



Ahmedabad
University

CSE 523 Machine Learning

Progress Report:

Toxic Comments Classification

Group Details:

NaN-Prediction Pending

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

Tasks performed in the week:

- Data Cleaning
- Referred some already built classification models
- Testing out which base models would work for our data

Outcomes of the tasks performed:

- Got a better understanding of the data
- Generated a clean data with only required information
- Learned how to write SQL commands to clean the data
- Learned about the different layers we can give to our model, yet to finalize the same

Tasks to be performed in the upcoming week:

- Creating a baseline model
- Making improvements using intuition
- Testing other classification models

References:

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