EP3260: Machine Learning Over Networks
Computer Assignment 6
Instructors: Hossein S. Ghadikolaei, José Mairton B. da Silva Jr.
Due Date: May 6, 2020

Computer Assignment 6 - Communication Efficiency

Split the "MNIST" dataset to 10 random disjoint subsets, each for one worker, and consider SVM classifier in the form of $\min_{\mathbf{w}} \frac{1}{N} \sum_{i\in[N]} f_i(\mathbf{w})$ with $N = 10$. An alternative approach to improve communication-efficiency is to compress the information message to be exchanged (usually gradients – either in primal or dual forms). Consider two compression/quantization methods for a vector: (Q1) keep only K values of a vector and set the rest to zero and (Q2) represent every element with fewer bits (e.g., 4 bits instead of 32 bits).

a) Repeat parts a-b from CA5 using Q1 and Q2. Can you integrate Q1/Q2 to your solution in part c of CA5? Discuss.

b) How do you make SVRG and SAG communication efficient for large-scale ML?