MISQ Archivist

Finding People with Emotional Distress in Online Social Media: A Design Combining Machine Learning and Rule-Based Classification

Michael Chau, Tim M. H. Li, Paul W. C. Wong,
Jennifer J. Xu, Paul S. F. Yip, and Hsinchun Chen

Abstract

Many people face problems of emotional distress. Early detection of high-risk individuals is the key to prevent suicidal behavior. There is increasing evidence that the Internet and social media provide clues of people's emotional distress. In particular, some people leave messages showing emotional distress or even suicide notes on the Internet. Identifying emotionally distressed people and examining their posts on the Internet are important steps for health and social work professionals to provide assistance, but the process is very time-consuming and ineffective if conducted manually using standard search engines. Following the design science approach, we present the design of a system called KAREN, which identifies individuals who blog about their emotional distress in the Chinese language, using a combination of machine learning classification and rule-based classification with rules obtained from experts. A controlled experiment and a user study were conducted to evaluate system performance in searching and analyzing blogs written by people who might be emotionally distressed. The results show that the proposed system achieved better classification performance than the benchmark methods and that professionals perceived the system to be more useful and effective for identifying bloggers with emotional distress than benchmark approaches.

Keywords: Social media, emotional distress, suicide research, design science, classification