**Capstone Project**

**Assignment 2**

Course code: CSA 1643

Course : Data warehousing and data mining for data science

S.No: 10

Name : CH. Chandra naga sai kumar

Reg. No: 192210299

Slot : C

Title : Text mining for competitive intelligence in business

Assignment Release Date : 14/02/2024

Assignment Preliminary Stage ( Assignment 1 ) submission Date : 20/03/2024

Mentor Name : Asha

Mentor Phone number and Department : 9840562097(foreign languages dept)

Code for title Text mining for competitive intelligence in business

In R programming :

# Load necessary libraries

library(tidyverse)

library(lubridate) # For handling dates

# Define function for text segmentation

TextSeg <- function(text, lag) {

# Your text segmentation logic here

}

# Define function for feature extraction

MakeFeatureWordsDict <- function(all\_words\_tf\_dict, stopwords\_set, writewords\_set, lag, fea\_dict\_size) {

# Your feature extraction logic here

}

# Define function for text processing

TextFeature <- function(words\_feature, textseg\_list) {

# Your text processing logic here

}

# Define function for classifier training

ClassifierTrain <- function(train\_features, train\_class) {

# Your classifier training logic here

}

# Define function for text mining

MakeTextMining <- function(posts, time\_col, content\_col, source\_col, t\_status\_col, keyword\_col, country\_col, imp\_col, limit\_number,

lag, stopwords\_set, blackwords\_set, writewords\_set,

all\_words\_tf\_dict, all\_words\_idf\_dict, train\_datas, test\_speedup) {

# Your text mining logic here

}

# Define function for classification testing

MakeTextMining\_ClassifyTest <- function(posts, time\_col, content\_col, source\_col, t\_status\_col, keyword\_col, country\_col, imp\_col, limit\_number,

lag, stopwords\_set, blackwords\_set, writewords\_set,

all\_words\_tf\_dict, all\_words\_idf\_dict, train\_datas, test\_speedup) {

# Your classification testing logic here

}

# Define function for calendar processing

MakeTextMining\_Calendar <- function(posts, time\_col, content\_col, source\_col, t\_status\_col, keyword\_col, country\_col, imp\_col, limit\_number,

lag, stopwords\_set, blackwords\_set, writewords\_set,

all\_words\_tf\_dict, all\_words\_idf\_dict, train\_datas, test\_speedup) {

# Your calendar processing logic here

}

# Call the functions with appropriate arguments

# Example:

# result <- MakeTextMining(posts, time\_col, content\_col, source\_col, t\_status\_col, keyword\_col, country\_col, imp\_col, limit\_number,

# lag, stopwords\_set, blackwords\_set, writewords\_set,

# all\_words\_tf\_dict, all\_words\_idf\_dict, train\_datas, test\_speedup)

OUTPUT :



