**머신러닝 모델 이름**

**학번: 2018039**

**이름: 이두호**

**Github address:** Upload files · 2018039DuhoLee/homework\_2 (github.com)

1. **안전 관련 머신러닝 모델 개발의 목적** 
   1. 진동안전모델을 만들었는데, 이 모델을 이용해서 작업을 수행할때 안전한 진동과 불안전한 진동을 나누어 안전한 진동하에서 작업이 수해될 수 있도록 할 수 있다. 따라서 안전한 작업환경에서 작업을 수행하기 때문에 사고 발생율이 줄어든다.
   2. 학습 모델 활용 대상: 섬세한 작업을 수행하는 등 진동에 민감한 작업장
   3. 데이터의 어떠한 독립 변수를 사용하여 어떠한 종속 변수를 예측하는 지

: 데이터에서 사용된 독립변수는 주파수와 진폭이며, 이를 통해 현재 진동상황이 작업상황에서 안전한지 안전하지 않은 지를 예측합니다.

* 1. 개발의 의의: 현장에서 일하는 노동자들이 작업도중 억울한 부상을 당하지 않게 하기 위해서 모델을 개발했습니다. 안전하고 쾌적한 작업환경을 조성하기 위해서 개발했습니다.

1. **안전 관련 머신러닝 모델의 네이밍의 의미**

이번 머신러닝의 이름은 ‘우당탕탕 진동 판별기’입니다. 작업 중 우당탕탕 사고가 발생하지 않았으면 하는 마음에서 이름을 지어보았습니다.

1. **개발 계획**

1)데이터에 대한 요약 정리 및 시각화 : 주파수와 진폭으로 그래프 표시

2)데이터 전처리 계획: 데이터를 얻어오지 못해서 임의의 데이터로 진행.

3)어떠한 머신러닝 모델을 사용할 것인지 (해당 머신러닝 모델의 이론 추가)

: 랜덤 포레스트 분류기를 사용했고 이 모델은 RandomForestClassifier는 scikit-learn 라이브러리에서 제공되는 앙상블 기반의 머신러닝 모델 중 하나로 랜덤 포레스트는 여러 결정 트리를 생성하고 그들의 예측을 결합하여 과적합을 줄이고 안정적인 예측을 제공하는 알고리즘입니다.

4)머신러닝 모델 예측 결과가 어떠할 지

: 진동수와 진폭을 사용해서 안전한 진동정도와 비 안전한 진동정도를 표시할 것 같다.

5)사용할 성능 지표

: 정확도, 혼동행렬, 분류보고서가 사용 될 것 같다.

6)성능 검증 방법 계획 등

: 교차검증, 갠덤 포스트 모델을 사용하여 성능을 검사할 예정.

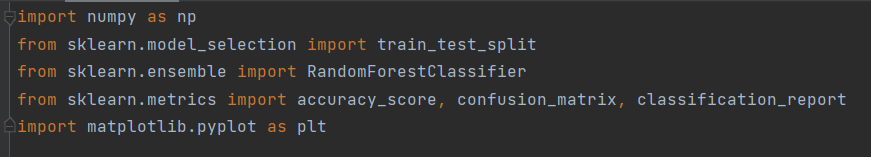
1. **개발 과정**

1. 필요한 라이브러리들을 부른다.

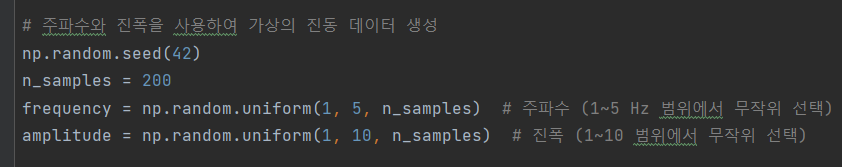
- NumPy는 수학 및 배열 연산에 사용.

- scikit-learn의 RandomForestClassifier는 랜덤 포레스트 모델을 제공

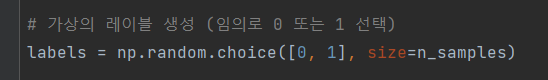
-\_score, confusion\_matrix, classification\_report등의 함수로 성능을 평가

- matplotlib.pyplot 시각화를 위해 사용 

2. 가상의 데이터 생성 np.random.uniform함수를 사용하여 주파수와 진폭을 무작위로 선택.

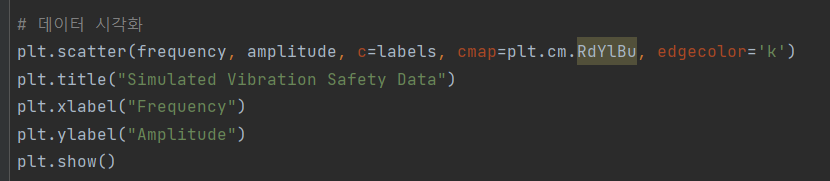


3. 가상의 레이블은 0 또는 1의 값을 무작위로 선택하여 생성하는데. 이는 이진 분류 문제에서의 레이블로 활용.

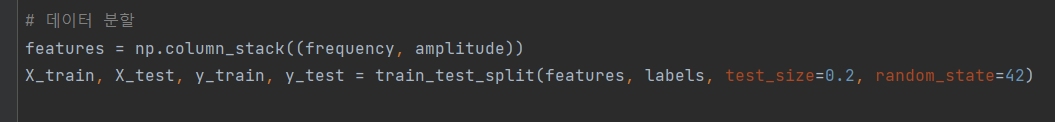


4.matplotlib.pyplot을 사용하여 주파수와 진폭을 가지고 생성한 데이터를 시각화하고,

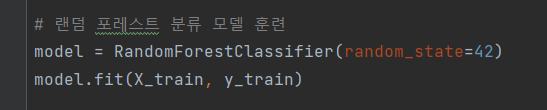
산점도를 사용하여 주파수와 진폭을 x, y 축에 표시 후 레이블에 따라 다른 색상을 부여하여 시각적으로 확인할 수 있도록 함.



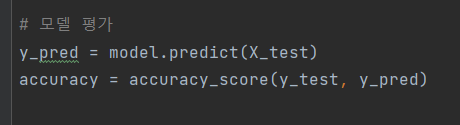
5.train\_test\_split함수를 사용하여 데이터를 훈련 세트와 테스트 세트로 분할하고 이렇게 나누어진 데이터를 통해 모델을 훈련하고 평가.



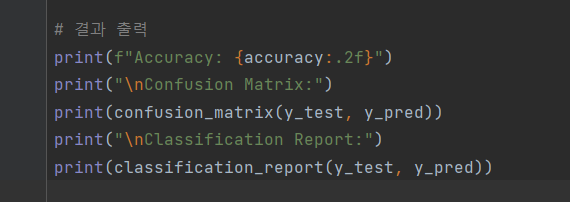
6.RandomForestClassifier를 사용하여 랜덤 포레스트 분류 모델을 생성하고, 훈련 데이터를 사용하여 모델을 훈련.



7.predict함수를 사용하여 테스트 데이터에 대한 예측을 수행하고, accuracy\_score함수를 사용하여 모델의 정확도를 계산.

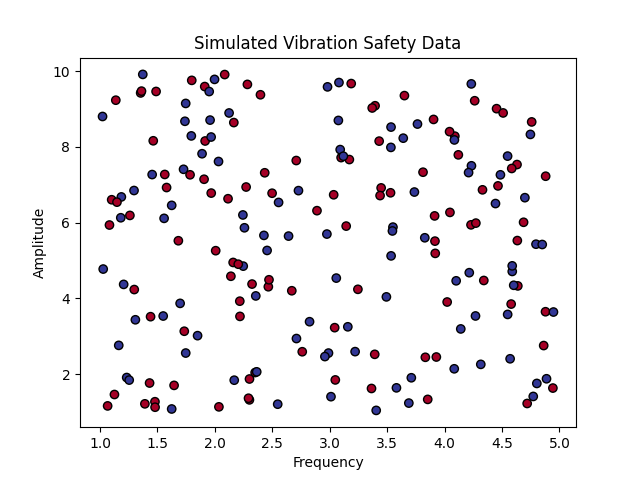


8.마지막으로, 모델의 성능을 출력 후 정확도, 혼동 행렬, 분류 보고서 등을 통해 모델의 성능을 평가.



개발과정에서 데이터를 불러오는데 오류가 발생해서 할 수 없이, 임의로 데이터를 생성한 후 개발을 진행했습니다.

결과의 시각화



1. **개발 후기**

**머신러닝의 활용가능성에 대해 알 수 있었다. 다시금 코딩의 중요성을 알게되었고, 또한 다시금 진짜 파이썬이 어렵다는 것을 알게 되었습니다. 교수님 한 학기동안 같이 할 수 있어서 너무 영광이였습니다. 제가 교수님 수업을 좀 더 빨아먹었어야 했는데 아쉽습니다. 참고로 저 남자고 여자친구도 있습니다. 오해 하지 않아주셨으면 합니다.**

**감사합니다.**

**(교수님 안재현 닮았어요!!)**