

Politics and big data. Nowcasting and forecasting elections with social media

Andrea Ceron, Luigi Curini y Stefano Maria Iacus
Routledge, Nueva York, 2017, 178 páginas

Paula Szewach

¿Es posible predecir resultados electorales usando información de redes sociales? Ceron, Curini y Iacus señalarían que sí. Evidentemente la respuesta a esta pregunta es mucho más compleja que un escueto sí. A lo largo de seis capítulos comprendidos en el volumen presentan una propuesta metodológica para pronosticar resultados de elecciones reduciendo el margen de las limitaciones a las que se enfrentó la investigación en ciencia política hasta el momento, así como evidencia empírica para sustentarla.

La creciente utilización de *Big Data* como fuente de información le ha permitido a las ciencias sociales comenzar a abordar viejas preguntas de investigación de manera novedosa, así como también han surgido nuevas preguntas a lo largo del proceso. El libro comienza introduciendo el panorama, logrando ubicar al lector en el estado del arte de la investigación que vincula redes sociales y fenómenos políticos como acción colectiva y políticas públicas. Hacia el final del apartado los autores se centran en crear una categorización de los distintos tipos existentes de predicciones electorales vinculados a datos de redes sociales, así como en señalar sus limitaciones.

El segundo capítulo sumerge al lector en la dimensión más puramente metodológica de la obra. Luego de una breve revisión de distintas técnicas de análisis de textos los autores presentan una descripción detallada de su propia técnica para diagnosticar resultados elec-

torales, una variante de *sentiment analysis*. Puede resultar un tanto denso para un lector no especializado, pero es notorio el esfuerzo por parte de los autores por resumir y simplificar los pasos a seguir. Incluso incorporan al final del capítulo las indicaciones para utilizar el paquete iSAX en R (software libre) que aplica el algoritmo que desarrollan con el código comentado. Como los *software* para correr los análisis suelen actualizarse con cierta periodicidad el código presentado y comentado en el libro podría rápidamente quedar obsoleto. De todos modos, su contraparte digital se encuentra en un repositorio abierto¹ en donde seguramente los autores presentarán las modificaciones necesarias. Para el lector especializado sugieren complementar la lectura del libro con algunos artículos en los que desarrollan en mayor profundidad los aspectos técnicos.

El tercer capítulo presenta los primeros estudios empíricos de monitoreo y pronóstico de resultados electorales valiéndose del método descrito en el capítulo anterior. El punto fuerte de esta sección es más que nada la ilustración del método en acción, así como la replicación en escenarios electorales diversos. Los investigadores analizaron elecciones en Francia, Estados Unidos e Italia entre abril de 2012 y diciembre de 2013. Cabe resaltar que se trata de elecciones de distinta naturaleza según lo que señala la li-

¹ <https://github.com/blogsvoices/iSAX>.

teratura en el área. Es decir, elecciones primarias no presentan las mismas dinámicas que elecciones generales y elecciones presidenciales se diferencian de legislativas, para dar algunos ejemplos. Esta distinción está presente en el libro y aparentemente el método se sostiene para los diferentes casos, aunque queda claro que es imprescindible tener conocimiento previo sobre el contexto de cada elección, tanto de las reglas del juego como de los actores involucrados y sus estrategias de campaña. En el capítulo cuarto también se presenta evidencia empírica con datos extraídos de redes sociales, pero en esta instancia se centran en monitorear dinámicas de campaña y tendencias de opinión pública. El gran atractivo que presenta esta aplicación metodológica es que permite obtener *feedback* de la efectividad de estrategias de campaña en tiempo real. En ese sentido, esta sección vincula las publicaciones de campaña en cuentas oficiales de partidos políticos con opiniones e intención de voto manifestada por ciudadanos para el caso de las elecciones generales en Italia en 2013.

A lo largo de los capítulos anteriores, los autores obtuvieron preponderantemente resultados exitosos valiéndose de su método de monitoreo y predicción. Sin embargo, el nivel de precisión presentó variaciones que los condujeron a profundizar en el análisis. Llegando al quinto capítulo, el libro presenta pruebas estadísticas para indagar en la precisión y robustez de los resultados obtenidos. La sección finaliza con los resultados de un metaanálisis, acaso la evidencia más interesante que despliegan en este apartado. Se trata de un estudio que valiéndose del análisis de una muestra de todas las predicciones de resultados electorales en base a datos de redes sociales disponibles da cuenta de las variables que afectan la calidad de las predicciones. Según

su estudio, la técnica aplicada para realizar la predicción modifica considerablemente el margen de error y el método iSA sugerido por ellos es el que conduce a errores más pequeños. Además, señalan que distintos atributos de los sistemas electorales inciden en la calidad de la predicción. Por ejemplo, sistemas electorales de representación proporcional mejoran las predicciones respecto a sistemas mayoritarios dado que los sistemas proporcionales habilitan más correspondencia entre preferencias sinceras y voto, mientras que los sistemas mayoritarios conducen a un aumento en el comportamiento electoral estratégico.

Como conclusión, la riqueza del libro recae en la reflexión cuidadosa en torno al *Big Data* y los métodos computacionales que progresivamente ocupan un lugar más importante en las ciencias sociales. En ese contexto, Ceron, Curini y Iacus presentan herramientas de calidad que se adaptan a los tiempos que corren y que pueden proveer al investigador información muy valiosa en tiempo real y a bajo costo. Todo esto sin descuidar la introspección sobre las implicancias de su contribución. Constantemente resaltan la importancia de que las ciencias sociales se guíen por marcos teóricos sólidos y se valgan de métodos informados por humanos. El mensaje que intentan transmitir los autores a lo largo de todo el libro se torna más explícito hacia el final, cuando insisten en que la investigación en ciencias sociales no puede ser un conjunto de herramientas novedosas, sino que se trata de una manera de ver el mundo y analizar los problemas que se presentan. La mejor receta para aplicar técnicas computacionales a las ciencias sociales es justamente que no haya una receta universal: el diseño de investigación no puede ser reemplazado por procesos automatizados.