

프로젝트 계획서

프로젝트 제목	챗봇	프로젝트 코드	K-Chatbot														
실행 조직	T3Q	실행 PM	석은수														
프로젝트 개요	<div><div>1. 프로젝트 명<ul style="list-style-type: none">- 본 프로젝트의 명칭은 'K-Chatbot' 이라 한다.</div><div>2. 프로젝트 기간<ul style="list-style-type: none">- 전체 프로젝트 기간: 2023.05.01. ~ 06.16 (총 90일간)</div><div>3. 프로젝트 배경<ul style="list-style-type: none">- 1인 가구, 개인주의의 증가로 인한 우울증, 외로움 등이 사회문제로 대두- 챗봇을 활용한 고객 응대 사이트 증가- 기업에서 고객센터, 교육 등 다양한 분야에서 사용 가능</div><div>4. 프로젝트 목적<ul style="list-style-type: none">- 가벼운 대화나 위로, 조언 등 감정적 교류- 한글로 서비스하기 적절한 QA 모델 구축</div><div>5. 프로젝트 기대효과<ul style="list-style-type: none">- 사람의 우울감 완화에 도움</div></div>																
프로젝트 범위	<div><div>1. 공감 및 대화형 챗봇 제공</div><div>2. 사람들에게 신선한 경험 제공</div></div>																
프로젝트 설계	<div><div>1. 데이터 수집</div><div>1.1. 공공데이터 사용</div><div>1.1.1. 데이터 출처: AI hub(https://aihub.or.kr/) 주제별 텍스트 일상 대화 데이터 인스타그램, 카카오톡, 페이스북, 밴드, 네이트온 등</div><div>1.2. 데이터 상세</div><div>1.2.1. 데이터 유형: 텍스트</div><div>1.2.2. 데이터 형식: txt</div><div>1.2.3. 데이터 구축년도: 2021년</div><div>1.2.4. 데이터 목록</div><table><tr><th>주제 영역</th><th>건수</th></tr><tr><td>KAKAO</td><td>89,805</td></tr><tr><td>FACEBOOK</td><td>9,908</td></tr><tr><td>INSTAGRAM</td><td>5,906</td></tr><tr><td>BAND</td><td>1,997</td></tr><tr><td>NATEON</td><td>1,998</td></tr><tr><td>총 계</td><td>109,614</td></tr></table></div>			주제 영역	건수	KAKAO	89,805	FACEBOOK	9,908	INSTAGRAM	5,906	BAND	1,997	NATEON	1,998	총 계	109,614
주제 영역	건수																
KAKAO	89,805																
FACEBOOK	9,908																
INSTAGRAM	5,906																
BAND	1,997																
NATEON	1,998																
총 계	109,614																

	<p>2. 데이터 전처리</p> <p>2.1. 개발툴: python 3.9</p> <p>2.2. 프레임워크: pandas, glob</p> <p>2.3. 구축 방안</p> <p>2.3.1. 질문과 답변으로 데이터 재구성</p> <p>2.3.2. 불용어 제거</p> <p>3. ChatBot 모델 구축</p> <p>3.1. 개발툴: python 3.9</p> <p>3.2. 프레임워크: pytorch, sklearn, transformer</p> <p>3.3. 모델: KoAlpaca</p> <p>3.4. 기능: 사용자가 문의하는 것에 대해 응답</p> <p>4. 협업</p> <p>4.1. 협업 툴: git, Notion</p>								
일정	<p>- 전체 일정: 첨부파일(간트 차트.xlsx) 참조</p> <p>- 세부 일정: 첨부파일(WBS.xlsx) 참조</p>								
역할	<table> <tr> <th>이름</th><th>역할</th></tr> <tr> <td>석은수</td><td> <ul style="list-style-type: none"> - PM - 회의록 작성 - 데이터 수집 및 분석 </td></tr> <tr> <td>손지현</td><td> <ul style="list-style-type: none"> - 계획서 작성 - 데이터 수집 및 분석 - 모델 개발 및 테스트 </td></tr> <tr> <td>이소정</td><td> <ul style="list-style-type: none"> - 일정 설계 - 모델 개발 및 테스트 - 발표 자료 제작 </td></tr> </table>	이름	역할	석은수	<ul style="list-style-type: none"> - PM - 회의록 작성 - 데이터 수집 및 분석 	손지현	<ul style="list-style-type: none"> - 계획서 작성 - 데이터 수집 및 분석 - 모델 개발 및 테스트 	이소정	<ul style="list-style-type: none"> - 일정 설계 - 모델 개발 및 테스트 - 발표 자료 제작
이름	역할								
석은수	<ul style="list-style-type: none"> - PM - 회의록 작성 - 데이터 수집 및 분석 								
손지현	<ul style="list-style-type: none"> - 계획서 작성 - 데이터 수집 및 분석 - 모델 개발 및 테스트 								
이소정	<ul style="list-style-type: none"> - 일정 설계 - 모델 개발 및 테스트 - 발표 자료 제작 								
제안 및 요구사항	<p>1. 자연어 처리를 원활하게 할 수 있는 워크스테이션 필요</p> <p>2. 자연어 처리 전문가 2명 이상 필요</p>								

첨부파일1. 간트 차트.xlsx

첨부파일2. WBS.xlsx