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

Harvard

Adversarial Examples Are Not Easily Detected | Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security (no date). Available at: https://dl.acm.org/doi/abs/10.1145/3128572.3140444 (Accessed: 28 May 2022).

Adversarial Machine Learning Attacks and Defense Methods in the Cyber Security Domain | ACM Computing Surveys (no date). Available at: https://dl.acm.org/doi/10.1145/3453158 (Accessed: 27 May 2022).

'Adversarial machine learning for cybersecurity and computer vision: : Current developments and challenges: WIREs Computational Statistics: Vol 12, No 5' (no date) *Wiley Interdisciplinary Reviews: Computational Statistics* [Preprint]. Available at: https://dlnext.acm.org/doi/abs/10.1002/wics.1511 (Accessed: 27 May 2022).

Biggio, B. and Roli, F. (2018) 'Wild patterns: Ten years after the rise of adversarial machine learning', *Pattern Recognition*, 84, pp. 317–331. doi:10.1016/j.patcog.2018.07.023.

Brachman, R.J. and Stone, P. (no date) 'Synthesis Lectures on Artificial Intelligence and Machine Learning', p. 169.

Goodfellow, I.J., Shlens, J. and Szegedy, C. (2015) 'Explaining and Harnessing Adversarial Examples'. arXiv. Available at: http://arxiv.org/abs/1412.6572 (Accessed: 27 May 2022).

Huang, L. *et al.* (2011) 'Adversarial machine learning', in *Proceedings of the 4th ACM workshop on Security and artificial intelligence*. New York, NY, USA: Association for Computing Machinery (AISec '11), pp. 43–58. doi:10.1145/2046684.2046692.

Laskov, P. and Lippmann, R. (2010) 'Machine learning in adversarial environments', *Machine Learning*, 81(2), pp. 115–119. doi:10.1007/s10994-010-5207-6.

Machine learning uncertainties with adversarial neural networks | SpringerLink (no date). Available at: https://link.springer.com/article/10.1140/epjc/s10052-018-6511-8 (Accessed: 27 May 2022).

McDaniel, P., Papernot, N. and Celik, Z.B. (2016) 'Machine Learning in Adversarial Settings', *IEEE Security Privacy*, 14(3), pp. 68–72. doi:10.1109/MSP.2016.51.

Tabassi, E. *et al.* (2019) *A taxonomy and terminology of adversarial machine learning*. preprint. doi:10.6028/NIST.IR.8269-draft.

Wiyatno, R.R. *et al.* (2019) 'Adversarial Examples in Modern Machine Learning: A Review'. arXiv. Available at: http://arxiv.org/abs/1911.05268 (Accessed: 28 May 2022).

Yuan, X. *et al.* (2019) 'Adversarial Examples: Attacks and Defenses for Deep Learning', *IEEE Transactions on Neural Networks and Learning Systems*, 30(9), pp. 2805–2824. doi:10.1109/TNNLS.2018.2886017.

Zhang, C. *et al.* (2020) 'Understanding Adversarial Examples From the Mutual Influence of Images and Perturbations', in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14509–14518. doi:10.1109/CVPR42600.2020.01453.

Zhao, Z., Dua, D. and Singh, S. (2018) 'Generating Natural Adversarial Examples'. arXiv. Available at: http://arxiv.org/abs/1710.11342 (Accessed: 28 May 2022).

Zheng, H. *et al.* (2020) 'Efficient Adversarial Training With Transferable Adversarial Examples', in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA: IEEE, pp. 1178–1187. doi:10.1109/CVPR42600.2020.00126.

Zizzo, G. *et al.* (2019) 'INVITED: Adversarial Machine Learning Beyond the Image Domain', in *2019 56th ACM/IEEE Design Automation Conference (DAC)*. *2019 56th ACM/IEEE Design Automation Conference (DAC)*, pp. 1–4.