

---

# Diseño de un algoritmo genético para la generación de conocimiento en el diagnóstico del Síndrome Autista

---

Yrma Cárdenas, Lic. Luis Guerra<sup>1</sup>, Lic. David Mauricio<sup>1</sup>

<sup>1</sup>Facultad de Ingeniería de Sistemas e Informática  
Universidad Nacional mayor de San Marcos

yrma.cardenas@gmail.com, luisg493@gmail.com, dms\_research@yahoo.com

---

## RESUMEN

El autismo es una enfermedad en la cual están presentes muchos síntomas y signos, por lo que es, a veces, muy complicado el poder diagnosticarla. En el presente trabajo se emplea el diseño de un algoritmo genético para construir un conocimiento médico parcial a partir de un número de hallazgos o síntomas presentes en el autismo. Dichos hallazgos son codificados internamente como cromosomas en forma binaria, los cuales son evaluados por una función fitness en términos de exactitud. Los mejores cromosomas son combinados y reproducidos en la siguiente generación. La importancia del presente tema radica en que es un área que ha sido poco explorada dentro del desarrollo en el campo médico automatizado en nuestro país.

**Palabras clave:** Autismo, cromosomas, algoritmo genético, hallazgos

## ABSTRACT

Autism is a condition in which they are present many symptoms and signs so it is sometimes very complicated to diagnose. This paper uses a genetic algorithm to construct a partial medical knowledge from a number of findings or symptoms present in autism. These findings are internally encoded as binary chromosomes, which are evaluated by a fitness function in terms of accuracy. The best chromosomes are combined and played back on the next generation. The importance of this issue is that it is an area that has been little explored within the development in automated medical field in our country.

**Keywords:**

## INTRODUCCIÓN

El rápido crecimiento de la información y las cantidades crecientes de datos han ocasionado que el análisis manual que realizaban los médicos para elaborar un diagnóstico se haya vuelto una tarea muy tediosa y muchas veces imposible. Muchos datos ocultos y además útiles no pueden ser reconocidos fácilmente por el médico especialista. El crecimiento acelerado de los datos requiere una manera automatizada de extraer el conocimiento útil, a través de metodologías computacionales de selección y de clasificación de la información.

La extracción del conocimiento de bases de datos médicas ayudan al médico en la toma de decisiones para poder diagnosticar un determinado tipo de enfermedad. El poder descubrir ese conocimiento, para poder extraer la clasificación de unas reglas, las cuales sean comprensibles por el experto médico. Simular el conocimiento médico y tratar de encontrar el mejor conjunto de reglas para diagnósticos médicos, mediante la optimización de los parámetros más relevantes. Este descubrimiento del conocimiento puede ser obtenido empleando numerosas técnicas de inteligencia artificial aplicadas a la medicina como la lógica difusa, las redes neuronales y los algoritmos evolutivos dentro de los cuales están contenidos los algoritmos genéticos.

Cabe indicar además que existen mezclas de estas técnicas las cuales se aplican para construir los sistemas inteligentes híbridos, que tienen un rendimiento inteligente muy superior a los sistemas inteligentes simples y una velocidad de procesamiento mucho mayor [1][2].

Estos sistemas inteligentes híbridos se vienen aplicando actualmente a varios dominios médicos con mucho éxito. Entre los sistemas híbridos empleados en el campo de la medicina tenemos: redes neuronales con lógica difusa, redes neuronales con algoritmos evolutivos, lógica difusa con algoritmos evolutivos, aprendizaje automático con lógica difusa, aprendizaje automático y algoritmos evolutivos y otras combinaciones más.

Una de las características de la Inteligencia Artificial es su buena adaptación para tratar información incompleta y/o imprecisa. La mayoría de técnicas de inteligencia artificial han sido desarrolladas para operar justamente bajo estas condiciones.

El proceso de extraer conocimiento mediante un sistema inteligente desde una base de datos médica es útil para la toma de decisiones del médico. Tradicionalmen-

te en la mayoría de los dominios, este análisis de datos se hacía mediante un proceso manual o semiautomático; uno o más analistas con conocimiento de los datos y con la ayuda de técnicas estadísticas proporcionaban resúmenes y se generaban informes o validaban modelos sugeridos manualmente por los expertos. Sin embargo, este proceso, en especial la generación de modelos es irrealizable conforme aumenta el tamaño de los datos y el número de dimensiones o parámetros se incrementa. Surge la necesidad de metodologías para el análisis inteligente de los datos que permitan descubrir un conocimiento útil a partir de ellos [1][2].

En las secciones de este artículo enfocamos los detalles concernientes al trabajo. En la sección 2, se describe el síndrome autista(S.A), en el apartado 3 sobre el algoritmo genético propuesto, finalmente conclusiones y referencias.

## SÍNDROME AUTISTA

La Asociación Americana de Autismo (AAA) propone la siguiente definición: el autismo es una discapacidad severa y crónica del desarrollo que aparece normalmente durante los tres primeros años de vida. Ocurre en 15 de cada 10,000 nacimientos y es más común en niños que en niñas. Está presente en todo tipo de razas, etnias y clases sociales de todo el mundo. No se conoce ningún factor en el entorno psicológico del niño como causa directa del autismo. Los síntomas causados por trastornos cerebrales afectan al niño, manifestándose como carencia de interacción social, habilidad para la comunicación, patrones de interés y patrones del comportamiento[6]. Un punto que caracteriza a los niños con autismo es que ellos son incapaces de poder seleccionar algún evento sea este o no importante. Ellos son saturados con frecuencia por muchos estímulos y de esta manera ellos adoptan comportamientos repetitivos, o comportamientos agresivos [6].

Se considera al autismo como una entidad multifactorial que tiene muchos desencadenantes. Investigaciones realizadas consideran que el autismo se debe a dificultades perinatales, a la rubéola congénita, al síndrome de down y a otras alteraciones[5]. Su diagnóstico debe basarse en la clínica. Se necesita de multidisciplinas para realizar el diagnóstico, debido a que se presentan varios problemas en las personas que lo padecen. El equipo debe incluir médicos, terapeutas del lenguaje, terapeutas ocupacionales, psicólogos y fisioterapeu-

tas. Es difícil el poder diferenciar al autismo desde una perspectiva clínica, debido a otros desordenes desintegrativos del desarrollo, ya que puede no ser reconocido fácilmente en individuos que están levemente afectados o en los que tienen impedimentos múltiples, debido a la presencia de un retraso grave o profundo del desarrollo, a alteraciones del lenguaje y a desórdenes de atención[6].

Según la Academia Americana de Neurología para identificar a un niño con autismo se requiere de dos planos de investigación, cada uno dirigido a un tratamiento diferente. Uno de vigilancia de rutina sobre el desarrollo que se realizará a todos los niños, donde se identificará a cualquier niño con desarrollo atípico, seguidos de aquellos en riesgo de autismo. El otro plano involucra una investigación más profunda de los niños identificados y diferencia el autismo de otros desórdenes del desarrollo [6].

Los criterios de las instituciones que indican acerca del diagnóstico del síndrome autista como son la Organización Mundial de la Salud (OMS), la Asociación Psiquiátrica Americana (APA), y el Instituto de Investigación del Autismo (ARI) identifican las características de esta enfermedad [3][4], damos a continuación, entre otros, lo que ellos indican.

#### **CRITERIOS DE LA OMS (CLASIFICACIÓN INTERNACIONAL DE ENFERMEDADES - CIE-10)**

1. Se requiere que al menos en una de las siguientes áreas haya habido retraso o un patrón anormal de funcionamiento con anterioridad a los tres años:
2. En el lenguaje receptivo y/o expresivo tal y como se utiliza para la comunicación social.
3. En el desarrollo de vínculos sociales delectivos y/o en la interacción social recíproca.
4. En el juego funcional y/o simbólico.
5. Alteración cualitativa en la interacción social recíproca.
6. Alteraciones cualitativas de la comunicación.
7. Patrones restringidos, repetitivos y estereotipados de conducta, intereses y actividades.
8. El cuadro clínico no es atribuible a otras variedades de trastornos profundos del desarrollo (síndrome de Asperger, síndrome de Rett, trastorno desintegrativo infantil) o a trastornos específicos del desarrollo del lenguaje receptivo con problemas socioemo-

cionales secundarios, trastorno reactivo de la vinculación, trastorno desinhibitorio de la vinculación, retraso mental con trastorno emocional/conductual asociado o a esquizofrenia de inicio infrecuente precoz.

Además de:

- Retraso mental significativo y/o grave
- Comportamiento autoagresivo.
- Temores, fobias.
- Trastornos del sueño y de la conducta alimentaria.

#### **Criterios diagnósticos de la Asociación Americana de Psiquiatría APA**

##### **(Manual Diagnóstico y Estadístico de los Trastornos Mentales - DSM-IV)**

- Alteración marcada en el uso de múltiples tipos de comunicación no verbal como la mirada directa ojo a ojo, expresión facial, posturas corporales y además para regular interacciones sociales.
- Falla para lograr relaciones apropiadas para su nivel de desarrollo.
- Carencia para mostrar disfrute espontáneo, intereses o vínculo con otras personas (por ejemplo: carencia para mostrar, brindar o señalar objetos de interés a otras personas).
- Carencia de reciprocidad social y emocional (ejm: no participar activamente en juegos simples sociales, prefiriendo actividades solitarias o involucrando a otros en actividades solo como simples herramientas o ayudas mecánicas).
- Retraso o ausencia total del desarrollo del lenguaje hablado (no acompañado por el intento de compensación con modos alternativos de comunicación como el lenguaje de gestos o mímico).
- En individuos con lenguaje adecuado, marcadas alteraciones en habilidad de iniciar o sostener una conversación con otros.
- Uso de lenguaje estereotipado y repetitivo o lenguaje idiosincrásico.
- Carencia de juegos de ficción espontáneo y variado, o de juego de imitación social adecuado al nivel evolutivo.
- Preocupación excesiva por un foco de interés (o varios) restringido y estereotipado, anormal por su intensidad o contenido.

- Apego aparentemente inflexible a rutinas o rituales específicos no funcionales.
- Manías motoras repetitivas y estereotipadas (ejm: aleteo o contorsión de manos o dedos, o movimiento complejos de todo el cuerpo).
- Preocupación persistente por partes de objetos. (ejm: ruedas de coches).

#### **Criterios diagnósticos del ARI (Autism Research Institute - teorías Dr. Bernard Rimlad)**

- Infección por candida
- Intolerancia a sustancias alimenticias
- Intolerancia al gluten y/o caseína
- Deficiencias de la enzima hepática Fenolsulfotransferasa (PST)
- Daño cerebral
- Deficiencias de enzima Dipeptidil
- Deficiencia de la enzima Adenilsuccinato
- Toxicidades
- Alteraciones del sueño.
- Conductas autoagresoras.

#### **ALGORITMO GENÉTICO PROPUESTO**

De los trabajos previos relacionados con el tema, mostramos algunos en los cuales se han empleado los algoritmos genéticos en el área médica: Magnetic Resonance Based Ventricle System Classification by Multi-Species Genetic Algorithm[7], A Distributed Genetic Algorithm for Parameters Optimization to Detect Microcalcification in Digital Mammograms[8], An Ischemia Diagnosis System Using Genetic Algorithms and Multicriteria Decision Analysis (HySMID) [9].

#### **3.1 Hallazgos y su codificación para el Síndrome Autista**

De acuerdo al diccionario de la Real Academia Española la “**hallazgo**” acción y efecto de **hallar**, y según el diccionario Océano, hallar es encontrar lo que se busca, observar, notar; estar alegre, enfermo, etcétera.

Un hallazgo o conjunto de hallazgos clínicos dan a conocer la naturaleza de una enfermedad o diagnóstico. En la **Tabla 1A y 1B** se codifican 24 hallazgos (mediante unos y ceros), se validan por las entidades OMS, APA, ARI, y pueden tomar valores diferentes para el síndrome autista, será “1” si la institución

médica considera el hallazgo, es decir si la persona presenta el síntoma, en caso contrario cero “0” si no lo presenta. Ejemplo: Tenemos que el hallazgo ( $h_{19}$ ): Deficiencia enzima Adenilsuccinato, es un hallazgo o síntoma que sí es considerado por el Instituto de Investigación del Autismo (ARI), la cual tendrá el valor de **1**. Mientras que para la OMS, y la APA tendrá el valor de **0** pues no la consideran relevante para el síndrome autista.

De acuerdo a dichas organizaciones, se han identificado 24 hallazgos para el diagnóstico de síndrome autista, y que denotamos como  $h_1, h_2, \dots, h_{24}$ . En las tablas 1A y 1B se denota con **0** ó **1**, para señalar que un hallazgo aparece o no respectivamente en el conocimiento establecido por la entidad.

#### **3.2 Cromosoma**

Un conocimiento para el diagnóstico de enfermedades está dado por una combinación de hallazgos que es reconocida o valuada por instituciones médicas mundiales, que si se verifican nos garantiza que el paciente presenta o no dicha enfermedad.

En la tabla N.º 2, se muestran algunas combinaciones de hallazgos como conocimientos dados por OMS, APA, ARI. Las tres primeras columnas  $K_1, K_2, K_3$  son algunas combinaciones obtenidas a partir de la OMS, las tres columnas siguientes  $K_4, K_5, K_6$  se han obtenido tomando la orientación de APA y las tres últimas combinaciones  $K_7, K_8, K_9$  tomando como guía lo dado por ARI. Cada columna representa una combinación de hallazgos para el diagnóstico de un paciente con síndrome autista. En las combinaciones se ha colocado unos y ceros, para identificar si la institución médica considera o no el hallazgo respectivo.

En el contexto de los algoritmos genéticos, para el presente trabajo, un individuo es dado por un conocimiento que permita hacer el diagnóstico del síndrome autista y, este es representado por un cromosoma de la siguiente forma:

$S_{ij}$ : Matriz de hallazgos, conocimientos.

#### **3.3. Fitness**

Pretendemos construir un conocimiento a partir de experiencias de diagnóstico de pacientes con síndrome autista, que proporcione un diagnóstico con menos equivocaciones.

	Hallazgos	OMS	APA	ARI
h1	Alteración marcada en el uso de múltiples tipos de comunicación no verbal como la mirada directa ojo a ojo, expresión facial, posturas corporales y además para regular interacciones sociales.	1	1	0
h2	Falla para lograr relaciones apropiadas para su nivel de desarrollo.	0	1	0
h3	Carencia para mostrar disfrute espontáneo, intereses o vínculo con otras personas (por ejemplo: carencia para mostrar, brindar o señalar objetos de interés a otras personas).	1	1	0
h4	Carencia de reciprocidad social y emocional (ejm: no participar activamente en juegos simples sociales, prefiriendo actividades solitarias o involucrando a otros en actividades solo como simples herramientas o ayudas mecánicas).	0	1	0
h5	Retraso o ausencia total del desarrollo del lenguaje hablado (no acompañado por el intento de compensación con modos alternativos de comunicación como el lenguaje de gestos o mímico).	1	1	0
h6	En individuos con lenguaje adecuado, marcadas alteraciones en habilidad de iniciar o sostener una conversación con otros.	0	1	0
h7	Uso de lenguaje estereotipado y repetitivo o lenguaje idiosincrásico.	1	1	0
h8	Carencia de juegos de ficción espontáneo y variado, o de juego de imitación social adecuado al nivel evolutivo.	1	1	0
h9	Preocupación excesiva por un foco de interés (o varios) restringido y estereotipado, anormal por su intensidad o contenido.	1	1	0
h10	Apego aparentemente inflexible a rutinas o rituales específicos no funcionales.	0	1	0
h11	Manías motoras repetitivas y estereotipadas (ejm: aleteo o contorsión de manos o dedos, o movimiento complejos de todo el cuerpo).	0	1	0
h12	Preocupación persistente por partes de objetos. (ejm: ruedas de coches).	0	1	0

Tabla N.º 1A. Hallazgos para el Síndrome Autista [Elaboración propia]

h13	Infección por candida	0	0	1
h14	Intolerancia a sustancias alimenticias	0	0	1
h15	Intolerancia al gluten y/o caseína	0	0	1
h16	Deficiencias de la enzima hepática Fenolsulfotransferasa (PST)	0	0	1
h17	Daño cerebral	0	0	1
h18	Deficiencias de enzima Dipeptidil	0	0	1
h19	Deficiencia de la enzima Adenilsuccinato	0	0	1
h20	Toxicidades	0	0	1
h21	Retraso mental significativo y/o grave	1	0	0
h22	Comportamientos autoagresivos.	1	0	1
h23	Temores, fobias.	1	0	0
h24	Trastornos del sueño.	1	0	1

Tabla N.º 1B. Hallazgos para el Síndrome Autista [Elaboración propia]

Sea

- I: Matriz de pacientes, hallazgos.
- Di: Vector diagnóstico de historias clínicas de pacientes "i" según médico.  $i = p$ , pacientes
- $X_{aj}$ : Variable de decisión que indica si un  $h_j$  se presenta o no en un  $K_a$ , donde se cumple que:
- $R_i$ : La respuesta del sistema para un paciente  $P_i$  en relación al síndrome autista y donde se cumple que:

Hallazgos	OMS			APA			ARI		
	K1	K2	K3	K4	K5	K6	K7	K8	K9
h1	1	0	1	1	1	0	0	1	1
h2	0	1	0	1	0	1	0	1	0
h3	1	1	0	1	0	0	0	0	1
h4	0	0	1	1	1	1	0	1	0
h5	1	1	0	1	0	1	0	1	1
h6	0	1	1	1	0	0	0	0	1
h7	1	0	1	1	0	1	0	1	1
h8	1	0	1	1	1	0	0	0	1
h9	1	0	1	1	0	1	0	1	0
h10	0	1	0	1	1	0	0	1	1
h11	0	1	1	1	0	1	0	1	0
h12	0	0	1	1	1	1	0	1	0
h13	0	1	1	0	0	1	1	0	1
h14	0	1	0	0	1	1	1	0	0
h15	0	1	0	0	1	0	1	0	1
h16	0	0	1	0	1	0	1	1	0
h17	0	1	0	0	0	1	1	0	1
h18	0	1	1	0	1	0	1	1	0
h19	0	0	1	0	0	1	1	0	1
h20	0	0	1	0	1	0	1	0	1
h21	1	1	0	0	0	1	0	1	0
h22	1	1	0	0	1	0	1	0	1
h23	1	1	0	0	1	1	0	1	0
h24	1	0	0	0	0	1	1	0	1

Tabla N.º 2. Algunas combinaciones de hallazgos como conocimientos de OMS, APA, ARI. [Elaboración propia].

Considerando  $R_i$  como la respuesta del sistema para un paciente  $P_i$  que presenta el cuadro del síndrome autista, tomará dos valores cero, o uno, y efectuamos como el producto de los valores en las historias clínicas por los valores en la variable de decisión que indica si

un hallazgo  $h_j$  se presenta en un conocimiento sobre la sumatoria de los valores de hallazgos encontrados para un conocimiento.

Cuando coincida los unos en  $X_i$  con los  $h_j$  encontraremos que el producto será igual a uno y la suma de esos unos serán en los lugares donde se verifica el hallazgo  $h = 1$ , si dividimos entre la sumatoria de los unos en  $K_j$  encontraremos el porcentaje de acierto del conocimiento con respecto al diagnóstico médico, de allí que proponemos el cálculo de  $R_i$  como se describe a continuación:

Coefficiente que representa la coincidencia de la historia clínica respecto al conocimiento.

Como todas las combinaciones del conocimiento están definidas para diagnosticar, entonces podemos definir la respuesta del sistema como el máximo de los coeficientes, así por ejemplo si el máximo es 1, significa que el sistema está diagnosticando que el paciente tiene síndrome autista (S.A), y de manera inversa si el máximo es cero significa que el sistema está diagnosticando que el paciente no tiene S.A, lo que expresamos en la siguiente ecuación:

$$R_i = \text{Max}_{1 \leq m} \{R_i^m\}$$

Al desear hallar el error del sistema para un paciente  $P_i$ , sería:

$$E_i = \begin{cases} R_i, D_i = 0 \\ 1 - R_i, D_i = 1 \end{cases}$$

Respecto a la muestra de pacientes  $p$ , el error promedio del sistema sería:

, lo que presentamos como fitness.

### Algoritmo genético propuesto

#### 3.5 Operadores

##### 3.5.1 Selección M Ruleta ( $S'_p, S'_x$ )

Dos cromosomas padres son seleccionados. Uno de ellos de acuerdo al método de la ruleta el que alcanza mayor probabilidad de selección en la población y el otro cromosoma padre, de manera aleatoria-random.

##### 3.5.2 Crossover ( $S'_p, S'_x$ )

Pedir un punto de corte, y unir parte del cromosoma  $S'_p$  a la diferencia de  $S'_x$ ; y de manera similar, con el mismo

punto de corte: parte de  $S'_x$  unir a la otra parte de  $S'_y$ , de tal manera que se obtienen dos nuevos cromosomas hijos (dos nuevos conocimientos) por el cruce .

### 3.5.3 Mutación ( $S'_y$ , $S'_x$ )

Modificar en algún gen 0 por 1 ó 1 por 0 para producir diversidad en la población, usando probabilidad.

### 3.5.4 Aceptación nuevos descendientes ( $S'_y$ , $S'_x$ )

Aceptar aquellos cromosomas que son buenos para diagnóstico.

### 3.5.5 Reemplazar nueva poblacion continua Algoritmo ( $S'_y$ , $S'_x$ , $S'_x, \dots S'_y$ )

Reemplazar nuevos descendientes en la población, y eliminando los de menor fitness, logrando así, una nueva población de cromosomas conocimientos para diagnóstico de S.A.

## CONCLUSIONES

De acuerdo a lo revisado en la literatura especializada podemos afirmar que no hay conocimiento unificado para el síndrome autista. Razón por la cual en el presente trabajo se diseña un AG para generar conocimiento y obtener un mejor diagnóstico presuntivo del S.A. El algoritmo propuesto presenta un fitness que se utiliza para hallar el valor de esta función en cada cromosoma conocimiento de la población; del proceso, cromosomas conocimiento no son seleccionados hasta que se tenga un error mínimo tendiendo a cero.

## REFERENCIAS BIBLIOGRÁFICAS

[1] Georgios Dounias. Hybrid Computational Intelligence in Medicine. University of the Aegean. Dept. of Business Administration, Greece.

[2] Athanasios Tsakonas, George Dounias. Hybrid Computational Intelligence Schemes in Complex Domains. University of the Aegean. Greece. 2000.

[3] The ICD- 10 (International Classification Diseases-10th version) Classification of Mental and Behavioural Disorders – Diagnostic criteria for research – World Health Organization – Geneva 1993.

[4] Diagnostic and Statistical Manual of Mental Disorders – DSM-IV-TR (4th version) American Psychiatric Association.

[5] Rapin I. Autism. The New Journal of medicine 1997;337(2):97-104.

[6] Ritvo ER, Freeman BJ, Pingree C, et.al The UCLA University of UTA epidemiologic survey of autism: prevalence. Am J. Psychiatry 1989; 146:194-9.

[7] Jacob Levman, Javad Alirezaie, Gul Khan , Magnetic Resonance Based Ventricle System Classification by Multi-Species Genetic Algorithm., Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON, Canada, 2000.

[8] Alessandro Bevilacqua, Renato Campanini, Nico Lanconelli : A Distributed Genetic Algorithm for Parameters Optimization to Detect Microcalcification in Digital Mammograms, Department of Electronics, Computer Science and Systems, University of Bologna, viale Risorgimento, 2003.[]

[9] Papaloukas, Y Goletsis, DI Fotiadis , A. Likas , LK Michalis : HySMID: An Ischemia Diagnosis System Using Genetic Algorithms and Multicriteria Decision Analysis, Department of Cardiology, Medical School, University of Ioannina, Ioannina, Greece, 2006.