kaggle

josef

September 14, 2018

library(ggplot2) # Data visualization  
library(readr)   
  
  
  
# import the train dataset  
train = read.csv("C:/Users/User/Desktop/My old files/Recent folder Free desktop/Recent Analysis/josef/train.csv")  
head(train)

## id  
## 1 0000997932d777bf  
## 2 000103f0d9cfb60f  
## 3 000113f07ec002fd  
## 4 0001b41b1c6bb37e  
## 5 0001d958c54c6e35  
## 6 00025465d4725e87  
## comment\_text  
## 1 Explanation\nWhy the edits made under my username Hardcore Metallica Fan were reverted? They weren't vandalisms, just closure on some GAs after I voted at New York Dolls FAC. And please don't remove the template from the talk page since I'm retired now.89.205.38.27  
## 2 D'aww! He matches this background colour I'm seemingly stuck with. Thanks. (talk) 21:51, January 11, 2016 (UTC)  
## 3 Hey man, I'm really not trying to edit war. It's just that this guy is constantly removing relevant information and talking to me through edits instead of my talk page. He seems to care more about the formatting than the actual info.  
## 4 "\nMore\nI can't make any real suggestions on improvement - I wondered if the section statistics should be later on, or a subsection of ""types of accidents"" -I think the references may need tidying so that they are all in the exact same format ie date format etc. I can do that later on, if no-one else does first - if you have any preferences for formatting style on references or want to do it yourself please let me know.\n\nThere appears to be a backlog on articles for review so I guess there may be a delay until a reviewer turns up. It's listed in the relevant form eg Wikipedia:Good\_article\_nominations#Transport "  
## 5 You, sir, are my hero. Any chance you remember what page that's on?  
## 6 "\n\nCongratulations from me as well, use the tools well. Â Â· talk "  
## toxic severe\_toxic obscene threat insult identity\_hate  
## 1 0 0 0 0 0 0  
## 2 0 0 0 0 0 0  
## 3 0 0 0 0 0 0  
## 4 0 0 0 0 0 0  
## 5 0 0 0 0 0 0  
## 6 0 0 0 0 0 0

names(train)

## [1] "id" "comment\_text" "toxic" "severe\_toxic"   
## [5] "obscene" "threat" "insult" "identity\_hate"

dim(train)

## [1] 159571 8

# clean the text  
library(tm)

## Loading required package: NLP

##   
## Attaching package: 'NLP'

## The following object is masked from 'package:ggplot2':  
##   
## annotate

library(SnowballC)  
corpus = VCorpus(VectorSource(train$comment\_text))  
corpus = tm\_map(corpus, content\_transformer(tolower))  
corpus = tm\_map(corpus, removeNumbers)  
corpus = tm\_map(corpus, removePunctuation)  
corpus = tm\_map(corpus, removeWords, stopwords())  
#corpus = tm\_map(corpus, stemDocument)  
corpus = tm\_map(corpus, stripWhitespace)  
  
# Creating the Bag of Words model  
dtm = DocumentTermMatrix(corpus)  
dtm = removeSparseTerms(dtm, 0.994)  
dataset = as.data.frame(as.matrix(dtm))  
head(dataset)

## â€“ â€” â€¢ able absolutely accept acceptable accepted according account  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## accurate across act action actions actual actually add added adding  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 1 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## addition address admin administrator administrators admins advice afd  
## 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## ago agree allow allowed almost alone along already also although always  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## american among another answer anyone anything anyway apparently appear  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## appears appreciate appropriate area arent argument arguments around  
## 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0  
## 4 1 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## article articles ask asked asking ass assume attack attacks attempt  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 1 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## attention august author automatically available avoid aware away back  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## bad ban band banned based become behavior believe best better bias  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## biased big bit black block blocked blocking book books box bring british  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## call called calling came can cant care case cases category cause certain  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 1 0 0 0 0 0  
## 4 0 0 0 0 1 1 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## certainly chance change changed changes changing check cheers citation  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 1 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## citations cite cited city claim claims clear clearly close come comes  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## coming comment comments common community company complete completely  
## 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## concerns conflict consensus consider considered contact content contest  
## 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## context continue contribs contribute contributing contributions copy  
## 1 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0  
## copyright correct country couple course create created creating criteria  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## criticism current currently date day days deal dear death debate decide  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 1 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## decided definition delete deleted deleting deletion described  
## 1 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0  
## description despite details didnt difference different directly disagree  
## 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## discuss discussed discussion dispute disruptive doesnt done dont doubt  
## 1 0 0 0 0 0 0 0 1 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## due earlier early edit edited editing editor editors edits either else  
## 1 0 0 0 0 0 0 0 0 1 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 1 0 0 0 0 1 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 1  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## email encyclopedia end english enjoy enough entire entirely entry error  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## especially etc even ever every everyone everything evidence exactly  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 1 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## example except exist existing explain explaining explanation external  
## 1 0 0 0 0 0 0 1 0  
## 2 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## fact facts fair faith false family far feel field file film finally find  
## 1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0 0  
## fine first five fix follow following form found four free friend friends  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 1 0 0 0 0 1 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## fuck fucking full future game gave general generally get gets getting  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## give given gives giving god goes going gone good google got government  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## great group guess guidelines guy guys hand happen happened happy hard  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 1 0 0 0 0 0 0  
## 4 0 0 1 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## hate havent head heard hell hello help helpful hes hey high highly  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 1 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## historical history hope hours however human idea ill image images  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## important improve inappropriate include included including inclusion  
## 1 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0  
## incorrect indeed indicate info information instead interest interested  
## 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0  
## 3 0 0 0 1 1 1 0 0  
## 4 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## interesting internet involved irrelevant isnt issue issues ive january  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 1  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## job john july june just keep kind know knowledge known knows lack  
## 1 0 0 0 0 1 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 1 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 1 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## language large last later law lead learn least leave left less let lets  
## 1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 2 0 0 0 0 0 0 0 1 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0 0  
## level life like likely line link linked links list listed little live  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 1 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## living long longer look looked looking looks lost lot love made main  
## 1 0 0 0 0 0 0 0 0 0 0 1 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## major majority make makes making man manual many march material matter  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 1 0 0 0 0 0  
## 4 0 0 1 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## may maybe mean meaning means meant media members mention mentioned  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 2 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## merely message messages might mind mine mistake months move moved much  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## multiple music must name names national necessary need needed needs  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 1 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## neither neutral never new news next nice none nonsense notability  
## 1 0 0 0 1 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## notable note nothing notice noticed now npov number obvious obviously  
## 1 0 0 0 0 0 1 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## official often okay old one ones online open opinion opinions order  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## original others otherwise outside page pages paragraph part particular  
## 1 0 0 0 0 1 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 1 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 1 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## particularly party past people per perhaps period person personal  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## personally picture piece pillars place placed play please point points  
## 1 0 0 0 0 0 0 0 1 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 1 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## policies policy political position possible possibly post posted posting  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## pov power present pretty previous probably problem problems process  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## produce project proof proper proposed prove provide provided public  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## published purpose put putting quality question questions quite quote  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## rather read readers reading real really reason reasons recent recently  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 1 0 0 0 0  
## 4 0 0 0 0 1 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## record redirect refer reference references regarding regards related  
## 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 2 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## release relevant reliable remember removal remove removed removing reply  
## 1 0 0 0 0 0 1 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 1 0 0 0 0 0 1 0  
## 4 0 1 0 0 0 0 0 0 0  
## 5 0 0 0 1 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## report request requesting research respect respond response rest result  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## revert reverted reverting review right rights rule rules run said  
## 1 0 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 1 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## sandbox saw say saying says school science search second section  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 1  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## sections see seem seems seen sense sentence separate series serious  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 1 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## seriously set several shit short shouldnt show shows side sign  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## significant similar simple simply since single site sites situation  
## 1 0 0 0 0 1 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## small someone something sometimes soon sorry sort source sourced sources  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## speak specific specifically speedy standard start started state stated  
## 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0  
## statement statements states status stay still stop story stuff stupid  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## style subject subjects suggest summary support supposed sure system tag  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 1 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## tagged tags take taken taking talk talking tell template term terms test  
## 1 0 0 0 0 0 1 0 0 1 0 0 0  
## 2 0 0 0 0 0 1 0 0 0 0 0 0  
## 3 0 0 0 0 0 1 1 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 1 0 0 0 0 0 0  
## text thank thanks thats theory therefore theres theyre thing things  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 1 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 1 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## think thinking third though thought three thus tildes time times title  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 1 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0  
## today together told took top topic totally towards tried true truth try  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## trying tutorial two type unblock understand understanding unfortunately  
## 1 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0  
## 3 1 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0  
## united university unless uploaded upon use used useful user username  
## 1 0 0 0 0 0 0 0 0 0 1  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 1 0 0 0 0  
## users uses using usually utc valid vandalism vandalize various version  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 1 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## video view views violation wait want wanted wants war warning wasnt way  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 1 0 0 0  
## 4 0 0 0 0 0 1 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0  
## web website week welcome well went whatever whats whether white whole  
## 1 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 2 0 0 0 0 0 0  
## wiki wikipedia wikipedian wikipedias will wish within without wont word  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## words work worked working works world worth wouldnt write writing  
## 1 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0  
## written wrong wrote yeah year years yes yet youd youll youre youve  
## 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0 0 0  
## 3 0 0 0 0 0 0 0 0 0 0 0 0  
## 4 0 0 0 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0

names(dataset)

## [1] "â€“" "â€”" "â€¢" "able"   
## [5] "absolutely" "accept" "acceptable" "accepted"   
## [9] "according" "account" "accurate" "across"   
## [13] "act" "action" "actions" "actual"   
## [17] "actually" "add" "added" "adding"   
## [21] "addition" "address" "admin" "administrator"   
## [25] "administrators" "admins" "advice" "afd"   
## [29] "ago" "agree" "allow" "allowed"   
## [33] "almost" "alone" "along" "already"   
## [37] "also" "although" "always" "american"   
## [41] "among" "another" "answer" "anyone"   
## [45] "anything" "anyway" "apparently" "appear"   
## [49] "appears" "appreciate" "appropriate" "area"   
## [53] "arent" "argument" "arguments" "around"   
## [57] "article" "articles" "ask" "asked"   
## [61] "asking" "ass" "assume" "attack"   
## [65] "attacks" "attempt" "attention" "august"   
## [69] "author" "automatically" "available" "avoid"   
## [73] "aware" "away" "back" "bad"   
## [77] "ban" "band" "banned" "based"   
## [81] "become" "behavior" "believe" "best"   
## [85] "better" "bias" "biased" "big"   
## [89] "bit" "black" "block" "blocked"   
## [93] "blocking" "book" "books" "box"   
## [97] "bring" "british" "call" "called"   
## [101] "calling" "came" "can" "cant"   
## [105] "care" "case" "cases" "category"   
## [109] "cause" "certain" "certainly" "chance"   
## [113] "change" "changed" "changes" "changing"   
## [117] "check" "cheers" "citation" "citations"   
## [121] "cite" "cited" "city" "claim"   
## [125] "claims" "clear" "clearly" "close"   
## [129] "come" "comes" "coming" "comment"   
## [133] "comments" "common" "community" "company"   
## [137] "complete" "completely" "concerns" "conflict"   
## [141] "consensus" "consider" "considered" "contact"   
## [145] "content" "contest" "context" "continue"   
## [149] "contribs" "contribute" "contributing" "contributions"   
## [153] "copy" "copyright" "correct" "country"   
## [157] "couple" "course" "create" "created"   
## [161] "creating" "criteria" "criticism" "current"   
## [165] "currently" "date" "day" "days"   
## [169] "deal" "dear" "death" "debate"   
## [173] "decide" "decided" "definition" "delete"   
## [177] "deleted" "deleting" "deletion" "described"   
## [181] "description" "despite" "details" "didnt"   
## [185] "difference" "different" "directly" "disagree"   
## [189] "discuss" "discussed" "discussion" "dispute"   
## [193] "disruptive" "doesnt" "done" "dont"   
## [197] "doubt" "due" "earlier" "early"   
## [201] "edit" "edited" "editing" "editor"   
## [205] "editors" "edits" "either" "else"   
## [209] "email" "encyclopedia" "end" "english"   
## [213] "enjoy" "enough" "entire" "entirely"   
## [217] "entry" "error" "especially" "etc"   
## [221] "even" "ever" "every" "everyone"   
## [225] "everything" "evidence" "exactly" "example"   
## [229] "except" "exist" "existing" "explain"   
## [233] "explaining" "explanation" "external" "fact"   
## [237] "facts" "fair" "faith" "false"   
## [241] "family" "far" "feel" "field"   
## [245] "file" "film" "finally" "find"   
## [249] "fine" "first" "five" "fix"   
## [253] "follow" "following" "form" "found"   
## [257] "four" "free" "friend" "friends"   
## [261] "fuck" "fucking" "full" "future"   
## [265] "game" "gave" "general" "generally"   
## [269] "get" "gets" "getting" "give"   
## [273] "given" "gives" "giving" "god"   
## [277] "goes" "going" "gone" "good"   
## [281] "google" "got" "government" "great"   
## [285] "group" "guess" "guidelines" "guy"   
## [289] "guys" "hand" "happen" "happened"   
## [293] "happy" "hard" "hate" "havent"   
## [297] "head" "heard" "hell" "hello"   
## [301] "help" "helpful" "hes" "hey"   
## [305] "high" "highly" "historical" "history"   
## [309] "hope" "hours" "however" "human"   
## [313] "idea" "ill" "image" "images"   
## [317] "important" "improve" "inappropriate" "include"   
## [321] "included" "including" "inclusion" "incorrect"   
## [325] "indeed" "indicate" "info" "information"   
## [329] "instead" "interest" "interested" "interesting"   
## [333] "internet" "involved" "irrelevant" "isnt"   
## [337] "issue" "issues" "ive" "january"   
## [341] "job" "john" "july" "june"   
## [345] "just" "keep" "kind" "know"   
## [349] "knowledge" "known" "knows" "lack"   
## [353] "language" "large" "last" "later"   
## [357] "law" "lead" "learn" "least"   
## [361] "leave" "left" "less" "let"   
## [365] "lets" "level" "life" "like"   
## [369] "likely" "line" "link" "linked"   
## [373] "links" "list" "listed" "little"   
## [377] "live" "living" "long" "longer"   
## [381] "look" "looked" "looking" "looks"   
## [385] "lost" "lot" "love" "made"   
## [389] "main" "major" "majority" "make"   
## [393] "makes" "making" "man" "manual"   
## [397] "many" "march" "material" "matter"   
## [401] "may" "maybe" "mean" "meaning"   
## [405] "means" "meant" "media" "members"   
## [409] "mention" "mentioned" "merely" "message"   
## [413] "messages" "might" "mind" "mine"   
## [417] "mistake" "months" "move" "moved"   
## [421] "much" "multiple" "music" "must"   
## [425] "name" "names" "national" "necessary"   
## [429] "need" "needed" "needs" "neither"   
## [433] "neutral" "never" "new" "news"   
## [437] "next" "nice" "none" "nonsense"   
## [441] "notability" "notable" "note" "nothing"   
## [445] "notice" "noticed" "now" "npov"   
## [449] "number" "obvious" "obviously" "official"   
## [453] "often" "okay" "old" "one"   
## [457] "ones" "online" "open" "opinion"   
## [461] "opinions" "order" "original" "others"   
## [465] "otherwise" "outside" "page" "pages"   
## [469] "paragraph" "part" "particular" "particularly"   
## [473] "party" "past" "people" "per"   
## [477] "perhaps" "period" "person" "personal"   
## [481] "personally" "picture" "piece" "pillars"   
## [485] "place" "placed" "play" "please"   
## [489] "point" "points" "policies" "policy"   
## [493] "political" "position" "possible" "possibly"   
## [497] "post" "posted" "posting" "pov"   
## [501] "power" "present" "pretty" "previous"   
## [505] "probably" "problem" "problems" "process"   
## [509] "produce" "project" "proof" "proper"   
## [513] "proposed" "prove" "provide" "provided"   
## [517] "public" "published" "purpose" "put"   
## [521] "putting" "quality" "question" "questions"   
## [525] "quite" "quote" "rather" "read"   
## [529] "readers" "reading" "real" "really"   
## [533] "reason" "reasons" "recent" "recently"   
## [537] "record" "redirect" "refer" "reference"   
## [541] "references" "regarding" "regards" "related"   
## [545] "release" "relevant" "reliable" "remember"   
## [549] "removal" "remove" "removed" "removing"   
## [553] "reply" "report" "request" "requesting"   
## [557] "research" "respect" "respond" "response"   
## [561] "rest" "result" "revert" "reverted"   
## [565] "reverting" "review" "right" "rights"   
## [569] "rule" "rules" "run" "said"   
## [573] "sandbox" "saw" "say" "saying"   
## [577] "says" "school" "science" "search"   
## [581] "second" "section" "sections" "see"   
## [585] "seem" "seems" "seen" "sense"   
## [589] "sentence" "separate" "series" "serious"   
## [593] "seriously" "set" "several" "shit"   
## [597] "short" "shouldnt" "show" "shows"   
## [601] "side" "sign" "significant" "similar"   
## [605] "simple" "simply" "since" "single"   
## [609] "site" "sites" "situation" "small"   
## [613] "someone" "something" "sometimes" "soon"   
## [617] "sorry" "sort" "source" "sourced"   
## [621] "sources" "speak" "specific" "specifically"   
## [625] "speedy" "standard" "start" "started"   
## [629] "state" "stated" "statement" "statements"   
## [633] "states" "status" "stay" "still"   
## [637] "stop" "story" "stuff" "stupid"   
## [641] "style" "subject" "subjects" "suggest"   
## [645] "summary" "support" "supposed" "sure"   
## [649] "system" "tag" "tagged" "tags"   
## [653] "take" "taken" "taking" "talk"   
## [657] "talking" "tell" "template" "term"   
## [661] "terms" "test" "text" "thank"   
## [665] "thanks" "thats" "theory" "therefore"   
## [669] "theres" "theyre" "thing" "things"   
## [673] "think" "thinking" "third" "though"   
## [677] "thought" "three" "thus" "tildes"   
## [681] "time" "times" "title" "today"   
## [685] "together" "told" "took" "top"   
## [689] "topic" "totally" "towards" "tried"   
## [693] "true" "truth" "try" "trying"   
## [697] "tutorial" "two" "type" "unblock"   
## [701] "understand" "understanding" "unfortunately" "united"   
## [705] "university" "unless" "uploaded" "upon"   
## [709] "use" "used" "useful" "user"   
## [713] "username" "users" "uses" "using"   
## [717] "usually" "utc" "valid" "vandalism"   
## [721] "vandalize" "various" "version" "video"   
## [725] "view" "views" "violation" "wait"   
## [729] "want" "wanted" "wants" "war"   
## [733] "warning" "wasnt" "way" "web"   
## [737] "website" "week" "welcome" "well"   
## [741] "went" "whatever" "whats" "whether"   
## [745] "white" "whole" "wiki" "wikipedia"   
## [749] "wikipedian" "wikipedias" "will" "wish"   
## [753] "within" "without" "wont" "word"   
## [757] "words" "work" "worked" "working"   
## [761] "works" "world" "worth" "wouldnt"   
## [765] "write" "writing" "written" "wrong"   
## [769] "wrote" "yeah" "year" "years"   
## [773] "yes" "yet" "youd" "youll"   
## [777] "youre" "youve"

dim(dataset)

## [1] 159571 778

dataset$toxic = NULL  
dataset$severe\_toxic = NULL  
dataset$obscene = NULL  
dataset$threat = NULL  
dataset$insult = NULL  
dataset$identity\_hate = NULL  
dim(dataset)

## [1] 159571 778

# attachh the output variables  
dataset$toxic = train$toxic  
dataset$severe\_toxic = train$severe\_toxic  
dataset$obscene = train$obscene  
dataset$threat = train$threat  
dataset$insult = train$insult  
dataset$identity\_hate = train$identity\_hate  
dim(dataset)

## [1] 159571 784

# create a multiclass variable  
attach(dataset)

## The following object is masked from package:base:  
##   
## version

dataset$target = ifelse(toxic==1,0,NA)  
dataset$target = ifelse(severe\_toxic==1,1,dataset$target)  
dataset$target = ifelse(obscene==1,2,dataset$target)  
dataset$target = ifelse(threat==1,3,dataset$target)  
dataset$target = ifelse(insult==1,4,dataset$target)  
dataset$target = ifelse(identity\_hate==1,5,dataset$target)  
dataset$target = ifelse(is.na(dataset$target),6,dataset$target)  
  
table(dataset$target)

##   
## 0 1 2 3 4 5 6   
## 5666 41 2233 163 6717 1405 143346

#remove na rows  
  
dataset = dataset[-which(dataset$target==6),]  
table(dataset$target)

##   
## 0 1 2 3 4 5   
## 5666 41 2233 163 6717 1405

dim(dataset)

## [1] 16225 785

# import the test datasets  
test = read.csv("C:/Users/User/Desktop/My old files/Recent folder Free desktop/Recent Analysis/josef/test.csv")  
test\_labels = read.csv("C:/Users/User/Desktop/My old files/Recent folder Free desktop/Recent Analysis/josef/test\_labels.csv")  
  
# clean the text  
  
corpus = VCorpus(VectorSource(test$comment\_text))  
corpus = tm\_map(corpus, content\_transformer(tolower))  
corpus = tm\_map(corpus, removeNumbers)  
corpus = tm\_map(corpus, removePunctuation)  
corpus = tm\_map(corpus, removeWords, stopwords())  
#corpus = tm\_map(corpus, stemDocument)  
corpus = tm\_map(corpus, stripWhitespace)  
  
# Creating the Bag of Words model  
dtm = DocumentTermMatrix(corpus)  
dtm = removeSparseTerms(dtm, 0.994)  
test\_dataset = as.data.frame(as.matrix(dtm))  
dim(test\_dataset)

## [1] 153164 685

# create a multiclass variable  
attach(test\_labels)

## The following objects are masked from dataset:  
##   
## identity\_hate, insult, obscene, severe\_toxic, threat, toxic

test\_dataset$target = ifelse(toxic==1,0,NA)  
test\_dataset$target = ifelse(severe\_toxic==1,1,test\_dataset$target)  
test\_dataset$target = ifelse(obscene==1,2,test\_dataset$target)  
test\_dataset$target = ifelse(threat==1,3,test\_dataset$target)  
test\_dataset$target = ifelse(insult==1,4,test\_dataset$target)  
test\_dataset$target = ifelse(identity\_hate==1,5,test\_dataset$target)  
test\_dataset$target = ifelse(is.na(test\_dataset$target),6,test\_dataset$target)  
table(test\_dataset$target)

##   
## 0 2 3 4 5 6   
## 1710 931 65 2825 712 146921

#remove na rows  
  
test\_dataset = test\_dataset[-which(test\_dataset$target==6),]  
table(test\_dataset$target)

##   
## 0 2 3 4 5   
## 1710 931 65 2825 712

dim(test\_dataset)

## [1] 6243 686

# match the variables of the traning\_set with those of the test\_dataset  
dim(test\_dataset)

## [1] 6243 686

dim(dataset)

## [1] 16225 785

common\_cols = intersect(colnames(dataset),colnames(test\_dataset))  
  
# create a dataset with 666 common variables with the test\_dataset and train dataset.  
  
  
dataset1= dataset[common\_cols]  
dim(dataset1)

## [1] 16225 666

test\_dataset1 = test\_dataset[common\_cols]  
dim(test\_dataset1)

## [1] 6243 666

# Encoding the target feature as factor  
dataset1$target = dataset$target  
dataset1$target = factor(dataset1$target)  
# Splitting the dataset into the Training set and Test set  
#library(caTools)  
library(caret)

## Loading required package: lattice

set.seed(123)  
split = createDataPartition(dataset1$target, p=0.7, list = FALSE)  
train\_set = dataset1[split,]  
test\_set = dataset1[-split,]  
dim(train\_set)

## [1] 11361 666

dim(test\_set)

## [1] 4864 666

# Applying PCA  
library(e1071)  
pca = preProcess(x = train\_set, method = 'pca', pcaComp = 5)  
train\_set = predict(pca, train\_set)  
test\_set = predict(pca, test\_set)  
test\_dataset2 = predict(pca, test\_dataset1)  
  
dim(train\_set)

## [1] 11361 6

dim(test\_set)

## [1] 4864 6

dim(test\_dataset2)

## [1] 6243 6

# Fitting Neural netweork to the Training set  
  
set.seed(300)  
ctrl = trainControl(method="cv",number=3)  
nn\_classifier = train(target~.,data=train\_set,method="nnet",trControl=ctrl,na.action=na.omit)

## # weights: 18  
## initial value 13716.570069   
## iter 10 value 9657.915686  
## iter 20 value 9577.388101  
## iter 30 value 9539.265965  
## iter 40 value 9536.944859  
## iter 50 value 9536.051762  
## iter 60 value 9532.071825  
## iter 70 value 9531.324862  
## final value 9531.315873   
## converged  
## # weights: 42  
## initial value 12920.789267   
## iter 10 value 9698.779659  
## iter 20 value 9539.097591  
## iter 30 value 9518.086977  
## iter 40 value 9506.131169  
## iter 50 value 9493.829427  
## iter 60 value 9490.433959  
## iter 70 value 9488.978797  
## iter 80 value 9488.406499  
## iter 90 value 9488.369327  
## iter 100 value 9488.196582  
## final value 9488.196582   
## stopped after 100 iterations  
## # weights: 66  
## initial value 17853.601314   
## iter 10 value 9556.426603  
## iter 20 value 9503.924827  
## iter 30 value 9492.573596  
## iter 40 value 9486.719560  
## iter 50 value 9482.903938  
## iter 60 value 9480.081611  
## iter 70 value 9477.753654  
## iter 80 value 9472.646555  
## iter 90 value 9470.188816  
## iter 100 value 9469.416451  
## final value 9469.416451   
## stopped after 100 iterations  
## # weights: 18  
## initial value 12477.145449   
## iter 10 value 9700.081816  
## iter 20 value 9587.376499  
## iter 30 value 9537.636666  
## iter 40 value 9535.498886  
## iter 50 value 9535.375747  
## iter 60 value 9534.437489  
## iter 70 value 9534.349434  
## iter 70 value 9534.349345  
## iter 70 value 9534.349344  
## final value 9534.349344   
## converged  
## # weights: 42  
## initial value 11210.807846   
## iter 10 value 9679.837234  
## iter 20 value 9545.452617  
## iter 30 value 9510.035287  
## iter 40 value 9505.425683  
## iter 50 value 9502.806559  
## iter 60 value 9501.657690  
## iter 70 value 9501.448149  
## iter 80 value 9501.277142  
## iter 90 value 9501.091178  
## iter 100 value 9501.067822  
## final value 9501.067822   
## stopped after 100 iterations  
## # weights: 66  
## initial value 17553.861335   
## iter 10 value 9621.960013  
## iter 20 value 9532.358934  
## iter 30 value 9510.799176  
## iter 40 value 9499.956384  
## iter 50 value 9493.892099  
## iter 60 value 9488.765283  
## iter 70 value 9485.758720  
## iter 80 value 9483.875337  
## iter 90 value 9481.833896  
## iter 100 value 9480.774007  
## final value 9480.774007   
## stopped after 100 iterations  
## # weights: 18  
## initial value 15415.088881   
## iter 10 value 10346.308281  
## iter 20 value 9735.563463  
## iter 30 value 9615.360820  
## iter 40 value 9612.360776  
## iter 50 value 9556.393625  
## iter 60 value 9547.568426  
## iter 70 value 9545.129619  
## iter 80 value 9536.582363  
## iter 90 value 9531.988055  
## iter 100 value 9531.330265  
## final value 9531.330265   
## stopped after 100 iterations  
## # weights: 42  
## initial value 14878.182522   
## iter 10 value 9655.984897  
## iter 20 value 9530.223864  
## iter 30 value 9505.907634  
## iter 40 value 9501.347156  
## iter 50 value 9496.708662  
## iter 60 value 9491.991395  
## iter 70 value 9489.094505  
## iter 80 value 9487.690058  
## iter 90 value 9487.190584  
## iter 100 value 9486.772507  
## final value 9486.772507   
## stopped after 100 iterations  
## # weights: 66  
## initial value 15607.887228   
## iter 10 value 9584.476838  
## iter 20 value 9514.558211  
## iter 30 value 9496.401330  
## iter 40 value 9491.103535  
## iter 50 value 9486.068483  
## iter 60 value 9479.345808  
## iter 70 value 9473.364550  
## iter 80 value 9469.608772  
## iter 90 value 9466.861287  
## iter 100 value 9464.358857  
## final value 9464.358857   
## stopped after 100 iterations  
## # weights: 18  
## initial value 14036.707167   
## iter 10 value 9995.553847  
## iter 20 value 9548.974902  
## iter 30 value 9541.293536  
## iter 40 value 9533.894314  
## iter 50 value 9533.176780  
## iter 60 value 9532.487839  
## iter 70 value 9532.223158  
## iter 80 value 9531.748417  
## iter 90 value 9530.485654  
## final value 9530.484645   
## converged  
## # weights: 42  
## initial value 15184.927759   
## iter 10 value 9542.203882  
## iter 20 value 9499.026270  
## iter 30 value 9485.369711  
## iter 40 value 9477.538245  
## iter 50 value 9470.811026  
## iter 60 value 9467.877748  
## iter 70 value 9466.795596  
## iter 80 value 9465.399146  
## iter 90 value 9462.059959  
## iter 100 value 9459.002666  
## final value 9459.002666   
## stopped after 100 iterations  
## # weights: 66  
## initial value 11962.970594   
## iter 10 value 9544.312702  
## iter 20 value 9494.308045  
## iter 30 value 9478.260986  
## iter 40 value 9470.080212  
## iter 50 value 9459.306190  
## iter 60 value 9454.411109  
## iter 70 value 9452.114443  
## iter 80 value 9447.597955  
## iter 90 value 9446.354344  
## iter 100 value 9444.896237  
## final value 9444.896237   
## stopped after 100 iterations  
## # weights: 18  
## initial value 16576.923432   
## iter 10 value 9757.202604  
## iter 20 value 9597.560327  
## iter 30 value 9543.390623  
## iter 40 value 9520.923903  
## iter 50 value 9508.791506  
## iter 60 value 9505.064081  
## final value 9504.699171   
## converged  
## # weights: 42  
## initial value 11660.588816   
## iter 10 value 9559.018785  
## iter 20 value 9504.908056  
## iter 30 value 9492.231660  
## iter 40 value 9482.516880  
## iter 50 value 9474.832086  
## iter 60 value 9472.842604  
## iter 70 value 9472.351728  
## iter 80 value 9472.221972  
## iter 90 value 9472.179579  
## iter 100 value 9472.168868  
## final value 9472.168868   
## stopped after 100 iterations  
## # weights: 66  
## initial value 11636.755859   
## iter 10 value 9667.371891  
## iter 20 value 9508.044214  
## iter 30 value 9487.116774  
## iter 40 value 9477.731339  
## iter 50 value 9469.340761  
## iter 60 value 9463.093290  
## iter 70 value 9460.650216  
## iter 80 value 9458.951622  
## iter 90 value 9458.293025  
## iter 100 value 9458.020778  
## final value 9458.020778   
## stopped after 100 iterations  
## # weights: 18  
## initial value 14920.551295   
## iter 10 value 9761.958260  
## iter 20 value 9578.332422  
## iter 30 value 9565.529540  
## iter 40 value 9561.648568  
## iter 50 value 9557.795079  
## iter 60 value 9554.724837  
## iter 70 value 9527.385551  
## iter 80 value 9513.104884  
## iter 90 value 9507.171126  
## iter 100 value 9502.560707  
## final value 9502.560707   
## stopped after 100 iterations  
## # weights: 42  
## initial value 15758.804565   
## iter 10 value 9578.409340  
## iter 20 value 9493.237600  
## iter 30 value 9483.967253  
## iter 40 value 9476.095096  
## iter 50 value 9471.578048  
## iter 60 value 9469.972146  
## iter 70 value 9469.095905  
## iter 80 value 9468.090658  
## iter 90 value 9466.760767  
## iter 100 value 9466.149659  
## final value 9466.149659   
## stopped after 100 iterations  
## # weights: 66  
## initial value 14585.413533   
## iter 10 value 9602.769354  
## iter 20 value 9515.858903  
## iter 30 value 9476.244373  
## iter 40 value 9465.380889  
## iter 50 value 9456.515488  
## iter 60 value 9449.792777  
## iter 70 value 9443.325025  
## iter 80 value 9437.992879  
## iter 90 value 9435.859630  
## iter 100 value 9434.998655  
## final value 9434.998655   
## stopped after 100 iterations  
## # weights: 18  
## initial value 16614.079567   
## iter 10 value 9915.052188  
## iter 20 value 9692.005022  
## iter 30 value 9688.557404  
## iter 40 value 9687.619890  
## iter 50 value 9684.451456  
## iter 60 value 9676.163837  
## iter 70 value 9626.699643  
## iter 80 value 9526.915982  
## iter 90 value 9506.298721  
## iter 100 value 9504.244223  
## final value 9504.244223   
## stopped after 100 iterations  
## # weights: 42  
## initial value 14267.511770   
## iter 10 value 9611.051650  
## iter 20 value 9502.240573  
## iter 30 value 9487.567496  
## iter 40 value 9478.840137  
## iter 50 value 9476.335775  
## iter 60 value 9475.654730  
## iter 70 value 9474.909717  
## iter 80 value 9472.129191  
## iter 90 value 9471.153285  
## iter 100 value 9469.585867  
## final value 9469.585867   
## stopped after 100 iterations  
## # weights: 66  
## initial value 17351.127144   
## iter 10 value 9535.273174  
## iter 20 value 9488.511774  
## iter 30 value 9472.757411  
## iter 40 value 9464.797448  
## iter 50 value 9459.866761  
## iter 60 value 9458.066390  
## iter 70 value 9456.506831  
## iter 80 value 9455.241378  
## iter 90 value 9453.872892  
## iter 100 value 9451.766269  
## final value 9451.766269   
## stopped after 100 iterations  
## # weights: 18  
## initial value 15435.738593   
## iter 10 value 9794.711672  
## iter 20 value 9614.467739  
## iter 30 value 9516.805403  
## iter 40 value 9508.648321  
## iter 50 value 9507.133639  
## iter 60 value 9506.254680  
## final value 9506.253653   
## converged  
## # weights: 42  
## initial value 13322.679825   
## iter 10 value 9597.398007  
## iter 20 value 9510.813387  
## iter 30 value 9493.303719  
## iter 40 value 9483.726535  
## iter 50 value 9480.650422  
## iter 60 value 9478.192894  
## iter 70 value 9476.433448  
## iter 80 value 9475.661714  
## iter 90 value 9475.482838  
## iter 100 value 9475.440025  
## final value 9475.440025   
## stopped after 100 iterations  
## # weights: 66  
## initial value 22193.449112   
## iter 10 value 9550.519188  
## iter 20 value 9493.057707  
## iter 30 value 9485.575348  
## iter 40 value 9479.715609  
## iter 50 value 9474.305498  
## iter 60 value 9470.169430  
## iter 70 value 9468.107319  
## iter 80 value 9466.654706  
## iter 90 value 9465.694688  
## iter 100 value 9464.341796  
## final value 9464.341796   
## stopped after 100 iterations  
## # weights: 18  
## initial value 13496.043075   
## iter 10 value 10448.290724  
## iter 20 value 9690.251569  
## iter 30 value 9690.232289  
## final value 9690.232078   
## converged  
## # weights: 42  
## initial value 15214.971496   
## iter 10 value 9593.568573  
## iter 20 value 9496.666807  
## iter 30 value 9485.548005  
## iter 40 value 9479.545302  
## iter 50 value 9472.575342  
## iter 60 value 9469.315168  
## iter 70 value 9468.547501  
## iter 80 value 9468.231715  
## iter 90 value 9468.082523  
## iter 100 value 9467.960981  
## final value 9467.960981   
## stopped after 100 iterations  
## # weights: 66  
## initial value 12429.001287   
## iter 10 value 9582.857618  
## iter 20 value 9500.176950  
## iter 30 value 9486.375223  
## iter 40 value 9473.109231  
## iter 50 value 9465.826034  
## iter 60 value 9457.460023  
## iter 70 value 9451.534821  
## iter 80 value 9447.687135  
## iter 90 value 9445.491675  
## iter 100 value 9444.514314  
## final value 9444.514314   
## stopped after 100 iterations  
## # weights: 42  
## initial value 21301.219172   
## iter 10 value 14397.231638  
## iter 20 value 14293.149431  
## iter 30 value 14264.124240  
## iter 40 value 14247.191610  
## iter 50 value 14240.336167  
## iter 60 value 14238.250986  
## iter 70 value 14236.969233  
## iter 80 value 14235.848979  
## iter 90 value 14235.394975  
## iter 100 value 14235.319125  
## final value 14235.319125   
## stopped after 100 iterations

y\_pred <- predict(nn\_classifier, newdata=test\_set[-1])  
#head(y\_pred)  
# Making the Confusion Matrix  
cm = table(test\_set[, 1], y\_pred)  
confusionMatrix(cm)

## Confusion Matrix and Statistics  
##   
## y\_pred  
## 0 1 2 3 4 5  
## 0 836 0 0 0 863 0  
## 1 2 0 0 0 10 0  
## 2 267 0 0 0 402 0  
## 3 8 0 0 0 40 0  
## 4 632 0 0 0 1383 0  
## 5 99 0 0 0 322 0  
##   
## Overall Statistics  
##   
## Accuracy : 0.4562   
## 95% CI : (0.4421, 0.4703)  
## No Information Rate : 0.6209   
## P-Value [Acc > NIR] : 1   
##   
## Kappa : 0.1091   
## Mcnemar's Test P-Value : NA   
##   
## Statistics by Class:  
##   
## Class: 0 Class: 1 Class: 2 Class: 3 Class: 4 Class: 5  
## Sensitivity 0.4534 NA NA NA 0.4579 NA  
## Specificity 0.7142 0.997533 0.8625 0.990132 0.6573 0.91345  
## Pos Pred Value 0.4921 NA NA NA 0.6864 NA  
## Neg Pred Value 0.6815 NA NA NA 0.4254 NA  
## Prevalence 0.3791 0.000000 0.0000 0.000000 0.6209 0.00000  
## Detection Rate 0.1719 0.000000 0.0000 0.000000 0.2843 0.00000  
## Detection Prevalence 0.3493 0.002467 0.1375 0.009868 0.4143 0.08655  
## Balanced Accuracy 0.5838 NA NA NA 0.5576 NA

# Fitting a naive baeys model to the Training set  
nb\_classifier = train(target~.,data=train\_set,method="nb",trControl=ctrl,na.action=na.omit)

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 98

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 114

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 122

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 129

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 154

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 199

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 234

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 302

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 304

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 319

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 338

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 381

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 493

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 580

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 600

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 688

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 847

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 858

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 879

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 909

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 930

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 949

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1042

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1081

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1087

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1116

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1178

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1272

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1335

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1381

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1390

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1452

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1487

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1524

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1527

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1532

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1538

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1548

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1555

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1559

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1621

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1646

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1659

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1673

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1971

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2076

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2129

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2156

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2175

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2191

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2196

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2207

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2247

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2256

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2286

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2307

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2391

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2451

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2557

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2627

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2730

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2736

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2872

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2883

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2941

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2942

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2949

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3102

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3126

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3299

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3392

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3415

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3427

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3458

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3474

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3519

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3747

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3786

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 98

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 114

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 122

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 129

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 154

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 199

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 234

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 302

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 319

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 338

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 381

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 580

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 681

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 847

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 930

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1042

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1087

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1272

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1335

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1381

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1452

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1487

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1527

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1548

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1555

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1621

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1634

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1646

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1659

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1673

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1887

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1971

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2032

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2129

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2156

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2196

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2247

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2286

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2451

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2557

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2627

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2730

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2872

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2883

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2941

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2949

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3102

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3126

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3299

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3392

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3427

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3458

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3519

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3747

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 100

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 162

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 289

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 313

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 347

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 415

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 462

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 668

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 674

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 694

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 772

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 836

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 948

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 962

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1023

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1060

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1061

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1092

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1131

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1162

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1207

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1212

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1233

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1245

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1252

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1254

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1282

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1297

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1325

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1407

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1419

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1450

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1464

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1583

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2011

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2127

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2249

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2270

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2427

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2650

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2753

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2985

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3030

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3032

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3299

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3473

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3476

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3561

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 162

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 223

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 289

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 347

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 415

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 462

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 531

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 618

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 668

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 694

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 772

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 836

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1023

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1060

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1061

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1131

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1233

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1245

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1282

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1297

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1325

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1407

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1419

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1450

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1507

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2011

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2249

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2427

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2650

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2985

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3032

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3299

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3473

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3561

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 112

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 118

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 120

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 175

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 195

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 247

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 311

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 442

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 454

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 540

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 608

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 623

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 642

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 679

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 715

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 778

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1021

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1101

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1162

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1173

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1182

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1185

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1237

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1374

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1396

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1458

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1531

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1557

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1770

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1848

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1909

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1951

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2047

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2070

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2213

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2229

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2260

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2295

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2414

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2429

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2671

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2718

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2740

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2881

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2912

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2922

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2970

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3003

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3027

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3062

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3089

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3121

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3132

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3134

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3142

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3188

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3237

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3325

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3397

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3422

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3523

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3597

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3617

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3635

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 112

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 120

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 175

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 195

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 247

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 311

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 442

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 454

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 608

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 623

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 642

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 679

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 715

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1021

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1101

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1162

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1173

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1182

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1185

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1374

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1396

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1458

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1557

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1951

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2047

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2070

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2229

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2260

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2295

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2414

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2431

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2718

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2740

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2881

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2922

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3003

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3027

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3062

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3121

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3132

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3188

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3397

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3422

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3523

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3635

# Fitting a random forest to the Training set  
#rf\_classifier = train(target~.,data=train\_set,method="rf",trControl=ctrl,na.action=na.omit)  
  
  
# Predicting the Test set results  
#which(colnames(test\_set)=="target")  
y\_pred <- predict(nb\_classifier, newdata=test\_set[-1])

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 114

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 184

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 186

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 208

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 240

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 293

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 354

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 420

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 458

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 473

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 601

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 625

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 632

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 670

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 845

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 863

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1002

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1136

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1207

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1243

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1455

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1466

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1496

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1514

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1546

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1575

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1694

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1720

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1728

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2104

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2124

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2185

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2293

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2306

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2323

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2367

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2395

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2434

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2500

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2538

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2600

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2605

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2622

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2685

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2772

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2787

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2790

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2908

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2941

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3069

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3254

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3306

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3327

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3547

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3558

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3586

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3601

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3654

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3770

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3821

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3822

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3876

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3911

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3998

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4076

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4142

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4315

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4416

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4472

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4523

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4547

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4813

#head(y\_pred)  
# Making the Confusion Matrix  
cm = table(test\_set[, 1], y\_pred)  
confusionMatrix(cm)

## Confusion Matrix and Statistics  
##   
## y\_pred  
## 0 1 2 3 4 5  
## 0 842 4 9 0 844 0  
## 1 5 0 0 0 7 0  
## 2 260 1 1 0 407 0  
## 3 14 0 0 0 34 0  
## 4 695 4 5 0 1311 0  
## 5 108 0 2 0 311 0  
##   
## Overall Statistics  
##   
## Accuracy : 0.4428   
## 95% CI : (0.4288, 0.4569)  
## No Information Rate : 0.5991   
## P-Value [Acc > NIR] : 1   
##   
## Kappa : 0.0913   
## Mcnemar's Test P-Value : NA   
##   
## Statistics by Class:  
##   
## Class: 0 Class: 1 Class: 2 Class: 3 Class: 4  
## Sensitivity 0.4376 0.000000 0.0588235 NA 0.4499  
## Specificity 0.7085 0.997528 0.8621828 0.990132 0.6390  
## Pos Pred Value 0.4956 0.000000 0.0014948 NA 0.6506  
## Neg Pred Value 0.6581 0.998145 0.9961859 NA 0.4373  
## Prevalence 0.3956 0.001850 0.0034951 0.000000 0.5991  
## Detection Rate 0.1731 0.000000 0.0002056 0.000000 0.2695  
## Detection Prevalence 0.3493 0.002467 0.1375411 0.009868 0.4143  
## Balanced Accuracy 0.5731 0.498764 0.4605032 NA 0.5444  
## Class: 5  
## Sensitivity NA  
## Specificity 0.91345  
## Pos Pred Value NA  
## Neg Pred Value NA  
## Prevalence 0.00000  
## Detection Rate 0.00000  
## Detection Prevalence 0.08655  
## Balanced Accuracy NA

# predict on the test dataset  
y\_pred = predict(nb\_classifier, newdata=test\_dataset2,type="prob")

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 292

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 619

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 625

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 909

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 964

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1012

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1115

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1237

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1336

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1570

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1758

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1782

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 1838

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2342

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2823

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 2921

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3380

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3447

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3512

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3542

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3780

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3859

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 3966

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4032

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4081

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4133

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4198

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4457

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4480

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4725

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4788

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 4846

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 5021

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 5269

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 5538

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 6087

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 6146

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with  
## observation 6167

#y\_pred = predict(xgb\_classifier, newdata=data.matrix(test\_dataset2),reshape=TRUE)  
#y\_pred = as.matrix(y\_pred)  
#head(y\_pred)