**FUTURE ENHANCEMENT:**

Use a different machine learning algorithm. The study used a variety of machine learning algorithms, but the best performing algorithm was AdaBoost. Other algorithms, such as deep learning models, may be able to achieve better performance.

Use a more rigorous evaluation methodology. The study used a simple accuracy metric to evaluate the performance of the models. More rigorous metrics, such as precision, recall, and F1 score, can be used to get a more complete picture of the models' performance.

Conduct a user study to evaluate the models' effectiveness. The study did not evaluate the models' effectiveness with real users. A user study can be conducted to get feedback from users on how well the models are able to detect offensive language