Text mining in R

Nik Bear Brown

In this lesson we'll learn regular expressions and how to use the 'tm' package in R.

# Additional packages needed

To run the code in M08\_Lesson\_02.Rmd you may need additional packages.

* If necessary install the followings packages.

install.packages("RTextTools");  
install.packages("tm");  
install.packages("wordcloud);

library(RTextTools)

## Loading required package: SparseM  
##   
## Attaching package: 'SparseM'  
##   
## The following object is masked from 'package:base':  
##   
## backsolve

library(tm)

## Loading required package: NLP

library(wordcloud)

## Loading required package: RColorBrewer

# Data

We will be using the [Wisconsin Diagnostic Breast Cancer Dataset from UC Irvine](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)), as well as Dr Suess quotes.

# Regular Expressions and Grep

In theoretical computer science and formal language theory, a [regular expression](https://en.wikipedia.org/wiki/Regular_expression) (abbreviated regex or regexp and sometimes called a rational expression) is a sequence of characters that define a search pattern, mainly for use in pattern matching with strings, or string matching, i.e. "find and replace"-like operations. The concept arose in the 1950s, when the American mathematician Stephen Kleene formalized the description of a regular language, and came into common use with the Unix text processing utilities ed, an editor, and [grep](https://en.wikipedia.org/wiki/Grep) (global regular expression print), a filter.

grep is a command-line utility for searching plain-text data sets for lines matching a regular expression. Grep was originally developed for the Unix operating system, but is available today for all Unix-like systems and is built in to languages like python and Perl.

# Regular Expressions Examples

Basic regex syntax

. Normally matches any character except a newline.   
  
When you match a pattern within parentheses, you can use any of $1, $2, ... later to refer to the previously matched pattern.   
  
+ Matches the preceding pattern element one or more times.   
? Matches the preceding pattern element zero or one times.   
\* Matches the preceding pattern element zero or more times.   
| Separates alternate possibilities.   
  
\w Matches an alphanumeric character, including "\_"; same as [A-Za-z0-9\_] in ASCII, and  
[\p{Alphabetic}\p{GC=Mark}\p{GC=Decimal\_Number}\p{GC=Connector\_Punctuation}]   
  
\W Matches a non-alphanumeric character, excluding "\_";  
same as [^A-Za-z0-9\_] in ASCII, and  
[^\p{Alphabetic}\p{GC=Mark}\p{GC=Decimal\_Number}\p{GC=Connector\_Punctuation}]   
  
\s Matches a whitespace character,  
which in ASCII are tab, line feed, form feed, carriage return, and space;   
  
\S Matches anything BUT a whitespace.   
  
\d Matches a digit;  
same as [0-9] in ASCII;   
  
\D Matches a non-digit;  
  
^ Matches the beginning of a line or string.   
  
$ Matches the end of a line or string.

Some simple regex examples

{^[-+]?[0-9]\*\.?[0-9]+([eE][-+]?[0-9]+)?$} # Floating Point Number  
  
{^[A-Za-z]+$} # Only letters.  
  
 {^[[:alpha?:]]+$} # Only letters, the Unicode way.  
  
 {(.)\1{3}} $string {\1} result # Back References   
  
(\[0-9]{1,3})\.(\[0-9]{1,3})\.(\[0-9]{1,3})\.(\[0-9]{1,3}) # IP Numbers

# Regular Expressions in R

letters

## [1] "a" "b" "c" "d" "e" "f" "g" "h" "i" "j" "k" "l" "m" "n" "o" "p" "q"  
## [18] "r" "s" "t" "u" "v" "w" "x" "y" "z"

grep("[a-z]", letters)

## [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23  
## [24] 24 25 26

grep("[A-Z]", letters)

## integer(0)

grep("[AB]", letters)

## integer(0)

grep("[AB]", letters,ignore.case = TRUE)

## [1] 1 2

grep("[a-zA-Z]", letters)

## [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23  
## [24] 24 25 26

grep("[azAZ]", letters)

## [1] 1 26

seuss <- c("You have brains in your head.",  
 "You have feet in your shoes.",   
 "You can steer yourself any direction you choose.",   
 "You're on your own.",   
 "And you know what you know.",   
 "And YOU are the one who'll decide where to go...",   
 "- Dr. Seuss")  
grep("you", seuss)

## [1] 1 2 3 4 5

if(length(i <- grep("in", seuss)))  
 cat("'in' appears at least once in\n\t", seuss[i], "\n")

## 'in' appears at least once in  
## You have brains in your head. You have feet in your shoes.

i # 1 and 2

## [1] 1 2

seuss[i]

## [1] "You have brains in your head." "You have feet in your shoes."

## Modify all 'a' or 'b's; "\" must be escaped   
s<-gsub("([ab])", "\\1-z\\1\_s", "abc vs, ABC - Oh, A B C. It's easy as, 1 2 3. As simple as, do re mi. a b c, 1 2 3")  
s

## [1] "a-za\_sb-zb\_sc vs, ABC - Oh, A B C. It's ea-za\_ssy a-za\_ss, 1 2 3. As simple a-za\_ss, do re mi. a-za\_s b-zb\_s c, 1 2 3"

s<-gsub("([ab])", "-z\_s", "abc vs, ABC - Oh, A B C. It's easy as, 1 2 3. As simple as, do re mi. a b c, 1 2 3")  
s

## [1] "-z\_s-z\_sc vs, ABC - Oh, A B C. It's e-z\_ssy -z\_ss, 1 2 3. As simple -z\_ss, do re mi. -z\_s -z\_s c, 1 2 3"

s<-gsub("([ab])", "\\a\\b", "abc vs, ABC - Oh, A B C. It's easy as, 1 2 3. As simple as, do re mi. a b c, 1 2 3")  
s

## [1] "ababc vs, ABC - Oh, A B C. It's eabsy abs, 1 2 3. As simple abs, do re mi. ab ab c, 1 2 3"

stop.words <- c("The", "for", "are",  
 "to", "your",   
 "to", "and", "the",   
 "is", "your", "and", "is",  
 "it", "for", "all", "its", "it's")  
  
seuss.re<-regexpr("[ ]+",seuss)  
seuss.re

## [1] 4 4 4 7 4 4 2  
## attr(,"match.length")  
## [1] 1 1 1 1 1 1 1  
## attr(,"useBytes")  
## [1] TRUE

seuss.re<-gregexpr("[ ]+",seuss)  
seuss.re

## [[1]]  
## [1] 4 9 16 19 24  
## attr(,"match.length")  
## [1] 1 1 1 1 1  
## attr(,"useBytes")  
## [1] TRUE  
##   
## [[2]]  
## [1] 4 9 14 17 22  
## attr(,"match.length")  
## [1] 1 1 1 1 1  
## attr(,"useBytes")  
## [1] TRUE  
##   
## [[3]]  
## [1] 4 8 14 23 27 37 41  
## attr(,"match.length")  
## [1] 1 1 1 1 1 1 1  
## attr(,"useBytes")  
## [1] TRUE  
##   
## [[4]]  
## [1] 7 10 15  
## attr(,"match.length")  
## [1] 1 1 1  
## attr(,"useBytes")  
## [1] TRUE  
##   
## [[5]]  
## [1] 4 8 13 18 22  
## attr(,"match.length")  
## [1] 1 1 1 1 1  
## attr(,"useBytes")  
## [1] TRUE  
##   
## [[6]]  
## [1] 4 8 12 16 20 27 34 40 43  
## attr(,"match.length")  
## [1] 1 1 1 1 1 1 1 1 1  
## attr(,"useBytes")  
## [1] TRUE  
##   
## [[7]]  
## [1] 2 6  
## attr(,"match.length")  
## [1] 1 1  
## attr(,"useBytes")  
## [1] TRUE

## trim trailing white space  
str <- " You have brains in your head "  
str

## [1] " You have brains in your head "

str<-sub("[ ]+$", "", str) ## trailing spaces only  
str

## [1] " You have brains in your head"

str<-sub("^[ ]+", "", str) ## leading spaces only  
str

## [1] "You have brains in your head"

sub("[[:space:]]+$", "", str) ## white space, POSIX-style

## [1] "You have brains in your head"

sub("\\s+$", "", str, perl = TRUE) ## Perl-style white space

## [1] "You have brains in your head"

## capitalizing  
txt <- "You have brains in your head"  
gsub("(\\w)(\\w\*)", "\\U\\1\\L\\2", txt, perl=TRUE)

## [1] "You Have Brains In Your Head"

gsub("\\b(\\w)", "\\U\\1", txt, perl=TRUE)

## [1] "You Have Brains In Your Head"

txt2 <- "You have feet in your shoes."  
gsub("(\\w)(\\w\*)(\\w)", "\\U\\1\\E\\2\\U\\3", txt2, perl=TRUE)

## [1] "YoU HavE FeeT IN YouR ShoeS."

sub("(\\w)(\\w\*)(\\w)", "\\U\\1\\E\\2\\U\\3", txt2, perl=TRUE)

## [1] "YoU have feet in your shoes."

## Decompose a URL into its components.  
url <- "http://nikbearbrown/machine/learning/"  
m <- regexec("^(([^:]+)://)?([^:/]+)(:([0-9]+))?(/.\*)", url)  
m

## [[1]]  
## [1] 1 1 1 8 0 0 20  
## attr(,"match.length")  
## [1] 37 7 4 12 0 0 18

regmatches(url, m)

## [[1]]  
## [1] "http://nikbearbrown/machine/learning/"  
## [2] "http://"   
## [3] "http"   
## [4] "nikbearbrown"   
## [5] ""   
## [6] ""   
## [7] "/machine/learning/"

## parts of a URL:  
URL\_parts <- function(u) {  
 m <- regexec("^(([^:]+)://)?([^:/]+)(:([0-9]+))?(/.\*)", u)  
 parts <- do.call(rbind,  
 lapply(regmatches(u, m), `[`, c(3L, 4L, 6L, 7L)))  
 colnames(parts) <- c("protocol","host","port","path")  
 parts  
}  
URL\_parts(url)

## protocol host port path   
## [1,] "http" "nikbearbrown" "" "/machine/learning/"

# Spilt text  
seuss

## [1] "You have brains in your head."   
## [2] "You have feet in your shoes."   
## [3] "You can steer yourself any direction you choose."  
## [4] "You're on your own."   
## [5] "And you know what you know."   
## [6] "And YOU are the one who'll decide where to go..."  
## [7] "- Dr. Seuss"

seuss[1]

## [1] "You have brains in your head."

sp<-strsplit(seuss[1], "a") # "You have brains in your head."  
s<-"You have brains in your head. "  
sp<-strsplit(seuss[1], " ")   
sp

## [[1]]  
## [1] "You" "have" "brains" "in" "your" "head."

sp<-strsplit(s," ")   
sp

## [[1]]  
## [1] "You" "" "" "" "have" "" ""   
## [8] "brains" "" "in" "" "your" "head."

sp<-strsplit(s,"[ ]+")   
sp

## [[1]]  
## [1] "You" "have" "brains" "in" "your" "head."

sp<-strsplit(s,"[ ]+",perl = TRUE)   
sp

## [[1]]  
## [1] "You" "have" "brains" "in" "your" "head."

# Text Mining in R (tm package)

The main structure for managing documents in tm is a so-called Corpus, representing a collection of text documents. Custom print() methods are available which hide the raw amount of information since displaying a lot of text can hide other data.

## Transformations

Once we have a corpus we typically want to modify the documents in it, e.g., stemming, stopword removal, et cetera. In tm, all this functionality is subsumed into the concept of a transformation. Transformations are done via the tm\_map() function which applies (maps) a function to all elements of the corpus.

## Creating Term-Document Matrices

A common approach in text mining is to create a term-document matrix from a corpus. In the tm package the classes TermDocumentMatrix and DocumentTermMatrix (depending on whether you want terms as rows and documents as columns, or vice versa) employ sparse matrices for corpora.

## Functions for Term-Document Matrices

The tm package has a number of operations for Term-Document Matrices (like clustering, classifications, etc.).

For example, given a Term-Document matrix called

* findFreqTerms(tdm, 5) # finds terms with all least 5 counts
* findAssocs(tdm, "brain", 0.8) # finds associations (i.e., terms which correlate) with at least 0.8 correlation for the term "brain"

# Introduction to the tm Package  
seuss <- c("You have brains in your head.",  
 "You have feet in your shoes.",   
 "You can steer yourself any direction you choose.",   
 "You're on your own.",   
 "And you know what you know.",   
 "And YOU are the one who'll decide where to go...",   
 "Today you are you, that is truer than true. There is no one alive who is youer than you. - Dr. Seuss",  
 "Don't cry because it's over. Smile because it happened. - Dr. Seuss",   
 "Unless someone like you cares a whole awful lot, nothing is going to get better. It's not. - Dr. Seuss",  
 "A person's a person, no matter how small. - Dr. Seuss",   
 "Today is your day! Your mountain is waiting so get on your way! - Dr. Seuss",   
 "You're never too old, too wacky, too wild, to pick up a book and read to a child. - Dr. Seuss",  
 "The more that you read, the more things you will know. The more that you learn, the more places you'll go. - Dr. Seuss",  
 "Oh the things you can find if you don't stay behind! - Dr. Seuss",  
 "And turtles, of course. all the turtles are FREE. As turtles and, maybe, ALL creatures should be. - Dr. Seuss",  
 "With your head full of brains and your shoes full of feet, you're too smart to go down any not-so-good street. - Dr. Seuss",  
 "You're on your own. And you know what you know. And you're the one who'll decide where to go. - Dr. Seuss",   
 "Sometimes the questions are complicated and the answers are simple. - Dr. Seuss",  
 "You do not like them. So you say. Try them! Try them! And you may! - Dr. Seuss",  
 "You have brains in your head, you have feet in your shoes. You can steer yourself any direction you choose. - Dr. Seuss",  
 "If things start happening, don't worry. Don't stew. Just go right along. You'll start happening too. - Dr. Seuss",   
 "So be sure when you step. Step with care and great tact and remember that life's a Great Balancing Act. - Dr. Seuss",   
 "If you never did you should. These things are fun, and fun is good. - Dr. Seuss",  
 "I have heard there are troubles of more than one kind. Some come from ahead and some come from behind. But I've bought a big bat. I'm all ready you see. Now my troubles are going to have troubles with me! - Dr. Seuss",  
 "Will you succeed? Yes! You will, indeed! (98 and 3/4 percent guaranteed. - Dr. Seuss",   
 "You're off to Great Places!",  
 "Today is your day!",  
 "Your mountain is waiting,",  
 "So... get on your way!",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "So be sure when you step, Step with care and great tact. And remember that life's A Great Balancing Act. And will you succeed? Yes! You will, indeed! ( and ? percent guaranteed) Kid, you'll move mountains.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "You'll get mixed up, of course, as you already know. You'll get mixed up with many strange birds as you go. So be sure when you step. Step with care and great tact and remember that Life's a Great Balancing Act. Just never forget to be dexterous and deft. And never mix up your right foot with your left.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "Oh the places you'll go! There is fun to be done! There are points to be scored. There are games to be won. And the magical things you can do with that ball will make you the winning-est winner of all.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "Things may happen and often do to people as brainy and footsy as you",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "Congratulations!",  
 "Today is your day.",  
 "You're off to Great Places!",  
 "You're off and away!",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "It's opener, out there, in the wide, open air.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "You have brains in your head.",  
 "You have feet in your shoes.",  
 "You can steer yourself any direction you choose.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "And when you're alone there's a very good chance",  
 "you'll meet things that scare you right out of your pants",  
 "There are some, down the road between hither and yon,",  
 "that can scare you so much you won't want to go on.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "You will come to a place where the streets are not marked.",  
 "Some windows are lighted. but mostly they're darked.",  
 "But mostly they're darked.",  
 "A place you could sprain both your elbow and chin!",  
 "Do you dare to stay out? Do you dare to go in?",  
 "How much can you lose? How much can you win?",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "Fame you'll be famous, as famous as can be, with everyone watching you win on TV, Except when they don't because sometimes they won't..",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "When you're in a Slump,",  
 "you're not in for much fun.",  
 "Un-slumping yourself",  
 "is not easily done.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "You won't lag behind, because you'll have the speed.",  
 "You'll pass the whole gang and you'll soon take the lead.",  
 "Wherever you fly, you'll be best of the best.",  
 "Wherever you go, you will top all the rest.",  
 "Except when you don't.",  
 "Because, sometimes, you won't.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "Onward up many a frightening creek, though your arms may get sore and your sneakers may leak. Oh! The places you'll go!",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "So...be your name Buxbaum or Bixby or Bray",  
 "or Mordecai Ali Van Allen O'Shea,",  
 "you're off to Great Places!",  
 "Today is your day!",  
 "Your mountain is waiting.",  
 "So...get on your way!",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "And will you succeed?",  
 "Yes! You will, indeed!",  
 "( (98 and 3/4 percent guaranteed.)",  
 "KID, YOU'LL MOVE MOUNTAINS!",  
 "Out there things can happen, and frequently do,",  
 "To people as brainy and footsy as you.",  
 "And when things start to happen, don't worry, don't stew.",  
 "Just go right along, you'll start happening too!",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "You can get so confused",  
 "that you'll start in to race",  
 "down long wiggled roads at a break-necking pace",  
 "and grind on for miles across weirdish wild space,",  
 "headed, I fear, toward a most useless place.",  
 "The Waiting Place...",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "Kid, you'll move mountains.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",  
 "Oh, the places you'll go! There is fun to be done!",  
 "There are points to be scored. There are games to be won.",  
 "And the magical things you can do with that ball",  
 "will make you the winning-est winner of all.",  
 "Fame! You'll be as famous as famous can be,",  
 "with the whole wide world watching you win on TV.",  
 "Except when they don't",  
 "Because, sometimes they won't.",  
 "I'm afraid that some times",  
 "you'll play lonely games too.",  
 "Games you can't win",  
 "'cause you'll play against you.",  
 "-- Dr. Seuss, Oh, The Places You'll Go!",   
 "- Dr. Seuss")  
seuss.corpus <- Corpus(DataframeSource(data.frame(seuss)))  
seuss.corpus

## <<VCorpus>>  
## Metadata: corpus specific: 0, document level (indexed): 0  
## Content: documents: 116

inspect(seuss.corpus)

## <<VCorpus>>  
## Metadata: corpus specific: 0, document level (indexed): 0  
## Content: documents: 116  
##   
## [[1]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 29  
##   
## [[2]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 28  
##   
## [[3]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 48  
##   
## [[4]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 19  
##   
## [[5]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 27  
##   
## [[6]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 48  
##   
## [[7]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 100  
##   
## [[8]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 67  
##   
## [[9]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 102  
##   
## [[10]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 53  
##   
## [[11]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 75  
##   
## [[12]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 93  
##   
## [[13]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 118  
##   
## [[14]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 64  
##   
## [[15]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 109  
##   
## [[16]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 122  
##   
## [[17]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 105  
##   
## [[18]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 79  
##   
## [[19]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 78  
##   
## [[20]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 119  
##   
## [[21]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 112  
##   
## [[22]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 115  
##   
## [[23]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 79  
##   
## [[24]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 216  
##   
## [[25]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 84  
##   
## [[26]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 27  
##   
## [[27]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 18  
##   
## [[28]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 25  
##   
## [[29]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 22  
##   
## [[30]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[31]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 205  
##   
## [[32]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[33]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 304  
##   
## [[34]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[35]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 201  
##   
## [[36]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[37]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 68  
##   
## [[38]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[39]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 16  
##   
## [[40]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 18  
##   
## [[41]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 27  
##   
## [[42]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 20  
##   
## [[43]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[44]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 46  
##   
## [[45]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[46]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 29  
##   
## [[47]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 28  
##   
## [[48]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 48  
##   
## [[49]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[50]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 48  
##   
## [[51]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 57  
##   
## [[52]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 53  
##   
## [[53]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 51  
##   
## [[54]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[55]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 58  
##   
## [[56]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 52  
##   
## [[57]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 26  
##   
## [[58]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 50  
##   
## [[59]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 46  
##   
## [[60]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 44  
##   
## [[61]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[62]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 135  
##   
## [[63]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[64]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 23  
##   
## [[65]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 27  
##   
## [[66]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 20  
##   
## [[67]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 19  
##   
## [[68]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[69]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 52  
##   
## [[70]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 57  
##   
## [[71]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 45  
##   
## [[72]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 43  
##   
## [[73]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 22  
##   
## [[74]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 30  
##   
## [[75]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[76]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 119  
##   
## [[77]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[78]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 42  
##   
## [[79]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 33  
##   
## [[80]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 27  
##   
## [[81]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 18  
##   
## [[82]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 25  
##   
## [[83]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 21  
##   
## [[84]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[85]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 21  
##   
## [[86]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 22  
##   
## [[87]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 34  
##   
## [[88]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 27  
##   
## [[89]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 47  
##   
## [[90]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 38  
##   
## [[91]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 57  
##   
## [[92]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 48  
##   
## [[93]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[94]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 23  
##   
## [[95]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 28  
##   
## [[96]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 47  
##   
## [[97]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 50  
##   
## [[98]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 44  
##   
## [[99]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 20  
##   
## [[100]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[101]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 27  
##   
## [[102]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[103]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 50  
##   
## [[104]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 57  
##   
## [[105]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 48  
##   
## [[106]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 44  
##   
## [[107]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 43  
##   
## [[108]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 49  
##   
## [[109]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 22  
##   
## [[110]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 30  
##   
## [[111]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 26  
##   
## [[112]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 29  
##   
## [[113]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 19  
##   
## [[114]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 31  
##   
## [[115]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 39  
##   
## [[116]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 11

seuss.corpus[1]

## <<VCorpus>>  
## Metadata: corpus specific: 0, document level (indexed): 0  
## Content: documents: 1

writeLines(as.character(seuss.corpus[1]))

## list(list(content = "You have brains in your head.", meta = list(author = character(0), datetimestamp = list(sec = 16.2074890136719, min = 10, hour = 4, mday = 13, mon = 9, year = 115, wday = 2, yday = 285, isdst = 0), description = character(0), heading = character(0), id = "1", language = "en", origin = character(0))))  
## list()  
## list()

writeLines(as.character(seuss.corpus[1:3]))

## list(list(content = "You have brains in your head.", meta = list(author = character(0), datetimestamp = list(sec = 16.2074890136719, min = 10, hour = 4, mday = 13, mon = 9, year = 115, wday = 2, yday = 285, isdst = 0), description = character(0), heading = character(0), id = "1", language = "en", origin = character(0))), list(content = "You have feet in your shoes.", meta = list(author = character(0), datetimestamp = list(sec = 16.2113900184631, min = 10, hour = 4, mday = 13, mon = 9, year = 115,   
## wday = 2, yday = 285, isdst = 0), description = character(0), heading = character(0), id = "2", language = "en", origin = character(0))), list(content = "You can steer yourself any direction you choose.", meta = list(author = character(0), datetimestamp = list(sec = 16.2115540504456, min = 10, hour = 4, mday = 13, mon = 9, year = 115, wday = 2, yday = 285, isdst = 0), description = character(0), heading = character(0), id = "3", language = "en", origin = character(0))))  
## list()  
## list()

# Eliminating Extra Whitespace  
seuss.clean<-tm\_map(seuss.corpus, stripWhitespace)  
# stemDocument  
seuss.clean.stem<-tm\_map(seuss.clean, stemDocument)

## Warning in mclapply(content(x), FUN, ...): all scheduled cores encountered  
## errors in user code

writeLines(as.character(seuss.clean.stem[1]))

## list("Error in loadNamespace(name) : there is no package called 'SnowballC'\n")  
## list()  
## list()

# "You have brains in your head."  
# Convert to Lower Case  
seuss.clean.lc <- tm\_map(seuss.clean, content\_transformer(tolower))  
writeLines(as.character(seuss.clean.lc[1]))

## list(list(content = "you have brains in your head.", meta = list(author = character(0), datetimestamp = list(sec = 16.2074890136719, min = 10, hour = 4, mday = 13, mon = 9, year = 115, wday = 2, yday = 285, isdst = 0), description = character(0), heading = character(0), id = "1", language = "en", origin = character(0))))  
## list()  
## list()

# Remove Stopwords  
seuss.clean <- tm\_map(seuss.clean.lc, removeWords, stopwords("english"))  
writeLines(as.character(seuss.clean.lc[1]))

## list(list(content = "you have brains in your head.", meta = list(author = character(0), datetimestamp = list(sec = 16.2074890136719, min = 10, hour = 4, mday = 13, mon = 9, year = 115, wday = 2, yday = 285, isdst = 0), description = character(0), heading = character(0), id = "1", language = "en", origin = character(0))))  
## list()  
## list()

# Building a Document-Term Matrix  
seuss.tdm <- TermDocumentMatrix(seuss.clean, control = list(minWordLength = 1))  
seuss.tdm

## <<TermDocumentMatrix (terms: 277, documents: 116)>>  
## Non-/sparse entries: 577/31555  
## Sparsity : 98%  
## Maximal term length: 16  
## Weighting : term frequency (tf)

inspect(seuss.tdm[11:33,1:33])

## <<TermDocumentMatrix (terms: 23, documents: 33)>>  
## Non-/sparse entries: 14/745  
## Sparsity : 98%  
## Maximal term length: 9  
## Weighting : term frequency (tf)  
##   
## Docs  
## Terms 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24  
## ahead 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1  
## air. 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## ali 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## alive 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## allen 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## alone 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## along, 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## along. 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0  
## already 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## answers 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0  
## arms 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## away! 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## awful 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## balancing 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0  
## ball 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## bat. 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1  
## behind, 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## behind! 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0  
## behind. 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1  
## best 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## best. 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## better. 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## big 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1  
## Docs  
## Terms 25 26 27 28 29 30 31 32 33  
## ahead 0 0 0 0 0 0 0 0 0  
## air. 0 0 0 0 0 0 0 0 0  
## ali 0 0 0 0 0 0 0 0 0  
## alive 0 0 0 0 0 0 0 0 0  
## allen 0 0 0 0 0 0 0 0 0  
## alone 0 0 0 0 0 0 0 0 0  
## along, 0 0 0 0 0 0 0 0 0  
## along. 0 0 0 0 0 0 0 0 0  
## already 0 0 0 0 0 0 0 0 1  
## answers 0 0 0 0 0 0 0 0 0  
## arms 0 0 0 0 0 0 0 0 0  
## away! 0 0 0 0 0 0 0 0 0  
## awful 0 0 0 0 0 0 0 0 0  
## balancing 0 0 0 0 0 0 1 0 1  
## ball 0 0 0 0 0 0 0 0 0  
## bat. 0 0 0 0 0 0 0 0 0  
## behind, 0 0 0 0 0 0 0 0 0  
## behind! 0 0 0 0 0 0 0 0 0  
## behind. 0 0 0 0 0 0 0 0 0  
## best 0 0 0 0 0 0 0 0 0  
## best. 0 0 0 0 0 0 0 0 0  
## better. 0 0 0 0 0 0 0 0 0  
## big 0 0 0 0 0 0 0 0 0

# Operations on Term-Document Matrices  
# Frequent Terms and Associations  
findFreqTerms(seuss.tdm, lowfreq=3)

## [1] "act." "balancing" "brains" "can" "care"   
## [6] "choose." "come" "day!" "direction" "dr."   
## [11] "except" "famous" "feet" "fun" "games"   
## [16] "get" "go!" "go." "great" "indeed!"   
## [21] "just" "kid," "know." "life's" "may"   
## [26] "mountain" "move" "much" "never" "oh,"   
## [31] "one" "percent" "places" "places!" "remember"   
## [36] "right" "seuss" "seuss," "shoes." "sometimes"  
## [41] "start" "steer" "step" "succeed?" "sure"   
## [46] "things" "today" "troubles" "way!" "whole"   
## [51] "will" "will," "win" "yes!"

# which words are associated with "brains"?  
findAssocs(seuss.tdm, 'brains', 0.30)

## $brains  
## head. --good feet, full head head, shoes smart street.   
## 0.70 0.49 0.49 0.49 0.49 0.49 0.49 0.49 0.49

# Word Cloud  
m <- as.matrix(seuss.tdm)  
# calculate the frequency of words  
v <- sort(rowSums(m), decreasing=TRUE)  
myNames <- names(v)  
d <- data.frame(word=myNames, freq=v)  
wordcloud(d$word, d$freq, min.freq=3)



# Text Mining in R (RTextTools package)

This example comes directly from the RTextTools package. To find the [original link](http://www.rtexttools.com/blog/classifying-cancer-as-benign-or-malignant-using-rtexttools)

[RTextTools](http://www.rtexttools.com/) created a text classification to identify breast cancer masses as benign or malignant. Using the Wisconsin Diagnostic Breast Cancer Dataset from UC Irvine, [RTextTools](http://www.rtexttools.com/) wrote a script that trains eight classifiers on characteristics such as clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, and mitoses. When run on the data, the classifiers were able to achieve up to 96% recall accuracy on a randomly sampled training set of 200 patients and test set of 400 patients.

*Citation: Classification via Supervised Learning. R package version 1.3.9.* [[*http://CRAN.R-project.org/package=RTextTools*](http://CRAN.R-project.org/package=RTextTools)](http://CRAN.R-project.org/package=RTextTools)

# Reference: http://www.rtexttools.com/blog/classifying-cancer-as-benign-or-malignant-using-rtexttools  
  
# Classifying Breast Cancer as Benign or Malignant Using RTextTools - RTextTools: a machine learning library for text classification http://www.rtexttools.com/1/post/2012/02/classifying-cancer-as-benign-or-malignant-using-rtexttools.html  
# http://www.rtexttools.com/how-to-cite.html  
# GET THE BREAST CANCER data FROM http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/wdbc.names  
data <- read.csv("http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/breast-cancer-wisconsin.data",header=FALSE)  
head(data)

## V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11  
## 1 1000025 5 1 1 1 2 1 3 1 1 2  
## 2 1002945 5 4 4 5 7 10 3 2 1 2  
## 3 1015425 3 1 1 1 2 2 3 1 1 2  
## 4 1016277 6 8 8 1 3 4 3 7 1 2  
## 5 1017023 4 1 1 3 2 1 3 1 1 2  
## 6 1017122 8 10 10 8 7 10 9 7 1 4

data <- data[-1]  
head(data)

## V2 V3 V4 V5 V6 V7 V8 V9 V10 V11  
## 1 5 1 1 1 2 1 3 1 1 2  
## 2 5 4 4 5 7 10 3 2 1 2  
## 3 3 1 1 1 2 2 3 1 1 2  
## 4 6 8 8 1 3 4 3 7 1 2  
## 5 4 1 1 3 2 1 3 1 1 2  
## 6 8 10 10 8 7 10 9 7 1 4

# ADD TEXTUAL DESCRIPTORS FOR EACH MASS CHARACTERISTIC FOR THE DOCUMENT-TERM MATRIX  
thick <- as.vector(apply(as.matrix(data[1], mode="character"),1,paste,"clump",sep="",collapse=""))  
size <- as.vector(apply(as.matrix(data[2], mode="character"),1,paste,"size",sep="",collapse=""))  
shape <- as.vector(apply(as.matrix(data[3], mode="character"),1,paste,"shape",sep="",collapse=""))  
adhesion <- as.vector(apply(as.matrix(data[4], mode="character"),1,paste,"adhesion",sep="",collapse=""))  
single <- as.vector(apply(as.matrix(data[5], mode="character"),1,paste,"single",sep="",collapse=""))  
nuclei <- as.vector(apply(as.matrix(data[6], mode="character"),1,paste,"nuclei",sep="",collapse=""))  
chromatin <- as.vector(apply(as.matrix(data[7], mode="character"),1,paste,"chromatin",sep="",collapse=""))  
nucleoli <- as.vector(apply(as.matrix(data[8], mode="character"),1,paste,"nucleoli",sep="",collapse=""))  
mitoses <- as.vector(apply(as.matrix(data[9], mode="character"),1,paste,"mitoses",sep="",collapse=""))  
training\_data <- cbind(data[10],thick,size,shape,adhesion,single,nuclei,chromatin,nucleoli,mitoses)  
head(training\_data)

## V11 thick size shape adhesion single nuclei chromatin  
## 1 2 5clump 1size 1shape 1adhesion 2single 1nuclei 3chromatin  
## 2 2 5clump 4size 4shape 5adhesion 7single 10nuclei 3chromatin  
## 3 2 3clump 1size 1shape 1adhesion 2single 2nuclei 3chromatin  
## 4 2 6clump 8size 8shape 1adhesion 3single 4nuclei 3chromatin  
## 5 2 4clump 1size 1shape 3adhesion 2single 1nuclei 3chromatin  
## 6 4 8clump 10size 10shape 8adhesion 7single 10nuclei 9chromatin  
## nucleoli mitoses  
## 1 1nucleoli 1mitoses  
## 2 2nucleoli 1mitoses  
## 3 1nucleoli 1mitoses  
## 4 7nucleoli 1mitoses  
## 5 1nucleoli 1mitoses  
## 6 7nucleoli 1mitoses

head(data)

## V2 V3 V4 V5 V6 V7 V8 V9 V10 V11  
## 1 5 1 1 1 2 1 3 1 1 2  
## 2 5 4 4 5 7 10 3 2 1 2  
## 3 3 1 1 1 2 2 3 1 1 2  
## 4 6 8 8 1 3 4 3 7 1 2  
## 5 4 1 1 3 2 1 3 1 1 2  
## 6 8 10 10 8 7 10 9 7 1 4

# [OPTIONAL] SUBSET YOUR data TO GET A RANDOM SAMPLE  
training\_data <- training\_data[sample(1:699,size=600,replace=FALSE),]  
training\_codes <- training\_data[1]  
training\_data <- training\_data[-1]  
# CREATE A TERM-DOCUMENT MATRIX THAT REPRESENTS WORD FREQUENCIES IN EACH DOCUMENT  
# WE WILL TRAIN ON THE Title and Subject COLUMNS  
matrix <- create\_matrix(training\_data, language="english", removeNumbers=FALSE, stemWords=FALSE, removePunctuation=FALSE, weighting=weightTf)  
# CREATE A container THAT IS SPLIT INTO A TRAINING SET AND A TESTING SET  
# WE WILL BE USING t(training\_codes) AS THE CODE COLUMN. WE DEFINE A 200   
# ARTICLE TRAINING SET AND A 400 ARTICLE TESTING SET.  
container <- create\_container(matrix,t(training\_codes),trainSize=1:200, testSize=201:600,virgin=FALSE)  
# THERE ARE TWO METHODS OF TRAINING AND CLASSIFYING data.  
# ONE WAY IS TO DO THEM AS A BATCH (SEVERAL ALGORITHMS AT ONCE)  
models <- train\_models(container, algorithms=c("MAXENT","SVM","GLMNET","SLDA","TREE","BAGGING","BOOSTING","RF"))  
results <- classify\_models(container, models)  
# VIEW THE RESULTS BY CREATING ANALYTICS  
analytics <- create\_analytics(container, results)  
# RESULTS WILL BE REPORTED BACK IN THE analytics VARIABLE.  
# analytics@algorithm\_summary: SUMMARY OF PRECISION, RECALL, F-SCORES, AND ACCURACY SORTED BY TOPIC CODE FOR EACH ALGORITHM  
# analytics@label\_summary: SUMMARY OF LABEL (e.g. TOPIC) ACCURACY  
# analytics@document\_summary: RAW SUMMARY OF ALL data AND SCORING  
# analytics@ensemble\_summary: SUMMARY OF ENSEMBLE PRECISION/COVERAGE. USES THE n VARIABLE PASSED INTO create\_analytics()  
analytics

## An object of class "analytics"  
## Slot "label\_summary":  
## NUM\_MANUALLY\_CODED NUM\_CONSENSUS\_CODED NUM\_PROBABILITY\_CODED  
## 2 262 260 258  
## 4 138 140 142  
## PCT\_CONSENSUS\_CODED PCT\_PROBABILITY\_CODED PCT\_CORRECTLY\_CODED\_CONSENSUS  
## 2 99.23664 98.47328 95.80153  
## 4 101.44928 102.89855 93.47826  
## PCT\_CORRECTLY\_CODED\_PROBABILITY  
## 2 95.41985  
## 4 94.20290  
##   
## Slot "document\_summary":  
## MAXENTROPY\_LABEL MAXENTROPY\_PROB SVM\_LABEL SVM\_PROB GLMNET\_LABEL  
## 1 2 0.9997896 2 0.9737916 2  
## 2 2 0.9999984 2 0.9944157 2  
## 3 2 0.9999598 2 0.9700087 2  
## 4 2 0.9999996 2 0.9972273 2  
## 5 2 0.9999963 2 0.9914334 2  
## 6 2 0.9999994 2 0.9956296 2  
## 7 2 0.9999994 2 0.9937439 2  
## 8 4 0.5399080 4 0.6719558 4  
## 9 4 0.9999353 4 0.9902735 4  
## 10 4 0.9999753 4 0.9888515 4  
## 11 2 0.9999705 2 0.9675052 2  
## 12 4 0.9999966 4 0.9919136 4  
## 13 4 1.0000000 4 0.9999968 4  
## 14 4 0.9999985 4 0.9882190 4  
## 15 2 0.9999970 2 0.9919114 2  
## 16 2 0.9999972 2 0.9877060 2  
## 17 2 0.9999924 2 0.9855771 2  
## 18 4 0.9999996 4 0.9963133 4  
## 19 2 1.0000000 2 0.9946599 2  
## 20 4 0.9980831 4 0.9809666 4  
## 21 2 0.9999998 2 0.9971608 2  
## 22 4 0.9999465 4 0.9680775 4  
## 23 2 0.9999968 2 0.9934604 2  
## 24 2 0.9999968 2 0.9934604 2  
## 25 2 0.9998488 2 0.9401435 2  
## 26 2 0.6619296 4 0.5266271 2  
## 27 2 0.9999988 2 0.9948131 2  
## 28 2 0.9999980 2 0.9914596 2  
## 29 2 0.9993221 2 0.8793026 2  
## 30 2 0.9999450 2 0.9669219 2  
## 31 4 0.9178325 4 0.8663224 4  
## 32 4 0.9999994 4 0.9952153 4  
## 33 4 1.0000000 4 0.9999994 4  
## 34 2 0.9999575 2 0.9601978 2  
## 35 2 0.9999998 2 0.9971608 2  
## 36 4 0.9999998 4 0.9898368 4  
## 37 2 0.9999996 2 0.9972273 2  
## 38 4 0.9999999 4 0.9935908 4  
## 39 4 0.9999997 4 0.9958483 4  
## 40 4 0.9999966 4 0.9937208 4  
## 41 4 1.0000000 4 1.0000000 4  
## 42 4 1.0000000 4 0.9999997 4  
## 43 4 0.9999977 4 0.9948245 4  
## 44 2 0.9999995 2 0.9949554 2  
## 45 2 0.8702899 2 0.6315226 2  
## 46 4 0.9989115 4 0.9301341 4  
## 47 4 1.0000000 4 0.9999943 4  
## 48 2 0.9995978 2 0.9304279 2  
## 49 2 0.9999732 2 0.9803041 2  
## 50 4 0.9999892 4 0.9826776 4  
## 51 4 0.9997029 4 0.9361982 4  
## 52 2 0.9999598 2 0.9700087 2  
## 53 2 0.9999996 2 0.9961354 2  
## 54 2 0.9999924 2 0.9855771 2  
## 55 2 0.9999899 2 0.9878254 2  
## 56 2 0.9999993 2 0.9941116 2  
## 57 2 0.9982654 2 0.9275215 2  
## 58 2 0.9999994 2 0.9937439 2  
## 59 2 0.9999996 2 0.9972273 2  
## 60 2 0.9999352 2 0.9733764 2  
## 61 2 0.9999988 2 0.9930373 2  
## 62 2 0.9916784 2 0.8102376 2  
## 63 2 0.9999991 2 0.9967035 2  
## 64 2 0.9999912 2 0.9882081 2  
## 65 2 0.9999963 2 0.9914334 2  
## 66 2 0.9999999 2 0.9968719 2  
## 67 2 0.9999912 2 0.9882081 2  
## 68 2 0.9999622 2 0.9726902 2  
## 69 2 0.9999486 2 0.9746320 2  
## 70 4 0.9999986 4 0.9943501 4  
## 71 4 0.9994438 4 0.9543832 4  
## 72 2 0.9999997 2 0.9970227 2  
## 73 4 0.9995908 4 0.9766465 4  
## 74 2 0.9947131 2 0.8529127 2  
## 75 4 0.9970381 4 0.9684242 4  
## 76 4 1.0000000 4 0.9999942 4  
## 77 4 0.9995107 4 0.9923158 4  
## 78 2 0.9999878 2 0.9803431 2  
## 79 2 0.9994619 2 0.8277231 2  
## 80 2 0.9999976 2 0.9661949 2  
## 81 2 0.9999996 2 0.9969743 2  
## 82 4 0.9999996 4 0.9836614 4  
## 83 2 0.9999912 2 0.9882081 2  
## 84 2 0.9999998 2 0.9971608 2  
## 85 2 0.9999993 2 0.9970405 2  
## 86 4 0.9997711 4 0.9607393 4  
## 87 2 0.9999994 2 0.9948434 2  
## 88 4 0.9581145 4 0.7780398 4  
## 89 2 0.9954589 2 0.9095049 2  
## 90 2 0.9999602 2 0.9720491 2  
## 91 2 0.8047728 2 0.7226327 2  
## 92 2 0.9999968 2 0.9934604 2  
## 93 2 0.9999996 2 0.9972273 2  
## 94 4 0.9999985 4 0.9934327 4  
## 95 2 0.9999980 2 0.9914596 2  
## 96 4 0.9999886 4 0.9901645 4  
## 97 2 0.9999998 2 0.9971608 2  
## 98 4 0.9982924 4 0.9457608 4  
## 99 2 0.9998576 2 0.9586110 2  
## 100 2 1.0000000 2 0.9993400 2  
## 101 2 0.9995978 2 0.9304279 2  
## 102 4 0.9999964 4 0.9936011 4  
## 103 4 0.9999557 4 0.9840837 4  
## 104 4 0.9999921 4 0.9884739 4  
## 105 4 0.9999917 4 0.9924840 4  
## 106 2 0.9999847 2 0.9810159 2  
## 107 2 0.9995978 2 0.9304279 2  
## 108 2 0.9998488 2 0.9401435 2  
## 109 2 0.9999999 2 0.9986368 2  
## 110 4 0.9999998 4 0.9933608 4  
## 111 2 0.9999985 2 0.9908711 2  
## 112 2 0.9999980 2 0.9914596 2  
## 113 4 0.9999797 4 0.9839727 4  
## 114 2 0.9999994 2 0.9948434 2  
## 115 4 0.9999994 4 0.9931201 4  
## 116 4 0.9913091 4 0.9284185 4  
## 117 4 1.0000000 4 0.9969290 4  
## 118 4 0.9997839 4 0.9895352 4  
## 119 2 0.9999998 2 0.9967144 2  
## 120 2 0.9999952 2 0.9931852 2  
## 121 2 0.9999848 2 0.9864804 2  
## 122 2 0.9999912 2 0.9882081 2  
## 123 2 0.9997517 2 0.9302881 2  
## 124 2 0.9999999 2 0.9966732 2  
## 125 2 0.9995403 2 0.8962485 2  
## 126 4 0.9999908 4 0.9734910 4  
## 127 2 0.9999988 2 0.9948131 2  
## 128 4 1.0000000 4 0.9999998 4  
## 129 4 0.9999647 4 0.9444318 4  
## 130 2 0.9999996 2 0.9972273 2  
## 131 4 0.9999967 4 0.9944656 4  
## 132 2 0.9999998 2 0.9971608 2  
## 133 4 0.9998739 4 0.9660707 4  
## 134 2 1.0000000 2 0.9962133 2  
## 135 2 0.9999994 2 0.9937439 2  
## 136 2 0.9997714 2 0.9334486 2  
## 137 2 0.9992269 2 0.9095265 2  
## 138 4 0.9997907 4 0.9743641 4  
## 139 2 0.9999998 2 0.9971608 2  
## 140 4 0.9999985 4 0.9965169 4  
## 141 2 0.8917743 2 0.8304403 2  
## 142 2 0.9999936 2 0.9660702 2  
## 143 2 0.9999924 2 0.9855771 2  
## 144 2 0.9999996 2 0.9972273 2  
## 145 2 0.9999994 2 0.9948434 2  
## 146 4 0.9948378 4 0.8248679 2  
## 147 2 0.9999988 2 0.9948131 2  
## 148 4 0.9999974 4 0.9966806 4  
## 149 4 0.9999929 4 0.9844120 4  
## 150 4 1.0000000 4 0.9999990 4  
## 151 2 1.0000000 2 0.9985558 2  
## 152 2 1.0000000 2 0.9964243 2  
## 153 2 0.9999876 2 0.9833515 2  
## 154 4 1.0000000 4 0.9974444 4  
## 155 2 0.9999799 2 0.9732664 2  
## 156 2 0.8646740 2 0.8478779 2  
## 157 2 0.9999998 2 0.9971608 2  
## 158 4 0.9999984 4 0.9970666 4  
## 159 2 0.9999999 2 0.9950556 2  
## 160 4 1.0000000 4 1.0000000 4  
## 161 4 0.9999910 4 0.9862631 4  
## 162 2 0.9883690 2 0.7715877 2  
## 163 2 0.9999998 2 0.9971608 2  
## 164 2 0.9999769 2 0.9160336 2  
## 165 4 0.9999930 4 0.9954575 4  
## 166 2 0.9999999 2 0.9985675 2  
## 167 2 0.9999622 2 0.9726902 2  
## 168 4 0.9992913 4 0.9507533 4  
## 169 2 0.5888366 2 0.7528653 2  
## 170 2 0.9999994 2 0.9948434 2  
## 171 2 0.9999968 2 0.9934604 2  
## 172 4 0.9999829 4 0.9880122 4  
## 173 2 0.9999998 2 0.9971608 2  
## 174 2 0.9665045 2 0.5908551 2  
## 175 2 0.9996943 2 0.9585156 2  
## 176 2 0.9999950 2 0.9885200 2  
## 177 4 1.0000000 4 0.9999995 4  
## 178 2 0.9994266 2 0.9540936 2  
## 179 2 0.9999996 2 0.9972273 2  
## 180 2 0.9999999 2 0.9984855 2  
## 181 4 0.9999999 4 0.9999894 4  
## 182 4 1.0000000 4 0.9966843 4  
## 183 4 0.9819447 4 0.7094687 4  
## 184 4 1.0000000 4 0.9999994 4  
## 185 2 0.9999978 2 0.9915951 2  
## 186 4 1.0000000 4 0.9959639 4  
## 187 4 0.9998499 4 0.9278612 4  
## 188 2 0.9926749 2 0.8432224 2  
## 189 2 0.9999996 2 0.9969743 2  
## 190 4 0.9999991 4 0.9894750 4  
## 191 4 0.9999991 4 0.9940003 4  
## 192 2 0.8096780 4 0.6956986 4  
## 193 2 0.9999968 2 0.9934604 2  
## 194 2 0.9999988 2 0.9948131 2  
## 195 4 0.7757843 4 0.5791690 4  
## 196 2 0.9999874 2 0.9873546 2  
## 197 2 0.9999997 2 0.9969167 2  
## 198 4 1.0000000 4 0.9967980 4  
## 199 2 0.9999994 2 0.9961280 2  
## 200 4 0.9999078 4 0.9779876 4  
## 201 4 0.9999701 4 0.9178155 4  
## 202 2 0.9999968 2 0.9934604 2  
## 203 2 0.9804316 2 0.5178053 4  
## 204 4 0.9985392 4 0.9724056 4  
## 205 4 0.9999899 4 0.9599116 4  
## 206 2 0.9999922 2 0.9836112 2  
## 207 2 0.9995840 2 0.9401313 2  
## 208 4 0.7938303 4 0.7104616 4  
## 209 2 0.9999722 2 0.9772132 2  
## 210 2 0.9999963 2 0.9914334 2  
## 211 4 0.9999569 4 0.9907252 4  
## 212 4 0.9999915 4 0.9864139 4  
## 213 4 0.9999842 4 0.9586309 4  
## 214 2 0.9995978 2 0.9304279 2  
## 215 2 0.9999352 2 0.9733764 2  
## 216 2 0.9999998 2 0.9971608 2  
## 217 2 0.9999994 2 0.9937439 2  
## 218 4 0.9997445 4 0.9016469 4  
## 219 2 0.9998823 2 0.9662408 2  
## 220 2 0.9999994 2 0.9961280 2  
## 221 2 0.6634950 4 0.6032912 2  
## 222 2 0.9999988 2 0.9948131 2  
## 223 2 1.0000000 2 0.9986088 2  
## 224 2 0.9999980 2 0.9914596 2  
## 225 2 0.9999997 2 0.9874782 2  
## 226 2 1.0000000 2 0.9971812 2  
## 227 4 0.9999976 4 0.9926391 4  
## 228 2 0.9998474 2 0.9424238 2  
## 229 2 0.9999431 2 0.9743472 2  
## 230 4 0.9999352 4 0.9816116 4  
## 231 4 0.9999477 4 0.9844647 4  
## 232 4 0.9999888 4 0.9795805 4  
## 233 2 1.0000000 2 0.9985037 2  
## 234 4 0.9999737 4 0.9895857 4  
## 235 2 0.9999968 2 0.9934604 2  
## 236 4 0.9999115 4 0.9572661 4  
## 237 2 0.9999299 2 0.9327201 2  
## 238 4 1.0000000 4 0.9999999 4  
## 239 2 0.9999996 2 0.9972273 2  
## 240 2 0.9996608 2 0.8882903 2  
## 241 2 0.9999029 2 0.9710006 2  
## 242 2 0.9674614 2 0.6471001 2  
## 243 4 1.0000000 4 0.9999992 4  
## 244 2 0.9999988 2 0.9948131 2  
## 245 4 0.9999999 4 0.9999938 4  
## 246 2 0.9999653 2 0.9639247 2  
## 247 4 1.0000000 4 0.9999998 4  
## 248 4 1.0000000 4 0.9999998 4  
## 249 4 0.9771148 4 0.8075743 2  
## 250 2 0.9999999 2 0.9968887 2  
## 251 4 0.5160729 4 0.5143533 2  
## 252 4 1.0000000 4 0.9965383 4  
## 253 4 1.0000000 4 0.9966266 4  
## 254 2 0.9999968 2 0.9934604 2  
## 255 2 0.9999732 2 0.9803041 2  
## 256 4 1.0000000 4 0.9999815 4  
## 257 2 0.9999995 2 0.9912431 2  
## 258 2 0.9999705 2 0.9675052 2  
## 259 4 0.9999915 4 0.9864263 4  
## 260 4 1.0000000 4 0.9949896 4  
## 261 4 1.0000000 4 0.9999998 4  
## 262 2 0.9999912 2 0.9882081 2  
## 263 4 0.9999998 4 0.9999915 4  
## 264 2 0.9999998 2 0.9915073 2  
## 265 4 1.0000000 4 0.9999993 4  
## 266 2 0.9999602 2 0.9720491 2  
## 267 4 1.0000000 4 0.9999857 4  
## 268 2 0.9999990 2 0.9937818 2  
## 269 2 0.9999988 2 0.9948131 2  
## 270 2 0.9999705 2 0.9675052 2  
## 271 2 0.9999849 2 0.9796387 2  
## 272 2 0.9999955 2 0.9906898 2  
## 273 4 1.0000000 4 0.9999981 4  
## 274 2 0.9999963 2 0.9914334 2  
## 275 2 0.9965066 2 0.8518310 2  
## 276 2 0.9999994 2 0.9948434 2  
## 277 2 0.9999988 2 0.9948131 2  
## 278 2 0.9999998 2 0.9971608 2  
## 279 2 0.9999994 2 0.9961280 2  
## 280 4 0.9999993 4 0.9955887 4  
## 281 2 0.9926749 2 0.8432224 2  
## 282 2 0.9998875 2 0.9587133 2  
## 283 2 0.9999409 2 0.9688903 2  
## 284 2 0.9961917 2 0.7685581 2  
## 285 4 0.9944354 4 0.9853465 4  
## 286 4 0.9983648 4 0.9268440 4  
## 287 2 0.9994590 2 0.9259786 2  
## 288 2 0.9999993 2 0.9970405 2  
## 289 4 1.0000000 4 0.9999996 4  
## 290 4 0.9999808 4 0.9868304 4  
## 291 2 0.9999952 2 0.9910806 2  
## 292 4 0.9963777 4 0.8311284 4  
## 293 2 0.9999998 2 0.9971608 2  
## 294 2 1.0000000 2 0.9967275 2  
## 295 4 1.0000000 4 0.9999960 4  
## 296 2 0.9999998 2 0.9975425 2  
## 297 4 0.8432012 4 0.5712581 2  
## 298 2 0.9998488 2 0.9401435 2  
## 299 2 1.0000000 2 0.9987144 2  
## 300 4 1.0000000 4 0.9999993 4  
## 301 2 0.9999924 2 0.9855771 2  
## 302 4 0.9999995 4 0.9968212 4  
## 303 2 0.9999918 2 0.9865319 2  
## 304 2 1.0000000 2 0.9985098 2  
## 305 2 0.9999663 2 0.9650437 2  
## 306 2 0.9999952 2 0.9931852 2  
## 307 2 0.9999732 2 0.9803041 2  
## 308 4 0.9733362 4 0.9405373 4  
## 309 2 0.9999029 2 0.9710006 2  
## 310 2 0.9999998 2 0.9971608 2  
## 311 2 0.9995978 2 0.9304279 2  
## 312 2 0.9999990 2 0.9937818 2  
## 313 4 1.0000000 4 0.9999993 4  
## 314 2 0.9999968 2 0.9934604 2  
## 315 2 0.9995840 2 0.9401313 2  
## 316 4 1.0000000 4 0.9999987 4  
## 317 2 0.9999480 2 0.9571446 2  
## 318 2 0.9999999 2 0.9905509 2  
## 319 2 0.9999999 2 0.9962683 2  
## 320 2 0.9990454 2 0.8580603 2  
## 321 2 0.9999955 2 0.9925683 2  
## 322 2 0.9995978 2 0.9304279 2  
## 323 2 0.9993972 2 0.9284658 2  
## 324 4 0.9999476 4 0.9744910 4  
## 325 2 1.0000000 2 0.9993400 2  
## 326 2 0.9999998 2 0.9971608 2  
## 327 4 0.9999999 4 0.9947185 4  
## 328 4 0.9999143 4 0.9915733 4  
## 329 2 0.9999996 2 0.9972273 2  
## 330 2 0.9999998 2 0.9971608 2  
## 331 2 0.9999988 2 0.9948131 2  
## 332 4 0.9999999 4 0.9903793 4  
## 333 2 0.9999963 2 0.9914334 2  
## 334 2 0.7498879 2 0.7278363 2  
## 335 4 1.0000000 4 0.9999955 4  
## 336 2 0.9997775 2 0.9216129 2  
## 337 4 0.9999449 4 0.9721311 4  
## 338 4 0.8850564 4 0.8573027 4  
## 339 2 0.9971880 2 0.9356755 2  
## 340 2 1.0000000 2 0.9985098 2  
## 341 4 1.0000000 4 0.9999905 4  
## 342 2 0.9999999 2 0.9968887 2  
## 343 4 0.9892643 4 0.8553069 2  
## 344 2 0.8877827 4 0.5524751 2  
## 345 2 0.9999980 2 0.9914596 2  
## 346 2 0.9999994 2 0.9948434 2  
## 347 2 1.0000000 2 0.9980957 2  
## 348 4 0.9999297 4 0.9954363 4  
## 349 2 0.9999952 2 0.9931852 2  
## 350 4 1.0000000 4 0.9951040 4  
## 351 4 0.9999998 4 0.9969799 4  
## 352 2 0.9999236 2 0.9586501 2  
## 353 2 0.9999899 2 0.9772971 2  
## 354 2 0.9961344 2 0.8138635 2  
## 355 2 0.9999963 2 0.9914334 2  
## 356 2 0.9999999 2 0.9968887 2  
## 357 2 0.7669177 4 0.5908602 2  
## 358 2 0.9999963 2 0.9914334 2  
## 359 2 0.9998588 2 0.9773248 2  
## 360 2 0.9999997 2 0.9953470 2  
## 361 4 0.9999733 4 0.9633401 4  
## 362 2 0.9999998 2 0.9971608 2  
## 363 2 0.9999957 2 0.9861191 2  
## 364 2 0.9999993 2 0.9970405 2  
## 365 2 0.9999705 2 0.9675052 2  
## 366 2 0.9999988 2 0.9948131 2  
## 367 2 0.9999450 2 0.9669219 2  
## 368 2 0.9999968 2 0.9934604 2  
## 369 2 0.9999949 2 0.9935684 2  
## 370 2 0.9999996 2 0.9972273 2  
## 371 2 0.8882970 2 0.7554219 2  
## 372 2 0.9999994 2 0.9937439 2  
## 373 2 0.9999990 2 0.9937818 2  
## 374 4 0.9999856 4 0.9826795 4  
## 375 2 0.9999732 2 0.9803041 2  
## 376 4 0.9961143 4 0.7588837 4  
## 377 2 0.9999991 2 0.9956934 2  
## 378 2 0.9999996 2 0.9968511 2  
## 379 2 0.9996681 2 0.9511352 2  
## 380 2 0.9999732 2 0.9803041 2  
## 381 2 0.9999912 2 0.9882081 2  
## 382 2 0.9999924 2 0.9855771 2  
## 383 2 0.9999783 2 0.9834545 2  
## 384 2 0.9999634 2 0.9823144 2  
## 385 2 0.5083342 4 0.6029048 2  
## 386 4 0.9990920 4 0.9669068 4  
## 387 4 1.0000000 4 0.9999967 4  
## 388 2 0.9994221 2 0.9527567 2  
## 389 2 0.9999998 2 0.9969791 2  
## 390 2 0.9999848 2 0.9848430 2  
## 391 2 0.9999991 2 0.9956934 2  
## 392 2 0.9999029 2 0.9710006 2  
## 393 4 0.8419980 4 0.7395693 4  
## 394 4 0.9999977 4 0.9702123 4  
## 395 4 1.0000000 4 0.9999950 4  
## 396 2 0.9999994 2 0.9937439 2  
## 397 2 0.9999988 2 0.9948131 2  
## 398 4 1.0000000 4 0.9999979 4  
## 399 4 1.0000000 4 0.9971304 4  
## 400 4 0.9999801 4 0.9950919 4  
## GLMNET\_PROB SLDA\_LABEL SLDA\_PROB TREE\_LABEL TREE\_PROB BAGGING\_LABEL  
## 1 0.9869330 2 1.0000000 2 1.0000000 2  
## 2 0.9972601 2 1.0000000 2 1.0000000 2  
## 3 0.9757082 2 1.0000000 2 1.0000000 2  
## 4 0.9965793 2 1.0000000 2 1.0000000 2  
## 5 0.9965793 2 1.0000000 2 1.0000000 2  
## 6 0.9868132 2 1.0000000 2 1.0000000 2  
## 7 0.9966107 2 1.0000000 2 1.0000000 2  
## 8 0.6799620 4 0.9970750 4 1.0000000 4  
## 9 0.9707861 4 0.9998728 4 1.0000000 4  
## 10 0.9937241 4 1.0000000 4 1.0000000 4  
## 11 0.9759265 2 1.0000000 2 1.0000000 2  
## 12 0.9945143 4 1.0000000 4 1.0000000 4  
## 13 0.9975950 4 1.0000000 4 1.0000000 4  
## 14 0.9568785 4 1.0000000 4 1.0000000 4  
## 15 0.9897269 2 0.9999998 2 1.0000000 2  
## 16 0.9886719 2 1.0000000 2 1.0000000 2  
## 17 0.9965793 2 1.0000000 2 1.0000000 2  
## 18 0.9962880 4 0.9999998 4 1.0000000 4  
## 19 0.9966107 2 1.0000000 2 1.0000000 2  
## 20 0.9705848 4 0.9999943 4 1.0000000 4  
## 21 0.9966107 2 1.0000000 2 1.0000000 2  
## 22 0.9919868 4 0.9999464 4 1.0000000 4  
## 23 0.9965793 2 1.0000000 2 1.0000000 2  
## 24 0.9965793 2 1.0000000 2 1.0000000 2  
## 25 0.9757082 2 1.0000000 2 1.0000000 2  
## 26 0.6616557 2 0.9348860 2 0.6000000 2  
## 27 0.9965793 2 1.0000000 2 1.0000000 2  
## 28 0.9966107 2 1.0000000 2 1.0000000 2  
## 29 0.9232695 2 0.9999972 2 1.0000000 2  
## 30 0.9757082 2 1.0000000 2 1.0000000 2  
## 31 0.8228600 4 0.9999945 2 0.6000000 4  
## 32 0.8905488 4 1.0000000 4 1.0000000 2  
## 33 0.9877224 4 1.0000000 4 1.0000000 4  
## 34 0.7675361 2 0.9999761 2 0.6666667 2  
## 35 0.9966107 2 1.0000000 2 1.0000000 2  
## 36 0.9902928 4 1.0000000 4 1.0000000 4  
## 37 0.9965793 2 1.0000000 2 1.0000000 2  
## 38 0.9567242 4 1.0000000 4 1.0000000 4  
## 39 0.9944764 4 1.0000000 4 1.0000000 4  
## 40 0.9666361 4 0.9999994 4 1.0000000 4  
## 41 0.9994470 4 1.0000000 4 1.0000000 4  
## 42 0.9964924 4 1.0000000 4 1.0000000 4  
## 43 0.9821854 4 0.9999999 4 1.0000000 4  
## 44 0.9966107 2 1.0000000 2 1.0000000 2  
## 45 0.7169406 2 0.9999847 2 0.6666667 2  
## 46 0.9193682 4 0.9999631 2 0.6000000 4  
## 47 0.9957271 4 1.0000000 4 1.0000000 4  
## 48 0.9757082 2 1.0000000 2 1.0000000 2  
## 49 0.9965793 2 1.0000000 2 1.0000000 2  
## 50 0.9477356 4 0.9999999 4 0.7142857 4  
## 51 0.8835993 4 1.0000000 4 1.0000000 4  
## 52 0.9757082 2 1.0000000 2 1.0000000 2  
## 53 0.9885678 2 0.9999995 2 1.0000000 2  
## 54 0.9965793 2 1.0000000 2 1.0000000 2  
## 55 0.9885678 2 0.9999999 2 1.0000000 2  
## 56 0.9965793 2 1.0000000 2 1.0000000 2  
## 57 0.9757082 2 0.9999997 2 1.0000000 2  
## 58 0.9966107 2 1.0000000 2 1.0000000 2  
## 59 0.9965793 2 1.0000000 2 1.0000000 2  
## 60 0.9862110 2 0.9999998 2 1.0000000 2  
## 61 0.9487627 2 0.9999990 2 1.0000000 2  
## 62 0.9790133 2 0.9999968 2 1.0000000 2  
## 63 0.9974583 2 1.0000000 2 1.0000000 2  
## 64 0.9965793 2 1.0000000 2 1.0000000 2  
## 65 0.9965793 2 1.0000000 2 1.0000000 2  
## 66 0.9968266 2 1.0000000 2 1.0000000 2  
## 67 0.9965793 2 1.0000000 2 1.0000000 2  
## 68 0.9885678 2 0.9999999 2 1.0000000 2  
## 69 0.9515163 2 0.9999998 2 0.6666667 2  
## 70 0.9794385 4 1.0000000 4 1.0000000 4  
## 71 0.9711781 4 0.9999996 4 0.7142857 4  
## 72 0.9863181 2 1.0000000 2 1.0000000 2  
## 73 0.8926867 4 0.9998506 4 1.0000000 4  
## 74 0.9885678 2 0.9993129 2 1.0000000 2  
## 75 0.9921559 4 0.9999978 2 0.6000000 4  
## 76 0.9983909 4 1.0000000 4 1.0000000 4  
## 77 0.9825499 4 1.0000000 4 1.0000000 4  
## 78 0.9868132 2 1.0000000 2 0.6666667 2  
## 79 0.7594470 4 0.5880868 2 0.6000000 4  
## 80 0.7852454 2 0.9999801 2 1.0000000 2  
## 81 0.9968557 2 1.0000000 2 1.0000000 2  
## 82 0.9530796 4 1.0000000 4 0.7142857 4  
## 83 0.9965793 2 1.0000000 2 1.0000000 2  
## 84 0.9966107 2 1.0000000 2 1.0000000 2  
## 85 0.9968266 2 1.0000000 2 1.0000000 2  
## 86 0.7783299 4 0.9942075 4 0.7142857 4  
## 87 0.9966107 2 1.0000000 2 1.0000000 2  
## 88 0.8990187 2 0.9653489 4 1.0000000 2  
## 89 0.9758258 2 0.9999993 2 1.0000000 2  
## 90 0.9868132 2 0.9999984 2 0.6666667 2  
## 91 0.5903965 4 0.8142822 2 0.6000000 4  
## 92 0.9965793 2 1.0000000 2 1.0000000 2  
## 93 0.9965793 2 1.0000000 2 1.0000000 2  
## 94 0.9701018 4 0.9999794 4 1.0000000 4  
## 95 0.9966107 2 1.0000000 2 1.0000000 2  
## 96 0.6191281 4 0.9999667 4 1.0000000 4  
## 97 0.9966107 2 1.0000000 2 1.0000000 2  
## 98 0.8071847 4 0.9999998 4 1.0000000 4  
## 99 0.9132987 2 0.9999996 2 1.0000000 2  
## 100 0.9972852 2 1.0000000 2 1.0000000 2  
## 101 0.9757082 2 1.0000000 2 1.0000000 2  
## 102 0.9255941 4 1.0000000 2 0.6000000 4  
## 103 0.9766617 4 1.0000000 4 1.0000000 4  
## 104 0.9918766 4 1.0000000 2 0.6000000 4  
## 105 0.9568785 4 0.9999999 4 1.0000000 4  
## 106 0.9862110 2 1.0000000 2 1.0000000 2  
## 107 0.9757082 2 1.0000000 2 1.0000000 2  
## 108 0.9757082 2 1.0000000 2 1.0000000 2  
## 109 0.9972601 2 1.0000000 2 1.0000000 2  
## 110 0.9871244 4 1.0000000 4 1.0000000 4  
## 111 0.9885678 2 1.0000000 2 1.0000000 2  
## 112 0.9966107 2 1.0000000 2 1.0000000 2  
## 113 0.9571043 4 0.9999996 4 1.0000000 4  
## 114 0.9966107 2 1.0000000 2 1.0000000 2  
## 115 0.9971634 4 1.0000000 4 1.0000000 4  
## 116 0.7737974 4 0.9999946 4 1.0000000 4  
## 117 0.9942812 4 1.0000000 4 1.0000000 4  
## 118 0.9352576 4 1.0000000 4 1.0000000 4  
## 119 0.9874133 2 1.0000000 2 1.0000000 2  
## 120 0.9968266 2 1.0000000 2 1.0000000 2  
## 121 0.9965538 2 1.0000000 2 0.6666667 2  
## 122 0.9965793 2 1.0000000 2 1.0000000 2  
## 123 0.9116388 2 0.9999722 2 1.0000000 2  
## 124 0.9965793 2 1.0000000 2 1.0000000 2  
## 125 0.9757082 2 0.9999999 2 1.0000000 2  
## 126 0.9922770 4 1.0000000 4 1.0000000 4  
## 127 0.9965793 2 1.0000000 2 1.0000000 2  
## 128 0.9968273 4 1.0000000 4 1.0000000 4  
## 129 0.9291437 4 1.0000000 4 1.0000000 4  
## 130 0.9965793 2 1.0000000 2 1.0000000 2  
## 131 0.9786748 4 0.9999849 4 1.0000000 4  
## 132 0.9966107 2 1.0000000 2 1.0000000 2  
## 133 0.5840371 4 1.0000000 4 1.0000000 4  
## 134 0.8791897 2 1.0000000 2 0.6000000 2  
## 135 0.9966107 2 1.0000000 2 1.0000000 2  
## 136 0.9079244 2 1.0000000 2 1.0000000 2  
## 137 0.7154638 4 0.6318544 4 0.7142857 2  
## 138 0.9894653 4 0.9992930 4 1.0000000 4  
## 139 0.9966107 2 1.0000000 2 1.0000000 2  
## 140 0.9914570 4 1.0000000 4 1.0000000 4  
## 141 0.9563946 2 0.9999990 2 1.0000000 2  
## 142 0.7976698 2 0.9999998 2 0.6000000 2  
## 143 0.9965793 2 1.0000000 2 1.0000000 2  
## 144 0.9965793 2 1.0000000 2 1.0000000 2  
## 145 0.9966107 2 1.0000000 2 1.0000000 2  
## 146 0.6072239 4 0.9989814 4 0.7142857 4  
## 147 0.9965793 2 1.0000000 2 1.0000000 2  
## 148 0.9701018 4 1.0000000 4 1.0000000 4  
## 149 0.9843055 4 1.0000000 4 1.0000000 4  
## 150 0.9983909 4 1.0000000 4 1.0000000 4  
## 151 0.9864424 2 1.0000000 2 1.0000000 2  
## 152 0.9966107 2 1.0000000 2 1.0000000 2  
## 153 0.9885678 2 1.0000000 2 1.0000000 2  
## 154 0.9957518 4 1.0000000 4 1.0000000 4  
## 155 0.9116388 2 0.9999993 2 1.0000000 2  
## 156 0.8303849 2 0.9301818 4 1.0000000 2  
## 157 0.9966107 2 1.0000000 2 1.0000000 2  
## 158 0.9387992 4 1.0000000 4 1.0000000 4  
## 159 0.9908284 2 1.0000000 2 0.6000000 2  
## 160 0.9975353 4 1.0000000 4 1.0000000 4  
## 161 0.9964013 4 1.0000000 4 1.0000000 4  
## 162 0.7204088 2 0.9999969 2 0.6666667 2  
## 163 0.9966107 2 1.0000000 2 1.0000000 2  
## 164 0.9073002 2 0.9999987 2 1.0000000 2  
## 165 0.9933663 4 0.9999999 4 1.0000000 4  
## 166 0.9974583 2 1.0000000 2 1.0000000 2  
## 167 0.9885678 2 0.9999999 2 1.0000000 2  
## 168 0.9896677 4 1.0000000 4 1.0000000 4  
## 169 0.8647687 2 0.9455414 2 1.0000000 2  
## 170 0.9966107 2 1.0000000 2 1.0000000 2  
## 171 0.9965793 2 1.0000000 2 1.0000000 2  
## 172 0.9307531 4 1.0000000 4 1.0000000 4  
## 173 0.9966107 2 1.0000000 2 1.0000000 2  
## 174 0.9554191 2 0.9999982 2 1.0000000 2  
## 175 0.9567787 2 1.0000000 2 1.0000000 2  
## 176 0.9862110 2 1.0000000 2 1.0000000 2  
## 177 0.9957271 4 1.0000000 4 1.0000000 4  
## 178 0.8097965 2 0.9987759 2 1.0000000 2  
## 179 0.9965793 2 1.0000000 2 1.0000000 2  
## 180 0.9974816 2 1.0000000 2 1.0000000 2  
## 181 0.9958034 4 1.0000000 4 1.0000000 4  
## 182 0.9877481 4 1.0000000 4 1.0000000 4  
## 183 0.5285795 2 0.9940232 2 0.6666667 4  
## 184 0.9984287 4 1.0000000 4 1.0000000 4  
## 185 0.9863181 2 1.0000000 2 1.0000000 2  
## 186 0.9913090 4 1.0000000 4 1.0000000 4  
## 187 0.8790029 4 0.9999889 4 1.0000000 4  
## 188 0.9860847 2 0.9999763 2 1.0000000 2  
## 189 0.9968557 2 1.0000000 2 1.0000000 2  
## 190 0.9828808 4 1.0000000 4 1.0000000 4  
## 191 0.8228600 4 0.9999997 4 1.0000000 4  
## 192 0.8689807 4 0.9996268 2 0.6000000 4  
## 193 0.9965793 2 1.0000000 2 1.0000000 2  
## 194 0.9965793 2 1.0000000 2 1.0000000 2  
## 195 0.7938157 4 0.9675846 2 0.6000000 4  
## 196 0.9693880 2 0.9999797 2 1.0000000 2  
## 197 0.9872978 2 1.0000000 2 1.0000000 2  
## 198 0.9870905 4 1.0000000 4 1.0000000 4  
## 199 0.9972601 2 1.0000000 2 1.0000000 2  
## 200 0.9616890 4 1.0000000 4 1.0000000 4  
## 201 0.9465640 4 1.0000000 4 1.0000000 4  
## 202 0.9965793 2 1.0000000 2 1.0000000 2  
## 203 0.6818246 4 0.9556302 4 0.7142857 4  
## 204 0.9568785 4 0.9999989 4 1.0000000 4  
## 205 0.5953231 4 0.9999995 4 1.0000000 4  
## 206 0.9759265 2 1.0000000 2 1.0000000 2  
## 207 0.9947650 2 1.0000000 2 1.0000000 2  
## 208 0.7229309 4 0.8047078 2 1.0000000 2  
## 209 0.9885678 2 0.9999998 2 1.0000000 2  
## 210 0.9965793 2 1.0000000 2 1.0000000 2  
## 211 0.9953945 4 1.0000000 4 1.0000000 4  
## 212 0.9716562 4 1.0000000 4 1.0000000 4  
## 213 0.9730755 4 1.0000000 4 1.0000000 4  
## 214 0.9757082 2 1.0000000 2 1.0000000 2  
## 215 0.9862110 2 0.9999998 2 1.0000000 2  
## 216 0.9966107 2 1.0000000 2 1.0000000 2  
## 217 0.9966107 2 1.0000000 2 1.0000000 2  
## 218 0.9684399 4 1.0000000 4 1.0000000 4  
## 219 0.9759265 2 1.0000000 2 1.0000000 2  
## 220 0.9972601 2 1.0000000 2 1.0000000 2  
## 221 0.8897483 2 0.9969031 2 1.0000000 2  
## 222 0.9965793 2 1.0000000 2 1.0000000 2  
## 223 0.9968557 2 1.0000000 2 1.0000000 2  
## 224 0.9966107 2 1.0000000 2 1.0000000 2  
## 225 0.9793629 2 1.0000000 2 1.0000000 2  
## 226 0.9965793 2 1.0000000 2 1.0000000 2  
## 227 0.9956044 4 1.0000000 4 1.0000000 4  
## 228 0.9370842 2 0.9999592 2 1.0000000 2  
## 229 0.9947650 2 1.0000000 2 1.0000000 2  
## 230 0.9953945 4 1.0000000 4 1.0000000 4  
## 231 0.9905633 4 0.9998762 2 0.6000000 4  
## 232 0.9871746 4 0.9999993 4 1.0000000 4  
## 233 0.9965793 2 1.0000000 2 1.0000000 2  
## 234 0.9961066 4 1.0000000 4 1.0000000 4  
## 235 0.9965793 2 1.0000000 2 1.0000000 2  
## 236 0.8854510 4 0.9999998 4 1.0000000 4  
## 237 0.9759265 2 1.0000000 2 1.0000000 2  
## 238 0.9980488 4 1.0000000 4 1.0000000 4  
## 239 0.9965793 2 1.0000000 2 1.0000000 2  
## 240 0.5903965 2 0.9737716 2 0.6000000 2  
## 241 0.9871983 2 0.9999997 2 1.0000000 2  
## 242 0.9073002 2 0.9997274 2 1.0000000 2  
## 243 0.9995848 4 1.0000000 4 1.0000000 4  
## 244 0.9965793 2 1.0000000 2 1.0000000 2  
## 245 0.9852430 4 1.0000000 4 1.0000000 4  
## 246 0.9757082 2 1.0000000 2 1.0000000 2  
## 247 0.9979431 4 1.0000000 4 1.0000000 4  
## 248 0.9991963 4 1.0000000 4 1.0000000 4  
## 249 0.8936967 2 0.9899921 2 0.6000000 2  
## 250 0.9864424 2 1.0000000 2 1.0000000 2  
## 251 0.7600283 4 0.9990062 4 0.7142857 4  
## 252 0.9958523 4 1.0000000 4 1.0000000 4  
## 253 0.9934142 4 1.0000000 4 1.0000000 4  
## 254 0.9965793 2 1.0000000 2 1.0000000 2  
## 255 0.9965793 2 1.0000000 2 1.0000000 2  
## 256 0.9920958 4 1.0000000 4 1.0000000 4  
## 257 0.9889324 2 1.0000000 2 1.0000000 2  
## 258 0.9759265 2 1.0000000 2 1.0000000 2  
## 259 0.9959268 4 1.0000000 4 1.0000000 4  
## 260 0.9664110 4 1.0000000 2 0.6000000 4  
## 261 0.9997123 4 1.0000000 4 1.0000000 4  
## 262 0.9965793 2 1.0000000 2 1.0000000 2  
## 263 0.9634930 4 1.0000000 4 1.0000000 4  
## 264 0.9868132 2 1.0000000 2 1.0000000 2  
## 265 0.9965019 4 1.0000000 4 1.0000000 4  
## 266 0.9868132 2 0.9999984 2 0.6666667 2  
## 267 0.9794385 4 1.0000000 4 1.0000000 4  
## 268 0.9965793 2 1.0000000 2 1.0000000 2  
## 269 0.9965793 2 1.0000000 2 1.0000000 2  
## 270 0.9759265 2 1.0000000 2 1.0000000 2  
## 271 0.9804614 2 0.9999995 2 1.0000000 2  
## 272 0.9885678 2 1.0000000 2 1.0000000 2  
## 273 0.9978463 4 1.0000000 4 1.0000000 4  
## 274 0.9965793 2 1.0000000 2 1.0000000 2  
## 275 0.5903965 2 0.9155565 2 0.6666667 2  
## 276 0.9966107 2 1.0000000 2 1.0000000 2  
## 277 0.9965793 2 1.0000000 2 1.0000000 2  
## 278 0.9966107 2 1.0000000 2 1.0000000 2  
## 279 0.9972601 2 1.0000000 2 1.0000000 2  
## 280 0.9729990 4 1.0000000 4 1.0000000 4  
## 281 0.9860847 2 0.9999763 2 1.0000000 2  
## 282 0.9885678 2 0.9999978 2 1.0000000 2  
## 283 0.9860847 2 1.0000000 2 1.0000000 2  
## 284 0.9635587 2 1.0000000 2 1.0000000 2  
## 285 0.9082031 4 1.0000000 4 1.0000000 4  
## 286 0.8862039 4 0.9999998 4 1.0000000 4  
## 287 0.9862110 2 0.9999960 2 1.0000000 2  
## 288 0.9968266 2 1.0000000 2 1.0000000 2  
## 289 0.9884343 4 1.0000000 4 1.0000000 4  
## 290 0.9973409 4 1.0000000 4 1.0000000 4  
## 291 0.9972601 2 1.0000000 2 1.0000000 2  
## 292 0.8487923 4 0.9998183 4 1.0000000 4  
## 293 0.9966107 2 1.0000000 2 1.0000000 2  
## 294 0.9863181 2 1.0000000 2 1.0000000 2  
## 295 0.9977156 4 1.0000000 4 1.0000000 4  
## 296 0.9972601 2 1.0000000 2 1.0000000 2  
## 297 0.7378430 2 0.7431518 2 1.0000000 2  
## 298 0.9757082 2 1.0000000 2 1.0000000 2  
## 299 0.9863181 2 1.0000000 2 1.0000000 2  
## 300 0.9993498 4 1.0000000 4 1.0000000 4  
## 301 0.9965793 2 1.0000000 2 1.0000000 2  
## 302 0.9970818 4 1.0000000 4 1.0000000 4  
## 303 0.9965854 2 1.0000000 2 0.6666667 2  
## 304 0.9972601 2 1.0000000 2 1.0000000 2  
## 305 0.9085830 2 1.0000000 2 1.0000000 2  
## 306 0.9968266 2 1.0000000 2 1.0000000 2  
## 307 0.9965793 2 1.0000000 2 1.0000000 2  
## 308 0.9835374 4 0.9999998 4 1.0000000 4  
## 309 0.9871983 2 0.9999997 2 1.0000000 2  
## 310 0.9966107 2 1.0000000 2 1.0000000 2  
## 311 0.9757082 2 1.0000000 2 1.0000000 2  
## 312 0.9965793 2 1.0000000 2 1.0000000 2  
## 313 0.9980945 4 1.0000000 4 1.0000000 4  
## 314 0.9965793 2 1.0000000 2 1.0000000 2  
## 315 0.9947650 2 1.0000000 2 1.0000000 2  
## 316 0.9959974 4 1.0000000 4 1.0000000 4  
## 317 0.9569178 2 0.9999921 2 0.6666667 2  
## 318 0.9725578 2 1.0000000 2 1.0000000 2  
## 319 0.9968266 2 1.0000000 2 1.0000000 2  
## 320 0.9757082 2 1.0000000 2 1.0000000 2  
## 321 0.9972397 2 1.0000000 2 0.6000000 2  
## 322 0.9757082 2 1.0000000 2 1.0000000 2  
## 323 0.9774303 2 1.0000000 2 1.0000000 2  
## 324 0.9037222 4 0.9989588 4 0.7142857 4  
## 325 0.9972852 2 1.0000000 2 1.0000000 2  
## 326 0.9966107 2 1.0000000 2 1.0000000 2  
## 327 0.9816307 4 1.0000000 4 1.0000000 4  
## 328 0.8689807 4 0.9999995 4 1.0000000 4  
## 329 0.9965793 2 1.0000000 2 1.0000000 2  
## 330 0.9966107 2 1.0000000 2 1.0000000 2  
## 331 0.9965793 2 1.0000000 2 1.0000000 2  
## 332 0.9944511 4 1.0000000 4 1.0000000 4  
## 333 0.9965793 2 1.0000000 2 1.0000000 2  
## 334 0.9563946 2 0.9999988 2 0.6666667 2  
## 335 0.9791838 4 1.0000000 4 1.0000000 4  
## 336 0.9759265 2 1.0000000 2 1.0000000 2  
## 337 0.8320607 4 0.9999873 4 1.0000000 4  
## 338 0.6212285 4 0.9962830 4 0.7142857 4  
## 339 0.9731345 2 0.9999930 2 1.0000000 2  
## 340 0.9972601 2 1.0000000 2 1.0000000 2  
## 341 0.9957271 4 1.0000000 4 1.0000000 4  
## 342 0.9864424 2 1.0000000 2 1.0000000 2  
## 343 0.7435776 2 0.8272626 2 1.0000000 2  
## 344 0.9071482 2 0.9995344 2 1.0000000 2  
## 345 0.9966107 2 1.0000000 2 1.0000000 2  
## 346 0.9966107 2 1.0000000 2 1.0000000 2  
## 347 0.9659076 2 1.0000000 2 0.6000000 2  
## 348 0.9817720 4 0.9999999 4 1.0000000 4  
## 349 0.9968266 2 1.0000000 2 1.0000000 2  
## 350 0.9807142 4 1.0000000 4 1.0000000 4  
## 351 0.9962293 4 1.0000000 4 1.0000000 4  
## 352 0.9140284 2 0.9999998 2 1.0000000 2  
## 353 0.9578776 2 0.9999997 2 0.6000000 2  
## 354 0.7435776 2 0.9994393 2 1.0000000 2  
## 355 0.9965793 2 1.0000000 2 1.0000000 2  
## 356 0.9864424 2 1.0000000 2 1.0000000 2  
## 357 0.9071482 2 0.9999214 2 1.0000000 2  
## 358 0.9965793 2 1.0000000 2 1.0000000 2  
## 359 0.9796883 2 0.9999999 2 1.0000000 2  
## 360 0.9869330 2 1.0000000 2 1.0000000 2  
## 361 0.9592802 4 0.9999195 2 0.6666667 2  
## 362 0.9966107 2 1.0000000 2 1.0000000 2  
## 363 0.9862110 2 1.0000000 2 1.0000000 2  
## 364 0.9968266 2 1.0000000 2 1.0000000 2  
## 365 0.9759265 2 1.0000000 2 1.0000000 2  
## 366 0.9965793 2 1.0000000 2 1.0000000 2  
## 367 0.9757082 2 1.0000000 2 1.0000000 2  
## 368 0.9965793 2 1.0000000 2 1.0000000 2  
## 369 0.9870809 2 1.0000000 2 1.0000000 2  
## 370 0.9965793 2 1.0000000 2 1.0000000 2  
## 371 0.8477905 4 0.9763385 2 0.6000000 2  
## 372 0.9966107 2 1.0000000 2 1.0000000 2  
## 373 0.9965793 2 1.0000000 2 1.0000000 2  
## 374 0.9393431 4 1.0000000 4 1.0000000 4  
## 375 0.9965793 2 1.0000000 2 1.0000000 2  
## 376 0.8096363 4 0.9999951 4 1.0000000 4  
## 377 0.9893882 2 1.0000000 2 1.0000000 2  
## 378 0.9872978 2 1.0000000 2 1.0000000 2  
## 379 0.9860847 2 0.9999994 2 1.0000000 2  
## 380 0.9965793 2 1.0000000 2 1.0000000 2  
## 381 0.9965793 2 1.0000000 2 1.0000000 2  
## 382 0.9965793 2 1.0000000 2 1.0000000 2  
## 383 0.9877579 2 0.9999984 2 0.6666667 2  
## 384 0.9965538 2 1.0000000 2 0.6666667 2  
## 385 0.7129202 4 0.9998936 4 0.7142857 2  
## 386 0.5241156 4 0.9998599 4 1.0000000 4  
## 387 0.9954705 4 1.0000000 4 1.0000000 4  
## 388 0.9869330 2 1.0000000 2 0.6666667 2  
## 389 0.9968557 2 1.0000000 2 1.0000000 2  
## 390 0.9460470 2 1.0000000 2 1.0000000 2  
## 391 0.9893882 2 1.0000000 2 1.0000000 2  
## 392 0.9871983 2 0.9999997 2 1.0000000 2  
## 393 0.8664449 2 0.5555316 4 0.7142857 2  
## 394 0.8822184 4 0.9999974 4 0.7142857 4  
## 395 0.9981262 4 1.0000000 4 1.0000000 4  
## 396 0.9966107 2 1.0000000 2 1.0000000 2  
## 397 0.9965793 2 1.0000000 2 1.0000000 2  
## 398 0.9983909 4 1.0000000 4 1.0000000 4  
## 399 0.9971634 4 1.0000000 4 1.0000000 4  
## 400 0.9548806 4 1.0000000 4 1.0000000 4  
## BAGGING\_PROB LOGITBOOST\_LABEL LOGITBOOST\_PROB FORESTS\_LABEL  
## 1 1.00 2 1.0000000 2  
## 2 1.00 2 1.0000000 2  
## 3 1.00 2 1.0000000 2  
## 4 1.00 2 1.0000000 2  
## 5 1.00 2 1.0000000 2  
## 6 1.00 2 1.0000000 2  
## 7 1.00 2 1.0000000 2  
## 8 0.68 4 0.8807971 4  
## 9 0.88 4 1.0000000 4  
## 10 1.00 4 1.0000000 4  
## 11 1.00 2 1.0000000 2  
## 12 0.88 4 1.0000000 4  
## 13 1.00 4 1.0000000 4  
## 14 0.80 4 1.0000000 4  
## 15 0.84 2 1.0000000 2  
## 16 1.00 2 1.0000000 2  
## 17 1.00 2 1.0000000 2  
## 18 1.00 4 1.0000000 4  
## 19 1.00 2 1.0000000 2  
## 20 0.64 4 0.9999999 4  
## 21 1.00 2 1.0000000 2  
## 22 0.96 4 1.0000000 4  
## 23 1.00 2 1.0000000 2  
## 24 1.00 2 1.0000000 2  
## 25 1.00 2 0.9999992 2  
## 26 0.84 4 0.9820138 4  
## 27 1.00 2 1.0000000 2  
## 28 1.00 2 1.0000000 2  
## 29 0.96 2 0.9999939 2  
## 30 1.00 2 0.9999999 2  
## 31 0.64 4 0.9999546 4  
## 32 0.52 4 1.0000000 4  
## 33 0.96 4 1.0000000 4  
## 34 0.88 2 0.9999999 2  
## 35 1.00 2 1.0000000 2  
## 36 0.76 4 1.0000000 4  
## 37 1.00 2 1.0000000 2  
## 38 0.92 4 1.0000000 4  
## 39 1.00 4 1.0000000 4  
## 40 0.76 4 1.0000000 4  
## 41 1.00 4 1.0000000 4  
## 42 1.00 4 1.0000000 4  
## 43 0.84 4 1.0000000 4  
## 44 1.00 2 1.0000000 2  
## 45 0.68 2 0.9996646 2  
## 46 0.72 4 0.9999999 4  
## 47 1.00 4 1.0000000 4  
## 48 1.00 2 0.9999939 2  
## 49 1.00 2 1.0000000 2  
## 50 0.76 4 1.0000000 4  
## 51 0.96 4 0.9999939 4  
## 52 1.00 2 1.0000000 2  
## 53 1.00 2 1.0000000 2  
## 54 1.00 2 1.0000000 2  
## 55 1.00 2 1.0000000 2  
## 56 0.96 2 1.0000000 2  
## 57 1.00 2 0.9999992 2  
## 58 1.00 2 1.0000000 2  
## 59 1.00 2 1.0000000 2  
## 60 0.96 2 1.0000000 2  
## 61 0.92 2 1.0000000 2  
## 62 0.96 2 1.0000000 2  
## 63 1.00 2 1.0000000 2  
## 64 1.00 2 1.0000000 2  
## 65 1.00 2 1.0000000 2  
## 66 1.00 2 1.0000000 2  
## 67 1.00 2 1.0000000 2  
## 68 1.00 2 1.0000000 2  
## 69 0.84 2 1.0000000 2  
## 70 1.00 4 1.0000000 4  
## 71 0.72 4 1.0000000 4  
## 72 1.00 2 1.0000000 2  
## 73 0.88 4 0.9999999 4  
## 74 0.96 2 0.9999992 2  
## 75 0.96 4 1.0000000 4  
## 76 1.00 4 1.0000000 4  
## 77 0.96 4 1.0000000 4  
## 78 1.00 2 1.0000000 2  
## 79 0.60 2 0.9996646 4  
## 80 0.76 2 1.0000000 2  
## 81 1.00 2 1.0000000 2  
## 82 0.88 4 1.0000000 4  
## 83 1.00 2 1.0000000 2  
## 84 1.00 2 1.0000000 2  
## 85 1.00 2 1.0000000 2  
## 86 0.88 4 0.9999992 4  
## 87 1.00 2 1.0000000 2  
## 88 0.76 4 0.9999546 2  
## 89 0.92 2 1.0000000 2  
## 90 0.96 2 1.0000000 2  
## 91 0.60 2 0.9999546 2  
## 92 1.00 2 1.0000000 2  
## 93 1.00 2 1.0000000 2  
## 94 0.84 4 1.0000000 4  
## 95 1.00 2 1.0000000 2  
## 96 0.72 4 0.9999992 4  
## 97 1.00 2 1.0000000 2  
## 98 0.60 2 0.5000000 4  
## 99 0.92 2 1.0000000 2  
## 100 1.00 2 1.0000000 2  
## 101 1.00 2 0.9999939 2  
## 102 0.96 4 0.9999999 4  
## 103 0.72 4 1.0000000 4  
## 104 0.92 4 1.0000000 4  
## 105 0.96 4 1.0000000 4  
## 106 0.96 2 1.0000000 2  
## 107 1.00 2 0.9999939 2  
## 108 1.00 2 0.9999992 2  
## 109 1.00 2 1.0000000 2  
## 110 1.00 4 1.0000000 4  
## 111 1.00 2 1.0000000 2  
## 112 1.00 2 1.0000000 2  
## 113 0.64 4 1.0000000 4  
## 114 1.00 2 1.0000000 2  
## 115 1.00 4 1.0000000 4  
## 116 0.80 4 1.0000000 4  
## 117 1.00 4 1.0000000 4  
## 118 1.00 4 0.9999992 4  
## 119 1.00 2 1.0000000 2  
## 120 1.00 2 1.0000000 2  
## 121 1.00 2 1.0000000 2  
## 122 1.00 2 1.0000000 2  
## 123 0.96 2 0.9999939 2  
## 124 1.00 2 1.0000000 2  
## 125 1.00 2 0.9999546 2  
## 126 0.84 4 1.0000000 4  
## 127 1.00 2 1.0000000 2  
## 128 1.00 4 1.0000000 4  
## 129 1.00 4 0.9999999 4  
## 130 1.00 2 1.0000000 2  
## 131 0.92 4 1.0000000 4  
## 132 1.00 2 1.0000000 2  
## 133 0.80 4 1.0000000 4  
## 134 0.88 2 1.0000000 2  
## 135 1.00 2 1.0000000 2  
## 136 0.96 2 0.9999992 2  
## 137 0.60 2 1.0000000 2  
## 138 0.96 4 1.0000000 4  
## 139 1.00 2 1.0000000 2  
## 140 0.96 4 1.0000000 4  
## 141 0.92 2 0.9999546 2  
## 142 0.72 2 1.0000000 2  
## 143 1.00 2 1.0000000 2  
## 144 1.00 2 1.0000000 2  
## 145 1.00 2 1.0000000 2  
## 146 0.68 2 0.5000000 4  
## 147 1.00 2 1.0000000 2  
## 148 1.00 4 1.0000000 4  
## 149 1.00 4 1.0000000 4  
## 150 1.00 4 1.0000000 4  
## 151 1.00 2 1.0000000 2  
## 152 1.00 2 1.0000000 2  
## 153 1.00 2 1.0000000 2  
## 154 1.00 4 1.0000000 4  
## 155 0.96 2 1.0000000 2  
## 156 0.64 2 0.9975274 2  
## 157 1.00 2 1.0000000 2  
## 158 1.00 4 1.0000000 4  
## 159 0.92 2 1.0000000 2  
## 160 0.96 4 1.0000000 4  
## 161 0.96 4 1.0000000 4  
## 162 0.76 2 0.9999546 2  
## 163 1.00 2 1.0000000 2  
## 164 0.84 2 0.9999546 2  
## 165 1.00 4 1.0000000 4  
## 166 1.00 2 1.0000000 2  
## 167 1.00 2 1.0000000 2  
## 168 0.96 4 1.0000000 4  
## 169 0.72 2 0.9996646 2  
## 170 1.00 2 1.0000000 2  
## 171 1.00 2 1.0000000 2  
## 172 1.00 4 1.0000000 4  
## 173 1.00 2 1.0000000 2  
## 174 0.96 2 0.8807971 2  
## 175 1.00 2 1.0000000 2  
## 176 1.00 2 1.0000000 2  
## 177 1.00 4 1.0000000 4  
## 178 0.84 2 0.9999992 2  
## 179 1.00 2 1.0000000 2  
## 180 1.00 2 1.0000000 2  
## 181 1.00 4 1.0000000 4  
## 182 1.00 4 1.0000000 4  
## 183 0.64 2 0.9996646 2  
## 184 1.00 4 1.0000000 4  
## 185 1.00 2 1.0000000 2  
## 186 0.92 4 1.0000000 4  
## 187 0.60 4 0.9975274 4  
## 188 0.88 2 0.9999999 2  
## 189 1.00 2 1.0000000 2  
## 190 0.88 4 1.0000000 4  
## 191 0.80 4 1.0000000 4  
## 192 0.60 4 0.9975274 4  
## 193 1.00 2 1.0000000 2  
## 194 1.00 2 1.0000000 2  
## 195 0.52 4 0.9975274 2  
## 196 0.96 2 1.0000000 2  
## 197 0.88 2 1.0000000 2  
## 198 1.00 4 1.0000000 4  
## 199 1.00 2 1.0000000 2  
## 200 0.92 4 0.9999999 4  
## 201 0.92 4 0.9999992 4  
## 202 1.00 2 1.0000000 2  
## 203 0.68 4 0.9975274 4  
## 204 0.76 4 1.0000000 4  
## 205 0.56 4 0.9996646 4  
## 206 1.00 2 1.0000000 2  
## 207 0.96 2 1.0000000 2  
## 208 0.52 4 0.9996646 4  
## 209 0.96 2 1.0000000 2  
## 210 1.00 2 1.0000000 2  
## 211 1.00 4 1.0000000 4  
## 212 1.00 4 1.0000000 4  
## 213 1.00 4 1.0000000 4  
## 214 1.00 2 0.9999939 2  
## 215 0.96 2 1.0000000 2  
## 216 1.00 2 1.0000000 2  
## 217 1.00 2 1.0000000 2  
## 218 0.92 4 0.9999992 4  
## 219 1.00 2 1.0000000 2  
## 220 1.00 2 1.0000000 2  
## 221 0.72 4 0.8807971 2  
## 222 1.00 2 1.0000000 2  
## 223 1.00 2 1.0000000 2  
## 224 1.00 2 1.0000000 2  
## 225 0.96 2 1.0000000 2  
## 226 1.00 2 1.0000000 2  
## 227 1.00 4 1.0000000 4  
## 228 0.72 2 1.0000000 2  
## 229 0.96 2 1.0000000 2  
## 230 1.00 4 1.0000000 4  
## 231 0.88 4 1.0000000 4  
## 232 0.80 4 1.0000000 4  
## 233 1.00 2 1.0000000 2  
## 234 0.96 4 1.0000000 4  
## 235 1.00 2 1.0000000 2  
## 236 0.56 4 0.9999999 4  
## 237 1.00 2 0.9999992 2  
## 238 0.96 4 1.0000000 4  
## 239 1.00 2 1.0000000 2  
## 240 0.60 2 0.9999999 2  
## 241 0.96 2 1.0000000 2  
## 242 0.76 2 0.9996646 2  
## 243 1.00 4 1.0000000 4  
## 244 1.00 2 1.0000000 2  
## 245 1.00 4 1.0000000 4  
## 246 1.00 2 1.0000000 2  
## 247 1.00 4 1.0000000 4  
## 248 0.96 4 1.0000000 4  
## 249 0.60 2 0.9975274 2  
## 250 1.00 2 1.0000000 2  
## 251 0.64 2 0.5000000 4  
## 252 1.00 4 1.0000000 4  
## 253 1.00 4 1.0000000 4  
## 254 1.00 2 1.0000000 2  
## 255 1.00 2 1.0000000 2  
## 256 0.92 4 1.0000000 4  
## 257 1.00 2 1.0000000 2  
## 258 1.00 2 1.0000000 2  
## 259 0.96 4 1.0000000 4  
## 260 0.88 4 1.0000000 4  
## 261 1.00 4 1.0000000 4  
## 262 1.00 2 1.0000000 2  
## 263 0.76 4 1.0000000 4  
## 264 0.92 2 1.0000000 2  
## 265 1.00 4 1.0000000 4  
## 266 0.96 2 1.0000000 2  
## 267 1.00 4 1.0000000 4  
## 268 1.00 2 1.0000000 2  
## 269 1.00 2 1.0000000 2  
## 270 1.00 2 1.0000000 2  
## 271 0.92 2 1.0000000 2  
## 272 1.00 2 1.0000000 2  
## 273 1.00 4 1.0000000 4  
## 274 1.00 2 1.0000000 2  
## 275 0.64 2 0.9999999 2  
## 276 1.00 2 1.0000000 2  
## 277 1.00 2 1.0000000 2  
## 278 1.00 2 1.0000000 2  
## 279 1.00 2 1.0000000 2  
## 280 1.00 4 1.0000000 4  
## 281 0.88 2 0.9999999 2  
## 282 0.96 2 0.9999999 2  
## 283 1.00 2 1.0000000 2  
## 284 0.96 2 0.9999546 2  
## 285 1.00 4 1.0000000 4  
## 286 0.76 4 1.0000000 4  
## 287 0.92 2 1.0000000 2  
## 288 1.00 2 1.0000000 2  
## 289 1.00 4 1.0000000 4  
## 290 0.96 4 1.0000000 4  
## 291 1.00 2 1.0000000 2  
## 292 0.76 4 0.9996646 4  
## 293 1.00 2 1.0000000 2  
## 294 0.96 2 1.0000000 2  
## 295 1.00 4 1.0000000 4  
## 296 1.00 2 1.0000000 2  
## 297 0.64 2 0.9975274 2  
## 298 1.00 2 0.9999992 2  
## 299 1.00 2 1.0000000 2  
## 300 1.00 4 1.0000000 4  
## 301 1.00 2 1.0000000 2  
## 302 1.00 4 1.0000000 4  
## 303 1.00 2 1.0000000 2  
## 304 1.00 2 1.0000000 2  
## 305 1.00 2 0.9999992 2  
## 306 1.00 2 1.0000000 2  
## 307 1.00 2 1.0000000 2  
## 308 0.92 4 1.0000000 4  
## 309 0.96 2 1.0000000 2  
## 310 1.00 2 1.0000000 2  
## 311 1.00 2 0.9999939 2  
## 312 1.00 2 1.0000000 2  
## 313 1.00 4 1.0000000 4  
## 314 1.00 2 1.0000000 2  
## 315 0.96 2 1.0000000 2  
## 316 0.96 4 1.0000000 4  
## 317 1.00 2 0.9999992 2  
## 318 0.88 2 1.0000000 2  
## 319 1.00 2 1.0000000 2  
## 320 1.00 2 0.9996646 2  
## 321 0.92 2 1.0000000 2  
## 322 1.00 2 0.9999939 2  
## 323 1.00 2 0.9999992 2  
## 324 0.96 4 0.9999992 4  
## 325 1.00 2 1.0000000 2  
## 326 1.00 2 1.0000000 2  
## 327 1.00 4 1.0000000 4  
## 328 0.88 4 1.0000000 4  
## 329 1.00 2 1.0000000 2  
## 330 1.00 2 1.0000000 2  
## 331 1.00 2 1.0000000 2  
## 332 0.76 4 1.0000000 4  
## 333 1.00 2 1.0000000 2  
## 334 0.88 2 0.9996646 2  
## 335 0.96 4 1.0000000 4  
## 336 1.00 2 0.9999999 2  
## 337 0.84 4 0.9999546 4  
## 338 0.68 2 0.9820138 4  
## 339 1.00 2 0.9999999 2  
## 340 1.00 2 1.0000000 2  
## 341 1.00 4 1.0000000 4  
## 342 1.00 2 1.0000000 2  
## 343 0.80 4 0.8807971 2  
## 344 0.84 2 0.9820138 2  
## 345 1.00 2 1.0000000 2  
## 346 1.00 2 1.0000000 2  
## 347 1.00 2 1.0000000 2  
## 348 0.96 4 1.0000000 4  
## 349 1.00 2 1.0000000 2  
## 350 0.96 4 1.0000000 4  
## 351 1.00 4 1.0000000 4  
## 352 0.92 2 1.0000000 2  
## 353 0.68 2 1.0000000 2  
## 354 0.84 2 0.9999939 2  
## 355 1.00 2 1.0000000 2  
## 356 1.00 2 1.0000000 2  
## 357 0.88 2 0.9820138 2  
## 358 1.00 2 1.0000000 2  
## 359 1.00 2 1.0000000 2  
## 360 1.00 2 1.0000000 2  
## 361 0.56 4 1.0000000 4  
## 362 1.00 2 1.0000000 2  
## 363 1.00 2 1.0000000 2  
## 364 1.00 2 1.0000000 2  
## 365 1.00 2 1.0000000 2  
## 366 1.00 2 1.0000000 2  
## 367 1.00 2 0.9999999 2  
## 368 1.00 2 1.0000000 2  
## 369 1.00 2 1.0000000 2  
## 370 1.00 2 1.0000000 2  
## 371 0.52 2 0.9999939 2  
## 372 1.00 2 1.0000000 2  
## 373 1.00 2 1.0000000 2  
## 374 0.92 4 0.9999999 4  
## 375 1.00 2 1.0000000 2  
## 376 0.88 4 0.9999546 4  
## 377 1.00 2 1.0000000 2  
## 378 1.00 2 1.0000000 2  
## 379 0.92 2 1.0000000 2  
## 380 1.00 2 1.0000000 2  
## 381 1.00 2 1.0000000 2  
## 382 1.00 2 1.0000000 2  
## 383 0.96 2 1.0000000 2  
## 384 0.96 2 1.0000000 2  
## 385 0.68 2 0.9820138 4  
## 386 0.68 4 0.9975274 4  
## 387 1.00 4 1.0000000 4  
## 388 1.00 2 1.0000000 2  
## 389 1.00 2 1.0000000 2  
## 390 1.00 2 1.0000000 2  
## 391 1.00 2 1.0000000 2  
## 392 0.96 2 1.0000000 2  
## 393 0.72 4 0.9999939 4  
## 394 0.84 4 1.0000000 4  
## 395 1.00 4 1.0000000 4  
## 396 1.00 2 1.0000000 2  
## 397 1.00 2 1.0000000 2  
## 398 1.00 4 1.0000000 4  
## 399 0.88 4 1.0000000 4  
## 400 0.96 4 1.0000000 4  
## FORESTS\_PROB MANUAL\_CODE CONSENSUS\_CODE CONSENSUS\_AGREE  
## 1 0.995 2 2 8  
## 2 0.995 2 2 8  
## 3 0.985 2 2 8  
## 4 1.000 2 2 8  
## 5 0.995 2 2 8  
## 6 0.985 2 2 8  
## 7 1.000 2 2 8  
## 8 0.760 4 4 8  
## 9 0.720 4 4 8  
## 10 0.995 4 4 8  
## 11 0.995 2 2 8  
## 12 0.945 4 4 8  
## 13 0.995 4 4 8  
## 14 0.905 4 4 8  
## 15 0.895 2 2 8  
## 16 0.970 2 2 8  
## 17 1.000 2 2 8  
## 18 0.950 4 4 8  
## 19 0.975 2 2 8  
## 20 0.880 4 4 8  
## 21 1.000 2 2 8  
## 22 0.895 4 4 8  
## 23 1.000 2 2 8  
## 24 1.000 2 2 8  
## 25 1.000 2 2 8  
## 26 0.605 4 2 5  
## 27 1.000 2 2 8  
## 28 1.000 2 2 8  
## 29 0.870 2 2 8  
## 30 0.995 2 2 8  
## 31 0.690 4 4 7  
## 32 0.755 2 4 7  
## 33 0.950 2 4 8  
## 34 0.875 2 2 8  
## 35 1.000 2 2 8  
## 36 0.890 4 4 8  
## 37 1.000 2 2 8  
## 38 0.955 4 4 8  
## 39 0.985 4 4 8  
## 40 0.765 2 4 8  
## 41 1.000 4 4 8  
## 42 0.965 4 4 8  
## 43 0.960 4 4 8  
## 44 0.975 2 2 8  
## 45 0.835 2 2 8  
## 46 0.750 4 4 7  
## 47 1.000 4 4 8  
## 48 0.985 2 2 8  
## 49 1.000 2 2 8  
## 50 0.885 4 4 8  
## 51 0.900 4 4 8  
## 52 0.985 2 2 8  
## 53 0.970 2 2 8  
## 54 1.000 2 2 8  
## 55 0.925 2 2 8  
## 56 0.985 2 2 8  
## 57 0.955 2 2 8  
## 58 1.000 2 2 8  
## 59 1.000 2 2 8  
## 60 0.980 2 2 8  
## 61 0.995 2 2 8  
## 62 0.965 2 2 8  
## 63 1.000 2 2 8  
## 64 1.000 2 2 8  
## 65 0.995 2 2 8  
## 66 0.985 2 2 8  
## 67 1.000 2 2 8  
## 68 0.965 2 2 8  
## 69 0.860 2 2 8  
## 70 1.000 4 4 8  
## 71 0.800 4 4 8  
## 72 0.995 2 2 8  
## 73 0.845 2 4 8  
## 74 0.850 2 2 8  
## 75 0.835 4 4 7  
## 76 0.985 4 4 8  
## 77 0.895 4 4 8  
## 78 0.980 2 2 8  
## 79 0.525 4 2 5  
## 80 0.700 2 2 8  
## 81 1.000 2 2 8  
## 82 0.845 4 4 8  
## 83 1.000 2 2 8  
## 84 1.000 2 2 8  
## 85 1.000 2 2 8  
## 86 0.780 2 4 8  
## 87 1.000 2 2 8  
## 88 0.535 4 4 5  
## 89 0.910 2 2 8  
## 90 0.980 2 2 8  
## 91 0.515 4 2 6  
## 92 1.000 2 2 8  
## 93 1.000 2 2 8  
## 94 0.890 4 4 8  
## 95 1.000 2 2 8  
## 96 0.715 4 4 8  
## 97 1.000 2 2 8  
## 98 0.865 4 4 7  
## 99 0.945 2 2 8  
## 100 0.985 2 2 8  
## 101 0.985 2 2 8  
## 102 0.935 4 4 7  
## 103 0.940 4 4 8  
## 104 0.900 4 4 7  
## 105 0.880 4 4 8  
## 106 0.995 2 2 8  
## 107 0.985 2 2 8  
## 108 1.000 2 2 8  
## 109 1.000 2 2 8  
## 110 0.980 4 4 8  
## 111 0.975 2 2 8  
## 112 1.000 2 2 8  
## 113 0.945 4 4 8  
## 114 1.000 2 2 8  
## 115 0.940 4 4 8  
## 116 0.845 4 4 8  
## 117 0.995 4 4 8  
## 118 0.905 4 4 8  
## 119 0.980 2 2 8  
## 120 1.000 2 2 8  
## 121 0.970 2 2 8  
## 122 1.000 2 2 8  
## 123 0.980 2 2 8  
## 124 0.990 2 2 8  
## 125 0.970 2 2 8  
## 126 0.925 2 4 8  
## 127 1.000 2 2 8  
## 128 1.000 4 4 8  
## 129 0.880 4 4 8  
## 130 1.000 2 2 8  
## 131 0.850 4 4 8  
## 132 1.000 2 2 8  
## 133 0.850 4 4 8  
## 134 0.770 2 2 8  
## 135 1.000 2 2 8  
## 136 0.960 2 2 8  
## 137 0.650 2 2 6  
## 138 0.865 4 4 8  
## 139 1.000 2 2 8  
## 140 0.955 4 4 8  
## 141 0.900 2 2 8  
## 142 0.785 2 2 8  
## 143 1.000 2 2 8  
## 144 1.000 2 2 8  
## 145 1.000 2 2 8  
## 146 0.610 4 4 6  
## 147 1.000 2 2 8  
## 148 0.960 4 4 8  
## 149 0.980 4 4 8  
## 150 1.000 4 4 8  
## 151 0.955 2 2 8  
## 152 0.990 2 2 8  
## 153 0.970 2 2 8  
## 154 0.990 4 4 8  
## 155 0.935 2 2 8  
## 156 0.665 2 2 7  
## 157 1.000 2 2 8  
## 158 0.945 4 4 8  
## 159 0.840 2 2 8  
## 160 0.975 4 4 8  
## 161 0.970 4 4 8  
## 162 0.840 2 2 8  
## 163 1.000 2 2 8  
## 164 0.805 2 2 8  
## 165 0.910 2 4 8  
## 166 1.000 2 2 8  
## 167 0.965 2 2 8  
## 168 0.975 4 4 8  
## 169 0.615 2 2 8  
## 170 1.000 2 2 8  
## 171 1.000 2 2 8  
## 172 0.885 4 4 8  
## 173 1.000 2 2 8  
## 174 0.875 2 2 8  
## 175 0.985 2 2 8  
## 176 1.000 2 2 8  
## 177 1.000 4 4 8  
## 178 0.715 2 2 8  
## 179 1.000 2 2 8  
## 180 1.000 2 2 8  
## 181 1.000 4 4 8  
## 182 0.945 4 4 8  
## 183 0.620 2 2 4  
## 184 0.975 4 4 8  
## 185 0.995 2 2 8  
## 186 0.950 4 4 8  
## 187 0.890 4 4 8  
## 188 0.950 2 2 8  
## 189 1.000 2 2 8  
## 190 0.945 4 4 8  
## 191 0.850 4 4 8  
## 192 0.685 4 4 6  
## 193 1.000 2 2 8  
## 194 1.000 2 2 8  
## 195 0.540 4 4 6  
## 196 0.840 2 2 8  
## 197 0.950 2 2 8  
## 198 0.985 4 4 8  
## 199 1.000 2 2 8  
## 200 0.975 4 4 8  
## 201 0.860 4 4 8  
## 202 1.000 2 2 8  
## 203 0.680 4 4 6  
## 204 0.870 4 4 8  
## 205 0.810 4 4 8  
## 206 0.995 2 2 8  
## 207 0.935 2 2 8  
## 208 0.640 4 4 6  
## 209 0.915 2 2 8  
## 210 0.995 2 2 8  
## 211 0.980 4 4 8  
## 212 0.895 4 4 8  
## 213 0.980 4 4 8  
## 214 0.985 2 2 8  
## 215 0.980 2 2 8  
## 216 1.000 2 2 8  
## 217 1.000 2 2 8  
## 218 0.940 4 4 8  
## 219 0.935 2 2 8  
## 220 1.000 2 2 8  
## 221 0.535 4 2 6  
## 222 1.000 2 2 8  
## 223 0.990 2 2 8  
## 224 1.000 2 2 8  
## 225 0.900 2 2 8  
## 226 0.960 2 2 8  
## 227 0.990 4 4 8  
## 228 0.720 2 2 8  
## 229 0.945 2 2 8  
## 230 0.985 4 4 8  
## 231 0.825 4 4 7  
## 232 0.840 4 4 8  
## 233 0.975 2 2 8  
## 234 0.965 4 4 8  
## 235 1.000 2 2 8  
## 236 0.790 4 4 8  
## 237 0.985 2 2 8  
## 238 0.985 4 4 8  
## 239 1.000 2 2 8  
## 240 0.625 4 2 8  
## 241 0.980 2 2 8  
## 242 0.785 2 2 8  
## 243 1.000 4 4 8  
## 244 1.000 2 2 8  
## 245 0.970 2 4 8  
## 246 0.985 2 2 8  
## 247 1.000 4 4 8  
## 248 0.975 4 4 8  
## 249 0.625 4 2 6  
## 250 0.975 2 2 8  
## 251 0.555 4 4 6  
## 252 0.995 4 4 8  
## 253 0.980 4 4 8  
## 254 1.000 2 2 8  
## 255 1.000 2 2 8  
## 256 0.965 4 4 8  
## 257 0.960 2 2 8  
## 258 0.995 2 2 8  
## 259 0.950 4 4 8  
## 260 0.915 2 4 7  
## 261 1.000 4 4 8  
## 262 1.000 2 2 8  
## 263 0.915 4 4 8  
## 264 0.935 2 2 8  
## 265 0.995 4 4 8  
## 266 0.980 2 2 8  
## 267 0.995 4 4 8  
## 268 1.000 2 2 8  
## 269 1.000 2 2 8  
## 270 0.995 2 2 8  
## 271 0.890 2 2 8  
## 272 0.970 2 2 8  
## 273 0.970 4 4 8  
## 274 0.995 2 2 8  
## 275 0.700 2 2 8  
## 276 1.000 2 2 8  
## 277 1.000 2 2 8  
## 278 1.000 2 2 8  
## 279 1.000 2 2 8  
## 280 0.920 4 4 8  
## 281 0.950 2 2 8  
## 282 0.935 2 2 8  
## 283 0.985 2 2 8  
## 284 0.915 2 2 8  
## 285 0.945 4 4 8  
## 286 0.880 4 4 8  
## 287 0.975 2 2 8  
## 288 1.000 2 2 8  
## 289 1.000 4 4 8  
## 290 0.985 4 4 8  
## 291 0.995 2 2 8  
## 292 0.725 4 4 8  
## 293 1.000 2 2 8  
## 294 0.940 2 2 8  
## 295 0.995 4 4 8  
## 296 1.000 2 2 8  
## 297 0.650 2 2 6  
## 298 1.000 2 2 8  
## 299 0.980 2 2 8  
## 300 0.985 4 4 8  
## 301 1.000 2 2 8  
## 302 0.970 4 4 8  
## 303 1.000 2 2 8  
## 304 0.990 2 2 8  
## 305 0.945 2 2 8  
## 306 1.000 2 2 8  
## 307 1.000 2 2 8  
## 308 0.890 2 4 8  
## 309 0.980 2 2 8  
## 310 1.000 2 2 8  
## 311 0.985 2 2 8  
## 312 1.000 2 2 8  
## 313 0.995 4 4 8  
## 314 1.000 2 2 8  
## 315 0.935 2 2 8  
## 316 0.990 4 4 8  
## 317 0.970 2 2 8  
## 318 0.720 4 2 8  
## 319 0.960 2 2 8  
## 320 0.980 2 2 8  
## 321 0.890 2 2 8  
## 322 0.985 2 2 8  
## 323 0.975 2 2 8  
## 324 0.825 4 4 8  
## 325 0.985 2 2 8  
## 326 1.000 2 2 8  
## 327 0.955 4 4 8  
## 328 0.890 4 4 8  
## 329 1.000 2 2 8  
## 330 1.000 2 2 8  
## 331 1.000 2 2 8  
## 332 0.900 4 4 8  
## 333 0.995 2 2 8  
## 334 0.895 2 2 8  
## 335 0.900 4 4 8  
## 336 0.880 2 2 8  
## 337 0.675 4 4 8  
## 338 0.680 4 4 7  
## 339 0.890 2 2 8  
## 340 0.990 2 2 8  
## 341 1.000 4 4 8  
## 342 0.975 2 2 8  
## 343 0.725 2 2 5  
## 344 0.825 2 2 7  
## 345 1.000 2 2 8  
## 346 1.000 2 2 8  
## 347 0.910 2 2 8  
## 348 0.940 4 4 8  
## 349 1.000 2 2 8  
## 350 0.925 4 4 8  
## 351 0.995 4 4 8  
## 352 0.970 2 2 8  
## 353 0.700 2 2 8  
## 354 0.855 2 2 8  
## 355 0.995 2 2 8  
## 356 0.975 2 2 8  
## 357 0.845 2 2 7  
## 358 0.995 2 2 8  
## 359 0.975 2 2 8  
## 360 0.995 2 2 8  
## 361 0.700 4 4 6  
## 362 1.000 2 2 8  
## 363 1.000 2 2 8  
## 364 1.000 2 2 8  
## 365 0.995 2 2 8  
## 366 1.000 2 2 8  
## 367 0.995 2 2 8  
## 368 1.000 2 2 8  
## 369 0.990 2 2 8  
## 370 1.000 2 2 8  
## 371 0.505 4 2 7  
## 372 1.000 2 2 8  
## 373 1.000 2 2 8  
## 374 0.930 4 4 8  
## 375 1.000 2 2 8  
## 376 0.705 4 4 8  
## 377 0.970 2 2 8  
## 378 1.000 2 2 8  
## 379 0.955 2 2 8  
## 380 1.000 2 2 8  
## 381 1.000 2 2 8  
## 382 1.000 2 2 8  
## 383 0.980 2 2 8  
## 384 0.985 2 2 8  
## 385 0.730 4 2 4  
## 386 0.665 4 4 8  
## 387 0.970 4 4 8  
## 388 0.995 2 2 8  
## 389 0.980 2 2 8  
## 390 0.920 2 2 8  
## 391 0.970 2 2 8  
## 392 0.980 2 2 8  
## 393 0.530 2 4 6  
## 394 0.785 4 4 8  
## 395 1.000 4 4 8  
## 396 1.000 2 2 8  
## 397 1.000 2 2 8  
## 398 1.000 4 4 8  
## 399 0.950 4 4 8  
## 400 0.945 4 4 8  
## CONSENSUS\_INCORRECT PROBABILITY\_CODE PROBABILITY\_INCORRECT  
## 1 0 2 0  
## 2 0 2 0  
## 3 0 2 0  
## 4 0 2 0  
## 5 0 2 0  
## 6 0 2 0  
## 7 0 2 0  
## 8 0 4 0  
## 9 0 4 0  
## 10 0 4 0  
## 11 0 2 0  
## 12 0 4 0  
## 13 0 4 0  
## 14 0 4 0  
## 15 0 2 0  
## 16 0 2 0  
## 17 0 2 0  
## 18 0 4 0  
## 19 0 2 0  
## 20 0 4 0  
## 21 0 2 0  
## 22 0 4 0  
## 23 0 2 0  
## 24 0 2 0  
## 25 0 2 0  
## 26 1 4 0  
## 27 0 2 0  
## 28 0 2 0  
## 29 0 2 0  
## 30 0 2 0  
## 31 0 4 0  
## 32 1 4 1  
## 33 1 4 1  
## 34 0 2 0  
## 35 0 2 0  
## 36 0 4 0  
## 37 0 2 0  
## 38 0 4 0  
## 39 0 4 0  
## 40 1 4 1  
## 41 0 4 0  
## 42 0 4 0  
## 43 0 4 0  
## 44 0 2 0  
## 45 0 2 0  
## 46 0 4 0  
## 47 0 4 0  
## 48 0 2 0  
## 49 0 2 0  
## 50 0 4 0  
## 51 0 4 0  
## 52 0 2 0  
## 53 0 2 0  
## 54 0 2 0  
## 55 0 2 0  
## 56 0 2 0  
## 57 0 2 0  
## 58 0 2 0  
## 59 0 2 0  
## 60 0 2 0  
## 61 0 2 0  
## 62 0 2 0  
## 63 0 2 0  
## 64 0 2 0  
## 65 0 2 0  
## 66 0 2 0  
## 67 0 2 0  
## 68 0 2 0  
## 69 0 2 0  
## 70 0 4 0  
## 71 0 4 0  
## 72 0 2 0  
## 73 1 4 1  
## 74 0 2 0  
## 75 0 4 0  
## 76 0 4 0  
## 77 0 4 0  
## 78 0 2 0  
## 79 1 2 1  
## 80 0 2 0  
## 81 0 2 0  
## 82 0 4 0  
## 83 0 2 0  
## 84 0 2 0  
## 85 0 2 0  
## 86 1 4 1  
## 87 0 2 0  
## 88 0 4 0  
## 89 0 2 0  
## 90 0 2 0  
## 91 1 2 1  
## 92 0 2 0  
## 93 0 2 0  
## 94 0 4 0  
## 95 0 2 0  
## 96 0 4 0  
## 97 0 2 0  
## 98 0 4 0  
## 99 0 2 0  
## 100 0 2 0  
## 101 0 2 0  
## 102 0 4 0  
## 103 0 4 0  
## 104 0 4 0  
## 105 0 4 0  
## 106 0 2 0  
## 107 0 2 0  
## 108 0 2 0  
## 109 0 2 0  
## 110 0 4 0  
## 111 0 2 0  
## 112 0 2 0  
## 113 0 4 0  
## 114 0 2 0  
## 115 0 4 0  
## 116 0 4 0  
## 117 0 4 0  
## 118 0 4 0  
## 119 0 2 0  
## 120 0 2 0  
## 121 0 2 0  
## 122 0 2 0  
## 123 0 2 0  
## 124 0 2 0  
## 125 0 2 0  
## 126 1 4 1  
## 127 0 2 0  
## 128 0 4 0  
## 129 0 4 0  
## 130 0 2 0  
## 131 0 4 0  
## 132 0 2 0  
## 133 0 4 0  
## 134 0 2 0  
## 135 0 2 0  
## 136 0 2 0  
## 137 0 2 0  
## 138 0 4 0  
## 139 0 2 0  
## 140 0 4 0  
## 141 0 2 0  
## 142 0 2 0  
## 143 0 2 0  
## 144 0 2 0  
## 145 0 2 0  
## 146 0 4 0  
## 147 0 2 0  
## 148 0 4 0  
## 149 0 4 0  
## 150 0 4 0  
## 151 0 2 0  
## 152 0 2 0  
## 153 0 2 0  
## 154 0 4 0  
## 155 0 2 0  
## 156 0 4 1  
## 157 0 2 0  
## 158 0 4 0  
## 159 0 2 0  
## 160 0 4 0  
## 161 0 4 0  
## 162 0 2 0  
## 163 0 2 0  
## 164 0 2 0  
## 165 1 4 1  
## 166 0 2 0  
## 167 0 2 0  
## 168 0 4 0  
## 169 0 2 0  
## 170 0 2 0  
## 171 0 2 0  
## 172 0 4 0  
## 173 0 2 0  
## 174 0 2 0  
## 175 0 2 0  
## 176 0 2 0  
## 177 0 4 0  
## 178 0 2 0  
## 179 0 2 0  
## 180 0 2 0  
## 181 0 4 0  
## 182 0 4 0  
## 183 0 2 0  
## 184 0 4 0  
## 185 0 2 0  
## 186 0 4 0  
## 187 0 4 0  
## 188 0 2 0  
## 189 0 2 0  
## 190 0 4 0  
## 191 0 4 0  
## 192 0 4 0  
## 193 0 2 0  
## 194 0 2 0  
## 195 0 4 0  
## 196 0 2 0  
## 197 0 2 0  
## 198 0 4 0  
## 199 0 2 0  
## 200 0 4 0  
## 201 0 4 0  
## 202 0 2 0  
## 203 0 4 0  
## 204 0 4 0  
## 205 0 4 0  
## 206 0 2 0  
## 207 0 2 0  
## 208 0 2 1  
## 209 0 2 0  
## 210 0 2 0  
## 211 0 4 0  
## 212 0 4 0  
## 213 0 4 0  
## 214 0 2 0  
## 215 0 2 0  
## 216 0 2 0  
## 217 0 2 0  
## 218 0 4 0  
## 219 0 2 0  
## 220 0 2 0  
## 221 1 2 1  
## 222 0 2 0  
## 223 0 2 0  
## 224 0 2 0  
## 225 0 2 0  
## 226 0 2 0  
## 227 0 4 0  
## 228 0 2 0  
## 229 0 2 0  
## 230 0 4 0  
## 231 0 4 0  
## 232 0 4 0  
## 233 0 2 0  
## 234 0 4 0  
## 235 0 2 0  
## 236 0 4 0  
## 237 0 2 0  
## 238 0 4 0  
## 239 0 2 0  
## 240 1 2 1  
## 241 0 2 0  
## 242 0 2 0  
## 243 0 4 0  
## 244 0 2 0  
## 245 1 4 1  
## 246 0 2 0  
## 247 0 4 0  
## 248 0 4 0  
## 249 1 2 1  
## 250 0 2 0  
## 251 0 4 0  
## 252 0 4 0  
## 253 0 4 0  
## 254 0 2 0  
## 255 0 2 0  
## 256 0 4 0  
## 257 0 2 0  
## 258 0 2 0  
## 259 0 4 0  
## 260 1 4 1  
## 261 0 4 0  
## 262 0 2 0  
## 263 0 4 0  
## 264 0 2 0  
## 265 0 4 0  
## 266 0 2 0  
## 267 0 4 0  
## 268 0 2 0  
## 269 0 2 0  
## 270 0 2 0  
## 271 0 2 0  
## 272 0 2 0  
## 273 0 4 0  
## 274 0 2 0  
## 275 0 2 0  
## 276 0 2 0  
## 277 0 2 0  
## 278 0 2 0  
## 279 0 2 0  
## 280 0 4 0  
## 281 0 2 0  
## 282 0 2 0  
## 283 0 2 0  
## 284 0 2 0  
## 285 0 4 0  
## 286 0 4 0  
## 287 0 2 0  
## 288 0 2 0  
## 289 0 4 0  
## 290 0 4 0  
## 291 0 2 0  
## 292 0 4 0  
## 293 0 2 0  
## 294 0 2 0  
## 295 0 4 0  
## 296 0 2 0  
## 297 0 2 0  
## 298 0 2 0  
## 299 0 2 0  
## 300 0 4 0  
## 301 0 2 0  
## 302 0 4 0  
## 303 0 2 0  
## 304 0 2 0  
## 305 0 2 0  
## 306 0 2 0  
## 307 0 2 0  
## 308 1 4 1  
## 309 0 2 0  
## 310 0 2 0  
## 311 0 2 0  
## 312 0 2 0  
## 313 0 4 0  
## 314 0 2 0  
## 315 0 2 0  
## 316 0 4 0  
## 317 0 2 0  
## 318 1 2 1  
## 319 0 2 0  
## 320 0 2 0  
## 321 0 2 0  
## 322 0 2 0  
## 323 0 2 0  
## 324 0 4 0  
## 325 0 2 0  
## 326 0 2 0  
## 327 0 4 0  
## 328 0 4 0  
## 329 0 2 0  
## 330 0 2 0  
## 331 0 2 0  
## 332 0 4 0  
## 333 0 2 0  
## 334 0 2 0  
## 335 0 4 0  
## 336 0 2 0  
## 337 0 4 0  
## 338 0 4 0  
## 339 0 2 0  
## 340 0 2 0  
## 341 0 4 0  
## 342 0 2 0  
## 343 0 2 0  
## 344 0 2 0  
## 345 0 2 0  
## 346 0 2 0  
## 347 0 2 0  
## 348 0 4 0  
## 349 0 2 0  
## 350 0 4 0  
## 351 0 4 0  
## 352 0 2 0  
## 353 0 2 0  
## 354 0 2 0  
## 355 0 2 0  
## 356 0 2 0  
## 357 0 2 0  
## 358 0 2 0  
## 359 0 2 0  
## 360 0 2 0  
## 361 0 4 0  
## 362 0 2 0  
## 363 0 2 0  
## 364 0 2 0  
## 365 0 2 0  
## 366 0 2 0  
## 367 0 2 0  
## 368 0 2 0  
## 369 0 2 0  
## 370 0 2 0  
## 371 1 2 1  
## 372 0 2 0  
## 373 0 2 0  
## 374 0 4 0  
## 375 0 2 0  
## 376 0 4 0  
## 377 0 2 0  
## 378 0 2 0  
## 379 0 2 0  
## 380 0 2 0  
## 381 0 2 0  
## 382 0 2 0  
## 383 0 2 0  
## 384 0 2 0  
## 385 1 4 0  
## 386 0 4 0  
## 387 0 4 0  
## 388 0 2 0  
## 389 0 2 0  
## 390 0 2 0  
## 391 0 2 0  
## 392 0 2 0  
## 393 1 4 1  
## 394 0 4 0  
## 395 0 4 0  
## 396 0 2 0  
## 397 0 2 0  
## 398 0 4 0  
## 399 0 4 0  
## 400 0 4 0  
##   
## Slot "algorithm\_summary":  
## SVM\_PRECISION SVM\_RECALL SVM\_FSCORE SLDA\_PRECISION SLDA\_RECALL  
## 2 0.98 0.94 0.96 0.98 0.96  
## 4 0.89 0.96 0.92 0.92 0.96  
## SLDA\_FSCORE LOGITBOOST\_PRECISION LOGITBOOST\_RECALL LOGITBOOST\_FSCORE  
## 2 0.97 0.96 0.95 0.95  
## 4 0.94 0.91 0.92 0.91  
## BAGGING\_PRECISION BAGGING\_RECALL BAGGING\_FSCORE FORESTS\_PRECISION  
## 2 0.96 0.96 0.96 0.97  
## 4 0.93 0.93 0.93 0.92  
## FORESTS\_RECALL FORESTS\_FSCORE GLMNET\_PRECISION GLMNET\_RECALL  
## 2 0.96 0.96 0.96 0.95  
## 4 0.94 0.93 0.91 0.92  
## GLMNET\_FSCORE TREE\_PRECISION TREE\_RECALL TREE\_FSCORE  
## 2 0.95 0.93 0.95 0.94  
## 4 0.91 0.91 0.87 0.89  
## MAXENTROPY\_PRECISION MAXENTROPY\_RECALL MAXENTROPY\_FSCORE  
## 2 0.96 0.95 0.95  
## 4 0.90 0.93 0.91  
##   
## Slot "ensemble\_summary":  
## n-ENSEMBLE COVERAGE n-ENSEMBLE RECALL  
## n >= 1 1.00 0.95  
## n >= 2 1.00 0.95  
## n >= 3 1.00 0.95  
## n >= 4 1.00 0.95  
## n >= 5 1.00 0.95  
## n >= 6 0.98 0.96  
## n >= 7 0.95 0.97  
## n >= 8 0.92 0.97

analytics@ensemble\_summary

## n-ENSEMBLE COVERAGE n-ENSEMBLE RECALL  
## n >= 1 1.00 0.95  
## n >= 2 1.00 0.95  
## n >= 3 1.00 0.95  
## n >= 4 1.00 0.95  
## n >= 5 1.00 0.95  
## n >= 6 0.98 0.96  
## n >= 7 0.95 0.97  
## n >= 8 0.92 0.97

# Assingment

Create regular expressions for the patterns below:

* Match any of the following punctuation characters in the ASCII table: !"#$%&'()+,
* Create one regular expression to match all common misspellings of calendar (see <https://en.wikipedia.org/wiki/Wikipedia:Lists_of_common_misspellings/C>)
* Create one regular expression to match any character except line breaks.
* You need to validate a ZIP code (U.S. postal code), allowing both the five-digit and nine-digit (called ZIP+4) formats. The regex should match 02115 and 02115-5515, but not 2115, 2115-5515, 21155515,021155515, etc..
* You need to validate a legit any password for your website. Passwords have the following complexity requirements: Length between 8 and 32 characters, ASCII visible and space characters only, One or more uppercase letters, One or more lowercase letters, One or more special characters (ASCII punctuation)
* Load the file M08\_tweets.csv (it is online at '<http://nikbearbrown.com/YouTube/MachineLearning/M08/M08_tweets.csv>')
* Do the following:
  + Extract a list of the top 9 users (e.g. @NikBearBrown)
  + Extract a list of the top 9 hashtags (e.g. #Bear)
  + Find the top 5 most positve tweets
  + Find the top 5 most negative tweets
  + Create a world cloud of 100 related tweets
  + Which tweets could be classified as game development?

Write up your report as an .Rmd file.

# Resources

* [Introduction to the tm Package Text Mining in R - CRAN](https://cran.r-project.org/web/packages/tm/vignettes/tm.pdf)
* [RDataMining.com: R and Data Mining](http://www.rdatamining.com/examples/text-mining)
* [Text Mining in R Tutorial: Term Frequency & Word Clouds](https://deltadna.com/blog/text-mining-in-r-for-term-frequency/)
* [Basic Text Mining in R](https://rstudio-pubs-static.s3.amazonaws.com/31867_8236987cf0a8444e962ccd2aec46d9c3.html)
* [Text Mining in R Tutorial: Term Frequency & Word Clouds ..](https://www.youtube.com/watch?v=lRTerj8fdY0)