

# Informative Ensemble of Multi-Resolution Dynamics Factorization Models

InnerPeace Team<sup>1,2</sup>

<sup>1</sup>Shanghai Jiao Tong University, <sup>2</sup>Hong Kong University of Science and Technology



## Introduction

- task: Yahoo! Music user rating prediction
- major challenges:
  - rich information(temporal information, taxonomy information)
  - large training data size(200M ratings,1M users, 0.6M items)

## Temporal Dynamics Modeling

- observations
  - item/user bias may change with time
  - users' preference may change with time
  - users' recent rating history is more important

## Time Dependent Bias

- time bin bias

$$a_u(d, t) = a_u + a_{u, Bin(d)} + a_{u, Bin(t)} \quad (1)$$

- tensor based bias

$$a_u(d, t) = w_u^T u(d) + w_u^T v(t) + u(d)^T v(t) \quad (2)$$

## Time Dependent Factor

- piecewise linear factor

$$p_u(d, t) = \frac{(d - d^-)p_u(d^-) + (d^+ - d)p_u(d^+)}{\delta_{date}} \quad (3)$$

- time center factor

$$p_u(t) = e^{-\beta_u |t - c(u)|} p_u^{(c)} \quad (4)$$

## Time Dependent Neighborhood

$$\hat{r}_{ui} = a_u + b_i + p_u^T q_i + |N(u, i; k)|^{-\frac{1}{2}} \sum_{j \in N(u, i; k)} e^{-\alpha_u |\Delta t|} (r_{uj} - \bar{b}_u) \quad (5)$$

## Taxonomical Information

- category and user artist bias

$$\hat{r}_{ui} = b_i + a_u + a_{u, Cat(i)} + a_{u, Art(i)} + p_u^T q_i \quad (6)$$

- taxonomy neighborhood

$$\hat{r}_{ui} = a_u + b_i + p_u^T q_i + w_i(r_{u, Art(i)} - \bar{b}_u) \quad (7)$$

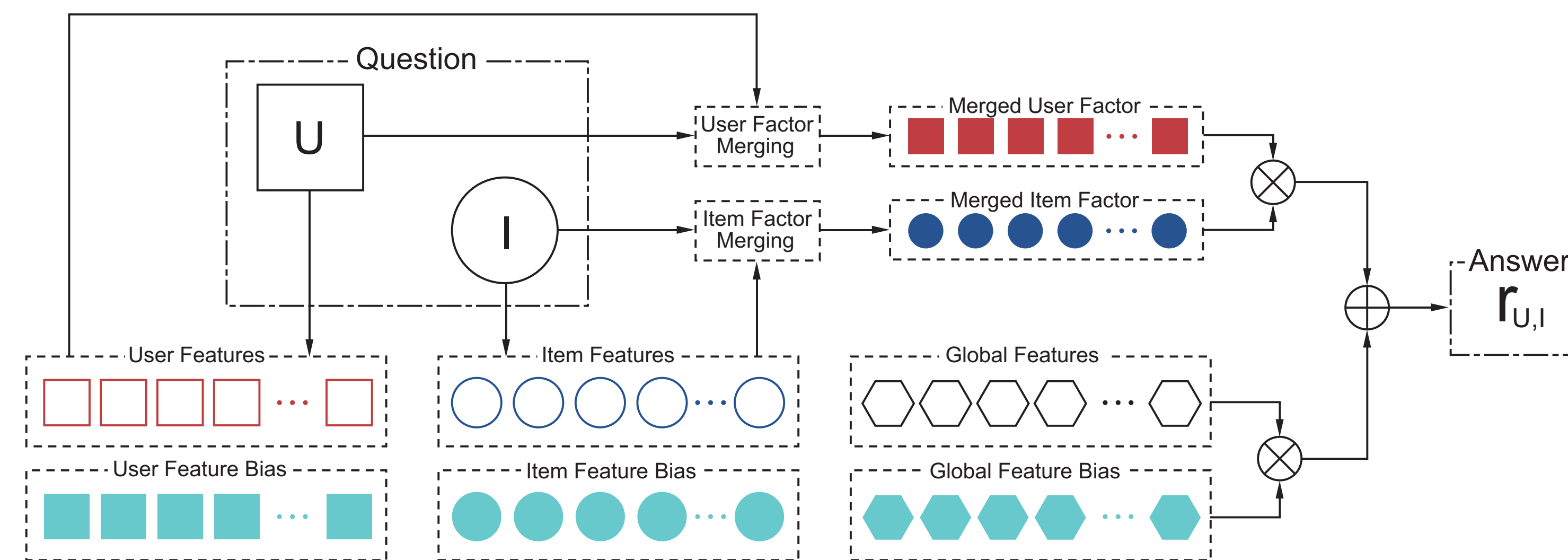
- taxonomy aware predictor

$$\hat{r}_{ui}(d, t) = \omega \cdot \check{r}_{ui}(d, t) + \frac{1 - \omega}{|\mathcal{A}_i|} \cdot \sum_{j \in \mathcal{A}_i} \check{r}_{uj}(d, t) \quad (8)$$

## Implementation

### Feature-based Matrix Factorization

- to solve the problem of so many variants of models to implement
- an abstract model that include most variants

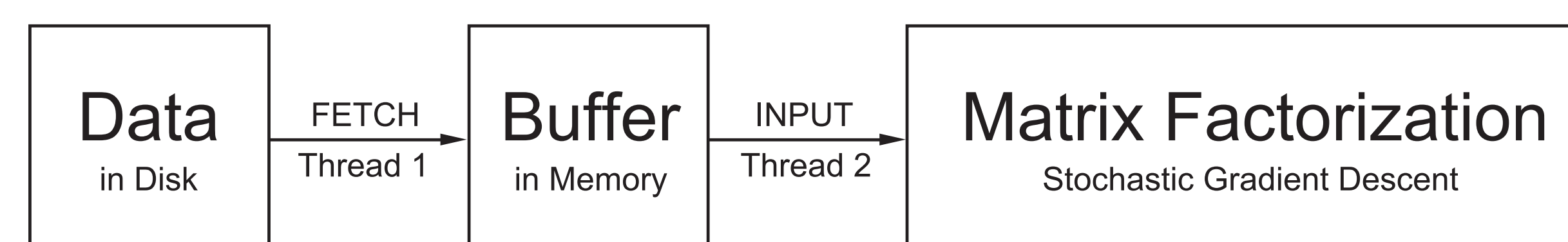


source code available at

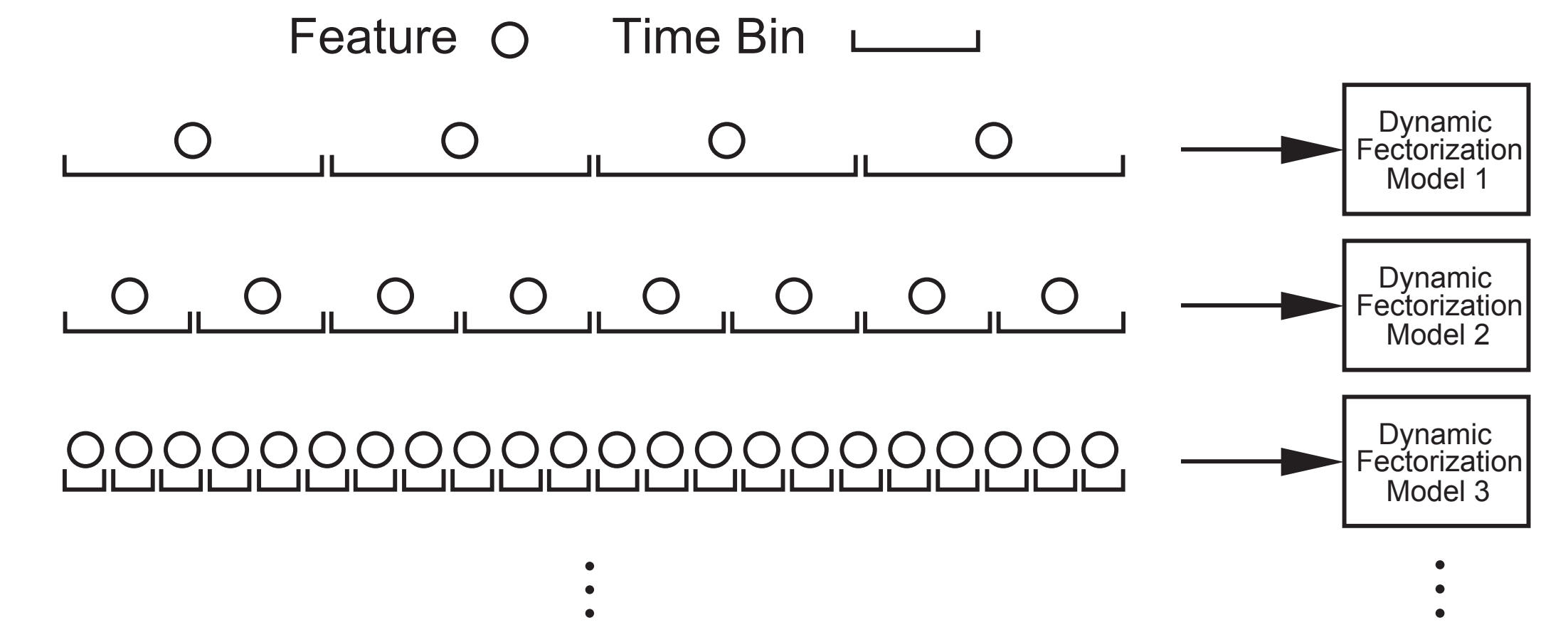
<http://apex.sjtu.edu.cn/apex/wiki/svdfeature>

### Input Buffering and Execution Pipeline

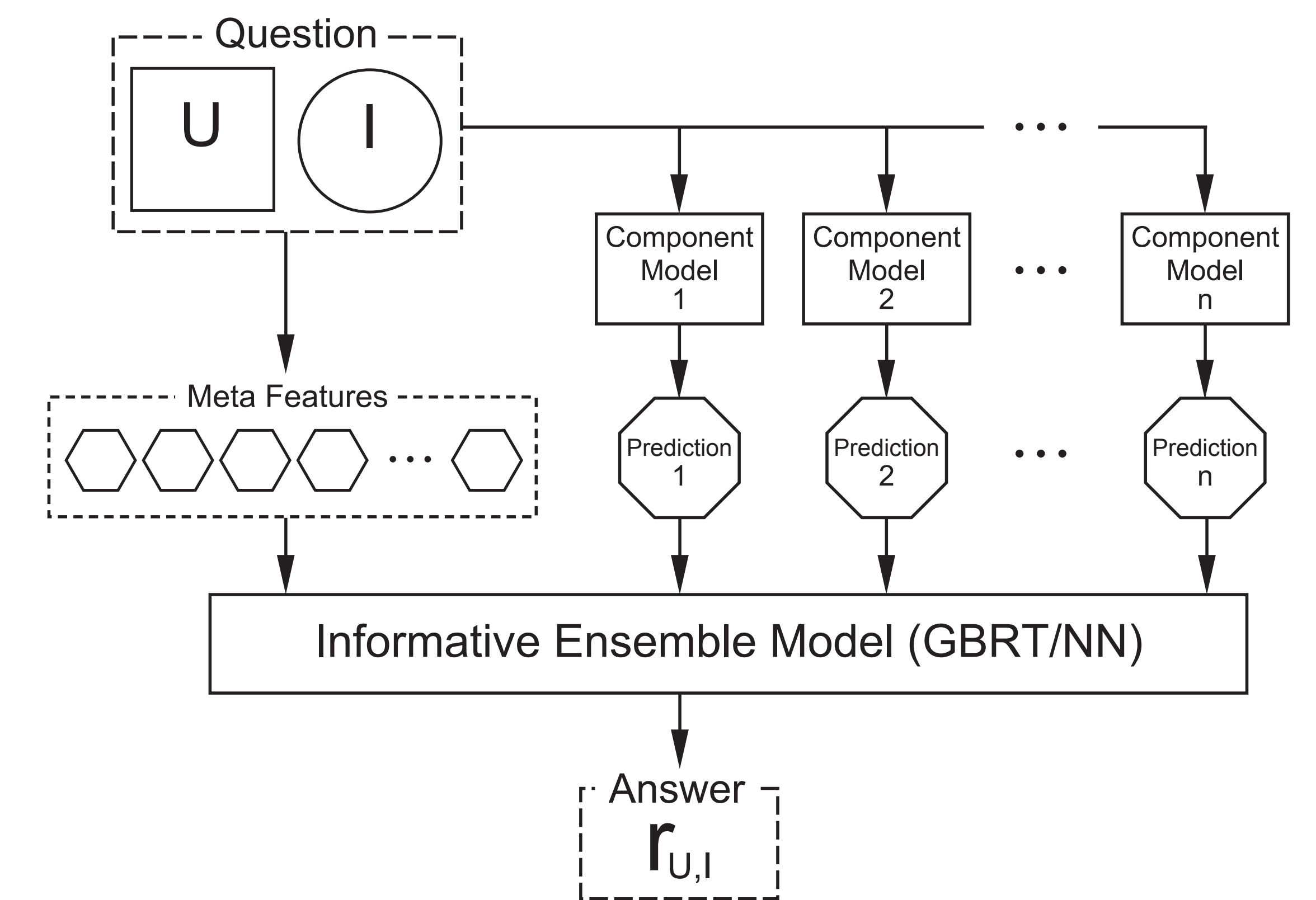
- to solve the problem of large-training data size
- do training efficiently without loading all data into memory



## Multi-Resolution Dynamics



## Informative Ensemble



- Examples of meta-features

- number of ratings of the user/item
- mean/variance of ratings of the user/item
- how far is prediction from training data in time
- number of available neighborhood information

## Experiment Result

Method	Test RMSE
best single method	22.06
informative ensemble(NN)	21.29
informative ensemble(GBRT)	21.32
final ensemble(NN+GBRT)	21.26