

# Modeling Impression Discounting In Large-scale Recommender Systems



**Pei Lee, Laks V.S. Lakshmanan**



**Mitul Tiwari, Sam Shah**



# LinkedIn By The Numbers

2



300M members

2 new members/sec



# Large-scale Recommender Systems at LinkedIn

3

- People You May Know
- Skills Endorsements

## PEOPLE YOU MAY KNOW



**Jay Kreps**

Principal Staff Engineer at LinkedIn

[Connect](#)



**Igor Perisic**

VP Engineering at LinkedIn

[Connect](#)



**Sam Shah**

Principal Engineer at LinkedIn

[Connect](#)



[See more »](#)

## What skills or expertise do your other connections have?



Does **DJ Patil** know about **Data Mining**?

[Endorse](#)



Does **Hirji Deihonte** know about **Entrepreneurship**?

[Endorse](#)



Does **Deep Nishar** know about **Business Strategy**?

[Endorse](#)



Does **Jay Kreps** know about **Distributed Systems**?

[Endorse](#)

[Endorse all 4](#)

[Close](#)

[See more](#)

# Recommender Systems at LinkedIn

4

## Connections

PEOPLE YOU MAY KNOW

-  **Matt Koenig**, product specialist @ linkedin [Connect](#)
-  **Raymond Ng**, Technical Lead & Engineering Manager at LinkedIn [Connect](#)
-  **Diego Buthay**, Senior Software Engineer at LinkedIn [Connect](#)

[See more >](#)

## News

LinkedIn Today: See all Top Headlines for You

-  **Slime Machine: SOPs Edition**
-  **HBR**
-  **Yahoo Co-Founder Jerry Yang**

## Skill Endorsements

What skills or expertise do your other connections have?

-  Does **Du Pall** know about **Data Mining**?
-  Does **Hirj Dethame** know about **Entrepreneurship**?
-  Does **Deep Nahar** know about **Business Strategy**?
-  Does **Jay Krups** know about **Distributed Systems**?

## Similar Profiles

Similar Profiles: Adrian Silvescu (50)

Similar profiles are based on similar experience, job function, and other criteria.

-  **Jaysimha Reddy Katukuri**, Senior Research Scientist at eBay

## Jobs You May Be Interested In



Jobs You May Be Interested In

## Talent Match



Talent Match

## CAP



CAP

## Companies

Recommendations, similar companies search, peer companies, and company browse maps, company products and services browse maps



Companies

## Related search



Related search

## Profile browse maps



Profile browse maps

## Behind the Scenes



Behind the Scenes

## Jobs browse maps



Jobs browse maps

## Ad matching engine

$pCTR = f(\text{member, creative, advertiser, context, inventory, DCTR})$

## Groups

Recommendations, similar groups search



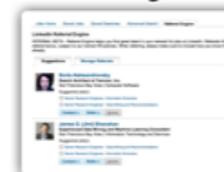
Groups

## Similar jobs



Similar jobs

## Referral Engine



Referral Engine



# What is Impression Discounting?

5

- Examples
  - ▣ People You May Know
  - ▣ Skills Endorsement
- Problem Definition

# Example: People You May Know

## People You May Know



**Deepak Agarwal**  
Director of Engineering at LinkedIn  
[+ Connect](#)



**Grace McGill**, Transforming myself, LinkedIn, and the world  
[+ Connect](#)



**Liz Li**, Senior Product Manager at LinkedIn  
[+ Connect](#)

[See more](#)

## People You May Know



**Reed Johnston**, HRIS Analyst, Workday at LinkedIn  
[+ Connect](#)



**Deepak Agarwal**  
Director of Engineering at LinkedIn  
[+ Connect](#)



**Garrett Vangelisti**, User Experience Design Recruiting...  
[+ Connect](#)

[See more](#)

## People You May Know



**Fan Chen**, Software Engineer at LinkedIn  
[+ Connect](#)



**Kalpana M.**, Senior software QA engineer  
[+ Connect](#)



**Deepak Agarwal**  
Director of Engineering at LinkedIn  
[+ Connect](#)

[See more](#)

**5 days ago:**  
Recommended, but not invited

**1 day ago:**  
Recommended again, but not invited

**Today:**  
If we recommend again, would you invite or not?

If you are not likely to invite, we better discount this recommendation.

# Example: Skills Endorsement

Does Darshan have these skills or expertise?

Social Media x Telecommunications x VoIP x

Type another area of expertise...

Endorse Skip

Recommended

Does Darshan have these skills or expertise?

Social Media x Telecommunications x VoIP x

Project Management x Wireless x Type another area of exper

Endorse Skip

Recommended Again

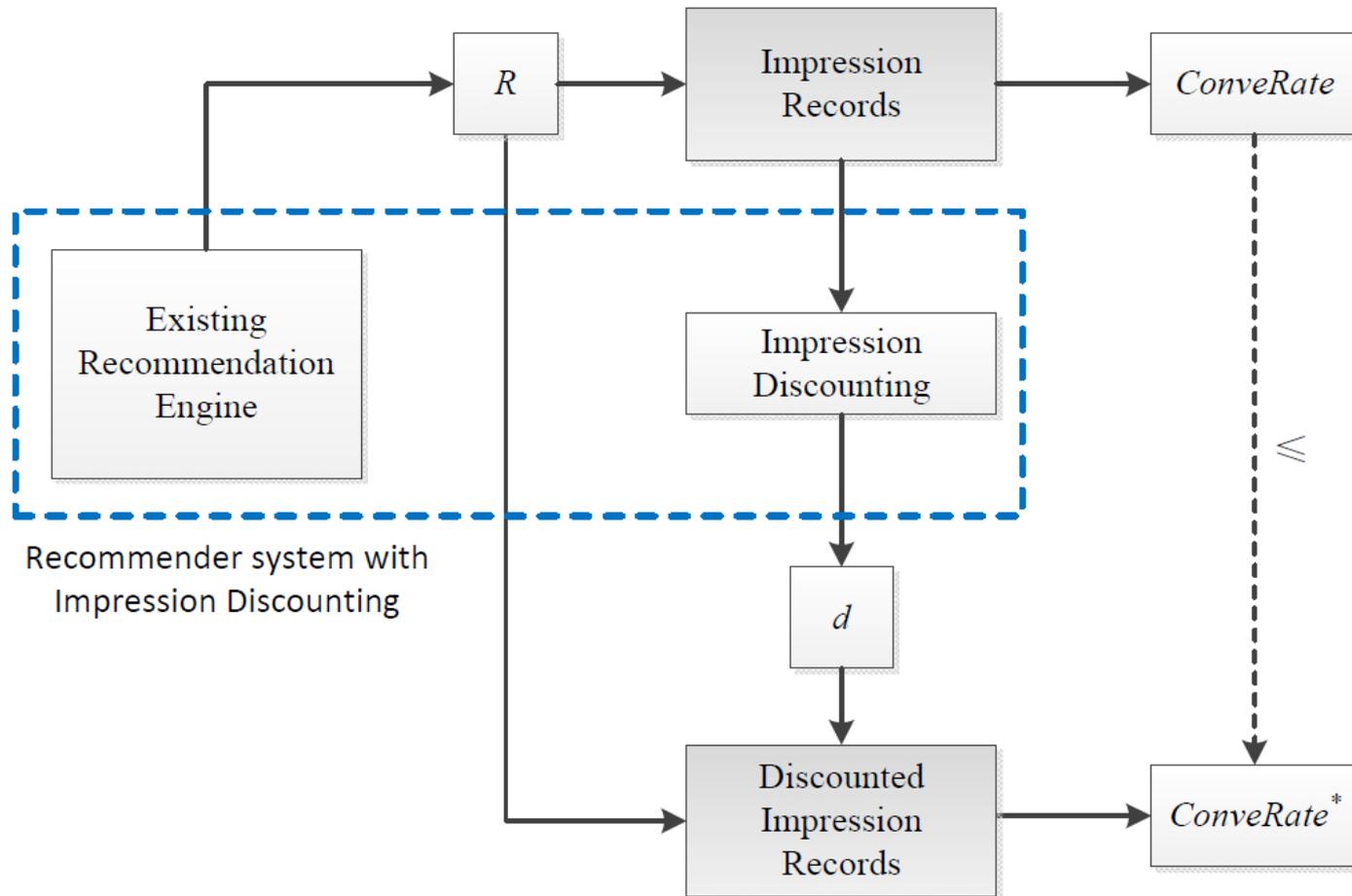
- If you are likely to endorse, we will show a skills suggestion again;
- Otherwise, we should leave this space for other skills suggestions

# Problem Definition

- Impressions: recommendations shown to user
- Conversion: positive action - invite, endorse, etc.
- In natural language:
  - ▣ We already impressed an item several times with no conversion. How much should we discount that item?
- More formally:
  - ▣ For an user  $u$  and item  $i$ , given an impression history  $(T_1, T_2, \dots, T_n)$  between  $u$  and  $i$ , can we predict the conversion rate for  $u$  on  $i$ ?

# Impression Discounting Framework

9



# Outline

10

- Define an impression
  - ▣ User, Item, Conversion Rate
  - ▣ Features
    - How many times the user saw this item?
    - When is the last time the user saw this item? ...
- Data analysis: impression features and conversion relationship
- Impression discounting model fitting
- Experimental evaluation

# Define an Impression

11

$$\mathcal{T} = (user, item, conversion, [behavior1, behavior2, \dots], t, R)$$

where:

- *user* is a user ID;
- *item* is a recommendation item ID;
- *conversion* is a boolean type to describe whether or not *user* takes an action on *item* in this impression;
- *behavior* is an observed feature of interaction;
- *t* is the time stamp of impression;
- *R* is the recommendation score of *item* to *user*, which is provided by the recommendation engine.

# Features on Impressed Items

12

- LastSeen (recency):
  - ▣ the day difference between the last impression and the current impression, associated with the same (user, item)
- ImpCount (frequency):
  - ▣ the number of historical impressions before the current impression, associated with the same (user, item)
- Position:
  - ▣ the offset of item in the recommendation list of user
- UserFreq:
  - ▣ the interaction frequency of user in a recommender system

# PYMK Dataset

- **Data:** 1.09 B impression tuples
- **Training dataset:** 80% of data
  - ▣ 0.55 billion unique impressions
  - ▣ 0.87 billion total impressions
  - ▣ 20 millions invitations
- **Testing dataset:** 20% of data
  - ▣ 0.14 billion unique impressions (3.7%)
  - ▣ 0.22 billion total impressions (2.4%)
  - ▣ 5.2 millions invitations

# Skills Endorsement Dataset

14

- Total dataset size: 190 million impression tuples
- **Training dataset:** 80% of data
- **Testing dataset:** 20% of data

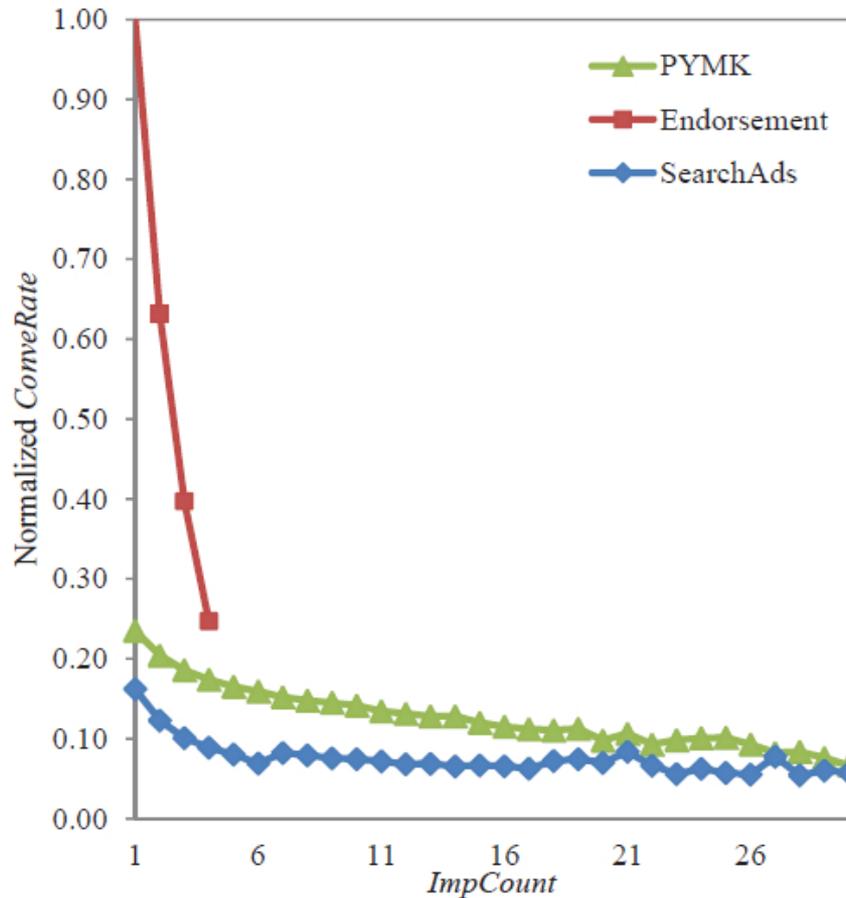
# Tencent SearchAds Dataset

15

- Publicly available for KDD Cup 2012 by the Tencent search engine
- Total Size: 150 million impression sequences
- CTR of the Ad at the 1st, 2nd and 3rd position: 4.8%, 2.7%, and 1.4%

# Conversion Rate Changes with Impression Count

16



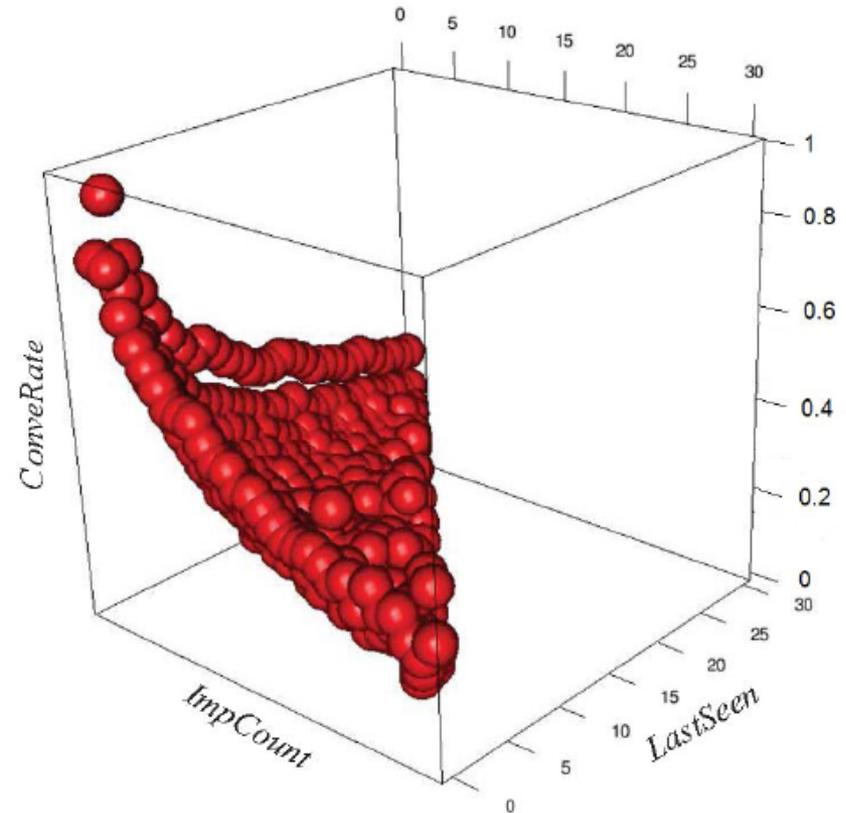
If we show a item to a user repeatedly, the conversion rate decreases

**Figure 3: The change of Normalized *ConveRate* with increasing *ImpCount* on three real-world impression data sets.**

# PYMK Invitation Rate Changes with Impression Count and Last Seen

17

- The conversion rate decreases with both ImpCount and LastSeen

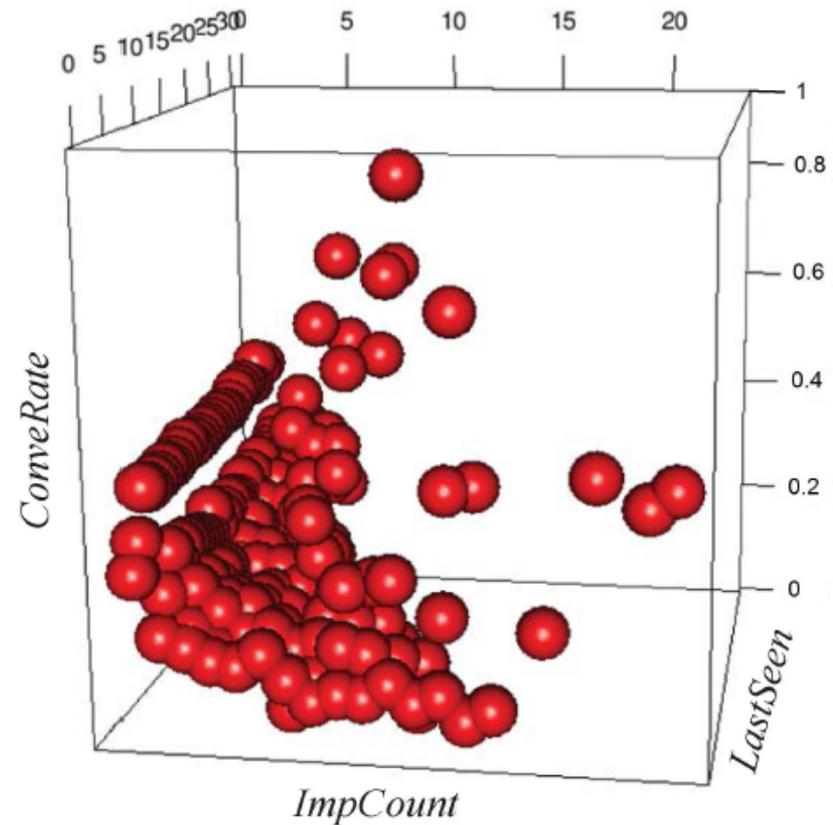


(a) PYMK 3-D

# Endorsement Rate Changes with Impression Count and Last Seen

18

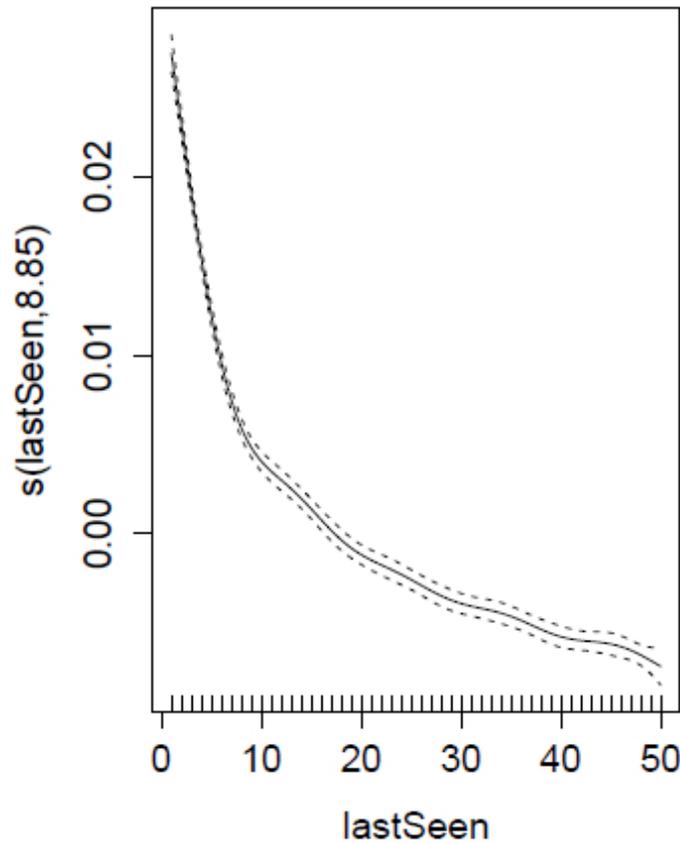
- Similar observation: The conversion rate also decreases with both ImpCount and LastSeen



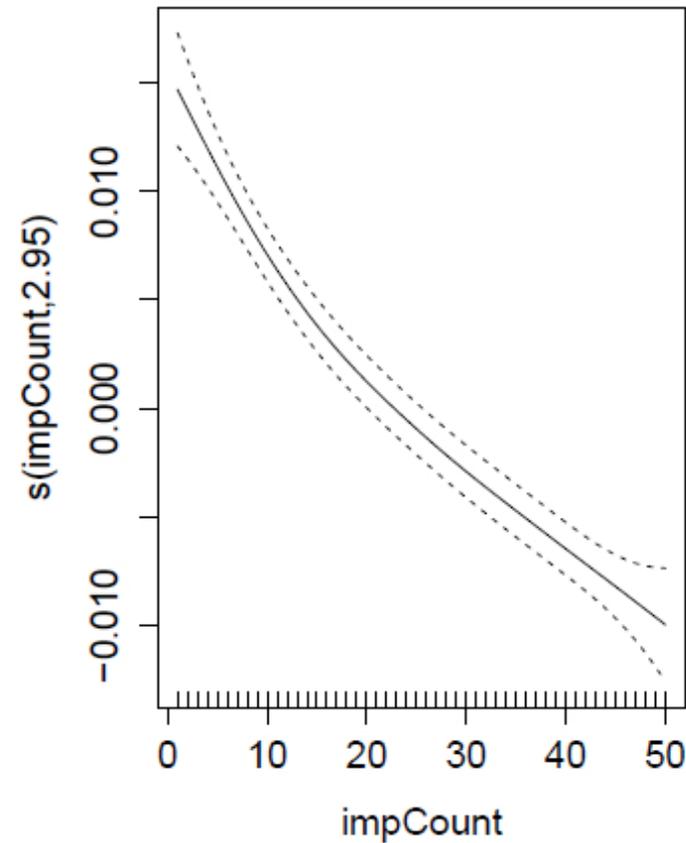
(b) Endorsement 3-D

# Correlation Confidence Analysis

19



(a) *LastSeen* Correlation



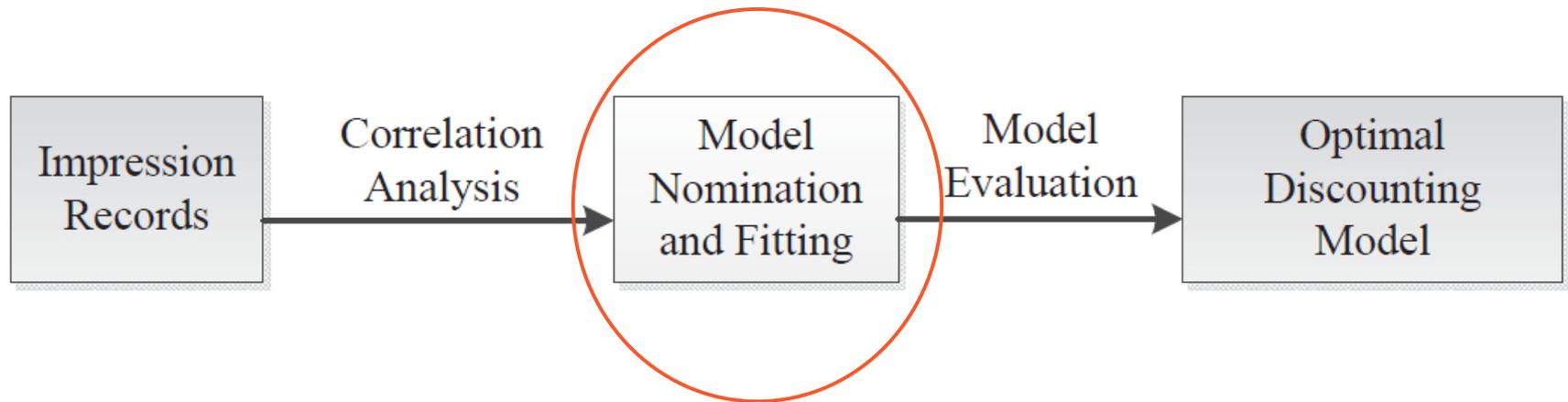
(b) *ImpCount* Correlation

Conversion rate has a strong, negative correlation with LastSeen and ImpCount

# Impression Discounting Model Fitting

20

## □ Workflow:



## □ Model fitting process:

- ▣ Fundamental discounting functions
- ▣ Aggregation Model
- ▣ Offline and online

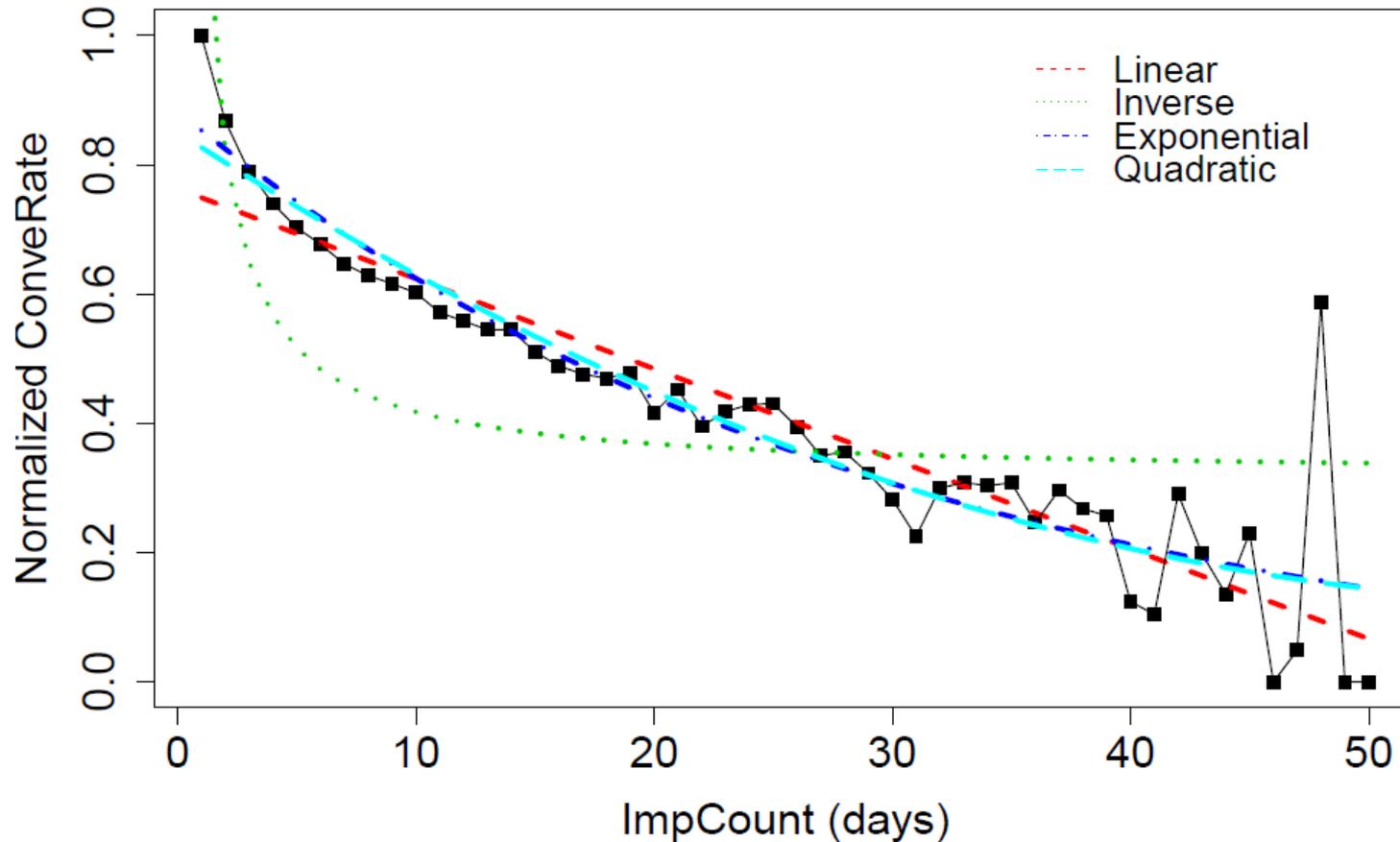
# Discounting Functions

21

- Linear Discounting:  $f_L(x) = \alpha_1 \cdot x + \alpha_2$ ;
- Inverse Discounting:  $f_I(x) = \frac{\alpha_1}{x} + \alpha_2$ ;
- Exponential Discounting:  $f_E(x) = e^{\alpha_1 \cdot x + \alpha_2}$ ;
- Quadratic Discounting:  $f_Q(x) = \alpha_1(x - \alpha_2)^2 + \alpha_3$ .

# Conversion Rate with ImpCount

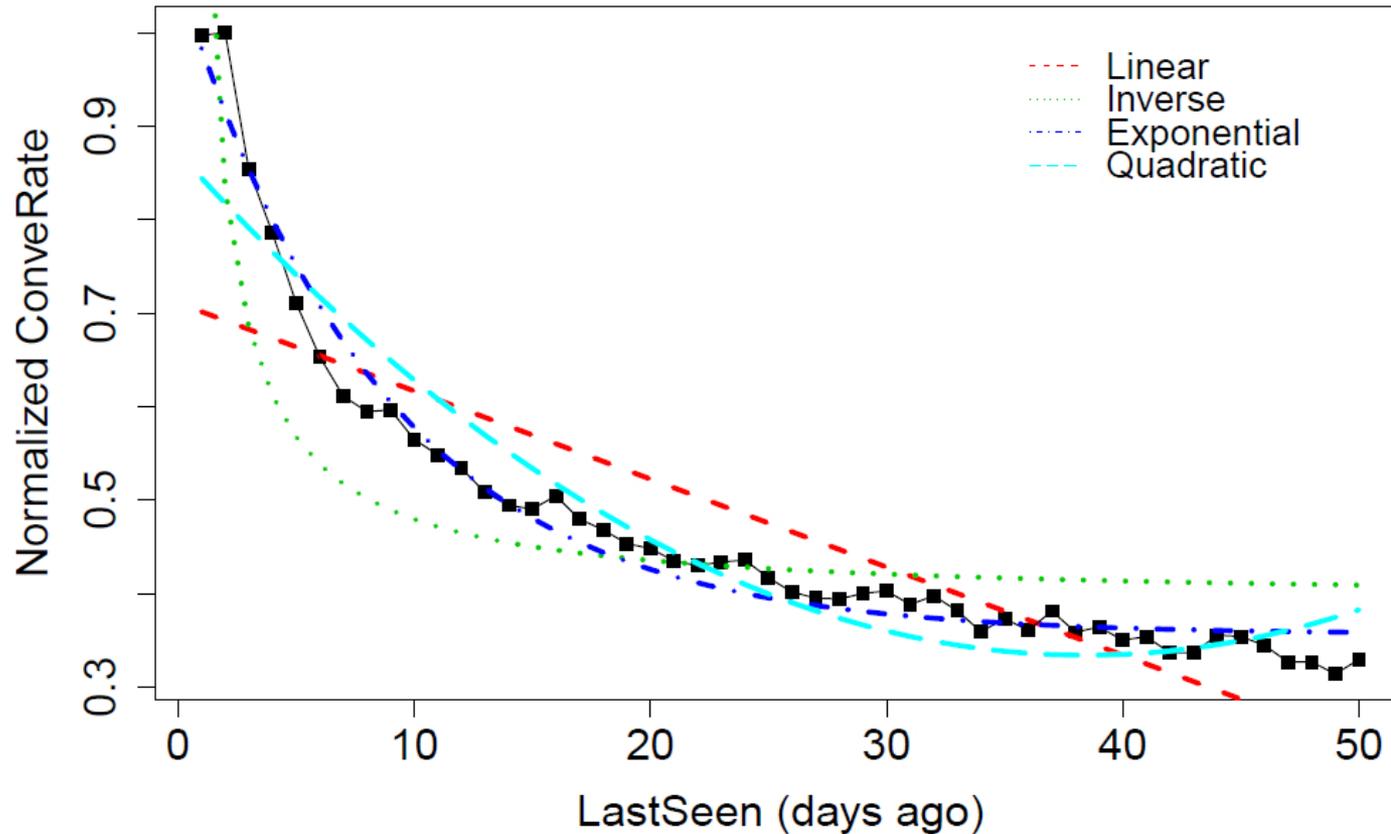
22



(b) Conversion rate vs. *ImpCount*

# Conversion Rate with LastSeen

23



(a) Conversion rate vs. *LastSeen*

# Aggregation Models

24

## □ Linear Aggregation

$$\hat{y} = \sum_{i=1}^m w_i f(X_i)$$

## □ Multiplicative Aggregation

$$\hat{y} = w \prod_{i=1}^m f(X_i)$$

$f(X_i)$  is one of  
discounting function  
given earlier

# Aggregation Models

25

## □ Linear Aggregation

$$\hat{y} = \sum_{i=1}^m w_i f(X_i)$$

## □ Multiplicative Aggregation

$$\hat{y} = w \prod_{i=1}^m f(X_i)$$

## □ Impression Discounting Factor

$$d = \frac{\tilde{y}}{\max \tilde{y}}$$

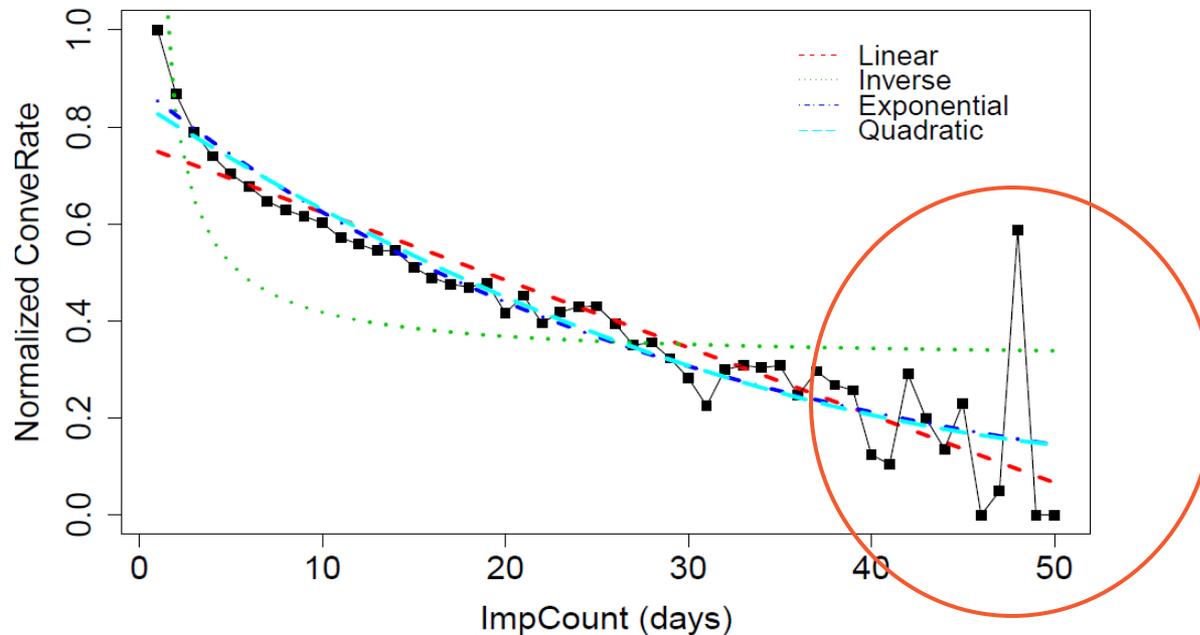
$f(X_i)$  is one of  
discounting function  
given earlier

$0 < d \leq 1$ ; will be  
applied to discount  
the recommendation

# Anti-Noise Regression Model

26

- The sparsity of observations increases variance in the conversion rate.
- Typically happens when the number of observations are relatively low.



(b) Conversion rate vs. *ImpCount*

# Density Weighted Regression

27

- Add a weight with square error
- Original RMSE:

$$RMSE^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y})^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{X}_i^T u)^2$$

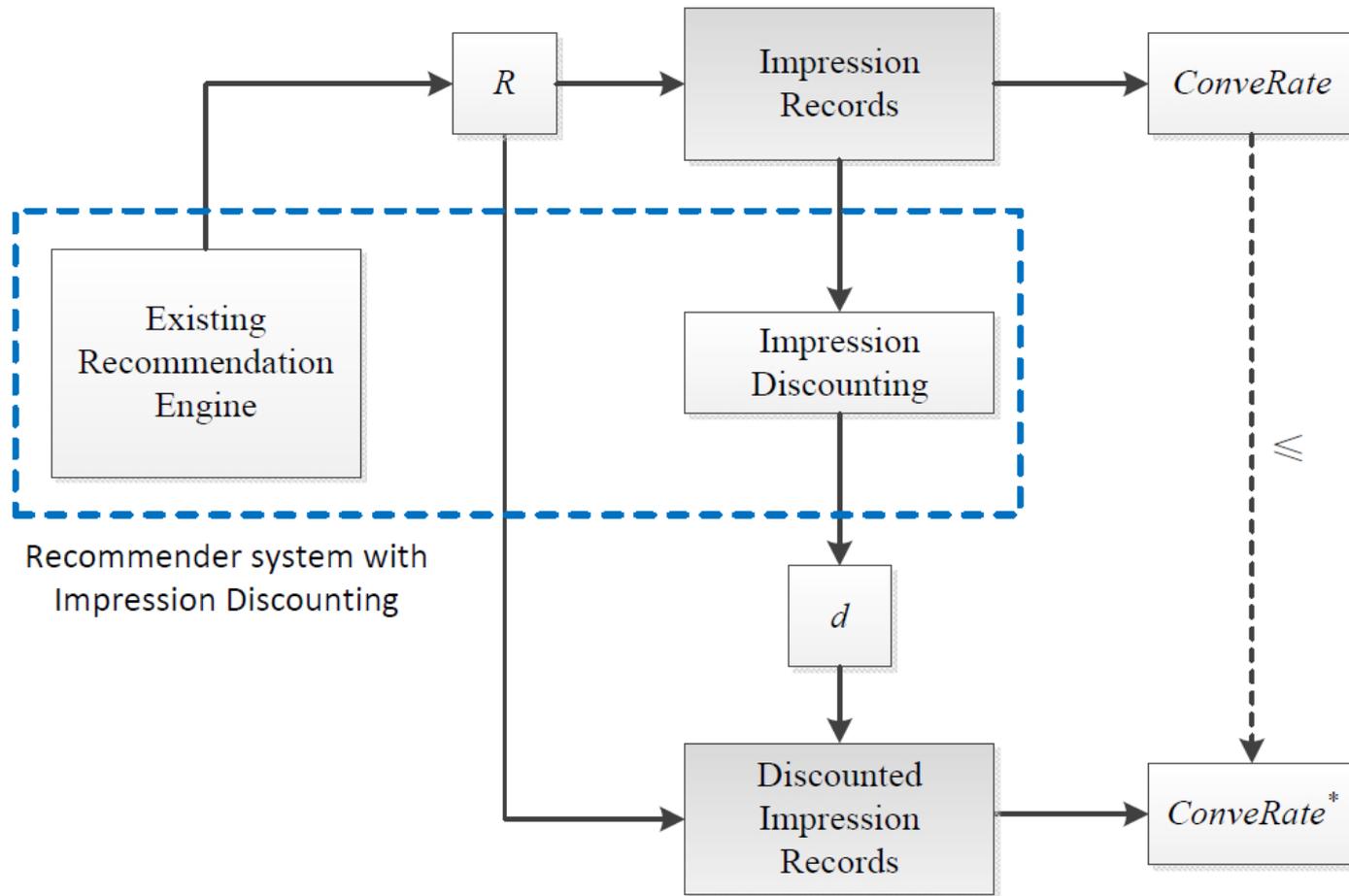
- Density weighted RMSE:

$$RMSE_v^2 = \frac{1}{n} \sum_{i=1}^n (v_i (y_i - \tilde{X}_i^T u))^2$$

$v_i$  is assessed by the number of observations in a small neighborhood of  $(X_i, y_i)$

# Impression Discounting Framework

28



# Quality Measure: Precision @k

29

- Precision at top k

$$P@k = \frac{\text{seqSize}(\text{conversion} = \text{true})}{k}$$

- Precision improvement for different behavior sets:

Behavior Set	Precision Improvement
<b>PYMK (<math>P@10</math>)</b>	
<i>LastSeen, ImpCount</i>	13.7%
<i>LastSeen, ImpCount Position, UserFreq</i>	31.3%
<b>Endorsement (<math>P@10</math>)</b>	
<i>LastSeen, ImpCount</i>	1.3%
<i>LastSeen, ImpCount Position, UserFreq</i>	3.4%
<b>SearchAds (<math>P@5, 10</math>)</b>	
<i>ImpCount</i>	0.53% ( $P@10$ )
<i>ImpCount, Position</i>	3.2% ( $P@10$ )
<i>ImpCount, Position</i>	6.87% ( $P@5$ )

More features yield  
higher precision

# Bucket Test on People You May Know

30

- Use behavior set (LastSeen, ImpCount)
- On LinkedIn PYMK online system

$X_1: LastSeen; X_2: ImpCount$	
Group A: Without Impression Discounting	
Group B:	Improvement
$\alpha_1 \cdot \alpha_2^{X_1} \cdot \alpha_3^{X_2}$	11.97% $\pm$ 0.2%
$(\frac{\alpha_1}{X_1} + \alpha_2) \cdot \alpha_3^{X_2}$	13.26% $\pm$ 0.2%
$(\alpha_1 \cdot X_1 + \alpha_2) \cdot \alpha_3^{X_2}$	12.18% $\pm$ 0.2%

**Table 3: The A/B test results of precision at top 10 of different impression discounting models on PYMK data set, with  $X = (LastSeen, ImpCount)$ .**

# Conclusion

31

- ❑ Impression Discounting Model
  - ❑ Learnt a discounting function to capture ignored impressions
  - ❑ Linear or multiplicative aggregation model
  - ❑ Anti-noise regression model
- ❑ Offline and Online evaluation (Bucket testing)

## PEOPLE YOU MAY KNOW



**Jay Kreps**

Principal Staff Engineer at LinkedIn

[Connect](#)



**Igor Perisic**

VP Engineering at LinkedIn

[Connect](#)



**Sam Shah**

Principal Engineer at LinkedIn

[Connect](#)



[See more »](#)

## What skills or expertise do your other connections have?



Does **DJ Patil** know about **Data Mining**?

[Endorse](#)



Does **Hirji Deihonte** know about **Entrepreneurship**?

[Endorse](#)



Does **Deep Nishar** know about **Business Strategy**?

[Endorse](#)



Does **Jay Kreps** know about **Distributed Systems**?

[Endorse](#)

[Endorse all 4](#)

[Close](#)

[See more](#)

# Questions?

32

- Related work in the paper
- Want to know more contact:  
[mtiwari@linkedin.com](mailto:mtiwari@linkedin.com)

