Judgment aggregation

acknowledgment: Ulle Endriss, University of Amsterdam

Lirong Xia

Rensselaer

Fall, 2016

Last class: Fair division

- Indivisible goods
  - house allocation: serial dictatorship
  - housing market: Top trading cycles (TTC)

Judgment aggregation: the doctrinal paradox

	Action p	Action q	Liable? ($p \wedge q$)
Judge 1	Y	Y	Y
Judge 2	Y	N	N
Judge 3	N	Y	N
Majority	**Y**	**Y**	**N**

- p: valid contract
- q: the contract has been breached
- Why paradoxical?
  - issue-by-issue aggregation leads to an illogical conclusion

Formal framework

- An agenda A is a finite nonempty set of propositional logic formulas closed under complementation ($[\varphi \in A] \Rightarrow [\sim\varphi \in A]$)
  - A = { p, q, ~p, ~q, p$\wedge$q}
  - A = { p, ~p, p$\wedge$q, ~p$\vee$~q}
- A judgment set J on an agenda A is a subset of A (the formulas that an agent thinks is true, in other words, accepts). J is
  - complete, if for all $\varphi \in A$, $\varphi \in J$ or $\sim\varphi \in J$
  - consistent, if J is satisfiable
  - S(A) is the set of all complete and consistent judgment sets
- Each agent (judge) reports a judgment set
  - $D = (J_1,\ldots,J_n)$ is called a profile
- An judgment aggregation (JA) procedure F is a function $(S(A))^n \rightarrow \{0,1\}^A$

Some JA procedures

- Majority rule
  - F(φ)=1 if and only if the majority of agents accept φ
- Quota rules
  - F(φ)=1 if and only if at least k% of agents accept φ
- Premise-based rules
  - apply majority rule on “premises”, and then use logic reasoning to decide the rest
- Conclusion-based rules
  - ignore the premises and use majority rule on “conclusions”
- Distance-based rules
  - choose a judgment set that minimizes distance to the profile

Premise-based approaches

- $A = A_p + A_c$
  - A_p=premises
  - A_c=conclusions
- Use the majority rule on the premises, then use logic inference for the conclusions
- Theorem. If
  - the premises are all literals
  - the conclusions only use literals in the premises
  - the number of agents is odd
- then the premise-based approach is anonymous, consistent, and complete

	p	q	$(p \wedge q)$
Judge 1	Y	Y	Y
Judge 2	Y	N	N
Judge 3	N	Y	N
Majority	**Y**	**Y**	**Logic reasoning Y**

Recommender systems

- Content-based approaches
  - based on user's past ratings on similar items computed using features
- Collaborative filtering
  - user-based: find similar users
  - item-based: find similar items (based on all users' ratings)

Applications

yelp

NETFLIX

amazon.com

YouTube

The Netflix challenge

- $1M award to the first team who can outperform their own recommender system CinMatch by 10%
- A big dataset
  - half million users
  - 17000 movies
  - a secret test set
- Won by a hybrid approach in 2009
  - a few minutes later another hybrid approach also achieved the goal

The problem

- Given
  - features of users i
  - features of items j
  - users' ratings $r_i(j)$ over items
- Predict
  - a user's preference over items she has not tried
    - by e.g., predicting a user's rating of new item
- Not a social choice problem, but has a information/preference aggregation component

Classical approaches

- Content-based approaches
- Collaborative filtering
  - user-based: find similar users
  - item-based: find similar items (based on all users' ratings)
- Hybrid approaches

Framework for content-based approaches

- Inputs: profiles for items
  - K features of item j
    - $w_j = (w_{j1}, \ldots, w_{jK})$
    - $w_{jk} \in [0,1]$: degree the item has the feature
  - the user's past ratings for items 1 through j-1
- Similarity heuristics
  - compute the user's profile: $v_i = (v_{i1}, \ldots, v_{iK})$, $v_{ik} \in [0,1]$
  - recommend items based on the similarity of the user's profile and profiles of the items
- Probabilistic approaches
  - use machine learning techniques to predict user's preferences over new items

Example

	Animation	Adventure	Family	Comedy	Disney	Bluesky	rate
	1	1	1	0	0	1	?
	1	1	0	1	0	1	9
	1	0	1	1	1	0	8
	1	1	1	0	1	0	7
$v =$	0.8	0.8	0.75	0.85	0.75	0.9	

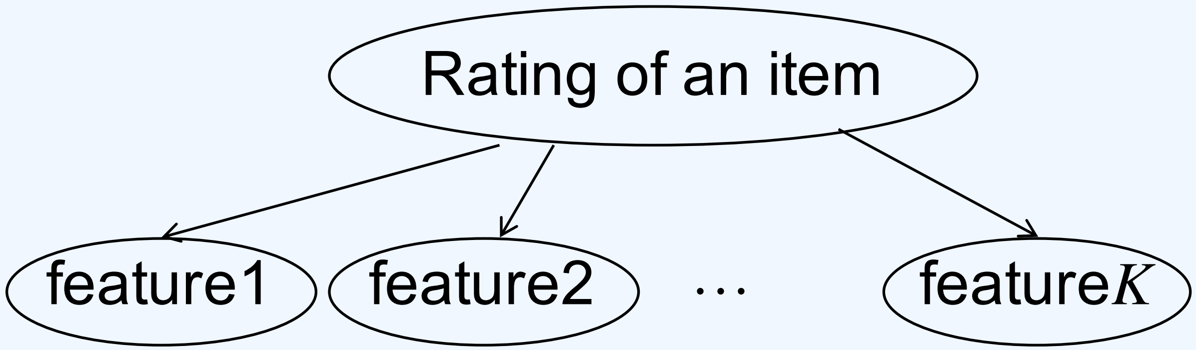
Similarity heuristics

- A possible way to define v_i
  - v_{ik} is the average normalized score of the user over items with feature k
- A possible way to define similarly measure
  - cosine similarity measure

$$\cos(v_i, w_j) = \frac{v_i \cdot w_j}{\| v_i \|_2 \| w_j \|_2} = \frac{\sum_{k=1}^{K} v_{ik} \cdot w_{jk}}{\sqrt{\sum_{k=1}^{K} v_{ik}^2} \sqrt{\sum_{k=1}^{K} w_{ik}^2}}$$

  - in the previous example, the measure is 0.68

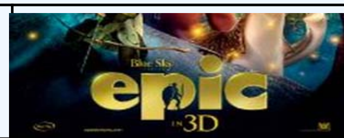
Probabilistic classifier



- Naïve Bayes model: suppose we know
  - $\Pr(r)$
  - $\Pr(f_k|r)$ for every r and k
  - learned from previous ratings using MLE
- Given $w_j = (w_{j1}, \ldots, w_{jK})$
  - $\Pr(r|w_j) \propto \Pr(w_j|r)\Pr(r) = \Pr(r)\,\Pi \Pr(w_{jk}|r)$
  - Choose r that maximizes $\Pr(r|w_j)$

Framework for collaborative filtering approaches

- Inputs: a matrix M.
  - $M_{i,j}$: user i's rating for item j

				
Alice	8	6	4	9
Bob	$\varnothing$	8	10	10
Carol	4	4	8	$\varnothing$
David	6	$\varnothing$	10	5

- Collaborative filters
  - User-based: use similar users' rating to predict
  - Item-based: use similar items' rating to predict

User-based approaches (1)

- Step 1. Define a similarity measure between users based on co-rated items
  - Pearson correlation coefficient between i and i^*
  - G_{i,i^*}: the set of all items that both i and i^* have rated
  - $\overline{M_i}$: the average rate of user i

$$sim(i,i^*) = \frac{\sum_{j \in G_{i,i^*}} (M_{ij} - \overline{M_i}) \cdot (M_{i^*j} - \overline{M_{i^*}})}{\sqrt{\sum_{j \in G_{i,i^*}} (M_{ij} - \overline{M_i})^2} \sqrt{\sum_{j \in G_{i,i^*}} M_{i^*j} - \overline{M_{i^*}})^2}}$$

User-based approaches (2)

- Step 2. Find all users $i*$ within a given threshold
  - let N_i denote all such users
  - let N_i^j denote the subset of N_i who have rated item j

User-based approaches (3)

- Step 3. Predict i's rating on j by aggregating similar users' rating on j

$$\hat{r}_i(j) = \frac{1}{|N_i^j|} \sum_{i^* \in N_i^j} r_{i^*}(j)$$

$$\hat{r}_i(j) = \frac{\sum_{i^* \in N_i^j} sim(i,i^*) r_{i^*}(j)}{\sum_{i^* \in N_i^j} sim(i,i^*)}$$

$$\hat{r}_i(j) = \overline{M}_i + \frac{\sum_{i^* \in N_i^j} sim(i,i^*)(r_{i^*}(j) - \overline{M}_{i^*})}{\sum_{i^* \in N_i^j} sim(i,i^*)}$$

Item-based approaches

- Transpose the matrix M
- Perform a user-based approach on M^T

Hybrid approaches

- Combining recommenders
  - e.g. content-based + user-based + item-based
  - social choice!
- Considering features when computing similarity measures
- Adding features to probabilistic models

Challenges

- New user
- New item
- Knowledge acquisition
  - discussion paper: preference elicitation
- Computation: challenging when the number of features and the number of users are extremely large
  - M is usually very sparse
  - dimension reduction