**테스트 계획 및 보고서**

**테스트 계획:**

**응급처치 상황과 관련된 질문(Positive)과 관련없는 질문(Negative)에 대한 답변에 대해 평가**

**1. 응급처치 상황과 관련된 질문에 대한 답변 평가**

**1. 뱀에 물렸어.**

**2. 동상 걸린 것 같아.**

**3. 다리 부러진 것 같아.**

**4. 손가락 부러졌어.**

**5. 동생 목에 사탕이 걸렸어.**

**6. 내가 오늘 산을 올랐는데 산을 오르다가 발목을 접질러서 굴러 떨어졌어.**

**2. 응급처치 상황과 관련없는 질문에 대한 답변 평가**

**1. 테슬라에 물렸어.**

**2. 발에서 냄새나.**

**3. 아 집가고 싶다.**

**4. 일론머스크가 미워.**

**5. 발 도봉산의 풍경에 대해 말해줘.**

**RAG, LangChain 적용 챗봇**

1. 응급처치 상황과 관련된 질문에 대한 답변 평가

|  |  |  |
| --- | --- | --- |
| Q | A | 평가 |
| 뱀에 물렸어. |  | O |
| 동상 걸린 것 같아. |  | O |
| 다리 부러진 것 같아. |  | O |
| 손가락 부러졌어. |  | O |
| 동생 목에 사탕이 걸렸어. |  | O |
| 내가 오늘 산을 올랐는데 산을 오르다가 발목을 접질러서 굴러 떨어졌어. |  | O |

1. 응급처치 상황과 관련없는 질문에 대한 답변 평가

|  |  |  |
| --- | --- | --- |
| Q | A | 평가 |
| 테슬라에 물렸어. |  | O |
| 발에서 냄새나. |  | X |
| 아 집가고 싶다. |  | O |
| 일론머스크가 미워. |  | O |
| 도봉산의 풍경에 대해 말해줘. |  | O |

**RAG 없는 gpt4o-mini**

1. 응급처치 상황과 관련된 질문에 대한 답변 평가

|  |  |  |
| --- | --- | --- |
| Q | A | 평가 |
| 뱀에 물렸어. |  | O |
| 동상 걸린 것 같아. |  | O |
| 다리 부러진 것 같아. |  | O |
| 손가락 부러졌어. |  | O |
| 동생 목에 사탕이 걸렸어. |  | O |
| 내가 오늘 산을 올랐는데 산을 오르다가 발목을 접질러서 굴러 떨어졌어. |  | O |

1. 응급처치 상황과 관련없는 질문에 대한 답변 평가

|  |  |  |
| --- | --- | --- |
| Q | A | 평가 |
| 테슬라에 물렸어. |  | X |
| 발에서 냄새나. |  | X |
| 아 집가고 싶다. |  | X |
| 일론머스크가 미워. |  | X |
| 도봉산의 풍경에 대해 말해줘. |  | X |

테스트 평가 :

RAG, LangChain 적용 챗봇:

* 응급처치 상황에 제대로 답변한 횟수 TP: 6 FN: 0
* 응급이 아닌 상황에 제대로 답변한 횟수 TN: 4 FP: 1
* 정밀도: 0.85
* 재현율: 1.0
* F-Score: 0.9189

단순 GPT4o-mini

* 응급처치 상황에 제대로 답변한 횟수 TP: 6 FN: 0
* 응급이 아닌 상황에 제대로 답변한 횟수 TN: 0 FP: 5
* 정밀도: 0.54
* 재현율: 1
* F-Score: 0.70