Introduction 00 Proposed Approach: E-NSDE 000 Experimenta 000

2024-11-11

ENSDE

Evidential Stochastic Differential Equations for Time-Aware Sequential Recommendation

Krishna Prasad Neupane, Ervine Zheng, and Qi Yu

Rochester Institute of Technology

December, 2024



Evidential Stochastic Differential Equations for Time-Aware Sequential Recommendation

Krishna Prasad Neupane, Ervine Zheng, and Qi Yu

Rochester Institute of Technology

December, 2024



▲口▶ ▲圖▶ ▲臣▶ ▲臣▶ 三臣 - のへの



Sequential recommender systems are designed to capture users' evolving interests over time to allow more personalized and timely suggestions. It has following key characteristics:

- **Time-Ordered Interactions:** Capture and analyze users' interactions in a specific sequence to reflect evolving interests.
- User Behavior Prediction: Predict the next item a user is likely to engage with, based on previous actions.
- Dynamic Preference Modeling: Track and adapt to users' changing preferences over time.
- **Context Awareness:** Often incorporate contextual factors, like time and location, for more relevant recommendations.
- Personalization: Tailor suggestions specifically to each user's unique behavior patterns.
- Real-Time Adaptability: Update recommendations in real-time as new user interactions occur.

Overview

Introduction

00

Proposed Approach: E-NSDE

Experimentation

Introduction

Proposed

E-NSDE

Key Challenges and Motivation

- Existing sequential recommendation models (e.g., BERT4Rec, SASRec, etc.) often assume uniform time intervals between user interactions, which fails to capture real-world, non-uniform patterns.
- This oversight can lead to inaccurate predictions, as user preferences may change significantly over varying interaction gaps.
- As interaction intervals lengthen, the uncertainty in predicting user behavior grows, making it harder to provide accurate recommendations.

Table: Impact of interaction interval

Interaction	Interval (Seconds)	Ranking \downarrow			
	,	BERT4Rec	E-NSDE		
$6 \rightarrow 7$	44	4	4		
13 ightarrow 14	623,591	24	16		
$116 \rightarrow 117$	62	6	3		
$150 \rightarrow 151$	896,291	56	18		

The unrealistic assumption of a uniform interaction interval could significantly impact the model's capability to capture users' continuously evolving behavior and subsequently hurt the recommendation performance as demonstrated in Table 1. ENSDE Introduction Key Challenges and Motivation

	ivation	d Motiva	hallenges	Key Ch	Key
odels me in I-work redict nterat ncerts ovide eractic	mendation r me uniform o capture re o inaccurate over varying rigthen, the t harder to p : Impact of in	tial recomment from assume ich fails to ca an lead to in sificantly over stervals lengt making it ha ma. Table: Im	Existing seep SASRec, etc interactions, This oversign may change As interactic behavior gro recommenda	E S Ir A b	
Fact	hanai (Seanh) .	Interaction Interaction			
4		67			
24	623,494	23 - 24 1			
	100 M	10. 10.			
	1 1		71 5	-	
1154	a or a liting	*******************	The diversity		
irty to seque	sodel's capal avior and su	act the mode	continuously	5	

・ロト・日本・日本・日本・日本・日本

Experimentation



We propose an Evidential Neural Stochastic Differential Equation (E-NSDE) framework as shown in Figure below.



- Integrates neural stochastic differential equations (NSDE) with evidential learning for time-aware uncertainty quantification.
- Includes a monotonic network to ensure a positive correlation between interaction gap and uncertainty.
- Leverages interaction and time-guided evidential uncertainty to maximize information gain through exploration of a large_item_pool

ENSDE Proposed Approach: E-NSDE Solution: E-NSDE Framework

2024-





The NSDE module generates fine-grained user and item representations as follows:

$$u(T) = u(t_0) + \int_{t_0}^T f(u(t), t; \psi) dt + \int_{t_0}^T g(u(t), t; \omega) dB_t$$
(1)

Experimentation

where $u(t_0)$ is the user's initial representation which aggregates a set of initially interacted items: $u(t_0) = agg(i_1..i_k)$ and T is the final time.

$$i(T) = i(t_0) + \int_{t_0}^T f(i(t), t; \psi) dt + \int_{t_0}^T g(i(t), t; \omega) dB_t$$
 (2)

where, $i(t_0)$ represents item's initial representation In Evidential module.

- The rating network utilizes the fine-grained user u_t and item i_t representations from the NSDE and predicts the score γ_(u_t,i_t).
- The monotonic network is designed in such a way that the increase in time interval (Δt), increases the output, i.e., the variance of the predicted rating.



E-NSDE Framework Con

predicted rating

 $\label{eq:response} \begin{array}{l} {\bf I} \mbox{ The NDE models generates for grained user and item representations as follow:} \\ {\bf s}(T) = s(t_0) + \int_0^T f(u(t), t; v) dt + \int_0^T g(u(t), t, \omega) d\theta_{a} \quad (1) \\ {\rm where } s(t_0) \mbox{ the user's initial representation which aggregates a set of initially interacted items: } u(t_0) = {\rm agg}(t_0-t_0) \mbox{ and } T \mbox{ is the final initial set of the set of the$

$$i(T) = i(t_0) + \int_{t_0}^{T} f(i(t), t; \psi)dt + \int_{t_0}^{T} g(i(t), t; \omega)dB_i$$
 (2)

where, $i(t_b)$ represents item's initial representation In Evidential module, The rating network utilizes the fine-grained user u_i and item i_i representations from the NSDE and predicts the score $\gamma_{(u_i, i_j)}$. The monotocic network is designed in such a way that the increase in time interval (dx_i) increases the estapsi, i.e., the variance of th Introduction 00 Proposed Approach: E-NSDE

Experimentation

Training: Evidential and Weighted Bayesian Personalized Ranking Loss

■ The regularized EDL loss with negative log-likelihood (L^{NLL}[f_☉]) to maximize the marginal likelihood and an evidential regularizer (L^R[f_☉]) to impose a high penalty on the predicted error with low uncertainty for each sequential update is:

 $\mathcal{L}_{EDL}(u_t, i_t) = \mathcal{L}^{NLL}[f_{\Theta}(u_t, i_t)] + \lambda \mathcal{L}^{R}[f_{\Theta}(u_t, i_t)]$ (3)

where λ is a regularization coefficient.

 To leverage the effective exploration for the long-term, we formulate weighted BPR loss which is computed from non-interacted (negative) that are similar to the user's future interacted items.

$$\mathcal{L}_{\text{WBPR}}(u_t, i_t) = \sum_{(u_t, i_t, j_t \in \mathcal{N}_t)} w_{(i_t, j_t)} \{-\ln\left[\sigma(\hat{r}_{(i_t, j_t)})\right]\}$$
(4)

where r̂_(it,jt) = r̂_(ut,it) − r̂_(ut,jt), σ(·) is the sigmoid.
The overall loss of the end-to-end model training is obtained by combining the EDL and WBPR loss:

$$\mathcal{L}(u_t, i_t) = \mathcal{L}_{\mathsf{EDL}}(u_t, i_t) + \zeta \mathcal{L}_{\mathsf{BPR}}(u_t, i_t) \quad \text{ for all } (5)$$



Proposed Approach: E-NSDE 000



DE:	xt-iten	1 recorts	mendat	ion task	ka.				
• Wi Wi	e split e lever ovide s	users b age the user pref	y 70% i actual ference Recomm	nto trai time of evolutio	in and 3 interac on. n perform	10% in t tions (i	est. • UNIX	timest.	итр)
1000	Mark .	Wate					~	Anne	
Ryssel M	-	199			-	100.000	100.000		
-	Lapone A	Last of the	1.000.000	1004110	1.801.000	1.001.000	1.001.01.02	1 Party and and a state of the	
-	111111111		1.000,0100 0.000,0100 0.000,0100 0.000,0100 0.000,0100 0.000,0100			1.000	1 100-21 107 1 100-21 107 2 100-21 2 100-21 2 100-21 2 100-21 2 100-21 2 100-21 2 100-21 2 100-21 2 100-		11000 11000 11000 11000 11000 11000
	1.4 1010	1007-010	1090-1103	1444-118	1001.000	1.000.000	1.001-01.011	1100.0104	11.00.
	1.000				100.000		100.000	100.000	-

mmendation Performance Comparison

Our experiment setting of sequential recommendation is based

- Our experiment setting of sequential recommendation is based on next-item recommendation tasks.
- We split users by 70% into train and 30% in test.

Recommendation Performance Comparison

Introduction

We leverage the actual time of interactions (in UNIX timestamp) to provide user preference evolution.

Table: Recommendation performance comparison

Category	Category Model		MovieLens-100K		Movielens-1M		Netflix		Amazon Book	
		P@5	nDCG@5	P@5	nDCG@5	P@5	nDCG@5	P@5	nDCG@5	
Dynamic MF	timeSVD++ CKF	$\substack{0.3842 \pm 0.015 \\ 0.3916 \pm 0.017}$	$\substack{0.3420\pm0.013\\ 0.3620\pm0.013}$	0.3917±0.016 0.3928±0.016	$\substack{0.3508 \pm 0.013 \\ 0.3552 \pm 0.015}$	$\substack{0.3756 \pm 0.016 \\ 0.3600 \ \pm 0.017}$	$\substack{0.2951\pm0.013\\0.2986\pm0.014}$	$\substack{0.3601 \pm 0.014 \\ 0.3823 \pm 0.016}$	$\substack{0.3128 \pm 0.012 \\ 0.3214 \pm 0.015}$	
Graph	NGCF LightGCN	$\substack{0.3859 \pm 0.014 \\ 0.4103 \pm 0.014}$	0.3662±0.012 0.3702±0.013	0.3978±0.016 0.4028±0.017	$\substack{0.3587 \pm 0.018 \\ 0.3632 \pm 0.015}$	0.3574±0.015 0.3617±0.013	$\substack{0.3167 \pm 0.017 \\ 0.3204 \pm 0.016}$	0.3574±0.012 0.3678±0.013	$\substack{0.3321 \pm 0.011 \\ 0.3382 \pm 0.012}$	
Sequential	CASER SASRec BERT4Rec S ³ -Rec CL4SRec SAR ResAct	$\begin{array}{c} 0.4096 \pm 0.012 \\ 0.4105 \pm 0.013 \\ 0.4149 \pm 0.014 \\ 0.4124 \pm 0.012 \\ 0.4210 \pm 0.016 \\ 0.4034 \pm 0.012 \\ 0.4366 \pm 0.014 \end{array}$	$\begin{array}{c} 0.3663{\pm}0.015\\ 0.3740\ {\pm}0.011\\ 0.3781{\pm}0.011\\ 0.3755{\pm}0.014\\ 0.3821{\pm}0.017\\ 0.3741{\pm}0.012\\ 0.3948{\pm}0.012 \end{array}$	$\begin{array}{c} 0.4021{\pm}0.014\\ 0.4112{\pm}0.015\\ 0.4163{\pm}0.012\\ 0.4134{\pm}0.013\\ 0.4205{\pm}0.013\\ 0.4023{\pm}0.014\\ 0.4286{\pm}0.014\\ \end{array}$	$\begin{array}{c} 0.3626 {\pm} 0.016 \\ 0.3708 {\pm} 0.017 \\ 0.3754 \ {\pm} 0.013 \\ 0.3715 \ {\pm} 0.014 \\ 0.3781 \ {\pm} 0.015 \\ 0.3747 {\pm} 0.014 \\ 0.3814 {\pm} 0.011 \end{array}$	$\begin{array}{c} 0.3658 \pm 0.013 \\ 0.3746 \pm 0.012 \\ 0.3793 \pm 0.011 \\ 0.3786 \pm 0.016 \\ 0.3814 \pm 0.016 \\ 0.3711 \pm 0.012 \\ 0.3867 \pm 0.012 \end{array}$	$\begin{array}{c} 0.3189{\pm}0.012\\ 0.3257{\pm}0.014\\ 0.3295{\pm}0.013\\ 0.3274{\pm}0.013\\ 0.3318{\pm}0.012\\ 0.3224{\pm}0.013\\ 0.3360{\pm}0.011 \end{array}$	0.3722±0.012 0.3812±0.014 0.3846±0.013 0.3725±0.014 0.3858±0.014 0.3658±0.012 0.3884±0.013	$\begin{array}{c} 0.3414 {\pm} 0.012 \\ 0.3445 {\pm} 0.012 \\ 0.3463 {\pm} 0.013 \\ 0.3350 {\pm} 0.012 \\ 0.3313 {\pm} 0.011 \\ 0.3320 {\pm} 0.014 \\ 0.3472 {\pm} 0.013 \end{array}$	
ODE	LT-OCF GRU-ODE	0.4267±0.013 0.4398±0.014	$\substack{0.3785 \pm 0.015 \\ 0.3902 \ \pm 0.017}$	0.4141±0.016 0.4275±0.013	$\substack{0.3673\ \pm 0.014\\0.3792\pm 0.012}$	0.3848±0.012 0.3994±0.013	$\substack{0.3313 \pm 0.013 \\ 0.3417 \pm 0.015}$	0.3841±0.014 0.3856±0.014	0.3416±0.012 0.3455±0.012	
Proposed	E-NSDE	$0.4711 {\pm} 0.015$	$0.4112 {\pm} 0.013$	$0.4551 {\pm} 0.011$	$0.3982 {\pm} 0.016$	$0.4194 {\pm} 0.013$	$0.3637 {\pm} 0.015$	$0.4021 {\pm} 0.014$	$0.3621 {\pm} 0.012$	

Experimentation

●óo



- Uncertainty vs. Interaction gap: We further investigate the impact of the user-item interaction gap and the corresponding uncertainty in providing important and diverse items in Table (a).
- Further, ablation study on key components of the E-NSDE model in Table (b).
 - Table: (a) Diverse recommendations by E-NSDE; (b) Ablation of key components

Model	Important Movies (Genre)	Future Movie's Genre	NSDE
GRU-ODE	Dead Man Walking ('Drama') Richard III ('Drama') Mad Love ('Romance')	'Thriller' 'Drama' 'Mystery'	~
E-NSDE	GoldenEye ('Thriller') Taxi Driver ('Drama') Twelve Monkeys ('Sci_Fi')	'Crime' 'Sci_Fi'	۰ ۱
	(a)		

Ablation Study

NSDE	EDL	WBPR .	Performance		
			P@5	nDCG@5	
~			0.4065	0.3677	
\checkmark	\checkmark		0.4523	0.3962	
\checkmark		\checkmark	0.4120	0.3715	
√	√	\checkmark	0.4711	0.4112	
		(b)			

Proposed Approach: E-NSDE Experimentation OOO

ENSDE Experimentation Conclusions

Conclusions

 We propose a novel time-aware sequential recommendation model called E-NSDE.

 It includes NSDE module that enables continuous, time-aware modeling of user preferences, without requiring uniform time intervals.

 Further, monotonic network ensures the relationship between time intervals and uncertainty is properly reflected—longer time intervals increase model uncertainty.

Provides SOTA performance in multiple real-world datasets.

More detailed information will be in the Poster with ID: 96864 (Poster Session 2).

- We propose a novel time-aware sequential recommendation model called E-NSDE.
- It includes NSDE module that enables continuous, time-aware modeling of user preferences, without requiring uniform time intervals.
- Further, monotonic network ensures the relationship between time intervals and uncertainty is properly reflected—longer time intervals increase model uncertainty.
- Provides SOTA performance in multiple real-world datasets.

Poster

Introduction

Conclusions

More detailed information will be in the Poster with **ID: 96864 (Poster Session 2)**.