



**Ahmedabad**  
**University**

# **CSE 523 Machine Learning**

**Progress Report:**

## **Toxic Comments Classification**

**Group Details:**

**NaN-Prediction Pending**

<b>Sr. No.</b>	<b>Name</b>	<b>Enrollment Number</b>
1	Aneri Dalwadi	AU1940153
2	Kairavi Shah	AU1940177
3	Nandini Bhatt	AU1940283
4	Mananshi Vyas	AU1940289

## Tasks performed in the week:

- We tried to observe the results of varying character pairs as well as more than one word for tf-idf matrix and analyzed model performance
- Studied about different methods of dimensionality reduction including PCA
- Tried to manually tune hyperparameters ( e.g. learning rate)

## Outcomes of the tasks performed:

- We were able to observe whether the relevance of words in comment increases or decreases by varying numbers of eigen values .
- Analyze the performance of models after fine tuning and how we could improve it

## Tasks to be performed in the upcoming week:

- We would be implementing more non-linear classifiers and check its accuracy on our dataset
- Working on end semester presentation and report

## References:

1. <https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge/data>
2. <https://smldar.com/>
3. <https://www.tidytextmining.com/tidytext.html>
4. <https://arxiv.org/ftp/arxiv/papers/1903/1903.06765.pdf>