

08기 정유경

3RD PROJECT; DATA ENGINEERING

각 테 일 추 천 서 비 스
웹 애 플 리 케 이 션 제 작 기



CODESTATES
AI BOOTCAMP

발표 순서

01

INTRODUCE

기획 의도 및
서비스 소개

+

02

DATA
INTRODUCE

모델링 과정
소개

=

03

PREVIEW
& DASHBOARD

서비스 시연

/

04

AND?

한계 및
보완 가능성

기획의도 및 서비스 소개

HOW WE CAN RECOMMEND
A DRINK TO CUSTOMERS?

SQL Query로 주성분 및 알코올 함량 여부에
따라 칵테일을 조회할 수 있고,
유사도 행렬을 활용한 머신러닝 모델로
칵테일을 추천받을 수도 있는
웹 어플리케이션

DATA ENGINEERING

1



USER TARGETING

- 칵테일에 관심이 있지만
- 어떻게 시작해야 할지 모르는
- BEGINNER DRINKER



DATABASE

- TheCocktailDB 데이터베이스 활용
- 유저 피드백을 수집하여 정확도 보정



SKILLS

Python for machine learning
SQLite/MongoDB for database building
html/css/javascript for front-end



RELEASE

- flask
- Heroku

추천을 위한 함수 구현

2

```
[34] def model(name, cosine_sim=cosine_sim):
      idx = indices[name] #indices 데이터프레임에서 칵테일 이름과 일치하는 인덱스 번호를 받아옴
      sim_scores = list(enumerate(cosine_sim[idx])) #유사도 점수를 코사인 유사도 방식으로 계산

      sim_scores = sorted(sim_scores, key=lambda x: x[1], reverse=True)
      sim_scores = sim_scores[1:4] #유사도 점수가 가장 높은 칵테일을 3개까지 출력

      cocktail_indices = [i[0] for i in sim_scores] #칵테일의 인덱스를 출력

      # Return the top 3 most similar movies
      return df['Name'].iloc[cocktail_indices]
```

```
[35] name = input()
```

Zippy's Revenge

```
[36] model(name, cosine_sim)
```

```
336    Kool-Aid Slammer
334    Kool First Aid
136    Berry Deadly
Name: Name, dtype: object
```

```
▶ model("Mojito", cosine_sim)
```

```
↳ 252    Frozen Mint Daiquiri
347          Limeade
114    Autodafé
Name: Name, dtype: object
```

모델링 과정 소개

① Data Wrangling

→ 총 41개의 컬럼 중 재료(Ingredient) 피처를 5개 선별

② 유사도 계산

→ 재료 피처를 모두 통합,
TF-IDF Vectorizer, linear_kernel을 활용하여
유사도 행렬을 구현

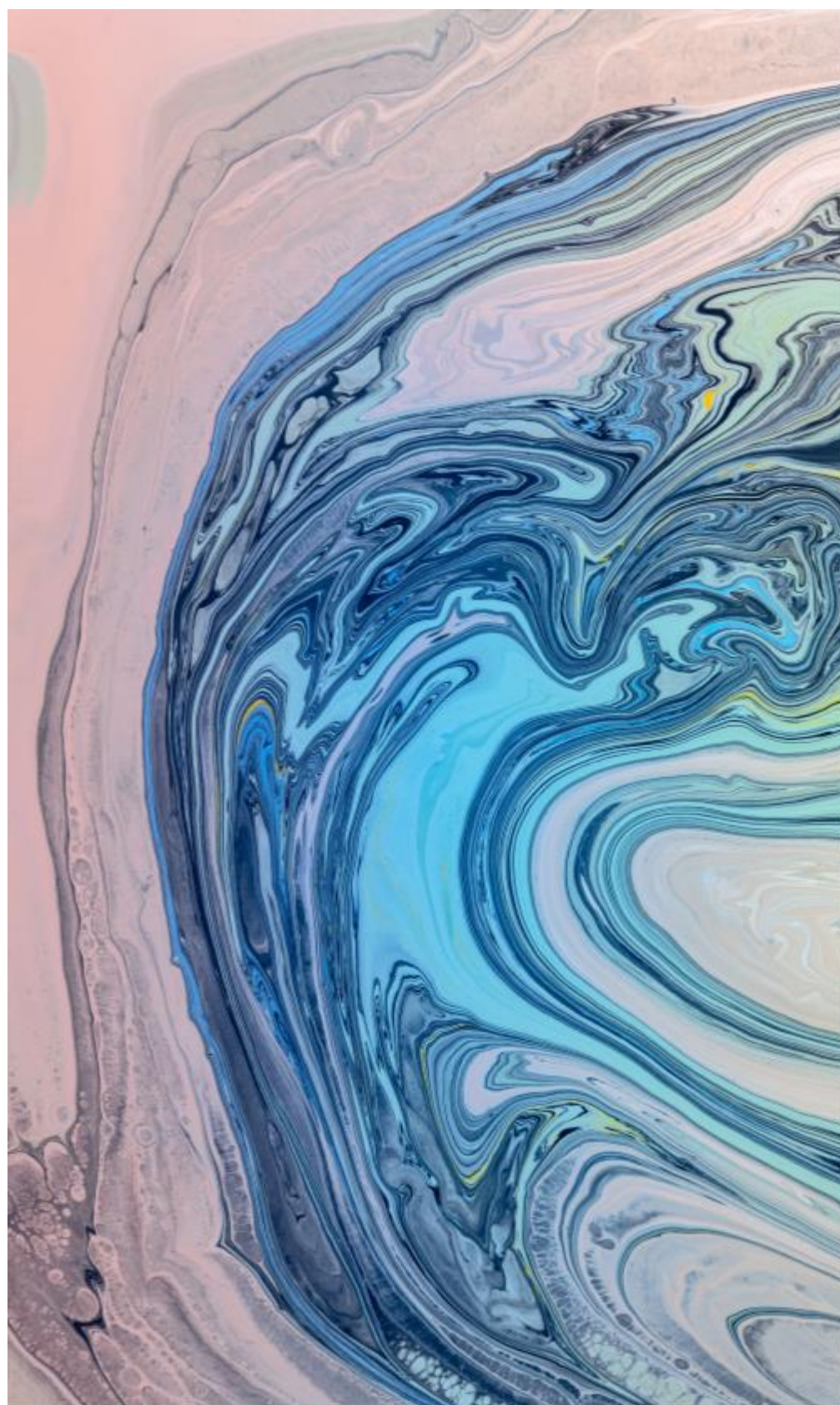
③ 모델 추출

→ 제작한 머신러닝 모델을 .pkl 형태의 파일로 부호화

한계와 보완가능성

DB 측면: API에서 자동 업데이트
공개 API가 업데이트될 때마다
자동으로 데이터를 가져오도록
애플리케이션을 설계할 수 있을까?

모델링 측면: 가중치 계산
칵테일에 포함된 재료의 비율에 따라
가중치를 다르게 부여하여
추천 정확도를 보정할 수 있을까?



YOUR NAME

THANK YOU.



CLASS NAME

YOUR NAME

CLASS OF

SPECIAL THANKS TO;

AIFIRST IN KAGGLE, THECOCKTAILDB (DATA)
MIRICANVAS, PIXABAY
(FREE IMGS & PRESANTATION FORMAT)