Project Proposal - Machine Learning, 2024/2025, 1st Semester
Doctoral Program in Electrical and Computer Engineering

Nelson Mateus Lorenzoni (up201800067@up.pt)
Pedro Vitor Soares Gomes de Lima (up202401143@up.pt)

November 14, 2024

Title: Machine Learning for Thermal Runaway Correlations in Li-Ion Batteries

1 Dataset

The Battery Failure Databank features data collected from hundreds of abuse tests conducted on commercial lithium-ion batteries. Methods of abuse include nail penetration, thermal abuse, and internal short-circuiting. This dataset provides the heat output from cells undergoing thermal runaway, the breakdown of heat from the cell casing and its ejected contents, as well as the mass of the cells before and after thermal runaway and the quantity of mass ejected from the cell [1].

2 Project idea

Thermal runaway in lithium-ion batteries poses a major safety risk in high-energy applications like electric vehicles and energy storage, as it triggers rapid temperature rise, gas release, and risks of combustion or explosion [2]. This study will focus on three variables: energy released as gas from the cell's top and bottom, energy release through conduction, and the relationship with cell mass and chemistry. The investigation will analyze input properties and their correlations with test result outcomes [3].

Therefore, we define our problem as a multioutput regression because we predict three numerical outputs. The main objective is to test different machine learning algorithms and choose the one that best describes our data behavior. We consider Linear Regression, K-Nearest Neighbors (KNN), and Tree-Based models, applying hyperparameter tuning to optimize performance. Evaluation metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2) score.

Regardless of the algorithm, we need to implement a preprocessing step to prepare the dataset for effective use. This includes checking for missing values and duplicates, transforming categorical variables, and normalization. In addition, we will perform feature selection to check for irrelevant, redundant, or noisy features that may not impact, or even negatively impact, the model's generalization capability.

3 Software

We will perform all experiments with Python 3.10 (or greater) and use Google Colab to work collaboratively. The main libraries include: *numpy* and *pandas* for data manipulation, *scipy* for statistics, *matplotlib* and *seaborn* for data visualization, and *scikit-learn* for machine learning algorithms and tools.

References

[1] National Renewable Energy Laboratory. *Battery Failure Databank*. Available at `https://www.nrel.gov/transportation/battery-failure.html`, last accessed on 13/11/2024.

[2] Matthew Sharp et al. "Thermal runaway of Li-ion cells: how internal dynamics, mass ejection, and heat vary with cell geometry and abuse type". In: *Journal of The Electrochemical Society* 169.2 (2022), p. 020526.

[3] Donal P Finegan et al. "The battery failure databank: Insights from an open-access database of thermal runaway behaviors of Li-ion cells and a resource for benchmarking risks". In: *Journal of Power Sources* 597 (2024), p. 234106.