

Clasificador Bayesiano de Filtro de Spam

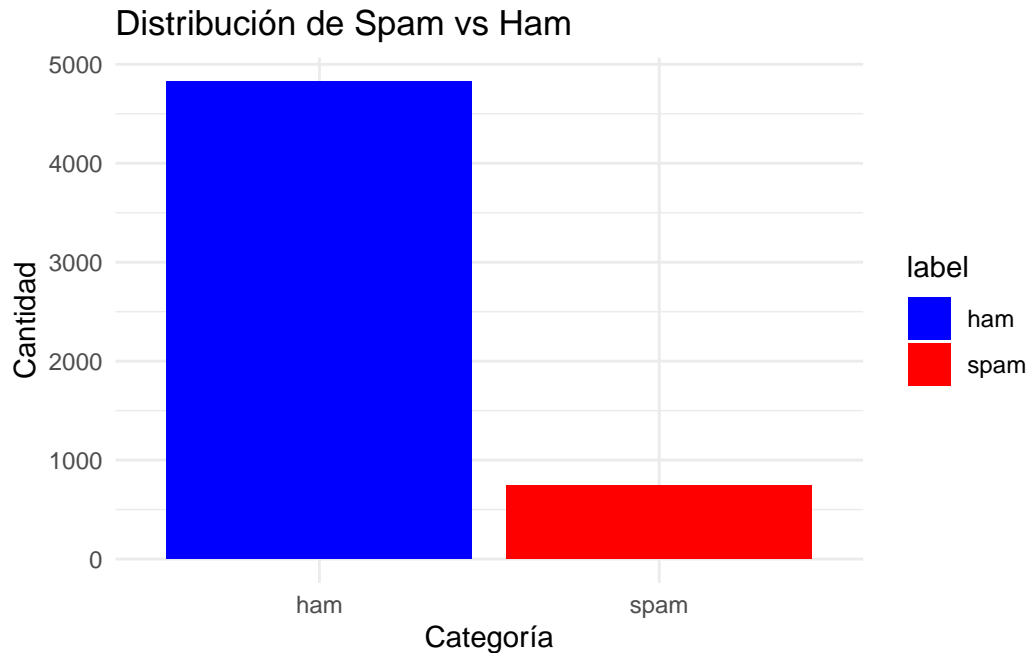
Tomás Ávila

Andrés León

```
library(tidyverse)
library(tm)
library(e1071)
library(SnowballC)
library(caTools)
library(caret)
library(pROC)
```

Análisis descriptivo

```
# Histograma
ggplot(df, aes(x = label, fill = label)) +
  geom_bar() +
  labs(
    title = "Distribución de Spam vs Ham",
    x = "Categoría",
    y = "Cantidad"
  ) +
  theme_minimal() +
  scale_fill_manual(values = c("ham" = "blue", "spam" = "red"))
```



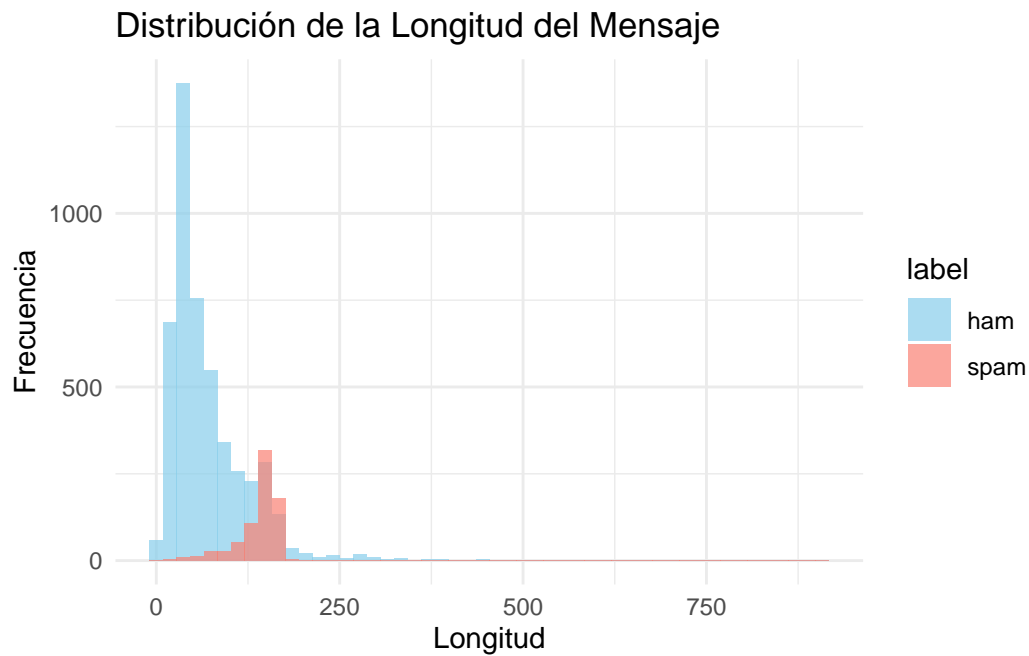
```
# Distribución porcentual
knitr::kable(df %>%
  count(label) %>%
  mutate(percentage = n / sum(n) * 100))
```

label	n	percentage
ham	4825	86.59368
spam	747	13.40632

```
# Longitud del mensaje
df <- df %>%
  mutate(message_length = nchar(message))

# Distribución de la longitud del mensaje
ggplot(df, aes(x = message_length, fill = label)) +
  geom_histogram(bins = 50, position = "identity", alpha = 0.7) +
  labs(
    title = "Distribución de la Longitud del Mensaje",
    x = "Longitud",
    y = "Frecuencia"
  ) +
```

```
theme_minimal() +
scale_fill_manual(values = c("ham" = "skyblue", "spam" = "salmon"))
```



Modelo regresión logística

```
# Corpus columna mensajes
corpus <- Corpus(VectorSource(df$message))

# Limpieza
corpus_clean <- corpus %>%
  tm_map(content_transformer(tolower)) %>%
  tm_map(removePunctuation) %>%
  tm_map(removeNumbers) %>%
  tm_map(removeWords, stopwords("english")) %>%
  tm_map(stemDocument) %>%
  tm_map(stripWhitespace)

# Document-Term Matrix (DTM)
dtm <- DocumentTermMatrix(corpus_clean)
```

```
# Reducir la dimensionalidad
sparse_dtm <- removeSparseTerms(dtm, 0.999)

dim(sparse_dtm)
```

```
[1] 5572 1208
```

```
X <- as.data.frame(as.matrix(sparse_dtm))
X$label <- df$label
```

```
set.seed(123)

split <- sample.split(X$label, SplitRatio = 0.80)
train_set <- subset(X, split == TRUE)
test_set <- subset(X, split == FALSE)
```

Call:

```
glm(formula = label ~ ., family = "binomial", data = train_set)
```

Coefficients: (13 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-3.861e+00	7.145e-01	-5.403	6.54e-08 ***
avail	-6.698e+01	3.269e+05	0.000	1.000
bugi	7.351e+01	6.735e+05	0.000	1.000
cine	5.623e+00	4.912e+05	0.000	1.000
crazi	-1.730e+01	3.619e+04	0.000	1.000
got	-1.940e+01	4.008e+04	0.000	1.000
great	-1.588e+01	1.167e+04	-0.001	0.999
point	-1.676e+01	1.242e+05	0.000	1.000
wat	-1.757e+01	1.931e+04	-0.001	0.999
world	-1.943e+01	5.265e+04	0.000	1.000
joke	-5.220e+01	2.692e+05	0.000	1.000
lar	2.707e+01	1.786e+05	0.000	1.000
wif	5.774e+00	2.074e+05	0.000	1.000
appli	-1.977e+01	2.291e+05	0.000	1.000
comp	-6.826e+01	4.837e+05	0.000	1.000
cup	-2.262e+01	1.565e+05	0.000	1.000
entri	1.352e+01	4.073e+05	0.000	1.000
final	-4.626e+01	3.290e+05	0.000	1.000
free	4.736e+00	3.152e+04	0.000	1.000

may	-9.839e+01	1.346e+05	-0.001	0.999
receiv	-3.745e+01	5.640e+04	-0.001	0.999
text	4.626e+01	6.340e+04	0.001	0.999
txt	2.270e+01	1.745e+05	0.000	1.000
win	1.371e+00	1.172e+05	0.000	1.000
wkli	2.189e+02	5.519e+05	0.000	1.000
alreadi	-3.405e+00	7.860e+04	0.000	1.000
dun	3.400e+01	1.977e+05	0.000	1.000
earli	-4.413e+01	2.798e+05	0.000	1.000
say	-5.532e+00	9.205e+04	0.000	1.000
around	-6.683e+00	1.758e+05	0.000	1.000
dont	-1.847e+01	8.527e+04	0.000	1.000
goe	-4.549e+00	6.099e+05	0.000	1.000
live	1.106e+00	7.252e+04	0.000	1.000
nah	-2.008e+01	9.540e+04	0.000	1.000
think	-1.429e+01	6.886e+04	0.000	1.000
though	1.827e+01	1.737e+05	0.000	1.000
usf	3.741e+01	3.059e+05	0.000	1.000
back	2.506e+00	5.866e+04	0.000	1.000
darl	1.450e+02	2.348e+05	0.001	1.000
freemsg	5.106e+01	7.832e+05	0.000	1.000
fun	1.377e+01	5.754e+04	0.000	1.000
hey	-5.665e+01	1.222e+05	0.000	1.000
like	-5.108e+01	8.371e+04	-0.001	1.000
now	-6.932e+00	3.456e+04	0.000	1.000
send	-1.736e+01	1.312e+04	-0.001	0.999
std	5.694e+01	4.192e+05	0.000	1.000
still	-5.248e+01	5.264e+04	-0.001	0.999
week	-1.334e+01	1.741e+05	0.000	1.000
word	1.515e+00	8.044e+04	0.000	1.000
xxx	2.131e+00	5.687e+04	0.000	1.000
`ã&`	1.147e+02	6.943e+04	0.002	0.999
brother	-1.819e+01	3.694e+04	0.000	1.000
even	5.285e+00	6.412e+04	0.000	1.000
speak	2.083e+00	1.774e+05	0.000	1.000
treat	-4.177e+00	1.457e+05	0.000	1.000
caller	5.875e+01	6.033e+05	0.000	1.000
copi	-3.062e+01	8.526e+05	0.000	1.000
friend	9.519e-02	1.920e+04	0.000	1.000
per	1.928e+00	2.407e+05	0.000	1.000
press	-1.285e+01	8.958e+04	0.000	1.000
request	-2.812e+01	3.773e+05	0.000	1.000
set	-1.423e+01	2.044e+05	0.000	1.000

call	-1.558e+01	4.497e+03	-0.003	0.997
claim	4.336e+01	2.326e+05	0.000	1.000
code	-6.469e+00	7.921e+05	0.000	1.000
custom	4.104e-01	2.271e+05	0.000	1.000
hour	7.586e+00	1.645e+05	0.000	1.000
network	1.794e+01	1.686e+05	0.000	1.000
prize	-2.948e+01	2.432e+05	0.000	1.000
reward	2.541e+01	2.318e+05	0.000	1.000
select	-1.865e+01	4.614e+04	0.000	1.000
valid	-2.372e+01	5.996e+05	0.000	1.000
valu	-2.266e+01	5.480e+05	0.000	1.000
winner	6.619e+00	3.844e+05	0.000	1.000
camera	-3.895e+01	3.350e+05	0.000	1.000
colour	2.854e+01	7.158e+05	0.000	1.000
entitl	3.293e+01	6.725e+05	0.000	1.000
latest	1.723e+01	2.628e+05	0.000	1.000
mobil	9.537e+00	7.981e+04	0.000	1.000
month	-1.104e+01	1.121e+05	0.000	1.000
updat	-1.159e+01	2.433e+05	0.000	1.000
anymor	-3.340e+01	3.163e+05	0.000	1.000
cri	7.915e+01	3.437e+05	0.000	1.000
enough	1.714e+00	1.572e+05	0.000	1.000
gonna	-3.039e+01	1.503e+05	0.000	1.000
home	-1.869e+01	2.721e+04	-0.001	0.999
ive	-2.080e+01	2.204e+05	0.000	1.000
soon	-1.277e+01	1.684e+05	0.000	1.000
stuff	-1.638e+01	1.504e+04	-0.001	0.999
talk	-1.669e+01	7.877e+03	-0.002	0.998
today	-1.746e+01	1.812e+04	-0.001	0.999
tonight	2.977e+00	1.559e+05	0.000	1.000
want	-1.977e+01	4.959e+04	0.000	1.000
cash	-2.830e+00	1.574e+05	0.000	1.000
chanc	-3.256e+00	1.215e+05	0.000	1.000
cost	9.627e+00	9.422e+04	0.000	1.000
day	8.260e+00	2.446e+04	0.000	1.000
info	1.317e+02	9.741e+05	0.000	1.000
pound	2.977e+01	1.488e+05	0.000	1.000
repli	-1.593e+01	1.201e+04	-0.001	0.999
pobox	-8.533e+00	4.970e+05	0.000	1.000
urgent	1.211e+01	5.348e+05	0.000	1.000
won	1.233e+02	3.619e+05	0.000	1.000
bless	-1.220e+02	6.134e+05	0.000	1.000
help	9.728e-01	5.856e+04	0.000	1.000

promis	2.213e+00	7.137e+05	0.000	1.000
right	-1.704e+01	2.091e+04	-0.001	0.999
search	-6.807e+01	3.199e+05	0.000	1.000
take	-1.864e+00	2.042e+04	0.000	1.000
thank	-3.760e+01	1.046e+05	0.000	1.000
time	-1.513e+01	7.693e+03	-0.002	0.998
will	-1.691e+01	1.919e+04	-0.001	0.999
wonder	5.063e+00	8.560e+04	0.000	1.000
wont	2.725e+00	1.191e+05	0.000	1.000
date	1.968e+01	1.030e+05	0.000	1.000
sunday	-2.497e+01	2.883e+05	0.000	1.000
click	-7.138e+01	3.715e+05	0.000	1.000
credit	3.439e+01	2.409e+05	0.000	1.000
link	1.129e+02	7.031e+05	0.000	1.000
messag	5.899e+00	5.138e+04	0.000	1.000
`next`	-2.712e+01	2.615e+05	0.000	1.000
use	-1.784e+01	9.940e+04	0.000	1.000
wap	3.829e+01	1.283e+05	0.000	1.000
watch	-2.170e+01	1.026e+05	0.000	1.000
make	-1.099e+01	6.736e+04	0.000	1.000
name	6.326e+01	7.738e+04	0.001	0.999
naughti	2.433e+01	9.440e+05	0.000	1.000
rememb	-2.648e+01	4.500e+04	-0.001	1.000
yes	-3.277e+01	4.130e+04	-0.001	0.999
feel	-3.724e+01	2.628e+05	0.000	1.000
fine	-1.988e+01	6.070e+04	0.000	1.000
way	-2.143e+01	6.320e+04	0.000	1.000
england	-4.603e+01	3.774e+05	0.000	1.000
miss	-1.467e+01	6.409e+03	-0.002	0.998
nation	5.064e+01	6.493e+05	0.000	1.000
news	1.339e+01	3.173e+05	0.000	1.000
team	6.861e+01	3.303e+05	0.000	1.000
serious	-8.178e+01	8.899e+04	-0.001	0.999
tri	-1.425e+01	1.122e+05	0.000	1.000
comin	-1.117e+01	4.232e+05	0.000	1.000
first	5.504e+00	2.297e+05	0.000	1.000
pay	7.285e+00	1.301e+05	0.000	1.000
stock	-4.819e+01	5.171e+05	0.000	1.000
aft	-3.746e+01	2.110e+05	0.000	1.000
ard	-1.213e+01	3.688e+05	0.000	1.000
finish	-3.261e+00	4.606e+04	0.000	1.000
lor	-1.881e+01	2.244e+04	-0.001	0.999
lunch	-1.932e+01	7.441e+04	0.000	1.000

smth	2.487e+01	5.587e+05	0.000	1.000
alright	-3.790e+01	8.265e+04	0.000	1.000
can	-2.205e+01	3.621e+04	-0.001	1.000
meet	-1.809e+01	3.496e+04	-0.001	1.000
eat	-1.085e+01	2.963e+05	0.000	1.000
get	-2.302e+01	4.955e+04	0.000	1.000
hungri	-6.696e+01	6.405e+05	0.000	1.000
just	-7.083e-01	1.249e+04	0.000	1.000
know	-2.094e+01	7.270e+04	0.000	1.000
lol	-1.900e+01	5.459e+04	0.000	1.000
mark	-4.604e+00	1.946e+05	0.000	1.000
pizza	-4.474e+00	4.607e+05	0.000	1.000
realli	5.293e+00	7.705e+04	0.000	1.000
sick	1.541e+00	1.278e+05	0.000	1.000
suck	4.120e+01	4.072e+05	0.000	1.000
tho	-2.548e+01	3.314e+05	0.000	1.000
turn	1.071e+01	2.797e+05	0.000	1.000
worri	6.571e+00	1.193e+05	0.000	1.000
alway	-1.315e+01	1.011e+05	0.000	1.000
bus	-1.888e+01	5.093e+04	0.000	1.000
catch	-3.350e+01	3.305e+05	0.000	1.000
dinner	-1.818e+01	3.662e+04	0.000	1.000
egg	2.008e+00	3.657e+05	0.000	1.000
fri	-2.930e+01	1.651e+05	0.000	1.000
left	8.879e+01	6.369e+04	0.001	0.999
love	-2.633e+01	7.859e+04	0.000	1.000
mom	8.255e+00	3.613e+05	0.000	1.000
tea	2.031e+00	1.450e+06	0.000	1.000
amp	-5.191e+01	2.513e+05	0.000	1.000
car	-1.871e+01	4.855e+04	0.000	1.000
ill	-2.805e+01	1.256e+05	0.000	1.000
let	1.051e+02	3.541e+04	0.003	0.998
room	-2.028e+01	1.001e+05	0.000	1.000
there	9.821e+00	3.679e+05	0.000	1.000
work	-1.971e+01	3.318e+04	-0.001	1.000
clear	-4.982e+00	3.276e+05	0.000	1.000
doesnt	-1.219e+01	5.017e+05	0.000	1.000
sure	-4.012e+00	7.490e+04	0.000	1.000
that	-2.739e+01	1.540e+05	0.000	1.000
wait	-1.851e+01	3.623e+04	-0.001	1.000
bad	-5.648e+01	2.598e+05	0.000	1.000
cheer	-3.019e+01	2.918e+05	0.000	1.000
till	6.106e+01	1.601e+05	0.000	1.000

yeah	-2.116e+01	1.542e+05	0.000	1.000
anyth	-3.049e+01	1.161e+05	0.000	1.000
tell	-1.644e+01	1.550e+04	-0.001	0.999
quick	-2.284e+01	4.716e+05	0.000	1.000
charg	1.595e+01	1.466e+05	0.000	1.000
confirm	1.683e+01	6.198e+04	0.000	1.000
pleas	1.507e+01	3.240e+04	0.000	1.000
rington	8.990e+01	2.578e+05	0.000	1.000
learn	-9.478e+01	1.492e+06	0.000	1.000
lesson	6.840e+01	3.326e+05	0.000	1.000
look	-1.808e+01	3.468e+04	-0.001	1.000
msg	1.885e+01	3.333e+04	0.001	1.000
yup	-2.375e+01	5.985e+04	0.000	1.000
done	-3.374e+01	1.234e+05	0.000	1.000
oop	-6.848e+01	5.756e+05	0.000	1.000
roommat	2.630e+01	5.317e+05	0.000	1.000
letter	3.110e+01	8.711e+04	0.000	1.000
see	-3.165e+01	3.618e+04	-0.001	0.999
decid	-1.703e+01	6.093e+05	0.000	1.000
hello	-1.849e+01	3.050e+04	-0.001	1.000
how	-3.765e+01	2.480e+05	0.000	1.000
invit	-5.611e-01	9.377e+04	0.000	1.000
saturday	2.092e+00	2.450e+05	0.000	1.000
tomo	-4.103e+01	5.309e+05	0.000	1.000
youd	1.287e+02	4.594e+05	0.000	1.000
abiola	2.549e+01	3.011e+05	0.000	1.000
ahead	-5.708e+01	1.140e+06	0.000	1.000
pls	-2.339e+01	8.427e+04	0.000	1.000
weekend	-4.180e+00	1.103e+05	0.000	1.000
crave	-5.588e+01	2.823e+05	0.000	1.000
forget	1.401e+02	1.704e+05	0.001	0.999
need	-1.836e+01	3.878e+04	0.000	1.000
sweet	-2.261e+01	1.869e+05	0.000	1.000
camcord	5.068e+01	9.932e+05	0.000	1.000
deliveri	6.149e+01	3.402e+05	0.000	1.000
nokia	-4.523e+01	1.979e+05	0.000	1.000
sms	1.401e+01	8.751e+04	0.000	1.000
tomorrow	-4.828e+00	7.868e+04	0.000	1.000
hope	-1.617e+01	1.354e+04	-0.001	0.999
inch	1.142e+02	2.704e+05	0.000	1.000
ltgt	-1.837e+01	2.290e+04	-0.001	0.999
man	-1.445e+01	1.727e+05	0.000	1.000
well	-1.955e+01	4.036e+04	0.000	1.000

didnt	-1.960e+01	2.557e+05	0.000	1.000
nigeria	2.213e+01	2.840e+05	0.000	1.000
ask	-1.982e+01	6.696e+04	0.000	1.000
bit	2.728e+01	2.634e+05	0.000	1.000
cant	-5.339e-01	4.890e+04	0.000	1.000
mayb	-2.036e+01	7.573e+04	0.000	1.000
tyler	3.333e+01	6.613e+05	0.000	1.000
hospit	-1.060e+02	1.093e+05	-0.001	0.999
kept	1.087e+02	9.625e+05	0.000	1.000
weak	2.312e+01	1.794e+07	0.000	1.000
class	-4.162e+01	8.504e+04	0.000	1.000
saw	7.067e+00	1.716e+05	0.000	1.000
almost	-2.558e+01	3.930e+05	0.000	1.000
half	1.996e+01	1.543e+05	0.000	1.000
run	-2.217e+00	1.388e+05	0.000	1.000
second	-6.782e+00	1.664e+05	0.000	1.000
usual	-1.753e+01	2.543e+05	0.000	1.000
whole	6.284e+01	4.947e+05	0.000	1.000
fyi	4.243e+01	4.420e+05	0.000	1.000
hes	3.444e+00	1.563e+05	0.000	1.000
morn	-3.552e+01	1.486e+05	0.000	1.000
place	8.304e-01	7.052e+04	0.000	1.000
best	3.245e+01	1.883e+05	0.000	1.000
give	-1.049e+01	6.015e+04	0.000	1.000
happi	-1.864e+01	4.148e+04	0.000	1.000
never	-7.593e-04	7.182e+04	0.000	1.000
offer	-5.671e+00	1.125e+05	0.000	1.000
realiz	-2.957e+02	6.873e+05	0.000	1.000
seem	-1.798e+01	2.363e+04	-0.001	0.999
sinc	-2.776e+01	2.013e+05	0.000	1.000
sorri	-2.007e+01	3.172e+04	-0.001	0.999
thought	-2.048e+01	1.601e+05	0.000	1.000
wow	-3.686e+01	1.939e+05	0.000	1.000
correct	-4.090e+01	1.161e+05	0.000	1.000
end	3.207e+00	1.180e+05	0.000	1.000
ice	1.367e+01	2.940e+05	0.000	1.000
new	2.539e+01	3.349e+04	0.001	0.999
play	2.486e+01	1.849e+05	0.000	1.000
find	2.637e+01	1.568e+05	0.000	1.000
yesterday	-1.754e+01	1.784e+05	0.000	1.000
cinema	8.234e+01	2.740e+05	0.000	1.000
congrat	3.359e+01	8.747e+04	0.000	1.000
etc	-1.343e+00	1.905e+05	0.000	1.000

pass	6.683e+00	3.291e+05	0.000	1.000
special	-1.782e+01	2.175e+04	-0.001	0.999
year	-2.481e+01	5.724e+04	0.000	1.000
later	-1.960e+01	5.234e+04	0.000	1.000
reach	-2.081e+01	5.259e+05	0.000	1.000
kill	2.807e+01	6.823e+05	0.000	1.000
move	-2.296e+01	1.540e+05	0.000	1.000
pain	-1.800e+01	3.385e+04	-0.001	1.000
pick	-1.863e+01	3.202e+04	-0.001	1.000
girl	2.416e-01	4.698e+04	0.000	1.000
good	-1.209e+02	3.708e+04	-0.003	0.997
situat	1.861e+01	8.501e+05	0.000	1.000
check	-2.356e+01	1.664e+05	0.000	1.000
part	9.043e-01	1.164e+05	0.000	1.000
come	-3.437e+01	3.458e+04	-0.001	0.999
forev	3.218e+01	4.810e+05	0.000	1.000
took	-1.877e+01	4.985e+04	0.000	1.000
cut	-1.517e+01	3.235e+05	0.000	1.000
doubl	7.173e+01	3.149e+05	0.000	1.000
hair	-1.336e+01	2.807e+05	0.000	1.000
nice	-3.857e+00	2.442e+05	0.000	1.000
said	-3.150e+01	2.790e+05	0.000	1.000
short	3.114e+01	2.562e+05	0.000	1.000
wun	1.559e+01	5.551e+05	0.000	1.000
award	3.315e+01	2.472e+05	0.000	1.000
bonus	-4.652e+01	2.883e+05	0.000	1.000
follow	-8.345e+01	2.874e+05	0.000	1.000
mob	4.418e+01	7.349e+05	0.000	1.000
recent	-1.776e+01	1.969e+05	0.000	1.000
review	4.011e+01	1.380e+06	0.000	1.000
frnds	1.081e+02	9.075e+05	0.000	1.000
rpli	-4.673e+01	1.583e+06	0.000	1.000
song	-2.892e+01	8.962e+05	0.000	1.000
complimentari	-8.788e+01	1.105e+06	0.000	1.000
dis	-7.027e+00	2.182e+05	0.000	1.000
trip	-1.074e-01	3.059e+05	0.000	1.000
hear	5.034e+01	5.323e+04	0.001	0.999
hee	-2.046e+01	1.126e+05	0.000	1.000
lucki	2.729e+01	4.684e+05	0.000	1.000
money	-5.674e+00	3.013e+05	0.000	1.000
save	1.413e+01	3.518e+05	0.000	1.000
babe	-1.900e+01	3.083e+04	-0.001	1.000
someth	-1.872e+01	4.796e+04	0.000	1.000

wanna	-1.632e+01	1.457e+04	-0.001	0.999
cool	-3.518e+01	1.345e+05	0.000	1.000
much	-2.732e+01	9.691e+04	0.000	1.000
peopl	-5.331e+01	5.922e+04	-0.001	0.999
oper	-3.540e+01	4.285e+05	0.000	1.000
job	-3.715e+01	1.760e+05	0.000	1.000
stop	1.626e+01	9.133e+04	0.000	1.000
one	-1.681e+01	1.313e+04	-0.001	0.999
real	6.076e+01	1.784e+05	0.000	1.000
ticket	-3.522e+01	1.874e+05	0.000	1.000
bed	-4.610e+00	8.009e+04	0.000	1.000
came	-1.837e+01	2.585e+04	-0.001	0.999
gotta	-1.910e+01	5.879e+04	0.000	1.000
start	-1.629e+01	1.433e+04	-0.001	0.999
download	-1.717e+01	3.406e+04	-0.001	1.000
wen	-4.705e+00	2.958e+05	0.000	1.000
close	1.562e+01	8.197e+04	0.000	1.000
stand	-2.070e+01	1.301e+05	0.000	1.000
`\u0089\u00b1ll`	-4.299e+01	1.485e+06	0.000	1.000
afternoon	-1.764e+01	3.019e+05	0.000	1.000
anoth	-1.427e+01	5.276e+05	0.000	1.000
havent	-1.811e+01	3.577e+04	-0.001	1.000
includ	-6.968e+01	7.652e+05	0.000	1.000
late	-1.574e+01	2.026e+05	0.000	1.000
mean	-2.664e+01	1.626e+05	0.000	1.000
night	1.118e+00	1.283e+05	0.000	1.000
spent	-2.911e+01	1.941e+05	0.000	1.000
hurt	-8.515e+00	2.159e+05	0.000	1.000
pleasur	1.049e+02	1.525e+06	0.000	1.000
rain	4.899e+01	1.887e+05	0.000	1.000
smile	-4.598e+00	2.578e+05	0.000	1.000
someon	-9.501e-01	2.877e+05	0.000	1.000
troubl	4.038e+01	2.492e+05	0.000	1.000
ampm	2.149e+01	4.828e+05	0.000	1.000
guarante	1.045e+01	4.393e+05	0.000	1.000
repres	-3.808e+02	9.489e+05	0.000	1.000
servic	5.680e+01	3.639e+04	0.002	0.999
buy	-1.667e+01	1.725e+04	-0.001	0.999
plan	-2.001e+01	5.217e+04	0.000	1.000
show	-1.939e+01	4.713e+04	0.000	1.000
collect	-2.628e+01	3.283e+05	0.000	1.000
password	-3.597e+00	1.483e+05	0.000	1.000
simpli	5.072e+00	7.218e+05	0.000	1.000

verifi	8.772e+01	3.694e+05	0.000	1.000
abt	-4.181e+01	1.726e+05	0.000	1.000
load	-5.615e+01	5.807e+05	0.000	1.000
loan	3.671e+00	3.267e+05	0.000	1.000
caus	-3.846e+00	3.010e+05	0.000	1.000
forgot	9.853e+00	2.317e+05	0.000	1.000
hol	9.997e+01	5.579e+05	0.000	1.000
prob	-2.040e+01	6.436e+04	0.000	1.000
shower	-3.008e+01	2.231e+05	0.000	1.000
els	-4.098e+00	2.537e+05	0.000	1.000
noth	-2.020e+01	6.974e+04	0.000	1.000
ave	1.843e+01	4.244e+05	0.000	1.000
long	-1.854e+01	4.228e+04	0.000	1.000
okay	-2.096e+01	2.422e+05	0.000	1.000
price	-1.788e+01	1.389e+05	0.000	1.000
drive	-3.416e+01	1.632e+05	0.000	1.000
gone	1.238e+00	1.039e+06	0.000	1.000
test	1.220e+01	4.995e+05	0.000	1.000
yet	-1.966e+01	7.224e+04	0.000	1.000
chang	-1.909e+01	5.717e+04	0.000	1.000
cuz	7.006e-01	1.086e+05	0.000	1.000
gave	-1.379e+01	7.706e+05	0.000	1.000
guess	8.890e-01	5.141e+04	0.000	1.000
locat	3.319e+01	9.666e+05	0.000	1.000
men	1.140e+02	7.547e+05	0.000	1.000
page	-3.629e+01	1.974e+05	0.000	1.000
your	-6.472e+01	7.374e+04	-0.001	0.999
dear	-3.336e+01	8.481e+04	0.000	1.000
life	-1.709e+01	2.128e+04	-0.001	0.999
lot	5.734e+00	9.470e+04	0.000	1.000
umma	5.873e-01	8.097e+05	0.000	1.000
birthday	-5.844e+00	1.146e+05	0.000	1.000
wish	-3.614e+01	2.947e+05	0.000	1.000
aight	-3.948e+01	1.540e+05	0.000	1.000
hit	2.522e+01	1.750e+05	0.000	1.000
address	-4.464e+01	7.159e+04	-0.001	1.000
comput	-9.654e+00	1.438e+05	0.000	1.000
consid	-7.141e+01	1.106e+06	0.000	1.000
isnt	-1.755e+01	2.703e+04	-0.001	0.999
better	4.947e+00	3.151e+05	0.000	1.000
lie	6.650e+01	3.723e+05	0.000	1.000
old	6.489e+00	2.435e+05	0.000	1.000
busi	-1.193e+01	6.221e+04	0.000	1.000

cos	6.991e+00	1.277e+05	0.000	1.000
mah	-1.620e+01	9.093e+05	0.000	1.000
thing	-6.027e+00	1.244e+05	0.000	1.000
contact	-2.980e+01	2.965e+05	0.000	1.000
draw	-2.121e+01	3.881e+05	0.000	1.000
hrs	-6.872e+01	3.441e+05	0.000	1.000
last	7.083e+00	1.309e+05	0.000	1.000
ppm	-3.990e+01	4.468e+05	0.000	1.000
anyway	1.632e+01	1.747e+05	0.000	1.000
eatin	-8.964e-01	1.251e+05	0.000	1.000
happen	-3.792e+01	1.096e+05	0.000	1.000
juz	1.018e+01	1.972e+05	0.000	1.000
apart	3.355e+01	2.212e+05	0.000	1.000
askd	-9.618e-01	2.418e+05	0.000	1.000
bday	-8.104e+01	2.044e+05	0.000	1.000
boss	-1.949e+01	7.068e+04	0.000	1.000
enter	1.389e+01	2.062e+05	0.000	1.000
felt	1.613e+02	3.754e+05	0.000	1.000
went	-2.251e+01	1.482e+05	0.000	1.000
flight	-4.881e+01	1.036e+06	0.000	1.000
holiday	1.462e+01	3.199e+05	0.000	1.000
inc	8.495e+01	5.498e+05	0.000	1.000
pmin	-9.005e+00	4.660e+05	0.000	1.000
friday	6.765e+01	2.590e+05	0.000	1.000
must	-3.830e+01	2.462e+05	0.000	1.000
direct	2.070e+01	2.038e+05	0.000	1.000
food	-3.930e+01	4.285e+05	0.000	1.000
inform	3.402e+00	3.548e+05	0.000	1.000
school	-1.699e+01	1.707e+05	0.000	1.000
uncl	-1.109e+01	2.990e+05	0.000	1.000
account	3.510e+00	1.375e+05	0.000	1.000
expir	1.909e+02	1.062e+06	0.000	1.000
identifi	-1.406e+02	4.746e+07	0.000	1.000
privat	-4.588e+01	3.048e+05	0.000	1.000
statement	8.133e+01	4.745e+07	0.000	1.000
unredeem	NA	NA	NA	NA
landlin	1.286e+02	2.849e+05	0.000	1.000
app	-3.116e+00	1.117e+05	0.000	1.000
match	-4.011e+00	6.894e+04	0.000	1.000
number	-2.439e+01	3.842e+04	-0.001	0.999
quot	-3.362e+01	2.139e+06	0.000	1.000
rate	5.085e+01	4.841e+05	0.000	1.000
standard	7.613e+01	2.354e+06	0.000	1.000

ii	-3.781e+01	2.141e+05	0.000	1.000
bother	3.366e+01	4.711e+05	0.000	1.000
sent	-1.836e+01	3.971e+04	0.000	1.000
bak	-1.267e+01	2.725e+05	0.000	1.000
del	-1.798e+02	1.489e+06	0.000	1.000
answer	-6.298e+01	1.897e+05	0.000	1.000
question	3.572e+01	1.193e+05	0.000	1.000
countri	-3.165e+02	1.175e+06	0.000	1.000
dvd	-4.466e+01	4.434e+05	0.000	1.000
player	-1.974e+01	8.088e+04	0.000	1.000
quiz	-1.133e+02	2.814e+05	0.000	1.000
soni	6.367e+01	9.267e+05	0.000	1.000
sunshin	1.493e+02	1.495e+06	0.000	1.000
top	-2.658e+01	2.719e+05	0.000	1.000
dog	-4.021e+01	4.212e+05	0.000	1.000
join	-1.386e+01	9.182e+04	0.000	1.000
txting	1.467e+01	1.781e+05	0.000	1.000
haf	-1.879e+01	5.008e+04	0.000	1.000
activ	-8.242e+01	5.108e+05	0.000	1.000
age	-1.866e+01	4.695e+04	0.000	1.000
chat	-3.065e+01	8.808e+04	0.000	1.000
msgs	5.548e+00	2.129e+05	0.000	1.000
youll	-4.676e+01	1.216e+05	0.000	1.000
lazi	6.247e+01	3.019e+05	0.000	1.000
lect	-2.336e+01	6.232e+05	0.000	1.000
type	-3.118e+00	4.796e+05	0.000	1.000
mail	-5.732e+01	1.841e+05	0.000	1.000
sir	-1.748e+01	2.617e+04	-0.001	0.999
gud	-9.038e+01	1.982e+05	0.000	1.000
heart	-2.003e+01	5.389e+04	0.000	1.000
littl	2.019e+01	1.564e+05	0.000	1.000
lovabl	2.606e+02	1.477e+06	0.000	1.000
person	-4.392e+01	7.608e+04	-0.001	1.000
tire	2.839e+01	1.022e+06	0.000	1.000
open	-2.023e+01	7.170e+04	0.000	1.000
what	-3.897e+01	1.254e+05	0.000	1.000
local	-5.930e+01	1.004e+06	0.000	1.000
ltd	1.551e+02	2.512e+05	0.001	1.000
luv	3.517e+01	1.141e+05	0.000	1.000
sexi	7.809e+01	1.893e+05	0.000	1.000
begin	4.535e+01	5.561e+05	0.000	1.000
hard	9.935e-01	7.806e+05	0.000	1.000
pray	-8.268e+01	3.556e+05	0.000	1.000

delet	-4.072e+01	9.940e+04	0.000	1.000
wine	-1.925e+01	4.344e+04	0.000	1.000
thk	-1.044e+01	1.836e+05	0.000	1.000
becom	9.824e+01	2.340e+05	0.000	1.000
dream	6.653e+01	2.186e+05	0.000	1.000
joy	-2.749e+01	4.005e+05	0.000	1.000
shirt	3.114e+01	1.795e+06	0.000	1.000
sometim	3.224e+01	6.262e+05	0.000	1.000
without	3.512e+00	1.224e+05	0.000	1.000
hous	-2.825e+01	1.932e+05	0.000	1.000
leav	-2.539e+01	1.504e+05	0.000	1.000
boy	8.176e+00	4.506e+04	0.000	1.000
arrang	1.521e+01	1.915e+06	0.000	1.000
everyon	6.561e+00	3.649e+05	0.000	1.000
keep	-1.748e+01	2.593e+04	-0.001	0.999
safe	1.511e+01	3.269e+05	0.000	1.000
hand	1.223e+01	1.542e+05	0.000	1.000
excit	1.537e+01	9.759e+05	0.000	1.000
spend	-1.123e+01	3.159e+05	0.000	1.000
frnd	9.985e+01	3.091e+05	0.000	1.000
content	7.237e+01	1.052e+05	0.001	0.999
goto	-7.731e+01	5.578e+05	0.000	1.000
order	5.028e+01	3.413e+05	0.000	1.000
wit	-9.715e+01	9.102e+05	0.000	1.000
fanci	7.005e+01	3.141e+05	0.000	1.000
also	-3.026e+01	2.485e+05	0.000	1.000
complet	-9.837e+00	1.053e+06	0.000	1.000
wast	-1.512e+01	2.206e+05	0.000	1.000
bank	-9.721e+00	3.408e+05	0.000	1.000
discuss	-2.067e+01	1.290e+05	0.000	1.000
hmmm	-3.519e+01	5.058e+05	0.000	1.000
liao	-1.573e+01	6.887e+04	0.000	1.000
muz	2.464e+01	1.925e+05	0.000	1.000
believ	1.599e+01	2.291e+05	0.000	1.000
hell	-5.217e+01	3.341e+05	0.000	1.000
bath	1.802e+01	1.399e+05	0.000	1.000
carlo	2.636e+01	1.873e+05	0.000	1.000
youv	1.030e+01	1.274e+05	0.000	1.000
smoke	1.813e+01	1.718e+05	0.000	1.000
stay	-2.052e+01	1.064e+05	0.000	1.000
til	1.659e+01	1.861e+05	0.000	1.000
worth	-7.652e+01	6.546e+05	0.000	1.000
log	-1.001e+02	5.193e+05	0.000	1.000

experi	5.215e+01	6.691e+05	0.000	1.000
spoke	1.323e+01	7.259e+05	0.000	1.000
wed	-2.005e+00	2.989e+05	0.000	1.000
especi	-8.311e+01	8.709e+05	0.000	1.000
lift	4.055e+01	1.846e+05	0.000	1.000
studi	-1.052e+02	9.740e+04	-0.001	0.999
trust	-7.139e+00	9.747e+04	0.000	1.000
guy	-1.824e+01	3.776e+04	0.000	1.000
boytoy	-2.177e+01	2.218e+05	0.000	1.000
net	4.464e+01	5.712e+05	0.000	1.000
toward	-4.578e+01	7.836e+05	0.000	1.000
where	3.115e+01	4.005e+05	0.000	1.000
awesom	-1.597e+01	2.419e+05	0.000	1.000
haha	-7.699e+01	2.183e+05	0.000	1.000
minut	-2.739e+00	2.371e+05	0.000	1.000
freephon	5.070e+01	1.211e+05	0.000	1.000
xmas	6.117e+01	1.202e+05	0.001	1.000
jus	4.220e+00	1.455e+05	0.000	1.000
sis	-1.950e+01	6.897e+04	0.000	1.000
deal	3.969e+00	4.618e+05	0.000	1.000
touch	-2.248e+01	1.376e+05	0.000	1.000
cours	-8.328e+00	9.336e+04	0.000	1.000
abl	-1.350e+01	2.790e+05	0.000	1.000
everi	9.718e+00	3.282e+05	0.000	1.000
howev	3.130e+01	6.094e+05	0.000	1.000
settl	-1.170e+02	4.042e+05	0.000	1.000
hav	-8.041e+01	5.473e+05	0.000	1.000
mrng	9.955e+01	3.630e+05	0.000	1.000
stori	4.436e+01	5.355e+04	0.001	0.999
dead	8.078e+01	2.937e+05	0.000	1.000
mrt	3.248e+01	1.900e+05	0.000	1.000
orchard	-9.763e+01	1.209e+05	-0.001	0.999
tmr	-2.225e+01	4.084e+05	0.000	1.000
kate	-1.692e+01	2.118e+05	0.000	1.000
found	-2.230e+00	1.219e+05	0.000	1.000
buck	7.251e+00	8.005e+05	0.000	1.000
colleg	-1.027e+02	6.253e+05	0.000	1.000
darlin	-2.485e+01	1.018e+05	0.000	1.000
balanc	4.388e+00	1.135e+06	0.000	1.000
ltdecimalgt	-6.028e+00	3.207e+05	0.000	1.000
success	-3.653e+00	1.301e+06	0.000	1.000
goodmorn	1.409e+01	2.009e+05	0.000	1.000
sleep	-3.623e+01	1.073e+05	0.000	1.000

dat	1.868e+01	2.050e+05	0.000	1.000
oredi	2.016e+00	2.302e+05	0.000	1.000
oso	3.451e+01	3.507e+05	0.000	1.000
bill	-1.233e+02	5.176e+04	-0.002	0.998
connect	1.225e+02	3.038e+05	0.000	1.000
big	-1.849e+01	4.272e+04	0.000	1.000
readi	-1.616e+01	2.482e+05	0.000	1.000
`break`	1.374e+01	5.267e+04	0.000	1.000
semest	1.134e+01	1.781e+05	0.000	1.000
leh	2.272e+01	2.004e+05	0.000	1.000
noe	-7.011e+01	2.227e+05	0.000	1.000
sound	-2.322e+00	1.663e+05	0.000	1.000
head	-6.251e+01	1.744e+05	0.000	1.000
past	2.614e+01	9.693e+05	0.000	1.000
slept	-3.936e+01	2.109e+05	0.000	1.000
easi	3.178e+00	6.334e+05	0.000	1.000
exam	-3.252e+01	1.980e+05	0.000	1.000
march	-4.161e+01	2.291e+05	0.000	1.000
regist	-3.231e+00	4.569e+05	0.000	1.000
file	2.152e+01	2.480e+05	0.000	1.000
import	-1.069e+01	1.708e+05	0.000	1.000
shop	3.077e+00	1.600e+05	0.000	1.000
system	-5.580e+01	2.141e+05	0.000	1.000
nite	-1.728e+01	2.358e+04	-0.001	0.999
optout	-3.323e+01	4.378e+05	0.000	1.000
appreci	1.863e+01	6.101e+05	0.000	1.000
partner	1.379e+02	2.890e+05	0.000	1.000
sign	1.065e+02	4.511e+05	0.000	1.000
star	-4.944e+01	9.966e+05	0.000	1.000
compani	-1.819e+01	1.285e+05	0.000	1.000
bcoz	1.979e+01	3.364e+05	0.000	1.000
teach	7.618e-01	1.693e+05	0.000	1.000
road	-4.231e+01	2.037e+05	0.000	1.000
side	-1.399e+01	1.637e+05	0.000	1.000
street	1.042e+00	1.764e+05	0.000	1.000
walk	-2.688e+01	5.200e+04	-0.001	1.000
batteri	-7.649e+00	2.998e+05	0.000	1.000
die	-9.335e+00	3.680e+04	0.000	1.000
flirt	3.489e+01	2.796e+05	0.000	1.000
wil	-4.603e+01	3.176e+05	0.000	1.000
argument	-8.374e+01	8.617e+05	0.000	1.000
kick	-2.482e+01	2.276e+05	0.000	1.000
lose	-4.113e+01	2.066e+05	0.000	1.000

admir	9.736e+00	1.182e+06	0.000	1.000
rreveal	1.844e+02	1.268e+06	0.000	1.000
secret	-9.056e+01	7.529e+05	0.000	1.000
specialcal	NA	NA	NA	NA
ufind	NA	NA	NA	NA
case	-7.552e+00	2.597e+05	0.000	1.000
laptop	-1.745e+01	6.827e+04	0.000	1.000
tel	-7.214e+00	3.467e+05	0.000	1.000
meant	-9.137e+01	1.014e+05	-0.001	0.999
everyth	-8.999e+00	2.328e+05	0.000	1.000
face	-2.448e+01	1.546e+05	0.000	1.000
thanx	-2.579e+01	9.863e+04	0.000	1.000
told	-5.932e+01	1.364e+05	0.000	1.000
uve	2.357e+02	1.129e+06	0.000	1.000
kalli	-3.266e+01	5.128e+05	0.000	1.000
fix	-5.299e+00	1.273e+05	0.000	1.000
goodnight	-7.316e+01	3.681e+05	0.000	1.000
wake	-2.068e+01	1.281e+05	0.000	1.000
congratul	-5.236e+01	5.086e+05	0.000	1.000
gift	1.174e+01	2.768e+05	0.000	1.000
music	-4.556e+01	3.802e+05	0.000	1.000
tncs	-1.085e+02	5.128e+05	0.000	1.000
voucher	-3.426e+01	6.231e+05	0.000	1.000
cal	2.975e-01	2.121e+05	0.000	1.000
hold	-2.144e+01	1.845e+05	0.000	1.000
min	5.677e-01	5.754e+04	0.000	1.000
angri	-1.112e+01	8.663e+05	0.000	1.000
care	-1.302e+01	8.438e+04	0.000	1.000
childish	-6.157e+01	7.285e+05	0.000	1.000
coz	-3.286e-01	9.404e+04	0.000	1.000
deep	4.908e+01	4.077e+05	0.000	1.000
dnt	-6.820e+01	8.270e+05	0.000	1.000
true	-8.231e+01	9.567e+04	-0.001	0.999
wid	4.911e+00	1.723e+05	0.000	1.000
ship	-3.349e+01	5.430e+05	0.000	1.000
lemm	6.511e+01	1.611e+05	0.000	1.000
expect	4.227e+01	1.177e+05	0.000	1.000
lover	-2.202e+01	1.603e+06	0.000	1.000
anytim	-1.958e+01	7.489e+04	0.000	1.000
handset	5.408e+01	8.970e+05	0.000	1.000
unlimit	3.627e+01	3.508e+05	0.000	1.000
video	5.953e+01	2.269e+05	0.000	1.000
park	-1.076e+02	8.202e+04	-0.001	0.999

disturb	-1.883e+01	5.111e+04	0.000	1.000
ring	-6.424e+01	2.632e+05	0.000	1.000
horni	3.264e+01	5.303e+05	0.000	1.000
hot	9.270e+01	1.667e+05	0.001	1.000
nake	-8.611e+01	1.853e+06	0.000	1.000
unsubscrib	-3.689e+00	4.216e+05	0.000	1.000
dint	1.882e+00	1.853e+05	0.000	1.000
wana	-2.381e+00	3.304e+05	0.000	1.000
box	-4.129e+00	2.295e+05	0.000	1.000
choos	-7.186e+01	4.385e+05	0.000	1.000
club	-1.018e+01	2.576e+05	0.000	1.000
qualiti	7.242e+01	4.469e+05	0.000	1.000
singl	1.684e+00	6.577e+05	0.000	1.000
sort	-1.485e+01	3.586e+05	0.000	1.000
blue	1.539e+01	9.469e+05	0.000	1.000
sunni	-2.650e+00	3.216e+05	0.000	1.000
might	-8.440e+01	1.005e+05	-0.001	0.999
celebr	-4.003e+01	5.474e+05	0.000	1.000
full	-1.957e+01	7.365e+04	0.000	1.000
swing	-7.880e+01	7.070e+05	0.000	1.000
far	-4.396e+01	1.053e+06	0.000	1.000
oki	-2.139e+01	8.215e+04	0.000	1.000
anybodi	4.675e+01	4.277e+05	0.000	1.000
throw	8.608e+00	5.748e+05	0.000	1.000
babi	-1.832e+01	3.869e+04	0.000	1.000
fone	5.590e+00	9.040e+04	0.000	1.000
sens	-1.874e+01	4.915e+04	0.000	1.000
stupid	-1.341e+01	3.036e+05	0.000	1.000
card	-1.811e+01	5.161e+05	0.000	1.000
loyalti	1.890e+02	7.362e+05	0.000	1.000
phone	-1.541e+01	9.258e+03	-0.002	0.999
sim	-2.077e+00	2.194e+05	0.000	1.000
unless	-3.849e+01	8.544e+05	0.000	1.000
bslvyl	5.839e+00	4.807e+05	0.000	1.000
plz	-2.635e+00	6.518e+04	0.000	1.000
rose	-1.656e+01	1.311e+06	0.000	1.000
teas	1.912e+01	1.515e+06	0.000	1.000
coffe	-2.107e+01	1.133e+05	0.000	1.000
high	3.275e+00	5.123e+05	0.000	1.000
somebodi	-3.508e+01	3.142e+05	0.000	1.000
imagin	2.067e+01	4.443e+05	0.000	1.000
shit	-2.770e+01	2.614e+05	0.000	1.000
somewher	-2.519e+01	2.919e+05	0.000	1.000

book	1.715e+01	1.844e+05	0.000	1.000
friendship	-2.125e-01	1.849e+05	0.000	1.000
hang	-1.491e+02	2.802e+05	-0.001	1.000
game	3.516e+01	4.726e+04	0.001	0.999
sport	-1.879e+01	5.006e+04	0.000	1.000
tone	3.661e+01	2.836e+05	0.000	1.000
arent	-2.364e+02	5.459e+05	0.000	1.000
key	-9.115e+00	2.275e+05	0.000	1.000
accept	5.486e+01	1.666e+05	0.000	1.000
sister	-1.140e+02	3.012e+05	0.000	1.000
definit	1.111e+02	2.609e+05	0.000	1.000
lost	-3.173e+00	1.136e+05	0.000	1.000
normal	-2.019e+01	9.857e+04	0.000	1.000
rest	-1.913e+01	3.575e+05	0.000	1.000
wot	-1.961e+01	5.219e+04	0.000	1.000
made	-1.841e+01	3.947e+04	0.000	1.000
dunno	1.003e+00	8.253e+04	0.000	1.000
power	-2.010e+01	9.666e+04	0.000	1.000
yoga	5.206e+00	1.048e+06	0.000	1.000
dude	-9.417e+00	1.148e+05	0.000	1.000
mths	-9.670e+00	1.793e+05	0.000	1.000
christma	2.105e+01	3.693e+05	0.000	1.000
kiss	4.530e+01	1.211e+05	0.000	1.000
merri	-9.393e+01	5.491e+05	0.000	1.000
pete	1.605e+01	8.196e+05	0.000	1.000
problem	-4.122e+01	2.922e+05	0.000	1.000
read	7.045e+01	1.840e+05	0.000	1.000
record	-5.019e+01	8.481e+05	0.000	1.000
track	-6.649e+01	7.816e+05	0.000	1.000
light	-4.642e+01	5.020e+05	0.000	1.000
movi	-1.622e+01	1.391e+04	-0.001	0.999
immedi	3.173e+01	5.628e+05	0.000	1.000
return	3.120e+01	5.539e+05	0.000	1.000
access	3.439e+01	1.562e+05	0.000	1.000
line	2.385e-01	3.536e+05	0.000	1.000
via	5.141e-02	6.213e+04	0.000	1.000
custcar	-1.578e+02	5.314e+05	0.000	1.000
pmsg	3.434e+01	4.003e+05	0.000	1.000
rcvd	1.167e+02	7.625e+05	0.000	1.000
valentin	-5.101e+01	3.966e+05	0.000	1.000
surf	-2.017e+01	1.437e+05	0.000	1.000
post	8.074e+01	1.316e+05	0.001	1.000
wiv	3.522e+01	3.720e+05	0.000	1.000

interest	-4.577e+00	1.804e+05	0.000	1.000
round	1.415e+01	1.084e+05	0.000	1.000
two	-5.165e+01	2.742e+05	0.000	1.000
yiju	-8.793e+00	6.333e+05	0.000	1.000
small	-1.819e+01	9.555e+04	0.000	1.000
txts	1.328e+01	3.904e+05	0.000	1.000
basic	5.307e+01	7.613e+05	0.000	1.000
ever	-1.734e+01	3.662e+04	0.000	1.000
fault	5.670e+00	3.308e+05	0.000	1.000
urself	1.668e+00	4.195e+05	0.000	1.000
figur	-1.997e-02	6.089e+05	0.000	1.000
jay	5.040e+01	1.447e+05	0.000	1.000
weed	-2.582e+01	3.439e+05	0.000	1.000
ago	-1.114e+00	7.514e+04	0.000	1.000
ish	-1.643e+01	4.996e+05	0.000	1.000
met	6.083e+01	6.589e+05	0.000	1.000
cashbal	2.091e+01	5.993e+05	0.000	1.000
cashin	NA	NA	NA	NA
current	-3.704e+00	3.263e+05	0.000	1.000
hgsuiteland	5.262e+00	4.821e+05	0.000	1.000
maxim	NA	NA	NA	NA
rowwjhl	NA	NA	NA	NA
moment	3.400e+01	6.378e+05	0.000	1.000
cold	-1.982e+01	1.880e+06	0.000	1.000
chikku	-5.733e+00	1.387e+05	0.000	1.000
forward	7.098e+00	8.156e+04	0.000	1.000
air	-8.590e+01	1.877e+05	0.000	1.000
breath	-5.168e+00	1.047e+06	0.000	1.000
bluetooth	3.421e+01	7.311e+05	0.000	1.000
mobileupd	-4.619e-01	8.021e+05	0.000	1.000
motorola	1.168e+02	8.201e+05	0.000	1.000
orang	9.716e+00	4.136e+05	0.000	1.000
discount	-1.260e+01	6.259e+06	0.000	1.000
cuddl	-8.169e+00	1.452e+06	0.000	1.000
woke	6.645e+00	7.013e+05	0.000	1.000
refer	-5.809e+01	5.447e+05	0.000	1.000
mei	5.159e+01	3.376e+05	0.000	1.000
seen	-2.950e+01	5.529e+05	0.000	1.000
sigh	1.737e+01	3.810e+05	0.000	1.000
bring	-4.425e+01	2.701e+05	0.000	1.000
mistak	4.352e+01	3.237e+05	0.000	1.000
project	2.129e+00	1.661e+05	0.000	1.000
guid	-6.788e+01	1.492e+05	0.000	1.000

quit	9.837e+00	3.109e+05	0.000	1.000
reason	-9.354e+01	6.799e+05	0.000	1.000
relax	1.428e+02	8.018e+05	0.000	1.000
slow	-2.057e+00	5.320e+04	0.000	1.000
coupl	4.066e+01	4.024e+05	0.000	1.000
rental	2.804e+01	8.415e+05	0.000	1.000
huh	-3.272e+01	4.436e+05	0.000	1.000
pilat	-3.185e+01	1.504e+06	0.000	1.000
sat	-7.139e+00	1.869e+05	0.000	1.000
offic	-3.684e+01	4.750e+04	-0.001	0.999
bout	2.738e+01	3.336e+05	0.000	1.000
actual	1.373e+01	2.028e+05	0.000	1.000
rock	-2.527e+01	1.175e+05	0.000	1.000
ass	-5.065e+01	7.000e+05	0.000	1.000
facebook	-6.526e+01	1.358e+05	0.000	1.000
pictur	4.901e+01	1.131e+06	0.000	1.000
put	2.881e+00	2.411e+05	0.000	1.000
god	1.653e+01	1.017e+05	0.000	1.000
india	-4.268e+01	4.340e+05	0.000	1.000
dai	-1.517e+00	8.954e+04	0.000	1.000
poli	-6.166e+00	3.168e+05	0.000	1.000
yep	-3.848e+01	2.268e+05	0.000	1.000
drink	-1.546e+01	4.666e+05	0.000	1.000
fullonsmscom	3.499e+01	4.650e+05	0.000	1.000
den	7.929e+00	1.310e+05	0.000	1.000
eve	5.764e+01	2.458e+05	0.000	1.000
yahoo	1.633e+01	2.857e+05	0.000	1.000
aha	-3.398e+00	9.570e+04	0.000	1.000
land	7.995e+00	1.075e+06	0.000	1.000
funni	-2.116e+00	5.851e+05	0.000	1.000
irrit	-1.455e+01	2.817e+05	0.000	1.000
voic	-1.452e+02	8.962e+05	0.000	1.000
shell	-2.146e+01	1.318e+05	0.000	1.000
mind	-1.556e+00	4.266e+04	0.000	1.000
vri	-4.234e+01	7.728e+05	0.000	1.000
wnt	4.104e+01	8.812e+05	0.000	1.000
fuck	-4.126e+01	2.302e+05	0.000	1.000
respond	-2.019e+01	1.009e+05	0.000	1.000
ldn	-6.543e+01	1.593e+06	0.000	1.000
vari	5.319e+01	1.795e+06	0.000	1.000
askin	1.746e+01	1.119e+06	0.000	1.000
dollar	2.205e+01	1.092e+06	0.000	1.000
ten	5.327e+01	5.805e+05	0.000	1.000

tough	-5.333e+01	1.028e+06	0.000	1.000
suppos	1.934e+00	1.365e+05	0.000	1.000
din	-1.609e+01	1.199e+05	0.000	1.000
group	-1.946e+00	1.621e+05	0.000	1.000
tot	-5.945e+00	9.893e+04	0.000	1.000
hmm	-4.428e+01	2.911e+05	0.000	1.000
kinda	1.443e+02	1.878e+05	0.001	0.999
detail	-2.030e+01	2.092e+05	0.000	1.000
previous	-1.485e+02	1.644e+06	0.000	1.000
welcom	5.763e+01	1.362e+05	0.000	1.000
beauti	-7.712e+01	6.389e+05	0.000	1.000
ure	-6.138e+01	2.636e+05	0.000	1.000
kind	-2.107e+00	1.804e+05	0.000	1.000
result	-1.395e+02	7.457e+05	0.000	1.000
ttyl	9.284e+01	3.429e+05	0.000	1.000
decis	-7.898e+00	7.791e+05	0.000	1.000
differ	8.025e-01	1.917e+05	0.000	1.000
thru	-5.248e+00	8.048e+05	0.000	1.000
tcs	8.793e+01	5.787e+05	0.000	1.000
enjoy	6.157e-01	3.938e+04	0.000	1.000
princess	-6.464e+01	2.813e+05	0.000	1.000
style	-2.138e+01	1.575e+05	0.000	1.000
mani	7.689e-01	1.963e+05	0.000	1.000
notic	1.035e+02	3.916e+05	0.000	1.000
sae	-1.527e+02	9.653e+05	0.000	1.000
tenerif	-1.580e+01	1.661e+05	0.000	1.000
scream	-5.211e+01	4.310e+05	0.000	1.000
remov	1.864e+01	4.626e+05	0.000	1.000
transfer	3.313e+01	2.365e+05	0.000	1.000
moan	7.014e+01	5.310e+05	0.000	1.000
nyt	1.137e+02	2.823e+05	0.000	1.000
cum	4.488e+01	1.433e+05	0.000	1.000
sec	1.133e+02	3.930e+05	0.000	1.000
condit	4.276e+01	3.312e+05	0.000	1.000
term	2.177e+00	2.423e+05	0.000	1.000
visit	6.931e+01	3.046e+05	0.000	1.000
depend	-2.048e+01	2.104e+05	0.000	1.000
meh	5.497e+01	2.756e+05	0.000	1.000
monday	-1.179e+02	3.751e+05	0.000	1.000
nope	-2.007e+01	9.137e+04	0.000	1.000
either	-1.028e+01	1.791e+05	0.000	1.000
fight	7.842e+01	3.757e+05	0.000	1.000
water	1.460e+00	2.109e+05	0.000	1.000

bore	3.954e+01	1.754e+05	0.000	1.000
outsid	-1.567e+01	9.567e+04	0.000	1.000
near	-1.945e+01	6.799e+04	0.000	1.000
rent	-1.939e+01	6.818e+04	0.000	1.000
charact	-1.603e+02	1.031e+06	0.000	1.000
matur	-7.948e+01	7.709e+05	0.000	1.000
opinion	3.923e+01	2.392e+05	0.000	1.000
silent	-2.208e+01	5.010e+05	0.000	1.000
simpl	-2.078e+01	1.308e+05	0.000	1.000
tat	-1.136e+02	3.590e+05	0.000	1.000
ipod	-4.073e+00	4.826e+05	0.000	1.000
pin	1.970e+01	1.384e+05	0.000	1.000
less	2.209e+00	5.957e+05	0.000	1.000
adult	5.183e+01	1.272e+06	0.000	1.000
children	5.985e+00	1.677e+06	0.000	1.000
shall	-5.271e+01	7.188e+04	-0.001	0.999
campus	6.247e+00	2.243e+05	0.000	1.000
attempt	-7.864e+01	7.068e+05	0.000	1.000
member	1.659e+01	3.570e+05	0.000	1.000
savamob	-4.757e+01	3.094e+06	0.000	1.000
sub	-1.252e+00	1.626e+05	0.000	1.000
unsub	-6.020e+01	2.940e+06	0.000	1.000
ladi	-1.891e+01	3.054e+05	0.000	1.000
pretti	5.714e-01	7.850e+04	0.000	1.000
within	1.109e+01	1.313e+06	0.000	1.000
across	-4.381e+01	9.268e+05	0.000	1.000
lead	2.490e+01	4.258e+05	0.000	1.000
sea	5.153e+01	8.052e+05	0.000	1.000
proabl	-4.126e+01	2.862e+05	0.000	1.000
fat	-3.966e+01	8.383e+05	0.000	1.000
finger	8.671e+01	6.421e+05	0.000	1.000
listen	9.368e+01	1.371e+05	0.001	0.999
marri	6.835e+01	8.470e+05	0.000	1.000
self	-5.423e+01	2.089e+06	0.000	1.000
mine	-2.027e+01	5.750e+05	0.000	1.000
rather	1.722e+00	2.152e+05	0.000	1.000
hotel	-3.605e+01	4.381e+05	0.000	1.000
omw	-1.264e+01	8.194e+04	0.000	1.000
remind	-1.586e+01	4.555e+05	0.000	1.000
hurri	5.440e+01	2.217e+05	0.000	1.000
warm	7.902e+01	2.634e+05	0.000	1.000
weight	5.249e+01	3.007e+05	0.000	1.000
cheap	1.163e+01	4.118e+05	0.000	1.000

`\u0089û÷s`	-1.120e+02	1.234e+05	-0.001	0.999
pic	3.373e+01	1.644e+05	0.000	1.000
aiyo	-5.593e+01	3.063e+05	0.000	1.000
fast	1.587e+01	1.927e+05	0.000	1.000
workin	-2.161e+01	1.695e+05	0.000	1.000
bare	-2.197e+01	6.254e+05	0.000	1.000
gym	-3.246e+01	3.541e+05	0.000	1.000
whatev	-5.194e+01	3.238e+05	0.000	1.000
daddi	3.123e+01	1.574e+06	0.000	1.000
clean	2.685e+01	1.360e+05	0.000	1.000
mum	-1.925e+01	6.065e+04	0.000	1.000
sch	-1.880e+01	2.726e+05	0.000	1.000
yar	5.905e+00	1.151e+05	0.000	1.000
door	-2.189e+01	1.768e+05	0.000	1.000
izzit	-1.700e+01	2.218e+05	0.000	1.000
kid	1.068e+01	1.515e+05	0.000	1.000
spree	-1.186e+02	3.743e+05	0.000	1.000
tscs	4.426e+01	5.057e+05	0.000	1.000
announc	-4.490e+01	3.712e+05	0.000	1.000
train	-2.079e+01	6.564e+04	0.000	1.000
present	-6.470e+01	9.741e+05	0.000	1.000
indian	1.583e+02	4.448e+05	0.000	1.000
noon	-1.185e+02	5.647e+05	0.000	1.000
neva	3.263e+00	1.507e+05	0.000	1.000
imma	6.642e+01	5.211e+05	0.000	1.000
daili	-1.034e+01	4.701e+05	0.000	1.000
euro	3.657e+01	6.429e+05	0.000	1.000
suit	2.365e+01	8.706e+05	0.000	1.000
futur	9.642e+01	3.361e+05	0.000	1.000
shuhui	-2.017e+01	5.722e+05	0.000	1.000
famili	-1.617e+00	1.077e+05	0.000	1.000
snow	-8.345e+01	2.473e+05	0.000	1.000
togeth	-1.289e+01	2.254e+05	0.000	1.000
weather	1.779e+01	2.618e+05	0.000	1.000
alex	-2.053e+01	3.382e+05	0.000	1.000
pub	-5.877e+01	3.907e+05	0.000	1.000
offici	-1.623e+02	1.190e+06	0.000	1.000
paid	2.375e+01	1.916e+05	0.000	1.000
pop	3.434e+00	2.353e+05	0.000	1.000
darren	8.549e+01	6.592e+05	0.000	1.000
longer	-3.928e+00	1.491e+05	0.000	1.000
area	5.173e+01	3.502e+05	0.000	1.000
vodafon	-8.995e+01	4.759e+05	0.000	1.000

`\u0089ü`	-4.067e+01	4.880e+05	0.000	1.000
holder	1.106e+02	1.106e+06	0.000	1.000
dad	-5.270e+01	1.840e+05	0.000	1.000
fetch	-1.073e+01	9.531e+04	0.000	1.000
login	2.623e+01	3.380e+05	0.000	1.000
email	-2.417e+01	5.936e+04	0.000	1.000
gentl	-1.092e+02	6.223e+05	0.000	1.000
tuesday	-8.173e+01	2.322e+05	0.000	1.000
drug	1.983e+01	8.701e+04	0.000	1.000
town	-1.391e+00	1.435e+05	0.000	1.000
evng	2.076e+01	6.699e+05	0.000	1.000
sun	-1.674e+01	1.800e+04	-0.001	0.999
paper	-1.227e+01	4.037e+05	0.000	1.000
prepar	-6.125e+01	4.141e+05	0.000	1.000
laugh	-1.263e+02	3.119e+05	0.000	1.000
opt	7.633e+00	3.321e+05	0.000	1.000
knw	-3.649e+01	2.452e+05	0.000	1.000
awak	5.792e+01	2.578e+05	0.000	1.000
bedroom	1.345e+01	4.776e+05	0.000	1.000
sex	1.357e+01	5.924e+05	0.000	1.000
king	5.684e+01	2.117e+05	0.000	1.000
bold	-2.512e+01	5.314e+05	0.000	1.000
torch	5.404e+00	2.502e+05	0.000	1.000
`\u0089ü`	3.905e+01	6.510e+05	0.000	1.000
idea	3.074e+01	1.735e+05	0.000	1.000
away	4.968e+01	1.003e+05	0.000	1.000
fantast	-1.202e+01	4.917e+05	0.000	1.000
aftr	2.747e+01	4.181e+05	0.000	1.000
sit	-5.571e+00	7.185e+05	0.000	1.000
student	-5.802e+01	5.379e+05	0.000	1.000
sell	1.282e+01	4.485e+05	0.000	1.000
exact	-1.956e+01	7.284e+04	0.000	1.000
yay	1.551e+01	2.550e+05	0.000	1.000
add	2.818e+01	2.393e+06	0.000	1.000
tonit	1.533e+00	1.477e+05	0.000	1.000
gal	2.525e+01	4.605e+05	0.000	1.000
rite	-2.193e+00	2.057e+05	0.000	1.000
glad	-3.828e+00	1.638e+05	0.000	1.000
eye	-6.520e+01	7.632e+05	0.000	1.000
subscrib	4.453e+01	7.897e+05	0.000	1.000
librari	2.665e+00	1.208e+05	0.000	1.000
holla	-6.923e-01	2.769e+05	0.000	1.000
anyon	6.878e+01	3.652e+05	0.000	1.000

yest	1.301e+02	4.287e+05	0.000	1.000
advanc	3.539e+01	3.596e+05	0.000	1.000
store	-4.096e+01	5.060e+05	0.000	1.000
summer	1.352e+01	3.636e+05	0.000	1.000
goin	-7.475e+00	3.249e+05	0.000	1.000
theyr	3.472e+01	7.546e+05	0.000	1.000
damn	-1.348e+02	6.471e+05	0.000	1.000
space	-7.852e+01	4.826e+05	0.000	1.000
sing	1.674e+02	1.034e+06	0.000	1.000
slave	4.217e+01	3.508e+05	0.000	1.000
shes	-1.889e+01	5.236e+04	0.000	1.000
omg	1.244e+00	2.548e+05	0.000	1.000
lttimegt	-3.943e+01	4.152e+05	0.000	1.000
mate	-1.232e+01	3.027e+05	0.000	1.000
normpton	3.666e+01	7.687e+05	0.000	1.000
wwq	NA	NA	NA	NA
wwwgetzedcouk	7.302e+01	1.289e+06	0.000	1.000
aint	-1.985e+01	8.506e+04	0.000	1.000
ugh	-3.647e+01	2.748e+05	0.000	1.000
bag	-2.263e+01	3.434e+05	0.000	1.000
model	1.909e+01	1.678e+05	0.000	1.000
boo	2.258e+01	3.937e+05	0.000	1.000
gettin	8.107e+01	2.113e+05	0.000	1.000
extra	-5.522e+00	6.674e+05	0.000	1.000
moon	1.242e+02	7.310e+05	0.000	1.000
zed	-8.233e+01	4.698e+05	0.000	1.000
reveal	7.161e+01	4.940e+05	0.000	1.000
skype	1.004e+02	2.281e+05	0.000	1.000
five	-6.095e+01	1.024e+06	0.000	1.000
understand	1.632e+01	1.358e+05	0.000	1.000
await	1.149e+02	6.268e+05	0.000	1.000
wan	-8.834e+01	7.674e+04	-0.001	0.999
poor	2.565e+01	2.544e+05	0.000	1.000
fall	4.541e+01	4.043e+05	0.000	1.000
tampa	-4.170e+01	3.378e+05	0.000	1.000
whenev	-1.025e+01	3.483e+05	0.000	1.000
yrs	5.584e+01	6.786e+05	0.000	1.000
asap	4.296e+01	6.710e+05	0.000	1.000
bird	6.647e+00	2.561e+05	0.000	1.000
user	9.836e+01	6.896e+05	0.000	1.000
drop	-1.980e+01	6.311e+04	0.000	1.000
otherwis	8.257e+00	9.810e+04	0.000	1.000
ntt	-1.897e+02	2.177e+06	0.000	1.000

travel	-1.206e+01	1.033e+06	0.000	1.000
mid	-8.192e+00	6.022e+05	0.000	1.000
knew	-1.283e+02	4.698e+05	0.000	1.000
midnight	-2.036e+01	1.101e+05	0.000	1.000
sup	-1.156e+01	1.888e+05	0.000	1.000
fail	-4.773e+01	7.207e+05	0.000	1.000
sad	-1.918e+01	5.960e+04	0.000	1.000
murder	1.402e+02	5.199e+05	0.000	1.000
due	1.356e+02	9.947e+05	0.000	1.000
possibl	-2.172e+01	1.185e+05	0.000	1.000
greet	3.261e+01	1.961e+05	0.000	1.000
plus	5.528e+00	3.802e+05	0.000	1.000
gay	-5.302e+01	9.602e+05	0.000	1.000
januari	-4.776e+01	2.611e+06	0.000	1.000
sale	3.431e+01	1.015e+05	0.000	1.000
piss	-5.689e+01	4.009e+05	0.000	1.000
wrong	1.921e+01	1.136e+05	0.000	1.000
photo	1.468e+01	1.806e+05	0.000	1.000
attend	-7.044e+01	6.828e+05	0.000	1.000
ate	-1.906e+00	1.344e+05	0.000	1.000
wherev	5.156e+01	5.344e+05	0.000	1.000
skxh	1.614e+02	2.406e+06	0.000	1.000
wear	-4.768e+01	1.851e+06	0.000	1.000
least	3.743e+01	2.353e+05	0.000	1.000
medic	1.093e+02	6.414e+05	0.000	1.000
score	2.865e+01	2.576e+05	0.000	1.000
heavi	4.041e+01	3.004e+05	0.000	1.000
truth	-1.023e+01	3.950e+05	0.000	1.000
brand	-1.891e+01	5.323e+04	0.000	1.000
wife	-2.169e+01	4.448e+05	0.000	1.000
earlier	-1.796e+01	3.520e+05	0.000	1.000
nxt	3.540e+01	5.101e+05	0.000	1.000
parti	-3.545e+00	6.973e+04	0.000	1.000
miracl	1.495e+02	9.587e+05	0.000	1.000
heard	8.336e+01	1.873e+05	0.000	1.000
contract	-2.584e+00	2.021e+05	0.000	1.000
frm	-1.711e+01	3.396e+05	0.000	1.000
don	-2.120e+00	1.728e+05	0.000	1.000
fill	1.978e+01	1.732e+05	0.000	1.000
loverboy	-9.497e+01	9.173e+05	0.000	1.000
fool	4.054e+00	2.198e+05	0.000	1.000
hiya	3.774e+01	7.350e+05	0.000	1.000
flower	1.801e+01	2.958e+05	0.000	1.000

support	1.116e+01	3.465e+05	0.000	1.000
bitch	-1.045e+00	1.381e+05	0.000	1.000
convey	-4.962e+00	7.211e+05	0.000	1.000
regard	-1.748e+01	7.043e+05	0.000	1.000
digit	2.902e+01	6.332e+05	0.000	1.000
sipix	-1.444e+02	1.453e+06	0.000	1.000
doctor	2.831e+01	2.948e+05	0.000	1.000
receipt	3.525e+01	3.440e+05	0.000	1.000
black	-5.604e+01	1.402e+06	0.000	1.000
aiyah	-4.382e+01	2.111e+05	0.000	1.000
httpwwwurawinnercom	-1.781e+02	6.108e+06	0.000	1.000
onto	1.633e+02	6.396e+06	0.000	1.000
surpris	-3.148e+01	6.599e+05	0.000	1.000
grin	-1.418e+01	1.794e+07	0.000	1.000
onlin	1.631e+00	1.184e+05	0.000	1.000
cancel	6.336e+00	3.101e+05	0.000	1.000
whos	-1.608e+01	2.107e+05	0.000	1.000
howz	-2.241e+01	1.529e+05	0.000	1.000
film	-4.468e+00	4.478e+04	0.000	1.000
luck	1.116e+01	2.168e+05	0.000	1.000
station	1.809e+01	3.980e+05	0.000	1.000
cute	-2.640e+01	6.882e+05	0.000	1.000
energi	5.936e+01	2.181e+05	0.000	1.000
chennai	-1.996e+01	6.305e+04	0.000	1.000
choic	6.458e+01	2.740e+07	0.000	1.000
deliv	6.567e+01	4.364e+05	0.000	1.000
arm	3.716e+01	2.783e+05	0.000	1.000
strong	3.499e+01	7.630e+05	0.000	1.000
nobodi	-2.071e+02	2.449e+06	0.000	1.000
doc	1.826e+01	1.609e+05	0.000	1.000
honey	-2.886e+01	4.205e+05	0.000	1.000
parent	1.327e+01	8.380e+04	0.000	1.000
cell	3.573e+01	4.172e+05	0.000	1.000
mode	-1.329e+02	9.924e+05	0.000	1.000
other	-5.802e+00	1.746e+05	0.000	1.000
tour	-4.068e+01	3.139e+05	0.000	1.000
alrit	-2.035e+01	1.516e+05	0.000	1.000
shd	4.066e+01	5.366e+05	0.000	1.000
fren	-4.250e+00	1.879e+05	0.000	1.000
super	-2.409e+01	1.138e+05	0.000	1.000
tht	1.775e+01	3.363e+05	0.000	1.000
croydon	1.609e+00	2.290e+06	0.000	1.000
fantasi	2.334e+01	2.220e+05	0.000	1.000

cover	-3.755e+01	5.224e+05	0.000	1.000
write	-1.658e+01	2.875e+05	0.000	1.000
fact	4.330e+01	7.530e+05	0.000	1.000
share	-6.654e+00	6.191e+05	0.000	1.000
vote	1.008e+01	1.554e+05	0.000	1.000
practic	-8.552e+00	4.289e+05	0.000	1.000
total	-1.775e+01	2.566e+05	0.000	1.000
lmao	-3.549e+01	1.722e+05	0.000	1.000
alon	5.476e+01	1.412e+05	0.000	1.000
natur	1.397e+01	4.607e+05	0.000	1.000
vomit	-7.012e+00	4.846e+05	0.000	1.000
internet	-3.760e+01	2.692e+05	0.000	1.000
arriv	-9.057e+01	6.613e+05	0.000	1.000
auction	3.678e+01	1.446e+05	0.000	1.000
difficult	-1.263e+01	3.042e+05	0.000	1.000
prefer	8.363e+01	4.637e+05	0.000	1.000
south	-8.182e+01	4.918e+05	0.000	1.000
gas	-3.782e+01	4.012e+05	0.000	1.000
excel	-7.955e+01	1.251e+05	-0.001	0.999
freak	-2.032e+01	8.621e+04	0.000	1.000
john	-1.186e+01	5.847e+05	0.000	1.000
instead	-2.696e+01	3.787e+05	0.000	1.000
london	6.700e+01	5.516e+05	0.000	1.000
wasnt	-2.625e+01	1.409e+05	0.000	1.000
respons	-2.253e+00	2.458e+05	0.000	1.000
lei	-1.915e+01	6.012e+04	0.000	1.000
lookin	-5.637e+01	7.441e+05	0.000	1.000
insid	6.166e+00	2.594e+05	0.000	1.000
remain	-1.906e+01	5.732e+04	0.000	1.000
father	-2.774e+00	4.375e+05	0.000	1.000
doin	3.755e+01	4.288e+05	0.000	1.000
slowli	2.511e+01	4.349e+05	0.000	1.000
list	-1.986e+01	8.590e+04	0.000	1.000
btnationalr	-5.233e+01	9.321e+05	0.000	1.000
wednesday	1.674e+01	3.103e+05	0.000	1.000
thnk	-3.348e+01	9.016e+05	0.000	1.000
excus	1.118e+01	1.384e+05	0.000	1.000
cook	-1.851e+01	3.671e+05	0.000	1.000
costa	NA	NA	NA	NA
`cost&pm`	NA	NA	NA	NA
maxmin	NA	NA	NA	NA
sol	NA	NA	NA	NA
stockport	NA	NA	NA	NA

	NA	NA	NA	NA
toclaim				
thinkin	-7.031e+00	2.421e+05	0.000	1.000
relat	1.079e+01	2.050e+05	0.000	1.000
mon	-3.044e+01	1.340e+05	0.000	1.000
role	1.087e+00	2.241e+05	0.000	1.000
theatr	1.032e+01	2.750e+05	0.000	1.000
ahmad	3.830e+01	9.828e+04	0.000	1.000
armand	3.158e-01	2.232e+05	0.000	1.000
med	2.849e+01	8.756e+05	0.000	1.000
vikki	1.295e+01	1.491e+05	0.000	1.000
manag	-1.996e+01	9.014e+04	0.000	1.000
pix	-8.415e+01	2.873e+05	0.000	1.000
colleagu	1.714e+02	8.019e+05	0.000	1.000
mood	-1.852e+01	4.397e+04	0.000	1.000
sed	-2.160e+01	9.503e+05	0.000	1.000
sofa	-1.812e+01	1.465e+06	0.000	1.000
tear	8.318e+01	2.357e+06	0.000	1.000
wors	-7.726e-01	2.868e+05	0.000	1.000
issu	4.116e+01	2.151e+05	0.000	1.000
pussi	-2.039e+01	1.121e+05	0.000	1.000
housemaid	-3.171e+02	2.985e+06	0.000	1.000
solv	1.534e+02	6.758e+05	0.000	1.000
budget	4.740e+01	9.543e+05	0.000	1.000
random	-4.013e+01	3.694e+05	0.000	1.000
cancer	1.584e+01	6.100e+05	0.000	1.000
feb	-3.972e+01	1.567e+06	0.000	1.000
ignor	-4.106e+00	2.643e+05	0.000	1.000
hug	-1.859e+02	2.369e+06	0.000	1.000
tariff	-1.781e+02	1.013e+06	0.000	1.000

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 3514.537 on 4457 degrees of freedom
 Residual deviance: 19.485 on 3262 degrees of freedom
 AIC: 2411.5

Number of Fisher Scoring iterations: 25

```
# Predicciones
predicted_probs <- predict(log_reg_model,
```



```

newdata = test_set, type = "response")

# Convertir a clases - umbral de 0.5
predicted_labels <- ifelse(predicted_probs > 0.5, "spam", "ham")
predicted_labels <- as.factor(predicted_labels)

# Matriz de confusión
confusionMatrix(data = predicted_labels, reference = test_set$label, positive = "spam")

```

Confusion Matrix and Statistics

```

              Reference
Prediction ham spam
      ham  856   43
      spam 109  106

      Accuracy : 0.8636
      95% CI : (0.842, 0.8832)
      No Information Rate : 0.8662
      P-Value [Acc > NIR] : 0.6247

      Kappa : 0.5041

      Mcnemar's Test P-Value : 1.348e-07

      Sensitivity : 0.71141
      Specificity : 0.88705
      Pos Pred Value : 0.49302
      Neg Pred Value : 0.95217
      Prevalence : 0.13375
      Detection Rate : 0.09515
      Detection Prevalence : 0.19300
      Balanced Accuracy : 0.79923

      'Positive' Class : spam

```

```

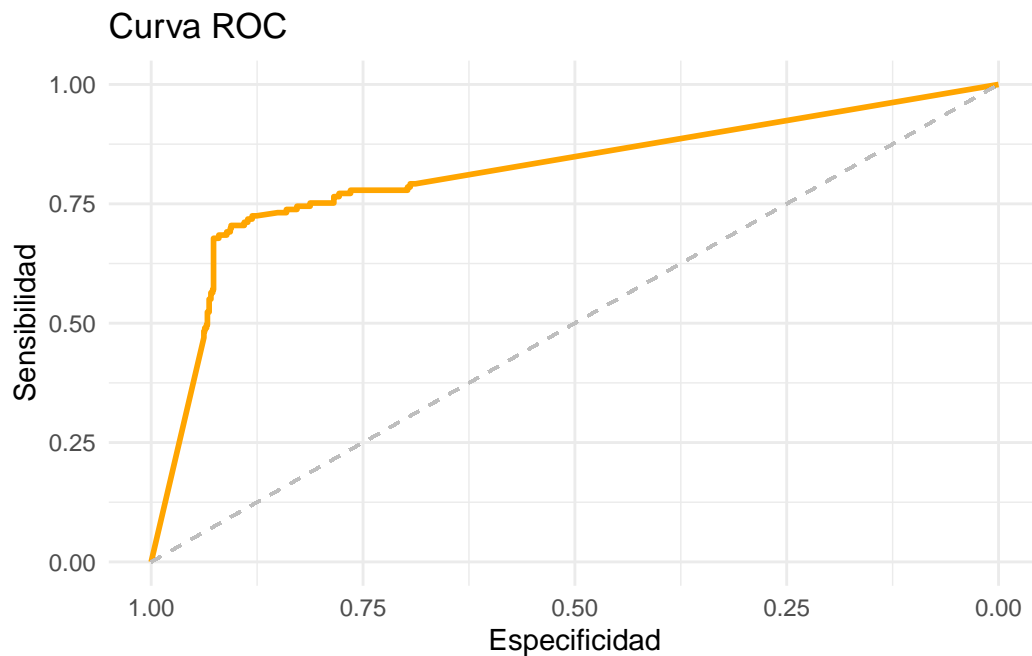
roc_obj <- roc(test_set$label, predicted_probs)

# AUC
auc_value <- auc(roc_obj); auc_value

```

Area under the curve: 0.8151

```
# Curva ROC
ggroc(roc_obj, colour = 'orange', size = 1) +
  geom_segment(aes(x = 1, xend = 0, y = 0, yend = 1), color="grey", linetype="dashed") +
  labs(
    title = "Curva ROC",
    x = "Especificidad",
    y = "Sensibilidad"
  ) +
  theme_minimal()
```



Modelo Naive-Bayes

```
# Entrenar modelo
X_train <- train_set[, -ncol(train_set)]
y_train <- train_set$label

nb_model <- naiveBayes(x = X_train, y = y_train, laplace = TRUE)
```

Confusion Matrix and Statistics

Reference
Prediction ham spam
ham 0 0
spam 965 149

Accuracy : 0.1338
95% CI : (0.1143, 0.1552)
No Information Rate : 0.8662
P-Value [Acc > NIR] : 1

Kappa : 0

McNemar's Test P-Value : <2e-16

Sensitivity : 1.0000
Specificity : 0.0000
Pos Pred Value : 0.1338
Neg Pred Value : NaN
Prevalence : 0.1338
Detection Rate : 0.1338
Detection Prevalence : 1.0000
Balanced Accuracy : 0.5000

'Positive' Class : spam

```
# Curva ROC
nb_predicted_probs <- predict(nb_model, newdata = X_test, type = "raw")

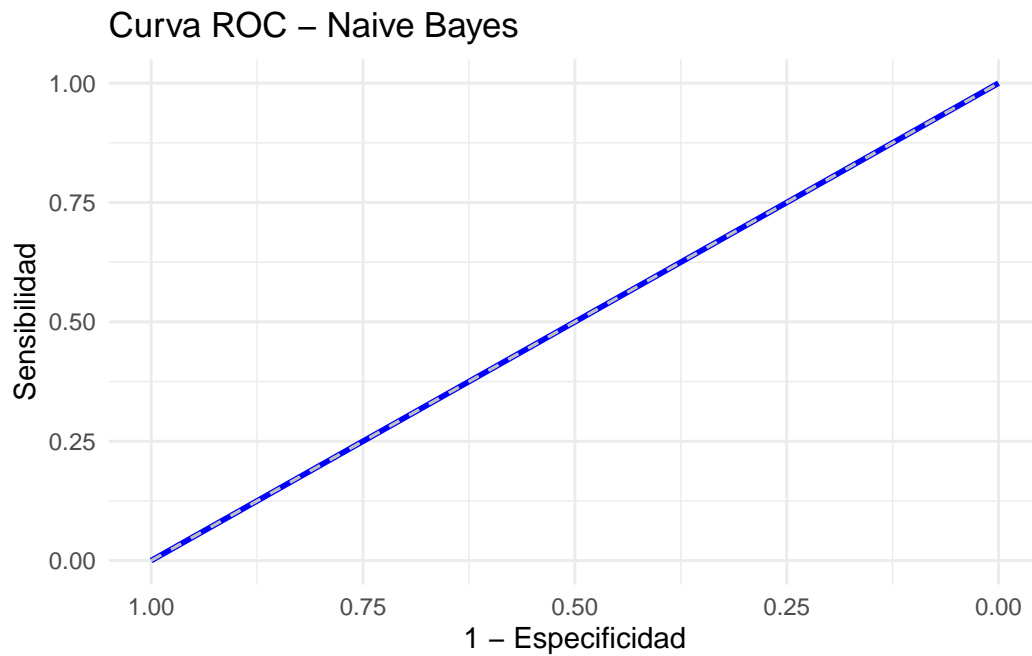
nb_roc_obj <- roc(y_test, nb_predicted_probs[, "spam"])

# AUC
nb_auc_value <- auc(nb_roc_obj)
print(paste("AUC del modelo Naive-Bayes:", round(nb_auc_value, 4)))
```

```
[1] "AUC del modelo Naive-Bayes: 0.5"
```

```
# Curva ROC
ggroc(nb_roc_obj, colour = 'blue', size = 1) +
  geom_segment(aes(x = 1, xend = 0, y = 0, yend = 1), color="grey", linetype="dashed") +
```

```
labs(
  title = "Curva ROC - Naive Bayes",
  x = "1 - Especificidad",
  y = "Sensibilidad"
) +
theme_minimal()
```



Naive-Bayes con otra prior

Se usara una prior de 50%/50%

```
# Entrenar modelo
nb_model_prior <- naiveBayes(x = X_train, y = y_train,
                             prior = c(ham = 0.5, spam = 0.5))

# Evaluar modelo
nb_predicted_labels_prior <- predict(nb_model_prior, newdata = X_test)
confusionMatrix(data = nb_predicted_labels_prior, reference = y_test, positive = "spam")
```

Confusion Matrix and Statistics

```

      Reference
Prediction ham spam
      ham    0    0
      spam 965  149

      Accuracy : 0.1338
      95% CI : (0.1143, 0.1552)
      No Information Rate : 0.8662
      P-Value [Acc > NIR] : 1

      Kappa : 0

      McNemar's Test P-Value : <2e-16

      Sensitivity : 1.0000
      Specificity : 0.0000
      Pos Pred Value : 0.1338
      Neg Pred Value :      NaN
      Prevalence : 0.1338
      Detection Rate : 0.1338
      Detection Prevalence : 1.0000
      Balanced Accuracy : 0.5000

      'Positive' Class : spam

```

Bernoulli Naive-Bayes

Esta variante no considera la frecuencia de las palabras, sino únicamente su presencia o ausencia.

```

X_bernoulli <- X %>%
  mutate(across(-label, ~as.factor(ifelse(. > 0, 1, 0))))

train_set_bernoulli <- subset(X_bernoulli, split == TRUE)
test_set_bernoulli <- subset(X_bernoulli, split == FALSE)

# Entrenar modelo
X_train_bernoulli <- train_set_bernoulli[, -ncol(train_set_bernoulli)]
y_train_bernoulli <- train_set_bernoulli$label

```

```

nb_model_bernoulli <- naiveBayes(x = X_train_bernoulli, y = y_train_bernoulli)

# Evaluar modelo
X_test_bernoulli <- test_set_bernoulli[, -ncol(test_set_bernoulli)]
y_test_bernoulli <- test_set_bernoulli$label
nb_predicted_labels_bernoulli <- predict(nb_model_bernoulli, newdata = X_test_bernoulli)

confusionMatrix(data = nb_predicted_labels_bernoulli, reference = y_test_bernoulli, positive

```

Confusion Matrix and Statistics

```

              Reference
Prediction ham spam
      ham  964   17
      spam    1  132

      Accuracy : 0.9838
      95% CI : (0.9746, 0.9904)
No Information Rate : 0.8662
P-Value [Acc > NIR] : < 2.2e-16

```

Kappa : 0.927

McNemar's Test P-Value : 0.000407

```

      Sensitivity : 0.8859
      Specificity : 0.9990
Pos Pred Value : 0.9925
Neg Pred Value : 0.9827
Prevalence : 0.1338
Detection Rate : 0.1185
Detection Prevalence : 0.1194
Balanced Accuracy : 0.9424

```

'Positive' Class : spam

```

nb_predicted_probs_bernoulli <- predict(nb_model_bernoulli, newdata = X_test_bernoulli, type
nb_roc_obj_bernoulli <- roc(y_test_bernoulli, nb_predicted_probs_bernoulli[, "spam"])

```

Setting levels: control = ham, case = spam

Setting direction: controls < cases

```
nb_auc_bernoulli <- auc(nb_roc_obj_bernoulli)
print(paste("AUC del modelo Bernoulli:", round(nb_auc_bernoulli, 4)))
```

```
[1] "AUC del modelo Bernoulli: 0.9861"
```

```
ggroc(nb_roc_obj_bernoulli, colour = 'darkgreen', size = 1) +
  geom_segment(aes(x = 1, xend = 0, y = 0, yend = 1),
               color = "grey", linetype = "dashed") +
  labs(
    title = "Curva ROC - Naive Bayes Bernoulli",
    x = "1 - Especificidad",
    y = "Sensibilidad"
  ) +
  theme_minimal()
```

Warning in `geom_segment(aes(x = 1, xend = 0, y = 0, yend = 1), color = "grey", : All aesthetic values must be the same length. Please consider using `annotate()` or provide this layer with data containing a single row.`

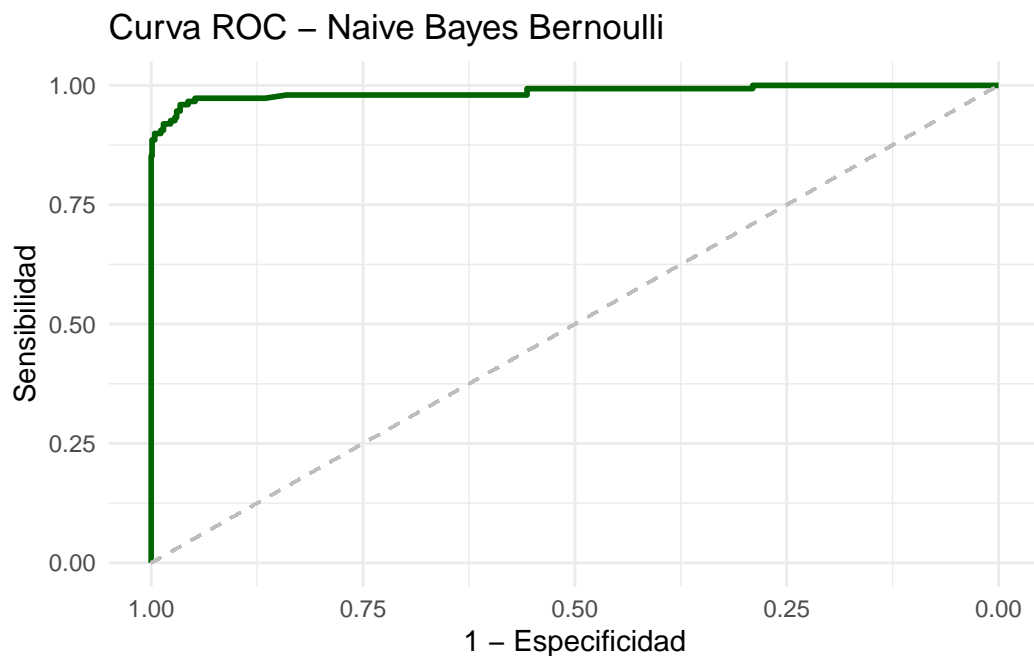


Tabla comparativa de resultados

```
# Valores AUC
# modelo Naive Bayes con Prior
nb_predicted_probs_prior <- predict(nb_model_prior, newdata = X_test, type = "raw")
roc_obj_prior <- roc(y_test, nb_predicted_probs_prior[, "spam"])
auc_value_prior <- auc(roc_obj_prior)

# modelo Naive Bayes Bernoulli
nb_predicted_probs_bernoulli <- predict(nb_model_bernoulli, newdata = X_test_bernoulli, type = "raw")
roc_obj_bernoulli <- roc(y_test_bernoulli, nb_predicted_probs_bernoulli[, "spam"])
auc_value_bernoulli <- auc(roc_obj_bernoulli)

# Metricas Matriz de confusión
log_reg_metrics <- confusionMatrix(data = predicted_labels, reference = test_set$label, positive = "spam")
nb_metrics <- confusionMatrix(data = nb_predicted_labels, reference = y_test, positive = "spam")
nb_prior_metrics <- confusionMatrix(data = nb_predicted_labels_prior, reference = y_test, positive = "spam")
nb_bernoulli_metrics <- confusionMatrix(data = nb_predicted_labels_bernoulli, reference = y_test, positive = "spam")

comparison_df <- data.frame(
  Modelo = c("Regresión Logística", "Naive Bayes Multinomial", "NB con Prior 50/50", "Naive Bayes Bernoulli"),
  Accuracy = c(
    log_reg_metrics$overall["Accuracy"],
    nb_metrics$overall["Accuracy"],
    nb_prior_metrics$overall["Accuracy"],
    nb_bernoulli_metrics$overall["Accuracy"]
  ),
  Precision = c(
    log_reg_metrics$byClass["Precision"],
    nb_metrics$byClass["Precision"],
    nb_prior_metrics$byClass["Precision"],
    nb_bernoulli_metrics$byClass["Precision"]
  ),
  Sensibilidad = c(
    log_reg_metrics$byClass["Sensitivity"],
    nb_metrics$byClass["Sensitivity"],
    nb_prior_metrics$byClass["Sensitivity"],
    nb_bernoulli_metrics$byClass["Sensitivity"]
  ),
  AUC = c(
    auc_value_prior,
    auc_value_bernoulli,
    auc_value_prior,
    auc_value_bernoulli
  )
)
```



```

    auc_value,
    nb_auc_value,
    auc_value_prior,
    auc_value_bernoulli
  )
)

knitr::kable(comparison_df,
              digits = 4)

```

Modelo	Accuracy	Precision	Sensibilidad	AUC
Regresión Logística	0.8636	0.4930	0.7114	0.8151
Naive Bayes Multinomial	0.1338	0.1338	1.0000	0.5000
NB con Prior 50/50	0.1338	0.1338	1.0000	0.5000
Naive Bayes Bernoulli	0.9838	0.9925	0.8859	0.9861