

Jocelyn Dunstan,
 académica de la UC,
 lidera el equipo
 interdisciplinario.



OSCAR VALENZUELA

El año pasado más de dos millones de personas estuvieron en la lista de espera de la salud pública, ya sea por una cirugía, consulta con un especialista o para atención GES, según informó al Congreso la Subsecretaría de Redes Asistenciales.

Para gestionar los recursos y bajar esa cifra es crucial cuantificar las enfermedades con mayor incidencia por región, sexo y edad. Un problema es manejar tal avalancha de datos; además, las listas de espera funcionan con palabras en la ficha de cada paciente que dificultan clasificar rápidamente su enfermedad.

"La sospecha diagnóstica se escribe en un texto que no está estructurado. Así, hay gente que tiene escrito un párrafo largo, hay gente que tiene una palabra o que tiene una sigla; es súper diverso cómo se anota ese campo", explica Jocelyn Dunstan Escudero, académica de Ciencias de la Computación de la Universidad Católica, quien lidera un equipo que -en conjunto con la U. de Chile- busca reducir las listas de espera.

"En medicina pasa que se abrevia mucho: escriben 'diabetes', 'diabetes mellitus', 'dm2'; encuentran mil formas de escribir lo mismo", ejemplifica la académica. Ahí es donde entra en escena la ciencia de datos: dentro de los dos millones de fichas puede detectar palabras claves, definir la enfermedad de cada persona y asignarle un código único a cada dolencia.

¿De qué sirve contar con ese nivel de información? "La lista de espera oftalmológica es la más grande del país, eso se sabe. Pero otra cosa es saber exactamente cuántas personas tienen problemas de refracción y así poder hacer un gráfico y decir 'en esta región está concentrado el tema, vamos a mandar una comitiva para saber cuántas personas necesitan lentes'", ilustra la ingeniera, quien

La especialidad de Data Science ya se enseña en las universidades como posgrado y también como carrera por sí sola.

también es investigadora del Instituto Milenio Fundamento de los Datos.

"Es muy potente no sólo detectar que hay una enfermedad, sino poder decir cuál es exactamente", agrega.

La máquina que aprende

En 2018 comenzaron a trabajar en el proyecto con un grupo interdisciplinario que ha incluido tesisistas y magíster de ciencia de datos y de áreas de la salud.

Al principio pidieron por ley de transparencia las listas en cada servicio de salud; ahí se toparon con que cada unidad tenía su propio criterio para guardar la información. Hoy ya tienen acceso a los datos del Depar-

tamento de Estadísticas e Información de Salud (Deis), del Ministerio de Salud.

"El entrenamiento, o sea cómo le enseñamos a la máquina a detectar qué es importante y qué no, lo hicimos con egresados de Medicina y Odontología", cuenta Dunstan.

Es complicado, admite, porque se necesita contar con lo que se llama corpus anotado, una palabra latina para designar un cuerpo de texto. "Tuvimos que tomar 10.000 notas de la lista de espera y las anotamos con 10 informaciones clave; por ejemplo, medicamentos, antecedentes familiares y genéticos", indica.

Con esa base comenzaron a clasificar. La tarea demoró dos años, pero ese recurso -único en el país- permitió automatizar el proceso de inteligencia artificial. "Después las redes aprenden a detectar esa información y ahora puedes aplicarlo", afirma la investigadora.

¿En qué va hoy el proyecto?

"Gracias a la colaboración con el Ministerio de Salud la idea es encontrar la forma en que esto se va a transformar en una herramienta, tanto para la Superintendencia de Salud como para la gente que gestiona la lista de espera. Estamos en ese trabajo; ya mejoramos la herramienta de detección, la estamos evaluando y en el fondo es cómo lo-

grar que se use".

Cruce de campos

El perfil multidisciplinario, como el del equipo que trabaja en las listas de espera, es una característica del trabajo con datos, aclara Cristian Candia, director del magíster en Data Science de la Universidad del Desarrollo (UDD).

"Alrededor del 50% de nuestros estudiantes provienen de alguna rama de la ingeniería y el resto se reparte en otras áreas del conocimiento. Han pasado por el programa gente de sociología, sicología, medicina, bibliotecología, es bien variado", asegura.

En el magíster estudian tres aristas: matemática, programación y dominio de aplicación de estas herramientas. "Hay mucho crecimiento en telecomunicaciones, también ingeniería, salud, marketing y en las ciencias sociales", comenta.

Sobre sueldos, estima que en el mercado un analista junior gana alrededor de \$1.600.000 mensuales y un experto puede llegar a \$3.500.000. Eso en ejecución, aclara, porque quienes gestionan proyectos de datos son mejor pagados.

El magíster en data science de la UDD comienza en junio y tiene matrículas abiertas (udd.cl, <https://bit.ly/3LLRnHx>).

Enfocado en hallazgos

En la Universidad Mayor imparten la carrera de Data Science, que está adscrita a la gratuidad ([umayor.cl](https://bit.ly/2mdO4QB), <https://bit.ly/2mdO4QB>). "Es relativamente nueva, este es el cuarto año que está trabajando y tiene una colaboración importante con los centros de investigación de la universidad", destaca Jorge Morales, director de la Escuela de Ingeniería.

En la práctica, dice, un científico de datos es un científico. "Lo que deben hacer es desarrollar procesos de análisis de datos complejos, desarrollan algoritmos que permiten encontrar hallazgos importantes como para poder pronosticar y sacar conclusiones relevantes en distintos ámbitos de investigación".

Hace unos años la carrera era asociada especialmente al retail y al manejo de información para ventas, recuerda. "Esto ha ido mutando y hoy día trabaja en recursos humanos, finanzas, operaciones o en investigaciones duras".