International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 05
April 2017

SOCIAL NETWORKING SITE DETECTING CYBERBULLYING WORDS

D. Sindhoori[1], E. Ramya Sree[2], G. Sharath Reddy[3],Dr. N. Venkatadri [4]
[1] B.Tech C.S.E TKREC Hyderabad Email: dhannarapusindhoori@gmail.com
[2] B.Tech C.S.E TKREC Hyderabad Email: ramyasree95errabelli@gmail.com
[3] B.Tech C.S.E TKREC Hyderabad Email: sharathreddy930@gmail.com
[4] Professor ,CSE Dept, TKREC, Hyderabad, TS-India, Email: nagala.venkat@gmail.com

1. ABSTRACT

In this present world we have come across many problems due to cyberbullying in social media which has become a major issue, especially in childrens, adolescents and young adults. In our project we have developed a new learning method to tackle this critical problem. We have developed techniques that make automatic detection of bullying messages in social media possible, and this could help to construct a healthy and safe social media environment. Our proposed method is able to exploit the hidden feature structure of bullying information and learn a robust and discriminative representation of text.

2.INTRODUCTION

Social Media is a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content. Via social media, people can enjoy enormous information, convenient communication experience and so on. However, social media may have some side effects such as cyberbullying, which may have negative impacts on the life of people, especially children and teenagers. Cyberbullying victimization rate ranges from 10% to 40%. In the United States, approximately 43% of teenagers were ever bullied on social media. The same as traditional bullying, cyberbullying has negative, insidious and sweeping impacts on children. The outcomes for victims under cyberbullying may even be tragic such as the occurrence of self-injurious behaviour or suicides. One way to address the cyberbullying problem is to automatically detect and promptly report bullying messages so that proper measures can be taken to prevent possible tragedies. Cyberbullying detection can be formulated as a supervised learning problem. A classifier is first trained on a cyberbullying corpus labeled by humans, and the learned classifier is then used to recognize a bullying message. Three kinds of information including text, user demography, and social network features are often used in cyberbullying detection . Since the text content is the most reliable,

our work here focuses on text-based cyberbullying detection.In the text-based cyberbullying detection, the first and also critical step is the numerical representation learning for text messages. In fact, representation learning of text is extensively studied in text mining, information retrieval and natural language processing (NLP). Bag-of-words (BoW) model is one commonly used model that each dimension corresponds to a term.

3.Proposed work

Three kinds of information including text, user demography, and social network features are often used in cyberbullying detection. Since the text content is the most reliable, our work here focuses on text-based cyberbullying detection.In this paper, we investigate one deep learning method named stacked denoising autoencoder (SDA). SDA stacks several denoising autoencoders and concatenates the output of each layer as the learned representation. Each denoising autoencoder in SDA is trained to recover the input data from a corrupted version of it. The input is corrupted by randomly setting some of the input to zero, which is called dropout noise. This denoising process helps the autoencoders to learn robust representation. In addition, each autoencoder layer is intended to learn an increasingly abstract representation of the input. In this paper, we develop a new text representation model based on a variant of SDA: marginalized stacked denoising autoencoders (mSDA), which adopts linear instead of nonlinear projection to accelerate training and marginalizes infinite noise distribution in order to learn more robust representations. We utilize semantic information to expand mSDA and develop Semantic-enhanced Marginalized Stacked Denoising Autoencoders (smSDA). The semantic information consists of bullying words. An automatic extraction of bullying words based on word embeddings is proposed so that the involved human labor can be reduced. During training of smSDA, we attempt to reconstruct bullying features from other normal words by discovering the latent structure, i.e. correlation, between bullying and normal words. The intuition behind this idea is that some bullying messages do not contain bullying words. The correlation information discovered by smSDA helps to reconstruct bullying features from normal words, and this in turn facilitates detection of bullying messages without containing bullying words.

4.System architecture

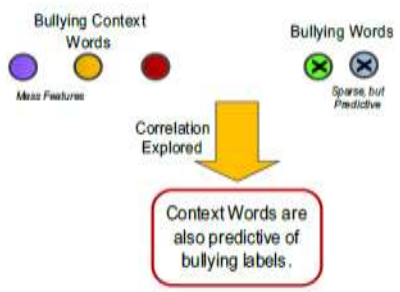

Fig System Architecture

The user and admin can use the software. The operations which are done are stored in the database. Admin will block the users who are indulged in using cyberbullying words.

5. Algorithm

Step 1: User gets registered initially

Step 2: User will then login if the entered details are right goto step4, if not goto step3

Step 3: Enter correct Email Id and Password

Step 4: User Homepage will be opened where user can send request,view request,update profile,view timeline images and send messages

Step 5: Admin logs in with his user name and password.

Step 6: Admin views his home page

Step 7: Admin will initially add all the cyberbullying words which will be compared with user messages.

Step 8: If any of the user will use cyberbullying words while sending messages, admin can view the list of people indulged in using cyberbullying words and can block them.

6.Result analysis

The role of user is 50% as he can send and receive messages and can post and view the images.

The role of admin is 50% as he can view the malicious users and can block them if the users use any of the cyberbullying words which are in database.

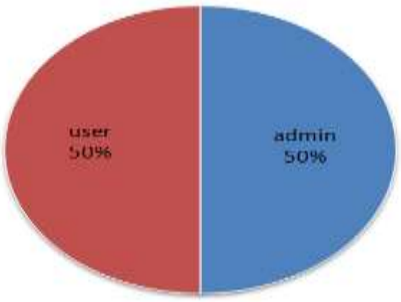

Fig: Result Analysis

7.Conclusion

In, This paper addresses the text-based cyberbullying detection problem, where robust and discriminative representations of messages are critical for an effective detection system. By designing semantic dropout noise and enforcing sparsity,we have developed semantic-enhanced marginalized denoising autoencoder as a specialized representation learning model for cyberbullying detection. In addition, word embeddings have been used to automatically expand and refine bullying word lists that is initialized by domain knowledge. The performance of our approaches has been experimentally verified through two cyberbullying corpora from social medias: Twitter and MySpace. As a

next step we are planning to further improve the robustness of the learned representation by considering word order in messages.

8. Future Enhancement

Our project, 'SOCIAL NETWORKING SITE DETECTING CYBERBULLYING WORDS' can be further enhanced. In our project we have used certain techniques to detect the cyberbullying words in order to reduce the cyber crimes in this present world. We can even include group chats which would be an added advantage to our project.But it would be more benifited if we can detect the cyber words even from an image as there is a chance to post an image which contains cyberbullying words and we can make it more secure.

References

1. R. M. Kowalski, G. W. Giumetti, A. N. Schroeder, and M. R.Lattanner, "Bullying in the digital age: A critical review and metaanalysis of cyberbullying research among youth." 2014.

2. Q. Huang, V. K. Singh, and P. K. Atrey, "Cyber bullying detection using social and textual analysis," in Proceedings of the 3rd International Workshop on Socially-Aware Multimedia. ACM, 2014, pp.3–6.

3. M. Dadvar, D. Trieschnigg, R. Ordelman, and F. de Jong, "Improving cyberbullying detection with user context," in Advances in Information Retrieval. Springer, 2013, pp. 693–696.

4. Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 35, no. 8, pp. 1798–1828, 2013.

5. J.-M. Xu, K.-S. Jun, X. Zhu, and A. Bellmore, "Learning from bullying traces in social media," in Proceedings of the 2012 conferenceof the North American chapter of the association for computational linguistics: Human language technologies. Association for Computational Linguistics, 2012, pp. 656–666.

6. V. Nahar, X. Li, and C. Pang, "An effective approach for cyberbullying detection," Communications in Information Science and Management Engineering, 2012.

7. M. Dadvar, F. de Jong, R. Ordelman, and R. Trieschnigg, "Improved cyberbullying detection using gender information," in Proceedings of the 12th -Dutch-Belgian Information Retrieval Workshop (DIR2012). Ghent, Belgium: ACM, 2012.

8. P. Baldi, "Autoencoders, unsupervised learning, and deep architectures," Unsupervised and Transfer Learning Challenges in Machine Learning, Volume 7, p. 43, 2012.

9. M. Chen, Z. Xu, K. Weinberger, and F. Sha, "Marginalized denoising autoencoders for domain adaptation," arXiv preprint arXiv: 1206.4683, 2012.

10. K. Dinakar, R. Reichart, and H. Lieberman, "Modeling the detection of textual cyberbullying." in The Social Mobile Web, 2011.

11. A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of social media," Business horizons, vol. 53, no. 1, pp. 59–68, 2010.

12. B. K. Biggs, J. M. Nelson, and M. L. Sampilo, "Peer relations in the anxiety–depression link: Test of a mediation model," Anxiety, Stress, & Coping, vol. 23, no. 4, pp. 431–447, 2010.

13. D. Yin, Z. Xue, L. Hong, B. D. Davison, A. Kontostathis, and L. Edwards, "Detection of harassment on web 2.0," Proceedings of the Content Analysis in the WEB, vol. 2, pp. 1–7, 2009.