In [0]:

```python
from pyspark.ml.feature import Tokenizer, StopWordsRemover, HashingTF, IDF, StringIndexer
from pyspark.ml import Pipeline
from pyspark.ml.evaluation import MulticlassClassificationEvaluator
from pyspark.sql.functions import col, when, size, length, rand, regexp_replace, asc, desc
from pyspark.ml.classification import NaiveBayes
```

In [0]:

```python
# The URL of the raw file on GitHub
github_raw_url = "https://raw.githubusercontent.com/swetha0404/datasets/main/bdma/spamham
text.csv"

# The destination path where you want to save the file in Databricks
dbfs_destination_path = "dbfs:/FileStore/tables/spamhamtext.csv"

# Using dbutils to read and write the file
dbutils.fs.cp(github_raw_url, dbfs_destination_path)
```

Out[48]: True

In [0]:

```python
# Load the dataset
df = spark.read.csv("dbfs:/FileStore/tables/spamhamtext.csv", header=True, inferSchema=True)
df.show()
```

```
+--------+--------------------+
|Category|             Message|
+--------+--------------------+
|     ham|Go until jurong p...|
|     ham|Ok lar... Joking ...|
|    spam|Free entry in 2 a...|
|     ham|U dun say so earl...|
|     ham|Nah I don't think...|
|    spam|FreeMsg Hey there...|
|     ham|Even my brother i...|
|     ham|As per your reque...|
|    spam|WINNER!! As a val...|
|    spam|Had your mobile 1...|
|     ham|I'm gonna be home...|
|    spam|SIX chances to wi...|
|    spam|URGENT! You have ...|
|     ham|I've been searchi...|
|     ham|I HAVE A DATE ON ...|
|    spam|XXXMobileMovieClu...|
|     ham|Oh k...i'm watchi...|
|     ham|Eh u remember how...|
|     ham|Fine if thats th...|
|    spam|England v Macedon...|
+--------+--------------------+
only showing top 20 rows
```

In [0]:

```python
df.select("Category").distinct().show()
```

```
+--------------------+
|            Category|
+--------------------+
|ham\tHI BABE UAWA...|
|                 ham|
|                spam|
|         ham\tYeah|
+--------------------+
```

In [0]:

```
# newdf = df.filter((df["Type"]) <= 20)
df.createOrReplaceTempView("messages")
filtered_df = spark.sql("SELECT * FROM messages WHERE LENGTH(Category) <= 5")
```

In [0]:

```
# filtered_df.select('Type').distinct().show()
df = filtered_df
df.printSchema()
```

```
root
 |-- Category: string (nullable = true)
 |-- Message: string (nullable = true)
```

In [0]:

```
df.show()
```

```
+--------+--------------------+
|Category|             Message|
+--------+--------------------+
|     ham|Go until jurong p...|
|     ham|Ok lar... Joking ...|
|    spam|Free entry in 2 a...|
|     ham|U dun say so earl...|
|     ham|Nah I don't think...|
|    spam|FreeMsg Hey there...|
|     ham|Even my brother i...|
|     ham|As per your reque...|
|    spam|WINNER!! As a val...|
|    spam|Had your mobile 1...|
|     ham|I'm gonna be home...|
|    spam|SIX chances to wi...|
|    spam|URGENT! You have ...|
|     ham|I've been searchi...|
|     ham|I HAVE A DATE ON ...|
|    spam|XXXMobileMovieClu...|
|     ham|Oh k...i'm watchi...|
|     ham|Eh u remember how...|
|     ham|Fine if thats th...|
|    spam|England v Macedon...|
+--------+--------------------+
only showing top 20 rows
```

In [0]:

```
df = df.orderBy(rand())
df.show()
```

```
+--------+--------------------+
|Category|             Message|
+--------+--------------------+
|     ham|Sorry, I'll call ...|
|     ham|This is all just ...|
|     ham|Sat right? Okay t...|
|     ham|Some friends want...|
|     ham|Hurry home u big ...|
|     ham|Sorry . I will be...|
|     ham|Joy's father is J...|
|     ham|Ok good then i la...|
|     ham|Ugh just got outt...|
|     ham|Tell me again wha...|
|     ham|Did u got that pe...|
|     ham|Abeg, make profit...|
|     ham|And now electrici...|
|     ham|Its on in engalnd...|
```

```
|    ham|Bugis oso near wa...|
|   spam|You are guarantee...|
|    ham|Then ur physics g...|
|    ham|"Finally it has h...|
|    ham|Dear we got  &lt;...|
|    ham|Reckon need to be...|
+--------+-------------------+
only showing top 20 rows
```

In [0]:

```
# Rename columns for convenience
df = df.withColumnRenamed("Message", "text").withColumnRenamed("Category", "label")
```

In [0]:

```
df = df.withColumn("text_1", regexp_replace("text", "[\W_]+", " "))
df = df.withColumn("text_2", regexp_replace("text_1", "^\s+|\s+$", ""))
df = df.withColumn("text_3", regexp_replace("text_2", "\s+", " "))
df.show(30)
```

```
+-----+-------------------+-------------------+-------------------+-----------------
---+
|label|               text|             text_1|             text_2|             text
_3|
+-----+-------------------+-------------------+-------------------+-----------------
---+
|  ham|Sorry, I'll call ...|Sorry I ll call l...|Sorry I ll call l...|Sorry I ll call l..
.|
|  ham|This is all just ...|This is all just ...|This is all just ...|This is all just ..
.|
|  ham|Sat right? Okay t...|Sat right Okay th...|Sat right Okay th...|Sat right Okay th..
.|
|  ham|Some friends want...|Some friends want...|Some friends want...|Some friends want..
.|
|  ham|Hurry home u big ...|Hurry home u big ...|Hurry home u big ...|Hurry home u big ..
.|
|  ham|Sorry . I will be...|Sorry I will be a...|Sorry I will be a...|Sorry I will be a..
.|
|  ham|Joy's father is J...|Joy s father is J...|Joy s father is J...|Joy s father is J..
.|
|  ham|Ok good then i la...|Ok good then i la...|Ok good then i la...|Ok good then i la..
.|
|  ham|Ugh just got outt...|Ugh just got outt...|Ugh just got outt...|Ugh just got outt..
.|
|  ham|Tell me again wha...|Tell me again wha...|Tell me again wha...|Tell me again wha..
.|
|  ham|Did u got that pe...|Did u got that pe...|Did u got that pe...|Did u got that pe..
.|
|  ham|Abeg, make profit...|Abeg make profit ...|Abeg make profit ...|Abeg make profit ..
.|
|  ham|And now electrici...|And now electrici...|And now electrici...|And now electrici..
.|
|  ham|Its on in engalnd...|Its on in engalnd...|Its on in engalnd...|Its on in engalnd..
.|
|  ham|Bugis oso near wa...| Bugis oso near wat |  Bugis oso near wat|  Bugis oso near wa
t|
| spam|You are guarantee...|You are guarantee...|You are guarantee...|You are guarantee..
.|
|  ham|Then ur physics g...|Then ur physics g...|Then ur physics g...|Then ur physics g..
.|
|  ham|"Finally it has h...| Finally it has h...|Finally it has ha...|Finally it has ha..
.|
|  ham|Dear we got  &lt;...|Dear we got lt gt...|Dear we got lt gt...|Dear we got lt gt..
.|
|  ham|Reckon need to be...|Reckon need to be...|Reckon need to be...|Reckon need to be..
.|
|  ham|What do u reckon ...|What do u reckon ...|What do u reckon ...|What do u reckon ..
.|
|  ham|Just come home. I...|Just come home I ...|Just come home I ...|Just come home I ..
.|
```

```
.|
|  ham|Ok set let u noe ...|Ok set let u noe ...|Ok set let u noe ...|Ok set let u noe ..
.|
| spam|Gr8 new service -...|Gr8 new service l...|Gr8 new service l...|Gr8 new service l..
.|
|  ham|U should have mad...|U should have mad...|U should have mad...|U should have mad..
.|
|  ham|Great escape. I f...|Great escape I fa...|Great escape I fa...|Great escape I fa..
.|
|  ham|Yo, you gonna sti...|Yo you gonna stil...|Yo you gonna stil...|Yo you gonna stil..
.|
|  ham|Nt only for drivi...|Nt only for drivi...|Nt only for drivi...|Nt only for drivi..
.|
|  ham|Prabha..i'm soryd...|Prabha i m soryda...|Prabha i m soryda...|Prabha i m soryda..
.|
|  ham|Are you not aroun...|Are you not aroun...|Are you not aroun...|Are you not aroun..
.|
+-----+-------------------+-------------------+-------------------+-----------------
---+
only showing top 30 rows
```

In [0]:

```
tokenizer = Tokenizer(inputCol="text_3", outputCol="words")
tokenized = tokenizer.transform(df)
tokenized.show()
```

```
+-----+-------------------+-------------------+-------------------+-----------------
---+-------------------+
|label|               text|             text_1|             text_2|             text
_3|              words|
+-----+-------------------+-------------------+-------------------+-----------------
---+-------------------+
|  ham|Sorry, I'll call ...|Sorry I ll call l...|Sorry I ll call l...|Sorry I ll call l..
.|[sorry, i, ll, ca...|
|  ham|This is all just ...|This is all just ...|This is all just ...|This is all just ..
.|[this, is, all, j...|
|  ham|Sat right? Okay t...|Sat right Okay th...|Sat right Okay th...|Sat right Okay th..
.|[sat, right, okay...|
|  ham|Some friends want...|Some friends want...|Some friends want...|Some friends want..
.|[some, friends, w...|
|  ham|Hurry home u big ...|Hurry home u big ...|Hurry home u big ...|Hurry home u big ..
.|[hurry, home, u, ...|
|  ham|Sorry . I will be...|Sorry I will be a...|Sorry I will be a...|Sorry I will be a..
.|[sorry, i, will, ...|
|  ham|Joy's father is J...|Joy s father is J...|Joy s father is J...|Joy s father is J..
.|[joy, s, father, ...|
|  ham|Ok good then i la...|Ok good then i la...|Ok good then i la...|Ok good then i la..
.|[ok, good, then, ...|
|  ham|Ugh just got outt...|Ugh just got outt...|Ugh just got outt...|Ugh just got outt..
.|[ugh, just, got, ...|
|  ham|Tell me again wha...|Tell me again wha...|Tell me again wha...|Tell me again wha..
.|[tell, me, again,...|
|  ham|Did u got that pe...|Did u got that pe...|Did u got that pe...|Did u got that pe..
.|[did, u, got, tha...|
|  ham|Abeg, make profit...|Abeg make profit ...|Abeg make profit ...|Abeg make profit ..
.|[abeg, make, prof...|
|  ham|And now electrici...|And now electrici...|And now electrici...|And now electrici..
.|[and, now, electr...|
|  ham|Its on in engalnd...|Its on in engalnd...|Its on in engalnd...|Its on in engalnd..
.|[its, on, in, eng...|
|  ham|Bugis oso near wa...| Bugis oso near wat |  Bugis oso near wat|  Bugis oso near wa
t|[bugis, oso, near...|
| spam|You are guarantee...|You are guarantee...|You are guarantee...|You are guarantee..
.|[you, are, guaran...|
|  ham|Then ur physics g...|Then ur physics g...|Then ur physics g...|Then ur physics g..
.|[then, ur, physic...|
|  ham|"Finally it has h...| Finally it has h...|Finally it has ha...|Finally it has ha..
.|[finally, it, has...|
|  ham|Dear we got  &lt;...|Dear we got lt gt...|Dear we got lt gt...|Dear we got lt gt..
.|[dear, we, got, l...|
|  ham|Reckon need to be   |Reckon need to be   |Reckon need to be   |Reckon need to be
```

```
|  ham|Reckon need to be...|Reckon need to be...|Reckon need to be...|Reckon need to be..
.|[reckon, need, to...|
+-----+-------------------+-------------------+-------------------+----------------
---+-------------------+
only showing top 20 rows
```

In [0]:

```
remover = StopWordsRemover(inputCol="words", outputCol="filtered")
removed = remover.transform(tokenized)
removed.show()
```

```
+-----+-------------------+-------------------+-------------------+----------------
---+-------------------+-------------------+
|label|               text|             text_1|             text_2|            text
_3|              words|           filtered|
+-----+-------------------+-------------------+-------------------+----------------
---+-------------------+-------------------+
|  ham|Sorry, I'll call ...|Sorry I ll call l...|Sorry I ll call l...|Sorry I ll call l..
.|[sorry, i, ll, ca...|[sorry, ll, call,...|
|  ham|This is all just ...|This is all just ...|This is all just ...|This is all just ..
.|[this, is, all, j...|     [creepy, crazy]|
|  ham|Sat right? Okay t...|Sat right Okay th...|Sat right Okay th...|Sat right Okay th..
.|[sat, right, okay...|[sat, right, okay...|
|  ham|Some friends want...|Some friends want...|Some friends want...|Some friends want..
.|[some, friends, w...|[friends, want, d...|
|  ham|Hurry home u big ...|Hurry home u big ...|Hurry home u big ...|Hurry home u big ..
.|[hurry, home, u, ...|[hurry, home, u, ...|
|  ham|Sorry . I will be...|Sorry I will be a...|Sorry I will be a...|Sorry I will be a..
.|[sorry, i, will, ...|[sorry, able, get...|
|  ham|Joy's father is J...|Joy s father is J...|Joy s father is J...|Joy s father is J..
.|[joy, s, father, ...|[joy, father, joh...|
|  ham|Ok good then i la...|Ok good then i la...|Ok good then i la...|Ok good then i la..
.|[ok, good, then, ...|[ok, good, later,...|
|  ham|Ugh just got outt...|Ugh just got outt...|Ugh just got outt...|Ugh just got outt..
.|[ugh, just, got, ...|[ugh, got, outta,...|
|  ham|Tell me again wha...|Tell me again wha...|Tell me again wha...|Tell me again wha..
.|[tell, me, again,...|     [tell, address]|
|  ham|Did u got that pe...|Did u got that pe...|Did u got that pe...|Did u got that pe..
.|[did, u, got, tha...|[u, got, persons,...|
|  ham|Abeg, make profit...|Abeg make profit ...|Abeg make profit ...|Abeg make profit ..
.|[abeg, make, prof...|[abeg, make, prof...|
|  ham|And now electrici...|And now electrici...|And now electrici...|And now electrici..
.|[and, now, electr...|[electricity, wen...|
|  ham|Its on in engalnd...|Its on in engalnd...|Its on in engalnd...|Its on in engalnd..
.|[its, on, in, eng...|[engalnd, telly, ...|
|  ham|Bugis oso near wa...| Bugis oso near wat |  Bugis oso near wat|  Bugis oso near wa
t|[bugis, oso, near...|[bugis, oso, near...|
| spam|You are guarantee...|You are guarantee...|You are guarantee...|You are guarantee..
.|[you, are, guaran...|[guaranteed, late...|
|  ham|Then ur physics g...|Then ur physics g...|Then ur physics g...|Then ur physics g..
.|[then, ur, physic...|  [ur, physics, get]|
|  ham|"Finally it has h...| Finally it has h...|Finally it has ha...|Finally it has ha..
.|[finally, it, has...|[finally, happene...|
|  ham|Dear we got  &lt;...|Dear we got lt gt...|Dear we got lt gt...|Dear we got lt gt..
.|[dear, we, got, l...|[dear, got, lt, g...|
|  ham|Reckon need to be...|Reckon need to be...|Reckon need to be...|Reckon need to be..
.|[reckon, need, to...|[reckon, need, to...|
+-----+-------------------+-------------------+-------------------+----------------
---+-------------------+-------------------+
only showing top 20 rows
```

In [0]:

```
new_df = removed.withColumn("label", when(col("label")=='spam', 0).when(col("label")=='h
am', 1))
new_df.show()
```

```
+-----+-------------------+-------------------+-------------------+----------------
---+-------------------+-------------------+
|label|               text|             text_1|             text_2|            text
```

```
|label|               text|             text_1|             text_2|               text
_3|               words|           filtered|
+-----+-------------------+-------------------+-------------------+-----------------
---+-------------------+-------------------+
|    1|Sorry, I'll call ...|Sorry I ll call l...|Sorry I ll call l...|Sorry I ll call l..
.|[sorry, i, ll, ca...|[sorry, ll, call,...|
|    1|This is all just ...|This is all just ...|This is all just ...|This is all just ..
.|[this, is, all, j...|    [creepy, crazy]|
|    1|Sat right? Okay t...|Sat right Okay th...|Sat right Okay th...|Sat right Okay th..
.|[sat, right, okay...|[sat, right, okay...|
|    1|Some friends want...|Some friends want...|Some friends want...|Some friends want..
.|[some, friends, w...|[friends, want, d...|
|    1|Hurry home u big ...|Hurry home u big ...|Hurry home u big ...|Hurry home u big ..
.|[hurry, home, u, ...|[hurry, home, u, ...|
|    1|Sorry . I will be...|Sorry I will be a...|Sorry I will be a...|Sorry I will be a..
.|[sorry, i, will, ...|[sorry, able, get...|
|    1|Joy's father is J...|Joy s father is J...|Joy s father is J...|Joy s father is J..
.|[joy, s, father, ...|[joy, father, joh...|
|    1|Ok good then i la...|Ok good then i la...|Ok good then i la...|Ok good then i la..
.|[ok, good, then, ...|[ok, good, later,...|
|    1|Ugh just got outt...|Ugh just got outt...|Ugh just got outt...|Ugh just got outt..
.|[ugh, just, got, ...|[ugh, got, outta,...|
|    1|Tell me again wha...|Tell me again wha...|Tell me again wha...|Tell me again wha..
.|[tell, me, again,...|    [tell, address]|
|    1|Did u got that pe...|Did u got that pe...|Did u got that pe...|Did u got that pe..
.|[did, u, got, tha...|[u, got, persons,...|
|    1|Abeg, make profit...|Abeg make profit ...|Abeg make profit ...|Abeg make profit ..
.|[abeg, make, prof...|[abeg, make, prof...|
|    1|And now electrici...|And now electrici...|And now electrici...|And now electrici..
.|[and, now, electr...|[electricity, wen...|
|    1|Its on in engalnd...|Its on in engalnd...|Its on in engalnd...|Its on in engalnd..
.|[its, on, in, eng...|[engalnd, telly, ...|
|    1|Bugis oso near wa...| Bugis oso near wat |  Bugis oso near wat|  Bugis oso near wa
t|[bugis, oso, near...|[bugis, oso, near...|
|    0|You are guarantee...|You are guarantee...|You are guarantee...|You are guarantee..
.|[you, are, guaran...|[guaranteed, late...|
|    1|Then ur physics g...|Then ur physics g...|Then ur physics g...|Then ur physics g..
.|[then, ur, physic...|  [ur, physics, get]|
|    1|"Finally it has h...| Finally it has h...|Finally it has ha...|Finally it has ha..
.|[finally, it, has...|[finally, happene...|
|    1|Dear we got  &lt;...|Dear we got lt gt...|Dear we got lt gt...|Dear we got lt gt..
.|[dear, we, got, l...|[dear, got, lt, g...|
|    1|Reckon need to be...|Reckon need to be...|Reckon need to be...|Reckon need to be..
.|[reckon, need, to...|[reckon, need, to...|
+-----+-------------------+-------------------+-------------------+-----------------
---+-------------------+-------------------+
only showing top 20 rows
```

In [0]:

```python
final_df = new_df.select("filtered", "label")
final_df.show()
```

```
+--------------------+-----+
|            filtered|label|
+--------------------+-----+
|[sorry, ll, call,...|    1|
|    [creepy, crazy]|    1|
|[sat, right, okay...|    1|
|[friends, want, d...|    1|
|[hurry, home, u, ...|    1|
|[sorry, able, get...|    1|
|[joy, father, joh...|    1|
|[ok, good, later,...|    1|
|[ugh, got, outta,...|    1|
|    [tell, address]|    1|
|[u, got, persons,...|    1|
|[abeg, make, prof...|    1|
|[electricity, wen...|    1|
|[engalnd, telly, ...|    1|
|[bugis, oso, near...|    1|
|[guaranteed, late...|    0|
```

```
|[guaranteed, late...|      0|
|   [ur, physics, get]|      1|
|[finally, happene...|      1|
|[dear, got, lt, g...|      1|
|[reckon, need, to...|      1|
+------------------+-----+
only showing top 20 rows
```

In [0]:

```
final_df = final_df.withColumnRenamed("filtered", "text")
display(final_df)
```

text	label
List(sorry, ll, call, later, meeting)	1
List(creepy, crazy)	1
List(sat, right, okay, thanks)	1
List(friends, want, drive, em, someplace, probably, take)	1
List(hurry, home, u, big, butt, hang, last, caller, u, food, done, m, starving, ask, cooked)	1
List(sorry, able, get, see, morning)	1
List(joy, father, john, john, joy, father, u, ans, ths, hav, lt, gt, iq, tis, ias, question, try, answer)	1
List(ok, good, later, come, find, c, lucky, told, go, earlier, later, pple, take, finish)	1
List(ugh, got, outta, class)	1
List(tell, address)	1

In [0]:

```
train, test = final_df.randomSplit([0.7, 0.3], seed = 200)
```

In [0]:

```
# train.show()
train.count()
```

Out[66]: 3874

In [0]:

```
# test.show()
test.count()
```

Out[67]: 1698

In [0]:

```
n = final_df.count()
temp = final_df.groupBy("label").count()
print("Count:", n)
```

Count: 5572

In [0]:

```
temp.show()
```

```
+-----+-----+
|label|count|
+-----+-----+
|    1| 4825|
|    0|  747|
+-----+-----+
```

In [0]:

```
labelCount = temp.rdd.map(lambda x: x[1]).collect()
print("labelCount:",labelCount)
```

labelCount: [4825, 747]

In [0]:

```
hamCount = labelCount[0]
spamCount = labelCount[1]
print("ham count: ", hamCount, "\nspam count: ", spamCount)
```

ham count: 4825
spam count: 747

In [0]:

```
priorham = hamCount / n
priorspam = spamCount / n

print("Prior of ham:", priorham, "\nPrior of spam:", priorspam)
```

Prior of ham: 0.8659368269921034
Prior of spam: 0.13406317300789664

In [0]:

```
trainInput = train.rdd.map(lambda x : [x[0], x[1]])

resultspam = trainInput.flatMap(lambda x: x[0] if x[1]==0 else []).filter(lambda x: x!=[
])
resultspam.collect()
```

Out[69]: ['07732584351',
 'rodger',
 'burns',
 'msg',
 'tried',
 'call',
 're',
 'reply',
 'sms',
 'free',
 'nokia',
 'mobile',
 'free',
 'camcorder',
 'please',
 'call',
 '08000930705',
 'delivery',
 'tomorrow',
 '07801543489',
 'guaranteed',
 'latests',
 'nokia',
 'phone',
 '40gb',
 'ipod',
 'mp3',
 'player',
 '500',
 'prize',
 'txt',
 'word',
 'collect',
 '83355',
 'tc',
 'llc',
 'ny',
 'usa',
 '150p',
 'mt',
```

'msgrcvd18',
'09066362231',
'urgent',
'mobile',
'07xxxxxxxxx',
'won',
'2',
'000',
'bonus',
'caller',
'prize',
'02',
'06',
'03',
'2nd',
'attempt',
'reach',
'call',
'09066362231',
'asap',
'0a',
'networks',
'allow',
'companies',
'bill',
'sms',
'responsible',
'suppliers',
'1',
'new',
'message',
'call',
'0207',
'083',
'6089',
'1',
'new',
'message',
'please',
'call',
'08712400200',
'1',
'new',
'message',
'please',
'call',
'08718738034',
'1',
'new',
'voicemail',
'please',
'call',
'08719181503',
'1',
'new',
'voicemail',
'please',
'call',
'08719181513',
'1',
'nokia',
'tone',
'4',
'ur',
'mob',
'every',
'week',
'txt',
'nok',
'87021',
'1st',
'tone'

'free',
'get',
'txtin',
'tell',
'ur',
'friends',
'150p',
'tone',
'16',
'reply',
'hl',
'4info',
'1',
'nokia',
'tone',
'4',
'ur',
'mob',
'every',
'week',
'txt',
'nok',
'87021',
'1st',
'tone',
'free',
'get',
'txtin',
'tell',
'ur',
'friends',
'150p',
'tone',
'16',
'reply',
'hl',
'4info',
'1',
'polyphonic',
'tone',
'4',
'ur',
'mob',
'every',
'week',
'txt',
'pt2',
'87575',
'1st',
'tone',
'free',
'get',
'txtin',
'tell',
'ur',
'friends',
'150p',
'tone',
'16',
'reply',
'hl',
'4info',
'1',
'polyphonic',
'tone',
'4',
'ur',
'mob',
'every',
'week',
'txt',
'pt2'

'87575',
'1st',
'tone',
'free',
'get',
'txtin',
'tell',
'ur',
'friends',
'150p',
'tone',
'16',
'reply',
'hl',
'4info',
'100',
'dating',
'service',
'cal',
'l',
'09064012103',
'box334sk38ch',
'18',
'days',
'euro2004',
'kickoff',
'u',
'kept',
'informed',
'latest',
'news',
'results',
'daily',
'unsubscribe',
'send',
'get',
'euro',
'stop',
'83222',
'1st',
'wk',
'free',
'gr8',
'tones',
'str8',
'2',
'u',
'wk',
'txt',
'nokia',
'8007',
'classic',
'nokia',
'tones',
'hit',
'8007',
'polys',
'nokia',
'150p',
'poly',
'200p',
'16',
'22',
'days',
'kick',
'euro2004',
'u',
'kept',
'date',
'latest',
'news',
'results',

'results',
'daily',
'removed',
'send',
'get',
'txt',
'stop',
'83222',
'2nd',
'attempt',
'contract',
'u',
'won',
'weeks',
'top',
'prize',
'either',
'1000',
'cash',
'200',
'prize',
'call',
'09066361921',
'2nd',
'time',
'tried',
'2',
'contact',
'u',
'u',
'won',
'750',
'pound',
'prize',
'2',
'claim',
'easy',
'call',
'08712101358',
'10p',
'per',
'min',
'bt',
'national',
'rate',
'2nd',
'time',
'tried',
'2',
'contact',
'u',
'u',
'won',
'750',
'pound',
'prize',
'2',
'claim',
'easy',
'call',
'08718726970',
'10p',
'per',
'min',
'bt',
'national',
'rate',
'2nd',
'time',
'tried',
'2',
'contact',
'u',

'u',
'u',
'won',
'750',
'pound',
'prize',
'2',
'claim',
'easy',
'call',
'087187272008',
'now1',
'10p',
'per',
'minute',
'bt',
'national',
'rate',
'2nd',
'time',
'tried',
'contact',
'u',
'u',
'won',
'1450',
'prize',
'claim',
'call',
'09053750005',
'b4',
'310303',
'cs',
'stop',
'sms',
'08718725756',
'140ppm',
'2nd',
'time',
'tried',
'contact',
'u',
'u',
'won',
'400',
'prize',
'2',
'claim',
'easy',
'call',
'087104711148',
'10p',
'per',
'minute',
'bt',
'national',
'rate',
'2p',
'per',
'min',
'call',
'germany',
'08448350055',
'bt',
'line',
'2p',
'per',
'min',
'check',
'planettalkinstant',
'com',
'info',
'c',

'text',
'stop',
'opt',
'3',
'free',
'tarot',
'texts',
'find',
'love',
'life',
'try',
'3',
'free',
'text',
'chance',
'85555',
'16',
'3',
'free',
'msgs',
'1',
'50',
'3',
'received',
'mobile',
'content',
'enjoy',
'400',
'xmas',
'reward',
'waiting',
'computer',
'randomly',
'picked',
'loyal',
'mobile',
'customers',
'receive',
'400',
'reward',
'call',
'09066380611',
'400',
'xmas',
'reward',
'waiting',
'computer',
'randomly',
'picked',
'loyal',
'mobile',
'customers',
'receive',
'400',
'reward',
'call',
'09066380611',
'44',
'7732584351',
'want',
'new',
'nokia',
'3510i',
'colour',
'phone',
'deliveredtomorrow',
'300',
'free',
'minutes',
'mobile',
'100',
'free'

'texts',
'free',
'camcorder',
'reply',
'call',
'08000930705',
'449050000301',
'won',
'2',
'000',
'price',
'claim',
'call',
'09050000301',
'4mths',
'half',
'price',
'orange',
'line',
'rental',
'latest',
'camera',
'phones',
'4',
'free',
'phone',
'11mths',
'call',
'mobilesdirect',
'free',
'08000938767',
'update',
'or2stoptxt',
'4mths',
'half',
'price',
'orange',
'line',
'rental',
'latest',
'camera',
'phones',
'4',
'free',
'phone',
'11mths',
'call',
'mobilesdirect',
'free',
'08000938767',
'update',
'or2stoptxt',
'cs',
'5',
'free',
'top',
'polyphonic',
'tones',
'call',
'087018728737',
'national',
'rate',
'get',
'toppoly',
'tune',
'sent',
'every',
'week',
'text',
'subpoly',
'81618',
'3',

'per',
'pole',
'unsub',
'08718727870',
'500',
'free',
'text',
'msgs',
'text',
'ok',
'80488',
'll',
'credit',
'account',
'500',
'new',
'mobiles',
'2004',
'must',
'go',
'txt',
'nokia',
'89545',
'collect',
'today',
'1',
'www',
'4',
'tc',
'biz',
'2optout',
'087187262701',
'50gbp',
'mtmsg18',
'500',
'new',
'mobiles',
'2004',
'must',
'go',
'txt',
'nokia',
'89545',
'collect',
'today',
'1',
'www',
'4',
'tc',
'biz',
'2optout',
'087187262701',
'50gbp',
'mtmsg18',
'txtauction',
'74355',
'xmas',
'iscoming',
'ur',
'awarded',
'either',
'500',
'cd',
'gift',
'vouchers',
'free',
'entry',
'2',
'r',
'100',
'weekly',
'draw'

'draw',
'txt',
'music',
'87066',
'tnc',
'8007',
'25p',
'4',
'alfie',
'moon',
'children',
'need',
'song',
'ur',
'mob',
'tell',
'ur',
'm8s',
'txt',
'tone',
'charity',
'8007',
'nokias',
'poly',
'charity',
'polys',
'zed',
'08701417012',
'profit',
'2',
'charity',
'8007',
'free',
'1st',
'week',
'no1',
'nokia',
'tone',
'4',
'ur',
'mob',
'every',
'week',
'txt',
'nokia',
'8007',
'get',
'txting',
'tell',
'ur',
'mates',
'www',
'getzed',
'co',
'uk',
'pobox',
'36504',
'w4',
'5wq',
'norm',
'150p',
'tone',
'16',
'80488',
'500',
'free',
'text',
'messages',
'valid',
'31',
'december',
'2005',
'83039'

'62735',
'450',
'uk',
'break',
'accommodationvouchers',
'terms',
'conditions',
'apply',
'2',
'claim',
'mustprovide',
'claim',
'number',
'15541',
'85233',
'free',
'ringtone',
'reply',
'real',
'88066',
'88066',
'lost',
'3pound',
'help',
'88066',
'lost',
'12',
'help',
'88800',
'89034',
'premium',
'phone',
'services',
'call',
'08718711108',
'accordingly',
'repeat',
'text',
'word',
'ok',
'mobile',
'phone',
'send',
'adult',
'18',
'content',
'video',
'shortly',
'asked',
'3mobile',
'0870',
'chatlines',
'inclu',
'free',
'mins',
'india',
'cust',
'servs',
'sed',
'yes',
'l8er',
'got',
'mega',
'bill',
'3',
'dont',
'giv',
'shit',
'bailiff',
'due',
'days',

'250',
'3',
'want',
'800',
'awarded',
'sipix',
'digital',
'camera',
'call',
'09061221061',
'landline',
'delivery',
'within',
'28days',
'cs',
'box177',
'm221bp',
'2yr',
'warranty',
'150ppm',
'16',
'p',
'p',
'3',
'99',
'awarded',
'sipix',
'digital',
'camera',
'call',
'09061221061',
'landline',
'delivery',
'within',
'28days',
'cs',
'box177',
'm221bp',
'2yr',
'warranty',
'150ppm',
'16',
'p',
'p',
'3',
'99',
'b4u',
'voucher',
'w',
'c',
'27',
'03',
'marsms',
'log',
'onto',
'www',
'b4utele',
'com',
'discount',
'credit',
'opt',
'reply',
'stop',
'customer',
'care',
'call',
'08717168528',
'babe',
'u',
'want',
'dont',
'u',

```
    u ,
'baby',
'im',
'nasty',
'thing',
'4',
'filthyguys',
'fancy',
'rude',
'time',
'sexy',
'bitch',
'go',
'slo',
'n',
'hard',
'txt',
'xxx',
'slo',
'4msgs',
'back',
'2',
'work',
'2morro',
'half',
'term',
'u',
'c',
'2nite',
'4',
'sexy',
'passion',
'b4',
'2',
'go',
'back',
'chat',
'09099726481',
'luv',
'dena',
'calls',
'1',
'minmobsmorelkpobox177hp51fl',
'bangbabes',
'ur',
'order',
'way',
'u',
'receive',
'service',
'msg',
'2',
'download',
'ur',
'content',
'u',
'goto',
'wap',
'bangb',
'tv',
'ur',
'mobile',
'internet',
'service',
'menu',
'bank',
'granite',
'issues',
'strong',
'buy',
'explosive',
'pick',
'members',
```

'members',
'300',
'nasdaq',
'symbol',
'cdgt',
'5',
'00',
'per',
'bears',
'pic',
'nick',
'tom',
'pete',
'dick',
'fact',
'types',
'try',
'gay',
'chat',
'photo',
'upload',
'call',
'08718730666',
'10p',
'min',
'2',
'stop',
'texts',
'call',
'08712460324',
'bloomberg',
'message',
'center',
'447797706009',
'wait',
'apply',
'future',
'http',
'careers',
'bloomberg',
'com',
'bored',
'housewives',
'chat',
'n',
'date',
'0871750',
'77',
'11',
'bt',
'national',
'rate',
'10p',
'min',
'landlines',
'bored',
'housewives',
'chat',
'n',
'date',
'0871750',
'77',
'11',
'bt',
'national',
'rate',
'10p',
'min',
'landlines',
'bored',
'speed',
'dating',
'try',

```
  'speedchat',
  'txt',
  'speedchat',
  '80155',
  'like',
  'em',
  'txt',
  'swap',
  'get',
  'new',
  'chatter',
  'chat80155',
  'pobox36504w45wq',
  '150p',
  'msg',
  'rcd',
  '16',
  'bought',
  'one',
  'ringtone',
  'getting',
  'texts',
  'costing',
  '3',
  ...]
```

In [0]:

```
resultham = trainInput.flatMap(lambda x: x[0] if x[1]==1 else []).filter(lambda x: x!=[]
)
resultham.collect()
```

```
Out[70]: ['',
  '',
  '1',
  '20',
  'call',
  'cost',
  'guess',
  'isnt',
  'bad',
  'miss',
  'ya',
  'need',
  'ya',
  'want',
  'ya',
  'love',
  'ya',
  '1',
  'cbe',
  '2',
  'chennai',
  '1',
  'go',
  'write',
  'msg',
  '2',
  'put',
  'dictionary',
  'mode',
  '3',
  'cover',
  'screen',
  'hand',
  '4',
  'press',
  'lt',
  'gt',
  '5',
  'gently',
  'remove'
```

```
'remove',
'ur',
'hand',
'interesting',
'1',
'go',
'write',
'msg',
'2',
'put',
'dictionary',
'mode',
'3',
'cover',
'screen',
'hand',
'4',
'press',
'lt',
'gt',
'5',
'gently',
'remove',
'ur',
'hand',
'interesting',
'1',
'go',
'write',
'msg',
'2',
'put',
'dictionary',
'mode',
'3',
'cover',
'screen',
'hand',
'4',
'press',
'lt',
'gt',
'5',
'gently',
'remove',
'ur',
'hand',
'interesting',
'1',
'tension',
'face',
'2',
'smiling',
'face',
'3',
'waste',
'face',
'4',
'innocent',
'face',
'5',
'terror',
'face',
'6',
'cruel',
'face',
'7',
'romantic',
'face',
'8',
'lovable',
'face',
'9',
```

```
'decent',
'face',
'lt',
'gt',
'joker',
'face',
'1',
'thing',
'change',
'sentence',
'want',
'2',
'concentrate',
'educational',
'career',
'im',
'leaving',
'1',
'thing',
'change',
'sentence',
'want',
'2',
'concentrate',
'educational',
'career',
'im',
'leaving',
'1apple',
'day',
'doctor',
'1tulsi',
'leaf',
'day',
'cancer',
'1lemon',
'day',
'fat',
'1cup',
'milk',
'day',
'bone',
'problms',
'3',
'litres',
'watr',
'day',
'diseases',
'snd',
'ths',
'2',
'u',
'care',
'2',
'babe',
'feel',
'lets',
'4get',
'try',
'cheer',
'fit',
'soo',
'muchxxlove',
'u',
'locaxx',
'2',
'half',
'years',
'missed',
'friendship',
'2',
'laptop',
```

```
'laptop',
'noe',
'infra',
'slow',
'lar',
'wan',
'fast',
'one',
'2',
'oz',
'guy',
'kinda',
'flaky',
'one',
'friend',
'interested',
'picking',
'lt',
'gt',
'worth',
'tonight',
'possible',
'2',
'sleeping',
'bags',
'1',
'blanket',
'paper',
'phone',
'details',
'anything',
'else',
'2',
'students',
'solved',
'cat',
'question',
'xam',
'5',
'3',
'2',
'lt',
'gt',
'9',
'2',
'4',
'lt',
'gt',
'8',
'6',
'3',
'lt',
'gt',
'7',
'2',
'5',
'tell',
'answer',
'u',
'r',
'brilliant',
'1thing',
'got',
'd',
'answr',
'2marrow',
'wed',
'lt',
'gt',
'2',
'aha',
'2mro',
'coming'
```

```
'coming',
'gym',
'machan',
'goodnight',
'3',
'pa',
'selected',
'4',
'let',
'go',
'bill',
'millers',
'4',
'oclock',
'mine',
'bash',
'flat',
'plan',
'4',
'tacos',
'1',
'rajas',
'burrito',
'right',
'5',
'2',
'work',
'timing',
'5',
'nights',
'nt',
'staying',
'port',
'step',
'liao',
'ex',
'545',
'lor',
'ya',
'go',
'4',
'dinner',
'together',
'6',
'get',
'ok',
'645',
'7',
'go',
'ok',
'na',
'7',
'lor',
'change',
'2',
'suntec',
'wat',
'time',
'u',
'coming',
'7',
'wonders',
'world',
'7th',
'6th',
'ur',
'style',
'5th',
'ur',
'smile',
'4th',
'ur',
'personality'
```

'personality',
'3rd',
'ur',
'nature',
'2nd',
'ur',
'sms',
'1st',
'ur',
'lovely',
'friendship',
'good',
'morning',
'dear',
'7',
'wonders',
'world',
'7th',
'6th',
'ur',
'style',
'5th',
'ur',
'smile',
'4th',
'ur',
'personality',
'3rd',
'ur',
'nature',
'2nd',
'ur',
'sms',
'1st',
'ur',
'lovely',
'friendship',
'good',
'morning',
'dear',
'7',
'wonders',
'world',
'7th',
'6th',
'ur',
'style',
'5th',
'ur',
'smile',
'4th',
'ur',
'personality',
'3rd',
'ur',
'nature',
'2nd',
'ur',
'sms',
'1st',
'ur',
'lovely',
'friendship',
'good',
'morning',
'dear',
'7',
'wonders',
'world',
'7th',
'6th',
'ur',
'style',

'style',
'5th',
'ur',
'smile',
'4th',
'ur',
'personality',
'3rd',
'ur',
'nature',
'2nd',
'ur',
'sms',
'1st',
'ur',
'lovely',
'friendship',
'good',
'morning',
'dear',
'8',
'min',
'go',
'lunch',
'aah',
'bless',
'arm',
'aah',
'cuddle',
'lush',
'd',
'need',
'lots',
'tea',
'soup',
'kind',
'fumbling',
'aathi',
'dear',
'aathi',
'dear',
'aathi',
'dear',
'abeg',
'make',
'profit',
'start',
'using',
'get',
'sponsors',
'next',
'event',
'abj',
'serving',
'staying',
'dad',
'alone',
'able',
'anything',
'able',
'sleep',
'abt',
'making',
'pics',
'bigger',
'accidentally',
'deleted',
'message',
'resend',
'please',
'account',
'refilled',
'successfully'

```
successfully',
'inr',
'lt',
'decimal',
'gt',
'keralacircle',
'prepaid',
'account',
'balance',
'rs',
'lt',
'decimal',
'gt',
'transaction',
'id',
'kr',
'lt',
'gt',
'actual',
'exam',
'harder',
'nbme',
'actually',
'decided',
'hungry',
'haven',
'left',
'yet',
'v',
'actually',
'fuck',
'whatever',
'find',
'excuse',
'tampa',
'point',
'january',
'though',
'actually',
'getting',
'ready',
'leave',
'house',
'actually',
'm',
'waiting',
'2',
'weeks',
'start',
'putting',
'ad',
'actually',
'mobile',
'full',
'msg',
'm',
'work',
'online',
'need',
'send',
'lt',
'gt',
'sent',
'msg',
'wil',
'explain',
'u',
'later',
'actually',
'nvm',
'got',
'hella',
'cash',
```

'still',
'lt',
'gt',
'ish',
'actually',
'sleeping',
'still',
'might',
'u',
'call',
'back',
'text',
'gr8',
'rock',
'sis',
'send',
'u',
'text',
'wen',
'wake',
'add',
'really',
'care',
'least',
'get',
'dude',
'fuck',
'hey',
'money',
'want',
'adult',
'learn',
'experience',
'real',
'danger',
'dont',
'like',
'peeps',
'using',
'drugs',
'dont',
'need',
'comment',
'agree',
'stop',
'thinkin',
'ipad',
'please',
'ask',
'macho',
'question',
'ah',
'poop',
'looks',
'like',
'ill',
'prob',
'send',
'laptop',
'get',
'fixed',
'cuz',
'gpu',
'problem',
'ah',
'poor',
'baby',
'hope',
'urfeeling',
'bettersn',
'luv',
'probthat'

'overdose',
'work',
'hey',
'go',
'careful',
'spk',
'2',
'u',
'sn',
'lots',
'lovejen',
'xxx',
'ah',
'see',
'lingo',
'let',
'know',
'wot',
'earth',
'finished',
'making',
'ah',
'well',
'confuses',
'things',
'doesn',
'ahhh',
'work',
'vaguely',
'remember',
'feel',
'like',
'lol',
'ahhhh',
'woken',
'bad',
'dream',
'u',
'tho',
'dont',
'like',
'u',
'right',
'didnt',
'know',
'anything',
'comedy',
'night',
'guess',
'im',
'aight',
'8',
'latest',
'probably',
'closer',
'7',
'jay',
'tyler',
'two',
'trips',
'aight',
'close',
'still',
'around',
'alex',
'place',
'aight',
'fuck',
'll',
'get',
'later',
'aight'

```
  'aight',
  'lemme',
  'know',
  'aight',
  'll',
  'head',
  'aight',
  'll',
  'hit',
  'get',
  'cash',
  'aight',
  'll',
  'text',
  'm',
  'back',
  'aight',
  'll',
  'work',
  'thanks',
  'aight',
  'm',
  'chillin',
  'friend',
  'room',
  'text',
  're',
  'way',
  'aight',
  'plan',
  'come',
  'later',
  'tonight',
  'aight',
  'see',
  'bit',
  'aight',
  'sorry',
  'take',
  'ten',
  'years',
  'shower',
  'plan',
  'aight',
  'sounds',
  'good',
  'want',
  'come',
  'aight',
  'text',
  'address',
  'aight',
  'text',
  'tonight',
  'll',
  'see',
  'aight',
  'thanks',
  'comin',
  'aight',
  'tomorrow',
  'around',
  'lt',
  'gt',
  'aight',
  've',
  'set',
  'free',
  'think',
  'text',
  'blake',
  'address',
  'occurs'
```

'occurs',
'm',
'quite',
'sure',
'm',
'thought',
'aight',
'yo',
'dats',
'straight',
'dogg',
'aiya',
'discuss',
'later',
'lar',
'pick',
'4',
'aiyah',
'e',
'rain',
'like',
'quite',
'big',
'leh',
'drizzling',
'least',
'run',
'home',
'aiyah',
'ok',
'wat',
'long',
'got',
'improve',
'already',
'wat',
'aiyah',
'sorry',
'lor',
'watch',
'tv',
'watch',
'forgot',
'2',
'check',
'phone',
'aiyah',
'wait',
'lor',
'u',
'entertain',
'hee',
'aiyar',
'dun',
'disturb',
'u',
'liao',
'thk',
'u',
'lots',
'2',
'aft',
'ur',
'cupboard',
'come',
'aiyar',
'hard',
'2',
'type',
'u',
'later',
'free',
'tell'

```
  'call',
  'n',
  'scold',
  'n',
  'tell',
  'u',
  'aiyar',
  'u',
  'poor',
  'thing',
  'give',
  'u',
  'support',
  'k',
  'jia',
  'll',
  'think',
  'u',
  'aiyo',
  'bit',
  'pai',
  'seh',
  'noe',
  'scared',
  'dun',
  'rem',
  'die',
  'hee',
  'become',
  'better',
  'lookin',
  'oredi',
  'leh',
  'aiyo',
  'cos',
  'sms',
  'neva',
  'reply',
  'wait',
  '4',
  'reply',
  'lar',
  'tot',
  'havent',
  'finish',
  'ur',
  'lab',
  'wat',
  'aiyo',
  'lesson',
  'early',
  'm',
  'still',
  'sleepin',
  'haha',
  'okie',
  'u',
  'go',
  'home',
  'liao',
  'den',
  'confirm',
  'w',
  'lor',
  'aiyo',
  'please',
  'got',
  'time',
  'meh',
  'aiyo',
  'u',
  'poor'
```

'thing',
    'u',
    'dun',
    'wan',
    '2',
    'eat',
    'u',
    'bathe',
    'already',
    'al',
    'moan',
    'n',
    'e',
    'thin',
    'goes',
    'wrong',
    'fault',
    'al',
    'de',
    'arguments',
    'r',
    'fault',
    'fed',
    'himso',
    'y',
    'bother',
    'hav',
    '2go',
    'thanx',
    'xx',
    'aldrine',
    'rakhesh',
    'ex',
    'rtm',
    'pls',
    'call',
    'urgent',
    'alex',
    'says',
    'ok',
    'ok',
    'allo',
    'braved',
    'buses',
    'taken',
    'trains',
    'triumphed',
    'mean',
    're',
    'b',
    'ham',
    'jolly',
    'good',
    'rest',
    'week',
    'almost',
    'see',
    'u',
    'sec',
    'already',
    'one',
    'guy',
    'loving',
    'alright',
    'babe',
    'justthought',
    'd',
    'sayhey',
    'u',
    'doin',
    'nearly',
    'endof'

```
'wk',
'offdam',
'nevamind',
'2hook',
'sn',
'uwant',
'm8',
'lovejen',
'x',
'alright',
'll',
'bring',
'see',
'like',
'lt',
'gt',
'mins',
'alright',
'll',
'head',
'minutes',
'text',
'meet',
'alright',
...]
```

In [0]:

```
trainham = resultham.map(lambda x : (x,1)).reduceByKey(lambda x,y : x + y).filter(lambda
x: x[0]!='')
trainham.take(10)
```

```
Out[71]: [('1', 41),
 ('20', 5),
 ('call', 163),
 ('cost', 8),
 ('guess', 24),
 ('isnt', 5),
 ('bad', 26),
 ('miss', 44),
 ('ya', 43),
 ('need', 115)]
```

In [0]:

```
trainspam = resultspam.map(lambda x : (x,1)).reduceByKey(lambda x,y : x + y).filter(lamb
da x: x[0]!='')
trainspam.take(10)
```

```
Out[72]: [('07732584351', 1),
 ('rodger', 1),
 ('burns', 1),
 ('msg', 37),
 ('tried', 13),
 ('call', 256),
 ('re', 10),
 ('reply', 75),
 ('sms', 32),
 ('free', 164)]
```

In [0]:

```
# Vocabulary Count
trainHamCount = trainham.count()
trainSpamCount = trainspam.count()

v = trainHamCount + trainSpamCount

print("trainHamCount: ", trainHamCount)
print("trainSpamCount: ", trainSpamCount)
print("v: ", v)
```

```
trainHamCount:  5602
trainSpamCount:  2326
v:  7928
```

In [0]:

```
trainHamProbability = trainham.map(lambda x : (x[0], (float)(x[1] + 1) / (float)(v + tra
inHamCount)))
trainSpamProbability = trainspam.map(lambda x : (x[0], (float)(x[1] + 1) / (float)(v + t
rainSpamCount)))
trainHamProbability.take(10)
```

```
Out[76]: [('1', 0.0031042128603104213),
 ('20', 0.0004434589800443459),
 ('call', 0.012121212121212121),
 ('cost', 0.0006651884700665188),
 ('guess', 0.0018477457501847746),
 ('isnt', 0.0004434589800443459),
 ('bad', 0.0019955654101995565),
 ('miss', 0.0033259423503325942),
 ('ya', 0.0032520325203252032),
 ('need', 0.008573540280857354)]
```

In [0]:

```
trainSpamProbability.take(10)
```

```
Out[77]: [('07732584351', 0.00019504583577140063),
 ('rodger', 0.00019504583577140063),
 ('burns', 0.00019504583577140063),
 ('msg', 0.0037058708796567194),
 ('tried', 0.001365320850399844),
 ('call', 0.025063389896625706),
 ('re', 0.0010727520967427347),
 ('reply', 0.007411741759313439),
 ('sms', 0.0032182562902282035),
 ('free', 0.01609128145114102)]
```

In [0]:

```
from pyspark.sql.functions import monotonically_increasing_id

#adding id column for easy access
testDataFrame = test.withColumn("id", monotonically_increasing_id())
testDataFrame.sort(testDataFrame.id.desc()).collect()
testData = testDataFrame.rdd.map(lambda x:(x["text"],x["label"],x["id"]))
```

In [0]:

```
testDataFrame.display()
```

text	label	id
List()	1	0
List()	1	1
List(08714712388, 10am, 7pm, cost, 10p)	0	2
List(1, finish, meeting, call)	1	3
List(1, new, message, please, call, 08715205273)	0	4
List(1, new, voicemail, please, call, 08719181513)	0	5
List(1, number, 2, gonna, massive, pain, ass, d, rather, get, involved, possible)	1	6
List(1, reach, home, call)	1	7
List(10, min, later, k)	1	8
List(1000, flirting, txt, girl, bloke, ur, name, age, eg, girl, zoe, 18, 8007, join, get, chatting)	0	9

In [0]:

```
testData.collect()
```

```
Out[91]: [([], 1, 0),
 ([], 1, 1),
 (['08714712388', '10am', '7pm', 'cost', '10p'], 0, 2),
 (['1', 'finish', 'meeting', 'call'], 1, 3),
 (['1', 'new', 'message', 'please', 'call', '08715205273'], 0, 4),
 (['1', 'new', 'voicemail', 'please', 'call', '08719181513'], 0, 5),
 (['1',
   'number',
   '2',
   'gonna',
   'massive',
   'pain',
   'ass',
   'd',
   'rather',
   'get',
   'involved',
   'possible'],
  1,
  6),
 (['1', 'reach', 'home', 'call'], 1, 7),
 (['10', 'min', 'later', 'k'], 1, 8),
 (['1000',
   'flirting',
   'txt',
   'girl',
   'bloke',
   'ur',
   'name',
   'age',
   'eg',
   'girl',
   'zoe',
   '18',
   '8007',
   'join',
   'get',
   'chatting'],
  0,
  9),
 (['1000',
   'girls',
   'many',
   'local',
   '2',
   'u',
   'r',
   'virgins',
   '2',
   'r',
   'ready',
   '2',
   '4fil',
   'ur',
   'every',
   'sexual',
   'need',
   'u',
   '4fil',
   'text',
   'cute',
   '69911',
   '1',
   '50p',
   'm'],
  0,
  10),
 (['1000',
   'winner',
   'guaranteed'
```

 'guaranteed',
  'caller',
  'prize',
  'final',
  'attempt',
  'contact',
  'claim',
  'call',
  '09071517866',
  '150ppmpobox10183bhamb64xe'],
 0,
 11),
([' 140',
  'ard',
  'rest',
  'ard',
  '180',
  'least',
  'price',
  '4',
  '2',
  'bedrm',
  '900'],
 1,
 12),
(['18',
  'days',
  'euro2004',
  'kickoff',
  'u',
  'kept',
  'informed',
  'latest',
  'news',
  'results',
  'daily',
  'unsubscribe',
  'send',
  'get',
  'euro',
  'stop',
  '83222'],
 0,
 13),
(['2', '2', '146tf150p'], 0, 14),
(['2', 'celebrate', 'b', 'day', 'y', 'else'], 1, 15),
(['2',
  'docs',
  'appointments',
  'next',
  'week',
  'm',
  'tired',
  'shoving',
  'stuff',
  'ugh',
  'couldn',
  'normal',
  'body'],
 1,
 16),
(['2004',
  'account',
  '07xxxxxxxxx',
  'shows',
  '786',
  'unredeemed',
  'points',
  'claim',
  'call',
  '08719181259',
  'identifier',
  'code'

   code ,
   'xxxxx',
   'expires',
   '26',
   '03',
   '05'],
  0,
  17),
 (['25p',
   '4',
   'alfie',
   'moon',
   'children',
   'need',
   'song',
   'ur',
   'mob',
   'tell',
   'ur',
   'm8s',
   'txt',
   'tone',
   'charity',
   '8007',
   'nokias',
   'poly',
   'charity',
   'polys',
   'zed',
   '08701417012',
   'profit',
   '2',
   'charity'],
  0,
  18),
 (['26th', 'july'], 1, 19),
 (['2nd',
   'time',
   'tried',
   '2',
   'contact',
   'u',
   'u',
   'won',
   '750',
   'pound',
   'prize',
   '2',
   'claim',
   'easy',
   'call',
   '08712101358',
   '10p',
   'per',
   'min',
   'bt',
   'national',
   'rate'],
  0,
  20),
 (['4mths',
   'half',
   'price',
   'orange',
   'line',
   'rental',
   'latest',
   'camera',
   'phones',
   '4',
   'free',
   'phone',
   '11mths'

```
   'imths',
   'call',
   'mobilesdirect',
   'free',
   '08000938767',
   'update',
   'or2stoptxt'],
 0,
 21),
([' 5p',
  '4',
  'alfie',
  'moon',
  'children',
  'need',
  'song',
  'ur',
  'mob',
  'tell',
  'ur',
  'm8s',
  'txt',
  'tone',
  'charity',
  '8007',
  'nokias',
  'poly',
  'charity',
  'polys',
  'zed',
  '08701417012',
  'profit',
  '2',
  'charity'],
 0,
 22),
(['6', 'ft', 'good', 'combination'], 1, 23),
(['6times'], 1, 24),
(['7', 'esplanade', 'mind', 'giving', 'lift', 'cos', 'got', 'car', 'today'],
 1,
 25),
(['8',
  'latest',
  'g',
  'still',
  'scrounge',
  'ammo',
  'want',
  'give',
  'new',
  'ak',
  'try'],
 1,
 26),
(['87077',
  'kick',
  'new',
  'season',
  '2wks',
  'free',
  'goals',
  'news',
  'ur',
  'mobile',
  'txt',
  'ur',
  'club',
  'name',
  '87077',
  'eg',
  'villa',
  '87077'],
 0,
```

 0,
 27),
(['aaooooright', 'work'], 1, 28),
(['absolutely',
  'love',
  'south',
  'park',
  'recently',
  'started',
  'watching',
  'office'],
 1,
 29),
(['accidentally', 'brought', 'em', 'home', 'box'], 1, 30),
(['account',
  'credited',
  '500',
  'free',
  'text',
  'messages',
  'activate',
  'txt',
  'word',
  'credit',
  '80488',
  'cs',
  'www',
  '80488',
  'biz'],
 0,
 31),
(['account', 'number'], 1, 32),
(['account',
  'refilled',
  'successfully',
  'inr',
  'lt',
  'decimal',
  'gt',
  'keralacircle',
  'prepaid',
  'account',
  'balance',
  'rs',
  'lt',
  'decimal',
  'gt',
  'transaction',
  'id',
  'kr',
  'lt',
  'gt'],
 1,
 33),
(['actually',
  'deleted',
  'old',
  'website',
  'm',
  'blogging',
  'magicalsongs',
  'blogspot',
  'com'],
 1,
 34),
(['actually',
  'first',
  'time',
  'went',
  'bed',
  'long',
  'spoke',
  'woke',

```
  woke ',
  '7',
  'night'],
 1,
 35),
(['actually', 'send', 'reminder', 'today', 'wonderful', 'weekend'], 1, 36),
(['affidavit',
  'says',
  'lt',
  'gt',
  'e',
  'twiggs',
  'st',
  'division',
  'g',
  'courtroom',
  'lt',
  'gt',
  'lt',
  'time',
  'gt',
  'll',
  'double',
  'check',
  'text',
  'tomorrow'],
 1,
 37),
(['aft',
  'finish',
  'lunch',
  'go',
  'str',
  'lor',
  'ard',
  '3',
  'smth',
  'lor',
  'u',
  'finish',
  'ur',
  'lunch',
  'already'],
 1,
 38),
(['ages', 'abj'], 1, 39),
(['ah',
  'well',
  'confuses',
  'things',
  'doesnt',
  'thought',
  'friends',
  'maybe',
  'wrong',
  'thing',
  'already',
  'sort',
  'invited',
  'tho',
  'may',
  'come',
  'cos',
  'money'],
 1,
 40),
(['aight', 'call', 're', 'close'], 1, 41),
(['aight', 'ill', 'get', 'fb', 'couple', 'minutes'], 1, 42),
(['aight', 'let', 'know', 're', 'gonna', 'around', 'usf'], 1, 43),
(['aight', 'll', 'ask', 'roommates'], 1, 44),
(['aight', 'll', 'grab', 'something', 'eat', 'text', 're', 'back', 'mu'],
 1,
 45)
```

45),
(['aight', 'pick', 'open', 'tonight'], 1, 46),
(['aight', 'rush', 'll', 'ask', 'jay'], 1, 47),
(['aight', 'still', 'want', 'get', 'money'], 1, 48),
(['aight',
  'text',
  're',
  'back',
  'mu',
  'll',
  'swing',
  'need',
  'somebody',
  'get',
  'door'],
 1,
 49),
(['aight', 'time', 'want', 'come'], 1, 50),
(['aight', 'well', 'keep', 'informed'], 1, 51),
(['ain',
  'answerin',
  'phone',
  'actually',
  'pretty',
  'reasonable',
  'hour',
  'm',
  'sleepy'],
 1,
 52),
(['aiyah', 'u', 'ok', 'already', 'lar', 'e', 'nydc', 'wheellock'], 1, 53),
(['aiyar', 'sorry', 'lor', 'forgot', '2', 'tell', 'u'], 1, 54),
(['aiyo',
  'u',
  'always',
  'c',
  'ex',
  'one',
  'dunno',
  'abt',
  'mei',
  'haven',
  'reply',
  'first',
  'time',
  'u',
  'reply',
  'fast',
  'y',
  'lucky',
  'workin',
  'huh',
  'got',
  'bao',
  'ur',
  'sugardad',
  'ah',
  'gee'],
 1,
 55),
(['alex',
  'knows',
  'guy',
  'sells',
  'mids',
  'south',
  'tampa',
  'think',
  'set',
  'like',
  '8'],
 1,
 56)

```
56),
(['already', 'sabarish', 'asked', 'go'], 1, 57),
(['already', 'squatting', 'new', 'way', 'walking'], 1, 58),
(['alright', 'm', 'good', 'night'], 1, 59),
(['alright', 're', 'hooked', 'guys'], 1, 60),
(['alright', 'took', 'morphine', 'back', 'yo'], 1, 61),
(['alright',
  'tyler',
  'got',
  'minor',
  'crisis',
  'home',
  'sooner',
  'thought',
  'asap'],
 1,
 62),
(['alrite',
  'hunny',
  'wot',
  'u',
  '2',
  '2nite',
  'didnt',
  'end',
  'goin',
  'town',
  'jus',
  'da',
  'pub',
  'instead',
  'jus',
  'chillin',
  'da',
  'mo',
  'bedroom',
  'love',
  'jen',
  'xxx'],
 1,
 63),
(['also', 'piece'], 1, 64),
(['also', 'remember', 'beads', 'come', 'ever'], 1, 65),
(['also',
  've',
  'sorta',
  'blown',
  'couple',
  'times',
  'recently',
  'id',
  'rather',
  'text',
  'blue',
  'looking',
  'weed'],
 1,
 66),
(['always', 'celebrate', 'ny', 'family'], 1, 67),
(['always', 'chat', 'fact', 'need', 'money', 'raise'], 1, 68),
(['always', 'saying', 'welp'], 1, 69),
(['amazing',
  'quote',
  'sometimes',
  'life',
  'difficult',
  'decide',
  'whats',
  'wrong',
  'lie',
  'brings',
  'smile',
  'truth'
```

```
  truth ,
  'brings',
  'tear'],
 1,
 70),
(['angry', 'happen', 'dear'], 1, 71),
(['annoying', 'isn'], 1, 72),
(['anybody',
  'number',
  'still',
  'haven',
  'thought',
  'tactful',
  'way',
  'ask',
  'alex'],
 1,
 73),
(['anyone',
  'calls',
  'treadmill',
  'say',
  'll',
  'buy',
  'make',
  'sure',
  'working',
  'found',
  'ad',
  'craigslist',
  'selling',
  'lt',
  'gt'],
 1,
 74),
(['anything', 'lor'], 1, 75),
(['anything', 'lor'], 1, 76),
(['anything', 'lor', 'go', 'go', 'lor'], 1, 77),
(['anything', 'lor', 'u', 'decide'], 1, 78),
(['anything', 'special'], 1, 79),
(['anything', 'specific', 'regards', 'jaklin', 'idk', 'fuck'], 1, 80),
(['anyway', 'many', 'good', 'evenings', 'u'], 1, 81),
(['apparently', 'bffs', 'carly', 'quick'], 1, 82),
(['ard', '515', 'like', 'dat', 'y'], 1, 83),
(['ard', '6', 'like', 'dat', 'lor'], 1, 84),
(['argh',
  '3g',
  'spotty',
  'anyway',
  'thing',
  'remember',
  'research',
  'province',
  'sterling',
  'problem',
  'free',
  'places',
  'looked'],
 1,
 85),
(['arms', 'fine', 'cardiff', 'uni'], 1, 86),
(['arrived', 'see', 'couple', 'days', 'lt', '3'], 1, 87),
(['arun', 'u', 'transfr', 'd', 'amt'], 1, 88),
(['ask', 'carlos', 'get', 'anybody', 'else', 'chip'], 1, 89),
(['ask', 'say', 'please', 'message'], 1, 90),
(['asked', 'anthony', 'dad', 'bf'], 1, 91),
(['asked', 'call', 'ok'], 1, 92),
(['asking', 'u', 'knw', 'nt', 'may', 'ur', 'frnds', 'classmates'], 1, 93),
(['aslamalaikkum',
  'insha',
  'allah',
  'tohar',
  'bessr',
```

   'beeen',
   'muht',
   'albi',
   'mufti',
   'mahfuuz',
   'meaning'],
  1,
  94),
 (['attended', 'nothing'], 1, 95),
 (['auction',
   'round',
   '4',
   'highest',
   'bid',
   '54',
   'next',
   'maximum',
   'bid',
   '71',
   'bid',
   'send',
   'bids',
   'e',
   'g',
   '10',
   'bid',
   '10',
   '83383',
   'good',
   'luck'],
  0,
  96),
 (['auntie', 'huai', 'juan', 'never', 'pick', 'phone'], 1, 97),
 (['awesome', 'lemme', 'know', 'whenever', 're', 'around'], 1, 98),
 (['awesome',
   'remember',
   'last',
   'time',
   'got',
   'somebody',
   'high',
   'first',
   'time',
   'diesel',
   'v'],
  1,
  99),
 (['aww',
   'must',
   'nearly',
   'dead',
   'well',
   'jez',
   'iscoming',
   'todo',
   'workand',
   'whilltake',
   'forever'],
  1,
  100),
 (['b',
   'floppy',
   'b',
   'snappy',
   'happy',
   'gay',
   'chat',
   'service',
   'photo',
   'upload',
   'call',
   '08718730666',
   '10p',

```
   'min',
   '2',
   'stop',
   'texts',
   'call',
   '08712460324'],
 0,
 101),
(['b', 'late', 'love', 'mum'], 1, 102),
(['b4', 'thursday'], 1, 103),
(['babe',
   'goes',
   'day',
   'sip',
   'cappuccino',
   'think',
   'love',
   'send',
   'kiss',
   'across',
   'sea'],
 1,
 104),
(['babe', 'lost'], 1, 105),
(['babe', 'need', 'advice'], 1, 106),
(['babe', 'talking', 'think', 'good', 'boy', 'missing', 'love'], 1, 107),
(['babes',
   'hope',
   'ok',
   'shit',
   'nights',
   'sleep',
   'fell',
   'asleep',
   '5',
   'm',
   'knackered',
   'm',
   'dreading',
   'work',
   'tonight',
   'thou',
   'upto',
   'tonight',
   'x'],
 1,
 108),
(['bad', 'girl', 'still', 'remember'], 1, 109),
(['badrith', 'chennai', 'surely', 'pick', 'us', 'competition'], 1, 110),
(['battery', 'mr', 'adewale', 'uncle', 'aka', 'egbon'], 1, 111),
(['bbq', 'sat', 'mine', '6ish', 'ur', 'welcome', '2', 'come'], 1, 112),
(['beauty',
   'life',
   'next',
   'second',
   'hides',
   'thousands',
   'secrets',
   'wish',
   'every',
   'second',
   'wonderful',
   'ur',
   'life',
   'gud',
   'n8'],
 1,
 113),
(['beauty', 'sleep', 'help', 'ur', 'pimples'], 1, 114),
(['becoz',
   'lt',
```

         'gt',
         'jan',
         'whn',
         'al',
         'post',
         'ofice',
         'holiday',
         'cn',
         'go',
         'fr',
         'post',
         'ofice',
         'got',
         'duffer'],
  1,
  115),
 (['big',
   'brother',
   'alert',
   'computer',
   'selected',
   'u',
   '10k',
   'cash',
   '150',
   'voucher',
   'call',
   '09064018838',
   'ntt',
   'po',
   'box',
   'cro1327',
   '18',
   'bt',
   'landline',
   'cost',
   '150ppm',
   'mobiles',
   'vary'],
  0,
  116),
 (['big',
   'difference',
   'lt',
   'gt',
   'versus',
   'lt',
   'gt',
   'every',
   'lt',
   'gt',
   'hrs'],
  1,
  117),
 (['birthday', 'feb', 'lt', 'gt', 'da'], 1, 118),
 (['bit',
   'ur',
   'smile',
   'hppnss',
   'drop',
   'ur',
   'tear',
   'sorrow',
   'part',
   'ur',
   'heart',
   'life',
   'heart',
   'like',
   'mine',
   'wil',
   'care',
   'u',

    u ,
    'forevr',
    'goodfriend'],
   1,
   119),
  (['block',
    'breaker',
    'comes',
    'deluxe',
    'format',
    'new',
    'features',
    'great',
    'graphics',
    'mobile',
    'buy',
    '5',
    'replying',
    'get',
    'bbdeluxe',
    'take',
    'challenge'],
   0,
   120),
  (['bloody',
    'hell',
    'cant',
    'believe',
    'forgot',
    'surname',
    'mr',
    'ill',
    'give',
    'u',
    'clue',
    'spanish',
    'begins',
    'm'],
   1,
   121),
  (['bloomberg',
    'message',
    'center',
    '447797706009',
    'wait',
    'apply',
    'future',
    'http',
    'careers',
    'bloomberg',
    'com'],
   0,
   122),
  (['boltblue',
    'tones',
    '150p',
    'reply',
    'poly',
    'mono',
    'eg',
    'poly3',
    '1',
    'cha',
    'cha',
    'slide',
    '2',
    'yeah',
    '3',
    'slow',
    'jamz',
    '6',
    'toxic',
    '8',

```
          o ,
 'come',
 'stop',
 '4',
 'tones',
 'txt'],
0,
123),
(['boo', 'm', 'way', 'moms', 'making', 'tortilla', 'soup', 'yummmm'], 1, 124),
(['boo', 'things', 'm', 'back', 'home', 'little', 'bored', 'already'],
 1,
 125),
(['booked', 'ticket', 'pongal'], 1, 126),
(['bookedthe', 'hut', 'also', 'time', 'way'], 1, 127),
(['borderline', 'yeah', 'whatever'], 1, 128),
(['bothering', 'trust', 'answers', 'pls'], 1, 129),
(['boy',
 'love',
 'u',
 'grl',
 'hogolo',
 'boy',
 'gold',
 'chain',
 'kodstini',
 'grl',
 'agalla',
 'boy',
 'necklace',
 'madstini',
 'grl',
 'agalla',
 'boy',
 'hogli',
 '1',
 'mutai',
 'eerulli',
 'kodthini',
 'grl',
 'love',
 'u',
 'kano'],
 1,
 130),
(['boy',
 'loved',
 'gal',
 'propsd',
 'bt',
 'didnt',
 'mind',
 'gv',
 'lv',
 'lttrs',
 'bt',
 'frnds',
 'threw',
 'thm',
 'd',
 'boy',
 'decided',
 '2',
 'aproach',
 'd',
 'gal',
 'dt',
 'time',
 'truck',
 'speeding',
 'towards',
 'd',
 'gal',
 'un'
```

```
   wıı ,
  '2',
  'hit',
  'd',
  'girl',
  'd',
  'boy',
  'ran',
  'like',
  'hell',
  'n',
  'saved',
  'asked',
  'hw',
  'cn',
  'u',
  'run',
  'fast',
  'd',
  'boy',
  'replied',
  'boost',
  'd',
  'secret',
  'energy',
  'n',
  'instantly',
  'd',
  'girl',
  'shouted',
  'energy',
  'n',
  'thy',
  'lived',
  'happily',
  '2gthr',
  'drinking',
  'boost',
  'evrydy',
  'moral',
  'd',
  'story',
  'hv',
  'free',
  'msgs',
  'd',
  'gud',
  'ni8'],
 1,
 131),
(['bring', 'got'], 1, 132),
(['bruce', 'b', 'downs', 'amp', 'fletcher'], 1, 133),
(['bruv', 'hope', 'great', 'break', 'rewarding', 'semester'], 1, 134),
(['btw',
  'regarding',
  'really',
  'try',
  'see',
  'anyone',
  'else',
  '4th',
  'guy',
  'commit',
  'random',
  'dude'],
 1,
 135),
(['buy',
  'blackberry',
  'bold',
  '2',
  'torch',
  'buy'
```

```
  'new',
  'used',
  'let',
  'know',
  'plus',
  'saying',
  'buy',
  'lt',
  'gt',
  'g',
  'wifi',
  'ipad',
  'saying',
  'lt',
  'gt',
  'g'],
 1,
 136),
([['buy',
  'space',
  'invaders',
  '4',
  'chance',
  '2',
  'win',
  'orig',
  'arcade',
  'game',
  'console',
  'press',
  '0',
  'games',
  'arcade',
  'std',
  'wap',
  'charge',
  'see',
  'o2',
  'co',
  'uk',
  'games',
  '4',
  'terms',
  'settings',
  'purchase'],
 0,
 137),
([['buzz',
  'hey',
  'love',
  'think',
  'hope',
  'day',
  'goes',
  'well',
  'sleep',
  'miss',
  'babe',
  'long',
  'moment',
  'together',
  'loving',
  'smile'],
 1,
 138),
([['call'], 1, 139),
([['call',
  '09095350301',
  'send',
  'girls',
  'erotic',
  'ecstacy'
```

   ecstacy',
   '60p',
   'min',
   'stop',
   'texts',
   'call',
   '08712460324',
   'nat',
   'rate'],
 0,
 140),
(['call', 'carlos', 'phone', 'vibrate', 'acting', 'might', 'hear', 'texts'],
 1,
 141),
(['call', 'freephone', '0800', '542', '0578'], 0, 142),
(['call',
  'germany',
  '1',
  'pence',
  'per',
  'minute',
  'call',
  'fixed',
  'line',
  'via',
  'access',
  'number',
  '0844',
  '861',
  '85',
  '85',
  'prepayment',
  'direct',
  'access',
  'www',
  'telediscount',
  'co',
  'uk'],
 0,
 143),
(['call', 'get', 'chance', 'plz', 'lt', '3'], 1, 144),
(['call', 'senthil', 'hsbc'], 1, 145),
(['call', 'tell', 'infront', 'call'], 1, 146),
(['call', 'times', 'job', 'today', 'ok', 'umma', 'ask', 'speed'], 1, 147),
(['called', 'said', 'choose', 'future'], 1, 148),
(['callin', 'say', 'hi', 'take', 'care', 'bruv'], 1, 149),
(['calls', 'messages', 'missed', 'calls'], 1, 150),
(['calls', 'messages', 'missed', 'calls'], 1, 151),
(['calls', 'messages', 'missed', 'calls'], 1, 152),
(['came', 'hostel'], 1, 153),
(['camera',
  'awarded',
  'sipix',
  'digital',
  'camera',
  'call',
  '09061221066',
  'fromm',
  'landline',
  'delivery',
  'within',
  '28',
  'days'],
 0,
 154),
(['campus', 'library'], 1, 155),
(['cant', 'pick', 'phone', 'right', 'pls', 'send', 'message'], 1, 156),
(['cant', 'pick', 'phone', 'right', 'pls', 'send', 'message'], 1, 157),
(['cant', 'pick', 'phone', 'right', 'pls', 'send', 'message'], 1, 158),
(['captain', 'room'], 1, 159),
(['career',
  'tel',
  'added'

   'u',
   'contact',
   'indyarocks',
   'com',
   'send',
   'free',
   'sms',
   'remove',
   'phonebook',
   'sms',
   'lt',
   'gt'],
 1,
 160),
(['carlos', 'll', 'minute', 'still', 'need', 'buy'], 1, 161),
(['carlos', 'says', 'll', 'mu', 'lt', 'gt', 'minutes'], 1, 162),
(['carlos', 'took', 'leave', 'minute'], 1, 163),
(['case', 'guess', 'll', 'see', 'campus', 'lodge'], 1, 164),
(['cha',
   'quiteamuzing',
   'scool',
   'babe',
   'probpop',
   'cu',
   'satthen',
   'hunny',
   '4brekkie',
   'love',
   'jen',
   'xxx',
   'psxtra',
   'lrg',
   'portions',
   '4',
   'please'],
 1,
 165),
(['chance',
   'might',
   'evaporated',
   'soon',
   'violated',
   'privacy',
   'stealing',
   'phone',
   'number',
   'employer',
   'paperwork',
   'cool',
   'please',
   'contact',
   'report',
   'supervisor'],
 1,
 166),
(['change', 'e', 'one', 'next', 'escalator'], 1, 167),
(['change', 'windows', 'logoff', 'sound'], 1, 168),
(['charges', 'transfer', 'charges', 'withdraw', 'anyhow', 'like'], 1, 169),
(['check', 'audrey', 'status', 'right'], 1, 170),
(['check', 'nuerologist'], 1, 171),
(['check', 'rooms', 'befor', 'activities'], 1, 172),
(['check', 'rooms', 'befor', 'activities'], 1, 173),
(['cheers', 'card', 'time', 'year', 'already'], 1, 174),
(['cheers',
   'lou',
   'yeah',
   'goodnite',
   'shame',
   'u',
   'neva',
   'came',
   'c',

```
   c ,
  'ya',
  'gailxx'],
 1,
 175),
(['cheers',
  'message',
  'zogtorius',
  've',
  'staring',
  'phone',
  'age',
  'deciding',
  'whether',
  'text'],
 1,
 176),
(['cheers',
  'u',
  'tex',
  'mecause',
  'u',
  'werebored',
  'yeah',
  'okden',
  'hunny',
  'r',
  'uin',
  'wk',
  'sat',
  'sound',
  'likeyour',
  'havin',
  'gr8fun',
  'j',
  'keep',
  'updat',
  'countinlots',
  'loveme',
  'xxxxx'],
 1,
 177),
(['chosen',
  'receive',
  '350',
  'award',
  'pls',
  'call',
  'claim',
  'number',
  '09066364311',
  'collect',
  'award',
  'selected',
  'receive',
  'valued',
  'mobile',
  'customer'],
 0,
 178),
(['class', 'hours', 'sorry'], 1, 179),
(['cme',
  'want',
  'go',
  'hos',
  '2morow',
  'wil',
  'cme',
  'got',
  'dear',
  'didnt',
  'say',
  'time'],
```

 ᴄɪᵐᵉ],
 1,
 180),
(['collecting',
  'ur',
  'laptop',
  'going',
  'configure',
  'da',
  'settings',
  'izzit'],
 1,
 181),
(['come'], 1, 182),
(['come', 'alive', 'better', 'correct', 'good', 'looking', 'figure'], 1, 183),
(['come', 'din', 'c', 'yup', 'cut', 'hair'], 1, 184),
(['come', 'get', 'b', 'quite', 'diff', 'guess', 'rite'], 1, 185),
(['come', 'lt', '25', 'n', 'pass', 'lar'], 1, 186),
(['come', 'mu', 're', 'sorting', 'narcotics', 'situation'], 1, 187),
(['come', 'n', 'pick', '530', 'ar'], 1, 188),
(['come', 'n', 'pick', 'come', 'immediately', 'aft', 'ur', 'lesson'], 1, 189),
(['come', 'room', 'come', 'house', 'cos', 'house', 'still', 'messy', 'haha'],
 1,
 190),
(['come', 'round'], 1, 191),
(['come',
  'slave',
  'going',
  'shell',
  'unconsciously',
  'avoiding',
  'making',
  'unhappy'],
 1,
 192),
(['come', 'tomorrow', 'di'], 1, 193),
(['come', 'u', 'got', 'nothing'], 1, 194),
(['confirm', 'eating', 'esplanade'], 1, 195),
(['congrats',
  '2',
  'mobile',
  '3g',
  'videophones',
  'r',
  'call',
  '09061744553',
  'videochat',
  'wid',
  'ur',
  'mates',
  'play',
  'java',
  'games',
  'dload',
  'polyh',
  'music',
  'noline',
  'rentl',
  'bx420',
  'ip4',
  '5we',
  '150pm'],
 0,
 196),
(['congrats',
  'great',
  'wanted',
  'tell',
  'tell',
  'score',
  'cos',
  'might',
  'make',

    'make',
    'relax',
    'motivating',
    'thanks',
    'sharing'],
 1,
 197),
(['congrats', 'kano', 'whr', 'treat', 'maga'], 1, 198),
(['congratulations',
  'thanks',
  'good',
  'friend',
  'u',
  'won',
  '2',
  '000',
  'xmas',
  'prize',
  '2',
  'claim',
  'easy',
  'call',
  '08718726978',
  '10p',
  'per',
  'minute',
  'bt',
  'national',
  'rate'],
 0,
 199),
(['congratulations',
  'ur',
  'awarded',
  '500',
  'cd',
  'vouchers',
  '125gift',
  'guaranteed',
  'free',
  'entry',
  '2',
  '100',
  'wkly',
  'draw',
  'txt',
  'music',
  '87066',
  'tncs',
  'www',
  'ldew',
  'com1win150ppmx3age16'],
 0,
 200),
(['congratulations',
  'ur',
  'awarded',
  '500',
  'cd',
  'vouchers',
  '125gift',
  'guaranteed',
  'free',
  'entry',
  '2',
  '100',
  'wkly',
  'draw',
  'txt',
  'music',
  '87066',
  'tncs',
  'www',

      www ,
     'ldew',
     'com1win150ppmx3age16'],
    0,
    201),
   (['contract',
     'mobile',
     '11',
     'mnths',
     'latest',
     'motorola',
     'nokia',
     'etc',
     'free',
     'double',
     'mins',
     'text',
     'orange',
     'tariffs',
     'text',
     'yes',
     'callback',
     'remove',
     'records'],
    0,
    202),
   (['convey', 'regards'], 1, 203),
   (['cooked', 'rather', 'nice', 'salmon', 'la'], 1, 204),
   (['cool', 'fun', 'practicing', 'making', 'babies'], 1, 205),
   (['cool', 'let', 'know', 'kicks', 'around', 'lt', 'gt', 'll', 'day'], 1, 206),
   (['cool', 'll', 'night', 'lemme', 'know', 're', 'around'], 1, 207),
   (['cool',
     'shall',
     'go',
     'see',
     'go',
     'tip',
     'anyway',
     'home',
     'got',
     'something',
     'drop',
     'later',
     'lets',
     'go',
     'town',
     'tonight',
     'maybe',
     'mum',
     'take',
     'us'],
    1,
    208),
   (['cool', 'text', 'head'], 1, 209),
   (['cool', 'text', 're', 'parked'], 1, 210),
   (['cos',
     'lar',
     'm',
     'ba',
     'dao',
     'ok',
     '1',
     'pm',
     'lor',
     'y',
     'u',
     'never',
     'ask',
     'go',
     'ah',
     'said',
     'u',
     'ask'

    'ask',
  'fri',
  'said',
  'u',
  'ask',
  'today'],
 1,
 211),
(['cos',
  'shopping',
  'wif',
  'darren',
  'jus',
  'n',
  'called',
  '2',
  'ask',
  'wat',
  'present',
  'wan',
  'lor',
  'started',
  'guessing',
  'wif',
  'n',
  'finally',
  'guessed',
  'darren',
  'lor'],
 1,
 212),
(['course', 'guess', 'god', 'got', 'hold', 'right'], 1, 213),
(['customer',
  'service',
  'annoncement',
  'new',
  'years',
  'delivery',
  'waiting',
  'please',
  'call',
  '07046744435',
  'arrange',
  'delivery'],
 0,
 214),
(['cute',
  'thought',
  'friendship',
  'necessary',
  'share',
  'every',
  'secret',
  'ur',
  'close',
  'frnd'],
 1,
 215),
(['cutest', 'girl', 'world'], 1, 216),
(['d',
  'like',
  'tell',
  'deepest',
  'darkest',
  'fantasies',
  'call',
  '09094646631',
  '60p',
  'min',
  'stop',
  'texts',
  'call',
  '08712460324'

          08712460324 ,
     'nat',
     'rate'],
    0,
    217),
   (['da',
     'birthdate',
     'certificate',
     'april',
     'real',
     'date',
     'today',
     'dont',
     'publish',
     'shall',
     'give',
     'special',
     'treat',
     'keep',
     'secret',
     'way',
     'thanks',
     'wishes'],
    1,
    218),
   (['da', 'car', 'park'], 1, 219),
   (['da', 'run', 'activate', 'full', 'version', 'da'], 1, 220),
   (['da', 'vijay', 'going', 'talk', 'jaya', 'tv'], 1, 221),
   (['da', 'vijay', 'going', 'talk', 'jaya', 'tv'], 1, 222),
   (['dad', 'gonna', 'call', 'gets', 'work', 'ask', 'crazy', 'questions'],
    1,
    223),
   (['dad', 'says', 'hurry', 'hell'], 1, 224),
   (['daily', 'text', 'favour', 'time'], 1, 225),
   (['damn', 'poor', 'zac', 'doesn', 'stand', 'chance'], 1, 226),
   (['dance', 'river'], 1, 227),
   (['dare', 'ask', 'luck', 'sorting', 'car'], 1, 228),
   (['dare', 'stupid', 'wont', 'tell', 'anything', 'hear', 'wont', 'talk'],
    1,
    229),
   (['darren',
     'saying',
     'dat',
     'u',
     'meeting',
     'da',
     'g

*** WARNING: max output size exceeded, skipping output. ***

ful',
     'get',
     'every',
     'morning',
     'thank',
     'god',
     'day',
     'like',
     'say',
     'lot',
     'tell',
     'later'],
    1,
    755),
   (['love', 'u', '2', 'little', 'pocy', 'bell', 'sorry', 'love', 'u'], 1, 756),
   (['love', 'want', 'flood', 'pretty', 'pussy', 'cum'], 1, 757),
   (['love',
     'ya',
     'try',
     'budget',
     'money',
     'better',
     'babe

```
   babe',
 'gary',
 'freak',
 'knew'],
1,
758),
(['loverboy',
 'keeps',
 'coming',
 'queen',
 'hmmm',
 'doesn',
 'ache',
 'speak',
 'miss',
 'desparately'],
1,
759),
(['lt', 'gt', 'guess', 'easy', 'enough'], 1, 760),
(['lt', 'gt', 'mca', 'conform'], 1, 761),
(['lt', 'gt', 'w', 'jetton', 'ave', 'forgot'], 1, 762),
(['lul',
 'im',
 'gettin',
 'juicy',
 'gossip',
 'hospital',
 'two',
 'nurses',
 'talking',
 'fat',
 'gettin',
 'one',
 'thinks',
 'shes',
 'obese',
 'oyea'],
1,
763),
(['m',
 'always',
 'yahoo',
 'messenger',
 'send',
 'message',
 'll',
 'get',
 'may',
 'send',
 'mobile',
 'mode',
 'sha',
 'll',
 'get',
 'reply'],
1,
764),
(['m', 'back', 'lemme', 'know', 're', 'ready'], 1, 765),
(['m', 'bus', 'love'], 1, 766),
(['m', 'chennai', 'velachery'], 1, 767),
(['m',
 'coming',
 'back',
 'thursday',
 'yay',
 'gonna',
 'ok',
 'get',
 'money',
 'cheers',
 'oh',
 'yeah',
 'everything',
```

     everything ,
  'alright',
  'hows',
  'school',
  'call',
  'work'],
 1,
 768),
(['m', 'coming', 'home', '4', 'dinner'], 1, 769),
(['m', 'coming', 'whatever', 'want'], 1, 770),
(['m',
  'diet',
  'ate',
  '1',
  'many',
  'slices',
  'pizza',
  'yesterday',
  'ugh',
  'm',
  'always',
  'diet'],
 1,
 771),
(['m', 'done', 'c'], 1, 772),
(['m', 'done', 'oredi'], 1, 773),
(['m',
  'eatin',
  'lor',
  'goin',
  'back',
  'work',
  'soon',
  'e',
  'mountain',
  'deer',
  'show',
  'huh',
  'watch',
  'b4',
  'liao',
  'nice'],
 1,
 774),
(['m', 'glad', 'following', 'dreams'], 1, 775),
(['m',
  'going',
  '4',
  'lunch',
  'wif',
  'family',
  'aft',
  'dat',
  'go',
  'str',
  '2',
  'orchard',
  'lor'],
 1,
 776),
(['m', 'going', 'try', '2', 'months', 'ha', 'ha', 'joking'], 1, 777),
(['m',
  'gonna',
  'home',
  'soon',
  'want',
  'talk',
  'stuff',
  'anymore',
  'tonight',
  'k',
  've',
  'cried'

```
        'enough',
        'today'],
       1,
       778),
      (['m', 'good', 'registered', 'vote'], 1, 779),
      (['m', 'great', 'aunts', 'anniversary', 'party', 'tarpon', 'springs'],
       1,
       780),
      (['m', 'home', 'please', 'call'], 1, 781),
      (['m', 'hungry', 'buy', 'smth', 'home'], 1, 782),
      (['m', 'leaving', 'house'], 1, 783),
      (['m', 'love', 'm', 'excited', 'day', 'spend', 'make', 'happy'], 1, 784),
      (['m', 'meeting', 'call', 'later'], 1, 785),
      (['m', 'meeting', 'call', 'later'], 1, 786),
      (['m', 'meeting', 'darren'], 1, 787),
      (['m',
        'nt',
        'goin',
        'got',
        'somethin',
        'unless',
        'meetin',
        '4',
        'dinner',
        'lor',
        'haha',
        'wonder',
        'go',
        'tis',
        'time'],
       1,
       788),
      (['m', 'office', 'call', 'lt', 'gt', 'min'], 1, 789),
      (['m', 'office', 'da'], 1, 790),
      (['m', 'ok', 'part', 'tomorrow'], 1, 791),
      (['m',
        'ok',
        'wif',
        'cos',
        'like',
        '2',
        'try',
        'new',
        'things',
        'scared',
        'u',
        'dun',
        'like',
        'mah',
        'cos',
        'u',
        'said',
        'loud'],
       1,
       792),
      (['m', 'outside', 'islands', 'head', 'towards', 'hard', 'rock', 'll', 'run'],
       1,
       793),
      (['m', 'reaching', 'home', '5', 'min'], 1, 794),
      (['m', 'really', 'sorry', 'lit', 'hair', 'fire'], 1, 795),
      (['m',
        'really',
        'sorry',
        'won',
        'b',
        'able',
        '2',
        'friday',
        'hope',
        'u',
        'find',
        'alternative'
```

   'alternative',
   'hope',
   'yr',
   'term',
   'going',
   'ok'],
 1,
 796),
 (['m',
   'saying',
   'haven',
   'explicitly',
   'told',
   'nora',
   'know',
   'someone',
   'm',
   'probably',
   'gonna',
   'bother'],
 1,
 797),
 (['m', 'serious', 'money', 'base'], 1, 798),
 (['m',
   'smoking',
   'people',
   'use',
   'wylie',
   'smokes',
   'much',
   'justify',
   'ruining',
   'shit'],
 1,
 799),
 (['m',
   'still',
   'cozy',
   'exhausted',
   'last',
   'night',
   'nobody',
   'went',
   'school',
   'work',
   'everything',
   'closed'],
 1,
 800),
 (['m',
   'still',
   'looking',
   'car',
   'buy',
   'gone',
   '4the',
   'driving',
   'test',
   'yet'],
 1,
 801),
 (['m',
   'taking',
   'derek',
   'amp',
   'taylor',
   'walmart',
   'm',
   'back',
   'time',
   're',
   'done',
   'leave'

```
  'leave',
  'mouse',
  'desk',
  'll',
  'text',
  'priscilla',
  'ready'],
 1,
 802),
(['m', 'town', 'll', 'jus', 'take', 'mrt', 'later'], 1, 803),
(['m',
  'used',
  'hope',
  'agents',
  'drop',
  'since',
  've',
  'booked',
  'things',
  'year',
  'whole',
  'boston',
  'nyc',
  'experiment'],
 1,
 804),
(['m', 'wait', 'till', '2', 'bus', 'pick'], 1, 805),
(['m', 'watching', 'live'], 1, 806),
(['m',
  'way',
  'home',
  'went',
  'change',
  'batt',
  '4',
  'watch',
  'go',
  'shop',
  'bit',
  'lor'],
 1,
 807),
(['m', 'wif', 'buying', 'tix', 'lar'], 1, 808),
(['m', 'work', 'please', 'call'], 1, 809),
(['m', 'workin', 'get', 'job'], 1, 810),
(['m', 'working', 'technical', 'support', 'voice', 'process'], 1, 811),
(['macha',
  'dont',
  'feel',
  'upset',
  'assume',
  'mindset',
  'believe',
  'one',
  'evening',
  'wonderful',
  'plans',
  'us',
  'let',
  'life',
  'begin',
  'call',
  'anytime'],
 1,
 812),
(['make', 'baby', 'yo', 'tho'], 1, 813),
(['make',
  'life',
  'stressfull',
  'always',
  'find',
  'time',
  'laugh',
```

    'laugh',
  'may',
  'add',
  'years',
  'life',
  'surely',
  'adds',
  'life',
  'ur',
  'years',
  'gud',
  'ni8',
  'swt',
  'dreams'],
 1,
 814),
(['make',
  'time',
  'tonight',
  'whenever',
  'cool',
  'text',
  'whenever',
  're',
  'around'],
 1,
 815),
(['man', 'carlos', 'definitely', 'coming', 'mu', 'tonight', 'excuses'],
 1,
 816),
(['many',
  'people',
  'seems',
  'special',
  'first',
  'sight',
  'remain',
  'special',
  'till',
  'last',
  'sight',
  'maintain',
  'till',
  'life',
  'ends',
  'sh',
  'jas'],
 1,
 817),
(['many',
  'people',
  'seems',
  'special',
  'first',
  'sight',
  'remain',
  'special',
  'till',
  'last',
  'sight',
  'maintain',
  'till',
  'life',
  'ends',
  'take',
  'cr',
  'da'],
 1,
 818),
(['march',
  'ending',
  'ready',
  'call',

```
    'call',
  'sure',
  'problem',
  'capital',
  'never',
  'complete',
  'far',
  'work',
  'ladies'],
 1,
 819),
(['mark',
  'taking',
  'forever',
  'pick',
  'prescription',
  'pain',
  'coming',
  'back'],
 1,
 820),
(['mark',
  'works',
  'tomorrow',
  'gets',
  '5',
  'work',
  'house',
  'meet',
  'u',
  'afterwards'],
 1,
 821),
(['match', 'started', 'india', 'lt', 'gt', '2'], 1, 822),
(['mathews', 'tait', 'edwards', 'anderson'], 1, 823),
(['may', 'b', 'approve', 'panalam', 'posts'], 1, 824),
(['maybe', 'find', 'something', 'else', 'instead'], 1, 825),
(['mean', 'come', 'chase', 'stated', 'watch', 'many', 'movies', 'want'],
 1,
 826),
(['mean', 'left', 'early', 'check', 'cos', 'm', 'working', '9', '6'], 1, 827),
(['mean', 'website', 'yes'], 1, 828),
(['means',
  'february',
  'april',
  'll',
  'getting',
  'place',
  'stay',
  'hustle',
  'back',
  'forth',
  'audition',
  'season',
  'since',
  'sister',
  'moved',
  'away',
  'harlem'],
 1,
 829),
(['means', 'still', 'think', 'teju'], 1, 830),
(['meant',
  'calculation',
  'lt',
  'gt',
  'units',
  'lt',
  'gt',
  'school',
  'really',
  'expensive',
  'started'
```

   'started',
   'practicing',
   'accent',
   'important',
   'decided',
   '4years',
   'dental',
   'school',
   'll',
   'nmde',
   'exam'],
  1,
  831),
 (['meet', 'corporation', 'st', 'outside', 'gap', 'see', 'mind', 'working'],
  1,
  832),
 (['meeting', 'da', 'call'], 1, 833),
 (['meeting',
   'rite',
   'll',
   'go',
   'home',
   'lor',
   'dun',
   'feel',
   'like',
   'comin',
   'ok'],
  1,
  834),
 (['meeting', 'u', 'work', 'tel', 'shall', 'work', 'tomorrow'], 1, 835),
 (['merry', 'christmas', 'babe', 'love', 'ya', 'kisses'], 1, 836),
 (['message',
   'free',
   'welcome',
   'new',
   'improved',
   'sex',
   'dogging',
   'club',
   'unsubscribe',
   'service',
   'reply',
   'stop',
   'msgs',
   '150p',
   '18'],
  0,
  837),
 (['message',
   'free',
   'welcome',
   'new',
   'improved',
   'sex',
   'dogging',
   'club',
   'unsubscribe',
   'service',
   'reply',
   'stop',
   'msgs',
   '150p',
   '18'],
  0,
  838),
 (['message',
   'important',
   'information',
   'o2',
   'user',
   'today',
   'lucky',

    lucky',
  'day',
  '2',
  'find',
  'log',
  'onto',
  'http',
  'www',
  'urawinner',
  'com',
  'fantastic',
  'surprise',
  'awaiting'],
 0,
 839),
(['message',
  'sent',
  'askin',
  'lt',
  'gt',
  'dollars',
  'shoul',
  'pay',
  'lt',
  'gt',
  'lt',
  'gt'],
 1,
 840),
(['message',
  'text',
  'missing',
  'sender',
  'name',
  'missing',
  'number',
  'missing',
  'sent',
  'date',
  'missing',
  'missing',
  'u',
  'lot',
  'thats',
  'y',
  'everything',
  'missing',
  'sent',
  'via',
  'fullonsms',
  'com'],
 1,
 841),
(['messages', 'phone', 'm', 'holding'], 1, 842),
(['met',
  'stranger',
  'choose',
  'friend',
  'long',
  'world',
  'stands',
  'friendship',
  'never',
  'ends',
  'lets',
  'friends',
  'forever',
  'gud',
  'nitz'],
 1,
 843),
(['midnight', 'earliest'], 1, 844),

(['might', 'ax', 'well', 'im'], 1, 845),
(['might',
  'want',
  'pull',
  'case',
  'plan',
  'spending',
  'much',
  'confidence',
  'derek',
  'taylor',
  'money',
  'management'],
 1,
 846),
(['mila',
  'age23',
  'blonde',
  'new',
  'uk',
  'look',
  'sex',
  'uk',
  'guys',
  'u',
  'like',
  'fun',
  'text',
  'mtalk',
  '69866',
  '18',
  '30pp',
  'txt',
  '1st',
  '5free',
  '1',
  '50',
  'increments',
  'help08718728876'],
 0,
 847),
(['misplaced',
  'number',
  'sending',
  'texts',
  'old',
  'number',
  'wondering',
  've',
  'heard',
  'year',
  'best',
  'mcat',
  'got',
  'number',
  'atlanta',
  'friends'],
 1,
 848),
(['miss'], 1, 849),
(['miss',
  'vday',
  'parachute',
  'double',
  'coins',
  'u',
  'must',
  'know',
  'well'],
 1,
 850),
(['missionary', 'hook', 'doggy', 'hook', 'standing'], 1, 851),
(['mm', 'ask', 'some', 'enough'], 1, 852)

(['mm', 'ask', 'come', 'enough'], 1, 852),
(['mm', 'asked', 'call', 'radio'], 1, 853),
(['mm',
  'umma',
  'ask',
  'vava',
  'also',
  'come',
  'tell',
  'play',
  'later',
  'together'],
 1,
 854),
(['mmm',
  'thats',
  'better',
  'got',
  'roast',
  'd',
  'b',
  'better',
  'drinks',
  '2',
  'good',
  'indian'],
 1,
 855),
(['mmmm',
  'fuck',
  'fair',
  'know',
  'weaknesses',
  'grins',
  'pushes',
  'knee',
  'exposes',
  'belly',
  'pulls',
  'head',
  'forget',
  'know',
  'wicked',
  'smile'],
 1,
 856),
(['mmmmmmm',
  'snuggles',
  'deep',
  'contented',
  'sigh',
  'whispers',
  'fucking',
  'love',
  'much',
  'barely',
  'stand'],
 1,
 857),
(['mobile',
  '10',
  'mths',
  'update',
  'latest',
  'camera',
  'video',
  'phones',
  'free',
  'keep',
  'ur',
  'number',
  'get',
  'extra',

  'extra',
  'free',
  'mins',
  'texts',
  'text',
  'yes',
  'call'],
 0,
 858),
(['mobile',
  '11',
  'months',
  'u',
  'r',
  'entitled',
  'update',
  'latest',
  'colour',
  'mobiles',
  'camera',
  'free',
  'call',
  'mobile',
  'update',
  'co',
  'free',
  '08002986030'],
 0,
 859),
(['mobile',
  '11',
  'months',
  'u',
  'r',
  'entitled',
  'update',
  'latest',
  'colour',
  'mobiles',
  'camera',
  'free',
  'call',
  'mobile',
  'update',
  'co',
  'free',
  '08002986030'],
 0,
 860),
(['mobile', 'club', 'choose', 'top', 'quality', 'items', 'mobile', '7cfca1a'],
 0,
 861),
(['mobile',
  'number',
  'won',
  '5000',
  'claim',
  'calls',
  'us',
  'back',
  'ring',
  'claims',
  'hot',
  'line',
  '09050005321'],
 0,
 862),
(['moby',
  'pub',
  'quiz',
  'win',
  '100',
  'high'

     'high',
     'street',
     'prize',
     'u',
     'know',
     'new',
     'duchess',
     'cornwall',
     'txt',
     'first',
     'name',
     '82277',
     'unsub',
     'stop',
     '1',
     '50',
     '008704050406',
     'sp',
     'arrow'],
  0,
  863),
 (['mode', 'men', 'left'], 1, 864),
 (['moment',
     'day',
     'value',
     'morning',
     'brings',
     'hope',
     'afternoon',
     'brings',
     'faith',
     'evening',
     'brings',
     'luv',
     'night',
     'brings',
     'rest',
     'wish',
     'u',
     'find',
     'today',
     'good',
     'morning'],
  1,
  865),
 (['money',
     'r',
     'lucky',
     'winner',
     '2',
     'claim',
     'prize',
     'text',
     'money',
     '2',
     '88600',
     '1million',
     'give',
     'away',
     'ppt150x3',
     'normal',
     'text',
     'rate',
     'box403',
     'w1t1jy'],
  0,
  866),
 (['moon',
     'come',
     'color',
     'dreams',
     'stars',
     'make',

```
    make ,
    'musical',
    'sms',
    'give',
    'warm',
    'peaceful',
    'sleep',
    'good',
    'night'],
 1,
 867),
(['motivate',
  'behind',
  'every',
  'darkness',
  'shining',
  'light',
  'waiting',
  'find',
  'behind',
  'every',
  'best',
  'friend',
  'always',
  'trust',
  'love',
  'bslvyl'],
 1,
 868),
(['much', 'buzy'], 1, 869),
(['much', 'buzy'], 1, 870),
(['much', 'dirty', 'fellow'], 1, 871),
(['much', 'r', 'willing', 'pay'], 1, 872),
(['much', 'textin', 'bout'], 1, 873),
(['much', 'u', 'trying', 'get'], 1, 874),
(['mum',
  've',
  'sent',
  'many',
  'many',
  'messages',
  'since',
  'got',
  'want',
  'know',
  'actually',
  'getting',
  'enjoy',
  'rest',
  'day'],
 1,
 875),
(['nah', 'help', 've', 'never', 'iphone'], 1, 876),
(['nah', 'man', 'car', 'meant', 'crammed', 'full', 'people'], 1, 877),
(['nah',
  'straight',
  'bring',
  'bud',
  'drinks',
  'something',
  'actually',
  'little',
  'useful',
  'straight',
  'cash'],
 1,
 878),
(['nan', 'sonathaya', 'soladha', 'boss'], 1, 879),
(['nannys', 'address'], 1, 880),
(['naughty',
  'little',
  'thought',
  'better'
```

   better',
  'flirt',
  'flirt',
  'n',
  'flirt',
  'rather',
  'loving',
  'someone',
  'n',
  'gettin',
  'hurt',
  'hurt',
  'n',
  'hurt',
  'gud',
  'nyt'],
 1,
 881),
(['ne',
  'thing',
  'interesting',
  'good',
  'birthday',
  'u',
  'wrking',
  'nxt',
  'started',
  'uni',
  'today'],
 1,
 882),
(['necessarily', 'respectful'], 1, 883),
(['need', 'details', 'online', 'job'], 1, 884),
(['needs', 'stop', 'going', 'bed', 'make', 'fucking', 'dealing'], 1, 885),
(['neft',
  'transaction',
  'reference',
  'number',
  'lt',
  'gt',
  'rs',
  'lt',
  'decimal',
  'gt',
  'credited',
  'beneficiary',
  'account',
  'lt',
  'gt',
  'lt',
  'time',
  'gt',
  'lt',
  'gt'],
 1,
 886),
(['neither',
  'sterm',
  'voice',
  'm',
  'studying',
  'fine',
  'sure',
  'thing',
  'resolved',
  'tho',
  'anyway',
  'fab',
  'hols'],
 1,
 887),
(['nervous', 'lt', 'gt'], 1, 888),
(['neshanth', 'tel', 'r', 'u'], 1, 889)

```
([ neshanch ,  tei ,  i ,  u ], 1, 889),
(['networking', 'job'], 1, 890),
(['networking', 'technical', 'support', 'associate'], 1, 891),
(['neva',
  'grumble',
  'sad',
  'lor',
  'hee',
  'buy',
  'tmr',
  'lor',
  'aft',
  'lunch',
  'still',
  'meetin',
  '4',
  'lunch',
  'tmr',
  'neva',
  'hear',
  'fr',
  'lei',
  'got',
  'lot',
  'work',
  'ar'],
 1,
 892),
(['neva', 'mind', 'ok'], 1, 893),
(['neva', 'tell', 'noe', 'm', 'home', 'da', 'aft', 'wat'], 1, 894),
(['new', 'years', 'plans'], 1, 895),
(['nice', 'day', 'today', 'love', 'dearly'], 1, 896),
(['nice',
  'line',
  'said',
  'broken',
  'heart',
  'plz',
  'cum',
  '1',
  'times',
  'infront',
  'wise',
  'll',
  'trust',
  'u',
  'good',
  '9t'],
 1,
 897),
(['nice', 'new', 'shirts', 'thing', 'wear', 'nudist', 'themed', 'mu'],
 1,
 898),
(['nice',
  'wait',
  'texting',
  'right',
  'm',
  'gonna',
  'pay',
  'ticket',
  'ya',
  'know'],
 1,
 899),
(['night',
  'ended',
  'another',
  'day',
  'morning',
  'come',
  'special',
  'way'
```

   'way',
   'may',
   'smile',
   'like',
   'sunny',
   'rays',
   'leaves',
   'worries',
   'blue',
   'blue',
   'bay',
   'gud',
   'mrng'],
 1,
 900),
(['night',
  'ended',
  'another',
  'day',
  'morning',
  'come',
  'special',
  'way',
  'may',
  'smile',
  'like',
  'sunny',
  'rays',
  'leaves',
  'worries',
  'blue',
  'blue',
  'bay',
  'gud',
  'mrng'],
 1,
 901),
(['night', 'night', 'see', 'tomorrow'], 1, 902),
(['nite'], 1, 903),
(['nobody', 'names', 'penis', 'girls', 'name', 'story', 'doesn', 'add'],
 1,
 904),
(['noice', 'text', 're'], 1, 905),
(['none', 'happening', 'til', 'get', 'though'], 1, 906),
(['none', 'nowhere', 'ikno', 'doesdiscount', 'shitinnit'], 1, 907),
(['nope',
  'll',
  'b',
  'going',
  '2',
  'sch',
  'fri',
  'quite',
  'early',
  'lor',
  'cos',
  'mys',
  'sis',
  'got',
  'paper',
  'da',
  'morn'],
 1,
 908),
(['nope', 'll', 'come', 'online'], 1, 909),
(['nope',
  'm',
  'going',
  'home',
  'go',
  'pump',
  'petrol',
  'lor'

     'lol',
     'like',
     'going',
     '2',
     'rain',
     'soon'],
    1,
    910),
   (['norm',
     'tomorrow',
     'finish',
     '4',
     '15',
     'cos',
     'st',
     'tests',
     'need',
     'sort',
     'library',
     'stuff',
     'point',
     'tomo',
     'got',
     'letter',
     'today',
     'access',
     'til',
     'end',
     'march',
     'better',
     'get',
     'move'],
    1,
    911),
   (['nothing',
     'jus',
     'tot',
     'u',
     'ask',
     'cos',
     'u',
     'ba',
     'gua',
     'went',
     'mt',
     'faber',
     'yest',
     'yest',
     'jus',
     'went',
     'already',
     'mah',
     'today',
     'going',
     'jus',
     'call',
     'lor'],
    1,
    912),
   (['nothing', 'ok', 'anyway', 'give', 'treat'], 1, 913),
   (['notice',
     'like',
     'looking',
     'shit',
     'mirror',
     'youre',
     'turning',
     'right',
     'freak'],
    1,
    914),
   (['nvm', 'ok'], 1, 915),
   (['nvm', 'take', 'ur', 'time'], 1, 916)

(['nvm', 'take', 'ur', 'time'], 1, 916),
(['nw', 'came', 'hme', 'da'], 1, 917),
(['o', 'fps'], 1, 918),
(['o', 'ic', 'lol', 'play', '9', 'doors', 'sometime', 'yo'], 1, 919),
(['o', 'played', 'smash', 'bros', 'lt', 'gt', 'religiously'], 1, 920),
(['o', 'shore', 'takin', 'bus'], 1, 921),
(['ofcourse', 'also', 'upload', 'songs'], 1, 922),
(['office', 'around', '4', 'pm', 'going', 'hospital'], 1, 923),
(['oh', 'asked', 'fun', 'haha', 'take', 'care'], 1, 924),
(['oh', 'brand'], 1, 925),
(['oh', 'forwarded', 'message', 'thought', 'send'], 1, 926),
(['oh',
  'get',
  'paid',
  'outstanding',
  'one',
  'commercial',
  'hasbro',
  'august',
  'made',
  'us',
  'jump',
  'many',
  'hoops',
  'get',
  'paid',
  'still'],
 1,
 927),
(['oh',
  'half',
  'hour',
  'much',
  'longer',
  'syria',
  'canada',
  'eh',
  'wow',
  'must',
  'get',
  'much',
  'work',
  'done',
  'day',
  'us',
  'extra',
  'time',
  'grins'],
 1,
 928),
(['oh', 'howda', 'gud', 'gud', 'mathe', 'en', 'samachara', 'chikku'], 1, 929),
(['oh', 'k', 'got', 'job', 'whats'], 1, 930),
(['oh', 'k', 'm', 'watching'], 1, 931),
(['oh', 'kay', 'sat', 'right'], 1, 932),
(['oh',
  'love',
  'soooo',
  'good',
  'hear',
  'omg',
  'missed',
  'much',
  'today',
  'm',
  'sorry',
  'problems',
  'provider',
  'thank',
  'tming'],
 1,
 933),
(['oh',
  'oh'

      'oh',
  'den',
  'muz',
  'change',
  'plan',
  'liao',
  'go',
  'back',
  'yan',
  'jiu'],
 1,
 934),
(['oh', 'oh', 'wasted', 'den', 'muz', 'chiong', 'sat', 'n', 'sun', 'liao'],
 1,
 935),
(['oh', 'ok', 'didnt', 'know', 'meant', 'yep', 'baby', 'jontin'], 1, 936),
(['oh', 'ok', 'prob'], 1, 937),
(['oh',
  'really',
  'perform',
  'write',
  'paper',
  'go',
  'movie',
  'home',
  'midnight',
  'huh'],
 1,
 938),
(['oh',
  'right',
  'ok',
  'll',
  'make',
  'sure',
  'loads',
  'work',
  'day',
  'got',
  'really',
  'nasty',
  'cough',
  'today',
  'dry',
  'n',
  'shot',
  'really',
  'help'],
 1,
 939),
(['oh', 'send', 'address'], 1, 940),
(['oh',
  'shit',
  'thought',
  'trip',
  'looooooool',
  'makes',
  'much',
  'sense',
  'grins',
  'sofa',
  'reference',
  'sleep',
  'couch',
  'link',
  'sent',
  'wasn',
  'went',
  'trip',
  'oh',
  'didn',
  'babe',
  'go',

  'go',
  'celebration',
  'rents'],
 1,
 941),
(['oh',
  'thkin',
  'goin',
  'yogasana',
  '10',
  'den',
  'nd',
  'go',
  '3',
  'den',
  'rush',
  'parco',
  '4',
  'nb',
  'okie',
  'lor',
  'u',
  'call',
  'ready'],
 1,
 942),
(['oh', 'usual', 'vijay', 'film', 'different'], 1, 943),
(['oh', 'yeah', 'diet', 'flew', 'window'], 1, 944),
(['oic', 'saw', 'tot', 'din', 'c', 'found', 'group', 'liao'], 1, 945),
(['ok'], 1, 946),
(['ok'], 1, 947),
(['ok'], 1, 948),
(['ok'], 1, 949),
(['ok'], 1, 950),
(['ok'], 1, 951),
(['ok'], 1, 952),
(['ok', 'also', 'wan', '2', 'watch', 'e', '9', 'pm', 'show'], 1, 953),
(['ok', 'april', 'cant', 'wait'], 1, 954),
(['ok',
  'b',
  'lovely',
  'u',
  'r',
  'sure',
  'think',
  'wot',
  'u',
  'want',
  'drinkin',
  'dancin',
  'eatin',
  'cinema',
  'u',
  'wot'],
 1,
 955),
(['ok', 'bag'], 1, 956),
(['ok', 'bus', 'come', 'soon', 'come', 'otherwise', 'tomorrow'], 1, 957),
(['ok', 'c', 'u'], 1, 958),
(['ok', 'careful', 'text', 'drive'], 1, 959),
(['ok', 'chinese', 'food', 'way', 'get', 'fat', 're', 'paying', 'lipo'],
 1,
 960),
(['ok', 'come', 'home', 'vl', 'nice', 'meet', 'v', 'chat'], 1, 961),
(['ok', 'come', 'n', 'pick', 'u', 'engin'], 1, 962),
(['ok', 'come', 'ur', 'home', 'half', 'hour'], 1, 963),
(['ok', 'da', 'already', 'planned', 'wil', 'pick'], 1, 964),
(['ok', 'enjoy', 'r', 'u', 'home'], 1, 965),
(['ok', 'finish', '6'], 1, 966),
(['ok', 'juz', 'receive'], 1, 967),
(['ok',
  'knackered',
  'came'

  'home',
  'went',
  'sleep',
  'good',
  'full',
  'time',
  'work',
  'lark'],
 1,
 968),
(['ok', 'll', 'ok', 'guess'], 1, 969),
(['ok', 'll', 'right', 'later'], 1, 970),
(['ok', 'lor'], 1, 971),
(['ok', 'lor'], 1, 972),
(['ok', 'lor'], 1, 973),
(['ok', 'lor', 'buy', 'wat'], 1, 974),
(['ok', 'lor', 'go', 'tog', 'lor'], 1, 975),
(['ok', 'lor', 'u', 'wan', 'go', 'look', '4', 'u'], 1, 976),
(['ok', 'm', 'gonna', 'head', 'usf', 'like', 'fifteen', 'minutes'], 1, 977),
(['ok', 'many', 'buy'], 1, 978),
(['ok', 'may', 'free', 'gym'], 1, 979),
(['ok', 'need', 'tell', 'anything', 'going', 'sleep', 'good', 'night'],
 1,
 980),
(['ok', 'ok', 'ok', 'whats', 'ur', 'todays', 'plan'], 1, 981),
(['ok', 'omw', 'castor'], 1, 982),
(['ok',
  'popped',
  'ask',
  'bout',
  'something',
  'said',
  'd',
  'around',
  'tonght',
  'wen',
  'girl',
  'comes'],
 1,
 983),
(['ok', 'problem', 'yup', 'm', 'going', 'sch', '4', 'rem', 'correctly'],
 1,
 984),
(['ok', 'r', 'meeting', 'later'], 1, 985),
(['ok', 'set', 'let', 'u', 'noe', 'e', 'details', 'later'], 1, 986),
(['ok',
  'stay',
  'night',
  'xavier',
  'sleeping',
  'bag',
  'm',
  'getting',
  'tired'],
 1,
 987),
(['ok', 'take', 'ur', 'time', 'n', 'enjoy', 'ur', 'dinner'], 1, 988),
(['ok',
  'thats',
  'cool',
  'either',
  'raglan',
  'rd',
  'edward',
  'rd',
  'behind',
  'cricket',
  'ground',
  'gimme',
  'ring',
  'ur',

      'closeby',
  'see',
  'tuesday'],
 1,
 989),
(['ok', 'thk', 'got', 'u', 'wan', '2', 'come', 'wat'], 1, 990),
(['ok', 'try', 'week', 'end', 'course', 'coimbatore'], 1, 991),
(['ok', 'u', 'tell', 'wat', 'time', 'u', 'coming', 'later', 'lor'], 1, 992),
(['okay',
  'lor',
  'wah',
  'like',
  'def',
  'wont',
  'let',
  'us',
  'go',
  'haha',
  'say',
  'terms',
  'conditions'],
 1,
 993),
(['okay', 've', 'seen', 'pick', 'friday'], 1, 994),
(['okey',
  'doke',
  'm',
  'home',
  'dressed',
  'cos',
  'laying',
  'around',
  'ill',
  'speak',
  'later',
  'bout',
  'times',
  'stuff'],
 1,
 995),
(['okie'], 1, 996),
(['okmail',
  'dear',
  'dave',
  'final',
  'notice',
  'collect',
  '4',
  'tenerife',
  'holiday',
  '5000',
  'cash',
  'award',
  'call',
  '09061743806',
  'landline',
  'tcs',
  'sae',
  'box326',
  'cw25wx',
  '150ppm'],
 0,
 997),
(['ola',
  'get',
  'back',
  'maybe',
  'today',
  've',
  'told',
  'direct',
  'link',
  'us',

```
      us ,
    'getting',
    'cars',
    'bids',
    'online',
    'arrange',
    'shipping',
    'get',
    'cut',
    'u',
    'partnership',
    'u',
    'invest',
    'money',
    'shipping',
    'takes',
    'care',
    'rest',
    'u',
    'wud',
    'b',
    'self',
    'reliant',
    'soon',
    'dnt',
    'worry'],
   1,
   998),
  (['olol',
    'printed',
    'forum',
    'post',
    'guy',
    'exact',
    'prob',
    'fixed',
    'gpu',
    'replacement',
    'hopefully',
    'dont',
    'ignore'],
   1,
   999),
  ...]
```

In [0]:

```python
tokenizedTestData = testData.flatMap(lambda x : [(word, x[2]) for word in x[0]])
tokenizedTestDF = tokenizedTestData.toDF(["word","id"])

#join the training Data for positive label with TokenizedtestData to get the Probability
of each word(By ID) in the ham class
#We are multiplying the prior probability of ham class with the probabilities of all the
words [P(A|B)]

trainHamProbabilityDF = trainHamProbability.toDF(["word", "hamProb"])
testHamProbability = tokenizedTestDF.join(trainHamProbabilityDF, tokenizedTestDF.word ==
trainHamProbabilityDF.word)
testHamProbability = testHamProbability.rdd.map(lambda x: (x["id"],x["hamProb"]))
testHamProbability = testHamProbability.reduceByKey(lambda x, y: x + y).map(lambda x:(x[
0], x[1] * priorham)).sortByKey()
testHamProbability.collect()
```

```
Out[92]: [(2, 0.0007040136804813849),
  (3, 0.01644831962579236),
  (4, 0.023232451455885702),
  (5, 0.020224393002919783),
  (6, 0.04460886684504775),
  (7, 0.022080429069643432),
  (8, 0.014848288533789212),
  (9, 0.03008058452965917),
  (10, 0.1979558467026294),
  (11, 0.011584225106102788)
```

```
(11, 0.011584225106102788),
(12, 0.026944523589333004),
(13, 0.07065737302285899),
(14, 0.029824579554938665),
(15, 0.030016583285979042),
(16, 0.042176815958520297),
(17, 0.01203223381186367),
(18, 0.05920115040411645),
(20, 0.148290881606685164),
(21, 0.030976601941180935),
(22, 0.05920115040411645),
(23, 0.012416241273944424),
(25, 0.025984504934131115),
(26, 0.026240509908851618),
(27, 0.031936620596382824),
(28, 0.00422408208288831),
(29, 0.012992252467065557),
(30, 0.00857616665313687),
(31, 0.008640167896816998),
(32, 0.0035840696460870502),
(33, 0.09452983691554596),
(34, 0.021312414145481923),
(35, 0.02560049747205036),
(36, 0.014272277340668077),
(37, 0.12083434806807768),
(38, 0.09849791402371376),
(39, 0.00032000621840062946),
(40, 0.04025678227479919),
(41, 0.016576322113152606),
(42, 0.018752364398276888),
(43, 0.025664498715730483),
(44, 0.016128313407391726),
(45, 0.03097660194118093),
(46, 0.008064156703695863),
(47, 0.016832327087873113),
(48, 0.031936620596382824),
(49, 0.04800093276009443),
(50, 0.027392532295093884),
(51, 0.00966418779569901),
(52, 0.02515248876628948),
(53, 0.06924934566189621),
(54, 0.08224159128961777),
(55, 0.15366698607598228),
(56, 0.021440416632842175),
(57, 0.016128313407391726),
(58, 0.011840223008082329),
(59, 0.03475267531830836),
(60, 0.00640012436801259),
(61, 0.009152177846258004),
(62, 0.021632420363882552),
(63, 0.09619386925122919),
(64, 0.0037760733771274283),
(65, 0.016768325844192983),
(66, 0.014976291021149458),
(67, 0.003840074620807554),
(68, 0.01388826987858732),
(69, 0.003392065915046673),
(70, 0.00947218406465863),
(71, 0.006656129342733093),
(72, 0.000704013680481384),
(73, 0.02092840668340117),
(74, 0.058113129261554431),
(75, 0.010176197745140019),
(76, 0.010176197745140019),
(77, 0.03891275615751655),
(78, 0.057345114337392804),
(79, 0.004672090788649195),
(80, 0.005312103225450449),
(81, 0.060801181496119606),
(82, 0.000704013680481384),
(83, 0.015424299972691034),
(84, 0.021952426582283184),
(85, 0.010433202719860523)
```

(85, 0.01043220271986032z),
(86, 0.0028160547219255394),
(87, 0.02592050369045099),
(88, 0.05305703101082437),
(89, 0.01894436812931727),
(90, 0.014464281071708454),
(91, 0.0028160547219255394),
(92, 0.02451247632948822),
(93, 0.0622732101007625),
(94, 0.0008320161678416366),
(95, 0.0016640323356832732),
(96, 0.03270463552054433),
(97, 0.008768170384177248),
(98, 0.019520379322438396),
(99, 0.03814474123335503),
(100, 0.007104138048493976),
(101, 0.05209701235562248),
(102, 0.014912289777469336),
(103, 0.0005760111931211331),
(104, 0.03500868029302887),
(105, 0.003712072133447302),
(106, 0.010688207694581025),
(107, 0.030528593235420055),
(108, 0.07347342774478455),
(109, 0.011392221375062041),
(110, 0.006848133073773471),
(111, 0.0015360298483230216),
(112, 0.037824735014954404),
(113, 0.02809654597557527),
(114, 0.015104293508509713),
(115, 0.054017049666602625),
(116, 0.06355323497436503),
(117, 0.09107376975681915),
(118, 0.03788873625863452),
(119, 0.10739408689525128),
(120, 0.03443266909990774),
(121, 0.07724950112191195),
(122, 0.00748814551057473),
(123, 0.050432980019939205),
(124, 0.024448475085808094),
(125, 0.04012877978743894),
(126, 0.0011520223862422662),
(127, 0.01644831962579236),
(128, 0.004352084570248561),
(129, 0.0063361231243324635),
(130, 0.1183382995645528),
(131, 0.2046759772890426),
(132, 0.011776228837143165),
(133, 0.007040136804813848),
(134, 0.011200217644022034),
(135, 0.017152333306273738),
(136, 0.10585805704692823),
(137, 0.044032855651926606),
(138, 0.054081050909706385),
(139, 0.010496203963540647),
(140, 0.030528593235420055),
(141, 0.017408338280994248),
(142, 0.010496203963540647),
(143, 0.0291845671181374),
(144, 0.042816832022004224),
(145, 0.010816210181941276),
(146, 0.026304511152531747),
(147, 0.0360967701435591),
(148, 0.00576011193121133l),
(149, 0.018496359423556385),
(150, 0.0017920348230435251),
(151, 0.0017920348230435251),
(152, 0.0017920348230435251),
(153, 0.0015360298483230216),
(154, 0.013440261172826644),
(155, 0.0005760111931211331),
(156, 0.02822454846293552),
(157, 0.02822454846293552)

```
(157, 0.02822454846293552),
(158, 0.02822454846293552),
(159, 0.0017280335793633992),
(160, 0.08849718765934O1),
(161, 0.02976057831125854),
(162, 0.044224859382967),
(163, 0.003968077108167806),
(164, 0.020160391759239657),
(165, 0.0249604850352491),
(166, 0.017152333306273738),
(167, 0.014464281071708454),
(168, 0.0017280335793633992),
(169, 0.010944212669301527),
(170, 0.005952115662251709),
(171, 0.0015360298483230216),
(172, 0.00212041041444155),
(173, 0.00212041041444155),
(174, 0.015360298483230216),
(175, 0.05817713050523444),
(176, 0.014208276096987951),
(177, 0.10912212047461467),
(178, 0.021824424094922936),
(179, 0.009920192770419514),
(180, 0.04915295514633669),
(181, 0.02636851239621187),
(182, 0.009984194014099644),
(183, 0.024576477573168342),
(184, 0.016768325844192983),
(185, 0.03059259447910018),
(186, 0.032960640495264835),
(187, 0.01516829475218984),
(188, 0.01990438678451915),
(189, 0.04198481585416259),
(190, 0.03788873625863452),
(191, 0.010176197745140017),
(192, 0.019136371860357645),
(193, 0.013184256198105935),
(194, 0.06873733571245522),
(195, 0.0009600186552018885),
(196, 0.04544088301288938),
(197, 0.029568574580218162),
(198, 0.0017920348230435251),
(199, 0.10694607818949038),
(200, 0.029888580798618794),
(201, 0.029888580798618794),
(202, 0.015744305945310972),
(203, 0.0008960174115217626),
(204, 0.0034560671587267984),
(205, 0.0036480708897671764),
(206, 0.06873733571245523),
(207, 0.036096701435591),
(208, 0.0862096752371296),
(209, 0.005376104469130575),
(210, 0.008704169140497122),
(211, 0.22688440884604633),
(212, 0.06310522626860413),
(213, 0.0197763842971589),
(214, 0.0212484129018018),
(215, 0.017344337037314115),
(216, 0.0028800559656056656),
(217, 0.04627289918073103),
(218, 0.039744772325358I8),
(219, 0.009088176602577876),
(220, 0.016000310920031475),
(221, 0.01721633454995387),
(222, 0.01721633454995387),
(223, 0.023104448968525447),
(224, 0.0035840696460870502),
(225, 0.0117122275934630A),
(226, 0.002048039797764O286),
(227, 0.0001280024873602518),
(228, 0.005824133174891456),
(229, 0.01523329599586996)
```

```
(229, 0.01525229599500900),
(230, 0.20576399843160478),
(231, 0.0010880211425621403),
(232, 0.0763534837103902),
(233, 0.013184256198105936),
(234, 0.020736402952360792),
(235, 0.034112662881507105),
(236, 0.01248024251762455),
(237, 0.10489803839172633),
(238, 0.021184411658121675),
(239, 0.026176508665171495),
(240, 0.026304511152531743),
(241, 0.07705749739087159),
(242, 0.014848288533789208),
(243, 0.019072237061667752),
(244, 0.016896328833155324),
(245, 0.006208120636972212),
(246, 0.07750550609663245),
(247, 0.10464203341700584),
(248, 0.006592128099052968),
(249, 0.05516907205226852),
(250, 0.024320472598447843),
(251, 0.02822454846293552),
(252, 0.007616147997934981),
(253, 0.016192314651071852),
(254, 0.105922205829060836),
(255, 0.009664187795699011),
(256, 0.031936620596382824),
(257, 0.08940973742113588),
(258, 0.024128468867407465),
(259, 0.03584069646087051),
(260, 0.054721063346507635),
(261, 0.0017920348230435251),
(262, 0.0009600186552018885),
(263, 0.010048195257779766),
(264, 0.035072681536708984),
(265, 0.0484489414658553),
(266, 0.039872277481271844),
(267, 0.032512631789503955),
(268, 0.03392065915046672),
(269, 0.023680460161646585),
(270, 0.0001280024873602518),
(271, 0.035776695217190937),
(272, 0.00640012436801259),
(273, 0.024576477573168345),
(274, 0.019584380566118522),
(275, 0.025024486278929226),
(276, 0.03795273750231466),
(277, 0.00953618530833876),
(278, 0.013312258685466186),
(279, 0.024896483791568974),
(280, 0.023232451455885702),
(281, 0.05555307951434928),
(282, 0.01811235196l475627),
(283, 0.02003238927l879406),
(284, 0.009856191526739388),
(285, 0.014912289777469332),
(286, 0.0009600186552018885),
(287, 0.011968232568183542),
(289, 0.023168450212205576),
(290, 0.011520223862422662),
(291, 0.08838571752225387),
(292, 0.027072526076693255),
(293, 0.17933148479171276),
(294, 0.026944523589333004),
(295, 0.037376726309193524),
(296, 0.09516984935234721),
(297, 0.06118518895820035),
(298, 0.0037760733771274274),
(299, 0.045440883012889385),
(300, 0.025088487522609348),
(301, 0.016704324600051286),
(302, 0.02432047259844784)
```

```
(302,  0.02432047259844784),
(303,  0.02432047259844784),
(304,  0.0012800248736025178),
(305,  0.05900914667307608),
(306,  0.062657217562844326),
(307,  0.004608089544969065),
(308,  0.007296141779534352),
(309,  0.02873655841237653),
(310,  0.06118518895820035),
(311,  0.06284922129388364),
(312,  0.009856191526739388),
(313,  0.04185681336680234),
(314,  0.019904386784519154),
(315,  0.028864560899736786),
(316,  0.005056098250729946),
(317,  0.00422408208288831),
(318,  0.016832327087873113),
(319,  0.0108802114256214),
(320,  0.013248257441786062),
(321,  0.08793770881649297),
(322,  0.04403285565192661),
(323,  0.005632109443851079),
(324,  0.005312103225450449),
(325,  0.0007680149241615108),
(326,  0.20806804320408928),
(327,  0.0017920348230435251),
(328,  0.025088487522609355),
(329,  0.015232295995869966),
(330,  0.04230482207256322),
(331,  0.042496825803603595),
(332,  0.02067240170868067),
(333,  0.015936309676351346),
(334,  0.36691913001816184),
(335,  0.02016039175923966),
(336,  0.2976057831125854),
(337,  0.018368356936196133),
(338,  0.028480553437656018),
(339,  0.006592128099052967),
(340,  0.0015360298483230216),
(341,  0.01888036688563714),
(342,  0.009280180333618257),
(343,  0.004608089544969065),
(344,  0.004608089544969065),
(345,  0.01939237683507815),
(346,  0.0035840696460870502),
(347,  0.0025600497472050357),
(348,  0.0028800559656056656),
(349,  0.005440105712810702),
(350,  0.003072059696646043),
(351,  0.011200217644022034),
(352,  0.018944368129317265),
(353,  0.012800248736025182),
(354,  0.029824579554938665),
(355,  0.06406524492380603),
(356,  0.04960096385209757),
(357,  0.06131319144556062),
(358,  0.062273210100076251),
(359,  0.01324825744178606),
(360,  0.03744072755287365),
(361,  0.028544554681336154),
(362,  0.012416241273944424),
(363,  0.009920192770419516),
(364,  0.02156841912020243),
(365,  0.038784753670156304),
(366,  0.03129660815958157),
(367,  0.038272743720715284),
(368,  0.03129660815958157),
(369,  0.09235379463042166),
(370,  0.06406524492380601),
(371,  0.03161661437798219),
(372,  0.03340864920102572),
(373,  0.03148861189062194),
(374,  0.03916876113237044)
```

(374, 0.039168761132237044),
(375, 0.02092840668340117),
(376, 0.013824268634907193),
(377, 0.017408338280994244),
(378, 0.02508848752260935),
(379, 0.02508848752260935),
(380, 0.149314901505733776),
(381, 0.149314901505733776),
(382, 0.02419247011108759),
(383, 0.014848288533789212),
(384, 0.0005760111931211331),
(385, 0.02105640917076142),
(386, 0.005632109443851079),
(387, 0.016128313407391726),
(388, 0.08992174737057687),
(389, 0.0010240198988820143),
(390, 0.0373127250655134),
(391, 0.01888036688563714),
(392, 0.037056720090792895),
(393, 0.014144274853307827),
(394, 0.017280335793633993),
(395, 0.035328686511429494),
(396, 0.07763350858399273),
(397, 0.07763350858399273),
(398, 0.025664498715730486),
(399, 0.02092840668340117),
(400, 0.013696266147546942),
(401, 0.018688363154596762),
(402, 0.0005760111931211331),
(403, 0.004160080839208184),
(404, 0.015104293508509713),
(405, 0.0249604850352491),
(406, 0.015360298483230216),
(407, 0.027328531051413755),
(408, 0.008640167896816998),
(409, 0.04345684445880548),
(410, 0.023296452699565824),
(411, 0.024064467623727336),
(412, 0.023040447724845325),
(413, 0.03462467283094811),
(414, 0.048128935247454686),
(415, 0.011584225106102788),
(416, 0.028800559656056657),
(417, 0.014144274853307824),
(418, 0.01984038554083903),
(419, 0.013632264903866816),
(420, 0.01171222759346304),
(421, 0.016384318338211223),
(422, 0.027008524833013133),
(423, 0.02688052234565288),
(424, 0.014592283559068706),
(425, 0.0405127872495197),
(426, 0.05126499618778085),
(427, 0.047040914104892535),
(428, 0.06854533198141481),
(429, 0.02438447384212797),
(430, 0.019776384297158903),
(431, 0.011904231324503417),
(432, 0.023680460161646582),
(433, 0.011648226349782912),
(434, 0.028160547219255396),
(435, 0.033664654175746225),
(436, 0.07372943271950505),
(437, 0.0726414115769429),
(438, 0.015552302214270593),
(439, 0.026240509990885162),
(440, 0.041920814610482464),
(441, 0.03315264422630522),
(442, 0.03641670765399163),
(443, 0.010816210181941276),
(444, 0.03328064671366547),
(445, 0.011648226349782912),
(446, 0.008960174115217625)

(446, 0.0089601411521762），
(447, 0.03219262557110333),
(448, 0.013056253710745683),
(449, 0.03244863054582383),
(450, 0.011904231324503418),
(451, 0.013632264903866816),
(452, 0.027072526076693255),
(453, 0.028544554681336154),
(454, 0.05113699370042059),
(455, 0.021120410414441546),
(456, 0.039552768594317805),
(457, 0.011904231324503417),
(458, 0.0748814551057473),
(459, 0.04832093897849505),
(460, 0.020352395490280034),
(461, 0.13561863535818677),
(462, 0.06188920263868176),
(463, 0.012736247492345053),
(464, 0.047040914104892535),
(465, 0.009728189039379139),
(466, 0.09235379463042168),
(467, 0.08761770259809236),
(468, 0.004928095763369694),
(469, 0.02310444896852545),
(470, 0.011456222618742535),
(471, 0.05542507702698902),
(472, 0.007232140535854227),
(473, 0.07052937053549875),
(474, 0.013824268634907193),
(475, 0.0002560049747205036),
(476, 0.08684968767393082),
(477, 0.005760111931211331),
(478, 0.026112507421491366),
(479, 0.02105640917076142),
(480, 0.02259243901908444),
(481, 0.008128157947375989),
(482, 0.04940896012105719),
(483, 0.03564869272983012),
(484, 0.07494545634942743),
(485, 0.01984038554083903),
(486, 0.1681952683913708),
(487, 0.02585650244677086),
(488, 0.013376259929146312),
(489, 0.051008991213060344),
(490, 0.028608555925016273),
(491, 0.031040603184861057),
(492, 0.0037760733771274274),
(493, 0.009344181577298381),
(494, 0.07225740411486213),
(495, 0.012288238786584173),
(496, 0.012224237542904047),
(497, 0.013632264903866816),
(498, 0.0026880522345652876),
(499, 0.017472339524674373),
(500, 0.011008213912981657),
(502, 0.03846474745175566),
(503, 0.015872308432671223),
(504, 0.026816521101972753),
(505, 0.09228979338674155),
(506, 0.031232606915901434),
(507, 0.028544554681336157),
(508, 0.01817635320515576),
(509, 0.016448319625792360),
(510, 0.029824579554938672),
(511, 0.006016116905931834),
(512, 0.007680149241615108),
(513, 0.010240198988820143),
(514, 0.00032000621840062946),
(515, 0.03494467904934875),
(516, 0.007040136804813848),
(517, 0.0358406964608705),
(518, 0.09190578592466078),
(519, 0.05664119056911414)

(519, 0.05864110085911414),
(520, 0.044160858139286864),
(521, 0.032576633033184074),
(522, 0.01932837559139802),
(523, 0.11053014783557741),
(524, 0.04108879844264082),
(525, 0.026112507421491363),
(526, 0.017600342012034625),
(527, 0.10329800729972319),
(528, 0.054017049666026266),
(529, 0.05337703722922499),
(530, 0.0022400435288044062),
(531, 0.022976446481165206),
(532, 0.002624050990885162),
(533, 0.0536330422039455),
(534, 0.036864716359752525),
(535, 0.07360143023214478),
(536, 0.021632420363882555),
(537, 0.013056253710745683),
(538, 0.038656751182796045),
(539, 0.11264218887702157),
(540, 0.043968854408246494),
(541, 0.08358562424624441),
(542, 0.018176353205155753),
(543, 0.0654732722847678),
(544, 0.07257741033326277),
(545, 0.05708910936267231),
(546, 0.0075521467542548565),
(547, 0.011968232568183542),
(548, 0.13549063287082652),
(549, 0.01760034201203462),
(550, 0.019712383053478777),
(551, 0.05606508946379028),
(552, 0.022528437775404316),
(553, 0.022400435288044068),
(554, 0.018240354448835882),
(555, 0.03654471014135189),
(556, 0.05344103847290512),
(557, 0.022528437775404316),
(558, 0.029568574580218165),
(559, 0.04684891037385215),
(560, 0.026240509908851618),
(561, 0.011200217644022034),
(562, 0.00998419401409964),
(563, 0.03244863054582383),
(564, 0.008192159191056115),
(565, 0.159043090545111284),
(566, 0.15897908930143273),
(567, 0.09587386303282858),
(568, 0.05740911558107293),
(569, 0.10048195257779766),
(570, 0.1705633144075355),
(571, 0.01689632833155324),
(572, 0.017984349474115376),
(573, 0.015616303457950717),
(574, 0.013696266147546944),
(575, 0.016832327087873113),
(576, 0.013248257441786062),
(577, 0.08179358942320088),
(578, 0.09209778965570116),
(579, 0.08390563046464505),
(580, 0.124546420201525),
(581, 0.038272743720071528),
(582, 0.03846474745175567),
(583, 0.06323322875596439),
(584, 0.013824268634907193),
(585, 0.017856346986755124),
(586, 0.05740911558107293),
(587, 0.09280180333618256),
(588, 0.03296064049526484),
(589, 0.018816365641957013),
(590, 0.044800870576088135),
(591, 0.03596869894823076)

(591, 0.03598698948230/0),
(592, 0.2001958902314338),
(593, 0.08563366404400845),
(594, 0.05740911558107293),
(595, 0.022592439019084445),
(596, 0.06643329093997069),
(597, 0.0042880833265684344),
(598, 0.039872774812718434),
(599, 0.00832016167841368),
(600, 0.012928251223385432),
(601, 0.021760422851242807),
(602, 0.053377037229224995),
(603, 0.020480397977640286),
(604, 0.02803254473189514),
(605, 0.011136216400341906),
(606, 0.0212484129018018),
(608, 0.04185681336680234),
(609, 0.01894436812931727),
(610, 0.04153680714840171),
(611, 0.005568108200170953),
(612, 0.06630528845261044),
(613, 0.006144119393292086),
(614, 0.0029440572092857913),
(615, 0.016192314651071852),
(616, 0.026944523589333007),
(617, 0.03968077108167806),
(618, 0.03968077108167806),
(619, 0.04716891659225278),
(620, 0.009344181577298381),
(621, 0.02835255095029577),
(622, 0.005632109443851079),
(623, 0.03769673252759416),
(624, 0.017472339524674373),
(626, 0.017344337037314115),
(627, 0.017088332062593616),
(628, 0.09305780831090306),
(629, 0.00640012436801259),
(630, 0.007424144266894604),
(631, 0.04390485316456637),
(632, 0.0056961106875312055),
(633, 0.04294483450936447),
(634, 0.06028917154667859),
(635, 0.014912289777469332),
(636, 0.00409607959552807),
(637, 0.007936154216335612),
(638, 0.12761847989817104),
(639, 0.004800093276009443),
(640, 0.05331303598554488),
(641, 0.011008213912981655),
(642, 0.18566760791604522),
(643, 0.13286658187994135),
(644, 0.0046720907886491905),
(645, 0.01715233330627374),
(646, 0.02944057209285791),
(647, 0.0063361231243324635),
(648, 0.02252843777540432),
(649, 0.00422408208288831),
(650, 0.014784287290109081),
(651, 0.05376104469130576),
(652, 0.0029440572092857913),
(653, 0.05728111309371268),
(654, 0.05177700613722185),
(655, 0.002368046016164658),
(656, 0.08409763419568543),
(657, 0.027648537269814387),
(658, 0.015680304701630843),
(659, 0.0215044178765223),
(660, 0.019072370611667752),
(661, 0.023872463892686956),
(662, 0.0229764464811652),
(663, 0.020096390515559528),
(664, 0.012288238786584173),
(665, 0.040704790980560075)

(665, 0.040704790980580075),
(666, 0.010432202719860522),
(667, 0.033984660394146854),
(668, 0.023232451455885702),
(669, 0.07244940784590251),
(670, 0.019776384297158903),
(671, 0.015680304701630085),
(672, 0.015552302214270597),
(673, 0.02656051612725225),
(674, 0.025792501203090738),
(675, 0.012608245004984801),
(676, 0.009728189039379137),
(677, 0.018240354448835882),
(678, 0.07097737924125962),
(679, 0.047040914104892535),
(680, 0.09376182199138444),
(681, 0.0091521778462580002),
(682, 0.0091521778462580002),
(683, 0.019840385540839025),
(684, 0.013376259929146313),
(685, 0.0005120099494410072),
(686, 0.05292902852346412),
(687, 0.01836835693619613),
(688, 0.055297074539628774),
(689, 0.0496649650957777),
(690, 0.01760034201203462),
(691, 0.04390485316456637),
(692, 0.09721788915011124),
(693, 0.028352550950295777),
(694, 0.024448475085808094),
(695, 0.0333446479573456),
(696, 0.03609670143559101),
(697, 0.0028160547219255394),
(698, 0.0028160547219255394),
(699, 0.0028160547219255394),
(700, 0.07718549987823184),
(701, 0.006656129342733093),
(702, 0.015552302214270593),
(703, 0.12704246870504995),
(704, 0.010816210181941278),
(705, 0.046144896693370078),
(706, 0.046144896693370078),
(707, 0.031936620596382824),
(708, 0.0424968258036036),
(709, 0.026624517370932372),
(710, 0.010944212669301527),
(711, 0.03142461064694181),
(712, 0.06700930213309182),
(713, 0.03302464173894496),
(714, 0.021696421607562678),
(715, 0.0634892337306849),
(716, 0.04185681336680234),
(717, 0.028928562143416908),
(718, 0.014912289777469332),
(719, 0.051521001162501336),
(720, 0.026368512396211877),
(721, 0.14093073858363722),
(722, 0.04268882953464398),
(723, 0.04096079595528058),
(724, 0.0012160236299223921),
(725, 0.0056961106875312055),
(726, 0.011264218887702158),
(727, 0.013504262416506563),
(728, 0.006784131830093344),
(729, 0.005760111931211331),
(730, 0.015360298483230216),
(731, 0.013952271122267447),
(732, 0.033216645469985345),
(733, 0.005248101981770325),
(734, 0.04185681336680234),
(735, 0.02201642782596331),
(736, 0.006912134317453597),
(737, 0.0091521778462580002)

(737, 0.0091521776846258002),
(738, 0.011648226349782912),
(739, 0.02105640917076142),
(740, 0.013568263660186692),
(741, 0.022912445237485073),
(742, 0.0800655558438375),
(743, 0.058305132992594697),
(744, 0.021184411658121675),
(745, 0.004352084570248561),
(746, 0.020672401708680666),
(747, 0.02214443031332356),
(748, 0.04716891659225278),
(749, 0.07987355211279713),
(750, 0.06508926482268804),
(751, 0.06508926482268804),
(752, 0.016832327087873113),
(753, 0.017536340768354496),
(754, 0.025408493741009983),
(755, 0.19488378700598344),
(756, 0.1333785918293824),
(757, 0.017024330818913490),
(758, 0.0212484129018018),
(759, 0.00780815172897536),
(760, 0.03276863676422446),
(761, 0.029248568361817533),
(762, 0.03148861189062194),
(763, 0.015232295995869963),
(764, 0.08908973120273524),
(765, 0.04083279346792032),
(766, 0.02752053478245414),
(767, 0.018816365641957013),
(768, 0.08556966280032832),
(769, 0.037504728796553775),
(770, 0.028928562143416908),
(771, 0.04601689420601052),
(772, 0.02278444275012482),
(773, 0.020992407927081295),
(774, 0.0499209700704982),
(775, 0.019392376835078152),
(776, 0.07404943893790568),
(777, 0.04396885440824649),
(778, 0.0625292150755483),
(779, 0.02918456711813741),
(780, 0.023808462649006837),
(781, 0.03974477232535818),
(782, 0.029632575823898288),
(783, 0.0209924079270813),
(784, 0.06310522626860415),
(785, 0.03609670143559101),
(786, 0.03609670143559101),
(787, 0.020224393002919783),
(788, 0.07001736058605772),
(789, 0.060353172790358726),
(790, 0.026304511152531743),
(791, 0.03507268153670899),
(792, 0.18291555443779983),
(793, 0.03366465417574622),
(794, 0.029696577067578424),
(795, 0.030336589504379678),
(796, 0.12921851099017417),
(797, 0.05817713050523444),
(798, 0.020800404196604092),
(799, 0.029376570849177785),
(800, 0.0430088357530446),
(801, 0.03526468526774937),
(802, 0.07942554340703624),
(803, 0.04313683824040486),
(804, 0.03475267531830837),
(805, 0.04147280590472158),
(806, 0.019968388028199283),
(807, 0.06342523248700477),
(808, 0.02156841912020243),
(809, 0.03648070889767176)

```
(809, 0.036480708897671704),
(810, 0.033728655419426344),
(811, 0.020672401708680666),
(812, 0.040960795595528057),
(813, 0.007616147997934981),
(814, 0.045312880552552913),
(815, 0.026752519858292627),
(816, 0.008512165409456745),
(817, 0.018240354448835882),
(818, 0.03020858701701942),
(819, 0.0249604850352491),
(820, 0.014016272365947574),
(821, 0.06297722378124387),
(822, 0.04569688798760989),
(824, 0.004800093276009442),
(825, 0.008064156703695863),
(826, 0.022528437775404316),
(827, 0.031104604428541186),
(828, 0.00499209700704982),
(829, 0.027136527320373378),
(830, 0.013888826987858732),
(831, 0.07968154838175674),
(832, 0.014272277340668077),
(833, 0.019136371860357642),
(834, 0.06899334068717572),
(835, 0.061953203882361860),
(836, 0.015680304701630850),
(837, 0.013696266147546942),
(838, 0.013696266147546942),
(839, 0.03820874247703515),
(840, 0.09497784562130683),
(841, 0.07680149241615104),
(842, 0.022272432800683813),
(843, 0.014592283559068706),
(844, 0.0002560049747205036),
(845, 0.009664187795699011),
(846, 0.019648381809798655),
(847, 0.07142538794702048),
(848, 0.03123260691590144),
(849, 0.0028800559656056656),
(850, 0.06726530710781232),
(851, 0.00089601741152176260),
(852, 0.015168294752189839),
(853, 0.012288238786584173),
(854, 0.03097660194118093),
(855, 0.050880988725700085),
(856, 0.02873655841237653),
(857, 0.015040292264829586),
(858, 0.05651309816955117),
(859, 0.072193402871182),
(860, 0.072193402871182),
(861, 0.0026880522345652876),
(862, 0.014528282315388582),
(863, 0.07353742898846465),
(864, 0.0022400435288044062),
(865, 0.09920192770419514),
(866, 0.052673023548743605),
(867, 0.03916876113223705),
(868, 0.02361645891796646),
(869, 0.005248101981770324),
(870, 0.005248101981770324),
(871, 0.0053121032254504490),
(872, 0.012800247736025180),
(873, 0.005888114418571583),
(874, 0.06643329093997069),
(875, 0.05600108822011016),
(876, 0.00780815172897536),
(877, 0.00748814551057473),
(878, 0.009024175358897752),
(879, 0.00044800870576088130),
(880, 0.0008960174115217626),
(881, 0.03315264422630521),
(882, 0.06899334068717572)
```

```
(882, 0.06899354068717572),
(883, 0.0002560049747205036),
(884, 0.011328220131382283),
(885, 0.0160643121637116),
(886, 0.1605791203934359),
(887, 0.02950457333653804),
(888, 0.029120565874457285),
(889, 0.05312103225450449),
(890, 0.001984038554083903),
(891, 0.0008320161678416366),
(892, 0.061761200015132148),
(893, 0.015488300970590467),
(894, 0.04512087679448876),
(895, 0.0044800870576088125),
(896, 0.026240509908851618),
(897, 0.0832656180278438),
(898, 0.009408182820978507),
(899, 0.04640090166809127),
(900, 0.055297074539628774),
(901, 0.055297074539628774),
(902, 0.01964838180979865),
(903, 0.0012160236299223921),
(904, 0.0046720907886491905),
(905, 0.006912134317453597),
(906, 0.016576322113152606),
(907, 0.00032000621840062946),
(908, 0.070657373022859),
(909, 0.023232451455885702),
(910, 0.08768170384177248),
(911, 0.06380923994908552),
(912, 0.14720286046428957),
(913, 0.02080040419604092),
(914, 0.01817635320515576),
(915, 0.013184256198105936),
(916, 0.024960485035249103),
(917, 0.00825616043473624),
(918, 0.0005760111931211331),
(919, 0.00825616043473624),
(920, 0.029888580798618794),
(921, 0.001984038554083903),
(922, 0.003840074620807554),
(923, 0.020416396733960163),
(924, 0.017856346986755124),
(925, 0.005440105712810702),
(926, 0.015680304701630846),
(927, 0.05408105090970638),
(928, 0.061377192689240745),
(929, 0.010944212669301527),
(930, 0.025344492497329858),
(931, 0.03168061562166232),
(932, 0.010816210181941278),
(933, 0.06444925238588678),
(934, 0.03424066536886736),
(935, 0.02278444275012482),
(936, 0.03347265044470585),
(937, 0.01920037310403777),
(938, 0.031104604428541193),
(939, 0.08729769637969173),
(940, 0.011520223862422662),
(941, 0.04832093897849504),
(942, 0.09875391899843423),
(943, 0.006592128099052968),
(944, 0.009024175358897752),
(945, 0.00857616665313687),
(946, 0.012928251223385432),
(947, 0.012928251223385432),
(948, 0.012928251223385432),
(949, 0.012928251223385432),
(950, 0.012928251223385432),
(951, 0.012928251223385432),
(952, 0.012928251223385432),
(953, 0.04134480341736133),
(954, 0.01958430566118526),
```

```
(954, 0.019584380566118520),
(955, 0.18221154075731844),
(956, 0.013248257441786062),
(957, 0.040320783518479314),
(958, 0.06220920885708237),
(959, 0.01657632211315261),
(960, 0.03705672009079289),
(961, 0.039360764863277435),
(962, 0.0790415359445548),
(963, 0.04364884818984586),
(964, 0.0282245484629352),
(965, 0.07443344639998642),
(966, 0.016128313407391726),
(967, 0.014464281071708456),
(968, 0.05107299245674047),
(969, 0.03865675118279604),
(970, 0.03404866163782698),
(971, 0.01964838180979865),
(972, 0.01964838180979865),
(973, 0.01964838180979865),
(974, 0.027392532295093884),
(975, 0.03750472879655377),
(976, 0.13600264282026753),
(977, 0.04652890415545153),
(978, 0.017856346986755128),
(979, 0.018240354448835882),
(980, 0.05491306707754802),
(981, 0.051905008624582104),
(982, 0.01344026117282644),
(983, 0.03628870516663139),
(984, 0.051585002406181475),
(985, 0.026496514883572124),
(986, 0.07475345261838705),
(987, 0.04064078973687995),
(988, 0.057217111850032545),
(989, 0.03616070267927113),
(990, 0.10464203341700586),
(991, 0.0195203793224384),
(992, 0.1399067186847552),
(993, 0.04454486560136762),
(994, 0.009792190283059263),
(995, 0.043776850067720611),
(996, 0.0009600186552018885),
(997, 0.02617650865171495),
(998, 0.2085160519098502),
(999, 0.00998419401409964),
(1000, 0.023168450212205576),
(1001, 0.007872152972655484),
(1002, 0.01184023008082329),
(1003, 0.013504262416506563),
(1004, 0.008384162922096492),
(1005, 0.034112662881507105),
(1006, 0.0014080273609627697),
(1007, 0.0020480397977640286),
(1008, 0.09420983069714534),
 ...]
```

In [0]:

```python
#join the training Data for spam label with TokenizedtestData to get the Probability of e
ach word(By ID) in the spam class
#We are multiplying the prior probability of spam class with the probabilities of all the
words [P(A|B)]

trainSpamProbabilityDF = trainSpamProbability.toDF(["word", "spamProb"])
testSpamProbability = tokenizedTestDF.join(trainSpamProbabilityDF, tokenizedTestDF.word =
= trainSpamProbabilityDF.word)
testSpamProbability = testSpamProbability.rdd.map(lambda x: (x["id"],x["spamProb"]))
testSpamProbability = testSpamProbability.reduceByKey(lambda x, y: x + y).map(lambda x:(
x[0], x[1] * priorspam)).sortByKey()
testSpamProbability.collect()
```

Out[93]: [(2, 0.000614488851990581)

```
Out[93]: [(2, 0.000614488691990381),
 (3, 0.004379867657269882),
 (4, 0.005831107388484678),
 (5, 0.005621919679480744),
 (6, 0.003961492239262013),
 (7, 0.004471387279959104),
 (8, 0.0007321569815137714),
 (9, 0.004942059625217957),
 (10, 0.0139763538078253383),
 (11, 0.006837823380066114),
 (12, 0.003360077575875701),
 (13, 0.0054911773361353286),
 (14, 0.0037653787762070824),
 (15, 0.0022618421036050434),
 (16, 0.0028763309988004102),
 (17, 0.0060272208656758685),
 (18, 0.009008145718981937),
 (19, 6.537115906372959e-05),
 (20, 0.01595056281155002),
 (21, 0.01172758593603309),
 (22, 0.0089689230235437),
 (23, 0.00018303924537844285),
 (25, 0.000496820808884345),
 (26, 0.0016473532084059857),
 (27, 0.008668215691850542),
 (28, 3.922269543823775e-05),
 (29, 0.00018303924537844285),
 (30, 0.0004445238816333612),
 (31, 0.007517683292328903),
 (32, 0.00045759811344610714),
 (33, 0.0005491177361353286),
 (34, 0.0008498250678284847),
 (35, 0.0004575981134461072),
 (36, 0.0010197900813941816),
 (37, 0.0021049513218520926),
 (38, 0.0035823395166923815),
 (39, 2.6148463625491837e-05),
 (40, 0.0004968208088843448),
 (41, 0.0035692652848796354),
 (42, 0.0008890477632667224),
 (43, 0.0005229692725098367),
 (44, 5.229692725098367e-05),
 (45, 0.0016473532084059857),
 (46, 6.537115906372959e-05),
 (47, 5.229692725098367e-05),
 (48, 0.0012028293267726242),
 (49, 0.002536400971672708),
 (50, 0.0005621919679480745),
 (51, 0.00014381654994020512),
 (52, 0.0007060085178882795),
 (53, 0.0017127243674697151),
 (54, 0.003608487980317873),
 (55, 0.0075438317559543945),
 (56, 0.0002222619408166806),
 (57, 0.0003791527225696963163),
 (58, 0.0006798600542627877),
 (59, 0.00040530118619512345),
 (60, 0.0002091877090039347),
 (61, 0.00026148463625491836),
 (62, 0.0002091877090039347),
 (63, 0.003935343775636521),
 (64, 3.922269543823775e-05),
 (65, 0.00011766808631471327),
 (66, 0.001595056281155002),
 (67, 2.6148463625491837e-05),
 (68, 0.0005360435043225827),
 (70, 0.00016996501356569694),
 (71, 0.00014381654994020512),
 (73, 0.0004183754180078693),
 (74, 0.0005098950406970908),
 (77, 0.000627563127011804),
 (78, 0.0015427593539040182),
 (79, 0.00018303924537844285)
```

```
(79, 0.0001830392453784428),
(81, 0.0016865759038442233),
(83, 0.00011766808631471326),
(84, 0.00019611347719118876),
(85, 0.002248767871792298),
(87, 0.0007060085178882795),
(88, 0.0015819820493422558),
(89, 0.0007583054451392631),
(90, 0.000797528140577501),
(91, 6.537115906372959e-05),
(92, 0.003477745662190414),
(93, 0.002967850621493323),
(95, 5.229692725098367e-05),
(96, 0.0033208548804374628),
(97, 0.0004968208088843448),
(98, 0.00045759811344610714),
(99, 0.0008367508360157387),
(100, 0.0001830392453784428),
(101, 0.011819105558722311),
(102, 0.00026148463625491836),
(103, 7.84453908764755e-05),
(104, 0.001111309704083403),
(105, 0.00013074231812745918),
(106, 0.00019611347719118878),
(107, 0.0003660784907568857),
(108, 0.001124383935896149),
(109, 0.00014381654994020512),
(110, 0.00013074231812745918),
(112, 0.0033470033440629546),
(113, 0.0019349863082863956),
(114, 0.001581982049342256),
(115, 0.0008367508360157388),
(116, 0.00881203224179075),
(117, 0.0003660784907568857),
(118, 2.6148463625491837e-05),
(119, 0.0059226270111739),
(120, 0.0031508898668717662),
(121, 0.0020134316991628712),
(122, 0.0015166108902785265),
(123, 0.009544189223304519),
(124, 0.0002745588680676643),
(125, 0.0007060085178882795),
(127, 0.0002745588680676643),
(129, 0.00011766808631471326),
(130, 0.00428834803458066),
(131, 0.009374224209738822),
(132, 7.84453908764755e-05),
(133, 0.00014381654994020512),
(134, 0.0002353361726294265),
(135, 0.0002876330998804101),
(136, 0.003307780648624717),
(137, 0.007583054451392631),
(138, 0.0007452312133265172),
(139, 0.003360077575875701),
(140, 0.009439595368802553),
(141, 0.00400071493470025),
(142, 0.0036477106756111),
(143, 0.011230765127148745),
(144, 0.004759020379839514),
(145, 0.003360077575875701),
(146, 0.006863971701691607),
(147, 0.0036869333711943486),
(148, 0.00016996501356569694),
(149, 0.0003660784907568857),
(150, 0.000627563127011804),
(151, 0.000627563127011804),
(152, 0.000627563127011804),
(154, 0.005543474288604268),
(156, 0.0016342789765932395),
(157, 0.0016342789765932395),
(158, 0.0016342789765932395),
(160, 0.006367150892807262),
(161, 0.0003660784907568857)
```

```
(161, 0.0003660784907388857),
(162, 0.00013074231812745918),
(163, 0.00011768086314713327),
(164, 0.00028763309988041023),
(165, 0.002026505930975617),
(166, 0.0020657286264138555),
(167, 0.0003922269543823775),
(169, 0.00019611347719118876),
(170, 0.00011768086314713327),
(171, 6.537115906372959e-05),
(172, 6.537115906372959e-05),
(173, 6.537115906372959e-05),
(174, 0.000313781563505902),
(175, 0.0021049513218520926),
(176, 0.0021441740172903306),
(177, 0.0038046014575090614),
(178, 0.0084328795192211177),
(179, 0.00013074231812745918),
(180, 0.0010982354722706572),
(181, 0.001412017035776559),
(182, 2.6148463625491837e-05),
(183, 0.00030070733169315617),
(184, 0.0005883404315735664),
(185, 0.0009936416177686897),
(186, 0.00028763309988041023),
(187, 0.00016996501356569694),
(188, 0.00016996501356569694),
(189, 0.0015558335857167643),
(190, 0.00018303924537844285),
(191, 2.6148463625491837e-05),
(192, 9.151962268922142e-05),
(193, 0.000156890781752951),
(194, 0.0016996501356569694),
(195, 2.6148463625491837e-05),
(196, 0.009844896554997675),
(197, 0.000797528140577501),
(198, 0.00013074231812745918),
(199, 0.013623349548881246),
(200, 0.010786241245515384),
(201, 0.010786241245515384),
(202, 0.008289062969280912),
(205, 0.00010459385450196733),
(206, 0.0007060085178882795),
(207, 0.0006014146633863123),
(208, 0.0016342789765932397),
(209, 0.0012289779039081161),
(210, 0.0013727943403383214),
(211, 0.006602487065436688),
(212, 0.0022618421036050434),
(213, 0.000222619408166806),
(214, 0.006184111647428819),
(215, 0.0018826893810354123),
(216, 3.922269543823775e-05),
(217, 0.009151962268922143),
(218, 0.001255126254023608),
(220, 6.537115906372959e-05),
(221, 0.00019611347719118876),
(222, 0.00019611347719118876),
(223, 0.0035300425894413973),
(224, 2.6148463625491837e-05),
(225, 0.001451239731214797),
(226, 0.0002484104044421724),
(228, 5.229692725098367e-05),
(229, 0.00023533617262942654),
(230, 0.007230050192448491),
(231, 0.00014381654994020512),
(232, 0.0089950714871169192),
(233, 0.000313781563505902),
(234, 0.0005621919679480744),
(235, 0.0037784529938835695),
(236, 0.00019611347719118878),
(237, 0.0079752814057755008),
(238, 0.0036215622121306187)
```

```
(238, 0.0036215622121506187),
(239, 0.0074523121332265172),
(240, 0.0029155536942423407),
(241, 0.0015558335857167645),
(242, 0.00026148463625491836),
(243, 0.0002484104044421725),
(244, 0.0002222619408166806),
(245, 0.0002222619408166806),
(246, 0.002026505930975617),
(247, 0.0063671508928072615),
(248, 0.00016996501356569694),
(249, 0.012786597128665508),
(250, 0.00570036507035722),
(251, 0.001150532399521641),
(252, 2.6148463625491837e-05),
(253, 0.0008498250678284846),
(254, 0.005621919679480744),
(255, 0.00018303924537844285),
(256, 0.000627563127011804),
(257, 0.003503894125815906),
(258, 0.0009282704587049601),
(259, 0.0032293352577482414),
(260, 0.0018696151492226662),
(262, 2.6148463625491837e-05),
(263, 0.0011897550949598785),
(264, 0.0010720870086445165),
(265, 0.0016081305129677476),
(266, 0.0004183754180078694),
(267, 0.0033600775758757012),
(268, 0.0007190827497010254),
(269, 0.0009936416177686897),
(271, 0.0007060085178882795),
(272, 0.00014381654994020512),
(273, 0.0004968208088843448),
(274, 0.0006144888951990581),
(275, 0.0006798600542627877),
(276, 0.000470672345258853),
(277, 0.0006667858224500418),
(278, 0.00018303924537844285),
(279, 0.0006798600542627876),
(280, 0.0005229692725098367),
(281, 0.0010851612404579112),
(282, 0.0004968208088843448),
(283, 0.0004968208088843448),
(284, 0.0002484104044421724),
(285, 0.008419805528740837),
(286, 0.00011766808631471327),
(287, 0.000483746577071599),
(289, 0.0004183754180078694),
(290, 7.84453908764755e-05),
(291, 0.001124383935896149),
(292, 0.00033993002713139387),
(293, 0.010067158495814356),
(294, 0.0005229692725098367),
(295, 0.0005229692725098367),
(296, 0.005465028897727793),
(297, 0.002510252508047216),
(298, 2.6148463625491837e-05),
(299, 0.007426163669639682),
(300, 0.0005752661997608205),
(301, 0.00013074231812745918),
(302, 0.0006014146633863123),
(303, 0.0006014146633863123),
(304, 0.005399657738664063),
(305, 0.0022226194081668062),
(306, 0.0016865759038442233),
(307, 0.0004968208088843448),
(308, 0.0006798600542627876),
(309, 0.005465028897727925),
(310, 0.010851612404579113),
(311, 0.0032554837213737333),
(312, 0.0002091877090039347),
(313, 0.0018565409174099204)
```

(313, 0.0018363409174699204),
(314, 0.0008628992996412305),
(315, 0.0004837465770715989),
(316, 7.84453908764755e-05),
(318, 0.0010197900813941816),
(319, 0.00014381654994020512),
(320, 0.0002222619408166806),
(321, 0.004484461511771849),
(322, 0.0006537115906372959),
(323, 0.0008628992996412305),
(324, 0.0001699650135656969),
(325, 5.229692725098367e-05),
(326, 0.007086233642508286),
(328, 0.00156890781752951),
(329, 0.0013204974130873377),
(330, 0.0012943489494618458),
(331, 0.0003399300271313938),
(332, 0.0016604274402187316),
(333, 0.0003399300271313938),
(334, 0.013610275317068496),
(335, 0.0005360435043225827),
(336, 0.010930057795455587),
(337, 0.0034123745031266845),
(338, 0.008406731055595628),
(339, 0.00013074231812745918),
(340, 0.0003660784907568857),
(341, 0.000954418922330452),
(342, 0.00043144964982061527),
(345, 7.84453908764755e-05),
(348, 5.229692725098367e-05),
(349, 0.00014381654994020512),
(350, 3.922269543823775e-05),
(351, 0.00030070733169315617),
(352, 0.0010197900813941816),
(353, 0.006171037415616073),
(354, 2.6148463625491837e-05),
(355, 0.003190112562310004),
(356, 0.01380638879425969),
(357, 0.013976353807825385),
(358, 0.01690498173388047),
(359, 0.005517325824978776),
(360, 0.010956206625908108),
(361, 0.009844896554997677),
(362, 0.006916268628942591),
(363, 0.0074000152060141895),
(364, 0.010237123509380056),
(365, 0.008445953751033863),
(366, 0.007962207173962264),
(367, 0.016578125938561825),
(368, 0.008642067228225052),
(369, 0.013910982648761654),
(370, 0.004379867657269883),
(371, 0.0031901256623100025),
(372, 0.00026148463625491836),
(373, 0.0015296851220912725),
(374, 0.0015558335857167643),
(375, 0.0007713796769520092),
(376, 0.0035431168212541435),
(377, 0.0005752661997608205),
(378, 0.0013466458767128296),
(379, 0.0013466458767128296),
(380, 0.005268915420536604),
(381, 0.005268915420536604),
(382, 0.00037915272256963163),
(383, 5.229692725098367e-05),
(384, 2.6148463625491837e-05),
(385, 0.0002484104044421724),
(386, 0.0004183754180786933),
(387, 0.0004706723452588531),
(388, 0.003686933371194349),
(390, 0.0005621919679480744),
(391, 0.0022879905672305353),
(392, 0.0007060085178882795)

```
(392, 0.0007060085178882793),
(393, 9.151962268922142e-05),
(394, 0.008733586850914275),
(395, 7.84453908764755e-05),
(396, 0.011806031326909563),
(397, 0.011806031326909563),
(398, 0.004759020379839514),
(399, 0.000797528140577501),
(400, 0.0007583054451392631),
(401, 0.0009674931541431978),
(402, 9.151962268922142e-05),
(403, 5.229692725098367e-05),
(404, 0.00035300425894413975),
(405, 0.0006406373588245499),
(406, 0.000313781563505902),
(407, 0.0005229692725098367),
(408, 0.0011897550949598785),
(409, 0.0005621919679480745),
(410, 0.0016473532084059857),
(411, 0.00045759811344610714),
(412, 0.0004183754180078694),
(413, 0.0010328643132069275),
(414, 0.002967850621493323),
(415, 0.0005621919679480745),
(416, 0.00019611347719118876),
(417, 0.0002222619408166806),
(418, 0.00033993002713139387),
(419, 0.000313781563505902),
(420, 0.0006144888951990581),
(421, 0.000156890781752951),
(422, 0.00028763309980041023),
(423, 0.0010720870086451653),
(424, 0.00088904776326672224),
(425, 0.000627563127011804),
(426, 0.0013727943403383216),
(427, 0.001268200485836354),
(428, 0.0013597201085255753),
(429, 0.00030070733169315617),
(430, 0.0003399300271313938),
(431, 0.00013074231812745918),
(432, 0.00049682080888443448),
(433, 0.00013074231812745918),
(434, 0.0004837465770715989),
(435, 0.003294706416811971),
(436, 0.0012682004858363541),
(437, 0.001712724367469715),
(438, 0.00010459385450196735),
(439, 0.0008236766042029928),
(440, 0.0015035366584657806),
(441, 0.0005098950406970909),
(442, 0.0008498250678284847),
(443, 0.00010459385450196735),
(444, 0.00045759811344610714),
(445, 0.0002091877090039347),
(446, 0.00028763309980041023),
(447, 0.0013204974130873377),
(448, 5.229692725098367e-05),
(449, 0.0008236766042029928),
(450, 0.00010459385450196735),
(451, 7.84453908764755e-05),
(452, 0.000156890781752951),
(453, 0.00028763309980041023),
(454, 0.003556191053066889),
(455, 0.000156890781752951),
(456, 7.84453908764755e-05),
(457, 7.84453908764755e-05),
(458, 0.0013727943403383214),
(459, 0.0009413446905177062),
(460, 0.00018303924537844285),
(461, 0.008197543346591691),
(462, 0.002052654394601109),
(463, 0.000313781563505902),
(464, 0.0009034294182607443)
```

(464, 0.009342941826074),

(465, 0.0034254487349394303),
(466, 0.004170679948265948),
(467, 0.0020526543946011093),
(468, 0.00010459385450196735),
(469, 0.00044452388163336112),
(470, 0.00019611347719118878),
(471, 0.0019611347719118875),
(472, 0.00010459385450196735),
(473, 0.0012812747176490999),
(474, 0.0012028293267726247),
(476, 0.009243481891611364),
(477, 0.00614488895199058),
(478, 0.00020918770900393467),
(479, 0.00037915272256963163),
(480, 0.00041837541800786933),
(481, 0.00018303924537844285),
(482, 0.004667500757150292),
(483, 0.00621026011105431),
(484, 0.00570036507035722),
(485, 0.0005491177361353286),
(486, 0.005543474288604269),
(487, 0.0007452312133265173),
(488, 0.00030070733169315617),
(489, 0.0015427593539040182),
(490, 0.006981639788006319),
(491, 0.001412017035776559),
(492, 7.84453908764755e-05),
(493, 0.0002614846362549183),
(494, 0.002222619408166806),
(495, 0.0013989428039638133),
(496, 0.00035300425894413975),
(497, 5.229692725098367e-05),
(499, 0.0011374581677088948),
(500, 0.000156890781752951),
(502, 0.0010067158495814356),
(503, 2.6148463625491837e-05),
(504, 0.0008890477632667224),
(505, 0.005726513533982711),
(506, 0.00023533617262942654),
(507, 0.00028763309988041023),
(508, 0.006785526310815131),
(509, 0.00048374657707159885),
(510, 0.0009544189223304521),
(511, 2.6148463625491837e-05),
(512, 7.84453908764755e-05),
(513, 0.000156890781752951),
(515, 0.00014381654994020512),
(516, 0.0006798600542627877),
(517, 0.009112739573483908),
(518, 0.005857255585211017),
(519, 0.0006537115906372958),
(520, 0.0006929342860755336),
(521, 0.0012289777903981161),
(522, 0.00028763309988041023),
(523, 0.004863614234341481),
(524, 0.0013335716449000836),
(525, 0.0007713796769520092),
(526, 0.0005883404315735662),
(527, 0.00835443412834464),
(528, 0.0013204974130873377),
(529, 0.0012420520222108623),
(530, 3.922269543823775e-05),
(531, 0.010485533913822228),
(532, 3.922269543823775e-05),
(533, 0.0010590127768324194),
(534, 0.0006275631270118041),
(535, 0.0007321569815137713),
(536, 0.00035300425894413975),
(537, 0.00011768086314713271),
(538, 0.0006275631270118041),
(539, 0.010864686636391858),
(540, 0.0026671432898901673)

(540, 0.002667145289601672),
(541, 0.005412731970476809),
(542, 0.005582696984042506),
(543, 0.001856540174099202),
(544, 0.0030985929396207825),
(545, 0.0020788028582266007),
(546, 6.537115906372959e-05),
(547, 9.151962268922142e-05),
(548, 0.006249482806492548),
(549, 0.0003399300271313938),
(550, 0.0002745588680676643),
(551, 0.0027325144488638967),
(552, 0.0005752661997608205),
(553, 0.0010197900813941816),
(554, 0.00045759811344610714),
(555, 9.151962268922142e-05),
(556, 0.0012420520222108623),
(557, 0.0004706723452588531),
(558, 0.0004968208088843448),
(559, 0.0012420520222108623),
(560, 0.0006798600542627876),
(561, 0.00018303924537844285),
(562, 0.000156890781752951),
(563, 0.0006667858224500418),
(564, 9.151962268922142e-05),
(565, 0.007753019464958328),
(566, 0.006327928197369023),
(567, 0.007033936715257303),
(568, 0.003033221780557052),
(569, 0.0026932917534256587),
(570, 0.008014504101213247),
(571, 0.0004968208088843448),
(572, 0.00022226194081668806),
(573, 0.0004968208088843448),
(574, 0.00040530118619512345),
(575, 0.0014643139630275428),
(576, 0.00019611347719118876),
(577, 0.004981282320656194),
(578, 0.0036215622121306187),
(579, 0.0028894052306168475),
(580, 0.004824391538903245),
(581, 0.0012159035585853702),
(582, 0.0005883404315735662),
(583, 0.004144531484640456),
(584, 0.004641352293524801),
(585, 0.0035954137485051272),
(586, 0.007321569815137714),
(587, 0.004589055366273817),
(588, 0.006184111647428819),
(589, 0.0010328643132069275),
(590, 0.004013789166512996),
(591, 0.000156890781752951),
(592, 0.008367508360157387),
(593, 0.0037000076030070947),
(594, 0.0018957636128481578),
(595, 0.000470672345258853),
(596, 0.0019611347719118875),
(598, 0.0006144888951990581),
(600, 0.00019611347719118876),
(601, 0.0004183754180786933),
(602, 0.0016996501356569694),
(603, 0.0006798600542627877),
(604, 0.0023795101899197566),
(605, 7.84453908764755e-05),
(606, 0.0016081305129677479),
(608, 0.0006929342860755336),
(609, 0.0004053011861951234),
(610, 0.0006667858224500417),
(611, 9.151962268922142e-05),
(612, 0.0034515971985649217),
(614, 6.537115906372959e-05),
(615, 6.537115906372959e-05),
(616, 0.007020862483444558)

(616, 0.00702080248544558),
(617, 0.0004968208088843448),
(618, 0.0004968208088843448),
(619, 0.003634636443943365),
(620, 9.151962268922142e-05),
(621, 0.0004445238816333612),
(622, 2.6148463625491837e-05),
(623, 0.010655498927387924),
(624, 0.00028763309988041023),
(626, 0.0014250912675893052),
(627, 0.0010459385450196734),
(628, 0.00227491633541779),
(629, 0.000797528140577501),
(630, 0.00019611347719118878),
(631, 7.84453908764755e-05),
(632, 0.00013074231812745918),
(633, 0.0024579555807962322),
(634, 0.0016996501356569692),
(635, 0.00018303924537844285),
(636, 0.000156890781752951),
(637, 0.00013074231812745918),
(638, 0.003569265284879635),
(639, 0.000156890781752951),
(640, 0.0013597201085255753),
(641, 0.00014381654994020506),
(642, 0.01181910555872231),
(643, 0.0047982430752777526),
(644, 0.00030070733169315606),
(645, 0.0013466458767128296),
(646, 0.0010590127768324194),
(647, 0.0002745588680676643),
(648, 0.000483746577071599),
(649, 3.922269543823775e-05),
(650, 0.0002484104044421724),
(651, 0.0016342789765932395),
(653, 0.0015427593539040182),
(654, 0.0016081305129677476),
(656, 0.002366435958107011),
(657, 0.0016473532084059857),
(658, 0.00011766808631471327),
(659, 0.0002745588680676643),
(660, 0.0033993002713139384),
(661, 0.00040530118619512345),
(662, 0.0003791527225696316),
(663, 0.0005229692725098367),
(664, 0.00018303924537844285),
(665, 0.00011766808631471327),
(666, 6.537115906372959e-05),
(667, 0.0009805673859559438),
(668, 0.00043144964982061527),
(669, 0.0026017721307364378),
(670, 0.0003791527225696316),
(671, 0.0008759735314539765),
(672, 0.00011766808631471327),
(673, 0.00030070733169315617),
(674, 0.00013074231812745918),
(675, 3.922269543823775e-05),
(676, 0.00014381654994020512),
(677, 0.0013989428039638133),
(678, 0.003752304530258078),
(679, 0.0010590127768324194),
(680, 0.0039484180074494267),
(681, 0.00013074231812745918),
(682, 0.00013074231812745918),
(683, 0.0003399300271313937),
(684, 0.00032685579531864794),
(685, 0.0016212047447804938),
(686, 0.0012159035585853702),
(687, 0.0005360435043225827),
(688, 0.0008367508360157388),
(689, 0.0009544189223304527),
(690, 0.0005229692725098367),
(691, 0.0004575981134610714)

```
(691, 0.0004575981134610714),
(692, 0.005582696984042505),
(693, 0.0004837465770715989),
(694, 0.0003007073169315617),
(695, 6.537115906372959e-05),
(696, 0.0027848113761148804),
(697, 2.6148463625491837e-05),
(698, 2.6148463625491837e-05),
(699, 2.6148463625491837e-05),
(700, 0.0052689154205366046),
(701, 0.000313781563505902),
(702, 0.0004183754180078694),
(703, 0.010224049277567308),
(704, 0.0006014146633863121),
(705, 0.00400071493470025),
(706, 0.00400071493470025),
(707, 0.001751947062907953),
(708, 0.0012812747176490999),
(709, 0.0005752661997608205),
(710, 0.00011766808631471326),
(711, 0.0015035366584657806),
(712, 0.004275273802767915),
(713, 0.0015819820493422558),
(714, 0.00030070733169315606),
(715, 0.0024187328853579946),
(716, 0.00019611347719118878),
(717, 0.0002091877090039347),
(718, 0.00014381654994020512),
(719, 0.0001699650135656969),
(720, 0.00019611347719118876),
(721, 0.008877403400854479),
(722, 0.001451239731214797),
(723, 0.0005491177361353286),
(724, 6.537115906372959e-05),
(725, 0.0010197900813941816),
(726, 0.0009021219950794683),
(727, 0.0012812747176490999),
(730, 0.000156890781752951),
(731, 0.00023533617262942654),
(732, 0.0021572482491030763),
(733, 6.537115906372959e-05),
(734, 0.0013074231812745915),
(735, 0.00040530118619512345),
(736, 9.151962268922142e-05),
(737, 0.0007190827497010254),
(738, 0.00022226194081668063),
(739, 0.0002222619408166806),
(740, 7.84453908764755e-05),
(741, 0.0003791527225696316),
(742, 0.002876330998804102),
(743, 0.0016473532084059857),
(744, 0.00192191207647365),
(745, 2.6148463625491837e-05),
(746, 0.0009021219950794682),
(747, 0.0005229692725098367),
(748, 0.009217333427985872),
(749, 0.002052654394601109),
(750, 0.001725798599282461),
(751, 0.001725798599282461),
(752, 0.0003399300271313938),
(753, 0.00035300425894413975),
(754, 0.0008106023723902468),
(755, 0.0053342865779600334),
(756, 0.005242766956911113),
(757, 0.000483746577071599),
(758, 0.000313781563505902),
(759, 0.00018303924537844285),
(760, 0.0002484104044421724),
(762, 0.00010459385450196735),
(763, 0.0003791527225696163),
(764, 0.005739587765795457),
(765, 0.00100671584958144356),
(766, 0.0003399300271313938)
```

(766, 0.0003399300271313938),
(767, 0.00024841040444421724),
(768, 0.0047851688434650055),
(769, 0.001477388194840289),
(770, 0.0005752661997608205),
(771, 0.0016081305129677476),
(772, 0.0008628992996412305),
(773, 0.00030070733169315606),
(774, 0.000810602372390247),
(775, 0.0003399300271313938),
(776, 0.003726156066632586),
(777, 0.0022618421036050434),
(778, 0.0011766808631471328),
(779, 0.00040530118619512345),
(780, 0.00035300425894413975),
(781, 0.004170679948265948),
(782, 0.0003399300271313938),
(783, 0.0002745588680676643),
(784, 0.000954418922330452),
(785, 0.0036084879803178734),
(786, 0.0036084879803178734),
(787, 0.00024841040444421724),
(788, 0.0020657286264413855),
(789, 0.00406608609376398),
(790, 0.0002745588680676643),
(791, 0.0005491177361353286),
(792, 0.0062625570038305295),
(793, 0.00040530118619512345),
(794, 0.0008890477632667224),
(795, 0.000313781563505902),
(796, 0.0051904700296601285),
(797, 0.0009282704587049601),
(798, 0.00028763309988041023),
(799, 0.0004183754180078694),
(800, 0.0005360435043225827),
(801, 0.0005360435043225827),
(802, 0.002549475203485454),
(803, 0.000483746577071599),
(804, 0.0006144888951990581),
(805, 0.002235693639979552),
(806, 0.0005098950406970908),
(807, 0.0019349863082863956),
(808, 0.00024841040444421724),
(809, 0.004170679948265948),
(810, 0.0010067158495814356),
(811, 0.00028763309988041023),
(812, 0.0040137891665512996),
(813, 0.000156890781752951),
(814, 0.0023141390308560276),
(815, 0.0017780955265334448),
(816, 6.537115906372959e-05),
(817, 0.0008367508360157389),
(818, 0.000980567385955944),
(819, 0.0036477106757561105),
(820, 0.0003268557953186479),
(821, 0.002000357467350125),
(822, 0.0020134316991628712),
(824, 0.00026148463625491836),
(825, 0.0003268557953186479),
(826, 0.0003922269543823775),
(827, 0.0004183754180786933),
(828, 0.00023533617262942654),
(829, 0.0005491177361353286),
(830, 0.00013074231812745918),
(831, 0.00026148463625491836),
(832, 0.0003399300271313938),
(833, 0.003360077557587570),
(834, 0.0005752661997608202),
(835, 0.0017780955265334446),
(836, 0.00020918770900039347),
(837, 0.007386940974201442),
(838, 0.007386940974201442),
(839, 0.005583696984042507)

(839, 0.0055826909840425007),
(840, 0.00040530118619512345),
(841, 0.004510609975397341),
(842, 0.0008498250678284847),
(843, 0.00037915272256963163),
(844, 2.6148463625491837e-05),
(845, 0.00014381654994020512),
(846, 0.0003660784907568857),
(847, 0.008864329169041733),
(848, 0.001451239731214797),
(849, 7.84453908764755e-05),
(850, 0.0022095451763540597),
(852, 5.229692725098367e-05),
(853, 0.003451597198564922),
(854, 0.0004183754180078694),
(855, 0.0023402874944815195),
(856, 0.0005752661997608205),
(857, 9.151962268922142e-05),
(858, 0.014969995425594075),
(859, 0.013937131112387148),
(860, 0.013937131112387148),
(861, 0.0027848113761148804),
(862, 0.004392941889082629),
(863, 0.009086591109858412),
(864, 6.537115906372959e-05),
(865, 0.0027848113761148804),
(866, 0.008916626096292714),
(867, 0.0008890477632667224),
(868, 0.0013727943403383214),
(871, 2.6148463625491837e-05),
(872, 0.0002876330998804101),
(874, 0.0024318071171707404),
(875, 0.0015689078175295103),
(876, 0.00040530118619512345),
(877, 7.84453908764755e-05),
(878, 0.0008628992996412305),
(880, 6.537115906372959e-05),
(881, 0.0007713796769520092),
(882, 0.0019742090037246332),
(884, 0.0001699650135656969),
(885, 0.0013204974130873377),
(886, 0.0007190827497010254),
(887, 0.000313781563505902),
(889, 0.0018303924537844283),
(891, 3.922269543823775e-05),
(892, 0.0014643139630275428),
(893, 9.151962268922142e-05),
(894, 0.0004314496498206153),
(895, 0.0007190827497010254),
(896, 0.0005752661997608205),
(897, 0.00313781563505902),
(898, 0.0006798600542627877),
(899, 0.0006537115906372958),
(900, 0.0008236766042029928),
(901, 0.0008236766042029928),
(902, 0.00040530118619512345),
(904, 0.00019611347719118878),
(905, 0.0013335716449000836),
(906, 0.0007583054451392631),
(908, 0.0022487678717922977),
(909, 7.84453908764755e-05),
(910, 0.002758662912489388),
(911, 0.002850182553517861),
(912, 0.0068378232380066114),
(913, 0.00018303924537844285),
(914, 0.000313781563505902),
(915, 5.229692725098367e-05),
(916, 0.0016604274402187316),
(918, 3.922269543823775e-05),
(919, 0.0002484104044421724),
(920, 7.84453908764755e-05),
(921, 3.922269543823775e-05),
(922, 6.537115906372959e-05)

```
(922, 6.33711390037293900-03),
(923, 0.0013858685721510671),
(924, 0.0003660784907568857),
(925, 7.84453908764755e-05),
(926, 0.0010197900813941816),
(927, 0.0018826893810354123),
(928, 0.0017519470629079532),
(929, 2.6148463625491837e-05),
(930, 0.00014381654994020512),
(931, 0.000313781563505902),
(932, 0.00010459385450196735),
(933, 0.0008498250678284847),
(934, 0.000627563127011804),
(935, 0.00018303924537844285),
(936, 0.0003791527225696316),
(937, 7.84453908764755e-05),
(938, 0.000470672345258853),
(939, 0.0013989428039638133),
(940, 0.0007321569815137714),
(941, 0.000797528140577501),
(942, 0.0071908274970102545),
(943, 5.229692725098367e-05),
(944, 2.6148463625491837e-05),
(945, 0.0005883404315735664),
(946, 5.229692725098367e-05),
(947, 5.229692725098367e-05),
(948, 5.229692725098367e-05),
(949, 5.229692725098367e-05),
(950, 5.229692725098367e-05),
(951, 5.229692725098367e-05),
(952, 5.229692725098367e-05),
(953, 0.002222619408166806),
(954, 0.00013074231812745918),
(955, 0.005569622752229762),
(956, 5.229692725098367e-05),
(957, 0.00026148463625491836),
(958, 0.0021572482491030763),
(959, 0.0012812747176490999),
(960, 0.0009805673859559438),
(961, 0.0006406373588245499),
(962, 0.0017650212947206987),
(963, 0.001581982049342256),
(964, 9.151962268922142e-05),
(965, 0.0020134316991628712),
(966, 0.00013074231812745918),
(967, 0.0003791527225696316),
(968, 0.0004452388163336126),
(969, 0.0002222619408166806),
(970, 0.000156890781752951),
(971, 5.229692725098367e-05),
(972, 5.229692725098367e-05),
(973, 5.229692725098367e-05),
(974, 0.00010459385450196735),
(975, 0.0003660784907568857),
(976, 0.004693649220775784),
(977, 0.0004968208088843448),
(978, 0.00014381654994020512),
(979, 0.002301064799043282),
(980, 0.0005229692725098367),
(981, 0.0016473532084059857),
(982, 5.229692725098367e-05),
(983, 0.0001699650135656969),
(984, 0.0015558335857167643),
(985, 0.000313781563505902),
(986, 0.0018696151492226662),
(987, 0.00045759811344610714),
(988, 0.0032424094895609876),
(989, 0.001738872831095207),
(990, 0.003608487980317873),
(991, 0.000810602372390247),
(992, 0.0034908198940031597),
(993, 0.0007060085178882794),
(994, 0.0001699650135656969)
```

```
(994,  0.00016996501356569694),
(995,  0.00045759811344610714),
(997,  0.0074523121332265172),
(998,  0.007478460596890665),
(999,  0.00023533617262942654),
(1000, 0.00065371159063722958),
(1001, 0.00010459385450196735),
(1002, 0.00018303924537844285),
(1003, 0.00016996501356569694),
(1004, 0.00010459385450196735),
(1005, 0.0012028293267726247),
(1007, 2.6148463625491837e-05),
(1008, 0.004366793425457136),
(1009, 0.0012682004858363541),
(1010, 0.00044452388163336126),
(1011, 0.0005883404315735662),
(1012, 0.000967493154143198),
(1014, 0.005974923938424884),
(1015, 0.00614488895199058),
(1016, 0.00023533617262942654),
(1017, 5.229692725098367e-05),
(1018, 0.01460391693483719),
(1019, 0.00035300425894413975),
(1020, 0.00011766808631471326),
(1021, 0.00037915272256963163),
(1022, 0.018656928796788423),
(1023, 2.6148463625491837e-05),
(1024, 0.00841980528740837),
(1025, 0.0006929342860755336),
(1026, 0.0006929342860755336),
(1027, 0.0005883404315735662),
(1028, 0.0005883404315735662),
(1029, 0.0005883404315735662),
(1030, 0.0005883404315735662),
(1031, 0.0005621919679480744),
(1032, 7.84453908764755e-05),
(1033, 0.00010459385450196733),
(1034, 0.0016735016720314773),
(1036, 9.151962268922142e-05),
(1037, 9.151962268922142e-05),
(1038, 9.151962268922142e-05),
(1039, 0.0016473532084059859),
(1042, 0.00030070733169315617),
(1043, 0.0013466458767128296),
(1044, 0.004928985393405211),
(1045, 0.004928985393405211),
(1046, 0.008171394882966199),
(1047, 0.008092949492089725),
(1048, 0.008092949492089725),
(1049, 0.008092949492089725),
(1050, 0.001255126254023608),
(1051, 0.0011374581677088948),
(1052, 0.0014250912675893052),
(1053, 0.0006667858224500418),
(1054, 0.0017650212947206989),
(1055, 0.00014381654994020512),
(1056, 0.0007060085178882795),
(1057, 0.0011766808631471328),
(1058, 0.0009805673859559438),
(1059, 0.0002353361726294265),
(1060, 0.0008498250678284847),
(1061, 0.0009413446905177062),
(1062, 0.000954418922330452),
(1063, 0.000222619408166806),
 ...]
```

In [0]:

```python
#Join the DFs so that we get (id, [hamProb, spamProb]) where each row shows the probabili
ty of word with ID in ham and spam class
combinedProb = testSpamProbability.join(testHamProbability).sortByKey("id")
combinedProb.collect()
```

```
Out[94]: [(2, (0.0006144888951990581, 0.0007040136804813849)),
 (3, (0.004379867657269882, 0.01644831962579236)),
 (4, (0.005831107388484678, 0.023232451455885702)),
 (5, (0.005621919679480744, 0.020224393002919783)),
 (6, (0.003961492239262013, 0.04460886684504775)),
 (7, (0.004471387279959104, 0.022080429069643432)),
 (8, (0.0007321569815137714, 0.014848288533789212)),
 (9, (0.00494205962521795, 0.03008058452965917)),
 (10, (0.013976353807825383, 0.1979558467026294)),
 (11, (0.006837823238066114, 0.011584225106102788)),
 (12, (0.003360077575875701, 0.026944523589333004)),
 (13, (0.005491177361353286, 0.07065737302285899)),
 (14, (0.003765378762070824, 0.029824579554938665)),
 (15, (0.0022618421036050434, 0.030016583285979042)),
 (16, (0.002876330998804102, 0.042176819585200297)),
 (17, (0.0060272208656758685, 0.01203223381186367)),
 (18, (0.009008145718981937, 0.05920115040411645)),
 (20, (0.01595056281155002, 0.14829088160685164)),
 (21, (0.01172758593603309, 0.030976601941180935)),
 (22, (0.0089689230235437, 0.05920115040411645)),
 (23, (0.00018303924537844285, 0.012416241273944424)),
 (25, (0.000496820808884345, 0.025984504934131115)),
 (26, (0.0016473532084059857, 0.02624050990885161)),
 (27, (0.008668215691850542, 0.031936620596382824)),
 (28, (3.922269543823775e-05, 0.00422408208288831)),
 (29, (0.00018303924537844285, 0.012992252467065557)),
 (30, (0.0004445238816333612, 0.00857616665313687)),
 (31, (0.007517683292328903, 0.008640167896816998)),
 (32, (0.00045759811344610714, 0.0035840696460870502)),
 (33, (0.0005491177361353286, 0.09452983691554596)),
 (34, (0.0008498250678284847, 0.021312414145481923)),
 (35, (0.0004575981134461072, 0.02560049747205036)),
 (36, (0.0010197900813941816, 0.014272277340668077)),
 (37, (0.0021049513218520926, 0.12083434806807768)),
 (38, (0.0035823395166923815, 0.09849791402371376)),
 (39, (2.6148463625491837e-05, 0.00032000621840062946)),
 (40, (0.0004968208088843448, 0.04025678227479919)),
 (41, (0.0035692652848796354, 0.016576322113152606)),
 (42, (0.0008890477632667224, 0.018752364398276888)),
 (43, (0.0005229692725098367, 0.025664498715730483)),
 (44, (5.229692725098367e-05, 0.016128313407391726)),
 (45, (0.0016473532084059857, 0.03097660194118093)),
 (46, (6.537115906372959e-05, 0.008064156703695863)),
 (47, (5.229692725098367e-05, 0.016832327087873113)),
 (48, (0.0012028293267726242, 0.031936620596382824)),
 (49, (0.002536400971672708, 0.04800093276009443)),
 (50, (0.0005621919679480745, 0.027392532295093884)),
 (51, (0.00014381654994020512, 0.00966418779569901)),
 (52, (0.0007060085178882795, 0.02515248876628948)),
 (53, (0.0017127243674697151, 0.069249345661896221)),
 (54, (0.003608487980317873, 0.08224159812896177)),
 (55, (0.0075438317559543945, 0.15366698607598228)),
 (56, (0.0002222619408166806, 0.021440416632842175)),
 (57, (0.00037915272256963163, 0.016128313407391726)),
 (58, (0.0006798600542627877, 0.01184023008082329)),
 (59, (0.00040530118619512345, 0.03475267531830836)),
 (60, (0.0002091877090039347, 0.00640012436801259)),
 (61, (0.00026148463625491836, 0.009152177846258004)),
 (62, (0.0002091877090039347, 0.021632420363882552)),
 (63, (0.003935343775636521, 0.09619386925122919)),
 (64, (3.922269543823775e-05, 0.0037760733771274283)),
 (65, (0.00011768808631471327, 0.016768325844192983)),
 (66, (0.001595056281155002, 0.014976291021149458)),
 (67, (2.6148463625491837e-05, 0.003840074620807554)),
 (68, (0.0005360435043225827, 0.01388826987858732)),
 (70, (0.00016996501356569694, 0.00947218406465863)),
 (71, (0.00014381654994020512, 0.006656129342733093)),
 (73, (0.0004183754180078693, 0.020928406668340117)),
 (74, (0.0005098950406970908, 0.05811312926155431)),
 (77, (0.000627563127011804, 0.038912756115751655)),
 (78, (0.0015427593539040182, 0.057345114337392804)),
 (79, (0.00018303924537844285, 0.0046720907886491905)),
```

(79, (0.00018303924537844285, 0.0046720908649190S)),
(81, (0.0016865759038442233, 0.060801181496119606)),
(83, (0.00011766808631471326, 0.01542429972691034)),
(84, (0.00019611347719118876, 0.021952426582283184)),
(85, (0.002248767871792298, 0.010432202719860522)),
(87, (0.0007060085178882795, 0.02592050369045099)),
(88, (0.0015819820493422558, 0.05305703101082437)),
(89, (0.0007583054451392631, 0.01894436812931727)),
(90, (0.000797528140577501, 0.014464281071708454)),
(91, (6.537115906372959e-05, 0.0028160547219255394)),
(92, (0.003477745662190414, 0.02451247632948822)),
(93, (0.002967850621493323, 0.0622732101007625)),
(95, (5.229692725098367e-05, 0.0016640323356832732)),
(96, (0.0033208548804374628, 0.03270463552054433)),
(97, (0.0004968208088843448, 0.008768170384177248)),
(98, (0.00045759811344610714, 0.019520379322438396)),
(99, (0.0008367508360157387, 0.03814474123335503)),
(100, (0.00018303924537844285, 0.007104138048493976)),
(101, (0.011819105558722311, 0.05209701235562248)),
(102, (0.00026148463625491836, 0.014912289777469336)),
(103, (7.84453908764755e-05, 0.0005760119312211331)),
(104, (0.001111309704083403, 0.03500680029302887)),
(105, (0.00013074231812745918, 0.003712072133447302)),
(106, (0.00019611347719118878, 0.010688207694581025)),
(107, (0.0003660784907568857, 0.030528593235420055)),
(108, (0.001124383935896149, 0.07347342774478455)),
(109, (0.00014381654994020512, 0.01139222137506241)),
(110, (0.00013074231812745918, 0.006848133073773471)),
(112, (0.0033470033440629546, 0.037824735014954404)),
(113, (0.0019349863082863956, 0.02809654597557527)),
(114, (0.001581982049342256, 0.015104293508509713)),
(115, (0.0008367508360157388, 0.05401704966602625)),
(116, (0.00881203224179075, 0.06355323497436503)),
(117, (0.00036607849075688564, 0.09107376975681915)),
(118, (2.6148463625491837e-05, 0.03788873625863452)),
(119, (0.0059226270111739, 0.10739408689525128)),
(120, (0.0031508898668717662, 0.03443266909990774)),
(121, (0.0020134316991628712, 0.07724950112191195)),
(122, (0.0015166108902785265, 0.00748814551057473)),
(123, (0.009544189223304519, 0.050432980019939205)),
(124, (0.0002745588680676643, 0.024448475085808094)),
(125, (0.0007060085178882795, 0.04012877978743894)),
(127, (0.0002745588680676643, 0.01644831962579236)),
(129, (0.00011766808631471326, 0.0063361231243324635)),
(130, (0.00428834803458066, 0.1183382995645528)),
(131, (0.009374224209738822, 0.2046759772890426)),
(132, (7.84453908764755e-05, 0.011776288837143165)),
(133, (0.00014381654994020512, 0.007040136804813848)),
(134, (0.0002353361726294265, 0.011200217644022034)),
(135, (0.0002876330998804101, 0.017152333306273738)),
(136, (0.003307780648624717, 0.10585805704692823)),
(137, (0.007583054451392631, 0.044032855651926606)),
(138, (0.0007452312133265172, 0.0540810509097063B5)),
(139, (0.003360077575875701, 0.010496203963540647)),
(140, (0.009439595368802553, 0.030528593235420055)),
(141, (0.00400071493470025, 0.017408338280994248)),
(142, (0.003647710675756111, 0.010496203963540647)),
(143, (0.011230765127148745, 0.0291845671181374)),
(144, (0.004759020379839514, 0.042816832022004224)),
(145, (0.003360077575875701, 0.010816210181941276)),
(146, (0.006863971701691607, 0.026304511152531747)),
(147, (0.0036869333711943486, 0.036096701435591)),
(148, (0.00016996501356569694, 0.0057601193122113331)),
(149, (0.0003660784907568857, 0.018496359423556385)),
(150, (0.000627563127011804, 0.0017920348230435251)),
(151, (0.000627563127011804, 0.0017920348230435251)),
(152, (0.000627563127011804, 0.0017920348230435251)),
(154, (0.005543474288604268, 0.01344026117282644)),
(156, (0.0016342789765932395, 0.02822454846293552)),
(157, (0.0016342789765932395, 0.02822454846293552)),
(158, (0.0016342789765932395, 0.02822454846293552)),
(160, (0.006367150892807262, 0.08844971876593401)),
(161, (0.00036607849075688857, 0.02976057831125854))

```
(161, (0.0003680784907508857, 0.02976057031125854)),
(162, (0.00013074231812745918, 0.044224859382967)),
(163, (0.00011766808631471327, 0.00396807710816806)),
(164, (0.00028763309988041023, 0.020160391759239657)),
(165, (0.0020265059309756717, 0.0249604850352491)),
(166, (0.0020657286264413855, 0.017152333306273738)),
(167, (0.0003922269543823775, 0.014464281071708454)),
(169, (0.00019611347719118876, 0.010944212669301527)),
(170, (0.00011766808631471327, 0.005952115662251709)),
(171, (6.537115906372959e-05, 0.0015360298483230216)),
(172, (6.537115906372959e-05, 0.002112041041444155)),
(173, (6.537115906372959e-05, 0.002112041041444155)),
(174, (0.000313781563505902, 0.015360298483230216)),
(175, (0.0021049513218520926, 0.05817713050523444)),
(176, (0.0021441740172903306, 0.014208276096987951)),
(177, (0.0038046014575090614, 0.10912212047461467)),
(178, (0.008432879519221117, 0.021824424094922936)),
(179, (0.00013074231812745918, 0.009920192770419514)),
(180, (0.0010982354722706572, 0.04915295514633669)),
(181, (0.001412017035776559, 0.026368512239621187)),
(182, (2.6148463625491837e-05, 0.00998419401409964)),
(183, (0.00030070733169315617, 0.024576477573168342)),
(184, (0.0005883404315735664, 0.016768325844192983)),
(185, (0.0009936416177686897, 0.03059259447910018)),
(186, (0.00028763309988041023, 0.032960640495264835)),
(187, (0.00016996501356569694, 0.01516829475218984)),
(188, (0.00016996501356569694, 0.019904386778451915)),
(189, (0.0015558335857167643, 0.04198481585416259)),
(190, (0.00018303924537844285, 0.03788873625863452)),
(191, (2.6148463625491837e-05, 0.010176197745140017)),
(192, (9.151962268922142e-05, 0.019136371860357645)),
(193, (0.000156890781752951, 0.013184256198105935)),
(194, (0.0016996501356569694, 0.06873733571245522)),
(195, (2.6148463625491837e-05, 0.0009600186552018885)),
(196, (0.009844896554997675, 0.04544088301288938)),
(197, (0.000797528140577501, 0.029685574580218162)),
(198, (0.00013074231812745918, 0.0017920348230435251)),
(199, (0.013623349548881246, 0.10694607818949038)),
(200, (0.010786241245515384, 0.029888580798618794)),
(201, (0.010786241245515384, 0.029888580798618794)),
(202, (0.008289062969280912, 0.015744305945310972)),
(205, (0.00010459385450196733, 0.0036480708897671764)),
(206, (0.0007060085178882795, 0.06873733571245523)),
(207, (0.0006014146633863123, 0.036096701435591)),
(208, (0.0016342789765932397, 0.0862096752371296)),
(209, (0.0012289779903981161, 0.005376104469130575)),
(210, (0.0013727943403383214, 0.008704169140497122)),
(211, (0.006602487065436688, 0.22688440884604633)),
(212, (0.0022618421036050434, 0.06310522626860413)),
(213, (0.0002222619408166806, 0.0197763842971589)),
(214, (0.006184111647428819, 0.0212484129018018)),
(215, (0.0018826893810354123, 0.017344337037314115)),
(216, (3.922269543823775e-05, 0.0028800559656056656)),
(217, (0.009151962268922143, 0.04627289918073103)),
(218, (0.001255126254023608, 0.03974477232535818)),
(220, (6.537115906372959e-05, 0.016000310920031475)),
(221, (0.00019611347719118876, 0.01721633454995387)),
(222, (0.00019611347719118876, 0.01721633454995387)),
(223, (0.0035300425894413973, 0.023104449968525447)),
(224, (2.6148463625491837e-05, 0.0035840696460870502)),
(225, (0.001451239731214797, 0.01171222759346304)),
(226, (0.0002484104044421724, 0.0020480397977640286)),
(228, (5.229692725098367e-05, 0.005824113174891456)),
(229, (0.00023533617262942654, 0.015232295995869966)),
(230, (0.007230050192448491, 0.20576399843160478)),
(231, (0.00014381654994020512, 0.0010880211425621403)),
(232, (0.0089950714871169192, 0.0763534837103902)),
(233, (0.000313781563505902, 0.013184256198105936)),
(234, (0.0005621919679480744, 0.020736402952360792)),
(235, (0.0037784529938835695, 0.034112662881507105)),
(236, (0.00019611347719118878, 0.012480242251762455)),
(237, (0.007975281405775008, 0.10489803839172633)),
(238, (0.003621562212131306187, 0.02118441165812167))
```

```
(238, (0.003621362212130618?, 0.021184411638121673)),
(239, (0.007452312133265172, 0.026176508665171495)),
(240, (0.0029155536942423407, 0.026304511152531743)),
(241, (0.0015558335857167645, 0.077057497390871159)),
(242, (0.00026148463625491836, 0.014848288533789208)),
(243, (0.0002484104044421725, 0.019072370061667752)),
(244, (0.0002222619408166806, 0.016896328331553324)),
(245, (0.0002222619408166806, 0.006208120636972212)),
(246, (0.0020265059309756617, 0.07750550609663245)),
(247, (0.0063671508928072615, 0.10464203341700584)),
(248, (0.00016996501356569694, 0.0065921280099052968)),
(249, (0.0127865987128655008, 0.055169072052268552)),
(250, (0.00570036507035722, 0.024320472598447843)),
(251, (0.0011505323995216411, 0.028224548462935552)),
(252, (2.6148463625491837e-05, 0.0076161479979349811)),
(253, (0.0008498250678284846, 0.016192314651071852)),
(254, (0.005621919679480744, 0.105922058290608361)),
(255, (0.00018303924537844285, 0.009664187795699011)),
(256, (0.0006275631270118041, 0.031936620596382824)),
(257, (0.003503894125815906, 0.08940973742113588)),
(258, (0.0009282704587049601, 0.024128468867407465)),
(259, (0.0032293352577482414, 0.03584069646087051)),
(260, (0.0018696151492226662, 0.054721063346507635)),
(262, (2.6148463625491837e-05, 0.0009600186552018885)),
(263, (0.0011897550949598785, 0.010048195257779766)),
(264, (0.001072087008645165, 0.035072681536708984)),
(265, (0.0016081305129677476, 0.0484489414658553)),
(266, (0.0004183754180078694, 0.039872774812718441)),
(267, (0.0033600775758757012, 0.032512631789503955)),
(268, (0.0007190827497010254, 0.033920659150466721)),
(269, (0.0009936416177686897, 0.023680460161646585)),
(271, (0.0007060085178882795, 0.03577669521719037)),
(272, (0.00014381654994020512, 0.006400124368012591)),
(273, (0.0004968208088843448, 0.024576477573168345)),
(274, (0.0006144888951990581, 0.019584380566118522)),
(275, (0.0006798600542627877, 0.025024486278929226)),
(276, (0.000470672345258853, 0.03795273750231466)),
(277, (0.0006667858224500418, 0.00953618530833876)),
(278, (0.00018303924537844285, 0.013312258685466186)),
(279, (0.0006798600542627876, 0.024896483791568974)),
(280, (0.0005229692725098367, 0.023324514558857021)),
(281, (0.0010851612404579112, 0.055553079951434928)),
(282, (0.0004968208088843448, 0.018112351961475627)),
(283, (0.0004968208088843448, 0.020032389271879406)),
(284, (0.0002484104044421724, 0.009856191526739388)),
(285, (0.00841980528740837, 0.014912289777469332)),
(286, (0.00011766808631471327, 0.0009600186552018885)),
(287, (0.000483746577071599, 0.011968232568183542)),
(289, (0.0004183754180078694, 0.023168450212205576)),
(290, (7.84453908764755e-05, 0.011520223862422662)),
(291, (0.001124383935896149, 0.08838571752225387)),
(292, (0.00033993002713139387, 0.027072526076693255)),
(293, (0.010067158495814356, 0.17933148479171276)),
(294, (0.0005229692725098367, 0.026944523589333004)),
(295, (0.0005229692725098367, 0.037376726309193524)),
(296, (0.005465028897727793, 0.09516984935234721)),
(297, (0.002510252508047216, 0.06118518895820035)),
(298, (2.6148463625491837e-05, 0.0037760733771274274)),
(299, (0.007426163669639682, 0.045440883012889385)),
(300, (0.0005752661997608205, 0.025088487522609348)),
(301, (0.00013074231812745918, 0.016704324600512861)),
(302, (0.0006014146633863123, 0.024320472598447843)),
(303, (0.0006014146633863123, 0.024320472598447843)),
(304, (0.0053996577386640063, 0.0012800248736025178)),
(305, (0.0022226194081668062, 0.05900914667307608)),
(306, (0.0016865759038442233, 0.062657217562843263)),
(307, (0.0004968208088843448, 0.004608089544969065)),
(308, (0.0006798600542627876, 0.007296141779534352)),
(309, (0.005465028897727925, 0.0287365584123765331)),
(310, (0.010851612404579113, 0.06118518895820035)),
(311, (0.0032554837213737333, 0.06284922129388364)),
(312, (0.0002091877090039347, 0.009856191526739388)),
(313, (0.001856540917409820?, 0.041856133680234?)),
```

```
(313, (0.001856509174099204, 0.0418681356680234)),
(314, (0.0008628992996412305, 0.0199043867845191154)),
(315, (0.0004837465770715989, 0.028864560899736786)),
(316, (7.84453908764755e-05, 0.005056098250729946)),
(318, (0.0010197900813941816, 0.016832327087873113)),
(319, (0.00014381654994020512, 0.0108802114256214)),
(320, (0.0002222619408166806, 0.013248257441786062)),
(321, (0.004484461511771849, 0.08793770881649297)),
(322, (0.0006537115906372959, 0.04403285565192661)),
(323, (0.0008628992996412305, 0.005632109443851079)),
(324, (0.0001699650135656969, 0.005312103225450449)),
(325, (5.229692725098367e-05, 0.0007680149241615108)),
(326, (0.007086233642508286, 0.20806804320408928)),
(328, (0.00156890781752951, 0.025088487522609355)),
(329, (0.0013204974130873377, 0.015232295995869966)),
(330, (0.0012943489494618458, 0.04230482207256322)),
(331, (0.0003399300271313938, 0.042496825803603595)),
(332, (0.0016604274402187316, 0.02067240170868067)),
(333, (0.0003399300271313938, 0.015936309676351346)),
(334, (0.013610275317068496, 0.366691913001816184)),
(335, (0.0005360435043225827, 0.02016039175923966)),
(336, (0.010930057795455587, 0.2976057831125854)),
(337, (0.0034123745031266845, 0.018368356936196133)),
(338, (0.008406731055595628, 0.028480553437656018)),
(339, (0.00013074231812745918, 0.006592128099052967)),
(340, (0.0003660784907568857, 0.0015360298483230216)),
(341, (0.000954418922330452, 0.01888036688563714)),
(342, (0.00043144964982061527, 0.009280180333618257)),
(345, (7.84453908764755e-05, 0.01939237683507815)),
(348, (5.229692725098367e-05, 0.0028800559656056656)),
(349, (0.00014381654994020512, 0.005440105712810702)),
(350, (3.922269543823775e-05, 0.003072059696646043)),
(351, (0.00030070733169315617, 0.011200217644022034)),
(352, (0.0010197900813941816, 0.018944368129317265)),
(353, (0.006171037415616073, 0.012800248736025182)),
(354, (2.6148463625491837e-05, 0.029824579554938665)),
(355, (0.003190112562310004, 0.06406524492380603)),
(356, (0.01380638879425969, 0.04960096385209757)),
(357, (0.013976353807825385, 0.06131319144556062)),
(358, (0.01690498173388047, 0.062273210100076251)),
(359, (0.005517325824978776, 0.01324825744178606)),
(360, (0.010956206258908108, 0.03744072755287365)),
(361, (0.009844896554997677, 0.028544554681336154)),
(362, (0.006916268628942591, 0.012416241273944424)),
(363, (0.0074000152060141895, 0.0099201927704199516)),
(364, (0.010237123509380056, 0.02156841912020243)),
(365, (0.008445953751033863, 0.038784753670156304)),
(366, (0.007962207173962264, 0.03129660815958157)),
(367, (0.016578125938561825, 0.038272743720715284)),
(368, (0.008642067228225052, 0.03129660815958157)),
(369, (0.013910982648761654, 0.092353794630042166)),
(370, (0.004379867657269883, 0.06406524492380601)),
(371, (0.0031901125623100025, 0.03161661437798219)),
(372, (0.00026148463625491836, 0.03340864920102572)),
(373, (0.0015296851220912725, 0.03148861189062194)),
(374, (0.0015558335857167643, 0.039168761132237044)),
(375, (0.0007713796769520092, 0.02092840668340117)),
(376, (0.0035431168212541435, 0.013824268634907193)),
(377, (0.0005752661997608205, 0.017408338280994244)),
(378, (0.0013466458767128296, 0.02508847752260935)),
(379, (0.0013466458767128296, 0.02508847752260935)),
(380, (0.005268915420536604, 0.14931490150573376)),
(381, (0.005268915420536604, 0.14931490150573376)),
(382, (0.00037915272256963163, 0.0241924701110875)),
(383, (5.229692725098367e-05, 0.014848288533789212)),
(384, (2.6148463625491837e-05, 0.0005760111931211331)),
(385, (0.0002484104044421724, 0.021056409170176142)),
(386, (0.00041837541800786933, 0.005632109443851079)),
(387, (0.0004706723452588531, 0.016128313407391726)),
(388, (0.003686933371194349, 0.08992174737057687)),
(390, (0.0005621919679480744, 0.0373127250655134)),
(391, (0.0022879905672305353, 0.01888036688563714)),
(392, (0.0007060085178882795, 0.03705672009079289))
```

(392, (0.0007060085178882795, 0.03703672009092893)),
(393, (9.151962268922142e-05, 0.014144274853307827)),
(394, (0.008733586850914275, 0.017280335793633993)),
(395, (7.84453908764755e-05, 0.035328686511429494)),
(396, (0.011806031326909563, 0.07763350858399273)),
(397, (0.011806031326909563, 0.07763350858399273)),
(398, (0.004759020379839514, 0.025664498715730486)),
(399, (0.000797528140577501, 0.020928406668340117)),
(400, (0.0007583054451392631, 0.013696266147546942)),
(401, (0.0009674931541431978, 0.018688363154596762)),
(402, (9.151962268922142e-05, 0.0005760111931211331)),
(403, (5.229692725098367e-05, 0.004160080839208184)),
(404, (0.00035300425894413975, 0.015104293508509713)),
(405, (0.0006406373588245499, 0.0249604850352491)),
(406, (0.000313781563505902, 0.015360298483230216)),
(407, (0.0005229692725098367, 0.027328531051413755)),
(408, (0.0011897550949598785, 0.008640167896816998)),
(409, (0.0005621919679480745, 0.04345684445880548)),
(410, (0.0016473532084059857, 0.023296452699565824)),
(411, (0.00045759811344610714, 0.024064467623727336)),
(412, (0.0004183754180078694, 0.023040447724845325)),
(413, (0.0010328643132069275, 0.03462467283094811)),
(414, (0.002967850621493323, 0.048128935247454686)),
(415, (0.0005621919679480745, 0.011584225106102788)),
(416, (0.00019611347719118876, 0.028800559656056657)),
(417, (0.0002222619408166806, 0.014144274853307824)),
(418, (0.00033993002713139387, 0.01984038554083903)),
(419, (0.000313781563505902, 0.013632264903866816)),
(420, (0.0006144888951990581, 0.01171222759346304)),
(421, (0.000156890781752951, 0.016384318382112223)),
(422, (0.00028763309988041023, 0.027008524833013133)),
(423, (0.0010720870086451653, 0.02688052234565288)),
(424, (0.0008890477632667224, 0.014592283559068706)),
(425, (0.000627563127011804, 0.0405127872495197)),
(426, (0.0013727943403383216, 0.05126499618778085)),
(427, (0.001268200485836354, 0.047040914104892535)),
(428, (0.0013597201085255753, 0.06854533198141481)),
(429, (0.00030070733169315617, 0.02438447384212797)),
(430, (0.0003399300271313938, 0.0197763842997158903)),
(431, (0.00013074231812745918, 0.011904231324503417)),
(432, (0.0004968208088843448, 0.023680460161646582)),
(433, (0.00013074231812745918, 0.011648226349782912)),
(434, (0.0004837465770715989, 0.028160547219255396)),
(435, (0.003294706416811971, 0.033664654175746225)),
(436, (0.001268200485836354, 0.07372943271950505)),
(437, (0.001712724367469715, 0.0726414115769429)),
(438, (0.00010459385450196735, 0.015552302214270593)),
(439, (0.0008236766042029928, 0.02624050990885162)),
(440, (0.0015035366584657806, 0.041920814610482464)),
(441, (0.0005098950406970909, 0.03315264422630522)),
(442, (0.0008498250678284847, 0.03641670765399163)),
(443, (0.00010459385450196735, 0.010816210181941276)),
(444, (0.00045759811344610714, 0.03328064671366547)),
(445, (0.0002091877090039347, 0.011648226349782912)),
(446, (0.00028763309988041023, 0.008960174115217625)),
(447, (0.0013204974130873377, 0.03219262557110333)),
(448, (5.229692725098367e-05, 0.013056253710745683)),
(449, (0.0008236766042029928, 0.03244863054582383)),
(450, (0.00010459385450196735, 0.011904231324503418)),
(451, (7.84453908764755e-05, 0.013632264903866816)),
(452, (0.000156890781752951, 0.027072526076693255)),
(453, (0.00028763309988041023, 0.028544554681336154)),
(454, (0.003556191053066889, 0.05113699370042059)),
(455, (0.000156890781752951, 0.021120410414441546)),
(456, (7.84453908764755e-05, 0.039552768594317805)),
(457, (7.84453908764755e-05, 0.011904231324503417)),
(458, (0.0013727943403383214, 0.0748145551057473)),
(459, (0.0009413446905177062, 0.04832093897849505)),
(460, (0.00018303924537844285, 0.020352395490280034)),
(461, (0.0081975433346591691, 0.13561863535818677)),
(462, (0.002052654394601109, 0.06188920263868176)),
(463, (0.000313781563505902, 0.012736247492345053)),
(464, (0.0009343429418260743, 0.047040914104892535)),

```
(464, (0.009034294182607439, 0.04704091410489253))),
(465, (0.0034254487349394303, 0.009728189039379139)),
(466, (0.004170679948265948, 0.09235379463042168)),
(467, (0.0020526543946011093, 0.08761770259809236)),
(468, (0.00010459385450196735, 0.004928095763369694)),
(469, (0.0004445238816333612, 0.02310444896852545)),
(470, (0.00019611347719118878, 0.011456222618742535)),
(471, (0.0019611347719118875, 0.05542507702698902)),
(472, (0.00010459385450196735, 0.007232140535854227)),
(473, (0.0012812747176490999, 0.07052937053549875)),
(474, (0.0012028293267726247, 0.013824268634907193)),
(476, (0.009243481891611364, 0.08684968767393082)),
(477, (0.00614488895199058, 0.005760119931211331)),
(478, (0.00020918770900393467, 0.026112507421491366)),
(479, (0.00037915272256963163, 0.02105640917076142)),
(480, (0.00041837541800786933, 0.02259243901908444)),
(481, (0.00018303924537844285, 0.008128157947375989)),
(482, (0.004667500757150292, 0.04940896012105719)),
(483, (0.00621026011105431, 0.03564869272983012)),
(484, (0.00570036507035722, 0.07494545634942743)),
(485, (0.0005491177361353286, 0.01984038554083903)),
(486, (0.005543474288604269, 0.1681952683913708)),
(487, (0.0007452312133265173, 0.02585650244677086)),
(488, (0.00030070733169315617, 0.013376259929146312)),
(489, (0.0015427593539040182, 0.051008991213060344)),
(490, (0.006981639788006319, 0.028608555925016273)),
(491, (0.001412017035776559, 0.031040603184861057)),
(492, (7.84453908764755e-05, 0.0037760733771274274)),
(493, (0.00026148463625491836, 0.009344181577298381)),
(494, (0.002222619408166806, 0.07225740411486213)),
(495, (0.00139

*** WARNING: max output size exceeded, skipping output. ***

 0.09209778965570116)),
(579, (0.0028894052306168475, 0.08390563046464505)),
(580, (0.004824391538903245, 0.124546420201525)),
(581, (0.0012159035585853702, 0.03827274372071528)),
(582, (0.0005883404315735662, 0.03846474745175567)),
(583, (0.004144531484640456, 0.06323322875596439)),
(584, (0.004641352293524801, 0.013824268634907193)),
(585, (0.0035954137485051272, 0.017856346986755124)),
(586, (0.007321569815137714, 0.05740911558107293)),
(587, (0.004589055366273817, 0.09280180333618256)),
(588, (0.006184111647428819, 0.03296064049526484)),
(589, (0.0010328643132069275, 0.018816365641957013)),
(590, (0.004013789166512996, 0.044800870576088135)),
(591, (0.000156890781752951, 0.03596869894823076)),
(592, (0.008367508360157387, 0.2001958902314338)),
(593, (0.0037000076030070947, 0.08563366404400845)),
(594, (0.0018957636128481578, 0.05740911558107293)),
(595, (0.000470672345258853, 0.022592439019084445)),
(596, (0.0019611347719118875, 0.06643329093997069)),
(598, (0.0006144888951990581, 0.039872774812718434)),
(600, (0.00019611347719118876, 0.012928251223385432)),
(601, (0.00041837541800786933, 0.021760422851242807)),
(602, (0.0016996501356569694, 0.053377037229224995)),
(603, (0.0006798600542627877, 0.020480397977640286)),
(604, (0.0023795101899197566, 0.02803254473189514)),
(605, (7.84453908764755e-05, 0.011136216400341906)),
(606, (0.0016081305129677479, 0.0212484129018018)),
(608, (0.0006929342860755336, 0.04185681336680234)),
(609, (0.0004053011861951234, 0.01894436812931727)),
(610, (0.0006667858224500417, 0.04153680714840171)),
(611, (9.151962268922142e-05, 0.005568108220170953)),
(612, (0.0034515971985649217, 0.066305288845261044)),
(614, (6.537115906372959e-05, 0.0029440572092857913)),
(615, (6.537115906372959e-05, 0.016192314651071852)),
(616, (0.007020862483444558, 0.026944523589333007)),
(617, (0.0004968208088843448, 0.03968077108167806)),
(618, (0.0004968208088843448, 0.03968077108167806)),
(619, (0.003634636443943365, 0.04716891659225278)),
(620, (9.151962268922142e-05, 0.009344181577298381)),
```

(620, (9.151962268922142e-05, 0.009344181577298381)),
(621, (0.0004445238816333612, 0.02835255095029577)),
(622, (2.6148463625491837e-05, 0.005632109443851079)),
(623, (0.010655498927387924, 0.03769673252759416)),
(624, (0.00028763309988041023, 0.017472339524674373)),
(626, (0.0014250912675893052, 0.017344370037314115)),
(627, (0.001045938545019674, 0.017088332062593616)),
(628, (0.00227491633541779, 0.09305780831090306)),
(629, (0.000797528140577501, 0.00640012436801259)),
(630, (0.00019611347719118878, 0.007424144266894604)),
(631, (7.84453908764755e-05, 0.04390485316456637)),
(632, (0.00013074231812745918, 0.005696110687531205)),
(633, (0.0024579555807962322, 0.04294483450936447)),
(634, (0.0016996501356569692, 0.06028917154667859)),
(635, (0.00018303924537844285, 0.014912289777469332)),
(636, (0.000156890781752951, 0.004096079595528057)),
(637, (0.00013074231812745918, 0.007936154216335612)),
(638, (0.003569265284879635, 0.12761847989817104)),
(639, (0.000156890781752951, 0.004800093276009443)),
(640, (0.0013597201085255753, 0.05331303598554488)),
(641, (0.00014381654994020506, 0.011008213912981655)),
(642, (0.01181910555872231, 0.18566760791604522)),
(643, (0.0047982430752777526, 0.13286658187994135)),
(644, (0.00030070733169315606, 0.004672090788649191905)),
(645, (0.0013466458767128296, 0.01715233330627374)),
(646, (0.0010590127768324194, 0.029440572092857911)),
(647, (0.0002745588680676643, 0.0063361231243324635)),
(648, (0.000483746577071599, 0.022528437775404322)),
(649, (3.922269543823775e-05, 0.00422408208288831)),
(650, (0.0002484104044421724, 0.014784287290109081)),
(651, (0.0016342789765932395, 0.05376104469130576)),
(653, (0.0015427593539040182, 0.05728111309371268)),
(654, (0.0016081305129677476, 0.05177700613722185)),
(656, (0.002366435958107011, 0.08409763419568543)),
(657, (0.0016473532084059857, 0.027648537269814387)),
(658, (0.00011766808631471327, 0.015680304701630843)),
(659, (0.0002745588680676643, 0.0215044178765223)),
(660, (0.0033993002713139384, 0.01907237061667752)),
(661, (0.00040530118619512345, 0.023872463892686956)),
(662, (0.0003791527225696316, 0.0229764464811652)),
(663, (0.0005229692725098367, 0.0200963905155559528)),
(664, (0.00018303924537844285, 0.012288238786584173)),
(665, (0.00011766808631471327, 0.040704790980560075)),
(666, (6.537115906372959e-05, 0.010432202719860522)),
(667, (0.0009805673859559438, 0.033984660394146854)),
(668, (0.00043144964982061527, 0.023324514558857702)),
(669, (0.0026017721307364378, 0.07244940784590251)),
(670, (0.0003791527225696316, 0.019776384297158903)),
(671, (0.0008759735314539765, 0.01568030470163085)),
(672, (0.00011766808631471327, 0.015552302214270597)),
(673, (0.00030070733169315617, 0.026560516127252225)),
(674, (0.00013074231812745918, 0.025792501203090738)),
(675, (3.922269543823775e-05, 0.012608245004984801)),
(676, (0.00014381654994020512, 0.009728189039379137)),
(677, (0.0013989428039638133, 0.018240354448835882)),
(678, (0.003752304530258078, 0.07097737924125962)),
(679, (0.0010590127768324194, 0.047040914104892535)),
(680, (0.003948418007449267, 0.09376182199138444)),
(681, (0.00013074231812745918, 0.009152177846258002)),
(682, (0.00013074231812745918, 0.009152177846258002)),
(683, (0.0003399300271339387, 0.019840385540839025)),
(684, (0.0003268557953186479494, 0.01337625999291461313)),
(685, (0.0016212047447804938, 0.0005120099494410072)),
(686, (0.0012159035585853702, 0.05292902852346412)),
(687, (0.0005360435043225827, 0.01836835693619613)),
(688, (0.0008367508360157388, 0.055297074539628774)),
(689, (0.000954418922330452, 0.0496649650957777)),
(690, (0.0005229692725098367, 0.01760034201203462)),
(691, (0.00045759811344610714, 0.04390485316456637)),
(692, (0.005582696984042505, 0.09721789915011124)),
(693, (0.000483746577070715989, 0.028352550950295777)),
(694, (0.00030070733169315617, 0.0244484750858808094)),
(695, (6.537115906372959e-05, 0.0333446479573456))

(695, (6.537115906372959e-05, 0.0333464079373456)),
(696, (0.0027848113761148804, 0.036096701435591101)),
(697, (2.6148463625491837e-05, 0.0028160547219255394)),
(698, (2.6148463625491837e-05, 0.0028160547219255394)),
(699, (2.6148463625491837e-05, 0.0028160547219255394)),
(700, (0.0052689154205366046, 0.07718549987823184)),
(701, (0.000313781563505902, 0.006656129342733093)),
(702, (0.0004183754180078694, 0.015552302214270593)),
(703, (0.010224049277567308, 0.12704246870504995)),
(704, (0.0006014146633863121, 0.010816210181941278)),
(705, (0.00400071493470025, 0.046144896669337078)),
(706, (0.00400071493470025, 0.046144896669337078)),
(707, (0.001751947062907953, 0.031936620596382824)),
(708, (0.0012812747176490999, 0.0424968258036036)),
(709, (0.0005752661997608205, 0.026624517370932372)),
(710, (0.00011768086314713262, 0.010944212669301527)),
(711, (0.0015035366584657806, 0.03142461064694181)),
(712, (0.004275273802767915, 0.06700930213309182)),
(713, (0.0015819820493422558, 0.03302464173894496)),
(714, (0.00030070733169315606, 0.0216964216075626783)),
(715, (0.0024187328853579946, 0.0634892337306849)),
(716, (0.00019611347719118878, 0.04185681336680234)),
(717, (0.0002091877090039347, 0.028928562143416908)),
(718, (0.00014381654994020512, 0.014912289777469332)),
(719, (0.0001699650135656969, 0.051521001162501336)),
(720, (0.00019611347719118876, 0.026368512396211187)),
(721, (0.008877403400854479, 0.14093073858363722)),
(722, (0.001451239731214797, 0.04268882953464398)),
(723, (0.0005491177361353286, 0.04096079595528058)),
(724, (6.537115906372959e-05, 0.0012160236299223921)),
(725, (0.0010197900813941816, 0.0056961106875312055)),
(726, (0.0009021219950794683, 0.011264218887702158)),
(727, (0.0012812747176490999, 0.013504262416506563)),
(730, (0.000156890781752951, 0.015360298483230216)),
(731, (0.00023533617262942654, 0.013952271122267447)),
(732, (0.0021572482491030763, 0.033216645469985345)),
(733, (6.537115906372959e-05, 0.005248101981770325)),
(734, (0.0013074231812745915, 0.04185681336680234)),
(735, (0.00040530118619512345, 0.02201642782596331)),
(736, (9.151962268922142e-05, 0.0069121343174535597)),
(737, (0.0007190827497010254, 0.0091521778462580002)),
(738, (0.00022226194081668063, 0.011648226349782912)),
(739, (0.0002222619408166806, 0.02105640917076142)),
(740, (7.84453908764755e-05, 0.013568263660186692)),
(741, (0.0003791527225696316, 0.022912445237485073)),
(742, (0.002876330998804102, 0.0800655558438375)),
(743, (0.0016473532084059857, 0.058305132992594697)),
(744, (0.00192191207647365, 0.021184411658121675)),
(745, (2.6148463625491837e-05, 0.004352084570248561)),
(746, (0.0009021219950794682, 0.020672401708680666)),
(747, (0.0005229692725098367, 0.02214443031332356)),
(748, (0.009217333427985872, 0.04716891659225278)),
(749, (0.002052654394601109, 0.07987355211279713)),
(750, (0.001725798599282461, 0.06508926482268804)),
(751, (0.001725798599282461, 0.06508926482268804)),
(752, (0.0003399300271313938, 0.016832327087873113)),
(753, (0.00035300425894413975, 0.017536340768354496)),
(754, (0.0008106023723902468, 0.025408493741009983)),
(755, (0.005334286579600334, 0.19488378700598344)),
(756, (0.005242766956911113, 0.1333785918293824)),
(757, (0.000483746577071599, 0.0170243308189349)),
(758, (0.000313781563505902, 0.0212484129018018)),
(759, (0.00018303924537844285, 0.00780815172897536)),
(760, (0.0002484104044421724, 0.03276863676422446)),
(762, (0.00010459385450196735, 0.03148861189062194)),
(763, (0.00037915272256963163, 0.015232295995869963)),
(764, (0.005739587765795457, 0.08908973120273524)),
(765, (0.0010067158495814356, 0.04083279346792032)),
(766, (0.0003399300271313938, 0.02752053478245414)),
(767, (0.0002484104044421724, 0.018816365641957013)),
(768, (0.004785168843465005, 0.08556966280032832)),
(769, (0.001477388194840289, 0.037504728796553775)),
(770, (0.0005752661997608205, 0.028928562143416908)),

```
(770,  (0.0005732001997008205, 0.020920002143410900)),
(771,  (0.00160813051296774476, 0.046016894200601052)),
(772,  (0.0008628992996412305, 0.02278444275012482)),
(773,  (0.00030070733169315606, 0.020992407927081295)),
(774,  (0.000810602372390247, 0.0499209700704982)),
(775,  (0.0003399300271313938, 0.019392376835078152)),
(776,  (0.003726156066632586, 0.07404943893790568)),
(777,  (0.0022618421036050434, 0.04396885440824649)),
(778,  (0.0011766808631471328, 0.062529215075483)),
(779,  (0.00040530118619512345, 0.029184567711813741)),
(780,  (0.00035300425894413975, 0.023808462649006837)),
(781,  (0.004170679948265948, 0.03974477232535818)),
(782,  (0.0003399300271313938, 0.029632575823898288)),
(783,  (0.0002745588680676643, 0.0209924079270813)),
(784,  (0.0009544189223330452, 0.063105226226860415)),
(785,  (0.0036084879803178734, 0.03609670143559101)),
(786,  (0.0036084879803178734, 0.03609670143559101)),
(787,  (0.0002484104044421724, 0.020224393002919783)),
(788,  (0.002065728626413855, 0.07001736058605772)),
(789,  (0.00406608609376398, 0.060353172790358726)),
(790,  (0.0002745588680676643, 0.026304511152531743)),
(791,  (0.0005491177361353286, 0.03507268153670899)),
(792,  (0.006262557038305295, 0.18291555443779983)),
(793,  (0.00040530118619512345, 0.03366465417574622)),
(794,  (0.0008890477632667224, 0.029696577067578424)),
(795,  (0.000313781563505902, 0.030336589504379678)),
(796,  (0.0051904700296601285, 0.12921851099017417)),
(797,  (0.0009282704587049601, 0.05817713050523444)),
(798,  (0.00028763309988041023, 0.02080040419604092)),
(799,  (0.00041837541800786694, 0.029376570849177785)),
(800,  (0.0005360435043225827, 0.0430088357530446)),
(801,  (0.0005360435043225827, 0.03526468526774937)),
(802,  (0.002549475203485454, 0.07942554340703624)),
(803,  (0.000483746577071599, 0.04313683824040486)),
(804,  (0.0006144888951990581, 0.03475267531830837)),
(805,  (0.002235693639979552, 0.04147280590472158)),
(806,  (0.0005098950406970908, 0.019968388028199283)),
(807,  (0.0019349863082863956, 0.06342523248700477)),
(808,  (0.0002484104044421724, 0.02156841912020243)),
(809,  (0.004170679948265948, 0.03648070889767764)),
(810,  (0.0010067158495814356, 0.033728655419426344)),
(811,  (0.00028763309988041023, 0.02067240170868666)),
(812,  (0.004013789166512996, 0.04096079595528057)),
(813,  (0.000156890781752951, 0.007616147997934981)),
(814,  (0.0023141390308560276, 0.04531288052552913)),
(815,  (0.0017780955265334448, 0.0267525198582927)),
(816,  (6.537115906372959e-05, 0.008512165409456745)),
(817,  (0.0008367508360157389, 0.018240354448835882)),
(818,  (0.0009805673859959944, 0.03020587701701942)),
(819,  (0.003647710675756110, 0.0249604850352491)),
(820,  (0.0003268557953186479, 0.014016272365947574)),
(821,  (0.002000357467350125, 0.062977223781243877)),
(822,  (0.0020134316991628712, 0.04569688798760989)),
(824,  (0.00026148463625491836, 0.004800093276009442)),
(825,  (0.0003268557953186479, 0.008064156703695863)),
(826,  (0.000392226954382377757, 0.022528437775404316)),
(827,  (0.00041837541800786933, 0.031104604428541186)),
(828,  (0.0002353361726294654, 0.00499209700704982)),
(829,  (0.0005491177361353286, 0.027136527320373378)),
(830,  (0.00013074231812745918, 0.01388826987858732)),
(831,  (0.00026148463625491836, 0.07968154838175674)),
(832,  (0.0003399300271313938, 0.014272277340668077)),
(833,  (0.003600775775875701, 0.019136371860357642)),
(834,  (0.0005752661997608202, 0.06899334068717572)),
(835,  (0.0017780955265334446, 0.06195320388236186)),
(836,  (0.0002091877090039347, 0.015680304470163085)),
(837,  (0.007386940974201442, 0.013696266147546942)),
(838,  (0.007386940974201442, 0.013696266147546942)),
(839,  (0.005582696984042507, 0.03820874247703515)),
(840,  (0.00040530118619512345, 0.09497784562130683)),
(841,  (0.004510609975397341, 0.07680149241615104)),
(842,  (0.0008498250678284847, 0.022272432800683813)),
(843,  (0.00037915272256963163, 0.014592283559068706)),
```

```
(843, (0.0003791527225696316?, 0.01439228535960876?)),
(844, (2.6148463625491837e-05, 0.0002560049747205036)),
(845, (0.00014381654994020512, 0.009664187795699011)),
(846, (0.0003660784907568857, 0.019648381809798655)),
(847, (0.008864329169041733, 0.07142538794702048)),
(848, (0.001451239731214797, 0.03123260691590144)),
(849, (7.84453908764755e-05, 0.0028800559656056656)),
(850, (0.0022095451763540597, 0.06726530710781232)),
(852, (5.229692725098367e-05, 0.015168294752189839)),
(853, (0.003451597198564922, 0.012288238786584173)),
(854, (0.0004183754180078694, 0.030976601941180933)),
(855, (0.0023402874944815195, 0.050880988725700085)),
(856, (0.0005752661997608205, 0.02873655841237653)),
(857, (9.151962268922142e-05, 0.015040292264829586)),
(858, (0.014969995425594075, 0.05651309816955117)),
(859, (0.013937131112387148, 0.072193402871182)),
(860, (0.013937131112387148, 0.072193402871182)),
(861, (0.0027848113761148804, 0.0026880522345652876)),
(862, (0.004392941889082629, 0.014528282315388582)),
(863, (0.009086591109858412, 0.07353742898846465)),
(864, (6.537115906372959e-05, 0.0022400435288044062)),
(865, (0.0027848113761148804, 0.09920192770419514)),
(866, (0.008916626096292714, 0.052673023548743605)),
(867, (0.0008890477632667224, 0.03916876113223705)),
(868, (0.0013727943403383214, 0.02361645891796646)),
(871, (2.6148463625491837e-05, 0.0053121032254450449)),
(872, (0.0002876330998804101, 0.01280024873602518)),
(874, (0.0024318071171707404, 0.06643329093997069)),
(875, (0.0015689078175295103, 0.05600108822011016)),
(876, (0.00040530118619512345, 0.00780815172897536)),
(877, (7.84453908764755e-05, 0.00748814551057473)),
(878, (0.0008628992996412305, 0.009024175358897752)),
(880, (6.537115906372959e-05, 0.0008960174115217626)),
(881, (0.0007713796769520092, 0.03315264422630521)),
(882, (0.0019742090037246332, 0.06899334068717572)),
(884, (0.0001699650135656969, 0.011328220131382283)),
(885, (0.0013204974130873377, 0.0160643121637116)),
(886, (0.0007190827497010254, 0.1605791203934359)),
(887, (0.000313781563505902, 0.02950457333653804)),
(889, (0.0018303924537844283, 0.0531210322545450449)),
(891, (3.922269543823775e-05, 0.0008320161678416366)),
(892, (0.001464313963027 5428, 0.06176120015132148)),
(893, (9.151962268922142e-05, 0.015488300970590467)),
(894, (0.0004314496498206153, 0.04512087679448876)),
(895, (0.0007190827497010254, 0.0044800870576088125)),
(896, (0.0005752661997608205, 0.026240509908851618)),
(897, (0.000313781563505902, 0.0832656180278438)),
(898, (0.0006798600542627877, 0.009408182820978507)),
(899, (0.0006537115906372958, 0.04640090166809127)),
(900, (0.0008236766042029928, 0.055297074539628774)),
(901, (0.0008236766042029928, 0.055297074539628774)),
(902, (0.00040530118619512345, 0.019648381809 79865)),
(904, (0.00019611347719118878, 0.0046720907886491905)),
(905, (0.0013335716449000836, 0.006912134317453597)),
(906, (0.0007583054451392631, 0.016576322113152606)),
(908, (0.0022487678717922977, 0.070657373022859)),
(909, (7.84453908764755e-05, 0.0232324514558857 02)),
(910, (0.002758662912489388, 0.08768170384177248)),
(911, (0.00285018253517861, 0.06380923994908552)),
(912, (0.006837823238066114, 0.14720286046428957)),
(913, (0.00018303924537844285, 0.020800404196040 92)),
(914, (0.000313781563505902, 0.01817635320515576)),
(915, (5.229692725098367e-05, 0.013184256198105936)),
(916, (0.0016604274402187316, 0.024960485035249103)),
(918, (3.922269543823775e-05, 0.0005760111931211331)),
(919, (0.000248104044421724, 0.00825616043473624)),
(920, (7.84453908764755e-05, 0.0298885807786188794)),
(921, (3.922269543823775e-05, 0.001984038554083903)),
(922, (6.537115906372959e-05, 0.003840074620807554)),
(923, (0.0013858685721510671, 0.020416396733960163)),
(924, (0.0003660784907568857, 0.017856346986755124)),
(925, (7.84453908764755e-05, 0.0054401057128810702)),
(926, (0.001019790081394181 6, 0.01568030470163084 6))
```

(926, (0.0010197900813941816, 0.015680304701630846)),
(927, (0.0018826893810354123, 0.05408105090970638)),
(928, (0.0017519470629079532, 0.061377192689240745)),
(929, (2.6148463625491837e-05, 0.010944212669301527)),
(930, (0.00014381654994020512, 0.025344492497329858)),
(931, (0.000313781563505902, 0.03168061562166232)),
(932, (0.00010459385450196735, 0.010816210181941278)),
(933, (0.0008498250678284847, 0.06444925238588678)),
(934, (0.000627563127011804, 0.03424066536886736)),
(935, (0.00018303924537844285, 0.02278444275012482)),
(936, (0.0003791527225696316, 0.03347265044470585)),
(937, (7.84453908764755e-05, 0.01920037310403777)),
(938, (0.000470672345258853, 0.031104604428541193)),
(939, (0.0013989428039638133, 0.08729769637969173)),
(940, (0.0007321569815137714, 0.011520223862422662)),
(941, (0.000797528140577501, 0.04832093897849504)),
(942, (0.0071908274970102545, 0.09875391899843423)),
(943, (5.229692725098367e-05, 0.006592128099052968)),
(944, (2.6148463625491837e-05, 0.009024175358897752)),
(945, (0.0005883404315735664, 0.008576166665313687)),
(946, (5.229692725098367e-05, 0.012928251223385432)),
(947, (5.229692725098367e-05, 0.012928251223385432)),
(948, (5.229692725098367e-05, 0.012928251223385432)),
(949, (5.229692725098367e-05, 0.012928251223385432)),
(950, (5.229692725098367e-05, 0.012928251223385432)),
(951, (5.229692725098367e-05, 0.012928251223385432)),
(952, (5.229692725098367e-05, 0.012928251223385432)),
(953, (0.002222619408166806, 0.04134480341736133)),
(954, (0.00013074231812745918, 0.019584380566118526)),
(955, (0.005569622752229762, 0.18221154075731844)),
(956, (5.229692725098367e-05, 0.013248257441786062)),
(957, (0.00026148463625491836, 0.040320783518479314)),
(958, (0.0021572482491030763, 0.06220920885708237)),
(959, (0.0012812747176490999, 0.01657632211315261)),
(960, (0.0009805673859559438, 0.03705672009079289)),
(961, (0.0006406373588245499, 0.039360764863277435)),
(962, (0.0017650212947206987, 0.07904153594495548)),
(963, (0.001581982049342256, 0.04364884818984586)),
(964, (9.151962268922142e-05, 0.02822454846293552)),
(965, (0.0020134316991628712, 0.07443344639998642)),
(966, (0.00013074231812745918, 0.016128313407391726)),
(967, (0.00037915272256963163, 0.014464281071708456)),
(968, (0.00044452388163336126, 0.05107299245674047)),
(969, (0.0002222619408166806, 0.03865675118279604)),
(970, (0.000156890781752951, 0.03404866163782698)),
(971, (5.229692725098367e-05, 0.01964838180979865)),
(972, (5.229692725098367e-05, 0.01964838180979865)),
(973, (5.229692725098367e-05, 0.01964838180979865)),
(974, (0.00010459385450196735, 0.027392532295093884)),
(975, (0.0003660784907568857, 0.03750472879655377)),
(976, (0.004693649220775784, 0.13600264282026753)),
(977, (0.0004968208088843448, 0.04652890415545153)),
(978, (0.00014381654994020512, 0.017856346986755128)),
(979, (0.002301064799043282, 0.018240354448835882)),
(980, (0.0005229692725098367, 0.05491306707754802)),
(981, (0.0016473532084059857, 0.051905008624582104)),
(982, (5.229692725098367e-05, 0.01344026117282644)),
(983, (0.0001699650135656969, 0.03628870516663139)),
(984, (0.0015558335857167643, 0.051585002406181475)),
(985, (0.000313781563505902, 0.026496514883572124)),
(986, (0.0018696151492226662, 0.07475345261838705)),
(987, (0.00045759811344610714, 0.04064078973687995)),
(988, (0.0032424094895609876, 0.057217111850032545)),
(989, (0.001738872831095207, 0.03616070267927113)),
(990, (0.003608487980317873, 0.10464203341700586)),
(991, (0.000810602372390247, 0.0195203793224384)),
(992, (0.0034908198940031597, 0.1399067186847552)),
(993, (0.0007060085178882794, 0.04454486560136762)),
(994, (0.0001699650135656969, 0.009792190283059263)),
(995, (0.00045759811344610714, 0.043776850067720611)),
(997, (0.0074523121133265172, 0.026176508665171495)),
(998, (0.007478460596890665, 0.2085160519098502)),
(999, (0.0003533617262942654, 0.009984194014099964)),

```
(999, (0.0002353361726294265, 0.009841401409904))),
(1000, (0.0006537115906372958, 0.023168450212205576)),
(1001, (0.00010459385450196735, 0.007872152972655484)),
(1002, (0.00018303924537844285, 0.01184023008082329)),
(1003, (0.00016996501356569694, 0.013504262416506563)),
(1004, (0.00010459385450196735, 0.008384162922096492)),
(1005, (0.0012028293267726247, 0.034112662881507105)),
(1007, (2.6148463625491837e-05, 0.0020480397977640286)),
(1008, (0.00436793425457136, 0.09420983069714534)),
(1009, (0.0012682004858363541, 0.021120410414441546)),
(1010, (0.00044452388163336126, 0.028992563387097034)),
(1011, (0.0005883404315735662, 0.021376415389162053)),
(1012, (0.000967493154143198, 0.035776695217190374)),
(1014, (0.005974923938424884, 0.049408960121057194)),
(1015, (0.00614488895199058, 0.005760119931211331)),
(1016, (0.0002353361726294265, 0.020224393002919783)),
(1017, (5.229692725098367e-05, 0.011264218887702158)),
(1018, (0.01460391693483719, 0.03001658328597905)),
(1019, (0.00035300425894413975, 0.018752364398276888)),
(1020, (0.00011766808631471326, 0.0016000310920031475)),
(1021, (0.00037915272256963163, 0.031680615621662314)),
(1022, (0.018656928796788423, 0.5470826309777161)),
(1023, (2.6148463625491837e-05, 0.031744616865342454)),
(1024, (0.00841980528740837, 0.014912289777469332)),
(1025, (0.00069293428607553336, 0.030400590748059803)),
(1026, (0.00069293428607553336, 0.030400590748059803)),
(1027, (0.0005883404315735662, 0.00576011193121133)),
(1028, (0.0005883404315735662, 0.00576011193121133)),
(1029, (0.0005883404315735662, 0.006720130586413218)),
(1030, (0.0005883404315735662, 0.006720130586413218)),
(1031, (0.0005621919679480744, 0.014272277340668077)),
(1032, (7.84453908764755e-05, 0.007360143023214478)),
(1033, (0.00010459385450196733, 0.0044800870576088125)),
(1034, (0.0016735016720314773, 0.05075298623833984)),
(1036, (9.151962268922142e-05, 0.004416085813928688)),
(1037, (9.151962268922142e-05, 0.004416085813928688)),
(1038, (9.151962268922142e-05, 0.004416085813928688)),
(1039, (0.0016473532084059859, 0.020864405439721043)),
(1042, (0.00030070733169315617, 0.007296141779534353)),
(1043, (0.0013466458767128296, 0.05996916532827796)),
(1044, (0.004928985393405211, 0.019712383053478777)),
(1045, (0.004928985393405211, 0.019712383053478777)),
(1046, (0.008171394882966199, 0.016832327087873113)),
(1047, (0.008092949492089725, 0.016832327087873113)),
(1048, (0.008092949492089725, 0.016832327087873113)),
(1049, (0.008092949492089725, 0.016832327087873113)),
(1050, (0.001255126254023608, 0.041664809635761954)),
(1051, (0.0011374581677088948, 0.02028839424659991)),
(1052, (0.0014250912675893052, 0.011968232568183542)),
(1053, (0.0006667858224500418, 0.01107221515666178)),
(1054, (0.0017650212947206989, 0.013888269878558732)),
(1055, (0.00014381654994020512, 0.0046720907886491905)),
(1056, (0.0007060085178882795, 0.062721218806552338)),
(1057, (0.0011766808631471328, 0.025984504934131115)),
(1058, (0.0009805673859559438, 0.04134480341736133)),
(1059, (0.0002353361726294265, 0.013504262416506565)),
(1060, (0.0008498250678284847, 0.013696266147546942)),
(1061, (0.0009413446905177062, 0.013632264903866816)),
(1062, (0.000954418922330452, 0.019904386784519154)),
(1063, (0.0002222619408166806, 0.01593630967635135)),
(1064, (0.00899507148716919, 0.07897753470127535)),
...]
```

In [0]:

```python
#Finding the actual prediction by comparing the hamProb and spamProb
prediction = combinedProb.map(lambda x:(x[0], 1 if x[1][0] > x[1][1] else 0)).toDF(["id"
,"predictions"])
finalPrediction = prediction.join(testDataFrame, prediction.id == testDataFrame.id)

finalPrediction = finalPrediction.select(col("predictions").cast("double"),col("label").c
ast("double"))
```

```
finalPrediction.display()
```

predictions	label
0.0	0.0
0.0	1.0
0.0	0.0
0.0	0.0
0.0	1.0
0.0	1.0
0.0	1.0
0.0	0.0
0.0	0.0
0.0	0.0

In [0]:

```
from pyspark.mllib.evaluation import BinaryClassificationMetrics, MulticlassMetrics

results = finalPrediction.rdd.map(lambda x : x)
results.take(10)
```

```
Out[97]: [Row(predictions=0.0, label=0.0),
 Row(predictions=0.0, label=1.0),
 Row(predictions=0.0, label=0.0),
 Row(predictions=0.0, label=0.0),
 Row(predictions=0.0, label=1.0),
 Row(predictions=0.0, label=1.0),
 Row(predictions=0.0, label=1.0),
 Row(predictions=0.0, label=0.0),
 Row(predictions=0.0, label=0.0),
 Row(predictions=0.0, label=0.0)]
```

In [0]:

```
metrics = MulticlassMetrics(results)
print("Confusion matrix:\n",metrics.confusionMatrix())
print("Model Accuracy = %s" % metrics.accuracy)
```

```
/databricks/spark/python/pyspark/sql/context.py:165: FutureWarning: Deprecated in 3.0.0.
Use SparkSession.builder.getOrCreate() instead.
  warnings.warn(
```

```
Confusion matrix:
 DenseMatrix([[ 210.,     5.],
             [1376.,     0.]])
Model Accuracy = 0.13199245757385292
```

In [0]: