



Programación Concurrente y Distribuida (CC65)

Tarea Académica 1

Secciones: Todas
Profesores: Todos

Instrucciones

- La tarea se desarrolla de manera individual.
- Toda respuesta escrita debe ser gramatical y ortográficamente correcta. La buena redacción es indispensable para otorgar el puntaje completo.
- Toda cita literal debe estar correctamente indicada y asociada a su correspondiente referencia.
- Las respuestas deben ser concisas, pero deben incluir la información suficiente para poder ser adecuadamente entendidas.

Contexto

Deep Learning (DL), ha experimentado un incremento en protagonismo en los últimos 7 años; en especial por el mejor aprovechamiento de sistemas computacionales con poder de procesamiento paralelo, como el procesamiento masivo paralelo heterogéneo, que combina el uso de CPU multicore y GPU con miles de cores.

Herramientas como OpenCL, CUDA, OpenMP, OpenMPI y otras, han permitido el desarrollo de gran variedad de poderosas bibliotecas para machine learning tales como Tensor Flow, Caffe, Theano, PyTorch y otros. Estas bibliotecas potencian el desarrollo de modelos cada vez más sofisticados, al punto que el estándar de rendimiento esperado de un modelo, es superar la capacidad del ser humano para realizar dicha tarea.

En la misma dirección, la aparición de hardware específico para deep learning como los TPU propuestos por Google, y propuestas similares por parte de Intel y NVidia, presentan un panorama en el que la programación paralela masiva se ha convertido en el presente de la programación. No obstante, llevar a producción estos sistemas construidos sobre algoritmos de deep learning, es un reto constante que requiere de un gran abanico de funcionalidades como lo son la auto escalabilidad horizontal, autocuración, lanzamientos automáticos, atomicidad, entre otros.

Descripción de la Tarea

La tarea académica consiste en encontrar 3 papers académicos (últimos 5 años), usando fuentes como Scopus, IEEE Explore, ACM Digital Library, etc. las cuales están disponibles en la universidad a través del centro de información (<http://recursosinvestigacion.upc.edu.pe/>). **El tema central de los papers debe ser Deep Learning. Deben describir el uso de aplicaciones, servicios, bibliotecas sobre infraestructura contenerizada para dar soporte a sistemas paralelos y distribuidos para procesamiento de los correspondientes modelos.**

Por cada paper deberá presentar una tabla con la siguiente información:

- Título del paper
- Autores
- Año de publicación
- Motivación
- Problema
- Propuesta de solución
- Algoritmos de deep learning usados

- Servicios, bibliotecas usadas y descripción de cómo el algoritmo es paralelizado
- Arquitectura física o diagrama de infraestructura contenerizada usada.
- Opinión crítica de su parte (estudiante) de las los beneficios o perjuicios del uso de paralelismo en dicho algoritmo.

Rúbrica de calificación

	Sobresaliente	Esperado	Deficiente
Calidad de los papers	Top ranking según Scimago o ranking similar. Autores afiliados a Universidades Top	Paper figura en simago o ranking similar pero no es top.	Paper no categorizado.
	3 puntos	1.5 puntos	0 puntos
Análisis	Identifica claramente la motivación, problema y propuesta de solución y describe las variables del problema.	Identifica superficialmente la motivación el problema y propuesta de solución.	No desarrollado o muy deficiente.
	5 puntos	2.5 puntos	0 puntos
Elementos técnicos	Identifica los algoritmos, bibliotecas, servicios utilizados, describe el tipo de algoritmo y su uso.	Identifica superficialmente los algoritmos, servicios y bibliotecas utilizadas.	No desarrollado o muy deficiente.
	6 puntos	3 puntos	0 puntos
Criterio del estudiante	Expresa con un lenguaje claro y conciso y a través de ejemplos los beneficios o desventajas de las implementaciones encontradas.	Usa un lenguaje limitado o identifica superficialmente las ventajas y desventajas de los algoritmos.	No desarrollado o muy deficiente.
	6 puntos	3 puntos	0 puntos