PoliMovie: a Feature-based Dataset for Recommender Systems

Mona Nasery
Politecnico di Milano
Via G.Ponzio 34/5
20133, Milan, Italy
mona.nasery@polimi.it

Mehdi Elahi
Politecnico di Milano
Via G.Ponzio 34
20133, Milan, Italy
mehdi.elahi@polimi.it

Paolo Cremonesi
Politecnico di Milano
Via G.Ponzio 34
20133, Milan, Italy
paolo.cremonesi@polimi.it

ABSTRACT

Many recommender systems enrich traditional user-rating datasets with side information about features of items (e.g., genre, director, cast) and build user models as estimates of users' interests on features. These models are usually evaluated based on their ability to discover items relevant to the users. However, as public datasets do not contain the explicit opinions of users on features of items, user models are never evaluated in terms of matching between estimated and true preferences on features. In this paper we present ongoing work aimed at filling this gap by using crowdsourcing to collect a dataset (*PoliMovie*) which contains the explicit preferences of users on both items and their attributes. We present some preliminary results based on the preferences collected from 341 users. The results confirm our initial intuition that traditional user models based on implicit user preferences on attributes do not match well with the explicit opinions of users on attributes: only 11% of the implicitly derived models are in agreement with the explicit opinion of users on attributes. These results are convincing enough to justify a much more extensive crowdsourced collection of data.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

Keywords

Recommender systems, Dataset, User-profile, Crowdsourcing, Movie-attributes

1. INTRODUCTION

In the recent years, a large amount of research effort has been devoted to developing recommender systems exploiting information beyond the traditional user-rating matrix. Most of these systems include side information describing the features of items (e.g., in the movie domain categories like genre, director, cast) and build user models as estimates of users' interests on features [12, 10]. Incorporating side information is especially useful with the cold start problem, when there is little or no rating history about items. Traditionally, these enriched user models are evaluated based on their ability to provide relevant recommendations of items. A number of publicly available datasets, such as MovieLens and Netflix, provides a ground truth for the preferences of users on items. However, as publicly available datasets do not contain the explicit opinions of users on features of items, user models are never evaluated in terms of matching between estimated and true preferences on features.

In this paper we present ongoing work aimed at filling this gap by using crowdsourcing to collect a dataset which contains the explicit preferences of users on both items and their attributes. The dataset is focused on the movie domain, and contains a catalog of more than 100 thousands movies and over 1 million distinct attributes such as genre, director, actors, country of production (Table 1).

We have developed *PoliMovie*, an online movie survey application which has been integrated with the Microworkers[1] crowdsourcing service to collect the opinions of users on movies and their attributes.

We present some preliminary results based on the preferences collected from 341 users verified as reliable (based on a quality check performed at the end of the survey). These results confirm our initial intuition that traditional user models based on implicit user preferences on attributes do not match well with explicit user models in which users are explicitly elicited to provide their opinions on attributes.

We show that, in many cases, the favorite movies selected by a user have attributes (e.g., actors, directors, genre) totally different from the attributes selected as favorites by the same user. Hence, a user may like "The Dark Knight" movie without choosing "Action, Crime, Drama" as favorite genre, "Christopher Nolan" as favorite director, and "Christian Bale" as favorite actor.

We show that more than 40% of the implicitly derived user models have a 0% overlap, in terms of attributes, with respect to explicitly elicited preferences. Only 11% of the models have an overlap of 100% (complete match). These results are convincing enough to justify a much more extensive crowdsourced collection of data.

The rest of this paper is organized as follows. In section 2, related work is presented. In section 3, we present PoliMovie dataset and the tool we used to collect it. Section 4, provides some statistical analysis of the dataset and finally section 5 is the conclusion.

2. RELATED WORK

Recommender Systems create user profiles, most often based on user interactions with items, which are intended to represent user preferences for items, and to capture these preferences at the level of item features (attributes). Indeed, recommendations are often generated by matching up the

[1] `http://microworkers.com/`

features of the user profile (i.e., a structured representation of her interests) against the features of the item. In order to do this, content-based recommender systems usually build a Vector Space Model (VSM), where each item is represented by an n-dimensional vector. Each dimension in this model represents an attribute from the overall set of attributes used to describe the item. Using this model, the system computes a relevance score that represents the user's degree of interest toward that item [6]. This also allows the recommender systems to produce explanations to recommendations and to naturally solve the new item problem [2].

Han and Karypis in [4] proposed a feature-based recommendation system for domains where there is not enough historical data to measure the similarity between users or items. Their idea was "Users who bought items with these features, also bought items with features like this". In [13] Karen H. L. Tso et al. proposed a Synthetic Data generator to produce user-item and user/item-attribute datasets. In [8], Ning and Karypis developed four algorithms that incorporates feature representation (as side information) for *top-N* recommendation systems and showed that their approach achieved performance improvements by incorporating side information compared to methods which only relied on user-item purchase profile. Gantner et al. proposed an attribute-aware MF model in which they map user or item attributes to latent features of matrix factorization to make predictions for new items [3].

There is also work that used user-defined features (e.g., tags) to generate recommendations. Shilad Sen et al. in [9] designed an algorithm that predicts users' ratings for movies based on their inferred tag preferences. They defined a user's preference for a tag as the user's level of interest in movies. They used 5 datasets that include movie ratings, movie clicks, tags applied by users, tags searched by users and tag preferences. In [11], Börkur Sigurbjörnsson and Roelof van Zwol analyzed how users tag photos and what kind of tags they provide by using random snapshots of Flicker containing 52 million photos. They also presented four different recommendation strategies to suggest a set of tags to the user that could be added to the photo. Federico Durao and Peter Dolog in [1] presented a personalized tag-based recommender system to offer similar web pages based on tag similarity. Pasquale Lops et al. in [7] presented STaR tag recommender system which merges collaborative approach with content-based technique to face the cold start problem of collaborative technique.

Very recently, Chenyi Zhang et al. in [14] proposed a feature-centric approach based on the principle that "Liking same features (as other users) leads to liking more same features". They predict ratings for item attributes using collaborative filtering technique and then used those feature ratings to predict item ratings. They used 4 tag datasets: *Delicious, Lastfm, DBLP,* and *Movielens.* Their idea is closest to what we are proposing. However, they used tag datasets and treated tags as attributes of items while tag and features are two different concepts. (Tags are user-defined terms which contain noise while attributes are defined by experts.) Moreover, tagging an item does not always mean that the user liked that item, for example, if the user assigns the tags "Paris", "Eiffel" to a photo of the "Eiffel tower", it does not necessarily means that she liked this place.

3. DATASET COLLECTION

Table 1: Characteristics of the primary dataset

Data	Number	info
Movies	100K	genre, year, number of ratings, average rating, country of production, full cast and crews
Stars	111K	name, year of birth, country of birth
Full cast	1M	name, year of birth, country of birth
Directors	40K	name, country of birth

In this section, we describe our ongoing work devoted to collecting explicits opinion from users on both items and attributes of movies. We fetched a list of 100k most popular movies from the IMDb[2] website (most popular according to the number of ratings), as well as some additional information on cast and directors. Table 1 shows some statistics about this primary dataset.

We have built a survey application called PoliMovie, which asks participants to answer 14 questions (all mandatory). The survey allows the system to collect demographics and movie preference data. Table 2 shows all the questions we ask in the PoliMovie survey. Answering all the questions in the survey is required, but there is possibility for users to select "No preference" or "leave blank" option for some of the questions such as demographic ones. Also, there are some questions which have minimum required answers such as "Which is your favorite movies?" for which users have to select at least 10 movies. Also, users have the possibility to either login with their email address or with their Facebook account.

Initially, we invited our friends and colleagues to complete the survey and provide some feedback. In our first experiment, PoliMovie survey contained 27 questions (including 7 demographic questions), asking users their opinion about various features of movies. However, later, we shortened the survey and kept only questions that we think are the most important ones. For instance, in the first version we asked users about their favorite movie writers or producers, but after collecting the primary users feedback and also consulting many friends and colleagues, we found that most users do not know many writers or producers or they do not have any preference about these movie features.

After the slight modification, we integrated the survey application to a crowdsourcing service, called "Microworkers", and had the survey completed by 341 users. As mentioned before, this data collection is still ongoing work and we will continue collecting data from more users.

Figure 1 shows a sample screenshot of the survey. We have adopted a number of best practice in order to simplify the survey as much as possible. Users can search for movies, actors and directors by any of the attributes. For instance, users can search for a movie by title, actors or director. We used auto-completion on text boxes to avoid typing errors and speedup the search process. The users are also able to save the form, quit and continue later starting from the last saved session.

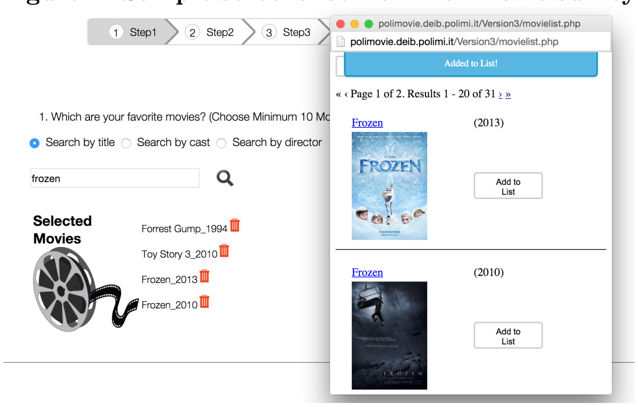
During the process of collecting the data using the crowdsourcing service, we have encountered a number of issues:

[2] http://www.imdb.com

Table 2: Questions included in PoliMovie survey

Question Type	Question		Possible Answers
	Id	question	
Movie questions	1	which are your favorite movies?	at least 10 movies should be selected by user
	2	who are your favorite actors?	at least 1 cast should be selected by user
	3	what are your favorite genres of movies?	23 genres, there is also the possibility to select “no preference”
	4	do you have any preference for gender of main actors of a movie?	male, female, no preference
	5	which is your favorite nationality for main actors of movies?	a set of 186 countries, there is also “leave blank” option
	6	where do you prefer for movie production country?	a set of 186 countries, there is also “leave blank” option
	7	who are your favorite directors?	at least 1 director should be selected by user
	8	which are the most important movie features for you when you want to select a movie to watch?	rating, actor, director, genre, production year, no difference. (maximum allowed: 2)
	9	which decade do you prefer to choose a movie from it to watch?	40’s, 50’s, 60’s, 70’s, 80’s, 90’s, 2000-2010, 2010 till now. there is also “no difference” option
	10	do you like to watch remakes of movies you liked in the past?	yes, no, no difference
Demographic	11	what is your gender?	male, female, would rather not say
	12	what is your age?	under 18, 18-30, 31-40, 41-60, over 60, would rather not say

Figure 1: Sample screenshot from PoliMovie survey



the first issue concerns the verification of the data quality, i.e., how to verify the reliability of the workers and to ensure that users did not answer their questions randomly. To tackle this issue, first, we described very clearly the task to workers through a user-friendly and an easy-to-use interface. Second, we used Ids of the workers as a unique identifier, in order to prevent them from re-submitting the form again.

Finally, we added a page at the end of the survey, where we repeated two of the questions (question 2 and 8). The answers to these *reliability* questions were compared with the previous answers to the same questions, in order to filter out unreliable users (e.g., workers with a mismatch between answers to reliability questions and answers to normal questions). To avoid cheating, we disabled navigation to previous pages, so that the workers could not return to previous pages to see their answers. We manually re-checked answers from unreliable workers. We made our best effort to treat the workers fairly and we rejected only those that we were sure they answered the questions randomly. We note that, in the analysis presented in this paper, we only considered the reliable answers.

4. PRELIMINARY RESULTS

This section presents preliminary results of some statistical analysis performed on this initial collection of data. Up to now, 1077 movies, 265 casts, 193 directors, and all genres have been selected as favorite at least by one user.

We have listed, in Table 3, the most important features, the users consider when making decision what to watch. As it can be seen, the most important features are “Cast” and the “Rating” of the movies. This is while the “Genre” and the “Year of Production” play the least important role in decision making for choosing which movie to watch.

Table 3: Users opinion on movie features

Movie feature	Level of importance
Casts	48.5 %
Rating	48.5 %
Directors	33.6 %
Genres	32.7 %
Year of production	7.0 %

We have computed the popularity scores of actors, directors and genres (i) based on explicit preferences of users

Table 4: Top 20 most popular actors. Names with * are present in one list only.

	Explicitly selected by users		Based on movies selected by users	
	Name	%pop	Name	%pop
1	Leonardo DiCaprio	16.2	Tom Cruise	14.7
2	Johnny Depp	10.5	Robert De Niro	14.2
3	Al Pacino	9.2	Leonardo DiCaprio	14.2
4	Robert De Niro	9.2	Brad Pitt	14.2
5	Brad Pitt	7.8	Johnny Depp	12.2
6	Tom Hanks	7.4	Tom Hanks	10.7
7	Scarlett Johansson	6.1	Jim Carrey	10.7
8	Jim Carrey	5.7	Al Pacino	10.7
9	Angelina Jolie*	5.7	Harrison Ford*	9.6
10	Natalie Portman*	5.2	Jessica Alba*	9.6
11	Tom Cruise	5.2	Scarlett Johansson	9.6
12	Maryl Streep*	4.8	Edward Norton	8.6
13	Arnold Schwarzenegger*	4.8	Robin Williams	8.6
14	Will Smith	4.3	Jason Statham*	8.6
15	Robin Williams	4.3	Daniel Radcliffe*	8.1
16	Jolia Roberts*	4.3	Rupert Grint*	8.1
17	Christoph Waltz*	3.9	Ben Stiller*	8.1
18	Robert Downey Jr.	3.9	Will Smith	7.6
19	Edward Norton	3.5	Robert Downey Jr.	7.6
20	John Travolta*	3.5	Morgan Freeman	7.6

Table 5: Top 20 most popular directors. Names with * are present in one list only.

	Explicitly selected by users		Based on movies selected by users	
	Name	%pop	Name	%pop
1	Christopher Nolan	16.4	Steven Spielberg	3.9
2	Quentin Tarantino	14.5	Quentin Tarantino	2.7
3	Steven Spielberg	13.6	Martin Scorsese	2.2
4	Martin Scorsese	10.7	Christopher Nolan	1.8
5	Woody Allen	10.3	Peter Jackson	1.4
6	James Cameron	8.9	Woody Allen	1.4
7	Clint Eastwood*	6.5	Mani Ratnam*	1.3
8	Peter Jackson	6.5	Michael Bay*	1.3
9	David Fincher	5.6	Stanley Kubrick	1.2
10	Stanley Kubrick	4.6	Robert Rodriguez*	1.2
11	George Lucas	4.6	Ridley Scott	1.2
12	Ridley Scott	4.6	David Fincher	1.2
13	David Lynch*	4.2	Tim Burton	1.1
14	Alfred Hitchcock*	3.7	Brian De Palma*	1.1
15	Tim Burton	3.2	George Lucas	1.1
16	Ron Howard*	3.2	Brett Ratner*	1.1
17	Joss Whedon*	2.8	James Cameron	1.1
18	Francis Ford Coppola	2.8	David Yates*	0.9
19	Roman Polanski	2.3	Roman Polanski	0.9
20	Joel Coen*	2.3	Chris Columnbus*	0.9

given directly to these features, and (ii) based on implicitly computing the scores according to the preferences of users given to the movies. Table 4 shows top 20 actors sorted according to their popularity scores, based on explicit or implicit user preferences. The column *pop* shows the percentage of popularity for each actor. As it can be seen, these two lists are actually different. There are some actors in the first list (based on explicit preferences) that are not present in the second (based on implicit preferences) or vice versa. These mismatches are shown with * beside the names. Some of the actors that appear in the first list, while missing in the second one are Natalie Portman, Maryl Streep, Angelina Jolie, Jolia Roberts. Some of the the actors that appear in the second list but not the first one are Harrison Ford, Jessica Alba, Daniel Radclie, Rupert Grint. Moreover, The popular actors who appear in both lists, do not locate in identical positions either in the ranking position or the movies. For instance, Tom Cruise has the 10th position in the first method while he is considered as the most popular actor in the second method.

Similar results can be seen for directors, as shown in table 5 and, for genres in table 6. Again the explicit preferences of users given for features of movies differ from their preferences inferred indirectly from their preferences given to movies. This is actually very interesting for genres (see table 6). We have noticed the most favorite movie genre, based on users' favorite movies, has to be "Drama". However, the most popular genre, selected indirectly by users, is totally the opposite genre, i.e. "Comedy". Indeed, the Drama genre appeared in rank position 6 in the second list, which shows the lower importance of this genre from the users' point of views.

In order to better show the difference between a user implicit profile based on the movies that users have selected

Table 6: Sorted top 20 most popular genres

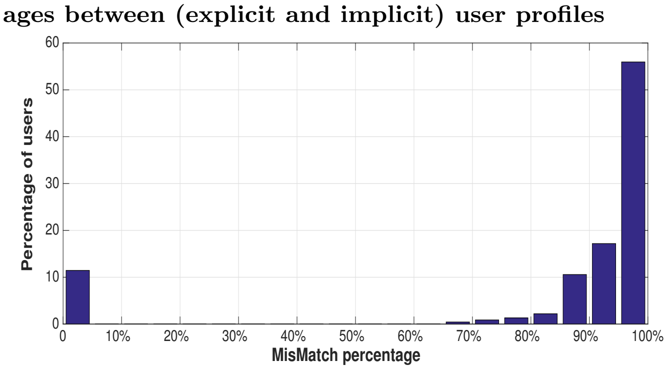
	Explicitly selected by users		Based on movies selected by users	
	Genre	%Popularity	Genre	%Popularity
1	Comedy	60.1	Drama	44.6
2	Action	55.1	Action	29.1
3	Adventure	45.6	Comedy	27.8
4	Sci-Fi	35.2	Adventure	20.9
5	Thriller	33.1	Crime	18.0
6	Drama	33.1	Thriller	17.7
7	Mystery	31.1	Romance	13.1
8	Romance	29.8	Sci-Fi	12.1
9	Crime	27.3	Horror	8.6
10	Animation	26.5	Mystery	8.5
11	Fantasy	24.8	Fantasy	8.1
12	Horror	24.4	Animation	7.0
13	History	19.0	Family	5.8
14	Documentary	17.8	Biography	5.6
15	War	16.1	War	2.4
16	Family	15.3	Sport	2.0
17	Biography	11.2	Musical	1.9
18	Sport	9.1	History	1.8
19	Western	8.7	Western	1.6
20	Musical	8.2	Music	1.3

as their favorites and the explicit user profile based on the direct preference of users on features of movies, we selected two users as a sample and compared their implicit and explicit user profiles. Table 7 and table 8 show the results for user_A and user_B respectively. The first column in both tables show the movie features we used to compare two profiles of the users. The second column shows the popular names (e.g., actors, directors or genres) based on the movies that user selected while the third column shows the names directly selected by user as her favorite feature (i.e., favorite actors/directors or genres). As can be seen in table 7, the implicit user profile based on the movies that user_A selected as her favorites have no overlap in terms of actors and directors with her explicit profile based on her direct preferences on these attributes.

Moreover, we compared the dissimilarity or mismatch between the two user models based on explicit and implicit user preferences. We measured the dissimilarity by using the Jaccard distance [5]. We have observed interesting results shown in Figure 2 where x axis is the percentage of mismatch, and y axis is percentage of users with certain mismatch between their explicit and implicit profiles. Our observations show that about 56% of users have at least 95% mismatch between their implicit and explicit user profiles and almost 86% users have at least 80% mismatch between their two types of profiles. Moreover, only about 11% of users have 0% mismatch, i.e, their two user profiles match 100% with each other. These observations confirm that user models based on implicit user preferences on attributes do not match well with explicit user models in which users are explicitly elicited to provide their opinions on attributes.

It worth noting that, in most of the cases, the size of implicit user profile is larger than the explicit ones. This is because in our crowd sourcing based data collection, we added a condition that each user has to select at least 10 movies. So, for example, a user may select only 2 actors as her favorite, and in her explicit profile, she will have only 2 actors. However, if she selects 10 movies as favorite and if each movie has approximately 2 stars, then the number of actors in her implicit profile will be 20.

Figure 2: Histogram showing the mismatch percentages between (explicit and implicit) user profiles



5. CONCLUSION

In this paper, we present ongoing work of data collection. The collected data would be included in a dataset called *PoliMovie*. The PoliMovie dataset is publicly available [3]. In contrast to the currently available datasets, it will included not only the preferences of users given to movies, but also the preferences of the users given for features (attributes) of movies, such as cast, genre and director. Such dataset can be very useful for researchers and the practitioners in the community of Recommender Systems since it will allow them to benchmark their feature-based recommendation algorithms using our dataset. To the knowledge of the authors, no such dataset has been so far made.

[3] through the link: `http://recsys.deib.polimi.it/polimovie-dataset/`

6. REFERENCES

[1] F. Durao and P. Dolog. A personalized tag-based recommendation in social web systems. *arXiv preprint arXiv:1203.0332*, 2012.

[2] M. Elahi, F. Ricci, and N. Rubens. Active learning strategies for rating elicitation in collaborative filtering: a system-wide perspective. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1):13, 2013.

[3] Z. Gantner, L. Drumond, C. Freudenthaler, S. Rendle, and L. Schmidt-Thieme. Learning attribute-to-feature mappings for cold-start recommendations. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pages 176–185. IEEE, 2010.

[4] E.-H. S. Han and G. Karypis. Feature-based recommendation system. In *Proceedings of the 14th ACM international conference on Information and knowledge management*, pages 446–452. ACM, 2005.

[5] M. Levandowsky and D. Winter. Distance between sets. *Nature*, 234(5323):34–35, 1971.

[6] P. Lops, M. De Gemmis, and G. Semeraro. Content-based recommender systems: State of the art and trends. In *Recommender systems handbook*, pages 73–105. Springer, 2011.

[7] P. Lops, M. De Gemmis, G. Semeraro, C. Musto, and F. Narducci. Content-based and collaborative techniques for tag recommendation: an empirical evaluation. *Journal of Intelligent Information Systems*, 40(1):41–61, 2013.

[8] X. Ning and G. Karypis. Sparse linear methods with side information for top-n recommendations. In *Proceedings of the sixth ACM conference on Recommender systems*, pages 155–162. ACM, 2012.

[9] S. Sen, J. Vig, and J. Riedl. Tagommenders: connecting users to items through tags. In *Proceedings of the 18th international conference on World wide web*, pages 671–680. ACM, 2009.

[10] Y. Shi, M. Larson, and A. Hanjalic. Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges. *ACM Comput. Surv.*, 47(1):3:1–3:45, May 2014.

[11] B. Sigurbjörnsson and R. Van Zwol. Flickr tag recommendation based on collective knowledge. In *Proceedings of the 17th international conference on World Wide Web*, pages 327–336. ACM, 2008.

[12] P. Symeonidis, A. Nanopoulos, and Y. Manolopoulos. Feature-weighted user model for recommender systems. In *Proceedings of the 11th International Conference on User Modeling*, UM '07, pages 97–106, Berlin, Heidelberg, 2007. Springer-Verlag.

[13] K. H. Tso and L. Schmidt-Thieme. Evaluation of attribute-aware recommender system algorithms on data with varying characteristics. In *Advances in Knowledge Discovery and Data Mining*, pages 831–840. Springer, 2006.

[14] C. Zhang, K. Wang, E.-p. Lim, Q. Xu, J. Sun, and H. Yu. Are features equally representative? a feature-centric recommendation. 2015.

Table 7: Implicit VS Explicit user profiles for a sample user_A. Names shown in bold exist in both user models

Feature	User Model based on movies selected by user	User Model based on explicit user preferences on features
Actor	Josh Zuckerman Sean Faris Miles Teller Simon Pegg Idris Elba Mike Myers Jake Gyllenhaal Matthew McConaughey Robert Downey Jr. Mark Ivanir Amanda Crew Amber Heard Melissa Benoist Martin Freeman Rinko Kikuchi Cameron Diaz Mary McDonnell Jessica Chastain Scarlett Johansson Mykel Shannon,Jenkins Clark Duke Djimon Hounsou J.K. Simmons Nick Frost Charlie Hunnam Eddie Murphy Jena Malone Chris Evans Scott Adkins Anne Hathaway	Ryan Reynolds Adam Sandler Hugh Jackman Christian Bale
Director	Christopher Nolan Damien Chazelle Jeff Wadlow Sean Anders Isaac Florentine Richard Kelly Edgar Wright Andrew Adamson Vicky Jenson Joss Whedon Guillermo del Toro	Michael Bay James Wan
Genre	Adventure Action Drama **Comedy** **Sci-Fi** Romance Crime Animation Sport Mystery Music	**Comedy** **Sci-Fi**

Table 8: Implicit VS Explicit user profiles for a sample user_B. Names in bold exist in both user models

Feature	User Model based on movies selected by user	User Model based on explicit user preferences on features
Actor	Charles Vanel Nelly Pappaert Dean Stockwell Peter Finch Lauren Bacall **Leslie Nielsen** Gary Sinise Carrie-Anne Moss Samuel L. Jackson Paul Meurisse Jacqueline Poelvoorde-Pappaert Nastassja Kinski William Holden Paul Bettany Julie Hagerty Robin Wright Laurence Fishburne Uma Thurman Brigitte Bardot Benoît Poelvoorde Harry Dean Stanton Faye Dunaway **Nicole Kidman** Robert Hays Tom Hanks Keanu Reeves John Travolta	Leonardo DiCaprio Scarlett Johansson Morgan Freeman **Leslie Nielsen** **Nicole Kidman**
Director	Lana Wachowski Rémy Belvaux Robert Zemeckis André Bonzel Benoît Poelvoorde Jim Abrahams Henri-Georges Clouzot David Zucker Jerry Zucker Lars von Trier Quentin Tarantino **Sidney Lumet** Andy Wachowski Wim Wenders	Martin Scorsese Frank Capra **Sidney Lumet** Paul Thomas Anderson David Fincher
Genre	**Drama** **Crime** **Comedy** **Action** **Sci-Fi** **Thriller** Romance	**Action** Adventure **Comedy** **Drama** **Crime** Western **Sci-Fi** **Thriller**