

6th International Conference on Smart Computing and Communications, ICSCC 2017, 7-8  
December 2017, Kurukshetra, India

## An Extensive study of Sentiment Analysis tools and Binary Classification of tweets using Rapid Miner

Vishal Vyas<sup>\*a</sup>, V.Uma<sup>b</sup>

<sup>a</sup>Vishal Vyas, Department Of Computer Science, Pondicherry University, Puducherry- 605014, India

<sup>b</sup>V.Uma, Department Of Computer Science, Pondicherry University, Puducherry- 605014, India

---

### Abstract

The online content includes conversation in social networking websites, tweets, blogs and various forums discussing occasions, people and everything which exist in this world. With the huge growth of the online content, high rated information is achievable. Customer feedbacks help organizations to improve their services by rectifying their drawbacks. Manual analysis of information in the era of big data would be a cumbersome task. Countless tools are available for mining information/sentiment from World Wide Web but the choice of tool is the biggest problem at present time. One should have an idea about the framework to evaluate tools. With the evaluation structure for the tools, this paper will compare twenty tools for text analysis with respect to their applications and extension availability for Sentiment Analysis (SA). An experiment is performed in Rapid Miner to derive Sentiment from tweets and accuracies of different algorithms are compared to find out the best performing one.

© 2018 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications

**Keywords:** Text mining; Sentiment analysis; Twitter; Rapid Miner

---

### 1. Introduction

The errand of Sentiment examination in social media content is troublesome because of variability and intricacy of dialect articulation. The not only complexity that makes the task difficult but also the availability of real-time content in a huge quantity is an issue. Manual sentiment analysis is an unmanageable task hence an efficient tool or intelligent system is the need of the hour. Organisations need an intelligent tool to formulate an excellent insight from web-based social networking information. Yet the analysis provided by the vast majority of the tools does not meet their expectations. Tools which monitor online content in social networking websites empower ventures to analyse clients

---

\* Corresponding author. Tel.: +91-7598593973.

E-mail address: [vyasvishaluni@gmail.com](mailto:vyasvishaluni@gmail.com)

supposition, inquiries at real time in a profoundly versatile manner and many more. Availability of tools to monitor online content has quickly expanded in recent years. Enterprises are confronted with the troublesome assignment of picking the right tool according to their necessity. There are numerous vendors who add new components to empower existing tool which helps clients to transform online networking information into significant knowledge. Machine learning (ML) algorithms such as support vector machine[1], decision tree[2], naive bayes[3] and many more acts as an important component while designing a tool. Various supervised[4] and unsupervised learning[5] machine learning classifications are used according to the tool's purpose. This results in a variety of tools but the user gets confused during the selection of a tool to empower their organisation with social media intelligence. These tools offer means for observing the online networking users, dissecting and measuring their sentiments toward a brand, prepare multifarious insight which ultimately helps organizations to improve or enhance their services. Concepts of choosing a tool for text analysis/SA are discussed in this paper. These concepts to evaluate tools comprises of seven areas which help in the selection of the best tool. Considering these concepts, any organisation approaching a new tool not only gets a right tool for getting insight from online media but also benefits the business to reach new heights [6]. Concepts such as Current offering, Workflow management, Market presence, Analysis, Engagement, Strategy and Influence are collectively called as a framework for evaluating a tool. Which tool is beneficial for an organization? To answer this question one must have the knowledge of the framework to evaluate intelligent tools. Section 2 of this paper provides a detailed explanation of a framework to evaluate a tool. In this section, all the necessary concepts which should be taken care of before approaching a new tool for sentiment analysis are clearly explained. In section 3 various tools are categorized with respect to parameters such as application, web sources and extensions available for Sentiment Analysis. The Categorization table provided will help researchers in getting a brief insight of available tools for sentiment analysis. Section 4 will present an experiment with twitter dataset using Rapid Miner to deduce sentiment (positive/negative).

### Nomenclature

SA - Sentiment analysis  
 SVM- Support Vector Machine  
 OSS - Open source software  
 NLP - Natural language processing

## 2. Framework to evaluate a tool

Choosing a tool from the market which at present has nearly two hundred tools is a difficult task. Which tool does what is identifiable but the question is that what thing/concepts should be taken care of while approaching a new tool. How it is different from other available tools which also perform the same task can only be known if certain concepts are considered. Concepts on which one should focus while approaching a new tool are Analysis, Current offering, Engagement, Influence, Strategy, Market presence and Workflow management. Considering these concepts one can identify an efficient tool which can help the enterprise. These seven concepts form a framework to evaluate a tool. In this list Current Offering, Strategy and Market presence are suggested by [Hofer-Shall et al][7]. Considering this list as insufficient [Stavarakantonakis et al][6] added more areas and made a detailed framework for the evaluation of tools.

### 2.1. Current Offering

Satisfaction of a customer should be the top priority of an enterprise. By satisfaction, it means the tools ability to record social media information, its examination and incorporation with the customers requirement. A customer should assess how well an enterprise set up and actualizes enquiries propose new extensions and associated social information with business statistics.

Download English Version:

<https://daneshyari.com/en/article/6900626>

Download Persian Version:

<https://daneshyari.com/article/6900626>

[Daneshyari.com](https://daneshyari.com)