

INFORME DE AVANCE DEL TRABAJO DE TITULACIÓN

SISTEMAS INFORMÁTICOS Y COMPUTACIÓN

PRACTICUM 4.2

Abril - Agosto 2020

Fecha Informe:	16 de July de 2020	Nro. Informe	2
Tema del trabajo:	Implementación de un Ambiente de Ciencia de Datos a través de la librería Open Source Rapids		
Objetivos del trabajo:	General: Implementación de un Ambiente de Ciencia de Datos a través de la librería Open Source Rapids Específicos: - Elaborar una investigación documentada sobre la librería Open Source Rapids. - Realizar analítica de grandes volúmenes de datos a través de uso de las características de la librería Open Source Rapids. - Implementar algoritmos de Machine Learning haciendo uso de la librería Open Source Rapids.		
Estudiante:	José Alberto Guarnizo Romero		
Nombres Director:	Mgr. René Elizalde	Departamento	Ciencias de la Computación y Electrónica
Sección Departamental/Unidad	Tecnologías Avanzadas de la Web y Sistemas Basados en el Conocimiento		

Estado General Proyecto					
Fecha Inicio Planificada	18 de October de 2019	Fecha Inicio	18 de octubre del 2019		
Fecha Fin Planificada	15 de julio de 2020	Fecha Fin	16 de julio del 2020		
Porcentaje Avance Actual	19774100 %	% Retraso	46 %		

Cronograma						
Nombre Fase/Capítulo	Aporte al TT (100%)	Fecha inicio planificada	Fecha fin planificada	Fecha cierre real	% Avance actual (100% por fase)	% Retraso
Marco Teórico sobre la librería Rapids practicum 4.1	100 %	18/10/2019	14/02/2020	14/02/2020	100 %	0 %
Estado del arte sobre la librería Rapids practicum 4.1	50 %	18/10/2019	17/02/2020	17/02/2020	100 %	0 %
Análisis de la Arquitectura y características de la librería Rapids practicum 4.1	100 %	13/01/2020	14/02/2020	14/02/2020	100 %	0 %
Uso de la librería Rapids para la analítica de grandes volúmenes de datos practicum 4.2	100 %	20/04/2020	05/07/2020	13/07/2020	85 %	15 %
Uso de la librería Rapids para la implementación de algoritmos machine learning practicum 4.2	100 %	08/05/2020	10/07/2020	15/07/2020	69 %	31 %

Riesgos		
Riesgos Identificados	Acciones Tomadas	Impacto
Máquina portátil personal no cumple con las prestaciones para el uso de la librería, en el aspecto de no tener una GPU incorporada y rendimiento para analizar bastante información (datos).	Pedir máquina portátil	Alto

Resultados obtenidos hasta la fecha:

- Documentación detallada de las características y funcionamiento de Rapids.
- Manual de instalación de la librería Rapids.
- Arquitectura Rapids analizada e implementada.
- Desarrollo y codificación de la librería Rapids.
- Comparativa de resultados de aceleración en el proceso de ETL, comparando Rapids y Pandas.

Indique los problemas que hayan ocasionado retraso en las actividades:

- No existe mucha documentación en vista que la librería es sumamente nueva.
- Máquina portátil personal no cumple con los prerrequisitos de instalación, para poder utilizar la librería y no tiene buen rendimiento de uso para analizar bastante información (datos).

Breve descripción de las actividades que se realizarán en los próximos 2 meses:

- Selección de los algoritmos machine learning, para comenzar a predecir.
- Analizar la biblioteca cuxfilter de Rapids, para la representación de resultados de forma gráfica.

Observaciones del Director:



José Guarnizo
Tesisista

Mgr. René Elizalde
Director de tesis

Ing. Audrey Romero
Tutora Practicum 4.2

Phd. Priscila Valdiviezo
Tutora Practicum 4.2