나라장터 적격가 예측 AI 결과 보고서

나라장터에 낙찰 완료된 데이터를 수집하여 AI 모델을 만들어 적격가를 예측하는 프로그램을 개발하였음

위 개발한 프로그램으로 나온 성능을 요약하면 다음과 같음

Test1. 학습 데이터 656건으로 학습 후

시뮬레이션 200건을 9번 반복한 결과는 다음 표와 같음

|  |  |  |
| --- | --- | --- |
| 건수 | 낙찰된 경우 | 낙찰안된 경우 |
| 200 | 39 | 161 |
| 200 | 37 | 163 |
| 200 | 46 | 154 |
| 200 | 46 | 154 |
| 200 | 45 | 155 |
| 200 | 4. | 160 |
| 200 | 44 | 156 |
| 200 | 38 | 162 |
| 200 | 41 | 159 |

낙찰된 확률은 20.88%로 즉 5번 정도 입찰을 하면 1번은 낙찰되는 결과를 얻을 수 있었음

Test2. Test1에서 125건을 추가로 수집하여 학습 데이터 781건으로 학습 후

시뮬레이션 200건을 9번 반복한 결과는 다음 표와 같음

|  |  |  |
| --- | --- | --- |
| 건수 | 낙찰된 경우 | 낙찰안된 경우 |
| 200 | 44 | 156 |
| 200 | 44 | 156 |
| 200 | 49 | 151 |
| 200 | 34 | 166 |
| 200 | 49 | 151 |
| 200 | 39 | 161 |
| 200 | 44 | 156 |
| 200 | 46 | 154 |
| 200 | 39 | 161 |

낙찰된 확률은 21.55%로 test1보다 0.7% 증가했으나 시뮬레이션 횟수를 반복하면 Test1과 비슷한 정확도가 나올 것으로 예상 됨… 즉 학습 데이터가 늘어나도 유의미한 성능 향상은 보여지지 않음

Test3. 학습 데이터 1047건으로 학습 후

시뮬레이션 200건을 3번 반복한 결과는 다음 표와 같음

|  |  |  |
| --- | --- | --- |
| 건수 | 낙찰된 경우 | 낙찰안된 경우 |
| 200 | 34 | 166 |
| 200 | 39 | 161 |
| 200 | 43 | 157 |

정확도는 19.33%로 Test1보다 정확도가 더 떨어짐

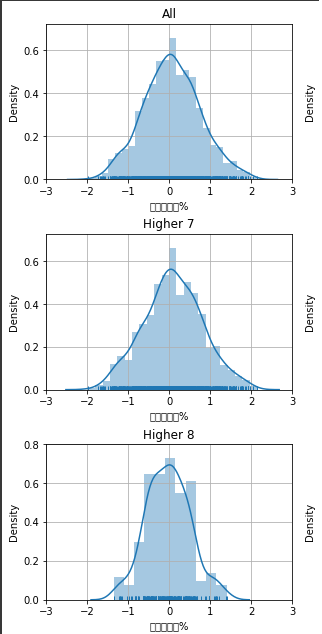
이는 시뮬레이션 횟수가 증가하면 Test1과 비슷한 결과로 수렴할 것으로 예측됨

학습 데이터 개수는 Test1과 비교하면 80%정도 증가하였으나 성능은 그대로인 것으로 보아

기초금액, 낙찰하한율, 기초금액기준 상위개수, 예가범위 등 나라장터에서 얻을 수 있는 데이터는 낙찰가와 상관관계가 낮음을 확인할 수 있었음.

본인이 딥러닝을 사용하지 않고 낙찰하한율, 예가범위를 고려하여 1800건을 시뮬레이션 해보았을 때 정확도가 21%정도로 사람과 딥러닝이 예측한 정확도의 차이는 거의 없음

이외 데이터 수집 후 분석하며 발견한 특이사항



1247건을 분석하여 예정가격과 기초금액기준 상위개수의 상관관계를 살펴본 결과

기초금액기준 상위개수가 7개인 경우 기초금액보다 조금 더 높은 금액이 예정금액으로 선정될 가능성이 높으며

기초금액기준 상위개수가 8개인 경우 기초금액보다 조금 더 낮은 금액이 예정금액으로 선정될 가능성이 높았음

이를 참고하여 투찰가를 선택하면 낙찰확률을 조금 더 높일 수 있을 것으로 예상됨