[붙임5] 2026년 소·중·대 산학협력프로젝트(캡스톤디자인) 결과보고서

2026년 전남대학교 소프트웨어중심대학사업
소·중·대 산학협력프로젝트(캡스톤디자인) 결과보고서

프로젝트명	RAG 기반 특정 도메인 자기소개서 직무 역량 자동 평가 시스템 설계 및 성능 비교 연구					
Github url 주소						
팀 명	독독(DockDock)			과제수행기간	2026. 3. 25. ~ 6. 23.	
지도교수	학 과	컴퓨터정보통신공학과		성 명	박수형	
프로젝트 수행인원 (※팀장은 첫줄에 기입)	학과(부·복수전공)	학년	학번	성명	연락처	E-Mail
	컴퓨터정보통신공학과	4	223321	박선주	010-6807-8651	rssn10@jnu.ac.kr
	컴퓨터정보통신공학과	4	223356	윤세희	010-6647-9575	ysh_1013@naver.com
참여 기업	기업명	멘토명		직위	연락처(HP)	E-Mail
	애드시티	조세근		대표이사	010-5583-5004	adct@adct.co.kr

위와 같이 2026년 전남대학교 소프트웨어중심대학사업
산학협력프로젝트(캡스톤디자인) 지원 프로그램 결과보고서를 제출합니다.

2026년 6월 20일

신청자명(대표학생) : 박선주

지도교수 : 박수형

전남대학교 소프트웨어중심대학사업단장 귀하

산학협력프로젝트(캡스톤디자인) 결과보고서(요약)

프로젝트명	RAG 기반 특정 도메인 자기소개서 직무 역량 자동 평가 시스템 설계 및 성능 비교 연구		
수행기간	2026. 3. 25. ~ 6. 23.	소요예산	161,584원
소요예산 세부내역	- 회의비: 60,600원 - SW활용비: 100,984원 - 재료비: -		

참여인원	구분	인원수	성명
	교수	1	박수형
	석박사과정	-	
	학부생	2	박선주, 윤세희
	기업체	1	조세근
	계	4	

추진배경

○ AI 기반 문서 평가 기술에 대한 사회적 수요가 증가
- 채용 시장에서 대량의 자기소개서를 일관된 기준으로 효율적으로 평가할 수 있는 자동화 기술의 필요성이 대두됨.
- 기존 키워드/단순 유사도 방식은 표현 다양성 미반영 및 짧은 요구사항과 긴 문서 간의 길이 불균형으로 인해 극단적인 예측 오차를 보이는 구조적 한계를 가진다. 따라서 신뢰성 있는 직무 표준 문서를 결합하고 LLM의 맥락적 추론을 활용하는 RAG 기반 시스템으로의 전환이 시급하다.

목표 및 내용

○ 백엔드 엔지니어 직무 자기소개서를 평가할 수 있는 RAG 기반 LLM 평가 시스템을 설계·구현
- NCS 문서를 지식베이스로 구축하고, 자기소개서를 항목별로 파싱하여 의미 기반 검색 및 LLM 평가를 수행하는 전체 파이프라인을 구현한다.
- 단순 점수 산출을 넘어 항목별 채점 근거와 기술 스택 식별 정보를 JSON 스키마 기반으로 출력하여 설명 가능한 평가 구조를 확립한다.

○ RAG 구성 요소별 성능 차이를 비교 실험을 통해 분석하고, 안정적인 RAG 구성 조합을 도출
- 청킹 전략(고정 길이/구조적/의미적), Pooling 방식(Mean/CLS/Max), Top-k 값을 변인으로 설정하여 총 27가지 조건에 대한 완전 요인 설계 기반 2단계 실험을 수행한다.
- 1차 표본(20건) 실험으로 상위 후보를 선별한 후, 2차 전체(100건) 합성 데이터셋 실험 및 실제 자소서(PDF) 검증을 거쳐 RMSE 15점 이하, Macro F1 0.80 이상의 정량적 목표 달성 여부를 검증한다.

기대효과

○ 자기소개서 직무 적합성 피드백 제공 가능
-NCS 표준에 기반한 구체적인 강·약점 분석 및 개선 제안 리포트(HTML)를 실시간으로 제공한다.

○ 다양한 직무 평가에 활용 가능
- 현재는 백엔드 직무에 특정하여 개발되었으나, 아키텍처 내 지식베이스 교체만으로 타 IT 직군이나 공기업 NCS 평가, 연구 제안서 심사 등 다양한 문서 평가 프레임워크로 확장 가능한 구조적 유연성을 가진다.

1. 프로젝트 개요

프로젝트명	**RAG 기반 특정 도메인 자기소개서 직무 역량 자동 평가 시스템 설계 및 성능 비교 연구**
주제영역	☑ 생활 ☐ 업무 ☐ 공공/교통 ☐ 금융/핀테크 ☐ 의료 ☐ 교육 ☐ 유통/쇼핑 ☐ 엔터테인먼트
기술분야	☐ IoT ☐ 모바일 ☐ 데스트톱 SW ☑ 인공지능 ☐ 보안 ☐ 가상현실 ☐ 빅데이터 ☐ 자동제어기술 ☐ 블록체인 ☐ 영상처리 ☐ 기타( )
성과목표	☐ 논문게재 및 포스터발표 ☐ 앱등록 ☑ 프로그램등록 ☐ 특허 ☐ 기술이전 ☐ 실용화 ☐ 공모전(*공모전명*) ☐ 기타( )

2. 프로젝트 추진배경

최근 채용 과정에서는 지원자의 자기소개서, 포트폴리오 등 비정형 문서를 기반으로 역량을 평가하는 방식이 확대되고 있으며, 이를 효율적으로 처리하기 위한 AI 기반 자동 평가 시스템의 필요성이 증가하고 있다. 특히 기업은 대량의 지원자를 짧은 시간 내에 평가해야 하기 때문에 기존의 수작업 기반 평가 방식에는 시간 및 인력 측면의 한계가 존재한다.

또한 기존의 자동 평가 시스템은 주로 키워드 기반 매칭 또는 단순 임베딩 유사도 기반 방식에 의존하고 있어, 문장의 맥락이나 직무 적합성과 같은 정성적 요소를 충분히 반영하지 못하는 문제가 있다. 이러한 한계로 인해 평가 결과의 일관성과 신뢰성이 낮아질 수 있다.

이에 따라 본 연구에서는 국가직무능력표준(NCS) 기반 직무 문서를 활용하여 평가 기준을 구조화하고, RAG(Retrieval-Augmented Generation)와 LLM을 결합한 문서 평가 시스템을 통해 보다 정교하고 설명 가능한 평가 구조를 구현하고자 하였다.

3. 프로젝트(주제) 목표 및 내용

본 연구의 목표는 백엔드 엔지니어 직무 자기소개서를 대상으로 RAG 기반 LLM 평가 시스템을 설계하고, 시스템 내부 구성 요소가 평가 성능에 미치는 영향을 분석하여 최적의 구성 조합을 도출하는 것이다.

이를 위해 먼저 NCS 기반 백엔드 직무 문서를 수집하여 지식베이스를 구축하고, 이를 기반으로 RAG 검색 환경을 구성하였다. 이후 자기소개서 평가를 수행하기 위해 문서 파싱 모듈, 임베딩 기반 검색 모듈, 평가 기준 검색 모듈, LLM 기반 평가 모듈로 구성된 전체 평가 파이프라인을 설계하였다. 또한 RAG 성능에 영향을 미치는 핵심 요소로 청킹 전략, Pooling 방식, Top-k 값을 설정하여 실험 변인으로 정의하였다. 청킹 전략은 고정 길이 청킹, 구조 기반 청킹, 의미 기반 청킹으로 구분하였으며, Pooling 방식은 Mean, CLS, Max 방식을 비교하였다. Top-k 값은 검색되는 문서의 개수를 조정하여 검색 범위가 평가 결과에 미치는 영향을 분석하였다.

실험은 완전 요인 설계를 기반으로 수행하였으며, 합성 자기소개서 데이터를 활용하여 다양한 조합의 성능을 비교하였다. 주요 평가 지표로는 점수 예측 오차를 측정하는 RMSE와 등급 분류 성능을 평가하는 F1-score를 사용하였다. 최종적으로 실험 결과를 기반으로 가장 안정적인 RAG 구성 조합을 선정하였다.

4. 시스템 구성 및 내용

본 연구에서 제안한 시스템은 백엔드 엔지니어 직무 자기소개서를 자동으로 평가하기 위한 RAG 기반 LLM 평가 시스템으로 구성된다. 전체 시스템은 크게 데이터 구축 단계, 검색 단계, 평가 단계로 구성된다.

먼저 데이터 구축 단계에서는 국가직무능력표준(NCS) 기반의 백엔드 직무 관련 문서를 수집하여 지식베이스를 구성하였다. 해당 문서는 평가 기준으로 활용되며, 문서 단위로 청킹되어 벡터 데이터로 변환된다. 이 과정에서 청킹 전략(고정 길이, 구조 기반, 의미 기반)에 따라 데이터 구조가 달라지도록 설계하였다.

다음으로 검색 단계에서는 임베딩 모델(KR-SBERT)을 활용하여 자기소개서 입력 텍스트와 지식베이스 간 유사도를 계산하고, Top-k 기준에 따라 관련 평가 기준 문서를 검색한다. 또한 Pooling 방식(Mean, CLS, Max)을 적용하여 청크 단위 벡터를 대표 벡터로 변환함으로써 검색 정확도에 미치는 영향을 분석하였다.

마지막 평가 단계에서는 검색된 평가 기준 문서를 LLM(GPT 기반 모델)에 입력하여 자기소개서를 평가하도록 설계하였다. LLM은 입력된 기준 문서를 바탕으로 점수 및 등급을 산출하며, 평가 근거를 함께 생성하도록 구성하였다.

전체 시스템 구조는 다음과 같은 흐름으로 요약된다.

자기소개서 입력 → 임베딩 생성 → NCS 지식베이스 검색 → Top-k 문서 추출 → LLM 평가 → 점수 및 등급 출력

5. 프로젝트 결과물에 대한 기술

본 연구의 주요 결과물은 RAG 기반 자기소개서 평가 시스템 구현 코드, 실험 데이터셋, RAG 설정 변수 실험 결과이다.

첫째, 시스템 구현 측면에서 자기소개서 입력을 자동으로 처리하고 평가하는 파이프라인을 구현하였다. 해당 시스템은 문서 파싱, 임베딩 생성, 유사도 검색, LLM 평가까지 전 과정을 자동화한 구조로 설계되었다.

둘째, 실험 데이터셋은 합성 자기소개서 데이터를 기반으로 구성되었으며, 실제 채용 데이터의 구조를 모사하여 다양한 평가 상황을 재현할 수 있도록 설계하였다. 모든 데이터는 사람이 직접 검수하여 평가 기준의 일관성을 확보하였다.

셋째, RAG 내부 구성 요소에 대한 실험 결과를 도출하였다. 청킹 전략, Pooling 방식, Top-k 값을 조합하여 완전 요인 실험을 수행하였으며, 각 조합별 성능을 RMSE 및 F1-score 기준으로 비교하였다. 그 결과, 고정 길이 청킹 + Mean Pooling + Top-k=5 조합이 가장 안정적인 성능을 보였다.

또한 일부 조건에서는 등급 분류 성능이 저하되는 현상이 관찰되었으며, 이는 RAG 기반 평가 시스템에서 검색 품질이 최종 평가 결과에 직접적인 영향을 미친다는 점을 확인한 결과이다.

구분	기능정의	세부기능 설명
파싱	자기소개서 파싱 및 분류	업로드된 자기소개서 문서 파일(PDF)에서 텍스트를 파싱하고 문항별 질문과 답변을 자동으로 분류하여 데이터 구조로 변환
지식베이스	NSC 직무 문서 정제	지식베이스 구축을 위해 수집된 백엔드 직무 NCS 문서의 노이즈를 제거
	텍스트 청킹 및 데이터베이스 적재	정제된 문서를 실험 조건(고정 크기, 구조 기반, 의미적)에 맞춰 분할, 분할된 청크를 ChromaDB에 저장
RAG 검색	연관 직무 기준 매칭	자소서 내용을 기반으로 쿼리를 생성하여 연관성이 높은 Top-k 청크를 추론
LLM 평가	직무 역량 맥락 추론	검색된 NCS 컨텍스트와 자소서 내용을 대조하여 직무 적합성을 종합 추론
	구조화된 결과 산출	프롬프트에 따라 점수, 등급, 강약점 피드백을 JSON 스키마 규격으로 생성
결과 출력	결과 리포트 시각화	산출된 평가 결과와 근거 데이터를 웹 형태(HTML)의 리포트로 화면에 출력

6. 프로젝트 진행내용

1) 참여인원 및 담당 역할

연번	소속학과	성명	수행역할 분담내용
1	컴퓨터정보통신공학과	박선주	데이터 적재 및 검색 파이프라인 구축 전담
2	컴퓨터정보통신공학과	윤세희	사용자 입력단 처리와 LLM 기반의 핵심 평가 엔진 구현 전담

2) 회의 및 SW멘토링 진행

번호	일시/장소	회의/멘토링 내용(상세히 작성)	관련 사진
1	2026. 4. 6. (19:00~) / 비대면	- 멘토링 개요 본 멘토링은 최종 성과 발표 및 결과물 도출을 위한 연구 방향 점검과 개선 사항 도출을 목적으로 진행되었다. 멘토는 멘토링을 수행한 경험을 바탕으로, 성과 발표에서 좋은 결과를 얻기 위한 실질적인 연구 설계 및 발표 준비 방향 제시 - 연구 방향 및 목표 설정 피드백 1. 단순 기능 구현이 아닌 명확한 목표 설정 필요 2. 연구 목표는 반드시 정량적 수치 필요(ex. 정확도, 정밀도 등 성능 지표) - 문제 정의 및 객관성 강화 1. 연구 배경에 대한 객관적 자료 확보 필요	

		2. 현재 연구 제안서 보완점: 근거 데이터 및 객관성 부족 - 테스트 데이터 및 성능 평가 기준 1. 테스트 데이터 구성 방식에 대한 명확한 정의 필요 2. 최소 기한을 정하고 테스트 결과 확보 필요 3. 성능 비교 연구 구조 필수(기존 방식 vs 제안 방식 비교)	
2	2026. 4. 16. (15:30~17:00) / 전남대학교 스토리움	- WBS 점검 및 이슈 정리 1. 테스트 데이터 생성: "나쁨" 레벨 문서 적정성 판별 필요 느낌, 테스트 데이터에 넣는 게 유의미한 결과를 가져올지에 관함. 2. 라벨 구성 비율 조정: 현재 훌륭, 좋음, 보통, 애매, 나쁨 총 5개의 라벨 존재. 테스트 데이터 각 2개씩 생성하도록 하였으나, 실제 문서의 수준 분포도와 유사하게 구성하는 것이 옳다 판단해 1, 3, 3, 2, 1 로 라벨 조정하도록함. 3. 테스트 데이터 적/부 판별 방법: 생성된 테스트 데이터에 관한 점수를 통해 적절한지 판단. 4. 각 알고리즘 구현 시작 - 중간 보고서 작성 내용 논의	
3.	2026. 4. 24. (14:00~16:00) / 전남대학교 스토리움	- WBS 점검 1. 생성할 테스트 데이터 문서의 양 재합의 → 도합 최소 100개 생성 2. 의미 기반 평가 방식 개선 방향 수정 – 파인튜닝 모델 비교는 후순위	
4.	2026. 5. 7. (15:00~17:00) / 전남대학교 스토리움	- WBS 재작성 - 작업 현황 공유 및 논의 1. 진행 완료: 문서 수집 방식 결정(저작권 확인 포함), 로그 자동 생성스크립트 작성, 문서 폴더 구조화, NCS 수집 세분화 조사/ 데이터 클래스 정의 및 테스트용 Mock 데이터 생성 완료 2. 주요 논의: 메타데이터 JSON 자동 생성 가능 여부 확인(구현 후순위), 구글 드라이브 업로드 후 공유 방식 확인 필요/ 실제 자기	

<table>
<tr>
<td></td>
<td></td>
<td>소개서와 RAG 내용 간의 유사도 결과값이 출력에 포함되도록 수정, RAG 완성될 때마다 연동 및 테스트 진행</td>
<td> </td>
</tr>
<tr>
<td>5.</td>
<td>2026. 5. 14.
(15:00~17:00)
/ 전남대학교 스토리움</td>
<td>- RAG 모듈 성능 실험 설계
단일 변인 실험 채택: 변인 간 교호작용을 고려한 실험이 현실적으로 불가능하다고 판단, 각 변인을 독립적으로 분석하는 설계 채택실험 순서 설정: 청킹 전략 → 임베딩 모델 파라미터 → top-k → 지식베이스 규모 (선택 사항)
<table>
<tr><td>항목</td><td>비교 내용</td></tr>
<tr><td>청킹 전략</td><td>고정 크기/구조적/의미적</td></tr>
<tr><td>임베딩 파라미터</td><td>Pooling 방식 변경(mean/cls/max)</td></tr>
<tr><td>검색 파라미터</td><td>top-k 값 조정(기본값: 3)</td></tr>
</table>
- 문서 수집 및 정제(파싱)
NCS 기반 데이터 수집: 출처 리스크(저작권)가 없고, 표준 문서로서 활용 가능/ SW아키텍처, 응용SW엔지니어링 등 8개 직종의 능력단위 레포트 및 직무기술서 등 4종류의 NCS 문서 수집데이터 가중치 적용 여부: 직무 순서나 수준에 따른 중요도가 명확하지 않고, 난이도와 중요도가 비례하지 않을 것으로 생각되어 가중치 부여가 힘들것으로 판단. 모든 문장에 동일 가중치 부여하기로(=가중치 없음) 결정</td>
<td> </td>
</tr>
<tr>
<td>6.</td>
<td>2026. 6. 1.
(19:00~)
/ 비대면</td>
<td>- 최종 발표 자료 점검 및 피드백
1. 제공된 '캡스톤디자인 프로젝트 가이드'를 참고하여 발표 구성 단계(문제 정의, 연구 방법, 실험 결과, 결론 등)에 따라 발표 자료를 체계적으로 구성할 수 있도록 안내받음
2. 발표 자료의 논리적 흐름과 시각적 구성(도표, 비교 결과, 시스템 구조도 등)에 대한 개선 방향을 제시받음
3. 발표 내용이 연구 목표 및 성능 평가 지표와 일관되도록 정리하는 과정에 대한 피드백 반영
4. 질의응답 대비를 포함하여 발표 리허설을 통해 전달력 및 이해도를 점검하고 보완할 수 있도록 지원</td>
<td> </td>
</tr>
</table>

7. 프로젝트 세부일정 및 내용

No.	작업 내용	3월				4월				5월				6월				담당자	비고
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4		
1	연구 기획 및 조사																		
2	초기 알고리즘 설계 및 구현																		
3	주제 변경 및 RAG 지식베이스 구축																		
4	평가 파이프라인 통합 및 테스트																		
5	종합 비교 실험 및 최종 점검																		
6	결과보고서 작성/제출																		

8. 결과물에 대한 향후 활용계획

본 연구 결과물은 단일 연구 프로젝트에 국한되지 않고 다양한 분야로 확장하여 활용할 수 있는 구조를 가진다.

우선 채용 분야에서는 기업의 서류 전형 자동화 시스템으로 활용 가능하다. 자기소개서를 직무 기준 문서와 비교하여 평가함으로써 1차 서류 검토 과정을 자동화할 수 있으며, 인사 담당자의 업무 부담을 줄이고 평가 기준의 일관성을 높일 수 있다. 특히 평가 결과와 함께 근거를 제공하는 구조는 기존 블랙박스형 AI 평가 시스템의 한계를 보완할 수 있다.

또한 API 형태로 확장할 경우 기존 채용 관리 시스템(ATS)과 연동하여 활용할 수 있으며, 기업 내부 인재 평가 시스템으로도 적용 가능하다.

향후에는 백엔드 직무뿐만 아니라 프론트엔드, 데이터 분석, AI 엔지니어 등 다양한 직무로 확장 가능하며, 평가 기준 문서만 교체하면 다른 문서 평가 문제에도 적용할 수 있다. 이를 통해 본 시스템은 범용 문서 평가 프레임워크로 발전할 수 있는 가능성을 가진다.

또한 장기적으로는 사용자 입력 기반 자동 직무 매칭 및 평가 시스템으로 확장하여, 자기소개서 작성 보조 도구 및 역량 진단 플랫폼으로 발전시킬 수 있을 것으로 기대된다.