Summary:

According to Vikas S Chavan and Shylaja S S in their paper “Machine Learning Approach for Detection of Cyber-Aggressive Comments by Peers on Social Media Network”. Their pre-processing consists of 2 main parts the “removing unwanted strings” and “correcting words”. Second part the standard feature extraction. They used they feature extraction methods for this part. First, N-gram method, they used the n-gram method to extract the tokens they are going to work on. Second, counting, in this part they are going to tokenize the comments and count the occurrence of every token. Third, TF-IDF, they used it to calculate the importance of the words according to their number of occurrences in the document.

Third part is the additional features. They used 2 methods in this part which are capturing pronouns and skip-gram. Fourth part is feature selection. For this part they used the chi- square which is a method that is used to select the best features to work on. Last part, classification. They used 2 classifiers for this which are the logistic regression and support vector machine. The logistic regression achieved 73.76 accuracy and 0.6 recall and 0.644 Precision. But for the support vector machine they achieved 77.65 accuracy and 0.58 recall and 0.7 precision's Got their Datasets from Kaggle. They have multiple drawbacks like in the pre-processing part they used to correct the words that is written to bypass their classifier according to their paper they are going to use the dictionary to make this work but if the model encountered a new word that isn’t available in the dictionary it is going to bypass it without alerting the user that there is a cyberbully in this comment. Also, they haven’t used any sentiment or contextual analysis to make sense of the words which is going to make a big problem of false positive because there could be 2 friends talking and use any word that is in the dictionary of the bullying words so this is going to raise alert flag without any actual bullying.