프로젝트명: 작업자 안전 위험 빅데이터 / AI 분석 웹 서비스

과정: 부산대학교 산학협력단 K-디지털 트레이닝

「AI 활용 빅데이터 분석 풀스택 웹 서비스 개발자 양성과정」

프로젝트팀

이름	전화번호	email
구민지(팀장)	010-4842-3268	yeskuminji@naver.com
김단우	010-5514-8086	qnfrmds2003@gmail.com
양철민	010-5351-3485	ycm7784@naver.com

수정일: 23.5.29 version

작업자 안전 위험 빅데이터 / AI 분석 웹 서비스 : 스마트 밴드, 헬멧, 벨트 센서 기반의 위험 행동 인식

1. 프로젝트 개요

· 배경

건설 현장에서의 낙상 등 안전 사고를 미연에 방지하기 위해 이미 일어난 사고 데이터를 수집하여 사고가 발생할 가능성이 높은 상황을 예측하여 조치를 취할 수 있도록 AI 프로그램을 개발하고자 합니다.

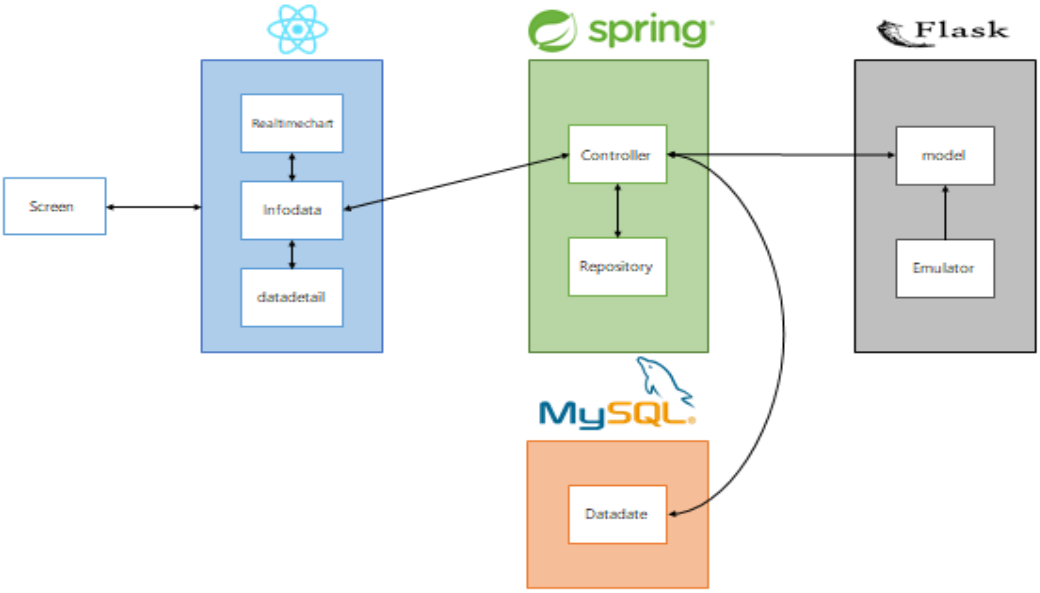
· 목표

- 스마트 밴드 또는 스마트 헬멧으로 데이터를 수집.
- 수집된 생체 데이터를 통한 작업자 상태 관리.
- 근로자의 Gyro sensor 데이터, 맥박, 체온을 통한 사고 특성 분석 및 예측.
- 위험 예측을 활용해서 안전관리자에게 위험관리자 위치 및 상태 알림.

· 팀 : WorkerSafety

구민지 (팀장)	Front-end(React)
김단우	Data Analysis(+Flask)
양철민	Back-end(SpringBoot)

- Project Structure



- 프로젝트 구현 환경

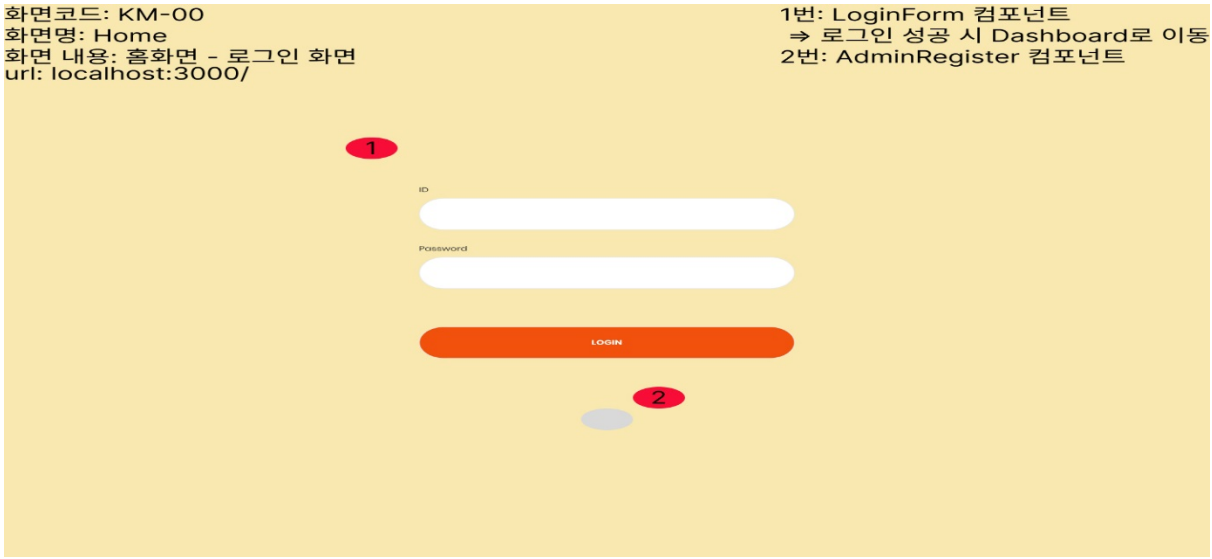
	언어 및 프레임워크	개발 환경	라이브러리, API
DA	Python : 3.10.10 Flask : 2.3.2	PyCharm	scikit-learn Matplotlib Pandas NumPy
BE	Springboot : 3.0.6 Java : 17	Eclipse	jpa
FE	Node.js : 18.13.0 JavaScript	VSCode	React : 18.2.0 react-router-dom : 6.11.0 axios : 1.4.0

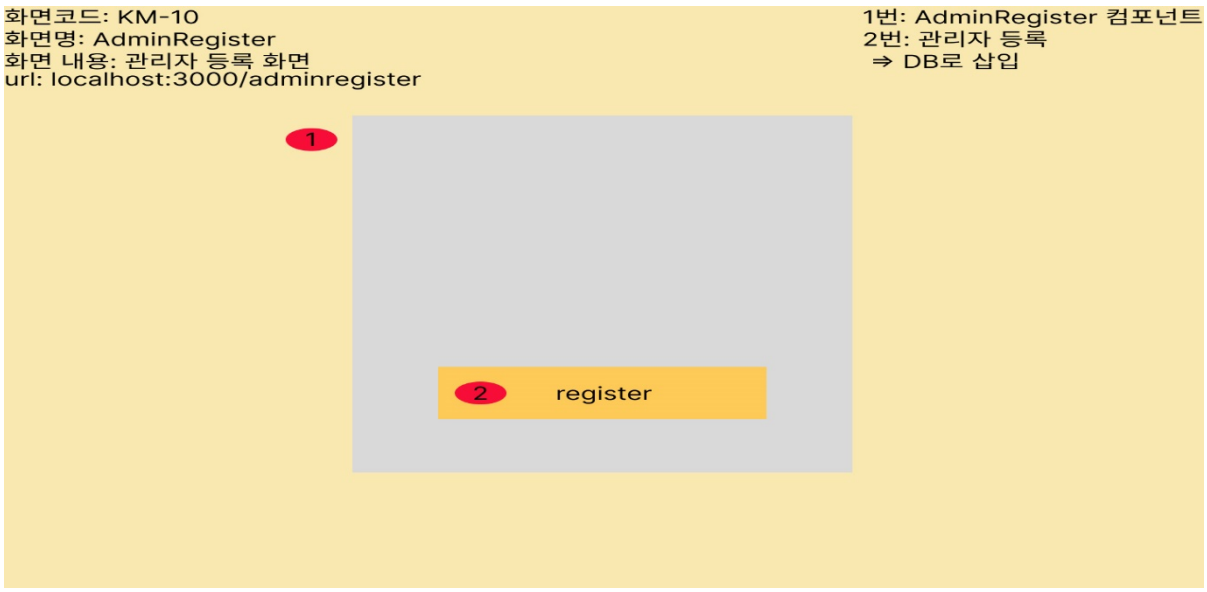
2. 요구 사항 명세

서비스	ID	요구사항명	요구사항 내용	날짜	버전
Data Analysis	ws-001	작업자 상태 분석 및 예측	작업자의 이상 맥박 감지	05.18	v0.5
	ws-002		작업자의 이상 체온 분석	05.18	v0.5
	ws-003		작업자의 이상 Gyro sensor 분석	05.18	v0.5
	ws-004		작업자 종합 데이터를 통한 사고 예측	05.18	v0.5
로그인	ws-005	로그인	ID, Password Form 제공	05.18	v0.5
지도창	ws-006	작업자 위치	GPS 데이터로 실시간 작업자 위치 표출	05.18	v0.5
	ws-007	작업자 상태	아이콘을 통해 위험상태 표출	05.18	v0.5
	ws-008	작업자 데이터	아이콘 선택 시 지도 아래에 상세 정보 표출	05.18	v0.5
상세창	ws-009	작업자 데이터	아이콘 선택 후 상세 보기 클릭	05.18	v0.5
목록창	ws-010	작업자 관리	작업자 목록 표출	05.18	v0.5
	ws-011		작업자 추가	05.18	v0.5
	ws-012		작업자 삭제	05.18	v0.5
	ws-013		작업자 검색	05.18	v0.5
	ws-014		작업자 정렬	05.18	v0.5
생체 데이터	ws-015	생체 데이터 표출	작업자 맥박 그래프 표출	05.18	v0.5
	ws-016		맥박 데이터 위험 범위 시 알림	05.18	v0.5
	ws-017		작업자 체온 그래프 표출	05.18	v0.5
	ws-018		체온 데이터 위험 범위 시 알림	05.18	v0.5
활동 데이터	ws-019	활동 데이터 표출	Gyro sensor 데이터 그래프로 표출	05.18	v0.5
	ws-020		Gyro sensor 데이터 위험 범위 시 알림	05.18	v0.5

3. 화면설계 및 기능명세서

화면코드	화면명	화면 내용	url	버전
KM-00	Home	홈화면: 로그인 화면	localhost:3000/	v0.5
KM-10	AdminRegister	관리자 등록 화면	localhost:3000/adminregister	v0.5
KM-20	Dashboard	관리자 화면	localhost:3000/dashboard	v0.5
KM-21	Dashboard	관리자 화면 - 아이콘 클릭 시	localhost:3000/dashboard	v0.5
KM-30	WorkerDetail	작업자 상세 화면	localhost:3000/workerdetail	v0.5

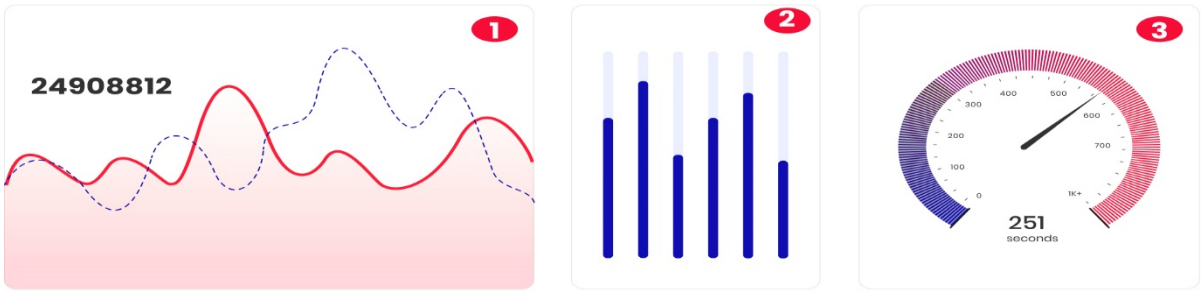




화면코드: KM-21
화면명: Dashboard
화면 내용: 관리자 화면 - 아이콘 클릭 시
url: localhost:3000/dashboard

1번: Data 컴포넌트
⇒ 작업자 생체 데이터, 활동 데이터 표출
2번: WorkerDetail 컴포넌트
⇒ 작업자 데이터 상세보기

아이콘을 누르면 표출되는 창



화면코드: KM-30
화면명: WorkerDetail
화면 내용: 작업자 상세 화면
url: localhost:3000/workerdetail

1번: 작업자 상세 데이터 (체온)
2번: 작업자 상세 데이터 (위험도)
3번: 작업자 상세 데이터
4번: 작업자 상세 데이터 (맥박)

4. RestAPI 설계

/login

Index	Method	URI	Description	Response(JSON)
1	POST	/login	로그인	로그인결과

/login

Index	Method	URI	Description	Response(JSON)
1	PUT	/logout	로그아웃	로그아웃 결과

/user

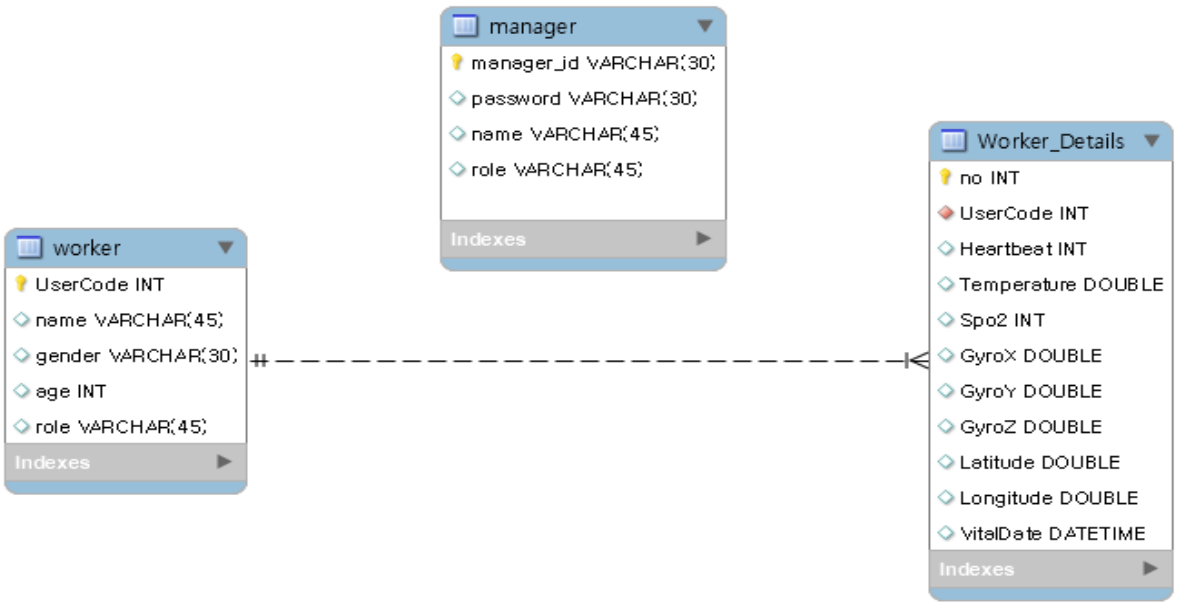
Index	Method	URI	Description	Response(JSON)
1	POST	/join	회원가입	회원가입 결과
2	DELETE	/delete	회원탈퇴	회원탈퇴 결과
3	PUT	/update	회원정보수정	회원정보수정 결과

/worker

Index	Method	URI	Description	Response(JSON)
1	GET	/List	작업자 조회	작업자 세부정보 결과
2	POST	/add	작업자 추가	작업자 추가결과 결과
3	DELETE	/delete	작업자 삭제	작업자 삭제결과 결과

5. DB 설계

- ERD



- DB 상세

Table	Type	Field	Remarks
manager	varchar	manager_id	Primary Key
	varchar	password	
	varchar	name	
	varchar	role	
worker	integer	UserCode	Primary Key
	varchar	name	
	varchar	gender	
	integer	age	
	varchar	role	
worker_Details	integer	no	Primary Key
	integer	UserCode	Foreign Key(worker)
	integer	Heartbeat	심장박도
	double	Temperature	체온
	integer	Spo2	산포도
	double	GyroX	자이로센서x축데이터

double	GyroY	자이로센서x축데이터
double	GyroZ	자이로센서z축데이터
double	Latitude	위도
double	Longitude	경도
datatime	VitalDate	시간

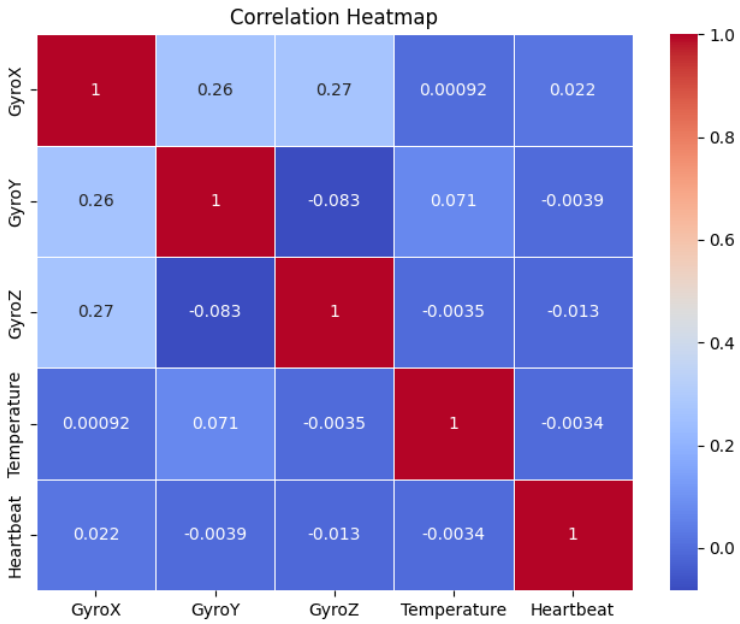
6. 데이터 분석

- 데이터 특징 : 시계열 데이터
- 데이터 출처
  - A. 강남앤인코누스 현장 수집 데이터
  - B. 외부 데이터
    - i. https://github.com/laxmimerit/Human-Activity-Recognition-Using-Accelerometer-Data-and-CNN
    - ii. https://sites.google.com/view/kfalldataset#h.6xeqx1x4k3qy
  - C. Dummy data 생성
    - i. 강남앤인코누스 현장 수집 데이터 : 체온, 맥박, 산소포화도
    - ii. https://github.com/laxmimerit/Human-Activity-Recognition-Using-Accelerometer-Data-and-CNN 데이터 GyroX, GyroZ, GyroY sensor 데이터
    - iii. 시간 데이터, UserCode 임의 생성

	UserCode	Label	GyroX	GyroY	GyroZ	Temperature	Heatbeat	Lat	Lng	No	time
0	8.0	Walking	5.41	13.21	-4.630918	36.8	94.0	35.235809	129.077699	1.0	00:00:00
1	8.0	Walking	6.74	8.16	0.953424	36.9	93.0	35.235969	129.077695	2.0	00:00:02
2	8.0	Walking	5.33	5.37	-2.792171	36.9	91.0	35.235999	129.077640	3.0	00:00:04
3	8.0	Walking	3.38	8.54	-1.525479	36.8	93.0	35.235998	129.077793	4.0	00:00:06
4	8.0	Walking	1.73	9.11	-0.994285	36.7	93.0	35.235855	129.077650	5.0	00:00:08

- 데이터 형태 : csv

- 데이터 내용 : 작업자별 Gyro sensor 데이터, 맥박, 체온, 위치, 시간
- 가설 설정
  A. 건설 현장에서는 근로자가 반복되는 작업을 하게 되는데, 이 과정에서 근로자의 자세나 움직임이 불규칙해지면 사고 발생 가능성이 높아질 것으로 가설을 설정합니다.
  B. 따라서, 근로자의 움직임 데이터와 체온, 맥박 변화를 사용하여 이상 감지 알고리즘을 구현하고, 일정 이상의 불규칙한 패턴이 발견될 경우 사고가 발생할 가능성이 높다는 가설을 설정합니다.
- 데이터 전처리 및 분석(탐색적 가시화)
  A. 수집된 데이터의 결측치를 제거합니다.
  B. 수집된 데이터에서 이상치를 탐지하고 처리합니다. 체온이 정상 범위를 벗어나는 경우 또는 맥박이 비정상적으로 높거나 낮은 경우, 급격한 변화가 생기는 경우를 탐지합니다.
  C. 데이터를 정규화를 수행하여 데이터를 가공합니다.
  D. 통계 분석: 통계 분석은 데이터 간의 관계를 이해하고, 특정 변수들의 패턴을 탐지하기 위해 사용되는 분석 방법입니다. 작업자 안전 예측을 위해 체온, 맥박, 그리고 Gyro sensor 데이터의 통계 분석을 수행할 수 있습니다. 특히 체온과 맥박의 상관관계를 확인하고, Gyro sensor 데이터의 표준편차를 계산하는 것이 중요



합니다.

i. 상관관계 분석:

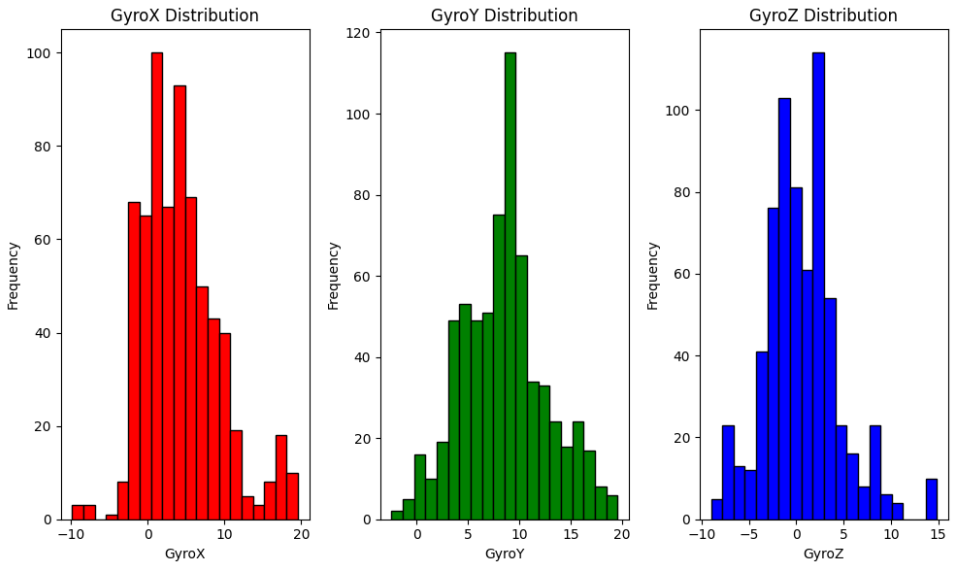
체온과 맥박 간의 상관관계를 확인하기 위해 상관계수를 계산합니다. 상관계수는 두 변수 간의 선형적인 관계의 강도와 방향을 나타내는 지표입니다. 상관계수는 -1부터 1까지의 범위를 가지며, 0에 가까울수록 두 변수 간의 관계가 약하고, 1 또는 -1에 가까울수록 강한 양의 또는 음의 상관관계를 나타냅니다.

```
Correlation between Temperature and Heartbeat: -0.0033593759045805927
```

ii. Gyro sensor 데이터의 표준편차 계산:

Gyro sensor 데이터의 표준편차를 계산하여 변동성을 확인합니다. 표준편차는 데이터의 분산 정도를 나타내는 지표로, 값이 크면 데이터가 평균에서 멀리 퍼져있음을 의미합니다. 따라서 Gyro sensor 데이터의 표준편차를 계산하여 작업자의 움직임의 변동성을 파악할 수 있습니다.

```
Standard Deviation (GyroX): 5.008231021012116
Standard Deviation (GyroY): 4.113663794972511
Standard Deviation (GyroZ): 3.9963879512903473
```



iii. 결론:

체온과 맥박의 상관관계 계수는 -0.0034로 매우 낮은 값을 가지고 있습니다. 이는 체온과 맥박 간에 선형적인 상관관계가 거의 없다는 것을 의미합니다.

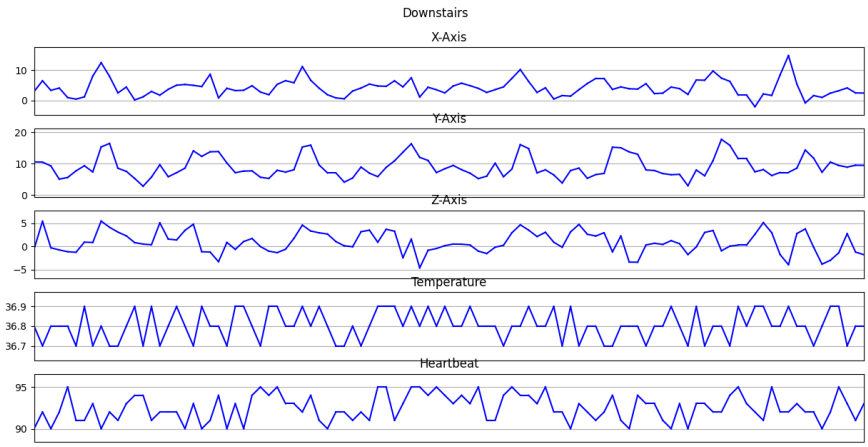
Gyro 센서 데이터의 표준편차는 각각 5.0082 (GyroX), 4.1137 (GyroY), 3.9964 (GyroZ)로 계산되었습니다. 표준편차는 데이터의 변동성을 나타내며, 값이 클수록 데이터가 평균값으로부터 더 많이 퍼져있음을 의미합니다. 따라서 GyroX 센서 데이터의 변동성이 가장 크고, GyroZ 센서 데이터의 변동성이 가장 작다고 볼 수 있습니다.

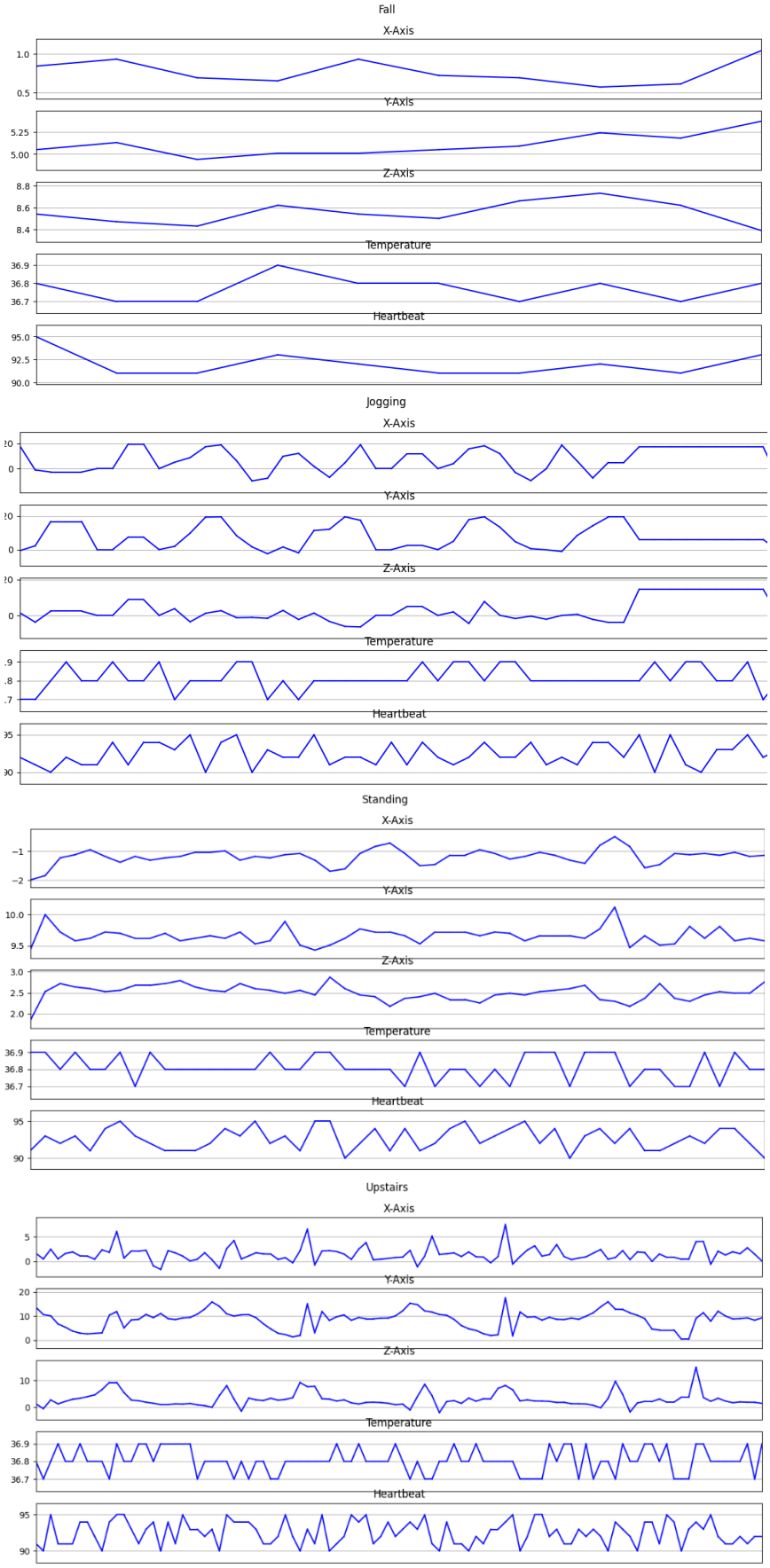
이러한 결과는 체온과 맥박 사이에는 강한 선형적인 관계가 없으며, Gyro 센서 데이터는 각각 다른 변동성을 가지고 있다는 것을 나타냅니다.

E. 시각화: 전처리된 데이터를 그래프로 시각화하여 패턴이나 상관관계를 확인합니다. 체온, 맥박 및 Gyro sensor를 각각 시계열 그래프로 표현하여 변화를 살펴봅니다. (시각화 데이터: UserCode_8)

i. 시계열 그래프:

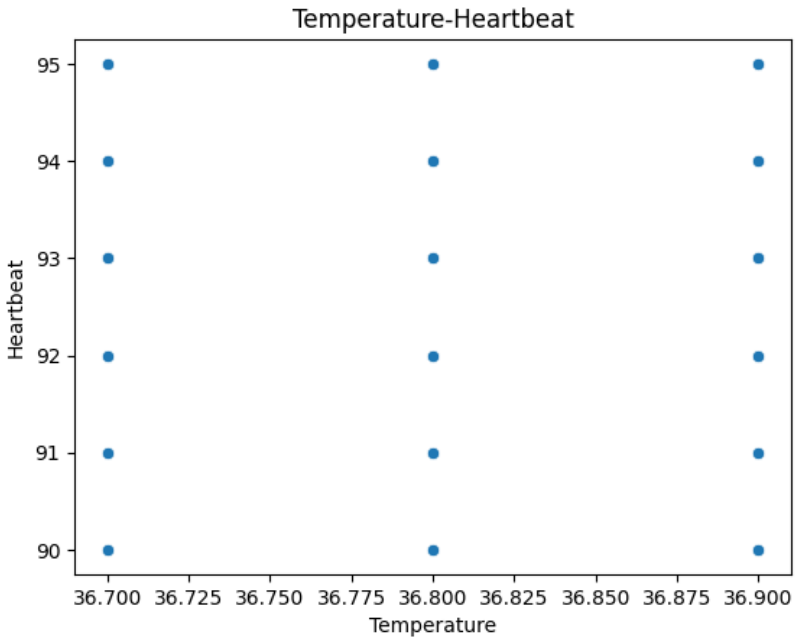
- 체온 변화 시각화: x축에는 시간, y축에는 체온 값을 나타내는 선 그래프를 사용하여 체온의 시간적 변화를 확인합니다. 이 그래프는 일정 시간 동안 체온이 어떻게 변하는지를 보여줍니다.
- 맥박 변화 시각화: x축에는 시간, y축에는 맥박 값을 나타내는 선 그래프를 사용하여 맥박의 시간적 변화를 확인합니다. 이 그래프는 일정 시간 동안 맥박이 어떻게 변하는지를 보여줍니다.
- Gyro sensor 시각화: x축에는 시간, y축에는 Gyro sensor 값을 나타내는 선 그래프를 사용하여 Gyro sensor의 시간적 변화를 확인합니다. 이 그래프는 일정 시간 동안 사람의 움직임이 어떻게 변하는지를 보여줍니다.




Fall
X-Axis
1.0
0.5
Y-Axis
5.25
5.00
Z-Axis
8.8
8.6
8.4
Temperature
36.9
36.8
36.7
Heartbeat
95.0
92.5
90.0
Jogging
X-Axis
Y-Axis
Z-Axis
Temperature
Heartbeat
Standing
X-Axis
−1
−2
Y-Axis
10.0
9.5
Z-Axis
3.0
2.5
2.0
Temperature
36.9
36.8
36.7
Heartbeat
95
90
Upstairs
X-Axis
5
0
Y-Axis
20
10
0
Z-Axis
10
0
Temperature
36.9
36.8
36.7
Heartbeat
95
90

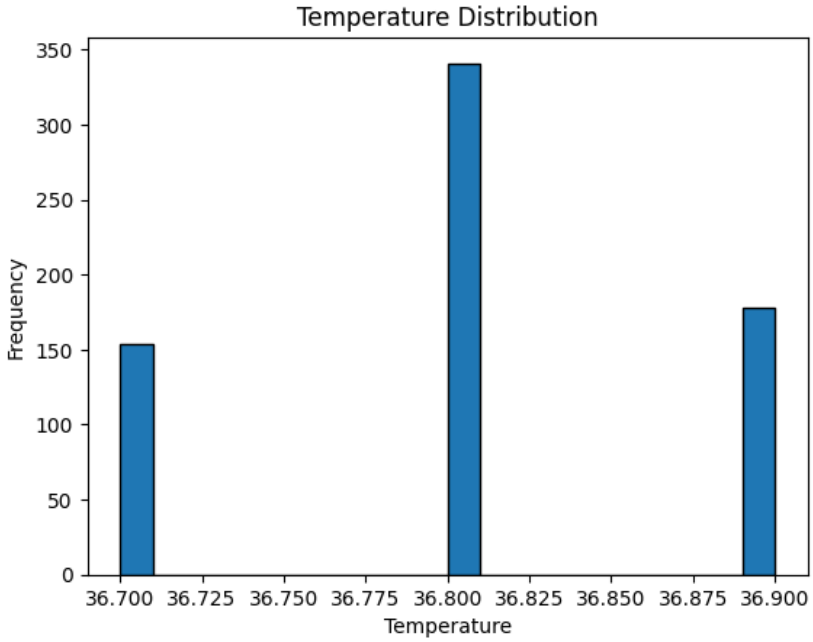
ii. 상관관계 그래프:

- 체온과 맥박 상관관계: 체온을 x축으로, 맥박을 y축으로 한 산점도 그래프를 그려 체온과 맥박 간의 상관관계를 확인합니다. 이 그래프는 체온과 맥박이 서로 어떤 관계를 가지고 있는지를 보여줍니다.

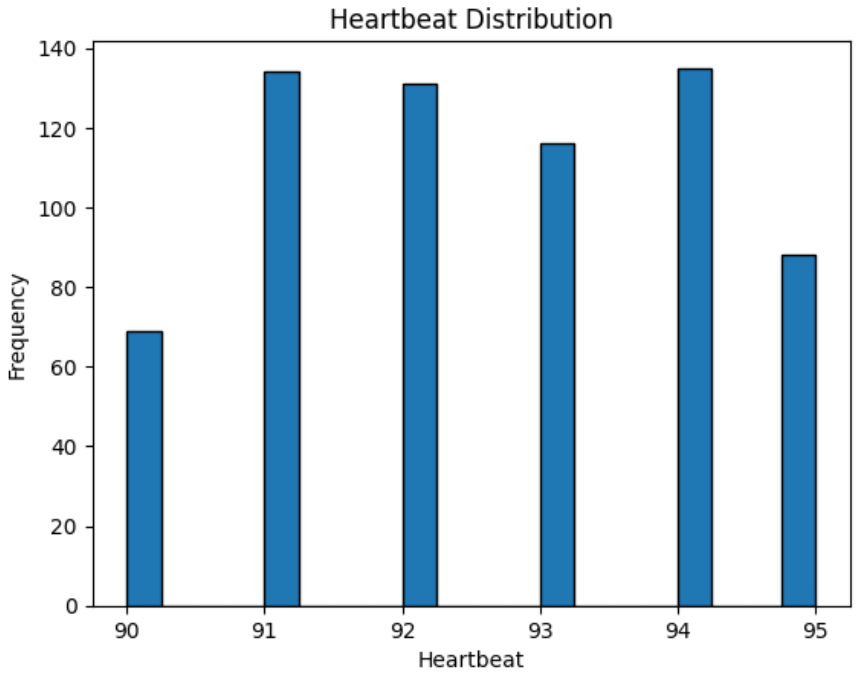


iii. 분포 그래프:

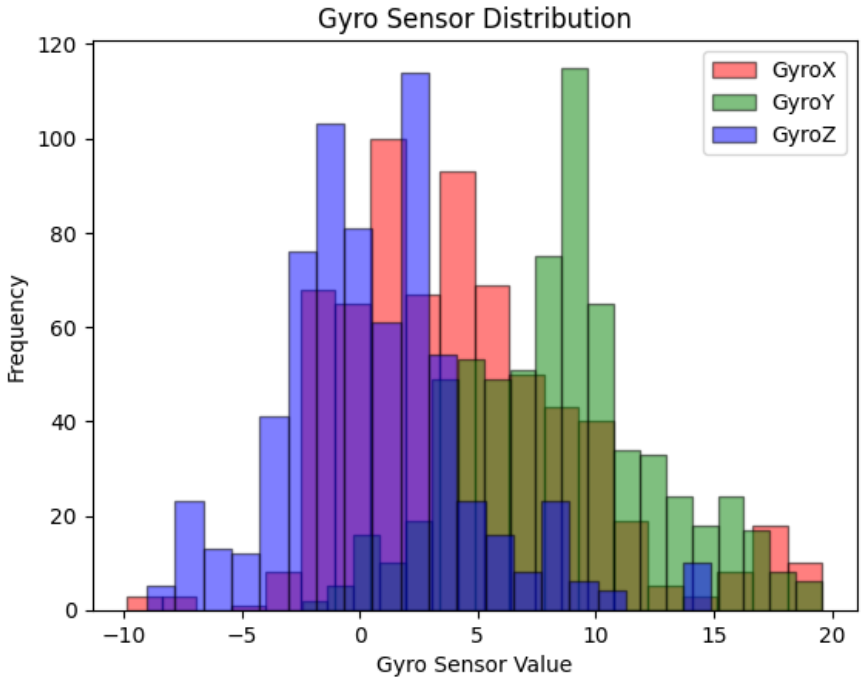
- 체온 분포: 체온 값을 x축으로 한 히스토그램을 그려 체온 분포를 확인합니다. 이 그래프는 체온 값의 분포가 어떻게 되는지를 시각적으로 보여줍니다



- 맥박 분포: 맥박 값을 x축으로 한 히스토그램을 그려 맥박 분포를 확인합니다. 이 그래프는 맥박 값의 분포가 어떻게 되는지를 시각적으로 보여줍니다.



- Gyro sensor 분포: Gyro sensor 값을 x축으로 한 히스토그램을 그려 Gyro sensor 값의 분포를 확인합니다. 이 그래프는 Gyro sensor 값의 분포가 어떻게 되는지를 시각적으로 보여줍니다.



- 모델링
  - 딥러닝 CNN(Conv1D)-LSTM(Long Short-Term Memory)

    CNN(Conv1D) 모델을 구축합니다. 입력 데이터로 맥박과 체온 데이터를 사용합니다. Conv1D 레이어를 사용하여 1차원 컨볼루션 연산을 통해 데이터의 지역적인 패턴을 추출합니다. 활성화 함수를 적용하여 비선형성을 추가하고,

    풀링 레이어를 사용하여 다운샘플링을 수행하고 중요한 정보를 유지합니다. CNN(Conv1D) 모델의 출력을 입력으로 사용하여 LSTM 모델을 구축합니다. LSTM 레이어를 사용하여 시계열 데이터의 패턴을 학습합니다. LSTM은 과거의 정보를 기억하고, 장기적인 의존성을 모델링할 수 있는 장점을 가지고 있어 시퀀스 데이터를 처리하는 데 적합합니다.

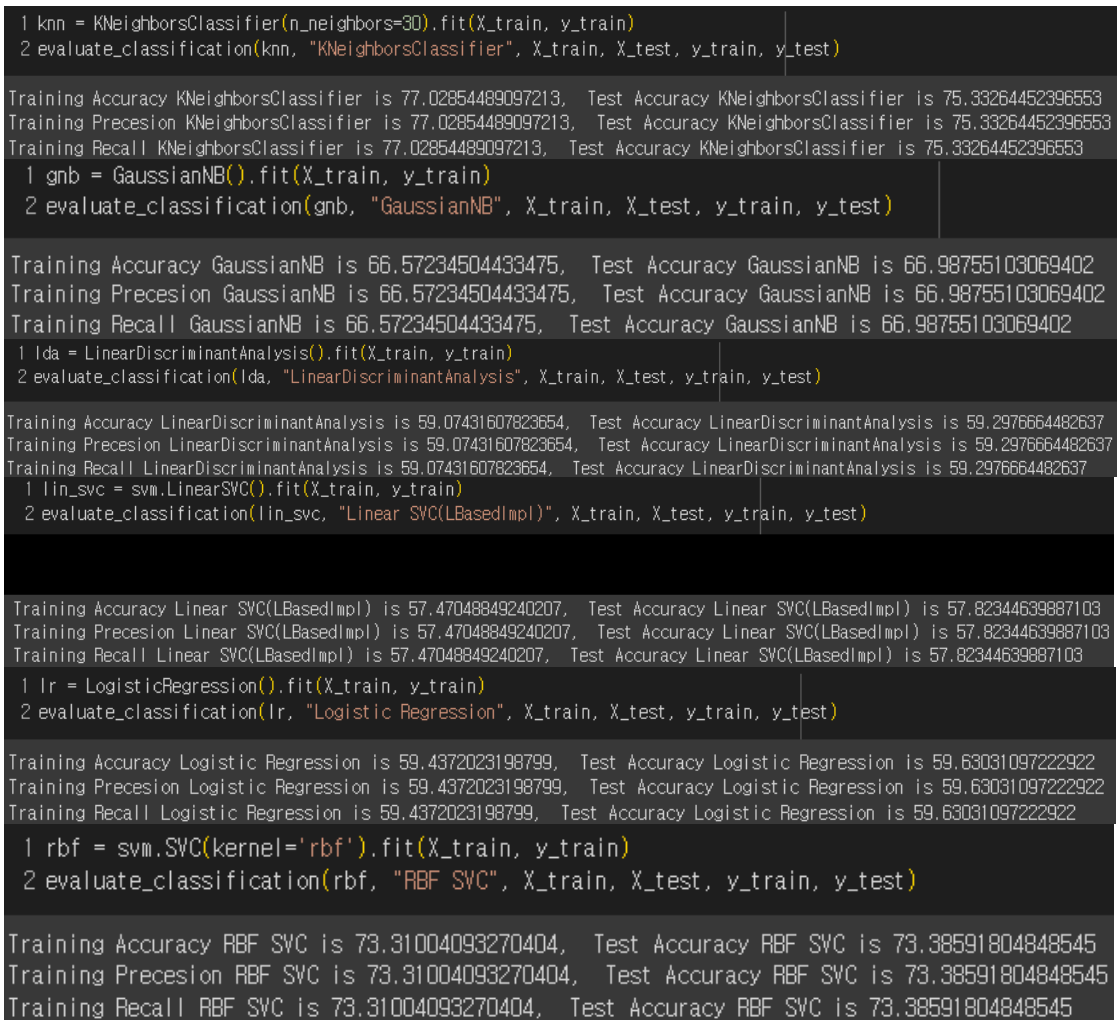
```
Epoch 100/100
17/17 [==============================] - 0s 6ms/step - loss: 0.4769 - accuracy: 0.8086 - val_loss: 0.5205 - val_accuracy: 0.8222
<keras.callbacks.History at 0x7faa9feaa320>
```

훈련 데이터에 대한 손실과 정확도는 각각 0.4769와 0.8086이며, 검증 데이터에 대한 손실과 정확도는 각각 0.5205와 0.8222입니다. 모델은 훈련 데이터에 대해서는 높은 정확도를 보이고 있으며, 검증 데이터에서도 일정 수준의 정확도를 유지하고 있습니다.

검증 데이터(validation data)에 대한 손실(loss)은 모델이 훈련 데이터와 검증 데이터 모두에서 얼마나 일반화되었는지를 나타냅니다. 이 경우 검증 데이터의 손실 값이 높습니다. 모델이 과적합(overfitting)되었을 수 있다는 것을 의미합니다.

- 머신러닝 모델

머신러닝을 사용하여 작업자의 안전 예측 모델들을 구축합니다. 맥박, 체온, Gyro sensor 데이터를 입력으로 사용하여 머신러닝 모델을 생성합니다.



KNeighborsClassifier이 train 77.0%, test 75.3%로 가장 높은 성능을 보입니다. 머신러닝 모델 사용시 KNeighborsClassifier 모델을 사용하여 하이퍼파라미터를 이용하여 모델을 개선합니다.

• 모델 평가 및 보완

모델의 성능을 평가하기 위해 근로자의 실제 사고 데이터, 활동정보와 모델이 예측한

사고 데이터, 활동 정보를 비교하여 모델의 정확도를 평가하고, 모델의 성능을 향상시키기 위해 추가적인 데이터나 특성을 수집하여 모델을 보완합니다.

7. 일정 계획

	1주	2주	3주	4주	5주	6주	7주	8주	9주	10주
DA	D1	D2		D3		D4	D5	D6	D7	
BE	B1	B2		B3		B4	B5	B6	B7	
FE	F1	F2		F3		F4	F5	F6	F7	

- DA
  - D1 : 계획
  - D2 : 데이터 전처리 & 머신러닝 기반 이상 감지 알고리즘 구현
  - D3 : 모델 개선
  - D4 : 플라스크를 활용한 API 작성
  - D5 : 신경망 기반 이상 감지 알고리즘 구현 (딥러닝, 텐서플로)
  - D6 : 모델 및 주요 서비스 개선
  - D7 : 테스트, 보완 및 발표

- BE
  - B1 : 계획
  - B2 : 모델 호출해서 REST API 생성
  - B3 : 스키마 기타 API 확정
  - B4 : 1차 구현, 개발 환경 구축, 로그인 시큐리티, 관리자 테이블, 각 서버 데이터 연결
  - B5 : 테스트 및 피드백
  - B6 : 2차 구현- 작업자 테이블, 각 서버의 요구·요청에 따른 API 작성, 데이터 저장

- B7 : 테스트, 보완 및 발표

- FE
  - F1 : 계획
  - F2 : UI/UX 설계 및 시각화
  - F3 : UI/UX 확정
  - F4 : 1차 구현 - 컴포넌트별 구현, 디자인 레이아웃 구현
  - F5 : 테스트 및 피드백
  - F6 : 2차 구현 - BE와 연결, 페이지 구현 및 완성
  - F7 : 테스트, 보완 및 발표