

Neural Aentive Session-based Recommendation

<https://arxiv.org/pdf/1711.04725.pdf>

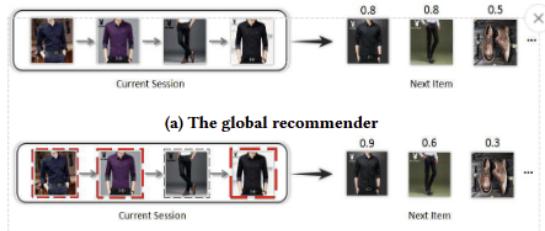
0. Introduction

- 기존 세션 기반 추천은 사용자의 현재 세션에서의 순차적 행동만 고려함
- 사용자가 세션 내에서 특정 목적을 가지고 행동하는 경우가 많지만 이를 반영하지 못함
- 단순 RNN 기반 모델만으로는 사용자의 세션 내 주요 목적 파악이 어려움
- 핵심 기여:
 - 순차적 행동 + 세션 목적을 동시에 고려하는 하이브리드 인코더 설계
 - Attention으로 세션 목적 추출 후 bi-linear matching으로 추천 점수 계산
 - 두 정보를 통합해 세션 표현 생성

1. Overview

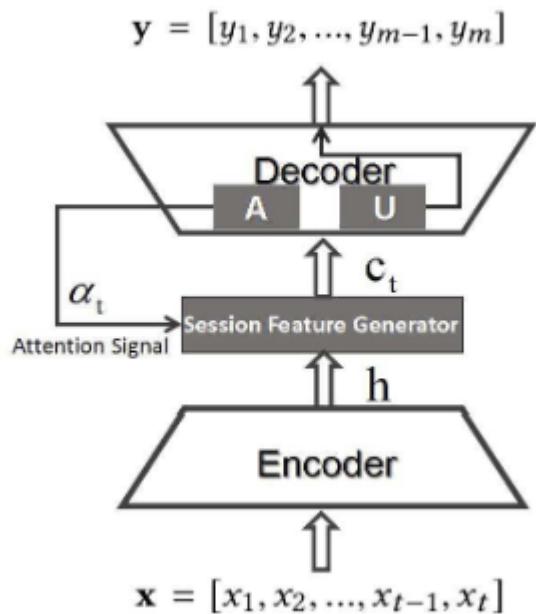
- 사용자의 세션 클릭 시퀀스를 Global Encoder와 Local Encoder로 분리하여 학습
- Model Structure:
 - Global Encoder: GRU 기반, 전체 행동 시퀀스 특징 추출
 - Local Encoder + Attention: 중요 행동에 가중치를 부여해 세션 목적 포착
 - Unified Representation: 두 인코더 출력을 결합해 세션 벡터 생성
 - Decoder: 후보 아이템과 세션 벡터 간 bi-linear matching으로 추천 점수 계산
- 클릭 기반 e-commerce 세션 추천

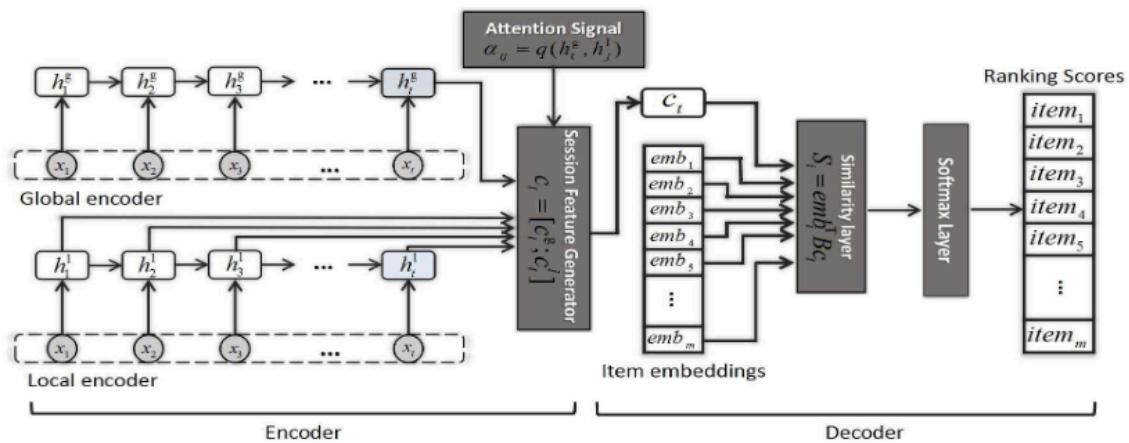
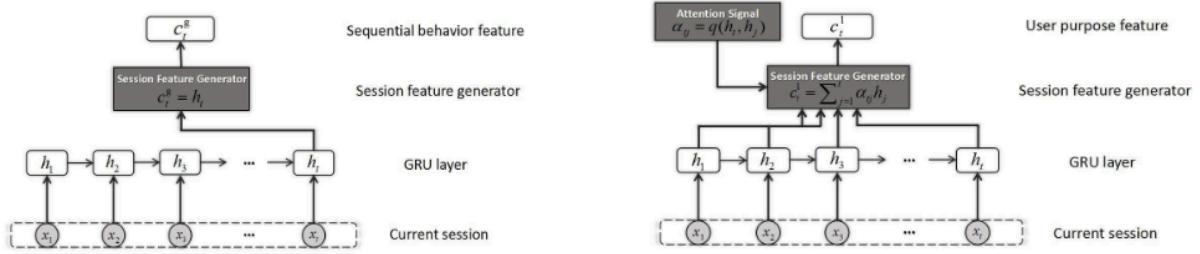
2. Challenges



- 세션 익명성: 사용자 프로필 없이 세션 내 정보만으로 추천
- Noise in Behavior: 순차 정보만으로 사용자 의도와 잡음을 구분하기 어려움
- Sequential + Purpose Capture: RNN만으로 순차 정보와 세션 목적을 동시에 반영하기 어려움
- 기존 한계: GRU 기반 세션 추천은 노이즈에 취약함

3. Method





- Global Encoder: GRU로 세션 시퀀스 처리, 마지막 hidden state를 전체 행동 특징으로 사용
- Local Encoder + Attention: 각 타임스텝 hidden state에 attention 가중치를 적용해 세션 내 중요 행동 임베딩 생성
- Unified Representation: Global과 Local 출력을 concatenate하여 세션 표현 생성
- Decoder: 세션 벡터와 item embedding 간 bi-linear matching으로 추천 점수 계산
- Training: cross entropy 기반 loss로 end-to-end 학습

4. Experiments

Datasets	all the clicks	train sessions	test sessions	all the items	avg.length
YOOCHOOSE 1/64	557248	369859	55898	16766	6.16
YOOCHOOSE 1/4	8326407	5917746	55898	29618	5.71
DIGINETICA	982961	719470	60858	43097	5.12

- Benchmarks: YOOCHOOSE, DIGINETICA

- 실험 설정: next-item 추천, Recall@20, MRR@20 측정
- 비교 모델: Improved GRU-Rec, RNN 기반 추천, 기타 baseline
- 평가 지표: top-k recall, MRR

5. Results

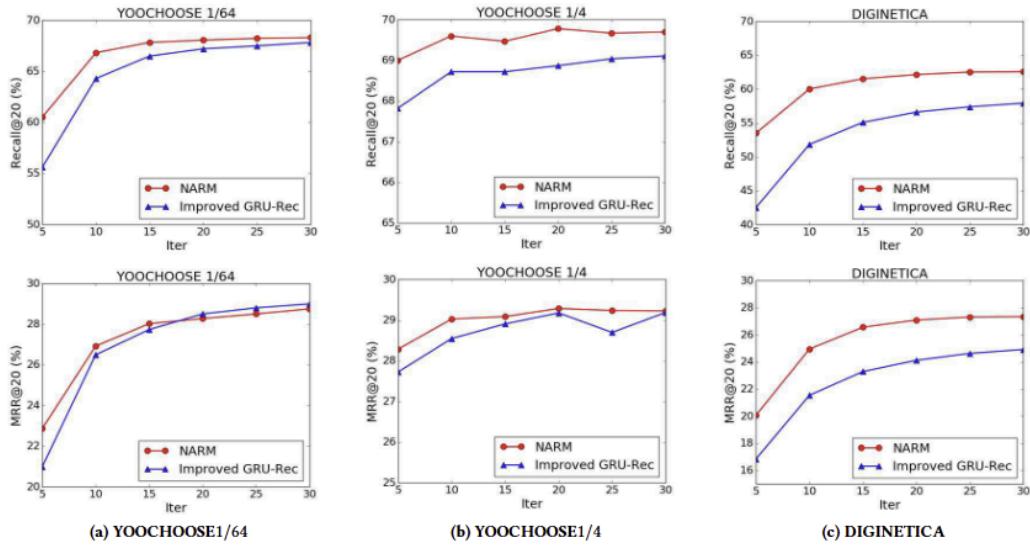
Table 2: The comparison of different decoders in NARM.

Decoders	YOOCHOOSE 1/64		YOOCHOOSE 1/4		DIGINETICA	
	Recall@20(%)	MRR@20(%)	Recall@20(%)	MRR@20(%)	Recall@20(%)	MRR@20(%)
Fully-connected decoder	67.67	29.17	69.49	29.54	57.84	24.77
Bi-linear similarity decoder	68.32	28.76	69.73	29.23	62.58	27.35

Table 3: Performance comparison of NARM with baseline methods over three datasets.

Methods	YOOCHOOSE 1/64		YOOCHOOSE 1/4		DIGINETICA	
	Recall@20(%)	MRR@20(%)	Recall@20(%)	MRR@20(%)	Recall@20(%)	MRR@20(%)
POP	6.71	1.65	1.33	0.30	0.91	0.23
S-POP	30.44	18.35	27.08	17.75	21.07	14.69
Item-KNN	51.60	21.81	52.31	21.70	28.35	9.45
BPR-MF	31.31	12.08	3.40	1.57	15.19	8.63
FPMC*	45.62	15.01	-	-	31.55	8.92
GRU-Rec	60.64	22.89	59.53	22.60	43.82	15.46
Improved GRU-Rec	67.84	29.00	69.11	29.22	57.95	24.93
NARM	68.32	28.76	69.73	29.23	62.58	27.35

* On YOOCHOOSE 1/4, we do not have enough memory to initialize FPMC. Our available memory is 120G.



(a) Performance comparison on YOOCHOOSE 1/64					DIGINETICA DATASET			
Models	d=50		d=100		Length	Baseline correct	NARM correct	Performance
<i>NARM_{global}</i>	67.26	26.95	68.15	28.37	1	8747	9358	+6.98%
<i>NARM_{local}</i>	67.07	26.79	68.10	28.38	2	6601	7084	+7.31%
<i>NARM_{hybrid}</i>	68.28	28.10	68.32	28.76	3	4923	5299	+7.63%
					4	3625	3958	+9.18%
					5	2789	3019	+8.24%
					6	2029	2202	+8.52%
					7	1520	1656	+8.94%
					8	1198	1295	+8.09%
					9	915	996	+8.85%
					10	690	753	+9.13%
					11	509	587	+15.32%
					12	411	459	+11.67%
					13	304	323	+6.25%
					14	243	260	+6.99%
					15	199	219	+10.05%
					16	149	165	+10.73%
					17	98	112	+14.28%
					18	88	93	+5.68%
					19	70	75	+7.14%

- 성능 향상: 모든 데이터셋에서 기존 baseline보다 성능 우수
- 긴 세션에서 성능 증가: 세션 길이가 길수록 Recall/MRR 향상 뚜렷
- Bi-linear vs FC Decoder: Bi-linear가 Recall 증가에 유리
- Baseline 비교: 기존 BPR/MF는 세션 추천에는 부적합, RNN + Attention이 유리

6. Insight

- Attention 기반 hybrid encoder는 세션 추천에서 목적/순차 정보를 동시에 반영
- 긴 세션이나 잡음이 많은 클릭 패턴에서 Robust
- 단순 RNN 대비 노이즈 분리 및 주요 목적 포착 능력이 성능 향상의 핵심 요인
- 후속 연구에서는 GNN 기반 세션 추천이나 Transformer 모델이 발전 방향