

다층 퍼셉트론(MLP)을 기반으로 한 전기로 출강 시점의 탄소 성분 예측 모델 개발

하혁준*, 신동현**, 김정우*, 조경제*, 가승호*, 김정현***, 이권영*

*한동대학교 기계제어공학부

**한동대학교 ICT 창업학부

***한동대학교 AI 융합교육원

e-mail:kylee@handong.edu

Multi-Layer Perceptron Based Model to Predict Carbon Content in Molten Steel at Tapping of Electric Arc Furnace

Hyukjun Ha*, Dong-hyen Shin**, Jungwoo Kim*, Kyoungje Cho*, Seung-ho Ga*, Junghyun Kim***, Kwon-Yeong Lee*

*School of Mechanical and Control Engineering, Handong Global University

**School of Global Entrepreneurship and ICT, Handong Global University

***School of Applied Artificial Intelligence, Handong Global University

요약

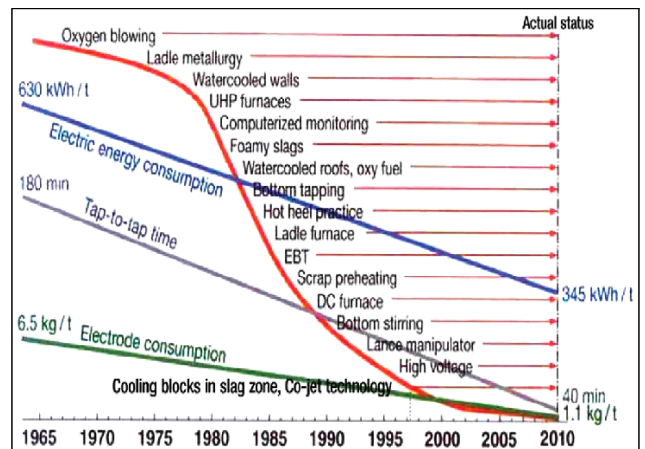
본 논문에서는 머신러닝 모델의 일종인 다층 퍼셉트론(MLP) 모델을 통하여 전기로 승온기 조업 중 용강 내의 탄소 함량을 예측하는 모델을 개발하였다. 모델 개발을 위해 데이터 수집, 데이터 전처리, 변수 선정, 다층 퍼셉트론 모델링을 차례로 진행하며 보다 효율적으로 전기로 공정을 관리하는 데이터 기반의 접근 방식을 제안하였다. 또한, 노드 개수에 따른 MLP 모델 결과를 비교 분석하여 최종적으로 노드 개수가 80개일 때, 검증 R^2 score가 0.877, 검증 RMSE가 2.291을 나타내는 모델을 생성하였다.

1. 서론

전기로 공정은 제강 공정 중의 하나로, 전 세계 철강 생산의 28% 비중을 담당하고 있다. 철강 산업은 전 세계 온실가스 배출량의 7%를 차지할 만큼 많은 양의 온실가스를 발생시킨다[1]. 파리 기후 협약 이후로 전 세계적으로 탄소 배출량을 줄이고자 하는 노력을 많이 하고 있고, 철강 산업 또한 마찬가지이다. 철강 산업에서 최근 전기로는 스크랩 재활용으로 인한 친환경적인 인식이 있으며, 탄소 절감 이슈에 대한 대응책으로 크게 주목받고 있다. 그러나 전기로는 용해의 주 열원으로 전기를 사용하기 때문에 많은 에너지를 필요로 한다. 이에 따라 전기로의 전력 감축에 대한 고민은 수십 년에 걸쳐 계속되었다. 그동안의 기술 개발에 따라 전기로의 성능은 크게 향상되었다. [그림 1]을 보면, 1965년대에는 1톤의 철강을 생산하는 데 630kWh의 전기에너지가 필요했지만, 2010년대에는 345kWh로 크게 줄었다. 또한, 고철을 로(Furnace)에 투입한 후 출강할 때까지 걸리는 시간을 측정하는 Tap-to-Tap Time(TTT)는 1965년대 180분에서 2010년대 40분으로 크게 단축되었다. 이처럼 예열 장치의 설치, 산소 가스 버너의 출력

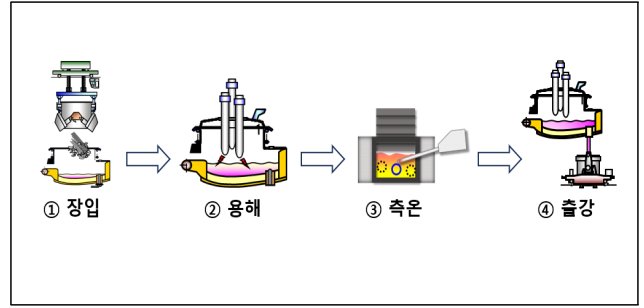
증대, 산소 및 탄소 송풍 랜스의 도입 등 전기로의 효율을 향상시키는 혁신적인 기술들이 전반적으로 도입되며 수십 년 동안 전기로의 비용과 에너지 사용량은 감소해 왔다.

그러나, 2010년 이후에는 전기로의 혁신이 포화상태에 이르러 전기로의 효율성과 생산성을 개선하기 위한 큰 틀에서의 일반적인 솔루션의 변화는 이루어지지 않았다. 이에 따라 연구자들은 전반적인 솔루션보다는 각각 개별의 전기로 특성에 맞는 솔루션을 제공하여 철강 생산 비용을 줄이고자 노력하였다.



[그림 1] 전기로 에너지 사용량 변천사

본 연구는 머신러닝 기반의 모델을 통하여 전기로 조업 중 용강 내의 탄소 함량을 예측하여 측온 과정을 최소화하고, 이를 통해 전력 소비를 감축하려는 목표를 가지고 있다. 기존 선행 연구들과의 주요 차이점은, 본 연구가 1차 측온 이후 탄소 함량의 정밀한 예측을 통해 승온기 과정에서 전기로의 운영을 최적화하려는 점에 있다. 이는 전력 소비의 효율적 관리 뿐만 아니라 환경적 영향을 줄이는 데에도 중요한 역할을 할 것으로 기대된다. 또한, 본 연구는 특정 머신러닝 알고리즘의 적용을 통해, 전기로의 복잡한 조업을 보다 정밀하게 모델링하고자 한다. 이는 기존 연구들이 주로 집중하였던 열역학적 접근이나 프로세스 최적화와는 다른, 데이터 기반의 접근 방식이다. 따라서, 본 연구의 접근 방식은 전기로의 운영을 보다 효율적으로 관리할 수 있는 새로운 가능성을 제시함과 동시에, 에너지 소비 및 탄소 배출 감소에 기여할 것으로 기대된다. 본 연구에서 타겟으로 삼은 전기로는 현대제철 포항공장의 AC타입, EBT 출강 방식의 120톤급 전기로이다.



[그림 2] 전기로 공정 진행 과정

이를 통해, 조업자들이 최적의 승온기 작업을 할 수 있도록 도와 측온 과정을 최소화하여 투입 전력에너지 및 탄소량을 감축하는 것 목적으로 한다. 효과적인 모델 개발을 위해 데이터 수집, 데이터 전처리, 변수 선정, 모델링, 탄소 예측의 과정 순으로 연구를 진행하였다.

2.2 데이터 수집 및 전처리

2.2.1 데이터 수집

데이터 수집은 현대제철 포항공장의 120톤 전기로 조업 데이터 관리 시스템을 통해 이루어진다. 수집한 데이터는 2023년 1월부터 2024년 2월 사이에 조업한 특수강 2761개 히트에 대한 조업 데이터이다. 조업 데이터는 4개의 그룹으로 분류할 수 있다. 첫 번째로, 센서데이터는 조업 중 센서를 통해 확보하며, 10초에 한 번씩 연속적으로 수집된다. 두 번째, 투입데이터들은 조업 중 스크랩과 부원료를 전기로에 투입한 시점에 수집되는 불연속적인 데이터로, 투입된 스크랩과 부원료의 정보를 담고 있다. 세 번째로, 측온데이터는 측온 시에 측온 센서를 통해 수집되는 불연속적인 데이터로, 용강 온도, 산소 농도, 탄소 함량에 관한 정보가 수집된다. 마지막으로, 가공 데이터는 추가적으로 필요한 정보를 얻기 위해, 앞의 세 종류의 데이터를 가공하여 새로 생성한 데이터이다.

2.2.2 도메인 지식 기반 데이터 필터링

정확하고 신뢰할 수 있는 모델을 만들기 위해서는 데이터 필터링을 통하여 모델 학습에 사용되는 데이터의 품질을 높일 필요가 있다. 따라서 본 연구의 전처리 단계에서는 전기로에서 추출한 raw 데이터를 필터링하는 과정을 거쳤다. 단순한 통계적 이상치 처리 방식으로는 조업의 복잡성을 반영할 수 없기에, 우선 아래의 [표 1]과 같이 도메인 지식 기반의 데이터 필터링을 먼저 수행하였다. 총 2761개의 히트 중 이 과정에서 1470개의 히트를 제거하였다.

2. 본론

2.1 방법론

본격적으로 방법론을 설명하기 이전에 전기로 공정에 대한 이해가 필요하다. 여러 철 스크랩을 수집하여 상품성이 있는 철을 생산하기까지의 과정 중 하나인 전기로 공정에서는 스크랩을 녹이고, 생산할 강종의 종류에 따라 대략적인 탄소 함량을 맞추는 작업을 진행하게 된다. 구체적인 조업 과정은 [그림 2]와 같다. 첫 번째로, 수집한 스크랩을 녹이기 위해 전기로 로체에 담은 장입 과정을 진행한다. 이후에는 3개의 전극봉을 활용하여 전력 에너지를 투입해 스크랩을 용해시킨다. 이때, 장입과 용해 과정은 스크랩의 파괴로 인해 2차 또는 3차로 나누어 진행한다. 스크랩의 용해가 어느 정도 완료되면 가탄량을 급격하게 증가시키며 슬래그포밍을 진행한다. 이후, 조업자들의 노하우로 용강 내부의 탄소 함량이 목표 기준에 도달했다고 생각되면 측온을 진행하게 되며, 이 과정에서 용강온도, 산소 농도, 탄소 함량을 확인한다. 목표한 강종별 기준에 도달하면 출강을 하고, 그렇지 않은 경우에는 기준에 맞추기 위한 추가적인 조업을 진행하고, 다시 측온하여 기준 도달 여부를 다시 확인한다.

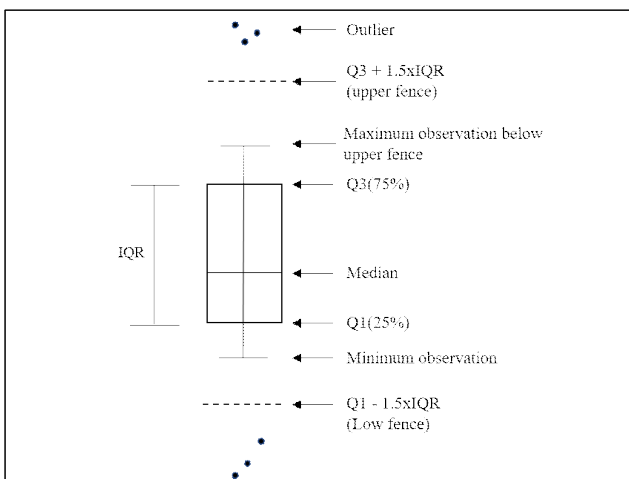
현재 현대제철 포항공장에서는 특수강 조업 중 50% 이상의 히트에서 3회 이상 측온을 진행하고 있다. 따라서, 본 연구는 1차 측온 정보와 이후 승온기 조업 과정 데이터를 활용하여 2차 측온 시점의 탄소 함량을 예측하는 다층 퍼셉트론 기반의 모델을 개발하고자 한다.

[표 1] 도메인 지식 기반 데이터 필터링 과정

순서	전처리 과정	제거 히트수
1	센서데이터 이상 히트 제거	459
2	조업 중 특이사항 발생 히트 제거	161
3	히트 시작시점 및 종료시점 이상 히트 제거	60
4	조업시간 이상 히트 제거	17
5	누적전력사용량 이상 히트 제거	17
6	측은 시점 이상 히트 제거	122
7	측은 데이터 오류 발생 히트 제거	628

2.2.3 Interquartile Range(IQR) 기법

도메인 지식 기반의 데이터 필터링을 거친 이후 남은 1291개의 히트에 대하여 Interquartile Range(IQR) 이상치 탐지 기법을 활용하여 더욱 정제된 데이터를 얻고자 하였다. 아래의 [그림 3]은 IQR에 대한 전반적인 설명을 담고 있다. IQR 이상치 탐지 기법에서는 차례로 쭉 나열된 데이터들의 중간 50%, 즉 25~75% 사이에 있는 값들을 IQR로 규정한다. 25%인 지점을 Q1, 75%인 지점을 Q3로 표현하며, IQR을 Q3-Q1으로 표현할 수 있다. 이때, $Q1-1.5 \times IQR$ 인 Low fence 지점과 $Q3+1.5 \times IQR$ 인 Upper fence 지점의 범위를 벗어나는 값들을 이상치로 규정하여 제거한다. 모든 변수에 대해 이 과정을 진행하였으며, 이 과정에서 총 1291개의 히트 중 892개의 히트가 제거돼 최종적으로 399개의 히트를 모델링에 활용하였다.



[그림 3] IQR 이상치 탐지

2.3 변수 선정

도메인 지식을 기반으로 [표 2]와 같이 2차 측은 시점의 탄소 함량에 영향을 줄 수 있는 변수들을 선정하였다.

[표 2] 최종 선정 변수

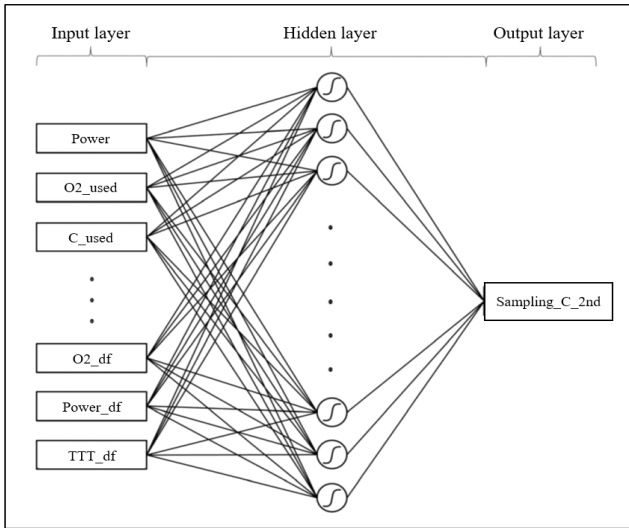
데이터 그룹	변수명	변수 설명
센서 데이터	Power	1차 측은 시점의 누적 전력 투입량
	O2_used	1차 측은 시점의 누적 산소 투입량
	C_used	1차 측은 시점의 누적 탄소 투입량
	Air_in	1차 측은 시점의 공냉덕트 입구 온도
	Air_out	1차 측은 시점의 공냉덕트 출구 온도
투입 데이터	C_comp	투입된 스크랩 내의 C 성분 함유량
	Si_comp	투입된 스크랩 내의 Si 성분 함유량
측은 데이터	Sampling_C	1차 측은 시점의 탄소 함량
	Sampling_O2	1차 측은 시점의 산소 농도
	T_iron_out	1차 측은 시점의 용강 온도
	Sampling_C_2nd	2차 측은 시점의 탄소 함량
가공 데이터	Power_pred	투입된 스크랩의 화학 성분을 기반으로 한 예상 전력 사용량
	C_df	1차 측은 시점 이후 탄소 추가 투입량
	O2_df	1차 측은 시점 이후 산소 추가 투입량
	Power_df	1차 측은 시점 이후 전력 추가 투입량
	TTT_df	1차 측은 시점부터 2차 측은 시점까지의 조업 시간

센서데이터 그룹에서는 전기로 내부 스크랩의 용해 정도와 성분에 영향을 줄 수 있는 Power, O2_used, C_used 변수를 선정하였고, 전기로 내부의 과열 정도에 관한 정보를 제공할 수 있는 Air_in과 Air_out 변수도 같이 선정하였다. 투입데이터 그룹에서는 투입 스크랩의 C와 Si 성분 함유량이 용강 내 탄소 함량과 밀접한 관련이 있기에 두 변수를 선정하였다. 측은 데이터 그룹에서는 1차 측은 시에 얻을 수 있는 용강 내 탄소 함량, 산소 농도, 용강 온도 데이터와 본 연구에서 예측하고자 하는 2차 측은 시의 탄소 함량을 변수로 선정하였다. 마지막으로 가공 데이터 그룹에서는 초기 투입 스크랩의 화학 성분을 기반으로 한 예상전력사용량을 변수로 선정하여 스크랩 정보를 반영하고자 하였으며, C_df, O2_df, Power_df, TTT_df 변수를 선정하여 1차 측은 이후부터 2차 측은까지의 조업에 관한 정보를 반영하였다.

2.4 다층 퍼셉트론(Multi-Layer Perceptron)

본 연구의 모델링에 사용한 머신러닝 기반의 모델은 다층 퍼셉트론(MLP)이다. MLP는 인공신경망의 일종으로 입력층과 출력층 사이에 하나 이상의 은닉된 중간층을 두어 비선형적으로 분리되는 데이터에 대해서도 학습이 가능하도록 한 모델이다. [그림 4]는 본 모델링에 사용된 MLP 모델을 도식

화한 그림이다. 다수의 입력 데이터들을 투입층에서 받고, 은닉층의 노드를 통과하여 출력층에서 데이터를 출력하게 된다. 이때에 은닉층은 최적의 비중과 편차를 찾는 역할을 한다. 은닉층의 층 개수와 노드 개수는 사용자가 따로 설정해 주어야 하는데, 본 모델에서는 하나의 은닉층을 사용하였고, 노드 개수를 조정하여 최적의 모델을 찾고자 하였다. 또한, 모델링 과정에서 K-fold 교차 검증을 활용하여 오버피팅을 피하고자 하였다.



[그림 3] 다층 퍼셉트론 도식화

2.5 모델링 결과 및 평가

모델의 평가지표로는 R-Squared score, RMSE(Root Mean Square Error)를 사용하였다. R-Squared score는 결정계수로 실제 관측값의 분산 대비 예측값의 분산을 계산하여 데이터 예측의 정확도 성능을 측정하는 지표이다. 0~1까지의 수로 나타내어지며 1에 가까울수록 더욱 잘 예측하는 모델이라고 말할 수 있다. R-Squared score는 아래의 [식 1]과 같이 표현할 수 있다.

$$R^2 score = \frac{\sum(\text{예측값} - \text{타겟평균값})^2}{\sum(\text{관측값} - \text{타겟평균값})^2} \dots (1)$$

RMSE는 평균 제곱근 오차로 예측값과 관측값의 차의 제곱된 값의 평균을 구하고, 구한 평균에 제곱근을 씌운 평가지표이다. RMSE는 아래의 [식 2]와 같이 표현할 수 있다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{예측값}_i - \text{관측값}_i)^2} \dots (2)$$

노드 개수의 변화에 따른 모델링 결과는 아래의 [표 3]과 같다. 최종적으로 성능이 제일 좋게 나온 노드 개수 80개일 때의 모델을 본 연구의 최종 모델로 선정하였다.

[표 3] 노드 개수에 따른 모델링 결과

노드 개수	훈련 $R^2 score$	훈련 RMSE	검증 $R^2 score$	검증 RMSE
60	0.796	3.126	0.707	3.529
70	0.848	2.693	0.828	2.705
80	0.886	2.339	0.877	2.291
90	0.820	2.931	0.796	2.947
100	0.856	2.627	0.828	2.704

3. 결론

현재 현대제철 포항공장에서는 승온기 조업 과정에서 측운 타이밍을 결정할 때 조업자의 경험과 노하우에 크게 의존하고 있으며, 이는 전기로의 효율성에 영향을 미칠 수 있다. 본 연구에서는 위의 문제를 해결하기 위하여 1차 측운 데이터와 이후 승온기 조업 데이터를 기반으로 2차 측운 시점의 탄소 함량을 예측하고자 하였다. 데이터 전처리, 변수 선정, 다층 퍼셉트론 모델링을 통하여 보다 효율적으로 전기로 공정을 관리하는 데이터 기반의 접근 방식을 제안하였다. 또한, 다층 퍼셉트론의 노드 개수에 따른 모델 결과를 분석하여 최종적으로 노드 개수가 80개일 때, 검증 $R^2 score$ 가 0.877, 검증 RMSE가 2.291을 나타내는 모델을 생성하였다. 수치적으로는 유의미한 모델을 생성하였으나, 사용 가능한 데이터의 부족으로 인해 모델의 검증이 완전히 완료되지는 않았다는 한계는 존재한다. 따라서 향후 연구에서는 보다 많은 데이터셋을 포함하여, 이 연구에서 제안한 모델을 더욱 견고히 발전시키는 것을 목표로 할 것이다.

참고문헌

[1] W. S. A. World Steel Association, "Steel statistical yearbook 2020", World Steel Association, vol. 6, 2020

[2] Song, Y., Ha, H., Lee, W., Lee, K. and Kim, J., "Data-Driven Approach Using Supervised Learning for Predicting Endpoint Temperature of Molten Steel in the Electric Arc Furnace", steel research int., 94: 2300143, May, 2023

[3] Manojlović, V., Kamberović, Ž., Korać, M., & Dotlić, M., "Machine learning analysis of electric arc furnace process for the evaluation of energy efficiency parameters" Applied Energy, 307, 118209, February 2022