

Nuevas tendencias inteligencia artificial y estadística

JOSE LUIS PIQUERAS GARCIA,
Capitán de Corbeta

D ICEN que la mar es el camino del pensamiento. ¿Habrán nacido grandes pensamientos en la mar? Todo aquél que haya tenido la oportunidad de navegar, habrá comprobado como en los ratos de calma la imaginación vuela y vuela sobre aquellos lugares, personas, hechos, proyectos, o quizás ideas, que están escondidas en nuestra mente y son precisamente en esos momentos cuando regresan al pensamiento.

Los que como yo hemos tenido la oportunidad de navegar y de efectuar hasta 19 horas seguidas de vuelo en los aviones antisubmarinos, hemos comprobado, que también hay momentos en los vuelos, donde de la memoria brotan los pensamientos, aunque como todo es función de la velocidad, la película de estos pensamientos en un vuelo pasa más rápida que en una navegación.

¿Porqué este inicio de un artículo que trata sobre Inteligencia Artificial (IA) y Estadística?

Muy sencillo, el pensamiento del ser humano, su forma de razonar, su conocimiento, la estructura de este conocimiento, etc. es lo que trata de implementar el hombre en una máquina, basándose en una de las ramas de esta nueva tecnología llamada IA, para que, no surja sólo en los momentos de calma, sino también, cuando se esté inmerso en una tempestad; y al igual que durante esos vuelos o esas navegaciones, de nuestra memoria surgen las ideas y los hechos pasados, así deberán surgir los datos de una base de datos informatizada, y una vez tratados estadísticamente, seamos capaces de sacar unas conclusiones o unos resultados, en resumen, basarnos en la historia pasada, bien para prevenir el futuro, bien para mejorar lo anteriormente realizado.

Dado que en la Revista de Aeronáutica y Astronáutica de julio de 1987, apareció un dossier sobre la IA, yo no voy a hablar sobre este tema, ya que quedó muy bien reflejada por sus autores la idea de lo que es y sus aplicaciones. Por el contrario, sí me quiero detener en una aplicación concreta, que no fue tratada en el citado "dossier" y es la influencia que puede tener esta técnica en la Estadística.

El nuevo horizonte que está abriendo la IA en las investigaciones estadísticas, trata de integrar el conocimiento previo para llegar a obtener una total aproximación al análisis de datos.

Al mismo tiempo, los investigadores de la IA están trabajando conjuntamente con los estadísticos en la búsqueda de soluciones para los problemas que engloban la incertidumbre. Uno de estos problemas es la representación de conclusiones probabilísticas sacadas de hechos que implican factores de certeza, gracias a unas reglas que no son del todo veraces.

Así como la matemática es una ciencia exacta, y las tablas estadísticas también, los resultados o conclusiones de una estadística no lo son, y siempre se dan con un margen de error, o dentro de unos límites.

La clave de la IA para su utilidad en la Estadística está en capacitar a los estadísticos para el estudio de las estrategias que se deben usar en el análisis de los datos.

El término estrategia aquí utilizado trata de resaltar una total aproximación coherente a la tarea analítica para trabajar con los datos. Sin embargo, no está generalmente aceptada la definición de este término, ya que según dos insignes investigadores como Daryl Pregibon y William A. Gale sugieren que una definición de este tipo debería responder a cuestiones tales como:

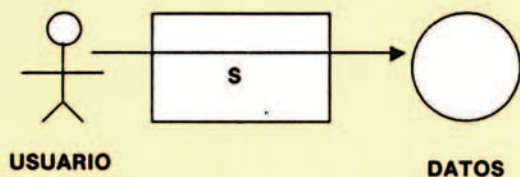
- ¿qué busco yo?
- ¿cuándo lo busco?
- ¿cómo lo busco?
- ¿porqué lo busco?
- ¿qué tengo que hacer para buscarlo?

Otros investigadores tales como Wayne Oldford y Steve Peters utilizan el término para definir el razonamiento utilizado por los expertos estadísticos en el desarrollo del análisis de los datos que nos deben resolver el problema.

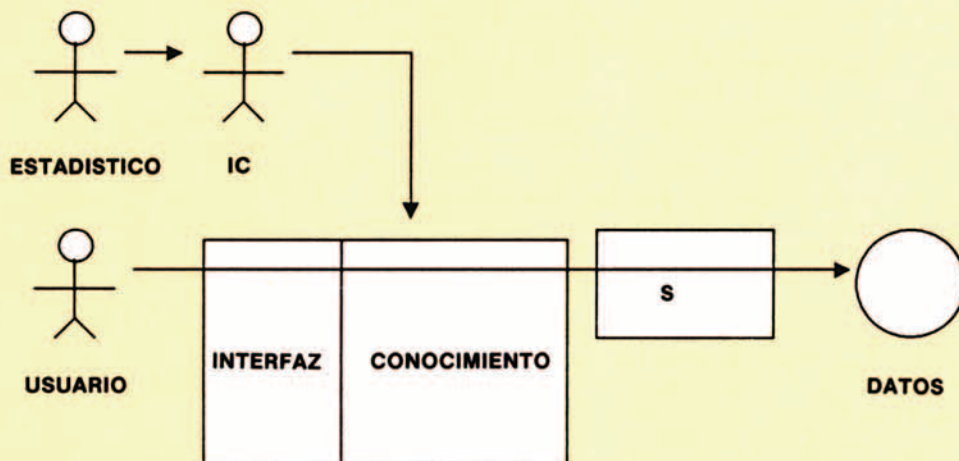
También, David Hand, le da un sentido diferente, definiéndolo como una descripción formal de las opciones, acciones y decisiones que deben elegirse para utilizar los métodos estadísticos durante la realización de un estudio.

Se necesita por lo tanto, crear un concepto de "estrategia estadística" más generalizado, para que puedan utilizarse por una mayor cantidad de personas los paquetes informáticos estadísticos existentes. En resumen, facilitar al usuario el manejo, elección e interpretación de datos estadísticos, por medio de la

PRACTICA DIARIA



SISTEMA REX



SISTEMA STUDENT

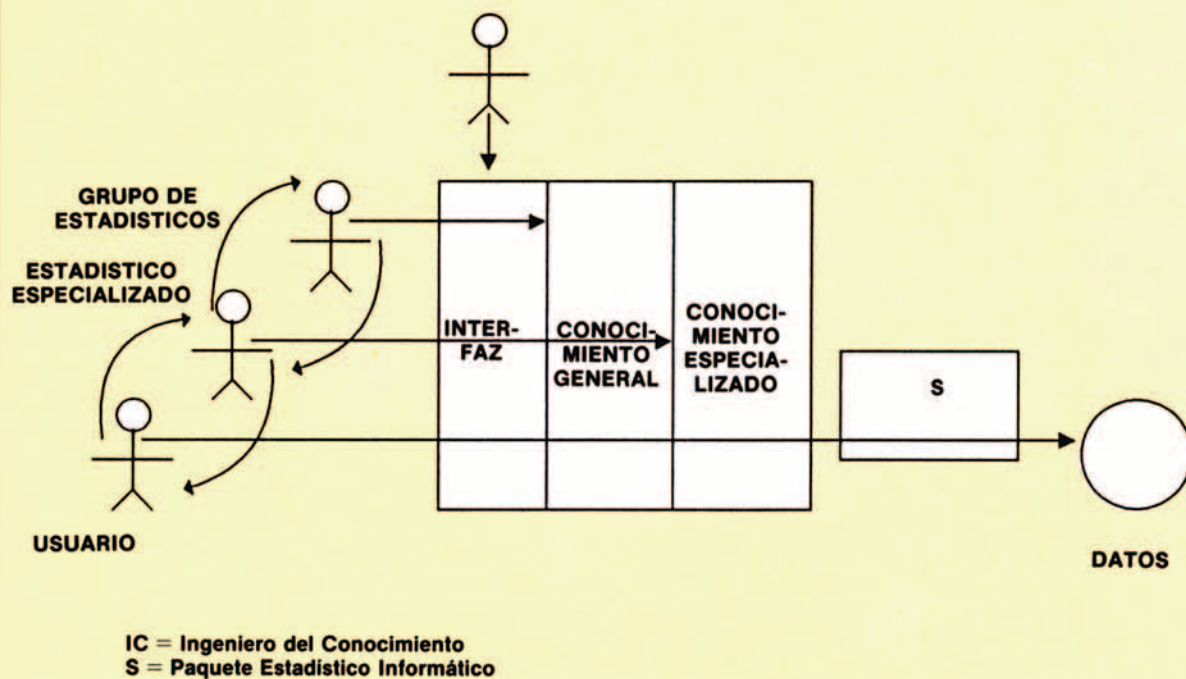


Figura 1.

mecanización de una "estrategia estadística". Las herramientas desarrolladas en materia de IA, nos van a permitir dicha mecanización.

En la última década, los investigadores de la IA, han demostrado que se pueden mecanizar los procesos simbólicos que hasta entonces estaban reservados solamente para la inteligencia humana. Estas técnicas incluyen entre otros, la comprensión del lenguaje escrito y oral y el reconocimiento de imágenes. En particular, los sistemas expertos realizados, han hecho posible concebir que la utilización de esta técnica puede ser muy útil en la Estadística.

La Estadística, ayuda a la IA, realizando los estudios de capacidades para una serie de eventos que luego serán utilizados en el razonamiento y en las Bases de Conocimiento de los Sistemas Expertos (SS.EE.).

Por ejemplo, del estudio del número de casos en los que se produce una u otra enfermedad, en los casos clínicos, se puede deducir el porcentaje probabilístico de cada una, ilustrando así el diagnóstico médico.

¿De qué forma ayuda la IA a la Estadística? Hay dos caminos, en los que actualmente se están realizando todos los esfuerzos para aplicar las técnicas de IA al análisis de datos. Uno de ellos está enfocado a facilitar el manejo de los complicados paquetes informáticos estadísticos, a todas aquellas personas que no tienen mucha experiencia con este manejo; el otro trata de que sea más productivo el uso para los profesionales de la estadística de los citados paquetes informáticos. Por lo tanto, un sistema que intenta ser de utilidad para los principiantes en la estadística, necesita representar, la estrategia estadística.

Así, por ejemplo, el sistema REX ("Regresión EXpert") está diseñado para representar la estrategia estadística con objeto de que puedan utilizarla personas no especialistas. El sistema aconseja al usuario en el análisis de problemas de regresión. Va guiándolo por medio de suposiciones regresivas, sugiriendo la transformación de los problemas cuando se violan las hipótesis y justificándolas cuando se le pregunte. También, interpreta los resultados intermedios y finales e instruye al usuario en los conceptos estadísticos.

Un sistema para ayudar a los profesionales de la estadística es el STUDENT. Está diseñado para permitir a los profesionales de la estadística construir un sistema de consulta basado en el conocimiento con la técnica del análisis de datos por medio de la selección y comparación de ejemplos y de respuestas a preguntas. Los estadísticos no necesitarán conocer ni la representación interna de la estrategia utilizada, ni como se escribe un programa basado en el conocimiento. Hablando en términos de IA, se diría que es una herramienta para desarrollar un Sistema Experto que efectúe el análisis de datos, a la medida de cada usuario, donde éste, lo único que tiene que saber es de estadística.

Conociendo pues, lo que realiza STUDENT y lo que realiza REX, se puede deducir que REX es una demostración de lo que STUDENT puede proveer.

Para una mejor comprensión de cómo actúan y cómo llegan a sus resultados sigamos la figura 1.

En la figura, los rectángulos representan las fuentes de conocimiento programadas, y las flechas a través de un rectángulo representan el conocimiento que debe utilizarse para trabajar sobre la entidad donde descansa la punta de la flecha.

En la práctica diaria, un usuario interactúa directamente sobre un paquete estadístico S, decidiendo lo que hay que hacer, seleccionando comandos e interpretando resultados. En resumen, utiliza a S para analizar los resultados.

Con el sistema REX, el usuario accede a S solamente a través del interfaz que se le proporciona. REX le va sugiriendo lo que hay que hacer, lo verifica, selecciona órdenes y ayuda al usuario a interpretar los resultados. Los usuarios no profesionales de la estadística encuentran más fácil el manejo del paquete estadístico.

Para ayudar al usuario de esta forma, dos actores aparecen en escena, un Ingeniero del Conocimiento (IC) y un estadístico. El primero es el encargado de crear el programa de la interfaz propiamente dicho y ayuda al segundo en su labor de introducir su conocimiento sobre estadística. Por lo tanto, el estadístico, debe de transmitir al IC su conocimiento. El IC pasará mucho tiempo ayudando al estadístico a representar su conocimiento de una forma que sea inteligible por el sistema.

Con el STUDENT el papel del IC será el de construir el motor de inferencia del sistema, que es el corazón del mismo y el que lo capacita para adquirir la estrategia estadística, cualquiera que ésta sea. Al principio, un grupo de estadísticos, proveerán al sistema de las mínimas estrategias sobre varias técnicas de análisis de datos. Ya se vislumbra el papel vital que juegan estos estadísticos en el sistema, pues deben introducir los primeros y más generales ejemplos que serán luego utilizados para ir eligiendo las diferentes estrategias. Durante este período es cuando se producen los cambios más drásticos en el motor de inferencia del sistema. Llegados a este punto, cuando el sistema tiene introducido ese conocimiento, se le irán añadiendo otros ejemplos para ir comprobando si va reaccionando correctamente. Después de este período de aprendizaje y entrenamiento el sistema será manejado por el usuario para su validación.

Con estas breves palabras he tratado de mostrar una de las múltiples aplicaciones de los Sistemas Expertos, subconjunto de la nueva tecnología llamada Inteligencia Artificial.

Los Sistemas Expertos no solamente revolucionarán lo que los usuarios podrán hacer sino que tendrán un gran impacto en la educación a todos los niveles.

Dice un proverbio chino: "Sus hijos deben vivir tiempos mejores". Durante los últimos 45 años, los analistas de datos que buscaban datos reales han vivido tiempos felices. En los 45 o 450 años venideros, los Sistemas Expertos que ahora comienzan a aparecer vivirán también tiempos mejores. ■