

REFERENCES

- Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv preprint arXiv:1707.02919*.
- Chalaris, M., Gritzalis, S., Maragoudakis, M., Sgouropoulou, C., & Tsolakidis, A. (2014). Improving quality of educational processes providing new knowledge using data mining techniques. *Procedia-Social and Behavioral Sciences, 147*, 390-397.
- Fatahillah, N. R., Suryati, P., & Haryawan, C. (2017, November). Implementation of Naive Bayes classifier algorithm on social media (Twitter) to the teaching of Indonesian hate speech. In *2017 International Conference on Sustainable Information Engineering and Technology (SIET)* (pp. 128-131). IEEE.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine, 17*(3), 37-37.
- Feldman, R., & Dagan, I. (1995, August). Knowledge Discovery in Textual Databases (KDT). In *KDD* (Vol. 95, pp. 112-117).
- Hevner, A., March, S. T., Park, J., & Ram, S. (2004). Design science research in information systems. *MIS quarterly, 28*(1), 75-105.
- Hotho, A., Nürnberger, A., & Paaß, G. (2005, May). A brief survey of text mining. In *Ldv Forum* (Vol. 20, No. 1, pp. 19-62).
- Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons, 54*(3), 241-251.

- Lee, K., Palsetia, D., Narayanan, R., Patwary, M. M. A., Agrawal, A., & Choudhary, A. (2011, December). Twitter trending topic classification. In *2011 IEEE 11th International Conference on Data Mining Workshops* (pp. 251-258). IEEE.
- Mahardhika, A. A., Saptono, R., & Anggrainingsih, R. (2015). Classification System for Customer Feedback and Recommended Solutions for Complaints at UPT Puskom UNS With the Naive Bayes Classifier Algorithm and Cosine Simiality. *ITSMART: Technology and Information Journal*, 4(1), 36-42.
- Mining, K. D. T. D. (1996). What Is Knowledge Discovery. *Tandem Computers Inc*, 253.
- Meyer, D., Hornik, K., & Feinerer, I. (2008). Text mining infrastructure in R. *Journal of statistical software*, 25(5), 1-54.
- Narkhede, S. (2018, May 9). *Understanding confusion matrix*. Towards Data Science. <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62?gi=628da9e73b23>
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *J. Inform*, 4(01).
- Obar, J. A., & Wildman, S. S. (2015). Social media definition and the governance challenge-an introduction to the special issue. *Obar, JA and Wildman, S.(2015). Social media definition and the governance challenge: An introduction to the special issue. Telecommunications policy*, 39(9), 745-750.
- Parlina, I., Arnol, M. Y., Febriati, N. A., Dewi, R., Wanto, A., & Lubis, M. R. (2019, August). Naive Bayes Algorithm Analysis to Determine the

Percentage Level of visitors the Most Dominant Zoo Visit by Age Category. In *Journal of Physics: Conference Series* (Vol. 1255, No. 1, p. 012031). IOP Publishing.

Pratama, B. Y., & Sarno, R. (2015, November). Personality classification based on Twitter text using Naive Bayes, KNN and SVM. In *2015 International Conference on Data and Software Engineering (ICoDSE)* (pp. 170-174). IEEE.

Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.

Ridwan, M., Suyono, H., & Sarosa, M. (2013). Application of Data Mining for Student Academic Performance Evaluation Using the Naive Bayes Classifier Algorithm. *journal EECCIS*, 7(1), 59-64.

Ruuska, S., Hämäläinen, W., Kajava, S., Mughal, M., Matilainen, P., & Mononen, J. (2018). Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behavioural processes*, 148, 56-62.

Qaiser, S., & Ali, R. (2018). Text mining: use of TF-IDF to examine the relevance of words to documents. *International Journal of Computer Applications*, 181(1), 25-29.

Satvika, G. A. J., Nasution, S. M., & Nugrahaeni, R. A. (2018, October). Determination of the Best Vehicle Pathway with Classification of Data Mining Twitter using K-Nearest Neighbor. In *2018 International Conference on Information Technology Systems and Innovation (ICITSI)* (pp. 72-76). IEEE.

Ray, S. (2017, September 9). *Commonly used machine learning algorithms (with python and r codes)*. Analytics Vidhya.

<https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>

Talib, R., Hanif, M. K., Ayesha, S., & Fatima, F. (2016). Text mining: techniques, applications and issues. *International Journal of Advanced Computer Science and Applications*, 7(11), 414-418.

Tseng, C., Patel, N., Paranjape, H., Lin, T. Y., & Teoh, S. (2012, August). Classifying twitter data with naive bayes classifier. In *2012 IEEE International Conference on Granular Computing* (pp. 294-299). IEEE.

Zulfikar, W. B., Irfan, M., Alam, C. N., & Indra, M. (2017, August). The comparison of text mining with Naive Bayes classifier, nearest neighbor, and decision tree to detect Indonesian swear words on Twitter. In *2017 5th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-5). IEEE.

Team, R. C. (2000). R language definition. *Vienna, Austria: R foundation for statistical computing*.