**FEATURE DETECTOR:**

Functions to be considered are :

* Is it part of the title tag
* Heading tags (from h1 to h6 ) : have a separate function for each heading tag
* Anchor tag
* url name
* tfidf-value

extract all the tags (this gives info on which word falls under which tag), get the tf-idf for each term.

let every word have a dictionary associated with it which basically has key-value pairs of {feature : yes /no(present or not)} //Let this be global . So wherever it is present : i.e. title tag, heading tag , etc the dictionary for that word is updated . let there be a rank field for every word : the more features associated with the word the higher the ranking