International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

ELSEVIER
SSRN

IJIEMR Transactions, online available on 5th Jan 2023. Link :http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=Issue 01

DOI: 10.48047/IJIEMR/V12/ISSUE 01/24

Title **Hybrid Technique for Opinion Mining in Cyber Space**

Volume 12, ISSUE 01, Pages: 246-255

Paper Authors

Navneet Kaur, Jaspreet Singh Budwal

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Hybrid Technique for Opinion Mining in Cyber Space

Navneet Kaur[1], Jaspreet Singh Budwal[2]

[1]Assistant Professor, Department of Computer Science and IT,
Lyallpur Khalsa College, Jalandhar, Punjab, India
Email:- saininavneet@gmail.com
[2]Computer Faculty, Department of Computer Science,
GSSS Hazara, Jalandhar, Punjab, India
Email:- js2100@gmail.com

Abstract
The opinion mining is widely used in the area of big data for grouping the opinions of people having different sentiments. It is also used to know the temperament of the people towards certain issue. Opinion mining has a significant importance due to overdependence on the internet and its services. Opinion mining is also called as sentiment analysis; it is a part of Natural Language Processing. It draws people's opinions, which includes assessments, attitudes, and emotions toward various topics and issues. This is a difficult task but extraordinarily beneficial for finding people's opinion. People and organizations are progressively hoping on belief for decision-making not just for policies but conjointly for day-after-day decisions, due to the rise of digital platforms in Internet adore social networking and journal websites. So, the bulk of the analysis centered on ways comparable to machine learning, lexicon-based approaches, Kansei approaches. This paper describes the hybrid approach for opinion mining in the field of cyber space, which is very effective and efficient as compared to the individual approaches.

Keywords: *Opinion Mining, Machine Learning, Lexicon based approach, Kansei Approach, Hybrid Approach, Cyber Security*

Introduction

In the recent years, significant research has been conducted into finding the sentiments of people in cyberspace mainly text-based information using opinion mining. To analyse and classify sentiments of people based on text; researchers used a variety of opinion mining techniques, like machine learning and a lexicon-based approach. Researchers found many drawbacks of these approaches in the case of opinion mining.

It gives an opportunity for other analysts to further explore new domain applications and propose new methods for use of opinion mining. This study elaborates various opinion mining techniques. Further use of Kansei approach with other approaches like machine learning, lexicon-based approaches is discussed. This will help in improving techniques for mining sentiment of people and emotions in cyberspace. Apart from this, study elaborates that how hybrid approach will help in improvement of cyber security. In present scenario cyber security is a major concern for the governments and as well as for common citizens.

In this paper Section II demonstrates and gives brief understanding of opinion mining and its various aspects. Section III discusses machine learning and Lexicon-based approaches, as well as Kansei techniques for opinion mining. Apart from this, basic detail of cyber security is included in section IV. Section V demonstrates a Hybrid Approach to Opinion Mining. Section VI explains the advantages of hybrid approach for Opinion Mining in the field of cyber security.

Opinion Mining

Opinions are always important and integral part of humans, because opinions describe and make significant impact on the conduct of human beings in the society. We as individuals always gives importance to the views and ideas of people who surrounds us. Their opinions influence our decisions in all aspects of life; it may be personal as well as professional decisions. Individuals and organizations often seek the opinion of friends and the general public regarding various issues and products [1]. Sentiment analysis is widely used in recent years. It has attained a cult status and attention in the field of research. This is mainly due to use of various social networking sites and other information-centric applications. In these sites the generated data is mainly in the form of textual in nature.

Opinion mining has been studied on a variety of topics such as movie reviews, game reviews, product reviews, mobile reviews, and more. Opinion mining is the computational study of people's opinions, emotions and attitudes towards a specific issue or topic, which can vary from individual to organizations, products or services. It is a technology for extracting opinions from unstructured text documents

that is closely related to data mining and uses machine learning techniques and lexicon-based approach to identify and analyses the sentiments expressed in a text document [2]. Political security, food security, economic security, human security, resource security, environmental security, border security, cyber security and health security etc. all are popular Opinion mining research topics [3]. Despite the fact that Opinion Mining has been extensively researched in the literature, new studies continue to emerge as available data grows and becomes more complex.

Techniques for Opinion Mining

Various methods have been applied by the researchers for the classification of opinions mainly comprises of text data. For classification algorithms, two methods are used namely supervised and unsupervised methods. One approach is the unsupervised called as lexicon based approach, in which the process is guided by rules and heuristics derived from linguistic knowledge. Another approach is called as supervised machine learning, where algorithms retrieve intrinsic information from previously labelled data to classify newer unlabelled data.

Polarity of the sentence is the main concern of the Sentiment classification approach. It expresses various sentiments like positive, negative, and neutral sentiment towards the topic. The by and large reason of text based opinion mining is to determine high-quality data and significant experiences from text, permitting businesses to form wise and accurate decisions. In general, the function of text mining is to obtain high-quality detail and information from text, allowing decision makers to take decisions on the basis of available results.

As a result, sentiment classification is often called as polarity determination. Artifacts reviews, news forums, newspaper articles and blogs all had various levels of their polarity determination factors.

A. Lexicon Based Approach

A sentiment lexicon approach is used to describe the polarity (positive, negative, and neutral) of textual content. In comparison to machine learning-based algorithms, this approach is more understandable and easier to implement. The disadvantage is that it necessitates the involvement of humans in the text analysis process. The lexicon-based approach is further subdivided into two types: dictionary-based (based on dictionary words, such as WordNet or other entries)

and corpus-based (based on corpus entries) (using corpus data, can further be divided into Statistical and Semantic approaches) [4][7].

B. Machine Learning

Machine learning algorithms are mainly of two types called as supervised and unsupervised learning methods. In the supervised learning, the training data comprises input vectors and it has corresponding target vectors, while in unsupervised learning, the training data mainly comprises of combination of input vectors but no corresponding target values.

Machine learning (ML), also known as deep learning, refers to processes that involve the creation and evolution of machine dictionaries that model human behaviour, thoughts, and responses. Data mining tasks includes supervised learning, unsupervised learning and sequential pattern mining [5][17].

For assessment arrangement, the machine learning strategy employs various grouping calculations. For survey content, grouping calculations such as Naive Bayes, SVM, and Decision Tree perform admirably. An alternative strategy such as asset-based conclusion mining has also been proposed by various analysts. It is based on assets such as SentiWordNet and WordNet estimations of word order. [1][6]

C. Hybrid Approaches

In order to perform opinion mining, researchers used a hybrid approach. When machine learning and lexicon-based approach is used individually, there are certain shortcomings in the case of opinion mining. If we combine two or more than two methods (hybrid approach), we can get better accuracy and results while extracting and classifying the opinion of the people. The shortcomings of machine learning and lexicon based approach can be compensated.

According to the reviewed research papers, the hybrid approach is often used by the majority of researchers for opinion mining of business products such as mobiles, cars, bikes and for other electronic products. Apart from said products it is also used in the field of education and medicine. In addition, we discovered that the Nave Bayes Classifier (NBC) and Support Vector Machine (SVM) are the most frequently used machine learning techniques in the hybrid approach. Apart from this the other most commonly used hybrid approach

comprises random forest, fuzzy rule-based system and deep learning.

D. Kansei Approach

The definition of Kansei Engineering as given in Nagamachi stated that Kansei Engineering is seen as "translating technology of a consumer's feeling and image for a product into design elements". Kansei Approach acts as an apparatus of semantic differential method usually applied to stimulate consumers' emotions or preferences towards a product or environment.

Kansei approach is capable of relating sensitivity, feeling and emotion. Khairul's study discussed the use of Kansei Engineering in the development of the Kansei Information Security Assessment (KISA). KISA helps in developing a framework for user emotional assessment in the cyber security domain. Kansei Engineering approach can helps in assessing human emotion factors in the designs and applications of cyber security policies for governments as well as for organizations. Various studies had used the Kansei Engineering to evaluate the concept of reliance in the information security domain by scrutinizing all kind of emotions especially positive and negative. These emotions can influence the concept of reliance. Another study was conducted focusing on user rage assessment for information security using the Kansei Engineering methodology. The artefacts were used as a specimen to assess the anger emotion in a chatbot. The study found that the Kansei approach could be used in other applications or research involving emotion assessment, such as in the information security domain. This lends support to the current study's use of Kansei for assessing political security-related emotions [8][9].

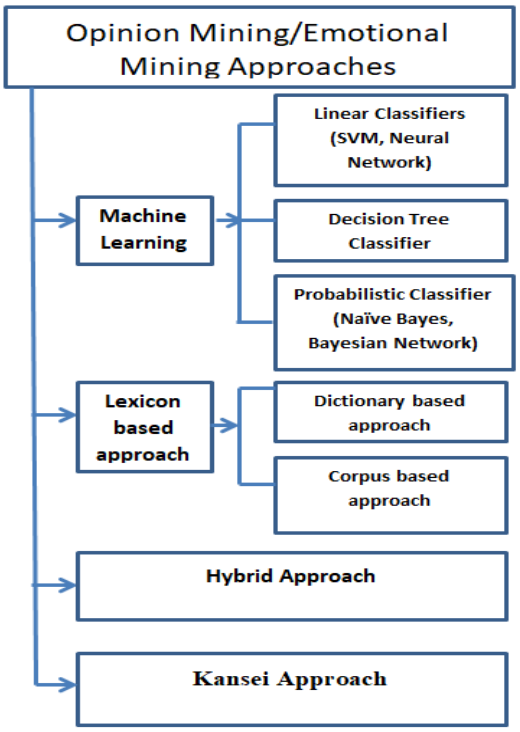

Figure 1: Opinion Mining Approaches

Cyber Security

Cyber security is a subject of concern for governments as well as for individuals. This concerns with the issues of diverse

cyber-attacks and devising defence strategies (i.e., counter measures) that preserve confidentiality, integrity and availability (CIA) of any digital and information technologies [10].

• Term Confidentiality is defined as "to prevent the disclosure of information to unauthorized persons or systems".

•Term Integrity is defined as "to prevent unauthorized modification/deletion of information".

• Term Availability is defined as "to ensure that the systems responsible for providing, storing, and processing information are accessible when needed by the user and for those who needs it".

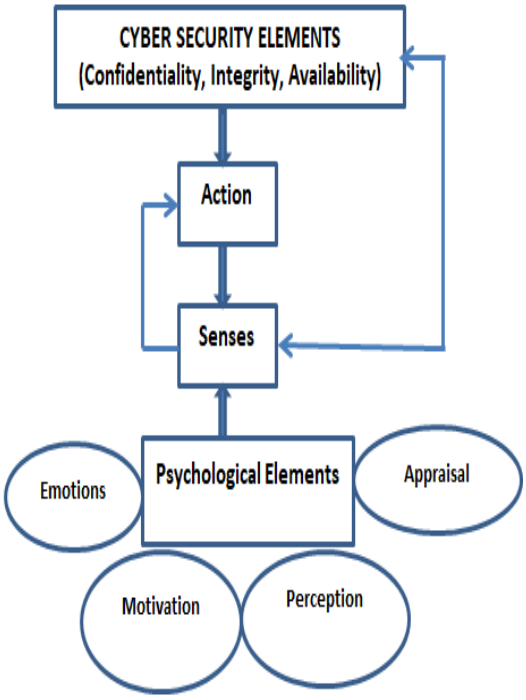

Figure 2: Relationship between Cyber Security Elements (CIA) and Psychological Elements

Hybrid Approach for Cyber Security

Opinion mining is a form of data mining that helps in extracting information from a large number of texts and data available on social media platforms about various subjects and topics, related to people's opinions and feelings. Several conclusions can be drawn from the reviewed articles. Opinion mining is done using techniques like machine learning, lexicon-based method and the Kansei approach are three different approaches, combined together are called as hybrid approach[16]. Many disadvantages and difficulties have been identified in numerous studies, in terms of opinion mining approaches, like the absence of distinct emotions in opinion mining, Machine learning techniques and lexicon-based approaches are being studied for their efficacy [11][12].

This study recommends using the Kansei approach as a hybrid strategy, which can be integrated with machine learning techniques and lexicon-based approaches. Despite the lack of a polarity score, this method can be overcome by integrating the Kansei methodology with machine learning techniques and a lexicon-based approach for dictionary establishment in the field of cyber security. The lexicon-based approach and machine learning

technique will aid in calculating the text polarity score and improving the correctness of the opinion outcome in case of cyber security [13].

As a result, this study introduces a new domain: opinion mining in cyber security utilizing hybrid technique.

According to the studies, when machine learning, lexicon-based approaches and the Kansei approach have been combined together, approach proved its potential to extract emotions of the people in opinion mining. However, because of the strong reliance on the existing library for scoring the views and emotions of words, the lack of domain-specific emotion words is a restriction for extracting emotions from text [14][15]. In case of cyber security various Kansei words such as Trouble, Insecurity, Anxious, Disturbed, Suspicious, Unsure, Sceptical, Denial, Astonished, Doubt, Concern, Frustrated, Appalling, Rage, Spite, Gloom, Disapproving, Furious, Nervousness, Confuse, can be used for finding the emotions and sentiments.

Accordingly, the Kansei approach can be used for development of a dictionary of domain-specific words on opinion mining related to the cyber security. This study recommends a hybrid approach for future research that combines the machine learning method, the lexicon-based approach with the Kansei approach to judge the opinion and emotions of the people specially related to cyber security. The Kansei approach can be helpful in resolving the emotions of the people in the cyber security domain, whereas machine learning and lexicon-based approaches can classify and predict people's opinions. This hybrid approach will help people belonged to various fields such as researchers, businesses, and governments to use the method to observe sentiments as well as emotions at the same time for cyber security purposes. The anticipated result of this combination would be an assessment of opinions and feelings of people, with the incorporation of the extremity score based on the cyber security components.

Table 1: Kansei Words

Kansei Words for Cyber Security		
Trouble	Insecurity	Anxious,
Disturbed	Astonished	Suspicious
Unsure	Sceptical	Confuse
Denial	Doubt	Concern
Frustrated	Rage	Spite
Gloom	Disapproving	Furious

Advantages of Hybrid Approach in Cyber Security

Various cyberspace activities, like fake news, spreading misinformation and hatred speech can leads to cyber security threats.

These sorts of cyber dangers can posture noteworthy dangers to national security. People included in such exercises may end up backstabbers incidentally since each internet client encompasses a unique persona, conclusion, religion, and feeling. They can accept these wrong bits of gossip readily or unwillingly and proceed to support and spread them to others. Human emotions and behaviours of this nature can influence cyberspace As a result, emotion is regarded as a critical mechanism for detecting threats to national defense.

Since the web encompasses a candidly well-off space where people can express their views, sentiments, assumptions, so due to this, connect between feelings, emotions and despise discourse in the internet is often evident. Related investigate on feeling within the field of national security found that fear and outrage impact legislative issues; this is one major viewpoint of national security. The relationship between feeling and national security components can be seen in how people respond to natural security issues. The application of opinion mining using hybrid approach in the domain of cyber security is extremely beneficial. The reason for this is presence of large amount of data in the form of text in online framework. It is easy and comfortable for individuals in cyberspace world to describe emotions through text and words, instead of other modes of expression. So, a large amount of text based data can be generated for finding the actual opinion of the people. The state of cyber security can be improved by using opinion mining to detect threats in cyberspace.

Conclusion

Opinion mining is an effective mechanism for defining sentiments and emotions based on the text in the case of cyber security. There are several disadvantages of opinion mining, but when we perform opinion mining using hybrid approach which comprises Machine Learning, Lexicon Based and Kansei Approach, the results will be far better and accurate.

According to the findings of this study, a combination of the machine learning method, the lexicon-based approach, and the Kansei approach will be a more viable mechanism in the field of cyber security. Furthermore, in order to expand the

research findings, a detailed analysis, including a thorough assessment and evaluation of hybrid approach is required. The results can also be used to create a cyber security dictionary that can be used for opinion mining in future.

References

[1] R.Akila, S.Revathi, "Opinion Mining on Food Services using Topic Modeling and Machine Learning Algorithms", 2020 6th International Conference on Advanced Computing & Communication Systems (ICACCS), pp. 1071-1076.

[2] Mwana Said Omar, Alexander Njeru, Samiullah Paracha,Muhammad Wannous, Sun Yi, "Mining Tweets for Education Reform", May 2017. DOI: 10.1109/ICASI.2017.7988441.

[3] Razali, N.A.M., Malizan, N.A., Hasbullah, N.A. et al. "Opinion mining for national security: techniques, domain applications, challenges and research opportunities", Journal of Big Data 8, 150 (2021). https://doi.org/10.1186/s40537-021-00536-5

[4] Azeema Sadia, Fariha Khan and Fatima Bashir, "An Overview of Lexicon-Based Approach For Sentiment Analysis", 2018 3rd International Electrical Engineering Conference (IEEC 2018)

[5] Haji Binali, Vidyasagar Potdar, Chen Wu, "A State Of The Art Opinion Mining and Its Application Domains", 2009 IEEE International Conference on Industrial Technology. DOI:10.1109/ICIT.2009.4939640.

[6] Ayça Deniz, Merih Angin, Pelin Angin, "Evolutionary Multiobjective Feature Selection for Sentiment Analysis", IEEE Access, vol. 9, pp. 142982-142996, 2021.

[7] Kumar Ravi a,b , Vadlamani Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications", Knowledge-Based Systems 89 (2015), Elsevier pp. 14–46.

[8] Noor Afiza Mat Razali, Nur Atiqah Malizan, Nor Asiakin Hasbullah, Norul Zahrah Mohd Zainuddin, Normaizeerah Mohd Noor, "Assessment of Emotion in Online News based on Kansei Approach for National Security", IJACSA, vol. 12, no. 7, pp. 361-371, 2021.

[9] Shuji Hashimoto, "Kansei Technology and Robotics", Kansei Engineering International Vol 8 No.1 pp. 11-14(2009).

[10] Julian Jang-Jaccard, Surya Nepal, "A survey of emerging threats in cybersecurity", Journal of Computer and System Sciences 80 (2014) pp. 973–993.

[11] Atsuhiro Yamada, Sho Hashimoto and Noriko Nagata, "A Text Mining Approach for Automatic Modeling of Kansei Evaluation from Review Texts", March 2018 Advances in Intelligent Systems and Computing. DOI:10.1007/978-981-10-8612-0_34

[12] A. Hadiana, A. A. Wahid, A. Sofyan, D. Hirawan, K. Patalia, and M. Z. Faruqi, "The Study of Information Support System for Decision Making based on Kansei Engineering", European Journal of Engineering and Technology Research, vol 6 Issue 5 pp. 94-98. DOI:http://dx.doi.org/10.24018/ejers.2021.6.5.2533

[13] Simon T. W. Schütte, Jorgen Eklund, Jan R.C. Axelsson and Mitsuo Nagamachi, "Concepts, Methods and Tools in Kansei Engineering", Theoretical Issues in Ergonomics Science pp 214-231 May 2004.

[14] Hsiao YH, Chen M-C, "Kansei Engineering with Online Content Mining for Cross-Border Logistics Service Design", 2016 https://doi.org/10.1109/iiai-aai.2016.12

[15] Hsiao Y-H, Chen M-C, Lin M-K, "Kansei Engineering with Online Review Mining for Hotel Service Development", 2017. https://doi.org/10.1109/iiai-aai.2017.12.

[16] Rosli RM, Lokman AM, Aris SRS, "Analysis of evoked emotions in extremist YouTube videos through Kansei evaluation". International Conference on Kansei Engineering & Emotion Research 2018:;740–747. Https://doi.org/10.1007/978-981-10-8612-0_77.

[17] Li Z, Tian ZG, Wang JW, Wang WM, "Extraction of affective responses from customer reviews: an opinion mining and machine learning approach", International Journal of Computer Integrated Manufacturing. 2020;33(7):670–685. https://doi.org/10.1080/0951192X.2019.1571240