DETERMINAR EL VOTO DE UNA PERSONA QUE NO HA RESPONDIDO A TODAS LAS PREGUNTAS

Un nuevo sistema aumenta la fiabilidad de las encuestas de opinión

Un nuevo sistema basado en Inteligencia Artificial permite determinar el voto de una persona que no ha respondido a todas las preguntas de una encuesta de opinión política. Desarrollado en la Facultad de Informática de la Universidad Politécnica de Madrid, tiene una fiabilidad superior al 90% y mejora sustancialmente los métodos actuales de imputación.

UPM 26/4/2011 10:09 CEST

Un nuevo sistema aumenta la fiabilidad de las encuestas de opinión. Imagen: Manolo Gómez

Investigadores de la Facultad de Informática de la Universidad Politécnica de Madrid han desarrollado una red neuronal difusa que es capaz de reconstruir conjuntos de datos incompletos a partir de un método de imputación numérico y categórico, mejorando sustancialmente los métodos actuales de imputación en encuestas de opinión. Los resultados de su investigación se han publicado en la revista especializada Neural Computing & Applications.

La ausencia de datos es un problema generalizado en la mayoría de las

encuestas que se realizan en los más diversos ámbitos. La falta de precisión en las encuestas de opinión políticas se debe en buena parte al tratamiento actual de las preguntas sin respuesta. La técnica más habitual para subsanar esta deficiencia es la imputación, mediante la cual se calculan los datos faltantes y se añaden a la muestra para obtener un conjunto más completo. Los métodos de imputación pueden inferir datos numéricos y categóricos.

Fiabilidad del 90%

Los investigadores Jesús Cardeñosa, del Departamento de Inteligencia Artificial de la Facultad, y Pilar Rey del Castillo, del Instituto de Estudios Fiscales del Ministerio español de Economía, han introducido una variación en el método de imputación numérico de Gabrys y Bargiela, consiguiendo de esta forma inferir datos categóricos.

Con este sistema, la imputación puede por ejemplo determinar el sentido del voto de una persona que no ha respondido a todas las preguntas de una encuesta de opinión con una fiabilidad próxima al 90%, lo que supone una mejora sustancial sobre los métodos actuales. Otras aplicaciones posibles de esta red neuronal se encuentran en el ámbito del diagnóstico médico o de otras encuestas que contienen variables categóricas.

El nuevo método de imputación comienza por definir las distancias entre categorías utilizando la lógica difusa. A continuación determina, con la ayuda de la red neuronal que aprende de cada caso, la ubicación de cada categoría en los diferentes espacios del conjunto de datos. Finalmente extiende la arquitectura de la red a todos los datos y al tratamiento de los datos que faltan.

Derechos: **Creative Commons**

Creative Commons 4.0
Puedes copiar, difundir y transformar los contenidos de SINC. Lee las condiciones de nuestra licencia