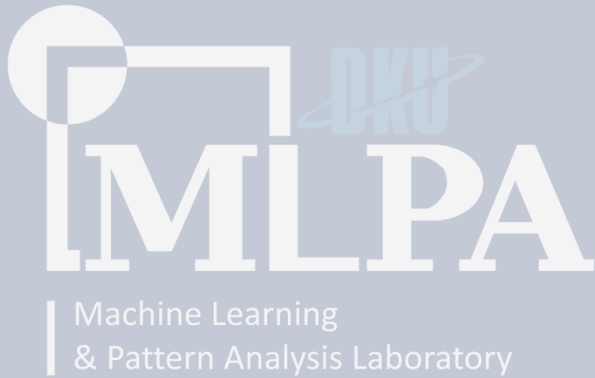


AI 기반 맛집 추천 챗봇 서비스

빅데이터기반분산-딥러닝혁신프로젝트
AI 서비스 기획서 발표

목차

- 1 고객 불편 탐색
- 2 서비스 목적 정의
- 3 인텐트 구체화
- 4 슬롯 구성
- 5 코퍼스 구축
- 6 핵심 엔티티 정의
- 7 성능 평가



Part 1 고객 불편 탐색

1. 고객 불편 탐색

지도 기반 정보의 한계

기존 지도 기반 서비스는 위치 정보를 중심으로 구성되어 있으며, 식당에 대한 구체적이고 신뢰 할 수 있는 정보를 충분히 제공하지 못함
이벤트 참여를 목적으로 하는 리뷰가 많으므로, 실제 방문자의 경험을 반영한 진실성 있는 리뷰를 찾기 어려움

정보 탐색 비효율성

사용자는 블로그, 리뷰 앱, SNS 등 여러 플랫폼을 넘나들어야 함

낮선 지역에서의 선택 어려움

여행객의 경우 낯선 환경에서 적합한 식당을 찾는데 어려움을 느낌
따라서 낮은 만족도를 경험하거나 탐색 비용이 과도하게 소모됨

Part 2 서비스 목적 정의

2.1. 고객 니즈 연계성

1. 챗봇 서비스

사용자가 자연어를 통하여 조건을 입력하고, 대화형 인터페이스를 통하여 식당을 탐색

예시) “서울에서 조용한 분위기의 혼밥 가능한 식당 찾아줘”

2. 리뷰 진실성 필터링 알고리즘

리뷰의 신뢰도를 높이기 위하여, 이벤트성 또는 광고성 리뷰를 자동으로 제거 및 실제 사용자 기반의 리뷰 요약 제공

리뷰 외에도 다양한 플랫폼에서 식당 관련 정보를 자동으로 수집 및 정제하여, 사용자가 한 번의 대화 내에서 통합된 탐색 경험을 얻음



사용자의 탐색 시간 최소화, 신뢰할 수 있는 정보를 바탕으로 만족도 높은 식사 경험 제공

2.2. 서비스 목표 설정

서비스의 성공적인 구현과 개선을 위하여 구체적인 목표를 설정

1. 대화 완결률

챗봇 이용자가 1회 대화 내에서 원하는 조건에 맞는 식당을 추천 받고 탐색을 마무리 하는 비율을 90% 이상으로 설정

2. 리뷰 진실성 확보

추천 대상 식당 리뷰 중 이벤트성 및 광고성 리뷰 비중을 10% 이하로 유지

3. 사용자 만족도

사용자 평가를 기반으로 한 만족도 조사 결과에서 평균 4.5점 이상 달성

2.3. 비즈니스 가치 창출

1. 식당 예약 시스템, 할인 쿠폰 제공, 방문 인증 리워드 시스템 등과 연계하여 부가 수익 구조 도입
2. 외식 업계 광고주와의 제휴를 통하여 사용자 맞춤형 식당 노출 또는 광고 기반 수익 창출
3. **사용자의 행동 데이터와 챗봇 상호작용 로그**의 서비스 부가적 가치로서의 연결
타겟 마케팅, 상권 분석, 지역 맞춤 전략 수립 등에 유용한 데이터로 활용
4. 다국어 기능 및 외국인 관광객을 위한 인터페이스를 추가하여 **글로벌 서비스로서의 경쟁력 확보**

Part 3 인텐트 구체화

3.1. 핵심 기능 정의

1순위 인텐트 (핵심 기능)

- **맛집 탐색 인텐트** : 지역, 분위기, 메뉴 조건 등을 자연어로 입력하면 해당 조건에 맞는 식당을 추천
- **리뷰 요약/분석 인텐트** : 특정 식당의 리뷰 수집, 광고성 리뷰 제거, 신뢰도 높은 후기만을 요약하여 제공

2순위 인텐트 (보조적 주요 기능)

- **위치 기반 인텐트** : 현재 위치 혹은 사용자가 지정한 위치를 기반으로 식당 정보 필터링 후 제공
- **조건 수정 및 재탐색 인텐트** : 기존에 입력한 조건을 수정, 다른 조건으로 재검색을 요청

3순위 인텐트 (부가 기능)

- **할인/혜택 안내 인텐트** : 특정 식당의 할인/혜택 정보를 안내
- **즐거찾기/공유 인텐트** : 추천받은 식당을 저장하거나 타인과 공유할 수 있도록 지원

3.2. 인텐트 세분화

<맛집 탐색 인텐트>

1. 조건 기반 탐색 인텐트

사용자가 입력하는 다양한 조건에 기반하여 해석하고 적합한 식당을 탐색

예시) “조용한 분위기의 식당 찾아줘”

2. 메뉴 키워드 탐색 인텐트

사용자가 특정 음식명이나 메뉴를 언급하면, 해당 메뉴가 주력인 식당을 추천

예시) “파스타 잘하는 집 추천해줘”

3. 시간대 기반 탐색 인텐트

조식, 브런치, 야식 등 특정 시간대에 최적화 된 식당을 추천하는 기능

예시) “야식 먹을 수 있는 곳 알려줘”

3.2. 인텐트 세분화

<리뷰 요약/분석 인텐트>

1. 긍/부정 리뷰 요약 이벤트

전체 리뷰 중 긍정적인 표현과 부정적인 표현을 분리하여 요약 제공

예시) “이 식당의 장점과 단점 알려줘”

2. 키워드 리뷰 요약 인텐트

사용자가 특정 키워드를 지정하면, 해당 키워드와 관련된 리뷰만 요약 제공

예시) “서비스에 대한 리뷰만 보여줘”

3. 리뷰 신뢰도 검토 인텐트

리뷰 작성자의 패턴이나 작성 양식을 분석하여, 해당 리뷰가 실제 방문자에 의해 작성된 것인지 판단하여 표시

예시) “이 리뷰 신뢰할 수 있어?”

3.2. 인텐트 세분화

<위치 기반 인텐트>

1. 목적지 기준 탐색 인텐트

사용자가 특정 목적지를 입력하면, 해당 위치 중심으로 주변의 식당을 탐색하여 추천

예시) “홍대 근처 맛집 추천해줘”

3.2. 인텐트 세분화

<조건 수정 및 재탐색 인텐트>

1. 조건 일부 수정 인텐트

기존에 입력한 조건 중 일부를 수정하여 다시 탐색

예시) “분위기는 그대로, 가격만 더 저렴한 곳으로”

2. 조건 추가 인텐트

탐색을 유지하면서 새로운 조건을 추가하여 필터링을 세분화

예시) “채식 가능한 곳으로 추가해줘”

3. 초기화 후 재탐색 인텐트

현재까지 입력된 모든 조건을 초기화하고 새롭게 탐색을 시작

예시) “처음부터 다시 찾아줘”

3.3. 기능 상호작용 설계

AI 기반 맛집 추천 챗봇 서비스는 사용자의 요청에 따라 다양한 인텐트가 개별적으로 작동하지만 실제 서비스에서는 이들 인텐트가 유기적으로 연결되어 하나의 대화 흐름을 구성

그 중 대표적인 대화 흐름 시나리오를 기반으로 각 기능 간 상호작용의 구조를 설명

따라서 아래와 같은 기능 간 연계성을 고려한 시나리오 중심의 설계가 필수적

시나리오 1 : 기본 조건 입력 -> 탐색 -> 리뷰 확인

시나리오 2 : 목적지 설정 -> 탐색 -> 조건 추가 -> 재추천

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

3.4. 인텐트 유효성 검증

정의된 인텐트 들이 실제 사용자 요구를 적절하게 반영하고 있는지 확인하기 위해, 4가지 인텐트 유효성 검증을 수행 예정

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

해당 유효성 검증은 단순한 기술 구현의 정확도를 확인하는 것을 넘어, 사용자의 경험의 품질을 개선하기 위한 중요한 기준으로 작용할 것으로 기대

Part 4 슬롯 구성

4.1. 슬롯 개요

슬롯은 AI 기반 챗봇이 사용자의 발화를 이해하고, 적절한 응답을 생성하기 위해 필수적으로 수집해야 하는 정보 항목

각 인텐트 실행 시 필요한 정보를 구조화된 형태로 저장하고, 대화 흐름 속에서 자연스럽게 이를 확보 할 수 있도록 **슬롯**이 설계



본 서비스에서는 인텐트별로 필요한 정보를 파악하여 주요 슬롯을 정의

슬롯마다 데이터 유형과 제약 조건을 설정하여 안정적인 정보 수집과 응답 생성을 보장

사용자의 경험을 고려하여, 명시적 질의 없이도 자연어 흐름 중에서 슬롯을 추출할 수 있도록 설계

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

4.4. 슬롯 설계 시 고려사항

슬롯은 단순한 데이터 수집을 넘어, 대화 자연스러움과 서비스 응답 품질을 좌우하는 핵심 요소이므로
본 서비스에서의 슬롯은 다음 4개의 설계 원칙을 적용

1. **명시적 질문 없이도** 자연어 대화에서 슬롯 정보를 추출 할 수 있어야 함
2. **슬롯 누락 시에는 확인** 질문을 통해 보완하도록 설계
예시) " 어디서 식사하실 예정이신가요? "
3. **제약 조건을 통해** 데이터 일관성과 품질을 유지해야 함
4. **서비스 특성에 따라 슬롯 우선** 순위를 설정하여, 탐색 우선 -> 보조 기능 -> 부가 기능 순으로 응답의 핵심정보를 확보

Part 5 코퍼스 구축

5. 고객 발화 구축

맛집 탐색 인텐트

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

리뷰 요약/분석 인텐트

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

5. 고객 발화 구축

위치 기반 인텐트

[장소:신촌] 근처 맛집 추천해줘.

[장소:강남역]에서 가까운 저녁 식당 있을까?

[장소:건대입구역] 주변에 늦게까지 하는 식당 알려줘.

[장소:내 위치] 기준으로 지금 갈 수 있는 집 알려줘.

오늘 [장소:서울역] 근처에서 [식사 시간대:점심] 약속 있어. 근처 괜찮은 곳?

조건 수정 및 재탐색 인텐트

방금 추천해준 곳보다 [가격대:더 저렴한] 데로 바꿔줘.

[분위기:분위기]는 그대로 두고 [메뉴 카테고리:한식]만 골라줘.

아까 추천한 집 중에 [메뉴 카테고리:채식] 가능한 곳 추가해줘.

조건 다 지우고 처음부터 다시 찾아줘.

지금 조건에서 [분위기:데이트 분위기]만 추가해줘.

5. 고객 발화 구축

할인/혜택 안내 인텐트

지금 [할인 여부:할인 중]인 식당만 알려줘.

[할인 여부:쿠폰] 받을 수 있는 맛집이 있을까?

[할인 여부:첫 방문자 혜택] 있는 데가 어디야?

이번 주 [할인 여부:할인]하는 [메뉴 카테고리:이탈리안 레스토랑] 있어?

[할인 여부:포인트 적립] 가능한 곳 알려줘.

즐거찾기/공유 인텐트

이 집 [행동:저장]해둘래.

친구한테 이 식당 추천하고 싶어. [행동:공유] 링크 줘.

나중에 또 보고 싶으니까 [행동:즐거찾기] 해줘.

이 집 다음에 또 가고 싶은데 [행동:저장]해줘.

가족 단톡방에 이 식당 정보 [행동:공유]할 수 있어?

Part 6 핵심 엔티티 정의

6.1. 엔티티 식별

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

각 엔티티는 사용자 발화 내에서 특정 패턴이나 단어군으로 등장하며
슬롯 추출 또는 인텐트 분류의 기준이 됨

6.2. 속성(Attribute) 정의

엔티티명	속성 또는 제약 조건
장소	최대 20자 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 15자, 사전에 정의된 음식 카테고리 내 일치 여부 확인
리뷰 키워드	서비스, 청결, 맛, 양, 분위기 등 주요 평가지표와 일치해야 함

각 엔티티는 고유한 속성 또는 제약 조건이 설정되며
이는 발화 해석 시 정제된 결과를 도출하는데 사용됨

6.3. 엔티티 간 관계 설정

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

엔티티 간 관계를 설정하면 단순한 키워드 매칭이 아닌
의도 흐름 기반의 복합 분류가 가능해짐

➡ 개인화된 사용자 경험을 제공하는데 필수적

6.4. 동의어 처리 전략

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

다양한 사용자 표현을 수용하기 위하여 동의어 처리 체계가 반드시 필요함
 동의어 매핑을 통하여 여러 표현이 하나의 동일한 의미로 연결 될 수 있도록 설계

Part 7 성능 평가

7.1. 모델 선택

의도 분류(Intent Classification)와 **엔티티 인식(NER)**은 본 서비스에서 핵심적인 자연어 처리 기능

이를 위하여 분류와 추출 능력에 강점을 가진 LLM 계열 경량 모델 중 **LLaMA3 2B**를 선택



WHY?

LLaMA3 2B 모델은 Meta에서 공개한 최신 경량 언어 모델로, 비교적 적은 자원으로도 뛰어난 추론 성능을 보임

사용자의 요구에 따라 다양한 Task로 **파인튜닝**이 가능한 구조를 가지고 있음

현재 모델은 영어 기반으로 사전학습 되어 있기에, 본 서비스에서는 **한국어 코퍼스를 기반으로 파인튜닝**을 계획

7.2. 학습 실행

모델 학습을 위해 정의된 인텐트 및 엔티티 정보를 기반으로 데이터를 구축하고, 아래와 같은 방식으로 학습을 수행
이때 모델 학습은 주로 Classifier와 NER Tagger로 구성됨

1. 전체 데이터셋을 **80%는 Train / 20%는 Validation** 으로 Split

2. 과적합 방지를 위한 기본 전략 적용

- Early stopping
- Dropout
- Oversampling or class weighting (불균형 클래스 대응)

3. Optimizer는 **Adam**, 손실 함수는 **CrossEntropy** 기준 사용

7.3. 성능 평가

훈련된 모델의 성능은 다음과 같은 4개의 정량 지표를 사용하여 평가

1. **정확도 (Accuracy)** : 전체 예측 중 맞춘 비율
2. **정밀도 (Precision)** : 예측한 결과 중 실제로 맞은 비율
3. **재현율 (Recall)** : 실제 정답 중 모델이 맞춘 비율
4. **F1 Score** : 정밀도와 재현율의 조화 평균

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

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision Value $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

7.4. 모델 개선 및 데이터 보강

성능 평가 결과를 바탕으로, 정확도가 낮거나 오분류 비율이 높은 인텐트 및 엔티티에 대해 다음 4가지 방식으로 개선을 진행

1. 실제 사용자 발화 로그 기반 신규 코퍼스 수집
2. 자주 혼동되는 의도 간 샘플 보장
3. 동의어, 지역어, 축약어 등 자연 발화를 반영한 도메인 확장
4. 지속적인 fine-tuning을 통한 챗봇 성능 고도화

또한, 추후 대규모 사용자 유입을 고려하여, 필요시 LLaMA3 2B의 상위 모델인 LLaMA3 8B로 업그레이드 할 수 있도록 학습 파이프라인을 유연하게 설계하고자 함

감사합니다