A home is one's safe haven.

NOT EVERY HOUSE IS A HOME

IF YOU KNOW ANYONE SUFFERING DOMESTIC ABUSE, REACH OUT.

GROUP A.

NATURAL LANGUAGE PROCESSING & TEXT MINING

Table of Contents

1   Introduction

Domestic Violence is not a pandemic, it's an epidemic. With Covid-19 ravaging the economy; increasing unemployment such crises are set to become much more frequent. Add another public health crisis to the toll of the new coronavirus: Mounting data suggests that domestic abuse is acting as an opportunistic infection, flourishing in the conditions created by the pandemic.

In France, reports of domestic violence have increased by 30%, In Spain, the emergency number for domestic violence received 18 per cent more calls in the first two weeks of lockdown than in the same period a month earlier

The United Nations Secretary-General, Antonio Guterres called for urgent action to combat the worldwide surge in domestic violence in saying, "I urge all governments to put women's safety first as they respond to the pandemic."

This is because domestic violence is an ongoing serious social problem worldwide (Xue et al. 2018). It is estimated that one-third of women worldwide have experienced some form of domestic violence by their intimate partner in their lifetime (WHO 2017).

As this problem has serious precautions on the physical and mental health of the victims (children, woman, man), a lot of research has been already conducted to analyse the nature of this social problem from interviews with victims, and official data like health records. Recently there has been significant effort put into social media data research. However, much of this research found has many limitations to create practical tools to deal with domestic violence. Within this project, we are looking utilizing topic modelling to assist with domestic violence data classification.

2   Short Literature Review

2.1   Topic Modelling

Topic modelling is a text mining technique which provides methods for identifying co-occurring keywords to summarize large collections of textual information. It helps in discovering hidden topics in the document, annotate the documents with these topics, and organize a large amount of unstructured data.

It should not be confused with Text classification which is a supervised machine learning problem, where a text document or article is classified into a pre-defined set of classes. Whereas topic modelling is the process of discovering groups of co-occurring words in text documents. These group co-occurring related words makes "topics". It is a form of unsupervised learning, so the set of possible topics are unknown. Thus, topic modelling can be used to solve the text classification problem. Topic modelling will identify the topics presents in a document" while text classification classifies the text into a single class.

There are several topic modelling techniques:

Latent Semantic Analysis (LSA) uses bag of word model, which results in a term-document matrix (occurrence of terms in a document). The rows represent terms and columns represent documents. LSA learns latent topics by performing a matrix decomposition on the document-term matrix using Singular value decomposition. As a result, LSA is typically used for dimension reduction or noise reducing.

Hierarchical Dirichlet Process (HDP), is a nonparametric Bayesian model for clustering problems involving multiple groups of data (Nathan and Scobell, 2012). Each group of data is modelled with a mixture, with the number of components being open-ended and inferred automatically by the model. Further, components can be shared across groups, allowing dependencies across groups to be modelled effectively as well as conferring generalization to new

Latent Dirichlet Allocation (LDA) is a generative probabilistic model for collections of discrete data such as a text corpus (Xue, Chen and Gelles, 2019). Moreover, LDA is a three-level hierarchical Bayesian model, in which each item of a collection is modelled as a finite mixture over an underlying set of topics (Subramani and O'Connor, 2018). Each topic is, in turn, modelled as an infinite mixture over an underlying set of topic probabilities. In the context of text modelling, the topic probabilities provide an explicit representation of a document.

There several uses cases where topic modelling has gained success:

- CV Summarization: It helps recruiters to evaluate CVs in a quick and efficient manner by reducing the effort in filtering many CVs.
- Search Engine Optimization: online articles, blogs, and documents can be tagged easily by identifying the topics and associated keywords, which can improve optimize search results. An example of this is the use of topic modelling by the New York Times to sort articles

- Recommender System Optimization: in this case recommender systems act as an information filter and advisor according to the user profile and previous history. It also helps to discover unvisited relevant content based on past visits.
- Improving Customer Support: Discovering relevant topics and associated keywords in customer complaints and feedback for examples product and service specifications, department, and branch details. Such information help company to directly rotated the complaint in the respective department.
- In the healthcare industry, topic modelling is used to extract useful and valuable information from unstructured medical reports. This information can be used for the patient's treatment and medical science research purpose. In the field of domestic violence identification/classification on online media, the development of NLP tools and techniques to effectively classify domestic violence intent/actions is still in its infancy.

(Subramani, Vu and Wang, 2018) explored the use intention mining as a binary classification problem (abuse or advice) with the use- case of abuse discourse. To address this problem, we extract rich feature sets from the raw corpus, using psycholinguistic clues and textual features by term-class interaction method. Machine learning algorithms were used to predict the accuracy of the classifiers between two different feature sets.

Their experimental results with high classification accuracy gave a promising solution to understand a big social problem through big social media and its use in serving information needs of various community welfare organizations. However, they observed two key challenges with intent classification on their Domestic Violence discourse. First, informal language use in short-text messages creates ambiguity to interpret user expressions and thus weakening term-class relationships. Second, sparsity of instances of specific intent classes in the corpus creates data imbalance.

(Subramani et al., 2019) investigated automated domestic violence crisis identification from Facebook posts based on deep learning.

2.2   Our Purpose

The online content generation has grown rapidly in scale. The unstructured and noisy character of such data has further added to an overall complexity of processing and utilizing available information. The mix of messages that are of personal nature in a form of mere opinions or empathetic thoughts along with general awareness promotions have greatly diminished the DVCS services efficacy.

With topic modelling, we are looking to identify the different topics or classes of the tweets or comments in reddit platform to make the large and unstructured data more organized in a way that will make it easier for NGOs, government officials or researchers to assess and get useful insights from to better analyze this crisis and come up with better courses of actions accordingly.

At present, there is no available benchmark dataset for domestic violence with pre-labeled data for multiclass classification, so it is a very manual intensive job. Topic modelling could help identify the different topics so it is easier to build a corpus later with labeled data that can serve as a training dataset for deep learning algorithms like CNNs and RNNs for classifying domestic violence texts.

2.3 Scope of Application

The findings from twitter and reddit data can be evaluated against the official reports regarding the DV prevalence and severity. The insights derived can be compared against official reports and official news about domestic violence.

Demographics Investigation: The metadata associated with the posts extracted from online forums can be a valuable source of information on users' demographics. Such knowledge can further enrich the findings, and lead to interesting insights on gender groups most likely to become either victims, or the abusers.

Geo-location Correlation: The geo-attributes available with Twitter data can be utilized for further study of the Abuse Types and Health Issues patterns across various geographical locations. The spatial information additionally allows for the correlation with regional development indicators (e.g. poverty rate, population density, education level etc.), to uncover potential inter-dependencies, and improve the prediction models.

3   Tool design

This is the architecture for our approach. Our platform will be like a repository that contains the results and insights from our sentiment analysis and topic modelling techniques to help develop better research around domestic violence. Also this platform can provide great insights for NGOs and government officials to help them with their policy making and strategies to help prevent and detect and fight against domestic violence.
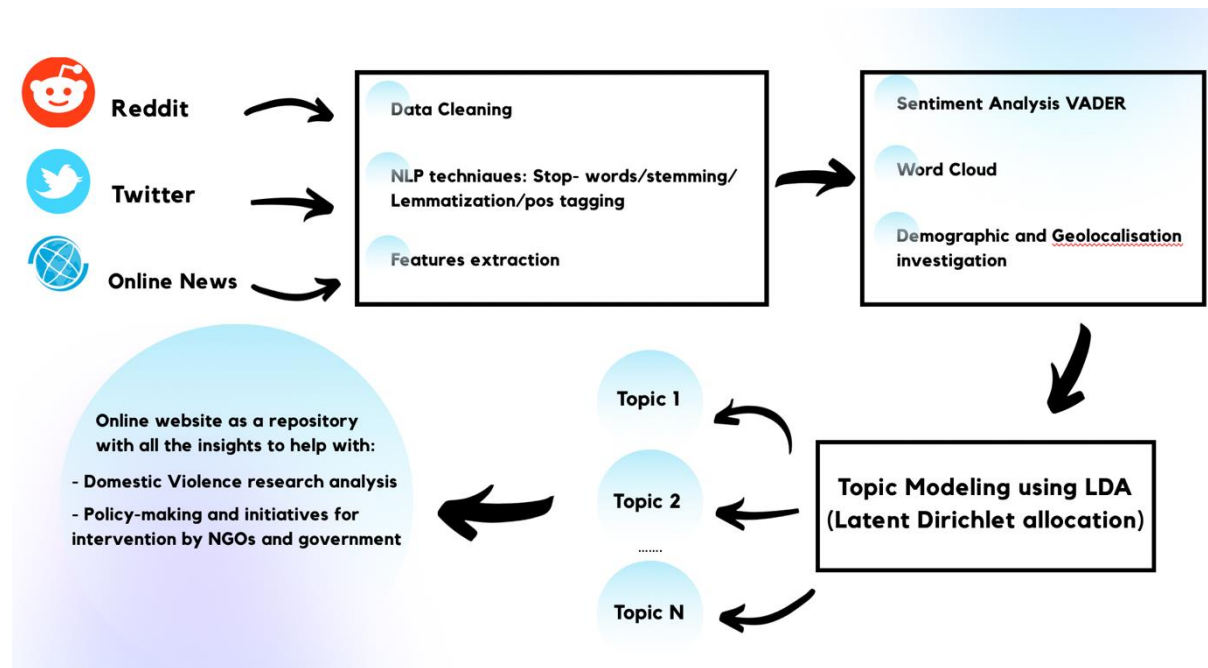
4   Modelling Methodology

4.1   Methodology

Data collection

Data is scraped online on twitter, reddit and from online news articles. Specific hashtags

Hashtags used for every file for twitter:

- 1 = METOO
- 2 = WHYISTAYED WHYILEFT
- 3 = HeForSheAtHome WomenCount GenerationEquality AntiDomesticViolenceDuringEpidemic Mask-19 WithHer SpotlightEndViolence
- 4 = staysafe domesticviolence DomesticAbuse DomesticViolence

4.2   Packages

VaDER

To perform sentiment analysis, we used

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. VADER uses a combination of A sentiment lexicon is a list of lexical features (e.g., words) which are generally labelled according to their semantic orientation as either positive or negative.

We decided to use ADER since it has been found to be quite successful when dealing with social media texts, NY Times editorials, movie reviews, and product reviews. This is because VADER not only tells about the Positivity and Negativity score but also tells us about how positive or negative a sentiment is.

Gensim and pyLDAvis

We used these packages in order to perform our topic modelling and pyLDAvis helped with creating the visuals of the clusters of topics we found. Those visuals are saved as HTML files and attached with this report.

4.3   Evaluation Metrics

Perplexity

This is a measure that indicates how "surprised" the algorithm is to see a term within a given topic (lower values indicating a better model) and is measured as the normalized log-likelihood of a held-out test set (Mishra, 2018). However, numerous authors have suggested that it is an ineffective metric because it is does not capture semantic information. For our analysis it will be calculated, but the output will not be used to evaluate the model.

Topic coherence

Topic Coherence measures score a single topic by measuring the degree of semantic similarity between high scoring words in the topic (Mishra, 2018). These measurements help distinguish between topics that are semantically interpretable topics and topics that are artefacts of statistical inference. This is a measure composed of four steps namely, segmentation of words subsets, probability estimation, confirmation measure then aggregation. Moreover, the measure it is calculated between 0 and 1, where 1 is the most coherent.

5   Findings
5.1   Reddit

Through all the comments, mainly women are seeking help, sharing their abuse stories, how they survived and how they want to help other people.

One of the main insights is that the most important topic is about family abuse, not relationship abuse, mentioning mother, father, relationship and child as well as kid, we assume that these people are young adults or teenagers living with their families.

Time is also a concept that is frequent among all topics, quoting "year", "month", "day" as important words describing the testimonies.

5.1.1   Recommendations
- Organizations and NGO should have initiatives and communication through Reddit since it is a space in which people share their testimonies, is a channel where the initiatives can be communicated to the right people.
- Partner abuse is highly related to partner abuse.

5.2   Twitter: All Data

Through all the tweets the main conversation is from organizations and movements to invite people to join the movements through hashtags and facts, we can also perceive not only organizations but sponsors.

Below are the relevant inferred topics:

Topic 1: #WHYILEFT, #METOO, #WITHHER - This movement revolves around raising awareness and acting with the organization - ipu parliament (global organization of national parliaments to promote peace, democracy and sustainability), also the sponsorship of women in sport.

Topic 2: #METOO, #believewoman - Victims sharing their stories and support of one another and sharing their survivor stories, organization - ipu parliament (global organization of national parliaments to promote peace, democracy and sustainability), also the sponsorship of women in sport.

Topic 3: #genderequality - Feminism and equality - Backed by the organization "everyday equality" to drive equal social opportunities in UK.

Topic 4: timesup - Campaign to report and end violence thus creating an equal future, promoting also job quality and activism.

Topic 5: Abuse regarding rape in family and work – organization like Orange the World: Generation Equality Stands against Rape.

Topic 6: Promotes discussion of psychological abuse, harassment and solution making.

Our final insights

On Twitter is about joining the conversation and promoting the initiatives and campaigns, identified movements:#whyistayed, #metoo,#timesup,#heforshe,#iwillgoout,#yesallwomen,#orangetheworld,#niunamenos ,#everydaysexism,#womenshould

5.2.1   Recommendations

An improvement that can be done in twitter is not only promote de initiatives and join the conversation but invite victims to share their story via Reddit, because there is no character limit, thus increasing Reddit´s testimonies and having more information regarding this subject.

5.3   Twitter Data: Top 5 Countries

After identifying the top 5 countries that are most engaging in the conversation about domestic violence on twitter, we wanted to further identify and analyze the topics in each of those countries using topic modeling.

India

The clusters of topics when filtering only on India are diverse:
   • Topic 1: The first topic is about gender equality and the laws against domestic violence. there is a mention of Prasad who is an Indian lawyer who recently wrote about male domestic violence act.
This topic is basically about having sometimes biased laws against domestic violence.
   • Topic 2: This topic is about assistance and help and society support.
Topic 3: The 3rd topic is about the pandemic and lockdown and how it is increasing the problem of DV.
   • Topic 4: same as topic 6 they both relate organizations, initiates, and government actions like
UNICEF, ncwindia (National Committee for Woman), pmoindia (prime minister of India who created a care fun for emergency and distress situations)
   • Topic 5 relates to topic 1 in the sense that it also focuses on "make home safe for man" as well as woman.

United Kingdom

UK topic are not only about organizations, but the topics revolve around mainly domestic abuse, trying to aid women, in the second topic inviting women to call immediately during COVID.

1. This topic is about the lockdown and emergency order by police.
2. Covid19 help: Aiding women to call during the actual crisis and trying to aid women.
3. Charities
4. community support and sharing stories of victims.
5. Aid during COVID promoting a helpline and communicating that support service remains available.
6. Same as topic 1 and 2

United States of America

- US Is the News country, where is more about the actual sharing of news and the story of the movement. With Joe Biden recent abuse scandal, the first 5 topics revolve around it, also there is a mention of "Alyssa Milano" American Actress that started the "metoo movement" and is in support of Biden.

- The only "Corona Lockdown mention" is "Covid" and #Stayathome, but there is no communication towards women or stories baout women in any of the hashtags, the only identified movement is #metoo and #believing_woman but realted to Joe Biden and Senator Tara.

Kenya

Topics clustered from filtering on Kenya are oriented towards gender-based violence and gender equality, and also initiates related to DV:
1. This topic is about gender equality in the case of domestic violence.
2. 2nd topic is about UNWoman organization and support to domestic violence victims.
3. This topic is about the domestic abuse and education
4. This is about initiatives related to DV like #withher that talks about creating a world where woman and girls are safe, and eachforequal movement that was created in international woman day 2020.
5. same as topic 6 it is about GBV(gender based violence) by the United Nations Population Fund (UNFPA) which is something Kenya has recently been shinning a light on.

Nigeria
Topics identified in Nigeria are as follows:

1. This topic is about movements like idefendher_heforshe and NGOs in Nigeria like Ceen foundation with the mission of promoting public safety, security and accessible justice.
2. The 2nd topic is about community sharing and support for victims.
3. COVID19 pandemic and lockdown.
4. This topic is about child marriage about domestic abuse.
5. Gender equality and UNWoman
6. Organizations and movements to provide aid and support to woman who suffer fromm DV during this pandemic like
@GlobalSpotlight that support women's organizations that are working to prevent and address violence against women.

All of these insights are organized in the website we created:

https://bfdelavega.wixsite.com/misitio

6 Limitations

One of the challenges with news articles is that unless it's opinionated, there is little room for a diverse identification of topics. But our results generated topics which had great separation.

But the noteworthy topics generated:

Topic 1: Children's services -> words most prevalent in the topic were "victim", "child", "resource", "home" and "risk". By identifying the percentage contribution of the most dominant document (see code), main stories are about the lack thereof services.

Topic 2: Survivor stories -> words most prevalent in the topic were "woman", "household", "victim", "survivor" and "impact".

Topic 3: Police response -> words most prevalent in the topic were "city", "police", "victim", "family" and "shelter".

As started previously, there were more topics generated but clearly identifying with the limited number of articles generated (only 26 articles), created a barrier to adequately class the topics. For future work, one can enrich the data with more news articles.

For visualizations, refer to News_Articles_Analysis.ipynb

7 Future Work

Based on the conclusions, the following recommendations for future work were made:

- Explore the use of GuidedLDA to pre-seed the data words which frequently occur.
- Increase data collected from online news articles and to further enrich corpus to possibly help with clearly identifying the topics.
- Diversify the datasources and media platform and improve the preprocessing techniques applied across the data collections
- Deal with the bias of creating our dataset possibly including more random sample data
- Incorporate temporal topic modelling to see changes of the topics over time.

8 References

Mishra, A. (2018) *Towards data science*, *Metrics to Evaluate your Machine Learning Algorithm*. Available at: https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234.

Nathan, A. J. and Scobell, A. (2012) 'How China sees America', *Foreign Affairs*, 91(5). doi: 10.1017/CBO9781107415324.004.

Subramani, S. *et al.* (2019) 'Deep Learning for Multi-Class Identification from Domestic Violence Online Posts', *IEEE Access*, 7, pp. 46210–46224. doi: 10.1109/ACCESS.2019.2908827.

Subramani, S. and O'Connor, M. (2018) 'Extracting Actionable Knowledge from Domestic Violence Discourses on Social Media', *ICST Transactions on Scalable Information Systems*, 5(17), p. 154807. doi: 10.4108/eai.29-5-2018.154807.

Subramani, S., Vu, H. Q. and Wang, H. (2018) 'Intent Classification Using Feature Sets for Domestic Violence Discourse on Social Media', *Proceedings - 2017 4th Asia-Pacific World Congress on Computer Science and Engineering, APWC on CSE 2017*, pp. 129–136. doi: 10.1109/APWConCSE.2017.00030.

Xue, J., Chen, J. and Gelles, R. (2019) 'Using Data Mining Techniques to Examine Domestic Violence Topics on Twitter', *Violence and Gender*, 6(2), pp. 105–114. doi: 10.1089/vio.2017.0066.