

REFERENCES

- Ahire, S. (2014), A survey of sentiment lexicons, *Google Scholar*.
- Arunava (2018), *IMDB Movie Reviews Dataset*.
- Bakken, P. F., Bratlie, T. A., Marco, C., dan Gulla, J. A. (2016), Political News Sentiment Analysis for Under-resourced Languages, *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pp. 2989–2996.
- Basu, S., Banerjee, A., dan Mooney, R. (2002), Semi-supervised clustering by seeding, *In Proceedings of 19th International Conference on Machine Learning (ICML- 2002*, Citeseer.
- Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., dan Varoquaux, G. (2013), API design for machine learning software: experiences from the scikit-learn project, *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pp. 108–122
- Goldberg, Y. dan Hirst, G., 2017, *Neural Network Methods in Natural Language Processing*, Morgan & Claypool Publishers.
- He, Y. and Zhou, D., 2011. Self-training from labeled features for sentiment analysis. *Information Processing & Management*, 47(4), pp.606-616.
- Kodžoman, 2020. *Pseudo-Labeling A Simple Semi-Supervised Learning Method - Data, What Now?*. [online] Data, what now?. Available at: <<https://datawhatnow.com/pseudo-labeling-semi-supervised-learning/>> [Accessed 25 May 2020].

- Liu, B., 2015, *Opinions, Sentiment, and Emotion in Text*, Cambridge University Press. Liu, B., 2015, *Sentiment analysis: Mining opinions, sentiments, and emotions*, Cambridge University Press.
- Low, J.S., Ghafoori, Z., Bezdek, J.C. and Leckie, C., 2019, August. Seeding on samples for accelerating k-means clustering. In *Proceedings of the 3rd International Conference on Big Data and Internet of Things* (pp. 41-45).
- Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., dan Potts, C. (June 2011), Learning Word Vectors for Sentiment Analysis, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Portland, Oregon, USA: Association for Computational Linguistics, pp. 142–150.
- Marr, B. (2019), *How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read*.
- Nigam, K., McCallum, A., dan Mitchell, T. (2006), Semi-supervised text classification using EM, *Semi-Supervised Learning*, pp. 33–56. 20
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cour-napeau, D., Brucher, M., Perrot, M., dan Duchesnay, E. (2011), Scikit-learn: Ma-chine Learning in Python, *Journal of Machine Learning Research* 12, pp. 2825–2830.
- Qian, L. and Zhou, G., 2010, October. Clustering-based stratified seed sampling for semi-supervised relation classification. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing* (pp. 346-355). Association for Computational Linguistics.

Triguero, I., García, S., dan Herrera, F. (2015), Self-labeled techniques for semi-supervised learning: taxonomy, software and empirical study, *Knowledge and In-formation systems* 42.2, pp. 245–284.

Wicaksono, A. F., Vania, C., Distiawan, B., dan Adriani, M. (2014), Automatically building a corpus for sentiment analysis on Indonesian tweets, *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computing*, pp. 185– 194.

Wu, J., Li, L. and Wang, W.Y., 2018. Reinforced co-training. *arXiv preprint arXiv:1804.06035*.