

[전공과목] 데이터 분석 및 활용 재가공 동영상 콘텐츠 선별 모델

2인 팀 프로젝트/ 데이터 수집, 분석/ ML(파이썬)

인사이트 및 결론

방영 요일(평일/주말), 평균 시청 연령(드라마에서만 한정), 방영 후 뉴스 기사 수가 여러 피처 중 재가공 동영상의 조회수와 가장 유의미한 관계가 있음을 확인함.

하지만 결론적으로는 모델의 평균오차값이 114-200으로 다소 큰 편에 속하였음. 때문에 시청자 반응에 대한 예측이 정확하다고 보기는 어려우나, 비교의 용도로는 충분히 활용 가능한 점을 확인함.

배운점

데이터 분석을 진행하기 앞서 연구하기로 한 내용이 얼마나 유의미한 결과를 가져다주고, 산업에 어떤 긍정적인 영향을 줄 수 있는지에 대한 고찰은 매우 중요한 요소임을 배웠습니다. 더불어 결과에 따른 유연한 의사결정을 통해 실제 산업에서의 활용 가능성을 확인할 수 있었던 좋은 경험이었습니다.

분석 배경

1. 빠르고 쉽게 정보를 습득하기에 용이한 '숏츠 콘텐츠'의 수요가 늘어나며, 방송사 또한 TV프로그램 재가공 영상을 클립형으로 생산하고 있음을 확인함.
2. 하지만 모든 콘텐츠를 생산하는 것은 비용적인 한계가 있음. 이 때, 대중들로부터 높은 관심을 받을 것이라 예측되는 콘텐츠를 선별하여 생산하는 것이 효율적이라고 판단함.
3. 이미 방영된 방송 관련 데이터를 활용하여 재가공 콘텐츠의 수요와 밀접한 관련이 있는 요인을 확인하고, 재가공 동영상 조회수에 대한 예측 모델을 개발하기로 함.

