



Ejercicio Formativo 4 Capítulo 3

Aspectos generales

- **Objetivos:** Aplicar los contenidos de modelos predictivos.
- **Lugar de entrega:** lunes 22 de septiembre a las 17:30 hrs. en repositorio privado.
- **Formato de entrega:** archivo Python Notebook (**C3E4.ipynb**) con el avance logrado para el ejercicio. El archivo debe estar ubicado en la carpeta **C3**. Utilice múltiples celdas de texto y código para facilitar el trabajo del cuerpo docente.
- **ULTRA IMPORTANTE:** todas las celdas utilizadas deben estar ejecutadas al momento de entregar el ejercicio, de modo que las salidas generadas sean visibles. En caso de no cumplir esto, su entrega no será considerada como validación del ticket de salida.

Descripción del problema

En este laboratorio trabajará con un conjunto de datos reales de accidentes de tránsito en el Reino Unido durante un año, compuesto por información tanto de los accidentes como de los vehículos involucrados. Los datos están divididos en dos archivos:

- **accidentes.csv:** cada fila de este archivo representa un único accidente de tránsito, identificado por la columna **Accident_Index**. El resto de las columnas describe diversos atributos y características del accidente.
- **vehiculos.csv:** cada fila de este archivo representa la participación de un vehículo único en un accidente en particular, describiendo variadas propiedades del vehículo y los pasajeros en sus columnas.

La relación entre cada vehículo y el accidente particular en que participó, está capturada mediante la columna `Accident_Index`.

Dado el tamaño de los datos, estos se encontrarán comprimidos en formato `.zip`. Para utilizarlos en Colab, luego de subirlos, debe ejecutar el comando `!unzip datos.zip` en una celda de código. Este comando solo debe ejecutarlo una vez, a menos que reinicie completamente Colab.

El tamaño y complejidad de la base hacen necesario integrar ambas fuentes y tomar decisiones de preparación y modelado que permitan extraer conclusiones útiles. El desafío no se limita a predecir, sino a usar los modelos como herramienta para entender mejor el fenómeno y generar conocimiento aplicable.

Misiones

1. **Factores de severidad:** diseñe un experimento que compare distintas formas de representar los datos y evalúe cómo estas influyen en la capacidad de modelar la severidad de los accidentes. Analice qué transformaciones y variables aportan más a la estabilidad y explicación de los resultados.
2. **Decisiones operativas:** construya un sistema que entregue **probabilidades calibradas** sobre un resultado operativo relevante, como por ejemplo, la respuesta policial, es decir, si un policía fue o no al lugar del accidente. Traduzca estas probabilidades en políticas de acción bajo restricciones de recursos. Evalúe la utilidad de estas políticas y discuta compromisos entre cobertura y riesgo.
3. **Índice de riesgo vial:** genere un índice espacio-temporal de riesgo a partir de los datos y modelos, construyendo visualizaciones y métricas que permitan entender dónde y cuándo se concentran los mayores riesgos. Explore escenarios alternativos modificando condiciones relevantes y analice cómo cambia el riesgo estimado.