

## SK네트웍스 Family AI과정 3기

# 데이터 수집 및 저장 프로젝트 기획서

- 산출물 단계 : 데이터 수집 및 저장
- 평가 산출물 : 프로젝트 기획서
- 제출 일자 : 2024.11.18
- 깃허브 경로 : <https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN03-FINAL-6Team>
- 작성 팀원 : 최연규, 김병수, 박지용

# 1. 프로젝트 개요

## 1.1 프로젝트 명

본 프로젝트의 명칭은 “**Tailor Link**”로, 고객사의 요구에 따라 챗봇 기능을 선택하고 맞춤형으로 구성할 수 있다는 의미를 담고 있습니다.

## 1.2 프로젝트 목표

본 프로젝트의 주요 목표는 기업 맞춤형 모듈형 **AI** 챗봇 상담 서비스를 개발하는 것입니다. 이를 위해 다음과 같은 세부 목표를 설정하였습니다:

- 기능 모듈화: 각 챗봇 기능을 독립적으로 모듈화하여 필요에 따라 선택 및 조정 가능하도록 설계합니다.
- 데이터 변경 용이성: 다양한 도입 분야에 맞춰 데이터 변경 및 관리가 간편하도록 시스템을 구현합니다.
- 맞춤형 제작: 고객사의 비즈니스 요구와 목표에 따라 챗봇 기능을 조합하여 최적화된 서비스를 제공합니다.

### 1.2.1 프로젝트 분류

대분류 : 고객 상담

중분류 : 자동차 산업

소분류 : 보험 / 정비 / 차량 관리

## 2. 프로젝트 제안 배경(문제 정의)

자동차 산업은 고객 서비스와 차량 유지보수를 통해 고객 충성도를 유지하고, 신차 판매 및 부가 서비스 제공을 통해 수익을 창출해야 하는 경쟁이 치열한 분야입니다. 그러나 현재의 고객 서비스 시스템은 다음과 같은 문제점들을 안고 있습니다.

### 2.1 고객 서비스 효율성 부족

- 고객센터는 반복적인 FAQ 문의로 인해 업무 부담이 과중되며, 이로 인해 고객 대기 시간이 길어지는 문제가 발생하고 있습니다.
- 차량 상태 모니터링이나 유지보수 관련 알림이 적시에 제공되지 않아 고객 경험이 저하되며, 이는 브랜드 이미지에 부정적인 영향을 미치고 있습니다.

### 2.2 고객 맞춤형 서비스 부재

- 고객의 차량 상태와 정비 이력을 기반으로 한 최적화된 신차 제안이나 유지보수 추천이 부족하여, 고객의 니즈를 충족시키지 못하고 있습니다.
- 고객의 위치와 날씨를 반영한 차량 관리 팁 제공이 이루어지지 않아 계절적 특성과 지역적 요구를 반영하지 못하고 있습니다.

### 2.5 수익 창출 기회 상실

- 차량 상태에 따른 신차 대체 제안이나 업셀링이 이루어지지 않아 판매 촉진 기회를 놓치고 있습니다.
- 정비 후 추가적인 부가 서비스 제안이 이루어지지 않아 추가 수익 창출 가능성을 충분히 활용하지 못하고 있습니다.

### 2.6 브랜드 이미지와 고객 충성도 약화

- 고객 맞춤형 차량 관리 및 유지보수 알림 부족으로 인해 고객의 차량 관리 경험이 저하되고, 이는 브랜드 이미지와 고객 충성도에 부정적인 영향을 미칩니다.
- 고객 지원 과정에서의 비효율성은 브랜드 신뢰도를 약화시키는 요인이 될 수 있습니다.

## 3. 해결 방안

차량 유지보수 및 서비스 과정에서 발생하는 다양한 불편함과 비효율성을 해결하기 위해 초거대 언어 모델(**LLM**) 기반의 에이전트 서비스와 맞춤형 상담 챗봇을 도입하는 방안을 제안합니다. 이를 통해 고객, 제조사, 딜러사 모두의 업무 효율성을 향상시키고, 사용자 경험을 대폭 개선할 수 있습니다.

### 3.1 상담 챗봇 활용 방안

#### 3.1.1 고객 측면

##### 1. 24/7 상담 서비스 제공

고객이 차량 상태, 정비 일정, 신차 옵션 등에 대해 언제든지 문의할 수 있도록 24시간 상담 서비스를 제공합니다.

##### 2. 개인화된 상담 제공

고객의 정비 이력 및 차량 데이터를 바탕으로 맞춤형 정보를 제공하여 서비스 신뢰성을 높입니다.

### 3.3 통합 솔루션 도입 기대 효과

1. 고객은 정비 및 유지보수 차량에 대한 문의의 간소화와 실시간 상담 서비스를 통해 더 큰 편리함을 느낄 수 있습니다.
2. 제조사와 딜러사는 데이터 기반 업무 자동화를 통해 운영 효율성을 높이고, 고객 만족도를 개선하며 추가 매출 기회를 확보할 수 있습니다.

## 4. 현상 및 시장조사

### 4.1 현재 시장

자동차 산업에서는 유지보수, 정비, 데이터 활용 및 고객 지원을 포함한 다양한 분야에서 초거대 언어 모델(LLM) 기반 솔루션의 도입이 가속화되고 있습니다. 이러한 솔루션은 고객 경험을 혁신하고 운영 효율성을 높이며, 새로운 매출 기회를 창출하는 데 주목받고 있습니다.

#### 4.1.1 시장 동향

##### 1. 고객 서비스 혁신

자동차 제조업체 및 딜러사는 LLM 기반 에이전트를 활용하여 고객 맞춤형 서비스를 제공하고 있습니다.

##### 2. 비용 절감 및 효율성 증대

반복적인 문의 및 업무를 자동화하여 인건비를 절감하고, 담당자는 고부가가치 업무에 집중할 수 있습니다.

- 사례: 정비 이력 관리 및 알림 자동화를 통해 정비 예약률을 상승시킵니다.

### 4.2 시장 기회

LLM 기반 솔루션의 도입은 자동차 산업 내 다양한 비효율성을 해소하고 새로운 성장 기회를 제공합니다.

##### 1. 고객 경험 개선

고객 맞춤형 서비스 제공으로 차량 관리의 편의성을 높이고, 고객 충성도를 강화할 수 있습니다.

##### 2. 업무 효율성 증대

정비 예약 및 부품 관리의 자동화로 제조사, 딜러사, 정비소의 운영 효율성을 높일 수 있습니다.

### 3. 기업 경쟁력 강화

LLM 솔루션을 통해 브랜드 충성도를 강화하고, 고객 관리 혁신을 통해 시장 점유율을 높일 수 있습니다.

## 5. 수집 데이터 및 시스템 구성

### 5.1 수집할 데이터 항목

LLM 기반 고객 지원 시스템의 성능을 극대화하기 위해 다음과 같은 데이터를 수집 및 관리합니다.

#### 5.1.1 고객 문의 및 상담 데이터

##### 1. FAQ(자주 묻는 질문) 목록

- 차량 구매, 유지보수, 사고 처리 등 고객이 자주 문의하는 질문과 답변을 정리합니다.
- 예: "보증 기간은 얼마나 되나요?", "정비소 예약 방법은?"

##### 2. 상담 시 주요 문의 주제 및 빈도

- 문의 유형(정비, 보험, 차량 관리 등)과 각 주제별 빈도를 수집합니다.
- 활용: 고객 지원 프로세스 개선 및 FAQ 업데이트.

##### 3. 고객 불만 사항 및 해결 사례

- 불만 접수 및 해결 과정을 기록하여 유사 사례 발생 시 신속히 대처합니다.
- 활용: 고객 불만 패턴 분석 및 서비스 개선.

##### 4. 상담 채널별 문의 유형

- 전화, 이메일, 채팅 등 채널별 주요 문의 내용을 파악하여 채널 특성에 맞는 서비스 제공.
- 활용: 상담 채널 효율성 분석 및 최적화.

### 5.1.2 차량 정보 데이터

#### 1. 차량 상세 스펙 및 옵션 정보

- 모델별 차량 스펙, 옵션, 가격 등 정보를 수집하여 고객 맞춤형 제안을 지원합니다.
- 활용: 신차 추천 및 유지보수 관련 정보 제공.

## 5.2 시스템 구성

LLM 기반 에이전트 및 챗봇 시스템은 다음과 같은 구성 요소로 설계됩니다:

### 5.2.1 데이터 수집 및 통합

- 데이터 소스:
  - 고객 문의 및 상담 데이터: CRM, 상담 로그
  - 차량 정보 데이터: 제조사 API, 차량 상태 데이터
  - 정비 및 서비스 데이터: 정비소 관리 시스템, 고객 피드백 데이터
- 데이터 통합 플랫폼:
  - 다양한 데이터 소스를 통합하여 LLM이 학습할 데이터셋을 구성.
  - 예: 데이터 레이크 또는 클라우드 기반 데이터베이스.

### 5.2.2 시스템 배포 및 운영

- 클라우드 기반 운영:
  - 고객 데이터와 차량 정보를 안전하게 관리하고, LLM 시스템을 클라우드 환경에서 운영.
- 데이터 분석 및 업데이트:
  - 고객 피드백과 차량 데이터를 지속적으로 분석하여 모델 업데이트 및 서비스 개선.

## 5.3 기대 효과

- 고객 만족도 증대: 맞춤형 정보를 제공하여 고객 경험 향상.
- 업무 효율성 향상: 정비 예약 및 유지보수 자동화를 통해 운영 효율 개선.
- 신규 매출 창출: 고객 맞춤형 서비스 및 신차 제안으로 수익 확대.

## 6. 모델링 방안

### 6.1 모델링 목표

- 고객, 보험사, 정비소 간의 원활한 소통과 업무 효율성을 위해 LLM(초거대 언어 모델) 기반의 챗봇 시스템을 구축.
- 데이터 기반으로 고객 요청을 분석하고, 빠르고 정확한 답변을 제공.
- 정비소 예약, 보험 상담 등 각 기능을 모듈화하여 시스템의 유연성을 높임.

### 6.2 세부 모델링 방안

#### 6.2.1 Retrieval-Augmented Generation (RAG)

- 구성: LLM과 벡터 검색 기반의 정보 검색 시스템 결합.
- 작동 방식:
  - 사용자의 질문을 벡터화하여 관련 데이터를 빠르게 검색.
  - 검색된 데이터를 바탕으로 LLM이 자연스러운 응답 생성.
- 활용 사례:
  - 고객이 사고 처리 관련 질문을 하면, 챗봇이 가장 관련성이 높은 FAQ나 고객 데이터에서 정보를 추출.
  - 보험 청구 절차에 대한 상세 정보를 제공.

#### 6.2.2 LangChain 기반 대화 흐름 관리

- 구성: LangChain을 활용하여 대화 흐름을 설계하고, 복잡한 고객 요청을 단계적으로 처리.
- 특징:
  - 다중 단계 대화를 구조화하여, 고객의 요구를 파악하고 추가 데이터를 요청.



### 6.2.3 LLM Fine-Tuning

- 목적: 자동차 산업과 관련된 도메인 지식 반영.
- 데이터:
  - 고객 문의 로그.
- 방법:
  - Hugging Face의 LLM(예: GPT-4 기반 모델)을 Fine-Tuning.
  - RAG와 연동해 데이터 기반 응답 정확성 향상.

### 6.2.4 추천 시스템

- 구성: 고객 데이터와 사고 이력 기반으로 맞춤형 추천 제공.
- 활용 사례:
  - 고객의 차량 모델과 사고 이력을 기반으로 정비소 추천.
  - 보험 갱신 시기 알림 및 추천.

### 6.2.5 데이터 파이프라인

- ETL 과정:
  - Extract: 고객 문의, 보험 정보, 정비소 데이터를 수집.
  - Transform: 데이터 전처리 및 벡터화.
  - Load: 분석 가능한 데이터베이스(MySQL)에 저장.
- 파이프라인 도구:
  - Python(Pandas, NumPy)와 Apache Airflow.

## 7. 사용 데이터

### 7.1 고객 문의 데이터

- FAQ 데이터: 사고 처리, 보험 청구, 정비소 예약 관련 자주 묻는 질문.
- 대화 로그: 고객의 과거 문의와 응답 기록.
- 만족도 데이터: 고객 서비스 피드백 및 평가.

### 7.2 차량 정보 데이터

- 차량 모델, 연식, 주행 거리, 사고 이력.
- 차량 유지보수 기록(정비 내역, 주요 수리 기록).
- 차량 부품 재고 상태(정비소별 재고 데이터).

### 7.3 정비소 데이터

- 정비소 목록 및 위치 정보.
- 정비 서비스 이력: 수리 내역, 비용 데이터.
- 정비소별 고객 피드백 및 평가.

### 7.4 외부 데이터

- 날씨 데이터: 사고 발생 패턴 분석에 활용.
- 도로 상황 데이터: 사고 원인 및 처리 시간 예측.
- 유사 사고 사례: 사고 유형 및 처리 방식 참고.

## 8. 기술스택

### 8.1 LLM

- **RAG (Retrieval-Augmented Generation):**
  - 사용자 요청 정보를 빠르게 검색하여 적절한 답변을 생성하는 기술로, 데이터 정확성과 응답 품질 향상
- **LangChain:**
  - 다양한 LLM 애플리케이션을 쉽게 통합할 수 있는 오픈소스 프레임워크로, 다양한 AI 기능을 확장 가능한 모듈로 구성
- **LangGraph:**
  - 언어 모델과의 상호작용을 그래프 형태로 구조화하여 복잡한 대화 흐름을 쉽게 관리
- **LangSmith:**
  - 챗봇의 대화 품질을 지속적으로 개선하기 위한 학습 관리 시스템
- **Python:**
  - 다양한 데이터 분석, 머신러닝 기능을 쉽게 구현할 수 있는 고급 프로그래밍 언어로 AI 모델 구현에 사용
- **Faiss:**
  - 대규모의 벡터 검색을 가능하게 하여 빠른 응답 제공

### 8.2 FrontEnd

- **Vue.js:**
  - 사용자가 직접 상호작용하는 웹 프론트엔드 구현에 사용

### 8.3 BackEnd

- **Java SpringBoot:**
  - 챗봇의 서버 사이드 로직을 구현하기 위한 프레임워크로, 안정성과 확장성이 뛰어난 구조 제공

## 8.4 기타

- **MySQL:**
  - 사용자 데이터와 고객 상호작용 데이터를 효율적으로 저장 및 관리하기 위한 관계형 데이터베이스 시스템
- **Docker:**
  - 애플리케이션 배포 및 실행을 위한 컨테이너 기술로, 다양한 환경에서 일관된 동작 보장
- **GitHub:**
  - 코드 관리와 협업을 위한 버전 관리 시스템
- **AWS:**
  - 클라우드 인프라스트럭처를 제공하여 서버 호스팅, 데이터 저장, 확장성 있는 서비스 구현 지원
- **Discord (디스코드)**
  - 음성, 영상, 텍스트 채팅이 가능한 커뮤니케이션 플랫폼으로, 팀원 간의 실시간 소통과 협업에 유용합니다. 특히, 개발자와 디자이너 간의 원활한 커뮤니케이션을 지원합니다.
- **Figma (피그마)**
  - 웹 기반의 디자인 협업 도구로, UI/UX 디자인 작업에 주로 사용됩니다. 실시간 협력이 가능하여 팀원들과의 원활한 소통을 지원합니다.
- **PowerPoint (파워포인트)**
  - 프레젠테이션 제작 도구로, 프로젝트 결과물이나 보고서 작성 시 활용됩니다. Figma에서 디자인한 내용을 PowerPoint로 내보내는 방법도 있습니다.
- **Notion (노션)**
  - 메모, 문서 작성, 프로젝트 관리 등 다양한 기능을 통합한 올인원 생산성 도구입니다. Figma와 연동하여 디자인 문서를 관리하거나, 팀 협업 시 활용할 수 있습니다.

## 9. R&R

최연규 : **PM / AI** 엔지니어

- 프로젝트 계획 수립 및 일정 관리
- LLM 기능 구현
- 시스템 설계 및 통합

김병수 : **AI** 엔지니어

- 기업 맞춤형 **RAG** 기반 솔루션 설계
- 프롬프트 최적화

박지용 : **AI** 엔지니어

- 데이터 분석을 통해 **LLM** 기반 모델이 처리할 데이터를 준비
- 데이터 수집 및 전처리
- 이미지 관련 **LLM** 기능 구현

박용주 : **BackEnd**

- 백엔드 개발
- API 개발 및 연동

정해린 : **FrontEnd**

- 프론트엔드 개발
- API 개발 및 연동