

# NLP Project DAY 2025



(주) 고누아이

윤수혁

류양환

문이환

신동찬



# NLP Project DAY 2025



## 팀원 소개

- 윤수혁 - 프로젝트 총괄, 모델 학습 및 시스템 통합
- 류양환 - 프론트엔드 개발, UI/UX 디자인
- 문이환 - 백엔드 개발, 데이터베이스 설계
- 신동찬 - 데이터 수집 및 전처리, 발표 자료 제작



# NLP Project DAY 2025



(주) 고누아이

1. 문제정의
2. 시장조사
3. 솔루션
4. 비즈니스모델



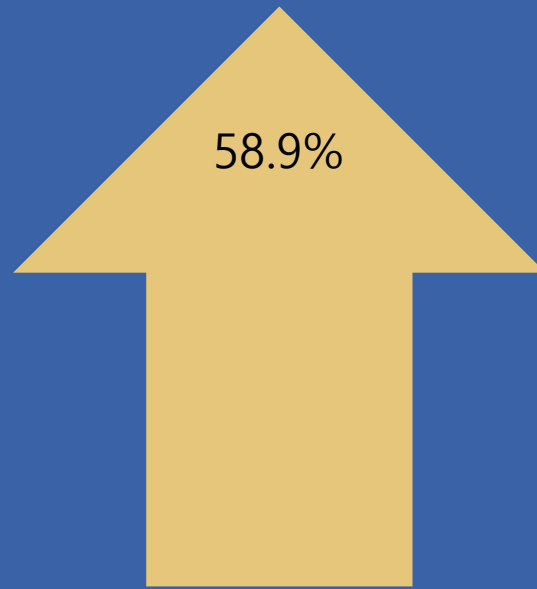
# NLP Project DAY 2025



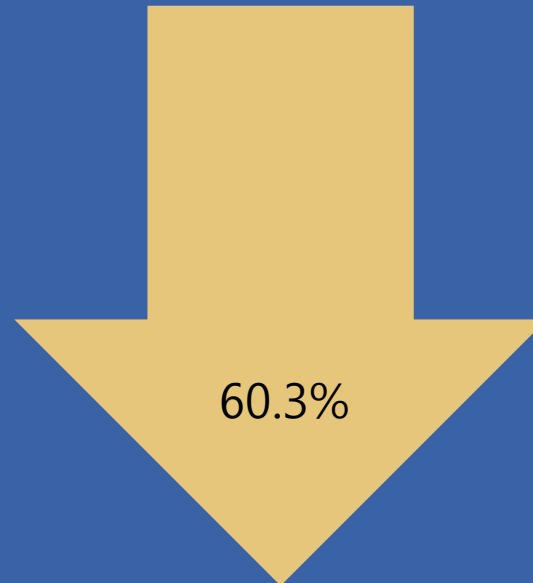
자사에 입사를 희망하게 된 구체적  
인 이유는 무엇이며, 입사 후 목표를  
기술해 주세요. (500자 이내)



# NLP Project DAY 2025



기준에 미달하지만  
잘 쓴 자기소개서



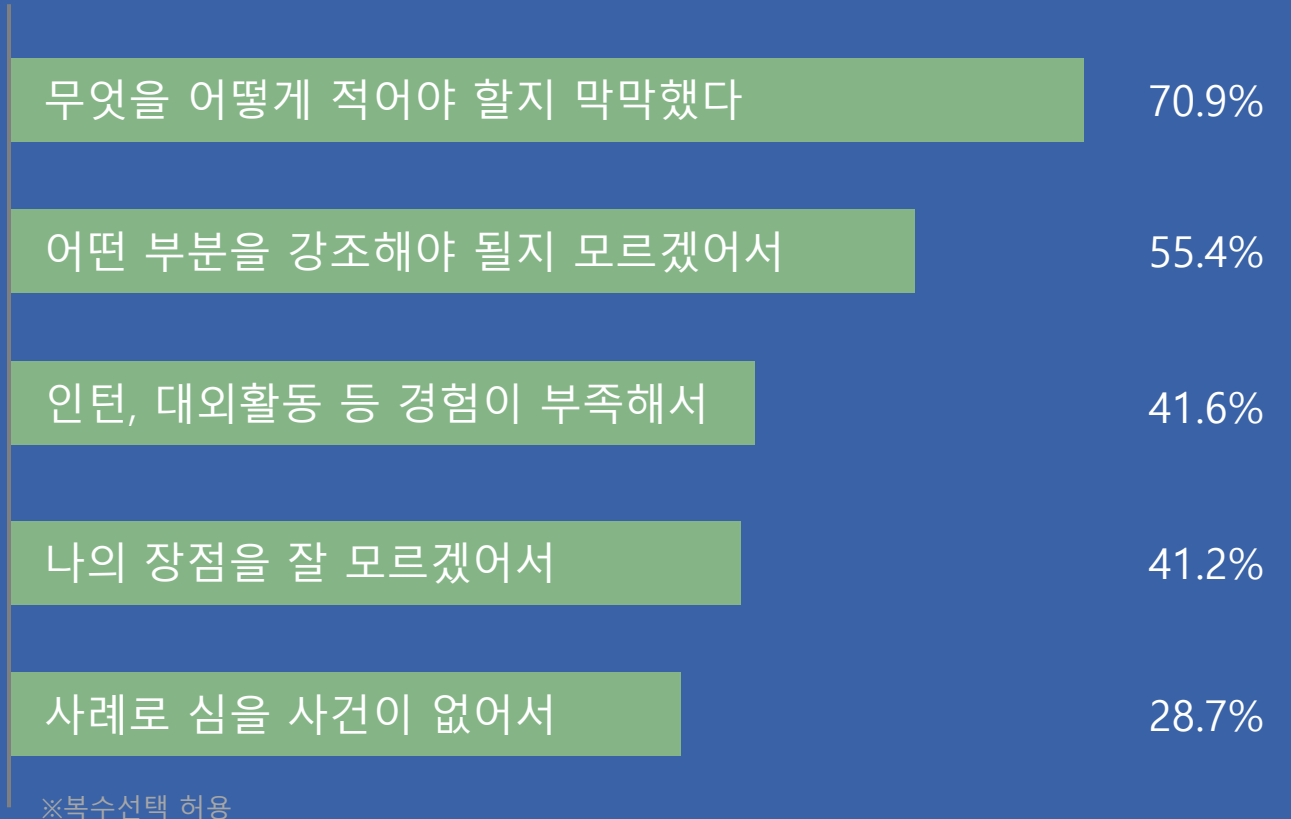
기준에 충족하지만  
못 쓴 자기소개서



# NLP Project DAY 2025



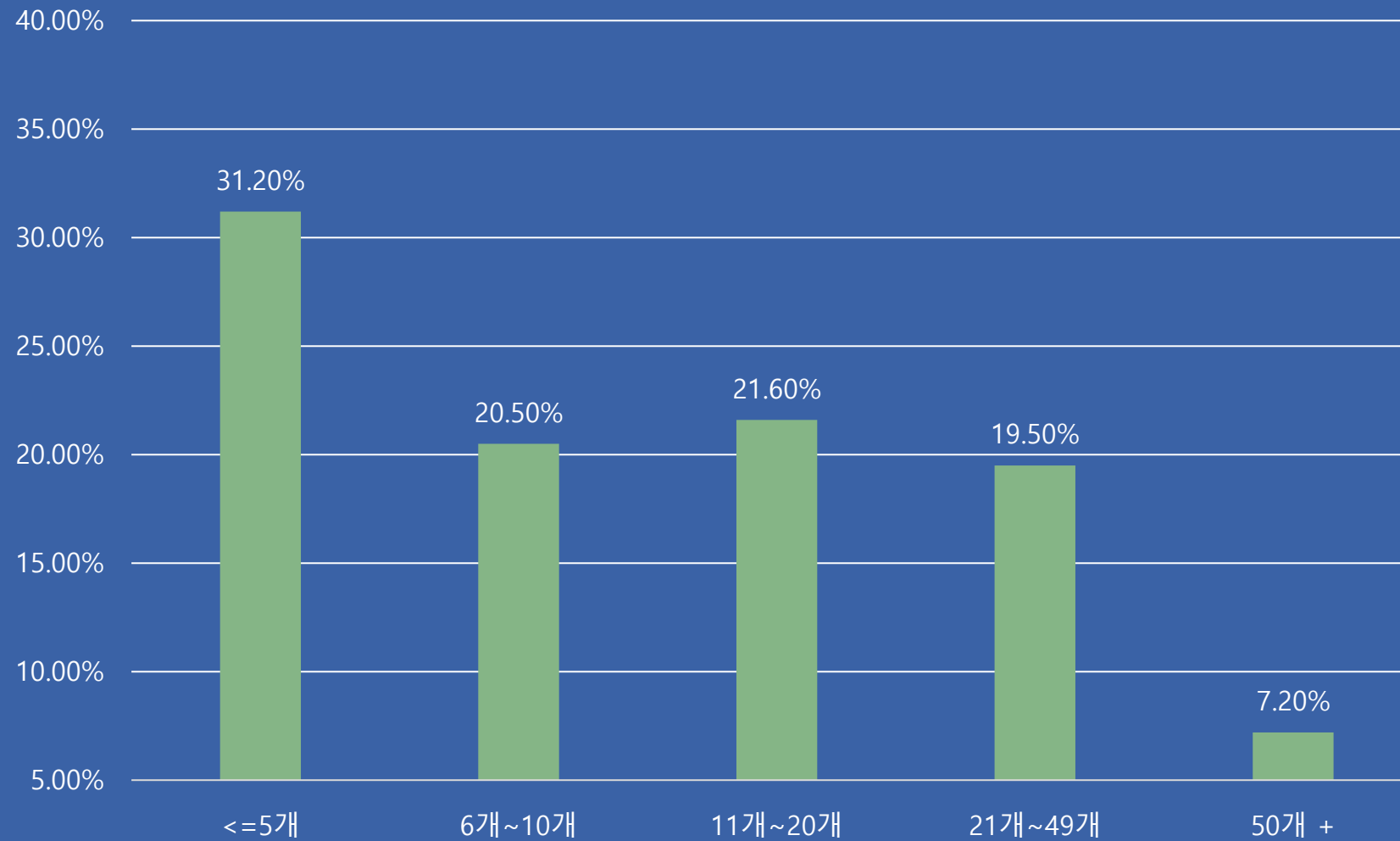
취준생 97%  
'자소서 준비, 막막한 적 있다'



# NLP Project DAY 2025



구직자별 평균 자소서 개수



# NLP Project DAY 2025



## 타 서비스 현황





- 경영·사무·금융
- 영업·고객상담
- 무역·유통
- 마케팅·광고·홍보
- 서비스
- IT·전기·전자·통신
- 생산·제조·정비
- 연구개발·설계
- 토목·건설
- 디자인
- 의료
- 미디어·예술·스포츠
- 교육
- 전문·특수직
- 공공·사회봉사

업종

- 서비스업
- 금융업
- 유통업
- IT·전기·전자
- 제조업
- 건설업
- 교육업
- 의료·제약업
- 미디어·광고·홍보
- 문화·예술·디자인업
- 공공기관·협회

기업형태

- 대기업
- 중견기업
- 중소기업
- 공기업
- 외국계
- 금융권
- 스타트업

지원분야

it

담당업무

app developer

기업명

왓차

경력

- 경력
- 신입

핵심키워드

- 신입 ✕
- 열정 ✕
- AI ✕
- Android ✕

추가

키워드를 입력해 주세요.

\* 핵심키워드는 여러분을 기업에 추천할 때 분류되는 키워드입니다.

작성방법(영상)

작성방법(이미지)

탁월한 사무능력에 배려의 마음을 더해 당사를 영입하겠습니다!

성장과정 ▾ 성장과정

이용하실 AI를 골라주세요.

- ☐ 직접쓰기
- ☒ ChatGPT
- ☐ GPT Pro
- ☐ PenGPT

AI에게 키워드나 문장을 입력해주세요 (최대 500글자)

생성하기

자기소개서 내용을 입력해주세요.

문장 번역하기 ▾

번역 하기

번역 되돌리기

글자수 : 0자(공백포함)

글자수 조정

자기 개발 계획

자기 개발 계획

이용하실 AI를 골라주세요.



직접쓰기



ChatGPT



GPT Pro



PenGPT

열정, mvvm모델, 빠른학습, 능숙한AI활용, ios까지공부할계획

AI에게 키워드나 문장을 입력해주세요 (최대 500글자)

생성하기

자기소개서 내용을 입력해주세요.



문장 번역하기



번역 하기

번역 되돌리기

글자수 : 0자(공백포함)

글자수 조정

## ChatGPT 추천문장

열정을 담아 미래를 설계하다

안녕하세요, 성장과 도전의 기회를 찾아 끊임없이 노력하는 예비 앱 개발자입니다. IT와 기술의 무한한 가능성에 매료되어 앱 개발자의 길을 선택하게 되었으며, 왓차라는 플랫폼에서 그 꿈을 펼치고자 지원하게 되었습니다.

우선, 저는 빠른 학습 능력을 기반으로 새로운 기술과 트렌드를 습득하는 데 자신이 있습니다. 특히, MVVM(Model-View-ViewModel) 아키텍처에 대한 깊은 이해를 바탕으로 효율적이고 유지보수 가능한 코드를 작성할 수 있는 능력을 갖추고 있습니다. MVVM 모델은 앱의 구조를 체계화하고, 클린 아키텍처를 구현하는 데 필수적인 요소로, 이를 통해 더욱 견고한 소프트웨어를 개발할 수 있다고 믿습니다.

또한, AI 기술의 활용에 능숙하여 데이터 기반의 의사결정과 추천 시스템 등에 이르는 다양한 IT 분야의 기술을 적재적소에 적용할 수 있습니다. 이러한 경험은 왓차가 사용자 경험을 최적화하고 더욱 개인화된 콘텐츠를 제공하는 데 기여할 수 있는 큰 강점이 될 것입니다.

더불어, 현재 iOS 개발에 대한 공부를 이어가고 있으며, 다양한 플랫폼에서 앱을 제공할 수 있는 역량을 갖추기 위해 노력하고 있습니다. 이러한 계획은 다방면의 사용자 경험을 제공함으로써 왓차의 글로벌 확장 및 시장 경쟁력 강화에 기여할 수 있는 밑거름이 될 것입니다.

저는 왓차라는 혁신적인 환경에서 제 열정을 발휘하여 유연하고 다재다능한 개발자로 성장할 준비가 되어 있습니다. 팀과 함께 새로운 기술을 탐구하고 도전 과제를 해결해 나가며, 왓차의 비전에 기여하고 싶습니다. 감사합니다.

이 문장 사용

# NLP Project DAY 2025



페르소나



# NLP Project DAY 2025



신동찬 (27, 남)  
구직자

## Needs



뻔한 자소서는 필요 없어 내가 지원하는 기업에 맞춤형 자소서가 필요해!



나는 글을 쓰는 게 겁이 나, 초안만이라도 빠르게 작성돼서 내가 수정할래!

## Pain Point



자기 경험을 정리하는 것 자체가 어려워서 어떻게 시작해야 할지 모르겠다.



한두 개 기업이 아니라 여러 개 기업에 지원해야 하다 보니 시간이 부족하다.

# NLP Project DAY 2025



## 솔루션 제안



# NLP Project DAY 2025



1

채용공고 크롤링

인크루트, 사람인, 잡코리아 등 플랫폼 제한 없는 크롤링

2

타겟 기업 크롤링

기업 공식사이트 크롤링을 통해 인재상, 사업방향을 파악함

3

사용자 편의성

타 서비스와 달리 가장 빠른 자소서 초안을 받아볼 수 있다.



자소서 생성을 위해 지원하려는 채용정보를 알려주세요!

공고 URL을 입력하세요.

지원하고자 하는 사이트 공식URL을 입력하세요.

나의 성격과 역경을 적어주세요

완성된 자기소개서가 여기에 표시됩니다.

생성하기

# NLP Project DAY 2025



사용한 Tool / 솔루션 / 서비스





# NLP Project DAY 2025



## 사용한 Tool / 솔루션 / 서비스

- 언어 모델: Llama-3-Motif-102B
- 프레임워크: Streamlit (프론트), FastAPI (백엔드)
- DB: Supabase
- 배포: Render, GitHub Actions
- 기타: HuggingFace Transformers, LangChain, V S Code



# NLP Project DAY 2025



## 모델링 소개



채용공고 크롤링은  
채용공고에서 전체 텍스트를 긁어옴

```
def save_to_file(text: str, filename: str = "output.txt"):
    """텍스트를 파일로 저장 (50자마다 줄바꿈 추가)"""
    try:
        # 텍스트를 50자마다 줄바꿈으로 포매팅
        formatted_text = format_text_by_line(text, line_length=50)

        with open(filename, "w", encoding="utf-8") as file:
            file.write(formatted_text)
        logging.info(f"결과가 파일에 저장되었습니다: {filename}")
    except Exception as e:
        logging.error(f"파일 저장 중 오류 발생: {str(e)}")

if __name__ == "__main__":
    url = "https://www.jobkorea.co.kr/Recruit/GI_Read/46478037?Oem_"

    try:
        result = fetch_and_clean_text(url)
        if result:
            # 전체 결과를 파일로 저장 (50자마다 줄바꿈 추가)
            save_to_file(result, "scraped_output.txt")

    except WebScrapingError:
        logging.error("스크래핑 작업 실패")
    except KeyboardInterrupt:
        logging.warning("사용자에 의해 작업 중단됨")
```

크롤링 로직은  
키워드 중심으로  
이런 키워드가 들어갈 만한

OR

이런 키워드가 들어가면 안될 것 같은  
URL을 선별해서 골라 들어가서  
내용을 긁어 가져온다

```
# 크롤링 제외할 URL 패턴
```

```
EXCLUDE_URLS = [
```

```
    "login", "signin", "signup", "register", "user", "account", "profile",  
    "password", "mypage", "session", "search", "apply", "cart", "recruit",  
    "faq", "help", "terms", "privacy", "support", "subsid", "policy",  
    "guide", "myform", "email", "phoneBook", "process", "return",  
    "로그인", "회원가입", "비밀번호", "계정", "아이디", "고객센터",  
    "문의", "이용약관", "개인정보처리방침", "약관", "법적고지", "자주 묻는 질문",  
    "도움말", "지원하기", "채용 공고", "채용 절차", "이메일 문의",  
    "채용 FAQ", "온라인 문의", "자주 하는 질문", "연락처"
```

```
]
```

```
# 크롤링할 주요 URL 패턴
```

```
TARGET_KEYWORDS = [
```

```
    "about", "vision", "mission", "values", "culture", "philosophy",  
    "our-story", "who-we-are", "strategy", "sustainability", "esg",  
    "corporate", "ethics", "principles", "history", "leadership",  
    "careers", "people", "insight", "story", "team", "life",  
    "company", "identity", "responsibility", "commitment",  
    "work", "growth", "innovation", "environment", "future",  
    "비전", "미션", "핵심가치", "철학", "가치", "목표", "전략",  
    "비전선언문", "기업소개", "윤리", "기업 윤리", "사회적 책임",  
    "지속 가능성", "환경", "윤리 강령", "리더십", "성장", "혁신",  
    "기업문화", "조직문화", "근무 환경", "일하는 방식", "기업가정신",  
    "팀워크", "경영이념", "인재상", "핵심 인재", "조직문화",  
    "채용 철학", "일하기 좋은 회사", "우리의 가치"
```

```
]
```

```
# 크롤링 후 제거할 단어 목록
```

```
EXCLUDE_WORDS = [
```

```
    "로그인", "아이디", "비밀번호", "회원가입", "검색", "채용공고",  
    "지원하기", "인재 등록", "공고", "상시지원", "채용 안내"
```

```
]
```

# NLP Project DAY 2025



## 서비스 구성도



# NLP Project DAY 2025



## 서비스 구성도

- [사용자 입력]
- ↓
- [프론트엔드 - Streamlit]
- ↓
- [백엔드 API - FastAPI]
- ↓
- [NLP 모델 (Llama-3)]
- ↓
- [DB 저장 및 응답 반환 - Supabase]



# NLP Project DAY 2025



## freemium

하루에 3번 제한 + 광고시청1회



크레딧 1원 = 200원

크롤링

1. 타겟 기업사이트 크롤링
2. 타겟기업 IR 크롤링
3. 타겟기업 뉴스 크롤링



+1  
+1  
+1



# NLP Project DAY 2025



향후 계획





# NLP Project DAY 2025



1

The VC 크롤링을 통해 해당기업 투자현황 분석

2

타겟기업 뉴스를 크롤링

3

GPT를 파인튜닝한다.



# NLP Project DAY 2025



## 자체 평가



# NLP Project DAY 2025



## 자체 평가

윤수혁

실제 모델 적용과 배포까지 경험할 수 있어 유익했다. 앱 개발과 다른 느낌의 웹 개발도 흥미로웠다. **9/10**

류양환

사이트 크롤링은 실패한 부분이 아쉬웠다. 보안의 이유로 막혀 있어 개발에 제한을 받아 아쉬웠다. **8/10**

문이환

2주의 시간 동안 개발을 완성했다. 차후에는 생성된 결과물 품질 평가 기준도 함께 설계하고 싶다. **8/10**

신동찬

사용자 중심 UI 설계가 부족했던 점은 아쉬웠다. 추후 시간이 추가된다면, 완벽하게 개발하고 싶다. **9/10**

# Thank you

## CONTACT US

**E-mail**                [sts05147@naver.com](mailto:sts05147@naver.com)

**Social Media**        [sts05147@naver.com](mailto:sts05147@naver.com)

**Phone**                +010-4584-6131

**Address**             2f, Gangnam-daero 94-gil 59, Gangnam-gu, Seoul, Republic of Korea