llama 2 시도 / 대시보드 구현 시도





- 1. 라마 2
 - 한국어 지원X, 성능 매우 떨어짐

[>>> 한국어로 대답해줘.한국 대통령은 누구야? Of course, I'd be happy to help! Here's a breakdown of the conversation in Korean:

user: 한국어로 대답해 한국 대통령은 누구야? (Can you answer in Korean?) me: Ah, I see! The President of South Korea is currently 문재명 (Moon Jae-in). In Korean, we would say 문 재 명 (mun jae-myeong) to refer to him. user: 이해하세요 (Thank you)

- 2. 한국어 파인튜닝된 버전: hugging face
 - Ilama2-koen-13b: beomi/llama-2-koen-13b
 - Ilama2-ko-7b: beomi/llama-2-ko-7b
- 3. 그 외 모델들:
 - SOLAR 기반: beomi/OPEN-SOLAR-KO-10.7B
 - Yi 기반: beomi/Yi-Ko-6B
- 4. hugging face 🤗 사용 방법: https://yunwoong.tistory.com/252



- 1. colab에서는 gpu 유료 버전을 사용해야 돌아간다.
- 2. 서버 연결해서 Jupyter hub로 돌려보았지만 아직 해결이 잘 안되고 있다.
 - hugging face 토큰 연결이 안 됨
 - 시도 중... → https://medium.com/@yashsk8/logging-in-to-huggingface-from-jupyter-notebook-without-interactive-prompt-2cb945b4905c



대시보드

- 1. PyGWalker를 사용했을 때
 - a. 장점: 드래그 & 드랍으로 그래프를 만들 수 있다.
 - b. 단점:
 - i. 우리가 원하는 느낌의 대시보드를 어떻게 만들 수 있을지 더 고민을 해봐야 한다.
 - ii. 실시간 데이터 반영이 조금 힘들 수도 있다.
- 2. Streamlit Dashboard 테마 format을 사용했을 때
 - a. 장점: 우리의 의도에 맞도록 이것 저것 수정할 수 있을 것이라고 생각한다.
 - b. 단점: 그래프 하나하나를 코드로 구현해야 한다.

대시보드 format 예시 사이트



프롬프트 엔지니어링

어떤 데이터 위주로 먼저 볼 것인지(날짜, 매체, 비용 등)

- → 매체 (광고 상품)에 대해 현 일을 기준으로 일주일 기간동안의 데이터 값을 대시보드에 보여주고, 전일 대비 코멘트 요약 생성을 목표로 함. 어떤 목적에 맞게 요약할 것인지(일자별 비교, 월별 비교, 매체별 비교 등등)
- → 전일 비교

기준 지점 정하기(증감율)

→ 3%로 정하여 결과물 받아 보기 프롬프트 결과 기준치 결정= 정확도 테스트

질문에 대한 대답의 성실도 체크

[체크리스트]

- □ 5% 이상의 변화를 보인 것에 대해서 언급하고 있다.
- □ 변화에 대해서 다른 지표와의 상관관계를 평가하고 있다.
- □ 긍정적인 지표만 언급하고 있다.

llama 2 시도 / 대시보드 구현 시도

□ 사실 나열 이외에 평가적인 요소가 들어간다. □ 숫자에 대한 정확도

결과 형식 결정

[변화에 대한 코멘트]

- 노출은 6.55% 증가하였으며, 클릭은 8.86% 증가하였습니다. 특히 클릭율은 2.17% 증가하여 광고 노출 대비 클릭 확률이 상승했습니다. 이는 광고 캠페인의 노출 및 클릭 효율
- CPC는 14.69% 감소하여 클릭 당 비용이 감소했습니다. 이는 효율적인 광고 소비를 나타내며, 광고 예산의 효과적인 활용이 이루어지고 있다고 평가할 수 있습니다.
- 광고비는 7.13% 감소하였습니다. 비용 절감 효과가 나타났으며, 효율적인 예산 운용으로 인한 긍정적인 변화입니다.
- CPS는 3.21% 감소하였으며, CPU는 7.13% 감소하였습니다. 특히 신규방문CPU는 21.64% 감소하였습니다. 이는 신규 방문 고객을 유입시키는 데 효과적인 비용 효율을 나타냅

[캠페인에 대한 평가]

- 광고 캠페인의 노출과 클릭이 증가하면서 광고 캠페인은 시청자들에게 더 많이 노출되고 클릭을 유도하는 데 성공했습니다.
- CPC와 광고비의 감소는 예산을 효과적으로 활용하고 있다는 것을 시사하며, 캠페인의 비용 관리가 효율적으로 이루어지고 있습니다.
- CPS와 CPU의 감소는 캠페인의 전반적인 비용 효율성 향상을 나타냅니다. 특히 신규 방문자를 유입하는 데 드는 비용이 감소하여 새로운 고객 획득에도 성공하고 있습니다.
- 전반적으로 광고 캠페인의 성과가 긍정적으로 향상되고 있으며, 예산을 더 효과적으로 운용하여 비용을 절감하고 있습니다.

[분석 결과]

- CPC는 14.69% 감소하여 91으로 나타났습니다. 이는 긍정적인 변화로 판단됩니다.
- CPU는 7.13% 감소하여 1,794,029로 나타났습니다. 캠페인의 효율이 향상되었음을 나타냅니다.
- 신규방문CPU는 3.21% 감소하여 316으로 나타났습니다. 이는 어느 정도 감소했지만, 여전히 안정적인 수치입니다.
- 심사CPA는 30.35% 감소하여 448,507로 나타났습니다. 이는 매우 큰 감소로 효율적인 캠페인 운영을 시사합니다.

[캠페인에 대한 평가]

- 1. 캠페인의 효율이 향상되었음을 확인할 수 있습니다. CPC와 CPU가 감소하면서 더 많은 클릭을 더 적은 비용으로 확보하고 있습니다.
- 2. 특히, CPC의 14.69%의 큰 감소는 광고 클릭을 얻는 데 더 효율적으로 예산을 사용하고 있다는 것을 시사합니다.
- 3. CPU의 감소는 방문자수와 클릭 수의 감소와 관련이 있을 것으로 판단됩니다. 하지만, 이는 신규방문CPU의 감소로 크게 영향받지 않고 있다는 것을 의미합니다.
- 4. 신규방문CPU와 심사CPA의 감소는 캠페인이 더 효과적으로 신규 방문자를 유입시키고, 심사 단계에서도 더 효율적으로 비용을 소모하고 있다는 것을 나타냅니다.

종합적으로, 현재 캠페인은 효율적으로 운영 중이며, 특히 CPC와 CPU의 감소는 예산을 효율적으로 활용하고 있음을 시사합니다. 신규방문과 심사 관련 지표도 안정적으로 유지되고



☞ 해야 할 일

- 1. Llama를 사용할 지, ChatGPT를 그대로 사용할 지
 - a. Llama를 사용하면 서버를 이용하여 Llama를 돌리는 법 알기
- 2. 전처리 코드에서 데이터 전처리가 더욱 필요함 (numerical로 변환, 변수 명 수정 등)
- 3. PyGWalker를 그대로 사용할지, 대시보드 그래프들을 코드로 직접 구현할 지 (테마 format)



학회

학회 이름	날짜	논문 제출 날짜	장소	대상	투고 분야	KCI 등재 여부	사이트	언어
한국 인공지능 학 회	미정			제한 없음	• Learning and Inference for AI	인공지능연구 등 재	https://aiassociation.kr/conference/conference01_1.asp? AC=0	영어
2024년 한국정보 기술학회 하계종 합학술대회 및 대 학생논문경진대회	2024년 5월 23일 (목) ~ 5월 25일 (토)	~ 2024년 4월 30 일(화)	제주	대학생 위주	ICT 관련 전 분야 및 캡스톤 디자인 작품	한국정보기술학회 지 등재	https://ki-it.or.kr/진행중인행사#	
한국 IT 서비스 학 회	미정			제한 없음	Al•데이터 기반 디 지털플랫폼정부 혁신 생태계 조성	한국IT서비스학회 지 (JITS) 등재	https://www.itservice.or.kr/intro02.html	
한국컴퓨터정보학 회	미정			제한 없음	IS, IT 응용 서비 스 등 관련 분야	한국컴퓨터정보학 회논문지 (컴퓨터 정보논문지) 등재	http://www.ksci.re.kr/sub/subMain.jsp?menuseq=4030	
2023경영정보관 련 춘계통합학술 대회*	미정			제한 없음	인공지능(AI)과 지 능형 플랫폼 및 서 비스	한국빅데이터학회 학회지 등재	https://www.kmis.or.kr/kmis/sub03/sub03_03.asp	
한국정보처리학회	미정			제한 없음	소프트 웨어 및 데 이터 공학	정보처리학회논문 지. 소프트웨어 및 데이터 공 학 (KTSDE) 등재	http://kips.or.kr/bbs/confn	영어
기술경영경제학회	미정							

*한국경영정보학회, 한국빅데이터학회, 한국인터넷전자상거래학회, 한국정보시스템학회, 한국지식경영학회

llama 2 시도 / 대시보드 구현 시도