1장

1. 표지 장
   1. 안녕하세요, 저는 1조의 발표를 맡은 1조 부팀장 유예슬 입니다
   2. 저희 1조의 멀티캠퍼스 파이널 프로젝트 주제는,
   3. 배달앱 입점 업체를 위한 리뷰 분석 및 답변 생성 서비스 입니다.
   4. 그러면 저희 조의 발표를 시작하겠습니다
2. 목차
   1. 1조 프로젝트의 목차를 간략히 설명 드리자면
   2. **1장 오버뷰, 개요 장**에서 : 프로젝트의 개요와 배경, 저희 조에 대해서 간략히 설명 드릴 예정이고,
   3. **2장, 3장 에서는** 저희가 기획한 서비스를 설명 드릴 예정입니다.
   4. 저희 서비스는 크게
   5. 답변 생성 서비스와, 리뷰 분석 서비스로 나뉘는데요, 각 장에서 서비스 개요와 모델 선정부터 모델 고도화 및 평가 까지 순차적으로 설명드릴 예정입니다.
   6. **그리고 마지막으로 4장에서** : 시현 및 q&a 시간을 가질 예정입니다
3. 팀원 소개
   1. 저희 팀원 소개를 하겠습니다.
   2. 우선 데이터 분석을 공부했다보니 AI가 변형시켜준 저희 사진을 활용해서, 소개장을 구성을 해보았습니다
   3. 저희 팀은 총 6명으로 구성되어 있는데요, 상우님께서 팀장을 맡아 프로젝트를 이끌어 주셨고, 제가 부팀장 역할과 발표를 맡았습니다.
   4. 형준님께서는 프로젝트 관리와 서비스 1 모델링그리고 발표를,
   5. 소현님께서는 서비스 1 모델링과 발표 준비 및 시현 파트를,
   6. 용민님께서는 서비스 서비스 2와 시현 파트를 맡아주셨고, 수인님께서는 데이터 수집과 전처리 부분을 도맡아서 작업해주셨습니다.
   7. 역할 분담이 있기는 했지만 전체적으로 서비스를 기획하는 과정에서 함께 문제 해결을 해가며 프로젝트를 진행하였습니다.
4. 기획 배경 1 - 소비자 리뷰가 소비 패턴에 지대한 영향 미치는 상황 
   1. 저희가 프로젝트 주제를 정하면서 집중했던 부분은 ‘배달앱 리뷰'라는 소재입니다
   2. 현재 시장에서, 소비자 리뷰가 소비 패턴에 지대한 영향을 미치고, 기업, 특히 자영업자들에게 리뷰 관리는 선택이 아닌 필수 사항이 되었습니다.
5. 기획배경 1 - 특히 배달앱의 경우, 사장님 댓글이 매출 상승 주요 요인 **중 하나로, 리뷰 뿐이 아니라 사장님의 성의 있는 답변이 매출을 상승 시키는 주요 요인이 되었습니다.** 
   1. 오른쪽 그래프를 보시면 배달의 민족 통계에 따르면 사장님이 모든 리뷰에 댓글을 달았을 때, 오픈 4개월 차에 **재주문율이 6.5배 상승한 변화가 있었다고 하며,**
   2. 이처럼 사장님의 댓글과 리뷰에 대한 반응이 매출과 직결되고 있음을 알 수 있었습니다
6. 기획배경 1 - 그러나 이렇게 사장님 댓글의 중요성이 커질 수록, 사장님들의 어려움도 커지고 있었는데요,
   1. 사장님들은 댓글 관리에 대한 스트레스를 받고, 고객과의 충돌까지 발생하는 경우가 있었으며,
   2. 리뷰 관리에 소요되는 시간 또한 많아 **사장님들의 30% 이상이 주 5일이상 댓글 관리에 시간을 쏟거나 월 480분을 소요하는 등,**
   3. **리뷰 관리에 소요되는 시간과 인건비가 적지 않아서** 가게 운영에 집중하기 어려운 상황이라는 통계들을 확인할 수 있었습니다.
7. 기획 배경 1 - 시장 조사 :
   1. 이렇게 고객과의 소통, 댓글이 중요성을 띄는 상황에서 사장님들을 위한 고객 리뷰 관련 기능들을 확인할 수 있었는데요,
   2. 크게 리뷰에 답변을 작성하는 기능과, 리뷰에 관한 통계를 제공하는 고객관리 기능을 확인할 수 있었습니다
8. 기획 배경 1 - **시장조사 : 비교 표 페이지** 
   1. 시장에 기 출시된 서비스들을 크게 1) 배달앱의 내부 기능 서비스 2) AI 리뷰봇 3) 리뷰 관리 서비스로 나눌 수 있었고,
   2. 배달앱의 경우에는,
      1. 답변 생성 기능에 있어서 - 자동 저장 문구 기능이 매크로 식으로 성의 없는 답변이라는 고객 인식이 있었으며,
      2. 리뷰 데이터의 정량적인 분석이 주문이나 별점 기반 통계인 점에서 피드백 활용도가 떨어졌습니다.
   3. 리뷰봇과 리뷰관리 서비스는
      1. 답변은 자동으로 생성을 하지만, 특히 리뷰봇의 경우 가게명만 바뀐 동일 양식의 답변이라는 인상을 주는 경우가 있었으며,
      2. 리뷰를 분석하는 관련 통계가 아예 없거나, 통계가 있더라도 주문 및 별점 기반 정량 통계로 비즈니스적으로 활용하는 측면에서 한계가 있다고 보았습니다.
9. 기획 배경 1 -이러한 배경에서, 저희 1조는 기존의 사장님들을 위한 리뷰 관련 서비스들을 보완하여
   1. 리뷰들에 대한 자연스럽고 성의 있는 답변을 생성하고, 리뷰에 대한 토픽 관련 분석을 제공함으로써,
   2. 보다 편리하고 효과적으로 리뷰를 관리하는 서비스를 기획하고자 하였습니다.
10. 기획 배경 1 - 저희 일정 및 작업관리로는,
    1. 프로젝트를 진행하는 2달 동안, 일정 관리는 노션을 활용하였고,
    2. 작업 관리는 Github를 활용, 작업 툴들은 이와 같습니다.

2장 : 개요에 대한 설명을 마치고, 본격적인 서비스, 답변 생성 서비스에 대한 설명은 형준님께서 해주시겠습니다.

3장

1. 표지
   1. 3장 발표를 이어가도록 하겠습니다
2. overview
   1. 우선 저희의 두번째 서비스인 리뷰 분석 서비스는, 단순 통계 자료가 아니라, 리뷰 텍스트를 분석함으로써 매출 전략에 도움이 될 수 있도록 기획되었습니다
   2. 리뷰 분석 서비스는 답변 생성시에 해당 리뷰의 토픽이 quality, quantity, service로 나뉘어 도출될 수 있도록 기획하였고,
   3. 리뷰 분석의 확장 서비스는, 부가적인 서비스로
   4. 리뷰의 토픽 뿐 아니라, 긍부정 지수 판단에 기반하여 추가적인 인사이트를 제공할 수 있도록 기획되었습니다.
3. 모델 선정
   1. 리뷰의 토픽을 분석하기 위해서, 저희가 활용한 모델은 Deberta 사전학습모델입니다.
      1. 이 모델은 시퀀스 문장에 대해서 입력된 라벨들이 적합한 라벨인지 여부에 대한 스코어를 반환하며, 이 모델을 통해서 멀티 라벨 학습을 진행할 수 있었습니다.
      2. 또한 이 모델은 Deberta 모델로, 버트 기반의 자연어 처리 모델보다도 문맥의 복잡성을 더 잘 이해할 수 있다고 판단되었습니다.
   2. 더불어 MoritzLauer 모델은, Deberta 모델 중에서도, 제로샷 Classification이 가능한 모델로, 이전에 학습되지 않은 데이터들에 대해서도 분류 기능을 잘 수행하는, 일반화 성능이 좋은 모델이기에
   3. 저희 토픽 분류 과제에 적합하다고 판단하였습니다.
4. 전처리
   1. 디버타 모델을 멀티라벨 학습시키기 위해,
      1. 우선 서비스 1의 Fine tuning에 활용된 요기요 데이터의 리뷰 데이터로부터
      2. 키위로 형태소를 추출하였으며,
      3. 5개의 요식업 카테고리를 결정하여, 각 카테고리에 포함되는 키워드를 할당하는 방식으로
      4. 리뷰에 라벨링을 하였습니다.
5. 모델링
   1. 모델링 또한 멀티라벨 형식으로 fine tuning 하였는데요,
   2. 하나의 리뷰에 대해서 두개 이상의 라벨을 반환하는 멀티 라벨 모델이므로,
   3. 이를 학습시키기 위해서 원핫 인코딩처리를 하였고,
   4. 최종 학습 결과 오른쪽 예시와 같이, 맛은 있는데 양이 적어요 -> 맛, 양을 반환하였습니다
6. 성능 고도화
   1. 모델학습 완료 후에는 Fine tuning 된 모델이 과적합 되지 않도록, 학습시키지 않은 Base의 디버타 제로샷 분류 모델과의
   2. 산술평균 앙상블을 통해 고도화 하였습니다.
   3. 저희는 각 모델의 라벨을 산출시에 5개였던(맛, 양, 가격, 서비스, 배달) 라벨을 3개로 축소하였는데요,
   4. 음식의 질과 양을 두개의 필수 라벨로 두고, 나머지 가격 / 배달 / 서비스 적인 측면을 모두 service의 평균값으로 처리하여 각 모델의 라벨을 산출하였습니다.
   5. 각 모델의 세가지 라벨의 스코어를 산술 평균 한 지수가 최종 score가 되었고,
   6. 이 score들 중에서 0.35 이상인 값들을 최종 라벨로 산출하였습니다.
   7. 예시
      1. 예시로 배터질것 같다는 리뷰 문장은 ‘양'이라는 표현이 직접 포함되지 않아 까다로운 문장으로 생각되었는데,
      2. 상대적으로 과적합된 파인튜닝 모델은 이 리뷰로부터 ‘맛' 만을 잡아낸 반면에,
      3. Base 모델은 양을 포함하여 둘을 함께 활용하여 조금 더 일반화된 결과를 도출할 수 있었습니다.
7. 다음으로, 리뷰 분석 단원의 마지막인 확장서비스 입니다.
   1. 확장 서비스는 총 세가지의 기능을 가지며, 개별 리뷰가 아닌 누적된 리뷰를 활용하여, 가게에 유의미한 통계를 제공하고자 했습니다.
      1. 먼저 유사도 기반 대표 부정리뷰 분석 (기능)은,
         1. 앞선 답변 생성모델에서 활용된 유사도 평가 방법을 활용하여, 5개의 부정적인 라벨과 유사도가 높은 리뷰들이 얼마의 비중을 차지하는지 확인하도록 했습니다.
      2. 두번째로, 긍부정 지수의 변화 추이 통계는,
         1. 가게 전체 리뷰의 긍정 score과 부정 score의 누적 추이를 주차별로 확인했습니다.
         2. 긍부정 감성 분석은 1장 답변 생성 증강을 위해 활용한 리뷰관련 sentiment 분류 모델인
         3. 요식업 리뷰 예측 사전학습 모델을 활용하여 수행했고,
         4. 리뷰별 누적 score 수치를 활용하여 시간의 흐름에 따라 가게가 전반적으로 고객으로부터 받고 있는 평가를 나타내고자 했습니다
      3. 마지막으로 고객별 온도 지수 및 유형 분석 (기능) 통계는,
         1. 긍부정 지수라는 개념을 당근마켓의 온도 개념처럼 차용해서, 고객이 가게에 대해서 가지는 호의의 정도인 ‘온도'지수로 바꾸어,
         2. K means 클러스터링을 활용하여 고객의 유형을 확인하고자 하였습니다.

4. 지금까지 리뷰 분석 서비스에 대한 설명을 마쳤습니다. 이제 마지막 4 장에서 서비스 시현을 해보도록 하겠습니다

저희는 스트림릿 서비스를 활용해서, 제공되는 서비스 형태를 예시적으로 구현을 해보았습니다

* 1. 우선 답변 생성 파트입니다
  + 예시 문장을 적어보겠습니다
  + >> 맛은 있는데 양이 적어요 / 배달이 늦어서 밥이 식었어요
  + >> 입력한 리뷰 문장에 대해서, 아래와 같이 답변을 생성하는 형태이며, 마지막에 해당 리뷰에 해당하는 라벨을 산출하도록 합니다
* 2. 두번째로 확장 서비스 파트 입니다.
  + 1. 유사도 분석
    - 우선 부정 리뷰토픽 통계에서는, 5개의 부정 라벨과 리뷰데이터들 간의 유사도를 분석하여 보여주도록 하였습니다
  + 2. 긍부정지수
    - 두번째로 긍부정 지수에서는, 긍정 리뷰의 누적, 부정 리뷰의 누적
    - 3개월 기간의 긍부정 지수 변화는 현시점 가게의 긍정지수, 부정지수를 확인할 수 있도록 하였고,
    - 3개년 기간의 누적 지수 변화로 전반적인 가게에 대한 인상을 확인할 수 있도록 했습니다.
  + 3. 고객 온도
    - 세번째로 고객 온도 지수 그래프입니다.
    - 1) 고객이 어떤 온도를 가지고 있는지, 해당 온도 유형에 확인이 될 수 있도록 하였습니다.
    - 긍부정지수의 평균 + idcount
    - 2) 고객 온도 지수를 기반으로 , 고객을 유형화 하였습니다. 단순히 온도 뿐아니라 리뷰를 함께
      * K means clustering 할 때, 고객의 활성 기간,
* 2037 개의 데이터 + 385개의 id를 확인할 수 있었고, 전체 데이터는 2019년부터 -> \_\_
* 중국집 가게
* 저희가 준비한 서비스에 대한 설명은 여기까지입니다.
* 이제 1조 q& a 시간을 가지도록 하겠습니다. 감사합니다.
* 추천 서비스 -> 실제 전문성이 있어야 된다고 함 / 추천이 왜 되어쓴ㄴ지
* 나는 우리 모델은 전문가고,

—--

Q&A 나올법한 질문들

* 모델 자체에서 반환하는 기능은 없었나? score만을 반환하는 거였나? / 맞음.
* 우리는 스코어를 활용하고, 거기서 룰베이스로 몇점이상을 라벨을 반환할 것인지를 결정했음
* 디버타 모델 확인 +
* 왜 quality, quantity, service 만으로는 -> 단순 누적 통계를 당연히 낼 수 있음 + 거기에서 어떤 내용이 추출되느냐가 더 중요한데, 전체 리뷰 안에서의 내용 파악을 위해서