|  |  |  |  |
| --- | --- | --- | --- |
| **팀명** | AI제조 팀 | **이름** | 남도형 |
| **일시** | 8/7 | **장소** | 온라인 |
| **프로젝트명** | 머신러닝을 활용한 스마트폰 불량품탐지 | | |
| **개별활동**  **내용** | Lsolation Forest는 데이터 포인트들 중에서 이상치를 탐지하는 머신러닝 알고리즘 중 하나이다. 이상치란 일반적인 데이터 패턴과 매우 다른 특이한 데이터 포인트를 말한다. 예를 들어, 신용카드 거래 데이터에서 도둑질로 인한 이상한 거래나 네트워크 트래픽 데이터에서 해킹 시도와 같은 이상한 활동이 이상치에 해당한다.    Random Fores를 사용한 이상치 탐지이다.    이상 관측치는 데이터의 분포/밀도로부터 멀리 떨어져 있는 관측치이다. 왼쪽 그림은 정상 관측치를 고립시키는 그림이다. 그림과 같이 정상 관측치는 다른 관측치들로부터 고립시키기 어렵다. 하지만 이상관측치는 다른 관측치들로부터 고립시키기 쉽다. 이상치를 판단하는 score를 고립시키는 split 횟수로 본다면 이상치를 판단할 수 있다.    모든 관측치가 고립되도록 split시키는 decision tree는 ltree이다.  Ltree의 Bagging 구조는 lsolation Forest이며 Aggregation은 여러 개의 ltree가 이상으로 탐지한 관측치를 다수결로 이상치를 정의하는 것이다. | | |
| **향후추진 계획** | LightGBM에 대해서 학습할 예정이다. | | |