

# 자소서 적합성 검출

AI야, 합격 시켜조

# 목차

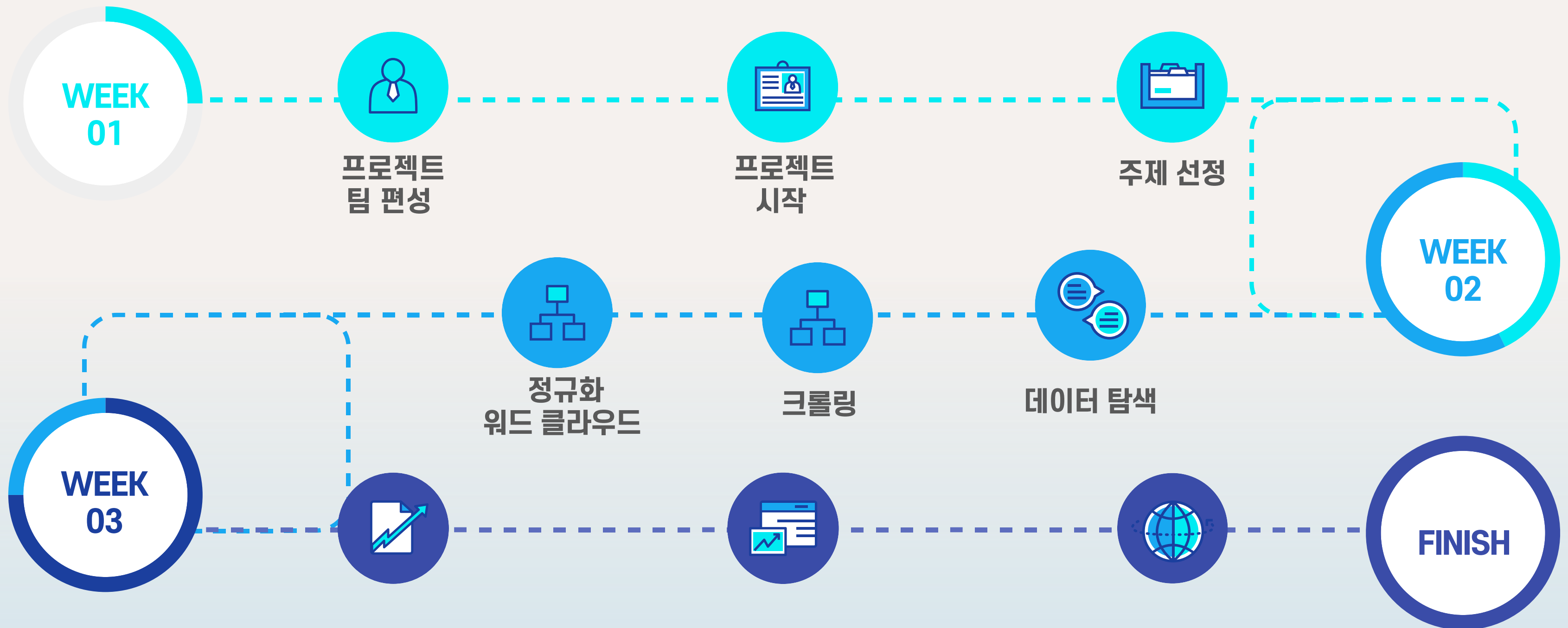
1. 조원소개
2. 프로젝트 일정
3. 주제선정 배경
4. 개발 공정

# 조원 소개

조장 : 서유진

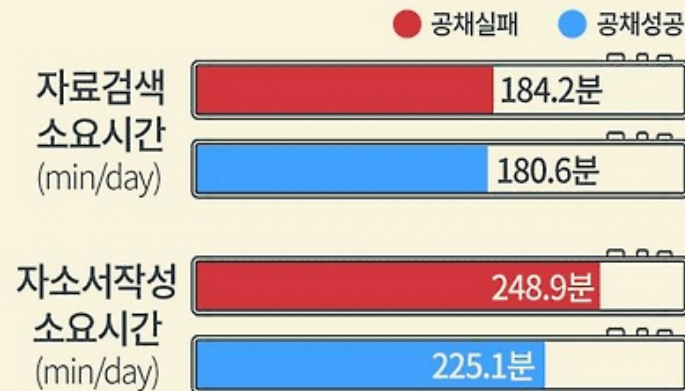
조원 : 김한솔, 이혁빈, 차다훈

# TIMELINE TITLE



## 공채시즌 취준생의 하루 봤더니... '자료검색3시간, 자소서작성4시간'

※ 2016 하반기 신입공채 취준생 681명 설문결과, 자료제공: 잡코리아



**JOBKOREA**

## 공채시즌 취준생의 하루

취준생 681명을 설문한 결과 취업 준비생들은 하루 평균 182.8분(3시간)을 자료검색에 쓰고, 242.4분(4시간)을 자기소개서 작성에 소비했다고 밝혔다.

자기소개서를 작성하는데 시간을 들인 사람 중 합격자(하루 평균 225.1분)보다 불합격자(248.9분)가 더 많았다

## 자소서 쓰면 안되는 단어

매우,엄청 -> 자극하, 대단히  
대개 -> 거의, 대부분  
그치만,하지만 -> 그러나,그럼에도  
그러니까 -> 따라서  
~라고 생각한다 -> ~로 생각된다. ~일 것이다  
~일지도 모른다 -> ~될 가능성이 있다  
어떻게해도 -> 어떠한 수단을 사용해도  
~재미있다 -> ~는 중요하다  
~가 잔뜩있다. -> ~가 다수 존재한다  
~하는 편이 좋을 것 같다 -> ~해야한다.  
~는 싫다 -> ~는 적절하지 않다  
합격시켜주세요. -> 기여할 수 있다고 생각된다  
~라는 건 사기다 -> ~가 진실이라는 증거는 없다  
~는 읽고 싶지 않다 -> ~를 정당하게 평가하는 것은 어렵다  
~는 읽지 않았다 -> ~의 평가는 정해지지 않았다  
이제 끝났다 -> 의논이 고착상태이므로 원점으로 돌아가 보기로 한다.  
하기 싫어졌다 -> 이만 펜을 내려놓기로 한다.  
~라 느낀다 -> ~라 추측된다  
~하는 편이 좋을 것 같다 -> ~해야한다 ~할 필요가 있다  
~는 없었다 -> ~라는 사실은 알려져 있지 않다.  
~라는 건 틀렸다. -> ~라는 주장은 잘못이다.  
이런얘기는 들은 적이 있다. -> 이러한 사례를 들 수 있다.  
다들 ~라고 한다 -> 일반적으로 ~라고 한다.  
다들 ~라고 생각한다 -> ~라고 널리 알려져있다.  
교과서에 ~라고 적혀있다 -> ~라는 것이 통설이다.  
연구했습니다만~는 모르겠네요. -> ~는 앞으로의 과제로 삼고싶다.  
[실정]으로

## 성남시 구직자, AI 자소서 '애스크토비' 쓴다



2024 성남시 채용박람회 방문객들이 그레이비랩의 인공지능(AI) 자소서 서비스 '애스크토비(ASKTobi)'를 체험하고 있다. /사진=그레이비랩

성남시 일자리지원센터에서 인공지능(AI) 자소서 서비스 '애스크토비(이하 ASKTobi)'로 기업 맞춤형 이력서를 작성할 수 있게 됐다.

이번 협력으로 성남시 일자리 지원센터는 ASKTobi를 활용해 구직자에게 보다 전문적인 자소서 작성 컨설팅을 제공하게 된다.

ASKTobi는 AI 기술을 통해 구직자가 기업에 본인을 어필할 수 있는 자기소개서 스토리 초안을 제공하는 커리어 관리 서비스다.

사용자는 본인의 이력 정보를 입력한 뒤 원하는 포지션을 검색하면 실시간 채용 공고를 확인할 수 있다. 자소서 질문 9가지가 제공돼 질문별 초안을 생성할 수 있다.

특히 대기업 공채의 경우 기업별 자소서 질문이 연동돼 자소서 준비에 직접적인 도움을 받을 수 있으며, 영문 기능을 제공해 해외 취업 준비에도 활용 가능하다



레포트컴퍼니

### 자기소개서 컨설팅 비용안내

| 구분       | 글자수     | 서비스비용   | 소요시간 | 추가비용                   |
|----------|---------|---------|------|------------------------|
| 자기소개서    | 500자 기준 | 40,000원 | 1~3일 | 빠른 서비스(1일 이내) + 10,000 |
| 1분 자기소개서 | 500자 기준 |         |      |                        |
| 경력기술서    | 500자 기준 |         |      |                        |

| 구분            | 글자수     | 서비스비용   | 소요시간 | 추가비용                   |
|---------------|---------|---------|------|------------------------|
| 대학·대학원 편입 자소서 | 500자 기준 | 50,000원 | 1~3일 | 빠른 서비스(1일 이내) + 10,000 |
| 대학·대학원 학업계획서  | 500자 기준 |         |      |                        |



## 취업실전전략

### 합격자소서

인적성·면접후기

인성역량검사

공기업 모의고사

직무인터뷰

## 역량테스트(MICT)

## 취업 TOOL

글자수 세기

어학점수 변환

사진크기조정

자소서 자동완성

자소서·면접 사전

연봉계산기

퇴직금계산기

실업급여계산기

학점계산기

온라인 도장

## 문서서식 자료

## 합격자소서 ⓘ

| 지원분야   | 근무형태         | 학력   | 30대 그룹별   | 기업형태  |   |
|--|--------------|--|---|---|---|
| 기획·전략<br>법무·사무·총무<br>인사·HR<br>회계·세무<br>마케팅·광고·MD<br>개발·디자인 | ←대분류를 선택하세요. | <input type="checkbox"/> 신입<br><input type="checkbox"/> 인턴 | <input type="checkbox"/> 고등학교 졸업<br><input type="checkbox"/> 대학 졸업(2,3년)<br><input type="checkbox"/> 대학교 졸업(4년)<br><input type="checkbox"/> 석/박사 졸업 | <input type="checkbox"/> 교보생명보험그룹<br><input type="checkbox"/> 농협그룹<br><input type="checkbox"/> 대림그룹<br><input type="checkbox"/> 대성그룹<br><input type="checkbox"/> 롯데그룹<br><input type="checkbox"/> 메리츠금융그룹 | <input type="checkbox"/> 30대그룹<br><input type="checkbox"/> 매출1000대기업<br><input type="checkbox"/> 대기업<br><input type="checkbox"/> 공공기관 공사협회<br><input type="checkbox"/> 외국계<br><input type="checkbox"/> 중소기업 |

선택한 검색조건 **323건** [개발·데이터-전체](#) × [초기화](#) [선택 조건으로 검색](#)

## 총 323건

[✓ 관심기업 합격자소서 보기](#) [✓ 전문가 분석 보기](#)

평점높은순 ▼

기업명, 자소서 항목 검색 🔍



**한국투자증권(주)** 2020년 하반기 | 신입 | 웹개발자

[자소서 작성](#)  
👁 150,434

Q. 성장과정:가족사항, 학창시절, 교우관계, 생활습관, 자신에게 크게 영향을 미친 사건 등을 포함하여 ...

A. 저희 가족은 회사원이신 아버지와 어머니 그리고 누나와 저 4인 가족입니다. 학창 시절 독서에 관심...

전문가 총평 ★★★★★ 5 | "본 자기소개서의 장점은 모든 질문에 대해 정확하게 답을 했다는 것입니다. 이 말은 회사와 의사소통이 잘 되는 지원자라는 말입니다. 자신... "



# 합격 자소서 게시판 -> 지원분야 : 개발·데이터 -> 전문가 분석 보기 -> 평점 높은순 : 총 323건



```
url_list = ['https://www.jobkorea.co.kr/starter/PassAssay?FavorCo_Stat=0&Pass_An_Stat=1&OrderBy=2&EduType=0&WorkType=0&schPart=10031&isSaved=1&Page='  
            + str(x) for x in range(1, 18)]
```

```
links = []
```

```
for url in url_list:
```

```
    driver.get(url)
```

```
    time.sleep(3)
```

```
    question_items = driver.find_elements(By.CSS_SELECTOR, 'div.item.question')
```

```
    for item in question_items:
```

```
        a_tag = item.find_element(By.TAG_NAME, 'a')
```

```
        href = a_tag.get_attribute('href')
```

```
        links.append(href)
```

```
driver.quit()
```

|   |
|---|
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/236140?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/236140?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/236159?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/236159?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/201630?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/201630?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/200697?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/200697?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/194528?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/194528?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/193261?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/193261?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/192318?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/192318?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/192093?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/192093?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/192016?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/192016?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/191179?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/191179?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/191434?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/191434?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/191435?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/191435?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |
| <a href="https://www.jobkorea.co.kr/starter/PassAssay/View/191439?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1">https://www.jobkorea.co.kr/starter/PassAssay/View/191439?Page=1&amp;OrderBy=2&amp;FavorCo_Stat=0&amp;schPart=10031&amp;Pass_An_Stat=1</a> |

```

def crawling_url(url):
    driver.get(url)
    response = driver.page_source
    one_page = bs4.BeautifulSoup(response, 'html.parser')

    # 불필요 요소 제거
    for p_element in one_page.find_all('p', {'class': 'txSp11Chk'}):
        p_element.extract()

    for span_element in one_page.find_all('span', {'class': 'sup'}):
        span_element.extract()

    for a_element in one_page.find_all('a'):
        a_element.extract()

    # 공백 제거 및 텍스트만 추출
    kwd_list = []
    for text in one_page.find_all('div', {'class': 'tx'}):
        kwd_list.append(text.get_text(separator=' ', strip=True))

    return kwd_list

```

```

korean_essay_lists = []
for kwd_url in essay_urls:
    kwd_list = crawling_url(kwd_url)
    korean_essay_lists.append(kwd_list)
    time.sleep(5)

```

```

1 "[함께 하는 삶] 대학 시절, 선거 개표 아르바이트를 하면서 따뜻함을 실천한 적이 있습니다. 아르바이트 당시 개표
2 "1. 객체지향프로그래밍 A+,프로그래밍 수업을 들으면서 프로그램을 만들고 수정하는 과정이 어릴 때부터 좋아하던 수
3 "고등학교를 졸업하고 대입을 다시 준비하게 되었습니다. 학원을 통해 부족한 공부를 보충하고 싶었지만 무엇보다 부
4 "1) 00 모바일 앱 서비스 개발 - 기관명 : (주)00 - 기간 : 2015.01 ~ 2015.09 - 근무부서/직위 : 서비스 개발/개발
5 "저는 금융 it의 발전 가능성이 매우 크다고 생각합니다. 최근 국내의 금융산업 쪽에서 it 규제가 많이 완화되 다양한
6 제 삶에서 가장 큰 변화를 준 사람은 바로 아버지입니다. 아버지는 제가 초등학교 때부터 고등학교 때까지 대략 10년
7 "저를 돌이켜보니 저는 항상 새로운 것을 배우는 걸 좋아했던 것 같습니다. 어렸을 때부터 운동을 좋아해서 다양한 동
8 """"슈어소프트테크""에 지원한 이유는 현재의 위치에 멈추지 않고 끊임없이 혁신을 준비하여 새로운 도약을 준비하는
9 "[위기를 기회로 만드는 것은 도전 정신이다.] 고등학교를 졸업하고 대학 입시를 다시 한 번 도전하였습니다. 학원을
10 "[선택과 집중] 저의 첫 번째 선택은 카페, 영화관, 호프집, 사무보조 등 다양한 아르바이트 경험 입니다. 학업과 경
11 "[열정, 호기심 그리고 성취감] 제가 S/W 개발자가 되려는 이유는 제 열정을 불타오르게 하기 때문입니다. 먼저 문제
12 저는 결과에 대해 책임을 지는 사람입니다. 저의 아버지는 제가 초등학교 때부터 고등학교 때까지 대략 10년 정도를
13 "매년 Las Vegas에서 개최하는 CES(세계가전박람회)를 Buyer자격으로 참관하여 최고수준의 기술 미리 체험할 수 있었
14 "타 금융투자 회사에서의 인턴활동은 저를 금융IT에 관심을 갖게 만들어 주었습니다. 금융에 대한 학습을 하며 IT인으로서
15 "인생은 혼자 살아가는 것이 아닙니다. 함께 사는 것입니다. 아무리 개인의 능력이 뛰어나도 결국은 서로 돕고 사는
16 "[ 유니폼] 단체 생활에서 소속감을 가지고 있는 유니폼은 저를 표현할 수 있는 단어라 생각합니다. 일례로 대학교 입
17 "[ERP의 시작과 경험] ** 소프트웨어 멤버십 소속으로 1년 가까운 시간을 지냈고 면접이 얼마 안 남은 상황이었습니
18 "저는 밝고 긍정적인신 부모님의 영향으로 어디서나 밝고 긍정적이며 다양성을 존중하는 열린 마음을 갖고 있습니다.

```

```

# 데이터 번역 및 출력
for row in rdr:
    if len(row[0]) >= 3000:
        dot_index = row[0][3000:2500:-1].index('.')
        x1, x2 = row[0][:3000-dot_index+1], row[0][3000-dot_index+1:]
        input_text = x1 # 번역할 텍스트
        input_box = driver.find_element(By.CSS_SELECTOR, "textarea#txtSource")
        input_box.clear()
        input_box.send_keys(input_text)
        time.sleep(20)

        translate_button = driver.find_element(By.CSS_SELECTOR, "button#btnTranslate")
        translate_button.click()

        # 번역 결과 대기
        output_element = WebDriverWait(driver, 30).until(EC.presence_of_element_located((By.CSS_SELECTOR, "div#txtTarget")))

        # 번역 결과 추출
        output = output_element.text
        translated_x1 = output

        input_text = x2 # 번역할 텍스트
        input_box = driver.find_element(By.CSS_SELECTOR, "textarea#txtSource")
        input_box.clear()
        input_box.send_keys(input_text)
        time.sleep(20)

        translate_button = driver.find_element(By.CSS_SELECTOR, "button#btnTranslate")
        translate_button.click()

        # 번역 결과 대기
        output_element = WebDriverWait(driver, 30).until(EC.presence_of_element_located((By.CSS_SELECTOR, "div#txtTarget")))

        # 번역 결과 추출
        output = output_element.text
        translated_x2 = output
        sentence = translated_x1 + translated_x2

        translated_list.append(sentence)

# 입력 상자 초기화
input_box.clear()

```

['My family consists of four members: my father, mother, older sister  
 ng books and continued to read books in a wide range of fields, especi  
 nd Western philosophy. As a result, I have acquired the ability to und  
 e also building up my knowledge of the humanities and trying to imitat  
 ding books and who have desirable moral senses and humility, and I stu  
 aving school after self-study at night, following the example of since  
 ized that by having a big dream, the criminal can be reborn as a great  
 have a lifelong dream to live a better society." "There is a case in w  
 ing data allocated during the internship program at the AI research in  
 building strategic competencies and overcoming them. Having the most i  
 ictive accuracy calculations by seven laggle competitions, but volunta  
 capabilities by overachieving work-level overachievement. As a result,  
 s.", The securities industry majored in business administration hoping  
 with data-based strategies and insights. As I decided on my career pat  
 latform that is integrated with customers\' daily lives based on data  
 g data analysis and AI implementation capabilities and SW app developm  
 the improvement of our Korean investment, Ministock and Bankies FinTec  
 We are applying to the Trading Development Department and the BIZ Devel

A word cloud featuring the words 'project' and 'member' in large, bold, purple and green fonts respectively. Surrounding these central words are various other terms in different colors and sizes, including 'experience', 'customer', 'process', 'system', 'using', 'service', 'first', 'result', 'learned', 'field', 'school', 'program', 'various', 'knowledge', 'application', 'order', 'programming', 'important', 'people', 'problem', 'learning', 'analysis', 'learn', 'future', 'working', 'development', 'student', 'always', and 'development'.

using experience always  
system process student  
project  
service  
first  
result  
learned  
field  
school  
program  
various  
knowledge  
application  
order  
programming  
important  
people  
problem  
learning  
analysis  
learn  
future  
working  
development

activities solution member program process  
communication learned department first number  
various project create  
problem knowledge  
information business people  
service. skill games  
experience customer result  
design users

[illegible]



| 전체          |      | 5점          |    | 1점            |    | 비교 단어    |
|-------------|------|-------------|----|---------------|----|----------|
| project     | 1205 | project     | 82 | project       | 73 |          |
| experience  | 925  | people      | 57 | system        | 54 | people   |
| first       | 846  | experience  | 56 | design        | 50 |          |
| system      | 809  | using       | 56 | experience    | 47 | program  |
| using       | 798  | design      | 42 | computer      | 42 | games    |
| members     | 717  | program     | 41 | school        | 41 | problem  |
| various     | 624  | games       | 40 | first         | 41 | member   |
| information | 599  | first       | 39 | graduate      | 40 | service  |
| learned     | 593  | problem     | 39 | safety        | 40 | various  |
| people      | 563  | members     | 38 | software      | 40 | skills   |
| working     | 519  | service     | 37 | using         | 39 | business |
| school      | 494  | system      | 35 | communication | 39 | working  |
| problem     | 473  | various     | 35 | knowledge     | 35 | learned  |
| customers   | 467  | knowledge   | 32 | information   | 35 |          |
| knowledge   | 464  | information | 32 | korea         | 35 |          |
| service     | 455  | learned     | 31 | application   | 35 |          |
| result      | 448  | working     | 31 | management    | 34 |          |
| process     | 443  | business    | 31 | related       | 33 |          |
| skills      | 432  | school      | 29 | research      | 32 |          |
| security    | 424  | skills      | 29 | programming   | 31 |          |



## 프로젝트 과정

1. 전문가 평점이 매겨진 개발 직군 자소서들 url 크롤링 - 완료
2. 크롤링 해온 url들에서 자소서 텍스트 크롤링 - 완료
3. 크롤링 해온 자소서 텍스트 번역해서 크롤링(파파고 이용) (한글로 직접 처리하기에는 성능이 너무 떨어지기 때문) - 완료
4. 번역된 자소서들을 전처리후 시각화 / EDA - 완료
5. 언어 벡터화 모델링 - 진행
6. 분류 모델링
7. 웹 애플리케이션화 및 앱 배포(streamlit or flask)

**감사합니다.**