

# Una herramienta automática para la clasificación y búsqueda de artículos de investigación en la web

Edgar Alan Calvillo-Moreno<sup>1</sup>, Ricardo Mendoza-González<sup>1</sup>, Jaime Muñoz-Arteaga<sup>3</sup>, Julio Cesar Martinez-Romo<sup>1</sup>, Miguel Vargas-Martin<sup>2</sup>, Laura Cecilia Rodriguez-Martinez<sup>1</sup>

<sup>1</sup> Instituto Tecnológico de Aguascalientes (México)  
<sup>2</sup> University of Ontario Institute of Technology (Canadá)  
<sup>3</sup> Universidad Autonoma de Aguascalientes (México)

## CASO DE ESTUDIO

Las Figuras 1 y 2 muestran un ejemplo de la interfaz de usuario de la herramienta propuesta, que fue utilizada para buscar y analizar 1000 artículos científicos extraídos de un repositorio local y la API de Springer [4]. Este es el flujo para obtener nuestra base de conocimiento:

1. Se utiliza una herramienta para extraer conocimiento de bibliotecas digitales.
2. La herramienta ejecuta el motor de clasificación sobre los artículos descargados.
3. El artículo se encuentra clasificado en la base de conocimiento listo para ser utilizado por el motor de búsqueda.

La demo de la aplicación está alojada en <http://engine-phdapp.rhcloud.com/> y está disponible para probar y puede ejecutar y replicar el caso de estudio.

El presente trabajo proporciona pruebas ejecutadas con un buen rendimiento en velocidad, precisión y uso de los recursos del servidor[5].

## REFERENCIAS

- [1] M.-G. Manuel, "Minería de texto: Un nuevo reto computacional," 3rd Int. Workshop Data Min. MINDAT-2001.
- [2] C. Denicia-Carral, M. Montes-y-Gómez, L. Villaseñor-Pineda, and R. García Hernández, "A Text Mining Approach for Definition Question Answering," Springer-Verl. Berl. Heidelberg. 2006.
- [3] C. E. Alan, P. Alejandro, M. Jaime, P. Julio, and F. Jesualdo, "Searching Research Papers Using Clustering and Text Mining," Conielectcomp 2013.
- [4] E. A. Calvillo Moreno, R. Mendoza González, J. Muñoz Arteaga, J. C. Martínez Romo, and S. Jiménez González, "Using Text Mining to Locate and Classify Research Papers," Math. Methods Comput. Tech. Intell. Syst. MAMECTIS 15, January 2015.
- [5] E. A. Calvillo Moreno, R. Mendoza-Gonzalez, J. Muñoz-Arteaga, J. C. Martínez Romo, M. Vargas Martín, L. C. Rodríguez Martínez, "A Bayesian classifier to search and classify research papers" Dyna New Technologies Enero 2015, vol. 2, no. 1, p.1-13. DOI: <http://dx.doi.org/10.6036/NT7657>.

DOI: <http://dx.doi.org/10.6036/7827>

El uso de repositorios de artículos de investigación permite a los investigadores y académicos alrededor del mundo obtener información en muchas áreas de la ciencia. Servicios como Springer, ACM Digital Library, Science Direct, Scopus, IEEE XPLore, entre otros disponibles en línea,

estos sólo consideran información general como título; Descripción; Resumen; autores; institución; área; y el tema [1].

El objetivo del trabajo es extraer información del contenido de los artículos de investigación en las librerías digitales[3], utilizando técnicas de minería de texto [5] para crear un motor que permite extraer conocimiento de cada artículo clasificado

y utilizar la información más importante de esos para crear una base de datos de conocimiento que facilitará la localización de artículos.

La diferencia es que otros trabajos similares utilizan un enfoque semiautomático para obtener una base de conocimiento y el trabajo propuesto se enfoca en la creación de una herramienta automática para clasificar y localizar documentos.

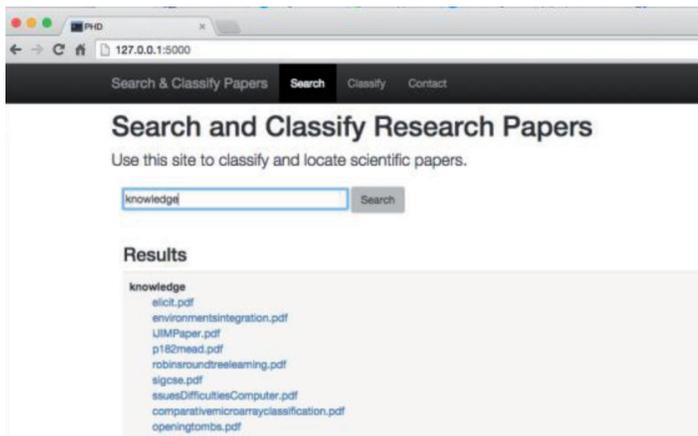


Fig. 1: Motor de búsqueda

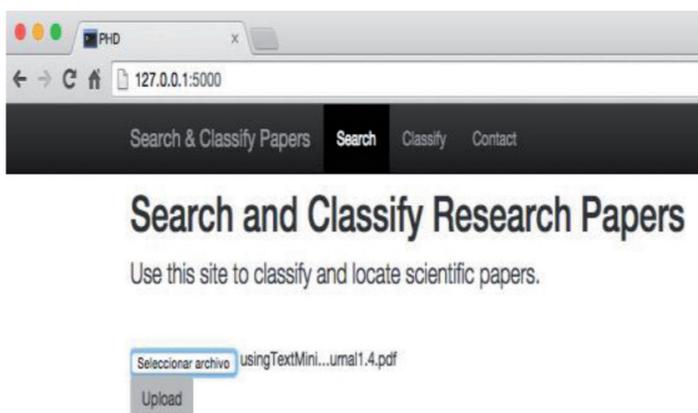


Fig. 2: Motor de Clasificación