

# Automated Blog Classification using Rule Extraction by Reverse Engineering the Neural Network

 K. Aruna Devi and T. Kathirvalavakumar

## Abstract

Blog content classification is a process of analyzing the blog posts and label with a predefined category which helps the search engine to improve searching and marketing. However, the effort is time consuming and error prone as the blogosphere is an open and growing domain. The state-of-the-art blog classification is influenced by the supervised methods which uses the ample corpus vocabulary as feature set demanding extensive memory space to classify. The intense usage of memory space affects the training and processing time of the classifier. To address the issue, Automated Blog Classification using Rule Extraction by Reverse Engineering the Neural Network (RxREN) is proposed which does multi stage feature reduction and classification using ANN. The proposed classification framework reduces the features in the first stage; patterns in the second stage and configure ANN according to the reduced dataset using N2PS pruning algorithm. In neural networks the knowledge generated from the input data is more precise but generally not descriptive. The symbolic rules can be derived to expose the knowledge built inside the ANN configuration. Hence using Reverse Engineering the Neural network (RxREN) the rule extraction is performed after the blog classification training. The extracted rules are used directly to classify the dataset. The proposed methodology results on a benchmarked dataset proved to be quite efficient in terms of average predictive accuracy and speed when compared with the existing methodology.

 Volume 11 | 04-Special Issue

 Pages: 1708-1719