Estimating Battery Charge Time: A Data Science And Machine Learning Proof Of Concept

Richard Lo

June 3, 2025

Abstract

Made for iPhones. And, yes, I know Battery Intelligence should be coming in iOS 19.

Problem Statement

We've all been there: you plug in your iPhone, eager to head out, only to realise you have no idea how long it will take to reach a usable charge. Will it be 15 minutes or an hour? This uncertainty is more than just a minor annoyance; it directly impacts our ability to plan our day, whether it's knowing if you can squeeze in a quick errand before a call, or deciding if you need to carry a power bank for an upcoming meeting. The current "charging percentage" on our devices, while informative about the current state, completely fails to provide a reliable estimate of remaining charge time.

Imagine a world where, upon plugging in your device, you're immediately presented with a precise message: "Charged in 27 minutes". This project aims to bring that vision to life. By using data science and machine learning, let's seek to develop a predictive model that provides an accurate battery charge time estimation, eliminating user frustration and enabling better planning. The end user will simply see a clear, estimate directly on their device's charging screen.

Project Plan

Our approach to tackling this problem begins with constructing a mathematical framework to model the remaining charging time. The core idea is to predict the time it takes to reach a desired charge level by understanding the rate of charge. We will use 80% as our target charge to promote good battery health. Lithium-ion batteries become strained when charged to 100% frequently.

We might want to just take measurements initially. Making a note that it might take roughly 45 minutes to charge from 20% to 80%. But what about all the other values? 37%, 42%, 69%, or 73%? That would be a real pain to measure each and every case 0% to 80%, then 1% to 80%, then 2% to 80%, then 3% to 80%, and so on up to 79% to 80%. We might spot that we can break up the problem to reduce overlap. Let's try that again taking measurements like this: 0% to 1%, then 1% to 2%, then 2% to 3%, then 3% to 4%, and so on up to 79% to 80%. This is better, but still leaves room for improvement. I certainly do not want to let my battery drain to 0% as this also hurts the health of the battery, much like charging to 100%. This also means we have missing data points we would have to predict. My conjecture is that it is easier to work with charging speed, rather than charging time, making the assumption that charging speed is constant from 0% to 50% where it may taper off as the charge reaches 80%. This will make data collection easier. And we will unify the quantities through the famous formula

$$\text{speed} = \frac{\text{distance}}{\text{time}}.$$

Mathematically, the total time $\tilde{T}(X)$ to charge from a current percentage X to a target percentage can be represented as a sum of infinitesimal time steps Δt_n for each infinitesimal change in charge Δx_n and empirical rate $s_n = \Delta x_n / \Delta t_n$

$$\tilde{T}(X) = \sum_{n=X}^{79} \Delta t_n = \sum_{n=X}^{79} \frac{\Delta x_n}{s_n} \sim \int_X^{80} \frac{1}{S(u)} du$$

where

$$S(X) = \frac{dX}{dt} \qquad T(X) = \int_X^{80} \frac{1}{S(u)} du \qquad X \in [0, 80)$$

will be our machine learning model to predict charging speed (percentage points per minute). Our primary goal is to accurately model the charging speed using empirical data.

Data Collection

I am the stakeholder and I want this project done fast. Data collection will be quick and dirty to get this done as fast as poss – Wait! I meant to say...

To build a robust model for $S(X)$, comprehensive data collection is paramount. For this project, the battery level was recorded after every minute. The charging speed can be approximated with

$$S_{\text{emp}}(X_n) \sim \frac{X_{n+1} - X_n}{(n+1) - n}$$

which contributes to data pairs

$$L = \{(X_n, S_{\text{emp}}(X_n))\}_n.$$

This dataset captures the instantaneous rate at which the battery charges at different levels, which is the direct input required to tune our $S(X)$ model. Here is the real data

$$\begin{aligned}\Big[&(20,1),(21,1),(22,3),(25,2),(27,2),(29,2),(31,2),\\ &(33,2),(35,1),(36,2),(38,2),(40,2),(42,1),(43,2),\\ &(45,1),(46,2),(48,2),(50,1),(51,2),(53,2),(55,1),\\ &(56,2),(58,1),(59,1),(60,2),(62,1),(63,1),(64,1),\\ &(65,1),(66,1),(67,1),(68,2),(70,1),(71,1),(72,1),\\ &(73,1),(74,.5),(75,1),(76,.5),(77,1),(78,.5),(79,1)\Big]\end{aligned}$$

where each point represents $(X, S_{\text{emp}}(X))$. The charging was done using a wired fast charger.

Control Group

To demonstrate the value of our predictive model, it is essential to compare it against existing approaches and note any weaknesses. Existing models are often a rudimentary estimate based on simple linear extrapolation of a predefined average, leading to significant inaccuracies.

Let's assume a constant charging rate $S(X) = a$ and see what we get. Fitting to the data yields

$$T(X) = \frac{80 - X}{a}, \qquad a = 1.39286$$

is the linear estimate model. The issue is that iPhones have lithium-ion batteries and will often charge quickly from 0-50% but then slows down drastically. So a linear model will wildly underestimate the remaining time. This clearly underscores the necessity and superiority of more sophisticated models using a data-driven, non-linear approach.

Model Exploration

The core of our project lies in accurately modelling the charging speed function, $S(X)$. This function is highly non-linear; batteries typically charge rapidly at lower percentages and then significantly slow down at higher percentages – i.e. the constant voltage or so-called "trickle charge phase". We will explore various machine learning regression models to capture this complex relationship:

$$
\begin{array}{rll}
\text{Linear} & S(X) = a - bX & T(X) = \dfrac{1}{b} \ln\left(\dfrac{a - bX}{a - 80b}\right) \\
\text{Reciprocal} & S(X) = \dfrac{a}{1 + bX} & T(X) = \dfrac{1}{a}\left(80 + 3200b - X - \dfrac{b}{2}X^2\right) \\
\text{Rational} & S(X) = \dfrac{a + bX}{c + dX} & T(X) = \dfrac{80d - Xd}{b} + \dfrac{bc - da}{b^2} \ln\left(\dfrac{a + 80b}{a + bX}\right) \\
\text{Exponential} & S(X) = a\exp(-bX) & T(X) = \dfrac{\exp(80b) - \exp(bX)}{ab} \\
\text{Shifted Exp.} & S(X) = a\exp(-bX) + c & T(X) = \dfrac{1}{bc} \ln\left(\dfrac{a + c\exp(80b)}{a + c\exp(bX)}\right) \\
\text{Logarithmic} & S(X) = a - b\ln(1 + cX) & T(X) = ??? \\
\text{Power Law} & S(X) = aX^{-b} & T(X) = \dfrac{80^{b+1} - X^{b+1}}{a(b+1)} \\
\text{Binomial} & S(X) = a\left(1 + \dfrac{X}{k}\right)^b & T(X) = \dfrac{k\left((1 - X/k)^{1-b} - (1 - 80/k)^{1-b}\right)}{a(1 - b)} \\
\text{Hyperbolic} & S(X) = a\tanh(b(c - X)) + d & T(X) = ??? \\
\text{Logistic} & S(X) = \dfrac{A}{1 + \exp(k(X - X_0))} &
\end{array}
$$

$$T(X) = \frac{80 - X}{A} + \frac{\exp(k(80 - X_0)) - \exp(k(X - X_0))}{Ak}$$

where it is carefully noted that troublesome values of b do not blow up the solutions for the Power Law and Binomial models. Singularities reduce the models to the Linear model. It is anticipated that the Logistic sigmoid curve might be an okay fit for the charging speed, as it is assumed that there is a higher speed at the constant current phase and a lower speed at the constant voltage phase.

Solution Fitting

The tool used to train the models, output the parameters, and draw graphs is Desmos. The data is input as a list

$$L = \left[(X, S_{\text{emp}}(X)),\ \ldots\right]$$

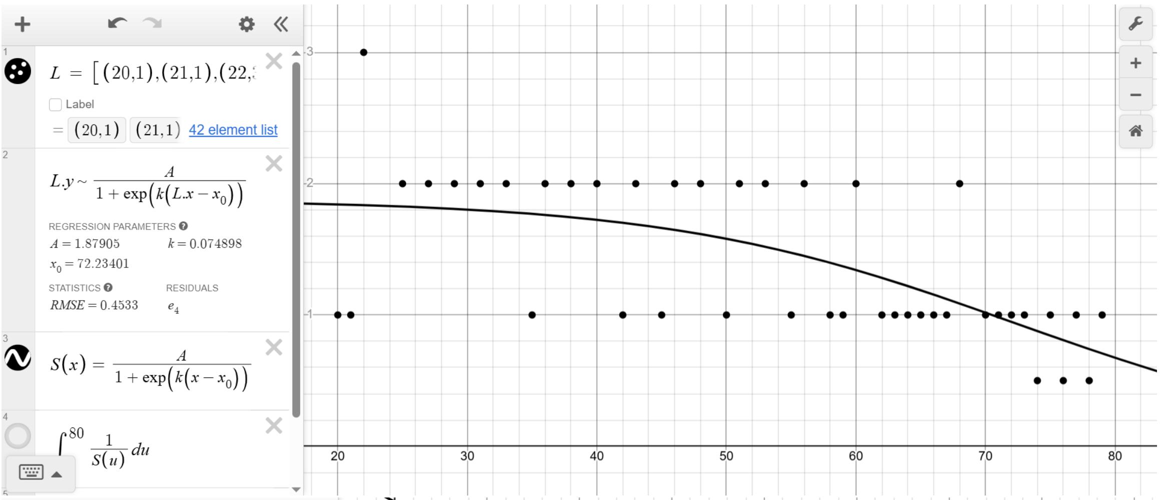
and the syntax

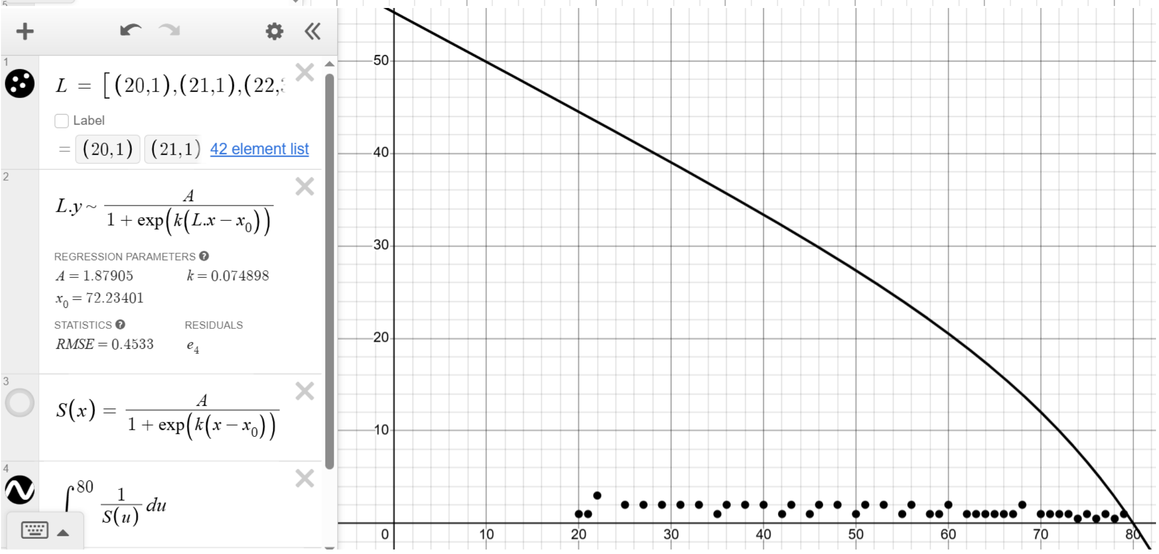
$$L.y \sim S(L.x)$$

fits the data and outputs the required parameters and the RMSE. The following two figures show Desmos being used for the Logistic sigmoid model, where

$$A = 1.87905 \qquad k = 0.074898 \qquad X_0 = 72.23401$$

The former shows $S(X)$ and the latter shows $T(X)$ in the following figures:





Selection and Implementation

Based on comprehensive evaluation, including qualitative fitting via Desmos graphing and quantitative performance metrics, we select the Logistic sigmoid model. The rationale for this selection is multifaceted:

- Physical Interpretation: The model inherently captures the observed physical charging characteristics. It accurately represents the fast charging at lower percentages (where $S(X)$ is higher) and the gradual deceleration of charging speed at higher percentages (as $S(X)$ decreases towards an asymptote), reflecting the battery's inherent chemical and electrical properties during the charging process.

- RMSE Performance: It demonstrated exceptional accuracy, yielding the second lowest RMSE of 0.4533 among all the models evaluated. This indicates a very close fit to the empirical charging speed data.

- Closed-Form Analytic Solution: Critically, the model has a closed-form analytic solution for $T(X)$. This eliminates the need for computationally intensive numerical integration, making real-time calculation significantly faster and more efficient, which is vital for on-device implementation.

While the Hyperbolic model exhibited the lowest RMSE value of 0.4529, it is quite minor. Its physical interpretation is less clear, and, critically, it lacked a closed-form analytic solution for $T(X)$, necessitating numerical integration which would complicate implementation and slow down on-device calculations. Thus, the Logistic model offered the optimal balance of accuracy, interpretability, and computation efficiency.

The Rational, Shifted Exponential, Logarithmic, and Binomial models also showed promising results, but were not selected due to either RMSE or not possessing a closed-form analytic solution.

The selected Logistic model for $S(X)$ is then used to calculate $T(X)$. This solution is then implemented as an automation within Apple's Shortcuts app. This application provides an on-demand estimate: when the device is plugged in and begins charging, the Shortcut automatically retrieves the current battery level, computes the remaining charge time using the derived $T(X)$ expression, and displays it as a notification for the user.

Results

The control group's linear extrapolation, based on a predetermined average charging rate, yielded inaccuracies of up to 10 minutes in its time-to-charge estimations. This significant deviation often led to false expectations and hindered effective planning, particularly during the critical final stages of charging where the actual rate slows considerably.

In stark contrast, our newly developed charge time curve, derived from the selected Logistic model, exhibited dramatically improved accuracy. Real-world testing showed that its predictions were typically within 1 to 2 minutes of the actual remaining charge time. This substantial reduction in error highlights the superior predictive power of a data-driven, non-linear modelling approach over traditional, simplistic methods. The enhanced accuracy provides users with reliable, actionable information, fulfilling the project's core objective of eliminating charging uncertainty.

The efficacy of our Logistic model for predicting battery charge time was clearly demonstrated when compared against the control group's simplistic estimation method.

Conclusion

This project successfully addressed the pervasive problem of unpredictable battery charge times, moving beyond simplistic estimations to provide users with precise, real-time insights. Our methodology systematically developed a data-driven solution, beginning with a clear mathematical framework and culminating in a practical, implemented application.

Specifically, we:

1. Formulated a clear plan to estimate $T(X)$, the remaining charge time, by leveraging the integral of the reciprocal of the charging speed $S(X)$, grounding our approach in robust mathematical principles.

2. Collected essential empirical data in the form of $(X, S_{\text{emp}}(X))$ points, capturing real-world charging characteristics of a device.

3. Explored and rigorously evaluated a diverse set of mathematical models for $S(X)$, such as Linear, Reciprocal, Rational, Exponential, Shifted Exponential, Logarithmic, Power Law, Binomial, Hyperbolic, and Logistic functions, also deriving or acknowledging their corresponding $T(X)$ solutions.

4. Conducted visual fitting analyses using Desmos to qualitatively assess how well each model aligned with the observed charging speed behaviour, complementing quantitative RMSE measurements.

5. Strategically selected the Logistic sigmoid model based on its strong physical interpretation of charging dynamics, its high accuracy (second lowest RMSE), and critically, its closed-form analytic solution for $T(X)$, which facilitates efficient on-device computation.

6. Successfully implemented the derived solution as an automation within Apple's Shortcuts app. This practical deployment allows the end-user to receive instant, accurate charge time notifications directly on their device upon plugging it in, transforming a frustrating unknown into a manageable predictable metric.

By bridging data science, mathematical modelling, and practical mobile automation, this project delivers a tangible enhancement to user experience, replacing guesswork with reliable, data-backed predictions for battery charge completion.

Disclaimer

It's important to acknowledge that this project was developed under significant time constraints. As such, while considerable effort was put into its development, I am not responsible for any inaccuracies that may arise from the model's predications.

Several areas present opportunities for future improvement and could impact the models' generalisability and long-term accuracy:

- Data Collection Limitations: The dataset was collected under specific conditions. More extensive and varied data collection would undoubtedly enhance the model.

- Unaccounted Variables: This model currently does not incorporate other crucial factors known to influence battery charging. These include, but are not limited to, device temperature, fluctuating voltage from different power sources, the impact of battery wear and age, specific charger specifications (e.g. wattage), and variations between different iPhone models. Including these parameters in future iterations would lead to a more robust and accurate predictor.

- Model Drift: Battery ageing is a continuous process that will inevitably cause model drift. This means the exact parameters derived here, along with the current model values, are not expected to accurately predict charging times for every iPhone indefinitely, nor for all iPhones out of the box. Regular recalibration or a more adaptive modelling approach would be necessary for sustained accuracy across diverse devices and over time.

This project serves as a proof-of-concept, demonstrating the viability of a data-driven approach to real-time charge time estimation. I look forward to iOS 19.