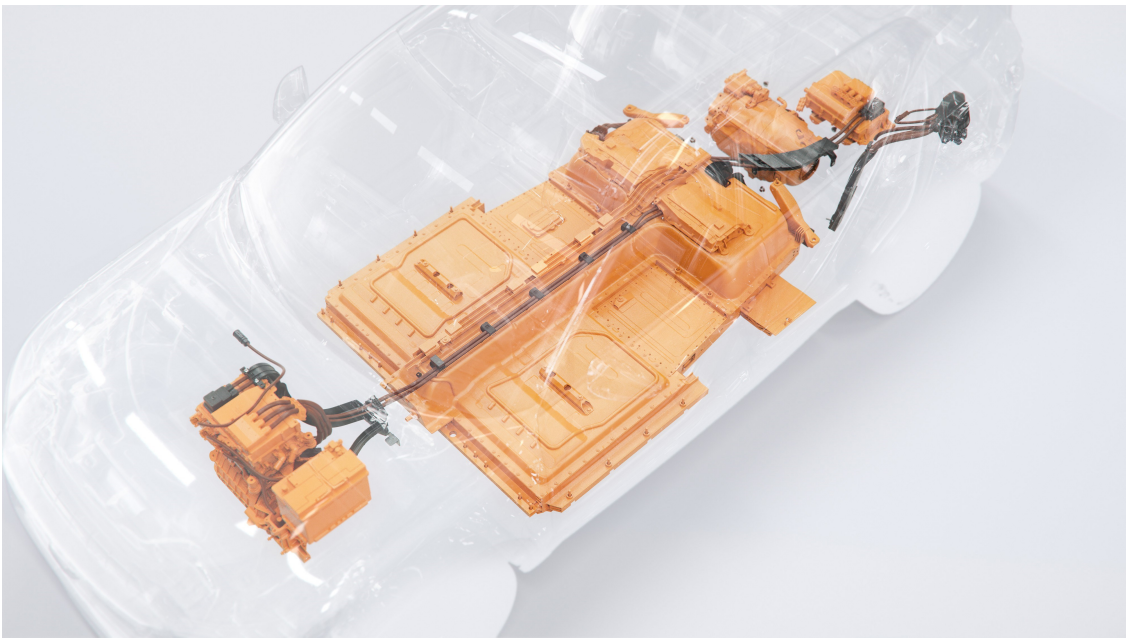
CHALMERS
UNIVERSITY OF TECHNOLOGY



# Predicting Battery Lifetime in a Fleet of Cars: An Application of a Transfer Learning-based Framework

Master's thesis in Systems, control and mechatronics

ANTON ALBINSSON, THOMAS TRINH

DEPARTMENT OF ELECTRICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2023
www.chalmers.se

Master's thesis 2023

# Predicting Battery Lifetime in a Fleet of Cars: An Application of a Transfer Learning-based Framework

ANTON ALBINSSON
THOMAS TRINH


CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2023

Predicting Battery Lifetime in a Fleet of Cars: An Application of a Transfer Learning-based Framework

ANTON ALBINSSON
THOMAS TRINH



Industry Supervisor: Johnny Ngu, Volvo Cars, Research & Development, Traction Battery Department
Examiner: Torsten Wik, Department of Electrical Engineering

Master's Thesis 2023
Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: The fully electric XC40 SUV – Volvo's commercial first electric car.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2023

Predicting Battery Lifetime in a Fleet of Cars: An Application of a Transfer Learning-based Framework
ANTON ALBINSSON
THOMAS TRINH
Department of Electrical Engineering
Chalmers University of Technology

# Abstract

The increasing adoption of electric vehicles (EVs) in recent years has put an increasing focus on an essential component in these vehicles: lithium-ion batteries. Accurately predicting the End-of-Life (EoL) of batteries is crucial for automotive manufacturers, such as Volvo Cars, to ensure performance and operational safety of their fleet of EVs.

Numerous methods exist today to predict EoL on EVs based on data from either a fleet of EVs or laboratory experiments. Predictions of EoL pose several challenges as empirical observations of EoL from EVs are limited, while the experiments in laboratory settings cannot fully replicate the operating conditions in a real-world setting.

This thesis explores a framework for predicting the EoL of a fleet of EVs by combining lab and fleet observation data using a network-based transfer learning model. The framework's purpose is to transfer degradation information from the lab data to real-world fleet data to predict EoL more accurately. As part of the framework, an ANN is first pre-trained on a dataset from one battery system and then fine-tuned on a dataset from another battery system.

To evaluate the framework's effectiveness, a two-part experimental study is conducted using lab datasets with degradation paths to EoL. The first part evaluates how the fine-tuning affects predictive performance, and the second part investigates the knowledge transfer between the pre-trained and fine-tuned model.

The experimental results demonstrate that the predictive performance improves with an adequate amount of fine-tuning data. Furthermore, fine-tuning the model also shifts the predicted degradation path closer to the ground truth, indicating that knowledge has successfully been transferred during the fine-tuning process. Furthermore, applying the framework with fine-tuning on a fleet of EVs, the model demonstrates reasonable degradation paths down to the EoL for the fleet, highlighting its ability to capture complex degradation mechanisms of batteries.

Keywords: end-of-life, electric vehicles, state-of-health, battery aging, battery degradation, machine learning, neural networks, transfer learning, fine-tuning, pre-training

# Acknowledgements

First of all, we would like to thank our industry supervisor Johnny Ngu and our colleagues, Robert Johansson and Veronica Menacho, for their valuable guidance in how we could transfer our knowledge from university into applications in the industry. We would also like to thank our examiner Torsten Wik, whose questions and feedback helped us fine-tune our approach to this fascinating problem.

The colleagues at Volvo Cars have been an invaluable source of feedback and insights into this domain and helped us to recharge our batteries during all the social activities in the office. We are also thankful for modern tools such as Grammarly and ChatGPT that are of great assistance in checking for grammatical and syntax errors in the text.

Finally, we would like to express our gratitude to our friends and family who have supported us with their energy and endless encouragement.

Anton Albinsson & Thomas Trinh, Gothenburg, June 2023

# List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

| | |
|---|---|
| ANN | Artificial Neural Network |
| BEV | Battery Electric Vehicle |
| EoL | End of Life |
| EV | Electric Vehicle |
| MAPE | Mean Absolute Percentage Error |
| RMSE | Root Mean Square Error |
| SoC | State of Charge |
| SoH | State of Health |
| SEI | Solid Electrolyte Interface |
| RNN | Recurrent Neural Network |
| RPT | Reference Performance Test |

# Contents

# List of Figures

# List of Tables

# 1
# Introduction

The transportation industry's reliance on fossil fuels causes significant carbon dioxide emissions [1]. To reduce carbon dioxide emissions, the EU has banned the sale of new petrol and diesel cars from 2035 [2]. One of the most promising approaches to reducing transportation emissions is the electrification of vehicles [3]. For automotive manufacturers, there has been a trend toward setting ambitious goals for selling Electric Vehicles (EVs). For example, the Swedish premium car manufacturer Volvo Cars, which this thesis is done in collaboration with, plans to sell only Battery Electric Vehicles (BEVs) cars from 2030 and onward [4]. As a result of this push towards EVs, the number of vehicles on the road powered by batteries is projected to increase significantly in the large automotive markets of China, the European Union, and the United States [5].

## 1.1 Background

The expected rise of EV adoption puts a focus on the battery, where the lithium-ion battery is a popular choice due to its high energy density and long lifespan [6]. Several important factors influencing consumer preferences for buying an electric vehicle, such as range and second-hand value, are related to the battery. Furthermore, batteries are costly to acquire and account for a significant part of an EVs total cost [7]. Therefore, understanding the battery's lifetime is of great interest from financial and environmental aspects for automotive manufacturers.

For the automotive industry, there are several benefits to having an accurate End of Life (EoL) estimation of batteries. Estimates of EoL are important for calculating warranty policies and insurance coverages. Predictions about future degradation of the battery can also be valuable when planning preventative maintenance. In addition, better estimations of a battery's health condition can be helpful when calculating the value on the second-hand market and when re-purposed for second-life applications, such as energy storage systems [8]. Increased knowledge of the battery aging mechanisms can also benefit the research of new batteries [9].

Various approaches to predicting EoL have been proposed, including empirical aging models, data-driven models, and physics-based models [10]. Each approach has limitations and challenges regarding development, implementation, and applicability. For instance, physics-based models require a resource-intensive modeling stage

involving battery cell disassembly and are bound by strict assumptions made of the battery system. On the other hand, data-driven approaches, like machine learning, are flexible and adaptable. They use large datasets to identify battery usage patterns without requiring extensive battery domain knowledge.

The surge in data collection for EVs has enabled the application of data-driven models [11]. This advancement has especially made machine learning models a research hotspot due to their flexibility and accurate EoL predictions [12]. However, the efficacy of these methods relies on the quality and quantity of collected battery data. A substantial amount of aging data is required to ensure accurate EoL predictions using data-driven models.

Field data, which refers to data collected from cars in the fleet, poses several challenges for developing accurate battery aging models using a data-driven approach. Limited labels of EoL due to the relatively short time that many EVs have been on the road makes it difficult to validate EoL predictions [9]. Additionally, field data introduces complexities related to sensor issues, data collection limitations, and uncontrolled operating conditions inherent in real-world usage. These complexities can result in noise, inconsistencies, and missing data, making developing accurate data-driven models using field data challenging.

To overcome the challenge of insufficient labeled data, researchers have explored the potential of transfer learning, a branch of machine learning that leverages data from other sources to compensate for the limited labeled data [13]. Transfer learning aims to enhance model performance on a specific task by leveraging knowledge acquired from related tasks. While transfer learning has shown promise in tasks like image and text classification, its application for time-series predictions, particularly in the domain of battery aging, is still relatively limited [14]. Nonetheless, the transfer learning framework offers a promising approach to mitigate the challenges of field data for modeling battery aging.

## 1.2 Previous work

A wide range of literature on EoL prediction for lithium-ion batteries uses data-driven approaches, primarily employing traditional machine learning techniques on datasets derived from laboratory tests [15]. For instance, Patil et al. [16] applied support vector regression to predict EoL using statistical features, including max, min, skewness, and kurtosis extracted from voltage, current and temperature curves obtained from lab data. While their approach showed promising results in terms of EoL prediction accuracy, it is worth noting that its applicability to a fleet of EVs may be limited due to solely utilizing data obtained in controlled laboratory settings.

However, in contrast to the many studies based on lab data, research on field data for battery lifetime prediction remains limited [9]. The scarcity can be attributed to the limited availability of publicly accessible datasets with real-world operating

conditions due to trade secrecy and industry competition concerns. Consequently, developing data-driven battery aging models and evaluating their performance encounter significant limitations when dealing with field data.

Transfer learning has been employed in modeling battery aging and has shown improvement in EoL predictions when there are limited EoL labels. This approach enables the transfer of knowledge between battery systems exhibiting similar degradation paths. For example, Che et al. [17] proposed a transfer learning framework that utilized a Gated Recurrent Neural Network to transfer relevant information between lithium-ion batteries with similar degradation paths. The model was initially trained on a dataset of batteries with EoL labels and then fine-tuned using only early degradation paths from a test battery. Using two data sources, instead of a single one, resulted in more accurate EoL predictions for the test battery. Another study conducted by Chehade and Hussein [18] employed a Gaussian Process Regression model, also leveraging transfer learning to enhance EoL predictions for battery cells with only early degradation data. The proposed method exhibited high prediction accuracy and robustness. Both these approaches showed improvements in EoL prediction compared to traditional machine learning models trained on a single data source, underscoring the potential of transfer learning.

Although the mentioned studies have made progress in modeling battery aging using transfer learning, there are still gaps and limitations that need to be addressed. A recent study by Liu et al. [14] emphasized the significance of tackling these gaps. Firstly, further exploration is needed to unlock the potential of transfer learning between batteries with different chemistries and degradation mechanisms. Secondly, there is a need to investigate the integration of laboratory and field data to develop more accurate EoL prediction models that account for real-world usage conditions. Lastly, it is essential to examine potential limitations and challenges associated with transfer learning, such as the impact of data quality and domain adaptation.

## 1.3 Project scope

In this thesis we have investigated the potential of using a transfer learning framework to generate EoL predictions within the domain of batteries for EVs. Expanding beyond the scope of most existing studies that solely focus on laboratory data, our project seeks to study a combination of field and laboratory data. The project scope includes three research questions that are explored using transfer learning models applied to lab and field datasets.

The three research questions are:

1. How much field data is needed for a transfer learning-based approach to improve predictive performance over traditional machine learning?
2. Are there any patterns of knowledge transfer between battery datasets during transfer learning?
3. How well does transfer learning perform when applied to a fleet of cars?

These research questions are explored by comparing a model trained on a single data source with a model trained on two data sources. The ideal comparison is between the lab and field datasets. However, due to limited degradation paths in the field dataset, EoL predictions are evaluated using two subsets of lab data.

The primary focus of this project is to investigate the performance of transfer learning over traditional machine learning rather than striving for optimal predictive performance. For this reason, a traditional machine learning model trained on a single data source is used as a baseline. The transfer learning model is based on and built upon this traditional machine learning model using an additional data source. Given this scope, activities focused on finding the best-performing model, such as exploring designs of model architectures and hyperparameter tuning, are not explored exhaustively.

The project scope includes a fleet dataset of BEVs and lab datasets provided internally by Volvo Cars. Detailed datasets of real operating conditions for a fleet of EVs are hard to find publicly available, and using this internal dataset from Volvo Cars enables the exploration of the research questions. One limitation is that the data is internal, and therefore figures derived from the data are shown in ratios or as normalized values for confidential reasons. The laboratory data has information about battery degradation to the EoL and is suitable for evaluating the effectiveness of transfer learning. The result from the evaluation of laboratory data helps assessing the usefulness of transfer learning to a fleet of cars.

## 1.4 Thesis outline

Chapter 2 introduces the theoretical concepts needed to understand the remaining chapters of this thesis. Chapter 3 outlines the different datasets used, and Chapter 4 describes the methodology used to perform the experiments. Results from the experiments are given in Chapter 5, and the discussion section follows in Chapter 6. Finally, Chapter 7 presents a summary and ideas for future work.

# 2
# Theory

In this chapter, the theoretical concepts necessary to be able to follow the rest of the report are presented. In Section 2.1, a background into batteries and their aging process is provided. In Section 2.2, a general description of machine learning followed by the specific method of neural networks is presented. Finally, in Section 2.3, the concept of transfer learning is introduced, and its areas of applications are covered.

## 2.1 Batteries

This section describes batteries, both in terms of their properties and the aging process.

### 2.1.1 Battery structure

The dominating battery type for EVs is lithium-ion batteries [19]. A battery cell in a lithium-ion battery consists of a set of principal components; positive electrode, negative electrode, separator, electrolyte, and current collectors. The separator divides the cell between the negative and positive electrode sides. It allows lithium-ions to migrate between the different electrodes while preventing the trespass of electrons, thus avoiding internal short circuit. The electrolyte is an ionic conductor and serves as the medium in which the ions travel between the electrodes. The current collectors' purpose is to enable the electrons to travel between the different sides of the cell via an external circuit [20].

A battery's cell chemistry commonly refers to the combination and ratio of materials used. This composition impacts the battery's properties, including its degradation rate under specific operating conditions. For example, certain cell chemistries may be more suitable for specific environments, such as high temperatures, while others may be preferable in colder climates [21].

During the discharge of the cell, lithium ions move from the negative electrode to the positive electrode through the separator, while electrons move as current via the current collectors and the electric circuit. The flow of electrons in the circuit connected between the two current collectors is the current. During the charging phase, the process is reversed, with lithium-ions and electrons moving in the opposite direction [22]. An illustration of the discharging cycle is shown in Figure 2.1.

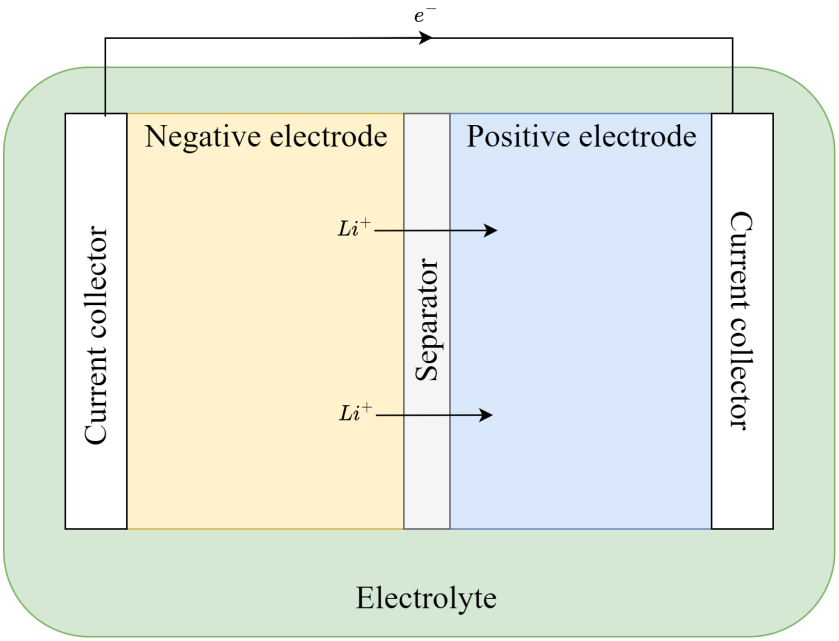



**Figure 2.1:** Overview of a lithium-ion battery during discharing.

Batteries for automotive applications are set up using many individual battery cells to achieve the necessary capacity and voltage for an electric vehicle. The battery cells are configured via serial or parallel connections to form a battery module, and connecting multiple battery modules form a battery pack. Cell configurations using parallel connections enable higher Ah capacity while connections in series provide the possibility for a higher supplied power [23].

### 2.1.2 Battery aging

Various degradation mechanisms impact a lithium-ion battery during charging and discharging as well as when the battery is at rest. Degradation mechanisms can have a mechanical or chemical origin and reduce the lifetime of the battery [24]. The State of Health (SoH) is used as an indicator to estimate the battery's overall condition. One way to assess the battery's SoH is by measuring the capacity at full charge. The SoH is then defined by

$$\text{SoH(t)} = \frac{Q_n^t}{Q_n^0} \times 100\ \% \tag{2.1}$$

where the $Q_n^t$ is the nominal capacity at time t and $Q_n^0$ is the initial nominal capacity.

For industrial applications, it is common to set a threshold of the SoH level that indicates when the battery has reached its EoL. In automotive applications, the threshold is commonly defined as a specific SoH value between 70 and 80%, indicating that the battery has significantly degraded and is no longer suited for its original application [25]. However, it might be used for other applications as the battery still has capacity.

There are many different degradation mechanisms that causes the decline of SoH, and two examples of mechanisms that have been the focus of intensive research are the solid electrolyte interface (SEI) growth and lithium plating [26]. SEI is a thin layer formed on the electrodes during the initial charge. While a thin layer has several beneficial purposes, such as protecting the electrode from the electrolyte, increasing the thickness of the SEI layer causes a loss of the lithium inventory and increased resistance. Lithium plating occurs with the formation of metallic lithium on the negative electrode during charging, which can result in loss of lithium inventory and active anode material. Both the loss of lithium inventory and active anode material degrades the capacity of the battery [27].

In a study by Yang et al. [26], the SEI growth was the dominant aging mechanism in the early stage, while in the later stage lithium plating became a more important mechanism. The study indicates that non-linear degradation can occur as a result of these degradation mechanisms and that they are influenced by other factors, such as temperature. For example, historical information about the storage temperature of the battery could be valuable to estimate its future degradation path.

The degradation mechanisms are complex, and one way to model the degradation process is to divide it into two parts based on if the aging is related to usage or time. *Cycling aging* is the part when degradation occurs as a result of charging and discharging the battery while *calendar aging* refers to the degradation as a result of time.

For the *calendar aging*, Barré et al. [24] describes two important conditions to attribute the rate of degradation caused by this aging process. The storage temperature is considered the primary condition, and extreme high and low temperatures can adversely affect the battery [24, 28]. The second condition of interest is the State of Charge (SoC) which refers to the amount of energy stored in the battery at a given time. These two conditions, temperature and SoC, are commonly explored together in a combination as their impacts on the degradation are not independent of each other [24].

*Cycling aging* are affected by the same conditions as the calendar aging as well as the battery's utilization mode. Two important characteristics of the utilization mode are the depth of discharge, indicating how much energy has been discharged, and the charging type. High levels of depth of discharge and fast charging are increasing the rate of degradation [24].

## 2.2 Machine Learning

Machine learning is a field where a program can perform a task using a model that learns from examples rather than explicitly instructing how the program should solve the task. Machine learning systems are sorted into different categories depending on the type of examples provided to the model. *Supervised learning* is a category where each example contains information about the desired output of the model, which is also known as the *target* or *label*. Machine learning systems are further classified based on the format for the desired output. Tasks that require a real-valued output to each example are known as *regression tasks* [29, Sec. 1].

In a supervised regression problem, each example $i$ consists of an input vector $x_i \in R^n$ and a target vector $y_i \in R$, where $n$ is the number of input features. The model is a function $f : R^n \rightarrow R$ that generates a prediction $\hat{y}_i$ based on the input $x_i$. The goal is to generate a prediction that is as close to the target as possible. A *performance measure* is an indicator of how good a prediction is with respect to the target [30, Sec. 5.1]. For regression tasks, there are numerous metrics to use, but two of the most common ones are *Root Mean Square Error (RMSE)* and *Mean Absolute Percentage Error (MAPE)* [31]. The RMSE is defined as

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(y_i - \hat{y}_i)^2}, \tag{2.2}$$

where $y_i$ is the desired target output, $\hat{y}_i$ is the model prediction and $k$ is the number of examples provided to the model. The MAPE is defined as

$$MAPE(y, \hat{y}) = \frac{1}{k}\sum_{i=1}^{k}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \cdot 100\%, \tag{2.3}$$

with the variables being the same as in Equation 2.2 [31].

For supervised learning, the model learns the function $f$ that converts the input vector $x$ to the target vector $y$ during the *training phase*, where the dataset used in this phase is called the *training set*. In the *training phase*, the model aims to find the parameters of the function $f$ that minimizes the error on the *training set* [29, Sec. 1]. The goal with supervised learning is generally not to find the model that achieves the best predictive performance on the training data but to find a model that can predict well on data it has not seen before. The idea is that the model learns patterns between the input vector $x$ and target vector $y$ and that these patterns also exist in data that is not part of the training data. One risk during the training phase is that the model detects and adjusts its parameters to specific patterns that only exist in the training data, which is known as overfitting. Therefore, the dataset is split into a training and test set to measure the model's ability to generalize to data unseen during the training phase. The training phase uses the training set, and the model is afterward evaluated on the test set, data that has not been seen before.

Allocating the examples in the data to either the training or the test set can be done randomly or systematically. The allocation method depends on whether the examples are independent or not. A case of a machine learning task where all examples are independent of each other is the classification of pictures [32, Sec. 14], such as classifying a picture as either having a cat or a dog. The examples do not depend on each other, so the dataset can be randomly split into a training and test set. One case of a machine learning task where the examples depend on each other is in time-series forecasting [32, Sec. 15], such as predicting future stock-market prices. The examples in this task can be the price of a stock in a day, where the opening price in one day is related to the closing price in the previous day. In this case, the allocation of examples is done systematically, setting up a specific date for the split and assigning days before this date into the training set and days after to the test set.

For many machine learning algorithms, additional settings allow for the improvement of learning during the training phase. These settings are known as algorithmic hyperparameters and are not part of the actual model parameters but control the learning behavior during the training phase. Finding the optimal hyperparameters can be computationally costly. Hence, a common approach is to define a range of potential values for each hyperparameter and randomly explore combinations of hyperparameter values in a grid search [30, Sec. 11.4]. The selection of hyperparameters can be evaluated by allocating data into an additional validation set, where predictions from different settings can be evaluated [29, Sec. 1].

### 2.2.1 Neural network

Neural networks are a category of machine learning models that have become popular for many applications, such as image recognition and natural language processing, due to their ability to detect complex and non-linear patterns in the data [32, Sec. 10]. A neural network consists of layers, each with a set of neurons. The layers are categorized into the input layer, which receives the input vector $x$, the output layer, which produces the output $\hat{y}$, and the hidden layer(s), which are the layers between the input and output layer [32, Sec. 10]. An example of an Artificial Neural Network (ANN) is shown in Figure 2.2.

The value of a neuron, $z_i$, at position $i$ in a hidden layer is defined as

$$a_i = \sum_{j=0}^{L} (w_{i,j} \cdot z_j) \tag{2.4}$$

$$z_i = h(a_i) \tag{2.5}$$

where $a_i$ is the activation of the neuron, $h$ is the activation function, $L$ is the number of neurons in the previous layer, $w_{i,j}$ is the weight coefficient between neuron $i$ in the current layer and neuron $j$ in the previous layer and $z_j$ is the value of the neuron $j$ in the previous layer. The coefficient $w_{i,0}$ is called the bias term as the $z_0$ is set to 1 for all $i$, where the other coefficients are known as the weights of the model [29, Sec.

5.1]. The purpose of the activation function is to introduce non-linearity between the layers to capture complex patterns. Some common activation functions are the sigmoid and tanh functions [29, Sec. 5.1].

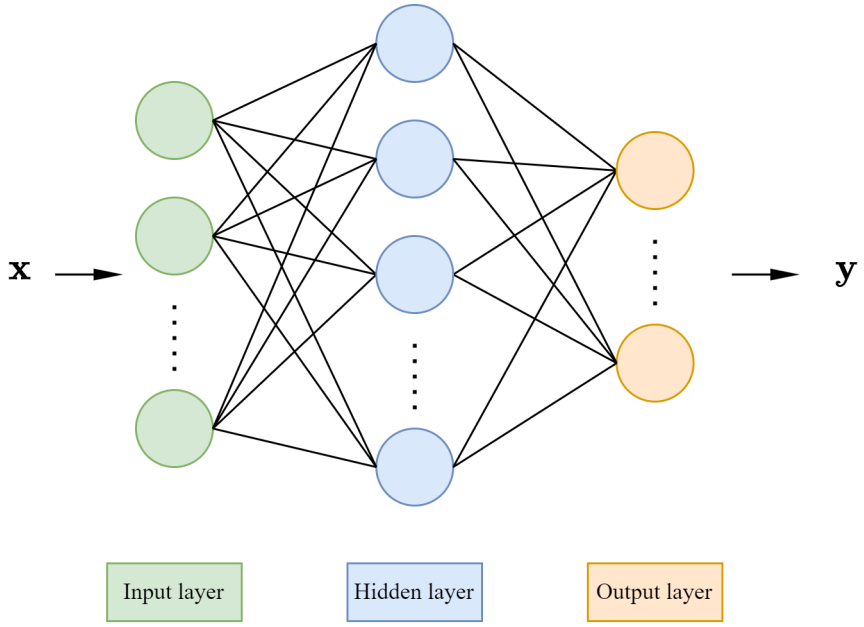



**Figure 2.2:** The layout of an Artificial Neural Network with an input vector of $\mathbf{x}$, a resulting output vector of $\mathbf{y}$ and one hidden layer.

During the training phase, the inputs $x_i$ are first passed through the network from the input layer via the hidden layers until the output layer generates a prediction $\hat{y}_i$. The loss for a single example, $l_i$, is calculated as

$$l_i = L(y_i, \hat{y}_i) \tag{2.6}$$

where $L$ is a loss function and $y_i$ is the target vector. The loss function is only used during the training phase of the model and can be different from the performance measure used when evaluating the model on the test set. After the loss is calculated, a step called *backpropagation* is performed, where the weights and biases of the neural network are updated based on the loss [30, Sec. 6.5].

Many settings must be explored during the training phase to find the optimal predictive performance. The model hyperparameters, such as the number of layers and neurons in each layer, are one of those settings. In addition, an extensive set of algorithmic hyperparameters can be explored, such as the learning rate, which influences how much the weights and biases are updated at each backpropagation. The training phase poses several challenges, such as tuning a large set of hyperparameters, high computational cost, and risk of overfitting [32, Sec. 11].

There are different types of techniques to reduce the risk of overfitting during the training phase. One set of techniques focuses on adding particular types of layers

to the model. Two examples of such layers are the *dropout* and *batch normalization* layers. The dropout layer randomly drops out a fraction of the neurons in each layer, except for the output layer, which remains the same. This method prevents the network from becoming too reliant on any individual neuron since if a neuron that the network relies heavily on is dropped, the model will make a poor prediction, resulting in a large loss. This results in the model learning a more robust representation that can perform well even if a few nodes are dropped out. The batch normalization layer normalizes the output from a layer for the batch in the training phase. One benefit of this approach is that it introduces noise during the training phase, which requires the model to learn more robust representations capable of handling noise [32, Sec. 11].

## 2.3 Transfer learning

*Transfer learning* is a concept that can be illustrated through a relatable example. Imagine two individuals learning to play different musical instruments [33]. One person has never played any instrument before, while the other has extensive experience playing the guitar. Now, suppose the person with guitar experience decides to learn the piano. In that case, they can transfer their knowledge of music theory, rhythm, and finger coordination from playing the guitar to quickly adapt to the piano. This ability to transfer knowledge and skills from one domain to another is the essence of transfer learning, and it has become a popular and promising area in machine learning.

In traditional machine learning, the data-driven algorithms rely on large amounts of labeled training data that closely resemble the target task or problem. However, obtaining such extensive labeled data in real-world scenarios can be challenging, time-consuming, and expensive [33]. This limitation poses a significant hurdle for applying traditional machine learning techniques effectively to many real-world problems.

To overcome this challenge, transfer learning provides a framework that enables the implementation of a data-driven approach even with insufficient labeled data for a given task. The idea is to leverage knowledge acquired from other related *tasks* or *domains* with sufficient labeled data. That is, by utilizing the information learned from a different but related problem, the performance of the machine learning models in real-world applications can be improved. A schematic diagram of the difference between the application of traditional machine learning and transfer learning is shown in Figure 2.3.

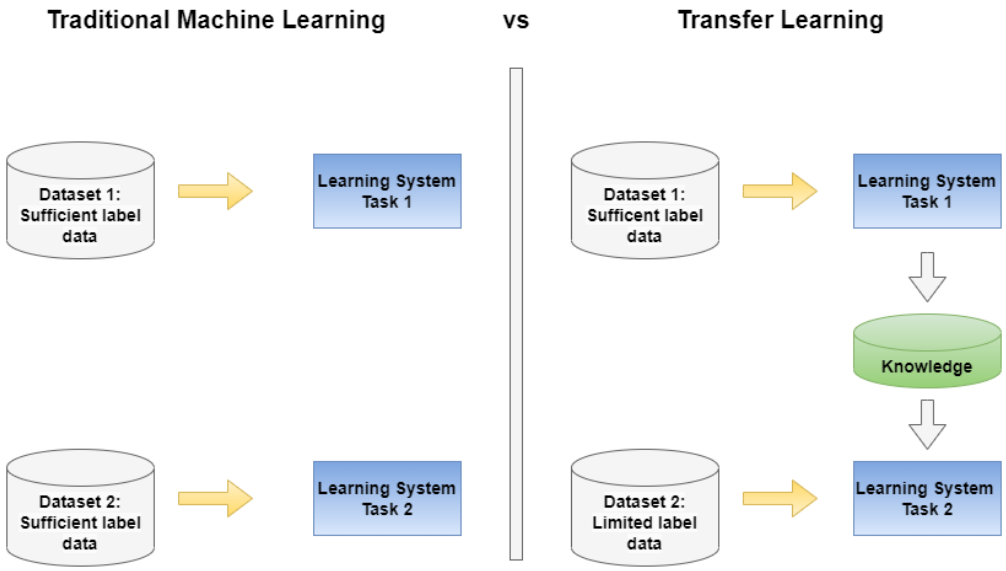



**Figure 2.3:** Schematic diagram illustrating the difference between traditional machine learning and transfer learning.

To illustrate how transfer learning works for a regression problem, consider the scenario to predict the human life expectancy for a generation born in the year 2000. Imagine there is data with three input values of each individual: the number of calories, alcohol, and sleep that the person is having, and one output value of the yearly health status of this person. The data is split into two datasets: source and target.

The source dataset contains information about people born in the year 1000. No one is alive anymore in this dataset, so perfect labels exist for all persons' lifespans. The target dataset contains information about people born in the year 2000. Everyone is alive in this data source, so no labels exist on any person's lifespan. One difference exists between the source and target datasets. In the year 1000, there was a problem with starvation, so eating more calories resulted in better health status. In the year 2000, there was instead a problem with obesity, so eating more calories resulted in worse health status.

Training a prediction model on just one of these datasets, either source or target, would be problematic for predicting the life expectancy of the generation born in the year 2000. Using only the source dataset for model training leads to learning the incorrect pattern that more calories are associated with positive health status. On the other hand, using only the target datasets is challenging since there are no labels about the potential lifespan of individuals.

Transfer learning could enable learning valuable patterns from both datasets. For example, a pre-trained model can be built from the source datasets to incorporate information about individuals' entire life span. A fine-tuned model could be trained on the target domain to, for example, learn the new pattern of calories being a negative indicator of health. This example is illustrated in Figure 2.4.

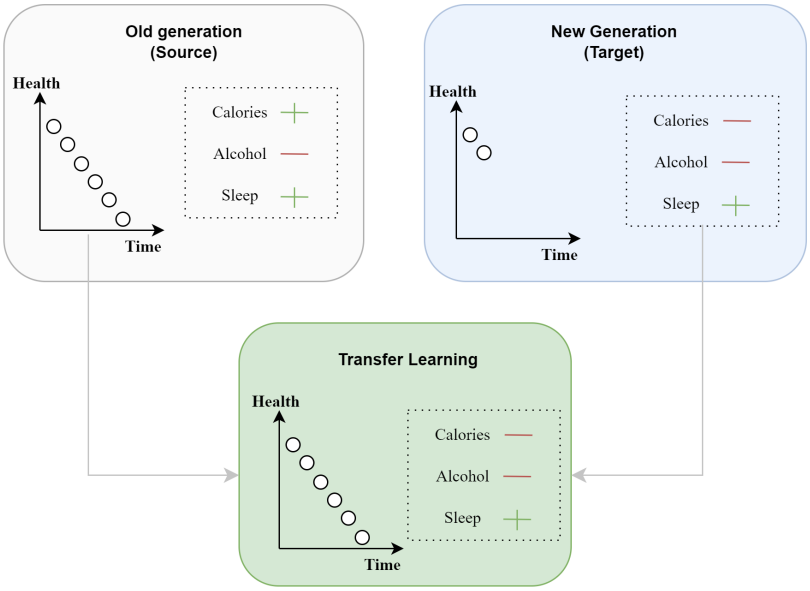



**Figure 2.4:** An example of transfer learning applied to a regression problem.

### 2.3.1 Definition

Transfer learning revolves around two key concepts: *domain* and *task* [34]. The domain refers to a source from which data is collected or generated. It comprises the input data and corresponding labels for training the pre-trained model. The domain is represented as $D = \{\chi, P(X)\}$, where $\chi$ is the feature space and $P(X)$ denotes the marginal probability distribution where $X = \{x_1, \ldots, x_n\} \in \chi$.

On the other hand, the task refers to a particular learning objective or problem. It represents the target of the learning process. A task is denoted as $T = \{Y, f(\cdot)\}$ and consists of two components: the label space $Y = \{y_1, \ldots, y_m\}$ and the objective predictive function $f(\cdot)$, which is to be learned from the pairs $\{x_i, y_i\}$.

Within transfer learning, a source domain denoted as $D_s = \{(x_{s_1}, y_{s_1}), \ldots, (x_{s_n}, y_{s_n})\}$, where $x_{s_i} \in \chi_s$ and $y_{s_i} \in Y_s$, represent the source instance and its corresponding class label. Additionally, $D_t = \{(x_{t_1}, y_{t_1}), \ldots, (x_{t_n}, y_{t_n})\}$ represents the target domain, where $x_{t_i} \in \chi_t$ is the target instance, and $y_{t_i} \in Y_t$ is its associated class label.

Transfer learning utilizes a source domain $D_s$ and its learning task $T_s$ for a learning task $T_t$ with a target domain $D_t$. The primary goal of transfer learning is to enhance the learning process of the target predictive function $f_t(\cdot)$ in $D_t$. This is achieved by leveraging the knowledge and information obtained from the source domain $D_s$ and its learning task $T_s$.

### 2.3.2 Negative transfer learning

*Negative transfer learning* occurs when knowledge gained from a source task has an adverse effect on the performance of a target task [33]. If the source task is not relevant or similar enough to the target task, it can result in a negative transfer. For instance, if the features or patterns learned from the source task are not applicable to the target task, noise might get introduced during training, leading to worse performance.

Several factors can contribute to negative transfer learning. One of the factors is the difference in the data distribution between the source and target tasks. If the data distributions are significantly dissimilar, the transferred knowledge may not generalize well to the target task. Another factor is the difference in representations learned by the models in the source and target tasks. Negative transfer can also occur if the source task has learned representations that are irrelevant to the target task.

### 2.3.3 Network-based transfer learning

Network-based transfer learning refers to reusing the partial or entire neural network that has been *pre-trained* in the source domain and transferring it to be part of a neural network used in the target domain [34]. Components that could be reused are the network structure, such as the layers and its corresponding parameters. A schematic diagram of network-based transfer learning is illustrated in Figure 2.5

A neural network is characterized by its multiple hidden layers, as mentioned in Section 2.2.1, which can capture intricate non-linear representations of data. The neural network can therefore be viewed as an abstraction process where the initial layers can be treated as a feature extractor. The extracted high-level features in the initial layers can then be reused on a different target task by freezing the initial layers and *fine-tuning* the parameters on the final layers. This enables the network model to leverage the knowledge captured from a related task or domain, leading to better performance, especially when the target task has limited label data.

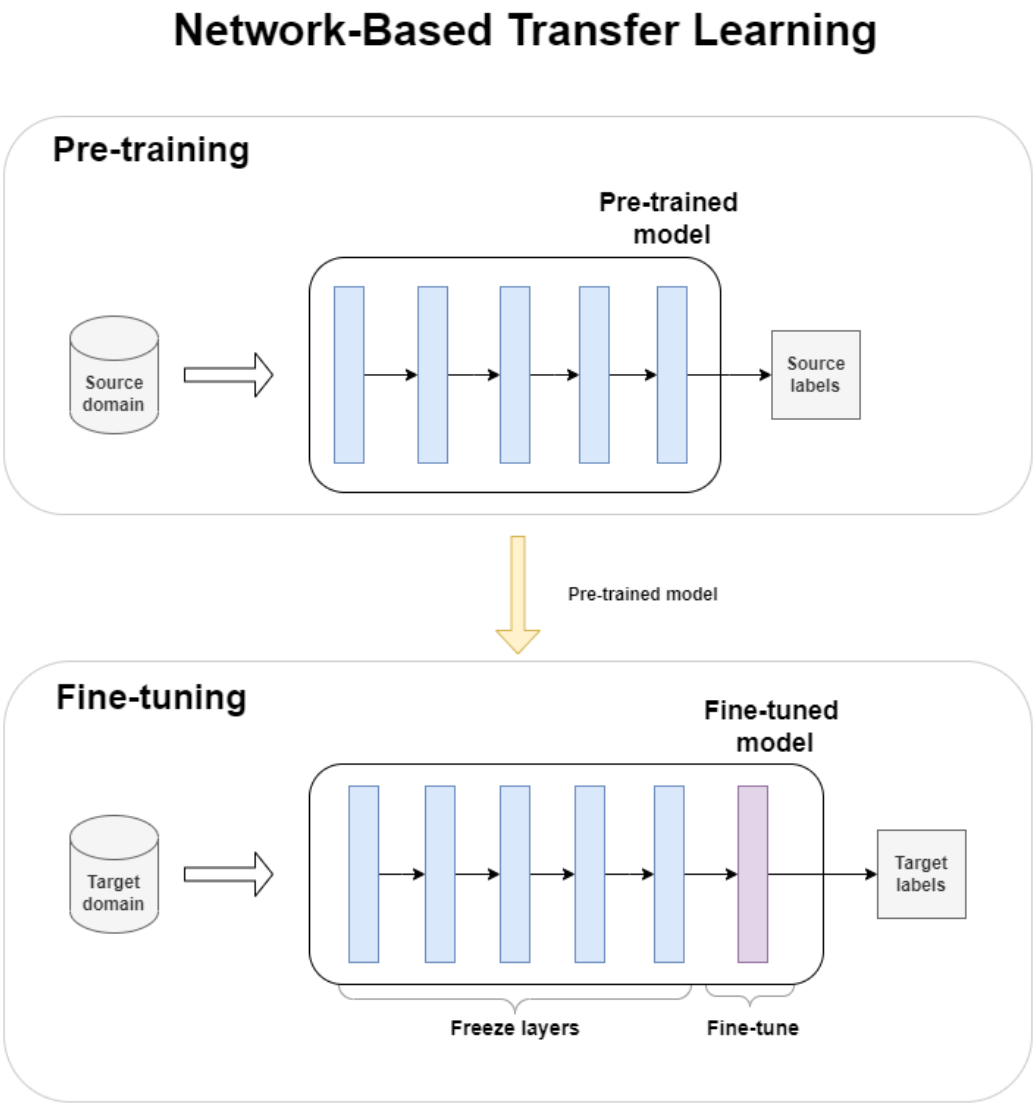



**Figure 2.5:** Schematic diagram illustrating an example of the pre-training and fine-tuning process in network-based transfer learning. In the pre-training stage, a neural network is used to train on the source domain. In the fine-tuning stage, the entire network from the pre-trained model is used, with its parameters frozen. Furthermore, an additional output layer is added on top of the network, which is then fine-tuned specifically for the target task.

# 3
# Datasets

The datasets have their origin in collections of sensor values from either a laboratory setting or from EVs in a fleet. Two types of laboratory data, one with raw data and the other with a processed version of the raw data, are presented separately in the two first sections. The data collected from EVs in a fleet is subsequently described in Section 3.3.

## 3.1 Raw lab data

Volvo Cars conducts laboratory tests on battery cells with different cell chemistries operating under *controlled conditions*. These tests can be categorized into cycling tests and calendar aging tests. In cycling tests, cells undergo continuous charge and discharge cycles according to predefined schedules. The cycling schedule may involve different SoC windows, such as charging/discharging between 25% and 65% or 6% and 96%, utilizing either fast charging or other charging methods. These tests are conducted in a setting with a constant temperature (e.g., 10°C, 20°C, 30°C). Calendar aging tests, on the other hand, involve storing the cells at a fixed SoC level (e.g., 40%, 70%, or 96%) at a constant temperature without performing any cycling.

Measurements for both calendar and cycling tests are obtained from Reference Performance Tests (RPTs) conducted during a testing phase to assess the performance and characteristics of the cells. During RPTs, various parameters are recorded, including the cells' capacity and resistance. The frequency of RPTs varies depending on the specific test.

A limitation of the raw data when constructing a data-driven model is that most tests are either cycling aging or calendar aging tests. There are few tests where battery cells undergo a schedule with cycling and extended storage periods, reflecting the degradation in real operating conditions. Furthermore, the number of conducted tests is limited due to time and resource constraints. The combined effect of calendar and cycling aging needs to be considered for the data-driven model to learn the patterns in real-world operating conditions. Nevertheless, these tests provide valuable insights into how batteries age under specific operating conditions. Furthermore, the measurements from the RPTs can be processed into full degradation paths to EoL for different simulated driving behaviors using a battery aging model.

## 3.2 Processed lab data

The *processed lab* data comprises a dataset containing battery degradation paths for a range of *simulated customer* profiles. Each simulated customer is characterized by four key parameters that reflect a simplified representation of their driving behavior. The parameters are average SoC level, usage of fast charging, average driving temperature, and annual mileage (see Table 3.1). The four parameters for a particular customer are assumed to be constant throughout the lifespan of the customer's battery. That is, the data does not account for the changes in driving behavior observed in real-world scenarios. For instance, Customer 1 will drive 10,000 km annually, never utilize fast charging, experience an average car temperature of -9 degrees Celsius, and maintain an average SoC level of 30%.

**Table 3.1:** Five customers and their driving parameters.

| **Customer** | **km/year** | **Fast charging ratio (%)** | **Temp (°C)** | **SoC Level (%)** |
|---|---|---|---|---|
| Customer 1 | 10000 | 0 | -9 | 30 |
| Customer 2 | 118000 | 0 | 30 | 30 |
| Customer 3 | 20000 | 10 | 20 | 30 |
| Customer 4 | 2000 | 25 | -9 | 30 |
| Customer 5 | 29000 | 25 | 30 | 30 |

For each of these four parameters, a discrete set of values has been selected, resulting in a total of 1,440 unique customer profiles (see Appendix A.1). For every unique customer, an SoH degradation path has been estimated from 100% to 70% using a model developed by Volvo Cars. This model has been constructed by using different tests in the raw lab data. By utilizing observations from calendar and cycling aging tests, the model combines aging factors to create a representation resembling real-world driving behavior. The degradation path to EoL for different customers is shown in Figure 3.1.

The processed lab data can be divided into six different datasets, each consisting of degradation paths for the 1440 customers for a certain cell chemistry. Two of the six cell chemistries are selected based on the occurrence of the most common cell chemistries in batteries for the fleet of cars. The cell chemistries are referred to as *lab data A* and *lab data B*. Comparison of the estimated EoL values of all 1440 customers for the two different cell chemistries is shown in Figure 3.2. Observe that cell chemistry A has a shorter average EoL than cell chemistry B.

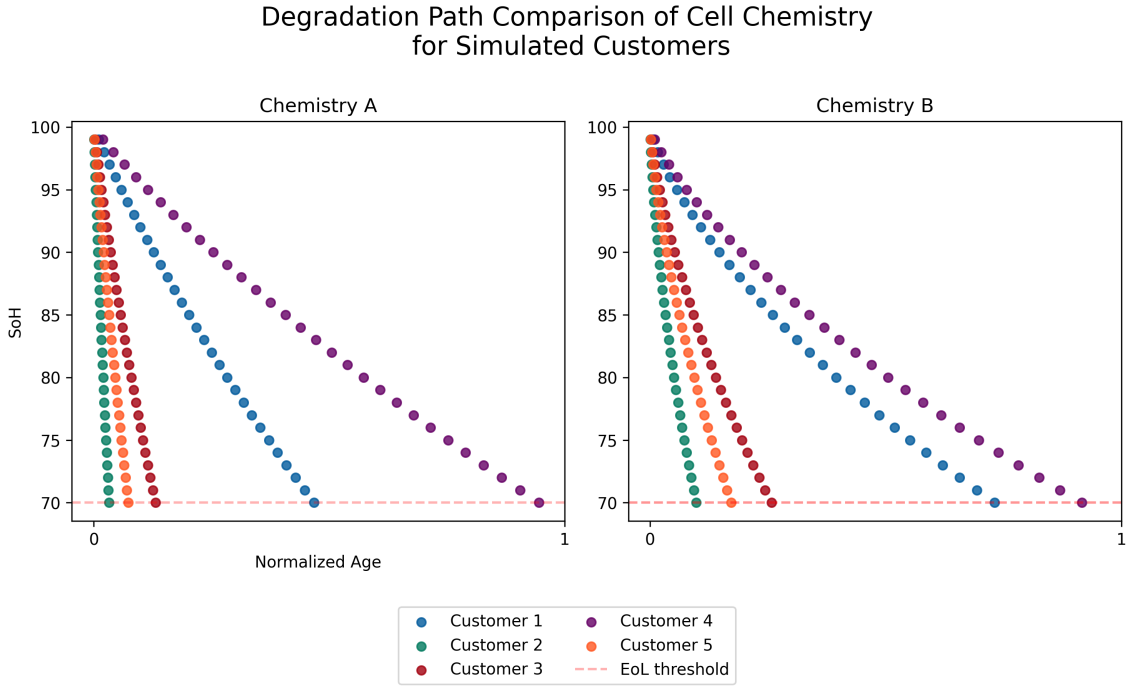



**Figure 3.1:** Degradation paths of five different customers for chemistry A and B.

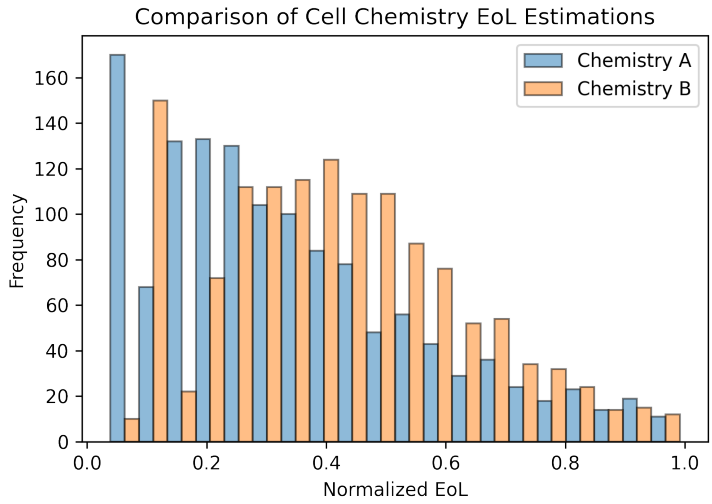



**Figure 3.2:** Histogram of the EoL estimation of all the customers for chemistry A and B.

There are limitations of the processed dataset as it has been generated using a Volvo Cars internal model. This processing of data can introduce biases and errors, which could propagate into subsequent models. Nevertheless, the dataset provides full degradation paths and EoL labels, which could be used to pre-train a data-driven model to learn the relationships between controlled driving patterns and battery degradation to EoL.

## 3.3 Field data

The *field dataset* consists of data collected from BEVs in Volvo Cars' customer fleet. This fleet contains cars from various parts of the world, resulting in a dataset containing information about battery degradation from a wide variety of environments. The cars are continuously collecting sensor values when in use. Storing raw sensor values in a car is problematic due to concerns about storage limits and the car owner's privacy protection. It is, therefore, common to transform raw sensor values, such as mapping raw values into a histogram with pre-defined bins, to avoid storing the raw values. For this reason, many of the signals extracted from the vehicle during a readout are not as the raw signals but in a transformed version, for example, as a histogram.

Readouts of the collected data are performed when the vehicle visits the workshops, typically for the yearly inspection, but other reasons might lead to more frequent visits for some vehicles. The signals extracted during a readout that is part of the field dataset are described in Table 3.2.

**Table 3.2:** The format of signals in the field dataset.

| **Signal** | **Format** |
|---|---|
| Production date | Timestamp |
| Readout time | Timestamp |
| State of Health | Numerical |
| Fast Charge | Numerical |
| Odometer | Numerical |
| Temperature | Histogram |
| State of Charge | Histogram |

An example of the variety of values within a signal is shown in Figure 3.3 with the temperature signal. A wide range of temperatures is observed in the dataset, resulting from a large fleet of cars in different regions and the collection of sensor values throughout all seasons. Between the lowest observed temperature, $T_{\text{min}}$, and the highest observed temperature, $T_{\text{max}}$, most of the values in the interval are observed. This dataset, therefore, provides information about how batteries perform in many different types of temperatures.

The collection of the lowest SoH value reached by each car in the fleet is shown in Figure 3.4. The lowest SoH value observed by any car in the fleet is $S_{min}$, with only a small amount of cars having values close to $S_{min}$. Volvo Cars are using the common practice in the automotive industry to define an SoH value of 70% as the EoL threshold. The value of $S_{\text{min}}$ is higher than 70%, which means that no cars in the dataset have reached the EoL, and consequently, there are no labels for EoL.

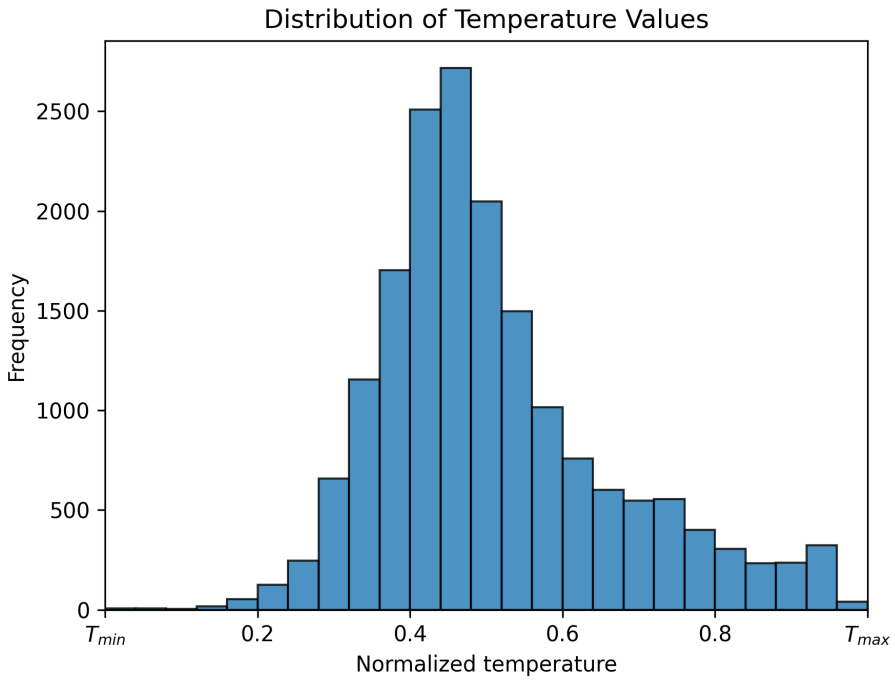



**Figure 3.3:** Histogram of the temperature values observed in the fleet normalized between the lowest $T_{min}$ and the highest $T_{max}$ observed values.

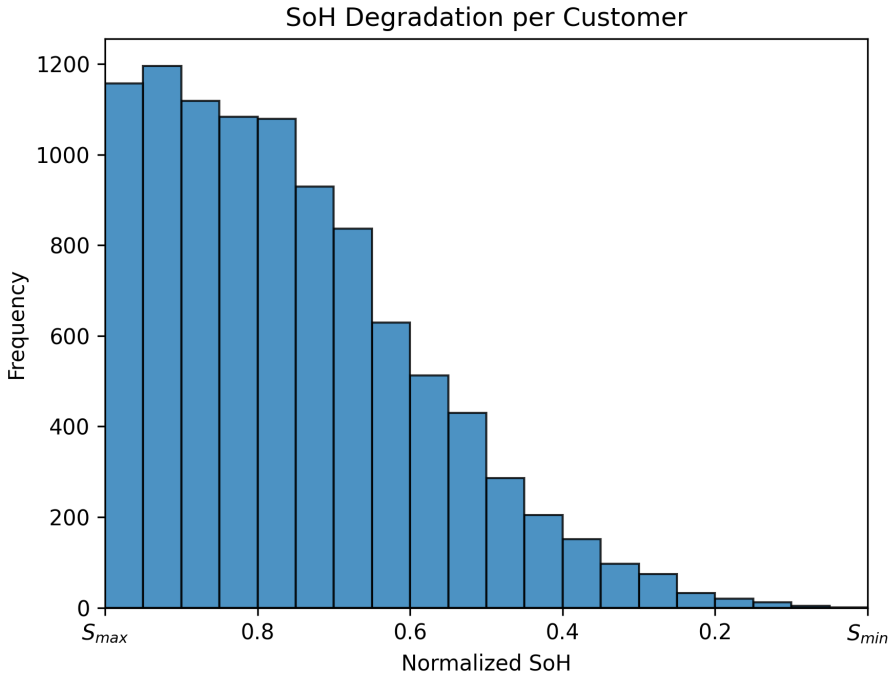



**Figure 3.4:** Histogram of the lowest SoH values observed per customer. Values are normalized between the highest SoH value $S_{max}$ and the lowest $S_{min}$.

The SoH signal is one of the cases where the value is not taken from a raw sensor value but is the result of computations in the vehicle. The computation of the SoH is based upon other types of raw sensor values, and the quality of the estimation is based on specific criteria on these raw sensor values. As a result, the SoH value in the field dataset is an estimation.

One challenge with the field dataset is that since many cars in the fleet only visit the workshop once a year, there are many cars with a limited amount of readouts. This leads to a sparse dataset, especially for signals such as the SoH, which are not stored in a histogram but only exist as a single numeric value during the readout. Another challenge is that storing signals in a histogram strips the information about the collection order for each value, which means that time-dependent information can only be seen between readouts but not within an individual readout.

# 4
# Methods

In this section, the methods used in the project are described. First, the steps in the data pre-processing for the lab and field datasets are explained are explained in Section 4.1. In Section 4.2 the transfer learning framework for EoL predictions is presented followed by the implementation details outlined in Section 4.3. Finally, in Section 4.4 the setup of the experiments are detailed.

## 4.1 Data pre-processing

The processed lab data described in Section 3.2 will hereby be referred to as lab data.

### 4.1.1 Lab data

A set of pre-processing steps are applied to the lab data to ensure the data's quality and quantity are sufficient for performing experiments on it. The pre-processing steps are outlined in Table 4.1.

**Table 4.1:** List of pre-processing steps for lab data.

| Step | Description |
|---|---|
| 1 | Clean data from outliers |
| 2 | Extend the degradation path |
| 3 | Generate training data |
| 4 | Add additional features |
| 5 | Split data used for pre-training |
| 6 | Split data used for fine-tuning |

The dataset undergoes an initial cleaning process to remove outliers in the form of simulated customers with extremely high EoL values, which are considered unlikely to occur in real-world data.

In the original dataset, the degradation path for each simulated customer ends at the SoH value of 70. Because of this, a model trained on this data would never observe the possibility that the degradation path can drop below 70. Never observing the possibility that the degradation path can drop below an SoH value of 70 can be problematic for models not well-suited for extrapolation. Ensuring that the model learns the possibility that the degradation path can drop below 70 is done by

extending the degradation path down to 65. A function developed by Volvo Cars was employed to fit a curve to the original data. This enabled the generation of data points down to 65.

Each simulated customer has a degradation path containing information about the battery's age when each SoH value is reached, resulting in one row of information for each SoH value observed. However, no explicit feature describes the differences between the placement of the points on the degradation path. The points' placement can be in terms of time and usage, such as an odometer value. This is important as the model needs examples in the training data that contain explicit information about the interval to future target point to facilitate the training phase.

A pre-processing step is done to compute the differences in placement between a point and all its subsequent points on the degradation path. This process is shown in Figure 4.1 with an example of four points.

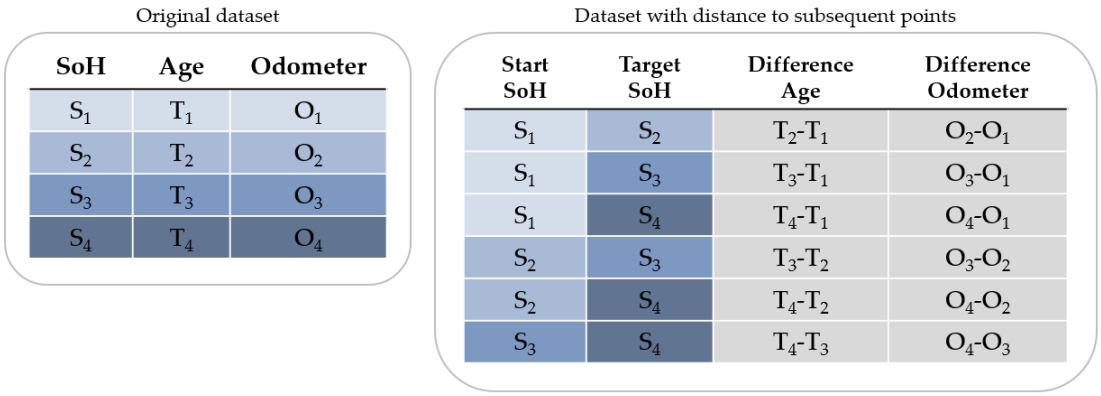



**Figure 4.1:** An example of four points illustrating the pre-processing step to compute the differences in placement to subsequent points.

The pre-processing step results in each pair of start and target points being one row of information, increasing the number of examples in the training data. The row of information contains the differences between the start and target points in age and odometer values.

The complete set of features for the input vector to the model is described in Table 4.2. All the original features describe the current state of the battery. However, the model also needs information about the difference in points placement to the target points on the degradation path. For this reason, the differences in age, measured in days, and odometer, measured in kilometers, are added as additional features to provide this information.

**Table 4.2:** Features in lab data.

| Type | Signal |
|---|---|
| Original | SoH |
| Original | Fast Charge |
| Original | Odometer |
| Original | Temperature |
| Original | State of Charge |
| Original | Age |
| Created | Days to Target |
| Created | Odometer to Target |

The final pre-processing step is splitting the data into subsets for use during the training and testing phase. The splitting process for a dataset differs depending on whether its part of the pre-training or fine-tuning process. The pre-training process uses lab data A, and the fine-tuning process uses lab data B.

As the dataset for the pre-training, lab data A is randomly split on simulated customers into three subsets; training, validation, and test dataset. The process is seen in Figure 4.2.

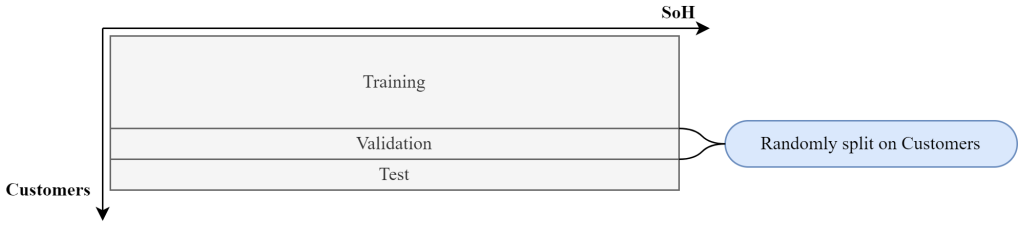



**Figure 4.2:** Splitting process for the dataset used in the pre-training phase.

The split of the dataset used for the fine-tuning processes requires defining two SoH limits, *SoH fine-tune amount* and *SoH starting point*. *SoH fine-tune amount* determines the length of the degradation path utilized during the training phase of the fine-tuning process. *SoH starting point* specifies the point on the degradation path from which EoL predictions are made in the testing phase.

Lab data B, serving as the dataset for fine-tuning, is first split randomly on simulated customers into two subsets, one for the training and testing phases. The subset for the training phase is systematically split on SoH according to the value of the *SoH fine-tune amount*. The subset for the testing phase is systematically split on SoH according to the value of the *SoH starting point*.

The values of *SoH fine-tune amount* and *SoH starting point* varies depending on the experiment settings, while the split on the customers is the same across the experiments. This process is seen in Figure 4.3.

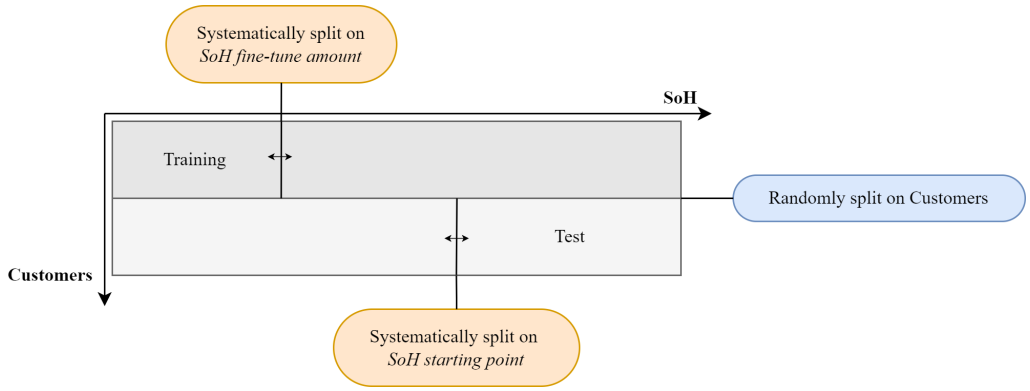



**Figure 4.3:** Splitting process for the dataset used in the fine-tuning phase.

### 4.1.2 Field data

To ensure the quality and suitability of the field data for a data-driven model, a series of cleaning and pre-processing steps are undertaken (see Table 4.3).

**Table 4.3:** Pre-processing steps for field data.

| Step | Description |
|---|---|
| 1 | Remove outliers outside thresholds |
| 2 | Remove cars with only a single SoH measurement |
| 3 | Remove cars with increasing SoH trend |
| 4 | Calculate km/year for each car |
| 5 | Fill missing odometer values |
| 6 | Histogram feature aggregation |
| 7 | Generate training data |
| 8 | Split data for pre-training |
| 9 | Split data for fine-tuning |

To identify outliers in the dataset, two thresholds are established. The first threshold is based on the observations from a calendar aging test, representing the slowest observed degradation scenario. SoH values above this threshold are considered implausible. The second threshold represents a simulated customer with an extremely fast degradation. SoH values below this threshold are regarded as unlikely. Additionally, cars exhibiting an increasing trend in battery health are removed from the dataset, as an increasing SoH contradicts the natural degradation pattern since batteries cannot regenerate SoH.

Missing or incomplete odometer readings are addressed by fitting a straight line from origo to its current odometer and age data. This fitted line has a slope value that represents the car's annual mileage and the value allows for filling the gaps. The process of filling the odometer feature is shown in Figure 4.4.

Original dataset

| SoH | Age | Odometer |
|---|---|---|
| $S_1$ | $T_1$ | NaN |
| $S_2$ | $T_2$ | $O_2$ |
| $S_3$ | $T_3$ | NaN |
| $S_4$ | $T_4$ | $O_4$ |

Dataset with filled odometer

| SoH | Age | Odometer | km/year |
|---|---|---|---|
| $S_1$ | $T_1$ | $K*T_1$ | K |
| $S_2$ | $T_2$ | $O_2$ | K |
| $S_3$ | $T_3$ | $K*T_3$ | K |
| $S_4$ | $T_4$ | $O_4$ | K |

**Figure 4.4:** An example showing how the missing odometer values are filled.

Each histogram feature is aggregated to a weighted average value, transforming the histogram into a scalar representation. For instance, for the histogram temperature feature, the aggregated scalar value represents the average temperature experienced by the car.

The data preprocessing steps are implemented to address data inconsistencies, eliminate outliers, and fill in missing values, ensuring that the field data reflects battery aging caused by degradation mechanisms rather than sensor fault or estimation errors.

## 4.2 Transfer learning framework

The transfer learning framework is applied to improve EoL prediction by harnessing insights and data from similar or related battery systems. The implementation consists of two main steps followed by each other: *pre-training* on the dataset with entire degradation paths and then *fine-tuning* the pre-trained model on a dataset with only early degradation paths. A schematic diagram of the transfer learning framework for EoL prediction is shown in Figure 4.5.

For the pre-training phase, the source dataset utilized contains the entire degradation for different customers. An ANN, further detailed in Section 4.3, is trained using this dataset. The model gets initialized with random weights, which are then adjusted in the training phase. A random grid search, described in Section 4.3.3, is used to find the model and algorithmic hyperparameters that yields the lowest RMSE score. The pre-training process enables the model to capture general degradation patterns and characteristic features specific to the source dataset.

Following the pre-training phase comes the fine-tuning step, where the pre-trained model is adapted to a target dataset comprising early degradation path for different customers. During fine-tuning, all model hyperparameters from the pre-trained model are reused except for the hyperparameter controlling which layers are frozen.

By selectively freezing specific layers and leaving others unfrozen, it enables the fine-tuned model to retain its ability to capture the essential degradation patterns learned during pre-training while still being able to adjust its representation to align with the degradation characteristics specific to the target dataset.

These two steps allow for transferring knowledge between two data sources from different domains. The transfer could, for example, be performed between lab and field or between batteries with two different cell chemistries. This transfer learning framework addresses the challenge of a battery target domain where there are only early degradation paths and no EoL labels by utilizing a source domain where full degradation paths to EoL exist.

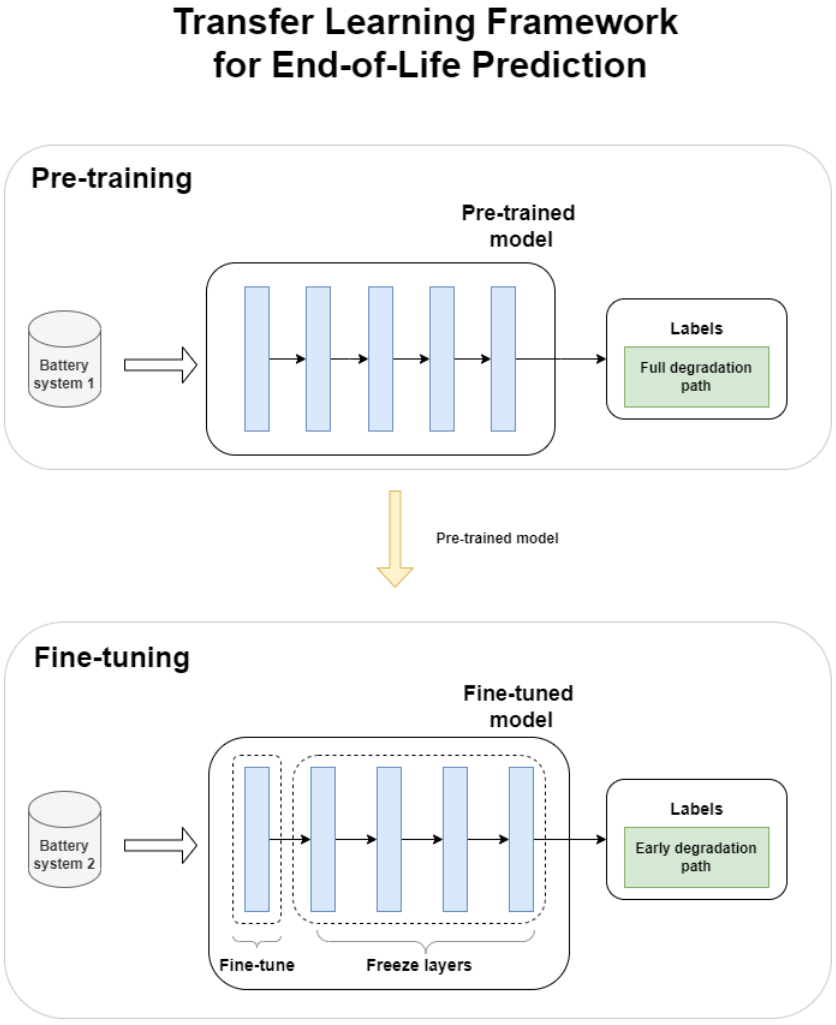



**Figure 4.5:** Schematic diagram illustrating the pre-training and fine-tuning process for EoL prediction of batteries. In the pre-training phase an ANN model is trained on a dataset with full degradation paths. In the fine-tuning stage, the entire network from the pre-trained model is used on a dataset with early degradation paths, where all layers except the first are frozen.

## 4.3 Implementation details

The model used in the transfer learning framework is an ANN consisting of fully connected layers. The input layer of the model has a size of eight that receives the input vector. The output layer of the model has a size of one, representing the predicted future SoH Value. The choice of specific properties of the hidden layers, such as the number of layers and neurons per layer, is described in Section 4.3.3 about hyperparameter selection. A visualization of the model layout is shown in Figure 4.6.

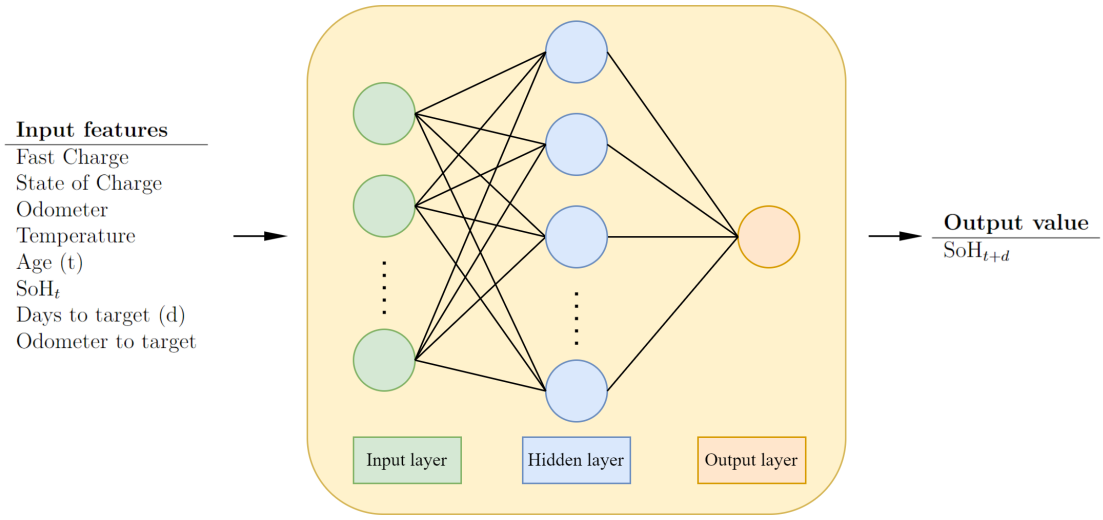



**Figure 4.6:** Layout of how the model converts the input vector into output value. $\mathbf{SoH_t}$ is the current SoH at age t and $\mathbf{SoH_{t+d}}$ is the predicted SoH at d days after t.

### 4.3.1 Prediction methodology

There are two prediction methods for the model, depending on whether the objective is to predict the EoL or, in the case of no labels, to predict a future SoH value.

The EoL value is measured as the age of the battery when it reaches the SoH threshold of 70, so predictions of EoL must be in the same age unit. Since the model's current output is an SoH value, the prediction method must be adjusted to facilitate predictions in an age unit. The variables required for the algorithm are described in Table 4.4.

The algorithm iterates over a loop to predict SoH over future points and returns an output of the age where SoH is predicted to be below 70%, which is described in detail below.

**Table 4.4:** Variables required for the EoL prediction algorithm.

| Variable | Description |
|---|---|
| t | Age at the starting point |
| $d'$ | Days to target |
| $o'$ | Odometer to target |
| $\Delta$ | Size of timestep increment |
| $\mathbf{x}$ | The input vector with the format: $\mathbf{x} = [x_1, x_2, x_3, x_4, x_5, x_6, d', o']$ where the input features of $x_1..x_6$ remains fixed in the algorithm. |

**Algorithm to generate EoL prediction**

$\hat{y} \leftarrow 100$ (arbitrary starting point $\geq 70$)
$d' \leftarrow 0$
**while** $\hat{y} \geq 70$ **do**
  $d' \leftarrow d' + \Delta$
  $o' = g(\mathbf{x}, d')$ ▷ Odometer to target calculated from new value of $d'$
  $\mathbf{x} = [x_1, x_2, x_3, x_4, x_5, x_6, d', o']$ ▷ Input vector is assigned values of $d'$ and $o'$
  $\hat{y} = f(\mathbf{x})$ ▷ Calculate new EoL prediction for input vector
  **if** $\hat{y} < 70$ **then**
    **Return:** t + d' ▷ Return the date predicted that EoL will be reached
  **end if**
**end while**

The function $g(x, d')$ estimates the odometer difference to the next point based on its current input vector and the days to target.

### 4.3.2 Model evaluation

A key objective in the aging of EVs domain is to estimate when the EoL is reached. The model's performance metric aims to assess how effectively the model achieves the objective. For evaluation, two types of errors can be computed, an EoL prediction error, $\epsilon_{\text{EoL}}$, and an SoH prediction error, $\epsilon_{\text{t+d}}$, defined by Equations 4.1 and 4.2

$$\epsilon_{\text{EoL}} = y_{\text{EoL}} - \hat{y}_{\text{EoL}} \tag{4.1}$$

$$\epsilon_{\text{t+d}} = y_{\text{t+d}} - \hat{y}_{\text{t+d}} \tag{4.2}$$

where $\epsilon_{\text{EoL}}$ refers to the EoL error in an age unit and $\epsilon_{\text{t+d}}$ refers to SoH error when predicting a point d days into the future from age t. Figure 4.7 shows the difference between these errors.

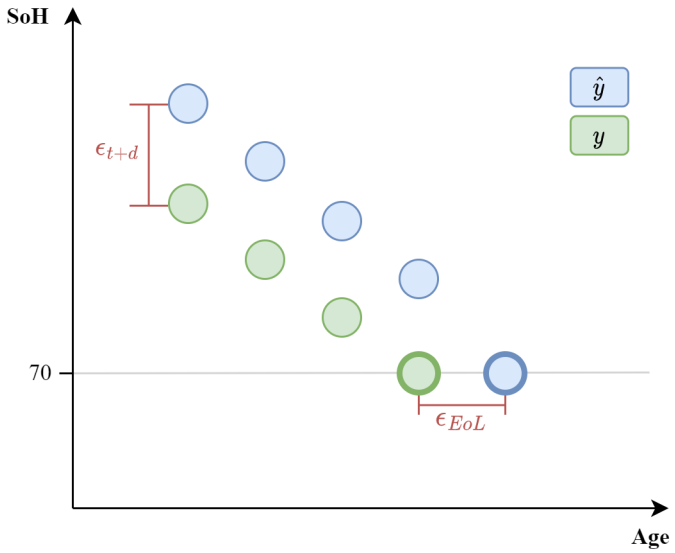



**Figure 4.7:** The model evaluation errors $\epsilon_{EoL}$ and $\epsilon_{t+d}$ are shown for an example degradation path.

In the case of lab data, the label $y_{EoL}$ is known. By applying the algorithm to generate EoL prediction described in Section 4.3.1, $\hat{y}_{EoL}$ can be computed, allowing for the calculation of the error $\epsilon_{EoL}$. This error measures the ability to predict EoL, which is the primary objective, and therefore it is used to evaluate the model's performance using the metrics RMSE and MAPE.

However, for field data, where $y_{EoL}$ is unknown, $\epsilon_{\text{EoL}}$ cannot be computed. Since the field data still has a few observations of SoH per customer, the prediction method of estimating $\epsilon_{\text{t+d}}$ for points further down the degradation path is used. While this evaluation does not directly measure the primary objective of predicting EoL, it serves as a valuable sanity check. Poorly predicting the SoH of future points, measured as $\epsilon_{\text{t+d}}$, can indicate that the model has problems predicting EoL, as the calculation is based upon future SoH predictions.

### 4.3.3 Hyperparameter selection

In order to effectively train models like ANN, it is important to optimize the *hyperparameters*, which are parameters configured prior to the learning process. These hyperparameters can influence the behavior and performance of the model. The hyperparameters used are detailed in Table 4.5.

The model hyperparameters used for the ANN are the number and size of layers within the network architecture, the dropout probability for the dropout layers, and the selection of the layers that should be frozen during training. Additionally, the algorithmic hyperparameters utilized are the batch size, the learning rate of the optimization algorithm, and the number of epochs.

**Table 4.5:** Model and algorithmic hyperparameters used in models.

| Type | Hyperparameter |
|---|---|
| Model | Number of layers |
| Model | Number of hidden neurons |
| Model | Dropout probability |
| Model | Type of layer frozen |
| Algorithmic | Learning rate |
| Algorithmic | Batch size |
| Algorithmic | Number of epochs |

The selection and combination of these hyperparameters significantly impact the model's overall performance. A hyperparameter random grid search is conducted on a pre-defined set of hyperparameter values to identify the optimal configuration. The models generated from different hyperparameter combinations are evaluated based on their RMSE scores on the validation set, and the model that achieves the lowest RMSE score is selected as the best-performing model.

## 4.4 Experiments

Three experiments that answers the research questions are described in this section.

### 4.4.1 The impact of fine-tuning on performance

This experiment evaluates how the amount of data provided during the fine-tuning phase affects the predictive performance for EoL estimation. To facilitate this, a pre-trained model is trained on lab data A and acts as the benchmark model. Several fine-tuned models are built based on the pre-trained model by fine-tuning on lab data B according to Section 4.2 for different amounts of *SoH fine-tune amount*.

The experiments utilize the *SoH fine-tune amount* values of {94, 90, 85, 80, 75, 70}. A visualization of how the value of *SoH fine-tune amount* affects the available data during the training phase is shown with three examples values in Figure 4.8

For the *SoH starting point*, the values selected are {94,92,90,88,86,84,82,80,78,75}. The selections omit values close to the curve's start or end and take every second value within the interval. How predictions in the testing phase are impacted by the *SoH starting point* are shown with three examples of starting points in Figure 4.9.

The experiment is performed in the following steps: First, the *SoH fine-tune amount* is selected, and the training subset of lab data B is systematically split based on this value. Next, the pre-trained model is fine-tuned based on this data subset. The resulting fine-tuned model generates predictions for each different *SoH starting point* in the test set and corresponding $\epsilon_{EoL}$ is computed. Finally, the EoL predictions are evaluated using the RMSE and MAPE performance metrics.

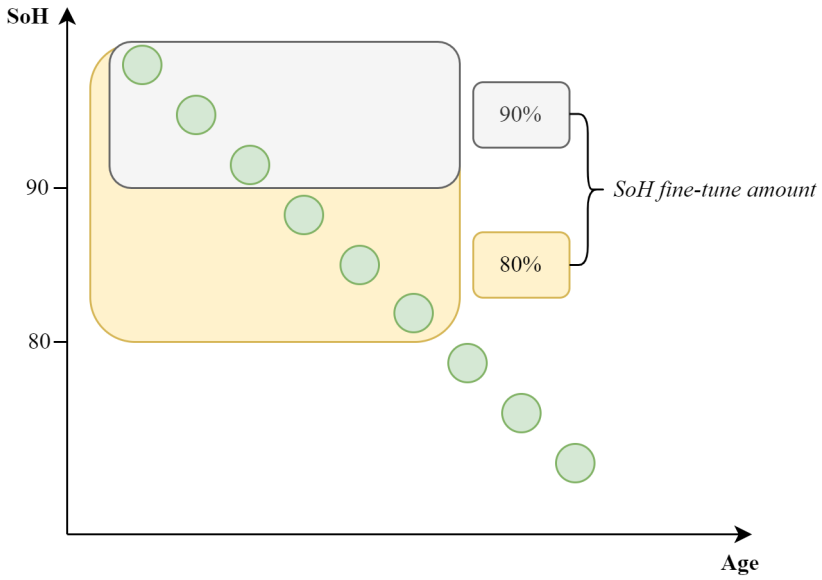



**Figure 4.8:** Example of the impact *SoH fine-tune amount* has on the available data of degradation paths for the training phase. A lower value on the *SoH fine-tune amount* indicates that a larger portion of the degradation path is included in the training phase, increasing the amount of data for the model to train on.

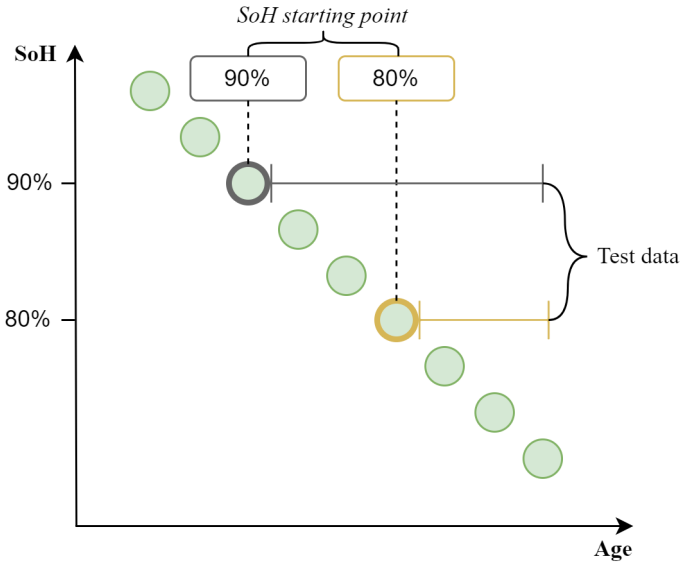



**Figure 4.9:** Example of the impact *SoH Starting point* has on the size of the degradation paths for the test data. A lower value on the *SoH Starting point* indicates that a smaller portion of the degradation path is unknown to the model during the testing phase.

### 4.4.2 Patterns of knowledge transfers between domains

This experiment aims to investigate the impact of fine-tuning further, focusing on analyzing if there are any patterns in the knowledge transfer between lab data A and lab data B. Patterns in this context refer to whether a model that is fine-tuned on lab data B makes predictions that form a trend of either over or underestimates of the ground truth.

Patterns are explored by comparing the predicted degradation paths of the fine-tuned and pre-trained model against two known degradation paths, the actual ground truth and a reference path. The ground truth path represents the actual degradation path for a simulated customer in lab data B. As described in Section 3.2, an identical configuration of driving behavior exists for simulated customers in lab data A and B. The reference path refers to the degradation path of the customer with identical driving behavior from lab data A.

A few specific simulated customers are selected to facilitate the visualizations of the degradation paths. To determine which simulated customer to choose, a selection criterion is established to enable prioritization of the customers. The selection criterion, $\delta$, is described in equation 4.3

$$\delta = \frac{FT - PT}{PT} \tag{4.3}$$

where the $FT$ is the fine-tuned and $PT$ is the pre-trained model's EoL predictions. Extreme values of $|\delta|$ indicate a significant relative difference in the prediction of the fine-tuned and pre-trained models, which is an interesting starting point to explore potential patterns.

Two types of patterns are examined in cases where $|\delta|$ is large. Firstly, an analysis is done on how the degradation paths of the fine-tuned and pre-trained predictions are positioned with respect to the actual ground truth path. Secondly, analysis is performed on how the fine-tuned and pre-trained predictions are positioned with respect to the reference path. A detailed list of experiment steps is described in Table 4.6.

**Table 4.6:** List of steps for selecting examples to visualize.

| Step | Description |
|---|---|
| 1 | Remove customers with unusually high or low values of EoL |
| 2 | Compute the selection criterion $\delta$ with Equation 4.3 |
| 3 | Extract the customers with the three highest and lowest values of $\delta$ |
| 4 | Extract the ground truth and reference degradation paths |

### 4.4.3 Fine-tuning on a fleet of cars

The transfer learning framework for EoL prediction on the fleet of cars is implemented by first pre-training on lab data B. The reason for choosing this dataset for

the pre-trained model is that many of the batteries use this cell chemistry in the fleet of cars. The pre-trained model is then fine-tuned on the field data to learn the specific degradation of the fleet of cars. The hyperparameters used for training are described in Appendix A.2. To predict EoL for a real customer in the fleet, the algorithm described in Section 4.3 is used with the customers' last measurements as input to the model.

Since the field data is incomplete and does not consist of entire degradation paths down to EoL it is not possible to evaluate the accuracy of EoL predictions on field data using the standard metric RMSE or MAPE. Nonetheless, other strategies can be employed to assess the performance and effectiveness of the transfer learning approach on the field data. Two analyses are performed to evaluate the performance: a *visual analysis* of EoL predictions for individual customers, and a *comparative analysis* on the available SOH data in the field.

To evaluate the effectiveness of transfer learning and the performance of the fine-tune model, a comparative analysis is done between the pre-trained and fine-tuned models using the available field data. This comparison assesses the relative improvement achieved through transfer learning. For the comparative analysis, both models are evaluated using the standard performance metric RMSE on test data that have not been used for training the model. The data split is performed as shown in Figure 4.2 where 80% of the customers is used for training and 20% is used for evaluation.

The visual analysis is performed to assess the behavior of the fine-tuned model under varying scenarios and data conditions. The model's validity is explored by visualizing different customers' predicted SoH values until EoL. This analysis could indicate whether the model has learned how different driving behavior affects battery degradation.

# 5
# Results

In this chapter, the results of the three experiments are presented. In Section 5.1, the results of different fine-tuning amount on performance is presented. In Section 5.2, results from the patterns of knowledge transfer experiment are provided. Lastly, in Section 5.2, the findings from applying the framework to the fleet of cars are unveiled.

## 5.1 The impact of fine-tuning amount on performance

The result of the predictive performance for different values of *SoH fine-tune amount* is shown in Figure 5.1, where the calculations of $\epsilon_{EoL}$ is from predictions made at a *SoH starting point* of 94. The models' performance for the first set of values of *SoH fine-tune amount* remains close to the pre-trained model. However, when the model is fine-tuned on a sufficient amount of data, the model improves significantly in the prediction performance.

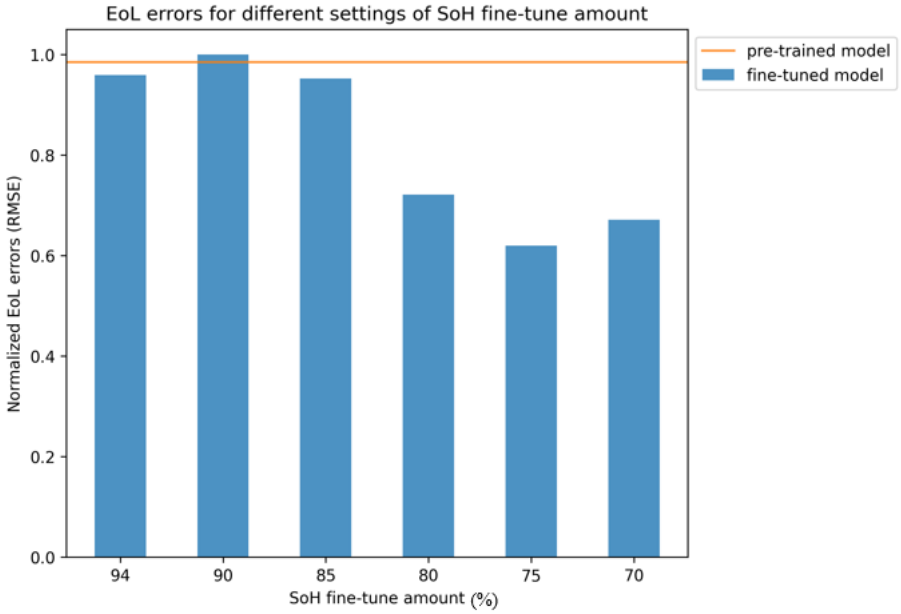



**Figure 5.1:** EoL errors for different settings of *SoH fine-tune amount* with an *SoH starting point* of 94.

The result for the predictive performance across different settings of *SoH starting point* is shown in Figure 5.2. Each model's accuracy for all *SoH starting point* is shown in Figure 5.3, and the best accuracy for each *SoH starting point* amongst all models is selected as the best fine-tuned model. The predictive performance improves for the pre-trained and the best fine-tuned model as the *SoH starting point* is set further down the degradation path. The error ratio between the best fine-tuned and pre-trained models falls within a range of 60-75%.

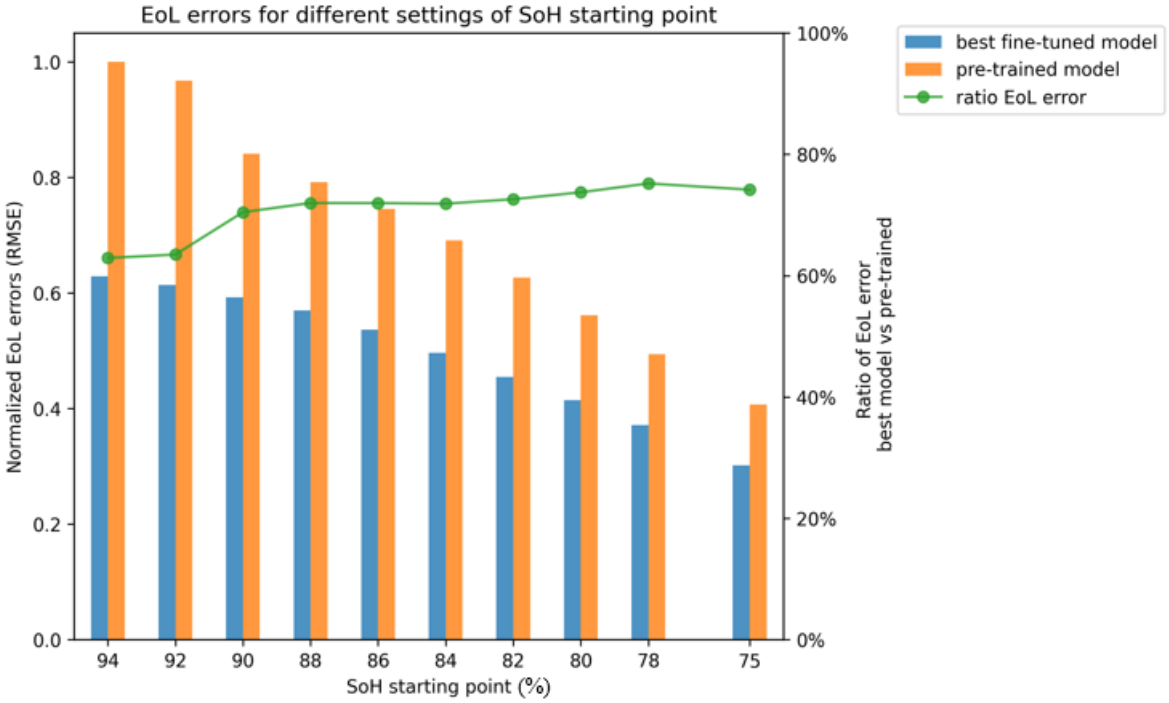



**Figure 5.2:** EoL errors for different settings of *SoH starting point* for the pre-trained model and the best fine-tuned model at each starting point. The error ratio between these models is shown as a line chart.

Results from all combinations of the *SoH fine-tune amount* and *SoH starting point* settings is visualized with RMSE error in Figure 5.3 and MAPE error in Figure 5.4. The results indicate a similar trend when measuring the errors in RMSE and MAPE. There is not one single fine-tuned model that stands out significantly superior to all others. Instead, different models, all having a *SoH fine-tune amount* between 70 and 80%, achieve the best results for various starting points.

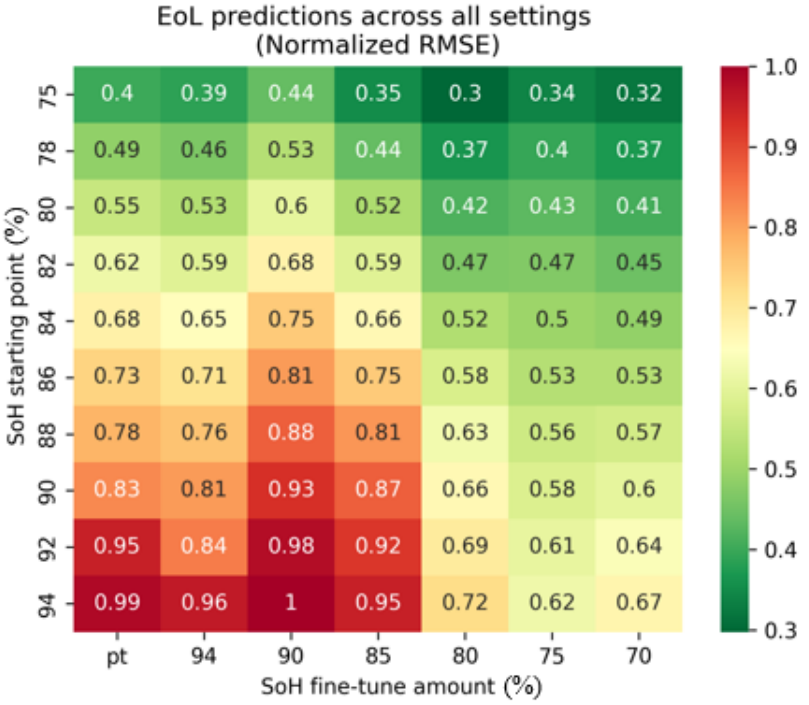



**Figure 5.3:** Heatmap of normalized RMSE scores for EoL prediction errors across all settings of *SoH fine-tune amount* and *SoH starting point*.

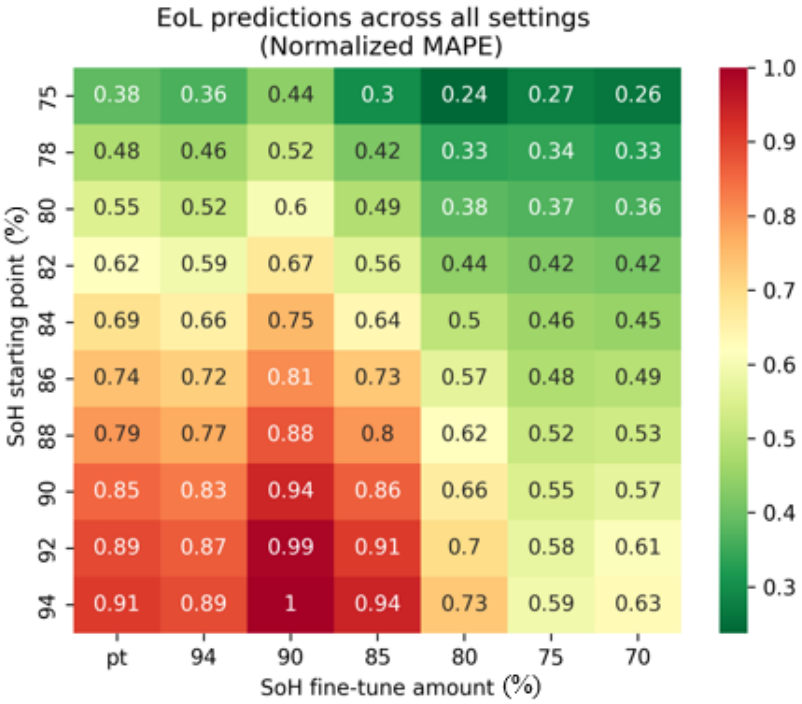



**Figure 5.4:** Heatmap of normalized MAPE scores for EoL prediction errors across all settings of *SoH fine-tune amount* and *SoH starting point*.

The EoL predictions at the *SoH starting point* of 94 from all the models are displayed in Figure 5.5. The result indicates that the models trained on larger amounts of fine-tuning data predict higher median EoL predictions. While the median EoL prediction is higher for these models, they seem less likely to make extreme EoL predictions, as there are fewer predictions above the normalized value of 0.8.

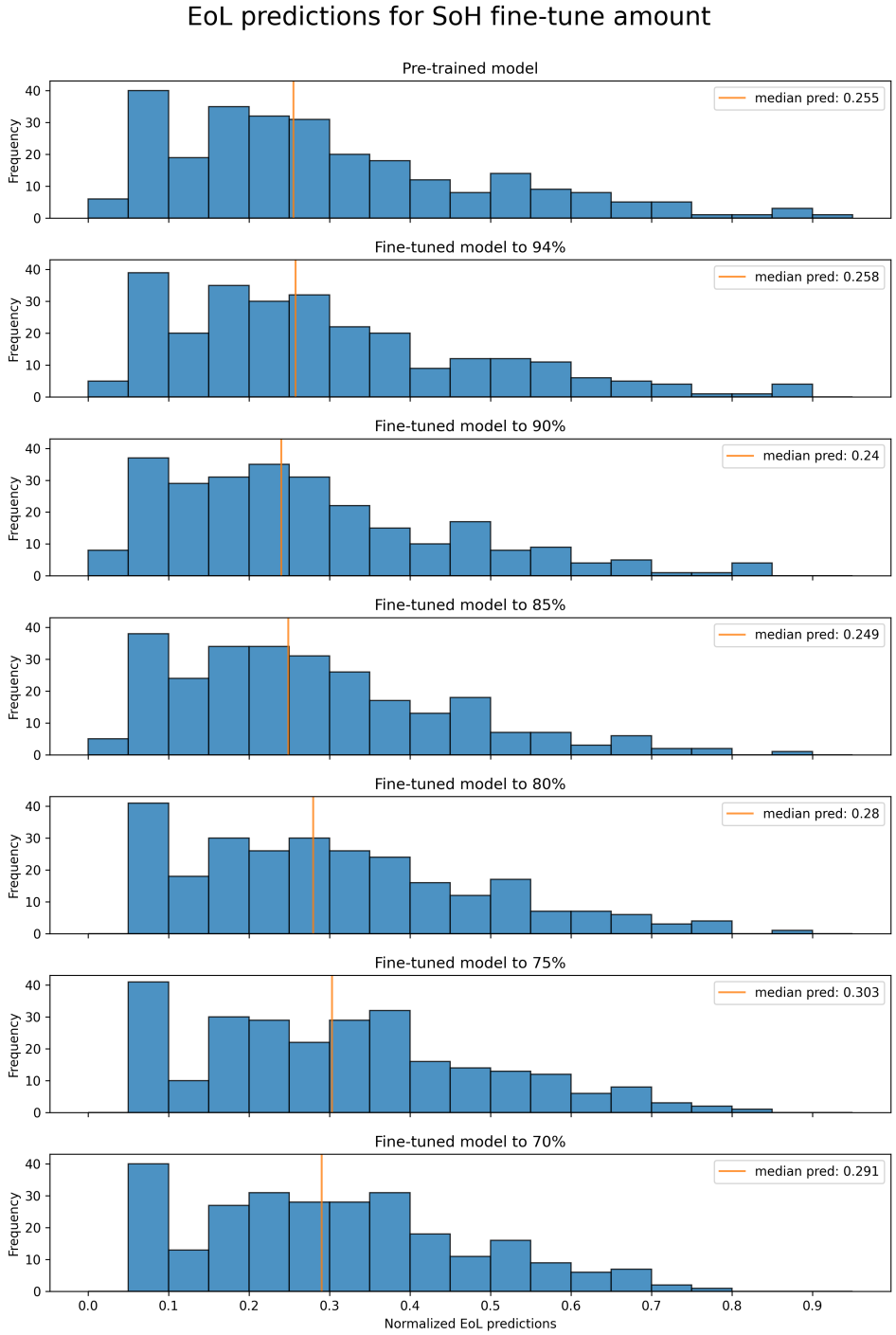



**Figure 5.5:** Histogram of EoL predictions for different settings of *SoH fine-tune amount* with an *SoH starting point* of 94%.

The $\epsilon_{EoL}$ from *SoH starting point* 94% across all the models are displayed in Figure 5.6. For all of the models, there are more underestimations, meaning that the model estimates an EoL lower than the ground-truth value, than overestimations. The amount of overestimations decreases as the model is fine-tuned on more data. The error distribution is also narrower for models with high amount of fine-tuning.

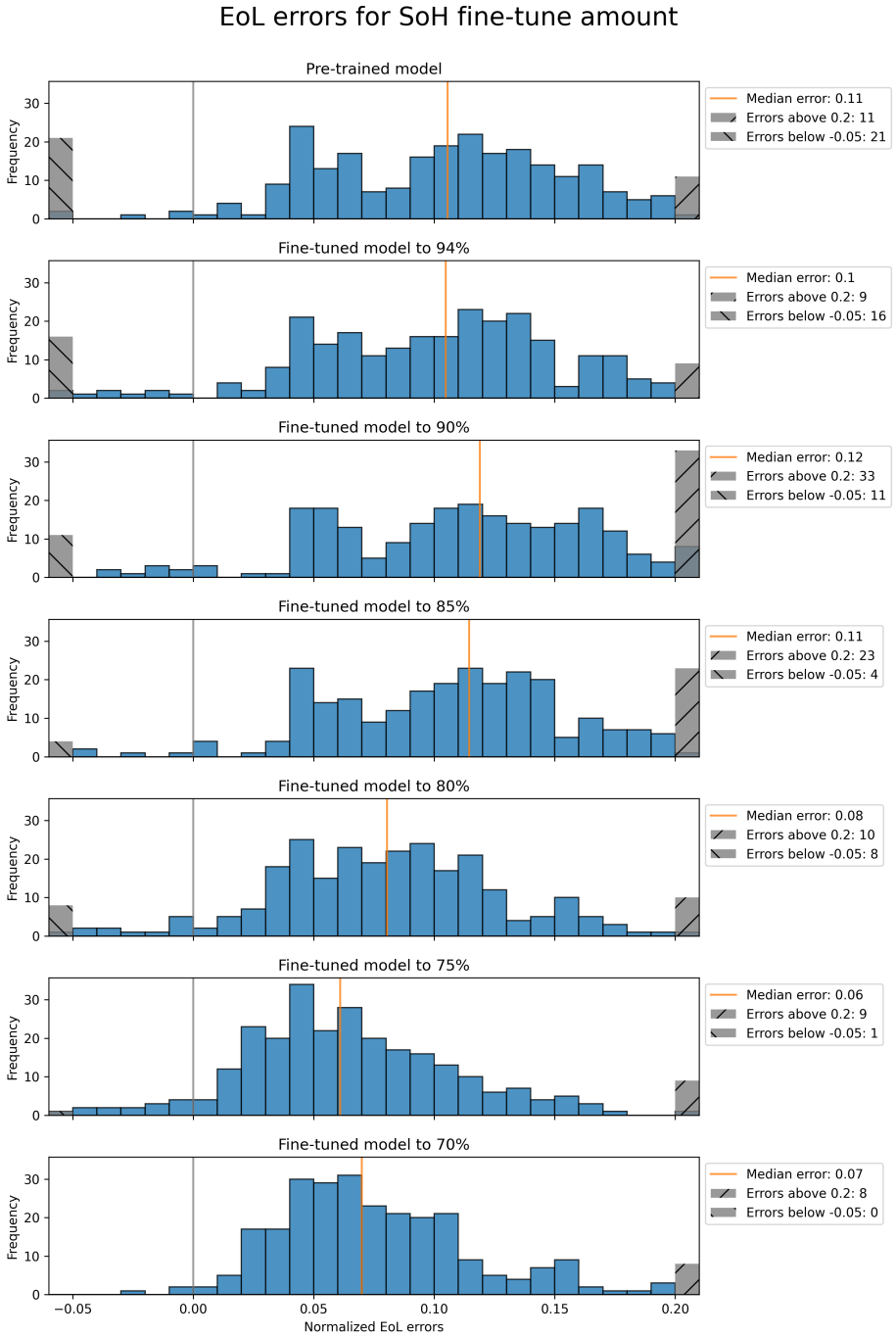



**Figure 5.6:** Histogram of EoL errors for different settings of *SoH fine-tune amount* with a fixed value of 94% on *SoH starting point*. The amount of errors below -0.05 and above 0.2 have been aggregated into one single bar each to improve visualization.

## 5.2 Patterns of knowledge transfer between domains

The three degradation paths belonging to the simulated customers with the highest $\delta$ are shown in Figure 5.7. The visualization contains predicted degradation paths from the pre-trained model and a fine-tuned model with *SoH fine-tune amount* of 80%, as well as the ground truth degradation paths from the simulated customer in lab data B and a reference degradation path from lab data A. The ground truth and reference degradation paths being different is a result of the different cell chemistry in lab data A compared to lab data B. The *SoH starting point* is set at 94%, and the entire ground truth degradation paths are displayed. The result shows that the fine-tuned model's predictions are closer to the ground truth prediction path compared to the predictions from the pre-trained model. Furthermore, the degradation path of the fine-tuned model is never below the pre-trained model's, even in the earlier phases of the degradation.

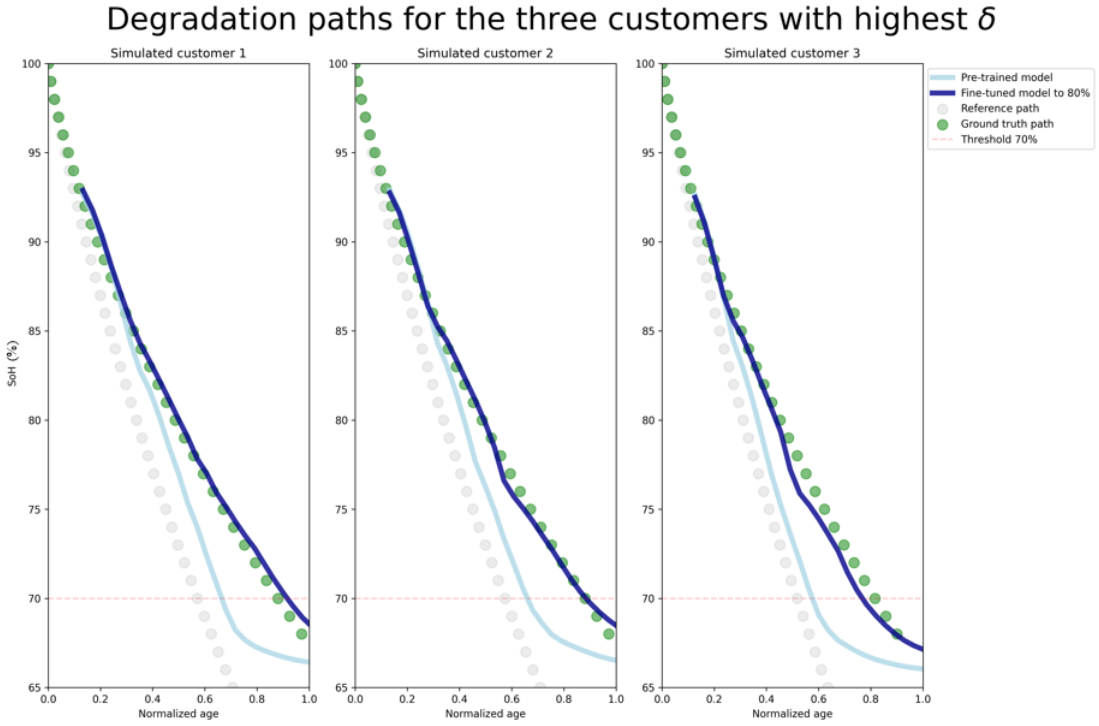



**Figure 5.7:** Degradation paths for the three customers with highest $\delta$ with model predictions from *SoH starting point* at 94%.

The degradation paths of the three simulated customers with the largest negative value of $\delta$ are shown in Figure 5.8. The fine-tuned model is significantly closer to the ground truth model for the two first customers, but for the third customer they are at a similar distance. In this case, the fine-tuned model is never above the pre-trained for any part of the degradation path.

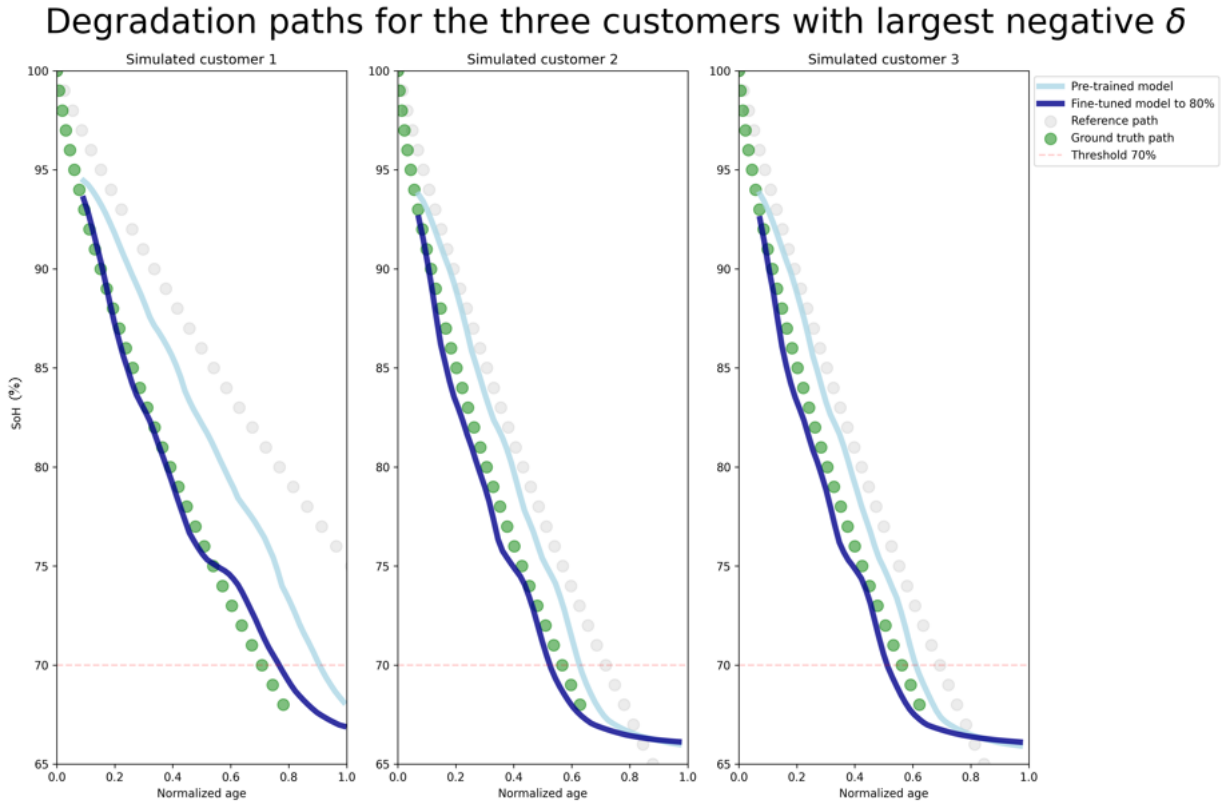



**Figure 5.8:** Degradation paths for the three customers with largest negative $\delta$ with model predictions from *SoH starting point* at 94%.

## 5.3 Fine-tuning on a fleet of cars

The comparative analysis on the test data for the pre-trained and fine-tuned model is shown in Table 5.1. The fine-tuned model attains a better performance compared to the pre-trained model. This indicates that the fine-tuning process on the field data can improve predictive performance.

**Table 5.1:** Comparison of RMSE loss on normalized test data between pre-trained and fine-tuned model.

| **Model** | Pre-trained | Fine-tuned |
|---|---|---|
| **Test RMSE** | 0.0030 | 0.0010 |

### 5.3.1 Predictions for individual customers

The driving behavior of three customers from the fleet is shown in Table 5.2. Notice the high annual mileage and usage of fast charging for Customer 3 compared to the other customers. Customers 1 and 2 have approximately equal annual mileage, however, Customer 1 operates at a higher temperature.

**Table 5.2:** Driving behavior of three different customers from the fleet. The customers' true value are hidden. Customer 1's values are set to variables and the relative proportion is calculated for customer 2 and 3.

| Customer | km/year | Fast charging ratio (%) | Temp (°C) | SoC Level (%) |
|---|---|---|---|---|
| Customer 1 | $k_1$ | $f_1$ | $t_1$ | $s_1$ |
| Customer 2 | $0.89k_1$ | $0.1f_1$ | $0.26t_1$ | $1.5s_1$ |
| Customer 3 | $16.35k_1$ | $12f_1$ | $0.87t_1$ | $1.25s_1$ |

The predicted degradation paths to EoL for the three customers are depicted in Figure 5.9. The predicted degradation paths for customers 1, 2 and 3 fall within the moderate, slow, and fast regions, respectively. Customer 3, who has significantly higher annual mileage and employs frequent fast charging compared to the other customers, is expected to experience more rapid battery degradation. On the other hand, customers 1 and 2, with similar annual mileage, have their predicted EoL fall into different regions. This significant difference in EoL prediction is correlated to the differences in fast charging frequency and average car temperature.
Customer 1 operates at a higher average temperature, likely contributing to its accelerated degradation.

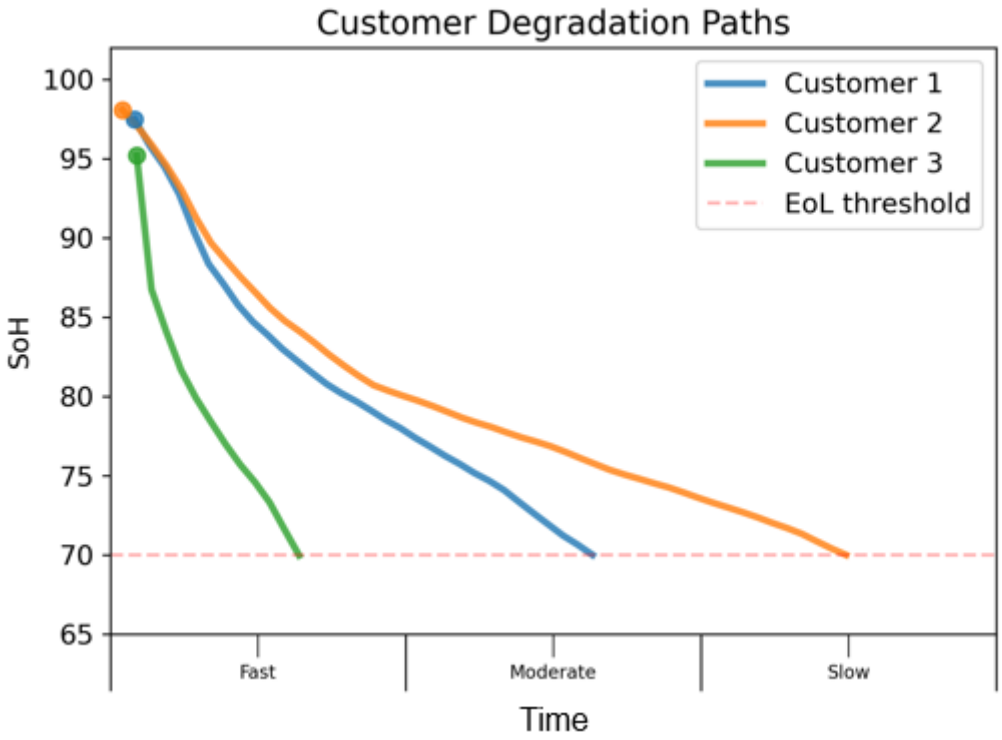



**Figure 5.9:** Predicted degradation path to EoL for three different customers from the fleet using the fine-tuned model. The battery age is divided into three sections: fast, moderate and slow as displayed on the x-axis. The region where a customer's predicted degradation path ends at the EoL classifies the rate at which the battery's health has degraded over its lifespan.

The driving behavior of two customers with relatively high annual mileage is shown in Table 5.3. Customer 5 drives one-third more kilometers per year than Customer 4, although it uses around a quarter as much fast charging and is located in a lower average temperature.

**Table 5.3:** Driving behavior of two different customers. Customer 4's values are set to variables and the relative proportion is calculated for Customer 5.

| Customer | km/year | Fast charging ratio (%) | Temp (°C) | SoC Level (%) |
|---|---|---|---|---|
| Customer 4 | $k_4$ | $f_4$ | $t_4$ | $s_4$ |
| Customer 5 | $1.32k_4$ | $0.19f_4$ | $0.65t_4$ | $s_4$ |

The predicted degradation paths for the two customers are illustrated in Figure 5.10. Initially, both customers exhibit a similar degradation pattern in their predicted paths until about an SoH of 85%. However, beyond this point, Customer 4's predicted degradation path has a faster rate of degradation compared to Customer 5. This is likely due to the significantly higher frequency of fast charging and the operation of the vehicle at a higher average temperature.

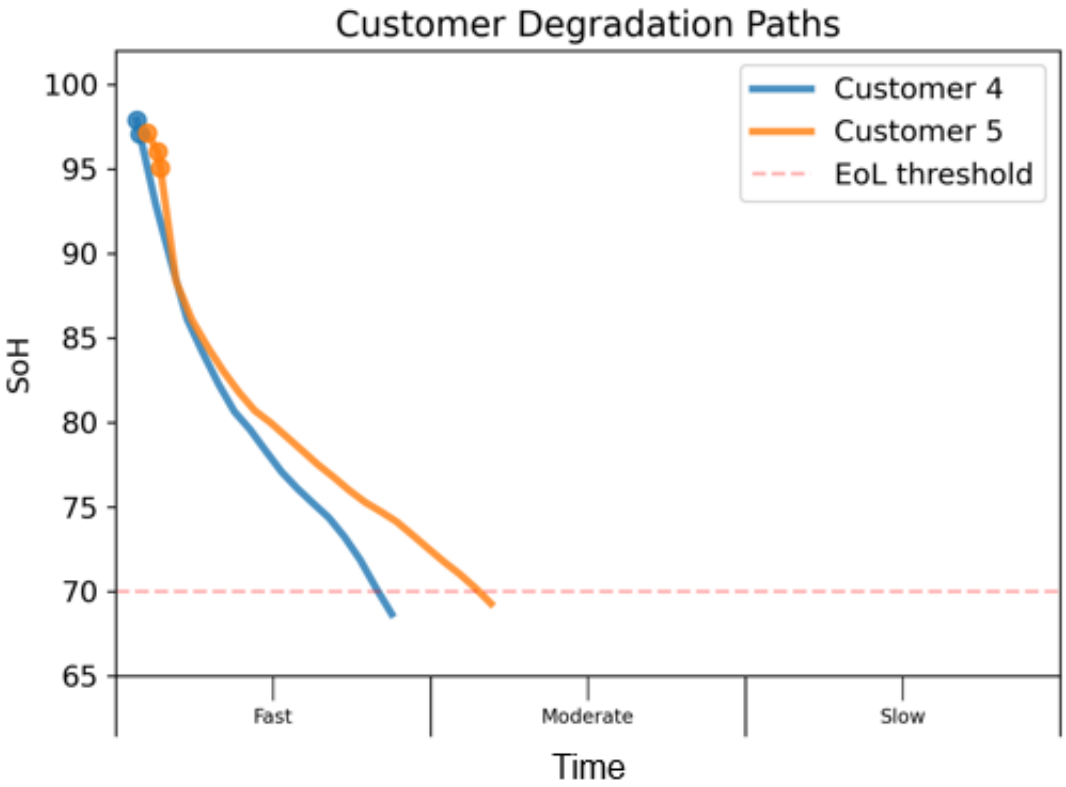



**Figure 5.10:** Predicted degradation path to EoL for two different customers using the fine-tuned model.

### 5.3.2 Prediction for the fleet of cars

The EoL predictions for the fleet of cars are displayed in Figure 5.11. The distribution of EoL predictions exhibits a bell-shaped curve where most customers have an EoL in the moderate region. Additionally, a noticeable number of customers are predicted to fall within the fast degradation region, while very few are predicted to have a slow degradation.

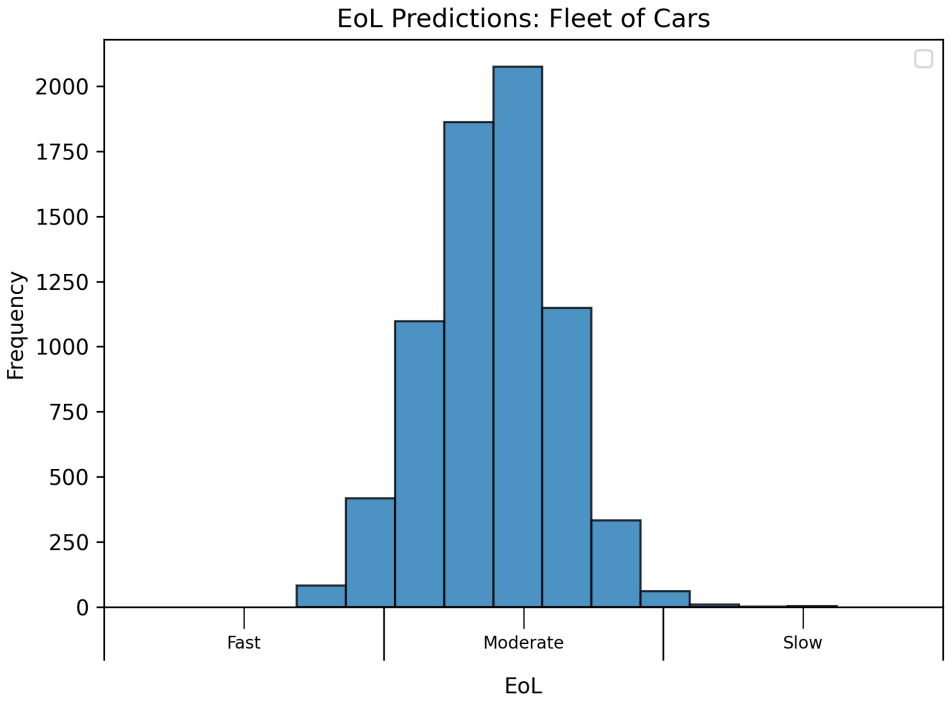



**Figure 5.11:** EoL predictions for the fleet of cars using the fine-tuned model.

# 6
# Discussion

In this chapter, the methodology and results are discussed. In Section 6.1, the limitations of the processed lab and field data are discussed. In Section 6.2 the results of the three experiments are discussed.

## 6.1 Discussion of data

The discussion first focuses on uncertainties in the data, followed by discussing preprocessing of histogram features.

### 6.1.1 Uncertainties in the data

In this domain of batteries for EVs, there are several sources of uncertainties from both the field and lab datasets. One source of uncertainty is that the laboratory collection methods in Section 3.1 for the raw lab data do not fully reflect the circumstances in which data from EVs in the fleet are collected. For example, cycle tests are obtained by performing charges and discharges of the battery at an accelerated rate compared to normal usage in an EV. Another difference is that since testing is performed on only a few cells, it does not fully replicate the setup in EVs with battery packs. In an EV, there might also be external factors such as vibration.

The processed lab data covered in Section 3.2 adds additional assumptions in order to obtain the simulated customers. Combining degradation information from calendar and cycle aging requires an assumption about potential path dependency between these two aging processes. Creating simulated customers requires assumptions about relevant driving parameters and the possibility of interpolating between these. For the field data presented in Section 3.3 there are uncertainties arising from the collection of sensor values in a vehicle on the road. As the calculation of SoH is based upon other sensor values, there could be discrepancies between the SoH in the field dataset and the actual condition of the battery in the vehicle.

These uncertainties lead to both the input and output vectors not fully reflecting the actual status of the battery's degradation. This is important to remember when interpreting the results; having access to estimations of these uncertainties could help build confidence intervals around the results.

### 6.1.2 Pre-processing histogram features

The data collected on temperature and SoC levels for customers in the fleet captures the amount of time spent by a car within a specific temperature or SoC window. However, by calculating the weighted average time spent in each bin, and using that as input to the model, valuable information of battery degradation may have been lost. As mentioned in Section 2.1.2, extreme temperatures can significantly impact battery degradation. Furthermore, the depth of discharge is an important factor to consider for cycling aging. Information about critical edge cases might not be kept after taking the aggregated mean of these histograms. To address this issue, a different fine-tuning approach could be employed by incorporating an adaptation layer that considers the time spent by the car within specific temperatures or SoC bins. This adjustment would make it possible for the model to learn the extreme temperatures and depth of discharge in real operating conditions, which could improve the accuracy of EoL predictions.

## 6.2 Discussion of results

The discussion of results is connected to the research questions. Section 6.2.1 discuss the question regarding fine-tuning amounts' impact on performance. Section 6.2.2 covers the results regarding knowledge transfer between domains. Finally, Section 6.2.3 presents and addresses the results for the predictions on the fleet of cars.

### 6.2.1 The impact of fine-tuning amount on performance

The potential of using transfer learning for EoL predictions depends on the value of *SoH fine-tune amount*, in other words, how much of the degradation path that is known in the target domain. For the extreme case where no part of the degradation path is known in the target domain, transfer learning is not possible since there is no data for the fine-tuning process. On the other hand, consider the opposite extreme case where the target domain contains a large set of customers with their full degradation path known. In this case, fully supervised learning directly on the target domain is a viable option instead of transfer learning. In the domain of battery aging in a fleet of EVs, neither of these extremes exist today, but rather the domain contains EVs where a part of the degradation path is known. Within this context, observing that fine-tuning on a partial degradation path can yield better predictive performance than the pre-trained model shows the potential of using transfer learning for EoL predictions.

The experiments conducted with different *SoH fine-tune amount* amounts reveal some interesting findings. It is observed that for fine-tuning amount at 94%, the improvement in RMSE score is relatively small. However, an unexpected observation is made when the SoH fine-tune amount is set to 90%; it results in a higher RMSE score than the pre-trained model. This increase in RMSE score seems to contradict the results obtained by a study also exploring different fine-tuning amounts when applying a transfer learning framework to predict EoL of batteries [17]. In the study,

the fine-tuning amounts $\{90, 80, 50, 30, 10\}$ were investigated, and the results from an average error of multiple experimental runs showed a monotonic decrease in error with more data to fine-tune on. That is, the RMSE score for fine-tuning amount at 90% could be due to inherent randomness in the training of the models. Training a model with a fine-tuning amount for multiple runs and taking the average RMSE loss of the runs would yield a more robust evaluation.

Another explanation of these results could be due to the fixed hyperparameters used in the different fine-tuning experiments. The decision to fix the hyperparameters was made to specifically evaluate the impact of the fine-tuning amount on performance while removing other factors. However, fixing the hyperparameters carries the risk of overfitting when dealing with a small amount of data being fine-tuned for too many epochs or with too-large batch size.

### 6.2.2 Patterns of knowledge transfer between domains

Results showing that predictions from a fine-tuned model are closer to the ground truth compared to the pre-trained model indicates that some knowledge about degradation characteristics in the target domain is transferred during the fine-tuning process.

An interesting observation is that the fine-tuned model's prediction is below both the pre-trained and ground truth in one of the examples. A potential factor to explain this is how the dataset is split in the fine-tuning phase. The dataset is first split randomly on customers, meaning that when the model makes EoL predictions for customers in the test data, it has never seen any part of its degradation path. This requires the model to learn degradation characteristics in the target domain from other customers instead. Depending on the similarity of the degradation for customers with comparable conditions, the knowledge transferred during the fine-tuning process might not be optimal in some scenarios. Exploring the fine-tuning process with a more extensive set of data and a broader set of operating conditions could be interesting to investigate.

The results show that the model is consistently underestimating the predicted EoL. There are many potential reasons why the model learns patterns that lead to systematic underestimations. As seen in Figure 3.2, the source domain has shorter EoL values than the target domain. With the fine-tuning process keeping most of the layers in the model frozen, many of the weights in the fine-tuned model are solely determined by the pre-training phase. While this is important to ensure that the model retains the knowledge of a full degradation path, there is naturally a risk that too many of the model's parameters are determined on the source domain.

### 6.2.3 Fine-tuning on a fleet of cars

The comparative analysis on the model performance reveals that the fine-tuned model obtained a lower RMSE loss compared to the pre-trained model. While the evaluation method does not directly assess the performance of predicting EoL, it provides insights into the model's generalizability to future SoH values and its ability to predict relative degradation within the observed timeframe. Although it is challenging to validate the accuracy of EoL predictions, these results demonstrate that fine-tuning enhances the predictive performance on field data. Furthermore, based on the findings from experiments in Section 5.1, collecting more field data will further improve the predictive performance and enable more accurate EoL predictions.

Comparing the driving behavior of the customers and their predicted degradation paths, the rate of the predicted paths agrees with the effect of the factors related to battery aging as described in Section 2.1.2. A customer with very high annual mileage is expected to have a faster degradation due to, for example, the increased cycles that the battery goes through. While higher annual mileage contributes to faster battery degradation, it does not necessarily mean that a customer that drives more will have a faster degradation which is illustrated by comparing predicted degradation paths for different customers. This emphasizes the need to consider multiple factors that affect battery degradation when predicting the lifespan of batteries in electric vehicles. That is, there is an interplay between many factors, such as the annual mileage, fast charging ratio, temperature, and SoC level. This interplay is seen to be, at least partly, captured by the network-based transfer learning model implemented.

# 7

# Conclusion

This thesis explored a transfer learning framework for predicting EoL of lithium-ion batteries for a fleet of EVs. The prediction of EoL in EV batteries posed three main challenges: no empirical EoL observations in the fleet, the inability of laboratory experiments to fully replicate real-world operating conditions, and that the collected data from the EVs contained noise, inconsistencies, and missing data.

To address the data quality issues, the lab and field data underwent several data preprocessing steps to obtain a dataset that reflected battery aging under real driving conditions. However, it is important to acknowledge that despite the steps, uncertainties still exist in both datasets. These uncertainties must be considered when applying the framework to the fleet of cars and making EoL predictions.

To overcome the challenges of the absence of EoL labels and the limitations of lab data, this thesis proposed a framework that combines lab and field data using a network-based transfer learning model. The experimental study evaluated the framework's effectiveness using lab datasets with degradation paths to EoL. The results showed improvement in predictive performance when increasing the fine-tuning data. However, the improvement was only significant at a certain fine-tuning amount. These findings suggest that implementing a transfer learning framework may be effective when the degradation paths of the batteries have sufficiently degraded.

Moreover, when applying the framework to the fleet of EVs, the fine-tuned model utilizing a combination of lab and field data outperformed the model trained solely on lab data. However, it is important to note that this evaluation was based on the existing degradation path and not specifically on EoL labels. Therefore, further validation is necessary to assess its generalizability to EoL predictions when EoL labels exist.

This thesis contributes a framework that combines lab and field data using transfer learning techniques to predict the EoL of batteries in EVs. The experimental findings provide evidence of the framework's effectiveness. As more field data is collected over time, there is great potential for improving the EoL predictions for the fleet of EVs.

## 7.1 Future work

Battery degradation is a mechanism with a temporal aspect; that is, it degrades over time, and the previous degrading state can affect the future state. However, due to the sparsity in the measurements of each car in the fleet, this limitation makes it difficult to model the temporal aspect over multiple data points. Therefore, an ANN with fully connected layers is chosen instead of, for instance, a Recurrent Neural Network (RNN). The selected ANN model does not inherently account for the temporal aspect of the data, as it lacks the ability to capture sequential information.

The implemented framework assumes that the customer is characterized by its parameter throughout the battery's lifetime. However, this assumption may not hold in real-world scenarios, as driving behavior can vary. With more data collected, an RNN can be implemented designed to handle sequential data and capture temporal dependencies. Such a model could then account for changes in driving behavior over the battery's lifetime and how it impacts future predictions. Previous studies implementing RNNs have demonstrated promising results in predicting the EoL of batteries [35, 36].

The project did not dive into understanding the predictions on individual customers in the fleet, partly due to only having early degradation paths and not being able to validate the results. However, an interesting future work could be to implement explainable AI to understand the predictions made by the model. Understanding different features' effects on the predicted degradation path using field data could be valuable to Volvo Cars as lots of resources go into experimental tests to better understand the factors that affect battery degradation.

The fleet of cars contains a wide variety of customers with different driving behavior from different parts of the world. Future work could investigate the field data further and look into how the customers could be segmented based on their driving behavior. By segmenting the customers, e.g., by how they drive and the climate that they are in, one model could be built for each group of customers which could yield a better performance than training a model on all the customers.

Finally, one aspect that could be of interest to future work is exploring more types of datasets. One way is investigating more cell chemistries to better understand knowledge transfer during the fine-tuning phase. It could also be valuable to use public datasets to apply the framework to, enabling comparison of results against other models.

# Bibliography


[1] David A Howey. "A challenging future for cars". In: *Nature Climate Change* 2.1 (2012), pp. 28–29.

[2] European Parlament. *EU ban on the sale of new petrol and diesel cars from 2035 explained*. Available at `https://www.europarl.europa.eu/news/en/headlines/economy/20221019STO44572/eu-ban-on-sale-of-new-petrol-and-diesel-cars-from-2035-explained`. Accessed on May 30, 2023. 2022.

[3] Gregory James Offer. "Automated vehicles and electrification of transport". In: *Energy & Environmental Science* 8.1 (2015), pp. 26–30.

[4] Volvo Cars. *Volvo Cars to be fully electric by 2030d*. Available at `https://www.media.volvocars.com/global/en-gb/media/pressreleases/277409/volvo-cars-to-be-fully-electric-by-2030`. Accessed on May 30, 2023. 2021.

[5] Hauke Engel, Russell Hensley, Stefan Knupfer, and Shivika Sahdev. "Charging ahead: Electric-vehicle infrastructure demand". In: *McKinsey Center for Future Mobility* 8 (2018).

[6] Mohammad A Hannan, MS Hossain Lipu, Aini Hussain, and Azah Mohamed. "A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations". In: *Renewable and Sustainable Energy Reviews* 78 (2017), pp. 834–854.

[7] Adrian König, Lorenzo Nicoletti, Daniel Schröder, Sebastian Wolff, Adam Waclaw, and Markus Lienkamp. "An overview of parameter and cost for battery electric vehicles". In: *World Electric Vehicle Journal* 12.1 (2021), p. 21.

[8] Robert Reinhardt, Ioannis Christodoulou, Santiago Gassó-Domingo, and Beatriz Amante García. "Towards sustainable business models for electric vehicle battery second use: A critical review". In: *Journal of environmental management* 245 (2019), pp. 432–446.

[9] Valentin Sulzer, Peyman Mohtat, Antti Aitio, Suhak Lee, Yen T Yeh, Frank Steinbacher, Muhammad Umer Khan, Jang Woo Lee, Jason B Siegel, Anna G Stefanopoulou, et al. "The challenge and opportunity of battery lifetime prediction from field data". In: *Joule* 5.8 (2021), pp. 1934–1955.

[10] Shunli Wang, Siyu Jin, Dan Deng, and Carlos Fernandez. “A Critical Review of Online Battery Remaining Useful Lifetime Prediction Methods”. In: *Front. Mech. Eng.* 7 (Aug. 2021). ISSN: 2297-3079. DOI: 10.3389/fmech.2021.719718.

[11] Dario Pevec, Jurica Babic, and Vedran Podobnik. “Electric Vehicles: A Data Science Perspective Review”. In: *Electronics* 8.10 (Oct. 2019), p. 1190. ISSN: 2079-9292. DOI: 10.3390/electronics8101190.

[12] Prabhakar Sharma and Bhaskor J. Bora. “A Review of Modern Machine Learning Techniques in the Prediction of Remaining Useful Life of Lithium-Ion Batteries”. In: *Batteries* 9.1 (Dec. 2022), p. 13. ISSN: 2313-0105. DOI: 10.3390/batteries9010013.

[13] Sinno Jialin Pan and Qiang Yang. “A Survey on Transfer Learning”. In: *IEEE Trans. Knowl. Data Eng.* 22.10 (Oct. 2009), pp. 1345–1359. ISSN: 1558-2191. DOI: 10.1109/TKDE.2009.191.

[14] Kailong Liu, Qiao Peng, Yunhong Che, Yusheng Zheng, Kang Li, Remus Teodorescu, Dhammika Widanage, and Anup Barai. “Transfer learning for battery smarter state estimation and ageing prognostics: Recent progress, challenges, and prospects”. In: *Advances in Applied Energy* 9 (Feb. 2023), p. 100117. ISSN: 2666-7924. DOI: 10.1016/j.adapen.2022.100117.

[15] Shahid A. Hasib, S. Islam, Ripon K. Chakrabortty, Michael J. Ryan, D. K. Saha, Md H. Ahamed, S. I. Moyeen, Sajal K. Das, Md F. Ali, Md R. Islam, Z. Tasneem, and Faisal R. Badal. “A Comprehensive Review of Available Battery Datasets, RUL Prediction Approaches, and Advanced Battery Management”. In: *IEEE Access* 9 (June 2021), pp. 86166–86193. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3089032.

[16] Meru A. Patil, Piyush Tagade, Krishnan S. Hariharan, Subramanya M. Kolake, Taewon Song, Taejung Yeo, and Seokgwang Doo. “A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation”. In: *Appl. Energy* 159 (Dec. 2015), pp. 285–297. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2015.08.119.

[17] Yunhong Che, Zhongwei Deng, Xianke Lin, Lin Hu, and Xiaosong Hu. “Predictive Battery Health Management With Transfer Learning and Online Model Correction”. In: *IEEE Trans. Veh. Technol.* 70.2 (Feb. 2021), pp. 1269–1277. ISSN: 1939-9359. DOI: 10.1109/TVT.2021.3055811.

[18] Abdallah A. Chehade and Ala A. Hussein. “A Collaborative Gaussian Process Regression Model for Transfer Learning of Capacity Trends Between Li-Ion Battery Cells”. In: *IEEE Trans. Veh. Technol.* 69.9 (June 2020), pp. 9542–9552. ISSN: 1939-9359. DOI: 10.1109/TVT.2020.3000970.

[19] Jens F Peters, Manuel Baumann, Benedikt Zimmermann, Jessica Braun, and Marcel Weil. “The environmental impact of Li-Ion batteries and the role of key parameters–A review”. In: *Renewable and Sustainable Energy Reviews* 67 (2017), pp. 491–506.

[20] Christopher D Rahn and Chao-Yang Wang. *Battery systems engineering*. John Wiley & Sons, 2013.

[21] Yuliya Preger, Heather M Barkholtz, Armando Fresquez, Daniel L Campbell, Benjamin W Juba, Jessica Romàn-Kustas, Summer R Ferreira, and Babu Chalamala. "Degradation of commercial lithium-ion cells as a function of chemistry and cycling conditions". In: *Journal of The Electrochemical Society* 167.12 (2020), p. 120532.

[22] Gregory L Plett. *Battery management systems, Volume I: Battery modeling*. Vol. 1. Artech House, 2015.

[23] MFR Zwicker, M Moghadam, W Zhang, and CV Nielsen. "Automotive battery pack manufacturing–a review of battery to tab joining". In: *Journal of Advanced Joining Processes* 1 (2020), p. 100017.

[24] Anthony Barré, Benjamin Deguilhem, Sébastien Grolleau, Mathias Gérard, Frédéric Suard, and Delphine Riu. "A review on lithium-ion battery ageing mechanisms and estimations for automotive applications". In: *Journal of Power Sources* 241 (2013), pp. 680–689.

[25] Andreas Podias, Andreas Pfrang, Franco Di Persio, Akos Kriston, Silvia Bobba, Fabrice Mathieux, Maarten Messagie, and Lois Boon-Brett. "Sustainability assessment of second use applications of automotive batteries: Ageing of Li-ion battery cells in automotive and grid-scale applications". In: *World Electric Vehicle Journal* 9.2 (2018), p. 24.

[26] Xiao-Guang Yang, Yongjun Leng, Guangsheng Zhang, Shanhai Ge, and Chao-Yang Wang. "Modeling of lithium plating induced aging of lithium-ion batteries: Transition from linear to nonlinear aging". In: *Journal of Power Sources* 360 (2017), pp. 28–40.

[27] Christoph R Birkl, Matthew R Roberts, Euan McTurk, Peter G Bruce, and David A Howey. "Degradation diagnostics for lithium ion cells". In: *Journal of Power Sources* 341 (2017), pp. 373–386.

[28] Joris Jaguemont, Loïc Boulon, and Yves Dubé. "A comprehensive review of lithium-ion batteries used in hybrid and electric vehicles at cold temperatures". In: *Applied Energy* 164 (2016), pp. 99–114.

[29] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*. Vol. 4. 4. Springer, 2006.

[30] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[31] Davide Chicco, Matthijs J Warrens, and Giuseppe Jurman. "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation". In: *PeerJ Computer Science* 7 (2021), e623.

[32] Aurélien Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.", 2022.

[33] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. "A Comprehensive Survey on Transfer Learning". In: *Proc. IEEE* 109.1 (July 2020), pp. 43–76. DOI: 10.1109/JPROC.2020.3004555.

[34] Jie Lu, Vahid Behbood, Peng Hao, Hua Zuo, Shan Xue, and Guangquan Zhang. "Transfer learning using computational intelligence: A survey". In: *Knowledge-Based Systems* 80 (May 2015), pp. 14–23. ISSN: 0950-7051. DOI: 10.1016/j.knosys.2015.01.010.

[35] Zhen Zhang, Wentao Zhang, Kuo Yang, and Shujing Zhang. "Remaining useful life prediction of lithium-ion batteries based on attention mechanism and bidirectional long short-term memory network". In: *Measurement* 204 (Nov. 2022), p. 112093. ISSN: 0263-2241. DOI: 10.1016/j.measurement.2022.112093.

[36] Yongzhi Zhang, Rui Xiong, He Hongwen, and Michael Pecht. "Long Short-Term Memory Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-Ion Batteries". In: *IEEE Trans. Veh. Technol.* PP (Feb. 2018), p. 1. DOI: 10.1109/TVT.2018.2805189.

[37] Wes McKinney. *pandas: a Foundational Python Library for Data Analysis and Statistics*. [Online; accessed 1. Jun. 2023]. 2011. URL: https://www.semanticscholar.org/paper/pandas%3A-a-Foundational-Python-Library-for-Data-and-McKinney/1a62eb61b2663f8135347171e30cb9dc0a8931b5.

[38] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. "Array programming with NumPy". In: *Nature* 585 (Sept. 2020), pp. 357–362. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2649-2.

[39] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, and Paul van Mulbregt. "SciPy 1.0: fundamental algorithms for scientific computing in Python". In: *Nat. Methods* 17 (Mar. 2020), pp. 261–272. ISSN: 1548-7105. DOI: 10.1038/s41592-019-0686-2.

[40] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. "Automatic differentiation in PyTorch". In: *OpenReview* (Oct. 2017). URL: https://openreview.net/forum?id=BJJsrmfCZ.

# A

# Appendix 1

## A.1 Processed lab data

For the processed lab data, a discrete set of values are selected which forms the 1440 customers. The discrete values for the different parameters are shown in Table A.1. By multiplying the number of values for each parameter together the number of customers becomes $8 \cdot 6 \cdot 6 \cdot 5 = 1440$.

**Table A.1:** The discrete set of parameters that forms all the simulated customer.

| **km/year** | **Fast charging ratio (%)** | **Temp (°C)** | SoC Level (%) |
|---|---|---|---|
| 2000 | 0 | -9 | 30 |
| 10000 | 10 | 8 | 50 |
| 20000 | 25 | 15 | 70 |
| 29000 | 50 | 20 | 85 |
| 46000 | 75 | 25 | 94 |
| 80000 | 100 | 30 | - |
| 118000 | - | - | - |
| 350000 | - | - | - |

## A.2 Implementation details

A cloud server is used to perform the data analysis, data preprocessing and model training. This cloud server enables the access to GPUs which speeds up the model training by parallelizing the computations. In total, three GPUs and 106 GB of memory is allocated as computational resources.

Furthermore, Python 3.8.8 is picked as the implementation language. For the data preprocessing, packages such as Pandas [37], NumPy [38] and SciPy [39] is used. The model building and training is set up in PyTorch using a integrated framework called lightningmodule [40].

## A.3 Hyperparameters for prediction model for fleet of cars

The hyperparameters to train the pre-trained and fine-tuned model for predicting EoL for the fleet of cars are seen in Table A.2.

**Table A.2:** Hyperparamter setup for transfer learning framework for combination of lab and field data. The abbrevations cc, bs, lr, ep and hn, lf, AEF stands for cell chemistry, batch size, learning rate, epoch,hidden neurons, layers frozen, all except first, respectively.

| **Model** | **Data** | **cc** | **bs** | **lr** | **ep** | **#layers** | **#hn** | **lf** |
|---|---|---|---|---|---|---|---|---|
| Pre-trained | Lab data | Cell B | 1028 | $1e^{-5}$ | 20 | 4 | 256 | None |
| Fine-tuned | Field data | - | 16 | $1e^{-4}$ | 30 | 3 | 256 | AEF |

**DEPARTMENT OF ELECTRICAL ENGINEERING**
**CHALMERS UNIVERSITY OF TECHNOLOGY**
Gothenburg, Sweden
www.chalmers.se


CHALMERS
UNIVERSITY OF TECHNOLOGY