Topic Recommender

Parameter

Topic cutoff = 10

Num of neighbours = 10

Sim function = topic-based

Success Rate

```
setwd("/Users/juri/Desktop/RFiles/Results_k10_n10/")
sr_r1 <-  read.csv("SR_Round1", sep="\t", header = F)
sr_r2 <-  read.csv("SR_Round2", sep="\t", header = F)
sr_r3 <-  read.csv("SR_Round3", sep="\t", header = F)
sr_r4 <-  read.csv("SR_Round4", sep="\t", header = F)
sr_r5 <-  read.csv("SR_Round5", sep="\t", header = F)
sr_r6 <-  read.csv("SR_Round6", sep="\t", header = F)
sr_r7 <-  read.csv("SR_Round7", sep="\t", header = F)
sr_r8 <-  read.csv("SR_Round8", sep="\t", header = F)
sr_r9 <-  read.csv("SR_Round9", sep="\t", header = F)
sr_r10 <-  read.csv("SR_Round10", sep="\t", header = F)
sr_tot <- rbind(sr_r1, sr_r2, sr_r3, sr_r4, sr_r5, sr_r6, sr_r7, sr_r8, sr_r9, sr_r10)
sr_tot$V3 <- NULL
sr_tot <- sr_tot %>% group_by(sr_tot$V1) %>% summarise(mean = mean(V2))

sr_tot <- sr_tot %>% rename(k = `sr_tot$V1`)
sr_tot
```

```
## # A tibble: 20 x 2
##        k  mean
##    <int> <dbl>
##  1     1 0.393
##  2     2 0.524
##  3     3 0.591
##  4     4 0.643
##  5     5 0.683
##  6     6 0.714
##  7     7 0.733
##  8     8 0.752
##  9     9 0.767
## 10    10 0.776
## 11    11 0.787
## 12    12 0.796
## 13    13 0.803
## 14    14 0.810
## 15    15 0.819
```

```
## 16    16 0.825
## 17    17 0.830
## 18    18 0.833
## 19    19 0.834
## 20    20 0.836
```

```
sr_tot$mean %>% summary()
```

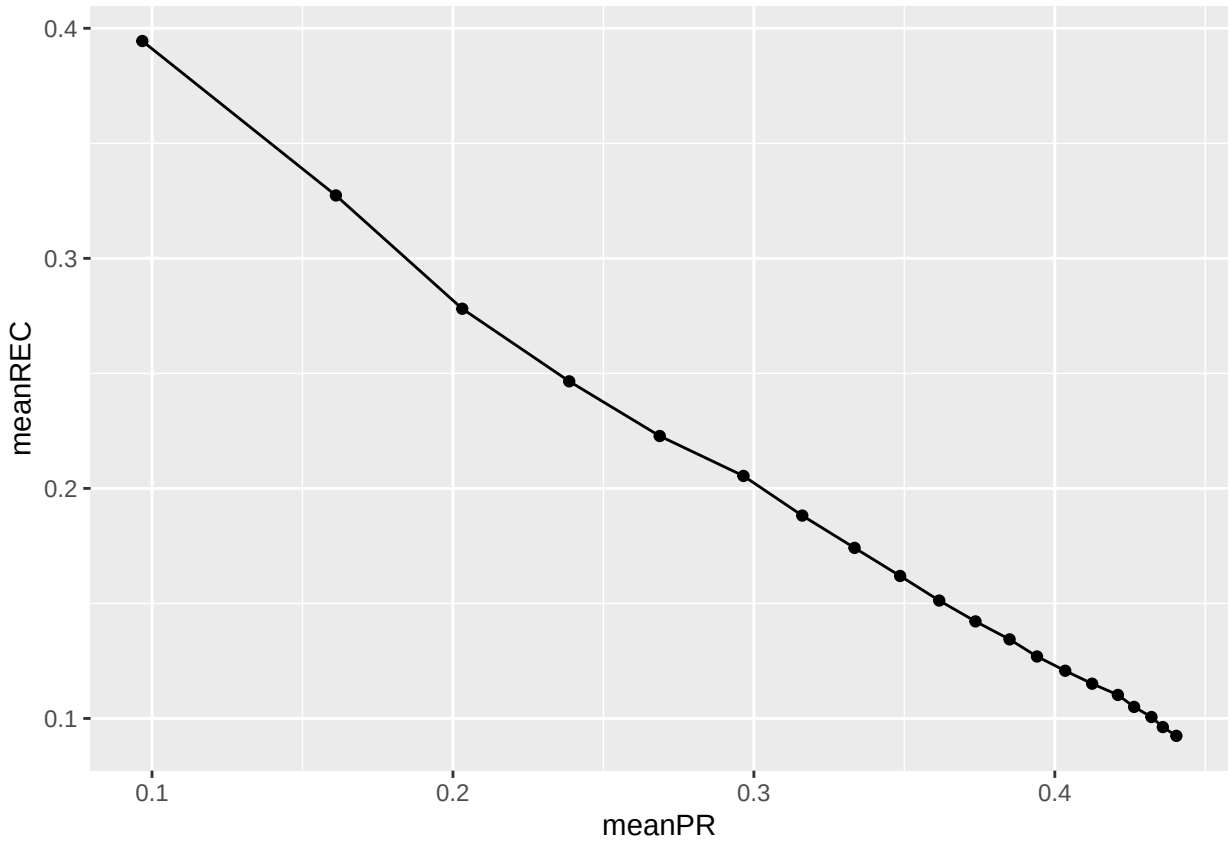
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3934  0.7061  0.7817  0.7375  0.8208  0.8360
```

Precision and Recall

```
setwd("/Users/juri/Desktop/RFiles/Results_k10_n10/")
pr_r1 <-  read.csv("PRC_Round1", sep="\t", header = F)
pr_r2 <-  read.csv("PRC_Round2", sep="\t", header = F)
pr_r3 <-  read.csv("PRC_Round3", sep="\t", header = F)
pr_r4 <-  read.csv("PRC_Round4", sep="\t", header = F)
pr_r5 <-  read.csv("PRC_Round5", sep="\t", header = F)
pr_r6 <-  read.csv("PRC_Round6", sep="\t", header = F)
pr_r7 <-  read.csv("PRC_Round7", sep="\t", header = F)
pr_r8 <-  read.csv("PRC_Round8", sep="\t", header = F)
pr_r9 <-  read.csv("PRC_Round9", sep="\t", header = F)
pr_r10 <-  read.csv("PRC_Round10", sep="\t", header = F)
pr_tot <- rbind(pr_r1, pr_r2, pr_r3, pr_r4, pr_r5, pr_r6, pr_r7, pr_r8, pr_r9, pr_r10)
pr_tot <- pr_tot %>% group_by(pr_tot$V1) %>% summarise(meanPR = mean(V2), meanREC = mean(V3)) %>% renam
pr_tot
```

```
## # A tibble: 20 x 3
##        k meanPR meanREC
##    <int>  <dbl>   <dbl>
##  1     1 0.0968  0.394
##  2     2 0.161   0.327
##  3     3 0.203   0.278
##  4     4 0.239   0.247
##  5     5 0.269   0.223
##  6     6 0.297   0.205
##  7     7 0.316   0.188
##  8     8 0.333   0.174
##  9     9 0.349   0.162
## 10    10 0.362   0.151
## 11    11 0.374   0.142
## 12    12 0.385   0.134
## 13    13 0.394   0.127
## 14    14 0.403   0.121
## 15    15 0.412   0.115
## 16    16 0.421   0.110
## 17    17 0.426   0.105
## 18    18 0.432   0.101
## 19    19 0.436   0.0962
## 20    20 0.440   0.0924
```

```
pr_tot %>% ggplot(aes(x=meanPR, y=meanREC)) + geom_line() + geom_point()
```



Precision summary

```
pr_tot$meanPR %>% summary()
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.09678 0.28959 0.36762 0.33745 0.41454 0.44040
```

Recall summary

```
pr_tot$meanREC %>% summary()
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.09244 0.11389 0.14671 0.17470 0.20976 0.39446
```