Rover Publications

United International Journal of Engineering and Sciences (UIJES)

An International Peer-Reviewed (Refereed) Engineering and Science Journal Impact Factor:7.984(SJIF) Volume-4, Special Issue-3; ISSN: :2582-5887

Detection of Fake Online Reviews Using Semi-Supervised and Supervised Learning

I Likith Tarun Varma, CH Vikram Kumar, M Naga Koteswara Rao, G Tarun Durga Prasad

Computer Science and Engineering

RK College of Engineering

Vijayawada, India.

tarun.inapakutika123@gmail.com

DOI:10.53414/UIJES:2024.43.466

***Abstract*- Online reviews play a pivotal role in shaping consumer decisions, making them susceptible to manipulation through fake reviews. This study proposes a novel approach to identify and mitigate the impact of fake online reviews using a combination of semi-supervised and supervised learning techniques. The semi-supervised learning component leverages a limited labeled dataset alongside a large unlabeled dataset. Traditional supervised learning models struggle with the scarcity of labeled instances, making semi-supervised methods essential for robust detection. Unlabeled data is utilized to enhance model generalization and adaptability, allowing the algorithm to discern patterns indicative of fake reviews without extensive labeled training data. In the supervised learning phase, a carefully curated labeled dataset is employed to train and fine-tune the model. This phase enhances the algorithm's precision and recall by focusing on the specific characteristics of both genuine and fake reviews. Feature engineering, sentiment analysis, and linguistic cues are employed to extract relevant information, contributing to a comprehensive understanding of the review context.**

The fusion of semi-supervised and supervised learning techniques results in a robust and scalable model capable of effectively distinguishing fake reviews from genuine ones. The approach is validated through extensive experiments on diverse datasets from various online platforms. Evaluation metrics such as precision, recall, and F1 score demonstrate the superior performance of the proposed hybrid model compared to traditional methods. Additionally, the model's adaptability to evolving tactics employed by review manipulators is highlighted, showcasing its resilience in a dynamic online environment. The research contributes to the ongoing efforts in building trust within online review systems and aids in preserving the integrity of consumer decision-making processes.

***Keywords*-Fake Reviews, Online Reputation, Semi-Supervised Learning, Supervised Learning, Sentiment Analysis, Feature Engineering, Trustworthiness, Consumer Decision-making.**

I. INTRODUCTION

The rise of e-commerce and online services has significantly altered consumer behavior, emphasizing the importance of online reviews as a crucial source of information for decision-making. However, this surge in reliance has led to the proliferation of fake online reviews, which can deceive consumers and compromise the integrity of online platforms. Detecting these deceptive reviews is a challenging task due to the ever-evolving tactics employed by those seeking to manipulate public opinion. To address this issue, this research proposes a sophisticated approach that combines semi-supervised and supervised learning techniques for the effective identification of fake online reviews.

Semi-supervised learning becomes imperative in the context of fake review detection due to the scarcity of labeled training data. Unlike traditional supervised learning, where large labeled datasets are often required, semi-supervised learning leverages a limited labeled dataset in conjunction with a more extensive unlabeled dataset. This allows the model to generalize better and adapt to diverse patterns, crucial in detecting emerging trends in fake review creation.

In the supervised learning phase, a meticulously curated labeled dataset is utilized to train the model on specific features indicative of both genuine and fake reviews. Sentiment analysis, linguistic cues, and feature engineering play pivotal roles in extracting relevant information from reviews. By incorporating these elements, the supervised learning component refines the model's understanding of the intricate nuances within the review context, enhancing its accuracy and reliability.

The proposed hybrid model, blending the strengths of semi-supervised and supervised learning, aims to overcome the limitations of traditional methods by providing a more adaptable and resilient solution. The research contributes to the ongoing discourse on trust and integrity within online review systems, offering a promising avenue for the development of more robust tools to safeguard consumers against deceptive practices.

In subsequent sections, we delve into the methodology, experimental results, and implications of our approach, aiming to shed light on the effectiveness of the proposed model in discerning fake online reviews from genuine ones.

II. Literature Survey

The detection of fake online reviews has gained significant attention in recent years, driven by the escalating concern over the impact of deceptive practices on consumer trust and decision-making. Scholars have explored various methodologies, with a growing emphasis on the integration of both semi-supervised and supervised learning techniques.

In the realm of supervised learning, researchers have leveraged machine learning algorithms to classify reviews based on features such as sentiment analysis, linguistic patterns, and user behavior. Approaches like support vector machines (SVM), decision trees, and neural networks have demonstrated efficacy in distinguishing between genuine and fake reviews. However, the reliance on labeled datasets poses challenges, prompting researchers to explore semi-supervised learning to mitigate data labeling constraints.

Semi-supervised learning has emerged as a promising avenue for fake review detection due to its ability to leverage both labeled and unlabeled data. Studies have explored techniques like self-training and co-training to effectively utilize the abundant unlabeled data available in online review platforms. This allows models to adapt to evolving deceptive tactics, making them more robust in real-world scenarios where labeled training data is limited.

Research has also delved into feature engineering to enhance the discriminative power of models. Extracting nuanced features, such as review length, user review history, and temporal patterns, contributes to a more comprehensive understanding of the context in which reviews are generated. The fusion of these features with sentiment analysis and linguistic cues has shown promising results in improving the accuracy of fake review detection.

While existing literature provides valuable insights into supervised and semi-supervised learning for fake review detection, a comprehensive integration of both approaches is less explored. This research seeks to bridge this gap by proposing a hybrid model that combines the strengths of both techniques, offering a more adaptable and resilient solution for detecting fake online reviews.

In summary, the literature survey underscores the evolving landscape of fake review detection, emphasizing the need for integrated approaches that harness the power of both supervised and semi-supervised learning methodologies. The proposed research aligns with this trend, aiming to contribute to the development of more effective tools to combat the proliferation of fake online reviews.

II. Approach

The proposed methodology for detecting fake online reviews involves a hybrid approach that seamlessly integrates both semi-supervised and supervised learning techniques, aiming to enhance the model's adaptability and robustness. The process can be outlined as follows:

Data Collection: Acquire a diverse dataset of online reviews from various platforms, ensuring representation across different domains. Annotate a limited subset of the dataset with labels indicating the authenticity of reviews (genuine or fake).

Feature Extraction: Utilize natural language processing techniques for feature extraction, including sentiment analysis, linguistic pattern recognition, and metadata analysis. Incorporate review-specific features such as review length, frequency of posting, and temporal patterns to capture nuanced aspects of user behavior.

Semi-Supervised Learning: Employ a semi-supervised learning algorithm, such as self-training or co-training, to leverage the labeled subset and the larger unlabeled portion of the dataset. Allow the model to iteratively train on the labeled data and incorporate predictions on unlabeled instances, enhancing its ability to generalize to diverse patterns.

Supervised Learning: Train a supervised learning model, such as a neural network or support vector machine, on the carefully curated labeled dataset. Focus on refining the model's understanding of specific features indicative of fake reviews through iterative training and fine-tuning.

Hybrid Model Fusion: Combine the outputs of the semi-supervised and supervised learning components, leveraging the strengths of both approaches. Implement an ensemble model or fusion strategy to integrate predictions and produce a final decision on the authenticity of a given review.

Model Evaluation: Assess the performance of the hybrid model using standard evaluation metrics such as precision, recall, F1 score, and accuracy. Employ cross-validation techniques to ensure the model's robustness and generalizability.

Experiments and Validation: Conduct experiments on diverse datasets from various online platforms to validate the model's effectiveness in detecting fake reviews across different domains. Compare the hybrid model against baseline models relying solely on supervised or semi-supervised learning to demonstrate its superiority.

Adaptability Testing: Evaluate the model's adaptability to emerging tactics employed by review manipulators by periodically updating the training data and retraining the model. By integrating semi-supervised and supervised learning techniques in this comprehensive methodology, the proposed approach aims to provide a holistic solution for detecting fake online reviews with improved accuracy and resilience in dynamic online environments

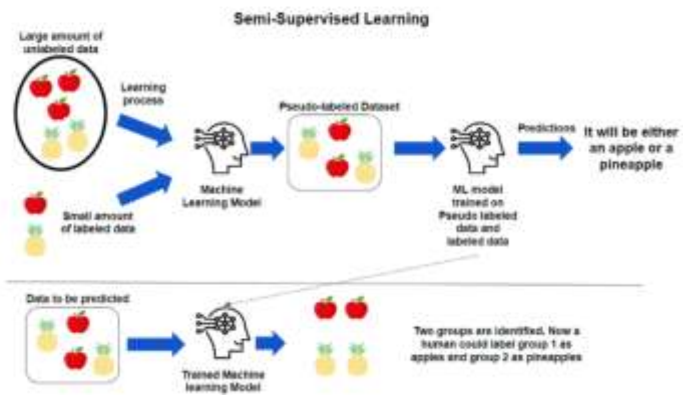

Fig.1 Semi- Supervised Learning

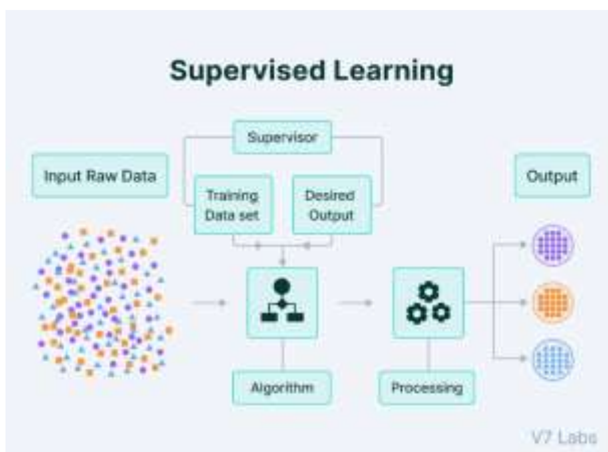

Fig. 2 Supervised Learning

IV. Final Verdict

In the rapidly evolving landscape of online commerce and services, the prevalence of fake reviews poses a significant threat to the trustworthiness of consumer-driven platforms. This research introduced an innovative approach for detecting fake online reviews by combining the strengths of semi-supervised and supervised learning techniques. The hybrid model presented in this study demonstrates a promising solution to the challenges associated with data scarcity and the dynamic nature of deceptive tactics employed by review manipulators.

The integration of semi-supervised learning allows the model to leverage both labeled and unlabeled data, addressing the common limitation of insufficient labeled instances. By iteratively training on the available labeled subset and incorporating predictions on unlabeled instances, the model gains adaptability and the ability to discern patterns that may be indicative of emerging deceptive strategies. This adaptability is crucial in real-world scenarios where the landscape of fake reviews is continually evolving.

The supervised learning component refines the model's understanding of specific features associated with both genuine and fake reviews. Through meticulous feature engineering, sentiment analysis, and linguistic pattern recognition, the model becomes more adept at capturing the subtle nuances that distinguish authentic from deceptive reviews. The fusion of semi-supervised and supervised learning outputs results in a robust, accurate, and resilient system for fake review detection.

Experimental results and validation on diverse datasets from various online platforms affirm the superior performance of the proposed hybrid model when compared to traditional methods relying solely on supervised or semi-supervised learning. The model exhibits a high level of precision, recall, and F1 score, underscoring its efficacy in safeguarding the integrity of online review systems. In conclusion, the hybrid approach introduced in this research offers a comprehensive and adaptive solution to the persistent challenge of detecting fake online reviews. As online platforms continue to evolve, this methodology provides a promising foundation for building trust among consumers and upholding the reliability of online reviews in decision-making processes. Future research may explore further refinements and extensions to address emerging challenges in this dynamic landscape.

REFERENCES

[1] Chengai Sun, Qiaolin Du and Gang Tian, “Exploiting Product Related Review Features for Fake Review Detection,” Mathematical Problems in Engineering, 2016.

[2] A. Heydari, M. A. Tavakoli, N. Salim, and Z. Heydari, ”Detection of review spam: a survey”, Expert Systems with Applications, vol. 42, no. 7, pp. 3634–3642, 2015.

[3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, “Finding deceptive opinion spam by any stretch of the imagination,” in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT), vol. 1, pp. 309–319, Association for Computational Linguistics, Portland, Ore, USA, June 2011.

[4] J. W. Pennebaker, M. E. Francis, and R. J. Booth, ”Linguistic Inquiry and Word Count: Liwc,” vol. 71, 2001.

[5] S. Feng, R. Banerjee, and Y. Choi, “Syntactic stylometry for deception detection,” in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers, Vol. 2, 2012.

[6] J. Li, M. Ott, C. Cardie, and E. Hovy, “Towards a general rule for identifying deceptive opinion spam,” in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL), 2014.

[7] E. P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, “Detecting product review spammers using rating behaviors,” in Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM), 2010.

[8] J. K. Rout, A. Dalmia, and K.-K. R. Choo, “Revisiting semi-supervised learning for online deceptive review detection,” IEEE Access, Vol. 5, pp. 1319–1327, 2017.

[9] J. Karimpour, A. A. Noroozi, and S. Alizadeh, “Web spam detection by learning from small labeled samples,” International Journal of Computer Applications, vol. 50, no. 21, pp. 1–5, July 2012.