

메모

▼ 시각화

- seaborn

seaborn.scatterplot - seaborn 0.12.2 documentation


Draw a scatter plot with possibility of several semantic groupings. The relationship between x and y can be shown for different subsets of the data using the hue , size , and style parameters. These parameters control what visual semantics are used to identify the different subsets.

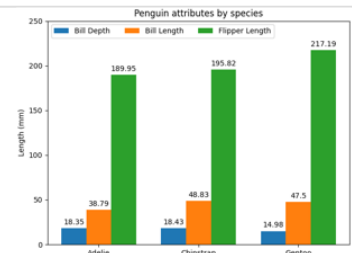
 <https://seaborn.pydata.org/generated/seaborn.scatterplot.html>

- matplotlib

Lines, bars and markers - Matplotlib 3.6.3 documentation

© Copyright 2002-2012 John Hunter, Darren Dale, Eric Firing,
Michael Droettboom and the Matplotlib development team;
2012-2023 The Matplotlib development team.

 https://matplotlib.org/stable/gallery/lines_bars_and_markers/index.html



사용자의 신체 정보에 잘 어울리는 데이터를 추천

개인정보보호관련에서 신경쓰면 좋을 것 같다.

▼ 프로젝트 배경/필요성

"인스타그램, 주요 쇼핑 플랫폼으로 변모" - ZDNet korea

@배지환[학부재학 / 산업경영공학부] : 유저 프로필을 재구성한다.

1. 추가 정보는 어떤 것을 어떻게 활용할 것이며,

2. 데이터 크롤링을 어떻게 할 것인지

3. 해시태그

해시태그가 없는 피드에 대해서 자동적으로 해시태그를 생성

이미지에서 키워드 추출..

남/녀 코디라는 해쉬태그로 인플루언서를 1차적으로 검색할 예정이기에 성별은 다른 과정 없이 확인 가능

한계점

@서재원[학부재학 / 산업경영공학부] : 이미지 인식 모델이 아쉽다. 다만 학습 양상을 봤을 때

1. 더 학습을 깊게 많이 시킬 수 있다면
2. 더 많고 다양한 데이터를 사용한다면

성능을 더욱 끌어올릴 수 있을 듯 싶다.

@신건우[학부재학 / 산업경영공학부] : 크롤링의 성능이 아쉽다. 더 많은 데이터를 효율적이고 안정적으로 수집할 수 있었다면, 프로젝트의 진행 속도를 더 빠르게 하고, 정확도 또한 좋아질 수 있었을 것이다.

@배지환[학부재학 / 산업경영공학부] : 모든 유저들이 균일한 피드를 올리지 않아 일반화를 했을 때 학습능력이 떨어진다.