

Archivos suplementarios

Apéndice 1

Cuadro 1. Breve descripción de las variables del conjunto de datos en estudio

Concepto	Variable	Tipo de variable	Naturaleza	Breve descripción
C1	Edad	Sociodemográfica	Numérica	Tiempo que transcurre desde el nacimiento
C2	Sexo	Sociodemográfica	Binaria	Características biológicas que diferencian en femenino o masculino
C3	Poliuria	Clínica	Binaria	Aumento del volumen de orina excretada
C4	Polidipsia	Clínica	Binaria	Aumento excesivo de la sed
C5	Pérdida repentina de peso	Clínica	Binaria	Disminución involuntaria del peso corporal
C6	Debilidad	Clínica	Binaria	Reducción de la fuerza física o muscular
C7	Polifagia	Clínica	Binaria	Aumento del consumo de alimentos por exceso de hambre
C8	Candidiasis genital	Clínica	Binaria	Infección fúngica en la zona genital
C9	Visión borrosa	Clínica	Binaria	Disminución de la agudeza visual
C10	Prurito	Clínica	Binaria	Picor o irritación de la piel que induce la necesidad de rascar
C11	Irritabilidad	Clínica	Binaria	Estado emocional relacionado con temperamento explosivo o fácil enojo
C12	Retraso en la cicatrización	Clínica	Binaria	Disminución de los mecanismos naturales para la regeneración de los tejidos
C13	Parálisis parcial	Clínica	Binaria	Pérdida parcial del movimiento muscular voluntario
C14	Rigidez muscular	Clínica	Binaria	Tensión muscular, que puede acompañarse de espasmos y dolor
C15	Alopecia	Clínica	Binaria	Caída anormal del cabello
C16	Obesidad	Clínica	Binaria	Acumulación excesiva de grasa corporal
C17	Clase	Diagnóstico	Binaria	Presencia o ausencia de diabetes

Cuadro 2. Representación de un mapa cognitivo difuso mediante una matriz de siete conceptos y sus relaciones

	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇
C ₁	0	0	0	0	0	0	W ₁₇
C ₂	W ₂₁	0	W ₂₃	0	0	0	W ₂₇
C ₃	0	0	0	0	W ₃₅	0	0
C ₄	0	0	W ₄₃	0	W ₄₅	0	0
C ₅	0	0	0	0	0	W ₅₆	0
C ₆	0	0	0	0	0	0	W ₆₇
C ₇	0	0	W ₇₃	0	0	W ₇₆	0

Cuadro 3. Hiperparámetros para el modelo de aprendizaje mapas cognitivos difusos

Técnica	Hiperparámetros	Opciones
Mapa cognitivo difuso	Población inicial	50, 100, 150, 200, 250, 300
	Función de activación	Sigmoide, tanh
	Función de inferencia	Kosko-estándar, Kosko-modificado, rescaled

Apéndice 2. Normalización min-máx

Se utilizó la normalización mín-máx para la edad, que fue la única variable numérica del conjunto de datos. Este procedimiento ayudó a mantener las variables en el mismo rango y, también, a optimizar el tiempo en la calibración de los modelos. Con la siguiente ecuación, se expresa la normalización mín-máx:

$$e_{\text{norm}} = \frac{e_i - e_{\min}}{e_{\max} - e_{\min}}$$

En donde: (e_{norm}) es la edad normalizada, (e_i) es la edad del individuo, ($e_{\min}$) es la edad mínima y ($e_{\max}$) es la edad máxima.

Apéndice 3. Explicación de un mapa cognitivo difuso

En la figura 2 (apéndice 4), se observa una estructura gráfica con siete nodos que representa un mapa cognitivo difuso simple. Los nodos son los conceptos (C) y la relación entre cada concepto (influencias) está representada por un peso (W), que se expresa por medio de una flecha dirigida desde un nodo de origen a un nodo de destino. El subíndice en la W de cada flecha indica la dirección de esa relación, es decir, W₁₇ representa la relación entre

el concepto 1 (C_1) y el concepto 7 (C_7). En el ámbito clínico, los nodos o conceptos suelen ser, por ejemplo, factores de riesgo asociados con una enfermedad, los síntomas de una enfermedad o las pruebas de laboratorio (1).

Otra forma de representar las relaciones entre los conceptos del sistema en estudio es mediante una matriz cuadrada, que almacena la información de las influencias o pesos entre los conceptos de un mapa cognitivo difuso (2). Como ejemplo, en el cuadro 2 del apéndice 1, se representa la matriz de pesos para el mapa cognitivo difuso de la figura 2 del apéndice 4.

Apéndice 4. Figuras adicionales

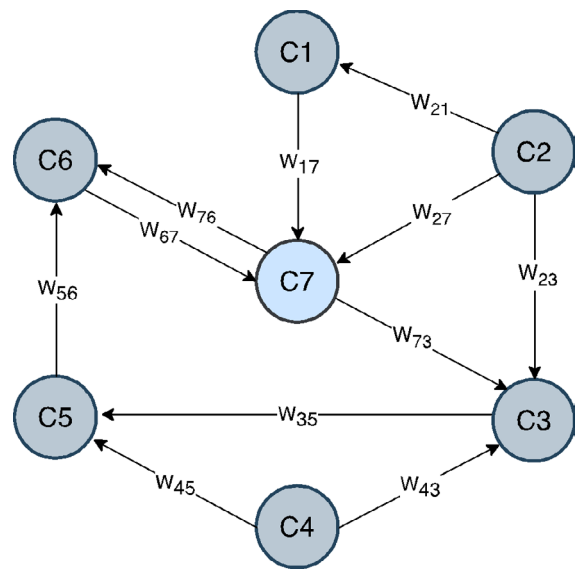


Figura 1. Representación gráfica de un mapa cognitivo difuso simple, con siete conceptos (nodos) y flechas que indican las relaciones entre ellos

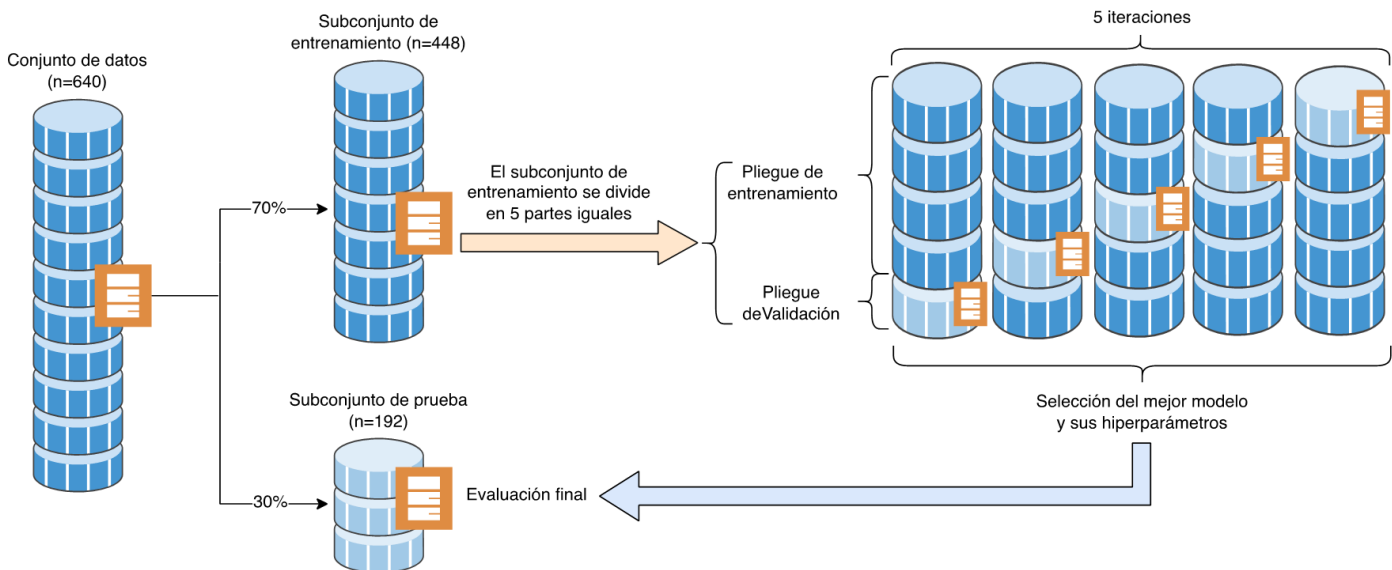


Figura 2. Esquema general de la validación cruzada de cinco repeticiones

Apéndice 5. Descripción del algoritmo *Particle Swarm Optimization* (PSO)

La optimización de los mapas cognitivos difusos con el algoritmo PSO (optimización por enjambre de partículas) se logra mediante el empleo de dos ecuaciones (3). La primera de ellas permite modelar la actualización de la velocidad de la partícula (en nuestro caso, la partícula es el mapa cognitivo difuso):

$$v_i(t+1) = v_i(t) + s_1 r_1 \cdot (W_1^{best} W_1(t)) + s_2 r_2 \cdot (W_1^{gbest} W_1(t))$$

En donde: v_i corresponde a la velocidad de la partícula; s_1 es el coeficiente cognitivo, mediante el cual la partícula i busca ubicarse en la posición donde ha obtenido resultados más exitosos; s_2 es el componente social (comportamiento colectivo), por medio del cual la partícula i busca alinearse a la mejor posición hallada por el enjambre hasta el momento; por su parte, r_1 y r_2 corresponden a valores aleatorios que siguen una distribución uniforme; las variables W_1^{best} y W_1^{gbest} corresponden a la mejor posición obtenida por una partícula en concreto y a la mejor posición obtenida por cualquier partícula del enjambre, respectivamente. La siguiente ecuación (4) permite modelar la posición actualizada de la partícula i (FCM), una vez actualizadas las velocidades, según la anterior ecuación:

$$W_i(t+1) = W_i(t) + v_i(t)$$

A partir de las anteriores operaciones, el algoritmo genera una matriz de pesos optimizada. El algoritmo PSO es eficaz, porque permite diseñar y entrenar modelos predictivos de mapas cognitivos difusos, empleando datos reales y artificiales, y superando los sesgos inducidos por la intervención humana durante el proceso (5).

Apéndice 6. Parámetros para evaluar el desempeño del modelo

El modelo desarrollado se evaluó según su capacidad para detectar la diabetes, mediante el empleo de los siguientes parámetros:

Exactitud: corresponde al porcentaje de ejemplos clasificados en forma correcta entre el número total de ejemplos clasificados. Entre mayor sea este porcentaje, mayor será el rendimiento total del modelo que se prueba.

$$\text{Exactitud} = \frac{VP + VN}{VP + FN + FP + VN}$$

De acuerdo con la fórmula, VP representa el valor de los verdaderos positivos; VN corresponde al valor de los verdaderos negativos; FN el valor de falsos negativos y FP el valor de falsos positivos.

Especificidad: es la medida de la capacidad que posee el modelo para clasificar los casos negativos frente a los verdaderamente negativos.

$$\text{Especificidad} = \frac{VN}{VN + FP}$$

Sensibilidad: es la medida de la capacidad que tiene el modelo para clasificar los casos positivos frente a los verdaderamente positivos.

$$\text{Sensibilidad} = \frac{VP}{VP + FN}$$

Referencias

1. Hoyos W, Aguilar J, Toro M. PRV-FCM: An extension of fuzzy cognitive maps for prescriptive modeling. *Expert Syst Appl.* 2023;23:1-15. <https://doi.org/10.1016/j.eswa.2023.120729>
2. Hoyos W, Aguilar J, Toro M. A clinical decision-support system for dengue based on fuzzy cognitive maps. *Health Care Manag Sci.* 2022;25:666-81. <https://doi.org/10.1007/s10729-022-09611-6>
3. Kennedy J, Eberhart, R. Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks.* Perth, WA, Australia, 1995, p. 1942-48. vol. 4. <https://doi.org/10.1109/ICNN.1995.488968>
4. Salmeron JL, Rahimi SA, Navali AM, Sadeghpour A. Medical diagnosis of rheumatoid arthritis using data driven PSO–FCM with scarce datasets. *Neurocomputing.* 2017;232:104-112. <https://doi.org/10.1016/j.neucom.2016.09.113>
5. Hoyos W, Aguilar J, Toro M. Federated learning approaches for fuzzy cognitive maps to support clinical decision-making in dengue. *Eng Appl Artif Intell.* 2023;123:1-15. <https://doi.org/10.1016/j.engappai.2023.106371>