به نام ایزد یکتا

تمرین هفتم درس روش پژوهش و ارائه

استاد: دکتر رضا صفابخش

تهیه کننده: بردیا اردکانیان

برسی و مهار مثال‌های خصمانه در یادگیری ماشین

Bibliography

Chicago (author-date)

"Adversarial Examples Are Not Easily Detected | Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security." n.d. Accessed May 28, 2022. https://dl.acm.org/doi/abs/10.1145/3128572.3140444.

"Adversarial Machine Learning Attacks and Defense Methods in the Cyber Security Domain | ACM Computing Surveys." n.d. Accessed May 27, 2022. https://dl.acm.org/doi/10.1145/3453158.

"Adversarial Machine Learning for Cybersecurity and Computer Vision: : Current Developments and Challenges: WIREs Computational Statistics: Vol 12, No 5." n.d. *Wiley Interdisciplinary Reviews: Computational Statistics*. Accessed May 27, 2022. https://dlnext.acm.org/doi/abs/10.1002/wics.1511.

Biggio, Battista, and Fabio Roli. 2018. "Wild Patterns: Ten Years after the Rise of Adversarial Machine Learning." *Pattern Recognition* 84 (December): 317–31. https://doi.org/10.1016/j.patcog.2018.07.023.

Brachman, Ronald J, and Peter Stone. n.d. "Synthesis Lectures on Artificial Intelligence and Machine Learning," 169.

Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. 2015. "Explaining and Harnessing Adversarial Examples." arXiv. http://arxiv.org/abs/1412.6572.

Huang, Ling, Anthony D. Joseph, Blaine Nelson, Benjamin I.P. Rubinstein, and J. D. Tygar. 2011. "Adversarial Machine Learning." In *Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence*, 43–58. AISec '11. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2046684.2046692.

Laskov, Pavel, and Richard Lippmann. 2010. "Machine Learning in Adversarial Environments." *Machine Learning* 81 (2): 115–19. https://doi.org/10.1007/s10994-010-5207-6.

"Machine Learning Uncertainties with Adversarial Neural Networks | SpringerLink." n.d. Accessed May 27, 2022. https://link.springer.com/article/10.1140/epjc/s10052-018-6511-8.

McDaniel, Patrick, Nicolas Papernot, and Z. Berkay Celik. 2016. "Machine Learning in Adversarial Settings." *IEEE Security Privacy* 14 (3): 68–72. https://doi.org/10.1109/MSP.2016.51.

Tabassi, Elham, Kevin J. Burns, Michael Hadjimichael, Andres D. Molina-Markham, and Julian T. Sexton. 2019. "A Taxonomy and Terminology of Adversarial Machine Learning." Preprint. https://doi.org/10.6028/NIST.IR.8269-draft.

Wiyatno, Rey Reza, Anqi Xu, Ousmane Dia, and Archy de Berker. 2019. "Adversarial Examples in Modern Machine Learning: A Review." arXiv. http://arxiv.org/abs/1911.05268.

Yuan, Xiaoyong, Pan He, Qile Zhu, and Xiaolin Li. 2019. "Adversarial Examples: Attacks and Defenses for Deep Learning." *IEEE Transactions on Neural Networks and Learning Systems* 30 (9): 2805–24. https://doi.org/10.1109/TNNLS.2018.2886017.

Zhang, Chaoning, Philipp Benz, Tooba Imtiaz, and In So Kweon. 2020. "Understanding Adversarial Examples From the Mutual Influence of Images and Perturbations." In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 14509–18. https://doi.org/10.1109/CVPR42600.2020.01453.

Zhao, Zhengli, Dheeru Dua, and Sameer Singh. 2018. "Generating Natural Adversarial Examples." arXiv. http://arxiv.org/abs/1710.11342.

Zheng, Haizhong, Ziqi Zhang, Juncheng Gu, Honglak Lee, and Atul Prakash. 2020. “Efficient Adversarial Training With Transferable Adversarial Examples.” In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 1178–87. Seattle, WA, USA: IEEE. https://doi.org/10.1109/CVPR42600.2020.00126.

Zizzo, Giulio, Chris Hankin, Sergio Maffeis, and Kevin Jones. 2019. “INVITED: Adversarial Machine Learning Beyond the Image Domain.” In *2019 56th ACM/IEEE Design Automation Conference (DAC)*, 1–4.