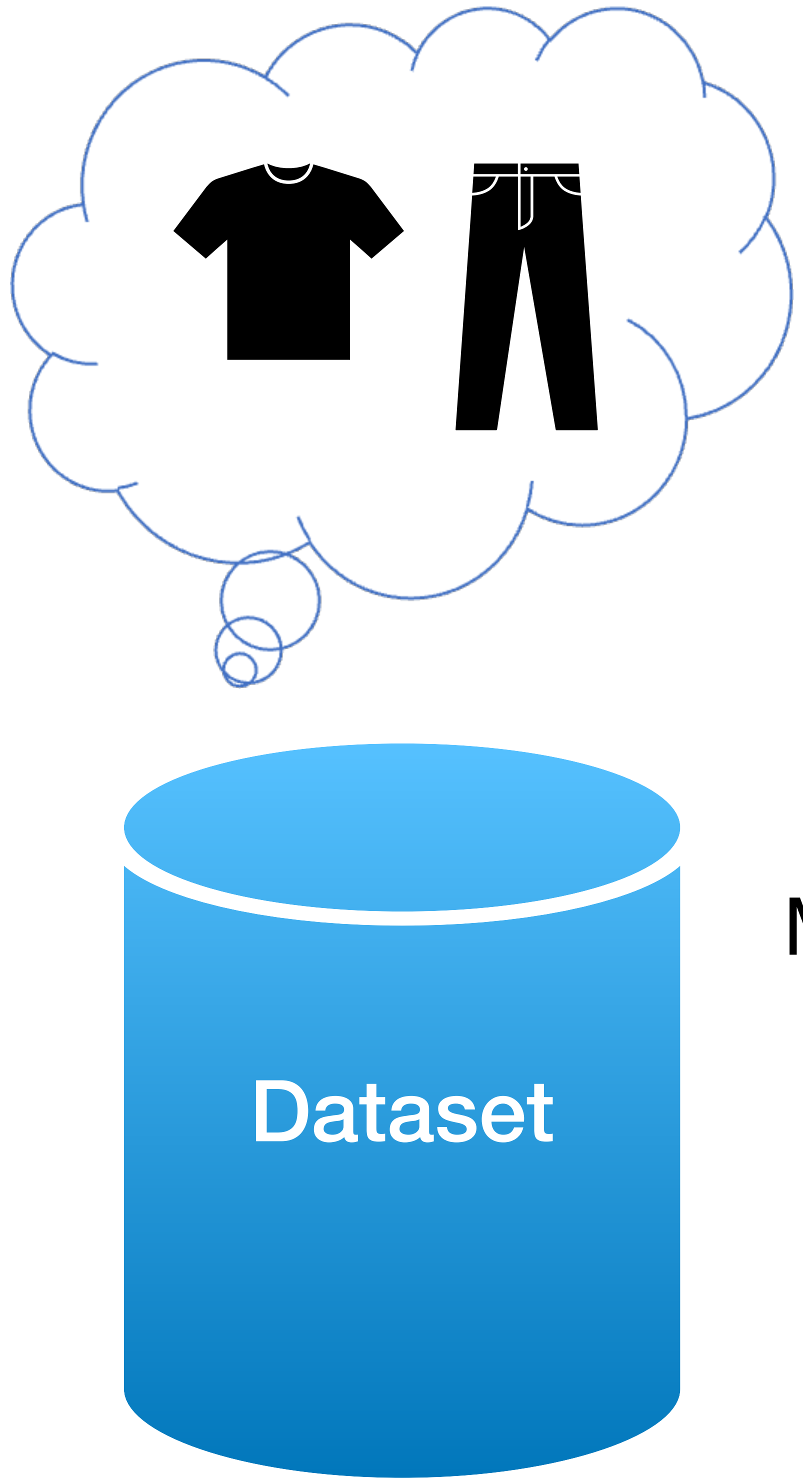
TimeKit: A Time-series Forecasting-based Upgrade Kit for Collaborative Filtering

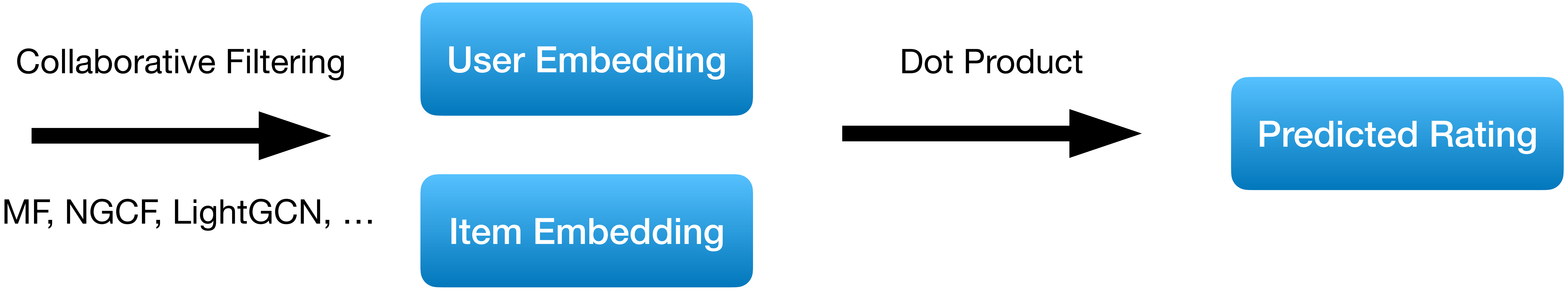
Yonsei University, South Korea

Seoyoung Hong, Minju Jo, Seungji Kook, Jaeeun Jung, Hyowon Wi, Noseong Park, Sung-Bae Cho

Introduction

Motivation (1/5)



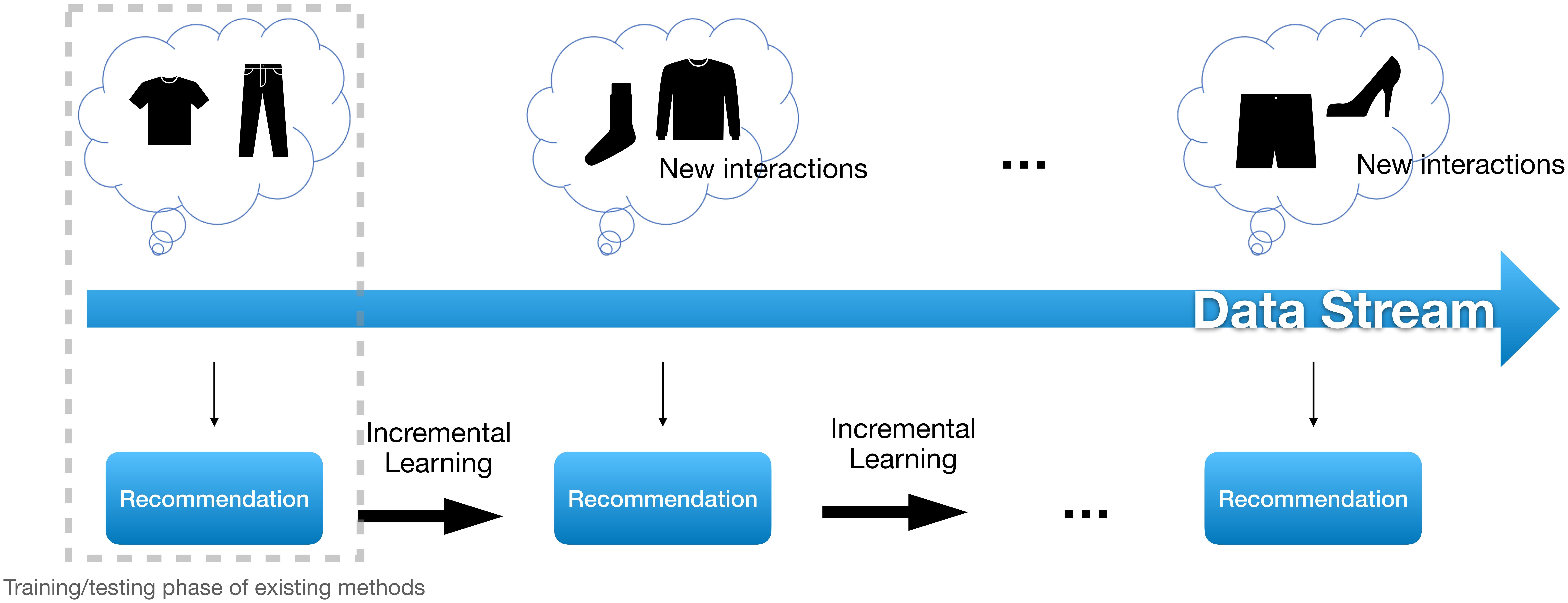


Existing collaborative filtering methods assume a single training!

Introduction

Motivation (2/5)

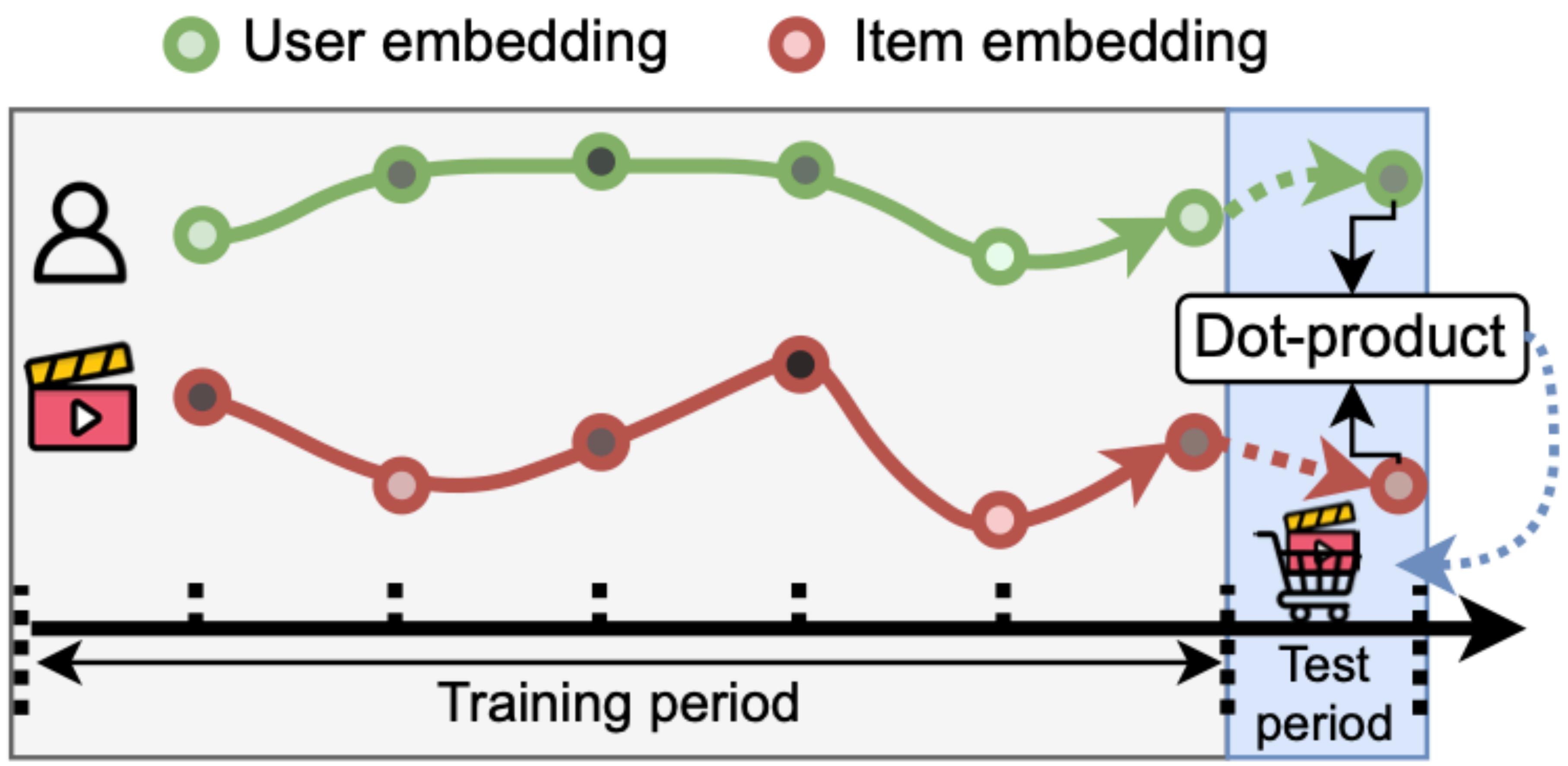
However, in the real world, we need to train a model periodically as new interactions are incremental!



Introduction

Motivation (3/5)

- A time-series of embedding vectors can be naturally defined as we need to periodically train a collaborative filtering algorithm.

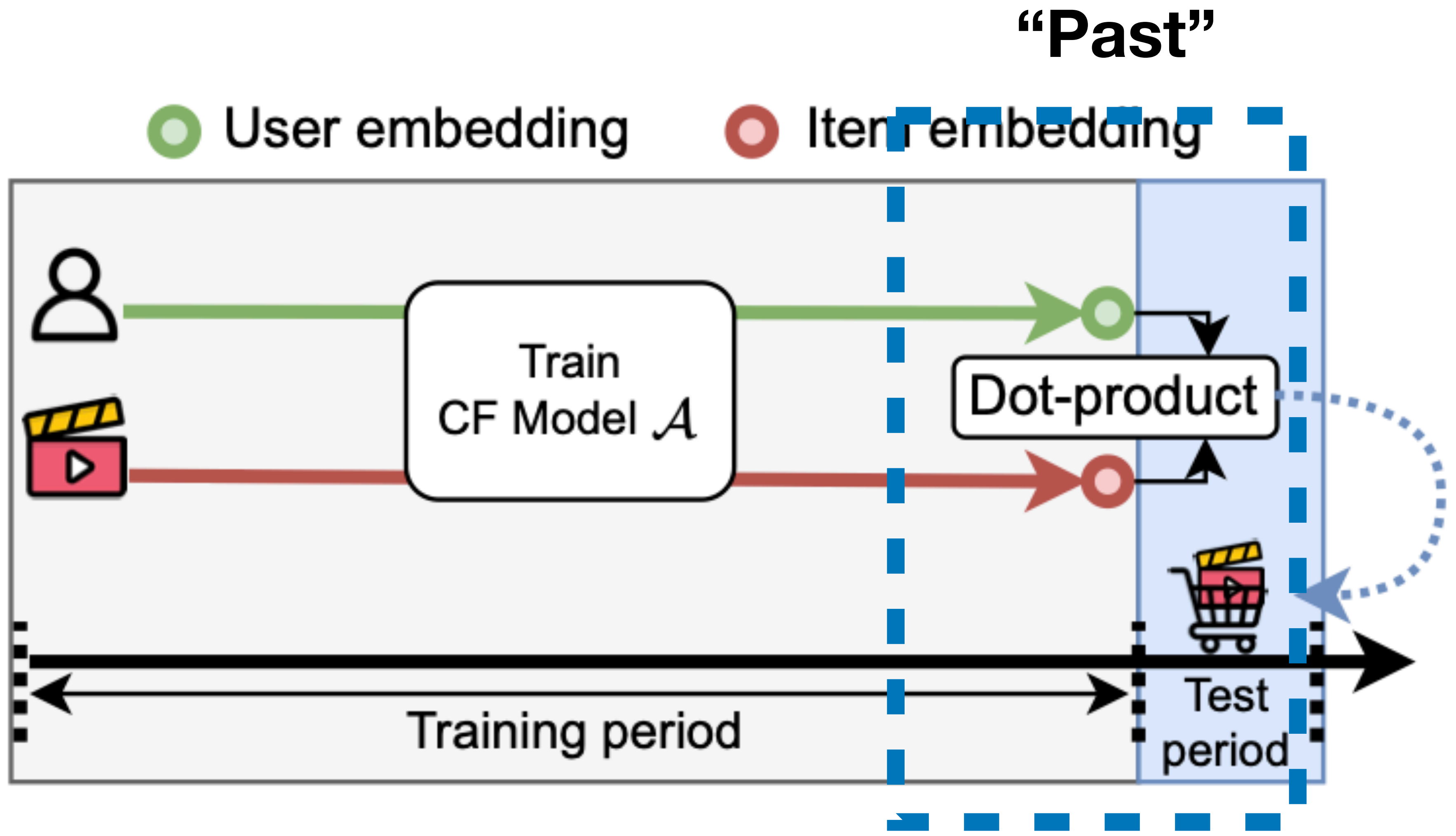


Proposed method, 'TimeKit'

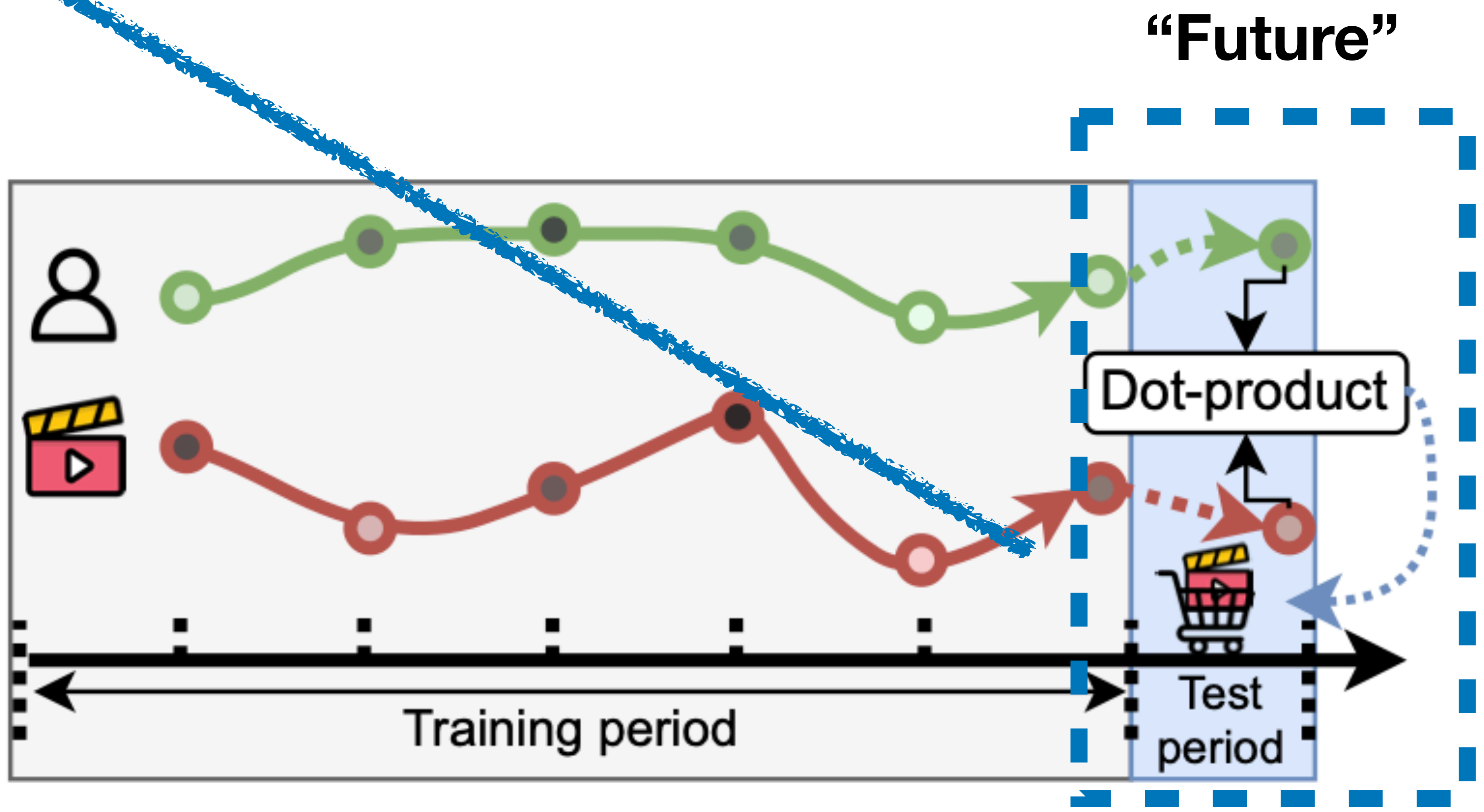
Introduction

Motivation (4/5)

- We propose forecasting <u>future embeddings</u> which describe latent behavioral patterns.



Existing collaborative filtering methods



Proposed method, ‘TimeKit’

Introduction

Motivation (5/5)

A Time-series Forecasting-based Upgrade Kit for Collaborative Filtering

'TimeKit'

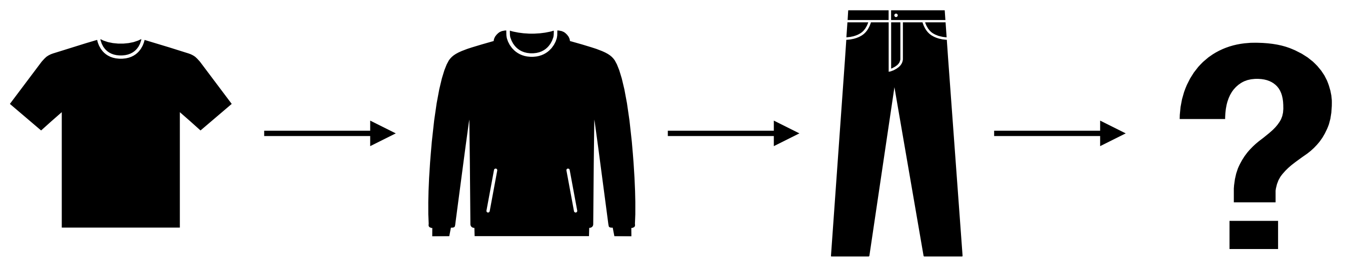
➡️ It is for an incremental recommendation and any CF methods can be used!

Related Work

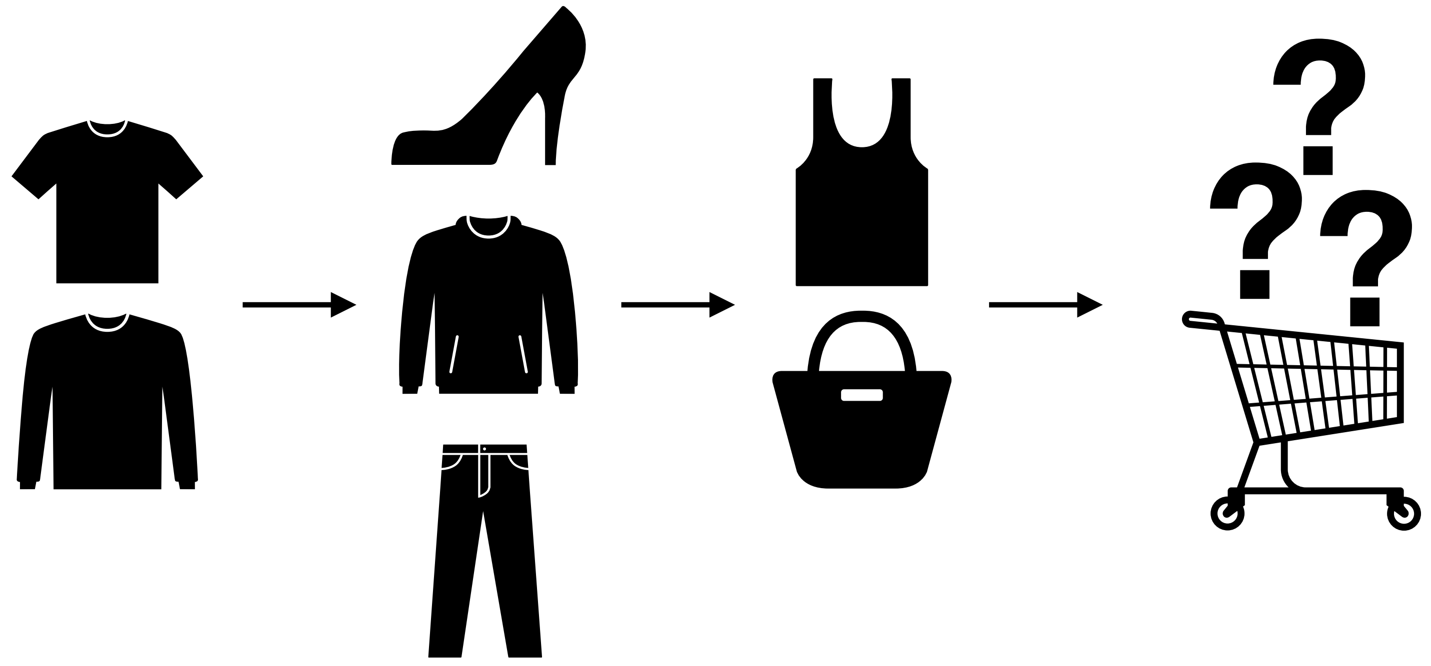
Sequential Recommendation

- The comparison between sequential recommendation and our proposed TimeKit

Sequential Recommendation



TimeKit



	Sequential Recommendation	TimeKit
What does model predict?	Next items	Embeddings
What is the input of model?	Sequence of items	Sequence of embeddings
What type of prediction?	Point-wise	Region-wise

Related Work

Neural Controlled Differential Equations (1/2)

- Neural Ordinary Differential Equations (NODEs) [1]
  - A continuous version of ResNet!
  - NODEs learn differential equations as a neural network.

$$\boldsymbol{z}(T) = \boldsymbol{z}(0) + \int_0^T f\big(\boldsymbol{z}(t);\, \boldsymbol{\theta}_f\big)dt$$

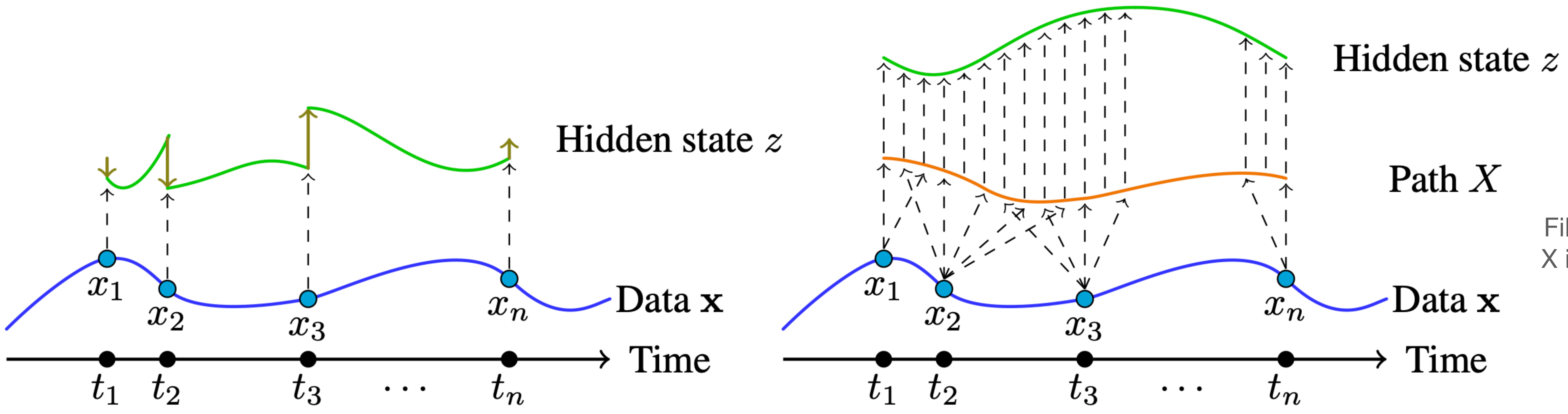
  - where $f(\mathbf{z}(t); \theta_f)$, which we call the ODE function, is a neural network to learn $\frac{d\mathbf{z}(t)}{dt}$.

[1] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural ordinary differential equations," Advances in neural information processing systems, 31, 2018.

Related Work

Neural Controlled Differential Equations (2/2)

- Neural Controlled Differential Equations (NCDEs) [2]
  - A continuous version of RNN!
  - Natural cubic splines were used to construct path X from the time-series x.



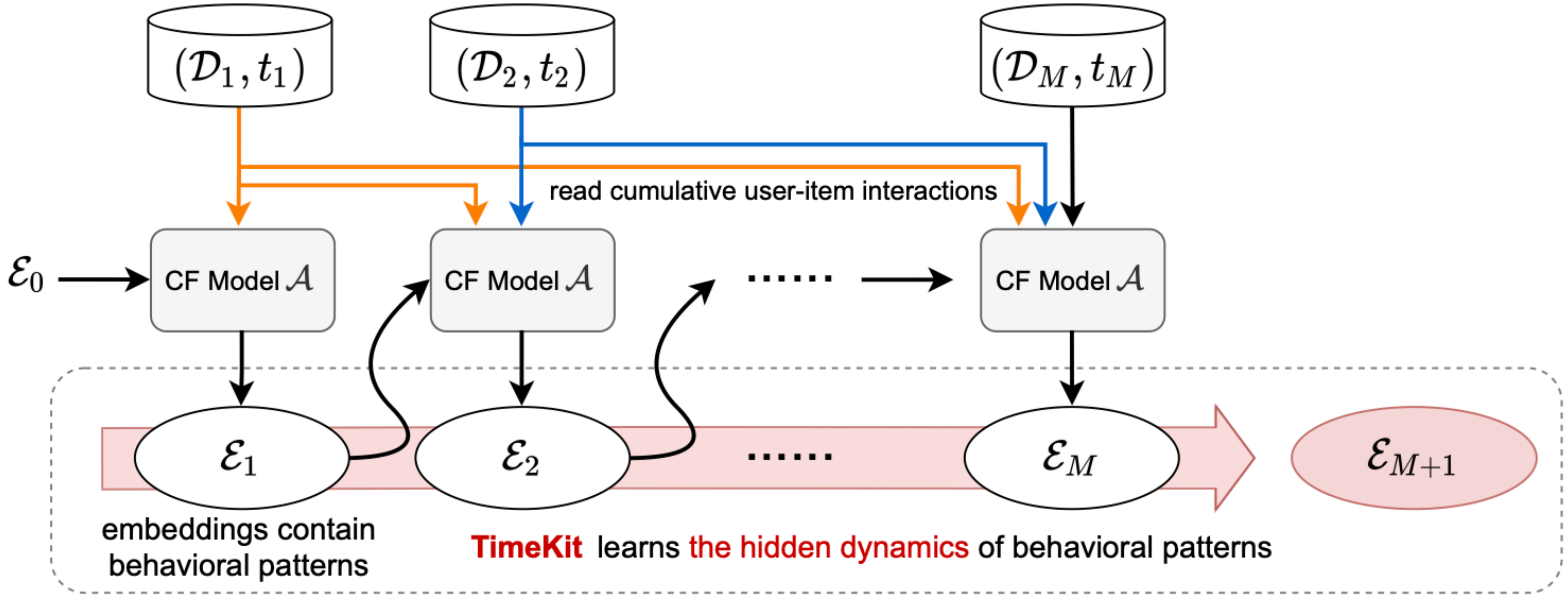
Fills missing values and ensures that X is twice continuously differentiable!

[2] P. Kidger, J. Morrill, J. Foster, and T. Lyons, "Neural controlled differential equations for irregular time series," Advances in Neural Information Processing Systems, 33, 6696-6707, 2020.

Method

Overall Workflow

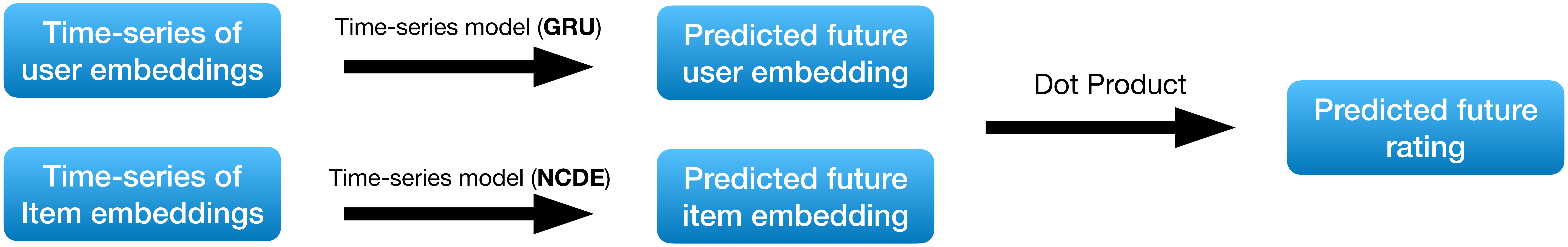
- The overall workflow of creating the time-series of embedding vectors and forecasting



Method

Embedding Forecaster

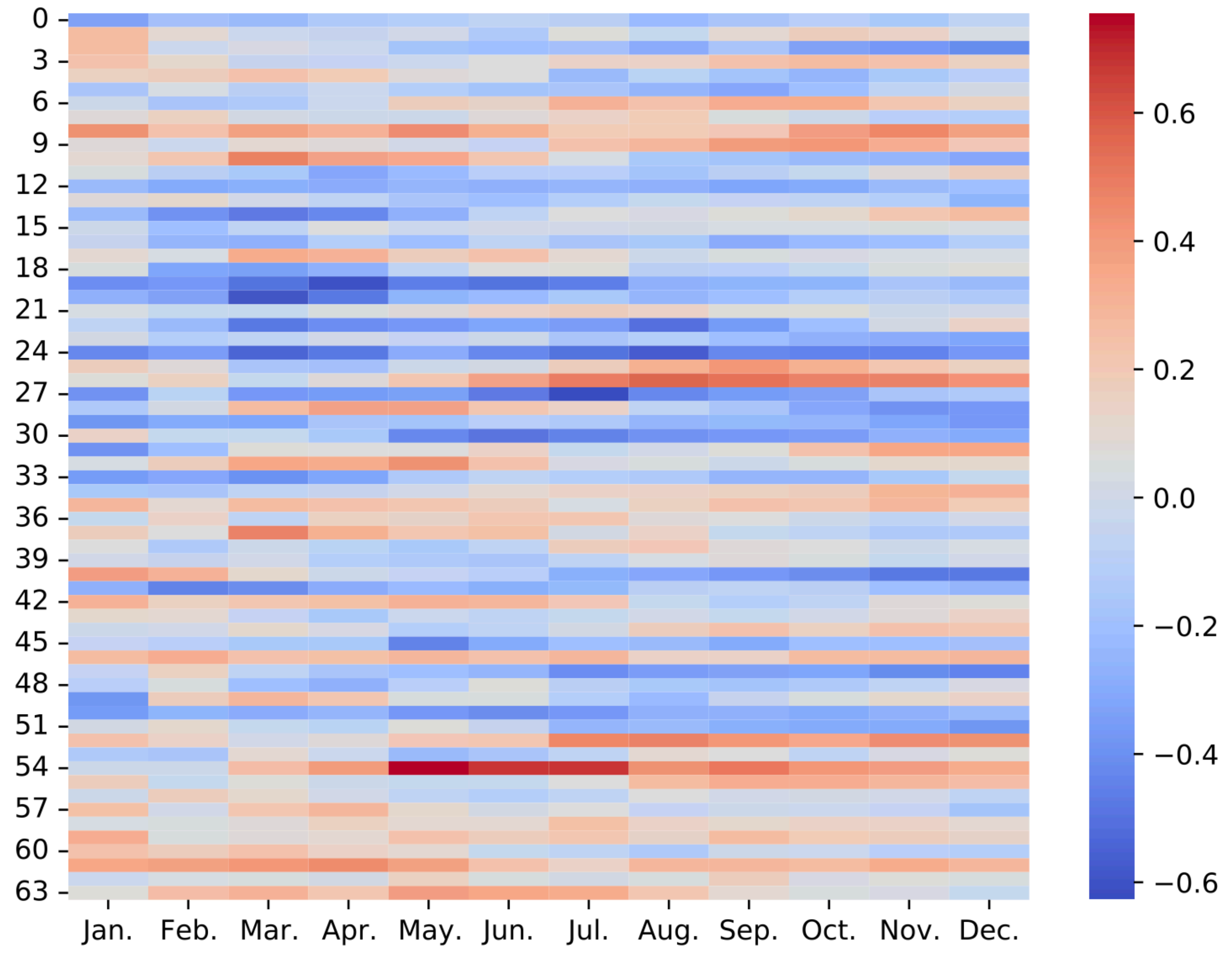
- From our pre-experiments, we figured out that item embedding vectors are more sensitive to noise, which is more difficult to predict.
- So we use a light-weighted model (GRU) for forecasting user embeddings and a high-performance model (NCDE) for forecasting item embeddings.



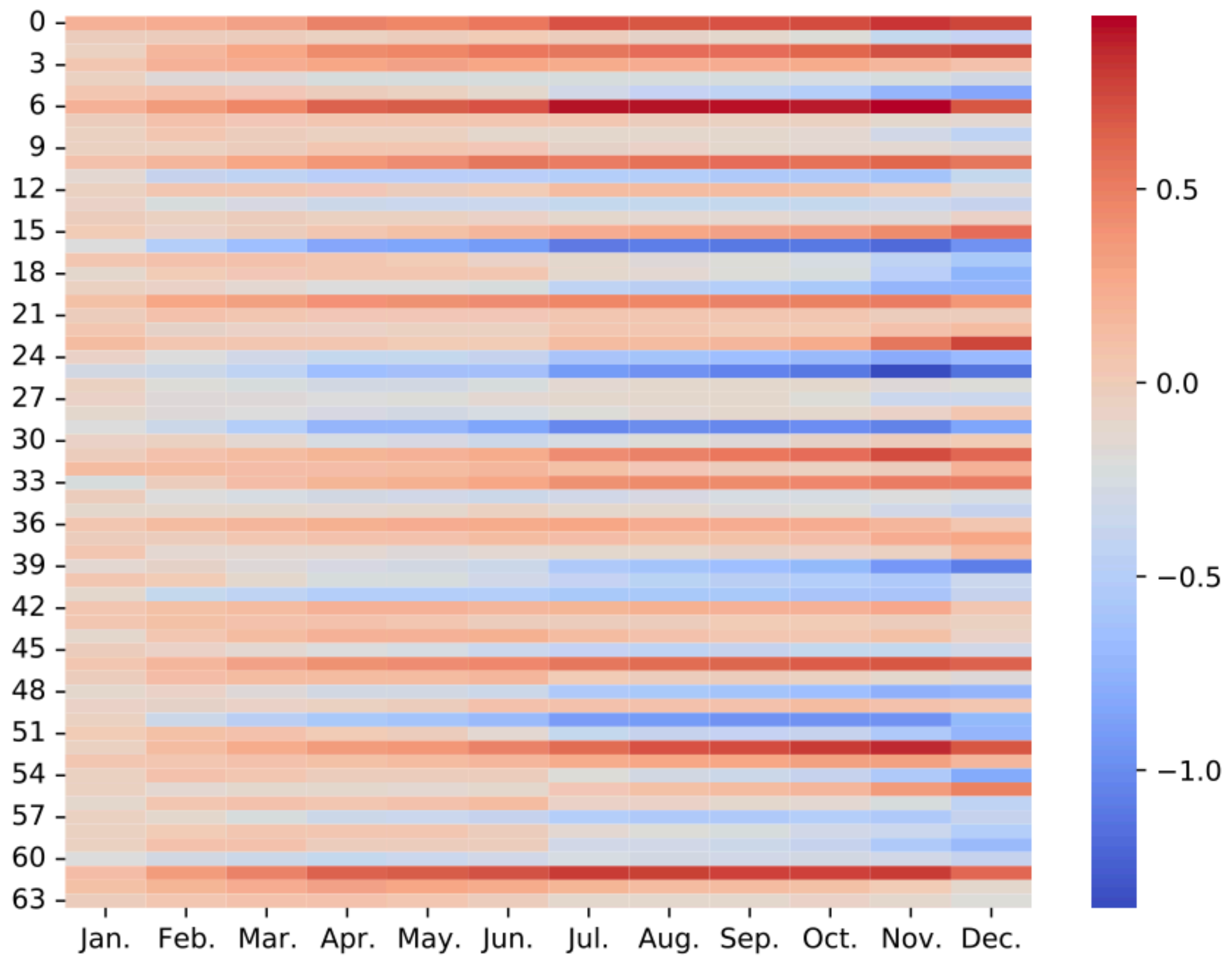
Experiments

The Visualization of Embedding Dynamics

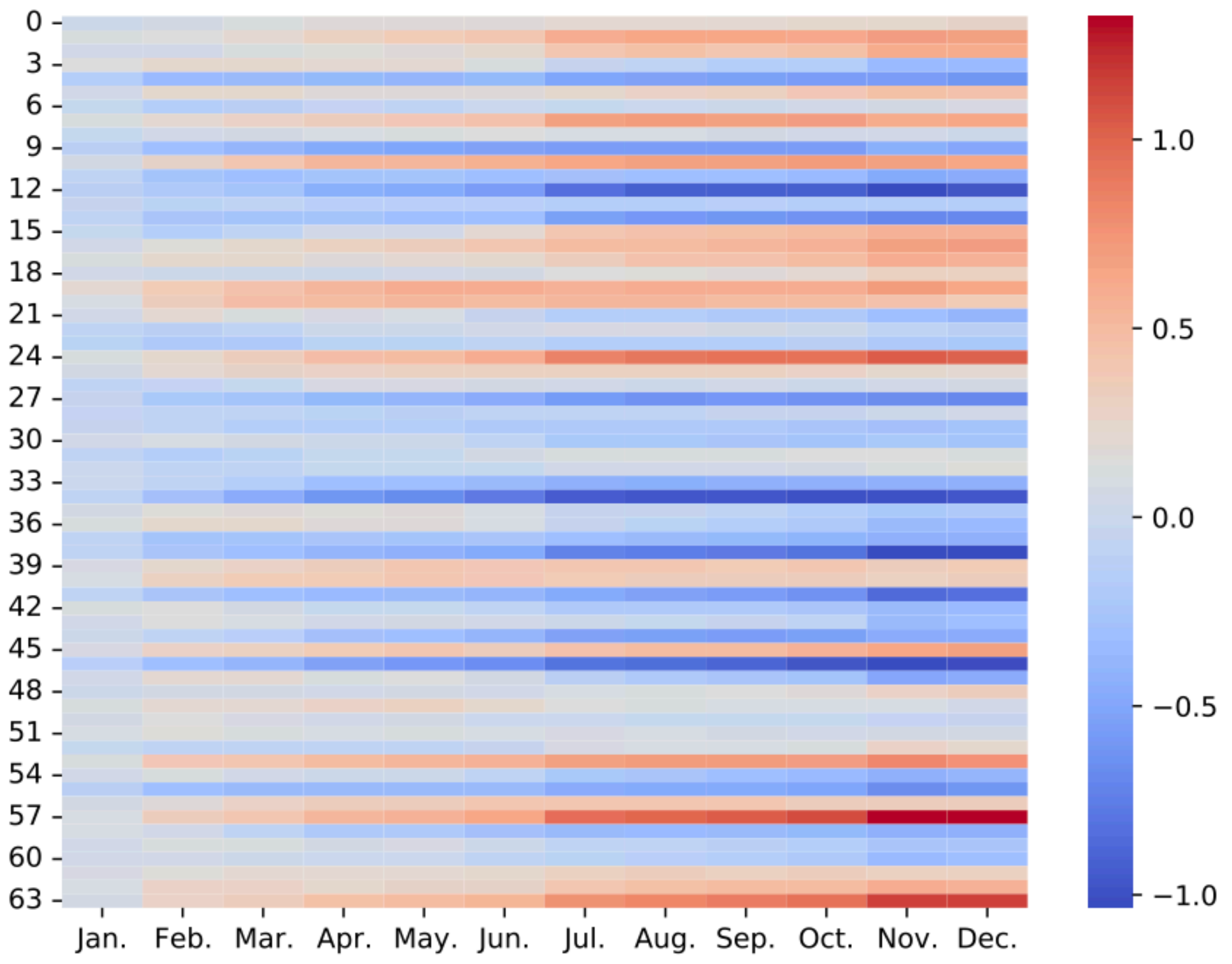
- We found that embedding vectors do not change randomly over time but are likely to change following hidden dynamics.



(a) User embedding example of LightGCN



(b) User embedding example of NGCF



(c) Item embedding example of NGCF

Experiments

Performance (1/2)

- Performance on Amazon-book dataset
  - Especially in the case of BPRMF + TimeKit, it even exceeds the original scores of LightGCN and LT-OCF.

Base CF algorithm	Base performance		+ TimeKit		Improvement (%)	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
BPR-MF	0.0496	0.0253	**0.0575**	**0.0303**	15.88	19.50
NGCF	0.0359	0.0181	**0.0471**	**0.0252**	31.12	39.58
LightGCN	0.0565	0.0294	**0.0667**	**0.0365**	17.93	24.17
LT-OCF	0.0573	0.0299	**0.0711**	**0.0393**	24.11	31.40

Experiments

Performance (2/2)

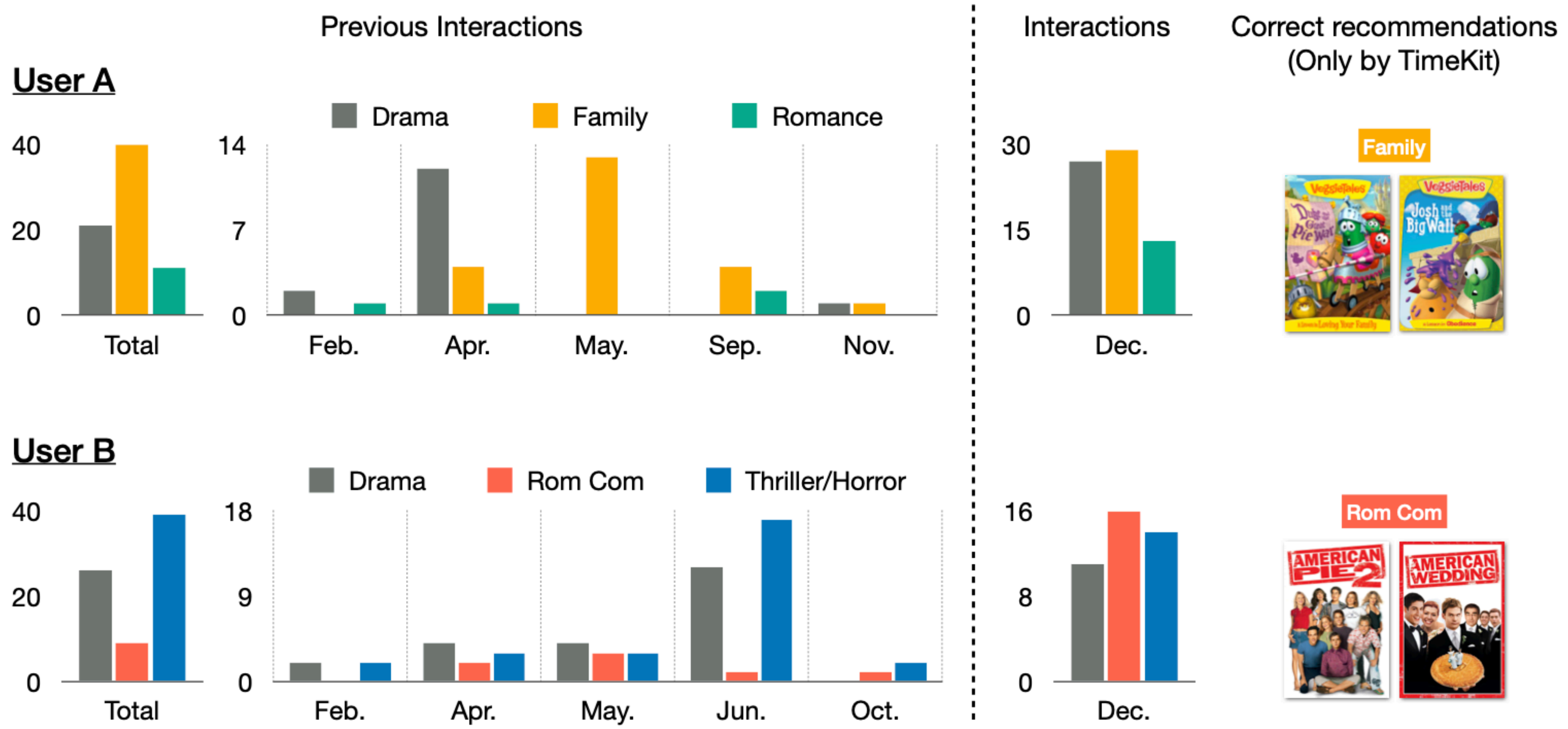
- Performance on Netflix dataset
  - TimeKit improves NGCF by 68% for Recall@20 and NDCG@20.

Base CF algorithm	Base performance		+ TimeKit		Improvement (%)	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
BPR-MF	0.0701	0.0405	**0.0776**	**0.0466**	10.74	15.07
NGCF	0.0608	0.0380	**0.1024**	**0.0642**	68.46	68.75
LightGCN	0.0787	0.0451	**0.0823**	**0.0480**	4.52	6.33
LT-OCF	0.0779	0.0446	**0.0899**	**0.0539**	15.41	20.92

Experiments

Case Study

- Our model is able to correctly recommend items because it is trained with the latent dynamics of the preference distribution changes over time.



Conclusion

- We presented a novel upgrade kit called TimeKit to forecast the future user/item embedding vectors, with which we will perform the collaborative filtering task.
- We uncover the hidden dynamics to accurately forecast the future embedding vectors.
- Various collaborative filtering algorithms are significantly improved when being upgraded with TimeKit.

Thank you!