

KURLY HACK FESTA 2022

CHALLENGE THE CODE / CHANGE YOUR KURLY

[과제 계획서] – 총 분량이 5매를 넘지 않도록 작성해주세요.

1. 참가자(팀/개인) 소개

팀 이름	데덕무침
팀원 정보	<p>이름/나이/역할/관련 이력 및 공모전 수상경험 등 특별한 점 소개</p> <ul style="list-style-type: none"> ● 장두혁/28/석사 - 역할 : 대형 E커머스 가격 크롤링 DB화, 머신러닝 분석 총괄 E mail : justin214@naver.com ● 박진용/26/학사 - 역할 : 네이버 온라인 스토어 가격 크롤링 데이터 DB화 및 테스트 E mail : mej1406@gmail.com ● 정병현/26/학사 - 역할 : 네이버 온라인 스토어 가격 크롤링 데이터 DB화 및 테스트 E mail : idealive7@gmail.com ● 강지윤/24/학사 - 역할 : 웹 프로그래밍 및 서버 E mail : xadamantiumx@naver.com
	2번 상품별 적정가를 모니터링할 수 있는 해결방안 제안
프로젝트 명	가격 선넘네?

1-1. 팀을 자유롭게 소개해주세요.

신선한 식품을 새벽에 배송한다는 메시지인 '샛별 배송'을 담고 마켓컬리는 2015년도에 등장했습니다. 코로나로 인한 다이어트 및 헬스와 사회 상의 변화로 인한 친환경적인 관심들이 다양한 연령층에서 나타나기 시작하면서 새벽 배송이라는 키워드와 더불어 이런 관심들을 충족시킬 수 있는 서비스를 가진 마켓컬리는 가파르게 성장했습니다.

그러나 저희 팀원들은 식품 구매 시 품질이 가장 중요하게 여기기에 마켓컬리를 애용했으나 식품이 아닌 타 제품 구매 시 다른 e커머스 기업들 대비 가격 경쟁력에 대해 의문을 표하고 있었습니다. 따라서 품질이 보장된 신선 식품 및 다양한 종류의 제품들을 새벽 배송으로 빠르게 만날 수 있다는 장점과 더불어 소비자로서 부담을 가지게 하는 가격 경쟁력 문제를 해결한다면 마켓컬리는 새벽 배송만이 아니라 배송 업계 전체에서 선두 주자로 성장할 것이라고 판단해 Kurly Hack Festa 2022에 참가했습니다.

저희 팀은 한성대학교의 isyslab 연구실 연구원들과 교내 캡스톤 프로젝트 팀으로 연결된 지인으로 장두혁(석사), 정병현(학사), 박진용(학사), 강지윤(학사)로 구성되어 있습니다. 연구실 연구 과제 및 프로젝트들을 함께 하면서 문제를 해결하고자 하는 의지와 다른 의견을 받아들이는 유연한 사고, 데이터 분석에 대한 지향하는 목표와 생각하는 방향이 잘 통하게 되어 금방 친해지면서 팀원으로 함께 활동하게 되었습니다.

모두 데이터 관련 분야에 큰 관심이 있어 데이터 + 덕후라는 말을 줄인 데덕과 짧은 기간 안에 금방 잘 어우러진 팀이라는 의미로 무침을 합하여 더덕무침과 어감이 비슷한 "데덕무침"으로 팀 이름을 결정했습니다.

2. 프로젝트 소개

2-1. 개발하고자 하는 서비스를 간략하게 소개해주세요.

개발하고자 하는 서비스의 기능, 그리고 해당 서비스가 왜 필요한 것인지, 어떠한 방식으로 문제를 해결할 것인지 등에 대해 요약하여 설명해 주세요.

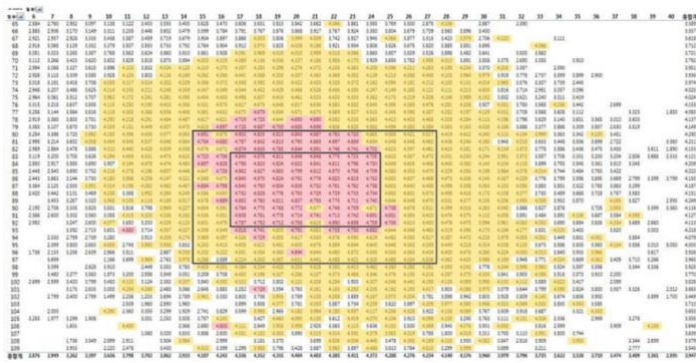
마켓컬리의 많은 상품들은 가격을 결정하는데 있어서 적절하게 형성 및 유지로 고객이 좋은 상품을 적당한 가격에 구매할 수 있도록 유지하고 있습니다. 하지만 타 커머스에 판매되는 동일한 상품 경우와 '사과', '김치'와 같은 하나의 카테고리 내에서 가격 편차가 큰 상품의 경우는 적정가를 찾는 것에 어려움이 있습니다. 따라서 그런 가격 편차가 큰 상품들의 적정가를 어떻게 찾을지에 대해 고민해보았습니다.

[1. 상품별 판매 적정가 추천 및 모니터링 대시보드 서비스 - 요약]

저희 팀은 1시간 단위로 다양한 타 e커머스들의 웹 크롤링을 통해 수집한 가격 정보와 상품별 주요 특징 기준으로 마켓컬리의 공급가와 타 e커머스 상품이 분석 결과로 현 시장 흐름 분석 결과와 상품별 적정가를 상품 현황에 따라 정할 수 있도록 웹 대시보드를 통해 제시하여 모니터링하며 서포트 하고자 합니다.

분석된 정보에 따라 히트맵, 분포도, 그래프로 해당 상품의 가격흐름에 따라 경쟁사보다는 가격이 저렴하면서 마켓컬리의 프리미엄 이미지를 유지할 수 있는 적절한 가격이 될 수 있는 기준을 정하고자 합니다.

[2. 실시간 상품별 가격 현황 분석과 시각화로 모니터링 기능]



가격 조정 유무에 있어 중요도 가장 높은 특징 항목(Y축)과 판매가/공급가(X축)를 기준으로 ABC 분석한 결과를 웹 대시보드의 시각화 그래프로 제공해, 모든 상품별 가격 현황 데이터를 직관적으로 분석할 수 있게 합니다.

기존 다양한 태블로 같은 시각화 툴도 있지만, 상품별 적정가 분석을 위한 실시간 시각화 분포 그래프와 마진율 등 분석 결과와 적정가 추천을 통합적으로 제공하므로써, 마켓컬리의 적정가 최종 가격 결정에 있어서 도움이 되고자 합니다.

<돼지출하목적의 등지방 두께 범위(상품 특징)에 따른 가격 ABC분석 분포 그래프 예시>



<ex> 실시간 상품별 적정가 추천 및 모니터링 대시 보드 샘플 예시>

[3. 타 e커머스 판매가 분석 따른 적정가 가격 추천 기능]

실시간의 상품의 상품가와 마켓컬리 공급가를 활용한 분석 분포도를 표현해 판매가 범위, 판매가와 공급가 차이간의 마진율범위와 대형 e커머스(이마트, 홈플러스, 롯데 쇼핑, SSG, 쿠팡) 보다 낮은 상품가 가격 범위의 교집합을 1시간 단위로 상품의 적정가를 추천하고자 합니다.

이를 통해, 상품에 대해서 대형 e커머스보다 낮은 가격으로 소비자의 부담을 줄이고, 품질이 보장된 마켓컬리 상품으로 공급가와 판매가 차이를 유지해 최저가가 아닌, 타 e커머스보다 경쟁력 있으며, 프리미엄 이미지를 유지할 수 있는 가격을 제시할 수 있습니다.

최저가가 아닌 품질 보장된 마켓컬리 상품 판매와 마켓컬리의 프리미엄 브랜드 이미지에 맞게 적정가가 되어야한다고 생각했습니다. 저희는 상품별 특징항목과 가격에 따른 ABC분석을 통해 정한 판매가 분류 범위(상위 30%, 상품별 평균 범위, 상위 70%), 타 e커머스의 판매가와 마켓컬리의 공급가 차이(마진율)의 평균치 범위, 대형 e커머스(이마트, 홈플러스, 롯데 쇼핑, SSG, 쿠팡) 판매가 이하 범위 3가지를 활용하여 시간대별로 적정가를 추천 및 모니터링하고자 합니다.

3가지 범위 기준

- 3-1. ABC 분석을 통해 정한 판매가 분류 범위 활용

크롤링한 가격 데이터와 머신러닝 기법으로 선정한 중요도가 가장 높은 픽처 항목 기준으로 ABC 분석을 통해 정한 범위를 상위 30%, 상품별 평균 범위, 상위 70%로 나누어 등급을 분류 범위를 정합니다.

- 1) 상품의 전체적 상품가격 현황 한눈에 파악하기 위해 범위별로 색깔로 분류해, 실시간으로 직관적으로 흐름을 파악하고자 합니다.
- 2) 대형 e커머스의 상품 판매가격들의 ABC분석을 통한 판매가 분류 범위에 속한 등급을 분석하며, 1차적으로 판매가의 상한가와 하한가를 정할 수 있습니다.

- 3-2. 타 e커머스의 판매가와 마켓컬리의 공급가 차이의 평균치 범위 활용

상품에 대한 크롤링한 타 e커머스 판매가와 마켓컬리의 공급가 차이로 상품에 대한 현 시장에서의 전반적인 마진율을 파악해, 마켓컬리의 상품 판매 적절가를 정하는데 서포트 할 수 있습니다.

- 1) ABC분석을 통해 정한 판매가 분류된 등급 범위와 타e커머스 판매가와 마켓컬리의 공급가 차이의 교집합으로 2차적으로 상한가와 하한가 범위를 좁혀 구체화 시킵니다.

- 3-3. 대형 e커머스 판매가 이하 범위 활용

같은 카테고리 상품에 대해서 대형 e커머스 판매가보다 낮지만, 품질이 보장된 마켓컬리 가격으로 소비자 부담을 줄여, 기존 다른 e커머스 고객들의 유입을 통해 마켓컬리 이용 횟수를 증가시킬 수 있습니다.

- 1) ABC분석을 통해 정한 판매가 분류 범위로 정한 상품의 상한가에 대해 최종적으로 범위를 좁힐 수 있습니다.

- 3-4. 상품에 따른 머신러닝 기법을 통한 중요한 피쳐(특징) 선택 알고리즘 활용

과거의 마켓컬리 각 상품 데이터를 통해, 시간대별 가격 조정 유무를 예측 모델을 설계하여 높은 정확도를 도출한 Feature의 중요도가 높은 항목을 선정해 상품별 가격 현황 분석과 시각화와 분석에 따른 각 시간대의 현시장에서 상품 적정가 가격 추천의 정확도를 높이 고자 합니다.

(개발 기간 때, 마켓컬리에서 과거 데이터 제공이 어려울 시에는 과일 카테고리 - 당도 또는 kg / 육류 -kg / 다음 같은 항목을 통해 피쳐를 대체하여 가격대 분포로 분석할 예정)

[4. 사내 협업 툴 Slack의 챗봇을 통한 가격 변동 필요성 알람]

일정 시간마다 크롤링한 상품가의 변동과 분석 결과가 이전보다 변동 폭이 정해진 차이보다 클 때, 자동적으로 담당 부서 팀원과 관련 부서에 챗봇을 통한 가격 변동 필요성에 대한 알림을 제공하여, 상품별 미팅이 현황에 따라 모든 팀원이 인지할 수 있도록 제공합니다.

2-2. 해당 과제를 선택한 이유와 관심을 갖게 된 계기를 말씀해주세요.

평소 쿨리 서비스를 이용해 보셨다면 사용 경험을 살려 말씀해 주셔도 좋고, 그렇지 않다면 평소 관심사 및 관심분야와 연관해 설명해 주셔도 좋아요. 평소 갖고 계셨던 문제의식을 통해 해당 과제를 왜 해결하고자 하는지 자유롭게 말씀해주세요.

마켓 쿨리를 이용할 때 소비자로서 제품을 구매할 때의 가격 경쟁력에 부담을 느껴 가격 책정에 대해 관심을 갖고 해커 톤의 2번(상품별 적정가를 모니터링할 수 있는 해결방안 제안)주제로 참가하게 되었습니다. 물가 인상의 여파로 배달 음식의 가격과 배달비가 증가하면서 기존에 배달 서비스를 자주 이용하던 고객들의 배달 음식 수요가 감소하고, 동시에 건강 및 다이어트에 대한 관심이 e-커머스에서 식재료를 직접 구매하는 고객을 증가시키고 있습니다. 고객들은 조금이라도 더 저렴하면서 퀄리티 좋은 상품을 구매하기 위해 여러 e-커머스 홈페이지를 비교해가며 구매합니다.

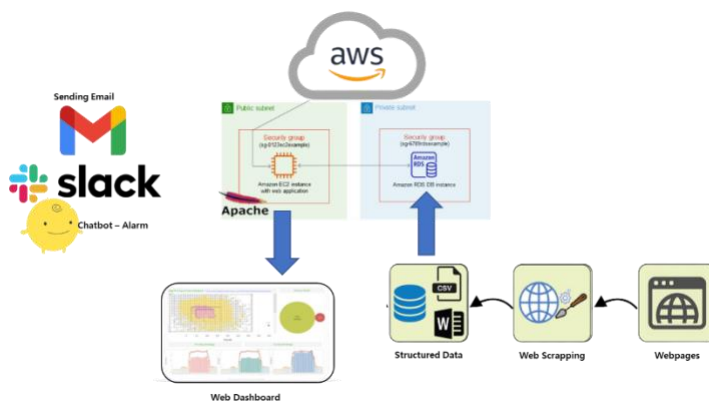
그런 고객들의 요구에 맞춰 대다수의 e-커머스 기업들은 본인들의 최저가 경쟁력을 광고를 통해 나타내곤 했습니다. 그러나 가격 결정이 단순하게 최저가로만 해결할 간단한 문제가 아니라 세부적인 품목마다 다른 부분이 있어 끊임없이 더 좋은 해결방법을 찾아볼 수 있는 분야로 판단했고, 가격 결정에 대한 소비자와 공급자 모두의 의문점을 마켓컬리에서 해결해내고자 해당 과제를 선택하게 되었습니다.

3. 프로젝트 개발 계획일

3-1. 서비스 아이디어를 구현하기 위한 개발 계획을 설명해주세요.

제안하는 아이디어를 구체적으로 어떻게 구현(개발)하고자 하는지, 사용하는 기술과 환경 등의 내용을 담아서 소개해주세요.

(예) 개발 형태(웹/앱), 개발언어, 아키텍처, 플로우 차트 등 자유롭게 계획을 설명해주세요.



<개발 서비스 플로우 차트>



<사용하는 개발 언어>

1. 마켓컬리의 카테고리별 상품에 대해서 네이버 온라인 스토어와 대형 타 e커머스 사이트의 해당 상품의 가격과 상품의 정보를 매시간별로 Python언어로 Selenium 라이브러리로 Chrome의 웹드라이버를 통해 웹 크롤링한다
2. 수집된 데이터를 데이터베이스로 구조화 시켜 AWS의 ubuntu기반 EC2 인스턴스에 연동된 RDS에 저장한다.
3. Python언어로 XGBoost, RandomForest, CatBoost 등을 활용해 상품별 시간대별 마켓컬리의 가격 조정 유무를 예측 모델을 설계하여 가격 조정 유무 예측에 대한 높은 정확도를 도출한 Feature들의 타겟에 대해서 중요도가 높은 특징 항목을 선정한다.

선정한 특징 항목(y)과 공급가/판매가(x)에 대해서 ABC분석기반으로 등급별 분포도 그래프를 산출한다.

4. EC2인스턴스에서 파이썬의 ABC 분석을 통해 정한 판매가 분류 범위, 타 e커머스의 판매가와 마켓컬리의 공급가 차이의 마진을 평균치 범위, 대형 e커머스 판매가 이하 범위 교집합을 구하는 알고리즘 프로그램을 실행시킨다.

5. 상품의 추천 적정가와 가격의 분포도와 마진율 분석 등 프로그램 결과물을 apache 웹서버를 통해 JavaScript, Html, php 언어를 활용해 웹형태의 대시보드에 출력한다.
6. 상품별로 과거의 해당 날짜와 시간대 인풋 폼을 클릭하여 해당 시간대의 분포도 그래프와 비교 분석 결과를 볼 수 있다.
7. 자동적으로 담당 부서 팀원과 관련 부서에 Slack 챗봇과 지메일을 통한 가격 변동 필요성에 대한 알림을 제공하여, 상품별 미팅이 현황에 따라 모든 팀원이 인지할 수 있도록 제공한다.

4. 프로젝트 기대 효과

4-1. 제안해 주신 서비스 혹은 아이디어의 차별화 전략 및 기대 효과를 말씀해주세요.

제안해 주신 서비스 혹은 아이디어가 기존의 서비스와 비교하였을 때 차별화되는 점을 중심으로 기대효과와 함께 설명해주세요. (예시) 어떤 타겟(ex. 컬리 고객, 물류센터 임직원 등)을 대상으로 어떠한 비즈니스 지표를 개선하여(ex. MAU 증가, 재구매율 증가, 풀필먼트 프로세스 정확도, 생산성 개선 등) 어떤 비즈니스 임팩트를 줄 수 있는지(ex. 매출 증대, 비용 절감 등) 설명해주세요

[타e커머스 고객들의 유입을 통한 매출 증대]

가격 크롤링을 통해 실시간 상품현황에 따른 상품이 분포 자동 분석과 타 대형 e 커머스 상품보다 저렴한 가격의 좋은 퀄리티 상품을 제공하여, 기존의 마켓컬리의 고객 뿐만 아니라 다른 이커머스를 사용하던 고객들이 경쟁사보다 더 저렴한 가격에 동일 물품을 동일한 퀄리티로 구매할 수 있게 되며 재구매율이 증가하고 장기 고객을 유치하는데 큰 도움이 되어 매출 증대를 기대해 볼 수 있을 것 입니다.

[상품의 공급가와 판매가 차이를 통한 매출 증대]

대형 e 커머스의 가격보다 낮지만, 질 좋은 상품에 대해 마켓컬리 브랜드 이미지에 걸맞는 프리미엄 상품 가격으로 매출 증대할 수 있습니다. 또한 최저가나 적절하지 못한 가격 책정은 기업의 브랜드를 훼손할 수 있다는 점에서 적절한 책정 기준선을 정해 모니터링 하는 일은 매우 중요하다고 생각이 들어 신뢰도 높은 기업이라는 이미지를 심어 줄 수 있을 것입니다.

[타 e 커머스의 가격과의 비교를 통한 합리성]

과거의 마켓 컬리 상품을 참고해 분석하기도 하지만 타 e 커머스의 가격을 포함해 분석함으로써 가격 결정에 대한 합리성 및 타당성이 부여될 수 있습니다.

[사내 협업 툴을 통한 알림]

변동성이 심하게 이뤄짐에 대한 모니터링을 알림을 통해 공유할 수 있어 모든 사내 직원들이 항상 주의하고 있지 않아도 소수의 마켓컬리의 직원의 관리로 상품에 대해서 모니터링할 수 있습니다.