

[1]안녕하십니까

우리에프아이에스 D&A플랫폼부 김태수 대리입니다

DAM, 빅데이터 마스터 과정을 진행하면서
익히고 만든 내용들을 소개드리려고 합니다.

[2]서순은 연수과정을 간략히 소개를 드리고,
연수 내용을 토대로 실습 해본 과제와 프로젝트를 수행한 내용을 말씀드리겠습니다.

[3]DAM 과정 소개입니다.

[4]간략히 DAM과정을 말씀드리면,
빅데이터 분산처리를 하는 하둡에서는
HDFS, 맵 리듀스, 플럼, sqoop, Hive를 학습했고 이들은 각각
하둡의 파일시스템, 병렬처리, 로그수집, ETL처리와 데이터 조회를 말합니다.

데이터 처리 엔진인 스파크에서는
RDD, SQL, 스트리밍, ML LIB을 학습했고
이들은 각각 데이터 연산, 조회, 실시간처리, 머신러닝 라이브러리입니다.

여기에 더해 실시간 데이터 스트리밍 모듈인 카프카도 학습 과정에 있었습니다.

[5] 이 내용들을 가지고 실습하며 서비스를 구현해보았는데요,
원데이 쿠킹 클래스 들었다고 식당을 차릴 순 없겠지만
실전을 통해 배운 내용을 검증해 볼 기회로써 수행하였습니다

[6]프로젝트 목표에는 여러가지가 있겠으나
하둡을 활용해 기 구축된 서비스를 저비용으로 오마주하여 구현하고,
멀티 플랫폼으로 noti피케이션을 주는 푸쉬서비스를 개발하는 것이 주가 되었습니다.

[7]여러 기사들처럼 핀테크 기업과 인터넷전문은행이 시중은행과 비대면 측면에서 대조되는 점들을 표출하고
있음을 주제 선정에 영향을 주었습니다.

[8]구축 프리뷰입니다.
시간이 많지는 않아, 시스템 구성과 오마주하였던 그룹 내 기 구축 서비스를 살펴보겠습니다.

[11]앱 리뷰를 마켓에서 게더링하고, 데이터 클렌징하여 데이터베이스에 저장합니다.

이 데이터로 텍스트를 머신러닝하여 앱 리뷰 댓글들을 클래스화 합니다.

이 정보를 다시 저장한 후 비동기적으로 웹서버를 통해 시각화하고
웹 브라우저를 통해 확인할 수 있게 서비스합니다.

그리고 주기적으로 메일 또는 메시지를 푸쉬해서 알림을 발생시킵니다.

[13]여러 번 말씀드렸던 기 구축된 그룹 내 서비스 입니다. 그래프를 제공하는 모습입니다.

[14]어떤 댓글이 있었는지 텍스트를 보여주는 기능도 있습니다.

[15]사내 메일로 1점짜리 평가에 대해 푸쉬가 옵니다.

[16]이를 오마주하여 아까 보여드렸던 시스템 구성도를 검하여 구축한 서비스를 소개드리겠습니다.

[17]웹서비스는 보고서를 기본적으로 제공하는데,
여기 url을 따라 직접 보실 수 있습니다.

사용자가 선택한 옵션에 따라 그래프를 보여주고

워드 클라우드라고 불리우는 그림 형태로 좋아하는 내용과 싫어하는 내용이 댓글에서
어떤 단어들이 많이 발생하였는지 보여줍니다.

[18]차등분석을 할 수 있습니다.

은행은 은행끼리,

카드사는 카드사끼리 다섯가지 지표의 그래프로 비교할 수 있고
비교대상이 되는 타사의 심볼형태로 워드 클라우드를 표현합니다.

[19]텍스트 분석을 동적으로 수행하는데,
각 문장이 긍/부정일지, 어떤 부서가 담당해야할지 머신 러닝과 룰 베이스로 판단합니다.

[20]멀티플랫폼 푸쉬인데, 스마트폰이나 웹 브라우저 등으로 오전 아홉시경에 푸쉬를 줍니다.
이렇게 푸쉬를 주면, 오피스 외의 공간에서도 그간의 이력까지 쉽게 확인할 수 있습니다.

[21]이러한 데이터 리니지와 배치 타임라인을 설계하여 작성된 서비스입니다.

[22]RDBMS와 하둡파일시스템간의 인터렉션도 구현하였습니다.

[23]랭크된 데이터를 푸쉬하는 프로세스도 필요해서 작성하였습니다

[24]운영자는 자원관리가 필요한데, 이러한 부분도 구성하였습니다 순서대로 서버,

[25]데이베이스,

[26]소스코드 서버버전,

[27]웹서버 로그입니다.

[28]마지막으로 ,기대효과입니다

[29]개선점을 자동 알림하고, 리뷰 이력을 조회함으로써
고객이 선호하는 서비스 수요를 예측하는 의사결정을 지원할 수 있다고 기대하고 있습니다.

[30]끝으로 웹서비스와 푸쉬 서비스 화면 시연 보여드리고 발표를 마치겠습니다
감사합니다