Hybrid recommendation systems based on bayesian network

Chenxi He, Shengqian Chen, Wenxue Zhang

March 16, 2016

Canonical weighted sum

Let X_i be a node in a BN, let $Pa(X_i)$ be the parent set of X_i, and Y_k be the k^{th} parent of X_i in the BN. By using a canonical weighted sum, the set of conditional probability distributions stored at node X_i are then represented by means of

$$Pr(x_{ij} \mid pa(X_i)) = \sum_{Y_k \in Pa(X_i)} w(y_{k,l}, x_{i,j})$$

where $w(y_{k,l}, x_{i,j})$ are weights(effects) measuring how this l^{th} value of variable Y_k describes the j^{th} state of node X_i.

$$\sum_{j=1}^{r} \sum_{Y_k \in Pa(X_i)} w(y_{k,l}, x_{i,j}) = 1$$

Related theorems

Theorem 1 $Pr(x_{a,s} \mid ev) = \sum_{j=1}^{m_{x_a}} \sum_{k=1}^{l_{Y_j}} w(y_{j,k}, x_{a,s}) \cdot Pr(y_{j,k} \mid ev)$

Theorem 2
if $F_k \notin Pa(I_j)$
$Pr(f_{k,1} \mid i_{j,1}) = Pr(f_{k,1})$

if $F_k \in Pa(I_j)$
$Pr(f_{k,1} \mid i_{j,1}) = Pr(f_{k,1}) + \frac{w(f_{k,1}, i_{j,1}) Pr(f_{k,1}(1 - Pr(f_{k,1}))}{Pr(i_{j,1})}$

where $Pr(i_{j,1}) = \sum_{F_k \in Pa(I_j)} w(f_{k,1}, i_{j,1}) Pr(f_{k,1})$

Algorithm(Now we've completed CB and we'll further develop CF and hybrid part later)

1.Content-based propagation:
-$ev_{cb} == I_j \quad Pr(i_{j,1} \mid ev) = 1$
Compute $Pr(F_k \mid ev)$ using Theorem 2
-Propagate to items using Theorem1.
-Propagate to A_{CB} and $U_i \in U_1^-$ using Theorem 1.

2.Collaborative propagation

3.Combine content-based and collaborative likelihoods at hybrid node A_H

4.Select the predicted rating.

example data

- features $\{f_1, f_2, f_3, f_4\}$
- movies $\{i_1, i_2, i_3\}$
- users $\{u_1, u_2\}$

table1: row:movies; column:features; entry:{0,1}

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

table2: row:user; column:movies; entry:{0,1-5}

$$\begin{pmatrix} 0 & 3 & 5 \\ 4 & 2 & 4 \end{pmatrix}$$

Here we try to predict what user1 will rate item1.

CB-part Algorithm Application in the previous data

First we need to calculate the canonical weight in two circumstance (item,feature) and (item,user)

```
table3
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 0.5693234 0.0000000 0.0000000 0.4306766
## [2,] 0.0000000 0.3058654 0.4627564 0.2313782
## [3,] 0.3627827 0.3627827 0.0000000 0.2744345
```

```
wui_rate3
```

```
## [1] 0.0 0.5 0.0
```

Now we can compute the $Pr(F_k \mid ev)$, here in this example $evidence_{cb}$ is I_1, so it's a vector $P(F_k \mid I_1)$

```
p.featuregivenev
```

CB-part Algorithm Application in the previous data

```r
t(p.itemgivenev)
```

```
##         [,1]      [,2]      [,3]
## [1,]       1 0.3247369 0.4935499
```

Finally we can get the probabiliy how the user will rate the item:

```r
names(p)=c("P(U_1,0 | EV_cb)","P(U_1,1 | EV_cb)","P(U_1,2
t(as.matrix(p))
```

```
##         P(U_1,0 | EV_cb) P(U_1,1 | EV_cb) P(U_1,2 | EV_cb)
## [1,]           0.9091434                0                0
##         P(U_1,4 | EV_cb) P(U_1,5 | EV_cb)
## [1,]                   0         0.246775
```