International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Sentiment Analysis for Bullying Word Detection in Social Network

Srilakshmi

[1]Vishweshwaraiah Technological University, Computer Science Engineering, UBDTCE, Davanagere

Abstract: ***As a side effect of increasingly popular social media, cyberbullying has emerged as a serious problem afflicting children, adolescents and young adults. Machine learning techniques make automatic detection of bullying messages in social media possible and this could help to construct a healthy and safe social media environment. In this meaningful research area, one critical issue is robust and discriminative numerical representation learning of text messages. In this paper, we propose a new representation learning method to tackle this problem. Our method named natural language processing and artificial intelligence. Sentiment analysis is one of the methods which classify the given sentence as positive, neutral and negative. Natural language processing which classifies the data using navy bayies classifier. We use the social media tweeter to find the bullying words and classify the words or text as positive, neutral and negative.***

Keywords: sentiment analysis, natural language processing, navies bayes classifier, artificial intelligence

1. Introduction

Social network is a collection of applications based on internet that makes the philosophical and technological basis of Web 2.0, and that enable the creation and interchange of content generated by user, through social network, humans can enjoy huge information, experience suitable communication and so on. But, social network may have some side effects such as cyber bullying, which may have negative effects on the humans life, usually kids and youths.

Cyber bullying can be explained as hostile, intended actions performed by a single person or a group of person communication via digital methods such as sending messages in social network and posting comments in social network against a victim. Different from traditional bullying that usually occurs at school during face-to-face communication, cyber bullying on social media can take place anywhere at any time. For bullies, they are free to hurt their peers feelings because they do not need to face someone and can hide behind the Internet. For victims, they are easily exposed to harassment since all of us, especially youth, are constantly connected to Internet or social media The same as traditional bullying, cyber bullying has negative, insidious and sweeping impacts on children. The outcomes for victims under cyber bullying may even be tragic such as the occurrence of self-injurious behavior or suicides.

One way to address the cyber bullying problem is to automatically detect and promptly report bullying messages so that proper measures can be taken to prevent possible tragedies. Previous works on computational studies of bullying have shown that natural language processing and machine learning are powerful tools to study bullying. A classifier is first trained on a cyber bullying 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 cyber bullying detection. Since the text content is the most reliable, our work here focuses on text-based cyber bullying detection.

2. System Design

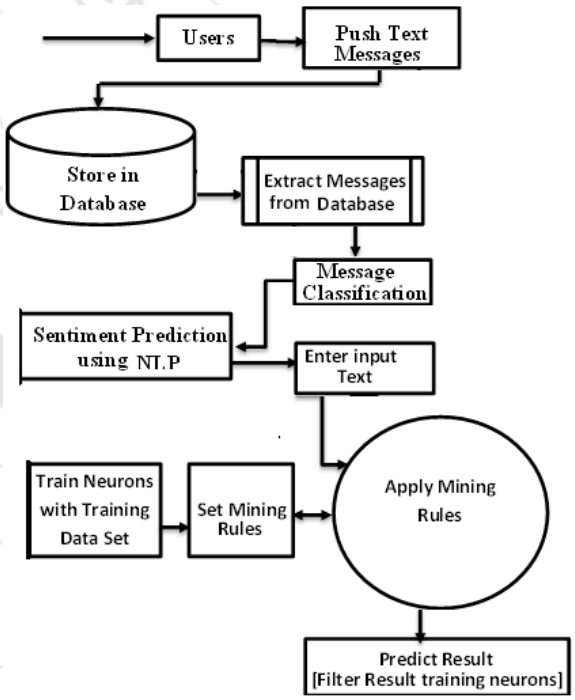


Figure 2.1: System Design for bullying words detection

In this design, users first enters a text messages, and this text messages are stored in database. From the database, all the messages that is entered by the users are extracted and the classification of the messages is done. Once classification of the messages is done then The Sentiment Prediction Is Done Using NLP. All the classified messages are given as Input and for these messages mining rules are applied to detect whether the messages entered by the user is authenticate. Neurons are trained using the training dataset. Finally result of the messages is predicted that is the entered messages are positive, negative or neutral.

3. Methodologies

- Natural Language Processing
- Sentiment Analysis
- Artificial Intellegence
- Weka

3.1 Natural Language Processing

The field of study that focuses on the interactions between human language and computers is called Natural Language Processing, or NLP for short. It sits at the intersection of computer science, artificial intelligence, and computational linguistics.

NLP algorithms are typically based on machine learning algorithms. Instead of hand-coding large sets of rules, NLP can rely on machine learning to automatically learn these rules by analyzing a set of examples (i.e. a large corpus, like a book, down to a collection of sentences), and making a statical inference. In general, the more data analyzed, the more accurate the model will be.

- Summarize blocks of text using Summarizer to extract the most important and central ideas while ignoring irrelevant information.
- Create a chat bot using Parsey Mc Parseface, a language parsing deep learning model made by Google that uses Point-of-Speech tagging.
- Automatically generate keyword tags from content using AutoTag, which leverages LDA, a technique that discovers topics contained within a body of text.
- Identify the type of entity extracted, such as it being a person, place, or organization using Named Entity Recognition.
- Use Sentiment Analysis to identify the sentiment of a string of text, from very negative to neutral to very positive.

3.1.1 Naïve Bayes Classifier

Naive Bayes classifiers are studying the classification task from a Statistical point of view. The starting point is that the probability of a class is given by the posterior probability P(C|D) given a training document D.
Here D refers to all of the text in the entire training set. It is given by D= (d_1, d_2, …, d_n) where $d_i$is the attribute of document D.

Using Bayes' rule, this posterior probability can be rewritten as:

$$P(C=c_i|D) = \frac{P(C=c_i)\cdot P(C=c_i)}{P(D)}$$

Since the marginal probability P(D) is equal for all classes, it can be disregarded and the equation becomes:

$$P(C = c_i|D) = P(D|C = c_i) \cdot P(C = c_i)$$

The document D belongs to the class C which maximizes this probability, so:

$$C_{NB} = argmax\ P(D|C) \cdot P(C)$$
$$C_{NB} = argmax\ P(d_1, d_2, \ldots., d_n) \cdot P(C)$$

Assuming conditional independence of the words , this equation simplifies to:

$$C_{NB} = argmaxP(d_1|C) \cdot P(d_2|C), \ldots \ldots cdotP(d_n|C) \cdot P(C)$$
$$C_{NB} = argmaxP(C) \cdot \prod_i P(d_i|C)$$

Here $P(d_i|C)$ is the conditional probability that word i belongs to class C. For the purpose of text classification, this probability can simply be calculated by calculating the frequency of word i in class C relative to the total number of words in class C.

$$P(d_i|C) = \frac{count(d_i|C)}{\sum_i count(d_i|C)}$$

3.2 Sentiment Analysis

Sentiment analysis (also known as **opinion mining**) refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials. Sentiment analysis is widely applied to reviews and social media for a variety of applications, ranging from marketing to customer service.

Generally speaking, sentiment analysis aims to determine the attitude of a speaker or a writer with respect to some topic or the overall contextual polarity of a document. The attitude may be his or her judgment or evaluation (see appraisal theory), affective state (that is to say, the emotional state of the author when writing), or the intended emotional communication (that is to say, the emotional effect the author wishes to have on the reader).

A basic task in sentiment analysis is classifying the polarity of a given text at the document, sentence, or feature/aspect level—whether the expressed opinion in a document, a sentence or an entity feature/aspect is positive, negative, or neutral. Advanced, "beyond polarity" sentiment classification looks, for instance, at emotional states such as "angry", "sad", and "happy".

3.3 Artificial Intelligence

Artificial intelligence (AI) is intelligence exhibited by machines. In computer science, the field of AI research defines itself as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of success at some goal. Colloquially, the term "artificial intelligence" is applied when a machine mimics "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving". As machines become increasingly capable, mental facilities once thought to require intelligence are removed from the definition. For example, optical character

recognition is no longer perceived as an example of artificial intelligence, having become a routine technology. Capabilities currently classified as AI include successfully understanding human speech, competing at a high level in strategic game systems, self-driving cars, intelligent routing in Content Delivery Networks, and interpreting complex data.

3.4 Weka

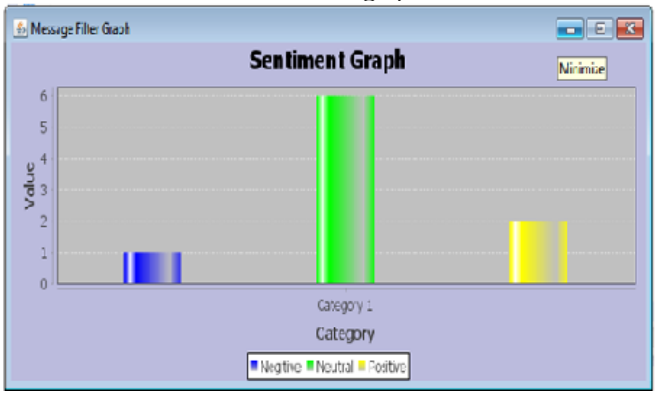
Weka contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to these functions. The original non-Java version of Weka was a front-end to modeling algorithms implemented in other programming languages, plus data reprocessing utilities in C, and a Make file-based system for running machine learning experiments. This original version was primarily designed as a tool for analyzing data from agricultural domains, but the more recent fully Java-based version, for which development started in 1997, is now used in many different application areas, in particular for educational purposes and research. Advantages of Weka include:

- Free availability under the GNU General Public License.
- Portability, since it is fully implemented in the Java programming language and thus runs on almost any modern computing platform.
- A comprehensive collection of data preprocessing and modeling techniques.
- Ease of use due to its graphical user interfaces.

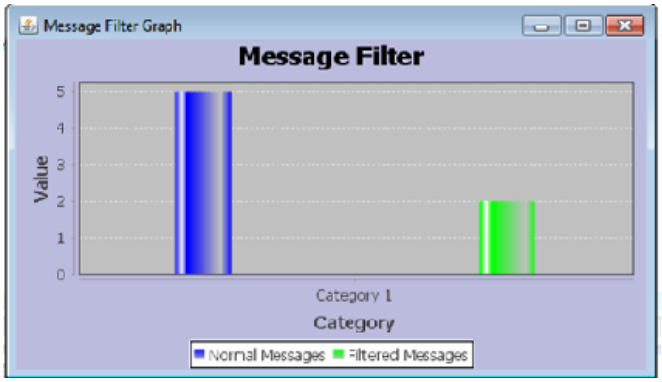
Weka supports several standard data mining tasks, more specifically,datapreprocessing, clustering, classification, regr ession, visualization, and feature selection. All of Weka's techniques are predicated on the assumption that the data is available as one flat file or relation, where each data point is described by a fixed number of attributes (normally, numeric or nominal attributes, but some other attribute types are also supported).

4. Results and Discussions

After applying all this methodologies into the twitter website which is created artificially the below results are shown, which classifies the given input text as positive ,negative and neutral shown in below sentiment graph.



The message filter graph shown below, which shows the percentage of normal messages and filtered messages shown in the below graph.



5. Conclusion

This addresses the text-based bullying detection problem in social network, where robust and discriminative representations of messages are critical for an effective detection system. By designing sentiment analysis for the given text document we can predict the user opinion in the social network. So we can create the healthy environment by avoiding bullying words in social network. The performance of our approaches has been experimentally verified through cyber bullying corpora social medias such as Twitter. As a next step we are planning to further improve the robustness of the learned representation by considering word order in messages.

References

[1] 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.R. Caves, Multinational Enterprise and Economic Analysis, Cambridge University Press, Cambridge, 1982. (book style)

[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 linkages to nontechnology aggression,” National Summit on 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. Routledge/Taylor & Francis Group, 2010.

Author Profile

Srilakshmi received MTech in Computer Science Engineering from University BDT College of Engineering in 2017 and B.E degree in Information Science & Engineering from Sri Siddhartha Institue of Technology in 2012, respectively.