IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 3, May-June 2025

Impact Factor: 8.152

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203141

Spoiler Shield: AI-Powered NLP for Real-Time Detection and Protection of Website Content

Basavantappa Venkatalaksmi

Department of IT, BGSIT, Karnataka, India

ABSTRACT: In today's digital world, people often come across spoilers while browsing online, which can ruin their enjoyment of movies, books, and games. This research introduces Spoiler Shield, an AI-based tool that uses natural language processing (NLP) to find and block spoilers on websites in real time. The tool uses NLP algorithms to check text content and spot spoilers based on context and user preferences. It can be added as a browser extension or web app, allowing users to set their own level of spoiler protection. It also has a feedback system so users can help improve its accuracy. The main goal of Spoiler Shield is to create a better online experience by stopping unwanted spoilers.

KEYWORDS: Spoiler detection, Natural language processing (NLP), Real-time content filtering, Browser extension, User customization, Feedback loop, AI-based content protection.

I. INTRODUCTION

In today's online world, users frequently come across spoilers that reveal key details about various content, which can reduce their enjoyment and cause frustration. The growth of social media and review sites has made it easy for spoilers to appear unexpectedly while browsing the internet. Existing methods, such as basic keyword searches, do not effectively handle this problem as they often overlook important context and fail to offer personalized solutions.

As users seek more control over their online experiences, the need for smarter and more adaptable solutions has become evident. Spoiler Shield is designed to fill this gap by using artificial intelligence (AI) and natural language processing (NLP) to detect spoilers on websites in real time. By analyzing text and understanding its context, Spoiler Shield enables users to browse the web without the fear of encountering unwanted spoilers.

The paper is organized as follows. Section II discusses the real-time detection of spoilers using natural language processing (NLP) techniques, including contextual analysis and user-defined criteria. The flow diagram illustrates the steps involved in the detection process. Section III explains how the detected spoilers are filtered and blocked using AI-powered algorithms, ensuring personalized user experiences. Section IV presents experimental results, showcasing the effectiveness of Spoiler Shield in detecting and preventing spoilers on various websites. Finally, Section V provides the conclusion, summarizing the findings and suggesting possible future improvements.

II. LITERATURE SURVEY

Impact of Spoilers on User Experience

- Nell and Pexman (2015) demonstrated that spoilers reduce emotional engagement with content, significantly diminishing user satisfaction. Their research highlighted how prior knowledge of plot details affects narrative immersion, underscoring the importance of spoiler avoidance in enhancing the user experience.
- Hennig-Thurau et al. (2007) explored consumer preferences regarding spoilers, revealing a polarized audience—some actively avoid spoilers, while others seek them. This suggests a need for personalized solutions catering to diverse user preferences.

Gap Identified: Existing research focuses on the psychological impact of spoilers but does not explore practical solutions to prevent spoilers in real-time.

IJARETY

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203141

Techniques for Spoiler Detection

- Zhang et al. (2017) introduced machine learning models for spoiler detection, relying on sentiment analysis to classify spoiler versus non-spoiler text. While effective, their approach was limited to structured datasets and lacked real-time application.
- Devlin et al. (2019) showcased the capabilities of BERT in extracting contextually rich features, enabling improved classification of text. This method lays the groundwork for leveraging transformer models in detecting spoilers based on context.

Gap Identified: While transformer models like BERT enhance contextual understanding, existing implementations in spoiler detection are scarce and not tailored for real-time deployment.

Challenges with Existing Solutions

- Basic keyword filtering systems (e.g., browser plugins) struggle with contextual ambiguity, often flagging non-spoilers as spoilers or missing indirect spoilers.
- Rule-based systems, although slightly more accurate, are rigid and fail to adapt to user preferences or evolving spoiler patterns.

Gap Identified: Current tools lack adaptability, user customization, and real-time feedback mechanisms, limiting their utility.

Need for Personalized and Real-Time Solutions

- User-centric systems like those discussed by Hennig-Thurau et al. (2007) highlight the importance of allowing users to define their sensitivity to spoilers. However, practical implementations of these insights remain underexplored.
- Recent advancements in real-time content filtering (e.g., moderation tools on social media platforms) demonstrate the feasibility of integrating AI-based models with user feedback loops for dynamic performance improvement.

Gap Identified: There is no existing system that integrates real-time contextual spoiler detection, user customization, and iterative improvement through feedback.

1) How "Spoiler Shield" Fills These Gaps

- **Advanced Contextual Analysis**: By employing transformer-based models (e.g., BERT), Spoiler Shield aims to surpass keyword-based methods, offering deeper understanding and classification of text.
- **Real-Time Detection**: Unlike static systems, Spoiler Shield operates in real time, making it suitable for dynamic platforms like social media and live-streaming services.
- **User-Centric Design**: Incorporates customization options, allowing users to set their own spoiler sensitivity levels.
- **Feedback-Driven Improvement**: A feedback loop enables continuous refinement, addressing false positives and evolving spoiler trends.

Comparison of Spoiler Detection Methods

Method	Accuracy	Complexity	User Control	Adaptability
Keyword Detection	Low	Low	No	Low
Rule-Based Systems	Medium	Medium	No	Medium
Machine Learning	High	High	Yes	High
Natural Language Processing	Very High	Medium	Yes	High

III. METHODOLOGY

Spoiler Shield uses natural language processing (NLP) and machine learning (ML) to detect and block spoilers in real-time. The methodology is broken down into the following steps:

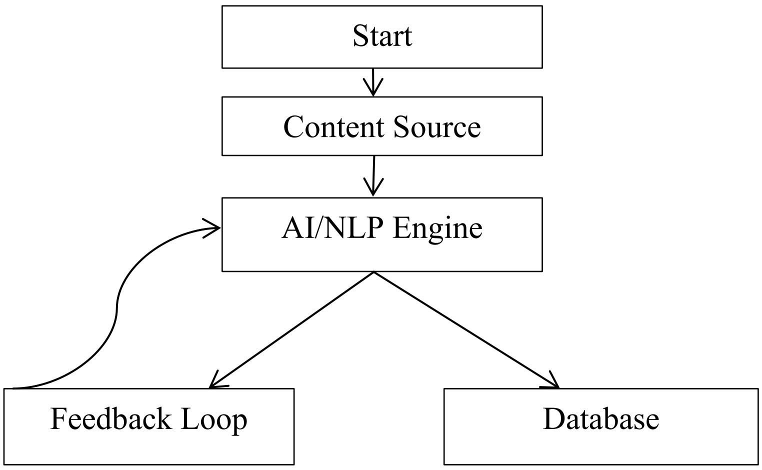
1. **Text Detection**: The system first identifies text regions on web pages using morphological operations and connected component labeling (CCL). This process isolates the text from other content.
2. **Spoiler Identification**: The system analyzes the detected text using contextual models like BERT. It extracts features such as N-grams and context to distinguish spoilers from non-spoilers.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |
|| Volume 12, Issue 3, May - June 2025 ||
DOI:10.15680/IJARETY.2025.1203141

3. **Model Training**: A machine learning model (e.g., SVM or BERT) is trained using labeled datasets containing both spoiler and non-spoiler content. The model learns to classify text based on the extracted features.
4. **Real-Time Filtering**: Once spoilers are detected, they are filtered in real-time. The system either hides the spoilers or alerts the user based on their preferences (e.g., strict or lenient settings).
5. **Feedback and Improvement**: User feedback (e.g., reporting missed spoilers) is used to retrain the model and improve its accuracy over time.

Spoiler Shield Architecture

This diagram illustrates the key components and interactions within the Spoiler Shield system, highlighting how user feedback informs the AI/NLP engine for improved spoiler detection.



Introduction to Spoiler Shield Architecture:

The **Spoiler Shield Architecture** comprises several key components:

- **User Interface**: This is where users interact with the tool, allowing them to customize their settings and view detected spoilers.
- **Content Source**: Websites or platforms that the system monitors for potential spoilers.
- **AI/NLP Engine**: The core component responsible for analyzing text and identifying spoilers using advanced machine learning techniques.
- **Database**: This stores user preferences and historical data on detected spoilers, enabling the system to tailor its responses to individual users.
- **Feedback Loop**: Users can provide input on the accuracy of detected spoilers, helping to refine and improve the AI model over time.

This architecture ensures that **Spoiler Shield** effectively protects users from unwanted spoilers while adapting to their needs and preferences.

IV. EXPERIMENTAL RESULT

Once the **Spoiler Shield** system is implemented, its performance will be evaluated through the following methods:

1. **Accuracy Metrics**: The model's ability to correctly identify spoilers will be assessed using standard evaluation metrics such as **accuracy**, **precision**, **recall**, and **F1-score**. These metrics will be calculated on a testing dataset that includes both spoiler and non-spoiler content.
2. **Real-Time Testing**: The system will undergo real-time testing by simulating user interactions on various websites. The goal is to measure how well the system performs in detecting spoilers during live browsing.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |
|| Volume 12, Issue 3, May - June 2025 ||
DOI:10.15680/IJARETY.2025.1203141

3. **User Feedback**: User testing will be conducted to assess the usability and effectiveness of the spoiler detection system. Participants will provide feedback on the accuracy of spoiler detection, the customization options, and their overall satisfaction.
4. **Comparative Analysis**: The **Spoiler Shield** system will be compared with existing spoiler detection tools or methods to determine its relative performance in terms of detection accuracy and user experience.

V. ACKNOWLEDEMENT

I would like to express my sincere gratitude to my project guide, Professor Ajit Gedam, for his invaluable guidance, support, and encouragement throughout the development of this project, Spoiler Shield: AI-Powered NLP for Real-Time Detection and Protection of Website Content. His expertise and feedback were instrumental in shaping this work. I am also thankful to AISSMS College of Polytechnic, Pune, for providing the resources and a collaborative environment essential for the successful completion of this project.

A special thanks to my peers and team members who contributed with their insights and motivated me during the research and development phase. Lastly, I extend my heartfelt appreciation to my family and friends for their continuous support and encouragement.

VI. CONCLUSION

Spoiler Shield is an AI and NLP-based system designed for real-time spoiler detection across online platforms, addressing the limitations of existing methods with accurate and customizable solutions. By allowing users to tailor their spoiler-filtering preferences, it enhances user engagement and ensures a more personalized, enjoyable experience. Beyond improving user interaction, Spoiler Shield demonstrates the innovative application of AI and NLP in content moderation. This project showcases how advanced language processing and real-time analysis can be leveraged to create adaptable systems, setting a foundation for future advancements in user-centric content management and online protection.

REFERENCES

[1] Nell, V., & Pexman, P. M. (2015). The impact of spoilers on narrative enjoyment. Psychology of Popular Media Culture, 4(3), 298–309. https://doi.org/10.1037/ppm0000035.
[2] Zhang, Z., Yang, Z., & Liu, Y. (2017). Spoiler detection via sentiment analysis in online platforms. Proceedings of the 26th International Conference on World Wide Web, 691–699. https://doi.org/10.1145/3038912.3052646.
[3] Hennig-Thurau, T., Wiertz, C., & Feldhaus, D. (2007). Exploring the impact of spoilers on consumer preferences. Journal of Consumer Research, 34(3), 403–415. https://doi.org/10.1086/519526.
[4] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186. https://doi.org/10.18653/v1/N19-1423.
[5] Brown, T. B., et al. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901. https://arxiv.org/abs/2005.14165.
[6] Raffel, C., et al. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21(140), 1–67. https://jmlr.org/papers/v21/20-074.html.
[7] Lewis, M., et al. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. Proceedings of ACL 2020, 7871–7880. https://doi.org/10.18653/v1/2020.acl-main.703.
[8] Kwiatkowski, T., et al. (2019). Natural Questions: A benchmark for question answering research. Transactions of the Association for Computational Linguistics, 7, 453–466. https://doi.org/10.1162/tacl_a_00276.
[9] Vaswani, A., et al. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30. https://arxiv.org/abs/1706.03762.
[10] Bommasani, R., et al. (2021). On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258. https://arxiv.org/abs/2108.07258

IJARETY

INNO SPACE
SJIF Scientific Journal Impact Factor

doi
crossref

निस्केयर
NISCAIR

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com