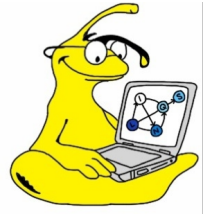
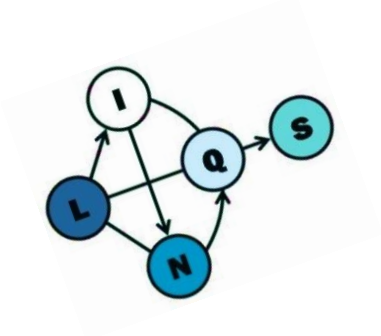
Collective Classification

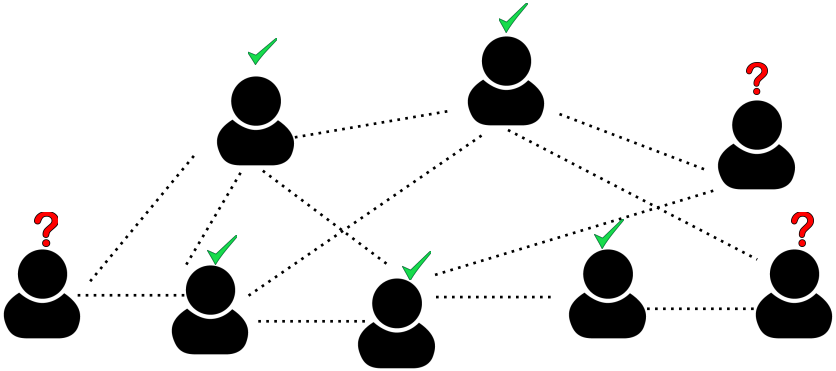
Eriq Augustine and Golnoosh Farnadi
UC Santa Cruz
MLTrain 2018



psl.linqs.org
github.com/linqs/psl

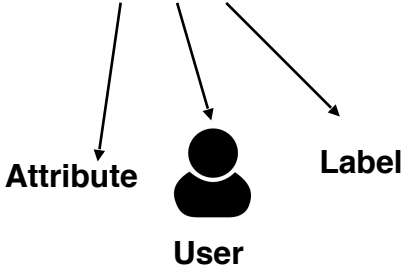


What is Collective Classification?

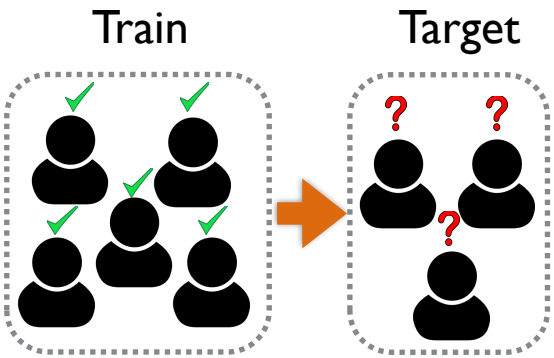


Attribute "**A**" of User "**U**" **is** Labeled "**L**"

```
Is(A,U,L)
```



What is Collective Classification?



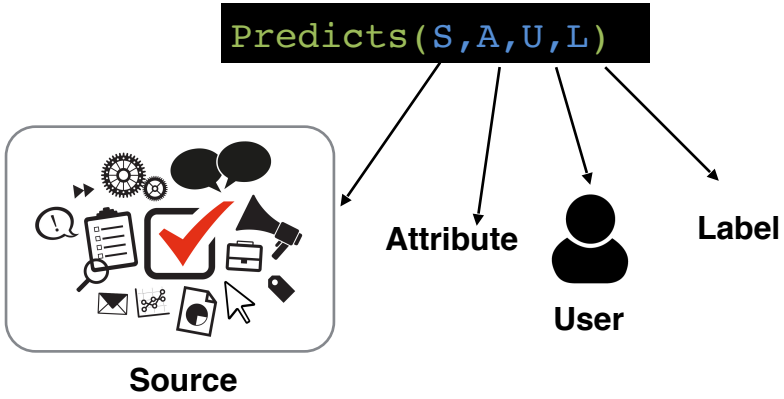
example:

```
Is(Gender,Alice,Female)
```

```
Is(Age,Bob,Young)
```

```
Is(Personality,Carol,Introvert)
```

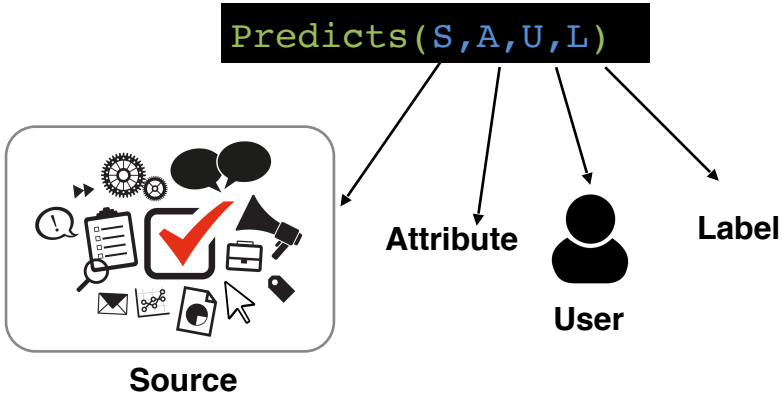
Local Predictor Rule



Source "**S**" **Predicts** Attribute "**A**" of User "**U**" **is** Labeled "**L**"

```
Predicts(S,A,U,L) -> Is(A,U,L)
```

Local Predictor Rule

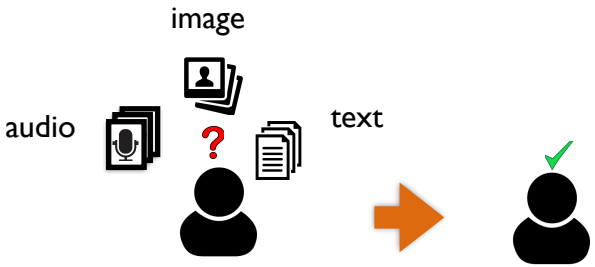


We collect training data to learn a predictive model, e.g. logistic regression

T_U	L_U
	0
⋮	⋮
	1

$$P(L|T)$$

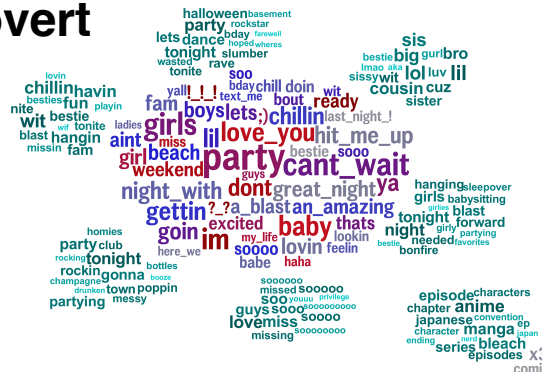
Local Predictor Rule



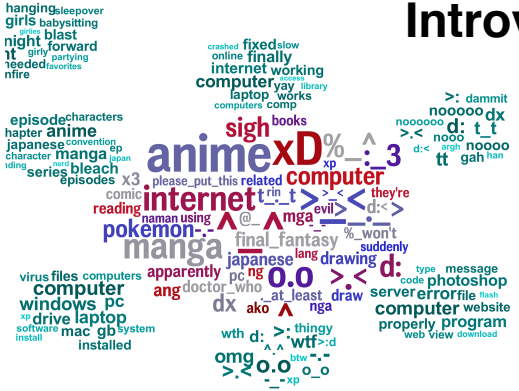
example:

```
Predicts(Txt,Personality,Alice,Int)
                    -> Is(Personality,Alice,Int)
```

Extrovert

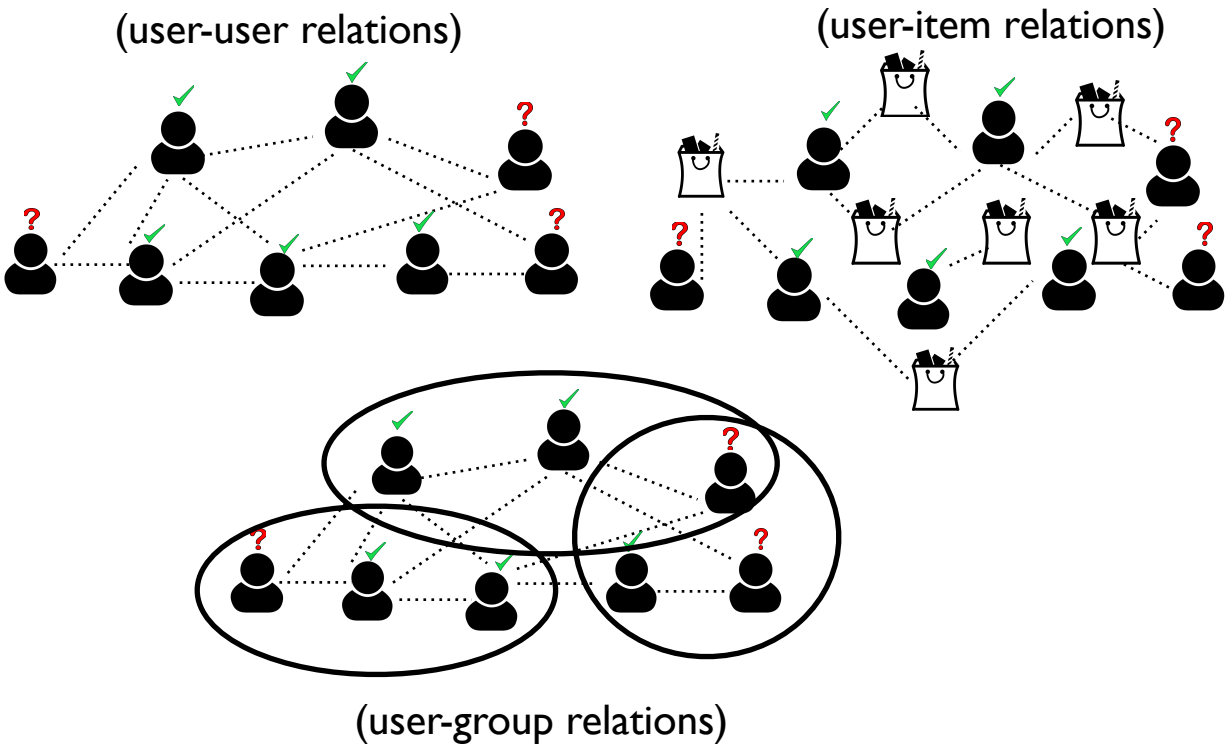


Introvert

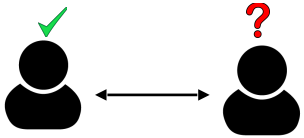


Park G, Schwartz HA, Eichstaedt JC, Kern ML, Kosinski M, Stillwell DJ, Ungar LH, & Seligman ME (2014). Automatic Personality Assessment Through Social Media Language. Journal of Personality and Social Psychology

Collective Rule



Collective Rule (1/3) (user-user relations)

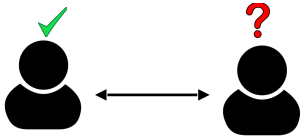


- Friend
- Follower
- Neighbour
- Spouse
- Idol
- Coauthor
- Colleague

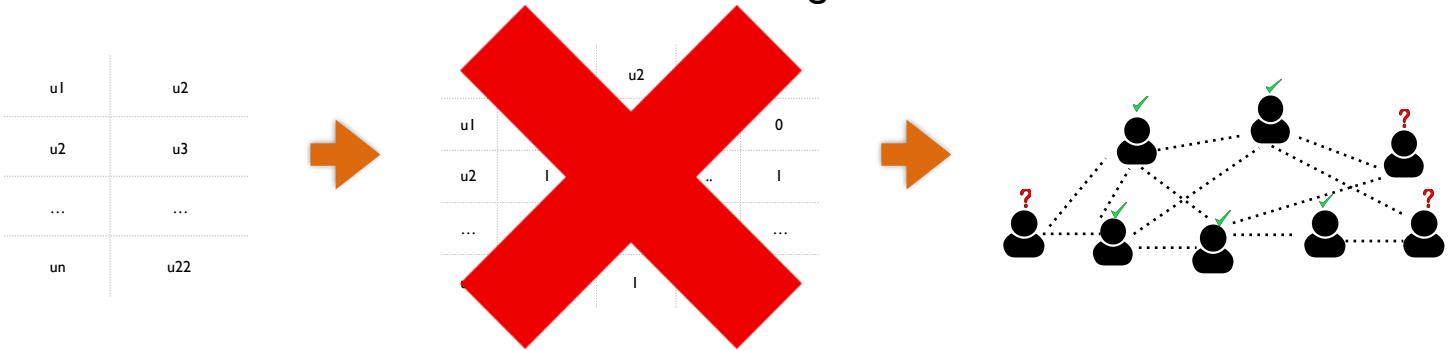
u1	u2
u2	u3
…	…
un	u22

	u1	u2	…	un
u1	1	1		0
u2	1	1	..	1
…	…	…		…
un	0	1	…	1

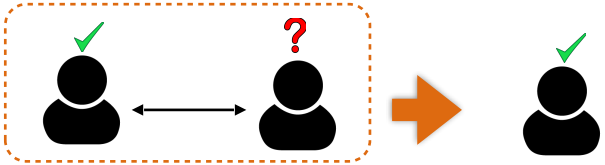
Collective Rule (1/3) (user-user relations)



- Friend
- Follower
- Neighbour
- Spouse
- Idol
- Coauthor
- Colleague

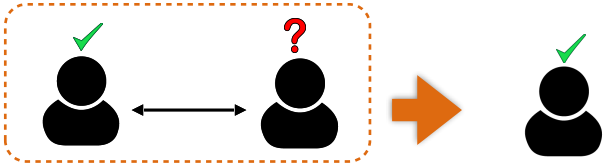


Collective Rule (1/3) (user-user relations)



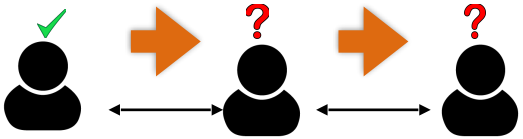
```
Friend(U1,U2) & Is(A,U1,L)-> Is(A,U2,L)
```

Collective Rule (1/3) (user-user relations)

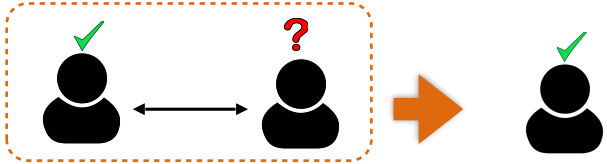


```
Friend(U1,U2) & Is(A,U1,L) -> Is(A,U2,L)
```

Open predicate



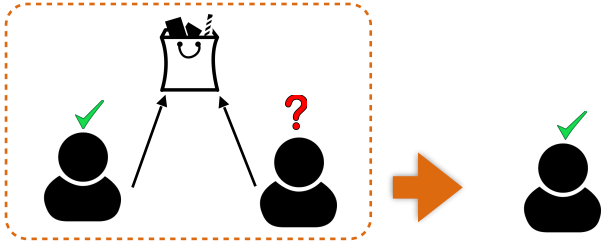
Collective Rule (1/3) (user-user relations)



example:

```
Friend(Alice,Carol) & Is(Personality,Carol,Ext)
                                -> Is(Personality,Alice,Ext)
```

Collective Rule (2/3) (user-item relations)



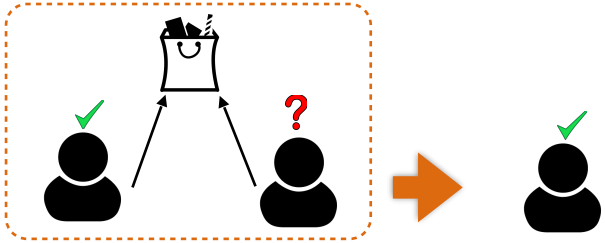
- Page likes
- Item rating
- Movie ratings

u1	I1
u2	I7
…	…
un	I8

	I1	I2	…	In
u1	1	1		0
u2	1	1	..	1
…	…	…		…
un	0	1	…	1

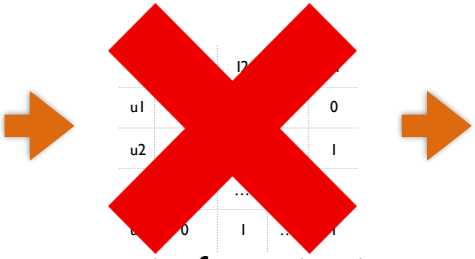
matrix-factorisation

Collective Rule (2/3) (user-item relations)

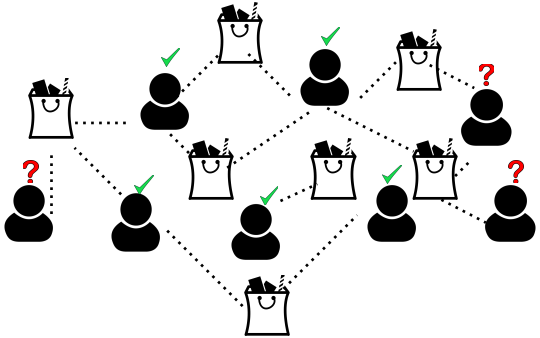


- Page likes
- Item rating
- Movie ratings

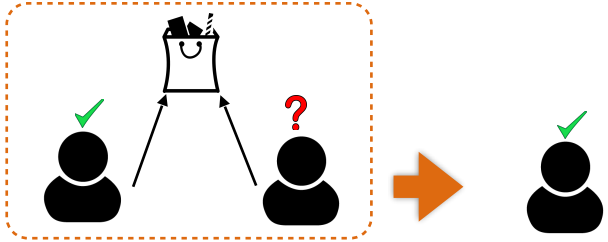
u1	I1
u2	I7
…	…
un	I8



matrix-factorisation

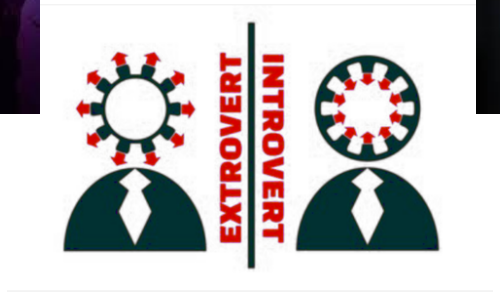


Collective Rule (2/3) (user-item relations)



```
Likes(U1,I) & Likes(U2,I) & Is(A,U1,L)-> Is(A,U2,L)
```

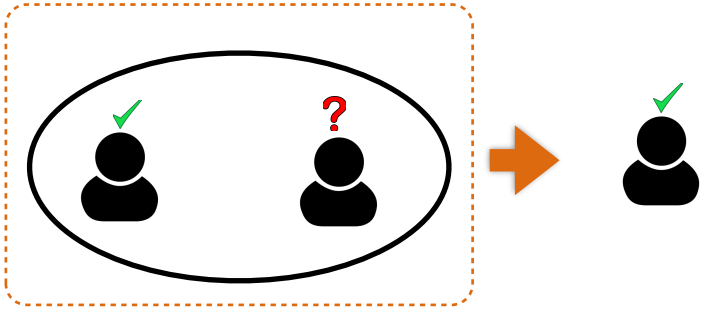
Collective Rule (2/3) (user-item relations)



example:

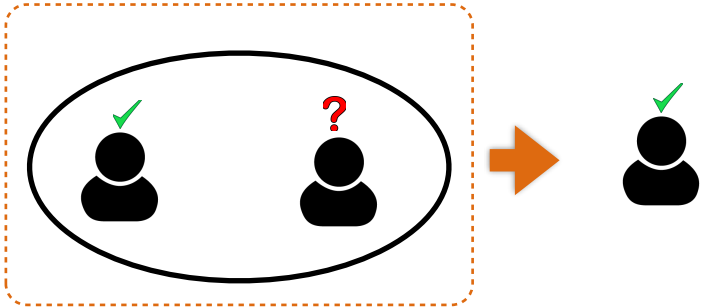
```
Likes(Alice,Partying) & Likes(Carol,Partying) &
     Is(Personality,Carol,Ext)-> Is(Personality,Alice,Ext)
```

Collective Rule (3/3) (user-group relations)



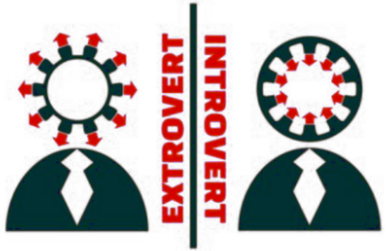
- groups
- clusters

Collective Rule (3/3) (user-group relations)



```
Joins(U1,G) & Joins(U2,G) & Is(A,U1,L)-> Is(A,U2,L)
```

Collective Rule (3/3) (user-group relations)



example:

```
Joins(Carol,Action-Movies) & Joins(Alice,Action-Movie) &
    Is(Personality,Carol,Ext)-> Is(Personality,Alice,Ext)
```

Hands on

- **Data**: Synthetic data, friendship links is a network whose degree distribution follows a power law, with 100 users, two local predictors, one set of joins relations and one set of likes relations.

- git clone https://github.com/linqs/psl-examples.git
- cd psl-examples/user-modeling/cli
- git checkout uai18

- Models
    - Local predictor (Text and Image)
    - Friendship
    - Likes
    - Joins
    - all

PSL Model for User Modeling

```
//Priors from local classifiers
  1: Has(U,S) & Predicts(S,A,U,L)-> Is(A,U,L)
  1: Has(U,S) & ~Is(A,U,L) -> ~Predicts(S,A,U,L)

//Collective Rules for relational signals

  1: Friend(U,V) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(U,V) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Friend(V,U) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(V,U) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & Is(A,V,L) -> Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & ~Is(A,V,L) -> ~Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & Is(A,V,L) -> Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & ~Is(A,V,L) -> ~Is(A,U,L)

//Ensure that user has one attribute
  1: Is(A,U,+L) = 1
```

Data file for User Modeling

predicates:

Predicts/4: closed
Friend/2: closed
Likes/2: closed
Joins/2: closed
Has/2: closed
Is/3: open

observations:

Predicts: ../data/local_predictor_obs.txt
Has: ../data/has_obs.txt
Friend: ../data/friend_obs.txt
Likes : ../data/likes_obs.txt
Joins : ../data/joins_obs.txt
Is : ../data/user_train.txt

targets:

Is : ../data/user_target.txt

truth:

Is : ../data/user_truth.txt

PSL Model for User Modeling

```
//Priors from local classifiers
  1: Has(U,S) & Predicts(S,A,U,L)-> Is(A,U,L)
  1: Has(U,S) & ~Is(A,U,L) -> ~Predicts(S,A,U,L)

//Collective Rules for relational signals

  1: Friend(U,V) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(U,V) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Friend(V,U) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(V,U) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & Is(A,V,L) -> Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & ~Is(A,V,L) -> ~Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & Is(A,V,L) -> Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & ~Is(A,V,L) -> ~Is(A,U,L)

//Ensure that user has one attribute
  1: Is(A,U,+L) = 1
```

local predictor

PSL Model for User Modeling

```
//Priors from local classifiers
  1: Has(U,S) & Predicts(S,A,U,L)-> Is(A,U,L)
  1: Has(U,S) & ~Is(A,U,L) -> ~Predicts(S,A,U,L)

//Collective Rules for relational signals

  1: Friend(U,V) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(U,V) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Friend(V,U) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(V,U) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & Is(A,V,L) -> Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & ~Is(A,V,L) -> ~Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & Is(A,V,L) -> Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & ~Is(A,V,L) -> ~Is(A,U,L)

//Ensure that user has one attribute
  1: Is(A,U,+L) = 1
```

Friend

PSL Model for User Modeling

```
//Priors from local classifiers
  1: Has(U,S) & Predicts(S,A,U,L)-> Is(A,U,L)
  1: Has(U,S) & ~Is(A,U,L) -> ~Predicts(S,A,U,L)

//Collective Rules for relational signals

  1: Friend(U,V) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(U,V) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Friend(V,U) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(V,U) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & Is(A,V,L) -> Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & ~Is(A,V,L) -> ~Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & Is(A,V,L) -> Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & ~Is(A,V,L) -> ~Is(A,U,L)

//Ensure that user has one attribute
  1: Is(A,U,+L) = 1
```

Likes

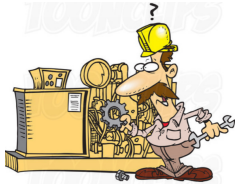
PSL Model for User Modeling

```
//Priors from local classifiers
  1: Has(U,S) & Predicts(S,A,U,L)-> Is(A,U,L)
  1: Has(U,S) & ~Is(A,U,L) -> ~Predicts(S,A,U,L)

//Collective Rules for relational signals

  1: Friend(U,V) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(U,V) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Friend(V,U) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(V,U) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & Is(A,V,L) -> Is(A,U,L)
  1: Likes(U,T) & Likes(V,T) & ~Is(A,V,L) -> ~Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & Is(A,V,L) -> Is(A,U,L)
  1: Joins(U,G) & Joins(V,G) & ~Is(A,V,L) -> ~Is(A,U,L)

//Ensure that user has one attribute
  1: Is(A,U,+L) = 1
```

Joins

Evaluation Result (Personality prediction-synthetic data)

Type of the model	AUC
Random	0.5
Local Predictor	0.811655
Friendship links	0.528963
Likes	0.724014
Joins	0.865880
All	0.777315



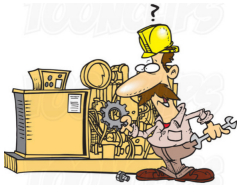
Rules' weights?

PSL Model for User Modeling

```
//Priors from local classifiers
  50: Has(U,S) & Predicts(S,A,U,L)-> Is(A,U,L)
  50: Has(U,S) & ~Is(A,U,L) -> ~Predicts(S,A,U,L)

//Collective Rules for relational signals

  1: Friend(U,V) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(U,V) & ~Is(A,V,L)-> ~Is(A,U,L)
  1: Friend(V,U) & Is(A,V,L)-> Is(A,U,L)
  1: Friend(V,U) & ~Is(A,V,L)-> ~Is(A,U,L)
  10: Likes(U,T) & Likes(V,T) & Is(A,V,L) -> Is(A,U,L)
  10: Likes(U,T) & Likes(V,T) & ~Is(A,V,L) -> ~Is(A,U,L)
  100: Joins(U,G) & Joins(V,G) & Is(A,V,L) -> Is(A,U,L)
  100: Joins(U,G) & Joins(V,G) & ~Is(A,V,L) -> ~Is(A,U,L)

//Ensure that user has one attribute
  1: Is(A,U,+L) = 1
```

Evaluation Result (Personality prediction-synthetic data)

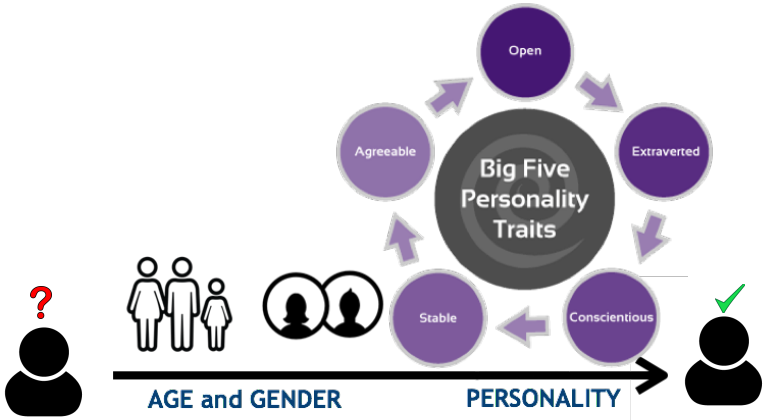
Type of the model	AUC
Random	0.5
Local Predictor	0.811655
Friendship links	0.528963
Likes	0.724014
Joins	0.865880
All	0.777315
All	0.913516



Other combinations?

Weight learning?

Knowledge Fusion Model for User Profiling Based on Multimedia and Multi-Relational User-Generated Content



- Personalised services
- Marketing and advertisement
- Law enforcement
- Employment selection

[Work in progress]

Predicting Users' Age, Gender and Big5 Personality traits

Task: Predicting Facebook Users':

Age, Gender and Big5 Personality traits (Extraversion (Ext), Agreeableness (Agr), Neuroticism (Neu), Openness (Opn), Conscientiousness (Con))

Using **Status updates**, **Profile Picture** and **Facebook Page Likes**

Data: ~6K Facebook users, ~49K Facebook pages and ~725K Page likes Relations

Results: Area under the curve (AUC), 10-fold CV

Model/Characteristic	Gender	Age	Opn	Con	Ext	Agr	Neu
Baseline	0.492	0.488	0.502	0.502	0.506	0.504	0.486
PSL-Textual	0.650	0.710	0.570	0.567	0.553	0.550	0.542
PSL-Visual	0.850	0.579	0.505	0.521	0.531	0.531	0.515
PSL-Relational	0.853	0.881	0.648	0.618	0.592	0.571	0.572
PSL-Fusion	**0.893**	**0.893**	**0.654**	**0.622**	**0.599**	**0.581**	**0.58**