

SEXUAL HARASSMENT CLASSIFICATION USING MACHINE LEARNING

PRESENTED BY:

**VIPASHA VAGHELA -
202211002**

**DREAMY PUJARA
- 202211005**

INTRODUCTION

- Sexual harassment is one of the biggest problems, it refers to unwelcome sexual advances, comments of a sexual nature, requests for sexual favors, and harassment or comments about an individual's gender, all of which create a hostile work environment.
- Here our objective is to automatically categorize each story submitted by a user. We will consider the user stories as our training data. We will try to build a Machine Learning model which will take these stories/descriptions as input and try to predict the categories of harassment.



MOTIVATION

- In previous decades opening up about your bad experiences, was considered as taboo, but in today's era Following the #MeToo movement we had a lot of people opening up about their sexual harassment incidents, but as with any internet viral movement, it faded with time. However, this plethora of information can be used effectively to automatically classify abuse incidents into appropriate categories, and this article would help us understand how we can reduce the incidences of Sexual Harassment using Machine Learning.

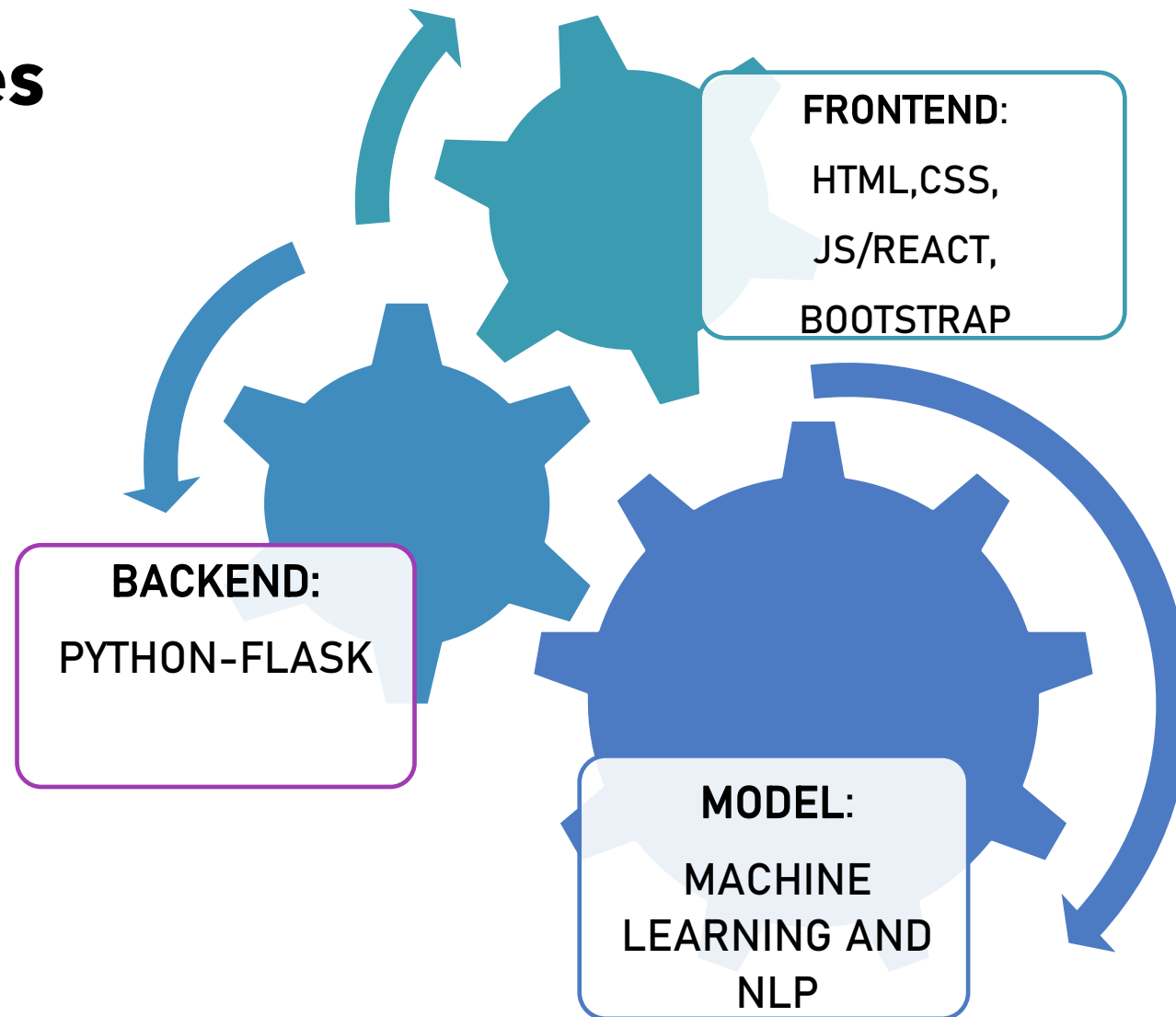


MACHINE LEARNING AND NLP: STATE-OF-ART



- Automatic classification of the type of sexual harassment will enable people who want to make a change in society to better analyze the data, and it is for this very reason that identifying Sexual Harassment using Machine Learning would prove to be a game changer and ensure no such unforeseen incidences take place.
- As we already have many personal stories about sexual abuse shared online through this case study we would make scientific use of the data.
- We would leverage NLP to correctly classify types of Sexual Harassment using Machine Learning. Single and multi-label classification is used for the task to help fill online incident reports automatically and classify, and summarize the data based on the type of abuse.

Tools and Technologies



PROJECT OUTLINE

DATASET ANALYSIS AND EDA OF DATA

**PREPROCESSING AND EMBEDDING OF THE TEXT
USING NLP**

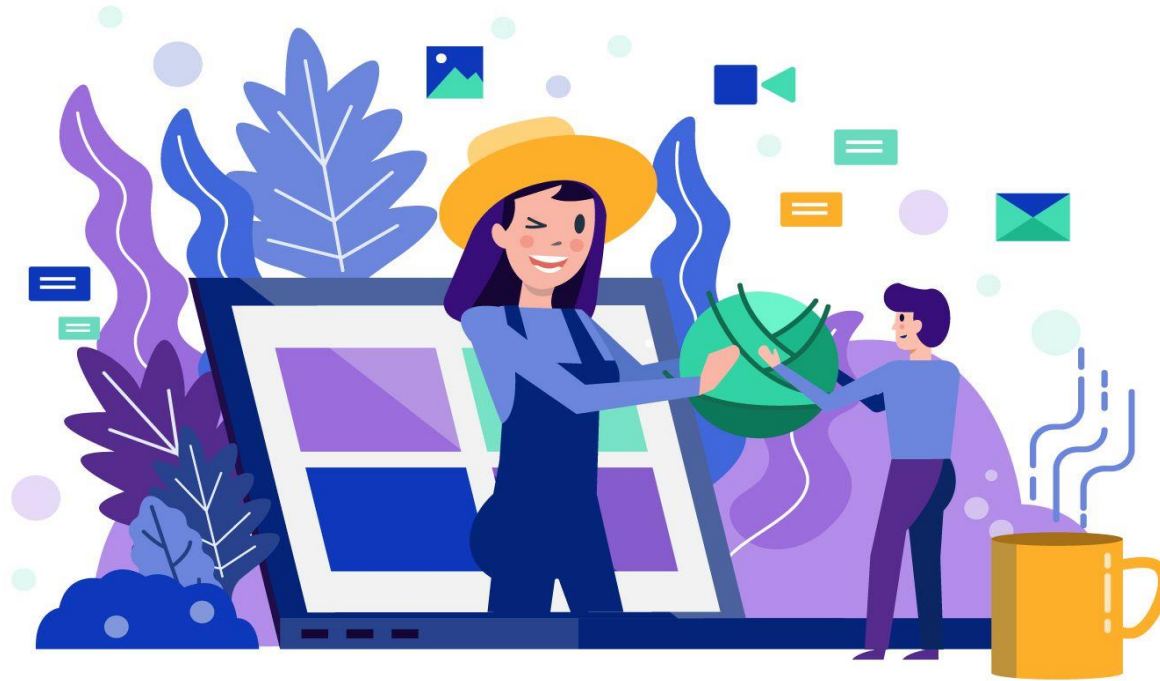
MODEL GENERATION

PERFORMANCE MATRICES

MODEL DEPLOYMENT AND APPLICATION

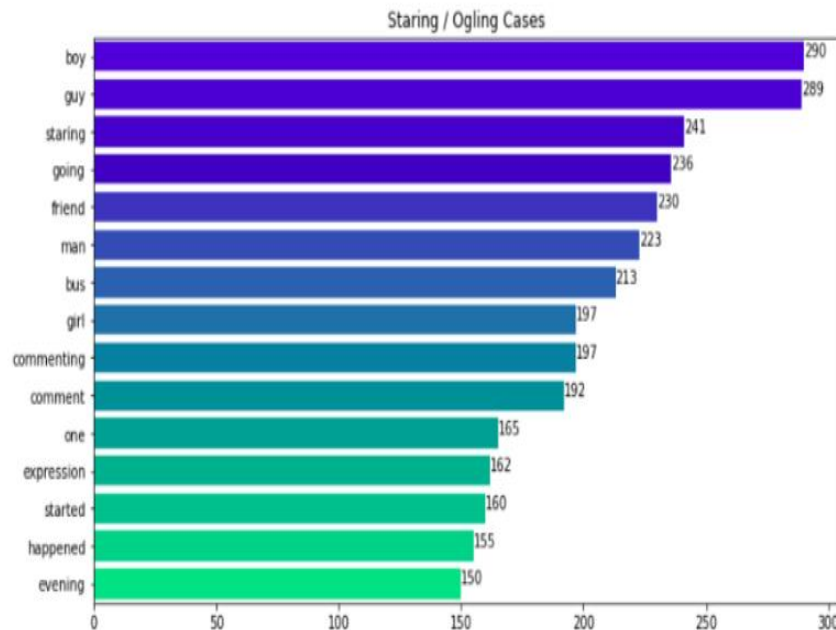


DEVELOPMENT WORK AND TIMELINE



STAGE-1

Frequent words and Wordcloud for Commenting Cases

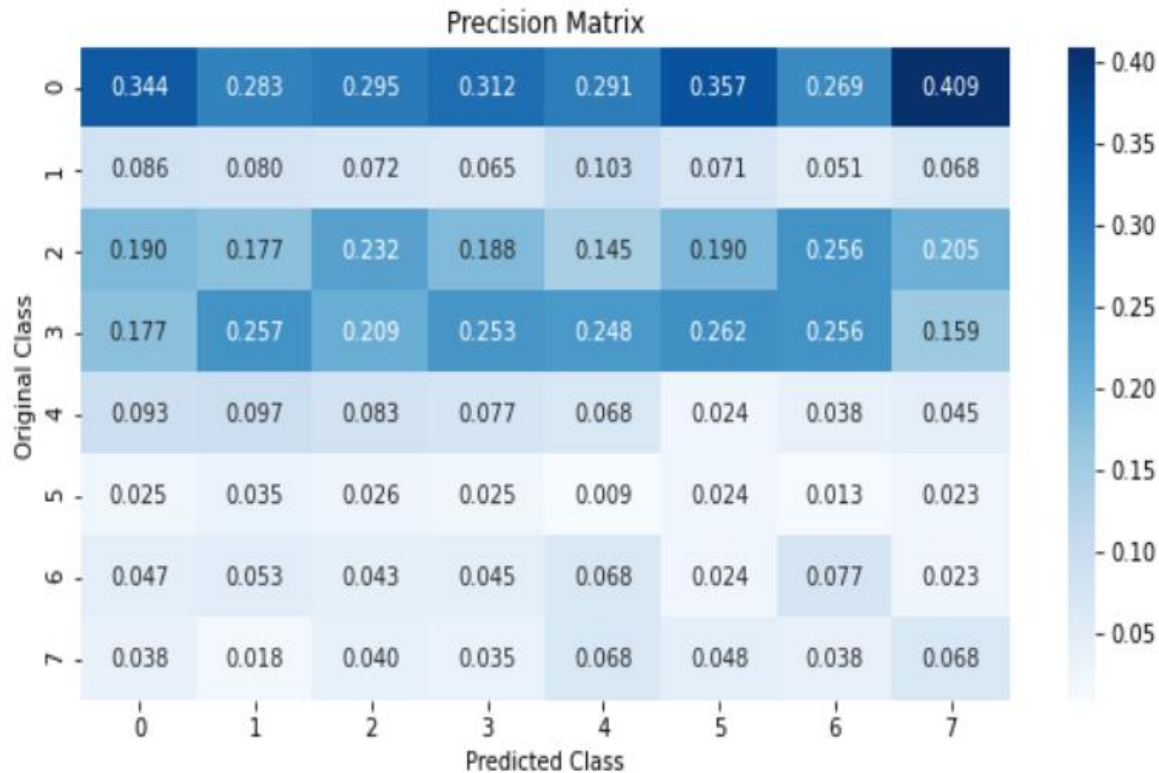


Frequent words and Wordcloud for Staring Cases



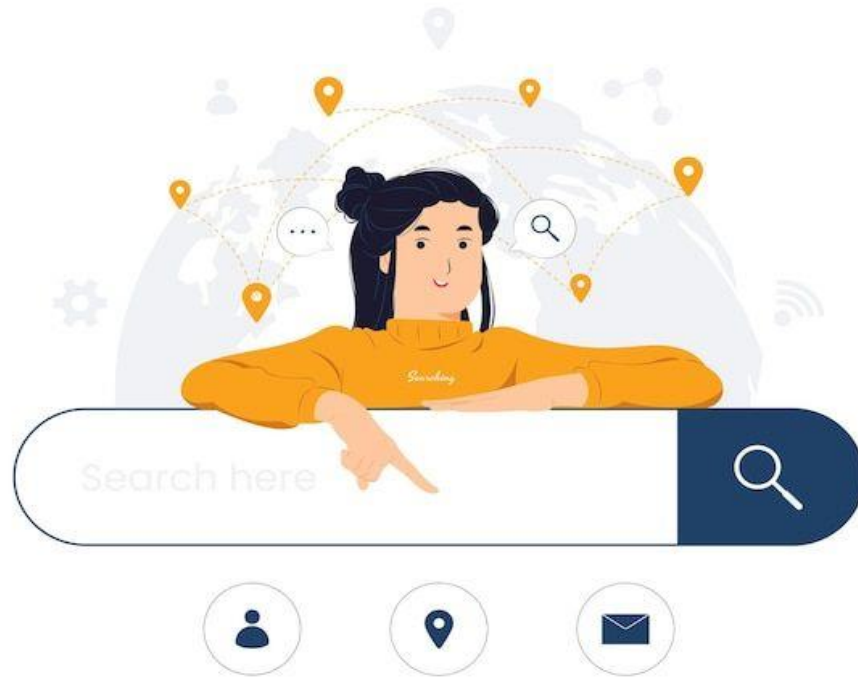
- In the first phase of our project, we are going to preprocess textual data using NLP and going to find some patterns in the data.
- Exploratory data analysis of data and visualization using different graphs and charts.

STAGE-2



- In the second phase we are going to develop a machine-learning model for classification, we are going to implement some core models with the estimation of accuracy.
- For the evaluation of our models, we are going to develop some performance matrices.

STAGE-3



- In the third phase, we are going to deploy our machine learning model in binary pickle format.
- We are going to integrate this deployed model into real-time applications using the flask framework.

Data Set Analysis

Data has been collected from [Safecity](#), which collects anonymous reports of crimes in public spaces. The top three most dense categories groping/touching, staring/ogling, and commenting, to use as our dataset out of a total of 13 different categories, as the others were more sparse. Each description may fall into none, some, or all of the categories. Data provided to us already have split for train test and validation sets, we have used the same combination of splits for our models.

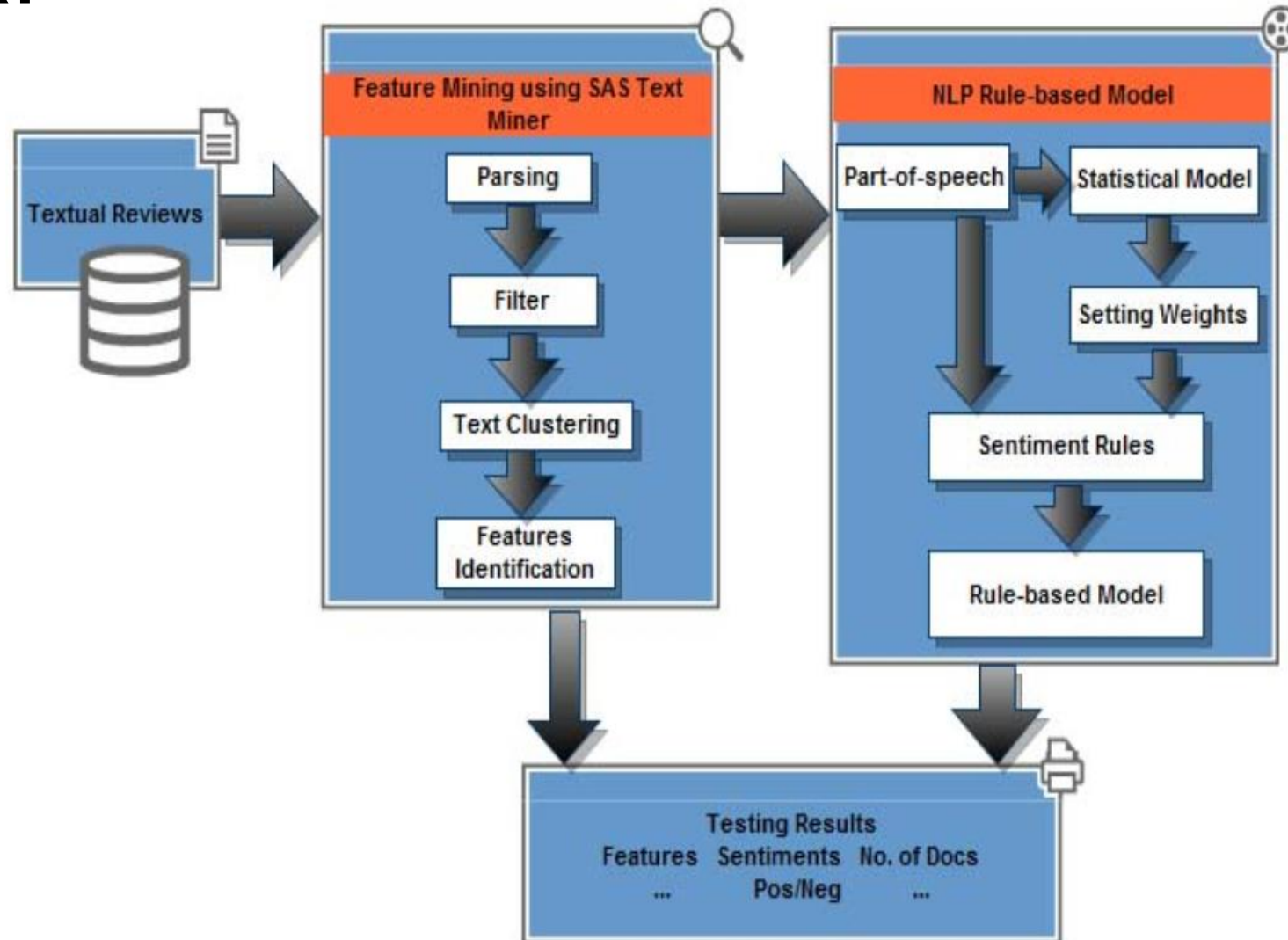
The training dataset has two columns

For single-label classification:-

1. Description of the incident:- We have data present in the English language.
2. Category:- Binary variable 1 indicating harassment 0 indicating none



FLOW-CHART

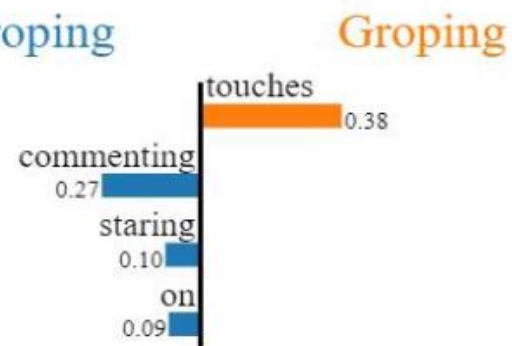
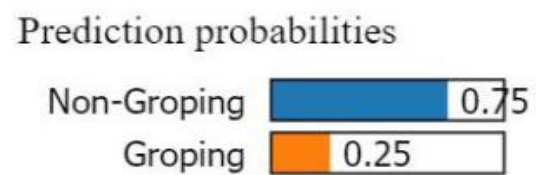
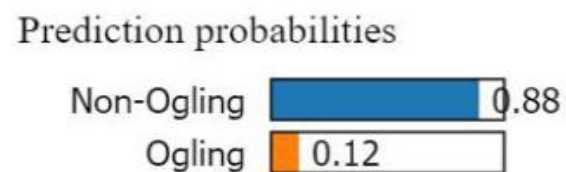
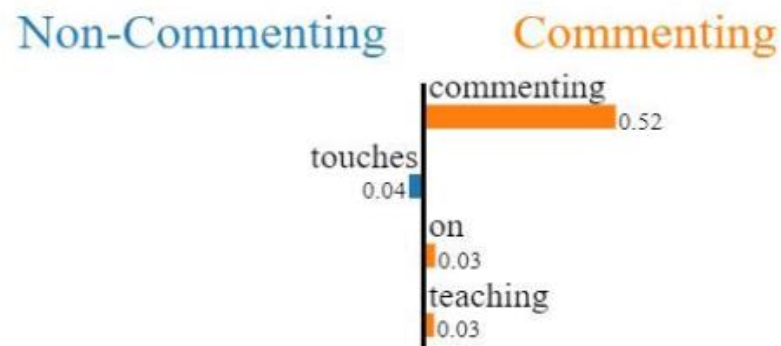
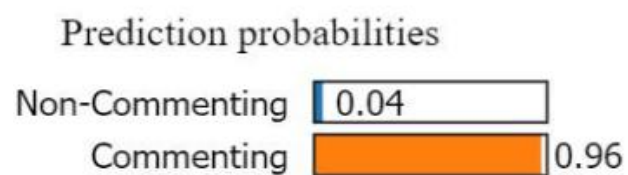


PSEUDO CODE

- 1.Loading dataset
- 2.helper functions
- 3.Text preprocessing and encoding
- 4.Find the best Algorithm using Hyperparameter tuning
- 5.Training best Algorithms
- 6.Feature Importance
- 7.Classifying new text



RESULTS



INTERFACE PROTOTYPE RESULT

Write Your Story Here...

When I was going to my school there were few boys who were commenting on me. when I reached school a teacher staring at me. There is also another non-teaching staff who touches me every day.



Share Your Story...Click Here!

Your story shared.

Commenting : 0.96

Ogling : 0.12

Groping : 0.25

Future Work

1. Get more data and retrain the models.
2. Tune the model with different hyper-parameters.
3. Use State-of-the-art algorithms like **BERT**
(**B**idirectional **E**ncoding **R**epresentation from **T**ransformer).



REVIEW WITH RESPECT TO NATIONAL AND INTERNATIONAL SCENARIOS

■ **References:**

- <https://www.analyticsvidhya.com/blog/2021/07/classifying-sexual-harassment-personal-stories/>
- <https://omdena.com/projects/sexual-harassment>
- <https://medium.com/omdena/automatic-classification-of-sexual-harassment-cases-a5f8cc604ab9>
- <https://mdiqbalbajmi00786.medium.com/sexual-harassment-stories-classification-using-nlp-1st-part-aab114e7015c>
- <https://omdena.com/blog/sexual-abuse-nlp/>

■ **Published results**

- https://www.researchgate.net/publication/353859973_Sentiment_Analysis-Based_Sexual_Harassment_Detection_Using_Machine_Learning_Techniques
- <https://arxiv.org/pdf/1809.04739.pdf>





THANK YOU!