

REFERENCES

Tf-idf weighting. (2009, April 7). Retrieved May 7, 2013, from [http://nlp.stanford.edu/
http://nlp.stanford.edu/IR-book/html/htmledition/tf-idf-weighting-1.html](http://nlp.stanford.edu/http://nlp.stanford.edu/IR-book/html/htmledition/tf-idf-weighting-1.html)

Tokenizing. (2011, August). Retrieved May 7, 2014, from <http://www.coli.uni-saarland.de/~schulte/Teaching/ESLLI-06/Referenzen/Tokenisation/schmid-hsk-tok.pdf>

Ahat, M., Amor, S. B., & Bui, M. (2009). *Document Classification with LSA and Pretopology*. Laboratoire des Usages en Technologies d'Information Numériques (LUTIN) .

Al-Shalabi, R. (2004). *Stop-Word Removal Algorithm for Arabic Language*.

Argamon, S., Whitelaw, C., Chase, P., & Dhawle, S. (2005). Stylistic Text Classification Using Functional Lexical Features.

Asy'arie, A. D., & Pribadi, A. W. (2009). Automatic News Articles Classification in Indonesian Language by Using Naive Bayes Classifier Method. *iiWAS*.

Basnur, P. W., & Sensuse, D. I. (2010). PENGKLASIFIKASIAN OTOMATIS BERBASIS ONTOLOGI UNTUK ARTIKEL BERITA BERBAHASA INDONESIA. *MAKARA JOURNAL OF TECHNOLOGY*.

Chee-Hong, Aixin, C., & Lim, S. E.-P. (2001). Automated Online News Classification with Personalization. *Proceedings of the 4th International Conference of Asian Digital Library*, 320-329.

Coporation, Intel. (2001). *OpenCv Reference Manual*.

Gebre, B. G., & Zampier, M. (2012). *Improving Native Language Identification with TF-IDF Weighting*. Cologne: University of Cologne.

Goller, C., & Löning, J. (n.d.). *Automatic Document Classification: A thorough Evaluation of various Methods*. München, Germany: Speech and Artificial Intelligence Labs.

Gupta, V., & Lehal, G. S. (2010). A Survey of Text Summarization Extractive Techniques. *JOURNAL OF EMERGING TECHNOLOGIES IN WEB INTELLIGENCE*, 260.

Kadiri, J. A., & Adetoro, N. A. (2012). Information Explosion and The Challenges of Information and Communication Technology Utilization in Nigerian Libraries and Information Centres. *Ozean Journal of Social Sciences* 5.

La Lei, G. Q. (2012). LDA boost classification: boosting by topics. *EURASIP*, 33.

Landauer, T. K., & Foltz, P. W. (1997). *An Introduction to Latent Semantic Analysis*.

Landauer, T. K., & Foltz, P. W. (1998). *An Introduction to Latent Semantic Analysis*.

Liu, M., & Yang, J. (2012). *An improvement of TFIDF weighting in text categorization*.

Minister of Industry Canada. (2003). *Survey Methods and Practices*. Ottawa, Canada.

Nigam, K. (1999). *Text Classification from Labeled and Unlabeled Documents Using EM*. Boston: Kluwer Academic Publishers.

Pidong, W. (2011, June 1). *Indonesian stop words list*. Retrieved April 7, 2014, from WANG PIDONG'S HOMEPAGE: <http://wangpidong.blogspot.com/2011/06/indonesian-stop-words-list.html>

Saad, M. K. (2010). *The Impact of Text Preprocessing and Term Weighting on Arabic Text Classification*.

Savov, I. (2009). *Latent Dirichlet Allocation for scientific topic extraction*.

Sriram, B., & Fuhry, D. (2010). *Short Text Classification in Twitter to Improve Information Filtering*. Geneva, Switzerland: SIGIR.

Vembunaryanan, J. (2013, October 27). *Tf-Idf and Cosine similarity*. Retrieved May 7, 2014, from Seeking Wisdom: <http://janav.wordpress.com/2013/10/27/tf-idf-and-cosine-similarity/>

Wang, L., & Wan, Y. (2011). Sentiment Classification of Document Based on Latent Sematic Analysis. *Springer-Verlag Berlin Heidelberg*.

Wijewickrema, C. M., & Gamage, R. (2013). An Ontology Based Fully Automatic Document Classification System Using an Existing Semi-Automatic System. *IFLA WLIC*.

Xia, T., & Chai, Y. (2011). An Improvement to TF-IDF: Term Distribution based Term Weight Algorithm. *Journal of Software*, 413.

