



CALIFORNIA SCIENCE & ENGINEERING FAIR 2018 PROJECT SUMMARY

Name(s) Advait Arumugam	Project Number S1501
Project Title Cybian: A Machine Learning Program Based on Naive Bayes to Classify and Mitigate Cyberbullying	
<p style="text-align: center;">Abstract</p> <p>Objectives/Goals The objective was to create a program, called Cybian, to detect cyberbullying in social media content based on a set of algorithms to (i) classify the sender, (ii) evaluate and detect the bullying message by using machine intelligence based on Naïve Bayes and the category of the sender, (iii) send the classified bullying message to the administrator for review, and (iv) either send feedback to present the message to the recipient or take action.</p> <p>Methods/Materials I developed Cybian, a conceptual model based on machine intelligence that comprises of a set of algorithms to (i) intercept social media/instant messages at the recipient device, (ii) categorize the messages into pre-defined groups based on the sender, and (iii) classify messages by performing sentiment analysis, extracting features of bullying content, and comparing it against that of labeled data sets. Cybian determines behavioral patterns of the messages using the Bayesian probabilistic formula in the context of cyberbullying. (iv) Depending on the sender category and the threshold value, Cybian decides whether to present the message to the recipient or send it to the administrator. (v) Part of Cybian also resides on the administrator's device to present the classified message for review and respond by either sending feedback to the recipient or discarding the message with appropriate action. (vi) Based on the administrator's feedback, Cybian presents the message to the recipient. I developed a simulator to test my model using Anaconda Python 3.6.4 and Spyder 3. Separate modules for (i) the sender, (ii) the recipient, and (iii) the administrator were developed. To check the accuracy of the Naïve Bayes Algorithm, I created confusion matrices to compare classification and misclassification in the form of True Positives, True Negatives, False Positives, and False Negatives by performing tests on five different test data sets.</p> <p>Results Data from the five confusion matrices showed that Cybian had a 77.4% accuracy in detecting bullying messages.</p> <p>Conclusions/Discussion Cybian introduces a unique concept of categorizing the social media/instant messages based on the sender, detecting bullying messages by machine learning techniques, sending bullying messages to an administrator to allow or reject the messages. The results show that Cybian effectively mitigates cyberbullying while maintaining the privacy of the recipient so all my design criteria were met.</p>	
Summary Statement Cybian, a hybrid software based on machine learning and human intervention successfully mitigates and ends the cycle of cyberbullying.	
Help Received I designed and programmed the algorithm myself after an internet search on techniques. I compiled many databases together for this project.	