REFERENCES

[1] J. Singh and J. Singh, "A survey on machine learning-based malware detection in executable files," *Journal of Systems Architecture*, p. 101861, 2020.

[2] A. Kumar, K. Kuppusamy, and G. Aghila, "A learning model to detect maliciousness of portable executable using integrated feature set," *Journal of King Saud University-Computer and Information Sciences*, vol. 31, no. 2, pp. 252–265, 2019.

[3] B. Quinto, *Next-Generation Machine Learning with Spark: Covers XGBoost, LightGBM, Spark NLP, Distributed Deep Learning with Keras, and More*, 1st ed. Apress, 2020.

[4] A. Shalaginov, S. Banin, A. Dehghantanha, and K. Franke, "Machine learning aided static malware analysis: A survey and tutorial," in *Cyber Threat Intelligence*. Springer, 2018, pp. 7–45.

[5] L. Yang, A. Ciptadi, I. Laziuk, A. Ahmadzadeh, and G. Wang, "Bodmas: An open dataset for learning based temporal analysis of pe malware," in *4th Deep Learning and Security Workshop*, 2021.

[6] P. Agrawal and B. Trivedi, "Machine learning classifiers for android malware detection," *Data Management, Analytics and Innovation*, vol. 1, p. 311, 2020.

[7] S. Ardabili, A. Mosavi, and A. R. Várkonyi-Kóczy, "Advances in machine learning modeling reviewing hybrid and ensemble methods," in *International Conference on Global Research and Education*. Springer, 2019, pp. 215–227.

[8] H. S. Anderson and P. Roth, "EMBER: An Open Dataset for Training Static PE Malware Machine Learning Models," *ArXiv e-prints*, Apr. 2018.

[9] A. Deng, H. Zhang, W. Wang, J. Zhang, D. Fan, P. Chen, and B. Wang, "Developing computational model to predict protein-protein interaction sites based on the xgboost algorithm," *International journal of molecular sciences*, vol. 21, no. 7, p. 2274, 2020.

[10] L. Li, R. Situ, J. Gao, Z. Yang, and W. Liu, "A hybrid model combining convolutional neural network with xgboost for predicting social media popularity," in *Proceedings of the 25th ACM international conference on Multimedia*, 2017, pp. 1912–1917.

[11] M. Ohsaki, P. Wang, K. Matsuda, S. Katagiri, H. Watanabe, and A. Ralescu, "Confusion-matrix-based kernel logistic regression for imbalanced data classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 9, pp. 1806–1819, 2017.

[12] P. S. Rajawat, D. K. Gupta, S. S. Rathore, and A. Singh, "Predictive analysis of medical data using a hybrid machine learning technique," in *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*. IEEE, 2018, pp. 228–233.

[13] L. Binxiang, Z. Gang, and S. Ruoying, "A deep reinforcement learning malware detection method based on pe feature distribution," in *2019 6th International Conference on Information Science and Control Engineering (ICISCE)*. IEEE, 2019, pp. 23–27.

[14] J. Kang and Y. Won, "A study on variant malware detection techniques using static and dynamic features," *Journal of Information Processing Systems*, vol. 16, no. 4, pp. 882–895, 2020.

[15] A. Burkov, *The hundred-page machine learning book*, 2019, oCLC: 1089445188.

[16] P. Sornsuwit and S. Jaiyen, "A new hybrid machine learning for cybersecurity threat detection based on adaptive boosting," *Applied Artificial Intelligence*, vol. 33, no. 5, pp. 462–482, 2019.

[17] B. T. Pham, I. Prakash, S. K. Singh, A. Shirzadi, H. Shahabi, D. T. Bui *et al.*, "Landslide susceptibility modeling using reduced error pruning trees and different ensemble techniques: Hybrid machine learning approaches," *Catena*, vol. 175, pp. 203–218, 2019.

[18] S. Raschka, "Mlxtend: Providing machine learning and data science utilities and extensions to python's scientific computing stack," *The Journal of Open Source Software*, vol. 3, no. 24, Apr. 2018. [Online]. Available: http://joss.theoj.org/papers/10.21105/joss.00638

[19] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, vol. 30, pp. 3146–3154, 2017.

[20] H. Zhang, S. Si, and C.-J. Hsieh, "Gpu-acceleration for large-scale tree boosting," *arXiv preprint arXiv:1706.08359*, 2017.

[21] A. Schonfeld, "man-group/dtale: Visualizer for pandas data structures," Oct 2021. [Online]. Available: https://github.com/man-group/dtale

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[23] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, "API design for machine learning software: experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 2013, pp. 108–122.