

이력서

각오 한마디	기본에 충실한 자세로 꾸준한 성과를 내겠습니다.				
지원분야	데이터 분석	지원구분	신입	희망연봉	회사 내규에 따름

☐ 기본정보

	성명	권용진	성별	남
	생년월일	2001. 05. 02	나이	26 세(만 25 세)
	주소	강원도 횡성군 횡성읍 북천서로 19 번길 9		
	휴대번호	010-2380-4683	E-Mail	qxas4550@naver.com
	병역사항	병역必(2021.05.31~2022.11.30)	취미	풋살, 야구(관람), 게임, 헬스

☐ 학력사항

기간	출신학교	전공	학위	학점
2020.03 ~ 2026.02	수원대학교	데이터 과학	학사	4.1/4.5
2017.03 ~ 2020.01	횡성고등학교	이과	-	-

☐ 보유기술

분야	보유기술	역량 소개
데이터베이스	SQL	MIMIC-IV 데이터에서 환자·입원·진단·ICU 테이블을 추출·결합하고 분석용 데이터셋 구성
프로그래밍	Python	pandas·NumPy 기반 데이터 전처리, 머신러닝 모델링, 웹 크롤링 및 시각화 수행
통계 분석	R	다중회귀분석, 상관관계 분석, 통계적 유의성 검정 및 지도 시각화 수행
공간 데이터 분석	QGIS	격자 생성, 좌표계 통일, 공간 조인 및 머신러닝용 공간 데이터셋 구축
데이터 시각화	Tableau	지역별 주요 지표와 분석 결과를 지도·차트 형태로 시각화
데이터 시각화	Power BI	이커머스 퍼널, 전환율, 매출·마진, 고객 세그먼트 성과 대시보드 구축
업무 자동화	Make	사용자 입력 기반 결과 생성·저장·전달 자동화 워크플로우 구축

☐ 교육이수내역

연수기간	연수기관	연수과정 및 내용
2025.08.26 ~ 2026.02.11 (6개월/920시간)	아이티윌	프로젝트 중심 빅데이터 융합 머신러닝 전문가 양성과정 - 무신사 상품 크롤러 / FotMob 데이터 크롤러 - 머신러닝을 활용한 경기도 싱크홀 위험도 예측 및 위험지도 구축

		- Willpang Growth Analytics: AARRR 기반 이커머스 성장 전략 설계 및 시뮬레이션 - 데이터분석 부트캠프 강의노트
2026.07.06 ~ 2026.08.16	데이터리언	GA4 데이터 분석 캠프 - 티스토리 블로그에 GA4·GTM 을 연동하고 사용자 행동 이벤트 수집 - 티스토리) GA4 캠프

□
 자격사항

취득일	종류 및 등급	발행처
2024.03.22	ADsP	한국데이터산업진흥원
2024.09.20	SQLD	한국데이터산업진흥원
2025.08.19	컴활 1 급	대한상공회의소
2026.06.12	정보처리기사	한국산업인력공단
2026.07.10	빅데이터분석기사	한국데이터산업진흥원
2026.06.30 ~ 2027.06.30	Google Analytics Certification	Google

□
 수상내역

행사명	시상일	팀명 / 제목	상세 설명
제 2 회 AI/빅데이터 분석 경진대회 - 화성시연구원 주관	최우수상 2025/11/20	유방암 환자의 예후 예측을 위한 병리 이미지와 전사체 기반 멀티모달 딥러닝 모델	병리이미지(WSI)와 전사체(Transcriptome)를 통합한 멀티모달 딥러닝 기반 유방암 예후 예측 모델(VA-GATSurv)을 제안하여, 단일 모달 대비 생존 예측 성능(C-index)을 유의미하게 향상시킨 연구

□
 아르바이트 및 기타활동

활동기간	소속명	담당	상세 설명
2024.05 ~ 2024.11	DNA(학부 내 동아리)	C2 팀원	- 코호트 분석 개념 블로그에 기록 - boostcourse 데이터 시각화를 위한 태블로 강의 수강 후 블로그에 기록 - 공공데이터분석 전공 수업 실습 자료 기반 Tableau 시각화 실습 - 티스토리) DNA C2 팀

□
 참여프로젝트

개발기간	구분	상세설명
24.03.15 ~ 24.06.05 (약 3 개월)	프로젝트명	▶ MIMIC-IV 데이터 분석을 통한 인사이트 도출 및 사망률 예측 5 명 팀프로젝트(수원대학교 학부 프로젝트)
	분석목적	- ICU 환자의 사망 위험 요인 분석 및 고위험 환자 식별을 위한 사망 예측 모델 구축

	분석방법	· SQL 기반 환자·입원·진단·ICU 테이블 추출 및 결합 · Python 을 활용한 데이터 전처리 및 Feature Encoding · Logistic Regression, Random Forest, LightGBM 등 모델 비교 · 교차검증과 ROC-AUC 기반 모델 성능 평가 · Voting Classifier 기반 최종 예측 모델 구축
	주요결과	✓연령·입원기간·체중·위험점수를 주요 사망 위험 요인으로 확인하고, 최종 Voting Classifier 가 테스트 데이터에서 약 91.6%의 정확도 기록
	담당업무	· 데이터 수집 및 전처리 · 머신러닝 모델링 및 시각화 · 분석 결과 정리 및 PPT 제작
24.06.17 ~ 24.08.30 (약 2 개월)	프로젝트명	▶ 경제 수준 차이에 따른 도서관 이에 미치는 영향 분석 4 명 팀프로젝트(도서관 데이터 활용 공모전)
	분석목적	- 지역 경제 및 도서관 운영 환경이 회원등록자 수에 미치는 영향 분석과 지역 간 이용 격차 원인 도출
	분석방법	· 국가도서관통계시스템, KOSIS, 복지포 등 공공데이터 수집 · Python·Excel 기반 데이터 정제 및 지역 단위 통합 · R 기반 다중회귀분석 및 통계적 유의성 검정 · 회귀계수, p-value, Adjusted R² 기반 결과 해석 · 지역별 주요 지표 지도 시각화
	주요결과	✓ 도서관 연면적·프로그램 강좌 수·학교 수·실업률·노후 건물 면적을 유의변수로 도출하고, 재구성 모형의 설명력을 Adjusted R ² 0.74 에서 0.79 로 개선
	담당업무	· 공공데이터 수집 및 전처리 · 회귀분석 및 시각화 · 결과 해석 및 보고서 작성
24.08. ~ 24.10. (약 2 개월)	프로젝트명	▶ 화성시 노인 교통사고 위험 지역 분석 및 예방 정책 제언 4 명 팀프로젝트(화성시 x 수원대 공공빅데이터 전문가 양성과정)
	분석목적	- 노인 교통사고 영향 요인 및 고위험 지역 분석을 통한 예방 정책 우선순위 제안
	분석방법	· Python·R 기반 데이터 정제 및 탐색적 분석 · 후진제거법을 활용한 유의 변수 선별 · Random Forest 기반 변수 중요도 분석 · XGBoost, Lasso, Ridge 모델 비교 및 5-Fold 교차검증 · 계층적 군집분석 및 Tableau 지도 시각화
	주요결과	✓약국·경로당·의료시설·도로면적 등을 주요 영향 변수로 도출하고, XGBoost 모델이 R ² 약 0.89 를 기록해 노인보호구역 확대·단속카메라 재배치가 필요한 고위험 지역 식별
	담당업무	· 데이터 수집 및 전처리 · 예측 모델링 및 군집분석 · 분석 결과 시각화 및 포스터 제작

24.09. ~ 24.12. (약 3 개월)	프로젝트명	▶ 지역별 생활자원순환지수 개발 및 군집 분석 5 명 팀프로젝트(수원대학교 학부 프로젝트)
	분석목적	- 지역별 재활용 효율과 폐기물 처리 특성을 종합한 생활자원순환지수 개발 및 군집별 개선 방향 도출
	분석방법	· 2019~2022 년 시·군·구 단위 공공데이터 통합 · EDA 및 Pearson 상관계수 기반 주요 변수 선정 · 가중합 및 정규화를 통한 생활자원순환지수 개발 · Elbow Method 기반 최적 군집 수 결정 · 계층적 군집분석 및 Tableau 지도 시각화
	주요결과	✓ 2019~2022 년 전국 시·군·구 데이터를 기반으로 생활자원순환지수(0~1)를 개발하고 4 개 유형으로 분류한 결과, Cluster 2 는 가장 높은 자원순환 효율, Cluster 3 은 가장 낮은 수준을 보여 군집별 맞춤형 개선 방향을 도출
	담당업무	· 공공데이터 수집 및 전처리 · 지수 개발 및 군집 분석 수행 · 분석 결과 해석 및 시각화 · 정책 제안 및 결과 정리
24.09. ~ 24.12. (약 3 개월)	프로젝트명	▶ AI 해외여행 플래너 서비스 기획 및 자동화 구현 5 명 팀프로젝트(수원대학교 학부 프로젝트)
	개발목적	- 사용자 여행 조건 기반 맞춤형 해외여행 일정 생성·저장·전달 자동화 서비스 구현
	개발방법	· Google Form 을 활용한 사용자 여행 설문 입력 수집 · Google Sheets 를 통한 설문 응답 데이터 저장 및 관리 · ChatGPT API(OpenAI)를 활용한 사용자 맞춤형 여행 일정 자동 생성 · Make 노코드 자동화 플랫폼을 활용한 전체 워크플로우 구성 · Discord 및 Gmail 을 통한 여행 계획 결과 자동 전달 · Notion Database 에 여행 계획 데이터 저장 및 관리
	개발효과	✓ 사용자 설문 응답을 통해 맞춤형 여행 일정 생성, Discord·Gmail 전달, 리마인더 발송 및 Notion DB 저장까지 전 과정을 자동화한 서비스 구현
	담당업무	· 서비스 기획 및 자동화 시나리오 설계 · Make 모듈 연동 및 워크플로우 구현 · OpenAI API 프롬프트 설계 · 결과 전달 및 데이터 저장 구조 설계 · 프로젝트 결과 정리 및 PPT 제작

25.03. ~ 25.06. (약 3 개월)	프로젝트명	▶ 유방암 환자 예후 예측을 위한 병리 이미지·전사체 기반 멀티모달 딥러닝 모델 3 명 팀프로젝트(수원대학교 캡스톤디자인)
	분석목적	- 병리 이미지와 전사체 데이터를 결합한 유방암 환자 생존 위험 예측 및 단일 모달 모델의 성능 한계 개선
	분석방법	· TCGA-BRCA 병리 이미지·전사체·생존 데이터 활용 · WSI 패치 분할 및 GAT 기반 이미지 표현 학습 · VAE Encoder 기반 전사체 잠재 표현 생성 · Cross-Attention 기반 모달리티 정보 융합 · Contrastive Loss 및 CoxPH Loss 기반 생존 분석
	주요결과	✓ 병리 이미지 단일 모델 C-index 0.75, 전사체 단일 모델 0.70 대비 멀티모달 VA-GATSurv 모델이 0.80 을 달성해 생존 예측 성능 개선
	담당업무	· 전사체 표현 학습을 위한 VAE 기반 선행 논문 스터디 및 모델 설계 방향 정리 · VAE Encoder 구조 구현 및 전사체 잠재 표현 생성 · 생성된 전사체 표현을 멀티모달 모델 입력으로 연계 · 연구 결과 정리 및 포스터 제작
25.10.13 ~ 25.10.22 (10 일)	프로젝트명	▶ 무신사 상품 크롤러 개인 프로젝트(아이티월)
	개발목적	- 무신사 카테고리별 상품 정보·이미지 자동 수집 및 Excel 보고서 생성 크롤러 개발
	개발방법	· Selenium 기반 로그인 및 브라우저 자동화 · BeautifulSoup 기반 상품 정보 파싱 · urllib 을 활용한 상품 이미지 다운로드 · pandas 기반 CSV·XLSX 파일 생성 · openpyxl·Pillow 기반 이미지 삽입 및 엑셀 서식 자동화
	개발효과	✓ 사용자가 원하는 상품 카테고리를 선택해 해당 상품 정보와 이미지를 자동 수집하고, Excel 보고서로 확인할 수 있는 기능 구현
	담당업무	· Selenium 기반 로그인 및 카테고리 선택 자동화 구현 · 상품 데이터 및 이미지 크롤링 로직 설계 및 구현 · CSV 및 XLSX 파일 저장 로직 구현 · Excel 이미지 삽입 및 서식 자동 조정 기능 개발
25.10.13 ~ 25.10.22 (10 일)	프로젝트명	▶ FotMob 데이터 크롤러 개인 프로젝트(아이티월)
	개발목적	- 리그·시즌·카테고리별 경기·순위·선수·팀·뉴스 데이터 자동 수집 및 Excel 저장 크롤러 개발

	개발방법	· Selenium 기반 브라우저 자동 제어 · WebDriverWait 기반 DOM 상태 감지 · 사용자 입력 기반 리그·시즌·카테고리 선택 · 경기, 순위, 선수·팀 통계 및 뉴스 데이터 파싱 · 섹션별 Excel 파일 자동 생성
	개발효과	✓ 사용자가 원하는 리그·시즌·카테고리를 선택해 경기·순위·선수·팀·뉴스 데이터를 자동 수집하고, 항목별 Excel 파일로 확인할 수 있는 기능 구현
	담당업무	· FotMob 크롤러 전체 구조 설계 및 구현 · 리그·시즌·카테고리·섹션 선택 인터페이스 로직 구현 · DOM 기반 섹션 감지 및 데이터 수집 함수 개발 · 경기, 순위, 선수·팀 통계, 뉴스 데이터 파싱 및 저장 · 크롤링 결과 엑셀 파일 자동 생성 및 관리

25.11.17 ~ 25.12.19 (약 1 개월)	프로젝트명	▶ 머신러닝을 활용한 경기도 싱크홀 위험도 예측 및 위험지도 구축 6 명 팀프로젝트(아이티윌)
	분석목적	- 경기도 250m 격자 단위 싱크홀 위험도 예측 및 예방 점검을 위한 공간 위험지도 구축
	분석방법	· 싱크홀 발생 이력과 인구·지반·지질·교통·건축물·하수도 관련 공공데이터 수집 · QGIS 기반 250m×250m 격자 생성 및 좌표계(EPSPG:5181) 통일 · Spatial Join 을 활용한 격자 단위 통합 데이터셋 구축 · 47 개 독립변수를 인구, 지반·지질, 교통·지하구조물, 건물, 하수도 인프라 등 5 개 범주로 구성 · 상관계수와 VIF 를 활용한 다중공선성 검토 · SHAP Feature Importance 와 PR-AUC 변화 기반 단계적 변수 선택 · Balanced Random Forest 와 Logistic Regression 모델 비교 · 클래스 불균형을 고려한 Recall·PR-AUC 중심 성능 평가 · Streamlit 기반 반응형 싱크홀 위험도 대시보드 및 로드뷰 기능 구현
	주요결과	✓ 경기도 전역을 약 13 만 개의 250m×250m 격자로 구성하고, 47 개 변수를 활용해 격자별 싱크홀 위험도를 예측하고, 최종 모델 Recall 0.78 을 확보해 예방 점검 우선지역을 확인할 수 있는 Streamlit 위험지도 구현
	담당업무	· 공공기관 싱크홀 관련 데이터 수집 및 전처리 · QGIS 기반 격자 생성 및 공간 조인 수행 · 머신러닝 모델 비교·평가 및 결과 해석 · ppt 제작 및 발표
26.01. ~ 26.02. (약 1 개월)	프로젝트명	▶ Willpang Growth Analytics: AARRR 기반 이커머스 성장 전략 설계 및 시뮬레이션 5 명 팀프로젝트(아이티윌)
	분석목적	- AARRR 기반 고객 행동·구매 퍼널·리텐션 분석 및 개선 시나리오별 매출·마진 변화 시뮬레이션

	분석방법	· Mockaroo 기반 고객·세션·이벤트·주문·주문상품 이커머스 데이터셋 설계 · 원가·마진 컬럼 설계 및 Gross Margin 계산 구조 구현 · Python(Pandas, NumPy) 기반 AARRR 단계별 KPI 정의 및 구매 퍼널 구성 · Activation 전환율과 병목 구간, Retention 및 고객 매출 집중도 분석 · TF-IDF 기반 리뷰 감성 분석을 통한 정성 데이터 정량화 · Traffic·Funnel·Retention 개선 시나리오별 매출 및 마진 시뮬레이션 · Power BI 기반 AARRR 5 단계 대시보드 구축 · RAG 기반 사내 데이터 분석 챗봇 'Willbot' 구현
	주요결과	✓ 1,966 개 세션에서 구매 473 건, 최종 전환율 24.06%를 기록했으며 Add to Cart → Checkout 구간을 핵심 병목으로 식별 ✓ 재구매 고객 74 명, 초기 유지율 18.83%로 구매 이후 리텐션 구조가 취약함을 확인 ✓ 총매출 \$25,702 중 상위 10% 고객이 43.78%를 차지하는 매출 편중 구조 도출 ✓ 단순 트래픽 증가보다 퍼널 전환율 개선과 리텐션 강화가 더 큰 마진 상승효과를 만드는 성장 전략임을 시뮬레이션으로 검증
	담당업무	· Mockaroo 를 활용한 이벤트 기반 이커머스 데이터 직접 설계 및 생성 · AARRR 분석 구조 설계 및 단계별 KPI 정의 · Power BI 대시보드 설계 및 시각화 · ppt 제작 및 발표