

Content_based_recommendation

February 12, 2019

```
In [2]: import Pkg; Pkg.add("Distances")
```

```
Resolving package versions...
Updating `~/julia/Project.toml`
[b4f34e82] + Distances v0.7.3
Updating `~/julia/Manifest.toml`
[b4f34e82] + Distances v0.7.3
[2a0f44e3] + Base64
[8ba89e20] + Distributed
[b77e0a4c] + InteractiveUtils
[8f399da3] + Libdl
[37e2e46d] + LinearAlgebra
[56ddb016] + Logging
[d6f4376e] + Markdown
[de0858da] + Printf
[9a3f8284] + Random
[9e88b42a] + Serialization
[6462fe0b] + Sockets
[2f01184e] + SparseArrays
[10745b16] + Statistics
[8dfed614] + Test
[4ec0a83e] + Unicode
```

```
In [7]: using Distances
```

```
In [9]: # toy data
```

```
um = [4 0 0 5 1 0 0;      # utility matrix
      5 5 4 0 0 0 0;
      0 0 0 2 4 5 0;
      0 3 0 0 0 0 3]

ip = [1 1 0 0 0 0 0;      # item profiles
      1 1 0 0 0 0 0;
      1 1 0 0 0 0 0;
      0 0 1 0 0 1 0;
      0 0 0 1 1 0 1;
```

```

0 0 0 1 1 0 1;
0 0 0 1 1 0 1]

up = zeros(4,7) # empty user profiles

# building user profile
for i=1:4 # num of users
    for j=1:7 # num of feature
        si = findall(um[i,:] .!= 0) # si is set of voted items
        up[i,j] = sum(ip[si,j])/length(si)
    end
end

sim = zeros(4,7)
for i=1:4
    for j=1:7
        sim[i,j] = sqrt(sum((up[i,:] - ip[j,:]).^2)) # Euclidean
    end
end

# measuring similarity between items and users
R = pairwise(Euclidean(),ip',up')

R
#R[1,:]'

```

```

Out[9]: 4E7 LinearAlgebra.Adjoint{Float64,Array{Float64,2}}:
 1.20185  1.20185  1.20185  1.20185  1.33333  1.33333  1.33333
 0.0      0.0      0.0      2.0      2.23607  2.23607  2.23607
 1.88562  1.88562  1.88562  1.49071  0.745356 0.745356 0.745356
 1.11803  1.11803  1.11803  1.80278  1.11803  1.11803  1.11803

```