

Predicción multifuente y multimétodo para dar soporte a la toma de decisiones frente a la COVID-19: El proyecto PredCov

Multi-source and multi-method prediction to support COVID-19 policy decision making: The PredCov Project

Editorial

La predicción de incidencia de SARS-CoV-2 resulta en la actualidad un proceso altamente complejo debido a que está determinado por un gran número de factores, como es la aparición de nuevas variantes del virus o el nivel de inmunidad a la enfermedad existente en la población. En una era post-pandemia, con una relajación de las medidas de distanciamiento social, la movilidad urbana y estacionalidad también juegan un papel importante en la forma en que la enfermedad se propaga. Otro factor significativo son las diversas intervenciones farmacéuticas y no farmacéuticas adoptadas en España, que han tenido un impacto en los hábitos de la población, y en la economía, existiendo sectores que han resultado más perjudicados o beneficiados por la pandemia.

El objetivo de este proyecto es desarrollar un sistema que permite combinar distintas fuentes de información para modelar futuros escenarios de propagación del COVID-19 (u otras enfermedades respiratorias similares) desde un enfoque amplio que combina aspectos epidemiológicos, sociales, económicos y de medios de comunicación.

Esto ha supuesto el desarrollo de metodologías de actuación innovadoras en el campo de la analítica de datos que permite poder hacer un seguimiento en tiempo real de pandemias y adelantar el éxito o fracaso de determinadas acciones políticas o sociales que se pueden emprender para disminuir su impacto. También se ha planteado simular escenarios alternativos de implementación de la vacuna contra la COVID-19 para dar soporte a los responsables de definir las políticas de vacunación más apropiadas.

El trabajo realizado en este proyecto abarca ocho líneas de investigación que se describen a continuación.

Objetivo del Proyecto PredCov: desarrollar un sistema que permite combinar distintas fuentes de información para modelar futuros escenarios de propagación del COVID-19

La primera línea ha realizado el diseño, desarrollo e implementación de una solución de digital de rastreo de contactos que mejora las propuestas realizadas durante la pandemia y mantiene garantías de privacidad. La solución desarrollada se basa en el uso de información de geolocalización masiva. Se ha llevado a cabo un análisis de movilidad de ciudadanos basado en datos de geolocalización GPS durante distintas fases de la pandemia en España e Italia, lo que ha permitido entender la modificación

de los patrones de movilidad y el efecto de las restricciones de distanciamiento social que fueron impuestas.

La segunda línea está centrada en realizar predicciones de incidencia de la COVID-19 en la Comunidad de Madrid mediante encuestas indirectas implantadas en el proyecto CoronaSurveys. En esta línea se ha estudiado aristas no descubiertas hasta ahora del Network Scale-Up Method (NSUM) y evaluando las campañas de vacunación y la efectividad de determinadas vacunas. Se ha conseguido un mapa de comparaciones basadas en simulaciones que ponen de manifiesto cuáles de los estimadores de incidencia de COVID-19 son los mejores y cómo es posible atajar los sesgos a los que están expuestos.

La tercera línea incluye el desarrollo de modelos de aprendizaje automático basados en la extracción de patrones de movilidad sobre datos de incidencia COVID-19 para realizar una estimación de la evolución de la pandemia. Se han desarrollado y validado modelos para la estimación de los flujos de tráfico en la ciudad de Madrid. También se han desarrollado modelos para la estimación de flujos de movimiento de personas entre zonas de Madrid usando información de movimientos en el servicio de BiciMAD. Estos modelos han sido utilizados para predecir la incidencia de COVID-19 en la Comunidad de Madrid.

La cuarta línea de investigación ha supuesto la mejora del simulador EpiGraph mediante un modelo social extendido que asigna a cada individuo una localización espacial tanto para su lugar de residencia como de trabajo. Estos valores, que han sido calculados para la ciudad de Madrid, totalizan 3.5 millones de individuos simulados.

También se ha extendido el simulador con un modelo de vacunación que incorpora distintos rangos de cobertura vacunal y se han simulado distintos escenarios de vacunación y uso de mascarillas teniendo en cuenta los resultados de otras líneas de este proyecto.

En la quinta línea los distintos modelos de predicción desarrollados en este proyecto se han integrado en un modelo combinado (ensemble forecast) que incluye aspectos novedosos, no contemplados por los modelos epidemiológicos actuales, como son la información de movilidad de la población extraída mediante la recopilación de trazas de Digital Contact Tracing, el rastreo de contactos, el análisis de los hábitos de la población y la capacidad de cumplimiento de las medidas de distanciamiento social. Las predicciones de este modelo permiten evaluar situaciones de riesgo (alta transmisibilidad del COVID-19) asociadas a eventos, a la percepción pública del riesgo, y a noticias falsas en temáticas de COVID-19.

La sexta línea de investigación realiza un análisis del papel de los medios de comunicación y redes sociales en la formación del debate y discurso de la población sobre la COVID-19, así como su influencia sobre la percepción pública del riesgo. Para llevarla a cabo se analizaron distintas las piezas informativas y posteriormente, mediante estudios de diseño experimental, se determinó la influencia de los medios de comunicación y redes sociales sobre la población durante una crisis de salud pública. De forma complementaria se desarrolló una encuesta nacional con una muestra representativa de todo el territorio y se elaboraron entrevistas en profundidad para conocer aspectos relevantes sobre el consumo de medios de comunicación y reticencia hacia la vacunación COVID-19.

Finalmente, en la octava línea de investigación se han analizado los efectos del shock de recolocación sectorial en función de la diversidad sectorial de la economía. Se han estudiado el gasto de consumo semanal de las tarjetas de crédito y débito, así como la evolución del empleo a través de la Muestra Continua de Vidas Laborales. Adicionalmente, se han presentado propuestas de rediseño de políticas de reducción del tiempo de trabajo utilizando microdatos de la muestra continua de la Encuesta de Población Activa para estimar las transiciones de los trabajadores en ERTE entre los sectores afectados y no afectados por la COVID-19, tanto en España como en la Comunidad de Madrid.

A modo de conclusión, el proyecto se ha completado cumpliendo con todas las expectativas y objetivos planteados inicialmente, los cuales han sido también ampliados de una forma importante. Este proyecto ha realizado una propuesta innovadora con un enfoque altamente multidisciplinar. Mediante la colaboración realizada en el proyecto se ha integrado información proveniente de múltiples fuentes de datos, muy diversas entre sí. Todos los resultados obtenidos han sido particularizados para la Comunidad de Madrid. El proyecto, además de los resultados y productos descritos anteriormente, ha servido para la creación de un equipo capaz de dar soporte a la toma de decisiones en la lucha contra la pandemia. Además, se ha creado una red de grupos de investigación e instituciones colaboradoras que permiten ampliar el alcance de los resultados producidos y enriquecer la información proporcionada en el proyecto. Hay que destacar que los resultados producidos son extrapolables a otras pandemias futuras y que los equipos del proyecto van a seguir manteniendo las líneas de investigación realizadas, por lo que seguirán

El proyecto PredCov ha cumplido con todos los objetivos planteados y ha realizado una propuesta innovadora con un enfoque altamente multidisciplinar

David E. Singh

Investigador principal del proyecto PredCov (Multi-source and multi-method prediction to support COVID-19 policy decision making); Departamento de Informática, Universidad Carlos III de Madrid, España.

 dexposit@inf.uc3m.es