

## DAFTAR PUSTAKA

- [1] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a Social Network or a News Media?," in *International World Wide Web Conference*, 2010, pp. 591–600.
- [2] X. Wang, D. E. Brown, and M. S. Gerber, "Spatio-Temporal Modeling of Criminal Incidents Using Geographic, Demographic, and Twitter-derived Information," in *ISI 2012 - 2012 IEEE International Conference on Intelligence and Security Informatics: Cyberspace, Border, and Immigration Securities*, 2012, pp. 36–41.
- [3] K. D. Bollacker, S. Lawrence, and C. L. Giles, "CiteSeer: An autonomous Web agent for automatic retrieval and identification of interesting publications," *Proc. Int. Conf. Auton. Agents*, no. January 1998, pp. 116–123, 1998.
- [4] J. Beel, B. Gipp, S. Langer, and C. Breitinger, "Research Paper Recommender Systems : A Literature Survey," *Int. J. Digit. Libr.*, vol. 17, pp. 305–338, 2016.
- [5] S. E. Robertson, S. Walker, S. Jones, M. M. Hancock-Beaulieu, and M. Gatford, "Okapi at TREC-3," *Proc. Third Text Retr. Conf. (TREC 1994)*, 1994.
- [6] S. Robertson and H. Zaragoza, "The probabilistic relevance framework: BM25 and beyond," *Found. Trends Inf. Retr.*, vol. 3, no. 4, pp. 333–389, 2009.
- [7] A. I. Kadhim, "Term Weighting for Feature Extraction on Twitter: A Comparison between BM25 and TF-IDF," *2019 Int. Conf. Adv. Sci. Eng. ICOASE 2019*, pp. 124–128, 2019.
- [8] R. Ghawi and J. Pfeffer, "Efficient Hyperparameter Tuning with Grid Search for Text Categorization using kNN Approach with BM25 Similarity," *Open Comput. Sci.*, vol. 9, no. 1, pp. 160–180, 2019.
- [9] A. M. Sari, "Tuning Parameter Dalam Bm25 Similarity Pada Temu Kembali Penyakit Akasia," *IPB Publ.*, 2017.
- [10] Z. Abbass, Z. Ali, M. Ali, B. Akbar, and A. Saleem, "A Framework to Predict Social Crime Through Twitter Tweets by Using Machine Learning," in *Proceedings - 14th IEEE International Conference on Semantic Computing, ICSC 2020*, 2020, pp. 363–368.
- [11] V. Mageshwari and I. Laurence Aroquiaraj, "An efficient feature extraction

- method for mining social media,” *Int. J. Sci. Technol. Res.*, vol. 8, no. 11, pp. 640–643, 2019.
- [12] A. S. Zakia, Indriati, and Marji, “Klasifikasi Jenis Kelamin Pengguna Twitter dengan menggunakan Metode BM25 dan K-Nearest Neighbor ( KNN ),” *Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 10, pp. 3331–3337, 2020.
  - [13] B. Gain, D. Bandyopadhyay, T. Saikh, and A. Ekbal, “IITP@COLIEE 2019: Legal Information Retrieval using BM25 and BERT,” in *Competition on Legal Information Extraction/Entailment (COLIEE 2019)*, 2019.
  - [14] J. Pardede, M. Gustiana Husada, and R. Riansyah, “Implementasi Dan Perbandingan Metode Okapi BM25 Dan PLSA Pada Aplikasi Information Retrieval,” *Itenas Repos.*, pp. 1–10, 2018.
  - [15] G. A. Tinega, P. W. Mwangi, and D. R. Rimiru, “Text Mining in Digital Libraries using OKAPI BM25 Model,” *Int. J. Comput. Appl. Technol. Res.*, vol. 7, no. 10, pp. 398–406, 2018.
  - [16] N. Gayatri Indah Reza *et al.*, “DBSCAN Algorithm: Twitter Text Clustering of Trend Topic Pilkada Pekanbaru,” *Soc. Technol. J. Phys. Conf. Ser.*, vol. 1363, p. 12001, 2019.
  - [17] S. Sun and K. Duh, “CLIRMatrix: A massively large collection of bilingual and multilingual datasets for Cross-Lingual Information Retrieval,” in *2020 Conference on Empirical Methods in Natural Language Processing*, 2020, pp. 4160–4170.
  - [18] N. Khamis and C. S. Tan, “Time of use Period Determination for Residential Customers in Peninsular Malaysia,” *Int. J. Eng. Adv. Technol.*, vol. 9, no. 2, pp. 4246–4249, 2019.
  - [19] Y.-S. Li and M.-L. Qi, “An Approach for Understanding Offender Modus Operandi to Detect Serial Robbery Crimes,” *J. Comput. Sci.*, vol. 36, p. 101024, Sep. 2019.
  - [20] O. Vaidya, S. Mitra, R. Kumbhar, S. Chavan, and R. Patil, “Comprehensive Comparative Analysis of Methods for Crime Rate Prediction,” *Int. Res. J. Eng. Technol.*, vol. 05, no. 02, pp. 715–718, 2018.
  - [21] Z. Hussain, “Criminal Behavior Analysis by Using Data Mining Techniques,” in *International Conference On Advances In Engineering, Science And Management*, 2012, pp. 656–658.
  - [22] X. Wang, D. E. Brown, and J. H. Conklin, “Crime Incident Association with Consideration of Narrative Information,” in *Systems & Information Engineering Design Symposium*, 2007.

- [23] G. Zhao, Y. Liu, W. Zhang, and Y. Wang, "TFIDF based Feature Words Extraction and Topic Modeling for Short Text," in *2nd International Conference on Management Engineering, Software Engineering and Service Sciences, Wuhan, China*, 2018.
- [24] H. Al-Saif and H. Al-Dossari, "Detecting and Classifying Crimes from Arabic Twitter Posts using Text Mining Techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 10, 2018.
- [25] D. S. A. Maylawati, M. A. Ramdhani, A. Rahman, and W. Darmalaksana, "Incremental Technique with Set of Frequent Word Item Sets for Mining Large Indonesian Text Data," in *2017 5th International Conference on Cyber and IT Service Management, CITSM 2017*, 2017, pp. 1–6.
- [26] W. T. H. Putri and R. Hendrowati, "Penggalian Teks dengan Model Bag of Words Terhadap Data Twitter," *J. Muara Sains, Teknol. Kedokteran, dan Ilmu Kesehat.*, vol. 2, no. 1, pp. 129–138, 2018.
- [27] O. Somantri, "Text Mining Untuk Klasifikasi Kategori Cerita Pendek Menggunakan Naïve Bayes (NB)," *J. Telemat.*, vol. 12(1), 2017.
- [28] R. T. Adek and S. Nasution, "Tweet Clustering in Indonesian Language Twitter Social Media using Naive Bayes Classifier Method," *Eurasian J. Anal. Chem.*, vol. 6, pp. 277–284, 2018.
- [29] G. Attigeri, M. M. Manohara Pai, R. M. Pai, and R. Kulkarni, "Knowledge Base Ontology Building for Fraud Detection Using Topic Modeling," in *Procedia Computer Science*, 2018, vol. 135, pp. 369–376.
- [30] K. H. dan H. A. Manusia, "Buku II KUHP Tindak Pidana," *Kementeri. Huk. Dan Hak Asasi Mns.*, pp. 50–186, 2005.
- [31] M. Button, C. Lewis, and J. Tapley, "Fraud typologies and victims of fraud Literature review," *Natl. Fraud Auth.*, 2009.
- [32] S. Abiteboul, P. Buneman, and D. Suciu, "Data on the Web: From Relations to Semistructured Data and XML," *Morgan Kaufmann; 1st Ed.*, pp. 1–245, 2014.
- [33] R. Sint, S. Schaffert, S. Stroka, and R. Ferstl, "Combining unstructured, fully structured and semi-structured information in semantic wikis," *CEUR Workshop Proc.*, vol. 464, no. January 2009, pp. 73–87, 2009.
- [34] R. Blumberg and S. Atre, "The Problem with Unstructured Data," in *DM Review*, 2003, vol. 13, pp. 42–46.
- [35] R. Feldman and I. Dagan, "Knowledge Discovery in Textual Databases (KDT)," *Int. Conf. Knowl. Discov. Data Min.*, pp. 112–117, 1995.

- [36] R. Feldman and J. Sanger, “The Text Mining Handbook,” *Cambridge Univ. Press*, 2007.
- [37] J. Han, M. Kamber, and J. Pei, “Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems),” *Morgan Kaufmann Publ.*, 2011.
- [38] V. Gupta and G. S. Lehal, “A survey of text mining techniques and applications,” *J. Emerg. Technol. Web Intell.*, vol. 1, no. 1, pp. 60–76, 2009.
- [39] V. Bijalwan, V. Kumar, P. Kumari, and J. Pascual, “KNN based machine learning approach for text and document mining,” *Int. J. Database Theory Appl.*, vol. 7, no. 1, pp. 61–70, 2014.
- [40] M. W. Berry and M. Castellanos, “Survey of Text Mining II: Clustering, Classification, and Retrieval,” *Springer Nat.*, pp. 1–240, 2008.
- [41] R. Elmasri and S. B. Navathe, “Fundamentals of Database Systems,” *New York Addison-Wesley*, p. 893, 2000.
- [42] S. Dang and P. H. Ahmad, “Text Mining : Techniques and its Application,” *IJETI Int. J. Eng. Technol. Innov.*, vol. 1, no. December 2014, pp. 22–25, 2014.
- [43] M. Hofmann, A. Chisholm, and H. Chisholm, “Text Mining and Visualization Case Studies Using Open-Source Tools,” *Chapman & Hall/CRC*, p. 290, 2015.
- [44] I. Feinerer, K. Hornik, and D. Meyer, “Text mining infrastructure in R,” *J. Stat. Softw.*, vol. 25, no. 5, pp. 1–54, 2008.
- [45] M. W. Berry and M. Castellanos, *Survey of Text Mining II Clustering, Classification, and Retrieval*. 2008.
- [46] M. Allahyari *et al.*, “A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques,” *KDD Bigdas*, 2017.
- [47] B. He and I. Ounis, “Term frequency normalisation tuning for BM25 and DFR models,” *Lect. Notes Comput. Sci.*, vol. 3408, no. 1, pp. 200–214, 2005.
- [48] S. Vijayarani, J. Ilamathi, and Nithya, “Preprocessing Techniques for Text Mining-An Overview Privacy Preserving Data Mining View project,” *Int. J. Comput. Sci. Commun. Networks*, vol. 5(1), pp. 7–16, 2015.
- [49] T. Y. Liu, “Learning to rank for Information Retrieval,” *Found. Trends Inf. Retr.*, vol. 3, no. 3, pp. 225–231, 2009.
- [50] A. H. Samudra, “Modus Operandi dan Problematika Penanggulangan Tindak

Pidana Penipuan Daring,” *Mimb. Huk.*, vol. 31, pp. 59–74, 2019.

- [51] S. Supardi, “Populasi dan Sampel Penelitian,” *Unisia*, vol. 13, no. 17, pp. 100–108, 1993.
- [52] Twitter, “Raw Data-Penipuan,” 2019. [Online]. Available: [https://twitter.com/search?q=penipuan+lang%3Aid+until%3A2019-10-31+since%3A2019-08-01+-filter%3Areplies&src=typed\\_query](https://twitter.com/search?q=penipuan+lang%3Aid+until%3A2019-10-31+since%3A2019-08-01+-filter%3Areplies&src=typed_query). [Accessed: 10-Feb-2020].
- [53] P. Monish, K. R. Ranjith, G. Varun, S. Sridhar, and B. Tech, “Crime Analysis: Probabilistic Prediction,” *Int. Res. J. Eng. Technol.*, vol. 04, no. 04, pp. 3163–3166, 2017.
- [54] I. Adiwijaya, “Text Mining dan Knowledge Discovery,” in *Kolokium bersama komunitas datamining Indonesia & soft-computing Indonesia*, 2006.
- [55] J. Asian, H. E. Williams, and S. M. M. Tahaghoghi, “Stemming Indonesian,” *Conf. Res. Pract. Inf. Technol. Ser.*, vol. 38, no. January, pp. 307–314, 2005.
- [56] L. Syarifah, “Cara Cleaning Data Text Menggunakan Package tm di Rstudio,” 2019. [Online]. Available: <https://lailatussyarifah32.wordpress.com/2019/02/19/menguak-cara-explorasi-text-text-mining-dengan-package-tm-di-rstudi-statistics02/comment-page-1/?unapproved=136&moderation-hash=2cdbe5a90d502004d6872835dd3806#comment-136>. [Accessed: 15-Mar-2020].
- [57] N. Salsabila, “Kamus Alay - Colloquial Indonesian Lexicon,” 2018. [Online]. Available: <https://github.com/nasalsabila/kamus-alay>. [Accessed: 13-Apr-2020].