



국내 최초 빅데이터 연합동아리

# BOAZ 16기 신입회원 모집

## 모집일정

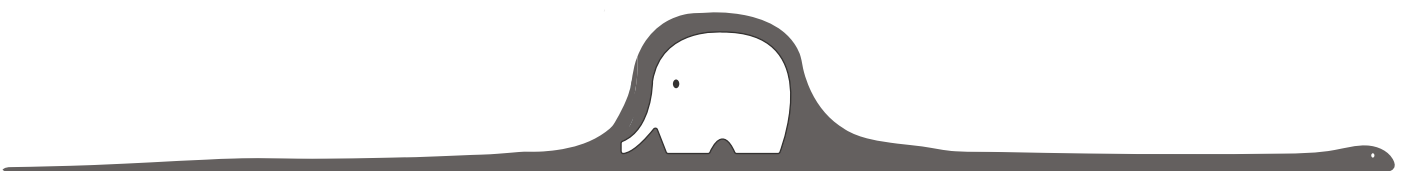
지원서 접수 기간	12.09 - 12.23
1차 합격 발표	12.28
면접 진행 기간	01.02 - 01.03
최종 합격 발표	01.04

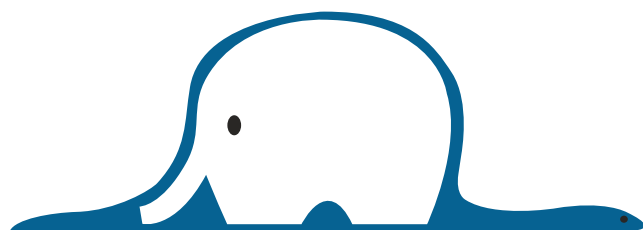
오리엔테이션*	01.07
컨퍼런스*	01.23

\*오리엔테이션과 BOAZ 컨퍼런스 불참 시 합격이 취소됩니다.

## 지원 방법 및 문의

메일	boazbigdata@naver.com
페이스북	<a href="http://fb.com/BOAZbigdata">http://fb.com/BOAZbigdata</a>
인스타그램	@boaz_bigdata





열정과 노력으로  
배움을 나누고  
꿈을 키워갈  
열여섯번째 주인공을  
찾습니다

#### Contact

블로그	<a href="http://blog.naver.com/boazbigdata">http://blog.naver.com/boazbigdata</a>
페이스북	<a href="http://fb.com/BOAZbigdata">http://fb.com/BOAZbigdata</a>
인스타그램	<a href="https://www.instagram.com/boaz_bigdata/">https://www.instagram.com/boaz_bigdata/</a>
대표메일	<a href="mailto:boazbigdata@naver.com">boazbigdata@naver.com</a>

# Index

## Greetings

지도 교수님 인사말

現 대표진 인사말

## Introduction

BOAZ 소개

## Curriculum

정규세션 전체

데이터분석 커리큘럼

데이터시각화 커리큘럼

데이터엔지니어링 커리큘럼

공동세션

## Activity

컨퍼런스

스터디

기타활동

## Appendix

전 활동회원들의 추천글

Q&A



2013. 09. 02	창립
2014. 01. 02	2기 활동 시작
2014. 06. 03	실습사례 발표회
2014. 07. 03	3기 활동 시작
2014. 12. 27	제 1회 BOAZ 빅데이터 컨퍼런스
2015. 01. 08	4기 활동 시작
2015. 07. 02	5기 활동 시작
2015. 07. 11	제 2회 BOAZ 빅데이터 컨퍼런스
2016. 01. 07	6기 활동 시작
2016. 01. 09	제 3회 BOAZ 빅데이터 컨퍼런스
2016. 07. 07	7기 활동 시작
2016. 07. 16	제 4회 BOAZ 빅데이터 컨퍼런스
2017. 01. 05	8기 활동 시작
2017. 01. 15	제 5회 BOAZ 빅데이터 컨퍼런스
2017. 07. 06	9기 활동 시작
2017. 07. 15	제 6회 BOAZ 빅데이터 컨퍼런스
2018. 01. 04	10기 활동 시작
2018. 01. 14	제 7회 BOAZ 빅데이터 컨퍼런스
2018. 07. 05	11기 활동 시작
2018. 07. 14	제 8회 BOAZ 빅데이터 컨퍼런스
2018. 01. 03	12기 활동 시작
2019. 01. 19	제 9회 BOAZ 빅데이터 컨퍼런스
2019. 07. 04	13기 활동 시작
2019. 07. 13	제 10회 BOAZ 빅데이터 컨퍼런스
2020. 01. 02	14기 활동 시작
2020. 01. 18	제 11회 BOAZ 빅데이터 컨퍼런스
2020. 07. 09	15기 활동 시작
2020. 08. 01	제 12회 BOAZ 빅데이터 컨퍼런스

# Greetings

## 지도교수님 인사말



### 이원석 교수님

Yonsei Univ Computer Science, College of Engineering

#### 학력

PhD in Computer Engineering, Purdue University, 1990  
MS in Computer Engineering, Purdue University, 1987  
BS in Computer Engineering, Boston University, 1985

#### 연구분야

Data Mining and Data Stream Mining  
Data Stream Management System [DSMS] Big Data  
[Map/Reduce, Hadoop, NoSQL]

다가오는 큰, 아주 큰 데이터 시대를 잘 준비하는 우리가 됩시다.

인류에게 증기기관의 발명은 산업화 사회를, 컴퓨터의 발명은 디지털 사회를, 인터넷의 발명은 정보화 사회를 만들었습니다. 우리는 이미 성숙한 정보사회에서 살아가고 있습니다. 앞으로 30년은 데이터 기술(DT)시대라고 합니다. 향후 20년 이내에 인류는 600억개의 스마트 기기들을 사용하면서 매일 15PB 분량의 사적인 정보를 발생시킬 것으로 예상됩니다. 이제 빅데이터의 발명은 데이터에 기반한 최적화 사회를 만들 것이라고 생각합니다. 빅데이터 기술이 선도하는 미래 사회는 사회가 작동하면서 발생하는 모든 데이터를 상세하게 모니터링하면서 개인의 가치와 대중의 가치를 최적으로 조화시키면서 각자에게 주어진 상황에 가장 적합하고 합리적인 방식으로 업무를 처리하는 최적화 사회가 될 것입니다.

빅데이터 기술은 정보화 사회 이후에 다가오는 미래 사회로의 중요한 변혁을 시작하는 도구입니다. 하지만 빅데이터를 잘 사용하기 위해서는 개인정보보호와 같이 해결해야 할 사회적 이슈들이 있습니다. 이제 우리나라는 빅데이터 기술의 역기능을 현명하게 해결하면서 방대한 양의 빅데이터를 분석하여 사회 모든 분야가 최적화되어 운영되는 미래 사회로 가는 새로운 탐험을 적극적으로 준비해야 합니다.

BOAZ, 빅데이터의 A에서 Z까지를 공부할 수 있는 대학생 동아리가 만들어진 지 벌써 7년이 넘어가고 있습니다. 그동안 생소하고 새로운 분야인 빅데이터를 공부하겠다고 모여서 열정적으로 시간과 노력을 함께 한 BOAZ 동아리 회원들에게 많은 찬사를 드리고 고마움을 느낍니다. 앞으로 미래를 준비하는 대학생의 열정으로 빅데이터의 A에서 Z까지를 탐구함과 동시에 빅데이터 활용으로 발생하는 사회적 이슈들을 현명하게 극복할 수 있도록 함께 노력해 주시기 바랍니다.

안녕하세요. BOAZ 14기 공동대표 김진호, 최윤정, 14기 분석대표 강민지입니다.

### 들어가며

다양한 기술들과 날로 쏟아지는 모델, 논문, ‘빅데이터’라는 용어로 시작된 데이터에 대한 트렌드는 하루가 다르게 발전하는 그리고 오픈소스의 형태로 급속도로 산업과 우리 생활에 녹아들고 있습니다. 이 글을 읽고 계신 여러분들 가운데서도, 빠르게 성장 중인 데이터 분야로의 진로를 치열하게 고민하고 공부하시는 분들이 많을 것입니다. BOAZ에 지원하기 위해, 혹은 BOAZ가 어떤 곳인지 궁금해서 이 책자를 열어 보신 여러분. 여러분께서는 과연 어떤 목표와 열정을 가지고 이 분야에 관심을 갖게 되셨는지 알고 싶습니다.

### BOAZ를 통해 얻었으면 하는 것들

BOAZ에 에서 활동하시게 될 회원 분들은 지식과 사람, 이 두 가지의 가치를 꼭 얻어 가셨으면 합니다.

BOAZ는 ‘데이터 사이언스’라는 학문을 배우고자 들어오는 곳임과 동시에 유사하지만 결코 완전히 같지는 않은 관심사를 지닌 사람들이 모여서 소통할 수 있는 공간입니다. BOAZ는 이러한 ‘커뮤니케이션 플랫폼’으로서 선배들과의 연락망을 구축하고 정기적으로 컨퍼런스톡이라는 행사 또한 진행하고 있습니다. 비단 데이터 관련 직무가 아니더라도, 다양한 분야의 인물들과 교류하며 시야를 넓힐 수 있다는 점은 그 어떤 동아리와 비교해도 밀리지 않는 BOAZ만의 큰 장점입니다. BOAZ를 통해 데이터 사이언스에 대한 지식은 물론이고, 세상을 보는 더 넓은 시야까지 얻어 가실 수 있게 되기를 진심으로 바랍니다.

### 같이 성장하는 가치

흔히 데이터 사이언스를 이야기하면 떠올리게 되는 통계학, 프로그래밍 등은 모두 핵심적인 역량이 맞습니다. 그러나 결코 이것이 전부라고 할 수는 없습니다. 어떤 도메인에 활용하느냐에 따라 요구되는 지식이 모두 다를 뿐만 아니라, 기술도 하루하루 발전하면서 필요한 능력도 점점 다양해지고 있습니다. 그렇기 때문에 다양한 전공, 다양한 출신의 학생들이 모여 각자의 시야를 공유하는 BOAZ가 필요한 것입니다. 이 분야를 독학함에 있어 공부 자체에는 큰 어려움이 없을지 모르지만, 접할 수 있는 지식의 폭은 지극히 한정적일 것입니다. 자신이 잘 아는 분야의 지식은 최대한 활발히 나누고, 그러면서 자신도 새로운 시선의 접근 방법을 배워갈 수 있는 것이 BOAZ가 자랑하는 ‘같이의 가치’입니다.

### 데이터 사이언티스트가 되기 위한 출발점

데이터 사이언티스트가 되기 위한 능력은 여러가지를 필요로 하며 분야에 따라 다르기도 합니다. BOAZ에서는 프로그래밍, 통계, 시각화, 머신러닝, 딥러닝 등을 커리큘럼에 따라 학습하고, 동아리원들의 자발적인 스터디를 통해 자신의 부족한 부분을 함께 공부합니다. 이를 바탕으로 BOAZ 멤버들은 여러 시행착오를 거쳐 프로젝트를 진행하며 6개월 마다 열리는 컨퍼런스에 참여하여 외부에 공개합니다. 또한, 여러 분야에서 열심히 활동하고 계신 선배들과 함께하는 활동을 통해 데이터 분야로의 진로를 구체적으로 그려보는 시간을 가지며 함께 이끌어 성장해 갑니다.

### 나가며

BOAZ는 2013년 창립되어 7년간 꾸준한 길을 걸어왔습니다. 그 시간 동안 BOAZ는 수없이 많은 세션과 다양한 주제의 컨퍼런스를 거치며 성장했습니다. BOAZ는 앞으로 만나게 될 16기 신입기수 여러분과 함께, 서로의 부족한 점을 메꿔주며 데이터 사이언티스트를 목표로 같이 걸어가고 싶습니다. BOAZ에서 함께 성장할, 새로운 16번째 동료가 되어 주실 여러분을 진심으로 환영합니다.



공동대표 김진호

국민대학교 빅데이터경영통계학과



공동대표 최윤정

국민대학교 빅데이터경영통계학과



분석대표 강민지

숙명여자대학교 컴퓨터과학전공



### BOAZ의 의미와 로고

분당 페이스북에는 2000만번 이상의 자료 공유가,  
트위터에는 30만번 이상의 트윗이,  
인스타그램에는 22만장 이상의 사진이,  
유튜브에는 72시간짜리 영상이,  
구글에는 하루 24테라바이트의 데이터가 업로드되고 있습니다.

더욱 가속화되고 있는 데이터의 증가 추이,  
규모를 가능할 수조차 없는 수많은 정보의 폭발.  
빅데이터는 미래 경쟁력의 핵심 자원으로 떠오르고 있습니다.

이렇게 많은 데이터로 어떤 일을 할 수 있을까요?  
어디에 어떤 방향으로 접근해야 유의미한 결과를 도출할 수 있을까요?  
빅데이터란 정말 ‘크고 많은’ 데이터만을 지칭하는 말이며,  
큰 데이터가 아니면 의미가 없는 것일까요?

BOAZ는 이러한 질문들을 끊임없이 던지면서  
토론과 실습을 통해 답을 찾아가려고 노력합니다.

BOAZ는 데이터로부터 긍정적인 가치를 이끌어 빅데이터 시대의  
인재로 거듭나고자 하는 대학생들이 모여 있는 동아리입니다.

BOAZ는 다양한 이공계/인문계/예술계열 학생들로 구성된  
가치관과 시각의 집합소입니다.



1. <어린 왕자>에 나오는 코끼리를 삼킨 보아뱀의 모습  
모두가 ‘모자’라고 할 때 ‘코끼리를 삼킨 보아뱀’을 상상할 수 있는 창의력  
코끼리(빅데이터)를 집어삼키는 보아뱀이 되겠다는 BOAZ의 도전 정신

2. BOAZ의 O 안의 차트  
데이터를 분석하고 시각화하여 병자들을 구했던 나이팅게일의 Rose  
Diagram 모티브

로고 디자인\_박소영

### 수상실적

서울시 희망광고 소재 대상 프로젝트 선정 〈데이터로 말하는 저상버스〉	김유나, 이성훈 (1기), 김진영, 박소영, 이주원 (2기)
교통데이터로 활용 공모전 네이버 특별상 〈국도로 떠나는 ‘휴’ 드라이브〉	김유나 (1기), 박소영, 이주원, 임제현 (2기)
제 13회 SAS 분석 챔피언십 입선	박인성 (3기), 은재현 (4기)
서울시 공공 데이터 아이디어 공모전 은상 〈장애인 이동 도움 알림, 장비서〉	윤일근 (3기)
제 13회 SAS 분석 챔피언십 입선	최진혁, 하윤정 (4기), 이주경 (5기)
서울시 공공 데이터 아이디어 공모전 동상 〈별보기 트래킹 어플리케이션, Star track〉	임민수, 신현화, 홍재현 (3기), 백재희, 한지희 (4기)
Kaggle USA Census Competition에서 Kaggle Staff's Favorites Script 수상 〈Look Over Korean Immigrants Life〉	최윤영 (3기), 박대한, 여새바그별, 이승진, 최진혁 (4기)
L.POINT 빅데이터 공모전 동상	이주경 (5기)
제 14회 SAS 분석 챔피언십 동상	최기준 (6기)
제 14회 SAS 분석 챔피언십 입선	전현우 (6기)
제 14회 SAS 분석 챔피언십 입선	여인연(7기)
제 4회 2016 빅콘테스트 챌린지리그 미래창조과학부 장관상	윤일근, 김한솔 (3기)
2017 한국축구과학회 국제컨퍼런스 대학생 아이디어 공모전 금상	장준환 (8기)
제 15회 SAS 분석 챔피언십 입선	김유경 (8기), 신소윤 (9기)
K-NEST Challenge 2018 입선	김완 (10기)
서울시 빅데이터 캠퍼스 공모전 우수상	구진모 (10기)
미래에셋 대우 빅데이터 페스티벌 장려상	김지연 (10기)
K-water 물관리분야 대국민 빅데이터 공모전 중화귀분석을 이용한 물 사용량 예측 장려상	구진모 (10기), 김정원 (11기)
2019 도서관 빅데이터 우수 활용 사례 및 아이디어 공모전 최우수상	고은비 (12기), 김태현 (12기), 전서연 (13기)



### 수상실적

2019 공공데이터 빅데이터 분석 공모전 대상	김명준 (13기), 박수현 (13기), 손유라 (13기), 전서연 (13기)
제 7 회 빅콘테스트 퓨처스리그 한국공항공사상	성민석 (12기), 최정만 (13기)
미래에셋 대우 빅데이터 페스티벌 금상	성민석 (12기)
Dacon AI프렌즈 시즌 1 온도 추정 경진대회 2위	성민석 (12기) , 문성민 (14기)
Tableau student viz contest 3위	전서연(13기), 윤채리(13기), 이성배(13기), 정동욱(13기), 지영우(13기)
Dacon AI 코로나 데이터 시각화 AI 경진대회 4위	전서연(13기), 윤채리(13기), 이성배(13기), 정소민(13기), 송인아(13기)
KED 빅데이터 시각화 공모전 대상	전서연(13기), 김원준(14기), 조하늘(14기)
KED 빅데이터 시각화 공모전 최우수상	강민지(14기), 장혜원(14기), 최윤정(14기)
2020 도서관 빅데이터 우수 활용 사례 및 아이디어 공모전 최우수상	고영희(13기), 이유진(13기), 김유진(14기), 김진호(14기)
kaggle Global Wheat Detection Bronze Medal	정연석(14기)
DACON 생체 광학 데이터 분석 AI 경진대회 7위	김태현(12기), 성민석(12기), 문성민(14기)
2020 금융 빅데이터 페스티벌 보험 과제 1위, 우수상	김태현(12기), 문성민(14기), 이다은(14기)

BOAZ는 16기부터 ‘데이터엔지니어링’ 부문이 추가되어 크게 세 부문으로 나누어 세션을 진행합니다.

데이터를 수집하여 분석하고 머신러닝 및 딥러닝 모델링을 실시하는 데이터분석 부문,

정보 전달의 효율성을 극대화하기 위해 주어진 데이터를 통해 인사이트를 도출하는 데이터시각화 부문,

그리고 데이터 적재부터 전처리, 모델링 결과의 활용까지 데이터 처리의 전반적인 흐름을 정의하는 데이터엔지니어링 부문입니다.

세 분야는 서로 다른 커리큘럼을 바탕으로 진행되며, 각 달에 한 번씩 공동세션을 통해 데이터 전반에 대한 다양한 생각을 나누고 친목을 도모합니다.

## 데이터분석 부문

Python을 이용한 프로그래밍 실습  
머신러닝, 딥러닝 학습  
통계이론, 머신러닝, 딥러닝 교재/강의 스터디  
데이터 분석 프로젝트 진행

데이터 수집, 전처리, 그리고 머신러닝 및 딥러닝을 통한 모델링까지 전반적인 데이터 분석을 아우르는 부문입니다. Base 회원들은 방학동안 데이터 분석, 머신러닝 및 딥러닝의 기초적인 개념을 익히고, 이를 이용해 프로그래밍 실습과 관련 kaggle 프로젝트를 진행합니다. 학기 중에는 방학동안 배운 내용을 바탕으로 세부 Task를 나눠 심화과정을 공부하게 됩니다. 15기의 세부 Task로는 이미지 분야 Detection, Generation과 텍스트 분야 Classification, Summarization, Topic Modeling, QA가 있었습니다. 세부 Task의 논문으로 들어가기 전, Vision 분야는 VGG, Resnet 등 이미지의 근간이 되는 논문과 NLP 분야는 워드임베딩 모델부터 Bert 등 각광받고 있는 논문들을 공통으로 학습한 후 각 Task에 맞는 논문을 공부하여 발표합니다. 논문 스터디를 마친 후에는 3주 간의 미니 프로젝트를 통해 다양한 모델들을 실제로 구현해보면서, 데이터 수집부터 모델링까지 딥러닝 프로젝트의 전반적인 과정을 익히게 됩니다.

## 데이터시각화 부문

Python을 이용한 프로그래밍 학습  
Tableau를 이용한 데이터 시각화 학습  
데이터 시각화 case study 및 토론  
데이터 시각화 프로젝트 진행

통계분석으로는 도저히 알 수 없는 데이터의 이야기를 끌어내 숨겨진 인사이트를 도출하는 것을 목표로 하는 부문입니다. 효과적으로 시각화된 데이터는 궁극적으로 뛰어난 전략 수립과 성공적인 판단으로 이어집니다. 데이터 시각화 부문 회원들은 방학 세션 동안 웹 크롤링을 통한 데이터 수집, 파이썬을 활용한 전처리, 그리고 시각화 툴 태블로(Tableau)를 다룸으로써 직관적이고 맛있는 시각화 방법뿐만 아니라 데이터의 신뢰성과 정확성을 위한 수집 및 전처리 방법과 통계 이론을 균형있게 학습하고, 이를 활용해 미니 프로젝트를 진행합니다. 이후 학기 중 세션에서는 자율 주제에 대한 데이터 시각화 프로젝트를 진행하고 발표함으로써 데이터를 통해 원하는 인사이트를 뽑아내고 이를 남들에게 제공하고 소통하는 방법을 배웁니다. 1년의 활동기간동안, BOAZ 시각화 부문 회원들은 본인만의 데이터에 관한 철학을 얻고, 나아가 개인의 개성을 살린 데이터 스토리텔링 능력을 갖춘 인재로 성장합니다.

## 데이터엔지니어링 부문

DBMS 및 SQL 이론 학습  
Hadoop 기초와 HDFS 이론 학습  
Docker를 이용한 Kafka 및 Spark 학습  
데이터 엔지니어링 프로젝트 진행

매일 1PB가 넘게 발생하는 빅데이터는 하나의 저장소에 저장하고 분석하는 것이 매우 어렵다는 점을 반영하여, 데이터베이스와 Hadoop을 통해 방대한 데이터의 효율적 관리와 운영을 추구하는 부문입니다. 데이터 엔지니어링 부문 회원들은 방학 세션 동안 각종 데이터베이스에 대한 개념 정립, 데이터베이스 내에서 데이터를 운용하는 SQL 구문, 그리고 Hadoop에 대한 이론을 배움으로써 빅데이터 저장 및 관리에 대한 전반적인 부분을 학습합니다. 이후 학기 중 세션에서는 Docker라는 가상환경 위에서 Kafka, Spark를 활용하여 기초적인 운영 및 SQL부터 Streaming, Machine Learning까지 방학 중 학습했던 내용을 실습하고 구현하며, 자율 주제의 프로젝트를 진행합니다. 1년의 활동기간 동안, BOAZ 엔지니어링 부문 회원들은 빅데이터에 관한 효율적인 관리 능력을 얻고, 나아가 파이프라인 구축을 통해 데이터 수집부터 ML/DL 분석까지의 과정에 대해 Real-Time 수준의 데이터 운용 실력을 키우게 됩니다.

# Curriculum

## 정규세션 전체

### 데이터분석 부문 (1 YEAR)

Base	방학세션	머신러닝 및 딥러닝 기초 이론을 배우고, 실습을 해보는 시간
	학기세션	방학 중 배운 머신러닝 알고리즘을 직접 실습해보고 딥러닝 논문을 읽고 구현해보는 시간
Adv	방학세션	신입기수와 함께 새로운 커리큘럼으로 세션을 진행하고, 컨퍼런스 프로젝트를 진행하는시간
	학기세션	

### 데이터시각화 부문 (1 YEAR)

Base	방학세션	Python을 이용한 프로그래밍 학습, Tableau를 이용한 데이터 시각화 학습, 케이스 스터디를 진행하는 시간
	학기세션	데이터 시각화 주제를 선정하여 데이터 분석부터 시각화까지 자율적인 프로젝트를 진행하는 시간
Adv	방학세션	신입기수와 함께 새로운 커리큘럼으로 세션을 진행하고, 컨퍼런스 프로젝트를 진행하는시간
	학기세션	

### 데이터엔지니어링 부문 (1 YEAR)

Base	방학세션	리눅스와 가상화, DBMS, SQL 등 데이터 파이프라인 구축을 위한 기초를 다지는 시간
	학기세션	분산 처리, 실시간 처리 오픈 소스를 직접 다뤄보고 데이터 운영 및 관리를 진행해보는 시간
Adv	방학세션	신입기수와 함께 새로운 커리큘럼으로 세션을 진행하고, 컨퍼런스 프로젝트를 진행하는시간
	학기세션	

# Curriculum

## 데이터분석 커리큘럼

### Base

#### 방학 세션

빅데이터를 다루기 위한  
기초지식을 배우는 시간  
머신러닝 및 딥러닝 기초 이론을 배우고  
실습을 해보는 시간

#### 기초 분석 방법론

Python을 이용하여 효과적인 데이터 전처리를 위해 필요한 데이터 구조에 대해 학습하고, 사용할 데이터를 가져와 재구성하는 방법을 배우는 시간입니다. 데이터 처리 및 활용에 효과적인 Numpy, Pandas, Seaborn 과 같은 다양한 Python 라이브러리에 대해 배우고, 여러가지 데이터에 직접 실습해 보면서 기초적인 데이터 처리를 학습하게 됩니다.

#### 머신 러닝 (Machine Learning) 기초

컴퓨터가 스스로 데이터를 분석해 의미 있는 결과를 도출하는 머신 러닝에 대해 학습하고 실제 사례에 적용해보는 시간입니다. Regression, Classification과 Regularization, Feature Analysis 등의 다양한 개념들을 공부합니다. 또한 배운 내용을 바탕으로 Kaggle에서 종료된 Competition을 통해 프로젝트를 진행합니다.

#### 딥러닝 (Deep Learning) 기초

Multi-Layer Perceptron(MLP), Backpropagation, Gradient Descent, Convolutional Neural Network (CNN), Recurrent Neural Network(RNN) 등 딥러닝의 기본이 되는 이론 과 기법들을 공부합니다. 이를 바탕으로 Pytorch, Keras, Tensorflow등의 딥러닝 프레임워크를 이용해 직접 구현해보는 시간을 갖습니다.

# Curriculum

## 데이터분석 커리큘럼

### Base

#### 학기 세션

방학 중 배운 머신러닝 알고리즘을  
직접 실습해보고 딥러닝 논문을 읽고  
구현해보는 시간

#### 딥러닝 (Deep Learning) 심화

방학동안 배운 딥러닝 이론을 바탕으로 NLP, Vision 등 관심 Task를 정해 집중적으로 공부합니다. 각 분야에서 대표적인 논문들을 중심으로 공부하게 되고, 이를 바탕으로 자신이 원하는 주제를 정하여 Tensorflow, Pytorch 등 딥러닝 프레임워크를 활용해 구현해 보는 시간을 갖습니다. 15기의 경우 물체 인식, 이미지 생성, 텍스트 분류, 텍스트 요약, 토픽 모델, 질의응답의 6개 조로 구성되어 각 Task를 공부하였고, 이를 마친 후에는 3주간의 미니 프로젝트를 통해 각자 원하는 주제를 결정하여 지금까지 공부해 온 다양한 딥러닝 모델들을 실제로 구현해보고 있습니다.

#### 자연어 처리

자연어처리(Natural Language Processing, NLP)는 자연언어의 의미를 분석하고 해석하는 분야로, 언어의 의미적/문법적 정보를 설명하고 문맥적인 의미를 보다 더 잘 파악하는 방향으로 발전하였습니다. 자연어처리는 단어의 출현빈도와 관계성을 파악하는 단순 텍스트 마이닝의 차원을 넘어 Machine Translation(기계번역), Sentiment Analysis(감성분석), Dialogue Generation(대화생성), Question Answering(질의응답)까지 응용되고 있습니다. 데이터 분석세션의 NLP팀은 4개의 세부Task로 나눠 지기 전에 말뭉치와 문법정보를 함축하기 위한 임베딩 기초 모델인 Word2vec, Seq2seq모델을 시작으로 자연어 처리 분야에서 각광받는 Attention, Transformer, Bert 모델까지 공통으로 공부한 후, 세부 분야로 나누어져 각 Task에서 뛰어난 성능을 보인 최신 NLP모델의 논문을 공부하고 있습니다.

#### 컴퓨터 비전

컴퓨터 비전은 딥러닝이 성공적으로 적용된 분야 중 하나로, 인공지능을 활용한 이미지 인식은 인간의 수준을 이미 뛰어넘었다고 할 수 있습니다. 컴퓨터 비전의 대표적인 분야로는 Image Classification(이미지 분류), Object Detection(물체 인식), Image Generation(이미지 생성)이 있습니다.

15기 데이터 분석세션의 Vision팀은 방학 세션에서 배운 CNN을 바탕으로 VGG, Resnet, Inception, DenseNet을 공통으로 공부하면서 컴퓨터 비전에 대한 기초를 탄탄히 다진 후 Object Detection과 Image Generation 두 분야로 나뉘어 세부 Task에 대해 집중적으로 학습하고 발표하는 시간을 갖습니다. 그 중 대표적으로 인식 팀은 RCNN과 YOLO, 생성팀은 VAE, GAN을 이용한 모델을 통하여 각 Task를 공부하고 있습니다.

# Curriculum

## 데이터분석 커리큘럼

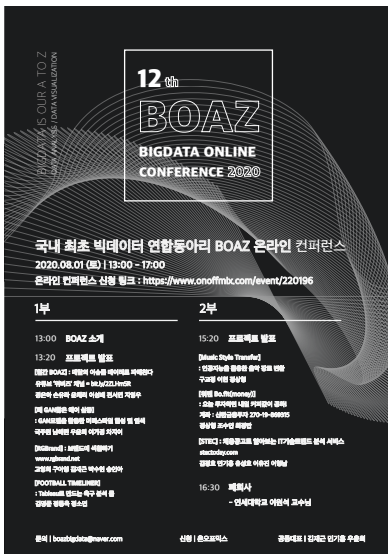
### Advanced

#### 방학 세션

신입 기수와 함께 새로운 커리큘럼으로 세션을 진행하고  
컨퍼런스 프로젝트를 진행하는 시간

### Advanced

#### 학기 세션



〈 제 12회 보아즈 빅데이터 컨퍼런스 〉

## 프로젝트 주제 선정

짧으면 두 달, 길면 반 년 동안 진행되는 프로젝트인 만큼 프로젝트 주제 선정은 매우 중요합니다.  
주제 선정은 BOAZ 회원들이 자발적으로 제시한 주제들을 바탕으로, 이후 관심이나 뜻이 맞는  
회원들끼리 팀을 꾸려 프로젝트를 진행하게 됩니다.

## 프로젝트 스터디

본격적인 프로젝트 진행에 앞서 팀끼리 모여 주제에 대한 분석 방식을 함께 공부하는 시간입니다.  
해당 주제에는 어떤 분석 알고리즘을 사용하는 것이 알맞는지, 그 분석 알고리즘은 어떻게 사용하  
야 최대의 성과를 낼 수 있는지에 대해 같이 논의하고 여러 참고 자료를 찾아보며 프로젝트가 수  
월하게 진행될 수 있도록 합니다.

## 프로젝트 진행

Advanced 세션에서 프로젝트를 진행하며 지식을 단순히 혼자 가지고 있는 것이 아닌, 다른 사  
람과 나누고 협력함으로써 시너지 효과가 나타나는 경험을 할 수 있습니다. 우리 팀의 프로젝트  
뿐만 아니라 다른 팀의 프로젝트에 대한 피드백을 같이 나눔으로써 BOAZ가 추구하는 '상생'의  
힘을 알 수 있습니다.

## BOAZ 빅데이터 컨퍼런스

회원들을 대상으로 한 리허설에서 뽑힌 우수한 최종 결과물들은 BOAZ에서 주최하는  
'BOAZ 빅데이터 컨퍼런스'의 무대에 설 수 있는 자격이 주어집니다. 이를 통해 1년 동안 BOAZ  
에서 배우고 익히고 적용한 것들을 대중들에게 공개하게 됩니다.

# Curriculum

## 데이터시각화 커리큘럼

### Base

#### 방학 세션

데이터 수집 및 전처리, 시각화 tool 활용,  
자율 과제 및 미니 프로젝트를 통해 전반적인  
데이터 시각화에 대한 이해를 넓히는 시간

#### 프로그래밍 학습

Python을 통해 데이터 수집 및 데이터 전처리를 위한 기초 프로그래밍을 학습합니다. 데이터를 수집하기 위한 웹크롤링과 데이터의 특성을 전반적으로 파악하기 위한 전처리 및 EDA 방법들을 실습합니다. 이를 통해 원하는 시각화를 구현하기 위해 최적의 데이터를 준비하는 방법을 배웁니다.

#### 방학 스터디

기존기수와 신입기수가 함께 태블로(Tableau)를 활용한 대시보드 설계를 구성하는 방법들을 실습합니다. 태블로란 Drag&Drop만으로 쉽게 차트, 그래프 등을 완성할 수 있는 유용한 시각화 툴이며, 대시보드(Dashboard)는 데이터 분석으로부터 얻어낸 중요한 인사이트들을 직관적으로 파악할 수 있도록 다양한 차트를 디자인 원칙에 따라 구성해 놓은 시각적 리포트입니다. 방학 스터디를 통해 태블로의 다양한 기능을 사용하여 데이터의 숨겨진 중요한 의미들을 빠르고 효과적으로 시각화하는 방법을 배웁니다.

#### 케이스 스터디

매주 과제를 통해 중요한 정보를 전달하기 위한 차트를 직접 구상하고 발표합니다.  
다양한 관점을 가진 회원들과 자유롭게 의견을 공유하고 시각화와 관련해 토론을 하는 시간을 가집니다. 이를 통해 문제를 다각도로 바라보는 힘을 기릅니다

주제 예시 : 맵 뷰를 활용한 여성 1인가구 분포 대비 안전 시설 필요도 시각화

# Curriculum

## 데이터시각화 커리큘럼

### Base

#### 학기 세션

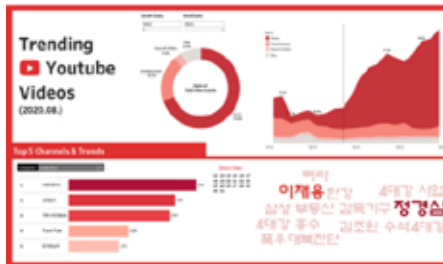
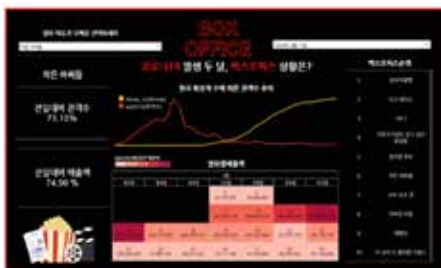
시각화 Tool 실습, 팀별 프로젝트를 통해  
데이터 분석부터 시각화까지의 전체적인  
과정을 진행하는 시간

### 데이터 분석 심화(Python)

방학동안 배운 데이터 전처리 기법을 바탕으로 더 심층적인 기법들을 학습하는 기간입니다. 다양한 데이터 유형별 전처리 기법을 실습하는 것부터, 간단한 머신러닝 및 자연어 처리를 비롯하여 자신의 관심사에 맞는 데이터 분석 기법을 공부할 수 있습니다.

### 팀별 프로젝트(Tableau)

방학 세션 및 다양한 활동을 기반으로 자유롭게 주제를 선택한 후, Base 팀별 프로젝트를 진행합니다. 데이터 수집부터 전처리, 분석 및 시각화까지 모든 단계를 수행하여 그 결과물을 충실히 보여줄 수 있는 콘텐츠까지 제작하게 됩니다. 이후 팀별 프로젝트의 결과물은 컨퍼런스에서 발표 및 전시합니다.



\* 아래 Tableau Public의 BOAZ 계정 링크에서 더 많은 대시보드를 보실 수 있습니다.  
<https://public.tableau.com/profile/boaz.visualization#!/>



# Curriculum

## 데이터시각화 커리큘럼

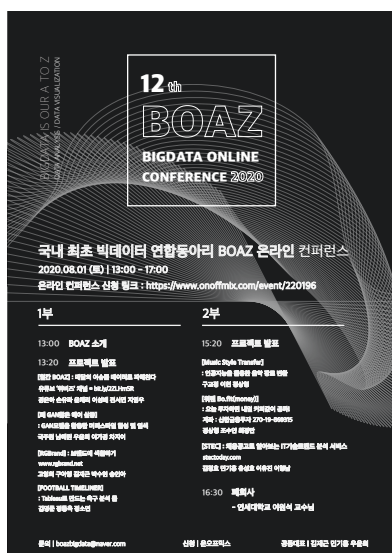
## Advanced

## 방학 세션

신입 기수와 함께 새로운 커리큘럼으로 세션을 진행하고  
컨퍼런스 프로젝트를 진행하는 시간

## Advanced

## 학기 세션



〈 제 12회 보아즈 빅데이터 컨퍼런스 〉

## 프로젝트 주제 선정

자신이 스토리텔링하고 싶은 데이터를 결정합니다. 시각화 프로젝트의 가장 큰 목표는 데이터에 관한 ‘자신의 철학’을 대중들에게 전달하는 것이기에, 프로젝트의 결과물이 사회에게 긍정적인 영향을 미칠 수 있을지 지속적으로 회원들과 논의합니다. 이를 통해 관심이 있는 회원들이 각자의 팀을 꾸려 프로젝트를 시작하게 됩니다.

## 프로젝트 설계

스터디를 진행하기 앞서 선정된 주제가 어떤 사회적 니즈를 가지고 있는지, 어떤 시각화 결과물이 최적으로 주제에 맞는지 설계하는 시간을 가집니다. 시각화를 볼 대상이 실제로 필요한 시각화는 무엇인지 고민하고 조사하며, 어떠한 데이터를 어떻게 활용해야 효율적인 시각화가 가능할지 논의합니다. 또한 웹, 동영상, 대시보드 등 다양한 형태로 시각화 방식을 토론하여 프로젝트 진행의 청사진을 그립니다.

## 프로젝트 진행

프로젝트 진행단계는 데이터 수집 -> 전처리 -> 시각화로 이뤄집니다. 이 모든 과정을 혼자서 아닌 BOAZ 내의 다른 팀원들과 함께 협력하며 진행하고, 매 단계가 끝날 때마다 다른 팀과의 피드백을 주고받음으로써 자신의 팀의 프로젝트를 되돌아볼 수 있게 합니다. 프로젝트가 끝나더라도 보다 확장성 있는 시각화를 위해 상품화까지 고려하며 프로젝트를 마무리합니다.

## BOAZ 빅데이터 컨퍼런스

회원들을 대상으로 한 리허설에서 뽑힌 우수한 최종 결과물들을 BOAZ에서 주최하는 ‘BOAZ 빅 데이터 컨퍼런스’의 무대에 설 수 있는 자격이 주어집니다. 이를 통해 1년 동안 BOAZ에서 배우고 익히고 적용한 것들을 대중들에게 공개하게 됩니다.

# Curriculum

## 데이터엔지니어링 커리큘럼

### Base

#### 방학 세션

리눅스와 가상화, DBMS, SQL 등

데이터 파이프라인 구축을 위한 기초를 다지는 시간

#### 리눅스 기초

리눅스(Linux)는 컴퓨터 OS 커널의 일종 혹은 리눅스 커널을 사용하는 운영 체제를 가리키며, 소스 코드가 공개되어 있는 대표적인 오픈 소스 소프트웨어입니다. BOAZ 가입 후 데이터 엔지니어링 부문 회원들은 먼저 원활한 데이터 파이프라인 처리를 위해 리눅스의 기초에 대해 배우게 됩니다. Shell Script 프로그래밍을 통한 실습으로 리눅스 명령어를 해석하는 기능의 Bash Shell부터 리눅스 파일 시스템의 구조 및 특징, 프로세스 및 패키지 관리 등 전반적인 리눅스의 기능에 대해 학습하게 됩니다.

#### 네트워크 기초

Network(Server&Client, Protocol, Routing, IP, Port, Mac Address)

네트워크는 두 대 이상의 컴퓨터를 서로 연결하여 데이터 통신을 제공하는 것을 의미합니다. 우리가 사용하는 컴퓨터에서 빅데이터가 저장되어 있는 데이터베이스에 접근하고 데이터를 처리하기 위해서는 네트워크가 필요합니다. 네트워크 기초를 학습함으로써 이후 데이터 엔지니어링 분야에서 많이 활용하는 가상화 기술에 대해 학습할 때, 높은 진입장벽을 낮춰줄 수 있는 시간입니다.

#### 데이터베이스 기초

RDBMS, ER-Model, 정규화 등 데이터베이스의 기본이 되는 이론을 공부합니다. 이를 바탕으로 SQL을 이용해 직접 데이터베이스에서 데이터를 구축하고 처리, 분석해보는 시간을 갖습니다. 그리고 확장성, 유연성, 가용성 등 빅데이터 시대의 요구사항을 충족시키지 못한 SQL의 한계로 등장한 NoSQL의 기초를 배우게 됩니다.

#### 가상화 기술

Docker의 특징, 명령어, Dockerfile, Container Orchestration(Kubemetes) 이론

대표적인 가상화 기술인 Docker에 대해 학습하는 시간입니다. 컨테이너, 가상 머신, 클라우드 환경 등의 가상화 기술에 대한 기초적인 이론에 대해 배우게 됩니다. 이후, 리눅스 커널 기반으로 구성되어 있는 Docker의 컨테이너 기술을 이용하여 실습을 진행합니다.

# Curriculum

## 데이터엔지니어링 커리큘럼

### Base

#### 학기 세션

Docker를 통해 Kafka와 Spark를 실행해보고  
데이터 운영 및 관리를 진행해보는 시간

#### 범용 빅데이터 저장소 구축

분산 컴퓨팅 기초, MapReduce, Yarn, HDFS, Hadoop Ecosystem

방학동안 배운 기초 이론을 바탕으로 빅데이터를 위한 범용 저장소 및 분석 플랫폼을 제공하는 하둡(High-Availability Distributed Object-Oriented Platform, Hadoop)을 공부하게 되고 하둡의 Map-Reduce, Yarn, HDFS 등의 이론을 통해 분산 컴퓨팅에 대한 기초를 학습하게 됩니다.

#### 인메모리(In-Memory) 분산 데이터 처리

Spark 특징, Job Scheduling, Streaming, ML

스파크(Spark)는 메모리 상에서 동작하는 분산데이터 분석 시스템으로서 하둡의 Map-Reduce 모델에 대해서 대화형 명령어 쿼리나 스트리밍, Machine Learning이 가능하도록 확장한 프레임워크입니다. Spark의 Application 실행 구조와 같은 특징과 스파크의 RDD 데이터 타입에 대한 이론 등을 배우고 스파크 인터페이스인 Spark Shell과 Zeppelin Notebook을 통해 데이터프레임을 생성하는 등의 실습을 진행하게 됩니다.

#### 실시간 메시지 처리

Kafka&Zookeeper, Consumer&Producer, 운영, Stream처리

링크드인(Linkedin)의 급성장으로 인해 발생하는 서비스 내부의 여러 이슈들을 해결하기 위해 탄생한 카프카(Kafka)는 통합된 데이터 처리 시스템의 부재로 발생하는 복잡도를 줄이기 위해 탄생했습니다. 모든 시스템으로 데이터를 전송할 수 있으며, 실시간으로 데이터 처리가 가능한 카프카는 데이터 파이프라인 구축에 있어 상당히 중요한 부분을 차지하고 있습니다. 해당 파트에서는 Consumer와 Producer가 분리되어 있는 카프카의 특징부터 카프카와의 직접적인 통신을 통해 카프카의 메타데이터 정보를 저장하고 상태를 관리하는 주키퍼(Zookeeper)에 대해 학습하는 시간을 가지게 됩니다. 이후, Docker와 Python을 활용한 카프카 실습을 통해 메시지와 토픽을 생성하고 전달해보는 등의 운영을 공부하게 됩니다.

# Curriculum

## 데이터엔지니어링 커리큘럼

### Advanced

#### 방학 세션

신입 기수와 함께 새로운 커리큘럼으로 세션을 진행하고  
컨퍼런스 프로젝트를 진행하는 시간

### Advanced

#### 학기 세션

### 프로젝트 주제선정

Advanced 세션이 진행되는 반 년 동안 프로젝트를 실시하기 때문에 프로젝트 주제 선정은 매우 중요합니다. 데이터 엔지니어링 부문의 경우 엔지니어링에 대한 심화 및 원활한 데이터 운용을 바탕으로 다른 부문과 연계하여 주제 선정이 이뤄질 수 있습니다. 이후, BOAZ 회원들이 자발적으로 제시한 주제들을 바탕으로 관심이나 뜻이 맞는 회원들끼리 팀을 꾸려 프로젝트를 진행하게 됩니다.

### 프로젝트 스터디

본격적인 프로젝트 진행에 앞서 팀끼리 모여 주제에 대한 분석 방식을 함께 공부하는 시간입니다. 해당 주제에는 어떤 파이프라인을 통해 운영하는 것이 알맞는지, 그 파이프라인을 어떤 식으로 구조화해야 최대의 성과를 낼 수 있는지에 대해 같이 논의하고 여러 참고 자료를 찾아보며 프로젝트가 수월하게 진행될 수 있도록 합니다.

### 프로젝트 진행

Advanced 세션에서 프로젝트를 진행하며 지식을 단순히 혼자 가지고 있는 것이 아닌, 다른 사람과 나누고 협력함으로써 시너지 효과가 나타나는 경험을 할 수 있습니다. 우리 팀의 프로젝트 뿐만 아니라 다른 팀의 프로젝트에 대한 피드백을 같이 나눔으로써 BOAZ가 추구하는 '상생'의 힘을 알 수 있습니다.

### BOAZ 빅데이터 컨퍼런스

회원들을 대상으로 한 리허설에서 뽑힌 우수한 최종 결과물들은 BOAZ에서 주최하는 'BOAZ 빅데이터 컨퍼런스'의 무대에 설 수 있는 자격이 주어집니다. 이를 통해 1년 동안 BOAZ에서 배우고 익히고 적용한 것들을 대중들에게 공개하게 됩니다.

# Curriculum

## 공동세션

한 달에 한 번씩 데이터분석, 시각화, 엔지니어링 부문 회원들이 모두 모여 빅데이터와 관련된 다양한 활동을 하는 자리를 가집니다. 공동세션에서는 빅데이터와 관련된 이슈로 토론을 하거나, 정규 세션에서 배우지 않는 새로운 빅데이터 분석 알고리즘에 대한 강연 등이 이루어집니다. 이 시간을 통해 빅데이터에 관한 다양한 견해를 나누고 배우며, 사회 전반에 걸치는 전체적인 동향을 파악할 수 있습니다.

공동세션은 평상시 정규 세션을 함께하지 못하는 Base와 Advanced, 두 활동 기수의 화합의 장이기도 합니다. 두 기수가 함께 어울려 빅데이터 이슈와 관련된 다양한 주제를 서로 공유하는 과정이며, 매 해 커리큘럼마다 BOAZ 내부 회원들의 의견을 수용하여 주제를 다르게 설정하고 있습니다. 또한 정규 세션을 진행하는 방식이 달라 자주 만나지 못하는 각 부문 회원들의 통합을 위한 자리이기도 합니다. 따라서 다양한 부문의 회원들에게 통합적인 지식을 제공하고 각 부문과 관련된 세부 내용을 다루기도 합니다.

지난 공동세션에서는 Kaggle, Colab, Slack, Python을 이용한 웹 크롤링등의 다양한 주제들로 진행되었습니다.

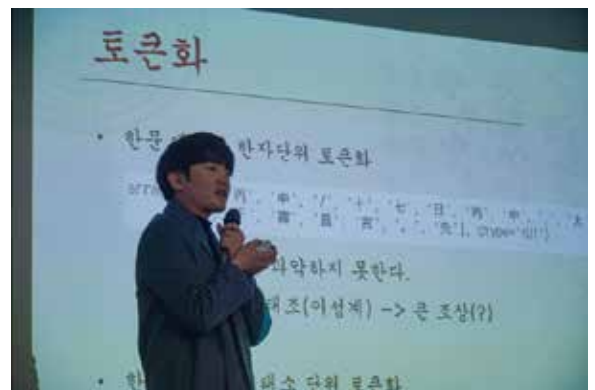


## BOAZ 빅데이터 컨퍼런스

컨퍼런스는 동아리원 모두가 자신의 모든 역량을 쏟아내어 이뤄낸 결과물입니다. 또한 동아리에서 활동하며 습득한 지식과 기법을 적용하여 새로운 응용 알고리즘에 도전을 하고, 어려움을 스스로 해결해 나가면서 성장해 나가는 발판이기도 합니다.

컨퍼런스는 일반적으로 지도교수님의 개회사로 시작되어 초청 연사분의 강연을 끝으로 마무리됩니다. 각 참가팀은 동아리원으로 구성되며, 동아리 회원분들은 매 발표 후 질문을 받고 답변함으로써 상호 피드백을 통한 발전을 모색합니다. 다양한 의견과 열띤 토론을 통해 BOAZ의 지향점인 '교류'와 '상생'에 한걸음 더 다가가고자 합니다.

\* 보다 자세한 컨퍼런스 내용은 컨퍼런스 책자를 확인해주세요 \*



## 1st BOAZ 빅데이터 컨퍼런스

- 빅데이터 활용 사례
- 데이터 분석을 활용한 효율적 사막화 방지
- 국도로 떠나는 ‘휴’ 드라이브  
제 3회 교통데이터 활용 공모전 정보디자인부문  
‘네이버 특별상’ 수상작
- 영화 흥행요소를 통해 알아보는 무비코드
- 감정용언과 음악코드를 통해 바라본  
우리나라 대중음악
- 데이터로 말하는 저상버스

## 2nd BOAZ 빅데이터 컨퍼런스

- 도서 분석을 통한 사람들의 심리&마음 파악
- 텔레마틱스 데이터를 이용한  
자동차 운전자 프로파일 생성
- 프리미어리그 축구팀 승자 예측
- 기상레이더를 활용한 강우 확률분포 예측
- 허브웨이 데이터를 활용한  
자전거 이용 시스템 시각화
- 지하철역에 숨겨진 운동의 기회를 시각화 :  
달까지 한걸음

## 3rd BOAZ 빅데이터 컨퍼런스

- 뉴스 데이터를 활용한 이슈 분석
- 2015 소주 트렌드북 :  
소주 관련 데이터 시각화와 인포그래픽
- 하이브리드 협업 필터링을 이용한  
맥주 추천 알고리즘
- 미국 Census 데이터 분석을 통해 알아보는  
한국인 이민자 라이프 스타일
- 사용자의 버킷리스트 분석 결과를  
시각화 이미지로 제공
- 대한약사회 데이터를 활용한  
비용 금기 의약품 시각화

## 4th BOAZ 빅데이터 컨퍼런스

- 통계로 본 서울시 유기동물 정보 시각화
- 음성데이터를 활용한 화자인식 구현 및  
보안에서의 활용
- CNN 이미지 인식 알고리즘을 활용한  
사칙연산 수행
- 모바일 SNS 사용자들의  
감성 용어 사전 제작 및 공인대상 감성 분석
- 회귀 분석 및 시뮬레이션 모델에 기반한  
웹서비스 제작
- 미국보다 한국 시장에서  
상대적으로 더 흥행할 영화의 사전 예측
- 소고기와 돼지고기에 대한 정보 시각화

## 5th BOAZ 빅데이터 컨퍼런스

웹툰의 OSMU 가능성 예측을 위한  
통계적 모델링

서울에 거주하는 20 - 30대  
나홀로족을 위한 라이프 가이드

서울 시내 간선도로 교통량을 활용한  
주요 지점 간 이동 시간 시각화

폐업률을 기준으로 한  
서울시 강남구의 창업 입점 생존율 분석

워드 임베딩과 인공지능망을 이용한  
개인 맞춤형 레시피 추천

유형 타입 분석을 통한 향수 추천

## 6th BOAZ 빅데이터 컨퍼런스

ITS와 유동인구 데이터를 활용한  
교통 예측 모델

베이지안 분석방법을 이용한 손상된  
이미지 복구

DeepTitle : 한국어 기사 자동 요약

위기탈출 경전철!  
- 신설 경전철 노선의 역별 수송수요 예측

## 7th BOAZ 빅데이터 컨퍼런스

좋니? 응, 좋아 - 이별가사 generator

Unpaired Image-to-Image Translation  
with CycleGAN

매의 눈 : 악성 댓글 분류 시스템

대학 생활 연구소 :  
대학별 Facebook '대나무숲' 토픽 분석

Img2Braille - 이미지 to 점자 번역기

어린이 지키는 노랑이

## 8th BOAZ 빅데이터 컨퍼런스

이젠 낚이지 마세요 - 낚시성 기사 분류기

감(感)으로 찍지 말고 감(GAM)으로 예측하자!

리뷰왕 답리뷰

YOLO WOW :  
You Only Look Once at What yOu Want

SNS 텍스트마이닝을 통한 브랜드 불만지수 구축

ABCD :  
Adjuvant of Breast Cancer Diagnosis



## 9th BOAZ 빅데이터 컨퍼런스

트위터 트렌드 분석을 통한  
신제품 개발  
- 새로운 맛을 찾아서

너의 기분 이모지(emoji)?  
: 감정기반 이모지 추천 시스템

AiRBnB  
: Aspect in Review 비교 & 분석

강화학습을 이용한  
주식 트레이딩 알고리즘

Deep Brother  
: 인공지능 출석체크

대학생 및 사회초년생을 위한  
여행 정보 제공 인터페이스  
(Web & Application)

## 10th BOAZ 빅데이터 컨퍼런스

영화 장르별 색채 시각화 및  
색채 조합 추천 시스템

2차가해의ナビ효과  
-정준영 동영상 사건

인공지능 영어 튜터  
: 자동으로 중요한 단어와 문장을 깊어주는 버트썸

밑바닥부터 시작하는 추천시스템  
-Trivago의 실제 로드 데이터로 구현

계약서 조항 자동 분류기  
:길고 어려운 계약서 쉽게 읽기!

## 11th BOAZ 빅데이터 컨퍼런스

BBM  
Boaz Barrier-free Map

코끼리(BOAZ)사서의  
도서 추천 솔루션

응답하라 1392  
트랜스포머 기법을 활용한 고문헌 번역

삐뽀삐뽀~ 댕다! AI 민중의 지팡이!  
딥러닝을 활용한 사회 범죄 예방 시스템

BOAZ Documents

## 12th BOAZ 빅데이터 컨퍼런스

월간 보아즈  
매달의 이슈를 데이터로 파헤친다

꽤 GAN찮은 헤어 살롱  
GAN 모델을 활용한 머리스타일 합성 및 염색

RGBrand : 브랜드에 색칠하기

FOOTBALL TIMELINER  
Tableau로 만드는 축구 분석 툴

Music Style Transfer  
인공지능을 활용한 음악 장르 변환

워렌 Bo.fit(money)

STEC  
: 채용공고로 알아보는 IT기술트렌드 분석 서비스



정규 세션을 진행하면서 자신이 부족하다고 생각하는 부분이나 더 공부하고 싶은 부분을 스터디를 통해 보충할 수 있습니다.



스터디 모집 기간 동안 원하는 분야와 공부하고 싶은 책을 추천하여 신청자를 받아 스터디 팀을 구성합니다.



다양한 기수와 다양한 부문의 활동 회원으로 구성된 스터디 팀은 8주 동안 진행될 스터디 계획과 운영 방식, 스터디 규칙을 자율적으로 정합니다.



보통은 해당 주차 계획된 분량을 각자 공부한 뒤 스터디에 참석하여 다른 팀원과 의견을 나누거나, 다른 팀원이 진행하는 강의를 들으며 부족한 부분을 질문하는 형식으로 이루어집니다.

## 〈2020년 하반기 스터디 목록〉

cs231  
Dacon  
GAN  
Google Analytics  
SQL  
Tableau  
머신러닝  
빅데이터분석기사  
추천시스템  
텍스트마이닝  
한국어임베딩



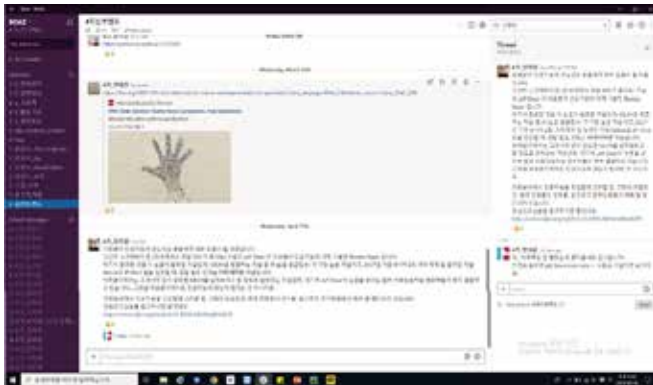
# Activity

## 기타활동



### 컨퍼런스 톡

BOAZ 내에서 진행되는 자율 방식의 미니 컨퍼런스입니다. 형식적인 자리에서 이루어지는 깊은 전문지식의 공유가 아닌, 다양한 전공의 서로 다른 방법론을 가진 BOAZ 회원들끼리 자신이 진행하는 일, 공부, 연구 등을 자유롭게 공유하는 자리입니다.



### 수료자들과의 교류

BOAZ를 수료한 회원들의 모임으로 Slack을 통해 지속적인 교류를 하고 있습니다. 구인구직 및 스터디, 최신 논문, 최신 트렌드 등에 대한 정보를 공유합니다.



### BOAZ 해커톤

BOAZ의 모든구성원들이 참여할 수 있는 해커톤입니다. 활동기수 뿐만 아니라 수료회원들도 참여하여 1박 2일 해커톤을 진행합니다. 조별로 주제를 선정하고, 데이터 분석 및 시각화를 진행하여 발표하는 자리입니다.

\*위 활동은 코로나19로 인해 진행되지 않았습니다



### 기업 방문

NEXON KOREA, Xlllab(씨이랩), 현대자동차, 한국기업 데이터, 밀버스

저희 BOAZ는 기업 방문을 통한 실무 체험도 하고 있습니다. 위 기업에 방문하여 실무에서의 데이터 분석 현황과 과정을 직접 경험하였습니다. 또한 인사담당자를 만나 채용 과정에 대해 질의응답하는 시간을 가졌습니다. 다양한 기업과 단체에서 먼저 연락이 오는 등 활발한 교류가 이루어지고 있습니다.

\*위 활동은 코로나19로 인해 진행되지 않았습니다

# Activity

## 기타활동

### 친목 활동

BOAZ의 또 다른 힘은 구성원 간의 유대감과 인적 자원에 있습니다. 학업 외의 시간에는 워크샵, 뒷풀이, 선배 기수와의 팀별 미션 등 각종 친목 도모를 위한 활동을 진행하고 있습니다.



### BOAZ X 패스트 캠퍼스

패스트캠퍼스는 프로그래밍, 데이터 사이언스, 마케팅, 파이낸스, 비즈니스 분야의 직무교육을 기획하는 기관입니다. BOAZ를 통해 패스트캠퍼스의 온라인/오프라인 강의 수강권과 조교모집 등을 지원받을 수 있습니다.





# Appendix

## 전 활동회원들의 추천글

### 7기 대표 구민수

‘빅데이터’는 어느새 우리 생활에 스며든지 수 년이 지났지만, 빅데이터의 잠재력과 활용성은 점점 더 증가하고 있습니다. 그렇기 때문에 성장속도와 트렌드 역시 빠르게 변화하고 있습니다. 빅데이터 동아리 ‘BOAZ’는 다양한 분야의 사람들이 모여 서로의 가치를 공유하고, 연구하여 발전할 수 있는 ‘정제소’입니다. BOAZ의 다양한 가치와 활동으로, 아이디어와 지식, 장점을 공유하고 결점을 보완함으로써 ‘원석’에서 진정한 ‘보석’으로 거듭나시길 바랍니다.

### 7기 대표 조수민

안녕하세요, BOAZ 7기 대표를 지냈던 조수민입니다. 저는 BOAZ에서 프로젝트를 1년간 진행하며 같은 꿈을 꾸고 있는 수많은 사람들과 어울려 놀기도, 고생하기도 했습니다. 그리고 현재 동고동락하고 있을 14기 후배님들과, 이들을 보며 같이 꿈을 키워 나갈 15기 후배님들을 한 발치 떨어져서 바라보고 있다면 부럽기도 합니다. 그 이유는 다양한 색깔을 가졌지만, 본성은 친근하고 정 많은 사람들이 항상 BOAZ에 찾아와준다는 점이라고 생각합니다. 후배님들도 재밌는 추억 많이 남기셔서 또 다른 후배님들에게 귀감이 될 수 있으면 좋겠습니다.

### 8기 대표 강희건

안녕하세요. 저는 8기 공동대표 강희건입니다. BOAZ에서 느낄 수 있는 매력은 전공 관련 없이 빅데이터에 관심 있는 사람들끼리 모여 프로젝트 및 스터디를 진행한다는 것이었습니다. 통계학, 공학, 상경계가 아닌 다양한 배경을 가진 사람들과 어울려서 프로젝트를 수행한 경험을 통해서 많은 것을 배울 수 있었습니다. 여러분들도 BOAZ에서 활동하면서 시야를 더 넓힐 수 있는 계기가 되었으면 합니다. 감사합니다.

### 8기 대표 김현중

BOAZ는 빅데이터 시대의 리더가 될 사람들이 모인 곳입니다. 빅데이터를 공부하기 위해서는 다양한 능력이 필요하고, 그 능력을 다른 친구들에게서 배울 수 있습니다. 저 역시 8기로 1년간 활동하면서 정말 대단한 친구들이 많다고 느꼈고, 더 많이 공부해야겠다는 생각이 듭니다. BOAZ는 빅데이터 광산에서 함께 보석을 찾아갈 광부를 기다립니다. BOAZ와 함께하는 광부들은 단순히 같이 공부하는 것을 넘어 그 이상의 무언가를 얻어갈 수 있을 것으로 확신합니다.

# Appendix

## 신입을 위한 Q&A

Q1. OT, 컨퍼런스중 하나만 참석이 가능한가요?

A1. OT, 컨퍼런스 둘 다 참석해야 합니다. 불참 시 불합격이 되니, 미리 일정 조정을 마쳐주세요.

Q2. 관련 지식이 없어도 혹은 다룰 줄 아는 데이터 분석 툴이 없어도 지원 가능한가요?

A2. 지원 가능합니다. 저희는 빅데이터에 관심과 열정이 있는 분들과 함께하고 싶습니다.

데이터 관련 기술이나 능숙하게 다룰 줄 아는 툴이 없다고 해도 본인이 가진 관심과 열정을 잘 표현하신다면 서류와 면접에서 좋은 결과를 기대하실 수 있을 것입니다. 또한 BOAZ의 일원이 되신다면 정규 세션 외에도 다양한 스터디를 통해 본인에게 부족한 부분을 충분히 채울 수 있습니다.

Q3. 면접 날짜를 선택할 수 있나요?

A3. 개인 사정으로 면접 불가능한 시간이 있으시다면 지원 시 지원서에 써주시면 됩니다. 면접 시간은 불가능한 시간을 최대한 고려해 조정됩니다. 그리고 면접 시간에 맞춰 면접 장소 대관과 면접관, 면접도우미 스케줄을 맞추기 때문에 면접 시간 공지 후 시간 변경은 지원자 간 시간 교환만 가능합니다. 참고로, 면접날짜는 1월 2일, 1월 3일 중 하루입니다.

Q4. 관련 학과가 아닌데 지원이 가능한가요?

A4. 지원 가능합니다. 현재 활동회원 중에도 언론홍보영상학과, 경영학과, 산업보안학과, 경제학과, 체육학과 등 다양한 전공자들이 있습니다.

Q5. 졸업 예정 / 입학 예정자도 지원 가능한가요?

A5. 지원 가능합니다. 현 동아리 회칙상 '현재 대학교, 대학원, 또는 전문대학에 재학 · 휴학 · 입학 예정이면서 최소 활동 기간인 1년 2term 동안 학생신분을 유지할 수 있는 자'에게 지원 자격이 주어집니다. 자세히 말씀 드리면 활동기간 첫 term(Base)뿐만 아니라 2term(Advanced)까지 학생 신분을 유지해야 합니다. 예를 들어, 졸업예정자는 대학원을 진학하거나 휴학이 예정되어 있는 자를 학생신분을 유지할 수 있는 자로 봅니다.

Q6. 경쟁률은 어느 정도인가요?

A6. 경쟁률은 기수/부문마다 모집인원과 지원인원이 매번 바뀌기 때문에 확실하게 말씀드리기 어렵습니다. 다만, 면접 대상자는 최종 선발 인원의 1.5배수 정도를 선발합니다.

Q7. 부문 간 중복지원이 가능한가요?

A7. 세 부문 중 하나의 부문에만 지원이 가능합니다.

Q8. 신입기수의 능력은 어느 정도 기대하시나요?

A8. 신입기수에게 기대하는 정해진 능력의 기준은 없다고 보시면 됩니다. 대신, 각 부문별로 주제에 대한 관심과 자신이 이를 통해 하고 싶은 것에 대해 기재해 주시면 합격에 도움이 될 것입니다. 참고로, 현 활동기수도 능력 분포는 다양합니다.

Q9. 각 부문에서 다른 부문의 공부를 함께하기 어려운가요?

A9. 각 부문 별 내용은 유기적인 관계로, 일부 과정은 중복으로 진행됩니다. 또한 시간 여유가 있으시다면 부문 병행이 가능합니다. 병행은 최대 두 세션까지 가능하며, 병행 세션도 과제와 학습 모두 본 세션처럼 열심히 참여해야 합니다. 병행은 Base때 한번 신청 가능합니다.

Q10. 현 기수는 평균적으로 어떤 수준에서 시작하여 지금 어느 정도 발전하셨나요?

A10. BOAZ는 정기적인 세션 진행 외에 스터디를 자율적으로 병행하여 진행하기 때문에 발전 정도를 일괄 측정하기 어렵습니다. 이 부분은 개인의 시간 투자에 따라 차이가 큼니다. 다만 매주 세션마다 발표를 세미나 형식으로 진행하기 때문에 기본적인 부분은 몸에 익히게 됩니다.

Q11. 정규 세션은 어떤 식으로 진행되나요?

A11. 각 부문의 Base 세션은 팀별 발제 형식으로 진행되며, Advanced 세션은 프로젝트 팀별로 자유롭게 진행됩니다. 아직 16기 커리큘럼은 논의 중으로, 변경될 수 있습니다.

Q12. 코로나19 이후 동아리 운영은 어떻게 진행되고 있나요?

A12. 지난 1년간 동아리 활동은 코로나19 방역 지침을 준수하여 대부분 정상적으로 진행되었습니다. 다만, 사회적 거리두기 격상으로 인해 세션, 컨퍼런스 등이 온라인으로 전환되었으며, 해커톤 및 기업 탐방의 경우 행사를 진행하지 않았습니다. 방역을 준수하면서 상황변화에 따른 유연한 대처를 통해 걱정은 덜고 알찬 활동이 되게끔 노력하는 BOAZ가 되겠습니다.

