



Movies Dataset

Julián Ferreira
Tiago Antunes



Improvements

- Add ratings to our dataset
- Implement a movie recommender
- Implement the search of words related

Adding the field rating to our dataset



MoreLikeThis - Evaluation I

- **Boost 1:** Title⁶ Director² Cast¹ Genre² Plot¹
- **Boost 2:** Title⁴ Director³ Cast² Genre³ Plot¹

q = Title:Batman

Movie	Title	Director	Cast	Genre	Plot
Batman re-returns	Yes	No	No	Yes	Yes
Batman forever	Yes	No	No	Yes	Yes
Batman begins	Yes	No	No	Yes	Yes
The Lego Batman movie	Yes	No	No	Yes	Yes
Batman and Robin	Yes	No	No	No	Yes

Movie	Title	Director	Cast	Genre	Plot
Batman re-returns	Yes	No	No	Yes	Yes
The invisible Ray	No	No	No	No	No
Batman forever	Yes	No	No	Yes	
Dracula's Daughter	No	No	No	No	No
Batman begins	Yes	No	No	Yes	Yes

Evaluation II

q = Title:Shutter Island

Movie	Title	Director	Cast	Genre	Plot
The Ring	No	No	No	No	Yes
Captain Newman, M.D.	No	No	No	No	Yes
Elephant Song	No	No	No	Yes	Yes
Half Light	No	No	No	Yes	Yes
Ring 2	No	No	No	No	Yes

Table 14: Shutter Island results with Boost 1

Movie	Title	Director	Cast	Genre	Plot
The Mangler	No	No	Yes	No	Yes
The Ring	No	No	No	No	Yes
The Silence of the Lambs	No	No	Yes	Yes	Yes
Little Boy	No	No	Yes	No	No
Captain Newman, M.D.	No	No	No	No	Yes

Table 15: Shutter Island results with Boost 2



Evaluation III

q = Title:The Godfather

Movie	Title	Director	Cast	Genre	Plot
The Godfather Part II	Yes	Yes	Yes	Yes	Yes
The Godfather Part III	Yes	Yes	Yes	Yes	Yes
Jane Austen's Mafia!	No	No	No	No	Yes
Redline	No	No	No	No	Yes
Heavy Traffic	No	No	No	No	Yes

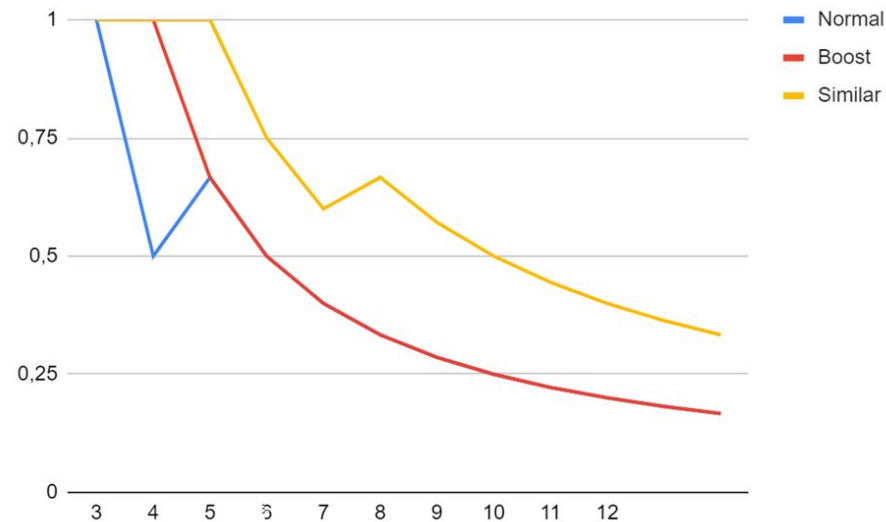
Table 13: The Godfather results with both Boosts

Search of “words related”

- Solr.SynonymGraphFilterFactory
 - synonym.txt

Option	Value
q	Plot:police AND Genre:action AND Release_Year:2000
qf (boost)	Plot:police^5 Genre:action^3 Releaseyear

Table 9: Arguments of query 4





Conclusions?

Possible future work?