AI Safety - SRILAB (ETH Zurich)

Zohar Cochavi

The AI Safety Lab: https://www.sri.inf.ethz.ch/research/safeai

Interesting Papers

- Fair Normalizing Flows (Balunović, Ruoss, and Vechev 2022)

Balunović, Mislav, Anian Ruoss, and Martin Vechev. 2022. "Fair Normalizing Flows." arXiv. https://arxiv.org/abs/2106.05937.