

김현우 김조현 박주연 송치영 정다빈



E.T.
The Europe - Travel Helper



1. Introduction

- 01 E.T. 소개
- 02 알고리즘 소개

2. 데이터 살펴 보기

3. loc2vec 만들기

- 01 Insight
- 02 모델 디자인

4. 결과

5. UI/UX workflow

- 01 Workflow
- 02 시연

6. Q & A

1

Introduction



주제 소개
E . T .

나만의 여행지를 추천해주는 **유럽 여행 도우미**



어디로 가야 할까?



여기로 가!





2

Loc2vec 만들
기

Insight



데이터수집

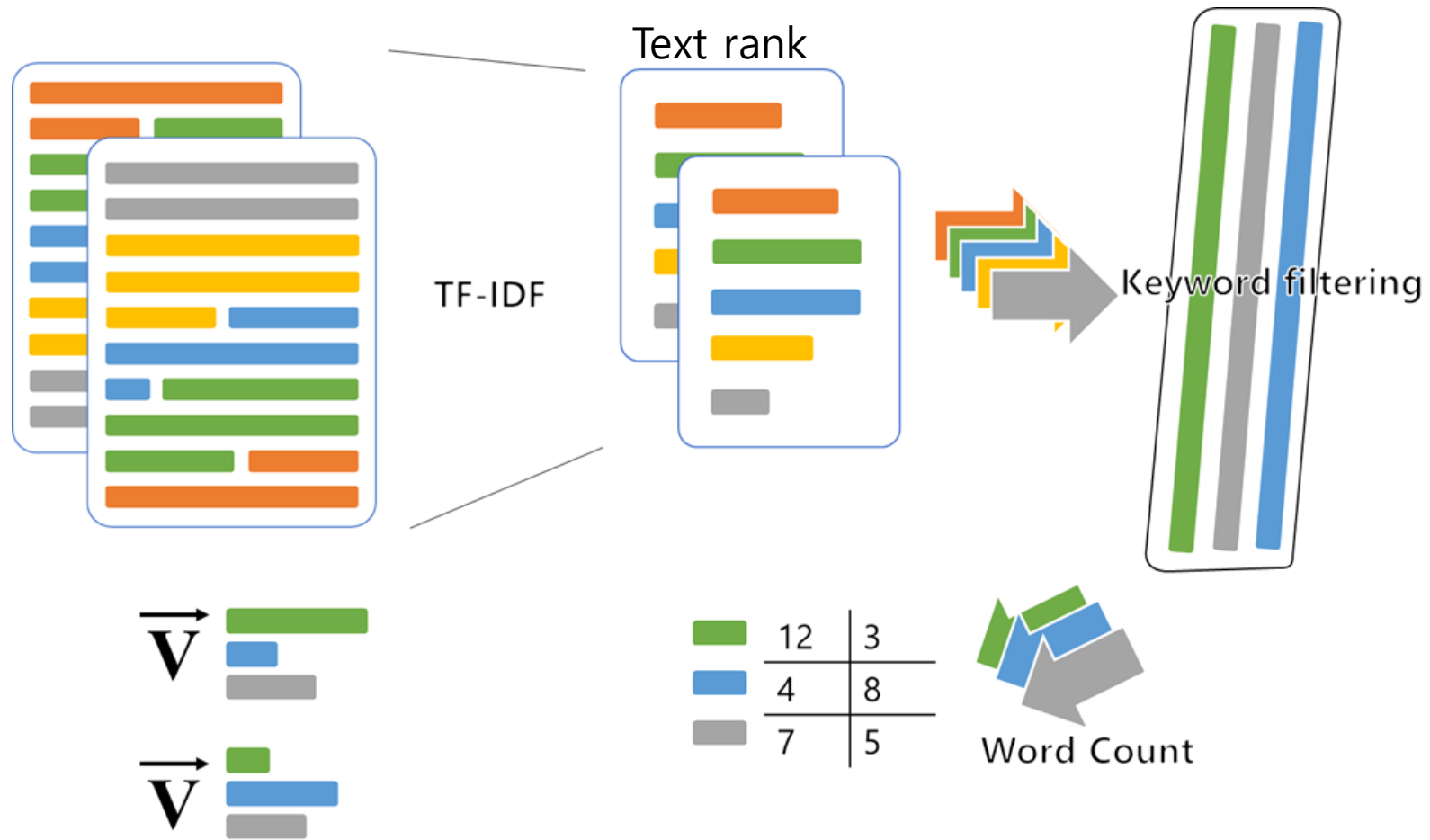
W I K I

- Wikipedia에서 스페인, 포르투갈 관광지 문서 수집
- 문서들을 바탕으로 관광지들간의 관계를 파악해 벡터화 (loc2vec)
- 추가적으로 위치와 거리 정보를 활용해 다음 관광지 추천

	place	nature	city	star	review	short_info	long_info	f_wiki	s_wiki
0	Buen Retiro Park	스페인	Madrid	4.6	24568	Vast 19th-century park with fountains	Expansive, 19th-century park with boating lake...	The Buen Retiro Park (Spanish: Parque del Buen...	The Buen Retiro Park (Spanish: Parque del Buen...
1	Royal Palace of Madrid	스페인	Madrid	4.4	6566	Regal 2,000-room palace, armory & garden	18th-century, ridge-top palace for state occas...	The Royal Palace of Madrid (Spanish: Palacio R...	The Royal Palace of Madrid (Spanish: Palacio R...
2	Plaza Mayor, Madrid	스페인	Madrid	4.3	17375	Madrid's vibrant main square	Cafes & restaurants line the arches of thi...	The Plaza Mayor (English Main Square) was buil...	The Plaza Mayor (English Main Square) was buil...
3	Museo Nacional Del Prado	스페인	Madrid	4.6	13871	World-class European art collection	Art museum with Velazquez, Goya and El Greco m...	The Prado Museum (Spanish pronunciation: [mu? s...	The Prado Museum (Spanish pronunciation: [mu? s...
4	Puerta del Sol	스페인	Madrid	4.3	26932	Vast pedestrianized public square	Public square with an equine statue of King Ca...	The Puerta del Sol (Spanish for "Gate of the S...	The Puerta del Sol (Spanish for "Gate of the S...

w o r k f l o w

1st trial

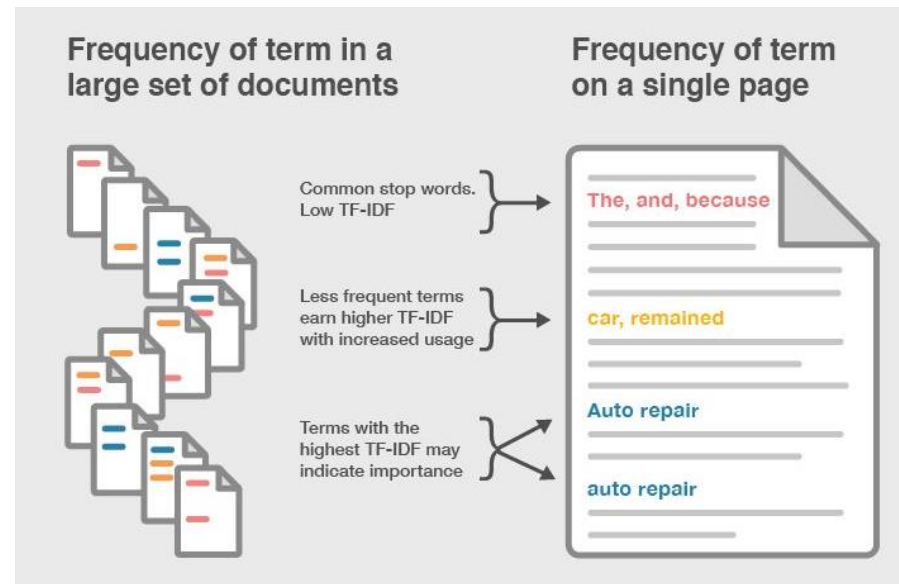


1 TF-IDF

문서 내에서 단어의 중요도를 파악하자

- 중요도가 높은 단어는 문장 내에서 자주 등장한다
- 한 단어가 문서 전체에 계속 등장하면 중요도가 떨어진다

→ 모든 문서에 흔하게 나타나는 단어는 제외시킴



2 Text Rank

Statistic

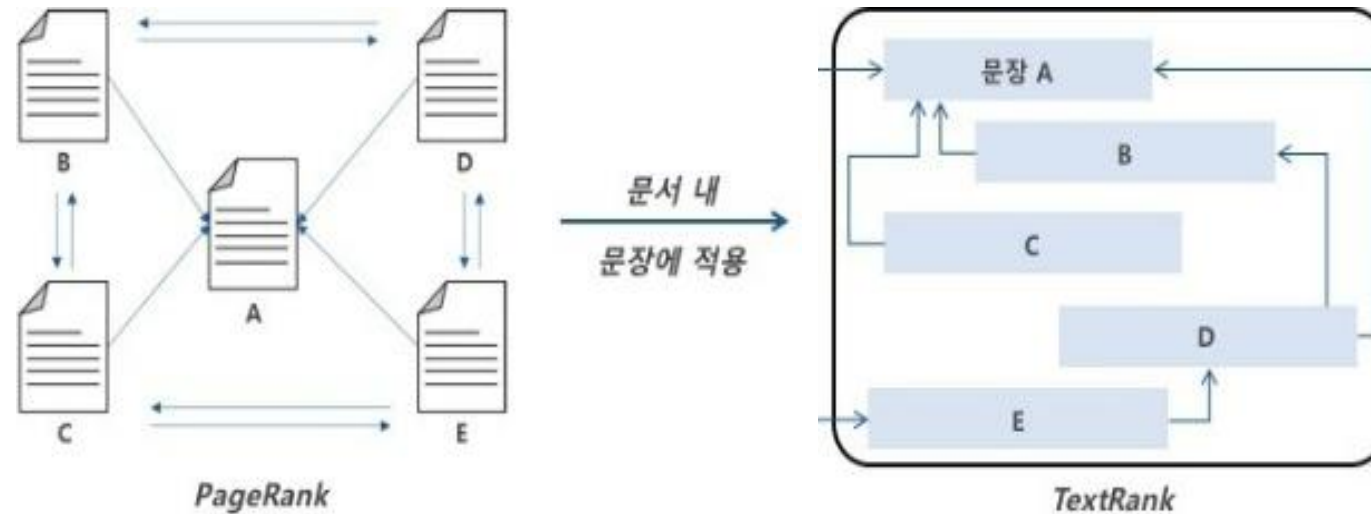
S

문서 내에서 키워드들을 뽑아내보자

- 중요도가 높은 단어는 여러 문장과 link되어 있다

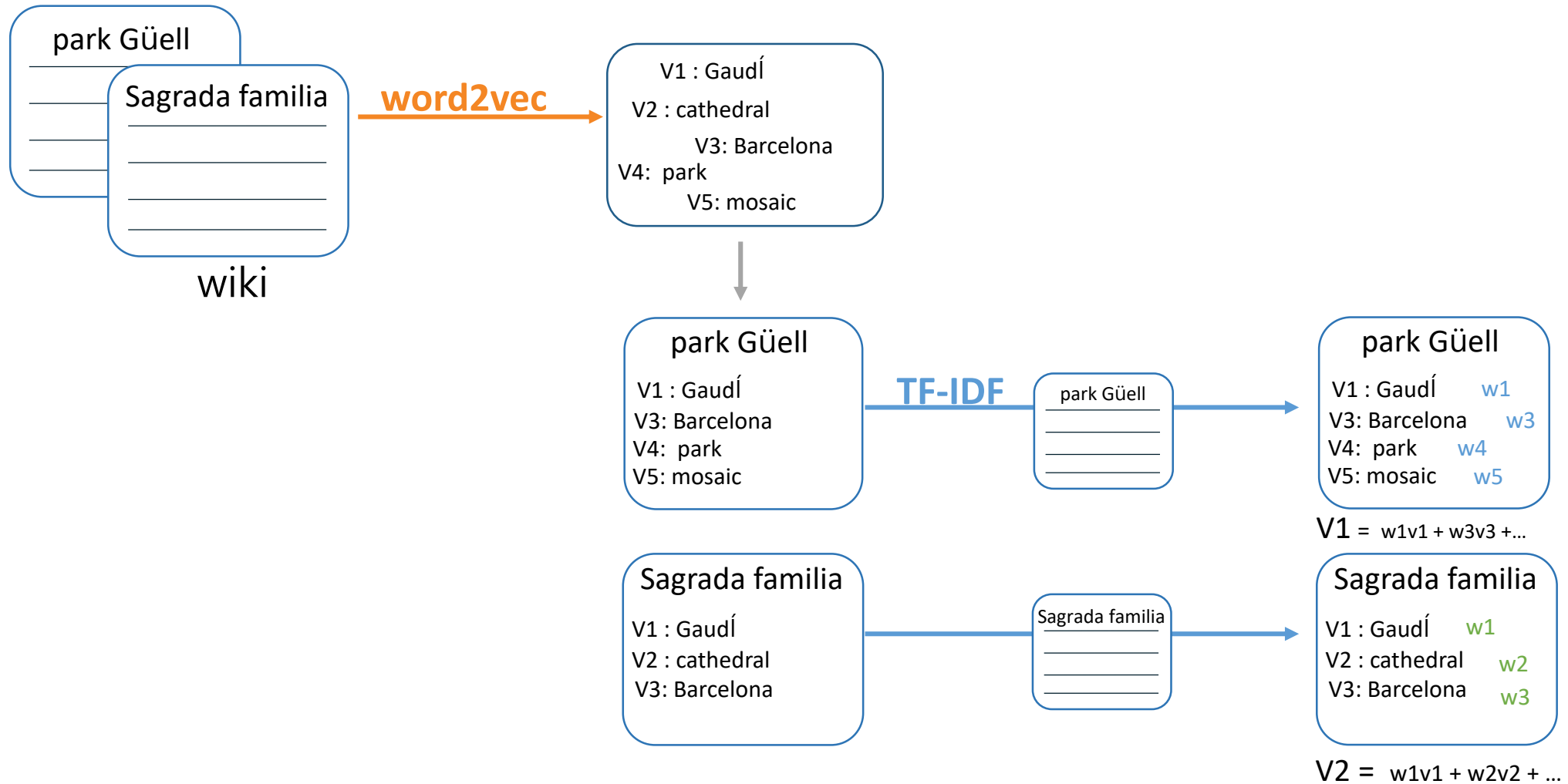
$$TR(V_i) = (1 - d) + d * \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} TR(V_j)$$

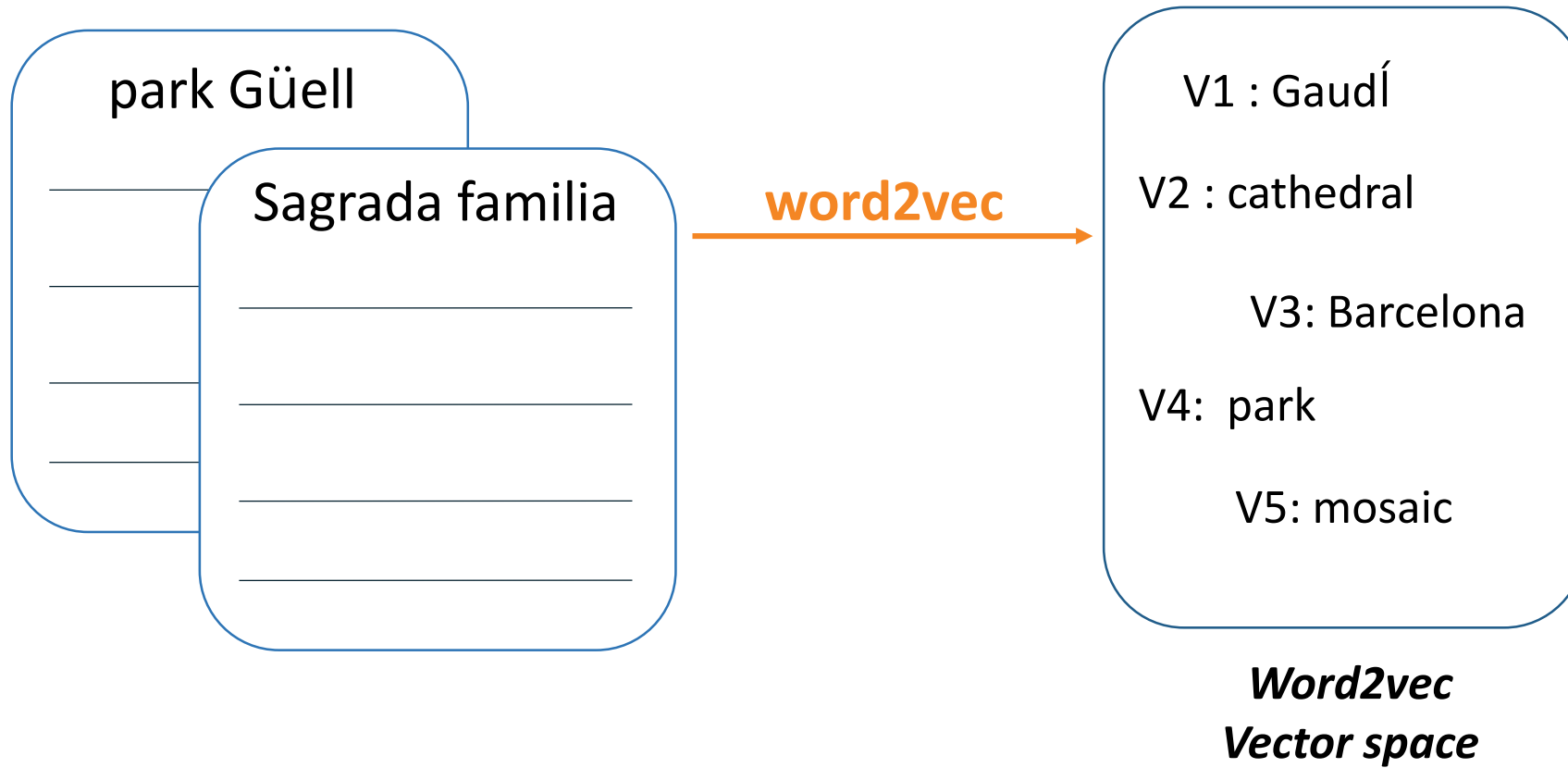
Wji: 단어 i와 j 사이의 가중치 ←TF-IDF



w o r k f l o w

2nd trial





1 word2vec

단어가 가지는 의미 자체를 벡터로 나타내자

- 같은 문장에 등장하는 단어는 의미가 비슷하다
- $\text{man} - \text{woman} + \text{king} = \text{queen}$

Input:
one document

Lorem ipsum dolor
sit amet, consete-
tur sadipscing elitr,
sed diam nonumy
eirmod tempor
invidunt ut labore
et dolore magna
aliquyam erat, sed
diam voluptua. At
vero eos et

word
vectors
↓

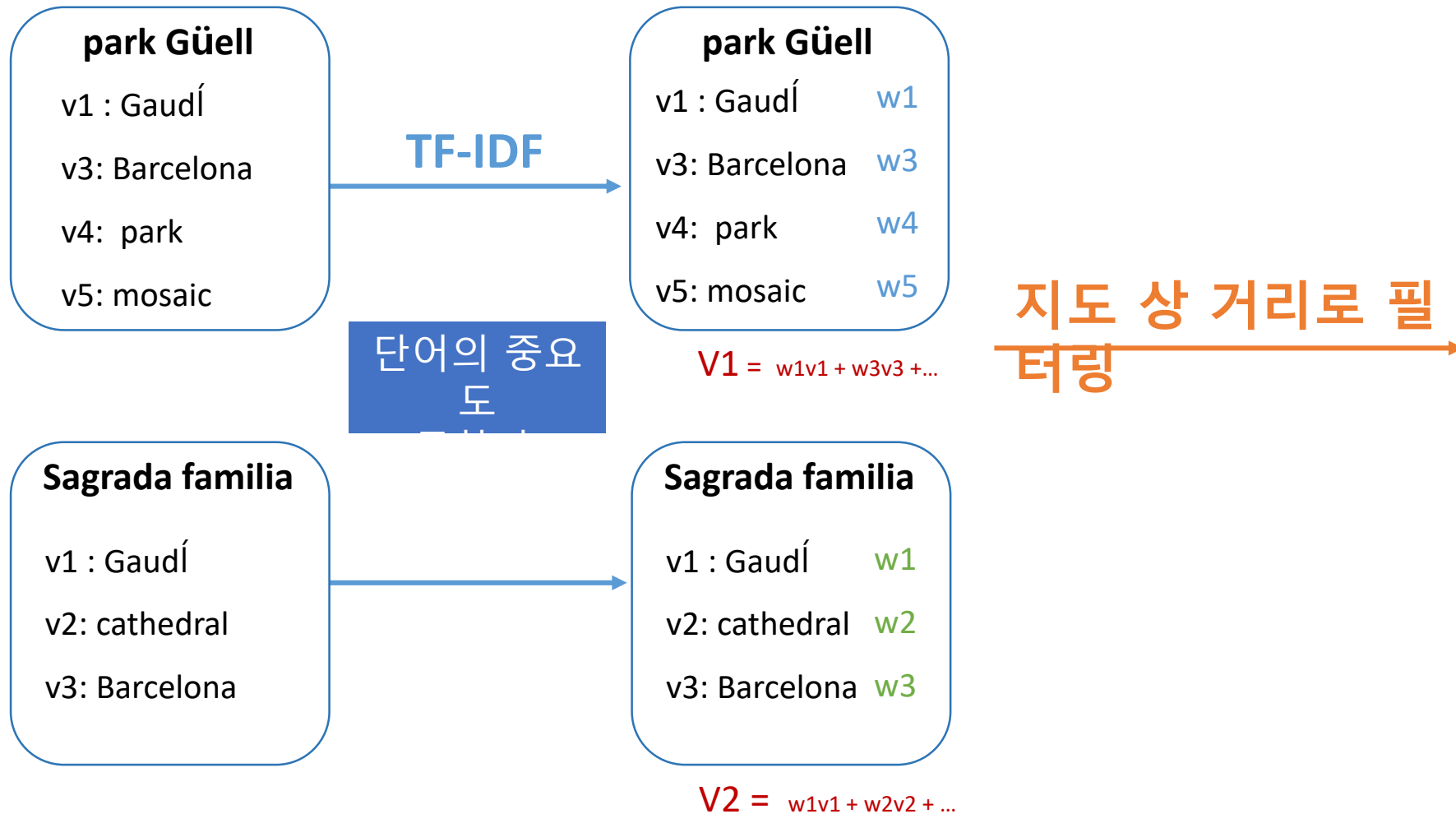
Model:



most_similar('france'):

spain	0.678515
belgium	0.665923
netherlands	0.652428
italy	0.633130

highest cosine
distance values
in vector space
of the nearest
words



4

여행지 추천하기

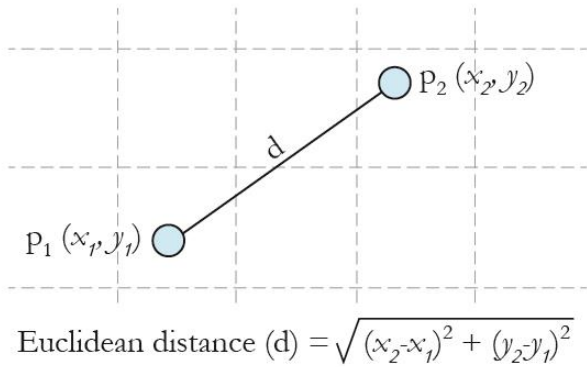


LOC2vec

Loc2vec 벡터 공간 안에서 거리가 가까울 수록 유사한 관광지



Euclidian distance



Euclidian distance를 이용해 query에서 가장 가까운 관광지 5개를 제시



Garden, park, palace

Parque de el Capricho



Madrid, gardens, park

Madrid Rio



Reasons, line, location

Plaza de Bib-Rambla



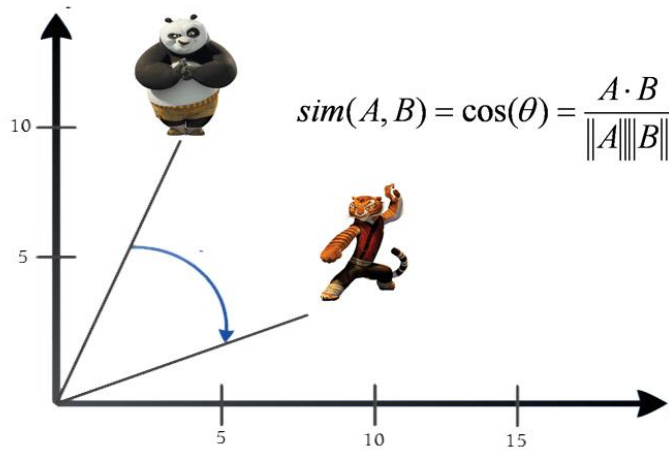
Building, line, city

0.17022871107400933

0.17460343664983302

0.18351190741545206

Cosine similarity



Cosine similarity를 이용해 query에서 가장 가까운 관광지 5개를 제시

Buen Retiro park



Garden, park, palace

Parque de el Capricho



Madrid, gardens, park

Parque de Atracciones de Madrid



Park, attractions, zone

Plaza de Espana



Plaza, espana, parque

0.46313844006755767

0.38834654597694782

0.314605153287057

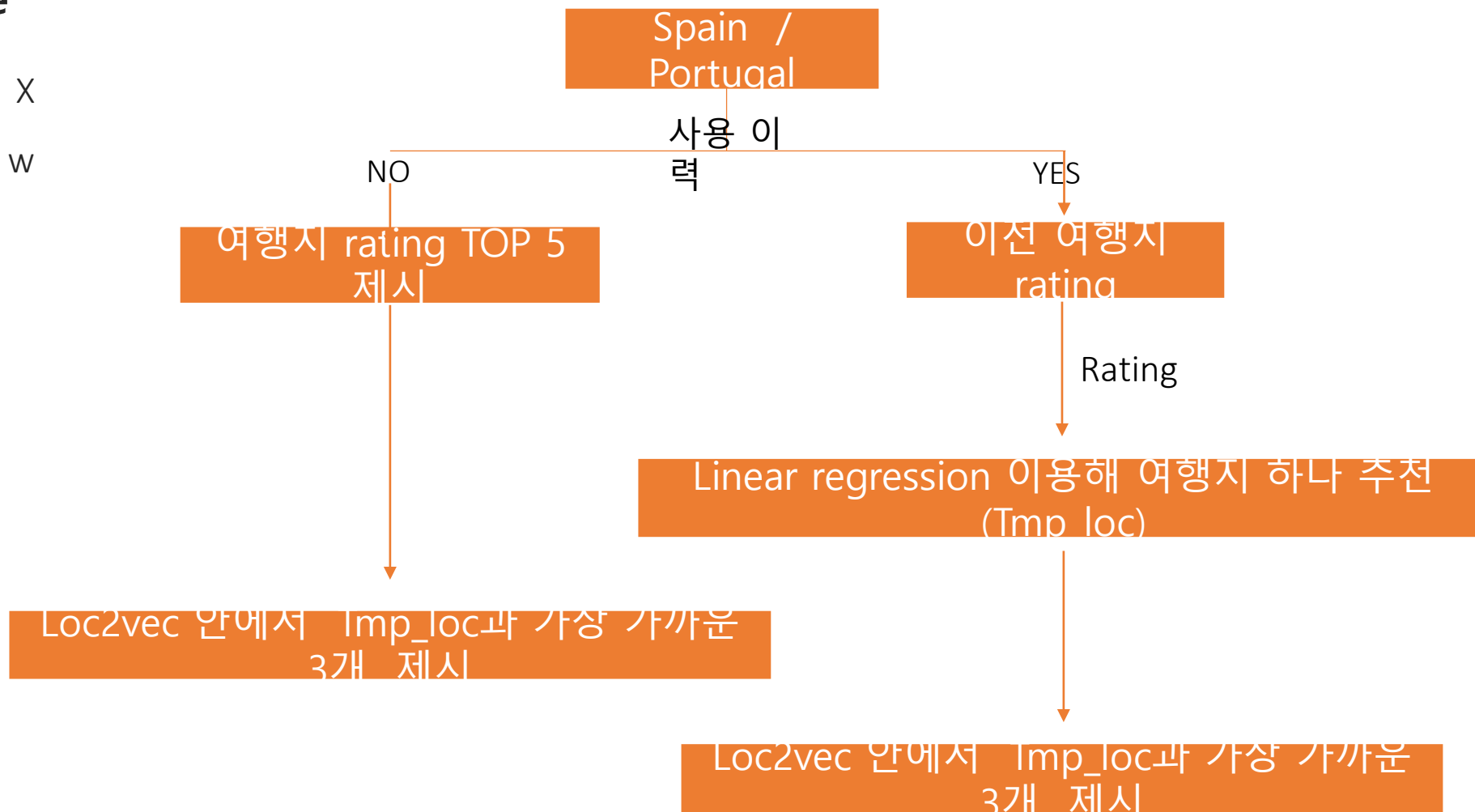
5

UI/UX workflow



service

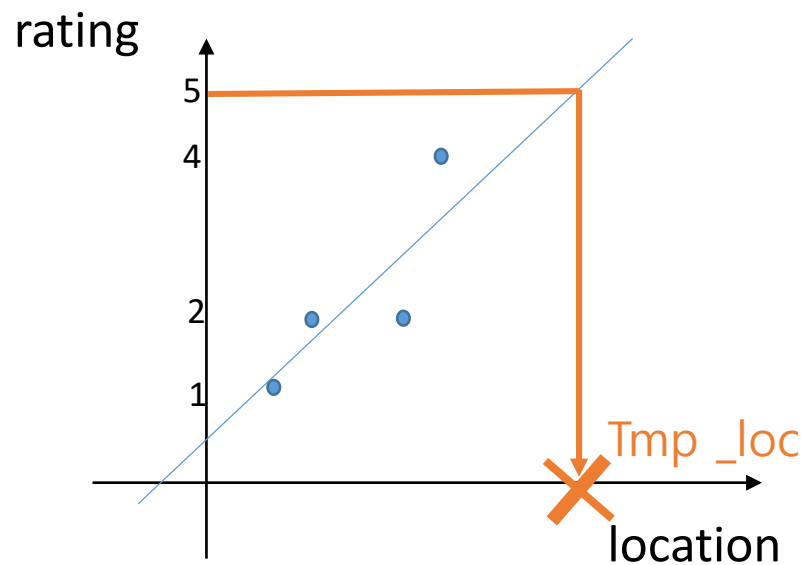
U I / U X
w o r k f l o w



1 Linear Regression

	X1	X2	X3	...	X100	Rating	Statistic
A	(a1	a2	a3	...	a100)	r1	S
B	(b1	b2	b3	...	b100)	r2	
C	(c1	c2	c3	...	c100)	r3	
D	(d1	d2	d3	...	d100)	r4	
	회귀 계수 β_1	β_2		...	β_{100}		

$$5 = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{100} x_{100}$$



service

U I / U X
w o r k f l o w

Royal palace of Madrid



Palace, royal, charles

Prada Museo



Historician,
nationalisation

Plaza Mayor, Madrid



Plaza, Philip, square

Buen Retiro park



Garden, park, palace

Puerta del sol



Square, city, plaza

Madrid Rio



Reasons, line, location

Gran via



Street, avenue



5

Q & A