

# Lab 6. USER-based collaborative filtering

## Building the recommendation system:

```
recommender_models <- recommenderRegistry$get_entries(dataType = "realRatingMatrix")
recommender_models$UBCF_realRatingMatrix$parameters

## $method
## [1] "cosine"
##
## $nn
## [1] 25
##
## $sample
## [1] FALSE
##
## $normalize
## [1] "center"
##
## $minRating
## [1] NA

recc_model <- Recommender(data = recc_data_train, method = "UBCF")
recc_model

## Recommender of type 'UBCF' for 'realRatingMatrix'
## learned using 111 users.

model_details <- getModel(recc_model)
names(model_details)

## [1] "description" "data"          "method"        "nn"           "sample"
## [6] "normalize"    "minRating"     "verbose"

model_details$data

## 111 x 332 rating matrix of class 'realRatingMatrix' with 10507 ratings.
## Normalized using center on rows.
```

## Applying the recommender model on the test set

```
n_recommended <- 6
recc_predicted <- predict(object = recc_model,
                          newdata = recc_data_test,
```

```
        n = n_recommended)

recc_predicted

## Recommendations as 'topNList' with n = 6 for 449 users.
```

## Explore results

```
recc_matrix <- sapply(recc_predicted@items,
                      function(x){ colnames(ratings_movies)[x] })

dim(recc_matrix)

## [1] 6 449

recc_matrix[, 1:4]

##      [,1]
## [1,] "Schindler's List (1993)"
## [2,] "L.A. Confidential (1997)"
## [3,] "Close Shave, A (1995)"
## [4,] "Vertigo (1958)"
## [5,] "Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb (1963)"
## [6,] "One Flew Over the Cuckoo's Nest (1975)"
##      [,2]
## [1,] "Silence of the Lambs, The (1991)"
## [2,] "Schindler's List (1993)"
## [3,] "Princess Bride, The (1987)"
## [4,] "Raiders of the Lost Ark (1981)"
## [5,] "Usual Suspects, The (1995)"
## [6,] "Shawshank Redemption, The (1994)"
##      [,3]
## [1,] "Silence of the Lambs, The (1991)"
## [2,] "Fargo (1996)"
## [3,] "Star Wars (1977)"
## [4,] "Swingers (1996)"
## [5,] "Willy Wonka and the Chocolate Factory (1971)"
## [6,] "Godfather, The (1972)"
##      [,4]
## [1,] "Usual Suspects, The (1995)"
## [2,] "Terminator 2: Judgment Day (1991)"
## [3,] "Sling Blade (1996)"
## [4,] "Terminator, The (1984)"
## [5,] "Shawshank Redemption, The (1994)"
```

```
## [6,] "Twelve Monkeys (1995)"
```

How many times each movie got recommended:

```
number_of_items <- factor(table(recc_matrix))

chart_title <- "Distribution of the number of items for UBCF"

qplot(number_of_items) + ggtitle(chart_title)
```

Top titles:

```
number_of_items_sorted <- sort(number_of_items, decreasing = TRUE)
number_of_items_top <- head(number_of_items_sorted, n = 4)
table_top <- data.frame(names(number_of_items_top), number_of_items_top)
head(table_top)
```

##	names.number_of_items_top.
## Usual Suspects, The (1995)	Usual Suspects, The (1995)
## Silence of the Lambs, The (1991)	Silence of the Lambs, The (1991)
## Sling Blade (1996)	Sling Blade (1996)
## Godfather, The (1972)	Godfather, The (1972)
##	number_of_items_top
## Usual Suspects, The (1995)	172
## Silence of the Lambs, The (1991)	131
## Sling Blade (1996)	126
## Godfather, The (1972)	125