Associate Professorship of Coding and Cryptography
TUM Department of Electrical and Computer Engineering
Technical University of Munich

Seminar

Gradient Compression

In many distributed and federated learning systems clients iteratively compute so-called gradient vectors based on their locally stored data and communicate them to a central entity. The gradient vectors are typically high dimensional, so transmitting them directly leads to undesirable amounts of data transmission, holding back the performance of the system.

To alleviate this issue, various gradient compression schemes have been proposed.

The student's task is to analyze and compare multiple proposed schemes based on their advantages and disadvantages. As a starting point, students can use [1, Section 3.2].

[1] S. Ouyang, D. Dong, Y. Xu, and L. Xiao, “Communication optimization strategies for distributed deep neural network training: A survey,” Journal of Parallel and Distributed Computing, vol. 149, pp. 52–65, Mar. 2021, doi: 10.1016/j.jpdc.2020.11.005.

Advisors

Christoph Hofmeister