



One Person, One Model, One World

⌵ Domain	RecSys
⋮ tag	
⌵ Conference / Journal	SIGIR
≡ Publish year	2021
📅 정리 날짜	@2024년 1월 21일

One Person, One Model, One World: Learning Continual User Representation without Forgetting

Summary

Background & Motivation

User modeling on same platform

- 하나의 플랫폼 내에서만 작동

Previous Generic representation of user

- Upstream recommendation task로 학습 후 transfer learning
- Dissimilar downstream task로 fine-tuning

- PeterRec
 - Task마다 개별적으로 model parameter 학습
 - Related downstream tasks에 대해 학습하여 parameter 따로 저장
 - Individual train이 storage overhead 발생시킴
 - task간에 substantial similarity가 있으므로, inefficient + suboptimal
- Single model learning
 - Parameter 따로 저장하지 않으므로 storage는 괜찮지만
 - Catastrophic forgetting 발생함
 - 이전의 well-trained task 잊어버림
- Multi-task learning
 - Related tasks에 학습 진행
 - Training data for all tasks가 inavailable할 수 있음
 - 새로운 task가 overlapping user 없을 수 있음

One Person, One Model, One World

- 한 사람당 하나의 Backbone network architecture로 표현
 - 얼마나 많은 role이건 관계없이 하나의 model로 표현
- Universal user representation + Continuous learning throughout lifespan
 - Quick learning + Immune to forgetting
 - Old task의 important만 재활용해서 학습
- Conure
 - Continual, or lifelong, User Representation learning framework
 - 전제로서 확인한 insight
 1. Single ANN sequentially task 학습하는 것은 catastrophic forgetting 발생 시킴
 2. Well-trained model에서 Unimportant parameter 지우는 것은 성능저하 X

Related Works

User Modeling and Recommendation

- UM: Process of obtaining the user profile: understanding of the user
 - User representation(Well trained user model)을 학습하는 것
 - Supervised learning기반 학습은 specific task만 잘 함
 - Generalize가 안됨
 - e.g. (Deep) Neural Factorization Machine
 - PeterRec: Self-supervised(unsupervised) pretraining
 - Pretraining: behavior sequences의 inherent relation을 modeling
→ user representation become universal
 - 일회성임
 - T1 → T2만 가능하고, 이후 T2→T3 등 무한히 사용 불가
 - Conure: Extension of PeterRec

Transfer Learning

- Two-stage training paradigm
 - Pretraining: Source dataset으로 학습
 - Conure: TCN(Temporal Convolutional Network)사용
 - Finetuning: Target dataset으로 학습, init시 pretrained param 사용
 - 네가지 종류
 1. Pretrained network의 Softmax layer만 finetuning as a feature extractor
 2. higher layer를 finetuning, bottom layers frozen
 3. 전체 pretrained model
 4. PeterRec같이 새로 추가된 adaptor network만

Continual Learning

- Lifespan 동안 지속적으로 학습할 수 있는 능력
- CV나 robot learning 등에서는 많이 연구됐지만, RecSys 분야에서는 task-level 없었음
 - RecSys에서는 same task 한해서만 좀 있었음
 - Conure: Different task에서 continual learning

Preliminaries

Task formation

- Task $\mathbf{T} = \{T_1, T_2, \dots, T_N\}$ 와 user $u \in \mathcal{U}$ 에 대해
 - user's interaction sequence $\mathbf{x}^u = \{x_0^u, \dots, x_n^u\}$ for $(u, \mathbf{x}^u) \in T_1$
 - supervised label $y \in \mathcal{Y}$ for $(u, y) \in T_i$
 - $T_i \in \{T_2, \dots, T_N\}$ 은 profile prediction (e.g., gender)나 item recommendation같은 task
 - user u 가 g 개의 y 가질 수 있음
 - Conure가 \mathbf{T} 전체에 대해 학습한 뒤, \mathbf{T} 내의 모든 task에 대해 한 개의 모델로서 사용 가능

Methodology

Questions

