

ELIMINATE THE CYBER BULLYING MESSAGES IN THE COMMUNICATION NETWORKS

MUMMIDI SATYA DURGA ^{#1}, NAKKINA SASIKIRAN ^{#2}

^{#1} B.Tech Student, Department of Computer Science and Engineering,
Dr.Lankapalli Bullayya College of Engineering (For Women),
Resapuvanipalem, Visakhapatnam-530013.

^{#2} M.Tech Scholar, Department of Computer Science and System Engineering,
Andhra University College of Engineering ,
Visakhapatnam, A.P, India.

ABSTRACT

The process of threatening, or harassing a kid or preteen or a young child by different child exploitation with internet using some rude messages is known as “CYBERBULLYING “.Therefore, it’s changing into a significant downside within the current OSN services by afflicting the kids, young adults with these rude messages. In this current thesis we tend to construct a machine learning approach for mechanically distinguishing the set of cyber bulled messages and take a look at to construct a healthy and safe communication in OSN. For this we tend to develop a completely unique technique Semantic-Enhanced Marginalized Denoising Auto-Encoder (smSDA) via linguistics extension of the popular deep learning model stacked denoising auto-encoder. Here the linguistics extension in the main attempt to maintain a Bag of Words (BoW) into the database, where the message is initially matched with the BoW and if any word is matched with this words, then message is treated as a CyberBulled message, and that will display separately in a cyber-bulled list apart from the normal list.

Key Words: Online Social Networks, Threatening, Afflicting, Marginalized, Semantic-Enhanced, Auto-Encoder, BoW, Cyber Bullying Detection.

I. INTRODUCTION

Data mining is the process of extracting the valuable information from a large data source and transform the raw data into meaningful information. From the below figure 1, we can clearly find out that for the process of data mining the data which should be taken as input will be collected from various resources like world wide web, database and so on. Once the input data is collected it will be then given to the process of data cleaning. Here in this data cleaning process the data will be cleaned and it will be processed in order to identify if there are any un-supervised data available in that input data.

Clustering algorithms are typically used for exploratory data analysis, where there is little or no prior knowledge about the data [4], [5]. This is precisely the case in several applications of *Computer Data Inspection*, including the one addressed in our work. From a more technical viewpoint, our datasets consist of unlabeled objects—the classes or categories of documents that can be found are *a priori* unknown. Moreover, even assuming that labeled datasets could be available from previous analyses, there is almost no hope that the same

classes (possibly learned earlier by a classifier in a supervised learning setting) would be still valid for the upcoming data, obtained from other computers and associated to different investigation processes. More precisely, it is likely that the new data sample would come from a different population.

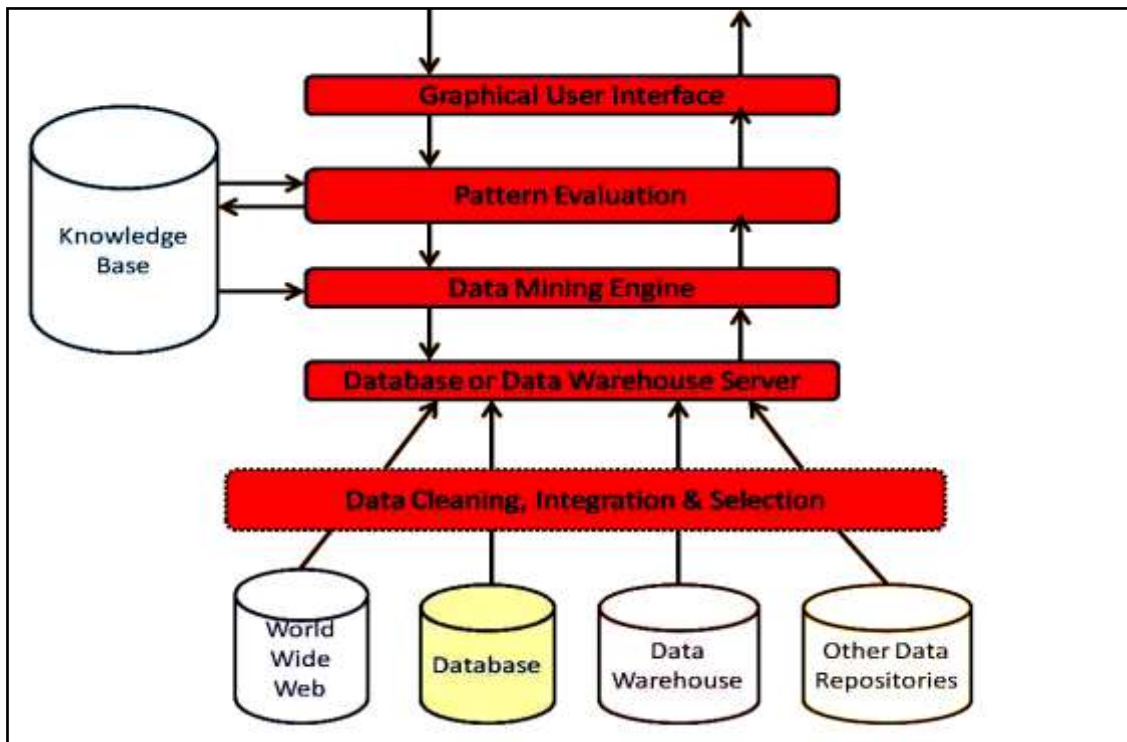


Figure 1: Represents the Basic Architecture of Data Mining

As we all know that Social Media, as defined in [6], 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 cyber bullying, which may have negative impacts on the life of people, especially children and teenagers. Cyber bullying can be defined as aggressive, intentional actions performed by an individual or a group of people via digital communication methods such as sending messages and posting comments against a victim. Different from traditional bullying that usually occurs at school during face-to-face communication, cyberbullying on social media can take place anywhere at any time.

II. LITERATURE WORK

In this section we mainly discuss about the background work that was carried out in finding the work that is related to cyberbullying detection based on the discriminative text representation. Now let us discuss about this in detail:

SCOPE

The main scope or aim for designing this current application or thesis is the term text mining, sometimes also referred to as text data mining, roughly equivalent to text analytics, refers to the process of deriving high-quality information from text data set. Normally from

statistical pattern learning we mainly try to derive the patterns and trends that are calculated from the high quality of information. As we all know that the process of extracting or structuring the input text by identifying the main features and removing the ir-related data from that main document and finally convert the document in a structured way is known as text mining. Also the term high quality in this paper clearly states that combination of some relevance and interestingness for the topic that was available in that conversation file. In the primitive text mining, there are a possibilities like to scan a set of documents written in a natural language and either model the document set for predictive classification purposes or populate a database or search index with the information extracted [6],[7].

Cyber bullying and internet predation occur combine with an extended period of time and across several technological platforms. In this passage we try to observe the current state of research in terms of cyber bullying and internet predation. Here we try to observe the several commercial products which claim to provide chat and social networking site monitoring for home use. The paragraph also provides a detailed summary of research into internet predation and cyberbullying. It reviews the technology that is available for capturing internet messenger (IM) and internet relay chat (IRC). The paragraph discusses the datasets that are currently available for research in the area. It surveys several research articles for both internet predation and cyber bullying detection, as well as provide a summary of the literature as it relates to legal issues. The paragraph finally offers conclusions and discusses opportunities for future research into this interesting and timely field.

III. THE PROPOSED SEMSDAE APPROACH

In this section we mainly discuss about the proposed **Semantic-Enhanced Marginalized Stacked Denoising Auto-Encoder** algorithm for detecting the cyberbullying messages over an OSN walls during communication. Now let us discuss about this proposed SEMSDAE Algorithm in detail as follows:

MOTIVATION

We first introduce notations used in our paper.

Let $D = \{w_1; : : : ; w_d\}$ be the dictionary covering all the words existing in the text corpus.

We represent each message using a BoW vector $x \in \mathbb{R}^d$.

Then, the whole corpus can be denoted as a matrix: $X = [x_1; : : : ; x_n] \in \mathbb{R}^{d \times n}$, where n is the number of available posts.

We next briefly review the marginalized stacked denoising auto-encoder and present our proposed Semantic Enhanced Marginalized Stacked Denoising Auto-Encoder.

A well known author like Mr.Chen et.al proposed a modified version of Stacked Denoising Auto-encoder that employs a linear instead of a nonlinear projection so as to obtain a closed-form solution for our proposed cyberbullying detection.

The basic idea behind denoising auto-encoder is to reconstruct the original input from a corrupted one $\sim x_1; : : : ; \sim x_n$ with the goal of obtaining robust representation.

In this model, denoising auto-encoder attempts to reconstruct original data using the corrupted data via a linear projection. The projection matrix can be learned as:

$$\mathbf{W} = \underset{\mathbf{W}}{\operatorname{argmin}} \frac{1}{2n} \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{W}\tilde{\mathbf{x}}_i\|^2 \quad (1)$$

where $\mathbf{W} \in \mathbb{R}^{d \times d}$. For simplicity, we can write Eq. (1) in matrix form as:

$$\mathbf{W} = \underset{\mathbf{W}}{\operatorname{argmin}} \frac{1}{2n} \operatorname{tr} \left[(\mathbf{X} - \mathbf{W}\tilde{\mathbf{X}})^T (\mathbf{X} - \mathbf{W}\tilde{\mathbf{X}}) \right] \quad (2)$$

Where $\tilde{\mathbf{X}} = [\tilde{x}_1; \dots; \tilde{x}_n]$ is the corrupted version of \mathbf{X} . It is easily shown that Equation. (2) is an ordinary least square problem having a closed-form solution:

IV. CONCLUSION

In this paper, we finally developed a novel SEMSDAE Approach for cyber bullying detection. By integrating BoW database we can refine bullying word lists that is initialized by expert knowledge. The performance of our approaches has been experimentally verified through our simulated application which has same appearance like our social network sites. By conducting various experiments on our proposed framework, we finally came to a conclusion that our proposed approach is robust and easy to represent with word embedding technique. If this was applied in future for all types of OSN sites, we can tell that no OSN site will have any abused or vulgar contents in the communication and all the chat will be free from vulgar or offensive content.

V. REFERENCES

- [1] 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.
- [2] A. Strehl and J. Ghosh, "Cluster ensembles: A knowledge reuse framework for combining multiple partitions," *J. Mach. Learning Res.*, vol. 3, pp. 583–617, 2002.
- [3] Hobbs, Jerry R.; Walker, Donald E.; Amsler, Robert A. (1982). "Proceedings of the 9th conference on Computational linguistics". 1: 127–32. doi:[10.3115/991813.991833](https://doi.org/10.3115/991813.991833).
- [4] Pyle, D., 1999. *Data Preparation for Data Mining*. Morgan Kaufmann Publishers, Los Altos, California.
- [5] 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.
- [6] M. Ybarra, "Trends in technology-based sexual and non-sexual aggression over time and linkages to nontechnology aggression," National Summit on Interpersonal Violence and Abuse Across the Lifespan: Forging a Shared Agenda, 2010.

- [7] G. Salton and C. Buckley, "Term weighting approaches in automatic text retrieval," *Inf. Process. Manage.*, vol. 24, no. 5, pp. 513–523, 1988.
- [8] S. Kotsiantis, D. Kanellopoulos, P. Pintelas, "Data Preprocessing for Supervised Learning", *International Journal of Computer Science*, 2006, Vol 1 N. 2, pp 111–117.
- [9] S. R. Jimerson, S. M. Swearer, and D. L. Espelage, *Handbook of bullying in schools: An international perspective*. Routledge/Taylor & Francis Group, 2010.
- [10] C. C. Paige and M. A. Saunders, "Lsq: An algorithm for sparse linear equations and sparse least squares," *ACM Transactions on Mathematical Software (TOMS)*, vol. 8, no. 1, pp. 43–71, 1982.
- [11] M. A. Saunders et al., "Cholesky-based methods for sparse least squares: The benefits of regularization," *Linear and Nonlinear Conjugate Gradient-Related Methods*, pp. 92–100, 1996.
- [12] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *Journal of the American statistical Association*, vol. 96, no. 456, pp. 1348–1360, 2001.

VI. ABOUT THE AUTHORS



MUMMIDI SATYA DURGA had completed her B.Tech in Computer Science and Engineering at Dr.Lankapalli Bullayya College of Engineering (For Women), Resapuvanipalem, Visakhapatnam in the year 2019. Her area of interests includes Data Mining, Programming Languages, and Web Technologies.



NAKKINA SASIKIRAN had completed his M.Tech in Computer Science and Engineering at Department of Computer Science and System Engineering, Andhra University College of Engineering, Visakhapatnam, in the year 2019. His area of interests includes Data Mining, Networking, Security and Operating Systems.