

André Flores up201907001

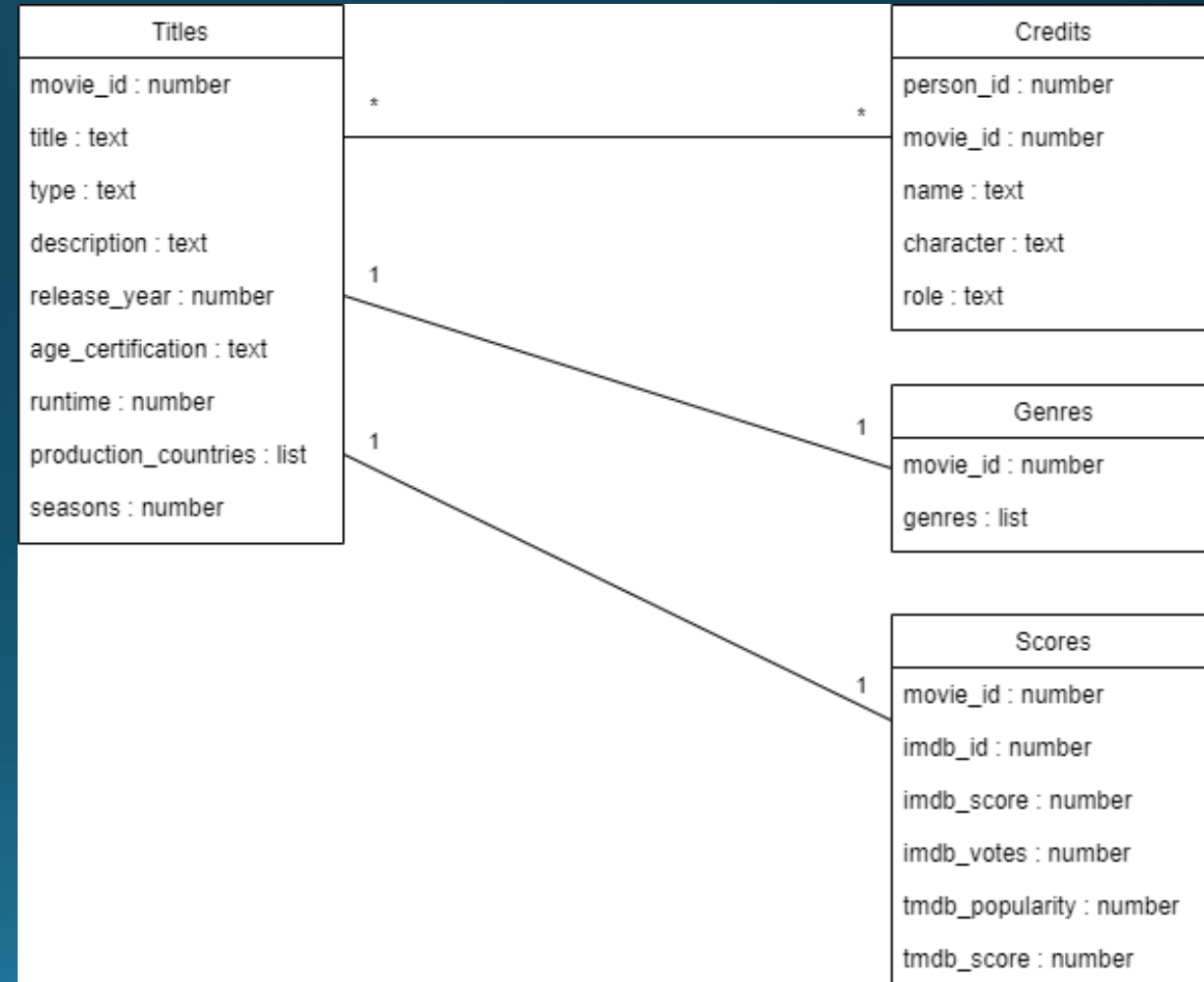
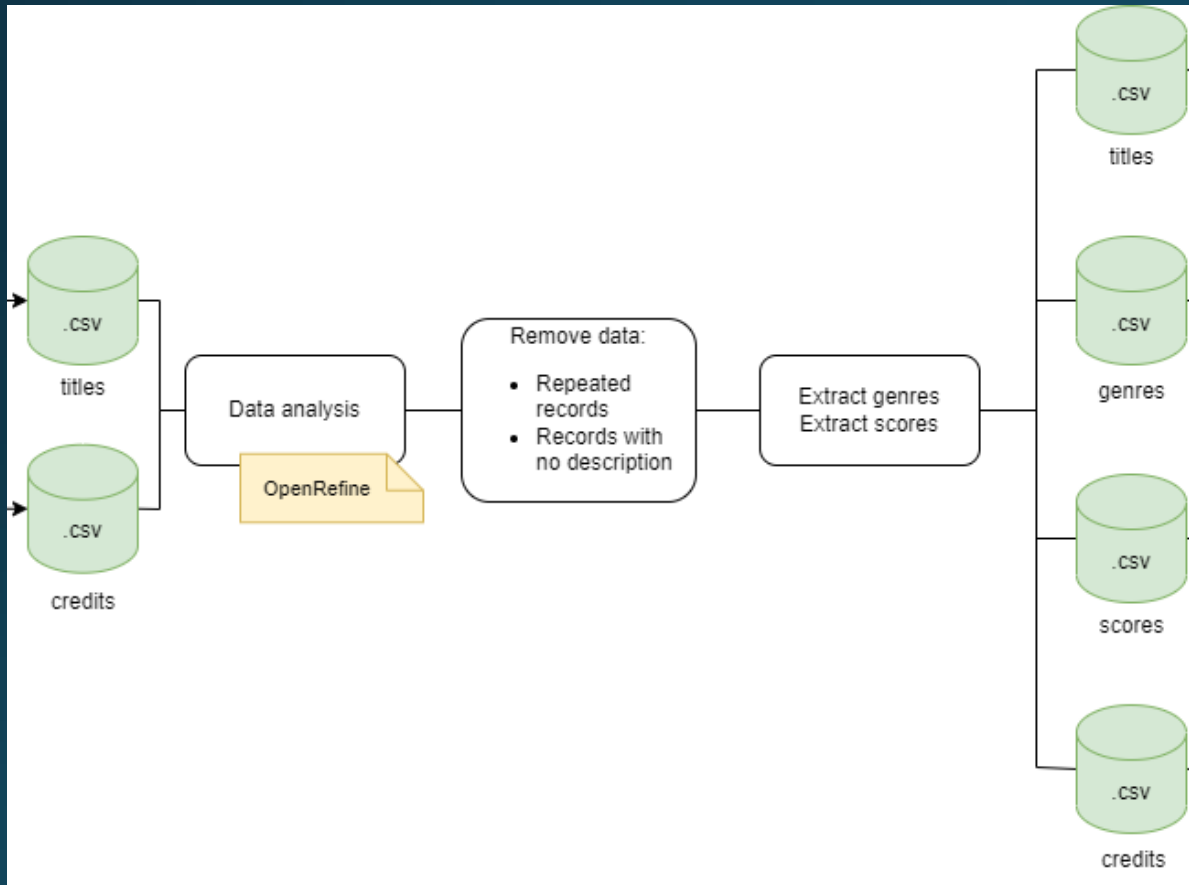
Guilherme Garrido up201905407

Luís Lucas up201904624

PRI 2022/23

RecFlix – Netflix titles recommendation

Previous work



Search scenarios

- What titles are on Netflix about cars?
- What titles are on Netflix related to Christmas and family?
- In what titles has Christina Hendricks participated since 2016?
- What titles are similar to “Fast & Furious Spy Racers”?
- Who participates in the movie “Bonnie and Clyde”?
- What actors are similar to George Clooney by genres of titles he’s participated in?

Collection

- Single **titles** collection to store our data
- Each document is either a **title** (name, type, description, etc.) or a **credit** (name, character, role, etc.)
- A **credit** is connected to its **title** using Solr's child documents functionality allowing us to search for one through the other
- Due to the way Solr works both **credits** and **titles** use the same schema

```
{
  "id": "tm288499",
  "title": "Roma",
  "type": "MOVIE",
  "description": "In 1970s Mexico City, two domestic wor",
  "release_year": 2018,
  "age_certification": "R",
  "runtime": 135,
  "genres": ["european",
    "drama"],
  "content_type": "parentDocument",
  "imdb_id": "tt6155172",
  "imdb_score": 7.7,
  "imdb_votes": 158129,
  "tmdb_score": 10.0,
  "tmdb_popularity": 12.745,
  "_version_": 1749469966138081280},
{
  "person_id": 3748,
  "name": "Robert De Niro",
  "character": "Travis Bickle",
  "role": "ACTOR",
  "id": "credit_tm84618_1",
  "_version_": 17494699907968327680
```

Indexing

- Titles:
 - title, description and genres
- Credits:
 - name and character
- Tokenizer:
 - ClassicTokenizer
- Filters:
 - ASCII Folding, Lower Case, English Minimal Stem, Classic, Porter Stem, Synonym Graph and Beider-Morse

Retrieval and evaluation #1

- Titles about cars

-	Normal	Boosted
Query string	cars	cars
Configuration	qf: title description	qf: title ³ description

TABLE II
QUERY 1 CONFIGURATION

-	Normal	Boosted
Recall	1	1
P@10	0.50	0.50
AP	0.90	0.90

TABLE XII
QUERY 1 EVALUATION

-	Normal		Boosted	
Rank	Title	R	Title	R
1	Fastest Car	X	Fastest Car	X
2	Car Masters: Rust to Riches	X	Car Masters: Rust to Riches	X
3	Comedians in Cars Getting Coffee	X	Comedians in Cars Getting Coffee	X
4	Top Gear	X	Top Gear	X
5	The Lies Within		The Lies Within	
6	Pui Pui Molcar		Pui Pui Molcar	
7	Okko's Inn		Okko's Inn	
8	Dirty Money		Dirty Money	
9	American Vandal		American Vandal	
10	Rust Valley Re- storers	X	Rust Valley Re- storers	X

TABLE III
QUERY 1 RESULTS

Retrieval and evaluation #3

- Titles Christina Hendricks participated in since 2016

-	Normal	Fuzzy
Query string	name:"Christina Hendricks"	name:"Christina~ "
Configuration	fq: release_year: [2016 TO *]	fq: release_year: [2016 TO *]

TABLE VI
QUERY 3 CONFIGURATION

-	Normal	Fuzzy
Recall	1	1
P@10	P@4=1	1
AP	1	1

TABLE XIV
QUERY 3 EVALUATION

-	Normal		Fuzzy	
Rank	Title	R	Title	R
1	Good Girls	X	War Dogs	X
2	Crooked House	X	Pottersville	X
3	Candy Jar	X	Wild District	X
4	Pottersville	X	Si saben cómo me pongo Pa qué me invitan	X
5			Paradise Lost	X
6			Little Boxes	X
7			The Dreamseller	X
8			Our Lovers	X
9			A Futile and Stupid Gesture	X
10			Valor	X

TABLE VII
QUERY 3 RESULTS

Retrieval and evaluation #4

- Titles similar to “Fast & Furious Spy Racers”

-	Normal	Boosted
Query string	A government agency recruits teen driver Tony Toretto and his thrill-seeking friends to infiltrate a criminal street racing circuit as undercover spies	A government agency recruits teen driver ² Tony Toretto and his thrill-seeking friends ² to infiltrate a criminal street racing ² <i>circuitas</i> undercover spies
Configuration	qf: title description	qf: title description

TABLE VIII
QUERY 4 CONFIGURATION

-	Normal	Boosted
Recall	1	1
P@10	0.40	0.40
AP	0.53	0.64

TABLE XV
QUERY 4 EVALUATION

-	Normal		Boosted	
Rank	Title	R	Title	R
1	Fast & Furious Spy Racers	X	Fast & Furious Spy Racers	X
2	The Last Laugh		Race: Bubba Wallace	
3	Guru		Rush	
4	Race: Bubba Wallace		Race	X
5	Turbo FAST		Race 2	X
6	Rush		Dragons: Race to the Edge	
7	Race 2	X	Motu Patlu Aur Khazaane Ki Race	
8	Race	X	Thomas & Friends: All Engines Go - Race for the Sodor Cup	
9	Hajwala 2: Mysterious Mission	X	Hajwala 2: Mysterious Mission	X
10	Dragons: Race to the Edge		Sugar Rush	

TABLE IX
QUERY 4 RESULTS

Evaluation

- Normal system MAP – 0.76
- Boosted system MAP – 0.83

Future work

- Improve precision and average precision
 - By improving document ranking
- Answer information need #6