Topic Recommender

Parameter

Topic cutoff = 15

Num of neighbours = 15

Sim function = topic-based

Success Rate

```
setwd("/Users/juri/Desktop/RFiles/Results_k15_n15/")
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 0.397
##  2     2 0.525
##  3     3 0.617
##  4     4 0.664
##  5     5 0.698
##  6     6 0.728
##  7     7 0.749
##  8     8 0.768
##  9     9 0.785
## 10    10 0.795
## 11    11 0.807
## 12    12 0.815
## 13    13 0.825
## 14    14 0.831
## 15    15 0.835
```

```
## 16    16 0.841
## 17    17 0.846
## 18    18 0.851
## 19    19 0.856
## 20    20 0.860
```

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

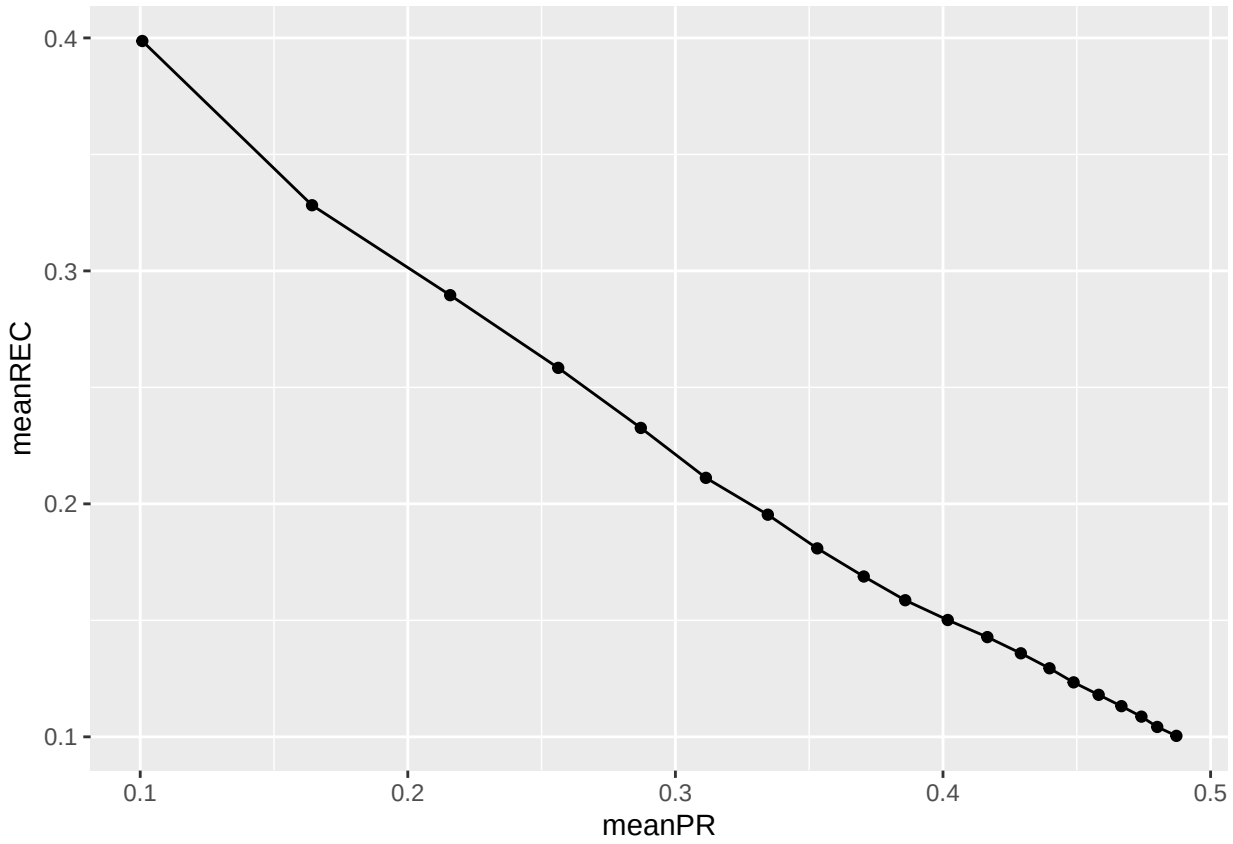
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3974  0.7207  0.8009  0.7546  0.8368  0.8604
```

Precision and Recall

```
setwd("/Users/juri/Desktop/RFiles/Results_k15_n15/")
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  0.101   0.399
##  2     2  0.164   0.328
##  3     3  0.216   0.290
##  4     4  0.256   0.258
##  5     5  0.287   0.233
##  6     6  0.311   0.211
##  7     7  0.335   0.195
##  8     8  0.353   0.181
##  9     9  0.370   0.169
## 10    10  0.386   0.159
## 11    11  0.402   0.150
## 12    12  0.417   0.143
## 13    13  0.429   0.136
## 14    14  0.440   0.129
## 15    15  0.449   0.123
## 16    16  0.458   0.118
## 17    17  0.467   0.113
## 18    18  0.474   0.109
## 19    19  0.480   0.104
## 20    20  0.487   0.100
```

```
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.
##  0.1008  0.3053  0.3938  0.3641  0.4512  0.4872
```

Recall summary

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

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1004  0.1220  0.1544  0.1824  0.2165  0.3987
```