



Manuel Gómez Rodríguez

Tenure Track Faculty en Max Planck Institute for Software Systems (Alemania)

Un modelo epidemiológico **espacio-temporal** para la Covid-19

Inicialmente interesado en la teoría de la señal y la codificación de la información, **Manuel Gómez** (Ingeniero de Telecomunicación por la Universidad Carlos III de Madrid) llegó a trabajar en una empresa de microelectrónica al terminar la carrera. Pero su interés por la investigación lo llevó a realizar un Master y más tarde su Tesis Doctoral en la Universidad de Stanford (EE.UU.), alternando esta última etapa con el trabajo en el Instituto Max Planck (Alemania). Y es aquí donde actualmente dirige un grupo de investigación centrado en las aplicaciones de la inteligencia artificial y, en particular, del “machine learning”. En abril, al tiempo que la pandemia alcanzaba su máxima virulencia en Europa, Manuel y su grupo presentaban ya resultados preliminares de su modelo epidemiológico para COVID-19.

Antes de nada, ¿cómo fue posible desarrollar el trabajo tan rápido, en sólo dos meses? ¿Obtuvisteis financiación específica para ello?

Teníamos experiencia en un proyecto que realicé hace dos años con dos

alumnos sobre estrategias de control de enfermedades, que nos ha servido a nivel técnico para modelar lo que ahora hemos hecho para la COVID-19. Así, aplicamos parte de ese conocimiento y la experiencia en las matemáticas del

modelo y en la utilización del *machine learning*. Pero sí es cierto que trabajamos a un ritmo frenético para hacer el código y sacarlo adelante. Por otra parte, el confinamiento por la pandemia ayudó a trabajar más rápido, con lo que



La predicción a largo plazo es muy complicada porque hay muchos factores en juego y demasiada aleatoriedad

en dos meses ya teníamos una versión preliminar. En cuanto a financiación, no tuvimos nada específico. Las personas que trabajan conmigo estaban financiadas y pudimos utilizar toda la infraestructura del Max Planck.

El objeto del proyecto es modelar matemáticamente tres de los elementos esenciales de los que están encima de la mesa: el impacto de los tests, del rastreo de contactos y del confinamiento. Con ello se podrían determinar las acciones menos restrictivas que pudieran mitigar la pandemia. ¿Puedes describir los elementos esenciales en los que se basa el funcionamiento del modelo propuesto? ¿Qué datos utilizáis?

La motivación inicial era hacer algo distinto a lo que hacía la gente de “ciencia de redes”, que mediante un grafo modelaba un determinado entorno y validaba el funcionamiento del *contact tracing*. Otros van un poco más allá y a un modelo similar añaden, por ejemplo, el efecto de la movilidad. Nosotros queríamos hacer algo más general, un modelo que particularizado a un entorno concreto permitiera hacer predicciones. Para eso necesitábamos información específica. Contactamos con Facebook que tiene estimaciones de densidad de población a muy bajo nivel (500x500 metros). También a través de Openstreet Maps pudimos obtener los sitios donde la gente va con más frecuencia (supermercados, colegios...). Por último, Alemania dispone de datos de población por franja de edad en cada municipio. Faltaban las trazas de movilidad real. Para ello hicimos un modelo sintético basado en la probabilidad de que personas de una determinada edad fueran a un determinado sitio de su zona.

Incluyendo el número de test realizados (información publicada diariamente) y los retrasos medios en la obtención de los resultados de los tests, ajustamos el modelo para que ofreciese lo que había ocurrido en una determinada zona, es decir, los casos reales de positivos CO-

VID en ese territorio en particular (datos también disponibles diariamente).

Esto lo hemos realizado ya para varias localidades de Alemania y Suiza, y se han obtenido los mismos números de contagios diarios en la simulación que los que ocurrieron en realidad.

Esto nos da confianza en que las simulaciones nos permitan inferir, por ejemplo, qué pasaría si cerrásemos las escuelas o la probabilidad de que alguien se contagie al ir a un supermercado.

¿Cómo de confiable es a la hora de predecir el comportamiento futuro?

La predicción a largo plazo es muy complicada porque hay muchos factores en juego y demasiada aleatoriedad. A lo que puede contribuir el modelo es, por ejemplo, en el caso de que haya diez casos importados por cada 100.000 habitantes qué es lo que puede pasar si se toman o no ciertas medidas. Pero esas cantidades (casos importados, personas que viajan) son muy difíciles de estimar. Al final las simulaciones dan información que permite decidir de forma cualitativa sobre pautas o políticas concretas y sus efectos.

¿Los datos que pudieran obtenerse de las apps de *contact tracing* ayudarían a enriquecer el modelo?

Sí, claro, con los datos de las aplicaciones de *contact tracing*, el modelo de movilidad mejoraría. Ahora bien, sólo aquellas que incluyan posicionamiento y precisamente eso es algo que ha estado en debate de cara a la privacidad.

De hecho, las apps de *contact tracing* que finalmente se están implantando y que priman privacidad, no ofrecen información que permita saber dónde se ha producido el contagio. Una persona puede saber que ha estado cerca de un positivo pero no dónde ha sido, con lo que es imposible determinar dónde se ha producido un brote. Es decir, esas apps, aunque útiles, no servirían mucho para crear modelos epidemiológicos más precisos. En ese sentido, el *contact tracing* “manual” que realizan los rastreadores puede ser más útil.

El modelo que habéis desarrollado utiliza técnicas de inteligencia artificial y “*machine learning*”. Los resultados que comentas ponen de manifiesto algunas de las limitaciones de esta tecnología, al menos en relación a la creación de modelos epidemiológicos.

Lo cierto es que las técnicas más “populares” en el ámbito de la inteligencia artificial no funcionan bien con problemas con excesiva aleatoriedad, como es este dominio. Un componente esencial para modelos predictivos es precisamente el modelo matemático. No vale la aproximación de “solo datos” (*model free*) para obtener un resultado. Aunque hay máxima atención (incluida financiación) al potencial de la inteligencia artificial, todavía hay una parte que podemos considerar “hype” y un gran camino por recorrer y áreas por estudiar (no solo automatizar tareas sino ayudar al humano a hacer tareas mejor). Quizás sería necesaria cierta divulgación al gran público al respecto. ■