Sentiment Analysis as an Indicator to Evaluate Gender disparity on Sexual Violence Tweets in South Africa

Jude Imuede
School of Computer Science and Applied Mathematics,
University of the Witwatersrand,
Johannesburg, South Africa.
judeimuede@gmail.com

Mpho Raborife
School of Consumer Intelligence and Information Systems,
University of Johannesburg,
Johannesburg, South Africa.
mraborife@uj.ac.za

Pravesh Ranchod
School of Computer Science and Applied Mathematics,
University of the Witwatersrand,
Johannesburg, South Africa.
pravesh.ranchod@wits.ac.za

***Abstract*—South Africa has one of the highest rates of sexual violence in the world, and with this technology era being the information age, it is not suprising that people are turning to social media to voice their views on this matter. As tweets on sexual violence grow continually, extracting insights from such data demands a robust, real-time, and scalable tools with flexibility in the database schema. Elasticsearch (`ES`) is an example of a free-license search engine written in Java and developed on Apache Lucene that meets the stated requirements. On the other hand, `Kibana` facilitates intuitive dashboard development, visual exploration, and real-time analysis of an index in `ES` through an intuitive graphical user interface. This study demonstrates how gender was inferred and evaluated through the integration of deep neural networks and Google's TensorFlow. We used the AFINN model to infer sentiment analysis as our measure of gender disparity in this instance. The `Indexer`, built of `Node.js`, and defined as the hub of the system connects with the Twitter streaming API to ingest tweets found within the boundaries of sexual violence. This system runs persistently with tweets through the `ES` search engine and visualised in `Kibana`.**

***Index Terms*—Twitter API, Tweet, Gender, Sentiment analysis, Sexual violence, TensorFlow, South Africa.**

I. Introduction

As a piece of a bigger project, we addressed the rampant nature of sexual violence in South Africa, and the robustness of our adopted deep neural network DNN technique emphasises that our approach is timely, viable and appropriate. Through the evolving understanding of gender-based violence (`GBV`), our study attempts to infer gender as a hidden demographic attribute from sexual violence tweets using sentiment analysis derived from AFINN model to disproportionately evaluate the gender classes (male and female) reporting on this situation. With that in mind, we trained a DNN using a pre-made estimator (DNNClassifier) from the TensorFlow high-level API (TF Estimator) to perform gender classification and prediction [1]. The AFINN model was containerised into a custom built application called Sentiment analyser. The AFINN model is a score-based list comprising of English words and implemented in Figure 3 to extract sentiment from a given tweet [2]. The approach used is polarised-oriented from English words in the AFINN model dictionary. The idea around this is that given a tweet text, the system will sum up weights for both positive, negative and neutral words. The sentiment class with the highest value is summed up and this reflects the sentiment of the text [3]. From there, we went on to visualise our results on the Web after training and serving the trained model on TensorFlow Serving (TF Serving) using services from Elasticsearch (`ES`) and `Kibana` through the injection of `Node.js`.

Twitter, unlike Facebook and Google+, does not provide a field for the gender of a user, making gender a hidden attribute that intelligently needs to be inferred through the implementation of machine learning (ML) algorithms. The name of a person tends to be the preferred indicator of gender in most languages [4]. Since, name data on Twitter may not be reliable, it is necessary to infer gender from other available user information. In recent years, mining Twitter data and inferring hidden attributes like gender, age, education status and ethnicity have had a limited representation in literature [4]. However, Mislove et al. [5], considered geographical distribution of Twitter users and highlighted gender and ethnicity in the United States of America (USA) based on a set of over one billion tweets collected between 2006 and 2009. They determined gender by matching the first name in the name field of the user profile with a list of popular baby first names born in the USA from data of the Social Security Administration. They discovered that while the Twitter population was male-biased, this bias decreased over the duration of data representation. This is a perceived limitation inferred from the study in that the consideration of revealing a person's actual gender resulted in under representation within a gender class.

`GBV` can be perpetrated by either a man or a woman. Our study is carried out considering this premise. Based on multiple tweets posted on Twitter relating to `GBV`, the main aim was to analyse such cases within the context of `GBV` so as to provide stakeholders such as police, prosecutors and the public

978-1-7281-4162-6/20/$31.00 ©2020 IEEE

with information that is useful in the execution of their jobs and interest. For example, policy making. Broadly speaking, sexual violence is one of the forms of GBV, which is described as a term that is usually used to capture violence that arises as a consequence of normative role expectations connected with each gender, along with unequal power relationships that are forcefully expressed mostly by a male on a female [6].

As an electronic medium of communication, social media can be used to identify widely discussed topics on any subject. Twitter, for example, can provide an overview of hot topics and issues [7]. Traditionally, offline analysis-related processes are recommended only when data collection is a one-time process. The trends of research as of now demand highly sophisticated real-time data analysis tools to tackle a range of challenges, especially the efficiency of non-stop streaming of data collection within the context of a NoSQL and relational database. As earlier stated in this paper, we introduced sentiment analysis through the implementation of Node.js as a basis to evaluate gender disparity between reported tweets on sexual violence in South Africa. Moreover, we implemented a DNNClassifier to predict the gender of a Twitter user. Extracting Twitter data on sexual violence tweets was possible through the Twitter streaming API which enabled us to develop a real-time and interactive system. ES was used as a storage repository and search engine, while Kibana (a plug-in to ES) facilitated intuitive dashboard creation and real-time data analysis.

A. Problem Identification

To our knowledge, the issue of sexual violence in South Africa has existed both physically and digitally for a long time, yet no effective solution in literatures provided a data-driven approach to help address this issue. By analysing and visualising tweets on sexual violence in real time, stakeholders such as law enforcement agencies can better decide where resources should be allocated as well as policies that will serve as counter-preventive measures. In addition, the development of an application that collects, stores, identifies and analyses this data is a pointer to pattern recognition of actionable intelligence. In particular, being able to infer demographic attributes and analyse sentiment from related tweets in the context of sexual violence would help isolate and mitigate the issue.

B. Motivation and Rationale

Registered users use the Twitter platform to share information and draw attention to issues that are important. Sexual violence is rampant in South Africa and as such, the issue needs serious attention [8]. Through tweets on platforms such as Twitter, sexually violated victims and those generally affected by it, often times make users and the general public aware of some of the ordeals they have encountered which pertain to GBV. Apart from giving us a perspective on what sexual violence means to different genders, inferring insights from such tweets can help in identifying some of the causes of sexual violence, the reasons why people do not speak out when sexually violated and may even shed some light as to why some individuals commit these atrocious crimes. Extracting and analysing such data could help in developing laws, which cannot be permeated by certain loopholes, and consequently, policies which help in the fight against GBV by protecting the vulnerable and abused as well as prosecuting perpetrators. Techniques and algorithms are therefore required to accurately infer such hidden user attributes from Twitter data or metadata in order to accurately classify tweets from a user demographic point of view. In particular, we try to infer gender as a demographic and hidden attribute through the invstigation of labelled data sets found online (Kaggle) using an Artificial Intelligence (AI) approach built of DNNs and TF Estimator. We understand that leveraging a Web-based services and design such as ES and Kibana would provide us an opportunity to see the prevalence of discussions around sexual violence on Twitter in real-time. Furthermore, we try to investigate the sentiment analysis as an indicator in evaluating gender disparity using AFINN model of English words which have been rated for valence by summing up the score of each word in a pool of text.

This study aims to make the following contributions;

1) provide a Web-based application, which attempts to assist on the shared frustrations expressed through tweets on GBV.
2) a distributive visual experience to present cases of sexual violence within cities in South Africa.
3) generated sentiment analysis. This would assist policy makers to make more informed decisions as per resource allocation to help solve the issue.

The roadmap for the rest of this paper is described as stated: In Section **II**, we presented the literature of related works conducted in this area of work. Sections **III** and **IV** presented the methods and implemented our proposed approach using Node.js. Further, we showed the results from evaluating model performance through a confusion matrix. Finally, in Sections **V** to **VI**, we discussed the results, and concluded with advise on future work.

II. Related Work

A. Sexual Violence

The violations of human rights are noted to be perpetrated against men and women [9]. The aftermath of these violations, however, varies according to the victim's gender. Studies such as [10] have shown that the aggressive acts carried out on women differ, exhibiting certain traits, and therefore warrant being classified. This asserts that violence in this form can be linked directly to the unfortunate circumstance of power being distributed unequally. In addition to this, the devaluation of women and the subordination to men is due to the asymmetric relationship between men and women in society. The notion that women are the weaker sex is what separates GBV from other violent forms of aggression and coercion.

Power relations between men and women in South Africa are often characterized through sexual violence and assault [11]. Violence against women is one of the highest forms of GBV in all provinces in South Africa [12]. Following the 1993 World Conference on Human Rights and the Declaration on the Elimination of Violence against Women, violence against women has been declared a major concern for public health, social policy, and human rights. It has been difficult to document the magnitude of violence against women as well as to produce reliable comparative data which can be used to guide policy and monitor progress in the fight against violence against women. We hope that our proposed system would serve a contributory purpose enough to assist policy makers with strategy formulation.

B. Sentiment Analysis

Bakliwal *et. al.* [13] introduced a method for finding opinions in a clustered group of tweets. With suggestive words being labeled as noisy, a list containing such words was used as one dataset while the other was built using emoticons. A method used for scoring polarity was proposed. Emphasis was given to types and levels of preprocessing which are necessary for good performance. The Naive Bayes method was utilised for the determination of the token polarity in tweets, this method performed well on both datasets, moreover, their process of mining sentiments from tweets was thorough. They used both the Stanford and Mejaj's dataset.

Pak and Paroubek [14] worked on a corpus of data (Twitter data) for analysing sentiment and opinion which resulted in the performance of linguistic analysis on the collected corpus. They were able to develop a sentiment classifier that was built to determine positive, negative and neutral sentiments. They used their proposed technique on English language and recommended it could be used with any other language. Their technique was observed to have performed more efficiently than previously proposed methods.

For this paper, we used a word list model which is benchmarked on [15]. This model is polarity-oriented and involves a pre-assigned list of words. Each word gets a positive or negative valence score, 5 being very positive and -5 very negative regardless of language.

C. Gender

This research study is driven by its positive impact in analysing sexual violence related cases from tweets in South Africa through the exploration of deep learning (DL) techniques and the application of web services for social good. In relation to Twitter data, we have seen from various literature reviews the application of DL techniques by [16], [17], [18], [19] and [20]. However, none of these stated reviews reflect issues similar to the South Africa context in terms of inferring gender from sexual violence tweets and the method used. The concept could be said to be the same in terms of its "deep learning" approach, but that is as far as the similarity goes. As a result, we look at a related work on gender prediction.

[21] used perceptron and Naive Bayes algorithms to identify the gender of Twitter users from 1 to 5 grams in tweets. Their work used streaming algorithms and made substantial use of gender prediction to manage speed and measure of tweet traffic. Features such as R-gram were implemented to represent streaming tweets because informal text (e.g. tweets) was not easily assessed using traditional dictionary methods. Multiple selection methods were used to select informative n-gram features as a large number of 1 to 5 grams requires only one subset to be used in the classification of gender. The Naive Bayes and Perceptron algorithms were 99% accurate. Six selection algorithms were implemented in extracting informative features and to improve the classification and run time of their gender prediction approach. The approach initiated through the stream mining perceptron and Naive Bayes performed relatively well to evaluate how effective these selected gender identification features on Twitter were. The perceptron functioned reasonably well with a very high precision of 97% and a balanced accuracy of 94% (this was outperformed by Naive Bayes with a score between 90% and 100% accuracy for all metrics). It is worth mentioning that perceived results from works of literature reviewed on gender relative to the accuracies achieved appeared to be results that were not from a partitioned sample.

III. Implementation, Methodology and Approach

In any case, deep learning systems are built to work with numerical values as input features. Therefore, tweets are represented in a way that reflects this constraint. In tackling this constraint in terms of text representation, we used word embedding as against the likes of bucketization, scaling, crossing features to convert tweets to Tensors of vector matrix before fitting to the model. Embeddings translate large sparse vectors into a lower-dimensional space that preserves semantic relationships. This is presented in Figure 1 below.

A. DNN Architecture

This study sets to describe our input in mining Twitter data for social good through the pipeline in Figure 1. Essentially, we applied mechanisms inspired by AI and driven by Web tools to design a streaming service on sexual violence tweets. We trained a DNNClassifier with 2 hidden layers to predict gender from labeled tweets on gender classification. The dataset used in training the gender predictor model was retrieved from Kaggle through the "Data for Everyone Library On Crowdflower" and is always available for the community, free of charge. Contributors to the project were asked to mainly analyse a Twitter profile and determine if the user was a male, a female or a brand (non-individual). The dataset presents 20,000 rows and columns, each with a username, a random tweet, account profile with image, location, link and sidebar colour. However, we mainly focused our attention on the subset of the features (tweet and profile description). The DNN is a 3-layer NN with two hidden layers of 500, 150 neurons respectively and an output layer with 3 neurons which represents the different gender classes (male and female) and brand (non-individuals).

B. Model Evaluation and Results

The model was evaluated to check the model's correctness during training. The data used was the data left off of parti-

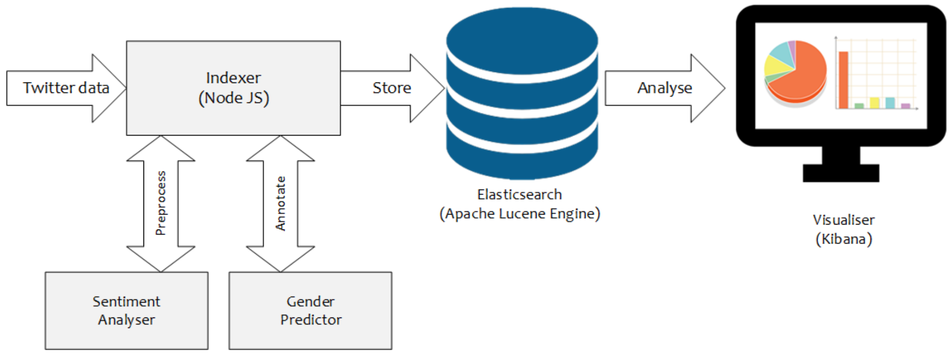

Fig. 1. Research pipeline showing flow of methods and processes

tioning before training called the testing set which was benchmarked around 20% from the whole set. This data was used for testing and evaluating the model's performance. We filtered the gender variable (male, female, brands and unknown) to align with our classification objective of tweets from only male and female. We used TensorBoard and confusion matrix to evaluate the model performance. The confusion matrix enables the percentage of properly and wrongly labeled instances to be visualised. Figure 2 demonstrates the bias of our classifier, and whether it makes sense to distribute labels. This is an expected performance from the classifier as tweets from both gender could have been misconstrued to have come from a male assuming it was from a female. Ideally, the biggest estimate of prediction should be spread across the diagonal. With regard to the analysis on the sentiment of a `GBV` reported instance on Twitter, A positive tweet is such a tweet that throws a positive sentiment following the analysis of all its expressions through tokenisation.

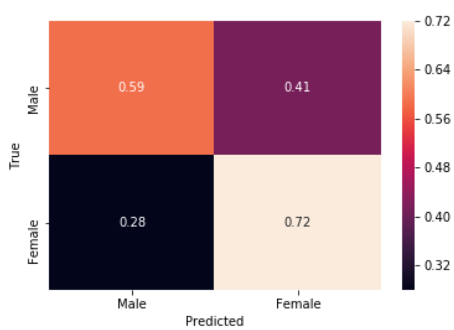

Fig. 2. Confusion matrix based on test data

Experimentally, we further evaluated the model's performance on gender prediction through TF Estimator. We connected the model's directory to TensorBoard to visualise the model's internals for analysis on accuracy and loss before it was passed into production through TensorFlow Serving. The loss curve decreased over a pre-defined number of steps. This pattern confirmed that the model generalised well on out-of-sample data when tested. The accuracy (68%) on the test set returned a good evaluation on the model from parameter and hyperparameter tuning. This is an expected condition for L1 & L2 regularisation.

We like to clarify that the Kaggle dataset used for training is not from South Africa as compared to the tweets currently streaming through `ES` and `Kibana` Web services. Hence, the variation in linguistic style which could have biased the classifier as shown in Figure 2.

C. Sentiment Analyser

The overall contextual clarity of a tweet document informed our choice to develop an application called Sentiment Analyser to help understand sentiment analysis within the context of this paper. The idea from this development will enable us to identify sexual violence tweets as an actual instance of sexual violence expressed by a user. Furthermore, It gives us an idea on the measure of disparity expressed by a male and a female on a tweet related to `GBV`. It runs on the `Node.js` runtime and was developed primarily in JavaScript for this paper. We built the application using the current version of the AFINN word list which contained more than 3,300+ words with a polarity score associated with each word. The background around this approach was clarified in Section I.

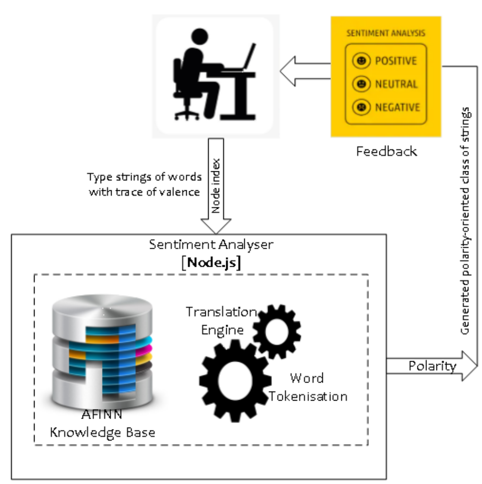

Fig. 3. Framework for sentiment analysis

We determined that a tweet dealt with `GBV` if it was negatively expressed on Twitter addressing an instance of a negative sentiment as measured from experiments in this paper.

IV. Evaluating Results from Kibana

As a possible limitation, the model gets to predict the gender from strings of text (the user's description and historical tweets from Twitter) it was fitted as an out-of-sample data. However, the system can be manipulated from changes in these key features which may result in a bias prediction. Then, this paper did not consider that. Figure 1 shows the `Indexer` serving as a hub of the system. The `Indexer` is a `Node.js` server-side application whose role is to connect to Twitter's streaming API and ingest tweets in real-time. At a high-level, the system will request data through Twitter API and then the Twitter API will respond back with requested data. Overall, the `Indexer` pre-processes and annotates real-time sexual violence tweets to derive sentiments and gender through user description and historical tweet.

Figure 4 below is a pie chart showing the percentage of tweets on sexual violence contributed by each gender class. 61.17% of the tweets were from females with 38.83% emanating from the male gender. These numbers showed a huge disparity between the genders tweeting about sexual violence for reasons which would have been based on sentiments. We addressed possible reasons for this in Section I.

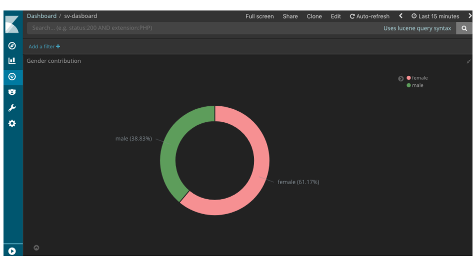

Fig. 4. A Pie chart showing contributions from male and female

In Figure 5, we present a sentiment heatmap as a graphical representation of data represented by a map from our system. Data is displayed in different shades of colours, the thickest colour representing the most prominent element and the lowest shade representing the least popular element. From our heatmap, we can see that the most prominent tweets were negative ones tweeted by females. Following this pattern, the least prominent tweets were those with a positive context which were tweeted by males.

Table I further suggested that a large proportion of female reported more on cases related to `GBV` by an estimate of 72.74% through tweets in comparison to their male counterpart who only reported 65.57% as stated earlier.

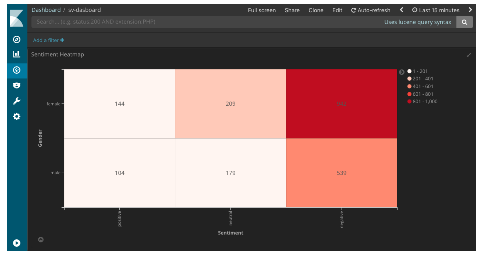

Fig. 5. A heat map showing sentiment count

TABLE I
Sentiment count from both gender classes

Gender	Sentiment	Estimated value (%)	Total sample (count)
Female	Positive	11.12	1295
	Neutral	16.14	
	Negative	72.74	
Gender	**Sentiment**	**Estimated value (%)**	**Total sample (count)**
Male	Positive	12.65	822
	Neutral	21.78	
	Negative	65.57	

This statement through the predicted labels of both gender classes showed that females expressed a negative sentiment of 72.74% versus males who showed a negative sentiment of 65.57%. This result reflects a good validation of our experiment as we would expect more negative sentiments from women than men, since women are disproportionately affected by sexual violence. These evaluations of disparity on sexual violence tweets by either gender are not an artefact of the system, but that these numbers make sense in the real-world because women are more negatively impacted by sexual violence. So, it is expected to have more comments that reflects negative sentiments from women. We attempt to clarify that the Twitter data powered through the Kibana visual functionality, has been streaming tweets on sexual violence within South Africa (since 2018 to date capturing each case of reported `GBV` related instances).

V. Discussion

From a pool of unstructured information found on Twitter, sexual violence tweets were chosen as a representative subject. The results presented in this paper have revealed that, by using an investigative approach and leveraging on deep neural network libraries coupled with the right processing of data, hidden attributes such as gender and sentiments could be inferred from any form of data. As expected, this information about sexual violence often spark insights about people's perceptions and the demographics affected. Consequently, the information about whether the people tweeting are victims or not became irrelevant in this instance. Thus, the data collected is still a valid social indicator because the effects of sexual violence cut across all societal boundaries. According to [22], this is useful piece of information from a sociological point of view in that the report of these cases on Twitter is not only valid if it comes from a victim but that it can equally be valid if it is comes from a friend of a victim or family of a victim.

VI. Conclusion and Future Work

The Estimator as a high-level API within TF module exposes a wide variety of ML algorithms through the integration of state-of-the-art ML models for large scale supervised or unsupervised problems. We have used this framework as a task-oriented interface in simplifying a unique use case in ML for data abstractions on gender classification. The Estimator simplifies ML applications through the usage of general-purpose high-level language as building blocks for approaches specific to any level of use case. For example, in a situation like medical imaging [23]. Emphasis is on simplicity of implementation, performance, documentation, and API consistency, encouraging its use in an academic space.

When working with a highly fluid, fast-moving domain like Twitter - populated by users who may unite around a topic and engage in volatile communication, it is crucial that we are thorough about the extent of insight to which extracting and analysing its data would reflect. Moreover, there are very few avenues for getting data collection about sexual violence; there is under reporting and most people are hesitant to share details of their ordeal with the authorities. However, people are more comfortable sharing on Twitter. In addition, this study has yielded interesting results, one of which is the disparity on negative sentiment expressed by the female gender class when reporting about tweets on sexual violence. This is a discovery that has scaled beyond the computer science academic frame but is a piece of information that would be of great use to a sociologist. Therefore, we are fundamentally producing a new database that will contribute to future studies that want to expand on this subject or related issues.

Acknowledgement

This study was funded under the Postgraduate Merit Award by the Financial Aid & Scholarships Office at the University of the Witwatersrand.

References

[1] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, 2016, pp. 265–283.

[2] F. Å. Nielsen, "A new anew: Evaluation of a word list for sentiment analysis in microblogs," *arXiv preprint arXiv:1103.2903*, 2011.

[3] K. Z. Aung and N. N. Myo, "Sentiment analysis of students' comment using lexicon based approach," in *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*. IEEE, 2017, pp. 149–154.

[4] C. Fink, J. Kopecky, and M. Morawski, "Inferring gender from the content of tweets: A region specific example." in *ICWSM*, 2012.

[5] A. Mislove, S. Lehmann, Y.-Y. Ahn, J.-P. Onnela, and J. N. Rosenquist, "Understanding the demographics of twitter users." *ICWSM*, vol. 11, no. 5th, p. 25, 2011.

[6] S. S. Bloom, "Violence against women and girls: a compendium of monitoring and evaluation indicators." 2008.

[7] J. Benhardus and J. Kalita, "Streaming trend detection in twitter," *International Journal of Web Based Communities*, vol. 9, no. 1, pp. 122–139, 2013.

[8] K. L. Dunkle, R. K. Jewkes, H. C. Brown, G. E. Gray, J. A. McIntryre, and S. D. Harlow, "Gender-based violence, relationship power, and risk of hiv infection in women attending antenatal clinics in south africa," *The lancet*, vol. 363, no. 9419, pp. 1415–1421, 2004.

[9] M. R. Decker, A.-L. Crago, S. K. Chu, S. G. Sherman, M. S. Seshu, K. Buthelezi, M. Dhaliwal, and C. Beyrer, "Human rights violations against sex workers: burden and effect on hiv," *The Lancet*, vol. 385, no. 9963, pp. 186–199, 2015.

[10] S. A. Basow, K. F. Cahill, J. E. Phelan, K. Longshore, and A. McGillicuddy-DeLisi, "Perceptions of relational and physical aggression among college students: Effects of gender of perpetrator, target, and perceiver," *Psychology of Women Quarterly*, vol. 31, no. 1, pp. 85–95, 2007.

[11] K. Wood and R. Jewkes, "Violence, rape, and sexual coercion: Ever yday love in a south african township," *Gender & Development*, vol. 5, no. 2, pp. 41–46, 1997.

[12] W. H. Organization *et al.*, "Violence against women= la violence contre les femmes," 1997.

[13] A. Bakliwal, P. Arora, S. Madhappan, N. Kapre, M. Singh, and V. Varma, "Mining sentiments from tweets," in *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*, 2012, pp. 11–18.

[14] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining." in *LREc*, vol. 10, no. 2010, 2010, pp. 1320–1326.

[15] L. K. Hansen, A. Arvidsson, F. Å. Nielsen, E. Colleoni, and M. Etter, "Good friends, bad news-affect and virality in twitter," in *Future information technology*. Springer, 2011, pp. 34–43.

[16] P. Badjatiya, S. Gupta, M. Gupta, and V. Varma, "Deep learning for hate speech detection in tweets," in *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 2017, pp. 759–760.

[17] D. T. Nguyen, S. Joty, M. Imran, H. Sajjad, and P. Mitra, "Applications of online deep learning for crisis response using social media information," *arXiv preprint arXiv:1610.01030*, 2016.

[18] L. Heck and H. Huang, "Deep learning of knowledge graph embeddings for semantic parsing of twitter dialogs," in *2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2014, pp. 597–601.

[19] A. Severyn and A. Moschitti, "Twitter sentiment analysis with deep convolutional neural networks," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2015, pp. 959–962.

[20] S. Yuan, X. Wu, and Y. Xiang, "A two phase deep learning model for identifying discrimination from tweets." in *EDBT*, 2016, pp. 696–697.

[21] Z. Miller, B. Dickinson, and W. Hu, "Gender prediction on twitter using stream algorithms with n-gram character features," *International Journal of Intelligence Science*, vol. 2, no. 04, p. 143, 2012.

[22] M. Stubbs-Richardson, N. E. Rader, and A. G. Cosby, "Tweeting rape culture: Examining portrayals of victim blaming in discussions of sexual assault cases on twitter," *Feminism & Psychology*, vol. 28, no. 1, pp. 90–108, 2018.

[23] V. Michel, A. Gramfort, G. Varoquaux, E. Eger, C. Keribin, and B. Thirion, "A supervised clustering approach for fmri-based inference of brain states," *Pattern Recognition*, vol. 45, no. 6, pp. 2041–2049, 2012.