

빅데이터기반분산-딥러닝혁신프로젝트

AI 서비스 기획서



빅데이터기반분산-딥러닝혁신프로젝트

1분반

소프트웨어학과

32204041, 32202698

정다훈, 오유석

1. 고객불편 탐색

현재 사용자가 일반적으로 사용하는 지도 기반 서비스는 위치 정보를 중심으로 구성되어 있어, 식당에 대한 구체적이고 신뢰할 수 있는 정보를 충분히 제공하지 못하는 문제점이 있다. 특히 리뷰 정보는 이벤트 참여를 목적으로 한 작성 사례가 많아, 실제 방문자의 경험을 반영한 진실성 있는 리뷰를 찾기 어렵다. 또한 사용자는 식당에 대한 정보를 얻기 위해 블로그, 리뷰 앱, SNS 등 여러 플랫폼을 넘나들어야 하며, 이로 인해 정보 탐색 과정이 매우 비효율적이다. 지역에 익숙하지 않은 여행객의 경우, 낯선 환경에서 적합한 식당을 찾는 데 어려움을 겪고, 결과적으로 낮은 만족도를 경험하거나 탐색 비용이 과도하게 소모되는 문제가 발생하고 있다.

2. 서비스 목적 정의

2-1. 고객 니즈 연계성

본 AI 기반 맛집 추천 챗봇 서비스는 사용자가 자연어를 통해 조건을 입력하고, 대화형 인터페이스를 통해 식당을 탐색할 수 있도록 설계되었다. 예를 들어 “서울에서 조용한 분위기의 혼밥 가능한 식당 찾아줘”와 같이 대화하듯 요청하면, 챗봇이 이를 해석하여 적합한 식당을 추천한다. 사용자의 조건에 부합하는 결과를 제공함으로써 검색 피로를 줄이는 것이 핵심이다.

또한 리뷰의 신뢰도를 높이기 위해, 리뷰 진실성 필터링 알고리즘을 도입하여 이벤트성 또는 광고성 리뷰를 자동으로 제거하고, 실제 사용자 기반의 리뷰만 요약 및 분석하여 제공한다. 이 외에도 여러 플랫폼(예: 블로그, SNS, 커뮤니티 등)으로부터 식당 관련 정보를 자동으로 수집하고 정제하여, 사용자가 한 번의 대화 내에서 통합된 탐색 경험을 얻을 수 있도록 구성한다.

이 서비스는 사용자가 원하는 조건에 맞는 식당을 쉽고 빠르게 찾고자 하는 실질적인 니즈와 밀접하게 연결되어 있다. 사용자는 탐색 시간을 최소화하고, 신뢰할 수 있는 정보를 바탕으로 만족도 높은 식사 경험을 추구하는 경향이 강하며, 본 서비스는 이러한 기대를 충족시키기 위한 기능을 중심으로 구성되어 있다.

2-2. 서비스 목표 설정

서비스의 성공적인 구현과 개선을 위해 다음과 같은 구체적 목표를 설정한다.

첫째, 챗봇 이용자가 1회 대화 내에 원하는 조건에 맞는 식당을 추천받고 탐색을 마무리하는 비율을 **90%** 이상으로 설정한다.

둘째, 리뷰 진실성 필터링을 통해 추천 대상 식당의 리뷰 중 이벤트성·광고성 리뷰 비중을 **10%** 이하로 유지한다.

셋째, 사용자 평가를 기반으로 한 만족도 조사 결과에서 평균 **4.5점 이상(5점 만점)**을 달성하는 것을 목표로 한다.

이러한 목표는 정량적 지표로서 서비스 성과를 객관적으로 측정하고, 향후 개선 방향을 도출하는 기준으로 활용될 수 있다.

2-3. 비즈니스 가치 창출

해당 챗봇 서비스는 단순한 탐색 도구를 넘어서, 외식 관련 산업과 연계된 비즈니스 모델로 확장 가능성이 높다.

첫째, 식당 예약 시스템, 할인 쿠폰 제공, 방문 인증 리워드 시스템 등과 연계하여 부가 수익 구조를 도입할 수 있다.

둘째, 외식 업계 광고주와의 제휴를 통해 사용자 맞춤형 식당 노출 또는 광고 기반 수익 창출이 가능하다.

셋째, 사용자 행동 데이터와 챗봇 상호작용 로그는 타겟 마케팅, 상권 분석, 지역 맞춤 전략 수립 등에 유용한 데이터로 활용될 수 있으며, 이는 서비스의 부가적 가치로 연결된다.

또한, 해당 서비스는 다국어 기능 및 외국인 관광객을 위한 인터페이스를 추가함으로써 글로벌 사용자로의 확장 기반을 마련할 수 있으며, 이를 통해 글로벌 AI 추천 서비스로서의 경쟁력을 확보할 수 있다.

3. 기능(인텐트) 구체화

3-1. 핵심 기능 정의

AI 기반 맛집 추천 챗봇 서비스의 목적을 달성하기 위해 필요한 기능들을 인텐트(Intent) 단위로 정의하고, 이들 인텐트의 중요도 및 사용 빈도를 고려하여 우선순위를 설정하였다. 인텐트는 사용자의 요청이나 의도를 기반으로 시스템이 수행해야 할 기능 단위를 의미하며, 각 인텐트는 대화 흐름 내에서 특정 역할을 수행한다. 본 서비스에서의 주요 인텐트와 그 정의는 다음과 같으며, 총 3단계의 우선순위로 나뉜다.

1순위 인텐트 (핵심 기능)

- **맛집 탐색 인텐트:** 사용자가 지역, 분위기, 메뉴 조건 등을 자연어로 입력했을 때, 해당 조건에 맞는 식당을 탐색하여 추천하는 기능이다. 서비스의 중심 기능으로서, 대부분의 사용자 요청이 이 인텐트를 통해 시작된다.
- **리뷰 요약/분석 인텐트:** 특정 식당에 대한 사용자 리뷰를 수집 및 분석하고, 광고성 리뷰를 제거한 뒤 신뢰도 높은 후기만을 요약하여 제공하는 기능이다. 사용자의 판단을 돕기 위한 필수 기능으로, 탐색 인텐트와 밀접하게 연동된다.

2순위 인텐트 (보조적 주요 기능)

- **위치 기반 인텐트:** 현재 위치 혹은 사용자가 지정한 위치를 기반으로 식당 정보를 필터링하여 제공하는 기능이다. 실시간 위치 기반 탐색과 지역 특화 추천에 활용된다.
- **조건 수정 및 재탐색 인텐트:** 기존에 입력한 조건을 수정하거나, 다른 조건으로 재검색을 요청할 때 사용되는 인텐트이다. 탐색 흐름의 유연성을 확보하기 위해 필요하다.

3순위 인텐트 (부가 기능)

- **할인/혜택 안내 인텐트:** 특정 식당의 할인 정보, 쿠폰 제공 여부, 이벤트 참여 가능성 등을 안내하는 기능이다. 사용자에게 부가적인 가치를 제공하는 기능으로, 필수적이지는 않으나 서비스 경쟁력 강화를 위해 포함된다.
- **즐거찾기/공유 인텐트:** 추천받은 식당을 저장하거나 타인과 공유할 수 있도록 지원하는 기능이다. 사용자의 행동 확장을 유도하고, 서비스 재이용률을 높이기 위한 목적을 가진다.

우선순위는 사용자 니즈, 기능 간 연계성, 구현 난이도 등을 종합적으로 고려하여 설정하였으며, 핵심 기능의 안정성과 품질 확보를 우선적으로 추진한 이후, 보조 및 부가 기능을 순차적으로 확대하는 방향으로 개발 일정을 계획한다.

3-2. 인텐트 세분화

챗봇 서비스의 각 인텐트는 사용자 의도를 보다 정밀하게 반영하기 위해 하위 인텐트로 세분화될 수 있다. 인텐트의 세분화는 단순한 기능 단위 구분을 넘어서, 실제 사용자 대화 흐름에서의 자연스러운 분기와 시나리오 설계를 가능하게 하며, 데이터 수집 및 모델 학습 효율성 확보에도 기여한다. 아래는 본 서비스에서 정의한 주요 인텐트별 세분화 인텐트와 그 설명이다.

1) 맛집 탐색 인텐트

- 조건 기반 탐색 인텐트: 사용자가 입력하는 다양한 조건(분위기, 가격대, 혼잡/모임 여부 등)에 기반하여 이를 해석하고 적합한 식당을 탐색하는 기능. 예: “조용한 분위기의 식당 찾아줘.”
- 메뉴 키워드 탐색 인텐트: 사용자가 특정 음식명이나 메뉴를 언급하면, 해당 메뉴가 주력인 식당을 추천하는 기능. 예: “파스타 잘하는 집 추천해줘.”
- 시간대 기반 탐색 인텐트: 조식, 브런치, 야식 등 특정 시간대에 최적화된 식당을 추천하는 기능. 예: “야식 먹을 수 있는 곳 알려줘.”

2) 리뷰 요약/분석 인텐트

- 긍·부정 리뷰 요약 인텐트: 전체 리뷰 중 긍정적인 표현과 부정적인 표현을 분리하여 요약 제공. 예: “이 식당의 장점과 단점 알려줘.”
- 키워드 리뷰 요약 인텐트: 사용자가 특정 키워드(예: 서비스, 분위기, 청결 등)를 지정하면, 해당 키워드와 관련된 리뷰만 요약 제공. 예: “서비스에 대한 리뷰만 보여줘.”
- 리뷰 신뢰도 검토 인텐트: 리뷰 작성자의 패턴이나 작성 양식을 분석하여, 해당 리뷰가 실제 방문자에 의해 작성된 것인지 판단하여 표시하는 기능. 예: “이 리뷰 신뢰할 수 있어?”

3) 위치 기반 인텐트

- 목적지 기준 탐색 인텐트: 사용자가 특정 목적지를 입력하면, 해당 위치 중심으로 주변의 식당을 탐색하여 추천하는 기능. 예: “홍대 근처 맛집 추천해줘.”

4) 조건 수정 및 재탐색 인텐트

- 조건 일부 수정 인텐트: 기존에 입력한 조건 중 일부를 수정하여 다시 탐색하는 기능. 예: “분위기는 그대로, 가격만 더 저렴한 곳으로.”
- 조건 추가 인텐트: 탐색을 유지하면서 새로운 조건을 추가하여 필터링을 세분화하는 기능. 예: “채식 가능한 곳으로 추가해줘.”
- 초기화 후 재탐색 인텐트: 현재까지 입력된 모든 조건을 초기화하고 새롭게 탐색을 시작하는 기능. 예: “처음부터 다시 찾아줘.”

3-3. 기능 상호작용 설계

AI 기반 맛집 추천 챗봇 서비스는 사용자 요청에 따라 다양한 인텐트가 개별적으로 작동하지만, 실제 서비스에서는 이들 인텐트가 유기적으로 연결되어 하나의 대화 흐름을 구성한다. 따라서 기능 간 연계성을 고려한 시나리오 중심의 설계가 필수적이며, 사용자 발화의 순차적 흐름에 따라 인텐트가 자연스럽게 전환될 수 있도록 구조화되어야 한다. 아래는 대표적인 대화 흐름 시나리오를 기반으로 각 기능(인텐트) 간 상호작용의 구조를 설명한다.

시나리오 1: 기본 조건 입력 → 탐색 → 리뷰 확인

이 시나리오는 사용자가 특정 조건(예: 혼밥 가능)에 따라 식당을 탐색한 후, 추천 결과 중 하나의 리뷰를 요청하는 흐름으로 구성된다. 이 과정에서는 먼저 ‘맛집 탐색 인텐트(조건 기반 탐색)’가 작동하고, 이어서 ‘리뷰 요약/분석 인텐트(긍·부정 리뷰 요약)’로 자연스럽게 전환된다. 이는 탐색 결과에 대한 신뢰도 확보 및 판단 보조 기능을 포함하며, 사용자가 탐색 → 이해 → 결정의 단계를 거칠 수 있도록 돕는다.

시나리오 2: 목적지 설정 → 탐색 → 조건 추가 → 재추천

이 시나리오는 사용자가 특정 목적지(지역)를 기준으로 식당을 탐색한 뒤, 조건을 추가하여 필터링을 정제하는 과정을 포함한다. 초기에는 ‘위치 기반 인텐트(목적지 기준 탐색)’가 작동하며, 이후 ‘조건 수정 및 재탐색 인텐트(조건 추가)’로 확장된다. 이 흐름은 사용자의 상황(약속 장소 등)에 따라 식당 선택 기준이 지역 중심으로 설정된 경우에 유효하며, 점진적인 필터링을 통해 탐색 정확도를 높이는 데 기여한다.

시나리오 3: 메뉴 기반 탐색 → 시간대 필터링 → 조건 일부 수정 → 리뷰 신뢰도 확인

이 시나리오는 사용자가 특정 메뉴(예: 라멘)와 시간대(예: 야식)를 기반으로 식당을 찾고, 이후 조건을 세밀하게 수정하거나 리뷰의 신뢰도를 확인하는 고도화된 흐름을 나타낸다. ‘맛집 탐색 인텐트(메뉴 키워드 탐색 + 시간대 기반

탐색)’가 첫 단계에서 작동하며, 사용자가 조건을 조정하면 ‘조건 수정 및 재탐색
인텐트(조건 일부 수정)’가 이어진다. 마지막으로, 리뷰 신뢰도를 판단하고자
하는 요청이 발생하면 ‘리뷰 요약/분석 인텐트(리뷰 신뢰도 검토)’가 작동한다. 이
시나리오는 조건의 정제와 검증 과정을 통해 사용자가 보다 확신을 가지고
결정을 내릴 수 있게 한다.

이처럼 각 인텐트는 단일 기능으로 분리되어 있으면서도, 실제 대화 맥락에서는
다층적이고 연속적인 흐름 속에서 연결된다. 서비스 설계 시 이러한 흐름을
고려하여 시나리오 기반의 인텐트 연결 구조를 마련하는 것이 중요하며, 사용자
경험의 자연스러움과 만족도를 크게 좌우하는 핵심 요소로 작용한다.

3-4. 인텐트 유효성 검증

정의된 인텐트들이 실제 사용자 요구를 적절하게 반영하고 있는지를 확인하기
위해, 향후 시제품 단계에서 인텐트 유효성 검증을 수행할 예정이다. 해당
검증은 각 인텐트 단위의 응답 정확도뿐만 아니라, 시나리오 흐름 내에서의 전환
자연성, 사용자 만족도, 대화 완료율 등을 중심으로 진행된다.

우선적으로는 샘플 사용자 대상의 테스트 및 대화 로그 분석을 통해, 다음과
같은 측면을 중심으로 평가가 이루어질 수 있다.

- 인텐트 분류 정확도: 사용자 발화를 올바른 인텐트로 매칭하는 비율
- 시나리오 완결률: 사용자가 추천을 받고 탐색을 완료하는 비율
- 인텐트 전환 유연성: 대화 흐름 내에서 인텐트 간 전환이 자연스럽게 이뤄지는 정도
- 사용자 만족도: 대화 종료 후 간단한 평가 또는 설문을 통해 수집된 만족도 지표

이와 같은 유효성 검증은 단순히 기술적 구현의 정확도를 확인하는 것을 넘어,
사용자 경험의 품질을 개선하기 위한 중요한 기준으로 작용할 것이며, 초기
테스트 결과를 기반으로 인텐트 구조의 수정, 누락 인텐트의 보완, 중복
인텐트의 통합 등을 추진할 계획이다.

4. 슬롯 구성

4-1. 슬롯 개요

슬롯(Slot)은 AI 기반 챗봇이 사용자의 발화를 이해하고, 적절한 응답을 생성하기 위해 필수적으로 수집해야 하는 정보 항목이다. 각 인텐트 실행 시 필요한 정보를 구조화된 형태로 저장하고, 대화 흐름 속에서 자연스럽게 이를 확보할 수 있도록 슬롯이 설계된다.

본 서비스에서는 인텐트별로 필요한 정보를 파악하여 주요 슬롯을 정의하였으며, 슬롯마다 데이터 유형과 제약 조건을 설정하여 안정적인 정보 수집과 응답 생성을 보장한다. 또한 사용자 경험(UX)을 고려하여, 명시적 질의 없이도 자연어 흐름 중에서 슬롯을 추출할 수 있도록 설계된다.

4-2. 슬롯 정의

슬롯명	데이터 유형	제약 조건	설명
식사 상대	텍스트	10자 이하	혼밥, 친구, 연인, 가족, 회식 등 식사 목적과 분위기를 유추할 수 있는 정보
장소	텍스트	최대 20자	현재 위치 혹은 목적지를 기반으로 추천 범위를 설정하는 정보
식사 날짜	날짜	오늘 이후 날짜만 허용	예약 여부나 시간대 추천을 위해 필요한 정보
식사 시간대	텍스트	‘아침’, ‘점심’, ‘저녁’, ‘야식’ 중 선택	시간대별로 적합한 식당을 분류하고 필터링하기 위한 정보
메뉴 카테고리	텍스트	최대 15자	한식, 중식, 일식, 양식, 디저트 등 음식 종류에 대한 사용자 선호 정보
분위기	텍스트	최대 15자	조용한, 캐주얼한, 고급스러운, 감성적인 등 식당 분위기에 대한 조건
가격대	숫자 또는 범주	0 이상 또는 저가/중가/고가	사용자의 예산을 고려한 추천을 위해 필요 (가격 조건 미설정 시 중간값 우선)
리뷰 관련 키워드	텍스트	최대 10자	청결, 맛, 서비스, 분위기 등 리뷰 요약 요청 시 어떤 관점 중심으로 제공할지를 판단

4-3. 인텐트별 슬롯 매핑

인텐트명	필요 슬롯 목록
------	----------

맛집 탐색 intent	식사 상대, 장소, 식사 날짜, 식사 시간대, 메뉴 카테고리, 분위기, 가격대
리뷰 요약/분석 intent	리뷰 관련 키워드
위치 기반 intent	장소
조건 수정 및 재탐색 intent	(기존 intent에 따라 수정되는 슬롯 일부)
할인/혜택 안내 intent	장소, 식사 날짜, 식사 시간대, 가격대
즐거찾기/공유 intent	장소, 메뉴 카테고리 (필요 시)

4-4. 슬롯 설계 시 고려사항

슬롯은 단순한 데이터 수집을 넘어, 대화 자연스러움과 서비스 응답 품질을 좌우하는 핵심 요소이다. 따라서 다음의 설계 원칙을 적용한다.

- 명시적 질문 없이도 자연어 대화에서 슬롯 정보를 추출할 수 있어야 하며,
- 슬롯 누락 시에는 확인 질문(예: “어디서 식사하실 예정인가요?”)을 통해 보완하도록 설계,
- 제약 조건을 통해 데이터 일관성과 품질을 유지하며,
- 서비스 특성에 따라 슬롯 우선순위를 설정하여, 탐색 우선 → 보조 기능 → 부가 기능 순으로 응답의 핵심 정보를 확보한다.

5. 고객발화(코퍼스) 구축

5-1. 맛집 탐색 intent

1. 여자친구랑 [식사 상대:연인] [장소:성수동]에 가서 식사할 건데, [식사 날짜:4월 3일 목요일] [식사 시간대:저녁]에 갈만한 식당 추천해줘.
2. [식사 상대:혼자] 하기 좋은 [분위기:조용한] [메뉴 카테고리:한식집] 있으면 알려줘.
3. [장소:강남]에서 [가격대:1만 원 이하]로 [식사 시간대:점심] 먹기 좋은 데 있을까?
4. [식사 날짜:주말]에 [식사 상대:부모님]이랑 [분위기:분위기 좋은] 레스토랑 가고 싶은데 추천해줘.
5. [식사 시간대:브런치] 먹을만한 [분위기:감성적인] [메뉴 카테고리:카페] 없어?

5-2. 리뷰 요약/분석 인텐트

1. 첫 번째 식당 리뷰 어땠는지 요약해줘.
2. 이 집 [리뷰 관련 키워드:서비스] 평가는 괜찮아?
3. 후기에 [리뷰 관련 키워드:청결] 문제가 있다는 말이 있는데 실제로 어때?
4. 여긴 [식사 상대:혼자] 하기 어떤지 [리뷰 관련 키워드:분위기] 좀 알려줘.
5. 이 식당 리뷰 [리뷰 관련 키워드:신뢰도] 신뢰할 수 있어?

5-3. 위치 기반 인텐트

1. [장소:신촌] 근처 맛집 추천해줘.
2. [장소:강남역]에서 가까운 저녁 식당 있을까?
3. [장소:건대입구역] 주변에 늦게까지 하는 식당 알려줘.
4. [장소:내 위치] 기준으로 지금 갈 수 있는 집 알려줘.
5. 오늘 [장소:서울역] 근처에서 [식사 시간대:점심] 약속 있어. 근처 괜찮은 곳?

5-4. 조건 수정 및 재탐색 인텐트

1. 방금 추천해준 곳보다 [가격대:더 저렴한] 데로 바꿔줘.
2. [분위기:분위기]는 그대로 두고 [메뉴 카테고리:한식]만 골라줘.
3. 아까 추천한 집 중에 [메뉴 카테고리:채식] 가능한 곳 추가해줘.
4. 조건 다 지우고 처음부터 다시 찾아줘.
5. 지금 조건에서 [분위기:데이트 분위기]만 추가해줘.

5-5. 할인/혜택 안내 인텐트

1. 지금 [할인 여부:할인 중]인 식당만 알려줘.

2. [할인 여부:쿠폰] 받을 수 있는 맛집이 있을까?
3. [할인 여부:첫 방문자 혜택] 있는 데가 어디야?
4. 이번 주 [할인 여부:할인]하는 [메뉴 카테고리:이탈리안 레스토랑] 있어?
5. [할인 여부:포인트 적립] 가능한 곳 알려줘

5-6. 즐겨찾기/공유 인텐트

1. 이 집 [행동:저장]해둘래.
2. 친구한테 이 식당 추천하고 싶어. [행동:공유] 링크 줘.
3. 나중에 또 보고 싶으니까 [행동:즐거찾기] 해줘.
4. 이 집 다음에 또 가고 싶은데 [행동:저장]해줘.
5. 가족 단톡방에 이 식당 정보 [행동:공유]할 수 있어?

6. AI 의도 분류를 위한 핵심 엔티티 정의

6-1. 엔티티(Entity) 식별

서비스 내에서 사용되는 주요 엔티티는 다음과 같다. 각 엔티티는 사용자 발화 내에서 특정 패턴이나 단어군으로 등장하며, 슬롯 추출 또는 인텐트 분류의 기준이 된다.

엔티티명	예시 표현
장소	강남, 성수동, 서울역, 집 근처, 내 위치, 목적지, 역세권 등
식사 시간대	아침, 점심, 저녁, 야식, 브런치, 늦은 밤 등
식사 날짜	오늘, 내일, 주말, 4월 5일, 이번 주 금요일 등
식사 상대	혼자, 친구, 연인, 여자친구, 남자친구, 부모님, 가족, 회사 동료 등
메뉴 카테고리	한식, 중식, 일식, 양식, 디저트, 라멘, 고기, 브런치 등
분위기	조용한, 감성적인, 고급스러운, 캐주얼한, 트렌디한 등
가격대	1만 원 이하, 저렴한 데, 저가, 중간 가격대, 고급스러운 등
리뷰 키워드	맛, 양, 서비스, 청결, 분위기, 친절도, 만족도, 후회 없음 등

할인/혜택 여부	쿠폰, 첫 방문 혜택, 할인 중, 포인트 적립, 이벤트 중, 무료 사이드 제공 등
----------	---

6-2. 속성 (Attribute) 정의

각 엔티티에는 고유한 속성 또는 제약 조건이 설정되며, 이는 발화 해석 시 정제된 결과를 도출하는 데 사용된다. 예를 들어 날짜는 미래 시점으로 한정되며, 가격은 숫자 또는 정성적 표현(저가, 고급 등) 모두 허용된다.

엔티티명	속성 또는 제약 조건
장소	최대 20 자 이내, 텍스트 기반, 행정구역 또는 지명 위주
식사 시간대	정해진 범주(아침, 점심, 저녁, 야식) 내에서만 유효
식사 날짜	오늘 이후의 유효한 날짜만 허용
가격대	0 이상 정수 or 범주형 표현(저가/중가/고가) 가능
메뉴 카테고리	최대 15 자, 사전에 정의된 음식 카테고리 내 일치 여부 확인
리뷰 키워드	서비스, 청결, 맛, 양, 분위기 등 주요 평가지표와 일치해야 함

6-3. 엔티티 간 관계 설정

엔티티 간 관계를 설정하면 단순한 키워드 매칭이 아닌, 의도 흐름 기반의 복합 분류가 가능해진다. 예를 들어 식사 상대가 "연인"일 경우, 추천해야 할 식당의 분위기는 "감성적인" 혹은 "고급스러운" 쪽으로 우선순위가 조정될 수 있다. 이러한 관계 기반의 추천은 개인화된 사용자 경험을 제공하는 데 필수적이다.

관계 유형	예시 설명
식사 상대 ↔ 분위기	연인 → 감성적인/조용한, 가족 → 편안한/넓은 등
식사 상대 ↔ 분위기	디저트 → 저가/중가, 고기류 → 중가/고가 등
장소 ↔ 식사 시간대	야간 시간대 → 역세권 위주, 점심 시간대 → 오피스 밀집 지역 등
식사 시간대 ↔ 메뉴 카테고리	아침 → 브런치/가벼운 음식, 저녁 → 정식/술안주 등
리뷰 키워드 ↔ 사용자 조건 우선순위	‘청결’ 중시하는 사용자 → 리뷰 요약 시 해당 항목 강조

6-4. 동의어 처리 전략

엔티티 정의 시에는 다양한 사용자 표현을 수용하기 위해 동의어 처리 체계가 반드시 필요하다. 예를 들어 ‘연인’이라는 엔티티는 ‘여자친구’, ‘남자친구’, ‘썸타는 사람’, ‘데이트 상대’ 등으로 표현될 수 있다. 동의어 매핑을 통해 여러 표현이 하나의 동일한 의미로 연결될 수 있도록 설계한다.

엔티티명	동의어 예시
식사 상대	연인 → 여자친구, 남자친구, 썸남, 썸녀, 데이트 상대 등
장소	내 위치 → 여기 근처, 지금 있는 곳, 현 위치, 현재 위치 등
할인 여부	할인 → 쿠폰, 프로모션, 이벤트 중, 첫 방문 혜택, 적립 이벤트 등
리뷰 키워드	청결 → 위생, 깔끔함, 더러움 없음, 깨끗함 등

7. 머신러닝 실행 및 성능 평가

AI 기반 챗봇의 의도 분류 및 슬롯 추출 정확도를 향상시키기 위해, 적절한 머신러닝 모델을 선정하고 데이터를 구성한 후 학습 및 성능 평가를 수행한다. 본 프로젝트에서는 과도한 연산 자원 없이도 우수한 성능을 발휘할 수 있는 경량 모델을 선택하되, 향후 지속적인 개선과 보강이 가능하도록 확장성을 고려하였다.

7-1. 모델 선택

의도 분류(Intent Classification)와 엔티티 인식(Named Entity Recognition, NER)은 본 서비스에서 핵심적인 자연어 처리 기능이다. 이에 따라, 분류와 추출 능력에 강점을 가진 LLM 계열 경량 모델 중 LLaMA3 2B를 선택하였다.

LLaMA3 2B는 Meta에서 공개한 최신 경량 언어 모델로, 비교적 적은 자원으로도 뛰어난 추론 성능을 보이며, 사용자 요구에 따라 다양한 태스크로 파인튜닝이 가능한 구조를 갖고 있다. 현재 모델은 영어 기반으로 사전학습되어 있으므로, 본 프로젝트에서는 한국어 코퍼스를 기반으로 파인튜닝을 수행할 계획이다.

7-2. 학습 실행

머신러닝 모델 학습을 위해 앞서 정의된 인텐트 및 엔티티 정보를 기반으로 데이터를 구축하고, 다음과 같은 방식으로 학습을 수행한다.

- 전체 데이터셋을 80% 학습용 / 20% 검증용으로 분할
- 과적합 방지를 위한 기본 전략 적용:

- Early stopping
- Dropout
- Oversampling or class weighting (불균형 클래스 대응)
- 옵티마이저는 Adam, 손실 함수는 CrossEntropy 기준 사용

모델 학습은 주로 의도 분류용 분류기(Classifier)와 토큰 단위 엔티티 인식기(NER Tagger)로 구성된다.

7-3. 성능 평가

훈련된 모델의 성능은 다음과 같은 정량 지표를 사용하여 평가한다.

- 정확도 (**Accuracy**): 전체 예측 중 맞춘 비율
- 정밀도 (**Precision**): 예측한 결과 중 실제로 맞춘 비율
- 재현율 (**Recall**): 실제 정답 중 모델이 맞춘 비율
- **F1 Score**: 정밀도와 재현율의 조화 평균

의도 분류와 엔티티 인식에 대해 각각 개별적으로 위 지표를 측정하며, **Confusion Matrix**를 통해 오분류 패턴을 시각적으로 분석한다.

7-4. 모델 개선 및 데이터 보강

성능 평가 결과를 바탕으로, 정확도가 낮거나 오분류 비율이 높은 인텐트 및 엔티티에 대해 다음과 같은 방식으로 개선을 진행한다.

- 실제 사용자 발화 로그 기반 신규 코퍼스 수집
- 자주 혼동되는 의도 간 샘플 보강
- 동의어, 지역어, 축약어 등 자연 발화를 반영한 도메인 확장
- 지속적인 미세 조정(fine-tuning)을 통해 챗봇 성능 고도화

또한, 추후 대규모 사용자 유입을 고려하여, 필요 시 **LLaMA3** 상위 모델(**LLaMA3 8B** 등)로 업그레이드할 수 있도록 학습 파이프라인을 유연하게 설계한다.