



Análisis neutrosófico para el diagnóstico de la hipertensión arterial a partir de un sistema experto basado en casos

Freddy E. Torres Cordero¹, Elizabeth de la C. Torres Rodríguez², José Manuel Chile Labrador³, Daylen de los Reyes Valdés⁴, Neilys González Benítez⁵

¹Especialista en Medicina Interna y Especialista en Cuidados Intensivos del Adulto Policlínica Docente Universitaria Pedro Borrás Astorga, Provincia, Pinar del Río, Cuba. Código Postal: 20100, E-Mail: freddyelie@nauta.cu ²Estudiante de 4to año de Medicina. Universidad Ernesto Che Guevara de la Serna, Pinar del Río, Cuba. Código Postal: 20100, E-Mail:

ivisanabel@nauta.cu

^{3 y 4} Estudiante de 4to año de Medicina. Universidad Ernesto Che Guevara de la Serna, Pinar del Río, Cuba. Código Postal: 20100, E-Mail: anaamalys@nauta.cu

⁵Doctora en Ciencias. Especialista en Meteorología, Centro Meteorológico Provincial de Pinar del Río, Cuba. Código Postal: 20100, E-Mail: neilysgonzalezbenitez@gmail.com

Resumen: El artículo tiene como objetivo diagnosticar la hipertensión arterial (HTA) a partir de un sistema experto basado en casos, realizado en el marco de un estudio para conocer la incidencia de la enfermedad en esta población de la provincia de Pinar del Río, Cuba. La muestra de la población estudiada estaba formada por 455 hombres y 394 mujeres, entre 18 y 78 años de edad. Los individuos fueron clasificados en normotensos (personas con presión arterial normal), prehipertensos (personas en riesgo de padecer HTA) e hipertensos. Se realizó un procesamiento estadístico en el que se emplearon técnicas multivariadas como el Análisis Discriminante y la Regresión Logística cuyos resultados, junto a los del Método del Triángulo de Füller, fueron utilizados en el sistema para jerarquizar los factores de riesgo de la HTA y obtener el grado de importancia (peso) de estos. Por medio de la técnica de segmentación CHAID se pudo reducir las comparaciones entre los casos haciendo más eficiente este proceso. La obtención de las funciones de comparación por rasgos para las variables continuas se obtuvo de la aplicación conjunta de un análisis de varianza (ANOVA) y el método TwoStep Cluster Analysis. La adaptación de la solución de los casos más semejantes se realizó con la aplicación de un modelo de recomendación neutrosófica que permite obtener las semejanzas entre los casos analizados y los casos nuevos a través del cálculo de la distancia ideal. Obtenidos los resultados se valida el sistema experto y se comprueba su efectividad en el diagnóstico de la hipertensión arterial ante los factores de riesgo que con frecuencia se presenta en la población.

Palabras Claves: Sistema experto basado en casos, hipertensión arterial, técnicas estadísticas, Neutrosófica, modelo de recomendación.

1. Introducción

La hipertensión arterial (HTA) es una enfermedad crónica caracterizada por un incremento continuo de la presión sanguínea en las arterias, también conocida como presión arterial. La presión arterial (PA) se clasifica en base a dos tipos de medidas: la sistólica y la diastólica, las cuales son expresadas, en ese orden, en forma de razón matemática, como, por ejemplo, 120/80 y se miden en milímetros de Mercurio (mmHg). Cuando la medida de la presión sistólica o diastólica está por encima de los valores aceptados como normales, el individuo puede clasificarse en la categoría de prehipertenso o hipertenso. La clasificación de la PA para adultos puede verse de forma detallada en la tabla 1. [1].

Tabla 1. Clasificación de la presión arterial para adultos

Categoría	PA sistólica (mmHg)	PA diastólica (mmHg)
-----------	---------------------	----------------------

Normotenso (con cifras de PA normales)	Menos de 120	Menos de 80
<i>Freddy E. Torres Cordero, Elizabeth de la C. Torres Rodríguez, José Manuel Chile Labrador, Daylen de los Reyes Valdés, Neily González Benítez. Análisis neutrosófico para el diagnóstico de la hipertensión arterial a partir de un sistema experto basado en casos.</i>		
Prehipertenso (en riesgo de padecer HTA)	120-139	80-89
Hipertenso		
Grado I	140-159	90-99
Grado II	160-179	100-109
Grado III	180 y más	120 y más

La HTA se diagnostica fundamentalmente cuando en tres ocasiones diferentes se observan cifras mayores o iguales a 140/90 mmHg [2]. Sin embargo, debido a la naturaleza multicausal de esta enfermedad, su diagnóstico no debe limitarse simplemente a la toma de la PA sistólica y diastólica, sino que deben tomarse en cuenta, el comportamiento de un grupo de factores de riesgo de los cuales se hará mención más adelante. La HTA, además de representar por sí misma una enfermedad, constituye un factor de riesgo para enfermedades cardiovasculares, cerebrovasculares y renales. Se considera, por tanto, uno de los problemas de salud pública más importantes y su prevalencia a nivel mundial se estima entre el 21 y el 30% de la población adulta general [3-5].

En encuestas realizadas en Cuba a nivel nacional esta prevalencia se ha estimado en alrededor de un 30% para zonas urbanas [1, 6]. La alta incidencia de la HTA en la población cubana se ha podido corroborar de manera particular en la provincia de Pinar del Río, Cuba en la que desde hace varios años se viene realizando una investigación con este propósito por un comité multidisciplinario de expertos. Tomando como base esta investigación previa, teniendo en cuenta la comprobada alta prevalencia de la HTA en la población y considerando además la diversidad de factores de riesgo a tener en cuenta en esta enfermedad, se comenzó a construir una herramienta de apoyo al médico comunitario con el objetivo de asistirlo en el diagnóstico de esta patología.

1.1. Sistema experto con razonamiento basado en casos

El uso de herramientas de Inteligencia Artificial (IA) para la toma de decisiones en el campo de la salud se encuentra muy generalizado actualmente. Esto se debe a que en este tipo de dominio del conocimiento la información que se maneja resulta difícil de formalizar por el alto grado de subjetividad asociado. Dentro de las herramientas de IA más utilizadas están las tecnologías basadas en el conocimiento y entre estas pueden mencionarse los Sistemas Basados en Reglas, las Redes Neuronales Artificiales, los Sistemas de Inferencia Borrosos y los Sistemas Basados en Casos [7-10].

Debido a las características del problema que se presenta y por el hecho de disponer de una base de casos diagnosticados por expertos médicos, es que se decide aplicar la técnica del razonamiento basado en casos pues esta apoya sus predicciones en ejemplos (casos) y en el que cada nuevo individuo a diagnosticar se compara con la base de referencia. En el presente trabajo se muestran los fundamentos de un Sistema Experto Basado en Casos para el diagnóstico de la HTA, el cual se basa en el uso de técnicas estadísticas [11-12].

Se utiliza para la adaptación de la solución de los casos más semejantes un modelo de recomendación neutrosófica que permite obtener las semejanzas entre los casos analizados y los casos nuevos a través del cálculo de la distancia ideal. Los modelos de recomendación son útiles en el proceso de toma de decisiones ya que proporcionan al usuario un conjunto de opciones que se espera satisfagan sus expectativas [13][1].

2. Materiales y métodos

La muestra de casos estudiados por el comité de expertos estuvo formada por 455 hombres y 394 mujeres, entre 18 y 78 años de edad, de la provincia de Pinar del Río, Cuba. Dicho comité confeccionó y aplicó un cuestionario (historia clínica) a la muestra estudiada, el cual incluyó un grupo de variables que se describen en la tabla 2.

Tabla 2. Descripción de las variables consideradas por el comité de expertos en el estudio

Variable	Símbolo	Descripción
Edad	edad	

Sexo	sexo	Femenino – Masculino
Raza	raza	Negra – Mestiza – Blanca
Índice de masa corporal	IMC	$IMC = \text{Peso(kg)}/\text{Talla(m)}^2$
Antecedentes patológicos personales	APP	asma (Sí - No), angina (Sí - No), infarto (Sí - No), insuficiencia cardíaca (Sí - No), enfermedad cerebrovascular (Sí - No) enfermedad renal (Sí - No), diabetes (Sí - No), dislipidemia (Sí - No), insuficiencia arterial (Sí - No).
Antecedentes patológicos familiares de HTA	APF-HTA	madre, padre, abuelos, tíos
Hábito de fumar	fuma	Sí – No
Ingestión de alcohol	alcohol	Sí – No
PA sistólica basal	PAS	-
PA diastólica basal	PAD	-
PA sistólica 1er min	PAS1	-
PA diastólica 1er min	PAD1	-
PA sistólica 2do min	PAS2	-
PA diastólica 2do min	PAD2	-
PA media 2do min	PAM	$PAM = (PAS2 - PAD2)/3 + PAD2$
Ácido úrico	acido_urico	-
Creatinina	Creatinina	-
Glucemia	glucemia	-
Colesterol Total	ColTot	-
Colesterol HDL	ColHDL	-
Colesterol LDL	ColLDL	-
Triglicéridos	trigliceridos	-
Estrés psicológico	estrés	Bajo – Medio - Alto

Las cifras de PA se tomaron en el marco de una prueba denominada Prueba del Peso Sostenido (PPS). La PAS y la PAD se obtuvieron en estado de reposo antes de comenzar la prueba. La PPS es una variante de una prueba denominada “handgrip”, que fundamenta su principio en analizar la reactividad vascular frente a un ejercicio físico isométrico y demostró ser de gran utilidad como método diagnóstico para estudios masivos de HTA debido a su sencillez y alto valor predictivo [14]. [2]Analizando los datos recogidos en la historia clínica y el criterio de los expertos, los individuos fueron clasificados en normotensos, pre hipertensos e hipertensos.

En la medida de la PA se utilizó el esfigmomanómetro digital OMRON modelo M6 COMFORT con el brazalete apropiado para cada caso. Los análisis complementarios fueron realizados en el laboratorio clínico del Policlínico Docente Pedro Borrás Astorga de la provincia de Pinar del Río, con un equipo automatizado marca HITACHI. [15- 17].

2.1. Análisis estadístico de la muestra

En el análisis de la muestra se utilizó el paquete estadístico SPSS 23. Entre los métodos de análisis de la estadística multivariada estuvo el Análisis Discriminante, la Regresión Logística y la técnica de segmentación CHAID. Las dos primeras permitieron validar el criterio de los expertos con respecto al grado de importancia o peso de cada variable para el diagnóstico, lo cual fue utilizado en la construcción del sistema experto; mientras que la técnica CHAID agrupó los sujetos de la base de casos, atendiendo a valores comunes en ciertas variables. En algunos casos esta agrupación clasificó directamente a los individuos, lo cual le aportó eficiencia y rapidez al programa. El criterio de comparación entre las variables continuas se obtuvo de la aplicación conjunta de un análisis de varianza (ANOVA) y el método *TwoStep Cluster Analysis*. [3]

2.2. Sistema basado en casos para el diagnóstico de la HTA

Un Sistema Basado en Casos para la Toma de Decisiones es un Sistema de Decisión S como el que se presenta en (1), y se define en términos de un par $(U, X \sqcup Y)$ donde U es un conjunto finito no vacío de objetos, eventos llamados casos, mientras X y Y son dos conjuntos finitos, no vacíos, de atributos o propiedades llamados rasgos predictores y objetivos respectivamente.[4] [5]

$$S = (U, X \cup Y) \quad (1)$$

Cada rasgo predictor $x_i \in X$ puede ser considerado una función, como la que aparece en (2), que mapea elementos de U en el conjunto $M \sqcup x_{i1}, x_{i2}, \dots, x_{ini} \sqcup$, el cual se denomina conjunto de valores del rasgo predictor X_i . $x_i: U \rightarrow M_i$

$$(2)$$

De manera análoga cada rasgo objetivo $y_j \in Y$ puede ser considerado una función que mapea elementos de U en el conjunto $N_j = \{y_{j1}, y_{j2}, \dots, y_{j\theta_j}\}$, que se denomina conjunto de valores del rasgo predictor y_j . Función que se representa a través de la expresión 3.

$$y_j: U \rightarrow N_j \quad (3)$$

Un Sistema Basado en Casos, está compuesto por:

- ↳ Una base de conocimientos o base de casos
- ↳ Módulo de recuperación de casos
- ↳ Módulo de adaptación de las soluciones

↳ Base de casos

La base de casos contiene las experiencias o casos con los cuales el sistema realiza sus inferencias y puede ser representada a través de una tabla, cuyas columnas son etiquetadas por variables o atributos que representan los rasgos predictores y objetivos, mientras que sus filas representan los casos. En la Tabla 3, se representa la estructura de la base de casos para el diagnóstico de la hipertensión arterial, el conjunto U está formado por los casos O_1, O_2, \dots, O_m el conjunto X por los atributos $\sqcup X_1, \dots, X_n \sqcup$, mientras que el conjunto Y lo integra el atributo Y_1 .

Tabla 3. Representación de la estructura de la base de casos

	Rasgos predictores			Rasgo objetivo
Casos	x_1	...	x_n	y_1
O_1	$x_1(O_1)$...	$x_n(O_1)$	$y_1(O_1)$
O_2	$x_1(O_2)$...	$x_n(O_2)$	$y_1(O_2)$
...
O_m	$x_1(O_m)$...	$x_n(O_m)$	$y_1(O_m)$

Para el diagnóstico de la hipertensión arterial, el conjunto U estaría integrado por los 849 sujetos de la muestra estudiada, el conjunto X lo integrarían los factores de riesgo (rasgos predictores) incluidos en la historia clínica y el conjunto Y estaría formado por el diagnóstico asignado a los casos, el cual, como se ha visto, puede ser normotenso, prehipertenso o hipertenso.

Para el diagnóstico de la hipertensión arterial a partir de los resultados almacenados en la Base de Casos confeccionada en el sistema experto se propone un modelo de recomendación basado en conocimiento utilizando el de números neutrosóficos de valor único permitiendo la utilización de variables lingüísticas [6]. El modelo de recomendación está compuesto por X que es el universo de discurso y un sistema de números neutrosóficos de valor único que se representan por A sobre X y a su vez ellos son un objeto que se representan como se muestra en la ecuación 4.

$$A = \{ \langle x, uA(x), rA(x), vA(x) \rangle : x \in X \} \quad (4)$$

Donde;

$uA(x): X \rightarrow [0,1]$, $rA(x): X \rightarrow [0,1]$ y $vA(x): X \rightarrow [0,1]$ con $0 \leq uA(x) + rA(x) + vA(x) \leq 3$ para todo $x \in X$.

El intervalo $uA(x)$, $rA(x)$ y $vA(x)$ representa las membrecías a verdadero, indeterminado y falso de x en A , respectivamente. Para obtener una respuesta rápida se utilizan los números neutrosóficos de valor único. Los números neutrosóficos de valor único, son de utilidad ya que permiten una mayor interpretabilidad de la información que se obtiene en términos lingüísticos útil para apoyar la toma de decisiones en general y en los sistemas de recomendación en particular. Estos números se expresan como se muestra en la ecuación 5.

$$A = (a, b, c) \quad (5)$$

Donde;

$$a, b, c \in [0,1], y + b + c \leq 3.$$

El modelo de recomendación que se propone está basado en conocimiento, ellos poseen la característica de realizar sugerencias haciendo inferencias sobre una situación determina y sus preferencias [18-21]. El enfoque basado en conocimiento se distingue en el sentido que usan conocimiento sobre cómo un objeto en particular puede satisfacer las necesidades del usuario, y por lo tanto tiene la capacidad de razonar sobre la relación entre una necesidad y la posible recomendación que se mostrará. [7]

El modelo de recomendación propuesto se basa en la construcción de perfiles de usuarios como una estructura de conocimiento que apoya la inferencia, la cual puede ser enriquecida con la utilización de expresiones que emplea lenguaje natural [20, 22]. El flujo de trabajo del modelo de recomendación para el diagnóstico de la hipertensión arterial a partir del sistema experto basado en casos se representa en la Figura 1, el mismo se fundamenta en la propuesta de Córdón [20, 23] para sistemas de recomendación basados en conocimiento permitiendo representar términos lingüísticos y la indeterminación mediante números neutrosóficos de valor único.

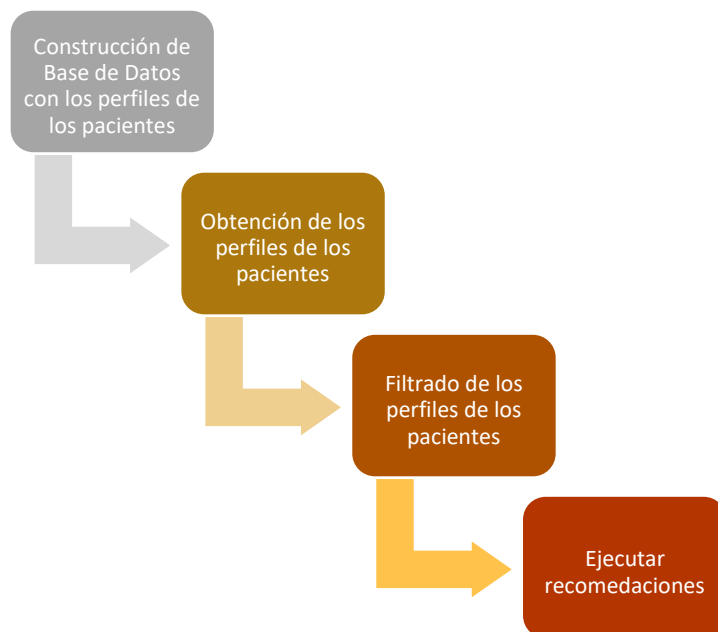


Figura 1. Flujo de trabajo del modelo de recomendación.

En la construcción de Base de Datos con los perfiles de los pacientes, cada una de los perfiles serán denotados como a_i y ellos serán descritas por un conjunto de síntomas, signos y factores de riesgo que conformarán el perfil de los pacientes. Ellos se representan matemáticamente como se muestra en la ecuación 6.

Para obtener la base de datos de los pacientes, el perfil de los pacientes es obtenido mediante números neutrosóficos de valor único [8]. De forma que $A^* = (A1^*, A2^*, \dots, An^*)$ es un vector de números neutrosóficos de valor único tal que $Aj^* = (aj^*, bj^*, cj^*)$ $j = (1, 2, \dots, n)$ y $Bi = (Bi1, Bi2, \dots, Bim)$ $(i = 1, 2, \dots, m)$ que son m vectores de n números de valores únicos tal que $Bij = (aij, bij, cij)$ $(i = 1, 2, \dots, m), (j = 1, 2, \dots, n)$ pudiéndose, entonces, calcular la distancia euclidiana como las Bi y A^* resultantes según refiere [24] a través de la ecuación 7.

$$d_i = \left(\frac{1}{3} \sum_{j=1}^n \left\{ (|a_{ij} - a_j^*|)^2 + (|b_{ij} - b_j^*|)^2 + (|c_{ij} - c_j^*|)^2 \right\} \right)^{\frac{1}{2}} \quad (i = 1, 2, \dots, m) \quad (7)$$

A partir de esta distancia euclidiana se puede definir una medida de similitud según [26] para el diagnóstico de la hipertensión arterial entre la base de casos del sistema experto y los casos que se modelan a través del sistema de recomendación basado en conocimiento utilizando el de números neutrosóficos de valor único permitiendo la utilización de variables lingüísticas [18]. En la medida en que la alternativa Ai sea más cercana al perfil de los pacientes (si) mayor será la similitud, permitiendo establecer un orden entre alternativas [27]. Los perfiles pueden ser obtenido de forma directa a partir de expertos, la ecuación para su obtención es la que se muestra en 8.

$$F_{aj} = \{v1^j, \dots, vk^j, \dots, vl^j\}, j = 1, \dots, n \quad (8)$$

Las valoraciones de las características de los perfiles, aj , compuesto por el conjunto de síntomas, signos y factores de riesgo se expresan utilizando la escala lingüística S , $v_k^j \in S$ donde $S = \{s1, \dots, sg\}$ es el conjunto de término lingüísticos definidos para evaluar las características ck utilizando los números neutrosóficos de valor único. Para esto los términos lingüísticos a emplear son definidos una vez descrito el conjunto de perfiles de los pacientes, a través de la ecuación 9, el resultado que se obtiene se guarda en la Base de Datos.

$$A = \{a1, \dots, aj, \dots, an\} \quad (9)$$

En la obtención del perfil de los pacientes se obtiene la información de cada paciente relacionada con los síntomas, signos y factores de riesgo, esta información se almacena en la Base de Datos y es en la Base de Datos el perfil de cada paciente que será comparado para recomendar, para obtener el perfil se utiliza la ecuación 10.

$$p_e = \{p_1^e, \dots, p_k^e, \dots, p_l^e\} \quad (10)$$

Dicho perfil estará integrado por un conjunto de atributos que son los síntomas, signos y factores de riesgo de los pacientes y el mismo se representa a través de la ecuación 11.

$$C_e = \{c1^e, \dots, c_{p_k}^e, \dots, c_l^e\} \quad (11)$$

Donde;

$$c_k^e \in S$$

El perfil de los pacientes puede ser obtenido mediante ejemplo o mediante el llamado enfoque conversacional o mediante ejemplos los cuales pueden ser adaptados según [9]. Después de obtenido el perfil de los pacientes, se realiza el filtrado de los perfiles de los pacientes que se realiza de acuerdo al perfil de cada paciente para encontrar cuáles son las más adecuadas para este lo cual es útil para calcular la similitud entre el

perfil de los pacientes, Pe y cada perfil aj registrado previamente en la base de datos. Para el cálculo de la similitud total se emplea la ecuación 12.

$$S_i = 1 - \left(\frac{1}{3} \sum_{j=1}^n \left((|a_j - a_i|)^2 + (|b_j - b_i|)^2 + (|c_j - c_i|)^2 \right)^{\frac{1}{2}} \right) \quad (12)$$

La función S calcula la similitud entre los valores de los atributos del perfil de usuario y la de los productos, aj [26]. Calculada la similitud entre el perfil de los pacientes en la Base de Datos y cada uno de los perfiles se ordenan de acuerdo a la similitud obtenida, representado por el vector de similitud que se representa en la ecuación 13.

$$D = (d1, \dots, dn) \quad (13)$$

Los mejores perfiles serán aquellos, que mejor satisfagan las necesidades del perfil de los pacientes, es decir con mayor similitud.

3. Resultados

El Análisis Discriminante jerarquizó los factores de riesgo en el siguiente orden: PAM, PAS2, PAD2, PAS1, PAD1, PAS, PAD, IMC, alcohol, dislipidemia y raza. Para diferenciar los prehipertensos de los normotensos, la Regresión Logística consideró el siguiente orden: sexo, fuma, PAS1, PAD1, PAS2, PAD2 y PAM. Sin embargo, para distinguir los prehipertensos de los hipertensos el orden fue: sexo, raza, enfermedad renal, PAS, PAD, PAS1, PAD1, PAS2, PAD2, PAM, IMC, acido_urico y ColHDL. La técnica de segmentación CHAID consideró la PAM y el IMC, como las variables más importantes para predecir la HTA. En la tabla 4 aparecen los factores de efectividad (cifras porcentuales) de estas técnicas.

Tabla 4. Factores de efectividad de las técnicas multivariadas

	Normotenso	Prehipertenso	Hipertenso
A Discriminante	85,7	93,8	85,1
R Logística	98,4	87,4	92,7
CHAID	99,8	92,7	95,5

Con la aplicación del Método del Triángulo de Füller, el cual es un método de comparaciones por parejas para tomar decisiones multicriterio a partir de expertos, se pudo obtener la importancia o peso de los rasgos predictores en la determinación del rasgo objetivo [29]. La tabla 5 muestra este resultado en el que las variables han sido ordenadas de forma descendente según el valor de su peso. Estas variables fueron las que finalmente se tomaron en cuenta en la construcción del sistema basado en casos.

Tabla 5. Variables obtenidas con el Método del Triángulo de Füller y sus pesos

<i>Variable</i>	<i>Peso</i>
PAM	0,69
PAS	0,63
PAD	0,63
PAS2	0,63
PAD2	0,63
PAS1	0,61
PAD1	0,58
IMC	0,53
Diabetes	0,53

<i>Variable</i>	<i>Peso</i>
Padres con HTA	0,53
Abuelos con HTA	0,53
Dislipidemia	0,50
ColTot	0,45
ColLDL	0,42
ColHDL	0,40
Glucemia	0,34
Triglicéridos	0,34
Estrés	0,32
Alcohol	0,29
Sexo	0,24
Raza	0,24
Fuma	0,21

La técnica de CHAID agrupó a los casos de forma que cada nodo, en la estructura jerárquica de árbol, contiene los casos que comparten el mismo rango de valores para una misma variable o grupo de variables. En la Figura 2 aparece una versión simplificada de la estructura del árbol obtenida. Este resultado se utilizó en la implementación del sistema, pues garantizaba que cada nuevo caso a diagnosticar, una vez ubicado en el nodo correspondiente del árbol, solo fuera comparado con los casos de ese nodo.

Los resultados obtenidos permitieron reducir notablemente el número de comparaciones a realizar y a su vez aporta una significativa eficiencia al programa al optimizar el número de comparaciones. En la tabla 6, obtenida del fichero de salida del SPSS, se muestran los resultados de clasificación de la técnica CHAID.

Tabla 6. Resultados de clasificación de la técnica de CHAID

<i>Observed</i>	<i>Predicted</i>			<i>Percent Correct</i>
	<i>Hipertenso</i>	<i>Prehipertenso</i>	<i>Normotenso</i>	
Hipertenso	210	9	1	95,5%
Prehipertenso	10	203	6	92,7%
Normotenso	0	1	409	99,8%
Overall Percentage	25,9%	25,1%	49,0%	98,6%

Clustering Method: CHAID
Dependent Variable: *morcodeg*

Las funciones de comparación por rasgos δ_i para las variables continuas, se construyeron con ayuda del paquete estadístico SPSS. En el caso de la variable PAD1 (PA diastólica 1er min) para el sexo femenino, el análisis de varianza (ANOVA) sugirió la formación de 4 grupos homogéneos. Esto aparece reflejado en la tabla 7 (salida

Freddy E. Torres Cordero, Elizabeth de la C. Torres Rodríguez, José Manuel Chile Labrador, Daylen de los Reyes Valdés, Neily González Benítez. Análisis neutrosófico para el diagnóstico de la hipertensión arterial a partir de un sistema experto basado en casos.

del SPSS). Con este resultado se aplicó el método *TwoStep Cluster Analysis*, resultado que se muestra en la tabla 8 (salida del SPSS) en el que se comprueba la formación de 4 clusters.

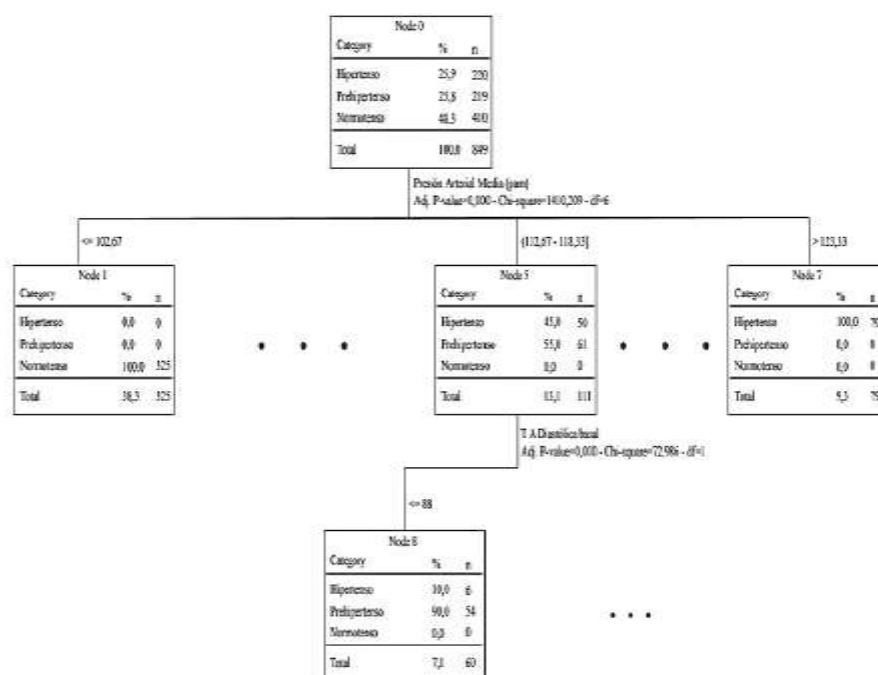


Figura 2 Versión simplificada de la estructura jerárquica de árbol obtenido por la técnica CHAID.

Tabla 7. Análisis de varianza en el caso de la variable PAD1 para el sexo femenino.

		<i>N</i>	% of Combined	% of Total
Cluster	1	137	30,1%	30,1%
	2	87	19,1%	19,1%
	3	144	31,0%	31,0%
	4	87	19,1%	19,1%
	Combined	455	100,0%	100,0%
	Total	455		100,0%

Tabla 8. Aplicación del método *TwoStep Cluster Analysis*

		<i>PA Diastólica (al 1er minuto)</i>	
		<i>Mean</i>	<i>Std. Deviation</i>
Cluster	1	89,10	1,971
	2	67,80	4,080
	3	79,81	1,044
	4	101,11	5,585
	Combined	84,39	11,412

De esta manera los dos valores a comparar fueron ubicados en los perfiles de los pacientes representados por vectores, de forma análoga se construyeron las restantes funciones para la comparación por rasgos y finalmente aplicar el modelo de recomendación propuesto. Los datos almacenados en los perfiles se representan como se muestra en la ecuación 14.

$$A = \{a_1, a_2, a_3, a_4, a_5\} \quad (14)$$

Descrito por el conjunto de atributos $C = \{c1, c2, c3, c4, \dots, c5\}$ es posible realizar la valoración a través

de la escala lingüística definida por [24]. Las valoraciones se almacenan por el sistema en la Base de Datos para luego realizar el cálculo de la similitud entre el perfil de los pacientes y los síntomas, signos y factores de riesgo de los pacientes que presentaron alteraciones de la HTA y que se almacenaron en la base de datos.

Para obtener la similitud previamente se ubicaron los nuevos pacientes con síntomas, signos y factores de riesgo como un nuevo caso en el nodo terminal correspondiente del árbol (Figura 1). Esta ubicación, permitió clasificar al caso directamente, dependiendo del nodo del árbol del que se tratase. Los casos ubicados en el nodo 1 (son aquellos con PAM menor o igual a 102,67), fueron diagnosticados directamente como normotensos, mientras que los casos ubicados en el nodo 7 (aquellos con PAM mayor a 123,33) fueron clasificados como hipertensos.

Cuando el caso quedó ubicado en otro nodo, se utilizó el algoritmo de recuperación ya descrito anteriormente, pero comparando el nuevo caso sólo con los que se encuentran en el nodo donde fue ubicado. Si los casos nuevos se ubican en el nodo 8, la comparación se realizaría con los 60 casos que la técnica de CHAID ubicó en ese nodo. Para determinados los casos más semejantes se calculó la distancia ideal, obteniéndose los resultados que se muestran en la tabla 9, que es la similitud entre los síntomas, signos y factores de riesgo de los pacientes con alteraciones de HTA y los perfiles de los pacientes ya afectados.

Tabla 9. Similitud obtenida.

a1 (Hipertensos)	a2 (Pre hipertensos)	a3 (Normotensos)
0.83	0.77	0.41

Después de obtenidos los resultados en la fase de recomendación se recomendará aquellos perfiles de pacientes que más se acerquen al perfil de los pacientes previamente almacenado en la Base de Datos. En este caso se recomienda atender los pacientes Pre hipertensos debido a que son los más próximos a padecer la enfermedad, convirtiéndose en una enfermedad crónica. El ordenamiento de los perfiles basado en la comparación es $\{a1, a2, a3\}$.

Conclusiones

En este artículo se muestran los resultados de aplicar técnicas de inteligencia artificial, como el razonamiento basado en casos, para el diagnóstico de la HTA. Se comprobó la gran utilidad del empleo de técnicas estadísticas en la construcción del sistema experto pues estas contribuyeron a determinar las variables a considerar, así como las funciones de comparación entre los rasgos. La construcción de una herramienta dedicada al diagnóstico de la HTA contribuye a hacer más eficaz este proceso y asiste de manera significativa al médico de la atención primaria de salud.

Aunque todos estos resultados han sido aplicados al diagnóstico de la HTA, la metodología seguida aquí puede extenderse a cualquier dominio y emplearse donde sea factible el uso de un sistema experto.

Se utiliza la neutrosofía y en particular un sistema de recomendación para recomendar los casos atender, una vez clasificados los pacientes y almacenados los síntomas, signos y factores de riesgo en la Base de Datos previamente creada.

Se realiza la recomendación, destacándose que los pacientes que deben ser atendidos son los pre hipertensos, debido a que pueden contener la HTA de forma crónica sino se tienen en cuenta los síntomas, signos y factores de riesgo que con frecuencia poseen.

El sistema experto que lleva por nombre *Tensoft-Expert*, fue validado y alcanzó una efectividad del 97%, efectividad que puede mejorar en la medida que la Base de Casos crezca y se enriquezca.

Referencias

- [2] M. D. P. Caballero, A. D. Herrera, J. A. Guerra, A. V. Vioa. Hipertensión Arterial. Guía para la prevención, diagnóstico y tratamiento. J. Quesada (editor). 1a ed. Ed. Ciencias Médicas. La Habana, Cuba. 2008. pp. 9-22.
- [3] <http://www.drscope.com/cardiologia/pac/arterial.htm>. Consultado el 13 de septiembre de 2010.

Freddy E. Torres Cordero, Elizabeth de la C. Torres Rodríguez, José Manuel Chile Labrador, Daylen de los Reyes Valdés, Neily González Benítez. Análisis neutrosófico para el diagnóstico de la hipertensión arterial a partir de un sistema experto basado en casos.

-
- [4] R. D. Feldman, G. Y. Zou, M. K. Vandervoort. "A simplified approach to the treatment of uncomplicated hypertension. A cluster randomized controlled trial". *Hypertension*. Vol. 53. 2009. pp. 646-653.
 - [5] D. L. Clement, M. L. De Buyzere, D. A. De Bacquer. "Prognostic value of ambulatory blood-pressure recordings in patients with treated hypertension". *J Med*. Vol. 348. 2003. pp. 2407-2415.
 - [6] K. Bjorklund, L. Lind, B. Zethelius. "Prognostic significance of 24-h ambulatory blood pressure characteristics for cardiovascular morbidity in a population of elderly men". *Hypertension*. Vol. 22. 2004. pp. 1691-1697. [6] http://scielo.sld.cu/scielo.php?pid=S0864-21251996000200007&script=sci_arttext. Consultado el 13 de septiembre de 2010.
 - [7] J. L. Kolodner. "An Introduction to Case-Based Reasoning". *Artificial Intelligence Review*. Vol. 6. 1992. pp. 3-34.
 - [8] F. Mariel. "CBR: A categorized bibliography". *The Knowledge Engineering Review*. Vol. 9. 1994. pp. 1-36. [9] A. Aamodt, E. Plaza. "CBR: foundational issues, methodological variations and systems approach". *AI Communications*. Vol. 7. 1996. pp. 1-22.
 - [10] W. Dubitzky. *Knowledge Integration in Case-Based Reasoning: A Concept-Centred Approach*. Thesis submitted an application for the degree of Doctor of Philosophy. Faculty of Informatics, University of Ulster, The United Kingdom. 1997. pp. 226-231.
 - [11] I. G. Martínez, R. E. Bello, Y. C. Bergolla. URS-HTA: Sistema de Razonamiento en condiciones de Incertidumbre para el diagnóstico de la HTA. II Simposio Internacional de Hipertensión Arterial. Ed. Feijoo. Villa Clara (Cuba). 2004. pp. 2-4.
 - [12] S. C. Rodríguez, G. C. Cardoso. Diagnóstico y detección de factores de riesgo de la HTA usando técnicas estadísticas. III Simposio Internacional de Hipertensión Arterial y I Taller sobre Riesgo Vascular. Ed. Feijoo. Villa Clara, Cuba. 2006. pp. 6-8.
 - [13] Leiva, J.L., et al., Realidad aumentada y sistemas de recomendación grupales: Una nueva perspectiva en sistemas de destinos turísticos. *Estudios y perspectivas en turismo*, 2014. 23(1): p. 40-59.
 - [14] A. P. de Armas, O. G. Blanco, H. C. Hernández, E. G. Rodríguez, Y. L. Carvajal, S. C. Rodríguez. Nuevos métodos para el pesquiasaje y el diagnóstico precoz de la hipertensión arterial esencial. III Simposio Internacional de Hipertensión Arterial y I Taller sobre Riesgo Vascular. Ed. Feijoo. Villa Clara, Cuba. 2006. pp. 2-4.
 - [15] P. Armario, R. Hernández del Rey, M. Martín. "Estrés, enfermedad cardiovascular e hipertensión arterial". *Medicina Clínica*. Vol. 119. 2002. pp. 23-29.
 - [16] R. S. Vasan, M. G. Larson, E. P. Leip, J. C. Evans, Ch. J. O'Donnell, W. B. Kannel, D. Levy. "Impact of HighNormal blood pressure on the risk of cardiovascular disease". *J Med*. Vol. 345. 2001. pp. 1291-1297.
 - [17] M. Benet, A. J. Yanes, L. J. González, J. J. Apolinaire, J. García del Pozo. "Criterios diagnósticos de la prueba del peso sostenido en la detección de pacientes con hipertensión arterial". *Medicina Clínica*. Vol. 116. 2001. pp. 645-649.
 - [18] Biswas, P., S. Pramanik, and B.C. Giri, TOPSIS method for multi-attribute group decision-making under single-valued neutrosophic environment. *Neural computing and Applications*, 2016. 27(3): p. 727-737.
 - [19] Dietmar Jannach, *Tutorial: Recommender Systems*, in *International Joint Conference on Artificial Intelligence Beijing, August 4, 2013*. 2013.
 - [20] Cordon, L.G.P., *Modelos de recomendación con falta de información. Aplicaciones al sector turístico*. 2008, Universidad de Jaén.
 - [21] Freire, J.B., et al., *Modelo de recomendación de productos basado en computación con palabras y operadores OWA [Aproduct recommendation model based on computing withword and OWA operators]*. *International Journal of Innovation and Applied Studies*, 2016. 16(1): p. 78.
 - [22] Herrera, F. and L. Martínez, *A 2-tuple fuzzy linguistic representation model for computing with words*. *Fuzzy Systems, IEEE Transactions on*, 2000. 8(6): p. 746-752.
 - [23] Arroyave, M.R.M., A.F. Estrada, and R.C. González, *Modelo de recomendación para la orientación vocacional basado en la computación con palabras [Recommendation models for vocational orientation based on computing with words]*. *International Journal of Innovation and Applied Studies*, 2016. 15(1): p. 80.
 - [24] Şahin, R. and M. Yiğider, A Multi-criteria neutrosophic group decision making metod based TOPSIS for supplier selection. *arXiv preprint arXiv:1412.5077*, 2014.
 - [25] Ye, J., Single-valued neutrosophic minimum spanning treeand its clustering method. *Journal of intelligent Systems*, 2014. 23(3): p. 311-324.
-

[28] Pérez-Teruel, K., M. Leyva-Vázquez, and V. Estrada-Sentí, *Mental models consensus process using fuzzy*

cognitivemaps and computing with words. Ingeniería y Universidad, 2015. **19**(1): p. 173-188.

- [29] Vázquez, M.Y.L., et al., *Modelo para el análisis de escenarios basados en mapas cognitivos difusos: estudio de caso en software biomédico*. Ingeniería y Universidad: Engineering for Development, 2013. **17**(2): p. 375-390. [28] Pérez, L., *Modelo de recomendación con falta de información*. Aplicaciones al sector turístico. 2008, Tesis doctoral. Universidad de Jaén.
- [30] M. Tabucanon. Multiple criteria decision making in industry. Ed. Elsevier. Amsterdam. 1988. pp. 233-299.
- [31] Hernández, N.B., et al., *Validation of the pedagogical strategy for the formation of the competence entrepreneurship in high education through the use of neutrosophic logic and Iadov technique*. Neutrosophic Sets & Systems, 2018. **23**.
- [32] Ricardo, J.E. and I.R.S. Vázquez, *LA EDUCACIÓN SEXUAL PARA PADRES DE NIÑOS CON RETRASO MENTAL, UNA VÍA PARA SU CONSOLIDACIÓN*. Revista Magazine de las Ciencias. ISSN 2528-8091, 2019. **4**(3): p. 137-144.
- [33] Hernandez, N.B., M.B.R. Cueva, and B.N.M. Roca, *Prospective analysis of public management scenarios modeled by the Fuzzy Delphi method*. Neutrosophic Sets & Systems, 2019. **26**.
- [34] VÁZQUEZ, M.L., N.B. HERNANDEZ, and F. SMARANDACHE, *MÉTODOS MULTICRITERIOS PARA DETERMINACIÓN DE LA EFECTIVIDAD DE LA GESTIÓN PÚBLICA Y EL ANÁLISIS DE LA TRANSPARENCIA*. Infinite Study.
- [35] Ricardo, J.E., et al., *EL ASSESSMENT CENTER PARA LA EVALUACIÓN DE LAS COMPETENCIAS ADQUIRIDAS POR LOS ESTUDIANTES DE NIVEL SUPERIOR*. Investigación Operacional, 2019. **40**(5): p. 638-643.
- [36] Ricardo, J.E., et al., *Neutrosophic model to determine the degree of comprehension of higher education students in Ecuador*. Neutrosophic Sets and Systems, 2019: p. 54.
- [37] Hernández, N.B., et al., *LA TOMA DE DECISIONES EN LA INFORMÁTICA JURÍDICA BASADO EN EL USO DE LOS SISTEMAS EXPERTOS*. Investigación Operacional, 2019. **40**(1): p. 131-139.
- [38] Hernandez, N.B. and J.E. Ricardo, *Gestión empresarial y posmodernidad*. 2018: Infinite Study.
- [39] Batista Hernández, N. and N. Valcárcel Izquierdo, *Determinación de la prefactibilidad en la aplicación de una estrategia pedagógica para la formación de la competencia Emprender en la educación preuniversitaria como contribución a la formación integral del estudiante*. Dilemas Contemporáneos: Educación, Política y Valores, 2018.