

# A Novel Efficient Classification Algorithm for Search Engines

Hanan Ahmed Hosni Mahmoud Abd Alla<sup>1</sup>  
*Information Technology Department  
College of Computer and Information Sciences  
King Saud University  
Hanan.hosni@yahoo.com*

Nadia Al-Ghreimil  
*Information Technology Department  
College of Computer and Information Sciences  
King Saud University*

## Abstract

*In this paper a new classification algorithm of Web documents into a set of categories, is proposed. The proposed technique is based on analyzing relationships between different documents and the terms they contain by producing a set of rules relating the category of the document, its terms and their frequencies. Each document is represented by a graph that correlates its most frequent combined words and its category. The relationships among these graphs and the documents' categories are captured. The proposed technique has three phases. The first phase is a training phase where human experts determines the categories of different web pages and articles and the supervised classification algorithm will combine these categories with appropriate weighted index terms according to the highest supported rules among the most frequent words. The second phase is the blind categorization phase where a web crawler will crawl through the World Wide Web to build a database that will be categorized according to the result of the first phase. This data base contains URLs and their categories. The third phase is applying the proposed graph representation technique on the whole set of documents per category to determine its final graph representation. The third phase will produce better classification rules because the sample size is larger with no additional cost of supervised categorization. Experiments using data sets collected from different Web portals are conducted.*

**Keywords:** Information Processing on the Web,  
Supervised Classification, Document Classification

1. On leave from Department of Computer and Systems Engineering,  
Faculty of Engineering, University of Alexandria, Egypt