Improving Users' Demographic Prediction via the Videos They Talk about

Yuan Wang, Yang Xiao, Chao Ma, and Zhen Xiao
Peking University, China
EMNLP2016, 3 November 2016

Table of Contents

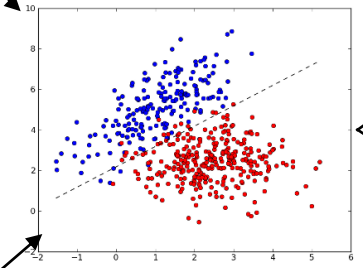
- Introduction
- Data
- Indirect Relationships between Users and Videos
- Evaluation
- Conclusion

Introduction

Introduction

Love, pretty, work, …

Sports, Financial, …, Car



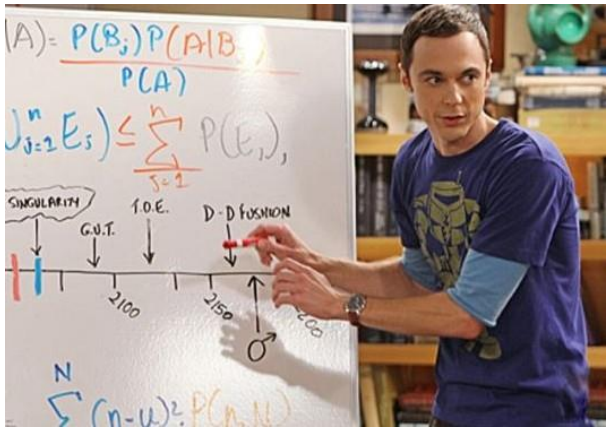
Female

Male

Introduction

- Mean Girls, Pretty Woman, The Devil Wears Prada
- House of Cards, Mission Impossible, NBA

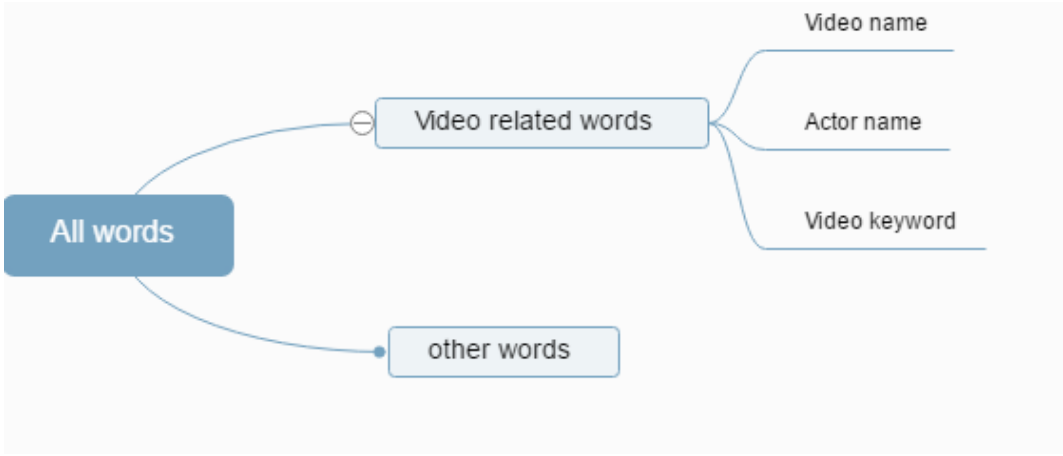
Introduction



"Will the Big Bang Theory last into the next century?"

"Sheldon is so cool, I love him!"

"Jim Parsons was nominated for another Emmy Award"



Introduction

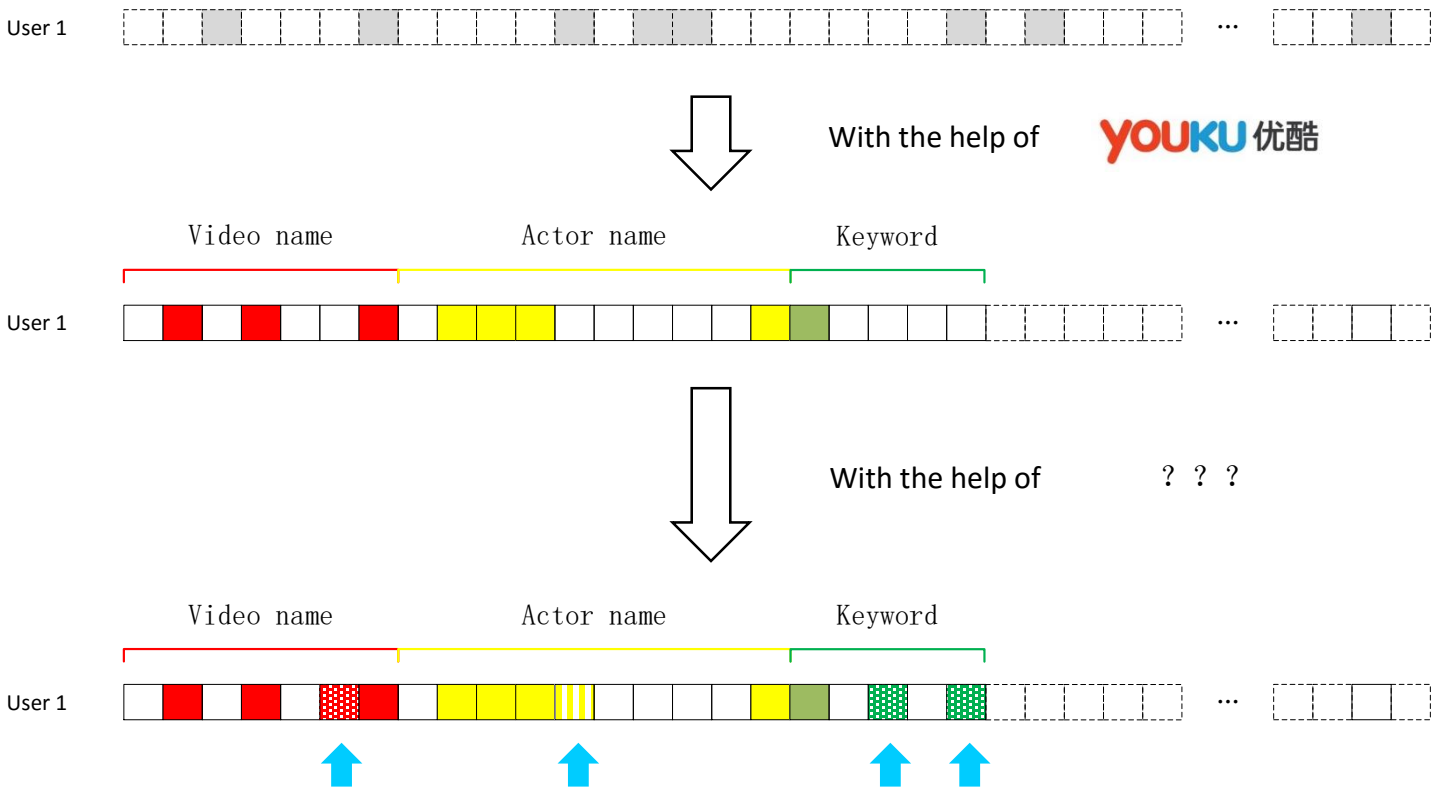
YOUKU 优酷

iQIYI 爱奇艺

腾讯视频 V.QQ.COM

搜狐视频

Introduction

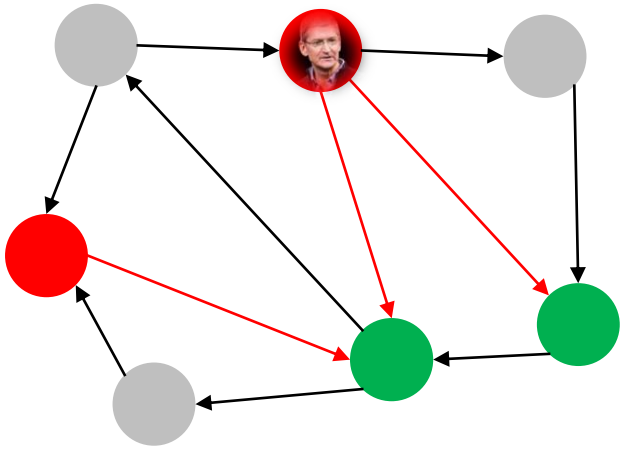


Data

Normal User

Verified User

User followed by Verified User



Attribute	Completion Rate	Categories
Gender	95.019%	Male, Female
Age	18.604%	Teenage (<18), Youngster (18-24), Young (25-34), Mid-age(>34)
Education BG	17.443%	University, Non-University
Marital Status	2.203%	Single, Non-Single

Table 1: Demographic attributes and corresponding categories

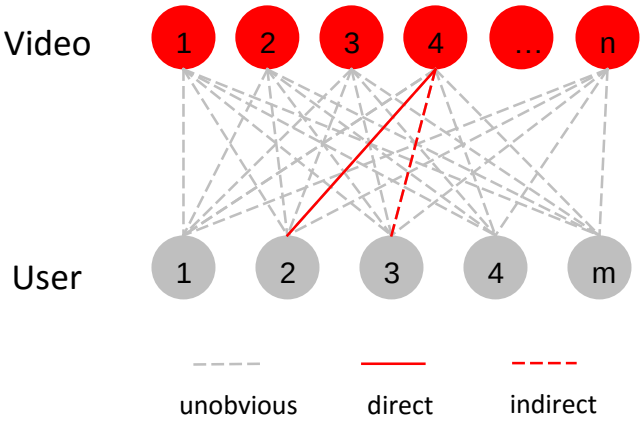
Data

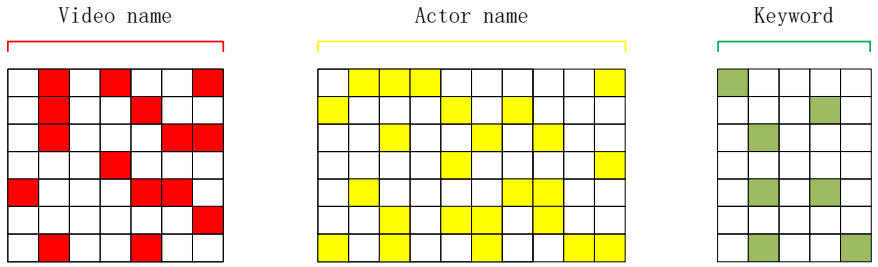
	Video	**Actor**	**Keyword**
Variety show	344	1007	2925
Movie	306	741	2049
TV	197	515	1302
Total	**847**	1422	4094

Table 2: Statics of video relevant information (There is an overlap between the three collections of actors and keywords.)

Discover Indirect Relationships

- Unobvious relationship
  - N*M pairs
- Direct relationship
  - User 2, “Will the Big Bang Theory last into the next century?”
- Indirect relationship
  - User 3 posts , “Sheldon is so cool, I love him!”





Discover Indirect Relationships

Step 1

$$P(v_n) = \frac{num(\text{users } watched\ the\ n_{th}\ video)}{num(users)}$$

$$P(w_{ni}|v_n) = \frac{num(\text{users } watched\ the\ n_{th}\ video\ and\ mentioned\ the\ ni_{th}\ keyword)}{num(\text{users } watched\ the\ n_{th}\ video)}$$

$$P(a_{nj}|v_n) = \frac{num(\text{users } watched\ the\ n_{th}\ video\ and\ mentioned\ the\ nj_{th}\ actor)}{num(\text{users } watched\ the\ n_{th}\ video)}$$

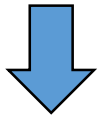
Step 2

$$P(v_n|W_m, A_k) = \frac{P(W_m, A_k|v_n) * P(v_n)}{P(W_m, A_k)}$$

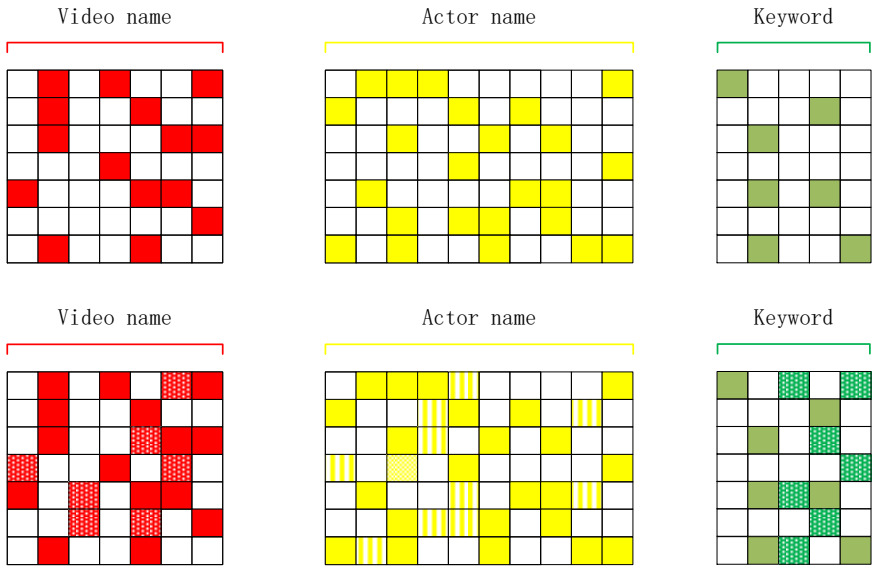
$$= \frac{\prod_{w_{ni} \in W_m} P(w_{ni}|v_n) * \prod_{a_{nj} \in A_k} P(a_{nj}|v_n) * P(v_n)}{P(W_m, A_k)}$$

Discover Indirect Relationships

Direct

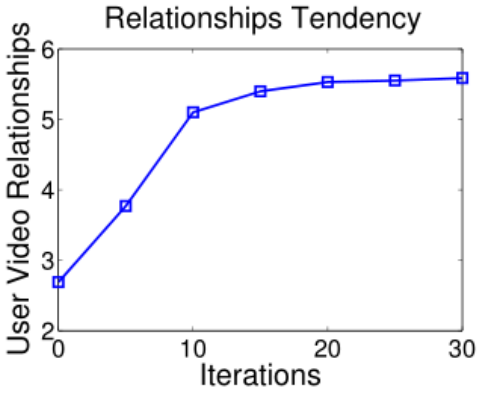


Direct + Indirect (more denser)



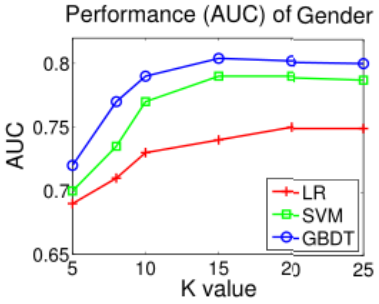
Two Baseline Model

Two Indirect Relationship Based Model

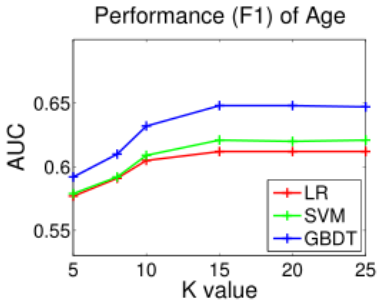


Discriminant Model

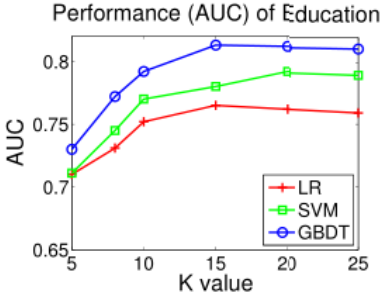
- Matrix Factorization[1], K=20
- LR[2], SVM[2], GBDT[3]



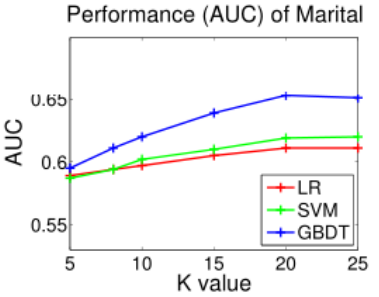
(a) AUC of Gender



(b) F1 of Age



(c) AUC of Education BG



(d) AUC of Marital Status

1 libFFM
2 liblinear
3 XGBoost

Generative Model

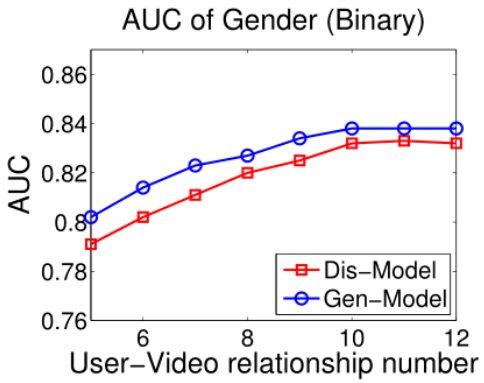
- Calculate video demographic tendency
- Calculate user demographic attribute
- Smooth the result

Evaluation

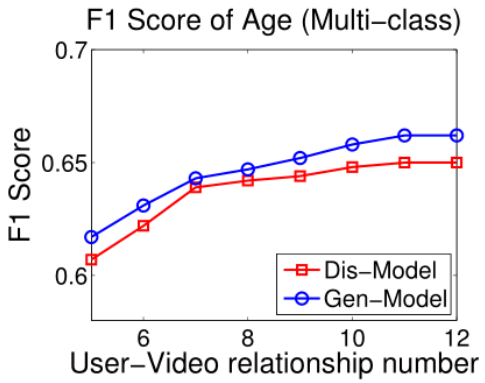
		Precision	Recall	F1	AUC
Gender 1	Dis-Baseline	0.720	0.714	0.717	0.730
	Dis-Model	0.786	0.779	0.783	**0.812** ↑ 11.2%
	Gen-Baseline	0.701	0.687	0.694	0.707
	Gen-Model	0.799	0.802	0.801	**0.825** ↑ 16.7%
Age 2	Dis-Baseline	0.569	0.541	0.554	*
	Dis-Model	0.642	0.653	**0.648** ↑ 16.8%	*
	Gen-Baseline	0.529	0.504	0.516	*
	Gen-Model	0.663	0.645	**0.654** ↑ 26.7%	*
Education BG 3	Dis-Baseline	0.707	0.716	0.711	0.730
	Dis-Model	0.788	0.801	0.795	**0.809** ↑ 11.1%
	Gen-Baseline	0.680	0.659	0.669	0.690
	Gen-Model	0.790	0.808	0.799	**0.812** ↑ 17.7%
Marital Status 4	Dis-Baseline	0.565	0.549	0.557	0.571
	Dis-Model	0.657	0.640	0.648	**0.659** ↑ 15.4%
	Gen-Baseline	0.572	0.550	0.560	0.581
	Gen-Model	0.682	0.691	0.687	**0.696** ↑ 19.8%

Table 3: Prediction accuracy based on users' video describing words. Classes have been balanced.

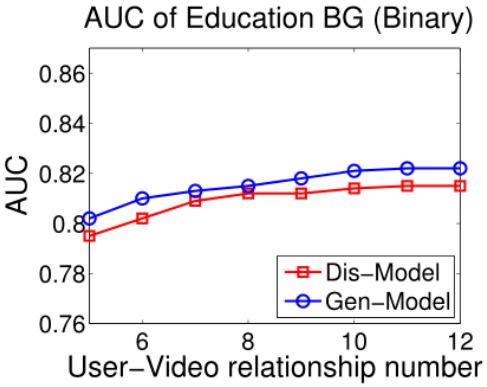
Evaluation



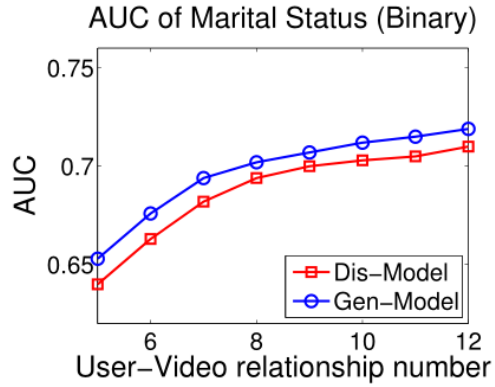
(a) AUC of Gender



(b) F1 Score of Age



(c) AUC of Education BG



(d) AUC of Marital Status

Evaluation

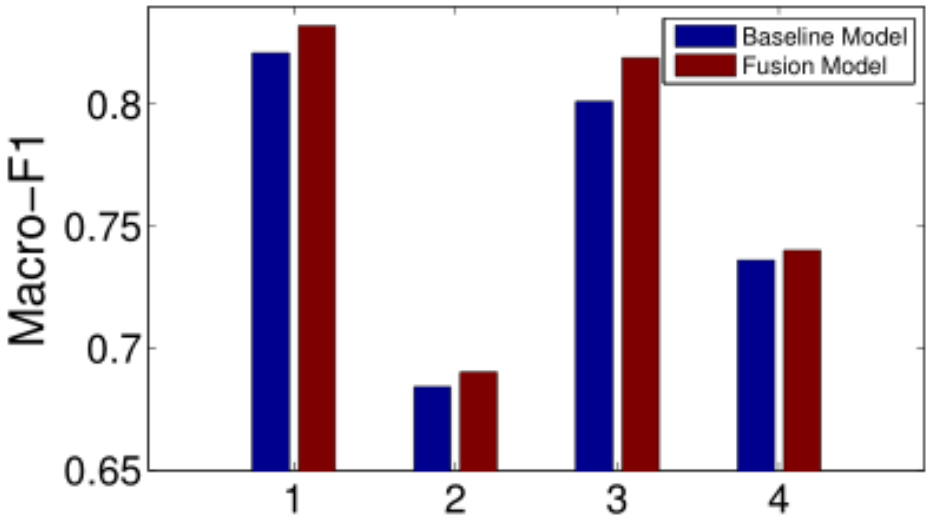


Figure 5: Results of Fusion Model evaluation (Macro-F1).

Conclusion

- Our motivation is that user's video related behavior is usually under-utilized on demographic prediction tasks.
- With the help of third-party video sites, we detect the direct and indirect relationships between users and video describing words, and demonstrate this effort can improve the accuracy of users' demographic predictions.
- To our knowledge, this is the first work which explores demographic prediction by fully using users' video describing words.
- This framework has good scalability and can be applied on other concrete features, such as user's book reading behaviors and music listening behaviors.

Thanks!