

Causalidad y probabilidad

Description

Desde hace alrededor de un siglo el manejo de la información se realizaba recopilando y analizando pequeñas parcelas de ésta. Los estadísticos efectuaban un muestreo sobre un segmento del universo de información existente, proyectando sus resultados sobre el conjunto. Es la esencia de las encuestas, en donde la opinión de un grupo aleatorio pero necesariamente reducido de personas, permite dar a conocer la tendencia general.

Desde hace alrededor de un siglo el manejo de la información se realizaba recopilando y analizando pequeñas parcelas de ésta. Los estadísticos efectuaban un muestreo sobre un segmento del universo de información existente, proyectando sus resultados sobre el conjunto. Es la esencia de las encuestas, en donde la opinión de un grupo aleatorio pero necesariamente reducido de personas, permite dar a conocer la tendencia general.

En la medida en que la recolección y manejo de la información constituían un proceso complejo, costoso y muy exigente en tiempo, la solución era la de focalizar la atención en una muestra reducida y aleatoria. Este planteamiento asumía que, dentro de ciertos márgenes menores de error, era posible inferir lo que sucedía en el conjunto a partir de esa muestra. La misma debía ser obtenida de manera ordenada y procesada con extremo cuidado.

A partir de lo micro se buscaba entender así lo que ocurría en lo macro. Encuestas pre electorales o a boca de urna, muestreos médicos, estadísticas económicas, partían todas del mismo supuesto. Ello brindaba la sensación de abarcar y comprender la totalidad. Sin embargo, como bien quedaba demostrado en una y otra oportunidad, tal sensación podía no ser más que una ilusión auto gratificante. Con cuanta frecuencia las encuestas no sólo se contradicen entre sí, sino que se equivocan rotundamente. O nuevos estudios médicos arrojan resultados a contracorriente de los anteriores.

Un par de ejemplos pueden ilustrar lo dicho. Según las encuestas a boca de urna, los Republicanos obtuvieron el 31% del voto latino en la elección estadounidense de 2016. No obstante, la encuestadora hispana “Latino Decisions” no aceptó tal cifra. Según sus propios resultados, Clinton obtuvo 79% del voto latino contra 18% para Trump. La diferencia, a su juicio, estribaba en que los “exit polls” que proporcionaron la cifra de 31% fueron obtenidos a partir de un número pequeño e insuficientemente representativo de precintos electorales. Ellos, por el contrario, se concentraron en las circunscripciones de presencia latina, trabajando de manera puntual este grupo (Gabriel Sanchez, Matt Barreto, “In record numbers, Latinos voted overwhelmingly against Trump”, Washington Post, November 11, 2016).

Otro ejemplo puede venir del mundo médico, donde la sabiduría convencional proveniente de estudios anteriores es frecuentemente revertida por otros nuevos. Es así que el Harvard Medical Review, de agosto de 2017, reporta que la noción de que los bebedores moderados de alcohol disfrutaban de mayor protección frente a las enfermedades cardíacas es falsa. Lo que ocurre, según demuestran estudios recientes, es que quienes están en condiciones de consumir alcohol sin restricciones médicas disfrutaban por definición de mejor salud. En otras palabras, los muestreos anteriores estaban colocando equivocadamente la carreta por delante de los caballos.

Sin embargo, el mundo de los muestreos puede estar pasando al desván de las antigüedades, ante la aparición de lo que en inglés se conoce como “big data”, y que aproximativamente se traduciría como información amplia. De acuerdo a la misma, todo el universo de información disponible en la esfera digital puede ser instantáneamente “accesado” y cotejado por la computación. El volumen de la información digitalizada es por lo demás de tal magnitud, que si toda la información allí disponible fuese colocada en discos CD, se podrían hacer cinco pilas que llegarían a la luna (Neil Cukier and Victor Mayer Schoenberg, “The Rise of Big Data”, Foreign Affairs, May/June 2013).

La esencia de este proceso es que, a través de la información masiva, se puede llegar a niveles de comprensión por completo inaccesibles para quienes sólo manejan fragmentos de aquella. En otras palabras, a partir de lo macro se puede comprender lo micro. Un buen ejemplo de ello lo encontramos en las funciones que cumple la célebre computadora Watson de IBM.

Luego de diversas proezas en materia de pensamiento analítico, Watson se ha adentrado en el campo del diagnóstico médico. Ni más ni menos, dicha computadora está en capacidad de pasar revista inmediata a todos los trabajos médicos publicados, a todas las convenciones médicas celebradas, a todos los diagnósticos médicos digitalizados para, luego de cotejar y analizar todas las fuentes, emitir un diagnóstico.

En efecto, la convergencia entre una memoria computacional de bajo costo, poderosas computadoras, algoritmos inteligentes, softwares altamente creativos y altas matemáticas, está permitiendo alterar de manera radical el manejo dado a la información. Ello está llevando, a la vez, a un cambio de paradigmas en esta materia. En lugar de buscar entender la causa de los fenómenos, lo importante pasa a ser comprender la correlación entre éstos. Esto, a no dudarlo, llevará también a un cambio en los procesos de pensamiento y análisis. Lo importante será la relación probabilística de las cosas y no su relación causa-efecto.

APARTADOSTEMATICOXEOGRAFICOS

Outros

ETIQUETAS

inteligencia artificial estadística soluciones

IDIOMA

Castelán

INVESTIGACION

Relacións Internacionais

Date Created

Setembro 4, 2017

Meta Fields

Autoria : 3733

Datapublicacion : 2017-09-04 00:00:00