



# Tutorial on mining of biomedical literature with the help of R Package

Vinaitherthan Renganathan

[www.vinaitherthan.com](http://www.vinaitherthan.com)

## Abstract

This paper provides step by step overview of process involved in mining of biomedical literature using R-Statistical Package. Abstract from PubMed database on a given topic are retrieved, stored, pre-processed using R programming codes. The resultant term document matrix is used to find association between terms and frequency of the terms in each document. Finally the clouds of words and clustering of documents are created using the R software to discover the association between the documents. The results from the process provided a step by step understanding of the retrieval of abstracts, pre-processing of abstracts and clustering of abstracts using the user based query term

**Keywords:** Biomedical, Clustering, Classification, R Software, Text mining

## 1. Introduction

This paper assumes that the readers have knowledge about text mining concepts and especially in biomedical domain. Those who are interested to get an overview of text mining and its application biomedical domain are encouraged to refer to the author's paper on text mining in Biomedical domain [1] - [Renganathan.V. \(2017\). Text Mining in Biomedical Domain with Emphasis on Document Clustering Healthcare informatics research, 23\(03\) Pages 141-146](#)

## 2. R Package

The R [2] package is an open source statistical computing software which is useful for carrying out various statistical tests and methods, graphics, text and data mining procedures. The R software can be downloaded from the software website [1]. The R software can be used in various integrated development environment (IDE) such R-Studio, Eclipse and StatET. This paper use R-Studio [3] IDE which is an open source software and can be downloaded from the R-Studio website [3].

The R software works with the concepts called packages which is a compilation user created codes and can be used to perform specific functions.

Please email to get the code for text analysis [vinaiweb@yahoo.com](mailto:vinaiweb@yahoo.com)