How to read a Paper

ISMLL

Mohsan Jameel

Outline

Seminar - Big Data Analytics

How to read a paper

- Like novel or newspaper stories, scientific articles needs to be read differently.
- Since they are not books designed for students sometimes they are not self contained and requires some research to be fully understood.
- Understand a paper for a researcher means to be able to implement the described algorithm.

Seminar - Big Data Analytics

How to read a paper

- Skim
- Re–read
- Analyze
- Summarize

Seminar - Big Data Analytics

Skim

First get the "Big picture" by reading the title, abstract, and introduction carefully: this will tell you the major findings and why they matter.

- Quickly scan the article without taking notes: focus on headings and subheadings
- Note the publishing date and conference/journal
- Note terms and parts you don't understand.
  Only with the bigger picture you will understand how much it is necessary to investigate something.

Seminar - Big Data Analytics

Re-read

Read the article again, asking yourself questions such as:

- What problems is the study trying to solve?
- Are findings well supported by evidence?
- Is the study repeatable? (i.e. is the article self contained?)
- If you do not understand take some time to find a brief explanation of what you are not understanding (one-two sentences).
- Is the paper innovative?

Seminar - Big Data Analytics

Interpret

- Examine graphs and tables carefully
- Try to interpret data first before looking at captions
- When reading the discussion and results look after key issues and new findings
- Make sure you have distinguished the main points. If not go over the text again.

Seminar - Big Data Analytics

Summarize

- Take notes and underline key points: it improves reading
- Decide what part of the paper needs to be expanded and how much.

Seminar - Big Data Analytics

Common paper structure

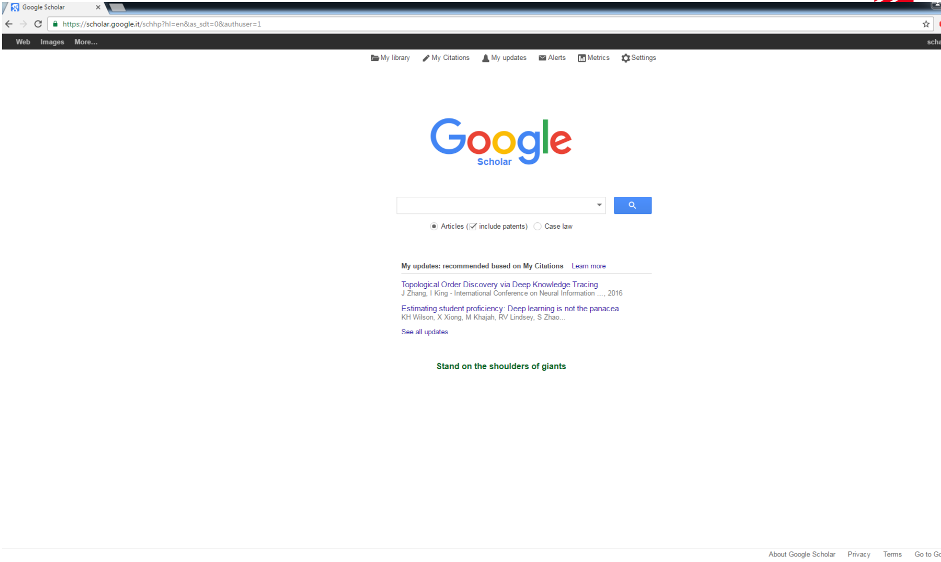
- Abstract
- Introduction
- State of the art
- Algorithms explanation
- Experiments
- Conclusions and future work
- References

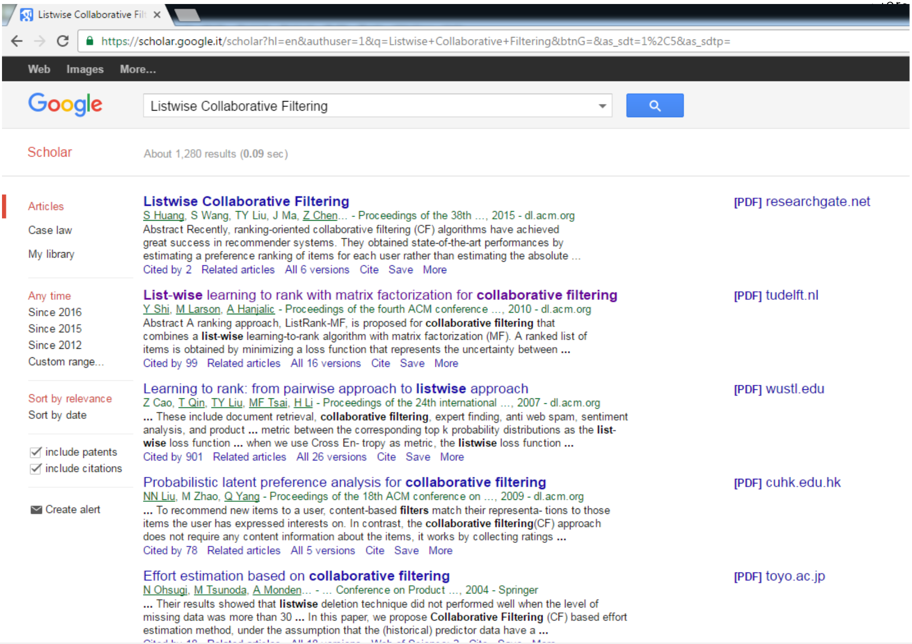
Seminar - Big Data Analytics

Let's take this paper as an example:

"Huang, S., Wang, S., Liu, T. Y., Ma, J., Chen, Z., and Veijalainen, J. (2015, August). **Listwise Collaborative Filtering.** In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 343-352). ACM."

Go to: www.scholar.google.com/


Google Scholar
https://scholar.google.it/schhp?hl=en&as_sdt=0&authuser=1
Web Images More...
My library My Citations My updates Alerts Metrics Settings
Google Scholar
Articles (include patents) Case law
My updates: recommended based on My Citations Learn more
Topological Order Discovery via Deep Knowledge Tracing
J Zhang, I King - International Conference on Neural Information ..., 2016
Estimating student proficiency: Deep learning is not the panacea
KH Wilson, X Xiong, M Khajah, RV Lindsey, S Zhao...
See all updates
Stand on the shoulders of giants
About Google Scholar Privacy Terms


Listwise Collaborative Filt
https://scholar.google.it/scholar?hl=en&authuser=1&q=Listwise+Collaborative+Filtering&btnG=&as_sdt=1%2C5&as_sdtp=
Web Images More...
Google
Listwise Collaborative Filtering
Scholar
About 1,280 results (0.09 sec)
Articles
Case law
My library
Any time
Since 2016
Since 2015
Since 2012
Custom range...
Sort by relevance
Sort by date
include patents
include citations
Create alert
Listwise Collaborative Filtering
S Huang, S Wang, TY Liu, J Ma, Z Chen... - Proceedings of the 38th ..., 2015 - dl.acm.org
Abstract Recently, ranking-oriented collaborative filtering (CF) algorithms have achieved great success in recommender systems. They obtained state-of-the-art performances by estimating a preference ranking of items for each user rather than estimating the absolute ...
Cited by 2 Related articles All 6 versions Cite Save More
[PDF] researchgate.net
List-wise learning to rank with matrix factorization for collaborative filtering
Y Shi, M Larson, A Hanjalic - Proceedings of the fourth ACM conference ..., 2010 - dl.acm.org
Abstract A ranking approach, ListRank-MF, is proposed for collaborative filtering that combines a list-wise learning-to-rank algorithm with matrix factorization (MF). A ranked list of items is obtained by minimizing a loss function that represents the uncertainty between ...
Cited by 99 Related articles All 16 versions Cite Save More
[PDF] tudelft.nl
Learning to rank: from pairwise approach to listwise approach
Z Cao, T Qin, TY Liu, MF Tsai, H Li - Proceedings of the 24th international ..., 2007 - dl.acm.org
... These include document retrieval, collaborative filtering, expert finding, anti web spam, sentiment analysis, and product ... metric between the corresponding top k probability distributions as the list-wise loss function ... when we use Cross En- tropy as metric, the listwise loss function ...
Cited by 901 Related articles All 26 versions Cite Save More
[PDF] wustl.edu
Probabilistic latent preference analysis for collaborative filtering
NN Liu, M Zhao, Q Yang - Proceedings of the 18th ACM conference on ..., 2009 - dl.acm.org
... To recommend new items to a user, content-based filters match their representa- tions to those items the user has expressed interests on. In contrast, the collaborative filtering(CF) approach does not require any content information about the items, it works by collecting ratings ...
Cited by 78 Related articles All 5 versions Cite Save More
[PDF] cuhk.edu.hk
Effort estimation based on collaborative filtering
N Ohsugi, M Tsunoda, A Monden... - ... Conference on Product ..., 2004 - Springer
... Their results showed that listwise deletion technique did not performed well when the level of missing data was more than 30 ... In this paper, we propose Collaborative Filtering (CF) based effort estimation method, under the assumption that the (historical) predictor data have a ...
[PDF] toyo.ac.jp

Seminar - Big Data Analytics

Abstract

- Brief introduction to the topic
- Brief introduction to paper achievements
- Brief summary of the experiments

Seminar - Big Data Analytics

Introduction

- Introduction to the topic
- Introduction of the main concepts
- Introduction of the main state of the art methods
- **State of the art limitations**
- **Hypotheses**
- **Contributions**

Seminar - Big Data Analytics

The Importance of Hypotheses

- It is not enough to describe some new technique or system, some claim about it must be stated and evaluated
- In experimental research, hypotheses typically take one of these two forms:
  - Technique/system X automates task Y for the first time
  - Technique/system X automates task Y better, along some dimension, than each of its rivals
- In theoretical papers, the hypotheses are the statements of theorems and the supporting evidence is their proofs

Seminar - Big Data Analytics

The Importance of Hypotheses

Technique/system X automates task Y better, along some dimension, than each of its rivals, where the dimensions are typically:

- **Behavior:** X has a higher success rate or produces better quality outputs than Y
- **Coverage:** X is applicable to a wider range of examples then Y
- **Efficiency:** X is faster or uses less space then Y
- **Dependability:** X is more reliable, safe or secure than its rivals
- **Maintainability:** X is easier to adapt and extend than its rivals
- **Usability:** Users find X easier to use than its rivals

Seminar - Big Data Analytics

State of the art / Related work

- Is a broad and shallow account of the field, which helps to place the contribution of the paper in context
- What are the rival approaches?
- What are the drawbacks of each?
  - One sentence per method. Is it clear enough?
- How has the battle between different approaches progressed?
- What are the major outstanding problems?

Seminar - Big Data Analytics

Algorithm Explanation

- First the authors introduce the algorithm from which they derived the new algorithm
- Then, the new algorithm is explained
- Contains:
  - Formulas
  - Pseudo code

Algorithm 1: The ListCF Algorithm

Input: An item set I, a user set U, and a rating matrix $R \in \mathbb{R}^{M \times N}$. A set of rated items $I_u \subseteq I$ by each user $u \in U$. The maximal number of iterations $maxIteration$ and error threshold ϵ.

Output: A ranking $\hat{\tau}_u$ of items for each user $u \in U$.

for $u \in U$ **do**
    for $v \in U$ *and* $u \neq v$ **do**
        $P_u, P_v \leftarrow \text{TopKProDist}(I_u, I_v, R)$ /* Eq.1 */
        $sim(u, v) \leftarrow \text{Similarity}(P_u, P_v)$ /* Eq.2 */
    end
    $N_u \leftarrow \text{SelectNeighbors}(\{sim(u, v)\}_{v \in U/u})$
end
for $u \in U$ **do**
    $t = 1$
    repeat
        $\varepsilon = 0$
        $\text{Initialize}(\varphi_u^0)$
        for $g \in \mathcal{G}_k^{T_u}$ **do**
            $\varphi_{u,g}^t \leftarrow \text{Update}(N_u, sim, R)$ /* Eq.8 */
            $\varepsilon += \sqrt{\sum (\varphi_{u,g}^t - \varphi_{u,g}^{t-1})^2}$
        end
        $t \leftarrow t + 1$
    until $t > maxIteration$ *or* $\varepsilon < \epsilon$;
    for $t \in T_u$ **do**
        $P(t) \leftarrow \text{Aggregation}(\{\varphi_{u,g}\}_{g \in \mathcal{G}_k^{T_u}})$
    end
    $\hat{\tau}_u \leftarrow \text{Ordering}(\{P(t)\}_{t \in T_u})$
end

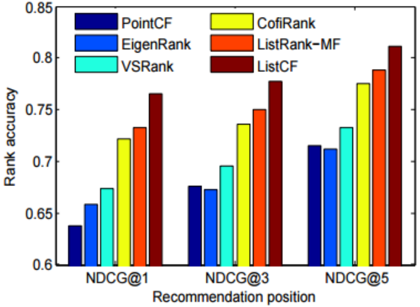
Seminar - Big Data Analytics

Experiments

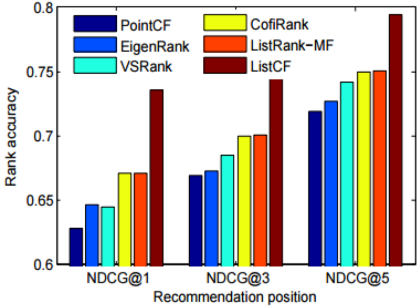
- **Evaluation System**
  - What were the system specifications? cores, nodes, connectivity
  - What was the technology used? MPI, Hadoop, HDFS, Spark, OpenMP
- **Dataset Explanation**
  - What are the available information?
  - What are the available statistics? E.g. number of users, items, sparsity etc.
- **Evaluation protocol**
  - How is the error of the algorithm computed?
  - Are there any other quantitative success measures?
- **Experiments**
  - Are the results statistically significant?

Table 2: Statistics on the three datasets.

	Movielens-1M	EachMovie	Netflix
#users	6,040	36,656	429,584
#items	3,952	1,623	17,770
#ratings	1,000,209	2,580,222	99,884,940
#ratings/user	165.6	70.4	232.5
#ratings/item	253.1	1589.8	5621.0
sparsity	93.7%	95.7%	98.7%



(a) Movielens-1M



(b) EachMovie

Seminar - Big Data Analytics

Conclusions

- Repeats the contributions pointing out specifically how the paper addressed it
- Include future works

Seminar - Big Data Analytics

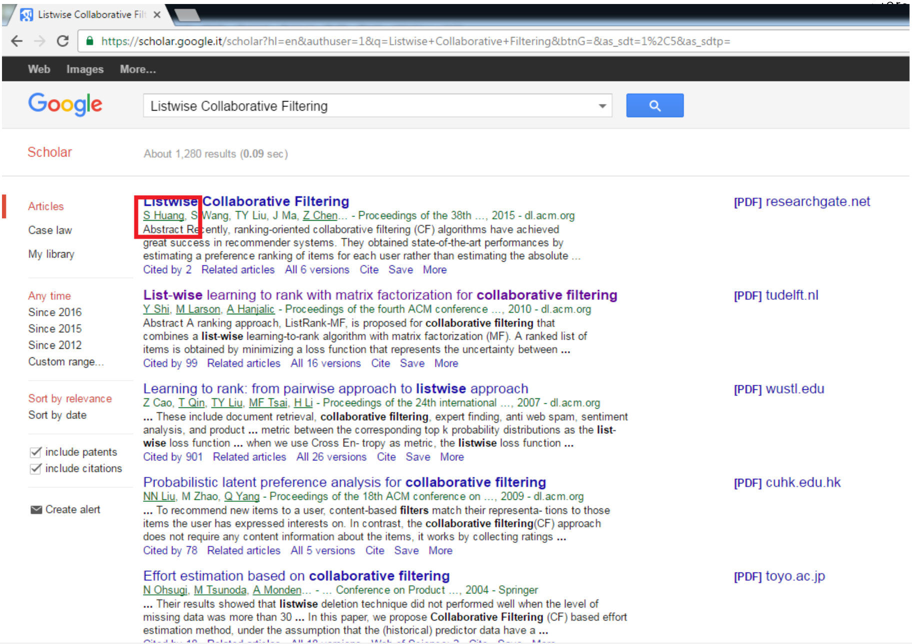
Finding additional material

- If you don't understand something..
- This is not a book, it happens...
  - Try to pose yourself a specific questions
  - Look online

Seminar - Big Data Analytics

Finding additional material

- A book explaining the algorithms
- A PhD thesis
- Tutorials
- Highly related state of the art papers


Listwise Collaborative Filt
https://scholar.google.it/scholar?hl=en&authuser=1&q=Listwise+Collaborative+Filtering&btnG=&as_sdt=1%2C5&as_sdtp=
Web
Images
More...
Google
Listwise Collaborative Filtering
Scholar
About 1,280 results (0.09 sec)
Articles
Case law
My library
Any time
Since 2016
Since 2015
Since 2012
Custom range...
Sort by relevance
Sort by date
include patents
include citations
Create alert
Listwise Collaborative Filtering
S Huang, S Wang, TY Liu, J Ma, Z Chen... - Proceedings of the 38th ..., 2015 - dl.acm.org
Abstract Recently, ranking-oriented collaborative filtering (CF) algorithms have achieved great success in recommender systems. They obtained state-of-the-art performances by estimating a preference ranking of items for each user rather than estimating the absolute ...
Cited by 2 Related articles All 6 versions Cite Save More
[PDF] researchgate.net
List-wise learning to rank with matrix factorization for collaborative filtering
Y Shi, M Larson, A Hanjalic - Proceedings of the fourth ACM conference ..., 2010 - dl.acm.org
Abstract A ranking approach, ListRank-MF, is proposed for collaborative filtering that combines a list-wise learning-to-rank algorithm with matrix factorization (MF). A ranked list of items is obtained by minimizing a loss function that represents the uncertainty between ...
Cited by 99 Related articles All 16 versions Cite Save More
[PDF] tudelft.nl
Learning to rank: from pairwise approach to listwise approach
Z Cao, T Qin, TY Liu, MF Tsai, H Li - Proceedings of the 24th international ..., 2007 - dl.acm.org
... These include document retrieval, collaborative filtering, expert finding, anti web spam, sentiment analysis, and product ... metric between the corresponding top k probability distributions as the list-wise loss function ... when we use Cross En- tropy as metric, the listwise loss function ...
Cited by 901 Related articles All 26 versions Cite Save More
[PDF] wustl.edu
Probabilistic latent preference analysis for collaborative filtering
NN Liu, M Zhao, Q Yang - Proceedings of the 18th ACM conference on ..., 2009 - dl.acm.org
... To recommend new items to a user, content-based filters match their representa- tions to those items the user has expressed interests on. In contrast, the collaborative filtering(CF) approach does not require any content information about the items, it works by collecting ratings ...
Cited by 78 Related articles All 5 versions Cite Save More
[PDF] cuhk.edu.hk
Effort estimation based on collaborative filtering
N Ohsugi, M Tsunoda, A Monden... - ... Conference on Product ..., 2004 - Springer
... Their results showed that listwise deletion technique did not performed well when the level of missing data was more than 30 ... In this paper, we propose Collaborative Filtering (CF) based effort estimation method, under the assumption that the (historical) predictor data have a ...
[PDF] toyo.ac.jp

Sheng-Lung Huang

National Taiwan University

Biomedical imaging, fiber, laser, crystal

Verified email at ntu.edu.tw

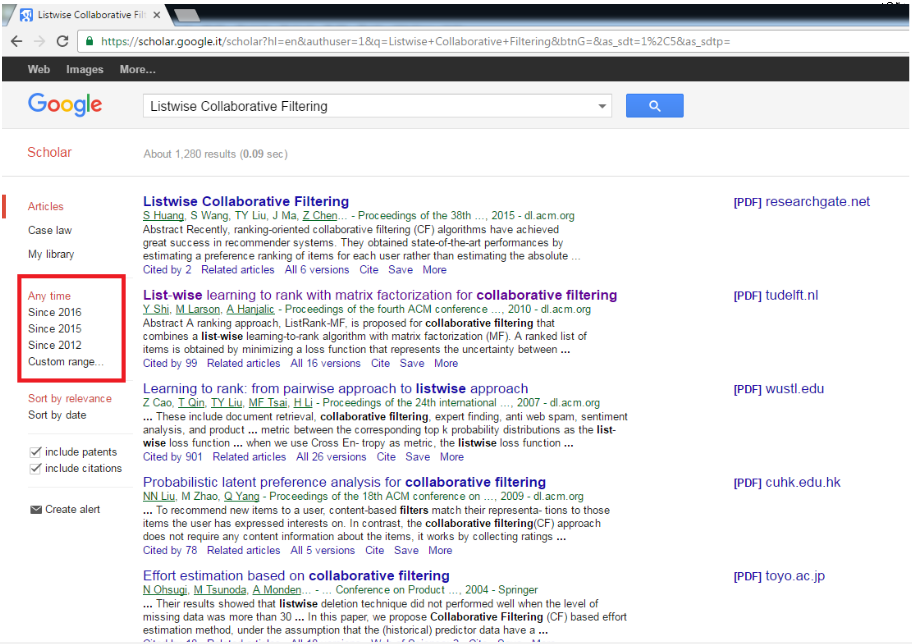
Follow

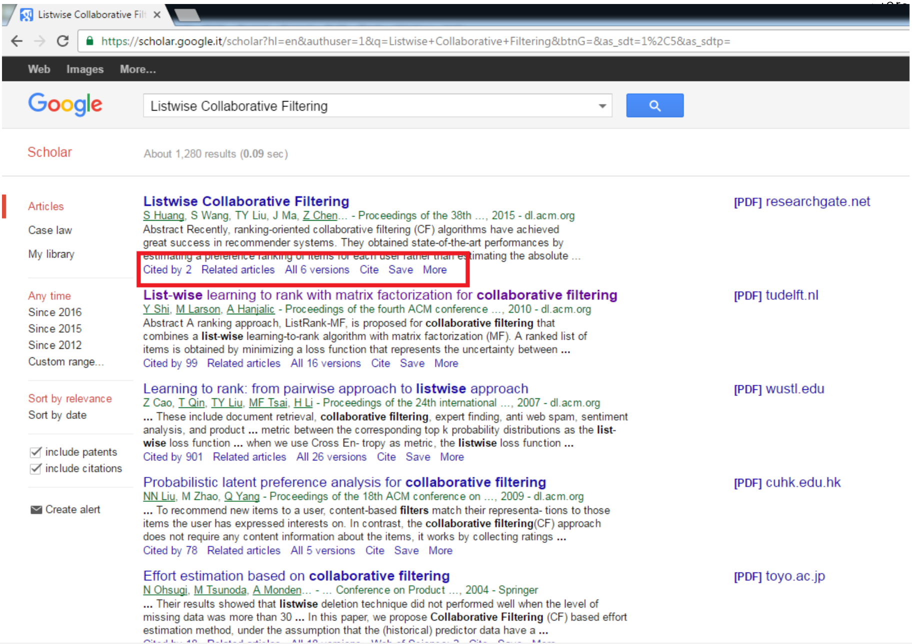
Title 1–20	Cited by	Year
Strain engineering and one-dimensional organization of metal–insulator domains in single-crystal vanadium dioxide beams J Cao, E Ertekin, V Srinivasan, W Fan, S Huang, H Zheng, JWL Yim, ... Nature nanotechnology 4 (11), 732-737	266	2009
Contrasting patterns of retinoblastoma protein expression in mouse embryonic stem cells and embryonic fibroblasts. P Savatier, S Huang, L Szekely, KG Wiman, J Samarut Oncogene 9 (3), 809-818	248	1994
Flooding-induced membrane damage, lipid oxidation and activated oxygen generation in corn leaves B Yan, Q Dai, X Liu, S Huang, Z Wang Plant and soil 179 (2), 261-268	227	1996

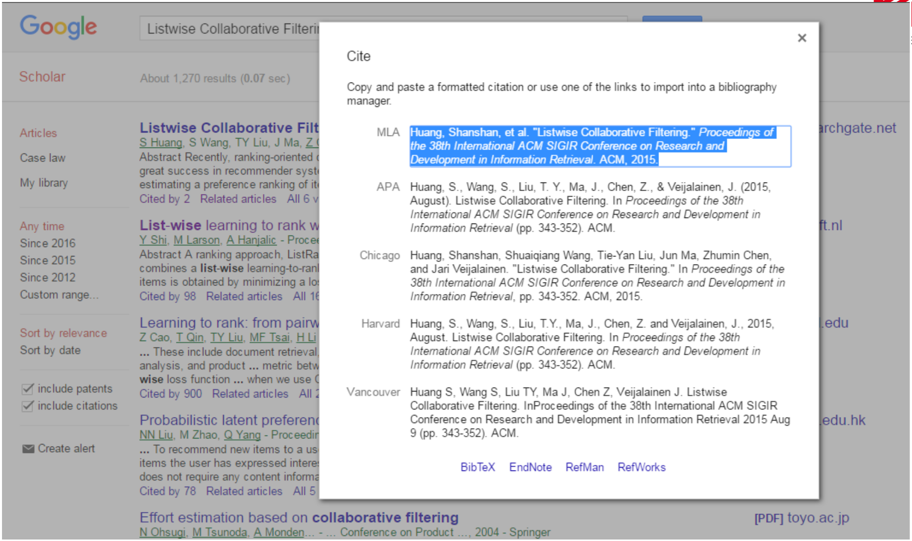
Google Scholar

Get my own profile

Citation indices	All	Since 2011
Citations	15313	10955
h-index	49	39
i10-index	408	298


Listwise Collaborative Fil
https://scholar.google.it/scholar?hl=en&authuser=1&q=Listwise+Collaborative+Filtering&btnG=&as_sdt=1%2C5&as_sdtp=
Web Images More...
Google
Listwise Collaborative Filtering
Scholar
About 1,280 results (0.09 sec)
Articles
Case law
My library
Any time
Since 2016
Since 2015
Since 2012
Custom range...
Sort by relevance
Sort by date
include patents
include citations
Create alert
Listwise Collaborative Filtering
S Huang, S Wang, TY Liu, J Ma, Z Chen... - Proceedings of the 38th ..., 2015 - dl.acm.org
Abstract Recently, ranking-oriented collaborative filtering (CF) algorithms have achieved great success in recommender systems. They obtained state-of-the-art performances by estimating a preference ranking of items for each user rather than estimating the absolute ...
Cited by 2 Related articles All 6 versions Cite Save More
[PDF] researchgate.net
List-wise learning to rank with matrix factorization for collaborative filtering
Y Shi, M Larson, A Hanjalic - Proceedings of the fourth ACM conference ..., 2010 - dl.acm.org
Abstract A ranking approach, ListRank-MF, is proposed for collaborative filtering that combines a list-wise learning-to-rank algorithm with matrix factorization (MF). A ranked list of items is obtained by minimizing a loss function that represents the uncertainty between ...
Cited by 99 Related articles All 16 versions Cite Save More
[PDF] tudelft.nl
Learning to rank: from pairwise approach to listwise approach
Z Cao, T Qin, TY Liu, MF Tsai, H Li - Proceedings of the 24th international ..., 2007 - dl.acm.org
... These include document retrieval, collaborative filtering, expert finding, anti web spam, sentiment analysis, and product ... metric between the corresponding top k probability distributions as the list-wise loss function ... when we use Cross En- tropy as metric, the listwise loss function ...
Cited by 901 Related articles All 26 versions Cite Save More
[PDF] wustl.edu
Probabilistic latent preference analysis for collaborative filtering
NN Liu, M Zhao, Q Yang - Proceedings of the 18th ACM conference on ..., 2009 - dl.acm.org
... To recommend new items to a user, content-based filters match their representa- tions to those items the user has expressed interests on. In contrast, the collaborative filtering(CF) approach does not require any content information about the items, it works by collecting ratings ...
Cited by 78 Related articles All 5 versions Cite Save More
[PDF] cuhk.edu.hk
Effort estimation based on collaborative filtering
N Ohsugi, M Tsunoda, A Monden... - ... Conference on Product ..., 2004 - Springer
... Their results showed that listwise deletion technique did not performed well when the level of missing data was more than 30 ... In this paper, we propose Collaborative Filtering (CF) based effort estimation method, under the assumption that the (historical) predictor data have a ...
[PDF] toyo.ac.jp


Listwise Collaborative Fil
https://scholar.google.it/scholar?hl=en&authuser=1&q=Listwise+Collaborative+Filtering&btnG=&as_sdt=1%2C5&as_sdtp=
Web Images More...
Google
Listwise Collaborative Filtering
Scholar
About 1,280 results (0.09 sec)
Articles
Case law
My library
Any time
Since 2016
Since 2015
Since 2012
Custom range...
Sort by relevance
Sort by date
include patents
include citations
Create alert
Listwise Collaborative Filtering
S Huang, S Wang, TY Liu, J Ma, Z Chen... - Proceedings of the 38th ..., 2015 - dl.acm.org
Abstract Recently, ranking-oriented collaborative filtering (CF) algorithms have achieved great success in recommender systems. They obtained state-of-the-art performances by estimating a preference ranking of items for each user rather than estimating the absolute ...
Cited by 2 Related articles All 6 versions Cite Save More
[PDF] researchgate.net
List-wise learning to rank with matrix factorization for collaborative filtering
Y Shi, M Larson, A Hanjalic - Proceedings of the fourth ACM conference ..., 2010 - dl.acm.org
Abstract A ranking approach, ListRank-MF, is proposed for collaborative filtering that combines a list-wise learning-to-rank algorithm with matrix factorization (MF). A ranked list of items is obtained by minimizing a loss function that represents the uncertainty between ...
Cited by 99 Related articles All 16 versions Cite Save More
[PDF] tudelft.nl
Learning to rank: from pairwise approach to listwise approach
Z Cao, T Qin, TY Liu, MF Tsai, H Li - Proceedings of the 24th international ..., 2007 - dl.acm.org
... These include document retrieval, collaborative filtering, expert finding, anti web spam, sentiment analysis, and product ... metric between the corresponding top k probability distributions as the list-wise loss function ... when we use Cross En- tropy as metric, the listwise loss function ...
Cited by 901 Related articles All 26 versions Cite Save More
[PDF] wustl.edu
Probabilistic latent preference analysis for collaborative filtering
NN Liu, M Zhao, Q Yang - Proceedings of the 18th ACM conference on ..., 2009 - dl.acm.org
... To recommend new items to a user, content-based filters match their representa- tions to those items the user has expressed interests on. In contrast, the collaborative filtering(CF) approach does not require any content information about the items, it works by collecting ratings ...
Cited by 78 Related articles All 5 versions Cite Save More
[PDF] cuhk.edu.hk
Effort estimation based on collaborative filtering
N Ohsugi, M Tsunoda, A Monden... - ... Conference on Product ..., 2004 - Springer
... Their results showed that listwise deletion technique did not performed well when the level of missing data was more than 30 ... In this paper, we propose Collaborative Filtering (CF) based effort estimation method, under the assumption that the (historical) predictor data have a ...
[PDF] toyo.ac.jp


Google
Listwise Collaborative Filteri
Scholar
About 1,270 results (0.07 sec)
Articles
Case law
My library
Any time
Since 2016
Since 2015
Since 2012
Custom range...
Sort by relevance
Sort by date
include patents
include citations
Create alert
Cite
Copy and paste a formatted citation or use one of the links to import into a bibliography manager.
MLA
Huang, Shanshan, et al. "Listwise Collaborative Filtering." Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2015.
APA
Huang, S., Wang, S., Liu, T. Y., Ma, J., Chen, Z., & Veijalainen, J. (2015, August). Listwise Collaborative Filtering. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 343-352). ACM.
Chicago
Huang, Shanshan, Shuaiqiang Wang, Tie-Yan Liu, Jun Ma, Zhumin Chen, and Jari Veijalainen. "Listwise Collaborative Filtering." In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 343-352. ACM, 2015.
Harvard
Huang, S., Wang, S., Liu, T.Y., Ma, J., Chen, Z. and Veijalainen, J., 2015, August. Listwise Collaborative Filtering. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 343-352). ACM.
Vancouver
Huang S, Wang S, Liu TY, Ma J, Chen Z, Veijalainen J. Listwise Collaborative Filtering. InProceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval 2015 Aug 9 (pp. 343-352). ACM.
BibTeX EndNote RefMan RefWorks
Listwise Collaborative Filt
List-wise learning to rank w
Learning to rank: from pairw
Probabilistic latent preferen
Effort estimation based on collaborative filtering
N Ohsugi, M Tsunoda, A Monden... - ... Conference on Product ..., 2004 - Springer
[PDF] toyo.ac.jp