www.komputersapiens.org



KOMPUTER SAPIENS

Revista de Divulgación de la Sociedad Mexicana de Inteligencia Artificial

Año 15
Volumen 2
Mayo - Agosto 2023

PREDICCIÓN DEL **RENDIMIENTO ACADÉMICO**

ATENUANDO **RUIDO** EN SEÑALES DE AUDIO

ECUACIONES DIFERENCIALES Y SUS APLICACIONES EN **LA VIDA REAL**

IDENTIFICACIÓN EFECTIVA DE **PATOLOGÍAS GINECOLÓGICAS**

ANOTACIÓN AUTOMÁTICA DE **IMÁGENES DE PLANTAS**

COMBATIR LA **PESCA ILEGAL** CON RECONOCIMIENTO DE PATRONES

RETOS DE LA IA PARA NO COMPROMETER LA **PRIVACIDAD DE LOS DATOS**

DETECCIÓN DE **NOTAS DE IDEACIÓN** SUICIDA

LA IA EN ASTRONOMÍA DE MULTIMESAJEROS, HECHOS Y PERSPECTIVAS.

ESTUDIO DE **ENFERMEDADES NEURODEGENERATIVAS** CON IA

IDENTIFICACIÓN DE ESPECIES DE **PLANTAS**

ISSN 2007-0691

9 772007 069007

©Komputer Sapiens, Año XV Volumen II, mayo-agosto 2023, es una publicación cuatrimestral de la Sociedad Mexicana de Inteligencia Artificial, A.C., con domicilio en Ezequiel Montes 56 s/n, Fracc. los Pilares, Metepec, Edo. de México, C.P. 52159, México, http://www.komputersapiens.org, correo electrónico: editorial@komputersapiens.org, tel. +52 (833)357.48.20 ext. 3024, fax +52 (833) 215.85.44. Impresa por Sistemas y Diseños de México S.A. de C.V., calle Aragón No. 190, colonia Álamos, delegación Benito Juárez, México D.F., C.P. 03400, México, se terminó de editar el 13 de mayo de 2023.
Reserva de derechos al uso exclusivo número 04-2009-111110040200-102 otorgado por el Instituto Nacional de Derechos de Autor. ISSN 2007-0691.
Los artículos y columnas firmados son responsabilidad exclusiva de los autores y no reflejan necesariamente los puntos de vista de la Sociedad Mexicana de Inteligencia Artificial. La mención de empresas o productos específicos en las páginas de Komputer Sapiens no implica su respaldo por la Sociedad Mexicana de Inteligencia Artificial.
Queda estrictamente prohibida la reproducción total o parcial por cualquier medio, de la información aquí contenida sin autorización por escrito de los editores.
Komputer Sapiens es una revista de divulgación en idioma español de temas relacionados con la inteligencia artificial.
Creada en LaTeX, con la clase **papertex** disponible en el repositorio *CTAN*: Comprehensive TeX Archive Network, http://www.ctan.org/
Indizada en el IRMDCT de CONACYT y en Latindex.

Directorio SMIA

Presidente	Ildar Batyrshin
Vicepresidente	Oscar Herrera Alcántara
Secretario	Francisco Hiram Calvo Castro
Tesorera	María de Lourdes Martínez Villaseñor
Vocales:	Obdulia Pichardo Lagunas
	Sabino Miranda Jiménez
	Roberto A. Vázquez Espinoza de los Monteros
	Hiram Erendin Ponce Espinosa
	Noé Alejandro Castro Sánchez
	Iris Iddaly Méndez Gurrola
	Omar Montaño Rivas
	Néstor Velasco Bermeo
	Gustavo Arroyo Figueroa
	Sofía Natalia Galicia Haro
	Juan Crisoforo Martínez Miranda
	Iván Olvera Rodríguez
	Leobardo Morales Tiburcio
	Gilberto Ochoa Ruiz
	David Eduardo Pinto Avendaño
	Antonio Marín Hernández
	César Jesús Núñez Prado

Komputer Sapiens

Director general	Ildar Batyrshin
Editora en jefe	Karina Mariela Figueroa Mora
Editor científico	Héctor Gabriel Acosta Mesa
Editores asociados	Marco Antonio Aguirre Lam
	Laura Cruz Reyes
Editores invitados	Marisol Flores Garrido
	Luis Miguel García Velázquez
Coordinadora de producción	Viridiana Mena Gómez
e-Tlakuilo	Karina Figueroa
	Rosa María Valdovinos
Estado del IArte	Reyna Carolina Medina Ramírez
	Jorge Rafael Gutiérrez Pulido
Sakbe	Claudia G. Gómez Santillán
	Laura Cruz Reyes
	Laura Gómez Cruz
IA & Educación	J. Julieta Noguez Monroy
	María Lucía Barrón Estrada
	María Yasmín Hernández Pérez
Deskubriendo Konocimiento	Alejandro Guerra Hernández
	Gildardo Sánchez Ante
Asistencia técnica	Alan G. Aguirre Lam
Corrección de estilo	Gilberto Rivera Zárate
	Marcela Quiroz Castellanos
	Sonia Lilia Mestizo Gutiérrez
	Guillermo de Jesús Hoyos Rivera
	Susana Aurora Velasco Montiel
	Nicandro Cruz Ramírez
	Leonardo Romero Muñoz
	Nelson Rangel Valdez
	Claudia Leticia Díaz González
	Luis Miguel García Velázquez
Portada	María Fernanda Hernández Arroyo

Directores Fundadores

Carlos Alberto Reyes García
Ángel Kuri Morales

Comité Editorial

Ildar Batyrshin
Oscar Herrera Alcántara
Jesús Favela Vara
Sofía Natalia Galicia Haro
Miguel González Mendoza
Raúl Monroy Borja
Eduardo F. Morales Manzanares
Leonardo Garrido Luna
Carlos Alberto Reyes García
Antonio Sánchez Aguilar
Luis Enrique Sucar Succar
Ángel Kuri Morales
José A. Martínez Flores
Juan Manuel Ahuactzin Larios
Manuel Montes y Gómez
Ofelia Cervantes Villagómez
Alexander Gelbukh
Grigori Sidorov
Laura Cruz Reyes
Ramon Brena Pinero
Juan Humberto Sossa Azuela
Félix Castro

Árbitros

Víctor Anaya-Muñoz
Carlos Astengo Noguez
Daniele Colosi
Rodrigo Cuéllar Hidalgo
Guillermo de Anda-Jáuregui
María del Río Francos
Karina Figueroa Mora
Marisol Flores-Garrido
Julián Fresán-Figueroa
Rita Fuentes-Aguilar
Gibran Fuentes-Pineda
Sandra García
Jorge González Ordiano
Héctor Alonso Guzmán Gutiérrez
Mónica Hernández
Delia Irazú Hernández Farías
Maribel Hernández Rosales
María Lárraga
Karen Lopez Castro
Adrian P Lopez-Monroy
Arturo Lopez Pineda
Beatriz Luna Olivera
Dulce Magaña
Leonardo Ignacio Martinez Sandoval
Adriana Menchaca-Mendez
Ricardo Menchaca Mendez
Rolando Menchaca-Mendez
Ivan Vladimir Meza Ruiz
Antonio Neme
Marco Peña Díaz
Andrea Quintanilla
Miguel Raggi
Gabriela Ramírez-de-la-Rosa
Caleb Rascon
Aaron Junior Rocha Rocha
Javier Rodríguez Herrejon
Christian Rubio-Montiel
Elias Ruiz
Ivete Sanchez
Gildardo Sanchez-Ante
Sergio Tinoco-Martínez

Contenido

Identificación efectiva de patologías ginecológicas aplicando Random Forest. Caso: Hospital de la mujer, Tabasco, México
Pág. 8 ⇒Adriana-Guadalupe López-Ramírez, Eddy Sánchez-DelaCruz y Cecilia-Irene Loeza-Mejía

Técnicas de aprendizaje automático para la predicción del rendimiento académico
Pág. 12 ⇒ Andrés Rico Páez

Redes Neuronales Convolucionales con Atención para Atenuar Ruido en Señales de Audio
Pág. 16 ⇒ Noel Zacarias-Morales, José Adán Hernández-Nolasco, Pablo Pancardo y J. Arturo Olvera-López

Redes neuronales físicamente informadas para aproximar la solución de ecuaciones diferenciales y sus aplicaciones en la vida real
Pág. 21 ⇒ Juan Daniel Meshir Vargas, Edgar Alejandro Guerrero Arroyo y Abel Palafox González

La inteligencia artificial como mecanismo de investigación en astronomía de multimesajeros, hechos y perspectivas
Pág. 27 ⇒ Alejandro Casallas-Lagos, Javier M. Antelis, Claudia Moreno y Michele Zanolin

Aplicación de la inteligencia artificial para el estudio de enfermedades neurodegenerativas
Pág. 33 ⇒ Alejandro Mateos-Moreno, Eddy Sánchez-DelaCruz y Luis Eduardo Maldonado-López

La inteligencia artificial como apoyo en la identificación de especies de plantas
Pág. 38 ⇒ Juan Augusto Campos-Leal, Inés F. Vega-López, Gerardo Beltrán-Gutiérrez, José Ramón López-Arellano y Arturo Yee-Rendón

Anotación automática de imágenes de plantas utilizando Inteligencia Artificial
Pág. 44 ⇒ Eduardo Díaz Gaxiola, Gerardo Beltran-Gutierrez, Inés F. Vega-López y Arturo Yee Rendón

Reconocimiento de Patrones de Artes de Pesca para Combatir la Pesca Ilegal
Pág. 50 ⇒ Ramón Aranda, Hugo Carlos, Angel Díaz-Pacheco, Miguel Á. Álvarez-Carmona y Ansel Y. Rodríguez-González

Retos de la IA para no comprometer la privacidad de los datos
Pág. 55 ⇒ Miguel Morales Sandoval, Heidy Marisol Marin Castro y Karina Figueroa Mora

Detección de notas de ideación suicida
Pág. 60 ⇒ María del Carmen García-Galindo, Ángel Hernández-Castañeda, René Arnulfo García-Hernández, Yulia Ledeneva y Rafael Cruz-Reyes

Columnas

Sapiens Piensa. Editorial pág. 2

e-Tlakuilo pág. 3

Estado del IArte pág. 6

Sakbe pág. 7

IA & Educación pág. 65

Deskubriendo Konocimiento pág. 67

Sapiens Piensa

Dra. Marisol Flores Garrido y Dr. Luis Miguel García Velázquez

Desde hace algunos años, la Inteligencia Artificial (IA) se ha integrado a nuestra vida diaria realizando todo tipo de tareas. Algunas veces, estas nuevas herramientas captan la atención del mundo entero y se insertan en el imaginario alimentado por la ciencia ficción, como ha sucedido con el desarrollo de asistentes virtuales y de ChatGPT. Otros algoritmos de IA de uso cotidiano reciben menor publicidad y, a pesar de ello -o precisamente por eso-, consiguen impactos mucho mayores: la autorización o el rechazo de un crédito en el banco, la curaduría de la información que consumimos en redes sociales, la identificación de patrones en los horarios y las rutas que usamos para conducir.

En medio de las fanfarrias y los mensajes desproporcionados que acompañan el anuncio de las nuevas aplicaciones de IA, puede ser difícil distinguir sus verdaderos alcances. También resulta complejo emitir opiniones informadas sobre aquello que, a pesar de ser asequible, no queremos construir. ¿Cuál es nuestro criterio para distinguir entre una tecnología nos acerque más al mundo que queremos habitar y una que alimente las desigualdades existentes? ¿Qué potencialidades son auténticamente maravillosas y cuáles son promesas vacías de la publicidad?

Abordar seriamente estas cuestiones requiere de una idea clara, informada, sobre el funcionamiento de la IA desde múltiples perspectivas, acompañada de una sensibilidad para valorar sus posibilidades, reconocer sus limitaciones y anticipar sus efectos al utilizarse en contextos concretos. A medida que la IA continúe con su avanzada al interior de nuestra rutina diaria, será más y más importante contar con medios que permitan establecer un puente de comunicación entre quienes desarrollan herramientas y quienes, en distintos sectores de la sociedad, enfrentan la encrucijada de abrazarlas o rechazarlas.

Celebramos la existencia de Komputer Sapiens y la vía que representa para abrir la discusión sobre el universo de posibilidades de la IA. Celebramos, además, un número que da la bienvenida a aplicaciones de áreas muy diferentes entre sí. En las páginas de esta edición encontrarán implementaciones de la IA que resultan útiles en campos como la astronomía, la modelación matemática, el procesamiento de audio y la anotación de imágenes; así mismo, hallarán ejemplos de los beneficios que la IA puede promover en el ámbito de la salud al emplearse en áreas tan diversas como la ginecología, la atención de enfermedades neurodegenerativas o la salud mental.

Este número también incluye aplicaciones de la IA en áreas que quizá no resultan tan evidentes a primera vista. ¿Cómo puede utilizarse para combatir la pesca ilegal, para evaluar la biodiversidad de una región a partir de imágenes muestra de plantas o para entender mejor factores que ponen en riesgo el rendimiento académico? Finalmente, cabe señalar que esta edición contribuye también a las reflexiones éticas sobre temáticas importantes en el ámbito de las tecnologías inteligentes, como es el caso de la protección de los datos personales ante la creciente tendencia para involucrar a terceras partes en el procesamiento automatizado.

El presente número, en la diversidad de temas que aborda, ilustra elocuentemente las posibilidades que encierra la IA. Más allá de un mero acercamiento hacia algunas aplicaciones concretas, consideramos que esta publicación nos invita a ensanchar la imaginación sobre lo que puede hacerse con esta tecnología, lo que nos parece fundamental para generar una conciencia crítica sobre el uso de la IA. Ampliar nuestra intuición sobre el futuro resulta clave para el presente, puesto que a la tecnología, a su desarrollo, sus aplicaciones concretas, a la retórica misma que la acompaña, las precede siempre la imaginación.

Esperamos que este número informe y, sobre todo, siembre ideas, alimente preguntas y contribuya a que nos preparemos mejor para cuestionar, construir e imaginar una versión del futuro que, por encima de la tecnología, ponga la vida en el centro de la conversación.✷



Marisol Flores Garrido estudió matemáticas aplicadas en la Universidad Autónoma de Coahuila. Posteriormente obtuvo una maestría en cómputo científico en la Universidad de British Columbia y un doctorado en ciencias computacionales en el Instituto Nacional de Astrofísica, Óptica y Electrónica. Es profesora en la Escuela Nacional de Estudios Superiores, Unidad Morelia (UNAM). Sus intereses de investigación se centran en el aprendizaje automático, el análisis de redes y la intersección entre la inteligencia artificial y las humanidades.



Luis Miguel García Velázquez es Ingeniero en Sistemas Computacionales graduado con Mención de Excelencia por el Tecnológico de Monterrey. Posteriormente obtuvo el grado de Maestría por el posgrado conjunto de la Universidad Michoacana de San Nicolás de Hidalgo y la Universidad Nacional Autónoma de México. Se doctoró en Ciencias Matemáticas por la Universidad Nacional Autónoma de México. Actualmente es profesor en la Escuela Nacional de Estudios Superiores, Unidad Morelia (UNAM). Sus intereses de investigación incluyen la modelación matemática y computacional para problemas interdisciplinares y la archivística computacional dentro de las humanidades digitales.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

e-Tlakuilo: Cartas de nuestros lectores

Karina Figueroa y Rosa María Valdovinos
etlakuilo@komputersapiens.org

En *Komputer Sapiens* nos hemos esforzado por estar "a sólo un click de distancia" a través de diferentes medios como Facebook, Twitter y correo electrónico. Les presentamos algunas de las preguntas que hemos recibido a través de estos medios.

Aprendizaje Automático para la determinación de factores que afectan la calidad del sueño.

Anónimo – Estudiante de nivel Licenciatura. (vía correo electrónico)

¿No se podría crear en vez de modelos totalmente personales o el otro modelo que trata de generalizar a todos los individuos, un punto medio donde se pueda hacer ciertos grupos donde pueda englobar más individuos que el personalizado para agilizar todo y así también obtener un R2 mayor a 0.7 que el modelo global?

Si, es una buena observación y si se ha considerado. Se ha estado trabajando en la definición de perfiles de usuario con características en común y de esa manera orientar los modelos a subgrupos. Aún no se tienen resultados precisos sobre la eficiencia del modelo en estos escenarios, consideramos que este tipo de estudio requiere de más tiempo y un conjunto grande de individuos, por lo que es un estudio más costoso, pero esperamos tener resultados en un futuro próximo.

¿No encontraron alguna relación con los principales FHS de cada individuo con su edad, estado civil, si es estudiante o trabaja?

En la primera fase del estudio no obtuvimos resultados que puedan relacionar los FHS con las características mencionadas, ya que finalmente el estudio se basó en las características psicofisiológicas individuales. Estamos trabajando en la siguiente fase orientada a subgrupos con características similares.

Alcances, límites y promesas en la implementación de sistemas de Inteligencia Artificial inspirados en teorías neuronales de la visión para el diagnóstico asistido basado en imágenes clínicas.

Anónimo – Estudiante de nivel Licenciatura. (vía correo electrónico)

¿Cree que algún día las IA superen en diagnóstico médico a los médicos?

Si bien parecería ser que en futuro las tecnologías de aprendizaje de máquinas y los sistemas de IA prometen llevar a cabo diagnósticos con un mayor índice de especificidad y sensibilidad en comparación con los agentes humanos, hoy en día esto ya es una realidad. Por ejemplo, algunos estudios experimentales han comprobado que diferentes arquitecturas de IA/AP son lo suficientemente sensibles para diagnosticar afecciones dérmicas como melanomas que, en algunos casos, requerirían ser examinadas mediante punciones u otras pruebas de verificación. Esto deja en evidencia que hoy en día la IA/AP supera algunas de las capacidades de diagnóstico de los agentes humanos.
Sin embargo, los sistemas de IA/AP operan mediante la computación de información que es generalmente procesada a través de reglas deductivas o de inducción, y el diagnóstico médico no siempre es producto de este tipo de razonamientos. Por el contrario, como ha discutido la filósofa mexicana Atocha Aliseda, gran parte del diagnóstico médico que se lleva a cabo por agentes humanos es producto de razonamientos abductivos, que son un área en donde la IA/AP no tiene un buen desempeño. Este tipo de razonamiento gira en torno a la suposición y a la creación de hipótesis, que son habilidades particularmente desarrolladas entre los agentes humanos.
Mientras que no existan arquitecturas de IA/AP capaces de intuir, difícilmente podremos esperar que las máquinas superen los alcances del diagnóstico médico llevado a cabo por personas. Esto debido a que el gasto computacional requerido para probar todas las hipótesis posibles sería demasiado elevado y porque, a final de cuentas, replicar el pensamiento abductivo de los agentes humanos (a veces referido como sentido común) sigue sin ser una prioridad de los desarrolladores informáticos (como ha evidenciado Erik J. Larson en el no poco polémico libro *The Myth of Artifical Intelligence*).

¿Cuál cree que es el futuro de la IA en el diagnóstico médico?

Las tecnologías de IA/AP son hoy en día utilizadas ampliamente en el área biomédica e incluso son herramientas que ya forman parte de las agendas de desarrollo y comercialización de dispositivos médicos. Por ejemplo, en los Estados Unidos, la FDA ha aprobado una buena cantidad de tecnologías que usan IA/AP.
Ante este panorama podemos intuir tres principales ca-

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

minos futuros:

1. Uno en el que se desarrollen sistemas inteligentes con un alto nivel de sensibilidad, precisión y exactitud, y que, además, demuestren la seguridad en su uso. Esto desembocaría en hacer de la IA/AP herramientas comunes en todos los niveles de atención al paciente o, incluso, en la reconfiguración de las prácticas médicas para el autodiagnóstico u otras derivas clínicas que no podemos vislumbrar desde nuestro panorama tecnológico.
2. Uno en el que estas tecnologías sean abandonadas en favor de otras prácticas. Esto podría ocurrir por el creciente interés de las entidades regulatorias por analizar las implicaciones bioéticas en el uso de los sistemas o por no contar con la infraestructura necesaria para el desarrollo tecnológico, entre muchas otras causas.
3. O aquel en el que los sistemas inteligentes sustituyan algunas de las prácticas médicas y de imagenología centradas en el diagnóstico, pero manteniendo la autoridad médica en la verificación de los datos y emisión de un diagnóstico.

El tercer camino es, con notable evidencia, aquel que podemos esperar con mayor confianza. Sin embargo, es importante tener en cuenta que el diagnóstico médico es una práctica situada, por lo que difícilmente podríamos imaginar un futuro en el uso de estas tecnologías homogéneo en todos los contextos culturales y territoriales.

La brecha técnica y digital, aunada a los intereses de la industria médica por capitalizar sus propios desarrollos tecnológicos, pueden representar un riesgo para la democratización en el uso de la IA/AP con fines clínicos, limitando el acceso a estos sistemas y creando desigualdad en la calidad de los diagnósticos. Por ello, aunque estas tecnologías puedan parecer prometedoras a largo plazo, es imposible pensarlas sin tomar en cuenta las coyunturas políticas, económicas y sociales de los contextos en los que serán implementadas.

¿Cuál cree que fue el diferenciador o beneficio de usar IA para el diagnóstico de COVID en la pandemia a diferencia de usar los métodos comunes de diagnóstico?

Una de las ventajas más evidentes que tenían algunas arquitecturas de IA/AP durante la pandemia por COVID-19 es que eran mucho más económicas que algunas otras alternativas, como las pruebas PCR. Los modelos utilizados habían sido previamente entrenados, por lo que el diagnóstico era rápido, eficiente y barato.
Estas cualidades derivaban también en un menor tiempo de espera para que un paciente fuera atendido, reduciendo asimismo el tiempo de exposición a potenciales fuentes de contagio y mejorando la atención hospitalaria. Sin embargo, es importante tener en cuenta que, a pesar de que diferentes tecnologías de IA/AP fueron utilizadas como primer medio de atención, en realidad estos sistemas no jugaron un papel protagónico en el diagnóstico y tratamiento de pacientes. La relativa novedad de estos métodos de diagnóstico representó una limitante para su uso extensivo, mientras que la premura hizo imposible la capacitación extensiva de personal para el uso de los sistemas.
Una implementación efectiva de la IA/AP en eventos futuros de esta naturaleza requerirá contar con personal capacitado y con la infraestructura adecuada para situarla como recurso en la línea de atención primaria.

Cómo contar personas usando WiFi y ML.

Anónimo – Estudiante de nivel Licenciatura. (vía correo electrónico)

¿Cómo contar personas usando WiFi y ML?

Al haber personas dentro de una señal WiFi es decir, entre un transmisor y receptor de señales WiFi, ocurren cambios o deformaciones en estas señales. Desde un punto de vista numérico, las señales son diferentes si hay una, dos, tres o n-cantidad de personas; inclusive son diferentes si las personas hacen actividades diversas como correr, caminar o sentarse. Los datos WiFi o información deben ser capturadas y matemáticamente procesadas con el propósito de que los algoritmos de ML los reciban como entradas, generen los modelos matemáticos y realicen el conteo de personas.

¿Cómo se asegura de no invadir la privacidad de las personas con el método?

Hasta el momento no es posible identificar o extraer información sensible de las personas que están ubicadas en un sistema de conteo WiFi. Hay estudios que permiten extraer información de tipo biométrico como frecuencia cardiaca o de respiración. Pero esto difícilmente puede asociarse a una persona en particular. Por lo que la privacidad de las personas está perfectamente protegida con este tipo de desarrollo tecnológico.

¿Qué se requeriría para que fuera aplicado al seguimiento de personas desaparecidas?

La tecnología de sensado remoto vía WiFi esta diseñada para funcionar en una región o área pequeña, en la escala de metros cuadrados. La detección de personas desaparecidas puede ser implementada utilizando la tecnología celular; que en la realidad ya existe. Las redes celulares

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

registran la "huella"que dejan los teléfonos celulares al conectarse a éstas. Si la persona esta en posesión del teléfono celular, ésta puede ser ubicada. Una posibilidad pudiera ser el uso de modelos de ML que detecten cambios en la trayectoria habitual de las personas por medio de la ubicación del teléfono; y enviar las alertas correspondientes.



© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Estado del IArte

Jorge Rafael Gutiérrez Pulido (@jrgpulido) y Reyna Carolina Medina Ramírez
estadoiarte@komputersapiens.org

ChatGPT y el procesamiento de lenguaje natural

Turquía es un país situado en Eurasia, que comparte fronteras con Grecia y Bulgaria al oeste, Georgia al norte, Armenia, Azerbaiyán y el mar Caspio al este, Irán y Irak al sur y el mar Mediterráneo y Siria al oeste. La capital de Turquía es Ankara y su ciudad más grande y conocida es Estambul. La población de Turquía es de aproximadamente 83 millones de personas y la moneda oficial es la lira turca. La mayoría de la población es musulmana y el idioma oficial es el turco ...

Las ondas gravitacionales son una forma de radiación gravitacional que se propaga a través del espacio-tiempo. Fueron teorizadas por primera vez por Albert Einstein en su teoría de la relatividad general en 1915, pero no fueron detectadas hasta 2016. Las ondas gravitacionales son causadas por la aceleración de objetos masivos, como estrellas binarias o agujeros negros. Cuando estos objetos se mueven, deforman el espacio-tiempo alrededor de ellos, causando ondulaciones que se propagan a través del universo ...

Soy un modelo de lenguaje desarrollado por OpenAI, una organización de investigación en inteligencia artificial con sede en San Francisco, California. Fui entrenado con un gran corpus de texto en internet para responder a preguntas y realizar otras tareas relacionadas con el lenguaje. Mi objetivo es proporcionar información precisa y útil a mis usuarios, y estoy diseñado para ser un recurso útil en una amplia variedad de tareas y aplicaciones, desde la respuesta a preguntas generales hasta la ayuda en tareas más específicas ...

Los párrafos anteriores son extractos de las respuestas obtenidas por una inteligencia artificial, conocida como *chatGPT* (Chat-Generative Pre-Trained), creada por la empresa OpenIA. El Chat GPT construyó el discurso anterior a partir de las siguientes preguntas: *dame algunos datos acerca de Turquía, háblame de las ondas gravitacionales, dame algunos datos acerca de tí.*

ChatGPT ha roto récords de adopción. Se estimó que en enero del 2023 alcanzó los 100 millones de usuarios activos apenas dos meses después de su lanzamiento. A TikTok le tomó 9 meses lograr lo mismo, cuando a Instagram dos años y medio. Esta tecnología y otras relacionadas, por ejemplo BERT de Google o ERNIE de Baidu, son consideradas estado del arte en el área de procesamiento de lenguaje natural, específicamente de los sistemas basados en **preguntas y respuestas**. Hacen uso de aprendizaje profundo y son entrenadas con miles de millones de textos, en casi todos los idiomas, sobre todo tipo de temas. Éstas tecnologías pueden generar artículos, ensayos, chistes, incluso poesía, no sólo en inglés y español, sino también en francés, alemán, italiano, portugués, ruso, árabe, chino, entre otros. ChatGPT acaba de anunciar un servicio *plus*, de paga, para aquellos usuarios que requieran servicios más especializados. Es posible conocer más sobre la herramienta en su página web[1] (incluyendo distintos idiomas además del españo). Si es el caso, en sus respuestas incluye la nota aclaratoria respecto de la actualidad y vigencia de las mismas (2021).

En el ámbito de la educación, hay cierta preocupación, tanto que en muchas escuelas del mundo se ha restringido su uso. Incluso ya se ha creado software para detectar su uso, e.g. Coalition Technologies[2], Copyleaks[3], y el mismo OpenAI[4]. Sin embargo, es posible obtener cierta ventaja, por ejemplo, al aplicarlas como asistente o tutor educativo para la resolución de problemas en áreas como programación, matemáticas, filosofía, entre otras. Con lo anterior se contribuye al proceso de enseñanza-aprendizaje autogestivo, así como al colaborativo de nuestros estudiantes.

El panorama completo del impacto que tendrán estas tecnologías aún no lo tenemos. Por lo pronto están marcando tendencia, como en su momento los *buscadores* lo hicieron, permitiéndonos tener al alcance de la mano la información. Por otro lado, siguen presentes las cuestiones éticas ¿cómo limitar el acceso a la información? tanto para personas que no tienen la madurez para asimilarla, como para aquellas que puedan darle un mal uso a la información. Microsoft está muy interesado en estos[5]temas, pues está por integrar esta tecnología en su buscador *Bing.*✷

[1] https://chatgptonline.net
[2] https://coalitiontechnologies.com/ai-content-detector
[3] https://copyleaks.com/ai-content-detector-chrome-extension
[4] https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text
[5] https://blogs.microsoft.com/on-the-issues/2022/06/21/microsofts-framework-for-building-ai-systems-responsibly

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Sakbe

Laura Gómez Cruz, **Laura Cruz Reyes** y **Claudia Guadalupe Gómez Santillán**
sakbe@komputersapiens.org

ChatGPT: una herramienta revolucionaria de la IA

La inteligencia artificial (IA) ha dado lugar a nuevas tecnologías que están modificando rápidamente nuestra vida de una manera muy importante. Uno de los desarrollos más emocionantes son los ***modelos de lenguaje de IA*** ya que pueden generar respuestas creativas a preguntas relacionadas con la escritura de código de software, la solución de teoremas matemáticos, la traducción de textos, la compresión de lectura y muchas más.

Los modelos de lenguaje más avanzados son los llamados ***modelos de lenguaje grandes*** (**LLM**, por sus siglas en inglés), que están entrenados en cantidades masivas de datos para predecir y generar texto coherente. GPT-3 es un LLM desarrollado por OpenAI, que actualmente cuenta con más de 175 mil millones de parámetros. Otro LLM es ChatGPT, aunque es más pequeño (6 mil millones de parámetros), estableció en enero de 2023 el récord de la base de usuarios de más rápido crecimiento debido a su accesibilidad (es rápido y sencillo).



ChatGPT ⇒ Como otros LLMs diseñados para tareas de conversación y diálogo, extienden la capacidad de los bots conversacionales para procesar y responder preguntas de manera más precisa y natural a una gama más amplia de preguntas y temas. Han mostrado capacidad para: a) *comprender el lenguaje natural*, facilitando al usuario la interacción; b) *evolucionar* mediante algoritmos de aprendizaje automático -como redes de aprendizaje profundo-, mejorando sus respuestas a las interacciones y datos alimentados; y c) *comprender el contexto* de una conversación para responder de acuerdo con ese contexto. La siguiente conversación con ChatGPT ilustra de manera sencilla algunas capacidades del modelo, incluyendo un error comúnmente cometido en las referencias bibliográficas (marcado en azul). Algunas respuestas han sido recortadas para ahorrar espacio en el extracto y se indican con puntos suspensivos (. . .)..

Conversación con ChatGPT

Enseño diseño de algoritmos a estudiantes de Ingeniería en Sistemas Computacionales ¿Qué tópicos me recomiendas para esta materia?

Para un curso de diseño de algoritmos para estudiantes de ingeniería en sistemas computacionales, algunos de los tópicos recomendados pueden incluir:

1. Análisis de complejidad de algoritmos (notación O grande)
2. Estructuras de datos (listas, arreglos, árboles, grafos)…

Me puedes ordenar los temas anteriores por dificultad y relación entre ellos?

Claro, aquí te presento una posible ordenación de los temas anteriores por dificultad, comenzando por los más básicos y avanzando hacia los más complejos:

1. Estructuras de datos (listas, arreglos, árboles, grafos)
2. Algoritmos de búsqueda y ordenamiento (búsqueda binaria, ordenamiento de burbuja, quicksort, mergesort)…

En cuanto a la relación entre ellos, cada tema es independiente, pero algunos se basan en conceptos previos…

Escribe un código en lenguaje R para el algoritmo quicksort, y que funcione sin instalar paquetes especiales?

¡Por supuesto! Aquí te presento una posible implementación del algoritmo quicksort en R, sin la necesidad de instalar paquetes adicionales:

```
quicksort <- function(x) {
  if (length(x) <= 1) {
    return(x)
  } else {
    Pivot <- x[1]
    left <- which(x < pivot)
    right <- which(x > pivot)
    middle <- which(x == pivot)
    return(c(quicksort(x[left]), x[middle], quicksort(x[right])))
  }
}
```

En esta implementación, la función **quicksort** toma un vector **x** como entrada y devuelve el vector ….

Escribe por favor un código R con datos que invoque a la función quicksort anterior.

Claro, aquí te muestro un ejemplo de cómo podrías usar la función **quicksort** en R con un vector de ejemplo:

```
# Definimos un vector desordenado
x <- c(3, 5, 2, 1, 4)
# Aplicamos la función quicksort
x_sorted <- quicksort(x)
# Imprimimos el resultado
print(x_sorted)
```

Este código definiría un vector desordenado **x** con los valores **[3, 5, 2, 1, 4]**…

¿Cuáles son los mejores libros para aprender los temas que me recomendaste?

Aquí te dejo algunas recomendaciones de libros para aprender sobre los temas que te mencioné:

- **Algoritmos. Teoría y práctica**, de Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest y Clifford Stein…

Estos libros cubren …

Aunque con limitaciones, resulta asombrosa la variedad de aplicaciones que se han reportado utilizando ChatGPT, un modelo de lenguaje que avanza a pasos agigantados. Sin embargo, todo progreso tiene un costo, y la sociedad debe estar preparada para aprovechar esta tecnología disruptiva y afrontar sus implicaciones controvertidas y posibles riesgos. Si deseas saber más sobre este tema, te recomendamos consultar la siguiente literatura.

- Barcelona (2023). Los 6 errores más comunes de ChatGPT. *Revista La Vanguardia*. https://www.lavanguardia.com/tecnologia/20230323/8843482/estudio-open-ai-puestos-trabajo-amenazados-gpt-4-pmv.html.
- Morales, M. (2023). ChatGPT aún no cuenta buenos chistes. *Revista el País*. https://elpais.com/cultura/2023-03-29/chatgpt-aun-no-cuenta-buenos-chistes.html?event_log=oklogin.
- Rudnitsky, J. y Bergen, M. (2023). Elon Musk vs ChatGPT: pide ir más despacio con proyectos de inteligencia artificial. *Revista El Financiero*. https://www.elfinanciero.com.mx/tech/2023/03/29/elon-musk-vs-chatgpt-pide-ir-mas-despacio-con-proyectos-de-inteligencia-artificial/.
- Van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R. y Bocking, C. L. (2023). ChatGPT: five priorities for research. *Nature*, 614(7947), 224-226. doi: https://doi.org/10.1038/d41586-023-00288-7.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Identificación efectiva de patologías ginecológicas aplicando Random Forest. Caso: Hospital de la mujer, Tabasco, México

Adriana-Guadalupe López-Ramírez, Eddy Sánchez-DelaCruz y Cecilia-Irene Loeza-Mejía

Impacto del aprendizaje automático en el estudio de Patologías ginecológicas

En México, el diagnóstico de patologías ginecológicas es de gran importancia, dado que las mujeres entre 20 a 45 años padecen infecciones vaginales al menos una vez en su vida [1]. Por lo que, el objetivo de este estudio fue identificar (o categorizar) casos clínicos de los siguientes tipos de vaginitis: *gardnerella*, *candidiasis*, *tricomoniasis* y *chlamydia*. Se implementó un enfoque de Aprendizaje Automático, específicamente el algoritmo Random Forest para categorizar casos clínicos de vaginitis en un conjunto de datos de sujetos de estudio del Estado de Tabasco, obteniendo resultados óptimos en términos de sensibilidad y especificidad, con valores entre el 93 % al 100 %, además, validados por dos médicos.

Antecedentes

El uso de Inteligencia Artificial mediante algoritmos de Aprendizaje Automático (ML por sus siglas en inglés de *Machine Learning*) se ha extendido a nivel mundial en áreas como gobierno, industria, ámbito médico y cultura. En medicina, ML se ha aplicado en el diagnóstico de enfermedades, personalización de tratamientos y en el enfoque emergente de Salud 4.0 [2].

En México, el correcto diagnóstico de las patologías ginecológicas es un reto, principalmente en mujeres entre 20 a 45 años, en edad reproductiva, embarazadas, con diabetes o en tratamiento contra el cáncer [1].

Las patologías ginecológicas se pueden producir por bacterias, mediante contacto sexual sin preservativos, uso incorrecto de preservativos, químicos de las cremas o productos de higiene, desequilibrio de organismos, bajas defensas o estilo de vida [1, 3, 4]. El diagnóstico temprano es de gran importancia, pues errores en el diagnóstico pueden alargar o empeorar la infección y provocar infertilidad [1]; en el caso de mujeres embarazadas puede producir nacimientos prematuros, secuelas en los bebés o la muerte de estos [1,4]. La vaginitis produce síntomas de ardor, irritación y flujo anormal y es una de las causas más comunes de acudir con el médico [3, 5]. Además, aumenta los riesgos en el embarazo y riesgos de adquisición de virus de inmunodeficiencia humana, herpes simple y papiloma [3]. Aunado a lo anterior, la vaginitis puede ser un síntoma de gonorrea [4]. Sin embargo, en muchas ocasiones, la atención de estas pacientes puede no ser apropiada, por no tener un diagnóstico adecuado al no contar con acceso médico especializado, lo que impacta en la calidad de vida de una paciente, además, los costes en los tratamientos pueden ser altos. Así mismo, regularmente en centros de salud de comunidades rurales, se carece de ginecólogos, Por lo que, la contribución de este estudio es ayudar en el diagnóstico de vaginitis, utilizando ML, sirviendo como apoyo al ámbito médico. Para llevar a cabo la categorización se implementó el algoritmo Random Forest, el cual, ha mostrado en la literatura resultados competentes con aplicaciones en el ámbito médico [6].

Materiales y métodos del estudio

Para abordar la problemática expuesta, empleamos ML en la categorización de casos clínicos relacionados con vagitinis.

Adquisición de datos

Se adquirieron casos clínicos de 600 sujetos de estudio, entre 2018 a 2019, del Hospital de la Mujer en Villahermosa, Tabasco. Los casos clínicos describen cuatro tipos comunes de vaginitis:

- Gardnerella: es un desequilibrio en la micribiota saprófita normal de la vagina con una disminución de Lactobacillus spp. y un sobrecrecimiento poblacional de *Gardnerella vaginalis* y otras bacterias aerobias y anaerobias [7]. Entre las manifestaciones se incluyen flujo maloliente y prurito (picazón en la piel) [3].
- Candidiasis: infección fúngica ocasionada por el hongo *Candida albicans* [4] que se presenta general-

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

mente en la piel o las membranas mucosas, manifestada en individuos con defensas bajas [3] y afectando entre el 70 a 75 % de mujeres en edad fértil [8]. Los síntomas incluyen flujo, disuria (ardor o molestia al orinar), prurito vulvar y enrojecimiento [4, 8].

- Tricomoniasis: tiene alta prevalencia en mujeres prisioneras, además, puede estar relacionada con cervicitis y tiene altas tasas de reinfección [4]. Los factores de riesgo son tener varias parejas sexuales y no usar preservativo durante el acto sexual [4]. La tricomoniasis causa flujo abundante, picazón en los genitales y micción dolorosa en las mujeres [8].
- Chlamydia: infección bacteriana de transmisión sexual con mayor frecuencia en adolescentes y adultos jóvenes [4], aunque muchas de las personas con chlamydia no desarrollan síntomas, pueden infectar a otras mediante el contacto sexual.

Tabla 1. Descripción de datasets

Dataset	Casos (clases)	Registros	Etiqueta de clase
1	Gardnerella	100	1
	Candidiasis	75	0
2	Gardnerella	100	1
	Tricomoniasis	61	0
3	Gardnerella	100	1
	Chlamydia	100	0
4	Candidiasis	75	1
	Tricomoniasis	61	0
5	Candidiasis	75	1
	Chlamydia	100	0
6	Tricomoniasis	61	1
	Chlamydia	100	0
7	Gardnerella	100	1
	{Candidiasis, Tricomoniasis, Chlamydia}	236	0
8	Candidiasis	75	0
	{Gardnerella, Tricomoniasis, Chlamydia}	261	1
9	Chlamydia	100	0
	{Gardnerella, Tricomoniasis, Candidiasis}	236	1
10	Tricomoniasis	61	0
	{Gardnerella, Candidiasis, Chlamydia}	275	1

Tabla 2. Atributos de un dataset (Fragmento)

Flujo	Secreción	Prurito	Dolor abdominal	...	Clase
1	1	2	1	...	1
1	1	2	2	...	1
1	2	2	2	...	0
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
1	2	1	1	...	0

De los 600 casos clínicos adquiridos, se construyeron 10 datasets en formato de archivo delimitado por comas (CSV), que se describen en la Tabla 1. Cada uno de los datasets consistió en 10 atributos: flujo, secreción, prurito, dolor abdominal, dolor pelvis, disuria, olor, poliuria (más orina de la esperada), polaquiuria (necesidad de orinar muchas veces) y dispareunia (dolor genital recurrente que ocurre justo antes, durante o después de tener relaciones sexuales), más la etiqueta de clase, es decir el valor que indica el tipo de vaginitis, como se muestra en la Tabla 2, donde 1 o 0 indica el tipo de vaginitis (ver Tabla 1). Una vez construidos los datasets, se procedió a separarlos en dos subconjuntos: entrenamiento y prueba, con el criterio de muestreo 70/30, respectivamente.

Random Forest

Un algoritmo de ML que ha mostrado resultados competitivos en la categorización de enfermedades es Random Forest [6], el cual consiste en un ensamble de árboles de decisión [9]. Si lo queremos comparar con la manera de tomar decisiones de los humanos, Random Forest es similar al jurado de un evento, en este caso, cada uno de los jurados es un árbol de decisión y realiza un voto, al finalizar la opción que se elige es la decisión de la mayoría. En la Figura 1 se representa gráficamente la estructura general del algoritmo Random Forest.

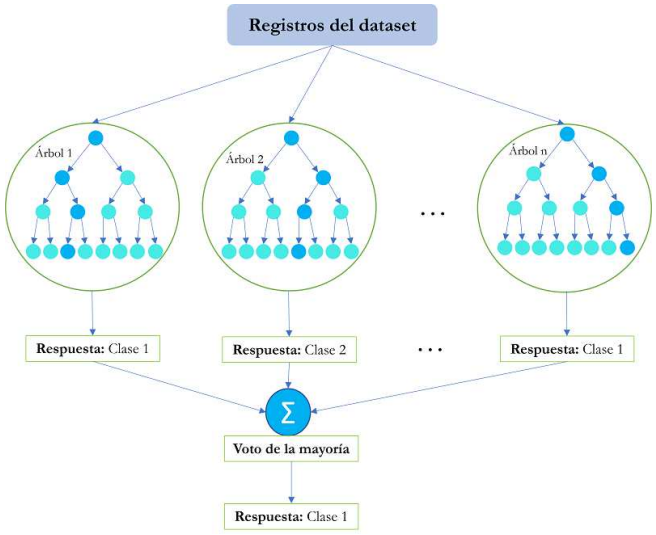


Figura 1. Estructura general de Random Forest.

El algoritmo Random Forest se implementó utilizando la herramienta de ML scikit-learn[1], las opciones de configuración que se establecieron fueron random_state=1, número de árboles=100, cantidad mínima de observaciones para dividir un nodo=2, observaciones mínimas que puede tener una hoja del árbol=1 y tareas en paralelo=1.

Configuración experimental

Los experimentos se llevaron a cabo en un equipo lenovo ideapad 330, procesador Intel(R) Core(TM) i3-7020U CPU @ 2.30GHz, memoria RAM 8.00 GB, sistema operativo de 64 bits, python 3.7 y software Excel para creación y visualización de datasets en formato CSV. En la Figura 2 se muestra el proceso metodológico realizado en el presente estudio, para la categorización de patologías ginecológicas.

[1] Disponible en `https://scikit-learn.org/stable/modules/ensemble.html`.

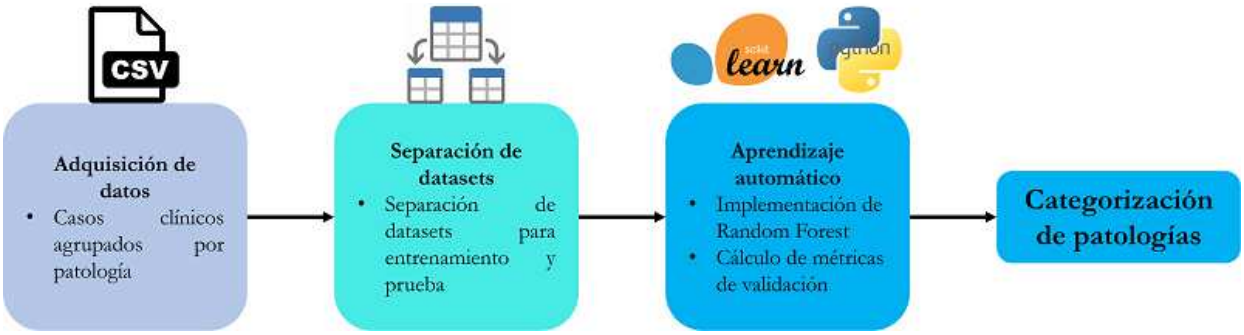


Figura 2. Proceso metodológico.

Resultados y Discusión

Para evaluar los resultados obtenidos en nuestro caso de estudio, utilizamos la matriz de confusión, sensibilidad y especificidad. La matriz de confusión es una tabla en donde se asocian los valores de las etiquetas de clase asignadas (o capturadas) por el experto y las etiquetas de clase estimadas mediante un algoritmo de ML. En el caso de una matriz binaria, es decir, donde se analizan dos clases, se obtienen cuatro valores: verdaderos positivos, verdaderos negativos, falsos negativos y falsos positivos. Los verdaderos positivos representan el total de recurrencias donde coincide el valor estimado mediante un algoritmo de ML y el que asignó el experto, mientras que los verdaderos negativos son aquellos donde coinciden los valores negativos categorizados correctamente. Por otra parte, los falsos positivos y falsos negativos representan los valores donde no se estiman correctamente los valores que ha asignado el experto. La sensibilidad representa la tasa de valores positivos estimados correctamente y la especificidad la tasa de valores negativos estimados correctamente [10].

En la Tabla 3 se muestran las matrices de confusión obtenidas para cada uno de los datasets en el subconjunto de prueba, en azul se representan los verdaderos positivos, en gris los verdaderos negativos, en verde los falsos negativos y en blanco los falsos positivos. En los datasets 1, 2, 4, 5, 6, 8 y 10 se obtuvieron valores de 1.0 de sensibilidad y 1.0 de especificidad. Mientras que, en el dataset 3, se obtuvo 0.93 de sensibilidad y 0.93 de especificidad. En el dataset 7, se logró 0.98 de sensibilidad y 0.93 de especificidad. Por otra parte, en el dataset 9 se obtuvo 0.96 de sensibilidad y 0.97 de especificidad.

De acuerdo con la médico Rubí López Álvarez, con Cédula Profesional no. 7574655 para ejercer en México, los resultados obtenidos los considera favorables. Cabe mencionar que ella sugiere aumentar la cantidad de casos para observar los cambios en las precisiones que resultaron ser un poco bajas; también mencionó abordar más a detalle el cuadro clínico contemplando otras posibles características. Así mismo, la médico Claudia Ivonne Ramírez Santiago con Cédula Profesional no. 11634974 para ejercer en México, manifiesta que los resultados obtenidos son favorables, considerando que sería interesante agregar más características extraídas de otros exámenes; además, ella considera que estudios como estos, enfocados al área de la salud en patologías o procesos clínicos, son de mucha utilidad para que ellos (los médicos), puedan tener otra perspectiva.

Tabla 3. Matrices de confusión

Matriz 1		Matriz 2		Matriz 3		Matriz 4		Matriz 5		Matriz 6		Matriz 7		Matriz 8		Matriz 9		Matriz 10	
23	0	19	0	29	2	17	0	29	0	31	0	67	1	23	0	28	1	17	0
0	30	0	30	2	27	0	24	0	24	0	18	2	31	0	78	2	70	0	84

Conclusiones y trabajos futuros

Mediante este estudio se observó que pueden converger disciplinas de la inteligencia artificial y del área médica, Se encontró, además, que la implementación del algoritmo Random Forest es adecuada en la categorización de patologías vaginales, ayudando a lograr un mejor diagnóstico y con esto un tratamiento adecuado en las patologías relacionadas con vaginitis.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Como trabajo a futuro se sugiere lo siguiente:

- Extender el estudio considerando otros hospitales de México.
- Realizar comparaciones de algoritmos de aprendizaje automático en otros casos médicos.
- Desarrollar un sistema experto, para uso funcional en centros de salud, donde no se cuente con un médico especialista. Además de aplicar encuestas de satisfacción a los usuarios del sistema.✷

Agradecimientos

Los autores agradecen al Hospital de la Mujer en Villahermosa, Tabasco y a los sujetos de estudio por proporcionar la información para llevar a cabo el estudio.

Los autores agradecen a las médicos Rubí López Álvarez y Claudia Ivonne Ramírez Santiago por el continuo apoyo para conducir el estudio y la retroalimentación de los resultados.

Los autores desean agradecer al CONACyT, en México, por los fondos para estudio de posgrados 2018-000068-02NACF-31312 y 2021-000018-02NACF-12228 otorgados a López-Ramírez y Loeza-Mejía, respectivamente.

REFERENCIAS

1. Secretaría de Salud (2018). 181. padecen mujeres infección vaginal al menos una vez en su vida. Recuperado el 15 de febrero de 2023, de `https://www.gob.mx/salud/prensa/181-padecen-mujeres-infeccion-vaginal-al-menos-una-vez-en-su-vida`.
2. Loeza-Mejía, C.-I., Sánchez-DelaCruz, E., Pozos-Parra, P., y Landero-Hernández, L.-A. (2021). The potential and challenges of health 4.0 to face covid-19 pandemic: a rapid review. *Health and Technology*, 11(6), 1321–1330.
3. Moreno-Tapia, J. G. y Abaca Castillo, E. P. (2019). Estudio microbiológico de agentes colonizantes e infectantes del flujo vaginal de mujeres que asisten al centro de salud José Dionisio Astaburuaga de Talca. Memoria de licenciatura, Licenciatura en Tecnología Médica, Universidad de Talca. Facultad de Ciencias de la Salud. Escuela de Tecnología Médica. Chile.
4. Workowski, K. A., Bachmann, L. H., Chan, P. A., Johnston, C. M., Muzny, C. A., Park, I., Reno, H., Zenilman, J. M., y Bolan, G. A. (2021). Sexually transmitted infections treatment guidelines, 2021. *MMWR Recommendations and Reports*, 70(4).
5. Aminimoghaddam, S., Batayee, S. S., Maghsoudnia, A., y Mohammadi, M. (2018). Update of vaginitis for obstetricians: A review. *Razi Journal of Medical Sciences*, 24(163), 48-–57.
6. Fuentes-Ramos, M., Sánchez-DelaCruz, E., Meza-Ruiz, I.-V., y Loeza-Mejía, C.-I. (2022). Neurodegenerative diseases categorization by applying the automatic model selection and hyperparameter optimization method. *Journal of Intelligent & Fuzzy Systems*, 42(5), 4759–4767.
7. quimica.es (2023). Gardnerella vaginalis. Recuperado el 28 de septiembre de 2022, de `https://www.quimica.es/enciclopedia/Gardnerella_vaginalis.html`.
8. Espitia-De-La-Hoz, F.-J. (2021). Evaluación de la eficacia y seguridad del policresuleno en el tratamiento de la vaginitis mixta, Armenia, Colombia, 2017-2019. Estudio aleatorizado. *Archivos de Medicina (Col)*, 21(1), 45-–59.
9. Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5—32.
10. Loeza-Mejía, C.-I., Biswal, R. R. y SánchezDelaCruz, E. (2019). Estado del arte de clasificación de nódulos pulmonares utilizando aprendizaje automático. *Research in Computing Science*, 7, 429—440.

SOBRE LOS AUTORES



Adriana-Guadalupe López-Ramírez es Ingeniera en Sistemas Computacionales por el Instituto Tecnológico Superior de Comalcalco y Maestra en Sistemas Computacionales por el Instituto Tecnológico Superior de Misantla. Actualmente es colaboradora de plataforma para control de avance en obras de construcción. Sus intereses de investigación son los sistemas inteligentes para la monitorización y evaluación del espacio de construcción mediante inteligencia artificial.



Eddy Sánchez-DelaCruz es Doctor en Sistemas Computacionales por la Universidad del Sur, Maestro en Sistemas Computacionales y Licenciado en Informática Administrativa por la Universidad Autónoma de Tabasco. El Dr. Sánchez-DelaCruz es miembro del Sistema Nacional de Investigadores, nivel I (destacando que obtuvo el nivel uno en su primera evaluación), es miembro de la Sociedad Mexicana de Inteligencia Artificial y miembro de la RedICA-CONACyT, es Perfil deseable PRODEP y es fundador y líder del Cuerpo Académico Inteligencia Artificial. Sus intereses de investigación son el aprendizaje automático y la ciencia de datos con un enfoque multidisciplinario. (www.eddsac.com).



Cecilia-Irene Loeza-Mejía es estudiante del Doctorado en Ciencias de la Ingeniería y Maestra en Sistemas Computacionales por el Instituto Tecnológico Superior de Misantla (ITSM), es Maestra en Ciencias Computacionales y Telecomunicaciones por el Instituto de Estudios Universitarios y es Ingeniera en Tecnologías de la Información y Comunicaciones por el ITSM. Es miembro de la Sociedad Mexicana de Inteligencia Artificial y colaboradora del Cuerpo Académico Inteligencia Artificial. Sus intereses de investigación son metodologías de aprendizaje automático y el descubrimiento de conocimiento. (cecilialoeza.com).

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Técnicas de aprendizaje automático para la predicción del rendimiento académico

Andrés Rico Páez

Introducción

Actualmente, los dispositivos móviles, las plataformas en internet y diferentes medios han propiciado una gran cantidad de información referente a características de los usuarios, la cual puede ser procesada con objetivos tales como ventas, publicidad, negocios, etc. En el campo educativo, se ha comenzado a generar cada vez mayor cantidad de datos debido a la posibilidad de almacenarlos de forma digital. Esto ha originado la posibilidad de utilizar dicha información para analizarla y encontrar patrones en los estudiantes para mejorar los procesos de enseñanza y aprendizaje, es decir, existe la potencialidad de ayudar en la disminución de problemas académicos tales como los altos índices de reprobación y la deserción estudiantil [1].

Una de las razones de la deserción estudiantil se debe a que el estudiante va obteniendo un bajo desempeño académico, ya sea en una o en muchas asignaturas. Es decir, conforme al estudiante se le van acabando las oportunidades de aprobar sus asignaturas tiende a desertar de la escuela [2].

El rendimiento académico es uno de los principales indicadores del éxito o el fracaso de un estudiante y depende de diversos factores [3]. De esta manera, es de interés de las instituciones educativas predecir, en la medida de lo posible, el rendimiento académico de estudiantes. En la literatura se han utilizado técnicas de aprendizaje automático para la predicción del rendimiento académico [4]. Los resultados obtenidos mediante este tipo de técnicas muestran cómo algunas características de los estudiantes influyen en la predicción del rendimiento académico [5]. El estudio de las características de estudiantes para la creación de modelos de predicción del rendimiento académico ha sido estudiado en años recientes [6]. De entre las características o factores que influyen en la predicción del rendimiento académico destacan el promedio actual del estudiante y las evaluaciones internas realizadas, es decir, las actividades académicas hechas dentro del curso. Para medir la efectividad de las predicciones se utiliza, típicamente, la exactitud de las predicciones, es decir, el total de predicciones correctas con respecto al total de las predicciones. En varios trabajos de predicción de rendimiento académico se ha observado cómo distintas técnicas de aprendizaje automático han sido mejores con respecto a la exactitud ya que depende del contexto y las características que se utilizan para hacer las predicciones [7]. Si bien es importante obtener una exactitud lo más alta que se pueda, también es de interés realizar la predicción lo más tempranamente posible con el fin de realizar algún tipo de intervención educativa en el estudiante, detectar estudiantes en peligro de reprobación o algún otro riesgo educativo.

Las técnicas de aprendizaje automático brindan la posibilidad de detectar estudiantes con bajo rendimiento académico.

El estudio presentado en este trabajo tiene por objetivo aplicar técnicas de aprendizaje automático para predecir el rendimiento académico de estudiantes de México.

El análisis de datos provenientes de contextos educativos ofrece potencial beneficio en la detección de deserción estudiantil o de éxito al finalizar sus estudios. La predicción del rendimiento académico es un tipo de análisis que se puede realizar a diferentes niveles de detalle, es decir, se puede predecir los resultados de tareas simples o de exámenes que abarquen un determinado periodo. De entre estos tipos de predicción, la que resulta de mayor interés, tanto para estudiantes y profesores, es la predicción del rendimiento académico al final de un curso debido a que representa la culminación de una fase tanto a nivel académico del estudiante como a nivel administrativo en la institución educativa. En México existen pocos estudios acerca de este tema como en [8], a pesar de que este tipo de herramientas ofrece la posibilidad de detectar estudiantes con problemas académicos, es decir, el planteamiento de metodologías para aplicar técnicas de aprendizaje automático en la predicción del rendimiento académico puede aportar información para la mejora del proceso educativo.

Metodología

En esta sección se propone una metodología para utilizar datos de estudiantes extraídos al principio de un curso y las calificaciones de actividades académicas durante el mismo con el propósito de predecir el rendimiento aca-

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

démico al final del curso como se muestra en la Figura 1.

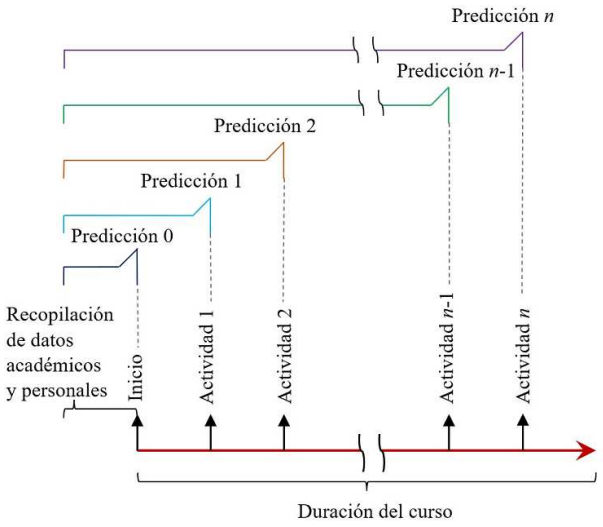


Figura 1. Metodología propuesta para predecir el rendimiento a partir de datos al inicio y durante el curso.

En esta metodología se realizan predicciones del rendimiento académico en un entorno educativo universitario en distintos momentos. La primera predicción se realiza al inicio o antes de que inicie el curso mediante los datos académicos y personales de estudiantes (predicción 0). La siguiente predicción (predicción 1) utiliza los datos anteriores y, además, la calificación de la actividad 1. Este proceso se repite, pero ahora incluyendo la actividad 2 (predicción 2), y así sucesivamente, hasta realizar la predicción n con todos los datos recabados hasta ese momento.

Para aplicar la metodología, se recopilaron 23 variables de 111 estudiantes universitarios de un curso de Matemáticas pertenecientes a una institución pública de México. Esta cantidad de registros ha sido utilizada con técnicas similares en otros estudios como en [9]. A los estudiantes se les informó que los datos recopilados serían únicamente para los objetivos planteados en este estudio.

En la Tabla 1 se muestran las primeras 9 variables recopiladas en este estudio, las cuales, se refieren a datos académicos, socioeconómicos y personales al inicio del curso, y aunque existen muchas otras variables que pueden influir en el rendimiento académico, se ha observado que, con el tipo de variables utilizadas en este artículo se han obtenido buenos resultados en la predicción del rendimiento académico en trabajos similares [7].

Las otras 14 variables corresponden a las calificaciones de 14 actividades académicas realizadas por las estudiantes distribuidas en tiempos similares a lo largo de un curso. Los valores posibles de estas actividades fueron: aprobada, reprobada y no presento. Finalmente, se recopiló la calificación final del curso teniendo los valores posibles de aprobado o reprobado.

Tabla 1.
Valores posibles de las variables recopiladas al inicio del curso.

Variables	Valores posibles
Escolaridad del padre	Primaria y secundaria, bachillerato, nivel superior o posgrado
Escolaridad de la madre	Primaria y secundaria, bachillerato, nivel superior o posgrado
Ingreso familiar	<$5000, entre $5000 y $10000, >$10000
Promedio obtenido en el nivel medio superior	Entre 0 y 7.4, entre 7.5 y 8.4, entre 8.5 y 10
Materias reprobadas actualmente	0, 1, ≥2
Promedio actual	Entre 0 y 7.4, entre 7.5 y 8.4, entre 8.5 y 10
Preferencia de estudio	Sólo, dúo, en grupo
Preferencia para realizar actividades	Sólo, dúo, en grupo
Frecuencia de estudio	Continuamente, una semana antes del examen, un día antes del examen

Técnicas de aprendizaje automático utilizadas

Naïve Bayes. Esta técnica construye el modelo predictivo por medio del cálculo de un conjunto de probabilidades a partir de los datos de entrenamiento. Es decir, el modelo predictivo está representado por un conjunto de probabilidades. Para la predicción de un registro en concreto se realizan varias operaciones con estas probabilidades de acuerdo con los valores posibles de las variables de dicho registro.

K vecinos más cercanos. Esta técnica realiza la predicción de un registro con respecto a la clase que más se repita de entre los k vecinos más próximos a dicho registro.

Árbol de decisión C4.5. Esta técnica consiste en un algoritmo para la construcción de un conjunto de condiciones organizadas en una estructura jerárquica en forma de árbol, de tal forma, que para obtener una predicción de un registro se siguen las condiciones o ramas que cumplen desde la raíz del árbol hasta alguna de las hojas.

Voto mayoritario. Esta técnica combina las predicciones de distintas técnicas de aprendizaje automático, y selecciona, la predicción que más se repita, es decir, la que tenga mayor cantidad de votos.

La técnica de voto mayoritario es una técnica combinada que utiliza técnicas individuales. En este estudio, se emplean como técnicas individuales las presentadas anteriormente. De esta manera, se realizan las predicciones con cada una de estas técnicas y se elige la que tenga

mayor número de votos. Se debe notar que la técnica de voto mayoritario requiere un mayor procesamiento de datos debido a que requiere realizar todos los modelos con las tres técnicas de aprendizaje automático y además realizar el conteo para la selección de la predicción.

Resultados

El análisis y procesamiento de datos se realizó con el apoyo del software libre Weka [10]. Éste contiene algoritmos para el análisis de datos y la construcción de modelos predictivos. Se debe notar que los resultados obtenidos en esta sección son válidos solamente para el conjunto de datos utilizado. No obstante, la metodología empleada puede ser aplicada a cualquier curso.

Para evaluar el desempeño de las técnicas de aprendizaje automático se utiliza la exactitud de las predicciones, es decir, la cantidad de predicciones correctas con respecto al total de las predicciones Para el cálculo de la exactitud se emplea la validación cruzada con 10 conjuntos, la cual consiste en dividir aleatoriamente los datos de entrenamiento en 10 conjuntos de datos. Uno de los conjuntos se utiliza como datos de prueba y el resto como datos de entrenamiento para construir el modelo y calcular la exactitud. El proceso se repite 10 veces con un conjunto diferente. Finalmente, se promedia la exactitud de los 10 casos. Esta metodología se realiza para cada una de las cuatro técnicas de aprendizaje automático en diferentes momentos del curso analizado. El curso duró 16 semanas y se recopilaron las primeras 14 actividades académicas realizadas una por semana. Es decir, inicialmente, se aplican las cuatro técnicas de aprendizaje automático con validación cruzada únicamente utilizando los datos académicos y personales al inicio del curso (datos iniciales). En la siguiente semana, se aplican las técnicas a los datos iniciales y la primera actividad. En la semana 2 se emplean los datos iniciales y las actividades 1 y 2. Así sucesivamente, hasta la semana 14 donde se utilizan los datos iniciales y las 14 actividades académicas. En la Figura 2 se presenta la exactitud de las predicciones con las técnicas de aprendizaje automático en diferentes momentos del curso.

Es posible predecir, con cierta exactitud, si un estudiante aprueba un curso.

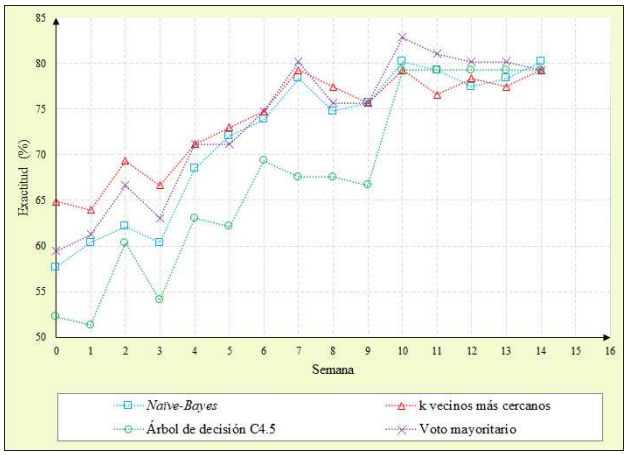


Figura 2. Exactitud de las técnicas de aprendizaje automático.

Se pueden observar que la exactitud de las predicciones con las diferentes técnicas de aprendizaje automático muestra diferentes valores a lo largo del curso. Sin embargo, de manera global, se puede apreciar una tendencia creciente de la exactitud de las predicciones durante curso. Es decir, la exactitud tiende a aumentar conforme se va añadiendo información a los modelos predictivos construidos con las técnicas de aprendizaje automático. Se puede observar en la semana 0, es decir, únicamente utilizando los datos académicos y personales iniciales que se obtienen exactitudes bajas, siendo la más alta de alrededor de 65 % con la técnica k vecinos más cercanos. También se puede notar que la técnica de aprendizaje automático k vecinos más cercanos tiene un desempeño en la exactitud de las predicciones superior en casi todos los casos en comparación a las demás técnicas desde el inicio del curso hasta la semana 6. Posteriormente, de la semana 6 hasta la semana 14, la técnica de voto mayoritario muestra una mejor exactitud en las predicciones en la mayoría de los casos.

En la Figura 2, el valor más alto de exactitud es de 82.88 % y se obtiene con la técnica voto mayoritario cuando se emplean los datos iniciales y las actividades de la 1 a la 10. La actividad 10 se realiza en la semana 10, poco más de la mitad de la duración del curso y cerca del final de este, de esta manera, realizar una predicción en la semana 10 puede no ser tan eficaz para una intervención académica en un estudiante como realizar una predicción en semanas anteriores. Sin embargo, sería conveniente mantener un compromiso entre una predicción temprana del rendimiento académico del estudiante y la exactitud de las predicciones. En este sentido, algunos valores de que mantienen dicho compromiso pueden ser en la semana 2 con una exactitud de alrededor de 70 % (k vecinos más cercanos) y en la semana 7 con una exactitud de aproximadamente 80 % (k vecinos más cercanos, voto mayoritario y Naïve Bayes).

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Discusión y conclusiones

En esta investigación de se aplicaron técnicas de aprendizaje automático para la predicción del rendimiento académico de estudiantes de una institución educativa de nivel superior en México. La metodología propuesta utiliza datos académicos y personales al inicio de un curso y actividades académicas durante el mismo. En los resultados obtenidos se puede notar que la exactitud de las predicciones presenta aumentos y disminuciones durante el curso. Sin embargo, de manera general, se puede observar una tendencia creciente de la exactitud con todas las técnicas de aprendizaje automático utilizadas. Esto se debe a que, al avanzar el curso, se va adicionando más información a los modelos predictivos construidos y tienden a tener una mayor exactitud. La exactitud de las predicciones tiende a ser mayor con la técnica k vecinos más cercanos en las primeras semanas curso, y en las últimas semanas, la técnica de voto mayoritario ofrece un mejor desempeño en la exactitud.

Es de interés hacer la predicción del rendimiento académico lo más tempranamente posible para realizar, de manera más oportuna las intervenciones académicas a los estudiantes. Además, también es importante que la exactitud de las predicciones sea lo más alta posible para que nos brinde confiabilidad en las predicciones. De esta manera, es conveniente mantener un compromiso entre la predicción temprana y su exactitud. En esta línea de pensamiento, se identificaron valores de exactitud de 70 % (k vecinos más cercanos) y 80 % (voto mayoritario) en las semanas 2 y 7, respectivamente. En este sentido, parece más relevante utilizar la técnica k vecinos más cercanos en las primeras semanas para una detección temprana de estudiantes con bajo rendimiento académico y manteniendo un compromiso de tener la exactitud de las predicciones lo más alta posible con las técnicas de aprendizaje empleadas y con los datos de estudiantes recabados.

Las metodologías de predicción del rendimiento académico, como la planteada en este estudio, permite ayudar a los estudiantes al posibilitarle a las instituciones educativas obtener predicciones lo suficientemente confiables para realizar algún tipo de intervención, lo más tempranamente posible, por parte del profesor y disminuir, en la medida de lo posible, la reprobación.

A pesar de los avances realizados en este trabajo, es conveniente mencionar algunos posibles temas de investigación en esta área. Primeramente, puede ser de interés emplear otras técnicas de aprendizaje automático diferentes a las utilizadas en este estudio. Además, se puede complementar lo anterior, empleando una mayor cantidad de datos de estudiantes. Adicionalmente, se pueden aplicar técnicas de aprendizaje automático a otros niveles educativos u otras modalidades con el fin de tratar de generalizar, en la medida de lo posible, los resultados obtenidos.✷

REFERENCIAS

1. Castillo, M., Gamboa, R. e Hidalgo, R. (2020). Factores que influyen en la deserción y reprobación de estudiantes de un curso universitario de matemáticas. *Uniciencia*, 34(1), 219-245.
2. Uribe, O. L. y Carrillo, S. (2014). Relación entre la lecto-escritura, el desempeño académico y la deserción estudiantil. *Entramado*, 10(2), 272-285.
3. Contreras, L., Fuentes, H. y Rodríguez, J. I. (2020). Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático. *Formación Universitaria*, 15(5), 233-246.
4. Narváez, A. F. (2023). Modelo Estadístico para determinar los factores académicos en los Resultados de las Pruebas Saber Pro. *Investigación e Innovación en Ingenierías*, 11(1), 3-21.
5. Saire, E. A. y Velarde, E. A. (2023). Modelo clasificador para personalizar ejercicios propuestos a estudiantes utilizando redes neuronales artificiales. *Publicaciones*, 53(2), 89-124.
6. Incio, F., Capuñay, D. y Estela, R. (2022). Artificial Neural Network Model to Predict Academic Results in Mathematics II. *Revista Electrónica Educare*, 27(1), 1-19.
7. Shahiri, A. M., Husain, W. y Rashid, N. (2015). A Review on Predicting Student's Performance using Data Mining Techniques. *Procedia Computer Science*, 72, 414-422.
8. Juárez, A., Cortés, J. y Coronilla, U. (2014). Aplicación de la inteligencia artificial en la sistematización de procesos educativos. Caso: Sistema de detección de riesgo escolar en ESCOM. *Revista Iberoamericana de Producción Académica y Gestión Educativa*, 1(1), 140-163.
9. Mueen, A., Zafar, B. y Manzoor, U. (2016). Modeling and Predicting Students' Academic Performance Using Data Mining Techniques. *International Journal of Modern Education and Computer Science*, 8(11), 36-42.
10. Abuhaija, B., Alloubani, A., Almatari, M., Jaradat, G. M., Abdallah, H. B., Abualkishik, A. M. y Alsmadi, M. K. (2023). A comprehensive study of machine learning for predicting cardiovascular disease using Weka and SPSS tools. *International Journal of Electrical and Computer Engineering*, 13(2), 1891-1902.

SOBRE EL AUTOR



Andrés Rico Páez es Doctor en Tecnología Avanzada por el Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada unidad Legaria del Instituto Politécnico Nacional (IPN). Es Profesor Titular en la Escuela Superior de Ingeniería Mecánica y Eléctrica unidad Zacatenco del IPN. Actualmente, es miembro del Sistema Nacional de Investigadores. Sus líneas de investigación se enfocan en la aplicación de la Ciencia de Datos e Inteligencia Artificial al área educativa.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Redes Neuronales Convolucionales con Atención para Atenuar Ruido en Señales de Audio

Noel Zacarias-Morales, José Adán Hernández-Nolasco, Pablo Pancardo y J. Arturo Olvera-López

Introducción

Los humanos tenemos la capacidad de entender los mensajes cuando estamos platicando con otra persona a pesar de que se encuentren presentes varias fuentes de sonido adicionales (motores funcionando, música, aves, etcétera), las cuales representan ruido para la comunicación. Por lo general, en un escenario donde debemos poner atención a una voz en particular, mitigar el ruido no representa un problema para nosotros. Esto es debido a que desde nuestro nacimiento continuamente mejoramos nuestra capacidad de atención selectiva [1]; esta capacidad nos permite seleccionar de forma natural una fuente de sonido específica (como una voz) y atenuar los ruidos que puedan estar presentes para comprender lo que nos están diciendo.

Desafortunadamente para los sistemas computacionales de procesamiento de audio, realizar esta tarea requiere de la integración de algoritmos que les brinden la capacidad para atenuar o eliminar diversas fuentes de audio que representan ruidos en la señal de voz. Uno de los casos que los investigadores encuentran más desafiante es cuando el audio a analizar es capturado en un ambiente complejo, es decir, cuando el tipo o ubicación de sonidos varía continuamente. Conversaciones reales en nuestro día a día representan el mejor ejemplo.

Con la finalidad de encontrar una solución a la problemática es común que diversas áreas de la ciencia (neurociencia, robótica o inteligencia artificial) busquen inspiración en las ciencias biológicas y cognitivas; una gran cantidad de los enfoques inspirados se concentran en el modelado de áreas cerebrales y fenómenos cognitivos como la memoria, la percepción, el lenguaje, la categorización y el aprendizaje utilizando métodos de procesamiento de información [2].

Dotar a los modelos computacionales de procesamiento de audio con la capacidad humana de selección y atención de una fuente de sonido específica, como la voz, podría permitir que los modelos computacionales puedan lidiar con las características variables de un ambiente complejo y que puedan ser aplicados a sistemas de procesamiento de audio para mejorar la capacidad de atención de asistentes virtuales, para detectar violencia verbal en una ubicación, o por cualquier otro sistema computacional que requiera utilizar la voz de una persona pero con la menor cantidad de ruidos, por decir algunos ejemplos (ver figura 1).

Recientemente, se ha avanzado en la resolución de problemas de atenuación o eliminación de señales de ruido en mezclas de audio con voz en escenarios cada vez más difíciles, gracias a los métodos de aprendizaje profundo (las redes neuronales artificiales) [3]. Un tipo de red neuronal artificial muy utilizado en este problema es la red neuronal convolucional (CNN, por sus siglas en inglés), que tienen la capacidad de capturar patrones en las señales de audio. Se ha reportado que las redes neuronales convolucionales son más eficaces que las redes neuronales artificiales de tipo perceptrón multicapa [4] y más eficientes que las redes neuronales artificiales recurrentes [5] cuando se aplican al procesamiento de señales de audio.

Redes neuronales convolucionales

El diseño de las redes neuronales artificiales se hizo tomando como referencia el comportamiento y estructura de las neuronas en el cerebro humano, de modo que se construye una red de nodos interconectados y organizados en capas.

Los modelos computacionales con capacidad de atención podrían procesar mejor los ambientes sonoros con características más complejas.

Las redes neuronales convolucionales son una subcategoría de redes neuronales artificiales y son llamadas así porque aplican una operación matemática, llamada convolución, que transforma a las capas en convolucionales. El objetivo de una capa convolucional en una red neuronal artificial es extraer características particulares a partir de los datos de entrada en forma de matrices o vectores, comprimiéndolos para reducir su tamaño inicial. Los datos de entrada proporcionados pasan a través de un conjunto de filtros, creando nuevos datos llamados características extraídas. Finalmente, los nuevos datos se concatenan en una matriz, o vector de características, que serían los datos de entrada para la siguiente capa de la red neuronal.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Los mecanismos de atención

En los últimos años, y con la finalidad de ser más eficientes, las CNN han incorporado un nuevo concepto proveniente del mundo biológico denominado .ªtención". La atención es un mecanismo cognitivo de procesamiento de señales de nuestro cerebro. Permite a nuestros cerebros captar eficazmente varias características informativas de los distintos estímulos sensoriales. La fusión de los modelos basados en el uso de redes neuronales artificiales y los mecanismos de atención han ayudado a los modelos a identificar y enfatizar las características más importantes y minimizar las que son menos útiles para el logro de atenuar el ruido.

Los mecanismos de atención empleados en las redes neuronales artificiales tuvieron su origen como una mejora a la arquitectura codificador-decodificador que se utilizaban en el procesamiento del lenguaje natural (procesamiento de texto). Este mecanismo y sus variantes fueron aplicados posteriormente a otras áreas como la visión por computadora y el procesamiento de voz.

Actualmente, uno de los mecanismos de atención más utilizados debido a su gran efectividad es la atención de múltiples encabezados (*Multi-Head Attention*, por su traducción en inglés). La atención de múltiples encabezados es el uso simultáneo de múltiples mecanismos de autoatención en paralelo para identificar la información más útil. Cada mecanismo de autoatención se enfoca en diferentes partes de la información y luego las une para conservar los datos más importantes o relevantes que el mecanismo de atención puede identificar [6].

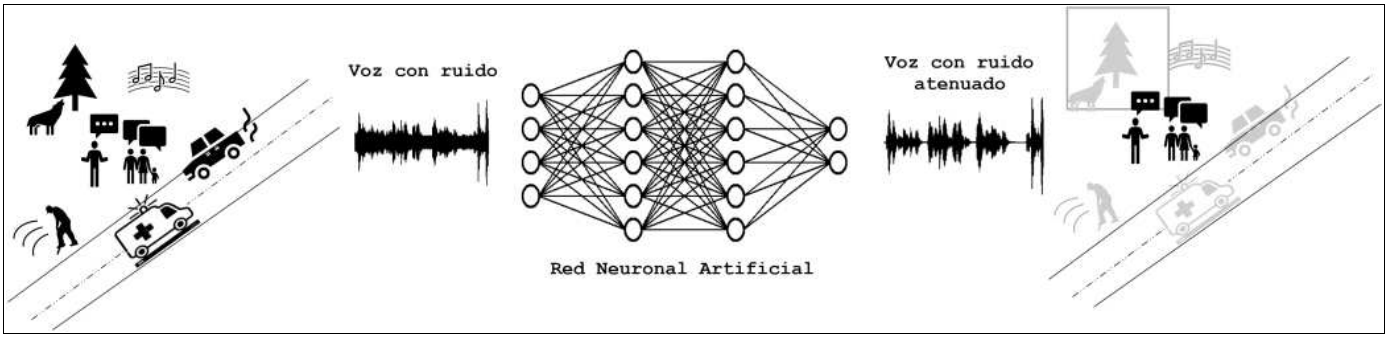


Figura 1. Representación gráfica del proceso de atenuación de ruidos utilizando redes neuronales artificiales.

Preparación de los datos

La señal de audio suele ser artificial en el entrenamiento de los modelos de redes neuronales artificiales ya que es necesario mezclar la señal de voz con las señales de ruido. Para emular los entornos naturales de ruido en las señales de voz, es necesario recopilar las señales de ruido de las bases de datos.

Estas bases de datos contienen sonidos ambientales de diferentes fuentes y en diferentes ambientes (ubicaciones físicas), donde el ruido agregado es la suma de sonidos con formas y magnitudes variables. El uso de sonidos ambientales repercute en la complejidad del problema de eliminar señales de ruido, ya que es difícil establecer un patrón en la señal. Los investigadores suelen utilizar valores de relación señal a ruido (SNR, por sus siglas en inglés) que van de +10 dB a -10 dB, donde una señal de audio a +10 dB podría ser un aula escolar (la señal de voz es mayor que la señal de ruido) y -10 dB podría ser una estación de tren (la señal de ruido es mayor que la señal de voz).

En el proceso de preparación de los datos, primero fue necesario identificar y obtener conjuntos de datos públicos que incluyeran señales de voz así como de ruidos en entornos interiores y al aire libre; a partir de estos conjuntos de datos se construyeron los datos de entrenamiento, validación y prueba para los dos modelos de redes neuronales artificiales convolucionales. Después de analizar las características y disponibilidad de los diversos conjuntos de datos públicos, se seleccionaron tres conjuntos de datos; como conjunto de datos de voz se utilizó TIMIT, y como conjunto de datos de ruido se utilizaron NoiseX-92 y DEMAND.

Una vez que se obtuvieron los datos a utilizar, se seleccionó la parte de entrenamiento del conjunto de datos TIMIT y el ruido de los conjuntos de datos NoiseX-92 y DEMAND para el conjunto de entrenamiento y validación. A continuación, se crearon cinco horas de mezclas de audio en clips de un minuto con SNR muestreados uniformemente entre -10 dB y 10 dB (con lo que los diferentes ruidos corrompen la señal de voz). Los clips de voz

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

y de ruido fueron elegidos aleatoriamente para crear las mezclas. Luego, se muestrearon los audios a 8 kHz para alimentar el modelo con las bandas de frecuencia más relevantes. Después, se calculó el espectro de la magnitud de la señal de audio mediante la transformada de Fourier de tiempo corto con un tamaño de 256 puntos de frecuencia, una ventana de longitud de trama de 32 ms y un solapamiento de 50 %. Por último, se normalizaron los datos de entrenamiento y validación con media cero y varianza unitaria para facilitar el proceso de entrenamiento.

Para evaluar los modelos se crearon mezclas de audio con SNRs uniformemente muestreados de -5 dB, 0 dB, 5 dB y 10 dB. La fase de la señal únicamente se conservó durante el proceso de predicción del modelo, para luego añadirla a la señal limpia estimada.

La figura 2(A) y 2(B) muestra la representación gráfica de la señal de voz en el dominio de tiempo y en el dominio de la frecuencia respectivamente; y en la figura 2(C) y 2(D) se muestra la misma señal de voz, pero corrompida con una señal de ruido ambiental.

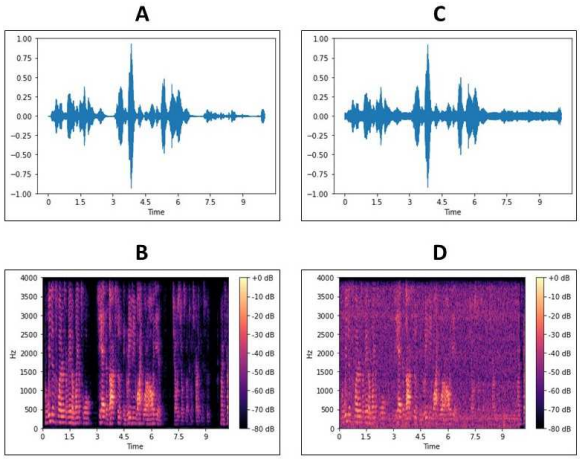


Figura 2. Representación gráfica de una señal de voz limpia y corrompida con ruido en el dominio de tiempo y frecuencia.

Arquitectura de la red neuronal

Se diseñaron e implementaron dos arquitecturas del modelo de red neuronal artificial de tipo convolucional, una arquitectura de tipo convolucional de una dimensión y una arquitectura de dos dimensiones.

La primera arquitectura, figura 3(A), tiene ocho capas convolucionales de una dimensión (Conv-1D) con una función de activación de unidad lineal rectificada paramétrica (PReLU, por sus siglas en inglés), que es una función matemática que se utiliza para determinar la salida de una neurona y permite que la red neuronal tome decisiones sobre qué tan activa o inactiva debe estar la neurona; seguidas por una capa lineal de 256 unidades sin función de activación. Se estableció el número de filtros en las primeras siete capas convolucionales en 64, con un solo filtro en la última capa de convolución. Además, se utilizaron filtros de tamaño 16 y capas de abandono con una tasa del 10 % en las capas uno, tres, cinco y siete (las capas de abandono, del inglés *dropout*, es una técnica de regularización para reducir el sobre-entrenamiento en redes neuronales artificiales). La segunda arquitectura, figura 3(B), tiene capas convolucionales de dos dimensiones (Conv-2D) con función de activación PReLU y una capa lineal final de 256 unidades sin función de activación. Se estableció el número de filtros en las primeras siete capas convolucionales en 64 (con un solo filtro en la última capa de convolución) pero con filtros de tamaño 4x4 en todas las capas convolucionales. Además, también se emplearon capas de abandono con una tasa del 10 % en las primeras siete capas convolucionales.

El módulo de atención fue incorporado antes de la capa lineal final en ambos modelos.

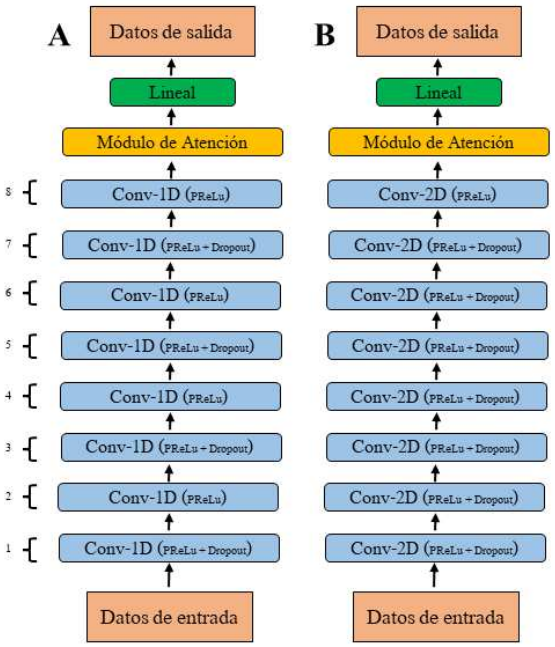


Figura 3. Arquitectura de las redes convolucionales de una dimensiones (A) y dos dimensiones (B).

Después de analizar las características del funcionamiento del sistema auditivo humano y de las teorías de la atención selectiva, se diseñó y construyó un módulo encargado de emular la capacidad de atención auditiva para ayudar a los modelos de redes neuronales artificiales a identificar las características más importantes de los datos para mejorar su rendimiento al atenuar el ruido.

El módulo de atención otorga mayor importancia a aquello que interesa con base en el objetivo establecido (la voz humana); tome como ejemplo cuando sostenemos una conversación con otra persona pero estamos en un lugar donde hay ruido ambiental (como música, tráfico,

animales, etc.), en este escenario nosotros somos capaces de comprender a la persona con la que conversamos porque centramos nuestra -atención- en su voz y descartamos todos los otros sonidos o ruidos que no son de nuestro interés. La función del módulo de atención es mejorar la capacidad de identificar y preservar la información de la voz humana y atenuar los ruidos ambientales del modelo de red neuronal convolucional.

El módulo de atención se basó en el uso de la atención de múltiples encabezados. Gráficamente, la arquitectura del módulo de atención se puede ver en la figura 4, y está compuesto de la siguiente forma:

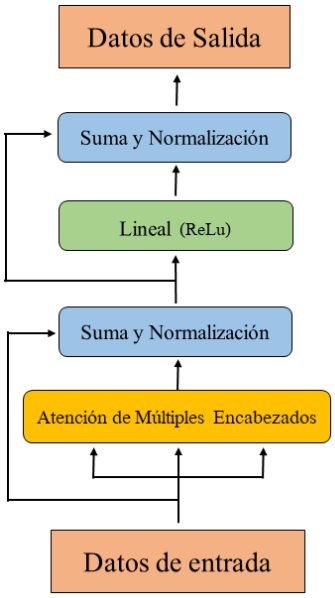


Figura 4. Arquitectura del módulo de atención.

1. Una capa de atención de múltiples encabezados (con cuatro encabezados).
2. Una capa de normalización, la cual suma los datos de entrada y salida de la capa de atención para después aplicar la normalización.
3. Una capa lineal con función de activación de unidad lineal rectificada (ReLu, por sus siglas en inglés).
4. Una capa de normalización, la cual suma los datos de entrada y salida de la capa lineal para finalmente aplicar nuevamente la normalización.

Resultados

La tabla 1 muestra los resultados de tres de las métricas estándar utilizadas habitualmente en la eliminación de ruido en las señales de voz: la evaluación perceptiva de la calidad de la voz (PESQ), la inteligibilidad objetiva a corto plazo (STOI), y la relación señal-distorsión invariable en escala (SI-SDR); en todas las métricas, cuanto más alto sea el valor obtenido, mejor. Los resultados se obtuvieron al promediar los valores de las métricas en audios con niveles de SNR de: -5 dB, 0 dB, 5 dB y 10 dB. La evaluación se realizó en los modelos convolucionales con y sin el módulo de atención para contrastar el resultado del impacto del módulo de atención y, además, los resultados se compararon con los resultados de otras investigaciones publicadas.

Tabla 1. Resultados promedio de las métricas PESQ, STOI y SI-SDR.

	PESQ	STOI	SI-SDR
CNN-1D	3.30	90.69 %	17.66
CNN-1D + atención	3.32	91.15 %	18.82
CNN-2D	3.30	89.99 %	16.57
CNN-2D + atención	3.32	90.61 %	18.07
U-FORMER [7]	2.93	92.01 %	-
DCN-SM [8]	2.83	90.40 %	-

En los resultados se aprecia que la inclusión del módulo de atención basado en la atención de múltiples encabezados mejoró en términos generales la capacidad de los dos modelos convolucionales utilizados. Respecto a los resultados de otras propuestas, se aprecia que con la métrica PESQ, los dos modelos aplicados obtienen mejores resultados, lo que se puede interpretar como que la percepción de la calidad de las señales de voz procesadas por nuestros modelos convolucionales es mejor que en otras propuestas; sin embargo, los resultados de la inteligibilidad objetiva de la señal de voz son más bajos que otras propuestas, según la métrica STOI.

Los resultados de nuestros experimentos también reflejan que la inclusión de un mecanismo de atención permite que las arquitecturas de los modelos de redes neuronales sean menos complejas sin sacrificar la eficiencia, ya que los dos modelos convolucionales utilizados tienen menos de dos millones de parámetros entrenables, a diferencia de otras propuestas en la literatura cuyos modelos llegan a tener decenas de millones de parámetros entrenables. Finalmente, consideramos esencial que los modelos evaluados fueran modelos de redes neuronales con arquitecturas sencillas. De este modo, concluimos que es posible crear modelos con arquitecturas sencillas (y con relativamente pocos parámetros entrenables) que obtengan resultados competitivos y consuman pocos recursos computacionales en su implementación.

Conclusiones

Aunque la mejora de la voz basada en el aprendizaje profundo ha demostrado ser muy eficiente al generar una señal de voz limpia con una calidad relativamente alta, todavía se considera que algunos ambientes con ruido son muy difíciles de procesar para una red neuronal arti-

ficial. En esta investigación se incorporó exitosamente un módulo de atención basado en la atención de múltiples encabezados en dos arquitecturas de red neuronal artificial convolucional (en una y dos dimensiones), los cuales mostraron mejoras en la tarea de atenuar y eliminar señales de ruido y mantener las señales de voz, tal como se demuestra cuando se evalúa con las métricas STOI, PESQ, SI-SDR.

En el futuro se planea continuar con la investigación en mejora de voz, incorporando una mayor variedad de señales de ruidos no estacionarios para permitir que el modelo se pueda evaluar en ambientes más complejos, así como proponer modificaciones del módulo de atención.✳

REFERENCIAS

1. Styles E. (2010) "Psicología de la atención". Centro de Estudios Ramón Areces.
2. Pfeifer, R., Iida, F. y Lungarella, M. (2014). Cognition from the bottom up: On biological inspiration, body morphology, and soft materials. *Trends in Cognitive Sciences* 18(8), 404-413.
3. Zacarias-Morales, N., Pancardo, P., Hernández-Nolasco, J. y Garcia-Constantino, M. (2021). Attention-Inspired Artificial Neural Networks for Speech Processing: A Systematic Review. *Symmetry*, 13(2)
4. Fu, S., Tsao, Y., Lu, X. y Kawai, H. (2017). Raw waveform-based speech enhancement by fully convolutional networks. *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 006-012.
5. Park, S. y Lee, J. (2016). A Fully Convolutional Neural Network for Speech Enhancement. *ArXiv*.
6. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. y Polosukhin, I. (2017). Attention is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000-6010.
7. Xu, X. y Hao, J. (2022). Improving Monaural Speech Enhancement with Multi-head Self and Cross Attention. *ArXiv*.
8. Pandey, A. y Wang, D. (2021). Dense CNN With Self-Attention for Time-Domain Speech Enhancement. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 29, 1270-1279.

SOBRE LOS AUTORES



Noel Zacarias-Morales es egresado del Doctorado en Ciencias de la Computación en la División Académica de Ciencias y Tecnologías de la Información en la Universidad Juárez Autónoma de Tabasco. El obtuvo su grado de Maestría en Administración de Tecnologías de la Información en la Universidad Juárez Autónoma de Tabasco en 2019. Sus intereses de investigación son el desarrollo de modelos basados en aprendizaje profundo (redes neuronales artificiales) aplicados al análisis y procesamiento de audio.



José Adán Hernández-Nolasco obtuvo el grado en Ingeniería Electrónica y Comunicaciones por la Universidad Autónoma de Nuevo León en 1996, la maestría en Ingeniería Electrónica (telecomunicaciones) por el Instituto Tecnológico y de Estudios Superiores de Monterrey en 2003, y el Doctorado en Óptica por el Instituto Nacional de Astrofísica, Óptica y Electrónica en 2012. Es Profesor Investigador de la Universidad Juárez Autónoma de Tabasco. Sus intereses de investigación incluyen el aprendizaje profundo, la lógica difusa, el IoT y la óptica.



Pablo Pancardo obtuvo la Maestría en Tecnologías de la Información en el Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM) en 1998, y el Doctorado en Ciencias de la Computación en la Universidad Juárez Autónoma de Tabasco (UJAT) en 2016. Actualmente es Profesor e Investigador del Grupo de Sensado y Procesamiento Inteligente de Datos de la UJAT. Sus intereses de investigación incluyen la inteligencia ambiental, la lógica difusa, el IoT y la HCI.



José Arturo Olvera-López es profesor de tiempo completo e investigador en la Facultad de Ciencias de la Computación en la Benemérita Universidad Autónoma de Puebla (Puebla, México). Sus intereses de investigación incluyen: reconocimiento de patrones, minería de datos, aprendizaje máquina, análisis digital de imágenes, y biometría, entre otras.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Redes neuronales físicamente informadas para aproximar la solución de ecuaciones diferenciales y sus aplicaciones en la vida real

Juan Daniel Meshir Vargas, Edgar Alejandro Guerrero Arroyo y Abel Palafox González

Introducción

En los últimos años, la Inteligencia Artificial, en particular las *Redes Neuronales Artificiales (ANNs)*, han sido utilizadas para resolver una gran cantidad de problemas en la industria; sin embargo, debido a su naturaleza inherente, también pueden ser utilizadas para resolver problemas matemáticos que enfrenten y modelen la realidad. Esto motiva el surgimiento de un nuevo modelo conocido como *Redes Neuronales Físicamente informadas (PINNs)*, una arquitectura de inteligencia artificial de aprendizaje profundo, derivada de las ANNs clásicas que, a diferencia de éstas, buscan implementar en su aprendizaje leyes físicas que gobiernan el comportamiento de los datos. Esto permite aproximar soluciones a *Ecuaciones Diferenciales (DEs)*, por sus siglas en inglés, y dar una alternativa para resolver estos problemas matemáticos que modelan situaciones reales. Para este propósito es preciso reconocer qué son las ecuaciones diferenciales y cuáles son sus aplicaciones en el mundo real.

Ecuaciones diferenciales y sus aplicaciones

Las DEs son expresiones matemáticas que tratan de describir la naturaleza del mundo que nos rodea. Buscan predecir el comportamiento de fenómenos físicos, estableciendo una relación entre una función con sus derivadas. Esto permite modelar, por ejemplo, las órbitas de los planetas y cuerpos celestes, facilitando el viaje del hombre al espacio; o modelar la propagación de epidemias como el COVID 19, ayudando a predecir la transmisión de la enfermedad; o entender cómo un murciélago puede "ver" a través del sonido, aprendiendo de ello para crear radares y sonares en aviones y submarinos, tal como en el sonar de la figura 1, en donde se utiliza la ecuación diferencial de onda reducida para recuperar la forma de un objeto desconocido D; entre muchas otras aplicaciones [1][2].

El principal estudio de DEs se centra en encontrar la solución que cumple la relación que establece una DE: la función de interés que estamos estudiando (el desplazamiento de un auto, las ganancias de una inversión, la temperatura de un comal sobre la estufa, etc). Sin embargo, en muchas ocasiones, hallar la expresión simbólica en términos matemáticos de la solución de una DE, no es posible o resulta ser muy complicado. Por consiguiente, en esos casos, se suele buscar una aproximación a la función mediante métodos numéricos, en donde los cálculos necesarios se realizan por medio de una computadora.

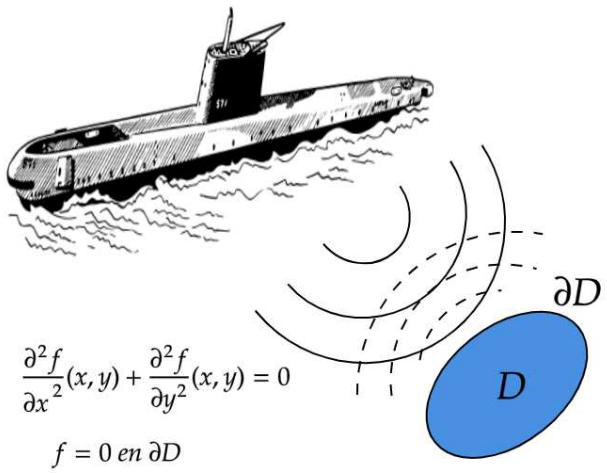


Figura 1. Aplicación de la ecuación diferencial de onda reducida al sonar para identificar la forma de un objeto desconocido D.

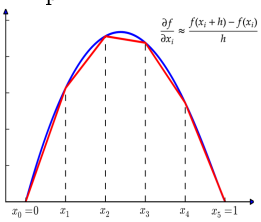
Las ecuaciones diferenciales son esenciales para comprender y modelar procesos naturales y físicos en el mundo real.

La idea de los métodos numéricos es expresar los términos simbólicos matemáticos como operaciones numéricas simples. Por ejemplo los métodos tradicionales para aproximar soluciones de DEs tales como *elementos finitos (FEM)*, *diferencias finitas (FD)* o *volúmenes finitos (FV)* [3], buscan aproximar la derivada analítica de una función por medio de la resta dos valores numéricos, como en la figura 2 *FD*. Además se debe elegir una canti-

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

dad finita de valores que pueden tomar las variables de la función, ya que una computadora no puede recibir una cantidad infinitas de valores para esas variables. Esto es conocido como *discretizar el dominio de la función*, figura 2 *FD*, donde para la variable x se toman 6 valores de x_0 a x_5 para aproximar $f(x)$, o en la figura 2 *FEM*, donde se toman 124 puntos del objeto D de la figura 1. Esta discretización permite representar el problema de la DE, que puede ser no lineal, como un sistema de ecuaciones lineales, en donde se tiene en cuenta la condición de la frontera del dominio. La frontera es el borde del dominio, por ejemplo, para la figura 1 el dominio es D y su frontera es el borde ∂D, con la condición de que en esta frontera la función se anula.

FD con 5 puntos de discretización.



FEM con 124 puntos de discretización.

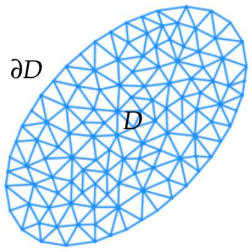


Figura 2. Método de diferencias finitas y método de elementos finitos.

Con estos métodos se puede aproximar soluciones de DEs; sin embargo, se debe invertir una gran cantidad de tiempo en la elección de la discretización del dominio de la función, en donde los valores de las variables deben de estar separados con cierta regularidad o cierta estructura, como en la figura 2 *FD* y *FEM*. La elección de esta discretización puede cambiar el comportamiento del resultado de la aproximación. En general, entre más puntos se tomen del dominio, mejor será la aproximación a la solución, pero esto implica un mayor costo de tiempo y recursos computacionales. Además, si se quiere aumentar la dimensión, es decir, si se aumenta el número de variables independientes en la función, el sistema de ecuaciones lineales cambia, haciéndose mas complejo. Por lo tanto, para problemas reales, en donde se requiere una gran cantidad de datos, estos métodos numéricos tradicionales resultan ser muy demandantes en términos computacionales, incluso utilizando el cómputo en paralelo.

Por otra parte, gracias a los avances tecnológicos que se han generado en las últimas décadas y con el procesamiento en paralelo, las ANNs se convierten en un modelo atractivo para solucionar DEs, ya que es posible trabajar con una gran cantidad de datos, optimizando el tiempo computacional y permitiendo aumentar la dimensión de los datos sin cambiar la estructura de las operaciones que se deben realizar. Pero ¿qué son las ANNs y cómo podemos utilizarlas para aproximar la solución de estos problemas matemáticos?

Redes neuronales artificiales (ANNs)

Las ANNs son un modelo de aprendizaje que trata de simular a las redes neuronales biológicas de nuestro cerebro, en donde se recibe un estímulo, dato de entrada, que pasa a través de unos nodos de procesamiento, y con base a esto, se genera una respuesta (dato de salida). Matemáticamente se puede representar una ANN como una función:

$$\mathcal{N}(E^i) = \hat{S}^i;$$

en donde el estímulo, se representa por el símbolo E^i, $\mathcal{N}$ lo procesa y $\hat{S}^i$ es la salida de la ANN. Una ANN, al igual que su contraparte biológica, cuenta con una unidad básica de procesamiento de la información llamada neurona; estas neuronas se ven representadas mediante círculos en la figura 3, cuya función es recibir la información, procesarla y enviarla a otras neuronas. Las neuronas se agrupan en capas, donde la primera capa (*Capa* 0), es llamada capa de entrada y cuenta con el número necesario de neuronas para recibir los datos de entrada E^i. La información procesada en esta capa es enviada a la siguiente, repitiendo este proceso hasta llegar a la última capa (*Capa* 3), conocida como capa de salida. La conexión entre dos neuronas es nombrada como *peso sináptico*, descrito en la figura 3 por w_{ij}^l, un valor que representa la magnitud de interacción entre cada neurona. Por último cada capa cuenta con una función de activación, que permite a la ANNs reflejar un comportamiento no lineal.

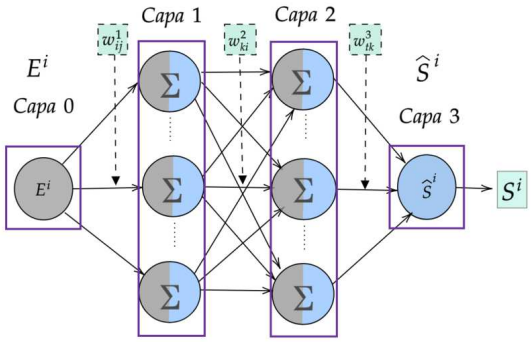


Figura 3. Representación de una ANNs.

Las ANNs son un modelo de aprendizaje supervisado, es decir, necesitan un conjunto de ejemplos de donde puedan aprender. Por ejemplo, en la figura 4, para enseñarle a una ANN a reconocer manzanas, se necesita un conjunto de ejemplos de frutas, con una etiqueta que indique si es una manzana o no. La respuesta de la ANN, se compara con la etiqueta con el fin de saber si se comete un error para corregir a la ANN, enseñándole a reconocer las manzanas.

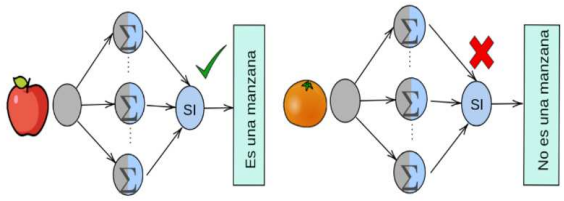


Figura 4. Ejemplos para entrenamiento de una ANNs.

El aprendizaje de la ANNs se puede expresar como calcular los pesos sinápticos adecuados que hagan que la salida de la ANN sea lo más parecida posible a la etiqueta real. Matemáticamente se puede expresar como encontrar los pesos que minimicen el error:

$$ENNs := \frac{1}{2}|N(E^i) - S^i|^2.$$

Las ANNs son utilizadas en muchas disciplinas debido a sus aplicaciones en: clasificación, reconocimiento de imágenes, reconocimiento del lenguaje, procesamiento de datos y más [4][5]. Además, en una publicación realizada en la década de los 80 por Cybenko [6], se muestra que las ANNs son capaces de aproximar funciones lineales y no lineales, reforzando el atractivo de su uso para representar la solución de una DE. Las ANNs clásicas ignoran las leyes físicas o reglas que puedan existir para modelar el comportamiento de los datos, si bien una red neuronal puede aprender la relación establecida por una DE, aproximar numéricamente sus coeficientes, o incluso determinar un número idóneo de capas o neuronas, son temas activos de investigación. Sin embargo, si se informa del conocimiento físico a las ANNs éstas serán capaces de aproximar soluciones de DEs.

Por otra parte si se añade la ley física al entrenamiento de una ANN, la capacidad predictiva del modelo puede mejorar. Se puede ver un ejemplo en la figura 5, en donde se quiere enseñar a una ANN a predecir la trayectoria de una flecha lanzada con un arco. En el aprendizaje se necesita un conjunto de ejemplos, los datos de entrada, los cuales pueden ser la masa de la flecha y la velocidad con la que es lanzada, y las etiquetas, que serán la trayectoria que recorre la flecha (representados por la trayectoria celeste en la figura 5). No obstante, se sabe que la trayectoria de las flechas se rige por la segunda ley de Newton. Una ANN clásica no tiene en cuenta esta ley, así que puede predecir, como trayectorias válidas, flechas que nunca caen. Con base a miles o millones de datos de ejemplo, podría aprender que todas las flechas caen en algún momento. Pero ¿y si no tenemos miles o millones de datos ? Si le enseñamos a la ANN la segunda ley de Newton, sabrá que todas las flechas caen, por consiguiente no se necesitarán muchos datos para predecir correctamente la trayectoria de las flechas, o representar de mejor forma la realidad. Esta es la esencia de informar físicamente a las ANNs.

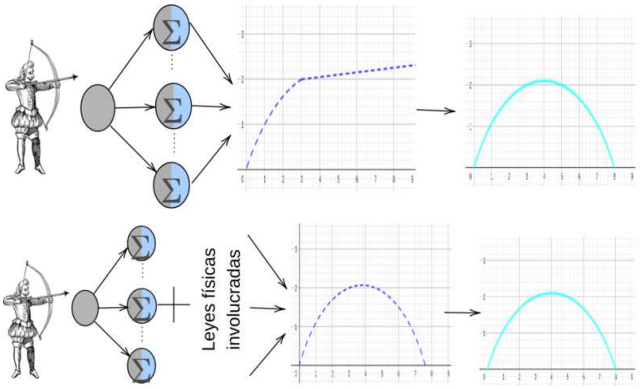


Figura 5. Ejemplos para entrenamiento de una ANNs.

Al agregar la ley física al aprendizaje de la ANN se pueden resolver modelos basados en DEs con aplicaciones a la vida real.

Uno de los pioneros en proponer la aproximación a la solución de DEs mediante ANNs es el artículo de Lagaris [7] de 1998, en donde se utilizan las ANNs para representar la solución de una DE, es decir, que ésta se puede describir por $\mathcal{N}(x_i)$, donde x_i son puntos del dominio de la función (datos de entrada E^i). Además Lagaris propone que dicha representación se anule en la frontera del dominio, para que las condiciones de frontera se cumplan siempre de manera exacta.

En Raissi [8] del 2019 se presenta el término y arquitectura de las PINNs para resolver DEs. En este artículo se plantea tomar puntos dentro y en la frontera del dominio de manera arbitraria, planteando el problema como uno en donde se debe minimizar el error de la ANN al

ser evaluada en dichos puntos con respecto a la DE y las condiciones en la frontera. Matemáticamente informar a una ANNs físicamente, se refiere a encontrar los pesos sinápticos adecuados que minimicen el error:

$$EPINNs := \frac{1}{2}|\mathcal{N}(\hat{x}_i) - y_i|^2 + \frac{1}{2}|\lambda(\mathcal{N}(x_i))|^2,$$

en donde x_i son puntos del dominio y $\hat{x}_i$ puntos de la frontera. Además λ representa el operador diferencial, esto es, la relación que debe cumplir la función a aproximar con respecto sus derivadas. Esta es la diferencia entre una ANN y una PINN, en donde el segundo sumando añade la ley física del problema, es decir, la información de cómo se deben comportar los datos, buscando minimizar el error $EPINNs$ en lugar del error $ENNs$.

Luego, Sirignano [9], muestra cómo se pueden utilizar las PINNs para resolver DEs en dimensiones altas, soluciones que tienen gran cantidad de variables independientes, algo que con otros métodos numéricos, como por ejemplo FEM, es muy complicado. Además, se muestra que al variar ligeramente el problema, no hay que resolver nuevamente la DE, sino evaluar la PINN ya entrenada con el nuevo conjunto de datos de las condiciones en la frontera. Esta es una de las ventajas principales de las PINNs sobre otros métodos numéricos tradicionales que no permiten una evaluación en un conjunto de datos diferente.

Las PINNs son capaces de aproximar soluciones de EDs, en dominios generales, sin necesidad de tomar puntos de forma regular en el dominio o su frontera

Aplicaciones de PINNs

Mencionaremos algunas de las aplicaciones de las PINNs para aproximar las soluciones a DEs.

La ecuación de Schrödinger es una DE que describe el comportamiento de una partícula de naturaleza ondulatoria de un sistema cuántico a través del tiempo. Es el semejante, en la cuántica, a la segunda ley de Newton en la mecánica clásica. Esta ecuación es de suma importancia ya que, por medio de ésta, se pueden estudiar muchos de los fenómenos que surgen en la mecánica cuántica, tal como el efecto túnel. Además tiene aplicaciones en la propagación no lineal de ondas en fibra óptica, ondas guiadas y ondas plasma [8]. Esta ecuación tiene soluciones complejas y varía en el tiempo; sin embargo, esto no es impedimento para utilizar PINNs.

En la figura 6 podemos ver la aproximación de la solución, en donde se describe el comportamiento de la magnitud de la solución en el espacio en dos tiempos particulares diferentes. El error relativo entre la solución y la aproximación es tan pequeño que la diferencia no se puede apreciar.

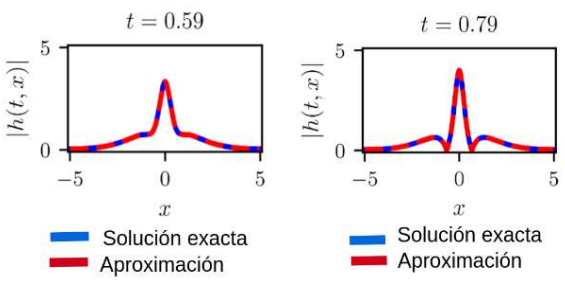


Figura 6. Aproximación a la solución de la ecuación de Schrödinger mediante PINNs tomada de [8].

Otro ejemplo para el cual se pueden utilizar PINNs es para aproximar la solución de **la ecuación de Navier-Stokes**. Ésta es una ecuación diferencial que describe el movimiento de un fluido. Es muy útil para describir muchos fenómenos que ocurren en la naturaleza y en la ingeniería, tales como: corrientes oceánicas, flujo de agua en una tubería, flujo de aire alrededor de un ala, presión de la sangre y el análisis de contaminación [8]. En un escenario realista, tal como el movimiento de un fluido incompresible a través de un cilindro circular, figura 7, se pueden utilizar PINNs para aproximar todo el campo de presión del mismo, como se ve en las figuras 8 y 9. Esto muestra la gran capacidad de las PINNs para trabajar con datos y escenarios reales, independientemente de la dimensión.

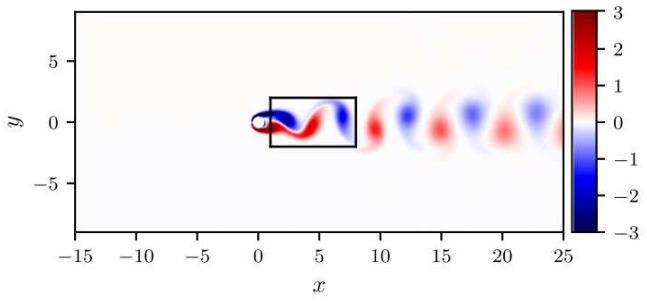


Figura 7. Flujo incompresible a través de un cilindro circular, tomada de [8].

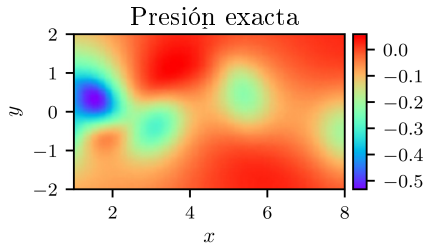


Figura 8. Solución exacta del campo de presión del movimiento de un fluido, tomada de [8].

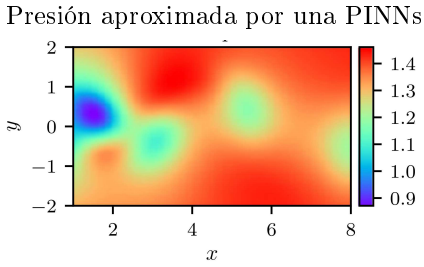


Figura 9. Aproximación mediante una PINNs al campo de presión de un escenario realista de movimiento de un fluido, tomada de [8].

Por último, si consideramos un objeto D como en la figura 10, para la **ecuación de la onda reducida** mencionada en la figura 1, y tomamos puntos aleatorios del dominio y su frontera, podemos aproximar la solución de cómo se va a comportar una onda dispersada por este objeto, tal como se verifica en la figura 11.

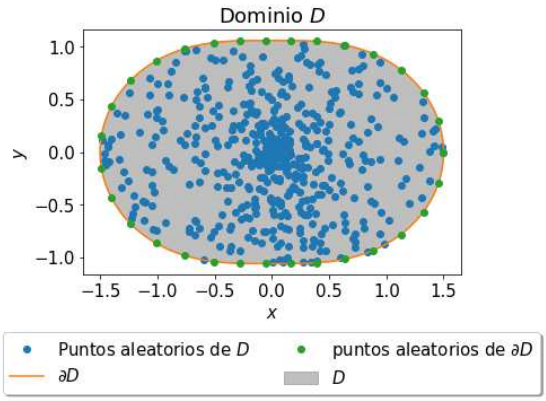


Figura 10. Objeto D con puntos aleatorios dentro y sobre su frontera.

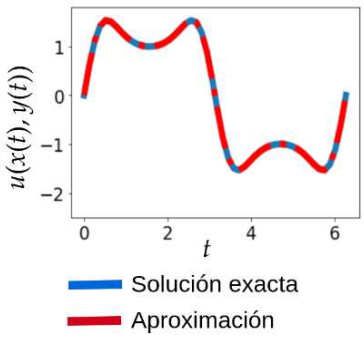


Figura 11. Aproximación a la solución de la ecuación de onda reducida a través de una PINN.

De manera similar a la aproximación de la ecuación de Schrödinger, la solución aproximada es tan buena que no se puede apreciar la diferencia entre ésta y la solución exacta.

Estas aplicaciones muestran la capacidad de las PINNs para aproximar soluciones de problemas reales con los que métodos tradicionales tienen dificultades.

Discusión sobre las ventajas de las PINNs

Mencionaremos aquí un listado de las ventajas, que consideramos, poseen las PINNs frente a la solución de una DE.

1. Se pueden tomar puntos de forma aleatoria del dominio y frontera de la función, sin invertir gran cantidad de tiempo en la elección de los puntos, lo que permite resolver en cualquier dominio, sin importar su complejidad. Este hecho no es posible en otros métodos tradicionales

2. Es posible trabajar con DEs de alta dimensión, funciones con muchas variables independientes, manteniendo la simplicidad del método.

3. Permiten evaluar cualquier punto del dominio o su frontera, sin importar si este punto está, o no, en los conjuntos de ejemplos elegidos para entrenar la PINN.

4. Se puede utilizar cómputo de alto rendimiento en el manejo de las PINNs, lo cual reduce drásticamente el costo computacional. Esto puede hacer que el programa pase de durar meses a horas.✷

Agradecimientos

Agradecemos las facilidades brindadas al proyecto "Modelación y pronóstico de fenómenos físicos mediante redes neuronales artificiales" por parte del Centro de Análisis de Datos y Supercómputo (CADS) de la Universidad

de Guadalajara, a través del uso de la Supercomputadora Leo Atrox.

REFERENCIAS

1. Braun, M. (1993). Differential Equations and Their Applications. Cuarta edición, Springer Science+Business Media, LLC.
2. Beilina, L. y Klibanov, M. (2013). Approximate global convergence in imaging of land mines from back scattered data, Applied Inverse Problems. Springer, New York.
3. LeVeque, R.J. (2005). Finite Difference Methods for Differential Equations. Primera Edición. University of Washington.
4. Bishop-Christopher, M. (2006). Pattern Recognition and Machine Learning. Springer Science+Business Media, LLC.
5. Murray, A. (1995). Applications of Neural Networks. Primera edición. Springer New York, NY.
6. Cybenko, G. (1989). Approximation by Superpositions of a Sigmoidal Function. Math. Control Signals Systems, 2, 303-314.
7. Lagaris, I., Likas, A. y Fotiadis, D. (1998). Artificial neural networks for solving ordinary and partial differential equations. *Institute of Electrical and Electronics Engineers* (IEEE), 9, 987-1000.
8. Raissi, M., Perdikaris, P. y Karniadakis, G.E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707.
9. Sirignano, J. y Spiliopoulos, K. (2019). Mean Field Analysis of Neural Networks: A Law of Large Numbers. *Mathematics of Operations Research*, 47(1), 120-152.

SOBRE LOS AUTORES



Juan D. Meshir Vargas es Licenciado en Matemáticas por la Universidad Distrital Francisco José de Caldas de Bogotá Colombia y actualmente cursa la Maestría en Ciencias Matemáticas en el Centro Universitario en Ciencias Exactas e Ingenierías CUCEI, de la Universidad de Guadalajara. Sus intereses en investigación se centran en Deep Learning, modelación numérica de Ecuaciones Diferenciales y Problemas inversos.



Edgar Alejandro Guerrero Arroyo recibió el grado de Doctor en Ciencias de la Computación en el Centro de Investigación en Matemáticas, CIMAT, en el 2016. Es profesor de tiempo completo en el Departamento de Matemáticas, del Centro Universitario en Ciencias Exactas e Ingenierías CUCEI, de la Universidad de Guadalajara. Sus líneas de investigación se centran en modelación numérica de ecuaciones diferenciales, optimización numérica y ciencia de datos.



Abel Palafox González doctor en Ciencias con orientación en Ciencias Computacionales en el Centro de Investigación en Matemáticas, CIMAT A. C. Es profesor investigador en el Departamento de Matemáticas, del Centro Universitario en Ciencias Exactas e Ingenierías CUCEI, de la Universidad de Guadalajara, y miembro del Sistema Nacional de Investigadores, Nivel I. Sus intereses científicos incluyen los Problemas Inversos y métodos numéricos para Ecuaciones Diferenciales Parciales.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

La inteligencia artificial como mecanismo de investigación en astronomía de multimensajeros, hechos y perspectivas

Alejandro Casallas-Lagos, Javier M. Antelis, Claudia Moreno y Michele Zanolin

Las ondas gravitacionales

Una de las predicciones más intrigantes y enigmáticas de la teoría general de la relatividad (TGR) propuesta por Einstein en 1916, corresponde a la existencia de las Ondas Gravitacionales (OG). Se trata de un fenómeno ondulatorio que puede describirse como perturbaciones del espacio-tiempo, producidas por eventos astronómicos de gran escala, y cuya detección resulta uno de los problemas más desafiantes e intrincados de la física contemporánea, esto, debido a la complejidad de su detección (Janka, 2021). La amplitud de una OG es equivalente a una millonésima parte del diámetro de un protón (del orden de 10^{-21} m a 10^{-23} m), lo cual hace que su exploración sea extremadamente dificultosa debido fundamentalmente a dos razones: 1.) Siendo un fenómeno de proporciones comparables con las dimensiones atómicas es altamente sensible a fuentes de ruido que pueden alterar fácilmente su medición, 2.) La sensibilidad que deben alcanzar los dispositivos experimentales para detectar tales eventos es altísima. A pesar de las dificultades técnicas descritas previamente, el 14 de septiembre del año 2015 la colaboración LIGO (*Laser Interferometer Gravitational wave Observatory*, por sus siglas en inglés) publicó el primer evento confirmado de detección de OG producida por la fusión de dos agujeros negros, este evento fue denominado GW150914, las siglas denotan el tipo de evento, GW para onda gravitacional por sus iniciales en inglés (*Gravitational Wave*) seguido de la fecha de detección (año/mes/día).

En este escrito presentamos una aplicación puntual de la inteligencia artificial en el contexto de la astronomía de multimensajeros.

Relatamos los descubrimientos más relevantes en el área, poniendo en contexto las perspectivas actuales y destacando los retos más relevantes para el futuro.

Supernovas: Las mayores explosiones observadas en la naturaleza.

Desde la primera detección de OG la colaboración LIGO ha confirmado más de 90 eventos que corresponden, al igual que GW150914, a OG producidas por sistemas binarios (es decir, pares de objetos astronómicos como estrellas o agujeros negros que orbitan respecto a un centro común antes de fusionarse), sin embargo, nuevas fuentes de radiación gravitacional están siendo estudiadas, y hacen parte de la siguiente generación fenómenos físicos asociados a OG que esperan ser detectados por LIGO en los próximos años. El principal objetivo de la colaboración LIGO está enfocado en la detección de OG producidas por explosiones estelares, conocidas como Supernovas (SN).

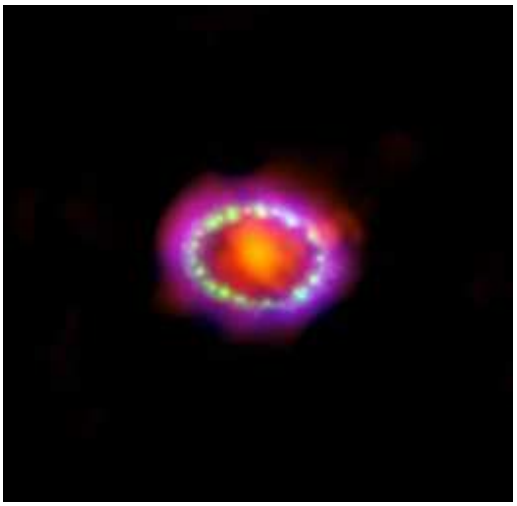


Figura 1. La supernova SN1987A, una supernova tipo II detectada en la Gran Nube de Magallanes (Burrows, 2017), una galaxia cercana a la Vía Láctea.

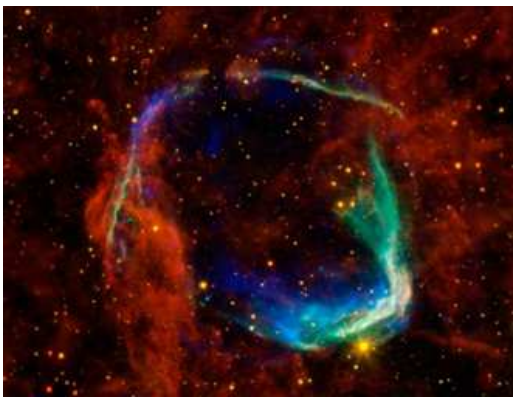


Figura 2. Remanente de supernova RCW86, acumulación de gas que se encuentra a muy altas temperaturas como consecuencia de la explosión de la estrella localizada en esa región del espacio (Williams, 2011).

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Esta clase de fenómenos físicos se presentan con baja frecuencia en nuestra galaxia, aproximadamente dos cada siglo (Janka 2012), sin embargo, han sido registradas, estudiadas y analizadas desde tiempos inmemoriales por astrónomos en diversas regiones del mundo, siendo la Supernova RCW-86 [ver figuras ??, ??] la primera registrada en el año 185 d.c por astrónomos chinos. Las supernovas son consideradas eventos astronómicos de gran impacto, la cantidad de energía liberada durante la explosión produce un brillo que ha llegado a ser millones de veces mayor que la luminosidad de nuestro sol (Janka 2020), permaneciendo visibles en el cielo por semanas e incluso años.

La supernova RCW-86 fue visible durante ocho meses en el cielo, y su luminosidad fue tan alta que incluso era visible a plena luz del día.

Caracterizadas por su inmensa luminosidad, liberación de energía y generación de OG, las supernovas se dividen en dos grandes grupos, supernovas tipo I y II, siendo las de tipo II más energéticas y con explosiones más violentas comparadas con aquellas de tipo I (Janka 2012). Al mecanismo de explosión de las supernovas tipo II se les conoce como colapso de núcleo, por lo cual son identificadas en la literatura científica como Colapso de Supernova (CSN) (Janka 2012). El 23 de febrero de 1987 los telescopios en tierra detectaron una fuente luminosa de alta intensidad cerca de la Vía Láctea, poco tiempo después los astrónomos confirmarían que este fenómeno correspondió a una supernova tipo II que en notación astronómica fue descrita como SN1987A (SN por supernova [Figura ??], 1987 por el año de descubrimiento y la letra A por la aparición en el cielo, fue la primera detectada). Previo a la detección en los observatorios astronómicos, otras fuentes de información ya reportaban una actividad inusual en el cielo, los observatorios de neutrinos informaban sobre un evento especial. Los neutrinos son partículas neutras (sin carga eléctrica), muy pequeñas, de hecho, su nombre proviene del italiano "neutrón pequeño", descubiertas en 1956 por Clyde Cowan y Frederick Reines. Ante las detecciones en los dos observatorios (astronómicos y de neutrinos) los científicos confirmaron que las explosiones por CSN son fuentes astronómicas que producen información óptica, visible en los telescopios, información de tipo energética detectable en los observatorios de neutrinos, y muy probablemente también OG (Janka 2012). Esta combinación de factores ópticos/energéticos/OG es denominada Astronomía de Multimensajeros y es en la actualidad un tópico de investigación de gran actividad, dadas sus implicaciones e interconexiones entre la física, la astronomía, la astrofísica y la computación.

Ondas gravitacionales producidas por colapso de supernova

A pesar de no haber alcanzado aún la primera detección de OG producidas por CSN, su estudio se ha desarrollado intensamente desde los años 60 del siglo pasado (Bailes 2021), permitiendo ampliar la cantidad de información disponible para obtener las piezas del rompecabezas que configura el entendimiento de este complejo fenómeno. Diversos grupos de investigación alrededor del mundo han hecho uso de los avances teóricos en la caracterización de CSN junto a robustas herramientas computacionales y métodos de simulación numérica, para predecir la forma asociada a una OG producida por un CSN (Bailes 2021). Gracias a las simulaciones numéricas, ahora sabemos que las OG producidas por CSN son esencialmente estocásticas, es decir, aleatorias, lo cual se traduce en que cada OG producida por un CSN es diferente a cualquier otra (Janka 2012). A pesar de este hecho, algunas características deterministas surgen de las simulaciones numéricas, es decir patrones repetitivos pueden determinarse dentro de la señal. En la figura 3 presentamos una forma de OG producida por CSN construida a través de simulaciones numéricas, y en la figura 4 una característica determinista que surge de las simulaciones numéricas denominada el modo gravitacional, o simplemente el modo g.

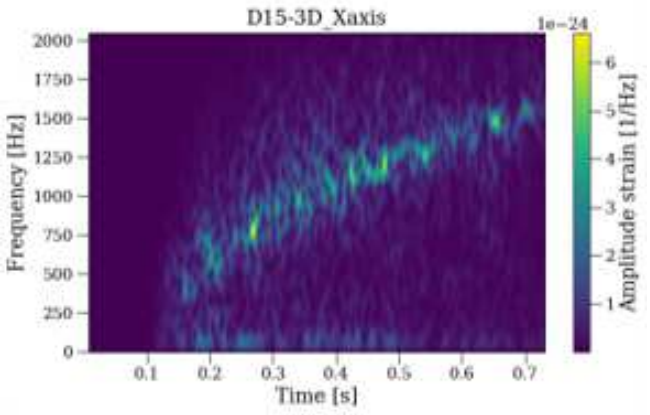


Figura 3. Onda gravitacional de CSN obtenida a través de simulaciones numéricas.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

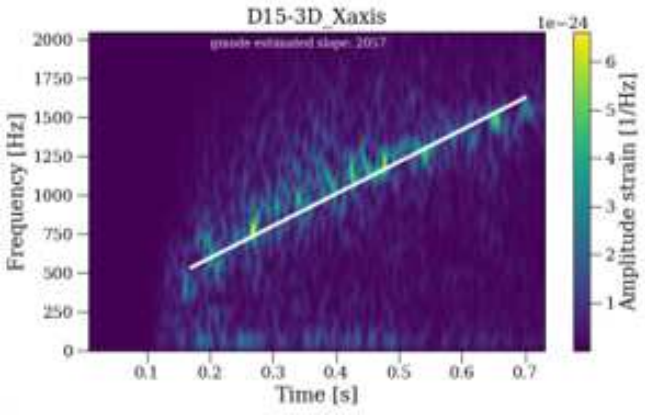


Figura 4. crecimiento de la frecuencia que forma el arco en colores verde y amarillo en las figuras 3 y 4, está presente en todas las simulaciones numéricas de OG de CSN. A este crecimiento se le denomina el modo gravitacional, en la figura 4 se presenta la estimación de la pendiente del modo g para una OG de CSN.

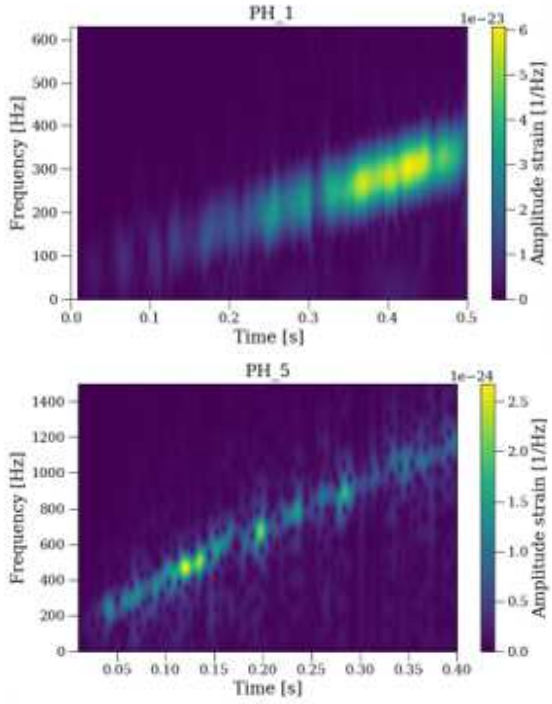


Figura 5. Corresponden a OG sintéticas que simulan la señal obtenida de un CSN. Estas formas de onda se caracterizan por presentar un crecimiento lineal en el espectrograma de la señal a diferencia de su contraparte obtenida a través de simulaciones numéricas como se ilustra en la figura 3.

Paralelamente con el desarrollo de las OG de CSN han surgido diferentes formas de OG sintéticas, es decir, señales que no representan una OG real. Estas señales son obtenidas de un modelo teórico que simula una OG de CSN y que presenta características particulares que sirven para estudiar los componentes de la señal presentes en ella, como por ejemplo el modo g. En la figura 5 presentamos una señal sintética de OG de CSN que tiene como característica primordial un modo g explícitamente lineal, a diferencia de su contraparte obtenida a través de simulaciones numéricas, define una pendiente que puede ser calculada directamente dada su geometría.

Uno de los objetivos principales de la astronomía de multimensajeros consiste en usar las propiedades físicas de luz, neutrinos u OG para obtener propiedades de la fuente que emite esta información (Antelis 2022). Dado que las OG producidas por CSN son de naturaleza estocástica, es necesario involucrar estrategias computacionales para poder extraer de manera eficiente información sobre las supernovas que dan lugar a las OG. En este sentido, las herramientas de inteligencia artificial (IA) han adquirido una relevancia particular en la última década dada su versatilidad, bajo costo computacional en comparación con otros modelos, variedad de arquitecturas y rendimiento.

Inteligencia artificial, aprendizaje automático y aprendizaje profundo

La IA busca emular, o al menos reproducir parcialmente, la inteligencia humana empleando computadoras, códigos de programación y algoritmos que imiten de forma aproximada los procesos de razonamiento que llevan al ser humano a resolver problemas (Yang 2022). Dentro de la IA se encuentra ubicado el Aprendizaje Automático (AA), cuyo principal fin consiste en desarrollar sistemas computacionales con la habilidad de aprender; a partir de la experiencia. Es decir, usando como herramienta fundamental el entrenamiento del sistema involucrando múltiples ejemplos (Yang 2022), este entrenamiento es puesto a prueba usando un nuevo conjunto de datos desconocido para la red (conjunto de prueba) con el fin de encontrar o caracterizar patrones contenidos en los datos. En este sentido el AA hace uso de dos diferentes tipos de datos: Entrenamiento y Prueba. Una herramienta ampliamente aplicada; es el Aprendizaje Profundo (AP). Este conjunto de estructuras se organiza en unidades denominadas Neuronas, capaces de enviar y recibir señales. Las neuronas se agrupan en estructuras llamadas Capas, que interactúan entre ellas a través de funciones de activación (imitando la función de la sinapsis en el cerebro). La sinapsis son impulsos eléctricos que activan las neuronas, desencadenando una larga cadena de impulsos químicos y eléctricos. El AP está inspirado en el funcionamiento del cerebro humano, y responde a la necesidad de encontrar soluciones a preguntas que involucran grandes cantidades de datos en tiempo real; estos problemas pueden clasificarse en dos grandes clases: Clasificación y Regresión, una red neuronal ejecuta una tarea de cla-

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

sificación cuando determina si los elementos contenidos en el conjunto de datos pertenecen a una categoría en específico. De otro lado, si el resultado final de la red consiste en obtener un número, se denomina un Regresor. A continuación, describiremos la estructura general de una red neuronal completamente conectada (RNCC) (Yang 2022), cuya aplicación es un elemento central en nuestro estudio:

- RNCC: Consiste en una estructura general de tres capas: capa de entrada, capas ocultas y capa de salida. Todas las neuronas en cada capa están conectadas entre sí con las neuronas de la capa siguiente (de allí su nombre), y son conocidas por resolver problemas complejos de clasificación o regresión, como por ejemplo, reconocimiento de imágenes y, predicción de costos asociados a bienes y servicios.

Redes Neuronales Artificiales y su Aplicación en la Astronomía de Multimensajeros

Con el desarrollo de las simulaciones numéricas que describen la forma de OG asociada a un CSN, han surgido paralelamente, diferentes herramientas computacionales para efectuar búsqueda y detección de este tipo de OG. Una de las más exitosas es: *Coherent Wave Burst* (cWB), un algoritmo que opera sobre los datos extraídos por LIGO y que detecta OG de CSN sin necesidad de tener un conocimiento previo de la OG que se pretende encontrar (Drago 2021). cWB genera para cada detección de OG de CSN un mapa de pixeles que contiene la información sobre la detección [ver Figuras 6, 7], en particular estos diagramas de pixeles describen como se acumula la frecuencia obtenida en un intervalo de tiempo en el cual la OG es detectada.

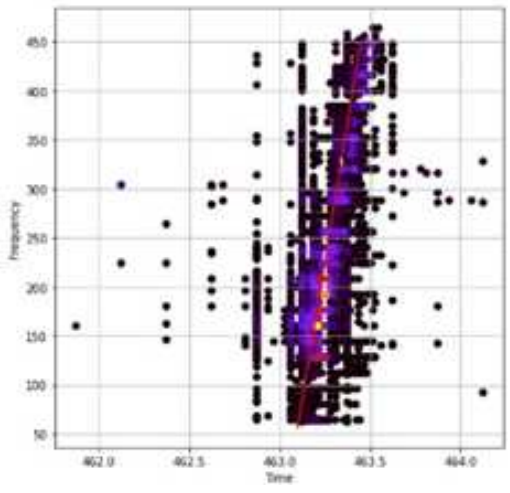


Figura 6. Diagrama de pixeles generado por cWB asociado a la detección de una OG de CSN.

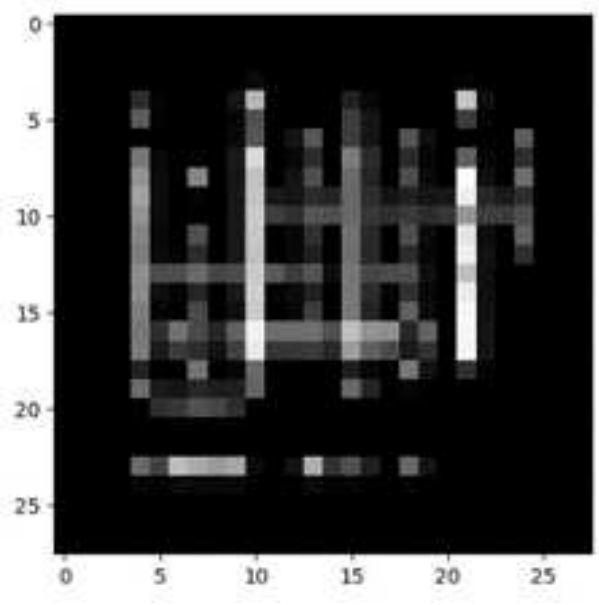


Figura 7. Procesamiento del diagrama de pixeles presentado en la figura **??**, esta imagen corresponde al diagrama en resolución 28x28, en espacio de color gris y negro.

Dado que contamos en la actualidad con cWB, una herramienta eficiente y confiable para la detección de OG de CSN, podemos preguntarnos, ¿Cómo hacer uso de las detecciones de OG de CSN generadas por cWB para hacer astronomía de multimensajeros? Para responder a esta pregunta presentamos un plan general basado en aprendizaje automático (Antelis 2022), particularmente en la aplicación de un modelo de red neuronal profunda completamente conectada que extrae información de la OG de CSN, en particular, esta red estima la pendiente del modo g presente en la señal de OG detectada por cWB. La figura 8 describe de manera general la metodología implementada en este escrito.

Para cumplir con la estimación de la pendiente del modo g desarrollamos la siguiente metodología pasos metodológicos:

1. Conjunto de datos de entrenamiento: Efectuamos detecciones de señales de OG sintéticas, y recolectamos todos los diagramas de pixeles generados por cWB, en estos mapas de pixeles se almacena la información del evento y su correspondiente pendiente del modo g.

2. Conjunto de datos de prueba: Efectuamos detecciones de señales de OG de CSN producidas por simulaciones numéricas, y recolectamos todos los diagramas de pixeles generados por cWB. A diferencia del contenido de las ondas sintéticas la pendiente del modo g aquí no es conocida.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

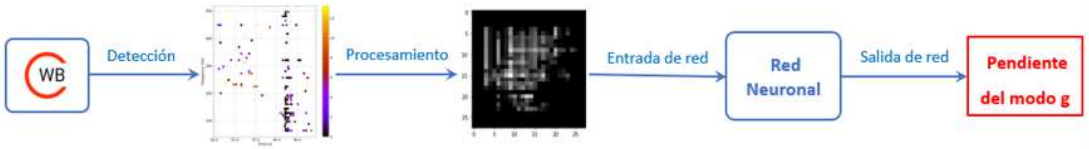


Figura 8. Esquema general de aplicación de la propuesta metodológica para estimar la pendiente del modo g presente en una OG de CSN. Partiendo de la izquierda: Buscamos detecciones en cWB, obtenemos la detección decodificada como un diagrama de pixeles, procesamos las detecciones generando imágenes de 28x28 pixeles de resolución en espacio de color negro y gris. Esta imagen servirá como entrada de la red neuronal profunda totalmente conectada, que producirá como salida la estimación de la pendiente del modo g.

3. Procesamiento de datos: Las detecciones de cWB, es decir, los mapas de pixeles, son procesados de tal forma que dichos mapas se transforman en imágenes planas, cuyo espacio de color es negro y gris, con resolución 28x28. Es decir, imágenes de dos dimensiones (alto y ancho), siendo cada dimensión de 28 pixeles. Esto es equivalente a tener un arreglo numérico de 784 componentes (que proviene de multiplicar las dimensiones alto x ancho, 28 x 28 = 784) en el cual se almacena toda la información del mapa de pixeles. Este procedimiento es efectuado para alimentar a la RNCC con un arreglo numérico que decodifica la información contenida en la detección. El resultado de este procesamiento se ilustra en la figura **??**

3. RNCC para estimar la pendiente del modo g: Usamos las formas sintéticas (gracias a su crecimiento lineal) para entrenar un algoritmo de inteligencia artificial, en particular una RNCC. Dotando a la red de múltiples ejemplos de pendientes sintéticas. Cuando la red ha sido entrenada, es aplicada sobre un conjunto de prueba que contiene OG de CSN generadas por simulaciones numéricas. La RNCC actúa entonces sobre ejemplos que no ha visto antes, y, con base en su entrenamiento, asocia el valor de la pendiente estimada para el modo g a cada OG de CSN. En la figura 8 se presenta la estructura general de la metodología implementada en este documento.

Discusión y resultados

En este documento ilustramos una forma de aplicar una metodología de aprendizaje profundo en el contexto de la astronomía de multimensajeros, estimando la pendiente del modo g contenida en una OG de CSN. En la figura 9 presentamos el panorama general de aplicación de nuestra metodología que involucra los conjuntos de entrenamiento, prueba, RNCC y la estimación del modo g. Como base, esta metodología, implementamos un conjunto de OG sintéticas cuya pendiente del modo g es fácilmente reconocible, extraíble y medible para entrenar la red neuronal. Usando las detecciones de cWB (Drago 2021) diseñamos una RNCC que usa como conjunto de prueba OG de CSN que se representan los modelos reales de detección en datos de LIGO. La implementación de esta metodología revela que es posible vincular la inteligencia artificial, en particular la arquitectura de RNCC, para estimar una propiedad determinista presente en una OG cuya naturaleza es esencialmente estocástica. ✷

REFERENCIAS

1. Drago, M. et al. (2021). Coherent WaveBurst, a pipeline for unmodeled gravitational-wave data analysis. *SoftwareX*, 14.

2. Janka, T. (2012). Explosion Mechanisms of Core-Collapse Supernovae. *Annual Review of Nuclear and Particle Science*, 62, 407-451.

3. Yang, X., Yan, J., Wang, W. et al. (2022). Brain-inspired models for visual object recognition: an overview. *Artif Intell*, 55, 5263–5311.

4. Antelis, J. y Cavaglia, M., Hansen, T., Morales, M. et al. (2022). Using supervised learning algorithms as a follow-up method in the search of gravitational waves from core-collapse supernovae. *Phys. Rev. D*, 105, 54-69.

5. Bailes, M., Berger, B.K., Brady, P.R. et al. (2021). Gravitational-wave physics and astronomy in the 2020s and 2030s. *Nat Rev Phys 3*, 344–366.

6. Burrows, A. et al. (2017). Supernova 1987A. Recuperado el 6 de agosto de 2017, de `https://www.nasa.gov/chandra/images/supernova-1987a.html`.

7. Pazos R., Pérez J., et al. (2005). A Domain Independent Natural Language Interface to Databases Capable of Processing Complex Queries. *Lecture Notes in Artificial Intelligence*, 3789, 833–842.

8. Williams, B. et al. (2011). Oldest Recorded Supernova. Recuperado el 7 de agosto de 2017, de `https://www.nasa.gov/multimedia/imagegallery/imagefeature2173.html`.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

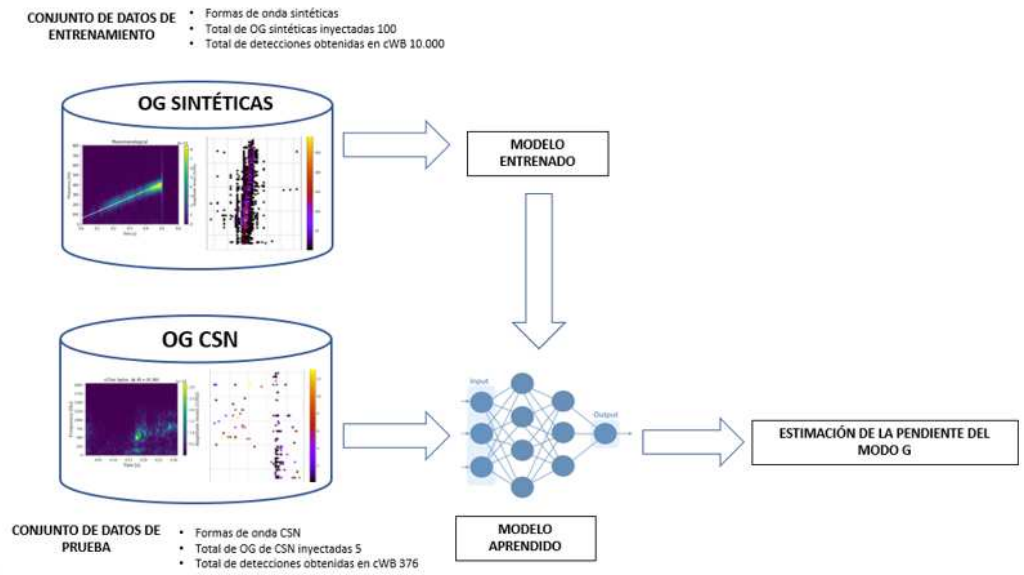


Figura 9. Ilustración general del mecanismo basado en IA para extraer la pendiente del modo g de una OG de CSN. En el bloque de imágenes a la derecha presenta: (arriba) el conjunto de datos de entrenamiento compuesto por 100 OG sintéticas, (abajo) el conjunto de prueba compuesto por 5 diferentes OG de CSN. El número total de detecciones en los datos de LIGO reportado por cWB corresponde a: 10,000 para el conjunto de entrenamiento y, 376 para el conjunto de prueba.

SOBRE LOS AUTORES



Alejandro Casallas-Lagos Es Ms en C en matemáticas de la Universidad de los Andes y Ms en C en Física de la Universidad Nacional de Colombia. Actualmente es candidato a doctor en física de la Universidad de Guadalajara - CUCEI. Su área de investigación está centrada en la caracterización de ondas gravitacionales producidas por colapso de supenova e inteligencia artificial.



Javier Mauricio Antelis Es doctor en ingeniería biomédica de la Universidad de Zaragoza España. Actualmente es profesor de la escuela de ingeniería del Tecnologico de Monterrey – campus Guadalajara. Es miembro de la colaboración LIGO y sus investigaciones están centradas en interfaces cerebro ordenador, teoría de la estimación en ondas gravitacionales por colapso de supernova y inferencia Bayesiana y simulaciones por método de Monte Carlo.



Claudia Moreno González Es doctora en física teórica CINVESTAV-IPN México. Actualmente es profesora de la Universidad de Guadalajara – CUCEI. Su campo de investigación se enfoca en relatividad general, teoría de campo y ondas gravitacionales. Es miembro de la colaboración LIGO y líder del grupo ondas gravitacionales de la UDG.



Michele Zanolin Es doctor en física de la Universidad Degli Studi de Parma Italia, posdoctorado de Massachussets Institute of Technology (MIT) y actualmente profesor de Embry Riddle Aeronautical University en Prescott - Arizona. Es líder mundial en el estudio de ondas gravitacionales producidas por colapso de supernova, miembro de la colaboración LIGO y sus investigaciones están centradas en relatividad general experimental.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Aplicación de la Inteligencia Artificial para el estudio de enfermedades neurodegenerativas

Alejandro Mateos-Moreno, Eddy Sánchez-DelaCruz y Luis Eduardo Maldonado-López

La importancia de la inteligencia artificial para el análisis de datos en ciencias de la salud

La Inteligencia Artificial (IA) es una técnica la cual se basa en el uso de algoritmos y bases de datos que analizan cierto porcentaje de confiabilidad. Durante los últimos años ha sido una herramienta fundamental para la obtención de diagnósticos más acertados y precisos, el tratamiento de enfermedades crónico-degenerativas y mantener control y bienestar del paciente, lo anterior se realiza mediante el análisis de bases de datos empleando herramientas informáticas con máquinas inteligentes.

En la actualidad existe una amplia gama de enfermedades neurodegenerativas, estas enfermedades afectan el sistema nervioso, ya sea a nivel central (cerebro y médula espinal) como periférico (nervios que convergen con los órganos del paciente). Dichos padecimientos, pueden afectar desde la conciencia, el pensamiento abstracto, la toma de decisiones, memoria, hasta el movimiento de las extremidades corporales. Algunos de ellos son los siguientes: enfermedad de Alzheimer, esclerosis lateral amiotrófica, ataxia, enfermedad de Huntington, demencia con cuerpos de Lewy, enfermedad de Parkinson y atrofia muscular espinal.

En el presente artículo mostraremos algunas de las formas mediante las cuales se puede emplear la IA para predecir, detectar y analizar la progresión de estos padecimientos.

Enfermedad de Alzheimer

La enfermedad de Alzheimer es el padecimiento mas común de demencia senil, la cual, tiene diferentes manifestaciones clínicas, principalmente la pérdida de memoria de forma degenerativa, el avance progresivo del Alzheimer hace evidente la estructura de un cerebro normal y uno enfermo (Fig 1). Su patología es multifactorial, es decir, que no hay solo un factor determinante para poder tratarla, estos van desde alteración en el funcionamiento de las neuronas encargadas de la memoria, de células conocidas como células gliales que le dan soporte a las anteriores, hasta factores genéticos que se pueden usar como marcadores en el diagnóstico.

Hasta la fecha no existe un tratamiento en particular para esta enfermedad que mitigue la demencia, sin embargo, pueden utilizarse mecanismos inteligentes (IA) para el tratamiento y predicción de estados tempranos de la neurodegeneración.

Tecnologías basadas en modelos estadísticos, así como técnicas de aprendizaje automático (machine learning) emplean algoritmos para el análisis de datos de imágenes de encefalogramas con el objetivo de diagnóstico y pronóstico de la patología. Lo anterior se realiza comúnmente a través de la observación de pérdida neuronal, disminución en los niveles de neurotransmisores como glutamato, análisis de fluido cerebroespinal, niveles de proteína beta-amieloide y de Tau hiperfosforilada.

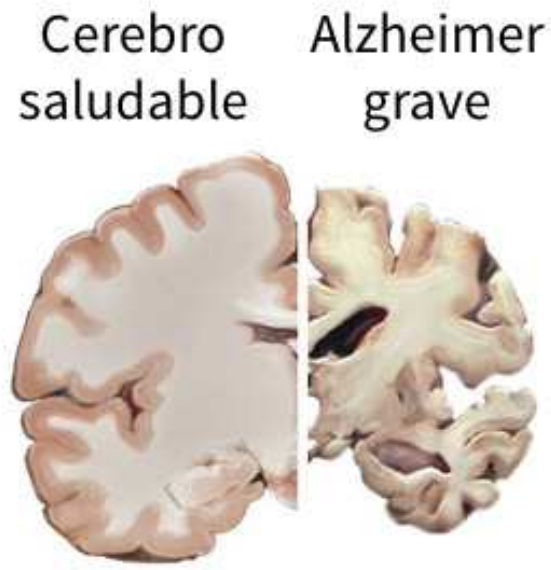


Figura 1. Diferencia entre un cerebro con y sin la enfermedad.

Esclerosis Lateral Amiotrófica

La Esclerosis Lateral Amiotrófica (ELA) es una enfermedad heterogénea altamente compleja la cuál incluye la degeneración de neuronas del cerebro y la médula espinal que controlan el movimiento de las extremidades corporales. Debido a que la esperanza de vida es bastante corta, se buscan alternativas al uso de medicamentos que predigan si un paciente va a desarrollar la enfermedad y el tiempo aproximado en el que vendrá la neurodegeneración. Los mecanismos de IA usan bases de datos provenientes de repositorios de ensayos clínicos para desarrollar algoritmos de aprendizaje automático. Un

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ejemplo de ello es el análisis de biomarcadores genéticos y moleculares asociados a la enfermedad. Estos hallazgos usando modelos matemáticos son muy importantes porque ayudarían en la formulación de medicamentos que eviten o disminuyan la degeneración neuronal, es decir usándolos como dianas terapéuticas y con ello la sintomatología. Así mismo existen otras técnicas basadas en IA, por ejemplo, "el eye tracking", el cual, consiste en un software desarrollado para seguir el movimiento ocular detectado a través de electrodos y una cámara, de tal manera que el paciente pueda comunicarse sin la necesidad del movimiento de del cuerpo, ya que este se pierde por completo durante la enfermedad. Estos software permiten una mejor calidad de vida del paciente, el cuál es bastante significativo (Fig 2).



Figura 2. Lesiones cerebrales de pacientes con ataxia mediante electroencefalografía.

Ataxia

La ataxia, es una enfermedad hereditaria que se caracteriza por alteraciones de la marcha (patrones de movimiento al caminar), ocasionando movimientos similares a un individuo completamente mareado. La patología principalmente consiste en alteraciones en la neurotransmisión de las neuronas Purkinje, que forman parte del cerebelo. Dado que este último se encarga de todos los movimientos de las extremidades desde los más delicados hasta los que requieren fuertes esfuerzos de la musculatura, es que este tipo de pacientes por lo regular permanecen en sillas especiales.

Recientemente, mediante el uso de sensores en distintas partes del cuerpo, se han formado bases de datos. El propósito de estas es permitir el diseño de algoritmos que diferencien la Ataxia de otras patologías que repercuten en el movimiento, como la enfermedad de Parkinson y Huntington.

Asimismo, existen estudios, los cuales miden frecuencias de neuronas del cerebelo que representan cambios en los niveles de iones como calcio y sodio. Estos trabajos analizan los datos mediante redes neuronales artificiales para clasificar sujetos sanos y con el padecimiento (Fig 3).

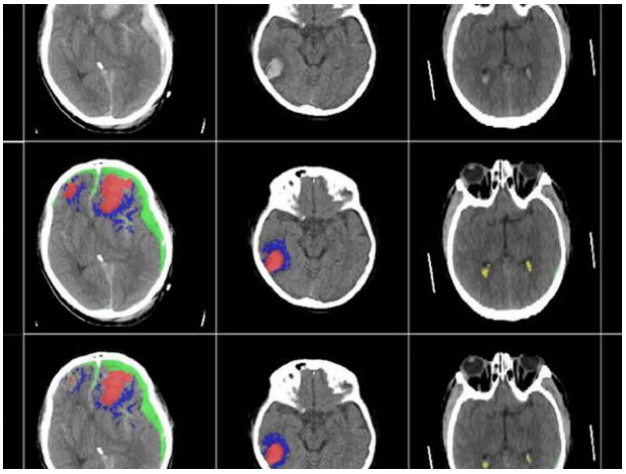


Figura 3. Uso de técnicas de imagenología para la creación de bases de datos.

Enfermedad de Huntington

La Enfermedad de Huntington (EH) es un padecimiento hereditario, el cuál se transmite por medio de un gen en específico. Dicha enfermedad ocasiona alteraciones motoras (movimientos involuntarios), y ya que existe muerte neuronal en diversas estructuras cerebrales, problemas de memoria y de toma de decisiones, llevando al paciente a condiciones bastante precarias. Al igual que la ataxia, para la clasificación de dicha enfermedad también se han empleado sensores y con dicha base de datos de las respuestas a través de patrones de movimiento se emplean algoritmos utilizando aprendizaje automático. De igual manera usando redes neuronales artificiales se ha logrado detectar la enfermedad por medio de marcadores de ADN, ya que existen dentro del mismo secuencias anormales que pueden ser fácilmente detectadas por técnicas moleculares y de forma similar por técnicas de encefalografía (Fig 4).

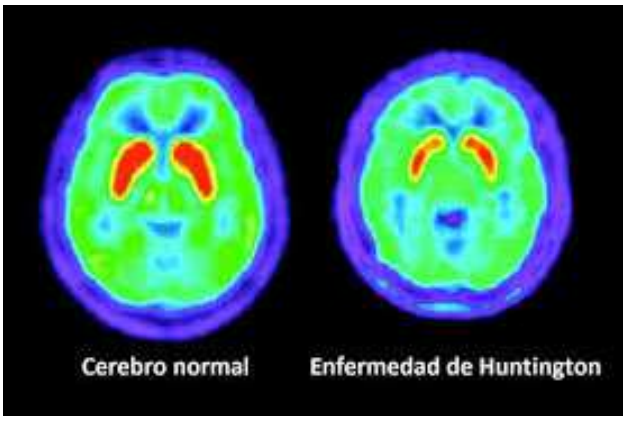


Figura 4. Encefalograma de cerebro normal y uno con la enfermedad.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Durante los últimos años el uso del aprendizaje automático ha beneficiado múltiples ámbitos de la salud, no solo de enfermedades neurodegenerativas.

Demencia por cuerpos de Lewy

A diferencia de la enfermedad de alzheimer, la demencia por cuerpos de lewy consiste en la formación de un exceso de depósitos en las terminaciones nerviosas de neuronas del sistema nervioso central de una proteína llamada alfa-sinucleína, dicha proteína representa el 1 % de las proteínas contenidas en las neuronas y que en bajas cantidades beneficia en el correcto funcionamiento del cerebro) (Fig 5), la cual cu acumulación daña las conexiones de las mismas, llevando a síntomas como pérdida de la memoria, alteraciones en el movimiento y en el pensamiento abstracto, es decir en la toma de decisiones. Este tipo de demencia es considerado el segundo padecimiento más frecuente después del alzheimer, es por ello que para los médicos resulta complicado hacer un diagnóstico certero si no se hacen técnicas invasivas.

Dado lo anterior, cobra relevancia el uso de algoritmos especializados en la diferenciación de esta patología. Entre los que se han empleado con éxito están las máquinas de vectores de soporte, los cuales son un conjunto de algoritmos de aprendizaje automático que mediante el análisis de datos obtenidos por tomografías (imágenes del funcionamiento cerebral) pueden determinar con mayor precisión si la demencia es por alzheimer o demencia por cuerpos de Lewy.

En el presente existen muy pocos estudios acerca del uso de técnicas de IA que analicen esta enfermedad, por lo que es un área importante a cubrir en la actualidad.

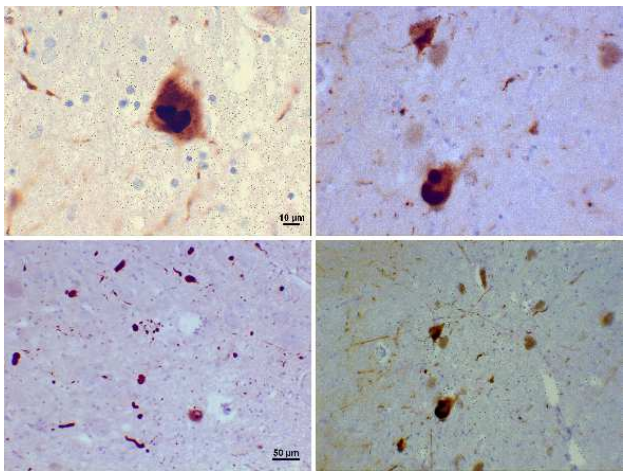


Figura 5. Cuerpos de Lewy en neuronas del sistema nervioso central.

Enfermedad de Parkinson

Por otro lado, la enfermedad de Parkinson es un padecimiento originado por la disminución de un neurotransmisor neuronal conocido como dopamina. La dopamina es vital para la comunicación neuronal, es decir que una neurona del cerebro se comunique con otras de diferentes estructuras cerebrales, originando disfunción en los patrones del movimiento como temblores, movimientos lentos, movimientos involuntarios, el habla y disfunción para retener objetos en las manos.

Dándole seguimiento a los estudios de IA, hay algunos que combinan aprendizaje automático con sensores (ubicados en diferentes regiones del cuerpo) para formar bases de datos, y contribuir en la clasificación de otras enfermedades que alteran el movimiento (como las mencionadas anteriormente). Por otro lado, también se tienen bases de datos que mediante creación de algoritmos, se ha logrado la creación de aplicaciones en teléfonos inteligentes (smartphones) que miden la frecuencia de la voz a través del micrófono, el golpeteo de los dedos en la pantalla (ya que la característica principal de la enfermedad son los temblores de las extremidades principalmente las manos), la marcha y el equilibrio gracias al acelerómetro del dispositivo (Fig 6).

Respecto al uso de estas técnicas de IA, hay más estudios que emplean diferentes sensores para poder diferenciar y darle seguimiento a la progresión de la enfermedad, sin embargo, se busca la técnica más adecuada para el paciente y que este no represente un gasto tan alto.

Atrofia muscular espinal

En última instancia en el seguimiento de enfermedades neurodegenerativas y algunas de las maneras en las que se ha empleado la IA para analizar bases de datos, tenemos a la atrofia muscular espinal. Su patología es muy similar a las anteriores, ya que hay muerte neuronal en regiones del cerebro y de la médula espinal que controlan el movimiento del cuerpo. La diferencia es que en ella existen 4 tipos que varían en el grado de gravedad y que puede presentarse a partir del nacimiento del individuo o hasta los 21 años aproximadamente (siendo este último de menor gravedad o tipo 4)). Algunos de los síntomas de los diferentes tipos engloban desde la dificultad de deglución hasta la incapacidad para caminar. Las técnicas de aprendizaje automático han logrado diferenciar los distintos tipos de atrofia muscular, ello por medio de

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

técnicas como la miografía por impedancia eléctrica, que mide el grado de lesión del tejido muscular, y de esta manera se obtienen grandes cantidades de información para poder aplicar algoritmos con altos porcentajes de exactitud. El objetivo de mostrar toda esta información de las diferentes enfermedades es exhibir la relevancia del empleo de la tecnología mediante el uso de la IA, por lo que se esperan grandes avances en este campo.

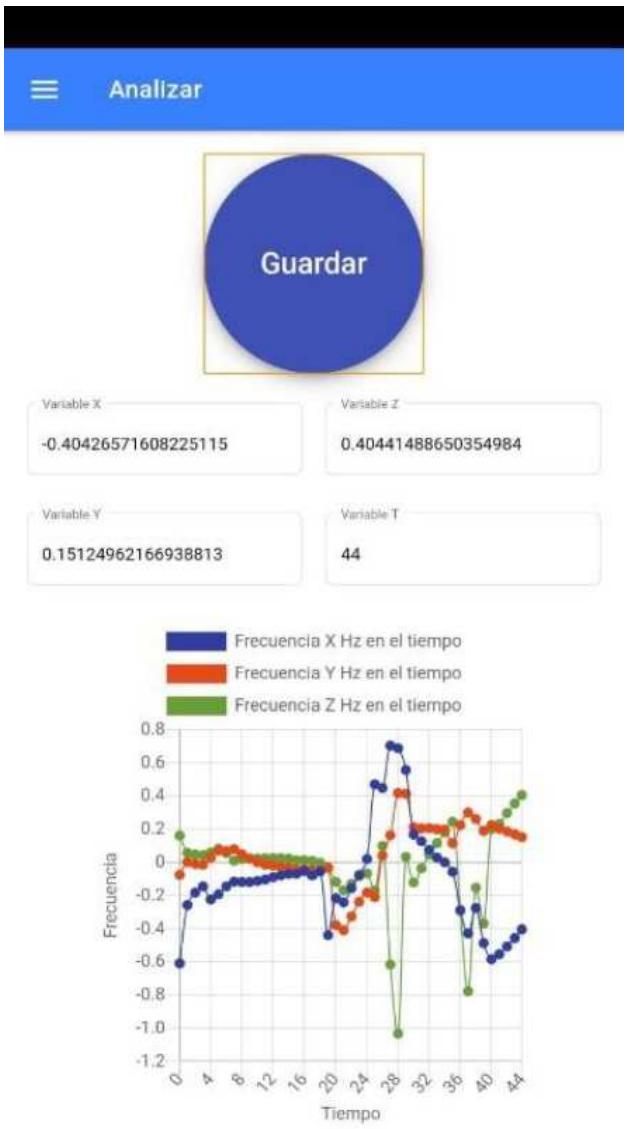


Figura 6. Ejemplo de la interfaz de un programa para el análisis de patrones del movimiento conocido como Parkinson Analyzer.

Conclusiones

De esta manera y a pesar de los grandes avances de la Inteligencia Artificial en el área de ciencias de la salud, particularmente en las neurociencias, aún existe un gran campo en el análisis de bases de datos de diferentes marcadores de otras enfermedades crónico-degenerativas como la diabetes, hipertensión, artritis, asma, cáncer, etc. Esto significa que el médico podría llevar a cabo diagnósticos más acertados, un mejor tratamiento y un mejor seguimiento de alguna patología.

Finalmente, durante los últimos 5 años y como se observa en las referencias ha existido un crecimiento también en la creación de algoritmos basados en aprendizaje automático, y también cada vez más empresas se están dedicando a que la eficiencia y exactitud (un mayor porcentaje de estos) de los mismos puedan ayudar en la detección, predicción y avance de ciertos padecimientos, todo ello hablando solamente en el área médica.✳

REFERENCIAS

1. Abbasi, S., Abbasi, A., Sarbaz, Y. y Ebrahimi, A. (2014). Analysis spectrum of normal and ataxia Purkinje cell output and classification using artificial neural network. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 26(4), E14.

2. Alfonso-Perez, G. y Caballero-Villarraso, J. (2022). Neural Network Aided Detection of Huntington Disease. *Journal of clinical medicine*, 11(8), 2110.

3. Chang, C. H., Lin, C. H. y Lane, H. Y. (2021). Machine learning and novel biomarkers for the diagnosis of Alzheimer's disease. *International Journal of Molecular Sciences*, 22(5), 2761.

4. Fuentes-Ramos, M., Sánchez-DelaCruz, E., Meza-Ruiz, I. V. y Loeza-Mejía, C. I. (2022). Neurodegenerative diseases categorization by applying the automatic model selection and hyperparameter optimization method. *Journal of Intelligent & Fuzzy Systems*, 1-9.

5. García, D., Flores, V. , López, J., Jiménez, E. y Acurio, E. (2019). Avances de la inteligencia artificial en salud. *Dominio de las Ciencias*, 5(3), 603-613.

6. Katako, A., Shelton, P., Goertzen, A. L., Levin, D., Bybel, B., Aljuaid, M. y Ko, J. H. (2018). Machine learning identified an Alzheimer's disease-related FDG-PET pattern which is also expressed in Lewy body dementia and Parkinson's disease dementia. *Scientific reports*, 8(1), 1-13.

7. Liu, X., Chen, K., Wu, T., Weidman, D., Lure, F. y Li, J. (2018). Use of multimodality imaging and artificial intelligence for diagnosis and prognosis of early stages of Alzheimer's disease. *Translational Research*, 194, 56-67.

8. McCombe, N., Joshi, A., Finn, D. P., McClean, P. L., Roberts, G., O'Brien, J. T. y Wong-Lin, K. (2022). Distinguishing Lewy Body Dementia from Alzheimer's Disease using Machine Learning on Heterogeneous Data: A Feasibility Study. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4929–4933.

9. Mohan, A., Sun, Z., Ghosh, S., Li, Y., Sathe, S., Hu, J. y Sampaio, C. (2022). A Machine-Learning Derived Huntington's Disease Progression Model: Insights for Clinical Trial Design. *Movement Disorders*, 37(3), 553-562.

10. Pancotti, C., Birolo, G., Rollo, C., Sanavia, T., Di Camillo, B., Manera, U. y Fariselli, P. (2022). Deep learning methods to predict amyotrophic lateral sclerosis disease progression. *Scientific Reports*, 12(1), 1-10.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

SOBRE LOS AUTORES



Dr. Alejandro Mateos-Moreno se encuentra como posdoctorante en el Instituto Tecnológico Superior de Misantla, Veracruz, México. Es Dr, en Investigaciones Cerebrales por la Universidad Veracruzana, Xalapa, Veracruz, México. Químico Farmacéutico Biólogo por la Universidad Veracruzana, Xalapa, Veracruz, México. Actualmente se encuentra trabajando en el análisis de datos para el desarrollo de algoritmos para predecir enfermedades neurodegenerativas.



Eddy Sánchez-DelaCruz es Doctor en Sistemas Computacionales por la Universidad del Sur, Maestro en Sistemas Computacionales y Licenciado en Informática Administrativa por la Universidad Autónoma de Tabasco. El Dr. Sánchez-DelaCruz es miembro del Sistema Nacional de Investigadores, nivel I (destacando que obtuvo el nivel uno en su primera evaluación), es miembro de la Sociedad Mexicana de Inteligencia Artificial y miembro de la RedICA-CONACyT, es Perfil deseable PRODEP y es fundador y líder del Cuerpo Académico Inteligencia Artificial. Sus intereses de investigación son el aprendizaje automático y la ciencia de datos con un enfoque multidisciplinario.



Luis Eduardo Maldonado López es Ingeniero Químico por la Universidad Juárez Autónoma de Tabasco y tiene maestría y doctorado en Nanociencias y Materiales por el Instituto Potosino de Investigación Científica y Tecnológica. Actualmente realiza una estancia posdoctoral en el Cuerpo Académico de Inteligencia Artificial del TecNM, Campus Misantla.



Imagen tomada de:
https://www.news-medical.net/news/20200128/New-AI-tool-predicts-disease-progression-and-evolution-in-neurodegenerative-disorders.aspx.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

La inteligencia artificial como apoyo en la identificación de especies de plantas

Juan Augusto Campos-Leal, Inés F. Vega-López, Gerardo Beltrán-Gutiérrez, José Ramón López-Arellano y Arturo Yee-Rendón

Introducción

En los últimos años, la tecnología ha tomado un rol importante para la realización de las actividades de un individuo en su día a día. Usamos dispositivos que nos permiten llevar a cabo tareas simples y complejas o bien, tareas repetitivas y específicas. En busca de resolver tareas de forma automática y sencilla, hay una constante búsqueda por generar las herramientas que nos permitan realizarlo. La tecnología ha permitido que diversas áreas de la ciencia utilicen sus conocimientos para obtener estas herramientas inteligentes. La inteligencia artificial (IA) es el área de las ciencias de la computación que permite generarlas a través del procesamiento de datos. Dichas herramientas han sido capaces de resolver tareas complejas para el ser humano, de forma eficiente y sencilla.

Una de las tareas en la que la inteligencia artificial ha apoyado es la identificación de especies de plantas. Este tipo de tarea es llevada a cabo por expertos bajo un proceso meticuloso y muy bien cuidado. En este proceso se identifican las características de un ejemplar para determinar a que grupo pertenece, estos grupos los determinan los expertos mediante características similares y puede ser a diferentes niveles, como familia, género o especie [1].

La identificación de especies de plantas a través de imágenes digitales no es un problema trivial para la IA, representa un reto extraordinario debido a las similitudes inter-especies y grandes variaciones intra-especie.

La idea detrás de este trabajo es mostrar que el uso de herramientas tecnológicas basadas en la IA, es una estrategia viable para auxiliar a expertos en el estudio y cuantificación de la biodiversidad de una región. El trabajo de registro y recolección de especies de plantas que se ha hecho durante años puede ser aprovechado por estas herramientas tecnológicas para entrenar modelos de identificación automática con altas tasas de aciertos.

Antecedentes

Durante muchos años, la tarea de identificación de especies de plantas se llevó a cabo de manera manual por taxónomos expertos, bajo proceso de laboratorio y observación. Este es un proceso totalmente biológico que no tenía nada que ver con la computación o con la inteligencia artificial. Sin embargo, Younis y sus colaboradores [2] mencionan que desde hace aproximadamente dos décadas fue propuesta la idea de generar herramientas inteligentes para la identificación de especies de plantas.

Las primeras herramientas inteligentes para este tipo de tarea se basaban en un proceso de aprendizaje automático, estas herramientas partían del aprendizaje supervisado, es decir, el experto había identificado previamente el ejemplar a caracterizar. Por lo tanto, conocía el dato de entrada y la respuesta que la herramienta debería de proporcionar. La información de entrada eran las características del ejemplar y estas características se tomaban directamente del ejemplar o de imágenes digitales. En el caso de la extracción de características de las imágenes digitales se usaban algoritmos diseñados por un humano experto para el proceso de extracción. Sin embargo, este proceso de extracción estaba limitado por la capacidad humana. Estas técnicas son conocidas como técnicas convencionales de aprendizaje de máquina.

Nilsback y sus colaboradores [3] presentaron trabajos para la identificación de especies de plantas que se basaban en técnicas convencionales de aprendizaje de máquina. Además, algunos conjuntos de datos de especies de plantas fueron propuestos para auxiliar a la creación de este tipo de herramientas inteligentes. Algunos contenían imágenes de hojas como Swedish Leaf y Flavia, con imágenes de 15 y 32 especies, respectivamente. También algunos otros conjuntos de imágenes de flores como Oxford 17 Flower y Oxford 102 Flower. El rendimiento de cada herramienta propuesta era probado en alguno de estos conjuntos para tener un punto de referencia en cuanto a los resultados.

Hasta el año 2012 los resultados no eran muy alentadores y la aplicación de estas herramientas en el mundo real aún no era viable. Sin embargo, con la aparición de

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

las redes neuronales convolucionales [4] (CNNs, por sus siglas en inglés), un tipo de técnica basada en el aprendizaje profundo, las tareas de identificación de especies de plantas usando imágenes digitales se volvieron viables. La principal diferencia entre el aprendizaje profundo y los métodos convencionales de aprendizaje de máquina es que el primero no tiene un proceso de extracción de características limitado por la capacidad humana.

En el aprendizaje profundo el proceso de extracción de características en una imagen digital es automático.

El proceso va mejorando en la extracción de características mediante el algoritmo de aprendizaje. Por lo tanto, una CNN aprende aquellas características que más le sirven para la identificación de especies de plantas. Sin embargo, para que el proceso de extracción de características pueda ser lo suficientemente bueno es necesario que durante el proceso de aprendizaje el modelo procese un gran número de imágenes. Cuando el número de imágenes en un conjunto de datos es limitado, se han aplicado técnicas de aprendizaje por transferencia para tratar de mejorar la tasa de aciertos de los modelos.

El año 2012 marcó un antes y un después con respecto a la identificación de imágenes. Fue la primera vez que las CNNs superaron a las técnicas convencionales de aprendizaje de máquina [4]. A partir de ese año han sido grandes los esfuerzos para la generación de herramientas inteligentes. En el ámbito de la identificación de especies de plantas ha quedado demostrado en trabajos como el de Sulc y colaboradores [5], donde a través de su propuesta basada en CNNs obtuvo el primer lugar en la edición 2018 del reto PlantCLEF, uno de los retos mas importantes a nivel mundial en este tipo de tarea.

En 2020, Little y sus colaboradores [6] presentaron un reto para fomentar el desarrollo de algoritmos para la identificación de especies de plantas de herbario. La competencia atrajo a 22 equipos a nivel mundial que ingresaron 254 modelos CNNs entrenados para la identificación de ejemplares de la familia *Melastomataceae*. Los cuatro mejores con una tasa de aciertos mayor a 88 %. En 2022, Campos-Leal y sus colaboradores [7] presentaron una propuesta basada en CNN para la identificación de 120 especies de plantas con flores nativas de México. En su trabajo presentan un modelo CNN simplificado que obtuvo un 91.92 % en tasa de aciertos y fue comparado con otros modelos basados en CNN. En 2023, Vega-López y colaboradores [8] realizaron un trabajo para la identificación de 202 especies de plantas de la flora mexicana usando cuatro órganos distintivos para la identificación, tallo, hoja, fruto y flor. Los autores usaron estrategias de aprendizaje por transferencia con lo que obtuvieron 86.97 % en tasa de aciertos.

Redes neuronales convolucionales

Las CNNs son un tipo de técnica del aprendizaje profundo capaces de extraer por sí mismas las características visuales que ayudan a discriminar las clases que se le presentan. Esto lo realizan mediante el procesamiento de imágenes a través de filtros digitales, cuyos coeficientes se infieren durante un proceso de entrenamiento. Este tipo de redes se conforman por tres tipos de capas, convolución, agrupamiento y clasificación. Una arquitectura de CNNs es formada por un apilamiento de este tipo de capas. Generalmente, los primeros dos tipos de capas forman un bloque que es utilizado para la extracción de características. La salida de este bloque es un vector de entrada a un clasificador. Este clasificador es compuesto por capas densas. Las arquitecturas reportadas en la literatura científica, tales como VGG-16, Inceptionv3 y ResNet50, hacen uso de diferentes configuraciones de estos bloques. La Figura 1 muestra la estructura de la arquitectura ResNet50, que está formada por 50 capas agrupadas en bloques que se encargan de la extracción automática de características y clasificación. Esta arquitectura tiene 25.6 millones de parámetros.

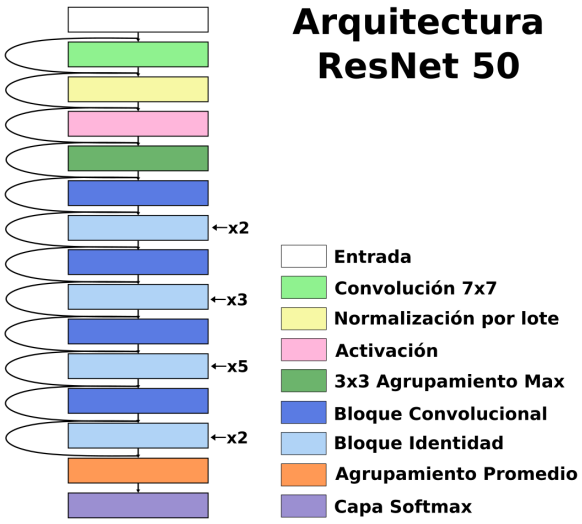


Figura 1. Diagrama de la arquitectura ResNet 50.

Las CNNs deben su nombre a su principal operación, la convolución. Una convolución es una operación matemática realizada por un filtro de tamaño $n \times n$ sobre

una región de una matriz de entrada. Cada elemento del filtro es multiplicado por el elemento correspondiente de la matriz. El filtro es desplazado por cada espacio disponible en la matriz, de izquierda a derecha y de arriba a abajo, a este proceso se le llama convolución. El número de celdas que avanza es determinado por un valor paramétrico llamado paso. Cuando el filtro se ha desplazado por toda la matriz de entrada se obtiene una matriz de activación, que será la entrada a procesar por una siguiente capa.

Las capas de agrupación tienen como propósito reducir el tamaño de las dimensiones $ancho \times alto$ de las matrices de activación. Similar a las capas convolucionales, estas capas desplazan un filtro de tamaño $n \times n$ (generalmente de 2×2) por una matriz de entrada para aplicar la agrupación por regiones. Los tipos de agrupación son: por valor mínimo, máximo y promedio, tomando el valor más representativo según sea el caso. Cuando el filtro se ha desplazado por toda la matriz de entrada es obtenida una matriz que será la entrada a procesar por la siguiente capa. Las capas densas son las capas completamente conectadas tradicionales de las redes neuronales artificiales. Estas capas se encuentran normalmente al final de las CNNs y se utilizan para realizar el proceso de clasificación. Las capas densas normalmente reciben como entrada el vector de características obtenido durante el proceso de extracción.

Entrenamiento

El entrenamiento de una arquitectura de CNNs es un proceso donde el conjunto de datos es tratado iterativamente para intentar descubrir características que permitan realizar la identificación. Durante este proceso se usa un algoritmo de aprendizaje llamado descenso de gradiente estocástico (SGD, por sus siglas en inglés). De esta manera, el problema de identificación es tratado como un problema de optimización que busca minimizar el error. Esto se hace mediante el ajuste de valores de los parámetros de la arquitectura. Durante el proceso de aprendizaje, la arquitectura procesa ejemplos de las categorías a identificar y ajusta los valores de los parámetros tratando de minimizar el error. Esto se hace mediante el algoritmo de retropropagación.

En las CNNs los valores de los parámetros se dividen en dos partes, filtros convolucionales y clasificador. Los primeros son los encargados de encontrar los patrones en las imágenes (características) y los segundos son los encargados de la clasificación. Para medir el comportamiento de la arquitectura durante el proceso de aprendizaje se mide la tasa de aciertos con respecto a las respuestas esperadas. Si la respuesta es la esperada contabilizamos un acierto, si no, entonces se utiliza una función de pérdida para medir que tan cerca o que tan lejos está la respuesta con respecto a la respuesta esperada. Para llevar a cabo el proceso de entrenamiento de las arquitecturas, el usuario proporciona un conjunto de valores que determinan el comportamiento de los algoritmos. Por ejemplo, controlan que tanto se modifica el valor de los parámetros, cuantas imágenes se procesan y cuantas iteraciones (o épocas) se observara por completo el conjunto de entrenamiento. El conjunto de estos valores son conocidos como hiperparámetros.

Durante el proceso de entrenamiento el conjunto de datos es utilizado y es dividido en tres subconjuntos, entrenamiento, validación y prueba.

El primer subconjunto contiene las imágenes que se procesarán y de las que el modelo aprenderá características, este conjunto debe ser el de mayor tamaño. El de validación es utilizado para medir el rendimiento del modelo cada vez que finaliza una iteración (o época), esto nos da una idea del comportamiento del modelo. El subconjunto de prueba es utilizado al final del entrenamiento, este lo utilizamos para obtener el rendimiento del modelo con un subconjunto de imágenes que no son vistas ni durante el entrenamiento ni la validación.

Aprendizaje por transferencia

El aprendizaje por transferencia (AT) es una técnica utilizada durante el entrenamiento para mejorar la tasa de aciertos de un modelo de clasificación. Esta técnica consiste en utilizar como punto de partida los valores de los parámetros de un modelo entrenando en un conjunto de datos a gran escala (generalmente con cientos de miles de imágenes). Esta técnica se basa en el supuesto de que un modelo entrenado en una tarea A puede ser entrenado de nuevo para tratar una tarea B, ajustando solo un poco los valores de los parámetros durante el entrenamiento. Esto es debido a que el modelo A ya tiene la capacidad para extraer características y los valores de los parámetros no necesitan ser modificados por completo, solo necesitan ajustarse a las nuevas características. El punto de partida más utilizado en el aprendizaje profundo son los valores de los parámetros de modelos entrenados en el conjunto de datos de ImageNet 2012, esto es debido a que este conjunto es utilizado como punto de referencia para medir la tasa de aciertos de los modelos de clasificación.

Conjunto de datos

En este documento, para ilustrar como la IA puede ser aplicada a la identificación de especies de plantas utili-

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

zamos dos conjuntos de datos como experimento ilustrativo. Para este trabajo elegimos Oxford 102 Flower [3] como un conjunto de datos de especies de plantas en su hábitat natural y un subconjunto de PlantCLEF2020 [9] como un conjunto de especies de plantas en un ambiente controlado. El conjunto Oxford 102 Flower, es un conjunto que contiene 8,189 imágenes de flores que corresponden a 102 especies de plantas que pueden ser encontradas en Reino Unido. En la Figura 2 mostramos algunas imágenes de ejemplo en el conjunto. Estas imágenes fueron tomadas en condiciones de campo por lo cual pueden existir variaciones en la luz, fondo, perspectiva y escala, por mencionar algunas. Este conjunto contiene entre 40 y 258 imágenes por especie.



Figura 2. Imágenes del conjunto de Oxford 102 Flower.

El otro conjunto de datos utilizado es un subconjunto de PlantCLEF 2020, el cual consiste en 10 especies de plantas de herbario con mayor representación. Para cada especie se tomó una muestra de 50 imágenes. Cada imagen tiene dimensión rectangular con orientación vertical. La Figura 3 muestra imágenes de ejemplo para cada especie en el conjunto.



Figura 3. Imágenes del conjunto de PlantCLEF 2020.

Las CNNs reciben como entrada una imagen digital cuadrada de un tamaño específico. Cuando la imagen no cumple con estás especificaciones, sufre de algunas alteraciones que pueden llevar a distorsionar el objeto a identificar que está contenido en la imagen, o hay un redimensionamiento modificando la información de la imagen. Estas alteraciones pueden tener un impacto negativo en la tasa de aciertos del modelo, debido a que puede cambiar la naturaleza de los objetos a identificar. El preprocesamiento de las imágenes puede evitar las alteraciones en una imagen. De esta manera podemos enfocar la atención del modelo y evitar distorsiones o alteraciones. El preprocesamiento consistió en la realización de dos recortes cuadrados por cada imagen. La Figura 4 muestra un ejemplo de este proceso. Cada recorte corresponde a una región de interés en la imagen. Para generar el primer recorte se tomó como referencia de inicio la esquina superior izquierda de la imagen sin procesar. Para el segundo recorte se tomó como referencia el centro de la imagen. De esta manera se obtuvieron dos regiones de interés cuadradas por cada imagen. En total formamos un conjunto de 1,000 imágenes con 100 imágenes para cada especie.

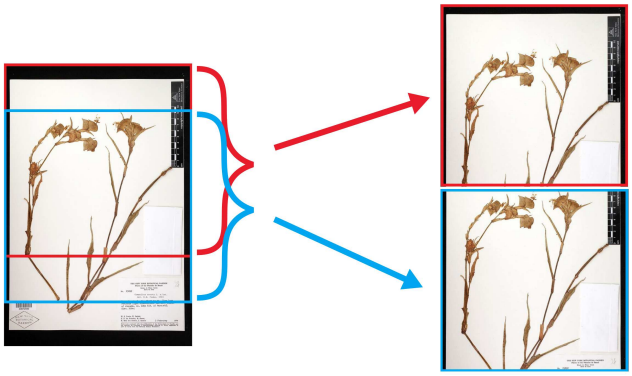


Figura 4. Imagen sin procesar y sus dos recortes.

Resultados

Utilizamos la arquitectura ResNet50 [10] para llevar a cabo nuestros experimentos. Entrenamos modelos de clasificación de la siguiente manera: modelos cuyos valores iniciales de sus parámetros fueron asignados aleatoriamente, conocido como entrenamiento desde cero, y modelos que fueron generados utilizando aprendizaje por trasferencia del reto ImageNet 2012, es decir, modelos cuyos valores iniciales de parámetros fueron valores obtenidos por modelos previamente entrenados. Además, utilizamos una técnica de validación cruzada de 10 partes para tener una mayor certeza estadística en los resultados. Los valores de los hiperparámetros que utilizamos están descritos en la Tabla 1.

Tabla 1.Valores de los hiperparámetros.

Parámetros	Desde cero	Con AT
Optimizador	SGD	SGD
Tasa de aprendizaje	0.01	0.0001
Decaimiento	0.005	0
Impulso	0.9	0.9
Tamaño de lote	16	16
Número de épocas	200	40

En nuestros experimentos reportamos el desempeño de los modelos a través de la tasa de aciertos, conocido en inglés como *Top-K accuracy*. *Top-K* se refiere al porcentaje de respuestas correctas en el conjunto de las K respuestas mejor clasificadas por el modelo. Por ejemplo,

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

para un indicador *Top-1*, un acierto se contabiliza cuando la respuesta del modelo es igual a la respuesta esperada. Los resultados se muestran en la Tabla 2. Los renglones uno y dos representan a los modelos entrenados usando el conjunto de Oxford 102 Flower. El resto corresponde a los modelos usando el subconjunto de PlantCLEF 2020. De acuerdo a los resultados de la tabla, observamos que los modelos que utilizan estrategias de aprendizaje por transferencia obtiene mejores desempeños que aquellos modelos entrenados desde cero.

Tabla 2. Modelos ResNet50.

Modelo	Dataset	Top-1
ResNet50	Oxford 102 Flower	46.60 %
ResNet50 + AT	Oxford 102 Flower	94.87 %
ResNet50	PlantCLEF 2020 (10)	84.10 %
ResNet50 + AT	PlantCLEF 2020 (10)	93.50 %

En las Figuras 5 y 6, se presentan las curvas de aprendizaje de los modelos entrenados usando el conjunto de datos de Oxford 102 Flower. Las curvas de aprendizaje permiten conocer el comportamiento de los modelos durante el entrenamiento, para ello se utiliza el subconjunto de validación que nos permite conocer cómo está generalizando (aprendiendo) el modelo.

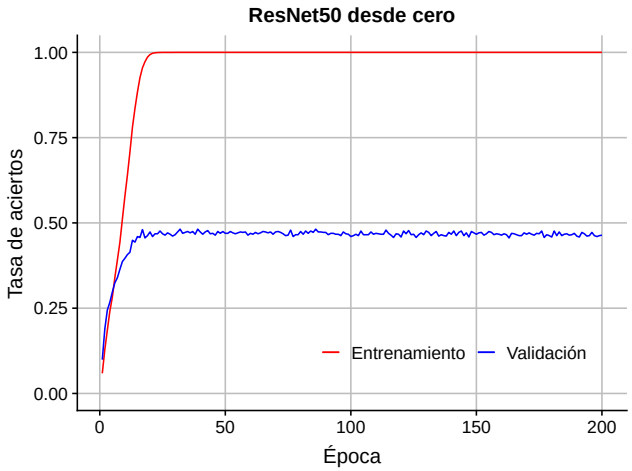


Figura 5. Modelo entrenado en Oxford 102 Flower.

La Figura 5 muestra la curva de aprendizaje de un modelo ResNet50 entrenado desde cero. El comportamiento presenta un amplio margen entre la tasa de aciertos del subconjunto de entrenamiento (línea de color rojo) y el subconjunto de validación (línea color azul). La diferencia entre las curvas es aproximadamente de 50 puntos porcentuales, lo que indica que el modelo obtiene buenos resultados durante el entrenamiento, pero no durante la validación, y por tanto, tampoco lo hará cuando se utilice con el conjunto de prueba. La Figura 6 muestra la curva de aprendizaje del modelo ResNet50 utilizando aprendizaje por transferencia. Observamos que el margen entre las curvas de aprendizaje es mucho menor con respecto al margen del modelo entrenado desde cero. El análisis de las curvas de aprendizaje entre el modelo entrenado desde cero y el que usa aprendizaje por transferencia, nos permite validar y explicar los resultados de los modelos presentados en la Tabla 2.

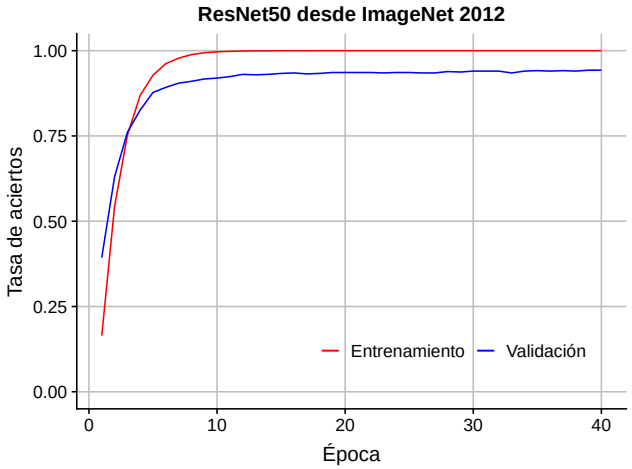


Figura 6. Modelo utilizando AT.

Conclusiones

La identificación de especies de plantas a través de herramientas tecnológicas basadas en IA ha mostrado buenos avances durante de la última década. Sin embargo, una considerable cantidad de datos de entrenamiento es requerida para obtener modelos de identificación automáticos con altas tasas de aciertos. En los últimos cinco años, los expertos en IA se han enfocado en utilizar las imágenes de especies de plantas de herbarios, pues la información recolectada a través de décadas está siendo digitalizada y es aprovechada para auxiliar a los expertos taxónomos en las tareas de identificación. En este trabajo llevamos a cabo experimentos ilustrativos para mostrar la utilidad de los modelos basados en IA utilizando dos conjuntos de datos públicos de imágenes de especies de plantas, imágenes de plantas en su hábitat natural e imágenes de plantas en herbario. Los resultados mostraron que los modelos de clasificación basados en IA son capaces de obtener altas tasas de aciertos en la identificación de especies de plantas mediante la utilización de técnicas de aprendizaje por transferencia. ✷

Agradecimientos

Los autores desean agradecer al Consejo Mexicano de Ciencia y Tecnología (CONACYT) por la beca otorgada al primer autor, así como por el financiamiento para el proyecto 291772 del fondo CONACYT-INEGI. De la

misma manera, los autores agradecen el apoyo económico brindado por la Universidad Autónoma de Sinaloa.

REFERENCIAS

1. Sulc, M. y J., Matas. (2014). Texture-Based Leaf Identification. En *Proc. Computer Vision-ECCV 2014 Workshops*, 185-200.
2. Younis, S., Weiland, C., et al. (2018). Taxon and Trait Recognition from Digitized Herbarium Specimens using Deep Convolutional Neural Networks. *Botany Letters*, 165(3-4), 377-383.
3. Nilsback, M. E. y Zisserman, A. (2008). Automated Flower Classification over a Large Number of Classes. En *Proc. 2008 Sixth Indian Conference on Computer Vision, Graphics and Image Processing*, 722-729.
4. Simonyan, K. y Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. En *Proc. of 3rd International Conference on Learning Representations*.
5. Sulc, M., Picek, L. y Matas, J. (2018). Plant Recognition by Inception Networks with Test-time Class Prior Estimation. En *Proc. of the Conference and Labs of the Evaluation Forum*, 2125.
6. Little, D. P., Tulig, M., et al. (2020). An Algorithm Competition for Automatic Species Identification from Herbarium Specimens. *Applications in Plant Sciences*, 8(6), e11365.
7. Campos-Leal, J. A., Yee-Rendón, A. y Vega-López, I. F. (2022). Simplifying VGG-16 for Plant Species Identification. *IEEE Latin America Transactions*, 20(11), 2330-2338.
8. Vega-López, I. F., Vega-Aviña, R., Delgado-Vargas, F., et al. (2023). Identificación de especies de plantas de la flora mexicana utilizando aprendizaje por transferencia a través de Inception-v4. *RDE Revista Internacional de Estadística y Geografía*, 14(1), 22-37.
9. Goëau, H., Bonnet, P. y Joly, A. (2020). Overview of Lifeclef Plant Identification Task 2020. En *Proc. of the CLEF 2020-Conference and Labs of the Evaluation Forum*, 2696(140).
10. He, K., Zhang, X., et al. (2016). Deep Residual Learning for Image Recognition. En *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 770-778.

SOBRE LOS AUTORES



Juan Augusto Campos Leal recibió su licenciatura en Sistemas Computacionales de la Universidad Autónoma de Occidente en 2008 y su maestría en Ciencias de la Computación en 2018 por la Universidad Autónoma de Sinaloa. Actualmente es estudiante de Doctorado en Ciencias de la Información en la misma universidad. Sus intereses de investigación actuales incluyen el aprendizaje profundo y las técnicas de aprendizaje automático.



Inés Fernando Vega-López es profesor de Ciencias de la Computación en la Universidad Autónoma de Sinaloa (Culiacán, México) desde 2004. Recibió su Doctorado en Ciencias de la Computación por la Universidad de Arizona en 2004. El Dr. Vega-López es miembro del Sistema Nacional de Investigadores Nivel I. Sus intereses de investigación actuales incluyen sistemas de bases de datos de alto rendimiento, análisis de datos a gran escala y aprendizaje automático.



Gerardo Beltrán-Gutiérrez es profesor de Ciencias de la Computación en la Universidad Autónoma de Sinaloa (Culiacán, México). Obtuvo su maestría en Ciencias de la Computación por la Universidad Autónoma de Sinaloa. Sus líneas de interés incluyen bases de datos, ciencia de datos y programación paralela y distribuida.



José Ramón López-Arellano es ingeniero en Sistemas de Información por el Instituto Tecnológico y de Estudios Superiores de Monterrey Campus Monterrey (ITESM). Obtuvo su maestría por EGADE Business School Monterrey y es Doctor en Gestión de las Organizaciones por las Universidades Autónoma de Nayarit, Autónoma de Sinaloa y Juárez del Estado de Durango. Actualmente, es Coordinador del Parque Científico Tecnológico (PCT) de la Universidad Autónoma de Sinaloa (UAS).



Arturo Yee-Rendón es profesor de Ciencias de la Computación en la Universidad Autónoma de Sinaloa. Obtuvo su maestría y doctorado en Ciencias de la Computación por el CINVESTAV-IPN en 2010 y 2015, respectivamente. El Dr. Yee-Rendón es miembro del Sistema Nacional de Investigadores Nivel I. Sus intereses de investigación actuales incluyen reconocimiento de patrones, técnicas de aprendizaje profundo, teoría de juegos y algoritmos de optimización (algoritmos genéticos).

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Anotación automática de imágenes de plantas utilizando Inteligencia Artificial

Eduardo Díaz Gaxiola, Gerardo Beltran-Gutierrez, Inés F. Vega-López y Arturo Yee Rendón

A lo largo de los años, se ha buscado automatizar diferentes tareas y emular el conocimiento humano mediante computadoras y equipos digitales. Es por esto que la Inteligencia Artificial (IA) ha tomado relevancia en diferentes áreas como en la salud, la biología, y la geodesia, entre otras áreas. Tareas como el detectar y clasificar autos que transitan por una calle, clasificar tipos de suelo mediante imágenes digitales o detectar diferentes enfermedades en imágenes médicas son tareas que pueden ser automatizadas utilizando la IA, en especial técnicas de aprendizaje profundo.

Hoy en día, los problemas de detección y clasificación de objetos en imágenes digitales han sido abordados utilizando redes neuronales convolucionales (una técnica de aprendizaje profundo). Este tipo de técnicas adoptan una estrategia de aprendizaje supervisado, donde todos los datos se encuentran previamente anotados. La anotación es el proceso de identificar los objetos contenidos en las imágenes y asignarles una etiqueta (clasificación). Para entrenar modelos basados en aprendizaje profundo, no solamente es necesario contar con una gran cantidad de datos, sino que estos datos se encuentren correctamente anotados. Para cumplir esta condición, los objetos de las imágenes deben encontrarse señalados correctamente y contar con una etiqueta que los distinga. Por lo tanto, un conjunto de datos correctamente anotado es esencial para generar modelos correctos.

La anotación de datos es un proceso que, dependiendo del problema a tratar y la cantidad de datos, puede llegar a ser tardado y repetitivo. Automatizar la anotación de datos ayudaría a generar conjuntos de datos que permitan generar modelos de clasificación en distintos problemas de clasificación como tipos de suelo en imágenes satelitales, autos que circulan por una carretera o especies de plantas.

La identificación de especies de plantas es uno de los problemas que la comunidad científica ha adoptado técnicas de aprendizaje profundo para darle solución. Este problema requiere la recopilación de un gran conjunto de datos que se obtienen a través de trabajos de campo realizados por distintos grupos de investigación, conformados en su mayoría por taxonomos expertos. El número de taxonomos expertos en el país que garanticen la autenticidad de las especies es limitado, por lo que lograr una identificación completa de las distintas especies del campo se vuelve una tarea complicada. En particular, es necesario procesar y analizar diferentes formas y texturas de especies de plantas. La tarea de identificación de plantas ha resultado ser un proceso difícil. La razón no es sólo por la gran cantidad de especies vegetales que existen (más de 1,600,000 aproximadamente), sino también por la gran diversidad que existe entre sus propiedades taxonómicas. Por esto, contar con sistemas computacionales que apoyen a la tarea de automatizar la identificación de especies de plantas ayudaría a eficientar la solución a este tipo de problemas.

PlantCLEF [1] es el reto anual más importante sobre identificación de imágenes de plantas. El reto recopila una gran colección de imágenes sin procesar para generar un conjunto de datos del reto. Otros conjuntos de datos públicos importantes para la identificación de plantas son Flavia [2], SmithSonian Leaf [3] y Swedish Leaf [4]. Estos conjuntos de datos se anotan manualmente, lo que requiere personal especializado y tiempo para realizar las anotaciones correspondientes de los objetos contenidos en las imágenes. Existen servicios para la anotación de datos (Google Cloud, Amazon SageMaker Data Labeling, CloudFactory, Oracle Cloud Infrastructure (OCI) Data Labeling) que ofrecen una solución donde el personal de la empresa realiza la anotación de los datos. Este tipo de servicios tienen dos problemas principales. El primer problema es el alto precio a pagar por grandes conjuntos de datos. Por ejemplo, si usamos Google Cloud para la anotación automática del conjunto de datos de ImageNet 2012 (es un conjunto de datos a gran escala de alrededor de 14 millones de imágenes), el precio por imagen trabajada es de $0.025, por lo que tendremos que pagar más de $496,000 dólares. El otro problema es la anotación de conjuntos de datos que tienen poca variación entre las diferentes clases. Por ejemplo, la anotación de la clase correcta entre dos especies con poca variación en las características visuales como el color, la forma o el tamaño no es una tarea fácil. Una alternativa a este tipo de servicios de anotación de datos es entrenar modelos basados en IA que permitan generar conjuntos de datos para el entrenamiento y validación de modelos, con una alta calidad y un bajo costo.

En este trabajo describimos un enfoque de anotación automática para imágenes de flores mexicanas descargadas de un sitio web público de especies de plantas, Naturalista. Para ello, entrenamos un modelo de detección de objetos basado en la arquitectura YOLO (You Only Look Once). YOLO es una técnica de IA, particularmente de aprendizaje profundo, para detectar y clasificar múltiples objetos en imágenes digitales.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

La anotación del conjunto de datos es una tarea esencial para las técnicas basadas en aprendizaje profundo.

Antecedentes

La anotación de datos es un proceso que requiere una gran cantidad de tiempo y recursos humanos. En grandes conjuntos de datos, la anotación manual hace a este proceso una tarea complicada [5]. Una propuesta para facilitar la anotación de datos es utilizar técnicas de aprendizaje de máquina y aprendizaje profundo que permitan generar modelos para automatizar la anotación de datos.

Técnicas de aprendizaje de máquina para la anotación automática

El aprendizaje de máquina ha demostrado dar soluciones a problemas como segmentación o clasificación de imágenes, por lo que distintos autores han optado por utilizar técnicas de aprendizaje de máquina para semiautomatizar la anotación de conjuntos de datos, planteando la anotación de datos como un problema de segmentación o clasificación. Una técnica de aprendizaje de máquina para automatizar la anotación de datos son las máquinas de vectores de soporte (SVM por sus siglas en inglés) [6]. Al utilizar las SVM en un problema de anotación en imágenes digitales, se puede dividir la imagen en distintas regiones, donde cada región representará una clase a anotar. Utilizando las SVM, podemos clasificar pixeles, para luego unir con pixeles contiguos de la misma clase para formar regiones. Estas regiones contienen objetos, donde una persona realizará la anotación de estas regiones propuestas. Otra técnica de aprendizaje de máquina utilizada para realizar la anotación de datos son los bosques aleatorios (*Random Forest* por su definición en inglés). Con esta técnica, diversos autores han reducido el tiempo de anotación de los datos, llegando a reducir una anotación de video manual de 3.1 minutos a 0.6 minutos aproximadamente [7]. Utilizando los bosques aleatorios como un clasificador, es posible clasificar datos de señales en múltiples clases.

En este trabajo, utilizamos una técnica de detección de objetos, basada en la arquitectura YOLO (You Only Look Once), para entrenar un modelo de detector que automatice la anotación de datos de especies de plantas con flores. YOLO [8] es una arquitectura de código abierto ampliamente utilizada en problemas de detección de objetos en imágenes, videos y tiempo real. El procesamiento de YOLO se basa en el uso de convoluciones. La convolución es una operación matemática que extrae información relevante de una imagen mediante multiplicación de matrices, utilizando para esto kernels o filtros.

En la Figura 1 se muestra un diagrama de la arquitectura de YOLO. Esta arquitectura consta de una capa convolucional seguida de cinco bloques convolucionales. Estos bloques se componen de una capa convolucional para reducir dimensiones y múltiples capas convolucionales para extraer características. En el tercer, cuarto y quinto bloque la salida se conecta a bloques de detección, detectando así objetos en pequeña, mediana y gran escala. En la arquitectura YOLO, la imagen de entrada se divide en una rejilla de $S \times S$ celdas. Si el centro de un objeto está en una celda de la rejilla, esa celda se utilizará para predecir el objeto. En cada celda se predice un cuadro delimitador B, incluyendo un valor de confianza y las probabilidades de las clases.

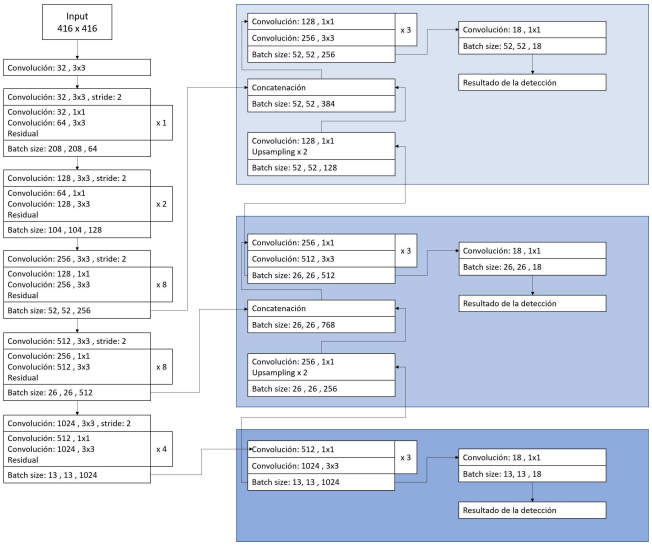


Figura 1. Arquitectura de YOLO [8].

YOLO utiliza una función que calcula la similitud entre el objeto verdadero y la predicción mediante la métrica de Intersección Sobre Unión (*IOU* por sus siglas en inglés). Definimos el objeto verdadero como el área anotada de la imagen correspondiente a un objeto. La métrica IOU calcula qué porcentaje del área predicha está sobre el objeto verdadero. Cuando más de un cuadro delimitador detecta el mismo objeto, YOLO utiliza el método de supresión no máxima (*NMS* por sus siglas en inglés) para seleccionar el mejor cuadro delimitador según el valor de confianza. El método NMS consiste en seleccionar el cuadro delimitador con el valor de confianza más alto y compararlo con los otros cuadros delimitadores utilizando la métrica IOU. Finalmente, se elimina los cuadros delimitadores que tengan un IOU más bajo que un umbral establecido.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Automatizamos la anotación en conjuntos de datos mediante un módulo de entrenamiento y un módulo de generación de datos.

Aportaciones

En esta sección describimos las distintas tareas para automatizar el proceso de anotación en conjuntos de datos. El flujo de trabajo se ilustra en la Figura 2. Utilizamos imágenes de la flora de México para entrenar nuestro modelo detector basado en YOLO. Las tareas se dividen en dos módulos, un módulo de entrenamiento del modelo detector y un módulo de generación de conjuntos de datos. El módulo de entrenamiento se conforma de la creación de un conjunto para entrenar nuestro modelo detector. Una vez creado el conjunto de entrenamiento, el siguiente paso es el entrenamiento del modelo basado en YOLO. El módulo de generación de conjuntos de datos se conforma de la detección de objetos, que consta de ejecutar el modelo basado en la arquitectura YOLO utilizando como información de entrada imágenes de plantas sin procesar y, por último, el modelo genera como información de salida, las coordenadas de los objetos detectados en las imágenes de entrada. Estas coordenadas son utilizadas para, que de manera automatizada, extraer imágenes de áreas de interés como recortes de la imagen original.

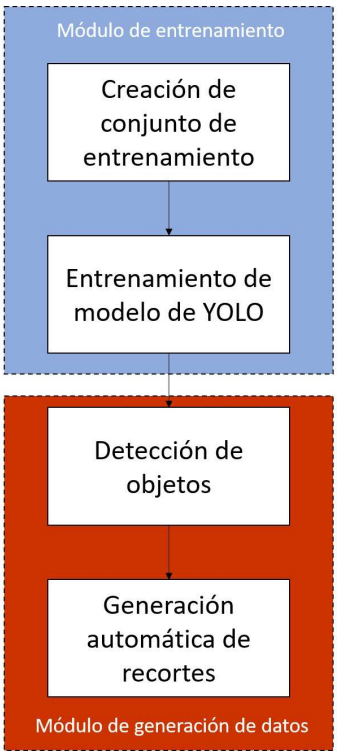


Figura 2. Flujo de trabajo de nuestra propuesta.

Creación del conjunto de entrenamiento

Creamos un conjunto de datos de imágenes de la flora mexicana, utilizando como órgano distintivo la flor de la planta. Este conjunto de datos es una recopilación de imágenes de plantas que fueron procesadas para realizar la anotación de las flores contenidas en ellas. En particular, las imágenes fueron recopiladas de Naturalista [9], una red social en línea para compartir información sobre biodiversidad. Las imágenes utilizadas en este trabajo tienen licencia que permite su descarga, uso y modificación. El conjunto de datos consta de imágenes a color de especies de flores de la flora mexicana, con 2,000 imágenes para entrenar y 161 imágenes para validar el modelo detector. Estas imágenes corresponden a 23 especies de plantas. El conjunto de datos está conformado de imágenes de flores con diferentes tamaños, formas y colores. En la Figura 3 mostramos ejemplos de imágenes del conjunto de datos recopilado. Este conjunto de datos fue etiquetado de manera manual para poder utilizarlo en el entrenamiento del modelo de detección.



Figura 3. Ejemplos de imágenes del conjunto de datos.

Para la anotación de las flores contenidas en las imágenes utilizamos el software labelImg para generar cuadros delimitadores. La información de un cuadro delimitador se conforma por un valor que representa la clase a la que pertenece el objeto, coordenada X y coordenada Y que corresponden a la esquina superior izquierda del cuadro y, además del ancho y alto del cuadro. La Figura 4 muestra un ejemplo de anotación de flores usando el software labelImg.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691



Figura 4. Ejemplo de anotación en imágenes de flores.

Entrenamiento de modelo de YOLO

Los experimentos se realizaron en una computadora con procesador Intel Xeon W-2133, 32 GB de RAM y una tarjeta de video NVIDIA GeForce GTX 1080. La computadora usó Ubuntu 18.04 como sistema operativo y la biblioteca CUDA toolkit 10.0, además de las bibliotecas de Python y el proyecto Darknet para el modelo YOLO V3.

Para el proceso de entrenamiento del modelo detector, utilizamos aprendizaje por transferencia del conjunto de datos COCO [10]. Este conjunto de datos cuenta con imágenes anotadas correspondientes a 80 clases. El aprendizaje por transferencia es el proceso de reutilizar los valores de los parámetros adquiridos de un modelo entrenado en un conjunto de datos de origen como punto de partida para nuevos modelos que serán entrenados utilizando otro conjunto de datos. Los hiperparámetros utilizados para entrenar el modelo detector fueron: número de épocas igual a 4,000, un tamaño de lote de 64, una subdivisión de 16 y una tasa de aprendizaje de 0.001.

Creación y automatización de recortes

El resultado de las predicciones de YOLO es un archivo con información con cuadros delimitadores. Usamos el valor de la coordenada X, la coordenada Y, el ancho, la altura y el valor de clase para crear un recorte del cuadro delimitador. Recortamos las imágenes en tamaño $N \times N$, donde N es el valor máximo entre el ancho o el alto del cuadro delimitador.

Para validar el modelo detector basado en YOLO, utilizamos un conjunto de datos con 161 imágenes en total (siete imágenes por clase) y 302 objetos en las imágenes.

En la Figura 5 se ilustra el flujo de trabajo aplicado a nuestro caso de estudio. El flujo de trabajo empieza utilizando imágenes sin procesar, donde se encuentren una o múltiples flores. Lo siguiente es utilizar estas imágenes como entrada a nuestro modelo basado en YOLO, el cuál nos dará como salida una imagen con los objetos detectados y los datos referentes a los objetos detectados. Por último, utilizaremos estos datos para generar así imágenes de solo flores a partir de las imágenes sin procesar.

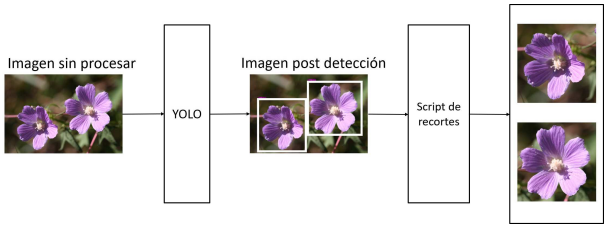


Figura 5. Flujo de trabajo aplicado a nuestro caso de estudio.

El modelo detector se evaluó utilizando los índices de precisión (1) y sensibilidad (2).

$$Precisin(P) = \frac{VP}{VP + FP} \quad (1)$$

$$Sensibilidad(S) = \frac{VP}{VP + FN} \quad (2)$$

Donde VP indica el número de verdaderos positivos (hay una flor en la imagen y el modelo la detecta correctamente), FP indica el número de falsos positivos (no hay flor en la imagen, pero el modelo la detecta) y FN indica el número de falsos negativos (el modelo no detecta una flor que en realidad hay en la imagen). La precisión es una medida que indica cuántas de las predicciones positivas realizadas son correctas. La sensibilidad es un índice del desempeño del modelo sobre el total de predicciones realizadas, es decir es la proporción de casos positivos que fueron correctamente identificados. El modelo obtuvo una precisión de 0.97 y una sensibilidad de 0.94.

La Figura 6 muestra ejemplos de anotaciones de flores que el modelo realizó en imágenes de prueba, con flores únicas y múltiples en la imagen. En las imágenes A-D de la Figura 6, el modelo detectó cada una de las flores de manera individual. En algunos casos, no es posible separar las flores en la imagen, por lo que algunas de las flores se traslapan (una flor o flores que cubran total o parcialmente a otras). Por lo tanto, asumimos como una detección correcta si el modelo detectó las flores traslapadas como una sola flor o si el modelo detectó algunas flores individualmente. Por ejemplo, en la Figura 6 E y F, el modelo detectó las flores traslapadas como una sola flor.

Las técnicas de IA son cada día más utilizadas para solucionar problemas de detección y clasificación de objetos en imágenes.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

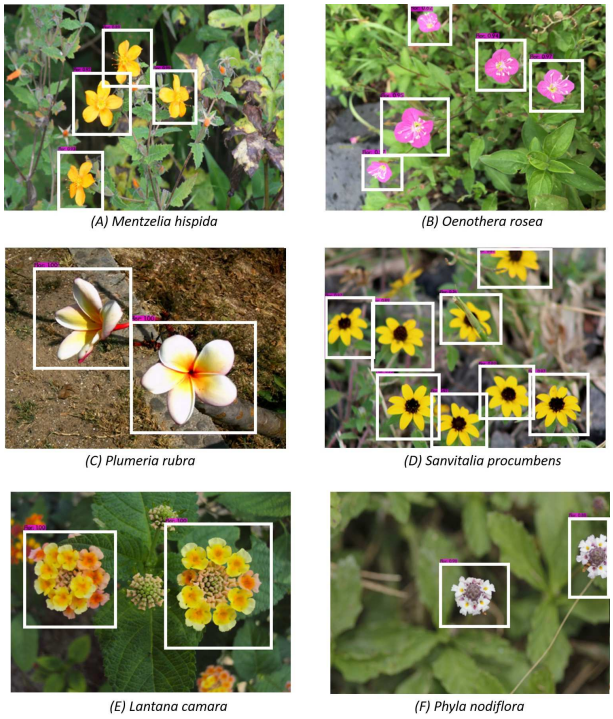


Figura 6. Ejemplos de detección de flores en imágenes digitales.

En la Figura 7 presentamos algunos casos en donde el modelo no fue capaz de detectar todas las flores que aparecen en las imágenes. En la Figura 7 A, notamos que algunas flores aparecen fuera de foco y el modelo no pudo detectarlas. En la Figura 7 B, notamos que hay flores traslapadas, en consecuencia el modelo no pudo separar correctamente todas las flores en la imagen, pero aún así hubo detecciones correctas.



Figura 7. Ejemplos de detecciones incompletas.

Además, realizamos un experimento para cuantificar el tiempo que se toma entre realizar el etiquetado de manera manual con respecto a realizarlo de manera automática. El etiquetado de manera manual de una imagen requiere en promedio 144 segundos. Ahora bien, con nuestra propuesta solo se necesitan, en promedio, de 0.266 segundos para procesar una imagen. Este análisis lo realizamos con personal que fue instruido por expertos para la anotación de regiones de interés (flores) en imágenes de plantas vivas. Aunque, sabemos que este análisis es subjetivo, pero nos brinda una idea de los tiempos de anotaciones manuales y automatizadas.

Por último, en la Figura 8 se presenta una muestra del conjunto de imágenes generados a partir de nuestra propuesta, en donde podemos ver distintas especies de plantas con colores y formas diferentes.



Figura 8. Ejemplos de elementos del conjunto de imágenes generado con YOLO.

Conclusiones

Las técnicas de IA son cada día más utilizadas para solucionar problemas de detección y clasificación de objetos en imágenes digitales. Uno de los mayores obstáculos al momento de utilizar estas técnicas es contar con un conjunto de datos anotado. Esto hace que la anotación de los datos sea una tarea fundamental para las técnicas de IA, particularmente de las técnicas de aprendizaje profundo. En este articulo, describimos una manera de aplicar técnicas basadas en IA para automatizar la anotación de conjuntos de datos y, poder así, generar conjuntos de datos adecuados para las técnicas de aprendizaje profundo. Nuestro modelo detector basado en YOLO realizó anotaciones correctas de diferentes especies de flores de muy diversos tamaños.

Agradecimientos

Los autores desean agradecer al Consejo Mexicano de Ciencia y Tecnología (CONACYT) por la beca otorgada al primer autor, así como por el financiamiento para el proyecto 291772 del fondo CONACYT-INEGI. De la

misma manera, los autores agradecen el apoyo económico brindado por la Universidad Autónoma de Sinaloa.

REFERENCIAS

1. Goeau, H., Bonnet, P., y Joly, A. (2021). Overview of PlantCLEF 2021: Cross-domain Plant Identification. En *Proc. CLEF 2021 - Conference and Labs of the Evaluation Forum*, 1422–14.
2. Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y.-X., Chang, Y.-F., y Xiang, Q.-L. (2007). A Leaf Recognition Algorithm for Plant Classification using Probabilistic Neural Network. En *Proc. of the IEEE International Symposium on Signal Processing and Information Technology*, 11–16.
3. Belhumeur, P. N., Chen, D., Feiner, S., Jacobs, D. W., Kress, W. J., Ling, H., Lopez, I., Ramamoorthi, R., Sheorey, S., White, S., y Zhang, L. (2008). Searching the World's Herbaria: A System for Visual Identification of Plant Species. En *Proc. of the Computer Vision – ECCV*, 116–129.
4. Soderkvist, O. (2001). Computer vision classification of leaves from swedish trees. Master's Thesis. Linkoping University
5. Fernandez, J. B., Venkatesh, G. M., Zhang, D., Little, S., y O'Connor, N. E. (2019). Semi-automatic Multi-object Video Annotation based on Tracking, Prediction and Semantic Segmentation. En *Proc. of the International Conference on Content-Based Multimedia Indexing (CBMI)*, 1–4.
6. Cusano, C., Ab, G., Ciocca, G., y Schettini, R. (2003). Image Annotation using Svm. En *Proc. of the SPIE - The International Society for Optical ENgineering*.
7. Schreiner, C., Torkkola, K., Gardner, M., y Zhang, K. (2006). Using machine learning techniques to reduce data annotation time. En *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, 2438–2442.
8. Redmon, J., Divvala, S., Girshick, R., y Farhadi, A. (2016). You Only Look Once: Unified, Real-time Object Detection. *En Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788.
9. Naturalista. (2023). `https://www.naturalista.mx`.
10. Tsung-Yi Lin, Maire M., Belongie S. J., Bourdev L. D., Girshick R. B., Hays J., Perona P., Ramanan D., Dolla P. y Zitnick C. L. (2014). Microsoft COCO: Common Objects in Context. *CoRR*. abs/1405.0312.

SOBRE LOS AUTORES



Eduardo Díaz-Gaxiola recibió el grado de Licenciado en Informática por la Universidad Autónoma de Sinaloa, en el 2018 y en el 2021 su grado de maestro en Ciencias de la Información. Actualmente, es estudiante del programa de doctorado en Ciencias de la Información de la Universidad Autónoma de Sinaloa. Entre sus líneas de interés se encuentran aprendizaje profundo, técnicas de aprendizaje máquina y visualización de redes neuronales convolucionales.



Gerardo Beltrán-Gutiérrez es profesor de Ciencias de la Computación en la Unviersidad Autónoma de Sinaloa (Culiacán, México). Obtuvo su maestría en Ciencias de la Computación por la Universidad Autónoma de Sinaloa. Sus líneas de interés incluyen bases de datos, ciencia de datos y programación paralela y distribuida.



Inés Fernando Vega-López es profesor de Ciencias de la Computación en la Universidad Autónoma de Sinaloa (Culiacán, México) desde el 2004. Recibió su Doctorado en Ciencias de la Computación por la Universidad de Arizona en 2004. El Dr Vega-López es miembro del Sistema Nacional de Investigadores Nivel I. Sus intereses de investigación actuales incluyen sistemas de bases de datos de alto rendimiento, análisis de datos a gran escala y aprendizaje automático.



Arturo Yee-Rendón es profesor de Ciencias de la Computación en la Universidad Autónoma de Sinaloa. Obtuvo su maestría y doctorado en Ciencias de la Computación por el CINVESTAV-IPN en el 2010 y 2015, respectivamente. El Dr. Yee-Rendón es miembro del Sistema Nacional de Investigadores Nivel I. Sus intereses de investigación actuales incluyen reconocimiento de patrones, técnicas de aprendizaje profundo, teoría de juegos y algoritmos de optimización (algoritmos genéticos).

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Reconocimiento de Patrones de Artes de Pesca como Herramienta para Combatir la Pesca Ilegal

Ramón Aranda, Hugo Carlos, Angel Díaz-Pacheco, Miguel Á. Álvarez-Carmona y Ansel Y. Rodríguez-González

Resumen

La pesca es una de las actividades más importantes en todo el mundo, ya que proporciona una fuente considerable de alimentos e ingresos económicos a nivel mundial. A pesar de que la actividad pesquera es de gran importancia, actualmente esta industria enfrenta diversas amenazas, principalmente la Pesca Ilegal, no Declarada y no Reglamentada (IUUF, por sus siglas en inglés: Illegal, Unreported, and Unregulated Fishing). La trazabilidad puede ayudar a combatir esa amenaza y lograr la sustentabilidad de la industria pesquera. Mediante el uso de sistemas de monitoreo de embarcaciones, es posible identificar la posición de un barco en tiempo real e inferir su actividad o comportamiento. Por esta razón, en este trabajo de divulgación se describen diferentes métodos y tecnologías basadas en Inteligencia Artificial (IA) para identificar automáticamente los comportamientos de las embarcaciones.

Introducción: La pesca

La pesca es una de las actividades más importantes en todo el mundo, ya que proporciona una fuente considerable de alimentos e ingresos económicos a nivel mundial. Los productos pesqueros como los peces, crustáceos, moluscos y otros animales acuáticos, excluyendo mamíferos acuáticos, reptiles, algas y otras plantas acuáticas, alcanzaron a nivel mundial 179 millones de toneladas en 2018 y 178 millones de toneladas en 2019 y 2020, de los cuales 156 millones de toneladas se destinaron al consumo humano; lo que equivale a un suministro anual estimado de 20.5 kg por persona.

El importe económico de la producción pesquera en 2020 tiene un valor total estimado de primera venta de 406 mil millones de dólares, y aproximadamente el 51 % de esta producción proviene de la pesca de captura en aguas marinas y continentales. La pesca de captura se ha mantenido bastante estable, alrededor de 90 millones de toneladas desde principios de la década de 1990, con algunas fluctuaciones interanuales en el rango de 3 a 6 millones de toneladas. Estas fluctuaciones son particularmente determinadas por la variabilidad del cambio climático, pero también de la productividad de los ecosistemas, el manejo de la intensidad pesquera y los recursos pesqueros [1].

En 2020, 156 millones de toneladas de la producción pesquera se destinaron al consumo humano

La Pesca Ilegal, no Declarada y no Reglamentada

A pesar de que la actividad pesquera es de gran importancia de nivel mundial, actualmente esta industria enfrenta diversas amenazas, principalmente la IUUF:

Pesca ilegal: se refiere a las actividades pesqueras realizadas en contravención de las leyes y reglamentos aplicables, incluidas las leyes y normas adoptadas a nivel regional e internacional.

Pesca no declarada: se refiere a las actividades pesqueras que no se declaran o se declaran erróneamente a las autoridades competentes en contravención de las leyes y reglamentos nacionales o los procedimientos de notificación de una organización regional de ordenación pesquera pertinente.

Pesca no reglamentada: tiene lugar en áreas o poblaciones de peces para las que no existen medidas de conservación o gestión aplicables y donde dichas actividades pesqueras se llevan a cabo de manera incompatible con las responsabilidades del Estado para la conservación de los recursos marinos vivos conforme al derecho internacional.

La IUUF es un problema mundial que amenaza los ecosistemas oceánicos y la pesca sustentable, ya que violan las normas de pesca nacionales e internacionales. También amenaza la seguridad económica y los recursos naturales que son fundamentales para la seguridad alimentaria mundial, y pone en desventaja a los pescadores y productores que respetan las leyes internacionales sobre pesca.

La industria pesquera enfrenta amenazas, principalmente la IUUF

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Trazabilidad de productos pesqueros

La trazabilidad es la capacidad de dar seguimiento a un producto en cada una de sus etapas, desde el lugar de origen de los recursos pesqueros en el océano hasta el punto de venta, considerando también el arte de pesca utilizado para capturarlo. De esta manera, la trazabilidad facilita el conocimiento de la identidad, la historia y la fuente de un producto o sus componentes. También facilita el conocimiento sobre el destino de un producto, o cualquier ingrediente contenido en él. En el sector de la pesca, la información recolectada se usa para asegurar que los productos provienen de orígenes que cumplen con las condiciones de inocuidad de los alimentos para asegurar que se apliquen las normas y leyes internacionales, así como también para asegurar que el pescado provenga de fuentes sostenibles o saber que embarcaciones siguen las reglas de conservación. La trazabilidad es necesaria para combatir amenazas como la IUUF y lograr la sustentabilidad de la pesca [2].

Para una trazabilidad completa es necesario monitorear las trayectorias de las embarcaciones. Hay dos sistemas que se utilizan principalmente para el seguimiento de embarcaciones: los sistemas de identificación automática (AIS, por sus siglas en inglés) y el sistema de monitoreo de embarcaciones (VMS, por sus siglas en inglés). AIS fue pensado principalmente como un sistema de prevención de colisiones para barcos, aunque recientemente, las señales AIS han sido capaces de monitorear embarcaciones de manera continua Por el contrario, VMS fue diseñado como un sistema de monitoreo y seguimiento de embarcaciones desde el principio.

Mediante el uso de datos geo-referenciados de las trayectorias por medio de sistemas VMS o AIS, es posible identificar la posición del barco en tiempo real e inferir su actividad. Sin embargo, inferir las actividades de las embarcaciones no es una tarea fácil. Por esta razón, para identificar automáticamente los comportamientos de las embarcaciones (por ejemplo, patrones de artes de pesca), se han propuesto diferentes métodos y tecnologías basadas en IA. Así, dada una trayectoria y algunas características (*velocidad*, *distancia desde la costa*, *distancia desde el puerto*, *rumbo*, *latitud* y *longitud*) se pueden analizar para obtener patrones de comportamientos, como por ejemplo de las diferentes artes de pesca.

La trazabilidad es necesaria para combatir amenazas como la IUUF y lograr la sustentabilidad pesquera

Artes de Pesca

Se denominan artes de pesca a los métodos utilizados en la captura y extracción de su medio natural de los peces u otras especies acuáticas como crustáceos, moluscos y otros invertebrados. La captura de peces se considera legal, si el arte de pesca usado es el adecuado y reglamentado para la especie a ser capturada. A continuación presentamos unos ejemplos de artes de pesca [1]:

Pesca de arrastre implica arrastrar una o más redes detrás de un barco de pesca, ya sea en el fondo del mar (arrastre de fondo) o en la columna de agua (arrastre pelágico o de media agua). Durante la pesca de arrastre, los barcos de pesca suelen reducir la velocidad y tratar de mantener una velocidad constante para mantener la tensión en la red arrastrada lo más uniforme posible.

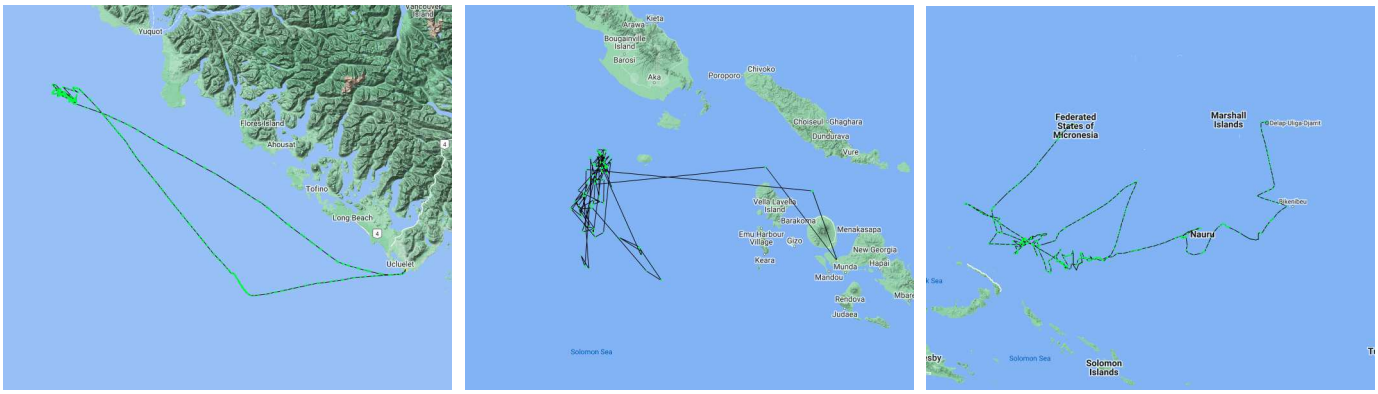
Pesca con palangre implica varias líneas de pesca (de hasta 100 km de longitud) equipadas con varios cientos o varios miles de anzuelos. Las líneas se pueden desplegar a varias profundidades con el uso de flotadores y líneas horizontales que se extienden a aguas más profundas. Cuando se inicia, el barco viaja solo un poco más lento que su velocidad de navegación mientras se establece la línea. Durante este tiempo, el barco se desplaza lentamente con la línea y/o coloca otras líneas en las cercanías.

Redes de cerco son redes largas desplegadas que cuelgan verticalmente de flotadores alrededor de cardúmenes de peces en o cerca de la superficie por el barco o por un esquife separado. Para evitar que los peces se escapen, el calado de la red de cerco debe ocurrir rápidamente y se realiza a altas velocidades con un promedio de alrededor de 10 nudos. Una vez que la red rodea la escuela por completo, la parte inferior de la red se cierra y se tira de la red. A la deriva con la red atada, los peces se recuperan y se transfieren a la embarcación.

La Figura 1 muestra ejemplos de trayectoria de embarcaciones con sistemas AIS empleando diferentes tipos de arte de pesca. Las trayectorias mostradas fueron tomadas de la base de datos llamada “Anonymized AIS training data”, obtenidad de la plataforma Global Fishing Watch[2]. Los datos de Global Fishing Watch están bajo las licencias *Creative Commons Attribution-ShareAlike 4.0* y *Apache 2.0*. Dicho conjunto de datos está diseñado para entrenar algoritmos de IA para detectar patrones de pesca.

[1] http://www.fao.org/fishery/
[2] https://globalfishingwatch.org/

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691



(a) Pesca por arrastre (b) Pesca con palangre (c) Redes de cerco

Figura 1. Ejemplos visual de trayectorias de embarcaciones usando diferentes artes de pesca.

Las artes de pesca son los métodos utilizados en la captura y extracción de su medio natural de los peces u otras especies acuáticas

IA y tecnologías para identificar artes de pesca

El reconocimiento automático de artes de pesca a través de las trayectorias de sistemas AIS o VMS, se ha abordado en varias investigaciones. Por ejemplo, los autores en [3] aplicaron tres metodologías diferentes (modelos oculto de Markov, minería de datos y filtro de mediana) para identificar los puntos de actividad pesquera de tres tipos de artes de pesca: arrastre, cerco y palangre. Los autores en [4] interpolaron datos de trayectoria AIS con un intervalo de tiempo regular y extrajeron la velocidad del barco y el cambio de rumbo para la clasificación de artes de pesca usando la técnica *map reduce* de big data. En [5], los autores propusieron utilizar un modelo de red neuronal convolucional, y se utilizaron un método basado en ventanas (vecinos) para segmentar la información de las trayectorias. Los autores en [6] combinaron la extracción de características derivadas de VMS y las caracterización especificas de artes de pesca, y utilizaron los métodos random forest y máquina de vector de soporte para clasificar las artes de pesca. Los autores en [7] propusieron un enfoque de red neuronal recurrente (RNN) para discriminar los tipos de embarcaciones pesqueras a partir de sus trayectorias. Ellos segmentan las trayectorias para calcular un vector de características como entrada para la RNN.

Para todos los métodos mencionados, es común que dado un conjunto de trayectorias se realicen las siguientes etapas:

1. Pre-procesamiento de datos,
2. el calculo no supervisado del vector características,
3. un paso de entrenamiento de los modelos,
4. y, finalmente, la identificación o clasificación de las artes de pesca.

Un trabajo reciente incluye un paso adicional que demostró buenos resultados: los autores en [8] proponen la inclusión de un paso de reducción de dimensionalidad por medio de un autocodificador al vector características. La reducción dimensional ha sido una técnica ampliamente utilizada para comprender datos de alta dimensionalidad, incluida la visualización, la identificación de patrones, la separabilidad y la estructura. Cuantas más características hay, mayor es el espacio de combinaciones de valores de dichas características, por lo que se necesitan más muestras para entrenar un modelo de IA. Así, el paso de reducción dimensional permite eliminar las características redundantes y el ruido, evitando el sobreajuste, disminuyendo la complejidad de los datos y preservando las diferencias entre diferentes artes de pesca.

En [8], los autores proponen calcular un vector de características formado por la media, mediana y sesgo de la velocidad, la distancia desde la costa, la distancia desde el puerto, el curso, la latitud y la longitud por cada trayectoria. De esta forma, caracterizan cada trayectoria a través de un vector de características con 18 variables (18–FV). Clasificando sobre el vector 18–FV directamente, los resultados en *exactitud* obtenidos usando dos diferentes clasificadores son 53 % para el clasificador Naive-Bayes y 86 % para el método de la Máquina de Soporte Vectorial. Pero cuando aplica el paso reducción de dimensionalidad, los resultados mejoran considerablemente: 95 % tando para Naive-Bayes como para la Má-

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

quina de Soporte Vectorial. Esto indica que la reducción de dimensionalidad ayuda a los clasificadores a tener un mejor desempeño.

Aunque muchos de los trabajos que estudian en el problema reconocimiento de artes de pesca se basan en el análisis de trayectorias provenientes de sistemas AIS o VMS, también existen tecnologías alternativas que están emergiendo para monitorear el comportamiento de las embarcaciones. Algunos ejemplos son:

- En [9], evalúan si los sensores más comunes integrados en los dispositivos móviles actuales (GPS, acelerómetro, giroscopio y el campo magnético) identifican las diferentes fases de la actividad de pesca de arrastre.
- En [10], los autores propusieron un sistema inteligente de videovigilancia que puede detectar automáticamente si y cuándo se producen actividades de pesca mediante el control visual de las actividades de la embarcación.

Conclusiones

Este trabajo de divulgación presentó diferentes enfoques novedosos basados en análisis de datos provenientes de sistemas AIS o VMS para el reconocimiento de artes de pesca. Este tipo de enfoques computacionales pueden ayudar a abordar estrategias de gestión pesquera más eficientes y así mejorar y facilitar la trazabilidad pesquera, el cual es un tema muy relevante para el desarrollo sustentable de esta actividad. Sin embargo, el análisis en tiempo real de los datos AIS o VMS aún no es obligatorio y necesita expertos monitoreando la información, lo cual no sería práctico. Las empresas y el gobierno deben comenzar a considerar estos enfoques de análisis automáticos en entornos operativos y utilizarlas en situaciones reales para validar sus capacidades y limitaciones, además de abrir la posibilidad de retroalimentación de otros expertos y enriquecer otras investigaciones que busquen mejorar el reconocimiento de los comportamientos de las embarcaciones pesqueras.✷

REFERENCIAS

1. FAO. (2022). World Food and Agriculture – Statistical Yearbook 2022. Rome. ISBN 978-92-5-136930-2.
2. He, J. (2018). From country-of-origin labelling (COOL) to seafood import monitoring program (SIMP): How far can seafood traceability rules go?. *Marine Policy*, 96, 163–174.
3. de-Souza, E. N., Boerder, K., Matwin, S. y Worm, B. (2016). Correction: Improving fishing pattern detection from satellite ais using data mining and machine learning. *Plos one*, 11(9), 1–20.
4. Park, J. W., Lee, K. M., y Kim, K. I. (2019). Automatic identification system based fishing trajectory data preprocessing method using map reduce. *International Journal of Recent Technology and Engineering*, 8, 352–356.
5. Kim, K. y Lee, K. M. (2020). Convolutional neural network-based gear type identification from automatic identification system trajectory data. *Applied Sciences*, 10(11), 4010-4031.
6. Marzuki, M. I., Gaspar, P., Garello, R., et al. (2018). Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories. *IEEE Journal of Oceanic Engineering*, 43(3), 689–699.
7. Pipanmekaporn, L. y Kamonsantiroj, S. (2020). A deep learning approach for fishing vessel classification from vms trajectories using recurrent neural networks. En *Proc. 2nd International Conference on Human Interaction and Emerging Technologies: Future Applications*, 135–141.
8. Carlos, H., Aranda R., De-Velasco, M. R., et al. (2022). Fishing Gear Pattern Recognition by Including Supervised Autoencoder Dimensional Reduction. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.
9. Galotto-Tebar, M. M., Pomares-Padilla, A., Czerwinski, I. A., et al. (2020). Using mobile device's sensors to identify fishing activity. *Journal of Marine Science and Technology*, 25(3), 978–989.
10. Perdigao, P., Lousa, P., Ascenso, J., et al. (2020). Visual monitoring of high-sea fishing activities using deep learning-based image processing. *Multimedia Tools and Applications*, 79(31), 22131–22156.

SOBRE LOS AUTORES



Ramón Aranda SNI Nivel 1. Recibió en 2008 el grado de Licenciado en Ciencias de la Computación por parte de la Facultad de Mátematicas de la UADY, en 2010 recibio el grado de Maestro en Ciencias de la Computación con especialidad en Mátematicas Industriales y en 2016 el grado de Doctor en Ciencias con Orientación en Ciencias de la Computación ambos posgrados por el CIMAT. De 2019 a la fecha, es investigador por México de CONACyT, actuamente comisiodo al CIMAT-Unidad Mérida.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691



Hugo Carlos SNI Nivel 1. Recibió el grado de Ingeniero en Sistemas Computacionales por parte de la Facultad de Ingenieria de la UAC. Obtuvo el grado de Maestro en Ciencias de la Computación con especialidad en Mátematicas Industriales y el grado de Doctor en Ciencias con Orientación en Ciencias de la Computación en el CIMAT. Actualmente se desempeña como Cátedra CONACYT en Centro Geo, en las áreas de Sensado Remoto y Geointeligencia Computacional.



Angel Díaz Pacheco. SNI nivel 1. Recibió en 2010 el grado de Ingeniero en Sistemas Computacionales por el Instituto Tecnológico de Veracruz, en 2013 el grado de Maestro en Sistemas Computacionales por el Instituto Tecnológico de Apizaco y en 2019 el grado de Doctor en Ciencias en el área de Ciencias de la Computación por el INAOE. Desde 2022, es profesor-investigador en la Universidad de Guanajuato, Campus Irapuato-Salamanca. Sus intereses de investigación incluyen el aprendizaje automático, Big Data, el cómputo suave, la lógica difusa y la aplicación de las anteriores en la investigación de constructos importantes de la investigación turística.



Miguel Á. Álvarez-Carmona. SNI nivel 1. Recibió en 2012 el grado de Ingeniero en Ciencias Computacionales por por la BUAP, en 2014 y 2019 el grado de Maestro y Doctor en Ciencias Computacionales por el INAOE donde se especializó en Inteligencia artificial y Procesamiento de Lenguaje Natural. Ha sido autor de diversos artículos en revistas y congresos nacionales e internacionales. De 2019 a la fecha, es investigador por México de CONACyT en CIMAT-Unidad Monterrey-UT3. Pertenece a la Asociación Mexicana de Procesamiento de Lenguaje Natural y a la Sociedad Mexicana de Inteligencia Artificial.



Ansel Y. Rodríguez González. SNI Nivel I. Recibió los grados de Licenciado en Ciencias de la Computación y Maestro en Ciencias Matemáticas en la Universidad de la Habana, Cuba en 2004 y 2007 respectivamente. Recibió el grado de Doctor en Ciencias Computacionales en el INAOE en 2011. Ha participado en 15 proyectos de investigación y desarrollo. Actualmente tiene una cátedra CONACyT en la Unidad de Transferencia Tecnológica (UT3) del CICESE. Sus intereses de investigación incluyen tópicos de inteligencia artificial como minería de datos, reconocimiento de patrones, aprendizaje por refuerzo y algoritmos evolutivos.



© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Retos de la Inteligencia Artificial para no comprometer la privacidad de los datos

Miguel Morales Sandoval, Heidy Marisol Marin Castro y Karina Figueroa Mora

Hoy en día existe una gigantesca recopilación de datos a través de casi todos nuestros movimientos en internet. Por ejemplo, mientras compramos, usamos las redes sociales, o realizamos algún trámite en línea, generamos datos digitales como nuestros nombres, direcciones, teléfonos, fotografías o información relacionada con compras en línea (productos, marcas, costos). Los algoritmos de Inteligencia Artificial (IA) generalmente usan estos datos para realizar predicciones u obtener algún conocimiento. Quienes realizan estas tareas de análisis saben que existe un reto enorme sobre el manejo y privacidad de los datos, ya sean personales (número de teléfono) o de carácter privado (salario). Por lo que, el desafío es: ¿cómo entrenamos un modelo de IA sin que esos datos estén vulnerables? El lector podría pensar, "simplemente quitémoslos" y eso sería lo ideal en un escenario de unos pocos registros pero cuando hablamos de cientos de miles de millones es una tarea titánica y prácticamente imposible, ¡primer gran reto! Por supuesto, aunque varía el alcance en cada país, en general cada vez más se desarrollan o actualizan regulaciones en torno a garantizar la privacidad del uso de datos. Por ello, este es un aspecto de gran relevancia para los algoritmos que usan datos para derivar conocimiento, como los de IA [1].

Por otro lado, un entrenamiento para un modelo de IA consiste básicamente en detectar un patrón en los datos. Esta técnica se utiliza para muchas tareas de minería de datos, clasificación, agrupamiento, etc. Por ejemplo, imagine que ud. que se desea realizar una tarea de agrupamiento de salarios de acuerdo con años de experiencia, dentro de un contexto o plano organizacional, para que con estos resultados se puedan realizar más adelante predicciones y una planeación financiera en relación con presupuesto de recursos humanos. Quitar ejemplos de personas con bajo o alto salario, por cuestiones de privacidad, sin duda afectará al modelo de agrupamiento y de predicción, y llevará a una toma de decisiones sin la información correcta. Por otro lado, quien calcule el modelo de agrupamiento o de predicción, estaría comprometiendo la privacidad de los datos de salarios, los cuales deberían ser privados.

En este artículo presentamos el problema relacionado con la privacidad de datos para ser usados en algoritmos de IA y cómo se puede abordar esta problemática desde un enfoque de solución basado en el cifrado homomórfico ("¡¿homomó....qué?!", ya le explicaremos). Esta tecnología le permite al propietario de los datos delegar las tareas de procesamiento y de análisis a un tercero, cifrando los datos de entrada sin que éstos pierdan su utilidad. Esto es, se garantiza que dicho algoritmo no aprenderá nada de los datos del propietario (básicamente, no sabrá de quién o de qué estamos hablando). Finalmente, presentamos algunos de los retos relacionados con esta problemática y oportunidades que se tienen en el plano de la investigación y del desarrollo tecnológico sobre este tema.

La confidencialidad, requisito para la privacidad de datos, debe restringir el acceso a datos solo a las entidades autorizadas. El cifrado de datos ha sido una herramienta efectiva para garantizar la confidencialidad.

Cifrado homomórfico de un texto

La criptografía es un campo de las matemáticas y de las ciencias computacionales, cuyos orígenes se remontan a antes de nuestra era. Iniciemos explicando que significa *cifrar* un texto: básicamente es una transformación de dicho texto. Imagine el lector que quiere cifrar el texto "*hola mundo*". El resultado se puede ver cifrado como "*3oij4s823ADe23*". Sin embargo, el texto "*holamundo*" podría ser cifrado como "*aLKQeirjq34*". Note que a pensar de que solo cambió una letra en ambos textos (el espacio), las transformaciones son completamente distintas. En principio, la transformación debe producir texto ilegible, ya que el principal objetivo del cifrado es ocultar el mensaje original (que si es legible). Existen diversas técnicas y algoritmos para cifrar, el más fácil y básico consiste en reemplazar una letra por otra, y nuestros textos cifrados se verían así "*kpñs qimfp*" y "*kpñsqimfp*". Cómo puede darse cuenta, en este ejemplo se ha usado un algoritmo débil pues se puede intuir que es casi el mismo mensaje. Actualmente, los cifradores son una pieza fundamental en la seguridad de la información digital. Algunos ejemplos de su amplio uso son: cuando

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

nos comunicamos a través de aplicaciones de mensajería instantánea, donde se nos informa que los mensajes intercambiados están cifrados; cuando realizamos compras en línea y accedemos a una página web que se nos dice que es segura porque la conexión está cifrada; o cuando pagamos usando una tarjeta bancaria. Los cifradores modernos son mucho más complejos que el que hemos mencionado en el ejemplo, pero el principio de funcionamiento es el mismo. Todos los cifradores, desde los antiguos hasta los modernos, hacen uso de una *llave* para realizar la transformación. La llave, en este caso, también puede verse como un conjunto de letras y generalmente es de un gran tamaño para evitar que pueda ser adivinada. Entonces, para cifrar se requieren 3 elementos: el mensaje, la llave, y el algoritmo de cifrado. Todo esto se muestra en la Figura 1. Una vez cifrado, un mensaje solo podrá restaurarse a su estado original mediante el descifrado, el cual revierte la transformación realizada por el cifrado.

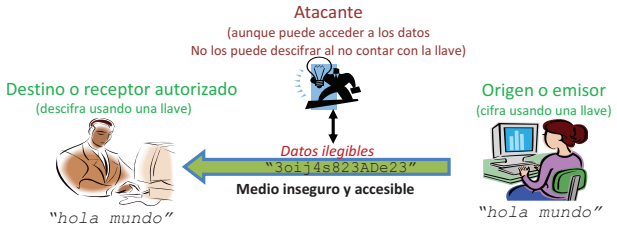


Figura 1. Modelo básico donde la criptografía garantiza la privacidad de los datos que se generan en un origen, se transmiten o se almacenan en un medio inseguro que puede ser accedido por terceros (atacante), pero que solo es accesible por entidades autorizadas que cuente con la llave para descifrarlos.

Comúnmente, el cifrado se visualiza como una caja que resguarda un mensaje, cerrada con un candado bajo llave. Nadie, excepto quien cuente con la llave que abre dicho candado, podrá abrir la caja y extraer el mensaje. En la práctica, esta analogía se implementa con algoritmos basados en principios matemáticos, particularmente en el álgebra abstracta. La criptografía tiene una relación muy estrecha con el origen de la computación y de la IA. Alan Turing, considerado como el padre de la IA, lideró el equipo que estuvo a cargo de descifrar los códigos de la máquina Enigma, un cifrador usado por los alemanes y sus aliados durante la Segunda Guerra Mundial. Turing y su equipo pudo descifrar los códigos de Enigma gracias a la creación de un dispositivo, que evolucionó para convertirse en el modelo de computación que se usa en todos los dispositivos de cómputo actuales.

El reto es mantener la privacidad: obtener conocimiento sin revelar nada de los datos de entrada.

El cifrado hace ilegible un texto, mensaje o dato, y en principio le quitaría utilidad a los datos en aras de mantener su privacidad. Pero, ¿se podrían seguir aplicando algoritmos de IA sobre los datos cifrados?. La respuesta, estimado lector, es que sí es posible, y se puede lograr usando un cifrado especial, llamado *cifrado homomórfico*.

Para la IA, la relevancia del cifrado homomórfico es que le puede proporcionar la capacidad de seguir realizando operaciones sobre los datos cifrados, de tal forma que el procesamiento resultante es equivalente al cifrado del procesamiento de los datos sin cifrarlos. Esto significa que es posible realizar cálculos sobre los datos sin tener acceso a éstos en claro, y así preservar la privacidad de los mismos. Este no es el caso de las técnicas de cifrado tradicionales, donde el resultado de operaciones que involucran datos cifrados no tiene una interpretación significativa (como nuestro primer ejemplo). Sin embargo, el cifrado homomórfico comparte muchos de los conceptos, términos y notación con la criptografía convencional [2].

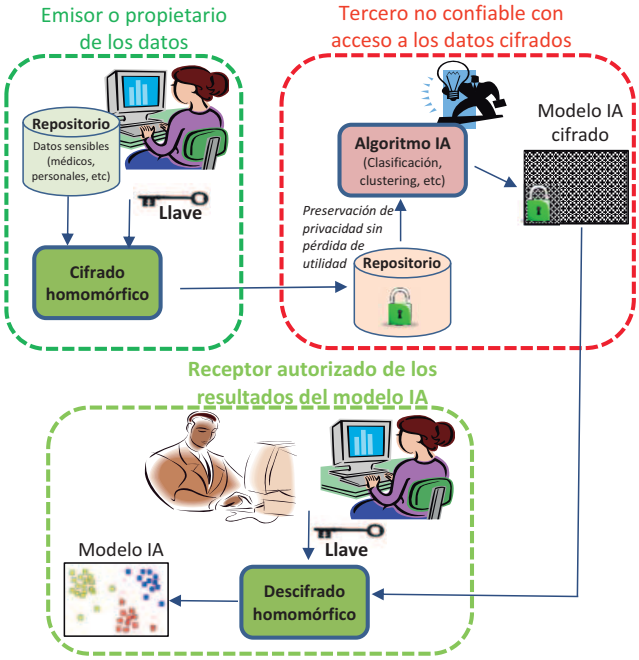


Figura 2. El cifrado homomórfico permite mantener la privacidad de los datos sin pérdida de utilidad para algoritmos de IA.

Al igual que en el cifrado convencional, en un cifrado homomórfico se tienen 3 operaciones: generación de llaves, una para cifrar y otra para descifrar, en algunos casos pueden ser iguales, cifrado (se usa la llave para cifrar) y descifrado (se deshace el proceso de transformación usando la llave para descifrar). En la Figura 2 se pueden ver estas 3 operaciones. Quien ejecuta el algoritmo de IA es el principal atacante a la privacidad

de datos que tienen un propietario. El algoritmo de IA debe seguir funcionando aún con los datos cifrados y obtener el modelo para el que está diseñado. Este modelo, cifrado, no tendría ningún significado para quien ejecuta dicho algoritmo: solo podrá ser descifrado por las entidades autorizadas, que puede ser el mismo propietario de los datos o algún otro.

El cifrado homomórfico supone que es posible definir dos operaciones: una que se aplica sobre los datos legibles y otra que se aplica sobre los correspondientes datos cifrados. La operación con los datos cifrados debe implicar la operación de manera indirecta con los datos originales. Veamos esto a través de un ejemplo. Suponga que los datos legibles son salarios. Ahora, suponga que cifrar estos datos consiste en realizar una exponenciación usando una base fija, por ejemplo, la base 3. Así, el cifrado del salario 1,000 sería 3^{1000} (un número muy grande). Por consiguiente, el descifrado de 3^{1000} debería ser 1,000. Evidentemente, un tercero tendría acceso al cifrado del sueldo, 3^{1000} en este caso, pero no tendría acceso al valor real que es 1,000 (ver Figura 2). Ahora, suponga que multiplicamos dos salarios cifrados (esto lo realizaría el algoritmo de IA, quien operaría sobre datos cifrados), por ejemplo 3^{1000} (el cifrado del salario 1,000) y 3^{1500} (el cifrado del salario 1,500), lo cual, usando la propiedad de exponentes sería $3^{1000} \times 3^{1500} = 3^{(1000+1500)}$. Es decir, *una multiplicación* de datos cifrados resultó en *una suma* de los correspondientes datos originales, y esta suma se obtuvo de manera indirecta, sin tener acceso a esos datos como tal. Para acceder al resultado de la suma (que sería el dato de interés para el propietario de los datos), se debe descifrar el resultado de la multiplicación (como en la Figura 2), que sería el valor (1,000 + 1,500). Esta es la idea básica del cifrado homomórfico. Nuestro querido lector puede corroborar el ejemplo previo dando algunos valores distintos para los salarios y para la base usada.

Desde luego, un aspecto de suma importancia del cifrado homomórfico es la seguridad. En nuestro contexto de la Figura 2, el cifrado debe ser robusto para evitar que el tercero no confiable (con gran capacidad de cómputo dicho sea de paso) pueda romper el cifrado. La seguridad del cifrado homomórfico está generalmente asociado a la dificultad de resolver un problema matemático, mismo que se usa para la generación de las llaves. Entre más difícil el problema a resolver, más seguro (más difícil de romper) es el cifrado. Uno de los problemas mayormente usados es el cálculo de logaritmos discretos o la factorización de números enteros grandes. Actualmente, estos problemas se consideran difíciles si se usan tamaños de números, al menos, de 2048 bits. Las operaciones de cifrado y descifrado involucran operaciones aritméticas con éstos números. Para darse una idea de lo grande que pueden ser estos números, 2048 bits equivale a un número entero de 620 dígitos decimales.

La seguridad de datos se refiere a garantizar el servicio de confidencialidad de los datos que se delegan a un tercero para análisis, y es un requisito para obtener privacidad.

Retos y oportunidades

En este ámbito, la privacidad se centra en los propietarios de los datos, que pueden ser individuos o grupos, con el objetivo de proteger su información privada mientras sus datos se utilizan por terceros para la construcción de modelos de IA procesados en la nube como *la minería de datos como servicio*, esto es, un tercero accede a un repositorio de datos, posiblemente sensibles, para analizarlos y extraer conocimiento de ellos, etc. Así, cuando la naturaleza de los datos sobre los que se aplicarán algoritmos de minería de datos es sensible, como se mostró en los ejemplos anteriores, es necesario resolver desafíos de privacidad.

De manera general, el uso de algoritmos de IA (de análisis para descubrir conocimiento), puede llevar a comprometer la confidencialidad y de privacidad de los datos cuando contengan información sensible o de acceso restringido, o cuando el uso de los datos esté sujeto a regulaciones de acceso.

Reto 1: Eficiencia y nivel de seguridad. El principal desafío del cifrado homomórfico es que los esquemas actuales solo admiten un número limitado de operaciones aritméticas, como en el caso del cifrado homomórfico completo, o totalmente homomórfico, donde es posible utilizar dos operaciones aritméticas: la suma y la multiplicación. Por otro lado, el cifrado homomórfico parcial únicamente soporta un tipo de operación, ya sea la suma o la multiplicación. Siendo la suma y multiplicación operaciones elementales, cualquier operación aritmética en algoritmos de inteligencia artificial podría implementarse a través de un cifrador homomórfico, aunque con algunas restricciones o con una penalización en la eficiencia.

Dicha eficiencia es afectada directamente por el nivel de seguridad del cifrador homomórfico, el cual, como se ha comentado, está asociado a la dificultad para resolver un problema matemático, la cual está relacionada con el tamaño de los números usados para las operaciones aritméticas de cifrado y descifrado. Para números de

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

2048 bits, como se recomienda actualmente, el impacto del costo de las operaciones del cifrador podría ser muy alto cuando se procesen grandes cantidades de datos.

Oportunidad. El despliegue eficiente de los métodos de cifrado homomórfico puede realizarse haciendo uso de técnicas de paralelismo u otros modelos de aceleración que permitan contar con métodos tanto seguros como eficientes, lo cual, constituye un área de oportunidad de investigación y desarrollo. Por ejemplo, un método de IA con preservación de privacidad de datos puede desplegarse bajo un enfoque hardware/software, donde las partes más demandantes computacionalmente se aceleren con hardware dedicado, o se implementen bajo modelos como el cómputo heterogéneo, el cual explota el poder de cómputo de todos los dispositivos disponibles en la infraestructura de cómputo al mismo tiempo, como núcleos del procesador, GPUs o FPGAs.

Reto 2: Viabilidad y aplicación en un escenario de cómputo cuántico. Los métodos criptográficos homomórficos se definen en estructuras algebraicas donde se establece un problema matemático, como la factorización de enteros o el cálculo de logaritmos discretos. La dificultad para resolver ese problema determina el nivel de seguridad del método criptográfico. Por ello, las estructuras algebraicas deben ser de tamaño suficiente para que el problema sea difícil de resolver. Esta suposición se ha mantenido hasta estos tiempos, ya que no existe un algoritmo eficiente (con complejidad polinomial) que pueda ejecutarse en una computadora convencional, que resuelva el problema en el que el cifrado homomórfico sustenta su seguridad. Sin embargo, ya se ha demostrado que, para el dominio del cómputo cuántico, si existen algoritmos de complejidad polinomial que pueden resolver prácticamente cualquiera de los problemas en los que se sustentan los algoritmos criptográficos actuales. Por ello, los métodos criptográficos homomórficos, sobre los que se construya la protección a la privacidad están amenazados en un escenario postcuántico. De acuerdo con expertos, se prevé que en 2030 se cuenten con computadoras cuánticas capaces de vulnerar cualquier esquema de cifrado actual.

Oportunidad. Desde 2006, se está desarrollado una nueva línea de investigación llamada Criptografía Post-Cuántica [3] (PQC, por sus siglas del Inglés). Esta comprende el desarrollo de algoritmos criptográficos resistentes a ataques de computadoras cuánticas. Actualmente, se han identificado distintas familias de algoritmos PQC, y se ha realizado recientemente trabajo de estandarización de esquemas PQC para cifrado [4], y sobre los cuales se pueden construir sistemas homomórficos, poco explorados aún. En octubre de 2022, concluyó la ronda 3 de evaluación y selección de algoritmos PQC por parte del NIST (*National Institute for Standadars and Technology*), que finalmente recomendó a la criptografía basada en látices como la más adecuada para cifrado de datos. Sin embargo, PQC aún no se ha explorado ampliamente en el contexto del cifrado homomórfico para la minería de datos como servicio. De ahí que ésta sea un área de oportunidad importante de investigación y desarrollo en los próximos años.

Reto 3: Disponibilidad de herramientas. Herramientas disponibles para análisis de datos mediante algoritmos de IA, con preservación de privacidad, aún son escasas. Si fuera posible garantizar la privacidad de los datos, se podrían disponer de herramientas en línea, en la nube, donde los propietarios de los datos puedan subir sus datos, procesarlos y obtener algún modelo de IA de interés (agrupación o clasificación por ejemplo), sin riesgo de que los datos de entrenamiento queden expuestos. Desde el punto de vista de la academia y de la investigación, este tipo de herramientas en la nube también pueden ser usadas para estudio y mejora de métodos de privacidad basados en cifrado homomórfico. Una herramienta como tal podría incluir más métodos que le permitan tener distintos alcances respecto a algoritmos de IA soportados y podría ser configurable para soportar un nivel de seguridad específico. Por ello, para promover la adopción de herramientas de seguridad y privacidad como servicio es necesario construir más herramientas, desde la perspectiva del usuario final y también del científico de minería de datos.

Oportunidad. El cuerpo del conocimiento en relación con el cifrado homomórfico y la criptografía postcuántica, así como de los modelos de aceleración bajo entornos de cómputo en la nube pueden converger para crear modelos de protección de privacidad robustos, eficientes y seguros, así como herramientas flexibles y configurables, que promuevan la adopción de servicios en la nube en los casos de uso más representativos, como algoritmos de agrupamiento o de clasificación.

Conclusiones

La privacidad de datos es un aspecto que ha cobrado relevancia en distintos ámbitos, motivada por la alta disponibilidad y acceso a los mismos, sobre todo cuando dichos datos son sensibles (datos médicos, personales, financieros, etc.). Actualmente, la garantía de privacidad está siendo incorporada en la legislación de cada país, por lo que se debe preservar la privacidad cuando los datos se generan, almacenan, transmiten o *cuando se usan.* Este último caso es el que compete al dominio de los algoritmos de IA, como se ha venido usando a lo largo del artículo y que se usan en minería de datos y en aprendizaje automático. Ante un escenario de cómputo en la nube, es cada vez más común disponer de algoritmos de análisis de datos como servicio. Una herramienta que permite garantizar la privacidad de los datos usados en escenarios como el del cómputo en la nube es la cripto-

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

grafía homomórfica, la cual permite operar con los datos cifrados y no con los datos originales. El resultado obtenido, al descifrarse por el propietario de los datos, será el mismo que se hubiera obtenido usando los datos sin cifrar. De esta forma, se pueden construir herramientas efectivas que salvaguarden la privacidad de los datos.

Sin embargo, no es suficiente que las herramientas en la nube sean efectivas, se requieren que sean eficientes y viables de operar en aplicaciones reales. En este sentido, se requiere aún resolver problemas abiertos en lo que se refiere a eficiencia, seguridad y practicidad de los sistemas homomórficos como servicios. Por ello, los modelos actuales de cifrado homomórfico deben estudiarse usando niveles de seguridad recomendables, sin pérdida de eficiencia y bajo un escenario de cómputo postcuántico, donde los esquemas de cifrado homomórfico basados en criptografía convencional no serán seguros.✷

REFERENCIAS

1. Bartneck, C., Lutge, C., Wagner, A. y Welsh, S. (2021). Privacy Issues of AI. *A*n Introduction to Ethics in Robotics and AI, Springer International Publishing, Cham, 61-70.
2. Almutairi, N., Coenen, F. y Dures, K. (2017). K-Means Clustering Using Homomorphic Encryption and an Updatable Distance Matrix: Secure Third Party Data Clustering with Limited Data Owner Interaction. *B*ig Data Analytics and Knowledge Discovery, DaWaK 2017. Lecture Notes in Computer Science. 10440.
3. Bernstein, D.J. (2009). Introduction to post-quantum cryptography. *P*ost-Quantum Cryptography, Springer Berlin Heidelberg, 1-14.
4. Alagic, G., Alperin-Sheriff, J., Apon, D., Cooper, D., Dang, Q., Kelsey, J., Liu, Y., Miller, C., Moody, D., Peralta, R., Perlner, R., Robinson, A. y Smith-Toneand, D. (2020). Status report on the second round of the NIST post-quantum cryptography standardization process. *T*he National Institute for Standards and Technology (NIST).
5. ur-Rehman, M.H., Yaqoob, I., Salah, K., Imran, M., Jayaraman, P.P. y Perera, C. (2019). The role of big data analytics in industrial Internet of Things. *F*uture Generation Computer Systems, 99, 247-259.
6. Almutairi, N., Coenen, F. y Dures, K. (2021). Secure third-party data clustering using SecureCL, data and multi-user order preserving encryption. *Expert Systems*, 38(7), e12581.
7. Qiu, G., Gui, X. y Zhao, Y. (2020). Privacy-Preserving Linear Regression on distributed data by homomorphic encryption and data masking. *I*EEE Access, 8, 107601-107613.

SOBRE EL AUTOR



Miguel Morales Sandoval es profesor-investigador adscrito a la Unidad Tamaulipas del Cinvestav. Es miembro del Sistema Nacional de Investigadores Nivel II. Sus intereses académicos y de investigación se centran en seguridad de datos. Recibió el doctorado en Ciencias de la Computación en 2008 del Instituto Nacional de Astrofísica, Óptica y Electrónica de México. Actualmente está enfocado en el desarrollo de esquemas de seguridad de datos con aplicaciones en la ciencia de datos, la nube y el Internet de las cosas.



Heidy Marisol Marin Castro es profesora e investigadora por México-CONACYT adscrita a la Facultad de Ingeniería y Ciencias de la Universidad Autónoma de Tamaulipas. Doctora en Ciencias de la Computación por el Centro de Investigación y de Estudios Avanzados del IPN, con intereses de investigación en ciencia de datos con énfasis en minería de procesos y de datos, así como aprendizaje máquina y gestión de información web. Es miembro del Sistema Nacional de Investigadores Nivel 1.



Karina Figueroa Mora es Doctora en Ciencias de la Computación por la Universidad de Chile, Ingeniera Electricista y Maestra en Ingeniería Eléctrica opción Sistemas Computacionales por la Facultad de Ingeniería Eléctrica de la Universidad Michoacana de San Nicolás de Hidalgo. Profesora e investigadora de la Facultad de Ciencias Físico Matemáticas "Mat. Luis Manuel Rivera Gutiérrez". Miembro del Sistema Nacional de Investigadores. Perfil PRODEP. Especialista en recuperación de información, diseño y análisis de algoritmos, bases de datos (de texto y métricas).

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ARTÍCULO ACEPTADO

Detección de notas de ideación suicida

María del Carmen García-Galindo, Ángel Hernández-Castañeda, René Arnulfo García-Hernández, Yulia Ledeneva y Rafael Cruz-Reyes

Hoy en día, el internet y las plataformas de redes sociales han remplazado a los medios de comunicación tradicionales, al permitir que usuarios de todo el mundo compartan de manera instantánea y abierta sus ideas, pensamientos, emociones, opiniones, experiencias, etc; expresándose a través de texto, audio o imagen.

Desafortunadamente, entre los temas discutidos por los jóvenes en las redes sociales se encuentran la ideación suicida, la muerte y los pensamientos autolesivos [1]. Esto indica que los jóvenes prefieren compartir sus pensamientos e ideas suicidas a través de notas en estos canales de comunicación, en lugar de buscar ayuda profesional o compartirlos con sus círculos sociales más cercanos.

Durkee et al. [2] indican que estos medios juegan un doble papel (factor de riesgo o protector) en la ideación suicida. Por un lado, el uso inadecuado de estas plataformas podría permitir a los usuarios acceder a contenidos que promueven el suicidio. Además, las redes sociales facilitan la difusión de pensamientos y pactos suicidas o el acceso a recursos materiales para realizar un intento de suicidio.

Por otro lado, otros estudios afirman que los contenidos de redes sociales pueden ser utilizados como factor protector y preventivo, debido a que permiten identificar patrones que frecuentemente aparecen en textos con contenido suicida [3][4].

La palabra suicidio proviene del latín *sui* que significa 'a si mismo' y *cadere*, que significa 'matar', es decir, muerte producida por uno mismo [5]. Este fenómeno ha existido desde la antigüedad modificando su significado con base a cada cultura.

El suicidio comprende diferentes etapas: la ideación suicida, seguida por la planeación y finalmente el intento (en el peor caso puede resultar en un acto consumado). La ideación suicida se refiere a los pensamientos y deseos intencionales de querer morir; en la planeación se definen los métodos para realizar un intento y finalmente, el intento constituye la conducta potencialmente autolesiva para concluir su vida [6].

La Organización Mundial de la Salud (OMS) estima que alrededor de 703,000 personas deciden quitarse la vida cada año a nivel mundial. Además, por cada suicidio consumado existen más intentos y esto convierte al suicidio en un grave problema de salud pública [1].

Existen muchas razones por las que una persona decide suicidarse, por lo que se han identificado diferentes factores de protección contra la ideación suicida. Uno de ellos es la regulación de emociones, debido a que un balance en las emociones permite disminuir el riesgo de un intento de suicidio [7].

Uno de los métodos de diagnóstico más antiguo para estudiar el suicidio, es el análisis de notas suicidas [8]. En el pasado estas notas eran escritas a mano y su estudio era posterior a la muerte del individuo; sin embargo, con el incremento y la influencia de las redes sociales, estas notas son escritas y compartidas digitalmente. No obstante, monitorear el contenido que se comparte en las plataformas digitales y detectar aquellas notas de ideación suicida de manera manual se hace ineficiente debido a la sobrecarga de información.

Trabajos relacionados

En los últimos años, se han desarrollado estrategias en diferentes áreas de investigación con el objetivo de prevenir y reducir el número de muertes por suicidio. Recientemente, estas estrategias de prevención se han centrado en el análisis del contenido de las redes sociales e internet debido a su influencia en los jóvenes.

En el área computacional, se han utilizado técnicas de aprendizaje automático y Procesamiento del Lenguaje Natural (PLN) para detectar textos con contenido de ideación suicida.

El Procesamiento de Lenguaje Natural es la disciplina computacional que utiliza el lenguaje natural para establecer la comunicación entre las personas y la computadora, de modo que las computadoras pueda entender las instrucciones proporcionadas por el humano [9][10]. Se entiende por lenguaje natural a la forma cotidiana que utilizamos los humanos para comunicarnos.

La mayoría de los enfoques computacionales aplicados en la detección de notas con contenido de ideación suicida se centran en el análisis lingüístico. Este tipo de análisis reconoce si el texto de estudio tiene relación a la tendencia suicida con base en las palabras utilizadas. Por ejemplo, es posible verificar si en el escrito existen palabras tales como: *suicidio*, *muerte*, *plan suicida* o *ideación suicida*. La presencia de estas palabras permiten identificar si el texto de estudio se trata de una nota suicida.

El análisis de emociones en notas que contienen ideas suicidas se ha abordado frecuentemente como una clasi-

[1] Información consultada del sitio oficial de la Organización Mundial de la Salud (https://www.who.int/es/news-room/fact-sheets/detail/suicide) el 12 de enero del 2023

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

ficación binaria, es decir, identifica si el texto contiene cierta emoción o no (ver Figura 1).

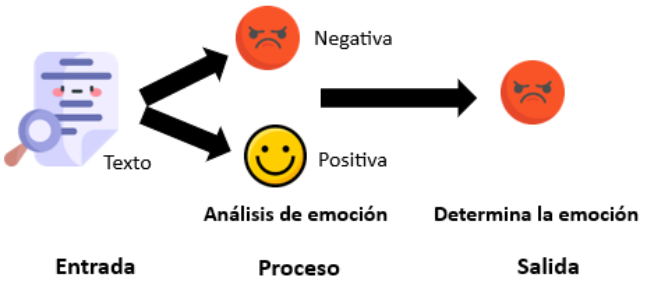


Figura 1. Clasificación binaria de emociones.

Las investigaciones actuales, han alcanzado resultados prometedores en identificar automáticamente notas de suicidio mediante técnicas de PLN y algoritmos de aprendizaje automático; sin embargo, la mayoría de estos estudios aplican clasificación directa. Este tipo de clasificación identifica que las palabras utilizadas en el texto de estudio sean propias de una nota suicida. La presencia de estas palabras determina que el texto es una nota suicida. El flujo general de la clasificación directa puede observarse gráficamente en la Figura 2.

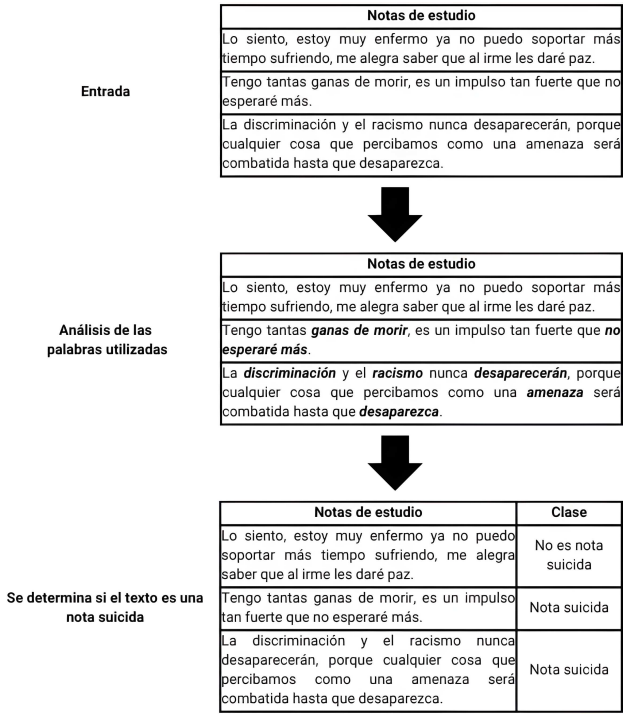


Figura 2. Flujo de la clasificación directa.

En consecuencia, los métodos basados en la clasificación directa pueden no ser útiles en la detección de notas de ideación suicida con contenido no explícito, debido a que algunas de ellas no presentan palabras relacionadas específicamente al suicidio.

A diferencia de los métodos basados en la clasificación directa, en este estudio se propone un método de clasificación en cascada. Este método primero identifica automáticamente la distribución de emociones, tales como: enojo, miedo, felicidad, optimismo, etc., en un texto. Por ejemplo, se enfoca en determinar el grado de felicidad que contiene un texto con base en las palabras que lo forman.

La clasificación en cascada trata de identificar la ideación suicida con respecto al estado emocional de la persona. Dicho estado, en este estudio se determina a través de textos digitales y no directamente del análisis físico-psicológico de la persona.

Por lo tanto, la metodología propuesta consta de dos clasificadores configurados en una arquitectura en cascada. Así, un primer clasificador obtiene la distribución emocional que muestra el grado en el que cada emoción está presente en el texto y después, un segundo clasificador toma la distribución emocional para identificar los textos suicidas. Ver Figura 3.

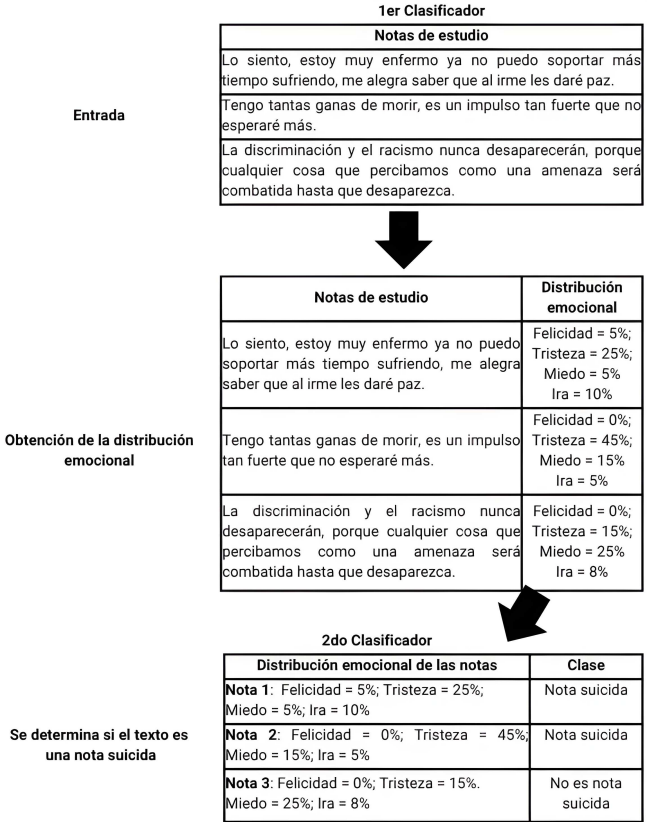


Figura 3. Metodología de la clasificación en cascada.

Datos

Una nota suicida podría permitir conocer el estado emocional, las motivaciones, los pensamientos, los deseos,

etc., de una persona en relación con el momento en que lleva a cabo su muerte.

La mayoría de los enfoques computacionales, desarrollados previamente, se han centrado en detectar automáticamente notas suicidas en idioma inglés.

Con el objetivo de comparar el desempeño del método de clasificación en cascada contra investigaciones previas, en este trabajo se crearon tres conjuntos de datos diferentes en idioma inglés.

Actualmente, el método propuesto analiza notas suicidas escritas en inglés; sin embargo, podría modificarse para estudiar este tipo de textos en cualquier idioma. Para ello, en general se requiere intercambiar los textos que analiza el sistema al lenguaje correspondiente.

Las técnicas de PLN y algoritmos de aprendizaje automático han identificado automáticamente notas de suicidio.

Así, para estudiar notas en otro idioma se necesitan nuevos conjuntos de datos. Por un lado, se requiere un conjunto de emociones y uno de notas suicidas. Por el otro, se requieren textos con dominio diferente al suicidio.

Las colecciones de datos, se utilizan para otorgarle al clasificador automático ejemplos reales de las características lingüísticas que identifican a las notas de ideación suicida.

Los conjuntos de datos creados se detallan a continuación. En la Figura 4, se puede observar un ejemplo resumido en idioma español de cada tipo de nota considerada en este estudio.

- Notas de suicidio consumado: esta colección consta de 155 notas de ideación suicida escritas por personas que consumaron el intento. En muchos de estos casos las personas no expresaron explícitamente sus intenciones. Estas notas se recolectaron de sitios web y periódicos en línea.

  Las textos de este conjunto pueden expresar emociones tales como: dolor, ira, miedo, etc; sin embargo, no en todos los casos se utilizan palabras asociadas al suicidio que indiquen explícitamente el deseo de morir, como puede observarse en la Figura 4.

- Notas de intenciones suicidas: la colección consiste en 453 notas de ideación suicida recolectadas manualmente de la red social Reddit de los canales: SuicideWatch y Depression. Para crear este conjunto de datos, se recolectaron notas de intenciones suicidas que presentaban términos y/o frases relacionadas al acto suicida, por ejemplo: *suicidio*, *ideación suicida*, *intento de suicidio*, *muerte*, *intento fallido*, etc.

En la red social Reddit las notas suicidas son bastante explícitas, marcando una clara diferencia entre las personas que expresan explícitamente su intención de morir, frente a otras que tienen el mismo deseo, pero no lo expresan abiertamente. Sin embargo, con estos textos se desconoce si el acto se consumó.

- El último conjunto de datos creado fue el de notas de otros dominios distintos al suicidio: esta colección consta de 255 notas que abarcan seis temas diferentes (cultura, economía, política, música, comida y deportes). Estas se recolectaron de la red social Reddit.

El conjunto de datos de dominios distintos al suicidio se consideró con el objetivo de tener contraejemplos de lo que no es una ideación suicida y para darle al clasificador el ejemplo de lo que no debe considerar como una nota de suicidio.

Clase	Nota
Suicidio consumado	Lo siento, estoy muy enfermo ya no puedo soportar más tiempo sufriendo, me alegra saber que al irme les daré paz.
Intención suicida	Tengo tantas ganas de morir, es un impulso tan fuerte que no esperaré más.
Otro dominio	La discriminación y el racismo nunca desaparecerán, porque cualquier cosa que percibamos como una amenaza para nosotros será combatida hasta que desaparezca.
Declaración final	Adiós. Te esperaré en el Cielo. Te estaré esperando. Estoy en paz con Dios. Lucha por el bien.
Emociones	A veces este genio se oscurece y se hunde en el pozo amargo de su corazón.

Figura 4. Ejemplos de las notas de estudio.

Para un análisis más profundo de las emociones implicadas en casos previos de muerte, se utilizó un conjunto de 453 notas de las declaraciones finales de presos sentenciados a muerte en el estado de Texas. El conjunto de datos está disponible en el sitio web oficial del Departamento de Justicia de Texas.[2]

[2] https://www.tdcj.texas.gov/death_row/dr_executed_offenders.html

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Este conjunto de notas se consideró para analizar las diferencias lingüísticas y emocionales de las personas que se quitan la vida, en comparación con las personas que han sido condenadas, las cuales no precisamente tienen el deseo de morir.

Uno de los pasos fundamentales de nuestra propuesta requiere proporcionar la distribución emocional que presenta cada texto de estudio. Para ello, se utilizó el CEASE (Corpus of Emotion Annotated Suicide notes in English) que consta de 1,001 frases, etiquetadas en 13 emociones.

Resultados

Para clasificar notas de ideación suicida, mediante el enfoque en cascada propuesto, se realizaron experimentos considerando diferentes niveles del lenguaje (léxico y semántico). En el nivel léxico, se analizó la importancia de las palabras en una colección de textos. A nivel semántico, se estudio el significado de la palabras en un texto considerando el contexto en el que se utilizaron.

Se observó que ambos enfoques de clasificación (directa y en cascada) obtienen un buen rendimiento, al clasificar notas de intención suicida contra escritos de personas sentenciadas a muerte. La clasificación directa logró un desempeño de 97.9 % de precisión. Mientras que la clasificación en cascada logró un desempeño de 96.57 % de precisión.

Esto se debe a que el vocabulario entre ambas clases es bastante variado, debido a que como se mencionó anteriormente las notas de intenciones suicidas obtenidas de Reddit son bastante explicitas, haciendo referencia al suicidio, a la muerte y a la vida. Por el contrario, si analizamos las palabras más frecuentes de los escritos de personas sentenciadas a muerte, no se encuentran palabras asociadas al acto de muerte. En este caso, se usan palabras más comunes como familia, agradecimiento, perdón, amor. Por tanto, la diferencia entre el vocabulario hizo que ambos tipos de clasificaciones pudieran diferenciar y clasificar eficientemente los textos obteniendo un buen desempeño.

Por el contrario, al clasificar notas de suicidio consumado y notas de no suicidio, se percibió que el vocabulario entre ambas no difiere mucho. Esto se debe a que las personas que se suicidan en muchos casos no utilizan palabras explícitas relacionadas al tema de muerte.

La clasificación entre las notas de suicidio consumado y notas de no suicidio mostró que el rendimiento del enfoque directo disminuye cuando no se tienen presentes palabras explícitas que permitan caracterizar a una de las dos clases. Por el contrario, el enfoque en cascada, que analiza primero la distribución emocional, obtiene un mejor resultado que el enfoque directo porque no depende de las palabras referentes al suicidio. En este caso, el enfoque de clasificación directa logró un desempeño de 90.85 % de precisión, mientras que el enfoque de clasificación en cascada logró un desempeño de 94.84 % de precisión.

Conclusiones

La tarea de la detección de notas de ideación suicida en contenidos de redes sociales es de gran importancia debido a que algunas personas con dicha tendencia prefieren recurrir a los medios digitales, en vez de compartirlo abiertamente con sus círculos sociales.

En muchos casos, la clasificación directa no basta para detectar notas de ideación suicida. En esta investigación, se desarrolló un método de clasificación en cascada, el cual superó el desempeño de la clasificación directa.

El método de clasificación en cascada resulta ser más robusto debido a que no depende exclusivamente de las palabras que refieren al suicidio. En cambio, aprovecha la distribución emocional y posteriormente la utiliza para determinar si el texto está relacionado con la ideación suicida.✷

REFERENCIAS

1. Abdulsalam, A. y Alhothali, A. (2022). Suicidal Ideation Detection on Social Media: A Review of Machine Learning Methods. arXiv preprint arXiv:2201.10515, 1-14.
2. Durkee, T., Hadlaczky, G., Westerlund, M. y Carli, V. (2011). Internet pathways in suicidality: A review of the evidence. *International Journal of Environmental Research and Public Health*, 8(10), 3938–3952.
3. Parrott, S., Britt, B. C., Hayes, J. L. y Albright, D. L. (2020). Social media and suicide: A validation of terms to help identify suicide-related social media posts. *Journal of evidence-based social work*, 17(5), 624–634.
4. Dieu, F. D. (2018). Cibersuicidio en adolescentes. Trabajo final de grado: Monografía, Licenciado en Psicología. Universidad de la República, Uruguay.
5. Jiménez-Treviño, L., Sáiz Martínez, P.A. y Bobes García, J. (2006). Suicidio y depresión. *Humánitas, Humanidades médicas*, 09, 1-21.
6. Ortega-González, M. (2018). Comportamiento suicida. Reflexiones críticas para su estudio desde un sistema psicológico. Primera edición. Qartuppi, S. de R.L. de C.V. Hermosillo, Sonora 83220 México.
7. Cha, C.B. y Nock, M.K. (2009). Emotional intelligence is a protective factor for suicidal behavior. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(4), 422-430.
8. Hernández, A., Macías-García, L.F. y Luna-Lara, M.G. (2011). Notas suicidas mexicanas. Un análisis cualitativo. *Pensamiento Psicológico*, 9(17), 33-42.
9. Gelbukh, A. (2010). Procesamiento de lenguaje natural y sus aplicaciones. *Komputer Sapiens*, Sociedad Mexicana de Inteligencia Artificial, 1, 6-11.
10. Cortez-Vásquez, A., Vega huerta, H., Pariona-Quispe, J. y Huayna, A.M. (2009). Procesamiento de lenguaje natural. *Revista de investigación de Sistemas e Informática*, 6(2), 45-54.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

SOBRE LOS AUTORES



María del Carmen García Galindo Recibió el título de Ingeniera en Software por la Universidad Autónoma del Estado de México, en 2021. Actualmente es estudiante de la Maestría en Ciencias Computacionales en la misma universidad. Sus intereses de investigación incluyen el procesamiento del lenguaje natural, el aprendizaje automático, el reconocimiento de patrones y minería de datos.



Ángel Hernández Castañeda recibió su Maestría y Doctorado en Ciencias de la Computación, con honores, por el Centro de Investigación en Computación (CIC) del Instituto Politécnico Nacional (IPN), en 2013 y 2017, respectivamente. Asimismo, es autor de diversas publicaciones en revistas internacionales de alto impacto. Actualmente es Profesor investigador de Tiempo Completo en la Universidad Autónoma del Estado de México y miembro del Sistema Nacional de Investigadores (SNI) de México. Sus intereses de investigación incluyen el procesamiento del lenguaje natural, la minería de datos, el aprendizaje automático y el reconocimiento de patrones.



René Arnulfo García Hernández recibió el título de Licenciado en Ingeniería en Sistemas Computacionales del Instituto Tecnológico de Toluca, México, en 2001; la Maestría en Ciencias de la Computación del Centro Nacional de Investigación y Desarrollo Tecnológico, México, en 2003; y el Doctorado en Ciencias de la Computación por el Instituto Nacional de Astrofísica, Óptica y Electrónica, México, en 2017. Es autor de más de 67 publicaciones y 3 libros. Es Investigador Nacional de Nivel II y Profesor de Tiempo Completo en la Facultad de Ingeniería de Software y del Postgrado de la Universidad Autónoma del Estado de México.



Yulia Ledeneva recibió grado de Maestría en Ciencias de la Computación del Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), México, en 2016; el Doctorado en Ciencias de la Computación del Centro de Investigación en Computación, IPN, México. Recibió la presea Lázaro Cárdenas (2009) de manos del presidente de México. Actualmente es profesora investigadora en la Universidad Autónoma del Estado de México y miembro del Sistema Nacional de Investigadores (SNI) de México. Es autora de más de 70 publicaciones. Sus principales intereses de investigación son: lingüística computacional, procesamiento del lenguaje natural, minería de texto, gráficos y algoritmos genéticos.



Rafael Cruz-Reyes MSc por el IT Toluca, México; Ingeniero en Comunicaciones y Electrónica por el Instituto Politécnico Nacional; Certificado en Diversas especialidades en Electrónica Aeronáutica por el Centro de Adiestramiento de Aviación Civil Internacional de OACI y la SCT México. Especializado en Gestión académica en comités de certificación de calidad educativa. Actualmente es académico en el Laboratorio de Inteligencia Artificial de la Universidad Autónoma del Estado de México. Sus intereses son la Inteligencia Artificial, Procesamiento de Lenguaje Natural y la aplicación de la Metodología de Investigación estructural para la Gestión de proyectos de Investigación en Ciencias computacionales.



Imagen tomada de:
https://consultorsalud.com/ia-inteligencia-artificia-riesgo-suicidio/.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

IA & Educación

Julieta Noguez, Lucía Barrón y Yasmín Hernández

El procesamiento de Lenguaje Natural y la Inteligencia Artificial ¿Abrimos la caja de pandora?

En noviembre de 2022, la empresa OpenAI lanzó al mercado la herramienta *ChatGPT* acaparando la atención de múltiples sectores de la sociedad, especialmente del sector científico. La empresa presenta a *ChatGPT* como un proceso de optimización de modelos de lenguaje para el diálogo, el cual es capaz de interactuar con el usuario a través de conversación y asegura que "El formato de diálogo hace posible que *ChatGPT* responda preguntas de seguimiento, admita sus errores, cuestione premisas incorrectas y rechace solicitudes inapropiadas" [1].

OpenAI ha desarrollado otros modelos de lenguaje como *InstructGPT* [2] que proporciona respuestas detalladas acordes a una consulta (promot) del usuario, este modelo ha servido como base para la creación de *ChatGPT*. El desarrollo de *InstructGPT* y *ChatGPT* está basado en una técnica llamada Aprendizaje por Refuerzo con Retroalimentación Humana (*Reinforcement Learning from Human Feedback*, *RLHF*, por sus siglas en inglés). *OpenIA* resume el proceso de entrenamiento del modelo conversacional en tres pasos:

1) Recopilar datos de demostración y ajuste fino supervisado (SFT) del modelo inicial: seleccionaron GPT-3.5 como modelo inicial y un grupo de entrenadores humanos de Inteligencia Artificial (IA) se encargó de generar respuestas para un conjunto de consultas (prompt), simulando las consultas del usuario y las respuestas que debería proporcionar el asistente de IA. El conjunto de datos generado se mezcló con el conjunto de datos de InstructGPT transformándolo en formato de diálogo.

2) Recolectar datos comparativos y entrenar el modelo de recompensa: usando el conjunto de datos generado en el paso 1, se seleccionaron consultas (prompt) y para cada una de ellas se usaron varias respuestas generadas con el modelo SFT; los etiquetadores clasificaron cada una de las respuestas generando un nuevo conjunto de datos etiquetado de acuerdo con la calidad de la respuesta. Este nuevo conjunto de datos fue utilizado para entrenar el modelo de recompensa.

3) Optimizar el modelo de recompensa utilizando el algoritmo de aprendizaje por refuerzo: se realizaron varias iteraciones para ajustar el modelo usando la optimización de política proximal.

Lo anterior puede visualizarse en la Figura 1.

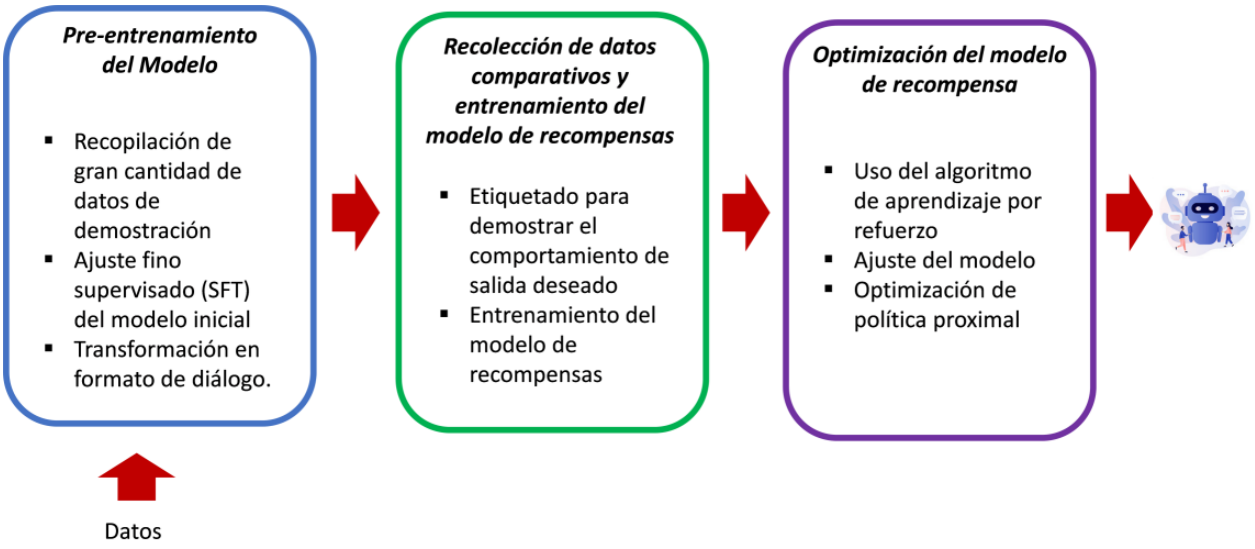


Figura 1. Principales procesos de entrenamiento del modelo conversacional de *ChatGPT*.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

El resultado fue una herramienta que permite al usuario una experiencia más cercana a una conversación humana, que además puede generar textos en diferentes estilos, traducir a varios idiomas, responder preguntas y entregar al usuario desde respuestas simples hasta la generación de textos largos y poéticos. También puede ayudar con tareas como redacción, investigación y servicio al cliente [3].

EDEM Tecnología [4] realizó algunas consultas al propio *ChatGPT*, obteniendo algunas de las siguientes repuestas: "Chat GPT se ha entrenado para realizar una amplia variedad de tareas relacionadas con el lenguaje natural. Esto lo convierte en una herramienta muy útil para diversas aplicaciones, como la generación automática de respuestas en un chatbot o la mejora de la precisión en los sistemas de búsqueda de información".

Se identifican cuatro puntos clave en los que el modelo de lenguaje *ChatGPT* puede ser utilizado con éxito:

- *Generación de texto:* el modelo puede ser utilizado para generar texto coherente y natural, ya sea en forma de historias, artículos o respuestas a preguntas.
- *Mejora de la precisión en los sistemas de búsqueda:* el modelo puede ayudar a mejorar la precisión en los sistemas de búsqueda de información, ya que puede comprender el contexto y la intención detrás de las consultas de los usuarios.
- *Desarrollo de chatbots:* el modelo puede ser utilizado para desarrollar chatbots que puedan mantener conversaciones naturales con los usuarios, respondiendo de manera coherente y precisa a sus preguntas.
- *Mejora del procesamiento del lenguaje natural:* el modelo puede ser utilizado para mejorar el procesamiento del lenguaje natural en diversas aplicaciones, como la traducción automática o la detección de sentimientos en el texto.

Sin embargo, la sorprendente capacidad para brindar respuestas precisas y coherentes en una simulación casi fiel de textos escritos por humanos también ha generado preocupación sobre posibles usos inapropiados de la IA. Por ejemplo, resulta inquietante que, según *OpenAI*, *ChatGPT* puede responder a instrucciones dañinas o exhibir un comportamiento sesgado. Por ejemplo, a solicitud del usuario, el bot podría crear en poco tiempo un borrador de programa maligno (software malicioso que generalmente se usa para ataques de piratería) o bien podría generar noticias falsas. Según el desarrollador de esta tecnología, se están haciendo esfuerzos para que el chatbot rechace este tipo de solicitudes inapropiadas.

En el ámbito de la educación, el fraude en trabajos académicos es otro punto de gran preocupación, debido a que en la medida en que la herramienta aprenda y se perfeccione, será muy difícil evaluar sí los textos que entregan los estudiantes son de autoría propia o provienen de una herramienta de IA. Para esto, la misma empresa *OpenIA* esta preparando un nuevo clasificador que sea capaz de indicar si un texto fue escrito por una herramienta de IA [5]. OpenIA considera importante reconocer los límites y el impacto en el aula de los generadores de texto que usan IA y se encuentra desarrollando un recurso para educadores sobre el uso de *ChatGPT*.✷

REFERENCIAS

1. ChatGPT: Optimizing Language Models for Dialogue. (2023). http://openai.com/blog/chatgpt (Consultado el 24 de enero de 2023).
2. L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, et al. (2022). Training language models to follow instructions with human feedback. *arXiv preprint arXiv:2203.02155*.
3. National Geographic. (2023). CIENCIA. ¿Qué es ChatGPT y para qué sirve? https://www.nationalgeographicla.com/ciencia/2023/01/que-es-chatgpt-y-para-que-sirve. (Consultado el 25 de enero de 2023)
4. Ortiz, P. (2023). CHAT GPT: QUÉ ES, PARA QUÉ SIRVE Y SU APLICACIÓN EN LA ECONOMÍA [EXPLICADO POR CHAT GPT]. EDEM Tecnología. https://edem.eu/chat-gpt-que-es-para-que-sirve-y-su-aplicacion-en-la-economia-explicado-por-chat-gpt/. (Consultado el 26 de enero de 2023)
5. New AI classifier for indicating AI-written text. https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text/. (Consultado el 31 de enero de 2023)

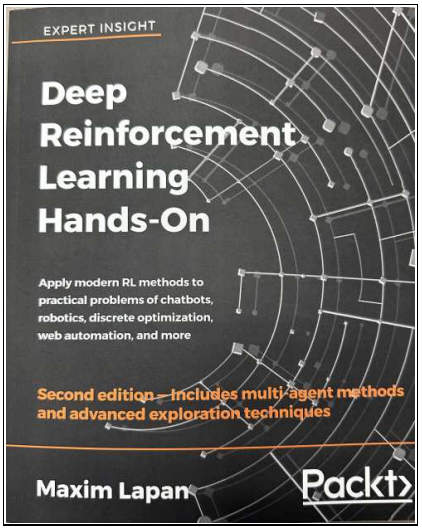
© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Deskubriendo Konocimiento

Alejandro Guerra Hernández y **Leonardo Garrido**
deskubriendokonocimiento@komputersapiens.org

Revisión de "Deep Reinforcement Learning Hands-On, 2^{nd} Edition"

Dr. Gildardo Sánchez-Ante
Tecnológico de Monterrey
Escuela de Ingeniería y Ciencias



Portada del libro.

Sin duda alguna que el aprendizaje computacional ha generado un gran interés en los últimos años, debido en buena medida a la cantidad y variedad de problemas a los que se ha aplicado. Es muy probable que un buen número de esas aplicaciones conciernan al enfoque del aprendizaje supervisado. Este tipo de aprendizaje requiere de casos resueltos para entrenar modelos. Sin embargo, existen otros dos tipos de aprendizaje: el aprendizaje no supervisado y el aprendizaje por refuerzo (*reinforcement learning* (RL), en inglés). Este último quizá sea el menos conocido, pero es una estrategia poderosa para resolver cierto tipo de problemas. La idea central en este paradigma es lograr que un agente aprenda a lograr un objetivo a través de escenarios de prueba y error. Normalmente se trata de situaciones del tipo de juegos en los que ese agente recibe recompensas o penalizaciones dependiendo de su desempeño. El aprendizaje por refuerzo suele utilizarse en entornos inciertos y complejos en los que resulta complicado obtener ejemplos resueltos como los que se usan en el aprendizaje supervisado. Una de las referencias obligadas en este tema es el famoso libro de Sutton y Barto[*], pero la primera edición de ese libro, que fue publicado a finales de los años noventa se concentra fundamentalmente en la teoría subyacente al aprendizaje por refuerzo. Ya hablaremos más adelante de la edición reciente del mismo. En esta ocasión dedicaremos este espacio para comentar sobre otro texto, que además resulta doblemente interesante pues no solo trata los temas del aprendizaje por refuerzo, sino que contempla su interacción con otro tema relevante y reciente como es el aprendizaje profundo o "*deep learning*".

El texto que en esta edición nos ocupa se titula "Deep Reinforcement Learning Hands On"[**]. Desde la portada, Maxim Lapan promete al lector que éste podrá aplicar los métodos del aprendizaje por refuerzo a problemas prácticos de chatbots, robótica, optimización discreta, y automatización web y otros. El libro en cuestión es la segunda edición, publicada en 2020 que extiende la primera edición en varios aspectos. Incluye seis nuevos capítulos para un total de 25 capítulos a lo largo de los cuales el autor explica los conceptos teóricos y ofrece laboratorios y desafíos prácticos que se resuelven con programación. Para ello, pone a disposición del lector un conjunto de códigos base, que puede obtenerse desde GitHub. Pudiera decirse que el libro presenta un buen balance entre la teoría y la práctica. Por un lado, cada capítulo describe la teoría usando las matemáticas adecuadas para entender el principio básico, sin ahondar en la explicación de demasiados detalles matemáticos que pueden ocasionar una lectura más densa. Por el otro lado, las implementaciones prácticas en código son una excelente oportunidad para reafirmar el conocimiento y brindan una retroalimentación casi inmediata de que lo que ese está aprendiendo se puede utilizar para resolver problemas reales. El código de apoyo al libro está desarrollado en Python, utilizando librerías como NumPy, PyTorch, PyTorch Ignite y PTAN, con

[*]Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
[**]Lapan, M. (2020). Deep Reinforcement Learning Hands-On. Birmingham, 2nd Ed., UK: Packt Publishing. ISBN: 978-1838826994.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

lo cual el lector puede desarrollar sus experimentos ya sea localmente en su propio computador, o en algún servicio en la nube, como por ejemplo Google Colab. Es de especial mención el uso de la librería Gym, de OpenAI. En la mayoría de los problemas de aprendizaje por refuerzo se requiere que el agente (o agentes) puedan interactuar con un entorno, lo que permite obtener recompensas para cada acción y retornos para episodios completos de interacción. Esos valores se utilizan para definir la política (estrategia) óptima para resolver el problema en cuestión. Gym provee de herramientas muy valiosas para definir un entorno. El capítulo 2 del libro ofrece una excelente introducción al uso de esta herramienta, y el capítulo 3 se enfoca a conceptos de aprendizaje profundo mediante PyTorch. Con los conceptos básicos del aprendizaje por refuerzo y de las herramientas de cómputo apropiadas para empezar a probar, los capítulos 5 a 8 presentan redes profundas Q (DQN por sus siglas en inglés). En el capítulo 9 se discuten formas de acelerar los métodos de aprendizaje por refuerzo, para en el capítulo 10 desarrollar un primer proyecto práctico completo relacionado con el mercado de acciones. El capítulo 11 se dedica a estudiar gradientes de políticas, en los capítulos 12 y 13 se estudia el método actor-crítico, para luego en el capítulo 14 desarrollar un segundo proyecto práctico orientado al abordaje de problemas de procesamiento de lenguaje natural. El capítulo 15 cubre la aplicación del RL en juegos interactivos. El capítulo 16 trata sobre otra aplicación relacionada con la navegación web. Los capítulos 17 y 19 toca el tema de espacios de acción continuos. El capítulo 18 cubre la aplicación de RL en problemas de robótica, que incluye la construcción de un pequeño dispositivo. El capítulo 20 trata de optimización, usando métodos diferentes de los gradientes. El capítulo 21 se dedica al estudio de varios métodos de exploración. Luego, en el capítulo 22 se presenta el enfoque libre de modelos. El capítulo 23 estudia AlphaGo Zero. El capítulo 24 por su parte trata sobre la aplicación de los métodos de RL a dominios de la optimización discreta, tomando como pretexto el cubo de Rubik como entorno. Finalmente, el capítulo 25 trata de la aplicación de RL en sistemas multi-agentes, un tema reciente y de gran relevancia.

Es así que en estas casi 800 páginas Lapan lleva al lector por un recorrido bastante amplio de métodos y problemas del aprendizaje por refuerzo. La secuencia de libro ayuda a que el conocimiento adquirido sea reforzado en la práctica y va llevando al lector desde escenarios relativamente simples a otros más complejos. Sin duda es un texto que es recomendable para formar parte de cualquier persona interesada en estos temas.✷

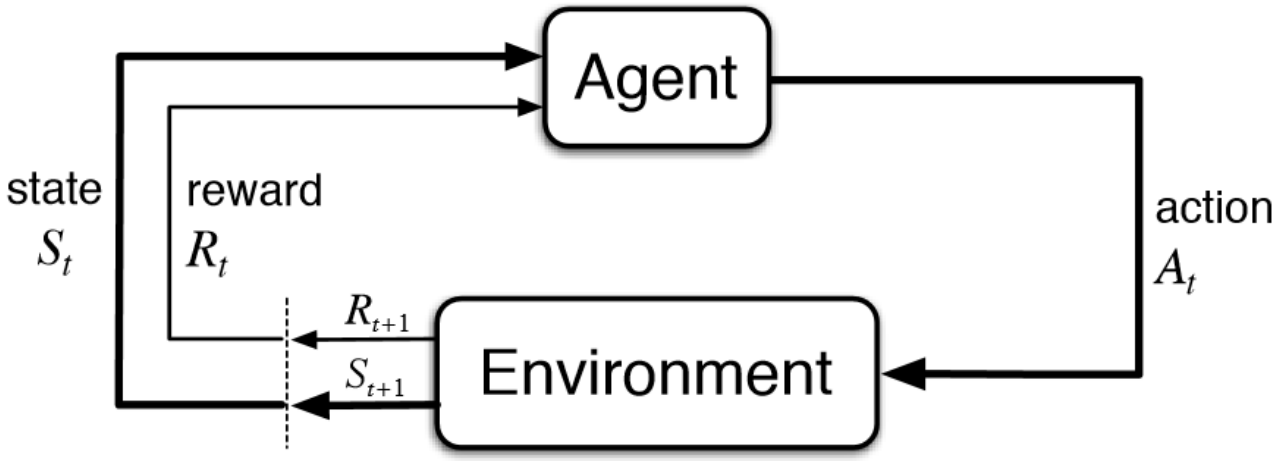


Imagen tomada de:
https://towardsdatascience.com/introduction-to-various-reinforcement-learning-algorithms-i-q-learning-sarsa-dqn-ddpg-72a5e0cb6287.

© 2023 - Sociedad Mexicana de Inteligencia Artificial ISSN 2007-0691

Si te cuidas **tú**, nos cuidamos **todos**

Todos podemos hacer algo para mitigar y contener la expansión de las enfermedades. Lo primero siempre es entender qué son, cómo se transmiten y cómo afectan.

En este caso, se trata del coronavirus SARS-COV2. Apareció en China en diciembre pasado y provoca una enfermedad llamada COVID-19, que se ha extendido por el mundo y fue declarada pandemia global por la Organización Mundial de la Salud.

Medidas de prevención



LAVA TUS MANOS CON AGUA Y JABÓN



ACUDE AL MÉDICO SOLO EN UNA URGENCIA



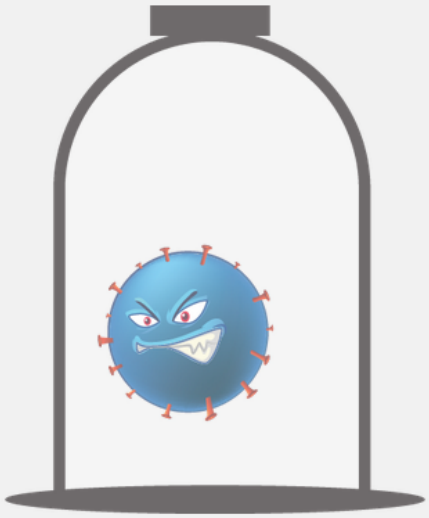
TOSE O ESTORNUDA EN LA PARTE INTERNA DE TU CODO



NO DIFUNDAS NOTICIAS FALSAS

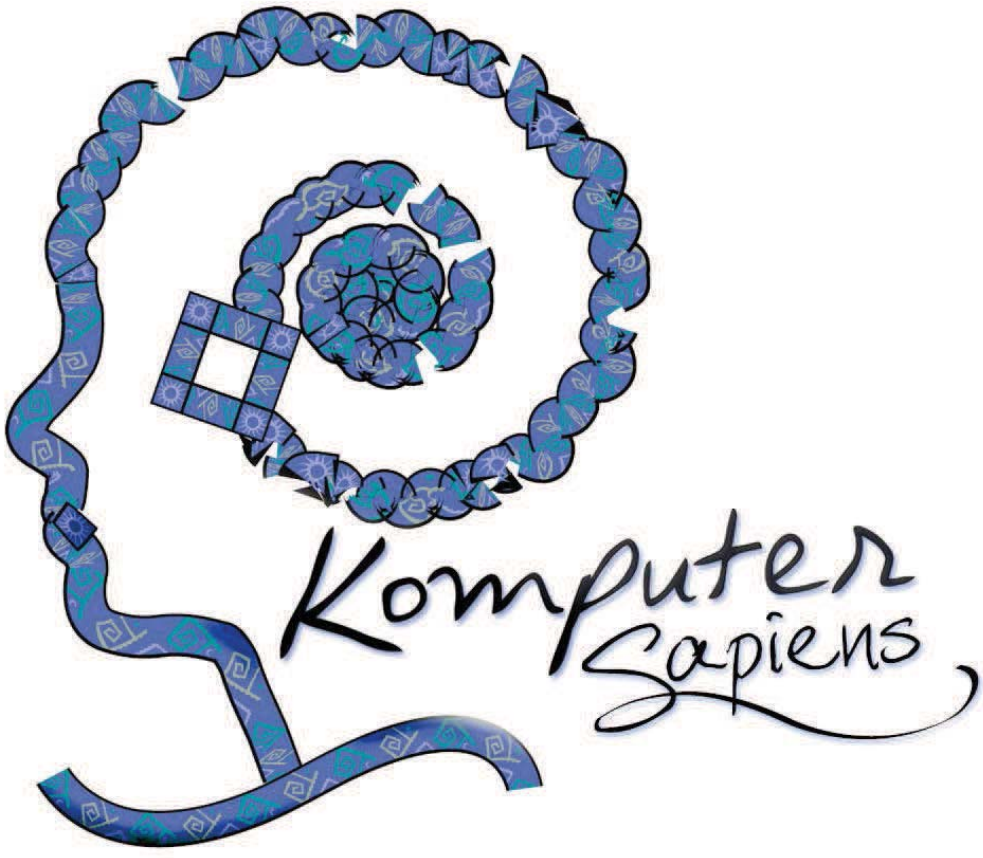
Atendiendo las normas de la Nueva Normalidad y cuidándonos unos a otros podremos salir adelante.

¡SI AISLAMOS AL VIRUS, LO VENCEREMOS!



Información tomada de: https://coronavirus.gob.mx

Indizada en el IRMDCT de CONACYT y en Latindex



¡Publique en Komputer Sapiens!

CONACYT
ÍNDICE DE REVISTAS MEXICANAS DE DIVULGACIÓN CIENTÍFICA Y TECNOLÓGICA (IRMDCT)

Komputer Sapiens solicita artículos de divulgación en todos los temas de Inteligencia Artificial, dirigidos a un amplio público conformado por estudiantes, académicos, empresarios, tomadores de decisiones y consultores. Komputer Sapiens es patrocinada por la **SMIA, la Sociedad Mexicana de Inteligencia Artificial**

SMIA
www.smia.org.mx

Instrucciones para autores e información general: http://www.komputersapiens.org
Síguenos en las redes sociales: www.facebook.com/Komputer.Sapiens, twitter.com/KomputerSapiens