

송예설yeseolsong

## who am i?



2012.03-2018.02

고려대학교 졸업 영어영문학과, 경영학과 (이중전공)

서울대학교 빅데이터 연구원 비즈니스 애널리틱스 교육 수료

**ず**하나금용티아이

**하나금융티아이** 금투지원팀 | 프로젝트 인턴

2017.06-2018.02

2017.12-2018.02

2016.11-2017.05

Growth Hackers

GrowthHackers in SNU PR 팀 | 팀장

2017.07-2018.06

Kakao Corp. 광고사업부문 비즈데이터분석팀 | 어시스턴트

2018.03-현재 재직중

**마이리얼트립** 사업개발팀 | 인턴 myreáltrip

"데이터지능의힘을 믿습니다"

스타트업에서의 인턴 경험을 통해 데이터 바다 속에서 빛나는 진주를 찾아내는 능력의 중요성을 깨달았습니다.

졸업 후 친구들 모두가 취업 전선에 뛰어들 때 무모하지만 설레는 목표를 가지고 첫 걸음을 내딛었습니다.

분석 교육 프로그램 수료, 분석 학회 활동, 자격증 취득, 인턴 활동을 통해 데이터 지능을 기르기 위해 온 열정을 쏟았습니다.

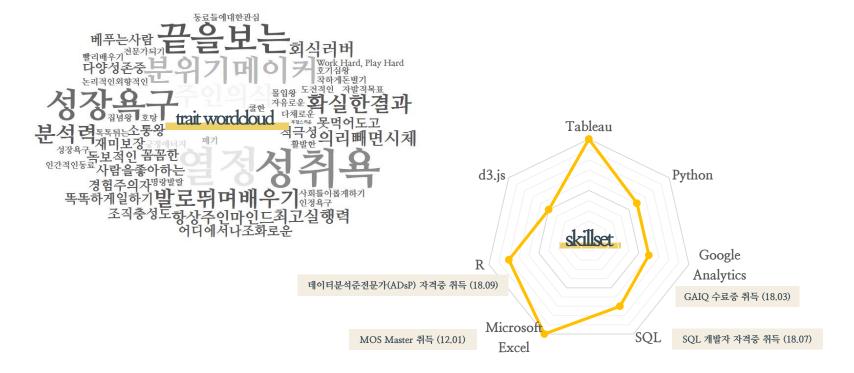
#### "데이터를읽어내는서비스기획자를꿈꿉니다"

그렇다고 제가 분석모델링을 하는 데이터 사이언티스트를 꿈꾸는 것은 아닙니다.

사람에 대한 큰 애정으로, 사람을 이롭게 할 서비스를 만드는 서비스 기획자를 꿈꿔왔습니다.

다만 데이터를 기반으로 의사결정을 하는 것은 비단 데이터 분석가만의 업무 영역은 아니라고 생각합니다.

도메인, 서비스, 유저에 대한 이해를 바탕으로 서비스에 유의미 한 데이터를 선별하고, 검증 가능한 인사이트를 기반으로 서비 스를 만들어 내는 서비스 기획자가 되겠습니다.



# data analysis portfolio

kakao

hanaTI

bigdata hackerthon

growthhackers

## kakao

광고사업부문비즈데이터분석팀 어시스턴트(18,03-현재재직중)

주업무및역할

태블로대시보드기획및제작

대시보드운영및보수

통계분석

#### 사용기술및프로그램





#### 느낀점

의미있는지표분석을위해서는서비스와상품에대 한충분한공부가 필수적이다!

대시보드사용자의분석니즈를 명확하게 이해하고, 사용자 입장이 되어꼭 필요한 가능만을 구현하자.

목표없는대시보드는 버려지기만한다. 충분한기획 과지속적인 피드백반영이 필수적이다.



Click me for further information!

#### 대시보드UX/UI기획





다양한 팀의 요청을 받고 충분한 사전 인터뷰를 통해 대상과 목표를 확실히 이해한 후, 사용자 UX 시나리오를 구상합니다.

#### 그 후에는

- 사용목적에 적합한 그래프 및 차트 선정
- 전달력 극대화를 위한 시각적 퀄리티
- 분석을 위한 인터렉션 기능

등 을 고려하여 **대시보드 UI 기획안**을 제작합 니다.

이를 기반으로 **사용자의 피드백을 받고 반영** 하는 과정을 반복합니다.

#### 분석용인터렉티브대시보드제작및보수





실무자들이 트렌드를 이해하고 분석하기 위해 **분석용 대시보드**를 요청합니다.

요청 실무자들의 니즈를 정확히 파악하고, 그에 맞는 분석을 용이하게 하기 위해 다양한 **인터레 티브 기능을 구현**합니다.

#### 프론트 작업을 담당한 대시보드:

- 광고상품별(톡광고, 검색광고, 로컬광고 등) KPI 관리 대시보드
- 영업팀, 영업담당자별 광고주/대행사/캠페인 KPI 관리 대시보드
- 광고주 및 대행사 이탈 탐지용 대시보드

#### 모니터링용대시보드제작및보수





임원 및 조직장들이 카카오 플랫폼의 다양한 서비스에서 나오는 전체적인 광고매출을 보 기 위한 **모니터링용 대시보드**를 제작합니다.

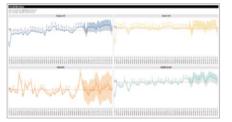
각 플랫폼의 **목표매출 및 달성여부 데이터를 관리**합니다.

사무실에 배치된 **TV 및 모바일용 대시보**드를 제작하여 임직원이 자주 데이터를 들여다볼 수 있도록 하는 것이 저희 팀의 목표입니다.

## 프론트 작업을 담당한 대시보드:

- 카카오 광고통합 목표매출 및 예상매출 관리 대시보드
- TV, Mobile용 일별 매출 대시보드
- 카카오 서비스별 매출 관리 대시보드

## 통계분석





서비스 의사결정을 돕기 위한 **통계분석**을 진행 합니다.

R프로그램을 사용하여 광고매출 시계열 데이터를 통해 각 광고상품에 대한 매출을 예측합니다.

또한 검색데이터를 통해 카테고리별 쿼리수가 급상승하는 키워드들을 **시각적으로 표현**합니다.

연령층별 유저들의 성향을 나타내는 키워드들에 대한 **테스트마이닝을 수행**하였고, 이를 기반으 로 신규검색광고상품 아이디어를 제안했습니다.

**태블로 프로그램에 R과 엑셀을 연동**시킴으로써 통계분석의 결과를 대시보드에 구현하였습니다.

# **\***하나금융티아이

**금투지원팀** 프로젝트인턴(17.12-1802)

주업무및역할

데이터추출및전처리(SQL)

분석모델링(R)

결과보고서작성

사용기술및프로그램



ORACLE'

느낀점

"Carbage In, Carbage Out." -분석에 있어서 전치리는 매우 중요하다

분석을 위한 분석이 아닌, 확실한 문제의식을 가지고 데이터에 접근하자!

분석코드를칠수있는프로그래밍실력보다, 쓸모있는인사이트를찾아내는데기반이되는도메 인자식이분석에더중요하다.



Click me for further information!

#### 문제정의및분석목표



#### 문제 정의:

- 주식상품을 분류하는데 산업, 거래시장 등 단순한 분류기준만이 존재
- 개인투자자의 설문답변에 의존하는 투자 자의 성향 파악 시스템
- 개인화되어있지 않은 추천 주식상품

#### 분석 목표:

- (1) 산업, 거래시장과 같은 단순한 분류기준 이 아닌, 주식상품의 복합적인 시계열 특 성을 고려하는 군집 분석
- (2) 개인투자자의 포트폴리오에 대한 성향 분석
- (3) 포트폴리오 성향에 따른 개인화 추천 알 고리즘 구현

## 데이터수집및전처리



#### 데이터 수집:

하나금융투자 원장 데이터베이스

- 주식상품 데이터 (종가, 산업 등)
- 개인투자자 주문내역 데이터 (상품, 빈도 등)
- 개인투자자 인구통계 및 설문답변 데이터

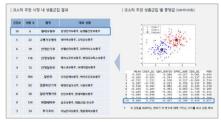
#### 변수 선택 및 생성:

- 장세흐름, 성장모멘텀, 기업가치측면에서 개 인투자자가 주식상품 매수 시 고려하는 기준
- 기준가의 평균, 변화량의 분산 및 첨도, 상장 주식수 대비 거래량 등

#### 이상치 처리:

 액면분할/병합, 자본감소 기업의 경우, 변화 전 가격으로 고려

#### 분석모델링





#### 주식상품 군집분석 모델링:

- 주식상품의 시계열 특성을 나타내는 변수 들에 대한 군집분석 실행
- Hierarchical, k-means, Gaussian mixture
- 해석이 용이한 k-means 선택
- 각 군집의 centeroid가 각 변수에 가지는 수치를 통해 군집 특성 파악

#### 개인투자자 성향분석 모델링:

- 개인투자자가 매매한 상품군집의 특성을 통해 성향을 파악하기 위한 Topic Modeling LDA 알고리즘 사용
- (1) 각 토픽을 구성하는 상품군집 특성 +
   (2) 토픽에 해당하는 고객들의 인구통계 및 설문응답 정보를 통해 성향유형별 인구 통계특성 추론 가능

## 분석결과





## 군집 분석 결과:

- 코스피, 코스닥, 코넥스, ETF, ETN, ELW, 총 6개의 거래시장에서 6963개의 개별상품을 41 개의 군집으로 축소
- 황제우량주, 일반저평가주, 투기주의 등 각 군 집의 특성을 나타낼 수 있는 레이블 할당

#### 성향 분석 결과:

• 'YOLO형', '꿈꾸는 소시민' 등 총 10개의 투사 성향 유형 도출

#### 추천 알고리즘 결과:

- 고객별 가장 큰 확률을 보이는 성향 파악 후, 선 호확률이 가장 높은 상품군집 파악
- 군집 안에서 예쌍수익률이 가장 높은 상품 추천

# data analysis portfolio

kakao

hanaTI

연평균 변화을 (CAGR) : 데이터의 첫 연도부터 마지막 연도

까지 일정한 비율로 성장(감소했다 고 가정했을 때의 그 성장/감소율 bigdata hackerthon

growthhackers



국토교통부주최로서울시공공데이터를 활용하여주택문제해결아이디어제안

## 사용기술및프로그램





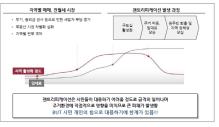
#### 성과





Click me for further information!

### 문제정의및분석목표





#### 문제 정의:

- 젠트리피케이션 과정에서 주거비용 및 임 대료가 상승하면서 원주민이 퇴출되고 지 역 정체성을 상실하며 장기적으로 지역경 제에 악영향을 끼침
- 서울시 공고데이터를 활용하여 젠트리피 케이션을 미리 예측하고 이에 대한 방지 대책 마려

#### 분석 목표:

- 젠트리피케이션 지수 및 유형 예측 모델 구형
- 유형에 따른 방지 대책 수립

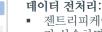
#### 데이터수집 및 전처리



## **데이터 수집 :**■ 서울시 공공데이터 포털

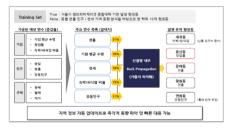
- 지역의 경제적, 문화적, 관광적, 소셜 특성을 나
- 타내는 데이터 수집

  전월세, 전입출, 창업률, 지하철역 유동인구 등
- 웹크롤러를 제작하여 각 행정동과 관련된 연관 검색어 크롤링 (소셜 변수)



- 젠트리피케이션 과정에서 주거비용 및 임대료 가 상승하면서 원주민이 퇴출되고 지역 정체성 을 상실하며 장기적으로 지역경제에 악영향을 끼침
- 서울시 공고데이터를 활용하여 젠트리피케이션
   을 미리 예측하고 이에 대한 방지 대책 마련

#### 분석모델링



유	형 분류 고	정	유형 분류 결과								
각 행정동 연관검색어 웹 크롤링 유형별 키워드 리스트 업		관광 관련 키워드			문화예술 관련 키워드			상업 관련 키워드			
		순위	495	79年4	순위	495	7월드 수	순위	495	7145	
		1	98	7	1	98	7	1	98	7	
연관검색이와			2	498	6	2	서원동	6	2	488	6
키워드 비교			3	삼선통	5	3	삼선종	5	3	상선용	5
유형별 순위			4	상황동	5	4	상형종	5	4	상원문	5
		5	878	5	5	848	5	5	948	5	
유형별 키워드			6	机料等	4	6	NAS	4	6	제화공	4
28	상업경제	259g	7	성수동	3	7	성수동	3	7	성수동	3
立号	시장	시장	8	音楽芸	3	8	988	3	8	888	3
88 88	부동산 데이트	전통 약집		985	2	9	985	2	9	Ves	2
문학	병송	문학	10	HOE	2	10	908	2	10	908	2
연출	용화관	도시관									
역적 거립	99	2621	<ul> <li>시계열 데이터가 아니기 때문에 변화주이는 알기 어려움</li> <li>사람들에게 알려진 지역명의 현정동명이 다를 수 있음 (e.g. '휴대'와 '저희동')</li> </ul>								
28	높게리	長町	<ul> <li>사람들</li> <li>유형별</li> </ul>	에게 얼룩집	: 지역명과 명 본복자의 자의	성용병이 다음 적인 판단에	함구 있음 () 의합	*0.843	VI-81	_	
-			. #88	NAPE	문의자의 자의	적인 원단에	의함				

#### 훈련 데이터 설계:

• 1값으로 서울시에서 공표한 젠트리피케이 션 발생지역 13개 행정동으로 설정, 0값으 로는 반대양상을 보이는 행정동으로 임의 선정

### 토픽 분석 모델링:

- 웹크롤러로 크롤링한 연관검색어에 대해 토픽 부석 실행
- 각 토픽에 해당될 확률이 하나의 변수로 사용

#### 신경망 알고리즘 모델링:

각 변수들의 가중치를 최적화시킬 수 있는
 역전파 신경망 알고리즘을 사용

#### 군집 분석 모델링

## 분석결과

각 등 각 항목마다 절대병의 크기가 데 우다를

· 전트리디웨이션은 어떤 지역이 고구에 비해 얼마나 앤디디션는지가 좋요

 때라서 모든 등이 같은 기간 등안 얼마 나 변화했는지 할 수 있도록 연평군 변 작물로 환산





## 분석 결과:

- 전출, 기업평균수명, 전체, 숙박/외식업 비율, 유동인구가 젠트피피케이션 지수에 가장 큰 영 향을 미침
- 젠트리피케이션 발생 유력 행정동으로 세곡동, 용산동, 문래동, 망원동, 연희동

#### 하계:

- 군집분석의 결과로 이미 발생한 지역들이 골고 루 분포되어있지 않다는 문제
- 이미 젠트리피케이션이 발생한 지역에서 실행하고 있는 젠트리피케이션 대책을 조사하고 평가한 후, 같은 양상을 보이는 행정동에 적용하려 했으나 실패

# Gronten Hackers

비즈니스데이터분석학회 | PR팀 팀장(17,07-18,06)

주업무및역할

대외홍보채널구축및관리

기업프로젝트수주지원

외부행사기획및운영

사용기술및프로그램



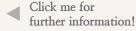


## 성과및느낀점

컨텐츠제작은쉽지않다.우리가전달하려는정보에 대해정확한이해없이는팔로워들의신뢰를떨어트 릴수있다.

페이스북페이지, 웹페이지, 깃헙등다양한홍보채 널을구축하는 것은 쉽지만, 주기적인 보수와 운영이 더욱 중요하다.





#### 페이스북페이지컨텐츠제작및운영



## 학회 페이스북 페이지 컨텐츠 제작 및 운영

- 페이스북 페이지 'GrowthHackers' 운영
- 현재 570명 팔로워 보유중 (임기 전 대비 팔로워수 900% 증가)
- 주 1회 카드뉴스 컨텐츠 업로드
- 매주 카드뉴스 기획을 위한 회의 진행

#### 정기 연재 컨텐츠

- 데이터 뉴비들을 위한 지침서
- 실무자에게 듣는 이야기
- 그로스 해킹
- 위클리 데이터 이슈

## 홍보채널구축및관리





#### 홈페이지 구축 및 관리

- 기업과 잠재적 신규멤버에게 대외적으로 학회에 대해서 알릴 수 있는 웹페이지 구축
- 홈페이지 UX/UI 및 컨텐츠 기획
- GA 연동

#### 깃헙 계정 구축 및 관리

- 학회원들의 학습 및 대외용 포트폴리오 구축을 도모하기 위한 깃헙 계정 구축
- 프로젝트 수주를 위한 대외용 포트폴리오로 서 활용
- 학회 교육 자료 저장소로 활용

#### 기업프로젝트수주지원



#### 프로젝트 제안서 기획 및 제작

- 프로젝트 제안을 위한 문서 제작
- 프로젝트 수주를 위한 아이데이션 회의 진행
- 총 4개의 프로젝트 수주 성공

## 외부행사기획및운영





그로스해커스 4기 모집 설명회

- **대외 행사 기획 및 진행** 학회를 알리기 위한 대외행사 기획
- 연사 초청 및 강연 기획
- 학회를 알리기 위한 외부강연 진행
- 리크루팅 설명회 기획 및 진행



## 송예설yeseolsong

Email songyeseol@gmail.com

Phone 010.6631.5229

Linkedin linkedin.com/in/yeseolsong

Webpage www.songyeseol.com