

Cheng-wei WANG, Teng-fei ZHOU, Chen CHEN, Tian-lei HU, Gang CHEN,
2020. HAM: a deep collaborative ranking method incorporating textual
information. *Frontiers of Information Technology & Electronic Engineering*,
21(8):1206-1216. <https://doi.org/10.1631/FITEE.1900382>

HAM: a deep collaborative ranking method incorporating textual information

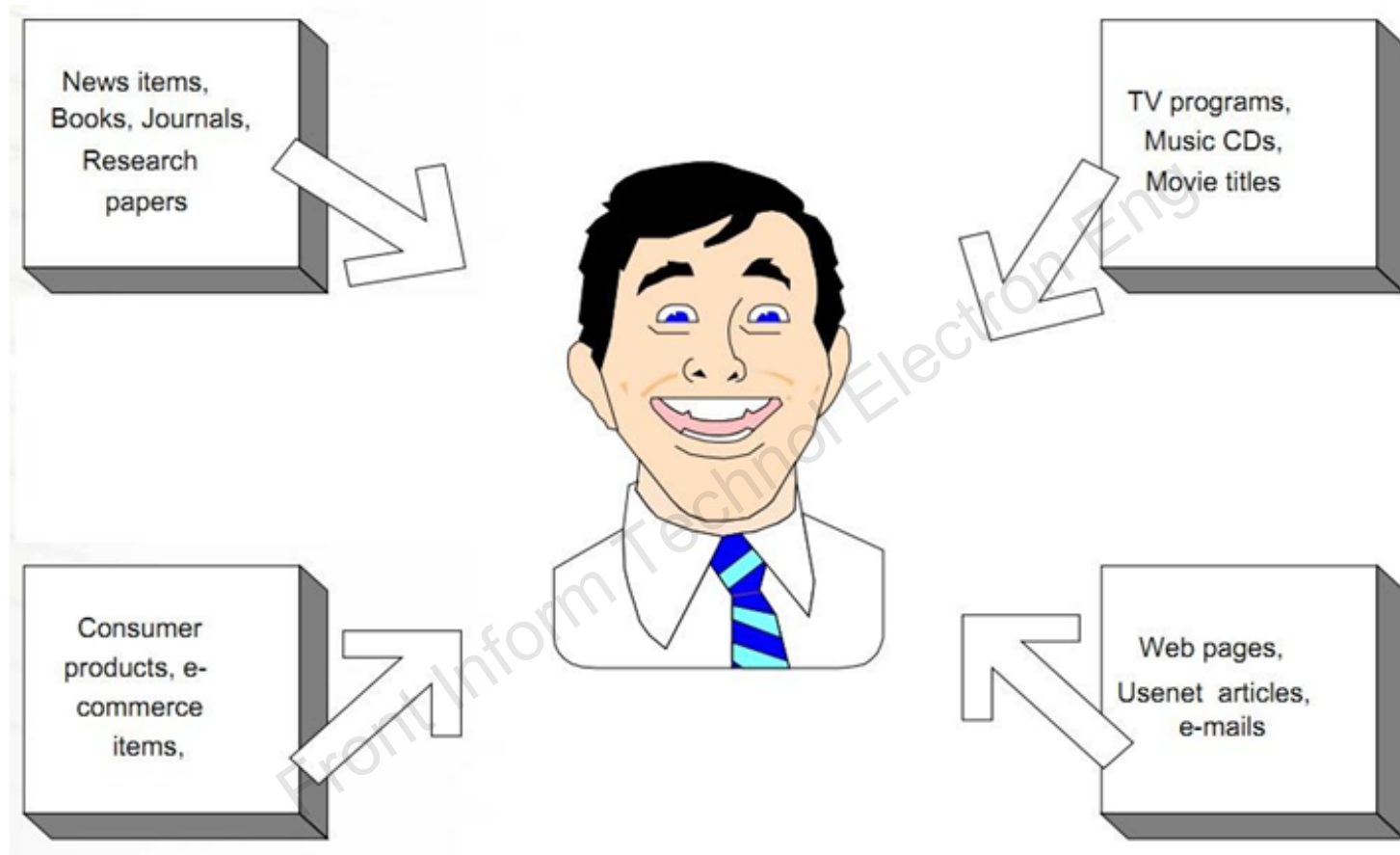
Key words: Deep learning; Recommendation system; Highway
network; Block coordinate descent

Corresponding author: Gang CHEN

E-mail: cg@zju.edu.cn

 ORCID: <https://orcid.org/0000-0002-7483-0045>

Background for a recommendation system

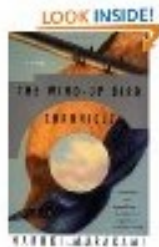


A recommendation system, or a recommender system (sometimes replacing “system” with a synonym such as platform or engine), is a subclass of information filtering system that seeks to predict the “rating” or “preference” a user would give to an item.

Background for a recommendation system (Cont'd)

- Amazon.com personalized book recommendation

Customers Who Bought This Item Also Bought



The Wind-Up Bird Chronicle:
A Novel

> Haruki Murakami

★★★★☆ (382)

Paperback

\$11.53



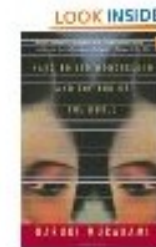
1Q84: 3 Volume Boxed Set
(Vintage International)

> Haruki Murakami

★★★★☆ (424)

Paperback

\$16.34



Hard-Boiled Wonderland and
the End of the World: A ...

> Haruki Murakami

★★★★☆ (165)

Paperback

\$10.63

Background for deep learning

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised, or unsupervised.

- In 1980, Fukushima Kunihiro proposed a new perceptron.
- In 1989, LeCun and other researchers applied the standard backpropagation algorithm proposed in 1974 to deep neural networks for handwritten postal code recognition.
- In 2012, Hinton significantly reduced the error rate of ImageNet image classification problems.
- In 2016, DeepMind's AlphaGo used deep reinforcement learning to train the Go AI to defeat the human world champion.

Background for deep learning based applications

Deep Learning-based Applications

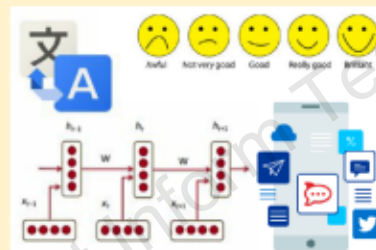
Social Network Analysis



Autonomous Driving



Natural Language Processing



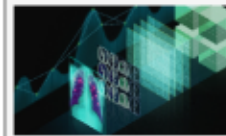
Sentiment Classification
Entity Extraction
Translation

Visual Data Processing



Computer Vision
Multimedia Data Analysis

Biomedicine



Disaster



Speech and Audio Processing



Speech Enhancement
Speech Recognition

Information
Retrieval



Content based recommendation

This method further fully mines the content information of users and items in addition to using feedback data

- content information: **text content**, picture content, audio content, video content

爱尔兰人 The Irishman (2019)



导演: 马丁·斯科塞斯
编剧: 查尔斯·布兰特 / 斯蒂文·泽里安
主演: 罗伯特·德尼罗 / 阿尔·帕西诺 / 乔·佩西 / 安娜·帕奎因 / 杰西·普莱蒙 / 更多...
类型: 剧情 / 传记 / 犯罪
制片国家/地区: 美国
语言: 英语
上映日期: 2019-09-27(纽约电影节) / 2019-11-01(美国点映) / 2019-11-27(美国网络)
片长: 210分钟
又名: 爱尔兰杀手(港) / 听说你刷房子了 / I Heard You Paint Houses
IMDb链接: tt1302006

想看 看过 评价: ☆☆☆☆☆

写短评 写影评 分享到

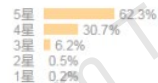
推荐

爱尔兰人的剧情简介 · · · · ·

《爱尔兰人》为马丁·斯科塞斯执导的传奇巨制，罗伯特·德尼罗、阿尔·帕西诺和乔·佩西主演。通过二战老兵弗兰克·希兰的视角，讲述了战后美国有组织犯罪的故事。弗兰克·希兰是一名骗子和杀手，曾经在 20 世纪最恶名昭彰的人物身边工作。该电影跨越数十年，记录了美国历史上最大的悬案之一，即传奇工会领袖吉米·霍法失踪案，以宏大的故事之旅，展现有组织犯罪的隐秘通道：其内部运作、仇敌以及与主流政治的瓜葛。

豆瓣评分

9.1 42159人评价



好于 99% 犯罪片
好于 98% 剧情片



Quaternary Science Reviews
Volume 207, 1 March 2019, Pages 13-36



用户的收藏记录

Earth system impacts of the European arrival and Great Dying in the Americas after 1492

Alexander Koch ^{a, *}, Chris Brierley ^a, Mark M. Mazlin ^a, Simon L. Lewis ^{a, b}

Show more

<https://doi.org/10.1016/j.quascirev.2018.12.004>

Get rights and contents

Under a Creative Commons license

open access

文章主要内容

Highlights

- Combines multiple methods estimating pre-Columbian population numbers.
- Estimates European arrival in 1492 lead to 56 million deaths by 1600.
- Large population reduction led to reforestation of 55.8 Mha and 7.4 Pg C uptake.
- 1610 atmospheric CO₂ drop partly caused by indigenous depopulation of the Americas.
- Humans contributed to Earth System changes before the Industrial Revolution.

Recommended articles 推荐列表

Inferring critical transitions in paleoecological ti...

Quaternary Science Reviews, Volume 207, 2019, pp. 49...

Purchase PDF

Hide details

M. Allison Stegner, Zak Ratajczak, Stephen R. Carpenter, John W. Williams

Inferring critical transitions in paleoecological time series with irregular sampling and variable time-averaging
Quaternary Science Reviews, Volume 207, 2019, pp. 49-63

Corena album archaeobotanical remains in west...

Quaternary Science Reviews, Volume 207, 2019, pp. 1-12

Purchase PDF

Hide details

Carmen M. Martínez-Varea, P. Pablo Ferrer-Gallego, M^a Dolores Raigón, Ernestina Badal, Inmaculada Ferrando-Pardo, Emilio Laguna, Cristina Real, Dídac Roman, Valentín Villaverde

Corena album archaeobotanical remains in western Mediterranean basin. Assessing fruit consumption during Upper Palaeolithic in Cova de les Cendres (Alicante, Spain)
Quaternary Science Reviews, Volume 207, 2019, pp. 1-12

On the eruption age and provenance of the Old...

Quaternary Science Reviews, Volume 207, 2019, pp. 64...

Purchase PDF

Hide details

S.D. Burgess, M.A. Coble, J.A. Vazquez, M.L. Coombs, K.L. Wallace

On the eruption age and provenance of the Old Crow tephra
Quaternary Science Reviews, Volume 207, 2019, pp. 64-79

1 2 Next >

Motivations and challenges

爱尔兰人的短评 ····· (全部 46755 条)

我要写短评

热门 / 最新 / 好友

桃桃林林 看过 ★★★★★ 2019-11-27

1023 有用

最后实在是，，，太伤感了。终究，我们敌不过时间，这才是最可怕的杀手。其实这是一部老年人心态的黑帮片，相比马丁之前的黑帮作品，少了狠辣，多了些温柔与情义，以及一个时代逝去之后的感伤。片子节奏很慢，入戏更慢，前面各种铺线案，差不多一个多小时之后才真正入戏。最后一小时才真正精彩，有犯罪的紧张感，还有最后的感伤。“找我的律师”“他死了”“谁干的”“癌症”。这样的台词，真是笑过之后又难过得想哭。能重看罗伯特... (展开)

Blularry 看过 ★★★★★ 2019-09-28

827 有用

啊啊啊！马丁斯科塞斯阿尔帕西诺罗伯特德尼罗跟我相距十米不到啊！！阿尔帕西诺宝刀未老，依旧是二十年前的那副神韵，罗伯特德尼罗在影片中变回了年轻时的模样，眼袋松弛一时让人不大习惯。想了想，一个学电影的人能跟教父相距不到十米，这辈子也值了！哎我怎么老是有这种没出息的想法。整个影片真的是属于那种回亿几十年迷影铁粉的一份大礼！所有你想看到的华丽调度和宏大场面，有起有伏的历史悬念与黑色幽默，还有影帝们返老还... (展开)

奕奕电影 看过 ★★★★★ 2019-11-28

676 有用

看完老马的《爱尔兰人》了，怎么说呢，他这是用《纯真年代》的方式拍了一部《教父》啊！让这些老家伙们可以借助减龄技术再次返老还童，重回那个逝去的黄金年代，有种时光倒流的感觉。三个半小时的电影，全程配着旁白，用意识流的叙事讲一个如此长时间跨度的故事，浓缩进一个老家伙的一生，这样的故事也只有这群老家伙们做出来才这么史诗这么可信吧。Netflix愿意花这么多钱，投一部这样的电影，也算是这个时代给影迷的馈赠吧... (展开)

- Noise in text content

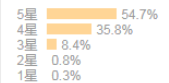
爱尔兰人 The Irishman (2019)



导演: 马丁·斯科塞斯
编剧: 查尔斯·布兰特 / 斯蒂文·泽里安
主演: 罗伯特·德尼罗 / 阿尔·帕西诺 / 乔·佩西 / 安娜·帕奎因 / 杰西·普莱蒙 / 更多...
类型: 剧情 / 传记 / 犯罪
制片国家/地区: 美国
语言: 英语 / 意大利语 / 拉丁语 / 西班牙语
上映日期: 2019-09-27(纽约电影节) / 2019-11-01(美国点映) / 2019-11-27(美国网络)
片长: 209分钟
又名: 爱尔兰杀手(港) / 听说你刷房子了 / I Heard You Paint Houses
IMDb链接: tt1302006

豆瓣评分

8.9 ★★★★★
163711人评价



好于 98% 犯罪片
好于 97% 传记片

想看 看过 评价: ☆☆☆☆☆

写短评 写影评 分享到

推荐

爱尔兰人的剧情简介 ·····

《爱尔兰人》为马丁·斯科塞斯执导的传奇巨制，罗伯特·德尼罗、阿尔·帕西诺和乔·佩西主演。通过二战老兵弗兰克·希兰的视角，讲述了战后美国有组织犯罪的故事。弗兰克·希兰是一名骗子和杀手，曾经在 20 世纪最恶名昭彰的人物身边工作。该电影跨越数十年，记录了美国历史上最大的悬案之一，即传奇工会领袖吉米·霍法失踪案，以宏大的故事之旅，展现有组织犯罪的隐秘通道：其内部运作、仇敌以及与主流政治的瓜葛。

□ Topic 1: Gangster □ Topic 2: love

- Multiple topics

Motivations and challenges (Cont'd)

Motivation I: mining textual content more effectively

Motivation II: designing a very deep neural structure

- The number of neural layers usually determines the neural model's performance. A recommendation system based on a deeper neural network is likely to outperform the shallow ones.
- Due to overparameterization, a deeper neural network generalizes better than shallow ones.

Challenge I

- Textual noise makes a neural network be prone to overfitting.
- It is difficult to identify topics in textual information by neural networks.

Challenge II

- Gradient vanishing/exploding issue.
- Data sparsity of the recommendation problem further exacerbates the instability issue of very deep networks.
- Most state-of-the-art recommenders are based on shallow neural networks.

HighwAy recoMmender (HAM)

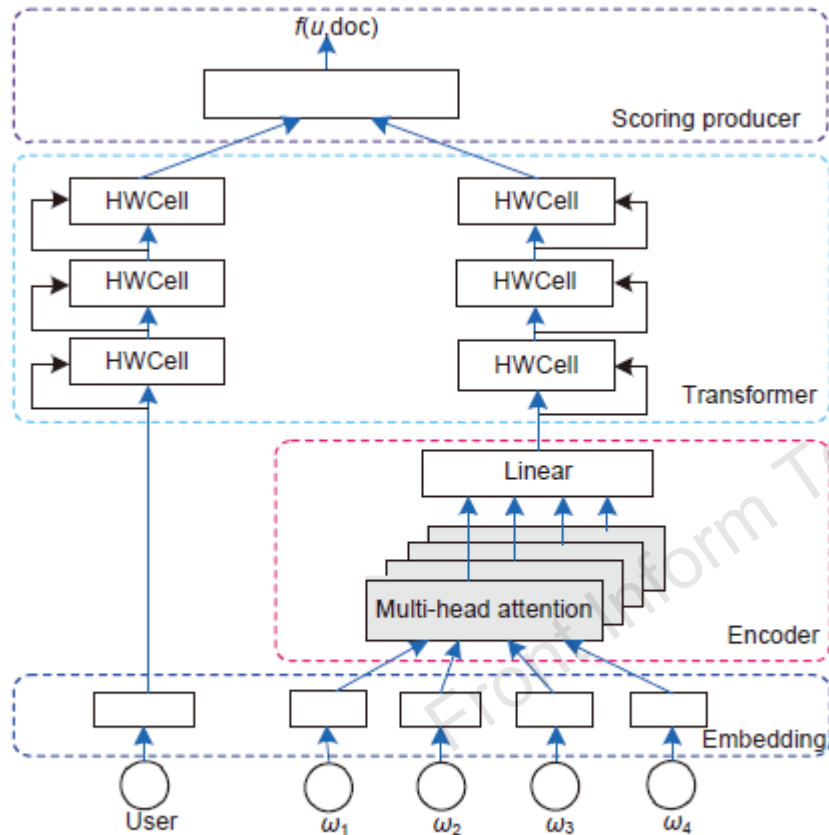


Fig. 1 Architecture of the HighwAy recoMmender (HAM)

Deep transformer

- Deep transformer enhances expressive power of the encoder
- The transformer is constituted of a series of HWCell
- The HWCell $H(x|G, K)$ can be computed by

$$\begin{cases} g = \sigma(Gx), \\ H(x|G, K) = g \odot x + (1 - g) \odot \tanh(Kx), \end{cases}$$

- where g acts as an information gate.
- The HWCell uses gated skip-connection to stabilize the gradient flow of the neural network
- This kind of neural structure can alleviate gradient vanishing/exploding issue of deep neural networks

Scoring module

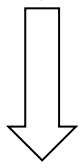
- Suppose the latent vectors of users and items are mapped into u_i and v_j respectively by the deep transformer
- The scoring module describes the match of users and items via:

$$f(i, \text{doc}_j) = u_i^T v_j.$$

Training the model

Train HAM via multitask learning

Ranking task + Word embedding task



$$\min_{\theta_f, C} \mathcal{L}_{\text{rank}}(\theta_f) + \lambda \mathcal{L}_{\text{embd}}(C, E).$$

Coordinate descent method

- The optimization problem can be transformed into the following problem by penalizing constrains

$$\min_{\theta_{\text{opt}}} -\frac{1}{N} \sum_{(i,j,j')} \ln \sigma((P_L[i])^T(Q_L[j] - Q_L[j']))) + \lambda \mathcal{L}_{\text{embd}}(C, E)$$

s.t.

$$\begin{cases} P_l[i] = H(P_{l-1}[i] | G_l, K_l), & i \leq m, 1 \leq l \leq L, \\ Q_l[j] = H(Q_{l-1}[j] | \tilde{G}_l, \tilde{K}_l), & j \leq n, 1 \leq l \leq L, \\ P_0 = P, \\ Q_0[j] = \text{Enc}(E[\text{doc}] | \theta_e), & j \leq n \end{cases}$$

$$\begin{aligned} \min_{\theta_{\text{opt}}} & -\frac{1}{N} \sum_{(i,j,j')} \ln \sigma((P_L[i])^T(Q_L[j] - Q_L[j']))) \\ & \underbrace{\hspace{10em}}_{\mathcal{L}_O(P_L, Q_L)} \\ & + \lambda \mathcal{L}_{\text{embd}}(C, E) + \sum_{l=1}^L \left\{ \underbrace{\frac{\gamma}{m} \|P_l - H_l(P_{l-1})\|^2}_{\mathcal{L}_{H(l)}(P_l, P_{l-1}, G_l, K_l)} \right. \\ & \left. + \underbrace{\frac{\gamma}{n} \|Q_l - \tilde{H}_l(Q_{l-1})\|^2}_{\mathcal{L}_{H(l)}(Q_l, Q_{l-1}, \tilde{G}_l, \tilde{K}_l)} + \underbrace{\frac{\gamma}{m} \|P_0 - P\|^2}_{\mathcal{L}_S(P_0, P)} \right. \\ & \left. + \underbrace{\frac{\gamma}{n} \|Q_0 - Q - E_{\theta_e}(\{\text{doc}_j\})\|^2}_{\mathcal{L}_E(Q_0, Q, E_{\text{doc}_j}, \theta_e)} \right\} \end{aligned}$$

Front Information Technology Electron Eng

Experimental results

Table 3 Performance of the different methods on the collaborative ranking task

Method	AUC			MAP		
	CiteULike	M1M	M10M	CiteULike	M1M	M10M
CMLST	0.9313±0.003	0.9104±0.013	0.9370±0.013	0.1217±0.002	0.1881±0.004	0.1753±0.006
CRAE	0.9206±0.005	0.9157±0.023	0.9682±0.024	0.0580±0.007	0.1856±0.007	0.1585±0.005
CDL	0.9120±0.004	0.9188±0.023	0.9365±0.036	0.1124±0.003	0.1432±0.005	0.1547±0.004
CDAE	0.9348±0.003	0.9107±0.014	0.9366±0.017	0.1211±0.006	0.1630±0.002	0.1712±0.004
CTR	0.9043±0.007	0.9109±0.012	0.9211±0.021	0.0603±0.002	0.1008±0.009	0.1148±0.003
HAM	0.9406±0.002	0.9239±0.009	0.9671±0.010	0.1481±0.008	0.2332±0.004	0.2249±0.007

The best results are in bold

Table 4 Performance of the different methods on the top- k ranking task

Method	Recall@5			Precision@5		
	CiteULike	M1M	M10M	CiteULike	M1M	M10M
CMLST	0.0937±0.002	0.0722±0.003	0.0667±0.013	0.1100±0.006	0.2291±0.004	0.1822±0.006
CRAE	0.0825±0.003	0.0731±0.003	0.0729±0.024	0.1308±0.004	0.3132±0.002	0.2260±0.007
CDL	0.0895±0.007	0.0758±0.006	0.0840±0.036	0.1404±0.006	0.3782±0.010	0.2773±0.008
CDAE	0.0925±0.004	0.0858±0.002	0.0767±0.017	0.1511±0.002	0.3680±0.009	0.2445±0.003
CTR	0.0363±0.001	0.0434±0.002	0.0520±0.013	0.0430±0.001	0.2059±0.003	0.1356±0.005
HAM	0.1151±0.002	0.0958±0.004	0.1038±0.010	0.1647±0.006	0.4142±0.006	0.3532±0.009

The best results are in bold

Experimental results (Cont'd)

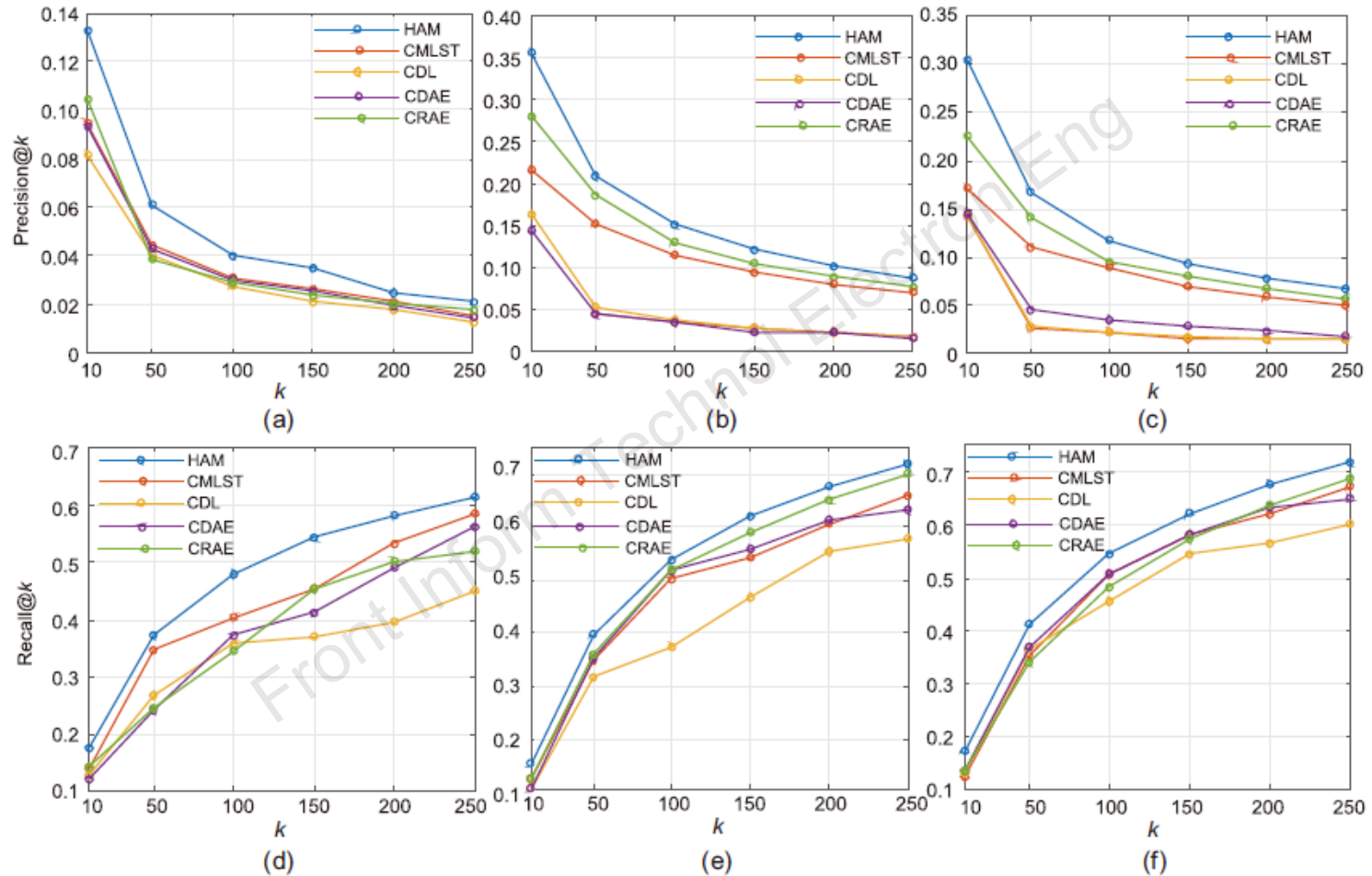


Fig. 2 Precision@ k of the different methods under different choices of k with respect to benchmark datasets CiteULike (a), M1M (b), and M10M (c), and Recall@ k of the different methods under different choices of k with respect to benchmark datasets CiteULike (d), M1M (e), and M10M (f)

Experimental results (Cont'd)

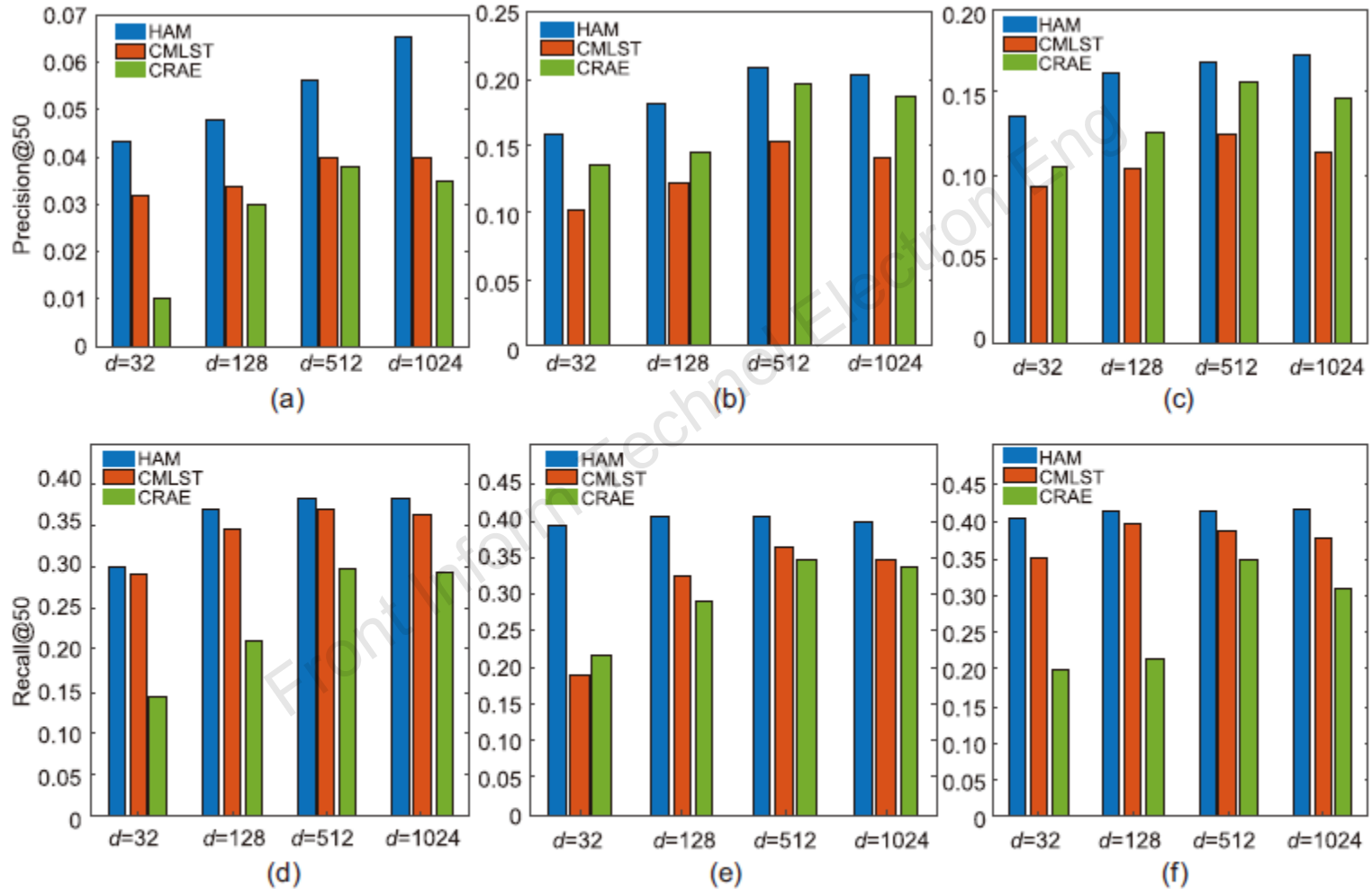


Fig. 3 Precision@50 of the different methods under different choices of dimension with respect to benchmark datasets CiteULike (a), M1M (b), and M10M (c), and Recall@50 of the different methods under different choices of dimension with respect to benchmark datasets CiteULike (d), M1M (e), and M10M (f)

Summary

- We designed a new neural recommendation framework based on the highway network to stabilize the gradient flow of the deep recommender.
- A multi-head attention encoder has been designed to exploit textual information for enhancing the recommendation performance.
- We proposed a novel BCD method to train the deeper recommender effectively.
- Empirical studies showed that the proposed method outperforms state-of-the-art methods significantly in terms of accuracy.



Cheng-wei WANG received the Ph.D. degree in Computer Science from Zhejiang University, Hangzhou, China. He is a data scientist at the Hangzhou Laohe Mountain Technology Corp. His research interests include machine learning, data mining, and recommendation system.



Gang CHEN received the Ph.D. degree in Computer Science from Zhejiang University, Hangzhou, China. He is a professor at the College of Computer Science and Technology, Zhejiang University. He is also Director of the Key Laboratory of Big Data Intelligent Computing of Zhejiang Province. His research interests include database management technology, intelligent computing based big data, and massive Internet system. He is a member of the ACM and a standing member of the China Computer Federation Database Professional Committee.