

# 류 두 영

Dooyeoung Ryu/ 柳斗榮

Data Analysis Python MySQL C# PHP HTML CSS JavaScript jQuery

## 기본사항

생년월일 : 1989. 07. 11  
주소 : 서울시 성동구 성수동2가 299-134  
이메일 : dooyeoung.ryu@gmail.com  
연락처 : 010-5061-2576  
웹사이트/블로그: <https://dooyeoung.github.io>

## 학력사항

2008. 02-2015.02 한림대학교 (컴퓨터공학 학사)  
2018. 01-2018.05 패스트캠퍼스 (데이터 사이언스 스쿨 7기)

## 경력사항

2015. 02-2017.12 아라엔지니어링 IT Team, 대리  
software 개발

- OTS(Operator Training Simulator) software 개발
  - HYSY, DeltaV, Synergi 제어 프로그램 개발
  - HMI instructor graphic module 개발
- PDMS utility software 개발
  - \*.rvm file converter 개발

## 자격사항 및 수상내역

2014. 08 정보처리기사 (한국산업인력공단)  
2007. 11 리눅스마스터 2급 (한국정보통신산업협회)  
2007. 01 E-Test Professional (한국산업인력공단)  
2005. 11 ITQ-Master (한국생산성본부)

## 보유 기술

능숙 Python, C#, MySQL, HTML  
활용가능 C++, Java, PHP, JQuery, Flask

## 프로젝트

Bike Sharing Demand (2018. 05-2018.05, 개인 프로젝트)  
날씨와 시간 데이터를 통해 자전거 대여량 예측  
[https://github.com/dooyeoung/bike\\_sharing\\_demand](https://github.com/dooyeoung/bike_sharing_demand)

- Data Preprocessing
  - 날씨별, 시간별 데이터 파악
- Data Analysis
  - Randomforest를 사용하여 분석
  - Gradient boost regressor를 사용하여 분석
  - Xgboostregressor를 사용하여 분석

Recommend coordination (2018. 04-2018.04, 팀프로젝트)  
웹 코디 이미지 색상 분석을 이용한 coordination 추천

[https://github.com/dooyeoung/codi\\_recommendation](https://github.com/dooyeoung/codi_recommendation)

- 데이터 크롤링 및 데이터베이스 구축
  - Phantomjs로 Mapssi 코디 데이터 크롤링
  - 이미지 주소, 색상 데이터 MySQL 저장
- 이미지 분석
  - RGB clustering으로 대표색상 추출
  - KNN으로 색상검색기능 구현
  - MLP로 색상추천기능 구현
- Flask 구축
  - bootstrap으로 레이아웃 구성
  - RGB값을 KNN을 사용하여 검색기능 구성

---

Walmart : Trip Type Classification (2018. 03-2018.04, 팀 프로젝트)  
구매 품목과 수량으로 고객을 38개로 분류

[https://github.com/dooyeoung/walmart\\_recruiting\\_trip\\_type\\_classification](https://github.com/dooyeoung/walmart_recruiting_trip_type_classification)

- Data Preprocessing
  - 상품 Missing value 처리
  - Type, 고객별 구매데이터 취합
- Data Analysis
  - Xgboost를 사용하여 classification 분석
  - LightGBM을 사용하여 classification 분석

---

Walmart : Sales in Stormy Weather (2018. 02-2018.03, 팀 프로젝트)  
날씨를 토대로 각 매장의 아이템의 판매량을 예측

[https://github.com/dooyeoung/walmart\\_recruiting2\\_stormy\\_weather](https://github.com/dooyeoung/walmart_recruiting2_stormy_weather)

- Data Preprocessing
  - weather data missing value 처리
  - train data store 별로 데이터 취합
- Data Analysis
  - OLS을 이용하여 Regression 분석

---

OLGA Extension 개발 (2017.05-2017.12, 아라엔지니어링/IT)  
UnisimDesign, OLGA 제어 프로그램 개발

- Unisim, Olga, Synergi 제어프로그램 개발
  - OPC Client 기능 구현
  - TCP/IP 기반으로 제어 모듈 개발

---

OTS HMI 개발 (2016.07-2017.05, 아라엔지니어링/IT)  
OTS 그래픽 제어 프로그램 개발

- HMI component 개발
  - GDI+ 선, 도형, 이미지 그리기, 그래프 모듈 기능구현
  - MariaDB 적용하여 데이터 관리

---

OMGP OTS 프로젝트 (2015. 08-2016.07, 아라엔지니어링/IT)  
OTS 제어 프로그램 개발

- 리팩토링
  - 메모리 사용 최적화
  - 디자인패턴, 클래스 재구성
  - MariaDB를 통한 데이터 관리

---

Chart component 개발 (2015.01-2015.07, 아라엔지니어링/IT)  
Component기획 및 개발

- C# chart component 개발
  - GDI+를 사용하여 개발 진행
  - Multiple Axis 기능 구현
  - 드래그를 이용한 Zoom 기능 구현
  - Mouse tracking 기능 구현

---

빈강의실 자리예약 시스템 (2017.03-2017.05, 개인 프로젝트)  
기획 및 개발

- PHP로 도서관 자리예약 시스템 개발
  - Bootstrap으로 레이아웃 디자인
  - 로그인기능 구현
  - 시간단위 예약기능 구현
  - 예약 연장 기능 구현

---

설문조사 서비스 (2015. 11-2015.12, 개인 프로젝트)  
기획 및 개발

- php를 이용한 설문조사 서비스 구현
  - Bootstrap으로 레이아웃 디자인
  - 설문양식 조회, 작성 기능 구현

## Q. 1) 데이터 분석과 관련된 경험

C# 프로젝트 진행 중 데이터베이스를 직접 설계하고 초당 10,000개 이상의 데이터를 처리하는 모듈 개발을 담당했었습니다. 데이터베이스 설계부터 관리를 직접 관리하며 데이터를 처리하였고 SQL 튜닝과 쿼리결과를 실시간으로 표시하는 모듈이었습니다.

직접 개발한 모듈과 기초적인 통계지식을 이용하여 데이터 현황을 보고할 계기가 있었습니다. 이 과정에서 데이터 분석에 흥미를 갖게 됐으며 전문적인 통계분석의 필요성을 다시 한번 느꼈습니다.

가장 먼저 데이터 분석 기초를 공부했습니다. 3개월간 인터넷 강의를 통해 선형대수, 미적분, 확률과 통계 기반을 다졌습니다. 데이터 분석 분야를 알아 갈 수록 욕심이 생겨 제대로 공부하기 위해 퇴사 후 패스트캠퍼스 데이터사이언스 스쿨 교육을 수강하여 5개월간 선형대수, 확률과 통계, 미적분, python의 numpy, pandas 활용방법, 딥러닝 등 전반적인 내용을 배웠습니다.

학습한 내용을 실제 데이터 분석에 적용해 보고자 kaggle 데이터 분석 프로젝트를 진행하였습니다. Walmart : Sales in Stormy Weather를 주제로 선정하였습니다. 기상 데이터를 통해 판매량을 예측하는 주제였습니다. 회귀분석으로 접근하였지만 기상데이터와 판매량간에 관계성 찾을 수 없어 분석에 큰 어려움이 있었습니다. 더 나은 분석을 위해 휴일, 블랙프라이데이 데이터를 추가하여 다시 회귀분석을 했지만 결과는 별반 다르지 않았습니다. 팀원들과 오랜 회의 끝에 낱짜 데이터에 집중하여 시계열 분석으로 방향을 바꾸어 더 나은 결과를 얻을 수 있었습니다.

다음으로 Walmart Recruiting: Trip Type Classification 주제를 선정하여 팀프로젝트를 진행했습니다. 물품의 분류와 구매수량, 반품수량으로 고객을 38개 type으로 분류하는 문제였습니다. 소수의 독립변수를 분석하여 type별 특징을 찾기위해 의미있는 독립변수로 변화하는 과정에서 많은 시간을 소비했습니다. 그 후 decision tree, random forest, XGboost, lightgbm model을 비교하여 더 나은 데이터 분류 결과를 얻었습니다.

데이터 분석을 모르던 평범한 개발자에서 무수히 많은 데이터 속에서 원하는 정보를 찾을 수 있는 데이터분석가로 거듭날 수 있었던 시간이었습니다.

## Q. 2) 새로운 것을 접목하여 문제를 개선했던 경험

아라엔지니어링 재직 당시 PDMS 프로그램의 3D file을 변환하는 업무를 맡았습니다. 기존 작업 방식은 PDMS 프로그램에서 추출된 1000개 이상 \*.rvm 3D file을 Autodesk의 Navisworks 프로그램을 통해 \*.fbx 3d file로 변환하는 작업이었습니다.

1000개의 파일을 한번에 하나씩 변환 할 수 밖에 없기에 많은 시간을 소비한다는 문제점 때문에 3D file 변환작업을 개선하기로 했습니다.

먼저 \*.rvm file을 분석하였습니다. \*.rvm file을 일주일동안 분석하여 3D 정보를 모두 추출할 수 있었습니다. 추출한 내용을 DB에 저장하여 원하는 정보를 추출 할 수 있게 하여 생산성을 높였습니다. 분석한 3D 정보는 fbx sdk를 사용하여 \*.fbx 로 변환 할 수 있도록 프로그램을 개발 하였습니다.

한달 이상의 시간이 소비 될거라 예상했던 작업을 데이터 파일 분석과 데이터베이스 활용 그리고 변환 프로그램 개발로 작업 시간을 많이 줄일 수 있었습니다.

데이터 파일을 분석하는 과정을 통해 raw데이터의 충분한 이해와 도메인 지식이 굉장히 중요하다는 점을 배웠습니다.

## Q. 3) 성장 과정

매사에 "쉽고 빠르게 같은 일을 해결할 수 있을까?"라는 생각을 하는 제 자신을 발견했습니다. 계획을 수립하여 일을 수행하는 과정을 몇 차례 반복하다 보니 효율적인 방법을 생각하는 시야를 갖게 되었습니다.

이러한 영향으로 효율을 생산하는 프로그래밍에 자연스레 관심이 생겼습니다. 처음으로 만든 프로그램은 시간표 비교 프로그램입니다. 수강을 원하는 과목을 입력하여 과목의 우선순위에 어떤 시간표가 최적의 시간표인지 결과를 출력하는 프로그램이었습니다. 이때 개발한 모듈로 저뿐만 아니라 친구들이 원하는 시간표를 미리 얻을 수 있었습니다. 나아가 효율을 생각해 무언가를 만들었다는 보람감을 얻을 수 있었습니다.

두번째 프로젝트는 학기 시작 전에 자취방을 구해야하는 상황에서 2주가량을 직접 방을 보고 인터넷에서 부동산 정보를 찾으며 자취방을 힘들게 구했던 적이 있습니다. 이때 한 고생을 반복 하고 싶지 않았으며 정보만 정리해도 다른 사람들에게 큰 도움이 되리라 생각했습니다. 부동산 중개인들과 얘기하여 원룸, 오피스텔 정보 뿐만아니라 온라인 커뮤니티의 정보를 한데 모아 학교주변 자취방 정보 100개 정도를 모을 수 있었습니다. 자취방 정보를 정리하여 데이터베이스에 저장하여 "학교주변 자취방정보검색" 프로젝트를 완성 하였습니다. 완성 후 정보갱신 문제로 유용한 서비스로 자리매김 하지 못하였지만 데이터를 수집하고 정리하는 가공하는 과정에 흥미를 느꼈습니다.

그 후로 교내에 설문조사를 진행 해야 할 상황에서 대면 설문조사방식을 개선하기 위해 서비스를 웹프로그래밍으로 개발했습니다. 이 프로젝트에서 설문조사 데이터를 기초적인 통계분석으로 의미있는 정보를 찾을수 있어 재미와 보람을 느낀 경험이었습니다.

## Q. 4) 성격 장단점

"할 수 있다", "꾸준함" 이 두 말로 제 성격을 말씀드릴 수 있습니다. 9살 무렵 두발자전거를 타기 위해 10일 동안 넘어지고 또 넘어지면서 자전거를 탈 수 있게 되었고 "할 수 있다" 라는 것을 처음 배우게 되었습니다. "할 수 있다" 와 "안되면 될때까지"를 몸으로 얻은 기억입니다. 그 뒤로 새로운 일에 도전할 때면 "할 수 있다" 와 "꾸준함"을 바탕으로 시작했습니다.

이런 성격은 코딩을 배울 때 그리고 프로젝트를 진행할 때 큰 도움이 되었습니다. 개발언어를 습득하기 위해서는 꾸준함이 필요했으며 프로젝트를 시작하기 위해서는 할 수 있다는 마음가짐으로 시작했습니다. 나아가 기능을 구현할 때는 안 되면 될 때까지라는 생각으로 프로젝트를 진행했습니다.

할 수 있다는 통해 얻은 도전정신과 꾸준함을 통해 얻은 성실함이 제 성격의 장점입니다.

성격이 내성적인 편이라 시선이 집중되는 것이 부담이었습니다. 거절도 잘 못 하고 주장도 잘 펼치지 못했습니다. 내성적인 성격이 불편하였으며 개선이 필요하다고 생각하고 있었지만 변화하기 힘들었습니다. 사회성을 키울 수 있는 첫 번째 계기는 군시절 다양한 사람들과 어울리면서 외향적인 사회성을 배울 수 있었습니다. 제대 후 학생회 활동을 통해 더 많은 사람들과 마주하면서 다양한 사회성을 키울 수 있었습니다. 제 능력을 발휘 할 수 있는 역할을 맡은 덕분에 자신감을 많이 얻게 되었으며 사회적 능력을 개발하였기에 내성적인 성격이 더 이상 불편하지 않았습니다.

사회적 능력을 개발한 후에는 내성적인 타인에게 먼저 다가가 편안하게 대할 수 있습니다. 제 자신이 내성적인 사람이었기에 타인을 배려하고 이해할 할 수 있는 덕목도 단점을 극복하는 과정에서 얻게된 부분입니다.

## Q. 5) 협업을 통해 문제를 해결한 경험

Walmart : Sales in Stormy Weather 분석시 팀프로젝트로 진행했습니다. walmart 2012년부터 2014년사이의 45개 지점별 111개 품목의 판매량이 날씨데이터의 영향력을 토대로 판매량을 예측하는 문제였습니다.

팀구성은 통계학 전공자와 팀을 이뤘습니다. 통계학을 전공한 팀원은 통계 분석의 흐름을 알고있지만 프로그래밍 경험이 부족했습니다. 그렇기에 파이썬을 이용한 분석은 서툴렀습니다. 반면 컴퓨터공학을 전공한 저는 프로그래밍이 가능했지만 통계분석 경험이 없었습니다. 서로가 서로에게 도움을 주며 분석을 진행했습니다. 통계분석의 흐름을 제시하면 그에 맞는 분석방법을 프로그래밍으로 구현하여 분석을 한단계 한단계 진행했습니다. 분석 시작전 처음 한 일은 날씨데이터의 누락값을 처리하는 작업부터 진행했습니다.

날씨데이터는 계절성을 갖기때문에 그래프로 흐름을 본 후 3일 간 데이터의 평균값으로 데이터를 복구했습니다. 너무 허술한 시계열분석을 이요한 처리라 생각했습니다. Moving Average 와 facebook의 prophet패키지를 사용하여 누락값을 처리했습니다.

프로젝트를 진행하면서 이론으로 배운 회귀분석을 실제 데이터에 어떻게 적용해야 하는지 어떻게 회귀모델을 만들어야 하는지, 만든 모델의 성능평가는 어떻게 해야하는지를 다시 공부 할 수 있었습니다. 추세와 계절성을 갖는 데이터는 시계열분석으로 접근해야한다는 것도 실제 데이터를 통해 다시 한번 깨달았습니다. 팀 프로젝트를 통해서 서로 의논하며 지식을 공유했기에 위 프로젝트를 진행하면 데이터 분석능력을 키울 수 있었습니다.