IRJCS: International Research Journal of Computer Science
Volume 13, Issue 04, April 2026
ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1304.18

Comparative Study of Machine Learning Platforms for Text Analytics

Dr.Puja M.Dadhe
Assistant Professor, Department of Computer Science,
Shivaji Science College, Nagpur, India
poojadadhe@gmail.com
https://orcid.org/0009-0008-8696-528X
Dr.Dileep S.Sadhankar
Assistant Professor, Department of Computer Science,
SFS College,Nagpur, India
dileep.sadhankar@gmail.com
https://orcid.org/0009-0001-9453-6265
Dr.Ravikant N.Jugele
Professor and Head, Department of Computer Science,
Shivaji Science College, Nagpur, India
rn_jugele@yahoo.com
https://orcid.org/0009-0004-9395-5247

Publication History
Manuscript Reference No: IJIRAE/RS/Vol.13/Issue04/CSAP26.APCS10098
Research Article | Open Access | Double-Blind Peer-Reviewed| Article ID: IJIRAE/RS/Vol.13/Issue04/CSAP26.APCS10098
Received:20,March 2025, Revised: 02, April 2026, Accepted: 22,April 2026,Published Online: 02, May 2026.
https://www.irjcs.com/volumes/Vol13/iss-04/18.APCS10098.pdf
Article Citation:Dr.Puja,Dr.Dileep,Dr.Ravikant(2026), Comparative Study of Machine Learning Platforms for Text Analytics ,IRJCS: International Research Journal of Computer Science, Volume 13, Issue 04 of 2026 pages 562-565
Doi:> https://doi.org/10.26562/irjcs.2026.v1304.18
BibTeX Key: Dr.Puja@2026 Comparative
IRJCS papers should be cited as IRJCS (International Research Journal of Computer Science, AM Publications, India 2026, ISSN 2393-9842, https://doi.org/10.26562/irjcs.2026.v1304.18
The journal's official abbreviation is IRJCS. Orcid: https://orcid.org/0009-0004-9398-7488

Abstract: In this paper, a comparative study of popular machine learning environments used for text analytics is discussed. These environments include Weka, RapidMiner, Python-based environments, and RStudio. For this purpose, a study of existing literature on text analytics environments is conducted. These environments are compared on the basis of usability, scalability, and flexibility, including their suitability for natural language processing. The comparison is conducted by keeping in view the distinction between GUI-based environments and programming-based environments. At the same time, research gaps and trends in text analytics environments are identified. These gaps and trends are beneficial for choosing suitable environments for sentiment analysis and text mining.

1. INTRODUCTION

The exponential rise in digital information creation through social media, online reviews, blogs, and discussion forums has witnessed a phenomenal rise in unstructured textual information. This rise in user-generated information is a reflection of their views, feelings, and sentiments about a product, service, event, or even a social issue. Thus, information extraction from this data has assumed a significant role in both academic and business environments. Sentiment Analysis, also known as Opinion Mining, has gained prominence in the domain of Natural Language Processing [2]. The primary focus of sentiment analysis is to identify the polarity of the text, which is normally categorized as positive, negative, or neutral. Sentiment analysis is a critical concept in many areas, and these areas include business intelligence, healthcare, politics, and social media monitoring. Organizations are increasingly using sentiment analysis to improve their business and achieve competitive advantage. With the widespread adoption of Web 2.0 tools, the importance of sentiment analysis has been further emphasized by the generation of vast amounts of user data. Over the years, the approach used in sentiment analysis has moved from lexicon-based approaches to machine learning and deep learning techniques [11][12]. In the conventional approach, lexicon-based approaches use predefined dictionaries containing sentiment-carrying words, whereas machine learning-based approaches use algorithms such as Naïve Bayes, SVM, Decision Trees, etc., for the purpose of text classification based on the extracted features. Recently, deep learning-based approaches have shown promising results in the context of improving the performance of the sentiment classification task [5][15]. However, the performance of the sentiment analysis task depends on various factors such as data pre-processing, feature extraction, etc.In order to use these techniques, various machine learning platforms and tools have been developed with different levels of complexity and functionality. GUI-based platforms like Weka and RapidMiner provide a friendly interface where users can use the tools for text analytics without the requirement of programming knowledge. These tools are best for beginners and can be used for educational purposes.

ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1304.18

On the other hand, programming-based tools like Python's scikit-learn and RStudio provide flexibility and can be used for more complex natural language processing. However, this is a challenge that researchers and practitioners face in choosing the most appropriate tool for their needs. Each of these platforms has different characteristics in terms of ease of use, performance, and flexibility in using advanced techniques like deep learning and big data processing. Although tools like GUI-based tools are easy to use, they are not flexible enough for more complex applications [10]. On the other hand, tools like programming-based tools are very flexible but require technical skills. It is therefore essential to conduct a comprehensive comparative analysis of these tools to understand their potential and applicability in different scenarios. In this paper, a systematic review and comparative analysis of some of the most popular machine learning platforms that can be used for text analytics will be conducted. Specifically, their potential in sentiment analysis will be considered. In this study, these tools will be reviewed and compared based on different factors like ease of use, flexibility, scalability, and support for natural language processing. In addition, gaps and trends in this field will be identified.

2. LITERATURE REVIEW

Sentiment analysis and text analytics have been extensively explored in the last two decades. Initially, sentiment analysis and text analytics were explored using conventional machine learning algorithms. Research in sentiment analysis and text analytics started with Bing Liu's work on laying the foundation for sentiment analysis using opinion mining and polarity detection techniques [1][8]. Moreover, Bo Pang and Lillian Lee conducted a thorough investigation of sentiment classification techniques using machine learning algorithms like Naïve Bayes and SVM for text classification [9].

The subsequent research was aimed at improving the accuracy of the sentiment classification results through the integration of feature engineering techniques. Various research works have demonstrated the effectiveness of techniques like term frequency-inverse document frequency, n-gram models, and part-of-speech tagging in improving the accuracy of the results [3]. Apart from these, there have been suggestions for the development of hybrid models based on lexicon-based methods and machine learning models to avoid the limitations associated with these techniques [4][7].

With the development in the field of machine learning, various tools have been developed to make text analytics more accessible. Weka has been widely used in academic research due to its easy-to-use graphical interface.

According to Yin, H., Li, S., & Lu, W. (2015), Weka offers a complete environment for data preprocessing, visualization, and evaluation, making it more suitable for academic purposes [18]. However, its limited capability in dealing with large-scale data has been identified as a limitation.

Similarly, RapidMiner has achieved popularity as a data science platform based on workflow, enabling users to develop analytical processes using a drag-and-drop interface. According to research, RapidMiner provides better scalability and integration capabilities compared to Weka. Nevertheless, it has limitations in being flexible enough for customized NLP tasks [6].

Contrary to GUI-based tools, Python and R environments have dominated recent trends. Python environments, such as scikit-learn, have achieved extensive support for machine learning algorithms, making them suitable for sentiment analysis tasks. Fabian Pedregosa et al. demonstrated the efficiency and scalability of scikit-learn for data processing and classification tasks [16]. In addition, the availability of libraries such as NLTK and spaCy has greatly improved Python's capabilities for NLP tasks [13].

Similarly, R Studio has gained popularity in statistical computing and text mining. Packages like tm and tidytext are available for efficient processing and analysis of text data. A research by Islam,S.,& Islam,S. indicates that R-based solutions are powerful in text mining and sentiment analysis-based applications [23].

Recent research has shown a trend towards deep learning techniques that have shown promising results in terms of performance in sentiment classification-based applications. Techniques like RNNs, CNNs, and transformer-based techniques like BERT have shown promising performance in this context [9]. However, these require high computational power and are available in programming environments rather than GUI-based tools.

Several comparative studies have been conducted to compare different sentiment analysis tools and techniques. Medhat et al. have presented a detailed survey of sentiment analysis techniques, highlighting the advantages and disadvantages of different approaches [1].

Similarly, Minaee et al. have presented a review of different text classification techniques based on deep learning approaches, highlighting the increasing role of deep learning in the field of NLP [12]. However, few studies have specifically addressed the comparative analysis of different machine learning platforms in terms of ease of use, flexibility, and efficiency.

Although significant advancements have been made in the field of sentiment analysis, there are still a number of challenges that need to be addressed. Some of the challenges include the analysis of sarcastic sentences, domain dependency, multilingual text, and large amounts of text data. However, a comparative analysis of different machine learning platforms in the context of text analysis has not yet been conducted, which highlights the need for a comparative analysis of different machine learning platforms used in text analysis.

3. OVERVIEW OF MACHINE LEARNING PLATFORMS

Machine learning platforms are very helpful in simplifying text analytics and sentiment analysis. These platforms vary in terms of their usability, flexibility, scalability, and their ability to support natural language processing. This section gives a brief overview of popular machine learning platforms, including Weka, RapidMiner, Python, and RStudio.

IRJCS: International Research Journal of Computer Science
Volume 13, Issue 04, April 2026
ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1304.18

3.1 Weka

Weka is machine learning software developed at the University of Waikato. It is a popular, open-source machine learning software that provides a graphical user interface that allows users to carry out data preprocessing, classification, clustering, and visualization without requiring programming knowledge. Some of the machine learning algorithms that Weka supports includes Naïve Bayes, decision trees, and support vector machines [22]. For text analytics, Weka supports several machine learning algorithms, including StringToWordVector, which is used for converting text data into numerical features [17]. Even though it is widely used for academic purposes, Weka has several limitations in terms of scalability and its ability to support natural language processing compared to other platforms [19][21].

3.2 RapidMiner

RapidMiner is a powerful data science tool that provides a visual workflow design experience through a drag-and-drop interface. It allows users to create analytical workflows by combining different operators. RapidMiner provides extensions for text mining, including features like tokenization, filtering, and feature extraction. It has better scalability and supports external data sources compared to Weka. However, it is inflexible compared to programming-based tools, especially for advanced NLP and deep learning techniques [20].

3.3 Python with scikit-learn

Python is one of the most popular programming languages for machine learning and text analytics. scikit-learn is a library that provides efficient implementations of classification algorithms, including Naïve Bayes, SVM, and logistic regression. In addition to scikit-learn, Python supports advanced NLP libraries like NLTK and spaCy, which provide features for text analytics, including tokenization, stemming, and named entity recognition. Python is also suitable for deep learning techniques, making it a good choice for sentiment analysis. However, it requires programming skills and has a higher learning curve compared to GUI-based tools [16][14].

3.4 R Studio

R Studio is an Integrated Development Environment for R language programming. It is commonly used for statistical analysis and computation. R Studio offers good support for text mining using tm and tidytext libraries. R Studio is best used for data visualization and exploratory data analysis [25]. It helps researchers to conduct a comprehensive statistical analysis of their data. However, like Python, R Studio also demands programming skills and is not easy to use for beginners [24].

4. COMPARATIVE ANALYSIS

A comparative analysis of machine learning platforms is essential to understand their strengths and limitations in text analytics applications. The comparison is based on key factors such as usability, flexibility, scalability, and NLP capabilities.

Feature	Weka	RapidMiner	Python	RStudio
Interface	GUI	GUI	Code	Code
Ease of Use	High	High	Medium	Medium
Flexibility	Low	Medium	High	High
NLP Capability	Basic	Moderate	Advanced	Advanced
Scalability	Low	Medium	High	High
Deep Learning	No	Limited	Yes	Yes

Table 1: Comparison of Machine Learning Platforms

Comparative analysis of machine learning platforms is necessary to identify the capabilities and limitations of machine learning platforms in text analysis applications. The comparison is done on the basis of various factors such as usability, flexibility, scalability, etc. Although many researchers have worked on the applications of sentiment analysis and text analysis, many gaps exist in the comparative analysis of machine learning platforms such as Weka, RapidMiner, Python using scikit-learn, and RStudio. The studies done by researchers mainly compare the algorithms rather than the platforms, which creates a lack of standardized benchmarking frameworks. The absence of standardized datasets, metrics, and setups makes it difficult to compare the machine learning platforms [20], [22].

The second major gap relates to the disparity between ease of use and the ability to perform sophisticated analysis. GUI-based tools like Weka and RapidMiner are easy to use but lack the ability to perform the latest advancements in deep learning, while tools like scikit-learn and RStudio provide the necessary flexibility and scalability at the cost of ease of use. Furthermore, the lack of large-scale data processing, multilingual analysis, and context-based sentiment analysis (sarcasm and irony) also needs to be addressed in all the tools [22].

These gaps in the tools indicate the need for a comparative analysis that not only compares the tools in terms of efficiency but also compares them in terms of ease of use, scalability, and the ability to perform sophisticated analysis using the latest advancements in NLP tools.

5. CONCLUSION

This paper has presented a comparative analysis of the most commonly used machine learning tools in the field of text analytics, including Weka, RapidMiner, Python with scikit-learn, and RStudio. Based on the analysis, it has been identified that the tools have different strengths and weaknesses in different contexts of application. GUI-based tools like Weka and RapidMiner are more suitable for new learners and academic purposes due to their ease of use and ease of implementation. However, the tools lack flexibility and the ability to perform advanced tasks like deep learning. Programming-based tools like Python and RStudio are highly suitable for the efficient handling of complex and large-scale text analytics tasks, though the tools are highly technical in nature.

REFERENCES

1. M.Medhat, A.Hassan, and H.Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, 2014.
2. B.Liu, Sentiment Analysis and Opinion Mining. Morgan & Claypool, 2012.
3. J.Cui et al., "Survey on sentiment analysis: evolution of research methods and topics," Artificial Intelligence Review, 2023.
4. M.Bordoloi and S.K.Biswas, "Sentiment analysis: A survey on design framework, applications and future scopes," Artificial Intelligence Review, 2023.
5. Marouane Birjali, Mohammed Kasri, Abderrahim Beni-Hssane,"A comprehensive survey on sentiment analysis: Approaches, challenges and trends," Knowledge-Based Systems, volume 226, 2021.
6. Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowledge-Based Systems, 89,14–46.https://doi.org/10.1016/j.knosys.2015.06.015
7. Joshi, M.,Prajapati, P.,Shaikh, A.,& Vala,V.(2017).A survey of sentiment analysis. International Journal of Computer Applications, 163(6), 34–38. https://doi.org/10.5120/ijca2017913552
8. Chandan, M. K., & Mandal, S. (2025).A comprehensive survey on sentiment analysis: Framework, techniques, and applications. Computer Science Review, 55, 100693. https://doi.org/10.1016/j.cosrev.2024.100693
9. B.Pang and L.Lee, "Opinion mining and sentiment analysis," Foundations and Trends in Information Retrieval, 2008.
10. I.H.,Frank, E.,Hall, M.A.,& Pal,C.J.(2016). Data mining: Practical machine learning tools and techniques (4th ed.). Morgan Kaufmann.
11. Silge, J., & Robinson,D.(2017). Text mining with R: A tidy approach. O'Reilly Media.
12. Minaee,S., Kalchbrenner,N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep learning-based text classification: A comprehensive review. ACM Computing Surveys, 54(3), 1–40. https://doi.org/10.1145/3439723
13. Yao,Jiawei.(2019). Automated Sentiment Analysis of Text Data with NLTK. Journal of Physics: Conference Series. 1187. 052020. 10.1088/1742-6596/1187/5/052020.
14. Lorincz, T.,& Bathó,A.(2019).Sentiment Analysis with TextBlob: A systematic review. Information Processing & Management, 56(5), 1797-1812.
15. Socher,R.,Perelygin,A.,Wu,J.Y.,Chuang, J.,Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. Proceedings of the 2013 conference on empirical methods in natural language processing, 1631-1642.
16. Pedregosa, F.,Varoquaux, G.,Gramfort, A.,Michel, V.,Thirion, B.,Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830.
17. Hamdi, A.,& Salah, H.(2019). A Weka-based Framework for Sentiment Analysis. Journal of Intelligent Systems, 28(1), 49-62.
18. Yin, H., Li, S., & Lu, W. (2015). A Hybrid Approach to Sentiment Analysis with Weka. Procedia Computer Science, 55, 693-700.
19. Rosner,T.,& Stabinger,S.(2016). Emotion Recognition in Social Media with Weka. In Proceedings of the 10th International Conference on Semantic Computing (pp. 176-179).
20. Rizal,A.,& Trihastuti,T.(2020). Comparison of Weka and RapidMiner in Sentiment Analysis. International Journal of Emerging Technologies in Learning, 15(12), 55-63.
21. Putri,N.,&Tjandrasa,H.(2020). Aspect-Based Sentiment Analysis of Customer Reviews Using Weka. Journal of Information Systems Engineering and Business Intelligence, 6(2), 110-118.
22. Sivasankari,S., &Anandhi, S. S. (2021). A Comparative Study of Weka and Python in Sentiment Analysis. In Advances in Communication, Devices and Networking (pp. 27-32). Springer, Singapore.
23. Islam,S.,& Islam,S.(2020). Sentiment Analysis in R Using Machine Learning Techniques. Journal of Information Systems Engineering and Business Intelligence, 6(2), 62-70.
24. Shahzad, A.,& Shahzad,M.A.(2020). Sentiment Analysis of Movie Reviews in R. Journal of Big Data, 7(1), 1-17.
25. Loh, N.M.C.,&Tham,J.Y.(2020). Sentiment Analysis of YouTube Comments Using a Hybrid Approach in R. International Journal of Computer Science and Information Security, 18(1), 141-149.