Cyberbullying Detection

Presented by:
Ariba Azam
Rimsha Zahid
Areeba Kabir

Table of contents

Introduction

Apply clustering and association rule mining techniques on a dataset related to bullying statement detection.

Data set:

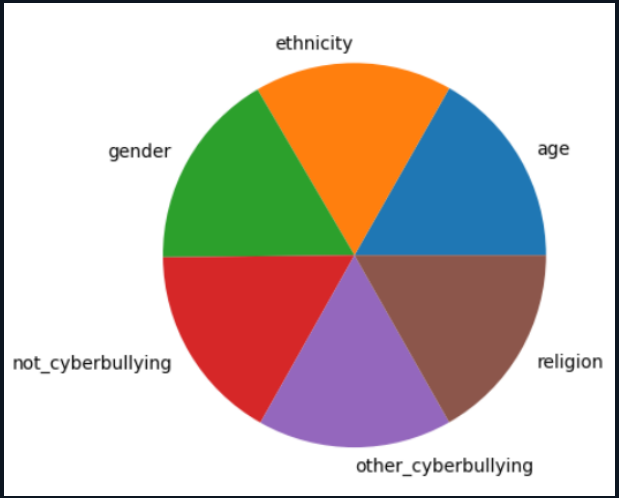
Having only 2 columns :

- tweet_text
- cyberbullying_type

Cyberbulling_types:

cyberbullying_type	
religion	7998
age	7992
gender	7973
ethnicity	7961
not_cyberbullying	7945
other_cyberbullying	7823

	tweet_text	cyberbullying_type
0	In other words #katandandre, your food was cra...	not_cyberbullying
1	Why is #aussietv so white? #MKR #theblock #ImA...	not_cyberbullying
2	@XochitlSuckkks a classy whore? Or more red ve...	not_cyberbullying
3	@Jason_Gio meh. :P thanks for the heads up, b...	not_cyberbullying
4	@RudhoeEnglish This is an ISIS account pretend...	not_cyberbullying

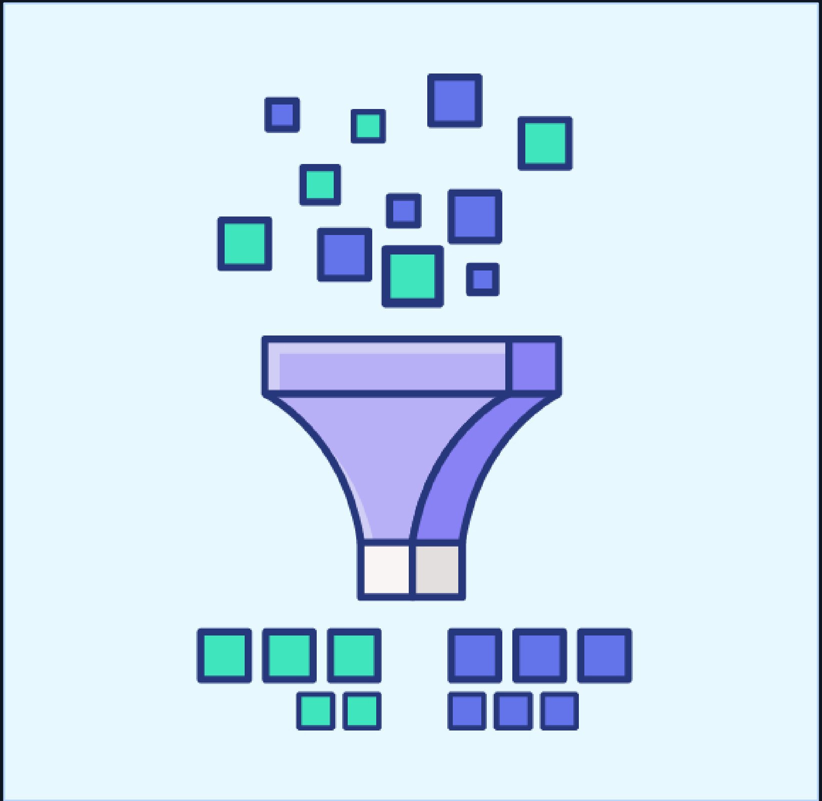


Data Cleaning:

```
def remove_punct(text):
def decontract(text):
def lower(text):
def remove_stopwords(text):
def smile_handle(word_list):
def lemmatize(words):
```

TF-IDF:

Converting the text data into numerical features



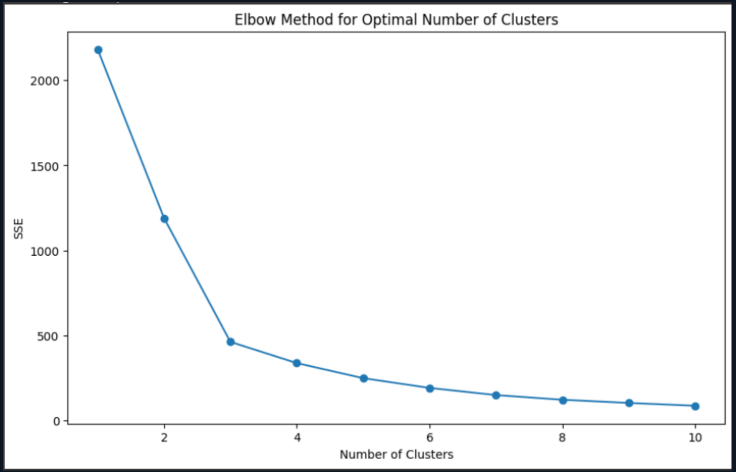
Cluster K-mean

K-Means

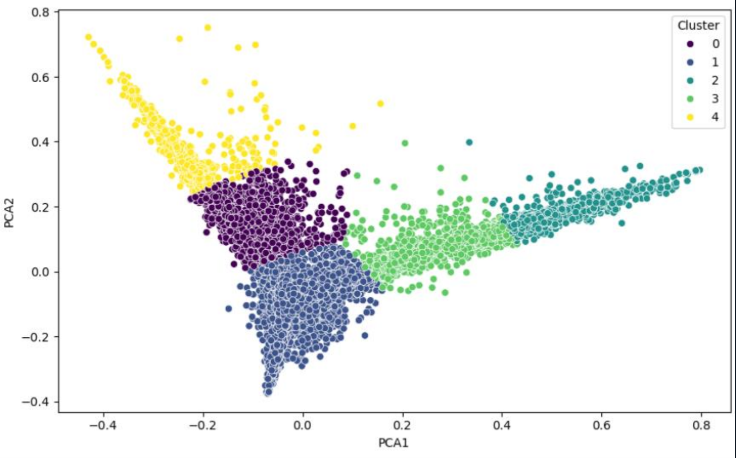
- Used for large data sets
- Efficient when we have to make equal no of clusters

Elbow method

- To find the optimal number of clusters, you look for the "elbow" point in the plot, where the SSE starts to decrease more slowly.
- Elbow appears is 3.
- After 3 clusters, the decrease in SSE becomes more gradual



K-Mean for k=5



- **Cluster Separation: Cluster Characteristics:** clusters 0 and 3 appear more densely packed, whereas clusters 1 and 4 are more widely distributed.
- **Influence of PCA Components:**

Association:

Apriori algorithm

Apriori algorithm

It operates by:

- identifying the frequent itemset in a dataset
- extending them to larger itemset if they meet a specified minimum support threshold.

Implementation:

```
frequent_itemsets_fp = fpgrowth(oht_df_bool, min_support=0.01, use_colnames=True)
```

```
[24] # Step 3: Generate Association Rules
     rules = association_rules(frequent_itemsets_fp, metric='confidence', min_threshold=0.5)
```

```
[25]

     # Step 4: Set thresholds and filter meaningful rules
     filtered_rules = rules[(rules['lift'] > 1) & (rules['confidence'] > 0.5)]
```

Implementation:

```
[26] # Interpret the results
     print(filtered_rules[['antecedents', 'consequents', 'support', 'confidence', 'lift']])
```

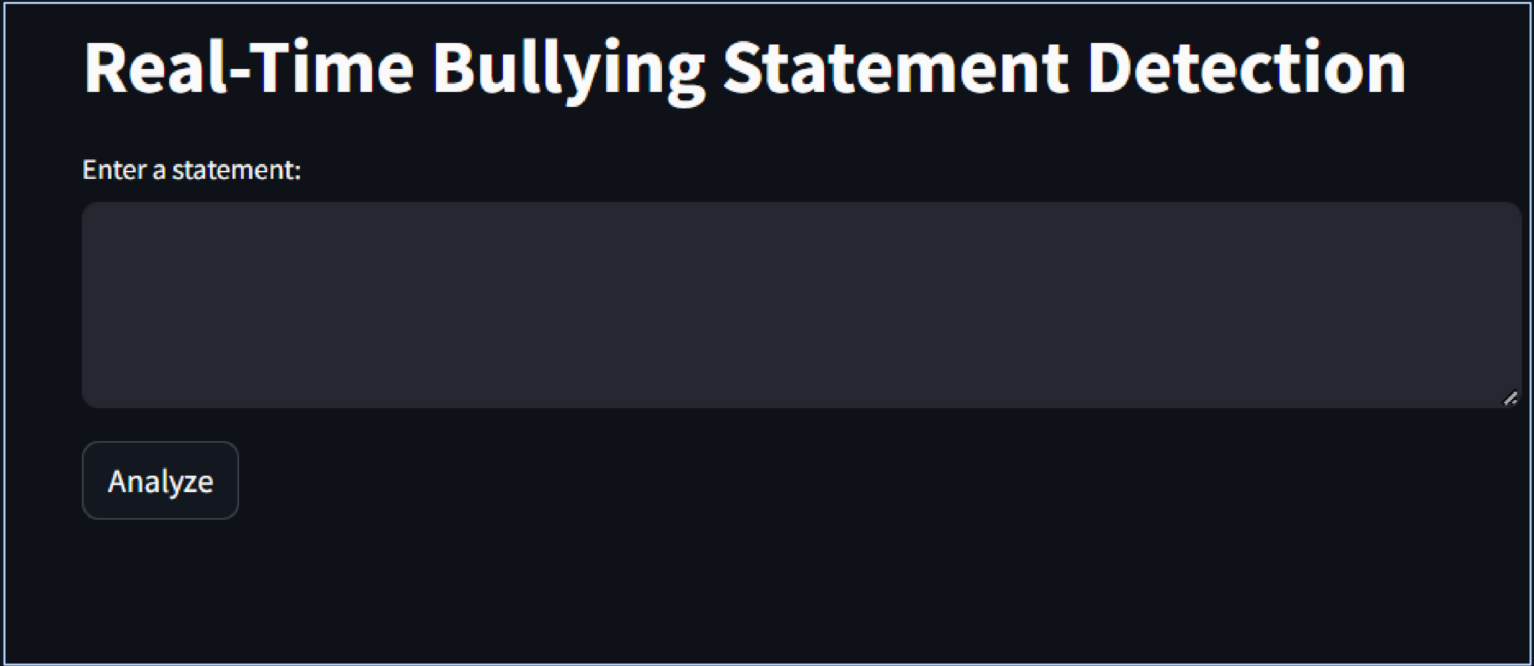
```
             antecedents                           consequents   support  confidence       lift
0        (school, not)                               (bulli)  0.017487    0.923588   4.478219
1          (not, bulli)                             (school)  0.017487    0.818449   4.819545
2        (school, but)                               (bulli)  0.025036    0.950637   4.609371
3          (bulli, but)                             (school)  0.025036    0.863965   5.087570
4       (school, like)                               (bulli)  0.034870    0.962384   4.666331
...              ...                                   ...       ...         ...        ...
1740       (rt, obama)      (tayyoung, nigger, fuck, dumb, as)  0.010211    0.868093  46.258186
1741        (dumb, rt)     (tayyoung, nigger, fuck, obama, as)  0.010211    0.563006  30.000974
1742          (rt, as)   (tayyoung, nigger, fuck, obama, dumb)  0.010211    0.762128  40.611647
1743       (obama, as)      (tayyoung, nigger, rt, fuck, dumb)  0.010211    0.503099  49.268595
1744        (tayyoung)    (nigger, rt, fuck, obama, dumb, as)  0.010211    0.523656  50.452925

[1745 rows x 5 columns]
```

Implementation:

High confidence and lift values indicate strong association. For example, (school, not) => (bulli) with a confidence of 0.923588 and a lift of 4.478219 shows a strong association between these terms.

The rules with antecedents containing terms like (rt, obama), (dumb, rt), and (tayyoung) have very high lift values (ranging from 30 to 50), indicating exceptionally strong associations in those contexts.

Prototype system:



Recommendation:

- **Cluster-Based Association Rule Mining:** First, use clustering
    - (e.g., K-Means or DBSCAN) to group similar instances.
    - Then, apply association mining techniques (e.g., Apriori or FP-Growth)

  Utilize **multimodal data**, including images and videos, to capture a broader spectrum of bullying behaviors.
- **Advanced Text Analysis Techniques:** (**NLP**) techniques, such as transformer-based models like **BERT** and **GPT**.
- **Contextual and Behavioral Analysis:** having more contextual information, such as users' past behaviors and engagement patterns,

Thanks!

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon** and infographics & images by **Freepik**