

새로운 노트

2025.05.27 화 오후 7:17 · 44분 36초

이태수

참석자 1 00:00

일단 지금 보이시는 화면이 저번 토요일 날 멘토링 했을 때 제가 내용을 적은 건데 일단 이 부분부터 시장조사 부분에서 멘토링 멘토님이 뭔가 조언을 주셨는데 저희가 이 내용을 지금 시장 조사하는 거는 이제 여기서 마무리하고 이 내용은 버린다는 얘기는 아니고 저희 중간 PPT 발표를 만들 거잖아요. 거기에 일단 반영을 하면 좋지 않을까 생각을 해서 일단은 남겨는 놔어요.

참석자 2 00:25

이거 어떻게

참석자 1 00:26

그러니까 저희가 시장에서 PPT를 따로 만들었었잖아요. 보니까 근데 그거에다 추가하는 것보다는 이제 저희가 내일인가 내일 모레부터 이제 중간 발표 PPT를 이제 만들기 시작할 거라서 거기에 이제 반영을 하면

참석자 2 00:38

그렇게 하면 그리고 저희들이 그거 하려고 한 게 원래 중간고사에 넣으려고 중간 발표 발표

참석자 1 00:46

하려고 만든 거여서 그리고 이제 바로 다음 게 시계 데이터인데 여기서 저희가 강사님이랑 멘토링을 하면서 그리고 저희가 좀 회의를 좀 많이 하면서 생각을 해봤는데 저희가 그 가입 고객 증감률로 시계열 데이터 모델을 해서 이렇게 하

면 좋지 않을까라고 생각을 했는데 저희가 모았던 CSV를 은사님한테 보여드렸는데 너무 속된 말로 단순하다고 해야 되나 컬럼이 너무 적다고 해가지고 좀 생각을 좀 바꿔야겠다.

시계열은 최대한 나중에 미루고 이상치 예측으로 하는 게 어떨까요

참석자 2 01:22

이상치 예측 이상치 예측을 어떻게 할

참석자 1 01:25

그래서 저희가 지금 생각을 해본 게 태수 님 한번 그 자료 한번 보여주실래요
이거 여기 올리셨나 데이터에 데이터에 여기 맞나요?

네 키워드 카나시 근데 키워드도

참석자 3 01:47

진짜 단순하게 나와

참석자 1 01:48

인터넷 자체가

참석자 1 01:58

내가 이거 풀어보겠지

참석자 1 02:14

다시 오면 다 50페이지

참석자 4 02:18

네

참석자 1 02:19

이게 해당 통신사의 해당 연도에 그리고 해당 연도에 대한 기사들이랑 그리고 출판사 그리고 키워드 같은 거예요.

이게 저희가 이거 선정을 어떻게 했냐면은 저희가 기존에 가지고 있던 데이터에서 이상치라고 판단되는 기준을 저희가 나름 세워가지고 이 증감률을 평균을 낸 다음에 그 평균보다 그 표준 차가 표준 편차라고 해야 되나 평균보다 2배 이상 증가하거나 감소한 해당 통신사에 데이터들만 따로 모아놔서 이 내용들 이 시기에 해당 통신사에 대한 기사들만 따로 뽑아봤어요.

그 키워드를 추출하려고 네. 그렇게 해서 나온 게 이 데이터이긴 합니다.

참석자 2 03:10

이거는 따로 어디서 본 거예요?

참석자 3 03:14

저희

참석자 4 03:16

딥 딥 리서치 API가 네. 거기에

참석자 3 03:19

API 써가지고 네이버 뉴스만 싹 긁어 왔습니다.

참석자 2 03:23

그렇게 크롤링을

참석자 1 03:24

네 그래서 이걸 왜 했냐 하면은 저희가 나중에 해당 달에 이거는 달 기준으로 모은 거라서 해당 달에 만약에 이런 이슈들과 비슷한 키워드들이 뉴스가 나온 다 그러면은 해당 통신사는 해당 월에 이상치가 있을 수 있는 확률이 높으니까 그런 식으로 그 인사이트를 주는 방식으로 가면 되지 않겠냐 여기까지가 일단 결론이긴 했습니다.

참석자 1 03:55

사실

참석자 2 03:57

지금 여러분들이 바빠서 여기에 신경을 많이 안 쓰면 좋겠긴 한데 네 그렇게 가면 더 단순해질 것 같아요.

더 단순해지는 거 그런 뉴스가 나오면 누구나 알 수 있는

참석자 4 04:08

네 맞아요.

참석자 3 04:09

근데 저희 강사님이 말씀해 주셨던 게 기능을 심화하지 말고 그냥 이런 기능이 있다는 것만 일단 보여줘라 어차피 상용화가 안 될 거기 때문에 그래서 시계열 데이터보다는 그냥 이상치 분석이 훨씬 나올 것 같다고 말씀해 주셨거든요.

참석자 2 04:23

이건 이상치 분석이 아닌가요?

참석자 3 04:28

예상치 분석을 하면 좋겠다 했는데 이제 저희가 거기서 생각해낸 게 여기까지 밖에 안 왔어

참석자 4 04:32

247 이렇게 배우질 않아가지.

참석자 2 04:36

네 그렇죠 그렇죠 네 배경 지식이

참석자 4 04:39

데이터 분석 위주로

참석자 1 04:40

네 그래서 생각을 하는 게 여기 이슈 같은 거를 따로 모아서 이상치 이상치 키워드 다큐먼트라고 하는 그런 걸 따로 만들 생각이긴 했는데 근데 박정우가

참석자 5 04:53

시스트가 아니지

참석자 3 04:55

그 기사에서 가장 키워드 많이 나온 애들이

참석자 4 04:59

제목이 박정호만 나오는데요. 박정호

참석자 3 05:01

그러니까 근데 그게 기사별로 키워드를 따 게 아니고 월에 제일 많이 나온 걸 그냥 옆에 월마다 넣어놓은 거예요.

3월이면 저거 나오고 19년도 보면 저거 나오고

참석자 4 05:13

박정호 기자 이름 아날까요? 기자 맞나요? 저는 SK텔레콤 사장

참석자 6 05:18

사장 그래서

참석자 2 05:24

키워드는 왜 넣으신 거예요? 그러면 이거 많이 나오는 거

참석자 1 05:27

나중에 저희가 생각을 했을 때 그 기사가 나온다 하더라도 똑같은 내용이 똑같은 제목의 기사로 나중에 나오지 않을 거니까 기사 제목이 똑같이 나올 수는 없으니까 그래서 해당 기사의 내용이 있는 키워드로 그 예상치 예측을 하자라는

참석자 2 05:45

혹시 이거 모델링을 어떻게 할지 그려보실 수 있어요.

모델이 그러니까 지금 이 뉴스 데이터로 저 증감률을 예측을 할 거라는 거죠.

참석자 1 05:58

네네 이 증감률이 이상치인지 아닌지 앞으로의 증감률의 변화가 이상치인지 아닌지

참석자 2 06:06

약간 잘 이해가 안 되는데

참석자 1 06:08

어

참석자 2 06:19

여기 그러면 여기 있는 뉴스 데이터를 가지고 아까 평균을 만든 그걸 예측을 할
거라는 거예요 아니면 어떤 걸 예측을 한다는 거죠?

답지가 뭐예요?

참석자 1 06:29

이 키워드들로 이 키워드들은 아까 전에 여기 보여드린 그 이상치 애네들에 해
당하는 기사들로

참석자 2 06:38

이걸 이상치라고 이제 판단을 한 게 정확하게 어떻게 이상치예요?

이게

참석자 4 06:41

증감률을 평균을 해가지고 그 증감률 평균 낸 증감률에 2배 이상 차이 나면은
그냥 여기 이상치라고 판단하도록 일단

참석자 2 06:53

2배 이상은 어디서 나온 거예요?

참석자 1 06:55

여기 지금 보이시는

참석자 4 06:57

전월에 비해서 증감한 거를 평균을 내가지고 평균 증감치 절대값으로 했을 때 2배 이상 나오면은

참석자 2 07:08

그러면 그 두 배를 판단할 기준이 뭐예요? 기준이 왜 두 배가 넘으면 이상치가

참석자 3 07:16

그냥 저희의 판단

참석자 2 07:18

저도 이제 궁금해서 이제 여쭙보는 거였어. 왜냐하면 어떻게 이렇게 나왔는지
저도 이제 알아야 되니까

참석자 1 07:25

2배 정도 늘어

참석자 2 07:27

뭐 제 생각에 강사님 말 듣고 하시면 될 것 같은데 네 사실 이런 거를 하려면 첫 번째로는 네 이 데이터를 이제 정규 분포나 이런 걸 그려봐서 네 정확하게 그 이상치에 대한 스레숄더를 찾아야 돼요.

실제 이상치가 그렇죠 그러니까 실제 찾아야 되고 그다음에 사실 이 이상치라는 걸 찾아낼 때는 제일 중요한 게 뭐냐면 답지를 만드는 거거든요.

그러니까 이게 이상이다 아니다라고 판단을 하는 거 근데 실제 공장을 가보면

어떤 데이터가 많을까요?

이상이 많을까요? 아니면 이상이 아닌 게 많을까요?

이상이 아닌 이상 아닌 이상이 아닌 게 많죠. 그러면 이제 언밸런스드한 데이터가 생성이 되잖아요.

네 그래서 어떤 걸 많이 하나면 그거를 실제 밸런스를 맞춰도 실제 우리가 가지고 있는 정상 좋은 데이터를 다 잃어버릴 수가 있잖아요.

참석자 2 08:21

그렇기 때문에 첫 번째로는 비정상인 거보다 그다음에 이제 정상인 데이터에 집중을 하는 게 첫 번째 그리고 두 번째는 이거를 앞으로 데이터들이 이런 게 많이 들어올 건데 자꾸 학습을 시켜야 되잖아요.

그쵸 학습을 하려면 이게 정상인지 비정상인지 미리 알고 거기에 레이블을 우리가 달아야 되잖아요.

그러면 이것도 엄청나게 돈이 많이 드는 작업이고 할 수가 없다.

그래서 어떤 걸 많이 하나면 이제 오토인 코드라는 모델이 있어요.

그러니까 이게 복원하는 옛날에 이제 사운드나 이제 이미지를 복원할 때 많이 쓰는 모델인데 거기에 정상 데이터만 학습을 시켜요.

그래서 나중에 모델이 어떻게 생겼는지 보시면 이렇게 갔다가 이렇게 다시 돌아오는 거거든요.

그래서 그게 다시 원복이 됐을 때 이제 이상치가 들어오면 이게 MSD나 이런 게 내가 생각 만들어 놓은 스레시홀드보다 좀 내려갔을 때 그냥 이거를 이상이라고 해버리는 거 원래 우리는 정상만 학습이 돼 있으니까 그래서 그런 식으로 판별을 많이 하거든요.

참석자 2 09:21

그래서 지금은 뭔가 이상치를 정하는 느낌보다는

참석자 1 09:27

그냥 좀 눈에 띄는 애들만 따로 뽑은 거예요.

참석자 2 09:29

그러면 이게 지금 한 달 주기로 항상 데이터들을 학습을 시켜서 어떻게 할 거라는 건가요?

참석자 1 09:40

네 저희는 데이터 소스 자체가 일단 한 달 단위로만 저희가 구할 수 있어서

참석자 4 09:45

제가 생각한 알고리즘이 하나 있긴 한데 일단 저희가 지금 화면에 화면처럼 뉴스 같은 것을 란다로 2주 주기로 해가지고 트리오를 쥐가지고 태빌리로 주마다 이슈가 되는 것을 여기 이런 컬럼에다가 정리해서 넣을 생각이었습니다.

근데 그러면은 예를 들어서 월 초에 테밀리로 기사 같은 거 트렌드 정리한 내용 이랑 내용을 인비딩을 해가지고 지금 이것들 다 임베딩을 했을 때 비슷한 이슈가 일어난 다를 찾아서 그때 증감률을 이용을 하면은 대략 이번 달에도 어느 정도 증감률이 나올 거다라고 예측할 수 있지 않을까 싶긴 했는데 비슷한 이슈가 있으면 비슷한 증감률이 나올 거다라는 생각으로 그냥 다 임베딩 시켜놓고

참석자 2 11:00

그럼 이것을 머신러닝 모델로 만들 때 여기서 어떤 컬럼을 이용을 하려고 했어요?

참석자 4 11:06

이거는 머시닝으로 되어있는

참석자 1 11:10

애는 머신러닝이라기보다는 그냥 비즈니스 분석 세 번째 에이전트용이라서 머신러닝 그냥 이탈자 텔레콤 이탈자 예측 거기서만 사용이 되고 여기서 머신러닝은 사용되지

참석자 2 11:21

머신러닝을 꼭 사용해야 되는 게 하나 있지 않았어요

참석자 1 11:24

사용하는 건데 여기서

참석자 4 11:25

일단 저희는 고정으로 준 게 이탈자 분석은 일단 고정으로 뒤키고 시계 쪽은 저희가 필수가 아니긴 한데 일단 시열 쪽을 해보기 위해서 어떻게 해결을 해야 될까요?

참석자 1 11:43

그러니까 비즈니스 분석 에이전트에서 세 가지 분야로 나뉘는데 첫 번째는 그 멘토님이 방금 궁금해하시 카테고리들 머신러닝 예측을

참석자 2 11:52

하는 그런 게 있고 그다음에 그거는 따로 있어

참석자 1 11:54

네네 그건 따로 있고 네네 세 번째 에이전트

참석자 2 11:58

그럼 이거는 어떤 식으로 그러면 만들려고 했어요

참석자 1 12:02

이에 이슈를 예측을 하는

참석자 2 12:06

그러니까 이슈로 예측을 한다는 게 모델링을 어떻게 하려고 했던 거예요?

참석자 4 12:10

그 그냥 지금 이거 데이터 조사할 때까지는 그런 세세한 알고리즘은 생각 안 하고 무작정 일단 이슈만 조사한 거여 가지고 제가 방금 말한 벡터 인베딩 해가지고 비슷한 이슈 찾아서 증감을 이용한다는 것도 제가 그냥 방금 생각난 거지

참석자 2 12:34

예 이해했습니다. 그것도 머신러닝 모델이죠. 그것도 딥러닝 모델이니까 엠베딩 자체가 딥러닝 모델이기 때문에 다르게

참석자 4 12:41

같은 경우에는 라우팅 쪽을 머신러닝으로 해가지고 하더라고요.

에이전트 분기하는 쪽을

참석자 2 12:50

분기를 머신러닝으로 왔다고 예를 들면 분기

참석자 1 12:56

인플루언서인데 그 인플루언서가 인사하는

참석자 4 12:59

일반 질문인지 아니면 목적이 있는 질문인지를

참석자 2 13:04

아마 머신러닝으로 변경 근데 그거를 그럼 머신러닝으로 판별할 때 거의 이제 이런 들어오면 인사를 어떻게 하는지 모르겠는데 예를 들면 STT로 하잖아요. 그 텍스트가 나올 거잖아요. 그럼 그 텍스트를 머신러닝 모델로 어떻게 구축을 하려고 하면 이거를 컬럼을 만들어야 되거든요.

네 그렇죠 그럼 tfidf라든지 그런 걸 만들어야 될 건데 그냥 머신러닝으로는 거의 불가능할 것 같은데 딥러닝 엠베딩을 쓰면 뭐 말은 되긴 하는데

참석자 4 13:33

예 벡터로 만든 다음에 그거 가지고

참석자 2 13:36

딥러닝 모델을 쓴다 그렇게 굳이 할 필요가 있나

참석자 4 13:45

그래서 일단 중간중간 계속 넣어야 되니까 판다라는 얘기 이해했어요.

이 오케이 오케이

참석자 2 13:52

일단 뉴스 데이터가 있다 우리가 네 첫 번째로는

참석자 4 14:00

비슷한 이슈로 증감률하는 건 어떻게 생각하십니까?

참석자 2 14:04

비슷한 이슈로 증감률을 만든다 이슈만으로 증감률을 판단합니다.
그것도 괜찮을 것 같긴 해요.

참석자 7 14:23

잘 모르겠네요. 내가 인증을 이렇게 했던 게 있는지를 확인해 봐

참석자 2 14:35

뭔가

참석자 2 14:42

너무 빠른 거가 나올 것 같아서

참석자 1 14:46

새로운 모델 나오면 그 가입자들 고

참석자 2 14:49

모델을 떠나서 안 쓸 것 같거든요.

참석자 1 14:53

네 사실 그래서 저희가

참석자 4 14:57

아예 12월 쪽을 좀 말이 안

참석자 1 14:59

네 빼버리는 것도 일단 최후의 보드로 생각을 하긴 했습니다.

참석자 3 15:03

사실 쓴다기보다는 그거 아니었어요. 저희 강사님이 말한 게 이상치는 내가 요청할 때 나오는 게 아니고 너 큰일 났어 하면 알려

참석자 1 15:12

맞아요. 근데 그게 큰일 났어 가지고 모두가 이미 알고 있어 알고 있는 게 문제긴 한데 맞아요.

참석자 5 15:20

공사 님이 말씀하신 거 하려면 우리가 내부 데이터가 진짜 있어가지고 그 추이를 볼 수 있어야 할 수 있을 것 같아서 맞아요.

참석자 2 15:30

아니면 이게 무슨 주제인지는 모르겠는데 그것도 비슷하긴 약간 주제 분류를 하는 주제 분류 그런 걸 또 생각을 해볼 수 있을 것 같고 주제 분류요.

네 그러니까 뭐 이 안에서도 신기술이라는 주제일지 아니면 새로운 핸드폰이 나온 건지 아니면 유심 같이 어떤 특정한 이슈가 발생한 거라든지 그런 것들을 나눠 가지고 그걸 뭘 그거 가지고 뭘 할까도 생각을 해야 되고

참석자 4 16:05

뭐냐 약간 타입 기사의 타입이라고 유형을 나눠가지고 그 유형별로 그 달에 뉴스가 얼마나 분포돼 있는지 얼마나 있는지 그런 식으로

참석자 2 16:19

그럴 수도 있죠.

참석자 2 16:27

원래 우리가 이걸 뭔가 보고서를 만들려고 하지 않았어요.

그거는 1번이었나 보고서는 머신러닝이었나 카테고리

참석자 1 16:36

지금 하고 있는 에이전트의 세 번째 거 네 지금 이거 두 번째 거고 비즈니스 분석 에이전트에서 세 가지 기능이 있었는데 방금 말

참석자 2 16:44

그러니까 세 번째 첫 번째가

참석자 1 16:46

첫 번째는 카테고리

참석자 2 16:47

하브 시각화 하는 게 있었잖아요. 텍스트에서

참석자 4 16:51

네네 맞아요. 네 맞아요. 맞아요.

참석자 2 16:53

그게 첫 번째예요. 텍스트 엑스ql 같은 거

참석자 1 16:57

세 번째 에이전트

참석자 2 16:59

네 첫 번째가 그거 시계열이고 두 번째가 이거고 세 번째가 텍스트에

참석자 4 17:03

네네네 맞습니다. 네네 네네

참석자 1 17:08

첫 번째가 시계열이고 첫 번째가 카테고리 컬

참석자 2 17:12

그런 거랄까 트렌드 분석 리포터 같은 거 만드는 거예요.

이걸 가지고 그래서 분석 예 그래서 뉴스 가지고 어 여기 이 뉴스를 가지고 어떤 키워드가 많이 나왔는지 같은 것들을 만들어서 그리고 저 뉴스를 요약본들을 만들어서 그거를 이제 LLM에 넣어서 그걸로 이제 보고서 리포팅을 이번 연도에 동향 3월의 동향 4월의 동향 이런 걸 만들어주는 거가 사실 그거는 실제적으로 많이 하거든요.

참석자 4 17:43

3월 뉴스 요약 4월 뉴스 요약

참석자 2 17:46

예를 들면 4월에 SK 텔레콤 이걸 제가 제가 통신사를 잘 몰라서 마음대로 하는 거예요.

4월에 지금 SKT가 주목할 만한 어떤 동량이 뭐가 있을까 이런 텔레콤 업계에서

중요할 만한 동향이 뭐가 있을까 이런 걸 분석해줘라고 LLL에 넣을 자료들을 만드는 거죠.

그래서 우리가 미리 키워드 같은 것들을 뽑아놓는 거고 키워드도 뽑아놓고 그 다음에 이제 거기서 또 몇 개를 뽑아야 되겠지 키워드가 있을 거고

참석자 2 18:25

키워드만 뽑는 것보다 아니 키워드를 뽑고 이 뉴스들을 제목들이 있을 거죠. 제목들만 이제 내용 말고 제목들만 이렇게 잘 넣어가지고 LLL 피딩을 해주고 그걸로 이제 3월에 이슈 리포트를 만들어주고 동량 리포트를 만들어주고 이런 거 하면 괜찮을 것 같은데

참석자 2 18:54

그러면 저런 뉴스 분석이 되잖아요. 보세요. 머신러닝은 일본이 하니까 더 이상 포함시킬 필요 없고 여기서는 그렇게 동양 전망 이런 걸 파악하는 뉴스 가지고 실제로 최근에 제가 프로젝트 했던 두 개 회사에서도 동양 전망에 대한 리포트를 배치로 하루마다 이제 아침마다 받기를 원하는 애가 있어요. 그래서 아침에 딱 일어나면 임원이 딱 출근했을 때 딱 키면 내가 그걸 만들어 가지고 오늘의 리포터가 이렇게 착 나오면 그것만 딱 읽을 수도 있게 그런 것들을 만들어 주는 거죠.

참석자 2 19:36

어떻게 생각하세요? 이 부분에 대해서 저희가 다른 아이디어도 괜찮습니다. 초은 내용은 할 수 있을지

참석자 1 19:45

할 수는 있을 것 같아요.

참석자 3 19:47

저는 통신사랑 어떻게 엮을지가 그냥 그냥 그냥 분석하는 거면 괜찮을 것 같은데 통신사 진짜 그냥 임원한테 주는 용도면은 상관없을 것 같은데

참석자 4 19:56

다른 아이디어인데

참석자 2 19:59

그거는 무조건 통신사를 엮든지 아니면 그거는 뭐 어떻게 파악을 하든지 타겟 오디언스를 어떻게 정하든지 오늘의 주요 이슈를 요약해 준다든지 그런 것들로 간다든지 여러 가지로 생각을 해보면 좋을 것 같아요.

참석자 1 20:12

큐 님 아이디어

참석자 6 20:15

방금 생각한 아이디어인데 이제 저희가 화면 공유 꺼드릴게요.

맞네요. 빅카인즈 여기서 이게 저희 예시 그러니까 저희 주제에 예시로 요 사이트를 이제 주셨거든요.

참석자 2 20:36

학원에서 네

참석자 6 20:38

그래서 여기 중에서 아까 그 뉴스 제목을 기반으로 이제 뭐 하신다 했었을 때 이게 좀 생각이 났거든요.

참석자 4 20:48

뉴스 기사마다 키워드 뽑아가지고 키워

참석자 6 20:51

키워드를 뽑아서 이제 이런 식으로

참석자 2 20:54

이런 것도 할 수 있겠죠. 근데 이제 키워드도 뽑을 방법이 여러 가지가 있을 건데 이런 것들을 미리 잡을 돌려가지고 배치로 이제 뉴스가 하루에 들어오면 그걸 LLM을 피딩해서 키워드를 뽑게 할 수가 있을 거고요.

그래서 매일 하루마다 주기로 도는 아니면 이제 TF IDF 같은 이제 텀을 다 뽑아가지고 그 텀의 개수를 이제 하나씩 하나씩 세워서 거기서 많이 나오는 키워드를 뽑을 수 있을 건데 그렇게 하려면 형태소 분석이 엄청 중요해요.

내가 원하는 대로 형태에서 분석이 잘 안 될 거예요.

그래서 아마 LLM으로 이런 특정한 스트럭처 아웃풋을 만들어서 뽑아내는 게 좀 제일 편하게 할 거예요.

그래서 많이 하지 말고 이렇게 한다고 생각을 하고 이걸 매일 배치 잡이 돕니다라는 시나리오로 아침에 이걸 보고서 리포팅으로 전 사람들이 받아볼 수가 있다 그런 식으로 해서 이런 식으로 이제 뽑아볼 수 있게 이런 따로 자동으로 돌려놓는 그러면 좋죠.

참석자 2 21:57

그걸 어떻게 뽑을지는 좀 차근차근 이제 우리가 결과를 보면서 한번 생각을 해보면 좋을 것 같고

참석자 2 22:09

실제로 이제 회사들은 이런 뉴스를 이제 미국이나 이런 한국 언론사에서 크롤링을 안 하고 사야 돼요.

API를 사가지고 그거를 이제 하루에 몇만 개 몇십만 개 이렇게 받아가지고 자꾸 만들거든요.

배치 작업을 아침마다 볼 수 있게

참석자 4 22:27

AWS 마켓플레이스에 있더라고요. 기본적으로

참석자 2 22:30

그런 것도 있죠. 근데 이제 그 회사들이 하는 요건은 뭐냐 하면 자기 회사에 주된 리포트를 만들어준다는 거 이제 그런 걸로 이제 하는 데도 있고

참석자 2 22:45

그렇게 한번 생각을 한번 해보시고 가면 좋을 것 같습니다.

네 알겠습니다.

참석자 1 22:51

그다음에

참석자 2 22:53

그리고 이런 거는 회사들도 좋아할 거예요.

참석자 1 23:00

이런 시나리오들 아침마다 동해

참석자 2 23:03

그래서 할 때 그냥 대충 하지 말고 좀 많이 고민을 해보면서 해보세요.

참석자 4 23:06

네 알겠습니다. 네 저거 트렌드 시각화하는 화면은 그냥 메인에다가 박아버리죠.
구현을

참석자 1 23:11

그래도 될 것 같아 네 그래도 될 것 같아요. 방금 딕 거기 디카인즈 이렇게 딱
그 키워드 나오듯이 그리고 또 요구 사항 기능서에서 저희가 그 기능을 했던 게
파이썬 2.0 버전 3.0 버전을 바꾼다든지 자바트 파이썬 뭐 그런 걸 저희가 기능
을 했었는데 이것을 강사님한테 보여드렸더니 이것은 멘토님도 아시겠지만 실
무 교수 뛰어주신 분들은 이런 격언이 있다고 그러더라고요.
오류 안 나면 수정하지 마

참석자 4 23:41

그런 게 있다고 그래가지고

참석자 2 23:44

그거는 이제 그렇지 않는데요. 회사마다 다르고 해주면 너무 좋아하죠.

참석자 4 23:51

강사님의 경우에는 한 번 시스템 구축해 놓으면은 거기서 모듈 같은 거나 아니
면은 파이썬 버전 같은 거는 변하면 안 되는 경우가 너무 많다고 해가지고

참석자 2 24:03

최근에 저희한테 많이 들어와요. 그런 기가 예를 들면 이제 자바 코드를 파이썬
으로 바꾼다든지 사실 자바 파이썬이 아니고 네 제일 많이 하는 게 코블 투 자
바로 많이 바꿨어 코블 투 바 그런 식으로 이제 사실 엄청 기대가 많이 들어와

요.

그리고 씨를 또 자바로 바꾸는 것도 많이 들어오거든요.

그거는 엄청 큰 기업들에서 요구들도 많고 그걸 이제 요즘 코드 어시스턴트에 대한 기능에 대한 이제 관심이 엄청 높아요.

그게 거의 1번이에요. 요즘은 그게 제일 높고 그다음에 이제 이게 어디서 이렇게 많이 생기냐면 이전 이제 카카오에 있을 때 실제로 이런 이 머신러닝 어프로치 같은 걸 쓰려면 파이썬 버전이 업그레이드가 돼야 되거든요.

그러면 이제 패키지를 못 들고 오니까요. 그래서 다 바꿔야 돼요.

그래서 그런 것들도 테크 기업에서 엄청나게 많이 해요.

그래서 이거는 뭐 회사 2회 회사 바이 회사인데 아무튼 최근에 관심이 많다.

참석자 4 25:11

이번에 호딩 페이스 둘러보다가 데브 스트리 하려고 미스트랄 코딩 특화로 파인 충전기 성능 엄청 좋다고 나와 있길래

참석자 2 25:21

요즘 코덱스 이런 거 나오고 코딩은 좀 너무 편해져가지고

참석자 4 25:26

이거를 지금 제가 vallm으로 런팻에다가 서빙을 해놨거든요.

참석자 2 25:33

그러면 좋죠 몇 몇 몇 빌리언 짜리에요. 24 24 이에 이 c2에서 올려놓은 거죠.

참석자 4 25:43

런팻으로 서버리스

참석자 2 25:47

런팻에 올렸다고 바로 써버리지 그러면 얼마 정도 나와요?

그럼 오픈ai보다 싸죠

참석자 4 25:52

이거 호출 안 하면은 애가 그냥 서버가 죽어 있어 가지고

참석자 2 25:56

다른 건 없고 호출하면

참석자 4 25:59

호출하면은 단가랑 좀 비교를 해 봐. 초당 이게 0.0003달러였던가

참석자 2 26:05

그러면 ofai보다 비싸네요.

참석자 1 26:10

어제 얼마 쓰셨는지

참석자 4 26:12

뭐냐? 어제 팔 하나를 안 죽이고 그냥 자버려가지고 죽 달러스

참석자 2 26:19

그러면 그래야겠다. 사실 이것도 챗gpt 같은 거 쓰면 제일 좋고 클로드 같은 거 쓰면 좋은데 이번에는 우리가 이 VLL 같은 것도 시도를 해보고 싶고 최근에 나온 여러 가지 모델이 있어서 자원을 효율적으로 써보고 싶어서 이런 걸 만들었

다.

사실 지금 비교를 해보면 요금이 이런데 이제 이거 24비 말고 나중에 8비 1비 3비 같은 경우도 나오기 때문에 그런 경우로 이제 우리가 나중에 서빙을 하고 이러면 돈이 덜 들 거기 때문에 그런 거를 피지빌리티를 체크해 보기 위해서 이걸 했다 라고 시나리오를 짜면 좋을 것 같아요.

그래서 우리는 자체 LRM도 써봤다 이런 식으로 오픈 소스 LM도 써봤다 이런 식으로 좀 어필을 하면 좋을 것 같아요.

참석자 4 26:58

지금 임베딩 올려놓은 서버하고 VLLM 올려놓은 서버 두 개 지금 올려놓고 사용은 안 하고 있긴 한데 임베딩 서버는 형태 어휘 분석 어휘 임베딩 있지 않습니까?

스파스 그거 서빙해가지고 하이브리드로 한번 임베딩 해보려고

참석자 2 27:29

하이브리드 서치를 해보고 싶다. 하이브리드 서치를 하려면 첫 번째로 이제 벡터 밀버스에 이제 넣을 때 그럼 넣을 거잖아요.

파인 콘으로 파인 콘으로 할 거예요. 그러면 아무튼 그러면 파인 코드로 하면 그걸 넣어야 되고 두 번째는 이제 이브젝트 템 매칭을 할 수 있는 알고리즘이 필요하잖아요.

하이브리드를 하려면

참석자 4 27:55

하이브리드 공식 문서에 아예 그냥 함수 예제가

참석자 2 27:59

네 그걸로 썼어요. 그 아마 bm 투니 5 건데 어디 거 참고했어요?

혹시 그래서 bm 투니5랑 가장 많이 하거든요. 거의 다 이제 이거 하이브리드는

아마 bm 2기 파이 쓸 거예요.

제가 한번 채워볼게요. 하이코는 제가 쓰지 않아서 진짜 진짜 인터넷만 진짜

참석자 2 28:47

업서트 설치 멈췄네.

참석자 4 28:53

번가락이 익숙하네. 진짜 뭐라고 말씀하

참석자 2 29:22

BMT니바이어 맞네요. 네 이렇게 하면 돼요.

참석자 2 29:32

엠베딩은 뭐 쓰시게요

참석자 4 29:34

저거는 따로 그것도 올려서 쓰고요. 댄스는 그냥 오픈 AI 인베딩 오픈 AI

참석자 2 29:43

스파스 여긴 bm 케이블을 치는데요. 제가 하이브리드 서치 아세요?

여러분들 대강은 아닌가요? 약간 이런 거라고 생각하시면 돼요.

하이브리드 서치가 우리 벡터 검색은 알죠? 네 배우셨죠?

네 벡터 검색을 하면 최근에는 이제 나중에 혹시 찾으실 수 있으면 콜버트 같은 거 찾아보시면 좋아요.

벡터 서치가 처음에는 이제 어떤 식으로 가장 많이 발전을 했냐면 예를 들면 도쿄 도쿄의 수도 이런 식으로 해가지고 거기 답을 이제 어떻게 매핑을 하는 작업을 했었는데 사실 이것들이 엄청나게 많아지면서 이 비슷한 문맥을 가진 거 비

숫하게 또 추천을 해 줄 때가 있어요.

예를 들면 트럼프 대변인 트럼프 대변인 누구 뭐 이런 식으로 거기에 대변인과 트럼프와 대변인이 비슷한 이런 걸로 바꿔주면 돼요.

그래서 퀘스천 앤서이 나온 것보다 퀘스천을 날리면 그 퀘스천과 비슷한 퀘스천을 매핑해 줄 때가 많아요.

그러니까 그런 경우도 많거든요.

참석자 2 30:55

그래서 콜버트 같은 경우는 퀘스천 그러니까 엔서들을 많이 이제 좀 튜닝을 시켜가지고 웹 페이지 같은 거 검색할 때 그런 것들을 좀 많이 안정성 있게 하려고 하는 건데 결과적으로 어떤 것들이 나오냐면 그렇게 가다 보면 이제 좀 도메인 스페시픽한 이제 검색을 할 때 예를 들면 제조업이라든지 법률이라든지 딱 특정한 이 단어는 여기 없을 것 같은데 하는 것을 찾지를 못하거든요.

언노운 워드로 떠요. 네 그래서 언노 워드로 뜨면 그렇게 판단을 못하고 그 뒤에 문맥만 가지고 판단을 하는데 그리고 예를 들면 어떤 태양광 패널이라는 제품이 있으면 그 태양광 패널 제품에 특정 코드 번호가 있을 거예요.

그런 것들도 다 트레이닝이 안 돼 있잖아요. 그래서 이 태양광 패널 번호 누르면서 이거 얼마야 이렇게 물어보는데 없단 말이에요.

그래서 그 태양광 패널에 대한 정확한 번호가 들어오면 그 번호랑 이그젝트 매치한 문서를 끌어올리는 거예요.

참석자 2 31:52

다른 태양광 번호 네 그래서 하이브리드 서치는 방금 말씀드린 벡터 엠베딩 플러스 톰 매칭을 해주는 특정 알고리즘이랑 이제 합치는 거거든요.

사실 웹 검색의 톰 대부분은 방금 제가 부자로 말씀드렸던 톰 매칭으로만 이루어져 있어요.

그래서 우리가 네이버나 구글에 검색할 때 자연으로 검색하지 않잖아요.

그 키워드로 하잖아요. 그게 우리가 네이버 검색이나 이런 데 이제 검색 엔진에

우리가 길들여진 거죠.

여기는 자연으로 치면 안 되고 텀으로 쳐야 답이 나오구나라고 그래서 그런 텀 매칭을 하는 방식이 원래 전통적인 방식이에요.

근데 이제 벡터 검색이 나오면서 그 두 개를 합치는 거죠.

그 두 개를 이제 웨이트트 썸이라고 해가지고 벡터 검색에 나온 수치에다가 0.5를 곱하고 하이브리드 서치에 나온 거에다가 bm 2m 5 나온 거서 0.5를 곱해서 각각의 가중이 0.50 대 해서 결과를 내주고 이런 식으로 이제 하거든요.

그래서 그런 부분이 하이브리드 서치입니다.

참석자 2 32:51

그래서 이런 리콜 리콜을 올리기 위한 작업을 이제 하는 것들이라고 생각을 하시면 돼요.

텀이

참석자 4 33:01

저희가 추가적으로 타인콘에서 저희가 25달러짜리 플랜을 결제를 했는데 파인콘을요 예 AWS 마켓 플레이스에서 파인콘 연동되길래 그러니까 AWS 쪽에서 30만 원을 쓸 수가 있어가지고

참석자 2 33:18

돈 주고 했다고 그러니까 자기 돈 주고 했다 그 정도까

참석자 4 33:24

이것도 근데 뭐 타임 폰 테스트해볼까 하다가 실수로 결제한 거여서

참석자 2 33:31

저 같은 경우는 AWS 많이 썼으면 그냥 ec2 같은 데 그냥 밀버스 토크로 받아가

지고 설치해가지고 쓰는 경우가 많을 것 같아요.

2.5 버전 이상은 하이루트 서치도 다 되고 하니까 근데 이제 벡터 베딩은 여러 가지 많이 써보면 좋아요.

근데 가장 트렌드로 많이 쓰이는 게 이제 밀버스나 그 이름을 까먹었다.

그 두 개가 많이 쓰이긴 하는데 여러 가지 많이 경험이 있으면 좋겠다.

다 비슷비슷합니다.

참석자 4 33:56

이게 10억 달러 크레딧으로 이게 여기 파인 콘 자체에서 서빙해 주는 리렌코 모델이라든가 인베딩 모델을 사용할 수가 있어.

그래서 리렌코 모델을 리렌코 모델은 일단 파인콘 거 bg m3 리랭크 모델을 쓸 수 있길래 이거 한번

참석자 2 34:20

근데 그 리베카 모델이 파이콘에서 하는 게 아니고 원래 있는 거잖아요.

참석자 4 34:23

근데 여기서 소프트에서 대신 서빙해 주고 있는

참석자 2 34:26

그럼 우리가 할 필요 없다.

참석자 4 34:29

아예 그냥 라이브러리에서 모듈만 불러오면 바로 써지는

참석자 2 34:33

그냥 리랭커 아무튼 썼다고 어필은 할 수 있겠네요.

예 그렇게 하면 좋아요. 솔직히 실제적으로 지금 오픈마켓에서 가장 많이 쓰이는 이제 멀티 딥고 엠베린이 BGM 3거든요.

그래서 그게 지금 스코도 가장 높고 많이 제일 차지 그거 아니면 e5 라지 멀티 리콜이라는 걸 쓰는데 그런데 이제 컨텍스트를 많이 못 들고 가요.

한 5012 토큰 청크마다 근데 bgm3 같은 건 아마 8192 정도가 들고 갈 거예요.

참석자 2 35:04

그다음에 또 뭐 간단히 진행을 하시죠. 네

참석자 1 35:11

모델링 방안 같은 거에서는 그냥 저희 프로젝트 기획서에서 PM 제가 해놨으니
까 나머지 분들도 AI 엔지니어라든지 또 프론트 엔지니어 그건 제가 나름대로
추가를 해 봤고요.

참석자 2 35:23

여기에 대해서 피드백은 따로 없었죠

참석자 1 35:25

네네네. 여기 제출 문서에는 피드백은 따로 없었습니다.

참석자 2 35:28

간단하게 적어도 돼요. 네네. 왜냐하면 결과적으로 솔루션 아키텍처 까가지고
다 하면 저런 거 했구나 이렇게 하는 거지 여기서 아무리 많이 적어 봤자

참석자 1 35:40

그리고 또 보여드릴 게 멘토님께서 이거 시스템 아키텍처 LLM 부분이 인베딩
부분 나뉘었다면 좋겠다고 하셔가지고

참석자 2 35:51

예쁘게 잘했네. 이제 잘 보이잖아요. 그러니까 이런 걸 이런 것들을 딱 할 줄 알아야 돼.

이제 어떤 느낌인지 알겠죠. 우리가 이제 이런 거를 딱 그러라 하면 어떤 느낌인지 알겠죠.

그 이전 거는 봤을 때 너희들만 이해할 수 있는 이런 거지만 지금은 딱 봤을 때 처음 보는 사람도 이해할 수 있잖아요.

그런 느낌 좋아요. 좋아요. 지금 딱 정확하게 그냥 잘 모르는 사람이 봐도 이해할 수 있겠다 이런 것들이 느껴지는 좋네.

참석자 4 36:26

애초에 그 bgm3가 댄스 모드하고 스파스 모드 둘 다 지원을 하더라고요.

이러면 제가 오픈 AI를 굳이 써야 하나 싶어가지고

참석자 2 36:37

오픈 AI m밴딩 그리고

참석자 4 36:39

한국어 한정에서는 일단 BGM셀이랑 비슷하다고 벤치마크를 막아주고 멀티 링

참석자 2 36:48

그래도 근데 벤치마크 성능만 다 믿을 수가 없어서 왜냐하면 실제로 우리가 챗gpt 했을 때 거기에서 지금 자기들이 자체 구축한 걸로 이렇게 서치를 하잖아요.

근데 누가 봐도 들고 오는 청크나 이런 것들이 자기들이 물론 잘 짜놔겠지만 백을 잘 들고 오잖아요.

뉴스나 이런 것들도 그래서 쓸 수 있으면 글 쓰는 게 나운데 오픈 소스를 많이 쓰면 또 좋은 게 체즈 피티를 썼다.

이거는 솔직히 뭔가 해결을 하는 것보다는 그냥 뭐라고 해야 되지 그냥 너무 성의없어 보여.

그냥 챗gpt 누구나 쓰는 거 아니야 그러니까 이거는 그냥 챗gpt 사가지고 사용 해가지고 너그들 뭐 했어 II은 뭐 해봤어?

GPT 썼는데요. 아무 어필이 안 돼요. 저희들 이제 오픈 소스에서 이거 넘пат으로 다 운영도 해보고 그다음에 이런 것도 서빙도 해가 가지고 썼습니다.

참석자 2 37:39

그리고 실제적으로 이런 모델 쓸 때 성능 평가까지 해가면서 이 성능 평가 세이트를 만들어야 돼요.

데이터 세트 이런 거 쓸 때 꼭 리즈너블하게 말해야 되거든요.

성능 평가 데이터 셋을 만들어서 이게 좋다고 하지만 그래도 우리 도메인은 이게 아닐 수도 있기 때문에 한 30개 정도 만들어서 테스트를 해봤는데 이게 더 좋았습니다.

그래서 눈치 스코어나 이른바 봤을때가 더 좋았기 때문에 결과적으로 이런 모델을 뽑았고 썼습니다.

그리고 실제적으로 오픈ai에는 우리가 쓰려고 하는 궁극적인 온프레미스 같은 것도 안 되고 결과적으로 이제 나중에 비용 효율적으로 하면 이런 모델들이 작은 모델들을 쓰면 좋기 때문에 이게 더 나을 거라고 생각을 했었습니다.

엄청 리즈 분들 리즈너.

참석자 2 38:17

그래서 저는 솔직히 채치 뷰티에서 기대서 쓴다고 하면 어떤 느낌이나면 LLM 개발자보다는 뭔가 에이전트 개발자나 아니면 툴 개발자 이런 식으로 가야 되는 느낌이고 그거 아닌 툴도 근데 사실 LLM 서빙도 중요하기 때문에 챗gpt는 그냥 개인용으로는 회사용으로는 좀 아쉬운 그런 느낌이라고 생각하시면 돼요.

알겠습니다. 잘 만들었어

참석자 4 38:45

저거를 뭐냐? qn 3 32b를 일단 너 팻에 올려놔는데 저거를 로컬이라고 적는 게 좋을까 아니면 그냥 연파 VLD m으로 정리를

참석자 2 39:02

뭐가 좋으면 의미가 있을까?

참석자 7 39:08

왕팔이라고 써도 크게 팔이라고 써도 상관없을 것 같긴 한데

참석자 2 39:15

근데 뭔가 어필을 하면 이제 로컬로 하면 좋은데 그냥 림팻으로 적는 게 나을 것 같아 림팻에 우리가 서빙했다라 서빙을 했다 이런 것들을 좀 강조할 수 있는 이름으로 좀 만들면 좋을 것 같아요.

참석자 3 39:31

두 개 다 있어야 되는 거 아니야 로컬도 있어야 되는 거 아니고

참석자 2 40:30

뭐라고 그러면 좋을까? 그냥 림팻만 적어도

참석자 4 40:35

이거

참석자 4 40:50

파인콘을 쓰면서 네. 스파스 벡터가 출력이 되긴 하는데 애가 컬럼이 제대로 안 만들어져 가지고 스파스 벡터가 일단은 제가 댄스 벡터 컬럼만 해가지고 댄스 쪽만 일단 파인콘에 사내 문서들 다 올려놓고 구현을 일단 적재를 해놨고

참석자 2 41:15

이거 럼팻 하지 말고 쉼벌 온드 럼팻 이런 식으로 하세요.

우리가 막 우리가 여기 돌리고 있다 이런 걸 강조를 써보는 이런 식으로 하는 게 나을 것 같아요.

참석자 1 41:28

서빙으로 할까요?

참석자 2 41:29

서베임으로 할까요? 이거 서브 러판 돼도 상관없을 것 같아요.

서브 서브 서버 브 맞지 브 아니면 호스피드 월러 할 때 그래서 어떤 것도 상관 없을 것 같습니다.

이렇게 하면 괜찮을 것 같습니다. 그다음으로 갈

참석자 1 41:53

그다음에는 이제 그리고 회의록 다큐먼트 한번 만들어 보라고 하셔 가지고.

네. 지금 저희가 하는 멘토링을 저희가 다 녹음해서 회의록을 텍스트를 만들어 봤어요.

그래서

참석자 2 42:05

잘하셨는데 그거는 어떻게 됐어요? 그래서 시장 조사는 이번에 했어요.

아직 안 나왔어요. 네 아직 안 오케이 오케이 그것도 이제 나중에 미리 좀 해놓으셔서 해놓으시는 게 좋을 거예요.

네 맞습니다.

참석자 2 42:27

누가 제일 잘하는 것 같아요. 거기 거기 팀들 중에 팀들 중에 저희

참석자 4 42:37

우리 팀 일단 다른 팀들은 멘토님이 간섭이 적습니다.

네 그래요. 네 그런거 듣는 거 보면 그런 것 같아요.

다른 멘토님들 중에 가장 열정적이죠. 그러면 다행인데

참석자 2 42:53

좋은 건가? 네 저희는 저희한테는 너무 좋아요. 그럼 다행이네요.

네 사실 저는 이게 좀 조심스러운 게 있어서 엄청 싫어하는 사람도 있어요.

이전 첫 팀은 맨 처음에 이제 멘토 받고 약간 뭐라 해야 되지 멘토님이랑 말할 때마다 너무 힘들다.

그러니까 그게 뭐냐면 이제 우리가 모르니까 뭔가 아는 척을 해야 될지 모르겠다.

근데 거기 강의에서 다 이걸 배웠다고 떠가지고 이제 이야기를 했는데 그런 경우도 있더라고요.

그래서 저는 이제 최대한 맞추려고

참석자 4 43:30

저희도 랩 그래프 배웠었는데 다들 까먹으셔가지고 이름을

참석자 2 43:36

배웠어요. 여기서

참석자 4 43:38

스테이트 그래프 가지 로드하고 어떻게 라우팅 버 뭐냐 조건 분기하는지 그거
까지만 배우고

참석자 2 43:49

이렇게 막 구체적으로 막 엄청 나눠서 짜려고는

참석자 4 43:53

그래서 어제서부터 냉그래프 스터디 1시간씩 한

참석자 2 43:57

좋아요. 이따 보여드릴 거긴 한데 좋아요.

참석자 1 44:00

네 매일매일 일단

참석자 2 44:01

그런 거 스터디 할 때도 잘 해가지고 회의록에 넣으세요.

참석자 1 44:03

스터디 회의록 따로 따로

참석자 2 44:05

좋네 좋네.

참석자 2 44:11

혹시 근데 지금 식사 안 하셔도 돼요 아니면 뭐 들고 올까요?

어떻게 하시는 거예요? 저는 아까 뭐 먹었어요? 저도 다른 분들은 좀 먹었어.

아 배고프시면 아니 저는 괜찮아요. 저는 여러분

참석자 1 44:21

저는 괜찮긴 합니다. 평소에는 잘 안 먹

참석자 2 44:25

그러면 좀 나중에 그러면 좀 쉴까요? 일단 쉬고 한 8시 20분쯤에 일 끝나죠.

clovanote.naver.com