**Name:** MD IQBAL BAJMI

**Email address:** mdiqbalbajmi00786@gmail.com

**Contact number:** 9701922101

**Anydesk address:** **Near Salman Plaza, Adam's Colony, Tolichowki**

**Years of Work Experience:** **0**

**Date:** **08th Apr 2021**

**Self Case Study -1:** SafeCity: Understanding Diverse Forms of Sexual Harassment Personal Stories

"After you have completed the document, please submit it in the classroom in the pdf format."

Please check this video before you get started:
https://www.youtube.com/watch?time_continue=1&v=LBGU1_JO3kg

## Overview

*** Write an overview of the case study that you are working on. ***(MINIMUM 200 words)*** ***

1. Sharing personal stories of sexual harassment increasing everyday on online platform like SafeCity. In order to fight against this type of Criminal Offence We are trying to build a Natural Language Processing model using the stories shared on online platform SafeCity. There are three main categories of sexual harassment recorded online.

   1. Groping

   2. Ogling

   3. Commenting

Our model will classify a given story(text) among the above categories. Story might belong to all the above classes.

So, We solve this problem in two ways:

I. Creating a Classifier for each category as a Simple Binary Classification.

II. Creating a Multi-label classifier for above categories because story may belong to all the above classes

2. Manually sorting and comprehending the stories shared online is a difficult task. So, We are using the power of NLP to help an activist to do real-change in society.

---

### Research-Papers/Solutions/Architectures/Kernels

*** Mention the urls of existing research-papers/solutions/kernels on your problem statement and in your own words write a detailed summary for each one of them. If needed you can include images or explain with your own diagrams. it is mandatory to write a brief description about that paper. Without understanding of the resource please don't mention it***

1. Research paper: **SafeCity: Understanding Diverse Forms of Sexual Harassment Personal Stories.**
2. URL of research paper: https://arxiv.org/pdf/1809.04739.pdf
3. Description:

   Above research paper was published by **Sweta Karlekar** and **Mohit Bansal** on 14 Sep 2018. In this paper, they proposed two ways to solve this problem, One is Single-label classification and second is Multi-label classification.

To solve this problem they used both Machine Learning and Deep Learning techniques for both Single-label and Multi-label classification.

**Single-label Classification:**

They use different ML and DL Models to solve Single-label Classification like:

a. Linear SVM
b. Gaussian NB
c. Logistic Reg
d. SVM
e. CNN
f. RNN
g. CNN-RNN

For evaluation of Single-label Classification they used accuracy-score for both ML and DL models.

*CNN-RNN won the game!*

**Multi-label Classification:**

For Multi-label they used Model like:

a. Random Forest
b. CNN
c. RNN
d. CNN-RNN(bidirectional + character encoding)

For evaluation of Multi-label Classification they used **Exact Match** and **Hamming Score.**

**Analysis of Model:**

They presented 3 ways to Analyse the Model.

1. **Word Embedding Visualization**(Word Cloud per categories)
2. **LIME Analysis(perturbation testing**): Can be applied to any ML Model. Provide Local Model Interpretability. Input should be interpretable by human beings like BOW or TF-IDF not a dense vector like Glove or Word2Vec
3. **First derivative Saliency**: illustrate which words of an input have the biggest impact on the final classification by taking the gradient of the final scores outputted by the neural network(S) with respect to embedding(E), given the true label(L), giving ∂SL(E)/∂E.

---

## First Cut Approach

*** Explain in steps about how you want to approach this problem and the initial experiments that you want to do. ***(MINIMUM 200 words)*** ***

*** When you are doing the basic EDA and building the First Cut Approach you should not refer any blogs or papers ***

1. **Dataset:** As we already have data in the **train**, **val**, and **test** set format given by **SafeCity**.

   **Preprocessing :**

   a. **Removal of html tags**(if any).
   b. Change each word into **lowercase**.

c. **stop word removal**.
d. **Stemming** (because fast and we are not using Word2Vec or Globe).
e. I will use **TF-IDF Vectorizer** to vectorize my dataset.

**Modeling :**

I will use Logistic Regression for Modeling because It works well in high-dimensional space(high dimension because of TF-IDF).

I will hyper-parameter tuning to get a better model on f1-score(to have control on both precision and recall).

**Evaluation:**

a. I will use AUC for binary-classification and a confusion matrix also.
b. I will use micro f1-score, and hamming distance for Multi-label classification. Exact Match will also be used.

**Interpretability:**

**a.** **Feature Importance** using coefficient value of Logistic Regression.
b. Word Embedding Visualization(Word Cloud)
c. LIME Analysis

---

**Notes when you build your final notebook**:

1. You should not train any model either it can be a ML model or DL model or Countvectorizer or even simple StandardScalar
2. You should not read train data files

3. The function1 takes only one argument “X” (a single data points i.e 1*d feature) and the inside the function you will preprocess data point similar to the process you did while you featurize your train data

    a. Ex: consider you are doing taxi demand prediction case study (problem definition: given a time and location predict the number of pickups that can happen)

    b. so in your final notebook, you need to pass only those two values

    c. def final(X):

    ```
    def final(X):
            preprocess data i.e data cleaning, filling missing values etc
            compute features based on this X
            use pre trained model
            return predicted outputs
    final([time, location])
    ```

    d. in the instructions, we have mentioned two functions one with original values and one without it

    e. final([time, location])   # in this function you need to return the predictions, no need to compute the metric

    f. final(set of [time, location] values, corresponding Y values)  # when you pass the Y values, we can compute the error metric(Y, y_predict)

4. After you have preprocessed the data point you will featurize it, with the help of trained vectorizers or methods you have followed for your train data
5. Assume this function is  like you are productionizing the best model you have built, you need to measure the time for predicting and report the time. Make sure you keep the time as low as possible
6. Check this live session: https://www.appliedaicourse.com/lecture/11/applied-machine-learning-online-course/4148/hands-on-live-session-deploy-an-ml-model-using-apis-on-aws/5/module-5-feature-engineering-productionization-and-deployment-of-ml-models