

# Chapter 13

## Ethical Considerations in Using Fuzzy Artificial Intelligence for Detecting Fake Reviews

**A. Firos**

 <https://orcid.org/0000-0003-4207-713X>

*Rajiv Gandhi University, India*

**Seema Khanum**

 <https://orcid.org/0000-0002-2933-2717>

*Indian Computer Emergency Response Team, MeitY, Electronics Niketan, India*

### **ABSTRACT**

*This chapter examines Fuzzy Artificial Intelligence (FAI) as a solution for detecting fake reviews, a growing concern in digital marketplaces. FAI combines fuzzy logic with artificial intelligence to assess the authenticity of reviews by analyzing linguistic variables and producing a desirability score that indicates the likelihood of a review being genuine. Unlike traditional models, FAI handles ambiguity, improving detection accuracy. Results show that FAI outperforms conventional methods, offering deeper insights into review authenticity. The chapter highlights FAI's role in enhancing online trust, protecting consumers, and ensuring reliable decision-making. With its ability to adapt to new data, FAI is crucial for maintaining the integrity of online marketplaces and creating a trustworthy digital environment.*

DOI: 10.4018/979-8-3693-4147-6.ch013

Copyright © 2025, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

# 1 INTRODUCTION

## 1.1 Generative Models for Assessing Relative Desirability for Fake Review Detection

Generative models have emerged as a powerful tool in the realm of artificial intelligence, particularly in the detection of fake reviews (Hu et al., 2021). These models, which can generate new data instances that resemble the training data, are being increasingly applied to assess the relative desirability of online reviews, distinguishing between genuine and fabricated content. This capability is crucial in maintaining the integrity of online platforms where user-generated reviews significantly influence consumer behavior.

The application of generative models in fake review detection centers around their ability to learn the complex patterns and nuances of genuine reviews. By training on a dataset composed of authentic reviews, these models can grasp the subtleties of human language, including sentiment, syntax, and style, which are often challenging for fake reviewers to replicate accurately. This understanding allows generative models to generate synthetic reviews that serve as a benchmark for assessing the authenticity of new reviews (Hu et al., 2022).

One of the key strengths of generative models in this context is their ability to deal with the ambiguity and variability inherent in human-generated content. Traditional binary classification models often struggle with the fuzzy boundaries between genuine and fake reviews, as the latter become increasingly sophisticated. Generative models, however, can navigate this complexity by evaluating the likelihood of a review being real based on its similarity to the learned patterns of authentic reviews.

Another significant advantage is the continuous learning capability of generative models. As they are exposed to more data over time, including newly identified fake reviews, they refine their understanding of what constitutes a genuine review. This dynamic learning process ensures that the models remain effective even as fake review tactics evolve, maintaining a high level of accuracy in the detection process.

The implementation of generative models also brings challenges, such as the need for substantial and diverse training data to avoid biases and ensure the model's generalizability (Bouramdane, 2023). Additionally, the computational complexity of training generative models requires significant resources, which might be a barrier for some applications. Despite these challenges, the potential benefits in terms of improving the reliability of online review ecosystems are considerable.

In conclusion, generative models represent a promising approach to enhancing the detection of fake reviews by assessing the relative desirability of user-generated content. Their ability to learn and mimic the complexity of genuine reviews, adapt to new patterns of deception, and navigate the ambiguity of human language positions

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/ethical-considerations-in-using-fuzzy-artificial-intelligence-for-detecting-fake-reviews/359648](http://www.igi-global.com/chapter/ethical-considerations-in-using-fuzzy-artificial-intelligence-for-detecting-fake-reviews/359648)

## Related Content

---

### Genetic Privacy: A European Design or Default?

Elsa Suptotand Margo Bernelin (2019). *Cyber Law, Privacy, and Security: Concepts, Methodologies, Tools, and Applications* (pp. 713-730).

[www.irma-international.org/chapter/genetic-privacy/228752](http://www.irma-international.org/chapter/genetic-privacy/228752)

### Unveiling Security Vulnerabilities in Generative AI

Geeta Sharma, Pooja Chopraand Souravdeep Singh (2025). *Generative Artificial Intelligence and Ethics: Standards, Guidelines, and Best Practices* (pp. 267-286).

[www.irma-international.org/chapter/unveiling-security-vulnerabilities-in-generative-ai/358935](http://www.irma-international.org/chapter/unveiling-security-vulnerabilities-in-generative-ai/358935)

### Privacy and Security: Safeguarding Personal Data in the AI Era

Geeta Sandeep Nadella, Hari Gonaygunta, Mohan Harishand Pawan Whig (2025). *Ethical Dimensions of AI Development* (pp. 157-174).

[www.irma-international.org/chapter/privacy-and-security/359642](http://www.irma-international.org/chapter/privacy-and-security/359642)

### A Study of Cyber Crime and Perpetration of Cyber Crime in India

Saurabh Mittaland Ashu Singh (2019). *Cyber Law, Privacy, and Security: Concepts, Methodologies, Tools, and Applications* (pp. 1080-1096).

[www.irma-international.org/chapter/a-study-of-cyber-crime-and-perpetration-of-cyber-crime-in-india/228769](http://www.irma-international.org/chapter/a-study-of-cyber-crime-and-perpetration-of-cyber-crime-in-india/228769)

### Using Crowd Sourcing to Analyze Consumers' Response to Privacy Policies of Online Social Network and Financial Institutions at Micro Level

Shaikha Alduaij, Zhiyuan Chenand Aryya Gangopadhyay (2019). *Cyber Law, Privacy, and Security: Concepts, Methodologies, Tools, and Applications* (pp. 492-516).

[www.irma-international.org/chapter/using-crowd-sourcing-to-analyze-consumers-response-to-privacy-policies-of-online-social-network-and-financial-institutions-at-micro-level/228742](http://www.irma-international.org/chapter/using-crowd-sourcing-to-analyze-consumers-response-to-privacy-policies-of-online-social-network-and-financial-institutions-at-micro-level/228742)