q-teda_order

Michael Pearson

7/02/2018

R Markdown

Do the combi thing for samples

```
tri_trigrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/tri_gram
tri_trigrams <- data.table(tri_trigrams)
combi_tri_s_s <- unite(tri_trigrams, bigrams, c("word1", "word2"), sep = " ")
rm(tri_trigrams)
write.csv(combi_tri_s_s,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_tri
rm(combi_tri_s_s)
tri_trigrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/tri_gram
tri_trigrams <- data.table(tri_trigrams)
combi_tri_ns_s <- unite(tri_trigrams, bigrams, c("word1", "word2"), sep = " ")
rm(tri_trigrams)
write.csv(combi_tri_ns_s,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_ti
rm(combi_tri_ns_s)
tri_trigrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/tri_gram
tri_trigrams <- data.table(tri_trigrams)
combi_tri_s_ns <- unite(tri_trigrams, bigrams, c("word1", "word2"), sep = " ")
rm(tri_trigrams)
write.csv(combi_tri_s_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_ti
rm(combi_tri_s_ns)
```

now for no stemming and no stopwords

```
tri_trigrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/tri_gram
tri_trigrams <- data.table(tri_trigrams)
combi_tri_ns_ns <- unite(tri_trigrams, bigrams, c("word1", "word2"), sep = " ")
rm(tri_trigrams)
write.csv(combi_tri_ns_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_
rm(combi_tri_ns_ns)
```

Now for the test batch

```
tri_trigrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20test/tri_gram_
tri_trigrams <- data.table(tri_trigrams)
combi_tri_ns_ns <- unite(tri_trigrams, bigrams, c("word1", "word2"), sep = " ")
rm(tri_trigrams)
write.csv(combi_tri_ns_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20test/combi_tri
rm(combi_tri_ns_ns)
```

quadgrams

```
quad_quadgrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/quad_g
quad_quadgrams <- data.table(quad_quadgrams)
combi_quad_s_s <- unite(quad_quadgrams, trigrams, c("word1", "word2", "word3"), sep = " ")
rm(quad_quadgrams)
```

```r
write.csv(combi_quad_s_s,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_q
rm(combi_quad_s_s)
quad_quadgrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/quad_
quad_quadgrams <- data.table(quad_quadgrams)
combi_quad_ns_s <- unite(quad_quadgrams, trigrams, c("word1", "word2", "word3"), sep = " ")
rm(quad_quadgrams)
write.csv(combi_quad_ns_s,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_
rm(combi_quad_ns_s)
quad_quadgrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/quad_
quad_quadgrams <- data.table(quad_quadgrams)
combi_quad_s_ns <- unite(quad_quadgrams, trigrams, c("word1", "word2", "word3"), sep = " ")
rm(quad_quadgrams)
write.csv(combi_quad_s_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_
rm(combi_quad_s_ns)
```

Now for ns_ns

```r
quad_quadgrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/quad_
quad_quadgrams <- data.table(quad_quadgrams)
combi_quad_ns_ns <- unite(quad_quadgrams, trigrams, c("word1", "word2", "word3"), sep = " ")
rm(quad_quadgrams)
write.csv(combi_quad_ns_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_
rm(combi_quad_ns_ns)
```

Now for test

```r
quad_quadgrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20test/quad_gra
quad_quadgrams <- data.table(quad_quadgrams)
combi_quad_ns_ns <- unite(quad_quadgrams, trigrams, c("word1", "word2", "word3"), sep = " ")
rm(quad_quadgrams)
write.csv(combi_quad_ns_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20test/combi_qu
rm(combi_quad_ns_ns)
```

quingrams

```r
quin_quingrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/quin_
quin_quingrams <- data.table(quin_quingrams)
combi_quin_s_s <- unite(quin_quingrams, quadgrams, c("word1", "word2", "word3", "word4"), sep = " ")
rm(quin_quingrams)
write.csv(combi_quin_s_s,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_qu
rm(combi_quin_s_s)
quin_quingrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/quin_
quin_quingrams <- data.table(quin_quingrams)
combi_quin_ns_s <- unite(quin_quingrams, quadgrams, c("word1", "word2", "word3", "word4"), sep = " ")
rm(quin_quingrams)
write.csv(combi_quin_ns_s,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_
rm(combi_quin_ns_s)
quin_quingrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/quin_
quin_quingrams <- data.table(quin_quingrams)
combi_quin_s_ns <- unite(quin_quingrams, quadgrams, c("word1", "word2", "word3", "word4"), sep = " ")
rm(quin_quingrams)
write.csv(combi_quin_s_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_
rm(combi_quin_s_ns)
```

Now for ns_ns

```
quin_quingrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/quin_g
quin_quingrams <- data.table(quin_quingrams)
combi_quin_ns_ns <- unite(quin_quingrams, quadgrams, c("word1", "word2", "word3", "word4"), sep = " ")
rm(quin_quingrams)
write.csv(combi_quin_ns_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20sample/combi_
rm(combi_quin_ns_ns)
```

Lastly, for test

```
quin_quingrams <- read.csv(file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20test/quin_gra
quin_quingrams <- data.table(quin_quingrams)
combi_quin_ns_ns <- unite(quin_quingrams, quadgrams, c("word1", "word2", "word3", "word4"), sep = " ")
rm(quin_quingrams)
write.csv(combi_quin_ns_ns,file = "/Users/mutecypher/Documents/Coursera/Capstone Project/20test/combi_qu
rm(combi_quin_ns_ns)
```

Including Stuff at the end