**Guía de trabajo de módulo y avances**

**Problema:** Para la entrega de este modulo cada grupo debe buscar un problema de analítica en salud y sus respectivos datos. Adicional a [kaggle](https://www.kaggle.com/datasets) pueden utilizar estos enlaces para buscar la descripción del problema y datos:

* UCI ML Repository: creado en 1987 por David Aha, +600 datasets donde filtrar por tipo de problema, tipo de datos, de características, etc. <https://buff.ly/3E5H3Wa>
* OMS: de las Naciones Unidas, datos de la salud de todo el mundo <https://buff.ly/3Wmt05X>
* Zdataset: selección según el problema que queremos practicar <https://buff.ly/3HSq5Ob>
* Data Is Plural: newsletter semanal con todo tipo de datos actuales de diversas materias <https://buff.ly/3vQUS55>
* Data .gov: base de datos publica del gobierno de EEUU donde podemos encontrar +300.000 conjuntos de datos <https://buff.ly/2C4Z0p9>

Tengan presente que la descripción del problema de negocio debe ser muy claro, generalmente en los sitios que se publican datos no se describe mucho de los problemas, si este el caso, la descripción del problema debe ser complementada por el grupo, no se pueden limitar sólo a decir la información del problema que esté en la página.

**El trabajo que se entrega a final del módulo debe contener:**

1. Un proyecto en GitHub con el código generado para la solución del problema, organizado, documentado y dónde se refleje la **contribución de todos los integrantes del grupo**.
2. Un informe con el diseño de la solución propuesta, los análisis realizados, y las propuestas con base en los análisis (El informe debe tener máximo 5 páginas letra tamaño 12). El informe es para escribir, principalmente análisis, se pueden hacer descripciones de lo que se hizo pero de manera muy breve y solo lo que sea necesario de explicar en el informe.

Se propone que los siguientes elementos sean analizados en el informe:

* 1. Diseño de solución propuesto
  2. Limpieza y transformación
  3. Análisis exploratorio
  4. Selección de algoritmos y técnicas de modelado
  5. Selección de variables
  6. Comparación y selección de técnicas
  7. Afinamiento de hiperparámetros
  8. Evaluación y análisis del modelo
  9. Despliegue del modelo

**Nota 1:** Dependiendo de la solución que el equipo proponga, que es libre, puede que se omitan pasos o se incluyan adicionales, en este caso, se debe analizar por qué no se incluyó el paso o por qué se adicionaron.

**Criterios de evaluación serán los siguientes:**

1. Descripción del problema 10%
2. Pertinencia y creatividad en la propuesta de solución 15%
3. Estructura (funciones, varios archivos, uso de sql), documentación y Justificación del código 20%
4. **Análisis de resultados** 35%.
5. Redacción adecuada del informe (incluye ortografía) 20%

**Nota2:** Se puede solicitar a cualquier grupo la sustentación de lo que se realizó en el trabajo y si el equipo no puede sustentarlo la nota será cero independientemente del producto entregado.

**Presentación de avances:**

* Cada grupo presentará avances en **una sola sesión por modulo**.
* Los avances presentados son de temas diferentes en cada sesión de la siguiente manera:

**Presentación de Avance 1:** Problema y diseño de la solución, selección de variables, selección de algoritmos.

**Presentación de Avance 2:** Problema y diseño de la solución, comparación de algoritmos y afinamiento.

* Cada día de avances se seleccionará aleatoriamente los grupos que presentarán. Todos los grupos que no hayan presentado hasta la fecha deben estar preparados para exponer.
* La presentación es de máximo 10 minutos.
* Se puede presentar en notebook o archivo de códigos.
* La presentación de avances también es para resolver dudas en el avance del proyecto.

**Los criterios de evaluación son:**

* Claridad y orden en la presentación (presentar lento y escoger bien los contenidos a mostrar)
* Grado de avance adecuado
* Compresión adecuada del problema
* Propuestas innovadoras en los desarrollos
* No tener errores conceptuales (mal uso o mala interpretación de conceptos o herramientas)