

## REFERENCES

- Ahsan, M. M., Raman, S., & Siddique, Z. (2023). *BSGAN: A Novel Oversampling Technique for Imbalanced Pattern Recognitions*.
- Alom, M. Z., Bontupalli, V., & Taha, T. M. (2015). *Intrusion Detection Using Deep Belief Networks*. In Proceedings of the 2015 National Aerospace and Electronics Conference (NAECON), Dayton, OH, USA (pp. 339-344). doi:10.1109/NAECON.2015.7443094.
- Arjovsky, M., Chintala, S., & Bottou, L. (2017). *Wasserstein gan*. ArXiv, abs/1701.07875.
- BBC News. (2018). *Cambridge Analytica: The story so far*. Retrieved from <https://www.bbc.com/news/technology-43465968>.
- Belarbi, O., Khan, A., Carnelli, P., & Spyridopoulos, T. (2022). *An Intrusion Detection System based on Deep Belief Networks*. arXiv preprint arXiv:2207.02117. Presented at the 4th International Conference on Science of Cyber Security (SciSec).
- Bourou, S., El Saer, A., Velivassaki, T.-H., Voulkidis, A., & Zahariadis, T. (2021). *A Review of Tabular Data Synthesis Using GANs on an IDS Dataset*. Information, 12, 375. doi:10.3390/info12090375.
- Chao, W., Wenhui, W., Jiahan, D., & Guangxin, G. (2021). *Research on Network Intrusion Detection Technology Based on DCGAN*. Proceedings of the 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China (pp. 1418-1422). IEEE. doi: 10.1109/IAEAC50856.2021.9390891.
- Chu, Y., Zhao, X., Zou, Y., Xu, W., Han, J., & Zhao, Y. (2018). *A Decoding Scheme for Incomplete Motor Imagery EEG With Deep Belief Network*. In Frontiers in Neuroscience, 12. doi: 10.3389/fnins.2018.00680.
- DominoAI. (2023). *Model Evaluation*. Retrieved from <https://domino.ai/data-science-dictionary/model-evaluation>
- Fan, C., Chen, M., Wang, J., & Wang, X. (2021). *A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data*. Frontiers in Energy Research, 9. <https://doi.org/10.3389/fenrg.2021.652801>
- Goodfellow, I., Bengio, Y., Courville, A. (2016). *Deep Learning*. Massachusetts: The MIT Press, 528-566.
- Han, H., Wang, W.-Y., & Mao, B.-H. (2005). *Borderline-SMOTE: A new oversampling method in imbalanced data sets learning*. In the International Conference on Intelligent Computing (pp. 878-887). Springer.

- Hinton, G. E., & Salakhutdinov, R. R. (2006). *Reducing the dimensionality of data with neural networks*. Science, 313(5786), 504–507.
- Hinton, G. E., Osindero, S., Teh, Y.-W. (2006). *A Fast Learning Algorithm for Deep Belief Nets*. In Neural Computation, 18(7), 1527–1554. doi: <https://doi.org/10.1162/neco.2006.18.7.1527>.
- Hamid, Y., Ranganathan, B., Journaux, L., & Muthukumarasamy, S. (2018). *Benchmark Datasets for Network Intrusion Detection: A Review*. International Journal of Network Security. Retrieved from [https://doi.org/10.6633/IJNS.2018xx.20\(x\).xx](https://doi.org/10.6633/IJNS.2018xx.20(x).xx)
- Lin, Z., Khetan, A., Fanti, G., & Oh, S. (2018). *PacGAN: The Power of Two Samples in Generative Adversarial Networks*. In Advances in Neural Information Processing Systems.
- Maji, S. (2018). *Building an Intrusion Detection System on UNSW-NB15 Dataset Based on Machine Learning Algorithm*. Medium. Retrieved from <https://medium.com/@subrata.maji16/building-an-intrusion-detection-system-on-unsw-nb15-dataset-based-on-machine-learning-algorithm-16b1600996f5>. Accessed: 2023.
- Montúfar, G. (2018). *Restricted Boltzmann Machines: Introduction and Review*. In Information Geometry and Its Applications (pp. 1-1). Springer. [https://doi.org/10.1007/978-3-319-97798-0\\_4](https://doi.org/10.1007/978-3-319-97798-0_4).
- Padarian, J., & Fuentes, I. (2019). *Word embeddings for application in geosciences: development, evaluation, and examples of soil-related concepts*. SOIL, 5, 177-187. <https://doi.org/10.5194/soil-5-177-2019>.
- Radford, A., Metz, L., Chintala, S., et al. (2016). *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks*. In the International Conference on Learning Representations, 1-16.
- Ravinda, R. (2023). *GeeksforGeeks. AUC-ROC Curve*. Retrieved from <https://www.geeksforgeeks.org/auc-roc-curve/>
- Roy, R. (2023). *Generative Adversarial Network (GAN)*. Geeksforgeeks. Retrieved from <https://www.geeksforgeeks.org/generative-adversarial-network-gan/>
- Sethi, A. (2020). *One Hot Encoding vs. Label Encoding using Scikit-Learn*. <https://www.analyticsvidhya.com/blog/2020/03/one-hot-encoding-vs-label-encoding-using-scikit-learn/>
- TURING. (2023). *Synthetic Data Generation: Definition, Types, Techniques, and Tools*. Retrieved from <https://www.turing.com/kb/synthetic-data-generation-techniques#based-on-the-statistical-distribution>.

- UNSW. (2017). *UNSW-NB15 Dataset*. Retrieved from <https://research.unsw.edu.au/projects/unsw-nb15-dataset>
- Vitter, Jeffrey S. (1985). *Random Sampling with a Reservoir*. ACM Transactions on Mathematical Software, 11(1), 37–57. doi:10.1145/3147.3165.
- Wheelus, C., Bou-Harb, E., & Zhu, X. (2018). *Tackling Class Imbalance in Cyber Security Datasets*. In Proceedings of the 2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, USA (pp. 229-232). IEEE. doi: 10.1109/IRI.2018.00041.
- Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). *Modeling Tabular Data using Conditional GAN*. arXiv preprint arXiv:1907.00503.
- Yang, J., Li, T., Liang, G., He, W., & Zhao, Y. (2019). *A Simple Recurrent Unit Model based Intrusion Detection System with DCGAN*. In IEEE Access (pp. 1-1). doi: 10.1109/ACCESS.2019.2922692.
- Yasar, K., & Laskowski, N. (2023). *Synthetic Data Definition*. TechTarget. Retrieved from [https://www.techtarget.com/searchcio/definition/synthetic-data#:~:text=Synthetic%20data%20is%20information%20that's,machine%20learning%20\(ML\)%20models](https://www.techtarget.com/searchcio/definition/synthetic-data#:~:text=Synthetic%20data%20is%20information%20that's,machine%20learning%20(ML)%20models).
- Zhang, G., Rui, X., Poslad, S., Song, X., Fan, Y., & Wu, B. (2020). *A Method for the Estimation of Finely-Grained Temporal Spatial Human Population Density Distributions Based on Cell Phone Call Detail Records*. Remote Sensing, 12, 2572. doi:10.3390/rs12162572.