

Parallel approaches to machine learning - A comprehensive survey

DJIEMBOU TIENTCHEU VICTOR NICO¹

¹*Université de Yaoundé I,
Faculté des Sciences,
Département d'Informatique,
Etudiant Master DS*

November 29, 2023



Faculté des Sciences

Outline

Introduction

Approches de parallélisation

- Approche de parallélisation général sur des architectures multicoeurs
- Parallélisation sur des GPUs
- Utilisation de la technique de map reduce

Conclusion

Introduction

Context

- temps d'exécutions exponentiel
- espace nécessaire grossissant
- multi-coeurs
- GPUs

motivation

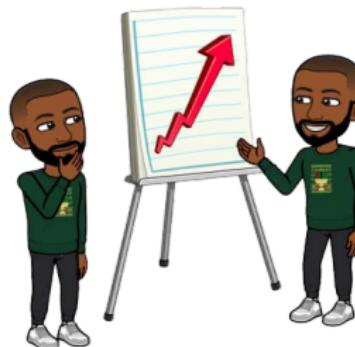
- comment établir la relation entre un problème données et les approches de parallélisation ?
- comment faire le choix d'une approche ?
- quels sont les contraintes d'implementation de ce choix ?



Introduction

Problème

En considérant un ensemble de catégories de problèmes ou encore un ensemble de modèles, le problème est de savoir quel est l'approche d'optimisation de la complexité sur les axes temporels et spatiales le plus recommandé selon la littérature



Approches de parrallélisation

Les 3 tendances

- Approche de parallélisation général sur des architectures multicoeurs
- Parallélisation sur des GPUs
- Utilisation de la technique de map reduce.



Approche de parallélisation général sur des architectures multicoeurs

Intérêts

- Parallélisme au niveau du thread
- Faible latence de communication
- Gestion dynamique des ressources

Figure: Processeur multi-coeur



Approche de parallélisation général sur des architectures multicoeurs

Quelques réalisations en terme de parallélisation

- Parallélisation des algorithmes basé sur Apriori
- Parallélisation des algorithmes basé sur DHP
- Parallélisation des algorithmes de certains algorithmes d'apprentissage automatique



Parallélisation sur des GPUs

Intérêts

- Architecture parallèle
- Accélération des opérations matricielles
- Mémoire vidéo dédiée

Figure: GPU



Utilisation de la technique de map reduce

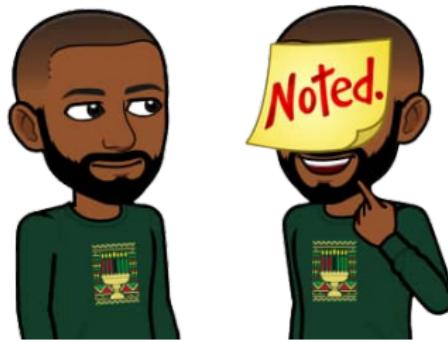
Intérêts

- Parallélisme
- Traitement distribué
- Tolérance aux pannes

Figure: Map Reduce



Conclusion



References

-  Kurt Thearling, *Massively parallel architectures and algorithms for time series analysis*, Lectures in Complex Systems, Addison-Wesley (1996).
-  Sujatha R Upadhyaya, *Parallel approaches to machine learning—a comprehensive survey*, Journal of Parallel and Distributed Computing **73** (2013), no. 3, 284–292.

Merci de votre attention

Victor DJIEMBOU¹

¹ *Université de Yaoundé I,
Faculté des Sciences,
Département d'Informatique,
Etudiant Master DS*