Algorithms for Distributed Machine Learning

Distributed machine learning, as the name suggests, refers to learning over data that is distributed amongst multiple machines. Each machine holds only a piece of the entire data, and the objective is to learn a parameter that minimizes a loss function defined over the entire data. The problem of distributed machine learning can be formulated as a special class of a more general distributed optimization problems.

This seminar course will discuss the *state-of-the-algorithms* for solving distributed optimization problems, in the context of distributed machine learning. The tentative timeline for the topics to be covered in the course is as follows:

Week #	Topic
1	Introduction to Distributed Optimization
2, 3	Existing Algorithms for Distributed Optimization
4	Convergence Analysis of Algorithms
5	Asynchronous Algorithms
6	Stochastic Approximations (for large-scale learning)
7	Introduction to Fault-Tolerance
8	Introduction to Privacy Challenges
9	Further Challenges in Distributed Learning
Remaining Weeks	Student Presentations

A partial list of the papers to be discussed are as follows. Other papers may be added, depending upon the progress of the course.

Partial List of Papers:

1. Tsitsiklis, John, Dimitri Bertsekas, and Michael Athans. "Distributed asynchronous deterministic and stochastic gradient optimization algorithms." *IEEE transactions on automatic control* 31.9 (1986): 803-812
2. Boyd, Stephen, et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers." *Foundations and Trends in Machine learning* 3.1 (2011): 1-122.
3. Nedic, Angelia, and Asuman Ozdaglar. "Distributed subgradient methods for multi-agent optimization." *IEEE Transactions on Automatic Control* 54.1 (2009): 48.

4. Terelius, Håkan, Ufuk Topcu, and Richard M. Murray. "Decentralized multi-agent optimization via dual decomposition." *IFAC Proceedings Volumes* 44.1 (2011): 11245-11251.
5. Bottou, Léon. "Online learning and stochastic approximations." *On-line learning in neural networks* 17.9 (1998): 142.
6. Bottou, Léon, Frank E. Curtis, and Jorge Nocedal. "Optimization methods for large-scale machine learning." *Siam Review* 60.2 (2018): 223-311
7. Su, Lili, and Nitin H. Vaidya. "Fault-tolerant multi-agent optimization: optimal iterative distributed algorithms." *Proceedings of the 2016 ACM symposium on principles of distributed computing*. ACM, 2016.
8. Blanchard, Peva, Rachid Guerraoui, and Julien Stainer. "Machine learning with adversaries: Byzantine tolerant gradient descent." *Advances in Neural Information Processing Systems*. 2017.
9. Gupta, Nirupam, and Nitin H. Vaidya. "Byzantine fault tolerant distributed linear regression." *arXiv preprint arXiv:1903.08752* (2019).
10. Huang, Zhenqi, Sayan Mitra, and Nitin Vaidya. "Differentially private distributed optimization." *Proceedings of the 2015 International Conference on Distributed Computing and Networking*. ACM, 2015.

Evaluation:

To evaluate the participation of students, the course will have weekly assignments and a final report. In the weekly assignments the students will be asked to submit a short summary of the topics covered in that week. For the final report, each student is expected to identify and choose a research problem in distributed machine learning. A list of potential research problems will be provided for reference. Students will be asked to submit a five-page report and give a short presentation on the chosen topic.