DAFTAR PUSTAKA

Abayomi-Alli, O., Misra, S., Abayomi-Alli, A., & Odusami, M. (2019). A review of soft techniques for SMS spam classification: Methods, approaches and applications. *Engineering Applications of Artificial Intelligence*, *86*, 197–212. https://doi.org/10.1016/j.engappai.2019.08.024

Abdulhamid, S. M., Abd Latiff, M. S., Chiroma, H., Osho, O., Abdul-Salaam, G., Abubakar, A. I., & Herawan, T. (2017). A review on mobile SMS spam filtering techniques. *IEEE Access*, *5*, 15650–15666. https://doi.org/10.1109/access.2017.2666785

Akbani, R., Kwek, S., & Japkowicz, N. (2004). Applying support vector machines to imbalanced datasets. In *Machine Learning: ECML 2004* (pp. 39–50). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-540-30115-8_7

Alghamdi, N., & Assiri, F. (2019). A Comparison of fastText Implementations Using Arabic Text Classification. In *Advances in Intelligent Systems and Computing* (pp. 306–311). Springer International Publishing. http://dx.doi.org/10.1007/978-3-030-29513-4_21

Almeida, T. A., Hidalgo, J. M. G., & Yamakami, A. (2011). Contributions to the study of SMS spam filtering. *Proceedings of the 11th ACM Symposium on Document Engineering - DocEng '11*. http://dx.doi.org/10.1145/2034691.2034742

Amalia, A., Sitompul, O. S., Nababan, E. B., & Mantoro, T. (2020, July). An Efficient Text Classification Using fastText for Bahasa Indonesia Documents Classification. *2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA)*. http://dx.doi.org/10.1109/databia50434.2020.9190447

Amir Sjarif, N. N., Mohd Azmi, N. F., Chuprat, S., Sarkan, H. M., Yahya, Y., & Sam, S. M. (2019). SMS Spam Message Detection using Term Frequency-Inverse Document Frequency and Random Forest Algorithm. *Procedia Computer Science*, *161*, 509–515. https://doi.org/10.1016/j.procs.2019.11.150

Anindya. (2020, October 7). FastText Word Embeddings Python implementation. ThinkInfi. https://thinkinfi.com/fasttext-word-embeddings-python-implementation/

Anna Octaviani, P., Wilandari, Y., & Ispriyanti, D. (2014). Penerapan metode klasifikasi support vector machine (svm) pada data akreditasi sekolah dasar (sd) di kabupaten magelang. *Jurnal Gaussian*, *3*(4).

Batuwita, R., & Palade, V. (2013). Class imbalance learning methods for support vector machines. In *Imbalanced Learning* (pp. 83–99). John Wiley & Sons, Inc. http://dx.doi.org/10.1002/9781118646106.ch5

Bhavsar, H., & Ganatra, A. (2012). A Comparative Study of Training Algorithms for

Supervised Machine Learning. *International Journal of Soft Computing and Engineering (IJSCE)*, *2*(4).

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146. https://doi.org/10.1162/tacl_a_00051

Bosaeed, S., Katib, I., & Mehmood, R. (2020, April). A Fog-Augmented Machine Learning based SMS Spam Detection and Classification System. *2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC)*. http://dx.doi.org/10.1109/fmec49853.2020.9144833

Budianto, A., Ariyuana, R., & Maryono, D. (2019). Perbandingan K-Nearest Neighbor (Knn) Dan Support Vector Machine (Svm) Dalam Pengenalan Karakter Plat Kendaraan Bermotor. *Jurnal Ilmiah Pendidikan Teknik Dan Kejuruan*, *11*(1), 27. https://doi.org/10.20961/jiptek.v11i1.18018

Cesconi, F. (2018, April 18). What is the main difference between word2vec and FastText? *Medium*. https://cesconi.com/what-is-the-main-difference-between-word2vec-and-fasttext-57 bdaf3a69ef

Chaudhary, A. (2020, June 21). *Amit Chaudhary*. Amit Chaudhary. https://amitness.com/2020/06/fasttext-embeddings/

Chen, Y. (2020). Mining of instant messaging data in the Internet of Things based on support vector machine. *Computer Communications*, *154*, 278–287. https://doi.org/10.1016/j.comcom.2020.02.080

Chia, D. (2018, December 6). An implementation guide to Word2Vec using NumPy and Google Sheets. *Derek Chia*. https://derekchia.com/an-implementation-guide-to-word2vec-using-numpy-and-go ogle-sheets/

Choi, J., & Lee, S.-W. (2020). Improving FastText with inverse document frequency of subwords. *Pattern Recognition Letters*, *133*, 165–172. https://doi.org/10.1016/j.patrec.2020.03.003

Dwiharyono, H., & Suyanto, S. (2022). Stemming for better indonesian text-to-phoneme. *Ampersand*, *9*, 100083. https://doi.org/10.1016/j.amper.2022.100083

Hasanah, N. A., Suciati, N., & Purwitasari, D. (2021). Identifying Degree-of-Concern on COVID-19 topics with text classification of Twitters. Register: Jurnal Ilmiah Teknologi Sistem Informasi, 7(1), 50. https://doi.org/10.26594/register.v7i1.2234

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2016). A Practical Guide to Support Vector Classification. *Department of Computer Science*.

Irmanda, H. N., & Ria Astriratma. (2020). Klasifikasi jenis pantun dengan metode support vector machines (SVM). *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *4*(5), 915–922. https://doi.org/10.29207/resti.v4i5.2313

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2017). Bag of tricks for efficient text classification. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. http://dx.doi.org/10.18653/v1/e17-2068

Karayiğit, H., İnan Acı, Ç., & Akdağlı, A. (2021). Detecting abusive Instagram comments in Turkish using convolutional Neural network and machine learning methods. *Expert Systems with Applications*, *174*, 114802. https://doi.org/10.1016/j.eswa.2021.114802

Kenter, T., Borisov, A., & de Rijke, M. (2016). Siamese CBOW: Optimizing word embeddings for sentence representations. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. http://dx.doi.org/10.18653/v1/p16-1089

Kowalczyk, A. (2017). *Support Vector Machines Succinctly*. Syncfusion. https://www.syncfusion.com/succinctly-free-ebooks/support-vector-machines-succinctly/introduction (Original work published 2017)

Kurnia, R. I., Tangkuman, Y. D., & Girsang, A. S. (2020). Classification of user comment using word2vec and SVM classifier. *International Journal of Advanced Trends in Computer Science and Engineering*, *9*(1), 643–648. https://doi.org/10.30534/ijatcse/2020/90912020

Kuyumcu, B., Aksakalli, C., & Delil, S. (2019, June 28). An automated new approach in fast text classification (fastText). *Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval*. http://dx.doi.org/10.1145/3342827.3342828

Li, Q., Peng, H., Li, J., Xia, C., Yang, R., Sun, L., Yu, P. S., & He, L. (2020). A Survey on Text Classification: From Traditional to Deep Learning. *ACM Transactions on Intelligent Systems and Technology*, *37*(4).

Liang, S. (2021). Comparative analysis of SVM, xgboost and neural network on hate speech classification. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(5), 896–903. https://doi.org/10.29207/resti.v5i5.3506

Miao, F., Zhang, P., Jin, L., & Wu, H. (2018, August). Chinese news text classification based on machine learning algorithm. *2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*. http://dx.doi.org/10.1109/ihmsc.2018.10117

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word*

Representations in Vector Space.

Moreno-Torres, J. G., Saez, J. A., & Herrera, F. (2012). Study on the impact of partition-induced dataset shift on k-fold cross-validation. *IEEE Transactions on Neural Networks and Learning Systems*, *23*(8), 1304–1312. https://doi.org/10.1109/tnnls.2012.2199516

Muhammad, P. F., Kusumaningrum, R., & Wibowo, A. (2021). Sentiment analysis using word2vec and long short-term memory (LSTM) for Indonesian hotel reviews. *Procedia Computer Science*, *179*, 728–735. https://doi.org/10.1016/j.procs.2021.01.061

Nawangsari, R. P., Kusumaningrum, R., & Wibowo, A. (2019). Word2Vec for Indonesian sentiment analysis towards hotel reviews: An evaluation study. *Procedia Computer Science*, *157*, 360–366. https://doi.org/10.1016/j.procs.2019.08.178

Pratama, A., Wihandika, R. C., & Ratnawati, D. E. (2018). Implementasi Algoritme Support Vector Machine (SVM) untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2*(4).

Pressman, R. (2010). *Software engineering: A practitioner's approach* (7th ed.). McGraw-Hill Education.

Rahmi, F., & Wibisono, Y. (2016). *Aplikasi Sms Spam Filtering Pada Android Menggunakan Algoritma Naïve Bayes*.

Raj, H., Weihong, Y., Banbhrani, S. K., & Dino, S. P. (2018). LSTM based short message service (SMS) modeling for spam classification. *Proceedings of the 2018 International Conference on Machine Learning Technologies - ICMLT '18*. http://dx.doi.org/10.1145/3231884.3231895

Reviantika, F., Azhar, Y., & Marthasari, G. I. (2021). Analisis Klasifikasi SMS Spam Menggunakan Logistic Regression. Jurnal Sistem Cerdas, 04(03), 155–160.

Rezalina, O. (2016). Perbandingan Algoritma Stemming Nazief & Adriani, Porter dan Arifin Setiono untuk Dokumen Teks Bahasa Indonesia. *Journal of Undergraduate Thesis, Universitas Muhammadiyah Jember*.

Rooy, N. (2018, March). *Word2vec from scratch with python and numpy*. Nathanrooy.Github.Io. https://nathanrooy.github.io/posts/2018-03-22/word2vec-from-scratch-with-python-and-numpy/

Sazany, E., & Budi, I. (2018, September). Deep learning-based implementation of hate speech identification on texts in indonesian: Preliminary study. *2018 International Conference on Applied Information Technology and Innovation (ICAITI)*. http://dx.doi.org/10.1109/icaiti.2018.8686725

Shafiabady, N., Lee, L. H., Rajkumar, R., Kallimani, V. P., Akram, N. A., & Isa, D. (2016). Using unsupervised clustering approach to train the Support Vector Machine for text classification. *Neurocomputing*, *211*, 4–10. https://doi.org/10.1016/j.neucom.2015.10.137

Sharaff, A. (2018). Spam detection in SMS based on feature selection techniques. In *Advances in Intelligent Systems and Computing* (pp. 555–563). Springer Singapore. http://dx.doi.org/10.1007/978-981-13-1498-8_49

Simarangkir, M. S. H. (2017). STUDI PERBANDINGAN ALGORITMA - ALGORITMA STEMMING UNTUK DOKUMEN TEKS BAHASA INDONESIA. *Jurnal Inkofar*, *1*(1). https://doi.org/10.46846/jurnalinkofar.v1i1.2

Sommerville, I. (2015). *Software Engineering, Global Edition*.

Subasi, A. (2013). Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. *Computers in Biology and Medicine*, *43*(5), 576–586. https://doi.org/10.1016/j.compbiomed.2013.01.020

Supriyadi, E., & Sensue, D. I. (2015). Optimasi algoritma support vector machine dengan particle swarm optimization dalam mendeteksi ketepatan waktu kelulusan mahasiswa : Studi kasus poltek lp3i jakarta "kampus depok." *Prosiding SNIT*.

Suyanto, S. (2018). *Machine Learning Tingkat Dasar dan Lanjut*. Penerbit Informatika.

Verma, T., Renu, R., & Gaur, D. (2014). Tokenization and filtering process in RapidMiner. *International Journal of Applied Information Systems*, *7*(2), 16–18. https://doi.org/10.5120/ijais14-451139

Vijayakumar, S., & Wu, S. (1999). Sequential support vector classifiers and regression. *Proc. International Conference on Soft Computing (SOCO'99), Genoa, Italy*.

Vishwakarma, A. K., Ansari, M. D., & Rai, G. (2020). SMS spam filtering using machine learning technique. In *Lecture Notes in Electrical Engineering* (pp. 689–701). Springer Singapore. http://dx.doi.org/10.1007/978-981-15-7961-5_66

Wahyudi, D., Susyanto, T., & Nugroho, D. (2017a). IMPLEMENTASI DAN ANALISIS ALGORITMA STEMMING NAZIEF & ADRIANI DAN PORTER PADA DOKUMEN BERBAHASA INDONESIA. *Jurnal Ilmiah SINUS*, *15*(2). https://doi.org/10.30646/sinus.v15i2.305

Wahyudi, D., Susyanto, T., & Nugroho, D. (2017b). Implementasi dan analisis algoritma stemming nazief & adriani dan porter pada dokumen berbahasa indonesia. *Jurnal Ilmiah SINUS*, *15*(2). https://doi.org/10.30646/sinus.v15i2.305

Wijayanti, R. A., Furqon, Muh. Tanzil, & Adinugroho, S. (2018). Penerapan Algoritme Support Vector Machine Terhadap Klasifikasi Tingkat Risiko Pasien Gagal Ginjal.

Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, *2*(10).

Xia, T., & Chen, X. (2021). A weighted feature enhanced Hidden Markov Model for spam SMS filtering. *Neurocomputing*, *444*, 48–58. https://doi.org/10.1016/j.neucom.2021.02.075

Young, J. C., & Rusli, A. (2019, August). Review and visualization of Facebook's fasttext pretrained word vector model. *2019 International Conference on Engineering, Science, and Industrial Applications (ICESI)*. http://dx.doi.org/10.1109/icesi.2019.8863015

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing [review article]. *IEEE Computational Intelligence Magazine*, *13*(3), 55–75. https://doi.org/10.1109/mci.2018.2840738

Yudi Widhiyasana, Transmissia Semiawan, Ilham Gibran Achmad Mudzakir, & Muhammad Randi Noor. (2021). Penerapan Convolutional Long Short-Term Memory untuk Klasifikasi Teks Berita Bahasa Indonesia. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, *10*(4), 354–361. https://doi.org/10.22146/jnteti.v10i4.2438