|  |
| --- |
| **「2020 빅콘테스트」데이터 분석 계획서** |

|  |  |  |  |
| --- | --- | --- | --- |
|  | \* 해당란에 체크 표시 | | |
| **참가분야** | □ 혁신아이디어분야 🗹데이터분석분야 | | |
| **세부분야** | □ 퓨처스리그 🗹 챔피언리그 \*데이터분석 분야에 한함 | | |
| **개인/팀여부** | 🗹 개인 □ 팀(구성원 명) | **개인/팀명** | 정경현 |
| **대표ID** | b990208@naver.com | | |

※ **5장 내외로** 목차는 준수하여 자유롭게 작성

|  |  |
| --- | --- |
| **분석 주제명** | 최적화편성표제시 |
| **분석 배경** | 홈쇼핑 특성 상 대개 TV매체를 통하여 접근하게 되고 홈쇼핑에서 판매되는 물품은 우리가 일상생활에 많이 사용하는 물품들, 생활과 밀접하게 관련된 물품이 다수이다. 또한, 외부요인(기상상황, 요일, 시간대 별)에 따라 구입하려 생각했던 물품들이 있을 거라고 생각됩니다. 그런 상황에서 바로 구입할 수 있도록 홈쇼핑을 편성하면 어떨까하는 생각이 들어서 분석을 시행하게 되었습니다. |
| **분석 내용**  **요약** | 아직 데이터탐색정도의 수준만 하고있는 상황이라 분석을 진행하지 못하였습니다, |
| **분석방법 및**  **계획** | -시간대에 따른 시청률의 상관관계가 있는지 파악  -외부요인(기상상황, 방송일시, 방송시간 등)에 따른 상품군 별 판매실적에 차이파악  앞서 계획한 분석에서 상품군 별 판매실적에 차이가 유의하다면  -기상상황에 따른 상품군 별 판매실적 차이파악  -방송일시(요일)에 따른 상품군 별 판매실적 차이파악  -방송시간에 따른 상품군 별 판매실적 차이파악  앞서 계획한 분석에서 유의미한 차이를 나타낸 요인을 포함한 모델 생성  (생성시 트레이닝셋과 테스트셋으로 구분)  -트레이닝셋으로 모델 생성  -테스트셋으로 정확도검사  정확도가 기준이상 나온다면  -분석결과와 모델링결과 토대로 기상상태, 요일, 시간별로 편성표 작성  추가사용데이터: 기상청’날씨누리’-기상데이터(지점, 지점명, 일시, 평균, 최저, 최고기온, 강수계속시간,1시간 최다강수량, 최대풍속, 최대풍속시간, 평균상대습도, 1시간 최다일사량, 합계일사량, 일 최심신적설, 일 최심적설, 안개 계속시간)  정확한 분석기법은 아직 정하지 못했지만 시계열, 회귀분석, 의사결정나무, 앙상블(배깅, 부스팅, 랜덤포레스트) 등을 이용하여 데이터 형식에 맞고 예측력이 좋은 분석기법을 선정하려고 합니다.  시각화방법은 여러가지 시각화를 보여주려는 분석결과에 맞게 기본그래프 이외에도 결과를 직접 구체적으로 탐색할 수 있도록 인포그래픽으로 나타내보려고 합니다, |
| **분석결과 활용 및 시사점** | 외부요인(기상변화, 요일, 계절)등에 따른 시간대별 시청률, 취급액의 변화에 따라 의미있는 차이가 나타나는 부분에 특정 상품군을 편성하고 요일별로 상품군의 편성에 의미가 있는 요일에 편성할 수 있도록 상품군에 따른 매출을 최대화 할 수 있는 요일, 시간대, 계절, 기상상태 등을 메뉴얼로 짜서 활용할 수 있도록 만들어보려고 합니다.  좀 더 구체적인 분석을 위해 구매한 고객들의 나이대, 성별 등의 고객정보를 알 수 있으면 더 좋았을 것 같습니다. 후에 더 구체적인 분석과 적용을 할 때 이용할 수 있으면 좋을 것 같습니다. |
|  |  |

※ 제출자료는 최종 출품작 평가시 활용될 수 있음