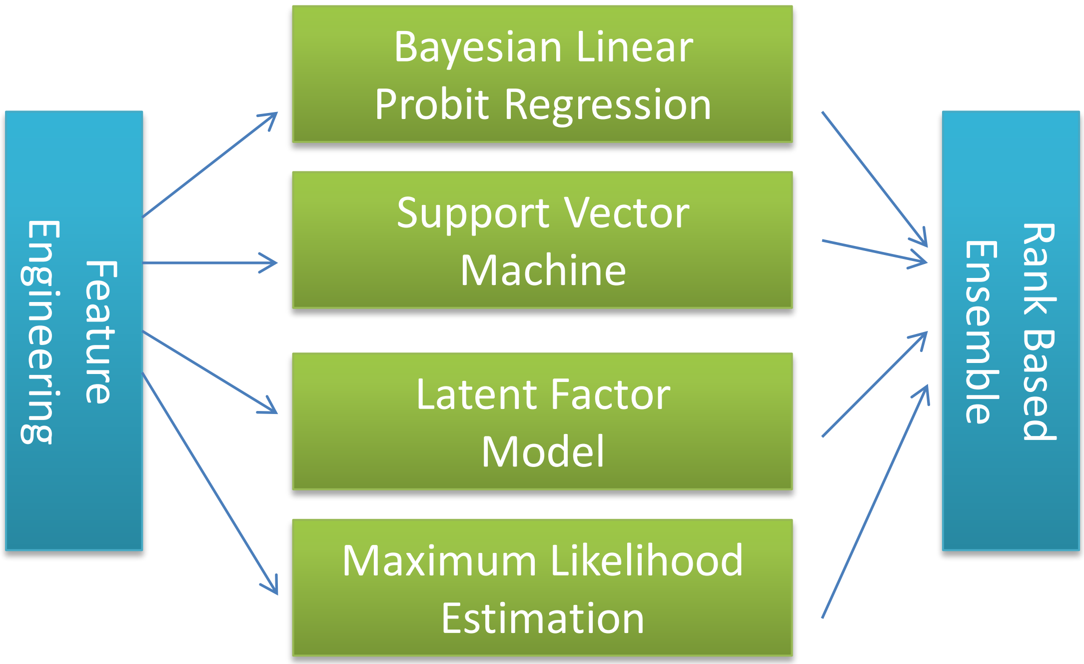
Click-Through Prediction for Sponsored Search Advertising with Hybrid Models

Chinese Academy of Science team

3rd place on the leaderboard in track 2

ACM KDD CUP Workshop 2012

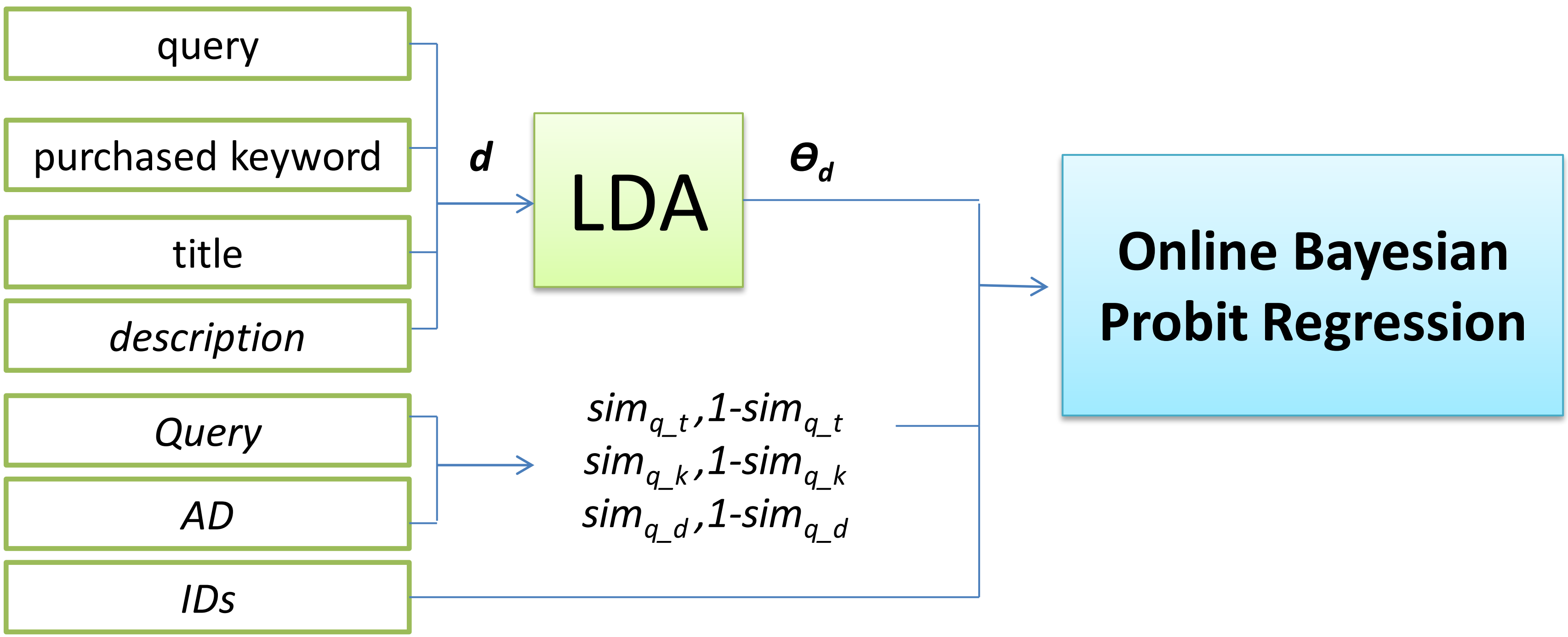
Ensemble method



Feature Engineering

- Original Features
  - Unique ID(ad, advertiser, query, user and etc.)
  - Ctr for each features
- Synthetic Features
  - 2-tuple features
  - Add position information to each feature
  - Token combination between features

Bayesian Linear Probit Regression



Gibbs sampling method to obtain the vector θ_d

Expectation propagation algorithm is used to get the approximate posterior

Bayesian Linear Probit Regression

- Features:
  - Discrete multi-value features

$$x := \left(x_1^T, \dots x_i^T\right)^T \text{and}$$

$$x_i = \begin{pmatrix} x_{i,j} \\ \vdots \\ x_{i,M_i} \end{pmatrix} \qquad \sum_{j=1}^{M_i} x_{i,j} = 1$$

- Labels:

$$\mathrm{y} \in \{-1,1\}$$

Bayesian Linear Probit Regression

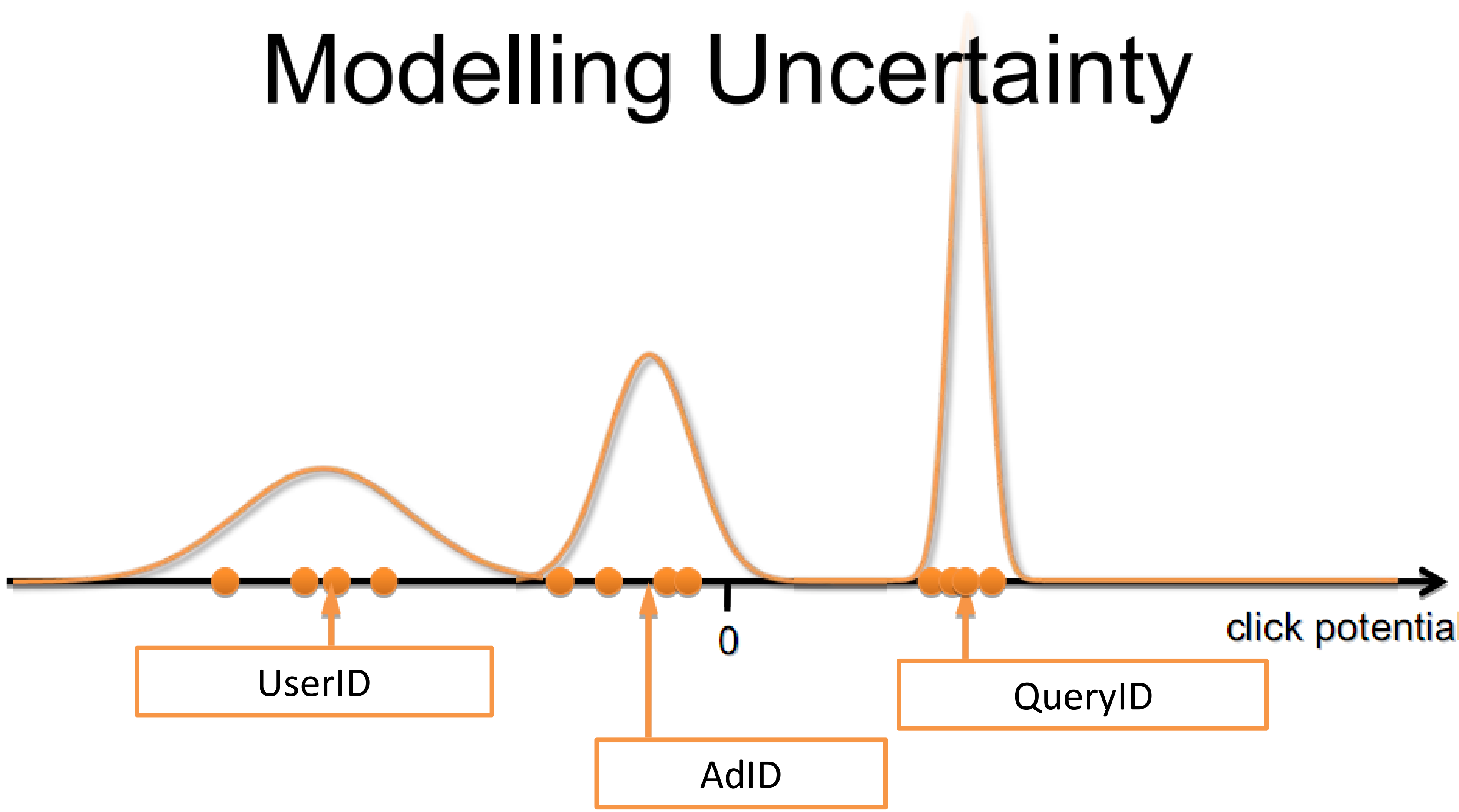
The Probit Regression

$$p(y|\boldsymbol{x}, \boldsymbol{w}) := \Phi\left(\frac{y \cdot \boldsymbol{w}^{\boldsymbol{T}} \cdot \boldsymbol{x}}{\beta}\right)$$

The Gaussian prior distribution over the weights

$$p(\boldsymbol{w}) = \prod_{i=1}^{N} \prod_{j=1}^{M_i} \aleph\left(w_{i,j}; u_{i,j}, \sigma_{i,j}^2\right)$$

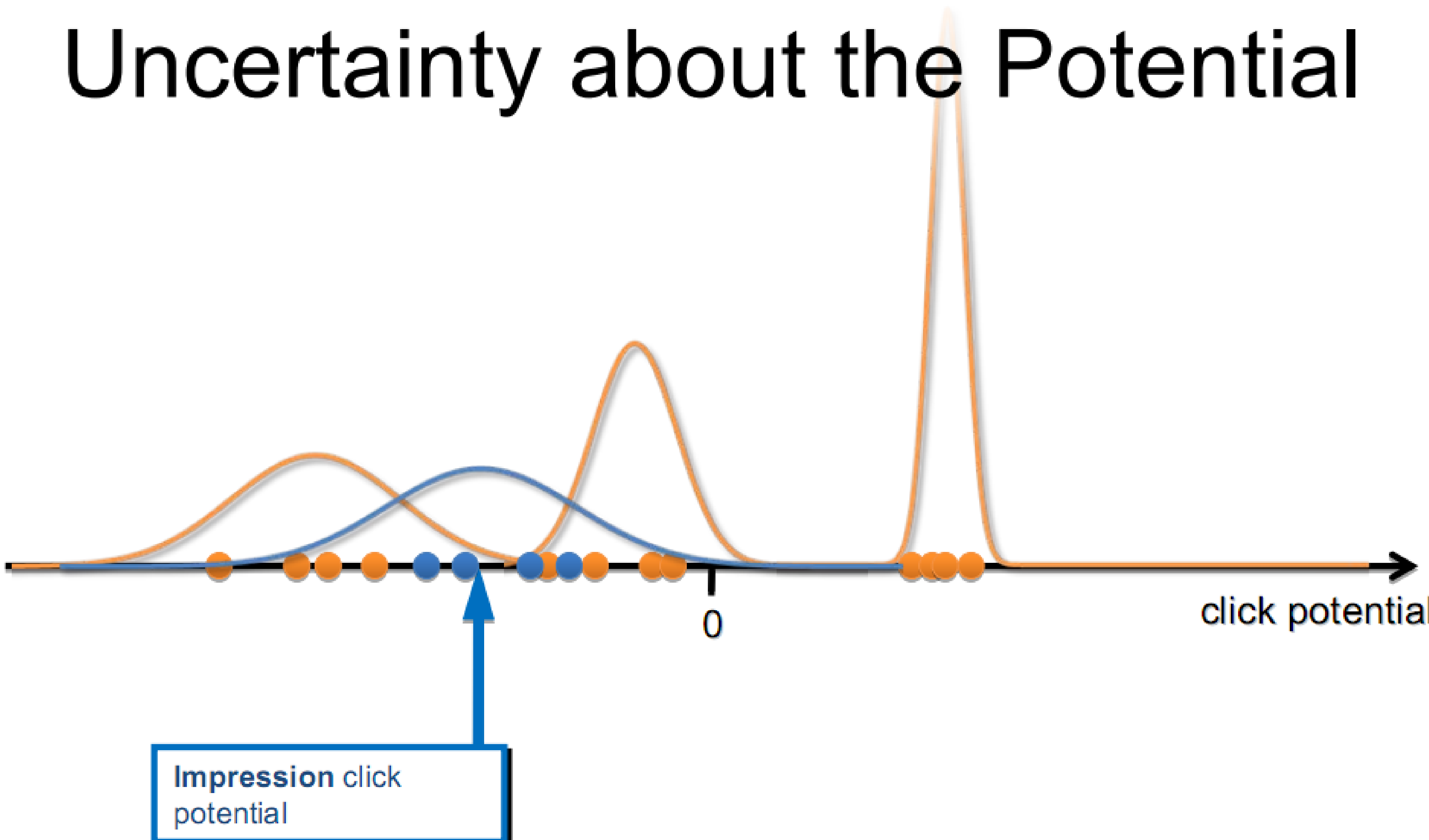
Bayesian Linear Probit Regression



--This page come from Thore Graepel's slides!

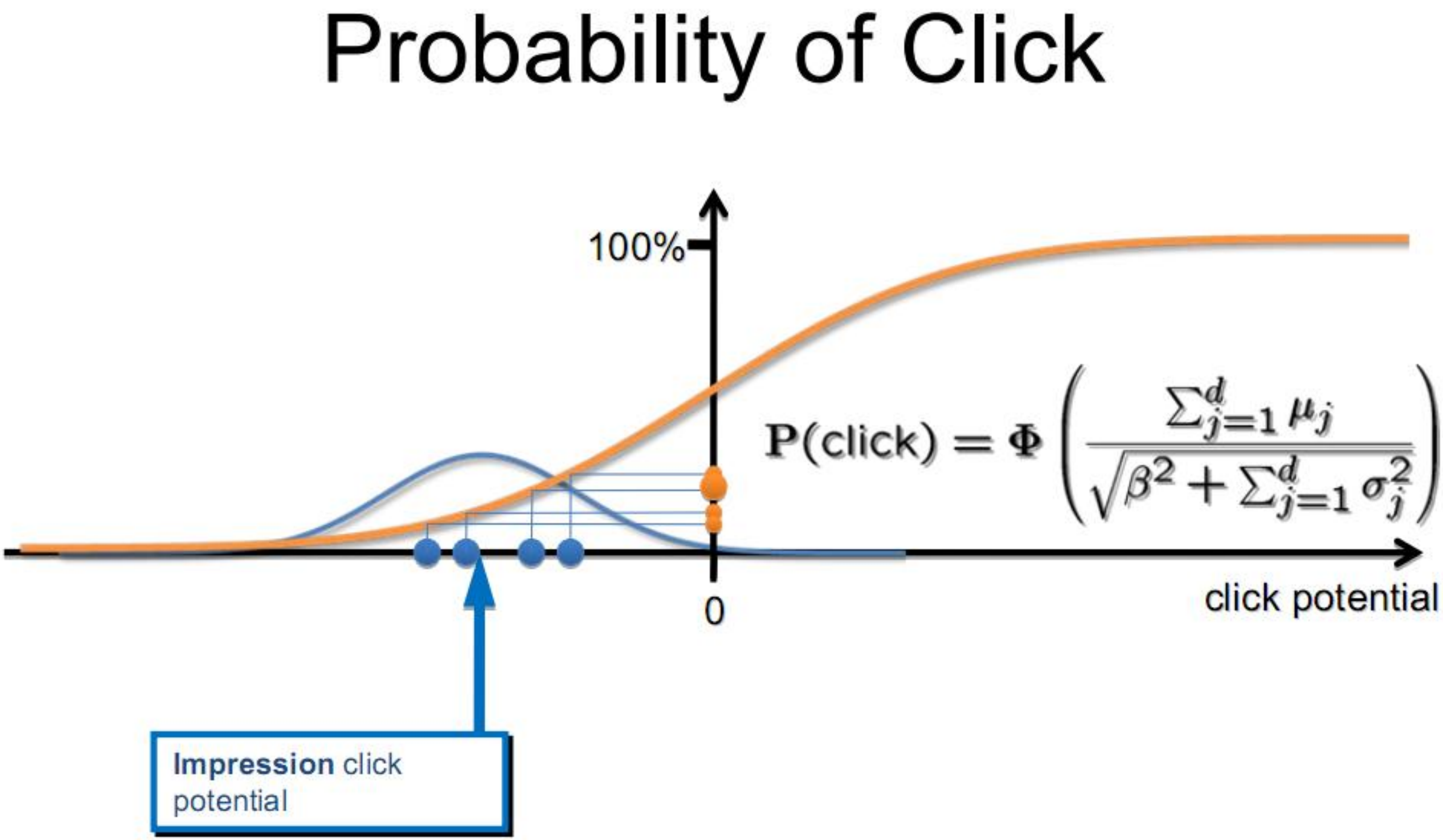
Bayesian Linear Probit Regression

Uncertainty about the Potential



--This page come from Thore Graepel's slides!

Bayesian Linear Probit Regression



--This page come from Thore Graepel's slides!

Bayesian Linear Probit Regression

$$\mathrm{p}(\boldsymbol{w}|\boldsymbol{x}, y) = \frac{p(y|\boldsymbol{x}, \boldsymbol{w}) \cdot p(\boldsymbol{w})}{\int p(y|\boldsymbol{x}, \boldsymbol{w}) \cdot p(\boldsymbol{w}) d\boldsymbol{w}}$$

Update $(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{x}, \mathrm{y}) \rightarrow (\widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{\sigma}}^2)$

$$\widetilde{\mu_{i,j}} = \mu_{i,j} + y x_{i,j} \cdot \upsilon(\frac{y \cdot \boldsymbol{x}^T \boldsymbol{\mu}}{\Sigma}) \qquad \widetilde{\sigma_{i,j}}^2 = \sigma_{i,j}^2 \cdot \left[1 - x_{ij} \cdot \frac{\sigma_{i,j}^2}{\Sigma^2} \cdot \omega\left(\frac{y \cdot \boldsymbol{x}^T \boldsymbol{\mu}}{\Sigma}\right)\right]$$

$$\Sigma^2 = \beta^2 + (\boldsymbol{x}^2)^T \boldsymbol{\sigma}^2 \qquad \upsilon(t) = \frac{\mathcal{N}(t; 0,1)}{\Phi(t; 0,1)}, \omega = \upsilon(t) \cdot [\upsilon(t) + t]$$

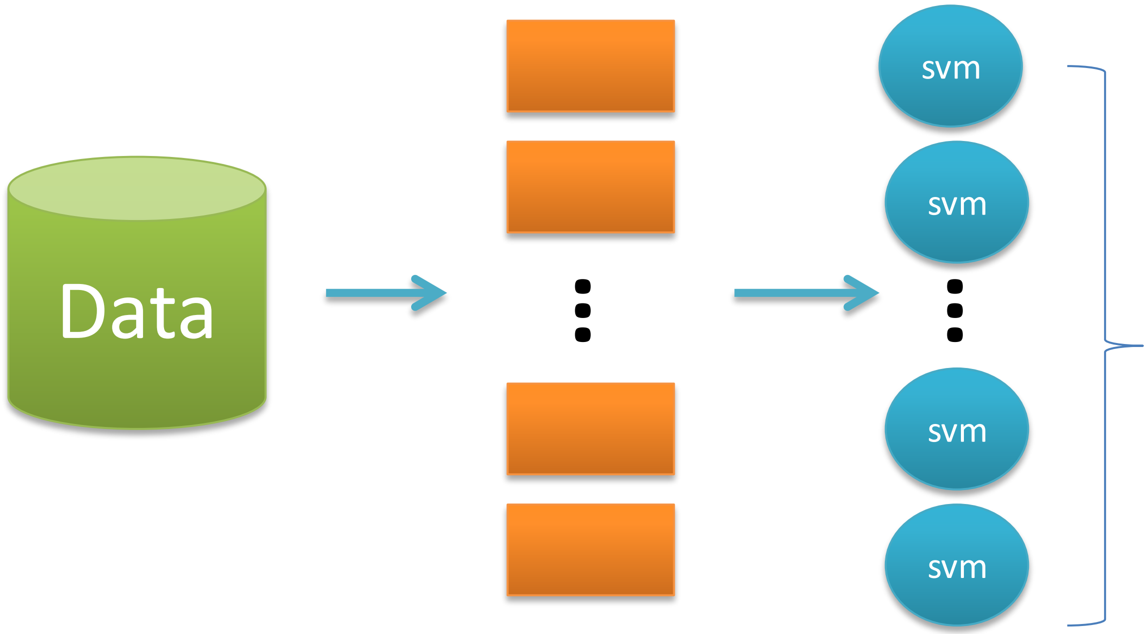
Predict $\mathrm{p}(\mathrm{y}|\boldsymbol{x}) = \Phi(\frac{y \cdot \boldsymbol{x}^T \boldsymbol{\mu}}{\Sigma})$

Support Vector Machine

- SVM^{perf} (http://svmlight.joachims.org/svm_perf.html)
  - Optimize AUC directly
  - Linear kernel

- Feature selection
  - × Features that do not exist in the Test set
  - × Features with low frequency (< 20) in the Training set

Support Vector Machine



$$\overline{score} = \frac{\sum_{i=1}^{10} score_i}{10}$$

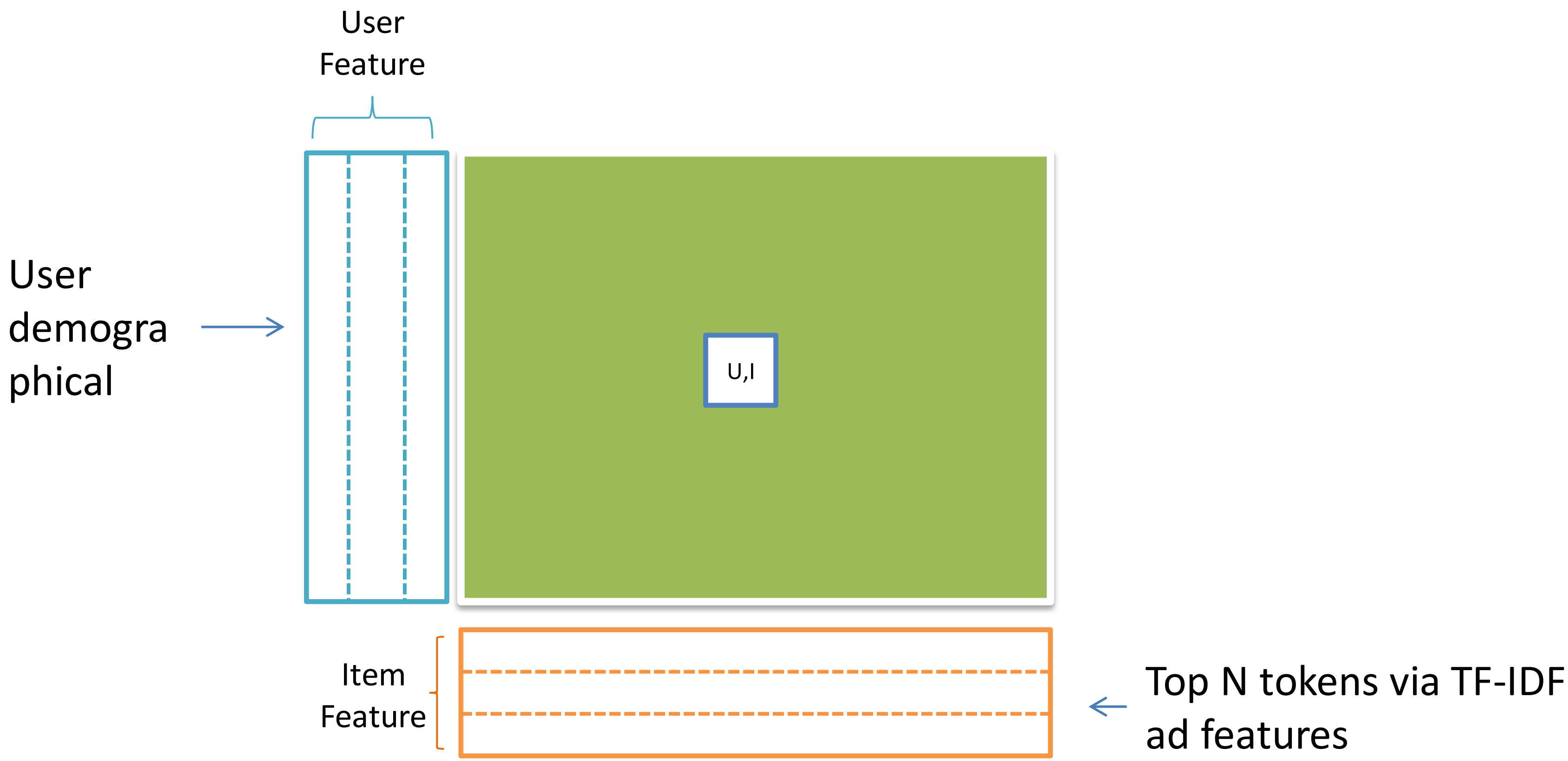
$$score_{svm} = \overline{score} \cdot user\_ctr$$

Randomly separate

10 sub data set

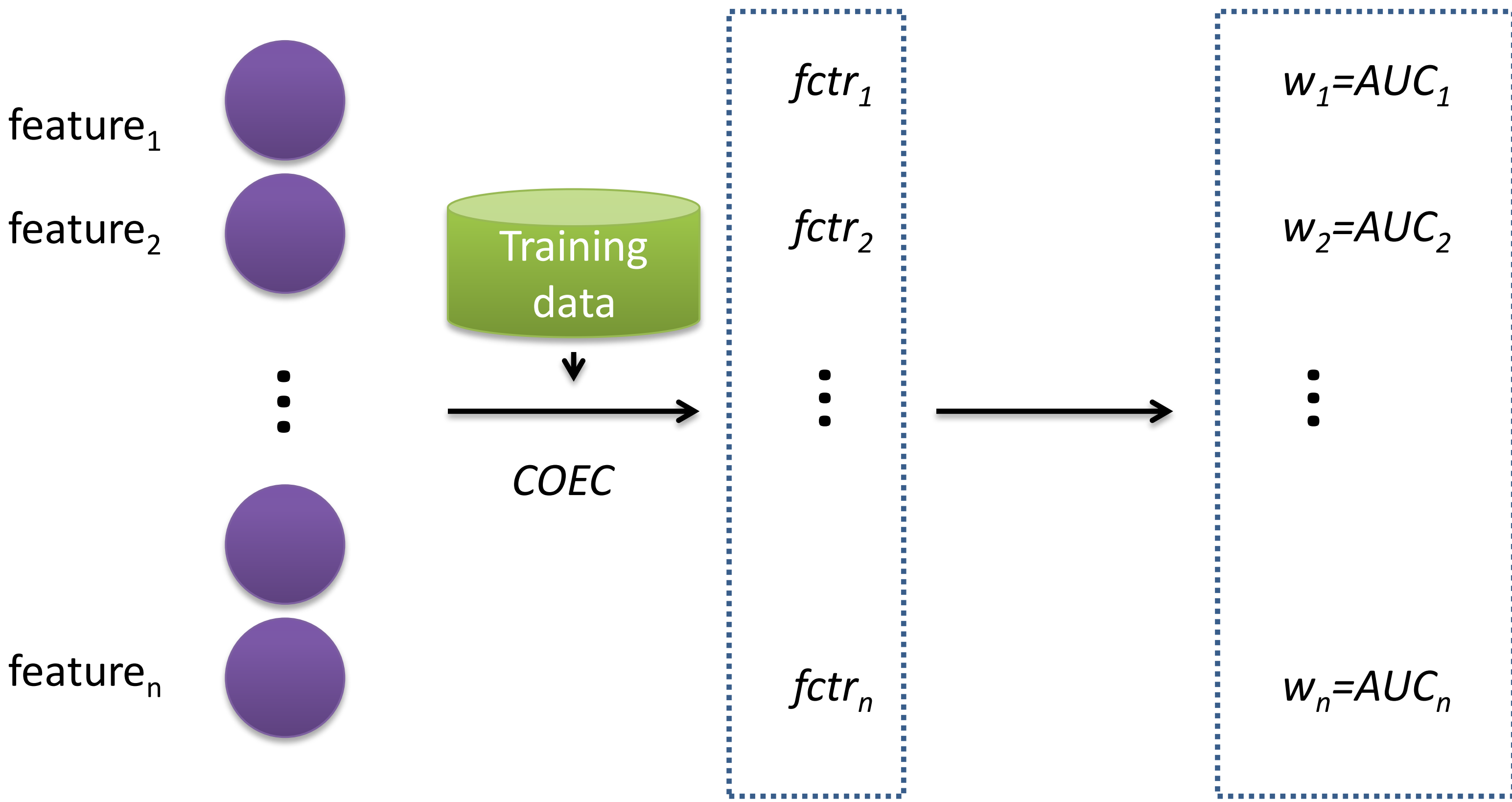
10 models

Feature-based Latent Factor Model



$$\tilde{r_{ui}} = \mu + \sum_{f_i \in F_i} b_{f_i} + \sum_{f_u \in F_u} b_{f_u} + \Big(\sum_{f_i \in F_i} q_{f_i}\Big)^T \Big(\sum_{f_u \in F_u} p_{f_u}\Big)$$

Feature-Based Maximum Likelihood Estimation



$$fctr_{ceoc} = \frac{\sum_{r=1}^{R} C_r}{\sum_{r=1}^{R} i_r \times CTR_r}$$

$$CTR = \sum_{i=1}^{|F|} (w_i \times fctr_i)$$

Feature-Based Maximum Likelihood Estimation

- Penalty Pattern
  - Null User Pattern
    - Average CTR for null user: 0.024926
    - Average CTR for the other: 0.038308

$$CTR' = \frac{CTR}{\log(\#IMP) + 1}$$

Rank-Based Blending

- Two problems:
  - Difference between instances are small
    - So the outlier can affect the final ranking
  - Differences on the pCTR value scale from different models
- Rank-Based Blending

$$\boldsymbol{rank_i = \frac{m}{\sum_{j=1}^{m} \frac{1}{rank_j()}}}$$

Summary

	AUC	Ratio
BPR	0.7902	1.6705%
SVM	0.7865	2.1488%
LFM	0.7710	4.2023%
MLE	0.7924	1.3883%
BLENDING	0.8034	0.0%

Thanks

We are all job seekers!