CORROSION SCIENCE AND TECHNOLOGY, Vol.24, No.4(2025), pp.269~280 pISSN 1598-6462 / eISSN 2288-6524
[Research Paper] DOI: https://doi.org/10.14773/cst.2025.24.4.269

기계학습에 기반한 AA1100 합금의 부식에 미치는 환경 영향 분석

정용주 · 김민태 · 서정우 · 이주희 · 장희진 · 김희수†
조선대학교 신소재공학과, 광주광역시 동구 조선대5길 60
(2025년 7월 07일 접수, 2025년 8월 04일 수정, 2025년 8월 17일 채택)

Machine Learning-Based Analysis of Environmental Effects on the Corrosion of AA1100 Alloy

Yong Ju Jeong, Mintae Kim, Jeongwoo Seo, Ju Hee Lee, Hee Jin Jang, and Hee-Soo Kim†
Department of Materials Science and Engineering, Chosun University 60, Chosundae 5-gil, Dong-gu, Gwangju, 61452, Republic of Korea
(Received July 07, 2025; Revised August 04, 2025; Accepted August 17, 2025)

This study explores the effects of three environmental chemical species—Na_2S, NaCl, and H_2O_2—on the corrosion behavior of AA1100 aluminum alloy using machine learning (ML) techniques. We collected experimental data through a full factorial design that included 27 solution conditions, from which we extracted two key electrochemical parameters: corrosion potential (E_{corr}) and corrosion current density (i_{corr}), We trained four regression algorithms—k-nearest neighbor, random forest, support vector regression, and extreme gradient boosting—on this data and compared their performance using root mean squared error. Among these models, the random forest algorithm demonstrated the most accurate predictions for both E_{corr} and i_{corr}, leading us to select it for further analysis. To assess the influence of each input variable, we employed SHAP (SHapley Additive exPlanations) analysis. Our findings revealed that NaCl made the greatest positive contribution to E_{corr}, while H_2O_2 significantly increased i_{corr}. In contrast, Na_2S had minimal impact on both corrosion indicators. SHAP interaction plots indicated that Cl^- and H_2O_2 had synergistic effects in accelerating corrosion, while Na_2S remained inert. This study highlights the effectiveness of interpretable ML approaches in analyzing complex corrosion systems and offers a data-driven perspective for corrosion prediction and material design.

Keywords: *Aluminum alloys, Corrosion, Polarization curve, Machine learning*

1. 서론

알루미늄 합금은 높은 강도 대비 낮은 밀도, 우수한 내식성, 높은 열전도성 및 가공성 등의 우수한 물리·화학적 특성으로 인해 항공우주와 자동차, 열교환기, 전자기기 등 다양한 산업분야에서 핵심적인 소재로 활용되고 있다 [1,2]. 그 중에서도 AA1100 합금은 99.0% 이상의 고순도 알루미늄으로 구성되어 있으며, 특히 우수한 열전도특성을 바탕으로 열교환기, 전기도체, 에어컨 및 자동차 냉난방 시스템 등의 응용분야에서 널리 적용되고 있다. 알루미늄 소재의 부식저항성은 주로 표면에 자연적으로 형성되는 부동태피막에 의해 결정된다. 이 부동태피막은 알루미늄의 표면을 안정화시키고 금속의 산화와 부식을 억제하는 역할을 수행한다. 그러나 이 피막의 안정성은 용액의 수소이온농도(pH)와 염화물이온(Cl^-), 황산이온(SO_4^{2-}), 산소 농도, 산화제의 존재 여부 등 다양한 환경인자에 의해 민감하게 영향을 받는다 [3–6]. 특히 도시 및 산업환경에서는 대기 중 염화물, 황산염, 산소 등의 다양한 오염물질이 존재하여 알루미늄 합금의 부식거동의 해석이 용이하지 않다.

본 연구팀의 이전 연구 [7]에서는 이러한 환경인자들이 AA1100 합금의 부식에 미치는 영향을 규명하고자, Na_2S와 NaCl, H_2O_2가 혼합된 수용액 내에서 동전위분극실험을 수행하였다. 이를 통해 부식전위(E_{corr})와 부식전류밀도(i_{corr}), 공식전위(E_{pit}) 등의 주요 전기화학적 부

†Corresponding author: heesoo@chosun.ac.kr
정용주: 학부생, 김민태: 학부생, 서정우: 학부생, 이주희: 박사과정, 장희진: 교수, 김희수: 교수

식특성을 측정하고, 각 인자의 농도변화에 따른 부식특성의 변화를 회귀분석을 통해 평가하였다. 특히, 실험계획법(design of experiments, DOE) [8]을 기반으로 한 반응표면분석법(response surface methodology, RSM) [9]을 활용하여 변수 간 상호작용 및 주요 인자의 효과를 정량적으로 분석하였다. RSM은 비교적 적은 실험 횟수로 변수들의 최적 조합을 도출할 수 있는 유용한 통계적 기법이나, 2차 회귀모형 이상의 복잡한 비선형 반응을 정확히 예측하는 데 한계가 있으며, 실험설계 외의 데이터에 대한 일반화에도 제약이 존재한다.

최근 이러한 한계를 극복하기 위한 방법으로 기계학습(machine learning) 기법이 다양한 재료공학 분야에서 활발히 적용되고 있다. 기계학습 기반의 회귀분석은 변수 간의 비선형적 관계 및 복잡한 상호작용을 효과적으로 모델링할 수 있으며, 기존 선형회귀법에 비해 우수한 예측 성능을 제공하는 장점을 가진다 [10–12]. 특히, 부식거동과 같이 다수의 환경인자가 복합적으로 작용하는 시스템에서는 기계학습이 더욱 유용한 도구로 평가되고 있다. 이에 본 연구에서는 선행연구에서 확보된 동전위분극실험 데이터를 기반으로 다양한 기계학습 회귀알고리즘을 적용하여, Na_2S, NaCl 및 H_2O_2 농도가 AA1100 합금의 부식 특성(E_{corr} 및 i_{corr})에 미치는 영향을 정량적으로 분석하고자 한다. 해석 가능한 인공지능(explainable artificial intelligence, XAI) [13] 기법을 도입하여 각 입력변수의 기여도를 정량적으로 평가함으로써, 복잡한 부식거동에 대한 보다 구체적인 해석 체계를 제공하고자 한다. 아울러, 다양한 기계학습 모델의 예측성능을 비교·평가함으로써 부식 예측 문제에 대한 최적의 모델 선택 기준을 제시하였다.

2. 기계학습

2.1 학습 데이터

본 연구에 사용된 데이터는 이전 연구 [7]에서 수행된 동전위분극실험의 결과를 기반으로 하였다. 실험은 알루미늄 합금인 AA1100 시편을 대상으로 수행되었으며, 세 가지 환경 변수(Na_2S, NaCl, H_2O_2)의 농도를 각각 세 수준으로 설정한 완전요인설계(full factorial design)

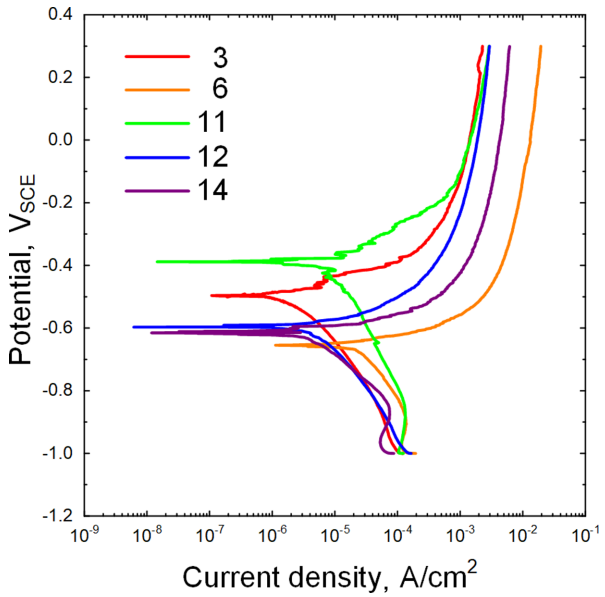

Fig. 1. Examples of polarization curves of AA1100 under various corrosive environments [7]; The numbers specified for the curves mean the experimental conditions listed in Table 1

Table 1. Chemical compositions of test solutions and corresponding E_{corr} and i_{corr} values for AA1100 corrosion analysis [7]

Solution no.	Na_2S (g/L)	NaCl (g/L)	H_2O_2 (mL/L)	E_{corr} (V_{SCE})	i_{corr} (A/cm^2)
1	0	0.3	0	-0.5870	6.5×10^{-7}
2	0	0.3	5	-0.4343	2.5×10^{-9}
3	0	0.3	10	-0.4961	3.0×10^{-10}
4	0	1	0	-0.8229	2.0×10^{-7}
5	0	1	5	-0.5644	1.4×10^{-9}
6	0	1	10	-0.6547	2.8×10^{-9}
7	0	3	0	-0.7039	6.9×10^{-8}
8	0	3	5	-0.6025	2.3×10^{-7}
9	0	3	10	-0.5546	4.6×10^{-9}
10	0.05	0.3	0	-0.7990	1.40×10^{-7}
11	0.05	0.3	5	-0.3882	9.9×10^{-7}
12	0.05	0.3	10	-0.5970	7.4×10^{-7}

Table 1. (Continued) Chemical compositions of test solutions and corresponding E_{corr} and i_{corr} values for AA1100 corrosion analysis [7]

Solution no.	Na_2S (g/L)	NaCl (g/L)	H_2O_2 (mL/L)	E_{corr} (V_{SCE})	i_{corr} (A/cm^2)
13	0.05	1	0	-0.8158	2.3×10^{-8}
14	0.05	1	5	-0.6154	4.2×10^{-7}
15	0.05	1	10	-0.5250	5.1×10^{-7}
16	0.05	3	0	-0.8586	5.0×10^{-8}
17	0.05	3	5	-0.6019	3.0×10^{-9}
18	0.05	3	10	-0.6260	8.7×10^{-7}
19	0.1	0.3	0	-0.7461	8.1×10^{-8}
20	0.1	0.3	5	-0.5971	3.6×10^{-7}
21	0.1	0.3	10	-0.5640	5.8×10^{-7}
22	0.1	1	0	-0.8512	5.6×10^{-8}
23	0.1	1	5	-0.5739	2.4×10^{-7}
24	0.1	1	10	-0.5958	8.8×10^{-7}
25	0.1	3	0	-0.7504	3.0×10^{-8}
26	0.1	3	5	-0.5806	3.8×10^{-7}
27	0.1	3	10	-0.6321	7.3×10^{-7}

을 통해 총 27개의 조합이 구성되었다. 각 화학종의 농도범위는, Na_2S의 경우 0 ~ 0.1 g/L이며 NaCl은 0.3 ~ 3 g/L, H_2O_2는 0 ~ 10 mL/L였다. 각 실험조합에 대하여 동일한 실험조건 하에서 분극시험을 진행하였으며, 해당 결과를 기계학습 모델 학습에 활용하였다.

동전위분극실험은 3전극 셀을 사용하여 수행되었으며, 포화감홍전극(saturated calomel electrode; SCE)을 기준전극으로, AA1100 시편을 작동전극으로 설정하였다. 분극곡선은 전위-전류 밀도 형태로 수집되었으며, E_{corr}는 전류밀도가 최소가 되는 지점으로부터, i_{corr}는 타펠외삽법(Tafel extrapolation method)을 적용하여 도출하였다. 대표적인 분극곡선은 Fig. 1에, 27개 실험조건 및 각 조건에 대응하는 E_{corr} 와 i_{corr} 값은 Table 1에 각각 제시하였다.

2.2 기계학습

본 연구에서는 E_{corr}와 부식전류밀도 i_{corr}를 예측하기 위해 총 네 가지 기계학습 기반 회귀 모델을 구축하고 그 결과를 비교하였다. 입력변수 X는 각 실험조건에 해당하는 세 가지 화학종(Na_2S, NaCl, H_2O_2)의 농도 값으로 구성되며, 출력변수 Y는 각각 E_{corr} 또는 i_{corr}로 설정하였다. 이 중 i_{corr}는 데이터 분포의 왜곡과 분산을 줄이기 위해 로그스케일로 변환하여 사용하였다.

본 연구에서 선정된 회귀모델은 K-최근접 이웃(K-Nearest Neighbors, K-NN) [14]과 서포트 벡터 회귀(Support Vector Regression, SVR) [15], 랜덤 포레스트(Random Forest, RF) [16], 고성능 그래디언트 부스팅 기법(eXtreme Gradient Boosting, XGBoost, XGB) [17]이다. 이들은 서로 다른 수학적 원리와 일반화 특성을 기반으로 하며, 부식과 같이 비선형성 및 변수 간 상호작용이 복잡하게 나타나는 시스템에 대한 예측성능을 비교하기 위해 사용하였다. 각 알고리즘의 이론적 배경과 구현 방식은 각각의 참고문헌 [14–17]에 제시되어 있으므로, 본문에서는 생략하였다.

예측성능을 나타내는 방법에는 여러가지 평가지표가 있는데, 본 연구에서는 평균제곱근오차(Root Mean Squared Error, RMSE)와 평균절대오차(Mean Absolute Error, MAE), 평균절대비율오차(Mean Absolute Percentage Error, MAPE)을 적용하여 각 모델을 비교하여, 각 평가지표가 최소값을 나타내는 기법을 최종 학습모델에 적용하였다.

각 모델에 대해서는 성능향상을 위해 GridSearchCV [18] 기법을 활용하여 주요 하이퍼파라미터(hyperparameter) 최적화를 수행하였다. 각 모델의 최적화된 하이퍼파라미터는 Table 2에 나타내었다. 또한 입력 데이터셋의 크기가 작기 때문에 데이터셋을 임의로 세 부분으로 나눈 후, 그중 한 부분을 검증데이터로, 나머지 부분을 훈련데이터로 사용하는 방법을 순환적으로 사용하는 K-겹 교차검증(K-fold cross validation) 방식을 적용하였다.

Table 2. Optimized hyperparameters for each regression model and prediction target

Model	Target	Scaler	Optimized hyperparameters
K-NN	E_{corr}	Robust	weights = uniform, n = 6, p = 1
	i_{corr}	Standard	weights = uniform, n = 5, p = 1
RF	E_{corr}	Standard	Max depth = 5, Max features = sqrt, n_estimators = 100, min samples leaf = 1
	i_{corr}	Standard	Max depth = 5, Max features = sqrt, n_estimators = 100, min samples leaf = 1
SVR	E_{corr}	Robust	C = 10, epsilon = 0.05, Gamma = scale, kernel = rbf
	i_{corr}	Standard	C = 10, epsilon = 0.01, Gamma = 0.1, kernel = rbf
XGB	E_{corr}	Robust	Learning rate = 0.01, Max depth = 3, n_estimators = 200
	i_{corr}	Robust	Learning rate = 0.01, Max depth = 3, n_estimators = 200

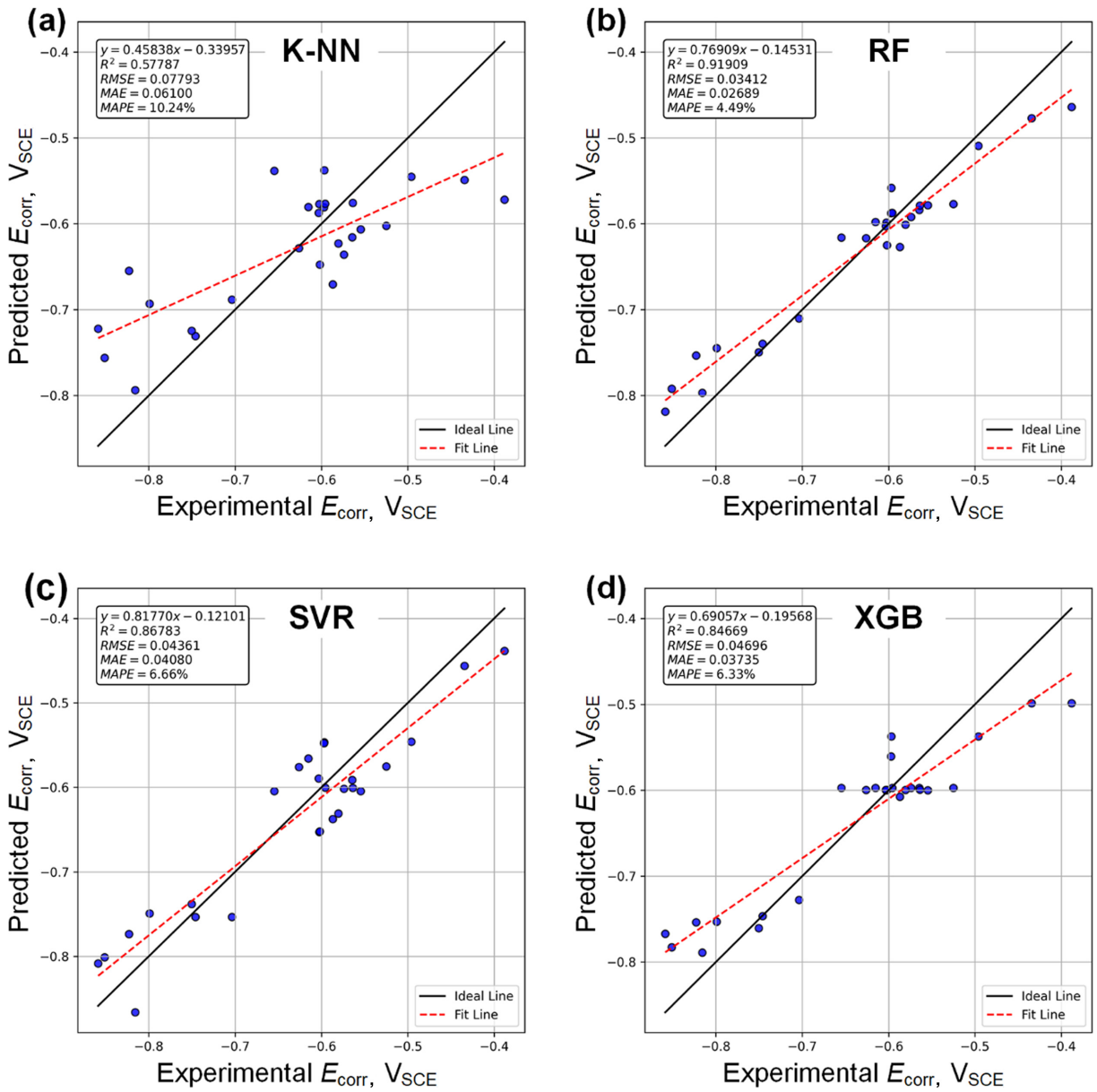

Fig. 2. Comparison between experimental and predicted values of E_{corr} for each regression model: (a) K-NN, (b) RF, (c) SVR, and (d) XGB

3. 결과 및 고찰

3.1 모델별 예측성능 비교

본 연구에서는 K-NN와 RF, SVR, XGB의 네 가지 회귀모델에 대해 각각 AA1100 합금의E_{corr}와 i_{corr}를 예측하고 그 성능을 비교하였다. Fig. 2는 각 회귀 모델별로 실험측정값과 예측값 사이의 상관관계를 나타낸 그래프로, 각 점은 특정 실험 조건에서의 실험값과 예측값의 쌍을 나타낸다. 검은색 점선은 이상선(ideal line)이며, 예측값과 측정값이 완전히 일치할 경우 위치하게 되는 기준선이다. 빨간색 점선은 선형회귀분석을 통해 도출된 최적적합선(best-fit line)으로, 해당 모델의 예측경향을 시각적으로 나타낸다. 모델의 예측 정확도를 정량적으로 평가하기 위해 회귀선의 기울기와 결정계수(R^2), RMSE, MAE, MAPE를 함께 분석하였다. 회귀선의 기울기는 예측 성능의 직접적인 평가지표로 해석되기에는 한계가 있으므로, 예측값과 측정값 사이의 전반적인 추세의 일치도를 시각적으로 보여주는 보조 지표로만 사용하였다. R^2는 회귀모델의 선형성 정도를 나타내는 지표로, 실제 측정값과의 직접적인 일치도를 평가하기 어

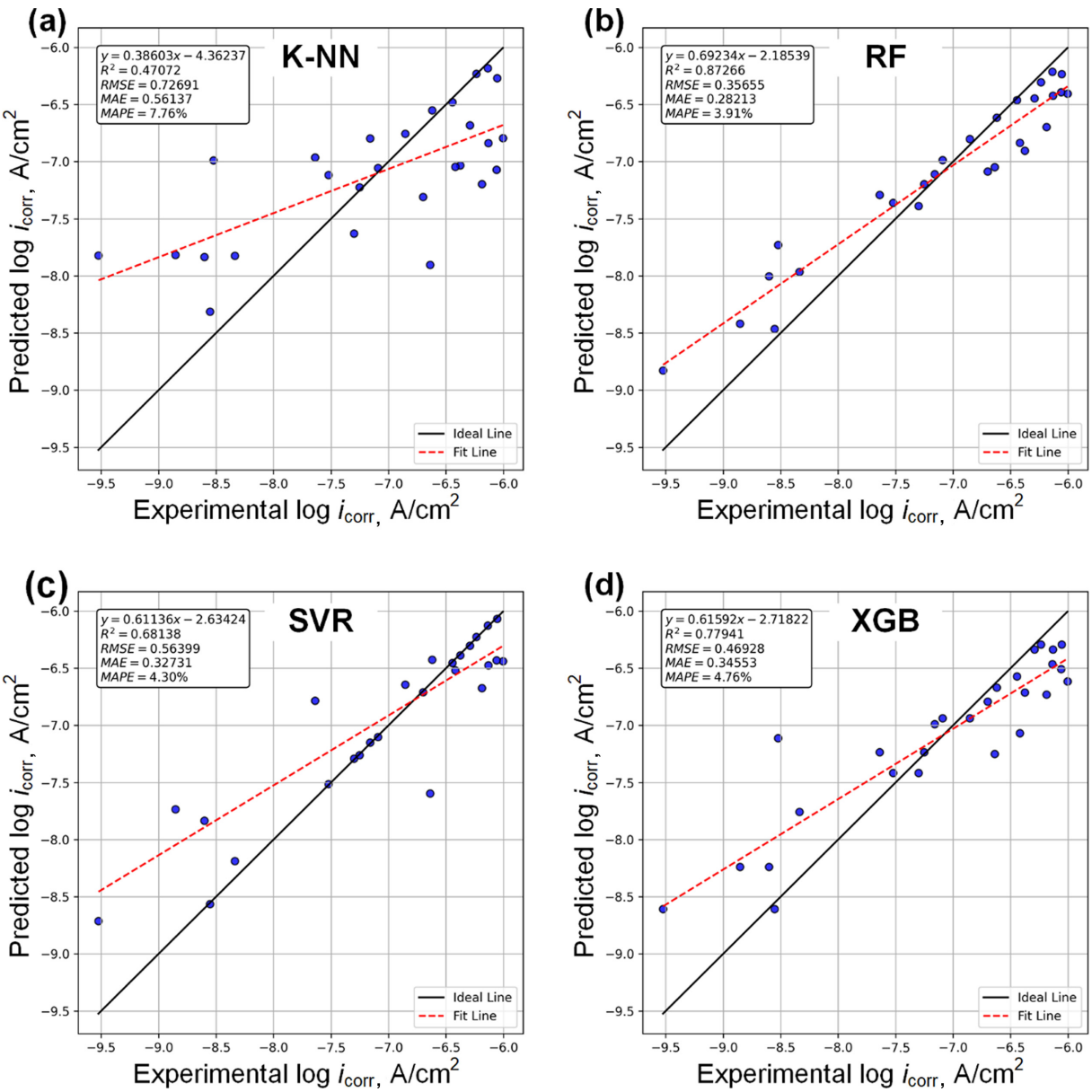

Fig. 3. Comparison between experimental and predicted values of i_{corr} for each regression model: (a) K-NN, (b) RF, (c) SVR, and (d) XGB

려우므로 본 논문에서는 이에 대한 고찰은 생략하였다. RMSE와 MAE, MAPE는 각 모델에 대하여 다른 수치가 나타났지만 전반적인 대소의 순위가 유사하게 나타났으므로, 편의상 RMSE로만으로 비교한 결과를 먼저 제시하였다. 그 이후에 RMSE와 MAE, MAPE의 수치 비교에 대한 고찰을 수행하였다.

Fig. 2에 나타난 기울기 측면에서 살펴보면, SVR의 회귀선 기울기가 0.8177로 1에 가장 근접하여 예측값이 이상선과 가장 유사한 경향을 보였음을 확인할 수 있다. 다음으로 RF의 기울기는 0.7691로, SVR에 비해 다소 낮지만 비교적 높은 예측 일관성을 나타냈다. 한편, RMSE 측면에서는 RF가 0.03412 V로 가장 낮은 값을 기록하였으며, 이는 SVR(0.04361 V)보다도 낮아 실험값과 예측값 사이의 오차가 가장 작고 고르게 분포되어 있음을 의미한다. RMSE는 예측값의 분산과 오차를 직접적으로 반영하는 지표이므로, E_{corr} 예측에 있어 RF 모델이 가장 안정적이고 신뢰도 높은 결과를 제공하는 모델로 판단된다.

반면, K-NN은 실험값의 전반적인 경향을 반영하는 데 한계를 보였으며, XGB 모델은 일반적으로 강력한 예측 성능으로 알려져 있음에도 불구하고 본 데이터셋에서는 일부 왜곡된 예측 경향이 나타났다. 이는 본 연구의 제한된 데이터 수와 변수 특성에 따른 모델 적합도 차이로 해석될 수 있다.

부식전류밀도에 대한 각 회귀 모델의 실험값-예측값의 비교는 Fig. 3에 나타내었다. 모든 모델에서 예측값과 이상선의 일치 정도가 E_{corr} 예측에 비해 낮게 나타났으며, 이는 i_{corr}의 데이터 분포가 비선형성이 강하고, 실험적 노이즈에 민감하며, 값의 범위가 넓기 때문으로 해석된다. 이 중 RF 모델은 회귀선의 기울기 0.6923과 RMSE 0.3566 A/cm^{2}로 가장 낮은 예측 오차를 보이며, i_{corr} 예측에서도 가장 높은 신뢰도를 나타냈다. XGB 모델은 기울기 0.6159, RMSE 0.4693 A/cm^{2}를 기록하여 RF 다음으로 우수한 성능을 보였다. 반면, SVR과 K-NN은 각각 기울기와 RMSE 측면에서 보다 낮은 정확도를 보이며, 실험값에 대한 예측 편차가 크게 나타났다.

E_{corr}와 i_{corr} 예측 성능을 종합적으로 비교한 결과, RF 모델은 두 변수 모두에서 회귀선 기울기가 1에 근접하고 RMSE가 가장 낮은 값을 보여, 전체적으로 가장 안정적이고 신뢰도 높은 회귀 성능을 갖는 것으로 평가되었다.

Fig. 4는 네 가지 회귀모델의 RMSE와 MAE, MAPE는 E_{corr}와 i_{corr} 각각에 대해 정리한 비교 도표이다. 앞서 언급한 것과 같이 각 평가지표는 모두 다른 값을 나타내지만, 모델별 대소관계를 크게 다르게 나타나지 않았다. 단지 MAE와 MAPE는 E_{corr}와 i_{corr}에 대한 순위가 RMSE의 순위와 다소 다르게 나타났다. 그럼에도 불구하고 RF 모델이 모든 평가지표에서 모두 가장 낮은 예측 오차를 기록했음을 시각적으로 확인할 수 있다. 이에 따라 본 연구에서는 이후의 XAI 분석을 RF 모델을 중심으로 수행하였다.

3.2 첨가 화합물의 주영향 분석

본 연구에서는 AA1100 합금의 동전위분극실험에서 얻은 E_{corr}와 i_{corr}에 대한 예측성능이 가장 우수하게 나타난 RF 회귀 모델을 기반으로 모델 예측에 대한 각 입력변수의 기여도를 정량적으로 분석하기 위해, SHAP (SHapley Additive exPlanations) 분석을 수행하였다.

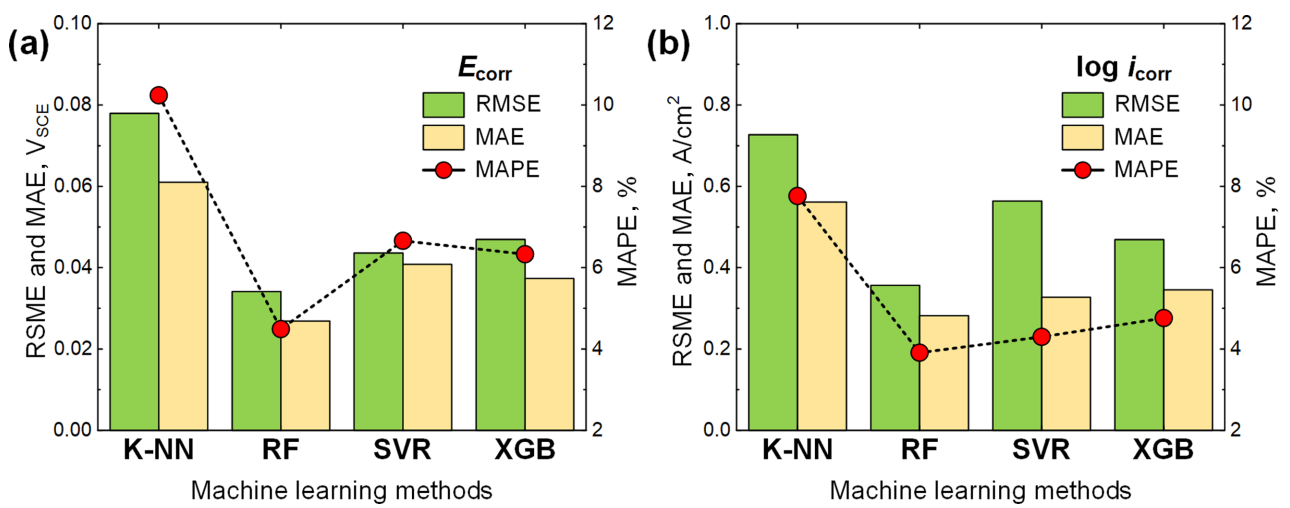

Fig. 4. Comparison of RMSE, MAE, and MAPE values for (a) E_{corr} and (b) i_{corr} obtained from different regression models. Connecting lines for MAPE are shown solely for visual guidance and do not imply continuous trends

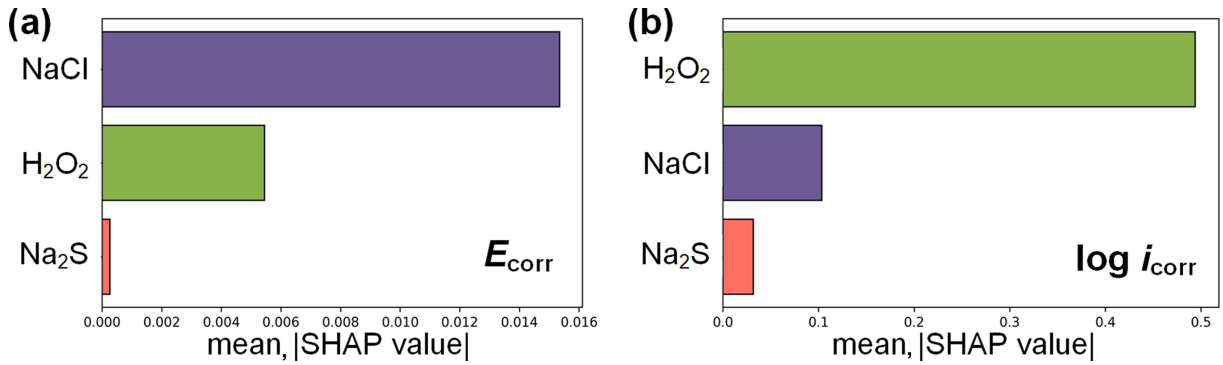

Fig. 5. Mean absolute SHAP values of input features for the prediction of (a) E_{corr} and (b) i_{corr} using the RF model. Each bar represents the average magnitude of SHAP values across all samples, indicating the overall importance of each feature in the model. NaCl and H_2O_2 were the most influential features for E_{corr} and i_{corr}, respectively

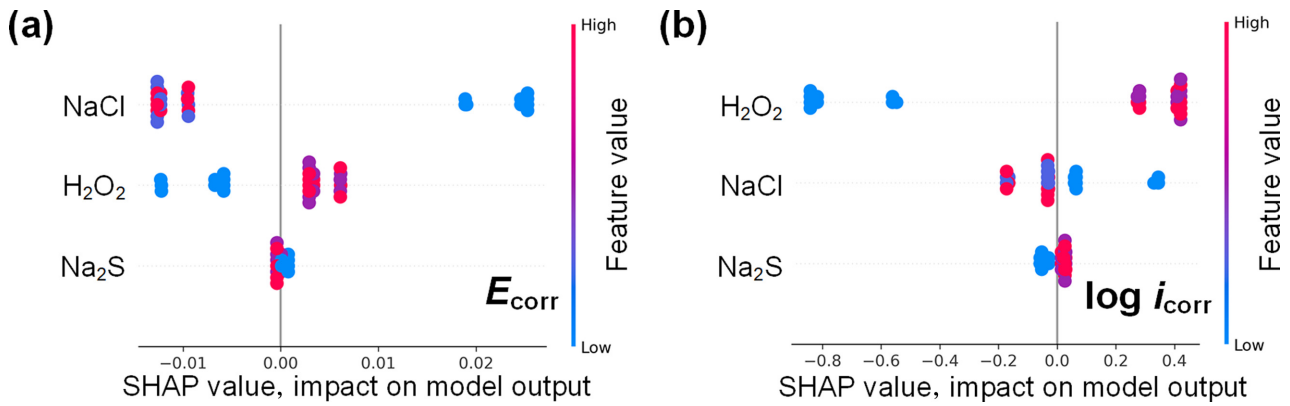

Fig. 6. SHAP summary plots for input features in the prediction of (a) E_{corr} and (b) i_{corr} using the RF model. Each dot represents a single prediction sample, where the *x*-axis indicates the SHAP value, and the color denotes the actual value of the feature (red: high, blue: low). For the E_{corr} (a), NaCl shows a predominantly positive influence, especially at high concentrations, while H_2O_2 tends to decrease E_{corr}. For the i_{corr} (b), H_2O_2 has the strongest positive contribution to the prediction, indicating its dominant role in accelerating corrosion rate. Na_2S exhibits relatively minor influence in both cases

SHAP은 게임이론에 기반한 설명 가능한 인공지능(XAI) 기법으로, 각 특성이 예측값에 미치는 영향을 직관적이고 정량적으로 설명할 수 있어 널리 활용되고 있다 [13].

두 목적변수(E_{corr}와 i_{corr})에 대한 SHAP 기반 변수중요도분석 결과는 Fig. 5에 나타내었으며, 이를 통해 각 변수의 상이한 기여 패턴을 확인할 수 있었다. Fig. 5a에 나타낸 바와 같이 E_{corr}에 대해서는 NaCl의 함량이 가장 큰 기여도를 보였다. 반면, H_2O_2의 기여도는 이보다 낮았으며, Na_2S의 영향은 미미한 것으로 나타났다. 한편 Fig. 5b에서 보는 바와 같이 i_{corr}에 대해서는 H_2O_2의 기여도가 가장 높게 나타났고, NaCl는 H_2O_2보다 기여도가 낮은 것으로 나타났다. 반면, Na_2S는 E_{corr}의 경우와 유사하게 매우 낮은 기여도를 보였다.

Fig. 6은 SHAP 값을 샘플 단위로 시각화하여 각 특성이 예측값에 미치는 영향 방향을 직관적으로 확인할 수 있도록 나타낸 그래프이다. 이 SHAP 요약 그래프(SHAP summary plot) 로서 Fig. 5의 단순 영향도 분석보다 좀더 자세한 변수의 영향에 대한 설명이 가능하다. SHAP 값은 모델 예측값에 대한 각 특성의 기여도를 나타내며, 색상은 해당 특성의 실험 범위 내의 특성값(feature value)을 의미한다. *X*축 방향으로 SHAP 값이 양(+)이면 예측값 증가에 기여함을, 음(–)이면 예측값 감소에 기여함을 나타낸다. SHAP의 값이 0이면 모델 예측에 미치는 영향이 거의 없다는 것을 나타낸다.

Fig. 6a의 E_{corr} 모델에서는 NaCl이 예측값 증가 방향으로 가장 크게 기여하는 변수로 나타났으며, 특히 고농도 NaCl 조건에서 양(+)의 SHAP 값이 두드러졌다. 반면 H_2O_2는 고농도 조건에서 음의 방향으로 작용하며, E_{corr}를 낮추는 경향을 보였다. Na_2S는 전반적으로 SHAP 값의 크기가 작고 중심에 밀집되어 있어, E_{corr}에서 상대적으로 낮은 영향력을 나타냈다. 한편 Fig. 6b의 i_{corr} 모델에서는 H_2O_2가 예측값 증가에 가장 크게 기여한 변

수로 나타났으며, 8 mL/L 이상의 고농도 H_2O_2에 해당하는 점들이 평균적으로 0.3 이상 높은 SHAP 값을 나타내었다. 반면 NaCl과 Na_2S는 SHAP 값이 중심 부근에 밀집되어 있어, i_{corr}에 미치는 영향이 제한적인 것으로 분석되었다.

NaCl의 농도(특성값) 증가에 따른 E_{corr}의 변화는 Fig. 6a에 제시된 바와 같이, Cl^- 이온의 농도가 증가할수록 AA1100 합금의 E_{corr}가 점진적으로 음(-)의 방향으로 이동함을 확인할 수 있었다. 이는 금속이 보다 활성화된 상태로 전이되었음을 의미하며, 부식반응이 더 쉽게 발생할 수 있는 조건이 형성된 것이라고 판단된다. Cl^- 이온은 Al 표면에 존재하는 보호산화피막(Al_2O_3)의 안정성을 저하시켜 피막을 국부적으로 파괴하며, 금속 표면을 직접적인 부식반응에 노출시킨다. 이로 인해 금속의 산화반응이 촉진되고, 전자전달 및 양이온의 용출이 용이해져 E_{corr}가 낮은 방향으로 이동한다고 보고되었다. 또한 Cl^- 이온은 미세한 결함이나 크랙 부위에 침투하여 국부부식(pitting corrosion)을 유도함으로써, E_{corr}의 안정성을 더욱 감소시키고 전체적인 부식 민감도를 증가시키는 것으로 해석된다 [19].

H_2O_2는 AA1100 합금의 부식 거동에 있어 중요한 산화제로 작용한다. 실험 결과에 따르면, H_2O_2가 포함된 환경에서는 E_{corr}가 음전위에서 양전위 방향으로 이동하여 금속의 열역학적 안정성이 향상되는 경향을 보인다. H_2O_2의 존재는 i_{corr}를 낮춘다고 보고되었으며 [20], 본 연구의 결과가 이와 일치함을 확인하였다. H_2O_2는 Al 합금 표면의 산화피막이 보다 안정적으로 형성되어 부식 반응을 억제하도록 하는 효과를 나타낸다 [20]. 최근 연구에서는 H_2O_2 농도를 70 g/L까지 증가시켰을 때 가장 치밀하고 균일한 산화피막이 형성되며, 이로 인해 전기화학적 부식 저항성이 현저히 향상됨이 보고되었다 [21]. 그러나 H_2O_2의 농도가 과도할 경우, 활성산소종의 생성과 함께 산화피막 내 결함이 발생하여 오히려 부식감수성을 증가시킬 수 있으므로, 최적농도 설정이 중요하다.

Na_2S은 금속 부식 환경에서 다양한 반응을 일으킬 수 있는 가능성이 있으나, Al 합금의 부식 거동에는 상대적으로 미미한 영향을 주는 것으로 본 연구진의 이전 논문 [7]에서 보고된 바 있으며, 본 연구에서도 같은 결과를 얻었다. Na_2S가 Al 표면에 존재하는 Al_2O_3의 안정성에 실질적인 영향을 미치지 않으며, 산화-환원 반응 또한 활성화시키지 않는다고 판단된다.

3.3 첨가 화합물의 상호영향 분석

Fig. 7은 SHAP 상호작용 분석을 통해, Na_2S, NaCl, H_2O_2 간의 상호작용이 AA1100 합금의 E_{corr}에 미치는 영향을 분석한 결과이다. 각 그래프는 특정 변수의 실제 농도에 따른 SHAP 값을 시각화한 것이며, 점의 색상은 해당 샘플에서의 다른 변수 농도를 나타낸다.

Fig. 7a와 b에 나타낸 바와 같이 E_{corr}에 대한 Na_2S의 SHAP 값은 전체적으로 0 근처에 밀집되어 있다. 이는 Na_2S가 E_{corr} 예측에서 단독으로도, 다른 변수와의 상호작용에서도 상대적으로 영향력이 낮은 변수임을 의미한다.

반면 NaCl은 Fig. 7c에서 농도가 1 g/L이상이면 0.3 g/L일 때에 비하여 E_{corr}를 하강시킨다. NaCl 농도가 높은 경우에는 H_2O_2 또는 Na_2S와의 상호효과가 두드러지지 않으나, NaCl이 0.3 g/L일 때는 H_2O_2 농도가 높은 경우에서 SHAP 값이 높게 나타났으며, 이는 산화제로 작용하는 H_2O_2의 영향이 E_{corr}를 낮추는 NaCl의 영향을 크게 상쇄한 것으로 볼 수 있다. Fig. 7d에서 Na_2S와 NaCl의 상호효과는 경향성이 분명하지 않아 상호효과를 판단할 수 없다.

Fig. 7e의 H_2O_2는 농도가 0일 때 SHAP 값이 음의 방향에 위치하며, H_2O_2의 농도가 증가하면 양의 값을 갖는다. 이는 H_2O_2가 산화제로 작용한 결과로 볼 수 있다. Na_2S 농도에 따른 상호효과는 판단하기 곤란하다(Fig. 7e). 반면 Fig. 7f에서 H_2O_2의 농도가 0일 때 NaCl의 농도가 높을 때 H_2O_2의 SHAP 값이 덜 음의 값으로 이동하는 것이 확인된다. 즉 H_2O_2를 첨가하지 않으면 첨가했을 때에 비하여 E_{corr}가 낮으며, NaCl은 H_2O_2 농도에 따른 E_{corr}의 차이를 줄인다. 즉 NaCl과 H_2O_2가 서로 상쇄 작용을 일으키는 것으로 볼 수 있다.

Fig. 8는 i_{corr}에 대한 SHAP 상호작용 결과를 나타낸 것이다. 먼저 Fig. 8a와 b의 Na_2S는 E_{corr}의 경우와 마찬가지로 상호작용에서도 기여도가 미미하여, 그 영향이 매우 제한적임을 알 수 있다.

Fig. 8c의 NaCl의 경우, 0.3 g/L일 때는 SHAP이 양의 값이며 H_2O_2의 농도가 낮을 때 더 큰 양의 값을 갖는다. 반면 1.0 g/L 이상에서 SHAP 값은 거의 0이거나 대체로 음의 방향에 분포하며, 특히 H_2O_2 농도가 낮은 조건에서 더 큰 음의 기여도를 나타냈다. 즉 NaCl의 농도가 낮거나 높은 경우 모두 H_2O_2의 농도가 높을 때 NaCl은 i_{corr}에 거의 영향을 끼치지 않았으며, H_2O_2의 농도가 낮을 때만 양 또는 음의 효과를 나타내는 것으로 해석된다. Fig. 8d에서도 NaCl 농도가 낮을 때와 높을 때 각각 약한 양 및 음의 SHAP 값을 나타내며, Na_2S와의 상호작용은 없는 것으로 보인다. Fig. 5b에서 확인한 바와 같이 NaCl이 i_{corr}에 미치는 영향은 H_2O_2보다 훨씬 작으므로, Cl^-가 i_{corr}에 미치는 영향은 미미한 것으로 보인다. 또는, Cl^- 이온이 단독으로 존재할 때는 i_{corr}을 억제하거나, 산화제가 충분하지 않은 조건에서 상대적으로 덜 활성화됨을 의미할 수 있다.

반면, Fig. 8e의 H_2O_2는 농도가 0일 때를 제외하고 모든 조건에서 강한 양(+)의 SHAP 값을 보였다. Na_2S와

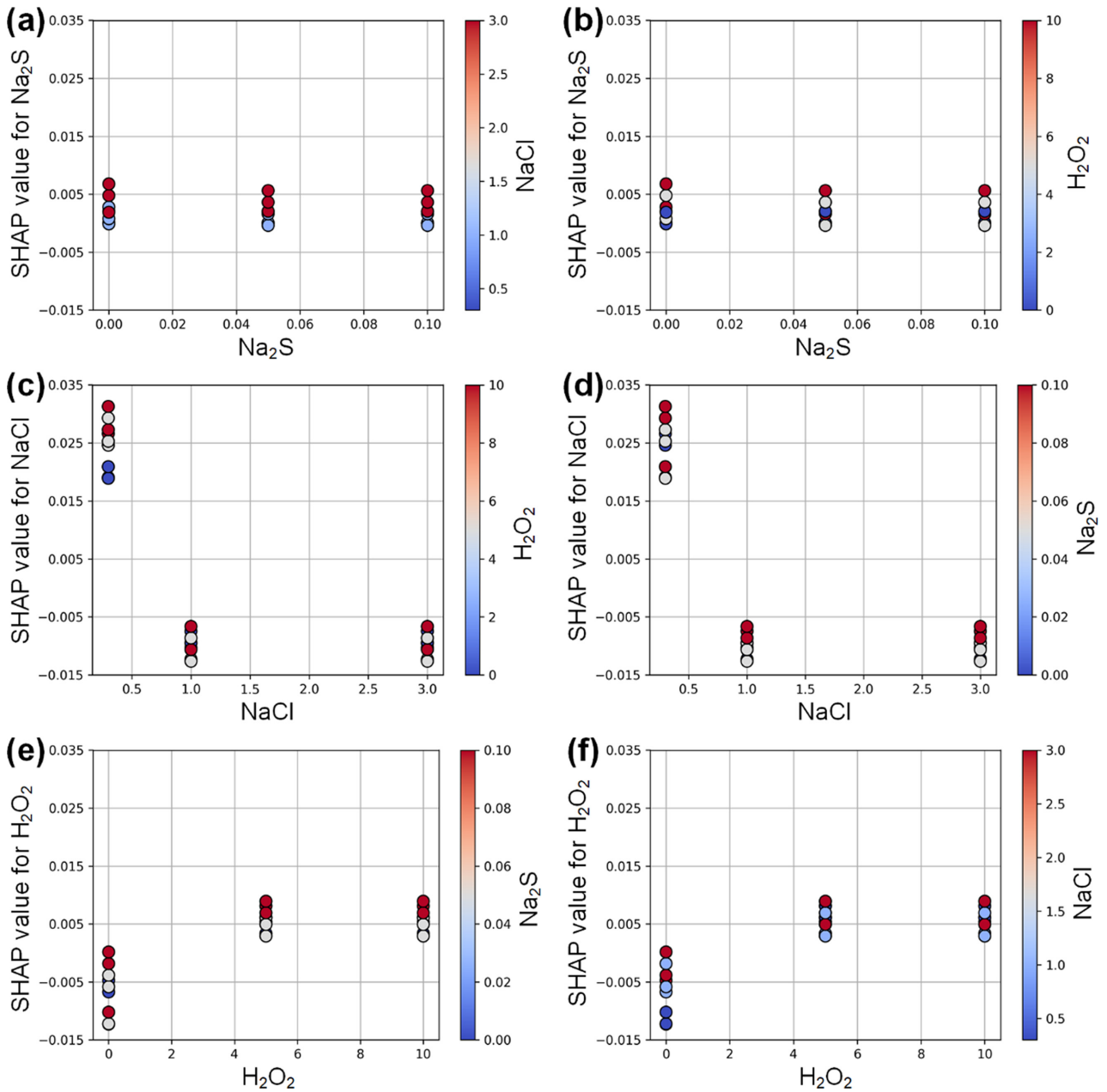

Fig. 7. SHAP dependence plots of input features for the prediction of E_{corr} using the RF model. Each plot shows the SHAP value of one feature, with point color representing the value of an interacting feature: (a) SHAP values of Na_2S, colored by NaCl concentration, (b) Na_2S colored by H_2O_2, (c) NaCl colored by H_2O_2, (d) NaCl colored by Na_2S, (e) H_2O_2 colored by Na_2S, and (f) H_2O_2 colored by NaCl. These plots illustrate both the marginal effects and interaction patterns among input variables in predicting E_{corr}

의 상호작용은 없는 것으로 보이나, Fig. 8f에서 NaCl 농도가 높은 경우 H_2O_2의 SHAP 값이 더욱 음 또는 양으로 커지는 양상을 보였다. 이는 두 인자가 상호작용하여 H_2O_2 농도가 높을수록 i_{corr}가 증가하는 효과를 강화하는 것으로 보인다. 즉, Cl^-는 부동태피막을 파괴하며, H_2O_2는 이로 인해 노출된 금속 표면의 산화 반응을 촉진하는 것으로 해석된다.

4. 결론

본 연구에서는 Na_2S와 NaCl, H_2O_2의 농도가 AA1100 알루미늄 합금의 E_{corr}와 i_{corr}에 미치는 영향을 정량적으로 분석하고자, 동전위분극실험 데이터 기반의 기계학습 회귀모델을 구축하였다. 총 27개의 실험 데이터를 바탕으로 K-NN와 RF, SVR XGB의 네 가지 회귀 알고리

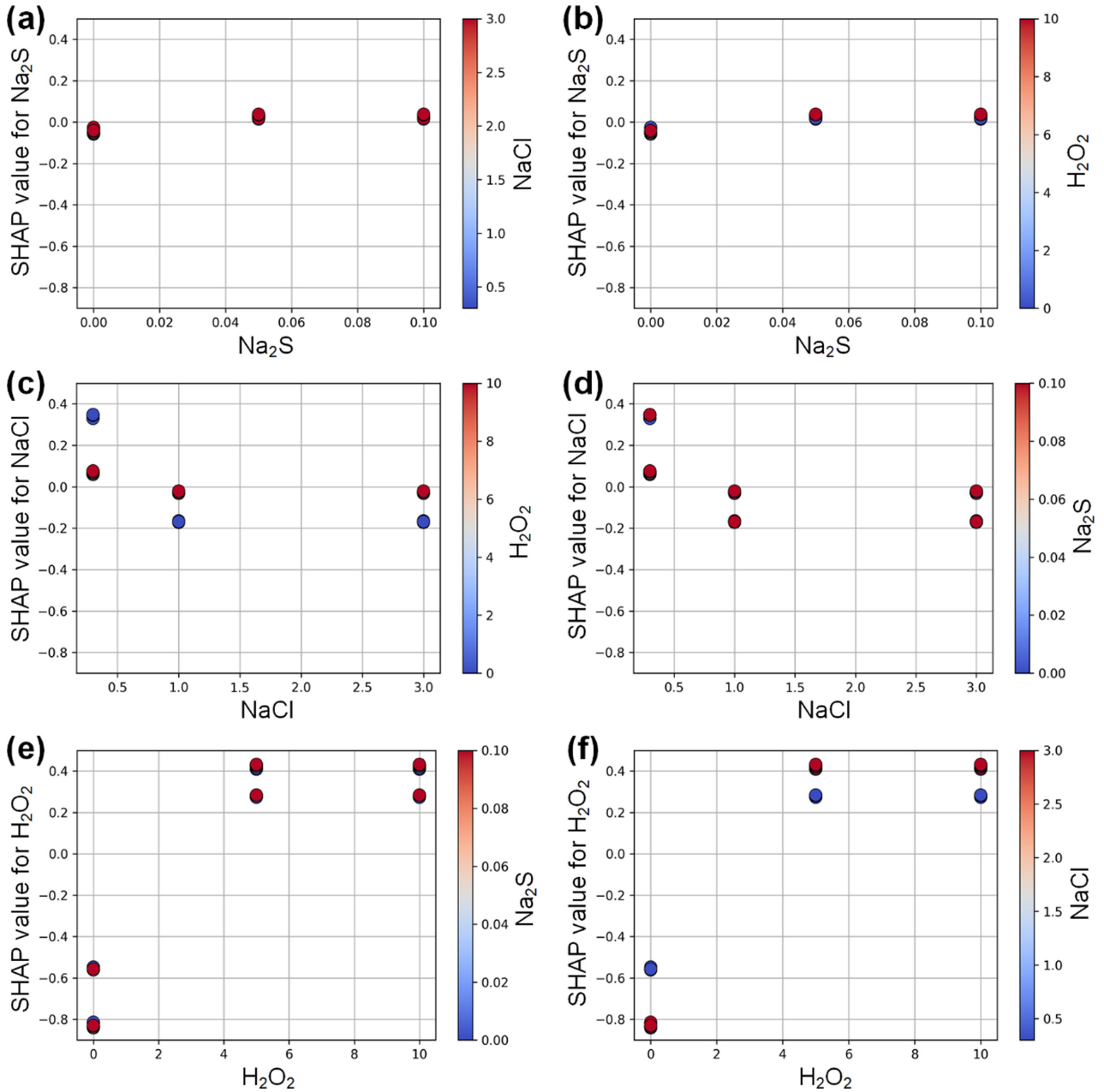

Fig. 8. SHAP dependence plots of input features for the prediction of i_{corr} using the RF model. Each plot shows the SHAP value of one feature, with point color representing the value of an interacting feature: (a) SHAP values of Na_2S, colored by NaCl concentration, (b) Na_2S colored by H_2O_2, (c) NaCl colored by H_2O_2, (d) NaCl colored by Na_2S, (e) H_2O_2 colored by Na_2S, and (f) H_2O_2 colored by NaCl. These plots illustrate both the marginal effects and interaction patterns among input variables in predicting i_{corr}

즘을 사용하여 E_{corr}와 i_{corr}를 예측비교하였다. 그 결과, RF가 두 가지 목적변수 모두에서 가장 낮은 RMSE와 가장 높은 예측안정성을 나타내었다. SHAP 기반 변수 기여도 분석을 통해 다음과 같은 주요 결과를 도출하였다:

NaCl은 E_{corr} 예측값을 증가시키는 방향으로 가장 크게 작용하였으며, 이것은 탈부동태에 의한 전위 하강 작용에 따른 것으로 생각된다.

H_2O_2는 i_{corr} 예측에서 예측값 증가에 가장 큰 영향을 미친 변수로, 산화 반응을 촉진하여 부식 전류 밀도를 높이는 주요 인자로 작용하였다.

상호작용 분석 결과, NaCl과 H_2O_2 간의 상호작용에서 부식 촉진의 시너지 효과가 관찰되었으며, 이는 실험적으로 알려진 전기화학적 메커니즘과 일치하였다.

Na_2S는 다른 변수들과의 상호작용에서도 영향도가 낮

아, 상대적으로 독립적이고 비활성적인 변수로 평가되었다.

본 연구는 XAI 기법을 활용하여 부식거동을 해석하고 예측하는 분석 방법을 제공하였으며, 복합적인 환경변수조건 하에서의 부식예측에 있어 기계학습 기반 모델의 적용 가능성을 제시하였다.

Acknowledgement

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2023-00217415).

References

1. W. S. Miller, L. Zhuang, J. Bottema, A. J. Wittebrood, P. De Smet, A. Haszler, and A. Vieregge, Recent development in aluminium alloys for the automotive industry, *Materials Science and Engineering: A*, **280**, 37 (2000). Doi: https://doi.org/10.1016/S0921-5093(99)00653-X
2. J. Hirsch, Recent development in aluminium for automotive applications, *Transactions of Nonferrous Metals Society of China*, **24**, 1995 (2014). Doi: https://doi.org/10.1016/S1003-6326(14)63305-7
3. Z. Szklarska-Smialowska, Pitting corrosion of aluminum, *Corrosion Science*, **41**, 1743 (1999). Doi: https://doi.org/10.1016/S0010-938X(99)00012-8
4. G. S. Frankel, Pitting Corrosion of Metals: A Review of the Critical Factors, *Journal of The Electrochemical Society*, **145**, 2186 (1998). Doi: https://doi.org/10.1149/1.1838615
5. P. Schmuki, From Bacon to barriers: a review on the passivity of metals and alloys, *Journal of Solid State Electrochemistry*, **6**, 145 (2002). Doi: https://doi.org/10.1007/s100080100219
6. F. Bolzoni, G. Contreras, L. Lazzari, M. Ormellese, E. Pérez, G. Re, and A. Benedetti, Influence of chloride and sulphate content on the performance of aluminium based galvanic anodes, *Corrosion Reviews*, **29**, 287 (2011). Doi: https://doi.org/10.1515/CORRREV.2011.017
7. J. H. Lee and H. Jang, Effects of Na_2S, NaCl, and H_2O_2 Concentrations on Corrosion of Aluminum, *Corrosion Science and Technology*, **18**, 312 (2019). Doi: https://doi.org/10.14773/CST.2019.18.6.312
8. D. C. Montgomery, *Design and Analysis of Experiments*, *10th ed.*, Wiley, Hoboken, NJ (2020).
9. A. I. Khuri and S. Mukhopadhyay, Response surface methodology, *WIREs Computational Statistics*, **2**, 128 (2010). Doi: https://doi.org/10.1002/wics.73
10. L. Coelho, D. Zhang, Y. Van Ingelgem, D. Steckelmacher, A. Nowé, and H. Terryn, Reviewing machine learning of corrosion prediction in a data-oriented perspective, *npj Materials Degradation*, **6**, 8 (2022). Doi: https://doi.org/10.1038/s41529-022-00218-4
11. S. R. Kalidindi and M. De Graef, Materials Data Science: Current Status and Future Outlook, *Annual Review of Materials Research*, **45**, 171 (2015). Doi: https://doi.org/10.1146/annurev-matsci-070214-020844
12. J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, Recent advances and applications of machine learning in solid-state materials science, *npj Computational Materials*, **5**, 83 (2019). Doi: https://doi.org/10.1038/s41524-019-0221-0
13. S. Lundberg and S.-I. Lee, A Unified Approach to Interpreting Model Predictions, *arXiv preprint* (2017). Doi: https://doi.org/10.48550/ARXIV.1705.07874
14. Z. Zhang, Introduction to machine learning: k-nearest neighbors, *Annals of Translational Medicine*, **4**, 218 (2016). Doi: https://doi.org/10.21037/atm.2016.03.37
15. S. Maldonado, R. Weber, and J. Basak, Simultaneous feature selection and classification using kernel-penalized support vector machines, *Information Sciences*, **181**, 115 (2011). Doi: https://doi.org/10.1016/j.ins.2010.08.047
16. G. Biau and E. Scornet, A random forest guided tour, *TEST*, **25**, 197 (2016). Doi: https://doi.org/10.1007/s11749-016-0481-7
17. T. Chen and C. Guestrin, XGBoost: A Scalable Tree Boosting System, *Proc. 22nd ACM SIGKDD Conf.*, p. 785, ACM, San Francisco, CA (2016). Doi: https://doi.org/10.1145/2939672.2939785
18. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, Scikit-learn: Machine Learning in Python, *arXiv preprint* (2018). Doi: https://doi.org/10.48550/arXiv.1201.0490
19. F. Mansfeld, Electrochemical impedance spectroscopy (EIS) as a new tool for investigating methods of corrosion protection, *Electrochimica Acta*, **35**, 1533 (1990). Doi: https://doi.org/10.1016/0013-4686(90)80007-B
20. F. Batmanghelich, M. B. Hariri, S. Sharifi-Asl, Y. Yaghoubinezhad, G. Mortazavi, and Y. Seo, Corrosion and time dependent passivation of Al 5052 in the presence of H_2O_2, *Metals and Materials International*, **22**, 609

(2016). Doi: https://doi.org/10.1007/s12540-016-5699-0
21. H. H. Veselivska, V. M. Hvozdetskyi, M. M. Student, Kh. R. Zadorozhna, and Yu. V. Dzioba, The Influence of the Electrolyte Composition for Hard Anodizing of Aluminum on Corrosion Resistance of Synthesized Coatings, *Materials Science*, **59**, 228 (2023). Doi: https://doi.org/10.1007/s11003-024-00767-w