International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 4 | ISSN : 2456-3307

Fuzzy Based Genetic Operators for Cyber Bullying Detection Using Social Network Data

[1]M.Devi [2]M. Chitra Devi

[1]Research Scholar, Department of Computer Science, Sakthi Arts and Science College For Women, Oddanchatram, Tamil Nadu, India

[2]PG Head & Associate Professor, Department of Computer Science, Sakthi Arts and Science College For Women, Oddanchatram, Tamil Nadu, India

ABSTRACT

Social media getting more and more popular in our day today life. By the popularity of the social media affects the people who involving into it. This makes the technology to work or to feel smarter and makes us lazier. On resulting to this robust and discriminative numerical representation learning of text messages is a critical issue. Hence here we propose a learning method to tackle this issue which is named as Semantic Enhanced Marginalized Denoising Auto Encoder (smsda). Semantic extension of the popular deep learning model stacked denoising auto encoder plays a major role in this method whereas semantic extension consists of semantic dropout noise and sparsity constraints. 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 structure of bullying information and learn a robust and discriminative representation of text. Comprehensive experiments on two public cyber bullying corpora (Twitter and myspace) are conducted, and the results show that our proposed approaches outperform other baseline text representation learning methods.

Keywords: Semantic Enhanced Marginalized Denoising Auto-Encoder, cyberbullying.

I. INTRODUCTION

Internet has become very popular and used around the world in our day to day life. By the growing of internet the cyber security is becoming the most important factor. Currently web 2.0 allows us to access the online related services and some users have been affected by the cybercrimes like cyber bullying experiences internationally. By these kinds of issues the growth of social media gets the negative impacts from the various users. We propose an effective predator and victim identification with semantic enhanced marginalized denoising auto-encoder approach to detect cyber-bullying message from social media through the weighing scheme of feature of selection. We present Model to extract the cyber bullying network, which is used to identify the most active cyber bullying predators and victims to ranking algorithms the existing filters generally work with the simple key word search and are unable to understand the Semantic meaning of the text. So we propose Semantic Enhanced Marginalized Denoising Auto-Encoder.

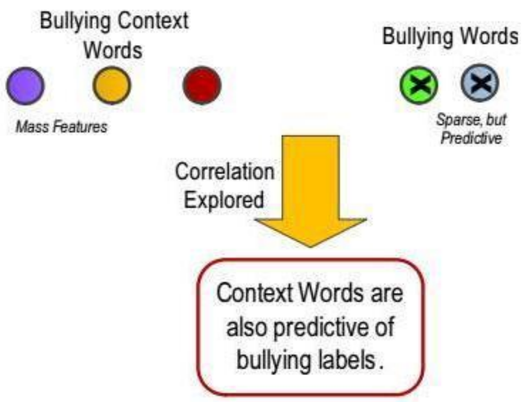

Figure-1: Cyberbullying

Cyberbullying is an increasingly important and serious social problem, which can negatively affect individuals. It is defined as the phenomena of using the internet, cell phones and other electronic devices to willfully hurt or harass others. Due to the recent popularity and growth of social media platforms such as Facebook and Twitter, cyberbullying is becoming more and more prevalent. It has been identified as a serious national health concern by the American Psychological Association 1 and the White House 2 . In addition to that, according to the recent report by National Crime Prevention Council, more than 40% of the teens in the US have been bullied on various social media platforms.

II. RELATED WORK

A Sexual Predator Identification competition took place for the first time at PAN-2012. Given a set of chat logs the participants had to identify the predators among all users in the different conversations or the part (the lines) of the conversations which are the most distinctive of the predator behavior. In conclusion, it is impossible to identify predators using a unique method but it is necessary the use of different approaches. Moreover the most effective method for identifying distinctive lines of the predator behavior in a chat log appeared to be those based on filtering on a dictionary or LM basis [4]. Yin et al., was the sole submission in the misbehavior detection task of CAW 2.0. Using three from the five datasets which were provided by the organizers of the content analysis workshop, they proposed a supervised learning approach for detecting harassment with a focus on detecting intentional annoyance. By employing a SVM classifier with the linear kernel and combining TF-IDF measure as local features, sentiment features, and contextual features of documents proved that identification of online harassment provide significantly improved performance when TF-IDF is supplemented with sentiment and contextual feature attributes. The results show improvements over the baselines. In a recent study on cyberbullying detection, Kontostathis et al., [6] taking a collection of posts from the website Formspring.me, which allows users to post questions anonymously (a question-answer website where users openly invite others to ask and answer questions) proposed a "bag-of-words" language model, which based on the text in online posts, in order to detect instances of cyberbullying. Moreover, they exploited a supervised machine learning called Essential Dimensions of LSI (EDLSI) approach in order to identify additional terms of cyberbullying in Formspring.me data. The data was labeled using a web service, Amazon’s Mechanical Turk.

The Mechanical Turk (MTurk) is a crowdsourcing Internet marketplace that enables individuals or businesses (known as Requesters) to co-ordinate the use of human intelligence to perform tasks that computers are currently unable to do. It is one of the sites of Amazon Wed Services. The goal was to identify the most commonly used cyberbullying term.

III. STATEMENT OF THE PROBLEM

This research is discussed as follows:

(i) Three kinds of information such as text, demography and social features are used for detecting the cyberbullying messages. Hence, text based cyberbullying detection framework is required.

(ii) Each autoencoder layer is intended to learn an increasingly abstract representation of the input.
(iii) Fuzzy rules are used for labeling the cyberbullying messages.
(iv) In addition to genetic algorithm is used for optimizing the parameters for labeling systems.
(v) The correlation information discovered by fuzzy rule generation helps to reconstruct bullying features from normal words, and this in turn facilitates detection of bullying messages without containing bullying words.

IV. CYBERBULLY ACTIVITIES

In the proposed framework for detecting cyberbully activities, following steps have been included:

- Data Pre-processing
- Feature Extraction
- FuzGen learning algorithm
- Naïve classifier technique

4.1. Data Pre-Processing

The data pre-processing is an important phase in representing data in feature space to the classifiers. Social network data are noisy, thus pre-processing has been applied to improve the quality of the research data and subsequent analytical steps, and this includes removing stop words, unwanted characters, etc.

4.2. Feature Extraction

This module is used for extracting the data required from the processed data. The part of speech for every word in the conversation is obtained using natural language processing technique and then features like Noun, Adjective and Pronoun are extracted from the tagged output and statistics on occurrence of word in the text are also extracted.

4.3. FuzGen Learning Technique

The learning module incorporates the adaptive component of the system by means of a GA with fuzzy set genes. GAs are adaptive search and optimization algorithms that work by mimicking the principles of natural genetics (Deb, 1996). In the proposed system, the function to be optimized is a hypothetical representation of cyberbully terms in the Social Network.

In the following, the elements of the GA model, namely: the fuzzy gene types and the GA operators are presented.

4.3.1. The Fuzzy Set Genes

A gene G is, G = (t, g, and c), where
t is frequency of the term,
g identifies the gene type and
c is a non-negative real number

When G(t=c), gene type represents the occurrences of a cyberbully term.
When G(t<c). T his gene type is completely satisfied by dataset that have no occurrences of the cyberbully term t.

When G(t≥c) . Genes of this type are satisfied completely by dataset with at least c occurrences of the cyberbully term t.

4.3.2. The GA Operators

Selection, crossover, and mutation are the genetic operators of evolutionary process. Choice of chromosomes from population to reproduce is done by selection. Using crossover an offspring chromosome is produced by taking sequences of genes from each of two parent chromosomes selected and combining them. The mutation is the random alteration of a gene in the chromosome selected.

4. 4 Advantages

The main advantages of this research are:
(i) These robust features are learned by reconstructing original input from corrupted (i.e., missing) ones. The new feature space can improve the performance of cyberbullying detection even with a small labeled training corpus.

(ii) These specialized modifications make the new feature space more discriminative and this in turn facilitates bullying detection.
(iii) Comprehensive experiments on real-data sets have verified the performance of our proposed model.

V. CYBERBULLYING ALGORITHM

Input: Conversation dataset from Social Network.

Step-1: Current population is assigned to the initial population.
Step-2: Evaluate the current population with the fuzzy rule set given as knowledge base.
Step-3: The fitness value of the current population is calculated using the function EvalPop ().
Step-4: The current population is considered as best population since it is the initial population.
Step-5: The fitness value of the current population is assigned as the best fitness value.
Step-6: The size of the term set retrieved from input is assigned as null.

// For Parent selection

Step-7: The size of the current term set is compared with the size of evolved term set, Ne, if the size of N is less than Ne, then the following steps takes place.
Step-8: The offspring population is initialized as null
Step-9: If the size of offspring population is less than current population then following steps will be executed
Step-10: Parents are selected by using the tournament selection mechanism and children are created by using mutation and cross over mechanism, where Tournament selection is a method of selecting an individual from a population of individuals in a genetic algorithm.

Step-11: Once the offspring population is created, it is joined to current population.
Step-12: End of while loop.
Step-13: Evaluate the Fuzzy rule set for the offspring population.
Step-14: Once the offspring population is created, it is joined to current population.
Step-15: Token competition is carried out to obtain the best individuals from the joint population.
Step16: The Joint population is assigned to the current population.
Step-17: The Fitness value of the joint population is calculated using the function EvalCurPop(). // Updating the Best fitness value and Best population for obtaining classified output.

Step-18: Fitness values of the current population are checked with the best fitness value. If the current fitness value is greater than following steps occur
Step-19: The best fitness value is updated with the current fitness value.
Step-20: The best population is updated with the current population.
Step-21: End of if loop.
Step-22: The size of the current term set is incremented.
Step-23: End of while loop.

Output: Identified Cyberbully terms and their type from the input dataset

VI. RESULT AND DISCUSSION

We used two social media datasets, namely Twitter and MySpace for the problem we study. Both datasets contain labeled social media post. Twitter is a micro blogging website which allows users to post 140 characters messages called "Tweets". The retweets are removed from the dataset. The posts in this dataset have been manually labeled as bully or normal. MySpace is a social networking website which allows a registered users to view pictures, read chat and check other users profile information. The MySpace dataset used in the experiments is crawled from MySpace's groups feature. Each post in the dataset is manually labeled as normal or bully.

	Twitter	Myspace
No. of posts	7321	3245
No. of Features	3709	4236
No. of Positive Posts	2102	950
No. of Negative Posts	5219	2295
No. of users	7043	1053
Average posts per user	1.04	2.98

Table-1: Verifying Sentimental score Distribution

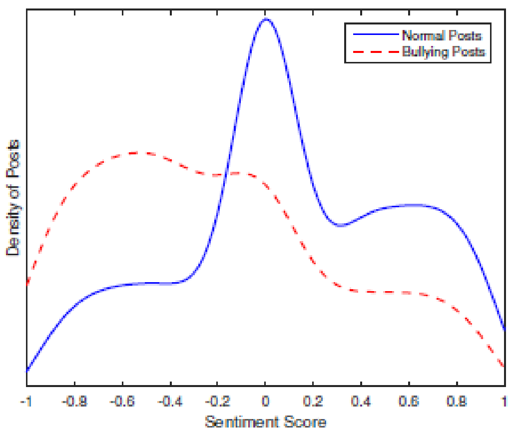

Figure-1: Sentiment Score Distribution of Normal Posts and Bullying Posts in the Twitter Dataset.

Figure-1 shows the sentiment score distribution of the normal and the bullying posts. In Figure-1, the X-axis shows the sentiment polarity score and Y-axis shows the density of users. From the Figure-1 we can observe that two distributions are centered around different mean values. This suggests that there is a clear difference between the sentiment of the normal posts and the bullying posts, and bullying posts tend to have more negative sentiment than normal posts. The sentiment distribution pattern is similar in MySpace dataset.

As the interest and utilization of OSNs are expanding on a regular routine, there emerges the need to fundamentally break down the networks in an efficient manner. The current issues are analyzed:

- Influence Propagation
- Community Detection
- Expert Finding
- Link Prediction
- Recommender systems
- Predicting trust and distrust among individuals
- Opinion mining
- Influence Propagation

(i) Influence Propagation

Domingos and Richardson gave the first algorithmic way to manage influence propagation. At that point, Kempe et al contemplated influence propagation so as to engender on two basic spread models, named Independent Cascade (IC) Model and Linear Threshold (LT) Model, which prompted the advancement of the Greedy Algorithm for influence propagation. They managed the influence propagation issue from an alternate point of view in particular various scalability issues. Chen et al. proposed another proliferation model like the greedy algorithm yet with a superior proficient result. Saito et al were the first to

concentrate how to take in the probabilities for the IC model from an arrangement of past propagations. Goyal et al additionally had made an investigation of the issue of learning impact probabilities utilizing an example of the General Threshold Model (GTC).

(ii) Community Detection

A relative examination on different group location calculations can be found in. Starting study on group or gathering identification was engaged predominantly on the connection structure of OSNs while disregarding the substance of social collaborations, which is likewise pivotal for exact and significant group extraction. It is just as of late that couple of analysts has tended to the issue of finding topically important groups from an OSN.

Pathak et al have proposed a Community-Author-Recipient Topic (Truck) model which utilizes both connection and content data for group location. Liu et al moreover have based a model taking into account Topic-Link Latent Dirichlet Distribution (LDA) however which works just with report systems. Zhao et al have tended to theme situated group discovery through social articles and join investigation in informal organizations. Sachan et al have proposed Topic User Community Model (TUCM) as Topic User Recipient Community Model (TURCM) which offers high time consumption.

(iii) Expert Finding

Analysis on expert ranking estimation is generally taking into account either domain based learning driven systems or space learning free systems or both. The expert ranking issue is likewise looked into on email communication relations. Zhang et al have proposed proliferation based methodology in view of Probabilistic Latent Semantic Analysis (PLSA) for expert finding in social organizations. Authors have utilized the RarestFirst and Enhanced Steiner calculations for expert finding while authors have changed the RarestFirst estimation and discovered the Simplified RareFirst (SRareFirst) estimation. Smirnova et al have proposed a client model for expert finding in light of objective client behavior. Jin et al discovered the ExpertRank calculation which depends on dissecting closeness and power for ranking experts in interpersonal organizations.

(iv) Link Prediction

Liben-Nowell and Kleinberg have managed link forecasting in interpersonal organizations however which works with just a static depiction of a system. Hasan et al have proposed a few characterization models for connection expectation which gives an examination of a few elements by demonstrating their rank of significance as acquired by distinctive estimation. Fouss et al have introduced a connection expectation system in light of a Markov-chain model of random walk however which does not scale well for huge databases. Zheleva et al have utilized a parallel algorithm in which family was utilized for connection expectation.

(v) Recommender systems

Recommender frameworks (RF) have created in parallel with the web. A decent overview on different RS can be found. They were at first in light of demographic, content based and collaborative sifting. Collaborative sifting is the most widely recognized system utilized for RS. Linden et al introduced their work on thing to item shared sifting for amazon.com suggestions. On the other hand, the development of RS has demonstrated the significance of half and half systems of RS, which blend diverse systems with a specific end goal to get the points of interest of each of them.

(v) Predicting Trust and Distrust among Individuals

Various orders have taken a gander at different issues identified with trust. The first errand was the EigenTrust estimation that expects to lessen the number of inauthentic record downloads in a P2P system. Guha et al proposed systems for engendering of trust and distrust, each of which is suitable in specific circumstances. PowerTrust is a trust proposal framework that totals the positive and negative feelings between the clients into the neighborhood trust scores, comparably to EigenTrust. Other work that studies an informal community with positive and negative feelings is introduced. DuBois et al introduced a paper for foreseeing trust and distrust in light of way likelihood in arbitrary diagrams. Kim et al have additionally proposed a technique for anticipating trust and distrust of clients in online networking sharing groups. Ortega et al proposed a novel framework planned to spread both positive and negative assessments of the clients through a system, in such way that the assessments from every client about others impact their worldwide trust score.

(vi) Opinion Mining

The majority of works in this examination concentrated on classifying texts as per their sentiment polarity, which can be positive, negative or neutral. Authors gave a top to bottom study of supposition mining and sentiment analysis. The issue was concentrated utilizing directed considering so as to learn logical feeling influencers, for example, invalidation (e.g., not and never) and contrary (e.g., yet and in any case). Wilson et al have considered a few distinctive learning estimations, for example, boosting, rule learning, and Support Vector Machines that can consequently recognize subjective and objective (impartial) dialect furthermore among weak, medium and strong subjectivity.

VII. CONCLUSION

The paper addresses the text-based cyber bullying detection problem, where we have developed semantic enhanced marginalized denoising auto encoder as a specialised illustration learning model for cyber bullying detection. In addition, word embeddings have been wont to automatically expand and refine bullying word lists that's initialized by domain information. The performance of our approaches has been experimentally verified through cyber bullying methods. As a next step we area unit coming up with to additional improve the strength of the learnedillustrationbyconsidering ordination in messages.

VIII. 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]. 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.

[3]. M. Ybarra, "Trends in technology-based sexual and non-sexual aggression over time And linkagest non technology aggression,"National Summiton Interpersonal Violence and Abuse Across the Lifespan: Forging a Shared Agenda, 2010.

[4]. 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.

[5]. S. R. Jimerson, S. M. Swearer, and D. L. Espelage, Handbook of bullying in schools: An international perspective. Rout ledge/Taylor & Francis Group, 2010.

[6]. G. Gini and T. Pozzoli, "Association between bullying and psychosomatic problems: A meta-analysis," Pediatrics, vol.123, no. 3, pp. 1059–1065, 2009.

[7]. A. Kontostathis, L. Edwards, and A. Leatherman, "Text mining and cybercrime," Text Mining: Applications and Theory. John Wiley & Sons, Ltd, Chichester, UK, 2010.

[8]. 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.

[9]. 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.

[10]. 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.

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

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

[13]. 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.

[14]. P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," The Journal of Machine Learning Research, vol. 11, pp. 3371–3408, 2010.

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