


Taller	Clasificación	
Asignatura	Computación Científica	
Caso	1	
Semestre	2025 - II	
Objetivo	Implementar la metodología ASUM – DM que permita diseñar modelos de clasificación a través de un caso particular de negocio.	

Ustedes hacen parte de un equipo de datos que trabajan en una clínica endocrinológica, en la que cuentan con la información de 253.680 individuos. Cuentan con los datos de auscultación de los pacientes obtenidos por medio de las consultas médicas. Ante esto, se les pide construir un modelo de clasificación que sea capaz de determinar si un paciente es diabético, prediabético o no diabético. El modelo debería contar con un desempeño superior al 85% en accuracy y con la mejor matriz de confusión y f1 score posible.

El conjunto de datos cuenta con el siguiente diccionario:

Variable	Descripción
Diabetes_012	Resultado del análisis clínico: - 0: sin diabetes - 1: prediabetes - 2: diabetes
HighBP	Tiene o no presión arterial alta: - 0: No - 1: Si
HighChol	Tiene o no colesterol alto: - 0: No - 1: Si
CholCheck	Chequeo del colesterol en los últimos 5 años: - 0: No - 1: Si
BMI	Índice de masa corporal
Smoker	Ha fumado al menos 100 cigarrillos en su vida: - 0: No - 1: Si
Stroke	Alguna vez ha tenido un ataque cardíaco: - 0: No - 1: Si
HeartDiseaseorAttack	Enfermedad coronaria: - 0: No - 1: Si
PhysActivity	Actividad física en los últimos 30 días: - 0: No - 1: Si

Fruits	Consume frutas una o más veces al día: - 0: No - 1: Si
Veggies	Consume vegetales una o más veces al día: - 0: No - 1: Si
HvyAlcoholConsump	Consume más de 14 bebidas alcohólicas por semana: - 0: No - 1: Si
AnyHealthcare	Alguna condición especial de salud: - 0: No - 1: Si
NoDocbcCost	Ha visitado al médico en el último año: - 0: No - 1: Si
GenHlth	Percepción general de salud: - 1: Excelente - 2: Muy buena - 3: Buena - 4: Aceptable - 5: Mala
MentHlth	De 0 a 100 calificación de la salud mental
PhysHlth	De 0 a 100 calificación de la salud física
DiffWalk	Dificultad para caminar: - 0: No - 1: Si
Sex	Género: - 0: Mujer - 1: Hombre
Age	Rango de Edad: https://www.icpsr.umich.edu/web/NAHDAP/studies/34085/datasets/0001/variables/AGEG5YR?archive=NAHDAP
Education	Nivel de educación: https://www.icpsr.umich.edu/web/NAHDAP/studies/34085/datasets/0001/variables/EDUCA?archive=nahdap
Income	Nivel de ingresos: https://www.icpsr.umich.edu/web/NAHDAP/studies/34085/datasets/0001/variables/INCOME2?archive=nahdap

Implemente los pasos metodológicos de ASUM – DM que le permita desarrollar de forma efectiva el proyecto de datos.

Construya todos los modelos discutidos en clase y elija el mejor modelo.

Condiciones de entrega:

- Presentación de 10 minutos.