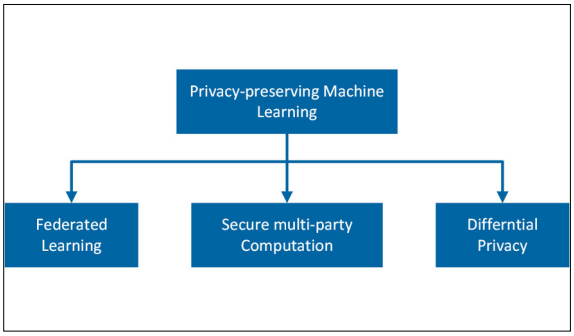
Sascha Jecklin

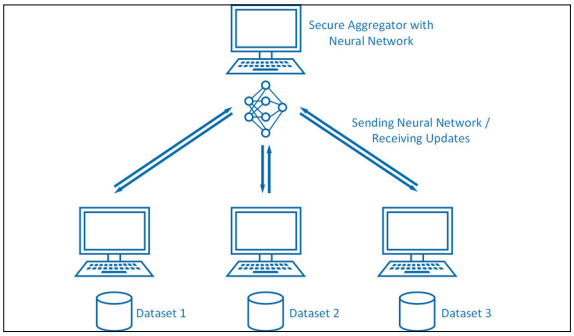
Graduate Candidate	Sascha Jecklin
Examiner	Prof. Dr. Guido Schuster
Co-Examiner	Fabian Riesen, Schaffhausen, SH
Subject Area	Sensor, Actuator and Communication Systems

Privacy-preserving Machine Learning Techniques

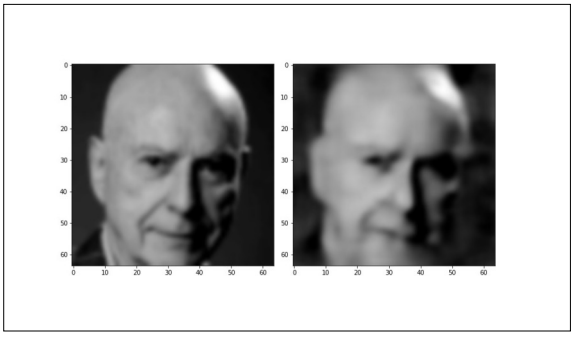
An Evaluation of the State of the Art



Privacy-preserving Machine Learning and its three main branches.
Own presentment



Working principle of Federated Learning.
Own presentment



Blackbox model-inversion attack is able to recreate input image (left) by intercepting a neural network layer.
Image on the left taken from CelebA dataset

Introduction: Machine learning is all about gaining knowledge from data. This data often contains some sensitive information. As machine learning becomes increasingly popular, privacy requires special attention. One might be willing to let a trusted machine process data. However, so far, not much research has been carried out on whether, and how machine learning can protect data owners. Furthermore, unlike in more established areas of software development, there are often no guarantees regarding the protection of intellectual property. Privacy-preserving machine learning (PPML) introduces methods to overcome both of those shortcomings. The aim of this work is to present and evaluate the state of the art of the working principles of PPML.

Approach: In a first step, the necessary basics of machine learning for PPML are discussed. This work then outlines attacker-models and potential threats related to machine learning. Afterwards, the three main components of PPML are explained in detail. These are namely federated learning, secure multi-party computation and differential privacy.
The second part deals with existing frameworks and displays potential pitfalls. In addition to that, practical examples are presented for all three PPML branches and some attacker models.

Result: This work shows that most existing frameworks only partially cover the branches of PPML. It also shows significant shortcomings regarding computation, that require further research. Nevertheless, this work outlines the importance of PPML and presents possible workarounds to overcome some of the shortcomings. Improvements of the frameworks are expected in the near future. The methods discussed in this work will then allow the transformation of traditional machine learning applications into privacy-preserving applications.