**Team Name:** Automat3rs

**Team Members:** Brenda Parnin, Rudy Martinez, Jose Fernandez, Arsalan Bhojani

**Submission Date:** May 7, 2021

**Overview**

For this project, we were tasked with creating a model that accurately predicts the classification of offensive language in Twitter tweets. To initiate this process, we were provided two files: a *train.tsv* dataset and a *test.tsv* dataset. These datasets included Twitter identification numbers, Twitter tweets, and labels that correspond to the tweet’s classification as NOT (Not Offensive), TIN (Targeted Insult), or UNT (Untargeted Insult).

**Describe your approach:**

1. Initial Steps / Setup
2. What model did you use?
3. What hyperparameters (e.g., C values) did you search over?
4. What features did you test (e.g., ungrams, unigrams+bigrams, lexicon features, other...)

**Describe your model selection procedure**

1. Here you should report validation scores of you model and potentially variants of your model. For instance, if you trained a model with 3 sets of features (A+B+C), report the results of A+B, A+C, and B+C
2. You should also perform an error analysis. Look at the predictions of your model. Find errors made by your model and try to explain why they were made (e.g., maybe your features don't handle negation very well).

**Assignments**

* Rudy –
* Brenda –
* Jose –
* Arsalan –