

박사학위논문

실시간 빅데이터 수집 및 AI 모델링 분석  
기반 스마트공장 고도화 연구

- IoT통합 기반의 빅데이터 수집 및 AI 모델링과  
SPC 분석 기술의 중심으로 -

2022년

한 성 대 학 교 대 학 원

스마트융합건설팅학과

스마트융합제품전공

김 재 표

박사학위논문  
지도교수 김승천

# 실시간 빅데이터 수집 및 AI 모델링 분석 기반 스마트공장 고도화 연구

- IoT통합 기반의 빅데이터 수집 및 AI 모델링과  
SPC 분석 기술의 중심으로 -

A Study on Advancement of Smart Factory Based on  
Real-Time Big Data Collection and AI Modeling Analysis

2021년 12월 일

한 성 대 학 교 대 학 원

스마트융합컨설팅학과

스마트융합제품전공

김 재 표

박사학위논문  
지도교수 김승천

# 실시간 빅데이터 수집 및 AI 모델링 분석 기반 스마트공장 고도화 연구

- IoT통합 기반의 빅데이터 수집 및 AI 모델링과  
SPC 분석 기술의 중심으로 -

A Study on Advancement of Smart Factory Based on  
Real-Time Big Data Collection and AI Modeling Analysis

위 논문을 공학 박사학위 논문으로 제출함

2021년 12월 일

한 성 대 학 교 대 학 원

스마트융합건설팅학과

스마트융합제품전공

김 재 표

김재표의 공학 박사학위 논문을 인준함

2021년 12월 일

심사위원장 노 광현 (인)

심 사 위 원 이 후진 (인)

심 사 위 원 김 영주 (인)

심 사 위 원 임 황용 (인)

심 사 위 원 김 승천 (인)

# 국 문 초 록

## 실시간 빅데이터 수집 및 AI 모델링 분석 기반 스마트공장 고도화 연구 - IoT통합 기반의 빅데이터 수집 및 AI 모델링과 SPC 분석 기술의 중심으로 -

한 성 대 학 교 대 학 원  
스 마 트 용 합 컨 설 팅 학 과  
스 마 트 용 합 제 품 전 공  
김 재 표

최근 스마트공장은 빅데이터와 AI 기술을 활용한 고도화의 추세로 전환하고 있다. 이에 따라 빅데이터 수집 및 처리 기술, IoT 통합 기술, 기계 학습 및 AI 모델링에 대한 선행 연구가 수행되어 왔다. 이러한 기존 문헌 검토를 바탕으로, 본 연구의 범위와 목적에 맞게 실시간 제조 데이터 수집, 분석 및 활용 시스템을 설계한 후, 기계 가공 업종으로 스마트공장 고도화를 설계하고 구현하였다.

첫째, 가공공정 스마트화를 위한 레이아웃 재배치 및 공정개선, 신설비, 측정검사시스템 도입을 통해 신공정을 설계하고 구현하였다. 신 공정 지그, 로봇 및 자동 컨베이어 이송장비, MCT 가공설비 5대, 부시 압력 측정 및 조립시스템, 공구 모니터링 및 검사시스템을 신규 도입 또는 설계, 제작, 설치하였다.

둘째, 실시간으로 제조 빅데이터의 수집 및 분석시스템을 설계하기 위해서,

먼저 표준 통신 프로토콜인 OPC-UA를 기반으로 한 IoT 통합 모듈로 실시간 데이터 수집을 위한 네트워크를 설계하고 구현하였다.

셋째, 데이터 소프트웨어는 3계층 구조와 스마트 클라이언트 개념의 애플리케이션 브라우저를 클라우드와 연계하여 설계하고 구현하였다. 신규 스마트화 공정에 대한 사용자의 요구사항을 반영한 수집 분석 장표를 만들고, 이를 반영하여 분석데이터 수집·저장을 위한 데이터베이스 및 추출 솔루션을 설계하고 구현하였다.

넷째, SPC 통계와 시각화 솔루션을 설계한 후 구현하였고 주요관리 변수와 AI 모델링으로 도출된 최상위 품질 영향변수에 대해서 교차 검증을 수행하였다.

다섯째, AI 모델링, 검증 및 시각화 분석을 설계하고 수행하였다. 변수들에 대한 상대적 상관분석을 실시하고, 알고리즘 간 예측정확도를 비교하였다. 주요 변수의 중요도에 대한 F1 스코어를 SPC 통계분석과 교차 검증한 결과, 부합한 것으로 판명되었다.

이상의 내용에서 실시간 데이터 수집, 저장을 위한 데이터베이스 및 추출 솔루션이 MES에 설계하고 구축되었다. 이 솔루션을 기반으로 빅데이터 분류, 저장을 통한 체계적인 데이터 축적 및 추출이 가능한 시스템이 구축되어 장기적인 측면에서 생산성과 품질 향상을 위한 기반이 되었다.

XGBoost 피처의 주요변수 중요도에 대한 F1 스코어가 SPC분석의 결과 값과 부합한 것으로 평가되었다. SPC 통계 검증으로 불량요인을 명확히 하고 AI 모델링의 예측결과에 대한 교차 검증을 수행하면서 사후 유지관리의 수행이 가능할 것으로 예상된다. AI 알고리즘의 예측 정확도에 대해서는, 향후에 데이터가 축적되고 학습이 지속적이고 반복적으로 이루어지면, 랜덤포레스트 알고리즘의 정확도가 높아질 것으로 예상된다.

**【주요어】** 빅데이터, AI모델링, 데이터 수집분석 시스템, IoT통합, SPC, 스마트공장 고도화, OPC-UA, MES, 기계가공, 애플리케이션 브라우저, 스마트 클라이언트, 클라우드, AI알고리즘, 통신프로토콜, 지도학습

# 목 차

<b>제 1 장 서론</b> .....	1
제 1 절 연구의 배경 .....	1
제 2 절 연구의 필요성 .....	7
제 3 절 연구의 범위와 목적 .....	10
1) 연구의 방법과 계획 .....	10
2) 연구의 범위 .....	11
3) 연구의 목적 .....	13
<b>제 2 장 관련 기술 및 선행연구</b> .....	14
제 1 절 빅데이터 수집 및 처리 연구 .....	14
1) 빅데이터 수집 및 처리 분석 관련 기술 .....	14
2) 빅데이터 수집 및 처리 선행연구 및 결과 .....	19
제 2 절 IoT 통합(Integration) 및 통신 프로토콜 연구 .....	31
1) IoT 통합 및 통신 프로토콜 관련 기술 .....	31
2) IoT 통합 및 통신 프로토콜 선행연구 및 결과 .....	37
제 3 절 기계학습 및 AI 학습 연구 .....	49
1) 기계학습 및 AI 학습 관련 연구 .....	49
2) 기계학습 및 AI 학습 선행연구 및 결과 .....	52
제 4 절 선행연구 기술 및 시사점 종합 .....	66
1) 빅데이터 수집 및 처리 선행연구 기술 및 시사점 .....	66
2) IoT 통합 관련 선행연구 기술 및 시사점 .....	67
3) 기계학습 및 AI 학습 관련 선행연구 기술 및 시사점 .....	67
4) 종합 정리 및 결과 .....	68
<b>제 3 장 제조데이터 수집분석 및 활용 시스템 설계</b> .....	69
제 1 절 제조데이터 수집분석 및 활용 시스템 설계 개요 .....	69
1) 제조데이터 수집, 분석 구조 설계 개요 .....	69
2) 기계가공 데이터 수집 분석 및 활용을 위한 설계 아키텍처 .....	71
3) 기계가공 스마트화 공정 선정 배경 .....	72
4) 기계가공 스마트화 공정 라인 및 데이터 수집 체계 설계 .....	74

제 2 절	스마트화 가공공정 상세 설계 .....	76
1)	제조공정 스마트화를 위한 공정 레이아웃 재배치 및 설계 .....	76
2)	부시 치수 및 압력 측정, 조립시스템 구축을 위한 설계 .....	77
3)	공구 모니터링 및 검사시스템 구축을 위한 설계 .....	79
제 3 절	제조데이터 수집, 분석 구조 상세 설계 .....	81
1)	데이터 수집을 위한 OPC-UA 기반 IoT 통합 설계 .....	81
2)	데이터 수집 SW 개발 .....	83
3)	데이터 수집 및 분석시스템 설계 .....	87
제 4 절	SPC 통계분석 & AI 모델링 분석 및 시각화 설계 .....	95
1)	SPC 통계분석 및 시각화 .....	95
2)	AI 모델링 분석 및 시각화 .....	97
제 5 절	제조데이터 수집분석 및 활용 시스템 종합 설계 .....	103
<b>제 4 장</b>	<b>제조데이터 수집분석 및 활용 시스템 구현 .....</b>	<b>105</b>
제 1 절	스마트화 가공공정 구현 .....	105
1)	제조공정 스마트화를 위한 공정 레이아웃 재배치 및 개선 구현 .....	105
2)	부시 치수 및 압력 측정, 조립 시스템 구축 .....	106
3)	공구 모니터링 및 검사시스템 구축 .....	108
제 2 절	제조데이터 수집분석 및 활용 시스템 구현 .....	112
1)	데이터 수집을 위한 OPC-UA 기반 IoT 통합 구현 .....	112
2)	데이터 수집 SW 구현 .....	114
3)	데이터 수집 및 분석시스템 개발 구현 .....	115
제 3 절	SPC 통계 & AI 모델링 분석 및 시각화 구현 .....	120
1)	SPC 통계 솔루션 및 시각화 구현 .....	120
2)	AI 모델링 분석 및 시각화 .....	122
3)	SPC 통계 솔루션과 연계한 통계분석 .....	141
제 4 절	제조데이터 수집 분석 및 활용시스템 구현의 성능 검증 .....	151
1)	SPC 통계 구현에 대한 분석 검증 .....	151
2)	AI 알고리즘 예측 정확도 성능 평가 .....	154
3)	XGBoost 피처의 중요도 모델에 대한 검증 .....	157
제 5 절	연구의 성과 .....	161
1)	빅데이터 수집분석 시스템의 설계 및 구현의 시사점과 성과 .....	161
2)	SPC & AI모델링 분석의 설계 및 구현의 시사점과 성과 .....	162
3)	연구의 기대효과 .....	162

제 5 장 결 론 .....	164
참 고 문 헌 .....	167
부 록 .....	174
ABSTRACT .....	184

## 표 목 차

[표 1-1] 스마트공장의 핵심기술 분류 및 정의 .....	4
[표 1-2] 연구의 목표 .....	13
[표 2-1] 데이터베이스의 종류별 차이점 .....	15
[표 2-2] 2018년 학회지에 사용한 분석 소프트웨어 리스트 .....	16
[표 2-3] 로봇 팔 데이터 .....	24
[표 2-4] 웹 클라이언트 및 타 솔루션과의 비교 .....	29
[표 2-5] 선행연구 주요 내역 .....	30
[표 2-6] 제조용 근거리 무선통신 기술 비교 .....	32
[표 2-7] LPWA관련 주요 무선 통신 기술 비교 .....	33
[표 2-8] HTTP, MQTT 및 CoAP의 특성 비교 .....	35
[표 2-9] 선행연구 주요 내역 .....	48
[표 2-10] 인공지능과 빅데이터 기술 분류표 .....	49
[표 2-11] AI학습 프로젝트 추진 단계와 절차 .....	51
[표 2-12] 공구마모도 특성 데이터 설명 .....	53
[표 2-13] 공구마모도 예측성능 분석 결과 .....	54
[표 2-14] 센서 구성표 .....	57
[표 2-15] 예측모델 성과지표 .....	58
[표 2-16] 모델별 속성 중요도 .....	58
[표 2-17] A공정의 가열온도, 주입속도, 공기흡입컵 크기감소량 .....	59
[표 2-18] 선행연구 주요 내역 .....	65
[표 2-19] 빅데이터 수집 및 처리 선행연구 기술 및 시사점 내역 .....	66
[표 2-20] IoT 통합 관련 선행연구 기술 및 시사점 내역 .....	67
[표 2-21] 기계학습 및 AI학습 관련 선행연구 기술 및 시사점 내역 .....	67
[표 3-1] 데이터 수집 항목 분석결과 .....	88
[표 4-1] 개선공정 정의 및 가공 범위 .....	106
[표 4-2] 공구 모니터링 및 검사시스템 주요 구현 기술 및 사양 .....	109
[표 4-3] 3차원 측정기 측정 결과; CON ROD CMM 검사성적서 .....	123

[표 4-4] CON ROD가공 > 공구부하량 .....	124
[표 4-5] 공구 사용 및 교환 이력 .....	124
[표 4-6] 변수의 정의 및 설명 .....	127
[표 4-7] 상관계수 분석 .....	128
[표 4-8] 공정능력 결과 해석 주요 기준 내용 .....	142
[표 4-9] 관리도 해석 주요 내용 .....	143
[표 4-10] 볼트홀 대칭도 Cpk 및 관리도 종합 분석표 .....	145
[표 4-11] 우측편향 y2 Cpk 및 관리도 종합 분석표 .....	148
[표 4-12] SPC 통계분석 종합 결과 .....	150
[표 4-13] 공정능력 결과 해석 주요 기준 내용 .....	151
[표 4-14] 관리도 해석 주요 내용 .....	151
[표 4-15] SPC통계분석 종합 요약표 .....	152
[표 4-16] 알고리즘 예측 정확도 수치 테이블 .....	156
[표 4-17] F1 스코어와 SPC 분석 비교표 .....	159

## 그림 목 차

[그림 1-1] 스마트공장 구축 수준별 정의 .....	2
[그림 1-2] 스마트공장 아키텍처 .....	8
[그림 1-3] 제조 데이터의 통신 이슈 .....	9
[그림 1-4] 단계별 제조데이터 활용모델 구축 계획 및 활용 도구 .....	11
[그림 1-5] 제조데이터 수집분석 및 활용 시스템 아키텍처 .....	12
[그림 2-1] 빅데이터 활용 프로세스 및 관련 기술 .....	14
[그림 2-2] 데이터 발생에서부터 시각화 플랫폼까지의 플로우 .....	16
[그림 2-3] 3계층 및 2계층 구조 아키텍처 .....	17
[그림 2-4] 스마트 클라이언트 특징 .....	18
[그림 2-5] 애플리케이션 브라우저 구조 .....	19
[그림 2-6] 빅데이터 분석플랫폼의 개념도 .....	20
[그림 2-7] 빅데이터 분석 플랫폼의 아키텍처 .....	21
[그림 2-8] 빅데이터 기반 센서 데이터 처리 및 분석 시스템 구조도 .....	22
[그림 2-9] 몽고DB의 로봇 팔 데이터 .....	24
[그림 2-10] 애저 플랫폼 기반 스마트팩토리 구성도 .....	26
[그림 2-11] SAP가 제안하는 스마트팩토리 분석 아키텍처 .....	26
[그림 2-12] SAP HANA 플랫폼 기반 스마트 팩토리 구성도 .....	27
[그림 2-13] 스마트 센서 아키텍처 및 형태 .....	31
[그림 2-14] 사물 인터넷 플랫폼 기술 특징 및 발전 방향 .....	34
[그림 2-15] 삼성 SDS의 Brightics IoT .....	34
[그림 2-16] OPC-UA 플랫폼 계층 구조 .....	36
[그림 2-17] PLCopen OPC-UA 프레임워크 개요도 .....	38
[그림 2-18] PLCopen OPC-UA 프레임워크에서 OPC-UA 정보 모델 .....	39
[그림 2-19] 시스템 구성도 .....	40
[그림 2-20] FANUC의 시퀀스 다이어그램 .....	41
[그림 2-21] SIEMENS의 시퀀스 다이어그램 .....	41
[그림 2-22] 저비용 통합 데이터 관리 시스템(LDMS) 아키텍처 .....	43

[그림 2-23] 실시간 모니터링 시스템 구성도 안 .....	44
[그림 2-24] IoT 디바이스 구성도 및 PC 실행 소프트웨어의 전체 구조 .....	45
[그림 2-25] IoT 게이트웨이의 프로토콜 레이어 구성도 .....	46
[그림 2-26] 이기종 네트워크 프로토콜 인터페이스 게이트웨이 구성도 .....	47
[그림 2-27] 훈련용 이미지 셋 .....	55
[그림 2-28] Proposed Method .....	55
[그림 2-29] 래피드마이너 활용한 데이터 전처리 .....	57
[그림 2-30] LSTM 전체 구성도 .....	60
[그림 2-31] 전체 학습 손실 값 그래프 .....	60
[그림 2-32] DBSCAN기반의 A공정 군집 분석 .....	61
[그림 2-33] window size 별 평균값 .....	62
[그림 2-34] 불량 예측 그래프와 RMSE .....	62
[그림 2-35] OLS 회귀 분석 결과표 .....	63
[그림 3-1] 제조 데이터 수집 및 분석 및 활용모델 .....	69
[그림 3-2] 기계가공 데이터 수집 및 활용을 위한 시스템 아키텍처 .....	71
[그림 3-3] 산업용 엔진 및 커넥팅 로드, 부시 관련 조립 부품 .....	72
[그림 3-4] 커넥팅 로드 홀 편심 .....	73
[그림 3-5] 현재 제조 공정도(9단계) .....	73
[그림 3-6] 커넥팅로드 가공 공차강화 공정 개선도 .....	74
[그림 3-7] 커넥팅로드 데이터 획득시스템 구축 구성도 .....	75
[그림 3-8] Con-Rod 가공공정 개선 전 레이아웃 .....	76
[그림 3-9] 부시 압입 압력 측정 및 전송시스템 구성도 .....	77
[그림 3-10] 부시 치수(내경) 측정 원리 .....	78
[그림 3-11] 공구 모니터링 및 검사시스템 구성도 .....	79
[그림 3-12] 공구 모니터링 데이터 추출 및 연관관계 분석 .....	80
[그림 3-13] IIoT 디바이스와 시스템 구성도 .....	81
[그림 3-14] 제조공정 데이터 수집 및 분석 시스템 네트워크 .....	82
[그림 3-15] 3계층 구조 및 엔터프라이즈 솔루션 프레임 구성 .....	83
[그림 3-16] 스마트 클라이언트 애플리케이션 브라우저 구조 .....	85

[그림 3-17] 데이터 수집 SW 기능 구성도 .....	86
[그림 3-18] 빅데이터 수집 및 분석시스템 구조도 .....	87
[그림 3-19] 가공, 품질 측정 데이터 수집 분석 장표 .....	89
[그림 3-20] 공구 정보 데이터베이스 ERD .....	90
[그림 3-21] 커넥팅로드 정보 데이터베이스 ERD .....	91
[그림 3-22] 커넥팅로드 정보를 나타내는 클래스 다이어그램 .....	92
[그림 3-23] 공구정보를 표현하는 클래스 다이어그램 .....	93
[그림 3-24] 공구 교환시기를 표현하는 클래스 다이어그램 .....	94
[그림 3-25] MES데이터 수집, SPC분석 및 시각화 .....	95
[그림 3-26] AI 솔루션 추론 분석, 시각화 및 절차 .....	97
[그림 3-27] 제조데이터 활용 시스템 아키텍처 .....	103
[그림 4-1] 가공공정 개선 후 레이아웃 설계 및 재배치 구현 .....	105
[그림 4-2] 신규 MCT 도입 및 로봇, 이송장치 설치 현황 .....	105
[그림 4-3] HMI 주요화면 .....	107
[그림 4-4] Con-Rod 가공공정 개선 후 설계 레이아웃 .....	107
[그림 4-5] 부시 조립체 모습 .....	108
[그림 4-6] 공구 모니터링 및 데이터 추출 시스템 .....	110
[그림 4-7] 공구 모니터링 및 검사시스템 레이아웃 .....	110
[그림 4-8] 공구 모니터링 및 검사시스템 설치 및 가동 현황 .....	111
[그림 4-9] OPC-UA 기반 통합 모듈 및 서버/클라이언트 .....	112
[그림 4-10] 데이터 수집 및 분석 시스템 네트워크 .....	113
[그림 4-11] MES 스마트 클라이언트 애플리케이션 브라우저 .....	114
[그림 4-12] 데이터쿼리 소스코드 일부 .....	115
[그림 4-13] 데이터 리포팅 소스코드 일부 .....	116
[그림 4-14] SPC 분석 소스코드 일부 .....	117
[그림 4-15] MES 내 커넥팅로드 측정검사 DB 및 추출 솔루션 .....	118
[그림 4-16] MES 내 공구 부하량 측정 DB 및 추출 솔루션 .....	119
[그림 4-17] MES 내 공구사용 관리 DB 및 솔루션 .....	119
[그림 4-18] MES DB의 추출 솔루션에서 SPC통계 모듈로 데이터 추출 .....	120

[그림 4-19] SPC통계 모듈프로그램의 각종 통계 구현 이미지 .....	121
[그림 4-20] 분석을 위한 데이터 셋 가져오기 .....	125
[그림 4-21] 이상치 확인 .....	125
[그림 4-22] 기종별(일자별) 측정항목별 평균, 최대, 최소 등 요약 정보 ..	126
[그림 4-23] 양품/불량 수 차트로 나타내기 .....	126
[그림 4-24] 변수별 산포도 .....	129
[그림 4-25] 변수별 히트맵 .....	129
[그림 4-26] 로지스틱 회귀 모델을 통한 1차 학습 및 정확도 결과 .....	130
[그림 4-27] SVM에 의한 학습과 OK, NG 확률에 의한 1차 예측결과 ..	130
[그림 4-28] 1차 랜덤포레스트 학습 .....	131
[그림 4-29] 표본 개수 증가 후 랜덤포레스트 학습 .....	131
[그림 4-30] 트리 깊이 증가 후 산출된 랜덤포레스트 1차 예측 정확도 ..	132
[그림 4-31] 로지스틱 회귀분석의 2차 예측 정확도 .....	133
[그림 4-32] SVM에 의한 학습과 OK와 NG 확률에 의한 2차 예측결과	134
[그림 4-33] 2차 랜덤포레스트 학습 .....	135
[그림 4-34] 표본 개수 증가 후 랜덤포레스트 학습 .....	135
[그림 4-35] 트리 깊이 증가 후 산출된 랜덤포레스트 2차 예측 정확도 ..	136
[그림 4-36] XGBoost 작업을 위한 관련 모듈 설치 작업 내용 .....	136
[그림 4-37] XGBoost 작업을 위한 변수 생성 작업 및 교차 검증 처리 ..	137
[그림 4-38] 학습 변수들에 대한 초기값 설정 .....	137
[그림 4-39] XGBoost 훈련 내용 .....	138
[그림 4-40] XGBoost의 정확도 .....	138
[그림 4-41] XGBoost를 통한 각 피처의 중요도 .....	139
[그림 4-42] 관리도 주요 경향 .....	142
[그림 4-43] 볼트홀 1차, 2차 시범생산 X바-R관리도 및 공정능력 .....	144
[그림 4-44] 볼트홀 3차, 4차 시범생산 X바-R관리도 및 공정능력 .....	144
[그림 4-45] 커넥팅 로드 볼트홀 P 관리도 .....	145
[그림 4-46] 우측 편홀 1차, 2차 시범생산 X바-R관리도 및 공정능력 .....	147
[그림 4-47] 우측 편홀 3차, 4차 시범생산 X바-R관리도 및 공정능력 .....	147

[그림 4-48] 커넥팅로드 우측 핀홀 P 관리도 .....	148
[그림 4-49] SPC통계분석 종합 그래프 .....	152
[그림 4-50] 성과평가의 기준 .....	155
[그림 4-51] 알고리즘 예측 정확도 수치 그래프 .....	156
[그림 4-52] XGBoost를 통한 각 피처의 중요도 .....	158
[그림 4-53] F1 스코어와 SPC분석 비교표 .....	159

## 부 록 목 차

[부록 1] 볼트홀 대칭도 1차, 2차 Cpk 및 X바-R관리도 상세 분석표 .....	174
[부록 2] 볼트홀 대칭도 3차, 4차 Cpk 및 X바-R관리도 상세 분석표 .....	175
[부록 3] 볼트홀 대칭도 P관리도 상세분석표 .....	176
[부록 4] 우측핀홀 y2 1차, 2차 Cpk 및 X바-R관리도 상세 분석표 .....	177
[부록 5] 우측핀홀 y2, 3차, 4차 Cpk 및 X바-R관리도 상세 분석표 .....	178
[부록 6] 우측핀홀 y2 P관리도 상세분석표 .....	179
[부록 7] 커넥팅로드 특정 날짜, 기간 조회 .....	179
[부록 8] 공구수명 분석을 위한 데이터 셋 불러오기 .....	180
[부록 9] 공구 23번에 대한 조회 .....	180
[부록 10] 공구수명 분석을 위한 데이터 셋 불러오기 .....	180
[부록 11] 공구 23번에 대한 날짜별 부하(max_tool_load) 조회 .....	181
[부록 12] features, label 데이터 준비 .....	181
[부록 13] max_tool_load 가공조건별 교환주기 확인 .....	181
[부록 14] max_x_torque 가공조건별 교환주기 확인 .....	182
[부록 15] max_y_torque 가공조건별 교환주기 .....	182
[부록 16] max_z_torque 가공조건별 교환주기 .....	182
[부록 17] 각 특징변수의 상관계수 확인 .....	183
[부록 18] 각 특징변수의 히트맵 .....	183

# 제 1 장 서 론

## 제 1 절 연구의 배경

스마트공장은 제품의 기획, 설계, 생산, 유통·판매 등 전 과정이 사물인터넷(IoT), 사이버 - 물리 시스템(CPS, Cyber Physical System), 임베디드 운영체제(ex. IoS) 등의 ICT와 융합하여 자동화 및 정보화되어 가치사슬 전체가 실시간 연동·통합됨으로써 생산성 향상, 에너지 절감, 인간 중심의 작업 환경을 구현하고, 최적비용 및 시간으로 고객 맞춤형 제품을 생산하는 공장을 말한다. 또한, 스마트공장은 현재의 소품종 대량생산, 생산자 주도에서 미래에는 다품종 유연생산, 소비자 주도로 생산방식의 전환이며, ICT를 활용하여 기존 제조업의 전 과정을 디지털화하고, 미래 첨단 산업으로 디지털 전환(Digital Transformation)함으로써 국가 산업구조를 혁신하기 위한 제반 활동을 의미하기도 한다.<sup>1)</sup>

해외 스마트공장 진행 현황을 살펴보면 독일, 미국, 일본 등 제조 기술 선진국뿐만 아니라, 중국 등 후발 주자 또한 제조업의 혁신을 위한 다양한 정책을 추진 중이다. 국가별 제조업 혁신 정책은 다양하나, 공통으로 첨단 제조 기술의 개발 및 스마트 제조가 정책에 포함되어 있다. 독일은 자국의 제조 기술력에 ICT(Information and Communications Technology)를 결합하여 제조 효율성 향상, 사업 모델 확장, 제품 경쟁력 강화 등을 도모하기 위한 Industry 4.0 정책을 추진하며 스마트 제조화를 선도하고 있다. 미국은 서비스업 중심의 경제 구조를 유지해오고 있다가, 경제 침체에 빠지면서 첨단 제조업의 육성을 통한 경제 회복을 위해 다양한 혁신 전략을 추진하고 있다. 미국의 첨단 제조 파트너십(Advanced Manufacturing Partnership, AMP) 정책은 첨단제조업의 혁신과 스마트 제조 시스템 구축을 위한 정책이다. 첨단 제조 파트너십은 국가 안보 및 산업에 관련된 제조 역량 강화, 첨단 소재 개발 및 상용화 기간 단축, 차세대 로

1) 정종필, 신광섭. (2020). 스마트팩토리 핵심기술 및 제조혁신 고도화 전략. 융합연구리뷰. Vol. 6, No. 12, pp. 4.

봇 개발, 혁신적이고 에너지 효율적인 제조 공정 개발 등을 주 내용으로 포함하고 있다. 일본은 장기 경제 정체 문제를 4차 산업혁명 기술의 적극적 활용을 통해 해결하려고 한다. 일본진흥 전략은 생산 인구 감소 상황에서 지속적인 경제 성장을 위한 산업 구조의 전환 및 혁신적 생산을 위한 정책으로, IoT, 빅데이터, 인공지능, 로봇의 활용을 중심 내용이다. 중국은 정부 차원의 중장기적 전략 수립 및 대규모 투자를 통해 빠른 속도로 스마트 제조 산업을 집중적으로 육성 중이다. 2015년에 발표된 중국제조 2025는 세계 제조 산업을 선도하기 위한 혁신정책으로, 성장 잠재력이 높은 10대 산업을 전략적으로 선정하고 집중적으로 투자를 추진하고 있다. 그중 주요 제조공정의 스마트화, 중점 산업 분야의 스마트 공장 및 디지털 작업 현장 구축이 중점 프로젝트로 꼽히고 있다.<sup>2)</sup>



[그림 1-1] 스마트공장 구축 수준별 정의

출처 : 중소벤처기업부

스마트공장 관련 국내 현황을 살펴보면, 중소 및 중견 제조기업에 스마트화를 확산하고자 스마트공장 보급·확산 사업이 진행되고 있다. 이는 제조업의 글로벌 경쟁력을 강화하고 변화하는 업계의 패러다임에 대응하기 위함으로, 시범 사업을 2014년 시작해 2017년까지 5,003개사, 2022년까지 30,000개사에 스마트팩토리 구축을 지원하는 것을 목표로 하고 있다. 그림 1-1과 같이 스마트공장을 구축한 국내 중소·중견기업은 2020년을 기준으로 스마트화 수준 ‘기초’의 비율이 76%로 전반적으로 기초수준에 머무르고 있다. 스마트공장 R&D는 핵심기술 분야를 애플리케이션, 센서 및 장비, 보안 인공지능, 보안 및 상호운용

2) 정해운. (2021) 스마트팩토리 솔루션. 혁신성장 품목보고서, IR협의회. pp. 3.

성, 플랫폼 등으로 구성하여 진행하여 왔다. 최근 5~6년간의 스마트공장 보급 확산사업에서는, 대다수 기초수준인 중소·중견기업 스마트팩토리의 수준을 제고하고 4차산업 핵심기술인 AI, 빅데이터 등을 활용한 스마트공장 고도화로 방향을 전환하고 있다. 대표적 예로서 AI 중소벤처 제조 플랫폼(KAMP, Korea AI Manufacturing Platform)은 중소기업 공장의 스마트화를 지원하는 제조 데이터 체계를 구축한다. 중소기업에서 생성되는 제조데이터를 저장, 분석, 및 활용해 중소기업들이 스마트공장 관련 데이터를 지속해서 분석하고, 알고리즘 표준화 및 공정 최적화 등을 통해 AI 솔루션을 구축할 수 있도록 돕는다. 인공지능 제조 플랫폼 구축은 3단계로 구분하여 추진된다.

1단계는 클라우드 인프라 구축 등 제조 데이터 활용 기반을 마련하여 현재 정부, 지역 중심의 추진체계를 민관 협력 추진 체계로 구축하도록 개선하는 것이고, 2단계는 솔루션 스토어를 통한 AI 팩토리를 확산하여 AI의 표준모델을 확대해 제조기업이 다양한 AI 솔루션을 활용할 수 있도록 솔루션 스토어를 구축하는 것이다. 그리고 3단계는 마이제조데이터 활성화 기반을 마련하여 수요와 공급을 매칭하는 공유경제를 구축하는 것이다. KGMP를 중심으로 개별 기업의 제조 문제와 해결방안을 찾아내고, 표준모델을 기반으로 AI 솔루션을 개발할 수 있도록 AI 전문가 컨설팅 및 실증을 지원하는 '5G+AI 스마트공장'을 1,000개 보급을 목표로 제조혁신 선도 사례를 확산시키는 방향으로 진행하고 있다. 마이제조데이터는 제조 데이터 활용 이익을 데이터 생산 제조 기업에 환원하는 서비스 모델로 호환 가능한 양질의 제조 데이터를 생성하고, 원활하게 공유할 수 있도록 제조 데이터 표준 및 공유 규범을 마련한다.<sup>3)</sup>

전 세계적으로 변화하고 있는 제조업 업계의 패러다임과 불확실성에 유연하게 대응할 수 있는 스마트공장 고도화 기반 기술 개발은 표 1-1과 같이 크게 혁신형장비, 융합플랫폼, ICT활용 제조응용, 차세대 네트워크 및 상호 보안 운용성의 네 가지로 구분할 수 있다.

---

3) 정종필. 전개논문, pp. 10-21.

[표 1-1] 스마트공장의 핵심기술 분류 및 정의

기술분류	핵심 기술	기술정의
혁신형 장비	제조공정 환경 모니터링 지능형 센서	제조업에서의 다양한 정보를 감지하고, 필요시 데이터 처리, 자동보정, 자가보정을 수행하는 스마트 센서(범용 스마트 센서 이외의 산업별 전용 센싱 디바이스의 개발도 필요)
	3D프린팅 제조융합 기술	다품종 소량생산 및 유연생산을 가능하게 하는 3D 프린터 및 후가공 장비들을 제조 공정과 융합한 기술
	고난도 제조공정용 지능 로봇	자동화가 어려운 공정에 대응하고, 신속한 생산라인의 재구성과 변동성에 대응이 가능하며, 인간 - 로봇 협력 생산을 통해 공정 유연성을 극대화할 수 있는 지능형 로봇
	산업용 디바이스 정보 모델	이산 / 프로세스 / 휴먼머신 등 다양한 팩토리 디바이스를 객체화 하는 핵심 기술인 산업용 디바이스 정보 모델 기술
차세대 네트워크	5G 네트워크 및 지능형 네트워크	저지연, 고신뢰 요구사항을 만족하는 시간 민감형 네트워크(TSN, Time Sensitive Network) 기반의 차세대 유선 통신과 5G 및 B5G 이동통신, 비면허대역 무선통신을 기반으로 하는 저지연, 고신뢰를 만족하는 무선 통신 기술이 융합된 차세대 산업용 네트워크 기술
융합 플랫폼	자율 지능 IoT 단말 플랫폼	팩토리 내의 다양한 개체(장비 / 설비 등)에 지능을 부여하여, 각 개체의 자율적 기능 수행을 위한 'IoT 지능화' 기술
	최적 자율 제조 플랫폼	팩토리 내 지능화된 개체들이 빅데이터 분석 및 인공지능 기술 등을 통해 최적의 생산과정을 진화시켜가는 제조 공통 플랫폼 기술
	사이버 - 물리 시스템(CPS)	스마트팩토리의 물리적 자원들을 가상공간에 모사하고 실시간 연동 및 제어가 가능하며, 가상공간에서 생산 시뮬레이션을 지원하기 위한 IoT / CPS 기술
	클라우드 제조 서비스 플랫폼	제품의 가치를 높이기 위해 제품과 관련된 서비스를 연계하는 제품 서비스 및 제품생산을 서비스 형태로 제공하기 위한 클라우드 기반 제조 서비스 플랫폼 기술
ICT 활용 제조 응용	데이터기반 공정설계 시뮬레이션	해석기술의 고속화 및 취득 데이터 기반(생산현장 데이터의 자동입력) 해석기술을 활용한 웹기반 클라우드형 시뮬레이션 기술
	AI 예측 기반 제조 응용	설비 예지보전, 품질관리, 공정관리 등의 제조 산업 현장에 적용 가능한 데이터 분석 기반 예측 기술
	실시간 품질검사 응용	X선 / 영상분석 등을 활용한 생산제품의 실시간 불량판정 자동화 기술 개발 및 불량판정 결과와 생산조건간의 연계 기술
	AR / VR 기반 작업 · 안전 지원	가상, 증강 현실 기술을 바탕으로 업무환경을 지원하는 작업자 지원 기술과 작업자 이상, 유해상황 감지 및 사전 대응을 위한 안전 지원 기술
	맞춤형 제조서비스	기업이 주문하는 제품에 필요한 생산 자원을 동적으로 구성하는 맞춤형 제조 기술
	지능형 유통 및 조달물류	가치 사슬을 연계하여 물류 흐름을 추적 관리하고, 물류 운영을 최적화하기 위한 기술
보안 및 상호 운용성	지능 보안 및 통신 네트워크 보호	사이버위협을 포함한 다양한 보안 위협 상황에 대응하기 위한 정보 보안 기술
	스마트팩토리 표준화 기술	제조 주체 또는 시스템 간 상호연동성 보장을 통해 스마트팩토리에 개방성을 제공하고 팩토리 간의 상호운용성을 보장하기 위한 표준 기술

\*출처 : 한국산업기술평가관리원,2018

혁신형 장비는 스마트 센서와 지능형 로봇, 3D 프린터 등의 최신 기술 장비를 산업 현장에 활용하여 사물인터넷(IoT), 스마트 센서, 디바이스 기반으로 자율 유연 생산 체계를 지원하는 기술이다. 산업용 3D 프린터는 국내 스마트공장 요소 기술들의 개별 시장 중 가장 작은 비중을 차지하는 시장이나, 점차 비중이 확대될 전망이다. 3D 프린팅 기술은 21세기 제조업의 패러다임 중 하나인 다품종 소량생산 및 사용자 개별 맞춤형 생산에 활용도가 높고, 시제품 개발 단계에서 특히 유용하게 사용된다. 스마트공장은 제조 공정의 로봇화라는 개념을 수반한다. 자동화된 전용기와 로봇들이 함께 작동하는 것이 어려워, 다양한 제조 환경에 유연하게 활용 가능한 로봇을 제작해 이를 중심으로 한 중심의 자동화를 추구하고 있다.

융합 플랫폼은 제조 전체 과정을 사이버와 물리적인 생산 자원들을 수직적으로 융합하고 제조와 서비스 영역의 수평적인 융합을 통해, 요약하면 제조와 인공지능의 융합으로 플랫폼을 제공하는 기술이다. 사이버 물리 시스템은 핵심 개념이 인지, 판단, 행동을 능동적이고 자율적으로 수행하며, 사이버 시스템과 물리적 시스템으로 통칭하는 기존의 실시간 임베디드 시스템이 확장된 개념이라고 볼 수 있다. 모든 사물을 연결하는 IoT와 데이터를 수집하는 플랫폼, 의사 결정을 지원하는 빅데이터 분석 기술, 인공지능 기술 등 다양한 기술들이 필요하다. 클라우드 컴퓨팅은, 데이터를 데이터 센터 격인 클라우드로 보내고, 해당 클라우드에서 정보를 저장하고 분석하는 것을 의미한다. 또 제조업에서도 데이터의 폭증이 두드러지게 나타나 데이터 송수신 시간이 증가하고, 많은 양의 데이터를 처리할 경우 데이터 병목 현상으로 인해 처리 시간이 지연되는 문제가 있어 보완하기 위해 나온 기술이 엷지 컴퓨팅이다. 클라우드에서는 기존 데이터들을 이용한 모델링을 수행하며, 실시간으로 수집된 데이터들에 맞는 최적의 모델을 엷지로 전송하여, 엷지에서 관련 데이터들을 처리할 수 있도록 한다. 엷지에서는 클라우드로부터 전송된 모델을 활용해 공정 및 설비 등 제조 현장에서 발생하는 문제 해결을 위한 피드백을 적용하며 실시간으로 데이터를 분석한다.

ICT 활용 제조응용기술은 ICT를 활용해 플랫폼 기반으로 해 새로운 가치를 창출할 수 있도록 한 응용기술이다. 빅데이터와 AI를 이용하여 제조 솔루션, 설비 등의 여러 생산 요소가 스스로 현장의 상황을 인식하고 판단하는 능력을 고

도화하여 자율적응형 환경을 구축하는 것이다. 제조과정 내의 생산 요소들이 생성하는 데이터들을 수집 및 분석하여 공정의 자율적인 운영과 최적화된 운영을 추구한다.

차세대 네트워크 및 상호 보안 운용성이다. 인터넷에 액추에이터, 센서, 제어기 등의 장비들을 연결할 수 있도록 실시간, 고신뢰, 저지연 등의 기술이 만족하는 것이 차세대 유무선 네트워크 기술이다. 사이버 보안은 주요 기반 시설로부터 가깝지 않은 곳에 있는 시스템을 효과적으로 제어하기 위해 필수적으로 활용되는 기술이다. 사이버 보안 기술은 네트워크, 데이터, 신원 및 접근, 엔드포인트, 그리고 클라우드 보안 기술들로 구성된다. 제조 현장에서 상호 연결된 기기들의 수가 급증하였으며 IoT를 통해 설계부터 생산, 유통, 서비스까지 각 프로세스가 가상공간상에 통합됨에 따라 정보 및 기술 유출의 위험성이 더욱 커지고 있다.

## 제 2 절 연구의 필요성

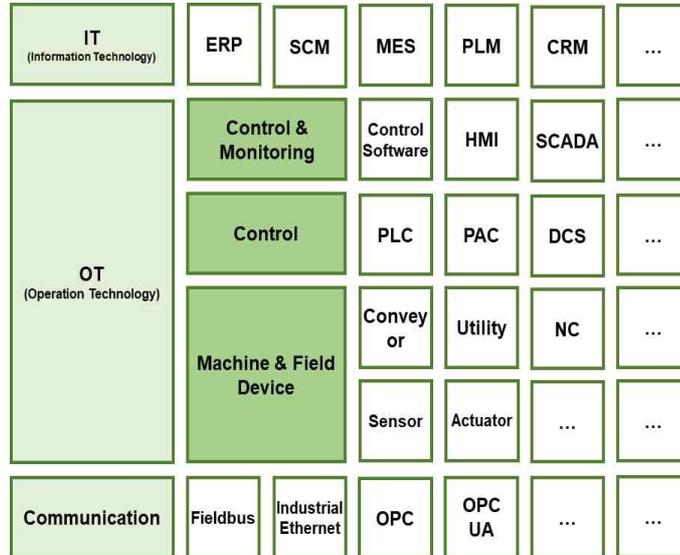
중소기업은 P, Q, C, D, F(생산성, 품질, 원가, 납기, 유연성)을 제고하기 위해 4차 산업에서 핵심기술이 더욱 중요하다고 볼 수 있다. 그러나 70% 이상의 중소기업이 기초수준에서 스마트팩토리의 수준을 높이기 위해서 우선 파급효과가 높을 것으로 예상되는 빅데이터수집 및 분석 기술, IoT와 스마트센서를 활용한 통합 기술, AI학습 모델 등을 활용한 스마트공장 고도화 기술을 연구할 필요가 있다

스마트공장의 일반적인 시스템 구조는 그림 1-2와 같이<sup>4)</sup> IT영역과 OT영역 그리고 통신 부분으로 구성할 수 있다. Information Technology(IT)영역은 ERP, SCM, MES, PLM, CRM 등으로 구성되고 Operation Technology(OT)영역은 Machine & Field Device(Sensor, Actuator, 설비 등) Control 장치(PLC, PAC, DSC 등), Control & Monitoring(HMI, SCADA 등)으로 구성한다. 통신영역은 Fieldbus, Ethernet, OPC, OPC-UA로 구성할 수 있다.<sup>5)</sup> 최근 스마트공장 추세는 IoT, CPS, 임베디드 등 최신 ICT기술이 급속도로 융합이 되고 있고 중소기업도 실시간 Big data 수집과 분석시스템이 확산하고 있다. 그 중 중요한 기술은 빅데이터 수집 및 분석 기술, IoT포함한 유무선 통합 통신프로토콜 기술, AI학습 모델 기술이 중요하다고 볼 수 있다. 빅데이터 플랫폼 데이터 생성, 수집, 저장, 처리, 분석 및 평가 단계 과정으로 일반적으로 진행되고 있다. 이들 단계과정이 효과적으로 진행되기 위해서 이에 부합한 공장 자동화 및 스마트화 수준이 충족되고, 특히 현장 설비 간의 최신 Network와 통신기술이 필요하다.

---

4) 김재표, 김승천. (2020). 중소기업 스마트공장 사전준비를 위한 이기종 시스템에 대한 네트워크 인터페이스 방안의 연구. 한국융합학회. Vol. 11, No 11, pp. 56.

5) 정지석. (2018). 4차 산업혁명 시대 ICT 기반 건설링 방법론. 한국건설링서비스협회. 3-24.



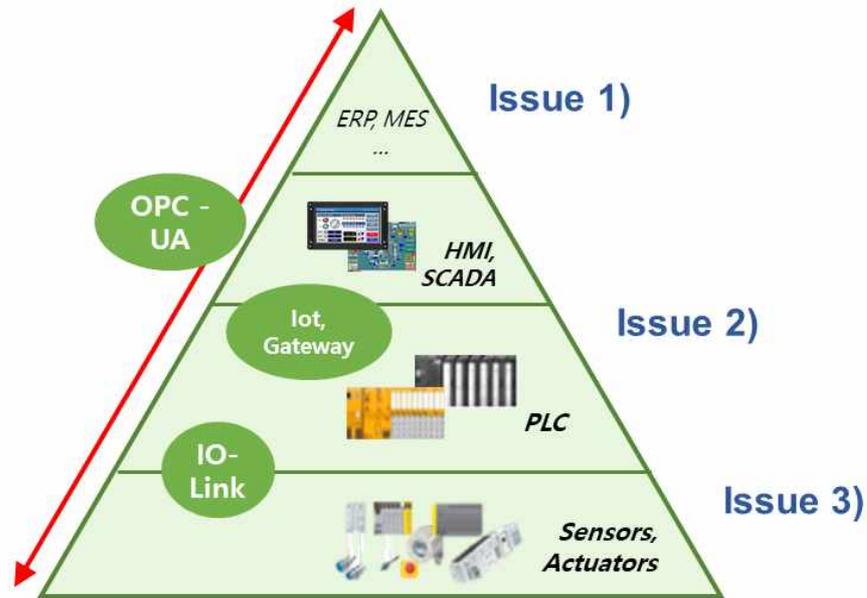
[그림 1-2] 스마트공장 아키텍처

생산제조 데이터 정보의 특징은 1) 제조실행 시 발생하는 RFID, 바코드, 온 습도, 압력 등 물리 화학적 정보의 다양한 종류와 아날로그 성격의 데이터, 2) 대량생산시스템에서 데이터 처리속도와 크기가 큰 유형, 3) 외부 네트워크와 단 절된 HMI/SCADA시스템, 4) 다양한 설비의 환경적 차이로 발생하는 데이터 포맷의 다양성, 5) 불필요한 연속적인 중복 데이터 및 반정형 데이터 활용 증가 등을 성격을 띠고 있다. 이런 정보의 특징으로 인해 실시간 운전 데이터의 통합 기술 부족, 빅데이터의 느린 속도 등의 문제가 발생할 가능성이 있다.<sup>6)</sup> 노후한 이기종 중심으로 네트워크의 실시간 데이터 처리와 통신 관점에서의 중요한 기술적 특성을 살펴보면 1) 송수신기 거리의 경로손실, 2) 송수신기의 반양방성 (half-duplex), 3) 물리계층의 오버헤드, 4) 채널 에러 발생 등이 있다. 이런 특 성으로 인해, 통신 패킷 전달과정에서 미승인 ID에 의한 공간적, 시간적 일관성 문제, 분배 토큰패싱 방식의 이동으로 의한 프로토콜 문제, 패킷의 충돌을 회피 할 경우 무선의 반양방성을 사용하는 등의 제약 등이 있다.<sup>7)8)</sup>

6) 이성준, 김춘경, & 이재덕. (2017). OPC UA를 이용한 스마트센서 시스템 개발. 대한전기학회 학술 대회 논문집, 1338-1340.

7) 문상국. (2006.). 산업기기 네트워크에서의 실시간 처리와 필드버스 통신의 문제점 분석. In Proceedings of KIIT Conference (pp. 303-306).

8) 문상국. (2007). 무선 통신 기술의 산업기기 응용에 대한 고찰. 한국정보통신학회논문지, 11(1),



[그림 1-3] 제조 데이터의 통신 이슈

중소기업에 예상되는 통신 관련 이슈는 그림 1-3과 같이<sup>9)</sup> 크게 1) IT 시스템이 구축된 경우 연계가 쉽지 않고 2) 이기종 시스템에서 생성되는 데이터 수집 및 통합이 어려우며 3) 다양한 필드버스 및 프로토콜이 존재하여 인터페이스가 쉽지 않다는 것이다.<sup>10)</sup> 중소기업은 이런 특성으로 OT, IT 등 각 영역 간의 데이터의 인터페이스가 되지 않고 실시간 처리도 어렵다. 대다수 중소기업이 통신과 OT영역의 문제점에 대한 사전준비 없이 IT 위주로 스마트공장을 도입해왔으나, 최근 스마트공장 고도화 사업을 중심으로 OPC-UA의 최신 통신프로토콜의 도입이 점차 늘어나고 있다. 이기종 설비와 OT, IT영역 간의 최신 통신 프로토콜과 네트워크 기반으로 실시간 데이터수집 및 분석시스템을 연구할 필요가 있다.

9) 김재표. 전계논문, pp. 57.

10) 권대욱. (2018). Smart Factory 구현을 위한 자동화 시스템의 Engineering 설계. 스마트팩토리 콘퍼런스 2018.

## 제 3 절 연구의 범위와 목적

### 1) 연구의 방법과 계획

#### 가) 연구 방법

제조데이터 수집분석 및 활용 시스템은 제조공정 스마트화의 단계, 빅데이터 수집, 처리 단계, SPC(통계적공정관리) 및 AI학습 분석 활용단계의 총 4단계로 구성되며, 상세 연구 방법은 다음과 같다.

제조공정 스마트화 단계에서는 생산성, 작업시간, 불량률 등에 대한 공정개선을 위해 도출한 파라미터로 자동화 설비와 지그 장치에 대한 사양 검토와 상세 설계를 진행한다. 가공불량의 측정방법을 검토한 다음, 적합한 스마트센서를 선정하고 측정, 검사시스템의 사양과 설계를 진행한다.

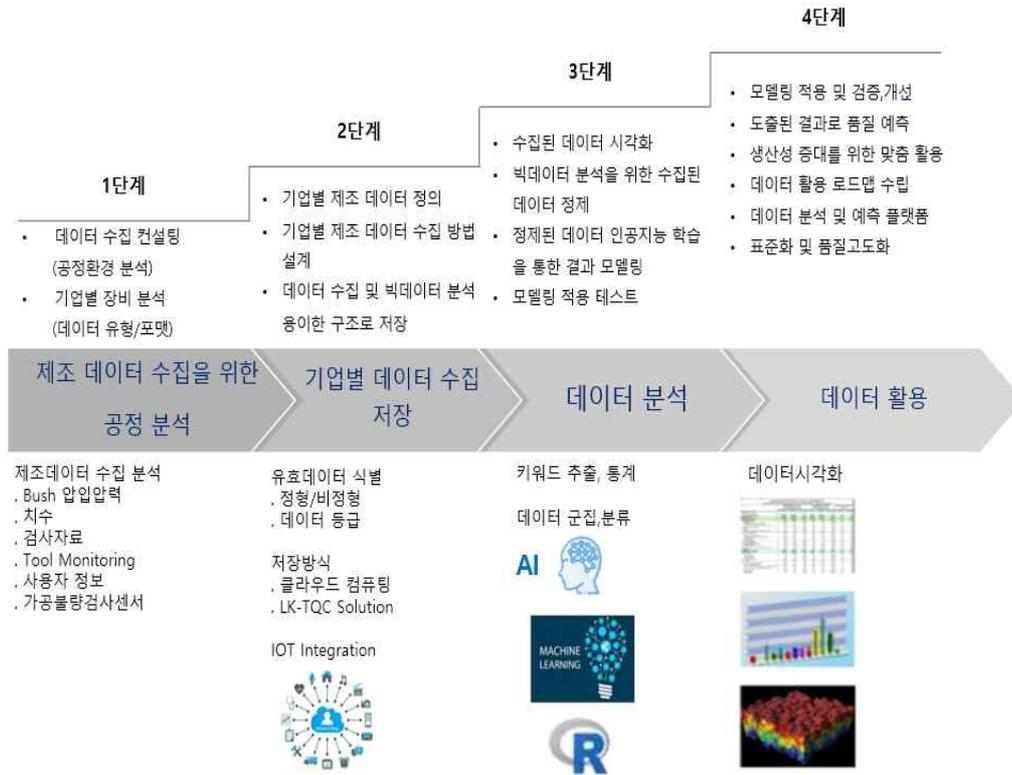
빅데이터 수집, 처리 단계에서는 스마트화 공정 설비와 측정검사 시스템에서 발생하는 데이터의 유형을 분석하고 스마트 센서, IoT 통합모듈 등으로 OPC-UA 기반의 네트워크를 통해서 추출 데이터를 수집하고 저장한다. 빅데이터 수집을 위한 SW를 개발하고 외부 클라우드 서비스와 연결하여 수집, 저장, 처리 시스템을 설계하고 구현한다.

SPC 및 AI 학습 분석 활용단계에서는 제조활용을 위한 빅데이터 분석과 활용방안을 수립한다. 볼트홀 대칭도, 살두께 등의 불량, 공구 교환시기 등과 같은 주요 관리 변수들의 품질 고도화와 공구수명 최적화를 위해, AI 모델링을 통해 품질과 공구교환시기에 대한 예측을 수행한다. AI 모델의 성능검증과 SPC의 통계분석으로 상호 비교 검증한 후 최적화된 방향으로 연구를 진행한다.

#### 나) 연구 계획

제조데이터 수집 및 활용시스템의 연구 계획은 그림 1-4와 같이 4단계로 구분하여 수립하였으며, 1단계는 제조데이터 수집을 위한 공정 분석 단계, 2단계는 공정분석을 통한 기업별 제조 데이터를 정의하고 수집하는 단계, 3단계는 수집된 데이터를 이용하여 기업에서 활용할 수 있도록 AI를 학습시키고 모델링하는 분석 단계이며, 마지막 4단계는 분석된 데이터를 기반으로 기업의 생산성을 높이기 위한 활용 방안을 도출하는 단계이다. 4단계의

활용방안 단계에서는 SPC 시각화와 AI 학습 모델 분석까지 진행한다.



[그림 1-4] 단계별 제조데이터 활용모델 구축 계획 및 활용 도구

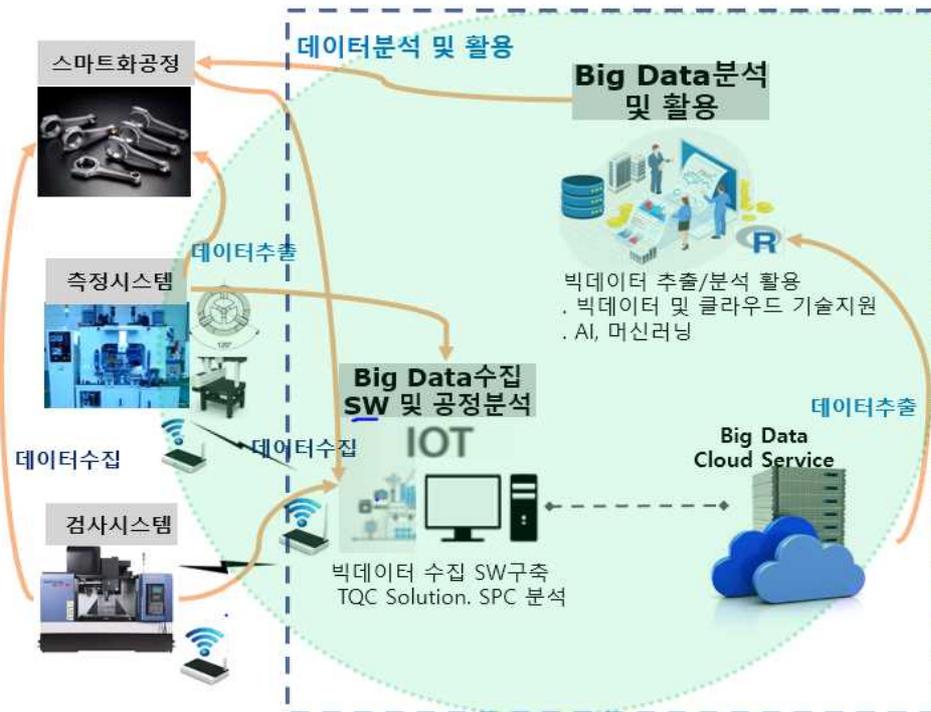
## 2) 연구의 범위

연구 범위는 다음 그림 1-5에서 확인할 수 있는, 제조데이터 수집분석 및 활용시스템에 대한 아키텍처이다. 첫째, 수집 가능한 제조·품질 데이터를 정의하고 연계공정을 분석한다. 생산현장의 가공·품질 데이터를 취득, 전송, 공동보관 및 활용하거나 통합분석 및 유사 데이터 비교 등 활용방안을 도출하고 적용한다.

둘째, 데이터를 수집, 저장·분류하면서 공동으로 분석하고 활용한다. 데이터 측정시스템의 데이터 수집 방안을 도출하고 저장·분류함으로써 데이터 간의 연계 포인트를 공동으로 분석 및 활용한다.

셋째, 빅데이터 분석 결과를 기반으로 품질 및 생산성 향상을 위한 대책안을 제시한다. IoT 통합 OPC-UA 기반의 제조공정 데이터 수집 SW 및 처리 시스템 개발, 클라우드 컴퓨팅 기술지원, SPC & AI 학습모델링 분석 및 시각화 등

을 활용해서 진행한다.



[그림 1-5] 제조데이터 수집분석 및 활용 시스템 아키텍처

넷째, AI모델링 분석으로 품질, 공구 수명에 대한 예측 활동이 가능하다. 그리고 MES SPC통계 솔루션으로 접근이 쉬운 실시간 통계분석이 가능해 기존 통계도구를 통한 분석보다 공정모니터링 및 사후 유지관리의 수준을 높일 수 있다. 실시간 SPC 통계분석과 AI 분석을 같이 진행하면, SPC 분석으로 AI모델링에 대한 사후 검증이 가능하다. 또한 AI 모델링 분석은 사전 예측 기능, SPC 분석은 사후 통계분석과 일상적 유지관리로서 역할을 분담할 수 있다. 본 연구에서도 SPC 통계분석과 AI모델링 분석 간의 비교검증을 하려고 한다. 이러한 과정을 통해 빠르고 상세하면서 정확한 분석 및 검증이 가능하고, 이는 가공공정의 고질적인 불량률의 감소로 이어져 품질고도화가 가능하며, 공구마모, 치수불량, 가공 정밀도 등 데이터 수집활동의 장기적인 축적을 통해 품질 안정화와 표준화를 달성할 수 있다.

이 연구는 기계가공 공정의 생산성 향상과 높은 가공정밀도를 위해, 커넥티드 로드 절삭가공 제조라인의 스마트공장 고도화를 설계하고 구현하려고 한다.

### 3) 연구의 목적

제조데이터 수집 분석 및 활용을 위한 전체 시스템 아키텍처로서, 그림 1-5와 같이 품질고도화 및 생산성 향상을 위한 통합 시스템 내 공정·품질관리 방안을 수립하고 적용하고자 한다.

제품의 가공조건에 영향을 미치는 모든 요소에 대한 데이터를 실시간 수집하는 시스템과 높은 가공정밀도의 공정라인을 통해 가공공정의 모니터링 체계를 도입하여, 표 1-2와 같이 IoT 통합, 빅데이터 소프트웨어, AI, 시각화 등의 IT 요소기술을 적용한 실시간 제조 빅데이터 수집, 처리 및 분석 시스템을 도입할 필요가 있다. 이를 위해 그림 1-5와 같은 연구수행 아키텍처와 표 1-2와 같은 IT 요소기술을 적용하고자 한다. OPC-UA를 기반으로 한 IoT 통합 기술을 적용한 공정 모니터링 체계와 실시간 제조데이터 수집, 처리, 분석 기술을 연구한 전체 시스템을 함께 설계와 구현하는 것을 연구 목표로 삼고자 한다.

[표 1-2] 연구의 목표

연구 목표		IT 요소 기술
Data 정의 및 수집, 저장 분류 연구	스마트화 공정	IIoT 디바이스 통합 모듈 데이터 수집대상 분석 및 분류 연구
	측정, 검사시스템	
Data 저장, 처리, 분석 기술 연구 및 개발	유무선 IoT 통합 프로토콜	OPC-UA표준 프로토콜 IoT통합 디바이스, 네트워크
	빅데이터 처리 SW	3계층 어플리케이션 브라우저, 오픈소스 데이터베이스
	TQC(SPC) 솔루션 및 시각화	통계분석 어플리케이션 오픈소스 시각화S/W
	AI학습 모델링 분석 및 시각화	MES TQC 어플리케이션 지도학습 알고리즘
	클라우드 컴퓨팅 기술지원	IaaS

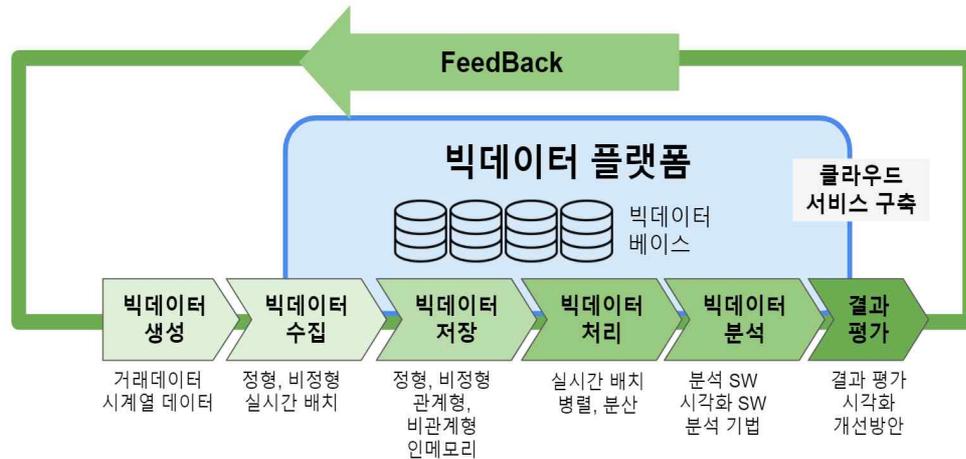
## 제 2 장 관련 기술 및 선행연구

### 제 1 절 빅데이터 수집 및 처리 연구

#### 1) 빅데이터 수집 및 처리 분석 관련 기술<sup>11)</sup>

##### 가) 빅데이터 수집 관련 DB 및 시각화 기술

일반적으로 빅데이터의 특성은 데이터의 양(Volume), 다양성(Variety), 속도(Velocity), 신뢰성(Veracity)이므로 스마트팩토리에서 빅데이터의 신뢰성 확보는 매우 중요하고, 이는 공장을 지능화하기 위한 머신러닝 및 인공지능의 기반이 된다. 스마트공장의 데이터의 경우, 현장 엔지니어의 경험을 잘 녹여내는 것이 중요하다. 설비, 공정을 비롯한 전반적인 현장의 특성들을 잘 반영한 질 좋은 데이터를 추출하기 위해서는 엔지니어의 노하우가 투영되어야 한다. 데이터 품질관리 인프라는 빅데이터를 효과적으로 활용하기 위한 전반적인 시스템을 말하는데, 품질 관리나 메타데이터 관리 등의 과정이 포함된다. 먼저 현장의 공정 및 설비 등이나 직원 및 고객 등이 생성하는 데이터를 모아 수집하고, 데이터의 구조에 따라 변환 등의 처리 과정을 거쳐 데이터베이스에 저장한다.



[그림 2-1] 빅데이터 활용 프로세스 및 관련 기술

11) 장원중, 조성인, 김수상, 김광용. (2018). 스마트팩토리에서 빅데이터 인프라 구축 방안에 관한 연구. Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology, Vol.8, No.10, pp.11-23.

그림 2-1과 같이<sup>12)</sup> 빅데이터를 구축하려면 빅데이터 생성, 수집, 저장, 처리, 분석, 평가 등의 일련의 과정을 거치고 클라우드 서비스와 연계된 데이터베이스 관련 기술도 필요하다. 저장된 데이터를 활용 목적에 맞게 처리·분석하는 다양한 빅데이터 기술들과 분석소프트웨어, 시각화소프트웨어, 운영 및 보안 기능을 갖춘 플랫폼 또한 필요하다. 빅데이터는 활용 목적에 적합한 데이터베이스에 저장하여야 한다. 다음 표 2-1은 데이터베이스 종류별 차이이다. 빅데이터 분석 소프트웨어로 플랫폼에 저장된 데이터를 분석하고, 해당 분석 결과는 시각화분석 도구로 시각화하거나 알고리즘을 활용해 고도화된 의사결정을 수행하는 자동화 분석 시스템을 구축한다.

[표 2-1] 데이터베이스의 종류별 차이점

데이터베이스 종류	일반적인 적용 분야	주요 제품
전통적인 관계형 데이터베이스 관리 시스템 (RDBMS)	온라인 트랜잭션 처리 (OLTP), 데이터 마트, 데이터 웨어하우스 자료 질의	Oracle, Microsoft SQL Server, IBM DB2, MySQL, ProgreSQL
분석용(Analytical) 데이터베이스 관리 시스템	대용량 데이터 분석용 자료 질의	ParStream, Vertifca, SybaseIQ
대용량 데이터 온라인 트랜잭션 처리(OLTP)	대용량 데이터 트랜잭션 처리	Aerospike, NuoDB, VoltDB
인메모리(In-memory) 데이터베이스 관리 시스템	비교적 저용량 데이터 온라인 트랜잭션 처리(OLTP) 및 자료 질의	HANA, Kognitio, Altibase
노에스큐엘(NoSQL)	문서/객체 저장 및 검색	MongoDB, DouchDB, MarkLogic
그래프 또는 RDF 데이터관리 시스템	그래프/관계 질의	Neo4j, SPARQLverse, Stardog
하둡(Hadoop) 데이터베이스 관리 시스템	주로 데이터 추출/변환/적재, 데이터 클린징 및 일부 분석	HBase, Impala, Splice Machine

다음의 표 2-2는<sup>13)</sup> 2018년 논문에 활용된 분석 소프트웨어의 리스트로, 사용된 정도를 기준으로 내림차순으로 나열되어 있다.

12) 장원중. 전계논문, pp. 19.

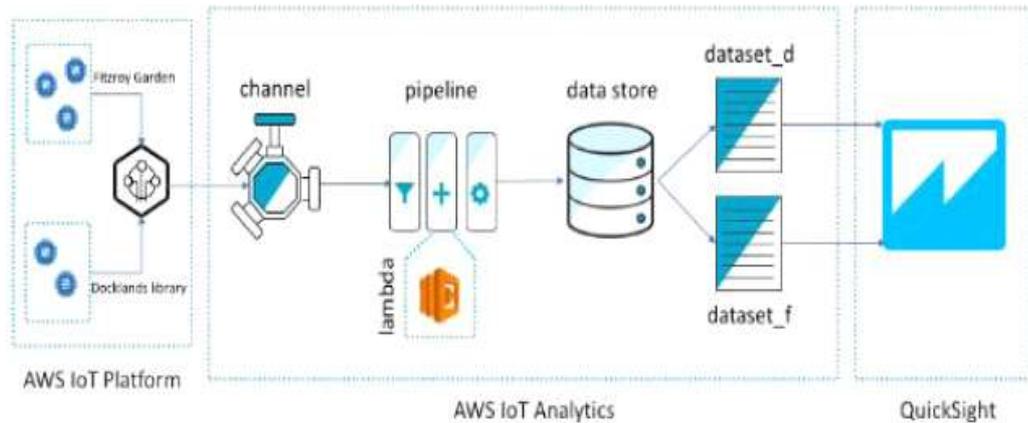
13) <http://r4stats.com/articles/popularity/>

[표 2-2] 2018년 학회지에 사용한 분석 소프트웨어 리스트

분석 소프트웨어 목록
SPSS Statistics, R, SAS, Stata, GraphPad Prism, MATLAB, Python, Apache Hadoop, Tensorflow, Java, Minitab, Statistica, SQL, Scikit Learn JMP, Systat, C, C++ or C#, Caffe, FORTRAN, Apache Spark, PyTorch, Statgraphics, Methemactics, IBM Watson, Weka, Theano

\* Google Scholar에서 1,700회 이상 인용된 소프트웨어에 대한 학술기사

최근 빅데이터 분석 및 시각화를 위한 플랫폼은 기계학습과 AI를 적용한 데이터 분석에까지 영역을 넓히고 있으며, 분석결과에 대한 정보전달 측면 뿐만 아니라 예측분석 능력까지도 연구 및 개발되고 있다. 그림 2-2는<sup>14)</sup> 데이터가 발생되어 수집 및 저장된 뒤 시각화되기까지의 흐름을 보여준다. 여기서 메타 데이터는 정형, 반정형, 비정형 데이터로 구분된다. 다양한 형식의 데이터를 SQL, NoSQL, Hadoop을 통해 분석 및 분류한 뒤 딥러닝 알고리즘을 접목해 의미 있는 분석결과를 도출하고 결과를 시각화와 도식화한다. 빅데이터 분석은 여러 출처로부터 추출, 처리 및 저장되어 있는 데이터를 활용해 학습하고 모델링하는 과정이기 때문에, 이전 단계들인 배치데이터 처리, 실시간 데이터 처리, NoSQL 데이터베이스, 분산 스토리지 등의 기술로부터 영향을 받는다.<sup>15)</sup>



[그림 2-2] 데이터 발생에서부터 시각화 플랫폼까지의 플로우

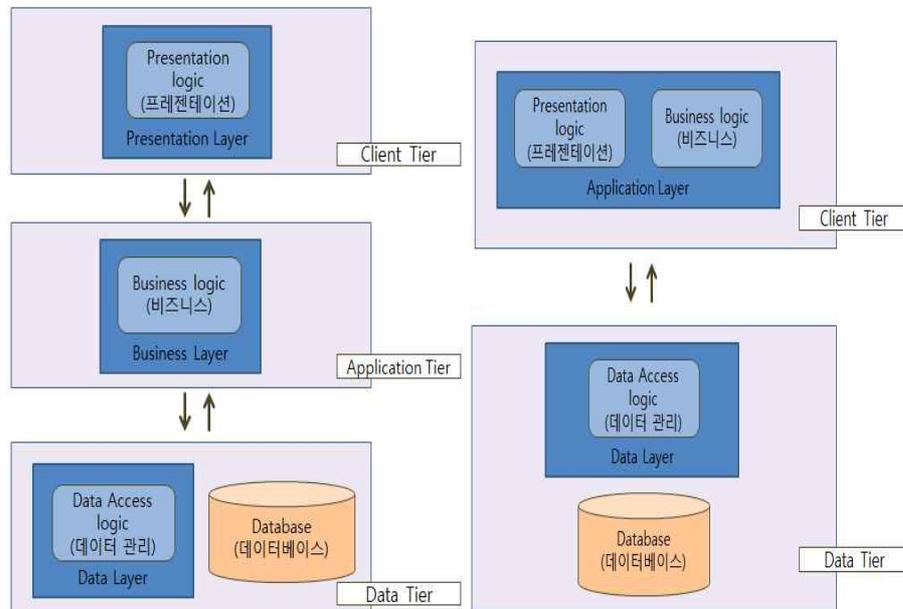
14) AWS(<http://aws.amazon.com>), 2019

15) 중소기업 전략기술로드맵 2021~2023. (2021) pp. 170~173

나) 3계층 구조 엔터프라이즈 솔루션

(1) 2, 3계층 구조 아키텍처 비교<sup>16)</sup>

3계층 구조는, 구조 아키텍처를 그림 2-3과 같이<sup>17)</sup> 프레젠테이션(클라이언트) 계층, 비즈니스(애플리케이션) 계층, 데이터 계층으로 나누며, 각 계층은 물리적으로 독립적이다. 프레젠테이션 계층은 사용자 인터페이스를 지원하고 인터넷 브라우저의 정적인 데이터를 제공한다. 이 계층은 프론트엔드라고도 하며 비즈니스 로직이나 데이터관리 코드를 포함하지 않는다. 비즈니스 계층은 정보 처리의 규칙을 포함하며 동적 데이터를 제공한다. 미들웨어 또는 백엔드라고도 하며 프레젠테이션 코드나 데이터관리 코드를 포함하지 않는다. 데이터 계층은 주로 데이터베이스를 의미한다. 데이터베이스 또는 파일 시스템을 접근 및 관리하며 주로 DB서버를 의미한다.



[그림 2-3] 3계층 및 2계층 구조 아키텍처

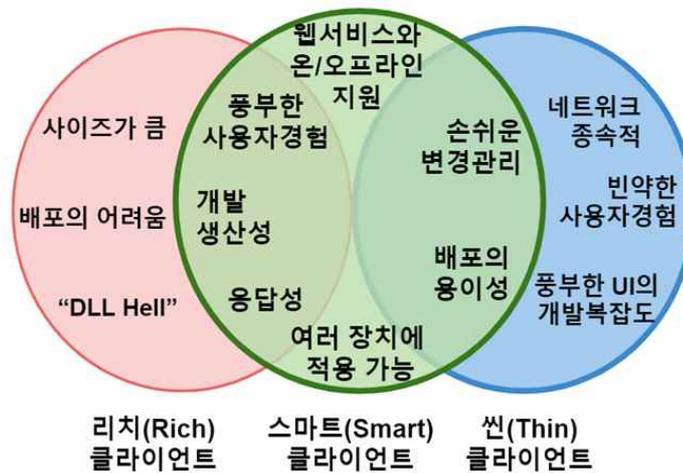
2계층 구조와의 장단점을 비교해보면, 3계층 구조의 경우 프레젠테이션 로직과 비즈니스/데이터 접근 로직을 별도로 작성하므로 2계층에 비해 개발이 복잡하다. 그러나 비즈니스 로직을 다양하게 구현할 수 있고, 로직을 모듈화해 클라이언트/서버와 웹 환경에 동시에 사용이 가능하다. 또한 동시 사용자 수가 증

16) <https://bkjeon1614.tistory.com/27>

17) 상계 웹사이트.

가해도 일정한 응답속도와 처리량을 보장하며 미들웨어에서 부하 분산, 큐잉 메커니즘을 통해 효율적으로 자원을 활용할 수 있다. 다양한 모니터링이 가능하여 관리 및 모니터링에 용이하다. 종합해보면 개발 편의성이 낮다는 단점이 있으나, 재사용성, 성능, 자원 활용, 시스템관리 등의 측면에서 회사 규모와 사용자 증가에서 3계층 구조가 우위에 있다고 볼 수 있다.

(2) 스마트 클라이언트 개념의 애플리케이션 브라우저<sup>18)</sup>



[그림 2-4] 스마트 클라이언트 특징

스마트 클라이언트는 그림 2-4와 같이<sup>19)</sup> 로컬 리소스를 활용하고 지능적으로 분산 데이터 원본에 연결함으로써 조정 가능하며 응답성이 뛰어난 다기능 대화형 환경을 제공하는 배포 및 관리가 쉬운 클라이언트 응용 프로그램이며, 리치 클라이언트와 씬 클라이언트의 이점을 결합하여 두 클라이언트 사이의 간격을 메운다고 할 수 있다.<sup>20)</sup>

18) 김영주, 최정민, 이병수. (2010) 엔터프라이즈 아키텍처를 위한 .Net 기반의 스마트 클라이언트 애플리케이션 브라우저 설계 및 구현. The Journal of Korean Institute of Information Technology. Vol. 8, No. 12, pp. 203-209.

19) 김영주. 상계논문, pp. 204.

20) Hill, D. (2004). The Architecture journal 4, Microsoft Corporation October 2004.



[그림 2-5] 애플리케이션 브라우저 구조

데스크톱 애플리케이션 브라우저는 그림 2-5와 같이 동등한 데스크톱 애플리케이션이 제공하는 기능과 유사한 기능을 제공한다. 애플리케이션 브라우저가 웹 브라우저와 구분되는 특징은 컴파일 시에 제공될 수 있는 기능들이 고정되지 않는다는 점이다. 아마존의 웹사이트를 예시로 들어보면, 브라우저는 현재 상품을 둘러보고 구매할 수 있는 애플리케이션을 실행하는데, 이를 가능케 하는 또 다른 무언가를 컴퓨터에 설치할 필요는 없다. 애플리케이션 브라우저도 이와 비슷하다.

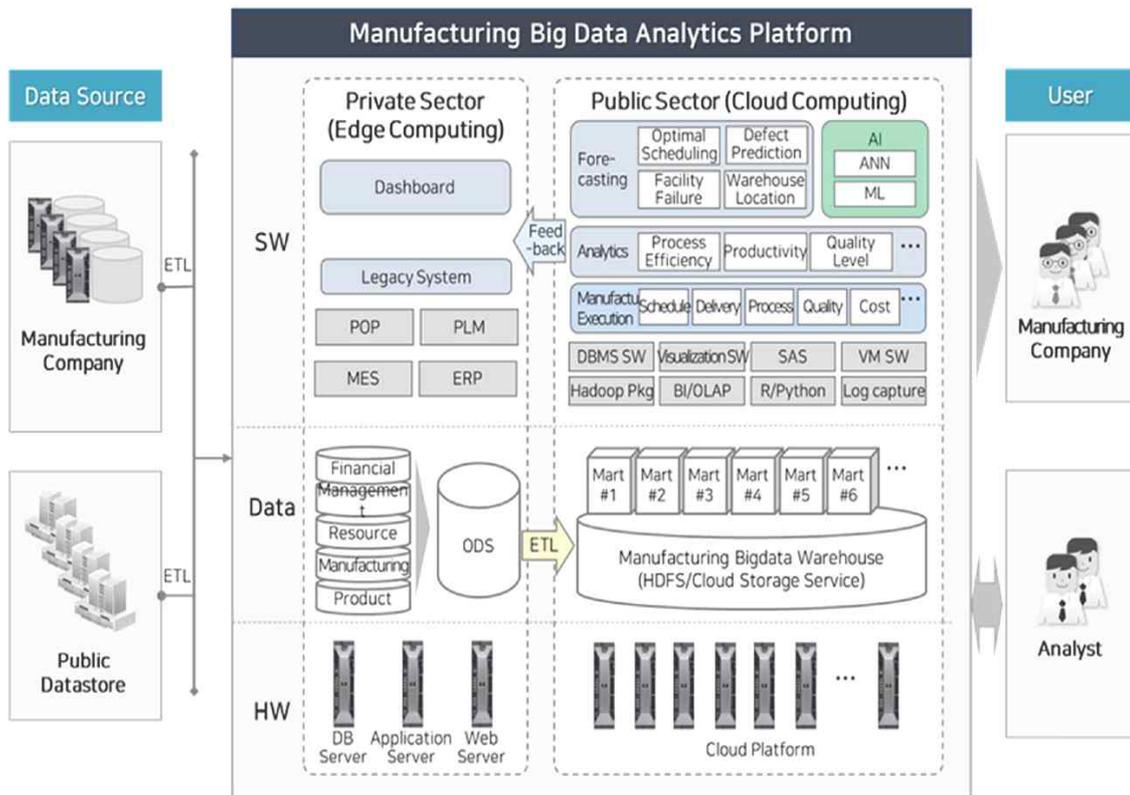
## 2) 빅데이터 수집 및 처리 선행연구 및 결과

### 가) 중소기업을 위한 빅데이터 분석 플랫폼 구축 사례에 대한 조사<sup>21)</sup>

본 연구에서 구축하려는 빅데이터 분석 플랫폼의 개념도는 그림 2-6과 같이, 서비스 공급자의 관점에서 수요기업에 맞춤형 서비스를 제공하고, 내부의 분석 전문가가 분석을 쉽게 할 수 있도록 지원하는 데 초점을 맞추고 있다. 데이터 소스는 개별 제조 기업의 기간계 시스템 데이터와 공개된 경제, 사회, 문화 지표 데이터 등이 될 수 있으며 데이터 추출, 전송, 적재 과정을 거쳐서 개별 영역(Private Sector)에 있는 ODS에 저장된다. ODS에 있는 데이터는 공통

21) 이선표, 고희석, 정희운, 박승범 (2020). 중소기업을 위한 제조 빅데이터 분석 플랫폼 구축. Journal of Information Technology and Architecture. Vol. 17. No. 3, pp. 239-253.

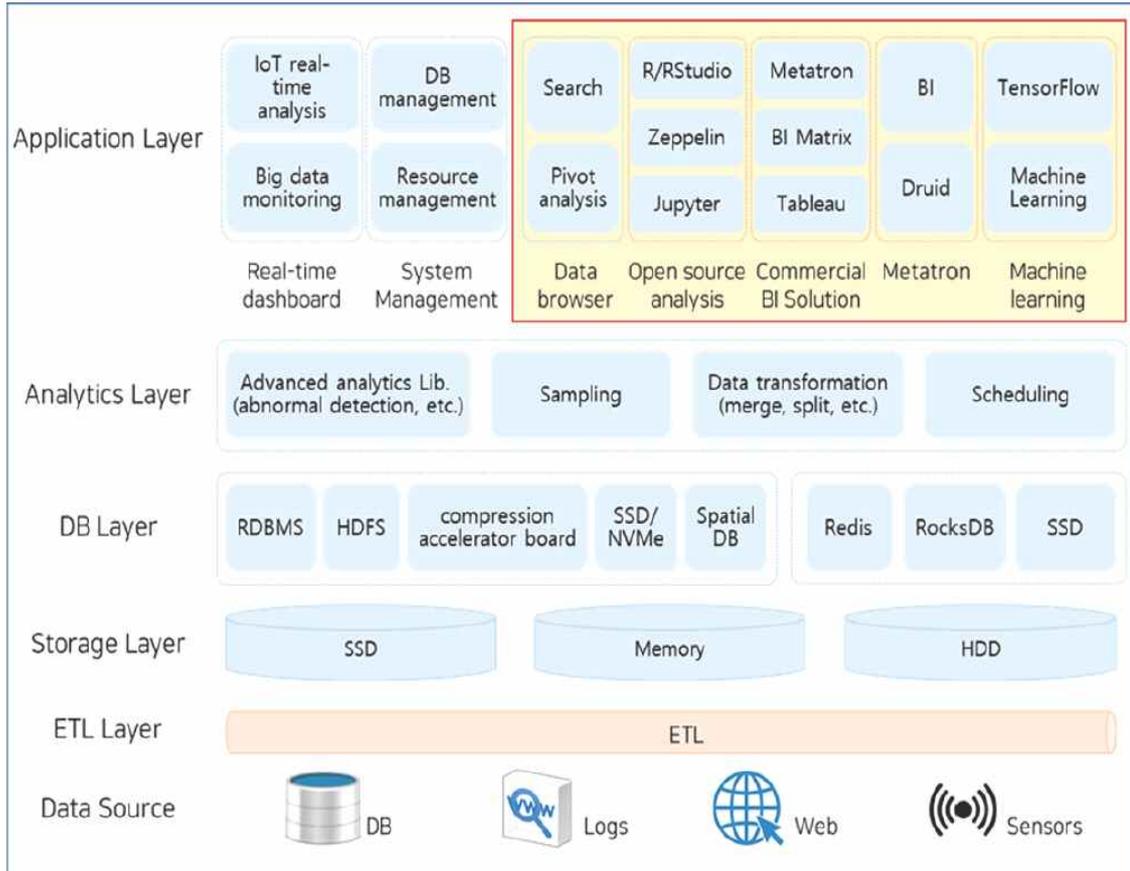
영역(Public Sector)에 있는 클라우드 데이터베이스에 적재가 되어 분석가는 다양한 도구와 알고리즘을 적용하여 분석을 수행한다. 개별 영역과 공통 영역을 구분한 이유 중 하나는, 아직 국내 중소 제조기업들은 클라우드 서비스의 보안이 취약할 것이라 막연하게 인식해 품질 및 제품 가공 레시피와 관련된 정보는 외부의 저장소에 저장하기를 꺼리는 경향이 있어, 사내에 저장하는 개념으로 구분한 것이다. 한편, 향후 서비스가 확대되었을 때 개별 영역 신속성을 요구하는 데이터를 주로 처리하는 엣지 컴퓨팅의 개념으로 활용될 것을 염두에 두었다.



[그림 2-6] 빅데이터 분석플랫폼의 개념도

그림 2-7은 플랫폼의 아키텍처를 표현한다. 해당 솔루션은 온디스크 (on-disk)와 인메모리(in-memory)를 함께 적용한 하이브리드 DB를 활용하고, 서비스 무정지 증설(scale-out)이 가능하며, 데이터 이중화에 따른 안정성을 보장하고 대용량 쿼리의 분산 처리를 지원한다는 특징이 있다. 또한 데이터 중복을 배제하여 최적의 하드웨어로 운용이 가능하고 이에 따른 인프라 자원 비용이 절감되며, RDBMS와 빅데이터 오픈소스를 결합하여 더욱 향상된 빅데이터 처리가 가능하다는 장점이 있다.

빅데이터 분석가의 관점에서는, 분석을 쉽게 하는 데이터 브라우저가 추가되어 원시 데이터를 직접 검색 또는 분석할 수 있는 기능을 제공하고 집계와 전처리가 가능하다. 상용 BI분석 도구와 오픈소스를 통합해서 한 플랫폼 안에서 사용하고, 분산시스템 적용을 위하여 분석코드의 프레임워크도 제공한다.



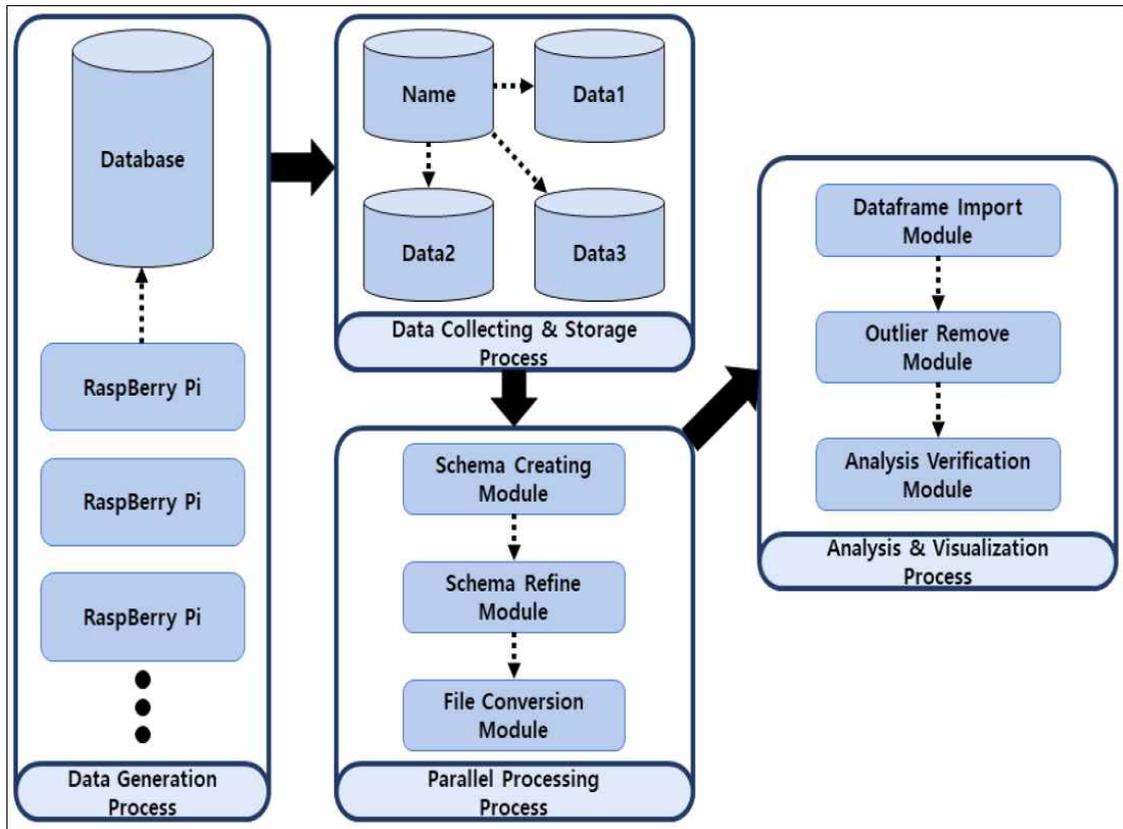
[그림 2-7] 빅데이터 분석 플랫폼의 아키텍처

#### 나) 빅데이터 기반 IoT 센서 데이터 처리 및 분석 사례에 대한 조사<sup>22)</sup>

빅데이터 기반 센서 데이터 처리 및 분석 시스템은 그림 2-8과 같이 데이터 생성 프로세스, 데이터 수집 및 저장 프로세스, 병렬 처리 프로세스, 분석 및 시각화 프로세스의 4가지로 설계 및 구현된다. 데이터 생성 프로세스의 하드웨어는 라즈베리 파이 B+이고, 데이터 생성 센서는 DHT11이다. 생성 데이터는 마리아(Maria)DB 서버로 자동으로 전송된다. 데이터베이스는 네이버에서 제공하

22) 신동진, 박지훈, 김주호, 광광진, 박정민, 김정준. (2019). IoT 환경을 위한 빅데이터 기반 센서 데이터 처리 및 분석. The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 19, No. 1, pp. 117-126.

는 mariaDB 서버를 이용하였다. 데이터 수집 및 저장 프로세스에서 빅데이터 수집 솔루션인 스킵(Sqoop)을 DB에 하둡(Hadoop)의 HDFS로 테이블 형태의 데이터를 저장하고 4개 노드로 데이터를 완전 분산모드 구성으로 분산하여 저장한다. 병렬 처리 프로세스에서는 하둡 프레임워크 상위에서 하이브(Hive)를 빅데이터 처리 솔루션을 이용하여 데이터 처리 및 정제가 가능하다. 분석 및 시각화 프로세스에서는 빅데이터 분석 및 시각화 솔루션의 R 프로그래밍을 활용하여 분석하고, 본문에서 구현한 시스템의 과정이 완료된다.



[그림 2-8] 빅데이터 기반 센서 데이터 처리 및 분석 시스템 구조도

본 논문에서는 IoT 환경에서 자주 사용되는 라즈베리파이를 이용하여 데이터를 생성하며, 생성되는 데이터를 데이터베이스로 전송하고, 빅데이터 처리 과정을 통해 정제된 온도와 습도 데이터의 관계를 검증하는 연구를 수행하였다. 라즈베리파이에서 발생하는 데이터는 작은 값이지만, 오랫동안 데이터를 생성하면 대용량의 데이터가 만들어지며, 데이터를 수집 및 저장하고 분석하기까지 사용돼 빅데이터 솔루션의 활용도가 증가하게 된다.

다) 스마트팩토리 실시간 제조설비의 빅데이터기반 처리 사례 조사<sup>23)</sup>

제조설비로부터 실시간으로 제조설비 데이터를 수집하여 분산 저장 및 처리하는 프로세스와 실시간 설비 데이터를 분석 및 시각화하는 프로세스를 포함한다. 대표적인 NoSQL(Not Only SQL)인 몽고(Mongo)DB는, 문서 기반의 데이터 저장소이기 때문에 확장성에 대한 RDBMS의 한계점을 해결할 수 있으며, 대용량 데이터 핸들링에 적절하고 저장 방식의 제한이 적다. 몽고DB는 정형, 비정형 데이터를 포함해 여러 형식의 데이터를 수용 가능하며, RDBMS의 전통적인 행 개념이 아닌 문서를 사용하는 문서지향 데이터베이스로 유연한 데이터 처리 모델로 사용된다. MTConnect는 다양한 설비에 대한 모니터링을 위한 소프트웨어 도구이다. Agent를 통해 복잡다양한 설비 데이터들을 합치고, 사용자에게 이를 보여주며 상호운용성을 증진할 수 있다. 실시간으로 이기종 설비 로그 데이터를 모니터링하려면, 큐의 구조로 내부 자료구조가 되어 있어, 이전 데이터는 삭제되므로 로그 데이터를 저장하기 위한 구조가 필요하다.

표 2-3에서 볼 수 있는 것처럼, MTConnect를 통해 로봇 데이터를 수집하여 CSV 파일로 저장하였다. 몽고DB를 사용하면 데이터를 원본 그대로 저장할 수 있으며, MySQL과 같은 RDBMS와 달리 테이블 스키마가 없기 때문에 MTConnect가 사용하는 SNMP 프로토콜을 저장하는 데 유리하다. MTConnect에서 수집한 CSV 파일을 mongoimport 명령어를 사용하여 몽고DB에 저장하였다.

---

23) 황승연, 신동진, 곽광진, 김정준, 박정민. (2019). 스마트팩토리를 위한 빅데이터 기반 실시간 제조설비 데이터 처리. The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 19, No. 5, pp.219-227.

[표 2-3] 로봇 팔 데이터

	A	B	C	D	E	F	G	H	I	J	K	L
1	Time	Goods	Exevution_Mode	Load	Vehicle_Position	Vehicle_Direction	J1_Angle	J1_Direction	J1_Spped	J2_Angle	J2_Direction	J2_Spped
26	2017-10-16 오후 1:11:38	elecmtor01	TESTING	EMPTY		0 FRONT	-25 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
27	2017-10-16 오후 1:11:38	elecmtor01	TESTING	EMPTY		0 FRONT	-26 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
28	2017-10-16 오후 1:11:38	elecmtor01	TESTING	EMPTY		0 FRONT	-27 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
29	2017-10-16 오후 1:11:38	elecmtor01	TESTING	EMPTY		0 FRONT	-28 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
30	2017-10-16 오후 1:11:38	elecmtor01	TESTING	EMPTY		0 FRONT	-29 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
31	2017-10-16 오후 1:11:38	elecmtor01	TESTING	EMPTY		0 FRONT	-30 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
32	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-31 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
33	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-32 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
34	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-33 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
35	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-34 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
36	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-35 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
37	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-36 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
38	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-37 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
39	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-38 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
40	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-39 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
41	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-40 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
42	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-41 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
43	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-42 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
44	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-43 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
45	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-44 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
46	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-45 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
47	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-46 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
48	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-47 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	
49	2017-10-16 오후 1:11:39	elecmtor01	TESTING	EMPTY		0 FRONT	-48 CW		288 UNAVAILABLE	UPWARD	UNAVAILABLE	

```

hadoop@hadoop-name: ~/Downloads
> db.machine.find().pretty()
{
  "_id" : ObjectId("5b570033650b984..."),
  "Time" : "2017-10-16 1:11:38",
  "Goods" : "elecmtor01",
  "Exevution_Mode" : "TESTING",
  "Load" : "EMPTY",
  "Vechicle_Position" : 0,
  "Vechicle_Direction" : "FRONT",
  "J1_Angle" : -1,
  "J1_Direction" : "CW",
  "J1_Spped" : 288,
  "J2_Angle" : "UNAVAILABLE",
  "J2_Direction" : "UPWARD",
  "J2_Spped" : "UNAVAILABLE",
  "J3_Angle" : "UNAVAILABLE",
  "J3_Direction" : "UPWARD",
  "J3_Spped" : "UNAVAILABLE",
  "J4_Angle" : "UNAVAILABLE",
  "J4_Direction" : "CW",
  "J4_Spped" : "UNAVAILABLE",
  "J5_Angle" : "UNAVAILABLE",
  "J5_Direction" : "UPWARD",
  "J5_Spped" : "UNAVAILABLE",
  "J6_Angle" : "UNAVAILABLE",
  "J6_Direction" : "CW",
  " : "UNAVAILABLE\r"
}

```

[그림 2-9] 몽고DB의 로봇 팔 데이터

그림 2-9에서 보는 것과 같이 몽고DB에 정상적으로 데이터가 저장된 것을 확인할 수 있다. 몽고DB에서 사용하는 고유키 값인 \_id는 데이터 생성 시에 자동으로 함께 생성된다. Time은 설비 데이터가 감지되는 시점이며, Goods는 설

비의 ID를 의미한다. Execution\_Mode는 동작 모드, Load는 로봇팔이 물건을 운반 중인 상태에 관한 여부를 나타낸다. Vechicle\_Position은 로봇팔을 운반하고 있는 기계의 위치값을 의미하고, Vechicle\_Direction은 로봇팔을 운반하고 있는 기계가 바라보는 방향을 의미한다. J1~J6은 로봇팔의 관절을 나타내며, Angle은 관절의 각도, Direction은 관절에 달린 모터의 회전 방향, Speed는 모터의 회전 속도를 의미한다. 초당 패킷 양을 산출하기 위해 몽고DB의 질의어(MQ)라는 변수에 저장하고, MQ변수를 통해 정제된 데이터를 result collection에 저장하였다.

라) 예지정비를 위한 클라우드 기반 IoT 플랫폼 구현 사례 조사<sup>24)</sup>

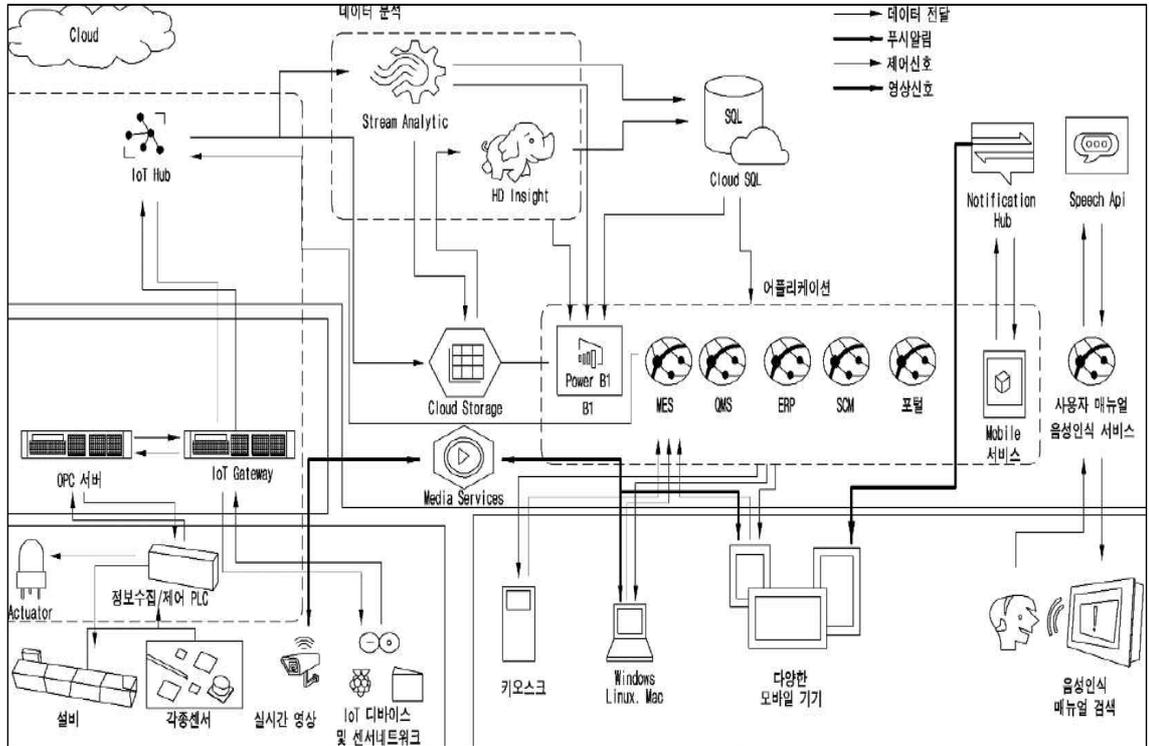
본 논문에서는 국내 중소기업의 경쟁력 강화를 위하여 그림 2-10과 같이 스마트팩토리 고도화에 필요한 클라우드 기반 IoT 플랫폼을 소비자 인터넷 대표 기업인 마이크로소프트의 클라우드 플랫폼인 애저(Microsoft Azure)와 산업 인터넷 글로벌 기업인 SAP의 HANA 플랫폼에서 각각 구현해 보고자 한다.

그림 2-11은 SAP가 제안하는 스마트팩토리 분석 아키텍처이다. 모두 3단계로 나누어져 있는데 1단계에서는 공정데이터의 수집과 저장을 목적으로 하고 있으며, 주요 내용은 공정 데이터의 실시간 전처리 및 데이터의 적재에 있다. 2단계는 데이터 분석 및 인사이트 확보 단계로, 데이터 시각화나 예측분석 등을 통해 인사이트를 확보하고 및 규칙 적용을 통한 실시간 대응 기반을 마련하는 과정이다. 마지막 3단계는 시스템과 콘텐츠 고도화 및 비즈니스 연동을 실시하는 단계이다.

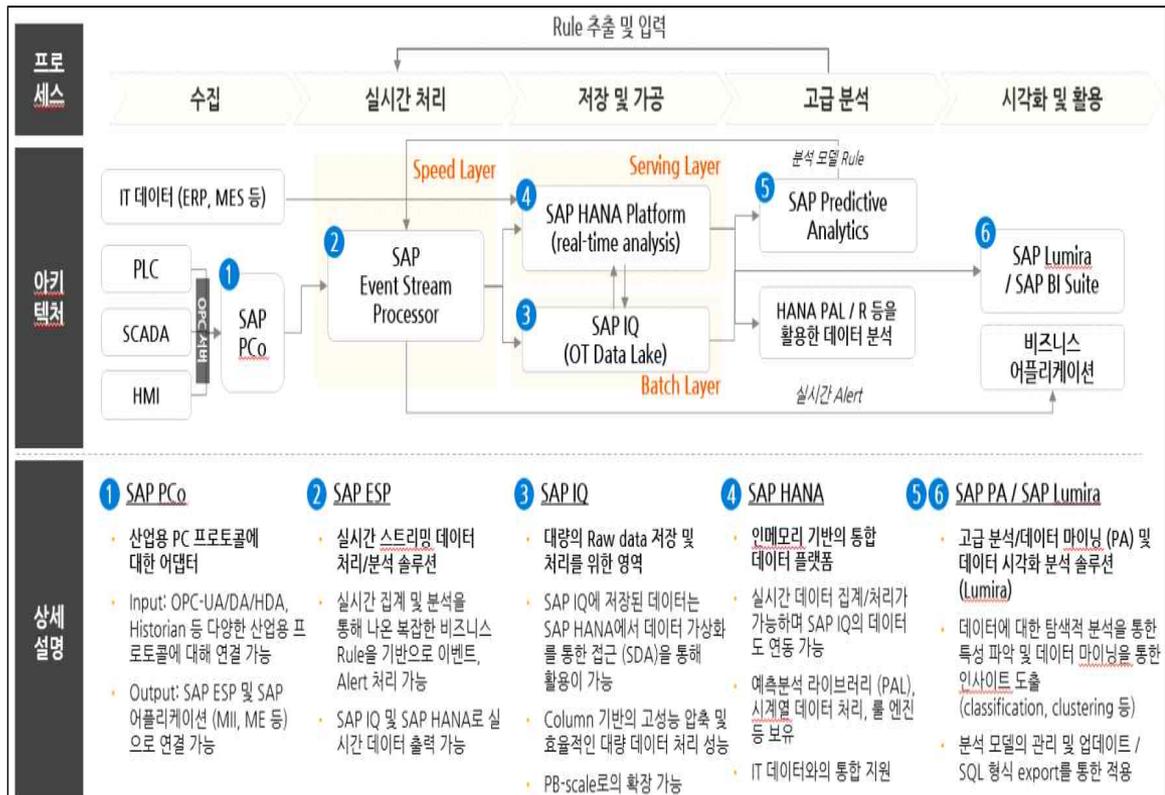
그림 2-12는 스마트팩토리 솔루션 전문기업에서 MES를 개발한 스마트공장 IoT 플랫폼 시스템 구성도를 나타낸 것인데, 이는 산업 인터넷 글로벌 기업인 SAP사의 클라우드 HANA 플랫폼의 최소기능제품 방식으로 개발하여 SAP에서 기술 인증을 받은 구성도이다.

---

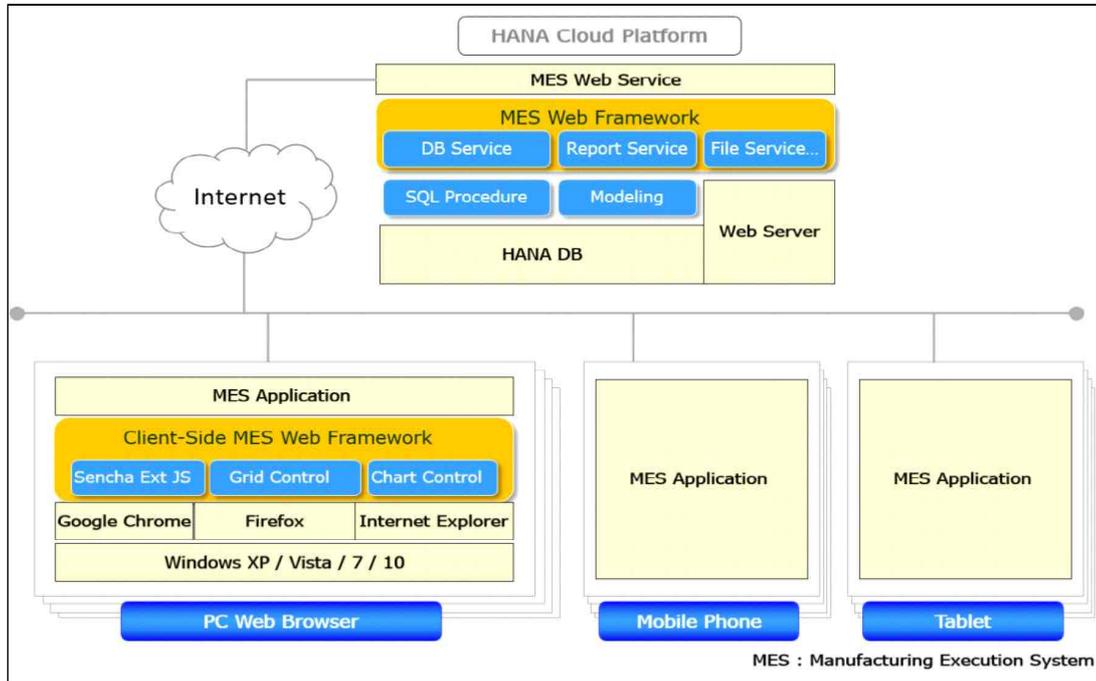
24) 손영욱, 김영국. (2018). 예지정비를 위한 클라우드 기반 스마트공장 IoT 플랫폼구현에 관한 연구. 한국창업학회



[그림 2-10] 애저 플랫폼 기반 스마트팩토리 구성도



[그림 2-11] SAP가 제안하는 스마트팩토리 분석 아키텍처



[그림 2-12] SAP HANA 플랫폼 기반 스마트 팩토리 구성도

마) 엔터프라이즈 아키텍처를 위한 .Net기반의 스마트클라이언트

애플리케이션 브라우저 설계 및 구현<sup>25)</sup>

유용한 인트라넷 기반 애플리케이션을 만들기 위해 필요한 것은 알맞은 웹 브라우저/저용 사용자 인터페이스와 서버 연결 방법이다. 이 용이한 배치는 인트라넷/엑스트라넷 애플리케이션을 효과적으로 구축하고 엔터프라이즈에서 이러한 종류의 애플리케이션이 지향되고 있다.<sup>26)27)</sup> 웹 서버에 연결한다는 점에서 웹 브라우저와 애플리케이션 브라우저는 유사하다. 애플리케이션 브라우저는 애플리케이션 서버의 기능을 사용하기 위해 애플리케이션 서버에 연결한다.<sup>28)</sup> 엔터프라이즈 아키텍처는 .Net 기반에서 프리젠테이션 계층의 스마트 클라이언트 애플리케이션 브라우저를 설계하고 구현하였다.

25) 김영주, 전계논문, pp. 203-209.

26) <http://msdn.microsoft.com/ko-kr/library/aa905318.aspx>

27) <http://msdn.microsoft.com/en-us/library/aa192611.aspx>

28) Reynolds, M. and Watson, K. (2002). *\_NET Enterprise Development in C#: 설계에서 배치까지* (임성준 역). 정보문화사.

### (1) 스마트 클라이언트

스마트 클라이언트는 응용 프로그램으로서 리치 클라이언트와 썬 클라이언트의 이점을 결합하여 두 클라이언트 사이의 간격을 메운다. 리치 클라이언트는 데이터를 로컬로 처리하기 때문에, 다른 애플리케이션 등의 주변 및 외부 환경에 대한 인식도가 낮고 사용자의 컴퓨터에만 격리되는 등 한계를 가지고 있다. 썬 클라이언트는 배포 및 업데이트가 중앙 웹서버에서만 이루어져야 하는 한계점이 있어, 네트워크에 항상 연결되어 있어야 하는 종속성이 많은 응용프로그램과 환경에 부담으로 작용한다.

표 2-4는 스마트 클라이언트와 웹 클라이언트 및 타 솔루션에 대한 성능 및 특징을 비교하였다.

### (2) 애플리케이션 브라우저

개발자들과 시스템 관리자들은 설치를 위한 일정한 단계들을 따라야 할 필요 없이 사용자들에게 새 기능을 제공할 수 있도록 허용한다.

애플리케이션 브라우저에서 다운로드 서버에 접근할 수 있게 완성된 프로그램을 게시하면 프로그램 배포는 이루어진다. .Net은 ActiveX DLL처럼 레지스트리에 설치하지 않고 접근할 수 있는 디렉터리에 복사하여 실행한다.

[표 2-4] 웹 클라이언트 및 타 솔루션과의 비교

비교항목	웹 클라이언트	스마트 클라이언트	비교항목	타 솔루션	스마트 클라이언트
네트워크 환경	인터넷 기반(불특정 다수가 사용 가능)	인트라넷 기반	기술 요소	Active X, Java Script, Applet등(Java와 Active X의 혼재)	Winform, Office(.Net 개발언어 지원)
개발성 및 유지보수	웹 형태로 구현해 하기 때문에 클라이언트 서버보다는 개발 공수가 최소한 2배 이상 들어가지만, 유지보수는 편리함.	클라이언트 서버 형태로 개발 생산성이 뛰어나고 배포는 웹 방식이기 때문에 향후 유지보수가 편리함.	개발성 및 유지보수	Script형태의 개발이 포함되어 있기 때문에 개발 공수 디버깅 및 유지 보수가 어려움.	단일 개발환경에서 일관된 언어로 개발하기 때문에 개발 생산성이 뛰어나고 향후 유지 보수가 편리함.
풍부한 GUI(Graphical User Interface)	웹 형태이기 때문에 화려한 GUI나 단축키 등을 구현하는데 쉽지 않음.	클라이언트 서버 형태로 다양한 그래픽 GUI의 구현이 가능하고 단축키 구현도 가능함.	보안	Active X를 사용하는 경우 개발자에게 무한대의 보안권한을 주기 때문에 보안에 취약 할 수 있음.	.NET 프레임워크에서 제공하는 강력한 보안 수준을 설정해서 보안을 유지할 수 있음.
손쉬운 배포와 관리	비즈니스 로직이 서버 쪽에 위치해 있어서 클라이언트는 브라우저만 있으면 응답을 받을 수 있고 애플리케이션 배포나 변경을 손쉽게 관리할 수 있음.	애플리케이션을 데스크톱에서 다운로드 받아서 처리하고 변경이 있으면 항상 업데이트 된 모듈을 적용해서 사용할 수 있고 배포와 변경을 손쉽게 관리 할 수 있음.	성능	Formula One, TrustForm, Curl등의 현재 타 솔루션은 .Net 기반보다 현격하게 늦은 응답속도	.Net 기반으로 웹 서비스 및 리모팅 등 다 사용할 수 있고 빠른 응답 속도.
오프라인 기능 제공	기능을 제공하지 못함	오프라인 기능을 제공하는 애플리케이션 블록을 무상으로 제공해서 손쉽게 구현가능함.	오피스 도구 연계	별도의 씨드 파티 컴포넌트 적용을 시켜야 함.	엑셀, 워드, 아웃룩을 Native하게 이용할 수 있는 오피스 도구제공.
			오프라인 기능 제공	부분적이고 제한적으로 사용가능	오프라인 기능을 제공하는 애플리케이션 블록을 무상으로 제공해서 손쉽게 구현가능함.

바) 빅데이터 수집 및 처리 선행연구 정리 및 결과

(1) 선행연구 주요 내용

빅데이터 수집 및 처리 관련 선행연구 주요 내용은 표 2-5와 같다.

[표 2-5] 선행연구 주요 내역

N	연구 제목	주요 내용	ICT, SW, 알고리즘 기술
1	중소기업 제조 빅데이터 분석 플랫폼 구축	축적데이터 분석하여 최적값 도출, 실시간 모니터링 등 빅데이터 분석 플랫폼 아키텍처 구성함	RDBMS+오픈소스+BI 분석도구 데이터브라우저
2	빅데이터 기반 센서 IoT데이터 처리, 분석	1. 데이터생성-> 2. 수집/저장-> 3. 병렬처리-> 4. 분석/시각화 프로세스 과정의 솔루션 활용함	1.라즈베리/MariaDB, 2.Sqoop /HDFS(Hadoop) 3.Hive, 4. R프로그래밍
3	빅데이터 기반 실시간 설비데이터 처리	실시간 제조설비 데이터를 수집, 분산, 저장하고 분석 및 시각화하는 프로세스를 구성함	MongoDB(대표적인 NoSQL), R프로그래밍 MTConnect(SNMP)
4	예지정비를 위한 클라우드 기반 IoT 플랫폼 구현	Microsoft Azure와 SAP의 HANA 플랫폼에서 구현 수집, 실시간 처리, 저장·가공, 고급분석, 시각화·활용 메뉴와 프로세스로 활용	Microsoft Azure, SAP HANA 플랫폼, MES, ERP, SCM, 클라우드 개발/서비스, 웹 브라우저, 프레임워크 서비스
5	.Net기반의 애플리케이션 브라우저 설계, 구현	.Net기반의 3계층 스마트 클라이언트 애플리케이션 브라우저 개념으로 엔터프라이즈 솔루션을 설계 구현함	Google.Net 플랫폼, 3계층 구조 스마트 클라이언트 애플리케이션 브라우저

(2) 시사점 및 결과

빅데이터 수집, 처리 및 분석시스템에서는 적용 SW도구는 다양하게 사용되고 있으나 그중에서도 마리아DB, 라즈베리, R프로그래밍 등 오픈소스 SW를 활용이 증가하고 있고 애저, SAP HANA, Google.Net 플랫폼 등 글로벌 플랫폼에서 오픈소스 SW와 함께 사용이 활성화되는 추세이다. 그리고 기업에 사용되고 있는 데이터나 애플리케이션도 웹브라우저 개념으로 개발자나 사용자의 맞춤형으로 제공되고 되고 있으며 효율적인 DB 운영이나 보안 측면을 위해서 2계층에서 3계층 구조로 확산하고 있다.

## 제 2 절 IoT 통합(Integration) 및 통신 프로토콜 연구

### 1) IoT 통합 및 통신 프로토콜 관련 기술

#### 가) 산업용 IIoT 기술 현황<sup>29)</sup>

##### (1) IoT 센서

제조 환경에서 발생하는 다양한 데이터를 ①생성(센서) → ②전달(네트워크) → ③통합·분석·제어(플랫폼)하기 위한 기술 중심으로 범위를 설정하고 크게 분류하면 IoT센서, IoT네트워크, IoT 플랫폼으로 나누어 볼 수 있다.

사물인터넷에서 사물, 즉 물리적인 영역을 측정하는 IoT 센서는, 현장에서 제조 및 생산과 관련된 상태를 감지하여 재고 현황 및 생산환경 변화 등의 데이터를 추출해 전달한다. 최근에는 분석 및 자동제어와 같은 기능을 직접 수행하기도 하며, 특정 변화를 단순 측정 및 감지하는 수준을 넘어 자가진단, 데이터 처리 등을 할 수 있는 스마트 센서 기술로 빠르게 전환되고 있다.



[그림 2-13] 스마트 센서 아키텍처 및 형태

SoC(System on Chip) 및 MEMS(Micro Electro Mechanical System) 기술의 발전과 MCU의 내장으로 인해 그림 2-13과 같이<sup>30)</sup> 데이터의 처리 및 저장, 자동보정 및 자가진단, 통신, 의사결정 등의 고도화 기능을 소화 가능한 스마트 센서를 중심으로 발전하는 것이다. 스마트 센서는 변화는 환경을 인지하는 센서, 데이터를 수집하고 분석 및 처리하는 MCU, 그리고 서버와 데이터를 주고받는 통신모듈로 구성되며, 이러한 구조가 하나의 반도체 내에 집적화되어 있다. 앞으로도 복합화, 저전력화, 소형화 등을 중심으로 센서의 진화는 빠르게 진

29) 나영식. (2020). 제조용 IoT. KISTEP 기술동향브리프. 2020-10호. pp.4-18.

30) 나영식. 상계자료, pp. 6.

행될 전망이다.

(2) IoT 네트워크

IoT 네트워크는 주로 제조 환경 네트워크에서의 통신기술로서, 주로 산업용 유선 통신망을 활용하며 고신뢰를 기반으로 한다. 필드버스(Fieldbus)와 산업용 이더넷(Ethernet)이 대표적인데, 이더넷은 이더넷아이피(Ethernet/IP), 프로피넷(PROFINET) 등의 모습으로 북미와 유럽에서 중심으로 사용되고 있으며, 국내에서는 RAPINET을 적용하고 있다. 이와 같이 다양한 종류의 유선 통신이 등장함에 따라 유연한 생산의 어려움이나 호환성 문제가 발생하는데, 이를 해결하기 위해 다른 종류의 프로토콜을 하나의 회선을 통해 전송 가능한 TSN(Time Sensitive Networking) 기술을 도입하고 있다. 무선 통신 기술의 발전은 유연성을 필요하는 설비들이나, 복잡다양한 장비와 설비를 구축한 공장에 유용하다.

네트워크 기술이 고도화됨에 따라 사용량의 증가세뿐만 아니라 저전력 및 저지연을 중심으로 기술 개발 또한 가속화되고 있고, 이는 제조용 근거리 무선 통신과 근거리무선통신으로 나누어 볼 수 있다.

[표 2-6] 제조용 근거리 무선통신 기술 비교

구분	802.11(Wi-fi)	Wireless HART	ISA 100.11a	블루투스
통신 속도	150-450Mbps	250Kbps	250Kbps	24Mbps
주파수	2.4GHz/5GHz	2.4GHz	2.4GHz	2.4GHz
표준규격	<ul style="list-style-type: none"> <li>• IEEE Std 802.11n</li> <li>• IEEE Std 802.11ac</li> </ul>	<ul style="list-style-type: none"> <li>• IEEE Std 802.15.4</li> <li>• IEC 62591</li> </ul>	<ul style="list-style-type: none"> <li>• IEEE Std 802.15.4</li> <li>• IEC 62734</li> </ul>	<ul style="list-style-type: none"> <li>• IEEE Std 802.15.1</li> </ul>
주요 활용	<ul style="list-style-type: none"> <li>• 휴대폰</li> <li>• PC 입출력장치</li> <li>• Industrial FA</li> </ul>	Industrial PA	Industrial PA	<ul style="list-style-type: none"> <li>• 휴대폰</li> <li>• PC 입출력장치</li> <li>• 무선 공유기</li> </ul>
특징	<ul style="list-style-type: none"> <li>• 디바이스 확장 불가</li> <li>• 250m까지 통신 가능</li> <li>• 추가 보안 개발 필요</li> </ul>	<ul style="list-style-type: none"> <li>• HART 기반</li> <li>• 표준 보안 사용 가능</li> </ul>	<ul style="list-style-type: none"> <li>• IPv6 패킷 기반</li> <li>• 표준 보안 사용 가능</li> </ul>	<ul style="list-style-type: none"> <li>• 페어링 필요</li> <li>• 끊김 현상 발생</li> </ul>

근거리 무선 통신 기술은 표 2-6과 같이 플랜트 영역에서는 와이파이와 블루투스를 많이 사용하며, 센서 및 액추에이터 영역에서는 ISA 100.11a와 WHART가 활용되고 있다. 하지만 이는 근거리에 주로 사용되기 때문에, 원거리 공장 간의 통신이나 지역 확대 등으로 통신 거리에 한계가 있다. IoT 전용 저전력 원거리 무선 통신 기술인 LPWA기술은 표 2-7과 같이 IoT 디바이스를 대상으로 한 이동통신망으로 수십 km 거리로 넓은 범위와 10년 이상의 긴 배터리 수명, 설치비용에 장점이 있다. 최근 5G 무선 통신 기술은 지금까지의 제조업에서 사용되어온 기존의 유무선 네트워크의 기술적 한계 극복으로서 최신의 초저지연 및 초연결 무선통신 기술이 국내외 대기업 중심으로 도입되기 시작하고 있다.

[표 2-7] LPWA관련 주요 무선 통신 기술 비교

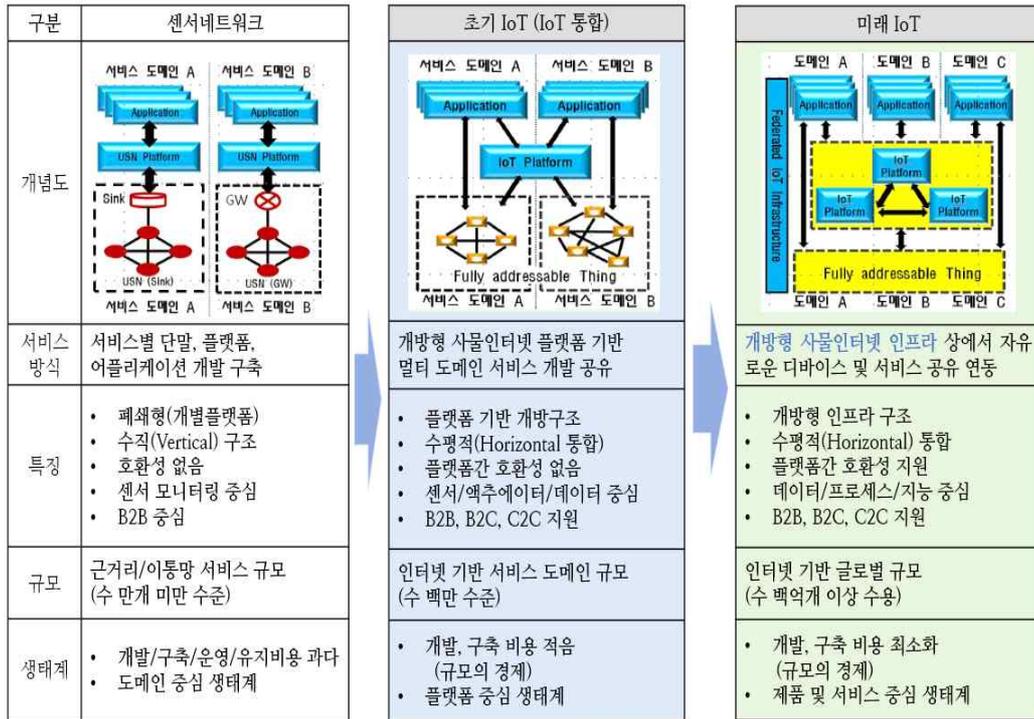
구분	LTE-M	NB-IoT	LoRa	SigFox
통신 거리	~11km	~15km	~11km	~13km
통신 속도	~1Mbps	~150kbps	~10kbps	~100bps
주파수대역	면허 대역 (1.4MHz)	면허 대역 (200KHz)	비면허 대역 (900MHz)	비면허 대역 (900MHz)
장점	<ul style="list-style-type: none"> <li>• 전국 서비스 가능</li> <li>• 통신 품질의 안정성</li> <li>• 기존 네트워크 활용</li> </ul>	<ul style="list-style-type: none"> <li>• 통신 품질의 안정성</li> <li>• 실내 커버리지 가능</li> </ul>	<ul style="list-style-type: none"> <li>• 저전력 장거리 통신</li> <li>• LTE-M 모듈의 1/4 가격</li> <li>• 저렴한 구축 비용</li> </ul>	<ul style="list-style-type: none"> <li>• 저전력 장거리 통신</li> </ul>
단점	<ul style="list-style-type: none"> <li>• 고가의 통신 모듈 가격</li> </ul>	<ul style="list-style-type: none"> <li>• LoRa에 비해 통신모듈 가격이 비쌈</li> </ul>	<ul style="list-style-type: none"> <li>• 비면허 대역으로 네트워크 불안정</li> </ul>	<ul style="list-style-type: none"> <li>• 전세계 한정적 인프라</li> <li>• 비면허 대역으로 네트워크 불안정</li> </ul>

### (3) IoT 통합 및 플랫폼

그림 2-14와 같이<sup>31)</sup> IoT 통합 및 플랫폼 기술을 살펴보면 하위 IoT 디바이스에서 입수한 정보를 최상위 애플리케이션에 전달하는 중간 연결다리 격인 시스템으로서, 디바이스로부터 수집된 데이터를 분석하고 생산과정을 관리 및 제

31) 나영식, 전계서, pp. 16.

어하여 상위 애플리케이션과 연계할 수 있는 1) 서비스 및 리소스 관리, 2) 네트워크 및 연결 관리, 3) 데이터 확보 및 처리 등의 기능을 시스템으로 구성한다. 최근에는 여러 가지 도메인 서비스에 대한 통합을 위해 상호호환성과 오픈소스를 기반으로 한 개방형 통합 플랫폼을 추구하고 개발이 이루어지고 있다.



[그림 2-14] 사물 인터넷 플랫폼 기술 특징 및 발전 방향



[그림 2-15] 삼성 SDS의 Brightics IoT

그림 2-15와 같이<sup>32)</sup> 삼성SDS의 “Brightics IoT Core”는 여러 종류의 통신 프로토콜뿐 아니라 IoT 국제표준을 지원하고, Brightics AI 플랫폼을 연계하여 IoT에서 수집한 데이터를 자연어처리(NLP, Natural Language Processing), 딥러닝 등의 기술로 분석하여 고도화된 기능을 제공하고 있다.

나) 통신 프로토콜 현황

(1) IoT 메시징 프로토콜<sup>33)</sup>

사물인터넷 메시징 프로토콜 기술을 살펴보면, 표 2-8과 같이 지금까지의 HTTP는 오직 하이퍼텍스트 교환을 위한 프로토콜로써 사물인터넷의 기능에 한계가 있어, CoAP와 MQTT는 사물인터넷 환경과 저전력 기기에 HTTP를 대체할 수 있는 대표적인 프로토콜이다. MQTT는 M2M과 사용하기 위한 경량화 메시지 전송 프로토콜이다. 작은 고정 헤더를 가지며, 적은 오버헤드로 패킷을 처리할 수 있어 낮은 대역폭 환경에서 저전력 디바이스에 사용될 수 있고 통신 구조는 브로커 기반의 구독/게시 모델이다. 서버의 개입 없이 사물 사이에 정보 공유를 통해 새로운 서비스를 생성하는 M2M 기반 사물인터넷 서비스에 적합한 구조이다. CoAP는 저전력, 고손실 네트워크 및 소형노드에 사용될 수 있는 경량 프로토콜이다. HTTP와 유사한 메시지 구조를 가지며, RESTful 구조의 요청/응답 모델이고 확장 옵션을 통해 부분적으로 게시/구독 모델을 지원한다.

[표 2-8] HTTP, MQTT 및 CoAP의 특성 비교

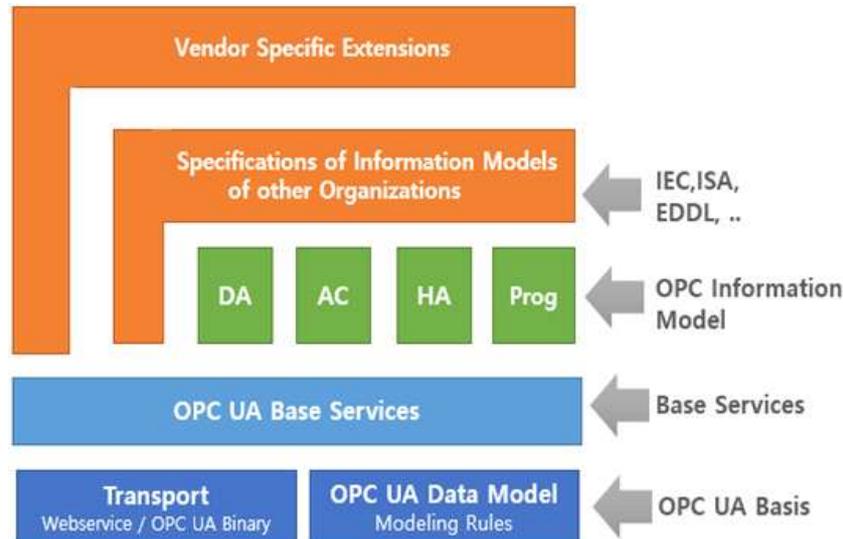
프로토콜	관련 단체	요청/응답	게시/구독	전송계층	보안
HTTP	IETF HTTP Working Group	지원	-	TCP, QUIC	TLS
CoAP	IETF CoRE Working Group	지원	부분 지원	TCP, UDP	TLS, DTLS
MQTT	OASIS	-	지원	TCP	TLS
MQTT-SN	OASIS	-	지원	UDP, Zigbee	DTLS

32) 나영식. 전계서, pp. 18.

33) 사물인터넷표준연구실. (2020). 사물인터넷 기술동향. IoT표준연구실 IT대학 컴퓨터학부 경북대학교 pp 12-13.

## (2) OPC-UA

OPC-UA(Open Platform Communication Unified Architecture) 프로토콜은 플랫폼에 의존적이지 않은 독립적 구성이 이상적이고 안정적인 메커니즘을 가진 통신 프로토콜이다.<sup>34)35)</sup> Industry 4.0의 구현을 위한 가장 이상적인 통신 프로토콜이라고 할 수 있다.



[그림 2-16] OPC-UA 플랫폼 계층 구조

그림 2-16은<sup>36)</sup> 구현된 OPC-UA의 계층 구조를 나타내고 있다. 제조공정에서 원활한 소프트웨어 기술의 융복합을 위해 표준기반 통신의 필요성이 높아지고 있으며, 이러한 추세에 따라 OPC-UA가 주목받고 있다. OPC-UA는 IEC 62541 표준에 설명되고 있으며 PLC, HMI, SCADA(Supervisory Control and Data Acquisition), MES 등 서로 다른 산업용 자동화 장치 간의 신뢰성이 있고 안전하며 상호 호환적인 데이터 교환을 가능하게 한다.<sup>37)</sup> 그림 2-16과 같이

34) OPC Unified Architecture Specification. (2015). Part 1: Overview and Concepts, OPC Foundation,

35) OPC Unified Architecture. (2016). Part 1: Pverview and Concepts, IEC 62541-1,

36) Kim, J., Kim, Y., Kim, S. (2021). A Study on Design of Real-time Big Data Collection and Analysis System based on OPC-UA for Smart Manufacturing of Machine Working. International Journal of Internet, Broadcasting and Communication Vol. 13 No. 4, pp. 122.

37) 이용수, & 정종필. (2014). 스마트 제조를 위한 AAS 와 OPC UA 기반 설비모니터링 시스템의 설계 및 구현. 한국인터넷방송통신학회 논문지, 21(2), 41-47.

OPC-UA의 플랫폼 독립적인 구조는 임베디드 플랫폼부터 고성능 시스템까지 다양한 규모의 시스템이 상호 접근 가능한 환경을 제공한다. OPC-UA 정보모델은 대상 시스템의 구조 및 데이터를 표준모델의 형태로 정의하고 기술할 수 있게 함으로써 서로 다른 도메인에서 시스템 간 표준통신을 가능하게 한다.<sup>38)</sup>

## 2) IoT 통합 및 통신 프로토콜 선행연구 및 결과

가) PLCopen OPC-UA 프레임워크의 설계, 구현 사례에 대한 조사<sup>39)</sup>

OPC-UA는 이종 기기들 간 표준화된 데이터 교환을 위해 정보 모델링의 개념을 채택한다. PLC 소프트웨어는 다른 제어기와 수평적으로 데이터를 교환하거나 SCADA, MES, ERP(Enterprise Resource Planning)와 같은 관리 시스템과 수직적으로 데이터를 교환할 필요가 있다.

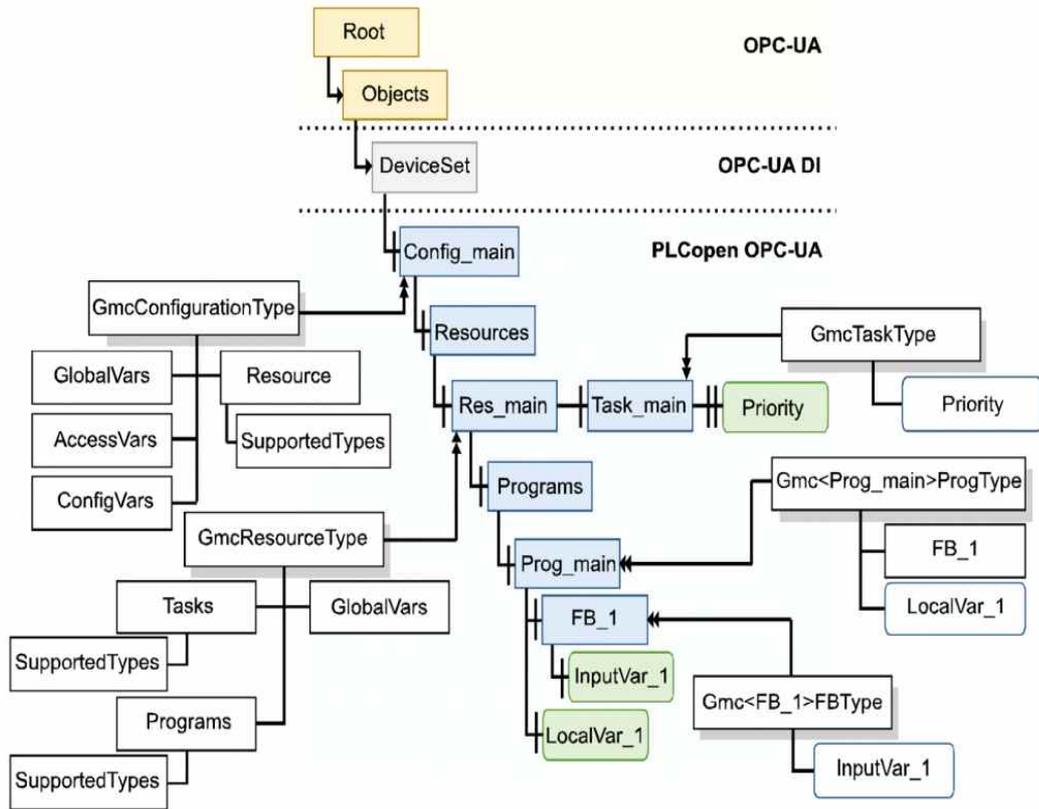
본 논문은 OPC-UA 통신으로 PLC 프로그램을 개발하는 목적의 PLCopen OPC-UA 프레임워크를 제시한다. 제안된 프레임워크의 개요는 그림 2-17과 같다. 해당 프레임워크는 PLCopen 표준 언어로 작성된 PLC 프로그램의 주요 데이터 구조 및 메소드 호출 정보를 포함하는 PLCopen OPC-UA 정보 모델 생성 기능과, 외부 OPC-UA 서버 데이터 및 메소드에 접근 가능한 PLC 프로그램의 작성을 위한 OPC-UA 클라이언트 함수 블록 라이브러리를 포함한다. PLCopen OPC-UA 프레임워크는 기존 PLC 프로그램 개발 환경인 Beremiz<sup>40)</sup>에 추가되어 OPC-UA 서버 및 클라이언트 기능을 제공한다. Beremiz는 IEC61131-3 표준을 준수하는 오픈소스 통합 개발 환경이다. 사용자는 이를 사용하여 PLC 프로그램을 작성할 수 있다. 개발 환경은 사용자가 작성한 프로그램을 타깃 PLC 장치(그림 2-17의 Target PLC)가 실행할 수 있는 응용(그림 2-17의 PLC executable)으로 빌드하고 해당 장치에 전송한다.

---

38) 이용민, 이원복, & 이승호. (2019). 스마트 팩토리 구축을 위한 프로토콜 개발. 전기전자학회 논문지, 23(3), pp. 1096-1099.

39) 오승엽, 김웅기, 성민영. (2018). PLC를 이용한 산업용 제어 응용을 지원하는 OPC-UA 프레임워크의 설계 및 구현. 한국정보과학회. Vol. 24, No. 10, pp. 513-526.

40) Kim, I., Kim, T., Sung, M., Tisserant, E., Bessard, L. and Choi, C. (2013). An Opensource development environment for industrial automation with EtherCAT and PLCopen motion control. Proc. of the IEEE 18th Conference on Emerging Technologies and Factory Automation (ETFA), IEEE, Cagliari, Italy, pp. 1-4.



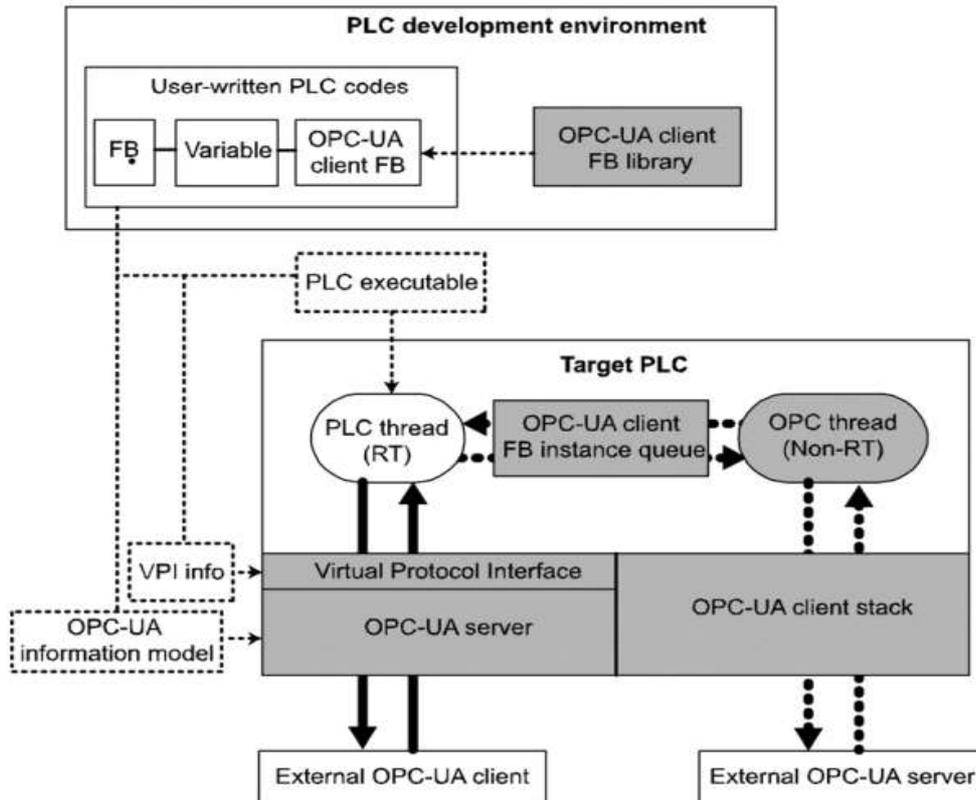
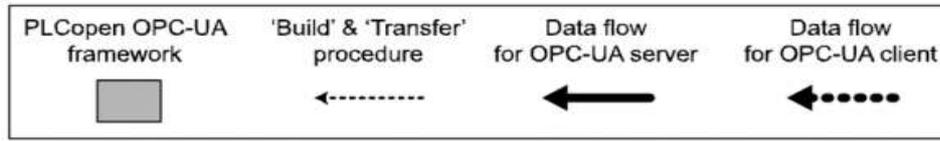
[그림 2-17] PLCopen OPC-UA 프레임워크 개요도

OPC-UA 서버와 클라이언트 기능을 지원하는 PLCopen OPC-UA 프레임워크의 구현을 위해 OpenOpcUa 오픈 스택을 사용하였다. 사용자가 작성한 PLC 프로그램은 프레임워크에 내장된 OPC-UA 서버를 통해 외부의 OPC-UA 클라이언트에게 PLC 프로그램의 변수 접근 및 메소드 원격 실행 등을 허용한다. 또한 내장된 OPC-UA 클라이언트를 통해 OPC-UA 네트워크에 접근하여 외부 환경에 능동적으로 대응할 수 있다.

프레임워크에 의해 생성되는 PLCopen OPC-UA 정보 모델 구조는 그림 2-18에서 확인할 수 있다. OPC-UADI(Device Integration) 객체인 DeviceSet에 PLCopen OPC-UA 객체들을 추가하였다. 외부 OPC-UA 클라이언트가 접근할 수 있도록 그림 2-18과 같이 PLCopen OPC-UA 프레임워크에 서버 기능을 말한다. 구현된 서버는 프레임워크로 OPC-UA 생성정보 모델을 활용하여 노드와 주소 공간들을 만들고, 외부 OPC-UA 클라이언트 요청을 처리한다.

본 논문에서는 산업 자동화 환경에 사용하고 있는 PLC에 표준 기반 통신

기술을 PLCopen OPC-UA 프레임워크로 적용하여 구현하였다.



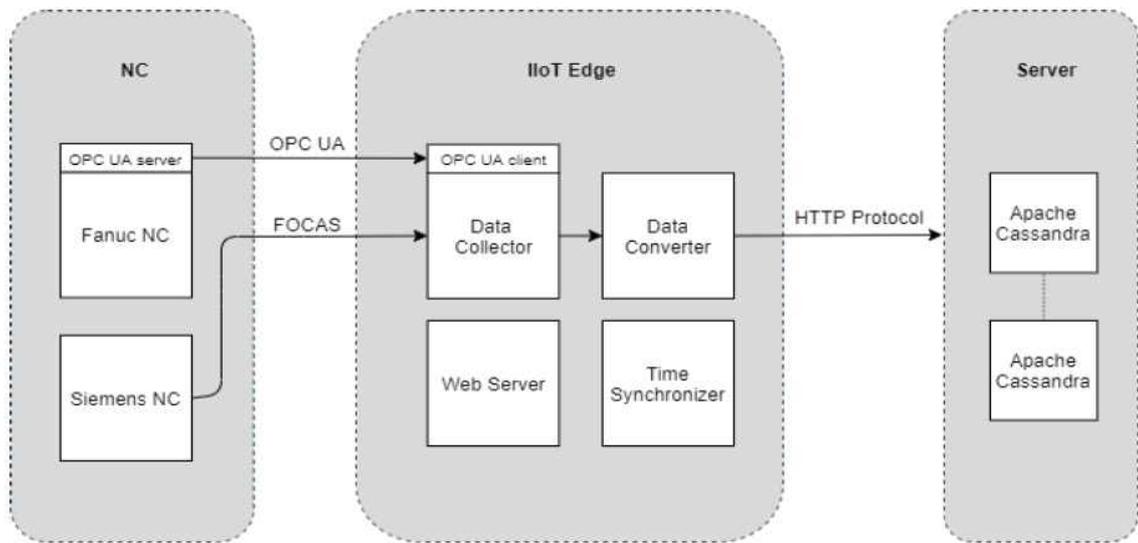
[그림 2-18] PLCopen OPC-UA 프레임워크에서 OPC-UA 정보 모델

나) 이기종 공작기계 데이터 수집을 위한 IIoT 엣지 디바이스 설계 및 구현에 대한 사례 조사<sup>41)</sup>

본 논문에서 제안하는 시스템 구성에서는, 그림 2-19와 같이 IIoT 엣지 디바이스가 OPC-UA와 FOCAS 라이브러리를 이용하여 공작기계의 데이터를 수집한다. IIoT 엣지 디바이스에서 클라우드로의 데이터 전송은 HTTP 방식으로 수행한다. 이때 네트워크 비용 및 클라우드 분석 대상 데이터 범위를 줄이기 위해 사전에 정의된 규칙에 의해 데이터를 필터링할 수 있다. OPC-UA는 기본적

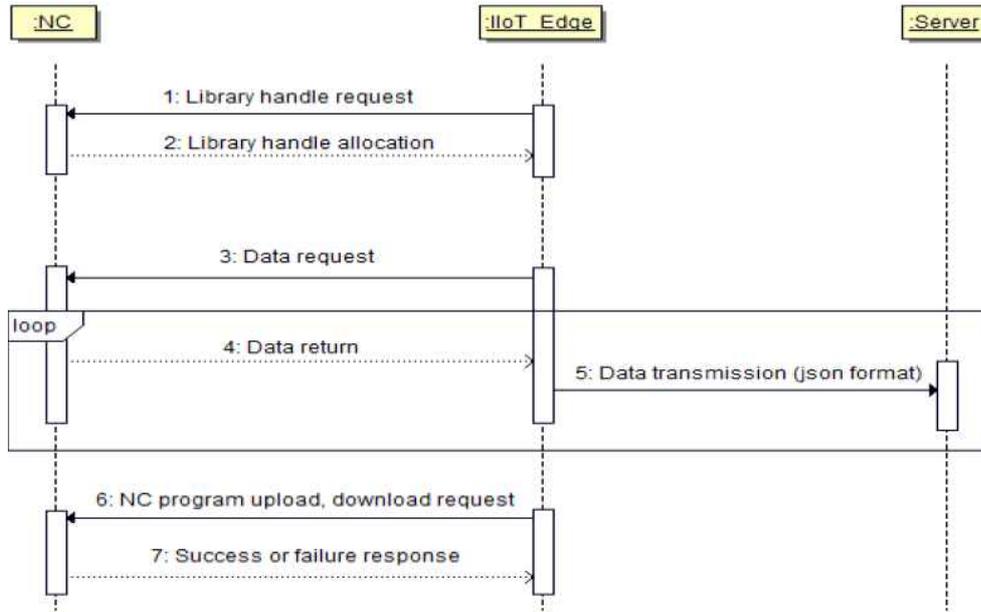
41) 유흥식, 박유현. (2021). 이기종 공작기계 데이터 수집을 위한 IIoT 엣지 디바이스 설계 및 구현. 인터넷전자상거래연구, 제21권, 1호, pp. 23-32.

으로 클라이언트/서버 구조를 가지며, 정보를 제공하는 쪽이 서버의 역할을 하고, 정보를 소비하는 쪽이 클라이언트 역할을 한다. 클라이언트와 서버는 모두 동일한 OPC-UA 스택을 가지며, 서비스를 이용하고 싶은 클라이언트는 언제든지 해당하는 OPC-UA 서버의 API를 호출하여 원하는 서비스 또는 정보를 이용할 수 있다. 본 논문에서는 OPC-UA 클라이언트는 IIoT 엣지 디바이스상에서 구현되고, OPC-UA 서버는 CNC 측에서 동작한다. OPC-UA 클라이언트 모듈로부터 요청받은 데이터는 OPC-UA 서버에서 제공되며, IIoT 엣지 디바이스는 관리 대상이 되는 생산설비로부터 각종 데이터를 수집하여 서버로 전송한다. FOCAS는 FANUC사의 NC(Numeric Controller)에서 데이터를 수집하기 위한 표준 프로토콜이다.

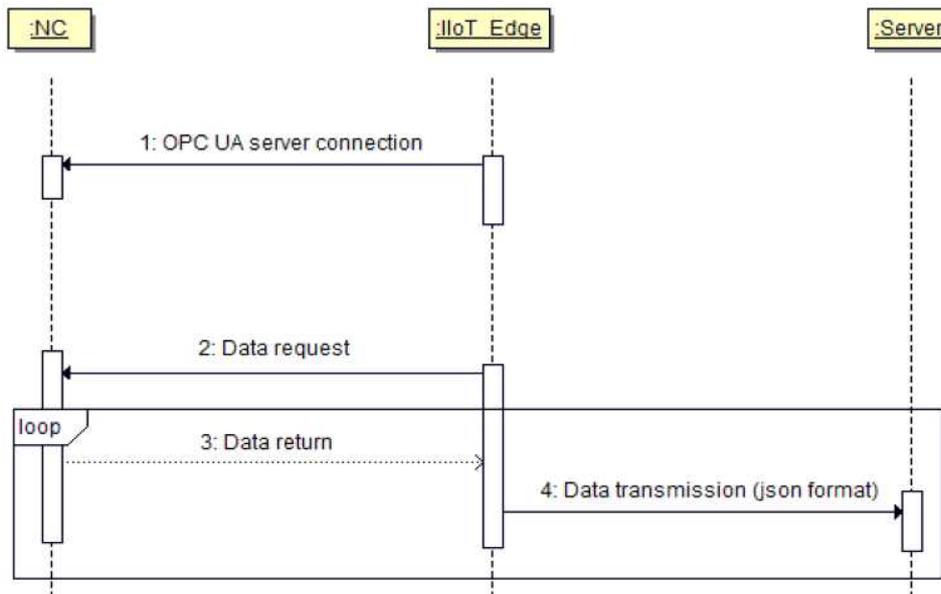


[그림 2-19] 시스템 구성도

그림 2-20, 21은 IIoT 엣지 디바이스의 동작을 시퀀스 다이어그램으로 표현하였다. IIoT 엣지 디바이스에서 CNC로 데이터에 대해 요청을 하면 미리 정해진 주기마다 CNC는 IIoT 엣지 디바이스로 CNC 데이터를 전송한다. 데이터에 따라 IIoT 엣지 디바이스는 CNC로부터 전송받은 데이터를 즉시 서버로 전송하기도 하지만, IIoT 엣지 디바이스에서 데이터를 특정 주기동안 모았다가 서버로 전송하기도 한다.



[그림 2-20] FANUC의 시퀀스 다이어그램



[그림 2-21] SIEMENS의 시퀀스 다이어그램

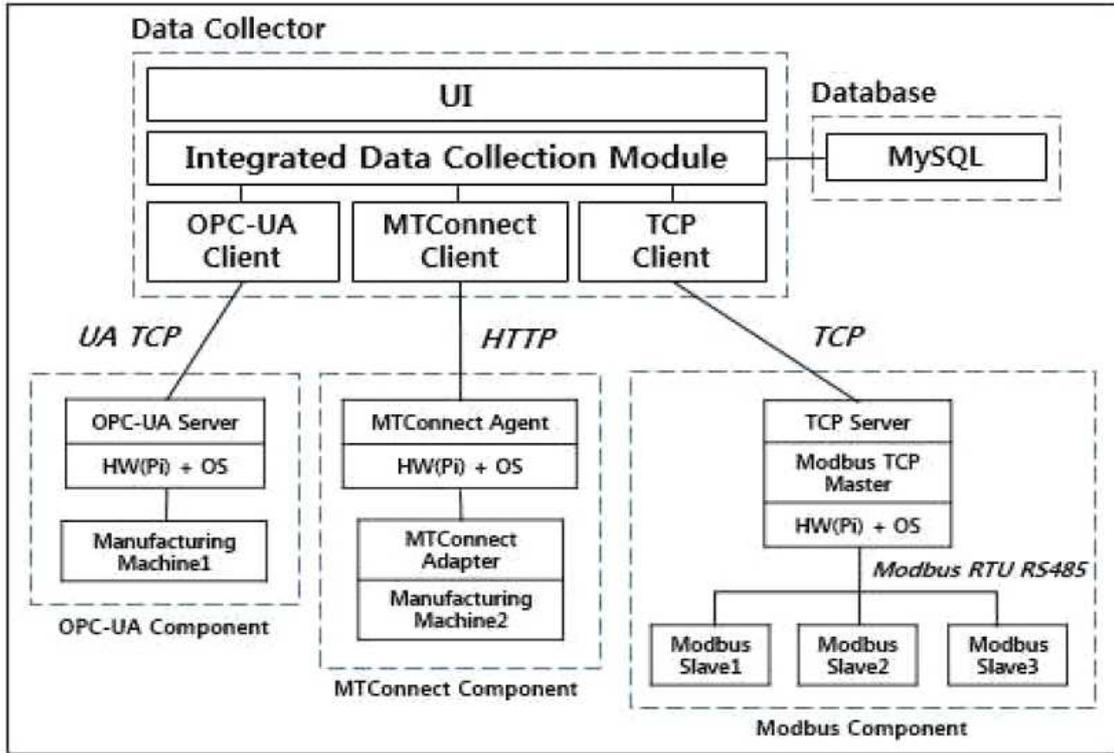
다) 소규모 스마트팩토리를 위한 저비용 통합 데이터 관리 시스템 구현 사례에 대한 조사<sup>42)</sup>

본 절에서는 실제 제조 설비와 데이터 수집 시스템 간의 연결에 많이 쓰이는 산업용 통신 프로토콜들의 특징을 정리한다. OPC-UA는 데이터 교환을 위해 안전하고 신뢰성 있는 상호 운용성 표준으로, 독립적인 플랫폼으로 구성할 수 있으며 인터넷 표준과 IP 기반의 프로토콜을 채택하고 있다. 오픈 소스를 활용하여 클라이언트와 서버 개발이 가능하다. 또한 OPC-UA 클라이언트와 다른 상용, 오픈 소스와 연결할 수 있다는 특징이 있다. MTConnect는 TCP/IP 기반의 표준으로, 설비 데이터를 제공하기 위해 확장 가능한 경량 프로토콜이다. 데이터를 요청하는 클라이언트와 데이터를 제공하는 서버인 에이전트로 구성되어 있다. XML 형태로 Agent에 데이터를 요청하면 응답받는다. Modbus은 산업용 필드버스에서 다수로 사용되는 통신 프로토콜이다. Modbus TCP는 상위 장치와 하위 장치 간의 TCP/IP 네트워크상에서 인터페이스하는 통신방법이다. Modbus TCP에는 슬레이브의 메모리에 마스터가 접근해서 데이터값을 읽어온다. 생산 설비와 데이터 수집 시스템에는 많이 사용되는 산업용 통신 프로토콜인 OPC-UA, MTConnect, Modbus를 이용했다. 각 프로토콜의 데이터 수집 서버는 저비용으로 시스템을 구축하기 위해서 라즈베리파이 상에 구현했다. LDMS는 Data Collector, 데이터베이스, 특정 프로토콜을 따르는 컴포넌트로 구성되어 있다. Data Collector는 Integrated Data Collection Module(이하 IDC 모듈)과 수집데이터를 확인할 수 있는 UI로 구성되어 있어 Python으로 구현하였다. IDC 모듈 첫 번째 기능은 OPC-UA 서버, MTConnect 에이전트, TCP 서버로부터 클라이언트를 통해 각각의 프로토콜에 맞게 데이터를 수집하는 것이다.

그림 2-22와 같이 IDC 모듈과 OPC-UA 서버는 UA TCP, MTConnect 에이전트와는 HTTP, TCP 서버와는 TCP를 통해 통신한다. 제조 설비, I/O 장치들과는 EtherCAT, SHDR, Serial 등의 방식으로 에이전트를 통해 각 프로토콜의 서버와 통신한다.

---

42) 정성일, 전인걸. (2018). 소규모 스마트팩토리를 위한 저비용 통합 데이터 관리 시스템 구현. 정보통신산업진흥원, No. S0608-18-1011.



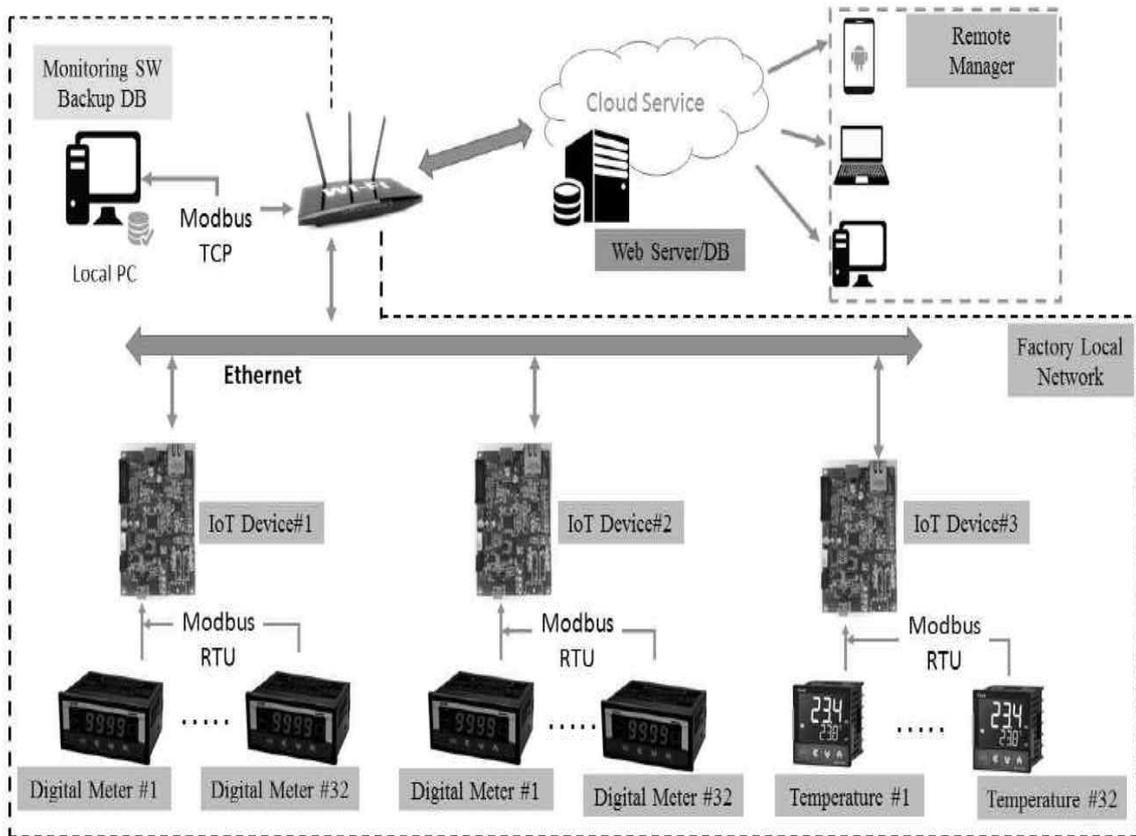
[그림 2-22] 저비용 통합 데이터 관리 시스템(LDMS) 아키텍처

라) IoT디바이스를 이용한 클라우드 플랫폼 기반의 실시간 공장  
모니터링 시스템 개발 사례에 대한 조사<sup>43)</sup>

제조업 분야의 실시간 모니터링 시스템을 위한 분산형 구조 기반의 임베디드 네트워킹 시스템을 제안한다. 온도센서, 전류센서 등과 같은 파라미터는 실시간 제어 시스템의 설계는 DC Power Supply의 가상 제조·공정장비로 발생하는 실제 센서데이터를 디지털 미터로 수집해서 클라우드 서버로 전송한다. 원격 모니터링이 가능한 클라우드 플랫폼 기반으로 임베디드 네트워킹을 구현하였다. 데이터 수집에 필요한 노드는 IoT 디바이스를 사용하고, 센서와 노드 간의 데이터 통신에는 RS485를 표준으로 하는 Modbus RTU를 이용한다. IP기반의 Modbus TCP를 이용하여 노드와 클라우드 서버 간의 데이터 전송한다. 그림 2-23은 개발한 임베디드 IoT 디바이스와 클라우드 간의 분산형 모니터링 시스템의 전체 구조로 나타낸다. 본 시스템에서 머신 또는 디바이스와 클라우드 간

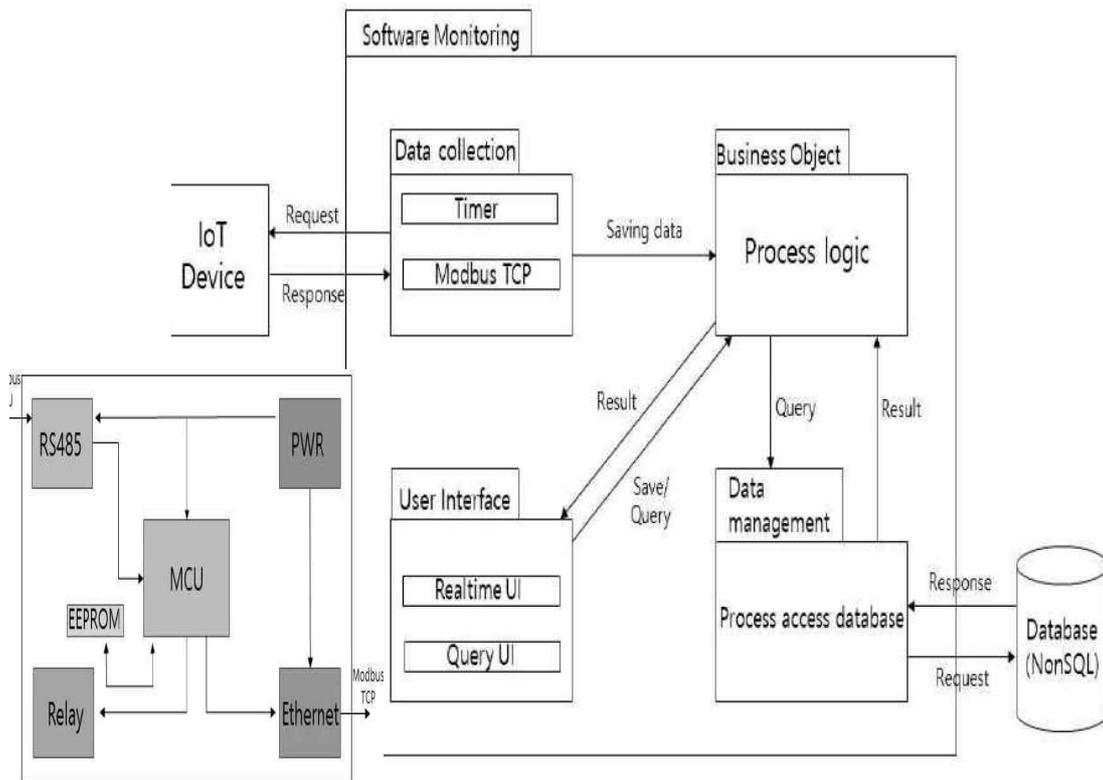
43) 박건수, 트란트렝틴, 당반치옌, 길기중, 신용빈, 최재원, 김종욱. (2018). IoT 디바이스를 이용한 클라우드 플랫폼 기반의 실시간 공장 모니터링 시스템 개발. 대한임베디드공학회, 제 13권 제 1호. pp. 25-32.

의 데이터 통신을 위해서 데이터 패킷을 변환하기 위해 컨버터 역할을 수행하는 IoT 디바이스를 개발하였다. IoT 디바이스의 MCU는 ARM Cortex-M3 STM32F10x를 사용하였다. 또한, 데이터통신의 USART 인터페이스 및 통신을 위해 여러 주변 장치들로 구성되어 있다. 그림 2-23에서 보이는 것처럼 IoT 디바이스는 RS485 표준의 Modbus RTU 전송방식으로 디지털 미터에서 데이터를 수집한다. 또한 이더넷 통신을 위해 데이터 패킷을 Modbus RTU에서 변환해 Modbus TCP로의 전송기능을 가지고 있다.



[그림 2-23] 실시간 모니터링 시스템 구성도 안

그림 2-24는 로컬 PC 상의 데이터 통신 시스템의 구조도이다. IoT 디바이스의 펌웨어로 제공된 데이터를 실시간으로 처리된 후 로컬 DB가 저장된 클라우드 서버로 전송된다. 또한 네트워크에서 IoT 디바이스에서 Modbus TCP를 통해 데이터가 수집되도록 설계되었다. 소프트웨어의 데이터 형식에 따라 DB 관리시스템은 데이터가 저장되고 쿼리 속도가 빠른 Non-SQL기반으로서의 몽고DB를 활용하였다.



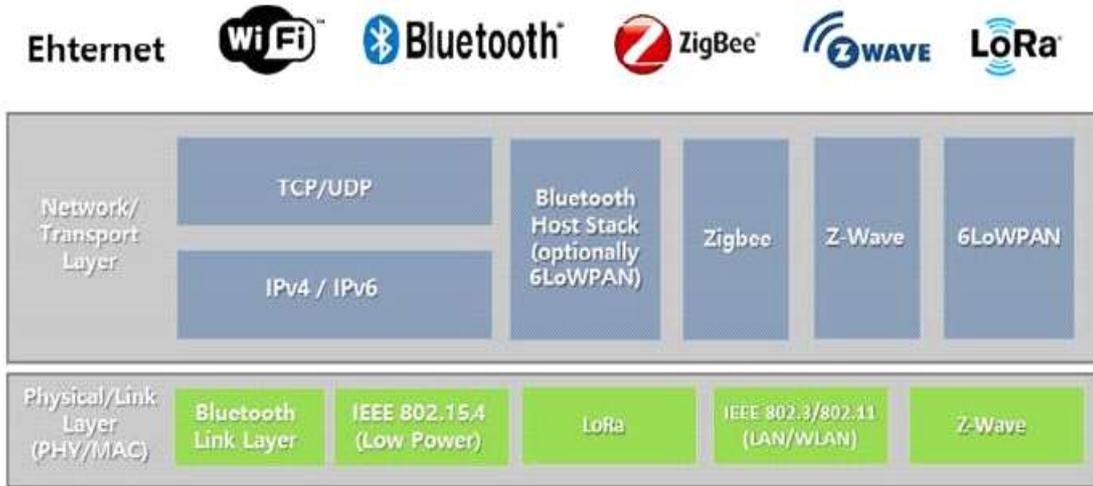
[그림 2-24] IoT 디바이스 구성도 및 PC 실행 소프트웨어의 전체 구조

마) IP기반 이기종 IoT 네트워크 연동 게이트웨이 사례 조사<sup>44)</sup>

OpenWrt는 무선랜 라우터 및 공유기(게이트웨이)를 위한 비실시간 리눅스 기반의 오픈소스 운영체제이다. 임베디드 보드는 게이트웨이 기능을 지원하기 위해서 제한된 자원(CPU, 메모리 등)을 패키지 관리 시스템인 opkg를 이용하여 자유롭게 필요한 패키지 또는 소프트웨어를 설치하였다. 이더넷, Wi-Fi 인터페이스를 OpenWrt에서 제공하는 패키지 모듈을 활용해서 연동시켰다. 6LoWPAN, IEEE 802.15.4, Bluetooth관련 패키지와 BLE, IEEE 802.15.4용 통신 동글 모듈을 활용해서 인터넷 연결성을 위해 IPv4 to IPv6 방식으로 지원하였다. IP 스택을 지원하지 않는 Z-Wave 및 LoRa와 같은 무선 네트워크 프로토콜은 미들웨어용 펌웨어를 구현하여 non-IP 패킷을 IP 패킷으로 변환할 수 있게 게이트웨이의 프로세서로 구동할 수 있도록 하였다.

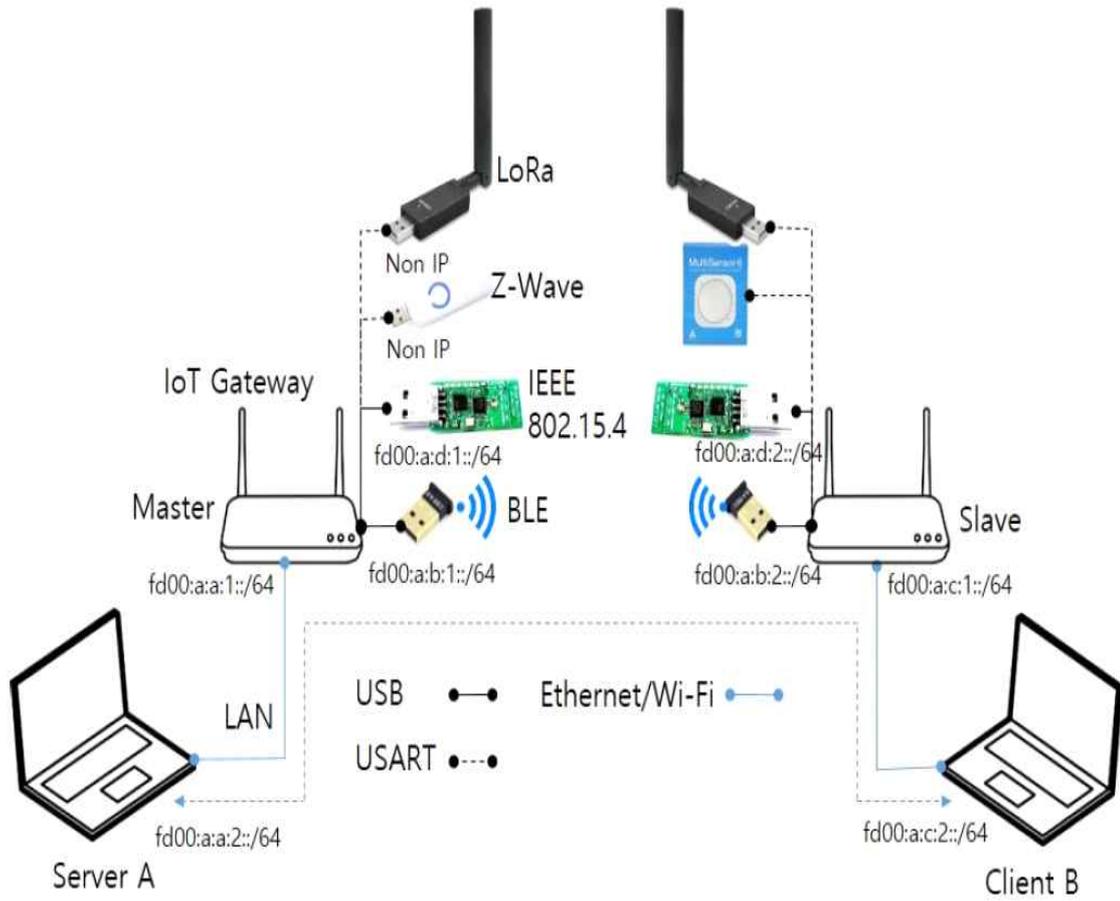
44) 강지현. (2019). IoT 빅데이터 수집을 위한 IP기반 이기종 네트워크 인터페이스 연동 게이트웨이. 한국정보통신학회논문지, vol. 3, no. 2, pp.173~178.

그림 2-25는 본 논문에서 각 계층에서 사용되는 네트워크 스택과 모듈이다.



[그림 2-25] IoT 게이트웨이의 프로토콜 레이어 구성도

그림 2-26과 같이 이기종 네트워크 프로토콜을 IP 기반으로 각 통신 모듈의 디바이스 드라이버, 호환 미들웨어, 프로토콜 스택을 OpenWrt IoT 게이트웨이에 탑재할 수 있게 I/O 드라이버 패키지를 추가하고 USB 허브 장치로 각각의 네트워크의 IoT 게이트웨이와 연결할 수 있도록 하였다. BLE와 IEEE 802.15.4의 경우 USB 타입의 동글 모듈을 사용하였고, IoT 게이트웨이에서 OpenWrt의 Bluetooth, 6LoWPAN패키지를 활용해서 IPv6 기반으로 통신할 수 있도록 하였다. Z-Wave와 LoRa의 경우 네트워크 계층표준이 없어서 AT-command를 지원하는 USB 타입 동글 모듈을 사용하였다. USB-to-Serial 인터페이스를 통해 IoT 게이트웨이로 전달되고, AT-command 내의 메시지를 파싱하여 IP 형태의 패킷으로 변환하여 미들웨어로 거쳐 백그라운드 프로세서로 동작할 수 있게 하였다. IoT게이트웨이는 OpenWrt를 기반으로 하며, Linux Ubuntu 16.04와 OpenWrt LEDE 버전을 사용하였다.



[그림 2-26] 이기종 네트워크 프로토콜 인터페이스 게이트웨이 구성도

바) IoT 통합 관련 선행연구 정리 및 결과

(1) 선행연구 주요 내용

IoT 통합관련 선행연구의 주요 내용은 표 2-9와 같다.

[표 2-9] 선행연구 주요 내역

No	연구 제목	주요 내용	ICT, SW, 알고리즘 기술
1	PLCopen OPC-UA 프레임워크 설계 및 구현	PLCopen OPC-UA 정보모델, OPC-UA 클라이언트 함수 블록 라이브러리로 프레임워크를 구성함	PLCopenOPC-UA, Beremiz(IEC61131-3 표준 오픈소스 통합 개발환경)
2	이기종 공작기계 IIoT 엣지 디바이스 설계, 구현	IIoT 엣지 디바이스가 OPC-UA와 FOCAS 라이브러리를 이용하여 데이터를 수집 및 전송함	OPC-UA, FOCAS(화낙 전용 통신 프로토콜) IIoT 엣지Device, HTTP
3	소규모 저비용 통합 데이터 관리 시스템 도입	IDC 모듈 아래 OPC-UA서버+UATCP, MTConnect+ HTTP, TCP서버+TCP로 LDMS(저비용 데이터 관리시스템) 구성함	OPC-UA, MTConnect Modbus/TCP IDC(통합수집모듈), SHDR, EtherCAT, Serial
4	IoT디바이스 클라우드 플랫폼 기반의 모니터링 시스템 개발	실시간 모니터링 시스템을 위한 분산형 구조 기반의 임베디드 네트워킹 시스템을 구성함	ARM Cortex-M3 RS485표준 Modbus RTU Non-SQL의 몽고DB RESTful API
5	IP기반 이기종 IoT 네트워크 연동 게이트웨이 구축	BLE, Zigbee, Z-Wave, LoRa 프로토콜로 인터페이스하여 송수신 데이터를 IP(Ethernet, Wi-Fi) 네트워크로 변환하는 기능을 지원하는 IoT 게이트 웨이를 OpenWrt 기반하여 설계함	OpenWrt(리눅스 기반 오픈소스 운영체제), IEEE802.15.4, BLE IPv4 to IPv6, 6LoWPAN, LoRa, Zigbee, Z-Wave, Linux Ubuntu 16.04

(2) 시사점 및 결과

표 2-9에서 보면 연구1은 PLC와 연계한 PLCopen OPC-UA 정보모델을 활용하였고, 연구 2는 이기종 공작기계와 연결하기 위해 IIoT 엣지 Device를 구현하여 OPC-UA 프로토콜을 사용하였다. 연구 3은 IDC(통합수집모듈) 아래 다른 통신프로토콜과 연계하기 위해 최신형 OPC-UA 오픈소스 표준토콜을 사용하고 있다. 최근 스마트공장의 통신네트워크의 프로토콜 사용은 OPC-UA와 결합하여 IoT 통합과 플랫폼 수준으로 진행하고 있다.

### 제 3 절 기계학습 및 AI 학습 연구

#### 1) 기계학습 및 AI 학습 관련 연구

인공지능 기술은 지식을 확보하려는 알고리즘을 연구하는 학습지능, 시각, 청각, 언어들과 같은 한 종류의 입력을 하고 지식을 확보하는 단일지능과 여러 형태의 입력을 통합하여 이해 판단하는 복합지능기술을 포함한다.<sup>45)</sup>

[표 2-10] 인공지능과 빅데이터 기술 분류표

중분류	소분류	세분류	요소기술
인공지능	학습지능	머신러닝	베이지안학습, 인공신경망, 딥러닝, 강화학습, 앙상블 러닝, 판단근거 설명
		추론/지식표현	추론, 지식표현 및 온톨로지, 지식처리
	단일지능	언어지능	언어분석, 의미이해, 대화 이해 및 생성, 자동 통역·번역, 질의응답(Q/A), 텍스트 요약 생성
		시각지능	영상처리및패턴인식, 객체인식, 객체탐지, 행동이해, 장소/장면이해, 비디오 분석/예측/요약, 시공간 영상 이해
		청각지능	음성분석, 음성인식, 화자인식/적용, 음성합성, 오디오색인 및 검색, 잡음처리 및 음원분리, 음향인식
	복합지능	행동/소셜지능	공간지능, 운동지능, 소셜지능, 협업지능
		상황/감정이해	감정 이해, 사용자 의도 이해, 뇌신호인지, 센서 데이터 이해, 오감 인지, 다중 상황 판단
		지능형 에이전트	에이전트 플랫폼, 에이전트 기술, 게임, 모방창작 지능
		범용인공지능	상식 학습, 범용 문제해결, 평생 학습, 도덕-윤리-법 지능
	빅데이터	빅데이터 처리, 유통	빅데이터 수집·유통기술
빅데이터 저장, 처리, 관리 기술			실시간 스트림 처리, 데이터 처리 및 관리, 데이터 라이프 사이클, 가상데이터 맵 관리, 데이터 처리 프레임워크
빅데이터 분석, 활용		빅데이터 분석, 예측 기술	심층분석, 실시간분석, 그래프분석, 예지분석, 예측분석, 지시적 분석, 인지분석, 시공간 복합분석, 분석모델링, 시뮬레이션, 데이터 기계학습/딥러닝, 이상치 검출
		빅데이터 활용, 시각화 기술	빅데이터 서비스, 빅데이터 응용, 사회변화 예측, 데이터 시각화, 분석 시각화, 대쉬보드

출처: ICT R&D 기술로드맵 2030-인공지능, 빅데이터 분야\_IITP\_ 2018.9

빅데이터 플랫폼기술과 연계한 인공지능과 빅데이터 기술 분류는 표 2-10

45) 권순선. (2020). 인공지능과 빅데이터 기술동향. TTA journal Vol.187, pp. 40~41.

과 같다. AI 해외기술동향을 살펴보면 구글의 TensorFlow, 마이크로소프트의 닷넷 기반의 Accord.NET, 바이두의 PaddlePaddle 같은 글로벌 플랫폼 업체는 다양한 인터페이스를 제한 없이 공개된 플랫폼에 제공하여, 플랫폼 내부의 복잡 다단한 구조에 대한 이해 없이도 이용할 수 있도록 오픈 플랫폼으로 개발을 진행하고 있다.

#### 가) AI 알고리즘 기술

AI 알고리즘 기술은 머신러닝과 학습할 수 있도록 다양하게 개발하면서 진보하고 있다. 크게 알고리즘을 지도학습(Supervised Learning), 비지도학습(Unsupervised Learning), 준지도학습(Semi Supervised Learning), 강화학습(Reinforcement Learning)으로 나누어 살펴보면 다음과 같다.

지도학습은 학습 데이터마다 레이블을 가지고 있으며 주어진 입력과 출력 쌍들을 매핑해주는 함수를 학습하는 것이다. 크게 분류, 회귀, 예측으로 나눌 수 있고 상세 알고리즘은 k-최근접 (k-Nearest Neighbors), 선형 회귀(Linear Regression, LR), 로지스틱 회귀(Logistic Regression), 서포트 벡터 머신(Support Vector Machines, SVM), 의사결정나무(Decision Tree)와 랜덤 포레스트(Random Forests, RF), 그래디언트 부스팅(Gradient Boosting, GB), XGBoost(Extreme Gradient Boosting, XGB). 신경망(Neural networks) 등이 있다.

비지도학습은 학습데이터가 레이블을 가지고 있지 않으며 입력만 있고 출력은 없는 상태에서 이루어지는 학습이다. 군집, 시각화와 차원축소, 연관규칙 영역으로 나누어 볼 수 있다. 상세 알고리즘은 군집(clustering) 영역에서는 k-평균(k-Means), 계층 군집 분석(Hierarchical Cluster Analysis, HCA), 기댓값 최대화(expectation Maximization) 등이 있다. 시각화(visualization)와 차원 축소(dimensionality reduction) 영역에서는 주성분 분석(Principal Component Analysis, PCA), 커널(kernel PCA), 특잇값 분해(Singular Value Decomposition, SVD), 부분 선형 임베딩 (Locally-Linear Embedding, LLE), 분산 스토캐스틱 근접 임베딩(t-SNE, t-distributed Stochastic Neighbor Embedding)이 있다. 연관 규칙 학습(Association rule learning)영역에서는 어

프라이어리(Apriori), 이클렛(Eclat)이 있다.

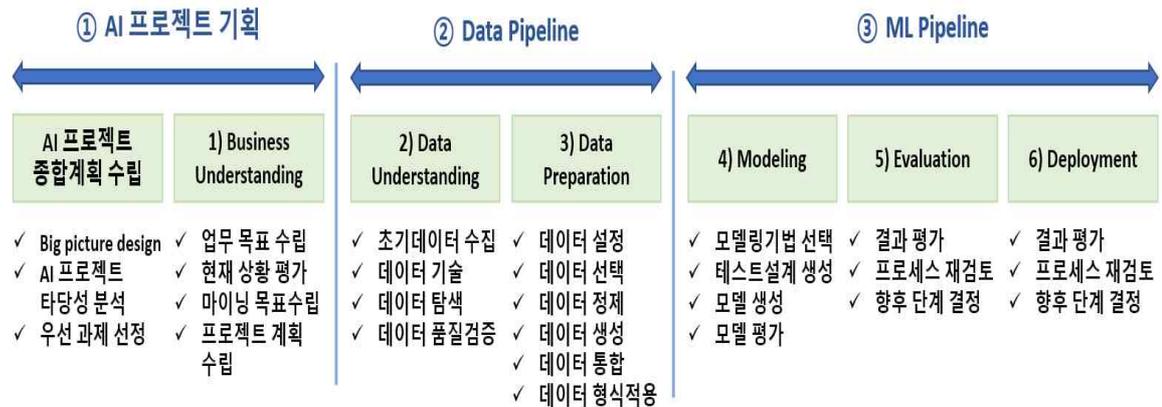
준지도학습은 학습데이터가 한정적이고 소량의 분류(레이블) 데이터를 가지고 있으며 비슷한 데이터끼리 묶는 클러스팅이 주로 사용되며 지도학습에 비해 학습하기 까다로운 편이다. 대부분의 준지도학습 알고리즘은 지도학습과 비지도학습의 조합으로 진행된다.

강화학습은 환경으로부터의 피드백을 기반으로 에이전트의 행동을 분석하고 최적화한다. 에이전트가 시간이 지나면서 가장 큰 보상을 얻기 위해 시나리오라고 부르는 최상의 전략을 스스로 학습한다. 시나리오는 주어진 상황에서 어떤 행동을 선택해야 할지 정의한다. 시행착오(Trial-and-error)와 지연 보상(delayed reward)은 다른 기법과 구별되는 강화 학습만의 특징이다.<sup>46)47)</sup>

#### 나) AI 학습 프로젝트 추진 단계와 절차

AI 학습 프로젝트 추진 단계와 절차를 살펴보면 표 2-11과 같이<sup>48)</sup> 프로젝트 계획 수립(기획), 데이터 이해 및 준비(Data pipeline), 모델 개발 및 배포(Model pipeline) 등 3단계로 구분하여 진행할 수 있다.

[표 2-11] AI학습 프로젝트 추진 단계와 절차



AI 프로젝트 계획 수립은 AI가 변화시킬 광범위한 영역을 종합적으로 고려하여 우선 추진과제를 선정한다.

46) 중소기업기술정보진흥원. (2021). 중소기업 전략기술로드맵, 2021~2023 인공지능. pp 26~27.

47) <https://tensorflow.blog/머신러닝-1장-2장/1-3-머신러닝-시스템의-종류/>

48) 이진휘. (2020). AI기술 동향과 오픈소스. 정보통신산업진흥원 이슈리포트, 2020-제3호. pp. 6.

데이터 파이프라인 단계에서는 데이터 획득 전략을 수립하며, 데이터 정제 (Ingest), 전환, 라벨링, 분석 등 전체 개발 단계 중 가장 많은 자원이 투입되는 데이터 전처리(pre-preparation) 작업이 이루어진다. 비즈니스 과제와 연계한 데이터 과학 및 엔지니어링 과제를 설정하며, 데이터 수집 방법과 획득모델을 고안한다. 또 데이터 접근방법, 대상, 유형, 포맷, 범위 등을 설정하고 데이터의 확보주기와 생성원천, 그리고 데이터 수집 검토기준을 마련한다. 데이터의 전처리 시에는 데이터 수집·저장, 라벨링, 정제, 전환 등 세부 작업을 수행하며, 일반적으로 데이터 분석, 관리 SW를 사용한다.

모델 파이프라인 단계는 처리된 데이터를 이용하여 모델을 개발하고 테스트 하여 운영 시스템에 사용되도록 배포하는 단계이다. 모델 개발은 학습 데이터와 알고리즘을 이용하여 다항식 또는 가중치와 같은 모델 매개변수를 변화하여 성능이 가장 좋은 모델을 찾는 과정이다. 모델 테스트로 성능을 측정하고 배포한 후 모니터링을 통해서 개선한다<sup>49)</sup>.

## 2) 기계학습 및 AI 학습 선행연구 및 결과

### 가) 기계학습을 이용한 CNC 공구 마모도 예측 연구에 대한 조사<sup>50)</sup>

지도학습 방법론을 이용하여 CNC 절삭 공구 마모도에 영향을 미치는 요인을 분석하고자 한다. 본 연구에서 방법론은 XGBoost와 다양한 커널 기법이 가능한 SVM, 전처리 과정과 배깅 방법으로 과적합 문제에 적합한 RF를 사용하였다. 본 연구들에서는 표 2-12와 같이 다양한 특성을 활용하여 X, Y, Z축의 회전력, 축력, 전압 등의 48가지 특성에 대한 분석을 통해 공구 마모도를 예측하였다. 18개의 시험을 독립적으로 수행하였다. 본 연구에서는 수집된 데이터를 사용하여 SVM의 경우 RBF(Radial Basis Function) 커널을 이용해서 모형을 구축하였다. 학습데이터를 모형 학습함으로써 최적의 모형을 선정하였다. 커널 기법이란 선형 분류기로 비선형 관계의 데이터를 차원을 고차원 공간으로 사영하여 모형에 적합으로 만드는 것이다. XGB 모형은 분류와 회귀로 구분되는데, 분류 방법으로 분석하였다.

---

49) 이진휘. 전계리포트, pp 6~10.

50) 이강배, 박성호, 성상하, 박도명. (2019). 기계학습 기법을 이용한 CNC 공구 마모도 예측에 관한 연구. Journal of the Korea Convergence Society, Vol. 10, No. 11, pp. 15-21.

XGB에서 중요한 초모수로서의 트리의 깊이가 모델의 복잡도를 결정한다. 과적합의 위험도 증가는 모델의 복잡도가 커지면서 트리의 깊이가 커지는 것이다. 과적합 정도를 다른 모형들과 비교 및 교차 검증을 통해 확인할 수 있다. 가지 수에 따라 XGB는 정확도에 차이가 크게 남을 알 수 있다. RF는 다수의 트리에서 평균 예측치를 출력함으로써 전체 데이터의 분산을 줄이는 것이 특징이다. 무작위로 구성된 트리 간의 상관관계를 최소화하면서 데이터의 크기나 분포의 영향을 적게 받는다. RF는 전처리가 단순하지만, 좋은 성능을 보인다.

[표 2-12] 공구마모도 특성 데이터 설명

no	Item	Explanation
1	X1_ActualPosition	Actual x position of the part (mm)
2	X1_ActualVelocity	Actual x speed of parts (mm / s)
3	X1_ActualAcceleration	Actual x acceleration of the part (mm / s / s)
4	X1_CommandPosition	X position of the part (mm)
5	X1_CommandVelocity	X part speed (mm / s)
6	X1_CommandAcceleration	X part acceleration (mm / s / s)
7	X1_CurrentFeedback	Current (A)
8	X1_DCBusVoltage	Voltage (V)
9	X1_OutputCurrent	Current (A)
10	X1_OutputVoltage	Voltage (V)
11	X1_OutputPower	Power (kW)
12	Y1_ActualPosition	Actual Y position of the part (mm)
13	Y1_ActualVelocity	Actual Y speed of parts (mm / s)
14	Y1_ActualAcceleration	Actual Y acceleration of the part (mm / s / s)
15	Y1_CommandPosition	Y position of the part (mm)
16	Y1_CommandVelocity	Y part speed (mm / s)
17	Y1_CommandAcceleration	Y part acceleration (mm / s / s)
18	Y1_CurrentFeedback	Current (A)
19	Y1_DCBusVoltage	Voltage (V)
...	...	...
23	Z1_ActualPosition	Actual Z position of the part (mm)
24	Z1_ActualVelocity	Actual Z speed of parts (mm / s)
...	...	...
48	Machining_Process	Current processing stage

분석 결과, SVM의 경우 특성의 수가 많고, 데이터의 분포가 일정하지 않은 문제로 인해 다른 모형에 비해 성능이 떨어졌다. XGB와 RF는 SVM에 비해 빠른 분석 속도와 정확도를 나타내었다. XGB와 RF는 덜 복잡한 전처리 과정이

장점이고 모형이 우수하다. 표 2-13과 같이 RF를 가장 우수한 모형으로 선정하였다.

기계학습 모형인 SVM, XGB, RF의 예측 성능 결과는 SVM이 가장 낮다. SVM의 경우 사용 특성이 많고, 각 특성의 수치가 달라서 좋지 않은 성능을 보여줬다. RF는 분석 시간과 정확도가 가장 우수한 성능을 보여 최종 모형으로 선정하였다.

[표 2-13] 공구마모도 예측성능 분석 결과

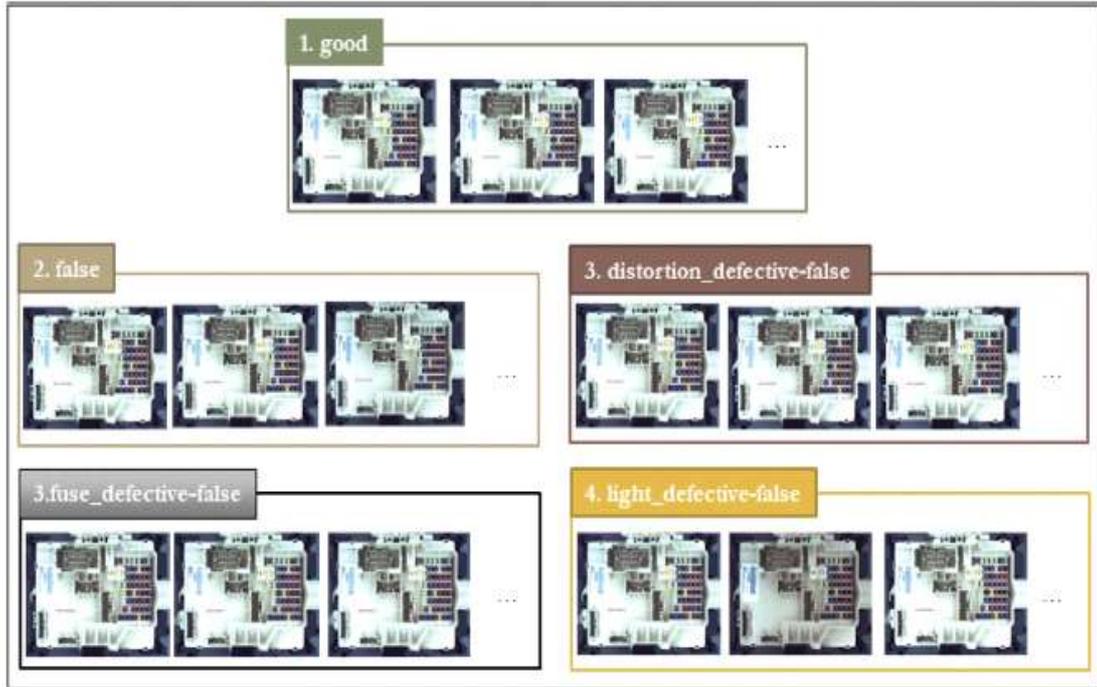
Division	Time(s)	Accuracy	Hyper Parameter
SVM	165	0.629	5
XGB	14	0.993	500
RF	1.29	0.993	50

나) CNN을 이용한 제조 공정데이터 분류 적용 연구에 대한 조사<sup>51)</sup>

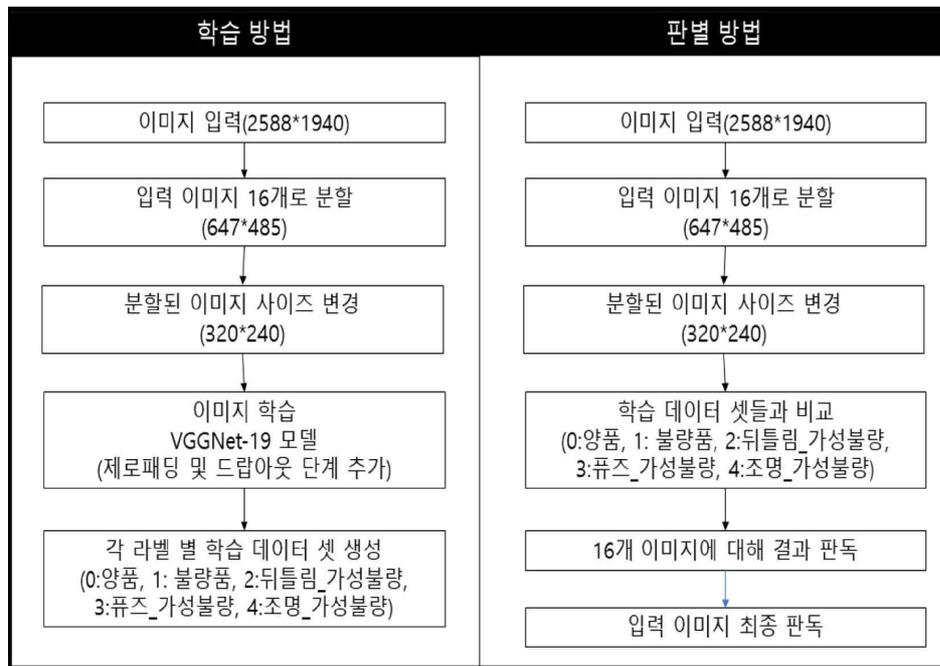
스마트팩토리 공정에서 나오는 이미지를 분석하기 위한 기법에 관해 설명한다. 제안하고자 하는 기법은 기존의 이미지에서 불량품과 양품만 구분하던 방식을 딥러닝을 이용하여 구별하고 불량품 중 양품인데 불량으로 판정한 가성불량품을 불량 원인에 따라 분류하는 방법을 제안하며, 불량 원인은 뒤틀림(왜곡), 퓨즈, 조명에 따라 라벨을 분류한다.

가성불량 판별 이미지 1,052개 중에 뒤틀림, 퓨즈, 조명에 대한 이미지들을 추출해서 뒤틀림 283개, 퓨즈 288개, 조명 304개 총 875개를 사용한다. 앞에 언급한 공정의 파생 이미지는 2588\*1940의 크기이다. 학습 이미지 셋은 그림 2-27과 같다. 이미지를 한 번에 학습시킬 수도 있지만, 너무 큰 이미지이기 때문에 정밀성을 고려하여 이미지를 16개로 분할하여 학습시키도록 한다. 이미지 분류는 16개로 분할하여 1품, 불량품, 가성불량(뒤틀림, 퓨즈, 조명)으로 결정한다.

51) 류가에, 류관희. (2018). 영상 데이터기반의 CNN을 이용한 제조 공정데이터 분류 적용에 대한 연구. Journal of Information Technology and Architecture, Vol. 15, No. 3, pp. 337~343.



[그림 2-27] 훈련용 이미지 셋



[그림 2-28] Proposed Method

제안한 방법의 전체 구성도는 그림 2-28과 같다. 2588\*1940의 제조 영상을 우선 이미지를 분할한다. 분할의 이유는 큰 이미지의 학습은 속도가 지연되

고 하드웨어가 낮은 사양은 학습이 되지 않는 경우가 있다. 이미지를 작은 사이즈로 줄여서 학습하는 방법도 있지만 이미지의 크기를 줄일 경우 제대로 된 특징을 추출하지 못할 수 있다는 위험성이 있다. 제조 공정에서는 정확도를 중요시하기 때문에, 따라서 2588\*1940 크기의 이미지를 분할해 총 16개의 647\*485의 이미지를 활용한다. 분할된 16개 이미지는 노이즈 및 조명에 강한 HSV 형식으로 변환한 후 320\*240 사이즈로 변경하는 전처리를 수행한다. 이렇게 전처리된 이미지는, 간단한 구조와 많은 특징을 빠르게 학습할 수 있는 장점을 가진 CNN을 활용한 VGGNet 모델로 학습한다. VGG-19 모델을 활용해서 컨볼루션 레이어 12번, 맥스풀링 레이어 4단계를 이용하여 구성하고 Dropout 단계를 추가하였다. 또한 제안 방법에 사용하는 이미지는 2588\*1940 크기의 이미지를 분할한 것으로 이미지 특징의 손실을 최소화하기 위해 컨볼루션 실행 시 제로패딩 기법을 사용한다.

실험결과는 화상검사 이미지 3,725개 중 양품 3,143개, 불량품 113개, 가성 불량 469(뒤틀림: 106, 퓨즈: 12, 조명: 288)개로 그림 36과 같이 98%의 높은 결과를 도출되었다, 분할된 16개 이미지는 1초라는 이른 시간에 분류가 완료되고 정확도 결과는 가성불량인 뒤틀림, 퓨즈 및 조명의 결과가 각각 91%, 95%와 89%의 정확도로 가성불량의 이미지들의 원인을 찾아 쉽게 분류해 가성불량을 줄임으로써 생산성 향상과 작업자가 가성불량 물품의 육안검사를 대폭 낮출 수 있었다.

#### 다) 센서와 머신러닝을 활용한 불량예측 방안 연구에 대한 조사<sup>52)</sup>

본 연구는 불량을 나타내는 수치 지표를 예측하는 것이기 때문에 수치 예측 기계학습을 사용하였다.

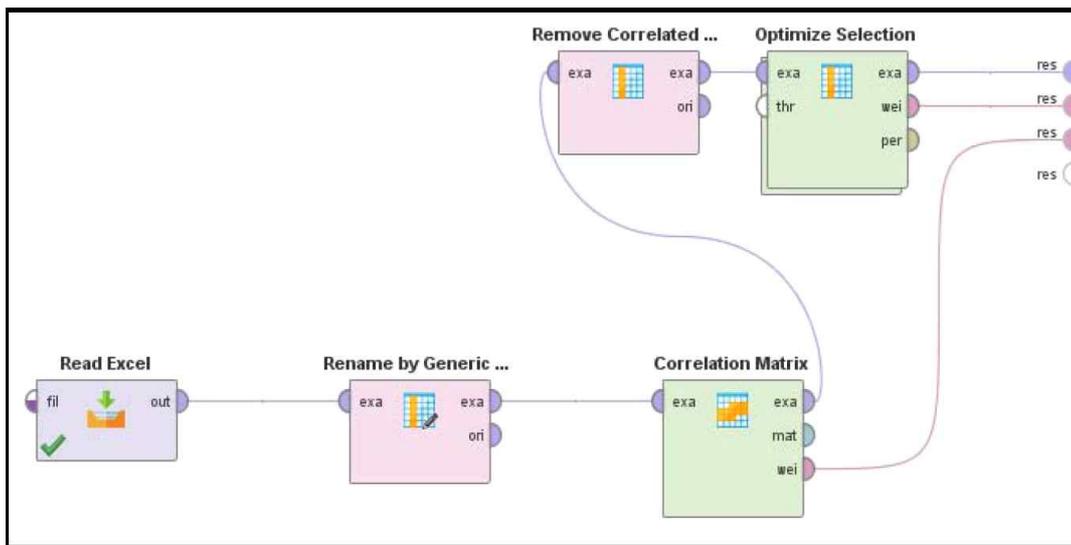
데이터 전처리와 예측 분석은 머신러닝 분석 도구인 래피드마이너를 사용하여 진행하였다. 표 2-14와 같이 제조공정 제품 무게에 큰 영향을 미치는 생산 설비 16개 부위에 센서를 부착하여 데이터를 수집하였고, 무게데이터의 변동계수 ( $CV=6/xbar$ )는 5 이상인 경우 이상치로 간주하였다. 그림 2-29는 래피드마이너를 활용한 데이터 모델링 전처리 과정을 보여준다. 연관 속성의 제거 기

52) 한무명초, 이충권, 김양석. (2019). 제조 공정에서 센서와 머신러닝을 활용한 불량예측 방안 연구. *Entrue Journal of Information Technology*, Vol.17, No.1, pp.89-98.

능을 이용하여 다중공선성의 문제가 발생 할 수 있는 속성을 제거하기 위해 상관계수 값이 '.90' 초과인 속성을 제거하였다. 그 결과, 센서 NO10, NO13, NO15, NO16번이 제거되었다.

[표 2-14] 센서 구성표

순번	센서 유형
NO 1	설비 1의 압력 센서
NO 2	설비 2의 진동 센서
NO 3	설비 3의 스피드 센서
NO 4	설비 1의 흡입 압력 센서
NO 5	설비 4의 진동 센서
NO 6	설비 5의 진동 센서
NO 7	설비 1의 흡입 센서
NO 8	설비 6의 흡입 센서
NO 9	설비 4의 갭(gap) 센서
NO 10	설비 6의 압력 센서
NO 11	설비 6의 갭(gap) 센서
NO 12	설비 7의 압력 센서
NO 13	설비 1의 온도
NO 14	설비 1의 습도
NO 15	설비 2의 온도
NO 16	설비 2의 습도



[그림 2-29] 래피드마이너 활용한 데이터 전처리

본 연구에서는 성과 검증을 위하여 k-NN 알고리즘을 사용하였고 모델링 프로세스를 설계하여 Optimize Parameters에서는 k-NN의 최적 유사도와 k값을 찾기 위해 파라미터를 설정하였다. Optimize Selection을 적용한 결과, 센서

No2, No3, No11, No14번이 제거되었고 최종적으로 일반속성 8개를 CV예측에 사용하였고. 검증 프레임워크는 10-중첩으로 교차 검증을 하였다.

스마트팩토리 공정의 센서에서 수집한 데이터와 공정의 결과물인 제품의 무게 데이터를 수집하여 제품의 품질에 영향을 미치는 요인을 찾는 방법을 제시하고자 하였다. 분석 결과를 종합해 보면 다음 표 2-15, 16과 같다. 첫째, LR, GBT, DL을 이용한 CV 예측 모델의 예측 결과는 DL이 가장 우수한 성과를 보였다. 둘째, 각 모델링 알고리즘에서 중요하게 판단되는 속성은 차이가 있다. NO6는 LR와 DL에서 가장 중요한 속성이지만 GBT의 경우 NO7이 가장 중요한 속성이고, NO6는 세 번째로 중요한 속성이었다.

연구결과표는 아래와 같다.

[표 2-15] 예측모델 성과지표

	LR	GBT	DL
MAE	0.127	0.118	<b>0.114</b>
RMSE	0.168	0.160	<b>0.149</b>
R <sup>2</sup>	0.571	0.611	<b>0.667</b>

[표 2-16] 모델별 속성 중요도

Attribute	LR	GBT	DL
NO1	0.000	0.008	0.124
NO4	0.000	0.183	0.161
NO5	0.000	0.016	0.042
NO6	0.313	0.164	0.285
NO7	0.117	0.218	0.039
NO8	0.000	0.013	0.117
NO9	0.000	0.010	0.001
NO12	0.030	0.035	0.076

라) 실시간 불량품 발생 원인 분석 및 설비 교체주기 예측 사례 조사<sup>53)</sup>

스마트공장 공정에서 사용되는 소모성 부품의 교체주기를 예측하기 위한 LSTM을 활용한 시계열 분석을 사용한다. 최종 불량률을 최소화하기 위한 최적

53) 황승연, 곽경민, 신동진. (2019). 빅데이터 기반 실시간 불량품 발생 원인분석 및 설비교체주기 예측  
The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 19, No. 6, pp.203-212.

의 조건을 찾는 군집 분석을 수행하기 위해 군집 분석 기법으로 DBSCAN(Density Based Spatial Clustering of Applications with Noise)을 사용한다. OLS(Ordinary Least Squares) 회귀 분석을 통해 각 공정에서 생성되는 데이터를 이용하여 불량률을 예측한다.

[표 2-17] 1) 가열온도 2) 주입속도 3) 공기흡입컵 크기감소량

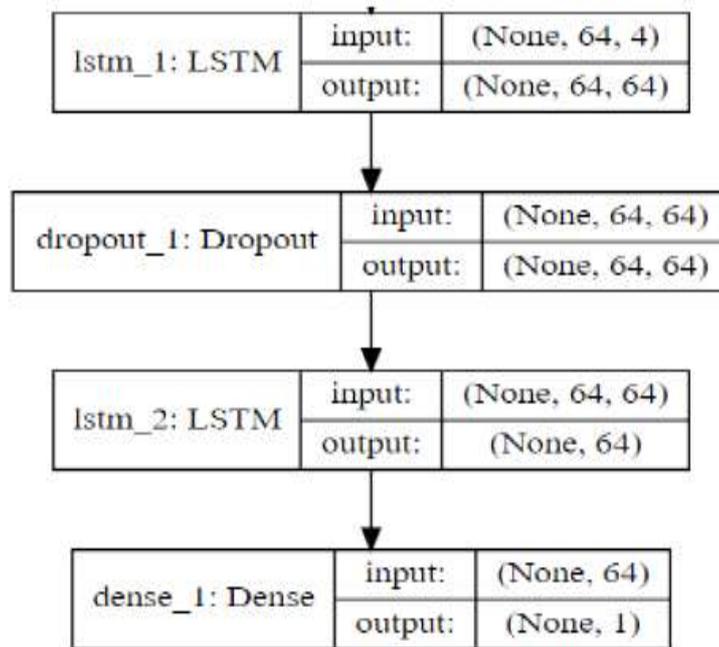
가열온도	제품완성도	주입속도	제품완성도	제품완성도	공기 흡입 컵의 크기 감소량
1000	50%	60~70	50%	100%	0.0001mm씩 감소
1001~1050	70%	71~80	70%	70%	0.0003mm씩 감소
1100~1300	100%	81~100	100%	50%	0.0005mm씩 감소
1301~1350	70%	101~110	70%		
1400	50%	111~120	50%		

완제품을 만들기 위해 A, B, C공정으로 구성된 시나리오를 활용하여 가상 데이터를 생성한다. 본 절에서는 데이터 생성기를 이용하여 생성되는 각 공정의 가상 데이터 생성 방법 및 조건에 대해 알아본다. 그중 A공정에서 대해서 살펴본다. 표 2-17 1), 2), 3)과 같다.

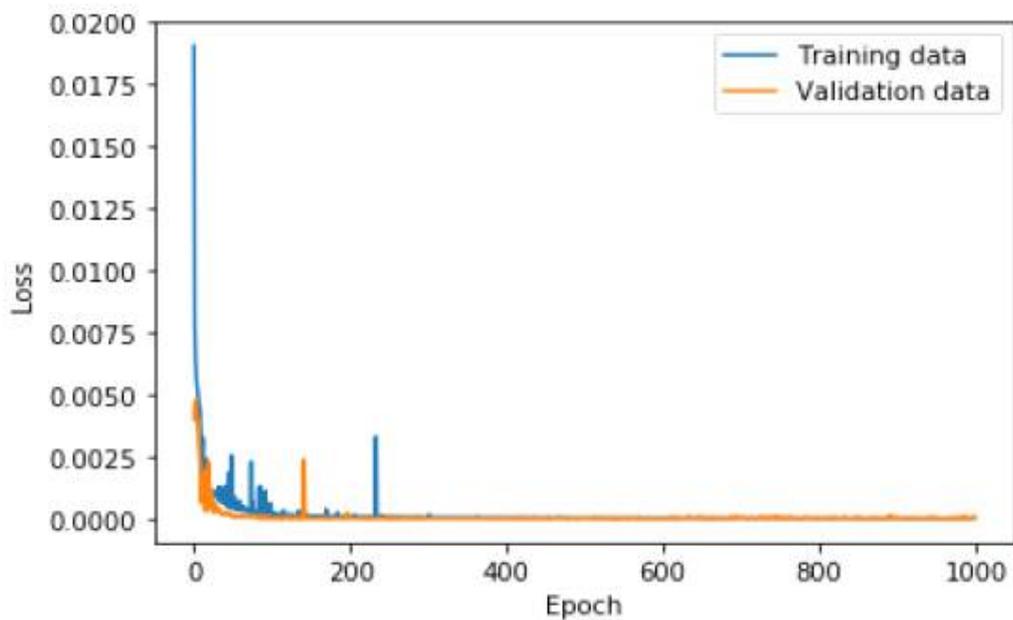
#### (1)시계열분석

시계열 분석을 이용한 교체주기 예측을 하기 위해 다변수, 다층 레이어를 이용한 LSTM을 이용하였다. 사용된 변수는 A공정의 가열온도, 주입속도, 공기흡입컵 제품완성도 4가지이고, 이를 이용하여 시간의 흐름에 따른 공기 흡입 컵의 크기를 예측한다. 손실함수는 평균 제곱근 편차(Mean-squared-error, MSE)를 사용하고, 옵티마이저(Optimizer)는 RMSProp을 사용하였다.

그림 2-30과 같이 구성된 레이어는 윈도 사이즈를 64개로 설정하고 그림 2-31을 통해 학습 손실(Train loss)과 검증 손실(validation loss)의 값이 같이 감소하는 것을 통해 과적합이 되지 않은 것을 알 수 있다. 예측된 결과의 손실이 0.00007로 아주 낮은 것을 확인할 수 있다.



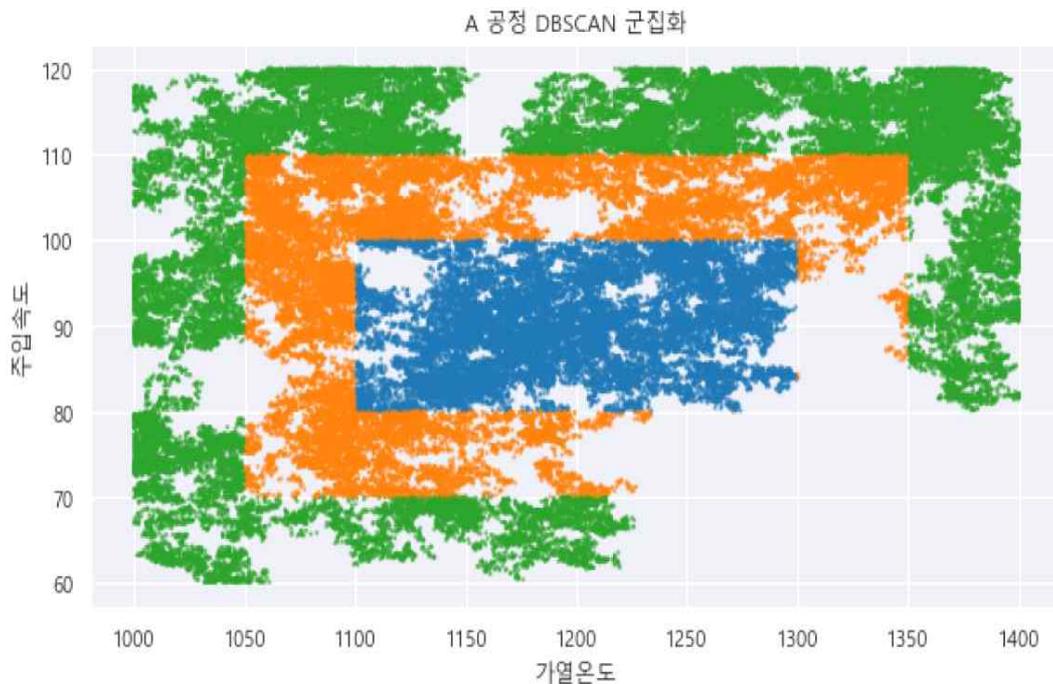
[그림 2-30] LSTM 전체 구성도



[그림 2-31] 전체 학습 손실 값 그래프

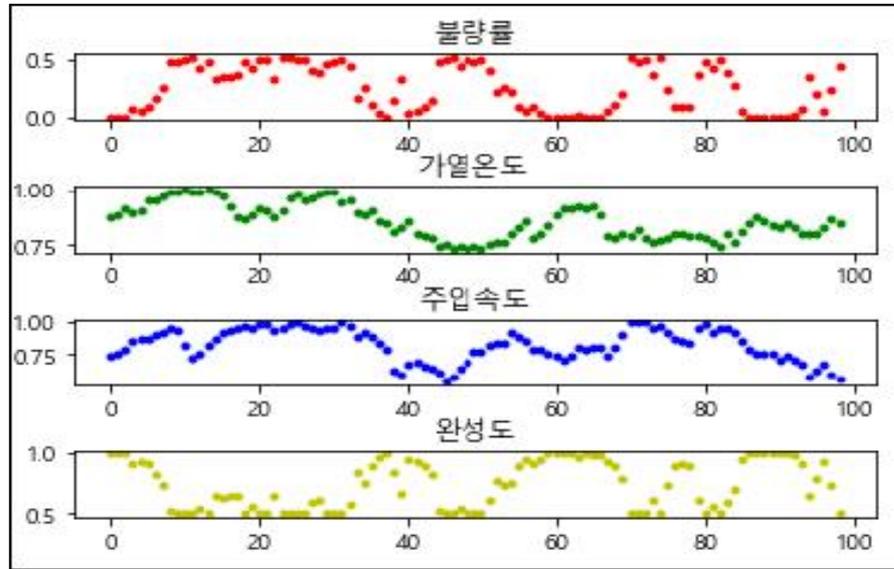
## (2) 군집분석

그림 2-32는 A공정을 DBSCAN 알고리즘을 이용하여 밀도 기반의 군집 분석 결과이다. 군집화를 위해 가열온도, 주입 속도, 제품완성도 3가지의 변수를 사용한다. 3개의 군집이 나타나는데, 각 군집은 제품완성도 별로 군집화되어 있으며, 각 군집은 제품완성도에 따른 불량 확률을 보인다. 이를 이용하여 각각의 군집의 평균 가열온도와 주입 속도를 계산해 최적의 제품 생산환경 분석이 가능하다. 파란색 군집은 100% 양품이 나오는 군집이고, 두 번째로 위치한 주황색 군집은 90% 양품을 갖는 확률이며, 세 번째로 위치한 초록색 군집은 50%의 양품을 갖는 군집이다. 분포도는 A공정의 제품완성도 분포로 50, 90, 100의 3가지 양상이다.



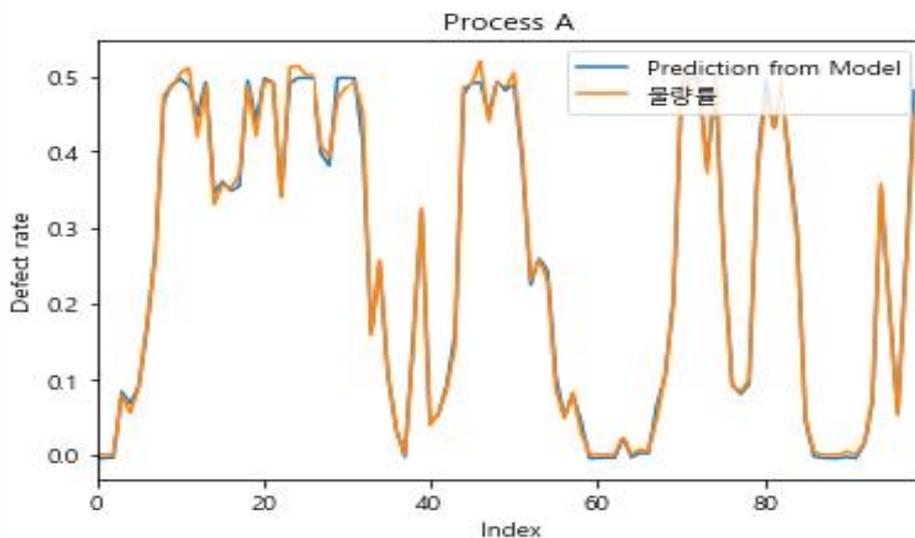
## (3) 각 공정 데이터를 이용한 회귀 분석

그림 2-33의 그래프는 총 10만 개의 데이터 중 불량률, 온도, 습도, 제품완성도를 윈도우 크기(window size)(1000개 지정)마다 평균을 내어 그래프로 나타낸 것이다. 이 중 독립변수는 가열온도, 주입속도, 제품완성도이며, 종속변수는 불량률에 해당한다. 불량률은 윈도우 크기의 데이터에서 불량 여부를 카운트하여, 윈도우 크기로 나눈 평균값이다.



[그림 2-33] window size 별 평균값

그림 2-34는 그림 2-35의 변수들을 기반으로 OLS 회귀분석을 진행하여, 실제 불량률과 예측 불량률을 시각화한 것이고 전체 RMSE 값은 0.0117이다. 그래프의 x축은 데이터의 인덱스를 의미하고, y축은 불량률을 의미한다. 주황색 그래프는 실제 데이터의 불량률을 나타내고, 파란색 그래프는 실제 데이터를 기반으로 불량률을 예측한 것이다. 분석 결과 실제 불량률과 예측한 불량률이 큰 차이가 없는 것을 확인할 수 있다.



RMSE Score : 0.01170

[그림 2-34] 불량 예측 그래프와 RMSE

그림 2-35는 OLS 회귀 분석 결과표이며, 이 중 결정계수(R-squared)는 주어진 데이터를 현재 모형이 얼마나 잘 설명하고 있는지를 나타내는 지표이다. 원도 크기가 1,000개일 때, 1,000개의 데이터의 가열온도, 주입속도의 평균이 1 증가하면 불량률은 각각 0.70, 0.28이 증가하고, 1,000개 데이터 제품완성도의 평균이 1씩 증가한다면, 불량률은 0.80 감소한다. 공정 B와 C도 같은 알고리즘을 적용하였다.

OLS Regression Results						
Dep. Variable:	불량률		R-squared:	0.944		
Model:	OLS		Adj. R-squared:	0.942		
Method:	Least Squares		F-statistic:	540.9		
Date:	Fri, 25 Oct 2019		Prob (F-statistic):	5.55e-60		
Time:	02:26:27		Log-Likelihood:	114.35		
No. Observations:	99		AIC:	-222.7		
Df Residuals:	96		BIC:	-214.9		
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
가열온도	0.7073	0.078	9.095	0.000	0.553	0.862
주입속도	0.2889	0.066	4.357	0.000	0.157	0.421
완성도	-0.8060	0.038	-21.475	0.000	-0.880	-0.731
Omnibus:	12.414	Durbin-Watson:	0.265			
Prob(Omnibus):	0.002	Jarque-Bera (JB):	13.289			
Skew:	0.878	Prob(JB):	0.00130			
Kurtosis:	3.368	Cond. No.	18.4			

[그림 2-35] OLS 회귀 분석 결과표

마) 온디바이스 기반 설비 예측 정비 구축 연구 구현사례<sup>54)</sup>

개방형 사물인터넷(IoT, Internet of Things) 플랫폼을 이용하게 되는데 이 사물인터넷 응용을 개발하는 데 있어 실시간 통신 기술은 필수적인 요소다. 실시간 소통으로서의 인스턴트 메시징 지원에는, MQTT 및 CoAP 프로토콜을 IoT 표준 프로토콜로 사용한다. IoT 기반 계측 센서, 설비 등록 기능, 데이터 구조 설정 기능과 설비 등록 도구의 제작으로 상세 설계 및 DB의 사용, UI/UX 개발로, 화면상으로 손쉽게 외부 시스템과 연결할 수 있도록 장비 등록과 설정 기능으로 구현된다. 자동으로 설비의 문제점을 파악하고, 해당 문제의 발생 원인 및 진전과정, 운전상의 결함 및 위험도 진단 등과 같이 전문가에게 맡겨야 하는 의사결정들을 대신해 현장의 엔지니어에게 전달하거나, 전문가가 더욱 정확한 진단을 내릴 수 있도록 지원한다. 도표 및 그래프 등을 통해 사용자가 정보를 보다 더 쉽게 파악할 수 있도록 가시화 기술을 지원하고, 이에 따라 의사결정이 더욱 빨라질 수 있도록 한다.

기계학습 기반의 알고리즘인 선형회귀(Linear Regression, LR), 서포트 벡터 머신(Support Vector Machine, SVM), KNN(K-Nearest Neighbor), GNB(Gaussian Naive Bayes)의 4개 알고리즘을 통해 학습 및 분류하였다. 스마트팩토리의 설비 예지보전 구축의 목적으로 설비 센싱 빅데이터 처리를 위한 플랫폼을 개발하고 가시화 기술을 개발하였다. 먼저, 실시간으로 센싱 생성 정보인 온도, 소음, 전류 등의 수집을 위해 MQTT 기반 실시간 수집이 가능한 기술과 플럼 기반 센싱 정보 수집모듈을 개발하여 대량의 데이터 수신과 저장할 수 있었다. 실시간으로 생성되는 센싱 빅데이터의 융합으로 설비 예지 보전을 분석하고 가시화 기능과 개발된 통합 플랫폼으로 다양한 분석 기능이 구현되어 통합 플랫폼에 활용될 수 있었다.

---

54) 이용환, & 서진형. (2020). 온-디바이스 엣지 컴퓨터 기반 설비예지보존 솔루션 구축에 관한 연구 (A Study on the Build of Equipment Predictive Maintenance Solutions Based on On-device Edge Computer). 한국컴퓨터정보학회논문지, 25(4), 165-172.

바) 기계학습 및 AI학습 관련 선행연구 정리 및 결과

(1) 선행연구 주요 내용

기계학습 및 AI학습 관련 선행연구의 주요 내용은 표 2-18과 같다.

[표 2-18] 선행연구 주요 내역

No	연구 제목	주요 내용	ICT/SW/알고리즘 기술
1	기계학습 기법을 이용한 CNC 공구 마모도 예측	지도학습 알고리즘 활용하여 CNC 공구 마모도 영향 요인 분석을 함 분석속도/정확도: RF>XGB>SVM	SVM: RBF커널활용 XGBoost:부스팅 방법 RF:쉬운 전처리, 배깅방법
2	CNN 이용한 제조 공정데이터 분류 적용	VGG-19모델, 3*3커널, 레이어 12번, 맥스풀링 4단계, Dropout/제로패딩 기법 추가하여 적용함	CNN을 이용한 VGGNet 모델 활용
3	센서와 머신러닝을 활용한 불량 예측 방안 연구	불량 지표의 예측을 위해 3개 알고리즘과 래피드마이너로 분석함	LR, GBT, DL 래피드마이너 :분석솔루션
4	실시간 불량품 발생원인 분석 및 설비 교체주기 예측	가열온도, 주입속도 제품완성도 변수를 3개 알고리즘으로 불량률 예측을 MapReduce 모델로 분석함	LSTM 시계열 분석 DBSCAN군집분석 기법 OLS 회귀분석 MapReduce 분석모델
5	PM Solutions based on device Edge computer	MQTT기반 실시간 수집모듈 개발, 저장하고, 4개 알고리즘으로 예지보전(PM) 분석 및 시각화함	CoAP, MQTT 프로토콜 KNN, LR, SVM, GNB

(2) 시사점 및 결과

상기 표에서 연속성 데이터를 취급하는 공정에서 기계학습 알고리즘을 살펴보면 불량률 예측에서는 LR, GBT, DL, LSTM DBSCAN, OLS를 사용하였고 공구수명과 설비보전 예측에서는 SVM, XGBoost, RF, KNN, LR, SVM, GNB를 활용하고 있다.

## 제 4 절 선행연구 기술 및 시사점 종합

앞에서 조사 연구한 빅데이터 수집 및 처리, IoT 통합, 기계학습 및 AI학습의 선행연구에서 선행연구 기술과 시사점 내용을 종합하면 표 2-19, 20, 21과 같다.

### 1) 빅데이터 수집 및 처리 선행연구 기술 및 시사점

[표 2-19] 빅데이터 수집 및 처리 선행연구 기술 및 시사점 내역

No	선행연구 제목	ICT, SW, 알고리즘 기술	시사점
1	제조 빅데이터 분석 플랫폼 구축	RDMS+오픈소스+BI 분석도구 데이터브라우저	빅데이터 수집 및 처리 분야에서는 다양한 SW가 사용되고 있으나 라즈베리파이, MariaDB, R프로그래밍 등 오픈소스 SW의 사용이 증가하고 있고 Microsoft Azure, SAP HANA, Google.Net 등 글로벌 플랫폼에서 오픈소스 SW와 함께 사용도 활성화되고 추세이다. 기업에서는 데이터나 애플리케이션도 웹브라우저 개념과 같이 개발자나 사용자의 맞춤형으로 제공되고 있으며 효율적인 DB운영이나 보안 측면을 위해서 2계층에서 3계층 구조로 확산하고 있다.
2	빅데이터 센서 IoT 데이터 처리, 분석	라즈베리/MariaDB, R프로그래밍, Sqoop /HDFS, Hive	
3	빅데이터 기반 실시간 설비데이터 처리	MongoDB(대표적인 NoSQL), R프로그래밍 MTConnect(SNMP)	
4	예지정비를 위한 클라우드 기반 IoT 플랫폼 구현	Microsoft Azure, SAP HANA 플랫폼, MES, ERP, SCM,클라우드개발/프레임워크서비스, 웹 브라우저	
5	.Net기반 애플리케이션 브라우저 설계구현	Google.Net 플랫폼, 3계층 구조, 스마트 클라이언트, 애플리케이션브라우저	

## 2) IoT 통합 관련 선행연구 기술 및 시사점

[표 2-20] IoT 통합 관련 선행연구 기술 및 시사점 내역

No	선행연구 제목	ICT, SW, 알고리즘 기술	시사점
1	PLCopen OPC-UA 프레임워크 설계 및 구현	PLCopen OPC-UA, Beremiz(IEC61131-3 표준 오픈소스 통합 개발환경)	표에서 선행연구1은 PLC에 PLCopen OPC-UA 생성 정보모델과 함수블록 라이브러리로 프레임워크를 적용하였고, 선행연구2는 이기종 공작기계와 연결하기 위해 IIoT 엿지 Device를 구현하였고 선행연구3은 IDC(통합수집모듈) 아래 다른 통신프로토콜과 연계하기 위해 최신형 OPC-UA 오픈소스 표준토콜을 사용하고 있다, 최근 스마트공장의 통신네트워크의 프로토콜은 점차 OPC-UA와 결합하여 IoT 통합과 함께 플랫폼 수준으로 진행하고 있다.
2	이기종 공작기계 IIoT 엿지 디바이스 설계구현	OPC-UA, FOCAS(화낙 전용 통신 프로토콜), IIoT 엿지Device, HTTP	
3	소규모 저비용 통합 데이터 관리 시스템 도입	OPC-UA, MTConnect Modbus/TCP IDC, SHDR, EtherCAT, Serial	
4	IIoT 디바이스 클라우드 모니터링 시스템 개발	ARM Cortex-M3, RS485표준, Modbus RTU, Non-SQL의 몽고DB, RESTful API	
5	IP기반 이기종 IoT 네트워크 연동 게이트웨이 구축	OpenWrt(리눅스 오픈소스) IEEE802.15.4, BLE, LoRa, 6LoWPAN, Zigbee, Z-Wave, IPv4 to IPv6,	

## 3) 기계학습 및 AI학습 관련 선행연구 기술 및 시사점

[표 2-21] 기계학습 및 AI학습 관련 선행연구 기술 및 시사점 내역

No	연구 제목	ICT/SW/알고리즘 기술	시사점
1	기계학습 기법을 이용한 CNC 공구 마모도 예측	SVM: RBF커널활용 XGBoost:부스팅 방법 RF:쉬운 전처리, 배깅방법	연속성 데이터를 취급하는 공정에서 기계학습의 지도학습 알고리즘이 주로 활용하고 있으며 세부적으로 살펴보면 불량률 예측에서는 LR, GBT, DL, LSTM DBSCAN, OLS를 사용하였고, 공구 수명과 설비보전예측에서 SVM, XGBoost, RF, KNN, LR, SVM, GNB를 활용하고 있다.
2	CNN 이용한 공정 데이터 분류 적용	CNN을 이용한 VGGNet 모델 활용	
3	센서와 머신러닝을 활용한 불량예측 방안 연구	LR, GBT, DL 래피드마이너(분석솔루션)	
4	실시간 불량품 발생원인 분석 및 설비 교체주기 예측	LSTM(시계열), DBSCAN(군집분석), OLS(회귀분석), MapReduce(분석모델)	
5	PM Solutions based on device Edge computer	CoAP, MQTT 프로토콜 KNN, LR, SVM, GNB PM(예방보전)	

#### 4) 종합 정리 및 결과

가) 빅데이터 수집 및 처리 선행연구에서는 SW 도구를 다양하게 사용해 왔으나, 글로벌 플랫폼과 함께 오픈소스 SW의 활용이 점차 증가하고 있다는 것을 고려할 필요가 있다. 기업에서는 웹브라우저 개념의 3계층 스마트 애플리케이션 브라우저 개념이 맞춤형으로 받고 있고 확산하고 있다.

나) IoT Integration 관련 선행연구에서는 PLC, 이기종 가공설비, 다른 통신프로토콜과 연결을 위한 생성 프레임워크, IIoT 엣지 Device, IDC(통합수집모듈)를 통해 최신형 OPC-UA 오픈소스 표준토콜을 사용하여 IoT Integration에 활용하고 있으며 최근에는 플랫폼 수준으로 진보하고 있다.

다) 기계학습 및 AI학습 관련 선행연구에서는 지도학습으로 지도학습 알고리즘을 주로 사용하고 있으며 불량률, 공구수명, 설비보전 예측에서는 LR, GBT, DL, LSTM DBSCAN, OLS, SVM, XGBoost, RF, KNN, GNB를 활용하고 있다.

라) 종합적으로 정리하면, 본 논문에서는 표 2-17, 5항의 선행논문과 같이 오픈소스의 Maria와 3계층 구조의 스마트 클라이언트 애플리케이션 브라우저 기능으로 설계하여 엔터프라이즈 솔루션으로 개발하고자 한다.

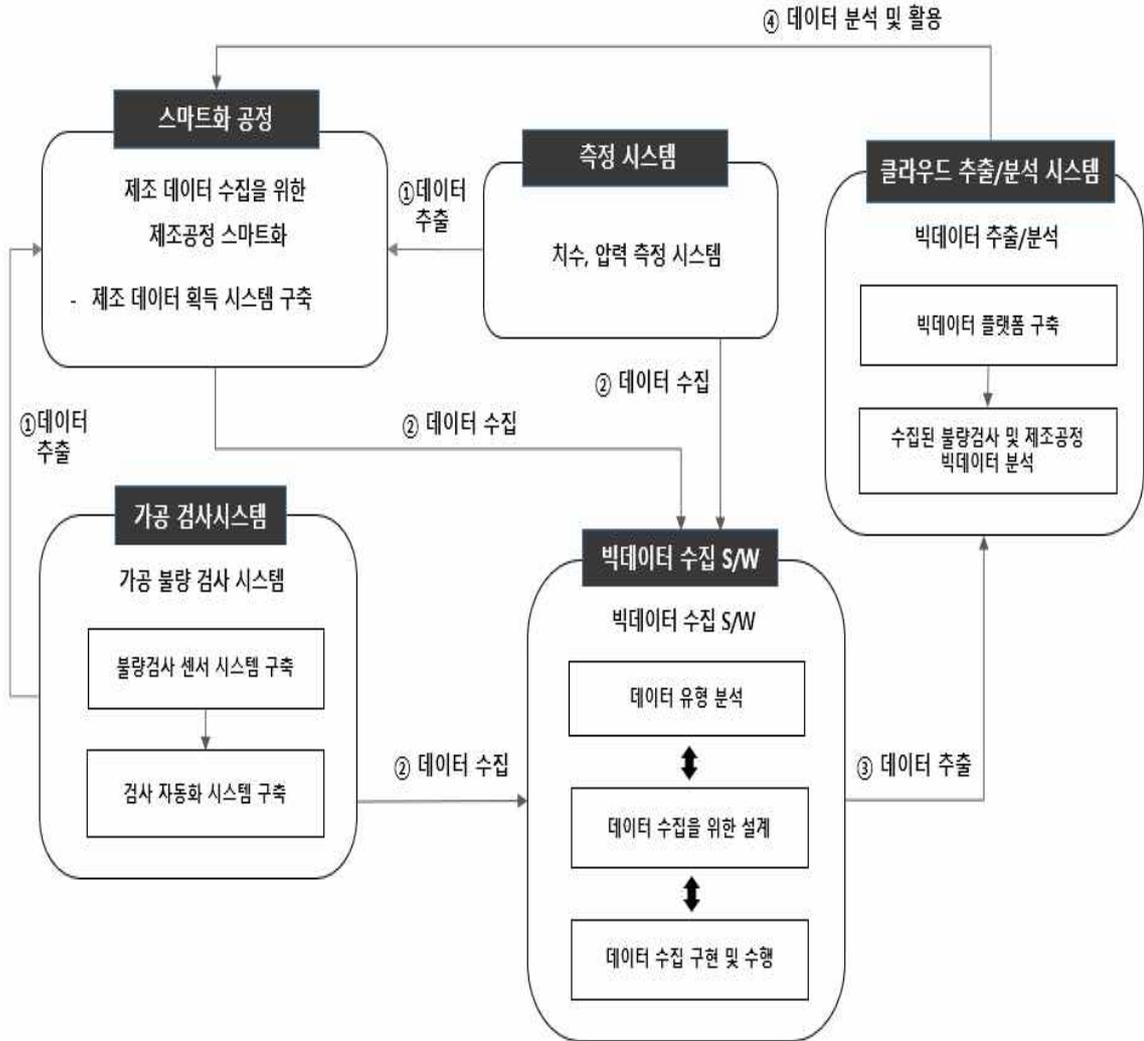
표 2-18, 2항의 선행논문과 같이 표준 통신 프로토콜이자 오픈소스인 OPC-UA로 IIoT 디바이스 Type으로 유무선 통합하여 OPC-UA를 기반한 IoT Integration으로 데이터 수집 및 처리 시스템을 개발하고자 한다.

표 2-19, 1, 3항의 선행논문과 같이 불량률과 공구수명에 대한 예측을 하기 위해서 지도학습 모델인 LR, SVM, RF, XGBoost 등의 알고리즘을 적용하여 데이터 예측 결과의 정확도를 검증하고 시각화를 솔루션으로 구현하고자 한다. 상시 모니터링 사후분석은 통계(SPC)분석 솔루션으로 진행한다.

# 제 3 장 제조데이터 수집분석 및 활용 시스템 설계

## 제 1 절 제조데이터 수집분석 및 활용 시스템 설계 개요

### 1) 제조데이터 수집, 분석 구조 설계 개요<sup>55)</sup>



[그림 3-1] 제조 데이터 수집 및 분석 및 활용모델

55) Kim, J., Kim, Y., Kim, S. (2021). A Study on Design of Real-time Big Data Collection and Analysis System based on OPC-UA for Smart Manufacturing of Machine Working, International Journal of Internet, Broadcasting and Communication Vol. 13, No. 4 pp. 122-127.

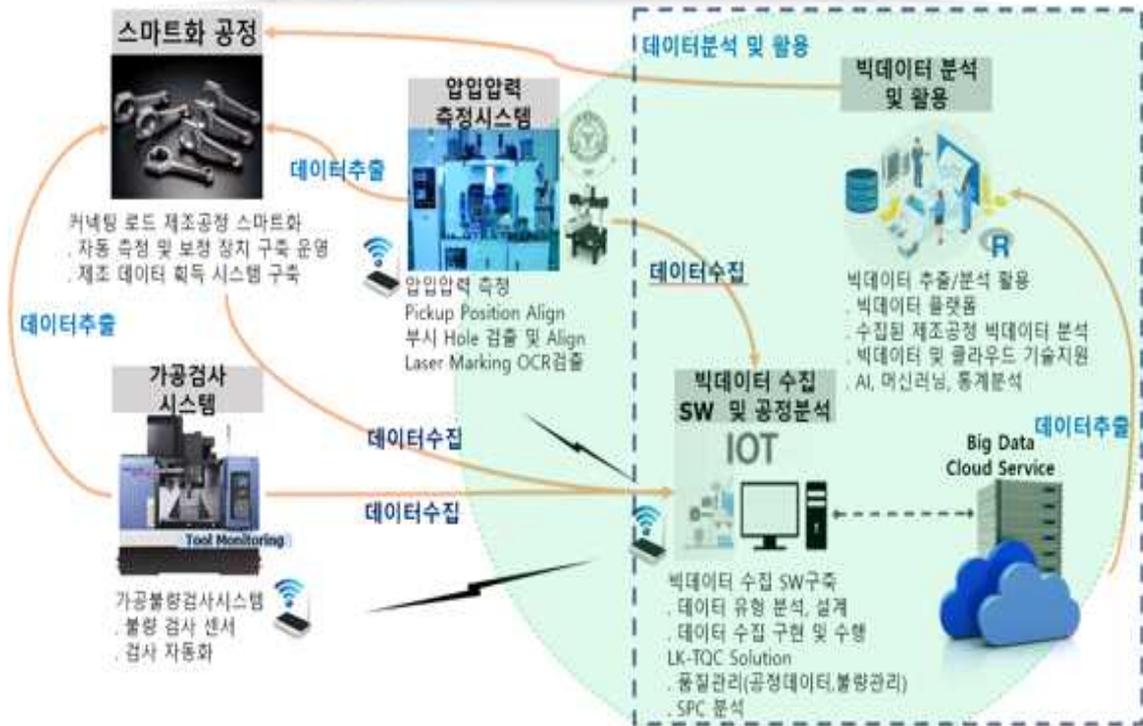
빅데이터 수집 및 처리 기술, IoT 통합 기술, 기계 학습 및 AI 모델링에 대한 선행연구를 수행해, 그림 3-1과 같이 제조 데이터 수집 및 분석 활용 모델에 적용하였다. OT, IT영역의 센서, 디바이스, 애플리케이션 등에서 나오는 데이터를 활용하여 수평 및 수직적으로 통합할 수 있는 OPC-UA를 기반으로 빅데이터 수집 및 분석 모델을 설계한다. 자동화된 제조공정과 각종 제조데이터에 대한 라이브러리를 구축하여 공동으로 활용한다. 첫째, 신규 측정, 검사시스템을 도입해 감소요인에 해당하는 제조데이터를 수집한다. 절삭 가공 데이터, 절삭공구 수명 데이터, 장비 정보 및 로그 데이터를 수집하는 스마트 제조공정을 추가하여 구축한다. 셋째, 데이터 수집 및 분석을 위해 OPC-UA를 기반한 IoT 통합, 제조공정 데이터 수집 및 처리시스템, 소프트웨어 개발 등 시스템을 설계한다. 클라우드 컴퓨팅 서비스와 연계하여 빅데이터 분석을 위한 SPC 통계분석과 AI학습 모델링을 적용한 제조데이터 수집 및 분석 활용 모델을 설계한다. IoT 통합, 제조공정 데이터 수집 및 관리시스템, 소프트웨어 개발 등 데이터 수집 및 분석을 위한 시스템을 구축하고, 빅데이터 분석 결과를 바탕으로 수요기관들에 품질고도화 및 생산성 향상에 대한 대책안 제시 및 클라우드 컴퓨팅 서비스 구축 활용에 대해 기술지원을 한다.

연구범위와 2절 1) 선정배경에서 언급한 대로 기계가공 공정을 대상으로 커넥티드 절삭가공 제조라인의 스마트공장 고도화를 설계를 적용하려고 한다.

## 2) 기계가공 데이터 수집 분석 및 활용을 위한 설계 아키텍처

빅데이터 수집 및 처리 기술, IoT 통합 기술, 기계 학습 및 AI 모델링에 대한 선행연구를 기반으로, 본 연구의 범위와 목적에 맞게 그림 3-1과 같이 실시간 제조 데이터 수집, 분석 및 활용 시스템의 개요에 대한 설계를 적용하였다. 그 다음, 기계가공 업종에 적합한 스마트공장 고도화를 그림 3-2와 같은 아키텍처로 설계하였다.

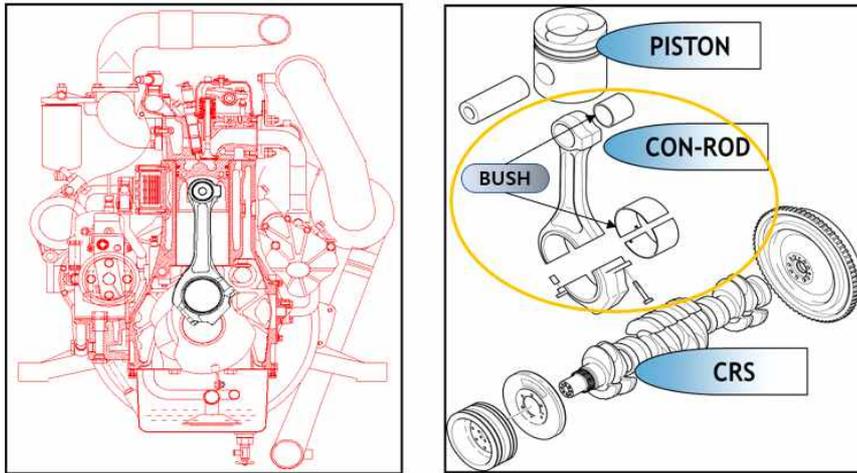
" 수집된 제조 데이터를 분석 AI+클라우드 시스템 활용으로 품질고도화 및 생산성 향상 "



[그림 3-2] 기계가공 데이터 수집 및 활용을 위한 시스템 아키텍처

### 3) 기계가공 스마트화 공정 선정 배경

스마트공장 고도화 대상 중에서도, 기계가공 스마트공장 업종을 선택하였다. 실시간 가공데이터의 수집 및 분석으로 고객사의 요구에 대응하고, 품질향상을 목적으로 스마트공장 고도화사업을 추진 중인 기계가공 업체 중에 그림 3-3과 같이 커넥팅로드를 가공생산 업체에 기계가공 스마트화를 위해 빅데이터의 IoT 통합을 기반한 데이터 수집 및 분석 기술을 적용할 계획이다.

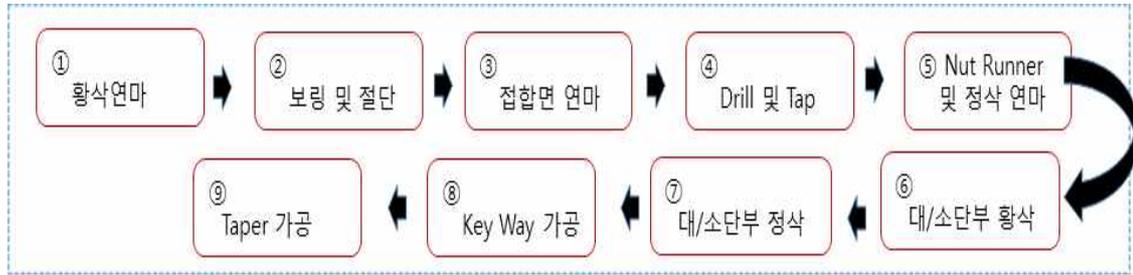


[그림 3-3] 산업용 엔진 및 커넥팅 로드, 부시 관련 조립 부품

산업용 디젤엔진에 배기량은 증가하지 않고 출력을 증대시키는 기술 등을 적용하기 위해, 기존 엔진 부품들의 내구성과 정밀도 제고 등의 과제와 미세먼지의 발생원으로 이슈가 되고 있는 디젤 엔진, 특히 대형 자동차용 엔진의 경우 그 내구성 및 정밀도에 대한 더 높은 품질의 필요성이 대두되고, 이에 따라 라인을 개선하거나 또는 새로운 공정과 생산기술의 개발이 요구되고 있다.



[그림 3-4] 커넥팅 로드 홀 편심



[그림 3-5] 현재 제조 공정도(9단계)

이를 위해 최근 커넥팅로드 제조공정에서는 기계가공 공차가 강화되었으나 이것만으로는 엄격해진 공차를 충족하기 쉽지 않고 또 생산수량과 증량의 관리도 어려워 공정 개선 또는 신 공정 개발이 요구되는 상황이다. 가공 공정과정에서 불량률이 생기면서 그림 3-4와 같이 커넥팅 로드 대단부의 볼트 홀 부분의 홀 편심이 현저하게 발생하고 있다. 이는 슈퍼차저 적용 시 엔진 연소압 증대에 따른 커넥팅 로드 내구성을 저하하는 주요 요인이다.

현재 상용차용 커넥팅 로드는 주강 소재를 기준으로 그림 3-5와 같이 9단계의 가공공정을 거쳐 완성품이 제작되는데, 공정별 가공품질이 완성품의 품질에 직접 영향을 미치기 때문에 전체적인 공정관리가 중요한 요소이다.

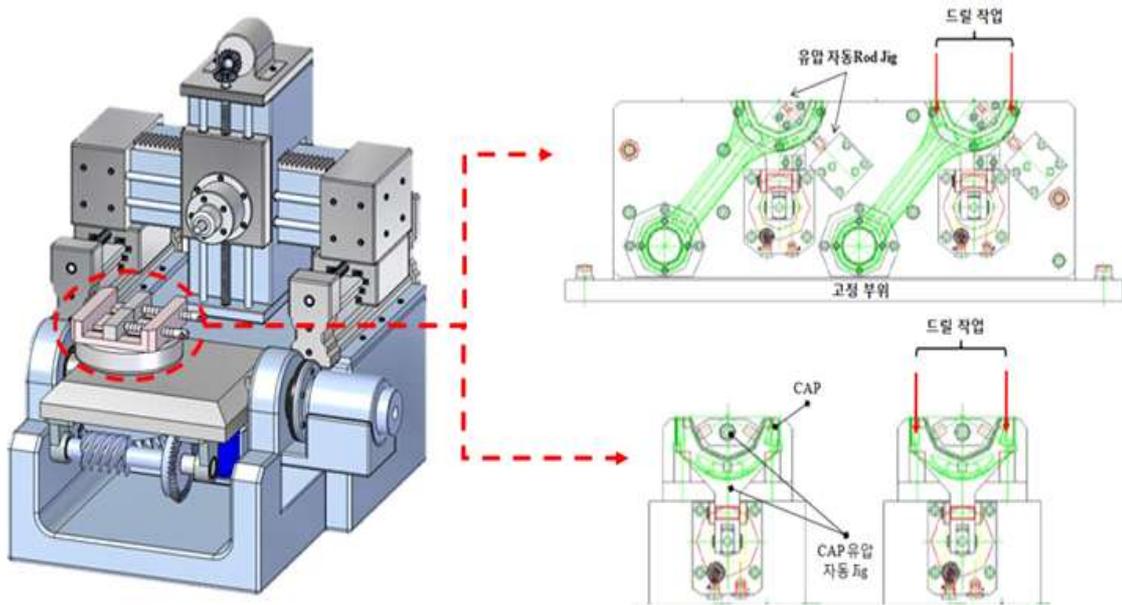
#### 4) 기계가공 스마트화 공정 라인 및 데이터 수집 체계 설계

커넥팅 로드 대단부의 홀 편심 문제로 전체적인 불량률이 발생하고 있어 원가절감 및 생산성 향상을 위해 불량률 감소의 대책이 필요하다. 이를 위해

첫째, 불량률 개선 및 공차 강화 부분의 충족을 위해 치공구 및 이송시스템을 아래 그림 3-6과 같이 1) 드릴구멍 및 탭 가공부의 정밀한 가공을 위한 고정을 개선 및 가공용 Tool을 도입하고 2) 기존 공정에서 발생하던 불량을 제거하기 위한 Cap & Rod 자동유압 지그를 제작하며 3) 드릴 및 탭작업 생산성 제고를 위해 Rod의 대단부 면을 항상 수평 유지 및 고정 후 드릴 및 탭 작업의 자동화 시스템을 구축하고, 4) 가공 품질정확도 및 생산성 제고를 위한 부시 치수 및 압입압력 측정시스템 구축 등 공정개선을 위해 설계한다.

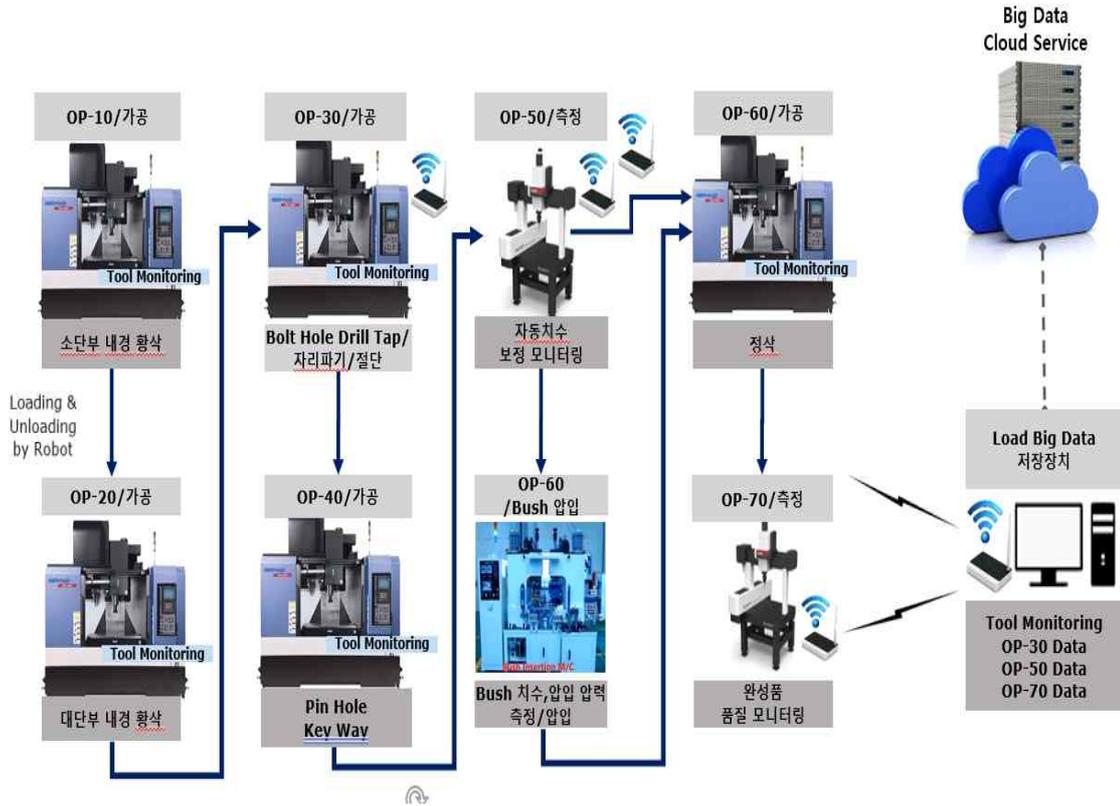
축을 통한 절삭가공 데이터(가공조건, 피삭재 정보, 가공 형태, 동작기계 정보, 절삭공구 정보, 절삭유 정보 등), 절삭가공 관련 기술정보, 절삭공구 수명

등 품질정보에 대한 데이터, 4)불량 발생을 사전에 방지하기 위해 황삭제품의 치수측정과 피드백을 통한 완성품의 치수 측정데이터를 확보한다.



[그림 3-6] 커넥팅로드 가공 공차강화 공정 개선도

둘째, 품질 안정화 및 생산성 향상을 위한 데이터로는, 1) 정밀한 가공 유지를 위해 사용 후 마모 공구의 교체 시기, 지그의 품질 안정성 및 정밀도 유지의 가능 여부를 판단할 수 있는 데이터, 2) 부시 압입 시 치수와 압입 압력에 대한 데이터, 3) 절삭가공의 최적화 및 품질 고도화를 위해 공구 모니터링 시스템 제조데이터 수집·분석·활용을 위해 그림 3-7과 같이 커넥팅 로드 제조와 관련하여 실시간으로 수집된 데이터들은 제품의 품질을 향상뿐만 아니라 생산성 향상에에도 활용한다.



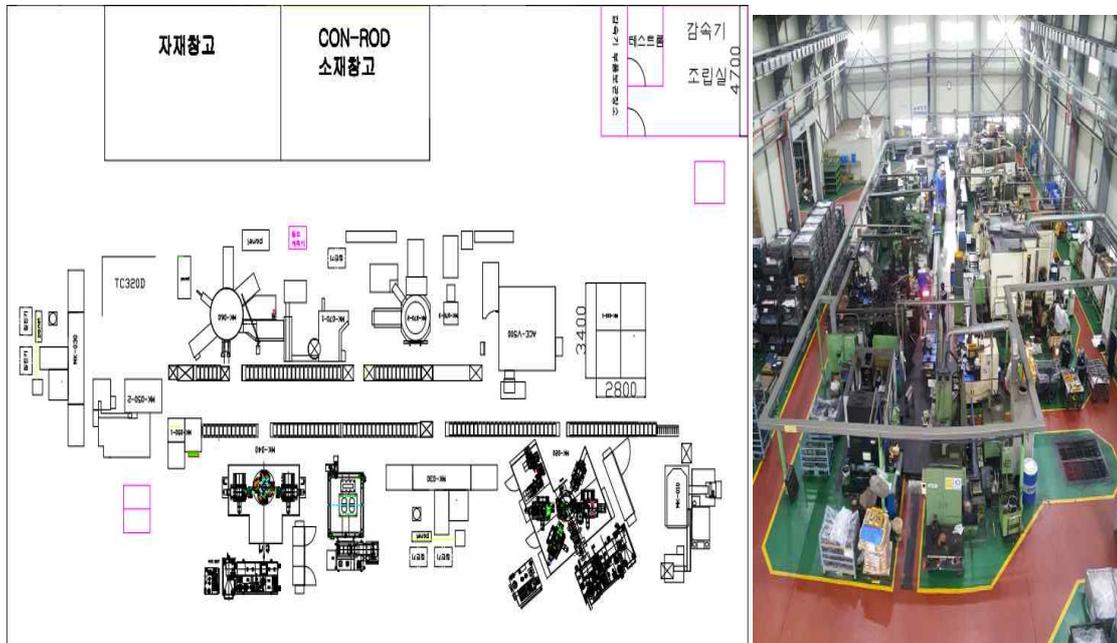
[그림 3-7] 커넥팅로드 데이터 획득시스템 구축 구성도

셋째, 주강소재인 제품의 가공공정에서의 공구마모, 치수불량, 치공구 정밀도 및 고정부 조건상태, 부시 치수 및 압입압력 변화 등 가공조건에 영향을 미치는 모든 요소에 대한 데이터 수집 및 분석을 통해 품질 안정화와 공정 표준화가 요구되고 있다. 커넥팅 로드 공정 모니터링을 통해 데이터 획득 시스템을 도입하여 수요처에서 요구하는 높은 가공정밀도의 공정라인을 구현하고 품질 및 생산성 향상을 위해 커넥팅 로드 절삭가공 공정을 스마트화 및 실시간 모니터링 체계로 실시간 제조데이터 수집 및 분석 플랫폼을 설계한다.

## 제 2 절 스마트화 가공공정 상세 설계

### 1) 제조공정 스마트화를 위한 공정 레이아웃 재배치 및 설계

가공공정의 문제점은 그림 3-8과 같이 현재의 생산라인이 수요처에서 요구하는 가공공정의 정밀도 수준에 충족하기에는 한계가 있다는 점이다. 특히 Rod와 Cap 조립부의 볼트홀 간 거리정밀도 문제, 소단부 부서 압입 공정에서 과도한 시간 소요 및 높은 불량빈도로 인해 생산성이 저하되고 있다. 개선대책으로 그림 3-8과 같이 수요처에서 요구하는 제품의 품질수준 충족과 생산성 향상을 위해 로봇생산 시스템과 생산제품의 품질을 실시간으로 모니터링할 수 있는 시스템으로 공정 레이아웃을 재배치한다. 설계 대상의 가공공정을 스마트화하여 Con-Rod 가공 신 공정을 구축을 아래와 같이 설계한다.



[그림 3-8] Con-Rod 가공공정 개선 전 레이아웃

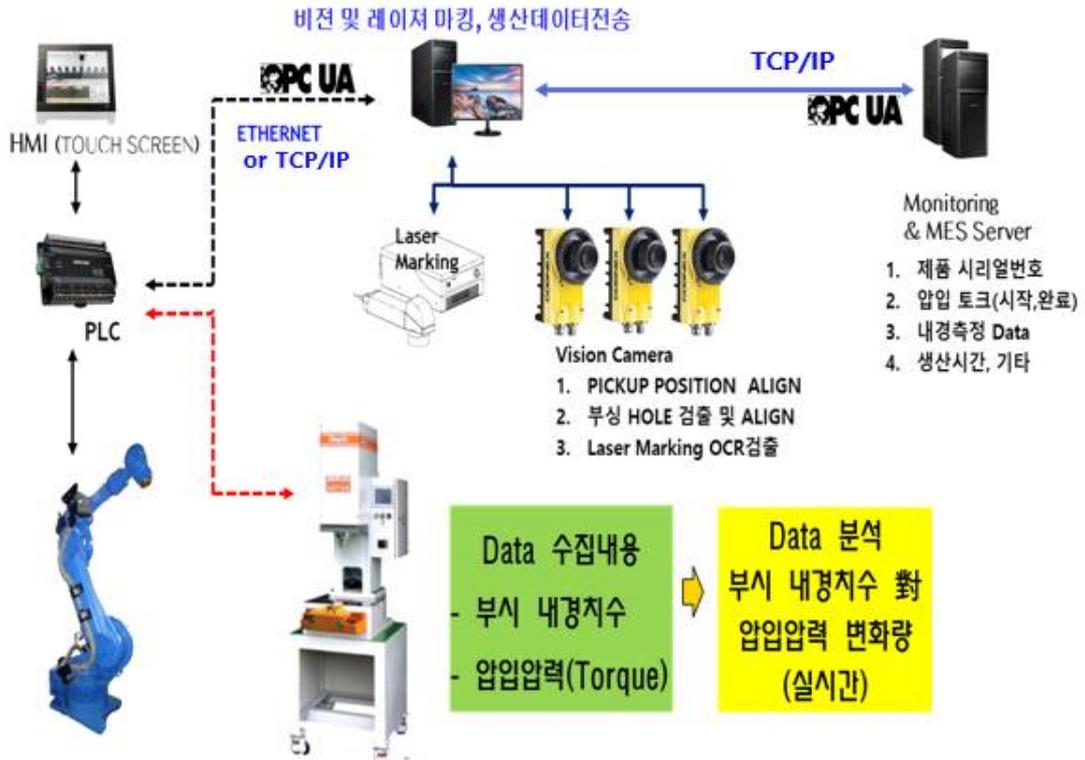
- 신규 수평 MCT 도입 및 설치
- 로봇 및 이송장치 설치
- 공정별 Jig-Fix

OP-20 공정 : Con-Rod 소단부 정삭 및 기준면 가공

OP-30 공정 : Con-Rod 대단부 자리파기 및 절단

- OP-40 공정 : Con-Rod 대단부 볼링 및 볼트홀 가공
- OP-50 공정 : Con-Rod 세레이션 및 조립면, 기준핀 가공
- OP-90 공정 : Con-Rod 대소단부 정삭가공

2) 부시 치수 및 압력 측정, 조립시스템 구축을 위한 설계<sup>56)</sup>



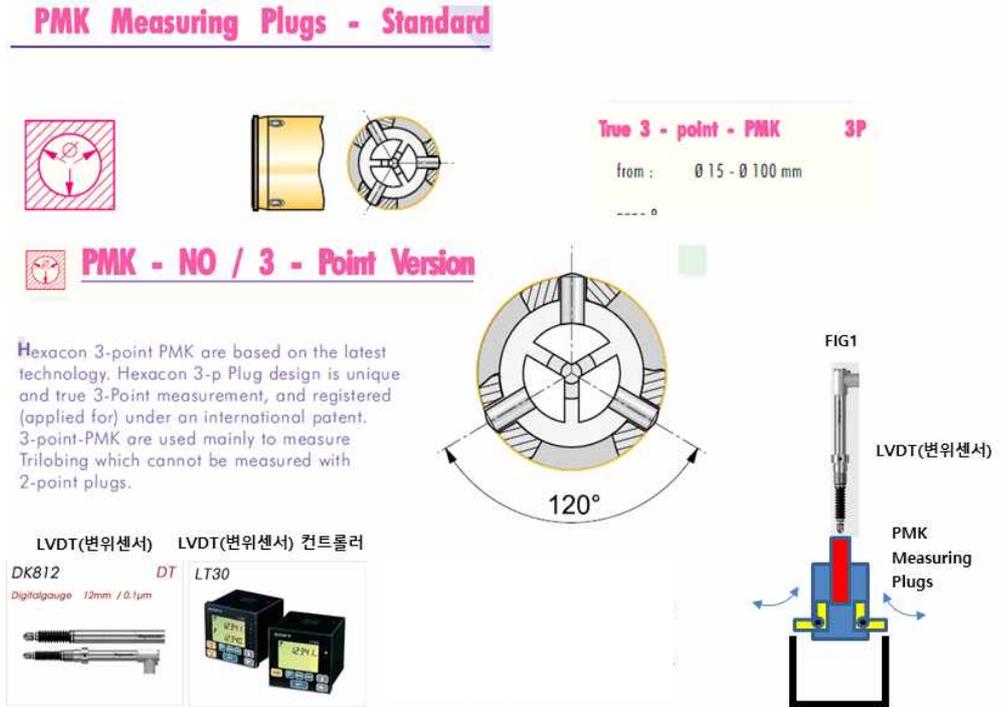
[그림 3-9] 부시 압입 압력 측정 및 전송시스템 구성도

부시 압입압력 측정시스템은 그림 3-9의<sup>57)</sup> 데이터전송 시스템 구성도와 같이 하드웨어는 크게 부시 측정을 위한 마이크로미터, 측정장치, 측정시스템으로 구성된다. 측정시스템은 LVDT(선형가변 변위변환기)의 변위센서로 내경의 변위량을 측정하고, 이 정보를 PLC 및 HMI로 비전카메라에 내재된 위치와 홀에 대한 검출 기반의 조정 기능과 레이저 마킹 정보로 검출하여 제품 시리얼 번호, 압입 토크(Torque) 내경, 생산시간 등의 데이터를 OPC-UA기반으로 한 TCP/IP통신으로 서버로 보낸다. 동시에 이 정보는 OPC-UA기반 이더넷으로

56) Kim, J. op.cit., pp 124.

57) Kim, J. ibid., pp 124.

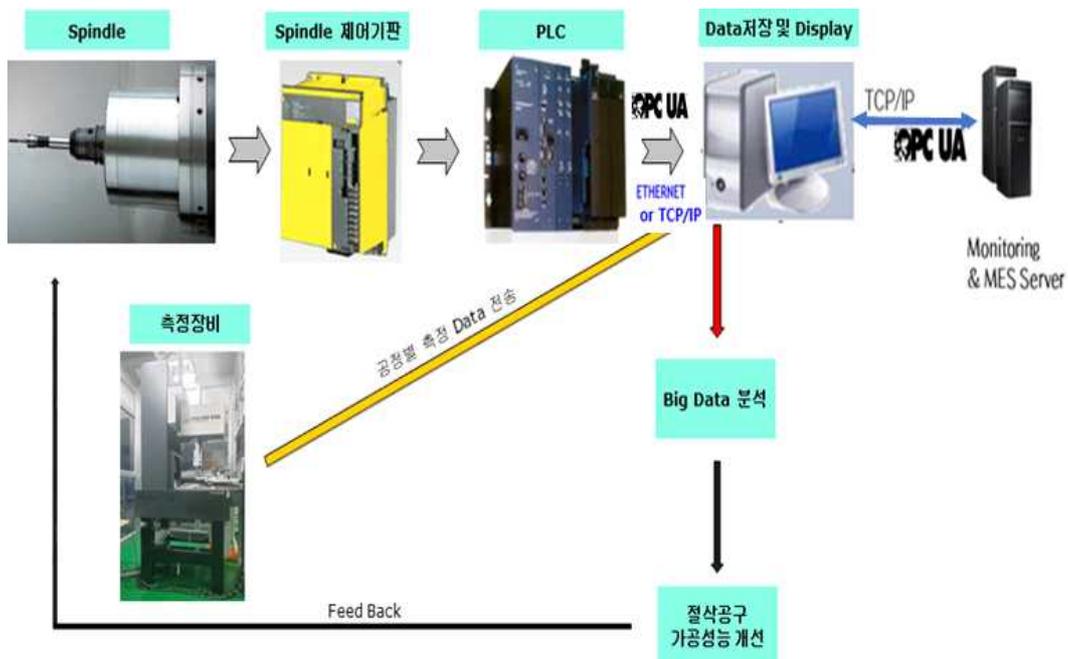
PLC를 거쳐 작업용 로봇에 전달되며 HMI에 작업상태가 표시된다. 부시 내경 측정원리는 그림 3-10과 같이 내경측정 치구가 측정하고자 하는 제품 홀에 삽입되면 베어링 치구가 밀려 들어가면서 상부에 있는 LVDT 변위센서를 밀어 올리고 변화된 양을 내경으로 측정한다.



[그림 3-10] 부시 치수(내경) 측정 원리

### 3) 공구 모니터링 및 검사시스템 구축을 위한 설계<sup>58)</sup>

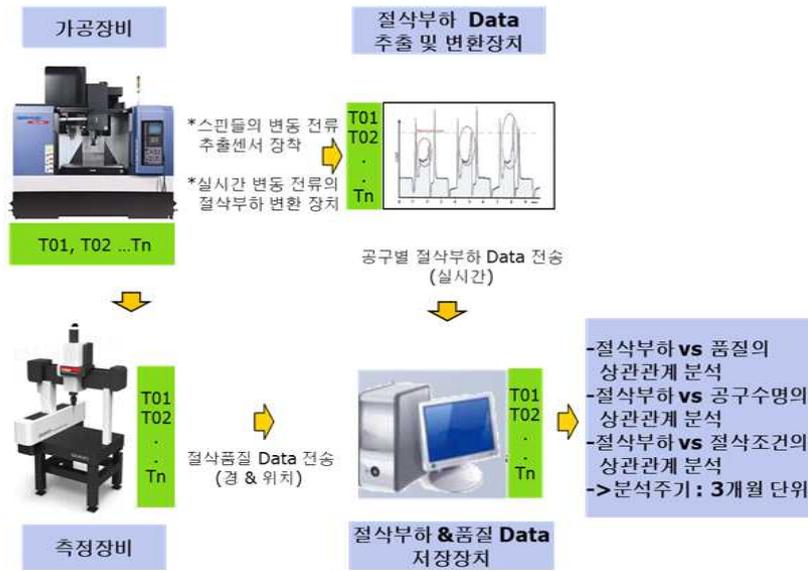
공구 모니터링 및 검사시스템의 원리는 그림 3-11과<sup>59)</sup> 같다. 공구 모니터링 및 추출 단계에서는 가공 시 설비 스피들(spindle)에 절삭 Tool을 장착한 후 스피들 제어기판의 센서를 통해 가공장비의 스피들에 걸리는 전류의 변동치를 추출한다. 실시간 변동데이터를 PLC에 전송한 다음 전류치 변동데이터를 뉴턴 단위로 변경한 실시간 절삭 부하데이터로 변환한다. 이 절삭 부하데이터는 실시간 공구의 상태를 관리하는 시스템의 저장장치로 전송하며 각 장비 본체에 구성한다. 측정검사단계에서 측정장비는 커넥팅로드의 대소단부, 홀의 위치, 경 등 데이터를 실시간으로 측정하기 위한 생산 라인용 가공품질 측정 시스템이다. 마지막으로 절삭부하와 가공측정 데이터를 OPC-UA기반의 이더넷, TCP/IP의 유무선 통신으로 공구별로 실시간 저장 및 조회를 할 수 있으며 서버에도 동시에 저장한다. 상관분석 등의 통계 결과로 공구 인선부 형상, 재질, 코팅 정도 등 가공공정 조건을 최적화하고 표준화한다.



[그림 3-11] 공구 모니터링 및 검사시스템 구성도

58) Kim, J. op.cit., pp 124.

59) Kim, J. ibid., pp 124.



[그림 3-12] 공구 모니터링 데이터 추출 및 연관관계 분석

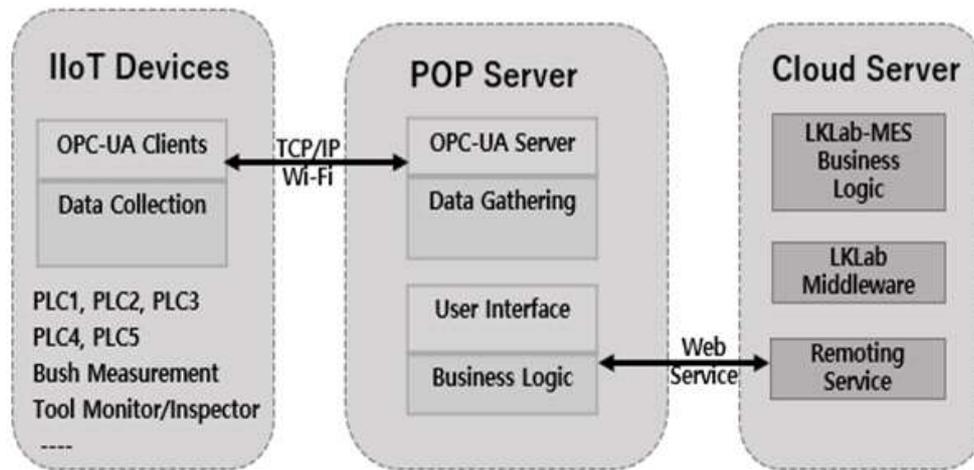
공구 모니터링 시스템은 그림 3-12와 같이 절삭부하 대비 가공품질, 공구수명, 절삭조건의 상관관계 분석을 한 후에 개선한다. 분석 주기별로 절삭공구 및 절삭조건의 최적화 작업을 진행한다.

### 제 3 절 제조데이터 수집, 분석 구조 상세 설계<sup>60)</sup>

#### 1) 데이터 수집을 위한 OPC-UA 기반 IoT 통합 설계

##### 가) OPC-UA 서버 / 클라이언트 설계

중소기업의 네트워크는 이기종 설비, OT, IT 등 각 영역 데이터 간의 인터페이스가 되지 않은 상태에서 자동화가 구축되어 있고, 이로 인해 실시간 처리도 어려워 데이터 수집 및 분석 시스템 구축에 한계가 있다. 대다수의 중소기업이 통신 및 OT 영역의 문제점에 대한 사전준비 없이 IT 위주로 스마트 공장을 도입해왔으나, 최근 스마트공장 고도화 사업을 중심으로 OPC-UA의 최신 통신 프로토콜의 도입이 점차 늘어나고 있다. OPC-UA는 플랫폼에 의존적이지 않은 독립적 구성이 가능한 프로토콜이며, 프로토콜 내부에 구현된 안정적인 통신 메커니즘을 제공할 수 있다. 따라서 이기종 설비와 OT, IT 영역 간에, OPC-UA를 기반으로 실시간 데이터 수집 및 저장 시스템을 설계할 필요가 있다.



[그림 3-13] IIoT 디바이스와 시스템 구성도

설계하고자 하는 IIoT(Industrial IoT) 디바이스<sup>61)</sup>는 그림 3-13과<sup>62)</sup> 같이 제조설비의 절삭조건의 원시데이터를 OPC-UA 산업용 표준 프로토콜을 활용

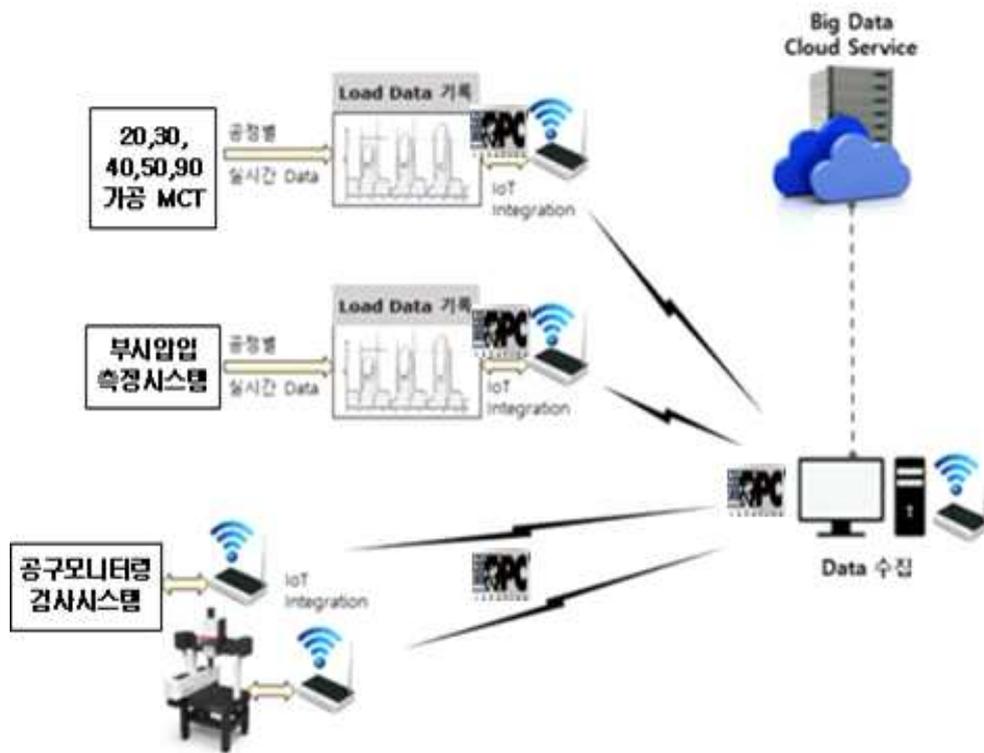
60) Kim, J. op.cit., pp 122.

61) 유흥식, 전계논문, pp 26.

62) Kim, J. ibid., pp 122.

하여 수집한다. TCP/IP과 와이파이를 포함한 유무선 네트워크 통해 IIoT 디바이스에서 POP(Point of Production) 서버로 수집된 데이터를 OPC-UA 서버 프로토콜을 통하여 데이터를 취합한다. 취합된 데이터는 POP 서버에서 애플리케이션 브라우저에 의해 웹 서비스를 통하여 클라우드 서버에 전달되어 제조현장 운영시스템인 MES의 비즈니스 로직에 의하여 수집데이터를 저장 관리한다. 수집 저장된 데이터는 분석 가공하여 활용한다.

나) OPC-UA 기반 IIoT 통합 네트워크 설계<sup>63)</sup>



[그림 3-14] 제조공정 데이터 수집 및 분석 시스템 네트워크

수집 가능한 제조·품질 데이터를 정의하고 연계공정을 분석하여 생산현장의 가공·품질 데이터의 취득, 전송, 공동보관 및 활용 등을 적용하고 측정가능 데이터의 수집방안과 데이터의 저장·분류방안 및 데이터 간의 연계 포인트를 공동으로 분석하고 공정에 활용하여 통합 시스템 내 유무선 네트워크로 구성된 통신

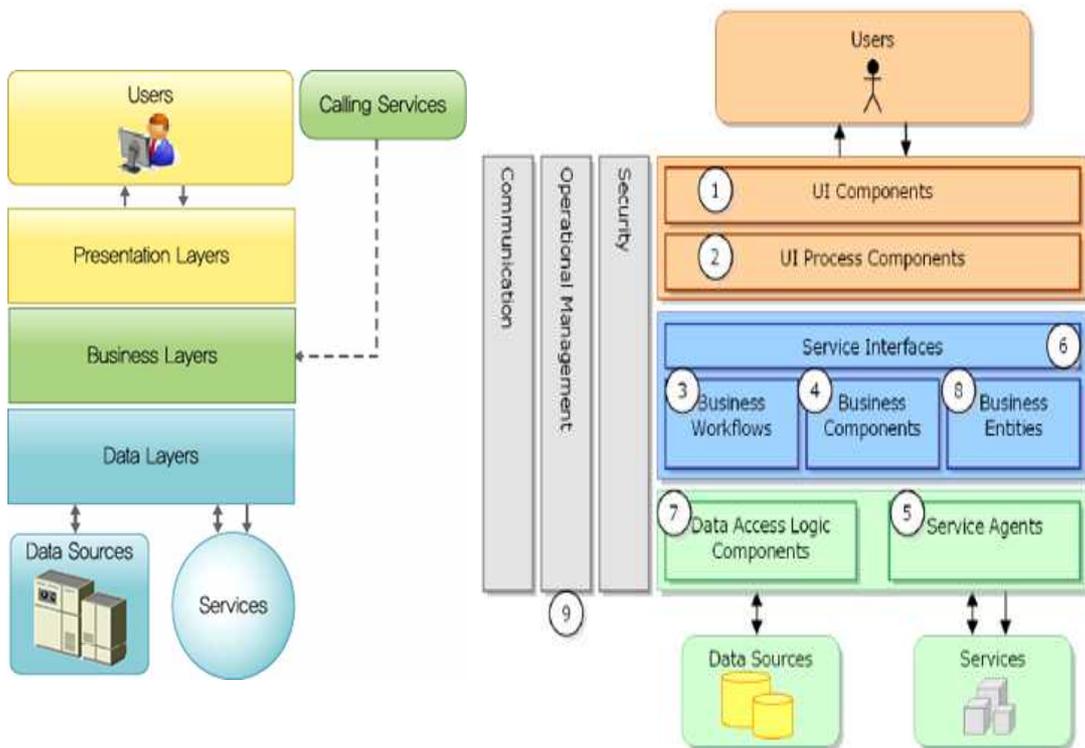
63) Kim, J. op.cit., pp 124.

모듈을 활용하고 공정 및 측정장비의 점점 IoT 통합 모듈을 설계한다.

제조현장의 절삭, 황삭 등 절삭가공설비 5대의 PLC에서 Rod와 Cap 조립부의 볼트 홀 간의 거리정밀도, 소단부 부시압입 공정에서의 시간소요 등의 품질 개선 생산성 향상을 위한 데이터를 수집한다. 부시압입 측정장비에서는 부시압입, 내경측정, 부시 조정 등의 정보를 수집하고, 공구 모니터링 및 검사시스템 등에서 절삭조건의 원시데이터를 OPC-UA 산업용 표준 프로토콜<sup>64)</sup>을 활용하여 그림 3-14와 같이<sup>65)</sup> OPC-UA 기반으로 네트워크를 설계한다.

## 2) 데이터 수집 SW 개발

가) 3계층(Tier) 스마트 클라이언트 애플리케이션 브라우저 SW설계<sup>66)</sup>



[그림 3-15] 3계층 구조 및 엔터프라이즈 솔루션 프레임 구성

64) OPC Unified Architecture. 전게서.

65) Kim, J. pp 123.

66) 김영주, 전계논문, pp. 203-209.

3계층 구조와 스마트 클라이언트 개념의 애플리케이션 브라우저로 설계한 엔터프라이즈 솔루션을 적용한다. 3계층 구조는 그림 3-15의 좌측<sup>67)</sup>과 같이 구성되며, 프레젠테이션(클라이언트) 계층, 비즈니스(애플리케이션) 계층, 그리고 데이터 계층의 3계층은 서로 물리적으로 독립적이며, 각 계층의 변경이 다른 계층에 의존하지 않는다. 프레임 구성 범위는 그림 3-15의 우측<sup>68)</sup>과 같으며, ①과②는 프레젠테이션 계층의 스마트클라이언트 애플리케이션 브라우저이다. ④, ⑥, ⑧은 애플리케이션 계층에서의 프레임워크의 클래스를 생성하는 클래스 빌더이다. ⑦의 엔터프라이즈 데이터 매니저(Enterprise Data Manager, EDM)는 데이터 계층의 데이터 액세스 로직 컴포넌트이며, ⑨의 시스템 매니저는 전체적인 시스템을 운영하고 관리한다.

애플리케이션 브라우저의 특별한 점은 웹 브라우저와 같이 브라우저가 컴파일될 때 사용자에게 유용한 기능이 고정되어 있지 않는다는 것이다. 시스템의 관리자 및 개발자들은 사용자가 새 기능을 받기 위해 일정한 설치 단계를 수행할 필요가 없도록 허용한다. 보안과 사용자 인증에 대한 애플리케이션의 주요 기능은 원격 액세스가 가능하게 하려면 전체 프로그램을 브라우저를 통해서만 실행할 수 있다. 전체 프로그램에서 브라우저를 실행하도록 하려면 개발 사용자가 인증을 통하여 브라우저는 어떤 애플리케이션 서버에 프로그램을 원격 호출할 것인지를 결정한다. 또한 각 사용자에게 개별 메뉴가 생성되기 때문에, 권한이 주어지지 않은 프로그램의 경우 접근 혹은 사용이 제한된다. 스마트 클라이언트 기능은 로컬 리소스를 이용하고 지능적으로 분산 데이터 원본에 연결되므로 조정할 수 있는 클라이언트 응용 프로그램이다. 스마트 클라이언트는 씰 클라이언트와 리치 클라이언트가 각각 가진 장점의 결합이라고 볼 수 있다.

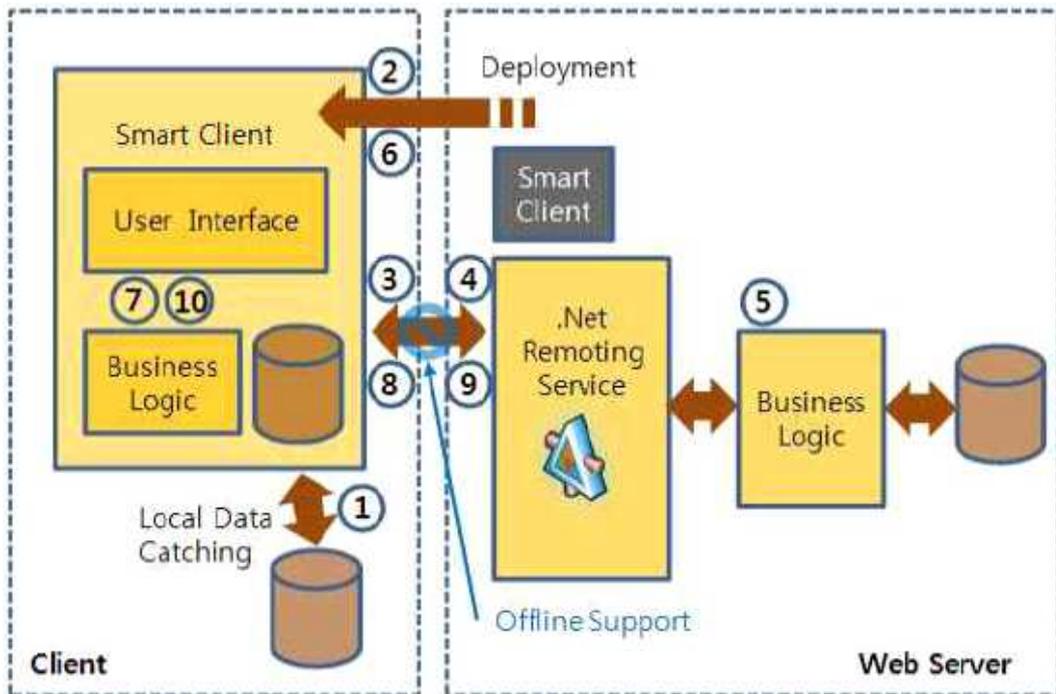
스마트 클라이언트 애플리케이션 브라우저의 구조는 그림 3-16과<sup>69)</sup> 같다. 애플리케이션은 원격 호출주소로 애플리케이션 서비스에 접근하여 3계층 환경을 구현한다. 또한 스마트클라이언트의 특징을 효과적으로 수행하려면 이 브라우저 내의 실행 애플리케이션의 주요 로직에서 애플리케이션 서버에서 실행하

67) 김영주. 상계논문, pp. 204.

68) 김영주. (2010). .Net 기반에서 엔터프라이즈 애플리케이션 프레임워크 구현. 인천대학교 대학원 박사학위 논문. pp 48.

69) 김영주. 전계논문, pp. 206.

도록 한다. 시스템의 실행 순서에 따라 요소 내용을 살펴보면 다음과 같다.



[그림 3-16] 스마트 클라이언트 애플리케이션 브라우저 구조

(1) 액세스 DB로부터 가져온 애플리케이션 브라우저를 실행하기 위한 속성은, 저장 DB의 중요정보로 IP주소, 웹 서버 주소, 시스템 등이 있다.

(2) 애플리케이션 브라우저에 관련된 DLL 파일은 웹 서버로부터 설치한다. 설치자료가 없거나 구버전이 아닐 경우 이 단계는 뛰어넘는다.

(3) 서버와 보안 채널을 오픈하고 서버에서 공개키를 받는다.

(4) 공개키로 유저명과 비밀번호를 암호화하여 서버 인증을 위해 전달한다.

(5) 서버는 개인 키로 복호화하여 시스템에서 관리하는 정보와 일치할 경우 토큰을 발행해서 클라이언트를 인증한다.

(6) 인정 절차가 통과되었다면, 공통 DLL파일이 없거나 구버전인 브라우저는 서버로부터 오는 애플리케이션에서 다운로드 받는다.

(7) 애플리케이션을 실행하려면 브라우저는 메뉴 정보를 서버의 애플리케이션에서 가져와 메뉴를 그린다.

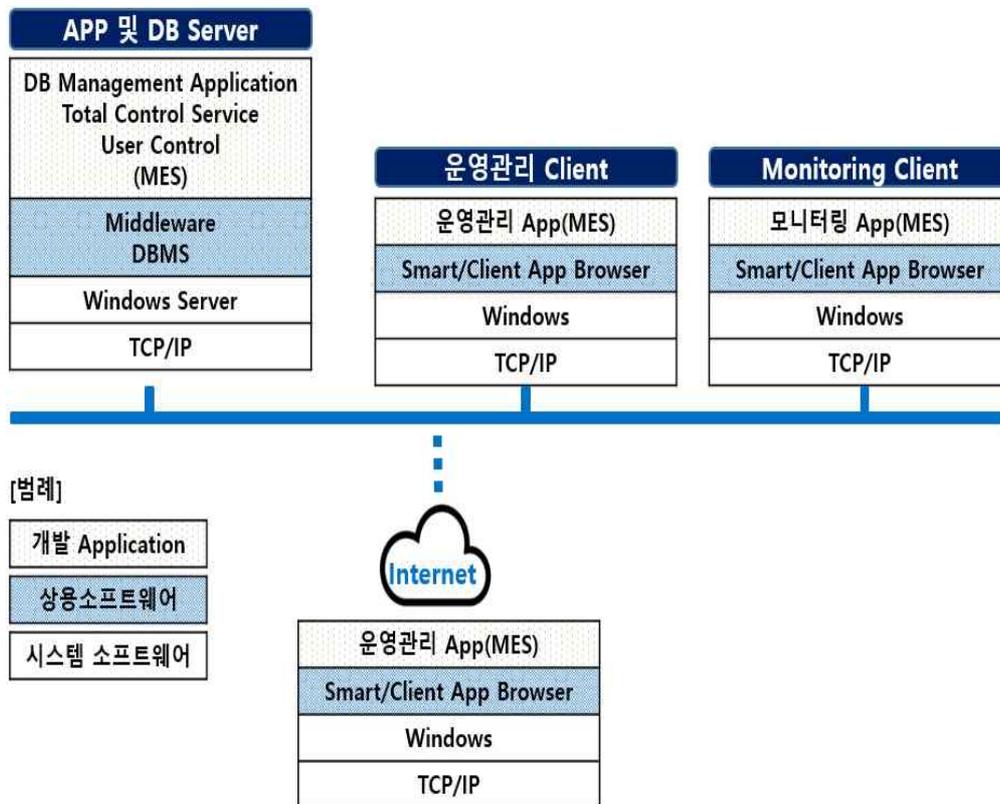
(8) 애플리케이션 서버의 원격 서비스를 사용하기 위해 서버에서 서비스 호출 주소 등의 필요한 정보를 가져와 원격 호출주소를 확정한다.

(9) 애플리케이션의 실행을 위해 사용자가 메뉴를 클릭했을 때, 실행 DLL에서 사용하는 참조 DLL이 구버전이거나 실행 폴더에 DLL이 실행 폴더에 없으면 서버에서 다운로드받는다. 참조 DLL을 잘못 게시할 경우 검증한다.

(10) 애플리케이션 프로그램을 실행한다.

### 나) 데이터 수집 SW 설계

데이터 수집 소프트웨어는 3계층 구조와 스마트 클라이언트 개념의 애플리케이션 브라우저로 설계한 엔터프라이즈 솔루션 아키텍처를 적용하여 설계한다.



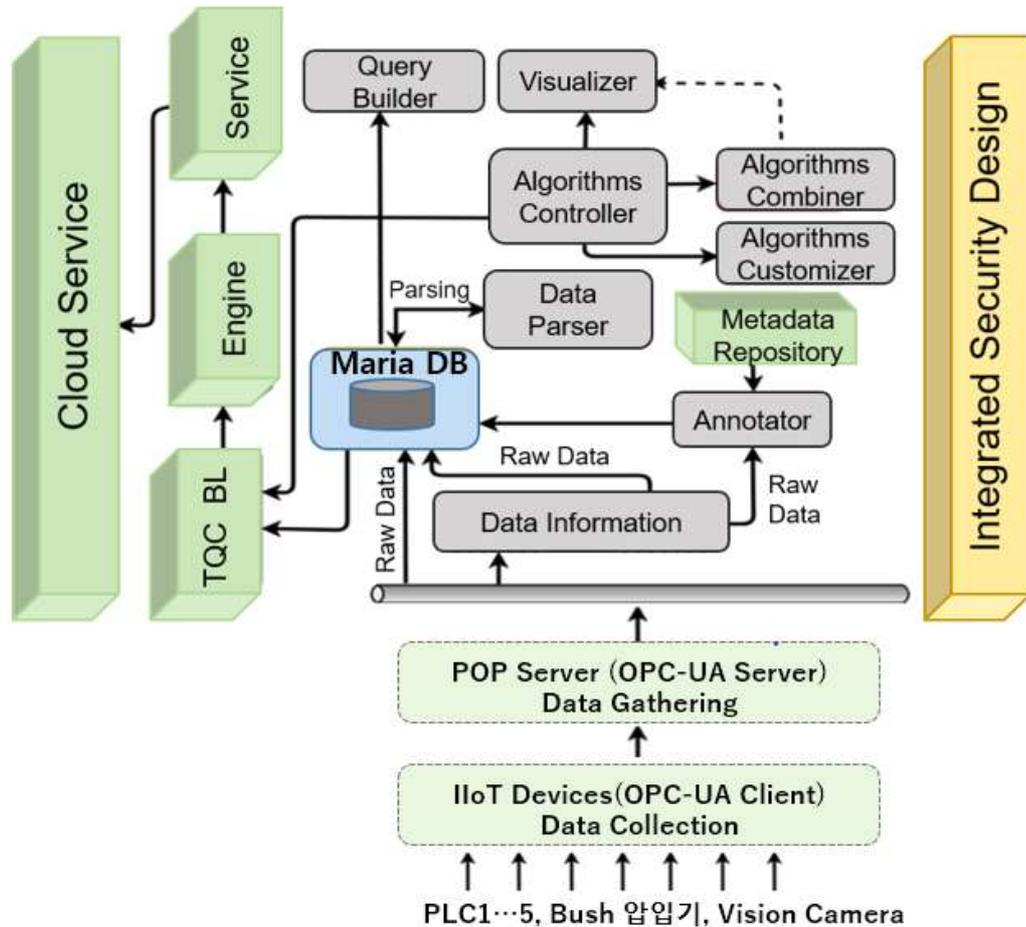
[그림 3-17] 데이터 수집 SW 기능 구성도

그림 3-17과 같이 DB 서버, 애플리케이션 서버, 웹서버 3개의 레이어로 구성되며, DB 서버는 애플리케이션 서버에 의해 접속할 수 있어 보안이 뛰어나며 애플리케이션 서버는 비즈니스 로직과 다양한 관리 서비스를 지원한다. 웹 서버는 접속자 및 배포관리·보안관리 등의 기능을 서비스한다. 클라이언트는 스마트 클라이언트 애플리케이션 브라우저를 통하여 서버에 접속 할 수 있으며, 접속

시 비즈니스 로직 등의 버전이 변경되었을 경우 자동 배포되어 클라이언트에 애플리케이션 프로그램이 다운로드된다. 배포된 애플리케이션 프로그램은 클라이언트에서 운영관리, 모니터링 등 작업을 진행한다.

### 3) 데이터 수집 및 분석시스템 설계

#### 가) 빅데이터 수집 및 분석시스템 구조 설계



[그림 3-18] 빅데이터 수집 및 분석시스템 구조도

빅데이터 수집 및 분석시스템은 그림 3-18과 같이<sup>70)</sup> 제조현장의 PLC 및 부시압입 측정 시스템, 공구 모니터링 및 검사시스템 등 IIoT 디바이스에서 추출된 원시데이터를 OPC-UA 프로토콜을 통하여 POP 서버에서 개더링하여 웹 서비스를 통하여 클라우드 서버로 전달한다. 클라우드 서버에 수집된 원시데이

70) Kim, J. op.cit., pp 126.

터는 그림 3-18의 데이터 Information, Annotator에 의해 가공되어 MES의 DB에 저장된다. DB에 저장된 품질 관련 데이터는 MES의 비즈니스 로직과 다양한 서비스 애플리케이션을 통하여 사용자에게 운영 및 모니터링 서비스를 제공한다.

SPC통계분석은 MES 시스템 품질 데이터에 의하여 구현되며, AI 학습도구를 활용한 시각화는 TQC(Total Quality Control) 비즈니스 로직을 통하여 데이터를 전달받아 데이터를 분석하고 알고리즘을 통하여 시각화 도구를 활용하여 시각화한다.

나) 빅데이터 수집 및 분석 상세 설계

(1) 데이터 수집 대상 항목의 분석

표 3-1의 MES 상에서의 데이터항목 수집 및 분석 내용과 같이 3개 설비·장비에서 추출한 항목의 데이터를 수집한 후 절삭부하와 품질요소, 공구수명, 절삭조건의 상관관계, 부시내경과 압입토크를 분석한다<sup>71)</sup>.

[표 3-1] 데이터 수집 항목 분석결과

구분		내역
데이터 수집	절삭가공 설비	Machine/Tool 정보, 치수 변화량, 공구부하량, 공구수명, 절삭부하 등
	부시압입 측정장비	부시내경, 줌쇠, 압입압력(Torque)
	Tool 검사장비	Con-Rod 황삭&정삭치수(Rod & Cap 볼트 홀, 대칭도, 대단부, 살두께, 내경 등)
데이터 분석	상관관계	절삭부하 vs 품질요소 절삭부하 vs 공구수명 절삭부하 vs 절삭조건 부시내경(줌쇠) vs 압입Torque

그 결과 값으로 가공장비를 보정하고 공정 조건을 최적화 및 표준화하여 품질 및 생산성을 관리한다. 이 공정제품의 절삭데이터에 의한 통계 결과 값을 불

71) Kim, J. op.cit., pp. 126.

량방지를 위해 유사 공정에 수평 전개한다. 부시내경 치수 대 압입압력 변화량은 부품 삽입기 제작에 적용하고 Tool가공 및 부하 데이터는 절삭조건 최적화를 위해 가공공정에 적용한다. 그림 3-19와 같이 가공 및 품질측정 데이터 수집 분석 장표의 설계는 다음과 같다. 부시 압입측정기, 커넥팅로드 가공검사 결과, 비전 검사 가공데이터 추출에 대한 데이터 수집 정보이다. 데이터 수집 항목을 분석한 장표를 근거로 MES에서 가공 설계한다.

**263 CON ROD CMM 검사성적서**

일련 번호	검사일자	기종	CAP													합부 판정
			대단	대단경위치		볼트홀 좌면 높이		볼트홀 위치_상,하			핀홀 위치_상,하			코킹홀		
			경	X	Y	상	하	경	Y	Z	경	Y	Z	Z		
			Ø74.3	2	59.6	36	36	Ø14.5	47	20.375	Ø6	47	7.375	13.375		
			+/-0.2	+/-0.05	+/-0.05	+0/-0.2	+0/-0.2	+/-0.1	+/-0.05	+/-0.05	+0.012 +0.000	+/-0.05	+/-0.05	+/-0.025		
1	2011241000	263													OK	
															NG	

평균값, 최대 최소값 표현 기종별  
측정항목별 검색기능 추가,

일련 번호	검사일자	기종	ROD																합부 판정			
			소단	대단	대단경위치		볼트홀 위치_상			볼트홀 위치_하			핀홀 위치_상			핀홀 위치_하				코킹홀		
			경	경	X	Y	축간거리	경	Y	Z	경	Y	Z	경	Y	Z	경	Y		Z	Z	
			Ø47	Ø74.3	2	59.6	232	Ø13.3	47	20.375	Ø13.3	47	20.375	Ø6	47	7.375	Ø6	47	33.375	32.375		
			+0.025 +0.009	+/-0.2	+/-0.05	+/-0.05	+0.100 +0.000	-0.050 +0.091	+/-0.05	+/-0.05	-0.050 +0.091	+/-0.05	+/-0.05	+0.012 +0.000	+/-0.05	+/-0.05	+0.012 +0.000	+/-0.05	+/-0.05	+/-0.025		
1	2011241000	263																			OK	
																						NG

**BUSH 압입기 품질 Data**

일련 번호	검사일자	기종	Marking번호		BUSH 압입 품질			합부판정
			CAP	ROD	소단경	압입력	압입시간	
1	2011241000	262						
2								

\*검색란 구성  
1)일련번호 범위 지정  
2)생산일자 범위 지정  
3)기종별 지정  
4)Marking 번호 지정 : CAP & ROD  
5)압입력 범위 지정  
\*\*최대값,최소값 및 압력 변화량 그래프화

**CON ROD VISION 검사기 DATA 추출**

일련 번호	생산일자	기종	Marking번호 인식		합부판정
			CAP	ROD	
1	2011241000	262			OK
2					

\*검색란 구성  
1)일련번호 범위 지정  
2)생산일자 범위 지정  
3)기종별 지정  
4)Marking 번호 지정 : CAP & ROD  
\*\* 로트번호 검색기능, 일자별 검색기능

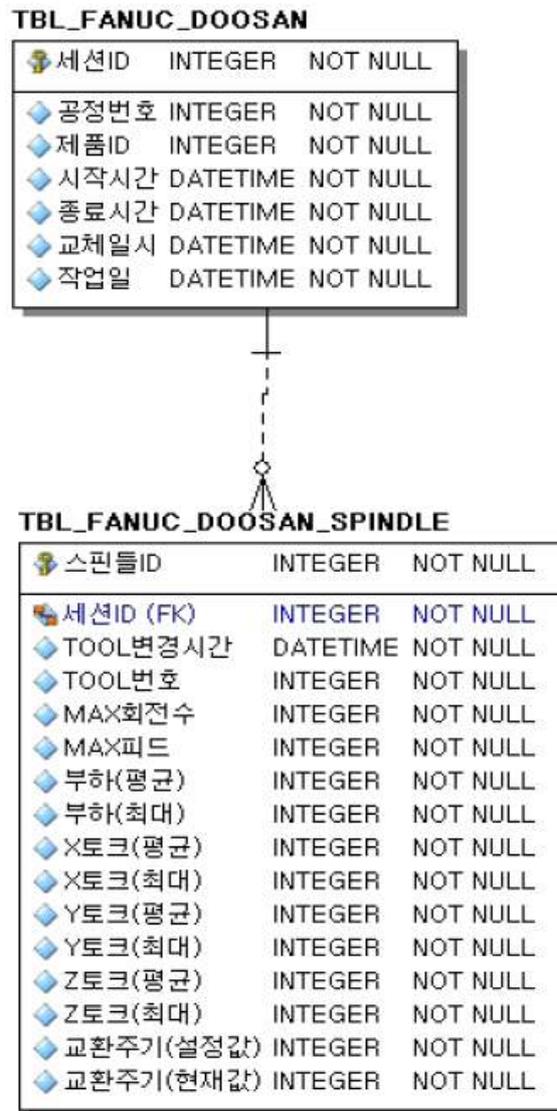
**CON ROD OP-50 가공 DATA 추출내용**

일련 번호	생산일자	기종	공구번호	절삭조건		절삭부하			공구교환 주기	공구수명
				회전수	절삭피드	최소	최대	평균		
1	2011241000	262	T01							
		266	T02							
			T03							
			T04							
			T05							
			T06							
			T07							
			T08							

\*검색란 구성  
1)일련번호 범위 지정  
2)생산일자 범위 지정  
3)기종별 지정  
4)공구번호 지정  
5)공구수명 범위 지정

[그림 3-19] 가공, 품질 측정 데이터 수집 분석 장표

(2) 데이터 수집 및 분석시스템 정보 데이터베이스 설계  
 (가) 정보 데이터베이스 ERD



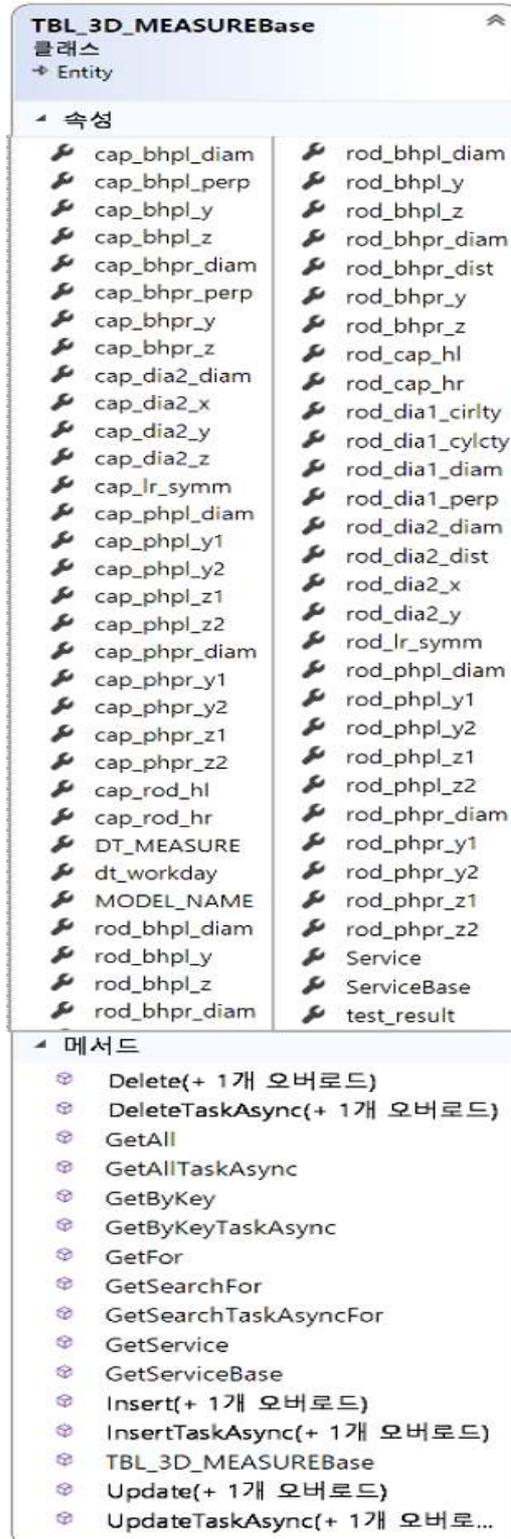
[그림 3-20] 공구 정보 데이터베이스 ERD

TBL\_3D\_MEASURE

모델	INTEGER	NOT NULL
작업일시	DATETIME	NOT NULL
작업일	DATETIME	NOT NULL
ROD_소단경_경	ROD_핀홀위치우_Y2	DECIMAL(20,4) NOT NULL
ROD_소단경_진입도	ROD_핀홀위치우_Z2	DECIMAL(20,4) NOT NULL
ROD_소단경_직각도	ROD_핀홀위치우_경	DECIMAL(20,4) NOT NULL
ROD_소단경_원통도	CAP_대단경_Y	DECIMAL(20,4) NOT NULL
ROD_대단경_Y	CAP_대단경_Z	DECIMAL(20,4) NOT NULL
ROD_대단경_X	CAP_대단경_X	DECIMAL(20,4) NOT NULL
ROD_대단경_경	CAP_대단경_경	DECIMAL(10,4) NOT NULL
ROD_대단경_축간거리	CAP_핀트위치우_Y	DECIMAL(20,4) NOT NULL
ROD_핀트핀위치좌_Y	CAP_핀트위치우_Z	DECIMAL(20,4) NOT NULL
ROD_핀트핀위치좌_Z	CAP_핀트위치우_경	DECIMAL(20,4) NOT NULL
ROD_핀트핀위치좌_경	CAP_핀트위치우_직각도	DECIMAL(20,4) NOT NULL
ROD_핀트핀위치우_Z	CAP_핀트위치좌_Z	DECIMAL(20,4) NOT NULL
ROD_핀트핀위치우_경	CAP_핀트위치좌_Y	DECIMAL(20,4) NOT NULL
ROD_핀트핀위치우_Y	CAP_핀트위치좌_경	DECIMAL(20,4) NOT NULL
ROD_핀트핀위치우_홀간거리	CAP_핀트위치좌_직각도	DECIMAL(20,4) NOT NULL
ROD_좌우핀트홀머칭도	CAP_좌우핀트홀머칭도	DECIMAL(20,4) NOT NULL
ROD_CAP결합면높이_좌	CAP_핀홀위치우_Y1	DECIMAL(20,4) NOT NULL
ROD_CAP결합면높이_우	CAP_핀홀위치우_Z1	DECIMAL(20,4) NOT NULL
ROD_핀홀위치좌_Y1	CAP_핀홀위치우_Y2	DECIMAL(20,4) NOT NULL
ROD_핀홀위치좌_Z1	CAP_핀홀위치우_Z2	DECIMAL(20,4) NOT NULL
ROD_핀홀위치좌_Y2	CAP_핀홀위치우_경	DECIMAL(20,4) NOT NULL
ROD_핀홀위치좌_Z2	CAP_핀홀위치좌_Y1	DECIMAL(20,4) NOT NULL
ROD_핀홀위치좌_경	CAP_핀홀위치좌_Z1	DECIMAL(20,4) NOT NULL
ROD_핀홀위치우_Y1	CAP_핀홀위치좌_Y2	DECIMAL(20,4) NOT NULL
ROD_핀홀위치우_Z1	CAP_핀홀위치좌_Z2	DECIMAL(20,4) NOT NULL
	CAP_핀홀위치좌_경	DECIMAL(20,4) NOT NULL
	CAP_ROD결합면_좌	DECIMAL(20,4) NOT NULL
	CAP_ROD결합면_우	DECIMAL(20,4) NOT NULL
	테스트결과	VARCHAR(2) NOT NULL

[그림 3-21] 커넥팅로드 정보 데이터베이스 ERD

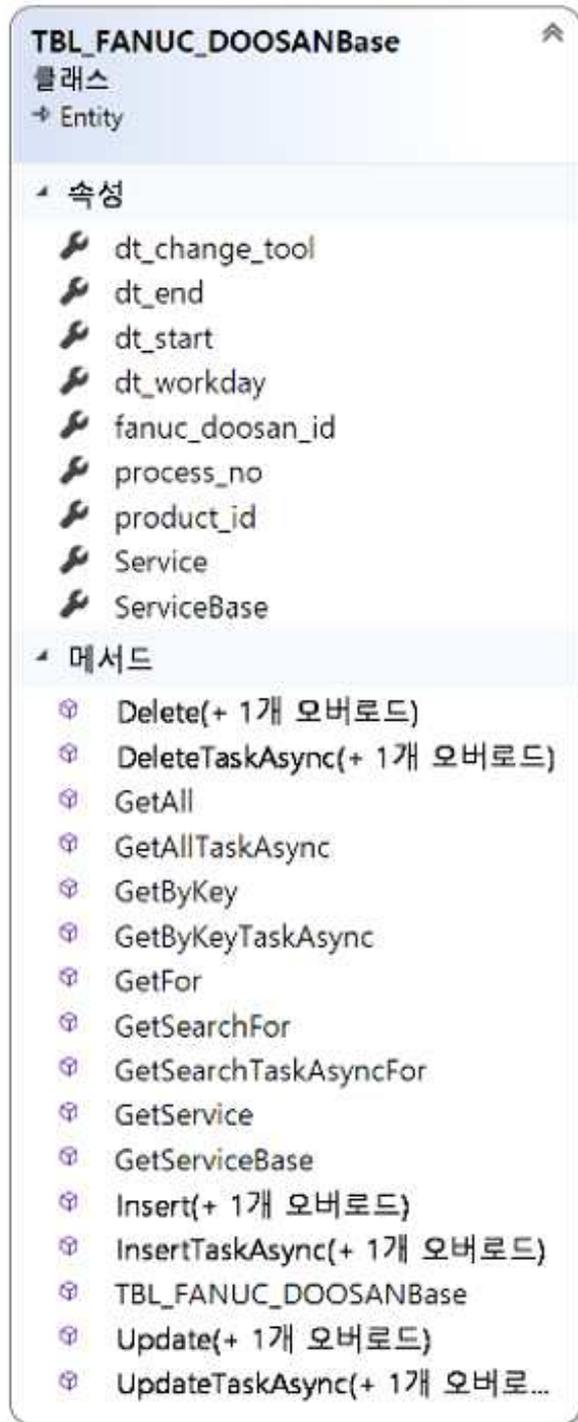
(나) 정보 클래스 다이어그램



[그림 3-22] 커넥팅로드 정보를 나타내는 클래스 다이어그램



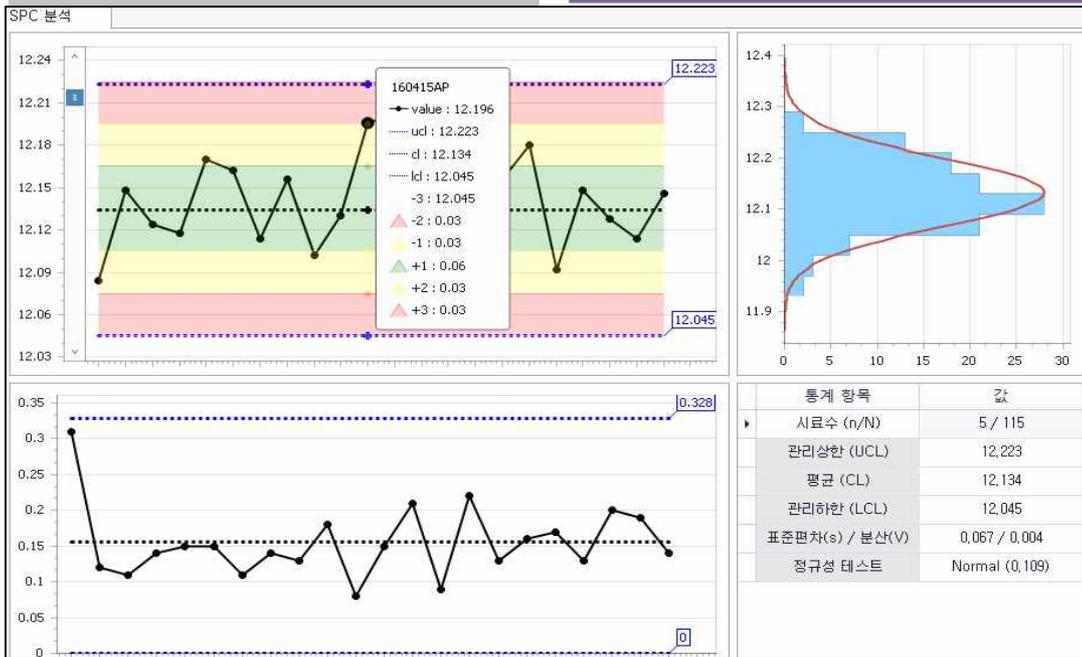
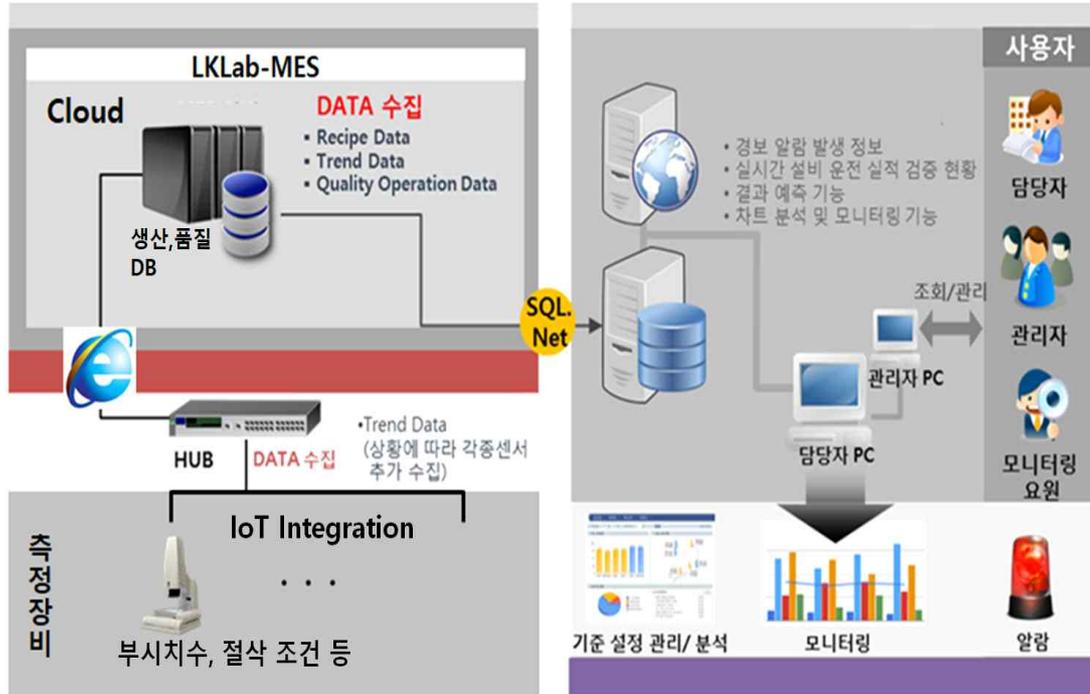
[그림 3-23] 공구정보를 표현하는 클래스 다이어그램



[그림 3-24] 공구 교환시기를 표현하는 클래스 다이어그램

## 제 4 절 SPC 통계분석 & AI 모델링 분석 및 시각화 설계

### 1) SPC 통계분석 및 시각화



[그림 3-25] MES 데이터 수집, SPC분석 및 시각화

1단계는 그림 3-25와 같이<sup>72)73)</sup> 데이터 저장, 분석 및 시각화 프로세스이다. 데이터 분석은 SPC통계로 공정능력(CPk),  $\bar{X}$ -R관리도, 관리도 등 일상적인 공정관리하는 통계를 적용하여 구현 및 설계한다. 아래 수식을 적용한 MES의 TQC솔루션을 기반으로 개발하여 실시간으로 저장, 분석, 시각화한다. MES 기반의 TQC 애플리케이션으로 일상적인 통계적 공정관리 차원에서 SPC 통계 분석을 수행하고 시각화 솔루션은 오픈소스와 FP(FarPoint) Spread를 활용한다.

품질관리는 공정품질의 변동축소에 관리의 초점을 맞추고 있다. 공정의 품질변동을 측정하고 변화하는 데이터를 통계적으로 분석 내지 해석하여 관리하는 통계적 기법을 통틀어서 통계적 공정관리(Statistical Process Control : SPC)라고 한다. SPC 주요도구로서 공정능력지수와 관리도이다.

가) 치우침도 있는 공정능력지수(CPk)

$$CPk = (1 - K)CP, \quad K = |M - \bar{X}| / (T/2) \quad * T: \text{규격공차}, M: \text{규격중심}$$

$$CP = (Su - SL) / 6\sigma \quad (\text{양측규격})$$

$$= (\bar{x} - SL) / 3\sigma = (Su - \bar{x}) / 3\sigma \quad (\text{단측규격})$$

나)  $\bar{X}$ -R관리도

$$\bar{X}\text{관리도 } UCL = \bar{\bar{x}} + A2R, \quad CL = \bar{\bar{x}}$$

$$\bar{X}\text{관리도 } LCL = \bar{\bar{x}} - A2R$$

$$R\text{관리도 } UCL = D4\bar{R}, \quad CL = \bar{R}$$

$$R\text{관리도 } LCL = D3\bar{R}$$

\*  $UCL, LCL$ : 상하한선,  $A2, D3, D4$ : 시료수에 따른 계수

다) P관리도

$$\begin{cases} UCL \\ LCL \end{cases} = \bar{P} \pm \sqrt{\frac{\bar{P}(1 - \bar{P})}{n}}$$

$$P = \frac{Pn}{n}, \quad \bar{P} = \frac{\sum Pn}{\sum n} = \frac{\sum Pn}{kn}$$

MES데이터 수집, SPC분석 및 시각화

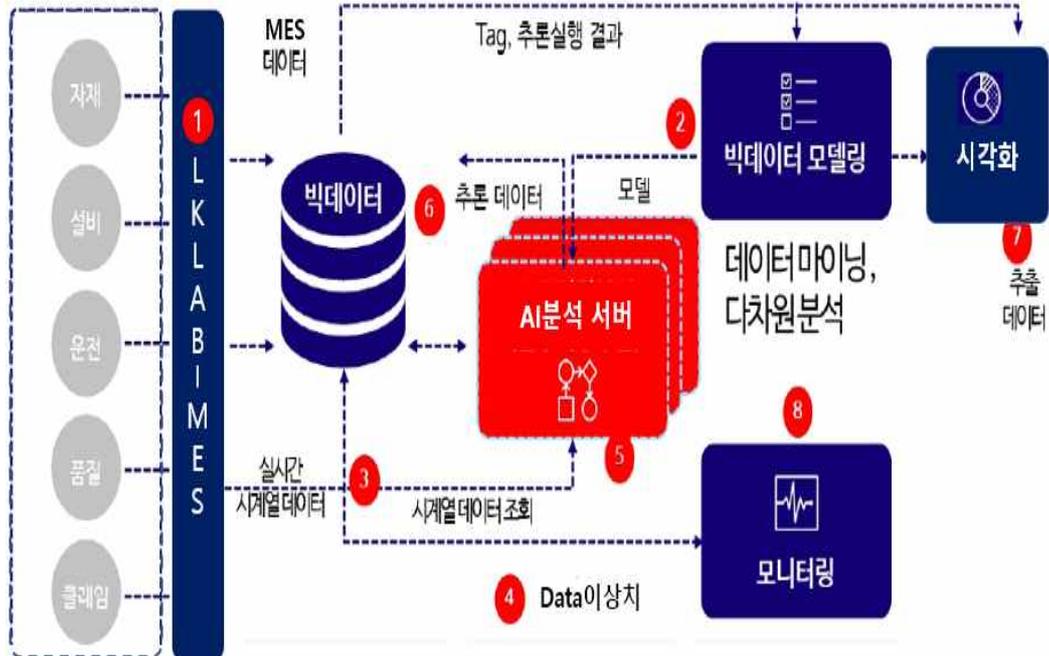
72) Kim, J. op.cit., pp 126.

73) <http://www.dataanalytics.co.kr/>

## 2) AI 모델링 분석 및 시각화

2단계로 AI학습 모델링 분석이다. 커넥팅로드의 기계가공 데이터를 지도학습 알고리즘을 통해 불량률 예측의 정확도를 비교한 뒤, 시각화를 설계하고 실시간 추론 설계와 분석절차를 수립한 후 MES과 연계한 AI솔루션을 설계한다.

### 가) AI솔루션 실시간 추론 설계와 분석절차



[그림 3-26] AI 솔루션 추론 분석, 시각화 및 절차

AI솔루션 실시간 추론 설계와 분석절차는 그림 3-26과 같이 다음과 같다. AI솔루션 실시간 추론 설계는 다양한 발생주기를 기준으로 시간(Time)으로 통합한다. AI모델링 분석 절차 단계이다. 공정, 설비, 판정데이터 등 제조데이터 현황을 분석한다. 데이터 탐색처리 단계에서 원시데이터의 이상치·결측치를 제거 및 정제해 전처리하고, 이상 유형별 원인분석과 품질예측 모델을 개발하여 실행하고 알람을 발생한다. 알람에 따른 조치가이드를 설정하고 트렌드를 시각화하여 상세분석으로 활용한다. 인공지능이 품질데이터에 대한 양품 및 불량품 판정을 내리고 유형별로 분류한다.

종합하면, 품질 추론 판정한 다음, AI 알고리즘은 해당 시간표기(Timestamp)의 시계열 데이터를 시계열 DBMS로 부터 추출한다. 빅데이터 분석을 위한 기초데이터로 활용할 뿐만 아니라 점진적으로 인공지능 추론모델을 최적화하여 재사용한다. 융합된 데이터를 기반으로 품질예측 및 이상증후 패턴을 인지하는 인공지능을 학습 시켜 인공지능 추론모델로 구현한다. 추론 데이터는 다시 데이터 저장소에 저장되고 추출된 추론 데이터를 시각화하면서 시계열 데이터를 실시간 모니터링한다.

#### 나) 구현을 위한 적용 알고리즘 정의 및 주요 수식

지도학습의 알고리즘인 선형 회귀(Logistic Regression, LR), SVM(Support Vector Machine), 랜덤포레스트(Random Forest, RF)로 불량률 예측 정확도를 비교하고 XGBoost (eXtream Gradient Boosting) 알고리즘을 이용하여 시각화를 구현하고자 한다.

##### (1) 선형회귀분석(LR)

학습데이터들의 특성과 관계 등을 파악한 후에, 미지의 입력데이터에 대해서 결과가 어떤 종류의 값으로(양품 혹은 불량) 분류될 수 있는지를 예측하는데 사용하는 분류 모델이다<sup>74)</sup>.

$$f(i) = \beta \cdot X_i$$

74) Freedman, D. A. (2009). Statistical Models: Theory and Practice. Cambridge University Press. pp26.

## (2) 로지스틱 모델

로지스틱 회귀(Logistic Regression)은 독립변수와 종속변수 사이의 관계를 나타낸 함수를 예측 모델에 활용하는 것이 목적으로, 선형 회귀 분석 개념과 유사하지만, 로지스틱 회귀는 선형회귀와는 달리 종속변수가 범주형인 경우에 활용된다. 즉 입력데이터에 대한 결과가 특정 분류로 출력되는 분류 기법으로 간주할 수 있다.

머신러닝 관점에서는 수치예측을 위해 통계학의 회귀분석 기법을 사용하여 훈련데이터를 학습시킨 뒤, 평가데이터를 통해 결과 값을 예측한다는 점에서 지도학습으로 회귀분석 기법을 다루고 있다. 합격/불합격 예측 프로그램을 개발하기 위한 다음 회귀분석의 모델의 식을 나타내었다.

$$y = f(x) + \epsilon$$

오차항으로서 기댓값이 0이고 분산이  $\sigma^2$ 이라고 가정한다. 회귀분석에서는 독립변수  $X$ 가 주어졌을 때 목적변수(혹은 반응변수)  $Y$ 의 조건부 기댓값을 특정 형태의 함수  $f(\cdot)$ 로 가정하고 데이터 세트로부터 해당 함수  $f(\cdot)$ 를 추정하게 된다. 선형회귀모형(linear regression model)의 식은 아래와 같다.

$$\hat{y} = w_0 + w_1x_1 + w_2x_2 + \dots + w_Dx_D = w_0 + w^T x$$

로지스틱 회귀분석을 위해서는, 우선 예측하고자 하는 목표변수  $Y$ 의 범주가 0, 1 두 가지만 있다고 할 때(이항 로지스틱 회귀모델을 가정) 목표변수  $Y$ 의 범주가 1이 될 확률을  $P(Y=1) = P(Y)$ 로 표기하면, 이를 회귀식으로 다음과 같이 표현하는 것에서 시작한다. 아래 식에서 좌변은 오즈(odds)라고도 불린다.

$$\frac{P(Y)}{1 - P(Y)} = \exp(B_0 + B_1X)$$

위 식의 좌변은 확률들의 비율이고 우변은 지수함수의 형태이므로,  $(0, \infty)$ 의 범위를 가지고 있음을 알 수 있다. 이에 좌변 및 우변 모두  $(-\infty, \infty)$ 값의 범위를 가지게 하기 위해 양변에 로그 함수를 취한다.

$$\log\left(\frac{P(Y)}{1 - P(Y)}\right) = B_0 + B_1X$$

위 식을 자세히 보면, 우변은 선형모델이므로  $(-\infty, \infty)$  범위의 값을 가지고, 좌변 역시  $(-\infty, \infty)$  범위의 값을 가짐을 알 수 있다.

좌변의  $\log\left(\frac{P(Y)}{1-P(Y)}\right)$  형태를 로짓(logit) 함수라고도 부른다.

로지스틱 회귀분석은 주어진 훈련 데이터에서 목표변수  $Y$ 가 범주 값 1을 가질 확률  $P(Y=1) = P(Y)$ 를 위의 식의 로지스틱 함수를 이용하여 모형을 수립하고 모수들을 추정하는 알고리즘이라고 할 수 있다. 이러한 모수추정에는 일반적으로 최대우도추정법(Maximum Likelihood Estimation, MLE)을 사용하고, 이는 수식을 변형하는 등의 해석학적 방법으로 직접 계산이 어렵기 때문에 일정한 초기값을 부여하고 난 뒤에 반복적인 계산을 통해 값을 계속해서 조정해가는 수치계산적인 기법을 활용하게 된다. 그림 3-24는 이러한 로지스틱 함수의 그래프 모양을 나타낸다. 그래프 형태에서 알 수 있듯  $X$ 축은  $-\infty$ 부터  $\infty$ 까지의 값을 가지고 있으며  $Y$ 축의 경우 목표변수  $Y$ 의 발생확률이므로, 0에서 1 사이의 범위 값을 가지며 그 결과 함수인  $P(Y=1)$ 는  $X$ 값의 증가에 따라 0에서 1로 'S자 곡선'의 형태를 보인다는 것을 알 수 있다.

### (3) SVM(서포트 벡터 머신)

회귀분석이나 분류에 주로 사용된다. 두 가지의 카테고리 중 한쪽에 속하는 데이터의 집합이 주어지면, 주어진 해당 데이터 집합을 활용해 새로운 데이터는 둘 중 어느 쪽에 속할지에 대해 판단하는 분류 모델을 만든다. 적절한 커널 함수  $k(x,y)$ 를 정의한 SVM 구조를 설계한다,

데이터의 분류가 두 그룹으로 나누어질 때 활용되며, 데이터 그룹과 학습 데이터의 공간 사이의 거리를 가장 멀게 만드는 직선식을 찾아 모델을 만든다. 직선식으로부터 가장 가까운 거리에 있는 각 데이터 그룹의 데이터가 서포트 벡터이며, 서포트 벡터 머신 알고리즘은 직선식에서 서포트 벡터까지의 거리인 마진을 가장 크게 하는 직선식을 도출하여 두 개 데이터 그룹을 최적으로 분류한다. 따라서 공구의 상태를 정상/이상의 두 개 집단으로 분류하는 데에 서포트 벡터 머신이 적합하다고 판단되었고, 알고리즘의 유효성을 k-fold 교차 검증법을 통해 학습 및 검증하였다.

앞서 실험에서 획득한 가속도 및 스피들 로드 센서 데이터를 각각 서포트 벡터 머신 알고리즘의 학습 데이터로 사용하였으며 정상상태 및 이상상태 데이

터의 레이블을 각각 0과 1로 하여 학습하였다<sup>75)</sup>.

$$(가) 선형 : D = \{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{-1, 1\}\}_{i=1}^n$$

$$(나) 원형 : \operatorname{argmax}_{(w,b)} \frac{1}{2} \|w\|^2$$

단,  $y_i(w \cdot x_i - b) \geq 1$ , for all  $1 \leq i \leq n$

$$(다) 쌍대 : \tilde{L}(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i^T x_j = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{(i,j)} \alpha_i \alpha_j y_i y_j k(x_i, x_j)$$

$$\text{단, } \alpha_i \geq 0, \sum_{i=1}^n \alpha_i y_i = 0, \text{ for all } 1 \leq i \leq n.$$

#### (4) RF(랜덤포레스트)

양상불 학습 기법의 하나로, 분류, 회귀분석 등에 사용되며, 학습의 과정에서 구성된 여러 개의 결정나무로부터 분류, 검출, 혹은 회귀분석 결과(평균 예측치)를 출력해 동작하여 다양하게 활용되고 있다. 결정 트리 하나보다 좋은 성능을 보이며 안정적이고 강력하다. 데이터 스케일을 조정할 필요가 없으나 고차원 희소 데이터에는 부적합할 때가 있다.

랜덤포레스트는 배깅 알고리즘의 변형된 하나의 형태라고 할 수 있으며, 배깅과 동일하게 학습데이터를 부트스트래핑하여 랜덤한 서브셋을 다수 생성한다. 하지만 배깅이 모델을 만드는 데에 학습데이터의 모든 독립변수를 사용하는 것과 달리, 랜덤포레스트는 부트스트랩 시에 모델 생성에 쓰일 독립변수를 선택하는 과정까지도 수행된다는 차이가 있다. 이러한 과정에서 랜덤 서브셋이 다수 생성되면 각 서브셋에 대해 서로 독립인 서브트리 모델이 만들어지고, 이렇게 생성된 서브트리들은 합쳐져 포레스트를 이룬다. 분류 단계에서는, 배깅과 마찬가지로 각 서브트리들의 분류 결과에서 가장 많이 채택된 범주를 모델 전체의 예측결과로 선정하는 방식(Majority Voting Process)으로 분류가 이루어진다.

랜덤 포레스트에 가장 큰 영향을 미치는 매개변수들은 다음과 같다<sup>76)</sup>.

(가) 포레스트의 크기 (트리의 개수) : T, 최대 허용 깊이 : D

(나) 임의성의 정도와 종류 :  $|T_j|/|T|$   $\rho = |T_j|$

75) Vapnik, V. (2000). Section 5.6. Support Vector Machines.

76) Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

(다) 노드분할 함수의 선택 :  $h(v, \theta_j) \in \{0, 1\}$

(매개변수  $\theta = (\phi, \psi, \gamma)$ 에 따라 결정)

노드분할 함수가 선형적일 때  $h(v, \theta_j) = [\gamma_1 > \phi(v) \cdot \psi > \gamma_2]$

노드분할 함수가 비선형일 때  $h(v, \theta_j) = [\gamma_1 > \phi(v)' \psi \phi(v) > \gamma_2]$

(라) 훈련목적함수의 정의 :  $\theta_j^* = \underset{\theta_j \in \gamma_j}{\operatorname{argmax}} I_j,$

$$I_j = I(S_j, S_j^L, S_j^R, \theta_j) = H(S_j) - \sum_{i \in \{L, R\}} \frac{S_j^i}{S_j} H(S_j^i).$$

(마) 실제 적용에서의 특징벡터의 선택

## (5) XGBoost

XGBoost는 의사결정나무 기반의 앙상블 방법론으로 Boosting 기법을 이용하여 분산 환경에서도 실행할 수 있도록 구현할 수 있어 빠르며, 회귀와 분류 문제를 모두 지원하며 시각화 기능으로 활용되고 있다.<sup>77)</sup>

$$y = m_1(x) + \text{error}(x) \quad \text{error}(x) = m_2(x) + \text{error}_2(x)$$

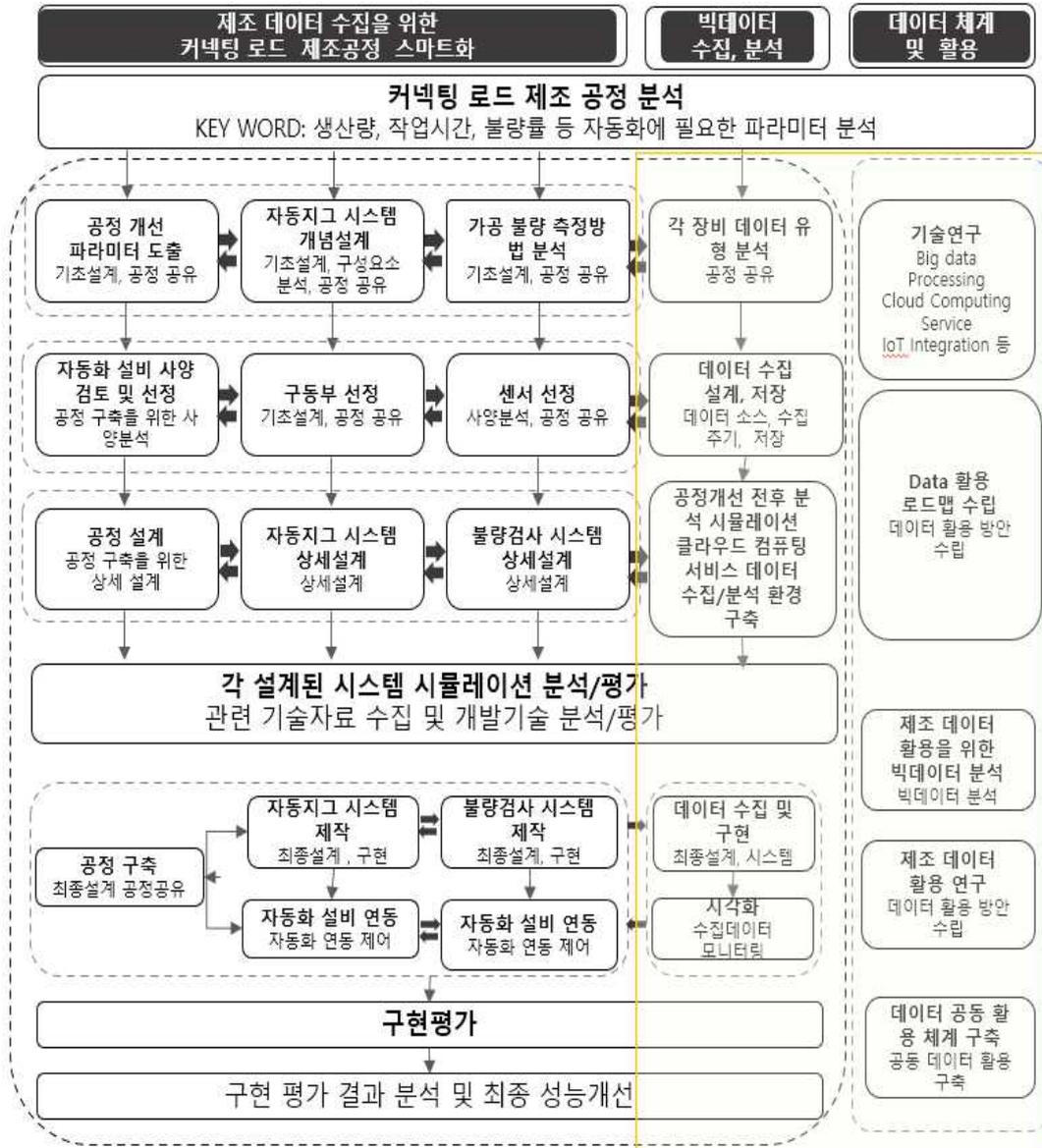
$$y = m_1(x) + m_2(x) + m_3(x) + \text{error}_3(x)$$

$$y = W_1 m_1(x) + W_2 m_2(x) + W_3 m_3(x) + \text{error}_3(x)$$

---

77) Chen, T., Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining pp. 785-794.

## 제 5 절 제조데이터 수집분석 및 활용 시스템 종합 설계



[그림 3-27] 제조데이터 활용 시스템 아키텍처

제조데이터 활용 시스템의 종합설계 아키텍처는, 그림 3-27과 같이 제조공정 스마트화의 단계와 빅데이터 수집분석 및 데이터 체계의 활용 단계로 크게 구분할 수 있고 상세 내용은 아래와 같다.

제조공정 스마트화 단계에서는 생산성, 작업시간, 불량률 등 공정개선을 위한 파라미터를 도출하여 자동지그 시스템 및 자동화 설비의 사양 검토 및 선정을 수행하고 자동화 공정과 자동지그 시스템의 상세 설계로 진행된다. 공정개선

파라미터 도출과 자동지그 시스템 개념 설계에서 가공불량의 측정방법에 대한 분석을 정립한다. 이어 불량 측정 방법을 위한 구동부에서 스마트센서를 선정하고 사양을 분석한 다음 불량시스템을 상세 설계한다.

빅데이터 수집분석 및 데이터 체계의 활용단계에서는 스마트화 공정 설비에서 발생하는 대상 데이터의 유형을 분석하고 스마트 센서 및 IoT 통합에 기반을 둔 추출 데이터의 소스, 수집, 저장 등 빅데이터 수집, 저장, 및 처리 시스템을 설계한다. 공정개선 전후 분석을 위한 예측(시뮬레이션)이나 외부 클라우드 서비스까지 연계하여 빅데이터 수집분석 시스템으로의 개발을 위한 데이터활용 방안을 마련해 데이터활용 로드맵을 수립한다.

빅데이터 수집 및 분석 시스템의 구현단계에서는 설계된 시스템과 기술을 분석 및 평가한다. 자동화 설비와 측정장비를 제작 및 구축해서 데이터 수집, 처리 및 시각화 분석을 진행하면서 제조활용을 위한 빅데이터 분석과 활용방안을 수립한다. 마지막으로 빅데이터 수집 분석시스템의 구현 평가를 통해서 최종 성능개선과 함께 데이터 공동활용 체계를 구축한다.

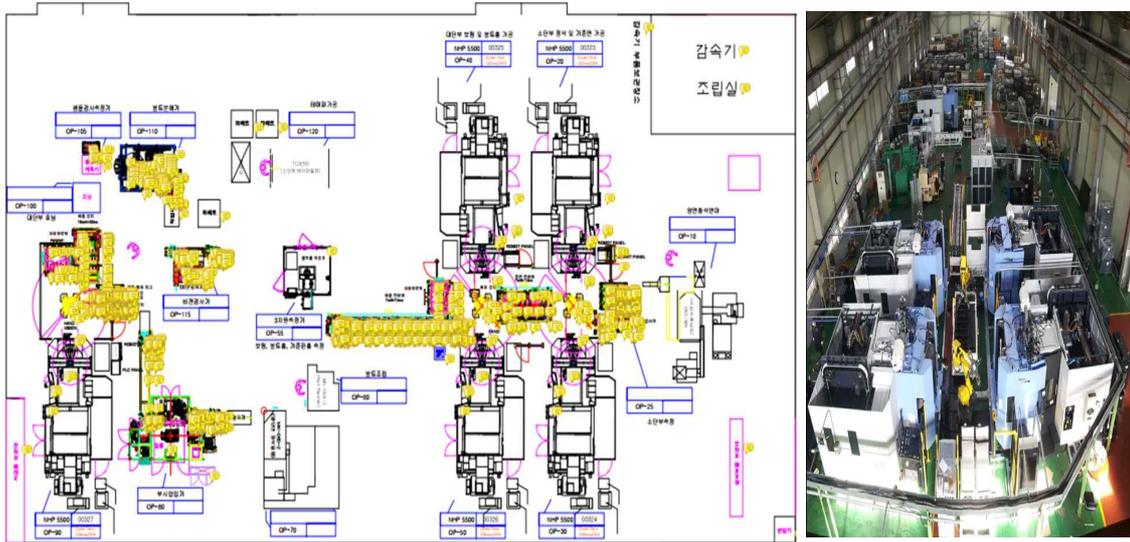
제조데이터 수집분석 및 활용시스템 종합 설계의 주요 내용을 요약해 보면 아래와 같다. 첫째, 수집 가능한 제조·품질 데이터를 정의하고 연계공정을 분석한다. 생산현장의 가공·품질 데이터를 취득, 전송, 공동보관 및 활용하거나 통합 분석 및 유사데이터 비교 등과 같은 활용방안을 도출 및 적용한다. 둘째, 데이터를 수집, 저장·분류하면서 공동으로 분석하고 활용한다. 데이터 측정시스템의 데이터 수집 방안을 도출하고 저장·분류함으로써 데이터 간의 연계 포인트를 공동으로 분석 및 활용한다. 셋째, 빅데이터 분석 결과를 기반으로 품질 및 생산성 향상을 위한 대책안을 제시한다. IoT 통합 기반의 제조공정 데이터 수집 및 관리 SW개발, 클라우드 컴퓨팅 서비스 기술지원, AI 학습모델링과 통계분석 기반 데이터 분석기법 등을 활용해서 방안을 마련한다.

# 제 4 장 제조데이터 수집분석 및 활용 시스템 구현

## 제 1 절 스마트화 가공공정 구현

### 1) 제조공정 스마트화를 위한 공정 레이아웃 재배치 및 개선 구현

가공공정 개선 후 레이아웃 설계 및 재배치 구현은 그림 4-1과 같다.



[그림 4-1] 가공공정 개선 후 레이아웃 설계 및 재배치 구현

### 가) 스마트화 라인 신규 공정 구현



[그림 4-2] 신규 MCT 도입 및 로봇, 이송장치 설치 현황

신규 수평 MCT 5대 도입 및 로봇, 이송장치를 설치한 신규 공정은 4-2와 같다.

나) 신가공 공정 도입을 위한 지그 설계 및 도입 구현

품질 및 생산성 향상을 위해 아래 표 4-1과 같이 Con-Rod 가공을 개선하기 위해, 공정별 가공 범위 등을 반영하여 지그를 설계하고 제작하는 개념으로 신 공정을 도입하여 신설 스마트화 라인과 연계하여 추진하였다.

[표 4-1] 개선공정 정의 및 가공 범위

공정	개선공정 정의	가공 범위	비 고
OP-20	소단부 정삭 및 기준면 가공	6개 공정, 소단경 4, 경사면 입구, 볼트홀 후면	*가공 예상시간 :2.5분/개 (가공시간:2.0, 비절삭시간:0.5) *스핀들 사양 : HSK-100A
OP-30	대단부 자리파기 및 절단	5개 공정, 볼트 좌면 2, 볼트홀 입구, Pin A/S홀, CAP절단	
OP-40	대단부 볼링 및 볼트홀 가공	14개 공정, 결합면 황정삭, Deburing부, 볼트홀, Pin홀, 대단경	
OP-50	조립면, 기준핀 및 세레이션 가공		
OP-90	대소단부 정삭가공	6개 공정, 소단경 측면, 대소단경 증삭, 정삭	

2) 부시 치수 및 압력 측정, 조립 시스템 구축

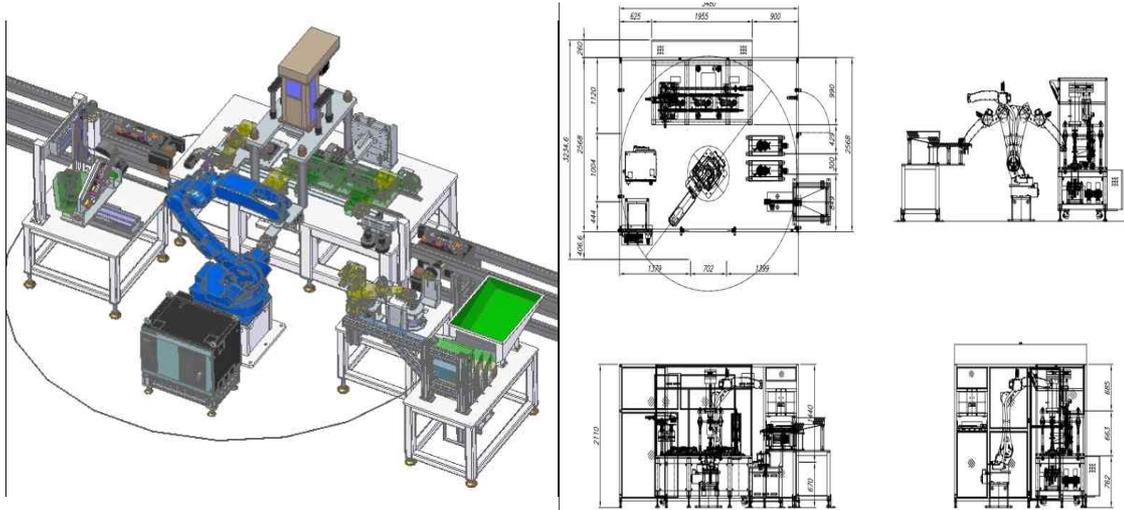
부시 압입압력 측정시스템은 제3장 2절 2)의 설계 개념으로 사용자의 요구 사항을 반영해 시스템을 구성하여 구현하였다. 그림 4-3과 같이 측정시스템의 주요 기능으로서 레이저 마킹문자로 제품시리얼 번호를 식별하여 내경측정치, 오차, 부시 최종거리, 압입압력 등 생산투입량 대비 합부수량을 모니터링하면서 운영되고 있다. 하드웨어 주요 기능으로는 부시 공급 컨베이어 제품감지, 커넥팅 로드 얼라인 기능, 부시 압입 상승 및 하강, 레이저 마킹 및 로봇 클램프, 컨베이어 공급 및 배출 기능이 동작하고 있다.



[그림 4-3] HMI 주요화면

가) 부시 치수, 압입압력 측정 및 조립시스템 제작 설계 구현

그림 4-4와 같이 조립체 도면과 레이아웃을 설계하여 제작, 설치하였다.



[그림 4-4] Con-Rod 가공공정 개선 후 설계 레이아웃

나) 부시 치수, 압입압력 측정 및 조립시스템 제작 구현

그림 4-9와 같이 부시 압입 시스템이 로봇, 컨베이어 장치와 연계되어 설치 후 가동되고 있다.



[그림 4-5] 부시 조립체 모습

3) 공구 모니터링 및 검사시스템 구축

실시간 가공품질 측정을 위한 생산 라인용 측정시스템을 표 4-2와 같이 사용자의 요구사항을 반영하여 구현하였다.

[표 4-2] 공구 모니터링 및 검사시스템 주요 구현 기술 및 사양

항 목	구현 기술 및 사양
측정 및 검사시스템	1) 측정범위 : 400mm(X축) x 300mm(Y축) x 200mm(Z축) 2) 측정내용 : 대/소단부, Bolt Hole, Pin Hole & Serration의 위치 및 경 3) 장비구성 : Measuring Bed, Main column Cross column & 측정 Software, Data 저장장치, 반입/반출 장치 4) 측정 Data 실시간 전송
Con-Rod 최적 절삭 Solution	1) 최적 절삭가공 조건 검토 및 설정 : Milling, Drilling, Boring & Reaming 등 2) Bolt 체결면 Milling Solution : 가공 시 절삭부하 최소화 및 가공시간 단축을 위해 사전 Drilling 공정 적용 3) Ø7 Pin service Hole Drilling Solution : 코너 R2.5부 가공 시 입구부 5mm 구간 회전수/절삭 피드 1/2↓, 내부쿨란트 off, 이후 정상조건으로 가공 4) 소단경 Finish Boring Solution ->진원도: 0.005 : 회전수/절삭피드는 100%, 내부쿨란트 off, 장비 본체의 툴 매거진 및 Chip 배출 Screw 정지 5) S5 기종 Bolt 체결 Hole의 Reaming Solution : 내부 Coolant Hole 적용으로 Chip 배출 용이, Shank 중간경 축소로 Chip 배출 공간 확보 등 절삭 Solution 구축

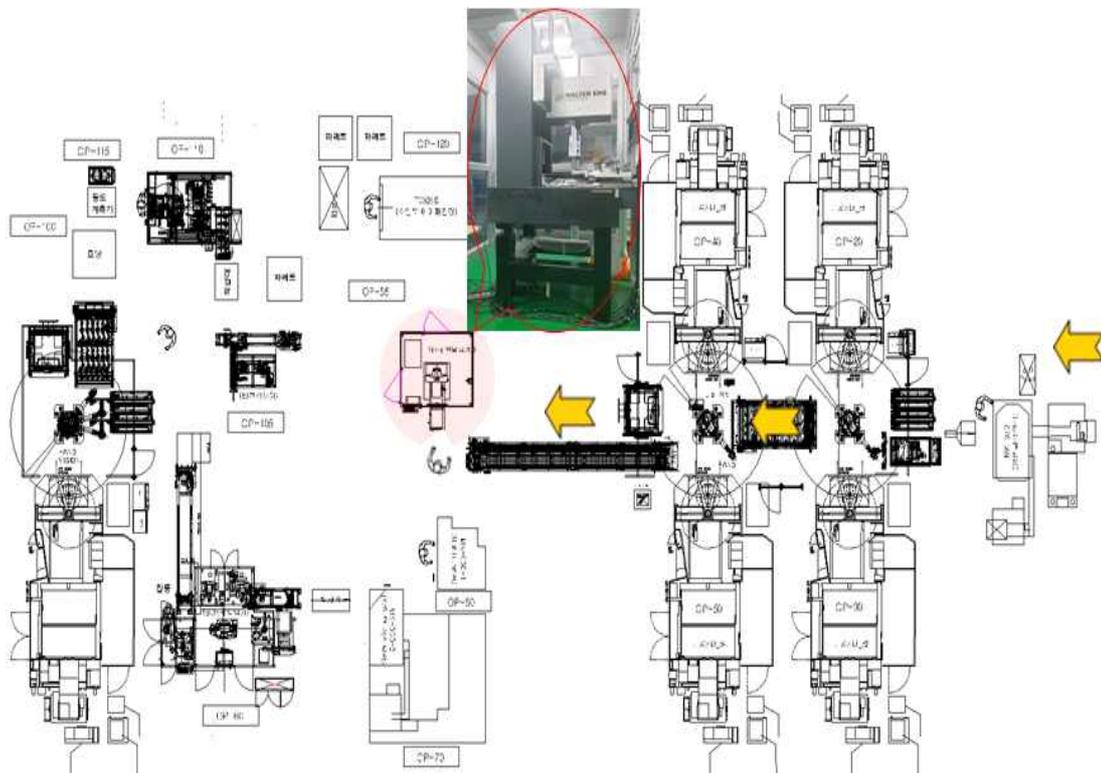
가) 공구 모니터링 및 검사시스템 설계 구현

제3장 2절 3)의 설계 개념에 따라 그림 4-7과 같이 공구 모니터링 검사시스템을 설계하여 설치하였다. 절삭기계 가공 시 그림 4-6과 같이<sup>78)</sup> 스피들에 걸리는 전류 변동치를 추출한 후 절삭 부하치로 변환한 다음, PLC로 전송해서 절삭부하로 변환한다. 실시간 공구의 상태를 관리하는 시스템으로 전송하며 각 장비 본체에 시스템을 구성하고 공구별로 실시간 저장되며 조회가 된다, 공구의 절삭 부하치에 대한 최대, 최소, 평균치 등 상관관계 분석을 통한 최적 피삭재의 가공품질, 공구수명, 최적 절삭조건 등을 관리하고 있다.

78) Kim, J. op.cit., pp. 124.



[그림 4-6] 공구 모니터링 및 데이터 추출 시스템



[그림 4-7] 공구 모니터링 및 검사시스템 레이아웃

나) 공구 모니터링 및 검사시스템 제작 및 설치 구현

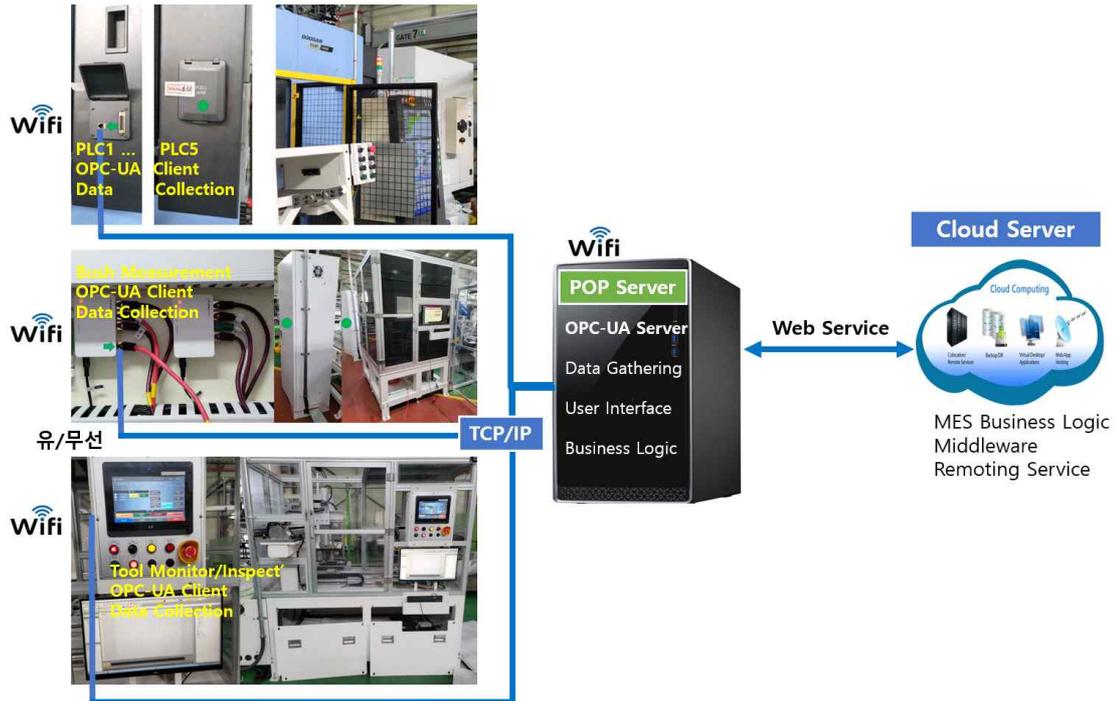
표 4-2, 그림 4-7의 설계에 의해 공구 모니터링 검사시스템이 그림 4-8과 같이 제작되어 설치 후 가동되고 있다.



[그림 4-8] 공구 모니터링 및 검사시스템 설치 및 가동 현황

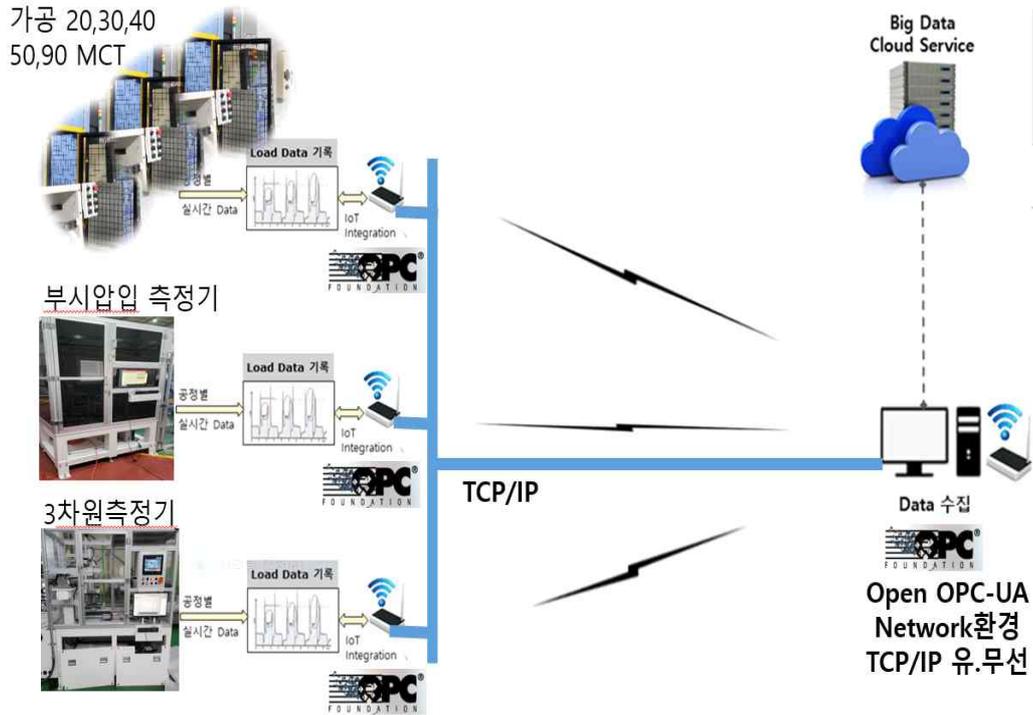
## 제 2 절 제조데이터 수집분석 및 활용 시스템 구현

### 1) 데이터 수집을 위한 OPC-UA 기반 IoT 통합 구현



[그림 4-9] OPC-UA 기반 통합 모듈 및 서버/클라이언트

OPC-UA를 기반으로 TCP/IP, 이더넷 등의 유무선 네트워크로 구성된 통신 모듈을 활용하고 공정 및 측정장비의 점점 IoT 통합 모듈을 개발하였다. 통신 방식 유연성을 적용해서 실시간 제조데이터 획득시스템을 구성한 다음, 그림 4-9는 IIoT 디바이스 개념으로 OPC-UA클라이언트와 데이터컬렉션으로 구성하였다. 신규 도입 중인 절삭가공설비, 부시 압입압력 측정장비, 공구 모니터링 및 검사시스템에서 발생한 데이터를 OPC-UA 기반 IoT 통합모듈에서 POP 서버로 인터페이스하여 데이터를 수집한다. 수집된 데이터는 클라우드 서버에서 MES에 저장 및 가공되어 SPC 통계분석이 수행되고 AI를 활용한 데이터 수집분석 시스템까지 영역을 확장해서 구현된다.



[그림 4-10] 데이터 수집 및 분석 시스템 네트워크

그림 4-10과 같이 OPC-UA기반을 둔 실시간 데이터 수집 분석시스템의 네트워크를 구현하였다. 절삭가공 정보로는 각 설비의 PLC에서 치수 변화량, 공구 부하량, 공구수명, 절삭부하, 공구 등 가공데이터를 획득하고, 부시 압입압력 측정기는 부시내경 치수와 압입 토크를, 공구 모니터링 및 검사 시스템은 황삭과 정삭가공 치수 데이터를 수집한다.

## 2) 데이터 수집 SW 구현

데이터 수집 소프트웨어는 3계층 구조이고 MES에서 스마트 클라이언트 개념의 애플리케이션 브라우저 설계를 적용하였다. DB 서버, 애플리케이션 서버, 웹서버 3개의 레이어로 구성하여 구현하였다.



[그림 4-11] MES 스마트 클라이언트 애플리케이션 브라우저

그림 4-11과 같이 해당 LKLab-MES 아이콘이 서버에 접속하는 스마트 클라이언트 애플리케이션 브라우저이다. 바탕화면의 LKLAB-MES 아이콘을 클릭하면 로그인할 수 있는 메뉴가 나온다.

### 3) 데이터 수집 및 분석시스템 개발 구현

#### 가) 분석데이터 수집, 데이터베이스 및 추출 프로그램 개발 구현

```
public DataSet GetbyWorkDate(DateTime fromDate,
DateTime toDate, string SPINDLEID, string TOOLID)
{
    string sql = "";
    // Sql Command
    System.Data.IDbCommand sqlSelectCommand =
EnterpriseManager.DbFactory.CreateDbCommand();
    sql +=
@"
SELECT *
FROM TBL_FANUC_DOOSAN_SPINDLE
WHERE dt_insert BETWEEN @fromDate and @toDate
and spindle_id like '%";
    if (SPINDLEID != "")
        sql += SPINDLEID;
    sql += "' and spindle_tool_no like '%";
    if (TOOLID != "")
        sql += TOOLID;
    sql += "' order by spindle_id, fanuc_doosan_id, dt_insert";

    sqlSelectCommand.CommandText = sql;

    sqlSelectCommand.CommandType = System.Data.CommandType.Text;

    System.Data.IDataParameter fromDateParam =
EnterpriseManager.DbFactory.CreateDataParameter("@fromDate", typeof(DateTime));
    fromDateParam.Value = fromDate;
    sqlSelectCommand.Parameters.Add(fromDateParam);
    System.Data.IDataParameter toDateParam =
EnterpriseManager.DbFactory.CreateDataParameter("@toDate", typeof(DateTime));
    toDateParam.Value = toDate;
    sqlSelectCommand.Parameters.Add(toDateParam);
    //
    // Call command
    DataSet dataSet = this.GetDataSet(sqlSelectCommand);
    sqlSelectCommand.Dispose();
    return dataSet;
}
```

[그림 4-12] 데이터쿼리 소스코드 일부

그림 4-12는 측정데이터 데이터베이스 쿼리 프로그램의 일부이다.

```

using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Drawing;
using System.Data;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows.Forms;

namespace Erp.Inspection.InspectionSheet.DongLim
{
    public partial class FANUCSPINDLE : UserControl
    {
        public FANUCSPINDLE()
        {
            InitializeComponent();
            Init();
        }

        private void Init()
        {
            fpSpread1.Sheets[0].AutoFilterMode
            = FarPoint.Win.Spread.AutoFilterMode.EnhancedContextMenu;
            dateTimeFromDate.Value = DateTime.Now.Date;
            dateTimeToDate.Value = DateTime.Now.Date;
        }

        . . .

        private void textBoxTool_KeyDown(object sender, KeyEventArgs e)
        {
            if (e.KeyCode == Keys.Enter)
                buttonSearch.PerformClick();
        }

        private void textBoxTool_KeyPress(object sender, KeyPressEventArgs e)
        {
            if (!(char.IsDigit(e.KeyChar) || e.KeyChar == Convert.ToChar(Keys.Back)))
            {
                e.Handled = true;
            }
        }
    }
}

```

[그림 4-13] 데이터 리포팅 소스코드 일부

그림 4-13은 측정데이터 데이터베이스 프로그램 일부이다.

```

/// <summary>
/// Chart-Xbar (Xbar-Chart의 Y축값.)
/// </summary>
private void ApplyChartXbar(int decimalPlaces)
{
    decimal maximum = 1;
    decimal minimum = 0;
    if (sheetView.Cells["X-UCL"].Text.Trim() != "#DIV/0!" && sheetView.Cells
        ["X-UCL"].Text.Trim() != "#REF!" && sheetView.Cells["X-UCL"].Text.Trim() != "")
    {
        maximum = Convert.ToDecimal(sheetView.Cells["X-UCL"].Value);
        minimum = Convert.ToDecimal(sheetView.Cells["X-LCL"].Value);
    }

    decimal range = maximum - minimum;
    decimal increment = range / 2;
    switch(decimalPlaces)
    {
        case 0:
            ...
        case 1:
            ...
        case 2:
            ...
        case 3:
            ...
        case 4:
            if (increment < 0.001m)
                increment = 0.001m;
            else
                increment = decimal.Round(increment, 3);
            maximum = decimal.Round(maximum, 3);
            minimum = decimal.Round(minimum, 3);
            max = maximum + increment * 2;
            min = minimum - (increment * 2);
            break;
    }

    foreach (FarPoint.Win.Spread.Chart.SpreadChart chart in
        fpSpread1.ActiveSheet.Charts)
    {
        if (chart.ChartName == "Chart-Xbar")
        {
            FarPoint.Win.Chart.YPlotArea yPlotArea =
                (FarPoint.Win.Chart.YPlotArea)chart.Model.PlotAreas[0];
            FarPoint.Win.Chart.ValueAxis valueAxis =
                (FarPoint.Win.Chart.ValueAxis)yPlotArea.YAxes[0];
            valueAxis.Minimum = Convert.ToDouble(min);
            valueAxis.Maximum = Convert.ToDouble(max);
            valueAxis.MajorUnit = Convert.ToDouble(increment);
            valueAxis.MinorUnit = Convert.ToDouble(increment);
        }
    }
}

```

[그림 4-14] SPC 분석 소스코드 일부

그림 4-14는 측정데이터 검색 조회 프로그램의 일부이다.

나) 분석데이터 수집, 저장 데이터베이스 및 추출 솔루션 구현

분석데이터 수집, 저장 데이터베이스 및 추출 솔루션에서 추출된 데이터로 MES의 TQC애플리케이션을 활용하여 SPC통계, AI모델링 분석 및 시각화를 할 수 있다. 그림 4-15~17과 같이 PoP 서버에서 분석대상 데이터를 MES의 DB(클라우드)에 수집, 저장한다.

The screenshot displays the '3D MEASURE' software interface. At the top, there are navigation tabs for '시스템', '기초정보 관리', 'BOM 관리', '작업지시 관리', '공정 관리', '생산실적 관리', '품질 관리', '입고/출고', '재고 관리', '설비 관리', and '청(W)'. Below this, the main window shows a data table for '3D MEASURE' with a filter for '작업일자' (Work Date) from 2021-10-01 to 2021-11-11 and a '기준' (Reference) of 265. The table has multiple columns, including '작업일시' (Work Time), '작업일' (Work Day), '소단량' (Sub-quantity), '대단량' (Main quantity), '불투출 위치 좌' (Defect location X), and '불투출 위치 우' (Defect location Y). The data rows contain numerical values for these parameters across various job numbers and dates.

[그림 4-15] MES 내 커빅팅로드 측정검사 DB 및 추출 솔루션

시스템 기초정보 관리 BOM 관리 작업지시 관리 공정 관리 생산실적 관리 품질 관리 입  
실행 프로그램 FANUC SPINDLE

FANUC SPINDLE(5.10.81.124)

**FANUC SPINDLE**

변경일자 2021-09-01 ~ 2021-11-11 스펀들 ID TOOL 번호

스핀들 ID	생선 ID	TOOL 번호	변경 시간	TOOL 번호	MAX 회전수	MAX 피드	부하(평균)	부하(최대)
1	80665	16228	2021-09-30 오전 12:16:00	15	1592	103923	0	10
2	80666	16228	2021-09-30 오전 12:16:00	5	835	92951	21	264
3	80667	16228	2021-09-30 오전 12:17:00	7	2871	103923	1	35
4	80668	16228	2021-09-30					
5	80669	16228	2021-09-30					
6	80670	16229	2021-09-30					
7	80671	16229	2021-09-30	10	26160	62	68429	79
8	80672	16229	2021-09-30	264	50870	62	133063	79
9	80673	16229	2021-09-30	35	28919	62	75641	79
10	80674	16229	2021-09-30	21	38160	62	99811	79
11	80675	16229	2021-09-30	27	8751	62	22894	79
12	80676	16230	2021-09-30	10	17324	32	30866	74
13	80677	16230	2021-09-30	37	12333	32	21972	74
14	80678	16230	2021-09-30	348	21181	32	37734	74
15	80679	16230	2021-09-30	23	24672	32	43954	74
16	80680	16230	2021-09-30	13	25893	32	46127	74
17	80681	16231	2021-09-30	168	10185	32	18141	74
18	80682	16231	2021-09-30	27	12583	17	26259	50
19	80683	16231	2021-09-30	233	12509	17	26101	50
20	80684	16232	2021-09-30	38	41360	17	86300	50
21	80685	16232	2021-09-30	40	16346	17	34106	50
22	80686	16232	2021-09-30	30	13426	17	28011	50
23	80687	16232	2021-09-30	31	16942	17	46151	72
24	80688	16232	2021-09-30	210	12394	17	33466	72
25	80689	16232	2021-09-30	221	26129	17	71163	72
26	80690	16233	2021-09-30	15	22355	62	58463	79
27	80691	16233	2021-09-30	319	50922	62	133174	79
28	80692	16233	2021-09-30	28	29571	62	77332	79
29	80693	16233	2021-09-30	20	38199	62	99894	79
30	80694	16233	2021-09-30	39	8760	62	22913	79
31	80695	16234	2021-09-30	12	17205	32	30641	74
32	80696	16234	2021-09-30	38	12349	32	21992	74
33	80697	16234	2021-09-30	366	21425	32	38154	74
136	10298	32	24416	32	43483	74	45147	34
27	12676	17	25925	32	46170	74	47939	34
208	12523	17	10298	32	16336	74	19037	34
43	40579	17	12676	17	26442	50	23166	37
		17	12523	17	26120	50	22883	37
		17	40579	17	84635	50	74145	37

[그림 4-16] MES 내 공구 부하량 측정 DB 및 추출 솔루션

시스템 기초정보 관리 BOM 관리 작업지시 관리 공정 관리 생산실적 관리 품질 관리 입  
실행 프로그램 FANUC PROCESS

FANUC PROCESS(5.10.81.124)

**FANUC PROCESS**

작업일자 2021-09-01 ~ 2021-11-11 공정번호

생선 ID	공정번호	제품 ID	시작시간	종료시간
1	16228	20	2021-09-30 오전 12:13:00	2021-09-30 오전 12:18:00
2	16229	30	2021-09-30 오전 12:15:00	2021-09-30 오전 12:20:00
3	16230	50	2021-09-30 오전 12:17:00	2021-09-30 오전 12:21:00
4	16231	40	2021-09-30 오전 12:18:00	2021-09-30 오전 12:22:00
5	16232	20	2021-09-30 오전 12:18:00	2021-09-30 오전 12:22:00
6	16233	30	2021-09-30 오전 12:20:00	2021-09-30 오전 12:25:00
7	16234	50	2021-09-30 오전 12:21:00	2021-09-30 오전 12:25:00
8	16235	90	2021-09-30 오전 12:16:00	2021-09-30 오전 12:26:00
9	16236	20	2021-09-30 오전 12:23:00	2021-09-30 오전 12:27:00
10	16237	40	2021-09-30 오전 12:22:00	2021-09-30 오전 12:27:00
11	16238	50	2021-09-30 오전 12:25:00	2021-09-30 오전 12:30:00
12	16239	30	2021-09-30 오전 12:25:00	2021-09-30 오전 12:30:00
13	16240	40	2021-09-30 오전 12:27:00	2021-09-30 오전 12:31:00
14	16241	90	2021-09-30 오전 12:26:00	2021-09-30 오전 12:31:00
15	16242	20	2021-09-30 오전 12:27:00	2021-09-30 오전 12:32:00
16	16243	50	2021-09-30 오전 12:30:00	2021-09-30 오전 12:34:00
17	16244	30	2021-09-30 오전 12:30:00	2021-09-30 오전 12:35:00
18	16245	40	2021-09-30 오전 12:31:00	2021-09-30 오전 12:36:00
19	16246	20	2021-09-30 오전 12:32:00	2021-09-30 오전 12:36:00
20	16247	90	2021-09-30 오전 12:31:00	2021-09-30 오전 12:36:00
21	16248	50	2021-09-30 오전 12:34:00	2021-09-30 오전 12:39:00
22	16249	40	2021-09-30 오전 12:36:00	2021-09-30 오전 12:40:00
23	16250	20	2021-09-30 오전 12:36:00	2021-09-30 오전 12:40:00
24	16251	30	2021-09-30 오전 12:35:00	2021-09-30 오전 12:40:00
25	16252	90	2021-09-30 오전 12:36:00	2021-09-30 오전 12:41:00
26	16253	50	2021-09-30 오전 12:39:00	2021-09-30 오전 12:43:00
27	16254	40	2021-09-30 오전 12:40:00	2021-09-30 오전 12:44:00
28	16255	90	2021-09-30 오전 12:41:00	2021-09-30 오전 12:45:00
29	16256	30	2021-09-30 오전 12:42:00	2021-09-30 오전 12:47:00
30	16257	20	2021-09-30 오전 12:40:00	2021-09-30 오전 12:48:00
31	16258	50	2021-09-30 오전 12:43:00	2021-09-30 오전 12:49:00
32	16259	90	2021-09-30 오전 12:45:00	2021-09-30 오전 12:50:00
33	16260	40	2021-09-30 오전 12:44:00	2021-09-30 오전 12:50:00

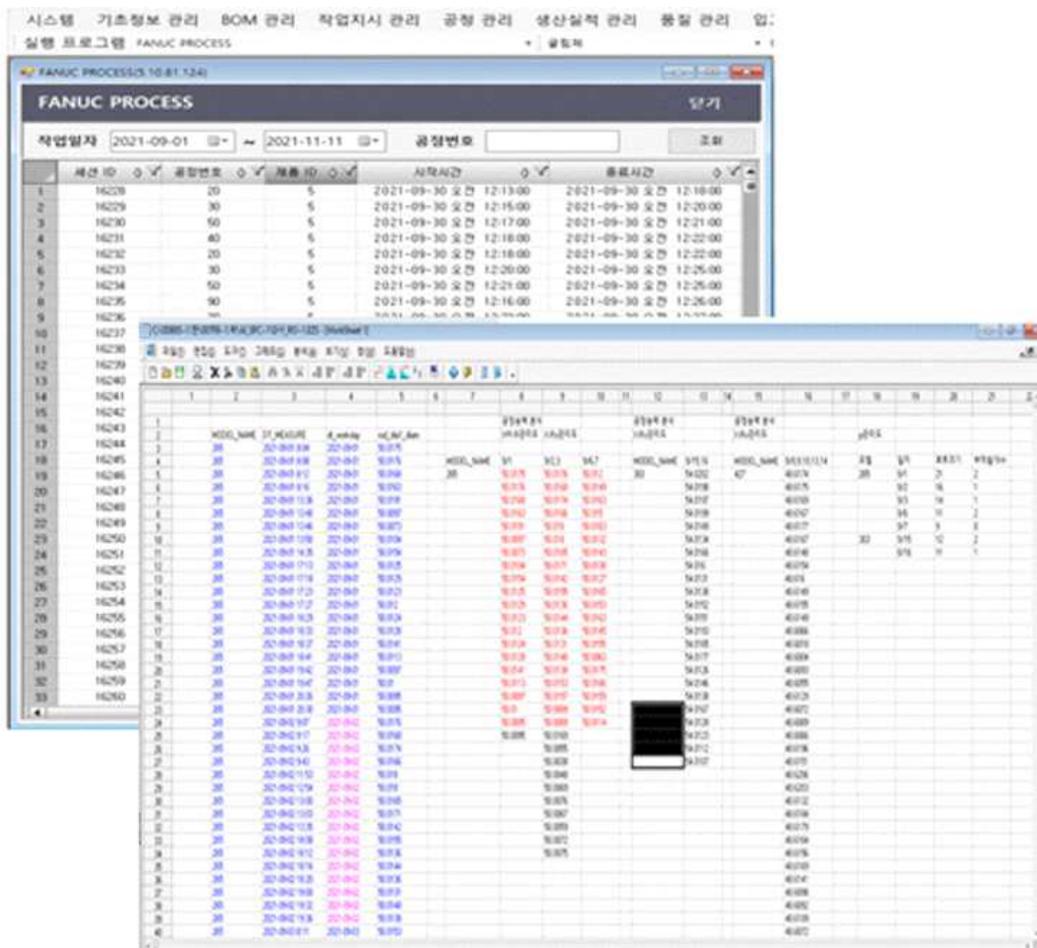
[그림 4-17] MES 내 공구사용 관리 DB 및 솔루션

### 제 3 절 SPC 통계 & AI 모델링 분석 및 시각화 구현

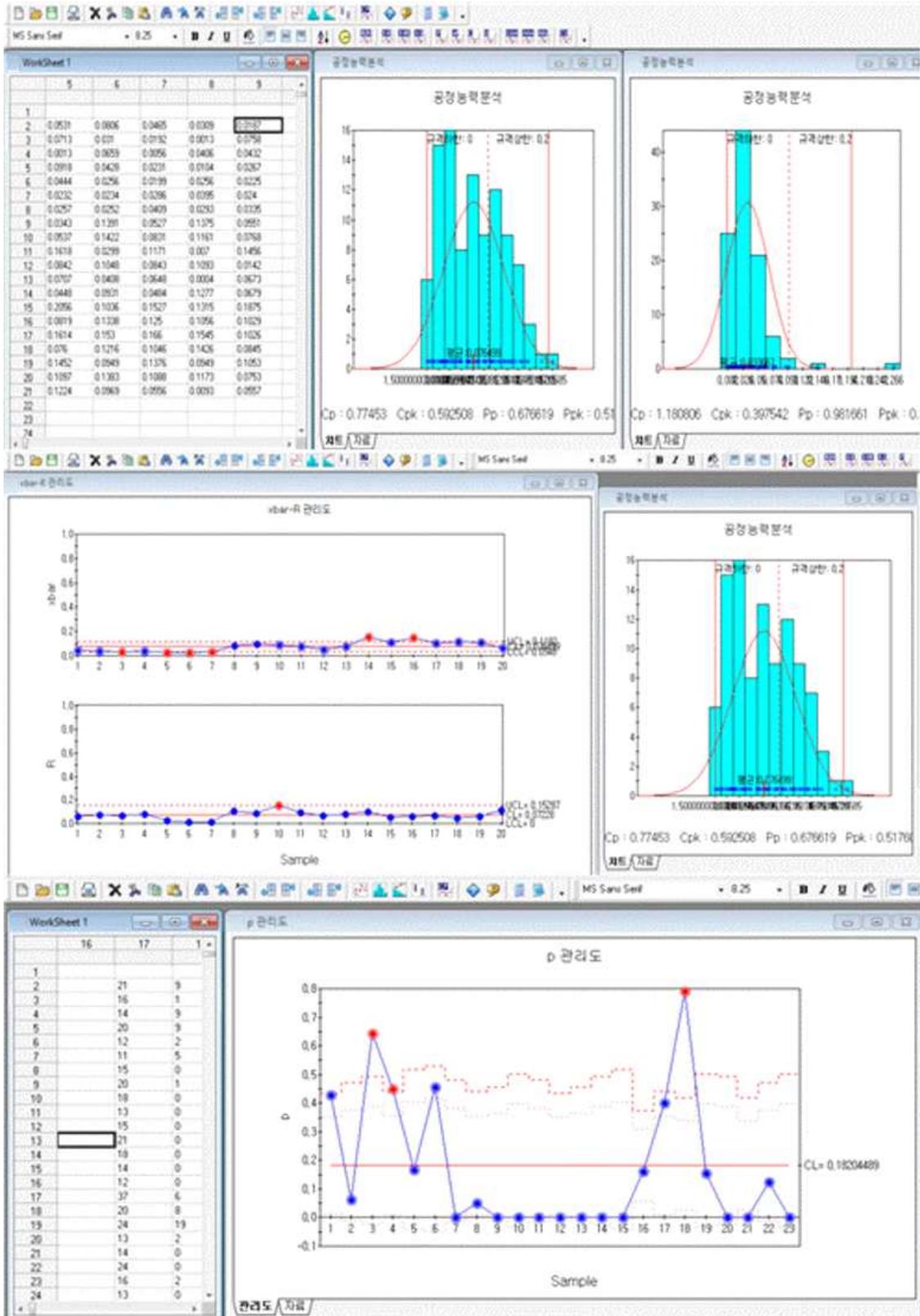
#### 1) SPC 통계 솔루션 및 시각화 구현

##### 가) MES솔루션 기반의 TQC 애플리케이션 SPC통계 모듈 구현

MES솔루션 기반의 TQC 애플리케이션으로 SPC통계를 구현하고 시각화 솔루션은 오픈소스와 FP(FarPoint) Spread를 활용하였다. 그림 4-18과 같이 MES DB에서 분석대상 데이터를 TQC 애플리케이션의 실시간 SPC 통계 프로그램으로 추출된다. SPC통계 및 시각화는 실제 공정능력지수(CPk)와  $\bar{X}$ -R관리도로 주로 사후 공정관리를 수행하고, 필요에 따라서 산점도, 선형회귀분석으로 간단한 예측분석을 한다. 제3장 3절 4)가와 같이 대표적인 수식을 적용하여 그림 4-19와 같이 SPC통계 모듈을 구성하였다.



[그림 4-18] MES DB의 추출 솔루션에서 SPC통계 모듈로 데이터 추출



[그림 4-19] SPC통계 모듈프로그램의 각종 통계 구현 이미지

## 2) AI 모델링 분석 및 시각화

XGBoost는 의사결정나무 기반의 앙상블 방법론으로 부스팅(Boosting) 기법을 이용하며 분산 환경에서도 실행할 수 있도록 구현할 수 있어 빠르며, 회귀와 분류 문제를 모두 지원한다. 선형회귀는 데이터의 분포를 가장 잘 설명하고 양품 생산을 위한 최적의 회귀모델을 도출하고, 각 설비의 값들의 최적치를 예측하는 데 사용한다.

### 가) 데이터 및 변수 선정

#### (1) 데이터 선정

절삭가공의 3가지(황삭-중삭-정삭) 공정에 대해 모든 절삭 조건을 다루는 것은 한계가 있으므로, 본 연구에서는 절삭가공의 3가지 공정에 의한 제품의 치수측정과 피드백을 통해 불량 발생을 사전에 방지하고 완성품의 치수측정 데이터 확보하여 수행하였다.

(가) 황삭 : Con-Rod 황삭 및 정삭품 치수(Rod & Cap 볼트홀 대칭도, 대단부 살두께, 내경 등등)

(나) 중삭 : 대단경 중삭 Boring & Back Chamfering, 소단경 중삭 Boring & Back Chamfering

(다) 정삭 : 대소단경 정삭 Reaming

학습에 사용된 데이터의 기간은 2021년 9월 01일부터 2021년 11월 05일까지이며, 약 440개의 테스트 시료를 활용하였다.

#### (2) 변수 선정

품질관리를 위한 합격/불합격, 공구수명 예측모형을 구축하기 위해, 본 데이터는 약 80여 개에 달하는 변수들에서 로짓 모형/판별분석 모형에 의하여 선택적 변수 선정 과정을 거쳐 공정여부와 관련된 유의한 분석변수를 추출하였다. 표 4-3~5와 같이 전체 440행의 데이터로 갖추었으며, 합격/불합격을 위한 총 10개의 변수와 공구수명을 위한 총 4개의 변수를 연구모델을 위한 독립변수로 활용하였다. 본 연구에서는 훈련용, 검증용 두 가지의 데이터셋이 구성되었으며 80%는 훈련용으로 나머지 20%는 검증용으로 사용하였다.

(가) 3차원 측정기 측정 결과; CON ROD CMM 검사성적서

파일 이름 : TBL\_3D\_MEASURE.csv

[표 4-3] 3차원 측정기 측정 결과; CON ROD CMM 검사성적서

구분	변수	설명	최소	최대		
ROD	기종	MODEL_NAME				
	검사일자	DT_MEASURE				
	일자	dt_workday				
	소단경	rod_dia1_diam	경			
		rod_dia1_cirly	진원도			
		rod_dia1_perp	직각도			
		rod_dia1_cylcty	원통도			
	대단경	rod_dia2_y	Y	59.6	+/-0.05	
		rod_dia2_x	X	2	+/-0.05	
		rod_dia2_diam	경	Ø74.3	+/-0.2	
		rod_dia2_dist	축간거리	232	+0.100 +0.000	
	볼트홀 위치 좌	rod_bhpl_y	Y			선택
		rod_bhpl_z	Z			
		rod_bhpl_diam	경			
	볼트홀 위치 우	rod_bhpr_z	Z			선택
		rod_bhpr_diam	경			
		rod_bhpr_y	Y			
		rod_bhpr_dist	홀간거리			
	좌우 볼트홀 대칭도	rod_lr_symm	좌우 볼트홀 대칭도			선택
	CAP 결합면 높이	rod_cap_hl	좌			
		rod_cap_hr	우			
핀홀 위치 좌	rod_phpl_y1	Y			선택	
	rod_phpl_z1	Z				
	rod_phpl_y2	Y				
	rod_phpl_z2	Z				
	rod_phpl_diam	경				
핀홀 위치 우	rod_phpr_y1	Y			선택	
	rod_phpr_z1	Z				
	rod_phpr_y2	Y				
	rod_phpr_z2	Z				
	rod_phpr_diam	경				
CAP	대단경	cap_dia2_y	Y			
		cap_dia2_z	Z			
		cap_dia2_x	X			
		cap_dia2_diam	경			
	볼트홀 위치 우	cap_bhpr_y	Y			선택
		cap_bhpr_z	Z			
		cap_bhpr_diam	경			
		cap_bhpr_perp	직각도			
	볼트홀 위치 좌	cap_bhpl_z	Z			선택
		cap_bhpl_y	Y			
		cap_bhpl_diam	경			
	좌우 볼트홀 대칭도	cap_lr_symm	좌우 볼트홀 대칭도			선택
	핀홀 위치 우	cap_phpr_y1	Y			선택
		cap_phpr_z1	Z			
		cap_phpr_y2	Y			
cap_phpr_z2		Z				
cap_phpr_diam		경				
핀홀 위치 좌	cap_phpl_y1	Y			선택	
	cap_phpl_z1	Z				
	cap_phpl_y2	Y				
	cap_phpl_z2	Z				
ROD 결합면	cap_rod_hl	좌				
	cap_rod_hr	우				
합부판정	test_result					

(나) CON ROD가공 > 공구부하량

파일 이름 : TBL\_FANUC\_DOOSAN\_SPINDLE.csv

[표 4-4] CON ROD가공 > 공구부하량

구분	변수	설명	최소	최대	비고
	spindle_id		.	.	
	fanuc_doosan_id	세션 ID	.	.	
	dt_insert	TOOL 변경 시간	.	.	
	spindle_tool_no	TOOL 번호	.	.	
	max_rpm	MAX 회전수	.	.	
	max_feed	MAX 피드	.	.	
	avg_tool_load	부하(평균)	.	.	
	max_tool_load	부하(최대)	.	.	선택
	avg_x_torque	X 토크(평균)	.	.	
	max_x_torque	X 토크(최대)	.	.	선택
	avg_y_torque	Y 토크(평균)	.	.	
	max_y_torque	Y 토크(최대)	.	.	선택
	avg_z_torque	Z 토크(평균)	.	.	
	max_z_torque	Z 토크(최대)	.	.	선택
	tool_set_count	교환주기(설정값)	.	.	
	tool_current_count	교환주기(현재값)	.	.	

(다) 공구 사용 및 교환 이력

파일 이름 : TBL\_FANUC\_DOOSAN.csv

[표 4-5] 공구 사용 및 교환 이력

구분	변수	설명	최소	최대	비고
	fanuc_doosan_id	세션 ID	.	.	
	process_no	공정번호	.	.	
	product_id	제품 ID	.	.	
	dt_start	시작시간	.	.	
	dt_end	종료시간	.	.	
	dt_change_tool	.	.	.	
	dt_workday	.	.	.	

## 나) 데이터 설명

### (1) EDA(Exploratory Data Analysis)

EDA란, 탐색적 데이터 분석을 의미한다. 본격적인 데이터 분석 전, 기본적인 표나 그래프를 보고 경향 등을 파악하는 매우 중요한 초기 분석의 단계이다

파일 이름 : TBL\_3D\_MEASURE.csv

	MODEL_NAME	DT_MEASURE	dt_workday	rod_dia1_diam	rod_dia1_cirity	rod_dia1_perp	rod_dia1_cylcty	rod_dia2_y	rod_dia2_x	rod_dia2_diam	rod_dia2_dist	rod_bhpl_y	rod_bhpl_z
0	265	2021-09-01 08:04:00	2021-09-01	50.0175	0.0025	0.0053	0.0034	67.4139	210.5990	87.0695	274.9175	12.5086	-21.3212
1	265	2021-09-01 08:08:00	2021-09-01	50.0176	0.0028	0.0033	0.0029	67.4684	210.6437	87.0896	274.9759	12.5056	-21.4043
2	265	2021-09-01 08:12:00	2021-09-01	50.0168	0.0035	0.0011	0.0035	67.4591	210.6507	87.1014	274.9849	12.4990	-21.3698
3	265	2021-09-01 08:16:00	2021-09-01	50.0163	0.0030	0.0015	0.0030	67.4818	210.6603	87.0736	274.9975	12.5268	-21.3735
4	265	2021-09-01 13:36:00	2021-09-01	50.0191	0.0018	0.0064	0.0025	67.4361	210.6380	87.0683	274.9683	12.4911	-21.2142
5	265	2021-09-01 13:40:00	2021-09-01	50.0097	0.0066	0.0019	0.0068	67.4634	210.6585	87.0781	274.9951	12.4977	-21.2599

[그림 4-20] 분석을 위한 데이터 셋 가져오기

다음은 이상치 확인을 위한 알고리즘이다.

```
y_null=analysisData.iloc[:, :-1].isnull
```

```
if y_null==0:
```

```
    print("결측치가 없습니다.\n")
```

```
else:
```

```
    print("결측치가 있습니다.\n")
```

```
    print(y_null, "\n")
```

```
[ ] y_null=analysisData.iloc[:, :-1].isnull
```

```
[ ] if y_null==0:
      print("결측치가 없습니다.\n")
    else:
      print("결측치가 있습니다.\n")
      print(y_null, "\n")
```

결측치가 있습니다.

<bound method DataFrame.isnull of	MODEL_NAME	DT_MEASURE	...	cap_rod_hl	cap_rod_hr
0	265	2021-09-01 08:04:00	...	36.3829	36.3886
1	265	2021-09-01 08:08:00	...	36.4379	36.4278
2	265	2021-09-01 08:12:00	...	36.3881	36.3941
3	265	2021-09-01 08:16:00	...	36.4117	36.4070
4	265	2021-09-01 13:36:00	...	36.3911	36.4005
...	...	...	...	...	...
127	427	2021-09-14 09:38:00	...	43.5338	43.5584
128	427	2021-09-14 13:00:00	...	43.4219	43.4249
129	427	2021-09-14 13:04:00	...	43.4637	43.4652
130	427	2021-09-14 13:22:00	...	43.4868	43.5128
131	427	2021-09-14 13:28:00	...	43.4859	43.5116

[132 rows x 56 columns]>

[그림 4-21] 이상치 확인

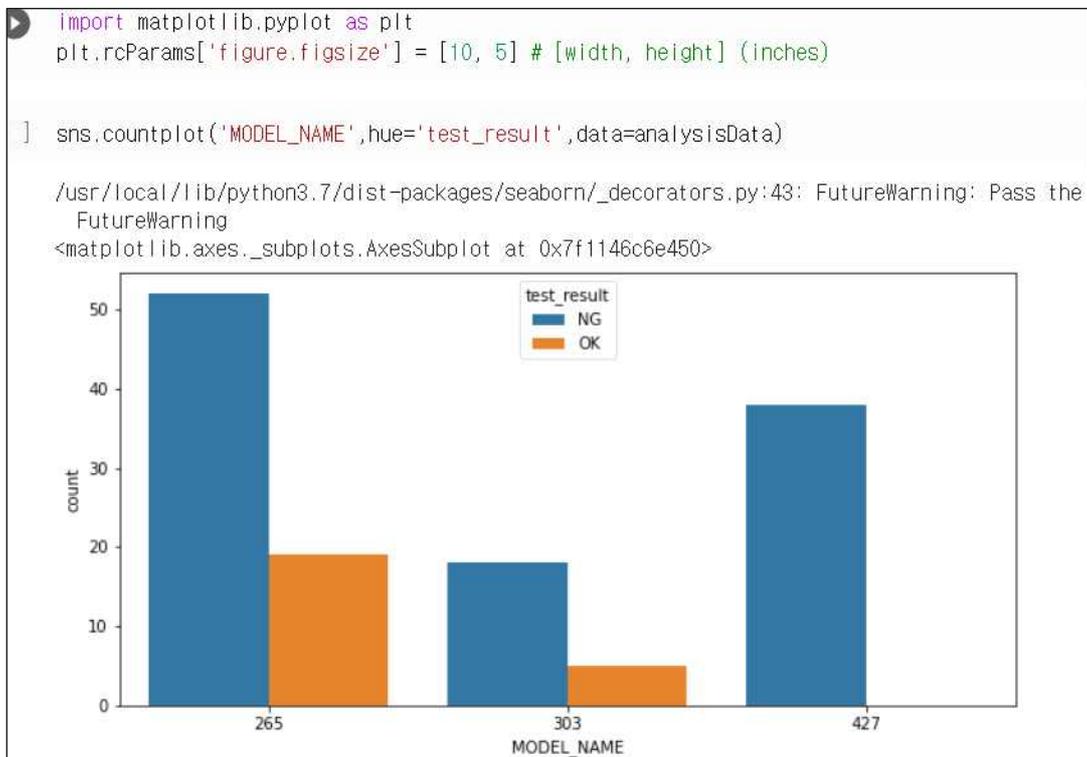
## (2) 정보 검색, 조회 기능 구현

### (가) 최대값, 최소값, 양품, 불량품 정보 검색

```
[ ] MachineType = 'MODEL_NAME' #@param {type:"string"}
SearchByMetric = 'rod_dia1_diam' #@param {type:"string"}
analysisData.groupby(MachineType)[SearchByMetric].describe()
```

	count	mean	std	min	25%	50%	75%	max
MODEL_NAME								
265	71.0	50.013115	0.003825	50.0038	50.010200	50.0141	50.01630	50.0191
303	23.0	54.015474	0.002904	54.0107	54.013250	54.0151	54.01720	54.0202
427	38.0	48.613024	0.005149	48.6004	48.608975	48.6149	48.61685	48.6206

[그림 4-22] 기종별(일자별) 측정항목별 평균, 최대, 최소 등 요약 정보



[그림 4-23] 양품/불량 수 차트로 나타내기

### (나) 특정일자, 기간 조회

특정일자, 기간에 대한 조회 및 소스코드는 부록 7을 참조할 수 있다.

(다) 공구 부하량 조회

공구번호 검색 시 시간, 일자, 월별 부하량 변동을 상태 조회할 수 있다. 공구부하량에 대한 조회 및 소스코드는 부록 8~11을 참조할 수 있다.

(3) 공구 수명 예측프로그램 개발

공구 마모 및 절삭력 비교를 위해, 공구동력계로 수집된 절삭력 데이터를 분석한다. 절삭부하는 공구 모니터링 시스템 통해 확보된 데이터로 분석하였으며, 가공 시 가공장비의 스피들에 걸리는 전류치를 추출하고 이것을 절삭부하로 변환하여 공구의 상태를 관리하는 시스템으로 각 장비 본체에 구성한다.

표 4-6은 변수의 정의와 설명을 나타내고 X는 특징값을 y는 라벨을 의미한다.

[표 4-6] 변수의 정의 및 설명

No	spindle_tool_no	TOOL 번호
X	max_tool_load	부하(최대)
	max_x_torque	X 토크(최대)
	max_y_torque	Y 토크(최대)
	max_z_torque	Z 토크(최대)
y	tool_current_count	교환주기(현재값)

공구 수명 예측프로그램에 대한 소스코드, 분석 및 시각화내용은 부록 12~18에서 참조할 수 있다.

(4) 정보검색, 공구수명 데이터 분석 및 예측 프로그램 유효성 검증

가속도 및 스피들로드리는 공구의 상태를 반영하는 특징적 요소로 알려져, 해당 데이터를 활용해 공구 상태 파악 및 진단하기 위하여 k-평균 군집, 서포트 벡터 머신 등의 기법들을 이용하였다. 이렇게 개발된 센서 데이터 및 기계학습 알고리즘을 이용한 공구 상태 진단 시스템의 유효성을 k-fold 교차 검증법을 통해 확인하였다.

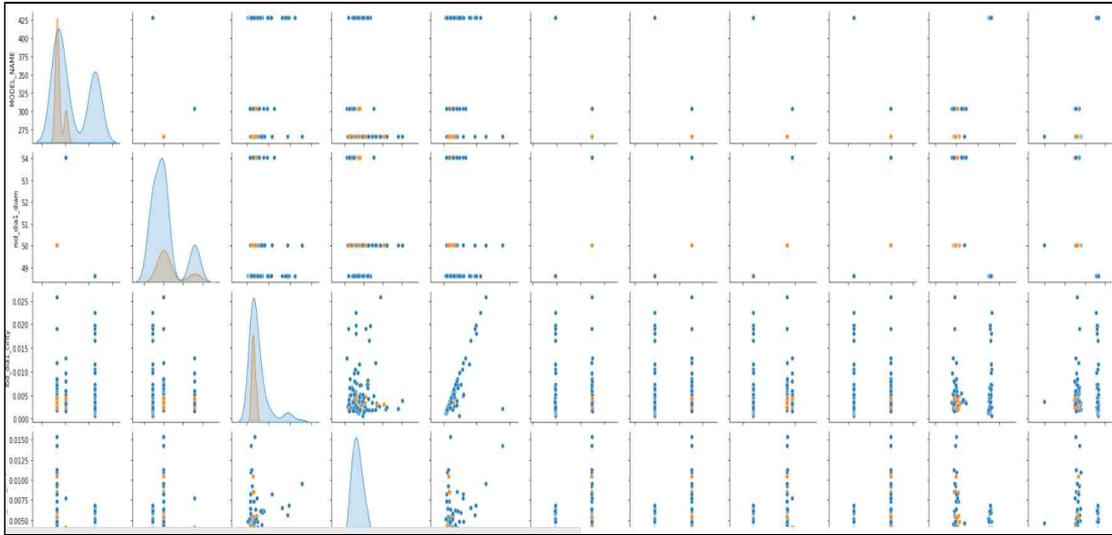
다) 변수 모델링 분석, 검증 및 시각화

(1) Cap/Rod 상관분석에 의한 독립변수의 상대적 중요도 비교

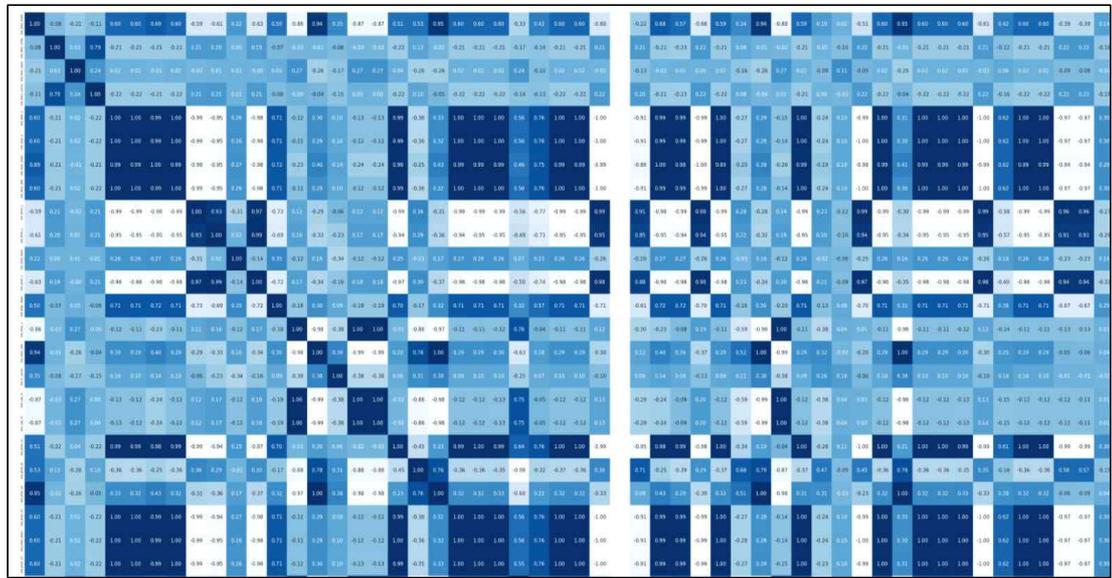
[표 4-7] 상관계수 분석

MODEL_NAME	rod_dia1_diam	rod_dia1_cirly	rod_dia1_perp	rod_dia1_cylcty	rod_dia2_y	rod_dia2_x	rod_dia2_diam	rod_dia2_dist
MODEL_NAME	1.000000	-0.427330	0.221178	-0.071658	0.217709	-0.979335	-0.980520	-0.951338
rod_dia1_diam	-0.427330	1.000000	-0.084549	-0.208313	-0.109966	0.601120	0.596575	0.685058
rod_dia1_cirly	0.221178	-0.084549	1.000000	0.033624	0.794478	-0.212453	-0.213891	-0.206337
rod_dia1_perp	-0.071658	-0.208313	0.033624	1.000000	0.242776	0.015338	0.018023	-0.012092
rod_dia1_cylcty	0.217709	-0.109966	0.794478	0.242776	1.000000	-0.215154	-0.216424	-0.212329
rod_dia2_y	-0.979335	0.601120	-0.212453	0.015338	-0.215154	1.000000	0.999934	0.993916
rod_dia2_x	-0.980520	0.596575	-0.213891	0.018023	-0.216424	0.999934	1.000000	0.993315
rod_dia2_diam	-0.951338	0.685058	-0.206337	-0.012092	-0.212329	0.993916	0.993315	1.000000
rod_dia2_dist	-0.980520	0.596574	-0.213891	0.018024	-0.216424	0.999934	1.000000	0.993315
rod_bhpl_y	0.972072	-0.588456	0.209800	-0.024821	0.208338	-0.990511	-0.990668	-0.983946
rod_bhpl_z	0.920025	-0.611181	0.196792	0.007685	0.212419	-0.949717	-0.949459	-0.949558
rod_bhpl_diam	-0.235680	0.224943	0.002582	0.008257	0.005409	0.259284	0.258231	0.266582
rod_bhpr_z	0.952642	-0.633011	0.194855	-0.001630	0.208062	-0.983633	-0.983189	-0.983425
rod_bhpr_diam	-0.678352	0.503307	-0.067236	0.045443	-0.084343	0.711979	0.711475	0.716709
rod_bhpr_y	-0.084568	-0.864716	-0.031030	0.269346	-0.000829	-0.118455	-0.112783	-0.226519
rod_bhpr_dist	-0.099413	0.942096	-0.009637	-0.255733	-0.039157	0.298358	0.292910	0.401125
rod_lr_symm	-0.021898	0.354723	-0.083622	-0.165533	-0.154015	0.095988	0.096297	0.137876
rod_cap_hl	-0.073258	-0.870358	-0.028482	0.268768	0.001694	-0.129711	-0.124048	-0.237554
rod_cap_hr	-0.072999	-0.870486	-0.028438	0.268775	0.001727	-0.129968	-0.124306	-0.237907
rod_phpl_y1	-0.995221	0.513465	-0.219159	0.044301	-0.218696	0.994362	0.994990	0.976836
rod_phpl_z1	0.539514	0.530671	0.129725	-0.260965	0.102893	-0.358287	-0.363635	-0.253867
rod_phpl_y2	-0.130069	0.949155	-0.020537	-0.257394	-0.048265	0.327142	0.321664	0.428253
rod_phpl_z2	-0.980373	0.596602	-0.213679	0.018505	-0.216010	0.999800	0.999873	0.993197
rod_phpl_diam	-0.980766	0.595571	-0.214487	0.018945	-0.216900	0.999927	0.999995	0.993170

일반적으로 공구 마모의 진행에 따라 절삭력, 가속도, 스피들로드르는 증가한다는 점이 알려져 있다. 또한 본 연구에서 진행한 두 차례의 반복실험에서 볼트 홀 위치와 핀홀 위치 값의 상관계수가 각각 0.994362 등으로 높은 상관관계를 보였으며, 공구의 상태를 반영하는 특징적인 변수로 활용될 수 있다는 점이 검증되었다. 다음 산포도와 히트맵은 위 내용을 보충 설명해 준다.



[그림 4-24] 변수별 산포도



[그림 4-25] 변수별 히트맵

(2) 알고리즘 예측 정확도 검증

로지스틱 회귀분석, 서포트 벡터 머신(SVM), 랜덤포레스트(RF) 알고리즘을 적용해 1, 2차를 걸쳐 실시하였으며 상세내역은 제4절 예측정확도 성능평가에서 기술하였다.

(가) 1차 예측 정확도

9월 시범생산 대상 중 9월 1~14일의 132개 데이터로 로지스틱회귀분석, SVM, RF를 적용해 1차 예측 정확도를 산정하였다.



### ③ 랜덤포레스트 예측 정확도

```
X=analysisData_pred.iloc[:,0:-1]
y=analysisData_pred.iloc[:,1:]

train_x, test_x, train_y, test_y = train_test_split(X, y, test_size = 0.2, random_s
print(train_x.shape, test_x.shape, train_y.shape, test_y.shape) # 데이터 개수 확인

(105, 37) (27, 37) (105, 1) (27, 1)

기본적인 randomforest모형

[ ] from sklearn.ensemble import RandomForestClassifier
    from sklearn.metrics import accuracy_score # 정확도 함수

clf = RandomForestClassifier(n_estimators=20, max_depth=5, random_state=0)
clf.fit(train_x, train_y)

predict1 = clf.predict(test_x)
print(accuracy_score(test_y, predict1))
```

[그림 4-28] 1차 랜덤포레스트 학습

랜덤포레스트는 트리(Tree)의 깊이를 5로 제한하고 20개의 표본추출(Sampling)을 통해 학습을 진행한다.

```
# sample 100개, tree depth - 20

clf = RandomForestClassifier(n_estimators=100, max_depth=20, random_sta
clf.fit(train_x, train_y)

predict2 = clf.predict(test_x)
print(accuracy_score(test_y, predict2))

0.8148148148148148
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:4: DataCor
after removing the cwd from sys.path.

# sample 300개, tree depth - 20

clf = RandomForestClassifier(n_estimators=300, max_depth=20, random_sta
clf.fit(train_x, train_y)

predict2 = clf.predict(test_x)
print(accuracy_score(test_y, predict2))

/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:4: DataCor
after removing the cwd from sys.path.
0.7777777777777778
```

[그림 4-29] 표본 개수 증가 후 랜덤포레스트 학습

표본 개수를 300으로 증가시킨 후 랜덤포레스트 학습 정확도가 0.77로 감소한 것을 보아, 표본 개수의 증가가 성능 향상을 보장하지는 않는다.

```
# sample 100개, tree depth - 100(max)

clf = RandomForestClassifier(n_estimators=100, max_depth=100, random_state=0)
clf.fit(train_x, train_y)

predict2 = clf.predict(test_x)
print(accuracy_score(test_y, predict2))

0.8148148148148148
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:4: DataConversionWarning: A column-vector
  after removing the cwd from sys.path.

# 중요
ok_data=[12.5268, -21.3735, 14.5868, -21.3857, 14.599, 109.9931, 0.0, 0.0415, 12.5222, -35.4199, 0.0045, 14.0465]
X_new = np.array([ok_data])
print("X_new.shape: {}".format(X_new.shape))
prediction = clf.predict(X_new)
print("예측: {}".format(prediction))
print("정답률 =", ac_score)
```

[그림 4-30] 트리 깊이 증가 후 산출된 랜덤포레스트 1차 예측 정확도  
트리의 깊이를 증가시킨 후 예측정확도가 0.81임을 알 수 있다.

## (나) 2차 예측 정확도

시범생산 대상 중 9월 1일~11월 5일 동안의 440개 데이터로 로지스틱 회귀 분석, SVM, RF를 적용해 2차 예측 정확도를 산정하였다.

### ① 로지스틱 회귀분석 예측 정확도

```
[ ] from sklearn.model_selection import train_test_split

Xtrain,Xtest,ytrain,ytest=train_test_split(X,y,test_size=0.2,random_state=2)
from sklearn.linear_model import LogisticRegression
log_reg=LogisticRegression()
log_reg.fit(Xtrain,ytrain)

LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
intercept_scaling=1, l1_ratio=None, max_iter=100,
multi_class='auto', n_jobs=None, penalty='l2',
random_state=None, solver='lbfgs', tol=0.0001, verbose=0,
warm_start=False)

[ ] predictions=log_reg.predict(Xtest)

from sklearn.metrics import confusion_matrix
cm=confusion_matrix(ytest,predictions)
print (cm)

[[22  0]
 [ 5  0]]

혼동행렬을 통한 정확도 계산 방법 22+0/22+0+5+0

▶ from sklearn.metrics import accuracy_score
accuracy_score(ytest,predictions)

0.8148148148148148
```

[그림 4-31] 로지스틱 회귀분석의 2차 예측 정확도

로지스틱의 혼동행렬을 통한 계산을 통해 정확도 0.81이 산출되었다.



### ③ 랜덤포레스트 예측 정확도

```
X=lectureData4.iloc[:,0:-1]
y=lectureData4.iloc[:,1:]

train_x, test_x, train_y, test_y = train_test_split(X, y, test_size = 0.2, random_state = 42)
print(train_x.shape, test_x.shape, train_y.shape, test_y.shape) # 데이터 개수 확인

(105, 53) (27, 53) (105, 1) (27, 1)
```

## 기본적인 randomforest모형

Random Forest는 트리의 깊이를 5로 제한하고 20개의 Sampling을 통해 학습을 진행한다. 그 결과

```
[ ] from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score # 정확도 함수

clf = RandomForestClassifier(n_estimators=20, max_depth=5, random_state=0)
clf.fit(train_x, train_y)

predict1 = clf.predict(test_x)
print(accuracy_score(test_y, predict1))

0.7407407407407407
```

[그림 4-33] 2차 랜덤포레스트 학습

랜덤포레스트는 트리의 깊이를 5로 제한하고 20개의 표본 추출을 통해 학습을 진행한다.

```
[ ] # sample 100개, tree depth - 20

clf = RandomForestClassifier(n_estimators=100, max_depth=20, random_state=0)
clf.fit(train_x, train_y)

predict2 = clf.predict(test_x)
print(accuracy_score(test_y, predict2))

0.8888888888888888
```

sample개수가 증가한다고 하여서 무작정 성능이 향상되지는 않는다.

```
[ ] # sample 300개, tree depth - 20

clf = RandomForestClassifier(n_estimators=300, max_depth=20, random_state=0)
clf.fit(train_x, train_y)

predict2 = clf.predict(test_x)
print(accuracy_score(test_y, predict2))

0.8518518518518519
```

[그림 4-34] 표본 개수 증가 후 랜덤포레스트 학습

표본 개수를 300으로 증가시킨 후의 랜덤포레스트 학습 정확도가 0.77로 감소한 것으로 보아, 표본 개수의 증가가 성능 향상을 보장하지 않는다.

```
[ ] # sample 100개, tree depth - 100(max)

clf = RandomForestClassifier(n_estimators=100, max_depth=100, random_state=0)
clf.fit(train_x, train_y)

predict2 = clf.predict(test_x)
print(accuracy_score(test_y, predict2))

0.8888888888888888

[ ] # 중요
ok_data=[50.0163, 0.003, 0.0015, 0.003, 67.4818, 210.6603, 87.0736, 274.9975, 12.5268]
X_new = np.array([ok_data])
print("X_new.shape: {}".format(X_new.shape))
prediction = clf.predict(X_new)
print("예측: {}".format(prediction))
print("정답률 =", ac_score)
```

[그림 4-35] 트리 깊이가 증가 후 산출된 랜덤포레스트 2차 예측 정확도  
 트리 깊이의 증가 후 랜덤포레스트의 예측정확도가 0.88로 산출되었다.

### (3) XGBoost 피처의 중요도 검증 및 시각화

```
[ ] !pip install xgboost

Requirement already satisfied: xgboost in /usr/local/lib/python3.7/dist-packages (0.90)
Requirement already satisfied: scipy in /usr/local/lib/python3.7/dist-packages (from xgboost) (1.4.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages (from xgboost) (1.19.5)

[ ] import xgboost as xgb ## XGBoost 불러오기
from xgboost import plot_importance ## Feature Importance를 불러오기 위함
import pandas as pd
import numpy as np
from sklearn.datasets import load_breast_cancer
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix, accuracy_score, precision_score, recall_score
from sklearn.metrics import confusion_matrix, f1_score, roc_auc_score
import warnings
warnings.filterwarnings('ignore')
```

[그림 4-36] XGBoost 작업을 위한 관련 모듈 설치 작업 내용

XGBoost를 통한 각 피처의 중요도 시각화를 위해서 먼저, 관련 모듈을 설치하고 훈련을 위해 특징변수와 라벨 변수 외 교차 검증을 위해 훈련 데이터셋을 생성한다.

```

X_features = analysisData_pred.iloc[:,0:-1]
y_label = analysisData_pred.iloc[:, -1]

# 전체 데이터셋을 학습용 80%, 테스트용 20%로 분할
X_train, X_test, y_train, y_test = train_test_split(X_features, y_label, test_size=0.2, random_state=156)
print(X_train.shape, X_test.shape)

(105, 37) (27, 37)

# 넘파이 형태의 학습 데이터 세트와 테스트 데이터를 DMatrix로 변환하는 예제
dtrain = xgb.DMatrix(data=X_train, label = y_train)
dtest = xgb.DMatrix(data=X_test, label=y_test)

```

[그림 4-37] XGBoost 작업을 위한 변수 생성 작업 및 교차 검증 처리

파이썬래퍼 XGBoost와 사이킷런래퍼 XGBoost의 가장 큰 차이는, 파이썬래퍼는 학습용 데이터셋과 테스트 데이터셋을 위해 별도의 DMatrix를 생성한다는 점이다.

```

# max_depth = 3, 학습률은 0.1, 예제가 이진분류이므로 목적함수(objective)는 binary:logistic(이진 로지스틱)
# 오류함수의 평가성능지표는 logloss
# 부스팅 반복횟수는 400
# 조기중단을 위한 최소 반복횟수는 100

params = { 'max_depth' : 3,
           'eta' : 0.1,
           'objective' : 'binary:logistic',
           'eval_metric' : 'logloss',
           'early_stoppings' : 100 }

num_rounds = 400

```

[그림 4-38] 학습 변수들에 대한 초기값 설정

학습 변수들에 대한 초기값을 설정한다. 파이썬래퍼 XGBoost에서 하이퍼파라미터를 XGBoost 모듈의 train() 함수에 파라미터로 전달한다. 사이킷런래퍼는 Estimator 생성자를 하이퍼파라미터로 전달한다. 조기 중단을 위한 early\_stopping\_rounds 파라미터를 설정하고 eval\_set과 eval\_metric이 함께 설정한다.

```
[ ] # train 데이터 세트는 'train', evaluation(test) 데이터 세트는 'eval' 로 명기
    wlist = [(dtrain, 'train'), (dtest, 'eval')]
    # 하이퍼 파라미터와 early stopping 파라미터를 train() 함수의 파라미터로 전달
    xgb_model = xgb.train(params = params, dtrain=dtrain, num_boost_round=num_rounds, evals=wlist)

[0]  train-logloss:0.624715  eval-logloss:0.643059
[1]  train-logloss:0.567584  eval-logloss:0.601902
[2]  train-logloss:0.516743  eval-logloss:0.566954
[3]  train-logloss:0.47498   eval-logloss:0.537929
[4]  train-logloss:0.436784  eval-logloss:0.52173
[5]  train-logloss:0.402406  eval-logloss:0.503046
[6]  train-logloss:0.372669  eval-logloss:0.492631
[7]  train-logloss:0.348571  eval-logloss:0.471721
[8]  train-logloss:0.323933  eval-logloss:0.463576
[9]  train-logloss:0.302418  eval-logloss:0.45853
[10] train-logloss:0.285296  eval-logloss:0.443484
[11] train-logloss:0.266954  eval-logloss:0.435493
```

[그림 4-39] XGBoost 훈련 내용

train( )을 통해 반복 학습을 수행하면 train-logloss와 eval-logloss가 지속적으로 감소한다.

```
pred_probs = xgb_model.predict(dtest)
print('predict() 수행 결과값을 10개만 표시, 예측 확률 값으로 표시됨')
print(np.round(pred_probs[:10], 3))

# 예측 확률이 0.5보다 크면 1, 그렇지 않으면 0으로 예측값 결정해 리스트 객체인 preds에 저장
preds = [ 1 if x > 0.5 else 0 for x in pred_probs]
print('예측값 10개만 표시: ', preds[:10])

predict() 수행 결과값을 10개만 표시, 예측 확률 값으로 표시됨
[0.005 0.335 0.002 0.003 0.002 0.014 0.055 0.131 0.002 0.001]
예측값 10개만 표시:  [0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

# 혼동행렬, 정확도, 정밀도, 재현율, F1, AUC 불러오기
def get_clf_eval(y_test, y_pred):
    confusion = confusion_matrix(y_test, y_pred)
    accuracy = accuracy_score(y_test, y_pred)
    precision = precision_score(y_test, y_pred)
    recall = recall_score(y_test, y_pred)
    F1 = f1_score(y_test, y_pred)
    AUC = roc_auc_score(y_test, y_pred)
    print('오차행렬:\n', confusion)
    print('\n정확도: {:.4f}'.format(accuracy))
    print('정밀도: {:.4f}'.format(precision))
    print('재현율: {:.4f}'.format(recall))
    print('F1: {:.4f}'.format(F1))
    print('AUC: {:.4f}'.format(AUC))

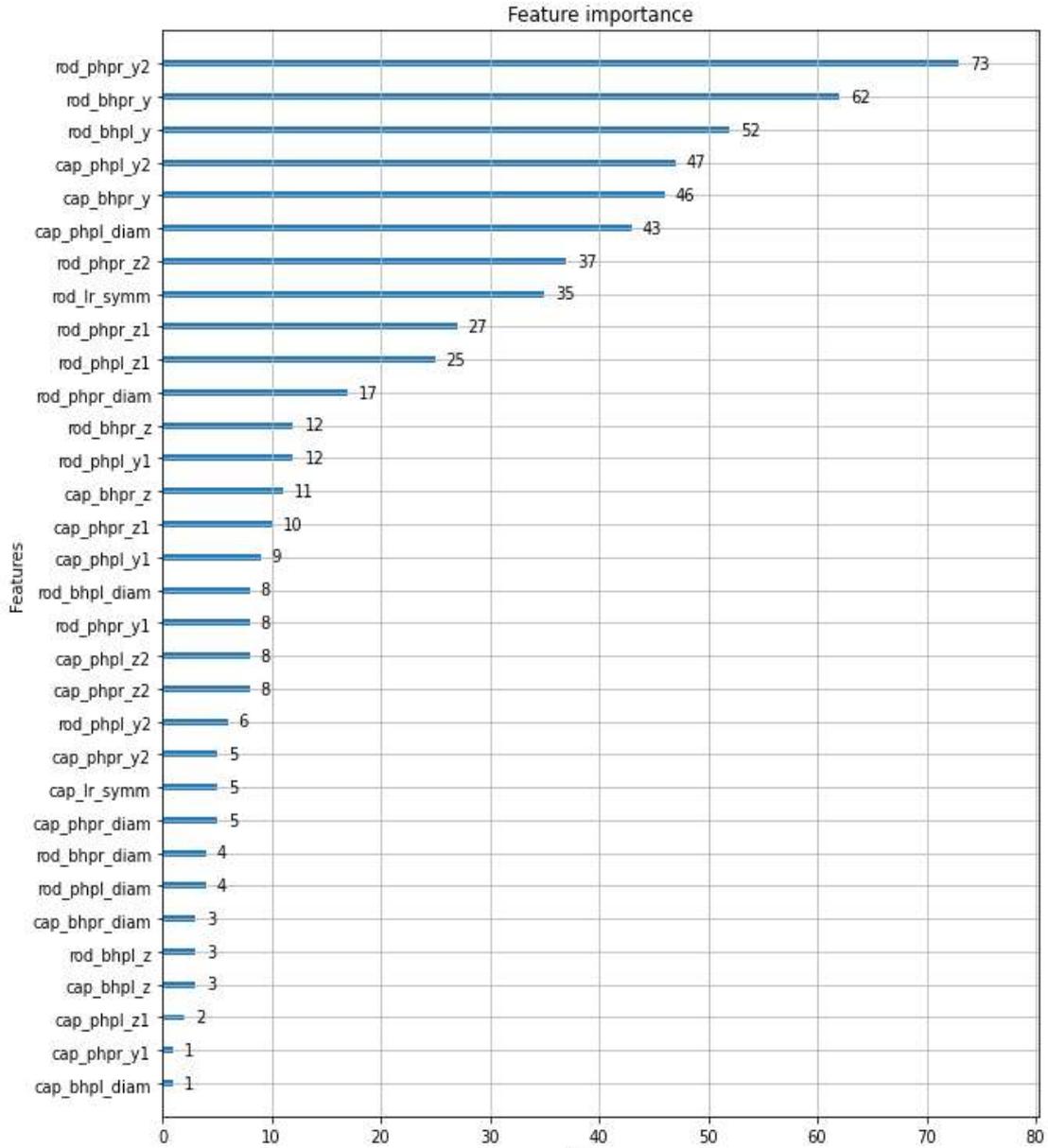
get_clf_eval(y_test, preds)

오차행렬:
[[19  1]
 [ 4  3]]

정확도: 0.8148
정밀도: 0.7500
재현율: 0.4286
F1: 0.5455
AUC: 0.6893
```

[그림 4-40] XGBoost의 정확도

XGBoost의 정확도는 0.814로 산출되었다.



[그림 4-41] XGBoost를 통한 각 피처의 중요도

그림 4-41을 보면, 합격/불합격 결정 요인으로 rod\_phpr\_y2(핀홀 위치 우/Y)와 rod\_bhpr\_y(볼트홀 위치 우)의 F1스코어가 73과 62로 가장 큰 중요한 역할을 하는 것을 확인할 수 있으며, 이 부분에 대한 관심과 관리가 향후 합격/불합격 관리에 결정적 요인이 됨을 알 수 있다.

라) AI 모델링 분석 결과

(1) 기술적 지표들과 일간 공정 변동성 상관분석을 실시한 결과 대부분의 기술적 지표와 공정 변동성 간에는 통계적으로 유의미한 상관관계가 있는 것으로 나타났다. 자료의 가용성과 연속성을 고려하여 커넥팅로드에 한정하여 일 단위로 방향성 예측을 했다

(2) 모델 성능 비교를 위해 로지스틱 회귀, 서포트 벡터 머신과 랜덤 포레스트를 1, 2차에 걸쳐 1차는 9월 1일~14일까지 132개, 2차는 9월 1일~11월 5일까지 440개 데이터로 알고리즘 예측정확도를 비교하였다.

(3) 1, 2차의 로지스틱 회귀분석, 서포트 벡터 머신은 거의 비슷한 예측 정확도를 보여 주고 있고, 랜덤포레스트 예측 정확도는 1, 2차 차수 간을 비교해 볼 때 2차가 정확도가 높고 샘플수 100, 트리-100 기준으로 정확도가 높음을 볼 수 있다. 세 모델을 종합적으로 살펴보면 랜덤 포레스트 배깅 모델의 성과가 가장 좋게 나타났다.

(4) F1 스코어의 상위 순위로 살펴보면 합격/불합격 결정 요인으로 rod\_phpr\_y2(우측 로드 핀홀 y2), rod\_bhpr\_y(우로드 볼트홀 y)의 두 변수가 중요한 역할을 하고 있는 것을 확인할 수 있다. 이 부분에 대한 불량 요인의 공정 개선 및 관리가 향후 합격/불합격 관리에 결정적 요인이 됨을 알 수 있다. F1 스코어가 가장 높은 rod\_phpr\_y2 등 주요 변수 등에 대해 SPC통계 분석과 비교 검증할 필요가 있다.

(5) 건전성 관리의 일환인 공구상태 진단 시스템의 개발을 위해, 공구의 마모도를 반영하는 특징적인 데이터 요소로 확인된 스피들 부하의 제공평균 제공근 값을 이용하여, 서포트 벡터 머신 및 k-평균 군집 등의 기계학습을 기반으로 한 알고리즘을 활용해 진단 알고리즘을 개발하였다. 또한 k-fold 교차 검증법을 통해 해당 알고리즘의 유효성을 검증 및 확인하였다.

(6) 공구 교환 시기의 적절한 단기방향성 예측에 대한 기술적 지표를 활용한 인공지능 모델의 유효성을 검증하였다.

### 3) SPC 통계 솔루션과 연계한 통계분석

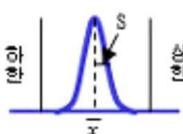
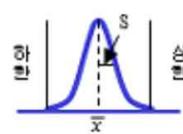
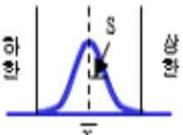
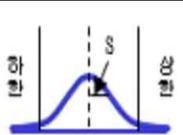
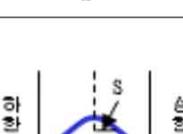
#### 가) 공정능력 및 관리도 해석 기준 주요내용

##### (1) 공정능력 및 관리도의 개요

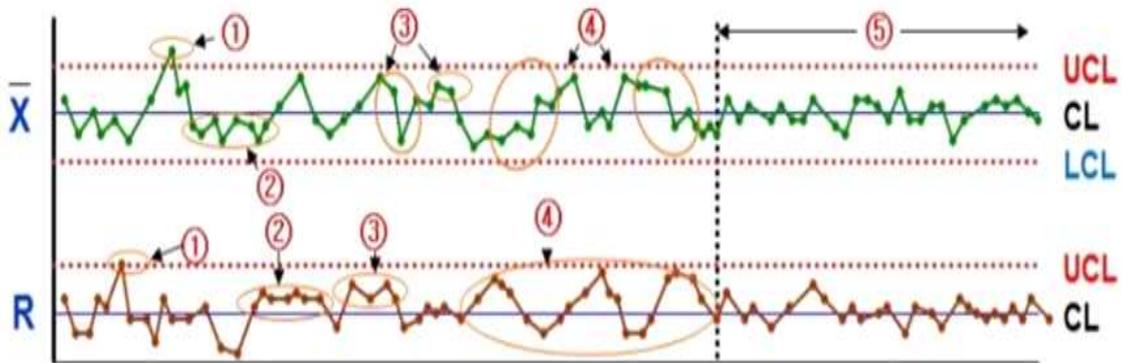
공정능력 조사는 공정변동의 원인을 연구하고 이상원인과 우연원인으로 분리하여 이상원인을 제거하기 위한 활동이고, 관리도는 공정의 품질변동을 측정하는 주요수단이다. 대부분 공정능력지수와 관리도는 함께 사용된다. CPk는 규격의 중심으로부터 실제 공정분포의 평균치의 치우침이 있을 경우의 공정능력 측정방법이고,  $\bar{X}$ -R관리도는 치수, 압력과 같이 계량치의 산포가 관리상한선 및 하한선 내에서 일상 공정관리 하는 도구이다. 1단계에서는 관리도를 이용하여 공정이 통계적 관리상태에 있는지를 판정한다. 2단계에서는 개략적인 공정능력(Cpk)을 산정한다. 공정능력 분석은 ① 공정상태를 파악하고 ② 주어진 규격을 인식하여 ③ 이들 공정상태와 규격을 비교하는 것이다. 관리도는 관리상한선(UCL)과 관리하한선(LCL) 내에 측정결과치가 있으면 관리상태로 간주하고, 이 점들이 관리한계선에 벗어나면 우연원인이 아닌 이상원인의 발생으로 본다. 공정능력 및 관리도에 대한 주요 해석내용은 아래의 표 4-8, 4-9, 그림 4-46과 같다.

(2) 공정능력 결과 해석 주요 기준 내용 79)

[표 4-8] 공정능력 결과 해석 주요 기준 내용

분포현상	공정능력 지수	등급	공정능력 판단	시정조치	비고	
					Cp	$\sigma$
	$Cp \geq 1.67$	0	매우 충분	- 들쭉날쭉이 약간 커져도 걱정할 필요가 없다. - 비용절감이나 관리의 간소화를 생각하도록 한다.	1.67	$\pm 5\sigma$
	$1.67 > Cp \geq 1.33$	1	충분	- 매우 이상적인 공정상태으로, 현 상태를 유지한다.	1.33	$\pm 4\sigma$
	$1.33 > Cp \geq 1.00$	2	양호함	- 공정관리를 확실히 하여 관리 상태를 유지한다. - Cp가 1에 가까워지면 불량 발생의 가능성이 있으므로 주의한다.	1.00	$\pm 3\sigma$
	$1.00 > Cp \geq 0.67$	3	부족	- 불량품이 생기는 상황으로, 전체 선별, 공정의 개선 및 관리가 필요하다.	0.67	$\pm 2\sigma$
	$0.67 > Cp$	4	매우 부족	- 불만족스러운 품질로서, 현황조사, 원인규명, 품질개선 등의 긴급대책이 필요하다. - 상한, 하한 규격값을 재검토한다.	0.33	$\pm 1\sigma$

(3) 관리도 해석 주요 내용



[그림 4-42] 관리도 주요 경향

79) <https://blog.daum.net/pldworld/2>

[표 4-9] 관리도 해석 주요 내용

번호	이름	점의 배열 관계	조치 사항
①	관리이탈	X관리도 : 평균치의 변화, R관리도 ; 산포의 변화	이상 원인을 조사하여 다시 일어나지 않도록 조치
②	크기 7 이상의 런(RUN)	중심선의 한 쪽으로만 연속하 여 7개 이상의 점이 있을 때	원인 파악, 유익한 정 보 입수
③	한계 가까이 있는 점	중심선으로부터 관리 한계선까 지 거의 2/3 이상 떨어진 곳 에 연속 3점 중 2점	공정의 산포가 증가하 였음을 보여줌
④	경향과 주기성 변동	점이 계속 위로 또는 아래로만 향하고 있을 때 또는 주기적 변동을 보일 때	원인 파악, 유익한 정 보 입수
⑤	안정상태(관리상태)	연속 25점 이상에서 ①~④까 지와 같은 현상이 없는 상태	공정 안정 및 규격 만 족.

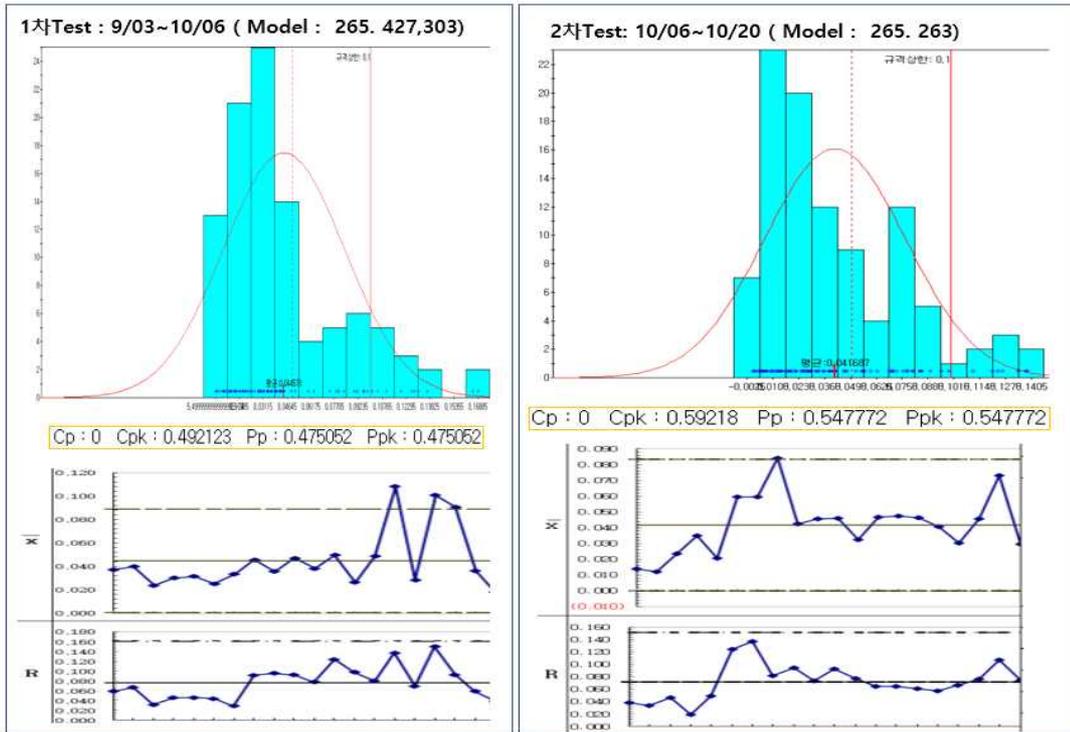
나)  $\bar{X}$ -R, P관리도 및 공정능력 분석 구현결과

(1) 시료 및 변수 선정

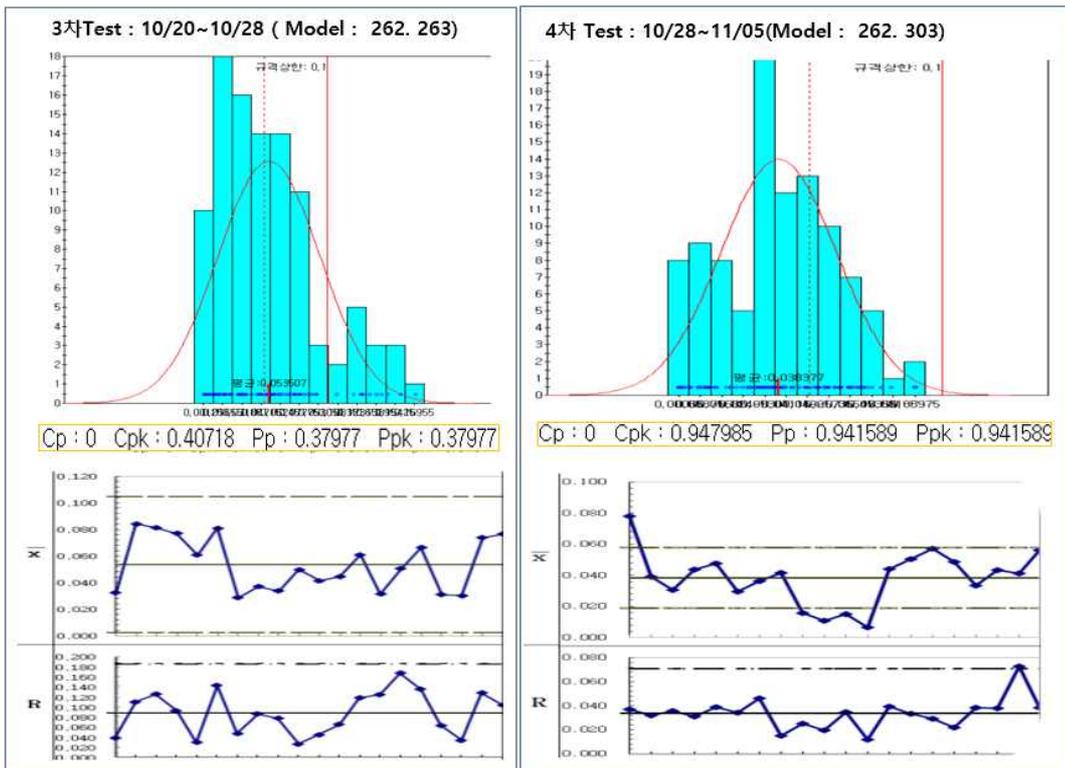
커넥팅 로드 제품 모델인 262, 263, 265, 303, 427을 시범생산(9.1~11.5)기  
간 중 총 시료수 440개 중 400개 데이터를 선정하였다. 기간별로 순차적으로  
구분하여 1~4차에 걸쳐 100개씩 층별하여  $\bar{X}$ -R, P관리도 및 공정능력통계 분  
석을 하였다.

통계 분석 대상의 주요 변수로는 2가지를 선정하였다. 주요 품질관리 변수  
요인인 커넥팅로드 볼트홀 대칭도와, 머신러닝의 XGBoost 알고리즘에서 피쳐  
(Feature) 중요도가 가장 높은 커넥팅로드 좌측 핀홀 y2 거리를 선정하였다.

(2) 커넥팅로드 볼트홀 대칭도 공정능력 및  $\bar{X}$ -R 관리도, P 관리도

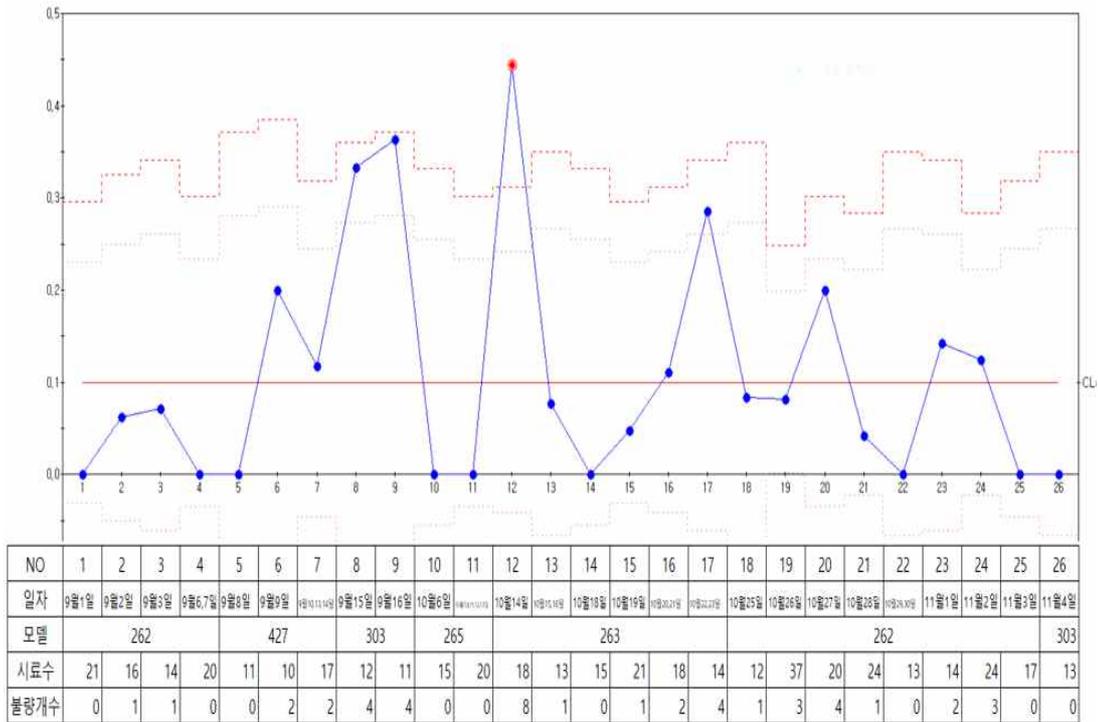


[그림 4-43] 볼트홀 1차, 2차 시범생산  $\bar{X}$ -R관리도 및 공정능력



[그림 4-44] 볼트홀 3차, 4차 시범생산  $\bar{X}$ -R관리도 및 공정능력

(3) 커넥팅 로드 볼트홀 대칭도 P관리도



[그림 4-45] 커넥팅 로드 볼트홀 P 관리도

(4) 볼트홀 대칭도의 공정능력 및  $\bar{X}$ -R, P관리도 상세 분석

[표 4-10] 볼트홀 대칭도 Cpk 및 관리도 종합 분석표

구분	공정능력			$\bar{X}$ -R관리도								계 (이상 해석 개소)
	Cpk	평균 값	표준 편차	관리 이탈		런 현상		한계 근접		주기, 경향		
				$\bar{X}$	R	$\bar{X}$	R	$\bar{X}$	R	$\bar{X}$	R	
1차	0.47	0.044	0.038	3		2	2		1	1		9
2차	0.57	0.041	0.035	1			1		1			3
3차	0.40	0.053	0.048			1	1					2
4차	0.94	0.038	0.021	1	1				1			3
P관리도				1				1				2

### (가) 공정능력 분석

커넥팅로드 볼트홀 대칭도의 공정능력지수(CPk)를 살펴보면 해석기준은 표 4-10과 같이 Cpk 0.67↓는 공정능력이 매우 부족, Cpk 0.67~1.0 부족한 편, 1.0~1.33 보통, 1.33~1.67 충분, 1.67↑ 우수로 분류된다. 표 4-10과 같이 시범생산 결과 1차 0.47, 2차 0.59, 3차 0.40, 4차 0.94로 결과가 도출되었다. 1~3차의 평균값, 표준편차가 0.041~ 0.041, 0.035~0.048이었으나 4차는 0.038, 0.021로 대폭 감소한 것을 볼 수 있다. 그림 4-43, 4-44와 같이 1~3차 까지 좌측으로 평균값이 몰려 있으나 조정 및 공정개선을 통해 4차에는 중앙으로 평균값이 이동하고 표준편차가 줄어든 것으로 볼 수 있다. 평균이 공정안정을 위해 최소  $CPk > 1.0$  이상으로 공정개선이 이루어져야 할 필요가 있다.

### (나) 관리도 분석

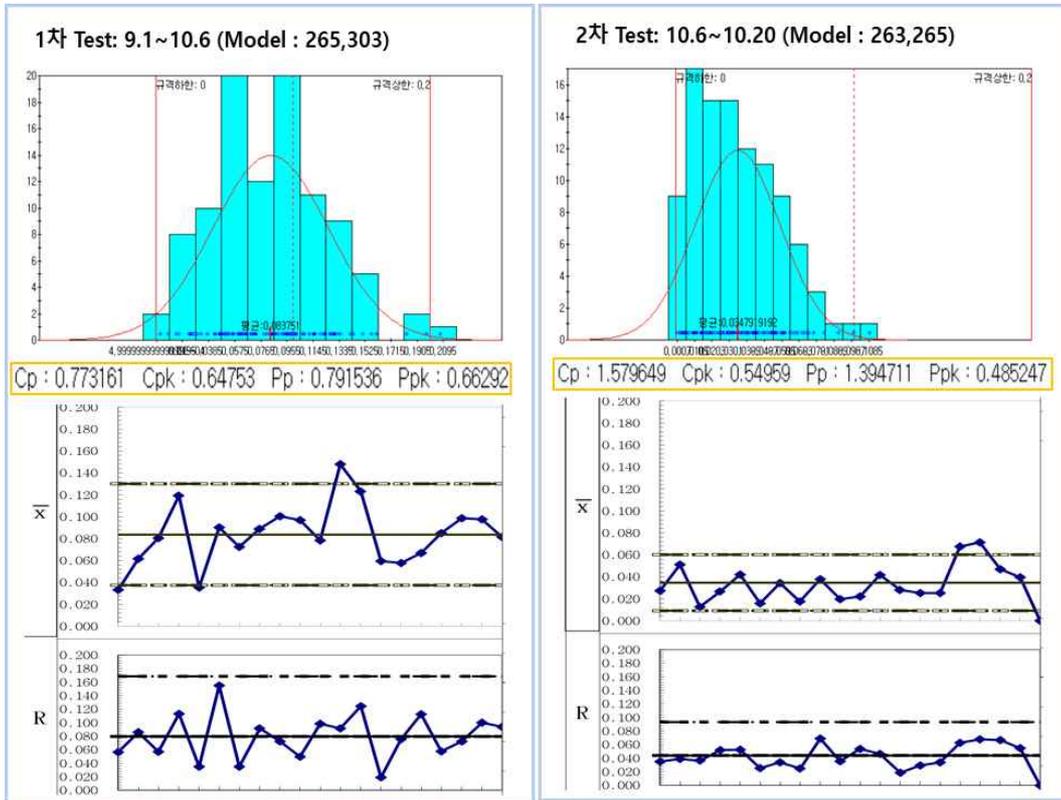
볼트홀 대칭도의 관리도 수준 조사를 살펴보면, 표 4-9와 같이 관리도의 이상요인에 대한 주요 해석은, ① 관리이탈은 상한 관리선을 이탈한 경우, ② 런 현상은 7개점 이상 점이 한쪽에 연속 발생한 경우, ③ 한계근접은  $2\sigma \sim 3\sigma$  내 3점중 2개 이상이 위치한 경우, ④ 경향은 규칙적인 상승 또는 하강이 발생하는 경우 등의 4가지이다.

$\bar{X}$ -R관리도는 표 4-10 종합분석표와 같이 4가지 이상요인의 주요 해석관점으로 보면 1차 9개소, 2차 3개소, 3차 2개소, 4차 3개소이다. 2, 3차와 비교해 볼 때 4차가 이상요인이 많다고 할 수 있으나 그림 4-43, 4-44와 같이 2, 3차는  $\bar{X}$ -R관리도의 관리상한선이 범위가 넓어서 Cp와 함께 비교해 볼 때 오히려 공정이 불안정하다고 볼 수 있다. 그러나 1차, 3차, 4차 테스트에서 관리상한선 이상으로 벗어나고 있고, R관리도를 포함 관리한계, 주기적 경향과 변동이 관찰되므로 원인파악 및 공정개선이 필요하다.

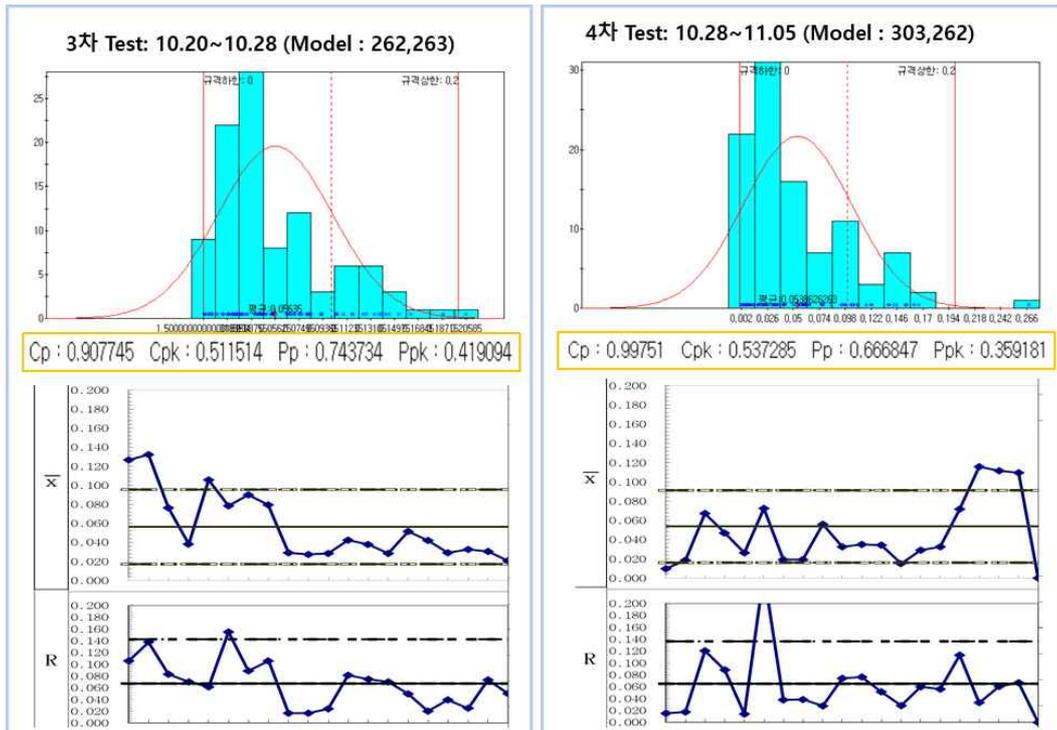
그림 4-45와 같이 P관리도는 일자별 모델을 고려하여 불일정한 시료 수에 적용하였다. P(불량율)관리도에서 1점이 관리상한선에 벗어나고 있고 1개소가 관리한계에 위치하므로 원인제거 후 공정관리가 필요하다.

공정능력,  $\bar{X}$ -R관리도, 관리도에 대한 개별 상세 분석 내역은 부록 1~3을 참조한다.

(1) 커넥팅로드 우측 핀홀 y2 공정능력 및  $\bar{X}$ -R관리도, P관리도

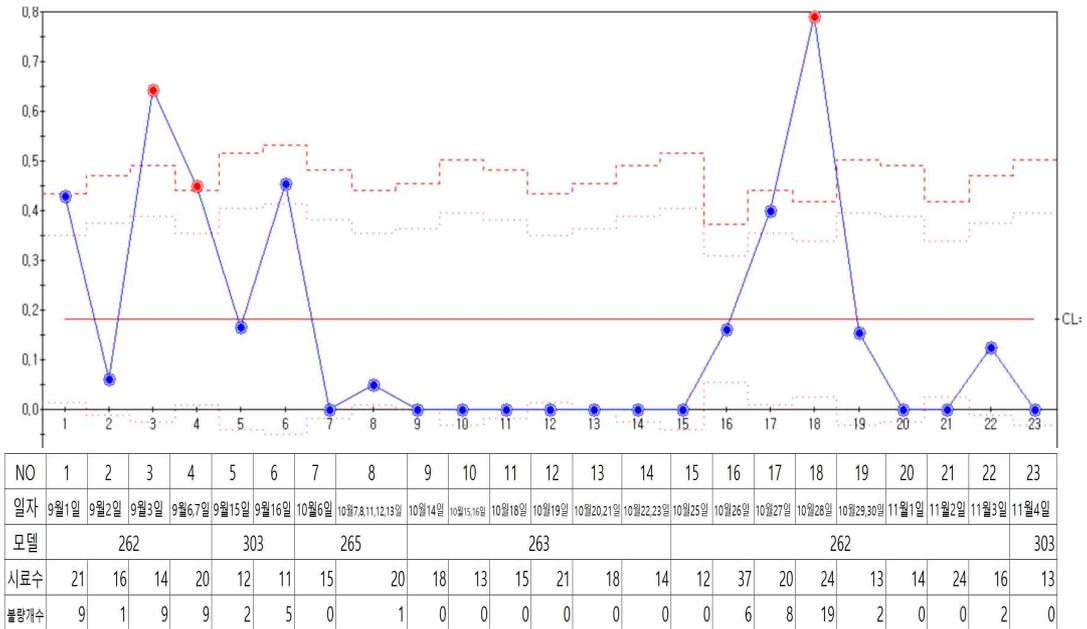


[그림 4-46] 우측 핀홀 1차, 2차 시범생산  $\bar{X}$ -R관리도 및 공정능력



[그림 4-47] 우측 핀홀 3차, 4차 시범생산  $\bar{X}$ -R관리도 및 공정능력

(2) 커넥팅로드 우측 핀홀 P관리도



[그림 4-48] 커넥팅로드 우측 핀홀 P 관리도

(3) 커넥팅로드 우측 핀홀 y2 공정능력 및  $\bar{X}$ -R, P관리도 상세 분석

[표 4-11] 우측핀홀 y2 Cpk 및 관리도 종합 분석표

구분	공정능력			$\bar{X}$ -R관리도								계 (이상 해석 개소)
	Cpk	CP	평균값	관리 이탈		런 현상		한계 근접		주기, 경향		
				$\bar{X}$	R	$\bar{X}$	R	$\bar{X}$	R	$\bar{X}$	R	
1차	0.64	0.77	0.083	3							1	4
2차	0.54	1.57	0.034	3	1					1	1	6
3차	0.51	0.90	0.056	3	1	1		3	1			9
4차	0.53	0.99	0.053	5	2			1	1			9
P관리도				3								3

#### (가) 공정능력 분석

커넥팅로드 우측편홀 y2의 공정능력지수(CPk)를 살펴보면 표 4-11과 같이 시범생산결과, 1차 0.64, 2차 0.54, 3차 0.51, 4차 0.53으로 결과가 도출되었다. 1~4차 전체 실제공정능력지수(Cpk)가 표 4-8과 같이 0.67 이하로 공정능력이 매우 부족함을 볼 수 있다. 그림 4-46, 4-47과 같이 1~4차의 시계열순으로 공정능력 정상화를 볼 수가 없다. 좌측으로 평균값이 몰려 있거나 공정변동 요인에 따른 공정개선이 미흡한 것으로 보인다. 표 4-11과 같이 잠재공정능력지수(Cp)보다 1~4차 실제공정능력지수(Cpk)가 현저하게 낮은 것은 공정평균과 많이 벗어나 있는 것으로 해석된다. 평균도 일정하지 않음을 볼 수 있다. 공정조정 및 공정개선을 통해 평균값이 중앙으로 이동하고 공정산포를 줄이는 활동이 필요하다. 공정안정을 위해 평균이 최소  $CPk > 1.0$ 으로 공정개선이 이루어져야 한다.

#### (나) 관리도 분석

우측 편홀 y2의 관리도의 이상요인을 보면, 표 4-11과 같이 1차 4개소, 2차 6개소, 3차 9개소, 4차 9개소이다. 특히 각 차수의 관리선 이탈이 3건~7건으로 전체적으로 이상요인이 많고, 공정관리가 부족하다고 볼 수 있다. 그림 4-46, 4-47과 같이 무엇보다도  $\bar{X}$ -R관리도의 관리선 이탈의 요인이 공정평균의 일치되도록 하고 공정산포를 줄이기 위해서 원인파악 및 체계적 공정개선 활동이 필요하다.

P관리도는 일자별 모델을 고려하여 불일정한 시료 수에 적용하였다. P(불량율)관리도에서 3점이 관리상한선에 벗어나고 있어 앞에서 언급한 내용과 연계하여 원인제거 후 공정관리가 필요하다.

공정능력,  $\bar{X}$ -R관리도, 관리도에 대한 개별 상세 분석 내역은 부록 4~6을 참조한다.

라) SPC통계분석 종합 결과

커넥팅로드 볼트홀 대칭도, 우측핀홀의 두 변수에 대한 1~4차  $\bar{X}$ -R관리도, P관리도 및 공정능력 등의 SPC 통계분석 종합 결과는 표 4-12와 같다.

상세 통계분석 결과는 4절 1) SPC 통계구현에 대한 분석 검증에서 기술하였다.

[표 4-12] SPC 통계분석 종합 결과

구 분	Cpk값	$\bar{X}$ -R관리도 이상점				계	
		관리 이탈	런현상	한계 근접	주기적 변동		
볼트홀 대칭도	1차	0.47	3	4	1	1	9
	2차	0.57	1	1	1		3
	3차	0.40		2			2
	4차	0.94	2		1		3
	p관리도 이상점		1		1		2
	계		7	7	4	1	19
우측 핀홀 y거리	1차	0.64	3			1	4
	2차	0.54	4			2	6
	3차	0.51	4	1	4		9
	4차	0.53	7		2		9
	p관리도 이상점		3				3
	계		21	1	6	3	31

## 제 4 절 제조데이터 수집 분석 및 활용시스템 구현의 성능 검증

### 1) SPC 통계 구현에 대한 분석 검증

#### 가) SPC 통계 개요 및 해석 기준

제3장 3절 4) 가)의 CPk,  $\bar{X}$ -R관리도, P관리도 수식을 적용한다.

[표 4-13] 공정능력 결과 해석 주요 기준 내용

등급	공정능력 지수	공정능력 판단	비고	
			Cp	$\sigma$
0	$Cp \geq 1.67$	매우충분	1.67	$\pm 5\sigma$
1	$1.67 > Cp \geq 1.33$	충분	1.33	$\pm 4\sigma$
2	$1.33 > Cp \geq 1.00$	양호함	1.00	$\pm 3\sigma$
3	$1.00 > Cp \geq 0.67$	부족	0.67	$\pm 2\sigma$
4	$0.67 > Cp$	매우부족	0.33	$\pm 1\sigma$

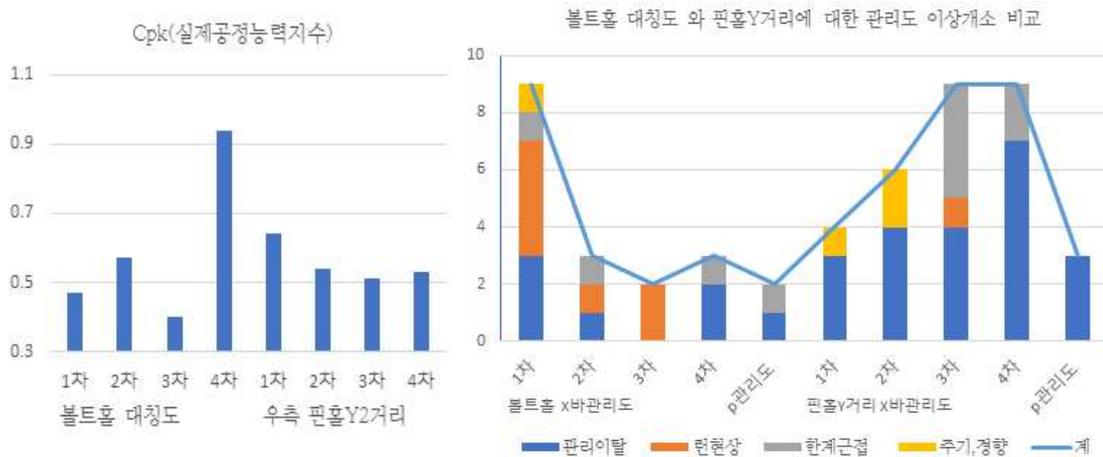
[표 4-14] 관리도 해석 주요 내용

번호	이름	점의 배열 관계
①	관리이탈	X관리도 : 평균치의 변화, R관리도 ; 산포의 변화
②	크기 7 이상의 런(RUN)	중심선의 한 쪽으로만 연속하여 7개 이상의 점이 있을 때
③	한계 가까이 있는 점	중심선으로부터 관리 한계선까지 거의 2/3 이상 떨어진 곳에 연속 3점 중 2점
④	경향과 주기성 변동	점이 계속 위로 또는 아래로만 향하고 있을 때 또는 주기적 변동을 보일 때
⑤	안정상태(관리상태)	연속 25점 이상에서 ①~④까지와 같은 현상이 없는 상태

나) SPC통계 분석 구현 평가

[표 4-15] SPC통계분석 종합 요약표

구 분	Cpk값	$\bar{X}$ -R관리도 이상점				계	
		관리 이탈	런현상	한계 근접	주기적 변동		
볼트 홀 대칭도	1차	0.47	3	4	1	1	9
	2차	0.57	1	1	1		3
	3차	0.40		2			2
	4차	0.94	2		1		3
	p관리도 이상점		1		1		2
	계		7	7	4	1	19
우측 핀홀 y거리	1차	0.64	3			1	4
	2차	0.54	4			2	6
	3차	0.51	4	1	4		9
	4차	0.53	7		2		9
	p관리도 이상점		3				3
	계		21	1	6	3	31



[그림 4-49] SPC통계분석 종합 그래프

커넥팅로드 제품 모델 5종을 시범생산 기간(9.1~11.5) 중 총 시료수 440개 중 400개 데이터를 선정하였다. 기간별로 순차적으로 1~4차에 걸쳐 100개씩 층별하여  $\bar{X}$ -R, P관리도 및 공정능력통계 분석을 하였다.

통계 분석 대상의 변수요인은 첫 번째로 주요 공정관리 변수인 볼트홀 대칭도의 요인을 선정하였고, 두 번째 변수로는 머신러닝의 XGBoost 알고리즘에서

피쳐 중요도가 가장 높은 좌측 핀홀 y2의 요인 2가지를 선정하였다.

두 변수에 대한 1~4차  $\bar{X}$ -R관리도, P관리도 및 공정능력 등의 SPC 통계분석 종합 결과는 표 4-15와 그림 4-49와 같다.

#### (1) 공정능력 분석

커넥팅로드 볼트홀 대칭도의 공정능력지수(CPk)가 그림 4-43, 44와 같이 시범생산 1~3차 0.40~0.59 4차 0.94로 결과가 도출되었다. 1~3차까지는 좌측으로 평균값이 몰려 있으나 조정 및 공정개선을 통해 4차에는 중앙으로 평균값이 이동하고 표준편차가 줄었다. 공정안정을 위해서는 CPk 평균이 최소 1.0 이상이 되도록 공정이 개선할 필요가 있다.

우측핀홀 Y2의 공정능력지수는 그림 4-46, 4-47과 같이 시범생산 결과 1~4차 0.51~0.64로 전체 차수 0.67 이하로 공정능력이 매우 부족함을 볼 수 있다. 시범생산 전체기간에 걸쳐 공정능력이 정상화라고 볼 수가 없고, 좌측으로 평균값이 몰려 있거나 공정변동 요인에 따른 공정개선이 미흡한 것으로 보인다. 잠재공정능력지수(Cp)보다 1~4차 실제 공정능력지수(Cpk)가 현저하게 낮은 것은 공정평균과 많이 벗어나 있는 것으로 해석되고 평균도 일정하지 않음을 볼 수 있다. 공정개선을 통해 중앙으로 평균값이 이동하고 공정산포를 줄이도록 하는 활동이 필요하다.

두 요인의 공정능력을 비교해 보면, 볼트홀 대칭도 요인은 공정개선을 통해 시간 경과에 따라 Cpk가 0.94로 안정화 되고 있으나, 우측 핀홀 y2는 여전히 Cpk 0.67 이하로 공정능력이 부족하고 공정평균의 이동과 공정산포를 줄이는 것이 중요할 것으로 보인다.

#### (2) 관리도 분석

볼트홀 대칭도의 관리도 수준은 공정능력과 연계해서 고찰해 보면  $\bar{X}$ -R관리도의 이상요인이 표 4-15와 같이 4차가 2, 3차에 비해 줄어들지 않았다고 할 수 있으나, 표 4-43, 44와 같이 2, 3차는  $\bar{X}$ -R관리도의 관리상한선이 범위가 넓어서 Cpk와 함께 조사해 볼 때 공정이 불안정하다고 볼 수 있다. P(불량율) 관리도도 관리상한선에 벗어나고 있고 관리한계에 위치하므로 관리한계, 주기

적 경향과 변동이 관찰되므로 원인제거 및 공정관리를 할 필요가 있다.

우측 편향  $y_2$ 의 관리도 수준 조사를 살펴보면 표 4-15와 같이  $\bar{X}$ -R관리도는 이상건수는 31건으로 특히 그 중 관리선이탈 21건은 전체 이상요인의 68%를 차지하여 관리이탈이 높은 비중이므로 공정관리가 매우 부족하다고 볼 수가 있다. 무엇보다도  $\bar{X}$ -R관리도의 관리선 이탈의 요인을 제거해 공정평균이 일치되도록 하고 공정산포를 줄이기 위해서 원인파악 및 체계적 공정개선 활동이 필요하다. P관리도도 관리상한선에서 3점이 벗어나고 있어 공정관리가 함께 필요하다.

두 요인의 이상요인을 비교해 보면, 표 4-15와 같이 볼트홀 대칭도는 19건, 우측편향  $y_2$ 는 31건으로, 볼트홀 대칭도는 우측편향  $y_2$  대비 61% 수준이고 관리이탈 건수는 31% 수준으로 우측편향보다 공정이 안정화되었다고 판단할 수 있다.

## 2) AI 알고리즘 예측 정확도 성능 평가

### 가) 혼동행렬(confusion matrix)의 개요 및 성과평가 기준<sup>80)</sup>

혼동행렬은 모델, 검사도구, 알고리즘의 진단, 분류, 판별, 예측 능력을 평가하기 위하여 고안된 표이며 오류행렬(error matrix)이라고도 한다. 검사도구의 성과평가 용도로 평가 기준으로 삼을 수 있으며 혼동행렬개념이 필요하다. 혼동행렬은 그림 4-50과 같이<sup>81)</sup> 예측이 참인데 실제로도 참인 경우(TP), 예측이 참인데 실제로는 거짓인 경우(FN), 예측은 거짓인데 실제로는 참인 경우(FP), 예측도 거짓이고 실제도 거짓인 경우(TN)의 4가지 상황을 빈도별로 정리한다.

---

80) Fawcett, T. (2006). An introduction to ROC analysis. Pattern recognition letters, 27(8), pp. 861-874.

81) <https://manisha-sirsat.blogspot.com/2019/04/confusion-matrix.html>

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	<b>Sensitivity</b> $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	<b>Specificity</b> $\frac{TN}{(TN + FP)}$
		<b>Precision</b> $\frac{TP}{(TP + FP)}$	<b>Negative Predictive Value</b> $\frac{TN}{(TN + FN)}$	<b>Accuracy</b> $\frac{TP + TN}{(TP + TN + FP + FN)}$

[그림 4-50] 성과평가의 기준

(1) 정확도(Accuracy) =  $\frac{(TP + TN)}{(TP + TN + FP + FN)}$  (오류율 = 1 - Accuracy)

(2) 정밀도(Precision) =  $\frac{TP}{TP + FP}$

(3) 재현율(Recall) {=민감도(Sensitivity)} =  $\frac{TP}{TP + FN}$

(4) 특이도(Specificity) =  $\frac{TN}{TN + FP}$

나) 알고리즘 예측정확도 성능 평가

(1) 1차는 9월 14일간 132개, 2차는 9월 1일~11월 5일까지 440개 데이터로 알고리즘 예측정확도를 표 4-16, 그림 4-51과 같이 비교하였다. 1, 2차로 지스틱 회귀분석과 SVM은 거의 비슷한 예측 정확도를 보여 주고 있다.

(2) 랜덤포레스트 예측 정확도는 1, 2차 차수 간을 비교해 볼 때 2차가 정확도가 높고 1차보다 표본 수 100, 트리-100 기준으로 정확도가 7% 높다. 즉 표본 수가 적을 때 예측정확도가 떨어짐을 알 수 있다. 랜덤포레스트를 통해 앙상블 처리함으로써 합/불 예측이 잘 되고 있음을 확인할 수 있었고 표본 수보다 트리 깊이의 예측정확도가 높다는 것을 볼 수 있다. 향후 표본 수가 충분하고

학습이 반복회수가 증가하면 랜덤포레스트의 정확도는 90% 선으로 예상해 볼 수 있다.

[표 4-16] 알고리즘 예측 정확도 수치 테이블

구분	로지스틱	SVM	RF			XGBoost	정확도 적용 기간 및 샘플 수	
			샘플수	100	300			100
			트리깊이	-20	-20	-100		
1차	0.81	0.82		0.81	0.77	0.81	-	9.1~14일 132개
2차	0.81	0.80		0.88	0.85	0.88	0.81	9.1~11.5일 440개



[그림 4-51] 알고리즘 예측 정확도 수치 그래프

(3) 로지스틱 회귀와 SVM을 이용했을 경우 1, 2차 알고리즘 간을 비교 해 볼 때 합/불 예측이 잘 이루어지지 않음을 볼 수 있다.

(4) XGBoost는 1차는 실험샘플 수가 적어 수행하지 않았고 2차 정확도 값은 랜덤포레스트보다 낮다고 볼 수 있으나 변수 중요도 검증차원에서 다른 검증방법인 F1 스코어로 그림 4-56과 같이 수행하였다.

### 3) XGBoost 피처의 중요도 모델에 대한 검증

#### 가) F1 스코어(F1 Score) 개요<sup>82)</sup>

정밀도와 재현율(민감도)을 활용하는 평가용 지수이다. 만일 TP = FP = FN으로 세 셀의 빈도가 동일하다면 F1 값은 무조건 0.5이다. TP 셀이 변화하지 않는 한에서, F1 값은 FP의 빈도와 FN의 빈도 총합이 같다면 두 거짓 셀의 빈도 격차가 아무리 크더라도 이를 반영하지 않는다. 반면 TP가 변화하지 않을 때 FP와 FN의 빈도 총합이 크면 클수록 F1 값은 감소한다.

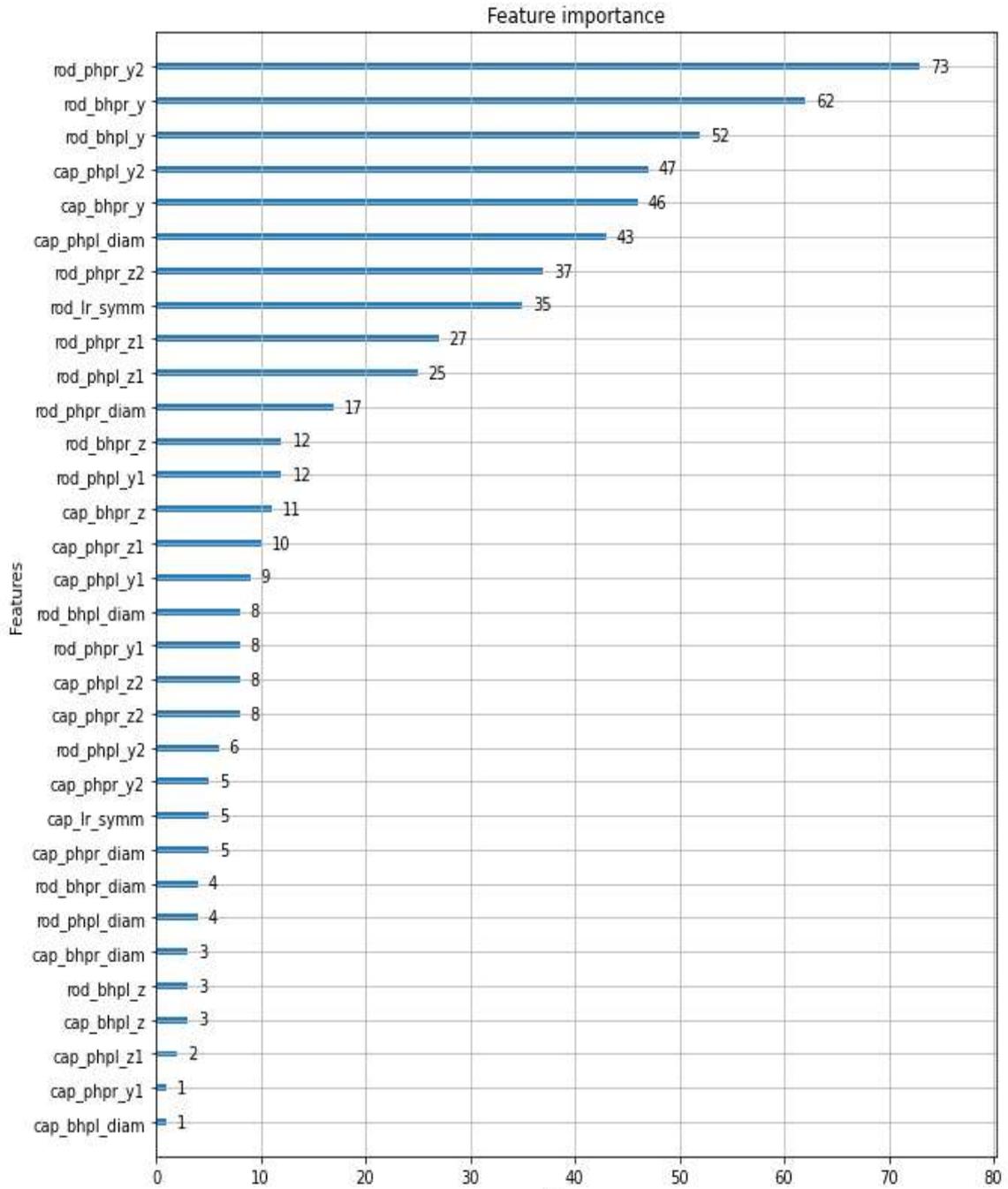
$$F1\text{스코어} = \frac{2 \times \text{정밀도} \times \text{재현율}}{\text{정밀도} + \text{재현율}} = \frac{2PR}{P+R}$$

F1 값의 논리는 조화평균에 입각한 것으로, 정밀도와 재현율 사이에 하나가 높아지면 다른 하나가 낮아지는 상황이 자주 발생하기에 이를 보정하기 위해 개발되었다. 본래는  $F\beta$  값의 일종으로,  $\beta$  값이 양수일 때는 재현율에, 음수일 때는 정밀도에 가중치를 부여한다. 따라서 F1 값은 정밀도와 재현율에 똑같이 1배의 가중치를 준  $F\beta$  값을 의미한다. 만일 재현율에 2배의 가중치를 주어 조화평균했다면 그때는 F2 값이 된다.

#### 나) XGBoost 피처의 중요도에 대한 시각화 구현 검증

그림 4-52을 보면, F1 스코어 40 이상을 상위 순위로 살펴보면 합격/불합격 결정 요인으로 rod\_phpr\_y2(우측 로드 핀홀y2), rod\_bhpr\_y(우로드 볼트홀 y), rod\_bhpl\_y(좌로드 볼트홀y)가 F1 스코어가 73, 62, 52로 가장 중요한 역할을 하고 있는 것을 확인할 수 있다. 주요 공정관리 변수인 rod\_lr\_symm(우측 로드 볼트홀 대칭도)도 F1 스코어가 43으로 6번째로 높은 순위를 차지하고 있다. 이 부분에 대한 불량 요인에 대한 공정 개선 및 관리가 향후 합격/불합격 관리에 결정적 요인이 됨을 알 수 있다. F1 스코어가 가장 높은 rod\_phpr\_y2 등 주요 변수 등에 대해 AI 알고리즘에 의한 피처 중요도 결과와 SPC통계분석과 비교 검증할 필요가 있다.

82) Fawcett, T. op.cit., pp. 861-874.



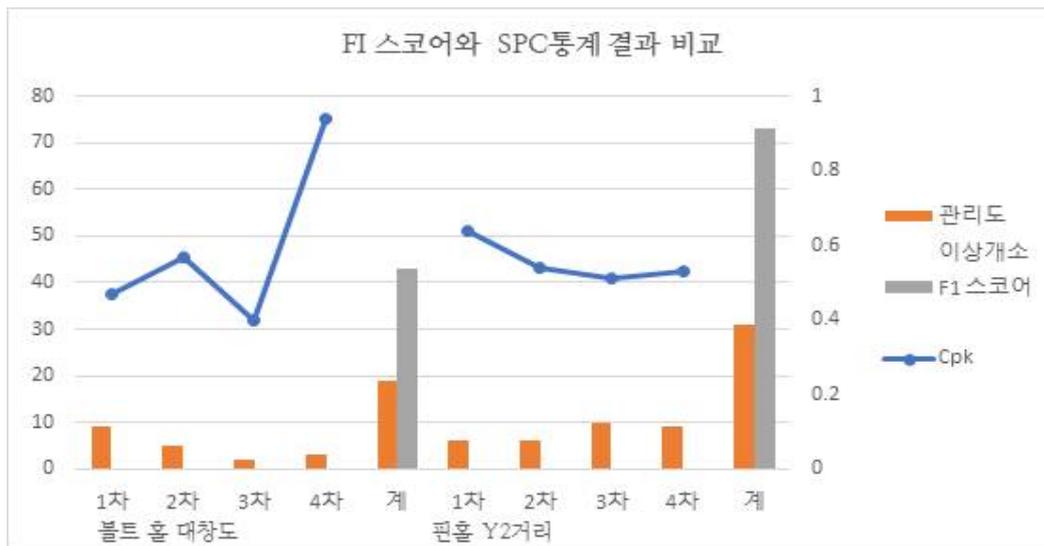
[그림 4-52] XGBoost를 통한 각 피처의 중요도

#### 4) XGBoost 피쳐모델 예측결과와 SPC통계분석의 상호 비교 검증

우선 F1 스코어가 가장 높은 rod\_phpr\_y2(우측 로드 핀홀y2)와 주요 공정 관리 변수인 rod\_lr\_symm(우측 로드 볼트홀 대칭도)에 대한 SPC분석(Cpk,  $\bar{X}$ -R/P관리도) 결과를 XGBoost의 피쳐 중요도 결과와의 비교한다. 비교 검증을 통해 XGBoost의 피쳐 시각화에 따른 중요도의 유효성을 검증한다.

[표 4-17] F1 스코어와 SPC 분석 비교표

구 분	XGBoost F1 스코어	SPC분석 결과		
		구분	Cpk	관리도 이상개소
우측 볼트 홀 대칭도	43	1차	0.47	9
		2차	0.57	5
		3차	0.4	2
		4차	0.94	3
		계	.	19
우측 핀홀 y2	73	1차	0.64	6
		2차	0.54	6
		3차	0.51	10
		4차	0.53	9
		계	.	31



[그림 4-53] F1 스코어와 SPC분석 비교표

우측 볼트홀 대칭도의 공정능력은 표 4-17, 그림 4-53과 같이 1~3차까지는 Cpk 0.67 이하로 공정능력이 부족하였으나, 4차에는 Cpk 0.94로 1.0에 근접할 정도로 정상 수준으로 접어들었다.  $\bar{X}$ -R, P관리도에서도 확인하듯 이상개소도 현저하게 줄어 안정화 수준으로 진행되는 것으로 보인다.

우측 핀홀 y2는 표 4-17, 그림 4-53과 같이 1~4차에 걸쳐 전차수가 Cpk 0.67이하로 공정능력이 매우 부족하며 관리도 이상개소도 증가 추세여서 공정이 불안정한 상태이다. 두 가지 변수요인을 비교해보면, 우측 볼트홀이 우측 핀홀 y2보다 공정능력이 안정적이고, 우측 볼트홀이 이상개소 합계가 19개인 데 비해 우측 핀홀 y2가 31개로 공정관리가 미비한 편이다.

앞에서 설명한 바와 같이 XGBoost의 피쳐 중요도의 예측결과인 F1 스코어도 표 4-17, 그림 4-53과 같이 우측 볼트홀 43, 우측 핀홀 73이 부합한다고 볼 수 있고 Cpk,  $\bar{X}$ -R, P관리도의 통계 결과로 비교 검증하였다,

XGBoost의 피쳐 중요도의 시각화 기능으로 많은 변수 주요변수를 빠르게 예측하여 감지한 후 SPC 통계로 해당 변수의 상세분석 및 검증을 통해 공정의 이상요인 및 변동요인을 찾아 공정개선 진행하는 사후 유지관리 수행하는 프로세스와 시스템을 구축하기 위한 방안과 역할 구분이 필요할 것으로 보인다.

## 제 5 절 연구의 성과

본 연구에서 제시된 제조데이터 수집분석 및 활용 시스템은, 스마트 제조 공정과 측정검사시스템에서 데이터를 추출한 후, 빅데이터 수집처리 시스템을 거쳐 클라우드 서비스와 연결한 SPC(통계적 공정관리) 및 AI 학습 분석을 활용하는 아키텍처로 설계하였다. 구현 적용 분야는 기계가공 공정의 스마트공장 고도화를 설계하고 구현하였으며 이에 따른 연구의 시사점과 성과는 다음과 같다.

### 1) 빅데이터 수집분석 시스템의 설계 및 구현의 시사점과 성과

가) 제조공정과 제조데이터에 대한 라이브러리를 구축하여 공동으로 활용이 가능하게 가공 데이터, 공구 수명 데이터, 장비 정보 및 로그 데이터를 수집하는 레이아웃 재배치와 측정검사 시스템을 설계한 후 스마트 제조공정을 완성하였고 실시간 데이터 수집 및 분석을 위한 제조환경의 인프라를 구축하였다.

나) IT, OT영역 간의 네트워크와 통신의 한계로 중소기업의 제조공정 스마트화에 제약이 있어 실시간 빅데이터 수집의 연구 한계가 있었으나, 본 연구는 가)의 내용과 함께 표준 통신 프로토콜 OPC-UA를 기반으로 한 클라이언트/서버 기능을 반영한 IIoT디바이스를 적용 설계하였고, IoT 통합 모듈 기술과 OPC-UA기반 네트워크를 적용하였다.

다) 클라우드 서비스와 연계한 빅데이터 수집 및 분석시스템의 구조와 3계층 구조의 스마트 클라이언트 애플리케이션 브라우저 개념을 적용한 데이터 수집 SW를 개발 및 설계하였다. 또 분석대상 데이터의 분석 장표를 기반으로 시스템 설계를 반영하고 최신 ICT기술을 적용하여 스마트제조 경쟁력이 향상되었다고 본다.

라) 상기 내용으로 구축한 MES 내 실시간 데이터베이스 및 추출 솔루션으로 빅데이터 분류와 저장을 통한 체계적인 데이터 축적 및 추출시스템의 구축함에 따라 장기적인 측면에서 생산성 향상과 품질 고도화를 위한 사전 기반이 마련되었다고 본다.

## 2) SPC & AI모델링 분석의 설계 및 구현의 시사점과 성과

가) MES의 TQC 솔루션에 Cpk값,  $\bar{X}$ -R 관리도 등의 SPC 수식과 시각화를 적용하여, SPC통계 솔루션에서 실시간의 분석 결과 값과 시각화도 함께 구현되었다. 해석 기준에 따른 알람기능과 같은 후속 조치 등으로 사후관리가 가능하며 SPC통계 솔루션의 설계와 구현으로 통계분석과 AI모델링 분석을 통한 상호 검증과 공정관리에 대한 역할 분담이 가능하다.

나) AI 예측정확도의 알고리즘은 로지스틱, 서포트 벡터 머신, 랜덤 포레스트로 설계하였고, AI 학습 결과, RF가 88%로 가장 높은 성능을 보였다. 향후 데이터가 축적되고 반복 학습이 지속해서 이루어지면, 랜덤 포레스트 알고리즘의 정확도가 90% 이상을 달성할 것임을 기대할 수 있다.

다) AI 모델링 시각화 분석을 XGBoost 알고리즘으로 설계하였고, 피처의 주요변수 중요도에 대한 F1 스코어가 SPC분석의 결과 값과 부합한 것으로 평가되었다. SPC통계의 검증으로 불량요인을 명확히 하고 AI모델링의 예측결과에 대한 교차 검증을 수행하면서 사후 유지관리의 수행이 가능할 것으로 예상된다.

라) 커넥팅로드의 합격/불합격과 공구 교환 시기의 적절한 예측을 위해 기존의 계량 기법을 활용하는 대신 추세와 변동성을 알 수 있는 실질적 지표를 입력 변수로 적용하여 품질 관리에서 기존 모델보다 좋은 성능을 기록하고 있는 인공지능 기반 모델을 적용하였다.

마) 모델 성능 평가를 위해 다양한 평가지표를 활용하고, 다양한 머신러닝 모델들과의 비교를 통해 모델의 객관적인 성능을 제시했다.

바) 앙상블 알고리즘의 경우, 대표적인 알고리즘 종류인 배깅, 랜덤포레스트, XGBoost를 선정해 설계하였고, 단일 알고리즘 모델인 의사결정나무 및 앙상블 알고리즘들을 비교해 성능이 가장 좋은 예측 모델을 선정하고자 하였다.

## 3) 연구의 기대효과

가) 지금까지는 수집 및 저장기술의 한계로 데이터의 표본 추출을 통해 SPC 통계도구 위주로 오프라인 개념의 통계분석을 진행해 왔으나, 빅데이터의 실시간 수집 및 저장이 가능해지면서 실시간 SPC 통계분석과 AI모델링 분석이 가능하게 되었으므로 SPC, AI모델링 분석 및 시각화 구현으로 생산성과 품질 향

상이 이어져 제조경쟁력까지 달성될 것으로 예상된다.

나) 실시간 수집, 분석 및 시각화 시스템 구축으로 품질에 대한 객관적 근거 확보와 빅데이터의 수집분석이 가능함에 따라 양산 품질의 안정화 및 품질의 고도화가 예상된다.

다) 커넥팅로드의 합격/불합격 예측을 통한 품질 관리 등 다양한 영역에서 인공지능기반 모델이 기존 계량 분석보다 나은 예측 성과를 보인다. 합격/불합격에 대해 예측을 하는 합격/불합격 선예지 시스템을 통해, 장비와 공구 등 영향을 미치는 요인들의 개선을 통한 품질 개선과 비용절감이 예상된다.

## 제 5 장 결 론

최근 스마트공장은 빅데이터와 AI 기술을 활용한 고도화의 추세로 전환하고 있다. 이에 따라 빅데이터 수집 및 처리 기술, IoT통합 기술, 기계 학습 및 AI 모델링에 대한 선행 연구가 수행되어 왔다. 이러한 기존 문헌 검토를 바탕으로, 본 연구의 범위와 목적에 맞게 실시간 제조 데이터 수집, 분석 및 활용 시스템을 설계한 후, 실시간 제조데이터 수집처리시스템에 기반한 SPC와 AI모델링 분석에 의한 활용시스템을 설계 및 구현하였다. 엄격한 공차와 내구성을 충족하기 위해서 가공정밀도를 높이는 품질 고도화와 생산성 향상을 목적으로 기계가공 분야의 스마트공장 고도화를 아래와 같이 설계하고 구현하였다.

첫째, 가공공정 스마트화를 위한 레이아웃 재배치 및 공정개선, 신설비, 측정검사시스템 도입을 통해 신 공정을 설계와 구현하였다. 신 공정 지그 5종, 로봇 및 자동 컨베이어 이송장비, MCT 가공설비 5대, 부시 압력 측정 및 조립시스템, 공구 모니터링 및 검사시스템을 신규 도입 또는 설계, 제작, 설치하였다. 가공설비에서 가공결과 값, 공구정보를 추출하고 부시 압력 측정 및 조립시스템에서 부시내경, 줌쇠, 압입압력 데이터를 발생시킨다. 공구 모니터링 및 검사시스템에서 커넥팅로드의 황삭, 정삭치수를 추출한 다음 OPC-UA 프로토콜에 기반을 둔 유무선 네트워크를 통하여 MES로 데이터를 전송한다.

둘째, 실시간으로 제조데이터의 수집 및 분석시스템을 설계하기 위해서, 먼저 표준 통신 프로토콜인 OPC-UA를 기반으로 한 IoT 통합 모듈로 실시간 데이터 수집을 위한 네트워크를 설계와 구현하였다. 가공설비, 부시 압력 측정 및 조립시스템, 공구 모니터링 및 검사시스템에서 추출한 데이터를 IIoT 디바이스로 보내고, OPC 기반 유무선 네트워크를 통해서 PoP 서버로 수집한다. 마지막으로 웹서비스를 통해서 클라우드 서비스로 저장, 분류한다.

셋째, 데이터 소프트웨어는 3계층 구조와 스마트 클라이언트 개념의 애플리케이션 브라우저를 클라우드와 연계하여 설계하고 구현하였다. 신규 스마트화 공정에 대한 사용자의 요구사항을 반영한 수집 분석 장표를 만들고, 이를 반영하여 분석데이터 수집·저장을 위한 데이터베이스 및 추출 솔루션을 설계하고 구

현하였다. 클라우드의 MES 비즈니스 로직에 의해서 TQC통계 애플리케이션으로 데이터가 추출되어 운영된다.

넷째, SPC 통계와 시각화 솔루션을 설계한 후 구현하였고 주요 공정관리 변수와 AI 모델링으로 도출된 최상위 품질 영향변수에 대해서 교차 검증을 수행하였다. Cpk값,  $\bar{X}$ -R, P관리도로 통계적으로 분석하였고 AI모델링에 대한 SPC 교차 검증을 진행하였다.

다섯째, AI 모델링, 검증 및 시각화 분석을 설계하고 수행하였다. 변수들에 대한 상대적 상관분석을 실시하고, 알고리즘 간 예측정확도를 비교하였다. 주요 변수의 중요도에 대한 F1 스코어를 SPC 통계와 교차 검증한 결과, 부합한 것으로 판명되었다.

상기 내용을 종합하면 첫째, 구축된 MES 내 실시간 데이터 수집, 저장을 위해 데이터베이스 및 추출 솔루션을 설계해서 구현해 활용 중이다. 빅데이터 분류, 저장을 통한 체계적인 데이터 축적 및 추출 등에 공동 활용할 수 있음으로 장기적인 측면에서 생산성과 품질 향상을 위한 사전 기반이 마련되었다고 볼 수 있다. 둘째, AI 모델링에서 실시간으로 구현되는 주요변수 중요도 및 시각화 분석 결과에 대해, SPC분석과 상호 비교를 통해 불량요인을 명확히 하고, AI모델링의 예측결과에 대한 교차 검증을 수행할 수 있다. 실시간으로 SPC통계와 시각화 솔루션에 대한 설계와 구현하면서 해석 기준에 따른 알람기능과 같은 후속 조치 등으로 사후 유지관리가 가능하다.

연구 진행 과정을 검토해 보자면, 데이터 수집 및 분석 시스템, SPC통계 솔루션, AI모델링 분석과 시각화까지 대부분이 설계하고 구현되었으나, 그 과정에서 아쉬운 점이나 한계를 되돌아보고 현재 기술 트렌드와의 비교 등을 검토해 향후 연구방향을 정립하고자 한다. 먼저 연구진행 과정 중 공정에서 설비 도입 및 설치와 시운전까지의 진도가 다소 지연되어 데이터의 충분한 양을 확보하는데 어려움이 있었다. 연구수행에 필요한 양은 확보할 수 있었고 공구 수명과 교환시기에 대한 예측정확도 검증과 예측프로그램 개발은 수행하였으나 유의한 결과 값의 도출에는 한계가 있었다. 공구정보 또한 커넥팅로드보다 충분한 데이터의 확보와 반복 학습이 필요하다. 커넥팅로드의 주요변수들에 대한 SPC통계 분석과 AI모델링 분석이 진행되어 이 과정에서 AI 모델링 솔루션은 설계 수행

과 구현은 되었으나 보완이 필요하다. 향후 커넥팅로드와 공구정보에 대한 충분한 데이터를 확보한 상태에서 지속적인 반복 학습이 이루어지면 예측정확도나 SPC 분석 결과는 개선될 것이라고 예상된다.

IoT 플랫폼 기술의 전 단계인 OPC-UA 기반의 IoT 통합을 포함한 유무선 네트워크로 MES과 클라우드까지 데이터를 전송하고 분류 및 저장한 다음, SPC 통계와 AI학습 모델링에 클라우드까지 연계한 솔루션을 설계와 구현하였다. AI학습 모델링과 SPC분석 실시 후 솔루션을 구현하고 별도로 교차 검증까지 진행하였으나 향후 최적화 및 보완이 필요하다. 최근에는 스마트센서의 IoT 기술, 유무선 통신과 결합한 IoT 통합 및 IoT 플랫폼의 등장으로 융합기술이 엣지 컴퓨팅, 클라우드서비스, 5G, TSN(Time Sensitive Network) 등의 최신기술과 융합하여 진보하고 있고, AI학습 모델링과 SPC분석 또한 최신기술과 연계된 융합 플랫폼으로서 사용자 맞춤 서비스가 제공되기 시작하고 있다.

향후 과제는 첫째, 공정의 안정화 후 충분한 데이터가 확보되고 최소한의 반복 학습이 이루어진 시기에 빅데이터 수집 및 분석시스템을 안정화하여 AI학습 모델 적용을 최적화하고 SPC 교차 검증 프로그램 및 솔루션을 추가로 고도화할 필요가 있다. 둘째, AI 학습 모델링 고도화를 통해 중소기업 규모에 적합한 특정 업종과 공정에 적용할 수 있는 CPS 모듈 개념의 모델 설계로 연구범위를 확장할 필요가 있다. 셋째, 공장 서버와 연동된 엣지 컴퓨팅과 클라우드, 5G, TSN 등의 최신기술을 연계한 스마트 고도화 2단계를 연구할 필요가 있다.

## 참 고 문 헌

### 1. 국내문헌

- 강지현. (2019). IoT 빅데이터 수집을 위한 IP기반 이기종 네트워크 인터페이스 연동 게이트웨이. 한국정보통신학회논문지, Vol. 3, No. 2, pp. 173~178.
- 권대욱. (2018). Smart Factory 구현을 위한 자동화 시스템의 Engineering 설계. 스마트팩토리 콘퍼런스 2018.
- 권순선.(2020). 인공지능과 빅데이터 기술동향. TTA journal Vol. 187, pp. 40~41.
- 김영주. (2010). .Net 기반에서 엔터프라이즈 애플리케이션 프레임워크 구현. 인천대학교 대학원 박사학위 논문, pp. 48.
- 김영주, 최정민, 이병수. (2010). 엔터프라이즈 아키텍처를 위한 .Net 기반의 스마트 클라이언트 애플리케이션 브라우저 설계 및 구현. The Journal of Korean Institute of Information Technology. Vol. 8, No. 12, pp. 203-209.
- 김원준. (2017). "로봇 및 산업용 장비의 통합 제어를 위한 DSP에서의 MODBUS 프로토콜 구현과 성능 검증", 서울과학기술대학교 석사학위 논문.
- 김재민, 김성렬, 자오밍하오, 유영환. (2018). "스마트공장을 위한 환경 데이터 수집시스템 설계," 한국정보과학회 한국컴퓨터종합학술대회(KSC).
- 김재표, 김승천. (2020). 중소기업 스마트공장 사전준비를 위한 이기종 시스템에 대한 네트워크 인터페이스 방안의 연구. 한국융합학회. Vol. 11, No 11, pp. 55-61.
- 나영식. (2020). 제조용 IoT. KISTEP 기술동향브리프. 2020-10호. pp. 4-18.
- 류가애, 류관희. (2018). 영상 데이터기반의 CNN을 이용한 제조 공정데이터 분류 적용에 대한 연구. Journal of Information Technology and

- Architecture. Vol. 15. No. 3, pp. 337~343.
- 문상국. (2006). 산업기기 네트워크에서의 실시간 처리와 필드버스 통신의 문제점 분석. In Proceedings of KIIT Conference, pp. 303-306.
- 문상국. (2007). 무선 통신 기술의 산업기기 응용에 대한 고찰. 한국정보통신학회논문지, 11(1), 157-162.
- 박건수, 트란트링턴, 당반치엔, 길기종, 신용빈, 최재원, 김종욱. (2018). IoT 디바이스를 이용한 클라우드 플랫폼 기반의 실시간 공장 모니터링 시스템 개발. 대한임베디드공학회, 제 13권 제 1호. pp. 25-32.
- 박태양. (2018). “산업용 이기종 엔지니어링 툴의 데이터 교환을 위한 Automation ML과 OPC UA 통합 데모시스템 구현”, 한양대학교 대학원 석사학위 논문.
- 사물인터넷표준연구실. (2020). 사물인터넷 기술동향. IoT표준연구실 IT대학 컴퓨터학부 경북대학교, pp. 12-13.
- 손영욱, 김영국. (2018). 예지정비를 위한 클라우드 기반 스마트공장 IoT 플랫폼 구현에 관한 연구. 한국창업학회
- 신동진, 박지훈, 김주호, 곽광진, 박정민, 김정준. (2019). IoT 환경을 위한 빅데이터 기반 센서 데이터 처리 및 분석. The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 19, No. 1, pp. 117-126.
- 신영환, 윤주식, 서순호, & 정종문. (2017). 딥러닝 및 증강현실을 이용한 재난 대응 역량 강화를 위한 네트워크 자원 확보 방안. 인터넷정보학회논문지, 18(5), 69-77.
- 오승엽, 김웅기, 성민영. (2018). PLC를 이용한 산업용 제어 응용을 지원하는 OPC-UA 프레임 워크의 설계 및 구현. 한국정보과학회. Vol. 24, No. 10, pp. 513-526.
- 유홍식, 박유현. (2021). 이기종 공작기계 데이터 수집을 위한 IIoT 엣지 디바이스 설계 및 구현. 인터넷전자상거래연구, 제21권, 1호, pp. 23-32.
- 이강배, 박성호, 성상하, 박도명. (2019). 기계학습 기법을 이용한 CNC 공구 마모도 예측에 관한 연구. Journal of the Korea Convergence Society,

Vol. 10, No. 11, pp. 15-21.

- 이선표, 고희석, 정희운, 박승범 (2020). 중소기업을 위한 제조 빅데이터 분석 플랫폼 구축. *Journal of Information Technology and Architecture*. Vol. 17. No. 3, pp. 239-253.
- 이성준, 김춘경, & 이재덕. (2017). OPC UA를 이용한 스마트센서 시스템 개발. *대한전기학회 학술대회 논문집*, 1338-1340.
- 이용민, 이원복, & 이승호. (2019). 스마트 팩토리 구축을 위한 프로토콜 개발. *전기전자학회 논문지*, 23(3), pp. 1096-1099.
- 이용수, 정종필. (2014). 스마트 제조를 위한 AAS 와 OPC UA 기반 설비모니터링 시스템의 설계 및 구현. *한국인터넷방송통신학회 논문지*, 21(2), pp. 41-47.
- 이용환, 서진형. (2020). 온-디바이스 엣지 컴퓨터 기반 설비예지보존 솔루션 구축에 관한 연구(A Study on the Build of Equipment Predictive Maintenance Solutions Based on On-device Edge Computer). *한국컴퓨터정보학회 논문지*, 25(4), pp. 165-172.
- 이진휘. (2020). AI기술 동향과 오픈소스. *정보통신산업진흥원 이슈리포트*, 2020-제3호, pp. 6-10.
- 장원중, 조성인, 김수상, 김광용. (2018). 스마트팩토리에서 빅데이터 인프라 구축 방안에 관한 연구. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, Vol.8, No.10, pp. 11-23.
- 정성일, 전인걸. (2018). 소규모 스마트팩토리를 위한 저비용 통합 데이터 관리 시스템 구현. *정보통신산업진흥원*, No. S0608-18-1011.
- 정종필, 신광섭. (2020). 스마트팩토리 핵심기술 및 제조혁신 고도화 전략. *융합연구리뷰*. Vol. 6, No. 12, 4, pp. 10-21.
- 정지석. (2018). 4차 산업혁명 시대 ICT 기반 컨설팅 방법론. *한국컨설팅서비스협회*, pp. 3-24.
- 정해윤. (2021) 스마트팩토리 솔루션. *혁신성장품목보고서, IR협의회*, pp3.
- 중소기업 전략기술로드맵 2021~2023. (2021). pp. 170~173

- 중소기업기술정보진흥원. (2021). 중소기업 전략기술로드맵, 2021~2023 인공지능. pp. 26~27,
- 한무명초, 이충권, 김양석. (2019). 제조 공정에서 센서와 머신러닝을 활용한 불량예측 방안에 대한 연구. *Entrue Journal of Information Technology*, Vol.17, No.1, pp. 89-98.
- 한승희. (2015). OPC UA 기반 IEC 61850 정보 모델 연동 기법, 서경대학교 대학원 석사학위 논문.
- 홍성일, 인치호. (2015). 저전력 무선통신 기반 다중 프로토콜 게이트웨이 시스템 설계. *한국통신학회논문지 제40권 제10호*.
- 황승연, 광경민, 신동진. (2019). 빅데이터 기반 실시간 불량품 발생 원인분석 및 설비교체주기 예측.
- 황승연, 신동진, 광광진, 김정준, 박정민. (2019). 스마트팩토리를 위한 빅데이터 기반 실시간 제조설비 데이터 처리. *The Journal of The Institute of Internet, Broadcasting and Communication*. Vol. 19, No. 5, pp. 219-227.
- <https://bkjeon1614.tistory.com/27>
- <http://www.dataanalytics.co.kr/>
- <http://msdn.microsoft.com/ko-kr/library/aa905318.aspx>
- <https://tensorflow.blog/머신러닝-1장-2장/1-3-머신러닝-시스템의-종류/>

## 2. 국외문헌

- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), pp. 5–32.
- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. pp. 785–794.
- Freedman, D. A. (2009). *Statistical Models: Theory and Practice*. Cambridge University Press. pp. 26.
- Hill, D. (2004). *The Architecture journal* 4, Microsoft Corporation October 2004.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), pp. 861–874.
- Hannelius, T., Salmenpera, M. and Kuikka, S. (2008). *Roadmap to adopting OPC UA*. Proc. of the 6th IEEE International Conference on Industrial Informatics (INDIN), pp. 756–761.
- Izaguirre, M. J. A. G., Lobov, A. and Lastra, J. L. M. (2011). *OPC-UA and DPWS interoperability for factory floor monitoring using complex event processing*. Proc. of the 9th IEEE International Conference on Industrial Informatics (INDIN), pp. 205–211.
- Kim, I., Kim, T., Sung, M., Tisserant, E., Bessard, L. and Choi, C. (2013). *An Opensource development environment for industrial automation with EtherCAT and PLCopen motion control*. Proc. of the IEEE 18th Conference on Emerging Technologies and Factory Automation (ETFa), IEEE, Cagliari, Italy, pp. 1–4.
- Kim, J., Kim, Y., Kim, S. (2021). A Study on Design of Real-time Big Data Collection and Analysis System based on OPC-UA for Smart Manufacturing of Machine Working, *International Journal of Internet, Broadcasting and Communication* Vol.13 No.4 pp. 122–127.
- Lange, J., Iwanitz, F. and Burke, T. J. (2010). *OPC, 4th Ed.*, VDE VERLAG GMBH,, Berlin, Germany.

- Reynolds, M. and Watson, K. (2002). .NET Enterprise Development in C#: 설계에서 배치까지(임성준 역), 정보문화사
- Moyne, J. and Armacost, M. (2017). *Big Data Analytics Applied to Semiconductor Manufacturing*. In 2017 Spring Meeting and 13th Global Congress on Process Safety. AIChE.
- Moyne, J. and Iskandar, J. (2017). *Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing*. Processes, 5, 39.
- Moyne, J., Schulze, B., Iskandar, J. and Armacost, M. (2016). *Next generation advanced process control: Leveraging big data and prediction*. In 2016, 27th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC) IEEE, pp. 191–196.
- OPC Unified Architecture Specification. (2015). *Part 1: Overview and Concepts*, OPC Foundation.
- OPC Unified Architecture. (2016). Part 1: Overview and Concepts, IEC 62541-1.
- Quinlan, J. R. (1997). C5.0. <http://www.rulequest.com/see5-info.html>.
- Ramsay, J. O., Hooker, G. and Graves, S. (2009). *Functional data analysis with R and MATLAB*: Springer Science & Business Media.
- Romero-Torres, S., Moyne, J. and Kidambi, M. (2017). *Towards Pharma 4.0: Leveraging Lessons and Innovation from Silicon Valley*. American Pharmaceutical Review, 5.
- Beudert, R. (2015). *Machinery Information Base Data Structure*. Presented at the ODVA 2015 Industry Conference.
- Schwarz, M. H. and Börcsök J. (2013). *A survey on OPC and OPC-UA: About the standard, developments and investigations*. Proc. of the XXIV International Symposium on Information, Communication and Automation Technologies (ICAT), pp. 1–6.
- SEMI, S. (2014). *E133-1014-SEMI Standard Specification for Automated*

*Process Control Systems Interface*. Milpitas, CA,(Semiconductor Equipment and Materials.)

Shang, C. and You, F. (2019). *Data analytics and machine learning for smart process manufacturing: recent advances and perspectives in the big data era*, *Engineering*, 5, pp. 1010 – 1016.

Wan, J., Yang, J., Wang, Z. and Hua, Q. (2018). *Artificial Intelligence for Cloud-Assisted Smart Factory*. *IEEE Access*, vol. 6, pp. 55419–55430.

Wuest, T., Irgens, C. and Thoben, K. D. (2014). An approach to monitoring quality in manufacturing using supervised machine learning on product state data. *Journal of Intelligent Manufacturing*, 25, pp. 1167 – 1180.

Vapnik, V. (2000). Section 5.6. Support Vector Machines.

AWS(<http://aws.amazon.com>), 2019

<https://manisha-sirsat.blogspot.com/2019/04/confusion-matrix.html>

<http://msdn.microsoft.com/en-us/library/aa192611.aspx>

<http://msdn.microsoft.com/ko-kr/library/aa905318.aspx>

<http://www.modbus.org/specs.php>

<http://r4stats.com/articles/popularity/>

# 부 록

[부록 1] 볼트홀 대칭도 1차, 2차 Cpk 및 X바-R관리도 상세 분석표

표본의 크기 100					표본의 크기 100						
평균					0.044573	평균					0.041687
분산					0.0015049697	분산					0.0012528476
표준편차					0.0388919035	표준편차					0.0354849701
Cp					*	Cp					*
Cpu					0.49212303	Cpu					0.59218029
Cpk					Cpk	Cpk					Cpk
Pp					0.47505175	Pp					0.54777182
Ppk					0.47505175	Ppk					0.54777182
부적합품률(ppm)					120000 ppm	부적합품률(ppm)					70000 ppm
잠재 부적합품률(ppm)					77056.17 ppm	잠재 부적합품률(ppm)					50158.84 ppm
장기 시그마수준					1.42515524	장기 시그마수준					1.64331546
단기 시그마수준					2.92515524	단기 시그마수준					3.14331546

xbar	관리상한	중심선	관리하한	판	xbar	관리상한	중심선	관리하한	판
0.0139	0.0833	0.041687	0.0001	G	0.0139	0.0833	0.041687	0.0001	G
0.0122	0.0833	0.041687	0.0001	G	0.0122	0.0833	0.041687	0.0001	G
0.0236	0.0833	0.041687	0.0001	G	0.0236	0.0833	0.041687	0.0001	G
0.0348	0.0833	0.041687	0.0001		0.0348	0.0833	0.041687	0.0001	
0.0206	0.0833	0.041687	0.0001		0.0206	0.0833	0.041687	0.0001	
0.0597	0.0833	0.041687	0.0001		0.0597	0.0833	0.041687	0.0001	
0.0594	0.0833	0.041687	0.0001		0.0594	0.0833	0.041687	0.0001	
0.0841	0.0833	0.041687	0.0001	*	0.0841	0.0833	0.041687	0.0001	*
0.0423	0.0833	0.041687	0.0001		0.0423	0.0833	0.041687	0.0001	
0.0455	0.0833	0.041687	0.0001		0.0455	0.0833	0.041687	0.0001	
0.046	0.0833	0.041687	0.0001		0.046	0.0833	0.041687	0.0001	
0.0324	0.0833	0.041687	0.0001		0.0324	0.0833	0.041687	0.0001	
0.0466	0.0833	0.041687	0.0001		0.0466	0.0833	0.041687	0.0001	
0.0474	0.0833	0.041687	0.0001		0.0474	0.0833	0.041687	0.0001	
0.0463	0.0833	0.041687	0.0001		0.0463	0.0833	0.041687	0.0001	
0.0408	0.0833	0.041687	0.0001		0.0408	0.0833	0.041687	0.0001	
0.0301	0.0833	0.041687	0.0001		0.0301	0.0833	0.041687	0.0001	
0.0455	0.0833	0.041687	0.0001		0.0455	0.0833	0.041687	0.0001	
0.073	0.0833	0.041687	0.0001		0.073	0.0833	0.041687	0.0001	
0.0295	0.0833	0.041687	0.0001		0.0295	0.0833	0.041687	0.0001	

R	관리상한	중심선	관리하한	판	R	관리상한	중심선	관리하한	관리도
0.038	0.15246	0.072085		0	0.038	0.15246	0.072085		0
0.0342	0.15246	0.072085		0	0.0342	0.15246	0.072085		0
0.0461	0.15246	0.072085		0	0.0461	0.15246	0.072085		0
0.0192	0.15246	0.072085		0	0.0192	0.15246	0.072085		0
0.0493	0.15246	0.072085		0	0.0493	0.15246	0.072085		0
0.1244	0.15246	0.072085		0 A	0.1244	0.15246	0.072085		0 A
0.1374	0.15246	0.072085		0 A	0.1374	0.15246	0.072085		0 A
0.0816	0.15246	0.072085		0 A	0.0816	0.15246	0.072085		0 A
0.095	0.15246	0.072085		0 A	0.095	0.15246	0.072085		0 A
0.0734	0.15246	0.072085		0 A	0.0734	0.15246	0.072085		0 A
0.0926	0.15246	0.072085		0 A	0.0926	0.15246	0.072085		0 A
0.0777	0.15246	0.072085		0 A	0.0777	0.15246	0.072085		0 A
0.0651	0.15246	0.072085		0	0.0651	0.15246	0.072085		0
0.0648	0.15246	0.072085		0	0.0648	0.15246	0.072085		0
0.0609	0.15246	0.072085		0	0.0609	0.15246	0.072085		0
0.0578	0.15246	0.072085		0	0.0578	0.15246	0.072085		0
0.0664	0.15246	0.072085		0	0.0664	0.15246	0.072085		0
0.0764	0.15246	0.072085		0	0.0764	0.15246	0.072085		0
0.107	0.15246	0.072085		0	0.107	0.15246	0.072085		0
0.0744	0.15246	0.072085		0	0.0744	0.15246	0.072085		0

[부록 2] 볼트홀 대칭도 3차, 4차 Cpk 및 X바-R관리도 상세 분석표

표본의 크기	100	표본의 크기	100
평균	0.053507	평균	0.038377
분산	0.0016569162	분산	0.0004735107
표준편차	0.0408080265	표준편차	0.0218152513
Cp	*	Cp	*
Cpu	0.4071796	Cpu	0.94798475
Cpk	Cpk	Cpk	Cpk
Pp	0.37977006	Pp	0.94158897
Ppk	0.37977006	Ppk	0.94158897
부적합품률(ppm)	140000 ppm	부적합품률(ppm)	0 ppm
잠재 부적합품률(ppm)	127286.91 ppm	잠재 부적합품률(ppm)	2365.75 ppm
장기 시그마수준	1.13931018	장기 시그마수준	2.82476691
단기 시그마수준	2.63931018	단기 시그마수준	4.32476691

xbar	관리상한	중심선	관리하한	관리도 판정코드
0.0324	0.1044	0.053507	0.0027	
0.0841	0.1044	0.053507	0.0027	
0.0811	0.1044	0.053507	0.0027	
0.0769	0.1044	0.053507	0.0027	
0.0605	0.1044	0.053507	0.0027	
0.0808	0.1044	0.053507	0.0027	
0.0285	0.1044	0.053507	0.0027	
0.037	0.1044	0.053507	0.0027	
0.0335	0.1044	0.053507	0.0027	
0.0498	0.1044	0.053507	0.0027	
0.0412	0.1044	0.053507	0.0027	
0.0444	0.1044	0.053507	0.0027	
0.0608	0.1044	0.053507	0.0027	
0.0313	0.1044	0.053507	0.0027	
0.0504	0.1044	0.053507	0.0027	
0.0661	0.1044	0.053507	0.0027	
0.0309	0.1044	0.053507	0.0027	
0.0303	0.1044	0.053507	0.0027	
0.0737	0.1044	0.053507	0.0027	
0.0765	0.1044	0.053507	0.0027	

xbar	관리상한	중심선	관리하한	관리도 판정코드
0.078	0.0577	0.038377	0.019	*
0.0391	0.0577	0.038377	0.019	
0.0304	0.0577	0.038377	0.019	
0.0437	0.0577	0.038377	0.019	
0.0474	0.0577	0.038377	0.019	
0.0295	0.0577	0.038377	0.019	
0.0362	0.0577	0.038377	0.019	
0.0416	0.0577	0.038377	0.019	
0.0156	0.0577	0.038377	0.019	*
0.0107	0.0577	0.038377	0.019	*
0.015	0.0577	0.038377	0.019	*
0.0065	0.0577	0.038377	0.019	*
0.0439	0.0577	0.038377	0.019	
0.0504	0.0577	0.038377	0.019	
0.057	0.0577	0.038377	0.019	
0.0486	0.0577	0.038377	0.019	
0.0334	0.0577	0.038377	0.019	
0.0434	0.0577	0.038377	0.019	
0.0411	0.0577	0.038377	0.019	
0.0562	0.0577	0.038377	0.019	

R	관리상한	중심선	관리하한	관리도 판정코드
0.0383	0.18642	0.08814	0	
0.1104	0.18642	0.08814	0	
0.1257	0.18642	0.08814	0	
0.0927	0.18642	0.08814	0	
0.0298	0.18642	0.08814	0	
0.1427	0.18642	0.08814	0	
0.0477	0.18642	0.08814	0	
0.086	0.18642	0.08814	0	
0.0777	0.18642	0.08814	0	
0.0263	0.18642	0.08814	0	
0.0447	0.18642	0.08814	0	
0.065	0.18642	0.08814	0	
0.1181	0.18642	0.08814	0	
0.1251	0.18642	0.08814	0	
0.1676	0.18642	0.08814	0	
0.1354	0.18642	0.08814	0	
0.0636	0.18642	0.08814	0	
0.0339	0.18642	0.08814	0	
0.1281	0.18642	0.08814	0	
0.104	0.18642	0.08814	0	

R	관리상한	중심선	관리하한	관리도 판정코드
0.0367	0.07088	0.033515	0	
0.032	0.07088	0.033515	0	
0.0354	0.07088	0.033515	0	
0.0309	0.07088	0.033515	0	
0.0389	0.07088	0.033515	0	
0.0343	0.07088	0.033515	0	
0.046	0.07088	0.033515	0	
0.0147	0.07088	0.033515	0	
0.0252	0.07088	0.033515	0	
0.0194	0.07088	0.033515	0	
0.0344	0.07088	0.033515	0	
0.0117	0.07088	0.033515	0	
0.0391	0.07088	0.033515	0	
0.0332	0.07088	0.033515	0	
0.0292	0.07088	0.033515	0	
0.0219	0.07088	0.033515	0	
0.0383	0.07088	0.033515	0	
0.0379	0.07088	0.033515	0	
0.0727	0.07088	0.033515	0	*
n 0384	n 07088	n 033515	n	

[부록 3] 볼트홀 대칭도 P관리도 상세분석표

	P	관리상한	중심선	관리하한	관리
1	0	0.2963961	0.1	0	
2	0.0625	0.325	0.1	0	
3	0.07142857	0.34053512	0.1	0	
4	0	0.30124612	0.1	0	
5	0	0.37136021	0.1	0	
6	0.2	0.38460499	0.1	0	
7	0.11764706	0.31828206	0.1	0	
8	0.33333333	0.35980762	0.1	0	
9	0.36363636	0.37136021	0.1	0	
10	0	0.332379	0.1	0	
11	0	0.30124612	0.1	0	
12	0.44444444	0.31213203	0.1	0	*
13	0.07692308	0.34961509	0.1	0	
14	0	0.332379	0.1	0	
15	0.04761905	0.2963961	0.1	0	
16	0.11111111	0.31213203	0.1	0	
17	0.28571429	0.34053512	0.1	0	
18	0.08333333	0.35980762	0.1	0	
19	0.08108108	0.24795909	0.1	0	
20	0.2	0.30124612	0.1	0	
21	0.04166667	0.28371173	0.1	0	
22	0	0.34961509	0.1	0	
23	0.14285714	0.34053512	0.1	0	
24	0.125	0.28371173	0.1	0	
25	0	0.31828206	0.1	0	
26	0	0.34961509	0.1	0	

[부록 4] 우측핀홀 y2 1차, 2차 Cpk 및 Xbar-R관리도 상세 분석표

표본의 크기	100	표본의 크기	100
평균	0.083751	평균	0.0347919192
분산	0.0017645146	분산	0.0005683272
표준편차	0.0421122018	표준편차	0.0238998151
Cp	0.77316075	Cp	1.57964857
Cpk	0.64752986	Cpk	0.54959005
Cpu	0.89879163	Cpu	2.60970708
Cpl	0.64752986	Cpl	0.54959005
Pp	0.79153623	Pp	1.39471093
Ppk	0.66291951	Ppk	0.4852467
Ppu	0.92015295	Ppu	2.30417516
Ppl	0.66291951	Ppl	0.4852467
부적합률(ppm)	10000 ppm	부적합률(ppm)	0 ppm
잠재 부적합률(ppm)	26249.94 ppm	잠재 부적합률(ppm)	72732.24 ppm
장기 시그마수준	2.76045885	장기 시그마수준	6.91252548
단기 시그마수준	4.26045885	단기 시그마수준	8.41252548

xbar	관리상한	중심선	관리하한	관리도
0.0333	0.1298	0.083751	0.0377	*
0.0616	0.1298	0.083751	0.0377	
0.0803	0.1298	0.083751	0.0377	
0.1193	0.1298	0.083751	0.0377	
0.0355	0.1298	0.083751	0.0377	*
0.0904	0.1298	0.083751	0.0377	
0.0724	0.1298	0.083751	0.0377	
0.0888	0.1298	0.083751	0.0377	
0.1006	0.1298	0.083751	0.0377	
0.0969	0.1298	0.083751	0.0377	
0.0784	0.1298	0.083751	0.0377	
0.1479	0.1298	0.083751	0.0377	*
0.1232	0.1298	0.083751	0.0377	
0.0596	0.1298	0.083751	0.0377	
0.0578	0.1298	0.083751	0.0377	
0.0667	0.1298	0.083751	0.0377	
0.0849	0.1298	0.083751	0.0377	
0.0985	0.1298	0.083751	0.0377	
0.0974	0.1298	0.083751	0.0377	
0.0814	0.1298	0.083751	0.0377	

xbar	관리상한	중심선	관리하한	관리도
0.0273	0.0608	0.0347919	0.0088	
0.0512	0.0608	0.0347919	0.0088	
0.0127	0.0608	0.0347919	0.0088	
0.0267	0.0608	0.0347919	0.0088	
0.0419	0.0608	0.0347919	0.0088	
0.0159	0.0608	0.0347919	0.0088	
0.0342	0.0608	0.0347919	0.0088	
0.0177	0.0608	0.0347919	0.0088	
0.0378	0.0608	0.0347919	0.0088	
0.0197	0.0608	0.0347919	0.0088	
0.0222	0.0608	0.0347919	0.0088	
0.0418	0.0608	0.0347919	0.0088	
0.0281	0.0608	0.0347919	0.0088	
0.0272	0.0608	0.0347919	0.0088	
0.0253	0.0608	0.0347919	0.0088	
0.0676	0.0608	0.0347919	0.0088	*
0.0714	0.0608	0.0347919	0.0088	*
0.0469	0.0608	0.0347919	0.0088	
0.0396	0.0608	0.0347919	0.0088	
0.0407	0.0608	0.0347919	0.0088	

R	관리상한	중심선	관리하한	관리도
0.0564	0.16885	0.079835	0	
0.086	0.16885	0.079835	0	
0.0568	0.16885	0.079835	0	
0.1131	0.16885	0.079835	0	
0.0352	0.16885	0.079835	0	
0.155	0.16885	0.079835	0	
0.0352	0.16885	0.079835	0	
0.0917	0.16885	0.079835	0	
0.0724	0.16885	0.079835	0	
0.0495	0.16885	0.079835	0	
0.0982	0.16885	0.079835	0	
0.0914	0.16885	0.079835	0	
0.1245	0.16885	0.079835	0	
0.0189	0.16885	0.079835	0	
0.0757	0.16885	0.079835	0	
0.1127	0.16885	0.079835	0	
0.0574	0.16885	0.079835	0	
0.0724	0.16885	0.079835	0	
0.1003	0.16885	0.079835	0	
0.0939	0.16885	0.079835	0	

R	관리상한	중심선	관리하한	관리도
0.0349	0.09519	0.045005	0	
0.0388	0.09519	0.045005	0	
0.0364	0.09519	0.045005	0	
0.0519	0.09519	0.045005	0	
0.0528	0.09519	0.045005	0	
0.0253	0.09519	0.045005	0	
0.034	0.09519	0.045005	0	
0.0246	0.09519	0.045005	0	
0.0692	0.09519	0.045005	0	
0.0355	0.09519	0.045005	0	
0.0535	0.09519	0.045005	0	
0.0464	0.09519	0.045005	0	
0.0188	0.09519	0.045005	0	
0.0292	0.09519	0.045005	0	
0.034	0.09519	0.045005	0	
0.0628	0.09519	0.045005	0	
0.0678	0.09519	0.045005	0	
0.0669	0.09519	0.045005	0	
0.055	0.09519	0.045005	0	
0.0623	0.09519	0.045005	0	

[부록 5] 우측핀홀 y2, 3차, 4차 Cpk 및 X바-R관리도 상세 분석표

표본의 크기 100					표본의 크기 100						
평균					0.05635	평균					0.0538626263
분산					0.0019986272	분산					0.0024860765
표준편차					0.0448189025	표준편차					0.0499864811
Cp					0.90774481	Cp					0.99751031
Cpk					0.5115142	Cpk					0.53728525
Cpu					1.30397541	Cpu					1.45773538
Cpl					0.5115142	Cpl					0.53728525
Pp					0.74373381	Pp					0.66684697
Ppk					0.419094	Ppk					0.35918129
Ppu					1.06837363	Ppu					0.97451264
Ppl					0.419094	Ppl					0.35918129
부적합률(ppm)					10000 ppm	부적합률(ppm)					10000 ppm
잠재 부적합률(ppm)					105000.8 ppm	잠재 부적합률(ppm)					142349.08 ppm
장기 시그마수준					3.20512088	장기 시그마수준					2.92353793
단기 시그마수준					4.70512088	단기 시그마수준					4.42353793
xbar	관리상한	중심선	관리하한	관리도 범위	xbar	관리상한	중심선	관리하한	관리도 범위		
0.1268	0.0953	0.05635	0.0174	*	0.0098	0.0918	0.05386265	0.016	*		
0.1323	0.0953	0.05635	0.0174	* I	0.0186	0.0918	0.05386265	0.016	H		
0.0763	0.0953	0.05635	0.0174	I	0.0673	0.0918	0.05386265	0.016	H		
0.0382	0.0953	0.05635	0.0174	H I	0.0469	0.0918	0.05386265	0.016	H		
0.1056	0.0953	0.05635	0.0174	* H I	0.0318	0.0918	0.05386265	0.016	H		
0.0785	0.0953	0.05635	0.0174	H I	0.0724	0.0918	0.05386265	0.016	G H		
0.09	0.0953	0.05635	0.0174	G H I D	0.0195	0.0918	0.05386265	0.016	G H		
0.0794	0.0953	0.05635	0.0174	G H I C D	0.0195	0.0918	0.05386265	0.016	G H		
0.029	0.0953	0.05635	0.0174	G H I A C D	0.056	0.0918	0.05386265	0.016	G		
0.0274	0.0953	0.05635	0.0174	G H I A C D	0.0327	0.0918	0.05386265	0.016			
0.0284	0.0953	0.05635	0.0174	G H I A C D	0.0351	0.0918	0.05386265	0.016			
0.0423	0.0953	0.05635	0.0174	G H I A C D	0.0345	0.0918	0.05386265	0.016			
0.0378	0.0953	0.05635	0.0174	H I A C D	0.0152	0.0918	0.05386265	0.016	*		
0.0283	0.0953	0.05635	0.0174	H I A C D	0.0292	0.0918	0.05386265	0.016			
0.0515	0.0953	0.05635	0.0174	A H I C D	0.0325	0.0918	0.05386265	0.016			
0.0421	0.0953	0.05635	0.0174	A H I C D	0.0719	0.0918	0.05386265	0.016			
0.0291	0.0953	0.05635	0.0174	A H I C D	0.1159	0.0918	0.05386265	0.016	*		
0.0327	0.0953	0.05635	0.0174	A C I D H	0.1114	0.0918	0.05386265	0.016	*		
0.0305	0.0953	0.05635	0.0174	A C I D H	0.1095	0.0918	0.05386265	0.016	*		
0.0208	0.0953	0.05635	0.0174	A C D H I	0.1476	0.0918	0.05386265	0.016	*		
R	관리상한	중심선	관리하한	관리도 범위	R	관리상한	중심선	관리하한	관리도 범위		
0.1056	0.1427	0.06747	0		0.0152	0.13894	0.0656931313	0			
0.1377	0.1427	0.06747	0		0.0176	0.13894	0.0656931313	0			
0.0829	0.1427	0.06747	0		0.1204	0.13894	0.0656931313	0			
0.0703	0.1427	0.06747	0		0.0879	0.13894	0.0656931313	0			
0.0614	0.1427	0.06747	0		0.0356	0.13894	0.0656931313	0			
0.1548	0.1427	0.06747	0	*	0.2513	0.13894	0.0656931313	0	*		
0.0885	0.1427	0.06747	0		0.0376	0.13894	0.0656931313	0			
0.1056	0.1427	0.06747	0		0.0383	0.13894	0.0656931313	0			
0.0169	0.1427	0.06747	0		0.0275	0.13894	0.0656931313	0			
0.017	0.1427	0.06747	0		0.0741	0.13894	0.0656931313	0			
0.0245	0.1427	0.06747	0		0.0762	0.13894	0.0656931313	0			
0.0814	0.1427	0.06747	0		0.0515	0.13894	0.0656931313	0			
0.0745	0.1427	0.06747	0		0.0284	0.13894	0.0656931313	0			
0.07	0.1427	0.06747	0		0.0597	0.13894	0.0656931313	0			
0.0497	0.1427	0.06747	0		0.0561	0.13894	0.0656931313	0			
0.0203	0.1427	0.06747	0		0.1131	0.13894	0.0656931313	0			
0.0396	0.1427	0.06747	0		0.033	0.13894	0.0656931313	0			
0.0251	0.1427	0.06747	0		0.0607	0.13894	0.0656931313	0			
0.0734	0.1427	0.06747	0		0.0666	0.13894	0.0656931313	0			
0.0502	0.1427	0.06747	0		0.0631	0.13894	0.0656931313	0			

[부록 6] 우측핀홀 y2 P관리도 상세분석표

	p	관리상한	중심선	관리하한	관리도 판정코드
1	0.0625	0.44910088	0.16842105	0	
2	0.64285714	0.46848041	0.16842105	0 *	
3	0.45	0.41946872	0.16842105	0 * G F	
4	0.16666667	0.4925222	0.16842105	0 G C E F	
5	0.45454545	0.50693366	0.16842105	0 G C E F	
6	0	0.45830593	0.16842105	0 A C D E F	
7	0.05	0.41946872	0.16842105	0 A C D E F	
8	0	0.43304853	0.16842105	0 A C D E F	
9	0	0.47980736	0.16842105	0 A C D E F	
10	0	0.45830593	0.16842105	0 A C D E F	
11	0	0.41341849	0.16842105	0 A C D E F	
12	0	0.43304853	0.16842105	0 A C D E F	
13	0	0.46848041	0.16842105	0 A C D E F	
14	0	0.4925222	0.16842105	0 A C D E F	
15	0.16216216	0.35299497	0.16842105	0 A C D E F	
16	0.4	0.41946872	0.16842105	0 C D E F	
17	0.79166667	0.39759517	0.16842105	0 * D E F	
18	0.15384615	0.47980736	0.16842105	0 D E F	
19	0	0.46848041	0.16842105	0 D E F	
20	0	0.39759517	0.16842105	0 D E F	
21	0.125	0.44910088	0.16842105	0 D E F	
22	0	0.47980736	0.16842105	0 D E F	

[부록 7] 커넥팅로드 특정 날짜, 기간 조회

```
import pandas as pd

DateMetric = 'dt_workday' #@param {type:"string"}
SearchDate = analysisData[DateMetric]

SearchDate.head(10)

0    2021-09-01
1    2021-09-01
2    2021-09-01
3    2021-09-01
4    2021-09-01
5    2021-09-01
6    2021-09-01
7    2021-09-01
8    2021-09-01
9    2021-09-01
Name: dt_workday, dtype: object

target_date1 = '2021-09-03' #@param {type:"string"}
target_date2 = '2021-09-05' #@param {type:"string"}
analysisData[(analysisData[dt_workday] >= target_date1) & (analysisData[dt_workday] <= target_date2)].head()

target_date1 = 2021-09-03
target_date2 = 2021-09-05

MODEL_NAME DT_MEASURE dt_workday red_dial_dia red_dial_cirly red_dial_perp red_dial_cycty red_dia2_v red_dia2_h red_dia2_dia red_dia2_dist red_bhp1_y red_bhp1_z red_bhp1_dia red_bhp1_z red_bhp2 red_bhp2_z
37    265    2021-09-03    2021-09-03    50.0153    0.0017    0.0044    0.0022    67.4351    210.6135    87.0826    274.9365    12.4424    -21.4142    14.6199    -21.3971
38    265    2021-09-03    2021-09-03    50.0157    0.0036    0.0046    0.0039    67.4337    210.6140    87.0825    274.9370    12.8882    -22.9908    13.7446    -22.0858
39    265    2021-09-03    2021-09-03    50.0089    0.0030    0.0043    0.0034    67.4368    210.6133    87.0758    274.9361    12.4482    -21.4107    14.6208    -21.3949
40    265    2021-09-03    2021-09-03    50.0089    0.0027    0.0032    0.0066    67.4810    210.6373    87.0690    274.9673    12.4216    -21.4404    14.6232    -21.4393
```

## [부록 8] 공구수명 분석을 위한 데이터 셋 불러오기

```
TBL_FANUC_DOOSAN_SPINDLE = 'https://raw.githubusercontent.com/looker2zip/bigStudy/main/TBL_FANUC_DOOSAN_SPINDLE.csv'
TBL_FANUC_DOOSAN_SPINDLE_csv = pd.read_csv(TBL_FANUC_DOOSAN_SPINDLE)
TBL_FANUC_DOOSAN_SPINDLE_csv.head(10)
```

	spindle_id	fanuc_doosan_id	dt_insert	spindle_tool_no	max_rpm	max_feed	avg_tool_load	max_tool_load	avg_x_torque	max_x_torque	avg_y_torque	max_y_torque	avg_z_torque
0	7076	1009	2021-09-02 10:02:29	3	578	103923	11	121	8275	80	15299	80	12054
1	7077	1009	2021-09-02 10:03:06	8	433	90329	41	181	3637	80	6725	80	5302
2	7078	1009	2021-09-02 10:03:31	17	1779	103923	3	91	5521	80	10209	80	8048
3	7079	1009	2021-09-02 10:04:28	23	3003	103923	1	29	2380	80	4400	80	3469
4	7080	1009	2021-09-02 10:05:48	21	1148	92735	1	23	1688	80	3121	80	2462

## [부록 9] 공구 23번에 대한 조회

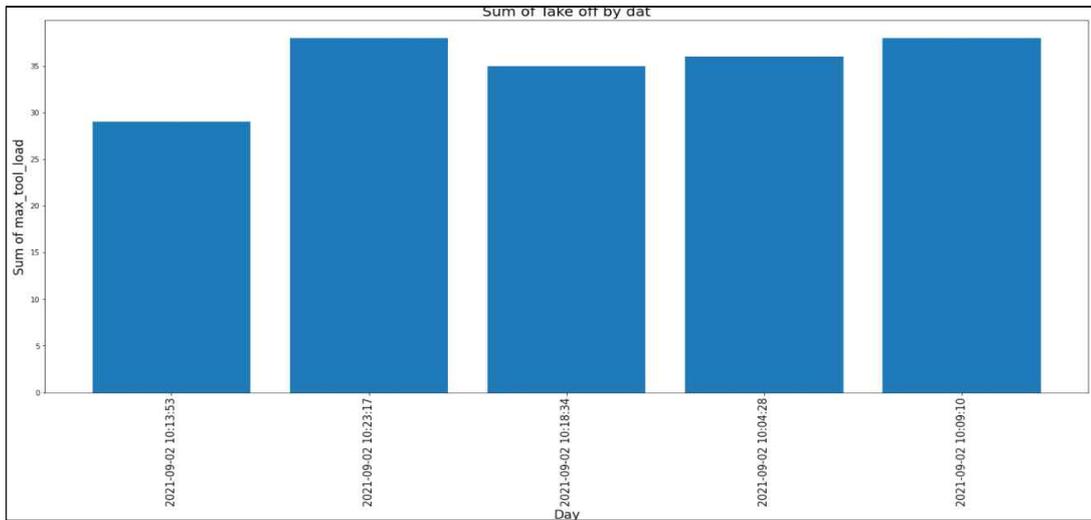
```
## 특정 공구번호
spindle_tool_no_search = '23' #@param {type:"string"}
TBL_FANUC_DOOSAN_SPINDLE_tool=TBL_FANUC_DOOSAN_SPINDLE_csv[TBL_FANUC_DOOSAN_SPINDLE_csv['spindle_tool_no']==spindle_tool_no_search]
TBL_FANUC_DOOSAN_SPINDLE_tool
```

	spindle_id	fanuc_doosan_id	dt_insert	spindle_tool_no	max_rpm	max_feed	avg_tool_load	max_tool_load	avg_x_torque	max_x_torque	avg_y_torque	max_y_torque	avg_z_torque
3	7079	1009	2021-09-02 10:04:28	23	3003	103923	1	29	2380	80	4400	80	3469
31	7107	1015	2021-09-02 10:09:10	23	3013	103923	1	38	2427	80	4486	80	4486
56	7132	1020	2021-09-02 10:13:53	23	3007	103923	1	35	2426	80	4483	80	4483
81	7157	1025	2021-09-02 10:18:34	23	3007	103923	1	36	2449	80	4526	80	4526

## [부록 10] 공구수명 분석을 위한 데이터 셋 불러오기

TBL_FANUC_DOOSAN_SPINDLE_csv.info()					TBL_FANUC_DOOSAN_SPINDLE_csv.info()				
<class 'pandas.core.frame.DataFrame'>					<class 'pandas.core.frame.DataFrame'>				
RangeIndex: 79574 entries, 0 to 79573					RangeIndex: 79574 entries, 0 to 79573				
Data columns (total 16 columns):					Data columns (total 16 columns):				
#	Column	Non-Null	Count	Dtype	#	Column	Non-Null	Count	Dtype
0	spindle_id	79574	non-null	int64	0	spindle_id	79574	non-null	int64
1	fanuc_doosan_id	79574	non-null	int64	1	fanuc_doosan_id	79574	non-null	int64
2	dt_insert	79574	non-null	object	2	dt_insert	79574	non-null	object
3	spindle_tool_no	79574	non-null	int64	3	spindle_tool_no	79574	non-null	int64
4	max_rpm	79574	non-null	int64	4	max_rpm	79574	non-null	int64
5	max_feed	79574	non-null	int64	5	max_feed	79574	non-null	int64
6	avg_tool_load	79574	non-null	int64	6	avg_tool_load	79574	non-null	int64
7	max_tool_load	79574	non-null	int64	7	max_tool_load	79574	non-null	int64
8	avg_x_torque	79574	non-null	int64	8	avg_x_torque	79574	non-null	int64
9	max_x_torque	79574	non-null	int64	9	max_x_torque	79574	non-null	int64
10	avg_y_torque	79574	non-null	int64	10	avg_y_torque	79574	non-null	int64
11	max_y_torque	79574	non-null	int64	11	max_y_torque	79574	non-null	int64
12	avg_z_torque	79574	non-null	int64	12	avg_z_torque	79574	non-null	int64
13	max_z_torque	79574	non-null	int64	13	max_z_torque	79574	non-null	int64
14	tool_set_count	79574	non-null	int64	14	tool_set_count	79574	non-null	int64
15	tool_current_count	79574	non-null	int64	15	tool_current_count	79574	non-null	int64
dtypes: int64(15), object(1)					dtypes: int64(14), object(2)				
memory usage: 9.7+ MB					memory usage: 9.7+ MB				

[부록 11] 공구 23번에 대한 날짜별 부하(max\_tool\_load) 조회

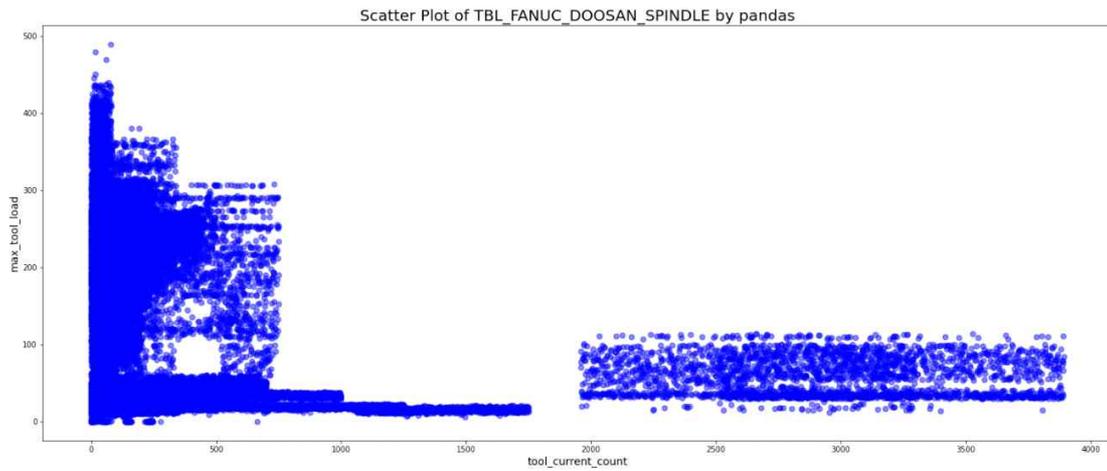


[부록 12] features, label 데이터 준비

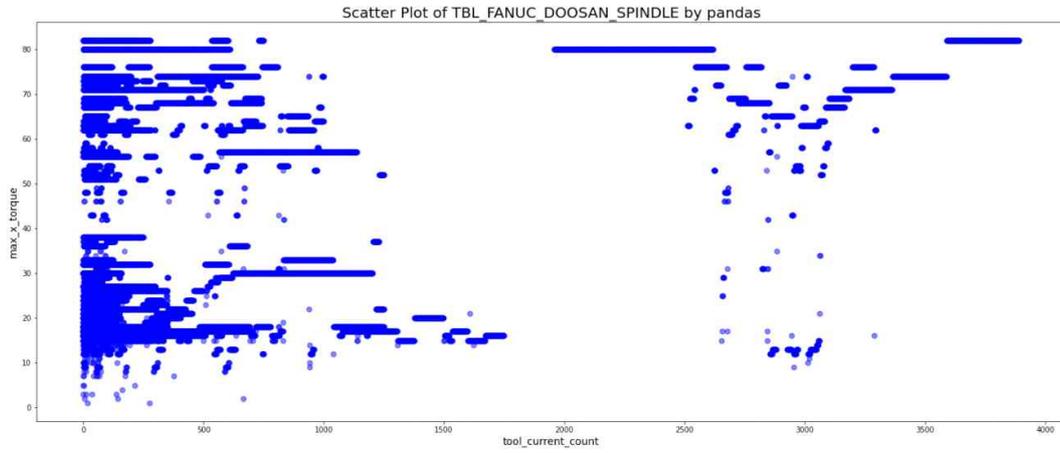
```
analysisData_TBL_FANUC_DOOSAN_SPINDLE=TBL_FANUC_DOOSAN_SPINDLE.csv.loc[:,['max_tool_load','max_x_torque','max_y_torque','max_z_torque','tool_current_count']]
analysisData_TBL_FANUC_DOOSAN_SPINDLE.head()
```

	max_tool_load	max_x_torque	max_y_torque	max_z_torque	tool_current_count
0	121	80	80	62	458
1	181	80	80	62	1
2	91	80	80	62	1958
3	29	80	80	62	175
4	23	80	80	62	16

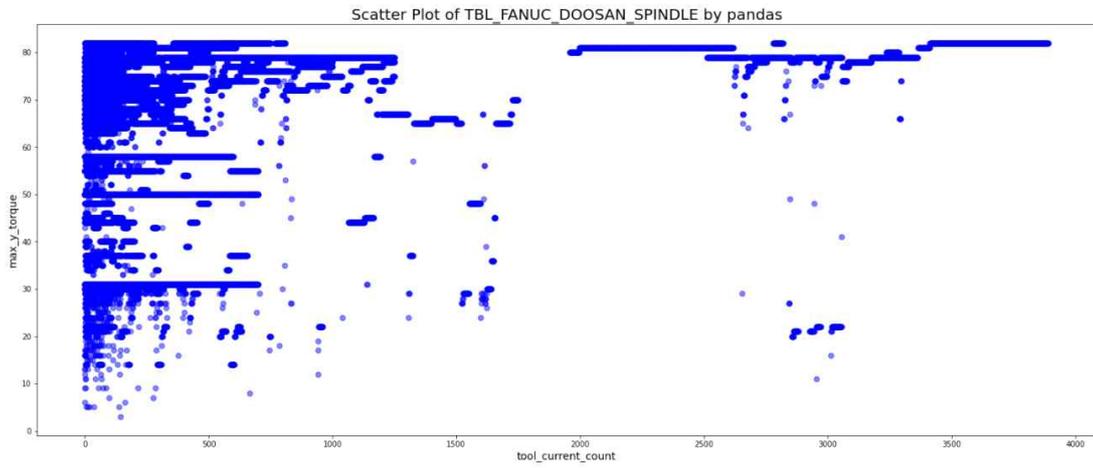
[부록 13] max\_tool\_load 가공조건별 교환주기 확인



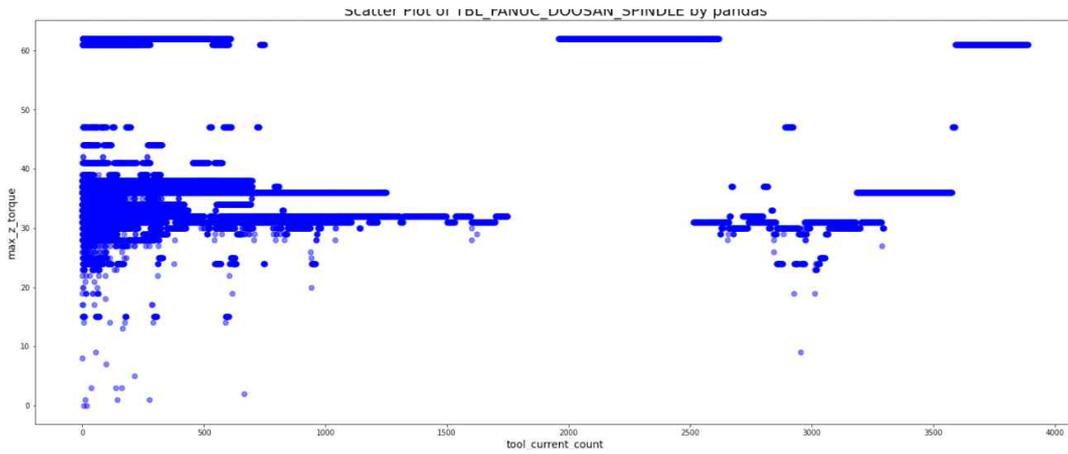
[부록 14] max\_x\_torque 가공조건별 교환주기 확인



[부록 15] max\_y\_torque 가공조건별 교환주기



[부록 16] max\_z\_torque 가공조건별 교환주기

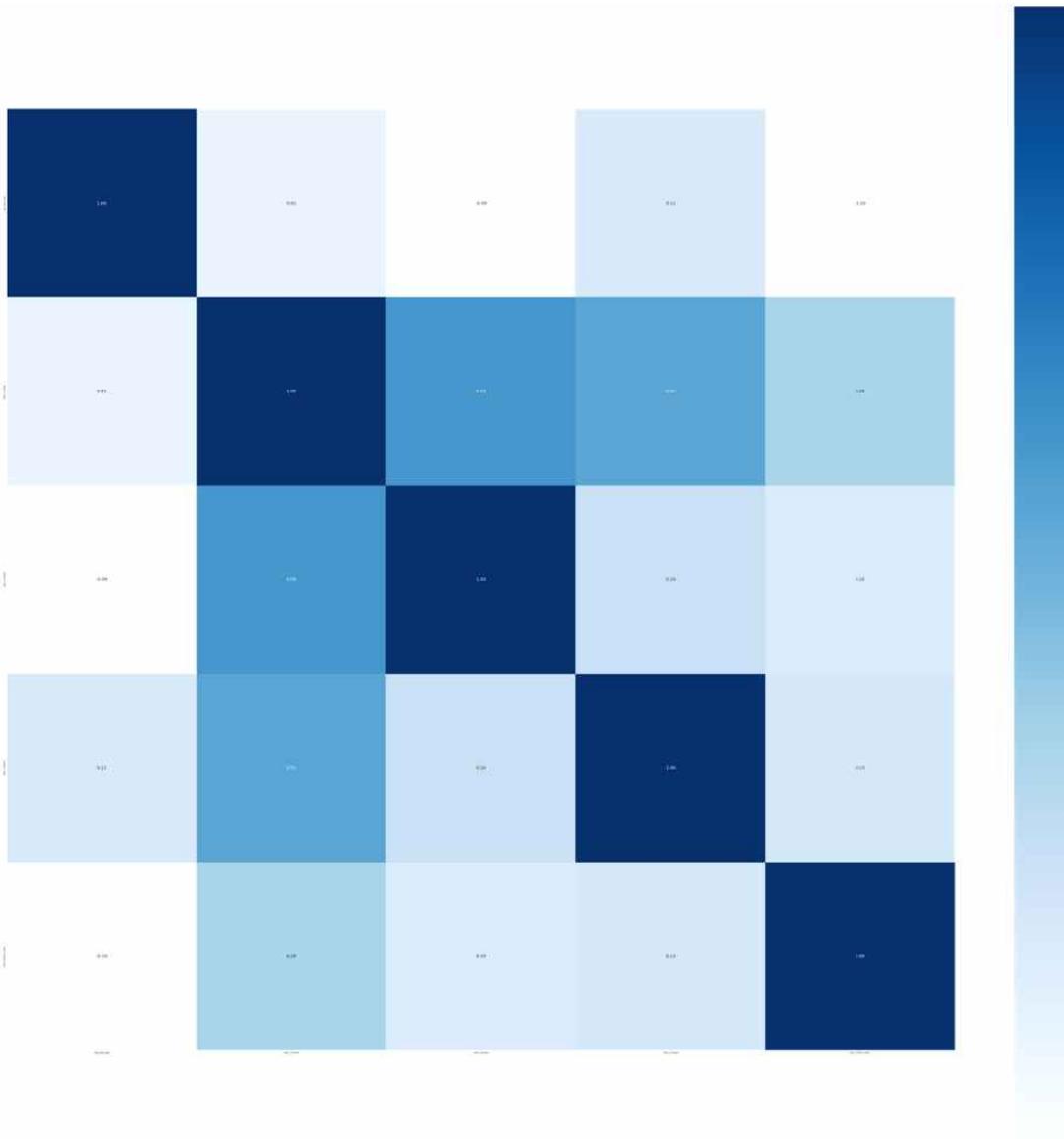


[부록 17] 각 특징변수의 상관계수 확인

```
analysisData_TBL_FANUC_DOOSAN_SPINDLE_corr=analysisData_TBL_FANUC_DOOSAN_SPINDLE_corr(method='pearson')
analysisData_TBL_FANUC_DOOSAN_SPINDLE_corr
```

	max_tool_load	max_x_torque	max_y_torque	max_z_torque	tool_current_count
max_tool_load	1.000000	0.013765	-0.090810	0.106239	-0.097571
max_x_torque	0.013765	1.000000	0.581499	0.512505	0.284835
max_y_torque	-0.090810	0.581499	1.000000	0.188793	0.099540
max_z_torque	0.106239	0.512505	0.188793	1.000000	0.128354
tool_current_count	-0.097571	0.284835	0.099540	0.128354	1.000000

[부록 18] 각 특징변수의 히트맵



## ABSTRACT

### A study on Advancement of Smart Factory Based on Real-Time Big Data Collection and AI Modeling Analysis

Kim, Jae-Pyo

Major in Smart Convergence Product

Dept. of Smart Convergence Consulting

The Graduate School

Hansung University

Recently, the trend of Smart Factory is shifting to the advancement by using Big Data and AI technologies. Accordingly, preceding research have been done on Big Data collection and processing technology, IoT integration technology, Machine Learning and AI modeling. Based on this literature review, this research designed real-time manufacturing data collection, analysis and utilization system which is suitable to the scope and purpose of this study, and then designed and implemented Smart factory advancement of machine working industry.

First, for the smartization of machining process, a new process was designed and built by rearranging and improving the layout. New process was constructed by newly introducing, designing, manufacturing, or installing new jig, robot and automatic conveyor, MCT processing machine, bush pressure measurement and assembly system, and tool monitoring and

inspection system.

Second, in order to design a real-time collection and analysis system of manufacturing data, a network for real-time data collection was priorly designed and implemented with an IoT integration module based on OPC-UA, the latest communication protocol.

Third, the data software is designed and implemented in connection with the cloud by applying the application browser design of the 3-Tier Structure and Smart Client concept. The user's requirements for the new smartization process are reflected on the collection and analysis table, and based on this, the database and extraction solution were implemented.

Fourth, SPC statistical analysis and visualization solutions were designed and implemented. Major variable managed in machining process and main feature derived from AI model were cross-validated by SPC statistical analysis and AI modeling.

Fifth, AI modeling, validation and visualization analysis were designed and performed. Relative correlation analysis of variables and comparison of prediction accuracy between algorithms were held. F1 score results on the importance of major variables are cross-validated with SPC analysis results.

From the above, a database for real-time data collection and storage and extraction solution were built in MES, and based on this solution, the system is able to systematically accumulate and extract data through big data classification and storage. In the long term, this builds the base for productivity and quality improvement.

F1 score of the XGBoost features was evaluated to be consistent with the results of the SPC analysis. It is expected that post-maintenance will be possible by clarifying defective factors through the validation of SPC statistics and performing cross-validation of the prediction results of AI modeling. As for the prediction accuracy of the AI algorithm, it can be expected that the accuracy of the Random Forest algorithm will increase if

data gets accumulated and training is continuously and repeatedly performed afterward.

**【Keyword】** Bigdata, AI Modeling, Data Collection and Analysis System, IoT Integration, SPC, Advancement of Smart Factory, OPC-UA, MES, Cloud, Machine Work, Application Browser, Smart Client, AI Algorithm, Communication Protocol, Supervised Learning