**SK네트웍스 Family AI과정 9기 4팀 PetMind 1**

**김도연, 김정훈, 김하늘, 이다인, 임수연**

**인공지능 학습 결과서**

* 개요
  + 제출일자: 2025-05-16
  + 작성자: 이다인, 임수연

| **실험 개요** | 벡터 DB 및 RAG 구축을 위한 임베딩 모델 선정을 위해 각 모델별 임베딩 후 정량적으로 측정/비교한다. |
| --- | --- |
| **모델 후보 선정**  **및**  **실험 과정** | **(1) 임베딩 모델 후보 리스트**   * 2024년 12월 공개 된 한글 텍스트 임베딩 모델 리더보드의 ‘Top-k 1’과 ‘Top-k 3’의 평균 스코어가 가장 높은 두 모델 ’**KURE-v1**’, ‘**BGE-m3-ko**’과 유료 모델인 ‘**text-embedding-3-small**’ 을 선정.   (<https://github.com/nlpai-lab/KURE?tab=readme-ov-file#mteb-ko-retrieval-leaderboard>)      **(2) 임베딩 모델 구축**   * FAISS를 이용하여 문서데이터를 임베딩하고 벡터DB 생성. * 생성한 벡터DB로 ‘MRR’, ‘Hit Ratio@K’를 통한 정량적 측정/비교.   **(3) 모델 실험**   * 동일한 문서 데이터에 적용하여 벡터화를 수행. * 동일한 질의-정답 쌍(Ground Truths)을 기준으로 검색 성능 비교. * 각 모델로 생성된 벡터 DB에 대해 동일한 질의 입력 후, 검색 결과의 정답 포함 여부를 기준으로 MRR (Mean Reciprocal Rank), Hit Ratio@1, Hit Ratio@5의 지표로 성능 평가. |
| **결과 비교** | **(1) 모델 비교 결과**   | **Model** | **MRR** | **Hit Ratio1** | **Hit Ratio5** | | --- | --- | --- | --- | | **text-embedding-3-small** | 0.593 | 0.56 | 0.67 | | **KURE-v1** | 0.444 | 0.33 | 0.56 | | **BGE-m3-ko** | 0.5 | 0.44 | 0.56 |   **(2) 최종 모델 선정**  ‘**text-embedding-3-small**’ 모델이 MRR 및 Hit Ratio@K 지표에서 가장 안정적이고 고른 성능을 보였으며, 특히 다양한 질의 유형에 대해 높은 검색 일관성을 유지한 OpenAI의 text-embedding-3-small 모델을 최종 임베딩 모델로 선정. |

| **실험 개요** | * 프롬프트 엔지니어링 적용 전과 후의 모델 응답 성능을 비교하여, 프롬프트 개선 효과를 정성적으로 분석한다. |
| --- | --- |
| **실험 대상** | * Qwen3-8B 동일 모델 * 동일한 10개 내외의 상담 질문 세트 * 서로 다른 프롬프트 버전 (기본버전 vs 개선버전) |
| **실험 설계** | * **설계 개요** * 반려견 상담 특성상 보호자의 질문은 단순 정보 요청부터 행동 교정, 감정적 토로까지 다양하다고 판단. * 따라서 챗봇 사용자의 질문을 **정보 탐색, 행동 교정, 감정 해소**로 분류하고 유형에 맞는 상담 프롬프트 적용. * **시스템 구성 흐름**  1. 사용자 질문 입력 2. 질문 분류 모델 응답 3. 상담 유형 결정 4. 해당 유형에 맞는 상담 프롬프트 적용 5. LLM 응답 생성 및 제공  * **시스템 프롬프트**  | 프롬프트 유형 | 프롬프트 설계 | | --- | --- | | 질문 분류 | 1. 사용자의 질문을 3가지 유형(행동 교정 / 지식 탐색 / 감정 공감) 중 하나로 분류하기 위해 설계 2. 분류 기준은 질문에 나타난 행동의 문제성, 정보의 일반성, 감정의 표현 여부를 기준으로 함 3. 이전 질문 / 이전 응답 / 이전 분류 결과까지 고려하여 맥락 기반 분류 가능 4. 출력 형식은 반드시 “카테고리: 행동 교정” 형식으로 고정 | | 행동 교정 | 1. 반려견의 문제 행동에 대해, 정보 기반 분석 → 해결책 제시 → 후속 질문의 3단계 응답 구조를 강제 2. 보호자의 입력만을 근거로 분석해야 하며, 추측이나 근거 없는 일반화 금지 3. 해결책은 반드시 1가지 방법만 명확히 제시해야 하며, 대안 나열 금지 4. 상담을 이어가기 위해 상황 파악 목적의 후속 질문 1개 포함 필수 | | 정보 탐색 | 1. 반려견 관련 일반적인 궁금증에 대해 정보를 제공하기 위해 설계 2. 보호자가 처음 듣는 내용도 이해할 수 있도록 쉬운 언어 사용을 요구 3. 의학적 판단이 포함된 정보는 반드시 제외 4. 정확한 정보만 전달, 이론이나 불확실성은 제거 | | 감정 공감 | 1. 반려견 양육 과정에서 보호자가 겪는 정서적 어려움(피로감, 상실감 등)에 대해 응답하기 위해 설계 2. 응답 흐름은 공감 → 감정 원인 정리 or 탐색 → 조언 제시의 3단계로 구성 3. 감정의 원인이 불분명한 경우, 감정 탐색을 위한 추가 질문 1~2개 포함 |  * **user 프롬프트**   + LLM이 단순한 질문만 보고 추론하지 않고, 반려견의 프로필과 보호자의 질문을 함께 인식할 수 있도록 프롬프트를 설계   + [보호자 질문]과 [반려견 프로필]을 조합한 구조로 구성 |
| **실험 방법** | * 결과 비교의 일관성을 위해 하나의 보호자 페르소나를 설정하고, 해당 상황에서 도출 가능한 질문을 3가지 선정 * **질문 구성**  | 분류 유형 | 질문 | 모델 답변의 방향 | | --- | --- | --- | | 행동 교정 | Q. 내가 출근 준비만 하면 메이가 짖기 시작하는데, 이건 어떻게 이해해야해? | 특정 행동에 대한 트리거 분석과 원인에 기반한 행동 해석을 수행하여 행동 관리 방안을 제시 | | 지식 탐색 | Q. 메이가 요즘 자꾸 베개에 머리를 비비는데, 이건 왜 그런거야? | 반려견의 머리 비비기 행동을 알레르기, 스트레스 등 일반적 원인 중심으로 설명하고, 필요 시 수의사 진료를 권장하는 방향으로 정보만 제공 | | 감정 공감 | Q. 메이를 다시 훈련사한테 맡기긴 했지만, 제가 너무 지쳐요. 나 없이 메이가 잘 지낼 수 있을까요? | 보호자의 심리적 피로와 불안감에 공감하고, 감정의 원인을 정리한 뒤 정서적 지지와 현실적인 위로를 제공 | |
| **실험 결과** | * **프롬프트 적용 전 / 후의 모델 답변 결과 비교**   + Case1) 행동 교정   + 프롬프트 적용 전      * + 프롬프트 적용 후      * + Case2) 지식 탐색   + 프롬프트 적용 전      * + 프롬프트 적용 후      * + Case3) 감정 공감   + 프롬프트 적용 전      * + 프롬프트 적용 후 |
| **결론** | * 응답 구조 체계화   + 모델 개선 전에는 하나의 응답에 여러 원인을 중첩적으로 나열하고 행동 분석과 해결책을 구분하지 않아 핵심 정보를 파악하기 어려웠던 반면, 개선된 모델은 프롬프트에 정의된 상담 구조(① 분석 → ② 해결책 → ③ 후속 질문)를 정확히 따르며, 메시지를 단계적으로 구조화 * 상담 목적 분리   + 개선 전에는 감정 공감, 정보 제공, 행동 분석이 명확히 구분되지 않아 질문 유형에 상관없이 유사한 톤의 응답이 생성되었지만, 개선 후에는 질문을 사전에 분류하고 그 결과에 따라 상담 방식이 분기되어 각 질문의 의도(행동 교정 / 정보 탐색 / 감정 공감)에 맞춘 응답 제공 * 대화 흐름 및 후속 유도   + 개선 전 응답은 후속 질문 없이 흐름이 단절되는 경우가 많았던 데 비해, 개선된 모델은 상담을 이어가기 위한 구체적이고 상황 중심적인 후속 질문을 포함함으로써 대화의 연속성 확보 |