

Computational Mathematics for AI: Numerical Methods and Distributed Computing for Deep Learning on Big Data

Pouya Ataei

August 26, 2024

1 Introduction

This document outlines the protocol for a systematic literature review (SLR) on computational mathematics for AI, focusing on numerical methods and distributed computing techniques for deep learning on big data. The review will follow the PRISMA guidelines (?) and Kitchenham's methodology for SLRs in software engineering (?).

2 Background

2.1 Rationale

As deep learning models grow in complexity and data volumes continue to increase, there is a critical need to understand and optimize the computational methods underpinning these systems. This review aims to synthesize current knowledge on numerical methods and distributed computing techniques specifically applied to deep learning in big data contexts.

2.2 Objectives

The primary objectives of this SLR are:

1. To identify and categorize state-of-the-art numerical methods used in deep learning for big data.
2. To evaluate the effectiveness of various distributed computing techniques for scaling deep learning to big data problems.

3. To compare these methods and techniques in terms of computational efficiency, scalability, and accuracy.
4. To identify emerging trends and future directions in this field.

3 Methods

3.1 Eligibility Criteria

3.1.1 Inclusion Criteria

- Studies published in the last 10 years (2019-2024)
- Peer-reviewed journal articles and conference papers
- English language publications
- Studies that directly address numerical methods or distributed computing for deep learning on big data

3.1.2 Exclusion Criteria

- Studies not focusing on big data scenarios
- Publications without clear methodological details
- Review papers (can be used for background but not included in the analysis)

3.2 Information Sources

The following electronic databases will be searched:

- IEEE Xplore
- ACM Digital Library
- arXiv
- Google Scholar
- Scopus

3.3 Search Strategy

The search string will be constructed using the following terms:

```
((("deep learning" OR "neural network") AND  
("big data" OR "large-scale") AND  
("numerical method*" OR "distributed computing" OR  
"parallel processing" OR "GPU acceleration" OR  
"federated learning")))
```

The search string will be adapted for each database as necessary.

3.4 Study Records

3.4.1 Data Management

References will be managed using Zotero. A spreadsheet will be used to track the selection process and reasons for exclusion.

3.4.2 Selection Process

Two reviewers will independently screen titles and abstracts. Full texts of potentially eligible studies will then be retrieved and independently assessed for eligibility by two reviewers. Disagreements will be resolved through discussion with a third reviewer.

3.4.3 Data Collection Process

A standardized, pre-piloted form will be used to extract data from the included studies. Two reviewers will extract data independently, with discrepancies resolved by consensus or by a third reviewer.

3.5 Data Items

For each included study, the following information will be extracted:

- Bibliographic information (authors, year, title, venue)
- Type of numerical method(s) or distributed computing technique(s) used
- Problem domain and dataset characteristics
- Performance metrics (e.g., training time, accuracy, scalability)

- Hardware and software environment
- Key findings and limitations

3.6 Outcomes and Prioritization

Primary outcomes will include:

- Effectiveness of numerical methods in terms of convergence speed and accuracy
- Scalability of distributed computing techniques
- Trade-offs between computational efficiency and model performance

Secondary outcomes will include emerging trends and identified research gaps.

3.7 Risk of Bias in Individual Studies

The quality of individual studies will be assessed using a modified version of the Quality Assessment Checklist for Technology Reports (?). Two reviewers will independently assess the quality of each included study.

3.8 Data Synthesis

A narrative synthesis will be provided of the findings from the included studies, structured around the types of interventions, target population characteristics, outcome measures and intervention content. Where possible, we will conduct:

3.8.1 Network Analysis

Using network analysis tools (e.g., Gephi), we will create a network graph where nodes represent papers and edges represent shared methods, techniques, or citations.

3.8.2 Meta-Analysis

For studies with comparable outcomes, we will perform a meta-analysis using a random-effects model. Forest plots will be used to visualize the combined effect sizes.

3.9 Meta-Bias(es)

To assess reporting bias, we will examine funnel plots for asymmetry when there are 10 or more studies in a meta-analysis.

4 Discussion

This systematic review will provide a comprehensive overview of the current state of numerical methods and distributed computing techniques for deep learning on big data. The findings will be interpreted considering the strength of evidence, applicability, and generalizability. Limitations of the review and the included studies will be discussed, and implications for future research will be outlined.