Battery Life Prediction Using Machine learning Model And Monte Carlo

Rakshit Raj Rajuladevi
Telecommunication Systems
Blekinge Institute of Technology
Karlskrona, Sweden
rara22@student.bth.se

Venkata Bhamidipati
Telecommunication Systems
Blekinge Institute of Technology
Karlskrona, Sweden
vebh22@student.bth.se

***Abstract*—In this work, a predictive model is developed to estimate the battery life of mobile devices based on user input and incorporate uncertainty analysis using Monte Carlo simulation. The Random Forest algorithm was selected as the predictive model due to its ability to handle complex, non-linear relationships in the data and through various tests among the models that suits better for our experiment . The model was trained on a dataset consisting of selected features such as screen time, charging cycles, and Bluetooth usage, which influence battery lifespan and its performance.**

To account for the inherent uncertainty in the input data and model predictions, utilizing Monte Carlo simulation. This allows generating a distribution of possible battery lifespan outcomes by incorporating the variability observed in input from the training data. To evaluate the variability in the model predictions and simulate several possible scenarios by modifying the input characteristics according to their respective distributions.

For a new user input, a single-point prediction was made using the trained Random Forest model. However, rather than relying only on this point estimate, we generated a range of possible lifespans by introducing variability based on the standard deviation observed from the Monte Carlo simulation. The output included the mean predicted lifespan, a 95% confidence interval, and a distribution of possible outcomes, providing a robust assessment of the likely range of battery performance for the user.

This approach offers a more comprehensive prediction that accounts for uncertainty and variability, allowing users to make informed decisions based on a range of possible outcomes rather than a single deterministic prediction. The combination of Random Forest and Monte Carlo simulation proved effective for quantifying uncertainty and enhancing the reliability of the prediction.

***Index Terms*—machine learning, mobile lifespan, Battery performance, uncertainty prediction ,Simulation.**

I. Introduction

In recent years, the use of mobile phones has increased dramatically, becoming an integral part of daily life for billions of people worldwide. Smartphones, equipped with a plethora of features such as high-resolution displays, GPS, and internet connectivity, are used for a variety of tasks, from communication and social media to professional work and entertainment. With this increased usage, there is a growing concern among users about the battery life of their mobile phones. As smartphones become more powerful and are used more intensively, users are increasingly aware of how their usage patterns—such as screen time, app usage, and connectivity—affect battery performance. [6] This concern is further amplified by the fact that battery degradation over time is inevitable, leading to a reduced lifespan of the device. [4] As a result, predicting battery life has become a critical need for both consumers and manufacturers to ensure optimal usage and longevity of mobile devices. [2]

Given the widespread concern over battery life, there is a clear need for predictive models that can accurately estimate how long a mobile phone's battery will last under varying conditions [3]. Predictive models use data about the phone's usage patterns and environmental factors to forecast battery life. These models can help users manage their usage more effectively and allow manufacturers to design more energy-efficient devices [4] [1]. Among the various types of predictive models, machine learning models have shown great promise. In particular, three types of machine learning models are often used for predicting battery life: Random Forest, Linear Regression, and Support Vector Machine (SVM) models.

Random Forest: This is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of their predictions for classification tasks or the mean prediction for regression tasks. It is known for its robustness to overfitting and ability to handle complex, non-linear relationships between features.

Linear Regression: A simple yet effective model that predicts the outcome based on the linear relationship between the dependent and independent variables. It is easy to implement and interpret, making it a popular choice for initial predictions.

Support Vector Machine (SVM): This model finds the hyperplane that best separates data into classes (for classification tasks) or fits a hyperplane that can predict a continuous variable. It is effective in high-dimensional spaces and when the number of dimensions exceeds the number of samples.

These models serve as powerful tools for predicting battery life, but they also come with their own sets of challenges [2] [5].

While machine learning models like Random Forest, Linear Regression, and SVM provide valuable predictions based on historical data, they primarily offer single-point estimates of battery life. This means that given a specific set of inputs, the model predicts a single output—such as the average or expected battery life. However, real-world scenarios are more complex and involve a great deal of variability and uncertainty. For example, different users may have different usage patterns that are not perfectly captured by the model, or external factors like temperature changes can affect battery performance [3] [5]. These single-point predictions do not account for the inherent randomness and fluctuations in everyday usage, leading to potential inaccuracies in battery life estimation.

To address the limitations of single-point predictions, Monte Carlo simulations can be employed. Monte Carlo simulation is a statistical technique that allows for the modeling of uncertainty in predictions by generating a distribution of possible outcomes based on random sampling of input variables. By incorporating variability in key factors such as screen time, charging cycles, and ambient temperature, Monte Carlo simulations provide a more comprehensive picture of the potential range of battery life outcomes. This approach not only gives a more realistic estimation of battery life but also helps in understanding the uncertainty and variability associated with the predictions. By using Monte Carlo simulations in conjunction with machine learning models like Random Forest, we can move beyond simple point estimates to provide users with a range of possible battery life scenarios, enhancing the utility and reliability of these predictions.

II. System Modelling

A. Overview And Architecture of System

1) Input Data and Dataset Collection: The data for this analysis is based on a combination of user behavior patterns and device specifications, with a focus on predicting battery life. The user data was generated by referencing research papers that explore smartphone usage patterns, battery degradation, and device performance across different environments. Additionally, data specific to various iPhone models, including iPhone 12, iPhone 13, iPhone 14, and iPhone 15, was collected and incorporated into the dataset.

For each iPhone model, key parameters such as screen time, charging cycles per week, Bluetooth usage, cellular data hours, GPS activity, ambient temperature, and battery capacity were considered. These factors play a crucial role in determining the lifespan of a smartphone battery. The battery capacity of each iPhone model was thoroughly researched to provide a realistic basis for our analysis. The dataset was created with this information to train and test machine learning models aimed at predicting battery life under varying user conditions.

2) Machine Learning Model for Prediction: To predict battery life, several machine learning models were explored, including Random Forest, Linear Regression, and Support Vector Machine (SVM). Each model was trained and tested on the collected dataset to evaluate its performance. The dataset was split into 70

The evaluation metrics used were Mean Squared Error (MSE) and R^2 Score:

Mean Squared Error (MSE): This metric measures the average of the squared differences between predicted and actual values. A lower MSE indicates a model with better predictive accuracy. R^2 Score: This metric indicates how well the model explains the variance in the data. A higher R^2 score suggests better model performance. Among the tested models, Random Forest demonstrated the lowest MSE and the highest R^2 score, making it the optimal choice for predicting battery life. Random Forest's ability to handle non-linear relationships and its robustness against overfitting made it suitable for this task.

To further improve the model's precision, two different dataset sizes were tested: 1000 and 2000 data points. Initially, training and testing the model on 1000 data points resulted in higher MSE, indicating lower precision. However, when the dataset was expanded to 2000 data points, the MSE decreased, showing a significant improvement in precision. Therefore, the final model for predicting battery life across iPhone models was trained and tested using 2000 data points for higher accuracy.

3) Monte Carlo Simulation: To overcome the limitations of single-point predictions from the Random Forest model, Monte Carlo Simulation was integrated. Monte Carlo simulations are used to model uncertainty by generating a distribution of possible outcomes based on random sampling. This approach helps capture the variability in battery life predictions that may arise due to differences in user behavior and environmental factors.

In this project, the Monte Carlo simulation was implemented by generating random data for each feature based on its mean and standard deviation. Specifically, a normal distribution (randn) was used to simulate variability in parameters such as screen time, charging cycles, and ambient temperature. By running these simulations and integrating the results with the Random Forest model, a range of battery life predictions was generated rather than a single value.

The simulation results were analyzed by calculating the 95% Confidence Interval (CI), which provides a range in which the true battery life is expected to fall with 95% certainty. The 95% CI is important because it accounts for the variability in the data and gives users a probabilistic range for battery life rather than an exact number. This interval allows for more informed decision-making by considering the inherent uncertainties in predicting battery lifespan.

4) Battery Life for New Users: For new users, the model is designed to provide battery life predictions based on the user's specific parameters, such as screen time, Bluetooth usage, and charging cycles. The user inputs these parameters, and the Random Forest model, trained on extensive datasets, generates a predicted battery life.

However, to provide a more comprehensive estimate, the Monte Carlo simulation is also applied. This allows the user to see not only the expected battery life but also the variability in the prediction. The battery life is predicted with an associated range, giving users an understanding of how their phone's battery may perform under various conditions. This range helps convey the uncertainty in the prediction, acknowledging that actual battery life may fluctuate within this interval depending on usage patterns and environmental factors.

B. Reasons For Considering This Design

Based on the some parameters i.e mean square error, R-square test the prediction model is fixed for the efficiency and accuracy of the prediction. Introducing Monte Carlo helps us to visualize the distribution of range of battery lifespan that could take place.This could help in predict the battery lifespan of the mobiles in various scenarios.

III. Results And Analysis

A. Comparison of Machine Learning Models

As a part of selecting the Machine Learning Model to decide which model is best suited for our experiment the parameters such as performance and efficiency are taken into consideration between the various machine learning models, including Random Forest, Linear Regression, and Support Vector Machine (SVM).The key evaluation metrics used for this comparison were Mean Squared Error (MSE) and the R^2 Score.Upon evaluation, the Random Forest model consistently demonstrated the lowest MSE and the highest R^2 score among the tested models. This clearly indicates its superior performance in predicting battery life accurately, making it the most suitable model for the task.

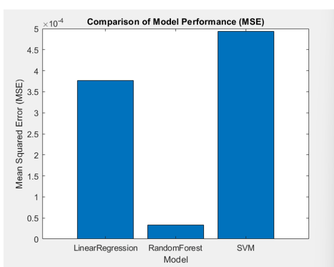

Fig. 1. Comparing Different ML models

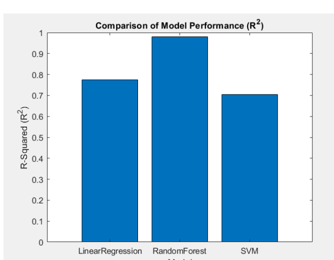

Fig. 2. Comparing R^2 Different ML models

Model	MSE	R^2
LinearRegression	0.00037714	0.77407
RandomForest	3.3551e-05	0.9799
SVM	0.00049327	0.70451

TABLE I
Performance comparison of different models

The Values that are illustrated in the table supports that comparative performance analysis of each model and justify the selection of Random Forest as the primary model. Therefore the Random Forest method is selected for the other data sets to train the Machine and integrating the Montecarlo simulation.

B. Impact of Dataset Size on Model Precision

The precision of the battery life prediction was also examined in relation to the size of the dataset. Initially, the model was trained and tested on a dataset containing 1000 data points.later the same training and testing on datasets containing 2000 data points.The comparision of both case values are also demonstrated in the table . The results showed a significant reduction in the MSE, demonstrating that the model's precision improved as the dataset size increased. This indicates that increasing the volume of training data helps the model capture a more comprehensive view of user behaviors and environmental factors, leading to better predictive accuracy. The above comparision was performed for dataset iphone12 considering datasizes 1000 and 2000. For the other iphone models 13,14,15 the simulation was conducted using the data sizes 2000.

Data Size	MSE
1000	4.864e-05
2000	3.2104e-05

TABLE II
Performance comparison of different datasizes considered

C. Dependence of Confidence Interval

The integration of Monte Carlo simulations with the Random Forest model allows for capturing the variability in battery life predictions by generating a 95% Confidence Interval (CI).If the Random Forest prediction lies within this 95% CI, it indicates that the model's prediction is consistent with the range of uncertainty generated by the Monte Carlo simulation.If not in any case their is some bias in the output generated. Generated CI is also demonstrated below to visualise that they are robust and inherent variability in the data for better precision .

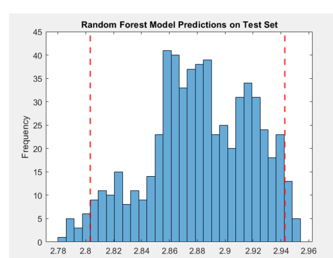

Fig. 3. 95% CI Random Forest Method

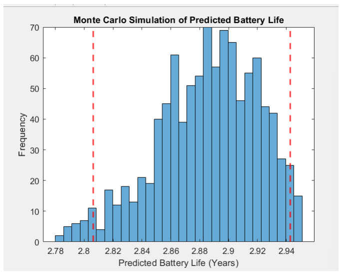

Fig. 4. 95% CI Montecarlo Simulation

Method	95% CI
Random Forest	2.8032 - 2.9427
Montecarlo Simulation	2.8061 - 2.9426

TABLE III
COMPARISON OF 95% CI

D. Comparative Analysis of Different iPhone Models

Another key analysis involved in the testing is providing same user input parameters for all iphone models considering Screen Time Hours: 5 Charging Cycles Per Week: 7 Bluetooth Hours: 2 Cellular Data Hours: 3 GPS Hours: 1 Ambient Temperature Celsius: 25 By varying Battery Capacity 2815mAh, 3240mAh, 3279mAh, 3349mAh across different iPhone models—iPhone 12, iPhone 13, iPhone 14, and iPhone 15 respectively. By running the Random Forest model with identical user conditions we can only achieve sigle point predictions. However, to provide a more comprehensive estimate, the Monte Carlo simulation is also applied. This allows the user to see not only the expected battery life but also the variability in the prediction. We could directly compare the battery performance of each iPhone model. This analysis aimed to determine which model offered the best battery life under similar usage conditions. The comparison of this iPhone models are illustrated in the table to showcase the users to decide which has better lifespan in this scenario.

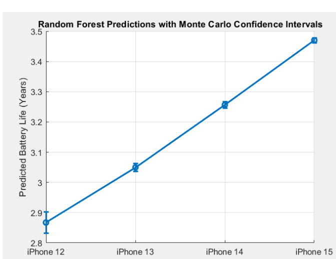

Fig. 5. Prediction of Lifespan in Years using Random Forest and Monte Carlo

iphone Model	Random Forest	Monte Carlo
12	2.8672 years	2.8318 - 2.9027 years
13	3.0495 years	3.0368 - 3.0622 years
14	3.2565 years	3.2453 - 3.2677 years
15	3.4702 years	3.4621 - 3.4784 years

TABLE IV
PREDICTION OF LIFE SPAN IN YEARS USING RANDOM FOREST AND MONTE CARLO

The results show case that iPhone 15 has the best suited for longevity.

IV. CONCLUSION

This Study successfully developed a robust model for predicting smartphone battery life by combining Random Forest machine learning with Monte Carlo simulations to address the inherent uncertainty in user behavior and environmental factors. By leveraging key parameters such as screen time, charging cycles, and ambient temperature, the model was able to predict battery life across different iPhone models with accuracy.

From results section the conclusion are driven in such a way that The Random Forest model was chosen due to its superior performance in handling non-linear relationships, as evidenced by its low Mean Squared Error and high R^2 score compared to other models like Linear Regression and SVM. As the dataset size increased, the model's precision improved, demonstrating the importance of comprehensive data. Furthermore, the integration of Monte Carlo simulations provided a probabilistic range of outcomes through the 95% Confidence Interval, ensuring that predictions were both accurate and reflective of real-world variability.

Ultimately, this approach aligns with the goals set out in the abstract by providing reliable, data-driven battery life predictions that account for uncertainty, making it a valuable tool for both users and manufacturers to optimize device usage and design.

V. FUTURE WORK

Based on Research done the conclusions are made but there are still various areas in which future researchers can work through to make this more promising . Some of the possible works that are suggested are as follows :

- **Power Consumption By Specific Apps**
  Currently this report only focuses on general factors that effect the battery life ,the integration of Power consumption due to specific apps should be included into the model as Different apps have varied power requirements and this can be a significant factor in determining a device's lifespan.
- **Collection of Real-time User Data**
  This process uses static datasets to predict the mobile life if the real-time data from users could improve the accuracy and usage- pattern of the users which helps in analysis of users up to data .
- **Hybrid Machine Learning Models**
  Single prediction model is used in this research , but hybrid modeling may increase accuracy .
- **Incorporating Dynamic Changes In Uncertainty**
  Static uncertainty is introduced in the system through the Monte Carlo but by incorporating the dynamic uncertainty the prediction can be made to the real-time user data .
- **Personalized Reports To Optimize Battery Life**
  Making the results obtained from the data as report helps

the user to change their daily habits to optimize their battery life.

REFERENCES

[1] Dutta Pramanik, Pijush Sinhababu, Nilanjan Mukherjee, Bulbul Padmanaban, Sanjeevikumar Maity, Aranyak Upadhyaya, Bijoy Holm-Nielsen, Jens Choudhury, Prasenjit. (2019). Power Consumption Analysis, Measurement, Management, and Issues: A State-of-the-Art Review of Smartphone Battery and Energy Usage. IEEE Access. 7. 182113-182172. doi:10.1109/ACCESS.2019.2958684.

[2] Chandra Sekhar, J N Domathoti, Bullarao Santibanez Gonzalez, Ernesto. (2023). Prediction of Battery Remaining Useful Life Using Machine Learning Algorithms. Sustainability. 15. 15283. doi:10.3390/su152115283.

[3] Antônio Sá Barreto Neto, Felipe Farias, Marco Aurélio Tomaz Mialaret, Bruno Cartaxo, Príscila Alves Lima, Paulo Maciel, Building energy consumption models based on smartphone user's usage patterns, Knowledge-Based Systems,Volume 213,2021,106680,ISSN 0950-7051, doi:10.1016/j.knosys.2020.106680.

[4] Carroll, Aaron Heiser, Gernot. (2010). An Analysis of Power Consumption in a Smartphone. Proc. 2010 USENIX conf., USENIX Assoc..

[5] Breda, Joseph Trivedi, Amee Wijesundara, Chulabhaya Bovornkeeratiroj, Phuthipong Irwin, David Shenoy, Prashant Taneja, Jay. (2019). Hot or Not: Leveraging Mobile Devices for Ubiquitous Temperature Sensing. BuildSys '19: Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. 41-50. doi:10.1145/3360322.3360856.

[6] https://www.honor.com/uk/blog/how-long-does-a-phone-battery-last

[7] https://medium.com/batterybits/comparison-of-open-datasets-for-lithium-ion-battery-testing-fd0de091ca2