CyberBullying

Michael Travers

Introduction

Thanks to the internet, bullying has become more widespread than ever, affecting the lives of all ages and races. From suicide, mental trauma, and issolation; cyberbullying has become a very serious phenomenon that needs to address to saves the minds and lives of millions of the now and future. The Effects on the victims can be life traumatizing for some giving , high rates of depression, physical pain, and some even develop eating disorders.

Motivation

46 percent of people ages 12 -20 has experience cyberbullying at least once

Problem becomes increasingly worse rising

75 percent of cyberbullying takes place on Facebook and 24 percent of cyberbullying takes place on instagram and twitter

Decreasing cyberbullying can help save lives and peoples mental trauma

Create algorithms to help classify and distinguish toxic and abusive messages through social media apps like facebook and twitter.

Causes

- cyberbullies usually have deeper issues within themselves. From bullying to getting popular and feeling some type of popularity,
- having poor relationships with their parents
- Difficulty with having empathy for peers.

Problem Statement

The Problem of cyberbullying is the effects it has on the people getting bullied, and to make matters worse sometimes people that get cyberbullied also have to deal with the torment at school, which can create a lot of mental issues including

- Depression
- Physical Pain
- Low Self Esteem
- Suicide

Humans don't have the efficiency to sort through millions of account and data to identify abuse or bullying

Solutions

Even with Policing being deployed on these platforms to prevent cyberbullying and blocking and hold suspects accountable, the issue is that these platforms have so many accounts it's hard to keep monitoring millions of users. As technology. As AI improves and allows deeper and more sophisticated machine learning that can help with identifying speech recognition and detect abusive behavior and allow authorities to be alerted. AI also helps censor content or allow content moderation that is considered as false information or abusive content. Data Analytics and AI over time will make the internet a more comfortable, safer and healthier space .

AI and Machine learning solutions

AI improves and allows deeper and more sophisticated machine learning that can help with identify speech recognition and detect abusive behavior and allow authorities to be alerted.

censor content or allow content moderation that is considered as false information or abusive content.

Data Analytics and AI over time will make the internet a more comfortable, safer and healthier space.

Machine Learning

Using machine learning a database can be made to track and detect cyberbullying

By using machine learning a algorithm is made to search for offensive languages or signs of cyberbullying including the use of words that usually has harmful meaning. This includes, swearing, negative adjectives, words of threats, or anything devaluing.

With the ability to search and detect harmful behavior or patterns of bullying it becomes possible to stop cyberbullying and ban or warn those who are suspected of it.

With data collection of all accounts a Classification of collected data was used with a SQL database queries and is compared to the classification

Text Mining and NLP

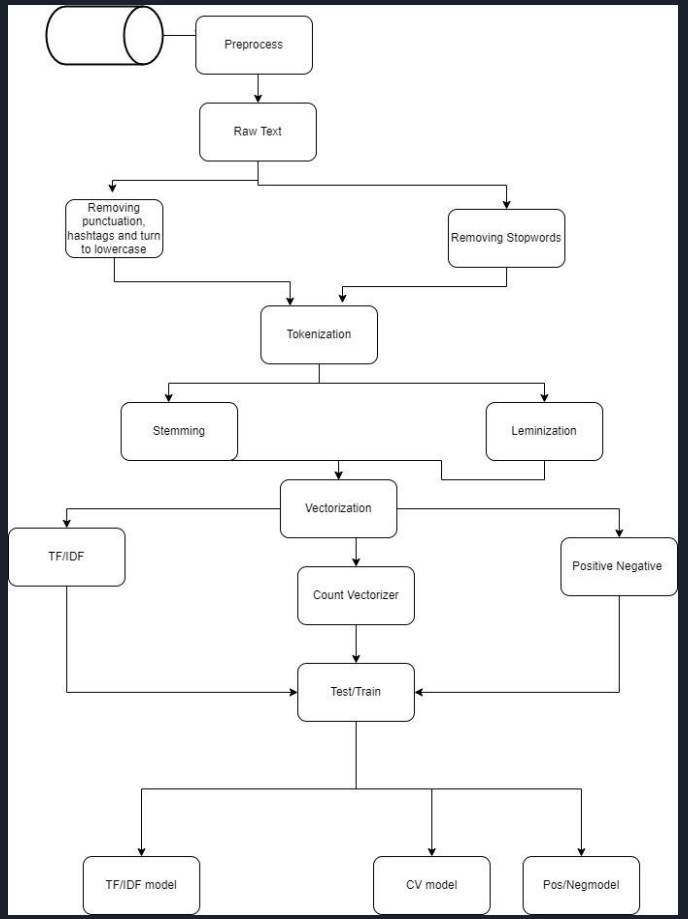
Text Mining is used in a multitude of devices and programs including, virtual assistants, Chat bots, Web search and Machine Translation. Computers are binary and with the Human language being diverse with complexities and ambiguities, text mining can make for a succession of letters so the document can make sense to the computer. Text mining can be split into 2 types of data, structured data where there's a structured format, a star rating system for example or a like and dislike button. And Unstructured data, where things like harmful messages, text reviews and relevant information.

UMI of methodology

Text Mining and NLP

This structure.

- Raw Text
- Normalization
- Tokenizing
- Vectorizing
- TD/IDF, Pos/Neg, Count Vect
- Train Test
- View model
- Review performance



Raw Text- Uploading csv file

Normalization - removed punctuation, turn to lowercase, removing stopwords

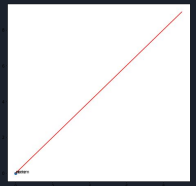
Tokenization- turned each tweet into a single

entity

Vectorize and apply count vect

Apply vectorization to TF/IDF

Create model of the Positive/Negative Frequencies



```
0    [aalwuhaib1977, muslim, mob, violenc, hindu, bangladesh, continu, 2014, islam, httptcoc1jbwjwurc]
1                                                                    [te4mnightm4r, httptco5ih7mkdbqg]
2    [jncatron, israjourisra, ampalestin, islamophobia, like, idea, naziphobia, islam, religion, hate...
3                                  [final, im, caught, sudden, death, cook, look, like, gonna, intens, mkr]
4                                                                    [carolinesind, herecomesfran, hug]
Name: clean_text, dtype: object
```

	2014	23	7	9	aalwuhaib1977	advic	aligharib	amaz	amp	ampalestin	approach	approv	arab
0	0.327902	0.000000	0.000000	0.000000	0.327902	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.304962	0.000000	0.000000	0.000000
3	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.291449
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.473947	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8	0.000000	0.308197	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9	0.000000	0.000000	0.283389	0.283389	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.283389	0.000000
10	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.274198	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.282331	0.000000	0.000000

	2014	23	7	9	aalwuhaib1977	advic	aligharib	amaz	amp	ampalestin	approach	approv	arab	athenahollow	aymannathem	bangladesh	better
0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	(
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(
2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	(
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(
6	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	(
7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	(
8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(
9	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	(
10	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	(

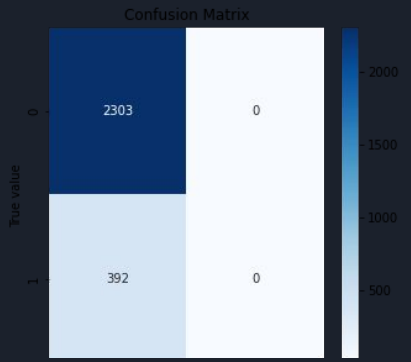
Techniques

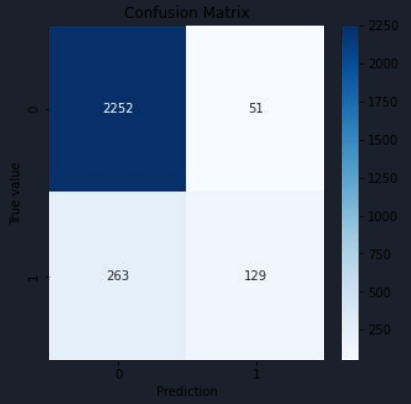
After vectorization is applied the data must be split and training with 80 percent training a

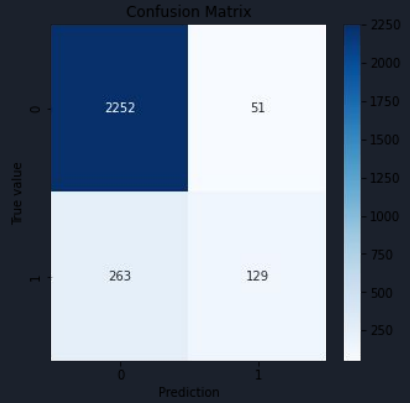
Once each model is trained a representation of each a model of the accuracy

```
Size of X_train: 10776
Size of y_train: 10776

Size of X_test: 2695
Size of y_test: 2695

Train proportion: 802
```

Positive/Negative	Count Vectorizer	TF/IDF
Accuracy: 85.45	88.35	88.35







Conclusion

Cyberbullying is a huge issue that needs to be addressed to save the well being of others but in order to combat cyberbullying a AI algorithm needs to be implemented to search for suspects. Using the Natural Language Process it is possible with text mining. Having the basic structure of importing raw text to be normalized and tokenized, Vectorizing to be so the computer can easily compute the data, then modeling and evaluating performance of the model we are able to identify and represent the percentage and users of using racism or suspected of cyberbullying with unique words. Looking that models created of the Positive Negative Frequencies, Count Vectorizing and Term Frequency/Inverse Document Frequency it shows that using the method of TF/IDF is the most useful as TF/IDF gives more value to words and using the count vectorizer only vectorize words in all the same way.

Refrences

[1]*A Multilingual System for Cyberbullying Detection: Arabic ...*
https://www.researchgate.net/profile/Chamoun-Maroun/publication/322730160_A_Multilingual_System_for_Cyberbullying_Detection_Arabic_Content_Detection_using_Machine_Learning/links/5a6eca350f7e9bd4ca6d658b/A-Multilingual-System-for-Cyberbullying-Detection-Arabic-Content-Detection-using-Machine-Learning.pdf.

[2]-Dean Chester, et al. “How Ai Can Help Fight Cyberbullying.” *TechTalks*, 6 Sept. 2019, https://bdtechtalks.com/2019/09/05/artificial-intelligence-online-bullying/.

[3]“11 Facts about Cyberbullying.” *DoSomething.org*, https://www.dosomething.org/us/facts/11-facts-about-cyber-bullying.

[4]Assistant Secretary for Public Affairs (ASPA). “What Is Cyberbullying?” *StopBullying.gov*, 27 Aug. 2021, https://www.stopbullying.gov/cyberbullying/what-is-it.

[5]“Effects of Cyberbullying: American SPCC - Negative Consequences of Cyberbullying.” *American SPCC*, 7 Oct. 2021, https://americanspcc.org/impact-of-cyberbullying/.

[6] Hinduja, Sameer, et al. “How Machine Learning Can Help Us Combat Online Abuse: A Primer.” *Cyberbullying Research Center*, 18 Sept. 2019, https://cyberbullying.org/machine-learning-can-help-us-combat-online-abuse-primer.

[7]McQuade, Samuel C., et al. *Cyber Bullying: Protecting Kids and Adults from Online Bullies*. Praeger, 2009.

[8]Talpur, Bandeh Ali, and Declan O’Sullivan. “Cyberbullying Severity Detection: A Machine Learning Approach.” *PLOS ONE*, Public Library of Science, https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0240924.

NLP

[9]Training, R-Tutorials. “Text Mining, Scraping and Sentiment Analysis with R.” *Udemy*, Udemy, https://www.udemy.com/course/r-social-media-mining-scraping-with-twitter/?gclid=CjwKCAiAhreNBhAYEiwAFGGKPHoH3XhlaZ1zuhH9ZywJPYdeHPPaN0TNfrjXpIDjAGvXVSNlEXKlDhoCxUAQAvD_BwE&matchtype=b&utm_campaign=LongTail_la.EN_cc.US&utm_content=deal4584&utm_medium=udemyads&utm_source=adwords&utm_term=_._ag_84297325372_._ad_532070152548_._kw_%2Btext%2B%2Bmining%2B%2Btraining_._de_c_._dm__._pl__._ti_kwd-649476312418_._li_9061285_._pd__._.

[10].Termonia, Benjamin. “Applied Text Mining and Sentiment Analysis with Python.” *Udemy*, Udemy, https://www.udemy.com/course/applied-text-mining-and-sentiment-analysis-with-python/#content.

[11]Trolley, Barbara, and Constance Hanel. *Cyber Kids, Cyber Bullying, Cyber Balance*. Corwin Press, 2010.

[12]“What Is Tokenization: Methods to Perform Tokenization.” *Analytics Vidhya*, 23 July 2021,

https://www.analyticsvidhya.com/blog/2019/07/how-get-started-nlp-6-unique-ways-perform-tokenization

[13]Wu, Jun. “Ai, Cyberbullying, and Social Media.” *Medium*, Towards Data Science, 12 July 2019, https://towardsdatascience.com/ai-cyberbullying-and-social-media-321d91d5b4ba.