Naive Bayes

Nik Bear Brown

In this lesson we'll learn the how to implement Naive Bayes in R.

# Additional packages needed

To run the code you may need additional packages.

* If necessary install the followings packages.

install.packages("tm");  
install.packages("wordcloud");  
install.packages("e1071");  
install.packages("gmodels");

library(tm)

## Loading required package: NLP

library(wordcloud)

## Loading required package: RColorBrewer

library(e1071)  
library(gmodels)

# Data

We will be using the SMS Spam Collection v.1, which is a public set of SMS labeled messages that have been collected for mobile phone spam research. It has one collection composed by 5,574 English, real and non-enconded messages, tagged according being legitimate (ham) or spam.

* from <http://www.dt.fee.unicamp.br/~tiago/smsspamcollection/>

data\_url <- 'http://nikbearbrown.com/YouTube/MachineLearning/M10/sms\_spam.csv'  
# SMS  
sms\_data <- read.csv(url(data\_url), stringsAsFactors = FALSE)  
str(sms\_data)

## 'data.frame': 5559 obs. of 2 variables:  
## $ type: chr "ham" "ham" "ham" "spam" ...  
## $ text: chr "Hope you are having a good week. Just checking in" "K..give back my thanks." "Am also doing in cbe only. But have to pay." "complimentary 4 STAR Ibiza Holiday or £10,000 cash needs your URGENT collection. 09066364349 NOW from Landline not to lose out!"| \_\_truncated\_\_ ...

# convert spam/ham to factor.  
sms\_data$type <- factor(sms\_data$type)  
str(sms\_data$type)

## Factor w/ 2 levels "ham","spam": 1 1 1 2 2 1 1 1 2 1 ...

table(sms\_data$type)

##   
## ham spam   
## 4812 747

# Naive Bayes

[Naive Bayes classifiers](https://en.wikipedia.org/wiki/Naive_Bayes_classifier) are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features. This means that under the above independence assumptions, the conditional distribution over the class variable C is:

where the evidence is a scaling factor dependent only on , that is, a constant if the values of the feature variables are known.

# Naive Bayes in R

# build a corpus using the text mining (tm) package   
sms\_corpus <- Corpus(VectorSource(sms\_data$text))  
print(sms\_corpus)

## <<VCorpus>>  
## Metadata: corpus specific: 0, document level (indexed): 0  
## Content: documents: 5559

inspect(sms\_corpus[1:3])

## <<VCorpus>>  
## Metadata: corpus specific: 0, document level (indexed): 0  
## Content: documents: 3  
##   
## [[1]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 49  
##   
## [[2]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 23  
##   
## [[3]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 43

# clean up the corpus using tm\_map()  
CleanCorpus <- tm\_map(sms\_corpus, content\_transformer(tolower))  
CleanCorpus <- tm\_map(CleanCorpus, removeNumbers )  
CleanCorpus <- tm\_map(CleanCorpus, removeWords, stopwords() )  
CleanCorpus <- tm\_map(CleanCorpus, removePunctuation )  
CleanCorpus <- tm\_map(CleanCorpus, stripWhitespace )  
inspect(sms\_corpus[1:3])

## <<VCorpus>>  
## Metadata: corpus specific: 0, document level (indexed): 0  
## Content: documents: 3  
##   
## [[1]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 49  
##   
## [[2]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 23  
##   
## [[3]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 43

inspect(CleanCorpus[1:3])

## <<VCorpus>>  
## Metadata: corpus specific: 0, document level (indexed): 0  
## Content: documents: 3  
##   
## [[1]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 29  
##   
## [[2]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 17  
##   
## [[3]]  
## <<PlainTextDocument>>  
## Metadata: 7  
## Content: chars: 13

# document-term sparse matrix  
sms\_dtm <- DocumentTermMatrix(CleanCorpus)  
sms\_dtm

## <<DocumentTermMatrix (documents: 5559, terms: 7868)>>  
## Non-/sparse entries: 42595/43695617  
## Sparsity : 100%  
## Maximal term length: 40  
## Weighting : term frequency (tf)

# training and test datasets  
sms\_data\_train <- sms\_data[1:4169, ]  
sms\_data\_test <- sms\_data[4170:5559, ]  
sms\_dtm\_train <- sms\_dtm[1:4169, ]  
sms\_dtm\_test <- sms\_dtm[4170:5559, ]  
sms\_corpus\_train <- CleanCorpus[1:4169]  
sms\_corpus\_test <- CleanCorpus[4170:5559]  
# check that the proportion   
prop.table(table(sms\_data\_train$type))

##   
## ham spam   
## 0.8647158 0.1352842

prop.table(table(sms\_data\_test$type))

##   
## ham spam   
## 0.8683453 0.1316547

# word cloud visualization  
wordcloud(sms\_corpus\_train, min.freq = 33, random.order = FALSE)

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): coming could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): someone could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): waiting could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): holiday could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): trying could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): guaranteed could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): thought could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): done could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): ready could not be fit on page. It will not be plotted.

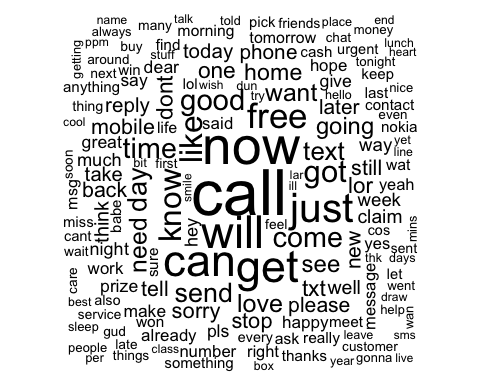
## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): haha could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): meeting could not be fit on page. It will not be plotted.

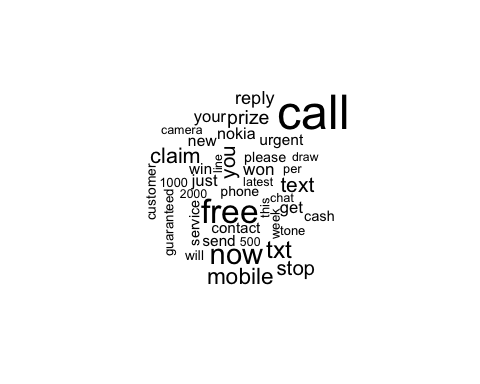
## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): tone could not be fit on page. It will not be plotted.

## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): camera could not be fit on page. It will not be plotted.

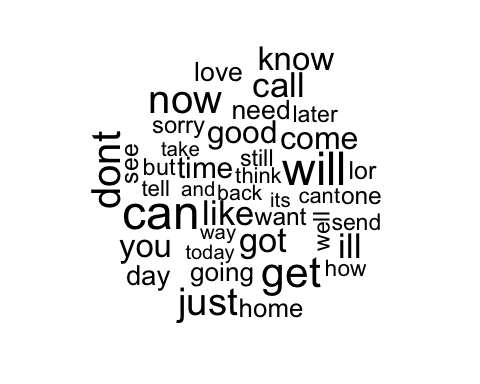
## Warning in wordcloud(sms\_corpus\_train, min.freq = 33, random.order =  
## FALSE): may could not be fit on page. It will not be plotted.



# subset the training data into spam and ham groups  
spam <- subset(sms\_data\_train, type == "spam")  
ham <- subset(sms\_data\_train, type == "ham")  
wordcloud(spam$text, max.words = 40, scale = c(3, 0.5))



wordcloud(ham$text, max.words = 40, scale = c(3, 0.5))



# indicator features for frequent words  
findFreqTerms(sms\_dtm\_train, 5)

## [1] "abiola" "able" "abt"   
## [4] "accept" "access" "account"   
## [7] "across" "activate" "actually"   
## [10] "add" "added" "address"   
## [13] "admirer" "advance" "aft"   
## [16] "afternoon" "age" "ago"   
## [19] "aha" "ahead" "aight"   
## [22] "aint" "air" "aiyo"   
## [25] "alex" "almost" "alone"   
## [28] "already" "alright" "also"   
## [31] "always" "angry" "announcement"   
## [34] "another" "ans" "answer"   
## [37] "anymore" "anyone" "anything"   
## [40] "anytime" "anyway" "apartment"   
## [43] "apply" "appreciate" "arcade"   
## [46] "ard" "area" "argument"   
## [49] "around" "arrive" "asap"   
## [52] "ask" "askd" "asked"   
## [55] "asking" "attempt" "auction"   
## [58] "available" "ave" "await"   
## [61] "awaiting" "awake" "award"   
## [64] "awarded" "away" "awesome"   
## [67] "babe" "babes" "baby"   
## [70] "back" "bad" "bag"   
## [73] "bank" "barely" "bathe"   
## [76] "battery" "bcoz" "bday"   
## [79] "beautiful" "become" "bed"   
## [82] "bedroom" "believe" "best"   
## [85] "better" "bid" "big"   
## [88] "bill" "birthday" "bit"   
## [91] "black" "blue" "bluetooth"   
## [94] "bold" "bonus" "boo"   
## [97] "book" "booked" "boost"   
## [100] "bored" "boss" "bother"   
## [103] "bout" "box" "boy"   
## [106] "boys" "boytoy" "break"   
## [109] "bring" "brings" "brother"   
## [112] "bslvyl" "btnationalrate" "bucks"   
## [115] "bus" "busy" "buy"   
## [118] "buying" "cabin" "call"   
## [121] "called" "caller" "callertune"   
## [124] "calling" "calls" "camcorder"   
## [127] "came" "camera" "campus"   
## [130] "can" "cancel" "cancer"   
## [133] "cant" "car" "card"   
## [136] "care" "carlos" "case"   
## [139] "cash" "cashbalance" "catch"   
## [142] "cause" "cell" "centre"   
## [145] "chance" "change" "charge"   
## [148] "charged" "charges" "chat"   
## [151] "cheap" "cheaper" "check"   
## [154] "checked" "checking" "cheers"   
## [157] "chennai" "chikku" "childish"   
## [160] "children" "choose" "christmas"   
## [163] "claim" "class" "clean"   
## [166] "close" "club" "code"   
## [169] "coffee" "cold" "colleagues"   
## [172] "collect" "collection" "college"   
## [175] "colour" "come" "comes"   
## [178] "comin" "coming" "comp"   
## [181] "company" "competition" "completely"   
## [184] "complimentary" "computer" "confirm"   
## [187] "congrats" "congratulations" "contact"   
## [190] "content" "contract" "cool"   
## [193] "copy" "correct" "cos"   
## [196] "cost" "costa" "costpm"   
## [199] "couple" "course" "cover"   
## [202] "coz" "crave" "crazy"   
## [205] "created" "credit" "credits"   
## [208] "cry" "cum" "cup"   
## [211] "currently" "custcare" "customer"   
## [214] "cut" "cute" "cuz"   
## [217] "dad" "daddy" "darlin"   
## [220] "darren" "dat" "date"   
## [223] "dating" "day" "days"   
## [226] "dead" "deal" "dear"   
## [229] "decide" "decided" "decimal"   
## [232] "deep" "definitely" "del"   
## [235] "delivery" "den" "depends"   
## [238] "details" "didnt" "die"   
## [241] "died" "different" "difficult"   
## [244] "digital" "din" "dinner"   
## [247] "direct" "dis" "discount"   
## [250] "disturb" "dnt" "doctor"   
## [253] "doesnt" "dog" "dogging"   
## [256] "doin" "don" "done"   
## [259] "dont" "door" "double"   
## [262] "download" "draw" "dream"   
## [265] "dreams" "drink" "drive"   
## [268] "driving" "drop" "dropped"   
## [271] "drug" "drugs" "dude"   
## [274] "due" "dun" "dunno"   
## [277] "dvd" "earlier" "early"   
## [280] "earth" "easy" "eat"   
## [283] "eatin" "eating" "either"   
## [286] "else" "email" "end"   
## [289] "ending" "ends" "energy"   
## [292] "england" "enjoy" "enough"   
## [295] "enter" "entered" "entitled"   
## [298] "entry" "envelope" "etc"   
## [301] "euro" "eve" "even"   
## [304] "evening" "ever" "every"   
## [307] "everyone" "everything" "exactly"   
## [310] "exam" "exams" "excellent"   
## [313] "exciting" "excuse" "expecting"   
## [316] "experience" "expires" "extra"   
## [319] "eyes" "face" "facebook"   
## [322] "fact" "family" "fancy"   
## [325] "fantastic" "far" "fast"   
## [328] "fat" "father" "fault"   
## [331] "feb" "feel" "feeling"   
## [334] "feels" "felt" "figure"   
## [337] "film" "final" "finally"   
## [340] "find" "fine" "fingers"   
## [343] "finish" "finished" "first"   
## [346] "fixed" "flag" "flat"   
## [349] "flower" "following" "fone"   
## [352] "food" "forever" "forget"   
## [355] "forgot" "forward" "forwarded"   
## [358] "found" "free" "freemsg"   
## [361] "freephone" "frens" "fri"   
## [364] "friday" "friend" "friends"   
## [367] "friendship" "frm" "frnd"   
## [370] "frnds" "full" "fullonsmscom"   
## [373] "fun" "funny" "future"   
## [376] "gal" "game" "games"   
## [379] "gap" "gas" "gave"   
## [382] "gay" "get" "gets"   
## [385] "gettin" "getting" "gift"   
## [388] "girl" "girlfrnd" "girls"   
## [391] "give" "glad" "god"   
## [394] "goes" "goin" "going"   
## [397] "gone" "gonna" "good"   
## [400] "goodmorning" "goodnight" "got"   
## [403] "goto" "gotta" "great"   
## [406] "grins" "guaranteed" "gud"   
## [409] "guess" "guy" "guys"   
## [412] "gym" "haf" "haha"   
## [415] "hai" "hair" "half"   
## [418] "hand" "handset" "happen"   
## [421] "happened" "happening" "happens"   
## [424] "happiness" "happy" "hard"   
## [427] "hate" "hav" "havent"   
## [430] "head" "hear" "heard"   
## [433] "heart" "heavy" "hee"   
## [436] "hell" "hello" "help"   
## [439] "hey" "hgsuitelands" "hit"   
## [442] "hiya" "hmm" "hmmm"   
## [445] "hmv" "hold" "holder"   
## [448] "holding" "holiday" "home"   
## [451] "hop" "hope" "hoping"   
## [454] "horny" "hospital" "hot"   
## [457] "hour" "hours" "house"   
## [460] "however" "hows" "howz"   
## [463] "hrs" "httpwwwurawinnercom" "huh"   
## [466] "hungry" "hurry" "hurt"   
## [469] "hurts" "ice" "identifier"   
## [472] "ill" "immediately" "important"   
## [475] "inc" "india" "info"   
## [478] "information" "inside" "instead"   
## [481] "interested" "invited" "ipod"   
## [484] "ish" "ive" "izzit"   
## [487] "january" "jay" "job"   
## [490] "john" "join" "joined"   
## [493] "joke" "joking" "jst"   
## [496] "jus" "just" "juz"   
## [499] "kate" "keep" "keeping"   
## [502] "kept" "kick" "kids"   
## [505] "kind" "kinda" "king"   
## [508] "kiss" "knew" "know"   
## [511] "knows" "knw" "lady"   
## [514] "land" "landline" "laptop"   
## [517] "lar" "last" "late"   
## [520] "later" "latest" "laugh"   
## [523] "lazy" "ldn" "learn"   
## [526] "least" "leave" "leaves"   
## [529] "leaving" "lect" "left"   
## [532] "leh" "lei" "less"   
## [535] "lesson" "lessons" "let"   
## [538] "lets" "liao" "library"   
## [541] "life" "lift" "light"   
## [544] "like" "liked" "line"   
## [547] "link" "listen" "little"   
## [550] "live" "lmao" "loads"   
## [553] "loan" "local" "log"   
## [556] "lol" "london" "long"   
## [559] "longer" "look" "lookin"   
## [562] "looking" "looks" "lor"   
## [565] "lose" "lost" "lot"   
## [568] "lots" "lovable" "love"   
## [571] "loved" "lovely" "lover"   
## [574] "loves" "loving" "loyalty"   
## [577] "ltd" "luck" "lucky"   
## [580] "lunch" "luv" "mad"   
## [583] "made" "mah" "mail"   
## [586] "make" "makes" "making"   
## [589] "malaria" "man" "many"   
## [592] "march" "mark" "married"   
## [595] "marry" "match" "matches"   
## [598] "mate" "mates" "maximize"   
## [601] "maxmins" "may" "mayb"   
## [604] "maybe" "mean" "meaning"   
## [607] "means" "meant" "medical"   
## [610] "meds" "meet" "meeting"   
## [613] "meh" "member" "merry"   
## [616] "message" "messages" "met"   
## [619] "midnight" "might" "min"   
## [622] "mind" "mine" "mins"   
## [625] "minute" "minutes" "miracle"   
## [628] "miss" "missed" "missing"   
## [631] "mistake" "moan" "mob"   
## [634] "mobile" "mobiles" "mobileupd"   
## [637] "mode" "mom" "moment"   
## [640] "moms" "mon" "monday"   
## [643] "money" "month" "months"   
## [646] "morning" "mother" "motorola"   
## [649] "move" "movie" "movies"   
## [652] "mrng" "mrt" "mrw"   
## [655] "msg" "msgs" "mths"   
## [658] "much" "mum" "music"   
## [661] "must" "muz" "nah"   
## [664] "naked" "name" "national"   
## [667] "nature" "naughty" "near"   
## [670] "need" "needs" "net"   
## [673] "network" "neva" "never"   
## [676] "new" "news" "next"   
## [679] "nice" "nigeria" "night"   
## [682] "nite" "nobody" "noe"   
## [685] "nokia" "noon" "nope"   
## [688] "normal" "normptone" "nothing"   
## [691] "notice" "now" "num"   
## [694] "number" "numbers" "nyt"   
## [697] "obviously" "offer" "offers"   
## [700] "office" "official" "okay"   
## [703] "okie" "old" "omg"   
## [706] "one" "ones" "online"   
## [709] "onto" "oops" "open"   
## [712] "operator" "opinion" "opt"   
## [715] "optout" "orange" "orchard"   
## [718] "order" "oredi" "oso"   
## [721] "others" "otherwise" "outside"   
## [724] "paid" "pain" "paper"   
## [727] "parents" "park" "part"   
## [730] "partner" "party" "pass"   
## [733] "password" "past" "pay"   
## [736] "paying" "people" "per"   
## [739] "person" "persons" "pete"   
## [742] "phone" "phones" "photo"   
## [745] "photos" "pic" "pick"   
## [748] "picked" "picking" "pics"   
## [751] "pictures" "pin" "pix"   
## [754] "pizza" "place" "plan"   
## [757] "planned" "planning" "plans"   
## [760] "play" "player" "players"   
## [763] "please" "pleasure" "plenty"   
## [766] "pls" "plus" "plz"   
## [769] "pmin" "pmsg" "pobox"   
## [772] "point" "points" "police"   
## [775] "poly" "polys" "poor"   
## [778] "possible" "post" "posted"   
## [781] "pound" "pounds" "ppm"   
## [784] "pray" "press" "pretty"   
## [787] "price" "princess" "private"   
## [790] "prize" "prob" "probably"   
## [793] "problem" "project" "promise"   
## [796] "pub" "put" "putting"   
## [799] "quality" "question" "questions"   
## [802] "quick" "quite" "quiz"   
## [805] "rain" "raining" "rate"   
## [808] "rates" "rather" "rcvd"   
## [811] "reach" "reached" "reaching"   
## [814] "read" "reading" "ready"   
## [817] "real" "really" "realy"   
## [820] "reason" "receipt" "receive"   
## [823] "recently" "records" "reference"   
## [826] "regards" "registered" "relation"   
## [829] "relax" "remember" "remind"   
## [832] "remove" "rent" "rental"   
## [835] "replied" "reply" "replying"   
## [838] "representative" "request" "rest"   
## [841] "review" "reward" "right"   
## [844] "ring" "ringtone" "rite"   
## [847] "road" "rock" "role"   
## [850] "room" "rose" "round"   
## [853] "rowwjhl" "rply" "rreveal"   
## [856] "run" "sad" "sae"   
## [859] "safe" "said" "sale"   
## [862] "sat" "saturday" "savamob"   
## [865] "save" "saw" "say"   
## [868] "saying" "says" "sch"   
## [871] "school" "screaming" "sea"   
## [874] "search" "sec" "second"   
## [877] "secret" "see" "seeing"   
## [880] "seems" "seen" "selected"   
## [883] "self" "sell" "semester"   
## [886] "send" "sending" "sense"   
## [889] "sent" "serious" "seriously"   
## [892] "service" "services" "set"   
## [895] "sex" "sexy" "shall"   
## [898] "share" "shd" "shop"   
## [901] "shopping" "short" "show"   
## [904] "shower" "shows" "sick"   
## [907] "side" "sight" "sign"   
## [910] "silent" "simple" "since"   
## [913] "single" "sipix" "sir"   
## [916] "sis" "sister" "sit"   
## [919] "sitting" "situation" "skxh"   
## [922] "slave" "sleep" "sleeping"   
## [925] "slept" "slow" "slowly"   
## [928] "small" "smile" "smiling"   
## [931] "smoke" "sms" "smth"   
## [934] "snow" "sofa" "sol"   
## [937] "somebody" "someone" "something"   
## [940] "sometimes" "somewhere" "song"   
## [943] "sony" "sonyericsson" "soon"   
## [946] "sorry" "sort" "sound"   
## [949] "sounds" "south" "space"   
## [952] "speak" "special" "specialcall"   
## [955] "specially" "spend" "spent"   
## [958] "spoke" "spree" "stand"   
## [961] "start" "started" "starting"   
## [964] "starts" "statement" "station"   
## [967] "stay" "staying" "std"   
## [970] "still" "stockport" "stop"   
## [973] "store" "story" "street"   
## [976] "study" "studying" "stuff"   
## [979] "stupid" "sub" "sucks"   
## [982] "summer" "sun" "sunday"   
## [985] "sunshine" "sup" "support"   
## [988] "supposed" "sure" "surely"   
## [991] "surprise" "sweet" "swing"   
## [994] "system" "take" "takes"   
## [997] "taking" "talk" "talking"   
## [1000] "tampa" "tariffs" "tcs"   
## [1003] "tea" "teach" "tear"   
## [1006] "tel" "tell" "telling"   
## [1009] "tells" "ten" "tenerife"   
## [1012] "terms" "test" "text"   
## [1015] "texting" "texts" "thank"   
## [1018] "thanks" "thanx" "thats"   
## [1021] "thing" "things" "think"   
## [1024] "thinkin" "thinking" "thinks"   
## [1027] "thk" "tho" "though"   
## [1030] "thought" "thru" "tht"   
## [1033] "thurs" "tick" "ticket"   
## [1036] "tickets" "til" "till"   
## [1039] "time" "times" "timing"   
## [1042] "tired" "tmr" "toclaim"   
## [1045] "today" "todays" "together"   
## [1048] "told" "tomo" "tomorrow"   
## [1051] "tone" "tones" "tonight"   
## [1054] "tonite" "took" "top"   
## [1057] "torch" "tot" "touch"   
## [1060] "tough" "tour" "towards"   
## [1063] "town" "track" "train"   
## [1066] "training" "transaction" "treat"   
## [1069] "tried" "trip" "trouble"   
## [1072] "true" "trust" "truth"   
## [1075] "try" "trying" "tscs"   
## [1078] "ttyl" "tuesday" "twice"   
## [1081] "two" "txt" "txting"   
## [1084] "txts" "type" "ufind"   
## [1087] "ugh" "uks" "ull"   
## [1090] "uncle" "understand" "unless"   
## [1093] "unlimited" "unredeemed" "unsub"   
## [1096] "unsubscribe" "update" "ure"   
## [1099] "urgent" "urself" "use"   
## [1102] "used" "user" "usf"   
## [1105] "using" "usual" "uve"   
## [1108] "valentine" "valentines" "valid"   
## [1111] "valued" "via" "video"   
## [1114] "vikky" "visit" "vodafone"   
## [1117] "voice" "vomit" "voucher"   
## [1120] "vouchers" "wait" "waiting"   
## [1123] "wake" "waking" "walk"   
## [1126] "walking" "wan" "wana"   
## [1129] "wanna" "want" "wanted"   
## [1132] "wants" "wap" "warm"   
## [1135] "waste" "wat" "watch"   
## [1138] "watching" "water" "wats"   
## [1141] "way" "weather" "wed"   
## [1144] "wednesday" "weed" "week"   
## [1147] "weekend" "weekends" "weekly"   
## [1150] "weeks" "welcome" "well"   
## [1153] "wen" "went" "whatever"   
## [1156] "whats" "whenever" "whole"   
## [1159] "wid" "wif" "wife"   
## [1162] "wil" "will" "willing"   
## [1165] "win" "wine" "winner"   
## [1168] "wins" "wish" "wishing"   
## [1171] "wit" "within" "without"   
## [1174] "wiv" "wkly" "wks"   
## [1177] "wnt" "woke" "won"   
## [1180] "wonder" "wonderful" "wont"   
## [1183] "word" "words" "work"   
## [1186] "workin" "working" "works"   
## [1189] "world" "worried" "worries"   
## [1192] "worry" "worse" "worth"   
## [1195] "wot" "wow" "write"   
## [1198] "wrong" "wwq" "wwwgetzedcouk"   
## [1201] "xmas" "xxx" "yahoo"   
## [1204] "yar" "yeah" "year"   
## [1207] "years" "yep" "yes"   
## [1210] "yesterday" "yet" "yoga"   
## [1213] "yup"

sms\_dict <- findFreqTerms(sms\_dtm\_train, 5)  
sms\_train <- DocumentTermMatrix(sms\_corpus\_train, list(dictionary = sms\_dict))  
sms\_test <- DocumentTermMatrix(sms\_corpus\_test, list(dictionary = sms\_dict))  
# convert counts to a factor  
convert\_counts <- function(x) {  
 x <- ifelse(x > 0, 1, 0)  
 x <- factor(x, levels = c(0, 1), labels = c("No", "Yes"))  
}  
# apply() convert\_counts() to columns of train/test data  
sms\_train <- apply(sms\_train, MARGIN = 2, convert\_counts)  
sms\_test <- apply(sms\_test, MARGIN = 2, convert\_counts)  
#----------------------------------------------------Training a model on the data --------------------  
sms\_classifier <- naiveBayes(sms\_train, sms\_data\_train$type)  
sms\_classifier

##   
## Naive Bayes Classifier for Discrete Predictors  
##   
## Call:  
## naiveBayes.default(x = sms\_train, y = sms\_data\_train$type)  
##   
## A-priori probabilities:  
## sms\_data\_train$type  
## ham spam   
## 0.8647158 0.1352842   
##   
## Conditional probabilities:  
## abiola  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## able  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 1.000000000 0.000000000  
##   
## abt  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## accept  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## access  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9946808511 0.0053191489  
##   
## account  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.97163121 0.02836879  
##   
## across  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## activate  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9929078014 0.0070921986  
##   
## actually  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 1.000000000 0.000000000  
##   
## add  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## added  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## address  
## sms\_data\_train$type No Yes  
## ham 0.996671290 0.003328710  
## spam 0.992907801 0.007092199  
##   
## admirer  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98404255 0.01595745  
##   
## advance  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## aft  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## afternoon  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## age  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9804964539 0.0195035461  
##   
## ago  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## aha  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## ahead  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## aight  
## sms\_data\_train$type No Yes  
## ham 0.99223301 0.00776699  
## spam 1.00000000 0.00000000  
##   
## aint  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## air  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## aiyo  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## alex  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## almost  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## alone  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.996453901 0.003546099  
##   
## already  
## sms\_data\_train$type No Yes  
## ham 0.98085992 0.01914008  
## spam 0.99822695 0.00177305  
##   
## alright  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 1.000000000 0.000000000  
##   
## also  
## sms\_data\_train$type No Yes  
## ham 0.986962552 0.013037448  
## spam 0.994680851 0.005319149  
##   
## always  
## sms\_data\_train$type No Yes  
## ham 0.98862691 0.01137309  
## spam 1.00000000 0.00000000  
##   
## angry  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## announcement  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## another  
## sms\_data\_train$type No Yes  
## ham 0.992787795 0.007212205  
## spam 0.998226950 0.001773050  
##   
## ans  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.992907801 0.007092199  
##   
## answer  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.992907801 0.007092199  
##   
## anymore  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## anyone  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.998226950 0.001773050  
##   
## anything  
## sms\_data\_train$type No Yes  
## ham 0.9850208 0.0149792  
## spam 1.0000000 0.0000000  
##   
## anytime  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.982269504 0.017730496  
##   
## anyway  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## apartment  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## apply  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9645390071 0.0354609929  
##   
## appreciate  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## arcade  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.994680851 0.005319149  
##   
## ard  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## area  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.991134752 0.008865248  
##   
## argument  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## around  
## sms\_data\_train$type No Yes  
## ham 0.98890430 0.01109570  
## spam 0.99822695 0.00177305  
##   
## arrive  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## asap  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9893617021 0.0106382979  
##   
## ask  
## sms\_data\_train$type No Yes  
## ham 0.98363384 0.01636616  
## spam 1.00000000 0.00000000  
##   
## askd  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## asked  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 0.991134752 0.008865248  
##   
## asking  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## attempt  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.96631206 0.03368794  
##   
## auction  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98404255 0.01595745  
##   
## available  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.998226950 0.001773050  
##   
## ave  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## await  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9716312057 0.0283687943  
##   
## awaiting  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## awake  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## award  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97340426 0.02659574  
##   
## awarded  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.95035461 0.04964539  
##   
## away  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.996453901 0.003546099  
##   
## awesome  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## babe  
## sms\_data\_train$type No Yes  
## ham 0.98696255 0.01303745  
## spam 0.98936170 0.01063830  
##   
## babes  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9964539007 0.0035460993  
##   
## baby  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 0.996453901 0.003546099  
##   
## back  
## sms\_data\_train$type No Yes  
## ham 0.97447989 0.02552011  
## spam 0.97340426 0.02659574  
##   
## bad  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 1.000000000 0.000000000  
##   
## bag  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## bank  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.998226950 0.001773050  
##   
## barely  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## bathe  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## battery  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## bcoz  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## bday  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## beautiful  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## become  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.998226950 0.001773050  
##   
## bed  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## bedroom  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## believe  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.99822695 0.00177305  
##   
## best  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 0.989361702 0.010638298  
##   
## better  
## sms\_data\_train$type No Yes  
## ham 0.99223301 0.00776699  
## spam 1.00000000 0.00000000  
##   
## bid  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.994680851 0.005319149  
##   
## big  
## sms\_data\_train$type No Yes  
## ham 0.99223301 0.00776699  
## spam 0.99822695 0.00177305  
##   
## bill  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## birthday  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 0.998226950 0.001773050  
##   
## bit  
## sms\_data\_train$type No Yes  
## ham 0.98945908 0.01054092  
## spam 1.00000000 0.00000000  
##   
## black  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## blue  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## bluetooth  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## bold  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## bonus  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.96808511 0.03191489  
##   
## boo  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## book  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.994680851 0.005319149  
##   
## booked  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## boost  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 1.0000000000 0.0000000000  
##   
## bored  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 0.998226950 0.001773050  
##   
## boss  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## bother  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## bout  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.998226950 0.001773050  
##   
## box  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9414893617 0.0585106383  
##   
## boy  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## boys  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.996453901 0.003546099  
##   
## boytoy  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## break  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.996453901 0.003546099  
##   
## bring  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 1.000000000 0.000000000  
##   
## brings  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## brother  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.99822695 0.00177305  
##   
## bslvyl  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## btnationalrate  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## bucks  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## bus  
## sms\_data\_train$type No Yes  
## ham 0.992787795 0.007212205  
## spam 1.000000000 0.000000000  
##   
## busy  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 0.998226950 0.001773050  
##   
## buy  
## sms\_data\_train$type No Yes  
## ham 0.988349515 0.011650485  
## spam 0.994680851 0.005319149  
##   
## buying  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## cabin  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## call  
## sms\_data\_train$type No Yes  
## ham 0.95228849 0.04771151  
## spam 0.55496454 0.44503546  
##   
## called  
## sms\_data\_train$type No Yes  
## ham 0.993619972 0.006380028  
## spam 1.000000000 0.000000000  
##   
## caller  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9787234043 0.0212765957  
##   
## callertune  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 1.0000000000 0.0000000000  
##   
## calling  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.996453901 0.003546099  
##   
## calls  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.985815603 0.014184397  
##   
## camcorder  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97695035 0.02304965  
##   
## came  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.998226950 0.001773050  
##   
## camera  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9521276596 0.0478723404  
##   
## campus  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## can  
## sms\_data\_train$type No Yes  
## ham 0.92871012 0.07128988  
## spam 0.95744681 0.04255319  
##   
## cancel  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.992907801 0.007092199  
##   
## cancer  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## cant  
## sms\_data\_train$type No Yes  
## ham 0.986407767 0.013592233  
## spam 0.994680851 0.005319149  
##   
## car  
## sms\_data\_train$type No Yes  
## ham 0.99223301 0.00776699  
## spam 1.00000000 0.00000000  
##   
## card  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.994680851 0.005319149  
##   
## care  
## sms\_data\_train$type No Yes  
## ham 0.988626907 0.011373093  
## spam 0.991134752 0.008865248  
##   
## carlos  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## case  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## cash  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.91312057 0.08687943  
##   
## cashbalance  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## catch  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## cause  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## cell  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## centre  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## chance  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.96453901 0.03546099  
##   
## change  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## charge  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.98936170 0.01063830  
##   
## charged  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## charges  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.996453901 0.003546099  
##   
## chat  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 0.950354610 0.049645390  
##   
## cheap  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.992907801 0.007092199  
##   
## cheaper  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9964539007 0.0035460993  
##   
## check  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.994680851 0.005319149  
##   
## checked  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## checking  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## cheers  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## chennai  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## chikku  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## childish  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## children  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## choose  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.984042553 0.015957447  
##   
## christmas  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.99822695 0.00177305  
##   
## claim  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.8457447 0.1542553  
##   
## class  
## sms\_data\_train$type No Yes  
## ham 0.99112344 0.00887656  
## spam 1.00000000 0.00000000  
##   
## clean  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## close  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 0.994680851 0.005319149  
##   
## club  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9787234043 0.0212765957  
##   
## code  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9592198582 0.0407801418  
##   
## coffee  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## cold  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## colleagues  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## collect  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.96808511 0.03191489  
##   
## collection  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.96276596 0.03723404  
##   
## college  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## colour  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.971631206 0.028368794  
##   
## come  
## sms\_data\_train$type No Yes  
## ham 0.957004161 0.042995839  
## spam 0.996453901 0.003546099  
##   
## comes  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 0.994680851 0.005319149  
##   
## comin  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## coming  
## sms\_data\_train$type No Yes  
## ham 0.9889043 0.0110957  
## spam 1.0000000 0.0000000  
##   
## comp  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9840425532 0.0159574468  
##   
## company  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 0.996453901 0.003546099  
##   
## competition  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## completely  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.998226950 0.001773050  
##   
## complimentary  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98404255 0.01595745  
##   
## computer  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.992907801 0.007092199  
##   
## confirm  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.998226950 0.001773050  
##   
## congrats  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.991134752 0.008865248  
##   
## congratulations  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9840425532 0.0159574468  
##   
## contact  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.914893617 0.085106383  
##   
## content  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9822695 0.0177305  
##   
## contract  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9929078014 0.0070921986  
##   
## cool  
## sms\_data\_train$type No Yes  
## ham 0.99112344 0.00887656  
## spam 0.99822695 0.00177305  
##   
## copy  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## correct  
## sms\_data\_train$type No Yes  
## ham 0.997780860 0.002219140  
## spam 0.994680851 0.005319149  
##   
## cos  
## sms\_data\_train$type No Yes  
## ham 0.98557559 0.01442441  
## spam 1.00000000 0.00000000  
##   
## cost  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.97340426 0.02659574  
##   
## costa  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## costpm  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## couple  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.998226950 0.001773050  
##   
## course  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## cover  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## coz  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## crave  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## crazy  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.996453901 0.003546099  
##   
## created  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## credit  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.987588652 0.012411348  
##   
## credits  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## cry  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## cum  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.991134752 0.008865248  
##   
## cup  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9929078014 0.0070921986  
##   
## currently  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.984042553 0.015957447  
##   
## custcare  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## customer  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.930851064 0.069148936  
##   
## cut  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## cute  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## cuz  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## dad  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## daddy  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## darlin  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## darren  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## dat  
## sms\_data\_train$type No Yes  
## ham 0.991955617 0.008044383  
## spam 1.000000000 0.000000000  
##   
## date  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.992907801 0.007092199  
##   
## dating  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9822695 0.0177305  
##   
## day  
## sms\_data\_train$type No Yes  
## ham 0.96255201 0.03744799  
## spam 0.97517730 0.02482270  
##   
## days  
## sms\_data\_train$type No Yes  
## ham 0.993619972 0.006380028  
## spam 0.978723404 0.021276596  
##   
## dead  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## deal  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## dear  
## sms\_data\_train$type No Yes  
## ham 0.98363384 0.01636616  
## spam 0.97695035 0.02304965  
##   
## decide  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## decided  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## decimal  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## deep  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## definitely  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## del  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## delivery  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9769503546 0.0230496454  
##   
## den  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 1.000000000 0.000000000  
##   
## depends  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## details  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.98581560 0.01418440  
##   
## didnt  
## sms\_data\_train$type No Yes  
## ham 0.991678225 0.008321775  
## spam 1.000000000 0.000000000  
##   
## die  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## died  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## different  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## difficult  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## digital  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9875886525 0.0124113475  
##   
## din  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## dinner  
## sms\_data\_train$type No Yes  
## ham 0.99334258 0.00665742  
## spam 1.00000000 0.00000000  
##   
## direct  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9858156028 0.0141843972  
##   
## dis  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 0.996453901 0.003546099  
##   
## discount  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9858156028 0.0141843972  
##   
## disturb  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## dnt  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## doctor  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## doesnt  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## dog  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## dogging  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.992907801 0.007092199  
##   
## doin  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## don  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## done  
## sms\_data\_train$type No Yes  
## ham 0.991678225 0.008321775  
## spam 0.992907801 0.007092199  
##   
## dont  
## sms\_data\_train$type No Yes  
## ham 0.97170596 0.02829404  
## spam 0.98404255 0.01595745  
##   
## door  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## double  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9822695035 0.0177304965  
##   
## download  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## draw  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.948581560 0.051418440  
##   
## dream  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## dreams  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## drink  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## drive  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## driving  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 0.99822695 0.00177305  
##   
## drop  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## dropped  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## drug  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## drugs  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## dude  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## due  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## dun  
## sms\_data\_train$type No Yes  
## ham 0.990291262 0.009708738  
## spam 1.000000000 0.000000000  
##   
## dunno  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## dvd  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9893617021 0.0106382979  
##   
## earlier  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## early  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 1.000000000 0.000000000  
##   
## earth  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## easy  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.980496454 0.019503546  
##   
## eat  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 1.000000000 0.000000000  
##   
## eatin  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## eating  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## either  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.989361702 0.010638298  
##   
## else  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## email  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.99822695 0.00177305  
##   
## end  
## sms\_data\_train$type No Yes  
## ham 0.99334258 0.00665742  
## spam 0.98581560 0.01418440  
##   
## ending  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9893617021 0.0106382979  
##   
## ends  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.992907801 0.007092199  
##   
## energy  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## england  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9893617021 0.0106382979  
##   
## enjoy  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 0.98404255 0.01595745  
##   
## enough  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.996453901 0.003546099  
##   
## enter  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.989361702 0.010638298  
##   
## entered  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.989361702 0.010638298  
##   
## entitled  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## entry  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9751773 0.0248227  
##   
## envelope  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.996453901 0.003546099  
##   
## etc  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.991134752 0.008865248  
##   
## euro  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.992907801 0.007092199  
##   
## eve  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.992907801 0.007092199  
##   
## even  
## sms\_data\_train$type No Yes  
## ham 0.989736477 0.010263523  
## spam 0.992907801 0.007092199  
##   
## evening  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## ever  
## sms\_data\_train$type No Yes  
## ham 0.993342580 0.006657420  
## spam 0.996453901 0.003546099  
##   
## every  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 0.953900709 0.046099291  
##   
## everyone  
## sms\_data\_train$type No Yes  
## ham 0.996671290 0.003328710  
## spam 0.996453901 0.003546099  
##   
## everything  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 1.000000000 0.000000000  
##   
## exactly  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## exam  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## exams  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## excellent  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## exciting  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.994680851 0.005319149  
##   
## excuse  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## expecting  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## experience  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## expires  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9751773 0.0248227  
##   
## extra  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## eyes  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## face  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 1.000000000 0.000000000  
##   
## facebook  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## fact  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.998226950 0.001773050  
##   
## family  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 1.000000000 0.000000000  
##   
## fancy  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.985815603 0.014184397  
##   
## fantastic  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.989361702 0.010638298  
##   
## far  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## fast  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.998226950 0.001773050  
##   
## fat  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## father  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## fault  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## feb  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## feel  
## sms\_data\_train$type No Yes  
## ham 0.98807212 0.01192788  
## spam 1.00000000 0.00000000  
##   
## feeling  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## feels  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## felt  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## figure  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## film  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.998226950 0.001773050  
##   
## final  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.976950355 0.023049645  
##   
## finally  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.99822695 0.00177305  
##   
## find  
## sms\_data\_train$type No Yes  
## ham 0.98890430 0.01109570  
## spam 0.96453901 0.03546099  
##   
## fine  
## sms\_data\_train$type No Yes  
## ham 0.991955617 0.008044383  
## spam 1.000000000 0.000000000  
##   
## fingers  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## finish  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 1.000000000 0.000000000  
##   
## finished  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## first  
## sms\_data\_train$type No Yes  
## ham 0.989181692 0.010818308  
## spam 0.992907801 0.007092199  
##   
## fixed  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## flag  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## flat  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## flower  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## following  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.992907801 0.007092199  
##   
## fone  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.994680851 0.005319149  
##   
## food  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## forever  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## forget  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.996453901 0.003546099  
##   
## forgot  
## sms\_data\_train$type No Yes  
## ham 0.993619972 0.006380028  
## spam 1.000000000 0.000000000  
##   
## forward  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## forwarded  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.994680851 0.005319149  
##   
## found  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## free  
## sms\_data\_train$type No Yes  
## ham 0.98751734 0.01248266  
## spam 0.76950355 0.23049645  
##   
## freemsg  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## freephone  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## frens  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## fri  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.996453901 0.003546099  
##   
## friday  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## friend  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.987588652 0.012411348  
##   
## friends  
## sms\_data\_train$type No Yes  
## ham 0.988072122 0.011927878  
## spam 0.994680851 0.005319149  
##   
## friendship  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## frm  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## frnd  
## sms\_data\_train$type No Yes  
## ham 0.997780860 0.002219140  
## spam 0.992907801 0.007092199  
##   
## frnds  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## full  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## fullonsmscom  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## fun  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 0.985815603 0.014184397  
##   
## funny  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## future  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## gal  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## game  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.985815603 0.014184397  
##   
## games  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98404255 0.01595745  
##   
## gap  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## gas  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## gave  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## gay  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## get  
## sms\_data\_train$type No Yes  
## ham 0.93814147 0.06185853  
## spam 0.89893617 0.10106383  
##   
## gets  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## gettin  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## getting  
## sms\_data\_train$type No Yes  
## ham 0.990568655 0.009431345  
## spam 0.992907801 0.007092199  
##   
## gift  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.978723404 0.021276596  
##   
## girl  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 0.996453901 0.003546099  
##   
## girlfrnd  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 1.0000000000 0.0000000000  
##   
## girls  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.991134752 0.008865248  
##   
## give  
## sms\_data\_train$type No Yes  
## ham 0.980859917 0.019140083  
## spam 0.991134752 0.008865248  
##   
## glad  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## god  
## sms\_data\_train$type No Yes  
## ham 0.992787795 0.007212205  
## spam 1.000000000 0.000000000  
##   
## goes  
## sms\_data\_train$type No Yes  
## ham 0.993619972 0.006380028  
## spam 1.000000000 0.000000000  
##   
## goin  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## going  
## sms\_data\_train$type No Yes  
## ham 0.966435506 0.033564494  
## spam 0.992907801 0.007092199  
##   
## gone  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## gonna  
## sms\_data\_train$type No Yes  
## ham 0.98918169 0.01081831  
## spam 1.00000000 0.00000000  
##   
## good  
## sms\_data\_train$type No Yes  
## ham 0.95700416 0.04299584  
## spam 0.98581560 0.01418440  
##   
## goodmorning  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## goodnight  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## got  
## sms\_data\_train$type No Yes  
## ham 0.952011096 0.047988904  
## spam 0.991134752 0.008865248  
##   
## goto  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.994680851 0.005319149  
##   
## gotta  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## great  
## sms\_data\_train$type No Yes  
## ham 0.98252427 0.01747573  
## spam 0.98226950 0.01773050  
##   
## grins  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## guaranteed  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.93617021 0.06382979  
##   
## gud  
## sms\_data\_train$type No Yes  
## ham 0.98779473 0.01220527  
## spam 1.00000000 0.00000000  
##   
## guess  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.991134752 0.008865248  
##   
## guy  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 1.000000000 0.000000000  
##   
## guys  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 0.994680851 0.005319149  
##   
## gym  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## haf  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## haha  
## sms\_data\_train$type No Yes  
## ham 0.990846047 0.009153953  
## spam 1.000000000 0.000000000  
##   
## hai  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## hair  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## half  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.976950355 0.023049645  
##   
## hand  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## handset  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## happen  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 1.000000000 0.000000000  
##   
## happened  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## happening  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## happens  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## happiness  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## happy  
## sms\_data\_train$type No Yes  
## ham 0.98169209 0.01830791  
## spam 0.99822695 0.00177305  
##   
## hard  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.994680851 0.005319149  
##   
## hate  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## hav  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## havent  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## head  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## hear  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 0.996453901 0.003546099  
##   
## heard  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.994680851 0.005319149  
##   
## heart  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 0.994680851 0.005319149  
##   
## heavy  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## hee  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## hell  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## hello  
## sms\_data\_train$type No Yes  
## ham 0.992233010 0.007766990  
## spam 0.994680851 0.005319149  
##   
## help  
## sms\_data\_train$type No Yes  
## ham 0.99334258 0.00665742  
## spam 0.97340426 0.02659574  
##   
## hey  
## sms\_data\_train$type No Yes  
## ham 0.978085992 0.021914008  
## spam 0.996453901 0.003546099  
##   
## hgsuitelands  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## hit  
## sms\_data\_train$type No Yes  
## ham 0.997780860 0.002219140  
## spam 0.996453901 0.003546099  
##   
## hiya  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## hmm  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## hmmm  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## hmv  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.994680851 0.005319149  
##   
## hold  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## holder  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## holding  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## holiday  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.953900709 0.046099291  
##   
## home  
## sms\_data\_train$type No Yes  
## ham 0.96588072 0.03411928  
## spam 1.00000000 0.00000000  
##   
## hop  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 1.0000000000 0.0000000000  
##   
## hope  
## sms\_data\_train$type No Yes  
## ham 0.981137309 0.018862691  
## spam 0.994680851 0.005319149  
##   
## hoping  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## horny  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## hospital  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## hot  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.980496454 0.019503546  
##   
## hour  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 1.000000000 0.000000000  
##   
## hours  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.989361702 0.010638298  
##   
## house  
## sms\_data\_train$type No Yes  
## ham 0.992787795 0.007212205  
## spam 0.998226950 0.001773050  
##   
## however  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## hows  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## howz  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## hrs  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.968085106 0.031914894  
##   
## httpwwwurawinnercom  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## huh  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## hungry  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## hurry  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## hurt  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## hurts  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## ice  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## identifier  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97695035 0.02304965  
##   
## ill  
## sms\_data\_train$type No Yes  
## ham 0.990568655 0.009431345  
## spam 1.000000000 0.000000000  
##   
## immediately  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## important  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.982269504 0.017730496  
##   
## inc  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## india  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## info  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9875886525 0.0124113475  
##   
## information  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9875886525 0.0124113475  
##   
## inside  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## instead  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## interested  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## invited  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.994680851 0.005319149  
##   
## ipod  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98049645 0.01950355  
##   
## ish  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## ive  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## izzit  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## january  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## jay  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## job  
## sms\_data\_train$type No Yes  
## ham 0.99112344 0.00887656  
## spam 1.00000000 0.00000000  
##   
## john  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## join  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.980496454 0.019503546  
##   
## joined  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## joke  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## joking  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## jst  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## jus  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 1.00000000 0.00000000  
##   
## just  
## sms\_data\_train$type No Yes  
## ham 0.94368932 0.05631068  
## spam 0.89893617 0.10106383  
##   
## juz  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## kate  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## keep  
## sms\_data\_train$type No Yes  
## ham 0.986685160 0.013314840  
## spam 0.991134752 0.008865248  
##   
## keeping  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## kept  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## kick  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## kids  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## kind  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## kinda  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.998226950 0.001773050  
##   
## king  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9964539007 0.0035460993  
##   
## kiss  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 1.00000000 0.00000000  
##   
## knew  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## know  
## sms\_data\_train$type No Yes  
## ham 0.95478502 0.04521498  
## spam 0.97163121 0.02836879  
##   
## knows  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.99822695 0.00177305  
##   
## knw  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## lady  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## land  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9751773 0.0248227  
##   
## landline  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9521276596 0.0478723404  
##   
## laptop  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## lar  
## sms\_data\_train$type No Yes  
## ham 0.99112344 0.00887656  
## spam 1.00000000 0.00000000  
##   
## last  
## sms\_data\_train$type No Yes  
## ham 0.98696255 0.01303745  
## spam 0.98049645 0.01950355  
##   
## late  
## sms\_data\_train$type No Yes  
## ham 0.98779473 0.01220527  
## spam 0.99822695 0.00177305  
##   
## later  
## sms\_data\_train$type No Yes  
## ham 0.97198336 0.02801664  
## spam 1.00000000 0.00000000  
##   
## latest  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.948581560 0.051418440  
##   
## laugh  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.998226950 0.001773050  
##   
## lazy  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## ldn  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## learn  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## least  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## leave  
## sms\_data\_train$type No Yes  
## ham 0.990568655 0.009431345  
## spam 0.996453901 0.003546099  
##   
## leaves  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## leaving  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## lect  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## left  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 1.000000000 0.000000000  
##   
## leh  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 1.000000000 0.000000000  
##   
## lei  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## less  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## lesson  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## lessons  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## let  
## sms\_data\_train$type No Yes  
## ham 0.98557559 0.01442441  
## spam 1.00000000 0.00000000  
##   
## lets  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## liao  
## sms\_data\_train$type No Yes  
## ham 0.991955617 0.008044383  
## spam 1.000000000 0.000000000  
##   
## library  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## life  
## sms\_data\_train$type No Yes  
## ham 0.986962552 0.013037448  
## spam 0.994680851 0.005319149  
##   
## lift  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## light  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## like  
## sms\_data\_train$type No Yes  
## ham 0.95256588 0.04743412  
## spam 0.98226950 0.01773050  
##   
## liked  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## line  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.948581560 0.051418440  
##   
## link  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.992907801 0.007092199  
##   
## listen  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.996453901 0.003546099  
##   
## little  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 0.996453901 0.003546099  
##   
## live  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.968085106 0.031914894  
##   
## lmao  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## loads  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## loan  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## local  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9929078014 0.0070921986  
##   
## log  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.989361702 0.010638298  
##   
## lol  
## sms\_data\_train$type No Yes  
## ham 0.98446602 0.01553398  
## spam 1.00000000 0.00000000  
##   
## london  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.994680851 0.005319149  
##   
## long  
## sms\_data\_train$type No Yes  
## ham 0.991678225 0.008321775  
## spam 1.000000000 0.000000000  
##   
## longer  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## look  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 0.994680851 0.005319149  
##   
## lookin  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## looking  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 0.98404255 0.01595745  
##   
## looks  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## lor  
## sms\_data\_train$type No Yes  
## ham 0.97087379 0.02912621  
## spam 1.00000000 0.00000000  
##   
## lose  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.99822695 0.00177305  
##   
## lost  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.998226950 0.001773050  
##   
## lot  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 1.000000000 0.000000000  
##   
## lots  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.99822695 0.00177305  
##   
## lovable  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## love  
## sms\_data\_train$type No Yes  
## ham 0.96837725 0.03162275  
## spam 0.98936170 0.01063830  
##   
## loved  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## lovely  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## lover  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.996453901 0.003546099  
##   
## loves  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## loving  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## loyalty  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## ltd  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## luck  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.994680851 0.005319149  
##   
## lucky  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.989361702 0.010638298  
##   
## lunch  
## sms\_data\_train$type No Yes  
## ham 0.99223301 0.00776699  
## spam 1.00000000 0.00000000  
##   
## luv  
## sms\_data\_train$type No Yes  
## ham 0.993619972 0.006380028  
## spam 0.992907801 0.007092199  
##   
## mad  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## made  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 0.998226950 0.001773050  
##   
## mah  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## mail  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## make  
## sms\_data\_train$type No Yes  
## ham 0.98280166 0.01719834  
## spam 0.98049645 0.01950355  
##   
## makes  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## making  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## malaria  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 1.000000000 0.000000000  
##   
## man  
## sms\_data\_train$type No Yes  
## ham 0.991400832 0.008599168  
## spam 1.000000000 0.000000000  
##   
## many  
## sms\_data\_train$type No Yes  
## ham 0.988349515 0.011650485  
## spam 0.994680851 0.005319149  
##   
## march  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## mark  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## married  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## marry  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## match  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.992907801 0.007092199  
##   
## matches  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## mate  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## mates  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## maximize  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## maxmins  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## may  
## sms\_data\_train$type No Yes  
## ham 0.992787795 0.007212205  
## spam 0.989361702 0.010638298  
##   
## mayb  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## maybe  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.996453901 0.003546099  
##   
## mean  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## meaning  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## means  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## meant  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## medical  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## meds  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## meet  
## sms\_data\_train$type No Yes  
## ham 0.985298197 0.014701803  
## spam 0.992907801 0.007092199  
##   
## meeting  
## sms\_data\_train$type No Yes  
## ham 0.990568655 0.009431345  
## spam 1.000000000 0.000000000  
##   
## meh  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## member  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.992907801 0.007092199  
##   
## merry  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## message  
## sms\_data\_train$type No Yes  
## ham 0.98723994 0.01276006  
## spam 0.96453901 0.03546099  
##   
## messages  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.985815603 0.014184397  
##   
## met  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.998226950 0.001773050  
##   
## midnight  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## might  
## sms\_data\_train$type No Yes  
## ham 0.993619972 0.006380028  
## spam 1.000000000 0.000000000  
##   
## min  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 0.982269504 0.017730496  
##   
## mind  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 0.998226950 0.001773050  
##   
## mine  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## mins  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.946808511 0.053191489  
##   
## minute  
## sms\_data\_train$type No Yes  
## ham 0.997780860 0.002219140  
## spam 0.991134752 0.008865248  
##   
## minutes  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 0.991134752 0.008865248  
##   
## miracle  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## miss  
## sms\_data\_train$type No Yes  
## ham 0.986962552 0.013037448  
## spam 0.992907801 0.007092199  
##   
## missed  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## missing  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.998226950 0.001773050  
##   
## mistake  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## moan  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## mob  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97163121 0.02836879  
##   
## mobile  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.847517730 0.152482270  
##   
## mobiles  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9822695 0.0177305  
##   
## mobileupd  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97695035 0.02304965  
##   
## mode  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## mom  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## moment  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## moms  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## mon  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## monday  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## money  
## sms\_data\_train$type No Yes  
## ham 0.989736477 0.010263523  
## spam 0.996453901 0.003546099  
##   
## month  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 0.992907801 0.007092199  
##   
## months  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.994680851 0.005319149  
##   
## morning  
## sms\_data\_train$type No Yes  
## ham 0.98307906 0.01692094  
## spam 1.00000000 0.00000000  
##   
## mother  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## motorola  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98404255 0.01595745  
##   
## move  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## movie  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## movies  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## mrng  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## mrt  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## mrw  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## msg  
## sms\_data\_train$type No Yes  
## ham 0.98945908 0.01054092  
## spam 0.96099291 0.03900709  
##   
## msgs  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.994680851 0.005319149  
##   
## mths  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97695035 0.02304965  
##   
## much  
## sms\_data\_train$type No Yes  
## ham 0.97864078 0.02135922  
## spam 0.99822695 0.00177305  
##   
## mum  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## music  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98049645 0.01950355  
##   
## must  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.992907801 0.007092199  
##   
## muz  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.994680851 0.005319149  
##   
## nah  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## naked  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## name  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 0.980496454 0.019503546  
##   
## national  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## nature  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## naughty  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## near  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## need  
## sms\_data\_train$type No Yes  
## ham 0.96588072 0.03411928  
## spam 0.98936170 0.01063830  
##   
## needs  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.998226950 0.001773050  
##   
## net  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## network  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.96985816 0.03014184  
##   
## neva  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## never  
## sms\_data\_train$type No Yes  
## ham 0.991955617 0.008044383  
## spam 0.998226950 0.001773050  
##   
## new  
## sms\_data\_train$type No Yes  
## ham 0.98502080 0.01497920  
## spam 0.90957447 0.09042553  
##   
## news  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.991134752 0.008865248  
##   
## next  
## sms\_data\_train$type No Yes  
## ham 0.990291262 0.009708738  
## spam 0.980496454 0.019503546  
##   
## nice  
## sms\_data\_train$type No Yes  
## ham 0.9889043 0.0110957  
## spam 1.0000000 0.0000000  
##   
## nigeria  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## night  
## sms\_data\_train$type No Yes  
## ham 0.980582524 0.019417476  
## spam 0.991134752 0.008865248  
##   
## nite  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 0.994680851 0.005319149  
##   
## nobody  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## noe  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## nokia  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9202127660 0.0797872340  
##   
## noon  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## nope  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## normal  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## normptone  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## nothing  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 0.994680851 0.005319149  
##   
## notice  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.994680851 0.005319149  
##   
## now  
## sms\_data\_train$type No Yes  
## ham 0.93980583 0.06019417  
## spam 0.75709220 0.24290780  
##   
## num  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## number  
## sms\_data\_train$type No Yes  
## ham 0.98640777 0.01359223  
## spam 0.96453901 0.03546099  
##   
## numbers  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9893617021 0.0106382979  
##   
## nyt  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.998226950 0.001773050  
##   
## obviously  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## offer  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.957446809 0.042553191  
##   
## offers  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9822695035 0.0177304965  
##   
## office  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.998226950 0.001773050  
##   
## official  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9911347518 0.0088652482  
##   
## okay  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## okie  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## old  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.99822695 0.00177305  
##   
## omg  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## one  
## sms\_data\_train$type No Yes  
## ham 0.97004161 0.02995839  
## spam 0.98936170 0.01063830  
##   
## ones  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.998226950 0.001773050  
##   
## online  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## onto  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## oops  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## open  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.99822695 0.00177305  
##   
## operator  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9822695 0.0177305  
##   
## opinion  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## opt  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9751773 0.0248227  
##   
## optout  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## orange  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.960992908 0.039007092  
##   
## orchard  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## order  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.982269504 0.017730496  
##   
## oredi  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## oso  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## others  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## otherwise  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## outside  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## paid  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## pain  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 1.000000000 0.000000000  
##   
## paper  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## parents  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## park  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## part  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 0.984042553 0.015957447  
##   
## partner  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## party  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## pass  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## password  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## past  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## pay  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 0.99822695 0.00177305  
##   
## paying  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## people  
## sms\_data\_train$type No Yes  
## ham 0.989459085 0.010540915  
## spam 0.996453901 0.003546099  
##   
## per  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.945035461 0.054964539  
##   
## person  
## sms\_data\_train$type No Yes  
## ham 0.99223301 0.00776699  
## spam 1.00000000 0.00000000  
##   
## persons  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## pete  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## phone  
## sms\_data\_train$type No Yes  
## ham 0.98307906 0.01692094  
## spam 0.92730496 0.07269504  
##   
## phones  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9840425532 0.0159574468  
##   
## photo  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## photos  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## pic  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.987588652 0.012411348  
##   
## pick  
## sms\_data\_train$type No Yes  
## ham 0.98502080 0.01497920  
## spam 0.99822695 0.00177305  
##   
## picked  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## picking  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## pics  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.982269504 0.017730496  
##   
## pictures  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## pin  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## pix  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## pizza  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## place  
## sms\_data\_train$type No Yes  
## ham 0.990568655 0.009431345  
## spam 0.992907801 0.007092199  
##   
## plan  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 1.000000000 0.000000000  
##   
## planned  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## planning  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## plans  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## play  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.98226950 0.01773050  
##   
## player  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9822695035 0.0177304965  
##   
## players  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## please  
## sms\_data\_train$type No Yes  
## ham 0.98307906 0.01692094  
## spam 0.92198582 0.07801418  
##   
## pleasure  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## plenty  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## pls  
## sms\_data\_train$type No Yes  
## ham 0.98196949 0.01803051  
## spam 0.98758865 0.01241135  
##   
## plus  
## sms\_data\_train$type No Yes  
## ham 0.996671290 0.003328710  
## spam 0.994680851 0.005319149  
##   
## plz  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## pmin  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97340426 0.02659574  
##   
## pmsg  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9787234 0.0212766  
##   
## pobox  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.95921986 0.04078014  
##   
## point  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## points  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9769503546 0.0230496454  
##   
## police  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.998226950 0.001773050  
##   
## poly  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98404255 0.01595745  
##   
## polys  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## poor  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## possible  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## post  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.992907801 0.007092199  
##   
## posted  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## pound  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9875886525 0.0124113475  
##   
## pounds  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.980496454 0.019503546  
##   
## ppm  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.93971631 0.06028369  
##   
## pray  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## press  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.994680851 0.005319149  
##   
## pretty  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## price  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.971631206 0.028368794  
##   
## princess  
## sms\_data\_train$type No Yes  
## ham 0.99334258 0.00665742  
## spam 1.00000000 0.00000000  
##   
## private  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9734042553 0.0265957447  
##   
## prize  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.8741135 0.1258865  
##   
## prob  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## probably  
## sms\_data\_train$type No Yes  
## ham 0.991678225 0.008321775  
## spam 1.000000000 0.000000000  
##   
## problem  
## sms\_data\_train$type No Yes  
## ham 0.99223301 0.00776699  
## spam 1.00000000 0.00000000  
##   
## project  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## promise  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## pub  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.996453901 0.003546099  
##   
## put  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.998226950 0.001773050  
##   
## putting  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## quality  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## question  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.996453901 0.003546099  
##   
## questions  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.994680851 0.005319149  
##   
## quick  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## quite  
## sms\_data\_train$type No Yes  
## ham 0.991955617 0.008044383  
## spam 1.000000000 0.000000000  
##   
## quiz  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98049645 0.01950355  
##   
## rain  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## raining  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## rate  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.973404255 0.026595745  
##   
## rates  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## rather  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## rcvd  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## reach  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 0.991134752 0.008865248  
##   
## reached  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## reaching  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## read  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## reading  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## ready  
## sms\_data\_train$type No Yes  
## ham 0.991678225 0.008321775  
## spam 0.992907801 0.007092199  
##   
## real  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 0.985815603 0.014184397  
##   
## really  
## sms\_data\_train$type No Yes  
## ham 0.98391123 0.01608877  
## spam 1.00000000 0.00000000  
##   
## realy  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## reason  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## receipt  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.994680851 0.005319149  
##   
## receive  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.966312057 0.033687943  
##   
## recently  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## records  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## reference  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.994680851 0.005319149  
##   
## regards  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## registered  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## relation  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## relax  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## remember  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 1.000000000 0.000000000  
##   
## remind  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## remove  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.992907801 0.007092199  
##   
## rent  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## rental  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9787234 0.0212766  
##   
## replied  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## reply  
## sms\_data\_train$type No Yes  
## ham 0.990846047 0.009153953  
## spam 0.891843972 0.108156028  
##   
## replying  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9911347518 0.0088652482  
##   
## representative  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## request  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## rest  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## review  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.994680851 0.005319149  
##   
## reward  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## right  
## sms\_data\_train$type No Yes  
## ham 0.982801664 0.017198336  
## spam 0.994680851 0.005319149  
##   
## ring  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 0.996453901 0.003546099  
##   
## ringtone  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.96276596 0.03723404  
##   
## rite  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 1.000000000 0.000000000  
##   
## road  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## rock  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## role  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## room  
## sms\_data\_train$type No Yes  
## ham 0.99223301 0.00776699  
## spam 1.00000000 0.00000000  
##   
## rose  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## round  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## rowwjhl  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## rply  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9911347518 0.0088652482  
##   
## rreveal  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## run  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 1.00000000 0.00000000  
##   
## sad  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## sae  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.96808511 0.03191489  
##   
## safe  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## said  
## sms\_data\_train$type No Yes  
## ham 0.98196949 0.01803051  
## spam 1.00000000 0.00000000  
##   
## sale  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9929078014 0.0070921986  
##   
## sat  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.998226950 0.001773050  
##   
## saturday  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.991134752 0.008865248  
##   
## savamob  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## save  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## saw  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## say  
## sms\_data\_train$type No Yes  
## ham 0.98002774 0.01997226  
## spam 0.99822695 0.00177305  
##   
## saying  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## says  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 1.00000000 0.00000000  
##   
## sch  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## school  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## screaming  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## sea  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## search  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 0.998226950 0.001773050  
##   
## sec  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## second  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## secret  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.984042553 0.015957447  
##   
## see  
## sms\_data\_train$type No Yes  
## ham 0.96809986 0.03190014  
## spam 0.97517730 0.02482270  
##   
## seeing  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## seems  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## seen  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## selected  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9680851064 0.0319148936  
##   
## self  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## sell  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## semester  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## send  
## sms\_data\_train$type No Yes  
## ham 0.97226075 0.02773925  
## spam 0.91134752 0.08865248  
##   
## sending  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.998226950 0.001773050  
##   
## sense  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## sent  
## sms\_data\_train$type No Yes  
## ham 0.98862691 0.01137309  
## spam 0.98581560 0.01418440  
##   
## serious  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## seriously  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## service  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.92553191 0.07446809  
##   
## services  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9787234043 0.0212765957  
##   
## set  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 0.998226950 0.001773050  
##   
## sex  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.985815603 0.014184397  
##   
## sexy  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.98049645 0.01950355  
##   
## shall  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 1.00000000 0.00000000  
##   
## share  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## shd  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## shop  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.992907801 0.007092199  
##   
## shopping  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 0.98936170 0.01063830  
##   
## short  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## show  
## sms\_data\_train$type No Yes  
## ham 0.995839112 0.004160888  
## spam 0.996453901 0.003546099  
##   
## shower  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## shows  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9539007092 0.0460992908  
##   
## sick  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## side  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## sight  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 1.0000000000 0.0000000000  
##   
## sign  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## silent  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## simple  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## since  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 0.996453901 0.003546099  
##   
## single  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.998226950 0.001773050  
##   
## sipix  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## sir  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.998226950 0.001773050  
##   
## sis  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## sister  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## sit  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## sitting  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## situation  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## skxh  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## slave  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## sleep  
## sms\_data\_train$type No Yes  
## ham 0.98945908 0.01054092  
## spam 1.00000000 0.00000000  
##   
## sleeping  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## slept  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## slow  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.998226950 0.001773050  
##   
## slowly  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## small  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## smile  
## sms\_data\_train$type No Yes  
## ham 0.992787795 0.007212205  
## spam 1.000000000 0.000000000  
##   
## smiling  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## smoke  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## sms  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 0.968085106 0.031914894  
##   
## smth  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## snow  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## sofa  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## sol  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## somebody  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.996453901 0.003546099  
##   
## someone  
## sms\_data\_train$type No Yes  
## ham 0.992787795 0.007212205  
## spam 0.982269504 0.017730496  
##   
## something  
## sms\_data\_train$type No Yes  
## ham 0.98446602 0.01553398  
## spam 1.00000000 0.00000000  
##   
## sometimes  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## somewhere  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## song  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.998226950 0.001773050  
##   
## sony  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9911347518 0.0088652482  
##   
## sonyericsson  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## soon  
## sms\_data\_train$type No Yes  
## ham 0.988626907 0.011373093  
## spam 0.994680851 0.005319149  
##   
## sorry  
## sms\_data\_train$type No Yes  
## ham 0.970873786 0.029126214  
## spam 0.996453901 0.003546099  
##   
## sort  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.998226950 0.001773050  
##   
## sound  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## sounds  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## south  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## space  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## speak  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 0.98758865 0.01241135  
##   
## special  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 0.98758865 0.01241135  
##   
## specialcall  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## specially  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## spend  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## spent  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## spoke  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## spree  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## stand  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## start  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 0.994680851 0.005319149  
##   
## started  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## starting  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.992907801 0.007092199  
##   
## starts  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## statement  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97695035 0.02304965  
##   
## station  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## stay  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 0.994680851 0.005319149  
##   
## staying  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## std  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## still  
## sms\_data\_train$type No Yes  
## ham 0.970873786 0.029126214  
## spam 0.992907801 0.007092199  
##   
## stockport  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## stop  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.877659574 0.122340426  
##   
## store  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9893617021 0.0106382979  
##   
## story  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## street  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.996453901 0.003546099  
##   
## study  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## studying  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## stuff  
## sms\_data\_train$type No Yes  
## ham 0.98973648 0.01026352  
## spam 0.99822695 0.00177305  
##   
## stupid  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## sub  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.987588652 0.012411348  
##   
## sucks  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## summer  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9929078014 0.0070921986  
##   
## sun  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.998226950 0.001773050  
##   
## sunday  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## sunshine  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9875886525 0.0124113475  
##   
## sup  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## support  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.996453901 0.003546099  
##   
## supposed  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## sure  
## sms\_data\_train$type No Yes  
## ham 0.98446602 0.01553398  
## spam 1.00000000 0.00000000  
##   
## surely  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## surprise  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.992907801 0.007092199  
##   
## sweet  
## sms\_data\_train$type No Yes  
## ham 0.99334258 0.00665742  
## spam 1.00000000 0.00000000  
##   
## swing  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## system  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## take  
## sms\_data\_train$type No Yes  
## ham 0.97614424 0.02385576  
## spam 0.97695035 0.02304965  
##   
## takes  
## sms\_data\_train$type No Yes  
## ham 0.997780860 0.002219140  
## spam 0.992907801 0.007092199  
##   
## taking  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.998226950 0.001773050  
##   
## talk  
## sms\_data\_train$type No Yes  
## ham 0.991123440 0.008876560  
## spam 0.996453901 0.003546099  
##   
## talking  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## tampa  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## tariffs  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## tcs  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.95921986 0.04078014  
##   
## tea  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## teach  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## tear  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## tel  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.998226950 0.001773050  
##   
## tell  
## sms\_data\_train$type No Yes  
## ham 0.97337032 0.02662968  
## spam 0.98049645 0.01950355  
##   
## telling  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## tells  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9964539007 0.0035460993  
##   
## ten  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.998226950 0.001773050  
##   
## tenerife  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## terms  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9893617021 0.0106382979  
##   
## test  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## text  
## sms\_data\_train$type No Yes  
## ham 0.98446602 0.01553398  
## spam 0.86524823 0.13475177  
##   
## texting  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## texts  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.975177305 0.024822695  
##   
## thank  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 0.998226950 0.001773050  
##   
## thanks  
## sms\_data\_train$type No Yes  
## ham 0.98668516 0.01331484  
## spam 0.98226950 0.01773050  
##   
## thanx  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 1.00000000 0.00000000  
##   
## thats  
## sms\_data\_train$type No Yes  
## ham 0.991400832 0.008599168  
## spam 1.000000000 0.000000000  
##   
## thing  
## sms\_data\_train$type No Yes  
## ham 0.98613037 0.01386963  
## spam 0.99822695 0.00177305  
##   
## things  
## sms\_data\_train$type No Yes  
## ham 0.98945908 0.01054092  
## spam 0.99822695 0.00177305  
##   
## think  
## sms\_data\_train$type No Yes  
## ham 0.97420250 0.02579750  
## spam 0.99822695 0.00177305  
##   
## thinkin  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## thinking  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## thinks  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.98404255 0.01595745  
##   
## thk  
## sms\_data\_train$type No Yes  
## ham 0.990291262 0.009708738  
## spam 1.000000000 0.000000000  
##   
## tho  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 1.000000000 0.000000000  
##   
## though  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 1.00000000 0.00000000  
##   
## thought  
## sms\_data\_train$type No Yes  
## ham 0.990568655 0.009431345  
## spam 0.998226950 0.001773050  
##   
## thru  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## tht  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## thurs  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9964539007 0.0035460993  
##   
## tick  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 1.000000000 0.000000000  
##   
## ticket  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 1.0000000000 0.0000000000  
##   
## tickets  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.994680851 0.005319149  
##   
## til  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 1.000000000 0.000000000  
##   
## till  
## sms\_data\_train$type No Yes  
## ham 0.996116505 0.003883495  
## spam 0.998226950 0.001773050  
##   
## time  
## sms\_data\_train$type No Yes  
## ham 0.96005548 0.03994452  
## spam 0.97340426 0.02659574  
##   
## times  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 1.000000000 0.000000000  
##   
## timing  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## tired  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## tmr  
## sms\_data\_train$type No Yes  
## ham 0.994729542 0.005270458  
## spam 1.000000000 0.000000000  
##   
## toclaim  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## today  
## sms\_data\_train$type No Yes  
## ham 0.97614424 0.02385576  
## spam 0.97340426 0.02659574  
##   
## todays  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 0.969858156 0.030141844  
##   
## together  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## told  
## sms\_data\_train$type No Yes  
## ham 0.98918169 0.01081831  
## spam 1.00000000 0.00000000  
##   
## tomo  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## tomorrow  
## sms\_data\_train$type No Yes  
## ham 0.98446602 0.01553398  
## spam 0.98758865 0.01241135  
##   
## tone  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.95035461 0.04964539  
##   
## tones  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97340426 0.02659574  
##   
## tonight  
## sms\_data\_train$type No Yes  
## ham 0.988904300 0.011095700  
## spam 0.996453901 0.003546099  
##   
## tonite  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## took  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## top  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.975177305 0.024822695  
##   
## torch  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## tot  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## touch  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.998226950 0.001773050  
##   
## tough  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## tour  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9946808511 0.0053191489  
##   
## towards  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## town  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 0.996453901 0.003546099  
##   
## track  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## train  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## training  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## transaction  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## treat  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 1.000000000 0.000000000  
##   
## tried  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.975177305 0.024822695  
##   
## trip  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.992907801 0.007092199  
##   
## trouble  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## true  
## sms\_data\_train$type No Yes  
## ham 0.996393897 0.003606103  
## spam 0.996453901 0.003546099  
##   
## trust  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## truth  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## try  
## sms\_data\_train$type No Yes  
## ham 0.990013870 0.009986130  
## spam 0.992907801 0.007092199  
##   
## trying  
## sms\_data\_train$type No Yes  
## ham 0.99334258 0.00665742  
## spam 0.97695035 0.02304965  
##   
## tscs  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.96808511 0.03191489  
##   
## ttyl  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## tuesday  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## twice  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## two  
## sms\_data\_train$type No Yes  
## ham 0.99334258 0.00665742  
## spam 1.00000000 0.00000000  
##   
## txt  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 0.81382979 0.18617021  
##   
## txting  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9858156028 0.0141843972  
##   
## txts  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98404255 0.01595745  
##   
## type  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## ufind  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9858156 0.0141844  
##   
## ugh  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## uks  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## ull  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## uncle  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## understand  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## unless  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## unlimited  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9875886525 0.0124113475  
##   
## unredeemed  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97695035 0.02304965  
##   
## unsub  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## unsubscribe  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9822695 0.0177305  
##   
## update  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9804964539 0.0195035461  
##   
## ure  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## urgent  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9131205674 0.0868794326  
##   
## urself  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## use  
## sms\_data\_train$type No Yes  
## ham 0.99334258 0.00665742  
## spam 0.98758865 0.01241135  
##   
## used  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.998226950 0.001773050  
##   
## user  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9893617021 0.0106382979  
##   
## usf  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## using  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## usual  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## uve  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9911347518 0.0088652482  
##   
## valentine  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## valentines  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## valid  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.96808511 0.03191489  
##   
## valued  
## sms\_data\_train$type No Yes  
## ham 0.999445215 0.000554785  
## spam 0.984042553 0.015957447  
##   
## via  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.996453901 0.003546099  
##   
## video  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9645390071 0.0354609929  
##   
## vikky  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## visit  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.989361702 0.010638298  
##   
## vodafone  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## voice  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## vomit  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## voucher  
## sms\_data\_train$type No Yes  
## ham 0.9997226075 0.0002773925  
## spam 0.9822695035 0.0177304965  
##   
## vouchers  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97695035 0.02304965  
##   
## wait  
## sms\_data\_train$type No Yes  
## ham 0.987794730 0.012205270  
## spam 0.996453901 0.003546099  
##   
## waiting  
## sms\_data\_train$type No Yes  
## ham 0.992510402 0.007489598  
## spam 0.978723404 0.021276596  
##   
## wake  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 1.000000000 0.000000000  
##   
## waking  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## walk  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 0.99822695 0.00177305  
##   
## walking  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## wan  
## sms\_data\_train$type No Yes  
## ham 0.990013870 0.009986130  
## spam 0.992907801 0.007092199  
##   
## wana  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## wanna  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 0.985815603 0.014184397  
##   
## want  
## sms\_data\_train$type No Yes  
## ham 0.96976422 0.03023578  
## spam 0.96631206 0.03368794  
##   
## wanted  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 0.99822695 0.00177305  
##   
## wants  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.998226950 0.001773050  
##   
## wap  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.98758865 0.01241135  
##   
## warm  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## waste  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## wat  
## sms\_data\_train$type No Yes  
## ham 0.98307906 0.01692094  
## spam 0.99822695 0.00177305  
##   
## watch  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## watching  
## sms\_data\_train$type No Yes  
## ham 0.99445215 0.00554785  
## spam 1.00000000 0.00000000  
##   
## water  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## wats  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## way  
## sms\_data\_train$type No Yes  
## ham 0.97780860 0.02219140  
## spam 0.99822695 0.00177305  
##   
## weather  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## wed  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## wednesday  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## weed  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## week  
## sms\_data\_train$type No Yes  
## ham 0.98751734 0.01248266  
## spam 0.94503546 0.05496454  
##   
## weekend  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.996453901 0.003546099  
##   
## weekends  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9911347518 0.0088652482  
##   
## weekly  
## sms\_data\_train$type No Yes  
## ham 1.00000000 0.00000000  
## spam 0.97163121 0.02836879  
##   
## weeks  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 0.978723404 0.021276596  
##   
## welcome  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.992907801 0.007092199  
##   
## well  
## sms\_data\_train$type No Yes  
## ham 0.975589459 0.024410541  
## spam 0.991134752 0.008865248  
##   
## wen  
## sms\_data\_train$type No Yes  
## ham 0.994174757 0.005825243  
## spam 1.000000000 0.000000000  
##   
## went  
## sms\_data\_train$type No Yes  
## ham 0.98945908 0.01054092  
## spam 1.00000000 0.00000000  
##   
## whatever  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## whats  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 0.998226950 0.001773050  
##   
## whenever  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## whole  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 1.000000000 0.000000000  
##   
## wid  
## sms\_data\_train$type No Yes  
## ham 0.997226075 0.002773925  
## spam 0.996453901 0.003546099  
##   
## wif  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## wife  
## sms\_data\_train$type No Yes  
## ham 0.99667129 0.00332871  
## spam 1.00000000 0.00000000  
##   
## wil  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## will  
## sms\_data\_train$type No Yes  
## ham 0.93342580 0.06657420  
## spam 0.93617021 0.06382979  
##   
## willing  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.998226950 0.001773050  
##   
## win  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.91134752 0.08865248  
##   
## wine  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## winner  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9751773 0.0248227  
##   
## wins  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.99822695 0.00177305  
##   
## wish  
## sms\_data\_train$type No Yes  
## ham 0.991955617 0.008044383  
## spam 0.996453901 0.003546099  
##   
## wishing  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## wit  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## within  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9893617021 0.0106382979  
##   
## without  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## wiv  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## wkly  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9822695 0.0177305  
##   
## wks  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9982269504 0.0017730496  
##   
## wnt  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## woke  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## won  
## sms\_data\_train$type No Yes  
## ham 1.000000 0.000000  
## spam 0.891844 0.108156  
##   
## wonder  
## sms\_data\_train$type No Yes  
## ham 0.997503467 0.002496533  
## spam 1.000000000 0.000000000  
##   
## wonderful  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## wont  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## word  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 0.966312057 0.033687943  
##   
## words  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 1.000000000 0.000000000  
##   
## work  
## sms\_data\_train$type No Yes  
## ham 0.981137309 0.018862691  
## spam 0.996453901 0.003546099  
##   
## workin  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## working  
## sms\_data\_train$type No Yes  
## ham 0.993897365 0.006102635  
## spam 1.000000000 0.000000000  
##   
## works  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## world  
## sms\_data\_train$type No Yes  
## ham 0.993065187 0.006934813  
## spam 0.998226950 0.001773050  
##   
## worried  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## worries  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 1.000000000 0.000000000  
##   
## worry  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 1.00000000 0.00000000  
##   
## worse  
## sms\_data\_train$type No Yes  
## ham 0.998335645 0.001664355  
## spam 1.000000000 0.000000000  
##   
## worth  
## sms\_data\_train$type No Yes  
## ham 0.9991678225 0.0008321775  
## spam 0.9804964539 0.0195035461  
##   
## wot  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.998226950 0.001773050  
##   
## wow  
## sms\_data\_train$type No Yes  
## ham 0.998613037 0.001386963  
## spam 0.996453901 0.003546099  
##   
## write  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## wrong  
## sms\_data\_train$type No Yes  
## ham 0.996948682 0.003051318  
## spam 1.000000000 0.000000000  
##   
## wwq  
## sms\_data\_train$type No Yes  
## ham 1.000000000 0.000000000  
## spam 0.991134752 0.008865248  
##   
## wwwgetzedcouk  
## sms\_data\_train$type No Yes  
## ham 1.0000000 0.0000000  
## spam 0.9893617 0.0106383  
##   
## xmas  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 0.98226950 0.01773050  
##   
## xxx  
## sms\_data\_train$type No Yes  
## ham 0.995006935 0.004993065  
## spam 0.980496454 0.019503546  
##   
## yahoo  
## sms\_data\_train$type No Yes  
## ham 0.998890430 0.001109570  
## spam 0.996453901 0.003546099  
##   
## yar  
## sms\_data\_train$type No Yes  
## ham 0.99778086 0.00221914  
## spam 1.00000000 0.00000000  
##   
## yeah  
## sms\_data\_train$type No Yes  
## ham 0.98141470 0.01858530  
## spam 0.99822695 0.00177305  
##   
## year  
## sms\_data\_train$type No Yes  
## ham 0.991678225 0.008321775  
## spam 0.992907801 0.007092199  
##   
## years  
## sms\_data\_train$type No Yes  
## ham 0.99556172 0.00443828  
## spam 0.98936170 0.01063830  
##   
## yep  
## sms\_data\_train$type No Yes  
## ham 0.998058252 0.001941748  
## spam 1.000000000 0.000000000  
##   
## yes  
## sms\_data\_train$type No Yes  
## ham 0.98474341 0.01525659  
## spam 0.97340426 0.02659574  
##   
## yesterday  
## sms\_data\_train$type No Yes  
## ham 0.995284327 0.004715673  
## spam 0.996453901 0.003546099  
##   
## yet  
## sms\_data\_train$type No Yes  
## ham 0.990568655 0.009431345  
## spam 0.996453901 0.003546099  
##   
## yoga  
## sms\_data\_train$type No Yes  
## ham 0.99889043 0.00110957  
## spam 1.00000000 0.00000000  
##   
## yup  
## sms\_data\_train$type No Yes  
## ham 0.99112344 0.00887656  
## spam 1.00000000 0.00000000

#---------------------------------------------- Evaluating model performance --------------------  
sms\_test\_pred <- predict(sms\_classifier, sms\_test)  
CrossTable(sms\_test\_pred, sms\_data\_test$type,  
 prop.chisq = FALSE, prop.t = FALSE, prop.r = FALSE,  
 dnn = c('predicted', 'actual'))

##   
##   
## Cell Contents  
## |-------------------------|  
## | N |  
## | N / Col Total |  
## |-------------------------|  
##   
##   
## Total Observations in Table: 1390   
##   
##   
## | actual   
## predicted | ham | spam | Row Total |   
## -------------|-----------|-----------|-----------|  
## ham | 1202 | 32 | 1234 |   
## | 0.996 | 0.175 | |   
## -------------|-----------|-----------|-----------|  
## spam | 5 | 151 | 156 |   
## | 0.004 | 0.825 | |   
## -------------|-----------|-----------|-----------|  
## Column Total | 1207 | 183 | 1390 |   
## | 0.868 | 0.132 | |   
## -------------|-----------|-----------|-----------|  
##   
##

#---------------------------------------------------Improving model performance ----------------------  
  
sms\_classifier2 <- naiveBayes(sms\_train, sms\_data\_train$type, laplace = 1)  
sms\_test\_pred2 <- predict(sms\_classifier2, sms\_test)  
CrossTable(sms\_test\_pred2, sms\_data\_test$type,  
 prop.chisq = FALSE, prop.t = FALSE, prop.r = FALSE,  
 dnn = c('predicted', 'actual'))

##   
##   
## Cell Contents  
## |-------------------------|  
## | N |  
## | N / Col Total |  
## |-------------------------|  
##   
##   
## Total Observations in Table: 1390   
##   
##   
## | actual   
## predicted | ham | spam | Row Total |   
## -------------|-----------|-----------|-----------|  
## ham | 1204 | 31 | 1235 |   
## | 0.998 | 0.169 | |   
## -------------|-----------|-----------|-----------|  
## spam | 3 | 152 | 155 |   
## | 0.002 | 0.831 | |   
## -------------|-----------|-----------|-----------|  
## Column Total | 1207 | 183 | 1390 |   
## | 0.868 | 0.132 | |   
## -------------|-----------|-----------|-----------|  
##   
##

# Resources

* [Data Mining Algorithms In R/Classification/Na?ve Bayes](https://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/Classification/Na%C3%AFve_Bayes)
* [Naive Bayes Classifier in r - YouTube](https://www.youtube.com/watch?v=-SeyrC4yZF4)
* [naive Bayes in R - Josh Walters](http://joshwalters.com/2012/11/27/naive-bayes-classification-in-r.html)
* [R: Naive Bayes Classifier](http://ugrad.stat.ubc.ca/R/library/e1071/html/naiveBayes.html)
* [R code: classification with Naive Bayes - Katrin Erk](http://www.katrinerk.com/courses/r-worksheets/r-code-classification-with-naive-bayes)