

Natural language processing Quiz 1

Shane Kao

Saturday, February 21, 2015

```
setwd("C:/Users/asus/Downloads/final/en_US")
twitter=readLines("en_US.twitter.txt")
blogs=readLines("en_US.blogs.txt")
news=readLines("en_US.news.txt")
data=c(twitter,blogs,news)
```

Question 1

The guy in front of me just bought a pound of bacon, a bouquet, and a case of

```
x="a case of"
data_v1=data[grepl(x,data)]
n=length(data_v1)
pred=function(y){
  data_v2=data_v1[grepl(paste(x,y,sep=" "),data_v1)]
  length(data_v2)/n
}
candidate=c("soda","pretzels","cheese","beer")
prob=apply(cbind(candidate),1,pred)
output=data.frame("candidate"=candidate,"probability"=prob)
output[order(output$probability,decreasing=TRUE),]
```

```
## candidate probability
## 4 beer 0.037567084
## 1 soda 0.001788909
## 2 pretzels 0.000000000
## 3 cheese 0.000000000
```

Question 2

You're the reason why I smile everyday. Can you follow me please? It would mean the

```
## candidate probability
## 1 world 0.153754470
## 2 most 0.014898689
## 3 universe 0.001787843
## 4 best 0.001787843
```

Question 3

Hey sunshine, can you follow me and make me the

```
## candidate probability
## 1 happiest 0.3802817
## 2 bluest 0.0000000
## 3 saddest 0.0000000
## 4 smelliest 0.0000000
```

Question 4

Very early observations on the Bills game: Offense still struggling but the

```
## candidate probability
## 2 crowd 0.0008476894
## 4 players 0.0003281378
## 1 defense 0.0001914137
## 3 referees 0.0000000000
```

Question 5

Go on a romantic date at the

```
## candidate probability
## 4 beach 0.0021965712
## 1 mall 0.0015489765
## 2 grocery 0.0010851587
## 3 movies 0.0005207012
```

Question 6

Well I'm pretty sure my granny has some old bagpipes in her garage I'll dust them off and be on my

```
## candidate probability
## 3 way 0.136363636
## 1 phone 0.007575758
## 2 motorcycle 0.000000000
## 4 horse 0.000000000
```

Question 7

Ohhhhh #PointBreak is on tomorrow. Love that film and haven't seen it in quite some

```
## candidate probability
## 3 time 0.870370370
## 4 years 0.004629630
## 1 thing 0.002314815
## 2 weeks 0.000000000
```

Question 8

After the ice bucket challenge Louis will push his long wet hair out of his eyes with his little

```
## candidate probability
## 3 fingers 0.0021356113
## 1 eyes 0.0016017085
## 2 ears 0.0005339028
## 4 toes 0.0000000000
```

Question 9

Be grateful for the good times and keep the faith during the

```
## candidate probability
## 2 hard 0.0001815102
## 4 bad 0.0001815102
## 1 sad 0.0000000000
## 3 worse 0.0000000000
```

Question 10

If this isn't the cutest thing you've ever seen, then you must be

```
## candidate probability
## 3 insane 0.0007470651
## 1 asleep 0.0002134472
## 2 insensitive 0.0000000000
## 4 callous 0.0000000000
```