

Web Mining – Service

Andreas Hotho, Gerd Stumme

All listed URLs and publications, together with additional references, are available online at <http://www.bibsonomy.org/group/kde/KI2007WebMining>. Further references can be found at <http://www.bibsonomy.org/tag/web+mining> and <http://www.bibsonomy.org/tag/kdd>. These lists can be extended by the community.

Books

- S. Chakrabarti: *Mining the Web: Discovering Knowledge from Hypertext Data*, Morgan-Kaufmann 2002.
- P. Baldi, P. Frasconi, P. Smyth: *Modeling the Internet and the Web: Probabilistic Methods and Algorithms*. Wiley, April 2003.
- B. M. Masand, M. Spiliopoulou (eds.): *Web Usage Analysis and User Profiling, International WEBKDD'99 Workshop, San Diego, California, USA, August 15, 1999, Revised Papers*, volume 1836 of *Lecture Notes in Computer Science*. Springer, 2000.
- B. Berendt, A. Hotho, D. Mladenic, M. van Someren, M. Spiliopoulou, G. Stumme (eds.): *Web Mining: From Web to Semantic Web*. LNAI 3209, Springer 2004.

Further references can be found at: <http://www.kdnuggets.com/publications/web-mining.html>

Conference and Workshop Series

No conference is targeting uniquely the Web Mining area, but all of the following conferences consider it as a major part.

- The *Intl. Conf. on Knowledge Discovery and Data Mining* (KDD) is always located in US and is the leading KDD conference (<http://www.kdd2007.com/>). More application oriented than, e.g., ICML.
- *Intl. Conf. on Machine Learning* (ICML) focuses more on theory (<http://oregonstate.edu/conferences/icml2007/>). ICML is located all over the world.
- The ECML/PKDD (European Conf. on Machine Learning and European Conf. on Principles and Practice of Knowledge Discovery in Databases) is the joint European counterpart to ICML and KDD <http://www.ecmlpkdd2007.org/>.
- The *IEEE Intl. Conf. on Data Mining* (ICDM) <http://www.cs.uvm.edu/~icdm/> is located all over the world. It focuses more on the theoretical aspects of data mining but discusses also application oriented issues.
- *Intl. Sunbelt Social Network Conferences*. Informal meeting point of the Social Network Analysis community. http://www.insna.org/INSNA/sunbelt_inf.html
- There is a series of (Semantic) Web Mining workshops co-located with KDD (called WEBKDD) and ECML/PKDD since 1999.

Upcoming Conferences and Workshops

- All conferences listed above also take place this year.
- Web Mining 2.0, Workshop at ECML/PKDD 2007, <http://www.kde.cs.uni-kassel.de/ws/ecmlpkdd2007/>,
- Bridging the Gap between Semantic Web and Web 2.0, Workshop at ESWC 2007, <http://www.kde.cs.uni-kassel.de/ws/eswc2007/>
- International Workshop on Data Mining in Web 2.0 Environments, Workshop at ICDM 2007, <http://www.kde.cs.uni-kassel.de/ws/Web2DM/>
- Web Mining and Social Network Analysis, Workshop at KDD 2007, <http://workshops.socialnetworkanalysis.info/websnakdd2007/>
- Prior Conceptual Knowledge in Machine Learning and Knowledge Discovery, Workshop at ECML/PKDD 2007, <http://keg.vse.cz/prick107/>
- Workshop On Data Mining For User Modeling at UM 2007, http://www.iit.demokritos.gr/um2007/ws_DMUM.php
- Conference of the German Classification Society on Data Analysis, Machine Learning, and Applications, 2008

Networks of Excellence

Knowledge Discovery in Ubiquitous Environments (KD^{ubiq}) is a EU FP7 Coordination Action settled around the emerging area of Ubiquitous Knowledge Discovery (<http://www.kdubiq.org/>). Its successor KNet, a Network of Excellence, lists also many related projects (http://www.kdnet.org/kdnet/control/related_projects). *Pattern Analysis, Statistical Modelling and Computational Learning* (PASCAL) is a EU FP6 Network of Excellence with a focus on statistical modelling and its applications (<http://www.pascal-network.org/>).

Organizations

- Fachgruppe Knowledge Discovery, Data Mining und Maschinelles Lernen (FG-KDML), former (FGML), <http://www.ais.fraunhofer.de/~wrobel/fgml/>
- International Society for Social Network Analysis (INSNA), <http://www.insna.org/>
- German Classification Society (GfKI) <http://www.gfki.de/>

Datasets

- EMail Spam Detection, <http://www.ecmlpkdd2006.org/challenge.html>
- Enron EMail, <http://www.cs.cmu.edu/~enron/>
- AOL query log, <http://www.gregsadetsky.com/aol-data/>
- BibSonomy Social Bookmarking, <http://www.bibsonomy.org/faq#faq-dataset-1>