



Computer Science Department

Examen INF5099 Apprentissage Distribué, 11 Mai 2021

Docteur Thomas MESSI

**Exercice 1 : Parallélisation de K-means avec Map Reduce (10 points).**

1. Algorithme séquentiel
  - (a) Rappeler le principe de K-Means
  - (b) Rappeler le fonctionnement de ce algorithme à l'aide du pseudo-code
2. Version parallèle
  - (a) Donner une stratégie permettant de le paralléliser.
  - (b) Donner la version parallèle de l'algorithme.
  - (c) Rappeler la loi d'Amdhal et dire quelle est la fraction parallélisable de votre code.
  - (d) Utiliser cette fraction pour donner le speedup maximal de votre implémentation.
  - (e) Réaliser une implémentation de l'algorithme en utilisant Map Reduce et posix thread.  
(On donnera juste l'essentiel : code des mappers, code des reducers).

**Exercice 2 : Parallélisation d'un algorithme d'apprentissage artificiel. (10 points)**

Considérez les algorithmes suivant : SVM, RNN et CNN

1. Faire un choix, puis donner une description de l'algorithme séquentiel (pseudo code). On suppose que vous utilisez la descente du gradient (batch, minibatch ou stochastique).
2. Dire comment paralléliser l'algorithme en parallisant uniquement la descente du gradient.
3. Donner les grandes lignes d'une implémentation avec posix threads.
4. Après avoir donné une estimation de la fraction parallélisable, donnez une estimation du speedup maximal.

..... Bon Courage ! .....