

Proyecto 42: tras el “conductoma” humano

La Dra. Laura Vargas-Parada es Investigadora Biomédica Básica egresada del Instituto de Investigaciones Biomédicas de la UNAM. Es escritora de ciencia freelance y colaboradora de diversos medios nacionales e internacionales. Además de escribir le encanta la docencia impartiendo cátedra por casi dos décadas en la Facultad de Ciencias de la UNAM. Desde 2016 dirige la oficina de comunicación del Centro de Ciencias de la Complejidad (C3). Esta publicación fue revisada por el comité editorial de la Academia de Ciencias de Morelos.

“Voy a comenzar con una historia de ciencia ficción”, dijo Christopher Stephens a la audiencia, el pasado 28 de noviembre, al iniciar la ponencia donde presentó algunos de los resultados de su investigación dentro de la 3.ª Semana de la Complejidad organizada por el Centro de Ciencias de la Complejidad (C3) de la UNAM.

La historia, explicó Stephens, coordinador de Ciencia de los Datos del C3 e investigador del Instituto de Ciencias Nucleares de la UNAM, se publicó en la serie The Hitchhiker Guide to Galaxy del escritor y guionista británico Douglas Adams y habla sobre una civilización avanzada que construyó una computadora, Deep Thought, para calcular la respuesta a la pregunta fundamental: el por qué del sentido de la vida, el universo y todo lo demás.

Eventualmente, la computadora llegó a una respuesta.

“La computadora dijo con una voz deprimida que a nadie le gustaría la respuesta”, relató Stephens. “Le respondieron ‘no nos importa, hemos esperado mucho para esta respuesta, la queremos de todas formas’. La computadora insistió que no quería dar la respuesta, que sería muy decepcionante. Finalmente, ante la insistencia respondió: 42. Todo mundo estaba desconcertado”.

La máquina explicó que la razón por la que no era posible comprender qué significa el 42, era porque había primero que comprender el verdadero sentido de la pregunta original. Los que preguntaron en primera instancia realmente no sabían qué era lo que estaban preguntando por lo que se requeriría una computadora todavía más poderosa para calcular la pregunta para la respuesta.

“Ahora les mostraré mi 42”, dijo Stephens. Un modelo predictivo para obesidad basado en más de 2,000 variables que están siendo estudiadas en distintas poblaciones de la UNAM

con apoyo de un grupo multidisciplinario de investigadores. El modelo parece dar buenos resultados predictivos, pero resulta ser equivalente a un 42 porque los investigadores aún no pueden comprender cómo estas variables se relacionan, en el detalle más fino, con los miles de factores directos e indirectos que influyen en la toma de decisiones.

OBESIDAD: LA EPIDEMIA SILENCIOSA

En la década de 1980, sólo 7% de los mexicanos eran obesos. Ese número se triplicó para 2016 (20.3%), de acuerdo al *Institute for Health Metrics and Evaluation* de la Universidad de Washington, EUA. Por su parte, un análisis publicado en la revista *Public Health Nutrition* estima que para 2050 la proporción de hombres y mujeres obesos en el país llegará al 54 y 37% respectivamente.

El sobrepeso se define como un índice de masa corporal (IMC) o proporción de peso a altura mayor o igual a 25 y menor a 30, mientras que la obesidad se define como un IMC igual o mayor a 30. El IMC se obtiene de dividir el peso en kilogramos entre el cuadrado de la estatura en metros (peso/altura x altura).

Esta forma de medir la obesidad tiene sus limitantes aunque sigue siendo una guía ampliamente utilizada como indicador de un exceso de peso. Factores como la edad, el sexo, etnicidad y masa muscular puede influir en la medida del IMC ya que esta medida no distingue entre exceso de grasa, músculo o masa esquelética.

Las consecuencias asociadas a la obesidad son diversas enfermedades, entre las cuales tenemos a las cardiovasculares, diabetes mellitus (reclasificada recientemente en 5 tipos), osteoartritis y algunos tipos de cáncer (endometrio, seno y colon). La diabetes es ahora la principal causa de muerte en nuestro país, con 80,000 muertes al año de acuerdo a la Organización Mundial de la Salud.

En 2010, el costo que representaron las enfermedades crónicas asociadas a la obesidad, para los sistemas de salud en el mundo, fue de un estimado de \$806 millones de dólares. De acuerdo a la tendencia ascendente de la obesidad, el pronóstico de estos costos podría elevarse a más \$1,700 millones para 2050. Para poner en contexto, esto equivale a la mitad de lo que se destina al sector agropecuario, silvicultura, pesca y caza en el país al 2018. Las autoridades de salud advierten

que si el problema no logra contenerse, probablemente disminuyan las expectativas de vida de las futuras generaciones. El problema radica en que las causas de la obesidad son complejas e involucran factores genéticos, de estilo de vida y psicológicos, entre otros. Algunos estudios han asociado el aumento de peso de la población al consumo de alimentos procesados, altos en sal, azúcar y grasas así como a la poca actividad física, pero aún estamos lejos de haber identificado todas las causas.

TRAS EL CONDUCTOMA

A pesar de la enorme inversión en recursos humanos y financieros, la epidemia de obesidad continúa en aumento, y junto con ella, la incidencia de las enfermedades crónicas asociadas. Hasta ahora, las iniciativas implementadas han tenido poco efecto en revertir la tendencia, lo que para Stephens y su equipo de colaboradores significa “que aún falta algo fundamental en nuestra comprensión de las causas y efectos”.

Con el supuesto de que la obesidad es una enfermedad causada por comportamientos particulares, Stephens y su equipo de trabajo proponen el concepto de “**conductoma**”: un conjunto completo de factores que influyen en el comportamiento humano y en la toma de decisiones. Han denominado a su proyecto **Proyecto 42**. El objetivo a largo plazo es obtener y modelar datos que permitan comprender mejor el conductoma humano con el fin de realizar pronósticos sobre el comportamiento.

A diferencia de otros “omas”, enfocados a aspectos moleculares como el **proteoma** (conjunto total de proteínas que se expresan en una célula o un organismo) o el **genoma** (todo el material genético de una célula u organismo), el conductoma contempla diversas escalas: desde lo molecular (la genética) a lo fisiológico y lo ambiental, “con un énfasis en cómo todos estos factores influyen en el comportamiento”.

En el contexto de la obesidad, el equipo de Stephens y sus colaboradores buscan enfocarse en cómo el comportamiento afecta el balance energético y el metabolismo, principalmente a través de la sobrealimentación, la desnutrición y el sedentarismo.

“Aunque el conductoma es enormemente multifactorial”, explicó Stephens, “la revolución en la ciencia de datos y los recientes desarrollos tecno-

lógicos asociados nos permiten poder comenzar a construirlo”.

UN MODELO PREDICTIVO PARA LA OBESIDAD



Los primeros resultados del modelo predictivo que presentó Stephens en su conferencia se basan en datos de un estudio de 1,076 individuos (académicos y no académicos de la UNAM). Se analizaron 2,524 variables incluyendo variables genéticas, epidemiológicas (incluyendo historia personal, historia familiar, nutrición, estilo de vida, conocimientos de salud, autoevaluación de la salud) y antropométricas/fisiológicas (como niveles de colesterol, insulina, hemoglobina, triglicéridos, etc.). Las variables se compilaron a través de cuestionarios y entrevistas personales, así como mediante estudios de laboratorio. Para elaborar los cuestionarios los investigadores se apoyaron en cuestionarios nacionales como ENSANUT.

Los datos personales incluyen preguntas desde a qué se dedica la persona o dónde vive, a variables sociales como cuántos amigos tiene, historia personal de salud, cómo se siente, preguntas sobre percepción de su salud e incluso sobre qué cosas sabe esa persona.

“Algunas de las preguntas que hacemos pueden sonar un poco raras”, explicó Stephens a la audiencia. Por ejemplo, cuántas calorías hay en un vaso con agua. Este tipo de pregunta re-

sulta muy predictiva para saber quién tiene obesidad, pero el problema que tienen los investigadores, su 42 retomando la analogía del principio, es que aún no tienen suficiente información para comprender por qué esta pregunta en particular

resulta predictiva. Con respecto a la población universitaria, el análisis arroja un dato interesante: los académicos tienen una tasa de obesidad del 14% mientras que en los trabajadores no académicos la tasa de obesidad alcanza el 35-40%. Esto es, los académicos tienen 3 veces menos obesidad que los no académicos. Nuevamente, esta primera aproximación indica algo, pero la explicación puede estar en muchísimas variables: el nivel de estudios, los recursos socioeconómicos, la edad, etc. Para los investigadores, el conductoma está esencialmente en tratar de determinar cómo todas esas variables se relacionan con el comportamiento.

Ese es el reto que enfrenta esta investigación. Si de las 2,524 variables sólo tres fueran importantes, los modelos computacionales podrían identificarlas fácilmente. Pero en un problema complejo como la obesidad, son muchas las variables que resultan relevantes por lo que es necesario comprender también cuál es la relación entre todas esas variables.

Para estudiar la base de datos, los investigadores utilizan un algoritmo de aprendizaje automático (*machine learning*). “Lo que hacemos es tomar todas

ACADEMIA DE CIENCIAS DE MORELOS, A.C.



¿Comentarios y sugerencias?, ¿Preguntas sobre temas científicos? CONTÁCTANOS: editorial @acmor.org.mx

las variables y analizarlas con este algoritmo para identificar las combinaciones de variables que se presentan en personas obesas y no obesas", dijo Stephens.

El aprendizaje automático o *machine learning* es una técnica



Joram Patiño Reyes © 2018

estadística (basada en algoritmos y modelos) que permite a las computadoras "aprender" al procesar grandes cantidades de datos, mejorando progresivamente su desempeño para realizar cierta labor. Para ello, el algoritmo construye un modelo matemático a partir de datos simples con los cuales puede hacer predicciones o decisiones sin tener que ser programado de forma específica para realizar esta tarea. Esta es la técnica detrás de las computadoras que pueden identificar caras o reconocer palabras habladas, por ejemplo.

En ese estudio el aprendizaje automático es fundamental para combinar múltiples variables y comparar sus contribuciones.

LA REVOLUCIÓN DE LOS DATOS

Para los investigadores, los primeros resultados obtenidos con la base de datos que han logrado compilar y los análisis asociados sugieren que es factible la construcción iterativa del conductoma, a la vez que permite identificar las principales oportunidades y desafíos.

La propuesta del grupo de investigación es que hay esencialmente un conductoma para cada decisión que toma-

mos. Y este conductoma está determinado por la influencia que múltiples factores tienen en particular para dicha acción. El conductoma es "extraordinariamente multifactorial y adaptativo". Aún no se cuenta con toda esa información pero la revolución de datos está ayudando a generarlos rápidamente.

Otras preguntas que los científicos tratan de resolver con la construcción del conductoma son: ¿qué nos lleva a tomar buenas o malas decisiones? ¿qué información es necesaria y qué información, acaso, es suficiente para decidir? ¿las malas decisiones están codificadas en nuestros genes? (esto es, si adaptaciones que fueron útiles evolutivamente en algún momento ahora nos resultan perjudiciales).

Stephens destaca que es importante enfatizar que "aunque los datos son una condición necesaria para construir el conductoma, no son suficientes, requieren nuevos marcos conceptuales y metodológicos para el análisis de estos datos".

Phillipe Cluzel, investigador de la Universidad de Harvard en Massachusetts, EUA, doctor en biofísica-física y quien no participa en esta investigación considera que: "Stephens y su equipo están trabajando en, posiblemente, uno de los problemas más cruciales del Siglo 21, esto es, comprender cómo todos los datos que tienen que ver con nuestra salud y nuestra forma de vida están conecta-

dos". Para el especialista en el estudio del comportamiento de células y organismos, y sus estructuras celulares, aplicando principios de la física y la ingeniería a sistemas biológicos: "si uno puede encontrar las claves de la conectividad entre estos datos, entonces podremos reducir el impacto en nuestra sociedad de algunos tipos de cáncer y de enfermedades como la obesidad".

Es ahí donde entran las ciencias de la complejidad con su aproximación integral a los problemas y la disponibilidad de nuevas técnicas de modelado como el aprendizaje automático (*machine learning*) y el modelado basado en agentes.

Definitivamente, concluyó Stephens, "los enfoques reduccionistas y disciplinarios tradicionales de la ciencia resultan inadecuados para enfrentar este tipo de problema".

En esta investigación colaboran, entre otros, investigadores del Centro de Ciencias de la Complejidad, Instituto de Ciencias Nucleares, Facultad de Ciencias, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Facultad de Medicina, Instituto de Fisiología Celular, Instituto de Investigaciones Biomédicas, Facultad de Ciencias Políticas y Facultad de Psicología de la UNAM así como del Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Instituto Nacional de Medicina Genómica, Instituto Nacional de Geriátrica, Hospital General de México y Hospital

Juárez de México.

Los investigadores esperan publicar los primeros resultados del "conductoma" de la obesidad durante 2019. Mientras tanto, en el caso de ser diagnosticado con sobrepeso u obesidad, es importante reflexionar en nuestro comportamiento y cuales podrían ser para cada uno, los factores que más influyen en la ganancia de peso, lo cual aparentemente es una tendencia en esta época del año.

Esta columna se prepara y edita semana con semana, en conjunto con investigadores morelenses convencidos del valor del conocimiento científico para el desarrollo social y económico de Morelos. Desde la Academia de Ciencias de Morelos externamos nuestra preocupación por el vacío que genera la extinción de la Secretaría de Innovación, Ciencia y Tecnología dentro del ecosistema de innovación estatal que se debilita sin la participación del Gobierno del Estado.

REFERENCIAS Y LECTURAS RECOMENDADAS

Nota: Una versión corta de este artículo se publicó en la sección de Noticias del Centro de Ciencias de la Complejidad (C3) de la UNAM. <https://www.c3.unam.mx/noticias/noticia76.html>

1. Adams, Douglas (2008) Guía del Autoestopista Galáctico. 6ª ED. Anagrama. Colección Contemporánea. 296p.
2. Murray, CJL & Ng, M (2014)

Nearly one-third of the world's population is obese or overweight, new data show. IHME. <https://bit.ly/1tXEqdf>

3. Rtveldz K, Marsh T, Barquera S, et al. Obesity prevalence in Mexico: impact on health and economic burden. *Public Health Nutr.* 2014;17(1):233-239. <https://bit.ly/2QUvBFS>

4. Rivera Dommarco, Juan Ángel Coord. (2013) Obesidad en México: recomendaciones para una política de Estado. Dirección General de Publicaciones y Fomento Editorial, UNAM. <https://bit.ly/2tGUQ4g>

5. Alcalde-Rabanal JE, Orozco-Núñez E, Espinosa-Henao OE, Arredondo-López A, Alcayde-Barranco L. The complex scenario of obesity, diabetes and hypertension in the area of influence of primary healthcare facilities in Mexico. *PLoS ONE.* 2018 Jan 25;13(1):e0187028 <https://bit.ly/2Lkqpn7>

6. IMSS. Calcula tu IMC. <https://bit.ly/2PFu1mD>

7. Hutson, M. Lifelong Learning. *Scientific American* 319, 5, 14-15 (November 2018) <https://bit.ly/2NJx10j>

8. Jones, N. How machine learning could help to improve climate forecasts. *Nature* 548, 379-380 (24 August 2017) <https://go.nature.com/2PGog8d>

Para actividades recientes de la Academia y artículos anteriores puede consultar: www.acmor.org.mx



Joram Patiño Reyes © 2018