Cyberbullying Classification using Text Mining

Summary:

This paper introduces a new technique to detect cyberbullying on the social media through using 2 famous classifiers which are Naive Bayes And SVM.

System Explanation:

They constructed a system that consists of 4 layers which are , Data Collection , Pre-Processing , Extraction and Classification.

Data Collection. They Used a data set from Kaggle which is a group of messages fromspring and it had Questions , Answers and Severity.

Pre-Processing. This stages they Made the data cleaning which exported it in 10 folders according to the severity class and removed the records that has characters under 15 and removed meaningless Words. And then they Tokenized the text into words . Then they transformed all the text to lower case, moving on they removed stop words and then Filter token that chooses words between 3-25 characters as less than 3 is stop word. Then stemming to make the word go back to it’s original form. Last they generated N-Grams they used the N-Grams from 2-5.

Feature Extraction. In this stage they used TF-IDF as a feature to calculate the weight of the word in the sentence.

For the Classification They used Naive Bayes and SVM with linear Poly RBF and Sigmoid Kernels.

Results:

Naive Bayes yields an average accuracy of 92.81%, SVM with a poly kernel yields an average accuracy of 97.11%.

Datasets:

Fromspring Dataset from Kaggle

Draw Backs:

They Removed the Stop words from the text which means they removed the pronouns and there could be text with any strong word but it isn’t referring to actual person and will be classified as cyberbullying.