AirSPACE 목표

"더 많은 이용자가 더 많은 SME를 만나게 하는 것"

User

이름: 김여자

연령/성별 : 23세 여자 현위치 : 가로수길 시간: 오후 12:30 취향: 우아한 분위기

주활동지역 : 서울 강남

선호하는 식당: 에비뉴로즈, 마노디셰프…

Connecting the dots



AirSPACE

Taste,
Region, Time, Age/Gender,
Theme, Intent…

SME



온다 이탈리아음식 ★4.72/5 * 방문자리뷰 173 * 블로그리뷰 103



연동 추천서비스 방문자 증가



Global 수준의 Local 추천기술



SME 노출 증대

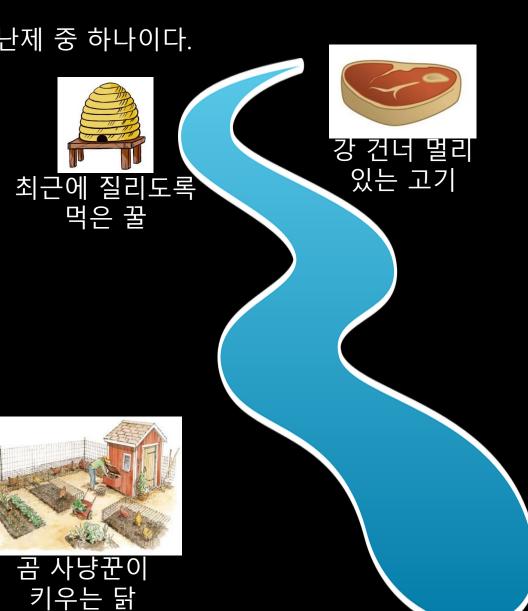
가장 좋아하는 걸 추천하면 될까?

"오늘 점심 뭐 먹지?"는 인류의 가장 어려운 난제 중 하나이다.



무엇을 먹어야 하지?





더 멀리 있지만

살짝 더 큰 고기

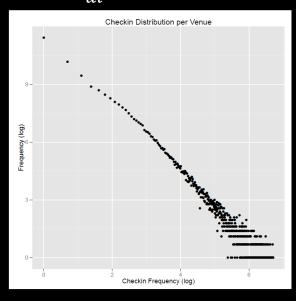
맛집 추천이 다른 추천과 차별화되는 주요 feature (Context)

 $Score_{user,item} = f(P_{ui}, d_{ui}, \mu_u, r_{ui})$

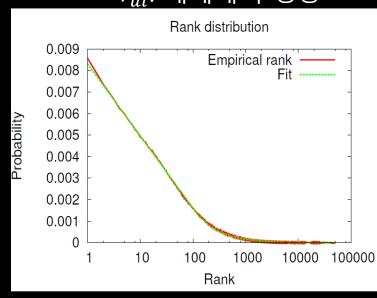
 P_{ui} = 유저의 맛집 preference d_{ui} = 유저와 맛집 사이의 거리 r_{ui} = 유저와 맛집 사이의 대체재

 $\mu_{y} = 유저의 활동 에너지 / 다양성 추구$

 d_{ui} : 거리의 영향



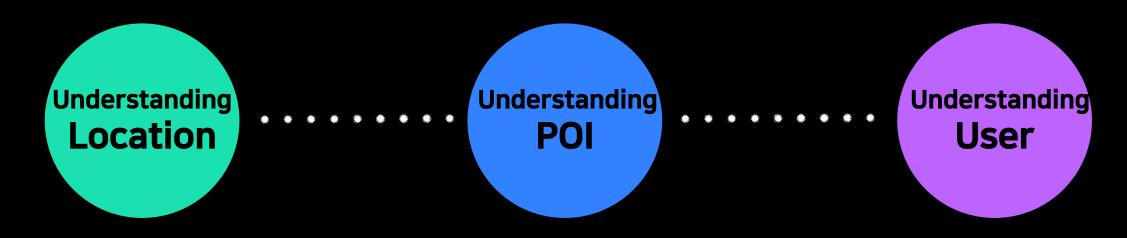
 r_{ui} : 대체재의 영향



μη: 다양성 추구

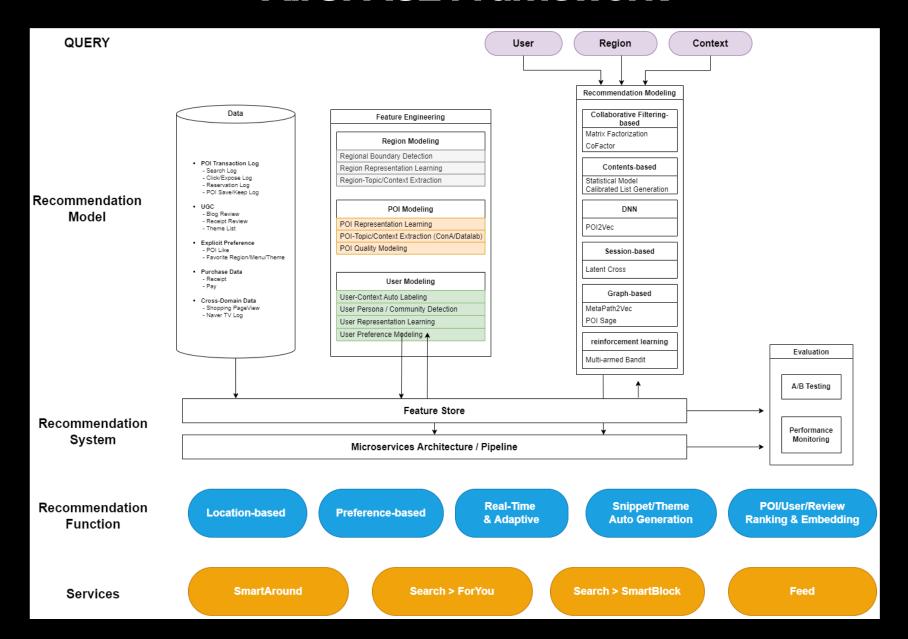


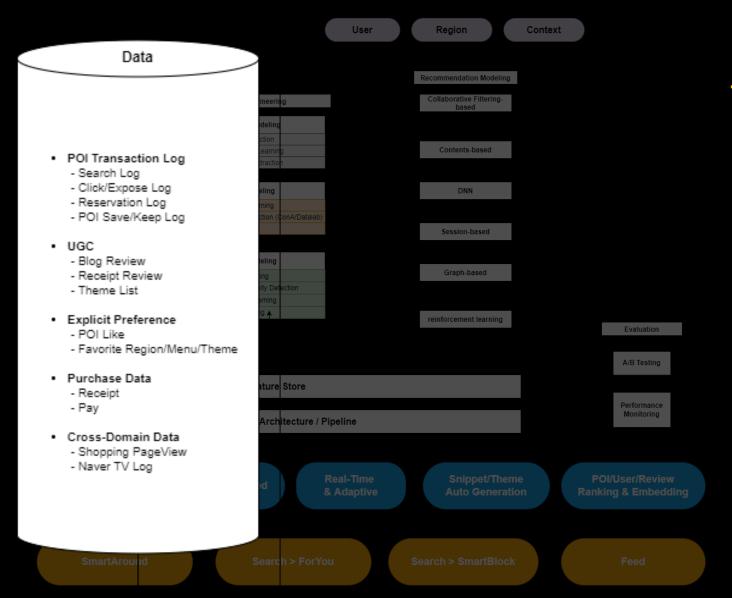
좋은 추천의 기반 = 지역, 장소, 유저를 이해하는 것



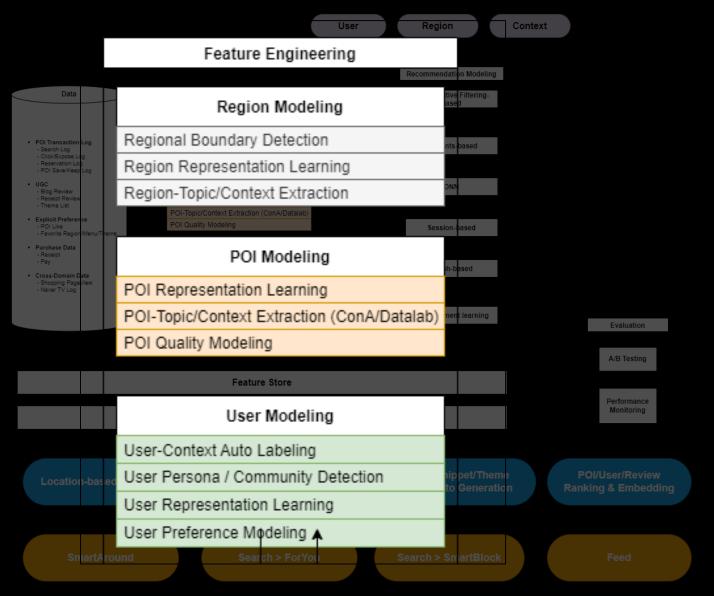
"이 주변 어디까지 가볼까?" "이 지역에서 사람들은 뭘 즐길까?" "이 장소는 어떤 사람들이 찾을까?" "이 장소는 어떤 장소와 어울릴까?" "이 장소는 언제 자주 찾을까?"

"나의 취향과 관심사에 맞는 장소는 어딜까?"



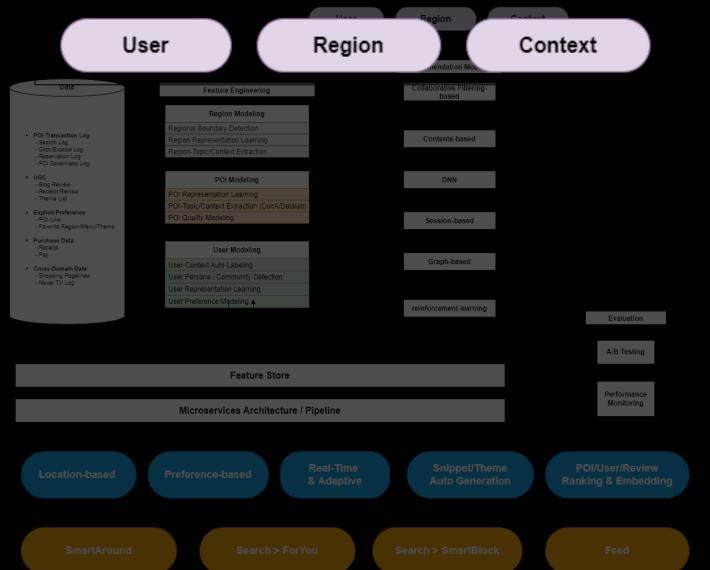


이용자로그, 피드백, 작성 UGC, Cross-domain Transaction(영수증,쇼핑 등) 등을 분석하여



이용자로그, 피드백, 작성 UGC, Cross-domain Transaction(영수증,쇼핑 등) 등을 분석하여

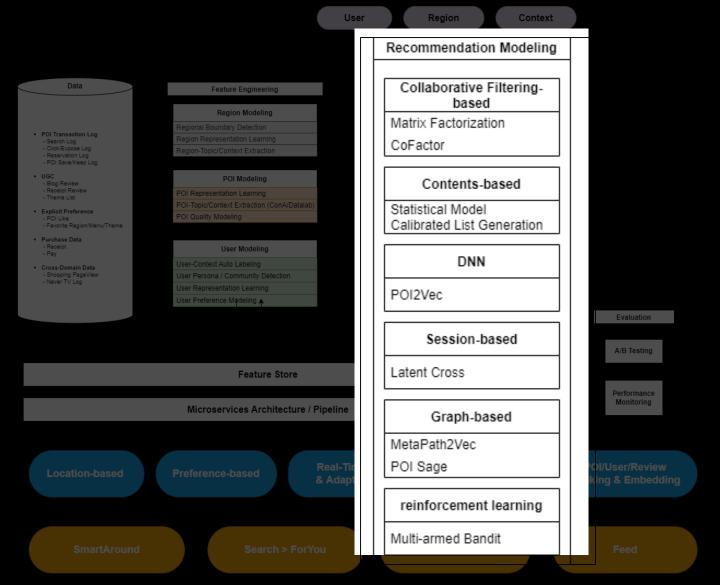
POI, User, Region, UGC별 주요 Feature와 이용자취향을 모델링하고



이용자로그, 피드백, 작성 UGC, Cross-domain Transaction(영수증,쇼핑 등) 등을 분석하여

POI, User, Region, UGC별 주요 Feature와 이용자취향을 모델링하고

이용자의 검색지역과 상황별 (위치,시간,취향등) 추천 시나리오에 맞게

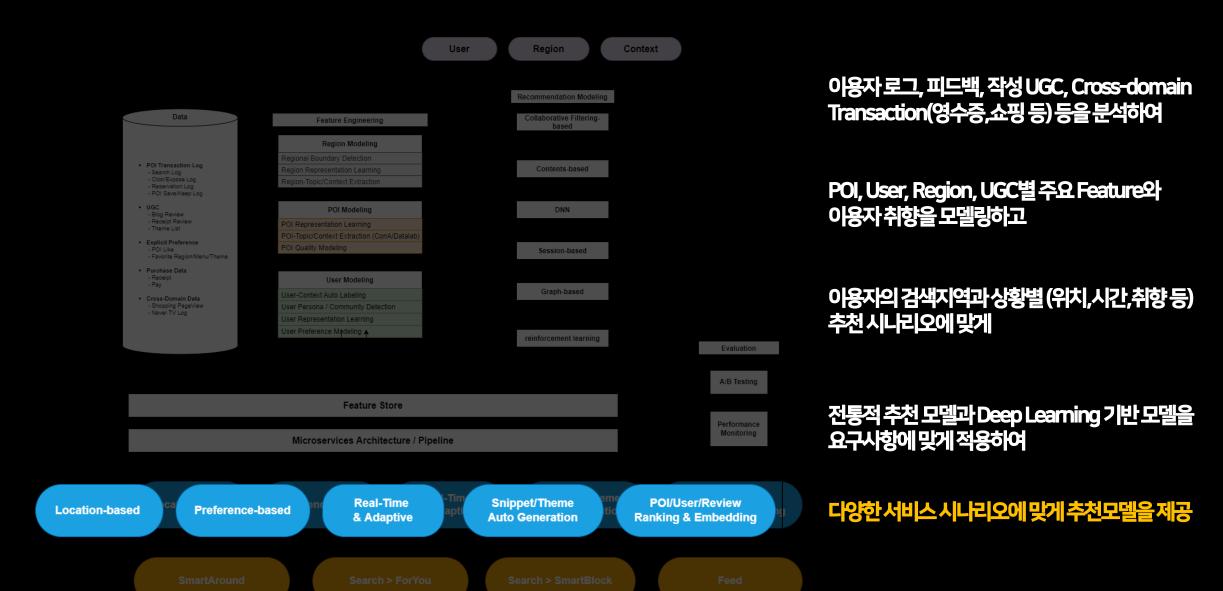


이용자로그, 피드백, 작성 UGC, Cross-domain Transaction(영수증,쇼핑 등) 등을 분석하여

POI, User, Region, UGC별 주요 Feature와 이용자취향을 모델링하고

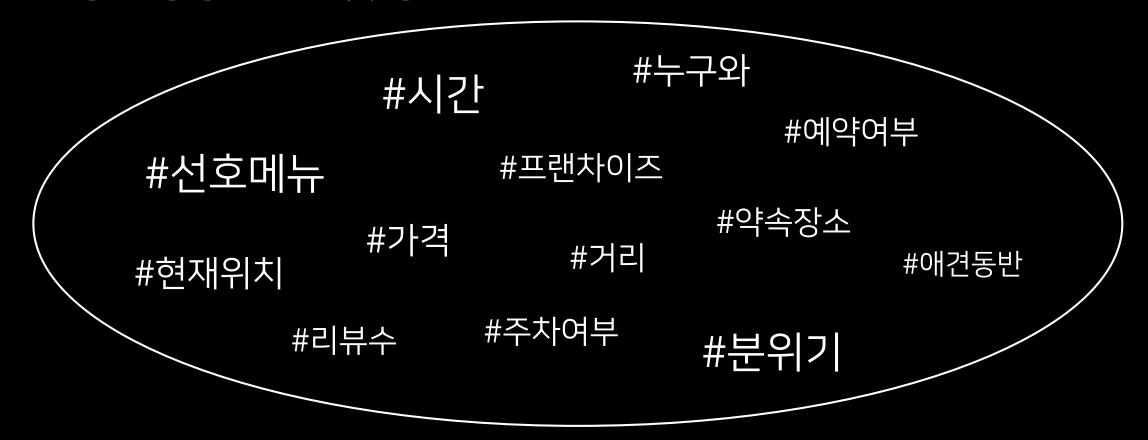
이용자의 검색지역과 상황별 (위치,시간,취향등) 추천 시나리오에 맞게

전통적 추천 모델과 Deep Learning 기반 모델을 요구사항에 맞게 적용하여



유저가 POI를 선택하는 법

POI에 고려되는 모든 것이 Context



Context를 종합적으로 고려하여 최적의 POI를 선택하게 된다!

Context를 고려한 POI 추천 기술

1. POI Feature Extraction

Non-embedding model

- 2. Poi2vec
- 3. Metapath2vec
- 4. Global Context Poi2vec
- 5. etc

Embedding model

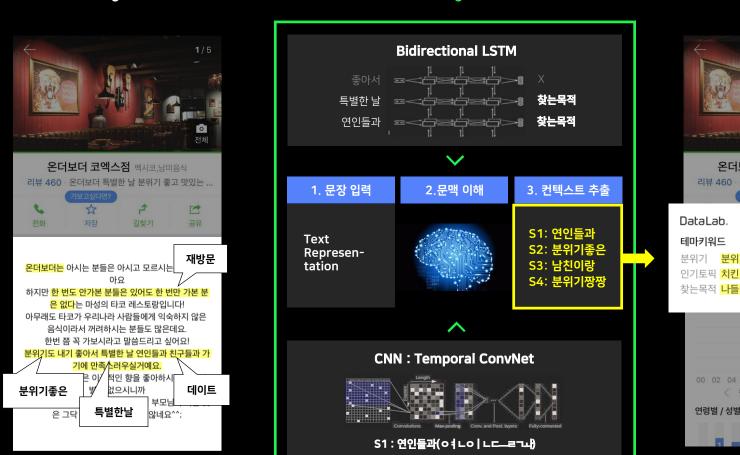
POI Feature Extraction: Keyword

POI에 대한 리뷰, 검색, 클릭을 요약하는 인공지능

POI에 대한 리뷰

Deep Learning

POI 토픽키워드





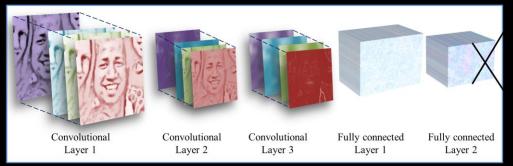
분위기 <mark>분위기좋은, 이국적, 화려한, 쾌적한</mark> 인기토픽 <mark>치킨, 레스토랑, 멕시코음식, 스테이크</mark> 찾는목적 나들이, 신선한, 데이트, 재방문, 특별한날



Convolutional Neural Networks (CNN) architecture

인간의 시신경과 같이 사물을 인지하는 인공지능 방법론





아이들과

(aideulgwa)

Text? idea!

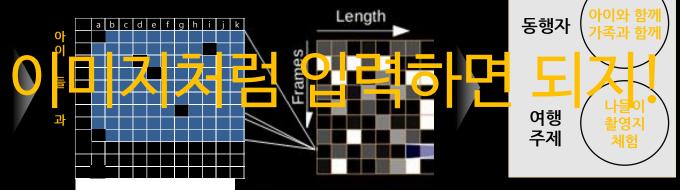


gogulyeo daejanggan maeuleun aideulgwa sajin jjiggiedo joheun gos ineyo ~ ^^

구리 고구려대장간마을에는 고구려 역사 문화 체험을 할 수 있는 체험장이 있는데 이곳에서 체험비는 고구려 와당은 2,000원, 전통활과 화 살은 5,000원입니다...^^:

대장간마을을 거진 다 둘러보았네요.. 그리 크 진 않지만 아이들과 간단히 나들이하기에는 좋

문장을 발음으로 변환하여 입력



97.67%

Long Short Term Memory (LSTM) architecture

문장의 앞뒤 패턴을 이해하는 인공지능 방법론

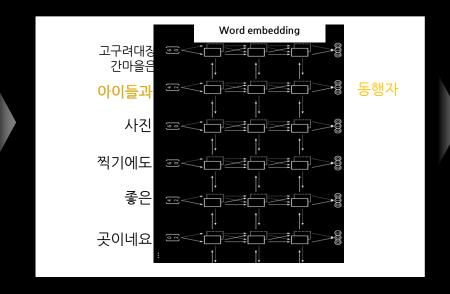


고구려대장간마을은 아이들과 사진 찍기에도 좋은 곳 같아요... 주변 곳곳이 사진을 찍기 좋 은 곳이네요~ ^^

구리 고구려대장간마을에는 고구려 역사 문화 체험을 할 수 있는 체험장이 있는데 이곳에서 체험비는 고구려 와당은 2,000원, 전통활과 화 살은 5,000원입니다...^^:

대장간마을을 거진 다 둘러보았네요.. 그리 크 진 않지만 아이들과 간단히 나들이하기에는 좋 은 곳 같아요..ㅎㅎ







95.74%

신규 테마 발굴에 초점

 오늘 점심 은 헤리 언냐 랑 강남 교자

 동행자 —

POI Feature Extraction : Popularity

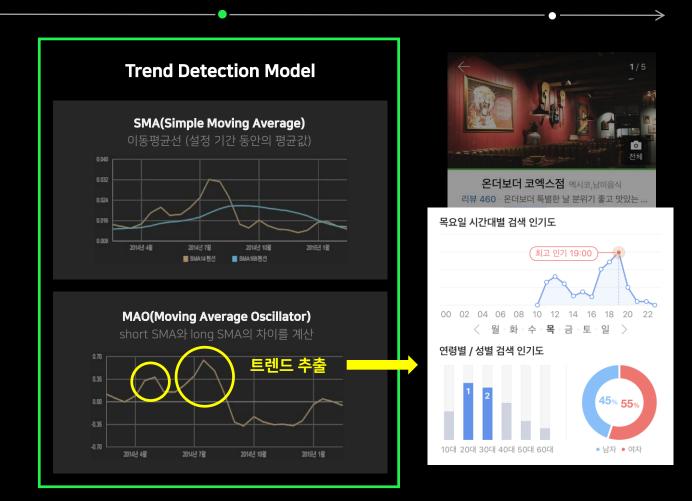
POI에 대한 리뷰, 검색, 클릭을 요약하는 인공지능

POI에 대한 검색,클릭로그

Statistical Model

시간/연령/성별 인기도





위치, POI, 사용자를 고려한 맞춤형 장소 추천 결과

서촌/12시/20대 초반 여자



서촌/12시/40대 남자



위치, POI, 사용자를 고려한 맞춤형 장소 추천 결과

서촌/9시/20대 초반 여자

서촌/**12시**/20대 초반 여자

서촌/15시/20대 초반 여자







POI Feature Extraction: Review Snippet

업체 랭킹만으로는 업체를 선택하기 어렵다는 의견이 다수 존재.

추천에 대한 신뢰도 향상을 위해 다양한 기술로 추천 이유를 나타낼 수 있는 요소들을 추출하여 서비스 적용 완료.

빅데이터 기반의 UP 레이블 추출

업체 검색 질의, 저장 추이, 최근 인기도 등을 고려한 인기 급상승 레이블 추출

빅데이터 기반 레이블

- · 어제 많이 검색된, 최근 많이 저장한
- · 지난주 이 지역 인기 TOP
- · 지난주 00으로 많이 검색된
- → 총 17개 레이블로 확장 완료 ("주말 점심 인기", "20대 선호" 등)

DB 기반 레이블

- · 00에 방영한
- · 미쉐린가이드 선정
- · 최근 오픈한 등





의미 기반 Snippet 추출

OverFeat에서 사용한 방법을 텍스트 CNN 모델에 적용하여 Heatmap으로 표현



음식점/카페,디저트/카페			생활	날,편의/미용	/미용실
키워드	CNN요약 빈도	앞문장요약 빈도	키워드	CNN요약 빈도	앞문장요약 빈도
과육	229	24	중화제	465	14
신맛나	87	11	샴푸실	1,769	141
망고라쉬	267	58	보호제	362	12

CoFactor 추천모델 기반 Vector 유사도 추출

MF 기반의 CoFactor 모델로, 이용자 및 POI 벡터를 학습, 추천된 업체와 이용자 벡터간 유사도 측정

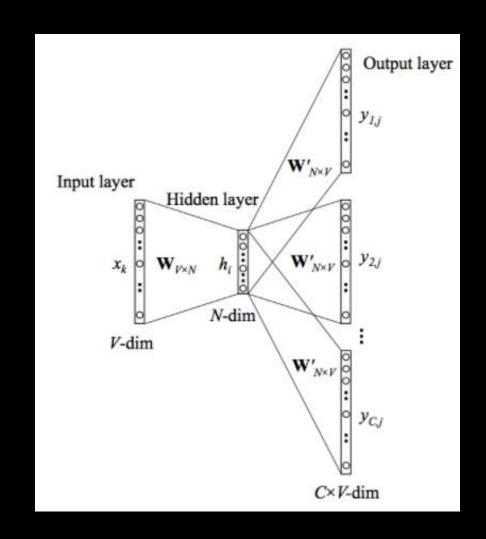
Word2Vec: Embedding-Based Model

Skip-gram: 중심단어로 주변단어를 예측

The quick fox jumps over the dog

The quick fox jumps over the dog

window



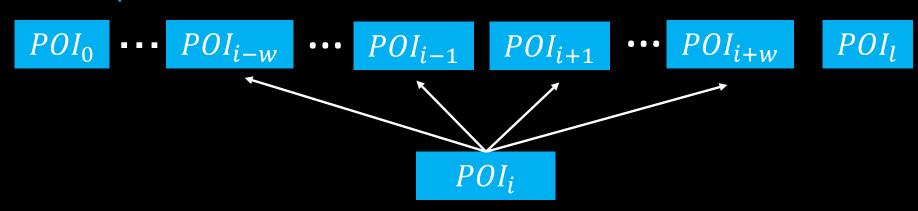
Understanding POI by click

POI2Vec: Embedding-Based Model



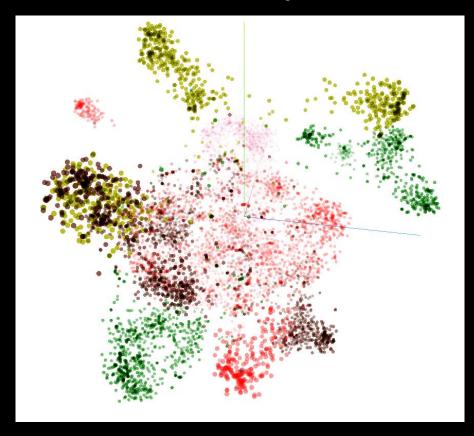
Session = 유저가 click 행동을 보이는 시간 단위 Sentence = session 내 poi click sequence

Click sequence



Understanding POI by Clicks Result of POI2VEC

Visualization by t-SNE



학습 성과

지 역

빨강 : 강남구

초록:서초구

연두 : 동작구

보라:분당구

지역 내의 유사한 업체를 잘 학습함

Understanding POI by Clicks Limitation of POI2VEC

지역이라는 Context만이 너무 강하게 학습된다!

- 1. 지역 이외의 다른 Context(ex-업종, 메뉴 등)를 잘 학습하지 못 한다.
- 2. 처음 가는 지역에 대한 추천의 결과가 좋지 못하다.

Understanding POI by Knowledge Graph Knowledge Graph with POI metadata

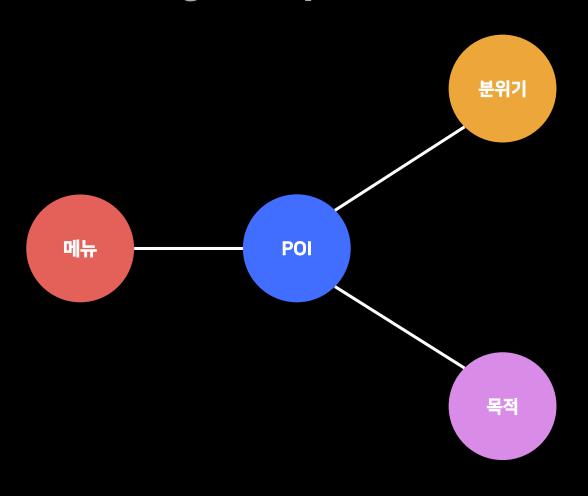
Knowledge Graph를 통해 이 문제를 해결하자!

- 1. 그래프는 노드 간의 관계 및 구조, 연결성에서 오는 정보를 보존함
- 2. 다양한 Type의 노드들을 통해, Graph를 다양하게 정의할 수 있음

Metapath2vec: POI Embedding by metadata

- 1. Knowledge Graph Generation
 - 다양한 Type의 Node들로 구성된 Heterogeneous Network를 구성함
- 2. Metapath Generation
 - 선행 지식을 기반으로, Knowledge Graph 내의 Path를 정의함
- 3. Node Embedding
 - Skip-gram 방식을 이용하여, Node들을 Embedding함

Knowledge Graph Generation



Heterogeneous Network

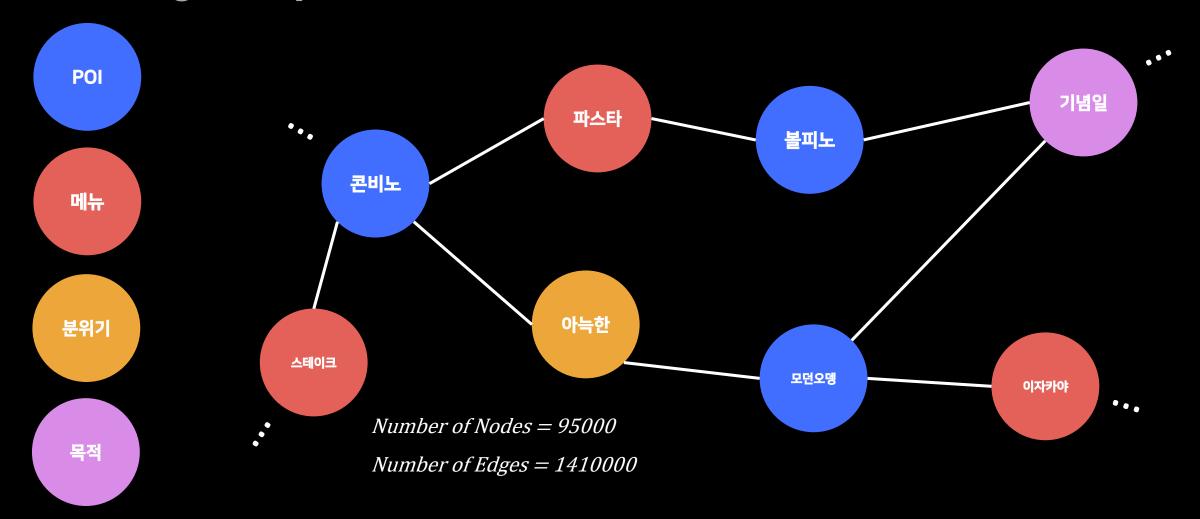
$$G = (V, E, T)$$

 $node\ v\ (V,T)\ \in\ set(POI, 메뉴, 분위기, 목적)$

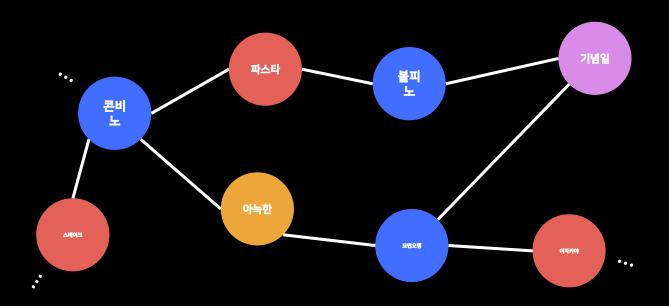
$$edge\ e\ (E) = (0,1)$$

Defined relation by ConA Model by LocalAI

Knowledge Graph Generation



Metapath Generation



Random walk
$$(V_i \rightarrow T_j) = \frac{P(V_i \rightarrow T_j)}{\sum_j P(V_i \rightarrow T_j)}$$

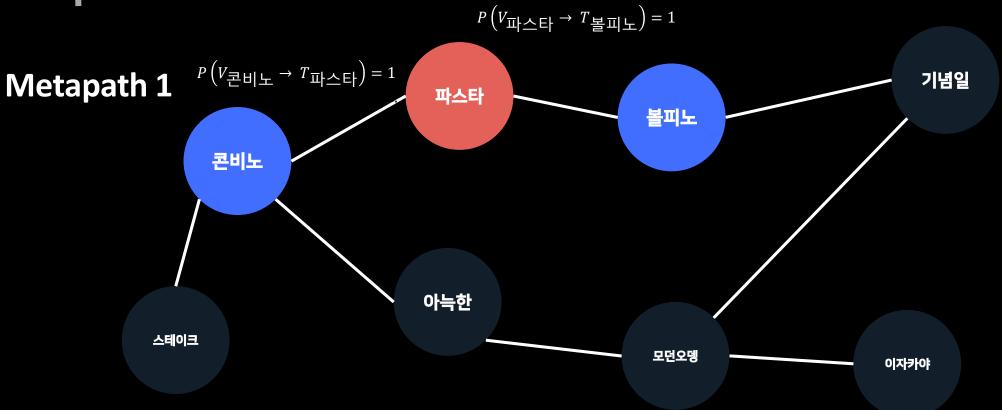
$$P(V_i \rightarrow T_j) = (0,1)$$
 $\begin{cases} 1, & \text{if (V, T) is occurred sequentially even once,} \\ 0, & \text{else} \end{cases}$

Train Dataset

- NAVER My PLACE Service
- Like Action

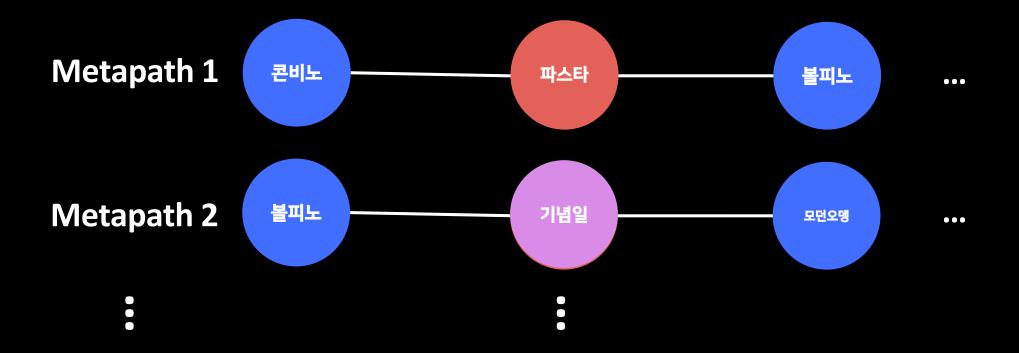
Then
$$P(Vi, Tj), P(Vj, Tk) = 1$$

Metapath Generation



User A가 콘비노와 볼피노를 함께 선호한 경우

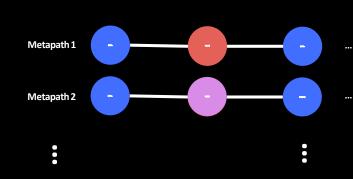
Knowledge Graph with POI metadata



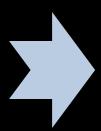
각 User가 선호한 POI간의 공통점을 기준으로 Metapath를 Generation함

Knowledge Graph with POI metadata

Metapath Generation



$$\theta = \operatorname*{argmax} \sum\nolimits_{(w,c) \in V_{pos}} \frac{1}{1 + \exp(-u_w^T v_c)} + \sum\nolimits_{(w,c) \in V_{neg}} \frac{1}{1 + \exp(u_w^T v_c)}$$



Sentence = Metapath

Center word(c) = Metapath 내의 모든 Node

Around word(w) = Center word앞 뒤의 window size 범위내의 Node

Node embedding (skip-gram)



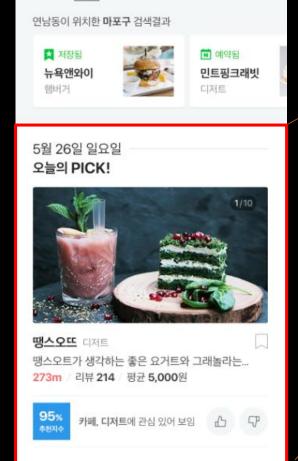






오늘의 Pick: History-based POI Recommendation

개인화 장소 추천 모델



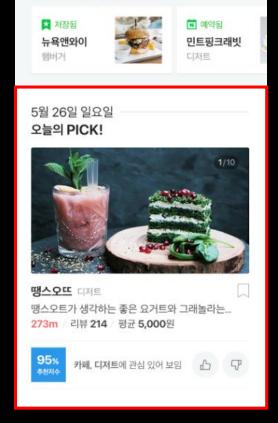
Model	특징	개요	
Co-Factor	MF 모델	기존 MF에 item간의 상관관계를 동시에 고려	Serendipity
POI2Vec	DNN 모델	이용자 성향 벡터와 가장 가까운 POI를 추천	User Coverage
Statistical Model	통계 모델	연령,성별 선호도 기반 추천	Cold Start Problem
LC	DNN 모델	시퀀스 기반 POI 추천	반응형 추천

오늘의 Pick: History-based POI Recommendation

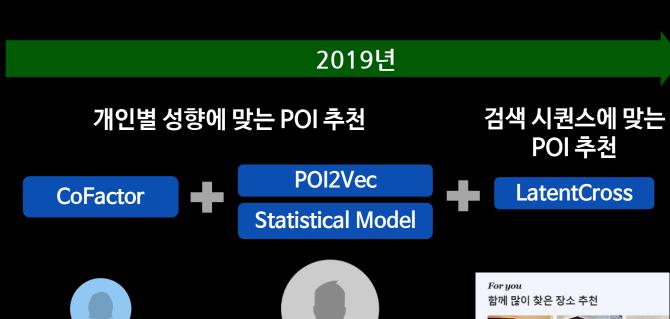
개인화 장소 추천 모델 Ensemble

Heavy Users

14%

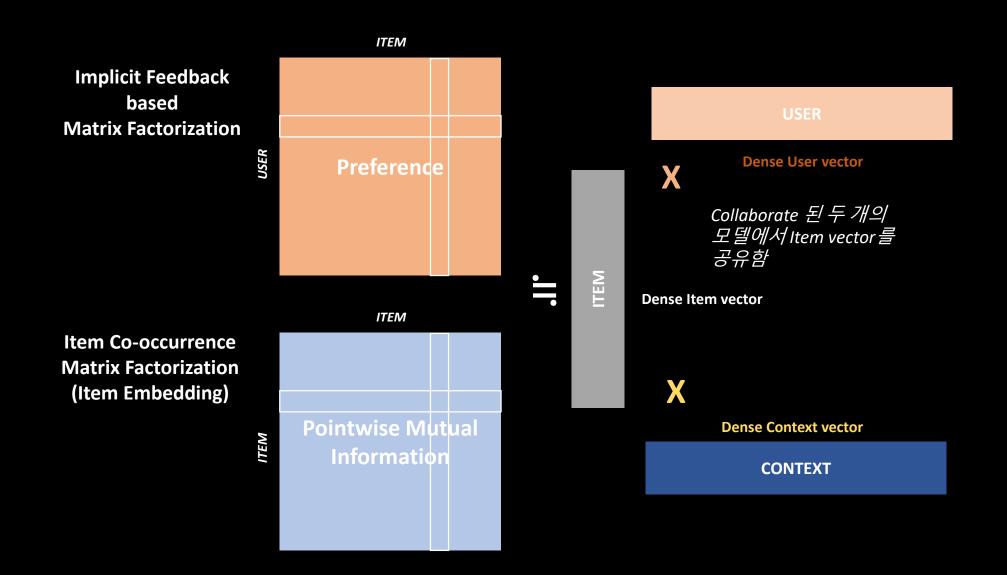


연남동이 위치한 마포구 검색결과





CoFactor: Factorization Meets the Item Embedding



CoFactor: Factorization Meets the Item Embedding

Rating Matrix

	호접몽 (분당)	하누비노 (분당)	달아래 (강남)	장서는날 (강남)
User 1 (분당)	0.43	1.26	0.10	
User 2 (강남)		0.12	0.52	1.02
User 3 (강남)	0.10		0.22	0.33
User 4 (분당)		0.50	0.12	
User 5 (분당)	0.13	0.26		0.05

Preference Modeling

- Click-based preference 추정 (Implicit Feedback)
- 1. Method 1
 - 가정 : n배 만큼의 click이 n배만큼의 preference를 갖음
 - 목표/의도 : User가 좋아하는 Item의 RMSE를 낮추겠다.
 - 문제점 : Popularity가 떨어짐

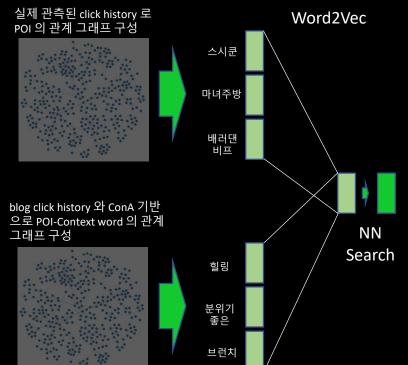
- 2. Method 2
 - 가정 : click은 confidence level 정도만 결정할 정도의 신뢰도만 갖음
 - 목표/의도 : Abuser에 강한 model을 만들겠다.
 - 문제점 : Serendipity가 떨어짐 / Item Coverage가 떨어짐

$$R(i,j) = p(i,j) \times c(i,j)$$

POI2Vec: Embedding-Based Model

User POI click history

유저 목록	유저의 click history			
User1	코이라멘	고쿠텐	스시쿤	토나리스 시
User2	마녀주방	88 브레드	에머이	배러댄피 브
User24045	가야밀면	금수복국	부산족발	할매국밥





POI-context word data from ConA

POI 목록	POI 에 tagging 된 단어 목록			
코이라멘	라멘	착한가격	혼밥	아담한
88 브레드	브런치	분위기 좋은	식빵	힐링

할매국밥	돼지국밥	향토음식	수요 미신히	분주한
			미역외	

POI2Vec: Embedding-Based Model

user

visited item vector's average and their nearest neighbor

레알차이나 호접몽 코이라멘

토나리스시

유명한

4.2

0.1

0.2

7.5

4.0

3.4

5.6

유저 검색기록

분위기	중식	고급	정자		
3.2	3.9	4.4	9.0	• • •	
0	4.1	0	10	•••	
0	0	1.2	11	•••	
1.2	0	0	8.0	•••	
- avaraga					

user vector

레알차이나

토나리스시

코이라멘

호접몽

average					
1.1	2.0	1.4	9.5	•••	2.5

user history 와 비슷한 맛집 목록

상해완탕 (정자1동 상해중식) 0.4942402243614197 (정자1동 이자까야) 키와미

0.48207470774650574

(정자1동 맥주집) **Gramps Ground**

0.46884194016456604 그래니살룬

(정자1동 파스타집)

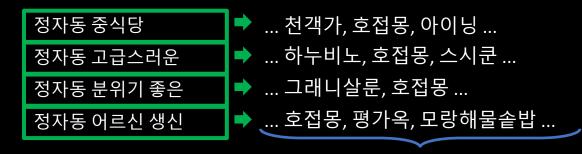
0.4571949243545532

천객가 (정자2동 중식)

0.4544779062271118

item (맛집)

co-click을 이용한 POI2vec



호접몽과 비슷한 맛집 목록

아이닝	(정자1동 중식)
0.8883786201477051 하누비노	(정자1동 비싼 한유
0.8703106045722961	
평가옥 분당점 0.8665239810943604	(정자2동 이북불고
0.8665239810943604 그래니살룬	(정자1동 파스타집
0.8661549091339111 모랑해물솥밥	(정자1동 한정식)
0.8563363552093506	
천객가	(정자2동 중식)
0.8559049367904663 미방	(정자1동 양갈비)
0.8495603203773499 스시쿤	(정자1동 스시)
0.8430420756340027 상해완탕	· (정자1동 상해중식
0.8427364230155945	

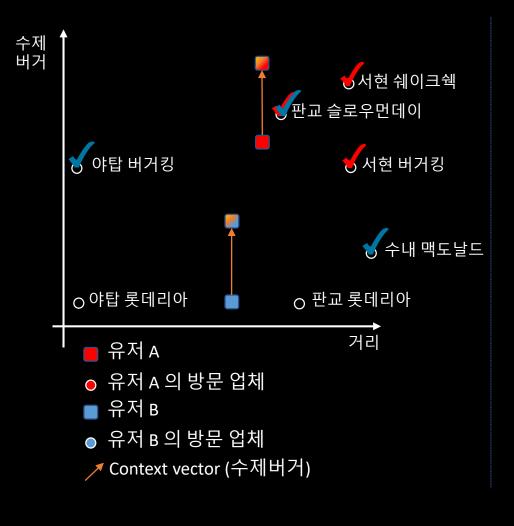
ITEM SENTENCES

Word2Vec Model

호접몽 vector (300 dim)



POI2Vec: Embedding-Based Model



유저 = 클릭한 업체의 평균 Context = train 된 embedding vector 유저 + Context = vector 합

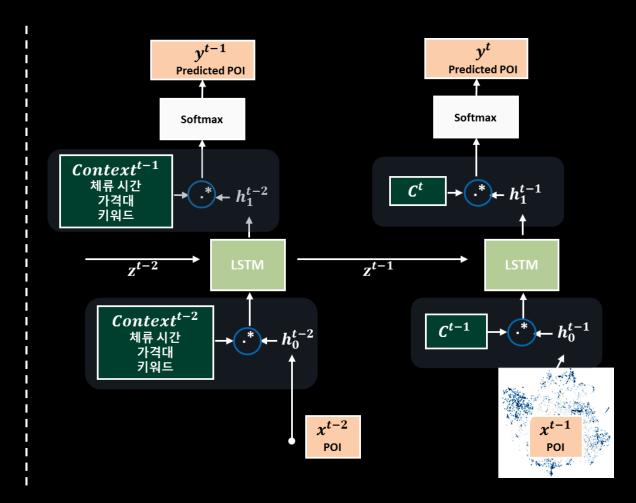
슬로우먼데이

0.6013346314430237

LatentCross: Session-based Recommendation

LSTM을 활용한 세션 기반 추천 모델





LatentCross: Session-based Recommendation

세션 기반 추천 모델

