


Chapter 10

A Framework for Detection of Hate Speech in Social Media: Achieving Societal Resilience

Nitish Kumar Ojha

 <https://orcid.org/0000-0002-2236-0766>

Amity University, Noida, India

Abhishek Vaish

Indian Institute of Information Technology, Allahabad, India

Shreyansh Dwivedi

Indian Institute of Information Technology, Allahabad, India

Brijesh Kumar Markandey

Indian Institute of Information Technology, Allahabad, India

Rania El-Gazzar

University of South-Eastern Norway, Norway

ABSTRACT

The authenticity and the orientation of information in social media require designing a complex information system to analyze the artifacts that may cause a positive or negative impact on the peace order of the global societal structure. Social cyber security is a new paradigm that is gaining momentum worldwide; every country is investing in capacity-building measures, R&D, and strategies to combat the menace of the problem caused by the abundance of manipulative information. Societal cyber security has become a multi-million-dollar market because economic change

DOI: 10.4018/979-8-3693-9904-0.ch010

has monetized data, and it is a top priority for every government to ensure that the digital society is secure and safe. Therefore, it is pertinent to find a framework that can help detect information that can negatively impact society at large. In this paper, we are using machine learning techniques for classification in a fusion dataset of manipulative information that has been concatenated with the benchmark dataset. The subclasses are real news without hate speech, real news with hate speech, fake news with hate speech, and fake news without hate speech. The application of this framework is to be used to mitigate the threat to society through the widespread use of manipulative information and help society to be resilient against such threats.

1 INTRODUCTION

Society is embracing the Digital infrastructure in its day-to-day activities with more positive outcomes. The dependence on information by individuals is increasing, this change has led to the shift to new emerging trends of cyber-attacks. Traditionally the sophistication of cyber-attacks was based on evolving techniques to target the technical systems causing loss and damage to ICT infrastructure but now the damages are more oriented towards damaging the social structure without compromising the technical system. The motive behind this shift could be linked with the abundance of information available, the ease of generating manipulative information & its widespread propagation, the potential of high-magnitude loss, and high valued stakeholders. In 2007, Estonia was at a standstill, fifty-eight websites were offline at once, an outrageous act that is linked with the disputed relocation of the Soviet-era Brownz Soldier monument. Though this has increased the preparedness level of Estonia but was seen as a whistle-blower to the international community to understand the ramifications of it and to bootstrap the efforts of preparing them for such innovative methods of invasive and non-invasive methods of technological destruction. Cyberspace has been used as a platform for crime against society, and the variety of such crimes is enormous from individual to group, from Organization to Government. Every stakeholder is a victim, and the impact is quite diverse. The recent trend of crime is the use of social media platforms as a launching pad against the Government in different forms e.g. destabilizing the law & order, Narratives against the government, using it to steal classified information about the nation, etc. Information in Cyberspace is witnessing exponential growth and monitoring the correctness of the information is practically not possible. The ease of publishing or uploading information in social media has not only led to the growth of the creation of digital content but has also led to the increase of manipulative content like fake news, content of hatred, and other disturbing content that harms society at large. It has been observed that one of the reasons for the successful operation of large-scale

20 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/a-framework-for-detection-of-hate-speech-in-social-media/371739

Related Content

Maximizing Impact: The Role of Influencer Marketing in Contemporary Tourism Management

Mohammad Badruddoza Talukder (2025). *Navigating Organizational Behavior in the Digital Age With AI* (pp. 277-304).

www.irma-international.org/chapter/maximizing-impact/364233

Intelligent Tutoring Systems for Cybersecurity Entrepreneurship: Cultivating AI-Augmented Risk Assessment Capabilities in Technology Startups

Abdelraouf Ishtaiwi (2026). *Enhancing Entrepreneurial Education and Training With AI* (pp. 81-116).

www.irma-international.org/chapter/intelligent-tutoring-systems-for-cybersecurity-entrepreneurship/409738

Attention-Driven Multi-Scale Clothing Detection Using an Enhanced SCS-YOLO Framework

Xuan Li (2025). *International Journal of Intelligent Information Technologies* (pp. 1-19).

www.irma-international.org/article/attention-driven-multi-scale-clothing-detection-using-an-enhanced-scs-yolo-framework/394108

Network Communication and Electronic Control Strategy of New Energy Vehicles Based on Cloud Platform in the IoT Environment

Yufeng Tang (2023). *International Journal of Ambient Computing and Intelligence* (pp. 1-15).

www.irma-international.org/article/network-communication-and-electronic-control-strategy-of-new-energy-vehicles-based-on-cloud-platform-in-the-iot-environment/318135

Transforming the Language Teaching Experience in the Age of AI: Ethical, Social, and Cultural Considerations in Implementing AI in Language Education

Reena Sudhakar Rao Satpute (2023). *Transforming the Language Teaching Experience in the Age of AI* (pp. 115-124).

www.irma-international.org/chapter/transforming-the-language-teaching-experience-in-the-age-of-ai/330379