

## 1. Descripción general

El objetivo de este proyecto es diseñar, prototipar y evaluar una solución basada en modelos de lenguaje natural considerando todas las etapas del proceso de MLOps. Para el proyecto se requiere:

1. Definir claramente el problema o pregunta de negocio y la justificación de por qué es relevante en su contexto.
2. Diseñar la solución basada en modelos de lenguaje para atender este problema (modelos de embedding, modelos entrenados para tareas específicas, finetuning, etc).
3. Identificar el conjunto de datos a emplear.
4. Implementar un prototipo de la solución diseñada.
5. Definir la evaluación de la solución. En este aspecto defina claramente:
  - a) Aspectos a evaluar de la solución.
  - b) Pruebas a realizar.
  - c) Métricas a evaluar.
  - d) ¿Por qué selecciona estos aspectos, pruebas y métricas?
  - e) ¿Qué limitaciones tienen estas pruebas y métricas?
  - f) ¿Cómo estas pruebas le permiten tomar decisiones sobre su solución?
6. Explorar los resultados de la solución, su alcance y limitaciones.
7. Responder a la pregunta: si hiciera el proyecto nuevamente, ¿qué habría hecho diferente?
8. Realizar una presentación de máximo 10 minutos respondiendo a los ítems anteriores y los resultados clave del proyecto.

## 2. Entregables

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

1. **(45 puntos)** Resultado 1: presentación de no más de 10 minutos con los resultados principales del proyecto donde se responda a cada uno de los ítems listados en la sección anterior.
2. **(45 puntos)** Resultado 2: código con el procesamiento de datos y modelos desarrollados que soporten los resultados reportados.

3. **(10 puntos)** Reporte de trabajo en equipo (max. 1 página): incluya un pequeño reporte de cómo se dividieron las tareas entre los miembros del equipo. Para cada miembro del equipo liste las tareas **específicas** desarrolladas en el **formato entregado** para este fin.
4. Soporte 1: enlace a los datos usados o el conjunto de datos.
5. Soporte 2: fuentes de procesamiento de datos y construcción y evaluación de modelos (archivos .py o .ipynb).

**Nota:** si bien el trabajo es en grupo (3 personas), la nota es individual, luego es necesario que cada miembro del equipo demuestre su contribución al proyecto.

### **3. Recomendaciones**

1. Sea conciso y eficiente en la presentación, asegúrese de dejar tiempo para resolver cada uno de los ítems listados en la Sección 1.
2. 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.
3. Empiece a trabajar prontamente y discuta con los profesores y tutores su avance y resultados.
4. Los datos pueden ser obtenidos de cualquier fuente, siempre y cuando se respeten los derechos de autor y se den los créditos necesarios.

### **Condiciones de entrega**

1. Envíe 1 (un) archivo .zip con los entregables definidos arriba.
2. El proyecto se debe realizar en grupos de 3 personas (los mismos grupos de los talleres) .
3. Medio de entrega: Bloque Neón.
4. Fecha y hora de entrega: viernes 27 de junio 23:59.
5. Las sustentaciones se realizarán en espacios de 20 minutos, la semana del 1 al 4 de julio, via Zoom.

**Fecha de entrega: viernes 27 de junio, hasta las 23:59.**