IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

Detecting Cyber-Bullying With Denoising Auto Encoder

Devi Venkata Sai Chandra Prasad.Morla[1], Sujith Mudraboyina[2], Sri Hari Babu Gole[3]

Student, Department of Computer Science, KL University, Vijayawada, India[1, 2]

Professor, Department of Computer Science, KL University, Vijayawada, India[3]

Abstract: Cyber bullying has emerged as a serious problem among the various effects of Social Media. It is afflicting children, adolescents and young adults. Machine learning techniques make automatic detection of bullying messages in social media is possible, and this could help to construct a healthy and safe social media environment. Robust and discriminative numerical representation learning of text messages is the critical issue in the research area. In this paper, we propose a new representation of learning method to tackle this problem. Our method is developed via semantic extension of the popular deep learning model stacked denoising auto encoder which is named as Semantic-Enhanced Marginalized Denoising Auto-Encoder (smSDA). The semantic extension consists of semantic dropout noise and sparsely constraints. Where the semantic dropout noise is designed based on domain knowledge and the word embedding technique. **Our proposed method is able to exploit the hidden feature i.e., structure of bullying information and learn a robust and discriminative representation of text.**

Keywords: Social Media, Bullying, Machine Learning, Safe Environment.

I. INTRODUCTION

Computer security (Also known as cyber security or IT Security) is information security which is applied to computers and networks. The field covers all the processes and mechanisms by which computer- grounded outfit, information and services are defended from unintended or unauthorized access, change or destruction. Protection from unplanned events and natural disasters is also included in the Computer Security. Else, in the computer assiduity, the term security-- or the expression computer security-- refers to ways for icing that data stored in a computer can not be read or compromised by any individuality without authorization. Utmost computer security measures involve data encryption and watchwords. Data encryption is the restatement of data into a form that's ungraspable without a decoding medium. A password is a secret word or phrase that restricts a user access to a particular program or system.

Fig. 1 Secure System

A. Working conditions and basic needs in the secure computing:

If you don't take basic steps to protect your work computer, you put it and all the information on it at risk.

- Physical security
- Access passwords
- Prying eye protection
- Anti-virus software

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

- Firewalls
- Software updates
- Keep secure backups
- Report problems

II. LITERATURE SURVEY

A. PAPER 1

Representation Learning: A Review and New Perspectives

AUTHORS: Y. Bengio, A. Courville, and P. Vincent

The success of machine learning algorithms generally depends on data representation, and we hypothesize that this is because different representations can entangle and hide more or less the different explanatory factors of variation behind the data. Although specific domain knowledge can be used to help design representations, learning with generic priors can also be used, and the quest for AI is motivating the design of more powerful representation-learning algorithms implementing such priors. This paper reviews recent work in the area of unsupervised feature learning and deep learning, covering advances in probabilistic models, auto-encoders, manifold learning, and deep networks. This motivates longer-term unanswered questions about the appropriate objectives for learning good representations, for computing representations (i.e., inference), and the geometrical connections between representation learning, density estimation and manifold learning.

B. PAPER 2

Users of the world, unite! The challenges and opportunities of Social Media

AUTHORS: A. M. Kaplan and M. Haenlein

The concept of Social Media is top of the agenda for many business executives today. Decision makers, as well as consultants, try to identify ways in which firms can make profitable use of applications such as Wikipedia, YouTube, Facebook, Second Life, and Twitter. Yet despite this interest, there seems to be very limited understanding of what the term ‘‘Social Media’’ exactly means; this article intends to provide some clarification. We begin by describing the concept of Social Media, and discuss how it differs from related concepts such as Web 2.0 and User Generated Content. Based on this definition, we then provide a classification of Social Media which groups applications currently subsumed under the generalized term into more specific categories by characteristic: collaborative projects, blogs, content communities, social networking sites, virtual game worlds, and virtual social worlds. Finally, we present 10 pieces of advice for companies which decide to utilize Social Media.

C. PAPER 3

Peer relations in the anxiety-depression link: test of a mediation model.

AUTHORS: B. K. Biggs, J. M. Nelson, and M. L. Sample

We employed a five-month longitudinal study to test a model in which the association between anxiety and depression symptoms is mediated by peer relations difficulties among a sample of 91 adolescents ages 14-17 (M=15.5, SD=.61) years. Adolescents completed measures of anxiety symptoms, depression symptoms, peer group experiences (i.e., peer acceptance and victimization from peers), and friendship quality (i.e., positive qualities and conflict). As hypothesized, Time 1 anxiety symptoms predicted Time 2 (T2) depression symptoms, and this association was mediated by T2 low perceived peer acceptance and T2 victimization from peers, both of which emerged as unique mediators when they were considered simultaneously in the model. Contrary to expectations, qualities of adolescents' best friendships at T2 did not emerge as mediators and were largely unrelated to symptoms of anxiety and depression. Implications of the findings include the importance of addressing peer relations difficulties, especially peer acceptance and victimization, in the treatment of anxiety and the prevention of depression among anxious youth.

III. METHODOLOGY

A. Existing System:

Previous works on computational studies of bullying says that Natural Language processing and Machine Learning are powerful tools to study bullying. Cyber bullying detection is also a supervised learning problem. Firstly, we take a classifier which is trained using the cyber bullying corpus data labeled by humans. Later bullying message recognition is done using that learned classifier.

Disadvantages:

- The first and also difficult step is the numerical representation learning for text messages.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

➢ Only a small portion of messages are left on the Internet, and most bullying posts are deleted due to protection of Internet users and privacy issues.

B. Tools Used:

Operating System	: Windows family.
Coding Language	: J2EE (JSP, Servlet, Java Bean)
Database	: MySql Server
IDE	: Netbeans.7.4
Web Server	: Tomcat 7

C. Proposed System:

The text, user demography, and social network features are the 3 kinds of information that we often use in cyberbullying detection. Our work here focuses on text-based cyberbullying detection because the text content is the most reliable. In this paper, we mainly focus on one deep learning method named stacked denoising autoencoder (SDA). SDA stacks several denoising autoencoders. And also concatenates the output of each layer as a learned representation. To recover the input data from a corrupted version of it, each denoising autoencoder in SDA is trained. By randomly setting some of the input to zero the input is corrupted, which is called dropout noise. This denoising process can be used by the autoencoders to learn robust representation. In addition, to learn an increasingly abstract representation of the input each autoencoder layer is intended. In this paper, based on a variant of SDA we develop a new text representation model: marginalized stacked denoising autoencoders (mSDA), which accelerates training and marginalizes infinite noise distribution by adopting linear instead of nonlinear projection, in order to learn more robust representations. The semantic information has bullying words. An automatic extraction of these words based on word embeddings is proposed. During training of smSDA, we discover the latent structure between bullying and normal words by attempting to reconstruct bullying features from other normal words.The reconstruction of bullying features from normal words is done with the help of correlation information discovered by smSDA, and this in turn facilitates detection of bullying messages without containing bullying words.

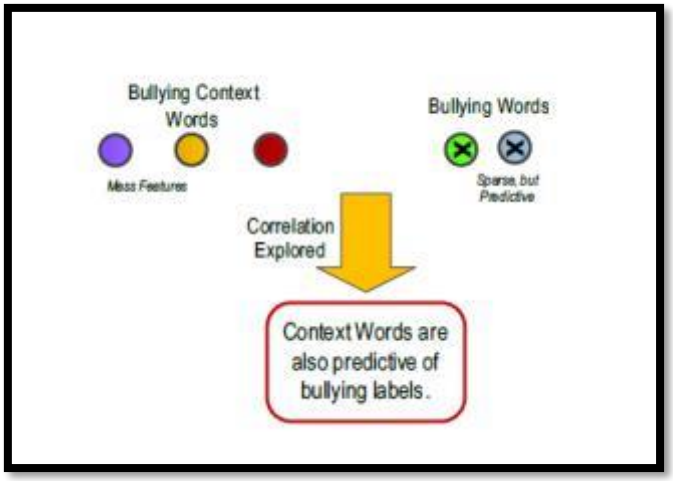

Fig. 2 System Architecture

IV. CODE

```
<div id="facebook-Bar">
<div id="facebook-Frame">
<div><font style="position: fixed; top: 20px; font-size:23px; font-weight:bold; color:#FFF; text-decoration:none; ">Cyberbullying Detection based on Semantic-Enhanced Marginalized Denoising Auto-Encoder </font></div>
<br></br><br><div id="header-main-right">
<div id="header-main-right-nav">
```

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

```
<form method="post" action="loginaction.jsp" id="login_form" name="login_form"> <table border="0" style="border:none">
<tr>
<td ><input type="text" id="email" placeholder="Your Email Id" name="mail" class="inputtext radius1" value=""></td>
<td ><input type="password" id="pass" placeholder="Password" name="pass" class="inputtext radius1" ></td>
<input type="hidden" value="1" name="status" />
<td ><input type="submit" class="fbbutton" name="login" value="Login" /></td>
<td>       </td>
</tr>
<tr>
<!-- <td><label><input id="persist_box" type="checkbox" name="persistent" value="1" checked="1"/><a href="" style="color:#ccc; text-decoration:none">forgot your password?</a></label>
</td>-->
<td><label><a href="" style="color:#ccc; text-decoration:none">forgot your password?</a></label></td>

</tr>
</table>
</form>
</div>
</div>
</div>
</div>
```

V. OUTPUT SCREENS

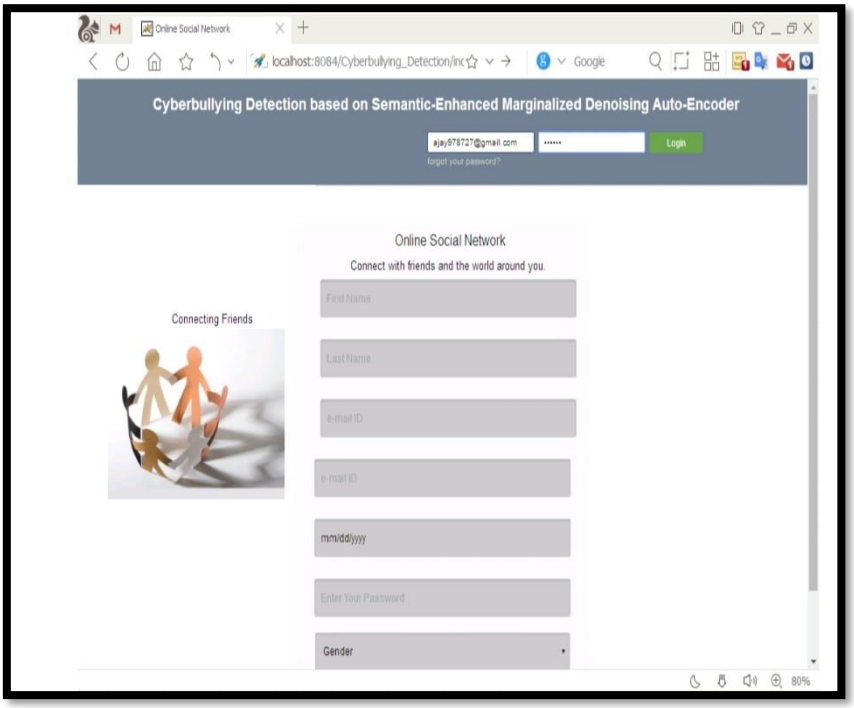

Fig. 3 Output Screen1

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

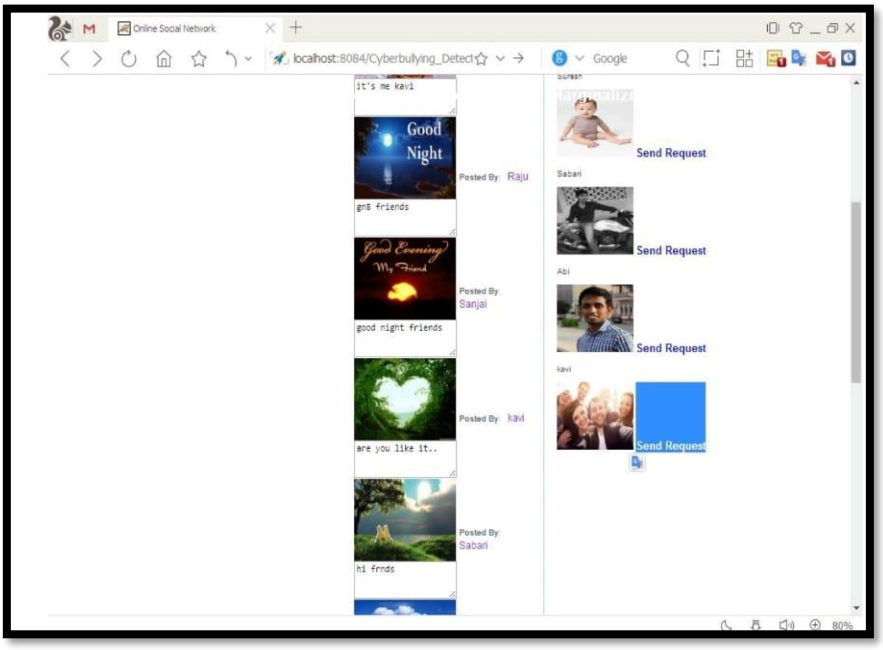

Fig. 4 Output Screen 2

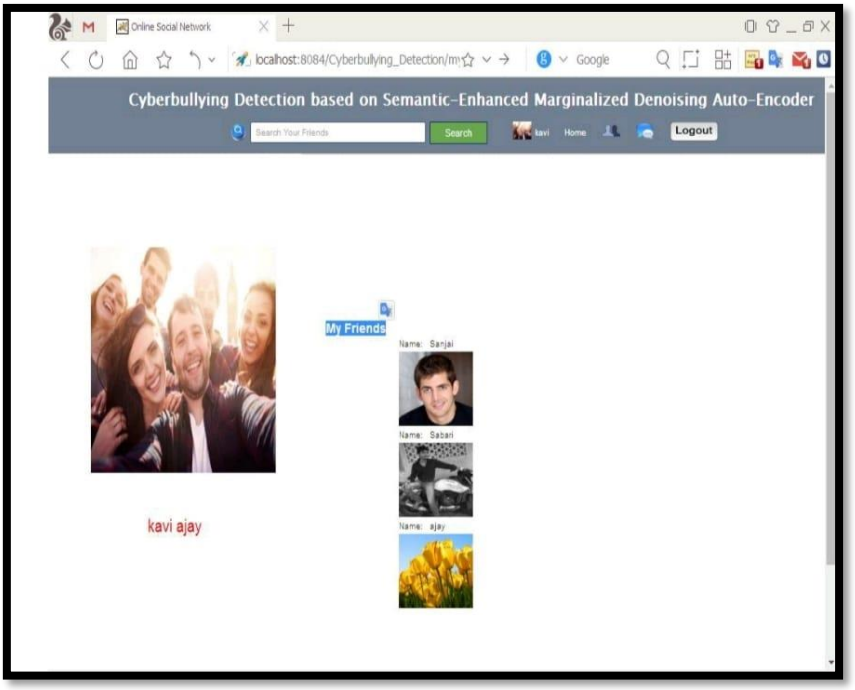

Fig. 5 Output Screen 3

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

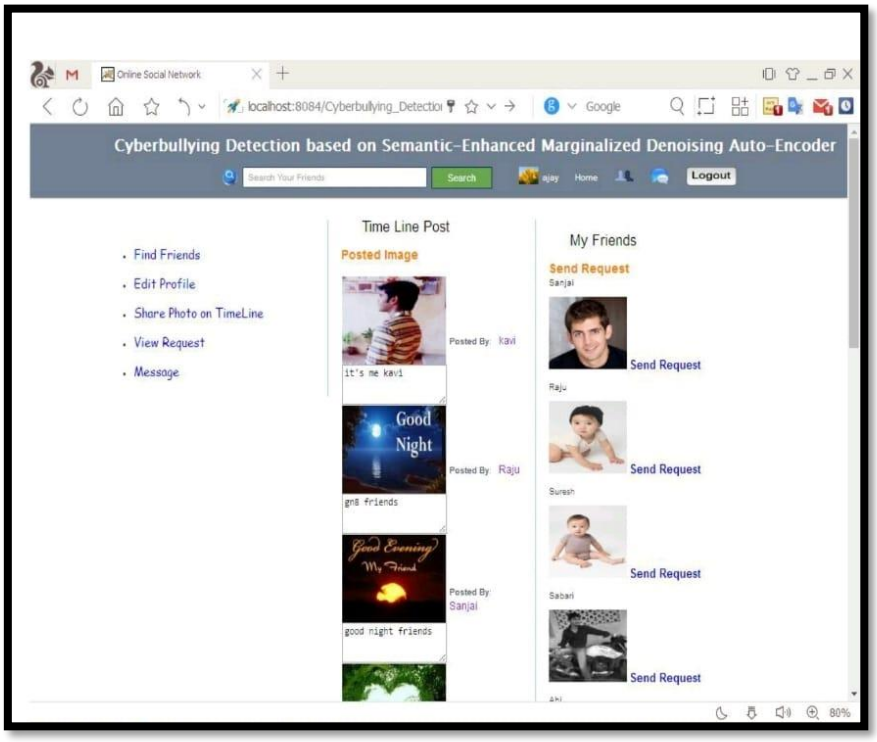

Fig. 6 Output Screen 4

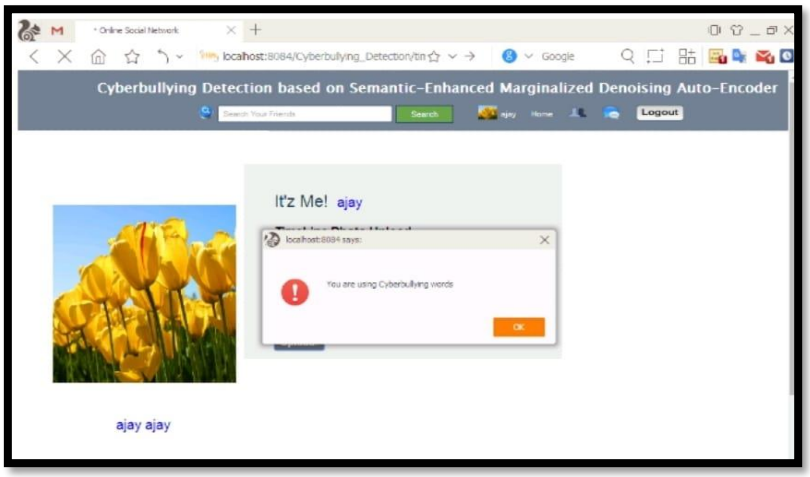

Fig. 7 Output Screen 5

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

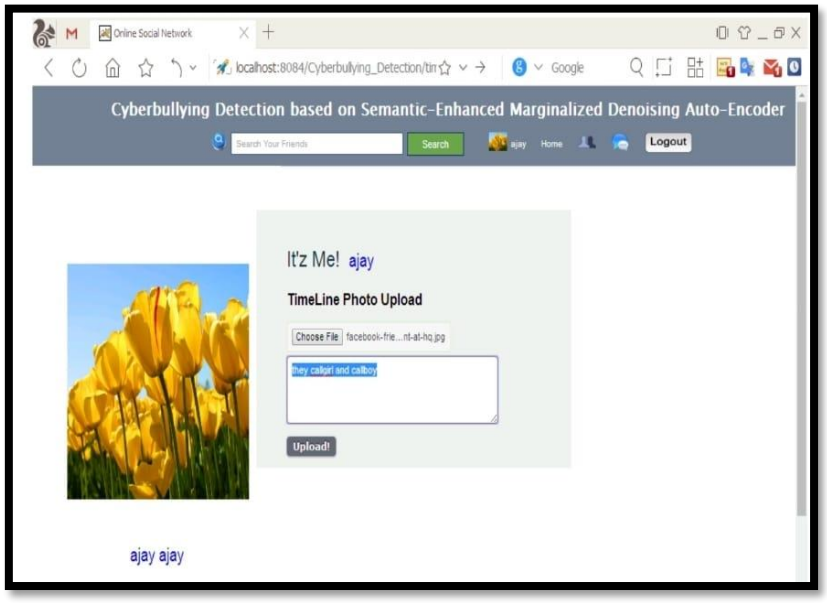

Fig. 8 Output Screen 6

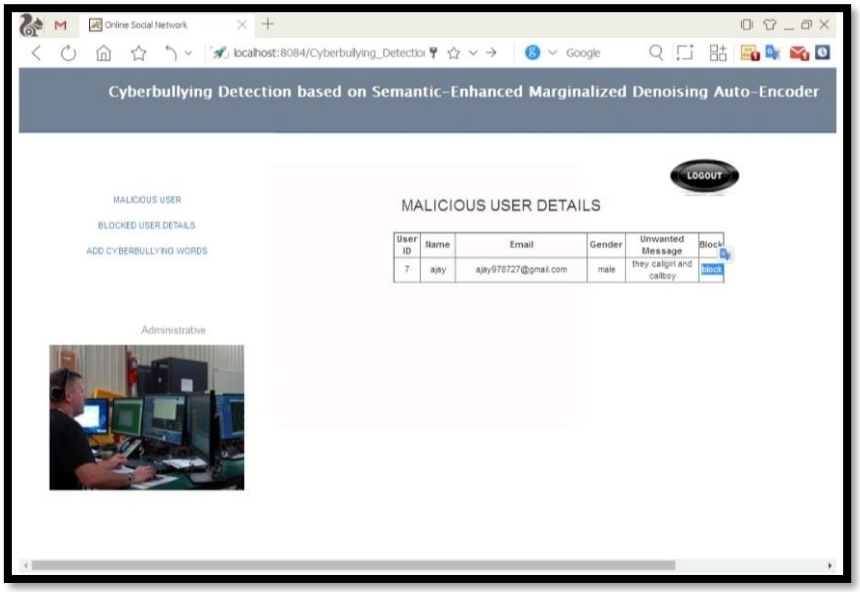

Fig. 9 Output Screen 7

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

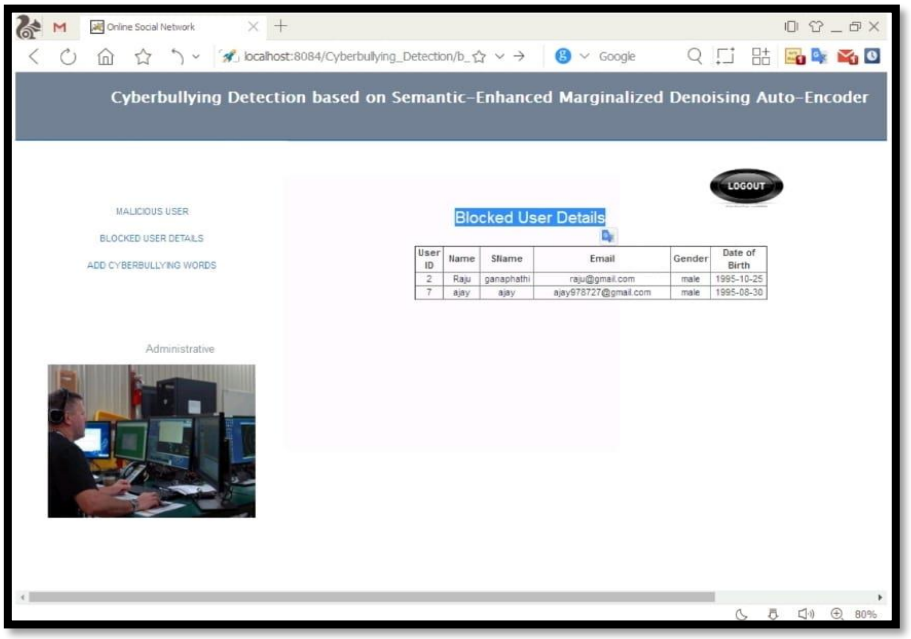

Fig. 10 Output Screen 8

VI. CONCLUSION & FUTURE WORK

This paper addresses the text-based cyberbullying detection problem, in which robust and discriminative representations of messages plays a major role for an effective detection system. We have developed semantic-enhanced marginalized denoising autoencoder as a specialized representation learning model for cyberbullying detection, By designing semantic dropout noise and enforcing sparsity. In addition, we use word embeddings for automatic expand and refine of bullying word lists that which 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.

REFERENCES

[1]. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," J. Mach. Learn. Res., vol. 3, pp. 993–1022, Jan. 2003.
[2]. T. Hofmann, "Probabilistic latent semantic analysis," in Proc. 15th Conf.
[3]. T. Hofmann, "Probabilistic latent semantic indexing," in Proc. 22nd Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, Berkeley, CA, USA, 1999, pp. 50–57.
[4]. C. Wartena and R. Brussee, "Topic detection by clustering keywords," in Proc. 19th Int. Workshop Database Expert Syst. Appl. (DEXA), Turin, Italy, 2008, pp. 54–58.
[5]. F. Archetti, P. Campanelli, E. Fersini, and E. Messina, "A hierarchical document clustering environment based on the induced bisecting k-means," in Proc. 7th Int.
Conf. Flexible Query Answering Syst., Milan, Italy, 2006, pp. 257–269. [Online]. Available:http://dx.doi.org/10.1007/11766254_22.
[6].C. D. Manning and H. Schütze, Foundations of Statistical Natural Language Processing. Cambridge, MA, USA: MIT Press, 1999.
[7]. M. Cataldi, L. Di Caro, and C. Schifanella, "Emerging topic detection on Twitter based on temporal and social terms evaluation," in Proc. 10th Int. Workshop Multimedia Data Min. (MDMKDD), Washington, DC, USA, 2010, Art. no. 4. [Online]. Available: http://doi.acm.org/10.1145/1814245.1814249.
[8]. W. X. Zhao et al., "Comparing Twitter and traditional media using topic models," in Advances in Information Retrieval. Heidelberg, Germany: Springer Berlin Heidelberg, 2011, pp. 338–349.
[9]. Q. Diao, J. Jiang, F. Zhu, and E.-P. Lim, "Finding bursty topics from microblogs," in Proc. 50th Annu. Meeting Assoc. Comput. Linguist. Long Papers, vol. 1. 2012, pp. 536–544.
[10]. H. Yin, B. Cui, H. Lu, Y. Huang, and J. Yao, "A unified model for stable and temporal topic detection from social media data," in Proc IEEE 29th Int. Conf. Data Eng. (ICDE), Brisbane, QLD, Australia, 2013, pp. 661–672.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11108

[11]. C. Wang, M. Zhang, L. Ru, and S. Ma, "Automatic online news topic ranking using media focus and user attention based on aging theory," in Proc. 17th Conf. Inf. Knowl. Manag., Napa County, CA, USA, 2008, pp. 1033–1042.

[12]. C. C. Chen, Y.-T. Chen, Y. Sun, and M. C. Chen, "Life cycle modeling of news events using aging theory," in Machine Learning: ECML 2003. Heidelberg, Germany: Springer Berlin Heidelberg, 2003, pp. 47–59.

[13]. J. Sankaranarayanan, H. Samet, B. E. Teitler, M. D. Lieberman, and J. Sperling, "TwitterStand: News in tweets," in Proc. 17th ACM SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst., Seattle, WA, USA, 2009, pp. 42–51.

[14]. O. Phelan, K. McCarthy, and B. Smyth, "Using Twitter to recommend real-time topical news," in Proc. 3rd Conf. Recommender Syst., New York, NY, USA, 2009, pp. 385–388.

[15]. K. Shubhankar, A. P. Singh, and V. Pudi, "An efficient algorithm for topic ranking and modeling topic evolution," in Database Expert Syst. Appl., Toulouse, France, 2011, pp. 320–330.

[16]. S. Brin and L. Page, "Reprint of: The anatomy of a large-scale hypertextual web search engine," Comput. Netw., vol. 56, no. 18, pp. 3825–3833, 2012.

[17]. E. Kwan, P.-L. Hsu, J.-H. Liang, and Y.-S. Chen, "Event identification for social streams using keyword-based evolving graph sequences," in Proc. IEEE/ACM Int. Conf. Adv. Soc. Netw. Anal. Min., Niagara Falls, ON, Canada, 2013, pp. 450–457.

[18]. K. Kireyev, "Semantic-based estimation of term informativeness," in Proc. Human Language Technol. Annu. Conf. North Amer. Chapter Assoc. Comput. Linguist., 2009, pp. 530–538.

[19]. G. Salton, C.-S. Yang, and C. T. Yu, "A theory of term importance in automatic text analysis," J. Amer. Soc. Inf. Sci., vol. 26, no. 1, pp. 33–44, 1975.

