

Radical Web Text Classification Using a Composite-Based Approach

Authors : Kolade Olawande Owoeye, George R. S. Weir

Abstract : The widespread of terrorism and extremism activities on the internet has become a major threat to the government and national securities due to their potential dangers which have necessitated the need for intelligence gathering via web and real-time monitoring of potential websites for extremist activities. However, the manual classification for such contents is practically difficult or time-consuming. In response to this challenge, an automated classification system called composite technique was developed. This is a computational framework that explores the combination of both semantics and syntactic features of textual contents of a web. We implemented the framework on a set of extremist webpages dataset that has been subjected to the manual classification process. Therein, we developed a classification model on the data using J48 decision algorithm, this is to generate a measure of how well each page can be classified into their appropriate classes. The classification result obtained from our method when compared with other states of arts, indicated a 96% success rate in classifying overall webpages when matched against the manual classification.

Keywords : extremist, web pages, classification, semantics, posit

Conference Title : ICCPET 2018 : International Conference on Countering Political Extremists and Terrorist Recruitment

Conference Location : Chicago, United States

Conference Dates : October 10-11, 2018