prod2vec

KDD2015

E-commerce in Your Inbox: Product Recommendations at Scale

Mihajlo Grbovic, Vladan Radosavljevic, Nemanja Djuric, Narayan Bhamidipati
Yahoo Labs
701 First Avenue, Sunnyvale, USA
{mihajlo, vladan, nemanja, narayanb}@yahoo-inc.com

Jaikit Savla, Varun Bhagwan, Doug Sharp
Yahoo, Inc.
701 First Avenue, Sunnyvale, USA
{jaikit, vbhagwan, dsharp}@yahoo-inc.com

RecSys '16 Proceedings of the 10th ACM Conference on Recommender Systems

Meta-Prod2Vec - Product Embeddings Using Side-Information for Recommendation

Flavian Vasile
Criteo
Paris
f.vasile@criteo.com

Elena Smirnova
Criteo
Paris
e.smirnova@criteo.com

Alexis Conneau *
Facebook AI Research
Paris
aconneau@fb.com

ECML-PKDD 2017

MRNet-Product2Vec: A Multi-task Recurrent Neural Network for Product Embeddings

Arijit Biswas, Mukul Bhutani and Subhajit Sanyal

Core Machine Learning, Amazon, Bangalore, India
barijit,mbhutani,subhajs@amazon.com

PAN XAING
2018.10.9

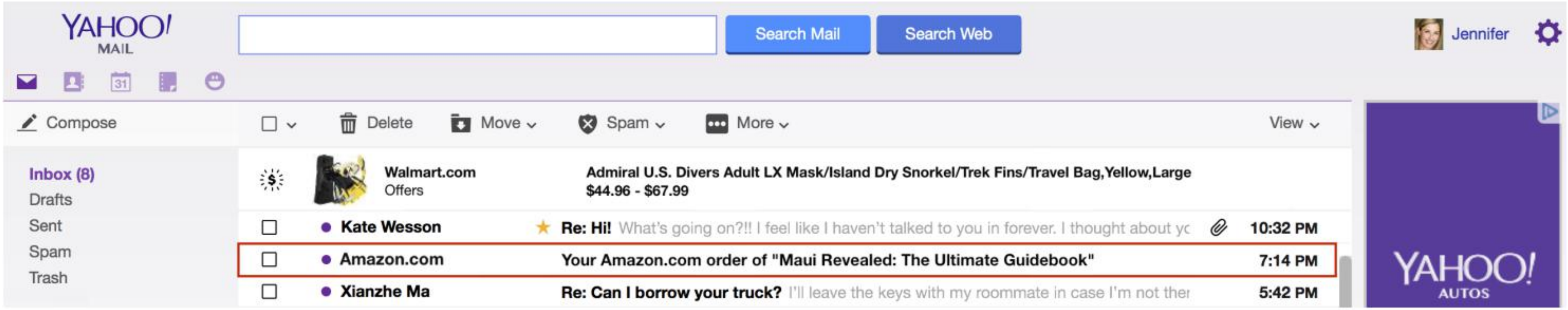
E-commerce in Your Inbox:
Product Recommendations at Scale

Mihajlo Grbovic, Vladan Radosavljevic,
Nemanja Djuric, Narayan Bhamidipati
Yahoo Labs
701 First Avenue, Sunnyvale, USA
{mihajlo, vladan, nemanja,
narayanb}@yahoo-inc.com

Jaikit Savla, Varun Bhagwan,
Doug Sharp
Yahoo, Inc.
701 First Avenue, Sunnyvale, USA
{jaikit, vbhagwan,
dsharp}@yahoo-inc.com

INTRODUCTION

- 10% : human-generated e-mails.
  22% : online shopping
- mail model
  - the info in the mail is formated
- mutiple source



RELATED WORK

- Mail Data
    - click or purchase
    - commercial domains
        - the domain inply the natural Community
    - CF(collaborative filtering)
        - context-based
        - user-based
- NLP
    - word2vec

Low-dimensional product embeddings

prod2vec

sentence: purchase sequence
words: products

maxmize(skip-gram):
get future content

$$\mathcal{L} = \sum_{s \in \mathcal{S}} \sum_{p_i \in s} \sum_{-c \leq j \leq c, j \neq 0} \log \mathbb{P}(p_{i+j} | p_i)$$

$$\mathbb{P}(p_{i+j} | p_i) = \frac{\exp(\mathbf{v}_{p_i}^\top \mathbf{v}'_{p_{i+j}})}{\sum_{p=1}^{P} \exp(\mathbf{v}_{p_i}^\top \mathbf{v}'_p)}$$

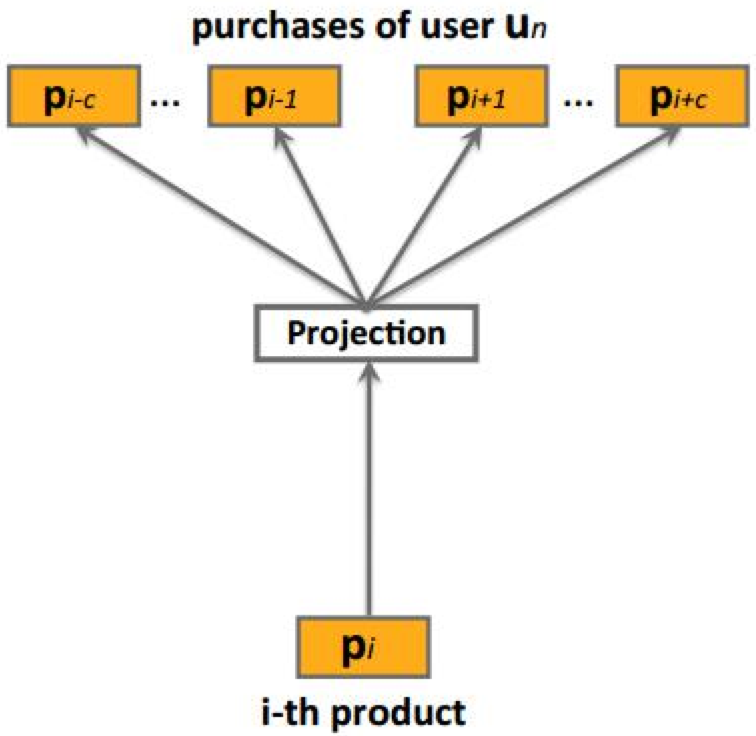
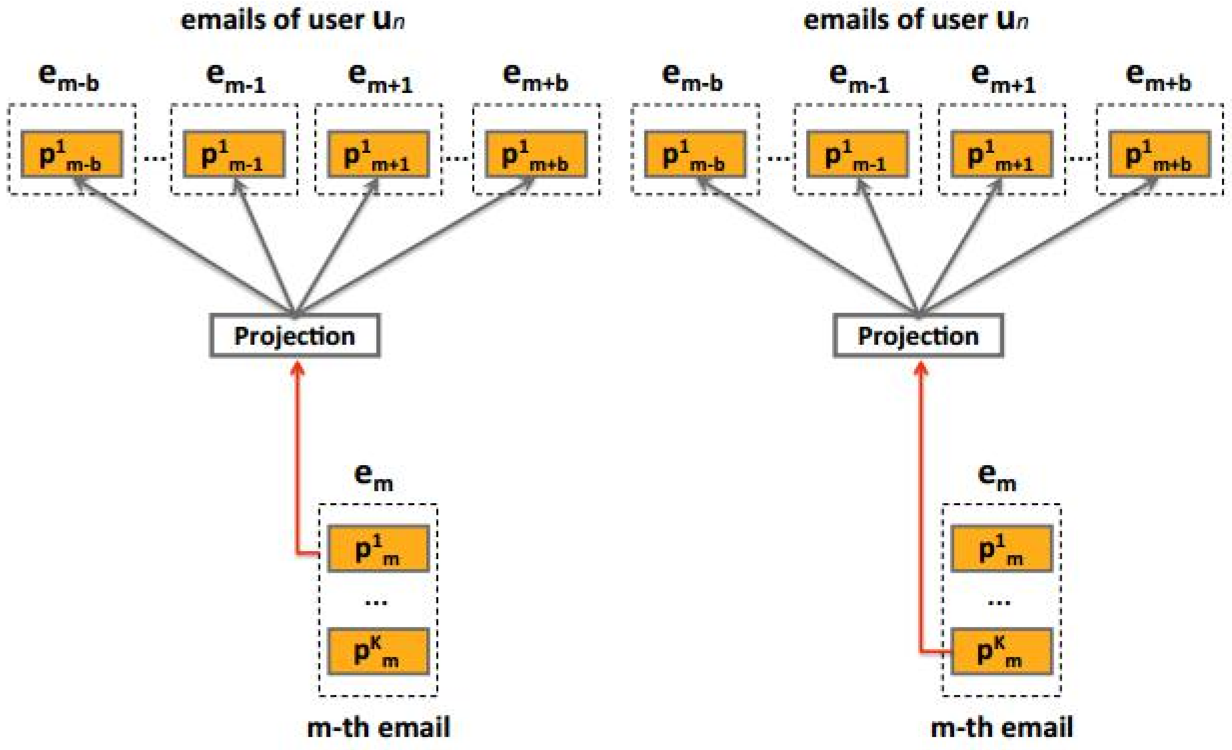


Figure 2: prod2vec skip-gram model

Low-dimensional product embeddings

bagged-prod2vec：

multiple products may be purchased at the same time



Probability $\mathbb{P}(e_{m+j}|p_{mk})$ of observing products from neighboring e-mail receipt e_{m+j}, $e_{m+j} = (p_{m+j,1} \dots p_{m+j,T_m})$,

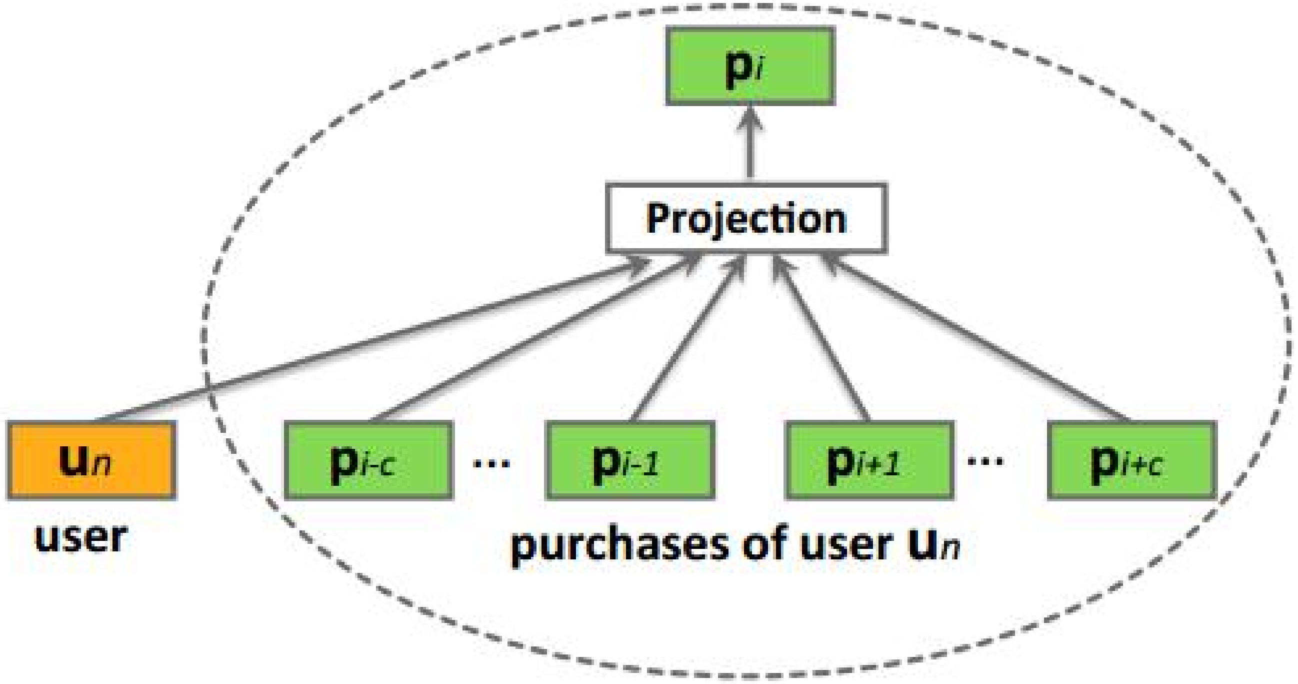
Product-to-product predictive models

prod2vec-topK：

K most similar products

prod2vec-cluster：

from cluster



User-to-product predictive models

global context: user

paragraph2vec

purchase sequences $\mathcal{S}$

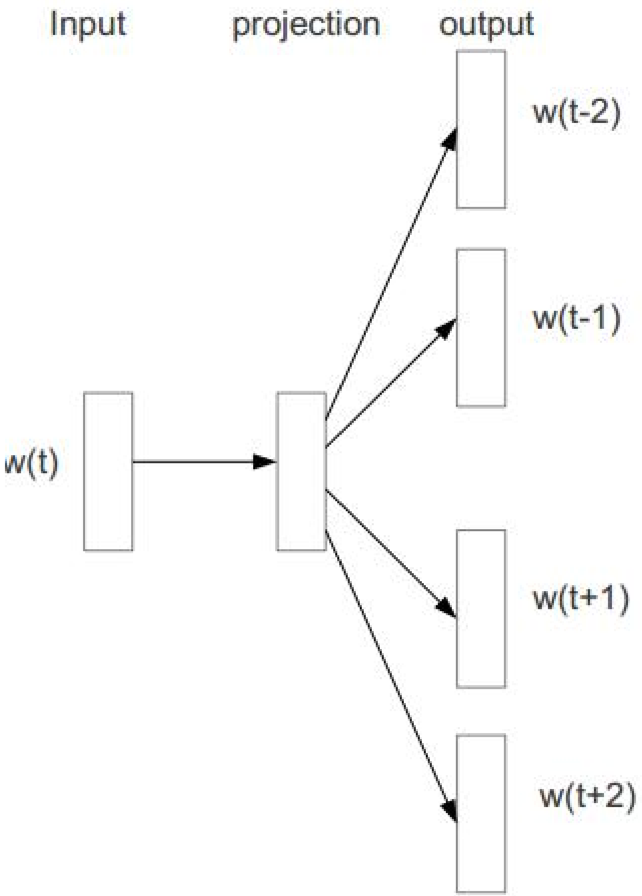
$u_n = (p_{n1}, p_{n2}, \ldots p_{nU_n})$

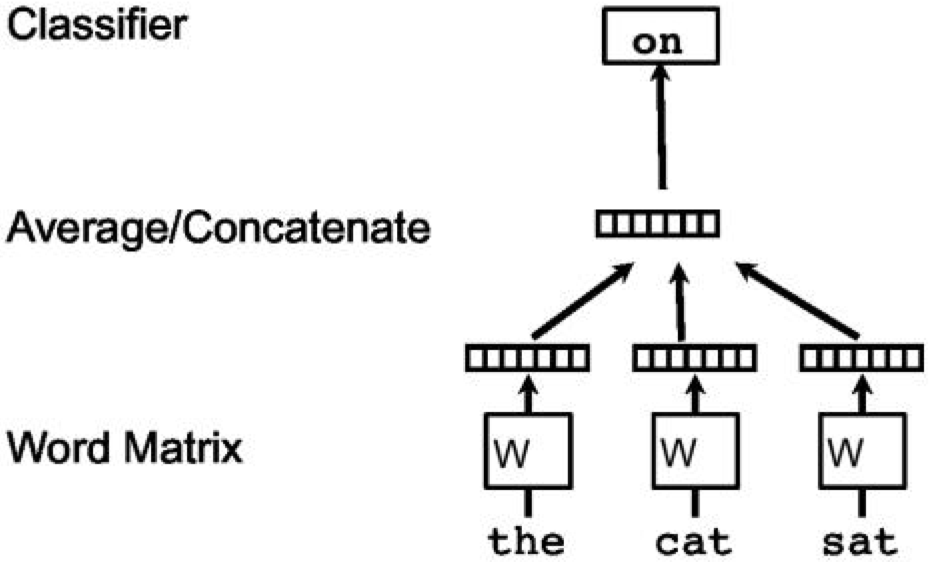
c is the length of the context for products

More specifically, objective of user2vec is to maximize the log-likelihood over the set $\mathcal{S}$ of all purchase sequences,

$$\mathcal{L} = \sum_{s \in \mathcal{S}} \Big(\sum_{u_n \in s} \log \mathbb{P}(u_n | p_{n1} : p_{nU_n}) + \sum_{p_{ni} \in u_n} \log \mathbb{P}(p_{ni} | p_{n,i-c} : p_{n,i+c}, u_n) \Big) \tag{3.5}$$

paragraph2vec





paragraph2vec

$$\frac{1}{T}\sum_{t=k}^{T-k}\log p(w_t|w_{t-k},...,w_{t+k})$$

The prediction task is typically done via a multiclass classifier, such as softmax. There, we have

$$p(w_t|w_{t-k},...,w_{t+k}) = \frac{e^{y_{w_t}}}{\sum_i e^{y_i}}$$

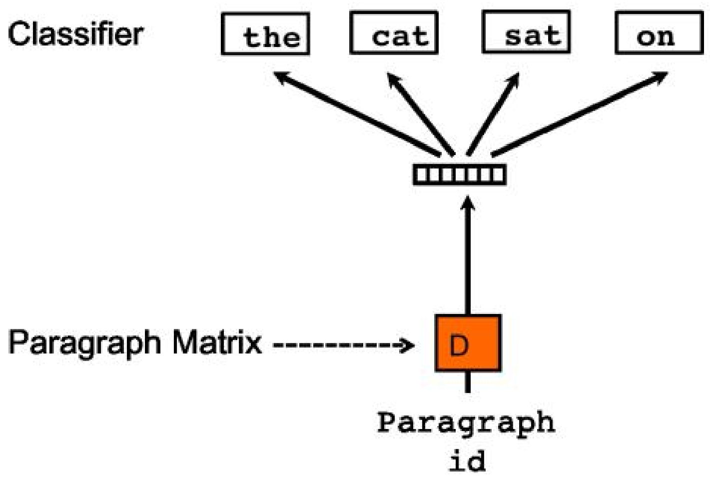
Each of y_i is un-normalized log-probability for each output word i, computed as

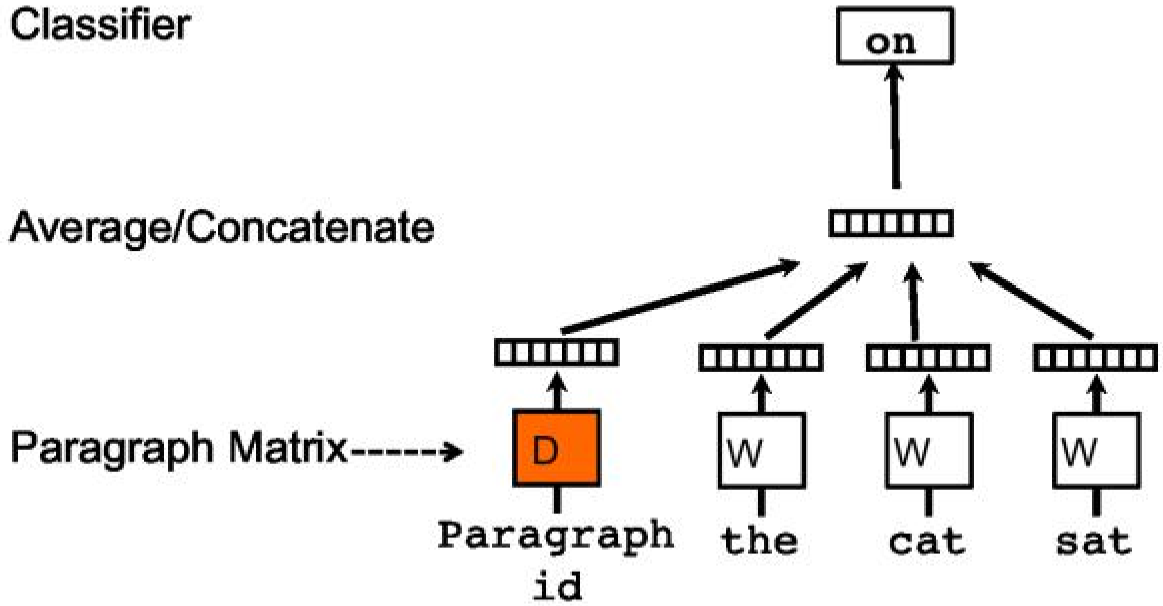
$$y = b + Uh(w_{t-k},...,w_{t+k};W) \tag{1}$$

DocVec:h is constructed from W and D specialization

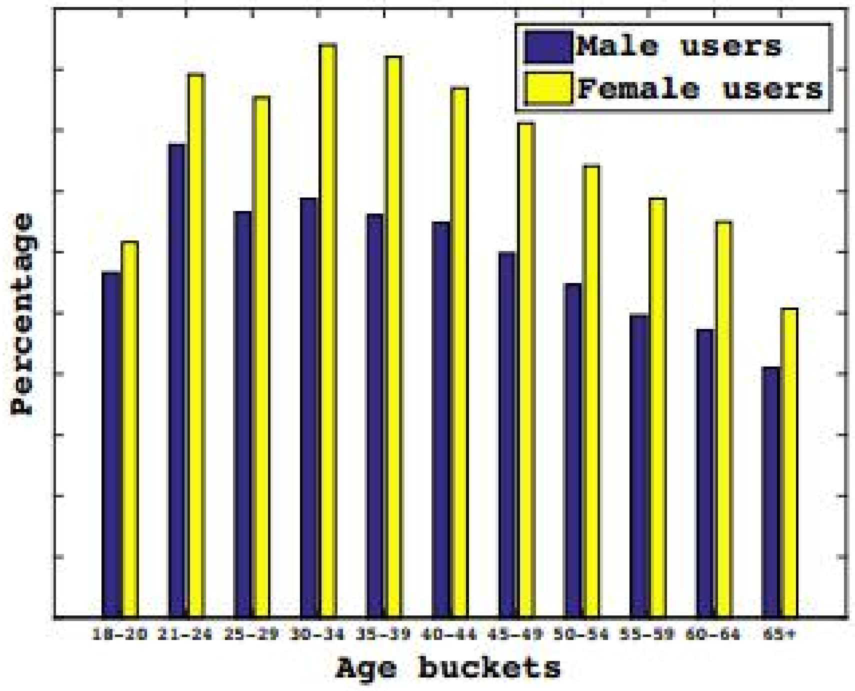
paragraph2vec

every paragraph is mapped to a unique vector, represented by a column in **matrix D**
every word is also mapped to aunique vector, represented by a column in **matrix W**

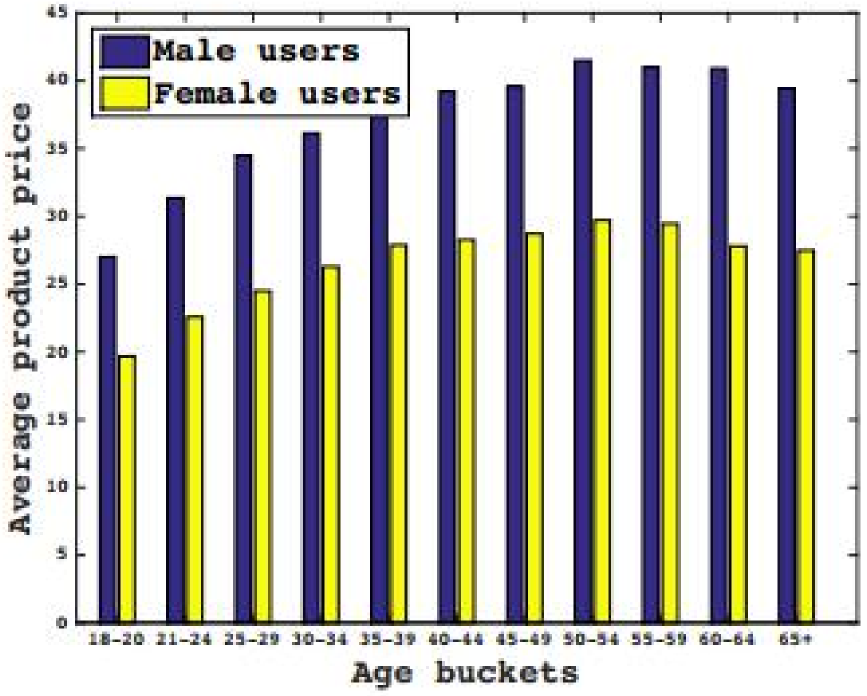




EXPERIMENTS

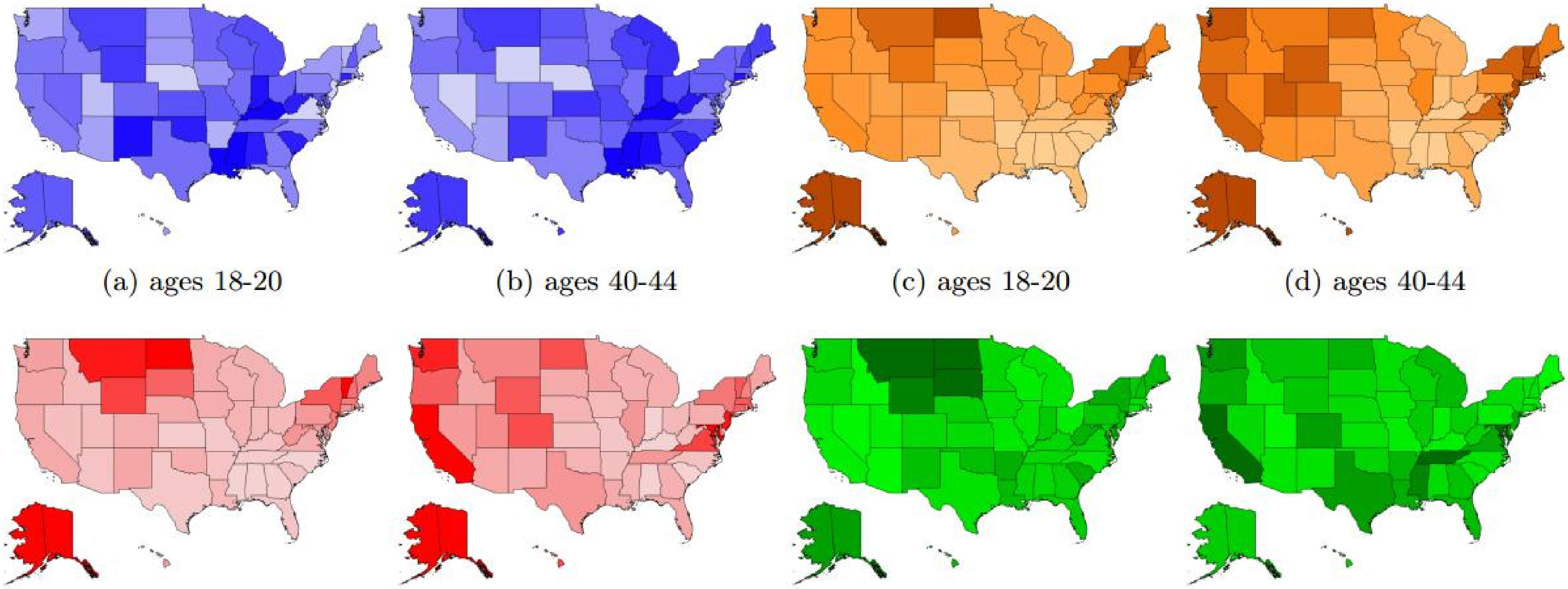


(a) Percentage of purchasing users among all online users



(b) Average product price

EXPERIMENTS



(a) ages 18-20

(b) ages 40-44

(c) ages 18-20

(d) ages 40-44

(e) ages 18-20

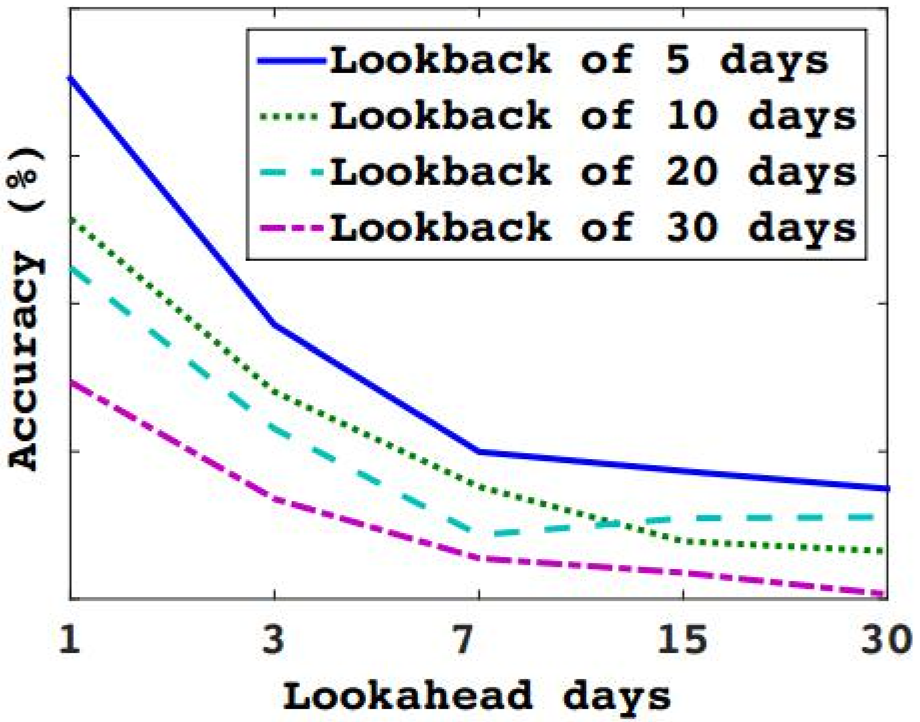
(f) ages 40-44

(g) ages 18-20

(h) ages 40-44

Figure 6: Purchasing behavior for different cohorts, dark color encodes higher values: (a, b) Percentage of shoppers among online users; (c, d) Average number of purchases per user; (e, f) Average amount spent per user; (g, h) Average product price

EXPERIMENTS



EXPERIMENTS

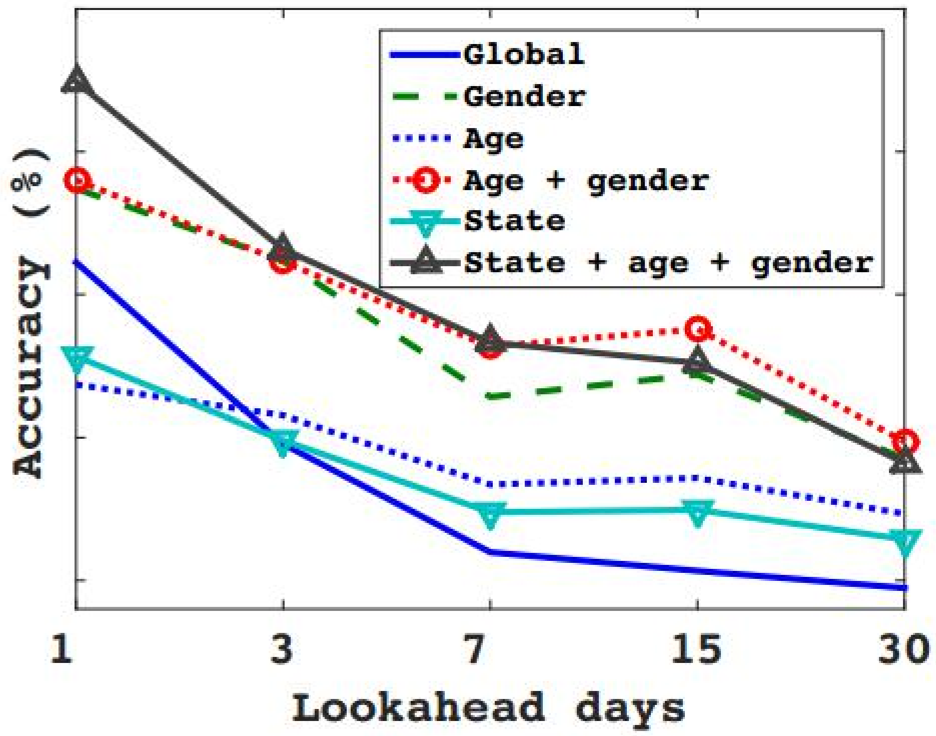


Figure 8: Prediction accuracy of popular products for different user cohorts

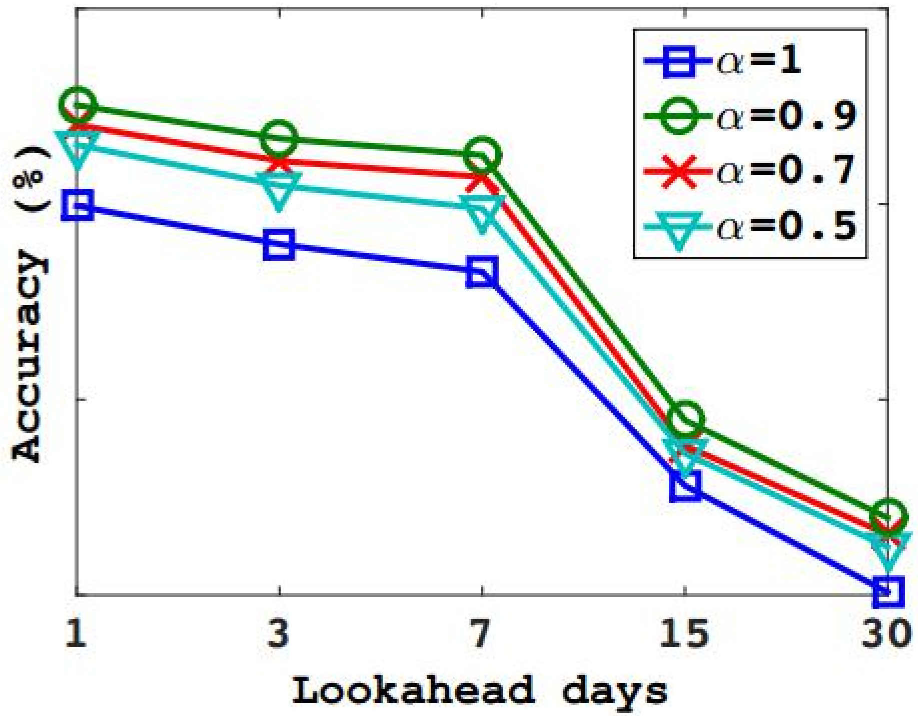
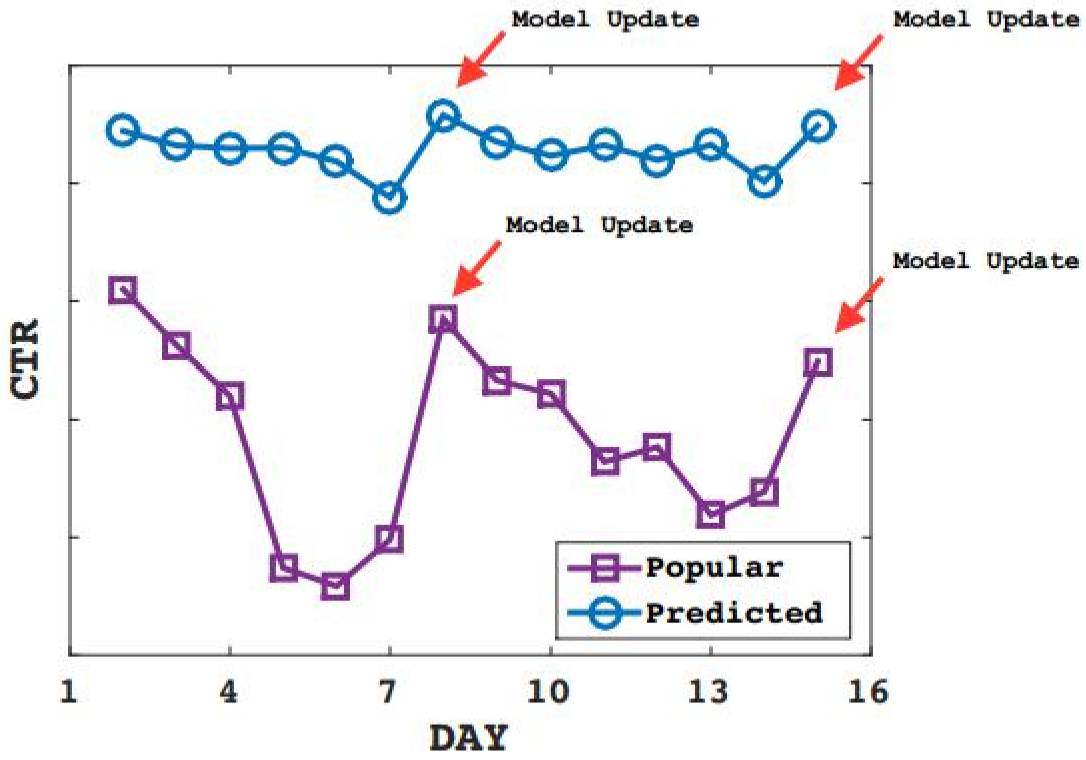


Figure 9: prod2vec accuracy with different decay values

different acc?

EXPERIMENTS



data time:60 days

REFERENCES

[1] E-commerce in Your Inbox:Product Recommendations at Scale
[2] Meta-Prod2Vec - Product Embeddings Using Side-Information for Recommendation
[3] MRNet-Product2Vec: A Multi-task RecurrentNeural Network for Product Embeddings
[4] Distributed Representations of Sentences and Documents