



Modelo predictivo de enfermedad cardiovascular basado en inteligencia artificial en la atención primaria de salud

>>> La aplicación de técnicas de inteligencia artificial en la prevención de enfermedades cardiovasculares es un avance significativo para identificar riesgo, permitir una intervención temprana y efectiva; lo que puede reducir la mortalidad y mejorar la calidad de vida de las personas afectadas..

>>> AUTORES

Jorge Baudilio Vega Abascal¹, Alberto Rubén Piriz Assa², Diego Nápoles Riaño³

1 Universidad de Ciencias Médicas Holguín, Policlínico Docente “José Ávila Serrano”. Velasco, Holguín, Cuba.

2 Hospital Pediátrico Provincial “Octavio de la Concepción de la Pedraja”. Holguín, Cuba.

3 Universidad de Ciencias Médicas Holguín. Holguín, Cuba.

>>> CORRESPONDENCIA

jvegaabascal28@gmail.com

Fuente: Revista Cubana de Medicina General Integral. 2023;39(3):e2768

>>> RESUMEN

Introducción: En Cuba y en el resto del mundo, las enfermedades cardiovasculares son reconocidas como un problema de salud pública mayúsculo y creciente, que provoca una alta mortalidad.

Objetivo: Diseñar un modelo predictivo para estimar el riesgo de enfermedad cardiovascular basado en técnicas de inteligencia artificial.

Métodos: La fuente de datos fue una cohorte prospectiva que incluyó 1633 pacientes, seguidos durante 10 años, fue utilizada la herramienta de minería de datos WEKA, se emplearon técnicas de selección de atributos para obtener un subconjunto más reducido de variables significativas, para generar los modelos fueron aplicados: el algoritmo de reglas JRip y el meta algoritmo *Attribute Selected Classifier*, usando como clasificadores el J48 y el

Tecnología escalable que acompaña su crecimiento

Módulo WEB, parte de la familia de NextLAB, que permite gestionar amigablemente a Pacientes, Doctores y Laboratorios derivantes



- Consulta de Resultados on line
- Ingresar órdenes en entorno Web
- Solicitar análisis a pie de cama



Detalle del módulo WEB.
Concentra la información del laboratorio en un solo sitio de internet.

p-WEB Brinda la posibilidad para que el paciente, desde cualquier lugar, acceda a sus resultados/ descargar/ imprimir, ingresando un usuario y clave de acceso.

i-WEB Módulo que permite la solicitud a pie de cama de nuevos análisis.

d-WEB Permite administrar la carga, el seguimiento y el resultado, siendo la mejor herramienta para los laboratorios derivantes.



SOFTWARE INTELIGENTE

NextLAB BY Genetrics S.A

Av. del Libertador 8630 6to Piso "1"

C1429EIB Núñez Buenos Aires

T. (+5411)52 63 02 75 Rot

F. (+5411)52 63 02 75 Ext 100

info@nextlab.com.ar

Multilayer Perceptron. Se compararon los modelos obtenidos y se aplicaron las métricas más usadas para clases desbalanceadas.

Resultados: El atributo más significativo fue el antecedente de hipertensión arterial, seguido por el colesterol de lipoproteínas de alta densidad y de baja densidad, la proteína c reactiva de alta sensibilidad y la tensión arterial sistólica, de estos atributos se derivaron todas las reglas de predicción, los algoritmos fueron efectivos para generar el modelo, el mejor desempeño fue con el *Multilayer Perceptron*, con una tasa de verdaderos positivos del 95,2 % y un área bajo la curva ROC de 0,987 en la validación cruzada.

Conclusiones: Fue diseñado un modelo predictivo mediante técnicas de inteligencia artificial, lo que constituye un valioso recurso orientado a la prevención de las enfermedades cardiovasculares en la atención primaria de salud.

Palabras clave: enfermedad cardiovascular; factores de riesgo; modelo predictivo; inteligencia artificial; aprendizaje automático; minería de datos; atención primaria de salud.

>>> INTRODUCCIÓN

Las enfermedades cardiovasculares son las enfermedades crónicas no transmisibles más comunes, responsables de 17,8 millones de muertes a nivel mundial; estas afectan, de manera creciente, a poblaciones en edad laboral y contribuyen desproporcionadamente a la pérdida de años potenciales de vida saludable y de productividad económica, situación que es reconocida como un problema de salud pública mayúsculo y creciente en el mundo y en Cuba.^(1,2)

En Cuba, en el año 2020, las enfermedades del corazón, ocuparon el primer lugar dentro de las causas de muerte, con una tasa de 267,3 por 100 000 habitantes, se incrementó también la mortalidad para las enfermedades cerebrovasculares.⁽³⁾

Las tablas de riesgo son métodos simplificados, basados en funciones matemáticas que modelan el riesgo de los individuos de distintas

cohortes de poblaciones seguidas, generalmente durante 10 años, en las que se establece un algoritmo matemático que permite obtener el porcentaje de sujetos que pueden desarrollar un evento cardiovascular.⁽⁴⁾

En diferentes países han sido desarrollados varios modelos de riesgo, que usualmente estiman el riesgo para un período de 10 años, usando factores de riesgo convencionales, como Framingham clásica y por categorías, REGICOR para España, SCORE, PROCAM y recientemente las tablas de la OMS para diferentes regiones del mundo, aprobadas para su uso en la población cubana, entre otras.⁽⁵⁾

La limitación más importante de las funciones de riesgo es su baja sensibilidad, de manera que gran parte de los acontecimientos cardiovasculares se presentan en el grupo de la población con riesgo intermedio, otras limitaciones son: necesitan recalibrarse cuando se usan directamente a poblaciones distintas de donde fueron diseñados, consideran un número limitado de factores de riesgo tradicionales o clásicos, sin incluir factores metabólicos y biomarcadores de inflamación, a pesar de que pueden ser un riesgo independiente de complicaciones vasculares, la definición del nivel de riesgo y los puntos de corte es arbitraria y no tienen en cuenta que el riesgo es un continuo, más que un proceso aditivo.^(1,4,5,6,7)

La obtención de conocimiento en un dominio de aplicación a partir de datos presentes es una práctica muy actual, el aprendizaje automático (*Machine Learning*, ML) es una subárea de la inteligencia artificial, que tiene como objetivo diseñar algoritmos que realizan el aprendizaje en los datos disponibles, son capaces de procesar grandes cantidades de datos y transformarlos en conocimiento, lo que permite tomar decisiones o acciones inteligentes.^(8,9) Varios estudios muestran que los modelos de predicción cardiovascular diseñados con algoritmos de machine learning exhiben un desempeño superior cuando son comparados con los modelos tradicionales y pueden predecir el riesgo cardiovascular eficazmente.^(10,11,12)

Hasta el momento, no se dispone de un modelo de predicción del riesgo de enfermedad

cardiovascular en la atención primaria de salud a partir de datos de seguimiento de cohortes de poblaciones cubanas.

El objetivo de la investigación fue diseñar un modelo predictivo para estimar el riesgo de enfermedad cardiovascular basado en técnicas de inteligencia artificial.

>>> MÉTODOS

La población de estudio de 35 a 74 años perteneciente al policlínico docente “José Ávila Serrano”, de Velasco, Holguín, en Cuba.

Los datos utilizados en la investigación provienen de la cohorte del estudio PredRCCG_aps, que incluyó 857 pacientes del sexo femenino y 776 del masculino, con edades de 35 a 74 años sin antecedentes de enfermedad cardiovascular a la inclusión, seleccionados aleatoriamente, que fue-

ron seguidos durante $10 \pm 1,3$ años, desde enero 2008 hasta diciembre de 2019; se realizaron evaluaciones clínicas anuales a cada paciente, la variable resultado fue el desarrollo de un evento cardiovascular mortal o no.⁽¹³⁾

La base de datos incluyó las siguientes variables predictoras: edad, sexo, hábito de fumar, antecedentes personales de diabetes mellitus e hipertensión arterial, índice de masa corporal, circunferencia abdominal, tensión arterial sistólica y diastólica, colesterol total, lipoproteínas de baja densidad colesterol (LDLc), lipoproteínas de alta densidad colesterol (HDLc), triglicéridos, ácido úrico, glicemia en ayunas y proteína c reactiva de alta sensibilidad (pcr-as o pcr-hs, high sensitivity).⁽¹³⁾

Fue utilizada WEKA, un software de código abierto, acrónimo de *Waikato Environment for Knowledge Analysis*, versión 3.9.5, que implementa machine learning y que tiene como objetivo brin-

Análisis multidisciplinarios de alta complejidad.

Clínico humano
Bromatológico
Veterinario
Agronómico
Bioanalítica
Industrial y Medio Ambiente

De Bahía Blanca para todo el país.

IACA
LABORATORIOS
www.iaca.com.ar

dar algoritmos de minería de datos.⁽¹⁴⁾

Cada caso quedó conformado por 16 atributos o rasgos predictores y un rasgo de decisión, Evento, donde se indica si el paciente desarrolló o no el evento cardiovascular, cuyo dominio es {Evento= SI, Evento= NO}, con dos clases de decisión: 1531 pacientes, no desarrollaron el evento cardiovascular y 102 sí.

La etapa Preprocesamiento consistió en la preparación y limpieza, se convierten o discretizan los datos validándolos, para un manejo estandarizado y homogéneo, permitiendo obtener mejores resultados según las características del algoritmo a utilizar, aplicándose filtros supervisados sobre los atributos o las instancias.

Para balancear ambas clases, fue aplicado el algoritmo supervisado dirigido a las instancias, SMOTE (*Synthetic Minority Over-sampling TEchnique*).⁽¹⁵⁾

Se emplearon técnicas de selección de atributos integradas en WEKA para obtener un subconjunto más reducido de variables significativas: CfsSubsetEval, con el método de búsqueda *Best First*, *Wrapper SubsetEval* con J48 y *Correlation Attribute Eval*, con el método Ranker.

Se ejecutaron los siguientes algoritmos de clasificación de WEKA: el algoritmo de clasificación de reglas JRip, los meta algoritmos de clasificación *Attribute Selected Classifier*, usando como algoritmos clasificadores el J48 y el algoritmo de clasificación de funciones *Multilayer Perceptron*, en ambos usando como evaluador el Cfs-SubsetEval y como método de búsqueda el *Best-First*.

Se compararon los modelos obtenidos y se aplicaron las métricas más usadas para clases desbalanceadas, basadas en la matriz de confusión como: la tasa de verdaderos positivos (TPrate, en inglés) y tasa de falsos positivos (FPrate, en inglés), en la clase de interés, que fue la minoritaria, es decir Evento=SI, el área bajo la Curva de Operación del Receptor (ROC) y la curva Precisión_Recall (PRC).

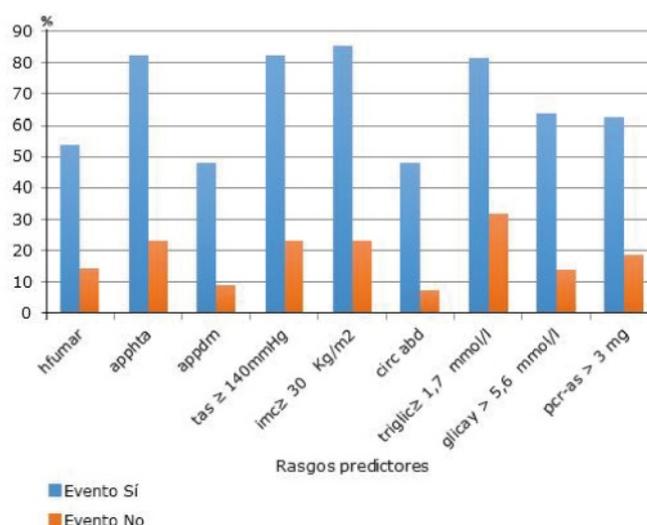
Al disponer de una sola base de datos, se utilizó el conjunto de datos para entrenamiento y para la validación se utilizó la validación cruzada con 10 particiones (10 *folds cross-validation*)

La investigación, desde el punto de vista ético, cumplió con la Declaración de Helsinki,⁽¹⁶⁾ se protegieron los derechos y confiabilidad de las personas en el estudio, la investigación fue revisada y aprobada por el Comité de Ética de la Investigación Clínica de la institución.

>>> RESULTADOS

Se emplearon técnicas de visualización de información para mostrar la distribución de los rasgos predictores seleccionados entre las 2 clases de decisión, la figura 1 refleja que la proporción fue superior en los pacientes que desarrollaron el evento cardiovascular, fue más marcada la diferencia en el antecedente de hipertensión arterial, las cifras de tensión arterial sistólicas mayor de 140 mm Hg, el índice de masa corporal mayor de 30 Kg/m² y de proteína c reactiva de alta sensibilidad (pcr-as) mayor de 3 mg/dl, se consideraron en los rasgos los valores considerados en riesgo o elevados, como las clases muestran un alto desbalance, el porcentaje fue superior en la clase minoritaria, Evento SÍ, que es la clase de interés.

>> Fig. 1- Proporción de rasgos predictores entre los pacientes que desarrollaron o no el evento cardiovascular.





Diestro

De Argentina al mundo.

Nuestros instrumentos informan millones de analitos al año alrededor del mundo, sin embargo cada diagnóstico es único.



PRESENTES EN MEDICA 2023

Düsseldorf, Alemania | **13-16 de noviembre**

diestroweb.com | (011) 4709 7707 |   @Diestro.Ar

Los algoritmos clasificadores (tabla 2) nos muestran que son efectivos para generar el modelo predictivo, al comparar las métricas en el test de entrenamiento y validación cruzada, en la clase minoritaria Evento = SI, el mejor desempeño se evidencia con el algoritmo *Multilayer Perceptron*.

>> **Tabla 2-** Evaluación del desempeño de los algoritmos clasificadores en los datos de entrenamiento y de validación cruzada

Datos	Algoritmo Clasificador	Tasa Verdaderos Positivos (%)	Tasa Falsos Positivos (%)	Área ROC	Área PRC
Datos de entrenamiento	JRip	96,1	2,8	0,975	0,961
	<i>Attribute Selected Classifier</i> 148	94,2	2,0	0,972	0,949
	<i>Attribute Selected Classifier/Multilayer Perceptron</i>	97,9	1,2	0,991	0,982
Validación cruzada con 10 particiones	JRip	94,8	3,6	0,965	0,942
	<i>AttributeSelected Classifier</i> 148	93,5	2,5	0,957	0,909
	<i>AttributeSelectedClassifier/Multilayer Perceptron</i>	95,2	3,3	0,987	0,970

>>> DISCUSIÓN

La obtención de conocimiento en un dominio de aplicación a partir de datos presentes se convierte en una práctica muy actual, han sido desarrolladas numerosas técnicas de descubrimiento del conocimiento desde la inteligencia artificial; los términos aprendizaje automático, minería de datos y de textos, nos resultan familiares, el conocimiento generado por ellas puede utilizarse para la construcción de modelos inteligentes y aumentar el conocimiento existente en dominios de aplicación como la predicción del riesgo de enfermedad cardiovascular.⁽¹⁶⁾

Con pocos atributos, los más significativos o los que tienen más peso, se obtiene una mejor clasificación y se evita el sobreajuste (*overfitting*) del modelo diseñado y que este sea más comprensible, ignorándose los atributos no significativos.⁽¹⁷⁾



+25 años de trayectoria

acompañando a los principales laboratorios de análisis clínicos de Argentina. **Sostenemos una política de innovación y excelencia continua** en materia de búsqueda, selección y comercialización de equipamiento médico de última generación.

HACÉ TU CONSULTA
ESCANEANDO EL QR



¡Conocé nuestra propuesta!

ventas@gematec.com.ar

www.gematec.com.ar



NUESTRO DESAFÍO

Ofrecer tecnologías innovadoras para hacer más eficiente cada área del laboratorio.

NUESTRA PASIÓN

Acompañar al laboratorio en cada etapa de su evolución.

NUESTRO COMPROMISO

Ser socios en el diagnóstico.

NUESTROS VALORES

Cordialidad, innovación y excelencia.

Dentro de los modelos de predicción, los árboles de decisión son los más utilizados por una serie de razones que los hacen especialmente atractivos como: la sencillez del modelo, la amplitud de implementaciones que existen, la rapidez de clasificación de nuevos patrones, la posibilidad de representarlos gráficamente aportando así una explicación de la división efectuada, la fácil interpretación, en caso de no ser excesivamente grandes y la posibilidad de obtener las reglas por las cuales asignamos cada una de las probabilidades.⁽¹⁷⁾

Como apreciamos, el mejor desempeño se alcanzó con el modelo predictivo creado por el algoritmo *Multilayer Perceptron*, en la red neuronal el aprendizaje se produce por medio de la retropropagación, el algoritmo compara el resultado de la capa de salida con el resultado deseado y asume que el error en la unidad de salida se debe a errores en las unidades conectadas con ella, y para corregirse realiza ajustes en los pesos asignados en la red desde la capa de salida hasta la capa de entrada, hacia atrás, de ahí el término de retropropagación.^(18,19)

Una de las limitaciones para la implementación práctica de estos sistemas de inteligencia artificial de tanto potencial como los modelos algorítmicos basados en redes neuronales artificiales, es que no hay suficiente comprensión de las reglas que generan, dificultando la interpretación, comportándose como cajas negras (*black boxes*, en inglés).^(8,20,21)

El aprendizaje automático es parte fundamental en un proceso de análisis predictivo, ya que proporciona las técnicas de análisis de datos mediante las cuales se pueden descubrir relaciones entre variables que en un principio pueden parecer insignificantes, o cuya significación es incierta, pero que tras la aplicación de estas técnicas puede descubrirse la trascendencia de las mismas.⁽²²⁾

La minería de datos es una ciencia experimental, existen una gran cantidad de técnicas de preprocesamiento, así como algoritmos de clasificación, pero no hay un método único, universal y mejor para un problema específico, necesitamos encontrar los métodos de aprendizaje que traba-

jen mejor ante un determinado problema.^(8,23,24)

No encontramos, en la revisión bibliográfica, modelos predictivos sobre riesgo cardiovascular en Cuba basados en la inteligencia artificial, lo que nos limitó hacer comparaciones, el modelo obtenido incorpora predictores noveles como la circunferencia abdominal, los triglicéridos, la glicemia en ayunas y la proteína c reactiva de alta sensibilidad, no contemplados en los modelos tradicionales.

Estos predictores noveles están involucrados en los mecanismos de resistencia a la insulina, inflamación y aterosclerosis, en ocasiones subclínica, muchas veces, el evento cardiovascular es mortal o no, la primera manifestación clínica.

La resistencia a la insulina incluye una variedad de trastornos en el metabolismo de lípidos, como aumento del nivel de triglicéridos plasmáticos y disminución de colesterol de HDL; los trastornos de la glucemia en ayuna, intolerancia a carbohidratos y, por último, la *diabetes mellitus* tipo 2, asociados a la obesidad, que incluye la abdominal, no siempre considerada por el médico en la atención primaria; la hipertensión arterial, teniendo como base un estado proinflamatorio crónico con niveles elevados de pcr-as relacionados directamente con la magnitud de la resistencia a la insulina.⁽¹³⁾

La inflamación desempeña un papel determinante en el inicio y progresión del proceso aterotrombótico, así como en sus manifestaciones clínicas, la determinación de los niveles de pcr-as determinados mediante un método inmunoturbidimétrico de alta sensibilidad permite detectar la inflamación de bajo grado característica de la aterosclerosis, su estabilidad por largo tiempo durante el almacenamiento, larga vida media y carencia de variación diurna, son factores que han contribuido a que este sea el biomarcador inflamatorio más extensamente evaluado en la actualidad.^(7,25)

Una limitación del estudio fue que no se cuenta con registros informatizados de datos de los pacientes en la atención primaria de salud y por las condiciones logísticas no se pudo disponer de



Analizador Multiparamétrico

Totalmente Automatizado

- Dispositivo individual de un solo uso que contiene todos los reactivos necesarios para realizar el ensayo.
- Capacidad multiparamétrica: Procesa hasta 30 diferentes pruebas por corrida.
- La velocidad permite obtener resultados simultáneos de diferentes paneles.
- El primer resultado se obtiene antes de 90 minutos.
- Volumen de muestra:
 - La muestra se dispensa manualmente. ELISA:
 - Mínimo de muestra 60 uL.
 - Fijación de complemento:
 - Mínimo de muestra 120 uL.



Enfermedades Infecciosas

ADENOVIRUS IgA
 ADENOVIRUS IgG
 BORDETELLA PERTUSSIS IgA
 BORRELIA IgG
 BORRELIA IgM
 CHIKUNGUNYA IgG
 CHIKUNGUNYA IgM
 CHLAMYDOPHILA PNEUMONIAE IgA
 CHLAMYDOPHILA PNEUMONIAE IgG
 CHLAMYDOPHILA PNEUMONIAE IgM
 CLOSTRIDIUM DIFFICILE A/B TOXINS
 CLOSTRIDIUM DIFFICILE GDH
 CYTOMEGALOVIRUS IgG
 CYTOMEGALOVIRUS IgG AVIDITY
 CYTOMEGALOVIRUS IgM
 DENGUE IgG
 DENGUE IgM
 DIPHTERIA IgG
 ECHINOCOCCUS IgG
 EPSTEIN-BARR EARLY ANTIGEN IgG
 EPSTEIN-BARR EARLY ANTIGEN IgM
 EPSTEIN-BARR EBNA IgG
 EPSTEIN-BARR VCA IgG
 EPSTEIN-BARR VCA IgM II
 HELICOBACTER PYLORI IgA

HELICOBACTER PYLORI IgG
 HSV1 SCREEN
 HSV2 SCREEN
 HERPES SIMPLEX 1 IgG Recombinant
 HERPES SIMPLEX 1+2 IgM
 HERPES SIMPLEX 2 IgG Recombinant
 INFLUENZA A IgA
 INFLUENZA A IgG
 INFLUENZA B IgA
 INFLUENZA B IgG
 LEGIONELLA PNEUMOPHILA
 LEGIONELLA PNEUMOPHILA 1 IgG
 LEGIONELLA PNEUMOPHILA 1-6 IgG
 LEGIONELLA PNEUMOPHILA IgM
 LEGIONELLA URINARY ANTIGEN
 MEASLES IgG
 MEASLES IgM
 MUMPS IgG
 MUMPS IgM
 MYCOPLASMA PNEUMONIAE IgA
 MYCOPLASMA PNEUMONIAE IgG
 MYCOPLASMA PNEUMONIAE IgM
 Parvovirus B19 IgG
 Parvovirus B19 IgM
 POLIOVIRUS IgG

RESPIRATORY SYNCYTIAL IgA
 RESPIRATORY SYNCYTIAL IgG
 RUBELLA IgG AVIDITY
 RUBELLA IgG
 RUBELLA IgM
 SYPHILIS SCREEN RECOMBINANT
 TETANUS IgG
 TICK-BORNE ENCEPHALITIS VIRUS
 TICK-BORNE ENCEPHALITIS VIRUS IgM
 TIROGLOBULIN HIGH SENSITIVITY
 TOSCANA VIRUS IgG
 TOSCANA VIRUS IgM
 TOXOCARA IgG
 TOXOPLASMA IgA
 TOXOPLASMA IgG AVIDITY
 TOXOPLASMA IgG
 TOXOPLASMA IgM
 TRACHOMATIS IgA
 TRACHOMATIS IgG
 TREPONEMA IgG
 TREPONEMA IgM
 VARICELLA IgG
 VARICELLA IgM
 25 OH VITAMIN D TOTAL

Autoinmunidad

ANA-8
 ANA-SCREEN
 ENA-6 S
 SM
 SS-A
 SS-B
 Scl-70
 Cenp-B
 Jo-1
 ds-DNA-G
 ds-DNA-M
 snRNP-C
 U1-70 RNP
 anti-CCP
 RF-G
 RF-M
 CALPROTECTIN
 CALPROTECTIN K
 CARDIOLIPIN-G
 CARDIOLIPIN-M
 BETA 2-GLYCOPROTEIN-G
 BETA 2-GLYCOPROTEIN-M
 DEAMIDATED GLIADIN-A
 DEAMIDATED GLIADIN-G
 GLIADIN-A

GLIADIN-G
 tTG-A
 tTG-G
 ASCA-A
 ASCA-G
 GBM
 MPO
 PR3
 TG
 a-TG
 a-TPO
 AMA-M2
 LKM-1
 INSULIN
 INTRINSIC FACTOR
 FSH
 LH
 PRL
 TSH
 ft4
 ft3
 TOTAL IgE

Fijación del Complemento

BORRELIA IgG
 BRUCELLA
 COXACKIE VIRUS A MIX
 COXACKIE VIRUS B MIX
 ECHO VIRUS N MIX
 ECHO VIRUS P MIX
 LEPTOSPIRA MIX
 LISTERIA MONOCYTOGENES
 PARAINFLUENZA MIX
 Q FEVER



BIODIAGNOSTICO

Av. Ing. Huergo 1437 P.B. "I" | C1107APB | CABA | Argentina | Tel./Fax: +5411 4300-9090
 info@biodiagnostico.com.ar | www.biodiagnostico.com.ar

una muestra más amplia de la población, actualmente se implementa el modelo para su aplicación práctica.

Los resultados pueden ayudar a mejorar la predicción de la enfermedad cardiovascular y muestran las capacidades de las técnicas de inteligencia artificial ante la gran cantidad de datos que no puede ser procesada y analizada con métodos tradicionales y nos lleva a hablar de la inteligencia artificial como una nueva forma de generar conocimiento; la medicina y la salud se han convertido en unas de las más prometedoras áreas para la aplicación de los sistemas de inteligencia artificial.⁽²⁶⁾

En conclusión, fue diseñado un modelo predictivo mediante técnicas de inteligencia artificial, lo que constituye un valioso recurso orientado a la prevención de las enfermedades cardiovasculares en la atención primaria de salud.

>>> APOORTE CIENTÍFICO

Es una de las primeras investigaciones que obtiene un modelo predictivo de riesgo de enfermedad cardiovascular usando técnicas de inteligencia artificial, basado en una cohorte de una población cubana e incorpora predictores noveles relacionados con la patogenia de la aterosclerosis y dota al médico de familia de una herramienta útil para la prevención de la enfermedad cardiovascular en la atención primaria de salud.

>>> CONFLICTO DE INTERESES

Los autores declaran que no existe conflicto de intereses.

>>> CONTRIBUCIÓN DE LOS AUTORES

Conceptualización: Jorge Baudilio Vega Abascal.
Curación de datos: Jorge Baudilio Vega Abascal.
Análisis formal: Jorge Baudilio Vega Abascal, Alberto Rubén Piriz Assa. Investigación: Jorge Baudilio Vega Abascal, Alberto Rubén Piriz Assa
Administración del proyecto: Jorge Baudilio Vega Abascal.
Supervisión: Jorge Baudilio Vega Abascal.

Visualización: Jorge Baudilio Vega Abascal, Alberto Rubén Piriz Assa.

Redacción – borrador original: Jorge Baudilio Vega Abascal.

Redacción – revisión y edición: Jorge Baudilio Vega Abascal, Alberto Rubén Piriz Assa, Diego Nápoles Riaño.

>>> FINANCIAMIENTO

Policlínico Docente “José Ávila Serrano”. Velasco, Holguín, Cuba.

>>> REFERENCIAS BIBLIOGRÁFICAS

- 1.WHO CVD Risk Chart Working Group. World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. *Lancet Glob Health*. 2019;11-11. DOI: [http://dx.doi.org/10.1016/S2214-109X\(19\)30318-3](http://dx.doi.org/10.1016/S2214-109X(19)30318-3)
- 2.Vega J, Guimará M, Vega L. Riesgo cardiovascular, una herramienta útil para la prevención de las enfermedades cardiovasculares. *Rev Cubana Med Gen Integr*. 2011 [acceso 14/01/2020];27(1):91-7. Disponible en: <http://scielo.sld.cu/pdf/mgi/v27n1/mgi10111.pdf>
- 3.Anuario Estadístico de Salud. MINSAP. Cuba. 2020. Disponible en: <http://files.sld.cu/bvscuba/files/2021/08/Anuario-Estadistico-Español-2020-Definitivo.pdf>
- 4.Vega J, Guimará MR, Garces Y, Vega LA, Rivas M. Predicción de riesgo coronario y cardiovascular global en la atención primaria de salud. *CM*. 2015 [acceso 14/03/2017];19(2):202-11. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1560-43812015000200003&lng=es
- 5.Pizarro R, Masson W. Estudios prospectivos poblacionales: las fortalezas de las cohortes históricas. *Acta Gastroenterol Latinoam*. 2020 [acceso 24/03/2021];50(4):382-7. Disponible en: <https://actagastro.org/estudios-prospectivos-poblacionales-las-fortalezas-de-las-cohortes-historicas/>
- 6.Elosua R. Las funciones de riesgo cardiovascular: utilidades y limitaciones. *Rev Esp Cardiol*. 2014 [acceso 14/01/2015];67(2):77-9. Disponible en: <http://www.revespcardiol.org/es/pdf/g0267565/S300/>
- 7.Shrivastava AK, Singh HV, Raizada A, Singh SK. C-reactive protein, inflammation and coronary heart disease. *The Egyptian Heart Journal*. 2015;67:89-97. DOI: <http://dx.doi.org/10.1016/j.ehj.2014.11.005>
- 8.Witten I, Frank E, Hall M. *Data Mining: Practical Machine Learning Tools and Techniques*. Tercera. San Francisco, California: Morgan Kaufmann, Elsevier; 2011. Disponible en: https://www.academia.edu/23331284/Data_Mining_Practical_Machine_Learning_Tools_and_Techniques_3rd_Edition
- 9.Mazon B, Pinta M, Redovran F. Desarrollo de competencias en Minería de Datos, una experiencia didáctica. En: *Sistematización de experiencias educativas innovadoras*. Primera Edición. Machala, Ecuador: UTMACH; 2020. p. 460. Disponible en: https://www.researchgate.net/publication/343558044_Desarrollo_de_competencias_en_Mineria_de_Datos_una_experiencia_didactica
- 10.Bannister C, Halcox J, Currie C, Preece A, Spasic I. A genetic programming approach to development of clinical prediction models: A case study in symptomatic cardiovascular disease. *PLoS ONE*. 2018;13(9):e0202685. DOI: <https://doi.org/10.1371/journal.pone.0202685>
- 11.Dimopoulos AC, Nikolaidou M, Caballero F, Engchuan W, Sanchez-Niubo A, Arndt H, et al. Machine learning methodologies versus cardiovascular risk scores, in predicting disease risk. *BMC Medical Research Methodology*. 2018;18(179):2-11. DOI: <https://doi.org/10.1186/s12874-018-0644-1>
- 12.Kim JO, Jeong YS, Kim JH, Lee JW, Park D, Kim HS. Machine Learning-Based Cardiovascular Disease Prediction Model: A Cohort Study on the Korean National Health Insurance Service Health Screening Database. *Diagnostics*. 2021 [acceso 14/07/2021];11(943). Disponible en: https://www.researchgate.net/publication/351854383_Machine_LearningBased_Cardiovascular_Disease_Prediction_Model_A_Cohort

_Study_on_the_Korean_National_Health_Insurance_Service_Health_Screening_Database

13.Vega J, Piriz A, Guimará M, Vega L, Caballero L, Nápoles D. PredRCG_aps: Predicción del riesgo cardiovascular global en la atención primaria de salud. CCM. 2022 [acceso 14/07/2022];26(2). Disponible en : <http://www.revcoemed.sld.cu/index.php/coemed/article/view/4202/2136>

14.WEKA 3.9.5. Disponible en: <http://www.cs.waikato.ac.nz/ml/WEKA/>

15.Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling TEchnique. Artificial Intelligence Research. 2002 [acceso 14/07/2003]; 341-78. Disponible en: <https://www.jair.org/index.php/jair/article/view/10302>

16.World Medical Association. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. JAMA. 2013 [acceso 14/01/2015];310(20):2191-4. DOI: <https://doi.org/10.1001/jama.2013.281053>

17.García MM, Rodríguez Y, Hernández A, Bello B, Filiberto Y, Rosete A, et al. Adquisición de conocimiento sobre la letalidad de la COVID-19 mediante técnicas de inteligencia artificial. Anales de la Academia de Ciencias de Cuba. 2020 [acceso 14/01/2021];10(3). Disponible en: <http://revistaccuba.sld.cu/index.php/revacc/article/view/891>

17.Arrejo Calviño HA. Métodos para la mejora de predicciones en clases desbalanceadas en el estudio de bajas de clientes (CHURN). La Coruña, España: Universidad de Vigo; 2016 [acceso 24/08/2021]. Disponible en : http://eio.usc.es/pub/mte/descargas/proyectosfinmaster/proyecto_1469.pdf

18.Ayala E, López R, Menéndez V. Modelos predictivos de riesgo académico en carreras de computación con minería de datos educativos. RED. 2021;21(26). DOI: <https://doi.org/10.6018/red.463561>

19.Gironés J, Casas J, Minguillón J, Caihuellas R. Minería de datos: modelos y algoritmos. Primera. Barcelona, España: UOC; 2017 [acceso 14/06/2020]. Disponible en: [https://dokumen.site/file/mineria-de-datos-modelos-y-](https://dokumen.site/file/mineria-de-datos-modelos-y-algoritmospdf-a5b39f02c4ae11)

[algoritmospdf-a5b39f02c4ae11](https://dokumen.site/file/mineria-de-datos-modelos-y-algoritmospdf-a5b39f02c4ae11)

20.Hailesilassie T. Rule Extraction Algorithm for Deep Neural Networks: A Review. International Journal of Computer Science and Information Security. 2016 [acceso 14/06/2020];14(7):376-81. Disponible en: <https://arxiv.org/abs/1610.05267>

21.Barredo Arrieta A, Díaz - Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion. 2020;58:82-115. DOI: <https://doi.org/10.1016/j.inffus.2019.12.012>

22.Espino Timón C. Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso [tesis]. España: Universidad de Cataluña; 2017 [acceso 14/06/2020]. Disponible en: <https://core.ac.uk/display/77515842>

23.Vaquero de Miguel M. Aproximaciones a la Explicación de Decisiones Algorítmicas: Inteligencia Artificial Explicable [tesis]. España: Universidad Politécnica de Madrid; 2020 [acceso 14/02/2021]. Disponible en: <http://oa.upm.es/63381>

24.González G. Agente Inteligente para Análisis de Datos de Carreteras [tesis]. España: Universidad Politécnica de Madrid; 2021 [acceso 14/11/2021]. Disponible en : https://oa.upm.es/68712/1/TFG_GONZALO_GONZALEZ_MORO.pdf

25.Sproston NR, Ashworth JJ. Role of C-Reactive Protein at Sites of Inflammation and Infection. Front. Immunol. 2018;9:754. DOI: <https://doi.org/10.3389/fimmu.2018.00754>

26.Romeo CM, Lazcoz G. Inteligencia artificial aplicada a la salud: ¿qué marco jurídico? 2020 [acceso 14/06/2021]. Disponible en: <https://www.fundacionmercksalud.com/wp-content/uploads/2020/03/1.3-IA-APLICADA-A-LA-SALUD.-Carlos-M.-Romeo-Guillermo-Lazcoz.pdf>

Biocientífica

Línea Schep

Fácil detección por PCR en tiempo real

En Biocientífica celebramos 40 años renovando el compromiso de hacer biotecnología para mejorar la calidad de vida. Conocé la línea completa de productos Schep para diagnóstico molecular, 100% I+D y producción argentina.

Schep Dengue Screen

Schep SARS-CoV-2 Multi-FAST



Schep Flurona Type

Schep BCR-ABL1 p210 Cuantiplex



#ADNdeInnovación

Conocé la línea completa de productos:

Detección del virus dengue: Schep Dengue Screen (IVD)

Monitoreo de leucemia:
Schep BCR-ABL1 p210 Cuantiplex (IVD)
Para leucemia mieloide crónica (LMC), leucemia mieloide aguda (LMA) y leucemia linfoblástica aguda (LLA)

Detección de los virus SARS-CoV-2, Influenza A y B:
Schep Flurona Type (IVD)

Detección de SARS-CoV-2:
Schep SARS-CoV-2 Multi-FAST (IVD)

Los kits incluyen todos los componentes necesarios para realizar la PCR en Tiempo Real.

Más información escanée acá:

+54 11 4857-5005
biocientifica.com.ar
ventas@biocientifica.com.ar

¡Seguinos!   

