

## 1. 개발 목적

이 보고서의 목적은 시간대별 사고 빈도를 예측하는 안전 관련 머신러닝 모델을 개발하고 해당 모델의 성능을 평가하는 것이다. 데이터의 독립 변수로 시간을 사용하여 종속 변수로 사고 빈도를 예측하는 모델을 통해, 특정 시간대에 발생할 수 있는 사고의 예측을 도출하고자 한다.

## 2. 개발 계획

### 2.1 데이터 요약 정리 및 시각화

사용한 데이터는 시간대와 사고 빈도에 대한 간단한 랜덤 데이터로 구성되어 있다.

데이터의 특성을 살펴보고, 시각화를 통해 시간과 사고 빈도 간의 관계를 확인했다.

### 2.2 데이터 전처리 계획

데이터는 시간과 사고 빈도로 이루어져 있어 추가적인 전처리가 필요하지 않았습니다.

### 2.3 머신러닝 모델 선택

랜덤 포레스트 회귀 모델을 선택했다. 이 모델은 다양한 특성에 대한 예측을 수행하며 안정적인 성능을 보이는 특징이 있다.

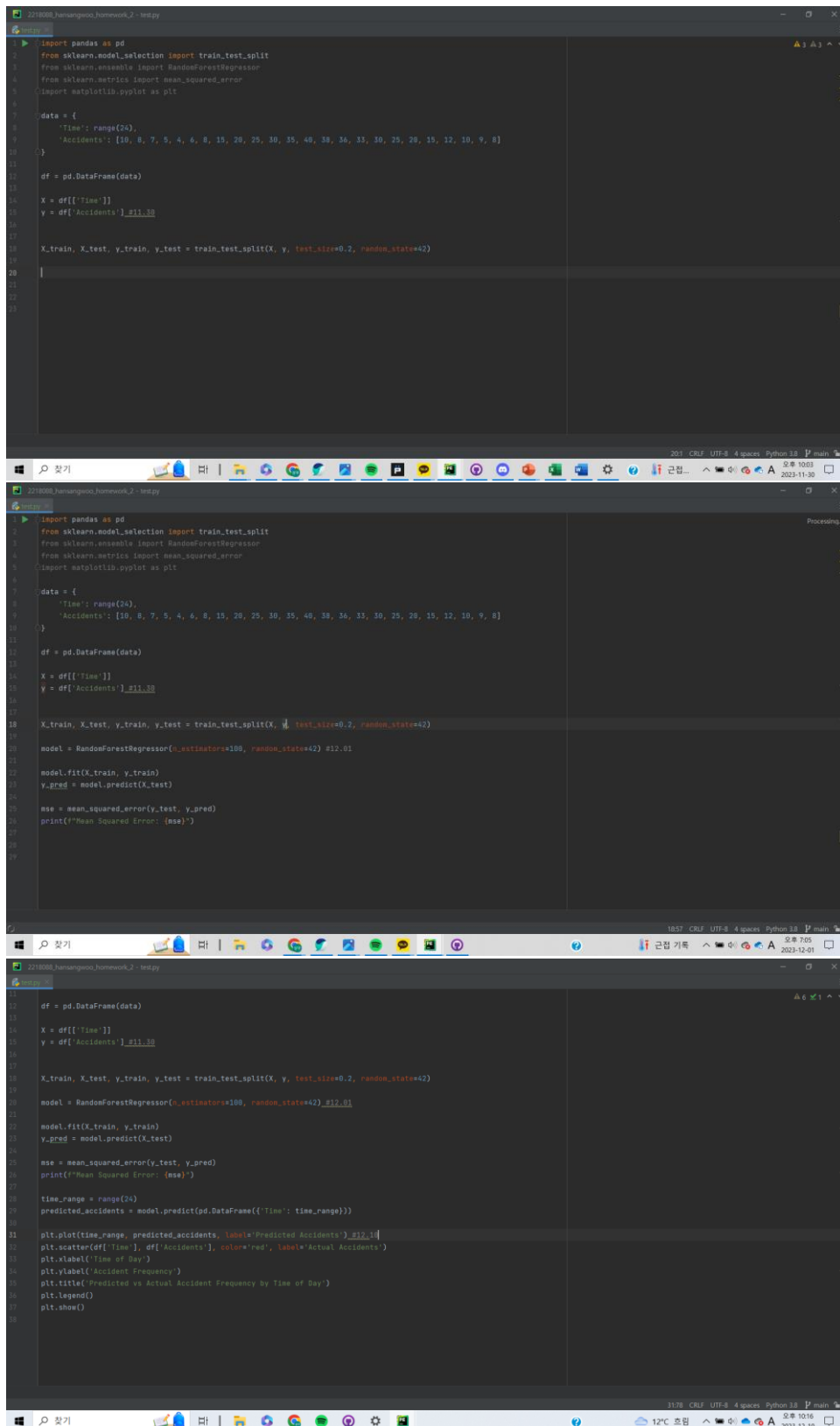
### 2.4 예측 결과 및 성능 지표

모델을 사용하여 시간대별 사고 빈도를 예측하고, 성능 평가를 위해 평균 제곱 오차(Mean Squared Error, MSE)를 사용했다.

### 2.5 성능 검증 방법 계획

학습 데이터와 테스트 데이터를 분리하여 모델의 일반화 성능을 평가했다.

## 3. 개발 과정 및 후기



### 3.1 학습 모델 개발 과정

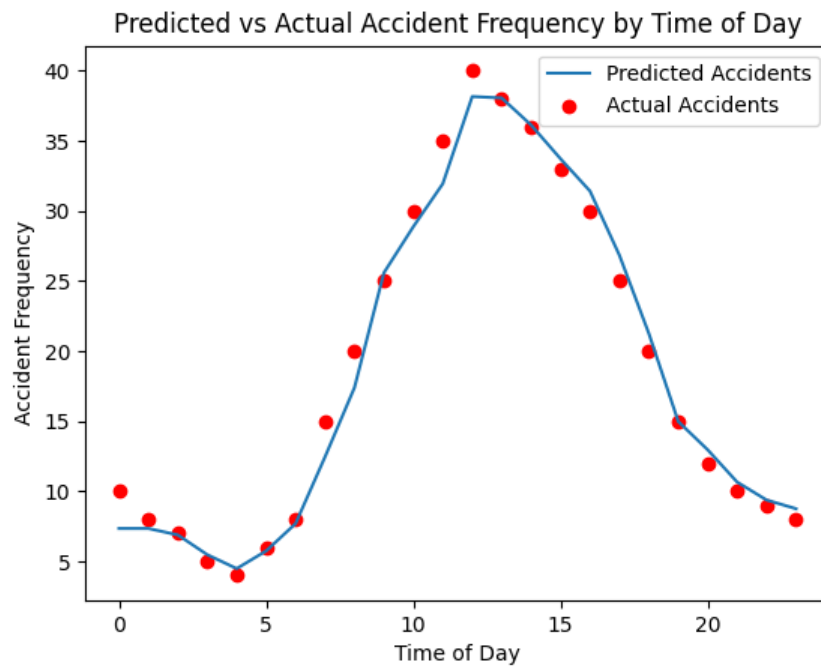
데이터를 통해 랜덤 포레스트 회귀 모델을 학습시켰습니다. 학습된 모델을 사용하여 시간대별 사고 빈도를 예측하고 시각화했다.

## 모델 개발과정

### 3.2 모델의 성능 평가

테스트 데이터를 사용하여 모델의 성능을 평가했다. 평균 제곱 오차를 통해 모델의 예측 정확도를 확인했다.

### 3.3 데이터 시각화



### 3.4 개발 후 느낀 점

랜덤 포레스트 모델은 간단하면서도 효과적인 모델로, 주어진 데이터에 대해 잘 작동했다.

모델의 성능을 높이기 위해 더 많은 데이터와 특성을 고려하는 것이 유용할 것으로 판단된다.

추가적인 튜닝 및 다른 모델과의 비교를 통해 모델 성능을 개선하는 방안을 고려할 필요가 있다.

이러한 개발 과정을 통해 안전 관련 머신러닝 모델을 설계하고 평가하는 방법을 익히고자 하였다.