A Comparative Analysis of Data Mining Techniques in Social Media Networks

A.C. Nanayakkara[1*], B.T.G.S. Kumara[2], and M.K.T. Rathnayaka[3]

[1]*Center for Computer Studies, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka*
[2]*Department of Computing and Information Systems, Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka*
[3]*Department of Physical Sciences and Technologies, Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka*
[*]*chethana@ccs.sab.ac.lk*

Social networks have gained exceptional attention within the last decade. Social network sites like Twitter, Facebook, YouTube, and LinkedIn are evolving in a speckled fashion. Users rely on social networks for both information and entertainment needs. Social media analytics with data mining technology could be an analysis axis cantered on extracting trends, patterns, and rules from the social media pool, to serve the people and organizations to have optimum choices concerning many disciplines. The traditional media analytical techniques appear obsolete and inadequate to gratify this immense array of unstructured social media knowledge characterized by three key problems namely; size, noise, and dynamism, predominantly shifting from the batch scale to the streaming one. The objective of this study is to investigate the data mining techniques that were used by social media networks between 2010 and 2018. The study demonstrates a systematic review of analysing trends and content analysis of studies within the field of social media analytics that were published in databases principally IEEE, Elsevier, ScienceDirect, and ResearchGate. Hundred articles were reviewed in this paper. Content analysis was implemented based on their approach, tools utilized, language, the dataset used, country, year, and nature of the experiment. Data mining techniques were utilized for retrieval of information, statistical modelling, and machine learning that engage data pre-processing, data analysis, and data interpretation. The review discovered that fifteen data mining techniques were employed in social media data while frequently used in Support Vector Machine, Bayesian networks, and Decision Tree. The study focused on assisting the involved analysers and educators to capture the research trends and problems associated with the Social media analytics process with future research initiatives.

Keywords: *Data mining, Data mining techniques, Social media, Social media analysis*