Mining

A text mining approach to Internet abuse detection The article gives a clear background of the widespread use of computers by companies to communicate with suppliers, clients, and employees and in exchange of data but there has emerged internet abuse at the work place in terms of using the internet for personal reasons. These acts include online chatting, gaming, shopping, investing, illegal downloading, online crimes, pornography, and cybersex. The results of these actions and measures taken by firms are well outlined in the article but there are shortcomings of these measures to deal with internet abuse at the workplace.

The article then develops the basis for the study as a determination of the effectiveness of text mining techniques for detecting Web-based Internet abuse at the work place as well as accuracy and efficiency of some techniques compared to others. This makes the reader have a comprehensive understanding of the article and how it will be organized making it a good and excellent article in terms of idea development. The article then reviews literature on internet abuse, approach and text mining techniques, empirical findings and results and the last part presented is the managerial implications and future research directions as envisaged by the authors. The article is organized in a manner to give the reader a clear understanding from the needs for the research, findings, and conclusions made as well as further research recommendations.

The authors give a clear exposition of the literature on internet abuse to bring the reader up to speed on the research undertaken on internet abuse in the work place. There are three major categories of literature reviewed by the authors including measures to prevent internet abuse behavior, measurements of evaluating the behavior and factors that augment the

behavior as well as the monitoring and blocking of employees internet usage. The main approach taken by the study is the detection of web-based abuse through text mining approach excluding non-web based abuse and e-mail abuse as a means to complement the existing techniques. The authors give a clear exposition to the reader of the methods of classification, data collection methods and preparation as well as the design taken to ensure the results reached are quality and supported by evidence.

After the results are analyzed using standard regression, standard deviations, Chi-test, and F-test, the conclusions and recommendations are presented at the end of the article. The analysis is in-depth to allow the reader to comprehend the steps taken to reach the conclusions made at the end of the article. The article makes conclusions including that text categorization as a promising approach for detecting internet abuse at the workplace as it yields accurate automatic classifiers of 99% accuracy in classification. This is a higher accuracy compared to 80% accuracy of commercial filtering products. The article also concludes that the effectiveness of technical solutions to solve internet abuse in the work place can be augmented using management solutions in terms of policy formulation on internet use and the management may use filtering tools as a means to solve internet abuse at the workplace. Through these conclusions, the article gives a clear exposition of the measures through which the management can benefit through using text mining approach in the detection of internet use at the workplace.

The article does an excellent work of giving recommendations for further research to augment use and understanding of measures to detect internet abuse at the workplace. Future research is proposed on the comparison of

the approach presented in this study and commercial filtering products as well as the ability to make continuous updates to classifiers. The other area is further research is to find out other measures to combine various techniques for better performance and evaluation of internet abuse detection at the workplace.

Reference

Chen-Huei Chou, Atish P. Sinha, and Huimin Zhao. A text mining approach to Internet abuse detection. Inf. Syst. E-Business Management 6(4): 419-439 (2008). Print.