

킵어스 포트폴리오 프로젝트 만들기

목차

1. 왜 포트폴리오가 중요한가 — 채용 관문의 현실
2. 상위 철학 — 리크루터 중심 최적화
3. Tip 1 — 스토리가 되는 프로젝트 주제 선택
4. Tip 2 — 시각적으로 돋보이는 프로젝트
5. Tip 3 — 열정 기반 프로젝트의 위력
6. Tip 4 — 흥미로운 데이터셋 확보 전략
7. Tip 5 — 완료 > 완벽, 완성을 증명하라
8. Tip 6 — 비즈니스 가치를 보여라
9. 하나의 킬러 프로젝트로 행동면접 돌파
10. 핵심 요약 및 실행 체크리스트

PART 1

포트폴리오 프로젝트, 왜 중요한가

포트폴리오 없는 지원서는 적신호

- 채용 매니저들이 공통적으로 지적한 최대 위험 신호: 포트폴리오 프로젝트 부재
- 특히 신입/전직자에게 치명적 — 증명할 것이 더 많기 때문임
- 포트폴리오의 3대 효과:
 - 실무 경험을 대체할 수 있는 가장 강력한 수단
 - 취업 시장에서 차별화의 핵심 도구
 - 면접에서 이야기할 거리를 만들어 줌
- 교실 지식을 현실 문제에 적용하는 실전 연습장 역할

PART 2

상위 철학 — 누구를 위해 만드는가

리크루터의 시선으로 설계하라

- 리크루터 = 면접 게이트키퍼 — 첫 관문을 여는 사람임
- 리크루터는 Jupyter Notebook 94번째 줄의 모델 선택을 보지 **않음**
- 실제 행동 패턴:
 - 이력서나 콜드메일의 프로젝트 설명을 10초 읽음
 - 운이 좋으면 링크를 클릭해 **그래픽/데모**를 봄
 - 30초 이내에 "흥미롭고 직무에 관련 있는가"를 판단함
- 핵심 원칙: "**트윗 가능한**" 프로젝트 — 140자로 본질 전달 가능해야 함

리크루터 최적화 vs 기술적 깊이

리크루터 최적화

- 이메일로 쉽게 설명 가능한 프로젝트
- 한 줄 요약이 즉시 와닿는 아이디어
- 시각적 결과물(GIF, 차트, 데모)
- 명확한 비즈니스 가치 또는 임팩트

기술적 깊이도 유지 가능

- 깔끔하게 포장된 프로젝트는 기술 면접관도 좋아함
- "이해하기 쉬운 프로젝트"가 "기술적으로 약한 프로젝트"는 아님
- 좋은 패키징 + 강한 기술력 = 양립 가능
- 첫날부터 리크루터를 염두에 두고 시작하면 결과물이 자연스럽게 향상됨

PART 3

Tip 1 — 스토리텔링이 되는 프로젝트

인간은 이야기에 반응하도록 설계됨

- 유발 하라리 *사피엔스*의 핵심 논지: 공유된 이야기가 인류 협력의 기반
- 리크루터와 면접관도 인간 — 스토리에 반응하게 진화함
- 스토리 구조를 갖춘 프로젝트의 이점:
 - 면접에서 유창하게 설명 가능
 - 면접관이 더 흥미를 느끼고 기억에 남음
 - 자연스럽게 논리적 흐름이 생김

프로젝트에 스토리 구조 적용하기

스토리 요소	프로젝트 적용
도입 (Build-up)	탐색적 데이터 분석(EDA)으로 배경과 동기 설정
갈등 (Conflict)	가설(Hypothesis) 제시 — 해결해야 할 질문
해결 (Resolution)	가설 검증 결과 공유 — 인사이트와 결론

- 이 3단계 구조를 따르면 프레젠테이션과 콜드메일 모두 강력해짐
- "무엇을 했는가"보다 "왜 이 질문이 중요한가" 에 집중할 것

스토리가 되는 프로젝트 아이디어 소스

뉴스 트렌드 활용

- 화제의 뉴스를 데이터 분석으로 연결
- 예: 코로나 — 우편번호별 백신접종률과 인구통계 상관관계 분석
- 예: 선거 — 어떤 인구통계 요인이 선거구 전환에 가장 큰 영향을 미치는가

데이터 저널리즘 참고

- NYT, FiveThirtyEight 등의 시각화를 출발점으로 활용

- 기존 분석을 더 세밀한 수준으로 확장

개인적 문제 해결

- 자신이 직접 겪는 불편을 해결하는 프로젝트
- 자신이 영웅(problem-solver)이 되는 내러티브

해커톤 사례

- 정신건강 문제를 다룬 해커톤 프로젝트 — 개인적 경험이 결합되어 강력한 서사를 형성함

PART 4

Tip 2 — 시각적으로 돋보이는 프로젝트

시각화가 주는 관심과 참여의 차이

- 텍스트만 있는 블로그 포스트 vs GIF/이미지 포함 포스트 → 참여율이 압도적으로 차이남
- SafeGraph 사례:
 - 새로운 지리공간 데이터셋 출시 시 GIF 포함 트윗이 훨씬 높은 참여율 기록
 - 콜드 이메일에서도 GIF 삽입 시 높은 반응률
- 추천 결과물 형태:
 - 인터랙티브 도구(웹앱, 대시보드)
 - GIF 또는 애니메이션
 - 한 장으로 결과를 요약하는 인포그래픽

시각화 능력은 실무 역량의 증거

- "왜 기술적 스킬 대신 시각화에 시간을 쓰나?"라는 의문에 대한 답:
 - 결과를 간결하고 정확하게 전달하는 능력 자체가 핵심 역량
 - 실무 데이터 사이언스 = 이해관계자를 설득하여 의사결정 유도
 - 복잡한 모델도 경영진에게 설명할 수 없으면 의미 없음
- 포트폴리오에서 시각화 역량을 보여주면:
 - 면접관에게 직무 적합성 신호를 보냄
 - "이 사람은 현업에서도 잘 소통하겠구나"라는 확신을 줌

PART 5

Tip 3 — 열정이 담긴 프로젝트의 힘

열정 프로젝트가 면접에서 만드는 효과

소통과 분위기

- 열정은 전염됨 — 즐겁게 말하면 면접관도 즐거워함
- 자연스러운 의사소통이 가능 → 긴장 속에서도 유창한 커뮤니케이터로 보임
- 열정에서 오는 **헤일로 효과**: 데이터 사이언스, 직무, 회사에 대한 열정으로 확장

주니어 지원자의 무기

- 초급 데이터 사이언티스트는 기업의 투자가 필요한 포지션
- 채용 매니저가 주니어에 투자하는 기준: **열정과 의욕**
- 열정 = 빠르게 배우고 성장할 것이라는 신호
- 기술력은 부족해도 관심과 의지가 더 강한 후보가 선택됨

열정을 프로젝트에 녹이는 실전 예시

관심 분야	프로젝트 아이디어
농구	NBA 데이터셋을 활용한 선수 성과 분석
음악	가사 기반 장르 분류 모델
넷플릭스	IMDB 데이터셋 기반 나만의 영화 추천 알고리즘
힙합	RapStock.io — 떠오르는 래퍼에 베팅하는 플랫폼

- Nick Singh의 RapStock.io 사례:
 - 힙합 음악에 대한 진정한 열정 → 프로젝트 설명 시 자연스럽게 열정 전달
 - 리크루터에게 데이터 사이언스와 알고리즘에 대한 열정으로 인식됨
- 추가 이점: 열정 프로젝트는 완성까지 가는 동력이 됨 (Tip 5와 연결)

PART 6

Tip 4 — 흥미로운 데이터셋 확보 전략

피해야 할 데이터셋과 그 이유

- 절대 사용 금지 목록:
 - Iris Plant 데이터셋
 - Titanic 생존자 분류 (Kaggle)
 - MNIST 손글씨 인식
- 문제점:
 - 수업 과제와 동일 → 리크루터가 이미 여러 번 본 프로젝트
 - 열정이 담기지 않은 프로젝트로 보임
 - 대학/부트캠프 채용 행사에서 같은 프로젝트를 반복해서 보는 리크루터의 피로감
- 주의가 필요한 데이터:
 - 주식 시세 데이터 (월스트리트 지원 아닌 한)
 - 트위터 피드 데이터 (정말 참신한 관점이 아닌 한)

흥미로운 데이터셋을 찾는 단계별 전략

• Step 1: 열정 분야 특정

- 나의 관심사와 목표 기업을 연결
- 예: 우주에 관심 + NASA 취업 희망

• Step 2: Kaggle에서 시작

- NASA의 소행성 분류 챌린지, 케플러 외행성 탐색 데이터셋
- 장점: 깨끗하고 포맷이 잘 됨, 스타터 노트북 존재
- 단점: 다른 지원자도 같은 데이터를 사용할 수 있음

• Step 3: 직접 스크래핑

- BeautifulSoup, Scrapy(Python) 또는 Rvest(R) 활용
- 장점: 프로젝트의 유일성 확보 + 프로그래밍 역량 입증
- 데이터 수집/정제 경험 = DS 전체 생애주기 이해 증명

• Step 4: 보완 데이터 결합

PART 7

Tip 5 — 완료가 완벽보다 낫다

"대략 맞으면 된다"의 진짜 의미

- 기술적 세부사항(아웃라이어 처리 등)은 면접 획득에 큰 영향 없음
- 리크루터가 빠르게 판단할 수 있는 것: 프로젝트의 완성도
- 핵심 원칙: Done > Perfect
- 프로젝트를 "마무리(wrapping up)"하는 데 추가 노력을 투자할 것
- 가능하면 프로젝트를 제품화(productionize) 하라

프로젝트 제품화의 구체적 방법

제품형 프로젝트

- 얼굴 사진으로 나이 예측하는 분류기 → 웹앱으로 전환하여 누구나 사진 업로드 가능하게
- 2단계: 신경망으로 다른 나이의 얼굴 변환 (FaceApp 유사)
- Snapchat, Instagram, TikTok에 콜드메일 가능

탐색형 프로젝트

- 동네 소득 vs 학군 품질 분석 → 인터랙티브 지도로 시각화
- D3.js(차트), Leaflet.js(지도), rCharts(R)

- 완성된 인터랙티브 지도를 Zillow, Opendoor에 콜드 메일

GitHub 공개 필수

- 코드를 GitHub에 공개 → 자부심의 신호
- Google Colab으로 노트북 공유 가능
- 아무도 안 봐도 링크 자체가 신뢰도 향상

핵심 메시지

- 포트폴리오 프로젝트는 공개되기 전까지 미완성
- 코드 공개 = "내가 진짜 이걸 했다"는 증거
- 이력서 조작 의혹 차단

PART 8

Tip 6 — 비즈니스 가치를 입증하라

비즈니스 언어로 말하기

- 최고의 포트폴리오는 비즈니스 가치를 구체적으로 입증함
- 특히 PhD 출신의 산업 전환 시 필수적 — 작은 기업/스타트업 지원 시 더욱 중요
- 비즈니스 지원 시 비즈니스 용어로 소통할 것
- 프로젝트에 반드시 포함할 것:
 - 명확한 비즈니스 추천안(recommendation)
 - 또는 매력적인 핵심 인사이트(takeaway)

비즈니스 가치를 수치로 보여주는 법

직접 지표가 있는 경우

- 절약된 비용(달러)
- 절약된 시간
- 생성된 매출
- 정확도 향상 수치

직접 지표가 없는 경우

- **사용량 지표**를 대리 변수로 활용:
 - 조회수(view counts)
 - 다운로드 수
 - 활성 사용자 수
- 이 지표들은 **완성도와 실제 가치 창출을 동시에 증명함**
- 예외: 데이터/스토리가 충분히 흥미로우면 비즈니스 가치 생략 가능

PART 9

하나의 킬러 프로젝트 전략

여러 개보다 하나에 올인하라

- 6가지 팁을 모두 적용하는 것은 엄청난 노력이 필요함
- 좋은 소식: 단 하나의 킬러 프로젝트에 에너지를 집중해도 충분함
- 하나의 프로젝트에 필요한 것:
 - 아름다운 시각화
 - 작동하는 데모
 - 사용량 통계 또는 비즈니스 가치 지표
- 여러 사이드 프로젝트를 홍보하는 것은 실패의 지름길 — 하나도 견인력 얻기 어려움

킬러 프로젝트로 행동면접 돌파하기

- 행동면접의 전형적 질문: "~했던 경험을 말해주세요"
- **같은 프로젝트를 반복 활용하면:**
 - 매번 배경 설명에 시간을 낭비하지 않음
 - 곧바로 **핵심 답변**으로 진입 가능
 - 더 깊은 수준의 기술적 도전을 다룰 수 있음
- "가장 어려웠던 데이터 문제는?" → 깊이 파고든 하나의 프로젝트가 최적의 답변

"왜 이 회사인가?"에 대한 완벽한 답변

- 열정 기반 프로젝트 = "말이 아니라 행동으로 보여주는" 관심 증명
- Nick Singh의 RapStock.io 활용 사례:

지원 기업 유형	강조한 프로젝트 측면
핀테크	주식시장/상품 가격 책정 측면
데이터 기업	Spotify 데이터 기반 래퍼 가격 알고리즘
마케팅/애드테크	Google Ads 캠페인 실패 후 광고에 대한 관심
스타트업	스타트업 창업자 경험, 빠른 실행과 배포
Facebook Growth	소비자 제품 성장, 엔지니어링+데이터+실험 결합

- 하나의 프로젝트로 다양한 기업에 맞춤형 답변 가능

PART 10

핵심 요약 및 실행 가이드

6가지 팁 한눈에 보기

Tip 1: 스토리

- 도입-갈등-해결 구조
- 뉴스, 개인 문제에서 착안

Tip 2: 시각화

- GIF, 인터랙티브 도구
- 시각 결과물 필수

Tip 3: 열정

- 관심사 기반 주제
- 헤일로 효과 활용

Tip 4: 데이터셋

- 식상한 데이터 회피
- 직접 스크래핑 추천

6가지 팁 한눈에 보기 (계속)

Tip 5: 완성도

- 제품화/인터랙티브
- GitHub 공개 필수

Tip 6: 비즈니스

- 구체적 가치 입증
- 사용량 지표 활용

올인 전략

- 하나의 킬러 프로젝트
- 행동면접 전체 커버

최종 목표

- 30초 안에 리크루터를 사로잡는 프로젝트

체크리스트

- ① 프로젝트 아이디어를 한 문장으로 설명할 수 있는가? (트윗 테스트)
- ② 스토리 구조(도입-갈등-해결)가 있는가?
- ③ 시각적 결과물(GIF, 차트, 데모)이 있는가?
- ④ 학생본인의 관심사가 반영되어 있는가?
- ⑤ Iris/Titanic/MNIST 같은 익숙한 데이터셋을 피했는가?
- ⑥ 프로젝트가 완성되어 있고, 코드가 GitHub에 공개되어 있는가?
- ⑦ 비즈니스 가치 또는 사용량 지표가 포함되어 있는가?