

Digital Tools and Methods

Transforming Literary Criticism

Sigrid Kim

Viennart Academy

Abstract

The advent of digital tools and methods has transformed literary criticism, enhancing traditional analysis through techniques such as text mining, sentiment analysis, and topic modeling. This paper explores the role of these tools in modern literary analysis, demonstrating their potential through case studies. Text mining, for example, helps researchers uncover hidden patterns in literary works, as seen in a study on the authorship of Shakespeare's works, where multiple authorship was suggested using machine learning techniques. Similarly, sentiment analysis allows for the identification of emotional trends in literature, exemplified by analyses of Jane Austen's novels and Pirandello's *War*. Additionally, topic modeling assists in summarizing large literary datasets, though it presents limitations in fiction, as shown by analyses of the Gutenberg corpus. These digital approaches not only provide new insights but also complement traditional methods, offering valuable quantitative data to support qualitative interpretations. However, further research is needed to refine these techniques and fully realize their potential in literary studies.

Introduction

Literary criticism can be described as the effort made to explain a piece of literary work, including aspects relating to its production, design, beauty, meaning, and strengths and weaknesses. Attesting to this, Piklu (2023) defined literary criticism as “the way to discover the elements a specific piece of literature is based on through the provision of different lenses through which we can easily delve into

literature” (p. 1). Similarly, Richards (2001) posits that literary criticism is an attempt to study, analyze, and evaluate imaginative literature. However, it is important to note that although anyone who expresses an opinion about a song, book, movie, or play is recognized as a critic, not everyone's opinion is founded upon reflection, analysis, thought, or consistently articulated principles (Richards, 2001). Furthermore, discussing the importance

of literary criticism, Piklu (2023) explains that it assists us in identifying or exposing the hidden policies and politics covered in a literary work's words. Additionally, the researcher posits that literary criticism helps broaden the reader's understanding so that the value of an author's work is understood.

Historically, early literary criticism was mainly based on the technical aspect of the writing and involved thoroughly reading a piece of literary work to analyze the diction, plot, character, etc. (Piklu, 2023). However, with the advent of technology, digital tools are now revolutionizing how we analyze literature (Mustifa & Lestari, 2023). Digital tools and methods, such as sentiment analysis, topic modeling, and text mining, are now employed in literary analysis. Thus, this essay aims to examine the use of these digital tools and methods in literary criticism, identifying the impact and potential through specific case studies.

Text Mining in Literary Criticism

Text mining refers to the process of semi-automatically extracting information and using techniques to mine knowledge from a large amount of unstructured data (Talib et al., 2016). It has emerged as a valuable way of analyzing literary works, especially in this digital age where the quantities of digital textual data are continually increasing (Hassani et al., 2020). Although Delen and Crossland (2008) say that text mining is meant to go along with human-driven literary analysis and not replace it, the benefits are seen in how it can help find valuable insights that might not be easy to find through human close reading. Furthermore,

text mining provides researchers and students with the ability to develop a sense of coherence, trends of research streams, and topic clustering among the continually increasing pool of literary works (Delen & Crossland, 2008).

Case Study

A notable example of how text mining has been beneficially employed in uncovering new insights into classic and contemporary literature is in the analysis of Shakespeare's works. Researchers have made several attempts to use text-mining techniques to examine various aspects of Shakespeare's works. For example, a researcher used text-mining techniques to mine the question of authorship in Shakespeare's works. According to Ruta (2019), over 154 sonnets and 38 plays are attributed to William Shakespeare, who is considered one of the greatest playwrights in the English language. Given the short span of time (24 years) in which these diverse masterpieces were written, the researcher sought to examine the possibility that Shakespeare's works were authored by multiple people. Thus, the researcher carried out a project using text mining techniques to criticize the issue of the authorship of Shakespeare's works. Using machine learning and text mining techniques such as principal component analysis with random forest classification, KMeans clustering, and linear regression, the researcher revealed that it is possible that multiple authors (including Thomas Middleton) wrote Shakespeare's works.

This case study indicates that literary criticism

is not limited to the analysis of the content of a particular piece of literary work but also extends to the criticism of the production process. Additionally, this case study also reveals the significant role that digital tools and methods play in uncovering new insights and findings that may hitherto be impossible or difficult using traditional literary criticism approach such as deep reading.

Sentiment Analysis in Literary Criticism

Drus and Khalid (2019) defined sentiment analysis as a process of using Natural Language Processing (NLP) to extract, convert, and interpret opinions from a text in order to classify such emotions as either negative, positive, or neutral sentiments. In a similar line of thought, Kharde and Sonawane (2016) also assert that sentiment analysis relates to “a process that automates mining of attitudes, opinions, views and emotions from text, speech, tweets, and database sources through Natural Language Processing (NLP)”(p.5). This process can also be described as opinion mining, appraisal extraction, and subjectivity analysis. Sentiment analysis has been employed by individuals in diverse sectors and fields (Sweta, 2024). In the field of literary studies, Rebora (2023) revealed that sentiment analysis is emerging as a widely discussed topic. According to this researcher, the use of sentiment analysis in literary studies is important, given the exponential growth of literature. Sentiment analysis will assist readers in choosing genres that are best suited to their interests (Rebora,

2023). Sentiment analysis has been used in multiple literary studies. For instance, Reagan et al. (2016) used sentiment analysis to identify the “basic shapes” of stories, Jacobs et al. (2020) used sentiment analysis to conduct a large-scale investigation of children’s literature; while Sprugnoli et al. (2020) used sentiment analysis to study attitudes and emotions in ancient texts in both contemporary and historical languages. In the aspect of fictional text, Adelman et al. (2019) state that sentiment analysis has been used for about 15 years now, especially in the aspect of plots and the development of characters. However, a notable example of the use of sentiment analysis is in the analysis of Jane Austen’s novels.

Case Studies

Katrin Olmann conducted a sentiment analysis of the writing of Jane Austen, a famous nineteenth-century novelist. The analysis was part of her university project and was aimed at criticizing the theme of irony in Austen’s writings. Olmann conducted this analysis based on a large corpus of 18th and 19th-century English novels as well as the Jane Austen Corpus (JAC). Based on the sentiment analysis, Olmann concluded that Austen’s novels can be utilized as a preliminary step when interpreting any literary work or as a pre-reading activity.

Furthermore, another case study of the use of sentiment analysis in literary criticism is the sentiment analysis of English literature using rasa-oriented semantic ontology conducted by

Sreejith et al. (2017). Utilizing the 'Navarasa' ontology' developed by the researcher, the study conducted the sentiment analysis of a short story - War, by Luigi Pirandello. Using sentiment analytical techniques, the study was able to map the emotion lexicons and identify the emotion keywords present in the raw text. Based on this analysis, the researchers discovered that the predominant sentiment of this short story is terror and sorrow, which were intermixed and interrelated in the short story. However, the researchers noted that future researchers and scholars could achieve better accuracy in the use of sentiment analysis to perform literary criticism by expanding their developed ontology to include more phrases and idioms. Thus, this case study revealed that digital tools can be used to save resources in the conduct of literary criticism, proving to be an effective way of identifying predominant themes and sentiments in a particular piece of literary work.

Topic Modeling in Literary Criticism

Topic modeling refers to a machine learning technique that utilizes unsupervised learning to produce a summary set of terms from a large set of texts (Kherwa & Bansal, 2018). Kherwa and Bansal (2018) further explain that the topic modeling technique has the ability to scan a set of documents, detect phrase and word patterns in it, and automatically cluster similar expressions and word groups that best characterize a set of text or documents. The significance of topic modeling lies in its ability

to combine different topics into a single, understandable structure. The 'topics' described in this context relate to a "recurring pattern of words that best represents a theme within the documents" (GeeksforGeeks, 2024). Additionally, topic modeling helps extract valuable insights and information from unstructured data, enhances the organization and retrieval of content, and enables the analysis and monitoring of trends (GeeksforGeeks, 2024).

Regarding the use of topic modeling in literary criticism, Uglanova et al. (2020) state that although topic modeling is acknowledged as a statistical tool for thematically decomposing texts, it only captures the statistical patterns in the structure of the object in reality. Thus, Uglanova et al. (2020) posit that given that structure itself is a relevant feature with an artistic function in literary texts, the sensitivity of the method for structure makes it less effective when applied to literary texts.

Case Studies

A notable case study of the use of topic modeling in literary criticism is "Topic Modeling Literary Interviews from The Paris Review" by Greene et al. (2024). The study was focused on using computer-assisted techniques to mine literary interviews. The researchers utilized a seven-decade dataset of literary interviews from The Paris Review. Topic modeling was conducted on these documents to identify the predominant themes dominating

the “culturally significant set of materials” while also examining the significance of topic modeling in conducting literary criticism. Consisting of about 404 in-depth interviews with about 399 different writers between 1953 and 2019, the Paris Review had a text range of between 1,284 and 17,135 words in total. The topic modeling conducted on these datasets by Greene et al. (2024) revealed topic models that demonstrated that while some interviewees speak about film as an industry, others speak of it as a practice. This uncovers the perspectives of literary writers on the film industry. This case study indicates that topic modeling can be applied in literary criticism to provide quantitative data that will complement the existing qualitative arguments, thereby revealing insights that might have been missed and validating existing perspectives.

Another interesting case study of the use of topic modeling in literary criticism is the project by Eisner et al. (n.d.) to find out the outcome of running topic models on literary texts. Given their observed neglect of fiction works in the application of topic models, the researchers decided to run a simple topic model to analyze the latent thematic structure of the texts in a subset of the Gutenberg corpus. The researchers found that although the topic modeling technique was able to discover interesting latent topics in the large dataset of texts, the topics identified did not significantly inform readers of the content of the works and were not exactly meaningful thematically.

Furthermore, as the topic models that emerged primarily consisted of names, the researchers concluded that such topics may not be of much use, especially in the organization of a large collection of literature. In summary, this case study highlights the limitations of digital tools, specifically topic modeling in literary criticism and analysis.

Conclusion

It can be observed from the arguments presented in this essay that the role of digital tools and methods in this digital age cannot be understated. In the field of literary works and literary analysis, digital tools such as text mining, sentiment analysis, and topic modeling have emerged as a significantly beneficial way of conducting literary criticism that may have been impossible or difficult using traditional methods. However, despite the arguments and findings presented in this essay, it is important to note that the application and adoption of digital tools and methods in literary studies and analyses warrants further attention and research. Thus, it is recommended that future research focus on examining how digital tools and methods such as text mining, machine learning, and sentiment analysis can be used to identify the future trends of literary works.

References

1. Adelman, B., Andresen, M., Begerow, A., Franken, L., Gius, E., & Vauth, M. (2019). Evaluation of a semantic field-based approach to identifying text sections about

- specific topics. DH 2019 - Digital Humanities Conference. <http://hdl.handle.net/11420/4538>
2. Delen, D., & Crossland, M. D. (2008). Seeding the survey and analysis of research literature with text mining. *Expert Systems with Applications*, 34(3), 1707-1720. <https://doi.org/10.1016/j.eswa.2007.01.035>
 3. Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. *Procedia Computer Science*, 161, 707-714.
 4. Eisner, E., Futoma, J., & Rice, D. (n.d.). Topic modeling on literary texts. Dartmouth College. <https://www.cs.dartmouth.edu/~lorenzo/teaching/cs174/Archive/Winter2013/Projects/FinalReportWriteup/joseph.d.futoma.14.pdf>
 5. GeeksforGeeks. (2024, May 14). Topic modeling—Types, working, applications. GeeksforGeeks. <https://www.geeksforgeeks.org/what-is-topic-modeling/>
 6. Greene, D., O'Sullivan, J., & O'Reilly, D. (2024). Topic modelling literary interviews from The Paris Review. *Digital Scholarship in the Humanities*, 39(1), 142-153. <https://doi.org/10.1093/lc/fqad098>
 7. Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1). <https://doi.org/10.3390/bdcc4010001>
 8. Jacobs, A. M., Herrmann, B., Lauer, G., Lüdtke, J., & Schroeder, S. (2020). Sentiment analysis of children and youth literature: Is there a Pollyanna effect? *Frontiers in Psychology*, 11, 574746. <https://doi.org/10.3389/fpsyg.2020.574746>
 9. Kharde, V., & Sonawane, P. (2016). Sentiment analysis of Twitter data: A survey of techniques. *International Journal of Computer Applications*, 139(11), 5-15. <https://doi.org/10.5120/ijca2016908625>
 10. Kherwa, P., & Bansal, P. (2018). Topic modeling: A comprehensive review. *ICST Transactions on Scalable Information Systems*, 0(0), 159623. <https://doi.org/10.4108/eai.13-7-2018.159623>
 11. Mustifa, A., & Lestari, L. A. (2023). Literary criticism in the digital age: Addressing the problems and opportunities of digital literature in EFL pedagogy. *English Review: Journal of English Education*, 11(1), 263-274. <https://doi.org/10.25134/erjee.v11i1.7137>
 12. Paul, P. K. (2023). Literary criticism and its significance. <https://doi.org/10.13140/RG.2.2.28890.72640>

13. Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5(1), 31. <https://doi.org/10.1140/epjds/s13688-016-0093-1>
14. Rebora, S. (2023). Sentiment analysis in literary studies: A critical survey. *DHQ: Digital Humanities Quarterly*, 17(3).
15. Richards, I. A. (2001). *Principles of Literary Criticism* (1st ed.). Routledge. <https://doi.org/10.4324/9781351223508>
16. Ruta, N. (2019). Text mining for the analysis of Shakespeare. Harvard University. <http://nickruta.com/shakespeare.pdf>
17. Sprugnoli, R., Passarotti, M., Corbetta, D., & Peverelli, A. (2020, May). Odi et Amo: Creating, evaluating and extending sentiment lexicons for Latin. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 3078-3086).
18. Sreejith, D., Devika, M. P., Tadikamalla, N. S., & Mathew, S. V. (2017). Sentiment analysis of English literature using rasa-oriented semantic ontology. *Indian Journal of Science and Technology*, 10(24), 1-11.
19. Sweta, S. (2024). Application of sentiment analysis in diverse domains. In *Sentiment Analysis and its Application in Educational Data Mining* (pp. 19-46). Springer Nature Singapore.
20. Talib, R., Kashif, M., Ayesha, S., & Fatima, F. (2016). Text mining: Techniques, applications and issues. *International Journal of Advanced Computer Science and Applications*, 7(11). <https://doi.org/10.14569/IJACSA.2016.071153>
21. Uglanova, I., Gius, E., Karsdorp, F., McGillivray, B., Nerghe, A., & Wevers, M. (2020). The order of things: A study on topic modelling of literary texts. *CHR*, (18-20).