Project Report

A Neural Network-Based Forecasting and Analysis of Battery State of Health (SoH) for Predictive Maintenance

February 2024

Sara Kim

1. Introduction

1.1 Project Motivation

The project is motivated by the critical importance of battery systems in various applications, ranging from portable electronics to electric vehicles and renewable energy storage. The longevity and reliable performance of batteries are paramount as they directly influence the overall efficiency and sustainability of these systems

1.2 Background

Battery state of health (SoH) is a pivotal metric that encapsulates the overall condition and performance of a battery over its operational lifespan. Understanding SoH is crucial for predicting remaining life, optimizing usage, and implementing timely maintenance strategies.

Capacity and Degradation:

Battery capacity, representing the amount of charge a battery can store, is a fundamental parameter in assessing its performance. Over time, batteries undergo degradation, a natural process that leads to a reduction in capacity. The relationship between capacity and degradation is a key aspect of studying battery behavior.

State of Health (SoH) Formulation

SoH is often defined as the ratio of the current capacity to the original capacity of a battery. Mathematically, SoH can be expressed as:

$$SoH = \frac{Current\ Capacity}{Original\ Capacity} \times 100\%$$

This formula provides a standardized measure to quantify the health status of a battery, facilitating comparative analyses across different units.

Predictive Modeling using Neural Networks

Sophisticated predictive modeling techniques are implemented in this project, encompassing both Long Short-Term Memory (LSTM) networks and Support Vector Regression (SVR). These two distinct methodologies have been selected for their complementary strengths in capturing temporal dependencies and managing non-linear relationships within time-series battery data.

LSTM Networks: LSTM networks, as a type of recurrent neural network (RNN), are utilized for their proficiency in sequential data analysis. The incorporation of memory cells and gates enables the capturing of long-term dependencies, rendering them particularly well-suited for modeling the evolving state of health (SoH) in batteries over time. The contextual information retained from preceding cycles facilitates the accurate prediction of future SoH values, a crucial aspect for the implementation of proactive maintenance strategies.

SVR: In addition to LSTMs, Support Vector Regression (SVR) is integrated as an alternative modeling technique. SVR, being a potent algorithm capable of handling both linear and non-linear relationships in data, operates by mapping data points into a higher-dimensional space. This enables the identification of an optimal hyperplane that best fits the data, thereby allowing for robust prediction of SoH values. The versatility inherent in SVR complements the strengths of LSTMs, contributing to the development of a well-rounded and comprehensive predictive modeling framework.

1.3 Dataset Description

The dataset employed in this study originates from the esteemed NASA Prognostics Center of Excellence (PCoE) and encompasses experimentation on Lithium-Ion (Li-Ion) batteries. The experimental procedures involved comprehensive charging and discharging protocols conducted under varying temperature conditions. The impedance of the batteries was meticulously recorded and serves as the primary criterion for assessing damage.

The raw dataset was initially provided in MATLAB format, offering a structured representation of the experiments. This format facilitates a meticulous capture of various experimental parameters, ensuring comprehensive data granularity. To enhance accessibility and versatility for analytical and modeling purposes, the dataset underwent a transformation process. The data was extracted from the MATLAB format and reformatted into CSV files, introducing greater compatibility with a diverse range of data analysis and machine learning tools.

The dataset is further organized into distinct CSV files for charging and discharging cycles, categorized by individual battery names. This categorical distinction allows for a more refined analysis, considering the nuanced behaviors exhibited during charge and discharge processes. Subsequently, these segregated datasets are harmoniously merged, subjected to meticulous cleaning procedures, and augmented with an additional parameter: State of Health (SoH). The integration of SoH is paramount, as it encapsulates the overarching health status of the batteries, offering a comprehensive metric for subsequent data analysis and machine learning model training.

The columns consist of:

- Voltage_measured: Represents the voltage measured during battery experiments.

- Current_measured: Denotes the current measured during battery experiments.
- Temperature_measured: Indicates the temperature recorded during battery experiments.
- Current_charge: Signifies the current applied during the charging phase of the battery.
- Voltage_charge: Represents the voltage observed during the charging phase of the battery.
- Time: Captures the elapsed time during the battery experiments.
- Capacity: Reflects the capacity of the battery at a specific point in time.
- id_cycle: Identifies the cycle number of the battery experiment.
- type: Specifies the type of battery operation, such as charging or discharging.
- ambient_temperature: Records the ambient temperature during the battery experiments.
- time: Represents the time elapsed during the battery experiments (possible duplicate, check data dictionary).
- battery_name: Identifies the unique name or label assigned to each battery in the dataset.

2. Exploratory Data Analysis (EDA)

2.1 Battery Capacity and SoH Evolution During Discharge Cycles

Implemented through Streamlit, our application empowers users to select a battery from the dataset and dynamically plot its State of Health (SoH) and capacity evolution per discharge cycle. Streamlit's interactive interface provides a seamless experience for insightful battery performance analysis.

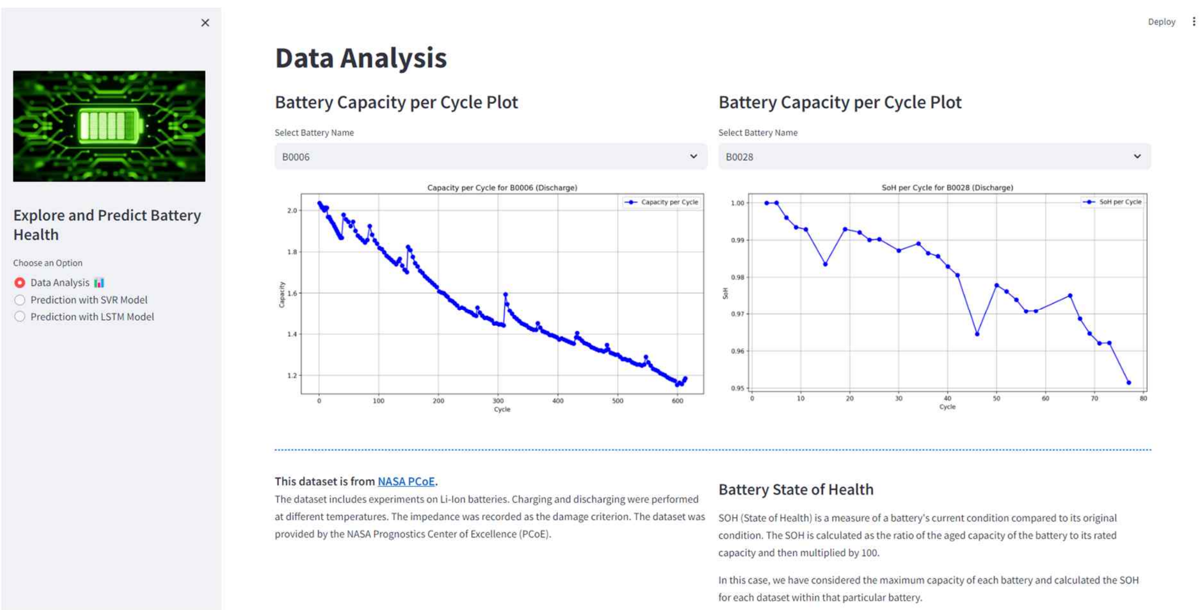
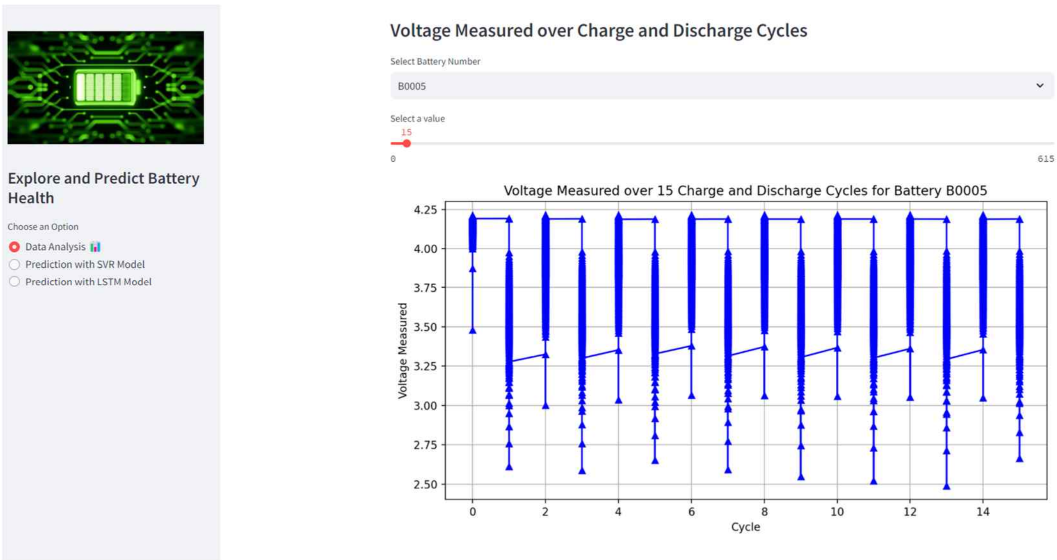
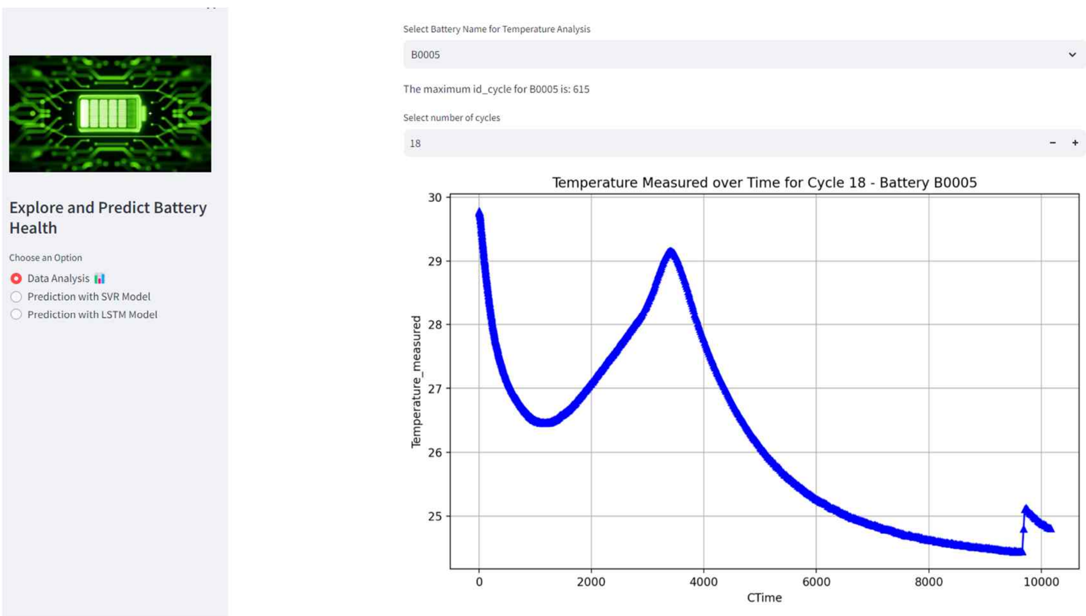


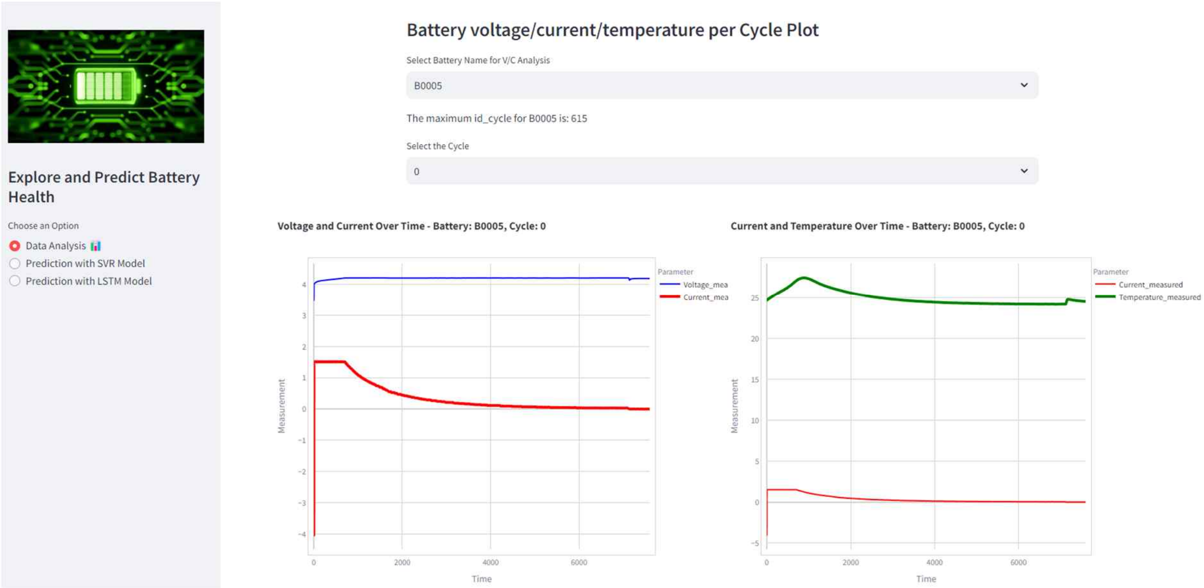
Figure 1: Capacity an SoH per Cycle Visualization

2.2 Voltage Measured, Temperature and Current Across Charge and Discharge Cycles

This section allows users to dynamically select a range of cycles for a specific battery, visualizing the measured voltage, current and temperature over the chosen cycles. Through Streamlit application, users can interactively explore how the voltage fluctuates across different charge and discharge cycles,.





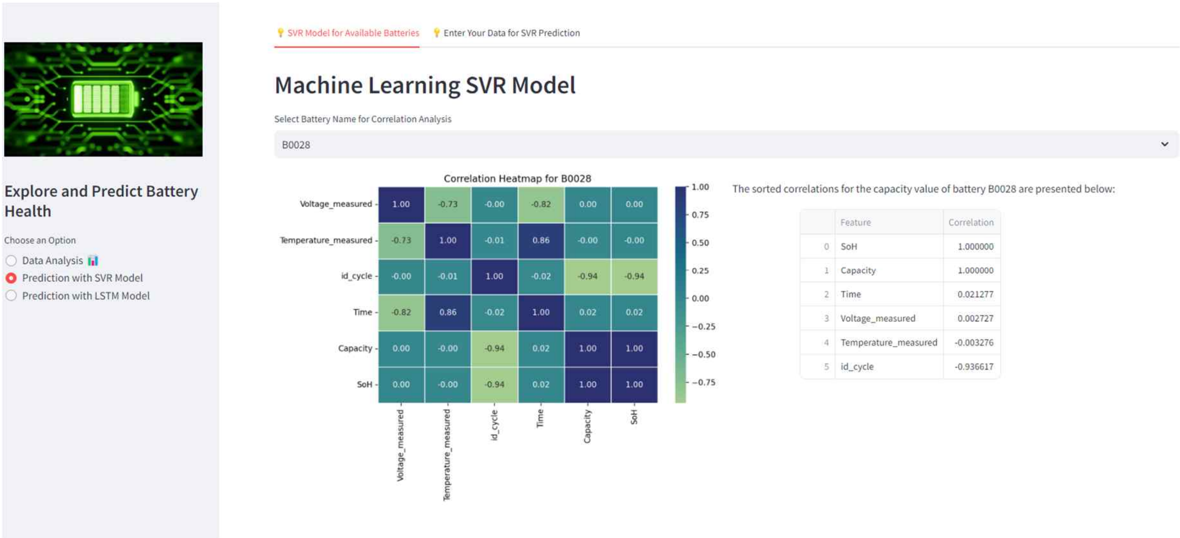


In charge cycles, rising temperature aligns with increased voltage and current, indicating heightened electrochemical activity. Conversely, discharge sees a temperature decline alongside reduced voltage and current. Abnormalities in these dynamics may signify inefficiencies or potential issues. Striking a balance in these relationships is pivotal for ideal battery performance, guiding the pursuit of enhanced energy storage systems and predictive maintenance strategies.

3. DL/ML Analysis:

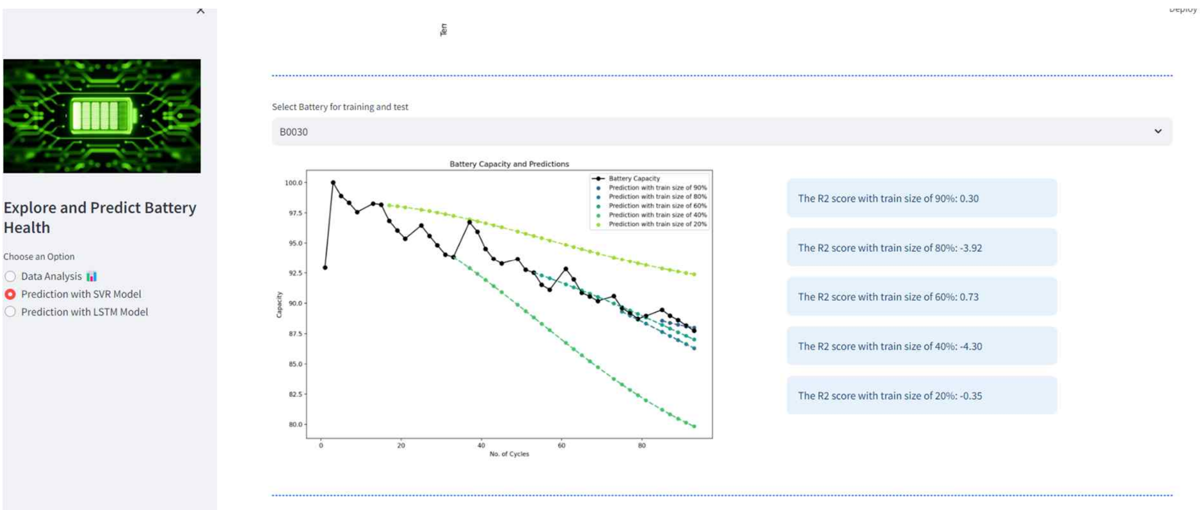
3.1 Exploring Feature Correlations with Correlation Heatmap

The machine learning model analysis in the Streamlit application incorporates a correlation heatmap, allowing users to explore the relationships between various features for a selected battery. The application presents a dropdown menu where users can choose the battery name for correlation analysis. Upon selection, the corresponding correlation heatmap is dynamically generated, providing a visual representation of feature correlations. This interactive feature enhances user engagement and facilitates a deeper understanding of the dataset, aiding in the interpretation of the machine learning models' performance.



3.2 Model Training and Evaluation by SVR model

In this section of the application, users can select a specific battery for training and testing machine learning models. The Streamlit interface offers a dropdown menu with available battery options. Once a battery is chosen, the application loads the dataset, extracts relevant features, and splits it into input (X) and output (Y) variables. The train_evaluate_plot function is then utilized to train the model, evaluate its performance, and generate plots showcasing the actual versus forecasted State of Health (SoH) for different ratios. The application automatically selects the best parameters for fitting the model,



3.3 Fine-Tuning Model Parameters: Customization Options for SVR model

This section of the application provides users with the ability to fine-tune the parameters of the Support Vector Regression (SVR) model. The Streamlit interface includes adjustable input fields for critical parameters such as the test size ratio, C parameter, epsilon, gamma, cache size, kernel type, max iterations, shrinking heuristic, tolerance for stopping criterion, and verbosity. Users can interactively set these parameters to customize the SVR model according to their preferences. The application then employs the train_evaluate_plot_param function to train, evaluate, and plot the model using the specified parameter values. Fine-tuning model parameters empowers users to optimize the predictive performance of the SVR model based on their specific dataset characteristics

3.4 Long Short-Term Memory (LSTM) Model: Training and Future Forecast

In this section of the application, the LSTM neural network model is trained and utilized for forecasting future battery performance. The training process involves scaling the input data, splitting it into training and testing sets, and configuring the LSTM architecture. The model is then compiled, trained using historical data, and saved for later use. Once trained, the LSTM model is applied to predict the State of Health (SoH) for future cycles, providing valuable insights into the anticipated behavior of the Li-Ion battery.