|  |  |  |  |
| --- | --- | --- | --- |
| **팀명** | AI제조 팀 | **이름** | 남도형 |
| **일시** | 8/10 | **장소** | 온라인 |
| **프로젝트명** | 머신러닝을 활용한 스마트폰 불량품탐지 | | |
| **개별활동**  **내용** | AD with PCA는 "Anomaly Detection with Principal Component Analysis"의 약자로, 주성분 분석을 사용하여 이상 감지를 수행하는 방법을 의미한다. 이 방법은 데이터의 차원 축소와 이상 감지를 결합하여 데이터의 패턴을 이해하고 이상 값을 식별하는 데 도움이 된다.    PCA는 다차원 데이터를 고차원 공간에서 주성분이라는 새로운 축으로 변환하는 기법이다. 주성분은 데이터의 분산을 최대화하는 방향으로 정의되며, 주요한 정보를 보존하는 더 낮은 차원의 표현으로 데이터를 압축한다. 이를 통해 데이터의 노이즈나 불필요한 변동성을 줄이면서 중요한 패턴을 강조할 수 있다.  AD with PCA는 4개의 절차를 따른다.  먼저 데이터 전처리로 데이터를 준비하고 특성을 표준화하여 각 특성의 스케일을 맞추는 등의 전처리 작업을 수행한다.  다음은 PCA 적용이다. 데이터의 주성분을 찾기 위해 PCA를 적용한다 PCA를 통해 원본 데이터의 차원을 줄이는데, 이렇게 줄인 차원에서 주성분은 데이터의 가장 큰 변동성을 설명하는 축으로 정의된다.  다음은 주성분 선택이다. 주성분은 주로 누적된 설명 분산을 기준으로 선택된다. 예를 들어, 주성분 1,2,3을 사용하여 전체 분산의 90% 이상을 설명할 수 있다면, 이 주성분을 선택한다.  마지막으로 4번째는 이상감지이다.  주성분 축으로 변환된 데이터를 사용하여 이상값을 감지한다. 이상 값을 식별하기 위해 주성분 축에서 원래 데이터까지의 거리를 계산하거나, 주성분 축에서의 분산 비율 등을 활용할 수 있다. 일반적으로 주성분 축으로부터 멀리 떨어진 데이터 포인트는 이상으로 간주될 수 있다. | | |
| **향후추진 계획** | LightGBM에 대해서 학습할 예정이다. | | |