DAFTAR PUSTAKA

Agrafiotis, I. et al., 2018. A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate. *Journal of Cybersecurity,* pp. 1-15.

Agustini, P., 2021. *Warganet Meningkat, Indonesia Perlu Tingkatkan Nilai Budaya di Internet.* [Online]
Available at: https://aptika.kominfo.go.id/2021/09/warganet-meningkat-indonesia-perlu-tingkatkan-nilai-budaya-di-internet/
[Accessed 15 October 2021].

Amini, A., Wah, T. . Y. & Saboohi, H., 2014. On Density-Based Data Streams Clustering Algorithms: A Survey. *Journal of Computer Science and Technology,* 29(1), pp. 116-141.

anaconda, 2017. *Anaconda Distribution Starter Guide.* [Online]
Available at: https://docs.anaconda.com/anaconda/user-guide/getting-started/
[Accessed 20 10 2021].

Anjani, N. H., 2021. *Cybersecurity Protection in Indonesia,* Jakarta: Center for Indonesian Policy Studies.

Bajtoˇs, T. et al., 2019. Analysis of the Infection and the Injection Phases of the Telnet Botnets. *Journal of Universal Computer Science,* 25(11), pp. 1417-1436.

Baykara, M. & Das, R., 2018. A novel honeypot based security approach for real-time intrusion detection and prevention system. *Journal of Information Security and Applications,* Volume 41, p. 105.

Butwall, M., Ranka, P. & Shah, S., 2019. Python in Field of Data Science: A Review. *International Journal of Computer Applications,* 178(49), pp. 20-24.

Chen, R.-C., Dewi, C., Huang, S.-w. & Caraka, R. E., 2020. Selecting critical features for data classification based on machine learning methods. *Journal of Big Data,* Volume 52, p. 7.

Collins, M., 2017. *Network Security Through Data Analysis.* 2nd ed. Sebastopol: O'Reilly Media, Inc..

Delen, D. & Demirkan, H., 2013. Data, information and analytics as services. *Decision Support Systems,* 55(1), pp. 359-363.

dionaea, 2020. *dionaea Documentation.* [Online]
Available at: https://media.readthedocs.org/pdf/dionaea/latest/dionaea.pdf
[Accessed 22 10 2021].

Enterprise, J., 2019. *Python untuk Programmer Pemula.* 1 ed. s.l.:Elex Media Komputindo.

Farjad, S. M. & Arfeen, A., 2020. *Cluster Analysis and Statistical Modeling: A Unified Approach for Packet Inspection.* Islamabad, 2020 International Conference on Cyber Warfare and Security (ICCWS), pp. 1-7.

Grout, J., Frederic, J., Corlay, S. & Bugnion, P., 2016. *ipwidgets: Interactive HTML Widgets.* [Online]
Available at: https://github.com/jupyter-widgets/ipywidgets
[Accessed 23 September 2021].

Hermawan, D. S., S. & Risqiwati, D., 2020. Analisa Real-Time Data log honeypot menggunakan Algoritma K-Means pada serangan Distributed Denial of Service. *Jurnal Repositor,* 2(5), p. 541.

Heydt, M., 2017. *Learning pandas.* 2 ed. s.l.:Packt Publishing.

Huang, Z., 1998. Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values. *Data Mining and Knowledge Discovery,* Volume 2, pp. 283-304.

Jang-Jaccard, J. & Nepal, S., 2014. A survey of emerging threats in cybersecurity. *Journal of Computer and System Sciences,* Volume 80, pp. 973-993.

Joshi, R., 2016. *Exsilio Solutions.* [Online]
Available at: https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/
[Accessed 25 10 2021].

Khandaker, M. S., Hussain, A. & Ahmed, M., 2018. Effectiveness of Hard Clustering Algorithms for Securing Cyber Space. *International Conference on Smart Grid and Internet of Things,* Volume 256, pp. 113-120.

Kox, H. L., 2013. Cybersecurity in the Perspective of Internet Traffic Growth. *MPRA Paper,* pp. 1-16.

Kulkarni, A., Chong, D. & Batarseh, F. A., 2020. 5 - Foundations of data imbalance and solutions for a data democracy. In: F. A. Batarseh & R. Yang, eds. *Data Democracy: Data Democracy: At the Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering.* s.l.:Academic Press, pp. 83-106.

Lehrer, C. et al., 2018. How Big Data Analytics Enables Service Innovation: Materiality, Affordance, and the Individualization of Service. *Journal of Management Information Systems,* Volume 35, pp. 424-460.

Lepenioti, K., Bousdekis, A., Apostolou, D. & Mentzas, G., 2020. Prescriptive analytics: literature review and research challenges. *International Journal of Information Management,* Volume 50, pp. 57-70.

Lester, J. N., Cho, Y. & Lochmiller, C. R., 2020. Learning to Do Qualitative Data Analysis: A Starting Point. *Human Resource Development Review,* 19(1), pp. 94-106.

Lu, X., Zhang, H. & Huang, K., 2020. Improving Investment Return Through Analyzing and Mining Sales Data. In: J. He, et al. eds. *Data Science.* Ningbo, China: Springer Nature Singapore Pte Ltd., pp. 112-118.

Mertz, D., 2021. *Cleaning Data foEffective Data Science: Doing the other 80% of the work with Python, R, and command-line tools.* s.l.:Packt Publishing.

Miranda-Calle, D. J., G., R. V., Dhawan, P. & Churi, P., 2021. Exploratory data analysis for cybersecurity. *World Journal of Engineering,* 18(5), pp. 734-749.

Molin, S., 2021. *Hands-On Data Analysis with Pandas.* 2 ed. Birmingham: Packt Publishing Ltd..

Moore, C. & Al-Nemrat, A., 2015. *An Analysis of Honeypot Programs amd the Attack Data Collected.* s.l., Springer International Publishing, p. 229.

Nelli, F., 2018. *Python Data Analytics.* 2 ed. s.l.:Apress.

Olding, P. B. & Wolfe, P., 2014. Inference for Graph and Networks Adapting Classical Tools to Modern Data. In: N. Adams & N. Heard, eds. *Data Analysis for Network Cyber-Security.* London: Imperial College Press, pp. 1-30.

Popat, S. K. & M., E., 2014. Review and Comparative Study of Clustering Techniques. *International Journal of Computer Science and Information Technologies,* 5(1), pp. 805-812.

Portalatin, M. et al., 2021. *Data Analytics for Cyber Risk Analysis Utilizing Cyber Incident Datasets.* Charlottesville, 2021 Systems and Information Engineering Design Symposium (SIEDS), pp. 1-6.

Randles, B. M., Pasquetto, I. V., Golshan, M. S. & Borgman, C. L., 2017. Using the Jupyter Notebook as a Tool for Open Science: An Empirical Study. *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL),* pp. 1-2.

Reyes, P. C., 2019. *Statistical learning with high-cardinality string categorical variables,* Palaiseau: Université Paris-Saclay.

Saikawa, K. & Klyuev, V., 2019. Detection and Classification of Malicious Access using a Dionaea Honeypot. *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS),* p. 845.

Santos, J. M. & Embrechts, M., 2009. *On the Use of the Adjusted Rand Index as a Metric for Evaluating Supervised Classification.* Berlin, Springer, Berlin, Heidelberg.

Sarker, I. H. et al., 2020. Cybersecurity data science: an overview from machine learning perspective. *Journal of Big Data,* 7(41), pp. 1-29.

Scitovski, R., Sabo, K., Martínez-Álvarez, F. & Ungar, Š., 2021. *Cluster Analysis and Applications.* Switzerland: Springer Nature Switzerland AG.

Sethia, V. & Jeyasekar, A., 2019. Malware Capturing and Analysis using Dionaea Honeypot. *2019 International Carnahan Conference on Security Technology (ICCST),* p. 3.

Shavitt, Y., 2011. A Geolocation Databases Study. *IEEE Journal on Selected Areas in Communications,* 29(10), pp. 2044-2057.

Stančin & Jović, A., 2019. An overview and comparison of free Python libraries for data mining and big data analysis. *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO),* p. 977.

Sumana, B. V. & Santhanam, T., 2016. Prediction of Imbalanced Data Using Cluster Based Approach. *Asian Journal of Information Technology,* 15(16), pp. 3022-3042.

Tartakovsky, G. A., 2014. Rapid Detection of Attacks in Computer Networks by Quickest Changepoint Detection Methods. In: N. Adams & N. Heard, eds. *Data Analysis for Network Cyber-Security.* London: Imperial College Press, pp. 33-34.

Toomey, D., 2017. *Jupyter for Data Science: Exploratory analysis, statistical modeling, machine learning, and data visualization with Jupyter.* s.l.:Packt Publishing.

Umargono, E., Suseno, E. J. & K, V. G. S., 2020. *K-Means Clustering Optimization using the Elbow Method and Early Centroid Determination Based-on Mean and Median.* s.l., Proceedings of the International Conferences on Information System and Technology (CONRIST 2019), pp. 234-240.

VanderPlas, J., 2016. *Python Data Science Handbook: Essential Tools for Working with Data.* 1 ed. s.l.:O'Reilly Media.

Wang, L., Jones & Randy, 2020. Big Data Analytics in Cyber Security: Network Traffic and Attacks. *Journal of Computer Information System,* pp. 1-8.

Waskom, M. L., 2021. seaborn: statistical data visualization. *Journal of Open Source Software,* 6(60), p. 3021.

Wes, M., 2017. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython.* 2 ed. s.l.:O'Reilly Media.

Wibawa, G. H. P., Sasmita, I. G. M. A. & Raharja, I. M. S., 2020. Analisis Data Log Honeypot Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Merpati,* 8(1).

Wulandari, W. A., 2018. Analisis Network Forensics menggunakan Honeypoy pada Jaringan Layanan Public Cloud Computing. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM),* 3(1), pp. 18-25.

Yasmin, M. S., 2020. *Exploratory Data Analysis of Phising Sites to Identify Most Important Features to Detect a Phising Site,* Dhaka: Computer Science and Engineering Daffodil Internation University.

Yugitama, R., Kartika Rachman, P. P. & S., 2020. Efisiensi Monitoring Honeypot Dengan Menggunakan Visualisasi Dan Otomatisasi Laporan Log Serangan. *Jurnal IT Media IT SMIK Handayani,* 10(3), pp. 245-252.