

Designing a Predictive Analytics Solution for Evaluating the Scientific Trends in Information Systems Domain

Babak Sohrabi, Iman Raeesi Vanani and Mohsen Baranizade Shineh

Webology, ISSN 1735-188X, vol. 14, n.1, 2017, p. 32-52

Los artículos a texto completo son proporcionados mayoritariamente en formatos electrónicos, pero todavía sigue siendo un desafío encontrar trabajos con una relación profunda más allá de las palabras claves y del contexto de los resúmenes. La identificación de artículos relacionados científicamente basándose en palabras claves y resúmenes ayuda a los investigadores a encontrar los contenidos deseados e informativos, pero requiere la utilización de una variedad de métodos de análisis de textos (minado) para dirigir la búsqueda. Este artículo investiga artículos sobre sistemas de información (IS) para agrupar y evaluar el campo de las recientes publicaciones sobre IS usando métodos de agrupación de textos, y después para predecir los grupos exactos de nuevos artículos usando métodos de clasificación. Esta aproximación a través de categorización y aprendizaje predictivo ayuda a los investigadores a encontrar los artículos más relevantes en sus investigaciones y tareas prácticas a través de un mecanismo automatizado. Los artículos fueron recogidos del repositorio Scopus. Los conjuntos de datos se limitaron a áreas específicas de investigación de IS recientes. Se aplicaron diferentes métodos de análisis de textos, como normalización de textos, procesamiento de lenguaje natural y algoritmos de agrupamiento, y los

resultados por cada grupo son evaluados por análisis extensivos de grupos identificados basados en la frecuencia de sus términos y frases clave. Después, diferentes algoritmos de clasificación son aplicados para conocer el actual agrupamiento y predecir los temas principales de los nuevos artículos publicados basándose en la aproximación de los resúmenes a los campos previamente aprendidos sobre IS. Esta predicción ayuda a los investigadores a identificar la usabilidad de muchos artículos para posteriores investigaciones.

Traducción del resumen de la propia publicación