RECOMMENDER SYSTEMS MEET LINKED OPEN DATA

Tommaso Di Noia
tommaso.dinoia@poliba.it

POLITECNICO DI BARI

Joint work with
Vito Claudio Ostuni, Roberto Mirizzi and Eugenio Di Sciascio
Recommender Systems

**A definition**
Recommendor Systems (RSs) are *software tools and techniques providing suggestions for items to be of use to a user.*


**Input Data:**
A set of users $U=\{u_1, ..., u_M\}$
A set of items $X=\{x_1, ..., x_N\}$
The rating matrix $R=[r_{u,i}]$

**Problem Definition:**
Given user $u$ and target item $i$
Predict the rating $r_{u,i}$
The rating matrix

<table>
<thead>
<tr>
<th></th>
<th>5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>3</th>
<th>??</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tommaso</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Enrico</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Sean</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Natasha</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Valentina</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>
The rating matrix (in the real world)

<table>
<thead>
<tr>
<th></th>
<th>Tommaso</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>??</td>
<td></td>
</tr>
<tr>
<td>Enrico</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sean</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natasha</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valentina</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
Collaborative Recommender Systems

Collaborative RSs recommend items to a user by identifying other users with a similar profile.
Content-based Recommender Systems

CB-RSs recommend items to a user based on their description and on the profile of the user’s interests.

User profile

- Item1, 5
- Item2, 1
- Item5, 4
- Item10, 5
- ...

Items

- Item1
- Item2
- ....
- Item100

Recommender System

Top-N Recommendations

- Item7
- Item15
- Item11
- ...
Knowledge-based Recommender Systems

KB-RSs recommend items to a user based on their description and domain knowledge encoded in a knowledge base.

User profile

Item1, 5
Item2, 1
Item5, 4
Item10, 5
...

Items

Item1
Item2
...
Item100

Item’s descriptions

Knowledge-base

Top-N Recommendations

Item7
Item15
Item11
...

Recommender System
Content-Based Recommender Systems

Content-Based Recommender Systems

Content-Based Recommender Systems

- Items are described in terms of attributes/features
- A finite set of values is associated to each feature
- Items representation is a (Boolean) vector
- Novelty, diversity, serendipity
Main CB RSs Drawback: Limited Content Analysis

No suggestion is available if the analyzed content does not contain enough information to discriminate items the user might like from items the user might not like. *

The quality of CB recommendations are correlated with the quality of the features that are explicitly associated with the items.

Need of domain knowledge!
We need rich descriptions of the items!

A solution based on Linked Data

Use Linked Data to mitigate the limited content analysis issue

- Plenty of structured data available
- No Content Analyzer required
Linked Data as structured information source for items descriptions

Rich items descriptions

<table>
<thead>
<tr>
<th></th>
<th>STARRING</th>
<th>DIRECTOR</th>
<th>SUBJECT+BROADER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat</td>
<td>Robert DeNiro</td>
<td>Al Pacino</td>
<td>Michael Mann</td>
</tr>
<tr>
<td></td>
<td>Heat</td>
<td></td>
<td>Heist films</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Crime films</td>
</tr>
</tbody>
</table>
SELECT count(?i) AS ?num ?c
WHERE {
    ?i a ?c .
    FILTER(regex(?c, "^http://dbpedia.org/ontology")) .
}
ORDER BY DESC(?num)

<table>
<thead>
<tr>
<th>num</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>956476</td>
<td><a href="http://dbpedia.org/ontology/Agent">http://dbpedia.org/ontology/Agent</a></td>
</tr>
<tr>
<td>763644</td>
<td><a href="http://dbpedia.org/ontology/Person">http://dbpedia.org/ontology/Person</a></td>
</tr>
<tr>
<td>572728</td>
<td><a href="http://dbpedia.org/ontology/Place">http://dbpedia.org/ontology/Place</a></td>
</tr>
<tr>
<td>387166</td>
<td><a href="http://dbpedia.org/ontology/PopulatedPlace">http://dbpedia.org/ontology/PopulatedPlace</a></td>
</tr>
<tr>
<td>348520</td>
<td><a href="http://dbpedia.org/ontology/Settlement">http://dbpedia.org/ontology/Settlement</a></td>
</tr>
<tr>
<td>333270</td>
<td><a href="http://dbpedia.org/ontology/Work">http://dbpedia.org/ontology/Work</a></td>
</tr>
<tr>
<td>277476</td>
<td><a href="http://dbpedia.org/ontology/OrganisationMember">http://dbpedia.org/ontology/OrganisationMember</a></td>
</tr>
<tr>
<td>277476</td>
<td><a href="http://dbpedia.org/ontology/SportsTeamMember">http://dbpedia.org/ontology/SportsTeamMember</a></td>
</tr>
</tbody>
</table>
Select the sub-graph you are interested in

```sparql
select ?m ?p ?o1
where{
  ?m dcterms:subject ?o.
  dbpedia:The_Matrix dcterms:subject ?o.
  ?m a dbpedia-owl:Film.
  filter(regex(?p,"^http://dbpedia.org/ontology/"))
}
order by ?m
```
Computing similarity in LOD datasets

- Brian Dennehy
- Drama
- Al Pacino
- Robert De Niro
- Heat
- Righteous Kill
- Starring
- Director
- John Avnet
- Michael Mann
- Serial killer films
- Heist films
- Crime films
- Broader
- Subject
- Genre
Vector Space Model for LOD

Righteous Kill
Heat
... ...

genre
subject/broader
director
starring

Righteous Kill
Heat

starring

Robert DeNiro
Brian Dennehy
Heat
Al Pacino
Righteous Kill
## Vector Space Model for LOD

<table>
<thead>
<tr>
<th>STARRING</th>
<th>Al Pacino (v1)</th>
<th>Robert De Niro (v2)</th>
<th>Brian Dennehy (v3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Righteous Kill (m1)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Heat (m2)</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

**Righteous Kill**

**Heat**

\[
W_{AlPacino,Heat} = tf_{AlPacino,Heat} \times idf_{AlPacino}
\]

<table>
<thead>
<tr>
<th>Righteous Kill (x1)</th>
<th>(w_{v1,x1})</th>
<th>(w_{v2,x1})</th>
<th>(w_{v3,x1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat (x2)</td>
<td>(w_{v1,x2})</td>
<td>(w_{v2,x2})</td>
<td>0</td>
</tr>
</tbody>
</table>
Cinemappy
Exploratory search

Recommend to Explore
Explore to Recommend

• After recommending a new item the users are allowed to explore new knowledge spaces they were not aware of
  – Suggest new paths to discover
• The exploration task enriches the user profile:
  – Exploration paths can be exploited to recommend further items to the users
  – Foster serendipitous discoveries
THANK YOU

Tommaso Di Noia
tommaso.dinoia@poliba.it

POLITECNICO DI BARI

Semantic Web
Tra ontologie e Open Data

Tommaso Di Noia, Roberto De Virgilio, Eugenio Di Sciascio, Francesco Donini

APOGEO