“Security is not my field, I’m a stats guy”: A Qualitative Root Cause Analysis of Barriers to Adversarial Machine Learning Defenses in Industry

Bibliographic Citation:

Jaron Mink, Harjot Kaur, Juliane Schmüser, Sascha Fahl and Yasemin Acar. “Security is not my field, I’m a stats guy”: A Qualitative Root Cause Analysis of Barriers to Adversarial Machine Learning Defenses in Industry. In 32nd USENIX Security Symposium, USENIX Security '23, Anaheim, CA, USA, August 9-11, 2023.

Link: https://www.usenix.org/conference/usenixsecurity23/presentation/mink

Abstract:

Adversarial machine learning (AML) has the potential to leak training data, force arbitrary classifications, and greatly degrade overall performance of machine learning models, all of which academics and companies alike consider as serious issues. Despite this, seminal work has found that most organizations insufficiently protect against such threats. While the lack of defenses to AML is most commonly attributed to missing knowledge, it is unknown why mitigations are unrealized in industry projects. To better understand the reasons behind the lack of deployed AML defenses, we conduct semi-structured interviews (n=21) with data scientists and data engineers to explore what barriers impede the effective implementation of such defenses. We find that practitioners' ability to deploy defenses is hampered by three primary factors: a lack of institutional motivation and educational resources for these concepts, an inability to adequately assess their AML risk and make subsequent decisions, and organizational structures and goals that discourage implementation in favor of other objectives. We conclude by discussing practical recommendations for companies and practitioners to be made more aware of these risks, and better prepared to respond.