

# Trip.com 기반 여행지 추천 시스템

여기 어때

윤희승, 이성찬, 이영호, 최명수

# 아이디어

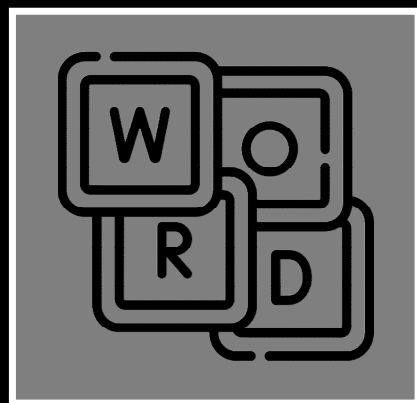


Trip.com의 명소 리뷰를 토대로  
사용자가 입력한 단어나 문장에 관련된 여행지를 추천하자

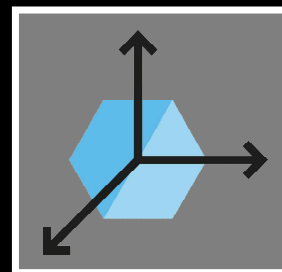
# 작업 순서



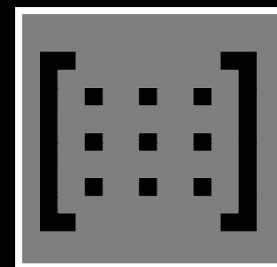
1. 주소 및 리뷰 크롤링



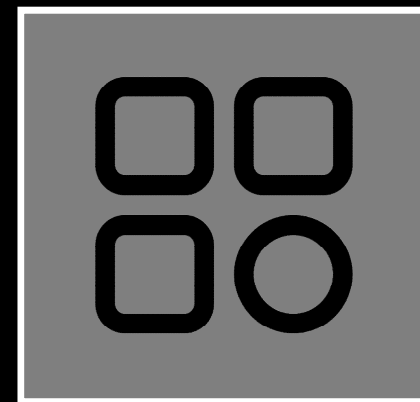
2. 데이터 전처리



3-1. Word2Vec 학습

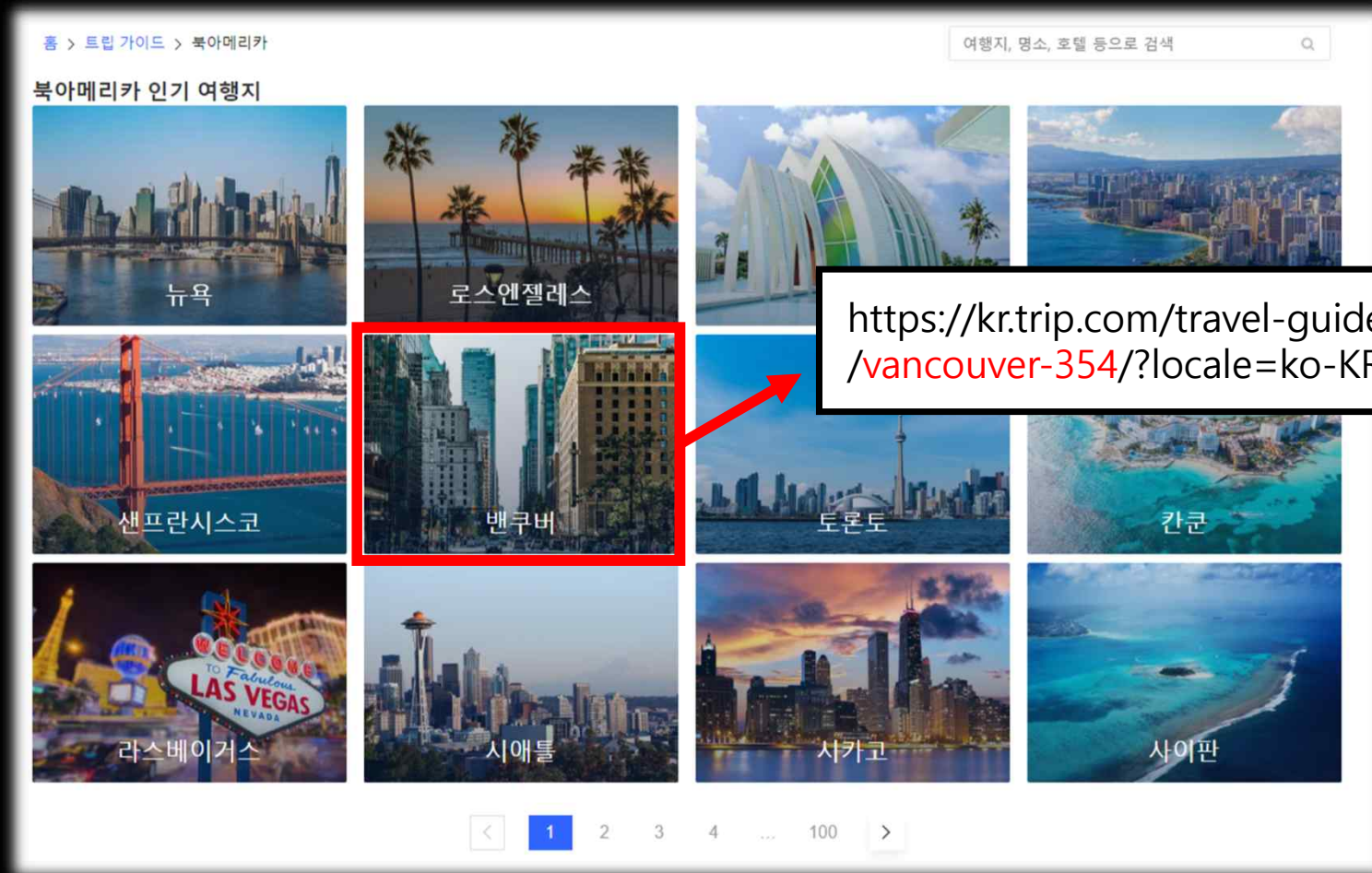


3-2. TFIDF 행렬 구축



4. UI 구현

# 1. 주소 및 리뷰 크롤링



# 1. 주소 및 리뷰 크롤링

<https://kr.trip.com/travel-guide/attraction/vancouver-354/tourist-attractions/comment-count-1011-100/?locale=ko-KR&curr=KRW>

캐나다 > 브리티시컬럼비아 > 그레이터 밴쿠버 A > 밴쿠버 > 밴쿠버명소

## 11월의 밴쿠버에서 가장 인기 있는 체험/액티비티(2023년에 업데이트됨)

현지 검색 결과만 보기 무료 입장 숨 전체 삭제

추천

여행자 평점

인기

판매량 (높음→낮음)

거리 (가까움→먼 거리)



### 1. 스탠리 공원

9.7 4.8/5 2,054 리뷰 2.9만건의 리뷰

Trip Best 1위 - 그레이터 밴쿠버 A 가볼만한곳 순위 도시 공원 인기 명소 국립공원

Syed Naushad "캐나다의 스탠리 공원은 매력적인 자연 오아시스입니다. 무성한 녹지, 깨끗한 호수, 우뚝 솟은 나무는 변화한 도시 밴쿠버에서 탈출할 수 있습니다. Seawall은 바다의 숨막히는 전경을 제공하..."

시내까지 거리: 5.3km

무료



### 2. 캐필라노 현수교

8.7 4.6/5 405 리뷰 1만건의 리뷰

Trip Best 1위 - 노스밴쿠버 가볼만한곳 순위 다리 하이킹

Lifecast "Capilano Suspension Bridge Park는 한 번 방문하기에 좋습니다. 잘 조직된 공원이며 시내에서 무료 셔틀 버스가 있습니다. 입장료는 조금 더 높습니다."

리뷰가 100개 이상 있는 명소만 검색

## 4.8/5 완벽해요! 2084건의 리뷰

모두 보기 (2,084)

최신순

사진 (1,349)

트립닷컴 인증 리뷰 (865)

공정적 (1,935)

부정적 (24)



### 5. 완벽해요!

매우 충격적인, 박물관은 3시간 동안 매우 크다, 많은 장소가 보지 않았다, 전시품은 매우 아름답다, 일부 문화 유물 복원에 아주 좋다, 박물관의 장면 배치와 디자인도 매우 아름답다, 제한된 전시 마네펀 / 대가 원래 방문자에게 두 사람의 그림을 함께 넣었다. 비교



작성일: 2023년 10월 31일



### 5. 완벽해요!

메트로폴리탄 미술관은 뉴욕시의 절대 보석입니다. 역사, 문화, 종교에 대한 열정을 지닌 문화 애호가로서 이 박물관은 지식과 영감의 천국입니다. 더 나은 점은 뉴욕거리에 무료라는 것입니다. 지역 주민들이 전 세계의 공대한 예술품과 유물을 탐험하도록 장려하는 환상적인 특전입니다. 신성한품을 겹다 보면 고대 문명에서 현대 걸작에 이르기까지 인간 창의성의 테피스트리에 빠져들게 될 것입니다. 박물관의 방대한 컬렉션에는 수세기, 문화 및 대륙에 걸쳐 있는 멋진 예술 작품, 조각품 및 역사적 유물이 포함됩니다. 전시회는 세심하게 선별되어...

모두 보기



## 2. 데이터 전처리

볼 것이 너무 많습니다

okt.pos() ↓ 형태소 분리

[['볼', 'Noun'], ['것', 'Noun'], ['이', 'Josa'], ['너무', 'Adverb'], ['많다', 'Adjective']]

↓ 불필요한 품사, 불용어, 한 글자 형태소 제거

많다

## 3-1. Word2Vec 학습

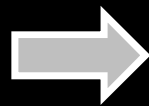
5번 이상 등장하는 단어만 취급

```
Word2Vec(tokens, vector_size=100, window=4, min_count=5, workers=12, epochs=100, sg=1)
```

[[ '많다', '당신', ... , '좋다'], ... , ['호바트', ... , '아름답다'] ]

각 리뷰 단어들을 모아놓은 리스트

작업 프로세스(Thread) 수



사용자가 입력한 문장/단어를  
유사한 단어들로 바꾸는 데 사용

## 3-2. TFIDF 행렬 구축

```
tfidf_matrix = tfidf_vectorizer.fit_transform(df['review'])
```

전처리 된 리뷰 Series

단어 (28206개)

리뷰  
(1201개)

|       |       |     |       |
|-------|-------|-----|-------|
| 0.029 | 0.027 | ... | 0.098 |
| 0.017 | 0.019 | ... | 0.027 |
| ...   | ...   | ... | ...   |
| 0.034 | 0.057 | ... | 0.015 |



cosine 유사도  
측정에 활용



## 4. UI 구현

The image shows a Windows-style window titled "Form" with a light gray background. It contains three main components: a dropdown menu on the left, a text input field in the center, and a button on the right. The dropdown menu is open, showing a list of countries in Korean. The text input field has a placeholder text. The button is labeled "추천!". Annotations with red circled numbers 1, 2, and 3 point to these elements respectively. A large black-bordered rectangle is positioned below the input field, with an arrow pointing to it from the text "나라 선택 후 검색 or 모든 나라에서 검색". Another arrow points from the button to the same rectangle. The text "관련된 명소 or 에러 메시지 출력" is centered within the rectangle.

Form

① 모든 나라

② 단어 혹은 문장을 입력하세요.

③ 추천!

나라 선택 후 검색 or 모든 나라에서 검색

관련된 명소 or 에러 메시지 출력

## 결과 - 1

모든 나라 ▾

공포

추천!

체코 / 쿠트나호라 / 세들레츠 납골당  
태국 / 방콕 / 퀴사오밥하기념연구소  
베트남 / 호치민시티 / 호치민전쟁박물관  
미국 / 피치 스프링스 / 그랜드캐니언스카이워크  
캐나다 / 노스밴쿠버 / 캐필라노 현수교  
미국 / 워싱턴 / 한국전 참전 용사 기념비  
일본 / 오사카 / 유니버설스튜디오재팬  
대한민국 / 제주 / 믿거나말거나박물관  
독일 / 베를린 / 학살된 유럽 유대인을 위한 기념물  
영국 / 런던 / 마담투소 밀랍인형관

## 결과 - 2

모든 나라 ▾

비빔밥

추천!

중국 / 타이베이 / 융캉제

중국 / 타이베이 / 닝샤야시장

대한민국 / 서울 / 광장시장

중국 / 타이베이 / 스린야시장

중국 / 타이베이 / 라오허제야시장

중국 / 타이베이 / 사대야시장

대한민국 / 서울 / 한국의집

중국 / 타이베이 / 리우허야시장

태국 / 방콕 / 실롬야시장

영국 / 런던 / 차이나타운

감사합니다