INITIAL PROJECT REPORT

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRONIC AND ELECTRICAL ENIGNERING

Primal-Dual Techniques for Distributed Multitask Learning

Author:
Zhiyu Ma (CID: 01844618)

Supervisor:
Dr S. Vlaski

Date: March 10, 2024

1 Aims of the Project

Traditional distributed multitask learning algorithms are observed to introduce bias and noise into the system, leading to increased instability and communication costs. The project is dedicated to developing an algorithm aimed at mitigating bias and minimizing noise, with the ultimate goal of enhancing system stability and reducing communication overhead.

To achieve this, there are some possible methods as shown below.

- **Task-specific Regularization**
  Implement task-specific regularization techniques to control overfitting and bias. By incorporating regularization terms tailored to each task, the algorithm can prevent the model from becoming overly biased towards certain tasks.

- **Dynamic Learning Rates**
  Introduce dynamic learning rate adjustments based on task performance. Adapting learning rates dynamically can help the algorithm prioritize tasks with higher uncertainty or those showing slower convergence, minimizing noise in the learning process.

- **Adaptive Weighting Mechanisms**
  Develop adaptive weighting mechanisms for task contributions. This involves dynamically adjusting the importance assigned to each task based on its current learning progress or relevance, mitigating biases introduced by tasks with disproportionately high influence.

- **Communication-Efficient Architectures** Explore communication-efficient architectures that minimize the amount of information exchanged between distributed nodes. This can help in reducing communication costs while maintaining the integrity of the multitask learning process.

2 Background Materials

For background research, there are several pertinent papers available.

2.1 Multitask Learning Over Graphs

(1) elucidates how prior knowledge about task relationships can be integrated into the adaptive mechanism to formulate more effective strategies for multitask learning. The author underscores the significance of selecting the optimal strategy, essentially choosing the task-relatedness model that best aligns with the underlying problem, while maintaining a balance between model adaptability and computational/communication constraints.

The focal point of the paper is multitask learning networks, particularly those involving parameter estimation tasks. These networks are designed to concurrently handle multiple tasks and optimize performance through the incorporation of prior knowledge. While the primary emphasis is on parameter estimation tasks, the author mentions other multitask perspectives, such as distributed detection, highlighting the widespread applicability of multitask learning across various domains and applications.

2.2 Networked Federated Multi-Task Learning

The primary conceptual contribution of (2) is the formalization of networked federated learning using generalized total variation (GTV) minimization as a regularizer. This formulation is highly flexible, allowing it to be combined with almost any parametric model, including Lasso or deep neural networks.

2.3 Fed-QSSL: A Framework for Personalized Federated Learning under Bitwidth and Data Heterogeneity

(3) propose a federated quantization-based self-supervised learning scheme (Fed-QSSL) designed to address heterogeneity in federated learning (FL) systems. At the client side, to handle data heterogeneity, they leverage distributed self-supervised learning while using low-bit quantization to meet the constraints imposed by local infrastructure and limited communication resources. On the server side, Fed-QSSL employs de-quantization, weighted aggregation, and re-quantization, ultimately creating models personalized to both the data distribution and specific infrastructure of each client's device.

References

[1] R. Nassif, S. Vlaski, C. Richard, J. Chen, and A. H. Sayed, "Multitask learning over graphs: An approach for distributed, streaming machine learning," **IEEE Signal Processing Magazine**, vol. 37, no. 3, pp. 14–25, 2020. pages 2

[2] Y. SarcheshmehPour, Y. Tian, L. Zhang, and A. Jung, "Networked federated multi-task learning," **Authorea Preprints**, 2023. pages 3

[3] Y. Chen, H. Vikalo, and C. Wang, "Fed-qssl: A framework for personalized federated learning under bitwidth and data heterogeneity," **arXiv preprint arXiv:2312.13380**, 2023. pages 3