

Abstract

- We construct E-ConvRec, an authentic Chinese dialogue dataset consisting of over 25k dialogues and 770k utterances, which contains user profile, product knowledge base (KB), and multiple sequential real conversations between users and recommenders.
- We explore conversational recommendation in a real scene from multiple facets based on the dataset and particularly design three tasks: user preference recognition, dialogue management, and personalized recommendation.
- We establish baseline results on E-ConvRec to facilitate future studies.

Introduction

- Users describe the sought products in broader terms and a casual way, resulting in difficulties in eliciting users' preferences.
- Accurately recognizing user preference words in casual utterances lays the foundation for providing high-quality recommendations.
- Customers usually proceed in a coarse-to-fine manner to gradually make their decisions during a conversation.
- Customer service staffs need to conduct effective interactions with them to collect required information or make a recommendation.
- There are massive personalized user profiles and product knowledge in the E-commerce domain.
- The auxiliary information makes it easy to trace connections from users to specific items and provides a high-quality recommendation for customers.

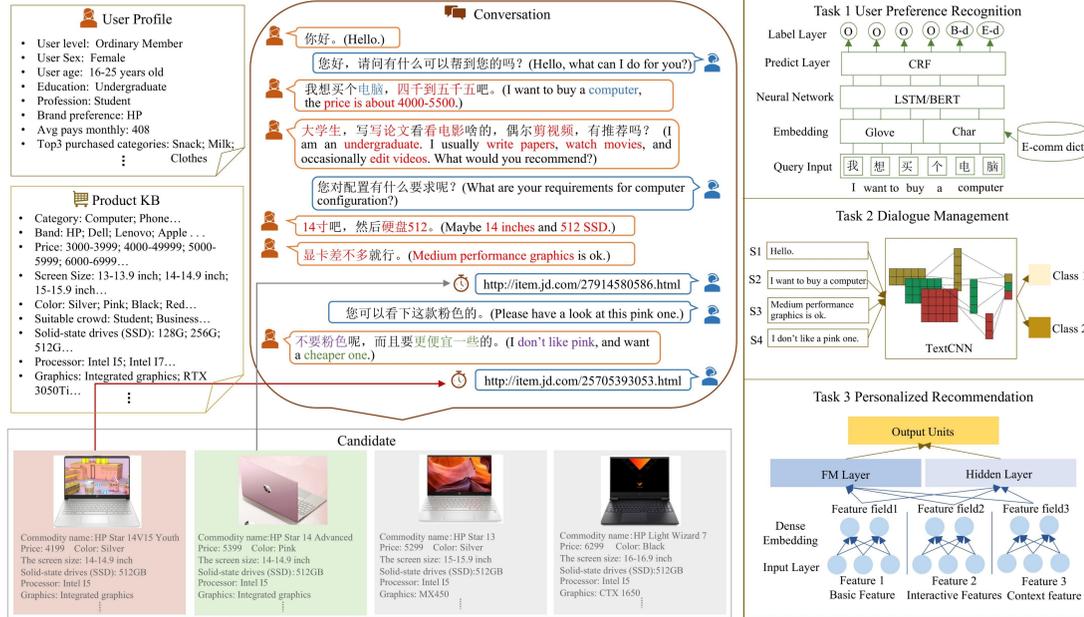
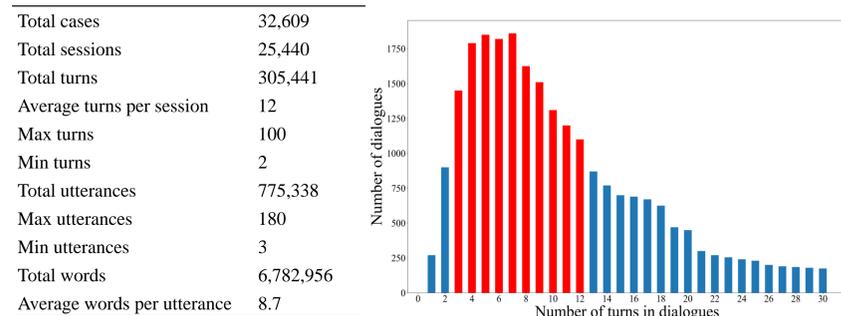


Figure1. A dialogue example for E-ConvRec. This dataset provides conversation flows from real scenario of E-commerce, user profile and product KB to enrich recommendation. Moreover, three sub-tasks of user preference recognition, dialogue management and personalized recommendation are devised to facilitate research on CRS.

Data Collection

Dialogue Flows

Our dialogue dataset contains conversations on pre-sales topics between users and customer service staff in an E-commerce scenario and pre-select from a broader set of dialogues.



User Profile

We provide 20 different types of user profiles.

- Attributes of users' personal information.
- Historical shopping activities.

User Profile	User-level	User sex	User age	Avg pay monthly	Top 3 order recently
	Ordinary Member	Female	16-25 years old	103.4	phone bill, snack, milk
	Golden Member	Male	36-45 years old	1566.0	phone, earphone, waist support

Product Knowledge Base

- KB information of products mentioned from the E-Commerce platform.
- Knowledge of other products in the same category and sell in the same online shop.

	Commodity ID: 25705393053 Category ID: 27943 Type: High-end light notebooks Screen sizes: 14.0-14.9 inches	Name Category: Computer System: Windows 10 Processor: Intel I7 Color: Silver
	Commodity Title: HP (HP) Star 14 /15 Youth edition Ultra-thin Student online class business games lightweight notebook [14.0 inches] I7-1165G7	
	Commodity ID: 100011762577 Category ID: 18628 Origin: China System: HarmonyOS 2	Name Category: Mobile Phone CPU: Kirin 9000 RAM: 8GB Color: Black
	Commodity Title: HUAWEI Mate 40 Pro 4G Note kirin 9000 flagship chip 8GB+128GB bright black HUAWEI mobile phone	

Task Formulation & Experiments

User Preference Recognition

Four kinds of preference words: **descriptive preference words**, **category words**, **comparative preference words**, and **negative preference words**.

Users' utterances
你们家有没有小女生用的那种口红? (Do you have the lipstick that girls wear in your shop?)
我不喜欢暗红色的。 (I don't like dark red.)
比之前的那个颜色淡吗? (Is it lighter than the previous one?)
请问有比iphone 性能好的安卓手机吗? (Is there any mobile phone with an Android system that is better performance than iphone.)
我不要黑色, 我要亮色的。 (I don't want a black one, I want a bright color one.)
麻烦你重新推荐一个低于300的手机, 不要贵的。(Please recommend another mobile phone with a price lower than 300. I don't want a too expensive one.)

- Sequence labeling task.
- Lexicon **E-comm dict** in the E-commerce domain with a vocabulary size of 722k.

Model	w/ CTB dict	w/ E-comm dict
LSTM + CRF	74.00	75.29
Simple-Lexicon	76.30	75.89
Multi-Digraph	76.37	77.40
FLAT	76.60	79.24
LEBERT	78.91	78.53

Dialogue Management

The system should make the proper recommendation at the golden time. We specifically focus on the task of recommendation timing prediction.

Method	Precision	Recall	F1
TextRCNN	72.61	48.44	58.11
TextRNN	69.86	57.78	63.25
TextCNN	70.10	58.62	63.85
TextCNN+Intent	72.08	59.73	65.33

- Binary text classification task.
- Label the intention for each query in the dialogues with an in-house intent classifier of the E-commerce domain.

Personalized Recommendation

The task of conversational recommendation is designed to judge whether the user will buy a candidate product based on user profile, product KB and conversation content

Model	Feature	AUC	T@1	T@5	T@10
DeepFM	BF	69.77	15.65	44.70	61.94
	BF+IF	80.19	33.56	63.85	76.28
	BF+CF	74.53	22.07	54.80	70.93
	BF+IF+CF	83.17	37.06	70.20	81.46
FGCNN	BF	70.73	16.87	45.50	62.81
	BF+IF	78.50	35.44	63.31	74.87
	BF+CF	73.63	22.02	53.83	69.39
	BF+IF+CF	80.82	37.28	68.24	78.74

- Click-through rate (CTR) prediction task.
- Basic Features (BF), Interactive Features (IF), Contextual Features (CF).

Case Study

User profile :	User level	User sex	User age	Avg pay monthly	Top3 purchased recently
	Golden Member	Male	26-35 years old	RMB: 6,072.98	watch/pen/earphone
Q:	您好。(Hello.)				
A:	您好, 请问有什么可以帮助您的吗? (Hello, what can I do for you?)				
Q:	你们店里有什么手机, 推荐一下? (I want to buy a mobile phone, what do you recommend?)				
A:	请问您对性能有什么需求么? (What are your requirements for performance?)				
Q:	没什么特别的需求, 我不打游戏, 最好是照相清晰的, 内存要大于200G。(I don't have any special requirements. I don't play games. It's better to take clear pictures and the memory should be more than 200G.)				
A:	https://item.jd.com/10036452410557.html				
Q:	颜色我要深色的, 不要金色的。(I would like a dark color, not golden.)				
Model	DeepFM+BF+CF	DeepFM+BF+IF	DeepFM+BF+CF+IF	Ground truth	
Recommend result					
	Price: 7488 RAM: 512G Color: Black Camera: 50 million pixels	Price: 6,488 RAM: 256G Color: Golden Camera: 50 million pixels	Price: 6,458 RAM: 256G Color: Black Camera: 50 million pixels	Price: 6,458 RAM: 256G Color: Black Camera: 50 million pixels	Price: 6,458 RAM: 256G Color: Black Camera: 50 million pixels

User profile :	User level	User sex	User age	Avg pay monthly	Top3 purchased recently
	Golden Member	Female	36-45 years old	RMB: 7,872.98	seafood/shoes/milk
Q:	在吗? (Anyone here?)				
A:	在的! 亲爱的客人, 我有什么可以帮助您的吗? (Of course! What can I do for you? My dear.)				
Q:	快过年了, 我需要一个新的冰箱, 推荐一下。(Spring Festival is coming. I need a new refrigerator. Help me recommend one.)				
A:	我们有一款新出的冰箱, 请您看一下。(We have a new refrigerator, please take a look.)				
A:	https://item.jd.com/10023613092198.html				
Q:	这款单开门的不错, 但是这只适合小两口, 我们过年一大家子人呢。(This single-door refrigerator is good, but this is only suitable for the couple, we have a big family.)				
A:	您的预算是多少呢? (What is your budget?)				
Q:	3000左右吧。(Around 3000.)				
Model	DeepFM+BF+CF	DeepFM+BF+IF	DeepFM+BF+CF+IF	Ground truth	
Recommend result					
	Price: 2,499 Capacity: 200-249L Color: Red Number of doors: 1	Price: 2,099 Capacity: 100L Color: Green Number of doors: 1	Price: 2,499 Capacity: 200-249L Color: White Number of doors: 1	Price: 2,899 Capacity: 500-549L Color: Black Number of doors: 2	

- Green box highlights the product mentioned in the conversation, and the orange one refers to ground truth