Alexander Mosiychuk

Cmpt 220

Project 2 proposal

Offensive Website Web Crawler

For my semester project, I intend to develop a WebCrawler that evaluates the offensiveness of an inputted website. In particular, the user will input a URL, which will then prompt the WebCrawler to scan the site along with any attached hyperlinks. To measure the offensive nature of an inputted site, I will likely implement some form of keyword system, searching for swear words, racial slurs, and other potential terms. Furthermore, if it within the realm of possibility, I might implement some form of image evaluation. However, my focus will be primarily centered around the evaluation of text based content. Besides simply collecting keywords, I also plan to implement some form of feedback to the user, displaying the risk of the site being offensive. My plan for this component of the project is to evaluate the amount and severity of the text which is recorded by the WebCrawler. Afterwards, I will use the evaluation to display the likelihood of a site being offensive. For example, a site containing a plethora of racial slurs would be evaluated with a high chance of being offensive. In contrast, a site containing only a handful of swear words would most likely be evaluated with a minimal likelihood of being offensive.