

Ant colony algorithm for web page classification

ABSTRACT

As the information contained in the web is increasing, organizing this information is a necessary requirement. Classification is a way to organize the web's information. While there is different ways to classify the web information, swarm intelligence is a new one in this field.

This paper investigates usage of a swarm intelligence algorithm in the field of the web page classification; Focusing on Persian web pages. Ant Miner II is the used algorithm. It also propose a simple text preprocessing technique to reduce the large numbers of attributes associated with web content mining, without dealing linguistic complications. The result shows Ant Miner II and proposed preprocessing technique are efficiency in the field of the web page classification.