

사출 성형 공정에서의 변수 최적화 방법론

정영진* · 강태호* · 박정인* · 조중연* · 홍지수*† · 강성우*

* 인하대학교 산업경영공학과

Methodology for Variable Optimization in Injection Molding Process

Jung, Young Jin* · Kang, Tae Ho* · Park, Jeong In* · Cho, Joong Yeon* · Hong, Ji Soo*† · Kang, Sung Woo*

* Department of Industrial Engineering, Inha University

ABSTRACT

Purpose: The injection molding process, crucial for plastic shaping, encounters difficulties in sustaining product quality when replacing injection machines. Variations in machine types and outputs between different production lines or factories increase the risk of quality deterioration. In response, the study aims to develop a system that optimally adjusts conditions during the replacement of injection machines linked to molds.

Methods: Utilizing a dataset of 12 injection process variables and 52 corresponding sensor variables, a predictive model is crafted using Decision Tree, Random Forest, and XGBoost. Model evaluation is conducted using an 80% training data and a 20% test data split. The dependent variable, classified into five characteristics based on temperature and pressure, guides the prediction model. Bayesian optimization, integrated into the selected model, determines optimal values for process variables during the replacement of injection machines. The iterative convergence of sensor prediction values to the optimum range is visually confirmed, aligning them with the target range. Experimental results validate the proposed approach.

Results: Post-experiment analysis indicates the superiority of the XGBoost model across all five characteristics, achieving a combined high performance of 0.81 and a Mean Absolute Error (MAE) of 0.77. The study introduces a method for optimizing initial conditions in the injection process during machine replacement, utilizing Bayesian optimization. This streamlined approach reduces both time and costs, thereby enhancing process efficiency.

Conclusion: This research contributes practical insights to the optimization literature, offering valuable guidance for industries seeking streamlined and cost-effective methods for machine replacement in injection molding.

Key Words: Injection Molding Process, Machine Learning, Process Variable Prediction

● Received 22 January 2024, 1st revised 13 February 2024, accepted 14 February 2024

† Corresponding Author(wltnghd5182@naver.com)

© 2024, The Korean Society for Quality Management

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-Commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

* 본 논문은 인하대학교의 지원에 의해 연구되었습니다.

1. 서론

사출 공정은 플라스틱 성형법 중의 하나로, 열가소성 수지를 가열하여 유동성을 가진 상태가 되었을 때 금형의 공동부에 가압 주입하여 금형 내에서 냉각시킨 후, 금형의 공동부에 상당하는 성형품을 만들어내는 방법이다(Ko and Cho, 2022). 플라스틱은 우수한 생산성과 경량성, 단열성 등의 장점이 있어 다양한 산업에서 금속을 대체하기 위한 재료로써 평가받고 있다(Yang et al., 2020). 이러한 플라스틱 사출 성형 산업은 제조 산업의 핵심 분야로써 국가에서 뿌리 산업으로 지정하여 육성, 관리하고 있을 만큼 중요한 산업이다(Kim and Hong, 2023). 하지만 국내 사출 성형 산업은 숙련된 기술자들의 노령화, 중소기업들로 이루어진 산업군, 신기술 개발에 대한 투자 부족 등의 이유로 어려움을 겪고 있는 실정이다(Ko and Cho, 2022). 기존의 사출 공정은 현장 엔지니어의 경험을 기반으로 최적의 초기 사출 공정 변수 값을 설정한다(Bae et al., 2021). 이러한 방법은 현장 엔지니어의 숙련도에 따라 서로 상이한 결과가 나오며, 최적의 초기 사출 공정 변수 값을 찾기 위해 많은 시간과 비용이 동반된다(Yang et al., 2020). 본 연구는 이러한 문제점을 해결하기 위해 사출 공정을 모델링하고, 베이지안 최적화를 통해 사출 공정 변수 값을 제시하는 것을 목적으로 한다. 이를 통해 사출 공정 변수 값을 설정하는 시간과 비용을 줄이고, 최종적으로 공정의 효율성 향상에 기여할 것으로 기대된다.

본 논문은 총 5개의 장으로 구성되어 있다. 제1장은 연구의 동기와 논문의 전체적 구성을 설명하는 서론으로 이루어진다. 제2장에서는 본 연구에 활용된 사출 공정 프로세스와 3가지 예측 모델(Decision Tree, Random Forest, XGBoost) 및 베이지안 최적화 프로세스에 대한 이론적 배경 및 선행연구를 논한다. 제3장에서는 사출 공정 최적의 초기 사출 공정 변수 값을 도출하는 방법론을 제안한다. 이후 제4장에서는 앞서 제시한 방법론을 LG전자의 사례에 적용한 결과 및 방법론의 성능에 대해 제시한다. 제 5장에서는 결론 및 본 연구에서 제안하는 방법론의 향후 연구에 대해 논한다.

2. 이론적 배경 및 선행연구

2.1 사출 공정 프로세스

사출 공정은 플라스틱 수지를 용융시킨 후 금형에 주입해 짧은 시간에 원하는 형상의 제품을 얻을 수 있는 제조 방식이다(Lee et al., 2022). 사출 공정 과정은 <Figure 1>과 같이 크게 충전 단계, 가압 단계, 보압 단계로 나뉜다. 첫째, Hopper에 용융 수지가 주입되면, 사출기의 Injection Machine을 통해 Mold의 Cavity 내로 수지가 유입된다. 이 단계를 충전 단계라고 하며, 수지가 Mold의 Cavity에 채워질 때까지 지속된다. 다음으로, 가압 단계는 Injection Machine의 Screw가 더욱 전진하여 Mold의 Cavity에 압력이 가해지는 단계이다. 이는 수지가 압축성의 재료이기 때문에 충전 단계에서 수지를 약 15%만큼 더 Cavity 내로 밀어 넣는 단계이다. 마지막으로, 보압 단계는 가압 단계 이후에도 Screw는 완전히 정지하지 않고 천천히 계속해서 전진하는 단계이다. 일반적으로 수지는 고체 상태로 변할 때 심한 요철 현상이 생기는데, 이는 Cavity와 성형품의 체적 차이 때문이다. 보압 단계는 상기와 같은 체적 차이를 보상하여 더 나은 품질의 제품을 성형하기 위해 진행된다.

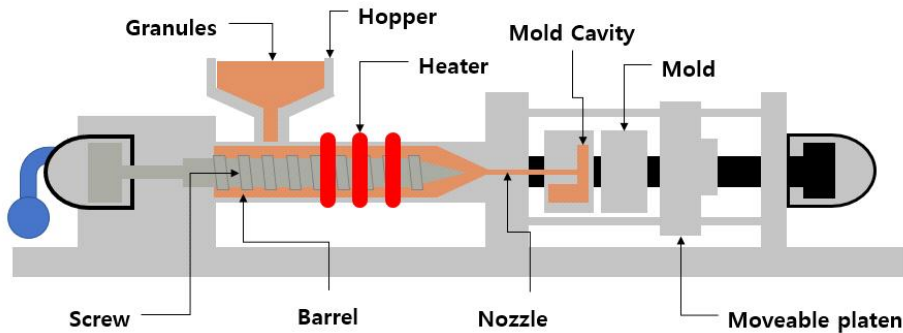


Figure 1. Principle of Injection Process

Yang et al.(2022)의 연구에서 레고 형태의 사출물을 대상으로 용융온도, 금형온도, 사출속도, 보압압력 등 8개의 공정 변수에 대하여 사출물의 x, y, z 축의 길이 및 질량, 압력을 예측하였다. 비선형 문제에 대해서 정확한 예측이 가능한 인공신경망(ANN)을 예측 모델로 채택하였다. 또한 랜덤 공정 변수 조합 100,000개로 예측한 길이 중 목표 길이 범위를 만족하는 공정 변수 조합만 선정하여 최적 공정 변수 조합을 도출하였다. 그 결과 최적 공정 변수 조합에 대한 사출물의 길이 값은 KUZMA Industrial Group의 치수 공차를 만족하였다. 하지만 ANN 모델은 예측에 대한 해석이 어려우며, 훈련 시간이 오래 걸린다는 단점이 존재한다. 또한, 목표로 하는 출력 변수가 많아질수록 최적 공정 변수 조합을 찾는 시간과 시도 횟수가 매우 늘어난다는 한계점이 존재한다.

사출 공정 변수 최적화를 위해 강화 학습을 활용한 연구도 지속되고 있다(Guo et al., 2019). 초고정밀 제품의 최적화 문제를 해결하기 위해, ANN 자체 예측 모델과 통합된 강화 학습 기반의 결정 시스템이 개발되어, 사출 공정에서 동적 최적화를 진행하였다. 훈련된 예측 모델은 R^2 가 0.998로 렌즈 두께를 정확하게 예측하였다. 이 예측 모델은 공정 변수 최적화를 사전 규칙이나 지식 없이 학습할 수 있도록 결정 요소를 훈련시키는 정보를 제공한다. 사출 공정 변수와 렌즈의 가장자리 두께 사이의 상관관계는 결정 요소에 의해 학습되며, 오프라인과 온라인 최적화에 모두 활용된다. 온라인 생산의 경우, Genetic Algorithm에 의해 최적화된 가장자리 두께의 최대 편차는 $+9\mu\text{m}$ 이고, 이 시스템에 의해 최적화된 최대 편차는 $+2\mu\text{m}$ 이다. 그러나 강화 학습을 기반으로 한 제안된 결정 모델은 이론적인 수렴을 보장하지 못하고 데이터 사용이 비효율적인 문제가 있다.

2.2 선행연구

Decision Tree는 데이터를 분할하는 기준에 따라 계층적으로 의사 결정 규칙을 구성하여 분류나 회귀 문제를 해결하는 알고리즘이다. 각각의 분할은 실제값과 예측값 간의 오차가 최소화되는 방향으로 이루어진다. 트리의 말단 노드에 도달하면 해당 클래스 레이블 또는 평균값을 예측값으로 사용한다. 이러한 Decision Tree를 활용하여 공정 개선 연구가 진행되고 있다(Hwang et al., 2021). 해당 연구는 자동차 Radiator Tank 부품의 사출 공정 변화를 예측하기 위해 여러 머신러닝 기법을 사용하였고, 이를 통해 최종적으로 형체력, 사출 압력과 같은 구속 조건을 만족하는 휨 개선 공정을 제안한다. 사용된 여러 머신러닝 기법들 중 Decision Tree 모델의 예측 성능이 가장 우수하였으며, 특히 RMSE(평균 제곱근 오차값)의 경우 평균 대비 1.83의 변동성을 가져 양호한 신뢰 수준을 나타냈다.

Random Forest는 여러 개의 의사 결정 트리를 구성하여 분류나 회귀 문제를 해결하는 알고리즘이다. 각각의 트리가 예측한 값들을 평균 또는 중앙값으로 결합하여 최종 예측값을 도출한다. 스마트폰에 사용되는 모바일용 렌즈의 사출 성형 조건의 최적값을 Random Forest를 활용하여 도출하는 연구가 진행되었다(Lee et al., 2022). 이 연구는

모바일용 비구면 플라스틱 렌즈의 최적 성형 조건을 예측하는 것을 목표로 하며, 렌즈 제품에 영향을 미치는 복굴절과 같은 현상을 최소화하기 위해 위상차를 줄이는 것을 목표로 한다. 위상차에 영향을 미치는 요인은 실험계획법을 사용하여 식별하고, 이러한 요인을 활용하여 위상차를 최소화하는 최적의 사출 성형 공정 변수값을 예측하였다. 이 과정에서 다양한 머신러닝 알고리즘을 사용하여 성능을 평가하였으며, Random Forest를 활용하여 렌즈의 위상차를 8.27%로 감소시켰다.

XGBoost 알고리즘이란, 단일 트리 모델을 순차적으로 결합하는 boosting 기법에 병렬 처리 기능을 추가하여 높은 예측력과 빠른 처리 속도를 장점으로 하는 알고리즘이다. 3D 프린터 출력 방식 중 하나인 전자빔 용해법(EBM) 공정 중 발생하는 결합 유무를 분류하는 연구가 진행되고 있다(Jeong, 2022). 전자빔 용해법(EBM)은 기존 선택적 레이저 용융법(PBF) 대비 빠르지만, 덜 정교하다는 단점이 존재한다. 따라서, 결합과 그 원인을 빠르게 찾는 것이 중요한데, XGBoost는 높은 분류 성능을 내는 앙상블 모델 중에서도 짧은 훈련 시간을 가지기 때문에 사용에 적합하다. 실험 결과 최대 99%의 높은 정확도(Accuracy)로 공정 결합 탐지에 성공했다. 선택적 레이저 용융(SLM) 방식을 이용한 티타늄 합금 부품 제작에서, 4개의 공정 변수에 따른 상대밀도를 XGBoost 모델을 통해 예측한 연구도 진행되고 있다(Zou et al., 2022). 선택적 레이저 용융(SLM)은 에너지 활용성, 흡수율, 진공 보호, 생산효율성 등의 이유로 주목받는 제조 방식이지만, 다양한 공정 변수의 넓은 설정 범위로 인해 부품의 밀도 편차가 큰 문제점이 있다. XGBoost 알고리즘은 이러한 비선형적 회귀 문제에서 높은 예측력을 가지기 때문에 사용에 적합하다.

3. 방법론

본 장은 3.1절 데이터 수집 및 전처리와 3.2절 예측 모델을 통해 3.3절 베이지안 최적화를 활용한 사출 공정 변수 최적화 방법론으로 구성된다. <Figure 2>와 같이 예측 모델 구축과 베이지안 최적화로 총 2단계로 구성된다.

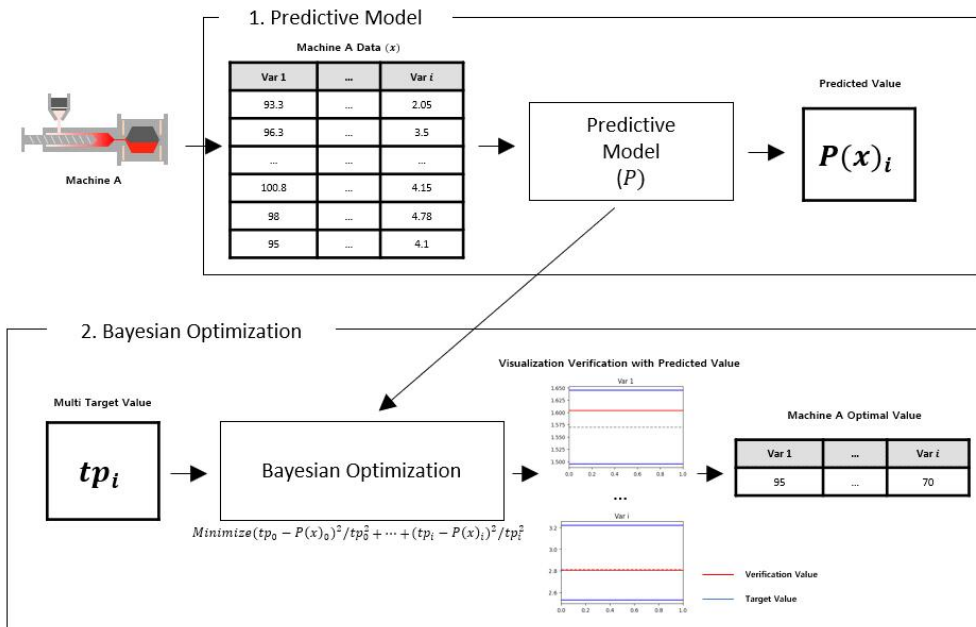


Figure 2. Overview of Variable Optimization in Injection Molding Process

먼저, 사출 공정 데이터를 기반으로 예측 모델을 구축한다. 목표에 적합한 변수를 최적화하기 위해서 여러 기법을 비교하여 예측 모델을 선택한다. 마지막으로 선택된 예측 모델을 기반으로 목표 값을 도출하기 위해 베이지안 최적화를 진행한다.

3.1 데이터 수집 및 전처리

사출 공정 데이터 전처리 과정으로써 같은 값을 가지는 변수는 학습에 유의미하지 않으므로 삭제한다. 또한, 사출 공정 내에서 금형 센서의 오류로 인한 결측치가 발생할 수 있다. 이러한 결측치는 예측 모델의 설명력이 떨어질 수 있으므로 삭제한다. 추가로 'Metering Distance'와 같이 금형 센서의 온도, 압력 등 다른 변수의 영향을 받는 변수도 삭제한다. 전처리가 완료된 데이터를 학습 데이터 80%, 검증 데이터 20%로 분할하고, 이후 모델 학습 및 예측을 수행하고, 이를 검증한다.

3.2 예측 모델

본 연구에서 제시하고자 하는 사출 공정 변수 최적화 제안을 위해 사출 공정 센서 예측 모델을 제시한다. 양품의 사출 조건에 적합한 독립변수를 최적화하기 위해 여러 모델을 비교하여 최고의 성능을 보이는 예측 모델을 선별한다. 이때 사용하는 모델은 Decision Tree, Random Forest, XGBoost이다.

3.2.1 Decision Tree

Decision Tree는 의사결정 규칙(Decision Rule)을 나무 구조로 도표화하여 목표로 하는 집단을 몇 개의 소집단으로 분류하거나 예측을 수행하는 방법이다(Yoo and Lee, 2021). Decision Tree 알고리즘은 반복적인 분류 규칙 생성 작업을 통해 더욱 효과적인 분류 규칙을 만들어가며, 이러한 의사결정 규칙을 나무의 구조로 나타낸다(Jeon and Cho, 2019). Decision Tree 모델의 분류 규칙을 만들기 위해서 CART, C5.0, CA HID, ID3 등 여러 가지 알고리즘이 존재한다(Yoo and Lee, 2021). 그 중 CART(Classification and Regression Trees) 알고리즘은 트리 모델을 회귀에 적용할 수 있다는 장점이 있다. 따라서 본 연구는 CART 알고리즘을 사용한 Decision Tree 모델을 통해 사출 공정 센서값을 예측한다. Decision Tree의 루트 노드(Root Node)에서는 사출 센서 예측값을 두개의 하위 노드로 분류한다. 이 때 노드의 분류 규칙은 수식 (1)에서 MSE_{total} 에 의해 정해진다. Sample은 노드에 포함된 데이터 샘플 수이다. \bar{y}_{left} 와 \bar{y}_{right} 는 두 하위 노드 속에 있는 실제 센서값의 평균이고, y_i 는 i 번째 실제 센서값이다. \bar{y}_{left} 와 y_i , \bar{y}_{right} 와 y_i 의 평균 제곱 오차값을 더함으로써 MSE_{total} 을 산출한다. 이 값이 최소화되도록 가지를 나누는 임계 공정 변수 값이 분류 기준이 된다. 더 이상 하위 노드를 나누지 않을 때, 마지막 노드들을 리프 노드(Leaf Node)라 하며, 이때의 \bar{y}_{left} 와 \bar{y}_{right} 들이 공정 변수 조건들에 대한 센서값의 예측값이다.

$$MSE_{total} = \frac{1}{sample} [\sum_{i \in left} (y_i - \bar{y}_{left})^2 + \sum_{i \in right} (y_i - \bar{y}_{right})^2] \quad (1)$$

3.2.2 Random Forest

Random Forest는 여러 개의 Decision Tree를 사용하여 성능을 향상시키는 앙상블(Ensemble) 학습 방법 중 하나이다(Breiman, 2001). 이때, 데이터를 무작위로 선택하여 다양한 Decision Tree를 생성하고, 각 트리의 예측값을 평균화하여 배깅(Bootstrap Aggregating) 기법을 사용한다. 기존 Decision Tree는 단일 트리를 사용하기에 트리의 깊이가 깊어지면 과적합(Overfitting) 위험이 증가하여 성능이 감소한다. 반면 Random Forest는 다수의 Decision Tree를 결합한 형태이므로 과적합 위험이 상대적으로 감소하는 이점이 있다(Hong and Jeon, 2023). 따라서 본 연구는 Random Forest 모델을 통해 공정 변수에 따른 사출 공정 센서값을 예측한다. Random Forest 모델의 예측값은 수식 (2)와 같이 나타낼 수 있다. 공정 변수와 사출 공정 센서값을 무작위로 선택해 N 개의 Decision Tree를 만든다. 각 트리의 센서 예측값을 $f_i(x)$ 라 할 때, 이 센서 예측값들의 평균 \hat{r} 이 최종 사출 공정 센서 예측값이 된다. 또한 본 모델은 Decision Tree를 기반으로 하기에 각각의 결정 트리는 수식 (1)과 같이 MSE_{total} 값이 최소화되도록 가지를 나누어 학습을 진행한다.

$$\hat{r} = \frac{1}{N_{trees}} \left(\sum_{i=1}^{N_{trees}} f_i(x) \right) \quad (2)$$

3.2.3 XGBoost

XGBoost는 Gradient Boosting에서 병렬 처리와 자체 과적합 방지 기능을 추가하여 높은 예측력과 처리 속도를 가진 트리 부스팅 앙상블 알고리즘이다. 트리 부스팅 알고리즘은 트리를 순차적으로 학습 및 변형함으로써 보다 효율적인 학습을 가능케 한다. 따라서 본 연구는 XGBoost 모델을 통하여 공정 변수에 대한 사출 공정 센서값을 예측한다. XGBoost는 개별 트리를 종합하여 최종 센서 예측값을 도출한다는 점에서 Random Forest 모델과 공통점을 가진다. 하지만 XGBoost는 수식 (3)과 같이 목적함수 $\iota(\Phi)$ 가 최소화되는 방향으로 트리를 순차적으로 생성하고 학습한다. 이때 $\hat{y}_i^{t-1} + f_t(x_i)$ 는 공정 변수 샘플 i 에 대한 $(t-1)$ 번째 트리의 센서 예측값 \hat{y}_i^{t-1} 과 t 번째 트리의 센서 예측값 $f_t(x_i)$ 를 더한 것이다. 이 값과 실제 사출 공정 센서값 y_i 의 평균 제곱 오차를 계산하여 손실함수 l 의 값을 구한다. $\Omega(f_t)$ 는 각 트리의 복잡도를 제어하기 위한 규제 항이다. 이는 각 트리의 리프노드 개수 T , 각 리프노드의 가중치 ω , 과적합 방지 계수 γ 와 λ 로 표현할 수 있다. 이 손실함수와 규제 항을 더하여 목적함수를 결정하고 이를 통해 생성된 개별 Tree를 수식 (4)와 같이 모두 더하여 최종 사출 공정 센서값 \hat{y} 을 도출한다.

$$\iota(\Phi) = \sum_i l(y_i, \hat{y}_i^{t-1} + f_t(x_i)) + \Omega(f_t)$$

$$\text{where } \Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (3)$$

$$\hat{y}_i = \sum_{t=1} f_t(x_i) \quad (4)$$

3.3 Bayesian Optimization

3.2절에서 제시된 예측 모델을 기반으로 사출기의 초기 최적 공정 변수값을 도출하기 위해 베이지안 최적화를 진행한다. 본 연구는 사출기의 성능에 맞춰 제품의 재현성을 확보를 목적으로 한다. 이를 위해 양품 데이터를 기반으로 한 최적의 센서값의 범위를 선정하고, 베이지안 최적화를 통해 사출 공정 변수 파라미터를 최적화한다. 최적의 센서값과 센서의 예측값 사이의 오차를 최소로 하는 목적함수는 수식 (5)로 나타낸다.

$$\text{Minimize } \frac{(tp_0 - P(x)_0)^2}{tp_0^2} + \dots + \frac{(tp_i - P(x)_i)^2}{tp_i^2} \quad (5)$$

tp_0 부터 tp_i 는 목표로 하는 센서의 값이고, $P(x)_0$ 부터 $P(x)_i$ 는 센서의 예측값을 나타낸다. 다중 사출기 조건을 고려하기 때문에 센서에 대한 제곱 오차값을 더하여 목적함수를 정의하였고, 이를 최소화하는 방향으로 사출 공정 변수 최적화를 진행하였다. 본 연구에서 대리 모델(Surrogate Model)과 획득 함수(Acquisition Function)를 활용하여 베이지안 최적화를 진행한다. 대리 모델이란 최적 사출 공정 변수값을 찾기 위한 탐색 과정에서 얻은 목적 함수의 관측값(observation)을 바탕으로 목적함수의 미지의 형태에 대한 확률적인 추정을 하는 모델이다(Lim et al., 2023). 본 연구에서는 대리 모델로 Gaussian Process(GP)를 사용하였다. 획득 함수란 목적함수값이 잠재적으로 높거나 불확실성이 큰 지점을 선정하는 함수로, 선정된 지점을 다음으로 평가하게 된다(Brochu et al., 2010). 본 연구에서는 획득 함수로써 Expected Improvement(EI)를 사용했고, Gaussian Process를 대리 모델로 이용할 경우 EI 함수는 수식 (6)으로 정의한다(Jones et al., 1998).

$$EI(x) = E[\max(f(x) - f(x^+), 0)] \quad (6)$$

수식 (6)은 현재 목적함수의 관측값 $f(x)$ 가 현재까지의 목적함수의 최댓값 $f(x^+)$ 보다 개선이 있으면 양수, 그렇지 않으면 0을 도출하는 분포의 기댓값 E 이다. 이 획득 함수의 값으로 다음 입력값을 도출한다.

4. 연구 실형

4.1 데이터 수집 및 전처리

본 연구에서 사용하는 데이터는 LG전자에서 제공한 40인치 TV 모니터 백 커버 생산 시의 사출 공정 변수 데이터로, 금형의 센서값을 23년 11월 8일, 9일, 13일에 걸쳐 기록한 데이터이다. 금형의 센서 위치는 <Figure 3>과 같은 형태로 구성되며, <Figure 3>의 금형 센서 위치는 LG전자의 보안 유지를 위해 임의의 위치에 표현한다. 센서값 중 T로 시작하는 온도 센서 관련 변수는 Detect(감지 시간), Max(최대온도), MaxDetect(최대온도 감지 시간)으로 세 가지 형태이고, P로 시작하는 압력 센서 관련 변수는 Max(최대압력), MaxDetect(최대압력 감지 시간)로 이루어져 있다. 우선, 데이터 전처리 과정으로 변수 선택 과정을 거친다. 제어 가능한 사출기 공정 변수와 이에 출력되는 금형의 압력/온도 변수를 제외한 나머지 변수들은 실험에 고려하지 않는다. 예를 들어, Actual Value인 'Max Clamping Force' 등의 16개 변수들은 현장에서 제어한 사출 공정 변수로 인해 출력되는 변수이므로 삭제한다. 또한, 'Suck-Back Velocity Set Value 1'와 같이 공정 변수의 변화와 관계없이 일정한 값을 가지는 변수는 삭제한다. 마

지막으로 ‘Cycle’, ‘Part No’와 같이 압력/온도 변수 예측에 관계없는 변수는 삭제한다.

본 연구에서 사용하는 독립변수인 사출 공정 변수 12개는 <Table 1>, 온도 종속변수는 <Table 2>, 압력 종속변수는 <Table 3>와 같다. 온도 특성 중 감지 시간 종속변수는 ‘T3_Detect’ 등의 14개 변수, 최대온도 종속변수는 ‘T3_Max’ 등의 14개 변수, 최대온도 감지 시간 종속변수는 ‘T3_MaxDetect’ 등의 14개 변수로 나타낸다. 압력 특성 중 최대압력 종속변수는 ‘P1_Max’ 등의 5개 변수, 최대압력 감지 시간은 ‘P1_MaxDetect’ 등의 5개 변수로 나타낸다. 실험에 사용된 데이터는 총 182개이다<Table 4>. 이후 전처리가 완료된 데이터를 학습과 검증을 위해 8:2로 분할하여 실험을 수행하며, 학습에 사용된 데이터는 146개, 검증에 사용된 데이터는 36개이다.

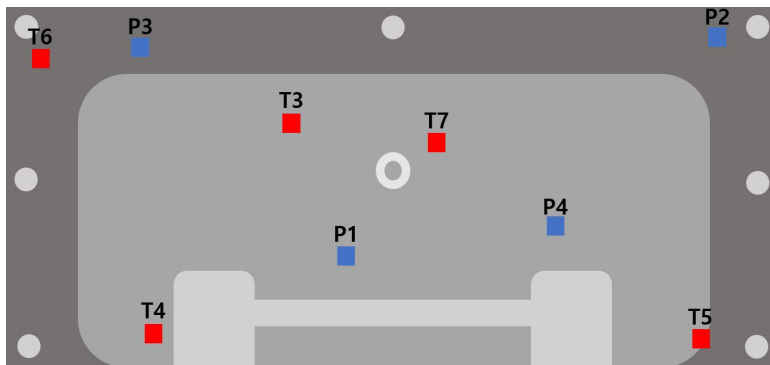


Figure 3. Example of Mold’s Pressure Sensor, Temperature Sensor Position

Table 1. Independent Variables of Injection Molding Process

Metering Position Set Value 3	Back Pressure Set Value 1	Metering RPM Set Value	Filling Process Set Position 1
Velocity Set Value during Filling 1	Velocity Set Value during Filling 2	VP Position Set Value	Packing Pressure Set Value 1
Packing Time Set Value 1	Barrel Temperature Set Value ZH	HRS Temperature Set Value	Mold Temperature

Table 2. Dependent Variables of Injection Molding Process (Temperature Sensor)

T3_Detect	T4_Detect	T5_Detect	T6_Detect
T7_Detect	T8_Detect	T9_Detect	T10_Detect
T11_Detect	T12_Detect	T13_Detect	T14_Detect
T15_Detect	T16_Detect	T3_Max	T4_Max
T5_Max	T6_Max	T7_Max	T8_Max
T9_Max	T10_Max	T11_Max	T12_Max
T13_Max	T14_Max	T15_Max	T16_Max
T3_MaxDetect	T4_MaxDetect	T5_MaxDetect	T6_MaxDetect
T7_MaxDetect	T8_MaxDetect	T9_MaxDetect	T10_MaxDetect
T11_MaxDetect	T12_MaxDetect	T13_MaxDetect	T14_MaxDetect
T15_MaxDetect	T16_MaxDetect		

Table 3. Dependent Variables of Injection Molding Process (Pressure Sensor)

P1_Max	P2_Max	P3_Max	P4_Max	P5_Max
P1_MaxDetect	P2_MaxDetect	P3_MaxDetect	P4_MaxDetect	P5_MaxDetect

Table 4. Injection Molding Process Data Frame

Index	Metering Position Set Value 3	Back Pressure Set Value 1	...	P3_MaxDetect	P4_MaxDetect	P5_MaxDetect
1	99.4	5	...	3.49	4.115	2.16
2	99.4	3	...	4.315	2.65	4.315
3	100.3	7	...	4.47	2.665	4.325
4	99.4	3	...	2.00	1.94	1.78
⋮	⋮	⋮	...	⋮	⋮	⋮
182	99.4	7	...	2.27	2.075	2.165

4.2 예측 모델 학습

공정 변수에 대한 사출 공정 센서값 예측 모델링을 위해 Decision Tree, Random Forest 및 XGBoost를 사용한다. 예측 모델들은 52개의 종속변수를 동시에 학습하고 예측할 수 있는 Multioutput Regressor를 활용하였다. 각 예측 모델에 대해서 12개의 공정 변수와 52개의 사출 공정 센서값 146행으로 구성된 학습 데이터(Train Data)가 입력된다. Decision Tree 모델의 경우 수식 (1)의 MSE_{total} 이 최소화될 수 있도록 특정 공정 변수의 임계값을 설정하고 이 값을 기준으로 두개의 하위 노드로 나누는 과정을 반복한다. 더 이상 하위 노드를 나누지 않을 때 리프노드들에 속해 있는 실측 센서값 행들의 평균이 해당 공정 변수 조건에 따른 예측 센서값으로 학습된다. 그 후 36행의 검증 데이터를 입력 시 12개 공정 변수의 행들은 학습되어 있는 트리의 분류 규칙에 따라 특정 리프 노드에 배정된다. 해당 리프 노드의 학습된 센서값이 예측 센서값이 된다.

Random Forest 모델의 경우 학습 데이터를 입력 시 사출 데이터 146행을 무작위로 복원 추출하여 100개의 Decision Tree를 생성한다. 각각의 트리는 Decision Tree의 센서값 학습 방법을 따른다. 그 후 검증 데이터 36행을 입력 시 100개의 Decision Tree에 대하여 각각의 예측 센서값이 도출된다. 이 값들을 수식 (2)와 같이 평균화하여 최종 예측 센서값을 구한다. XGBoost 모델의 경우 학습 데이터를 입력 시 수식 (3)과 같이 146행에 대한 실측 센서값과 예측값의 오차를 줄이는 방향으로 두 번째 트리를 생성한다. 이후 두 번째 트리도 예측한 센서값과 실측값의 오차를 측정하고, 이 오차를 줄이는 방향으로 세 번째 트리를 생성한다. 이러한 방식으로 10개의 트리를 순차적으로 생성하고 구조를 학습한다. 검증 데이터를 입력 시 12개의 공정 변수 행에 대한 예측 센서값을 도출한다. 본 연구에서는 예측 모델의 성능 평가 지표로써 결정 계수 (R^2)와 평균 절대 오차(Mean Absolute Error, MAE)를 사용하였다. MAE는 실측값과 각 실측값에 대응되는 모형의 예측값의 차이의 평균이다. 따라서 모형의 예측 성능이 좋을수록 실측값과 예측값의 차이가 줄어들게 되고, MAE는 작은 값을 가지게 된다.

검증 데이터(Test Data)에 대한 예측 모델 별 성능 결과는 <Table 5>와 같다. 전체 데이터 186행에 대하여 10점의 교차검증을 진행하였다. 또한 각 3가지 후보 모델에 대하여 5점의 교차검증을 통하여 하이퍼 파라미터를 최적화

하였다. 실험 결과는 종속변수의 5가지 특성인 ‘T_Detect’, ‘T_Max’, ‘T_MaxDetect’, ‘P_Max’, ‘P_MaxDetect’로 표현하고, 모든 종속변수를 포함한 결과는 ‘All Dependent Variables’로 표현한다<Table 5>.

실험 결과, 종속변수의 특성에 따른 5개 특성을 사용한 경우에는 R^2 와 MAE 기준 Decision Tree와 XGBoost 모델이 Random Forest 모델에 비해 미세하게 높은 성능을 보여주었다. 또한, 모든 종속변수 기준(‘All Dependent Variables’)으로 R^2 값은 XGBoost 모델이 0.81로 가장 높았으며, MAE 역시 XGBoost 모델이 0.77로 가장 좋은 성능을 보였다. 본 결과에 대한 모델 적합도를 <Figure 4>에 나타내었으며, X축은 실제 센서값, Y축은 예측 모델을 통한 센서 예측값이다.

Table 5. Results of Variables for Predictive Models

Predictive Models	T_Detect		T_Max		T_MaxDetect		P_Max		P_MaxDetect		All Dependent Variables	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
Decision Tree	0.96	0.01	0.90	0.45	0.47	0.20	0.98	6.08	0.88	0.07	0.80	0.78
Random Forest	0.96	0.02	0.88	0.49	0.51	0.20	0.96	9.60	0.86	0.11	0.81	1.12
XGBoost	0.96	0.01	0.90	0.44	0.47	0.20	0.98	6.20	0.88	0.07	0.81	0.77

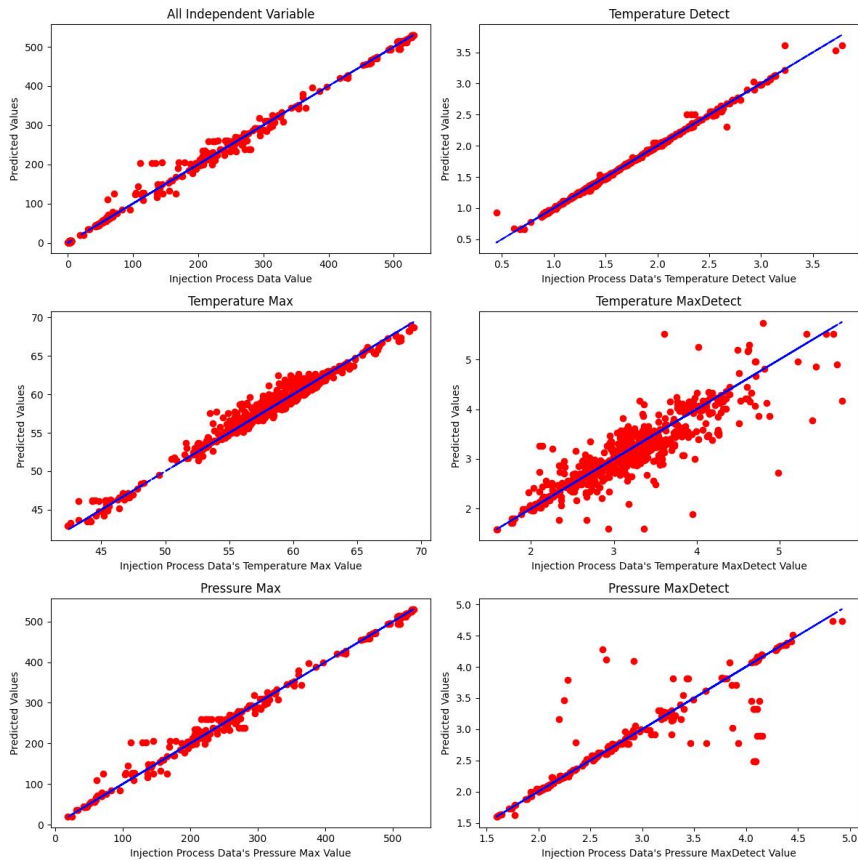


Figure 4. Comparison Between the Actual Values and Predicted Values of Sensors

4.3 베이지안 최적화

본 연구는 예측 모델을 사용하여 베이지안 최적화를 적용하였다. 먼저, 3.3절의 수식 (5)와 같이 목적함수를 구성한다. 또한, 대리 모델을 사용하여 획득 함수를 기반으로 다음 지점을 선정한다. 해당 과정을 반복하여 사출 공정 변수값을 도출하고, 도출된 12개의 사출 공정 변수를 예측 모델에 넣고 센서값을 예측한다. 3.3절의 수식 (5)에 적용 시, tp_0 부터 tp_{33} 은 최적 온도 센서 24개와 최적 압력 센서 10개의 평균값이고, $P(x)_0$ 부터 $P(x)_{52}$ 은 온도 센서 42개와 압력 센서 10개의 예측값을 나타낸다. 사출 공정 변수 탐색 범위는 선택한 예측 모델 학습 데이터 변수들의 최솟값과 최댓값으로 설정하였다.

대리 모델(Surrogate Model)과 획득 함수(Acquisition Function)를 활용하여 최적의 사출 공정 변수를 탐색하는 방식으로 진행한다. 먼저, 기존 사출기로부터 수집한 데이터를 예측 모델로 훈련시켜, 12개의 사출 공정 변수에 의해 생성된 52개의 금형 센서값을 예측한다. 다음으로, 기존 사출기와 신규 사출기 사이의 출력 비교를 통해 신규 사출기의 공정 변수와 센서값의 범위를 지정한다. 해당 범위 사이에서 획득 함수는 3.3절의 수식 (6)에 따라 목적 함수 값이 잠재적으로 높은 지점을 활용하거나 불확실성이 높은 지점을 탐색하여 다음 사출 공정 변수를 도출한다. 도출한 사출 공정 변수 값을 학습된 모델로 예측했을 때, 모든 센서값이 최적의 센서 범위 값의 허용 범위 내에 들어오면 사출 공정 변수 파라미터가 최적화가 되었다고 판단하고 반복을 중지한다. 만약 하나라도 센서값 중 범위를 벗어나는 값이 있으면 대리 모델을 업데이트하고 다시 획득 함수로 다음 사출 공정 변수를 도출한다. 최적 반복 중지를 목표로 하는 최적 센서값 범위는 <Table 6>과 같으며, 이는 '외관' 변수 중 완충인 종속변수 값들의 최솟값과 최댓값으로 나타났다. 이를 기반으로 베이지안 최적화를 통해 최적 사출 공정 변수를 도출한 결과는 <Table 7>과 같다.

Table 6. Optimal Range of Dependent Variable

Dependent Variable	Optimal Range	Dependent Variable	Optimal Range
T3_Detect	$1.005 \leq X \leq 2.380$	T15_Max	$50.90 \leq X \leq 66.30$
T4_Detect	$0.980 \leq X \leq 2.430$	T16_Max	$51.30 \leq X \leq 64.80$
T5_Detect	$2.010 \leq X \leq 3.780$	T3_MaxDetect	$2.255 \leq X \leq 5.400$
T6_Detect	$1.075 \leq X \leq 2.780$	T4_MaxDetect	$2.520 \leq X \leq 5.690$
T7_Detect	$0.615 \leq X \leq 1.585$	T5_MaxDetect	$3.340 \leq X \leq 5.805$
T8_Detect	$0.995 \leq X \leq 2.745$	T6_MaxDetect	$2.065 \leq X \leq 4.290$
T9_Detect	$1.060 \leq X \leq 2.325$	T7_MaxDetect	$1.575 \leq X \leq 3.235$
T10_Detect	$0.405 \leq X \leq 2.510$	T8_MaxDetect	$1.600 \leq X \leq 6.305$
T11_Detect	$0.405 \leq X \leq 3.000$	T9_MaxDetect	$2.240 \leq X \leq 4.265$
T12_Detect	$0.405 \leq X \leq 2.950$	T10_MaxDetect	$1.770 \leq X \leq 4.265$
T13_Detect	$1.620 \leq X \leq 3.140$	T11_MaxDetect	$2.165 \leq X \leq 4.015$
T14_Detect	$1.540 \leq X \leq 2.740$	T12_MaxDetect	$1.735 \leq X \leq 5.110$
T15_Detect	$0.995 \leq X \leq 2.695$	T13_MaxDetect	$2.715 \leq X \leq 4.655$
T16_Detect	$1.370 \leq X \leq 2.575$	T14_MaxDetect	$2.610 \leq X \leq 5.655$
T3_Max	$54.20 \leq X \leq 66.90$	T15_MaxDetect	$1.965 \leq X \leq 5.755$
T4_Max	$56.20 \leq X \leq 69.40$	T16_MaxDetect	$2.520 \leq X \leq 4.710$
T5_Max	$50.50 \leq X \leq 63.80$	P1_Max	$19.00 \leq X \leq 418.7$

Dependent Variable	Optimal Range	Dependent Variable	Optimal Range
T6_Max	54.30 ≤ X ≤ 69.00	P2_Max	386.0 ≤ X ≤ 553.7
T7_Max	54.00 ≤ X ≤ 68.40	P3_Max	41.80 ≤ X ≤ 363.1
T8_Max	55.60 ≤ X ≤ 68.40	P4_Max	138.6 ≤ X ≤ 445.5
T9_Max	52.70 ≤ X ≤ 66.70	P5_Max	194.3 ≤ X ≤ 498.4
T10_Max	52.50 ≤ X ≤ 66.40	P1_MaxDetect	2.000 ≤ X ≤ 4.230
T11_Max	38.80 ≤ X ≤ 53.70	P2_MaxDetect	1.595 ≤ X ≤ 3.300
T12_Max	49.00 ≤ X ≤ 69.10	P3_MaxDetect	1.720 ≤ X ≤ 4.925
T13_Max	50.10 ≤ X ≤ 64.80	P4_MaxDetect	1.640 ≤ X ≤ 4.210
T14_Max	51.80 ≤ X ≤ 66.60	P5_MaxDetect	1.620 ≤ X ≤ 4.340

Table 7. Independent Variable with Bayesian Optimization

Independent Variable	Value	Independent Variable	Value
Metering Position Set Value 3	95.73	VP Position Set Value	13.82
Back Pressure Set Value 1	3.275	Packing Pressure Set Value 1	43.34
Metering RPM Set Value 1	75.03	Packing Time Set Value 1	1.831
Filling Process Set Position 1	85.07	Barrel Temperature Set Value ZH	255.5
Velocity Set Value during Filling 1	23.30	HRS Temperature Set Value	237.0
Velocity Set Value during Filling 2	67.06	Mold Temperature	51.66

5. 결론

사출 공정은 열가소성 수지를 활용하여 대부분 플라스틱 제품을 제조하는 공정이다. 사출 공정에서 금형을 다른 사출기로 교체해야 하는 경우, 양품을 생산하기 위해 사출 공정 변수를 조정하는 과정은 중요하다. 하지만 기존 사출 공정 변수 조정 방식은 작업자들의 경험에 따라 Set Value 지정을 위한 실험 계획을 세우고 양품을 생산할 수 있도록 모든 Case에 대해 Full Factorial Design을 고려하여 진행하기에 작업자의 숙련도에 따라 상이한 결과를 도출할 수 있으며, 제품 품질이 불안정할 수 있다. 또한, 실험이 진행되는 동안 낭비되는 성형품으로 인한 비용적 문제와 최적 변수값 세팅을 위한 조정 시간의 지연으로 인한 시간적 문제 등 사출기 교체로 인한 생산 효율성이 크게 떨어진다. 본 연구는 사출기 교체 시 떨어지는 생산 효율성을 최소화하기 위한 방법으로 사출 공정 데이터의 예측 모델을 개발하고, 베이지안 최적화를 활용하여 최적의 사출 공정 변수 값을 도출하였다. 사출 공정 데이터의 공정 변수에 따른 센서 변수 52개를 온도 감지 시간, 최대온도, 최대온도 감지 시간, 최대압력, 최대압력 감지 시간으로 나눠 예측하였으며, 예측 모델이 실제 데이터와 큰 차이가 없음을 확인하였다. 또한, Grid Search보다 수행되는 속도가 빠르고, Random Search보다 정확한 결과를 도출할 수 있는 베이지안 최적화를 통해 최적의 사출 공정 변수 값을 도출하였으며, 최적화 결과의 검증은 위하여 시각화를 수행하였다.

본 연구에서 제시하는 방법론을 통하여 현장에서 사출 공정 변수 설정에 소요되는 시간과 변수 조정 과정에서의 낭비되는 성형품 생산 시 발생하는 비용을 줄일 수 있고, 최종적으로 작업자의 숙련도에 따라 좌우되는 공정 품질이 아닌 본 연구 방법론으로 인한 표준화된 Set Value 설정으로 공정 전반의 효율성 향상에 기여할 수 있을 것으로 보

이다. 본 연구팀은 향후 연구로 데이터의 증강 및 추가적인 모델의 하이퍼 파라미터 튜닝을 통해 예측 모델을 더욱 강화할 것이며, 새로운 사출 공정이나 사출기 교체 과정에 적용하여 본 연구 방법론을 확증할 예정이다.

REFERENCES

- Bae, J., Lee, W., Jee, H., Hong, B., and Lee, J. 2021. Simulation for Injection Molding of Insulation Spacers for Gas-Insulated Switches Using Thermosetting Epoxy Resin. *Journal of the Korean Institute of Electrical and Electronic Material Engineers* 34(6):426-432.
- Breiman, L. 2001. RandomForests. *Machine Learning* 45:5-32.
- Brochu, E., Cora, V. M., and De Freitas, N. 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*.
- Guo, F., Zhou, X., Liu, J., Zhang, Y., Li, D., and Zhou, H. 2019. A Reinforcement Learning Decision Model For Online Process Parameters Optimization From Offline Data In Injection Molding. *Applied Soft Computing* 85:105828.
- Hong, J., and Jeon, S. 2023. Prediction of Safety Grade of Bridges Using the Classification Models of Decision Tree and Random Forest. *Journal of The Korean Society of Civel Engineers* 43(3):397-411.
- Hwang, H., Cho, Y., Hwang, S., and Kim, S. 2022. Optimal Tire Design Using Machine Learning and Bayesian Optimization. *Journal of the Korean Institute of Industrial Engineers* 48(4):433-440.
- Hwang, S., Han, S., and Lee, H. 2021. A Study on the Improvement of Injection Molding Process Using CAE and Decision-tree. *Journal of the Korea Academia-Industrial cooperation Society* 22(4):580-586.
- Jeon, Y., and Cho, H. 2019. Model Based Hybrid Decision Tree. *Journal of the Korean Data And Information Science Society* 30(3):515-524.
- Jeong, J. 2022. Predicting Defects of EBM-Based Additive Manufacturing through XGBoost. *Journal of the Korea Institute of Information and Communication Engineering* 26(5):641-648.
- Jones, D. R., Schonlau, M., and Welch, W. J. 1998. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization* 13:455-492.
- Kim, D. and Hong, J. 2023. Improvement of Plastic Injection Molding Process of Mole Trap Parts using Design of Experiments. *Journal of the Korean Society of Manufacturing Technology Engineers* 32(2):85-91.
- Kim, K., and Kim, S. 2023. Hot-Rolled Steel Plate Thickness Prediction and Bayesian Optimization-Based Rolling Pattern Derivation. *Journal of the Korean Institute of Industrial Engineers* 49(2):167-175.
- Ko, M., and Cho, Y. 2022. Application of Deep Learning Algorithm for Defect Detection and Cause Analysis of Automotive Parts in Injection Process. *Journal of the Korean Society of Manufacturing Technology Engineers* 31(6):452-459.
- Lee, Y., Joo, J., Lim, S., Lee, J., and Han, S. 2022. Optimization of Molding Conditions for Mobile Lens using Injection Molding CAE and Machine Learning. *Journal of the Korean Society of Mechanical Technology* 24(4):652-658.
- Lim, S., Seo, H., and Yu, Y. 2023. Bayesian Optimization for 3D Source Localization from Multi-channel Time-series Signals. *Journal of the Korean Society for Nondestructive Testing*. 43(1):44-51.
- Wu, J., Chen, X. Y., Zhang, H., Xiong, L. D., Lei, H., and Deng, S. H. 2019. Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization. *Journal of Electronic Science and Technology* 17(1):26-40.
- Yang, D., Lee, J., Yoon, K., and Kim, J. 2020. A Study on the Prediction of Optimized Injection Molding Condition using Artificial Neural Network (ANN). *Transactions of Materials Processing* 29(4):218-228.

- Yang, D., Lee, J., Yoon, K., and Kim, J. 2022. A Study on the Practical Application of the Integrated ANN System for Manufacturing the Target Quality of the Injection Molded Product. *Korea–Australia Rheology Journal* 34(2):147–157.
- Yoo, J., and Lee, G., 2021. A Study on the Prediction Analysis of Quality Control Components and Control Criteria for Portuguese Red Wine using UCI Dataset: Focusing on the Decision Tree Techniques by CART Algorithm. *Korean Journal of Hospitality & Tourism* 30(6):239–255.
- Zou, M., Jiang, W. G., Qin, Q. H., Liu, Y. C., and Li, M. L. 2022. Optimized XGBoost Model with Small Dataset for Predicting Relative Density of Ti–6Al–4V Parts Manufactured by Selective Laser Melting. *Materials* 15(15):5298.

저자소개

- 정영진** 인하대학교 산업경영공학과 학사를 취득하고 현재 인하대학교 산업경영공학과 석박사통합과정에 재학 중이다. 주요 관심 분야는 데이터 사이언스이다.
- 강태호** 한밭대학교 정보통신공학과 학사를, 인하대학교 산업경영공학과 석사를 취득하였다. 주요 관심 분야는 빅데이터 분석, 품질 및 생산관리이다.
- 박정인** 인하대학교 산업경영공학과 학사를 취득하고 현재 인하대학교 산업경영공학과 석사과정에 재학 중이다. 주요 관심 분야는 데이터 사이언스, 머신러닝이다.
- 조중연** 현재 인하대학교 산업경영공학과 학사 재학 중이다. 주요 관심 분야는 데이터 사이언스, 머신러닝이다.
- 홍지수** 인하대학교 통계학과 학사를, 인하대학교 산업경영공학과 석사를 취득하고 현재 인하대학교 산업경영공학과 박사과정에 재학 중이다. 주요 관심 분야는 빅데이터 분석, 데이터 마이닝, 품질관리, 제조 최적화이다.
- 강성우** 현재 인하대학교 산업경영공학과 부교수로 재직 중이다. 인하대학교 산업경영공학과에서 학사를, 펜실베이니아 주립대학교 산업제조공학과에서 석사와 박사를 취득하였다. 관심 연구 분야는 빅데이터(3D 이미지, 텍스트 데이터) 프로세싱을 기반한 제품 설계, 공학 설계, 생산 장비 예측 진단 및 관리이다.