

Learning Data Model Definition and Machine Learning Analysis for Data-Based Li-Ion Battery Performance Prediction

Byoungwook Kim[†] · Ji Su Park^{††} · Hong-Jun Jang^{††}

ABSTRACT

The performance of lithium ion batteries depends on the usage environment and the combination ratio of cathode materials. In order to develop a high-performance lithium-ion battery, it is necessary to manufacture the battery and measure its performance while varying the cathode material ratio. However, it takes a lot of time and money to directly develop batteries and measure their performance for all combinations of variables. Therefore, research to predict the performance of a battery using an artificial intelligence model has been actively conducted. However, since measurement experiments were conducted with the same battery in the existing published battery data, the cathode material combination ratio was fixed and was not included as a data attribute. In this paper, we define a training data model required to develop an artificial intelligence model that can predict battery performance according to the combination ratio of cathode materials. We analyzed the factors that can affect the performance of lithium-ion batteries and defined the mass of each cathode material and battery usage environment (cycle, current, temperature, time) as input data and the battery power and capacity as target data. In the battery data in different experimental environments, each battery data maintained a unique pattern, and the battery classification model showed that each battery was classified with an error of about 2%.

Keywords : Lithium-ion Battery, Battery Performance Prediction, Machine Learning, Learning Data Model

데이터 기반 리튬 이온 배터리 성능 예측을 위한 학습 데이터 모델 정의 및 기계학습 분석

김 병 옥[†] · 박 지 수^{††} · 장 흥 준^{††}

요 약

리튬 이온 배터리는 사용 환경과 양극재 조합 비율에 따라 배터리의 성능이 좌우된다. 고성능 리튬 이온 배터리를 개발하기 위해서는 양극재 비율을 다양하게 변화시켜가면서 배터리를 제작하고 성능을 측정해야 한다. 하지만 모든 변수 조합에 대해 배터리를 제작하고 성능을 측정하기에는 많은 시간과 비용이 소모된다. 그렇기 때문에 최근에는 데이터 기반으로 인공지능 모델을 활용하여 배터리의 성능을 예측하고자 하는 연구가 활발히 진행되고 있다. 그러나 기존 공개 배터리 데이터는 동일한 배터리로 측정 실험을 하였기 때문에 양극재 조합 비율은 고정되어 있어서 데이터 속성으로 포함되지 않았다. 본 논문에서는 양극재 소재 조합 비율에 따른 배터리의 성능을 예측할 수 있는 인공지능 모델 개발에 필요한 학습 데이터 모델을 정의한다. 우리는 리튬 이온 배터리의 성능에 영향을 미칠 수 있는 요인을 분석하여 양극재 소재별 질량과 배터리 사용 환경을 입력데이터로, 배터리의 출력과 용량을 목적 데이터로 정의하였다. 공개 배터리 데이터 중에는 양극재 비율이 포함된 데이터가 없어 양극재 비율을 모두 동일한 값으로 설정한 제한된 데이터로 다중 선형회귀 분석, 서포트 벡터 회귀분석, 다중 로지스틱 회귀 분석, LSTM 분석을 수행하였다. 실험 환경이 다른 배터리 데이터에서 각각의 배터리 데이터는 고유한 패턴을 유지하였으며, 배터리 분류 모델은 각각의 배터리를 약 2%의 오차로 분류하는 것으로 나타났다.

키워드 : 리튬 이온 배터리, 배터리 성능 예측, 기계 학습, 학습 데이터 모델

1. 서 론

최근 전기자동차, 드론, 모바일기기 등의 수요 확대와 더불어 에너지 저장용 2차 전지 시장도 급성장하고 있다. 한 번의 충전으로 고전압으로 장시간 사용할 수 있는 배터리가 요구되면서 고성능, 고효율의 2차 배터리 개발 연구가 되고 있다. 2차 배터리는 납축전지, 니켈수소전지, 리튬 이온 전지 등 2차 배터리의 양극재 소재에 따라서 다양한 종류가 존

※ 이 논문은 동신대학교 학술연구비에 의하여 연구되었음. 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2021R1F1A1049387).

※ 2022년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과임(2021RIS-002).

† 정 회 원 : 강릉원주대학교 컴퓨터공학과 조교수

†† 총신회원 : 전주대학교 컴퓨터공학과 조교수

Manuscript Received : January 25, 2023

First Revision : February 6, 2023

Accepted : February 7, 2023

* Corresponding Author : Hong-Jun Jang(hongjunjang@jj.ac.kr)

재한다. 이 중에서 리튬 이온 배터리는 에너지 밀도가 높고, 충방전 수명이 우수하기 때문에 다양한 분야에서 사용되고 있다[1, 2]. 리튬 이온 배터리는 4대 핵심소재(양극재, 음극재, 분리막, 전해액)로 구성되며 배터리의 에너지 용량, 출력, 안정성 등 배터리 성능에 직접적인 영향을 미친다. 리튬 이온 배터리 4대 핵심소재 중에서 양극재는 배터리의 성능을 좌우하는 ‘용량’과 ‘전압’을 결정하는 핵심 소재이다. 배터리의 에너지 밀도는 배터리의 양극과 음극에 사용되는 소재의 종류와 비율, 배터리 내부 공간의 효율적인 사용에 따라 달라진다[3]. 따라서 배터리의 효율성, 안정성, 신뢰성이 보장된 우수한 성능의 배터리 개발에 대한 연구에 많은 관심이 집중되고 있다. 최적의 성능을 낼 수 있는 양극재 조합 비율을 찾거나 특정 온도나 출력 조건 등 배터리의 사용 환경에 따른 배터리 성능을 예측하는 연구가 수행되고 있다. 최적 성능의 양극재 비율을 찾기 위해서는 직접 배터리를 제작하고 성능을 측정해야 하는 실험을 수 없이 반복해야 하지만 직접 실험을 수행하기에는 비용과 시간이 소요된다는 한계점을 가지고 있다.

직접 제작 및 실험을 통해 배터리 성능을 측정하는 단점을 극복하고자 최근에는 배터리 성능 측정 데이터 기반으로 기계학습을 이용하여 배터리 성능을 예측하고자 하는 연구가 진행되고 있다. 공개된 배터리 노화 데이터나 충방전 데이터를 이용하여 리튬 이온 배터리의 잔여 유효 수명(Remaining Useful Life, RUL)이나 건강상태(State of Health, SOH)를 예측하는 하였다. 예측에 사용된 방법은 전통적인 기계학습 기법인 회귀분석[4], CNN (Convolutional Neural Network)[5, 6], RNN (Recurrent Neural Network)[7], LSTM (Long Short-Term Memory)[5, 6, 8, 9], GRU (Gated Recurrent Unit)[9]이 사용되었다. 그러나 기존의 딥러닝을 활용한 배터리 성능 예측 연구는 각 실험마다 양극재 소재와 비율이 고정된 배터리에서 생성된 데이터만을 대상으로 하여 배터리의 RUL이나 SOH를 예측하였다. 그렇기 때문에 양극재의 소재나 조합 비율이 달라질 경우 배터리 성능이 어떻게 변화하는지 예측하는데 한계가 있다. 양극재 소재 및 조합 비율에 따른 배터리 성능 예측을 위해 인공지능 모델을 학습시키기 위해서는 학습 데이터 수집이 선행되어야 한다. 그러나 아직까지 양극재 소재와 조합 비율이 포함되어 공개된 학습 데이터는 존재하지 않는다. 따라서 본 논문에서는 양극재 소재 및 조합 비율에 따른 배터리의 성능을 예측할 수 있는 모델 개발에 필요한 학습 데이터 구축을 위해 학습 데이터 모델을 정의한다. 일부 속성이 제한된 데이터이지만 공개된 데이터를 활용하여 전통적인 기계학습 방법으로 분석한 후 학습 데이터의 활용 가능성에 대해 논의하고자 한다. 본 연구의 기여도는 다음과 같다.

1) 양극재 소재 및 조합 비율에 따른 리튬 이온 배터리의 성능 예측 모델 개발에 필요한 학습 데이터의 모델을 정의하

였다. 인공지능 모델을 기반으로 배터리의 성능 예측에 대한 관심이 높아지면서 다양한 연구자들이 배터리 데이터를 공개하고 배터리 성능 측정에 관해서는 자세히 설명하고 있지만 실험에 사용한 배터리의 양극재 소재 및 조합 비율에 대한 데이터는 공개되지 않고 있다. 본 연구를 통해 배터리 데이터를 공개할 때 배터리 양극재 소재 및 조합 비율에 대한 정보도 중요하다는 것을 인식할 수 있을 것이다.

2) 배터리 실험 환경(사이클, 온도, 전류 등)에 따라 전통적인 기계학습 방법으로도 개별 배터리를 구별할 수 있음을 보였다. NASA Ames Prognostics Center에서 공개한 데이터를 활용하여 다중선형회귀와 서포트 벡터 회귀분석을 실시하여 배터리별로 고유의 패턴이 나타나는 것을 확인하였으며, 다중 로지스틱 회귀분석을 통해 입력 데이터에 따라 개별 배터리를 분류하는 성능이 있음을 보였다.

본 논문의 2장에서는 리튬 이온 배터리의 성능 측정에 중요한 변수인 양극재 소재 및 조합 비율에 대해 알아보고, 딥러닝을 활용한 리튬 이온 배터리 성능 예측의 최신 연구 동향에 대해 알아본다. 3장에서는 양극재 소재 및 조합 비율에 따른 배터리 성능 예측에 필요한 학습 데이터 모델을 정의한다. 4장에서는 실험에 사용된 데이터와 전통적인 기계학습 방법으로 분석한 결과를 제시한다. 5장에서는 배터리 학습 데이터에 양극재 소재 및 비율 데이터가 필요한 이유와 전체 연구 요약 및 결론에 대해 서술한다.

2. 관련 연구

2.1 양극재 조합

리튬이온전지에서 양극재로 사용되는 리튬은 단독으로 사용하기엔 불안정하기 때문에 Table 1과 같이 다른 금속 원소들과 조합하여 사용한다.

리튬 이온 배터리의 성능을 결정하는 핵심적인 소재는 LiCoO_2 (LCO)로 다른 물질과 합성하기 쉽고 전기화학적 성능도 우수하기 때문에 리튬 이온 배터리 개발 초기에 양극재 소재로 주로 사용되었다[10]. 그러나 코발트(Co)는 에너지 밀도도 높지 않고 가격이 비싸기 때문에 배터리의 에너지 밀도를 높이고 가격을 낮출 수 있는 새로운 양극재를 찾는 연구가 활발히 이루어지고 있다[11]. 이러한 연구 과정의 결과 니켈(Ni), 망가니스(Mn), 알루미늄(Al) 등이 새로운 양극재로

Table 1. Lithium-ion Battery Anode Material Synthetic Material

Metal element	Roles
Nickel(Ni)	Energy Density Determination
Cobalt(Co)	Increased stability
Manganese(Mn)	Increased stability
Aluminium(Al)	Power enhancement

활용되고 있다. 리튬 이온 배터리를 구성하는 양극재는 리튬(Li), 코발트(Co), 망가니스(Mn), 알루미늄(Al) 등의 소재를 조합하여 만들어지며 각 소재들의 조합 비율과 합성 조건(합성 온도 및 시간 등)에 따라 양극재의 성능이 달라지고 배터리의 성능에 영향을 준다[12]. 그렇기 때문에 최적의 용량과 출력을 낼 수 있는 배터리를 제작하기 위해서는 다양한 형태의 양극재 조합 비율 및 합성 조건에 따라 배터리를 제작한 후에 배터리의 성능을 측정해야 한다.

2.2 딥러닝 기반 리튬 이온 배터리 성능 예측

인공지능 기술이 발전하면서 데이터 기반 리튬 이온 배터리의 성능을 예측하는 연구들이 진행되고 있다.

Severson et al.[4]은 배터리의 열화(Degradation) 현상과 연관성이 변수를 도출하였다. 충방전 몇 사이클 후에 배터리의 SOH가 75% 이하가 되는지를 예측하는데 기존 배터리 실험 데이터의 속성을 이용하는 것이 아니라 $\Delta Q_{i-j}(V)$ 라는 변수(i 번째 사이클 때의 전압과 j 번째 사이클 때의 전압 차이)를 정의하고, 이 변수를 회귀분석을 이용하여 100사이클에서 9.1% 오류로 SOH가 75% 이하가 될 것을 예측하였다. 또한 4~5 사이클만에 4.9% 오류로 배터리가 평균보다 오랜 수명을 갖을지를 예측하였다.

Lim et al.[5]은 CNN과 LSTM을 결합한 하이브리드 모델을 이용하여 배터리 잔여 수명 예측을 연구하였다. 추정 오차를 줄이기 위해 시계열 신호의 노이즈를 분해하는 EMD(Empirical Mode Decomposition) 기법을 이용하여 데이터를 전처리하는 방법을 제안하였다. EMD를 적용한 방법인 EMD-CNN-LSTM 모델의 평균 추정 오차는 2.92%로 CNN-LSTM의 평균 추정 오차 3.23%보다 높은 성능을 나타낸다고 보고하였다.

Hong et al.[6]은 CNN과 LSTM을 결합한 LRCN(Long-term Recurrent Convolutional Network)을 이용하여 SOH(State of Health)를 추정하였다. CNN은 사이클 데이터에서 노화에 영향을 주는 특성을 추출하는데 활용되고, LSTM은 시계열 데이터를 분석하는 모델로 활용하였다. 기존의 CNN과 LSTM만으로 측정된 성능보다 LRCN의 정확도가 높은 것으로 보고하였다.

Kim et al.[7]은 리튬 이온 배터리의 온보드의 SOH 및 RUL 추정을 위한 딥러닝 기반 진단 프레임워크인 KIRNN(Knowledge-Infused Recurrent Neural Network)를 제안하였다. KIRNN은 경험적 모델과 심층 신경망 모델을 결합한 하이브리드 모델로 신경망의 확률론적 예측을 얻기 위해 몬테카를로 드롭아웃을 이용하여 추정의 견고성을 높였다. 베이스라인 딥러닝 모델보다 성능이 우수함을 제시하였다.

Park et al. [8]은 멀티채널 LSTM을 이용하여 RUL을 예측하는 기법을 제안하였다. 용량 재생 현상이 있는 경우에도 RUL을 추정하기 위해서 노화 데이터에 따라 달라지는 전압,

전류 및 온도 프로파일과 같은 여러 데이터를 고려하였다. 베이스라인 LSTM 모델보다 단일채널 LSTM 모델이 MAPE가 39.2% 성능향상을 보였고, 멀티채널 LSTM 모델은 63.7% 성능이 향상되었음을 보고하였다.

Jung et al.[9]은 딥러닝 알고리즘의 일종인 LSTM(Long Short-Term Memory)과 GRU(Gated Recurrent Unit)를 이용하여 리튬 이온 배터리의 잔여 유효 수명을 예측하였다. LSTM과 GRU를 이용하여 예측한 잔여 유효 수명은 모두 5% 이내의 오차를 나타냈으며, epoch 150, batch size 40일 때 오차는 1%수준으로 가장 낮게 나온 것으로 보고하였다.

딥러닝을 활용하여 배터리의 성능을 예측하는 연구는 대부분 한정된 배터리에서 측정된 데이터만을 이용하여 배터리의 RUL이나 SOH를 예측하는 연구로 제한된다. 즉, 각각의 실험에 사용된 배터리는 동일한 제조 공정을 거친 데이터로 모두 동일한 양극재 소재와 비율을 가진다. 따라서 양극재의 소재나 조합 비율이 바뀌는 경우는 예측 오차가 커질 수밖에 없다. 양극재 소재 및 조합 비율에 따른 배터리 성능을 예측하기 위한 인공지능 모델을 개발하기 위해서는 학습 데이터가 필요하나 아직까지 공개된 학습 데이터는 존재하지 않는다. 따라서 본 논문에서는 양극재 소재 및 조합 비율에 따라 배터리의 성능을 예측할 때 필요한 학습 데이터 구축을 위해 필요한 데이터 모델을 정의한다.

3. 데이터 모델

3.1 데이터 스키마 정의

학습 데이터는 Table 2에서 보는 바와 같이 입력데이터와 타깃데이터로 구성된다. 입력 데이터는 배터리의 성능에 영향을 미치는 요인인 양극재 소재의 양과 실험 환경으로 구성되며 목적 데이터는 출력(Power)과 용량(Capacity)으로 구성된다.

1) 입력 데이터

양극재 소재의 속성은 니켈(Nickel), 코발트(Cobalt), 망가니스(Manganese), 알루미늄(Aluminium)의 질량(g)으로 정의하였다. 양극재 소재 비율은 각 양극재 소재의 질량 비율로 계산할 수 있다. 단위를 질량으로 설정한 것은 동일한 비율이라 할지라도 총 질량이 달라지면 배터리 성능이 달라질 수 있기 때문에 단순히 비율로만 기록되면 양극재 소재의 전체 질량을 속성으로 추가시켜야 한다. 따라서 속성의 간소화를 위해 양극재 소재의 질량으로 데이터를 기록하고 어플리케이션 내에서 비율을 계산하여 활용하도록 속성을 정의하였다.

리튬 이온 배터리는 일반적으로는 정전류-정전압 방식의 충방전 실험을 통해 수집된 데이터로 성능을 측정한다. 충전은 완전 방전 상태의 배터리를 정해진 정전류(Constant

Table 2. Training Data Schema for Battery Performance Prediction

Type	Part	Attributes	Unit	Description
Input Data	Cathode	Lithium(Li)	g	Litium mass
		Nickel(Ni)	g	Nickel mass
		Cobalt(Co)	g	Cobalt mass
		Manganese(Mn)	g	Manganese mass
		Aluminium(Al)	g	Aluminium mass
	Setting	Cycle		Number of charge/discharge
		C-rate	C	Charge and discharge rates of a battery
		Temperature	℃	Air temperature
Timestamp			Battery performance measurement time	
Target Data	Performance	Power	W	Amount of electricity can use in 1 second
		Capacity	mAh	

Current; CC)에서 충전을 시작하고, 전압이 Charge Cut-off Voltage에 도달하면 정전압(Constant Voltage; CV)를 유지하면서 전류를 낮추어 배터리의 SOC가 100% 될 때까지 충전한다. 방전은 정해진 전류로 Discharge Cut-off Voltage가 될 때까지 정전류 모드로 수행한다. 충전과 방전을 한 번씩 하는 것을 1 사이클(Cycle)이라고 한다. 배터리 성능 측정 실험에서 측정할 수 있는 데이터를 속성으로 정의하였다. 기본적으로 배터리 실험에서 몇 번째 사이클인지에 대한 속성을 추가하였다. 충방전 시 흘리는 전류의 양은 C-rate로 설정하였다. 배터리가 놓인 대기의 온도(Temperature)도 배터리 성능에 많은 영향을 미치는 연구가 많다. 따라서 배터리의 온도도 속성으로 추가하였다. 실험이 진행되는 시간도 수집될 수 있도록 속성으로 추가하였다.

2) 목적 데이터

우리가 양극재 소재 및 조합 비율과 실험 환경 변수에 따라 예측하고자 하는 변수는 배터리의 출력(Power)과 용량(Capacity)이다. 출력은 배터리에서 1초 동안 꺼낼 수 있는 전기의 양으로 $W(\text{출력})=A(\text{전류})\times V(\text{전압})$ 으로 결정된다. 배터리를 사용하는 환경에 따라 전류의 양은 변화할 수 있으므로 출력은 전압을 예측하는 것으로 설정하였다. 용량은 배터리에 저장할 수 있는 전기의 양이다. 리튬 이온 배터리는 충방전 사이클을 거치며 열화(degradation)현상이 나타나며 용량이 감소되며 성능이 저하된다. 배터리 용량의 감소율에 따라 부가적으로 얻을 수 있는 성능 정보는 SOH이다. SOH는 Equation (1)과 같이 구한다. 배터리를 몇 번의 사이클을 진행한 후에 용량($Capacity_{cycle}$)을 최초 충전시켰을 때의 용량($Capacity_{init}$)으로 나눈 값이다.

$$SOH(\%) = \frac{Capacity_{cycle}}{Capacity_{init}} \quad (1)$$

4. 머신러닝 기반 성능 예측 및 분류

4.1 데이터셋

인공지능을 이용하여 리튬 이온 배터리 양극재 및 내부 저항에 따른 용량 및 출력 예측 모델을 개발하기 위해서는 인공지능 모델을 학습시킬 수 있는 학습 데이터가 필요하다. 양극재 비율 및 내부 저항에 따른 배터리 성능 예측을 위해서는 3장에서 정의한 데이터가 필요하나 양극재 비율 및 내부저항에 따른 배터리 성능을 측정할 수 있는 공개된 학습 데이터가 존재하지 않는다. 따라서 본 연구의 실험에서는 공개된 배터리 성능 측정 데이터 중에서 본 연구에서 정의한 데이터 스키마에 해당되는 데이터만으로 예측 모델 학습을 진행한다.

공개된 배터리 성능 측정 데이터 중에는 NASA Ames Prognostics Center에서 공개한 배터리 노화 데이터가 있다[13]. 여러 개의 배터리 데이터 중에 배터리 노화를 예측하기 위해 측정을 실험한 18650 리튬 이온 배터리 4개(B0005, B0006, B0007, B0018)로 구성된 데이터를 활용하였다. 데이터가 기록된 실험은 실온(23~24℃)에서 CC-CV (Constant Current-Constant Voltage) 충전 모드 및 CC (Constant Current) 방전 모드로 진행되었으며 Table 3과 같이 충방전 전류로 설정되었다. 충전 프로파일은 1.5A로 4.2V가 될 때까지 전류를 흘려주다 4.2V가 되면 전류를 20mA까지 줄이며 CC-CV 충전을 하였다. 방전 프로파일은 2A로 cut-off

Table 3. Test Setting of Battery

Battery No.	Constant Charge Current(A)	Charge Cut-off Voltage(V)	Discharge Current(A)	Discharge Cut-off Voltage(V)
B0005	1.5	4.2	2.0	2.7
B0006	1.5	4.2	2.0	2.5
B0007	1.5	4.2	2.0	2.2
B0018	1.5	4.2	2.0	2.5

voltage가 될 때까지 방전한다. 배터리마다 cut-off voltage는 각각 2.2~2.7V로 다르게 설정되었다.

데이터는 충전, 방전, 임피던스 측정 과정이 반복적으로 기록되어 있다. 본 실험에서는 방전과정에서 얻을 수 있는 용량을 이용하여 SOH를 예측해보는 실험을 실시하였다. 방전 데이터에는 다음의 데이터가 포함되어 있다. 용량(capacity), 배터리 단자 전압(voltage measured), 배터리 출력 전류(current measured), 배터리 온도(temperature measured), 측정 시간(time)이다. NASA 데이터에서 입력 데이터로 배터리 단자 전압, 배터리 출력 전류, 배터리 온도, 측정 시간을 독립변수로 사용하였다. NASA 데이터에는 양극재 질량이 포함되어 있지 않기 때문에 독립변수에 리튬, 니켈, 망가니스, 코발트 속성을 추가한 후에 값을 0으로 채운 후에 데이터를 생성하였다. 데이터는 전처리를 통해 모든 독립변수의 스케일을 [0,1]로 정규화하였다. 데이터에 기록된 충방전 순서에 따라 사이클을 카운트하여 싸이클 변수를 추가하였다. 이번 실험에서 예측할 변수는 용량으로 사이클마다 측정되는 용량을 통해 각 배터리의 SOH를 구한 후에 SOH를 예측하는 방식으로 진행하였다. 4개의 배터리 데이터를 각각 따로 분석하였으며, 모든 배터리 데이터를 하나로 통합한 후(Unified)에 분석하였다.

4.2 머신러닝

우리는 양극재 비율이 포함된 배터리 데이터를 대상으로도 기계학습을 이용하여 SOH를 예측하는데 문제가 없는지를 확인해보기 위해 3가지 전통적인 기계학습 알고리즘을 이용하여 분석한다. 분석에 사용된 기계학습 알고리즘은 다중 선형회귀 분석, 서포트 벡터 회귀분석, 다중 로지스틱 회귀분석이다. 본 실험에 사용된 데이터와 소스코드는 <https://github.com/Kim-ByoungWook/Battery>에서 확인할 수 있다.

1) 다중 선형회귀 분석

다중 선형회귀(Multiple linear regression)는 독립변수가 2개 이상인 경우 목적변수를 예측하는 분석 방법이다. 즉, 다수의 독립 변수와 하나의 종속 변수의 선형 관계를 모델링하는 것이다. 독립 변수는 입력 데이터에 해당되며 종속 변수

는 목적 데이터(SoH)에 해당된다. 다중 선형회귀 분석의 식은 Equation (2)와 같으며, x_i 는 독립변수이고 y 는 종속 변수이다.

$$y = W_1x_1 + W_2x_2 + \dots + W_nx_n \quad (2)$$

Table 4는 다중 선형회귀 분석을 하여 도출한 결정 계수를 보여주고 있다. 데이터셋 이름 뒤에 _mod 접미사가 붙은 것은 양극재 속성을 포함한 데이터이다. 우리가 사용한 데이터에는 양극재 속성이 포함되어 있지 않아 속성을 추가하였지만 모든 값을 0으로 설정하였기 때문에 이 실험에서는 두 종류의 데이터에 대한 차이는 보이지 않는다.

개별 배터리 데이터에 대한 결정계수(Determination Coefficient)는 1이 나왔다. 모든 배터리 데이터에서 1로 나온 것은 한 번의 방전과정에서 측정되는 전류량, 온도와 같은 데이터의 갯수는 다수(대략 5,000개)인 반면에 총 방출되는 전류 값은 하나이기 때문이다. 즉 용량은 하나의 값이고 이 값이 약 5,000개의 데이터에 동일하게 적용되기 때문에 예측이 높게 나왔다. 하지만 실험 조건이 다른 4개의 배터리 데이터가 혼합되면 (Unified) 결정계수는 0.9138으로 혼합되지 않은 데이터보다 낮아진 것으로 나타났다. 이를 통해 우리는 양극재의 비율과 실험 조건이 다른 데이터가 포함된다면 결정계수는 더욱 낮아질 수 있을 것으로 보인다.

2) 서포트 벡터 회귀분석

서포트 벡터 머신(Support Vector Machine; SVM)은 기계학습 알고리즘 중에서 분류와 회귀 분석에 사용되는 지도 학습 모델이다. 서포트 벡터 머신은 주어진 데이터를 분류할 때 분류되는 데이터의 그룹의 마진이 큰 경계선을 찾는 방식으로 데이터를 분류한다. 동일한 방법으로 서포트 벡터 회귀(Support Vector Regression; SVR)는 데이터를 잘 설명하는 일반화된 식을 찾는데 SVM과는 다르게 서로 다른 분류에 속한 관측치 사이에 마진이 최대가 되는 식을 찾는 것이다.

Table 5는 SVR 분석에서 MSE(Mean Squared Error)와 RMSE(Root Mean Square Error)값을 보여준다. 데이터 정

Table 4. Determination Coefficient in Multiple Linear Regression

Battery No.	Determination coefficient(R2)	Battery No.	Determination coefficient(R2)
B0005	1.0000	B0005_mod	1.0000
B0006	1.0000	B0006_mod	1.0000
B0007	1.0000	B0007_mod	1.0000
B0018	1.0000	B0018_mod	1.0000
Unified	0.9138	Unified_mod	0.9138

Table 5. MSE and RMSE in Support Vector Machine

Battery No.	MSE	RMSE
B0005	7.6215×10^{-24}	3.8107×10^{-24}
B0006	1.7674×10^{-24}	8.8371×10^{-25}
B0007	1.0914×10^{-19}	5.4574×10^{-19}
B0018	2.0096×10^{-9}	1.0048×10^{-9}
Unified	0.0022	0.0011

Table 6. Accuracy in Multinomial Logistic Regression Analysis

Battery No.	Accuracy
Unified	0.9795

구화로 인해 개별 데이터에 대한 오차가 0과 1사이이기 때문에 전반적으로 MSE와 RMSE 값은 작게 나온다. 개별 배터리에 대한 오차보다 통합 데이터에서는 MSE가 0.0022, RMSE가 0.0011로 상대적으로 높은 것으로 나타났다. 즉, 각각 배터리 데이터는 하나의 데이터로 인식하는 반면 데이터가 통합되면서 하나의 배터리 데이터라고 보기 어렵다. 두 개의 회귀 분석 결과를 통해 우리는 양극재의 비율과 실험 조건이 다른 다수의 데이터가 추가되면 각각의 배터리 데이터가 하나의 배터리 데이터인 것으로 인식될 수 있으며 전통적인 기계 학습 방법으로는 분석의 정확성이 낮아진다.

3) 다중 로지스틱 회귀 분석

다중 로지스틱 회귀 분석 (Multinomial logistic regression analysis)은 목적변수 범주가 3개 이상이며 명목 데이터일 때 사용하는 로지스틱 회귀분석이다. 목적 변수를 명목 데이터로 설정해야하기 때문에 앞의 두 개의 회귀분석과는 다르게 학습 데이터에서 배터리 No를 목적변수로 설정한 후에 독립변수만 주어졌을 때, 각각의 배터리를 예측할 수 있는지 분석해 보았다.

Table 6은 다중 로지스틱 회귀분석에서 정확도를 보여준다. 모든 데이터셋에서 각각의 배터리를 분류해야 하기 때문에 4개의 데이터를 통합한 데이터에 대해서만 분류를 수행하였다. 분류의 정확도는 0.979로 약 2%의 오차로 각 배터리를 분류하는 성능이 나왔다.

4) 장단기 메모리

장단기 메모리(Long Short Term Memory; LSTM)는 Recurrent neural network(RNN)이 긴 기간의 의존성을 다루기 어려운 단점을 보완하기 위한 신경망 구조이다. 즉, LSTM은 출력과 먼 위치에 있는 정보를 기억할 수 없다는 단점을 보완하여 장/단기 기억이 가능하게 설계되었다. 따라서 시계열 데이터 등 신경망에 입력되는 데이터에 시간 속성이 포함된 데이터를 분석하는데 LSTM이 주로 활용되고 있다 [14, 15].

본 연구에서도 LSTM을 이용하여 SOH를 예측하는 실험을 진행하였다. 실험에 사용된 LSTM 모델은 Fig. 1과 같이 3개의 LSTM층과 1개의 Dense층으로 구성되었다. LSTM층의 Units은 200으로, Dense층의 Units은 1로 설정하였다. 훈련 과정에서 epoch는 200, batch size는 25, drop out rate는 0.3으로 설정하였다. 모델 성능은 MSE로 평가하였다.

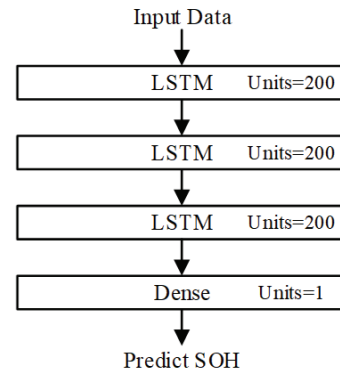


Fig. 1. LSTM Modes

‘B0005’ 데이터로 모델 훈련을 하였으며, ‘B0006’ 데이터로 검증을 실시하였다. LSTM 모델을 이용하여 SOH를 예측의 RMSE는 0.063으로 나타났다. 하이퍼 파라미터에 대한 최적화를 진행하지 않은 상태에서도 RMSE값이 낮게 나와 LSTM 모델을 통한 SOH 예측 모델도 양극재 조합 비율에 따른 성능 예측에 활용될 수 있을 것이다.

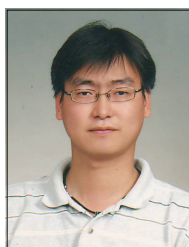
5. 결론 및 고찰

본 연구는 다양한 배터리의 사용환경과 양극재 조합 비율에 따라 고성능 리튬 이온 배터리 개발을 위해 직접 모든 배터리는 제작하고 실험해야 하는 문제점을 극복할 수 있는 양극재 소재 및 조합 비율에 따른 배터리 성능 예측 모델을 학습시킬 데이터 모델을 정의하였다. 학습 데이터 모델은 리튬 이온 배터리의 성능(출력 및 용량)에 영향을 미칠 수 있는 요인을 분석하여 양극재 소재별 질량과 배터리 사용 환경(사이클, 전류, 온도, 시간)으로 입력데이터로 설정하였으며, 출력과 용량을 목적 데이터로 설정하였다. 공개된 배터리 데이터 중에는 양극재 비율이 포함된 데이터가 없기 때문에 배터리 양극재 비율을 모두 동일한 값으로 설정한 제한된 데이터이지만 전통적인 기계학습 방법으로 성능을 예측하는데 활용될 수 있음을 보이기 위해 다중 선형회귀 분석, 서포트 벡터 회귀분석, 다중 로지스틱 회귀 분석을 수행하였다. 실험 환경이 다른 배터리 데이터에서 각각의 배터리 데이터는 고유한 패턴을 유지하였으며 입력 데이터로 각각의 배터리를 약 2%의 오차로 분류하는 것으로 나타났다.

본 연구에서는 배터리 성능에 영향을 미치는 요인으로 양극재 소재 및 조합 비율과 배터리 사용 환경에 대해만 다루었다. 하지만 배터리의 성능은 내부 저항에도 영향을 받기 때문에 내부 저항에 영향을 미치는 요인들도 추가적으로 연구할 필요가 있다. 또한 학습 데이터를 보완하여 양극재 소재 및 비율에 대한 실제 데이터를 활용하여 성능을 예측하는 알고리즘을 적용해 보아야 할 것이다.

References

- [1] A. Manthiram, "A reflection on lithium-ion battery cathode chemistry," *Nature Communication*, Vol.11, No.1550, pp.1-9, 2020.
- [2] L. Chen et al, "Estimation the internal resistance of lithium-ion-battery using a multi-factor dynamic internal resistance model with an error compensation strategy," *Energy Reports*, Vol.7, pp.3050-3059, 2021.
- [3] H. C. Lee, "Understanding and evaluating the performance and safety of lithium-ion batteries," *The Magazine of the IEIE*, Vol.34, Iss.12, pp.20-29, 2007.
- [4] K. A. Severson et al., "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, Vol.4, pp.383-391, 2019.
- [5] J.-Y. Lim, D.-H. Kim, T.-W. Noh, and B.-K. Lee, "Remaining useful life prediction for lithium-ion batteries using EMD-CNN-LSTM hybrid method," *The Transactions of the Korean Institute of Power Electronics*, Vol.27, No.1, pp.48-55, 2022.
- [6] S.-R. Hong, M. Kang, H.-G. Jeong, J.-B. Baek, and J.-H. Kim, "State of health estimation for Lithium-Ion batteries using Long-term Recurrent convolutional network," *The Transactions of the Korean Institute of Power Electronics*, Vol.26, No.3, pp.183-191, 2021.
- [7] S. W. Kim, K.-Y. Oh, and S. Lee, "Novel informed deep learning-based prognostics framework for on-board health monitoring of lithium-ion batteries," *Applied Energy*, Vol.315, pp.119011, 2022.
- [8] K. Park, Y. Choi, W. J. Choi, H.-Y. Ryu, and H. Kim, "LSTM-Based battery remaining useful life prediction with Multi-Channel charging profiles," *IEEE Access*, Vol.8, pp.20786-20798, 2020.
- [9] S.-J. Jung and J.-W. Hur, "Deep learning approaches to RUL prediction of Lithium-ion batteries," *Journal of the Korean Society of Manufacturing Process Engineers*, Vol.19, No.12, pp.21-27, 2020.
- [10] B. Xu, D. Qian, Z. Wang, and Y. S. Meng, "Recent progress in cathode materials research for advanced lithium ion batteries," *Materials Science and Engineering: R: Reports*, Vol.73, Iss.56, pp.51-65, 2012.
- [11] J. Xiao, N. A. Chernova, and M. S. Whittingham, "Influence of manganese content on the performance of LiNi_{0.9-y}Mn_yCo_{0.102} (0.45 ≤ y ≤ 0.60) as a cathode material for Li-Ion batteries," *Chemistry of Materials*, Vol.22, pp.1180-1185, 2010.
- [12] S. Y. Kim, S.-H. Choi, E. J. Lee, and J.-S. Kim, "Synthesis and electrochemical performance of Ni-rich NCM cathode materials for Lithium-Ion batteries," *Journal of the Korean Electrochemical Society*, Vol.20, No.4, pp.67-74, 2017.
- [13] B. Saha and K. Goebel, "Battery data set," NASA AMES Prognostics Data Repository, 2007.
- [14] H. Zhang, X. Wang, J. Chen, C. Wang, and J. Li, "D2D-LSTM: LSTM-Based path prediction of content diffusion tree in Device-to-Device social networks," *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol.34, No.1, pp.295-302, 2020.
- [15] P. Zhao, A. Luo, Y. Liu, J. Xu, Z. Li, F. Zhuang, V. S. Sheng, and X. Zhou, "Where to go next: A spatio-temporal gated network for next POI recommendation," *IEEE Transactions on Knowledge and Data Engineering*, Vol.34, Iss.5, pp.2512-2524, 2022.



김 병 옥

<https://orcid.org/0000-0001-5755-7510>

e-mail : bwkim@dsu.ac.kr

2007년 고려대학교 컴퓨터교육과(학사)

2015년 고려대학교 컴퓨터교육학과

(석·박사통합)

2018년 ~ 2021년 동국대학교 컴퓨터공학과
조교수

2021년 ~ 2023년 동신대학교 컴퓨터공학과 조교수

2023년 ~ 현 재 강릉원주대학교 컴퓨터공학과 조교수

관심분야 : Artificial Intelligence, Machine Learning,
Rechargeable Battery



박 지 수

<https://orcid.org/0000-0001-9003-1131>

e-mail : jisupark@jj.ac.kr

2013년 고려대학교 컴퓨터교육과(박사)

2015년 ~ 2018년: 충남대학교 초빙교수

2018년 ~ 2019년: 경기대학교 교수

2019년 ~ 2020년: 동국대학교 연구교수

2020년 ~ 현 재 전주대학교 컴퓨터공학과 조교수

2020년 ~ 2021년 한국정보처리학회 이사

2021년 ~ 현 재 JIPS/HCIS Managing editor

2021년 ~ 현 재 한국정보처리학회 상임부회장

2021년 ~ 현 재 KTSDE 편집위원장

관심분야 : 분산 시스템, 모바일 클라우드 컴퓨팅, 클라우드
컴퓨팅, 사물인터넷, 빅데이터, e-Learning



장 홍 준

<https://orcid.org/0000-0001-9979-6851>

e-mail : hongjunjang@jj.ac.kr

2008년 고려대학교 컴퓨터교육과(학사)

2019년 고려대학교 컴퓨터학과(박사)

2019년~2021년 한국과학기술정보연구원

선임연구원

2021년~현재 전주대학교 컴퓨터공학과 조교수

관심분야: 데이터베이스, 데이터마이닝, 기계학습