**안전한 승강**

**발표자: ???**

**팀원:2조원들**

**Github: seunghyun9999/~**

1. **안전 관련 머신러닝 모델 개발 관련 요약**
   1. 엘리베이터를 건설 현장의 리프트와 곤돌라 등 양중기에 접목하여IoT 센서가 한 달 동안 1분 단위로 캡처한 데이터를 분석해 실시간 모니터링을 통해 엘리베이터의 잠재적인 고장 가능성 및 고장 유무을 예측할 수 있는 모델 개발
2. **개발 목적**
   1. 모델 활용 대상: 엘리베이터의 고장 예측 및 사고 예방 개발 프로그램을 건설현장의 고위험장비인 리프트, 곤돌라 등 양중기에 적용시켜서 안전관리를 더욱 편하고 정밀하게 하는 것
   2. 개발의 의의: 건설현장의 대형화 및 복잡화로 인해 안전관리의 범위가 넓어짐에 따라 현장 내 모든 작업 데이터를 꼼꼼하게 파악하기는 어려운 실정임. 잠재적인 위험을 사전에 예측하기 위해서 장비의 상태와 작업 환경 데이터를 실시간으로 추적 및 모니터링해 사고를 예방하는 것이 필요함. 엘리베이터의 고장 예측 및 사고 예방 개발 프로그램을 활용 시 실시간으로 센서 데이터를 분석하여 승강 장비의 작동 패턴, 이상 징후, 고장 전조 등 식별이 가능하며, 잠재적인 고장 가능성이 있을 경우 사전 알림을 통해 파악 가능함. 따라서 승강 장비의 상태 파악으로 적절한 유지보수 시점 결정, 고장 예측 및 조기 경고 시스템으로 승강 장비의 안전성 확보 및 안전관리 효율성과 정확성 기대 가능.
   3. 데이터의 어떠한 독립 변수를 사용하여 어떠한 종속 변수를 예측하는지
3. **배경지식**
   1. 데이터 관련 사회 문제 설명: ‘승강기 사고 및 기술자 사고 통계'에 따르면 2019년부터 약 5년간 승강기 사고로 인한 사상자는 335명. 이 중 사망자는 27명, 부상자는 308명 수준으로 사고 발생률이 높음. 건설현장의 경우에도 고층에서 작업하거나, 무거운 자재를 들어 올리는 작업이 많기에 리프트와 곤돌라 같은 장비 사용이 잦음. 이러한 승강 장비 사용으로 인한 주요 사고사례로는 권과방지장치 미흡으로 과상승한 운반구로 인한 추락, 운행 전 점검 미흡으로 리프트 운반구 추락 등 추락 사고가 다수 발생하므로 사고 위험도가 매우 높음.
   2. 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**



