

REFERENCES

- [1] N. Burkart and M. F. Huber, “A Survey on the Explainability of Supervised Machine Learning,” *J. Artif. Intell. Res.*, vol. 70, pp. 245–317, Jan. 2021, doi: 10.1613/jair.1.12228.
- [2] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization,” p. 9.
- [3] A. B. Arrieta *et al.*, “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI.” arXiv, Dec. 26, 2019. Accessed: Nov. 14, 2022. [Online]. Available: <http://arxiv.org/abs/1910.10045>
- [4] Z. C. Lipton, “In machine learning, the concept of interpretability is both important and slippery,” *Mach. Learn.*, p. 28.
- [5] G. Montavon, S. Lapuschkin, A. Binder, W. Samek, and K.-R. Müller, “Explaining nonlinear classification decisions with deep Taylor decomposition,” *Pattern Recognit.*, vol. 65, pp. 211–222, May 2017, doi: 10.1016/j.patcog.2016.11.008.
- [6] M. D. Zeiler and R. Fergus, “Visualizing and Understanding Convolutional Networks,” in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., in Lecture Notes in Computer Science, vol. 8689. Cham: Springer International Publishing, 2014, pp. 818–833. doi: 10.1007/978-3-319-10590-1_53.
- [7] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, and W. Samek, “On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation,” *PLOS ONE*, vol. 10, no. 7, p. e0130140, Jul. 2015, doi: 10.1371/journal.pone.0130140.
- [8] G. Ras, N. Xie, M. Van Gerven, and D. Doran, “Explainable Deep Learning: A Field Guide for the Uninitiated,” *J. Artif. Intell. Res.*, vol. 73, pp. 329–397, Jan. 2022, doi: 10.1613/jair.1.13200.
- [9] M. D. Zeiler, G. W. Taylor, and R. Fergus, “Adaptive deconvolutional networks for mid and high level feature learning,” in *2011 International Conference on Computer Vision*, Barcelona, Spain: IEEE, Nov. 2011, pp. 2018–2025. doi: 10.1109/ICCV.2011.6126474.
- [10] K. Simonyan, A. Vedaldi, and A. Zisserman, “Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps.” arXiv, Apr. 19, 2014. Accessed: Nov. 18, 2022. [Online]. Available: <http://arxiv.org/abs/1312.6034>
- [11] P. Angelov and E. Soares, “Towards explainable deep neural networks (xDNN),” *Neural Netw.*, vol. 130, pp. 185–194, Oct. 2020, doi: 10.1016/j.neunet.2020.07.010.

- [12] P. P. Angelov and X. Gu, *Empirical Approach to Machine Learning*, vol. 800. in *Studies in Computational Intelligence*, vol. 800. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-02384-3.
- [13] P. Angelov and E. Soares, “EXPLAINABLE-BY-DESIGN APPROACH FOR COVID-19 CLASSIFICATION VIA CT-SCAN,” *Health Informatics*, preprint, Apr. 2020. doi: 10.1101/2020.04.24.20078584.
- [14] D. Gunning, “Explainable Artificial Intelligence (XAI)”.
- [15] S. M. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions”.
- [16] M. T. Ribeiro, S. Singh, and C. Guestrin, ““Why Should I Trust You?”: Explaining the Predictions of Any Classifier.” *arXiv*, Aug. 09, 2016. Accessed: Apr. 21, 2023. [Online]. Available: <http://arxiv.org/abs/1602.04938>
- [17] C. Rudin, “Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead.” *arXiv*, Sep. 21, 2019. Accessed: Apr. 23, 2023. [Online]. Available: <http://arxiv.org/abs/1811.10154>
- [18] P. Angelov and E. Soares, “Towards Explainable Deep Neural Networks (xDNN).” *arXiv*, Dec. 05, 2019. Accessed: Jul. 30, 2022. [Online]. Available: <http://arxiv.org/abs/1912.02523>
- [19] S. Lipovetsky and M. Conklin, “Analysis of regression in game theory approach,” *Appl. Stoch. Models Bus. Ind.*, vol. 17, no. 4, pp. 319–330, Oct. 2001, doi: 10.1002/asmb.446.
- [20] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, “A survey of content-based image retrieval with high-level semantics,” *Pattern Recognit.*, vol. 40, no. 1, pp. 262–282, Jan. 2007, doi: 10.1016/j.patcog.2006.04.045.
- [21] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”.