View Reviews

Paper ID
6233

Paper Title
Learning Privately over Distributed Features: An ADMM Sharing Approach

Reviewer #1

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words.
In this paper, the authors apply the alternating direction method of multipliers (ADMM) to solve the learning problem where features are distributed across workers. The authors also propose to add Gaussian noise to the messages sent to the master node so as to protect privacy.

2. [Relevance] Is this paper relevant to an AI audience?
Likely to be of interest to a large proportion of the community

3. [Significance] Are the results significant?
Moderately significant

4. [Novelty] Are the problems or approaches novel?
Somewhat novel or somewhat incremental

5. [Soundness] Is the paper technically sound?
Technically sound

6. [Evaluation] Are claims well-supported by theoretical analysis or experimental results?
Somewhat weak

7. [Clarity] Is the paper well-organized and clearly written?
Good

8. [Detailed Comments] Please elaborate on your assessments and provide constructive feedback.
My comments are as follows.
1. Above (1)-(2), the authors assume the form of the loss function, and claim that it incorporates several models. Please illustrate with one or two models.
2. The authors mention the computation complexity issue several times, but do not explicitly show the complexities of ADMM and gradient descent such that the comparison seems not convincing.
3. The authors claim that the analysis improves over that in (Hong 2016). What are the differences in the proofs?
4. The authors show the sources of Assumptions 2.1-2.3, but do not explain why these assumptions are reasonable.

9. [QUESTIONS FOR THE AUTHORS] Please provide questions for authors to address during the author feedback period.
My suggestions to the authors.
1. Be more concrete in discussing the loss function and the assumptions.
2. Explicitly show the complexity of ADMM and gradient descent.
3. Explain the differences in the proofs comparing to those in (Hong 2016).

10. [OVERALL SCORE]
6 - Marginally above threshold

Reviewer #2

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words.
The authors apply nonconvex ADMM to solve the feature distributed learning problem and show the convergence of their proposed algorithm. In addition, they integrate the differential privacy technique to the proposed algorithm. Preliminary experiments are encouraging.

2. [Relevance] Is this paper relevant to an AI audience?
Relevant to researchers in subareas only

3. [Significance] Are the results significant?
Moderately significant

4. [Novelty] Are the problems or approaches novel?
Somewhat novel or somewhat incremental

5. [Soundness] Is the paper technically sound?
Technically sound

6. [Evaluation] Are claims well-supported by theoretical analysis or experimental results?
Somewhat weak

7. [Clarity] Is the paper well-organized and clearly written?
Good

8. [Detailed Comments] Please elaborate on your assessments and provide constructive feedback.
1. In Theorem 1, the authors assume that D_m is a compact set. However, D_m is a linear operator!
2. The proposed algorithm 1 is a linearized version of [1] by linearizing the augmented terms for variables x_{m} and the convergence analysis techniques are directly adopted from [1], which limits the contribution of this paper.
3. The penalty parameter ρ highly depends on L, D_{m} and would greatly influence the numerical performance. More experiments should be provided to show the influence of the hyperparameter.
4. The feature dimensions of the tested datasets a9a and gisette are 123 and 5000, respectively, which is too small to demonstrate the advantage of the proposed algorithm for datasets with high dimensional feature spaces.

[1] Hong, Mingyi, Zhi-Quan Luo, and Meisam Razaviyayn. "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems." SIAM Journal on Optimization 26.1 (2016): 337-364.

9. [QUESTIONS FOR THE AUTHORS] Please provide questions for authors to address during the author feedback period.
see the comments above.

10. [OVERALL SCORE]
4 - Reject

Reviewer #3

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words.
The authors propose an ADMM sharing framework to approach risk minimization over distributed features.

2. [Relevance] Is this paper relevant to an AI audience?
Of limited interest to an AI audience

3. [Significance] Are the results significant?
Moderately significant

4. [Novelty] Are the problems or approaches novel?
Somewhat novel or somewhat incremental

5. [Soundness] Is the paper technically sound?
Technically sound

6. [Evaluation] Are claims well-supported by theoretical analysis or experimental results?
Somewhat weak

7. [Clarity] Is the paper well-organized and clearly written?
Satisfactory

8. [Detailed Comments] Please elaborate on your assessments and provide constructive feedback.
please see the "questions for the authors".

9. [QUESTIONS FOR THE AUTHORS] Please provide questions for authors to address during the author feedback period.
1. Unlike the Gauss-Seidel type update, the authors adopt Jacobian type update for x_m. However, this type of update have existed since the early development of the discipline. Although the authors name it as ''sharing'', the update scheme is not new in itself. Besides, updating variables in a Jacobian manner is equal to regard all x_m as a whole, thus the original problem (3) with M+1 variables is turned to a problem with only 2 variables: the convergence analyses in [Hong 2016] are naturally established. The authors should clarify this concern if I have mistaken it.

2. The assumption that each subproblem is strongly convex is relatively weak. The authors employ an l2-norm regularized logistic regression problem to verify their algorithm. Is there any more realistic problems that satisfy the assumptions given in the Assumption 1? If so, the authors should have applied the algorithm to those more realistic problems and data.

10. [OVERALL SCORE]
4 - Reject