

분석기획서

팀 명 : 아, 진짜요. 와 대박.

1.비즈니스 배경 및 문제 식별

<비즈니스 배경>

건설장비 대여 업체를 운영하며 장비 대여 전 작동오일 점검을 실시한다.

<문제 식별>

1. 장비 성능 저하

- 오일 상태가 나빠지면 장비의 윤활 및 냉각 기능이 저하되어 성능이 떨어지고 작업 효율이 감소한다.

2. 부품 마모 가속화

- 윤활 부족으로 부품 간 마찰이 증가하고, 부품 마모가 가속화되어 교체 주기가 빨라지고 유지보수 비용이 상승한다.

3. 과열 문제

- 오일의 열 전달 능력 저하로 장비 내부 온도가 상승하여 과열 문제가 발생하고, 장비 수명이 단축될 수 있다.

4. 고장 위험 증가

- 오염, 점도 변화, 산화, 물 혼입 등의 문제로 인해 장비 고장 위험이 증가하고, 적절한 유지보수가 이루어지지 않으면 가동 중단 발생 가능성이 높아진다.

5. 작동 효율 저하

- 오일 상태 악화로 인한 소음이 증가하며, 작업 속도가 저하될 수 있고 출력 감소 등으로 장비의 전반적인 효율성이 저하된다.

6. 고객과의 분쟁 발생 위험

- 오일 상태 이상으로 장비 고장 발생 시 고객과의 분쟁이 발생하여 작업 지연과 비용 증가로 이어진다.

결론

오일 상태에 영향을 미치는 요소들에 대해서 파악하고, 오일 상태를 분석하지 않으면 장비 성능 저하, 부품 마모 가속화, 과열 및 고장 위험이 커져 장비 운영에 심각한 문제가 발생할 수 있다. 이를 방지하기 위한 예방적 유지보수 및 오일 상태 진단이 필수적이다.

2.1 분석주제 확정

분석주제명	비즈니스 문제	분석방안	활용방안 및 기대효과
장비의 성능과 작동 효율성을 위해서 오일 상태를 판단할 모델을 개발한다.	장비 성능이 저하되고, 부품 마모가 가속화되며, 과열 문제가 발생한다. 또한, 고장 위험이 증가하고, 작동 효율이 저하된다. 이로 인해 고객과의 분쟁 발생 위험이 높아진다.	오일 상태에 영향을 미치는 요소들에 대한 데이터를 사용해서, 오일 상태를 예측하는 모델을 훈련한다. 불균형 데이터이므로, SMOTE 혹은 샘플링, 모델 가중치 조정으로 보완한다. CatBoost, XGBoost, 로지스틱회귀로 오일 상태를 분류하고, 영향력 있는 변수를 찾는다. 모델 성능은 ROC-AUC, G-Mean으로 평가하고 Optuna로 최적화한다.	오일 상태 점검과 사전 유지보수를 통해 장비 성능을 유지하고 고장 위험을 줄여 고객 만족도를 높인다. 이를 통해 장비 사용 수명 증가와 유지보수 비용 절감 효과를 기대할 수 있으며, 비가동 시간을 줄여 현장 운영 효율성을 높일 수 있다.

2.2 해결 방안 수립 및 주제 확정

주제: 오일의 이상 유무를 탐지하기 위한 오일 상태 판단 모델 개발

해결 방안

- 오일 상태 데이터 수집

- **데이터 수집 (기존 데이터 활용):** 기존에 저장된 오일 관련 데이터를 수집하고 이를 활용해 모델을 훈련한다.
- **센서 기반 실시간 모니터링 대신 기존 데이터 분석:** 실시간 센서 데이터 수집은 배제하고, 제공된 데이터셋을 분석해 오일 상태를 평가하는 데 집중한다.
- **오일 분석 시스템 구축**
 - **기존 데이터를 기반으로 분석:** 주기적인 오일 샘플 분석을 통해 오일의 화학적 성분 및 점도를 평가하는 대신, 이미 제공된 데이터를 활용해 머신러닝 모델을 훈련하여 오일 상태를 판단한다.
 - **AI 모델 기반 예측:** 분석 결과와 데이터를 바탕으로 머신러닝 모델을 개발해 오일 상태 이상 여부를 분류하는 시스템을 구축한다.
- **오일 상태 판단 모델 개발**
 - **이상 탐지 모델 구축:** 기존 오일 데이터를 이용해 머신러닝 알고리즘(CatBoost, XGBoost 등)으로 정상/이상 상태를 분류하는 모델을 개발한다.
 - **클래스 불균형 해결:** 제공된 데이터에서 이상 데이터가 적을 경우, 오버샘플링 기법(SMOTE)이나 클래스 가중치를 조정해 모델 성능을 개선하는 방식을 사용한다.
 - **모델 성능 평가 및 최적화:** 제공된 데이터에 맞춰 모델의 성능을 F1-Score, ROC-AUC, G-Mean 등을 통해 평가하고, 하이퍼파라미터 최적화 도구(Optuna 등)를 사용해 성능을 개선한다.

기대 효과

1. 실시간 오일 상태 모니터링으로 인한 이점

- **고장 예방:** 오일 상태를 주기적으로 모니터링하여 이상 징후를 조기에 발견하고, 예상치 못한 고장을 방지할 수 있다.
- **유지보수 비용 절감:** 불필요한 부품 교체나 점검을 줄이고, **예방 유지보수**를 통해 비용 효율적 관리를 할 수 있다.
- **장비 가동 시간 증가:** 예기치 않은 중단을 줄이고, 장비의 **연속적인 운영**을 보장한다.
- **작동 효율 향상:** 오일 상태가 이상적으로 유지되므로 장비의 마모가 줄고, 작업 효율이 향상된다.

2. 데이터 기반 유지보수 최적화

- **정확한 교체 시기 예측:** AI 모델을 통해 오일 교체 시기 예측, 장비의 성능과 수명을 최적화한다.
- **비정상 상태 조기 대응:** 이상 상태 감지 모델을 통한 즉각적인 유지보수로, 장비 성능 저하를 최소화한다.

3. 개선된 성능 평가

- **Precision, Recall, ROC-AUC, G-Mean** 등의 성능 지표를 활용하여 **이상 감지 성능 평가**가 가능하다.
- 소수 클래스(이상 상태)에 대한 **예측 신뢰 구간**을 통해 모델의 **예측 안정성** 평가를 할 수 있다.

3. 데이터 정의 및 식별

- 1) **컬럼명** : unknown1, unknown2, unknown3, unknown4, unknown5, unknown6, unknown7, unknown8, unknown9, unknown10, unknown11, unknown12, unknown13, unknown14, unknown15, unknown16, unknown17, target
- 2) **생성 규칙** : 기계별로 수집된 작동오일 관련 데이터
- 3) **관리 상태** : n년간 수집된 데이터

4. 분석정의

4.1 변수의 역할 정의

- **설명변수:** 오일 상태를 예측하기 위해 사용하는 다양한 측정 변수들이다. 이 프로젝트에서는 주어진 데이터 중 오일 상태에 영향을 미칠 수 있는 수치형 및 범주형 변수를 설명변수로 정의한다. 주요 변수로는 온도, 압력, 점도 등 다양한 오일 특성 데이터가 포함될 수 있다.
- **타겟변수 (종속변수):** target 컬럼은 오일의 이상 여부를 나타내는 변수로, 0(정상)과 1(이상)으로 구분되어 있다. 이 변수는 모델의 예측 대상이다.

4.2 분석기간 정의

2024년 10월 14일 ~ 2024년 10월 17일 (4일)

4.2.1 목표변수기간 정의

건설장비 대여 1개월 전 시점의 작동오일 정상여부를 분석한다.

4.2.2 활용기간 정의

건설장비 대여 1개월 전

4.2.3 설명변수기간 정의

건설장비 대여 1개월 전까지 n년간

4.3 평가 기준 정의

- **Precision, Recall:** 2종오류를 줄이기 위해 recall을 우선적으로 고려하여 사용한다.
- **ROC AUC:** TPR과 FPR을 고려하여 모델 성능을 비교하기 위해 사용한다.
- **G-Mean:** 클래스 불균형이 심하기 때문에 G-Mean을 통해 민감도와 특이도의 균형을 평가한다.

5. 분석 결과 활용 방안

1. 장비 성능 유지 및 고장 방지

- **장비 성능 최적화:** 장비 대여 1개월 전 작동오일 상태를 점검하여 장비가 정상 상태에서 최적의 성능을 유지할 수 있게 한다. 정상적인 오일 상태를 유지하면 장비의 윤활 기능이 원활히 작동해 부품 마모를 줄이고, 전체적인 장비 성능을 향상 시킬 수 있다.
- **고장 방지:** 오일 상태 데이터 분석을 통해 잠재적인 고장 원인을 사전에 발견할 수 있다. 예를 들어, 오일의 점도 변화나 산화 현상이 감지되면, 기계 부품의 마모나 과열이 진행되고 있다는 신호일 수 있다. 이를 조기에 파악해 유지보수를 통해 장비 고장을 방지할 수 있다.

2. 고장 예측 및 사전 대응

- **고장 발생 장비 예측:** 분석된 오일 데이터를 바탕으로 고장이 발생할 가능성이 있는 장비를 예측할 수 있다. 오일의 오염도, 점도 변화, 물 혼입 등 주요 변수를 통해 장비의 이상 상태를 조기에 감지하고 고장 가능성이 높은 장비를 미리 예측하여 대비할 수 있다.
- **고장 전 사전 유지보수:** 고장 위험이 감지된 장비에 대해 즉각적인 유지보수를 시행해, 실제 고장 발생 전 문제를 해결할 수 있다. 이를 통해 갑작스러운 장비 고장을 예방하고, 유지보수 비용 및 장비 중단에 따른 손실을 줄일 수 있다.

3. 예방적 유지보수 기반 효율성 증대

- **예방적 유지보수 계획 수립:** 예측 데이터를 기반으로 장비의 유지보수 계획을 미리 수립할 수 있다. 예방적 유지보수는 고장 발생 이후의 수리보다 비용이 적게 들며, 부품 교체나 오일 교환 시기를 정확히 파악하여 불필요한 유지보수를 줄일 수 있다.
- **유지보수 비용 절감:** 장비 고장으로 인해 발생하는 비용보다 사전 유지보수를 통해 문제를 해결함으로써 비용을 절감할 수 있다. 이는 장비 수명을 연장하고 운영 효율성을 극대화하는 데 기여한다.

4. 장비 가동률 향상

- **장비 가동 중단 최소화:** 고장 가능성이 높은 장비를 미리 파악하고 유지보수 시기를 조정함으로써 장비의 비가동 시간을 최소화할 수 있다. 이는 전체 프로젝트 일정에 차질을 줄이고, 장비 활용도를 높여 현장 운영의 효율성을 향상시키는 데 도움이 된다.