

사출 공정에서 불량 예측을 위한 데이터 분석 및 AI 적용 모델

이기성¹, 이종찬^{2*}

¹호원대학교 컴퓨터게임학부, ²군산대학교 컴퓨터정보공학과

Data Analysis and AI Model for Defect Prediction in the Injection Process

Gi-Sung Lee¹, Jong-Chan Lee^{2*}

¹Division of Computer & Game, Howon University

²Department of Computer Information Engineering, Kunsan National University

요약 사출 공정에서 제품의 품질은 사출기 내·외부 데이터에 따라 변화할 수 있고 이러한 변화는 용융 수지의 상태에 반영되므로 관련 데이터를 추출하여 제품의 품질과 공정을 관리하는 것이 필요하다. 현재 사출성형의 품질관리는 주로 과거의 생산 데이터에 기초한 통계적인 방법과 작업자의 경험에 의존하고 있지만, 사출성형 시에 복잡한 조건을 반영하기에는 현실적 어려움이 있으며 불량품을 양산할 가능성이 항상 존재하므로, 효율적인 데이터 관리 체계가 필요하다. 데이터의 수집은 물리적으로 한계가 없으나, 활용의 측면에서 수집한 데이터로부터 의미 있는 정보를 추출하고 제조 현장에 활용하기 위해서는 고수준의 분석작업이 필요하므로 데이터를 활용하고 있는 기업의 수가 적은 것이 현실이다. 본 연구에서는 실제 사출 공정에서 발생하는 품질 문제를 해결하기 위하여 수집된 주요 공정 데이터의 상관관계 분석을 수행하고, 다양한 기계학습 방법과 알고리즘을 활용하여 사출 공정 데이터로부터 불량 예측 모델을 만든다. 이를 통하여, 제조기업의 프레스 공정조건 최적화 및 품질 예측을 수행하고 휴먼 오류를 최소화함으로써 불량률을 개선할 수 있다.

Abstract In the injection process, product quality can change depending on the internal and external data of the injection machine, and these changes are reflected in the state of the molten resin, so it is necessary to manage the product quality and process by extracting these data. Currently, quality control of injection molding relies mainly on statistical methods and worker experience based on past production data, but it is difficult to reflect complex conditions during injection molding. There is always a possibility for mass production of defective products, so an efficient data management system is needed. Data collection is not physically limited, but the number of companies using data is still small because high-level analysis is required to extract meaningful information from the data collected in terms of utilization, and to use that information at the manufacturing site. In this study, to solve the quality problems arising from the actual injection process, we perform a correlation analysis of the collected major-process data, and we create a defect prediction model for injection-process data using various machine learning methods and algorithms. Through this, it is possible to minimize human error and improve the defect rate through optimization of press-process conditions and quality prediction by manufacturing companies.

Keywords : Big-data, AI, Manufacturing Data, Data Analysis, Injection Process

본 논문은 2022년도 호원대학교 지원에 의하여 출판되었음.

*Corresponding Author : Jong-Chan Lee(Kunsan National Univ.)

email: chan2000@kunsan.ac.kr

Received November 2, 2022

Revised December 6, 2022

Accepted December 7, 2022

Published December 31, 2022

1. 서론

산업현장에서 인건비 상승 부담을 최소화할 수 있는 방법으로 스마트 제조(Smart Manufacturing)가 주목 받고 있다. 스마트 제조는 각종 센서를 통해 수집된 데이터를 분석하여 제조 공정상의 불량률을 최소화하거나, 설비 이상을 사전 감지해 고장을 예방하고, 자동화된 통합 제조 공정의 구축에 활용한다. 이런 스마트 제조의 핵심기반 기술은 인공지능으로서, 현실 공간에서 수집된 데이터를 가상 공간에서 분석하고, 도출된 솔루션 및 의사결정 사항을 다시 현실 공간에 전달해 자동으로 기기를 제어하고 공정을 최적화하는 전 과정을 인공지능이 수행한다[1-3].

사출 산업은 자동화된 시스템으로부터 대량의 데이터가 생산되는 대표적인 제조업으로서, 불량 감소 및 생산성 향상을 위하여 정밀한 제조설비의 관리가 필요하다. 현장에서는 엔지니어의 경험과 직관으로 공정이 관리되고 있으나, 공정이 정밀화할수록 수작업으로 대량의 데이터를 분석하고 관리하는 것은 비효율적이다. 따라서 효율적인 공정 관리를 위하여 통계적 방법의 연구가 필요한 상황이다[4-6].

사출 공정 시에 생산되는 데이터는 설비나 공정, 또는 생산되는 제품에 대한 정보를 담고 있다. 따라서 데이터를 분석함으로써 설비 및 공정의 이상을 탐지하거나, 공정의 제어 조건을 산출하고 수율을 예측하는 것이 가능하다. 그러나 사출 공정에서 발생하는 데이터를 분석하고 활용할 수 있는 방법의 부재로 인하여 데이터를 적용한 사출 공정 관리가 어려운 실정이다[7-9].

사출 공정에서 발생하는 문제점을 파악하기 위해서는 우선, 공정에서 수집되는 빅데이터의 특성을 분석해야 한다. 사출기는 단계별 가공 상태에 영향을 받기 때문에 입력 변수와 출력 변수 간에 함수 관계가 존재하게 된다. 따라서 공정에서 발생하는 빅데이터는 비선형성(Non-linear)과 비정규성(Non-normal) 특징을 보이므로, 서로 다른 데이터 간에 선형 관계를 나타내는 상관관계(Correlation)와 시간의 흐름에 따라 연속적으로 관측되는 시계열 데이터의 자기 상관관계(Autocorrelation)여부의 확인이 필수적이다. 그리고 빅데이터를 분석하기 위해 변수 선택, 특징 추출, 데이터 샘플링 등 전처리 수행을 통해 데이터의 양을 줄이고 차원을 축소해야 한다. 제조 공정에서는 양품보다 불량품이 소수이고, 이에 따라 불균형하게 데이터가 수집되므로 데이터의 특성을 고려하여 표본을 추출해야 한다[10-12].

따라서 본 연구에서는 사출 공정의 품질 문제를 해결을 위하여, 불량품의 원재료 및 사출 정보를 수집하고 분석함으로써 불량 예측 AI 모델을 개발하였다. 즉 사출 공정에서 생산된 제품의 불량을 예측하기 위해서 AI 모델의 성능을 비교하고, 최종 모델을 도출하였다. 대부분의 생산라인의 경우 제품 공정조건의 변경은 현업의 경험과 감에 의존하므로, 본 연구를 통해 데이터 기반의 생산 공정 기술 및 기술 표준화로 생산성 및 품질 향상을 위한 초석을 마련할 수 있다.

2장에서는 데이터 분석 시스템 구조를 보인다. 3장에서는 데이터 분석을 위한 세부 방법을 기술하고, 4장에서 분석 결과를 보인다. 5장에서 결론을 맺는다.

2. 시스템 구조

본 논문은 제조설비에서 실시간으로 발생하는 데이터의 수집 및 전처리, 그리고 이를 분석하여 시각화함으로써 불량 예측 모델의 도출을 목표로 한다. Fig. 1과 같이 데이터의 수집, 전처리, 정밀한 분석 과정, 실제 현장의 공정 운영 조건에 설정하여 운영 최적화를 수행한다[1].

① 데이터 전처리

- 수집한 데이터로부터 결측치, 이상치, 데이터 샘플링을 수행한다.
- 서로 다른 규모의 자료들을 효율적으로 분석하기 위하여 데이터 정규화를 수행한다.
- 샘플링된 데이터의 라벨링을 수행한다.

② 데이터 분석

- 독립변수와 종속변수의 상관분석, 독립변수 간의 상관분석을 통하여 인자 간의 관련 유무를 파악한다.
- 딥러닝과 타 인공지능 모델의 예측을 시도하여 우수한 모델을 선택한다.
- 불량분석을 통하여 외부 상황에서의 생산 상태, 불량률을 개선한다.

3. 데이터 전처리 및 분석

3.1 데이터 전처리(Data Preprocessing)

3.1.1 결측값(Missing Value) 처리

사출 공정상의 데이터는 MCAR(Missing Completely

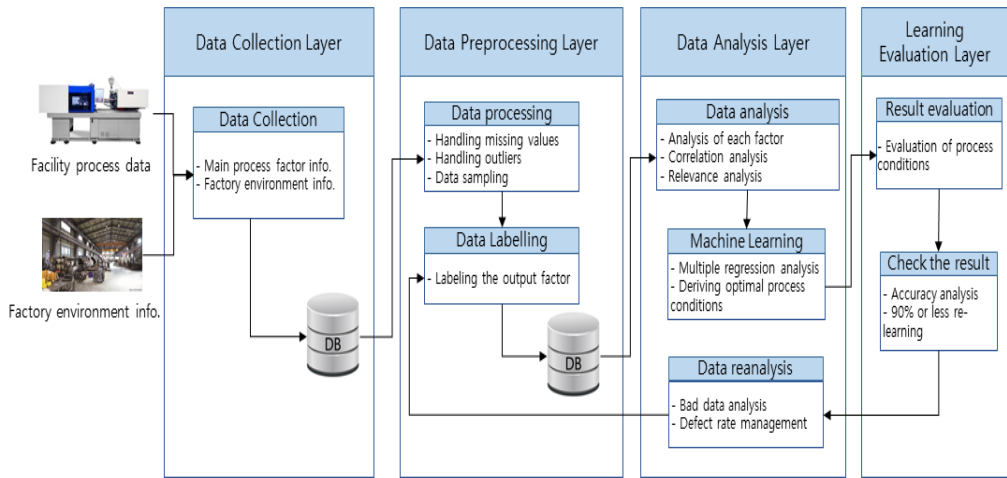


Fig. 1. AI-based data analysis procedure

At Random)이라고 가정하고 결측값을 처리한다. 결측값 처리 방법으로서 CQI(Completeness Quality Indicator)를 구하여 처리한다. Where $N(u)$ denotes the number of missing values, n denotes the total number of data.

$$cqi = (1 - N(u)/n) \times 100 \quad (1)$$

각 칼럼(Column)의 결측값의 비율이 30% 이상이면 데이터의 완전성이 떨어지기 때문에 해당 칼럼을 삭제한다. 데이터의 결측치를 확인하기 위해 isnull() 함수를 사용하고 sum() 함수를 이용하여 결측치 개수를 구한 후, 완전성 품질 지수를 구한다. 결측치가 10% 미만이면 결측치 데이터를 삭제하고, 10% 이상이면 공정 인자 값의 평균이 아닌 각 인자의 특성을 고려하여 값을 삽입하는 회귀 삽입(Regression Imputation)을 수행하여 CQI가 100%가 되도록 결측값 처리를 수행한다.

3.1.2 이상치(Outlier) 처리

원시 데이터(Raw data)는 변수마다 수치의 크기와 평균, 분산 등의 통계치가 다르고, 특히 지나치게 낮거나 높은 값인 이상치가 존재할 수 있다. 이상치로 인하여 변수들 사이의 보편적인 관계가 왜곡되고 통계 분석에 영향을 미칠 수 있으므로, 스케일링을 통해 모델의 성능을 재고하고 데이터를 정제한다.

이상치 탐지 방법으로는, 데이터 셋을 의사결정트리(Decision Tree) 형태로 표현해 정상값을 분리하는 IF(Isolation Forest)를 적용한다. 사출 공정상의 데이터가

정상 샘플이라는 가정하에 랜덤하게 칼럼을 선택하고, 선택된 칼럼의 최댓값과 최솟값을 분리하는 값을 랜덤으로 선택한 후, 의사결정트리를 연속으로 분기시키면서 데이터 관측치의 고립 정도에 따라 이상치를 판별한다.

이상치 처리 방법으로는 평균의 표준편차를 이용한다. 99.7% 이상 값 (lv) 혹은 이하 값(uw)을 이상치로 제거하거나 대체한다. Where μ denotes the mean, n denotes the total number of data, σ denotes the standard deviation.

$$lv = \mu - n \times \rho$$

$$uw = \mu + n \times \rho \quad (2)$$

3.2 데이터 분석(Data Analysis)

데이터 분석을 위하여, 우선, 분포를 확인하고, 변수 간의 관계를 확인하기 위하여 통계량을 파악한다. 우선 각 공정 인자의 특성을 분석하고, 각 상관관계 분석을 통하여 인자의 공정 내 역할과 사출 공정의 불량률 감소 조건을 분석한다.

3.2.1 통계량 파악

Fig. 2와 같이 수집된 데이터의 칼럼별 개수, 평균, 최솟값, 하위 25% 값, 중간값, 상위 25%, 최댓값을 확인한다. 우선 0과 1로 이루어진 데이터들을 확인해야 한다. Result의 평균을 보면 0.79 정도로 합격데이터가 많다는 것을 알 수 있다. 양품 데이터가 많지만, 불균형데이터는 아니므로 사용해도 무방하다. 그리고 과도하게 큰 값이나 작은 값인 이상치가 없는 것을 확인한다.

	MOLD_Temp	ITemp_WH	OTemp_WH	DTemp_RM	IPress_Injec	OPress_Injec	Temp_Fac	Humi_Fac	Result
count	1067.000000	1067.000000	1067.000000	1067.000000	1067.000000	1067.000000	1067.000000	1067.000000	1067.000000
mean	37.712277	175.011246	165.026992	200.425492	1378.833177	496.846298	27.787062	38.819021	0.799438
std	1.481806	0.590088	0.194200	3.039895	9.820065	11.365776	2.304367	8.091801	0.400609
min	32.000000	174.100000	164.500000	198.500000	1365.000000	471.000000	21.060420	23.368170	0.000000
25%	37.000000	174.900000	164.800000	199.900000	1372.000000	498.000000	26.116925	34.157735	1.000000
50%	38.000000	175.000000	165.000000	200.000000	1378.000000	500.000000	27.525720	40.126180	1.000000
75%	38.000000	175.100000	165.100000	200.100000	1388.000000	503.000000	29.298025	46.002235	1.000000
max	42.000000	193.300000	166.700000	220.800000	1400.000000	512.000000	33.303210	62.553680	1.000000

Fig. 2. Data Variable Structure

3.2.2 상관관계 파악

Fig. 3과 같이 상관관계 분석을 통하여 독립변수와 종속변수 사이에 선형적인 관계가 존재하는지를 확인할 수 있다. 각 사출 독립변수가 종속변수에 어떤 영향을 미치는가를 확인한다. 종속변수와 상관계수(Correlation Coefficient)가 1이나 -1에 근접하는 입력 인자들을 선정한다. Result는 Result는 Opress_Injec와 Ipress_Injec에 강한 양의 상관관계, OTemp_WH에 약한 양의 상관관계, MOLD_Temp에 약한 음의 상관관계가 있음을 알 수 있다.

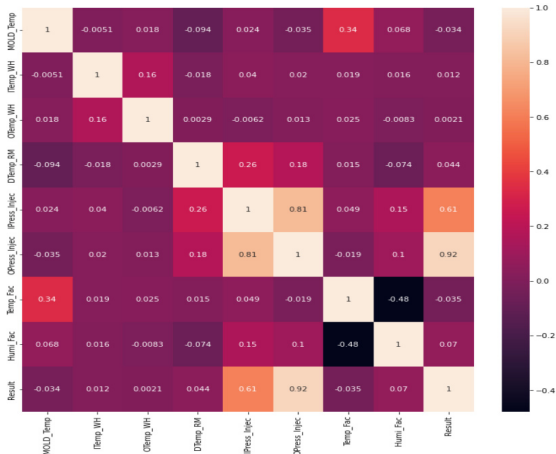


Fig. 3. Correlation matrix

3.2.3 다중공선성 파악

각 독립변수 간의 상관분석(Correlation Analysis) 수행 시에, 각 인자의 상관관계가 클수록 다중공선성(Multicollinearity) 문제가 발생한다. 다중공선성이 존재하면 독립변수의 공분산 행렬의 조건수(Conditional number)가 증가하고, 이는 부정확한 회귀 결과가 도출될 수 있다.

OLS Regression Results						
Dep. Variable:	Result	R-squared:	0.897			
Model:	OLS	Adj. R-squared:	0.896			
Method:	Least Squares	F-statistic:	1039.			
Date:	Wed, 09 Nov 2022	Prob (F-statistic):	0.00			
Time:	18:25:23	Log-Likelihood:	604.35			
No. Observations:	960	AIC:	-1191.			
DF Residuals:	951	BIC:	-1147.			
DF Model:	8					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	6.3140	3.664	1.723	0.085	-0.877	13.505
MOLD_Temp	0.0031	0.003	0.997	0.319	-0.003	0.009
ITemp_WH	0.0025	0.007	0.371	0.711	-0.011	0.016
OTemp_WH	-0.0307	0.022	-1.428	0.154	-0.073	0.012
DTemp_RM	-0.0121	0.001	-8.156	0.000	-0.015	-0.009
IPress_Injec	-0.0145	0.001	-18.763	0.000	-0.016	-0.013
OPress_Injec	0.0431	0.001	67.482	0.000	0.042	0.044
Temp_Fac	-0.0008	0.002	-0.354	0.724	-0.005	0.004
Humi_Fac	-0.0006	0.001	-1.034	0.301	-0.002	0.001
Omnibus:	790.176	Durbin-Watson:	1.948			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	17674.328			
Skew:	-3.715	Prob(JB):	0.00			
Kurtosis:	22.664	Cond. No.:	1.32e+06			

Fig. 4. OLS Regression Results and Multicollinearity

Fig. 4에서 Cond. No.(조건수)를 확인한다. 다중공선성이 있으면 하단 부분에 "[2] The condition number is large, 1.32e+06. This might indicate that there are strong multicollinearity or other numerical problems"를 확인할 수 있다.

다중공선성을 제거하기 위하여, Fig. 5와 같이 VIF(Variance Inflation Factor)을 확인한다. VIF는 독립변수를 다른 독립변수로 선형회귀한 성능을 나타낸 것으로서, VIF 값이 10 이상이면 일반적으로 다중공선성이 있다고 판단한다. i 번째 변수의 VIF는 Eq. (3)과 같다.

$$VIF_i = \frac{1}{1 - R_i^2} \quad (3)$$

R_i^2 는 다른 변수로 i 번째 변수를 선형회귀한 성능(결정 계수)이다. 다른 독립변수에 의존적일수록 VIF가 커진다.

VIF Factor	features
0	805.833010 MOLD_Temp
1	88260.416410 ITemp_WH
2	113775.408996 OTemp_WH
3	5122.756179 DTemp_RM
4	87077.684402 IPress_Injec
5	34986.874558 OPress_Injec
6	240.055077 Temp_Fac
7	37.863054 Humi_Fac
8	50.820579 Result

Fig. 5. VIF Factor

독립변수들을 Z-점수 정규화(Z-Score Normalization)를 이용하여 정규화하고 다중공선성을 다시 확인한다.

만일 다중공선성이 여전히 존재한다면 PCA(Principle Component Analysis)를 이용한 개각행렬(Diagonal Matrix)의 형태로 의존적인 성분을 삭제하고 공선성을 제거한다. PCA는 SVD(Singular Value Decomposition)를 적용하여 데이터 '주성분'을 추출하고, 이 주성분 데이터를 이용하여 분석을 진행한다. Fig. 6은 정규화와 PCA를 적용한 후 다중공선성을 확인한 결과이다. 조건수가 1.53이고, "[1] Standard Errors assume that the covariance matrix of the errors is correctly specified."로서 다중공선성이 제거된 것을 확인할 수 있다.

OLS Regression Results						
Dep. Variable:	Result	R-squared:	0.595			
Model:	OLS	Adj. R-squared:	0.592			
Method:	Least Squares	F-statistic:	238.4			
Date:	Tue, 08 Nov 2022	Prob (F-statistic):	7.64e-159			
Time:	16:06:54	Log-Likelihood:	-49.127			
No. Observations:	853	AIC:	110.3			
Df Residuals:	847	BIC:	138.7			
Df Model:	5					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.8030	0.009	91.153	0.000	0.788	0.820
principal component1	0.2181	0.006	33.427	0.000	0.201	0.226
principal component2	-0.0300	0.007	-4.259	0.000	-0.044	-0.016
principal component3	0.0037	0.008	0.467	0.640	-0.012	0.019
principal component4	-0.0288	0.008	-3.512	0.000	-0.045	-0.013
principal component5	0.0539	0.009	5.855	0.000	0.036	0.072
Omnibus:	38.721	Durbin-Watson:	1.938			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	24.831			
Skew:	-0.287	Prob(JB):	4.05e-06			
Kurtosis:	2.393	Cond. No.	1.46			
Notes:						
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.						

Fig. 6. OLS Regression Results after Normalization and PCA

3.2.4 불량 예측 모델

Keras package를 활용한 DNN(Deep Neural Network) 모델은 Fig. 7과 같다.

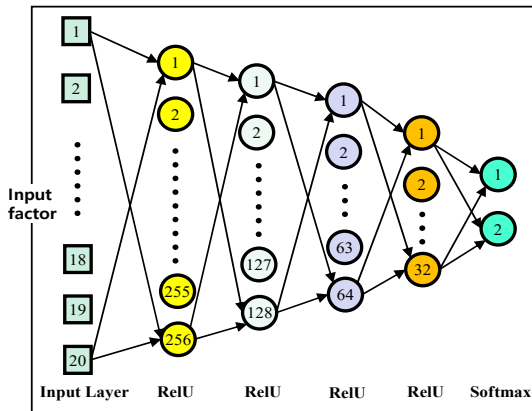


Fig. 7. DNN Model

Fig. 8과 같이 Sequential()은 순차적으로 신경망을 추가하는 모델을 정의한다. model.add()의 Dense(256)은 은닉층의 노드가 256개인 신경망을 의미하고, 'activation'은 은닉층의 활성화 함수로서 relu를 나타내며, 'input_dim'은 입력할 데이터의 노드 수로서 20을 지정한다. 출력층은 2 노드(양품, 불량품)이므로 Dense(2)이고, 출력층에서 출력 대상의 확률을 계산하기 위하여 소프트맥스 함수를 사용한다. 모델 학습 전에, compile 메소드를 통해서 학습 방법에 대한 환경설정이 필요하다. 정규화기(Optimizer)는 매개변수(가중치)의 최적값을 찾는 방법으로 AME(Adaptive Moment Estimation)을 적용한다. 손실 함수(loss Function)는 이진 교차엔트로피(binary_crossentropy)를 사용했으며, 실제 클래스 변수(0 또는 1)와 근접할 때 낮은 값을 가지고 반대의 경우에 큰 값을 가지게 된다. 분류 문제에 대해서는 metrics=['accuracy']로 설정한다. 기준은 문자열 식별자 또는 사용자 정의 기준 함수를 사용할 수 있다.

```

model = Sequential()
model.add(Dense(256, activation='relu', input_dim=20))
model.add(Dense(128, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(64, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(32, activation='relu'))
model.add(Dense(2, activation='softmax'))
model.compile(loss='(binary_crossentropy',
optimizer='adam', metrics=['accuracy'])
    
```

Fig. 8. Functions of the DNN Model

4. 성능 분석

Table 1에 제조 데이터 분석을 위한 주요 공정 데이터를 보인다[1]. 주요 인자를 중심으로 기술하였으며, 분석 결과를 제시한다.

데이터 전처리 및 분석된 데이터를 기반으로 DNN을 학습시키고 결과를 확인한다. 추가로 동일 데이터를 이용하여 SVM(Support Vector Machine), SGBost (Extreme Gradient Boosting), RF(Random Forest)를 학습시킨 후에, 각각의 혼동 행렬(Confusion Matrix)을 구하여 성능을 비교한다.

Table 1. Process factors for analysis

Classification	Category		
Data type	Text, Number		
Data Properties	Approximately 1400 numerical data of injection molding equipment		
Data collection method	1. PLC, Temperature and Humidity sensor 2. Receive cycle: By Lot, Every 20 seconds		
Main parameters	Mold surface temperature	°C	MOLD_Temp
	Input temperature to the water heater	°C	ITemp_WH
	Output temperature to the water heater	°C	OTemp_WH
	the drying temperature of the raw material	°C	DTemp_RM
	Input pressure to the injection device	kgf/cm ²	IPress_Injec
	Output pressure to the injection device	kgf/cm ²	OPress_Injec
	Temperature in the factory	°C	Temp_Fac
Humidity in the factory	g/m ²	Humi_Fac	

양품과 불량품의 분류 결과를 해석하기 위하여 혼동행렬(Confusion Matrix)을 이용한다. 혼동행렬에서 사용하는 주요 지표는 아래와 같다. 여기서, TP(True Positive)는 실제 양품을 양품으로 예측한 것을, FP(False Positive)는 실제 불량품을 양품으로 예측한 것을, FN(False Negative)은 실제 양품을 불량품으로 예측한 것을 의미한다.

- 정확도(Accuracy): 전체 데이터 중 실제값과 예측값이 같은 건수가 차지하는 비율로서 $\frac{TN+TP}{TN+FP+FN+TP}$ 이다.
- 정밀도(Precision)는 양품으로 예측한 경우 중 실제로 양품인 비율로서 $\frac{TP}{TP+FP}$ 이다. 즉 예측값이 얼마나 정확한가를 추정한다.
- 재현율(Recall)은 실제 양품인 것 중에서 양품으로 정확히 예측한 비율로서 $\frac{TP}{TP+FN}$ 이다. 실제 정답의 비율을 추정한다.
- F1 점수(Score)는 정밀도와 재현율의 조화평균으로 $2 \times \frac{Precision \times Recall}{Precision + Recall}$ 이다. 분류 클래스 간의 데이터가 불균형이 클 때 사용한다. 값이 1에 근접할수록 우수한 모델이다.

Fig. 8은 DNN 모델을 적용하여 양품과 불량품을 예측한 결과이다. 여기서 support는 각 라벨(양품(1), 불량

품(0))의 실제 샘플 개수이다. 정확도, 정밀도, 재현율은 0.92 ~ 1.0 사이의 값을 가지며, 정밀도가 1.0이라는 의미는 FP가 0건이라는 의미이고, 재현율이 1.0이라는 의미는 FN이 0건이라는 의미이다.

정밀도와 재현율은 서로 보완적인 지표로서, 모두 수치가 높고, 균등하다면 우수한 성능을 의미한다. 반면 정밀도(재현율)은 매우 높고, 재현율(정밀도)이 낮다면 이는 보완이 필요한 결과이다. Fig. 8과 같이 DNN의 정밀도는 양품 100%, 불량품 98%, 재현율은 양품 92%, 불량품 100%, F1 점수는 양품 96%, 불량품 99%로서, 정밀도와 재현율이 높고, 거의 균등하게 유지됨을 알 수 있다.

	precision	recall	f1-score	support
0	1.00	0.92	0.96	62
1	0.98	1.00	0.99	259
accuracy			0.98	321
macro avg	0.99	0.96	0.97	321
weighted avg	0.98	0.98	0.98	321

Fig. 9. Statistics by Class Variables of DNN

Table 2 ~ Table 5는 정제된 데이터를 기반으로 AI 분석을 적용한 결과이다. 321개의 데이터로 평가하였고, DNN의 경우, 정확도, 재현율 등 여러 평가지표에서 SVM, RF, XGBoost 보다 우수한 결과를 보인다.

Table 2. Confusion Matrix in DNN

Classification		Prediction	
		good	bad
Reality	good	259	0
	bad	5	57

Table 3. Confusion Matrix in SVM

Classification		Prediction	
		good	bad
Reality	good	252	7
	bad	15	47

Table 4. Confusion Matrix in RF

Classification		Prediction	
		good	bad
Reality	good	256	3
	bad	8	54

Table 5. Confusion Matrix in SGBost

Classification		Prediction	
		good	bad
Reality	good	256	3
	bad	11	51

5. 결론

전 세계적으로 제조업은 생산 비용 절감이 거세게 요구되고 있으며, 동시에 제품의 제조 원가와 공정 과정은 투명하게 공개되어 제품의 가격 대비 성능에 대한 기대치는 계속해서 높아지고 있다. 따라서 제조업 전반에 걸쳐서 이미 효율성을 갖추고 있는 생산 체계를 한 단계 더 효율화해야 하는 과제에 당면해 있다. 본 연구에서는 사출 공정의 품질, 생산성에 대한 이상 징후 발생 시 대응력을 높이고 품질 및 생산성을 최적화하기 위해 실시간으로 발생하는 제조 데이터를 분석하여 최적의 품질, 생산성 및 유연성을 가지는 사출성형 운영기술을 개발하고자 하였다.

추후 연구로서 품질의 불량 원인을 사후 분석이 아닌 제조 공정 중 실시간 분석으로 분석하는 방법을 제시하고자 한다. 그리고 사출 업체의 생산설비, 제품, 생산환경, 작업자로부터 얻어지는 빅데이터의 분석을 기반으로 설비의 최적 공정조건을 자동으로 결정하는 방법을 제시하고자 한다.

References

[1] G. S. Lee, J. C. Lee, "Data Analytics in Manufacturing", Journal of the Korea Academia-Industrial cooperation Society, vol. 23, no. 10, pp. 705-711, 2022.
DOI: <https://doi.org/10.5762/KAIS.2022.23.10.705>

[2] B. Wang, F. Tao, "Smart Manufacturing and Intelligent Manufacturing: A Comparative Review" Elsevier BV, vol. 7, no. 6, pp. 738-757, 2021.
DOI: <https://doi.org/10.1016/j.eng.2020.07.017>

[3] Y. Koren, X. Gu, W. Guo, "Reconfigurable Manufacturing Systems: Principles, Design, and Future Trends", Frontiers of Mechanical Engineering, vol. 13, pp. 121-136, 2018.
DOI: <https://doi.org/10.1007/s11465-018-0483-0>

[4] S. Y. Kim, "A Case Study of the Introduction of Smart Factory Operation Management (FOM) in the Fourth Industrial Revolution Era", Journal of Korean

Association of Computer And Accounting, vol. 16, no. 1, pp. 43-62, 2018.

DOI: <https://doi.org/10.32956/kaoca.2018.16.1.43>

[5] K. S. Noh, S. W. Park, "An Exploratory Study on Application Plan of Big Data to Manufacturing Execution System", The Society of Digital Policy and Management, vol. 12, no. 1 pp. 305-311, 2014.
DOI: <https://doi.org/10.14400/JDPM.2014.12.1.305>

[6] S. Liang, M. Rajora, X. Liu, C. Yue, P. Zou, L. Wang "Intelligent Manufacturing Systems: A Review", International Journal of Mechanical Engineering and Robotics Research, vol. 7, no. 3, pp. 324-330, 2018.
DOI: <https://doi.org/10.18178/ijmerr.7.3.324-330>

[7] J. G. Baek, "A Study on the Visualization of Facility Data Using Manufacturing Data Collection Standard", Research Report, Korea University, 2016.

[8] B. Wang, F. Tao, X. Fang, C. Liu, T. Freiheit, "Smart Manufacturing and Intelligent Manufacturing: A Comparative Review", Engineering, vol. 7, no. 6, pp. 738-757, 2020.
DOI: <https://doi.org/10.1016/j.eng.2020.07.017>

[9] S. Mittal, M. A. Khan, D. Romero, T. Wuest, "A Critical Review of Smart Manufacturing & Industry 4.0 Maturity Models: Implications for Small and Medium-sized Enterprises (SMEs)", Journal of Manufacturing Systems, vol. 49, pp. 194-214, 2018.
DOI: <https://doi.org/10.1016/j.jmsv.2018.10.005>

[10] D. Mourtzis, E. Vlachou, N. Milas, "Industrial Big Data as a Result of IoT Adoption in Manufacturing", Procedia CIRP, vol. 55, pp. 290-295, 2016.
DOI: <https://doi.org/10.1016/j.procir.2016.07.038>

[11] H. Y. Cho, "On-line Process Data-driven Diagnostics Using Statistical Techniques", Journal of the Korea Academia-Industrial cooperation Society, vol. 19, no. 3 pp. 40-45, 2018.
DOI: <https://doi.org/10.5762/KAIS.2018.19.3.40>

[12] D. Y. Lee, J. W. Park, S. H. Cho, J. S. Lee, "Maintainability Prediction of Guided Missile based on Machine Learning using Field Data", Journal of the Korea Academia-Industrial cooperation Society, vol. 23, no. 2 pp. 518-526, 2022.
DOI: <https://doi.org/10.5762/KAIS.2022.23.2.518>

이 기 성(Gi-Sung Lee)

[중신회원]



- 1996년 2월 : 송실대학교 컴퓨터 과학과 (공학석사)
- 2001년 2월 : 송실대학교 컴퓨터 과학과 (공학박사)
- 2001년 3월 ~ 현재 : 호원대학교 컴퓨터게임학과 교수

<관심분야>

모바일 멀티미디어, 네트워크 보안, 머신러닝

이 중 찬(Jong-Chan Lee)

[정회원]



- 1996년 8월 : 송실대학교 컴퓨터 과학과 (공학석사)
- 2000년 8월 : 송실대학교 컴퓨터 과학과 (공학박사)
- 2000년 10월 ~ 2005년 2월 : 한국전자통신연구원 선임연구원
- 2005년 3월 ~ 현재 : 군산대학교 컴퓨터정보공학과 교수

<관심분야>

블록체인, 빅데이터, 머신러닝