Linked Data Mining Challenge at Know@LOD
ESWC 2015, Portoroz, Slovenia, May 31st

Suad Aldarra
Emir Muñoz
Imagine a movie
Imagine a movie

Vampire Female
Imagine a movie

Vampire Female

Leather Pants
Imagine a movie

Vampire Female

Swords & Axes

Leather Pants
Imagine a movie

Vampire Female
Swords & Axes

Leather Pants
Chasing & Stabbing
Imagine a movie

Vampire Female

Swords & Axes

Based on a video game

Leather Pants

Chasing & Stabbing
Imagine a movie

Vampire Female

Swords & Axes

Based on a video game

Leather Pants

Chasing & Stabbing

Sequel
Imagine a movie

Vampire Female
Swords & Axes
Based on a video game

Leather Pants
Chasing & Stabbing
Sequel
Low budget
Imagine a movie

Vampire Female

Swords & Axes

Based on a video game

Rate R

Leather Pants

Chasing & Stabbing

Sequel

Low budget
Imagine a movie

- Vampire Female
- Swords & Axes
- Based on a video game
- Rate R
- Non famous actors
- Leather Pants
- Chasing & Stabbing
- Sequel
- Low budget
- Based on a video game Sequel
Imagine a movie

Vampire Female

Swords & Axes

Based on a video game

Rate R

Non famous actors

Leather Pants

Chasing & Stabbing

Sequel

Low budget
How to pick a good movie?
How to pick a good movie?

Extracting Features

- Freebase url
- Independent film
- Sequel Film
- Based on literature
- Release Date
How to pick a good movie?

Extracting Features

- Freebase url
- Independent film
- Sequel Film
- Based on literature
- Release Date

- Movie IMDB id
- Actor Date Of Birth
- Actor Gender
- Director Date Of Birth
- Director Gender
How to pick a good movie?  
Extracting Features

- Freebase url
- Independent film
- Sequel Film
- Based on literature
- Release Date
- Movie IMDB id
- Actor Date Of Birth
- Actor Gender
- Director Date Of Birth
- Director Gender
- OMDB API
  - MPAA rating
  - Genre
  - Directors
  - Actors
  - language
  - Country
  - Runtime

- How to pick a good movie?

- Extracting Features

- Freebase
- DBpedia
- OMDB API

- How to pick a good movie?

- Extracting Features

- Freebase
- DBpedia
- OMDB API

- How to pick a good movie?

- Extracting Features

- Freebase
- DBpedia
- OMDB API

- How to pick a good movie?

- Extracting Features

- Freebase
- DBpedia
- OMDB API
How to pick a good movie?
Extracting Features

- **DBpedia**
  - Freebase url
  - Independent film
  - Sequel Film
  - Based on literature
  - Release Date

- **Freebase**
  - Movie IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender

- **OMDB API**
  - MPAA rating
  - Genre
  - Directors
  - Actors
  - language
  - Country
  - Runtime

- **IMDb**
  - Budget
  - Gross
  - Actor Awards
  - Director Awards
  - Plot keywords
How to pick a good movie?
Extracting Features

- **DBpedia**: Freebase url, Independent film, Sequel Film, Based on literature, Release Date
- **Freebase**: Movie IMDB id, Actor Date Of Birth, Actor Gender, Director Date Of Birth, Director Gender
- **OMDB API**: MPAA rating, Genre, Directors, Actors, language, Country, Runtime
- **IMDb**: Budget, Gross, Actor Awards, Director Awards, Plot keywords, Critics Textual Reviews

**How to pick a good movie?**

**Extracting Features**

- **DBpedia**: Freebase url, Independent film, Sequel Film, Based on literature, Release Date
- **Freebase**: Movie IMDB id, Actor Date Of Birth, Actor Gender, Director Date Of Birth, Director Gender
- **OMDB API**: MPAA rating, Genre, Directors, Actors, language, Country, Runtime
- **IMDb**: Budget, Gross, Actor Awards, Director Awards, Plot keywords, Critics Textual Reviews

**How to pick a good movie?**

**Extracting Features**

- **DBpedia**: Freebase url, Independent film, Sequel Film, Based on literature, Release Date
- **Freebase**: Movie IMDB id, Actor Date Of Birth, Actor Gender, Director Date Of Birth, Director Gender
- **OMDB API**: MPAA rating, Genre, Directors, Actors, language, Country, Runtime
- **IMDb**: Budget, Gross, Actor Awards, Director Awards, Plot keywords, Critics Textual Reviews
How to pick a good movie?
Extracting Features

- **DBpedia**
  - Freebase url
  - Independent film
  - Sequel Film
  - Based on literature
  - Release Date

- **Freebase**
  - Movie IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender

- **OMDB API**
  - MPAA rating
  - Genre
  - Directors
  - Actors
  - language
  - Country
  - Runtime

- **IMDb**
  - Budget
  - Gross
  - Actor Awards
  - Director Awards
  - Plot keywords

- **Metacritic**
  - Critics Textual Reviews

- **Release Date**
  - Released_weekend
  - Released_weekday

**Extracting Features**

- **How to pick a good movie?**
  - Extracting Features
How to pick a good movie?
Extracting Features
How to pick a good movie?
Extracting Features

- #Female Actors
- #Male Actors
- #Actors>50
- #Actors<30
- #Actors30-50
- Released_weekend
- Released_weekday
- Released_Date
- Release_Date
- Based on literature
- Sequel Film
- Independent film
- Freebase url
- Movie IMDB id
- Actor Date Of Birth
- Actor Gender
- Director Date Of Birth
- Director Gender
- Director Awards
- Directors
- Actors
- language
- Country
- Runtime
- Budget
- Gross
- MPAA rating
- Genre
- Actors
- Plot keywords
- Critics Textual Reviews
How to pick a good movie?

Extracting Features

- **DBpedia**
  - #Female Actors
  - #Male Actors
  - #Actors>50
  - #Actors<30
  - #Actors30-50
  - Released_weekend
  - Released_weekday

- **Freebase**
  - Freebase url
  - Independent film
  - Based on literature
  - Release Date
  - Sequel Film

- **OMDB API**
  - Movie IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender
  - Genre
  - Directors
  - Actors
  - Language
  - Country
  - Runtime

- **IMDb**
  - MPAA rating
  - Budget
  - Gross
  - Actor Awards
  - Director Awards
  - Plot keywords
  - Critics
  - Textual Reviews

- **MetaCritic**
  - Critics Textual Reviews
How to pick a good movie?
Extraction Features

- **DBpedia**
  - Common Language
  - #Female Actors
  - #Male Actors
  - #Actors > 50
  - #Actors < 30
  - #Actors 30-50
  - Released_weekend
  - Released_weekday
  - Release Date
  - Based on literature
  - Sequel Film
  - Independent film

- **Freebase**
  - Freebase url
  - Movie IMDB id
  - Actor IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director IMDB id
  - Director Date Of Birth
  - Director Gender

- **OMDB API**
  - MPAA rating
  - Genre
  - Director
  - Actors
  - Language
  - Country
  - Runtime

- **IMDb**
  - Budget
  - Gross
  - Actor Awards
  - Director Awards
  - Plot keywords
  - Critics
  - Textual Reviews

**How to pick a good movie?**

**Extracting Features**

- **Release Date**
- **Actors**
- **Actors Gender**
- **Actors Number**
- **Actors Type**
- **Actor Awards**
- **Actors Country**
- **Actor Runtime**
- **Actor Language**
- **Actor Director**
- **Actor Movie**
- **Actor Movie Genre**
- **Actor Movie MPAA rating**
- **Actor Movie Budget**
- **Actor Movie Gross**
- **Actor Movie Director Awards**
- **Actor Movie Director**
- **Actor Movie Language**
- **Actor Movie Country**
- **Actor Movie Runtime**
- **Actor Movie Plot keywords**
- **Actor Movie Critics**
- **Actor Movie Textual Reviews**
- **Actor Movie Review**
- **Actor Movie Review Textual**
- **Actor Movie Review Critics**
- **Actor Movie Review Textual Critics**
- **Actor Movie Review Textual Textual**
- **Actor Movie Review Textual Critics Textual**
How to pick a good movie?

Extracting Features

**DBpedia**
- Freebase url
- Independent film
- Sequel Film
- Based on literature
- Release Date

**Freebase**
- Movie IMDB id
- Actor Date Of Birth
- Actor Gender
- Director Date Of Birth
- Director Gender

**OMDB API**
- MPAA rating
- Genre
- Directors
- Actors
- language
- Country
- Runtime

**IMDb**
- Budget
- Gross
- Actor Awards
- Director Awards
- Plot keywords

**Metacritic**
- Critics Textual Reviews

- #Female Actors
- #Male Actors
- #Actors>50
- #Actors<30
- #Actors30-50

- #Directors>50
- #Directors<30
- #Directors30-50

- Common Country
- Common Language

- Released_weekend
- Released_weekday
How to pick a good movie?

Extracting Features

- DBpedia
  - Freebase url
  - Independent film
  - Based on literature
  - Sequel Film
  - Released_date

- Freebase
  - Movie IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender

- OMDB API
  - MPAA rating
  - Genre
  - Directors
  - Actors
  - Language
  - Country
  - Runtime

- IMDb
  - Budget
  - Gross
  - Actor Awards
  - Director Awards
  - Plot keywords

- Metacritic
  - Critics Textual Reviews

- Common Country
- Common Language
- #Female Actors
- #Male Actors
- #Actors>50
- #Actors<30
- #Actors30-50
- #Directors>50
- #Directors<30
- #Directors30-50

How to pick a good movie:

1. Extracting Features
2. DBpedia
3. Freebase
4. OMDB API
5. IMDb
6. Metacritic

Common Country
Common Language
#Female Actors
#Male Actors
#Actors>50
#Actors<30
#Actors30-50
#Directors>50
#Directors<30
#Directors30-50

How to pick a good movie?

- Common Country
- Common Language
- #Female Actors
- #Male Actors
- #Actors>50
- #Actors<30
- #Actors30-50
- #Directors>50
- #Directors<30
- #Directors30-50
How to pick a good movie?

Extracting Features

- **DBpedia**
  - Freebase url
  - Independent film
  - Based on literature
  - Release Date
  - Sequel Film
  - #Actors>50
  - #Actors<30
  - #Actors30-50
  - #Female Actors
  - #Male Actors

- **Freebase**
  - Movie IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender
  - #Directors>50
  - #Directors<30
  - #Directors30-50

- **OMDB API**
  - MPAA rating
  - Genre
  - Directors
  - Actors
  - language
  - Country
  - Runtime
  - #Actors>50
  - #Actors<30
  - #Actors30-50

- **IMDb**
  - Budget
  - Gross
  - Actor Awards
  - Director Awards
  - Plot keywords
  - Critics Textual Reviews

- **metacritic**
  - Short Movie
  - Long Movie

- **Common Country**
- **Common Language**

- **Released_weekend**
- **Released_weekday**

- **DATE**
- **SCENE**
- **TAKE**
How to pick a good movie?

Extracting Features

- DBpedia
  - Freebase URL
  - Sequel Film
  - Based on literature
  - Release Date
  - Released_weekend
  - Released_weekday

- Freebase
  - Movie IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender

- OMDB API
  - MPAA rating
  - Genre
  - Directors
  - Actors
  - Country
  - Runtime

- IMDb
  - Budget
  - Gross
  - Actor Awards
  - Director Awards
  - Plot keywords
  - Critics Textual Reviews

- Metacritic

- #Female Actors
- #Male Actors
- #Actors>50
- #Actors<30
- #Actors30-50
- #Directors>50
- #Directors<30
- #Directors30-50

- High Budget
- Low Budget

- Common Country
- Common Language

- How to pick a good movie?

- Extracting Features

- Released_weekend
- Released_weekday

- Short Movie
- Long Movie
How to pick a good movie? 
Extracting Features

- **DBpedia**
  - **Freebase url**
  - **Independent film**
  - **Based on literature**
  - **Release Date**
  - **Sequel Film**
  - **Actors**
  - **Actors>50**
  - **Actors<30**
  - **Actors30-50**

- **Freebase**
  - **Movie IMDB id**
  - **Actor Date Of Birth**
  - **Actor Gender**
  - **Director Date Of Birth**
  - **Director Gender**

- **OMDB API**
  - **MPAA rating**
  - **Genre**
  - **Directors**
  - **Actors**
  - **language**
  - **Country**
  - **Runtime**

- **IMDb**
  - **Budget**
  - **Gross**
  - **Actor Awards**
  - **Director Awards**
  - **Plot keywords**

- **metacritic**
  - **Critics Textual Reviews**

- **Common Country**
  - **Common Language**

- **#Female Actors**
  - **#Male Actors**

- **#Directors>50**
  - **#Directors<30**
  - **#Directors30-50**

- **How to pick a good movie?**
  - Extracting Features

- **Release Date**
  - **Released_weekend**
  - **Released_weekday**

- **Short Movie**
  - **Long Movie**

- **HighBudget LowBudget**
  - **Gross>Budget**

- **Common Language**
  - **Common Country**
How to pick a good movie?
Extracting Features

- Common Country
- Common Language
- #Actors>50
- #Actors<30
- #Actors30-50
- #Female Actors
- #Male Actors
- #Directors>50
- #Directors<30
- #Directors30-50

- Released_weekend
- Released_weekday
- Released_date

- Released_weekend
- Released_weekday
- Released_date

- DBpedia
  - Freebase url
  - Independent film
  - Sequel Film
  - Based on literature
  - Release Date

- Freebase API
  - Movie IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender

- OMDB API
  - MPAA rating
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender

- IMDb
  - Budget
  - Gross
  - Genre
  - Actors
  - Directors
  - Director Awards
  - Plot keywords

- Metacritic
  - Critics Textual Reviews
  - Common Language
  - Common Country

- HighBudget LowBudget
- Gross>Budget
- Actors Oscar/Golden Globe Win/Nominated

- How to pick a good movie?
- Extracting Features

- Released_weekend
- Released_weekday
- Released_date

- Released_weekend
- Released_weekday
- Released_date

- DBpedia
  - Freebase url
  - Independent film
  - Sequel Film
  - Based on literature
  - Release Date

- Freebase API
  - Movie IMDB id
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender

- OMDB API
  - MPAA rating
  - Actor Date Of Birth
  - Actor Gender
  - Director Date Of Birth
  - Director Gender

- IMDb
  - Budget
  - Gross
  - Genre
  - Actors
  - Directors
  - Director Awards
  - Plot keywords

- Metacritic
  - Critics Textual Reviews
  - Common Language
  - Common Country

- HighBudget LowBudget
- Gross>Budget
- Actors Oscar/Golden Globe Win/Nominated
How to pick a good movie?

Extracting Features

- **DBpedia**
  - Freebase url
  - Independent film
  - Based on literature
  - Release Date

- **Freebase**
  - Sequel Film
  - Actor Date Of Birth
  - Director Date Of Birth

- **OMDB API**
  - Movie IMDB id
  - Actor Gender
  - Director Gender
  - Director Date Of Birth
  - Runtime
  - Genre

- **IMDb**
  - MPAA rating
  - Budget
  - Gross
  - Actor Awards
  - Director Awards
  - Plot keywords

- **metacritic**
  - Critics Textual Reviews

### Features
- Common Country
- Common Language
- #Female Actors
- #Male Actors
- #Directors>50
- #Directors<30
- #Directors30-50
- #Actors>50
- #Actors<30
- #Actors30-50

### How to Pick a Good Movie?

1. **Release Date**
   - Released_weekend
   - Released_weekday

2. **Runtime**
   - Short Movie
   - Long Movie

3. **Budget**
   - High Budget
   - Low Budget

4. **Gross**
   - Gross>Budget
   - Gross<Budget

5. **Actors**
   - Oscar/Golden Globe Win/Nominated

6. **Directors**
   - Oscar/Golden Globe Win/Nominated
How to pick a good movie?

Extracting Features

- Common Country
- Common Language
- #Actors>50
- #Actors<30
- #Actors30-50
- #Female Actors
- #Male Actors
- #Directors>50
- #Directors<30
- #Directors30-50

- Freebase url
- Independent film
- Sequel Film
- Based on literature
- Release Date

- Movie IMDB id
- Actor Date Of Birth
- Actor Gender
- Director Date Of Birth
- Director Gender

- OMDB API
  - MPAA rating
  - Genre
  - Director Awards
  - Director Gender
  - Actors
  - language
  - Country
  - Runtime

- IMDb
  - Budget
  - Gross
  - Actors
  - Director Awards
  - Plot keywords
  - Critics Textual Reviews

- metacritic
  - Critics
  - Textual Reviews

- HighBudget LowBudget
- Gross>Budget
- Actors Oscar/Golden Globe Win/Nominated
- Directors Oscar/Golden Globe Win/Nominated
- #Good Keywords
- #Bad Keywords
- #Mostly Good
- #Mostly Bad

- How to pick a good movie?
- Extracting Features

- Released_weekend
- Released_weekday

- Short Movie
- Long Movie
How to pick a good movie?

Extracting Features

- **DBpedia**
  - Movie url
  - IMDB id
  - Released_date
  - Released_weekend
  - Released_weekday

- **Freebase**
  - Freebase url
  - Independent film
  - Sequel Film
  - Based on literature

- **OMDB API**
  - Movie
  - Actor Date Of Birth
  - Director Date Of Birth
  - Actor Gender
  - Director Gender
  - Director Date Of Birth
  - MPAA rating
  - Genre
  - Country
  - Runtime

- **IMDb**
  - Movie
  - Actor
  - Actors
  - Director
  - Directors
  - Budget
  - Gross
  - MPAA rating
  - Critics Textual Reviews

- **Metacritic**
  - Critic Reviews
  - Sentiment

- **Common Country**
- **Common Language**
- **#Female Actors**
- **#Male Actors**
- **#Directors>50**
- **#Directors<30**
- **#Directors30-50**
- **#Actors>50**
- **#Actors<30**
- **#Actors30-50**
- **High Budget**
- **Low Budget**
- **Gross>Budget**
- **Actors Oscar/Golden Globe Win/Nominated**
- **Directors Oscar/Golden Globe Win/Nominated**
- **#Good Keywords**
- **#Bad Keywords**
- **#Mostly Good**
- **#Mostly Bad**
- **Common Language**
- **Common Country**
- **Short Movie**
- **Long Movie**
How to pick a good movie?
Extracting Features
How to pick a good movie?

Training Classifier

- 241 Features
- RDF Knowledge Base *(SPARQL)*
- Weka Tool
- Decision Tree Algorithm *(Best Performance, dealing with nominal/numeric features, easy visualised)*

**Accuracy For Training Set** 94% *(1503/2000)*

<table>
<thead>
<tr>
<th></th>
<th>Good</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Good</strong></td>
<td>Tp: 752</td>
<td>Fn: 49</td>
</tr>
<tr>
<td><strong>Bad</strong></td>
<td>Fp: 48</td>
<td>Tn: 751</td>
</tr>
</tbody>
</table>
And the Oscars goes to ..

Accuracy For Test Set

91.75%
Behind The Scenes

Decision Tree Diagram

Critics Negative Reviews
- <=0.3
  - Genre: Documentary +1 (352)
- >0.3
  - Language: English +1 (22)
    - Genre: Romance +1 (3)
      - #Bad Keywords -1 (8)
  - Language: German
    - #Actors Age <30 -1 (653/12)
    - Release Date: Weekend +1 (7)

Critics Positive Reviews
- <=0.4
  - #Good Keywords
- >0.4
  - #Good Keywords
    - Language: English
      - Genre: Romance +1 (3)
        - #Bad Keywords -1 (8)
    - Language: German
      - #Actors Age <30 -1 (653/12)
      - Release Date: Weekend +1 (7)
## Behind The Scenes

### Good Keywords
1. frustration  
2. melancholy  
3. very little dialogue  
4. looking out a window  
5. film director  
6. sin  
7. reference to Friedrich Nietzsche  
8. old friend  
9. moral ambiguity  
10. dressing

### Bad Keywords
1. critically bashed  
2. based on video game  
3. Taser  
4. pepper spray  
5. worst picture razzie winner  
6. spin off from video game  
7. physical comedy  
8. hung upside down  
9. female vampire  
10. dark heroine

### Common Keywords
1. weapon  
2. tourist  
3. spider  
4. sexual abuse  
5. Santa Claus  
6. rome italy  
7. queen  
8. mentor  
9. hollywood California  
10. black cop
Ranked Features

1) critics negative review
2) critics positive review
3) good keywords
4) bad keywords
5) country: USA
6) genre: Documentary
7) language: English
8) mostly Good Keywords
9) mostly Bad Keywords
10) MPAA: PG-13

Only 3 features from linked data in the top-10

- Linked Data is not enough alone
- DBpedia needs quality improvement and more interlinking