

1. Descripción General

El objetivo de este proyecto es desarrollar uno o varios modelos de grafos empleando los conceptos y herramientas vistos en clase para la solución de un problema en un contexto real. Para el proyecto se requiere:

1. Definir claramente el problema, su justificación de por qué es relevante en su contexto, y por qué un modelo de grafos es útil para el problema.
2. Identificar el conjunto de datos a emplear.
3. Diseñar e implementar un modelo de grafos usando los datos identificados para el problema planteado.
4. Realizar una exploración de hiperparámetros para determinar el mejor valor para cada uno.
5. Realizar una presentación de máximo 10 minutos con los pasos y resultados clave [**opcional**] del proyecto.

2. Entregables

Como resultados de las tareas anteriores deberá entregar los siguientes resultados y soportes:

- (+10 puntos) Resultado 1: presentación de no más de 10 minutos con los resultados principales del proyecto [**opcional**].
- (45 puntos) Resultado 2: código con el procesamiento de datos y modelos desarrollados que soporten los resultados reportados.
- (45 puntos) Resultado 3: Informe con todos los detalles de la solución, el análisis y las conclusiones. El informe no podrá exceder las 10 páginas, deberá enviarse en formato PDF y deberá ser autocontenido.
- (10 puntos) Reporte de trabajo en equipo: incluya un pequeño reporte de cómo se dividieron las tareas entre los miembros del equipo. Para cada miembro del equipo liste las tareas desarrolladas.
- Soporte 1: Informe con todos los resultados de la solución.
- Soporte 2: enlace a github de los datos usados o el conjunto de datos.
- Soporte 3: fuentes de procesamiento de datos y construcción y evaluación de modelos (archivos .py o .ipynb).

Nota: si bien el trabajo es en grupo (de 4 o 5 personas), la nota es individual, luego es necesario que cada miembro del equipo demuestre su contribución al proyecto.

3. Recomendaciones

- Sea conciso y eficiente con el espacio. Ni el reporte ni la presentación deben ser largos. Al contrario, en un buen reporte cada gráfica y afirmación importa, y en una buena presentación cada diapositiva cuenta.
- Es un trabajo en equipo. Defina los ítems de trabajo, asígnelos entre los miembros del equipo, defina fechas de entrega y revisión interna. Discuta los resultados, observaciones y conclusiones. Priorice tareas y resultados a incluir.
- Empiece a trabajar prontamente y discuta con el profesor su avance y resultados.
- Los datos pueden ser obtenidos de cualquier fuente, siempre y cuando se respeten los derechos de autor y se den los créditos necesarios. Algunas fuentes de datos para problemas particulares son:

1. SNAP Datasets: Stanford Large Network Dataset Collection [3].
2. SuiteSparse Matrix Collection[2]. Colección de conjuntos de datos adaptada a Matrix Market por la Universidad de la Florida a partir de SNAP Dataset[3].
3. SuiteSparse Matrix Collection (formalmente la colección de matrices dispersas de la Universidad de Florida[4]) Puede verse desde el repo [SuiteSparse Matrix Collection](#) del profesor [Tim Davis](#). Está es un conjunto ampliamente utilizado de puntos de referencia de matriz dispersa recopilados de una amplia gama de aplicaciones de grafos.
4. Network Repository[5] contiene varias redes y permite obtener una visualización de grafos de varios niveles para aprovechar las propiedades de los gráficos locales y globales, así como funciones y herramientas adicionales. El sitio de descargas los puede encontrar en [Network Datasets](#)[6]
5. PyTorch Geometric Datasets, que viene desde la perspectiva del [Geometric Deep Learning](#) apodado así por algunos investigadores[1]. Los datasets se encuentran en [PyTorch Geometric Datasets](#)

4. Condiciones de entrega

1. Envíe el link de Github del repositorio con los soportes respectivos. El link se debe enviar al correo del docente, no se recibirán repositorios que no enviaron el link al correo.
2. El proyecto se debe realizar en grupos de 4 a 5 personas. **No se recibirán proyectos en grupos de menos de 4 personas.**
3. Medio de entrega: correo electrónico sergioa.mora@javeriana.edu.co
4. Fecha y hora de entrega: viernes 20 de octubre 23:59.

Referencias

- [1] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges, 2021.
- [2] Tim Davis and University of Florida. SuiteSparse Matrix Collection adaptation from SNAP. <https://sparse.tamu.edu/SNAP>, June 2014.
- [3] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>, June 2014.
- [4] University of Florida. SuiteSparse: University of Florida Sparse Matrix Collection. <https://sparse.tamu.edu/>, June 2014.
- [5] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.
- [6] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *AAAI*, 2015.