



KOMPUTER SAPIENS

Revista de Divulgación de la Sociedad Mexicana de

Año 13
Volumen 1
Enero-Abril 2021

Aplicaciones de la **Inteligencia Artificial**

Aprende en el mundo de **CARLA**

Algoritmo Memético para el
Problema de Ventas por Internet
con Costos de Envío (**MAIShOP**)

La Realidad Extendida como
herramienta clave para el futuro

Análisis espacial de la **red
de transporte metropolitano**
de Tijuana



ISSN 2007-0691



Análisis de
**sentimientos y
emociones** en
redes sociales

Heurísticas de Selección
para **Programación Genética**
Semántica para resolver
Problemas de Clasificación

Un **clasificador basado**
en transformaciones
de espacio y selección
de modelos

Redes Neuronales
Convolucionales para
Piloteo Autónomo
en Carrera de Drones

Detección de
emociones
en los medio
sociales



©Komputer Sapiens, Año XIII Volumen I, enero-abril 2021, 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 15 de febrero de 2021.

Reserva de derechos al uso exclusivo número 04-2009-111110040200-102 otorgado por el Instituto Nacional de Derechos de Autor. ISSN 2007-0691.

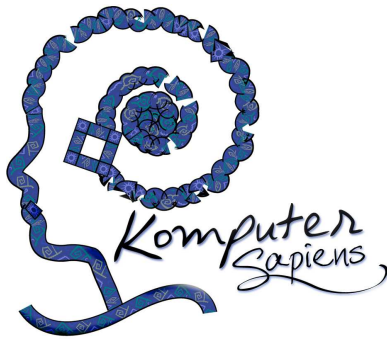
Los artículos y columnas firmados son responsabilidad exclusiva de los autores y no reflejan necesariamente los puntos de vista de la Sociedad Mexicana de Inteligencia Artificial. La mención de empresas o productos específicos en las páginas de Komputer Sapiens no implica su respaldo por la Sociedad 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.

Presidente	Directorio SMIA	Directores Fundadores
Vicepresidente	Félix A. Castro Espinoza	Carlos Alberto Reyes García
Secretario	Ildar Batyrshin	Ángel Kuri Morales
Tesorero	Oscar Herrera Alcántara	
Vocales:	María de Lourdes G. Martínez Villaseñor	Comité Editorial
	Sofía Natalia Galicia Haro	Félix A. Castro Espinoza
	Sabino Miranda Jiménez	Jesús Favela Vara
	Roberto A. Vázquez Espinoza de los Monteros	Sofía Natalia Galicia Haro
	Hiram Erendin Ponce Espinosa	Miguel González Mendoza
	Francisco Viveros-Jiménez	Oscar Herrera Alcántara
	Noé Alejandro Castro Sánchez	Raúl Monroy Borja
	Obdulio Pichardo Lagunas	Eduardo F. Morales Manzanares
	Omar Montaña Rivas	Leonardo Garrido Luna
	Antonio Marín Hernández	Carlos Alberto Reyes García
	Néstor Velasco Bermeo	Angélica Muñoz Meléndez
	Iris Iddaly Méndez Gurrola	Antonio Sánchez Aguilar
	Gustavo Arroyo Figueroa	Luis Enrique Sucar Succar
	Leobardo Morales Tiburcio	Ángel Kuri Morales
	Iván Olvera Rodríguez	José A. Martínez Flores
	Juan Martínez Miranda	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
		Árbitros
		Efren Mezura Montes
		Héctor Rodríguez-Rangel
		Marcela Quiroz Castellanos
		Julietta Noguez Monroy
		Ramón Zatarain Cabada
		Raúl Oramas Bustillos
		Daniela Moctezuma
		Gilberto Ochoa-Ruiz
		Karina Mariela Figueroa Mora
		María Auxilio Medina Nieto
Director general	Komputer Sapiens	
Editora en jefe	Félix A. Castro Espinoza	
Editor científico	Karina Mariela Figueroa Mora	
Editores asociados	Héctor Gabriel Acosta Mesa	
	Marco Antonio Aguirre Lam	
	Laura Cruz Reyes	
Editores invitados	María Lucía Barrón Estrada	
	José Antonio Camarena Ibarrola	
	María Yasmín Hernández Pérez	
Coordinadora de producción	Viridiana Mena Gómez	
Estado del IArte	Reyna Carolina Medina Ramírez	
	Jorge Rafael Gutiérrez Pulido	
Sakbe	Claudia G. Gómez Santillán	
	Laura Cruz Reyes	
	Laura Gómez Cruz	
IA & Educación	María Yasmín Hernández Pérez	
	María Lucía Barrón Estrada	
	J. Julieta Noguez Monroy	
Deskubriendo Konocimiento	Alejandro Guerra Hernández	
	Leonardo Garrido Luna	
Asistencia técnica	Alan G. Aguirre Lam	
Corrección de estilo	Gilberto Rivera Zárate	
	Marcela Quiroz Castellanos	
	Ruth Esmeralda Barreda Guajardo	
	Sonia Lilia Mestizo Gutiérrez	
	Guillermo de Jesús Hoyos Rivera	
	Susana Aurora Velasco Montiel	
	Nicandro Cruz Ramírez	
	Claudia Leticia Díaz González	
	Leonardo Romero Muñoz	
	Héctor Joaquín Fraire Huacuja	
	Nelson Rangel Valdez	
Edición de imagen	Laura Gómez Cruz	
Portada	Magdaleno Granados Olvera, El Pez Gordo Design	



Contenido

Algoritmo Memético para el Problema de Ventas por Internet con Costos de Envío (MAIShOP)

Pág. 6 ⇒ por Miguel A. García Morales, Héctor J. Fraire Huacuja, Laura Cruz Reyes, Claudia G. Gómez Santillán, Nelson Rangel Valdez y Ma. Lucila Morales Rodríguez.

Aprende en el mundo de CARLA

Pág. 10 ⇒ por Pablo Isaac Macias-Huerta, Rafael Cabañas-Rocha, Carlos Natanael Lecona Valdespino y Guillermo Santamaría-Bonfil.

Análisis espacial de la red de transporte metropolitano de Tijuana

Pág. 15 ⇒ por Bélgica J. Valencia García y Adán Hiraes-Carbajal.

La Realidad Extendida como herramienta clave para el futuro

Pág. 20 ⇒ por Kathy Castro.

Heurísticas de Selección para Programación Genética Semántica para resolver Problemas de Clasificación

Pág. 24 ⇒ por Claudia N. Sánchez y Mario Graff.

Un clasificador basado en transformaciones de espacio y selección de modelos

Pág. 26 ⇒ por José Ortiz-Bejar, Eric S. Téllez y Mario Graff.

Redes Neuronales Convolucionales para Pilotaje Autónomo en Carrera de Drones

Pág. 29 ⇒ por Leticia Oyuki Rojas Pérez y José Martínez Carranza.

Análisis de Sentimientos y Emociones en Redes Sociales

Pág. 32 ⇒ por Navonil Majumder, Alexander Gelbukh y Soujanya Poria.

Detección de emociones en los medios sociales

Pág. 34 ⇒ por Muhammad Hammad Fahim Siddiqui y Alexander Gelbukh.

Columnas

Sapiens Piensa.
Editorial [pág. 2](#)

Estado del IArte [pág. 4](#)

Sakbe [pág. 5](#)

IA & Educación [pág. 36](#)

Deskubriendo
Konocimiento [pág. 38](#)

Sapiens Piensa

M. Lucía Barrón E., José A. Camarena I., Karina Figueroa y Yasmín Hernández

Este número presenta temas de vanguardia relacionados directamente a problemas con los que vivimos día a día, por ejemplo, la red de transporte público en una ciudad congestionada, la optimización de compras por internet, la interacción entre el usuario y las plataformas e-learning y cómo la realidad virtual y extendida han permeado en casi todos los quehaceres cotidianos de este mundo.

Miguel García et al. nos muestran en *Algoritmo memético para el problema de ventas por internet con costo de envío* una propuesta de solución al problema de optimización de compras por internet. El comercio electrónico continúa creciendo, y la selección de productos no es una tarea trivial cuando se quiere optimizar las compras por internet de un cliente que trata de conseguir un conjunto de productos. Los autores proponen un algoritmo memético y muestran cómo el desempeño de su propuesta está ligado a la configuración de sus parámetros

El año 2020 marcó precedente respecto al uso de la tecnología en la educación. *Aprende en el mundo de CARLA* es un artículo donde Santamaría et al. presentan un método para desarrollar un agente conversacional en realidad virtual con analítica llamado CARLA. Mediante el método propuesto es posible construir un chatbot que

interactúa con el usuario a través de diálogos hablados, con el fin de proporcionar apoyo al usuario en una plataforma E-Learning. La intención de los autores es que este software sea de muy bajo costo o libre y lo suficientemente robusto para operar sin intervención humana.

Tijuana es la 6ª ciudad más grande de México y al igual que en otras grandes ciudades, sus principales avenidas y arterias suelen estar congestionadas por el tráfico vehicular. Valencia et al. presentan un estudio donde realizan un *Análisis espacial de la red de transporte metropolitano en Tijuana*. Con este estudio se busca caracterizar las propiedades estructurales e identificar la ubicación de nodos importantes. El estudio presenta características interesantes sobre la movilidad del transporte público en esa ciudad.

Finalmente, Castro nos presenta en *La realidad extendida como herramienta clave para el futuro* cómo la realidad virtual ha transformado distintas áreas: medicina, arquitectura, industria, educación entre otros. En este artículo se presenta, además, una explicación clara sobre la diferencia entre conceptos como realidad virtual, realidad aumentada y realidad mixta.

Esperamos que los disfrutes y nos envíes tus comentarios a través de nuestras redes sociales. *



Dra. María Lucía Barrón Estrada es profesora e investigadora del Tecnológico Nacional de México en el Instituto Tecnológico de Culiacán y pertenece al Sistema Nacional de Investigadores. Sus áreas de interés son el aprendizaje electrónico y los sistemas inteligentes en la educación.



Dr. José Antonio Camarena Ibarrola es profesor e investigador de la Facultad de Ingeniería Eléctrica de la Universidad Michoacana de San Nicolás de Hidalgo, pertenece al Sistema Nacional de Investigadores. Sus áreas de interés son el procesamiento de señales, el reconocimiento de patrones y los procesos estocásticos.



Dra. Karina Mariela Figueroa Mora es profesora e investigadora de la Facultad de Ciencias Físico-Matemáticas de la Universidad Michoacana de San Nicolás de Hidalgo, pertenece al Sistema Nacional de Investigadores. Sus áreas de interés son algoritmos en general, recuperación de información, bases de datos métricas, así como desarrollo de software para la enseñanza de las matemáticas.



Dra. María Yasmín Hernández Pérez es profesora e investigadora del Tecnológico Nacional de México en el Centro Nacional de Investigación y Desarrollo Tecnológico. Pertenece al Sistema Nacional de Investigadores. Sus áreas de interés son sistemas tutores inteligentes, minería de datos educativos, computación afectiva y procesamiento de lenguaje natural.

Sapiens Piensa - Invitados

Félix Castro Espinoza y Miguel González Mendoza

En este volumen, la Sociedad Mexicana de Inteligencia Artificial invitó a la comunidad a considerar la convocatoria al premio José Negrete 2020 a mejores Tesis de Maestría y Doctorado, defendidas en alguna Institución de Educación Superior o Centro de Investigación en México. El resultado de la convocatoria se dió a conocer en el marco del *19th Mexican International Conference on Artificial Intelligence, MICA I 2020*. Con mucho agrado les compartimos lo que podrán encontrar los lectores en estos trabajos tan interesantes.

El resumen presentado por Navonil Majumber y Alexander Gelbukh se intitula *Análisis de sentimientos y emociones en redes sociales*. Este trabajo fue ganador de 1er lugar en tesis de Doctorado. Nos presentan cómo es posible analizar los sentimientos en las publicaciones que son parte del día a día en las redes sociales. Sin duda, dado que el Internet es uno de los ejes de la tecnología moderna, el conocer el estado emocional de los usuarios es muy importante pues muchos aspectos de nuestras actividades cotidiana dependen de esto.

Claudia Sánchez y Mario Graff nos presentan un resumen del 2o lugar en tesis de Doctorado con el título *Heurísticas de Selección para Programación Genética Semántica para resolver Problemas de Clasificación*. En este trabajo nos muestran tres heurísticas de selección de padres basados en los vectores semánticos de los individuos y en otras propiedades.

Para el resumen del 3er lugar en tesis de Doctorado titulado *Un clasificador basado en transformaciones de espacio y selección de modelos*, presentado por José Ortiz-Bejar et al. nos muestran esta tarea de aprendizaje automático supervisado la cual trata de predecir una variable categórica a partir de un conjunto de elementos etiquetados. Los autores presentan una comparación contra 16 clasificadores ampliamente utilizados.

El trabajo *Redes Neuronales Convolucionales para Pilotaje Autónomo en Carrera de Drones* de Leticia Oyuki y José Martínez fue el ganador del 1er lugar en tesis de Maestría. Los autores nos adentran a un interesante mundo donde las carreras de drones se han convertido en un deporte popular. Nos presentan los retos que tiene este tipo de carreras y algunos algoritmos que se utilizan para este fin.

El último resumen presentado se titula *Detección de emociones en los medios sociales*. Este trabajo fue ganador del 3er lugar en la tesis de Maestría por Muhammad Hammad y Alexander Gelbukh. En este resumen nos muestran la posible categorización de 11 emociones siendo un problema de clasificación de etiquetas múltiples.

Invitamos a los lectores a conocer los mejores trabajos del 2020 en posgrado sobre Inteligencia Artificial en México.



Dr. Félix Castro Espinoza tiene Maestría y Doctorado en Inteligencia Artificial por la Universidad Politécnica de Cataluña, UPC, en Barcelona España. Es profesor-investigador en el Centro de Investigación en Tecnologías de Información y Sistemas, CITIS, de la Universidad Autónoma del Estado de Hidalgo, UAEH. Presidente actual de la Sociedad de Inteligencia Artificial, miembro del grupo de Investigación en Soft Computing de la UPC y líder del Cuerpo Académico Consolidado de Computación Inteligente de la UAEH. Sus líneas de investigación incluyen: Inteligencia Artificial, Soft Computing, Data Mining, Algoritmos de Extracción Automática de Reglas, Reducción de la Dimensionalidad en Sistemas Complejos, Lógica Difusa, e-Learning y su relación con la Inteligencia Artificial. Félix Castro ha dirigido alrededor de 35 tesis de doctorado, maestría y licenciatura, ha publicado alrededor de 90 artículos en revistas indizadas, capítulos de libro y artículos de congresos internacionales y nacionales



Dr. Miguel González Mendoza cuenta con Maestría y Doctorado en Inteligencia Artificial del Institut National des Sciences Appliquées de Toulouse, Francia, en 2000 y 2004 respectivamente. Se desempeña como profesor-investigador en el Tecnológico de Monterrey, México. Interés en aplicaciones de Machine Learning, Data Management y Visión por Computadora, áreas en las que ha asesorado 11 tesis doctorales y 28 tesis de maestría. Autor de más de 100 artículos en JCR, congresos y capítulos de libro. El Dr. González Mendoza lidera varios proyectos de investigación en México y la Unión Europea (como local chair), es expresidente de la Sociedad Mexicana de Inteligencia Artificial (2016-2018). Miembro Regular de la Academia Mexicana de Ciencias, Miembro Regular de la Academia Mexicana de Computación. Young Scientist en el Foro Económico Mundial.

Estado del IArte

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

Aplicaciones de la Inteligencia Artificial

En esta ocasión abordaremos los avances más recientes en el ámbito de las aplicaciones de la IA. Iniciaremos con las aplicaciones que identifican el estado de ánimo de las personas, pasando hacia los trabajos multidisciplinarios de imitación de órganos del cuerpo humano y aplicaciones en el ámbito médico, hasta el desarrollo de aplicaciones relacionadas con la pandemia ocasionada por el virus SARS-CoV2 (COVID-19).

El interés de la comunidad académica de IA hacia el modelado de las emociones, ha permitido desarrollar software capaz de identificar el estado de ánimo de las personas. Existe un vasto número de aplicaciones, destacando algunos trabajos de las redes sociales en línea que podrían detectar conductas suicidas. Sin embargo, difícilmente verán la luz debido al actual debate acerca de la privacidad. Rana el Kaliouby, cofundadora de Affectiva.com una de las IA más exitosas en este ámbito, ha enfatizado la importancia de *humanizar* las tecnologías haciendo uso de IA.

Algunos de los esfuerzos más importantes de la comunidad de IA se están llevando a cabo con el fin de emular los principales órganos del cuerpo humano. El ojo es un órgano bastante sofisticado. En la Universidad del Estado de Oregon, USA, John Labram y su equipo de investigadores hicieron uso de sensores ópticos para reproducir la habilidad del ojo humano de percibir cambios en su campo visual. En China, investigadores de la Universidad de Ciencia y Tecnología de Hong Kong han creado un ojo artificial basado en perovskita, un material conductor de la luz usado en celdas solares, y han logrado procesar patrones de luz en tan sólo 19 milisegundos. Esto es la mitad del tiempo que le toma a un ojo humano. Un equipo multidisciplinario de las Universidades de Ciencia y Tecnología de Pohang, Corea, y de Stanford, USA, están trabajando en crear una piel *electrónica* lo más parecida a la piel humana. También, TrinamiX, subsidiaria de BASF, está desarrollando soluciones para el cuidado de la piel a través de sensado molecular combinando tecnologías inalámbricas y teléfonos con microprocesadores Snapdragon de Qualcomm los cuales incorporan una IA que se comunica con la plataforma de Qualcomm para procesar los datos capturados por el teléfono.

Investigadores de la Universidad Rice y del Centro Anderson del Cáncer en Texas, han incorporado IA a un microscopio para determinar en tiempo real, si el tejido de cáncer ha sido removido exitosamente durante una cirugía. *DeepDOF* permite reducir costos y tiempos en el proceso de estudiar las secciones de tejido extraído. En la fase de entrenamiento de la aplicación se hizo uso de una base de datos de 1,200 imágenes.

OSTAR Meditech Corp ha desarrollado una tecnología para detectar enfermedades del corazón en etapas tempranas. Incorpora técnicas de IA para analizar datos de los pacientes y detectar condiciones anormales del corazón. El algoritmo hace uso de diagramas de espectro BPM/ECG. Esta empresa de Taiwan ganó el premio R&D 2020 en el área de ciencia y tecnología el año pasado. Para saber más puede consultar (en inglés) <https://www.rdworldonline.com>.

Respecto a la pandemia, en el Colegio de Ciencias de la Universidad de Purdue, USA, un grupo de investigadores y médicos están haciendo uso de la IA para estudiar con mayor rapidez y precisión los datos generados por una red de instituciones de salud, con la finalidad de identificar drogas que permitan combatir la enfermedad, haciendo especial énfasis en el impacto que tienen los medicamentos actuales sobre el COVID-19 mediante el análisis, selección y clasificación de las respuestas de los pacientes a los diversos tratamientos. En nuestro país, también se ha hecho lo propio a través de la Alianza de Inteligencia Artificial (AIA) del Conacyt, universidades públicas y Centros de Investigación Conacyt, han creado *pandemiahacks*, robots para evaluar síntomas, herramientas de diagnóstico rápido para monitorear, valorar y evitar contagios; aplicaciones de apoyo psicológico así como de detección molecular, entre otras.

Sin duda, las aplicaciones de IA para imitar sofisticados órganos del cuerpo humano, el desarrollo de trabajos en el ámbito médico y para combatir la pandemia provocada por la COVID-19, así como la creación de nuevas aplicaciones se incrementarán notablemente y por supuesto, algunas de ellas seguirán causando controversia.*

Sakbe

Laura Cruz Reyes, Claudia Guadalupe Gómez Santillán y Laura Gómez Cruz

sakbe@komputersapiens.org

Aplicaciones de la IA

La IA es una de las áreas de investigación que en estos tiempos está siendo aplicada en el desarrollo de diversas herramientas de hardware y software, con las que todos de alguna forma interactuamos por medio de los dispositivos electrónicos. Para los que les gustaría consultar algunas definiciones de esta área de conocimiento ac-

tualizadas les recomendamos revisar el siguiente artículo <https://www.mujeresconsejeras.com/5-claves-de-la-inteligencia-artificial-para-el-2020/>; en el cual la autoras hacen una recopilación de algunas temáticas claves que se necesitan conocer como lo son: los tipos de IA, aplicaciones de la IA en las empresas, como empezar a aplicar la IA, entre otras.

**MUJERES
CONSEJERAS.COM**

Después de actualizar algunos conceptos, que apoyaran para comprender y tener una introducción rápida en las tendencias de la IA y el software que se están aplicando actualmente en esta área de conocimiento a través de las siguientes recomendaciones: <https://cio.com.mx/cinco-tendencias-de-inteligencia-artificial-en-2020/>, describe las tendencias a seguir para conocer las temáticas actuales en las que se pueden apoyar a las empresas, por ejemplo algunas tendencias son: Automatización robótica de procesos (RPA), Procesamiento de lenguaje natural (PNL), Entorno de trabajo de IA de código abierto, entre otras; Continuamos con https://www.capterra.com/sem-compare/artificial-intelligence-software?gclid=CjwKCAiA_eb-BRB2EiwAGBnXXqaCXNGi9YtoZZEKsabE65o0QIXSYfbJSE-4rdHeImWCruGJtNibNIBoCOAMQAvD_BwE. Es un sitio que usa software de IA, que se enfoca en hacer comparaciones entre diferentes productos de software que realizan la misma actividad; es decir ayuda a las empresas a buscar el mejor software recomendado para alguna actividad específica. Las comparaciones las realiza con base en las características que se describen de los productos de software.

Finalmente aprovechando el inicio de año, les compartiremos las tendencias de la IA para este 2021. Cabe destacar que como mencionan en <https://cio.com.mx/nueve-tendencias-tecnologicas-para-2021/> A diferencia de otros años las tendencias se ven influidas por el fuerte impacto provocado por la situación de crisis global generada por la pandemia de la COVID-19, la cual ha llevado a muchas empresas a tener que adaptarse a las condiciones cambiantes para componer el negocio futuro. Dentro de las tendencias podemos encontrar: Internet



de los comportamientos, Nube distribuida, informática que mejore la privacidad, malla de ciberseguridad, entre otras. Otro sitio comenta que: “la pandemia de la COVID-19 ha funcionado como catalizador del desarrollo de esta disciplina”, y ubican dentro de las tendencias más importantes a el Big Data, la analítica de información y a los modelos predictivos. https://www.quanticsolutions.es/inteligencia-artificial/tendencias-en-inteligencia-artificial-para-2021.*



ARTÍCULO ACEPTADO

Algoritmo Memético para el Problema de Ventas por Internet con Costos de Envío (MAIShOP)

Miguel A. García Morales, Héctor J. Fraire Huacuja, Laura Cruz Reyes, Claudia G. Gómez Santillán, Nelson Rangel Valdez y Ma. Lucila Morales Rodríguez

Introducción

En la actualidad, el comercio electrónico se ha convertido en parte importante de la sociedad. La implementación de tecnología novedosa y en constante crecimiento ha hecho inevitable adaptarse a esta evolución [1].

Un caso significativo de este escenario, dada su implicación económica, es la compra por Internet, representada en la Figura 1. Como los costos de las tiendas virtuales son sensiblemente bajos en comparación con los físicos, cada vez más minoristas e incluso individuos ofrecen sus productos en línea.

“Por lo tanto, para un usuario en particular en busca de la mejor oferta, mirar a través de todos los rangos de las tiendas y tener en cuenta todos los rangos de productos y costos de envío asociados, puede ser muy difícil de administrar” [2].

Para los vendedores, la principal ventaja es que sus ofertas están disponibles para un público más amplio, sin la mayoría de los costos asociados, como alquiler, impuestos, mantenimiento y publicidad.

Los clientes tienen la comodidad de poder hacer sus compras desde cualquier lugar, en cualquier momento, con mejores precios y una gama más amplia de productos, siempre que tengan acceso a Internet [2].

En el problema de optimización de compras por Internet (IShOP) [3] se supone que un cliente con una lista de compras requiere adquirir los productos en un conjunto de tiendas en línea al menor costo posible.

En [4] se propuso el problema y se demostró que es NP-duro.

En [2] se propone un algoritmo de procesamiento celular que simula el procesamiento paralelo de dos o más procesos de búsqueda en el espacio de soluciones y actualmente se le considera el mejor algoritmo del estado del arte de IShOP.

En el presente artículo, se propone un algoritmo memético (MAIShOP) que a diferencia del algoritmo del estado del arte utiliza una representación vectorial para la representación de las soluciones candidatas, un mecanismo que acelera el cálculo de la función objetivo y para cada generación se actualiza la población dejando solo la mejor solución y recalculando el resto de la población.

Estos elementos se diseñaron con el propósito de lograr una mejora significativa en el desempeño del algoritmo. Para validar este nuevo enfoque se realiza un estudio

comparativo del desempeño del algoritmo propuesto contra el del algoritmo del estado del arte.

En los experimentos computacionales se usó un conjunto amplio de instancias estándar y los resultados muestran una clara superioridad del algoritmo memético. Para soportar estadísticamente los resultados se aplica la prueba de hipótesis no paramétrica de Wilcoxon.



Figura 1. Representación esquemática del IShOP.

Definición del problema

Un cliente desea comprar por internet un conjunto de n productos N , los cuales se pueden adquirir en un conjunto de m tiendas disponibles M . El conjunto N_i contiene los productos disponibles en la tienda i , cada producto $j \in N_i$ tiene un costo de c_{ij} y un costo de envío d_i .

El costo de envío se cobra solo si se compran uno o más productos en la tienda i . El problema consiste en minimizar el costo total de la compra de todos los productos en N (costo de los productos más los gastos de envío).

Formalmente el problema consiste en determinar una partición de los productos a comprar en las diferentes tiendas $X = (X_1, \dots, X_m)$, de modo que $X_i \subseteq N_i$ y que se compren todos los productos $\bigcup_{i=1}^m X_i = N$ y que minimice el costo total:

$$F(X) = \sum_{i=1}^m \left(\sigma(|X_i|) d_i + \sum_{j \in X_i} c_{ij} \right)$$

Dónde: $|X_i|$ es la cardinalidad del contenedor X_i y $\sigma(x) = 0$ si $x = 0$ y $\sigma(x) = 1$ si $x > 0$.

El valor objetivo de la solución I se calcula de la siguiente manera:

$$F(I) = \sum_{i=1}^m d_i + \sum_{j=1 \vee j \text{ tal que } I(j)=i}^n (c_{ij})$$

Si la lista de productos que se va a comprar está formada por N productos y hay M tiendas disponibles, entonces una solución se representa en un vector de longitud N , el cual contiene para cada producto la tienda donde se va a comprar.

Finalmente, la estructura tradicional del algoritmo memético se modificó de tal manera que al término de cada generación la población se actualiza dejando solo la mejor solución y regenerando el resto de la población. La idea es que la mejor solución transmita sus características a las siguientes generaciones y conjuntamente con la regeneración del resto de los individuos, se logre un mejor balance de la intensificación y la exploración del algoritmo.

Estructura general del algoritmo memético

En el algoritmo se usa una representación vectorial de las soluciones que, a diferencia del algoritmo del estado del arte, reduce la complejidad del procesamiento de una solución de $O(n^2)$ a $O(n)$.

La búsqueda local asigna cada producto a la tienda en que resulta más conveniente su compra, este proceso se realiza de manera intensiva en el algoritmo. Con el propósito de hacer más eficiente la localización de la tienda más conveniente, se incorpora un mecanismo que acelera el cálculo de la función objetivo.

Selección por torneo

Se seleccionan aleatoriamente p individuos de la población Pop y el más apto es el que pasa a la siguiente generación. En el algoritmo memético se realiza un torneo binario (con $p = 2$) sobre la población Pop . Cada solución participa en dos torneos y la solución ganadora se selecciona y se inserta en $NewPop$. De esta manera, cualquier solución en Pop puede tener como máximo dos copias en $NewPop$. La Figura 2 ilustra el proceso anterior descrito.

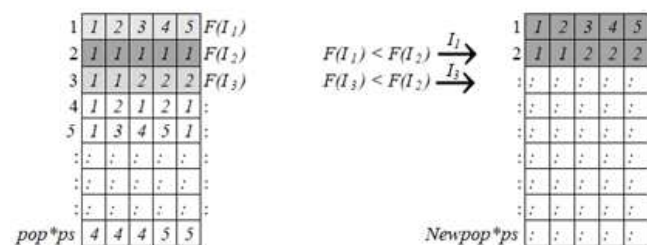


Figura 2. Torneo binario.

Operador de cruza

Este operador se aplica a un par de soluciones ($padre_1$ y $padre_2$), seleccionados de manera secuencial hasta que

se hayan seleccionado un porcentaje de los individuos de la población [5]. Para generar la solución $hijo_1$, primero se toma la mitad inicial del $padre_1$ y se une con la segunda mitad del $padre_2$. Posteriormente para formar la solución $hijo_2$, se toma la mitad inicial del $padre_2$ y se une con la segunda mitad del $padre_1$ [6].

En el operador de cruza, el punto $\lfloor N/2 \rfloor$ ó $\lceil N/2 \rceil$ se selecciona como punto de cruce.

Operador de mutación

Se diseñó un método de mutación heurística, considerando el siguiente procedimiento:

Seleccionamos $ps * er$ soluciones y por cada una de estas se genera un número al azar, si este valor generado al azar es menor que $ps * pm$ se procede a generar todas las posibles combinaciones de asignación (producto, tienda), se evalúan todas estas combinaciones y se selecciona la mejor.

Búsqueda local

Este procedimiento comienza seleccionando una solución de $IntermediatePop$, siendo está un vector de valores asociados que contiene las tiendas j asignadas en la secuencia de selección de productos X_{old} con su costo, que incluye c_{ij} y d_i del multiconjunto N_i .

Posteriormente se verifica para cada producto, desde que tienda podríamos comprarlo para reducir el costo total de la selección de productos X .

Cuando finaliza la búsqueda de esa tienda, se mueve a un nuevo producto. Este proceso continúa hasta que todas las tiendas j en X hayan sido revisadas.

Concluye cuando todas las X hayan sido asignadas en $ChildPop$.

Algoritmo memético

La Figura 3 describe la estructura general del algoritmo MAIShOP. El algoritmo, comienza definiendo los parámetros tales como tamaño de la población inicial ps , porcentaje de cruza pc , probabilidad de mutación pm , porcentaje de elitismo er , criterio de terminación $timelimit$, enseguida se genera una población inicial (paso 2) utilizando una heurística de construcción.

En la fase generacional (paso 5-15), se aplica el operador de torneo binario para la selección en el paso 6. Las mejores soluciones $ps * er$ (elitismo) de pop se copian en $IntermediatePop$ en el paso 7.

En $ps * pc$ de las soluciones restantes en $NewPop$, se aplica el operador de cruza para generar más soluciones para $IntermediatePop$ y las soluciones restantes se copian tal como están en $IntermediatePop$ en el paso 8.

Se supone que pc se selecciona de una manera que $ps * pc$ es menor o igual que el número de soluciones restantes en $NewPop$.

En el paso 9, los individuos en $IntermediatePop$ sufren mutación excepto aquellos individuos que fueron se-

leccionados en el paso 6 usando elitismo. En el paso 10, se aplica la búsqueda local a *IntermediatePop* para mejorar las soluciones.

En el paso 11, se obtiene la mejor solución local (*BestLocal*) de *ChildPop* y se compara con la solución global (*BestGlobal*) en el paso 12, además se calcula el tiempo de ejecución (*timesolution*) una vez evaluado este paso, si *BestLocal* tiene un menor costo entonces reemplaza a *BestGlobal*, en el paso 13 la mejor solución local (*BestLocal*) es copiada a la población inicial y todas las demás soluciones son recalculadas, se calcula el tiempo de ejecución total del algoritmo para cada generación en el paso 14 y finalmente se retorna la mejor solución global (*BestGlobal*) con el cálculo de tiempo de ejecución en la cual se encontró esta solución en el paso 16.

```

Input:
    timelimit: Tiempo límite en la ejecución del algoritmo
Output:
    BestGlobal: Mejor solución global
    timesolution: Tiempo de la solución
1:  Se definen los siguientes elementos: tamaño de la población inicial ps,
    porcentaje de cruce pc, probabilidad de mutación pm, porcentaje de
    elitismo er, criterio de terminación timelimit
2:  Generar población inicial pop de tamaño ps
3:  Determinar BestLocal en pop.
4:  BestGlobal = BestLocal.
5:  while no se cumpla el criterio de terminación timelimit do
6:      Aplicar operador de torneo binario en pop para obtener NewPop
7:      ps.er soluciones élite de NewPop se pasan a IntermediatePop
8:      Entre los individuos restantes en NewPop, se aplica el operador de
        cruce a ps*pc soluciones seleccionadas al azar y las soluciones restantes
        se copian tal como están a IntermediatePop
9:      Aplique el operador de mutación en IntermediatePop con
        probabilidad pm excepto a las soluciones élite.
10:     Para cada individuo en IntermediatePop, se realiza una mejora
        local asociando a cada producto la tienda que reduce el costo del
        producto, el individuo mejorado se coloca en ChildPop.
11:     Se obtiene la mejor solución (BestLocal) de ChildPop
12:     Se determina si el valor objetivo de la mejor solución local
        (BestLocal) es menor que la solución global (BestGlobal), si es así se
        actualiza la solución global. Se calcula el tiempo de ejecución del
        algoritmo hasta este paso (timesolution).
13:     La mejor solución local (BestLocal) se copia a pop y todos los
        demás individuos son generados nuevamente.
14:     Se calcula el tiempo de ejecución del algoritmo para cada
        generación (totaltime).
15: end while
16: return (BestGlobal, timesolution)

```

Figura 3. Algoritmo memético (MAIShOP).

Experimentos computacionales

La plataforma de hardware y software en que se realizaron los experimentos incluye: un procesador Intel Core 2 Duo a 2.53 GHz. con 4 GB de RAM y Java versión 1.8.0_231.

El nombre de las instancias especifica su tamaño, *m* representa el número de tiendas y *n* la cantidad de productos. Se crearon tres conjuntos de instancias de diferentes tamaños. Un conjunto de instancias pequeñas, con 3 subconjuntos: *3n20m*, *4n20m* y *5n20m*. Un conjunto de instancias medianas, con tres subconjuntos: *5n240m*, *5n400m* y *50n240m*. Un conjunto de instancias grandes con tres subconjuntos: *50n400m*, *100n240m* y *100n400m*. Cada subconjunto contiene 30 instancias.

El objetivo de los experimentos computacionales que se realizaron fue comparar el desempeño del algoritmo

memético (MAIShOP) y el algoritmo de procesamiento celular (Pcell).

Cada instancia se resolvió realizando 30 corridas independientes con cada uno de los algoritmos y se determinó el valor objetivo de la mejor solución encontrada y el tiempo de dicha solución.

Para determinar el desempeño de los algoritmos se calcularon los valores promedio del valor objetivo y del tiempo de solución; y las desviaciones estándar.

Se aplicó la prueba de hipótesis no paramétrica de Wilcoxon para determinar la significancia estadística de las diferencias observadas en calidad y eficiencia.

La Tabla 1 muestra los resultados obtenidos con cada grupo de 30 instancias. Para cada algoritmo se muestra el promedio y la desviación estándar (subíndice) del valor objetivo y del tiempo de solución.

En las celdas de Pcell se indica si hay diferencias significativas a favor del algoritmo Pcell (↓) o a favor del algoritmo MAIShOP (↑) y si no hay diferencias significativas entre los dos algoritmos (—).

La significancia se determinó con una confiabilidad del 95 %. Las celdas sombreadas corresponden al algoritmo que obtuvo el mejor resultado en calidad o eficiencia.

Tabla 1. Desempeño comparativo de los algoritmos.

Instancia	Pcell (FO)	MAIShOP (FO)	Pcell (Time)	MAIShOP (Time)
3n20m	63.03 _{0.483} ↑	62.76 _{2.57E-14}	0.00049 _{0.0006} ↑	1.2222E-05 _{6.21E-05}
4n20m	78.83 _{0.40} ↑	78.73 _{2.45E-14}	0.0018 _{50.0012} ↑	6.6667E-06 _{4.26E-05}
5n20m	102.56 _{0.80} —	102.09 _{0.0043}	0.0027 _{30.0014} ↑	0.00016 _{0.0002}
5n240m	75.43 _{0.77} ↓	76.32 _{1.67}	0.2736 _{0.0957} ↑	0.0669 _{0.0627}
5n400m	69.98 _{1.72} —	70.55 _{1.68}	0.7900 _{20.2882} ↑	0.2425 _{0.2139}
50n240m	1058.56 _{103.06} ↑	505.47 _{17.60}	16.4441 _{3.0266} ↓	22.5010 _{7.5014}
50n400m	951.08 _{119.35} ↑	438.74 _{24.09}	48.6607 _{25.3042} ↑	33.1063 _{3.8351}
100n240m	2349.94 _{159.48} ↑	993.70 _{38.25}	73.8573 _{39.7564} ↑	50.4910 _{5.6348}
100n400m	2213.90 _{173.70} ↑	801.51 _{35.03}	181.9521 _{95.3004} —	176.1110 _{41.5457}

Conclusiones

En este artículo, se propone un algoritmo memético (MAIShOP) que incorpora una serie de mecanismos innovadores: utiliza una representación vectorial para la representación de las soluciones candidatas, un mecanismo que acelera el cálculo de la función objetivo y para cada generación se actualiza la población dejando solo la mejor solución y recalculando el resto de la población. Estos mecanismos permiten resolver eficientemente el problema de IShOP.

Para validar este nuevo enfoque se realizó un estudio comparativo del desempeño del algoritmo propuesto contra el del mejor algoritmo del estado del arte.

Los resultados de los experimentos realizados muestran que el algoritmo propuesto supera claramente en calidad y eficiencia al mejor algoritmo del estado del arte. En calidad mejora significativamente en 6/9 conjuntos y en eficiencia mejora significativamente en 7/9 conjuntos.*

REFERENCIAS

1. Musial, J., Pecero, J., Dorronsoro, B., Blazewicz, J. (2014). Internet Shopping Optimization Project (IShOP). *European IST Projects*, 16.
2. Lopez-Loces, M.C., Musial, J., Pecero, J.E., Fraire-Huacuja, H.J., Blazewicz, J., y Bouvry, P. (2016). Exact and heuristic approaches to solve the Internet shopping optimization problem with delivery costs. *International Journal of Applied Mathematics and Computer Science*, 26(2), 391-406.
3. Musial, J. (2013). Applications of Combinatorial optimization for online shopping. (Doctoral dissertation, University of Luxembourg, Luxemburgo, Luxemburgo).
4. Gen, M., y Cheng, R. (2000). Genetic algorithms and engineering optimization. *John Wiley & Sons, Inc.*, EEUU.
5. Holland, J. (1975). Adaptation in natural and artificial systems: an introductory analysis with application to biology. *Control and artificial intelligence*.
6. Umbarkar, A.J., y Sheth, P.D. (2015). Crossover operators in genetic algorithms: a review. *ICTACT journal on soft computing*, 6(1).

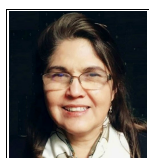
SOBRE LOS AUTORES



Miguel Ángel García Morales es un estudiante de Maestría en el Departamento de Estudios de Posgrado e Investigación en el Tecnológico Nacional de México-Instituto Tecnológico de Cd. Madero. El obtuvo su grado de licenciatura en Sistemas Computacionales en el Tecnológico Nacional de México-Instituto Tecnológico de Cd. Madero en 2008. Su interés de investigación incluye optimización inteligente.



Héctor Joaquín Fraire Huacuja es Profesor investigador del Tecnológico Nacional de México Campus Cd. Madero. Doctor en Ciencias de la Computación por el Centro Nacional de Investigación y Desarrollo Tecnológico. Reconocido durante 15 años consecutivos como miembro el SNI del CONACYT, recientemente recibió el más alto nivel de reconocimiento, investigador de Nivel III. Su principal área de interés es la optimización inteligente de problemas NP-Duros con uno múltiples objetivos.



Laura Cruz Reyes es profesora en el Instituto Tecnológico de Ciudad Madero, doctora en ciencias de la computación del Centro Nacional de Investigación y Desarrollo Tecnológico y maestra en ciencias de la computación del Instituto Tecnológico y de Estudios Superiores de Monterrey. Es miembro del Sistema Nacional de Investigadores de México en el nivel III y líder del cuerpo académico de Optimización Inteligente ante PRODEP. Sus áreas de interés incluyen optimización inteligente, heurísticos, algoritmos evolutivos, aprendizaje automático, decisión multicriterio y logística.



Claudia Guadalupe Gómez Santillán Obtuvo el grado de doctor en Tecnología Avanzada en el Politécnico Nacional de México. Actualmente, es Profesora Titular en el Instituto Tecnológico de Ciudad Madero. Sus intereses de investigación son la toma de decisiones difusa, el modelado de preferencias, los sistemas difusos, los sistemas de soporte de decisiones, el proceso de optimización, la inteligencia computacional, el descubrimiento de conocimientos, minería de datos, reconocimiento de patrones e ingeniería de software.



Nelson Rangel Valdez es Doctor en Ciencias de la Computación egresado del Laboratorio Nacional de Tecnologías de la Información del CINEVESTAV-Unidad Tamaulipas. Actualmente es Catedrático CONACYT adscrito al Instituto Tecnológico de Ciudad Madero, desarrollando el proyecto "Optimización de Problemas Complejos". Sus áreas de interés son Optimización Inteligente e Inteligencia Artificial. Ha desarrollado algoritmos multicriterio para Programación de Horarios, Ajuste de Parámetros Preferenciales, Construcción de Arreglos de Cobertura, y Diseño de Agentes.



María Lucila Morales Rodríguez recibió su grado de Doctora en Inteligencia Artificial por la Universidad Paul Sabatier, en Toulouse, Francia en 2007. En 2014 realizó un PostDoctorado en el Centro Europeo de Realidad Virtual en Brest, Francia. Sus líneas de investigación principales son la Toma de Decisiones, Aprendizaje Máquina y Agentes Virtuales Inteligentes. Es profesor-investigador del Instituto Tecnológico de Ciudad Madero, donde trabaja principalmente en el modelado comportamental de Agentes Virtuales Inteligentes Socio-Emocionales.

ARTÍCULO ACEPTADO

Aprende en el mundo de CARLA

Pablo Isaac Macias-Huerta, Rafael Cabañas-Rocha, Carlos Natanael Lecona Valdespino y Guillermo Santamaría-Bonfil

Resumen

Las interfaces humano-computadora utilizadas en entornos educativos virtuales imponen a los alumnos, además del tema de aprender, la necesidad de habituarse a ellas. No obstante, las actuales tendencias tecnológicas ofrecen formas de interactividad más naturales como el lenguaje hablado. En este trabajo se presenta un método para desarrollar un agente conversacional en realidad virtual con analítica, llamado CARLA. Este método permite la construcción de un chatbot que proporciona apoyo a un usuario dentro de una plataforma de E-Learning a través de diálogos hablados. El análisis recae en la capacidad del sistema de analizar la entrada de voz del usuario. Por el momento, CARLA se compone de tres elementos: la comprensión del lenguaje natural, el manejador de diálogos, y el generador de lenguaje natural. CARLA es evaluada en términos de la detección correcta de interacciones válidas e inválidas utilizando el caso de educación en sistemas de energía en un ambiente de realidad virtual. Los resultados muestran que CARLA logra identificar con alta precisión interacciones válidas, y en menor medida interacciones inválidas.

Introducción

De acuerdo con Google Trends (<http://bit.ly/eLearningTrend>), debido a la pandemia de COVID-19, las tecnologías de E-Learning han alcanzado su mayor pico de interés en los últimos 14 años. Desde la capacitación tradicional hasta la corporativa, estas tecnologías educativas son útiles en la mitigación de la propagación del virus ya que no se requiere la presencia de los estudiantes en el aula. Las tecnologías de E-Learning abarcan desde plataformas aisladas como entornos de Realidad Virtual (VR), hasta un alto nivel de ecosistemas de aprendizaje entrelazados [8].

La VR permite el despliegue de entornos de aprendizaje contextualizado con muchos beneficios: facilita el aprendizaje a través de la exploración y la práctica repetitiva, mejora la motivación y compromiso para aprender de los estudiantes, aplicable en equipos computacionales convencionales (y de bajo costo) hasta de alta sofisticación [7]. Desafortunadamente, la mayoría de los sistemas de VR utilizan Interfaces Humano-Computadora (HCI) como teclado y mouse los cuales limitan la interactividad, mientras que interfaces más sofisticadas requieren habituarse a ellas, una carga cognitiva adicional al proceso de aprendizaje. Afortunadamente, con la creciente madurez de las tecnologías de inteligencia artificial apli-

cadadas al Procesamiento del Lenguaje Natural (NLP), las posibilidades de integrar la conversación y el discurso en el E-Learning están recibiendo mayor atención. Como resultado existen aplicaciones que incluyen tutoría, respuesta a preguntas, práctica de conversación para el aprendizaje de idiomas, agentes pedagógicos, compañeros de aprendizaje, diálogos para promover la reflexión, entre otros [3].

Es por eso que proponemos un agente conversacional en realidad virtual con analítica, CARLA por sus siglas en inglés, para su uso en ambientes de aprendizaje. El objetivo no solo es construir un chatbot con software de bajo costo y/o libre, sino que además sea lo suficientemente robusto como para operar sin la necesidad de demasiados ajustes humanos.

El resto de este documento describe los materiales requeridos para CARLA como el ambiente virtual y los componentes del agente conversacional, un caso de estudio y su evaluación en términos de interacciones válidas e inválidas. Finalmente, se presenta brevemente conclusiones del proyecto, y se analizan futuras mejoras y cambios en la plataforma.

Materiales

Hay muchas formas de construir un agente conversacional, como complementos para redes sociales hasta robots humanoides de servicio como Pepper, creados industrialmente para necesidades comerciales [6]. En este sentido, CARLA se propone como un método para el desarrollo de un chatbot hablado en realidad virtual utilizando tecnologías libres y de bajo costo. En lo siguiente, hablamos del ambiente de aprendizaje, enfatizando los componentes del software utilizados para el NLP.

Entorno de aprendizaje en realidad virtual

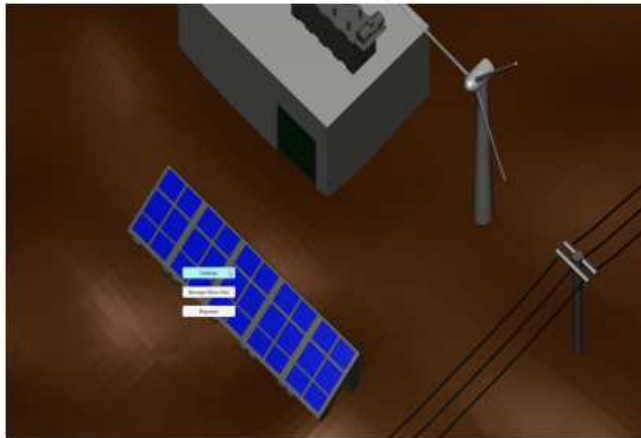
El entorno de aprendizaje se basa en realidad virtual no inmersiva desarrollada con Unity 3D. Este es un motor de VR para el desarrollo de videojuegos serios y casuales. Es un motor gratuito y compatible con múltiples sistemas operativos.

El ambiente de aprendizaje es el área de educación en sistemas de energía. Este entorno consiste en escenas 3D en VR de tres tipos: (1) menús, (2) instalaciones navegables y (3) escenas de comprensión del equipo. Los primeros están diseñados para proporcionar a los alumnos información sobre las instalaciones disponibles y cómo

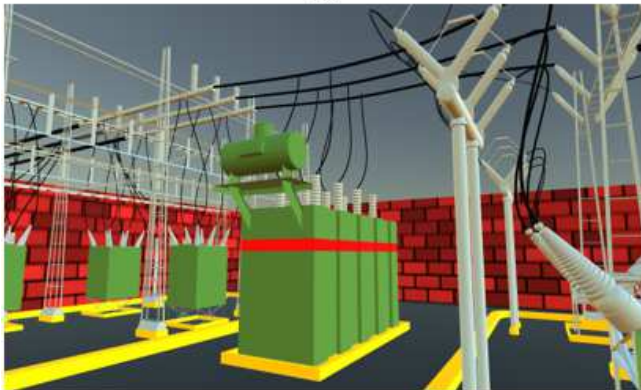
navegar por el entorno. Las segundas presentan varias instalaciones de sistemas de energía, que se pueden manipular y explorar para comprender conceptos de elec-

tricidad y fuentes energéticas. Por ejemplo, en la Figura 1a se muestra un menú, mientras que 1b se presenta una versión estilizada de una subestación eléctrica.

CARLA es una interfaz con la cual se puede manipular y conocer objetos en un mundo virtual usando expresiones verbales.



(a)



(b)

Figura 1. Entorno de aprendizaje en VR para CARLA. En (a) el menú del escenario de una micro red; en (b) una subestación eléctrica.

Componentes del Chatbot

Un chatbot es un sistema informático que funciona como HCI entre usuarios y el programa a través de texto o enunciados hablados. Se compone de 3 elementos principales: Comprensión del Lenguaje Natural (NLU), manejador de diálogos (DM) y Generador de Lenguaje Natural (NLG) [2]. Estos tres componentes y el flujo de los mensajes se muestran en la Figura 2.

NLU es un proceso de inteligencia artificial que sirve para extraer del mensaje hablado la intención. Por intención nos referimos a los datos, la información o la manipulación del entorno de aprendizaje que el usuario desea lograr. Por tanto, la NLU analiza los enunciados,

primero pasando del habla al texto y luego produce una estructura lingüística que el manejador de diálogos pueda interpretar. En el caso de CARLA, la NLU se compone por el API de Google Speech y el kit de herramientas de análisis de Stanford CoreNLP [5]. El primero, es una herramienta web ofrecida por Google que permite transformar un mensaje de voz a texto, cuyo costo es de 0.006 USD por 15 segundos de grabación. El segundo, es un conjunto de herramientas de NLP desarrollado por Stanford CoreNLP Group [5] el cual permite la división de oraciones en palabras, identificando categorías gramaticales (POS), sustantivos, valores numéricos y de tiempo, análisis de sentimientos, etc. Soporta 6 idiomas, aunque las funciones soportadas dependen del idioma elegido.

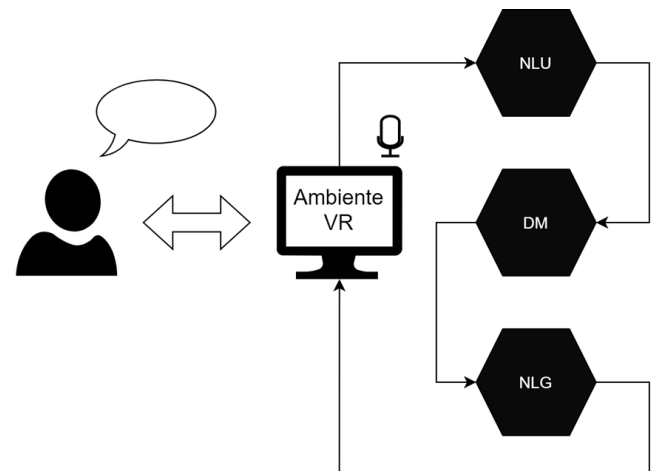


Figura 2. Arquitectura de CARLA. Los mensajes son provistos por un usuario al ambiente de VR los cuales son procesados por el sistema para sintetizar la respuesta adecuada.

El manejador de diálogos es el módulo que analiza las intenciones para determinar qué respuesta o respuestas debe producir el sistema [2]. Una vez que se ha extraído el significado del enunciado, se realiza un análisis de intención comparando las palabras filtradas con las diferentes intenciones válidas. Si la intención es válida, modifica uno o más objetos en el entorno de realidad virtual, además de proporcionar información hablada. Si no se reconoce la intención, CARLA informará al usuario de esta situación. Por el momento, esta validación

está codificada de forma fija en el sistema, pero podría automatizarse mediante algoritmos de aprendizaje automático [9]. Para el caso de estudio se diseñaron 15 intenciones válidas:

```

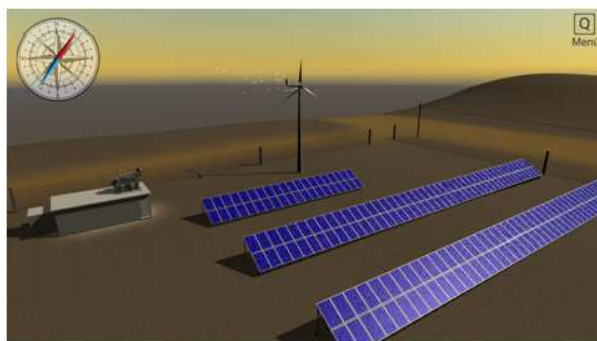
mostrar / qué es un panel solar
explicar cómo funciona un panel solar
mostrar / qué es un transformador
explicar cómo funciona un transformador
mostrar / qué es un paquete de baterías
explicar cómo funciona un paquete de baterías
mostrar / qué es un aerogenerador
gracias
ayuda
rotar (izquierda y derecha)
detener la rotación
acercar
alejarse.
zoom normal.
salir.

```

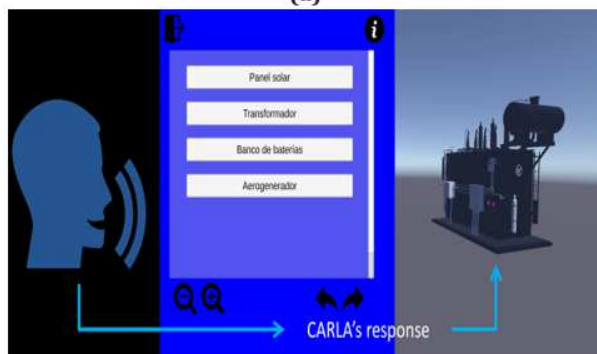
La última etapa es la NLG donde se sintetiza la respuesta hablada de CARLA. Para esto, obtenemos el texto de la respuesta codificada (aunque podría ser auto-generada mediante técnicas de redes neuronales) para transformarla a un mensaje de audio a través del SDK de texto a voz de Azure. Esta herramienta cuenta con más de 140 voces diferentes en más de 45 idiomas y variantes, por lo que proporciona voces muy expresivas que suenan como humanas. En particular, proporciona mediante técnicas de redes neuronales la capacidad de personalizar varios estilos de habla como chat o servicio al cliente, e incluso emociones como la alegría y empatía.

Unas últimas palabras sobre la parte de análisis de CARLA son merecidas. Aunque, por el momento, todos los análisis se centran en la NLP, se espera la integración con un modelo de dominio de conocimiento formal (como un mapa conceptual, clasificación de diálogos, extracción de datos de seguimiento de uso, etc.) en un futuro próximo.

La interfaz CARLA usa: Comprensión y generación del lenguaje natural, un manejador de diálogos y Unity3D.



(a)



(b)

Figura 3. En (a) una microrred del mundo real con simulación eólica y fotovoltaica; en (b) la escena de comprensión del equipo.

Experimentación

A continuación se presentan el caso de estudio y los resultados experimentales con respecto al desempeño del reconocimiento de intenciones válidas e inválidas.

CARLA en la educación en sistemas de energía

En la Figura 3 se muestra una microrred del mundo real, donde el alumno pueda manipular la generación de energía eólica y solar. Las escenas de comprensión del equipo presentan componentes de energía en forma de catálogo donde los usuarios pueden explorar de forma adicional los componentes del escenario, así como una explicación general de los fenómenos físicos bajo los cuales operan, dependiendo del comando dado.

Desempeño del reconocimiento de intenciones de CARLA

Probamos las capacidades de CARLA en términos de reconocimiento de intenciones válidas e inválidas. Una intención válida (un verdadero positivo -TP-) se identifica y se atiende dado que un conjunto específico de palabras estén contenidas dentro de un enunciado del usuario, mientras que una intención inválida (un verdadero negativo -TN-) es una solicitud que no puede ser atendida. En el caso del primero, una intención válida sería “CARLA, ¿qué es un panel solar?”; en el caso de este últi-

mo una intención inválida sería “CARLA, ¿por qué brilla el sol?”; en ambos casos, CARLA debe actuar en consecuencia de lo contrario incurrirá en un error (un falso

negativo -FN- si no se reconoce una intención utilizando el conjunto de palabras adecuado, o un falso positivo -FP- si se identifica una intención cuando no debería).

CARLA logra un nivel de detección elevado de intenciones, con posibles aplicaciones a sistemas de tutoría, conocimiento y apoyo.

Las pruebas incluyeron una muestra de 5 personas, 3 hombres y 2 mujeres, con una edad promedio de 37.4 años. Cada sujeto de prueba dijo 40 frases, de estas 10 corresponden a intenciones inválidas que tenían combinaciones de letras o pronunciaciones similares a intenciones válidas. Los resultados se muestran en la Tabla 1.

Como puede verse, la precisión del modelo es muy alta con un 92 %, con una tasa de detección de verdaderos positivos del 97 % y una tasa de detección de verdaderos negativos del 78 %. Esto es consecuencia de la madurez de las tecnologías de inteligencia artificial utilizadas para el NLP, las cuales permiten que un enfoque simple como el usar conjuntos de palabras codificadas de forma fija, tiene una alta tasa de detección de intenciones válidas, pero es menos robusto para las intenciones no válidas.

Tabla 1. Matriz de confusión de CARLA.

		Actual Value		total
		Prediction outcome		
	total	140 True Positives	12 False Positives	152
		4 False Negatives	44 True Negatives	48
		144	56	

Conclusión y discusión

En este trabajo, proponemos a CARLA como un método para el desarrollo de un agente conversacional educativo en realidad virtual con capacidades de NLP. Los experimentos con respecto a las principales capacidades del chatbot son prometedores. Por el momento, sólo analizamos el reconocimiento de intenciones válidas e inválidas con un enfoque simple. En términos de precisión general el desempeño fue muy satisfactorio. Sin embargo, las intenciones válidas se pueden activar utilizando enunciados no válidos con combinaciones de letras o pronunciaciones similares. Por lo tanto, las próximas mejoras a CARLA estarán enfocadas en la aplicación de modelos de aprendizaje automático para impulsar el reconocimiento de intenciones. Es importante notar que las funcionalidades de CARLA sólo están disponibles con una conexión a Internet activa, lo que puede crear interrupciones en la

disponibilidad del servicio y afectar la percepción que los usuarios tengan del agente [4]. De igual forma, para detectar estados emocionales relevantes para el aprendizaje, se profundizará en trabajos futuros el incluir el reconocimiento de emociones faciales mediante técnicas de Deep Learning las cuales destacan en esta tarea. Estas acciones mejorarán las interacciones de CARLA y permitirán personalizarlas para satisfacer las necesidades específicas de los usuarios. En este sentido, CARLA está destinada a convertirse, en algún momento del futuro cercano, en un Asistente Virtual Inteligente (IVA), un tutor o compañero de aprendizaje.*

REFERENCIAS

- Macias-Huerta, P., Santamaría-Bonfil, G., Ibáñez, M.B. (2020). CARLA: Conversational Agent in virtual reality with Analytics. En *13th Workshop on Intelligent Learning Environments*. Ciudad de México, México, 1-8.
- Galitsky, B. (2019). Developing Enterprise Chatbots. *Springer International Publishing*. Cham, Switzerland ISBN 978-303-004-298-1.
- Kerry, A., Ellis, R., Bull, S. (2008). Conversational agents in E-Learning. En *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, 169-182.
- Klopfenstein, L.C., Delpriori, S., Malatini, S., Bogliolo, A. (2017). The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms. En *Proceedings of the 2017 Conference on Designing Interactive Systems*, 555-565.
- Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S.J., McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. En *Association for Computational Linguistics (ACL) System Demonstrations*, 55-60.
- Pandey, A.K., Gelin, R. (2018). A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robotics & Automation Magazine*, 25(3), 40-48.
- Santamaría-Bonfil, G., Ibáñez, M.B., Pérez-Ramírez, M., Arroyo-Figueroa, G., Martínez-Álvarez, F. (2020). Learning analytics for student modeling in virtual reality training systems: Lineworkers case. *Computers & Education*, 151, 103871.
- Santamaría-Bonfil, G. (2019). Towards an Interactive Learning Ecosystem: Education in Power Systems. *Research in Computing Science*, 149(1), 6-7.
- Schuermans, J., Frasincar, F. (2019). Intent Classification for Dialogue Utterances. *IEEE Intelligent Systems*, 35(1), 82-88.
- Chung, H., Lee, S. (2018). Intelligent Virtual Assistant knows Your Life. arXiv preprint arXiv:1803.00466, Vol. abs/1803.00466.

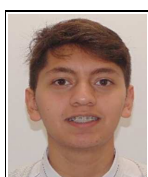
SOBRE LOS AUTORES



Pablo Isaac Macias Huerta es estudiante del CECyT 9 “Juan de Dios Bátiz” en la carrera de Técnico en Programación. Ganó el “Best Paper Award 2020” en el 13th WILE (Workshop on Intelligent Learning Environments) con el artículo “CARLA: Conversational Agent in viRtual ReaLity with Analytics.” Participante en los Encuentros Académicos Interpolitécnicos de Geometría y Trigonometría y Física I, obteniendo el quinto y vigésimo segundo lugar respectivamente. Participó en la estancia del XXV Verano de la Investigación Científica y Tecnológica del Pacífico con el proyecto presentado en WILE. Sus intereses de investigación incluyen el desarrollo de Videojuegos y Tecnologías Interactivas.



Rafael Cabañas Rocha es estudiante del CECyT 9 “Juan de Dios Bátiz”. Estudia la carrera Técnico en Programación. Es egresado de la Escuela Secundaria Anexa a la Normal Superior de México, institución con la cual participó en la Olimpiada de Matemáticas e Historia a nivel secundaria. A sus 6 años participó en la Olimpiada del Conocimiento quedando en segundo lugar a nivel regional. Sus intereses están enfocados al estudio y desarrollo de sistemas computacionales, desarrollo de videojuegos, inteligencia artificial y ciencia de datos.



Carlos Natanael Lecona Valdespino es un estudiante del CECyT 9 “Juan de Dios Bátiz” en la carrera de Técnico en Programación. Participó en la competencia de robótica FIRST en 2019 donde obtuvo el “Rookie Inspiration Award” y en la estancia del XXV Verano de la Investigación Científica y Tecnológica del Pacífico colaborando en el proyecto “Modelo de Inteligencia Artificial para la Detección de Enfermedades Cardiovasculares”. Sus intereses de investigación incluyen visión artificial y procesamiento del lenguaje natural.



Guillermo Santamaría Bonfil es Doctor en Ciencias Computacionales y catedrático CONACYT asignado al Instituto Nacional de Electricidad y Energías Limpias. Sus líneas de investigación son la ciencia de datos en problemas educativos y de sistemas de potencia, modelado y pronóstico de secuencias temporales, complejidad y análisis no lineal. Es SNI nivel 1, con publicaciones en revistas internacionales, congresos nacionales e internacionales, capítulos de libros, así como la co-dirección de estudiantes de maestría y doctorado en programas PNPC. Cuenta con experiencia en la industria en compañías como IBM y BBVA.

La inteligencia artificial conversacional se refiere a tecnologías, como chatbots o asistentes de voz, con las que los usuarios pueden hablar. Utilizan grandes volúmenes de datos, aprendizaje automático y procesamiento de lenguaje natural para ayudar a imitar las interacciones humanas, reconociendo las entradas de voz y texto y traduciendo sus significados en varios idiomas.

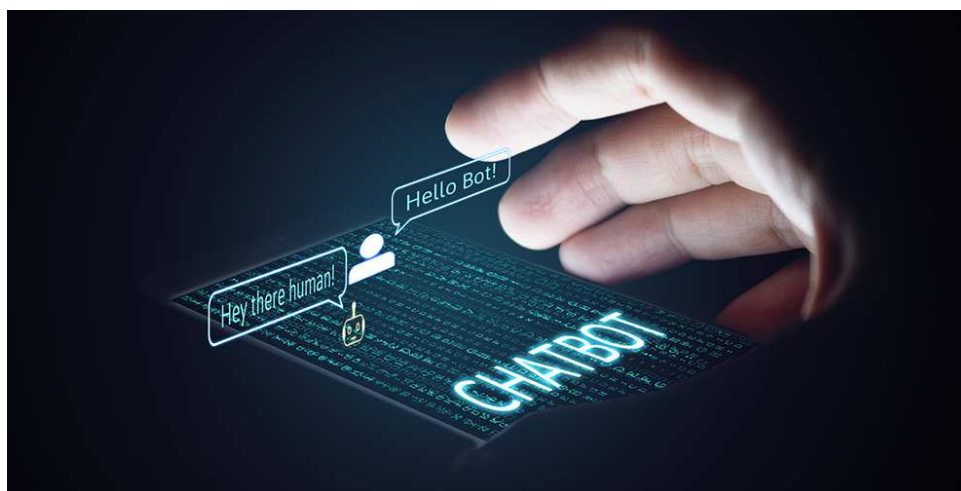


Imagen tomada de: <https://www.ibm.com/>.

ARTÍCULO ACEPTADO

Análisis espacial de la red de transporte metropolitano de Tijuana

Bélgica J. Valencia García y Adán Hiraes-Carbajal

Resumen

Tijuana es la sexta ciudad más grande de México. Debido a sus características socio-geográficas, las principales avenidas y arterias de la ciudad suelen estar congestionadas. En el 2006 y a partir del 2018, las flotas de transporte público de Tijuana han sido sometidas a una importante renovación, al reemplazar vehículos de pequeña capacidad, camionetas y taxis por otros de mayor capacidad (autobuses). Sin embargo, a nuestro conocimiento, no se han analizado las propiedades de su red de transporte. Este estudio busca caracterizar las propiedades estructurales e identificar las ubicaciones de nodos importantes. Modelamos la red de transporte como un grafo dirigido ponderado. Localizamos nodos importantes analizando cada nodo en función de tres criterios: grado, distancia y atributos basados en rango. Usamos ordenamiento no-dominado para localizar nodos en el frente de Pareto 1 que maximicen todos los criterios. Entre los principales hallazgos, encontramos que para un viajero que se traslade a lo largo del diámetro o el radio de la red, se requieren en promedio 16 y 10 transbordos de ruta correspondientemente. Adicionalmente, cuatro de nueve delegaciones concentran los nodos más importantes. Mostrando que la red necesita de una reestructuración.

Introducción

Tijuana, al ser parte de la aglomeración de la megalópolis San Diego - Tijuana, cuenta con una red de transporte compleja que se adapta a las necesidades inmediatas de su población. Sus características sociodemográficas son muy dinámicas, experimenta: alto volumen de tráfico nacional, transfronterizo y entre municipios; altos intercambios internacionales en los puertos vehiculares entre Estados Unidos y México, los cuales producen largas filas y congestionamiento en vías aledañas; altos flujos migratorios, entre otras características [1]. El fenómeno de la inmigración ha acelerado el crecimiento de la ciudad, con más del 50 % de sus asentamientos de origen irregular, mayoritariamente en la periferia [1, 2]. Por lo que la configuración topológica de Tijuana está constituida por asentamientos humanos regulares e irregulares.

A nuestro conocimiento, la *Red de Transporte Masiva* (RTM) de Tijuana no ha sido analizada. Estudios previos se han centrado en comprender aspectos socioculturales y políticas públicas, p. ej. movilidad urbana sostenible [3], la calidad del servicio [4, 5], el suministro de la red

[6], la relación entre el transporte urbano y medio ambiente [7] y la política pública y gobernanza de la red de transporte [4,5]. En este estudio, nuestro enfoque es analizar la estructura de la red de transporte e identificar las propiedades de sus nodos.

Tres clases de rutas operan sobre la red de transporte de Tijuana, propiamente: Red de Transporte Masivo (RTM), el cual opera vehículos con una capacidad promedio de 30 pasajeros; taxis, con una capacidad promedio de 12 pasajeros; y Autobuses de Tránsito Rápido (ATR), con capacidades de 50 y 80 pasajeros. De cada clase, existen 129, 157 y 22 rutas respectivamente. En este estudio analizamos la RTM que resulta de fusionar 104 rutas independientes a la clase transporte masivo. El conjunto de datos fue facilitado por las oficinas de la Dirección Municipal de Transporte Público de la ciudad.

El modelo de red

La red de transporte está compuesta por un conjunto de coordenadas geodésicas y calles que las rutas atraviesan. Se distinguen dos tipos de calles: bidireccionales, las cuales permiten el tráfico en dos direcciones, y unidireccionales. No se considera la capacidad vial. Las rutas están preestablecidas y operan en circuito. Un modelo de coordenadas geodésicas es una intersección, una ubicación dentro de una ruta o una ubicación dentro de una calle curva. Referimos el término coordenada geodésica como nodo y calle como arista.

Formalmente, modelamos la red de transporte como un grafo dirigido ponderado $G = (v, \varepsilon)$, con v el conjunto de nodos y ε el conjunto de aristas, $v \neq \emptyset$. El número de elementos en v y ε están denotados como v y ε . Un Nodo $n_i \in v$ está definido por la tupla (φ_i, λ_i) , con φ_i y λ_i las coordenadas geodésicas (latitud y longitud). Una representación matricial para modelar la adyacencia entre nodos está dada por A , una matriz cuadrada de $N \times N$ con entradas a_{ij} igual a 1 si la arista $e_{ij} \in \varepsilon$, caso contrario cero. El peso de la arista w_{ij} representa la longitud del tramo de calle entre los nodos i y j . Dicha unidad es representada en metros. La fórmula Haversine [8] es utilizada para calcular la distancia ortodrómica entre los nodos i y j , es decir la distancia más corta entre dos puntos en la superficie terrestre.

El grado k_i del nodo i es el número de aristas incidentes al nodo y es definido como $k_i = \sum_{j \in v} a_{i,j}$. Para nodos dirigidos, el grado entrante está denotado como $k_i^{in} =$

$\sum_{j \in \mathcal{V}} a_{j,i}$ y el grado saliente como $k_i^{out} = \sum_{j \in \mathcal{V}} a_{i,j}$. El conjunto de nodos incidentes a i es denotado como \mathcal{N}_i . Los conjuntos de nodos predecesores y sucesores inmediatos se definen como \mathcal{N}^{in} y \mathcal{N}^{out} correspondientemente.

Sea el camino más corto (o geodésico) entre los nodos i y j , tal que $i \neq j$. Sea $L = \frac{1}{V(V-1)} \sum_{i,j \in \mathcal{V}, i \neq j} d_{ij}$ el camino más corto promedio definido como la longitud geodésica media entre todos los pares de nodos. Sea n_{jk} el número de todas las rutas más cortas entre los nodos j y k . Finalmente, sea $n_{jk}(i)$ el número de caminos más cortos entre los nodos j y k que atraviesan el nodo i .

Por último, el número de triángulos en el grafo se define como $\delta(G)$ y los tripletes (posibles triángulos) con dos vínculos dirigidos entre tres nos es definido como $\tau(G)$. Para un resumen exhaustivo de métricos para grafos se refiere al lector a [9].

Procesamiento de datos

Una planificación inadecuada de los sistemas de transporte urbano genera congestión, rutas altamente redundantes y desde la perspectiva de la tenencia de las rutas, una distribución desigual de la riqueza. Una planificación minuciosa requiere evaluar si las políticas de transporte aplicadas dan como resultado una infraestructura de movilidad sostenible, robusta y justa. Un enfoque cuantitativo hacia tal fin implica modelar la red de transporte, evaluar sus propiedades estructurales y dinámicas, y analizar los resultados observados.

Al más alto nivel de granularidad, métricas globales de red analizan relaciones complejas entre datos con el objetivo de caracterizar y encontrar similitudes entre las características descubiertas y modelos de grafos conocidos. Métricas de granularidad más fina, revelan nodos y propiedades de enlace del *Grafo Dirigido Acíclico* (DAG), p. ej. el grado de nodo, el grado de intermediación (*Betweenness*), entre otras propiedades. En este trabajo, nos enfocamos en caracterizar la topología a nivel global y nodo. Evaluamos la calidad de la red aplicando las métricas que se resumen en la Tabla 1. Debido a limitaciones de espacio, remitimos al lector a las referencias [9-13] para su interpretación.

Con el fin de facilitar resultados reproducibles y verificables, el conjunto de datos, las métricas, el flujo de procesamiento de datos y el entorno de desarrollo de anaconda del proyecto están disponibles a través del sitio del proyecto (ver https://github.com/ahiralesc/tijuana_vehicle_network_analysis).

La red de transporte se construyó concatenando 104 de las 129 rutas independientes. Las 25 rutas restantes que no se incluyeron corresponden a rutas con datos incompletos. Por lo tanto, la importancia de los nodos puede disminuir en todos los nodos que atraviesan estas rutas en nuestros resultados. El tamaño de V es de 20,411 nodos. Las rutas comparten hasta el 80 % de los

nodos. Los nodos que comparten una ubicación o están cerca unos de otros, es decir, dentro del ancho de la calle, fueron fusionados. Con esta acción el tamaño de V fue reducido a 4461 nodos no redundantes. La red de transporte se editó digitalmente con la aplicación JOSM Ver. 15238, un editor extensible para Open Street Map. El conjunto de rutas resultante se almacenó en formato geojson y está disponible para su descarga en el sitio del proyecto.

Tabla 1. Métricas a nivel red y nodo.

	Métrica	Expresión
Centralidad	Grado	$k_i = \sum_{j \in \mathcal{V}} a_{i,j}$
	Knn	$k_{nn,i} = \frac{1}{k_i} \sum_{j \in \mathcal{N}_i} k_j$
	Rango de página	pr_i , ver [10,18]
	Centralidad promedio de intermediación	$b_i = \sum_{j,k \in \mathcal{V}, j \neq k} \frac{n_{jk}(i)}{n_{jk}}$
	Promedio de grado de cercanía	$c_i = \frac{1}{\sum_{j \in \mathcal{V}, i \neq j} d_{ij}}$
Red	Coef. de asortatividad	r , ver [11]
	Diámetro	$d = \max_{j \in \mathcal{V}} d_{ij}$
	Transitividad	$t = \frac{3\delta(G)}{\tau(G)}$

El flujo de procesamiento de los datos consta de las siguientes fases: *fase de preparación de datos*, el cual carga y fusiona las rutas de transporte. La distancia ortodrómica entre los nodos también se calcula durante esta fase. Los nodos y aristas son almacenados aplicando memoria asociativa; *fase de preparación del grafo*, los diccionarios son transformados a un grafo dirigido; *fase de aplicación de métricas*, se aplican las métricas globales y de nivel de nodo seleccionadas por el usuario; y la *fase de análisis estadístico y visualización de datos*. El análisis estadístico se realizó mediante la aplicación de estadísticos de primer orden. Todos los datos procesados son almacenados en formato CSV (valores separados por comas) comprimidos (pickle). La importancia del nodo se evalúa categorizando las métricas de nodo en tres clases, propiamente fundamentadas en el concepto de: grado (deg); distancia y vecindario (dist); y rango (rank). Las métricas de la primera clase son \bar{k}_i , k_i^{in} , k_i^{out} , k_{nn} , de la segunda b_i y c_i , y de la tercera clase pr_i . Las métricas de cada clase se ponderan por igual y se combinan linealmente. La técnica de clasificación no-dominada se utiliza para extraer nodos que residen en el primer frente de Pareto [14]. Los nodos extraídos corresponden a los nodos con mayor importancia.

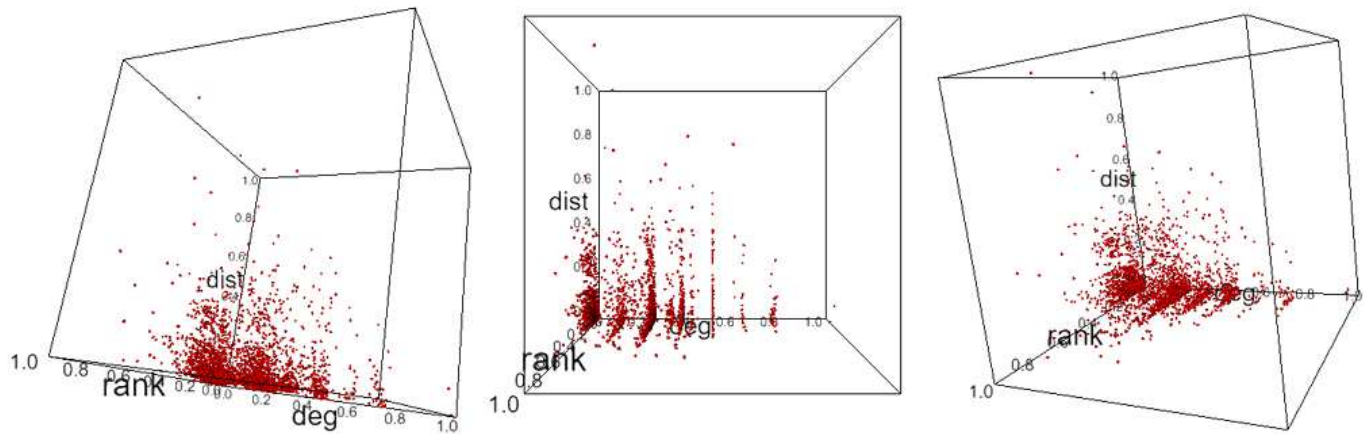


Figura 1. Linealización de métricas en espacio 3D. Fuente: Elaboración propia basada en [14].

Resultados

Los resultados se presentan en dos secciones. La primera sección describe las propiedades generales de la red. La segunda sección identifica los nodos de mayor importancia en acorde con el criterio de selección propuesto.

Propiedades de la red

Las propiedades a nivel de red se resumen en la Tabla 2. Se obtuvo un diámetro d o excentricidad máxima de 190. En promedio, un viajero que se traslada a lo largo del diámetro requiere 16 transbordos de ruta. Esa cantidad es cuatro veces mayor que los transbordos de la red de autobuses de Río de Janeiro [15]. El número de transbordos se puede minimizar a 10, sin embargo, aumenta la longitud del traslado. La excentricidad mínima (o radio) contiene 99 aristas. En promedio, se necesitan 10 transbordos de autobús para atravesar la ciudad a través de su radio, aunque el número de transbordos se puede minimizar a 6 a expensas de aumentar la longitud del viaje. El plan de desarrollo municipal de Tijuana 2017 informa que en promedio 4 transbordos de autobús son realizados por el 46 % de los usuarios de RTM [6], pero no existe ninguna declaración sobre la longitud del viaje. Se requerirían dos transbordos adicionales si la longitud de esos viajes fuesen igual a la del radio. Note, el anterior análisis supone que los transbordos ocurren en ubicaciones donde ocurre intersección de rutas. No supone la existencia de paraderos oficiales en tales ubicaciones.

Tabla 2. Propiedades globales de la red.

r	d	t
0.37	190	0.01

La asortatividad de la RTM r es de 0.37. Tal valor sugiere que la red no está centralizada en pocos nodos

concentradores (o hubs) y que éstos están interconectados. Los nodos hub a menudo poseen un grado de 7-9 y constituyen aproximadamente el 1.2 % de los nodos de la red. La transitividad de la red t es baja (0.01). En general, una red con una transitividad baja experimenta trayectos largos [16]. Estos resultados confirman que el diámetro de la red es grande.

Importancia del nodo

La Figura 1 muestra las puntuaciones ponderadas de importancia de los nodos según la linealización aplicada. Los nodos con menor importancia se agrupan cerca del origen, mientras que los de mayor importancia tienden al origen del cubo unitario. Los resultados muestran que la mayoría de los nodos se agrupan cerca del origen dentro del rango de $[0, 0.4]$. Dentro de esta región se forma un patrón similar a una nube. Sin embargo, a medida que los nodos se alejan del origen, la distribución de puntos comienza a inclinarse hacia el eje grado. Esto sugiere que la importancia del nodo está sesgada por el grado del nodo. Encontramos un total de 20 nodos no-dominados en el primer frente de Pareto. Estos corresponden a los nodos más importantes de la red de transporte (según la técnica propuesta).

Los 10 nodos más importantes se representan con marcadores azules. Ver la Figura 2. Los nodos en color rojo corresponden a los últimos 10 nodos con la segunda mayor importancia. Estos dos conjuntos pertenecen al primer frente de Pareto. El primer conjunto de nodos se concentra en 4 de las 9 delegaciones que componen la ciudad, propiamente: La Mesa, La Presa Abelardo, L. Rodríguez, La Presa Este y San Antonio de los Buenos. Las áreas que contienen estos nodos se amplían para una mejor visualización. Intuitivamente, uno esperaría que las grandes intersecciones concentren muchas rutas. Una inspección visual de las áreas restantes muestra que se

produce una gran concentración de rutas en calles principales de regiones con topología tipo malla, ver recuadros acompañantes de la Figura 2.

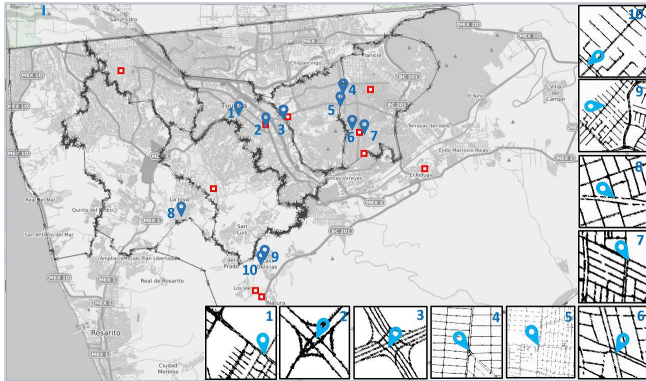


Figura 2. Nodos en el primer frente de Pareto. Fuente: Elaboración propia basada en [14].

La Tabla 3 lista los valores promedio de cada métrica correspondiente a los primeros 10 frentes de Pareto. Cuantitativamente, los nodos en los primeros cuatro frentes de Pareto son muy similares. Por lo tanto, comparten un nivel análogo de importancia. Acumulativamente, los cuatro primeros frentes suman un total de 164 nodos.

Tabla 3. Medias de las métricas.

Frente	k_j	b_i	c_i	$k_{m,i}$	pr_i
1	5.350	0.065	0.017	2.950	0.001
2	4.974	0.060	0.017	2.893	0.000
3	4.905	0.053	0.017	2.734	0.000
4	4.943	0.047	0.017	2.421	0.000
5	4.721	0.040	0.016	2.197	0.000
6	4.375	0.036	0.016	2.270	0.000
7	4.264	0.032	0.016	2.267	0.000
8	4.151	0.030	0.016	2.181	0.000
9	4.020	0.026	0.016	2.104	0.000
10	4.010	0.023	0.016	2.021	0.000

La longitud promedio de trayectoria a partir cualquier nodo alcanzable en el primer frente de Pareto es de 54. Tal longitud es menor que el radio por 45 aristas. Los nodos en el primer frente de Pareto pueden ser buenos candidatos para paradas de autobús y atraer una mayor proporción de viajes aumentando la centralidad de la red [17]. Sin embargo, es necesario evaluar su viabilidad, debido a que las ubicaciones actualmente modelan intersección de rutas.

La centralidad promedio de intermediación es baja. Aproximadamente el 6 % de todas las trayectorias más cortas pasan por un nodo en los conjuntos Pareto 1 y 2. Se encontró una centralidad máxima de intermediación de 0.21. Sin embargo, esto ocurre en un solo lugar. Incluso si ampliamos el rango de 0.1 a 0.21, solo el 1 % de los

nodos categorizados como importantes tienen este grado de centralidad.

El promedio de grado de cercanía no es significativamente alto, 0.017 y 0.016, esto podría explicarse porque: las rutas tienden a correr en paralelo a lo largo de las principales avenidas; transitan por glorietas; o siguen los circuitos en los suburbios.

Conclusiones

En este trabajo, analizamos las propiedades estructurales y de importancia de nodo de la Red de Transporte Masivo (RTM) de Tijuana modelándola como un grafo dirigido acíclico ponderado. En nuestro modelo, los pesos de las aristas corresponden a la distancia ortodrómica entre los nodos adyacentes, mientras que los nodos corresponden a coordenadas geodésicas. Se propone un flujo de procesamiento de datos de cuatro fases. La importancia de los nodos se evalúa categorizando las métricas de los nodos en tres clases, propiamente fundamentadas en el concepto de: grado; distancia y vecindad; y rango. Realizamos una combinación lineal ponderada de métricas por clase. Aplicamos la técnica de clasificación no-dominada para encontrar nodos en cada frente de Pareto.

Los resultados muestran que un viajero que se traslada a lo largo del diámetro o radio de la red requiere en promedio 16 y 10 transbordos de ruta, respectivamente. La red es asortativa, aproximadamente el 1.2 % de sus nodos experimentan un alto grado de concentración de rutas. En acorde con el criterio de asignación de importancia, los nodos importantes se concentran en 4 de las 9 delegaciones que componen la ciudad. Los nodos en el frente de Pareto 1 pueden reducir el radio por 45 aristas.

A nuestro conocimiento, este es el primer estudio que aborda el análisis de las propiedades estructurales de la red de transporte metropolitano de Tijuana. En estudios posteriores transformaremos el modelo de red a su representación multigrafo, ya que perdemos conocimiento de rutas independientes con la representación actual. También planeamos evaluar la robustez de la red con el objetivo de identificar las ubicaciones y eventos que pueden hacer que la red produzca particiones y sea inoperable.

Compartirnos el conjunto de datos, metodología y todos los instrumentos que construimos para el tratamiento de datos y análisis de la red de transporte en el sitio del proyecto. Para aplicar los anteriores instrumentos a otros estudios basta: (i) modelar las rutas de transporte como un grafo dirigido utilizando la aplicación JOSM; (ii) almacenar el grafo en formato geojson; (iii) y procesarlo utilizando el cuadernillo Jupyter que proporcionamos. Instrucciones adicionales para la instalación del ambiente anaconda y del cuadernillo están publicadas en el sitio del proyecto, ver [https://github.com/ahiralesc/tijuana_vehicle_network_analysis].*

REFERENCIAS

1. Ley, J. (2006). Movilidad y transporte público: El arreglo urbano en Tijuana. *Revista Universitaria UABC*, 4(53), 8-13.
2. Alegría, T. y Ordóñez, G. (2005). Legalizando la ciudad. Asentamientos informales y procesos de regulación en Tijuana. El Colegio de la Frontera Norte, México. ISBN: 968-7947-36-5, 170.
3. Avelar, M. (2014). Sistema de transporte de Tijuana: análisis de su proceso de modernización e implicaciones en la movilidad urbana. Un acercamiento desde la sustentabilidad. Tesis de maestría, Maestría en Desarrollo Regional. El Colegio de la Frontera Norte, México.
4. Castro, R. (2012). Procesos de concesión del transporte público en Tijuana, y efectos en la calidad del servicio. Tesis de maestría, Maestría en Desarrollo Regional. El Colegio de la Frontera Norte, México.
5. Millán, L.C. (2018). ¿Transporte público o mi automóvil?: Factores socioculturales determinantes en la elección modal de desplazamiento en Tijuana, B.C. Tesis de maestría, Maestría en Administración Integral del Ambiente. El Colegio de la Frontera Norte, México.
6. Instituto Metropolitano de Planeación IMPLAN. (2017). Plan municipal de desarrollo 2017-2019. Recuperado el 12 de Diciembre de 2019, de <https://www.implantijuana.org/informacion/C3%B3n/planes-y-programas/pmd-2017-2019/>.
7. Ramos, D. (2014). Relación entre forma urbana, movilidad a través de transporte motorizado y emisiones de GEI en Tijuana, B.C. Tesis de maestría, Maestría en Administración Integral del Ambiente. El Colegio de la Frontera Norte, México.
8. Chopde, N.R. y Nichat, M.K. (2013). Landmark based shortest path detection by using A* and haversine formula. *International Journal of Innovative Research in Computer and Communication Engineering*, 1(2), 298-302.
9. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. y Hwang, D. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5), 175-308.
10. Page, L., Brin, S., Motwani, R. y Winograd, T. (1998). The PageRank citation ranking: Bringing order to the Web. En *Proceedings of the 7th International World Wide Web Conference*. Brisbane, Australia, 161-172.
11. Newman, M.E.J. (2003). Mixing patterns in networks. *Physical Review E*, 67(2), 026126.
12. Foster, J.G., Foster, D.V., Grassberger, P. y Paczuski, M. (2010). Edge direction and the structure of networks. En *Proceedings of the National Academy of Sciences of the United States of America*, 107(24), 1-6.
13. Schank, T. y Wagner, D. (2005). Approximating clustering coefficient and transitivity. *Journal of Graph Algorithms and Applications*, 9(2), 265-275.
14. Coello, C., Veldhuizen, D. y Lamont, G. (2002). Evolutionary Algorithms for Solving Multi-Objective Problems. *Series Genetic and Evolutionary Computation*. Springer US. ISBN 978-1-4757-5184-0, 576.
15. Pumar, L., Barbastefano, R. y Carvalho, D. (2018). Bus transport network analysis in Rio de Janeiro based on topological models using social networks. ArXiv. <https://doi.org/10.31219/osf.io/azv63>.
16. Shanmukhappa, T., Ho, I., W.-H. y Tse, C.K. (2018). Spatial analysis of bus transport networks using network theory. *Physica A: Statistical Mechanics and its Applications*, Elsevier, 502, 295-314.
17. Scheurer, J., Curtis, C. y Porta, S. (2007). Spatial Network Analysis of Public Transport Systems: Developing a Strategic Planning Tool to Assess the Congruence of Movement and Urban Structure in Australian Cities. En *Australasian Transport Research Forum*. Melbourne, Australia, 19.
18. Langville, A.N. y Meyer, C.D. (2005). A Survey of Eigenvector Methods for Web Information Retrieval. *SIAM Review*, 47(1), 135-161.

SOBRE LOS AUTORES



Belgica J. Valencia García obtuvo el grado de Maestra en Ingeniería e Innovación en CETYS Universidad en la ciudad de Tijuana Baja California, en 2020. Ingeniera Industrial egresada de la Universidad Autónoma de Baja California. Sus intereses incluyen análisis espacial, movilidad urbana y transporte público.



Adán Hiraes Carbajal obtuvo el grado de Doctor en Ciencias de la Computación y Maestro en Ciencias de la Computación en 2012 y 2001 correspondientemente, en el Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE) en Baja California. Con especialidad en problemas combinatorios de calendarización y administración de recursos. Actualmente, es profesor-investigador en CETYS Universidad. Sus intereses de investigación incluyen: optimización, aprendizaje profundo e internet de las cosas.

ARTÍCULO ACEPTADO

La Realidad Extendida como herramienta clave para el futuro

Kathya Castro

La Realidad Extendida engloba todos aquellos elementos reales y virtuales combinados en un entorno generado por computadora y dispositivos móviles.

Poco a poco el mundo se ha ido transformando, y es gracias a la tecnología que muchos sectores como el industrial, el de la construcción, el inmobiliario, y el de la educación han ido evolucionando. Imaginemos una transición tecnológica que permita visualizar el mundo desde una perspectiva mucho más amplia, en el cual muchos negocios, sin importar el rubro al que pertenezcan, logren dar ese salto tecnológico hacia el futuro.

En el presente artículo se ilustrará como es que la Realidad Extendida puede facilitar dicho salto; sin embargo, para comprender de una mejor manera como es que esta será aplicada a distintas áreas, es necesario asimilar su significado y todos los componentes que la integran. El concepto de Realidad Extendida integra tanto las formas de XR (*Extended Reality*, por sus siglas en inglés) ya desarrolladas (de las cuales se hablará a más adelante) como aquellas que se establecerán en el futuro [1].

Es muy común confundir las tecnologías que forman parte de la XR; sin embargo, sus diversos componentes se diferencian e identifican principalmente por la relación que tienen entre el mundo real y el virtual, esto con un grado de inmersión diferente entre cada tecnología dependiendo del medio por el cual se haga uso de esta tecnología; dicho medio puede ser a través de lentes, equipos especiales, dispositivos móviles, entre otros.



Figura 1. Representación de la Realidad Aumentada en el sector de construcción.

Para comenzar con las tecnologías que integran la XR se hablará de lo que es la Realidad Virtual o VR (*Virtual Reality*, por sus siglas en inglés) siendo esta la más conocida; debido principalmente al gran diferenciador que ofrece, lo que la hace mucho más fácil de identificar por todos aquellos interesados en el tema sin importar su nivel de conocimiento respecto a dicha tecnología. La VR se define como un ambiente con escenas, situaciones y objetos de apariencia real, generado a través de tecnología informática, la cual brinda al usuario una sensación de inmersión dentro de dicho ambiente, esto por medio de la percepción de varios de sus sentidos, ya que la VR ofrece tanto una experiencia visual como una experiencia sonora, en la cual el usuario (al ser parte de la VR) se olvida del mundo real, logrando crear una nueva “realidad”; es por ello que dicho entorno únicamente se contempla y es posible a través de dispositivos tecnológicos (i.e., lentes o cascos de VR) [2].



Figura 2. Representación de la Realidad Virtual en actividades recreativas.

Si bien la VR no es una tecnología que muchas personas conozcan (y que se podría decir que acaba de ser introducida en el país por un número reducido de empresas que a nivel nacional brindan dichos servicios), la

VR tuvo su primer desarrollo en 1836, cuando Charles Wheatstone inventó el estereoscopio, un dispositivo capaz de crear una ilusión de profundidad en una imagen basándose en dos fotografías en extremo parecidas. Después de este dispositivo se comenzaron a generar distintos aparatos que tuvieran como finalidad acercar al usuario a un ambiente externo creando una nueva realidad. Otro de los dispositivos antiguos que permitieron cimentar el camino hacia la VR como actualmente la co-

nocemos fue el Sensorama, introducido en la década de los 50's. El Sensorama fue una máquina que reproducía películas en 3D; se diferenciaba del cine en que el usuario no sólo veía la película, sino que también la experimentaba, esto con base a distintos estímulos que el Sensorama presentaba, siendo estos la generación de olores, sonido envolvente y la atracción del usuario a la película por medio de sus imágenes 3D, las cuales en ese tiempo eran toda una novedad.

La Realidad Virtual facilita el proceso educativo, revolucionando sectores como el de la industria, el de construcción y el inmobiliario.

Es así que después de muchos años y dispositivos (que dejaban mucho que desear en cuanto a una experiencia inmersiva se refiere), en 2012 la marca Palmer Luckey reveló un dispositivo que lograba un ángulo de visión de 90°, los cuales son conocidos como Oculus Rift, dicho dispositivo logró captar la atención de Facebook, haciendo que el gigante de las redes sociales comprara la marca, convirtiéndose así en una de las empresas proveedoras y creadoras de ambientes de VR en el mundo[3].

Recientemente la VR ha sido usada como forma de entretenimiento a través de distintos videojuegos que permiten al usuario sentirse como si estuviera viviendo en carne propia una película de acción y aventura; sin embargo, en la actualidad, varias empresas nacionales (al igual que diversos sectores a nivel mundial) han optado por darle a la VR un enfoque relacionado con el aprendizaje y la capacitación (por ejemplo, en el sector industrial y de salud). Un ejemplo de lo anterior es el proceso de capacitación de personal por medio de la VR en las industrias con la finalidad de que aprendan de una manera mucho más simple y didáctica los procesos a desempeñar en su área de trabajo. Al aplicar esta tecnología se disminuyen los riesgos de accidentes en casi un 50 %, las empresas ahorran tiempo y dinero en entrenamiento, ya que, al hacer uso de un software de VR, los trabajadores pueden equivocarse sin tener repercusiones negativas que afecten a la empresa, sin mencionar que se maneja un autoaprendizaje, lo que eficientiza los tiempos de la empresa. Cada trabajador tiene una atención personalizada y no se tiene que contratar a alguien responsable de capacitar al personal, lo que no sólo representa un ahorro para la empresa, sino que también se tiene un mayor control del personal, además de que se erradican por completo las posibles mañas aprendidas, que por lo regular se transmiten en las capacitaciones industriales.

Otra de las aplicaciones más comunes que se le ha dado a la VR es en el sector salud, esto tanto para pacientes como para doctores, los cuales son entrenados envolviéndolos en un ambiente en donde puedan visualizar las

posibles situaciones que se puedan presentar en la práctica (operaciones, tratamientos contra enfermedades, entre otras), logrando resultados impresionantes que aseguraban que, en el caso de los cirujanos, el entrenamiento con VR reducía la duración de la cirugía y mejoraba el desempeño quirúrgico de los cirujanos. Mientras que la VR se aplica a los pacientes a manera de terapia que funciona para contrarrestar algunos efectos o síntomas de enfermedades, fobias y trastornos como la ansiedad, hiperactividad, estrés, entre otras.



Figura 3. Representación de la Realidad Aumentada en el sector educativo.

Por otro lado, el segundo componente de la XR es la Realidad Aumentada o AR (*Augmented Reality*, por sus siglas en inglés), la cual se refiere a la tecnología que permite superponer elementos virtuales sobre nuestra visión de la realidad, añadiendo capas de información visual sobre el mundo real que nos rodea a través de dispositivos tecnológicos móviles como Smartphones, Tablets y monóculos; la manera en como la AR sobrepone objetos en el entorno del usuario, mostrando la información digital en tiempo real la convierte en una herramienta clave para la optimización de distintos procesos, algunos de ellos serán revisados a continuación.

La AR avanzaba en base al desarrollo de la máquina Sensorama; sin embargo, los cimientos que forjaron el camino para la creación de la AR se dieron en 1973 cuando Myron W. Krueger realizó la primera instalación

de AR, que combinaba cámaras de video con un sistema de proyección, con la finalidad de establecer un entorno interactivo capaz de responder a los movimientos de los usuarios, todo esto por medio de sombras [4].

El término de AR es aplicado por primera vez en la década de los 90's por el investigador Thomas Caudell; sin embargo, el término fue acuñado por L.B. Rosenberg al desarrollar un dispositivo capaz de asesorar al usuario en el desempeño de tareas dentro del ejército estadounidense, específicamente la fuerza aérea [5].

Una vez que conocemos el contexto histórico de la AR, es que podemos partir y descubrir más acerca de sus aplicaciones en la actualidad. Una de las aplicacio-

nes que las empresas le han dado a la AR es en el sector inmobiliario; a través de una tarjeta de presentación que sirve como marcador, el usuario puede escanear el marcador con su Smartphone o Tablet, para que en el aparezca la propiedad desde una vista tanto externa como interna, además de que ofrece la posibilidad de conocer el inmueble con y sin muebles, aportando una vista con distintas tonalidades de luz a modo de vista diurna o nocturna, con la finalidad de mostrarle a posibles compradores como es que su dinero se verá invertido en una propiedad, sin siquiera tener la propiedad físicamente terminada, lo que no sólo ahorra costos a las inmobiliarias, sino que atrae a muchos más clientes potenciales.

La Realidad Mixta permite que el usuario interactúe con objetos reales dentro de un mundo virtual



Figura 4. Representación de la Realidad Aumentada en el sector inmobiliario.

Una de las maneras en como actualmente la AR es aplicada es por medio de filtros de Instagram y Facebook, los filtros se han convertido en la nueva herramienta del marketing digital, la cual permite a las empresas mejorar el branding de su marca a través de una campaña publicitaria indirecta capaz de captar la atención del cliente, generando interés en el target de la marca, y se obtienen como resultados: una conexión más profunda del usuario con la organización, mejores rendimientos en campañas de marketing, y un incremento de la marca en el *top of mind* de los clientes (tanto actuales como potenciales). Esta clase de estrategias resulta en un aumento en las compras y fidelización hacia la marca, lo cual es de vital importancia en tiempos como los que vivimos (en donde la nueva generación de consumidores es difícil de convencer y atraer), por lo que los filtros de AR suponen una manera dinámica, innovadora y atractiva de conectar con los consumidores. Por último, el tercer componente de la XR es la Realidad Mixta o MR (*Mixed Reality*, por sus siglas en inglés), la cual se refiere a la combinación en-

tre el mundo real y el virtual por medio de dispositivos electrónicos. La MR se puede definir como una mezcla entre la VR y la AR, por lo que combina los aspectos más representativos de ambas, unificando la experiencia para facilitar la interacción con el usuario. Algunos de los aspectos de la VR que son incorporados dentro de la MR es que permite al usuario sumergirse en entornos completamente digitales, mientras que incorpora el factor de la AR de incorporar objetos digitales en un entorno real; es así como la MR permite unir ambos conceptos con la finalidad de que el usuario interactúe con objetos reales dentro de un mundo virtual o, por el contrario, reproducir elementos virtuales en el entorno real, esto sin la necesidad de usar un marcador (a diferencia de la AR).

En la década de los 90's fue cuando por primera vez surge el término de MR, el cual tuvo su primera aparición en el ensayo *Augmented Reality: A class of displays on the reality-virtuality continuum*, de los autores Paul Milgram, Haruo Takemura, Akira Utsumi y Fumio Kishino [6]; es aquí donde se presentó el concepto MR como un todo que incluye a la AR y a la VR, y es a partir de dicho ensayo que los investigadores se han dedicado a darle mucho más énfasis a esta tecnología debido a la amplia gama de posibilidades que la MR ofrece, con mucho menos limitantes y restricciones (a diferencia de la VR y la AR).

Actualmente, diversas empresas dedicadas a sectores muy diferentes entre sí, como el arte, la moda, el diseño, etc., han hecho uso de esta tecnología con la finalidad de acercar sus productos y servicios al público de una manera innovadora, un ejemplo de ello es la empresa sueca de muebles más famosa del mundo, IKEA, la cual ha desarrollado una aplicación para dispositivos móviles en la cual se presenta la opción de escanear cualquier superficie del inmueble del usuario con la finalidad de presentar uno de sus productos de manera digital por encima de

la superficie, permitiendo que el cliente potencial se dé una idea de cómo el producto luciría en su propiedad,

sin la necesidad de comprarlo, lo que resulta en una herramienta que beneficia a ambas partes.

La Realidad Extendida engloba todos aquellos elementos reales y virtuales combinados en un entorno generado por computadora y dispositivos móviles.



Figura 5. Representación de la Realidad Aumentada en el sector educativo.

Muchas empresas han estado usando el concepto de catálogo digital por medio de filtros de MR, estas varían desde tiendas de lujo que ofrecen productos como zapatos por medio de esta tecnología, dando un sentido de estatus mucho mayor a la marca; de igual manera, empresas en el ámbito de la construcción, el sector inmobiliario e incluso la decoración de interiores, hacen uso de dicha tecnología, un ejemplo es la superposición de pisos en una superficie cualquiera, que le permite al usuario visualizar cómo es que el piso se vería, llamando la atención de los clientes y generando una imagen corporativa positiva al ofrecer un servicio de venta que va mucho más allá de lo convencional. Esta ventaja favorece que la empresa se posicione en la mente de los clientes tanto actuales como potenciales.

De igual manera, es importante tener en cuenta que

la MR también se puede implementar en el sector del entretenimiento, esto gracias a la facilidad de acceso que brinda a los usuarios a través de sus dispositivos móviles, un ejemplo de esto es Pokémon Go, el videojuego a través de la MR que superpone a sus famosos personajes para que el usuario pueda interactuar directamente con ellos.

Los ejemplos y las aplicaciones que se revisaron a lo largo de este artículo son solo un pequeño vistazo a todas las posibilidades que el mundo de la Realidad Virtual, Aumentada y Mixta ofrecen. Es importante expandir el conocimiento respecto a estas tecnologías ya que son las herramientas clave para llevar a las distintas industrias al siguiente nivel a través de un salto tecnológico. Bienvenido al futuro, bienvenido al mundo de la XR.*

REFERENCIAS

1. Burdea, G. y Coiffet, P. (2014). *Virtual Reality Technology*. Wiley-interscience.
2. Mihelj, M., Novak, D. y Begus, S. (2003). *Virtual Reality Technology and Applications*. Springer, ISBN 978-94-007-6909-0.
3. Sherman, W.R y Craig, A.B. (2003). *Understanding virtual reality*. Morgan Kaufman Publishers.
4. Reichardt, J., et al., eds. 2008. *White Heat Cold Logic: British Computer Art 1960-1980*. Cambridge, MA:MIT Press.
5. Muller, S. (1999). *Virtual Reality and Augmented Reality*. En *International Conference on Visual Computing*, Goa, India.
6. Milgram, P, Takemura, H., Utsumi, A., Kishino, F. (1994). *Augmented Reality: A class of displays on the reality-virtuality*. *Photonics for Industrial Applications*.

SOBRE LA AUTORA



Kathya Castro Es estudiante de Mercadotecnia Internacional en la Universidad Politécnica de San Luis Potosí. Parte del departamento de marketing de FYWARE (empresa potosina de Realidad Extendida).

ARTÍCULO INVITADO

Heurísticas de Selección para Programación Genética Semántica para resolver Problemas de Clasificación

Claudia N. Sánchez y Mario Graff

Introducción

Programación Genética (PG) [1] es una de las técnicas más recientes de la familia de los Algoritmos Evolutivos. Su uso ha sido muy útil para resolver problemas de aprendizaje supervisado, tanto regresión como clasificación. Específicamente, este documento utiliza PG para resolver problemas de clasificación. Algunos ejemplos de este tipo de problemas son los siguientes: reconocimiento de dígitos, predecir si un paciente tendrá o no un segundo ataque de corazón, o identificar la emoción en un texto corto [2]. Los primeros clasificadores basados en PG aparecen alrededor del año 2000 [3, 4]. Recientemente han generado modelos competitivos con clasificadores del estado del arte, como EvoDAG [5], TPOT[6], o M4GP[7]. Nuestro sistema de PG, EvoDAG [5] hace uso de los operadores genéticos recientes [8], además de estar inspirado en la implementación eficiente de [9]. Un ejemplo de un modelo generado por EvoDAG se muestra en la Figura 1, donde la clasificación es ascendente a partir de los nodos inferiores.

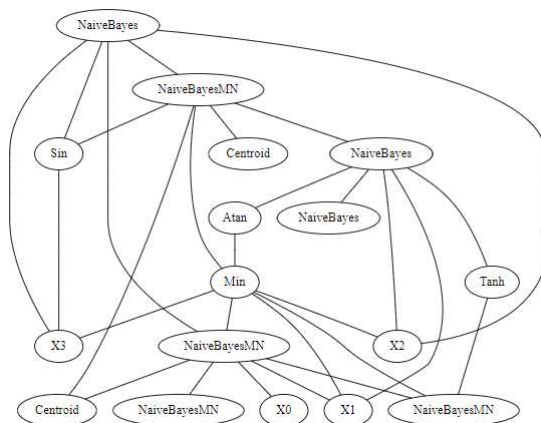


Figura 1. Modelo de clasificación generado con EvoDAG.

Metodología

PG genera un modelo de clasificación mediante la combinación de funciones matemáticas y también clasificadores como *Naive Bayes* y el Centroides más Cercano. Adicionalmente, consideramos diseñar operadores de cruce que utilicen las propiedades de dichas funciones matemáticas y clasificadores, además de la semántica de los individuos. Por lo anterior diseñamos tres heurísticas

de selección de padres para la función suma y para los clasificadores *Naive Bayes* y Centroides más Cercano. **Selección de Padres basada en Similitud Coseno.** La primera heurística usa el absoluto de la similitud coseno para seleccionar individuos padres con vectores de semántica lo más cercano a ángulos rectángulos. **Selección de Padres basada en el Coeficiente de Correlación de Pearson.** La segunda heurística usa el coeficiente de Pearson para seleccionar individuos padres con vectores de semántica no correlacionados. **Selección de Padres basada en Exactitud.** La tercera heurística selecciona padres minimizando la similitud que hay entre los vectores de semántica. Generalmente la exactitud es utilizada para medir el desempeño de los clasificadores, pero en este caso la idea es medir que tan similares son dos individuos entre sí, ya que cada individuo puede verse como un clasificador en sí mismo.

Experimentos y Resultados

Las heurísticas propuestas de selección, además de la técnica clásica de selección basada en aptitud y selección aleatoria, fueron probadas en 30 conjuntos de datos tomados del repositorio de la Universidad de California de Irvine. Para medir la eficiencia de los clasificadores se utilizó la métrica macro-F1. Los resultados arrojaron que las tres heurísticas son mejores que la técnica clásica de selección basada en aptitud. La heurística de selección que dio mejores resultados fue la basada en exactitud, además de ser estadísticamente mejor que la técnica de selección basada en aptitud. La prueba estadística utilizada fue Wilcoxon.

Para complementar los resultados, en la Figura 2 se puede observar una comparación de los clasificadores generados por EvoDAG y las heurísticas de selección propuestas (en azul) contra clasificadores clásicos y las técnicas de autoaprendizaje TPOT [6] y autosklearn [10]. Se puede ver que la heurística de selección basada en exactitud (EvoDAG acc) mejora significativamente el rendimiento de EvoDAG con la técnica clásica de selección basada en aptitud (EvoDAG fit). De hecho, también la selección aleatoria (EvoDAG rnd) mejora los resultados de la técnica clásica de selección. Sólo TPOT arrojó mejores resultados que EvoDAG, sin embargo, el tiempo de entrenamiento en promedio de la técnica TPOT es 50 veces más lenta que EvoDAG.

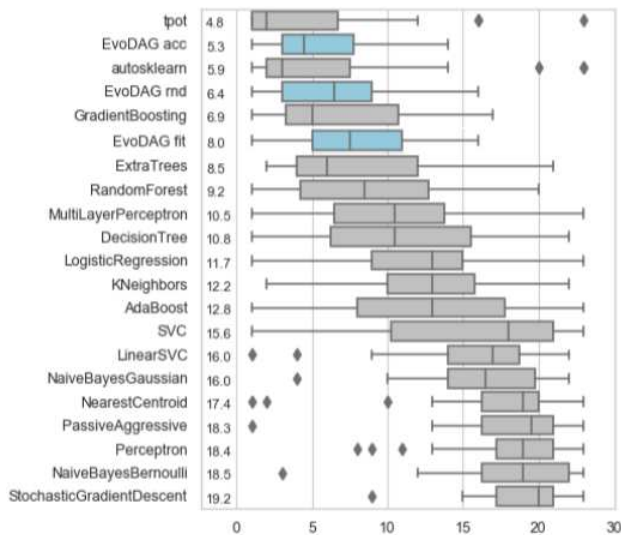


Figura 2. Resultados de diferentes clasificadores en base a la métrica macro-F1. .

Conclusiones

En esta investigación se presentan tres heurísticas de selección para PG basadas en los vectores semánticos de los individuos y en las propiedades de la función suma y los clasificadores Naive Bayes y Centroides más Cercano, las cuales han resultado mejores que la técnica clásica de selección basada en aptitud. Esto demuestra que es importante explorar propiedades en las funciones para diseñar operadores que mejoren el rendimiento de PG.*

REFERENCIAS

1. Koza, J., Bennett, F., Andre, D., Keane, M. y Dunlap, F. (1997). Automated synthesis of analog electrical circuits by means of genetic programming. *IEEE Transactions on Evolutionary Computation*, 1(2), 109-128.
2. Friedman, J. y Hall, P. (2007). On bagging and nonlinear estimation. *Journal of Statistical Planning and Inference*, 137(3), 669-683.
3. Loveard, T. y Ciesielski, V. (2001). Representing classification problems in genetic programming. En *Proc. 2001 Congress on Evolutionary Computation*, 1070-1077.
4. Muni, D., Pal, N. y Das, J. (2004). A novel approach to design classifiers using genetic programming. *IEEE Transactions on Evolutionary Computation*, 8(2), 183-196.
5. Graff, M., Tellez, E., Miranda-Jiménez, S. y Escalante, H. (2016). EvoDAG: A semantic genetic programming python library. En *Proc. 2016 IEEE International Autumn Meeting on Power, Electronics and Computing*, 1-6.
6. Olson, R., Urbanowicz, R., Andrews, P., Lavender, N. y Moore, J. (2016). Automating biomedical data science through tree-based pipeline optimization. En *Proc. European Conference on the Applications of Evolutionary Computation*, 123-137.
7. La Cava, W., Silva, S., Danai, K., Spector, L., Vanneschi, L. y Moore, J. (2019). Multidimensional genetic programming for multiclass classification. *Swarm and Evolutionary Computation*, 44, 260-272.
8. Moraglio, A., Krawiec, K., Johnson, C. (2012). Geometric semantic genetic programming. En *Proc. International Conference on Parallel Problem Solving from Nature*, 21-31.
9. Castelli, M., Silva, S. y Vanneschi, L. (2015). A C++ framework for geometric semantic genetic programming. *Genetic programming and evolvable machines*, 16(1), 73-81.
10. Feurer, M., Klein, A., Eggenberger, K., Springenberg, J., Blum, M. y Hutter, F. (2015). Efficient and robust automated machine learning. *Advances in neural information processing systems*, 2962-2970.

SOBRE LOS AUTORES



Claudia N. Sánchez es investigadora de la Universidad Panamericana campus Aguascalientes. Obtuvo su doctorado en Ciencias de Datos en INFOTEC Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación bajo la supervisión del Dr. Mario Graff. Sus áreas de interés son: aprendizaje de máquina, computación evolutiva, optimización, y visión computacional.



Mario Graff es investigador del Consejo Nacional de Ciencia y Tecnología (CONACYT) comisionado a INFOTEC. Miembro del Sistema Nacional de Investigadores Nivel I. Obtuvo su doctorado de la Escuela de Ciencias Computacionales e Ingeniería Electrónica de la Universidad de Essex trabajando bajo la supervisión del Profesor Riccardo Poli. Sus intereses son computación evolutiva, programación genética, aprendizaje de máquina, procesamiento de lenguaje natural y series de tiempo.

ARTÍCULO INVITADO

Un clasificador basado en transformaciones de espacio y selección de modelos

José Ortiz-Bejar, Eric S. Téllez y Mario Graff

Introducción

El problema de clasificación es una tarea de aprendizaje supervisado, la cual consiste en predecir una variable categórica a partir de un conjunto de elementos previamente etiquetados. Por ejemplo, determinar si un correo electrónico es o no spam, determinar si un paciente padece o no cierta enfermedad, identificar si un gato aparece o no en una imagen, clasificar si un tweet es agresivo o gracioso.

A los especialistas de las diferentes áreas les interesa obtener modelos que sean capaces de asignar la categoría correcta de manera automática a ejemplos nuevos. Por ejemplo, saber si el correo que acaba de llegar es o no spam. Una opción ampliamente utilizada son los modelos de aprendizaje de máquina (coloquialmente conocidos como clasificadores). Un clasificador automático es capaz aprender a categorizar ejemplos nuevos a partir de un conjunto de datos previamente etiquetado. Generalmente el conjunto de datos de dónde se aprende es etiquetado por expertos en el dominio que se aplica (e.g., médicos, psicólogos, lingüistas, biólogos, etc).

El resto del documento presenta una introducción al problema de clasificación de datos desde una perspectiva del aprendizaje computacional. Se presenta un clasificador novedoso basado en transformaciones de espacio y selección de modelos. Finalmente, una comparación del modelo propuesto con 16 clasificadores sobre 9 conjuntos de datos binarios.

Modelos de clasificación

Para ilustrar la idea de un modelo de clasificación, considere la Figura 1a, que es la representación gráfica de objetos con dos atributos. Por ejemplo, podrían representar la edad y el peso de una persona. Mientras que el color de la estrella o del triángulo representa la categoría, es decir, podría ser si padece o no una enfermedad, su nivel socio-económico, etcétera. Por otro lado, la frontera del clasificador (la línea que divide la zona de clasificación en dos colores) separa y agrupa los elementos según las clases en cuestión.

De forma breve podemos decir que el problema de clasificación consiste en determinar la frontera del modelo. El problema se vuelve complejo cuando los elementos a clasificar presentan un gran número y heterogeneidad de atributos, la densidad de los ejemplos por clase, el número de clases y el número de ejemplos, la calidad del etiquetado y muestreo, las fronteras no son especificables

de manera lineal, entre otras características. El lector interesado puede profundizar en la literatura especializada [1, 2, 5].

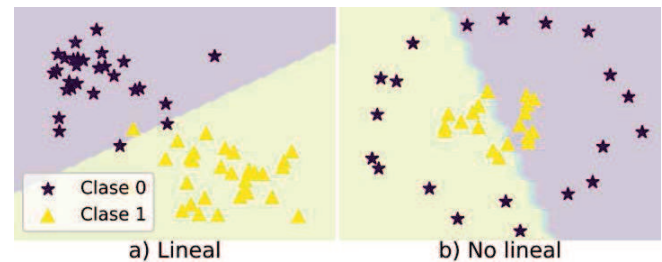


Figura 1. Problemas de clasificación.

En general, los modelos lineales requieren menos recursos computacionales, son simples de resolver e interpretar; sin embargo, muchos problemas requieren de modelos más complejos. Por ejemplo, en los datos en la Figura 1b; una frontera lineal en el mejor de los casos asignaría la mitad de los elementos a la clase incorrecta, lo cual sería similar clasificar los elementos lanzando una moneda al aire.

Una estrategia ampliamente utilizada para reducir las limitaciones de los métodos lineales es incrementar el número de atributos. En este sentido una de las técnicas más utilizadas son los métodos basados en funciones (conocidas como *núcleos* o *kernels* [1, 2]). Estas funciones proyectan los datos en dimensiones mayores donde es posible encontrar un modelo lineal eficiente. Por ejemplo, para el caso de los datos dispuestos en dos círculos concéntricos es posible hacerlo proyectando los datos en tres dimensiones como en la Figura 2b o bien mediante el uso directo de clasificadores no lineales (ver Figura 2a).

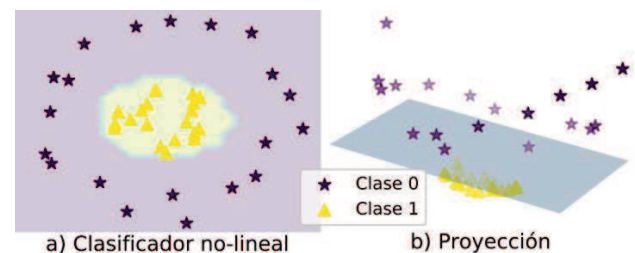


Figura 2. Enfoques de solución de problemas de clasificación no linealmente separables.

Un clasificador basado en proyecciones y selección de modelos

En este trabajo [3,4], se parte de la premisa de que aprovechando la distribución de los datos y un clasificador no lineal simple, es posible crear una proyección que facilite la tarea de clasificación. La propuesta está basada en la exploración de un espacio de modelos extenso para seleccionar los que mejor se adapten a una tarea determinada. El esquema resultante es denominado como **KMS**, del inglés *Kernel Model Selection*.

El espacio de modelos está definido por todas las combinaciones de diferentes conjuntos de estructuras, entre los cuales se encuentran: funciones para medir qué tan similares son dos elementos, funciones para determinar un subconjunto de elementos que caractericen a los ejemplos etiquetados, funciones kernel que proyectan los datos en un nuevo espacio, y clasificadores internos que funcionen sobre el espacio proyectado. El método propuesto intenta adaptarse a cada problema seleccionando elementos, que mediante el uso de un kernel pueda generar una proyección que contribuya a separar las clases, i.e., como la de la Figura 2b. Para una descripción precisa del método ver [3,4].

En este artículo, se muestra una comparación de KMS con 16 clasificadores disponibles en la librería **scikit-learn** de Python [6], sobre 9 bases de datos de clasificación binaria. Dichas bases de datos son ampliamente usadas en la literatura, ver [3,4,7] para una descripción más detallada del origen y distribución de los datos. Estos datos son llamadas “conjuntos de caracterización” en [3,4].

Se generó un ranking basado en el desempeño que cada clasificador tiene sobre cada uno de los 9 problemas. Para evaluar el desempeño de los distintos modelos, se utilizó la *razón balanceada del error*, BER por sus siglas en inglés. Dado un conjunto de prueba etiquetado y las predicciones dadas por un modelo para estos ejemplos, BER está definido como el promedio de la razón de errores por clase. De manera más detallada:

$$BER = \frac{1}{\#clases} \sum_{c \in \text{clases}} \frac{\# \text{ errores de predicción en } c}{\# \text{ elementos verídicos en clase } c};$$

BER tiene la ventaja de que considera el efecto de un conjunto de datos desbalanceado.

La Figura 3 muestra el ranking resultante de cada uno de los modelos. El ranking se calculó como el promedio de la posición de cada clasificador en cada uno de los 9 problemas evaluados, a menor BER mejor posición. La gráfica de caja indica el rango de posiciones que ocupó cada clasificador; siendo KMS el modelo con el mejor desempeño estando entre el primer y noveno lugar. El segundo mejor fue la máquina de soporte vectorial con kernel gaussiano (SVM RBF).

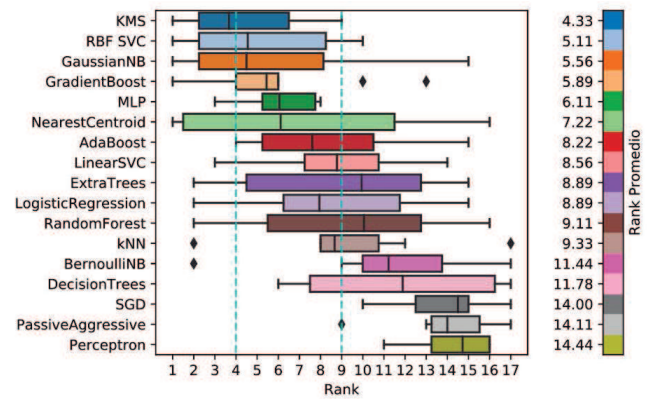


Figura 3. Comparativa de KMS con 16 clasificadores del estado del arte.

Conclusiones

En este trabajo se realizó una comparación contra 16 clasificadores ampliamente utilizados sobre 9 conjuntos de datos de clasificación binaria utilizados en la literatura del área de aprendizaje automático. En [3,4] se presenta un análisis extendido sobre más de 30 conjuntos de datos y 17 diferentes clasificadores. La comparación consiste en promedios de ranking que muestran que KMS es una alternativa competitiva con otros enfoques de clasificación del estado del arte. El lector interesado puede consultar algunas aplicaciones de KMS en [3, 8, 9].*

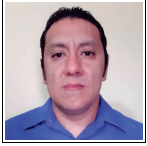
REFERENCIAS

1. Cristianini, N. y Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. *Cambridge university press*.
2. James, G., Witten, D., Hastie, T. y Tibshirani, R. (2013). An introduction to statistical learning. *Springer*.
3. Ortiz-Bejar, J. (2020). A Family of Classifiers based on Feature Space Transformations and Model Selection. Tesis de doctorado, Doctorado en Ciencias en Ciencia de Datos, INFOTEC Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación, Sede Aguascalientes.
4. Ortiz-Bejar, J., Téllez, E.S., Graff, M., et al. (2020). Improving k Nearest Neighbors and Naïve Bayes classifiers through space transformations and model selection. *IEEE Access*, 8, 221669-221688.
5. García-Gil, D., Luengo, J., García, S. y Herrera, F. (2019). Enabling smart data: noise filtering in big data classification. *Information Sciences*, 479, 135-152.
6. Pedregosa, F., Varoquaux, G., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
7. Rätsch, G., Onoda, T. y Müller, K.R. (2001). Soft margins for AdaBoost. Recuperado el 30 de agosto de 2017, de <http://theoval.cmp.uea.ac.uk/matlab/default.html>.
8. Pellegrin, L., Loyola-González, O., Ortiz-Bejar, J., et al. (2019). Image Annotation as Text-Image Matching: Challenge Design and Results. *Computación y Sistemas*, 23(4), 1305-1321.

9. Ortiz-Bejar, J., Tellez, E.S., Graff, M., et al. (2017). I3GO+ at RICATIM 2017: A semi-supervised approach to determine the relevance between images and text-annotations. En *IEEE International Autumn Meeting on Power, Electronics and Computing*, 1-6.

10. Haixiang, G., et al. (2017). Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications*, 73, 220-239.

SOBRE LOS AUTORES



José Ortiz-Bejar obtuvo el grado de doctor en Ciencias en Ciencia de Datos en INFOTEC Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación en 2020. Desde 2007 es académico en la Universidad Michoacana de San Nicolás de Hidalgo. Sus intereses de investigación incluyen clustering y aplicaciones de machine learning a problemas de categorización de texto y sistemas eléctricos.



Eric S. Tellez obtuvo el grado de Doctor en Ciencias en Ingeniería Eléctrica de la Universidad Michoacana de San Nicolás de Hidalgo en 2012. Desde 2014 es Catedrático CONACyT asignado a INFOTEC Aguascalientes. Es miembro del Sistema Nacional de Investigadores nivel I (2019-2021). Sus intereses de investigación incluyen búsqueda por similitud, inteligencia computacional, y clasificación automática de texto.



Mario Graff obtuvo el grado de Doctor en Filosofía por la Universidad de Essex en 2011. Desde 2014 es Catedrático CONACyT asignado a INFOTEC Aguascalientes. Es miembro del Sistema Nacional de Investigadores nivel I (2018-2021). Sus intereses de investigación incluyen programación genética, inteligencia computacional, y clasificación automática de texto.

***Scikit-Learn** es una biblioteca de Python de código abierto que implementa una variedad de algoritmos de aprendizaje automático, preprocesamiento, validación cruzada y visualización mediante una interfaz unificada.*

Para saber más consulte: <https://scikit-learn.org>

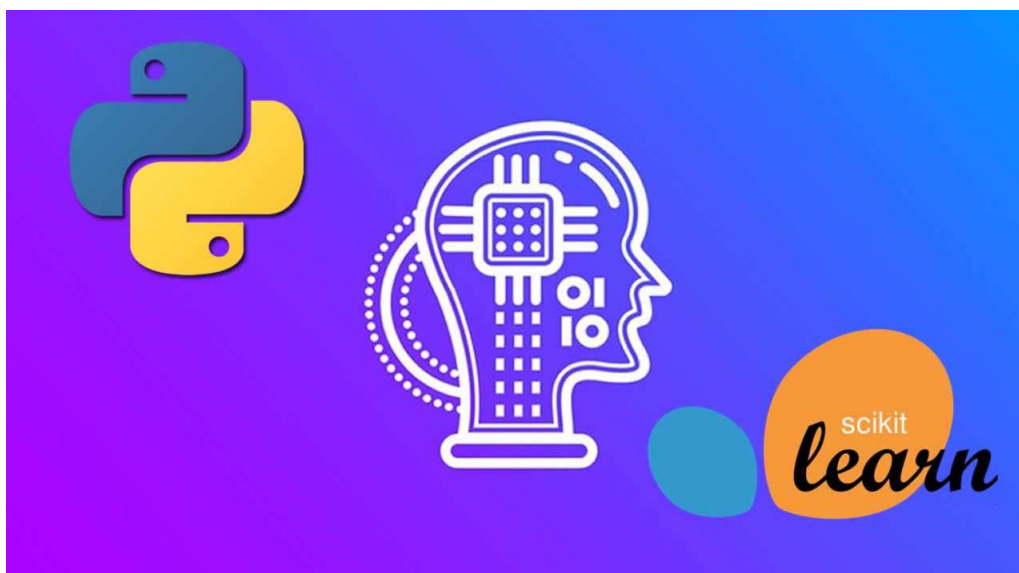


Imagen tomada de: <https://i.morloh.com>

ARTÍCULO INVITADO

Redes Neuronales Convolucionales para Pilotaje Autónomo en Carrera de Drones

Leticia Oyuki Rojas Pérez y José Martínez Carranza

DeepPilot es piloto artificial basado en redes neuronales convolucionales que controla a un dron para volar de manera autónoma a través de una pista de carreras.

Las carreras de drones se han convertido en un deporte popular, ya que muestran la capacidad humana para controlar un vehículo en un entorno altamente dinámico utilizando sólo información visual. En las carreras de drones cada vehículo es controlado de manera remota por un piloto, quien recibe una transmisión en vivo desde una cámara a bordo del vehículo. El vehículo debe recorrer una pista compleja de manera exitosa lo más rápido posible y para ello el piloto requiere años de entrenamiento. Este deporte ha motivado a competencias de carrera de drones autónomos o en inglés Autonomous Drone Racing (ADR), cuyo objetivo es desarrollar un piloto artificial, es decir que un dron pueda navegar con éxito en una pista de carreras mediante la interpretación continua de su entorno sin intervención humana.

El ADR plasma algunos de los problemas centrales destacados en robótica e inteligencia artificial, puesto que el desarrollo de un dron totalmente autónomo presenta diversos desafíos que abarcan el modelado dinámico, la percepción visual, la generación de trayectorias, el control, la localización y mapeo. La primera competencia de carrera de drones autónomos, fue lanzada en el IEEE IROS 2016 [1], y continuó organizándose en IROS 2017 [2], 2018 y 2019. Posteriormente se lanzaron otras competencias internacionales: el AlphaPilot [3], organizado por Lockheed Martin en colaboración con la Drone Racing League, y Game of Drones (GoD) [4], organizado por Microsoft y la Universidad de Stanford. Un rasgo distintivo de las competencias IROS ADR y AlphaPilot es que los drones competidores deben realizar el procesamiento a bordo. En la Figura 1 se muestran los circuitos de carreras de cada una de las ediciones del ADR en IROS. También se muestra una escena de la competencia de AlphaPilot y por último el entorno virtual de la competencia Game of Drones.

Actualmente en las competencias de ADR, el dron vuela a través de puertas cuadradas color naranja en una secuencia predefinida. El dron que vuele a través de más

puertas gana la carrera, y si dos drones llegan a la misma puerta o completan la pista, el que realizó el menor tiempo es el ganador. Por ello, a lo largo de estos años, los equipos participantes han propuesto diferentes soluciones con un conducto común: detección de puertas, localización de drones y control de vuelo estable. Recientemente, se ha utilizado el aprendizaje profundo o en inglés Deep Learning (DL) para la detección de la puerta y la localización del dron respecto a la puerta. Sin embargo, en competencias recientes como Game of Drones, celebrado en NeurIPS 2019, se han propuesto soluciones en las que el DL desempeña un papel más significativo.



Figura 1. Competencias internacionales de Carreras de Drones.

Motivados por esto último, mencionaremos tres artículos basados en redes neuronales convolucionales o en inglés Convolutional Neural Network (CNN): "A Temporal CNN-based Approach for Autonomous Drone Racing" [5], "DeepPilot: A CNN for Autonomous Drone Racing" [6] y "Autonomous Drone Racing with an Opponent: A First Approach" [7]; los cuales demuestran que es posible desarrollar un piloto artificial, mediante el uso de un conjunto de datos. Estos datos permiten que un dron aprenda los 4 comandos de vuelo, a partir de un vuelo

manual realizado por un piloto en una pista de carreras. En la que a cada imagen se asocia un grupo de valores, los cuales corresponden a roll (alabeo), pitch (cabeceo), yaw (orientación) y altitud (altura). Cabe señalar que este conjunto de datos no tiene una detección explícita de las puertas. Las tres redes presentadas en los artículos antes mencionados, toman las imágenes de la cámara del dron como entrada y predicen los comandos de vuelo como salida. Estos comandos de vuelo representan: la posición angular de la estructura del cuerpo del dron en los ángulos de roll y pitch, produciendo así un movimiento de traslación en esos ángulos; la velocidad de rotación en el ángulo de yaw; y la velocidad vertical referida a la altitud, como la muestra la Figura 2. Los valores de estos 4 comandos de vuelo, predichos por la CNN, se pasan al controlador interno del dron, permitiendo así al dron navegar de forma autónoma a través de las puertas de la pista de carreras. Para ello, se asume que la siguiente puerta se hace visible inmediatamente después de que la actual haya sido cruzada.

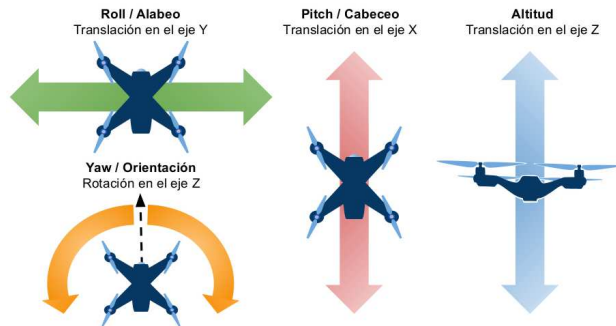


Figura 2. Comandos de vuelo roll, pitch, yaw y altitud. Los cambios en cada uno de los ángulos produce los siguientes movimientos: roll o alabeo, traslación en el eje Y; pitch o cabeceo, traslación en el eje X; yaw u orientación, rotación en el eje Z; y altitud el desplazamiento vertical sobre el eje Z.

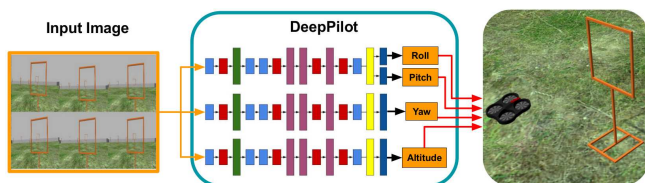


Figura 3. Diagrama general del sistema de control de la red DeepPilot.

En la Figura 3, se muestra el diagrama general del sistema control del artículo "DeepPilot: A CNN for Autonomous Drone Racing" [6], el cual ha sido implementado para guiar de manera autónoma a un dron a través de la pista de carreras. En este sistema de control el vehículo es controlado totalmente por la CNN y no utiliza información de sensores para mantener la altura y la

orientación, es decir que el vehículo reacciona a situaciones que ha aprendido previamente mediante un conjunto de ejemplos. En este trabajo se presentan evaluaciones en entornos simulados donde la CNN se ejecuta varias veces con éxito para probar la repetibilidad de la solución propuesta. También se presenta una evaluación completa de un enfoque temporal, que consiste en la creación de una imagen de mosaico, con fotogramas consecutivos de la cámara, que se pasa como entrada a la CNN.

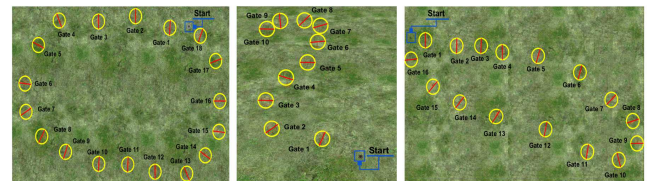


Figura 4. Vista Superior de las pistas de carreras simuladas para evaluar la red DeepPilot. Enlace al vídeo: <https://www.youtube.com/watch?v=Qo48pRCxM40>.

En la Figura 4, se muestra una vista superior de las pistas de carreras con las que DeepPilot fue evaluado, en donde las puertas tienen diferentes alturas y variaciones de orientación, a diferencia del trabajo "A Temporal CNN-based Approach for Autonomous Drone Racing"[5]; en el cual las puertas tienen una altura fija y la orientación siempre es paralela a la cámara del dron. Ambos trabajos sostienen que el enfoque temporal ayuda a aprender la tendencia de movimiento del dron con respecto a la puerta, actuando así como una memoria local que aprovecha la predicción de los comandos de vuelo. Los resultados reportados indican que este piloto artificial basado exclusivamente en DL es factible para ser usado en el desafío ADR.

Dados los resultados de la comunidad científica surge la pregunta ¿Un piloto artificial podrá competir contra uno o más adversarios en la misma pista de carreras?. Para dar respuesta a esto la competencia Game of Drones planteó un par de tareas en las que se mencionaba un oponente. Sin embargo, los competidores se centraron sólo en completar la pista lo más rápido posible, dejando al oponente siempre atrás desde el principio. Por lo tanto no se mostró interacción entre los competidores.

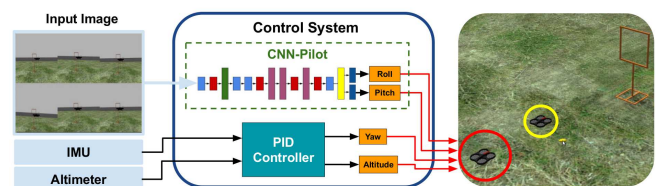


Figura 5. Diagrama general del sistema de control de la red CNNPilot. Enlace al vídeo: <https://www.youtube.com/watch?v=0KJg4Jy71K4>.

Motivado por esto en el artículo "Autonomous Drone Racing with an Opponent: A First Approach"[7], se presenta una primera propuesta de un piloto artificial que imita la toma de decisiones de un piloto humano ante la presencia de un adversario. En la Figura 5, se muestra el diagrama general del sistema de control. En el cual se usa una CNN llamada CNNPilot para controlar los comandos de control en roll y pitch. Además se mantiene fija la altura y la orientación. En este trabajo se presenta una evaluación en una pista con cinco puertas colocadas en zig-zag. Entre cada puerta se encuentra un adversario estático que obstruye el paso del dron para cruzar la puerta más próxima. Los resultados reportados demostraron que CNNPilot era capaz de controlar el dron para que volara de forma autónoma hacia la siguiente puerta y predecir los comandos de vuelo correspondientes para evadir al oponente, cuando este estuviera en su camino. Después de la evasión, el sistema de control era capaz de poner al dron en camino para navegar hacia la siguiente puerta. Finalmente, con este artículo queremos resaltar que el DL y las CNN son técnicas prometedoras para el desarrollo de un piloto artificial. Además el desarrollo de estos algoritmos pueden tener utilidad en aplicaciones civiles, por ejemplo: búsqueda y rescate, entrega de paquetes, inspección de infraestructura, inspección de bodegas, no solo entretenimiento.*

REFERENCIAS

1. Moon, H., Sun, Y., Baltes, J. y Kim, S. J. (2017). The IROS 2016 competitions [competitions]. *IEEE Robotics and Automation Magazine*, 24(1), 20-29.
2. Moon, H., Martinez-Carranza, J., Cieslewski, T., Faessler, M., Falanga, D., Simovic, A., y de Croon, G. (2019). Challenges and implemented technologies used in autonomous drone racing. *Intelligent Service Robotics*, 12(2), 137-148.
3. LockheedMartin (2019). Lockheed Martin And Drone Racing League Launch Groundbreaking AI Innovation Challenge. Recuperado el 9 de Noviembre de 2020, de <https://news.lockheedmartin.com/2018-09-05-Lockheed-Martin-and-Drone-Racing-League-Launch-Groundbreaking-AI-Innovation-Challenge>.
4. Escalante, H. J., y Hadsell, R. (2020, Agosto). NeurIPS 2019 Competition and Demonstration Track: Revised selected papers. En *NeurIPS 2019 Competition and Demonstration Track*, 1-12.
5. Rojas-Perez, L.O. y Martinez-Carranza, J. (2019, Noviembre). A temporal cnn-based approach for autonomous drone racing. En *2019 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED UAS)*, IEEE, 70-77.
6. Rojas-Perez, L.O. y Martinez-Carranza, J. (2020, Agosto). Deeppilot: A CNN for Autonomous Drone Racing. *Sensors*, 20, 4524.
7. Rojas-Perez, L.O., Martinez Carranza, J. (2020). Autonomous drone Racing with an Opponent: A First Approach. *Computación y Sistemas*.

SOBRE LOS AUTORES



Leticia Oyuki Rojas Pérez es estudiante de doctorado en la Coordinación de Ciencias Computacionales en el Instituto Nacional de Astrofísica Óptica y Electrónica. Obtuvo su grado de maestra en Ciencias en Ciencias de la Computación en el Instituto Nacional de Astrofísica Óptica y Electrónica en 2020. Sus intereses de investigación incluyen robótica e inteligencia artificial.



José Martínez Carranza es Investigador Titular B en la Coordinación de Ciencias Computacionales en el Instituto Nacional de Astrofísica Óptica y Electrónica. En 2018 fue distinguido como Honorary Senior Research Fellow en el Departamento de Ciencias Computacionales en la Universidad de Bristol y en 2015 ganó la Newton Advanced Fellowship, otorgada por la Royal Society, ambas instituciones en el Reino Unido. Es miembro del Sistema Nacional de Investigadores Nivel I. Su principal tema de investigación es sobre drones inteligentes, para lo cual desarrolla técnicas basadas en visión computacional, robótica e inteligencia artificial.

ARTÍCULO INVITADO

Análisis de sentimientos y emociones en redes sociales

Navonil Majumder, Alexander Gelbukh y Soujanya Poria

Introducción

Internet es uno de los ejes de la tecnología moderna. Varios aspectos de nuestras actividades cotidianas dependen de Internet –los viajes (Uber), compras (Amazon), entretenimiento (Netflix, YouTube), educación (Coursera), finanzas (banca por internet), etc.

Los servicios de comercio electrónico, como Amazon, han permitido a las personas con acceso a Internet hacer compras con solo presionar un botón, ahorrando horas para productividad y ocio. La proliferación de dispositivos inteligentes ha permitido que plataformas de redes sociales, como Facebook, Instagram y YouTube tengan gran influencia en la opinión pública. Los usuarios publican en estas plataformas sus opiniones sobre los temas importantes como política, deportes, alimentación, productos, etc. Estas opiniones son compartidas en gran cantidad, cada día, en forma textual o audiovisual.

Esta enorme cantidad de datos está siendo aprovechada por grandes empresas para impulsar sus ventas e ingresos. Los servicios de comercio electrónico, como Amazon, hacen recomendaciones a los clientes basadas en sus compras anteriores y sus opiniones sobre esas compras. Los diseñadores de productos a menudo quieren que el usuario retroalimente sobre varios aspectos de sus productos para mejorarlos o reducir sus inconvenientes. Esto ayuda a las empresas en la toma de decisiones cruciales basadas en las opiniones de sus usuarios.

En las plataformas como Facebook o Twitter, los usuarios a menudo interactúan entre sí expresando sus opiniones. Esto conduce a una conversación entre varias partes. El procesamiento de todas estas conversaciones, multimodales o textuales, requiere de algoritmos escalables para extraer información afectiva de ellas.

Las tareas principales

Los problemas que se discuten en este artículo son el análisis de los sentimientos y la detección de las emociones. Los métodos que se presentan son aplicables a ambos problemas, ya que están estrechamente relacionados. La tarea de la detección de las emociones es más precisa que la detección de los sentimientos. Además, exploramos la extensión multimodal de ambos problemas.

La tarea del análisis de sentimientos consiste en asignar una etiqueta de opinión adecuada a un fragmento de texto, ya sea un enunciado, párrafo o documento. Normalmente son usadas tres etiquetas: POSITIVO (“la pantalla es preciosa”), NEGATIVO (“el teléfono es demasiado

voluminoso”) y NEUTRO (“está alimentado por un Intel Core procesador i5”).

Similar a la clasificación del sentimiento, la clasificación de emociones es la asignación de una etiqueta de emoción apropiada a un fragmento de texto, ya sea un enunciado, párrafo o documento, por ejemplo: FELIZ (“*hoy conseguí un gran aumento junto con la promoción*”), TRISTE (“*es una pena que John se lastimara el tobillo antes de la expedición*”), NEUTRO (“*DHL es un servicio de mensajería*”), ENOJADO (“*¡no vuelvas a llamar aquí!*”), EXCITADO (“*tengo que terminar esta tesis lo antes posible*”), FRUSTRADO (“*terminar este borrador está tomando más tiempo del que esperaba*”).

A menudo, un solo enunciado u oración contiene más de un sentimiento, de múltiples aspectos del objeto. Como tal, la extracción de esos aspectos y su correspondiente sentimiento son dos tareas diferentes. Podemos ilustrar el análisis de sentimientos basado en aspectos de la siguiente manera: el texto “*Frenchie’s tiene comida mediocre, pero excelente servicio*” resulta en la salida “*comida: NEGATIVO, servicio: POSITIVO*”.

A diferencia de la clasificación del sentimiento o emociones en el texto, la clasificación multimodal se ocupa de la información visual y acústica, además de la información textual. Es particularmente importante para la extracción del sentimiento o emociones de los vídeos o sonido, donde las modalidades adicionales pueden ayudar a la clasificación. El principal desafío de cualquier tarea multimodal es la fusión de modalidades.

La clasificación de emociones y sentimientos conversacionales proporciona información sobre el hablante, así como cada turno particular del hablante, para ayudar a la clasificación.

Algunos métodos recientes

La fusión de información de múltiples modalidades es el proceso más crucial en el análisis de sentimiento multimodal. Recientemente han sido propuestos numerosos esquemas de fusión multimodal [1, 7, 8, 9]. Estos métodos sólo fusionan o codifican las características de la modalidad en una función unificada. Un paso más allá es mapear o decodificar la función fusionada para recuperar las funciones unimodales originales, lo que mejora la retención de la información y obliga al modelo de fusión a retener las características específicas e invariantes de la modalidad al mismo tiempo. Dado que esta estrategia tiene distintas partes de codificador y decodificador, se

puede utilizar como codificador para diseñar aún mejores estrategias de la fusión [4].

La mayoría de los algoritmos para el análisis de sentimientos basado en aspectos, lo que no siempre es correcto. Por ejemplo, “*su comida era mucho mejor que su servicio”* tiene dos aspectos, comida y servicio. Debido al uso de la palabra *mejor*, el sentimiento sobre comida depende del sentimiento sobre servicio y viceversa. Por lo tanto, se emplean las redes de memoria [6] para modelar la dependencia entre los aspectos dentro de la misma oración, lo que resulta en rendimiento mejorado sobre el estado del arte [2].

En cuanto a la clasificación de emociones y sentimientos en las conversaciones entre dos o varias personas, la mayoría de los trabajos existentes no utilizan la información del hablante de forma eficaz y escalable. Hemos propuesto métodos basados en las redes neuronales recurrentes que perfilan al hablante individual mientras la conversación avanza. Como tal, son capaces de manejar conversaciones multipartitas en una forma escalable. Así como de manejar un número arbitrario de oradores en una conversación sin volver a entrenar al modelo [3].

Conclusión

El análisis de sentimientos y emociones brinda grandes beneficios a las empresas y a los usuarios, permitiendo mejoras constantes a los productos y servicios a través

de la retroalimentación de sus usuarios. Este campo está evolucionando rápidamente y hay mucho en él más allá de lo que pudimos mencionar [5].*

REFERENCIAS

1. Majumder, N., Hazarika, D. et al. (2018). Multimodal sentiment analysis using hierarchical fusion with context modeling. *Knowledge-Based Systems*, 161:124–133.
2. Majumder, N., Poria, S. et al. (2018). IARM: Inter-aspect relation modeling with memory networks in aspect-based sentiment analysis. En *EMNLP 2018*, 3402–3411.
3. Majumder, N., Poria, S. et al. (2019). DialogueRNN: An attentionrnn for emotion detection in conversations. En *AAAI 2019*, 33, 6818–6825.
4. Majumder, N., Poria, S. et al. (2019). Variational fusion for multimodal sentiment analysis. abs/1908.06008.
5. Poria, S., Hazarika, D. et al. (2020). Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research. En *IEEE Transactions on Affective Computing*.
6. Sukhbaatar, S., Szlam, A. et al. (2015). End-to-end memory networks. En *NIPS 2015*, 2440–2448.
7. Zadeh, A., Chen, M. et al. (2017). Tensor fusion network for multimodal sentiment analysis. En *EMNLP 2017*, 1103–1114.
8. Zadeh, A., Liang, P.P. et al. (2018). Memory fusion network for multi-view sequential learning”. En *AAAI 2018*, 5634–5641.
9. Zadeh, A., Liang, P.P. et al. (2018). Multi-attention recurrent network for human communication comprehension. En *AAAI 2018*.

SOBRE LOS AUTORES



Navonil Majumder obtuvo los grados de Maestro y Doctor en Ciencias de la Computación en el Centro de Investigación en Computación (CIC) del Instituto Politécnico Nacional (IPN) en 2017 y 2020, respectivamente. Fue galardonado con la Presea Lázaro Cárdenas dos veces, por sus estudios de Maestría y Doctorado. Actualmente es un investigador en la Universidad de Tecnología y Diseño de Singapur (SUTD). Sus intereses de investigación abarcan el procesamiento del lenguaje natural, la comprensión del diálogo, el análisis de sentimientos y el procesamiento multimodal del lenguaje.



Alexander Gelbukh obtuvo el grado de Doctor en Ciencias de la Computación en el VINITI, Rusia. Actualmente es profesor-investigador del Centro de Investigación en Computación (CIC) del Instituto Politécnico Nacional (IPN), México y profesor honorífico de la universidad Amity, India. Es miembro de la Academia Mexicana de Ciencias, Investigador Nacional de México (miembro del SNI) de nivel 3 (el más alto). Sus intereses científicos incluyen el procesamiento del lenguaje natural, la lingüística computacional y la inteligencia artificial.



Soujanya Poria obtuvo el grado de Doctor en Ciencias de la Computación en la Universidad de Stirling, Reino Unido. Sus principales áreas de interés de investigación son el procesamiento del lenguaje natural y el análisis de sentimientos. Actualmente es profesor asistente en la Universidad de Tecnología y Diseño de Singapur (SUTD), forma parte del Instituto de Computación de Alto Rendimiento (IHPC), ASTAR, como investigador senior. Antes de unirse a SUTD, trabajó en la Universidad Tecnológica Nanyang (NTU), donde obtuvo la prestigiosa beca postdoctoral presidencial de la NTU.

ARTÍCULO INVITADO

Detección de emociones en los medios sociales

Muhammad Hammad Fahim Siddiqui y Alexander Gelbukh

La detección de emociones en los medios y redes sociales tiene aplicaciones en los ámbitos de la inteligencia de negocios, salud y seguridad. Las arquitecturas neuronales modernas del aprendizaje profundo ofrecen excelentes resultados en esta tarea.

Introducción

Las tecnologías de información actuales nos brindan a todos la oportunidad de expresar nuestros sentimientos en varios medios, llamados plataformas sociales, tales como Facebook o Twitter. Estas plataformas sociales tienen cientos de millones de usuarios que interactúan entre sí e intercambian contenido en la forma de texto, audio y video. Este contenido contiene emociones expresadas por los usuarios, las cuales se desencadenan por varias razones. En una de nuestras investigaciones más recientes [7] hemos abordado el tema de la detección y clasificación de estas emociones en once categorías predefinidas.

La detección y clasificación de las emociones en los medios y redes sociales brinda grandes beneficios y tiene aplicaciones importantes [2]. En el ámbito comercial, es parte de la inteligencia de negocios y toma de decisiones, así como de los sistemas de recomendación. En el ámbito político y de seguridad pública, ayuda a los partidos políticos y al gobierno a entender mejor las opiniones y los sentimientos de los ciudadanos. Tiene también aplicaciones en el ámbito de educación y salud, para monitorear el estado emocional de los alumnos y pacientes.

Trabajos relacionados

La tarea de clasificación de emociones en procesamiento del lenguaje natural se define como la detección y clasificación de las emociones humanas, tales como miedo, tristeza, alegría, sorpresa, etc., en un texto escrito. Las aplicaciones de una tarea como esta van desde el análisis de las emociones del cliente de un negocio hasta la evaluación de los sentimientos de reacción de las personas después de una noticia política. Es una tarea popular de procesamiento del lenguaje natural, y se ha trabajado mucho al respecto [1].

Sin embargo, esta tarea, tal como la abordamos nosotros, tiene una naturaleza diferente. La investigación realizada previamente era, en su mayoría, vista como una clasificación binaria, es decir, la detección y clasificación de una sola emoción por el texto dado. También hay trabajos publicados previamente en los cuales la cla-

sificación se realiza en torno a un total de seis o siete emociones [3, 5]. A diferencia de estos trabajos, en nuestra investigación hemos ido más allá, a un total de once emociones, lo que convierte la tarea en un problema de clasificación de etiquetas múltiples (de mayor complejidad).

La solución que le hemos dado requirió del uso de los métodos más modernos del incruste de palabras (*word embedding* en inglés) en una arquitectura moderna de una red neuronal profunda con memoria bidireccional y mecanismos de atención, la cual usamos para la clasificación a categorías múltiples.

Los datos

El conjunto de datos disponible para la tarea fue proporcionado por la serie de competencias internacionales SemEval 2018, a saber, la tarea 1: detección del afecto en tuits [6], y consistió en un número considerable de tuits (más de diez mil). En cada tuit los organizadores de la competencia marcaron manualmente las emociones expresadas por el autor en este tuit. Nuestra tarea, entonces, consistió en entrenar un clasificador que desarrollamos con una parte de estos datos, y el desempeño de nuestro clasificador fue evaluado en otra parte del conjunto de datos.

Hallazgos y resultados

Para abordar este problema, hemos realizado una amplia gama de experimentos utilizando el aprendizaje profundo con las redes neuronales profundas con las atenciones y con el enfoque de transferencia de aprendizaje.

Se observó un aumento considerable de la precisión con el uso de una nueva técnica de preprocesamiento de texto, la cual incluía características estilísticas.

Dos de los enfoques utilizados en nuestra investigación superaron la mejor precisión (57.4 %) previamente conocida para esta tarea. Estos dos enfoques incluían un modelo con la memoria bidireccional Bi-LSTM con un marco de atención múltiple, con una atención por cada una de las once emociones (ver la Figura 1). Esta aproxi-

mación fue capaz de lograr una precisión de 58.1 %, con un 0.5 % de ganancia sobre el estado del arte.

El otro enfoque utilizaba el aprendizaje por transferencia, con el entrenamiento del modelo moderno de incrustes contextuales RoBERTa [4] con nuestros datos. Este enfoque fue capaz de lograr una precisión del 61.2 %, equivalente a una ganancia del 3.8 % sobre el estado del arte. Con esto, nuestra investigación produjo un nuevo punto de referencia de última generación para esta tarea.

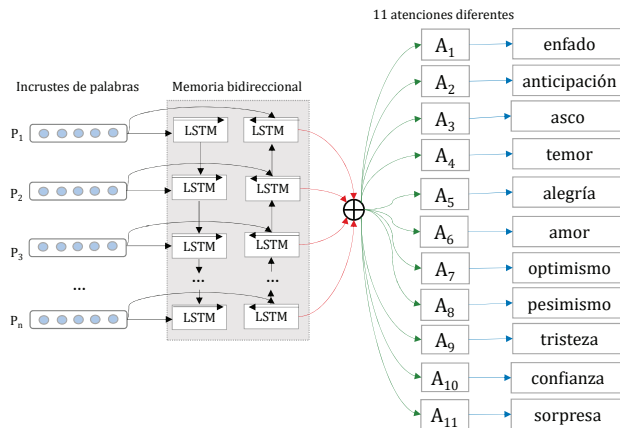


Figura 1. La arquitectura de la red neuronal profunda con la memoria bidireccional y atención múltiple usada para nuestros experimentos en clasificación de las emociones en el texto.

Conclusiones

La tarea de la detección de emociones en los medios y las redes sociales es de gran importancia para las aplicaciones en los ámbitos de los negocios y comercio electrónico, política y gobierno, seguridad, educación y salud, entre otros.

La detección de emociones con clasificación fina es un reto para los clasificadores automáticos. En nuestra investigación, hemos desarrollado dos métodos basados en el aprendizaje profundo, los cuales superaron el estado del arte actual en esta tarea. Estos métodos fueron basados en un preprocesamiento adecuado de texto, en una arquitectura moderna de la red neuronal profunda con memoria bidireccional y una atención por cada clase por separado, así como en los incrustes de palabras con el algoritmo RoBERTa [4] basado en transformadores neuronales [8].*

REFERENCIAS

1. Alswaidan, N., El Bachir Menai, M. (2020). A Survey of State-of-the-Art Approaches for Emotion Recognition in Text. *Knowledge and Information Systems*, 62, 2937–2987.
2. Cambria, E., Poria, S., Gelbukh, A., Thelwall, M. (2017). Sentiment Analysis Is a Big Suitcase. *IEEE Intelligent Systems*, 32(6), 74–80.
3. Ghosal, D., Majumder, N., Poria, S., Chhaya, N., Gelbukh, A. (2019). DialogueGCN: A Graph Convolutional Neural Network for Emotion Recognition in Conversation. *EMNLP-IJCNLP 2019*, 154–164.
4. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv:1907.11692.
5. Majumder, N., Poria, S., Hazarika, D., Mihalcea, R., Gelbukh, A., Cambria, E. (2019). DialogueRNN: An Attentive RNN for Emotion Detection in Conversations. *AAAI 2019*, 6818–6825.
6. Mohammad S B-M, Bravo-Marquez F, Salameh M, Kiritchenko S (2018) “SemEval-2018 Task 1: Affect in Tweets”. *SemEval 2018*, pp. 1–17.
7. Siddiqui, M.H.F. (2020). Social Media Emotion Detection. Tesis de Maestría. Instituto Politécnico Nacional, México, 60.
8. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I. (2017). Attention Is All You Need. arXiv:1706.03762.

SOBRE LOS AUTORES



Muhammad Hammad Fahim Siddiqui obtuvo el grado de Maestro en Ciencias de la Computación en el Centro de Investigación en Computación (CIC) del Instituto Politécnico Nacional (IPN) en 2020. Ha recibido el premio por mejor tesis de Maestría nacional sobre la inteligencia artificial (tercer lugar) de la Sociedad Mexicana de Inteligencia Artificial (SMIA). Actualmente está empezando sus estudios de Doctorado en la Universidad de Ottawa en Canadá. Sus intereses de investigación abarcan el procesamiento del lenguaje natural, preprocesamiento del texto, preprocesamiento del texto, clasificación del sentimiento y emociones y procesamiento del lenguaje natural con el aprendizaje profundo.



Alexander Gelbukh obtuvo el grado de Doctor en Ciencias de la Computación en el VINITI, Rusia. Actualmente es profesor-investigador del Centro de Investigación en Computación (CIC) del Instituto Politécnico Nacional (IPN), México y profesor honorífico de la universidad Amity, India. Es miembro de la Academia Mexicana de Ciencias, Investigador Nacional de México (miembro del SNI) de nivel 3 (el más alto). Sus intereses científicos incluyen el procesamiento del lenguaje natural, la lingüística computacional y la inteligencia artificial.

IA & Educación

Yasmín Hernández, Lucía Barrón y Julieta Noguez
iaeducacion@komputersapiens.org

Premio UNESCO para el uso de la inteligencia artificial en la educación

La inteligencia artificial ha propiciado el desarrollo de sistemas computacionales para resolver problemas de diversas áreas. De manera particular, en educación se han desarrollado herramientas para apoyar el aprendizaje de los estudiantes; sin embargo, la educación plantea desafíos importantes, tales como la personalización de la enseñanza y la mejora de los procesos de aprendizaje.

Entre las aplicaciones de inteligencia artificial (IA) en educación destacan los sistemas tutores inteligentes, los agentes pedagógicos animados, los juegos educativos, y las interfaces conversacionales, las cuales apoyan e impulsan la innovación en la educación y hacen el aprendizaje más personalizado y, por lo tanto, en principio, más efectivo. La IA provee herramientas y algoritmos para personalizar el aprendizaje a las necesidades particulares de los estudiantes con el fin de proporcionarles los medios precisos que les permitan vencer dificultades en el aprendizaje y avanzar en su educación.

La IA ha estado en la agenda educativa de diferentes países y de organizaciones internacionales desde hace algún tiempo. El Premio UNESCO Rey Hamad Bin Isa Al-Khalifa para el uso de las tecnologías de información y comunicación (TIC) en la educación distingue los enfoques innovadores en el uso de las nuevas tecnologías para ampliar las oportunidades educativas y de aprendizaje permanente para todos. Este premio se estableció en 2005 para galardonar a las personas y organizaciones que implementan proyectos destacados y promueven el uso creativo de tecnologías para mejorar el aprendizaje, la enseñanza y el desempeño educativo general en la era digital [1].

Cada año, el premio tiene un tema que aboga por el uso responsable y ético de las tecnologías de la información y la comunicación. En la edición 2019 del Premio, el tema fue el *uso de la IA para innovar la educación, la enseñanza y el aprendizaje*. En esta ocasión, la UNESCO estableció como objetivo el promover aplicaciones efectivas y éticas de la IA en la educación que estén disponibles de manera universal. En la edición 2019, fueron premiados el programa de habilidades de escritura Letrus de Brasil y la herramienta para detección de dislexia *DyTECTive* de España.

DyTECTive, desarrollado por la empresa *Change Dyslexia* de España, es una herramienta para detectar la dislexia en español. Este programa proporciona ejercicios

basados en juegos para estudiantes con dislexia, ayudándolos a superar las dificultades de escritura y lectura. La aplicación presenta una combinación de técnicas de gamificación y aprendizaje automático predictivo. Hasta la fecha, más de 270,000 niños en 43 países diferentes se han beneficiado de un diagnóstico de dislexia [2].

El desarrollo de *DyTECTive* es el resultado de un equipo de científicos de diversas áreas tales como inteligencia artificial, interacción humano-computadora, lingüística, pediatría, psicología y neuropsicología, entre otros. Este equipo ha desarrollado diferentes modelos con el fin de encontrar diferencias en el comportamiento de los niños, y así detectar la dislexia. Uno de estos se trata de un modelo estadístico para clasificar a lectores utilizando un clasificador con base en la técnica máquinas de soporte vectorial y datos de seguimiento ocular (*eye tracking*). Este estudio sugiere que el seguimiento ocular tiene el potencial de usarse para diagnosticar la dislexia en el futuro. Los autores argumentan que los rastreadores oculares son cada vez más asequibles y que leer un texto en silencio es menos intrusivo que las pruebas en los diagnósticos actuales [3]. Por otro lado, con el fin de mejorar la detección de la dislexia, también desarrollaron *MusVis*, un juego basado en web que mide la reacción de los niños durante el juego. *MusVis* tiene un componente visual y un componente auditivo. En la Figura 1 (izquierda) se muestra un ejemplo de la parte auditiva de *MusVis*. Se pide al participante que encuentre dos señales auditivas iguales haciendo clic en las tarjetas de sonido. En la Figura 1 (derecha) se muestra un ejemplo de la parte visual del juego. Adicionalmente, llevaron a cabo un estudio en el que recolectaron datos de la interacción con *MusVis*. Para construir el modelo predictivo procesaron estos datos con los algoritmos *Extra Trees*, *Random Forest* y *Gradient Boost* [4].

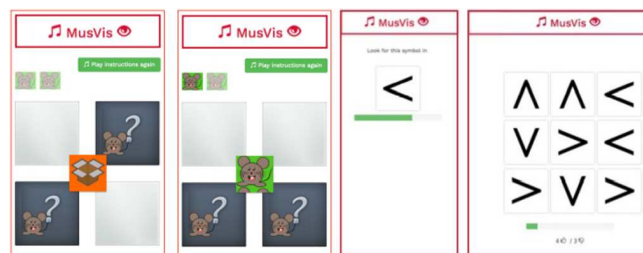


Figura 1. Parte auditiva (Izquierda) y parte visual (derecha) de *MusVis* [4].

El otro ganador de la edición 2019 del premio TIC en la educación, el programa *Letrus Writing Skills*, desarrollado por la empresa Letrus de Brasil, utiliza técnicas de IA y gamificación para la retroalimentación automática sobre la escritura de los estudiantes. *Letrus* tiene como objetivo reducir el analfabetismo funcional en el idioma portugués mejorando las oportunidades de practicar para los estudiantes y profesores de secundaria. Este programa identifica patrones de escritura y apoya las habilidades de escritura a través de la retroalimentación automática y humana. Los estudiantes reciben comentarios inmediatos del programa y comentarios más detallados de los evaluadores humanos. Desde 2017, más de 65,000 estudiantes han utilizado el programa en los 26 estados de Brasil [2]. En la Figura 2 se muestra a estudiantes interactuando con *Letrus*.



Figura 2. Estudiantes interactuando con *Letrus* [5].

Letrus, *Dytective* y *MusVis* son ejemplos de tecnologías educativas que adaptan los materiales educativos a las necesidades particulares de los estudiantes, monitorean desempeño, habilidades y avance; de esta manera predicen el desarrollo de los estudiantes, permitiendo que los estudiantes aprendan a su propio ritmo. Lo anterior

representa una gran oportunidad para quienes tienen necesidades de educación especial o requieren un aprendizaje personalizado.

El Premio UNESCO para el uso de TIC en la educación es muy importante para que las organizaciones y gobiernos consideren las ventajas de las tecnologías inteligentes en la educación y las pongan al alcance de la población, de tal manera que se vuelvan asequibles y sostenibles y puedan apoyar la educación en los retos del futuro. En la edición 2020 del premio, nuevamente se incluye a la IA con el tema: *el uso de la IA para mejorar la continuidad y la calidad del aprendizaje*, y en abril del 2021 se dará a conocer a los ganadores.*

REFERENCIAS

1. UNESCO (2020a). ICT in Education Prize. recuperado el 15 de diciembre 2020 de <https://en.unesco.org/themes/ict-education/ict-education-prize>.
2. UNESCO (2020b). Laureates from Brazil and Spain to receive UNESCO King Hamad Bin Isa Al-Khalifa Prize for their use of AI in education. recuperado el 15 de diciembre 2020 de <https://en.unesco.org/news/laureates-brazil-and-spain-receive-unesco-king-hamad-bin-isa-al-khalifa-prize-their-use-ai>.
3. Rello, L. y Ballesteros, M. (2015). Detecting Readers with Dyslexia Using Machine Learning with Eye Tracking Measures. En *proc. 12th Web for All Conference*, W4A '15. Florencia, Italia. pp. 1-8.
4. Rauschenberger, M., Baeza-Yates, R. y Rello, L. (2020). Screening Risk of Dyslexia through a Web-Game using Language-Independent Content and Machine Learning. En *proc. 17th International Web for All Conference*, W4A '20. Taipei, Taiwan. pp. 1-12.
5. UNESCO (2020c). Brazil programme awarded UNESCO Prize for using AI to improve writing skills. recuperado el 15 de diciembre 2020 de <https://en.unesco.org/news/brazil-programme-awarded-unesco-prize-using-ai-improve-writing-skills>.



Las TICs en Educación.

Imagen tomada de: <https://en.unesco.org/themes/ict-education/ict-education-prize>.

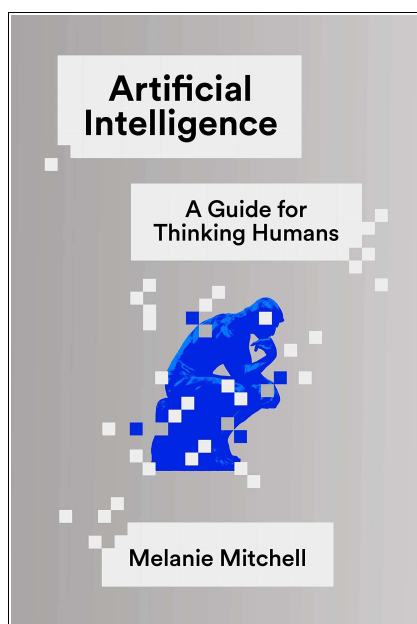
Deskubriendo Konocimiento

Alejandro Guerra Hernández y Leonardo Garrido
deskubriendokonocimiento@komputersapiens.org

Revisión de “Artificial Intelligence: A Guide for Thinking Humans” *

Dr. Alejandro Guerra-Hernández

Universidad Veracruzana
Centro de Investigación en Inteligencia Artificial



Portada de la revista.*

Desde el triunfo de Deep Blue sobre Kasparov en 1996, las victorias de los sistemas propuestos por la Inteligencia Artificial (IA) sobre sus oponentes humanos han sido ampliamente publicitadas. Conocidos son los casos de Watson ganando en 2011 el concurso televisivo de preguntas y respuestas Jeopardy; y AlphaGo derrotando a Lee Sedol en un encuentro de cinco partidas de Go en 2016. Ante estos acontecimientos, las reacciones de algunos de los involucrados resultan muy ilustrativas de nuestras preocupaciones. Ken Jennings, uno de

los campeones destronados de Jeopardy, declaró más tarde con humor —Al igual que Kasparov antes que yo, me gano la vida bastante bien con mi nueva carrera de perdedor humano profesional, misma que lo ha llevado a dar numerosas pláticas sobre su experiencia y ser el próximo conductor del concurso. Fan Hui, el campeón europeo de Go derrotado un año antes por AlphaGo, declaró al perder el segundo juego de su encuentro —No fue un movimiento humano. Nunca he visto a un humano hacer ese movimiento. Tan hermoso, hermoso, hermoso. IBM comunicó en su momento, incluso antes de que Watson ganara Jeopardy, que éste será capaz de responder a preguntas clínicas de manera mucho más eficiente que la mente humana. ¿Cuáles serán los efectos laborales de una IA con capacidades superiores a las humanas? ¿Puede una IA ser creativa? ¿Es Watson tan inteligente como para aplicar sus competencias en el dominio médico? ¿Estamos, como argumenta y promueve Ray Kurzweil desde Google, a las puertas de la Singularidad; o por el contrario, solo pasamos por una etapa más de triunfalismo desmedido, antesala de otro invierno de la IA?

Pues bien, Melanie Mitchell nos propone una guía para humanos pensantes que dé respuesta a la

cuestión de cual es el verdadero estado de la IA en estos momentos. ¿Qué pueden hacer las computadoras hoy? ¿Qué podemos esperar que hagan en las siguientes décadas? Ante tamaña empresa, es bueno saber que la autora cuenta con una amplia experiencia como investigadora en las áreas de visión computacional, algoritmos genéticos, razonamiento analógico, ciencias cognitivas y autómatas celulares; como docente, donde solo por citar un ejemplo, su curso de introducción a la complejidad en el Santa Fe Institute ha sido tomado por más de 25000 estudiantes; y como divulgadora, habiendo escrito previamente el premiado “*Complexity : AGuidedTour*” (Oxford University Press, 2011).

La guía está organizada en cuatro partes. La primera de ellas, Antecedentes, introduce al lector en las ideas que originaron a la IA en sus vertientes simbólica y sub-simbólica, con énfasis en el funcionamiento del perceptrón, cuyo entendimiento será necesario más adelante para introducir las redes neuronales y su éxito en las tareas mencionadas al principio de esta reseña. El resto de la primera parte introduce una serie de cuestiones que estarán presentes en el resto del libro: la inteligencia general versus la acotada; la cuestión de si una

*Mitchell, Melanie. Artificial Intelligence: A Guide for Thinking Humans. Farrar, Straus and Giroux. New York, NY, USA, 2019.

máquina puede pensar y la prueba de Turing; la singularidad de Kurzweil y el progreso exponencial en la computación. De manera que el lector pueda situarse en el contexto en el que se revisará cómo es que la IA ha conseguido reproducir habilidades tan propias de la inteligencia humana, como la percepción, el lenguaje, la toma de decisiones, el sentido común y el aprendizaje. Tres tecnologías son de especial interés y serán revisadas en las siguientes partes del libro: el aprendizaje profundo, los grandes datos y el creciente poder de cómputo.

La segunda parte, *Mirando y viendo*, comienza por dar cuenta de lo compleja y poderosa que es nuestra percepción visual y lo difícil que ha resultado emularla artificialmente, al menos hasta la revolución del aprendizaje profundo. Siguiendo la tónica de la primera parte, las redes neuronales convolucionales son introducidas en detalle, resaltando su poder de clasificación y su uso en el etiquetado de imágenes. Posteriormente, Mitchell se embarca en establecer sus limitaciones, a saber: la necesidad de diseñar su arquitectura manualmente; la gran cantidad de datos etiquetados por humanos necesarios para su entrenamiento; sus extrañas y poco humanas maneras de fallar debido a sobre-ajustes con los datos; su opacidad; y su vulnerabilidad a los llamados ejemplos de confrontación (falsificaciones indistinguibles al ojo humano, que provocan que la red se equivoque). Todo esto da lugar a una discusión sobre la necesidad de una IA ética y digna de confianza, que realmente aprenda por sí misma y, al igual que los humanos, sea más confiable por que puede razonar acerca de la situación en que se encuentra y planear para el futuro con grandes dosis de sentido común.

La tercer parte del libro, *Aprendiendo a jugar*, aborda el uso del aprendizaje por refuerzo profundo en juegos y simulaciones y cómo es que estas investigaciones nos han llevado a máquinas con capacidades de juego superiores a las humanas. Se presentan en detalle los desarrollos de DeepMind en los juegos de Atari, el ajedrez y el go, así como el noble origen de estos esfuerzos, en el jugador de damas chinas de Arthur Samuel. Mitchell reconoce el avance que el uso de una señal refuerzo supone con respecto a la necesidad de contar con millones de ejemplos etiquetados manualmente. A las limitaciones identificadas en las redes profundas, añade la dificultad de generalizar lo aprendido, el conocido problema de transferencia, y la dificultad inherente de entrenar a un agente en el mundo, fuera de un juego o una simulación.

La cuarta parte del libro, *La Inteligencia Artificial se encuentra con el Lenguaje Natural*, inicia con un ejemplo de las dificultades de la lectura de comprensión, el reconocimiento del habla y la clasificación de sentimientos. La tecnología revisada en esta parte son las redes neuronales recurrentes y la representación de las palabras como vectores que sirvan de entrada a las redes recurrentes. Posteriormente, se revisan los avances en traducción automática y los sistemas de pregunta-respuesta, tan evidentes en nuestros teléfonos, así como sus limitaciones.

La última parte del libro tiene que ver con lo que el filósofo Gian-Carlo Rota llama *La Barrera del Significado*, que captura una idea persistente a lo largo del libro: los humanos, en una forma profunda y esencial, entendemos la situación en la que nos encontramos; mientras que ningún sistema propuesto por la IA tiene esa capacidad aún. Es-

to se ve revelado por naturaleza tan inhumana de sus fallos, quizás tan sorprendente como la belleza sus soluciones. Mitchell considera algunas ideas para remediar estas limitaciones, la hipótesis del entendimiento como simulación y la relevancia de la abstracción y el razonamiento analógico, donde abunda un poco en su propio trabajo de investigación. Aunque el libro está dedicado a sus padres, bien podría estarlo a Douglas Hofstadter tan presente en la obra de Mitchell. De hecho, en homenaje a éste, la última parte cierra a la manera del Gödel, Escher, Bach, con una serie de preguntas, respuestas y especulaciones que incluyen: ¿Qué tan pronto los autos autónomos serán algo común? ¿La IA provocará desempleo masivo? ¿Puede una computadora ser creativa? ¿Qué tan lejos estamos de crear inteligencia general de nivel humano? ¿Qué tan atemorizado debería usted estar de la IA? ¿Cuáles son los grandes problemas que la IA aún no resuelve?

El libro es una lectura interesante para el público general que desea enterarse de los últimos avances de la IA y sus alcances. Es también una lectura imprescindible para quien se inicia seriamente en el estudio de esta disciplina y necesita entender mejor los límites de su disciplina. También resulta imprescindible, evidentemente, para quien toma decisiones sobre la adopción de tecnologías de la IA en sus actividades cotidianas y profesionales. De manera sorprendente, el texto es accesible y útil en igual medida, para estos tipos de lector. En un esfuerzo considerable, la renuncia al uso de formalismos a lo largo del texto, no repercute en su rigor, en su elegancia, ni en su profundidad.*



Durante la fase tres de la epidemia de COVID-19, el número de contagios crece de forma rápida, es por eso que debes quedarte en casa para reducir el número de casos que se pueden presentar.

El coronavirus es muy contagioso y, si hay más personas en las calles, seguramente habrá muchos enfermos que saturarían los hospitales.

¡No habría forma de ayudar a todos!

En cambio, si respetamos la sana distancia, habrá menos casos y más capacidad para que todos reciban atención.

Quédate en casa



Y no olvides:



Lavarte continuamente las manos con agua y jabón, al menos durante 20 segundos



Estornudar o toser en el ángulo interno del codo



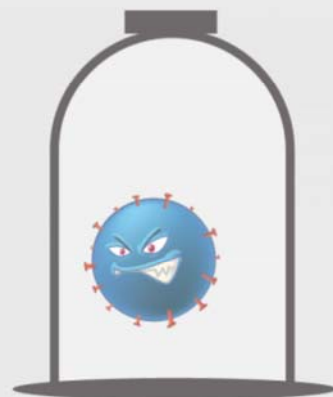
Mantener tu casa ventilada y limpia

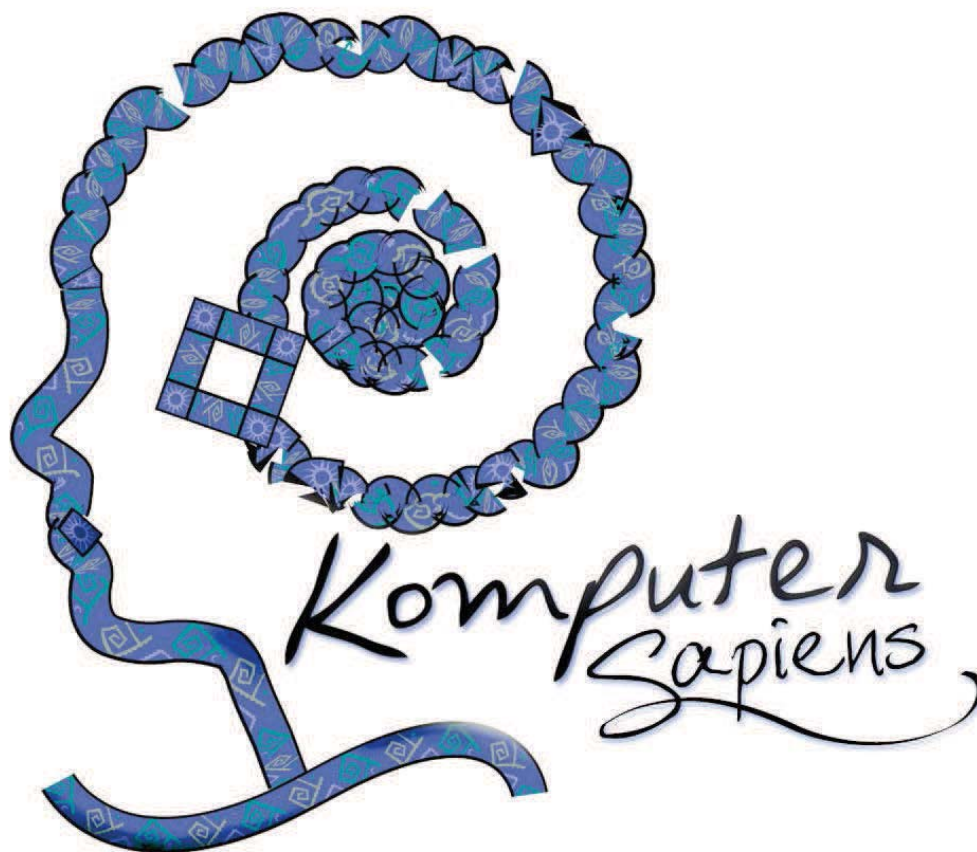


Desinfectar los utensilios y superficies de uso común

Quedándote en casa y cuidándonos unos a otros podremos salir adelante.

¡SI AISLAMOS AL VIRUS, LO VENCEREMOS!





¡Publique en Komputer Sapiens!



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



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

