

2023년 09월 19일 AI 엔지니어링 <JEOLMI'S FAMILY>

TEAM MEMBERS -









배지원 신유진 이신우 김혜진 우미림 정소이







JEOLMI'S FAMILY 구성 및 역할

훈련생	역할	담당
배지원	팀장/모델링	데이터 수집 및 전처리, 프롬프트 엔지니어링을 통한 QA셋 생성, BERT/LLM BERT/LLM 모델 파인튜닝, AWS EC2를 활용하여 모델 관련 flask서버 배포 포
신유진	팀원/모델링	서비스 적용 기대 효과 예측 GPT 싱글-턴 프롬프트를 활용 QA 셋 생성
이신우	팀원/모델링	모델링/QA 학습용 데이터 조사 및 선정 GPT 싱글-턴 프롬프트 활용 QA 셋 생성
김혜진	팀원/프론트	웹서비스 디자인 기획, JSP를 사용한 프론트엔드 화면 개발
우미림	팀원/백엔드	공지사항 및 회원정보 CRUD 구현, 회원가입 및 인증/인가 기능 구현, AWS(RDS/EC2) 활용 아키텍처 설계
정소이	팀원/백엔드	회원가입 유효성 검사, 가계부/챗봇 구현, 프론트앤드 화면 개발

CONTENTS

 1

 서론

 서비스 개요

02 웹 구현 웹 사이트 소개

03모델 구현챗봇 구현 과정 소개

04 참조문헌

하락하고 있는 청소년 금융이해력

그 이유는?

• 금융이해력이란

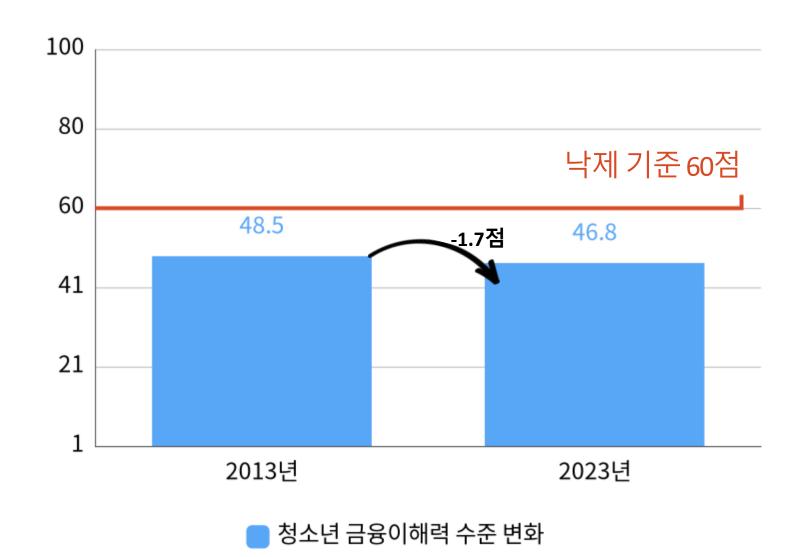
일상적 금융거래를 이해하고 실제 생활에서 금융 지식을 활용하며, 금융 선택에 따른 책임을 이해하는 능력

- 코로나 바이러스 이슈 코로나 바이러스 이후 금융교육의 빈도와 관심 하락
- 미흡한 교육
 학교 교육은 사실상 금융이해력 향상에 별다른 도움이 되지 않음
- 금융교육양극화현상

수도권 대비 지방 중/소도시의 학생들의 금융이해력 저하로 인한 평균 하락

청소년금융이해력수준

- 청소년금융교육협의회에서 고등학교 2학년 717명을 대상으로 금융이해력 확인 설문 실시
- 그결과,2023년금융이해력 평균점수는 "46.8점" -2013년도보다1.7점**하락**
 - 낙제점수(60점)를 크게 밑도는 수준
- 조기금융교육이미흡한상황에서이해력수준마저 크게 낮아최근 급증하는 사회초년생들의 금융사기 피해가앞으로도 더욱 늘어날수 있다는 우려증가



가계부 작성의 필요성

조기 예산관리방법 정립

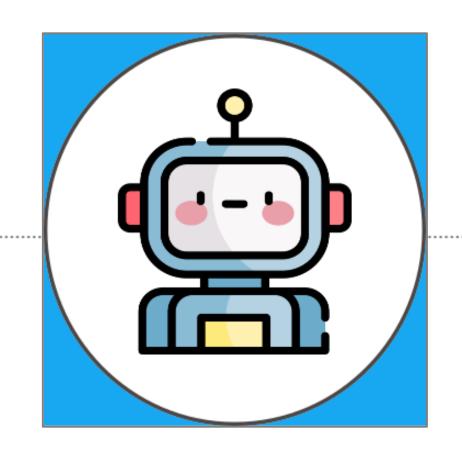
- 초·중·고 생활금융 교사용지도서에서 예산관리도구로 가장 먼저 가계부(용돈기입장)를 추천
- 용돈기입장,가계부등을통해예산을관리한다면 효과적인조기금융개념확립
 - -> 안정적으로 경제생활을 꾸려나갈 수 있음
- 예산:수입과지출을 함께 고려하여 세운 계획



Chat Bot 사용 이점

24시간 365일, 시간에 구애 받지 않고 사용 가능한 편리함

정보 제공을 통해 고객을 교육시킬 수 있는 교육 활용성



인건비를 절감하며 자동화 서비스를 제공할 수 있는 효율성

긴 대기 시간을 피하고 바로 응답 받을 수 있는 실시간 서비스

기대효과

금융지식이 올바르게 정착하며 금융상품 및 은행에 대한 개념 이해 향상

부분 상호작용으로 시너지 효과 기대 금융회사를 통해 저축하기를 희망할 것으로 기대

주관적 규범

계좌 개설에 대한 부모님의 지지

지각 된 행동통제

자립적 지출과 저축 목표달성을 위한 저축 객관적 금융지식

소득과 지출의 이해 화폐관리의 이해 저축과 투자의 이해 신용의 이해 금융시장의 이해 행동의사

규칙적인 용돈 수령 돈에 관한 가족과의 대화

가계부 기록 저축여부 지출 계획 세우기

행동

우리서비스는 Chat Bot을 통한 금융지식학습과 가계부를 활용한 예산관리가 가능해 청소년 금융이해력을 높일 것으로 기대됩니다.



Chat Bot은 지루한 이미지인 금융의 흥미 유발을 돕고, 직접 금융 궁금증을 풀 수 있는 경험을 제공함 가계부 작성 비율이 상대적으로 낮은 청소년들에게 가계부 작성을 제시해 조기에 예산관리방법을 정립할 수 있도록 도움

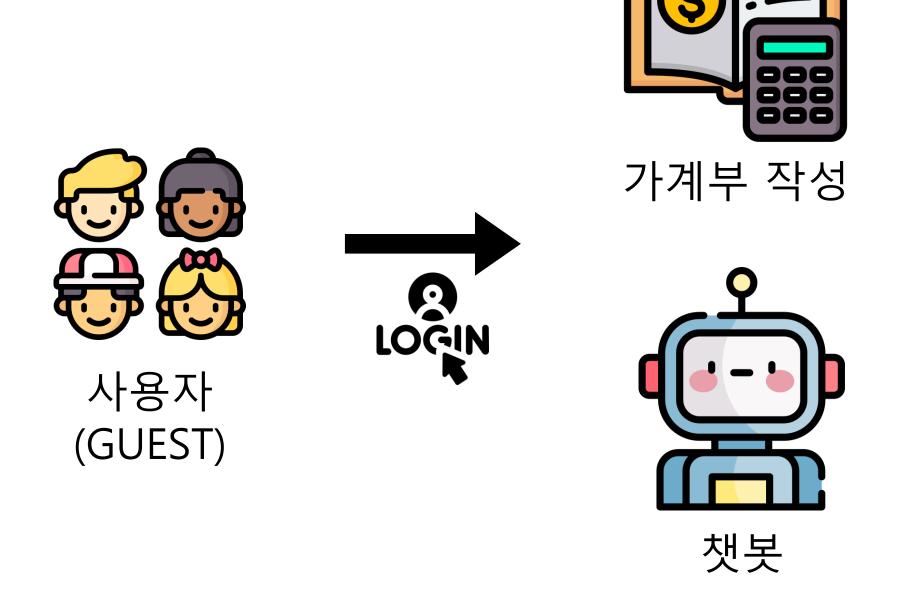
잠재적 고객 10대를 타깃한 금융서비스의 탄생

프로젝트 수행 절차

2023. 08. 0 2 ~ 2023. 09. 18

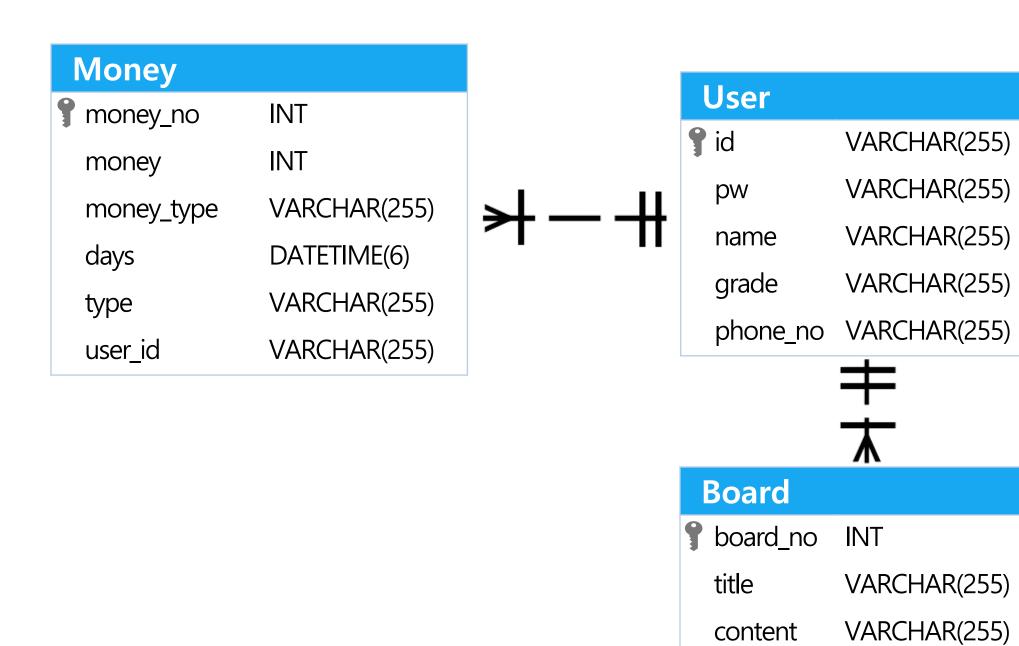
구분	기간	활동	비고
사전 기획	8/3(목)~8/9(수)	- 프로젝트 주제 확정 - 기획안 및 요구사항 산출물 작성	
데이터 수집	8/10(목)~8/14(월)	- 외부 데이터 수집 - Chat GPT를 활용한 데이터 추가 생성	
데이터 전처리	8/14(월)~8/18(금)	- 데이터 정제 및 정규화	
모델링	8/18(금)~9/18(화)	- 모형 구형	- 최적화, 오류 수정
서비스 구축	8/9(수)~9/18(화)	- AWS(EC2, RDS) 구축 - 웹 서비스 기능 구현	- 최적화, 오류 수정
총 개발기간	8/2(수)~9/18(화)		

1. 웹 서비스 소개





2. 데이터베이스 ERD



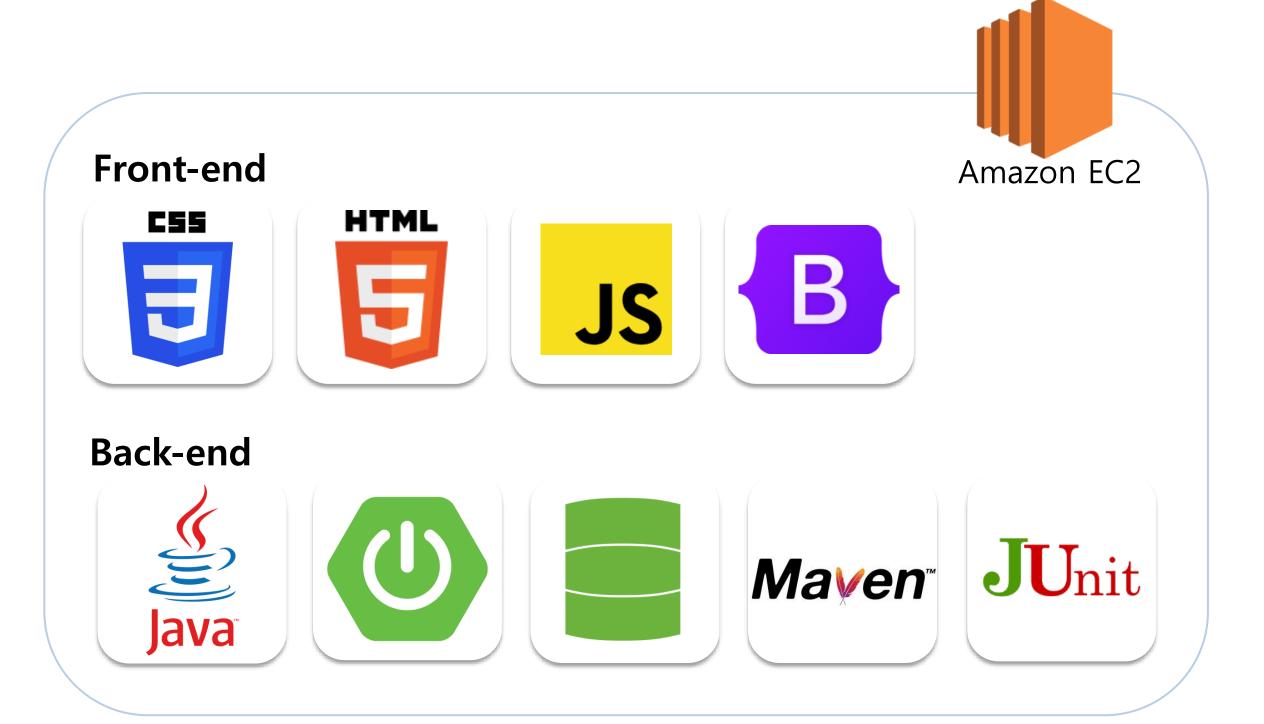
VARCHAR(255)

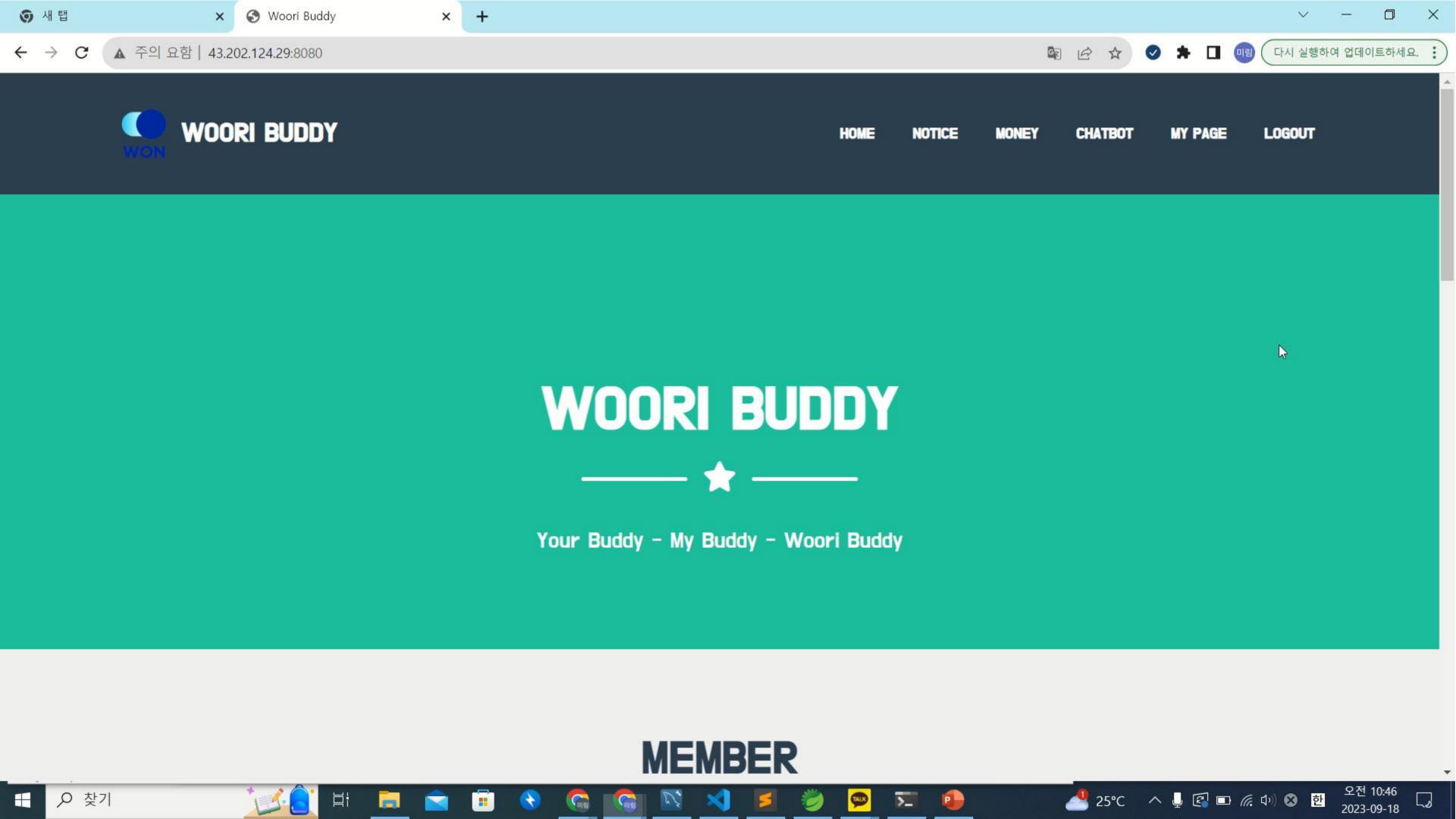
user_id

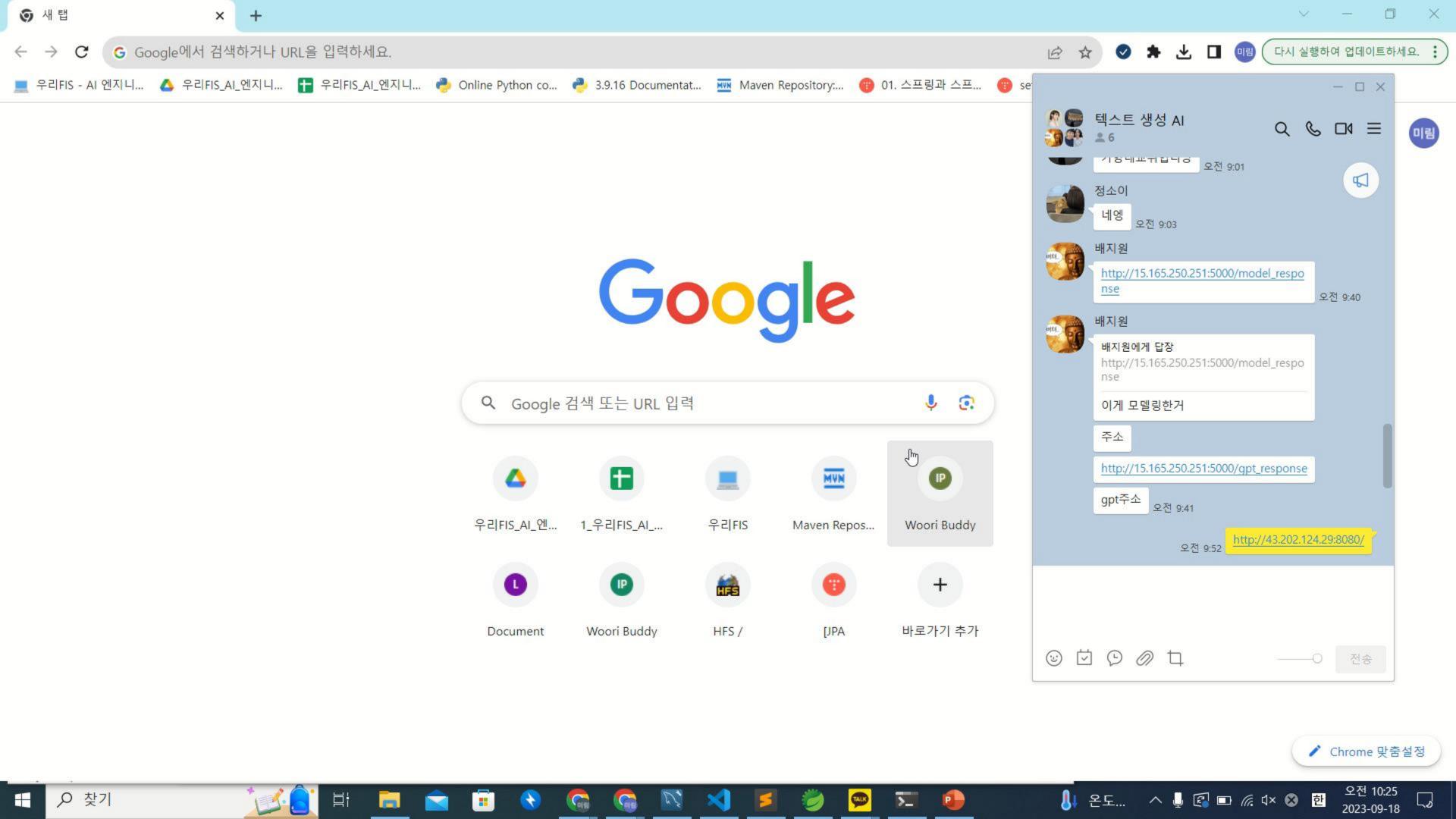
	Chat bot	
	chatbot_no	INT
	input	VARCHAR(100)
\(output	VARCHAR(100)
	best_context	VARCHAR(1000)
	feedback	VARCHAR(5)
	user_id	VARCHAR(20)

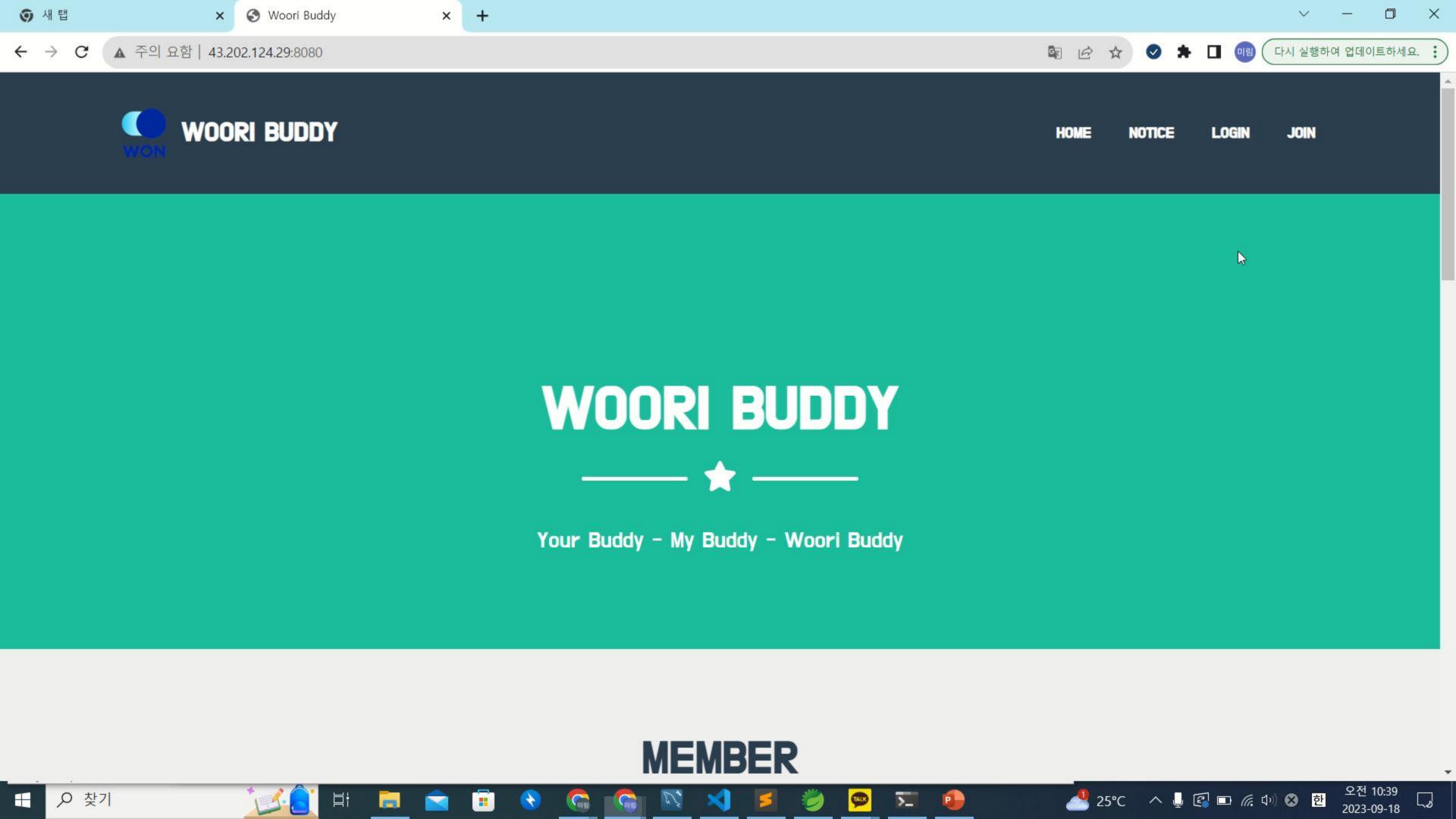
3. 웹 기술 스택

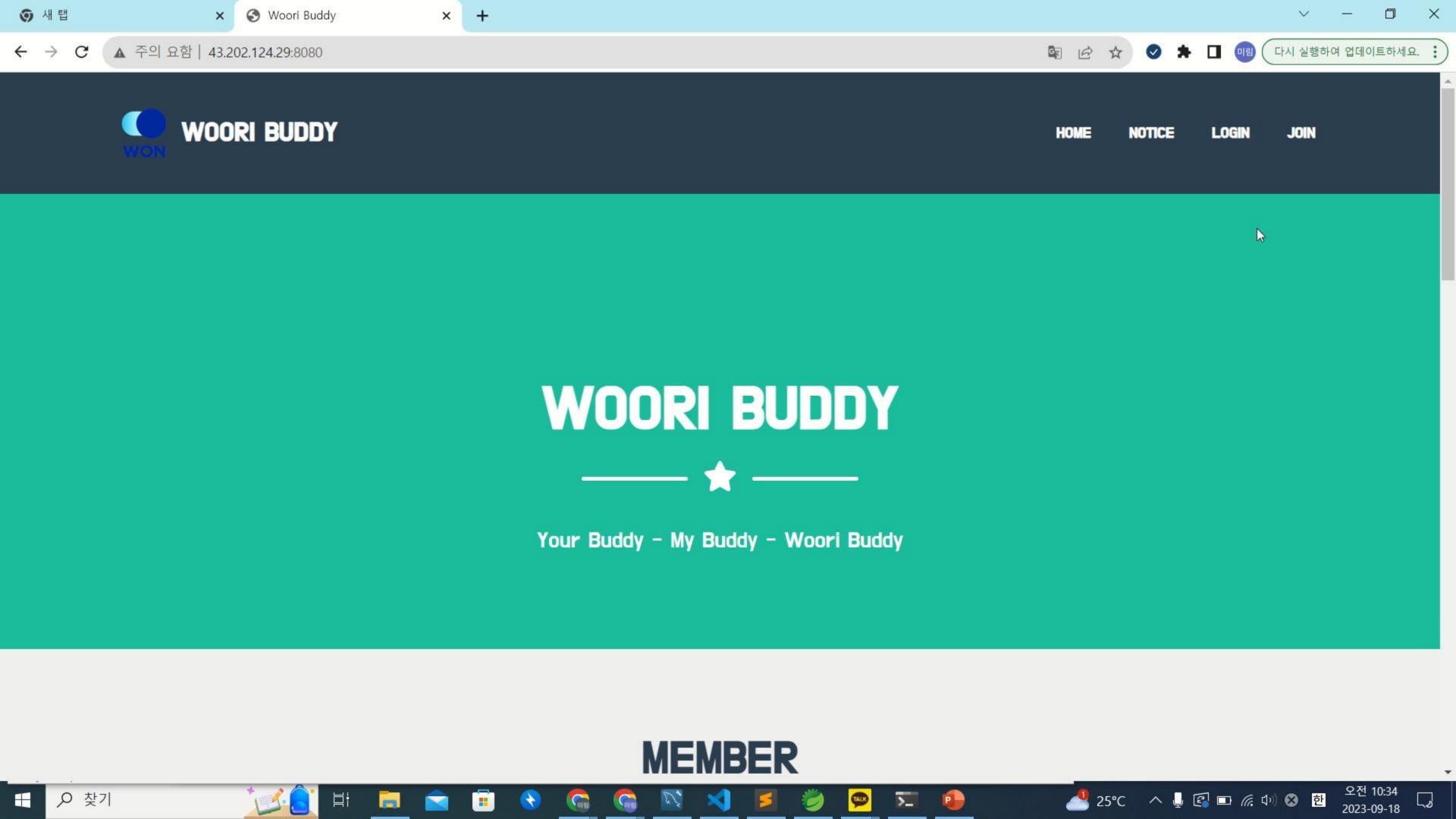


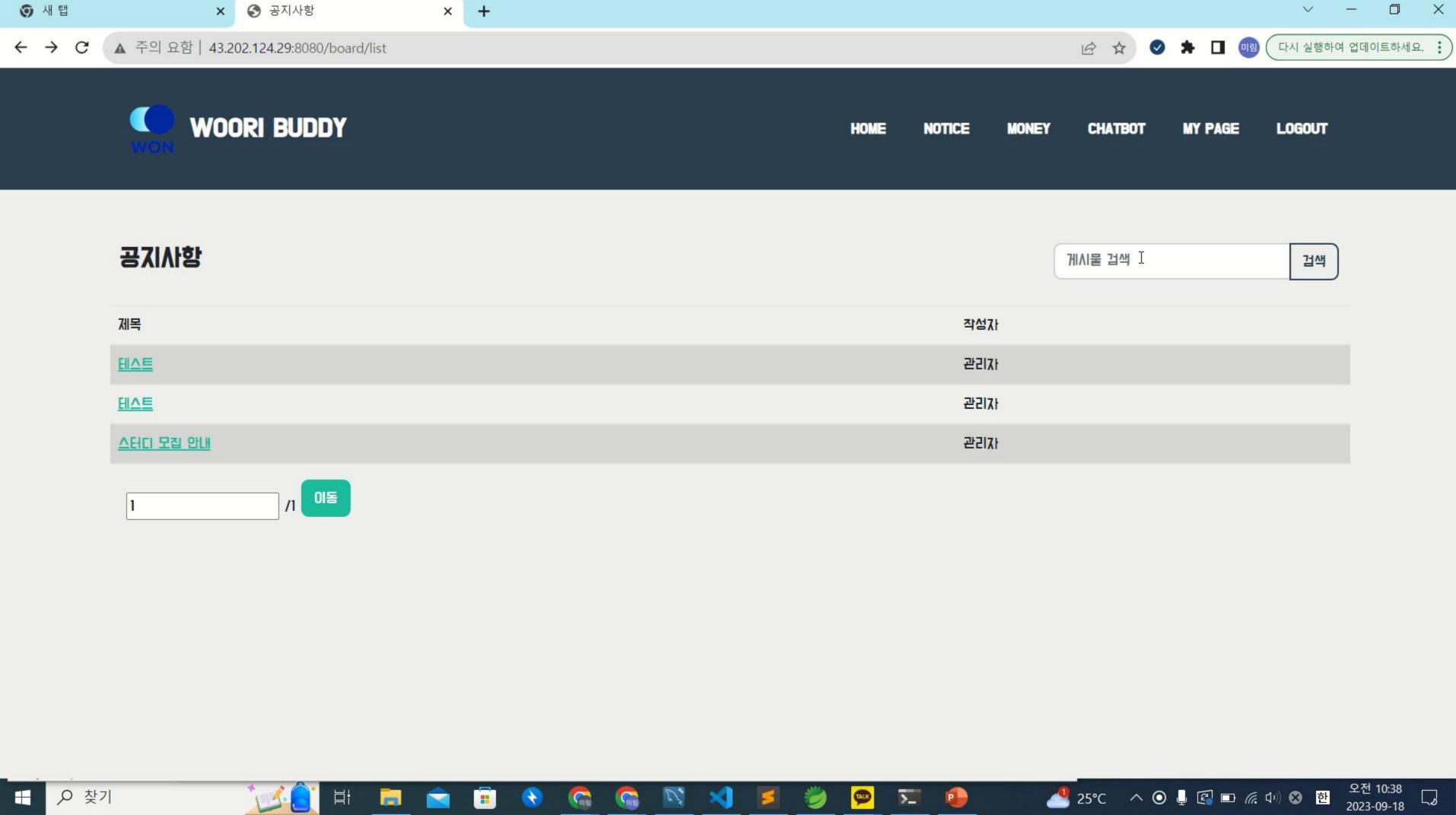


















































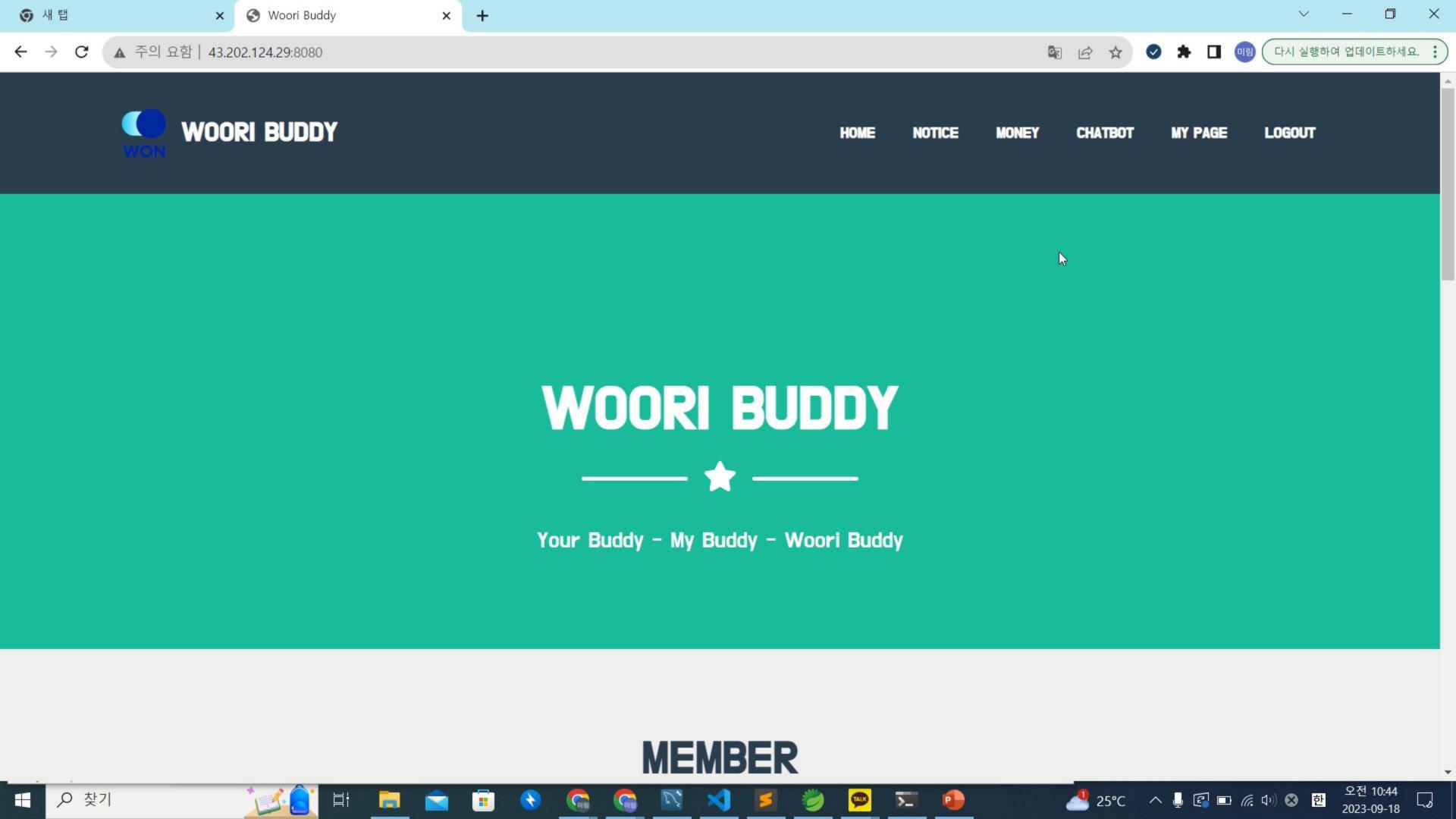












1. 서비스 프레임워크

모델링과 관련된 전반적인 서비스 프레임워크는 다음과 같음

1. 데이터 수집 및 전처리

2. DB 구축

3. 모델링

4. GPT - 4



G X Google

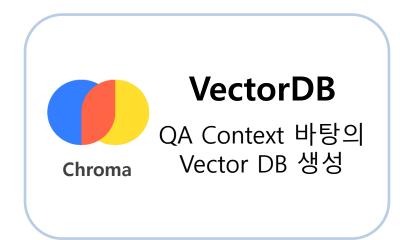


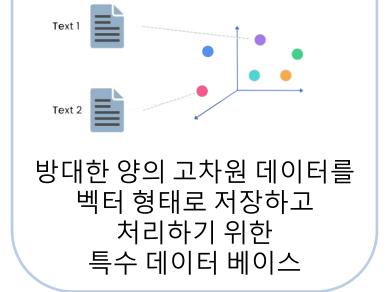




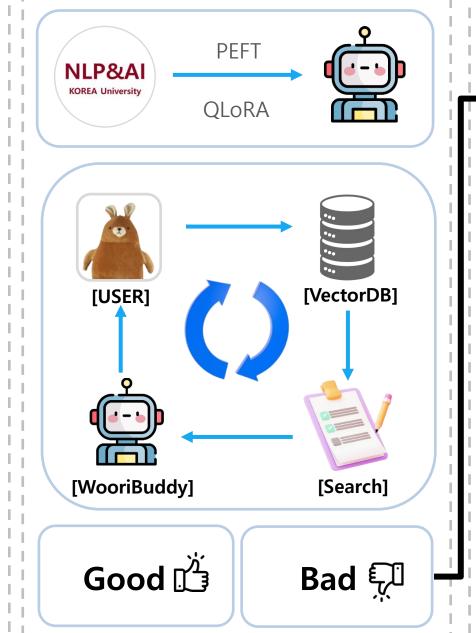


약 44K의 **QA** context data





VectorDB?





금융, 법률 문서 기계독해 데이터

(context_id	context	question	answer_text	answer_start	answer_end
0	2415370	`22.3월말 보험회사 대출채권 현황₩n1. 개요 ₩n□22.3월말 보험회사 대출채권	경기 변동성이 커지는 것에 대응하기 위해 어떻게 해	연체율 등 대출 건전성 지표를 지속 모니터링	432	455
1	2415370	`22.3월말 보험회사 대출채권 현황₩n1. 개요 ₩n□22.3월말 보험회사 대출채권	어떤 방법으로 경기 변동성이 커지는 것을 대처해	연체율 등 대출 건전성 지표를 지속 모니터링	432	455
2	2415371	2. 보험영업(수입보험료)₩n□ 생보사 : 120조 5,457 억원으로 전년 대비 9,5	보험회사가 커지는 불확실성에 대해 어떻게 준비 해야 돼	장기 수익성 및 자산운용 리스크 관리를 강화하고 재무건전성을 확보	466	501
3	2415371	2. 보험영업(수입보험료)₩n□ 생보사 : 120조 5,457 억원으로 전년 대비 9,5	불확실성이 커지는 보험회사는 어떤 방법으로 대 비를 해야 돼	장기 수익성 및 자산운용 리스크 관리를 강화하고 재무건전성을 확보	466	501
4	2415372	□금융감독원은 「2022년 금융감독 온라인 업무설명 회」를 유튜브 및 페이스북을 통해	금감원은 어떤 방법으로 온라인 업무설명회를 열 기로 했어	유튜브 및 페이스북을 통해	32	45

18755	2699778	5. 지역의 새로운 금융지원체제 확충방안₩n창업·투 자 활성화를 위한 지역의 금융지원	보증 중심의 금융관행에서 탈피하여 어떤 조치를 취해서 기술을 기반으로 한 사업화자금	기술금융제도를 재편	231	240
8756	2699779	크라우드펀딩을 통한 자금조달₩n크라우드펀딩은 아 이디어만으로 자금조달이 가능하며 아이	투자자가 금융사기 때문에 손해를 입을 가능성을 감소시키려면 어떻게 해야 돼	크라우드펀딩플랫폼 관리자들의 적격투자대상 창조 기업 선정 노력 등의 모니터링이 강화	467	511
18757	2699779	크라우드펀딩을 통한 자금조달₩n크라우드펀딩은 아 이디어만으로 자금조달이 가능하며 아이	어떤 방법으로 투자자가 금융사기로 인한 손실을 볼 가능성을 줄일 수 있어	크라우드펀딩플랫폼 관리자들의 적격투자대상 창조 기업 선정 노력 등의 모니터링이 강화	467	511
18758	2699780	셋째, 코넥스(KONEX)시장 상장을 창조기업의 자금조 달 수단으로 충분히 활용할 필	여러 가지 무형자산들을 어떻게 원활하게 상업화 시킬 수 있어	적절한 가치평가를 통해	430	441
18759	2699780	셋째, 코넥스(KONEX)시장 상장을 창조기업의 자금조 달 수단으로 충분히 활용할 필	어떤 방식으로 다양한 무형자산들을 순조롭게 상 업화시킬 수 있어	적절한 가치평가를 통해	430	441
18760 row	s × 6 column	is				

데이터	데이터 내용	데이터 수집 및 전처리
	• 금융 텍스트 데이터를 대상으로 한 MRC	
Al 🚰 Hub	QA학습 구축 데이터	• [수집] AI 허브에서 다운로드
금융, 법률 문서 기계독해 데이터	• 데이터 중 금융 분야 데이터만 사용	• [전처리] csv로 파일 형식 변경
	(총 18,760개)	

대학생을 위한 실용 금융 텍스트 데이터

	question	context	answer	answer_start	answer_end
0	생산물시장이 뭐야?	경제활동은 우리 생활에 필요한 재화와 서비스를 생산· 분배·소비하는 활동을 말	생활에 필요한 재화와 서비스가 거래되는 시장	237.0	261.0
1	생산물시장에서 공급자는 누구야?	경제활동은 우리 생활에 필요한 재화와 서비스를 생산· 분배·소비하는 활동을 말	기업은 생산물시장에서는 공급자	118.0	134.0
2	정부가 시장 경제를 위해 뭘 해?	정부는 국방, 치안, 사법제도를 유지하고 국·공립학교의 운영과 같은 수많은 공공서비	정부는 국방, 치안, 사법제도를 유지하고 국·공립학교의 운 영과 같은 수많은 공공서비	0.0	51.0
3	직접투자가 뭐야?	우리 기업이 생산한 생산물을 외국으로 수출하기도 하고 석유, 식량, 원자재 같은 것	회사를 설립하거나 인수하여 직접 운영	221.0	241.0
4	직접투자의 장점이 뭐가 있을까?	우리 기업이 생산한 생산물을 외국으로 수출하기도 하고 석유, 식량, 원자재 같은 것	직접 투자는 자본이 들어오는 것뿐만 아니라 기술도 같이 들어오는 장점이 있다	255.0	297.0
		***		***	***
1080	주민번호 도용을 확인할 수 있는 무료 방법은 무엇인가?	NICE평가정보의 마이크레딧, SCI(서울신용평가정보)의 사이렌24, KCB(코리아	주민번호 클린센터(clean.kisa.or.kr)를 이용하여 주민번호 이용내역을 조	130.0	184.0
1081	금융소비자 보호에서 가장 큰 이슈로 부각되고 있는 것은 무엇인가요?	최근 금융소비자 보호에서 가장 큰 이슈로 부각되고 있 는 것 중 하나가 금융사기이다	금융사기	39.0	43.0
1082	어떤 현실을 이용한 대학생 대상 사기가 증가하고 있나 요?	최근 금융소비자 보호에서 가장 큰 이슈로 부각되고 있 는 것 중 하나가 금융사기이다	취업난, 등록금 부담 등 어려운 현실을 이용한 대학생 대 상 사기	93.0	128.0
1083	피해예방을 위해 어떤 경우에도 금전의 이체를 요구하거 나, 금융거래정보를 수집하지 말	이와 같이 다양한 금융사기에 피해를 입지 않으려면 아 래와 같은 피해예방 요령을 명심	공공기관 및 금융회사	60.0	71.0
1084	전화나 문자 메시지를 통한 대출 광고에 응하면 어떤 사기 업체의 대출 광고에 노출될	이와 같이 다양한 금융사기에 피해를 입지 않으려면 아 래와 같은 피해예방 요령을 명심	사기업체의 대출 광고	387.0	398.0
1085 m	ows × 5 columns				

데이터	데이터 내용	데이터 수집 및 전처리
금융감독원 FINANCIAL SUPERVISORY SERVICE 대학생을 위한 실용 금융 텍스트 데이터 터	• 대학생을 위한 실용 금융 데이터(총 1,085개) 개)	 [수집] 웹 스크래핑 BS4, requests를 활용한 context 데이터 크롤링 [전처리] csv 형식으로 파일 형식 변경

2-1 데이터 수집 및 전처리

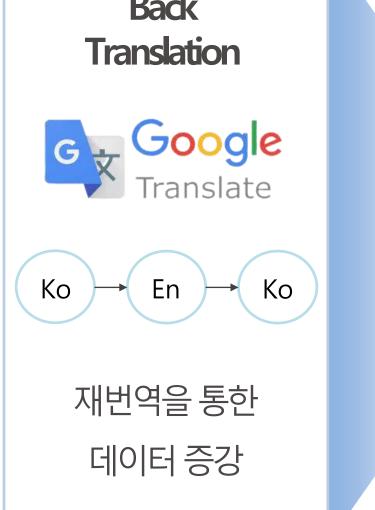
Al Hub에 올라온 SQuAD형태의 '금융, 법률 문서 기계 독해 데이터' 외에도 직접 생성한 QA셋을 사용함

GPT 3.5 turbo 기반의 싱글-턴 프롬프트를 사용하여 청소년이 물어볼 법한 QA셋을 생성함

이를 통해 생성된 웨일 위치 보세국과 배우나 PI를 통해 재번역하여 증강했음



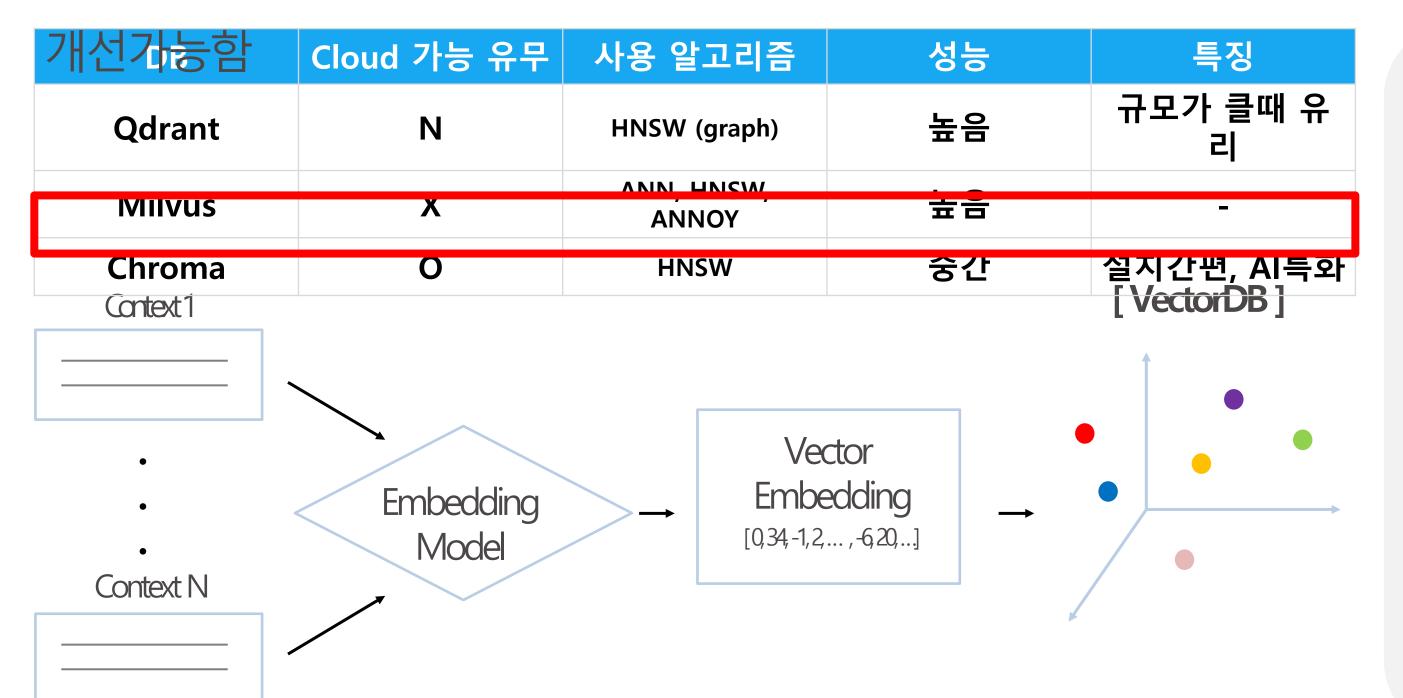




약 44K의 QA context data

2-2 DB 구축

VectorDB를 사용할 경우 문서 데이터를 벡터 형태로 저장하고, 유사성 검색 작업을 빠르게 할 수 ₩₩F 벡터 간 코사인 유사도 계산 시 발생하던 계산비용 및 부정확한 계산 결과 문제를



[VectorDB]

이미지, 텍스트 데이터 등 대규모 고차원 벡터 데이터를 효율적으로 저장하고 유사성 검색에 최적화된 데이터베이스 시스템

현 프로젝트에서는 사용하는 문서양이 많지않음 설치가 간단하며 AI에 특화된 ChromaDB 선택

3-1 BERT 계열 QA 챗봇(1)

처음에는 BERT계열의 QA챗봇을 제작하여 사용하고자 했음

금융 도메인에 특화된 KB-ALBERT를 비롯하여 다양한 BERT계열의 모델을 학습한 결과 성능은

Freined model

	F1	EM
KO-BERT	79.15	51.75
KoELECTRA - Base - v3	92.13	84.14
KB-ALBERT	93.45	84.83

· KO - BERT

SKT Brain에서 구축한 BERT 기반 기계번역 모델 (양방향 자연어 처리 모델)

· KoELECTRA - Base - v3

한국어 자연어 처리를 위한 작은 규모의 사전 학습된 언어 모델 (효율적인 학습과 뛰어난 한국어 텍스트 처리 성능 제공)

· KB - ALBERT

KB국민은행이 제공하는 경제/금융 도메인에 특화된 한국어 ALBERT 언어모델

모델 성능 판단

	F1	EM
KO-BERT	67.93	52.80
KoELECTRA - Base - v3	90.75	84.42
KB-ALBERT	91.94	85.13

모델 설정

max_seg_length = 512 max_query_length = 64 doc_stride = 128 batch_size = 16 learning_late = 3e-5 epochs = 10

3-2 BERT 계열 QA 챗봇(2)

그러나 실제 챗봇을 사용한 결과 답변이 정성적으로는 정확하지 않다는 문제가 발생함이는 사용한 데이터셋의 편향성과 BERT계열 모델 자체의 한계로부터 기인한 것으로 예상됨

[EM 정성적 평가]

정답

기업이 비교적 거액의 장기자금을 일시에 대량으로 조달하기 위하여 발행하는 일종의 차용증서이다.

예측

발행하는 일종의 차용증서이다.

답변이 정확하게 동일하지 않기도 하며, 외부 질문에 취약함

문제 1

데이터 셋의 한계

데이터 셋이 작고 깔끔하지 않다면 모델이 충분한 다양성과 복잡성을 학습하기 어려워 정확도가 떨어질 수 있음

문제 2

모델 구조의 한계

BERT 계열의 모델은 텍스트의 구조와 문맥을 이해하는 데에는 뛰어나지만, 학습하지 않은 외부지식이나 추론을 필요로 하는 질문에는 한계가 있음

4-1 LLM 모델 챗봇(1)

모델 및 데이터셋으로부터 발생한 문제를 해결하고자 LLM을 파인튜닝한 QA 챗봇을 제작하고자 此M은 BERT계열 모델보다 더 큰 데이터셋으로 학습되었기에 문제 해결에 용이할 것으로 판단됨

[언어에 대한 이해도 개선]

더 큰 데이터 셋으로 학습되었기에 BERT보다 더 깊은 언어 이해능력을 갖고 있음

[데이터 셋에 대한 문제 개선]

이미 다양한 문맥과 패턴을 학습했기에 데이터 셋의 패턴과 정보를 빠르게 인식 하면서도 일시적인 노이즈에 안정적

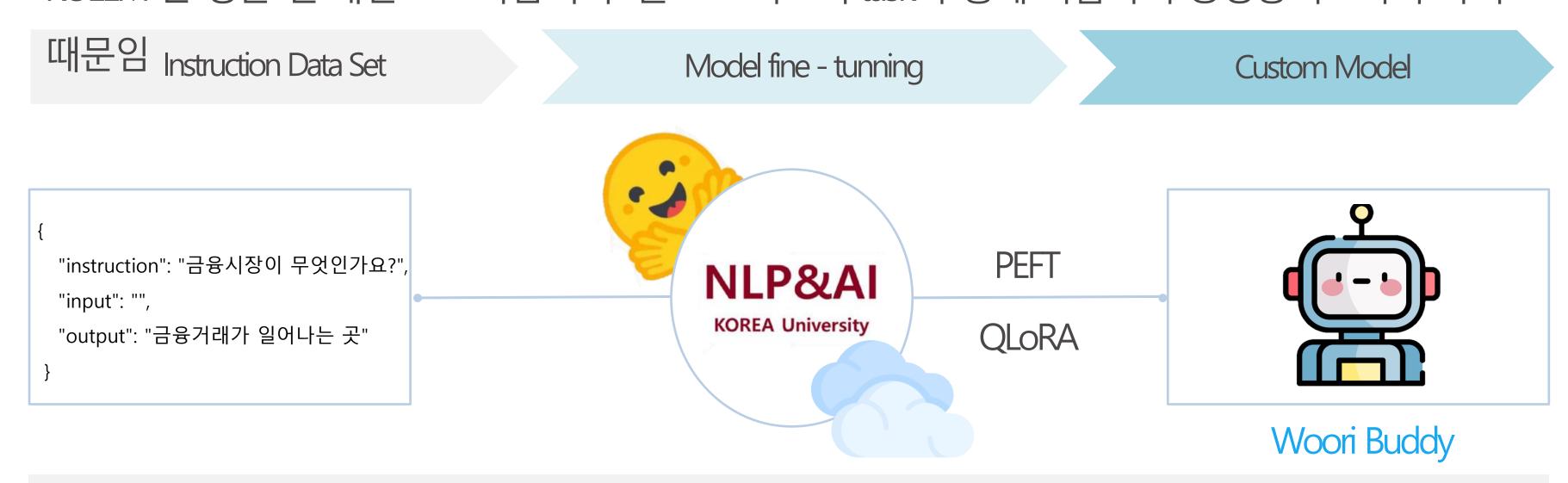
[문맥에 대한 이해도 개선]

더 큰 데이터 셋으로 학습 되었기에 주어진 문맥을 더 잘 이해하며 복잡한 질문에 대한 답변 가능

▶ sLLM을 파인 튜닝한 QA 챗봇 생성

4-1 LLM 모델 챗봇(2)

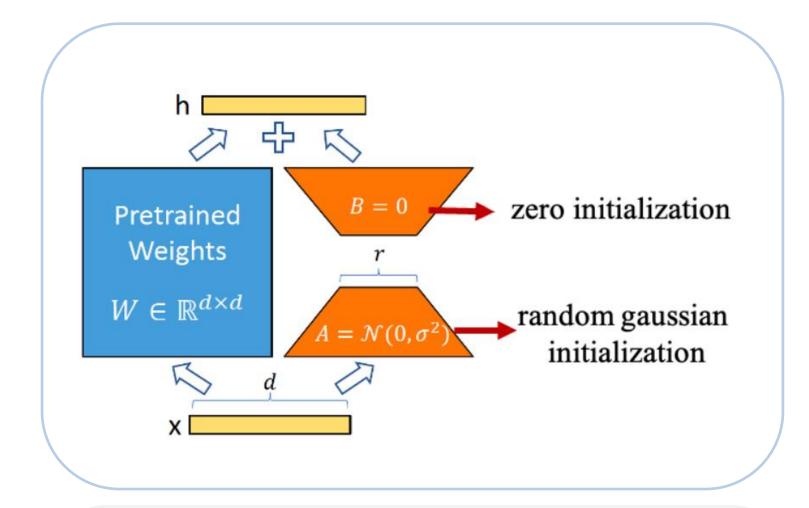
Base LLM은 고려대학교 NLP & AI 연구실과 HIAI 연구소가 개발한 KULLM을 사용함 KULLM 는 싱글-턴 패턴으로 학습되어 현 프로젝트의 task 수행에 적합하며 생성능력도 우수하기



컴퓨팅 자원 이슈로 인해 BackBone모델이 Polyglot-ko 5.8B인 KULLM모델을 QLoRA방법을 사용하여 파인튜닝

4-1 LLM 모델 챗봇(3)

LLM 모델을 full-finetuning시 막대한 time, resource cost가 발생하여 등장한 PEFT



[LoRA] Low-Rank Adaptation

고정된 weights를 갖는 pretrained model에 학습이 가능한 rank decomposition 행렬을 삽입

Fine-tuning과
Prefix tuning의 장점을 결합

PEFT
PEFT
PEFT

Parameter Efficient Fine Tuning

모델의 전체 파라매터/가중치가
아닌 일부 가중치만 파인튜닝하는
방법
일부 파라미터만을
파인튜닝하여
저장공간과 계산 능력 대폭
줄임

4-bit Quantized language model into Low-Rank Adaptation

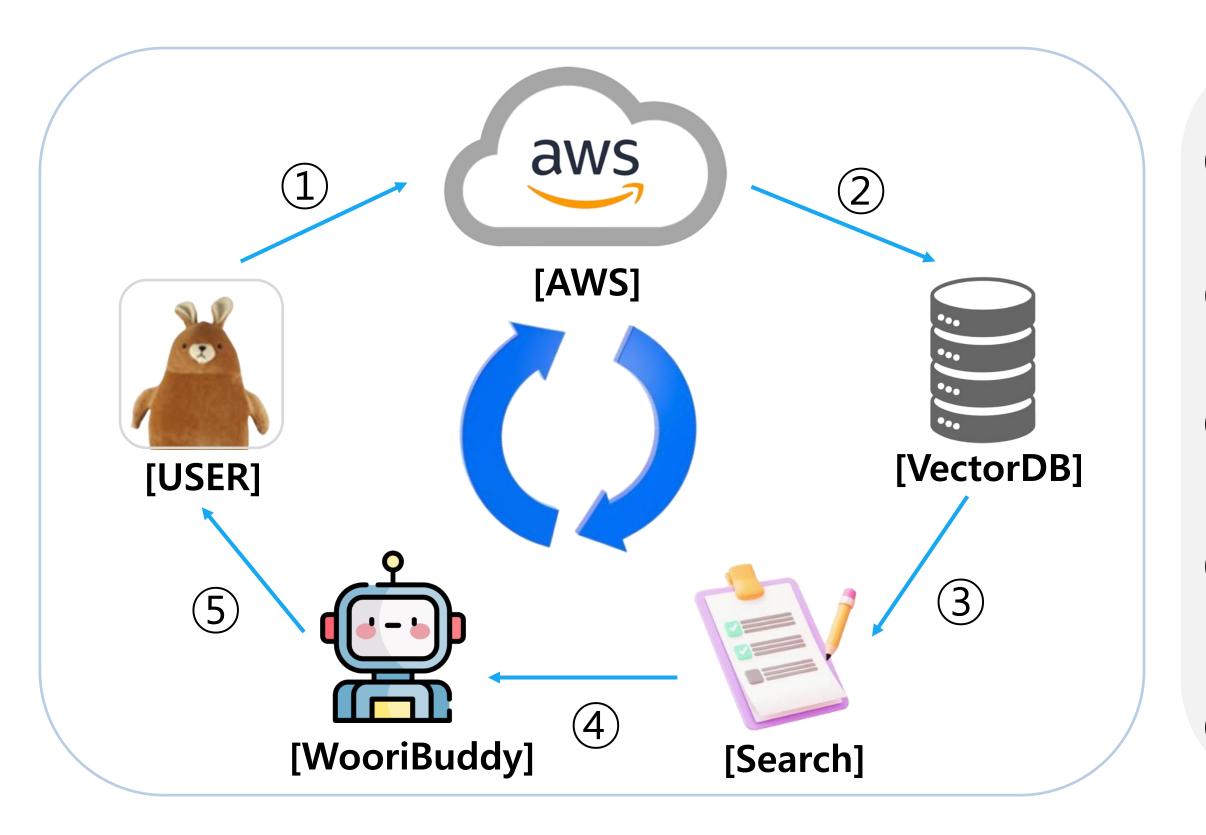
4-bit로 양자화된 모델을 가져와서 LoRA 기법을 활용해 파인튜닝하는 방법

학습 파라미터 수를 크게 줄여 기존 대비 메모리 사용량이 1/3이 됨

QLoRA

4-2 서비스 로직(1)

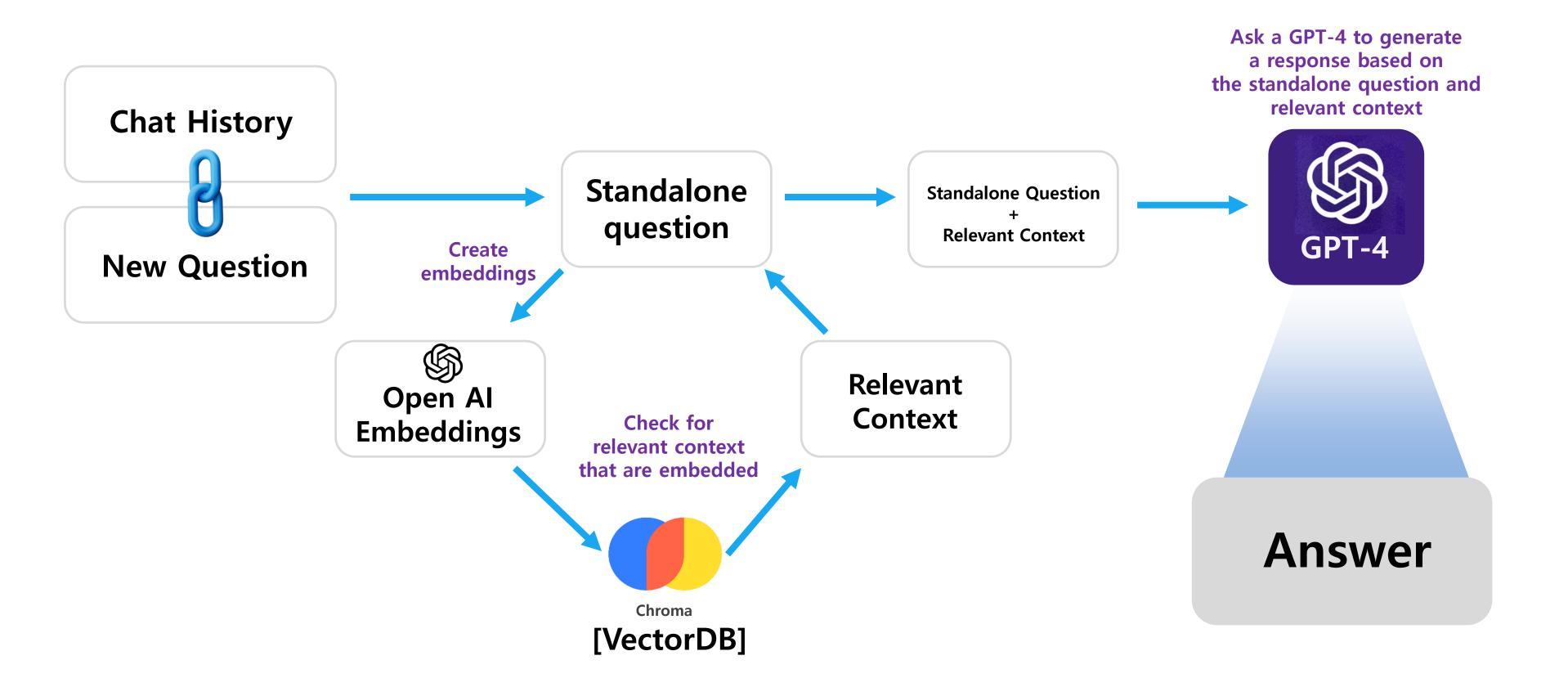
QA챗봇인 Woori Buddy를 활용하는 서비스 로직은 다음과 같음



- ① AWS에 올린 모델 서버에 사용자의 질문을 POST 형태의 API 요청을 보냄
- ② 사용자의 질문을 임베딩을 통해 벡터화
- ③ 임베딩을 통해 벡터화된 사용자의 질문과 유사한 Context를 찾음
- ④ Context를 LLM에게 전달하여 Context를 기반으로 답변을 생성하게 함
 - → 할루시네이션
- (5) 생성완헬문을 사용자에게 반환

4-2 서비스 로직(2)

만약 사용자가 질문에 대한 답변이 불만족스러운 경우 RetreiverQA를 활용하여 답변을 보완



WOORI BUDDY chatbot 안녕하세요. WOORI BUDDY chatbot 입니다. 궁금한 점이 있으신가요? Type message send

5. 문제 해결 리포트

Context를 참고할 때 사용했던 TF-IDF 알고리즘은 구현이 쉽지만 정확도가 낮다는 문제가 있었음 그래서 해당 문제점을 개선 시키고자 HNSW 알고리즘이 적용된 VectorDB를 사용했음

[알고리즘 및 DB관련 비교표]

	TF-IDF	HNSW
복잡도	하	상
비용	연산비용	초기구축비용
검색성능	하	상
장점	구현이 용이함	고차원, 대용량 데이터에 적합
단점	고연산 비용 및 맥락을 잘 이해 X	높은 초기 구축 비용

개선 이해도 개선

→ 질문의 맥락과 Context의 맥락을 이전보다 잘 파악

인덱싱 속도 개선

→ 질문과 유사한 → Context를 찾는데 소모시간이 감소함

문제 초기 구축 비용

선별적인 벡터화점진적으로 데이터증진

충분한 데이터 확보

▶ 데이터 증강 및 적절한 외부데이터 발굴

자체평가의견

훈련생	역할	평가의견
배지원	팀장/모델링	LLM을 파인튜닝하는 다양한 방법론을 찾아봤다는 점에서 의의가 깊다. 또한 학습시키 시키는 과정에서 컴퓨팅 자원을 상당히 소모했는데, 그러다 보니 실무에 간다면 LLM LLM학습과 관련된 비용 계획을 구체적으로 세워둬야겠다는 것을 느꼈다.
신유진	팀원/모델링	생성형 중 QA AI를 다루는 경험은 정보나 코드를 찾는 과정에서 흔치 않은 경험이라는 라는 것을 깨닫게 돼, 좀 더 흥미를 갖고 프로젝트에 임할 수 있었다.
이신우	팀원/모델링	BERT, LLM 모델 계열의 전반적 모델링 프로세스를 경험하며 보람을 느꼈다. 이번엔 딥 딥러닝 구축에 완벽히 기여하지 못해 아쉬웠지만 다음에는 주체적으로 도전하고자 한다.
김혜진	팀원/프론트	이번 프로젝트에는 프론트 엔드를 담당했지만, 다음에는 백엔드까지 해서 풀스택 개발을 해보고 싶다.
우미림	팀원/백엔드	제대로 구현하지 못한 기능들이 많아 아쉽다(가계부 수정, 휴대폰번호 인증, 회원탈퇴 팝업), 생각보다 매끄럽고 예쁘게 구현하기 어렵다
정소이	팀원/백엔드	간단한 기능처럼 보여도 개발자들의 노고가 들어가 있다는 것을 이번 개발을 통해 깨달았다.

참조문헌

- [청소년금융문맹위기] 전금융위원장이바라본'청소년금융문맹'해법은? http://www.fq.or.kr/board_info/517314
- 10년전보다점수더떨어졌다...청소년금융이해력형편없다는데 http://www.fq.or.kr/board_info/515985
- 금융, 법률 문서 기계독해 데이터

https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=71610

- 대학생을 위한 실용 금융 https://terms.naver.com/list.naver?cid=58438&categoryld=58438
- 김민정,고은희,서가연,and 김은미. "청소년 금융교육 효과평가 모형 검증 씽크머니 프로그램 참여 학생을 대상으로." 소비자학연구 34.1 (2023): 17-39.

Woori FISA

AI엔지니어링 JEOLMI'S FAMILY