

E. Dritsas, I.E. Livieris, K. Giotopoulos and K. Theodorakopoulos. [An Apache Spark implementation for graph-based hashtag sentiment classification on Twitter](#)

. In

Proceedings of ACM 21th Panellenic Conference on Informatics (PCI'18)
, 2018.



Abstract - Sentiment Analysis has been extensively investigated in recent years as a method of human emotions' classification to specific events, products, services etc. It is considered as a very important problem, especially for organizations or companies who want to know the consumers' view about their products and services. In combination with the evolution of social media, it has been established as an interesting domain of research. Through social media, people tend to express their opinions or feelings, such as happiness or sadness on a daily basis. Thus, the vast amount of available data has made the existing solutions inappropriate and the need for automated analysis methods is imperative. In this paper, we examine sentiment polarity analysis on Twitter data in a distributed environment, known as Apache Spark. More specifically, we propose three classification algorithms for tweet level sentiment analysis in Spark due to its suitability for Big Data processing against its predecessors, MapReduce and Hadoop.