대사.

안녕하십니까 카드 거래 데이터 기반 고객맞춤형 자산관리서비스팀

1학기 최종발표 시작하도록 하겠습니다.

=============================================

먼저 발표 순서를 먼저 말씀 드리도록 하겠습니다.

첫번째로는 프로젝트 소개, 두번째로는 주차별로 진행한 프로젝트 내용, 그리고 마지막으로는 프로젝트 결과에 대해서 발표하겠습니다.

=============================================

프로젝트 주제를 소개하겠습니다.

4인조로 팀이 꾸려지고 저희의 공통점을 추리다보니,

모두가 금융 쪽으로 진출을 원했고 금융데이터를 다뤄보며,

핀테크 기업에 입사하고 싶은 공통점들이 있었습니다.

그래서 저희는 금융 데이터 분석 및 시각화 능력 향상과 금융서비스 기획 및 개발이라는 목표를 세웠습니다.

다음은 프로젝트 주제에 대해서 간략히 설명 드리도록 하겠습니다.

주제는 카드거래 데이터 기반 고객 맞춤형 자산관리서비스이며,

소비습관 관리를 시작으로 알뜰하게 재테크를 하고싶은 분들을

위한 자산을 내 손으로 직접 관리하는 자산관리 모바일 앱을 만들기로 했습니다.

크게 두가지의 서비스를 구현할 예정입니다.

첫번째로는 사용자 과거소비내역을 분석해서 소비내역 시각화를 제공하고 분석된 소비패턴을 기반으로 과소비 항목 체크와 알림을 전송하며 사용자가 원하는 지출관리 항목을 정해 해당 항목만 관리할 수 있는 서비스와

매달 정한 투자 금액과 과소비를 막아서 생긴 여유 자금으로 투자를 권유하며

* 현재 경기를 바탕으로 자산 분배 비율을 조정하고,
* 사용자가 관심있게 투자할 만한 분야를 추천해주는 서비스입니다.

=============================================

프로젝트 주제 선정을 위한 자료조사입니다.

짠테크라는 말이 유행할 정도로 모두들 영끌 투자를 하고 있습니다.

그렇기에 자산관리야 말로 요즘 트렌드라고 생각을 하였습니다.

또한 현재 금융 서비스에서 불편한 점들을 기반으로 기획했습니다.

=============================================

그래서 저희가 직접 자산관리 서비스에 관해서 MZ세대 대상으로 설문조사를 하였습니다.

대부분이 시중은행이 아닌 핀테크 회사들에서 이체를 하고

거래내역을 확인 하며, 자산관리를 하는 것을 알 수가 있었습니다.

저희는 2030들이 시중은행 앱을 사용하지 않고 왜 핀테크서비스을 쓰는가?

이 점에 집중을 했습니다.

수수료 없이 이체가 쉬워지며 핀테크기업 만의 특색은 약화되었지만 추가적인 서비스를 도입해 아직 많은 사용자가 핀테크 서비스를 사용하고 있습니다. 그래서 저희는 각각의 자산관리서비스에서 장점을 뽑아서 서비스를 구현해보았습니다

=============================================

다음은 프로젝트 내용입니다.

2에서 4주차 동안은 텍스트 전처리 및 데이터 시각화를 연습을 했습니다.

* Tidyverse, KoNLP, RmecanKo 등의 패키지와 형태소분석, 텍스트분석의 정량적 지표 등에 대해 학습하는 시간을 가졌고
* KoNLP 패키지를 활용하여 ppt 아래쪽 기사 본문을 보면

네이버 기사의 댓글을 긁어와서 filter를 거쳐서 보여주고 오른쪽 사진을 보면 형태소 단위로 토큰화하고 단어출현빈도, 동시출현빈도 등을 계산하여 연관분석 기법과 연관지어 적용해보았습니다.

=============================================

또한 국민은행, 우리, 농협, 카카오뱅크 카드 데이터를 수집하기 위해서 설문조사를 하여 희망자에 관해서 데이터를 수집했고,

카드사별 데이터에서 이상치들을 제공하거나 중복되는 결과를 제거한 데이터 900여개에 대해서 전처리를 진행하고 시각화 하여 소비분석 레포트에 넣을 항목을 찾았습니다.

시각화한 항목에는 월별 사용금액, 일별 사용금액, 건수별 사용 금액, 이용가맹점 분류, 시간대별 사용 금액입니다.

=============================================

텍스트 전처리 및 데이터 시각화를 연습하면서 생긴 문제로는

* 은행별로 제공하는 데이터와 형식이   
  매우 달라서 전처리 작업이 까다로웠습니다.
* 사용자에게 보일 표준화된 분석결과와 시각화 결과를 마련할 필요가 있었습니다.

또한 이용가맹점 분류의 기준이 필요했고, 텍스트 타입의 이용가맹점을 자동으로 분류하는 분석 알고리즘이 필요했습니다.

=============================================

다음은 5~9주차에 한 카드 데이터 분류 알고리즘 구현에 관한 내용입니다.

분류 기준 정의를 2030 세대에 특화된 우리만의 분류기준으로 독자적으로 정의했습니다.

2030 세대의 지출이 집중되는 배달, 온라인 쇼핑, 오프라인 쇼핑, 택시, 편의점 운동등을

독자적인 카테고리로 정의를 했으며, 이후 고정비 지출이나 구독 서비스 관리 부분도

고려했습니다.

분류 알고리즘으로는 첫 layer에서 텍스트 전처리를 통해 태그 값이 명확한 카테고리 순으로

분류를 하며, 두번째, 세번째 layer에서는 여러가지 오픈 API와 크롤링으로 긁어온 텍스트에

TF점수를 계산하여 상위 10개의 키워드를 추출하여 분류했습니다.

다만, 우리은행 홈페이지에서 받은 거래내역항목은 온전한 텍스트가 아닌 항목들이 많았기에

짧은 시간내에 보완할수 없어 83%의 낮은 정확도가 나왔다고 판단을 했습니다.

=============================================

분류의 알고리즘의 예를 들자면

아래는 ‘돈수백 강서점’이라는 거래처명을 분류해 둔 카테고리에 맞게 분석해 보겠습니다.

‘돈수백 강서점’의 경우는 명확한 태그 값이 없고, 네이버지도 검색 결과 값도 없었습니다.

이런 경우는 네이버 블로그와 네이버 웹문서에서 크롤링한 후, 형태소 분석기를 이용하였고

추출된 키워드로 식비 카테고리로 분류할 수 있었습니다

=============================================

다음으로는

저희가 해커톤 기간에 맞추어서 구현한 기술 프레임워크입니다.

고객이 안드로이드 앱을 통해 서비스에 접근하면 앱들은 데이터에 접근하기 위해 RestAPI를 호출하여 람다함수를 실행합니다.

lambda함수에서는 DB에 데이터가 추가 됨에 따라 s3를 통해 업로드 된 분석 모델을 적용시키고 이는 카테고리별로 분류 및 지출을 분석 하게 합니다. 이후 사용자는 앱에서 다양하게 가공된 데이터 볼 수 있게 됩니다.

=============================================

5~9주차 동안 진행 하였던 우리은행 온택트 해커톤에서 저희의 서비스인 위비리치 서비스를 보여드리도록 하겠습니다.

첫 화면인 MY 계좌입니다.

이 화면에서는

우리은행 OpenAPI의 금융데이터조회서비스를 사용하여

사용자의 실시간 계좌상태, 거래내역 등을 확인할 수 있습니다.

또한 여러 계좌의 잔액을 한눈에 살펴볼 수 있으며

사용자가 설정해둔 금액 이하의 잔액을 가질 경우 알림을 받게 됩니다.

=============================================

위비리치 앱 내의 캘린더 layout입니다.

직접 수집한 우리은행 거래내역 데이터를 기반으로 구축한 분류 알고리즘을 실시간 거래데이터\*에 적용하여 카테고리별로 분류된 내역을 확인할 수 있습니다.

또한 자동분류 결과가 만족스럽지 않은 경우 직접 거래 항목을 수정할 수 있으며 이는 추후 분류 알고리즘에 반영될 수 있습니다.

=============================================

위비리치 앱의 리포트 layout입니다.

주별, 월별 생성되는 소비 리포트를 확인할 수 있으며

지난 주와 비교된 이번 주의 요일별 누적 사용금액,

가장 많이 소비한 카테고리 항목, 가장 많이 지출한 요일/시간,

사용자와 소비가 비슷한 그룹 내 카테고리별 지출 비교등을 그래프와 함께 확인하여 소비패턴을 개선할 수 있습니다.

=============================================

마지막으로는 사용자가 이용하는 우리카드와 분석된 소비패턴을 비교하여 나에게 더 적합한 카드 및 기타 금융상품을 추천해주는 서비스입니다.

보유 카드의 실적과 받을 수 있는 혜택, 추천 카드로 사용자가 받을 수 있는 혜택을 보여주어 데이터에 기반한 자연스러운 금융상품 노출로 이어질 수 있도록 하였습니다.

=============================================

다음은 실제 우리은행 온택트 해커톤에 제출햇던 시연영상입니다.

계좌 정보 조회를 클릭하면 계좌에 대한 정보를 받아오며,

이용내역 조회를 클릭하면 우리은행 계좌에서 이용내역 들을 받아와서 보여줍니다.

(쉬다가)

또한 통장 잔고가 정한 기준 이하이게 된다면 알림이 뜨도록 설정을 해뒀습니다.

아래 하단바의 캘린더를 누르게 되면 해당 날짜의 거래내역이 시각화 해서 보여지고

분류알고리즘을 이용해서 분류를 해서 카테고리를 나누고 해당하는 카테고리를 가격 밑에 시각화 했습니다.

(쉬다가)

또한 카테고리를 클릭하여 혹시나 알고리즘의 오류 때문에 잘못 분류된것은 수정을 하여 고칠수가 있습니다.

다음으로는 아래 하단바에서 리포트를 누르게 되면 각 주차별로

사용한 금액과 가장 많이 소비한 카테고리, 가장 많이 지출한 시간, 사용자와 비슷한 그룹끼리의 비교 등을 시각화해서 보여주는 리포트가 있습니다.

또한 아까 수정한 카테고리가 제대로 들어간 것을 알수가 있고

사용자의 리포트를 기반으로 카드추천 서비스를 구현해줄수도 있습니다.

=============================================

우리은행 온택트 해커톤을 참여하여 서류합격을 하고 최종심사를 본 결과 금융의 전반적인 지식이 부족하여 주제 접근이 잘못되어 좋은 결과를 내지 못했고, 금융상품, 주식/부동산 등 금융경제 전반에 대한 공부가 부족하여 이번 우리은행 온택트 해커톤에서 좋은 결과를 내지 못했던 것 같습니다. 그래서 저희는 금융에 관한 공부와 시각화 표현 부분 개선을 위해 한국 부동산 데이터 시각화 경진대회를 참가하게 되었습니다.

대회의 목표로는 다음과 같습니다.

현재 자산 시장에서 우세한 한국 부동산의 움직임이 쉽고 명확하게 보이도록 시각화

부동산 데이터로부터 사회 현상을 설명하는 인사이트 발굴

외부 데이터를 적극 활용한 인사이트 발굴

파이썬 및 R의 시각화 툴을 활용한 데이터 분석 학습 입니다.

=============================================

한국 부동산 데이터에서 주택매매, 전세 가격과 혼인인구 변동의 관계를 시각화해서 살펴보았습니다. 시각화 해본 결과 혼인건수가 감소함에 따라서 주택의 수요가 늘어나고 그에 따라 매매 가격이 증가한 것을 알수가 있었습니다.

아파트 실거래가격지수와 도시 인구이동 간의 관계를 시각화 해보았습니다.

2014년부터 부동산 가격이 지속적으로 증가함에 따라 서울시의 전출 인구도 많아졌으며, 주택수요가 늘어 주택가격이 급등하는 추세가 나타난 것으로 보이며 대개 12개월 정도의 차이를 두고 상관관계가 나타납니다.

=============================================

또한 전세값이 오르는 이유에 대해서 시각화를 하고나니 금리와

주택공급수는 그래프 확인결과

전세가격에 미치는 영향이 크지 않은 것으로 판명되고 매매가격지수와 그래프가 유사한 것으로 보아 매매가격에 의해 연쇄적으로

전세값이 오르는 것으로 판단이 됩니다.

또한 이탈자금에 초점을 맞추어서 부동산을 시각화 해보았습니다.

**왼쪽 그래프는 원화예금, 저축성 예금, 정기예금, 저축예금, 정기적금 순으로 한국은행 홈페이지에서 데이터를 가져와 그래프를 그렸습니다.**

**시각화를 해봄으로서** 저축성예금 자산과의 관계에서 부동산 가격추이를 예측하기란 어려운 것을 알 수 있습니다

=============================================

초기 계획 대비 경과를 비교해 보겠습니다.

3,4월에는 초기 계획에 맞게 데이터 수집 및 전처리와 데이터분석 알고리즘에 대해서 공부를 하고 데이터 기반 시각화를 구현 하였으며,

5,6월에 구현하려고 했던 서비스를 우리은행 온택트 해커톤에 참가를 하게 되면서 그 계획을 실현할 수 있었습니다.

우리은행 해커톤이 끝나고 피드백을 하면서 알게 된 부족한 부분들을 부동산 시각화 경진대회를 나가면서 채웠습니다.

=============================================

프로젝트에 대해 아쉬운 점과 개선방향에 대해서 발표하겠습니다.

수상을 한 팀중 대상을 받은 팀은 투명하고 간편한 기능으로 아이들의 용돈 관리 서비스를 구현 하였고

최우수상 팀은 사회초년생을 대상으로 한 아르바이트 입지급 서비스를 구현하였습니다.

수상 팀들과 비교분석을 해보자면, 수상 팀들은 타깃이 명확하고 기존에 없던 참신한 서비스를 구현하였고 저희는 주제 접근 방식에 문제가 있었다고 정리했습니다.

또한 분류 정확도에 있어서 제공된 데이터가 불안정하여 정확도를 높이는데 한계가 있었고 실제로 기업에서는 어떤 식으로 분류를 하는지에 대한 조사가 더 필요한 것 같습니다.

또한 금융에 대한 전반의 지식이 부족하여 해커톤 주제를 잡는데 어려움이 있었고 앞으로 자산관리, 금융상품, 주식/부동산 등의 금융경제 전반에 대한 꾸준한 공부가 필요한 것 같습니다.

끝으로 발표를 마치겠습니다.

감사합니다.