

토지 가치 예측 프로젝트

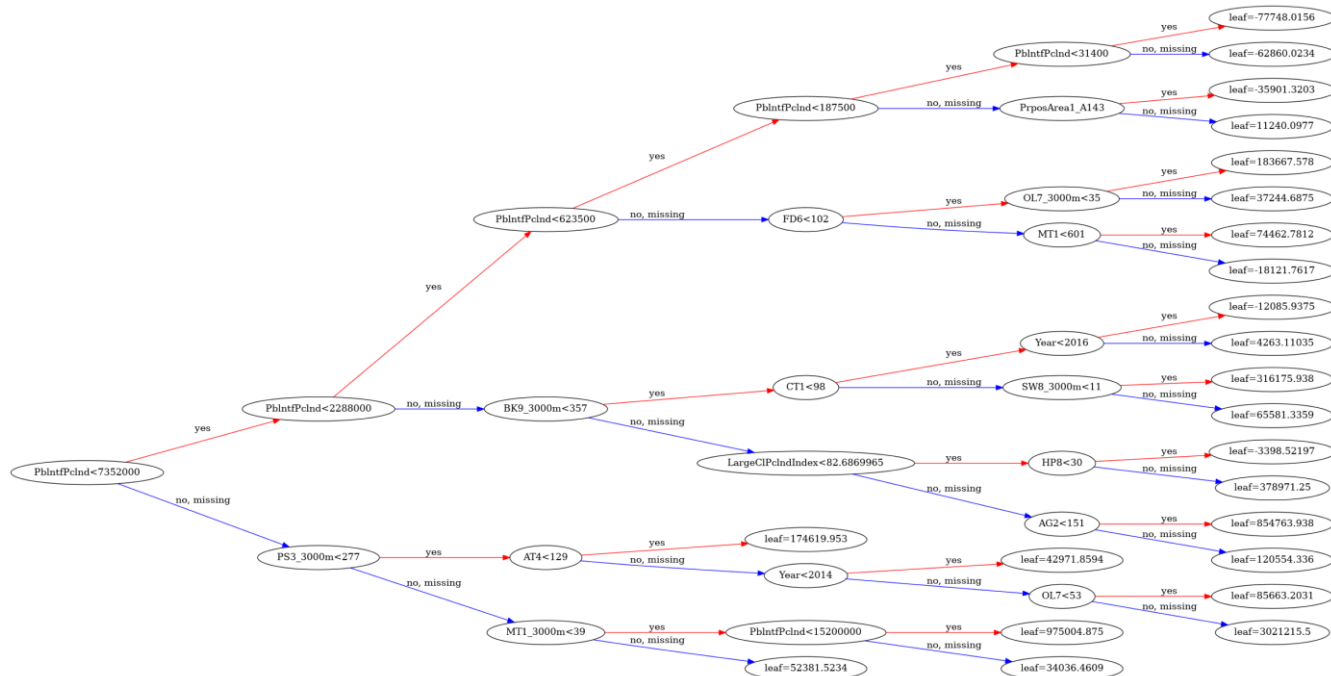
토지 가치 예측을 위한 AI 기반 모델 개발



진행 상황

1. 학습 데이터셋 추가
2. GridSearchCV를 이용하여 하이퍼파라미터 튜닝 (아직까진 유의미한 성능 향상은 없었음)
3. ChatGPT 결정 트리 기반 가치판단

ChatGPT 결정 트리 기반 가치판단



ChatGPT 결정 트리 기반 가치판단

== 결정트리 ==

Tree 0:

```
0:[PblntfPclnd<4728000] yes=1,no=2,missing=2
  1:[PblntfPclnd<1592000] yes=3,no=4,missing=4
    3:[PblntfPclnd<419000] yes=7,no=8,missing=8
      7:[PblntfPclnd<131000] yes=15,no=16,missing=16
        15:[MT1_3000m<31] yes=25,no=26,missing=26
          25:leaf=-197736.266
          26:leaf=-3778.89014
        16:[PblntfPclnd<326000] yes=27,no=28,missing=28
          27:leaf=-120212.977
          28:leaf=-47671.3516
      8:[PblntfPclnd<623500] yes=17,no=18,missing=18
        17:[CE7<59] yes=29,no=30,missing=30
          29:leaf=69068.4141
          30:leaf=-22785.1016
        18:[AG2_500m<20] yes=31,no=32,missing=32
          31:leaf=80791.4375
          32:leaf=260245.531
    4:[PblntfPclnd<2288000] yes=9,no=10,missing=10
      9:[PPI<98.3199997] yes=19,no=20,missing=20
        19:leaf=60958.2812
        20:[LandUsePlans_UQV700(3)<1] yes=33,no=34,missing=34
          33:leaf=427507.938
          34:leaf=2423.22998
```

ChatGPT 결정 트리 기반 가치판단

토지 가격은 여러 요소에 의해 측정되지만 이 경우에는 주어진 결정트리를 통해 예측된 가격입니다. 결정 트리는 일련의 질문을 통해 해결하려는 문제에 대한 결정을 내리는 모델입니다. 이 경우에는, 주어진 정보에 따라 토지 가치를 평가하는데 사용됩니다.

토지 항목은 개별공시지가에 따라 분류되었고, 해당 값은 3,757,000원으로 주어진되었습니다. 이는 처음의 분류기준인 4,728,000원보다 작고, 다음 분류 기준인 1,592,000원보다 크므로, 두 번째 노드로 넘어갑니다. 여기에서는 토지 항목이 2,288,000원보다 작음에 따라, 9번 노드로 이동합니다. 9번 노드에서는 생산자물가지수(PPI)가 사정된 분할 값인 98.32보다 크기 때문에, 20번 노드로 이동합니다.

20번 노드에서는 이용계획 'LandUsePlans_UQV700(3)'의 값이 1보다 작아야 하지만, 우리 데이터에서 이 필드는 명시되어 있지 않으므로 이 결정은 불확실합니다. 그러므로 34번 노드로 이동하여 2,423.22998의 leaf 값을 얻게 됩니다.

따라서, 감정평가사는 제곱미터당 가격이 약 2,423.23원이라고 평가했습니다. 이 토지의 면적이 291.3m²이므로, 최종적으로 평가된 토지 가격은 약 706,131,362.8원 (2,423.23원/m² x 291.3m²)입니다. 이렇게 결과적으로 해당 토지의 가치를 평가하는 방식이 결정트리를 이용한 가치 예측 방식입니다.

TODO:

1. 학습 데이터셋 양 늘리기
2. 모델 개선 (하이퍼파라미터 튜닝)
3. 결정 트리 내용을 단순화 할 방안이 있는지 고민해보려고 함
4. 주변의 비슷한 토지의 실거래 내용과의 비교를 통한 분석도 해볼 예정