

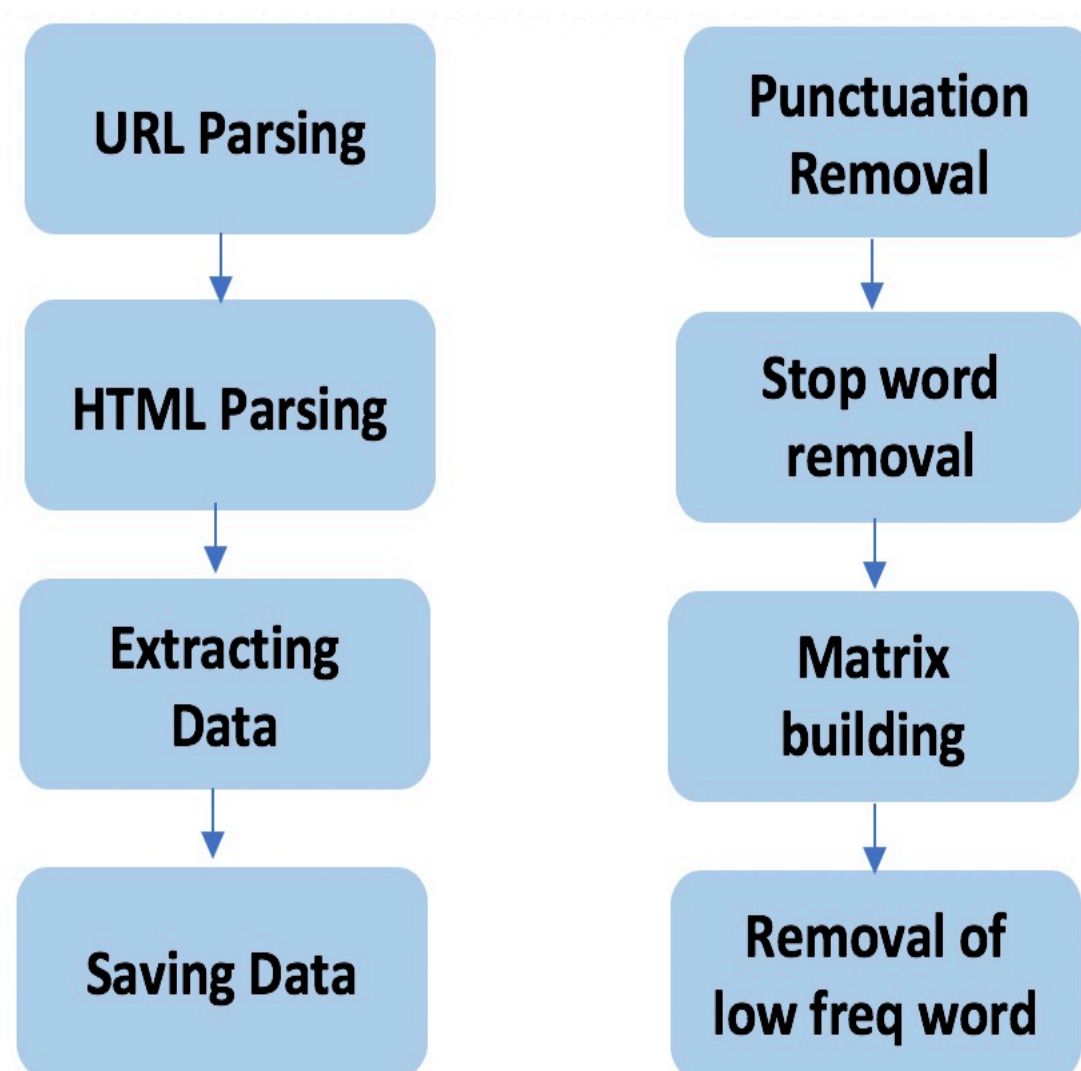
Sushil Suresh Gaidhankar  
MSc in Big Data

## Abstract:

The aim of this project is to analyze the customer reviews posted on the hotel website. Sentimental analysis, text mining, scraping, semi automation, supervised learning is the field of study which analyses customer reviews. Text mining is performed automatically through R programming. Scraping the customer reviews could be used to improve the service to be offered by the hotel. Sentimental analysis is performed on text based data to a particular keyword which is present in the text of reviews.



- Find the tags which holds the customer reviews and also find the HTML tags which holds the overall best rating.
- Now extract the relevant tags which consist the customer reviews and the rating.
- Now the data is being saved in an csv format for better understanding.
- Text mining is performed on the hotel website automatically through the use of R programming.
- Further steps are performed on the scraped data i.e punctuation removal, stopword removal, matrix building, removal of low freq words.
- Finally the decision tree is generated.



## Methodology:

- Fetch the coded URL and navigate to the hotel website webpage.



## Conclusion:

- Performed sentimental text analysis on the customer reviews and build the matrix and decision tree which will help the hotel to improve their services.