**인공지능 기반 전기차 배터리 가격 예측**

**1. 프로젝트 개요**

중고차의 가격은 많은 데이터가 누적되어 있고, 시장에서의 거래 이력도 많아 그 가격의 예측이 충분히 가능하였다. 주행거리와 연식, 그리고 사고 이력 등을 기반으로 그 가격이 평가된다. 그런다 전기차는 그 역사도 아직 짧고, 중고거래가 아직 활발하지 않다. 전기차 가격의 40~50% 정도를 차지하는 전기차 배터리는 향후 recycle 및 reuse를 위해 그 가격의 예측이 무엇보다 필요하고 중요하다.

**2. 프로젝트 목적**

배터리의 성능과 상태에 따라 그 가치가 평가되어야 한다. 이는 단순히 주행거리와 생산 년도 만으로는 평가하기엔 차량 배터리에 영향을 미치는 요인이 너무 많다. 본 프로젝트 에서는 차량 출고가격의 50%가 배터리의 초기 가치라고 정의하고, 배터리 주요 feature와 거래 가격을 기반으로 전기차 배터리 가격 예측 모델을 만들어 보고자 한다.

**3. 프로젝트 내용**

3-1. 주요 feature의 정의 및 데이터 수집

- 배터리의 현재 상태를 특정할 수 있는 주요 feature를 아래와 같의 정의한다

|  |  |
| --- | --- |
| 항목 | 설명 |
| 차량 모델 | 차량 모델명 |
| 차량 연식 | 차량의 연식 |
| 차량 등록일 | 차량 등록원부 상 등록일자 |
| 배터리 팩 용량 | 배터리 팩의 크기 |
| 배터리 셀 개수 | 배터리 셀 개수 |
| 배터리 구조 | 배터리 셀 배열 구조 |
| 차량 옵션 | 상세 옵션 |
| 출고 가격 | 출고 가격 |
| 공인 연비 | 공인 연비 |
| 주행 거리 | 누적 주행 거리 (km) |
| 셀 균일도 | 각 셀의 상태 |
| SoH | 배터리 퇴화 상태 |
| 평가연비 | 누적 평균 연비 |
| F-P1 | 미공개 (패턴관련1) |
| F-P2 | 미공개 (패턴관련2) |
| F-S1 | 미공개 (산술관련1) |
| F-S2 | 미공개 (산술관련2) |

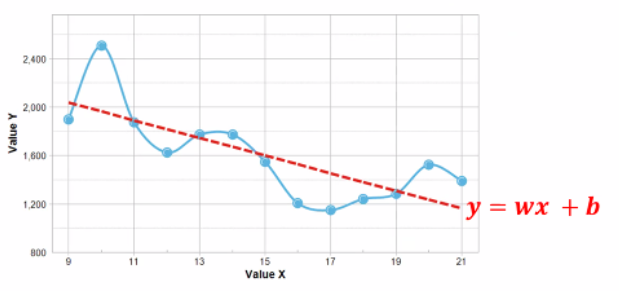
- 학습 데이터는 관련 회사의 차량 데이터 및 차량정보 API를 통해 스크래핑된 기반으로 진행하였다.

보안상의 이유로 본 문서 및 프로젝트 결과서에는 첨부하지 못하였으며, 본 보고서에는 임의의 data로 모델 구축을 하는 과정으로 작성하였다.

3-2. Training 계획

1. 다중 선형 회귀.

다중 선형 회귀 분석을 이용하여 모델을 학습해 보고자 하였다. 여러 데이터의 특성(feature)를 사용하여 예측 모델을 만드는 분석 모델이 가장 적절하다고 판단하였기 때문이다.



[그림 1] 선형 회귀 분석

다중 선형 회귀의 가설 함수는 다음과 같다.

(Feature의 수만큼 weight와 x 설정)

1. Preprocessing and feature Scaling  
   다양한 feature를 적용하는 만큼 각 feature의 type도 다르고 스케일 차이가 크다. 즉, feature의 값의 범위(range)를 동일하게 맞춰주어야 한다. 그렇지 않을 경우 모델들이 제대로 학습되지 않을 가능성이 있다. Type은 모두 숫자로 변환하고, feature들이 평균0, 표준편차 1 값이 되도록 스케일 조정을 하여 모델이 학습이 잘 될 수 있도록 전처리를 해줘야 한다.
2. Regularization  
   모델이 과대적합이 되지 않도록, 모델을 단순화 시켜야 한다. 이를 위해 weight를 0 또는 0에 가깝게 만드는 규제 적용이 필요하다.

L1규제와 L2규제가 있으며, L2 규제는 전체 가중치 크기를 0에 가깝게 만들어 모델을 단순화 시키는 것으로, 모든 weight를 같은 비율로 감소시킨다. 이는 weight가 클수록 더 크게 감소하게 만들어 준다.

**4. 주요 사용기술**

- 언어 : Python

- 모델 : 다중 선형 회귀

- 프레임워크 : Tensorflow

**5. 주요 소스코드**

- Git hub source code 참조  
https://github.com/booyonce/AI\_Mid