

LLM 기반의 질의 생성 및 RAG 기법을 활용한 숙박 추천 알고리즘 연구

A Study on Accommodation Recommendation Algorithms Using LLM-based Query Generation and RAG Methods

요 약

현대 사회에서 숙박 애플리케이션의 등장은 숙박업소를 편리하게 예약하고, 나아가 만족도 높은 여행을 할 수 있도록 도와주는 핵심 도구이다. 하지만, 많은 숙박 애플리케이션이 핵심어 기반의 숙박업소 추천 알고리즘을 사용하고 있으며, 이러한 추천 시스템은 사용자의 어휘 다양성에 따른 매칭 한계와 표현 유연성 부족 문제를 갖기 때문에, 자연어 기반의 검색 기능 지원이 필요한 상황이다. 본 연구에서는 자연어 형식의 사용자 질의 기반 숙박업소 검색을 목적으로 하는 거대 언어 모델 기반의 숙박업소 추천 알고리즘을 제안한다. 제안하는 알고리즘은 숙박업소 검색에 관한 예상 질의를 생성하고, 이를 벡터화하여 RAG의 성능을 향상시키는 전략을 취하며, 공공API를 통해 수집된 벤치마크 시험에서 우수한 예측 정확도를 보여주었다. 이러한 결과는 자연어 질의의 다양성과 상황 맥락을 반영하여, 보다 직관적이고 접근성 높은 추천 서비스 개발의 원천 기술이 될 것으로 기대된다.

1. 서론

현대 사회에서 Airbnb, 야놀자, 여기 어때 등의 숙박 애플리케이션은 여행 일정을 효과적으로 계획할 수 있도록 지원하는 핵심 도구로 자리 잡았다[1]. 이러한 O2O(online-to-offline) 서비스를 통해 사용자는 숙박업소를 편리하게 예약할 수 있으며, 이를 통해 여행 일정을 효율적으로 조정 및 만족도 높은 여행을 다녀올 수 있게 되었다.

많은 숙박업소 검색 시스템은 핵심어(keyword) 기반의 추천 시스템 알고리즘을 이용한다[2,3]. 그런데, 이러한 추천 시스템에서는 사용자가 시스템에서 다루지 않는 핵심어로 검색을 시도하는 경우 적절한 숙박업소를 추천하는 것이 어렵다는 한계가 있다. 더욱이, 추천 시스템에서 지원하는 핵심어의 개수는 제한적인 반면에, 사용자들의 어휘는 몹시 풍부하기 때문에, 효과적인 검색 서비스를 제공하기 위해서는 사용자가 자연어 질의를 통해 숙박업소를 검색할 수 있는 기술이 필요하다.

최근 대규모 언어 모델(large language model, LLM)을 활용하여 자연어 질의응답 시스템을 도입하는 산업이 늘어나고 있다[4]. 야놀자와 같이 일부 숙박 애플리케이션에서는 LLM을 활용하여 숙박업소에 대한 리뷰 요약 또는 간단한 추천 기능을 제공하고 있으나, 이는 보조적인 수준이며 LLM을 활용한 자연어 질의 기반의 숙박 검색 서비스는 부재한 상황이다.

사용자의 자연어 질의 기반 숙박업소 추천 시스템을 구현하기 위해 검색 증강 생성(retrieval-augmented generation, RAG)이 활용될 수 있다. RAG는 사용자의 질의에 대해 관련 문서를 먼저 검색한 후(retrieval), 검색 결과에 기반한 응답을 생성(generation)하는 LLM 기반 기술로, 정보 검색의 정확성과 자연어 응답의 유연성을 동시에 확보할 수 있게 한다[5]. 사용자 정의 문서를 기반으로 검색·생성 과정을 수행하기 때문에, 특정 도메인에 대한

지식을 반영한 응답이 가능하다는 장점이 있다. 숙박 추천 시스템의 경우, 숙박업소 정보(예: 위치, 시설, 정책 등) 또는 사용자 리뷰 데이터를 문서로 제공 및 RAG를 통해 자연어 질의에 대한 적절한 숙박 추천 답변을 얻을 수 있다. 하지만, 숙박업소 정보와 사용자 질의는 문장 구성이 상이하기 때문에 벡터(vector) 유사도 기반으로 동작하는 RAG에서는 두 자연어 문서 간의 유사도가 낮게 평가되어 효과적인 숙박업소 추천이 어렵다.

이러한 한계점을 극복하기 위해, 본 연구에서는 숙박 정보를 임베딩(embedding)하는 것 대신에, 숙박 정보 기반의 예상 질의(expected question)를 생성하고 예상 질의를 임베딩하여 향상된 벡터 유사도 계산 및 효과적인 숙박업소 추천 알고리즘을 제안한다. 제안하는 알고리즘은 사용자가 숙박업소 검색시 사용할 것으로 예상되는 질문들을 LLM을 통해 생성 및 VectorDB에 저장하고, 이후 실제 질의(query)와 저장된 질의들이 벡터 유사도를 평가함으로써 사용자에게 정밀하고 상황 맞춤형 속소를 추천할 수 있다.

2. 제안하는 방법

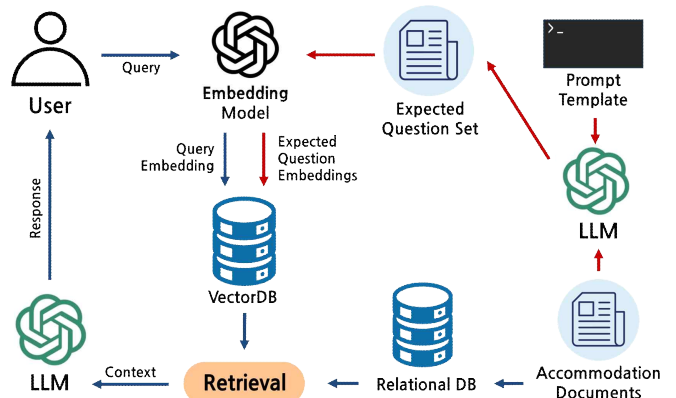


그림 1. 예상 질의 생성 및 RAG 기반 숙박 추천 알고리즘

제안하는 방법은 크게 3단계로, LLM을 통해 예상 질의를 생성하고, VectorDB에 예상 질의를 저장하며, RAG를 통해 실제 질의와 유사한 질의를 검색 및 해당 예상 질의에 대응하는 숙박업소를 추천 결과로 반환한다 (그림 1).

2.1. 예상 질의 생성

숙박업소 정보로부터 예상 질문을 자동 생성하기 위해 숙박업소의 지역, 시설, 상황 정보 등을 다양하게 활용하도록 고안된 프롬프트 템플릿(prompt template)을 이용한다. 숙박업소에 대한 자연어 형식의 개요가 주어지면 그림 2와 같이 본 연구에서 사전에 작성한 예상 질문 예시 및 few-shot 기법 [6]을 통해 해당 숙박업소에 대한 예상 질문들이 생성된다. 본 연구에서는 숙소 개요마다 5개의 예상 질문을 생성하도록 하였다.

2.2. 예상 질의 저장

생성된 예상 질의들은 문장 단위로 분리 및 정제되고, 숙박업소 추천 기능을 수행할 LLM의 임베딩 모델을 이용하여 문장들을 벡터화 및 VectorDB에 저장한다. 이러한 문서 임베딩 활용 전략을 통해 사용자 질의와 유사한 예상 질의를 탐색함으로써 사용자에게 알맞은 숙박업소를 제시할 수 있다.

2.3. RAG 기반 숙박업소 검색

예상 질의 기반의 숙박업소 추천을 위해 RAG는 예상 질문 임베딩을 포함하는 VectorDB와 숙박업소 개요를 포함하는 Relational DB(RDB)를 이용한다. 구체적으로, 입력된 사용자 질의에 대한 예상 질의와의 비교를 통해 유사도가 높은 숙박업소들을 검색하고 [7], 유사도가 높은 숙박업소들의 메타데이터(개요, 주소 등)를 RDB로부터 조회함으로써 사용자에게 추천하는 숙박업소 정보를 제공한다.

3. 실험 방법

3.1. 벤치마크 데이터집합

제안하는 알고리즘의 성능 평가를 위해 한국관광공사가 제공하는 오픈API인 TourAPI(v4.0)¹⁾를 이용하여 숙박업소 데이터를 수집하였다. 수집된 데이터에서 숙소명과 개요 항목만 추출하였으며, 해당 데이터를 통해 사전 질의 생성 및 벤치마크 데이터집합을 구축하였다.

제안하는 알고리즘을 평가하기 위해서는 숙박업소 검색에 관한 시험 질의(test query) 및 시험 질의에 대한 검색결과가 올바른지를 판단하기 위한 레이블(label) 데이터가 필요하다. 본 연구에서는 그러한 벤치마크 데이터집합을 구축하기 위해 각 숙박업소 개요에서의 핵심어를 정답 레이블로 정의하였으며, 핵심어를 이용하여 시험 질의를 생성하였다.

3.2. 핵심어 추출 알고리즘

본 연구에서는 숙박업소 개요에서 핵심어 추출을 위해

다음은 숙소에 대한 개요입니다. 이 숙소에 대해 사용자가 실제로 숙소 추천을 요청할 때 할 수 있는 질문을 5가지 생성해주세요. 질문은 실제 사용자가 이 숙소를 추천받기 위해 조건을 말하는 방식으로 구성되어야 합니다.

[개요]

{overview} # 숙소 개요

[예상 질문 리스트]

- 1.서면 근처에 롯데백화점과 가까운 숙소를 추천해 줄 수 있어?
- 2.고령자와 장애인 분들이 머물기 편한 숙소가 있니?
- 3.신불산에 가기 편리한 숙소를 찾고 있어요. 추천해주세요.
- 4.부산 수영구에 출장 갈 일이 있어서, 고급스러우면서도 청결한 숙소를 찾고 있습니다. 비즈니스 여행에 적합한 모델 추천 부탁드립니다.
- 5.부산에 출장을 가는데, 편안하게 쉴 수 있는 특실을 찾고 있습니다. 추천해주실 만한 곳이 있나요?

그림 2. 개요 기반 예상 질의 생성을 위한 프롬프트 템플릿

당신은 관광 숙소 추천 시스템에서 사용자 질의를 생성하는 전문가입니다. 아래 핵심어(지명, 장소, 시설 등)를 포함하여 자연스럽고 실제 사용자가 말할 법한 질문을 생성해주세요.

조건:

- 핵심어는 반드시 모두 포함되어야 합니다.
- 아래 TF-IDF 점수표는 키워드의 중요도를 나타내며, 상위 핵심어들이 이미 선택되어 있습니다.
- 핵심어는 질문 속에 자연스럽게 녹아들게 하세요. 단순 나열은 피해주세요.
- 사용자는 해당 장소 근처의 숙소를 찾고 있는 상황입니다.

예시:

TF-IDF 점수표:

부산: 0.342, 자갈치시장: 0.271

선택된 핵심어: 부산, 자갈치시장

→ 질문: 부산에서 자갈치시장 근처 괜찮은 숙소 있어?

TF-IDF 점수표:

{tfidf_str}

선택된 핵심어: {joined_keywords}

→ 질문:

그림 3. 핵심어 기반 시험 질의 생성을 위한 프롬프트 템플릿

TF-IDF(term frequency-inverse document frequency) [8], KoNLPy[9], 그리고 ChatGPT4o를 활용하였다. 구체적인 추출 절차는 다음과 같다. 개요 데이터에 TF-IDF를 적용하여 빈도 점수와 함께 문서 내의 단어들을 추출하고, KoNLPy를 이용하여 명사 단어만 선별하고, ChatGPT를 이용하여 불용어(stopwords) 제외 및 고유명사인 단어들만을 골라내었다. 이후 각 개요에 대해 빈도 점수가 높은 상위 고유명사 단어들을 해당 개요의 핵심어로 지정하였다. 본 핵심어 추출 알고리즘을 통해 총 339개의 숙소에 대한 핵심어 데이터를 확보하였다.

3.3 시험 질의 생성 알고리즘

본 연구에서는 각 숙소의 개요로부터 수집된 핵심어들을 하나의 집합으로 모으고, 해당 집합으로부터 1~3개의 핵심어를 무작위로 추출하여 시험 질의를 생성하였다. 추출된 핵심어 기반의 시험 질의 생성을 위해 ChatGPT4o를

1) <https://www.data.go.kr/data/15101578/openapi.do>

표 1. RAG 기반 숙박업소 추천 알고리즘 정확도 비교

Benchmark Dataset	N=1			N=2			N=3		
	Acc ₁	Acc ₃	Acc ₅	Acc ₁	Acc ₃	Acc ₅	Acc ₁	Acc ₃	Acc ₅
Baseline	0.40	0.48	0.49	0.35	0.46	0.50	0.35	0.42	0.46
Ours	0.50	0.62	0.63	0.41	0.56	0.56	0.36	0.5	0.51

이용하였으며, 그림 3과 같은 프롬프트 템플릿을 통해 시험 질의를 생성하였다. 생성 과정에서는 질의 구조의 일관성 유지 및 질문 내 핵심어 필수 포함을 위해 few-shot[6] 학습이 아닌 one-shot 학습 방식을 채택하였다. 또한, 불필요하게 일반화된 표현이나 핵심어 누락을 방지하는 규칙도 설정하였다. 본 시험 질의 생성 알고리즘을 통해 무작위 추출 크기 N=1,2,3별로 벤치마크 데이터집합을 구축하였으며, 각 데이터집합은 100개의 시험 질의를 가지도록 하였다.

3.4. 성능 평가 지표

벤치마크 데이터집합에서의 알고리즘 성능 평가를 정량적으로 수행하기 위해 top-K 정확도(Acc_k)를 사용하였다. 여기서 K는 시험 질의에 대해 알고리즘이 추천하는 숙박업소의 개수이며, 본 연구에서는 K=1,3,5 상황별로 평가하였다. 알고리즘 예측의 정답 여부는 예측된 K개의 숙박업소 중에서 정답 레이블인 N개의 핵심어를 모두 포함하는 경우가 있는지로 판단된다. 즉, 각 시험 질의에 대한 K개 예측 결과 중에 해당 시험 질의에 대응되는 N개 핵심어를 모두 포함하는 숙박업소가 적어도 1개가 있다면 그 예측은 정답인 것으로 판단된다. 따라서 top-K 정확도는 전체 시험 질의 개수 중 정답인 경우의 비율로 계산된다.

4. 실험 결과

본 연구에서 제안하는 예상 질의 기반 추천 시스템의 우수성을 평가하기 위해, 일반적인 RAG 방식인 숙박업소 개요 기반의 추천 시스템을 비교 모델(baseline)로 선택하였다. 표 1은 3개의 벤치마크 데이터집합에 대한 제안하는 알고리즘과 비교 알고리즘의 top-K 정확도 성능 비교 결과를 보여준다. 개요를 기반으로 하는 비교 모델은 K=1,3,5 모든 경우에서 정확도 50%를 넘지 못하는 제한적인 성능을 보였다. 반면에, 제안하는 알고리즘은 맞추어야 하는 핵심어 개수(N)가 3개까지 늘어나더라도 50% 이상의 예측 정확도를 기록하였으며, 기존 모델보다 우월한 성능을 보여주었다.

5. 결론

본 연구에서는 숙박업소 정보를 직접 임베딩하는 기존 방식 대신에 해당 정보 기반으로 생성된 예상 질의를 임베딩하는 방식을 제안하였다. 제안하는 알고리즘은 RAG 구조 내에서 예상 질의를 통해 사용자 질의와의 의미적 연관성을 강화하며, 그 결과로 기존 개요 기반 방식보다 향상된 추천 정확도 성능을 보였다. 이러한 결과는 자연어

기반 질의응답 환경에서 새로운 벡터화 전략을 제시함으로써, 다양한 표현의 사용자 질의에도 유연하게 대응할 수 있는 추천 시스템 구현 가능성을 보여준다. 추후 연구에서는 제안 기술을 활용하여 다양한 연령층과 디지털 환경에 익숙하지 않은 사용자들도 쉽게 활용할 수 있는 자연어 기반 추천 서비스를 개발하여 기술에 대한 실증 테스트를 수행할 계획이다.

참고문헌

- [1] 박정훈, and 김승인. "숙박 애플리케이션 재이용에 영향을 끼치는 요인과 개선 방향 제안-야놀자와 여기어때를 중심으로." 디지털융복합연구 16.8 (2018): 293-299.
- [2] 김민기, and 박두순. "개인성향과 협업필터링을 이용한 숙박업소 추천 시스템." 한국정보처리학회 학술대회논문집 24.1 (2017): 525-528.
- [3] 홍영민, and 박영덕. "NLP 기반 여행 리뷰 분류 및 추천 시스템 설계." 한국정보처리학회 학술대회논문집 30.1 (2023): 636-638.
- [4] 정천수. "LLM 애플리케이션 아키텍처를 활용한 생성형 AI 서비스 구현: RAG 모델과 LangChain 프레임워크 기반." 지능정보연구 29.4 (2023): 129-164.
- [5] 이광우, and 김수균. "국내 기업을 위한 RAG 구조 기반 질의응답시스템 설계." 한국컴퓨터정보학회논문지 29.7 (2024): 81-88.
- [6] 서상현, and 김준태. "퓨샷 관계 분류 성능 향상을 위한 메타 데이터 증강." 정보과학회 컴퓨팅의 실제 논문지 28.4 (2022): 256-261.
- [7] 이수빈, and 배호. "쿼리와 가상 문서 임베딩 결합을 활용한 제로샷 밀집 기반 검색 성능 향상 연구." 정보처리학회 논문지 (KTSDE) 14.3 (2025): 161-171.
- [8] 박대서, and 김화중. "TF-IDF 기반 키워드 추출에서의 의미적 요소 반영을 위한 결합벡터 제안." 한국정보기술학회논문지 16.2 (2018): 1-16.
- [9] Park, Eunjeong L., and Sungzoon Cho. "KoNLPy: Korean natural language processing in Python." Annual Conference on Human and Language Technology. Human and Language Technology, 2014.