

캡스톤디자인I 중간보고서

제 목	국문	시각 기반 질의응답에서 지식 베이스의 활용		
	영문	Knowledge Base for Visual Question Answering		
진 행 상 황	제출요망일시	<주요 기능 : 모델 학습> 1. NIA에서 제작한 지식기반 시각질의응답 데이터셋 분석 2. 일반적인 시각질의응답 모델 생성 3. 지식 그래프 임베딩으로 학습하는 방안 생성 4. 지식을 활용하는 시각질의응답 모델 생성 5. UI 구축 6. KCC 2023 논문 제출		
	진행상황	기존의 시각질의응답 데이터셋 조사 완료, 시각질의응답 관련 논문 조사 완료, NIA의 지식 기반 시각질의응답 데이터셋 분석 완료, 일반적인 시각질의응답 베이스라인 모델 생성 완료, 지식 그래프를 임베딩 형태로 변환하는 방안 구축, 지식 그래프 임베딩을 활용한 시각질의응답 모델 생성 완료, 두 모델을 학습하여 정확도 측정 및 비교 완료, 진행한 실험들을 바탕으로 KCC 2023 논문 제출 완료. 요구사항 분석과 사용자 편의성에 집중하여 챗봇의 기본 구조를 설계완료		
산출물	요구사항 정의서(별첨 1), 중간보고서(별첨 2)			
팀 구성원	학년	학 번	이 름	연락처(전화번호/이메일)
	4	20181796	김민준	010-8484-3956 / 20181796@edu.hanbat.ac.kr
	4	20202364	송지현	010-6889-3887 / 20202364@edu.hanbat.ac.kr
	4	20181620	송승우	010-9164-6572 / 20181620@edu.hanbat.ac.kr
컴퓨터공학과와 프로젝트 관리규정에 따라 다음과 같이 요구사항 정의서와 중간보고서를 제출합니다 <div style="text-align: center;">2023년 4월 28일</div> <div style="text-align: right;"> 책임자 : 김민준 (인) 지도교수 : 장한얼 (인) </div>				

김민준
장한얼인

[별첨1]

프로젝트명 : 시각 기반 질의응답에서 지식 베이스의 활용

소프트웨어 요구사항 정의서

Version 1.0

개발 팀원 명(팀리더):김민준
송지현
송승우

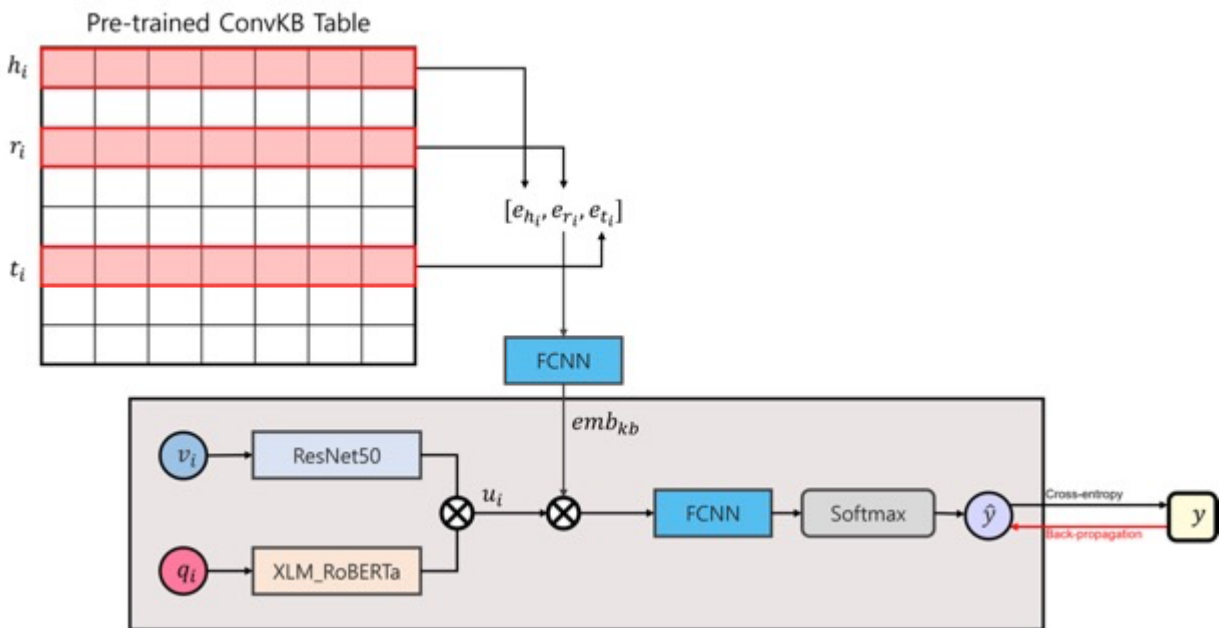
대표 연락처:010-8484-3956
e-mail: mjkmain20@gmail.com

목차

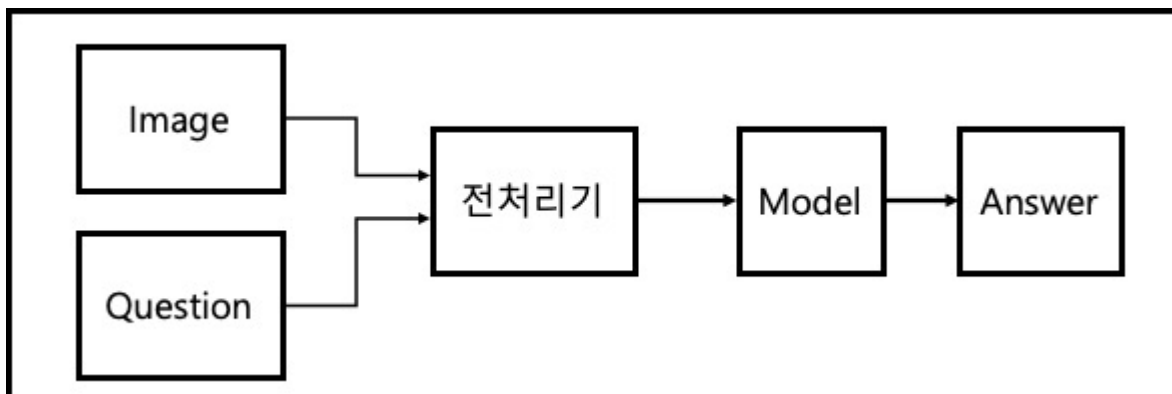
1. 개요
2. 시스템 장비 구성요구사항
3. 기능 요구사항
4. 성능 요구사항
5. 인터페이스 요구사항
6. 데이터 요구사항
7. 테스트 요구사항
8. 보안 요구사항
9. 품질 요구사항
10. 제약 사항
11. 프로젝트 관리 요구사항

1. 시스템 개요

- 외부 지식을 활용하는 시각 기반 질의응답 시스템 베이스라인 모델 구조



- 시각 기반 질의응답 시스템 학습 파이프라인



2. 시스템 장비 구성요구사항

요구사항 고유번호		ECR-001		
요구사항 명칭		장비 요구사항		
요구사항 분류		시스템 장비구성 요구사항	응락수준	필 수
요구사항 상세 설명	정의	모델 학습 장비		
	세부 내용	<ul style="list-style-type: none">- 장비 품목 : GPU (NVIDIA A100 80GB)- 장비 수량 : 4개- 장비 기능 : 모델의 연산 속도를 높인다.- 장비 성능 및 특징 : 총 메모리 80GB- 시간 제약사항 : 해당 사항 없음- 자원 제약사항 : 해당 사항 없음- 장애 처리 : 해당 사항 없음		

3. 기능 요구사항

요구사항 고유번호		SFR-001
요구사항 명칭		AI 모델 개발
요구사항 분류		기능
요구사항 상세 설명	정의	지식기반 시각질의 응답 모델 개발
	세부 내용	<ul style="list-style-type: none">- CNN 및 BERT 기반의 멀티 모델 개발- 지식그래프를 임베딩 형태로 변환하는 방안 구축- 이미지 외의 외부지식이 필요로하는 질의에 답변할 수 있도록 할 것

4. 성능 요구사항

요구사항 고유번호		PER-001		
요구사항 명칭		처리 속도 및 시간		
요구사항 분류		성능 요구사항	응락수준	필 수
요 구 사 항 상 세 설 명	정의	처리 속도 및 시간		
	세부 내용	이미지 및 자연어를 입력으로 받아 모델이 답변을 추론하는 시간을 의미함		

5. 인터페이스 요구사항

요구사항 고유번호		SIR-001		
요구사항 명칭		대화창		
요구사항 분류		사용자 인터페이스	응락수준	필 수
요 구 사 항 상 세 설 명	정의	대화창 구현		
	세부 내용	<ul style="list-style-type: none"> - 대화창에는 사용자와 시각 질의 응답 시스템 간의 대화가 표시될 수 있도록 인터페이스 구현 - 사용자의 입력 창과 시각 질의 응답 시스템의 출력 창을 구분하여 인터페이스에 표시 		

6. 데이터 요구사항

요구사항 고유번호	DAR-001		
요구사항 명칭	전처리된 데이터		
요구사항 분류	데이터	응락수준	필 수
요 구 사 항 상 세 설 명	- NIA에서 만든 지식기반 시각질의응답 데이터셋에서 지식그 래프를 가지고 있는 데이터 수집 후 csv파일로 저장		

7. 테스트 요구사항

요구사항 고유번호	TER-001		
요구사항 명칭	성능 테스트		
요구사항 분류	테스트	응락수준	필 수
요 구 사 항 상 세 설 명	- 일반적인 시각질의응답 모델과 지식기반 시각질의응답 모델 의 성능을 비교하여 평가		

8. 보안 요구사항

요구사항 고유번호	SER-001
요구사항 명칭	보안지침 준수
요구사항 분류	보안
요구사항 상세 설명	<ul style="list-style-type: none"> - 사용한 데이터 세트를 만든 NIA의 보안 및 저작권 관련 데이터 이용정책에 따라 개발이 수행되어야

9. 품질 요구사항

요구사항 고유번호	QUR - 001	
요구사항 명칭	데이터 품질관리	
요구사항 분류	품질	
요구사항 상세 설명	정의	품질관리(기술 관점)
	세부 내용	<ul style="list-style-type: none"> - 질의 유형 별 데이터 수량은 균일하고 적절한 비율을 유지해야 한다. - 학습 데이터는 중복이 없어야 한다. - 학습 데이터 및 검증, 테스트 데이터는 학습에 방해되지 않도록 이미지, 질의에 대하여 결측치가 없는지 검토해야한다. - 데이터에는 질의, 이미지, 지식 그래프가 포함되어야 한다.

10. 제약 사항

요구사항 고유번호	COR-001
요구사항 명칭	시스템 개발과 설계 및 구현 제약사항
요구사항 분류	제약사항
요구사항 세부내용	<ul style="list-style-type: none">• 현재 보유하여 활용 가능한 H/W, S/W를 최대한 활용함• 대부분의 인공지능 모델 개발에 사용되는 Python(언어), PyTorch (프레임워크)를 사용함

11. 프로젝트 관리 요구사항

요구사항 고유번호		PMR-001
요구사항 명칭		프로젝트 관리
요구사항 분류		프로젝트 관리
요구사항 상세 설명	세부 내용	<ul style="list-style-type: none"> - 세부 작업 분할구조 : <ol style="list-style-type: none"> 1. 분석 <ul style="list-style-type: none"> - 지식 그래프를 사용하기 위한 KGE기법 조사 2. 데이터 수집 및 전처리 <ul style="list-style-type: none"> - KVQA에서 사용할 데이터셋 전처리 3. 시스템 설계 <ul style="list-style-type: none"> - 지식그래프를 이용하기 위한 KGE 구현, CNN-BERT 모델 기반의 시각질의 응답 시스템 개발 4. 실험 <ul style="list-style-type: none"> - 일반적인 시각질의응답 데이터와 외부지식이 필요로 하는 시각질의 응답 데이터를 테스트데이터로 함. 일반적인 시각질의응답 모델 VQA과 지식기반 시각질의응답 모델 KVQA을 테스트하여 두 모델에 대한 정확도를 측정 - 프로젝트 수행조직에 대한 구성, 역할 김민준: 논문 및 자료조사, 데이터셋 전처리, 모델 코드 작성 송승우 : 논문 및 자료조사, 데이터셋 전처리, 모델 코드 작성 송지현 : 논문 및 자료조사, UI 개발

[별첨2]

중간보고서

1. 요구사항 정의서에 명시된 기능에 대하여 현재까지 분석, 설계, 구현(소스코드 작성) 및 테스트한 내용을 기술하시오.

- 분석 : 데이터 탐색적 분석
 - 이미지 당 질문 수 통계
 - 질문 길이 (Question length) 통계
 - 답변 길이 (Answer length) 통계
 - 질문 타입 (Question type) 통계
 - 외부 지식 통계
- 설계 : KBVQA 모델 설계
 - 외부 지식을 예측하며 학습하는 KBVQA 모델 설계
 - Contrastive Learning을 적용하는 모델 설계
- 구현 : KBVQA 모델 구현
 - Baseline 모델 : 모든 외부지식을 알고 있다고 가정한 후, ground-truth 외부지식을 모델의 입력으로 활용하는 KBVQA Baseline 모델 구현 완료
 - Triple prediction을 통해 이미지, 질의 쌍에 관련된 외부지식을 예측하는 모델을 활용하여 예측한 Triple을 기반으로 학습을 진행하는 KBVQA 모델 구현 완료

2. 프로젝트 수행을 위해 적용된 추진전략, 수행 방법의 결과를 작성하고, 만일 적용과정에서 문제점이 도출되었다면 그 문제를 분석하고 해결방안을 기술하시오.

- 추진 전략 : 매주 담당 교수님에게 프로젝트 진행 사항을 발표하며, 교수님의 피드백과 팀원들과 토의함으로써 수월하게 진행함
- 수행 방법 : 각자 장비 환경에 맞추어 효율적으로 업무를 분담하고, 또한 '구글의 공유 플랫폼'을 적극적으로 사용하여 효율성을 높일 것
- 추진 전략 및 수행 방법의 결과 : 업무 효율이 높았으며, 역할 분담을 함으로써 모든 팀원의 참여도가 높았음
- 팀원의 책임 및 역할 수행에 대한 결과 : 업무를 균등하게 분담함으로써 원활하게 프로젝트 수행하여 큰 문제는 없었음
- 프로젝트 일정계획에 맞추지 못한 경우의 문제점과 해결 방안 : 모델 실험에 관련해서 환경 설정 문제에 봉착해 일정계획에 차질이 생겼지만, 원래 계획에 맞추기 위해 교수님과 추가 미팅함으로써 문제해결 방안을 알게 됐고 모든 팀원이 실험을 다시 진행하여 다음 일정계획에 차질이 없도록 함
- 요구사항 변경관리 : 변경 사항 없음

캡스톤 디자인 | 중간보고서 채점표

평가도구	평 가 항 목	평 가 점 수				
		1	2	3	4	5
중간 보고서 및 실행 결과	1. 요구사항 정의서(기능, 성능, 인터페이스 등)가 구체적으로 작성되었는가?					
	2. 요구분석, 설계 산출물(모델, 프로토타입 등)의 내용이 충실한가?					
	3. 설계 및 구현 문제를 위해 적용한 이론, 문제해결 방법이 제시되었으며 그 적용이 적합한가?					
	4. 구현된 소프트웨어(또는 이와 동등한 하드웨어 시스템)가 버그 없이 실행되었는가?					
	5. 구현된 소프트웨어(또는 이와 동등한 하드웨어 시스템)의 성능 요구사항은 충족되었는가?					
도구활용	6. 설계 및 구현을 위해 도구가 적절히 활용되었는가?					
	7. 도구의 활용수준(능숙도)은 프로젝트 수행에 적합한가?					
팀원의 업무 및 역할	8. 팀원의 업무분담에 따른 역할 및 협력이 충실히 이루어졌는가? (평가자에 의한 질의)					
	9. 프로젝트 중간 진척상황에 대해 팀원이 충분히 인지하고 있는가?(평가자에 의한 질의)					
합계						
*검토 의견(최종완료 때까지 보완해야할 점에 대해 작성 요망) <div style="height: 150px; border: 1px solid black; margin-top: 5px;"></div>						
심사위원(소속):		(이름)			(인)	