

Weekly Team Meeting: Every Friday

5:00-6:00 PM Our github repo: [acmucsd-projects/wi23-ai-team-1 \(github.com\)](https://github.com/acmucsd-projects/wi23-ai-team-1)

Team Meeting Time: 2/24/23 at 5pm in CSE Basement

Attendees: Arnav Modi Aniket Gupta Steven Shi Vivian Liu Jeffrey Lee Jimmy Ying

What have we done so far

- Cleaning:
 - Filtering by number of words
 - Filtering non-english chars
 - Trimming whitespace
- Exploration
 - Graphs based on the number of words
 - Frequencies of toxic tags
- Initial steps of tokenization
 - Using tensorflow to create a `vectorize_layer` that separates words in the comment by whitespaces, forces everything to lowercase, and tokenizes to ints

What is the point of this meeting, what are we going to discuss

Trying to optimize our development workflow, looking at how we're collaborating.

More preprocessing stuff:

https://keras.io/api/layers/preprocessing_layers/core_preprocessing_layers/text_vectorization/

What will we do going forward

- 1) Watch a video on tensorflow and nltk