



PIPELINING + MODELING + VISUALIZATION

PRESENTATION

2021. 05. 15. SAT | YBIGTAY

실시간 SNS 광고 판별기

2021-1 YBIGTA CONFERENCE 주제 발제.

17기 김희진, 이학림

: Advertisement

01 INTRO

Motivation

프로젝트 주제를 선정하게 된 배경

'후기일까, 광고일까?'



02 주제

Project Topic

프로젝트 주제에 대해 설명합니다.

'실시간 SNS 광고 판별기'



03 계획

How to do?

프로젝트 계획, 대략적인 가이드라

인, 예상되는 시나리오



04 함께해요!

Welcome!

이런 분과 함께하고 싶어요!



CONTENTS



무분별한 SNS 광고, 후기일까, 광고일까?

한국소비자원 조사(2016)에 의하면,
응답자 2명 중 1명 꼴로 'SNS를 통해 하루 평균 최소 6편 이상의 광고를 본다'고 응답
14.2%는 SNS 광고로 인해 직접적인 피해를 경험했다고 응답

바이럴 마케팅(Viral Marketing)
상품 후기를 조작해 소비자를 현혹시키기도 한다.

특히, 제품 출시 주기가 짧고 트렌드가 빠르게 변하는 **화장품 업계**의 경우
빠른 확산이 가능한 SNS 마케팅을 선호한다.
대표적인 예시로, 페이스북 광고를 통해 입소문을 타면서 창업 3년 만에 650억원의 매출
을 올린 화장품 스타트업도 있다.





1. 내가 읽고 있는 이 글이 광고인지 아닌지 분류할 수 있지 않을까?
2. 광고주로부터 부탁 받은 광고는 어느 정도 내용이 비슷하지 않을까?
3. 지금 현재 SNS 광고에 판치는(?) 제품을 알고 싶다!

[Topic] 실시간 SNS 광고 판별기

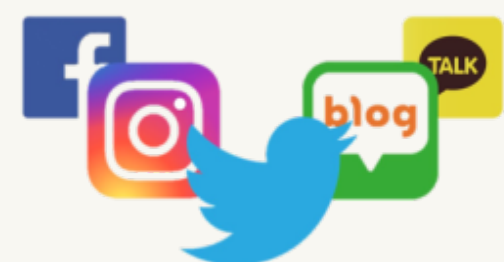
- ☒ SNS API, 크롤러로부터 게시글/사진 수집
- ☒ 실시간으로 광고글 탐지 모델링
- ☒ 일정 시간마다, 현재 SNS에 판치는 광고 제품을 랭킹으로 제공



PROJECT ARCHITECTURE

2021-1 YBIGTA CONFERENCE 주제 발제

2021. 05. 15. SAT | 실시간 SNS 광고 판별기



SNS API or SNS 크롤러
로부터 게시물/사진 수집



Spark Structured
Streaming 활용



Stream Processing

실시간으로 데이터 처리
Spark ML을 활용한 광고 판별 모델링



DB 적재

추가적으로 하고 싶은 것



airflow를 통한 워크플로 관리



광고글에 자주 등장하는 용어,
광고 사진만의 특성 분석
NLP, Vision Modeling



Batch Processing

일정 시간마다, 현재 SNS에 판치는
광고 제품을 Ranking
배치 자동화



Visualization

Tableau를 활용한 시각화

+tableau



이런 분들과 함께하고 싶습니다!

2021-1 YBIGTA CONFERENCE 주제 발제

2021. 05. 15. SAT | 실시간 SNS 광고 판별기

- ☒ YBIGTA 3팀과 모두 협업할 수 있는 프로젝트를 경험하고 싶다!
DE (Pipelining) + DS (Modeling) + DA (Analysis, Visualization)
- ☒ 스트림처리, 배치처리.. 말로만 들어봤지 둘 다 하는 프로젝트 한 번도 안해봤다! 해보고 싶다!
- ☒ (For DE) Spark, Kafka 직접 경험해보고 싶다.
Spark Structured Streaming, Spark ML, Kafka, Airflow 등 세션에서 배운 스택 활용 예정
- ☒ (For DS) SNS로부터 실시간으로 들어오는 글/사진에 대한 모델링을 해보고 싶다.
정말 기초실습인 '스팸분류기'처럼 쉬운 모델에서부터 복잡한 NLP 모델까지 다 적용할 수 있을 것 같아요!
- ☒ (For DA) Tableau 시각화에 관심이 많다. 광고 도메인 분석해보고 싶다!
- ☒ SNS 광고에 속아 돈을 날린 적이 있다..
- ☒ 멋진 사람들과 프로젝트 하고싶따^____^