



Mini-project
수집한 데이터를 통한 머신러닝 알고리즘 적용

2024254012 배인호

데이터 설명 – 배전반(36895개)

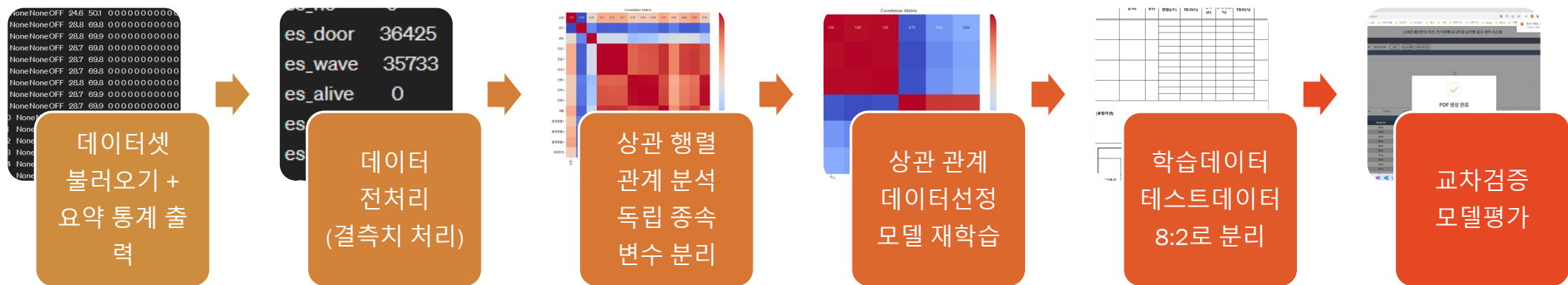
배전반 신뢰성 시험시 데이터

| es_no | es_door | es_wave | es_alive | es_temp | es_hum | es_ep1 | es_ep2 | es_ep3 | es_ec1 | es_ec2 | es_ec3 | es_pw_f | es_thd1 | es_thd2 | es_thd3 | es_max |
|-------|----------|-----------|----------|---------|--------|--------|--------|--------|--------|--------|--------|---------|---------|---------|---------|--------|
| 29517 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 21 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8019 |
| 29518 | None | None | R,S | 19.4 | 58.4 | 22 | 13 | 12 | 217 | 214 | 217 | -1 | 4 | 2 | 4 | 7811 |
| 29519 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 21 | 216 | 214 | 216 | 1 | 4 | 2 | 4 | 8023 |
| 29520 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 21 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8025 |
| 29521 | None | None | R,S,T | 19.4 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8053 |
| 29522 | None | None | R,S,T | 19.4 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8046 |
| 29523 | None | S Wave 20 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8009 |
| 29524 | None | S Wave 20 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8036 |
| 29525 | None | S Wave 20 | R,S,T | 19.3 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8034 |
| 29526 | None | None | R,S,T | 19.4 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8038 |
| 29527 | None | None | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8022 |
| 29528 | None | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8050 |
| 29529 | None | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8053 |
| 29530 | None | P wave 5~ | R,S,T | 19.4 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8056 |
| 29531 | B | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8059 |
| 29532 | B | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8068 |
| 29533 | B | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8074 |
| 29534 | B.Sensor | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8091 |
| 29535 | B.Sensor | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 218 | 215 | 217 | 1 | 4 | 2 | 4 | 8088 |
| 29536 | B.Sensor | P wave 5~ | R,S,T | 19.4 | 58.3 | 22 | 21 | 22 | 218 | 214 | 217 | 1 | 12 | 4 | 4 | 8080 |
| 29537 | B.Sensor | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 218 | 214 | 217 | 1 | 12 | 4 | 4 | 8080 |
| 29538 | B | P wave 5~ | R,S,T | 19.4 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8046 |
| 29539 | B | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8055 |
| 29540 | B | P wave 5~ | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8034 |
| 29541 | B | P wave 5~ | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8020 |
| 29542 | B | P wave 5~ | R,S,T | 19.4 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8039 |
| 29543 | B | P wave 5~ | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8005 |
| 29544 | B | P wave 5~ | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8024 |
| 29545 | B.Sensor | P wave 5~ | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8024 |
| 29546 | B.Sensor | P wave 5~ | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8013 |
| 29547 | B | P wave 5~ | R,S,T | 19.4 | 58.2 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8028 |
| 29548 | B | None | R,S,T | 19.4 | 58.4 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8024 |
| 29549 | None | None | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8024 |
| 29550 | None | None | R,S,T | 19.4 | 58.2 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8029 |
| 29551 | None | None | R,S,T | 19.3 | 58.2 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8036 |

수집한 데이터 설명

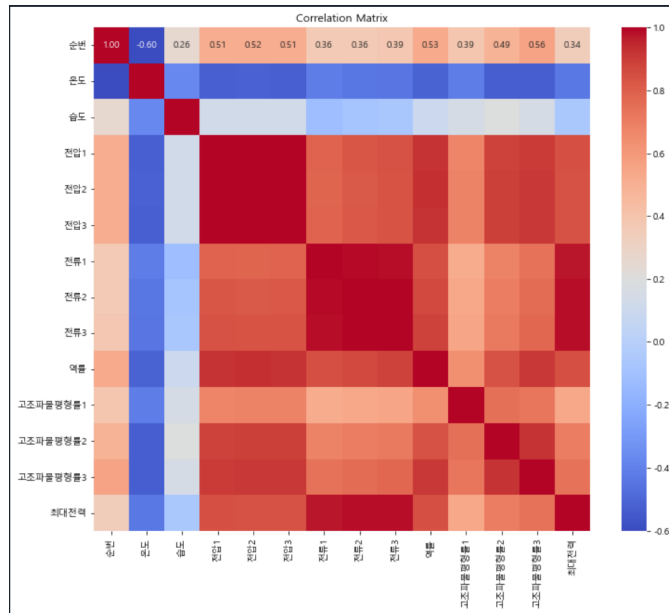
| 순번 | 도어열림 | 지진가속도센서값 | 활선상태 | 온도 | 습도 | 전압1 | 전압2 | 전압3 | 전류1 | 전류2 | 전류3 | 역률 | 고조파불평형률1 | 고조파불평형률2 | 고조파불평형률3 | 최대전력 |
|-------|----------|-------------|-------|------|------|-----|-----|-----|-----|-----|-----|----|----------|----------|----------|------|
| 29513 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 4 | 4 | 8066 |
| 29514 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 4 | 4 | 8047 |
| 29515 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 4 | 4 | 8031 |
| 29516 | None | None | R,S,T | 19.3 | 58.3 | 17 | 21 | 17 | 216 | 214 | 216 | 1 | 4 | 2 | 4 | 6920 |
| 29517 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 21 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8019 |
| 29518 | None | None | R,S | 19.4 | 58.4 | 22 | 13 | 12 | 217 | 214 | 217 | -1 | 4 | 2 | 4 | 7811 |
| 29519 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 21 | 216 | 214 | 216 | 1 | 4 | 2 | 4 | 8023 |
| 29520 | None | None | R,S,T | 19.3 | 58.3 | 22 | 21 | 21 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8025 |
| 29521 | None | None | R,S,T | 19.4 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8053 |
| 29522 | None | None | R,S,T | 19.4 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8046 |
| 29523 | None | S Wave 20 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8009 |
| 29524 | None | S Wave 20 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8036 |
| 29525 | None | S Wave 20 | R,S,T | 19.3 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8034 |
| 29526 | None | None | R,S,T | 19.4 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8038 |
| 29527 | None | None | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8022 |
| 29528 | None | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8050 |
| 29529 | None | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8053 |
| 29530 | None | P wave 5~10 | R,S,T | 19.4 | 58.4 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8056 |
| 29531 | B | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8059 |
| 29532 | B | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8068 |
| 29533 | B | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8074 |
| 29534 | B.Sensor | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 4 | 2 | 4 | 8091 |
| 29535 | B.Sensor | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 218 | 215 | 217 | 1 | 4 | 2 | 4 | 8088 |
| 29536 | B.Sensor | P wave 5~10 | R,S,T | 19.4 | 58.3 | 22 | 21 | 22 | 218 | 214 | 217 | 1 | 12 | 4 | 4 | 8080 |
| 29537 | B.Sensor | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 218 | 214 | 217 | 1 | 12 | 4 | 4 | 8080 |
| 29538 | B | P wave 5~10 | R,S,T | 19.4 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8046 |
| 29539 | B | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8055 |
| 29540 | B | P wave 5~10 | R,S,T | 19.3 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8034 |
| 29541 | B | P wave 5~10 | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8020 |
| 29542 | B | P wave 5~10 | R,S,T | 19.4 | 58.3 | 22 | 21 | 22 | 217 | 214 | 217 | 1 | 12 | 4 | 4 | 8039 |
| 29543 | B | P wave 5~10 | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8005 |
| 29544 | B | P wave 5~10 | R,S,T | 19.4 | 58.3 | 22 | 21 | 21 | 217 | 214 | 216 | 1 | 12 | 4 | 4 | 8024 |

학습과정 - 회귀, 의사결정트리

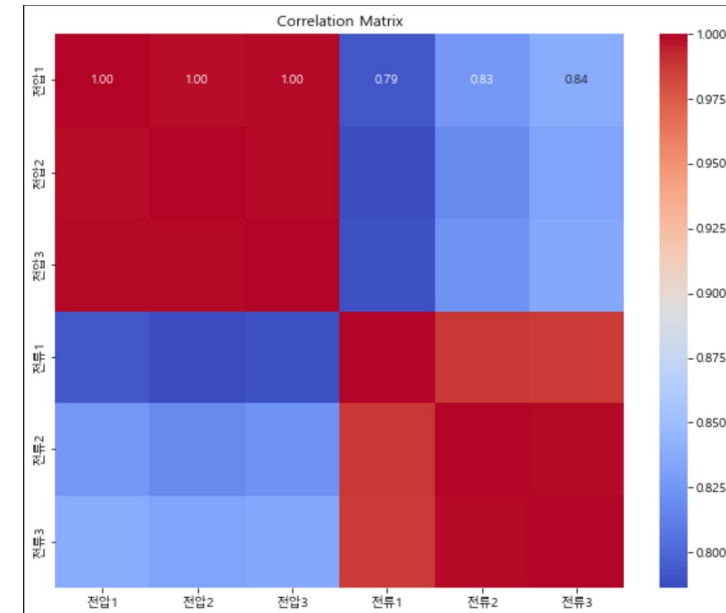


사용한 모델 – RandomForest

전체 독립변수들의 히트맵 시각화



상관관계 높은 변수들 히트맵 시각화



성능평가 – 평균 제곱 오차(MSE)

파이썬 코드 실행

```
import seaborn as sns

# 데이터셋 불러오기
data = pd.read_excel('배전반_데이터_역행수정.xlsx')

# 도어열림 열의 결측치 처리
door_mode_value = data['도어열림'].mode()[0]
data['도어열림'].fillna(door_mode_value, inplace=True)

# 지진가속도센서값 열의 결측치 처리
sensor_mode_value = data['지진가속도센서값'].mode()[0]
data['지진가속도센서값'].fillna(sensor_mode_value, inplace=True)

# 독립 변수와 종속 변수 분리
X = data[['진입1', '진입2', '진입3', '진류1', '진류2', '진류3']]
y = data['최대전력']

# 학습 데이터와 테스트 데이터 분리
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# 머신러닝 모델 선택 및 학습(RandomForestRegressor)
model_rf = RandomForestRegressor(n_estimators=100, random_state=42)
model_rf.fit(X_train, y_train)

# 모델 평가
y_pred_rf = model_rf.predict(X_test)
mse_rf = mean_squared_error(y_test, y_pred_rf)
print('Mean Squared Error (RandomForest):', mse_rf)

# 교차 검증(RandomForest)
scores_rf = cross_val_score(model_rf, X_train, y_train, cv=5, scoring='neg_mean_squared_error')
print('Cross-validated MSE (RandomForest):', -scores_rf.mean())

# 상관 행렬 계산
correlation_matrix = X.corr()

# 한글 폰트 설정
plt.rcParams['font.family'] = 'Nalgun Gothic'
plt.rcParams['axes.unicode_minus'] = False

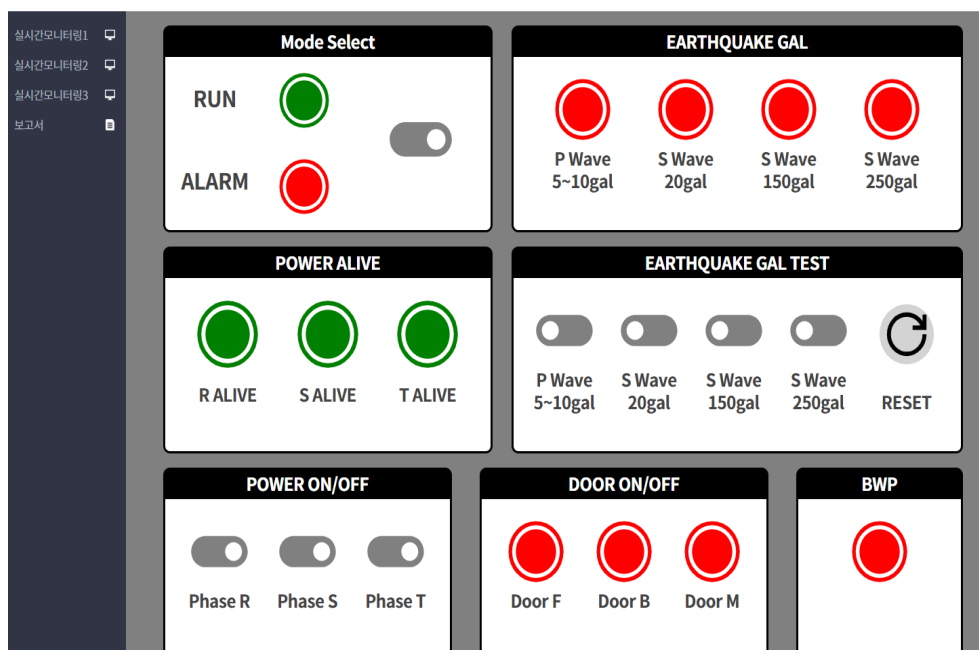
# 상관 행렬 히트맵 그리기
plt.figure(figsize=(10, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt=".2f", xticklabels=correlation_matrix.columns, yticklabels=correlation_matrix.columns)
```

MSE 결과값 및 예측값 정확도

```
MSE (Mean Squared Error)_RandomForest : 11972.1819029
교차검증(Cross-validated MSE)_RandomForest : 5254.2469
Percentage Difference: 20468    -0.065532
15306    -0.074754
25226    -0.341728
31744     0.183107
14601     0.427194
...
23920    -0.022826
18770    -0.118544
18429    -0.320926
36555     0.032870
19501     0.093638
Name: 최대전력, Length: 3659, dtype: float64
Accuracy Percentage(예측값 정확도) 98.551516807871
```


실제 현장 응용 - 센서데이터, 이상탐지

웹화면을 통한 배전반 상태 확인



규칙 조건을 통한 배전반 상태 알림

```

import pandas as pd

# 엑셀 파일에서 데이터 불러오기
data = pd.read_excel("배전실_데이터_측정수측.xlsx")

# 도어가 열렸는지 확인하는 함수
def check_consecutive_door_opening(data, start_index, consecutive_count):
    for i in range(start_index, start_index + consecutive_count):
        if i >= len(data):
            return False
        if data.iloc[i]["도어열림"] not in ['F', 'M', 'B', 'sensor']:
            return False
        return True

# 도어가 열렸을 때의 조건 확인 및 알람 기록
def check_door_condition(data, start_index, consecutive_count = 10):
    # 연속된 도어열림 상태를 확인할 횟수
    if check_consecutive_door_opening(data, start_index, consecutive_count):
        for i in range(start_index, start_index + consecutive_count):
            row = data.iloc[i]
            if row['속도'] >= 29 or row['속도'] >= 65:
                if row['속도'] >= 29:
                    print(f"도어가 연속해서 열렸으며 속도가 29 이상입니다. (총: {i+1})")
                if row['속도'] >= 65:
                    print(f"도어가 연속해서 열렸으며 속도가 65 이상입니다. (총: {i+1})")
        return True
    return False

# 화재상태가 RST 중 어떤 값인지 확인하고 도어가 열렸을 때 위험신호로 감지하는 함수
def check_power_condition(data, start_index):
    consecutive_count = 10 # 연속된 도어열림 상태를 확인할 횟수
    if check_consecutive_door_opening(data, start_index, consecutive_count):
        for i in range(start_index, start_index + consecutive_count):
            row = data.iloc[i]
            if 'M' in row['화재신'] or 'S' in row['화재신'] or 'T' in row['화재신']:
                print(f"도어가 연속해서 열렸으며 RST 중 하나이상 화재상태가입니다. (총: {i+1})")
        return True
    return False

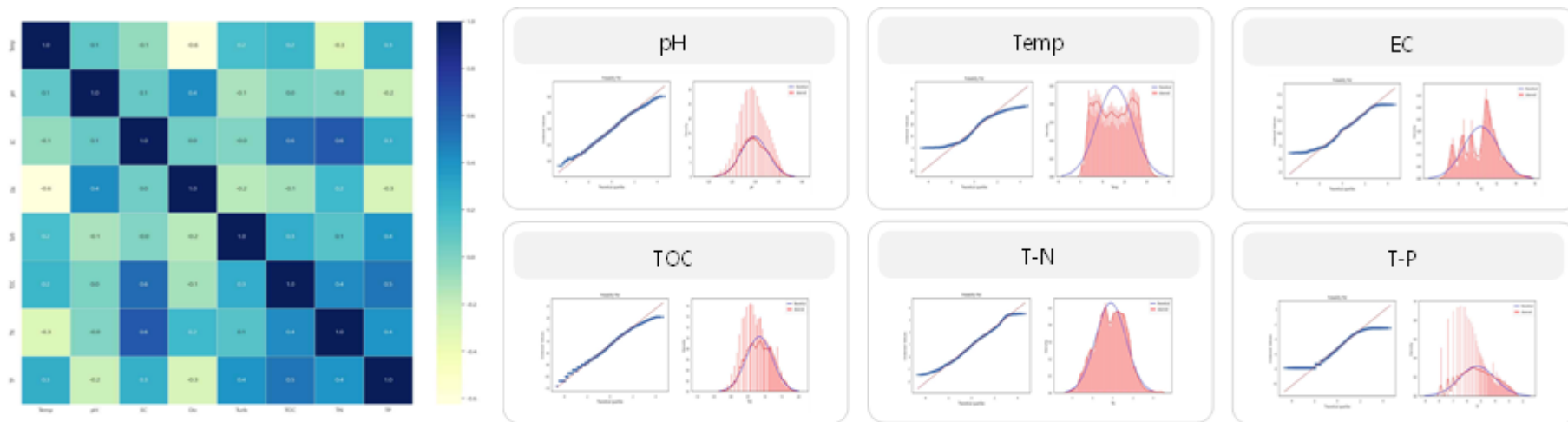
# 알람 횟수를 기록할 변수
alert_count = 0

# 데이터 프레임의 각 행에 대해 조건을 검사하고 알람을 줄
def check_alarm_condition(data, i):
    if check_door_condition(data, i):
        alert_count += 1
    if check_power_condition(data, i):
        alert_count += 1

print("알람을 준 횟수:", alert_count)

```

프로젝트 수행을 통해서 얻은 경험



- 입력 데이터(Feature) - 신규센서 설치 후 확보된 실시간 원본 데이터를 기반으로 결측치, 이상치 및 상관계수를 분석 머신 러닝으로 지도 학습
- 출력 데이터(Label) - 실시간 측정이 상대적으로 어려운 고농도 구간 3개의 변수(TOC, TN, TP)를 출력 데이터로 선정

수질분석 데이터에도 적용하여 상하수도 관련 인자 값의 상관관계 분석에도 활용 가능

프로젝트 수행을 통해서 얻은 경험



기존자료를 통한 학습 및 테스트를 통하여 예측 값과 실제 값을 비교 가능



상관행렬을 통한 다양한 변수 간의 상관관계 확인



공공 데이터 포털 활용
주요지표 탐색
(금융, 날씨, 교통, 고용, 교육, 물류 등)