JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 31 No. 10, pp. 281-288, October 2026
https://doi.org/10.9708/jksci.2026.31.10.281

Performance Evaluation of an AI Classification Model for Predicting Laser Cladding Process Outcomes

DongGug Yoo*

*Professor, Industry-Academic Cooperation Foundation, Sunchon National University, Suncheon, Korea

[Abstract]

This study re-evaluated five classifiers for crack occurrence in the supplied laser-cladding dataset (1,000 records; 500 crack-positive and 500 crack-negative cases). A 12-predictor model including post-process coating properties was compared with a six-predictor model restricted to candidate pre-process inputs. The same stratified 700/300 training/test split and five training folds were used across all analyses. Imputation, normalization and dummy encoding were estimated within each training fold. Model selection used cross-validated ROC AUC, and a common probability threshold of 0.5 was applied. The cross-validation-selected pre-process SVM achieved a test AUC of 0.848 (95% CI 0.806–0.890), accuracy of 0.740, and F1-score of 0.738. The 12-predictor unpenalized logistic model achieved an AUC of 0.996 and accuracy of 0.983, but exhibited complete separation in the training data and nonconvergence in the original fitting configuration. The performance gap demonstrates the importance of predictor availability at the intended prediction time; it does not establish causal effects or industrial deployment validity. Dataset provenance, label acquisition, sample independence and external validation remain necessary before operational conclusions can be drawn.

▸ **Key words**: Laser cladding, crack classification, predictor timing, data leakage, cross-validation

[요 약]

본 연구는 제공된 레이저 클래딩 데이터 1,000건(균열 Yes/No 각 500건)을 이용하여 5개 분류모델을 재평가하였다. 공정 후 측정특성을 포함한 전체 12개 변수 모델과 사전 확보 가능한 후보 6개 변수 모델을 구분하고, 동일한 계층화 학습/테스트 분할(700/300건)과 학습 세트 내부 5-fold 교차검증을 적용하였다. 결측치 대치, 표준화 및 더미 변수화는 각 학습 폴드에서만 추정하였으며, 분류 임계값은 모든 모델에 0.5를 적용하였다. 교차검증으로 선택된 사전 입력 SVM은 테스트 AUC 0.848(95% 신뢰구간 0.806–0.890), 정확도 0.740, F1 0.738을 보였다. 전체 변수 비정규화 로지스틱 회귀는 AUC 0.996, 정확도 0.983이었으나, 학습 자료에서 완전분리와 기본 적합의 비수렴이 확인되어 안정적인 계수 추론의 근거로 사용하지 않았다. 이러한 차이는 공정 후 특성을 이용한 분류 성능을 공정 전 예측 성능으로 해석해서는 안 됨을 보여준다.

▸ **주제어**: 레이저 클래딩, 균열 분류, 예측변수 확보 시점, 데이터 누수, 교차검증

• First Author: DongGug Yoo, Corresponding Author: DongGug Yoo
*DongGug Yoo (arahan4you@scnu.ac.kr), Industry-Academic Cooperation Foundation, Sunchon National University
• Received: 2026. 08. 10, Revised: 2026. 10. 05, Accepted: 2026. 10. 11.

Copyright © 2026 The Korea Society of Computer and Information
http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945
www.kci.go.kr

I. Introduction

금속 가공 표면처리 공정 단계의 일부로 활용되는 레이저 클래딩 공정은 금속 표면에 분말 소재를 레이저를 활용하여 내마모성과 표면 장력의 활용성이 뛰어난 코팅층을 형성하는 첨단 제조 기술이다. 레이저 클래딩 기술의 활용 요소를 살펴보면, 자동차, 전력에너지, 항공우주 등 핵심 산업 분야에서 부품의 내마모성 및 내식성 등 표면 성능을 향상시키는 데 사용되는 첨단 표면 처리 기술을 말한다[1].

그러나 공정상의 특성을 고려하여 레이저 출력, 주입 속도, 빔 크기 등 다수의 공정 변수(x, 퓨쳐 값)에 의해 용융풀의 온도 구배 및 응고 속도가 결정되며, 그 결과 코팅층에 필연적으로 잔류 응력이 발생 한다[2]. 이로 인해 잔류 응력이 재료의 항복 한계를 초과할 경우 균열(y, 타깃)이 발생하여 코팅의 기계적 안정성과 수명을 심각하게 저하 영향을 시키므로, 금속 표면 공정 조건을 세밀하게 최적화 하여 균열(타깃 변수)발생을 사전에 방지하는 것이 제조 품질 확보하고 예지 보존 상태를 위한 최우선 과제로 인식되고 있다[3]. 이와 관련하여 선행 연구의 특징을 살펴보면, 실험자의 경험이나 통계적 실험계획법(DOE)을 통해 최적화 모형을 구현하여, 금속 표면 처리 작업 공정 조건에 제공하였다. 이러한 전통적인 접근 방식은 해당 변수의 특성이 복잡하게 얽힌 공정 단위 퓨쳐들 간의 비선형적 상관관계를 파악하는 데 한계가 명확 할 수밖에 없는 단점이 초래된다. 가장많은 부분의 문제는 수많은 공정 데이터 변수 조합을 모두 실험해야 하므로 시간과 고비용을 요구이다. 최근 인공지능의 예측 모델을 기반으로 제조 현장의 데이터 특성 간의 난제를 해결하기 위해 공정 데이터 기반의 인공지능(AX) 및 딥러닝(DL) 기법이 비선형 관계를 효율적으로 모델링하고, 높은 분석 정확도를 제공하는 최적화 모델이 연구되고 있다[4]. 선행 연구를 기반으로 본 연구는 레이저 클래딩 공정 데이터의 균열 발생 예측과 공정 최적화 모형을 지원하는 모델구축을 목표로, 머신러닝 모델 성능평가를 기반한 평가 모델을 제안한다.

본 연구에 사용한 T100.csv는 전라남도 순천시에 소재한 ㈜미르테크가 광양제철소 관련 레이저 클래딩 업무를 수행하는 과정에서 축적한 공정 및 품질관리 자료를 바탕으로 연구용으로 구성된 데이터셋이다. 포스코 협력 과정에서 레이저 클래딩 공정 자료를 축적하였으며, 원자료가 13개 항목과 30,000개 공정 기록으로 구성되어 있으며, 본 연구가 제공받은 T100.csv는 이 가운데 1,000개 행과 13개 열로 정리된 분석 자료이다. 본 연구에서는 각 행을 독립 시편으로 단정하지 않고 하나의 공정 관측 기록으로 취급하였다. 수집 및 생성 과정은 레이저 출력, 주입속도, 빔 크기, 분말 및 기판 조건과 공정 후 코팅 특성, 균열 발생 여부가 함께 기록되어 있다. 공정 조건과 재료 조건은 작업 기록에서, 층두께·밀착력·경도·기공률·내마모성·내식성은 공정 후 품질 측정 기록에서 정형화된 것으로 판단된다. 이 구분은 변수의 측정 시점과 데이터 누수 가능성을 평가하기 위한 것이며, 구체적인 센서·시험기, 측정 표준, 측정 횟수와 집계 방법은 원자료 제공자의 작업표준서에서 확인되어야 한다. 분석에서는 공정 후 측정값을 포함한 Full12와 공정 전에 확보 가능한 후보만 사용한 Pre6를 분리하였다. 높은 예측 성능과 우수한 해석력을 제공하는 Logistic Regression, XGBoost (Extreme Gradient Boosting), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest, 분류 알고리즘을 적용하였다. 연구를 통하여, 공정 변수와 코팅층 특성 데이터(x)를 입력으로 하여 균열 발생 여부(y)를 정확하게 예측하는 분류 모델을 구축하고, 모델 기반의 순열 중요도(permutation importance) 분석을 통해 각 모델의 예측에 기여한 변수를 정량적으로 비교하였다. 이 중요도는 모델의 예측 기여도이며, 균열 발생에 대한 인과 효과를 의미하지 않는다[4]. 본 연구의 결과는 레이저 클래딩 공정 데이터를 기반으로 공정 이해도를 높이고, 궁극적으로 데이터 기반 예측 모델링을 통하여, 스마트 제조 시스템의 구축과 제조 데이터 품질 지수 관리 모델에 정량적인 가이드라인을 제공할 것으로 기대된다.

II. Preliminaries

1. Related works

1.1 Trend in laser cladding

실무 공정 데이터 기반의 레이저 클래딩 공정에서 발생하는 균열 결함을 최소화하기 위한 연구는 국내외적으로 중요한 영역이다. 근본적인 원인으로 균열 발생은 코팅층 내부에 누적되는 잔류 응력이며, 분류를 하면, 열응력, 조직응력, 구속응력의 세 가지 주요 범주로 구분된다[2]. 특이점으로는, 응력은 기판과 피복재의 열팽창 계수 차이, 용융층의 고화 과정에서의 미세구조 변화, 또는 외부 기계적 구속 등에 의해 복합적으로 작용하여 코팅 품질을 저하시킨다.

초기 제어 방법으로는 예열, 서냉 등의 열처리 기법이나 적층 패턴 변경을 통해 잔류 응력을 완화시키는 사후 처리가 주도 하였다, 그러나, 추가 비용이나 생산성 저해를 문

www.kci.go.kr

제로, 공정 변수를 사전에 최적화하여 균열 발생 자체를 예방하는 예지보전의 궁극적인 목표로 전환되었다[3].

예지보전할 선행적 제어를 위해 최근 국내외적으로 실험계획법 (DOE)과 최적화 알고리즘을 결합하는 연구가 활발하다. 특히, 국내 산업 환경에서는 제조 혁신의 일환으로 스마트 팩토리 도입이 가속화됨에 따라, 복잡한 비선형 공정의 품질 지표를 예측하고 공정 변수를 선정하는 데이터 기반의 접근법이 필수적인 방법론으로 대두되고 있다. 머신 러닝 기법으로 데이터를 학습하여 철강 표면 클래딩의 매개변수를 신속하게 최적화하고, 소수 검증용 데이터만으로도 다중 품질 관리 지표(경도, 기공도 등)를 높은 정확도로 예측하는 연구가 검증 효율성을 입증하는 추세다.

모델의 고도화를 적용하지 않은 연구, 머신러닝 적용 초기에는 인공신경망 (ANN)이나 단순 유전 알고리즘 등 단일 기법에 의존하는 경향이 있어, 예측 모델의 해석력이나 새로운 데이터에 대한 일반화 성능에 한계를 보였다[5,7(6)]. 이러한 단점을 극복하고 AI 모델의 연구 방법론 신뢰성을 확보하기 위해, 최근에는 XGBoost (Extreme Gradient Boosting)종류의 트리 기반 앙상블 학습기가 주목받고 있다. 뛰어난 분류 예측 성능을 제공하며, 변수 중요도 분석을 통해 모델의 다수결 원칙이 적용된 의사결정 논리를 정량적으로 도출함으로써, 데이터를 기반한 변수간의 상관관계 및 인과관계가 레이저 클래딩 공정 데이터에서 효율향상이라는 부가적인 목표를 달성하고 있다[8(7)].

따라서, 본 연구는 이러한 국내외 연구 동향을 바탕으로, 분류 예측 모델의 성능 평가를 중심으로 XGBoost를 포함한 5개 분류 알고리즘을 동일한 조건에서 학습하고, 모델 간 성능 비교 및 시각화를 통하여 레이저 클래딩 공정의 균열 발생 여부(타깃 변수)를 최적화 하여 분류 예측하는 모델을 제시하는 데 주안점을 두었다[6(8)]. 따라서, AI 모델의 분류 예측 성능을 정량적으로 검증하고, 주요 변수의 예측 기여를 비교하여 데이터 기반 공정 최적화의 실무 가이드 적용 가능성을 제시하는 데 기여 할 것이다.

III. The Proposed Scheme

1. Data General Status and Preprocessing

1.1 Datasets and preprocessing

분석 대상은 제공된 T100.csv의 1,000행, 13열이다. 입력 12개와 균열 발생 여부(Yes/No)로 구성되며 각 클래스는 500건이다. 원자료의 수집·생성 과정, 균열 판정 방법, 시편·배치별 반복 측정 여부는 확인되지 않았다. 따라서 독립적인 실제 공정실험 1,000건으로 단정하지 않는다. 원자료에는 빈 값이 없지만 기공율에 음수 3건이 관측되었다. 음수 기공율은 비율의 정의에 따라 분석 사본에서 결측 처리하고 학습 폴드 내 중앙값으로 대치하였다. 층두께·경도·기공율·밀착력·내마모성·내식성은 공정 후 측정특성이므로, 이를 포함한 분류 성능은 사전 예측 성능과 구분한다. 본 연구의 독립변수 중 경도(HV), 기공율(%), 밀착력(MPa) 등은 공정 종료 후 측정되는 코팅 '결과' 특성으로, 균열 발생과 동시 또는 사후에 결정될 가능성이 있어 예측 시점 기준으로 데이터 누수(data leakage) 소지가 있다[9]. 이에 따라 사전 입력 후보는 레이저 출력, 주입속도, 빔 크기, 입자 크기, 분말 혼합비 및 기판 재질의 6개로 한정하였다. 층 두께는 표 1의 분류처럼 공정 후 측정값이므로 이 모델에서 제외 하였다. 결측치 대치와 표준화는 학습 폴드에서 추정하였다.

Table 1. Input Variables and Observed Ranges in the Original Dataset

Variable	Unit*	Observed range or level	Timing classification
Laser power	W	1379.366-4926.366	Candidate pre-process input
Feed rate	mm/s*	5.000-15.000	Candidate pre-process input
Beam size	mm	2.000-8.000	Candidate pre-process input
Layer thickness	μm	103.528-412.155	Post-process measurement
Particle size	μm	0.470-19.339	Candidate pre-process input
Adhesion strength	MPa	31.005-90.983	Post-process measurement
Hardness	HV	793.108-1729.358	Post-process measurement
Porosity	%	-0.451-4.994	Post-process measurement
Wear resistance	Relative index	0.100-1.000	Post-process measurement
Corrosion resistance	Relative index	0.100-1.000	Post-process measurement
Powder mixture	Category	75Cr_3C_2-25NiCr, 80WC-20Cr, 90W_2C-10Ni	Candidate pre-process input
Substrate material	Category	Inconel, Stainless, Titanium	Candidate pre-process input

타깃 변수인 균열 발생 여부는 균열 발생 시 "Yes", 미발생 시 "No"로 기록되었으며, 전체 1000개 데이터 중 Yes와 No가 각각 500개로 1:1의 균형 잡힌 클래스 비율(Class Balance)을 가진다. 이는 공정 데이터 기반의 모델 학습 시 특정 클래스에 편향되지 않고 안정적으로 분류 경계를 학습할 수 있는 유리한 조건으로 활용하고, 모델의 AUC, ROC

www.kci.go.kr

Curve 성능 평가를 위해 caret 패키지를 사용하였으며, 분석에서는 Yes 클래스를 긍정 클래스(Positive Class)로 지정하여 이진 분류 예측 모델링 분석을 적용하였다.

본 연구에서는 데이터셋을 5가지 주요 분류 알고리즘(Logistic Regression, SVM, Random Forest, KNN, XGBoost)에 적용하기 위해 다음과 같은 전처리 및 데이터 분할을 체계적으로 수행하였다.

결측 처리와 표준화, 범주형 변수의 더미화는 분할 후 각 학습 폴드 안에서 추정하고 검증 폴드와 테스트 자료에 적용하였다.

데이터 분할은 전체 1000 개 중 70%(700개)는 모델 학습 및 교차검증에 사용하였고, 나머지 30% (300개)는 최종 성능 평가를 위한 홀드아웃 테스트 세트로 사용하였다. 이 분할은 균열 발생 비율의 불균형이 없도록 계층적 추출(stratified sampling) 방법을 적용하여 학습 및 테스트 세트 모두 1:1 비율을 유지하도록 구성하였다.

변수 변환 (전처리)는 수치형 변수는 단위와 범위가 상이한 수치형 변수(예: 레이저출력, 층두께, 밀착력 등)는 학습 안정성 및 각 변수의 기여도 균형을 위해 평균, 분산의 스케일로 표준화 (Center and Scale)하는 과정 실행.

범주형 변수는 기판 재질과 같은 범주형 변수는 R의 caret_패키지 train_함수 내에서 One Hot Encoding 형태의 더미 변수화가 학습 시 자동으로 수행되도록 하였다.

최종 데이터셋 구성으로 중점화 한 것은 표준화된 예측변수 (X)와 팩터 타입의 목표 변수 (Y)를 재결합하여 train_data와 test_data를 최종적으로 구성하였으며, 이는 5가지 알고리즘의 공통 입력 데이터로 사용되었다.

1.2 Experimental configuration and model learning

본 연구의 핵심 목표는 최적의 성능과 해석력을 갖는 분류 예측 모델을 구축하는 것이다. 이를 위해 우수한 일반화 성능과 Feature Importance 해석 능력이 검증된 R 기반 Logistic Regression, XGBoost(Extreme Gradient Boosting), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest, 분류 알고리즘 라이브러리를 사용하여 균열 발생 분류 모델을 적용하였다. 모델의 신뢰성과 성능을 극대화하기 위해 다음과 같은 전략적 학습 및 평가 단계를 진행했다.

교차 검증 및 학습 제어 설정으로 모델의 과적합을 방지 목적과 모델 성능 지표의 객관적으로 평가하기 위해 “R의 caret패키지”를 활용한 5-fold 교차 검증 전략을 채택하여 적용하였다. 학습 제어 설정 (fitControl)은 다음과 같은 주요 요소를 포함한다.

- method="cv", number=5, 5개의 폴드(Fold)로 나누어 학습 및 검증을 반복하는 교차 검증 방식 적용했다.
- metric="ROC" 모델의 성능 지표를 ROC 곡선 아래 면적 (AUC) 기준으로 평가하도록 설정하여, 분류기의 전반적인 판별 능력을 최적화하였다.
- summaryFunction=twoClassSummary, AUC 및 Sensitivity (Recall), Specificity, Recall의 보수를 계산하도록 지정하여, 이진 분류 모델의 성능을 다각도로 평가하였다.
- classProbs=TRUE: 모든 모델이 예측 결과로 클래스 확률값을 산출하도록 강제하여, AUC 계산의 필수 조건을 충족시켰다.

Table 2. Summary of Classification Algorithms and Features

Algorithm	Function/Instruction	Features
Logistic Regression	glm (family="binomial")	As the most basic classifier assuming linearity, it provides linear separability of data and standard performance.
SVM	svmRadial	A non-linear classifier utilizing kernel tricks, evaluating the ability to model complex decision boundaries.
KNN	knn	A non-parametric classifier, evaluating performance that reflects the local structure of the data.
RF	ranger	A Bagging-based ensemble model, evaluating stable performance through variance reduction capabilities.
XGBoost	xgbTree	A Boosting-based ensemble model evaluated under the same protocol as the other four models; it showed a high test AUC with Full12 (0.985) but the lowest test AUC with Pre6 (0.808).

표2에서 분류 알고리즘 및 하이퍼파라미터 구성으로 연구의 밀집도와 가중치 부분을 고려한 선형성, 비선형성, 앙상블 기반의 특성을 고루 갖춘 총5가지 대표 분류 알고리즘을 선정하고 학습을 수행하여,분류 모델의 성능을 객관적인 관점에서 비교 분석하였다. 특정 알고리즘을 핵심 모델로 미리 정하지 않았으며, 각 입력 구성의 모델은 교차검증 AUC로 선택하였다. XGBoost 하이퍼파라미터 상세 설정으로 살펴볼 모델중 XGBoost의 성능 극대화 및 과적합 방지를 위해, 교차 검증을 통해 하이퍼파라미터를 조정하고, 부분 샘플링 비율 (subsample) 0.5는 각 트리가

www.kci.go.kr

무작위로 절반의 데이터만 사용하게 하여 과적합을 제어하는 핵심적인 정규화 역할을 수행했다.XGBoost 학습 시에는 verbose = FALSE로 설정하여 메시지 출력을 억제하고 학습 효율을 높였다.

1.3 Supplementary Methods: Reproducibility and Statistical Evaluation

난수 시드 123으로 계층화 70:30 분할을 수행했다. 학습 700건과 테스트 300건은 각각 Yes/No가 350/350건, 150/150건이다. 모든 입력 구성과 모델에 같은 행 인덱스를 사용했다. 학습 세트의 동일한 5개 검증 폴드에는 각 140건(Yes/No 각 70건)을 배정했다. 테스트 자료는 모델·하이퍼파라미터·임계값 선택에 사용하지 않았다.

표3은 각 교차검증 학습 폴드에서 중앙값 대치, 영분산 연속변수 제거, 평균 0·표준편차 1 표준화, 새 범주 및 결측 범주 처리, K−1 더미 변수화, 영분산 입력 제거를 순서대로 적합했다. 전처리 recipe를 caret::train에 전달해 검증 폴드 정보가 학습에 사용되지 않도록 했다. 최종 모델의 전처리는 전체 학습 세트에 다시 적합한 뒤 테스트 세트에 적용했다.

Table 3. Hyperparameter search spaces

Model	Search space / fixed settings
Logistic	glm, binomial; unpenalized; maxit=25
SVM	svmRadial; sigma={0.01,0.05,0.1}; C={0.25,1,4,16}; scaling=FALSE
RF	ranger; mtry={2,4,6}; min.node.size={1,5,10}; gini; trees=500; threads=1
KNN	k={3,5,7,9,11,13,15}
XGBoost	rounds=500; depth={2,3,4}; eta=0.1; gamma=0; colsample={0.8,1}; min_child_weight=1; subsample=0.5; binary:logistic; threads=1

각 5개 폴드의 평균 ROC AUC를 최대화하여 튜닝과 주 모델을 선택했다(Table 5·6의 CV AUC). 모든 모델에서 P(Yes)≥0.5를 Yes로 분류했다. SVM 역시 기본 클래스 반환값이 아니라 예측확률에 이 임계값을 적용했다. Accuracy, Precision, Recall, Specificity, F1은 동일한 테스트 혼동행렬로 계산했다. 이 임계값의 현장 최적성이나 확률 보정은 확인되지 않았다.

AUC는 Yes 확률이 클수록 양성인 방향으로 계산했다. AUC의 95% 신뢰구간에는 DeLong 방법, 정확도 구간에는 정확 이항 방법을 사용했다. 정밀도·재현율·특이도·F1 구간은 고정된 테스트 예측에서 클래스별 층화 부트스트랩 2,000회로 산출했다. 학습과 튜닝을 재수행하는 불확실성은 이 구간에 포함되지 않는다.

Table 4. Selected values for Full12 and Pre6

Model	Full12	Pre6
Logistic	parameter=none	parameter=none
SVM	sigma=0.01; C=16	sigma=0.01; C=0.25
RF	mtry=2; gini; min.node.size=1	mtry=2; gini; min.node.size=10
KNN	k=15	k=15
XGBoost	rounds=500; depth=3; eta=0.1; gamma=0; colsample=0.8; min_child=1; subsample=0.5	rounds=500; depth=4; eta=0.1; gamma=0; colsample=1; min_child=1; subsample=0.5

같은 테스트 행의 모델별 ROC AUC는 paired DeLong 검정으로 비교하고 다중 비교에는 Holm 보정을 사용했다. 비교는 각 입력 구성 안에서 교차검증 AUC로 선택된 모델과 나머지 4개 모델 사이에 수행하였고(입력 구성별 4회), Holm 보정은 입력 구성별로 적용하였다. 비교 대상, AUC 차이, z 통계량, 원 p값과 Holm 조정 p값은 Table 8에 제시하였다. 이는 외부 검증이 아닌 동일 데이터 내부 비교다.

1.4 Performance Evaluation Process and Outcome Analysis Strategy

정량적 성능 지표 산출은 평가 지표는 이진 분류 모델의 표준 척도를 사용했으며, AUC를 중심으로 정확도, 정밀도, 재현율, 특이도, F1을 산출했다.

- 정확도 (Accuracy): 전체 예측 데이터 중 맞게 분류된 비율.
- 정밀도 (Precision): TP/(TP+FP)이다. 거짓 양성률 FP/(FP+TN)과 구별하며, 실제 오경보 빈도는 운영 자료의 균열 비율과 임계값에 따라 달라진다.
- 재현율 (Recall / Sensitivity): 실제 균열 발생 샘플 중 모델이 놓치지 않고 검출해낸 비율 (TP/(TP+FN)). 누락률을 검증하는 척도이다.
- 특이도 (Specificity): 실제 비균열 샘플 중 비균열로 분류한 비율 (TN/(TN+FP)).
- F1: 정밀도와 재현율의 조화평균 (2TP/(2TP+FP+FN)).
- AUC (Area Under the ROC Curve) FPR에 대한 TPR의 곡선 면적으로, 분류기의 종합적인 판별력을 나타내는 핵심 지표이다.

1.5 Permutation Importance

변수 중요도는 검증 폴드 순열 중요도(validation-fold permutation importance)로 산출하였다. 사용 지표는 ROC AUC 감소량이다. 즉, 검증 폴드에서 한 변수의 값을 무작위로 뒤섞은 뒤 AUC를 다시 계산하고, 원래 AUC에

www.kci.go.kr

서 뺀 값을 그 변수의 중요도로 하였다. 대상 모델은 트리 기반 앙상블인 RF와 XGBoost이며, 각 입력 구성(Full12, Pre6)에서 Table 4의 최종 하이퍼파라미터로 각 학습 폴드에 적합한 모델을 사용하였다. 범주형 변수는 더미 변수가 아니라 원변수(original variable) 단위로 순열하였다. 순열은 검증 폴드마다 10회 반복하여 평균하였고, 5개 검증 폴드의 평균과 표준편차를 Fig. 3에 제시하였다. 테스트 자료는 중요도 산출에 사용하지 않았다. 이 지표는 해당 모델의 예측 기여를 나타내며, 변수의 영향 방향이나 인과 효과를 뜻하지 않는다.

Table 5. Full-predictor test performance (Full12; n=300) Model

Model	AUC	Accuracy [95% CI]	Precision	Recall	Specificity	F1
Logistic	0.996	0.983 [0.962, 0.995]	0.987	0.980	0.987	0.983
SVM	0.995	0.967 [0.940, 0.984]	0.961	0.973	0.960	0.967
KNN	0.934	0.853 [0.808, 0.891]	0.858	0.847	0.860	0.852
RF	0.958	0.880 [0.838, 0.915]	0.885	0.873	0.887	0.879
XGBoost	0.985	0.947 [0.915, 0.969]	0.947	0.947	0.947	0.947

Table 6. Pre-predictor test performance (Pre6; n=300)

Model	AUC	Accuracy [95% CI]	Precision	Recall	Specificity	F1
Logistic	0.848	0.730 [0.676, 0.779]	0.738	0.713	0.747	0.725
SVM	0.848	0.740 [0.686, 0.789]	0.743	0.733	0.747	0.738
KNN	0.820	0.720 [0.666, 0.770]	0.732	0.693	0.747	0.712
RF	0.817	0.740 [0.686, 0.789]	0.743	0.733	0.747	0.738
XGBoost	0.808	0.733 [0.679, 0.783]	0.754	0.693	0.773	0.722

12개(Full 12) 예측변수를 사용한 분석에는 공정 후에 측정한 코팅층 특성이 포함되어 있다. 로지스틱 회귀의 AUC는 0.996, 정확도는 0.983이었으며, SVM의 AUC는 0.995, 정확도는 0.967이었다. 이 값들은 공정이 끝난 뒤 확보할 수 있는 측정값을 이용한 분류 성능을 나타낸다. 벌점을 적용하지 않은 로지스틱 회귀에서는 완전분리와 비수렴이 발생했으므로, 그 계수를 안정적인 효과 크기로 해석해서는 안 된다. 정밀도는 TP/(TP+FP)이고 거짓 양성률은 FP/(FP+TN)로, 두 지표는 서로 다르다. Table 5·6, Fig. 1(a)·(b) 및 Fig. 2(a)·(b)는 모두 동일한 분석 결과이다. 다만 Full12의 결과는 공정 후 측정값을 포함하므로, 이 결과만으로 공정 조건의 변경이 균열 발생에 미치는 효과를 입증할 수 없다.

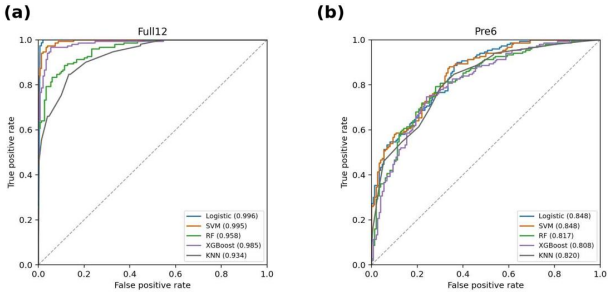


Fig. 1. ROC curves using (a) Full12 and (b) Pre6 predictors

그림1의 ROC 곡선은 전체 변수(Fig. 1(a))와 사전 입력 후보(Fig. 1(b))의 평가를 구분한다. 전체 변수의 높은 판별력은 공정 후 측정값을 포함한 결과이므로 공정 전 예측 성능으로 해석할 수 없다. Table 5·6의 테스트 AUC를 기준으로 각 입력 구성의 모델을 비교하였다. 각 곡선의 AUC와 95% 신뢰구간 수치는 Table 5·6에 제시하였다.

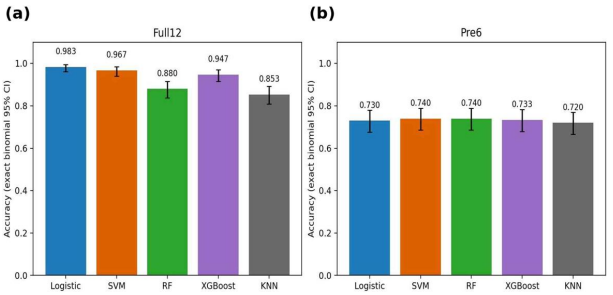


Fig. 2. Test accuracy using (a) Full12 and (b) Pre6 predictors

그림 2는 전체 변수(Fig. 2(a))와 사전 입력 후보(Fig. 2(b))의 테스트 정확도 및 정확 이항 95% 신뢰구간을 나타낸다. 모든 모델에 P(Yes)=0.5를 적용하였다. 수치는 Table 5·6의 Accuracy [95% CI] 열과 같다.

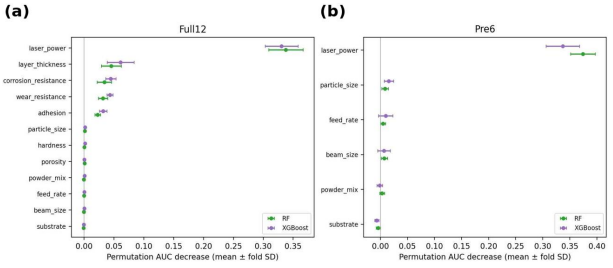


Fig. 3. Validation-Fold Permutation Importance by Original Variable ((a) Full12, (b) Pre6)

www.kci.go.kr

그림 3의 점(point)은 5개 검증 폴드 평균 AUC 감소량이며, 가로 막대는 폴드 간 표준편차이며 신뢰구간이 아니다. 각 폴드에서는 10회 순열을 평균하였다. 산출 방법은 III장 1.5절에, 수치는 Table 7에 제시하였다.

Full12(Fig. 3(a))에서는 RF와 XGBoost 모두 레이저 출력의 AUC 감소량이 가장 컸다(약 0.32–0.33). 그다음은 층두께, 내식성, 내마모성, 밀착력 순이었으며(약 0.02–0.06), 모두 공정 후 측정값이다. 입자 크기, 경도, 기공율, 분말 혼합비, 주입속도, 빔 크기, 기판 재질은 0에 가까웠다. Pre6(Fig. 3(b))에서도 레이저 출력이 가장 높았고(RF 약 0.37, XGBoost 약 0.33), 입자 크기(약 0.01–0.02)가 뒤를 이었다. 주입속도와 빔 크기는 0.01 미만이었고, 분말 혼합비와 기판 재질은 0 이하였다. 즉, 두 입력 구성 모두에서 RF·XGBoost의 예측 기여는 레이저 출력에 집중되었다. 다만 중요도는 RF와 XGBoost에 대해서만 산출하였으므로, Pre6에서 선택된 SVM의 예측 근거를 직접 설명하지는 않는다.

IV. Conclusions

1. Analysis Results and Interpretation

1.1 Experimental Configuration and Model Learning

본 연구는 공정 후 측정값을 포함한 전체 12개 변수의 분류와 사전 입력 후보 6개 변수의 분류를 구분했다. 전체 변수의 높은 성능은 공정 전 균열 예측 또는 공정 최적화의 근거가 아니다. 사전 입력 후보 SVM의 테스트 AUC는 0.848, 정확도는 0.740이다. 전체 변수 비정규화 로지스틱 회귀의 AUC는 0.996, 정확도는 0.983이었으나 완전분리와 비수렴으로 계수의 효과 크기를 해석하지 않았다.

동일한 계층화70:30 분할을 사용했다. 학습 세트 내부 5-fold 교차검증으로 모델을 선택하고 홀드아웃 300건에서 최종 성능을 산출했다. 결측치 대치, 표준화 및 범주형 변수의 더미화는 각 학습 폴드에서 추정했으며, 분류 임계값은 0.5였다. 테스트 자료는 외부 검증 자료가 아니다. 결과는 제공된 자료 내부의 분류 성능으로 한정된다. 두 입력 집합의 차이는 입력 확보 시점의 차이를 보여주지만 공정 조건 변경의 효과를 뜻하지 않는다.

2. Test performance and limitations

모든 모델은 동일한 계층화 학습·테스트 분할(700/300건)과 학습 세트 내부 5-fold 교차검증으로 평가하였다. 전체 12개 변수(Full12)에는 공정 후 측정값이 포함되며, 사전 입력 후보 6개 변수(Pre6)는 공정 전에 확보 가능한 정보만 사용한다. Full12 SVM의 테스트 AUC와 정확도는 각각 0.995와 0.967이었고, Pre6 SVM은 각각 0.848과 0.740이었다. 두 결과는 입력 시점이 달라 같은 목적의 성능으로 해석할 수 없다.

Full12 로지스틱 회귀의 테스트 혼동행렬은 TP 147, FP 2, FN 3, TN 148이며, 정밀도 0.987, 재현율 0.980, 특이도 0.987, F1 0.983이다. 정밀도는 TP/(TP+FP), 거짓 양성률은 FP/(FP+TN)이다. 이 모델의 거짓 양성률은 2/(2+148)=0.013이다. 이러한 값은 공정 후 측정값을 포함한 내부 테스트 결과이며, 현장의 오경보 빈도나 공정 전 경고 성능을 보장하지 않는다.

Pre6 SVM의 혼동행렬은 TP 110, FP 38, FN 40, TN 112이며, 정밀도 0.743, 재현율 0.733, 특이도 0.747, F1 0.738이다. 균열 양성 150건 중 40건을 놓쳤고, 비균열 150건 중 38건을 양성으로 분류하였다. 시험 자료의 균열 비율은 1:1이고 분류 임계값은 0.5이므로, 실제 균열 빈도가 다른 생산 환경에 이 수치를 그대로 적용할 수 없다. F1 역시 오경보와 균열 누락의 운영 비용을 반영하지 않는다.

2.1 Evaluation and Future Research

그림 3의 순열 중요도는 검증 폴드에서 변수를 교란했을 때의 AUC 감소량으로, 해당 모델의 예측 기여를 나타낸다[11]. 변수의 영향 방향이나 인과 효과를 뜻하지 않는다. 층두께와 내마모성은 공정 후 측정값이므로 사전 제어 변수로 취급할 수 없으며, 레이저 출력의 변경이 균열을 감소시킨다는 결론도 현재 분석만으로는 도출할 수 없다.

원자료의 수집·생성 과정, 균열 판정 기준, 시편·배치별 반복 측정 여부가 확인되지 않았다. 따라서 행 단위 분할의 낙관 편향 가능성이 남아 있고, 현재 성능은 단일 데이터셋의 내부 평가로 한정된다. 향후 시편·배치 식별자를 확인하여 반복 측정이 있으면 그룹 단위로 재분할하고, 독립 공정·시간대의 자료로 외부 검증해야 한다. 실제 균열 빈도와 오경보·누락 비용을 반영하여 임계값과 확률 보정을 평가한 뒤, 공정 조건 변경의 효과는 별도 실험으로 검증해야 한다. 공정 최적화 시스템 통합으로 본 연구에서 검증된 AI 분류 모델을 실시간 센싱 데이터와 연동된 공정 모니터링 시스템에 통합하여, 균열 발생 위험을 실시간으로 예측하고 피드백 제어를 통해 공정 변수를 자동 조정하는 지능형 제조 시스템 구축에 기여한다.

www.kci.go.kr

ACKNOWLEDGEMENT

This paper was supported by Sunchon National University Research Fund in 2025.(Grant number: 2025-0313)

REFERENCES

[1] Y. Feng et al., “Influence of residual stress on the cracking of the laser-cladded high-hardness coating” Materials Letters, 349, 134715, 2023. DOI: 10.1016/j.matlet.2023.134715.

[2] Y. Cao et al., “Laser Cladding Performance and Process Parameter Optimization for Fe90 Alloy,” Metals, 14(12), 1432, 2024. DOI: 10.3390/met14121432

[3] C. Zhao et al., “Study on the reduction of residual stress in laser cladding layers through groove texture,” Scientific Reports, 14, 2024. DOI: 10.1038/s41598-024-66793-5

[4] B. L. Chua, H. J. Lee, D. G. Ahn, J. G. Kim, “Influence of process parameters on temperature and residual stress distributions of the deposited part by a Ti-6Al-4V wire feeding type direct energy deposition process,” Journal of Mechanical Science and Technology, 32, 5363-5372, 2018. DOI: 10.1007/s12206-018-1035-6

[5] F. Yao and L. Fang, “Thermal Stress Cycle Simulation in Laser Cladding Process of Ni-Based Coating on H13 Steel,” Coatings, 11(2):203, 2021. DOI: 10.3390/coatings11020203

[6] J. Yu, W. Sun, H. Huang, W. Wang, Y. Wang, and Y. Hu, “Crack Sensitivity Control of Nickel-Based Laser Coating Based on Genetic Algorithm and Neural Network,” Coatings, 9(11), 728, 2019. DOI: 10.3390/coatings9110728

[7] Y. Xv, Y. Sun, and Y. Zhang, “Prediction Method for High-Speed Laser Cladding Coating Quality Based on Random Forest and AdaBoost Regression Analysis,” Materials, 17(6), 1266, 2024. DOI: 10.3390/ma17061266.

[8] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. da P. Francisco, J. P. Basto, and S. G. S. Alcalá, “A systematic literature review of machine learning methods applied to predictive maintenance,” Computers & Industrial Engineering, 137(106024), 2019, DOI: 10.1016/j.cie.2019.106024.

[9] S. Kaufman, S. Rosset, C. Perlich, and O. Stitelman, “Leakage in Data Mining: Formulation, Detection, and Avoidance,” ACM Transactions on Knowledge Discovery from Data, 6(4), 1-21, 2012. DOI: 10.1145/2382577.2382579

[10] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, “Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach,” Biometrics, 44(3), 837-845, 1988. DOI: 10.2307/2531595

[11] L. Breiman, “Random Forests,” Machine Learning, 45(1), 5-32, 2001. DOI: 10.1023/A:1010933404324

www.kci.go.kr

Authors



DongGug Yoo received the M.S. and Ph.D. degrees in Accounting from Chonnam National University, Korea, in 1991 and 1995, respectively. He completed the Ph.D. coursework in Economics at Chonnam National University in 2008. He served as the President of Gwangju Technopark from 2010 to 2014 and the President of Jeonnam Technopark from 2018 to 2022. Since 2023, he has been serving as the Chairman of the Board of Directors at the Korea Regional Economy & Industry Research Institute. Since 2025, he has been a Professor at Sunchon National University and the Chairman of the Board of Directors at the Research Institute for Machinery Industry Carbon Neutrality & Circular Technology. His research interests include artificial intelligence and secondary batteries.