

Árboles de decisión ID3 para el diagnóstico de apendicitis aguda en niños

Iván Sánchez Martínez¹, Leticia Flores Pulido¹, Alfredo Adán Pimentel²,
J. Federico Ramírez Cruz¹, María del Rocío Ochoa Montiel¹

¹ Universidad Autónoma de Tlaxcala, Facultad de Ciencias Básicas, Ingeniería y
Tecnología,
Apizaco, Tlaxcala, México

² Universidad Autónoma de Tlaxcala, Facultad de Ciencias de la Salud, Zacatelco,
Tlaxcala, México

{computoisz,aicitel.flores,ma.rocio.ochoa}@gmail.com,{aadanpimentel,
federico_ramirez}@yahoo.com.mx

Resumen. La apendicitis aguda es una enfermedad de difícil diagnóstico en ancianos y niños, en la mayoría de los casos se presenta un dolor en el ombligo que aumenta conforme al tiempo y se dirige hacia el cuadrante inferior derecho. Por su diversidad sintomatológica, existen propuestas de diagnóstico que se limitan a unos cuantos síntomas, signos y laboratorios. La apendicitis aguda por ser una de las enfermedades mortales y más comunes en el ser humano, ha sido abordada a nivel computacional para desarrollar modelos o conjuntos de reglas para diagnosticar de manera oportuna. En este trabajo se propone la creación de un árbol de decisión basado en arquitectura ID3 para la detección de apendicitis aguda en niños entre los 4 y 15 años en el Hospital General Regional “Lic. Emilio Sánchez Piedras” (Tzompantepec, Tlaxcala, México). Los datos para construir los árboles de decisión ID3 se recolectaron de manera prospectiva utilizando dos escalas de puntuación diagnóstica (escala de Alvarado y puntuación de apendicitis pediátrica), incluyendo algunas variables demográficas y de confirmación de la enfermedad.

Se recolectó un total de 41 casos los cuales fueron analizados y base para la construcción de dos árboles de decisión ID3. Los resultados muestran un buen desempeño en el diagnóstico de apendicitis, sin embargo es necesario recolectar una mayor cantidad de datos para maximizar el grado de precisión y factibilidad.

Palabras clave: Apendicitis aguda, árbol de decisión ID3, diagnóstico computarizado, niños.

1. Introducción

La patología quirúrgica más frecuente en abdomen es la apendicitis. La apendicitis es actualmente el procedimiento quirúrgico de urgencia más común

en el mundo. Uno de cada 15-20 mexicanos presentará apendicitis aguda en algún momento de su vida. La apendicitis aguda representa la patología quirúrgica más común en la infancia y se presenta en 1-2 casos por cada 10,000 niños menores a 4 años y 25 casos por cada 10,000 en niños entre 4 y 17 años de edad. La sospecha y diagnóstico de apendicitis aguda se basa predominantemente en la clínica. El diagnóstico incorrecto o tardío aumenta el riesgo de complicaciones como infección de herida quirúrgica (8 - 15%), perforación (5 - 40%), abscesos (2 - 6%), sepsis y muerte (0.5 - 5%). El diagnóstico retardado incrementa costos en el servicio de urgencias y hospitalarios. Ésta enfermedad es muy común y sigue siendo un problema de diagnóstico y representa un reto para todos los médicos que atienden al paciente con sintomatología, a pesar de la experiencia y los diferentes métodos de diagnóstico clínico y paraclínico llegando a desconcertar hasta al mejor de los médicos [10].

Existen muchos intentos para mejorar el diagnóstico uno de ellos es el análisis en laboratorio, en cuanto a: el conteo de glóbulos blancos, neutrófilos y proteína C reactiva [7]. Más su valor diagnóstico es limitado. Otras técnicas más sofisticadas son las radiografías, ecografías y tomografías computarizadas, estas últimas tienen un alto riesgo de radiación ionizante innecesaria para el paciente [28], inaccesible en ciertas unidades médicas y de un costo elevado.

Existen pruebas menos invasivas que han demostrado buena precisión diagnóstica como son los sistemas de puntuación diagnóstico como la escala de Alvarado [1] y la puntuación de apendicitis pediátrica [23].

En la medicina es posible aplicar métodos alternativos de diagnóstico, debido a la gran cantidad de padecimientos involucrados, las sintomatologías y los pacientes. Lo ideal sería que los médicos pudieran contar con el apoyo de una herramienta que les permita analizar los datos sintomatológicos de cada uno de sus pacientes para poder determinar con base en casos anteriores, el diagnóstico más acertado, lo cual representaría un soporte y ayuda para el médico [3].

Una alternativa para la predicción y clasificación de datos, que es utilizada ampliamente en el área de la inteligencia artificial son los árboles de decisión. Diferentes estudios han demostrado que esta técnica es útil en el área médica principalmente en la toma de decisiones [3,27,26]. En el caso de la apendicitis otros autores han utilizado técnicas similares [13,18,25,20,28] aunque la mayoría de ellos, cae en el error de generalizar su muestra lo cual en el campo real sus modelos pueden ser ambiguos y no eficaces en ciertas categorías de la muestra, en este caso los niños.

En este artículo, se propone construir una herramienta de diagnóstico computacional utilizando árboles de decisión ID3 para el diagnóstico de apendicitis aguda en niños entre los 4 y 15 años.

Para la recolección de datos se hizo uso de encuestas en base las escalas de Alvarado y puntuación de apendicitis pediátrica con el fin de determinar que escala es la mejor opción para nutrir la base de conocimientos de un árbol de decisión ID3. El espacio de recolección fue el área de urgencias del Hospital General Regional "Lic. Emilio Sánchez Piedras" (Tzompantepec, Tlaxcala, México) donde se recolectaron un total de 41 casos.

La programación, implementación y prueba de los árboles de decisión ID3 se realizó en Python [21] y Weka (Waikato Environment for Knowledge Analysis, que en español se traduce como “Entorno para análisis del conocimiento de la Universidad de Waikato”) [12].

Los resultados de este trabajo utilizando las escalas de evaluación y árboles de decisión son satisfactorios, sin embargo, es necesario un número mayor de casos para mejorar la exactitud diagnóstica de nuestro proyecto.

2. Conceptos básicos

2.1. Apendicitis aguda

La inflamación aguda del apéndice vermiforme, apendicitis, es una causa frecuente de abdomen agudo (dolor abdominal intenso de aparición súbita). La apendicitis es un proceso evolutivo, secuencial, de allí las diversas manifestaciones clínicas y anatomopatológicas que suele encontrar el cirujano y que dependerán fundamentalmente del momento o fase de la enfermedad en que es abordado el paciente [14].

2.2. Escalas de diagnóstico

Las escalas de diagnóstico o sistemas de puntuación clínica consisten en dar algunos valores previamente establecidos por estudios estadísticos a ciertos síntomas, signos pertinentes y relevantes [17].

En 1986, el Dr. Alvarado publicó la escala de Alvarado [1] uno de los sistemas de puntuación de apendicitis más conocidos y estudiados. La escala de Alvarado incluye tres síntomas, tres signos físicos y dos parámetros de laboratorio. Dicha escala fue nombrada MANTRELS, en representación de la primera letra de las 8 variables. A cada variable se le asigna una valoración numérica dependiendo si el paciente estudiado la presenta o no [4]. Las puntuaciones de 1-4 son negativas (No apendicitis); cuando la puntuación se encuentra entre el rango 5-6 se recomienda llevar a cabo un análisis exhaustivo del paciente mientras que las puntuaciones 7-10 se consideran positivas (Apendicitis). Ver Tabla 1.

La PAS (Pediatric Appendicitis Score, que en español se traduce como Puntuación de Apendicitis Pediátrica) es una puntuación que fue reportada por primera vez por el Dr. Samuel en la Revista de Cirugía Pediátrica en 2002. La PAS fue orientada exclusivamente a la población pediátrica, es por ello que se toma en cuenta en este estudio. Los resultados de esa investigación fueron: sensibilidad de 1, especificidad de 0.92 valor predictivo positivo de 0.96 y valor predictivo negativo de 0.99 mostrando esta herramienta como una forma sencilla pero estructurada, útil y bien fundamentada que puede mejorar la capacidad diagnóstica frente a un paciente con sospecha de apendicitis [19]. Las variables y sus especificaciones se muestran en la Tabla 1.

Tabla 1. Escala de Alvarado y puntuación de apendicitis pediátrica.

VARIABLES CLÍNICAS	ESCALA DE ALVARADO	PUNTAJES DE APENDICITIS PEDIÁTRICA
Migración del dolor hacia el cuadrante inferior derecho	1	1
Anorexia	1	1
Nauseas/vómitos	1	1
Hipersensibilidad en el cuadrante inferior derecho	2	2
Dolor al rebote a la palpación (Signo de Blumberg)	1	1
Temperatura elevada*	1	1
Recuento de leucocitos mayor de 10,000 por mm ³	2	1
Neutrofilia mayor de 75%	1	1
Toz/percusión/hipersensibilidad al dolor al movimiento en el cuadrante inferior derecho		2
Total	10	10
Recomendaciones		
	1-4 Apendicitis negativa	1-3 Apendicitis negativa
	5-6 Diagnostico dudoso (Posible)	4-7Muy probable apendicitis
	7-8 Probable apendicitis	8-10 Muy probable apendicitis
	9-10 Muy probable apendicitis	

* Temperatura elevada o fiebre generalmente se define como mayor o igual que 37.3° C (91. 2°F) para la escala de Alvarado Mayor o igual que 37. 3° C (99.2°F) o 38.0° C (100.4° F) para PAS.

2.3. Árboles de decisión

Los arboles de decisión es una de las técnicas de aprendizaje inductivo supervisado no paramétrico, se utiliza para la predicción y se emplea en el campo de inteligencia artificial, donde a partir de una base de datos se construyen diagramas de construcción lógica, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren en forma repetitiva para la solución de un problema [26].

Propiedades de los Árboles de decisión.

Una de las propiedades de esta técnica es que permite una organización eficiente de un conjunto de datos, debido a que los árboles son construidos a partir de la evaluación del primer nodo (raíz) y de acuerdo a su evaluación o valor tomado se va descendiendo en las ramas hasta llegar al final del camino (hojas del árbol), donde las hojas representan clases y el nodo raíz representa todos los patrones de entrenamiento los cuales se han de dividir en clases. Los sistemas que implementan arboles de decisión tales como ID3 son muy utilizados en lo que se refiere a la extracción de reglas de dominio. Este método (ID3) se construye a partir del método de Hunt. La heurística de Hunt, consiste escoger la característica más discriminante del conjunto X, luego realizar divisiones recursivas del conjunto X, en varios subconjuntos disyuntos de acuerdo a un atributo seleccionado [26].

2.4. Algoritmo ID3

El Algoritmo ID3 fue desarrollado por J. Ross Quinlan en 1986 [22] el cual es considerado un algoritmo seminal, ya que de aquí se derivan muchos algoritmos para la construcción de árboles de decisión.

Este algoritmo se basa en la teoría de la información, desarrollada en 1948 por Claude Elwood Shannon [24]. En particular se utiliza la noción de entropía

descrita en la teoría de la información, para ver que tan aleatorio se encuentra la distribución de un conjunto de ejemplos sobre las clases a las que pertenecen [15].

El objetivo preciso del algoritmo es aprender a partir de la diferencia que existe entre los datos para analizar, esto es, un procedimiento de divide y vencerás, que maximiza la información obtenida, la cual se utiliza como una métrica para seleccionar el mejor atributo que divida los datos en clases homogéneas [3]. A continuación se muestra el pseudocódigo del algoritmo ID3. Ver Algoritmo 1.1.

Algoritmo 1.1: Pseudocódigo del algoritmo ID3 [11]

ID3 (Ejemplos¹, Atributo-objetivo², Atributos³)

¹ **Ejemplos** son los datos de entrenamiento.

² **Atributo-objetivo** es el atributo cuyo valor va a ser precedido por el árbol y que toma valores positivos o negativos.

³ **Atributos** es una lista con otros atributos que pueden ser ensayados o candidatos a ser elegidos para ser la raíz de este árbol.

— Inicio

— Si todos los Ejemplos son positivos, devolver un nodo etiquetado con +

— Si todos los Ejemplos son negativos, devolver un nodo etiquetado con -

— Si Atributos está vacío, devolver un nodo etiquetado con el valor más frecuente de Atributo-objetivo en Ejemplos.

— En otro caso:

– Sea A el atributo de Atributos que MEJOR clasifica Ejemplos

– Crear **Árbol**, con un nodo etiquetado con A.

– Para cada posible valor v de A, hacer:

- Añadir un arco a **Árbol**, etiquetado con v.

- Sea Ejemplos (v) el subconjunto de Ejemplos con valor del atributo A igual a v.

- Si Ejemplos (v) es vacío

* Entonces colocar debajo del arco anterior un nodo etiquetado con el valor más frecuente de Atributo-objetivo en Ejemplos.

* Si no, colocar debajo del arco anterior el subárbol

ID3 (Ejemplos (v), Atributo-objetivo, Atributos - A).

—Devolver **Árbol**

3. Desarrollo

3.1. Contexto de la investigación

Este trabajo fue de tipo correlacional de carácter prospectivo, longitudinal y observacional. Se llevó a cabo en el Área de urgencias del Hospital General Regional “Lic. Emilio Sánchez Piedras” (Tzompantepec, Tlaxcala, México) de febrero a noviembre de 2015.

Criterios de selección Para recolectar los datos de nuestro estudio se tomaron en cuenta una serie de criterios de inclusión, exclusión y eliminación con el fin de regular nuestra muestra, Ver Tabla 2.

Tabla 2. Criterios de selección de pacientes.

Criterios de inclusión	Criterios de exclusión
<ul style="list-style-type: none">- Se incluyeron todos los pacientes mayores de 4 años y menores de 15 años de edad con dolor abdominal.- Pacientes con diagnóstico clínico de apendicitis efectuado por el médico en turno.- Pacientes que aceptaron participar en el estudio de forma libre y voluntaria.- Pacientes de ambos sexos.	<ul style="list-style-type: none">- Pacientes con dolor abdominal con el antecedente de ser post operados de apendicitis.- Pacientes que tengan diagnóstico de apendicitis sin estudios de laboratorios.- Pacientes con escala incompleta.
Criterios de eliminación	
<ul style="list-style-type: none">- Pacientes que no cuenten con resultados de biopsia.	

3.2. Recopilación de datos

Los pacientes entre 4 y 15 años que acudieron al servicio de urgencias del Hospital General Regional *Lic. Emilio Sánchez Piedras* (Tzompantepec, Tlaxcala, México) con dolor abdominal agudo de febrero a noviembre de 2015 se incluyeron en este estudio. Pacientes con más de 15 años, con antecedentes de post operación apendicular o encuestas incompletas fueron excluidos del estudio.

La técnica de recolección de datos fue la encuesta, ya que permitió obtener los datos de modo rápido y eficaz. La Figura 1, muestra nuestras encuestas.

Estas encuestas fueron divididas en tres áreas: Datos de cabecera, datos de escala y datos de confirmación. Los datos de cabecera se utilizaron para describir a la población y el tiempo de evolución de la enfermedad. Los datos de escala se utilizaron para construir nuestros árboles de decisión ID3.

Para la selección de las variables nos basamos en la literatura de 2 escalas de puntuación diagnóstica, con el fin de determinar que puntuación tiene mejores resultados en el diagnóstico de apendicitis aguda en niños. Las escalas utilizadas fueron la escala de Alvarado [1] y Puntuación de Apendicitis pediátrica (PAS) [23].

Las variables de diagnóstico son para poder confirmar la afección y si es el caso, saber en qué fase de evolución se encontraba el miembro retirado. Además de saber el tiempo que pasan los pacientes con apendicitis aguda desde su llegada

Arboles de decisión ID3 para el diagnóstico de apendicitis aguda en niños

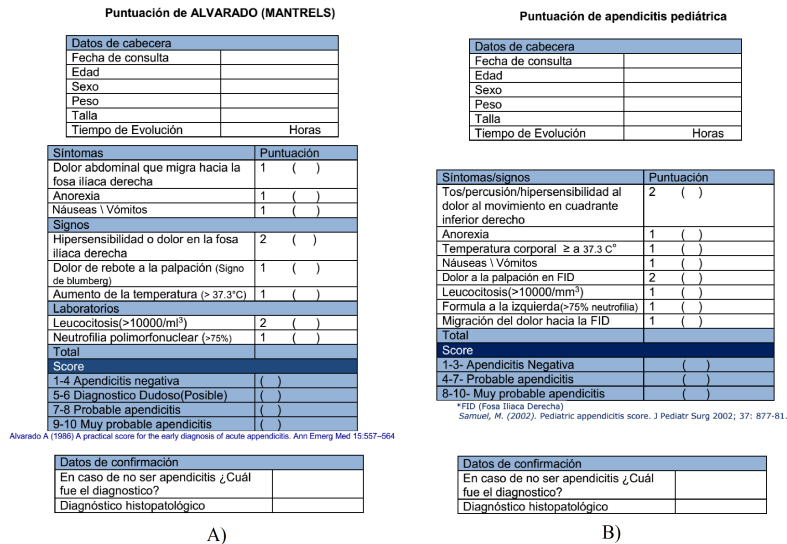


Fig. 1. Herramientas para la adquisición de datos A) En base a escala de Alvarado y B) En base a Puntuación de Apendicitis Pediátrica (Encuestas).

hasta su cirugía. El diagnóstico de los pacientes fue clasificado en dos categorías: apendicitis y no apendicitis.

Estas encuestas fueron realizadas por los médicos internos de urgencias con colaboración con los médicos internos de quirófano de manera prospectiva.

Para que existiera una variabilidad en la muestra se decidió aplicar la encuesta de manera intercalada, es decir, una semana Alvarado y otra semana PAS.

Para darle veracidad al estudio se obtuvieron los resultados del análisis histopatológico de todos los miembros del estudio con cirugía realizada y se compararon con los resultados de la encuesta. El diseño del estudio fue aprobado por la junta de revisión institucional local y el consentimiento informado del paciente no era necesario.

3.3. Construcción de los árboles

Para analizar y crear el árbol de decisión ID3 utilizamos el software WEKA. Los parámetros de configuración fueron los predeterminados por el software. Para probar los resultados del clasificador se manejó el algoritmo de validación cruzada de 10 iteraciones. Weka por ser un software de análisis solo nos permite crear y evaluar el clasificador, por ello se decidió crear una interfaz gráfica con el lenguaje de programación Python con el fin de analizar de manera más exhaustiva el desarrollo de los árboles de decisión ID3.

El software generado con Python cuenta con 3 apartados principales: un módulo de importación de archivos *.csv*; una ventana de análisis de datos y

generación de reglas; y un apartado de pruebas. Como se puede ver en la Figura 2, la ventana de visualización de datos permite analizar una base de conocimientos; obtener las entropías y ganancias de manera global y de los primeros atributos del conjunto de datos. Además, generar las reglas de un árbol de decisión ID3.

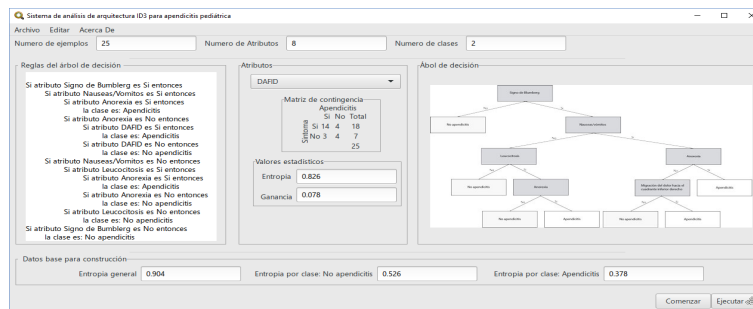


Fig. 2. Ventana de análisis de datos y generación de reglas.

En el apartado de pruebas se examina la precisión del clasificador ID3 a través de las reglas generadas. Este módulo nos permite cargar un conjunto de datos de prueba, en un archivo *.csv* y examinarlo con el clasificador generado. Este entorno nos proporciona las siguientes pruebas diagnósticas: sensibilidad, especificidad, tasas de falsos positivos, tasa de falsos negativos, el valor predictivo positivo, el valor predictivo negativo y la precisión global. Además, nos permite visualizar el resultado del análisis a través de una matriz de confusión y un gráfico de pastel que muestra el porcentaje de casos clasificados.

Para construir este software fue necesario ocupar Glade 3 [8] y las siguientes librerías externas de python: GTK+3 [9], numpy [16] y matplotlib [6].

4. Resultados

Se recolectaron un total de 41 casos, 25 de ellos fueron reportados con nuestra herramienta de recolección de datos en base a la escala de Alvarado y los 16 restantes en base a la puntuación de apendicitis pediátrica. La edad promedio de la población fue de 8 años con una tendencia a variar por debajo o por encima de 4 años; nuestra población tuvo una mayor influencia en casos de varones con un 63% contra un 37% de mujeres; el peso promedio fue de 32.26 kilogramos con una variación de 17 kilogramos; la talla promedio fue de 124.55 centímetros con una variación de 35 centímetros; el tiempo de evolución promedio de la enfermedad hasta la atención médica fue de 35.39 horas con una variación de 67 horas; el 63% de los paciente sufrieron automedicación antes de asistir con un médico; y el tiempo promedio de espera de un paciente desde su llegada al servicio de urgencias hasta su intervención quirúrgica de 19 casos reportados fue de 5:32 horas con una variación de 0.129 horas. En el caso de

Arboles de decisión ID3 para el diagnóstico de apendicitis aguda en niños

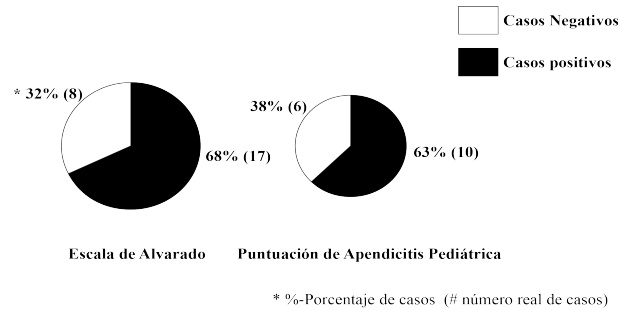
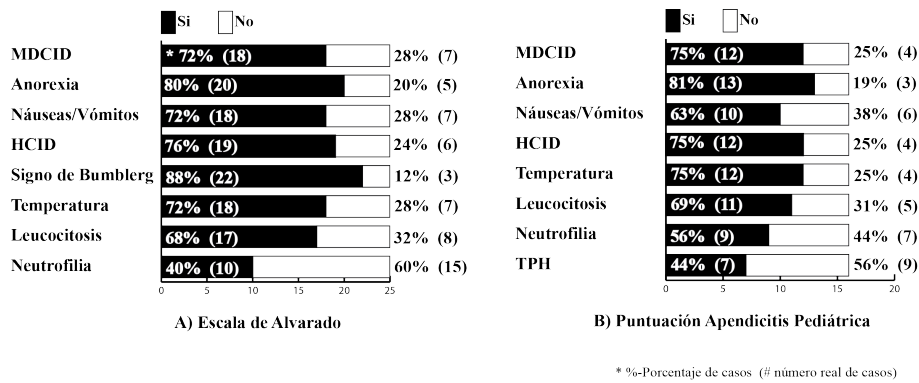


Fig. 3. Comparación de muestras recolectadas con escalas de diagnóstico médico.



MDCID - Migración del dolor hacia el cuadrante inferior derecho.
 HCID - Hipersensibilidad en el cuadrante inferior derecho.
 TPH - Toz/percusión/hipersensibilidad al dolor al movimiento en el cuadrante inferior derecho.

Fig. 4. Porcentaje de pacientes que presentaron signos, síntomas y laboratorios de acuerdo a cada escala de diagnóstico médico. A) Datos con respecto a la escala de Alvarado y B) Datos con respecto a la Puntuación de Apendicitis Pediátrica.

los datos recolectados con las escalas de diagnóstico médico se obtuvo un 68% de casos positivos (Apendicitis Aguda) con la escala de Alvarado y 63% con la Puntuación de Apendicitis Aguda, obteniendo un mayor número de casos positivos para entrenar los arboles de decisión, Ver Figura 3. La descripción porcentual de cada escala se presenta en la Figura 4. Las etiquetas (Si - No) se refieren a la presencia de dicho síntoma. En el análisis de datos se encontró que los síntomas que tenían un mayor peso en la ponderación de la escala se presentan de manera diferente en nuestra muestra. En la escala de Alvarado el síntoma signo de blumberg (88%) e hipersensibilidad en el cuadrante inferior derecho (76%), estas variables se encuentran en los primeros lugares de incidencia. En la puntuación de apendicitis los síntomas predominantes se presentan de manera diferente, la hipersensibilidad en el cuadrante inferior derecho tiene un valor 75% y toz/percusión/hipersensibilidad al dolor al movimiento en el cuadrante inferior

derecho (44 %) tiene una incidencia menor dentro de los sujetos de estudio, sin embargo, el síntoma Anorexia se presenta en ambas escala con un porcentaje mayor o igual a 80 %

El resultante del árbol de decisión ID3 en base a la escala de Alvarado fue un árbol binario de 4 niveles, 13 nodos, 12 ramas, 7 hojas y profundidad 5. Los resultados para este clasificador utilizando validación cruzada de 10 iteraciones fueron: sensibilidad (0.824), especificidad (0.625), razón de probabilidad positiva (2.20), razón de probabilidad negativa (0.28), precisión global (76 %) y área bajo la curva (0.724).

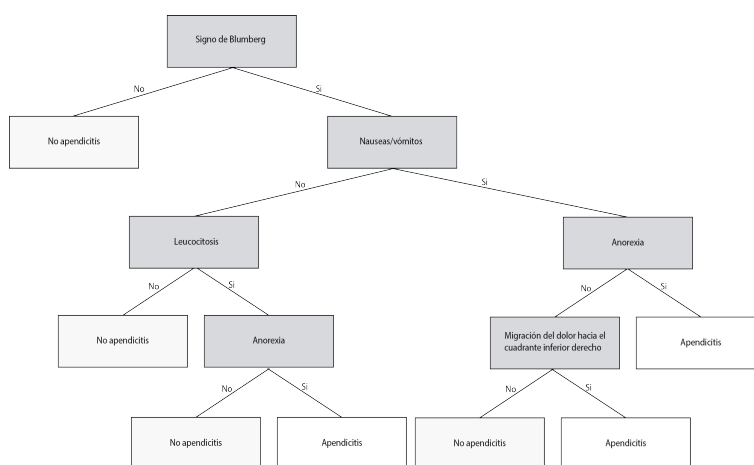


Fig. 5. Árbol de decisión ID3 en base a la escala de Alvarado.

En el caso del árbol de decisión ID3 en base a la puntuación de apendicitis pediátrica, Ver Figura 6, es un árbol binario de 2 niveles, 5 nodos, 4 ramas, 3 hojas y profundidad 3. Los resultados para este clasificador fueron los siguientes: sensibilidad (0.91), especificidad (1), razón de probabilidad positiva (0), razón de probabilidad negativa (0.09), precisión global (93.75 %) y área bajo la curva (0.917).

5. Conclusiones y trabajo futuro

La apendicitis es una enfermedad muy común en los seres humanos. A pesar de los grandes avances tecnológicos existen tasas de perforación muy altas, la capacidad de diagnóstico se complica para los especialistas en edades pediátricas y en edades avanzadas. En zonas alejadas de la ciudad se carecen de métodos de diagnóstico de muchas enfermedades entre ellas la apendicitis aguda. En este trabajo se aborda el problema del diagnóstico de apendicitis aguda en infantes por medio de métodos computacionales inteligentes como los arboles de decisión ID3.

Tabla 3. Tabla comparativa de los trabajos relacionados

Autores	Población	Metodo	Escala de puntuación	Sensibilidad	Especificidad	País
[27]	532	Árboles de decisión C5.0	Alvarado	0.95	0.80	Taiwán
[28]	156	RNA Perceptrón multicapa/Backpropagation	Lintula	0.97	1	Turquía
[5]	100	RNA Perceptrón multicapa/Backpropagation	Alvarado	0.92	0.81	México
[20]	60	RNA Perceptrón multicapa/Backpropagation	Alvarado	1	0.97	Reino Unido
[18]	911	RNA ART1	Propia (Diferentes escalas)	P1: 0.85	P1: 0.87	Finlandia
				P2: 0.85	P2: 0.89	
				P3: 0.91	P3: 0.73	
				P4: 0.46	P4: 0.80	
				P5: 0.79	P5: 0.78	
	RNA SOM		P1: 0.83	P1: 0.90		
			P2: 0.88	P2: 0.77		
			P3: 0.62	P3: 0.82		
			P4: 0.55	P4: 0.87		
			P5: 0.55	P5: 0.83		
			P1: 0.88	P1: 0.89		
			P2: 0.88	P2: 0.91		
			P3: 0.85	P3: 0.90		
			P4: 0.68	P4: 0.79		
			P5: 0.65	P5: 0.85		
			P1: 0.88	P1: 0.88		
			P2: 0.86	P2: 0.89		
			P3: 0.85	P3: 0.86		
			P4: 0.64	P4: 0.85		
			P5: 0.83	P5: 0.92		
[13]	180	Bosques aleatorios 200 arboles	[2]	0.94	1	Taiwán
		RNA Perceptrón multicapa/Backpropagation	con exclusión de sangre oculta en la orina y la hemoglobina.	0.94	0.85	
		Máquinas de vectores de soporte		0.91	1	
		Regresión logística		0.91	0.62	
[25]	2230	RNA Perceptrón multicapa/Backpropagation	Propia	.92		India
		Clasificador bayesiano		.87.		
Sánchez et al., 2016	41	Árbol de decisión ID3	Alvarado y Puntuación de Apendicitis Pediátrica	P1=0.82 P2=0.91	P1=.63 P2= 1	México

P (Prueba) y RNA (Red Neuronal Artificial); P - Se tomó el valor menor y mayor de cada prueba.

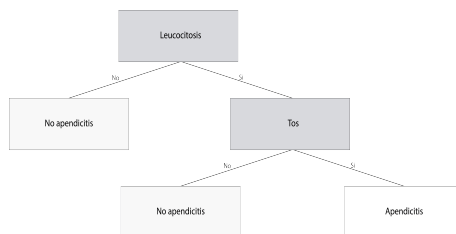


Fig. 6. Árbol de decisión ID3 en base a la puntuación de apendicitis pediátrica.

En el caso de la recolección de los datos, las escalas de diagnóstico médico fueron una elección acertada que nos permitieron reducir el número de variables clínicas y que por ende se reflejó en el número de datos a procesar. Uno de los principales problemas en este estudio fue la disposición del personal para recolectar datos, problemas como la pérdida de encuestas, el no seguimiento de los casos y la atención de los pasantes hacia la investigación fueron algunas limitantes. Aunque el promedio de afecciones al año no era bastante se considera que el modelo mejoraría si se hubieran obtenido más casos.

En el desarrollo de este trabajo se consideraron diferentes plataformas para el análisis de datos, optando por Weka, una opción eficiente y eficaz para obtener modelos de clasificación, a pesar de estas características, este software carece de una herramienta para exportar las reglas del árbol de decisión ID3 y ejecutar pruebas. Con el software que se desarrolló en Python se consiguió traducir las reglas, hacer pruebas y obtener algunas métricas de validación.

En el caso de la eficiencia de los clasificadores, el mejor de los modelos fue el árbol en base a la puntuación de apendicitis pediátrica con un 91 % de sensibilidad, 100 % de especificidad, precisión global del 93.75 % y un área bajo la curva de 0.917. Es importante mencionar que aunque existe una población menor de la puntuación de apendicitis pediátrica (16 casos) se obtuvieron buenos resultados en comparación con la escala de Alvarado donde se obtuvo menor precisión. La diferencia de los clasificadores fue de 17.5 de acuerdo con su precisión global.

Como se puede ver en la tabla 3, los trabajos que analizamos y que abordan la misma enfermedad tienen los siguientes resultados. Con la técnica de redes neuronales artificiales [28,5,20,18,13,25] tienen un desempeño que va del 87 % al 100 % de sensibilidad y en el caso de la especificidad va del 85 % al 100 %. Para el caso de los árboles de decisión y bosques aleatorios [27,13] su especificidad es de 94 % y especificidad de 80 % a 100 %. En el caso de clasificadores bayesianos [25] la especificidad es de 87.87 %. Con esto podemos determinar que nuestros clasificadores se encuentran en un rango promedio con respecto a trabajos relacionados e incluso en algunos casos superándolos en algunas métricas de eficiencia, aunque, cabe mencionar que dichas técnicas tienen otras especificaciones como el tipo de aprendizaje, es decir, RNA LVQ y SOM [18].

De acuerdo a los valores de asertividad de los árboles de decisión ID3 los resultados son prometedores y muy buenos a nivel cuantitativo. Sin embargo,

consideramos que se necesita un mayor número de ejemplos para poder certificar la calidad diagnóstica de nuestro trabajo y que nos permita pasar a la siguiente fase de pruebas en humanos. Consideramos que se cumplieron los objetivos de esta investigación, es necesario seguir con este trabajo para maximizar el grado de efectividad y que pueda ser utilizada por personal médico.

Para trabajos futuros lo ideal será abarcar más hospitales del estado o cooperar con nosocomios fuera de nuestra región para obtener un número mayor de ejemplos. Además, utilizar una sola escala de diagnóstico y de esta forma no segmentar nuestra muestra. Por la aproximación que obtuvimos con nuestros datos, lo ideal sería utilizar como base la Puntuación de Apendicitis Pediátrica. Sería necesario plantear una nueva herramienta de recolección de datos pensando a futuro con el fin de probar otras técnicas como lógica difusa y de esta forma generar una herramienta multipropósito para construir con los datos recolectados diferentes modelos con diversas técnicas que nos permitan ir más allá de clasificar (Apendicitis o No apendicitis), es decir, la fase en la que se encuentra el paciente y por supuesto poder diagnosticar otras patologías con cuadro sintomatológicos parecidos a la de la apendicitis aguda, utilizando la base de conocimientos generada.

Agradecimientos. Agradecemos a la Universidad Autónoma de Tlaxcala, la Secretaría de Salud del estado de Tlaxcala, al Hospital General Regional “Lic. Emilio Sánchez Piedras” (Tzompantepec, Tlaxcala, México), al Dr. Jorge Castro Pérez (Jefe de Enseñanza del HGRES), al Dr. Guadalupe Sánchez Pastrana (Médico Patólogo del HGT) y por supuesto a los médicos pasantes del año 2015 que nos apoyaron en la recolección de datos. Además a Fundación Telmex.

Referencias

1. Alvarado, A.: A practical score for the early diagnosis of acute appendicitis. *Annals of emergency medicine* 15(5), 557–564 (1986)
2. Andersson, R.: Meta-analysis of the clinical and laboratory diagnosis of appendicitis. *British journal of surgery* 91(1), 28–37 (2004)
3. Barrientos, E., Cruz, N., Acosta, H.: Árboles de decisión como herramienta en el diagnóstico médico. *Revista Médica de la Universidad Veracruzana* 9(2), 19–24 (2009)
4. Chavarría, B.C., Corral, D.M., González, S.V.: Efectividad de la tabla mantrel y sus modificaciones: detección de apendicitis en niños menores de 13 años. *Avances [Pediatria]* 1(1), 20–25 (2013)
5. Chavarría, B.C., Gutiérrez, S.U.: Mejoramiento del diagnóstico de apendicitis aguda en niños usando tabla de valoración mantrel a través de redes neuronales (perceptrón multi-capas). *Revista digital de posgrado, investigación y extensión del Campus Monterrey* (62) (Abril 2003)
6. Dale, D., Droettboom, M., Firing, E., Hunter, J.: Matplotlib. *SciPy* (2015), <http://matplotlib.org/contents.html>
7. Echániz, J.S., García, M.L., Ronco, M.V., Raso, S.M., Fernández, J.B., Alvarez-Buhilla, P.L.: Valor diagnóstico de la proteína c reactiva en las sospechas de apendicitis aguda en la infancia. *An Esp Pediatr* 48(5), 470–474 (1998)

8. GNOME-Project: Manual del diseñador de interfaces de usuario glade. GNOME-developer (2005), <https://developer.gnome.org/glade/stable/>
9. GNU: The python gtk+ 3 tutorial. GNU Free Documentation License 1.3 (2016), <https://python-gtk-3-tutorial.readthedocs.org/en/latest/#>
10. González, J., López, G., Cedillo, E., Juárez, M., González, D., López, J., González, R.: Guía de práctica clínica apendicitis aguda. Guías clínicas de la asociación mexicana de cirugía general (2014), <http://amcg.org.mx/images/guiasclinicas/apendicitis.pdf>
11. Gutiérrez, M., Martín, F., J, R.: Aprendizaje inductivo: árboles y reglas de decisión (2012), <https://www.cs.us.es/cursos/ia2/temas/tema-02.pdf>
12. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. ACM SIGKDD explorations newsletter 11(1), 10–18 (2009)
13. Hsieh, C.H., Lu, R.H., Lee, N.H., and Min Huei Hsu, W.T.C., Yu-Chuan, Li, J.: Novel solutions for an old disease: diagnosis of acute appendicitis with random forest, support vector machines, and artificial neural networks. Surgery 149(1), 87–93 (Enero 2011)
14. Moore, K.L., Dailey, A.F., Agur, A.M.: MOORE Anatomía con orientación clínica. Edición en español, Wolters Kluwer Health (2013)
15. Moreno, B., Kinney, R.: Minería sobre grandes cantidades de datos. Master's thesis, Universidad Autónoma Metropolitana (Noviembre 2009), http://mcyti.izt.uam.mx/archivos/Tesis/Generaci%F3n2007/ICR_Benjam%EDnMoreno.pdf
16. NumPy-Community: Numpy reference. SciPy (2015), <http://docs.scipy.org/doc/numpy-1.10.1/user/>
17. Ohmann, C., Yang, Q., Franke, C.: Diagnostic scores for acute appendicitis. abdominal pain study group. The European journal of surgery = Acta chirurgica 161(4), 273–281 (1995)
18. Pesonen, E., Eskelinen, M., Juhola, M.: Comparison of different neural network algorithms in the diagnosis of acute appendicitis. International journal of bio-medical computing 40(3), 227–233 (Enero 1996)
19. Pinilla, A.: Aplicación de una escala diagnóstica de apendicitis en niños: estudio prospectivo. trabajo presentado para optar al título de especialista en cirugía pediátrica. universidad nacional de colombia (2013), <http://www.bdigital.unal.edu.co/11423/>
20. Prabhudesai, S., Gould, S., Rekhraj, S., Tekkis, P., Glazer, G., Ziprin, P.: Artificial neural networks: Useful aid in diagnosing acute appendicitis. World Journal of Surgery 32(2), 305–309 (Septiembre 2008)
21. Python: Python 2.7.11 Documentation. Python Software Foundation (2016), <https://docs.python.org/2.7/>
22. Quinlan, J.R.: Induction of decision trees. Machine learning 1(1), 81–106 (1986)
23. Samuel, M.: Pediatric appendicitis score. Journal of pediatric surgery 37(6), 877–881 (2002)
24. Shannon, C.E.: A mathematical theory of communication. Bell System Technical Journal 27, 379–423, 623–656 (Julio, Octubre 1984)
25. Sivasankar, E., Rajesh, R.S., Venkateswaran, S.R.: Diagnosing appendicitis using backpropagation neural network and bayesian based classifier. International Journal of Computer Theory and Engineering 1(4), 358–363 (Octubre 2009)
26. Solarte, G., Soto, J.: Árboles de decisiones en el diagnóstico de enfermedades cardiovasculares. Scientia et Technica 3(49), 104–109 (2011)

27. Ting, H.W., Wua, J.T., Chan, C.L., Lin, S.L., Chen, M.H.: Decision model for acute appendicitis treatment with decision tree technology. a modification of the alvaro scoring system. *Journal of the chinese medical association* 73(8), 401–406 (Agosto 2010)
28. Yoldas, O., Mesut, T., Karaca, T.: Artificial neural networks in the diagnosis of acute appendicitis. *The american journal of emergency medicine* 30(7), 1245–1247 (Septiembre 2012)