

Seminar

# Deep Learning with Differential Privacy

Differential privacy [1] is a security notion that is widely used in data analytics. A differentially private algorithm guarantees that the privacy of an individual is not harmed while it is still possible to learn about a population.

This concept can be transferred to the domain of machine learning. In this setting, model is trained based on potentially sensitive data. For classification tasks for example, the trained model is stored on untrusted devices. Although only the trained model and not the data itself is stored, it was shown, however, that the model can still provide information about individual training data samples. Thus, a learning algorithm is required that preserves the privacy of training data samples. Such a differentially private learning algorithm has been introduced in [2].

[1] Dwork, Cynthia, and Aaron Roth. "The algorithmic foundations of differential privacy." Foundations and Trends® in Theoretical Computer Science 9.3–4 (2014): 211-407.

[2] Abadi, Martin, et al. "Deep learning with differential privacy." Proceedings of the 2016 ACM SIGSAC conference on computer and communications security. 2016.

[3] Geyer, Robin C., Tassilo Klein, and Moin Nabi. "Differentially private federated learning: A client level perspective." arXiv preprint arXiv:1712.07557 (2017).

## Prerequisites

Prior knowledge on

- machine learning
- probability theory and statistics

## Contact

Luis Maßny (luis.massny@tum.de)

## Advisors

Luis Maßny