|  |  |
| --- | --- |
| Birthday | 1996.11.18 |
| Email | once9891@gmail.com |
| Mobile | 01020733148 |
| Address | 서울특별시 금천구 |
| 병역사항 | 필 |

**이름 김경수**

**데이터분석가 , 통계**

**소개 / About Me**

넓은 시야를 가지고 있으며 , 논리적인 검증을 통해 결과를 향해서 거침없이 달려가는 진취적인 사람입니다.

**학력**

* 2015.03.02 ~ 2023.08.21 수원대학교 데이터과학부 졸업

**자격증**

* ADSP ( 2023.03.22 취득 )
* 빅데이터분석기사( 2023.12.22취득 )
* SQLD ( 2023.12.15 취득 )

**교육 내용**

* 2023 - 07 - 19 ~ 2024 - 01 - 13 빅데이터분석가 양성과정 ( 한국품질재단 )

**수상 내역**

* 2009 중등수학 올림피아드 장려상
* 빅데이터분석가 양성과정 최종 프로젝트 대상

**어학 능력**

* TOEIC 600점

**기술 스택 / Skill Set**

**기능 구현 등의 사용 경험이 있는 Skill Set**

|  |  |
| --- | --- |
| 구분 | Skill |
| Programing Languages | Python : 데이터 분석 및 다양한 시각화  다양한 클래스와 함수 정의 및 기본적인 로직 구현  R : 워드클라우드 시각화  DJango : Python에서 만든 로직을 이용하여 Django에 장착 후 웹서비스 구현 가능 |
| Framework/ Library | numpy : 배열을 이용하여 자유롭게 인덱싱 가능  Pandas : 분석한 자료들을 합치거나 나눠서 xml , csv로 출력가능  Pytorch : Pytorch를 이용하여 딥러닝 클래스 구현 및 모델링 가능  Sklearn : 다양한 모델을을 기반으로 효율적인 분류와 예측  matplotlib , Seaborn : 분석한 자료들의 상관관계 , 대소관계를 보기 쉽게 시각화  PIL : 이미지를 DataFrame 같은 곳에 추가 가능 |
| Server | SQL : SQL에서 기본적인 Model을 만들고 , Python으로 연결하여 Python에서 DataBase 조작  SQLite |
| Tooling/ DevOps |  |
| Environment |  |
| ETC | PPT , Word |

**과제 프로젝트 경험** *\* 혼자 혹은 팀 구성으로 프로젝트를 수행한 경험을 최신순으로 작성*

**프로젝트 요약(제목)**

|  |  |
| --- | --- |
| 작업 기간 | 2023.12.11 ~ 2024.1.12 |
| 인력 구성 | 팀장(본인) , 팀원 , 팀원 , 팀원 |
| 프로젝트 목적 | 직장인들이 병원을 갈 시간이 부족하여 , 자신의 머리 상태를 인터넷만 된다면 가볍게 자가진단 할 수 있는 웹서비스 구현 |
| 주요업무  및  상세역할 | - AI 허브에 있는 유형별 두피 이미지를 통해 4개의 클래스로 분류하는 모델을 생성  - 웹서비스를 이용하는 이용자들의 연령 , 성별 , 두피 이미지 그리고 결과를 이용하여 SLITE3파일로 데이터 저장 후 서버에서 DB를 관리하도록 하는 시스템 구현  - Python 코드와 웹서비스를 하기위한 사이트와 연결 중 경로 문제 해결  -> os path 이용 |
| 사용언어  및 개발환경 | - Window 10 64bit  - Google Colab  - Django  - SQLite3  - Python |
| 느낀 점 | 딥러닝 모델 구현 중 , 가장 큰 벽이었던 것은 GPU자원의 제한이었습니다. Integer , Catagory 데이터들을 분류 및 회귀하는 데에는 문제가 없었지만 , 이미지 분류는 메모리와 성능에 크게 영향을 받았기 때문에 , 기본적인 CPU사양으로는 어려움이 있었습니다. 이 문제를 해결하기 위하여 Google Colab Pro+를 구매하여 주어진 500컴퓨팅단위에서 최대의 성능을 뽑아내었습니다.  Colab에서 모델을 구현한 후 , Pycharm으로 옮겨서 사용자들의 이미지를 이용하여 Predict하는 것을 구현하고 , Jango와 연결하는 과정이 마치 퍼즐처럼 하나하나 맞춰졌습니다. 하나하나 맞추는 과정은 경로문제나 변수문제에 어려움이 있었지만 , os의 절대경로를 이용하여 이를 해결했을 때가 이 프로젝트 과정중에 가장 재미있었던 부분이었던 것 같습니다. |
| 참고자료 | <https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=216>  https://github.com/FEELGOODRAIN/KimGyeongSu/blob/main/Deep%20Learning/project/ScalpSpectra%20%EB%B0%9C%ED%91%9C%EC%9E%90%EB%A3%8C.pdf |

**프로젝트 요약(제목)**

|  |  |
| --- | --- |
| 작업 기간 | **2023-10-10 ~ 2023-10-20** |
| 인력 구성 | 팀장(본인) , 팀원 , 팀원 , 팀원 |
| 프로젝트 목적 | 서울시 안에 있는 장애인들의 편의시설 현황을 분석하여 취약점 , 개선점 제시 |
| 주요업무  및  상세역할 | - 서울시 전체에 있는 지도를 가져와서 어디에 장애인 편의시설이 있는지 Marker로 구현  - 장애인 시설들의 현황 csv를 이용하여 어떤 키워드가 가장 많이 다뤄졌는지 WordCloud로 구현  - 편의시설의 개수 및 분포들을 다양하게 시각화 |
| 사용언어  및 개발환경 | -Window 10 64bit  - Python  - Google Colab |
| 느낀 점 | 모델링만 하는게 아닌 단순히 데이터를 분석하고 이 프로젝트를 보는 비전공자인 시청자들을 이해시키는 과정이 많이 필요했습니다. 그래서 표를 그릴 때 , 최대한 전문용어를 줄이고 수치로만 표현하여 한눈에 보면 분포를 알 수 있게끔 구현했습니다. 지도시각화에서 Marker를 찍을 때 , 위도 경도로 변환하는 과정에서 살짝 어려웠는데 , geopy 패키지를 이용한 함수를 구현하여 도로명주소를 위도 , 경도로 바꾸어서 해결했습니다. 이 프로젝트에서는 도로명주소를 위도 , 경도로 변환하는 함수가 있다는 것을 알게 되어 신선했습니다. |
| 참고자료 | <https://github.com/FEELGOODRAIN/KimGyeongSu/blob/main/Machine%20Learning/Project/%EC%A7%80%EB%8F%84%EC%8B%9C%EA%B0%81%ED%99%94%20%ED%94%84%EB%A1%9C%EC%A0%9D%ED%8A%B8.ipynb>  <https://github.com/FEELGOODRAIN/KimGyeongSu/blob/main/Machine%20Learning/%ED%8F%AC%EB%A0%88%EC%8A%A4%ED%8A%B8%20%ED%94%84%EB%A1%9C%EC%A0%9D%ED%8A%B8.pptx> |

자기소개서

**[자기소개] 본인에 대해 자유롭게 기술해주시기 바랍니다.**

수치들을 가지고 논리적으로 접근하고 분석하여 결과를 도출해내는 것을 좋아하는 사람입니다.

초등학교 때부터 수학을 좋아하고 , 그리고 수학적 처리능력이 주변 친구들보다 뛰어난 편이었기 때문에

고등학교 때까지 수학쪽에서는 바라보는 시선이 다른 친구들보다 넓었습니다. 그래서 대학교도 수학과로

입학했었습니다. 수학과에서 교육과정을 이수하다가 전역하고 복학을 해보니 수학과가

데이터과학부라는 학부로 흡수가 되었습니다.

그래서 수학과에서 갑자기 데이터과학부로 바뀌어 3학년 , 4학년 2년동안 1학년부터 4학년까지의 과정

을 수료해야 했었습니다. 여기에서 2년동안 데이터분석 관련 강의들을 들었습니다.

데이터과학을 배워보니 수학 , 통계와 밀접한 관련이 있었습니다. 저는 수학을 좋아하기도 하고 , 결과까지

도달하는 과정을 좋아하기 때문에. 데이터분석도 여러 줄 코딩하여 실행했을 때 결과가 나오는 쾌감

이 있는 것 같습니다. 수학에서도 증명까지의 과정에서 치밀하게 틀을 만들어야 하기 때문에. 이 부분에서

는 데이터의 정제 및 분석과 연관이 있다고 생각했습니다. 교육과정 이수 중에 '지도시각화' 라는 부분을

해봤었는데. 단순히 돌아다니는 데이터를 가져와서 하는 것이 아니라. 직접 데이터들을 정제하고 점점

나아지는 시각화 모델들을 보면서 뿌듯함을 느꼈습니다.

수학에서의 증명과 데이터의 모델링은 중간에 어디에서 잘못되면 값이 아예 틀어져버리는 공통점이 있습

니다. 그래서 하나의 작업을 수행할 때 꼼꼼하게 검토해야 하고, 다른 팀원들의 시선을 빌려 알아채지

못한 개선점들을 찾아가나가는 것이 재미있었습니다.

**[성격기술] 본인의 성격의 장 · 단점을 기술해주시기 바랍니다.**

성격을 고체 , 액체로 분류한다면 , 액체에 가까운 성격입니다. 그래서 기본적으로 고정관념이 없고 , 틀에 박혀있지 않아서 모든 의견을 수용하고 , 실천할 수 있습니다. 서로의 의견들을 모아서 저의 의견에 전혀 맞지 않더라도 옮은 의견이 제시되었을 때 , 화내지 않고 더 좋은 방향으로 함께 가는 성향입니다.

대학교 때 , 응원단장을 했었는데 , 이때 응원단이 해체될 위기에 처했었습니다. 이유는 밤늦게까지 하는 연습에 비해 성취감이 상대적으로 적었기 때문입니다. 이런 위기를 극복하기 위해 단원들과 힘을 합쳐서 새로운 프로젝트(안무)를 만들고 , 공연을 다니면서 성취감을 올리고 , 응원단의 홍보도 추가적으로 해서 다음 기수의 단원 희망자들까지 생겨 해체위기에서 벗어날 수 있었습니다.

위와 같이 어떠한 팀에 속해있을 때 , 팀이 하향세라면 다른곳으로 도망치지 않고 , 될때까지 끊임없이 도전하는 도전정신이 있습니다.

단점은 한 프로젝트에 착수하면 그 프로젝트가 완성될 때까지 놓지 않으려 하는 경향이 있습니다. 머신러닝으로 프로젝트를 만들 때 , 텍스트마이닝을 해서 워드클라우드를 구현하려 했는데 , 생각대로 잘 되지 않고 결과의 시각화가 모자랐었습니다. 모자랐던 이유는 패키지와 모듈에 대한 지식 부족이었다고 생각하는데 , 그때는 그런 걸 몰랐기 때문에 , 어떻게 하면 시각화를 더 잘 표현할 수 있을지 그 주제 하나만 가지고 이틀동안 끌었었습니다.

이러한 단점을 개선하기 위하여 일의 우선순위를 정해놓는 습관을 기르고 있습니다. 중요한 일을 먼저 하고 , 그런 다음 탐구적인 활동은 조금 뒤로 미뤄놓아서 업무의 효율을 높이는게 팀원들과도 서로 윈윈하는 방법이라고 생각합니다.

**[역량기술] 본인이 관심 있는 IT 분야 혹은 관련 경험을 자유롭게 기술해 주십시오**

**.**

빅데이터분석가 양성과정을 이수하는 동안 프로젝트를 5개 했었습니다. 첫 프로젝트는 R을 활용한 프로젝트였는데 , 여기에서는 데이터 수집과 정제 역할을 했었습니다. 주제를 선정하고 , 그 주제에 맞는 데이터 수집, 그리고 데이터를 가져오고 시각화 및 모델링에 적합하게 정제하는 역할이었고 , 팀원 3명이서 서로 부족한 점을 보완하면서 프로젝트를 완성했습니다.

두 번째 프로젝트는 개인 프로젝트였고 , SQL과 Python을 연동한 시스템 만들기입니다. 여기에서는 SQL에 사원의 기본정보에 대한 DataBase를 생성해서 파이썬 환경으로 가져오고 , 파이썬에서 기본적인 함수와 알고리즘을 만들어 정보를 입력하면 서버 내에 있는 DataBase에 저장하고 , 삭제하고 , 수정하는 작업을 수행하게 했습니다.

세 번째 프로젝트는 팀 프로젝트였고, 데이터 시각화에 대한 프로젝트였습니다. 이 프로젝트가 비전공자가 보기에도 쉽게 해석이 가능하도록 만드는 프로젝트였기에 , 아주 재미있게 했습니다. 빅데이터캠퍼스에서 여러 데이터셋들을 가져와 데이터들의 전처리를 하고 , 전처리한 데이터들을 목적에 맞게 합치고 나눠서 시각화에 용이하게 하는 역할을 했습니다. 그리고 이 프로젝트는 제가 팀장이었고 , 팀원이 3명 있었는데 , 각자의 역할을 효율적으로 나누고 그 역할에서 막히는 쪽에 대한 비전을 제시해주면서 본격적인 팀 프로젝트에 대해 배우는 계기가 되었고 , 제 역할인 지도시각화와 워드클라우드를 처음 시도해보는 프로젝트가 되었습니다.

네 번째 프로젝트는 개인 프로젝트였고 , 머신러닝으로 모델링하는 프로젝트였습니다. 주제가 완전 자유였기 때문에. 머신러닝으로 회귀 및 분류모델을 구축하고 , 이 모델을 기반으로 솔루션을 제공하는 순서로 진행했습니다. 제가 했던 프로젝트는 여러 곳에 있는 물들을 가져와서 그 물 안에 있는 갖가지 요소들을 활용하여 이 물이 먹을수 있는지 없는지를 분류하는 것이었습니다. 여기에서는 물이 먹으면 안되는데 , 먹어도 된다고 결론을 내버리면 큰일나기 때문에 . 이런 척도를 조절하는 작업도 수행했습니다.

마지막 프로젝트는 팀으로 하는 최종 프로젝트였고 , 딥러닝을 포함하여 지금까지 배운 것들을 모두 활용하는 프로젝트였습니다. 주제는 두피 스펙트럼(두피 진단)이었습니다. 저는 이 팀의 팀장이었고 , 주제를 정한 후 역할을 나누는 과정에서 일반적인 팀프로젝트처럼 발표 , ppt , 수집 이런 것을 나누지 않고 , 각자 자신 있는 부분들을 조사 후 , 역할을 잘하는 부분으로 나누어 업무의 효율성을 높여주었고 , 제작 과정에서 팀원들의 지식을 공유하는 회의시간을 가졌습니다.

최종 프로젝트를 제작하는 과정에서는 어려움이 다른 프로젝트보다 많았습니다. 이미지를 이용하여 딥러닝을 수행하는 모델이다 보니까 , 컴퓨터 자체의 성능 문제가 있어서 , Colab Pro+를 결제해서 진행했었는데 , Colab Pro+보다 좋은 GPU를 제공받으면 모델의 성능을 더 높일 수 있었을텐데 자원이 제한적이어서 Colab Pro+에서 주어진 자원으로 최대의 효율을 낼 수 있게 모델링을 수행했습니다.

두 번째 어려움은 웹서비스를 구현해야 하는데 , Python 파일과 모델들을 Jango에 부착하는 작업이 어려웠습니다. 웹 서버는 다른 디렉토리를 사용했기 때문에 , Pycharm 내에서 경로를 설정했던 데로 하면 인식하지 못했습니다. 그래서 다방면으로 탐구해본 결과 , os 패키지를 이용하여 그 디렉토리의 절대경로를 활용하여 안정적으로 Jango에 부착하는 데에 성공했습니다.

최종 프로젝트는 팀으로 하는 프로젝트에서 서로의 지식을 나누는 회의와 각자의 특기를 살려 역할 분담을 해서 업무효율을 높이는 데에 중점을 두었기 때문에 , 현업으로 나가기 전 좋은 팀 프로젝트 방향성이 되어주었습니다.

**[지원동기] 지원하게 된 구체적인 이유는 무엇이며, 입사 후 목표는 무엇인지 기술해주십시오.**

다양한 니즈의 변화를 수용하고 도전하여 성장하고 싶습니다.

지금까지 배우고, 앞으로도 배울 데이터분석 기술을 활용하여 제가 하고싶었던 것은. 수치형 , 논리형으로 되어있는 데이터들을 가공 , 정제하여 현재 상황이 어떠한지 , 그리고 미래에 어떻게 될지 , 무엇이 필요할지 등을 전공자가 아닌 누구나 보기 편하게 만들어서 업무의 효율성을 높이는 일을 하고 싶었습니다. 공고의 담당업무를 보면 고객사가 어떤 것을 원할 때 , 그것을 알맞게 풀어서 설명해주는 업무를 담당하는 것으로 이해했습니다. 그리고 수학을 기본적으로 전공했기 때문에 이 회사에서 원하는 담당업무를 아주 알맞게 소화할 수 있을 것 같아서 지원했습니다.

그리고 입사를 하게 되어 업무를 할 때 , 다른 사람들과 팀을 이뤄서 변화되는 문화에 빠르게 적응하여 최고의 솔루션을 도출하는 팀원이 되고 싶습니다.

**1분 자기소개**

안녕하세요 , OOO에 지원한 OOO입니다.

저는 세상에서 물같은 사람이 되고 싶습니다.

제가 물이라고 비유한 이유는 물은 모든 곳에 필요하고 , 어디든지 갈 수 있고 , 어디에서든 흡수될 수 있습니다. 그리고 가장 중요한 점은 물도 고이면 썩습니다. 데이터쪽 업무는 특히 흐르지 않으면 급격히 썩고 물로써의 가치를 결국 잃어버리게 됩니다.

빅데이터분석가 양성과정에서 장애인 편의시설 프로젝트를 진행할 때 , 지금까지 배워왔던 모든 것을 활용하여 수분이 이리저리 스며드는 것처럼 여러 부분으로 탐구하는 시간을 가졌었습니다. 이 시간을 가짐으로 인해서 워드클라우드와 지도시각화 등 그래프뿐만이 아닌 지도와 워드클라우드를 이용하여 이해를 보다 쉽게 할 수 있었습니다.

저는 움직임을 멈추지 않고 어디든지 탐험하고 도전하며 온도가 높아지면 하늘도 가보고 낮아지면 엄청나게 큰 바다에 잠겨도 보고 하는 그런 변화 가능성이 무궁무진한 사람이 되고 싶습니다.