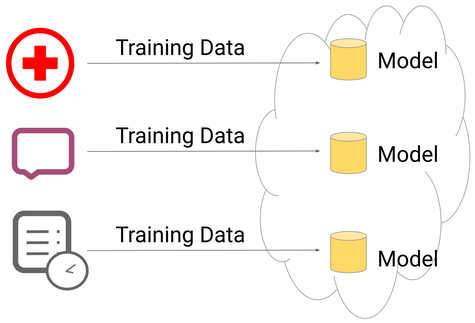
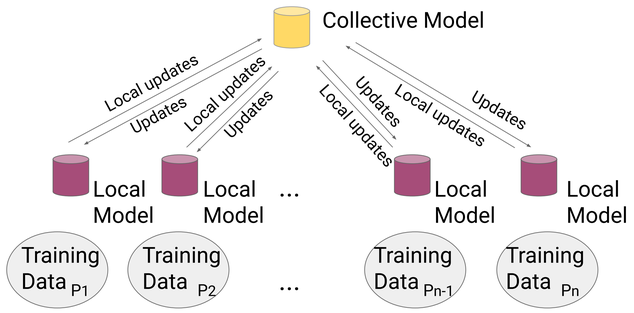
Private and Secure Collaborative Machine Learning

Bailey Kacsmar, Florian Kerschbaum

Machine Learning as a Service



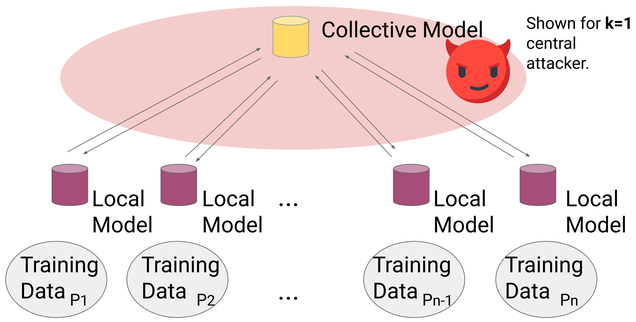
Collaborative Training



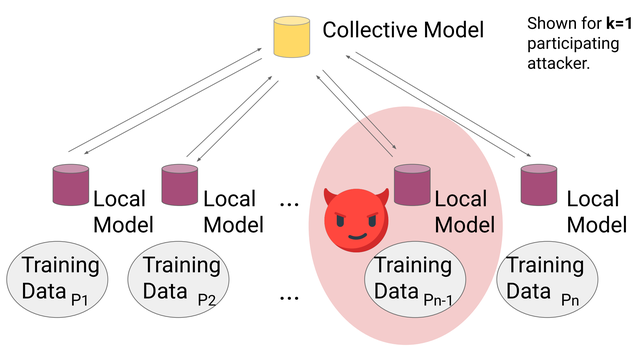
Privacy and Security

Goal: Adversaries possessing the collective model *and* who may have participated in training that model should be unable to learn any *additional* information about the training data used.

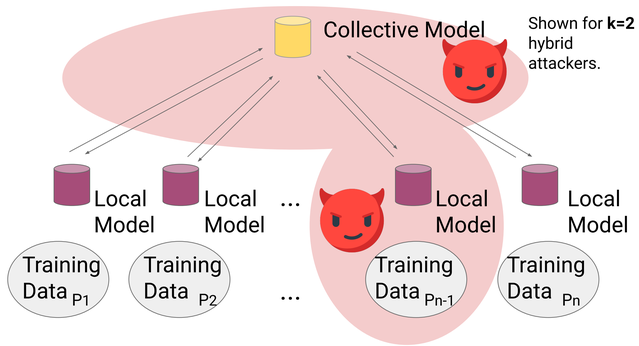
Coalition Type I



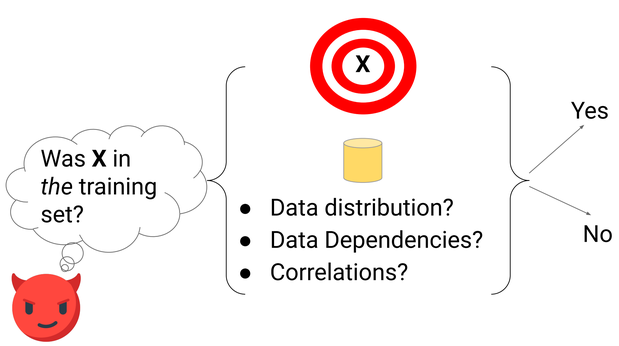
Coalition Type II



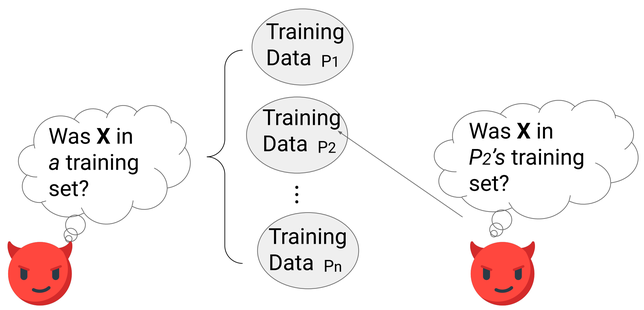
Coalition Type III



Membership Inference Attack



Membership Inference Type I & II



References

1. L. Melis, C. Song, E. De Cristofaro, V. Shmatikov. ”Exploiting Unintended Feature Leakage in Collaborative Learning." In 2019 IEEE Symposium on Security and Privacy IEEE, 2019.
2. M. Nasr, R. Shokri, A. Housmandr. ”Comprehensive Privacy Analysis of Deep Learning: Passive and Active White-box inference Attacks against Centralized and Federated Learning." In 2019 IEEE Symposium on Security and Privacy IEEE, 2019.
3. R. Shokri and V. Shmatikov. ”Privacy-Preserving Deep Learning." In Proceedings of the 22nd ACM SIGSAC CCS. ACM, 2015.

bkacsmar@uwaterloo.ca