Combat Sexual Abuse via Machine Learning

Vyoma Patel[1], Bhavin Gor[2], Preyanshu Sukhadia[3], Prasham Mody[4]

[1]vyoma.p@ahduni.edu.in
[2]bhavin.g@ahduni.edu.in
[3]preyanshu.s@ahduni.edu.in
[4]prasham.m@ahduni.edu.in

***Abstract*— One of the most sensitive and major concerns across globe-Women Safety. The early identification of the crucial issue of sexual harassment is vital for mitigating the risk. This can certainly be worked upon using Machine Learning techniques. The model aims to use classification algorithms based on abusing categories with the prediction of most prevailing locations. This data can be helpful for the government for taking necessary actions, safeguarding the nation and also alert females from such unavoidable incidents.**

***Keywords*— sexual harassment | machine learning | classification | analysis | prevention | violence**

I. INTRODUCTION

The proposed project definition quests about how Machine Learning can provide new ways to overcome this gender-based violence.

This is one of the crucial problems to be dealt and necessary actions should be taken.

Artificial Intelligence has been prevalent in today's era for handling and mitigating such criminal incidents.

The AI model is a collaborative technique by **Omdena and Safecity** project.

This model leverages open-sourced data as well as Safecity's database to build heatmaps, and safe routes using Nearby Search, Directions API and Grid Coverage techniques.

The Safecity platform acts as a crowdsourced place where individuals can share their personal stories of sexual assaults and rapes.

II. LITERATURE SURVEY

The prior research led to the use of several NLP techniques to easily aggregate and classify the labels. Deep learning algorithms like CNN and RNN were used for building and generating single and multi-label classification. Also, some textual data of Twitter was used for sentimental analysis like generating word clouds associated with such sexual assault cases. This indeed helped to aware citizens. But the scope was limited, and a rather complex approach.

III. IMPLEMENTATION

First of all, we acquired the dataset from the safe-city. After acquiring the dataset we preprocessed the dataset and performed EDA and did statistical and graphical data representation. Then we labelled the categories so that we can apply different algorithms to the data. As most of the important data is in text form we implemented TF-IDFvectorizer and countVectorizer so that we can now use that data for text feature extraction. Found out the most correlated unigrams and bi-grams from the description of each category which can help us ease the process of classifying according to different categories based on given description. Implemented different algorithms i.e., LogisticRegression, RandomForestClassifier, LinearSVC, and MultinomialNB. Generated a confusion matrix for Actual and Predicted categories for the LinearSVC algorithm.

IV. RESULTS

As a result of preprocessing, we removed all the irregularities from the dataset.

From the EDA and visualization, we found some interesting patterns in the data.

After applying different commonly used machine learning algorithms we were able to classify the category from the description. We got the highest accuracy from the LogisticRegression which was 50% and got almost similar accuracy for LinearSVC as well.

V. CONCLUSION

We got considerably less accuracy from all the applied machine learning algorithms. Almost half of the results we got lead to faulty predictions which need to be worked upon to get a better outcome. Therefore now our main focus will be on retraining the model by exploring some other features like PCA for dimensionality reduction, hyperparameter tuning so that the applied model gives better output metrics.

REFERENCES

[1] Karlekar, S. and Bansal, M., 2018. *SafeCity: Understanding Diverse Forms of Sexual Harassment Personal Stories*. [online] Aclweb.org. Available at: <https://www.aclweb.org/anthology/D18-1303.pdf> [Accessed 17 March 2021].

[2] CHANDRA, V. and SRINATH, R., 2020. *Analysis of Women Safety using Machine Learning on Tweets*. [online] Irjet.net. Available at: <https://www.irjet.net/archives/V7/i6/IRJET-V7I6427.pdf> [Accessed 17 March 2021].

[3] Pandey, R., Purohit, H., Stabile, B. and Grant, A., 2018. *Distributional Semantics Approach to Detect Intent in Twitter Conversations on Sexual Assaults*. [online] Arxiv.org. Available at: <https://arxiv.org/pdf/1810.01012.pdf> [Accessed 17 March 2021].

[4] D, H., 2021. *ANALYSIS OF WOMEN SAFETY IN INDIAN CITIES USING DEEP LEARNING ON TWEETS*. [online] Ijsdr.org. Available at: <https://www.ijsdr.org/viewpaperforall.php?paper=IJSDR2006070> [Accessed 17 March 2021].

[5] Sánchez-Medina, A., Galván-Sánchez, I. and Fernández-Monroy, M., 2020. *Applying artificial intelligence to explore sexual cyberbullying behaviour*. [online] sciencedirect.com. Available at: <https://www.sciencedirect.com/science/article/pii/S2405844020300633> [Accessed 17 March 2021].

[6] Chu, T., Jue, K. and Wang, M., 2017. *Comment Abuse Classification with Deep Learning*. [online]Web.stanford.edu. Available at: <https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1174/reports/2762092.pdf> [Accessed 17 March 2021].

[7] Yadav, B., Sheshikala, M., Swathi, N., Chythanya, K. and E, S., 2020. *ShieldSquare Captcha*. [online] Iopscience.iop.org. Available at: <https://iopscience.iop.org/article/10.1088/1757-899X/981/2/022042/pdf> [Accessed 17 March 2021].