Decision Tree models to outperform a single Decision Tree [21].

Ensemble methods may be categorized into three main categories: Bagging, Boosting, and Stacking. Bagging trains many base models independently and creates final predictions by averaging their outputs. Boosting sequentially improves models and tries to correct errors of previous predictions at each step, reducing the model's bias. Stacking combines predictions from different models, using them as input for the final model to generate final predictions. Each method applies different strategies to enhance the model's robustness and performance on data, thus improving the model's reliability and accuracy [22].

3. Results

3.1. Exploratory Data Analysis (EDA) Studies on the Dataset

In this study, datasets from two different flights named Flight 92 and Flight 129 were thoroughly examined. The data from both flights were evaluated separately using exploratory data analysis (EDA) techniques. The analysis process was designed to understand the operational dynamics and potential anomalies specific to each flight.

The examined datasets contain 'CYCLE' numbers ranging from 1 to 285 for each flight, representing different time intervals during the flights. In total, the examined dataset for Flight 92 consists of 46,170 rows and 5 columns, and for Flight 129 it consists of 46,170 rows and 5 columns as well. Figure 1 illustrates the frequency distribution of current values.

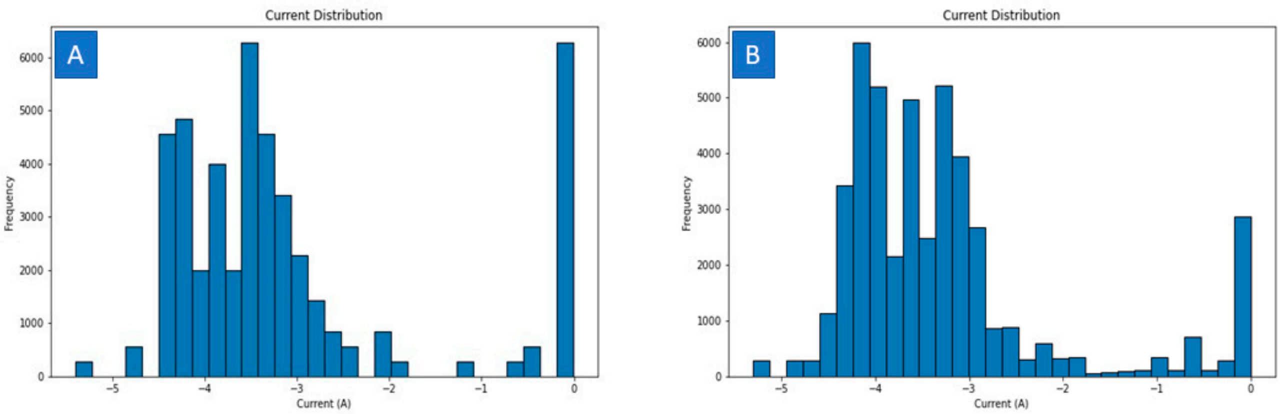

Figure 1. The frequency distribution of current values: (**A**) Flight 92, (**B**) Flight 129.

The current values provided in Figure 1 were classified into different intervals ranging from −5 amps to 0 amps. The shape of the histogram indicates that the majority of current values are concentrated within specific intervals. Particularly, the frequency between −3 and −4 amps shows the highest peak, with approximately 6000 observations within this range for Flight 92, while for Flight 192, the peak frequency occurs near −4 amps. This distribution concludes with a second peak near –1 amp, where approximately 6000 frequency values were observed. On the other hand, lower frequencies were observed between −1 and −2 amps intervals. The correlation matrix of the data is provided in Figure 2.

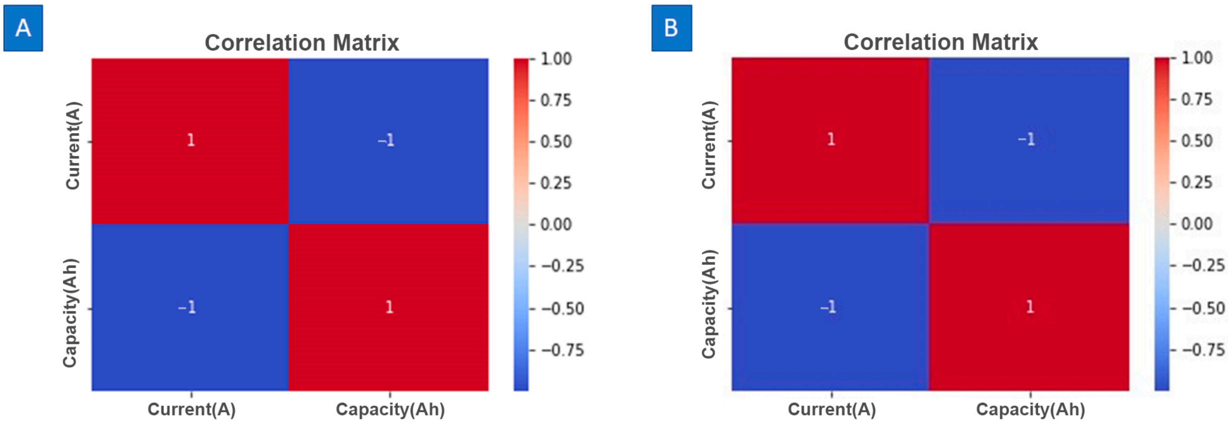

Figure 2. Correlation matrix of (**A**) Flight 92 and (**B**) Flight 129.

The correlation matrix provided in Figure 2 represents the correlation coefficients between current and capacity. These coefficients are shown as –1, indicating a perfect negative correlation between current and capacity. A perfect negative correlation indicates that as one variable increases, the other variable decreases in equal proportion (Equation (3)). This means that an increase in current corresponds to a decrease in capacity, or an increase in capacity corresponds to a decrease in current. This correlation analysis is especially helpful for further understanding the battery capacity degradation effect with respect to the current (A) feature values, accurate SoH modeling, improving predictive maintenance, and improving the accuracy for capacity prediction on models applied in this study.

$$Capacity(Ah) = \beta_0 + \beta_1 * Current(A) \tag{3}$$

According to the correlation matrix, there is a strong negative relationship between "Current (A)" and "Capacity (Ah)" (correlation coefficient of -1). In Equation (3), Capacity (Ah) represents the battery capacity, while Current (A) refers to the current. β_0 is the intercept term, indicating the capacity value when the current is zero. β_1, on the other hand, represents the slope of the relationship between current and capacity, and according to the correlation matrix, this slope is negative.

In this case, the correlation coefficient of -1 implies that the value of β_1 will also be negative. This means that as the current increases, the capacity decreases, resulting in a negative slope for the relationship. Therefore, a clear consistency can be observed between the correlation matrix presented in Figure 2 and Equation (3).

3.2. Dataset Splitting

When examining the data from both flights, it was observed that there were no missing values. The data were divided as shown in Table 1. The training set contains the data used in the learning process of the model. Each observation contains two features (columns). These features represent various dynamics and operational parameters of the flight. Training the model on these data allowed the algorithm to learn from the flight data and generate predictions that might be applied to future data. The "Holdout" set, on the other hand, was a separate data group allocated to evaluate the performance of the model, which was not used during the training process. This set was used to test how the model performed with data it had not seen during training.

Table 1. Dataset splitting values.

Flight 92		**Flight 129**	
Training Set	Holdout Set	Training Set	Holdout Set
30,934.2	15,237.2	30,934.2	15,237.2

3.3. Model Development

The process, starting with the preparation and preprocessing of the dataset, encompassed tasks such as standardization and normalization of the data required for training the models. These steps are of critical importance to enable the model to learn more effectively and make more accurate predictions. Subsequently, prediction models were developed using various regression models and Ensemble Learning techniques. These techniques included LR, Lasso, RR, and ElasticNet regression models. Each model was calibrated using specific error minimization techniques and regularization methods. The hyperparameters of the developed models are provided in Table 2.

Table 2. Model hyperparameters for LR, Lasso, ElasticNet and RR.

Model	**Hyperparameter**	**Value**
LR	fit_intercept normalize Positive	True False False
Lasso	Alpha fit_intercept random_state max_iter	0.0005 True 42 1000
ElasticNet	Alpha l1_ratio random_state	0.0005 0.9 42
RR	Alpha fit_intercept Normalize	2.0 True False

In Table 2, for each type of model, the important hyperparameters and their values are provided. For the LR model, the fit_intercept parameter was set to True, indicating that the model would compute the intercept term. Additionally, the Positive parameter was set to False, allowing coefficients to take negative values. For the Lasso Regression model, the Alpha value was set to 0.0005, indicating the amount of L1 regularization applied in the model. The fit_intercept was also set to True to compute the intercept term of the model. The random_state was set to 42 to determine the randomness of the model, ensuring its reproducibility. Moreover, the max_iter was set to 1000 to allow a sufficient number of iterations during the fitting process of the model. In the RR model, the Alpha value was set to 2.0, indicating the strength of the L2 regularization applied. The hyperparameter settings of each model significantly affected its performance and generalization ability. Therefore, the parameters were carefully selected and tuned.

Development of Ensemble Learning Models

Within the scope of the study, models were developed using regression models by employing Bagging, Boosting, and Stacking methods [21]. In the Bagging process, a Bagging regressor was built on top of a base estimator. The hyperparameters and regression models used in developing Ensemble models are provided in Tables 3 and 4.

Table 3. Model hyperparameters for bagging, boosting and stacking.

Model	Hyperparameter	Value	Predictions
Bagging	n_estimators max_samples bootstrap	10 1.0 True	LR, Lasso, ElasticNet, RR
Boosting	Table 4	Table 4	
Stacking	cv passthrough final_estimator	5 True linear_regression	ElasticNet, RR, Gradient Boosting Regressor

Table 4. Model hyperparameters for GBR, XGBoost and LightGBM.

Boosting Model	Hyperparameter	Value
Gradient Boosting Regressor (GBR)	n_estimators	3000
	learning_rate	0.05
	max_depth	4
	max_features	Sqrt
	min_samples_leaf	15
	min_samples_split	10
	loss	Huber
eXtreme Gradient Boosting Regressor (XGBoost)	colsample_bytree	0.4603
	gamma	0.0468
	learning_rate	0.05
	max_depth	3
	min_child_weight	1.7817
	n_estimators	2200
	reg_alpha	0.4640
	reg_lambda	0.8571
	subsample	0.5213
Light Gradient Boosting Machine Regressor (LightGBM)	num_leaves	5
	learning_rate	0.05
	n_estimators	720
	max_bin	55
	bagging_fraction	0.8
	bagging_freq	5
	feature_fraction	0.2319
	feature_fraction_seed	9
	bagging_seed	9
	min_data_in_leaf	6

In Table 3, the hyperparameters specified for three different Ensemble Learning techniques—Bagging, Boosting, and Stacking—are provided. Key parameters and their values for each technique, along with the summary of prediction models used with these techniques, are outlined. For the Bagging method, the fundamental parameters were specified as n_estimators, max_samples, and bootstrap. The max_samples value was set to 1.0, indicating that the maximum number of samples used for training each base estimator was all samples. The prediction models used in this method were listed as LR, Lasso, ElasticNet, and RR. The Stacking technique combined predictions from different prediction models to produce final predictions through a new meta-model. The cv parameter represented the number of folds used for cross-validation and was set to 5 in this case. The passthrough parameter was set to True, allowing the original features of base estimators to be directly passed through to the final predictor. LR was used as the final estimator, and predictions made with this method utilized ElasticNet, RR, and GBR models. The parameters and values specified for Boosting were provided in Table 5 in this table. Boosting was based on the principle of sequentially improving weak predictors to build a strong model.

Table 5. Comparison of the results of machine learning models.

Model	Flight 92		Flight 129	
	RMSE (%)	MAE (%)	RMSE (%)	MAE (%)
LR	0.04	0.03	0.97	0.20
Lasso	4.56	3.41	4.32	2.98
ElasticNet	4.11	3.07	3.91	2.69
RR	0.04	0.03	0.97	0.20

4. Experimental Results

This section thoroughly examines experimental analyses and the results obtained from evaluating the performance of different regression models. LR, Lasso, ElasticNet, and RR models were utilized within the scope of this study. Each model was trained and tested on appropriate datasets with predefined hyperparameters. RMSE and MAE metrics were employed to measure model accuracy. The predictive performances of the models were comparatively assessed.

4.1. Results of Machine Learning Techniques

During the analysis process, potential risks such as overfitting were considered alongside the predictive capabilities of the models on real data. Specifically, it was observed that the Lasso and ElasticNet regression models provided more stable results by reducing model complexity through regularization terms. Increasing the alpha parameter in RR significantly influenced the model's robustness and generalization ability, improving its adaptation to different datasets. The results of the models developed using machine learning techniques are provided in Table 5.

A performance evaluation conducted on four different regression models (LR, Lasso, ElasticNet, and RR) is presented in Table 5. The results were measured on two different flight datasets named Flight 92 and Flight 129. The LR and RR models exhibited significantly low RMSE and MAE values for both flight datasets. These results indicated high prediction accuracy and good fit of these two models to the data. The findings suggested that the performance of models might vary on datasets from flights with similar characteristics, emphasizing the importance of careful model selection based on flight attributes. This analysis provided important insights into assessing the suitability and generalizability of the model.

4.2. Results of Ensemble Modeling in Machine Learning

The findings obtained from applying ensemble modeling techniques in machine learning and the performance of these approaches were examined. Ensemble modeling aimed to improve overall prediction success by combining multiple prediction models. These techniques worked towards balancing individual model errors and reducing issues such as overfitting to provide more robust and reliable predictions. Within the scope of this study, ensemble methods such as Bagging, Boosting, and Stacking were used, and the effects of these methods on different datasets were analyzed comparatively. The results obtained are provided in Table 6.

Table 6. Result of ensemble modeling in machine learning.

Model	Flight 92		Flight 129	
	RMSE (%)	MAE (%)	RMSE (%)	MAE (%)
Bagging	2.20	1.64	2.22	1.46
GBR	0.14	0.03	0.67	0.10
XGBoost	24.17	19.87	24.42	18.36
LightGBM	1.21	0.23	1.73	0.65
Stacking	0.03	1.64	0.66	1.46

The Bagging method aimed to reduce the model's variance to achieve more stable results, while Boosting focused on correcting errors of individual predictors and thereby reduces the model's bias in this process. The Stacking method, on the other hand, combined predictions obtained from various models through a meta-regressor, ensuring the best possible balance of these predictions. Each method offered different advantages for specific data structures and types of problems.

The experimental results indicated that ensemble models generally provided predictions with higher accuracy compared to individual regression models. Particularly in noisy and complex datasets, the performance of ensemble methods was significantly superior (Figure 3). Moreover, models obtained through the application of these techniques have demonstrated better generalization ability on new and unseen data, enhancing the model's robustness. Figure 3 compares the performances of various machine learning models for the 92nd and 129th flights with their RMSE and MAE values. The GBR and Stacking models had the lowest error rates for both flights and provided the most accurate predictions. In particular, the GBR model gave the most successful results with RMSE 0.14 and MAE 0.03 for the 92nd flight, while it also exhibited very low errors for the 129th flight. On the other hand, the XGBoost model had the highest RMSE and MAE values for both flights and showed the lowest performance compared to the other models. The Bagging and LightGBM models also showed good performance with acceptable error rates.

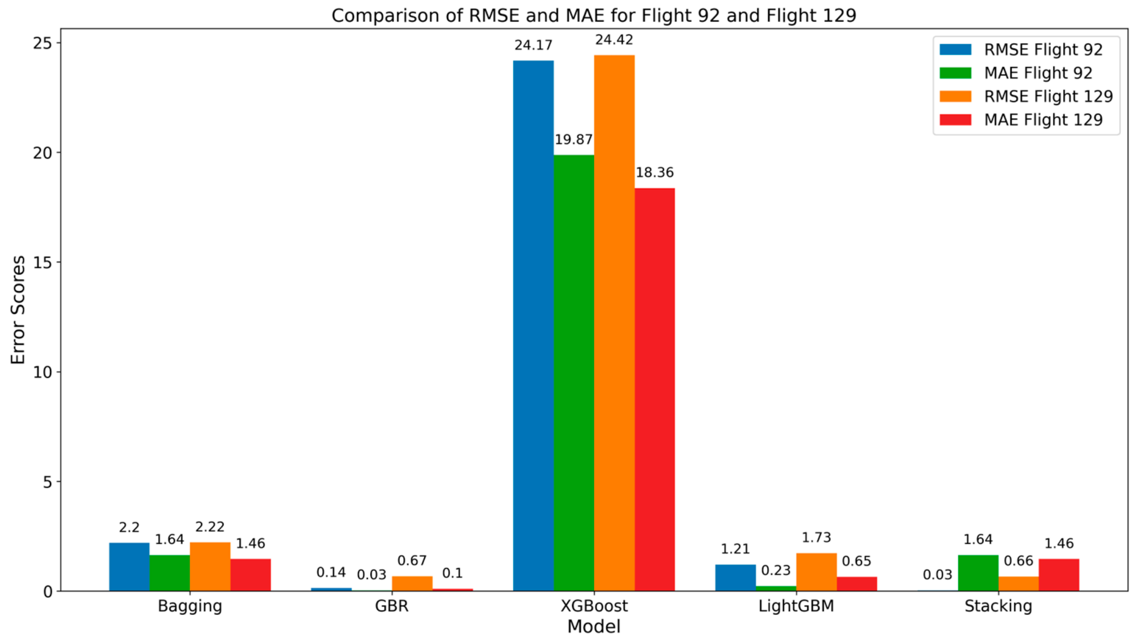

Figure 3. Comparison of RMSE and MAE for Flight 92 and Flight 129.

Consequently, ensemble modeling techniques offered an effective strategy in managing model complexity and maximizing prediction success in machine learning projects. These approaches optimized performance by striking a good balance between various prediction models and leveraging the strengths of each model. The results showed that the choice of hyperparameters and model selection played a crucial role in prediction accuracy. While each regression model offered different advantages for specific data structures and prediction needs, comprehensive testing of these models allowed for the selection of the most suitable model tailored to the application domain. This experimental study provided valuable insights into the development of advanced modeling techniques and the determination of the best model specific to the data.

5. Discussion

This section evaluates the comparative performance of various machine learning techniques applied to battery state estimation, as summarized in Table 7. This analysis highlighted the relative effectiveness of different approaches. Such comparative information with similar studies in the literature provided important insights to advance the field of battery management systems by identifying the most promising methodologies and areas for further research and improvement.

Table 7. Studies in the literature and model performances.

Reference	Method *	Chemistry	Dataset	MAE	RMSE
[23]	RLS	90 Ah $LiFePO_4$	OCV, SOC, and temperature	1.8%	2.3%
[24]	CNN + GRU (FUDS)	1.3 Ah 18650 NMC	Voltage, current, temperature	1.26%	1.54%
[25]	IB-ELM	2 Ah 18650 $LiFePO_4$	Charge, NASA dataset, aging test, voltage, current, temperature	0.010–0.034	0.014–0.039
[26]	LSTM	2.9 Ah 18650 NMC	Voltage, current, temperature	1.39%	1.7%
[27]	GRU	2.3 Ah 26650 $LiFePO_4$	Voltage, temperature	0.49%	0.64%
[28]	DNN	2.3 Ah 26650 $LiFePO_4$	Voltage, current, temperature	-	3.68%
[29]	SVR	2 Ah 18650 $LiFePO_4$	Oxford and NASA dataset	-	3.62% and 2.49%
Proposed Method	Ensemble Modeling	1.5 Ah 18650 NMC	Voltage, current, discharge capacity	0.03%	1.64%

* RLS: recursive least squares, CNN + GRU: convolutional neural network-gated recurrent unit, IB-ELM: improved blinex-extreme learning machine, LSTM: long short-term memory, DNN: deep neural network, SVR: support vector regression.

In Table 7, various machine learning models, ranging from simple regressions to complex deep learning architectures, were evaluated for their accuracy in predicting the SoH status of LIBs. The RLS method applied to the 90 Ah $LiFePO_4$ battery showed reasonable accuracy with MAE and RMSE values of 1.8% and 2.3%, respectively. In particular, the combined CNN+GRU model trained under the Federal Urban Driving Schedule (FUDS) provided improved performance with lower error measures (1.26% MAE and 1.54% RMSE). On the other hand, the performance of the DNN and Support Vector Regression (SVR) models, although generally robust, exhibited higher RMSE values under certain conditions, highlighting potential limitations in their adaptability or tuning.

Our proposed Ensemble Modeling approach, which combines multiple predictive models to leverage the strengths of individual models, achieved an MAE of only 0.03 and an RMSE of 1.64. This approach not only improved the prediction accuracy but also provided robustness to various operating conditions and dataset variabilities, validating the potential of ensemble methods in complex system predictions such as battery states.

The ensemble modeling techniques examined in this study, Bagging, Boosting, and Stacking, were applied to different regression problems, and the performance of each method was comprehensively evaluated. The results shed light on how these techniques might impact the overall success of the model and when each technique should be preferred. It was observed that the Bagging method is particularly effective in datasets with high variance problems. This method increased sample diversity and ensured independence among base predictors, thus minimizing the risk of overfitting. However, Bagging may sometimes be insufficient in reducing bias, which may lead the model to make biased errors in certain cases.

Boosting was used to improve weak predictors and correct sequential errors. While this method balanced low bias and high variance, it might increase complexity in some cases, leading to overfitting. Boosting was suitable for situations designed to manage low bias and variance that needed improvement. Stacking optimized performance by combining the strengths of various prediction models and integrating these predictions through a meta-regressor. This approach maintained generalization ability while increasing model complexity and synergies obtained by combining various prediction models enhanced overall prediction success. The application of Stacking might be highly effective when correct parameters and predictors were chosen, but incorrect implementation might complicate the model's interpretability and practical applicability.

6. Conclusions

The in-depth analysis of ensemble modeling techniques discussed in this study demonstrated how these approaches might be effectively used in various machine learning problems. While Bagging, Boosting, and Stacking methods offered advantages in certain data structures and learning scenarios, the challenges and limitations encountered during their application were also considered. Stacking stands out, especially with a 0.03 and 0.66 RMSE performance.

The Bagging methodology was effective in improving overall performance by reducing the model's variance, particularly in variable datasets where it prevented overfitting, but falls short in low bias problems. Boosting achieved noticeable successes, especially when working with weak learners, by enhancing model accuracy through sequential improvements, but carries risks of overfitting in complex data structures. Stacking created a robust meta-model by integrating predictions from different models and achieved high success rates through synergy from the combination of various prediction models.

In conclusion, the findings of this study shed light on the evolution of ensemble modeling techniques and their future applications. It was expected that with the integration of artificial intelligence algorithms, processes such as automatic adjustment of model parameters and dynamic model selection would be automated. The development of stronger hybrid models through the combination of different ensemble techniques offers the potential to achieve better results in complex data structures. The integration of ensemble modeling approaches with deep learning frameworks requires more attention to their environmental and ethical dimensions, especially regarding data privacy and algorithmic bias issues. In conclusion, ensemble modeling techniques will continue to be a significant part of ongoing developments and innovations in the field of machine learning. These techniques play a critical role in optimizing model performance and producing generalizable solutions on larger datasets, both in academic research and industrial applications.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/batteries10100371/s1. Table S1: Algorithm created for Flight No. 92 in battery test system; Table S2: Algorithm created for Flight No. 129 in battery test system.

Author Contributions: Conceptualization, E.B. and S.O.; methodology, S.O. and E.B.; software, S.O.; validation, B.Ç., S.O., and A.A.; formal analysis, B.Ç.; investigation, B.Ç. and S.O.; resources, A.A.; data curation, S.O.; writing—original draft preparation, E.B. and S.O.; writing—review and editing, E.B. and S.O.; visualization, S.O.; supervision, E.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The research was conducted collaboratively by the Battery Research Group at Sivas University of Science and Technology and MOBILERS team at Sivas Cumhuriyet University.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chen, C.; Zheng, Z.; Xu, T.; Guo, S.; Feng, S.; Yao, W.; Lan, Y. YOLO-based UAV technology: A review of the research and its applications. *Drones* **2023**, *7*, 190. [CrossRef]
2. Hayat, S.; Yanmaz, E.; Muzaffar, R. Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 2624–2661. [CrossRef]
3. Saravanakumar, Y.N.; Sultan, M.T.H.; Shahar, F.S.; Giernacki, W.; Łukaszewicz, A.; Nowakowski, M.; Holovatyy, A.; Stępień, S. Power sources for unmanned aerial vehicles: A state-of-the art. *Appl. Sci.* **2023**, *13*, 11932. [CrossRef]
4. Hong, D.; Lee, S.; Cho, Y.H.; Baek, D.; Kim, J.; Chang, N. Least-energy path planning with building accurate power consumption model of rotary unmanned aerial vehicle. *IEEE Trans. Veh. Technol.* **2020**, *69*, 14803–14817. [CrossRef]
5. Avanzini, G.; de Angelis, E.L.; Giulietti, F. Optimal performance and sizing of a battery-powered aircraft. *Aerosp. Sci. Technol.* **2016**, *59*, 132–144. [CrossRef]
6. Berecibar, M.; Gandiaga, I.; Villarreal, I.; Omar, N.; Van Mierlo, J.; Van Den Bossche, P. Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renew. Sustain. Energy Rev.* **2016**, *56*, 572–587. [CrossRef]
7. Oyucu, S.; Dümen, S.; Duru, İ.; Aksöz, A.; Biçer, E. Discharge capacity estimation for Li-ion batteries: A comparative study. *Symmetry* **2024**, *16*, 436. [CrossRef]
8. Andrioaia, D.A.; Gaitan, V.A.; Culea, G.; Banu, I.V. Predicting the RUL of Li-ion batteries in UAVs using machine learning techniques. *Computers* **2024**, *13*, 64. [CrossRef]
9. Zhang, H.; Zhou, M.; Lan, X. State of charge estimation algorithm for unmanned aerial vehicle power-type lithium battery packs based on the extended kalman filter. *Energies* **2024**, *12*, 3960. [CrossRef]
10. Shibl, M.M.; Ismail, L.S.; Massoud, A.M. A machine learning-based battery management system for state-of-charge prediction and state-of-health estimation for unmanned aerial vehicles. *J. Energy Storage* **2023**, *66*, 107380. [CrossRef]
11. Chopra, I. Small UAS and delivery drones: Challenges and opportunities the 38th Alexander A. Nikolsky honorary lecture. *J. Am. Helicopter Soc.* **2021**, *66*, 1–41. [CrossRef]
12. Hassan, T.J.; Jangula, J.; Ramchadra, A.R.; Sugunaraj, N.; Chandar, B.S.; Rajagopalan, P.; Rahman, F.; Ranganathan, P.; Adams, R. UAS-guided analysis of electric and magnetic field distribution in high-voltage transmission lines (Tx) and multi-stage hybrid machine learning models for battery drain estimation. *IEEE Access* **2024**, *12*, 4911–4939. [CrossRef]
13. Oyucu, S.; Doğan, F.; Aksöz, A.; Biçer, E. Comparative analysis of commonly used machine learning approaches for Li-ion battery performance prediction and management in electric vehicles. *Appl. Sci.* **2024**, *14*, 2306. [CrossRef]
14. Oyucu, S.O.; Ersöz, B.; Sagiroglu, S.; Aksöz, A.; Biçer, E. Optimizing Lithium-ion battery performance: Integrating machine learning and explainable AI for enhanced energy management. *Sustainability* **2024**, *16*, 4755. [CrossRef]
15. Tina, G.M.; Ventura, C.; Ferlito, S.; De Vito, S. A state-of-art-review on machine-learning based methods for PV. *Appl. Sci.* **2021**, *11*, 7550. [CrossRef]
16. Rincy, T.N.; Gupta, R. Ensemble learning techniques and its efficiency in machine learning: A survey. In Proceedings of the 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 28–29 February 2020. [CrossRef]
17. Rodrigues, T.A.; Patrikar, J.; Choudhry, A.; Feldgoise, J.; Arcot, V.; Gahlaut, A.; Lau, S.; Moon, B.; Wagner, B.; Scott Mathews, H.; et al. In-flight positional and energy use data set of a DJI matrice 100 quadcopter for small package delivery. *Sci. Data* **2021**, *8*, 6–13. [CrossRef] [PubMed]
18. Maulud, D.; Abdulazeez, A.M. A Review on linear regression comprehensive in machine learning. *J. Appl. Sci. Technol. Trends* **2020**, *1*, 140–147. [CrossRef]
19. Ranstam, J.; Cook, J.A. LASSO regression. *Br. J. Surg.* **2018**, *105*, 1348. [CrossRef]
20. Chamlal, H.; Benzmane, A.; Ouaderhman, T. Elastic net-based high dimensional data selection for regression. *Expert Syst. Appl.* **2024**, *244*, 122958. [CrossRef]
21. Banfield, R.E.; Hall, L.O.; Bowyer, K.W.; Kegelmeyer, W.P. A comparison of decision tree ensemble creation techniques. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 173–180. [CrossRef]
22. Satoła, A.; Satoła, K. Performance comparison of machine learning models used for predicting subclinical mastitis in dairy cows: Bagging, boosting, stacking and super-learner ensembles versus single machine learning models. *J. Dairy Sci.* **2024**, *107*, 3959–3972. [CrossRef] [PubMed]
23. Duong, V.H.; Bastawrous, H.A.; See, K.W. Accurate approach to the temperature effect on state of charge estimation in the $LiFePO_4$ battery under dynamic load operation. *Appl. Energy* **2017**, *204*, 560–571. [CrossRef]
24. Huang, Z.; Yang, F.; Xu, F.; Song, X.; Tsui, K.L. Convolutional gated recurrent unit-recurrent neural network for state-of-charge estimation of lithium-ion batteries. *IEEE Access* **2019**, *7*, 93139–93149. [CrossRef]
25. Ma, W.; Cai, P.; Sun, F.; Wang, X.; Gong, J. State of health estimation for Lithium-ion batteries based on extreme learning machine with improved blinex loss. *Int. J. Electrochem. Sci.* **2022**, *17*, 221170. [CrossRef]
26. Ma, L.; Hu, C.; Cheng, F. State of charge and state of energy estimation for lithium-ion batteries based on a long short-term memory neural network. *J. Energy Storage* **2021**, *37*, 102440. [CrossRef]
27. Xiao, B.; Liu, Y.; Xiao, B. Accurate state-of-charge estimation approach for lithium-ion batteries by gated recurrent unit with ensemble optimizer. *IEEE Access* **2019**, *7*, 54192–54202. [CrossRef]

28. How, D.N.T.; Hannan, M.A.; Lipu, M.S.H.; Sahari, K.S.M.; Ker, P.J.; Muttaqi, K.M. State-of-charge estimation of Li-ion battery in electric vehicles: A deep neural network approach. *IEEE Trans. Ind. Appl.* **2020**, *56*, 5565–5574. [CrossRef]
29. Tian, J.; Xiong, R.; Shen, W. State-of-health estimation based on differential temperature for lithium ion batteries. *IEEE Trans. Power Electron.* **2020**, *35*, 10363–10373. [CrossRef]