Data Analytics Seminar-1

ISMLL

Prof. Dr. Dr. Lars Schmidt Thieme, Mofassir Arif

Outline

Seminar Details

Text mining Analysis

Finding additional material

Seminar -Text Analysis and Application

Introduction

- The Process of deriving high-quality information from text.
- To turn text into data for analysis through the application of Natural Language Processing techniques.
- Aim of the course is to give an entry level exposure to the machine learning techniques and their uses.
- When? Tuesday 14:00-16:00
- Location: H-2 (Main Campus)

Seminar -Text Analysis and Application

Seminar tasks and activities:

- One paper per person about a topic and a presentation day are assigned
- Prepare a presentation in a small group (3 students):
  - The group has to prepare a presentation:
  - The presentation must be submitted in pre-final version to Mofassir Arif (arifmo@uni-hildesheim.de) one week in advance
  - If the presentation is not well done, part of it, or the complete presentation, will be canceled (Students will be informed a few days in advanced)
  - Peer Review: 3 of your peers will receive the presentation anonymously and their feedback will be referred back to you

Seminar -Text Analysis and Application

Grading

- Presenting the work to the class (50% of the mark)
- Submission of the Summary Paper due 4 weeks after term break (50% of the mark)

Seminar -Text Analysis and Application

Each group member has to prepare a presentation which consists of four parts:

- Introduce the topic
- Summarize the papers (This is the main part)
- Underline differences and similarities of the algorithms

It is important to:

- Involve the audience, will be counted as part of the mark
- Not omit crucial parts of the paper such as the evaluation, the algorithms, the baselines, etc.
- Try to provide your own interpretation of the models

Seminar -Text Analysis and Application

The group presents the topic

- The students will present 60 minutes (20 minutes each)
- After that 30 minutes for questions and answers
- If you don't present you will get a 5.0 as a presentation mark and that automatically results in a failed exam.

Seminar -Text Analysis and Application

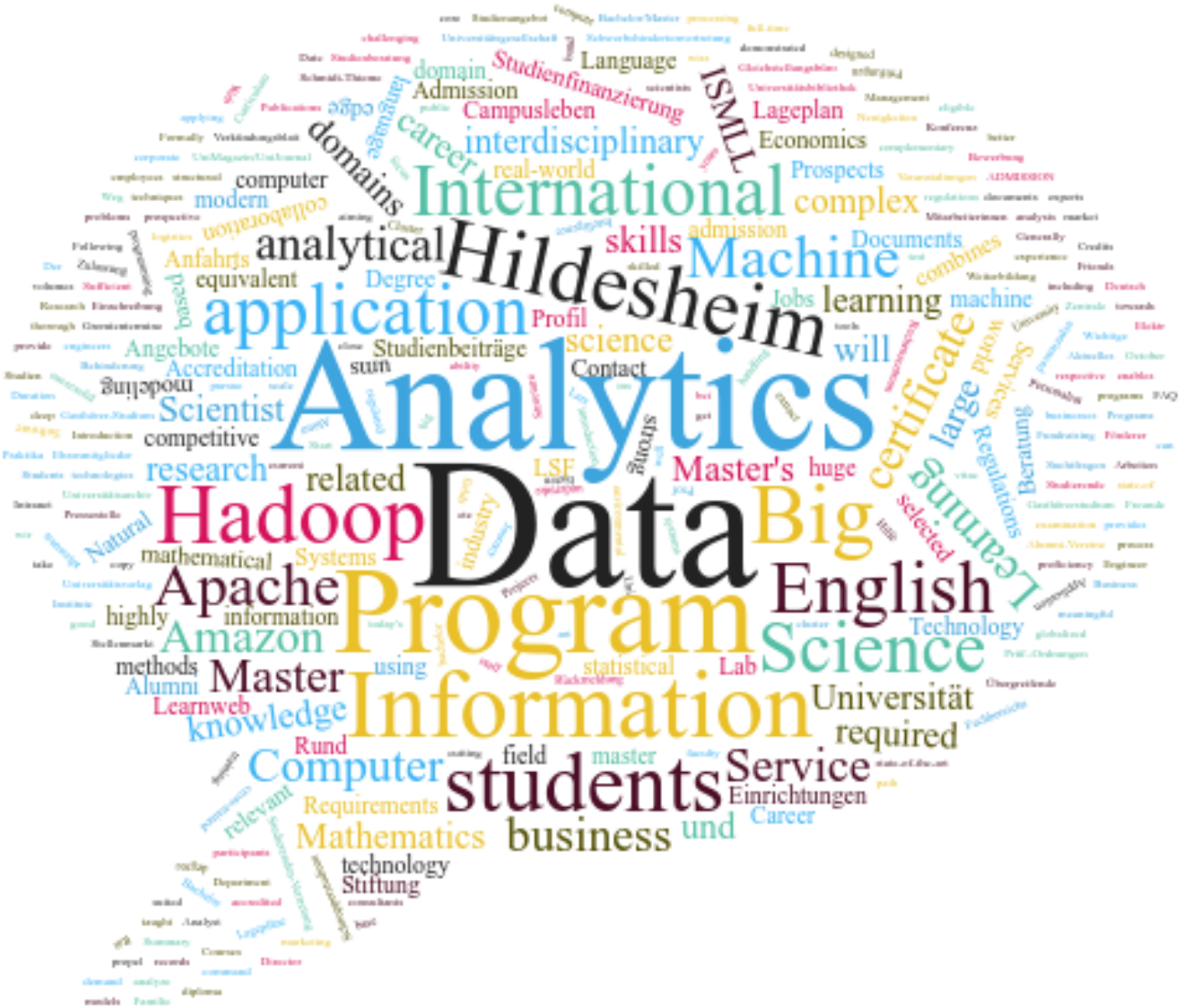
Summary Paper:

- Will be a paper like document, one for each participant, of exactly 15 pages (not one more not one less)
  - Introduce the topic
  - Summarize the paper (This is the main part)
  - Underline differences and similarities of the algorithms of your group
  - Argument why your method is or is not the best of the similar ones seen.
- Submit three hard copies and one digital copy to our secretary (hinzemelching@ismll.uni-hildesheim.de)
- A template will be provided
- More details in the next lecture

Seminar -Text Analysis and Application

Semester Plan

- Two meetings about:
  - Paper reading how to
  - Summary Paper writing how to
- Weekly presentations
- Submission of the Summary Paper
- **Attendance**: You can only miss 2 presentations.


Data
Analytics
Hadoop
Program
Information
Science
International
Hildesheim
Machine
application
Big
English
Apache
Amazon
Master
Computer
students
Service
business
Mathematics
Learning
certificate
Language
Admission
Campusleben
Studienfinanzierung
ISMLL
Lageplan
Economics
interdisciplinary
career
domains
domain
Prospects
complex
real-world
computer
modern
analytical
skills
admission
Documents
Anfahrts
equivalent
Degree
Profil
Jobs
learning
machine
science
Studienbeiträge
Contact
will
Angebote
Accreditation
modelling
Scientist
competitive
research
related
LSF
Master's
huge
Natural
mathematical
Systems
industry
highly
information
Technology
methods
using
statistical
Lab
Alumni
Learnweb
knowledge
Universität
required
Rund
field
master
Requirements
relevant
Einrichtungen
Career
und
technology
Stiftung
Beratung
Regulations
Services
large
selected

Seminar -Text Analysis and Application

A: Machine learning in automated text categorization

Survey Paper and a must read for everyone

Themes

- **Fundamentals**
  - B-1: Stochastic gradient descent training for L1-regularized log-linear models with cumulative penalty
  - B-2: Curriculum Learning
  - B-3: Combined Regression and Ranking

Seminar -Text Analysis and Application

Themes

- **Text Categorization**
  - C-1: Text Categorization with Support Vector Machines. How to Represent Texts in Input Space?
  - C-2: Effective Use of Word Order for Text Categorization with Convolutional Neural Networks
  - C-3: Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification

Seminar -Text Analysis and Application

Themes

- **Text Categorization**
  - D-1: An Effective Approach to Enhance Centroid Classifier for Text Categorization
  - D-2: Inductive learning algorithms and representations for text categorization
  - D-3: Character-level Convolutional Networks for Text Classification

Seminar -Text Analysis and Application

Themes

- **Sentiment Analysis**
  - E-1: Thumbs up?: sentiment classification using machine learning techniques
  - E-2: Twitter as a Corpus for Sentiment Analysis and Opinion Mining
  - E-3: Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts

Seminar -Text Analysis and Application

Themes

- **Sentiment Analysis**
  - F-1: Recognizing contextual polarity in phrase-level sentiment analysis
  - F-2: OpinionMiner: a novel machine learning system for web opinion mining and extraction
  - F-3: Coooolll: A Deep Learning System for Twitter Sentiment Classification

Seminar -Text Analysis and Application

Themes

- **Sentiment Analysis**
  - G-1: Twitter Sentiment Classification using Distant Supervision
  - G-2: Active learning for imbalanced sentiment classification
  - G-3: Context-Sensitive Twitter Sentiment Classification Using Neural Network

Seminar -Text Analysis and Application

Themes

- **Applications**
  - H-1: PTE: Predictive Text Embedding through Large-scale Heterogeneous Text Networks
  - H-2: FastXML: a fast, accurate and stable tree-classifier for extreme multi-label learning
  - H-3: Large-scale Multi-label Learning with Missing Labels

Seminar -Text Analysis and Application

Themes

- **Applications**
  - I-1: A Machine Learning Approach to Twitter User Classification
  - I-2: Broadly Improving User Classification via Communication-Based Name and Location Clustering on Twitter
  - I-3: Twitter-Based User Modeling for News Recommendations

Seminar -Text Analysis and Application

Themes

- **Applications**
  - J-1 Web-Search Ranking with Initialized Gradient Boosted Regression Trees
  - J-2: Mining text snippets for images on the web
  - J-3: Smart Reply: Automated Response Suggestion for Email

Seminar -Text Analysis and Application

Themes

- **Applications**
  - K-1: A system to grade computer programming skills using machine learning
  - K-2: Top-k Multiclass SVM
  - K-3: Robust Top-k Multi-class SVM for Visual Category Recognition

Seminar -Text Analysis and Application

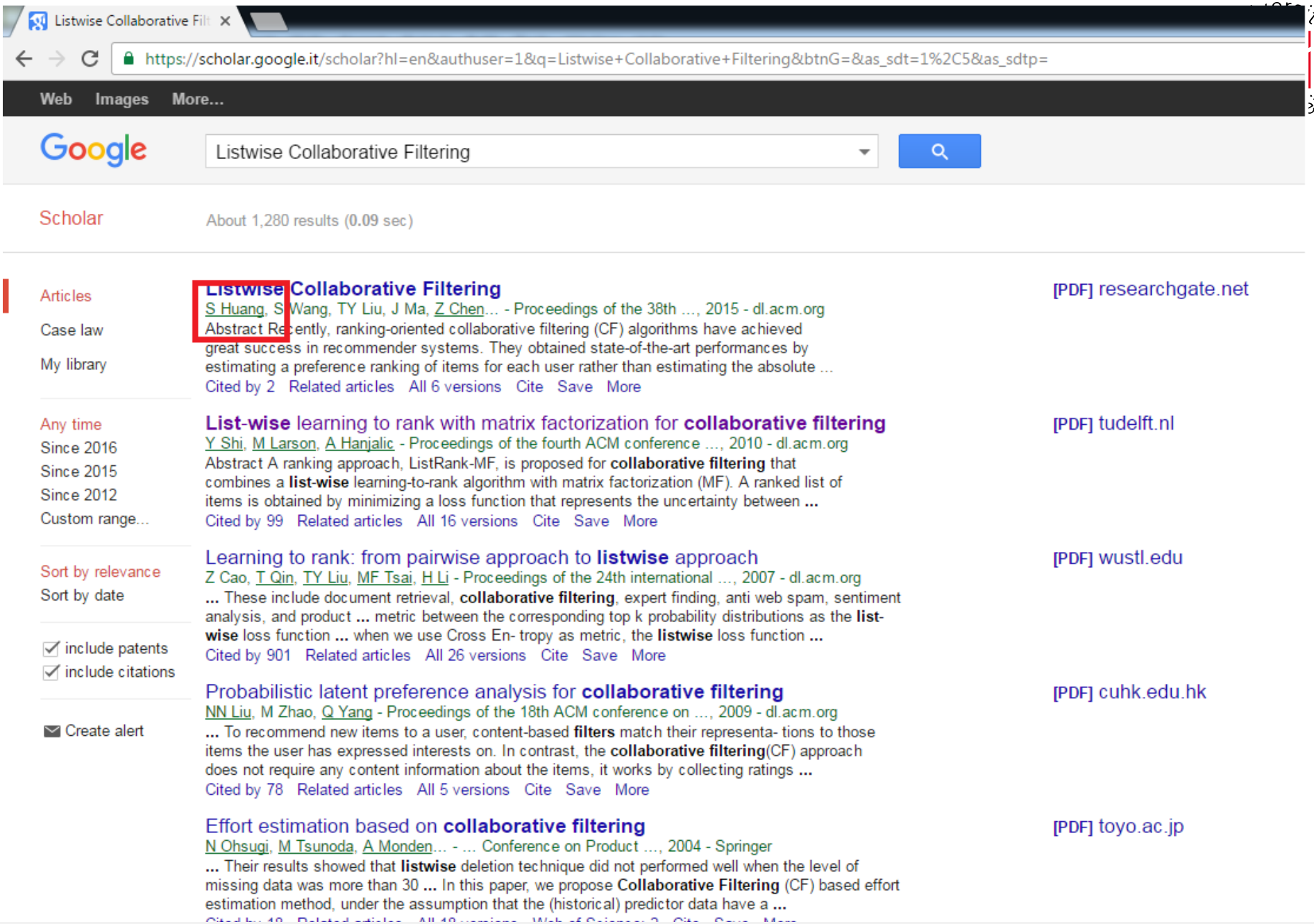
Finding additional material

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

Seminar -Text Analysis and Application

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

Follow

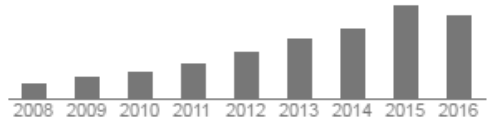
National Taiwan University
Biomedical imaging, fiber, laser, crystal
Verified email at ntu.edu.tw

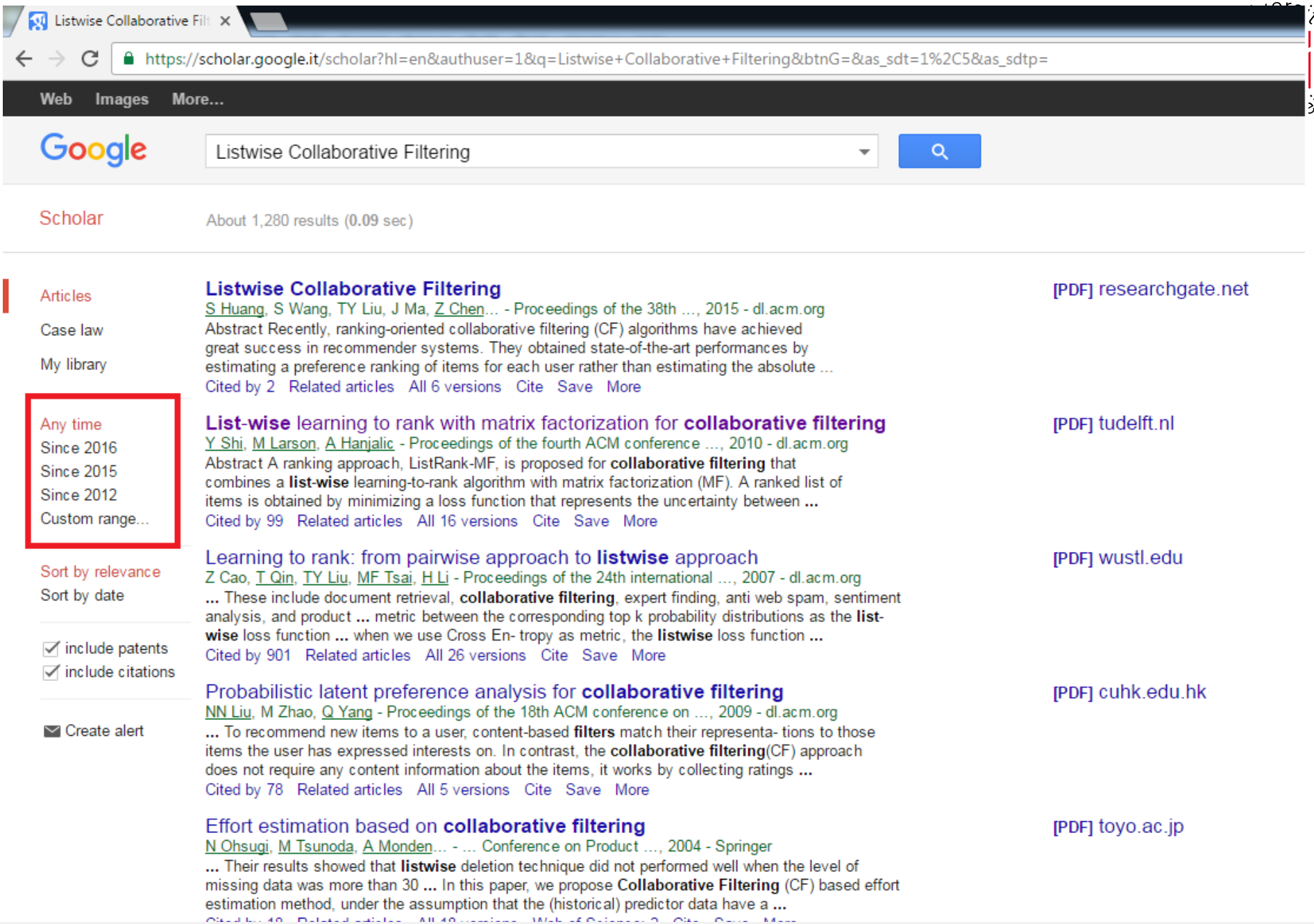
Title 1–20	Cited by	Year
Strain engineering and one-dimensional organization of metal–insulator domains in single-crystal vanadium dioxide beams J Cao, E Ertekin, V Srinivasan, W Fan, S Huang, H Zheng, JWL Yim, ... Nature nanotechnology 4 (11), 732-737	266	2009
Contrasting patterns of retinoblastoma 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

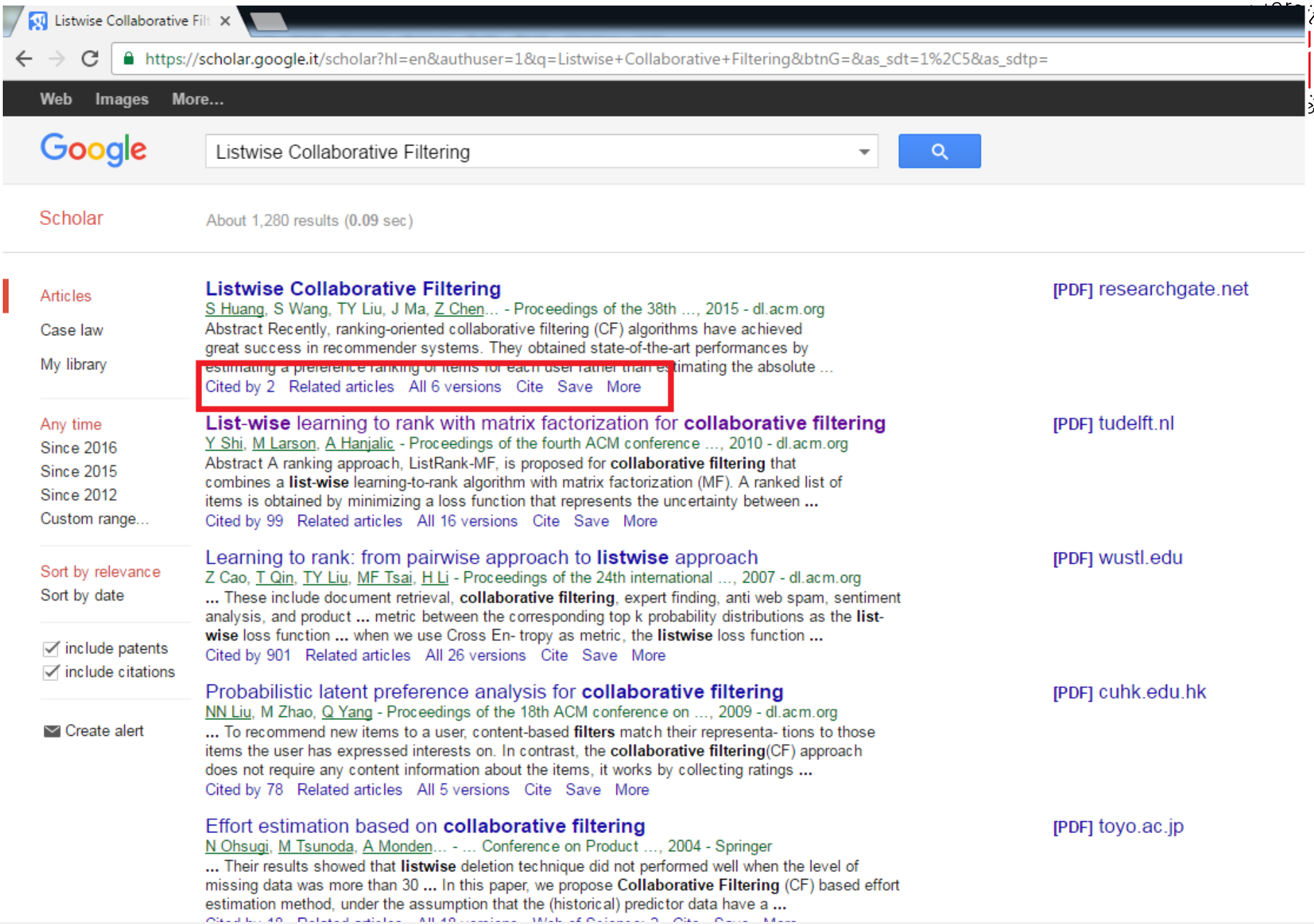
Google Scholar

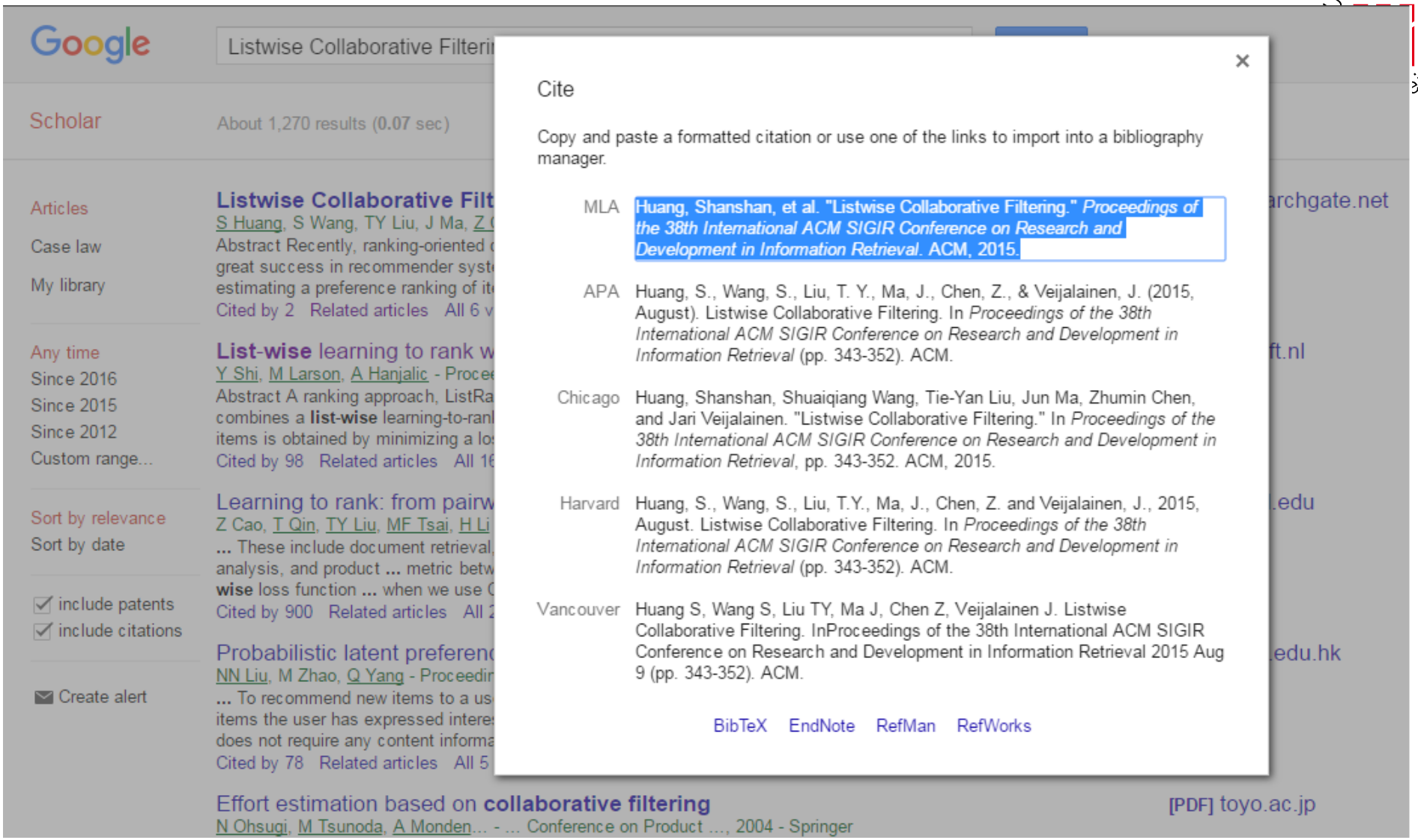
Get my own profile

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




Listwise Collaborative 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


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


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
Listwise Collaborative Filt
S Huang, S Wang, TY Liu, J Ma, Z
Abstract Recently, ranking-oriented
great success in recommender syst
estimating a preference ranking of it
Cited by 2 Related articles All 6 v
List-wise learning to rank w
Y Shi, M Larson, A Hanjalic - Proce
Abstract A ranking approach, ListRa
combines a list-wise learning-to-ran
items is obtained by minimizing a lo
Cited by 98 Related articles All 16
Learning to rank: from pairw
Z Cao, T Qin, TY Liu, MF Tsai, H Li
... These include document retrieval,
analysis, and product ... metric betw
wise loss function ... when we use
Cited by 900 Related articles All 2
Probabilistic latent preferenc
NN Liu, M Zhao, Q Yang - Proceedir
... To recommend new items to a us
items the user has expressed interes
does not require any content informa
Cited by 78 Related articles All 5
Effort estimation based on collaborative filtering
N Ohsugi, M Tsunoda, A Monden... - ... Conference on Product ..., 2004 - Springer
[PDF] toyo.ac.jp
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

Seminar -Text Analysis and Application

Tutor Information

Mofassir ul Islam Arif
arifmo@uni-hildesheim.de
C206
Open Hours: Thursdays 14:00-16:00