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 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:

$$p(C_k|x_1, \dots, x_n) = \frac{1}{Z} p(C_k) \prod_{i=1}^{n} p\left(x_i|C_k\right)$$

where the evidence $Z = p(\mathbf{x})$ is a scaling factor dependent only on $x_1, \dots, x_n$, 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.
```
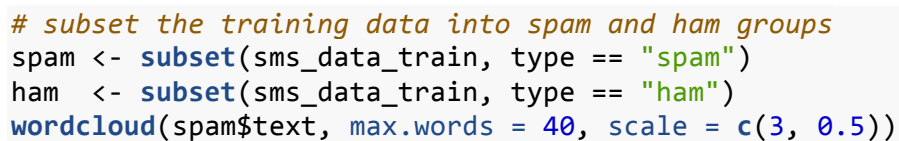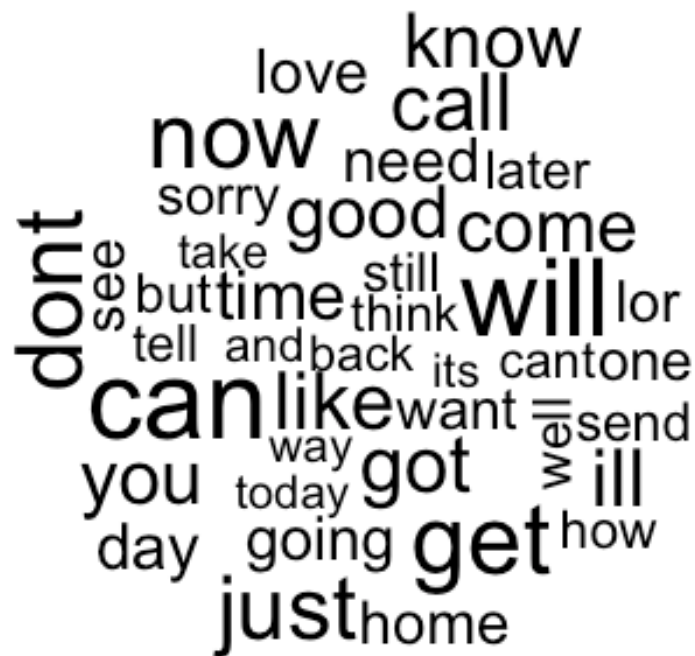
```
# subset the training data into spam and ham groups
spam <- subset(sms_data_train, type == "spam")
ham  <- subset(sms_data_train, type == "ham")
wordcloud(spam$text, max.words = 40, scale = c(3, 0.5))
```

reply
yourprize call
camera nokia urgent
new
claim please draw
win you won per
1000 just latest text
2000 phone chat
free get cash
contact tone
send 500
will now txt
mobile stop

```
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
```

```r
#------------------------------------------------    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 |           |
## -------------|------------|------------|-----------|
##
##
```

```r
#-------------------------------------------------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

- Naive Bayes Classifier in r - YouTube

- naive Bayes in R - Josh Walters

- R: Naive Bayes Classifier

- R code: classification with Naive Bayes - Katrin Erk