J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr
Copyright ⓒ 2023 KSII

머신러닝 기반 수소 충전소 에너지 수요 예측 모델☆

Machine Learning-based hydrogen charging station energy demand prediction model

황 민 우[1] 하 예 림[1] 박 상 욱[1*]
MinWoo Hwang Yerim Ha Sanguk Park

요 약

수소 에너지는 높은 에너지 효율로 열과 전기를 생산하면서도 온실가스와 미세먼지 등 유해물질 배출이 없는 친환경 에너지로서, 전 세계적으로 탄소중립으로의 전환을 위한 핵심으로 주목받고 있다. 특히 스마트 수소에너지는 경제적이고 지속 가능하며, 안전한 미래 스마트 수소에너지 서비스로써 수소 에너지의 기반 시설이 디지털로 통합되어 '데이터' 기반으로 안정적으로 운영되는 서비스를 의미한다. 본 논문에서는 데이터 기반 수소 충전소 수요예측 모델 구현을 위해 강원도 내 설치되어 있는 수소 충전소 3곳(춘천, 속초, 평창)을 선정, 수소 충전소의 수요공급 데이터를 확보하였고, 머신러닝 및 딥러닝 알고리즘 7개를 선정하여 총 27종 입력 데이터(기상데이터+수소 충전소 수요량)로 모델을 학습하였고, 평균 제곱근 오차(RMSE)로 모델을 평가하였다. 이를 통해 본 논문에서는 최적의 수소 에너지 수요공급을 위한 머신러닝 기반 수소 충전소 에너지 수요 예측 모델을 제안한다.

☞ 주제어 : 스마트수소에너지, 탄소중립, 에너지 수요예측, 머신러닝, 수소 충전소

ABSTRACT

Hydrogen energy is an eco-friendly energy that produces heat and electricity with high energy efficiency and does not emit harmful substances such as greenhouse gases and fine dust. In particular, smart hydrogen energy is an economical, sustainable, and safe future smart hydrogen energy service, which means a service that stably operates based on 'data' by digitally integrating hydrogen energy infrastructure. In this paper, in order to implement a data-based hydrogen charging station demand forecasting model, three hydrogen charging stations (Chuncheon, Sokcho, Pyeongchang) installed in Gangwon-do were selected, supply and demand data of hydrogen charging stations were secured, and 7 machine learning and deep learning algorithms were used. was selected to learn a model with a total of 27 types of input data (weather data + demand for hydrogen charging stations), and the model was evaluated with root mean square error (RMSE). Through this, this paper proposes a machine learning-based hydrogen charging station energy demand prediction model for optimal hydrogen energy supply and demand.

☞ keyword : Smart hydrogen energy, carbon neutrality, energy demand forecasting, machine learning, hydrogen charging station

1. 서 론

2030 탄소중립 실현을 위해 필수적인 것은 에너지, 환경, ICT 등 모든 영역이 협력하여 배출되는 CO_2를 줄이는 일이다. 다수의 영역에서 CO_2 배출을 줄이는 것은 사실상 불가능한 상황이기 때문에 CO2 배출 비중이 가장 높은 에너지 분야에서 CO_2 감축은 매우 중요하다. 에너지 분야에서 에너지 절감을 위해서는 에너지가 낭비되거나 부족하여 생기는 에너지 효율 저하를 최소화하는 것이다. 한 가지 방법은 에너지의 수요와 공급을 일치시키는 것이다. 이를 위해서는 에너지 절감 서비스 구축을 통한 '데이터' 차원에서 에너지 관련 '수요 데이터' 및 '에너지 공급' 데이터를 수집하고 이를 예측하여 에너지 수요 & 공급을 일치시키는 것이다. 인공지능 기반 데이터 분석 및 예측 기술은 이러한 기술을 구현하는데 매우 중요한 기술 중 하나다. 본 논문에서는 순차 데이터 즉, 시계열 데이터를 분석 및 예측하기 위한 인공지능 기술을 소개하고 시계열 데이터의 전력 수요 & 공급 예측 정확도를 높이기 위한 예측모델을 제시한다.

본 논문에서는 에너지 분야에서 '수소 분야'에 대해서

1 Department of Electronic, Information & Communication Engineering, Kangwon National University, 25913, Korea
* Corresponding author (sanguk@kangwon.ac.kr)
[Received 03 February 2023, Reviewed 02 March 2023, Accepted 08 March 2023]
☆ 본 논문은 2022년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과임(2022RIS-005), 또한 2022년도 강원대학교 대학회계 학술연구조성비로 연구하였음.

다룬다. 최근 '에너지안보'가 대두됨에 따라 수소에너지는 높은 에너지 효율로 열과 전기를 생산하면서도 온실가스와 미세먼지 등 유해물질 배출이 없는 친환경 에너지로서, 전 세계적으로 탄소중립으로의 전환을 위한 핵심으로 주목받고 있다. 특히 스마트수소에너지 서비스는 경제적이고 지속 가능하며, 안전한 수소 에너지 서비스로 수소 신재생 에너지의 기반 시설이 디지털로 통합되어 '데이터' 기반으로 안정적으로 운영되는 서비스를 의미한다. 본 논문에서는 강원도 내 설치되어 있는 수소 충전소 3곳(춘천, 속초, 평창)을 선정하여 수소 충전소의 수요 공급 최적화를 위한 머신러닝 & 딥러닝 기반 수소 충전소 에너지 수요 예측 모델을 제안한다.

2. 관련 연구

본 논문은 수소 충전소의 에너지 수요 & 공급 최적화를 위한 머신러닝 & 딥러닝 기반 수소 충전소 에너지 수요 예측 모델 구현에 대한 논문이다. 현재 전기자동차 충전소의 수요를 예측하기 위한 기술은 많이 나오고 있다. 하지만 수소의 경우 수소충전소의 수요를 예측하기 위한 기술은 아직까지 미비한 상황이다.

논문 [1] 에서는 시계열 데이터 분석을 통하여 전기 에너지 수요를 예측하는 모형 및 방법에 대한 연구를 수행하였다. 이를 위하여 "A" 테크노파크에서 사용한 전력 사용량 시계열 데이터를 ARIMA 모형에 적용하였다. 확정된 모형을 약 2.5시간 정도의 시차에 대한 전력사용량 예측에 적용하였으며, 예측한 전력사용량의 평균상대오차와 평균제곱오차는 각각 2-5% 그리고 1-3kWh임을 보였다. 논문 [2]는 한국전력공사의 EV 충전 데이터를 이용하여 충전소 단위의 일별최대부하를 예측하는 LSTM(Long Short-Term Memory) 신경망 모델을 제안하였다. LSTM(Long Short-Term Memory)을 사용해 신경망 모델을 설계 및 개발한다. 성능평가는 지수평활법(exponential smoothing)과 SARIMA(Seasonal Auto Regressive Integrated Moving Average), TBATS를 사용하였다. 논문 [3]은 에너지 인터넷을 위한 GRU기반 전력사용량 예측을 제안하였다. GRU(Gated Recurrent Unit)을 기반으로 딥러닝 모델을 제안하고, 전력사용량 데이터를 토대로 예측 성능을 분석한다. 제안한 학습 모델의 예측 성능과 기존의 Long Short Term Memory (LSTM) 인공지능 네트워크 기반의 전력량 예측 성능을 비교하며, 성능평가는 MSE(Mean Squared Error), MAE(Mean Absolute Error), Forecast Skill Score, NMBE(Normalized Root Mean Squared Error)를 사용하였다. 논문 [4]는 매장 에너지 절감을 위한 LSTM 기반의 전력부하 예측 시스템 설계를 제안하였다. LSTM(Long Short-Term Memory) 기반의 전력 부하 예측 시스템을 설계했다. 관련 연구로는 순환신경망 알고리즘(RNN, Recurrent Neutral Network), 장단기메모리 알고리즘이 있다. 결측 데이터는 MICE를 통해 inputation을 거친 데이터 셋으로 분석을 진행하였다. 그리고 K-means 클러스터링을 통해 패턴화를 진행하였다. 논문 [5]는는 에너지인터넷에서 ID-CNN과 양방향 LSTM을 이용한 에너지 수요 예측 방법을 제안하였다. CNN (Convolution Neural Network)과 BLSTM (Bidirectional Long Short-Term Memory)을 융합한 1D-ConvBLSTM (1Dimention- Convolution and Bidirectional LSTM)을 제안하였다.

지금까지 제시한 에너지 수요 예측 기술은 대부분 전기 에너지를 대상으로 진행되었고, 이미 전기 에너지 기반 수요예측 기술은 구현되어 있다. 하지만 탄소중립을 위한 신재생에너지 중요성 증대로, 수소 에너지 활용에 대한 중요성이 증가하고 있고, 수소라는 에너지의 특성에 맞는 수요 예측 기술이 필요하다. 따라서 본 논문에서는 수소 충전소를 대상으로 수소 에너지 수요 예측 방법에 대한 기술을 제안한다. 다음은 본 논문에서 사용된 머신러닝 및 딥러닝 관련기술 개요를 나타낸다.

3. 머신러닝 관련 기술

3.1 머신러닝 관련 기술

의사 결정 트리는 다단계 또는 계층적 의사 결정 시스템 또는 트리와 같은 구조를 기반으로 한다. 트리는 루트 노드(모든 데이터 포함), 내부 노드 세트(분할) 및 터미널 노드 세트(리프)로 구성된다. 의사 결정 트리 구조의 각 노드는 하나의 클래스 또는 일부 클래스를 나머지 클래스와 구분하는 이진 결정을 내린다. 처리는 일반적으로 리프 노드에 도달할 때까지 트리 아래로 이동하여 수행된다. 이를 하향식 접근 방식이라고 한다[6].

랜덤 포레스트는 여러 개의 독립적이고 상관관계가 없는 결정 트리가 있는 앙상블 접근 방식이다. 이러한 상관관계가 없는 트리는 결정 트리에 임의성을 주입하여 모델이 정확한 일반화를 달성하는 데 도움이 된다[7].

극도로 무작위화된 트리(또는 추가 트리 알고리즘은 비교적 최근의 기계 학습 기술로 랜덤 포레스트 알고리즘의 확장으로 개발되었으며 데이터 세트에 과적합될 가

능성이 적다. 추가 트리(ET)는 다음과 같은 원리를 사용한다. 랜덤 포레스트 및 기능의 임의 하위 집합을 사용하여 각 기본 추정기를 훈련한다. 그러나 노드를 분할하기 위해 해당 값과 함께 최상의 기능을 무작위로 선택한다. ET는 전체 교육 데이터 세트를 사용하여 각 회귀 트리를 훈련한다. 반면에, RF는 부트스트랩 복제본을 사용하여 모델을 훈련한다[8].

Gradient Boosting Algorithm(GBM)은 회귀 또는 분류 분석을 수행할 수 있는 예측 모델로, 예측 모델의 앙상블 방법론의 부스팅 계열에 속하는 알고리즘이다. Gradient Boosting Algorithm은 테이블 형태의 예측에서 엄청난 성능을 보이며, 머신러닝 알고리즘 중에서 가장 높은 예측 성능을 가지는 것으로 알려져 있다[9].

Hist Gradient Boosting Algorithm은 Gradient Boosting의 주요 문제는 모델 훈련 속도 문제를 해결한다. 이것은 특히 수만 개의 행이 있는 대규모 데이터 세트에서 모델을 사용할 때 문제가 된다. 연속 입력 변수를 수백 개의 고유한 값으로 이산화하면 앙상블에 추가되는 트리의 훈련 속도를 크게 높일 수 있다. 이 기술을 구현하고 이 변환의 입력 변수에 훈련 알고리즘을 맞추는 그래디언트 부스팅 앙상블을 히스토그램 기반 그래디언트 부스팅 앙상블이라고 한다[10].

3.2 딥러닝 관련 기술

CNN 알고리즘은 첫 번째로 입력 데이터로 가중함수를 가지고 있는 kernal이 일정한 흐름으로 순회하며 다수의 합성곱을 병렬로 연산한다. 두 번째는 병렬로 연산한 값이 활성화 함수를 거치게 되며 입력 데이터에 대한 특징을 검출하여 feature map으로 출력된다. 세 번째는 Pooling 계층에서 Pooling 함수를 사용해 feature map에 검출된 특징 데이터를 줄여주는 작업을 진행한다. 마지막으로 CNN 계층과 Pooling 계층으로 추출한 데이터 세트를 Fully Connected 계층을 통해 행렬로 구성된 데이터를 열벡터 배열로 데이터 구조를 변경해 데이터 특징을 분류한다.

RNN이란 시계열 데이터와 같이 순차적인 정보를 처리하는 딥러닝 알고리즘으로, 입력값은 시간에 따라 순차적으로 입력층을 통해 은닉층으로 들어간다. 은닉층의 뉴런은 각각의 시점에서 바로 이전 시점의 은닉층의 뉴런에서 나온 값을 계속해서 자신의 입력으로 보내는 재귀적 활동을 한다. 이 두 값을 기반으로 입력이 모두 들어갈 때까지 출력값을 출력하는 과정이 반복된다.

LSTM 알고리즘은 RNN 알고리즘의 구조상 데이터양이 많으면, 과거의 시계열 데이터가 반영되지 않는 문제점을 해결하기 위해 개발된 알고리즘이다. RNN 알고리즘에서 은닉층을 정보 흐름을 제어하는 게이트 시스템인 망각 게이트, 입력 게이트,출 력게이트로 변형하여 오랜 과거 시계열 데이터를 현재의 출력층에 반영시킨다.

BRNN 알고리즘은 RNN 알고리즘의 한 종류로 기존의 RNN 알고리즘을 응용한 딥러닝 알고리즘이다. 입력계층, 포워드계층, 백워드계층, 출력계층으로 BRNN 알고리즘이 구성되며 포워드계층과 백워드계층은 RNN 알고리즘의 은닉층의 역할을 한다.

BLSTM 알고리즘은 BRNN 알고리즘과 LSTM 알고리즘을 결합한 딥러닝 알고리즘이다. BLSTM 알고리즘은 입력계층, 망각게이트, 입력게이트, 출력게이트로 변형한 포워드계층, 망각게이트, 입력게이트, 출력게이트로 변형한 백포워드계층, BLSTM 출력계층으로 구성된다.

본 논문은 3절에서 제시된 머신러닝 & 딥러닝 기술을 활용하여 스마트수소에너지 구현을 위한 수소 충전소 수요 예측 기술에 대해서 제안한다. 다음 4절은 에너지 수요 예측 모델 구현의 필요성에 대한 내용이다.

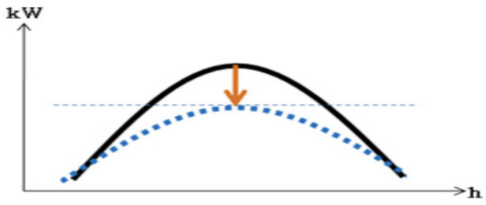
4. 에너지 수요예측모델 구현의 필요성

4.1 에너지 수요공급 관리 개념

수요공급 관리란 소비자의 에너지 사용 패턴을 변화시켜 최소의 비용으로 에너지 수요를 충족시키기 위한 모든 활동이며 공급 측의 관리와 대응되는 개념이다. 수요관리에는 크게 부하관리(Load Management)와 효율향상(Energy Efficiency)으로 나누어진다. 부하관리란 최대수요와 최저수요의 차이를 축소시켜 부하 평준화 도모와 공급설비 이용효율 향상이 주목적이다. 효율향상이란 고효율기기 보급으로 에너지이용 효율을 향상시켜 에너지를 절감하는 것이다. 에너지 수요공급의 핵심용어에는 최대수요 억제, 최대수요 이전, 기저부하 증대, 에너지 효율향상가 있다. 하단의 내용은 각 핵심용어에 대한 설명이다.

4.1.1 최대수요 억제(Peak Clipping)

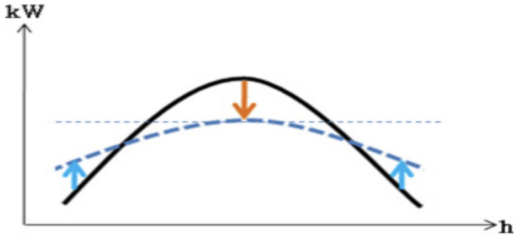
그림 1와 같이 계절별, 시차별 최대수요를 낮추는 것으로 가장 대표적인 수요관리 유형이다. 예를 들면 긴급절전, 건물냉반방 원격관리, 최대전력관리장치 보급 등이 있다[11].



(그림 1) 최대수요 억제 그래프
(Figure 1) Maximum Demand Suppression Graph

4.1.2 최대수요 이전(Peak Shifting)

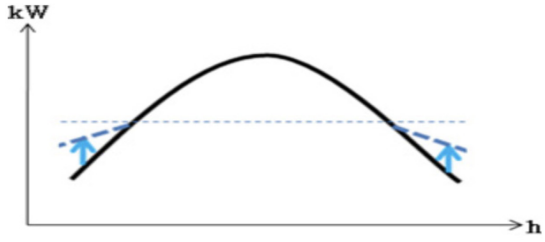
그림 2과 같이 피크시간대 에너지수요를 경부하시간대로 이전, 최대수요는 내리고 최저수요는 올리는 방식을 사용한다[11].



(그림 2) 최대수요 이전 그래프
(Figure 2) Peak Demand Conversion Graph

4.1.3 기저부하 증대(Valley Filling)

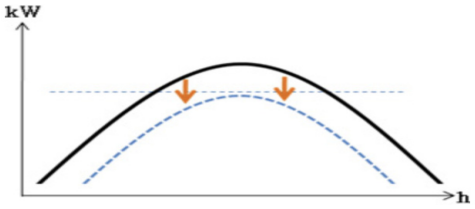
그림 3과 같이 경부하 시간대의 수요를 증대시켜 설비이용률을 높이고 판매전력량을 증대시키는 방식이다[11].



(그림 3) 기저부하 증대 그래프
(Figure 3) Baseload Growth Graph

4.1.4 에너지 효율향상(Energy Efficiency)

그림 4과 같이 고효율기기 보급, 절전정보 제공 등을 통한 전기이용효율 향상으로 에너지를 절약하는 방식이다[11].



(그림 4) 에너지 효율향상 그래프
(Figure 4) Energy Efficiency Improvement Graph

4.2 에너지 수요예측모델 구현의 필요성

에너지 수급 특히, 전기에너지는 생산과 동시에 소비가 이루어지는 물리적 특성을 갖고 있다. 이는 안정적인 전력공급을 위해서는 사전에 전력수요를 정확하게 예측해야함을 의미한다. 전력수요 예측은 설비투자, 수급안정, 구입전력비에 직결되는 중요한 사안이며 이는 국가경제에 큰 영향을 미친다. 단기적인 측면에서 전력수요를 과다하게 예측하면 전력시장에서의 가격이 상승하게 되고, 수요관리 비용이 상승하게 된다. 반대로 전력수요를 과소하게 예측하면 전력수급 불안이 야기되고 이는 전력시장에서 부가적인 제약발전 정산비용이 상승하게 되는 악영향을 불러일으킨다. 이를 통해 우리는 전력수요를 정확하게 예측하는 것이 얼마나 중요한지 알 수 있다.

최근 전력수요의 경우 경제성장과 더불어 지속적으로 증가하고 있고 여름철과 겨울철의 전력 성수기의 최대수요는 변동폭 및 오차율이 커지고 있다. 최근 전력수요와 같은 경우 주변 환경이 기존의 예상과는 다르게 변화되면서 수요 예측의 불확실성이 커지고 있다. 따라서 보다 정확한 전력수요를 예측하기 위해서는 새로운 전력수요 예측방법론을 찾아 변화하는 환경에 적응할 필요가 있다.

에너지 수요는 특히 기상상황의 변화와 밀접한 상관관계를 가지고 있다. 특히 최근 들어 냉·난방기기 보급의 지속적 확대에 따른 전력수요의 기상 민감도는 크게 증가하고 있으며 지구 온난화로 인해 기후패턴 변화에 따라 전력수요 예측의 어려움이 더욱 커지고 있는 추세이

다. 특히 에너지 수요에 영향을 주는 요인인 기온, 습도, 풍속, 강수, 일조 등을 활용하여 다양한 상관관계에 대한 분석을 통해 에너지수요예측 정확도를 높일 필요가 있다.

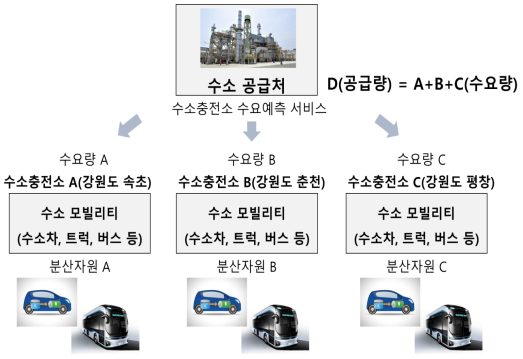
5. 수소 충전소 에너지 수요 예측 모델 구현

지금까지 제시한 에너지 수요 예측 기술은 대부분 전기 에너지를 대상으로 진행되고 있고, 2절에서 제시한 바와 같이 이미 전기 에너지 기반 수요예측 기술은 구현되어 있다. 하지만 탄소중립을 위한 신재생에너지 중요성 증대로, 수소 에너지 활용에 대한 중요성이 증가하고 있고, '수소' 라는 에너지의 특성에 맞는 수요 예측 기술이 필요하다. 이를 위해서는 전기 에너지와 수소 에너지의 특성을 잘 파악해야 한다.

(표 1) 수소 충전소 수요/공급 현황 (4월)
(Table 1) Demand/supply status of hydrogen charging station

일자	구입량(kg)	충전대수(대)	충전량(kg)
2022-04-01	222.99	15	65.977
2022-04-02		27	105.764
2022-04-03	219.301	16	62.695
2022-04-04		3	13.473
2022-04-05		10	40.441
2022-04-06		8	32.27
2022-04-07		12	50.646
2022-04-08		20	86.203
2022-04-09	213.32	38	156.133
2022-04-10		24	81.209
2022-04-11	190.443	8	34.29
2022-04-12		8	32.287
2022-04-13		8	34.982
2022-04-14		7	28.645
2022-04-15	197.807	15	61.09
2022-04-16		22	90.419
2022-04-17		17	64.7
2022-04-18		9	33.983
2022-04-19	224.66	12	50.318
2022-04-20		6	20.489
2022-04-21		16	63.749
2022-04-22		26	89.044
2022-04-23	230.082	18	51.987
2022-04-24		19	80.587
2022-04-25		7	23.96
2022-04-26		10	45.689
2022-04-27		13	57.113
2022-04-28	201.998	13	55.583
2022-04-29		16	67.729
2022-04-30		24	67.877

전기 자동차 충전소와 수소 자동차 충전소를 예를 들면, 전기 에너지는 실시간으로 전력 그리드를 통해 공급되는 전기를 충전소에서 충전하게 된다. 하지만 수소 에너지는 실시간으로 공급되는 수소 그리드가 존재하지 않는다. 생산된 수소는 수소 탱크로 공급받아 지역마다 분산되어 있는 수소 충전소에 저장하게 된다. 이렇게 저장된 수소를 수소 충전소에서 각 사용자들이 공급받아 사용하게 된다. 그렇기 때문에 각 충전소마다 소요되는 정확한 충전량(수요량)을 예측하지 못하면 수소를 제때 공급하지 못하는 블랙아웃 현상(표 2.)이 발생하게 된다. 본 논문에서는 이러한 현상을 방지하기 위해 기상데이터를 활용한 수소충전소 수요 예측 방법을 제안하였고, 그림 6 표 1과 같이 강원도 지역(속초, 춘천, 평창)의 3개의 도시에서 발생한 7개월 분량 수소 충전소 데이터를 활용하였고, 수소 충전 데이터와 기상 데이터를 활용하여 머신러닝 & 딥러닝 기반으로 학습된 모델을 통해 일별 수소 충전량을 예측하였다.



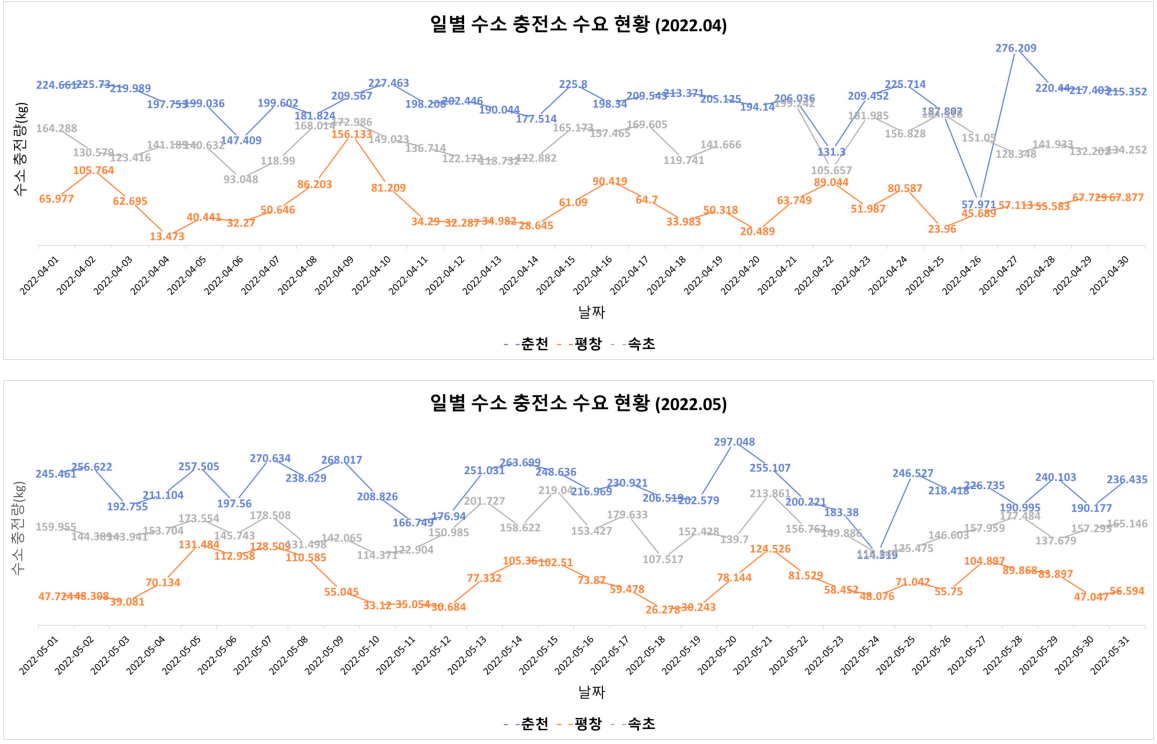
(그림 5) 수소 충전소 수요예측 서비스 개략도
(Figure 5) Schematic diagram of hydrogen charging station demand forecasting service

그림 5와 같이 이는 수소 공급처(수소를 생산하고 공급하는 발전소 등의 사업자)와 수소 수요처(수소 충전소) 간의 정확한 수요를 예측하여, 에너지가 부족하거나 낭비되는 현상을 제거하여 에너지 사용 효율을 극대화하기 위한 모델이다.이를 위해서는 과거에 거래된 각 충전소의 수소 충전 내역 데이터가 필요하다. 강원도 지역(속초, 춘천, 평창) 내에 위치한 수소 충전소에서 7개월 분량의 수

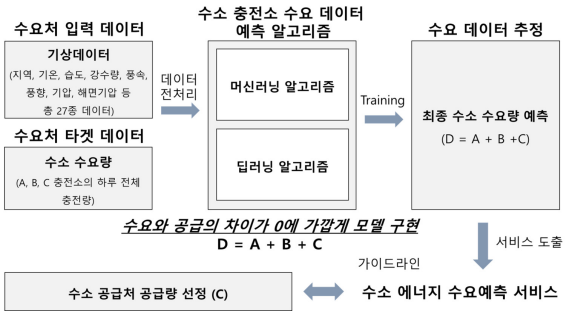
(표 2) 재고 소진으로 인한 블랙아웃 현황 (8월)
(Table 2) Blackout status due to out of stock (August)

일자	구입량(kg)	충전량(kg)	비고
2022-08-01	51	48.594205	재고소진 운영중단
2022-08-02	0	0	재고소진 운영중단
2022-08-03	0	0	재고소진 운영중단
2022-08-04	120	126.71761	수소교체
2022-08-05	61	65.201932	
2022-08-06	23	22.579091	재고소진 운영중단
2022-08-07	0	0	재고소진 운영중단
2022-08-08	79	83.631136	수소교체
2022-08-09	68	69.391023	
2022-08-10	43	45.721023	
2022-08-11	9	6.3379545	재고소진 운영중단
2022-08-12	3	3.115	재고소진 운영중단
2022-08-13	115	120.2833	수소교체

소 충전량 데이터를 확보하였고, 본 수소 충전량 데이터를 타겟 데이터로 사용하였다. 그리고 기상청으로부터 약 27종의 기상 데이터(지역, 기온, 습도, 압력 등 27종 데이터)를 확보하여 기상데이터와 수소 충전소 충전량 간의 어떠한 연관성이 있는지 분석하였다. 확보한 데이터를 통해 그림 7과 같이 머신러닝 및 딥러닝 모델을 학습하여 타겟데이터(하루 총 수소 자동차 충전량)을 예측하는 수소 충전소 하루 수요량 예측 모델을 구현하였다. 표 3은는 입력된 기상데이터 및 수소 충전 데이터 현황을 나타낸다. 데이터 결측치는 0으로 입력하였고 또는 불필요한 데이터는 제거하여 데이터 전처리를 진행하였다.



(그림 6) 강원도 (춘천, 평창, 속초) 일별 수소 충전소 수요 현황 (4월~5월, 지면 관계상 6~10월은 생략)
(Figure 6) Gangwon-do (Chuncheon, Pyeongchang, Sokcho) Demand for Hydrogen Charging Stations by Day (April-May)

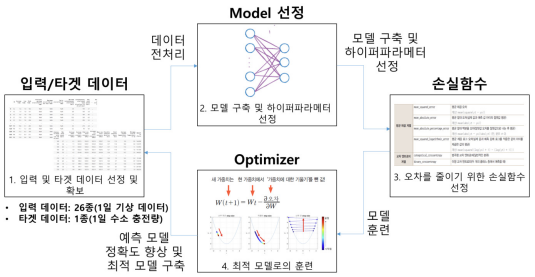


(그림 7) 수소 충전소 데이터 수요예측 서비스 구현 방안
(Figure 7) Implementation method for hydrogen filling station data demand forecasting service

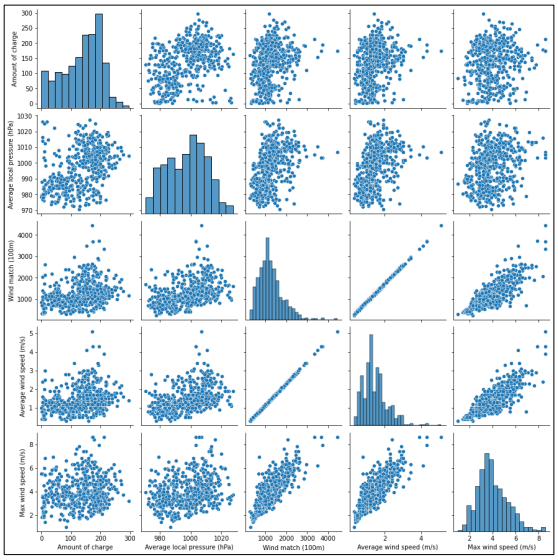
(표 3) 입력 데이터 종류
(Table 3) Input data type

구분	데이터 종류
기상 데이터 (입력 데이터)	• 지역(춘천, 속초, 평창) • 일별 평균 기온 • 최저 기온 • 최대 기온 • 10분 최다 강수량(mm) • 1시간 최다강수량(mm) • 일강수량(mm) • 최대 순간 풍속(m/s) • 최대 순간 풍속 풍향(16방위) • 최대 풍속(m/s) • 최대 풍속 풍향(16방위) • 평균 풍속(m/s) • 풍정합(100m) • 최다풍향(16방위) • 평균 이슬점온도(°C) • 최소 상대습도(%) • 평균 상대습도(%) • 평균 증기압(hPa) • 평균 현지기압(hPa) • 최고 해면기압(hPa) • 최저 해면기압(hPa) • 평균 해면기압(hPa) • 가조시간(hr) • 합계 일조시간(hr) • 평균 지면온도(°C) • 최저 초상온도(°C)
수소 충전소 데이터 (타겟 데이터)	• 일별 수소 충전량

데이터는 인공지능 모델 훈련을 위해 입력 데이터와 타겟 데이터로 분류되었다. 모델 구현 및 훈련은 대표적인 클라우드 기반 툴인 구글 Colaboratory를 사용하였고, 27개의 입력데이터와 타겟 데이터 간의 상관관계를 도출하여 타겟 데이터와 가장 상관도가 높은 5개 데이터를 추출하여(그림 9) 상관도 그래프를 작성하였다. 타겟 데이터인 수소 충전량 데이터와 가장 상관도가 높은 데이터는 평균 현지기압(hPa), 풍정합(100m), 평균 풍속(m/s), 최대 풍속(m/s)으로 도출되었다.



(그림 8) 머신러닝 및 딥러닝 기반 예측 모델 구현 프로세스
(Figure 8) Machine learning and deep learning-based predictive model implementation process



(그림 9) 데이터 연계성 분석
(Figure 9) Data linkage analysis

그림 8는 머신러닝 및 딥러닝 기반 예측 모델 구현 프로세스를 나타낸다. 아래 리스트는 모델 구현에 선정된 인공지능 알고리즘 리스트를 나타낸다.

- 머신러닝 알고리즘:
  - 선형회귀(Linear Regression),
  - 결정트리(Decision Tree Regressor),
  - 랜덤포레스트(Random Forest Regressor),
  - 엑스트라트리(Extra Trees Regressor),
  - 그레디언트부스팅(Gradient Boosting Regressor),
  - 확률적 그레디언트부스팅 (Hist Gradient Boosting Regressor)

- 딥러닝 알고리즘
  - Deep Neural Network (DNN): 9 Hidden Layer, 총 215,441개 파라메터 개수)

- 성능 평가: 평균 제곱근 오차
  - Root Mean Square Error (RMSE) =

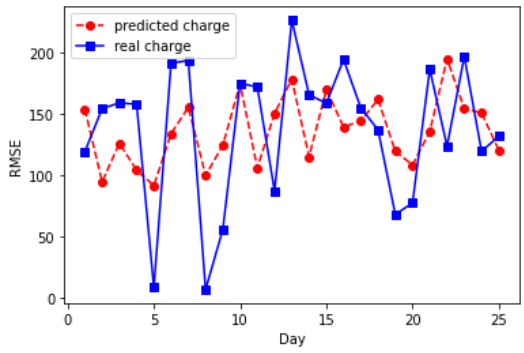
$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Real\ Data - Forecased\ Data)^2}$$

표 4는 성능 평가 결과를 나타낸다.

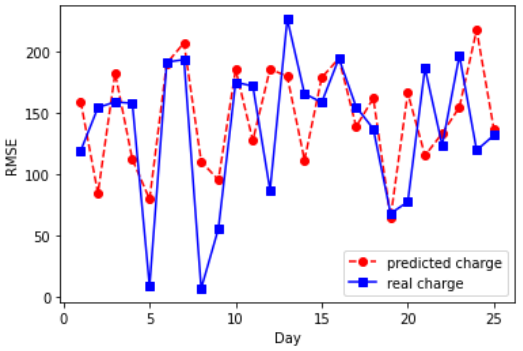
(표 4) RMSE 성능 평가 결과
(Table 4) RMSE performance evaluation results

구분	RMSE
선형회귀 (Linear Regression),	50.91
결정트리 (Decision Tree Regressor)	77.26
랜덤포레스트 (Random Forest Regressor)	55.10
엑스트라트리 (Extra Trees Regressor)	54.69
그레디언트부스팅 (Gradient Boosting Regressor)	53.35
확률적 그레디언트부스팅 (Hist Gradient Boosting Regressor)	52.56
DNN	51.61

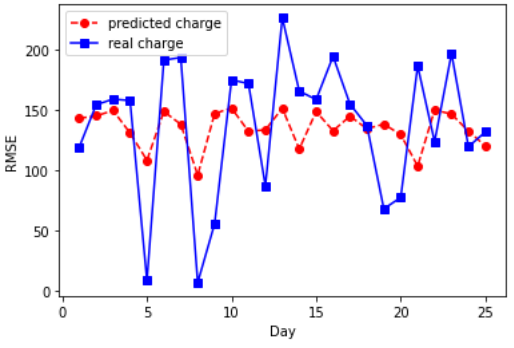
다음 그림 10, 11, 12는 총 25일 동안 수소 충전소 충전량 예측 그래프를 나타낸 그래프이고 낮은 오차(RMSE)를 산출한 3개의 모델을 선정하였다.



(그림 10) 선형회귀 예측 그래프 (50.91 RMSE)
(Figure 10) Regression Prediction Graph



(그림 11) 확률적 그레디언트 부스팅 예측 그래프 (52.56 RMSE)
(Figure 11) Stochastic Gradient Boosting Prediction Graph



(그림 12) DNN 예측 그래프 (51.61 RMSE)
(Figure 12) DNN prediction graph

오차(RMSE)가 낮은 모델 3개를 선정하면 선형회귀 모델, DNN 모델, 확률적 그레디언트 부스팅 모델이라 할 수 있다. RMSE는 평균 제곱근 오차로써 실제 데이터와 예측된 데이터의 오차를 평가할 때 쓰이는 성능 평가 방법이다. RMSE는 단위가 %가 아니기 때문에 0~100% 사이의 숫자값에 의해 평가되지 않고, 절대적인 값에 의해 평가된다. 즉 본 논문에서 측정된 50.91~77.26 RMSE 수치는 오차를 나타내기 위한 절대적인 수치이고, 실험 대상에 따라 값이 클수도 있고, 작을수도 있다. 이를 통해 보았을 때 가장 오차가 낮은 모델은 선형 회귀모델로써 머신러닝 모델에서 가장 기본적인 모델이라 할수 있다. 가장 기본적인 모델임에도 불구하고 가장 낮은 오차를 보이는 높은 성능을 보이는 이유는 모델 훈련에 사용한 데이터의 양(7개월 분량의 일별 데이터)이 많지 않고, 데이터의 복잡성 또한 낮기 때문이다. 향후 데이터의 양이 증가하고 데이터의 종류가 다양해지면 예측 성능은 더욱 높아질 것으로 예상되고, 그때의 가장 높은 성능을 보이는 모델은 DNN 모델이 될 것이다. 왜냐하면 딥러닝 모델은 복잡하고 다수의 데이터 체계에서 강한 성능을 보이기 때문이다.

5. 결 론

스마트 수소 에너지 서비스를 실질적으로 구현하기 위해서는 다수의 데이터, 다수의 데이터 종류 및 이에 따른 의미있는 데이터 체계가 필요하다. 데이터의 확보에 따라 정확도 높은 시계열 기반의 에너지 수요예측 모델을 구현할 수 있다. 수소 관련 에너지 수요는 기상의 영향을 많이 받을 수 있다는 가정하에, 본 연구에서는 풍속, 습도, 일조량 등의 기상데이터를 수소 에너지 수요 예측 모델 구현에 활용하였다. 특히 정확도 높은 에너지 수요예측 모델 구현을 위해서는 다수의 기상 데이터 + 에너지 수요 데이터 및 정확도 높은 시계열 기반의 모델이 필요하다.

본 연구에서는 스마트 수소 에너지 서비스를 구현하기 위해 정확도 높은 시계열 데이터 기반 수소 에너지 수요 예측 모델을 구현하여 에너지 운영 효율화 달성을 위한 연구를 진행하였다. 이를 통해 향후 2030 탄소중립을 실현하기 위한 초석을 다지고자 한다.

참고문헌(Reference)

[1] J. K. Chang-Yong Lee, "A Study on the Short-term Prediction of Power Consumption by using the ARIMA Model," Journal of The Korean Data Analysis Society (JKDAS), vol. 19, pp. 1349-1362, 2017. http://doi.org/10.37727/jkdas.2017.19.3.1349

[2] B. L. Haesung Lee, Hyun Ahn, "An LSTM Neural Network Model for Forecasting Daily Peak Electric Load of EV Charging Stations," Korean Society for Internet Information, vol. 21, no. 5, pp. 119-127, 2020. https://doi.org/10.7472/jksii.2020.21.5.119

[3] Y.-G. S. Dong-gu Lee, Is-sac Sim, Yu-Min Hwang,Sooh-wan Kim, Jin-Young Kim, "Prediction of Power Consumptions Based on GatedRecurrent Unit for Internet of Energy," Journal of IEEE Korea Council, vol. 23, no. 1, pp. 120-126, 2019. https://doi.org/10.7471/ikeee.2019.23.1.120

[4] Y. S. Jongseok Choi, "LSTM-based Power Load Prediction System Design for Store Energy Saving" Korea Information Electron Communication Technology, vol. 14, no. 4, pp. 307-313, 2021. https://doi.org/10.17661/jkiiect.2021.14.4.307

[5] Y. G. S. Ho Cheul Jung, Donggu Lee, Soo Hyun Kim Yu Min Hwang, Issac Sim, Sang Keun Oh, Seung-Ho Song and Jin Young Kim, "Prediction for Energy Demand Using 1D-CNN andBidirectional LSTM in Internet of Energy," Institute of Korean Electrical and Electronics Engineers, vol. 23, no. 1, pp. 134-142, 2019. https://doi.org/10.7471/ikeee.2019.23.1.134

[6] M. Xu, P. Watanachaturaporn, P. K. Varshney, and M. K. J. R. S. o. E. Arora, "Decision tree regression for soft classification of remote sensing data," vol. 97, no. 3, pp. 322-336, 2005. https://doi.org/10.1016/j.rse.2005.05.008

[7] Z. El Mrabet, N. Sugunaraj, P. Ranganathan, and S. J. S. Abhyankar, "Random forest regressor-based approach for detecting fault location and duration in power systems," vol. 22, no. 2, p. 458, 2022. https://doi.org/10.3390/s22020458

[8] V. John, Z. Liu, C. Guo, S. Mita, and K. Kidono, "Real-time lane estimation using deep features and extra trees regression," in Image and Video Technology: 7th Pacific-Rim Symposium, PSIVT 2015, Auckland, New Zealand, November 25-27, 2015, Revised Selected Papers 7, Springer, pp. 721-733, 2016.
https://doi.org/10.1007/978-3-319-29451-3_57

[9] J. Brownlee. "A Gentle Introduction to the Gradient Boosting Algorithm for Machine Learning."
https://machinelearningmastery.com/gentle-introduction-gradient-boosting-algorithm-machine-learning/ (accessed 2023.01.31).

[10] J. Brownlee. "Histogram-Based Gradient Boosting Ensembles in Python."
https://machinelearningmastery.com/histogram-based-gradient-boosting-ensembles/ (accessed 2023.01.31.).

[11] "Demand Management System Guide – Demand Management Techniques and Types", Korea Electric Power Corporation
https://home.kepco.co.kr/kepco/CY/K/htmlView/CYKAHP001.do?menuCd=FN0207/ (accessed 2023.03.09)

◐ 저 자 소 개 ◑

황 민 우(MinWoo Hwang)
2021년~현재 강원대학교 정보통신공학과(공학사)
관심분야 : 데이터베이스, 인공지능
E-mail : has0327@naver.com

하 예 림(Yerim Ha)
2021년~현재 강원대학교 정보통신공학과 재학 중
관심분야 : 데이터베이스, 인공지능
E-mail : dlalqnwk1081@naver.com

박 상 욱(Sanguk Park)
2009년 중앙대학교 전자전기공학부(공학사)
2017년 중앙대학교 전자전기공학과(공학박사)
2017년~2022년 중앙대학교 산학협력단 연구전담교수
2022년~현재 강원대학교 전자정보통신공학과 조교수
관심분야 : 스마트수소에너지서비스, 인공지능, 디지털트윈
E-mail : sanguk@kangwon.ac.kr