**안전한 승강**

**발표자: ???**

**팀원:2조원들**

**Github: seunghyun9999/~**

1. **안전 관련 머신러닝 모델 개발 관련 요약**
   1. 프로젝트에 관한 전체 내용을 요약

= IoT 센서가 한 달 동안 1분 단위로 캡처한 데이터를 분석하여 엘리베이터가 고장 났거나 고장이 나려고 하는지 예측할 수 있는 모델을 개발

1. **개발 목적**
   1. 머신러닝 모델 활용 대상: 엘리베이터로서 개발한 이 프로그램을 건설현장의 리프트 곤돌라 등 인양작업을 하는 장비들에 적용시켜서 안전관리를 더욱 편하고 정밀하게 하는 것
   2. 개발의 의의: 모델 개발 시 어떠한 가치를 창출할 것으로 예상하는지 등

= 안전관리자의 업무가 상당히 많다 주요 업무인 사업장 순회점검을 하면서 승강기에 대하여 항상 많은 데이터를 꼼꼼히 보기 란 어려운 일이다

하지만 우리가 개발한 프로그램을 이용하면 iot로서 실시간으로 데이터를 전송받아 승강장비가 고장 날 상황인지 바로 파악이 가능하고 고장이 일어 날 것이라고 예상된다면 미리 알람을 통해 알아 차릴 수 있다. 이로서 승강 장비에 대한 안전관리 시간 감소와 정확도 향상을 기대해 볼 수 있다.

* 1. 데이터의 어떠한 독립 변수를 사용하여 어떠한 종속 변수를 예측하는지

= 데이터 분석 글은 뒤에 있음

1. **배경지식(말그대로 우리 프로젝트를 처음보는 사람이 이해할 수 있도록 배경 말해주기)**
   1. 데이터 관련 사회 문제 설명

= 승강설비의 고장에 관해서는 대개 큰 사고로 이어진다~~~ (이런 느낌)

* 1. 머신러닝 모델 관련 설명 등

= 머신 러닝을 진행한 디시전트리와 로지스틱 회귀 방식은 이런거다

각각 이러한 특징이 있다.

1. **개발 내용**
   1. 데이터에 대한 구체적 설명 및 시각화
      1. 데이터 개수, 데이터 속성 등

= 데이터 개수 44640 => 30일은 44640분

그리고 독립변수 12개 종속변수1개

* + 1. 데이터 간 상관관계 설명 등

= 첫번째로 시간은 종속변수와의 관계성이 아예 없으므로 제외

두번째로 Sensor1과 Vibrations의 상관계수: 0.9999999986071603

RPM과 Sensor6의 상관계수: 0.9980335313388843

둘은 매우 높은 상관관계를 나타냄 센서 1 센서 6 제외가능

각 센서에 대한 출처가 너무 분명하지 않고 status와의 관계도 동일 하기 때문에 rpm과Vibrations만 채택하고 센서 1, 6은 제외시킨다

세번째로

* 1. 데이터에 대한 설명 이후, 어떤 것을 예측하고자 하는지 구체적으로 설명
     1. 독립변수, 종속변수 설정

= 사실 데이터를 보면 결과가 비정상일 때 많은 정보들이 0을 나타낸다 이를 보완하기 위해 자주 0을 보이는 데이터 모두 0일때의 행을 제외시키고 그 외 결과가 비정상이지만 다른 데이터 들이 0이 아닌 정보를 가지고 있는 경우에 가중치를 더해줘서 딥러닝을 진행

* + 1. 우리가 예측하고자 하는 정보

= 위에서 제시하는 독립변수를 대입하며 현재 엘리베이터가 고장의 위험이 있는가

* 1. 머신러닝 모델 선정 이유
     1. 설명한 데이터를 기반으로 머신러닝 모델 선정 이유 설명

= 기본적으로 이번 데이터의 결과가 정확히 정상 혹은 비정상으로 나눠지기 때문에 디시전 트리와 로지스틱 회귀방식이 적합하다고 생각하였다

* + 1. 성능 비교를 위한 머신러닝 모델 선정 이유

= ????

* 1. 사용할 성능 지표
     1. 머신러닝 모델의 성능을 평가하기 위해 사용하는 성능 지표에 관한 설명 등

=

* + 1. 성능 지표 선정 이유 등

=

1. **개발 결과(내가 해야하는거)**
   1. 성능 지표에 따른 머신러닝 모델 성능 평가
      1. 수치 자료 및 시각화 자료를 사용
         1. MAE(평균 정대 오차), RMSE(평균 제곱근 오차), MSE(평균 제곱 오차), Accuracy(정확도), 오차행렬(긍정, 부정 예측 파악인데 나쁘지 않아) 등
         2. KFold 결과
      2. 다른 머신러닝 모델과 성능 비교
   2. 머신러닝 모델의 성능 결과에 대한 해석
2. **결론**
   1. 머신러닝 모델 개발에 관한 간략한 요약 및 결과 설명
   2. 개발 의의 등
   3. 머신러닝 모델의 한계

**데이터 분석**

데이터 수 44640개

빨간 글씨가 0이 막 나오는 글씨임

1. 시간: 2020-01-01 00:00 부터 1분씩 2020-01-31 23:59 까지(딥러닝에 필요 없음)
2. !온도: 상관관계 높게 나옴 (엘리베이터 고장 시 센서가 작동하지 않아서 0으로 나와 그래보임)
3. 습도: 43262행부터 막판에 센서가 망가졌는지 이상하게 나옴
4. !=RPM: 2가 될 때 즈음 이상해지다가 0이나옴 (한발 늦음 의미가 좀 적음)
5. -진동: 2가 될 때 즈음 이상해지다가 0이나옴 (한발 늦음 의미가 좀 적음)
6. 압력: 항상 일정함 (의미가 적음)
7. -센서1: 고장난 경우 0이 되는데 그러다가 갑자기 지혼자 돌아옴ㅋㅋ
8. !센서2: 결과 값이 2인 경우 빈공간으로 나옴 그 외 모두 값이 나옴
9. 센서3:
10. 센서4:
11. !센서5: 결과 값이 2에 가까워질 때 값이 낮아짐
12. !=센서6:
13. 결과: 0=정상 1=고장 직전에 알려주는거 2=고장

2가 나오는 행 17158~18101=943, 24513~27622= 3109 총 4052행텍스트, 스크린샷, 사각형, 패턴이(가) 표시된 사진

자동 생성된 설명

**Sensor1과 Vibrations의 상관계수: 0.9999999986071603**

**RPM과 Sensor6의 상관계수: 0.9980335313388843**

**둘은 매우 높은 상관관계를 나타냄**

텍스트, 전자제품, 스크린샷, 소프트웨어이(가) 표시된 사진

자동 생성된 설명

**이 데이터가 기본적으로 시계열을 추가하여 분석하도록 짜여있다**

**그래서 내가 보기엔 그냥 딥러닝 해봤다가 중요하지 이걸로 진짜로 프로그램짜서 내가 돈잘벌고 잘살겠다는 게 아니니까**

**각 데이터 상관관계 잘 분석했고 그 분석한 값들이 이러해서 내가 이렇게 변경해서 딥러닝을 시켰다**

**!!전처리를 이렇게나 잘혔다 이런거**

**그래서 정확도가 이정도 나왔고 그래서 요로코롬 되었습ㄴ이다.**

**이렇게 하고 그냥 시각화를 좀 잘하는 느낌으로 가야할 듯**

**시각화를 어찌해야하냐**

**시각화를 해야 하는 것**

1. **각 독립변수 별로 종속변수와의 관계를 나타내는 시각화**

* **레드줄스에 나와있는 차트 정도**

1. **딥러닝 예측 결과에 관한 시각화**

* **산점도? 너무 겹치긴 함…**

**그리고 해야 하는 일들**

1. **템플릿 만들기**
2. **리드 미 꾸미기**
3. **Ppt? 만들거니?**
4. **발표? 준비?**
5. **코드 다듬기**

* **엑셀로 전환을 해주거나 클라스 같은 배웠던 것을 활용하여 무언가를 만들거나 알림, 정상이지만 비정상에 얼마나 가까운지 보여주는 거 + 데이터 요약 하기 보여주는 거 디스크립트 그거**