Recommendation via user’s personality and extracted product features from text

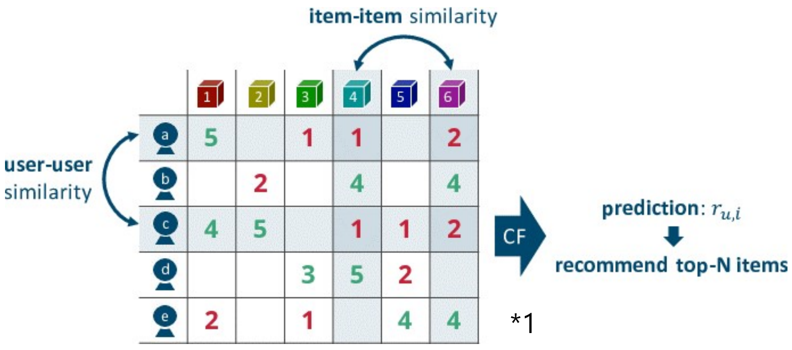
B9EM1013 Yosuke Sakai

※スライドの数式中にいくつか誤りがあ
ることを直前に発見しました。
また、不確かな情報が含まれていること
をあらかじめ謝罪いたします。

Motivation

- Recommendation … Selecting information or products which are beneficial for users, and presenting these to users (Kamishima 2007)
- There are mainly two types of recommendation system
  - Collaborative filtering

    …Calculating consumer similarity or product similarity from only evaluation history or purchase history. And predicting user i's evaluation (rating) against item j, then recommend top N items based on predicted rating.



	1	2	3	4	5	6
a	5		1	1		2
b		2		4		4
c	4	5		1	1	2
d			3	5	2	
e	2		1		4	4

*1

  - Content based filtering

    …Extracting feature vectors of products and users. And using these vectors, predict user i's evaluation (rating) against item j, then recommend top N items based on predicted rating.

	1	2	3	4	5	6
a	5		1	1		2
b		2		4		4
c	4	5		1	1	2
d			3	5	2	
e	2		1		4	4

$$\underset{\beta_b}{arg\ min} \sum_{j}^{m_b} \left(r_{bj} - \boldsymbol{\theta_j \beta_b} \right)^2$$

r_{bj}: rating of product j by user b

$\boldsymbol{\theta_j}$: feature vector of product j

m_b: number of products which user b evaluated

---

*1 https://takuti.me/note/two-decades-of-amazon-recommender/

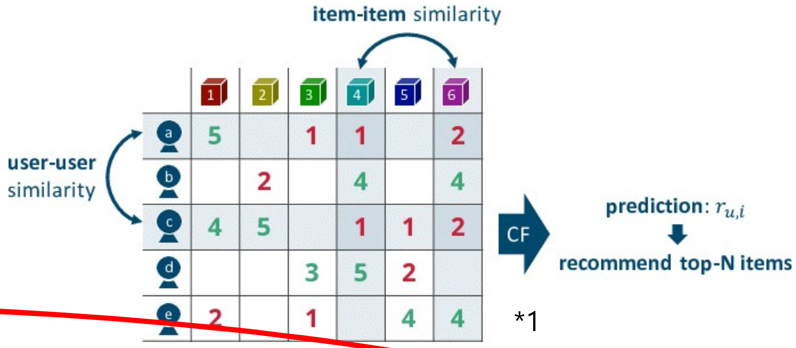
Motivation

- Recommendation … Selecting information or products which are beneficial for users, and presenting these to users (Kamishima 2007)
- There are mainly two types of recommendation system
  - Collaborative filtering

    …Ca y or product similarity from only evaluation history or purchase history.

    And predicting user i's evaluation (rating) against item j, then recommend top N items based on predicted rating.

Main content of today's presentation



	1	2	3	4	5	6
a	5		1	1		2
b		2		4		4
c	4	5		1	1	2
d			3	5	2	
e	2		1		4	4

  - **Content based filtering**

    …Extracting feature vectors of products and users.

    And using these vectors, predict user i's evaluation (rating) against item j, then recommend top N items based on predicted rating.

	1	2	3	4	5	6
a	5		1	1		2
b		2		4		4
c	4	5		1	1	2
d			3	5	2	
e	2		1		4	4

$$\underset{\beta_b}{arg\ min}\sum_{j}^{m_b}\left(r_{bj}-\boldsymbol{\theta_j\beta_b}\right)^2$$

r_{bj}: rating of product j by user b

$\boldsymbol{\theta_j}$: feature vector of product j

m_b: number of products which user b evaluated

*1 https://takuti.me/note/two-decades-of-amazon-recommender/

Motivation

- Some of characteristics of two methods
  - Collaborative filtering
    - ○ We don’t have to have domain knowledge.
    - × We cannot recommend to new user or cannot recommend new item.
    - × Recommendation without interpretation about why she likes it.
  - Content based filtering
    - ○ We can recommend new user or new item.
    - ○ Recommendation with interpretation about why she likes it.
    - × We have to have domain knowledge to extract feature vector.
    - × It is difficult to construct enough features to describe product.
- Because of trend of “explainable recommendation”, content based filtering is attracting attention.

Motivation

- Some researches dealt with the problem of content based filtering that "Difficulty for constructing enough features" and "Necessity for domain knowlege "
- They extracted features from texts automatically. And using extracted features instead of feature given explicitly, like "genre", they made recommendation.

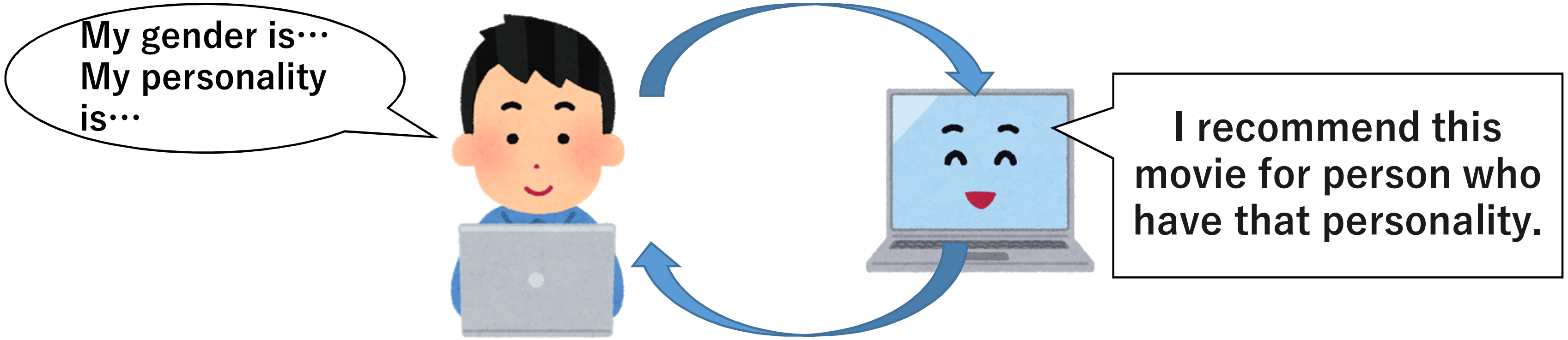
	Method	Used information	Research features
Büschken and Allenby (2016)	Topic model	･Review text ･Rating	They extracted product features from UGC and estimated rough preference parameters simultaneously.
Ansari et al. (2018)	Topic model	･Review text (word tag) ･Product genre ･Rating	They extracted product features from UGC and estimated individual preference parameters simultaneously.
Toubia et al. (2019)	Topic model + Hierarchical probit model	･Synopses of movies ･Product genre ･Rating	Firstly, they extracted product features relating psychological theme from synopses. Then, they used these feature and product genre for recommendation.

Motivation

- Problem of previous research
  - Above researches did not consider “user information”, like demographics or personality.
  - There may be interaction effects between extracted features and user information.

Ex) Users who have high “Sympathy” personality unlike entertainment products
having “Darkness” feature. On the other hands,
users who have high “Ingenuity” personality like these. (Rentflow et al. 2011)

- If we can reveal interaction effects, we can make recommendation against new users effectively with only user information (Pan 2015)



Motivation

Outline of my research

- Extracting product’s features which we cannot get explicitly from review text by topic model.
- Modeling user evaluation by hierarchical regression model using above feature and user’s personality.

Purpose of my research

- Using extracted features, making recommendation with higher accuracy than using explicitly feature, “genre”.
- By revealing rough preference for each personalities, making recommendation against even new users (have not evaluated) with higher accuracy than case of absence of it.

What’s “topic model”?

- Büschken and Allenby (2016), Ansari and Zhang (2018), Toubia et al. (2019) extracted features of products from review text by topic model.
- “Topic model” is method for dimension reduction or summary.

	文書 1	文書 2	文書 3	文書 4
雇用	0	2	0	0
景気	0	0	1	0
政策	1	0	0	0
総理	0	1	0	0

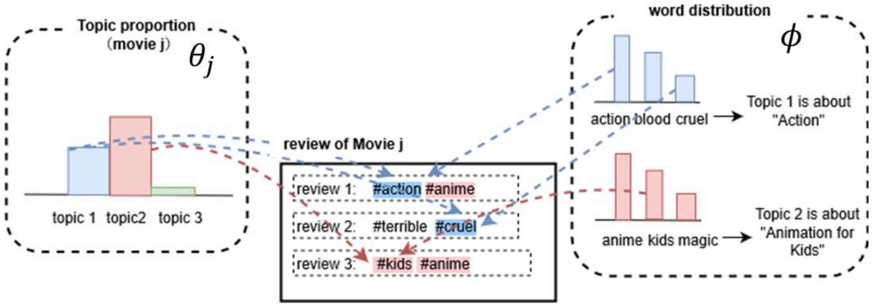
Bag of words 表現

	topic 1	topic 2	topic3
雇用	0.01	0	0
景気	0.01	0	0
政策	0	0.1	0
総理	0	0.2	0

単語分布 (Φ)

	文書 1	文書 2	文書 3	文書 4
topic 1	0.1	0	0.3	0.1
topic 2	0	0	0	0.5
topic3	0	0.3	0.2	0

トピック分布(Θ)



- “Topic” is used as features of products.

Empirical study

- <u>Outline of empirical study</u>
  - Empirical study is divided into two stage
  - Data are reviews of movies. So, in this study, item is “movie”.

Stage 1

- Extraction of movie’s features by topic model from review text
  - Interpreting means of topics through $\boldsymbol{\phi_k}$, and $\boldsymbol{\theta}$
  - $\boldsymbol{\theta_j}$ are used as feature of movie j

Stage 2

- Modeling user’s rating by hierarchical regression model of Ansari et al.(2000)
  - $r_{ij} = \boldsymbol{\theta}_{\boldsymbol{j}}'\boldsymbol{\beta_i} + \boldsymbol{X_i}'\boldsymbol{\beta_j} + \boldsymbol{\epsilon_{ij}} \quad \boldsymbol{\epsilon_{ij}} \sim N_n\left(0, \sigma^2\right)$
    $\boldsymbol{\beta_i} = \boldsymbol{\Delta X_i} + \boldsymbol{\lambda_i} \quad \boldsymbol{\eta_i} \sim N_K(0, \boldsymbol{\Lambda})$
    $\boldsymbol{\beta_j} = \boldsymbol{\gamma_j} \quad \boldsymbol{\gamma_j} \sim N_p(0, \boldsymbol{\Gamma})$

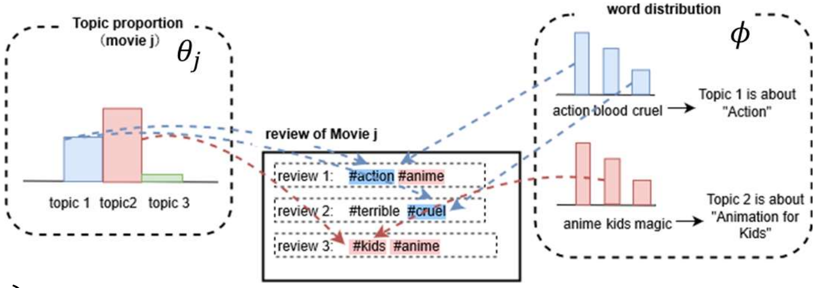
r_{ij} : rating of movie j by user i
$\boldsymbol{\theta_j}$: Topic proportion of movie j $\left(\boldsymbol{\theta_j} \in \mathbb{R}^{\boldsymbol{K}}\right)$
$\boldsymbol{X_i}$: Personality of user i $(\boldsymbol{X_i} \in \mathbb{R}^{\boldsymbol{p}})$

Empirical model

Stage 1

- Extraction of movie's feature by topic model from reviews.

1. For トピック　$k = 1,・・・\mathrm{K}$
    (*a*) 単語分布を生成　$\boldsymbol{\phi_k} \sim Dirichlet(\boldsymbol{\beta})$
2. For documents $d = 1,・・・D$
    (*a*) トピック分布を生成　$\boldsymbol{\theta_d} \sim Dirichlet(\boldsymbol{\alpha})$
3. For documents $d = 1,・・・D$
    (*a*) For 単語　$\mathrm{n} = 1,・・・\mathrm{N}$
        (i)トピックを生成　$z_{dn} \sim Multinomial(\boldsymbol{\theta_d})$
        (ii)単語を生成　$w_{dn} \sim Multinomial(\boldsymbol{\phi_{z_{dn}}}\ \ )$



- I employed collapsed Gibbs sampling as estimation method

full condition posterior distribution.

$$p(z_{dn} = k|\boldsymbol{z_{-dn}}, \boldsymbol{w}) \propto \frac{N_{kv}^{-dn}+\beta}{\Sigma_v(N_{kv}^{-m}\ \ +\beta)} * \frac{N_{mk}^{-n}+\alpha}{\Sigma_k(N_{mk}^{-n}+\alpha)}$$

Empirical model

Stage 2

- Hierarchical regression model of Ansari et al. (2000) modeled →
  - Fixed effects of user information and product information, interaction of these (**Δ**)
  - Random effects of user and products.($\boldsymbol{\lambda_i}$, $\boldsymbol{\gamma_j}$)

$$r_{ij} = \boldsymbol{\theta_j}'\boldsymbol{\beta_i} + \boldsymbol{X_i}'\boldsymbol{\beta_j} + \epsilon_{ij} \quad \epsilon_{ij} \sim N\left(0, \sigma^2\right)$$

$$\boldsymbol{\beta_i} = \boldsymbol{\Delta}\boldsymbol{X_i} + \boldsymbol{\lambda_i} \quad \boldsymbol{\lambda_i} \sim N_K(0, \boldsymbol{\Lambda})$$

$$\boldsymbol{\beta_j} = \boldsymbol{\gamma_j} \quad \boldsymbol{\gamma_j} \sim N_p(0, \boldsymbol{\Gamma})$$

r_{ij} : rating of movie j by user i

$\boldsymbol{\theta_j}$: Topic proportion of movie j $\left(\boldsymbol{\theta_j} \in \mathbb{R}^{\boldsymbol{K}}\right)$

$\boldsymbol{X_i}$: Personality of user i $(\boldsymbol{X_i} \in \mathbb{R}^{\boldsymbol{p}})$

- Ansari et al. (2000) used product's genre and user's demographics instead of topic proportion and personality.

Empirical model

Stage 2

	Prior distribution	Posterior distribution (Full condition)
$\boldsymbol{\beta}_i$	$p(\boldsymbol{\beta}_i \vert \boldsymbol{\Delta}, \boldsymbol{\Lambda}, \boldsymbol{X}_i) \sim N_K(\boldsymbol{\Delta X}_i, \boldsymbol{\Lambda})$	$p(\boldsymbol{\beta}_i \vert \boldsymbol{\Delta}, \boldsymbol{\Lambda}, \boldsymbol{\theta}_i^*, \boldsymbol{X}_i, \boldsymbol{r}_i^*) \sim N_K(\widetilde{\boldsymbol{\beta}}_i, \frac{(\boldsymbol{\theta}^{*T}_i \boldsymbol{\theta}_i^*)}{\sigma^2} + \boldsymbol{\Lambda}^{-1})$ $\widetilde{\boldsymbol{\beta}}_i = \left(\frac{\left(\boldsymbol{\theta}^{*T}_i \boldsymbol{\theta}_i^*\right)}{\sigma^2} + \boldsymbol{\Lambda}^{-1}\right)^{-1} \left(\frac{\left(\boldsymbol{\theta}_i^{*T} \widetilde{\boldsymbol{r}}_i\right)}{\sigma^2} + \boldsymbol{\Lambda}^{-1}(\boldsymbol{\Delta X}_i)\right)$
$\boldsymbol{\beta}_j$	$p\left(\boldsymbol{\beta}_j \middle\vert \boldsymbol{\Gamma}\right) \sim N_p(0, \boldsymbol{\Gamma})$	$p\left(\boldsymbol{\beta}_j \middle\vert \boldsymbol{\Gamma}, \boldsymbol{X}_j^*, \boldsymbol{r}_j^*\right) \sim N_p(\widetilde{\boldsymbol{\beta}}_j, \boldsymbol{\Gamma}^{-1})$ $\widetilde{\boldsymbol{\beta}}_j = (\boldsymbol{\Gamma}^{-1})^{-1} \left(\frac{\left(\boldsymbol{X}_j^{*T} \widetilde{\boldsymbol{r}}_j\right)}{\sigma^2}\right)$
σ^2	$p(\sigma^2) \sim IG(\frac{\nu_{0e}}{2}, \frac{s_{0e}}{2})$	$p\left(\sigma^2 \vert \boldsymbol{\beta}_i, \boldsymbol{\beta}_j, \boldsymbol{\Theta}, \boldsymbol{X}, \boldsymbol{r}\right) \sim IG(\frac{\nu_{0e} + N}{2}, \frac{s_{0e} + s^2}{2})$ $s^2 = \sum_{n=1}^{N} r_n - \boldsymbol{\theta}_{j^n} \boldsymbol{\beta}_{i^n} + \boldsymbol{X}_{i^n} \boldsymbol{\beta}_{j^n}$

※1 $\widetilde{\boldsymbol{r}}_i$ is obtained by stacking all the elements $\widetilde{r_{ij}}$ for user i $\widetilde{r_{ij}} = r_{ij} - X_i \beta_i$

※2 $\boldsymbol{\theta}_i^*$ is obtained by stacking the all $\boldsymbol{\theta}_j$ for user i

Empirical model

Stage 2

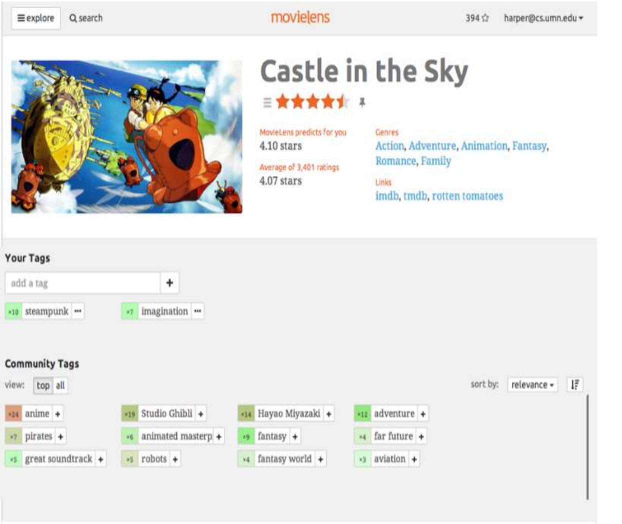
	Prior distribution	Posterior distribution (Full condition)
$\boldsymbol{\Delta}$	$p(\boldsymbol{\Delta}\|\,\boldsymbol{\Lambda}\,) \sim N_{K*p}(\mathrm{vec}(\overline{\boldsymbol{\Delta}}), \boldsymbol{\Lambda} \otimes A_{\Delta})$	$p(\boldsymbol{\Delta}\|\,\boldsymbol{\Lambda}\,, \boldsymbol{\beta}_{\boldsymbol{i}}, \boldsymbol{X}) \sim N_{K*p}(\tilde{d}, \boldsymbol{\Lambda} \otimes (\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X} + A_{\Delta}))$ $\tilde{d} = \mathrm{vec}(\widetilde{D}), \qquad \widetilde{D} = (\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X} + A_{\Delta})^{-1}(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{\beta}_{\boldsymbol{user}} + A_{\Delta}\overline{\boldsymbol{\Delta}})$
$\boldsymbol{\Lambda}$	$p(\boldsymbol{\Lambda}) \sim IW(\nu_{0l}, \boldsymbol{\Lambda}_{\boldsymbol{0}})$	$p(\boldsymbol{\Lambda}\|\boldsymbol{\beta}_{\boldsymbol{user}}) \sim IW(\nu_{0l} + N_{user}, \boldsymbol{\Lambda}_{\boldsymbol{0}} + \boldsymbol{S}'_{\boldsymbol{\Lambda}})$ $\boldsymbol{S}'_{\boldsymbol{\Lambda}} = \sum_{\boldsymbol{n=1}}^{N_{user}} (\boldsymbol{\beta}_{\boldsymbol{n,user}} - \boldsymbol{\Delta}\boldsymbol{X}_{\boldsymbol{n}})$
$\boldsymbol{\Gamma}$	$p(\boldsymbol{\Gamma}) \sim IW(\nu_{0g}, \boldsymbol{\Gamma}_{\boldsymbol{0}})$	$p(\boldsymbol{\Gamma}) \sim IW(\nu_{0g}, +N_{item}, \boldsymbol{\Gamma}_{\boldsymbol{0}} + \boldsymbol{S}'_{\boldsymbol{\Gamma}})$ $\boldsymbol{S}'_{\boldsymbol{\Gamma}} = \boldsymbol{\beta}_{\boldsymbol{item}}^{\boldsymbol{T}}\boldsymbol{\beta}_{\boldsymbol{item}}$

- Using posterior distribution, I employed Gibbs sampling for estimation of parameters.

Data for empirical study

Stage 1

- Review data of movies (“Movielens 10M”)
  - 7,038 movies with 71,060 word tags which given by users.
  - Tags were given in 2005/12~2009/01
  - Each movies have several following 18 genres.
    - *“Action“, "Adventure“, "Animation“, "Children's“, "Comedy" ,"Crime" , "Documentary“, "Drama“, "Fantasy“, "Film-Noir“, "Horror“, "Musical“, "Mystery“, "Romance“, "Sci-Fi“, "Thriller“, "War“, "Western"*



Castle in the Sky

MovieLens predicts for you
4.10 stars

Average of 3,401 ratings
4.07 stars

Genres
Action, Adventure, Animation, Fantasy, Romance, Family

Links
imdb, tmdb, rotten tomatoes

Your Tags

add a tag

×10 steampunk
×7 imagination

Community Tags

view: top all

sort by: relevance

×24 anime
×19 Studio Ghibli
×14 Hayao Miyazaki
×12 adventure
×7 pirates
×6 animated masterp
×9 fantasy
×4 far future
×5 great soundtrack
×5 robots
×4 fantasy world
×3 aviation

Stage 2

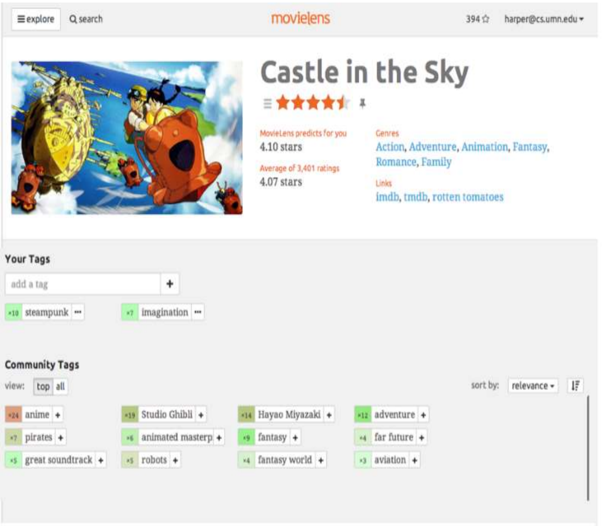
- Review data of movies (“Personality 2018” provided by grouplens)
  - Ratings (★×0.5 ~★×5 in 10 revels) about 7,038 movies given by users
  - Ratings were given in 1997/09~2019/01
  - User’s personalities are given as score (1~7 by 0.5 point) by following viewpoints
    - *“Openness”, “Agreeableness”, "Emotional stability”, “Conscientiousness”, “Extraversion”*   →called “The Big five personality”

Data for empirical study

Stage 1

- Review data of movies (“Movielens 10M”)
  - 7,038 movies with 71,161 word tags which given by users.
  - Tags were given in 2005/12~2009/01
  - Each movies have several following 18 genres.
    - *“Actio* **Common** *e“, "Animation“, "Children's“, "Comedy", "Crime", "Documentary“, "Drama“, "Fantasy“, "Film-Noir“, "Horror“, "Musical“, "Mystery“, "Romance“, "Sci-Fi“, "Thriller“, "War“, "Western”*

Different



Castle in the Sky

MovieLens predicts for you
4.10 stars

Average of 3,401 ratings
4.07 stars

Genres
Action, Adventure, Animation, Fantasy, Romance, Family

Links
imdb, tmdb, rotten tomatoes

Your Tags
×10 steampunk
×7 imagination

Community Tags
×24 anime
×7 pirates
×5 great soundtrack
×19 Studio Ghibli
×6 animated masterp
×5 robots
×14 Hayao Miyazaki
×9 fantasy
×4 fantasy world
×12 adventure
×4 far future
×3 aviation

Stage 2

- Review data of movies (“Personality 2018” provided by grouplens)
  - Ratings (★×0.5 ~★×5 in 10 revels) about 7,038 movies given by users
  - Ratings were given in 1997/09~2019/01
  - User’s personalities are given as score (1~7 by 0.5 point) by following viewpoints
    - *“Openness”, “Agreeableness”, ”Emotional stability”, “Conscientiousness”, “Extraversion”* →called “The Big five personality”

Data for empirical study

- What's "The Big five personality" ?
  - 5 viewpoints for measuring one's personality which are often used in psychology.

	Relating Feature	Opposite Feature
Openness	Open to new experience, complex	Conventional, Uncreative
Agreeableness	Sympathetic, warm	Critical, Quarrelsome
Emotional stability	Anxious, easy to upset	Calm, Emotional stable
Conscientiousness	Dependable, self-disciplined	Disorganized, Careless
Extraversion	Extraverted, enthusiastic	Quiet

- Dataset which is used in this research ask user to evaluate their personality at above viewpoints as score (1~7 by 0.5 point). High score means high tendency.*

*Only emotional stability is reverse.

Data for empirical study

- Data for stage2 is collected by Nguyen et al. (2018)
- They asked below questions to Movielens users.
- Personality scores of every items are given as average value between a positive question and reversed one.

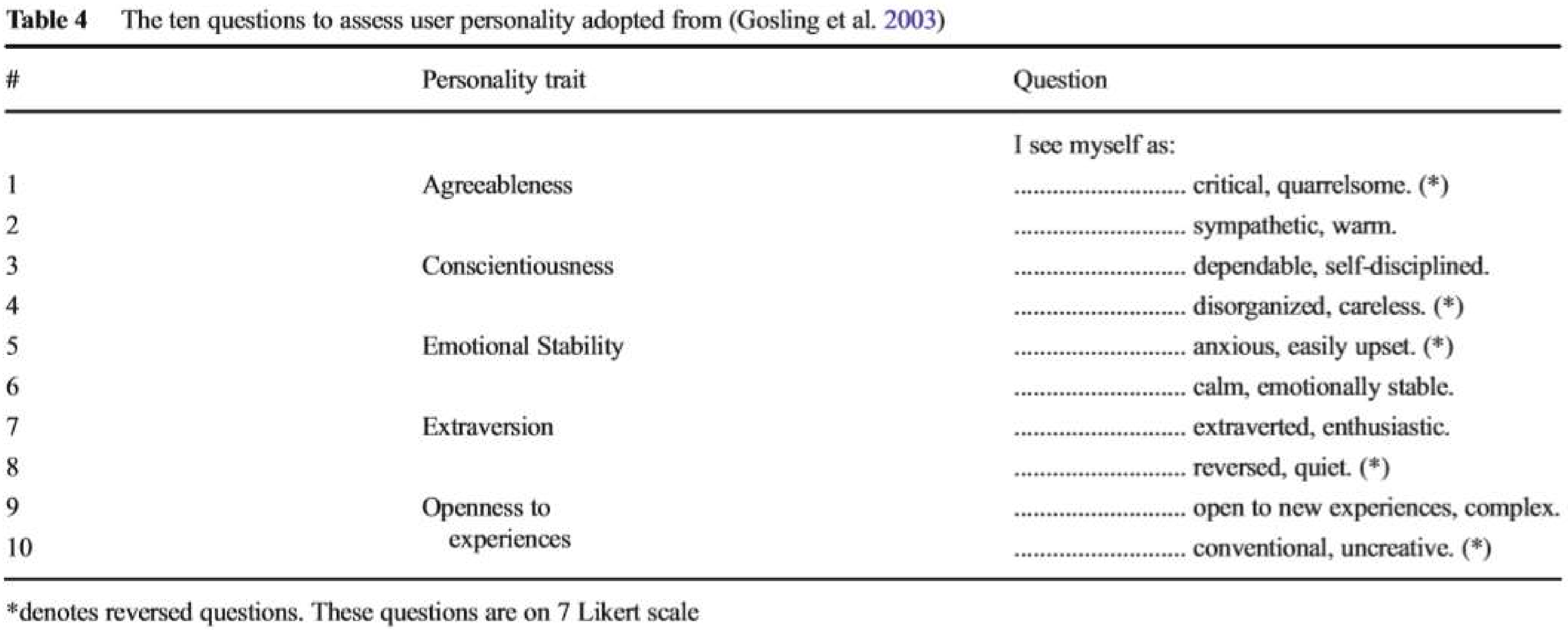


Table 4 The ten questions to assess user personality adopted from (Gosling et al. 2003)

#	Personality trait	Question
		I see myself as:
1	Agreeableness	.... critical, quarrelsome. (*)
2		.... sympathetic, warm.
3	Conscientiousness	.... dependable, self-disciplined.
4		.... disorganized, careless. (*)
5	Emotional Stability	.... anxious, easily upset. (*)
6		.... calm, emotionally stable.
7	Extraversion	.... extraverted, enthusiastic.
8		.... reversed, quiet. (*)
9	Openness to experiences	.... open to new experiences, complex.
10		.... conventional, uncreative. (*)

*denotes reversed questions. These questions are on 7 Likert scale

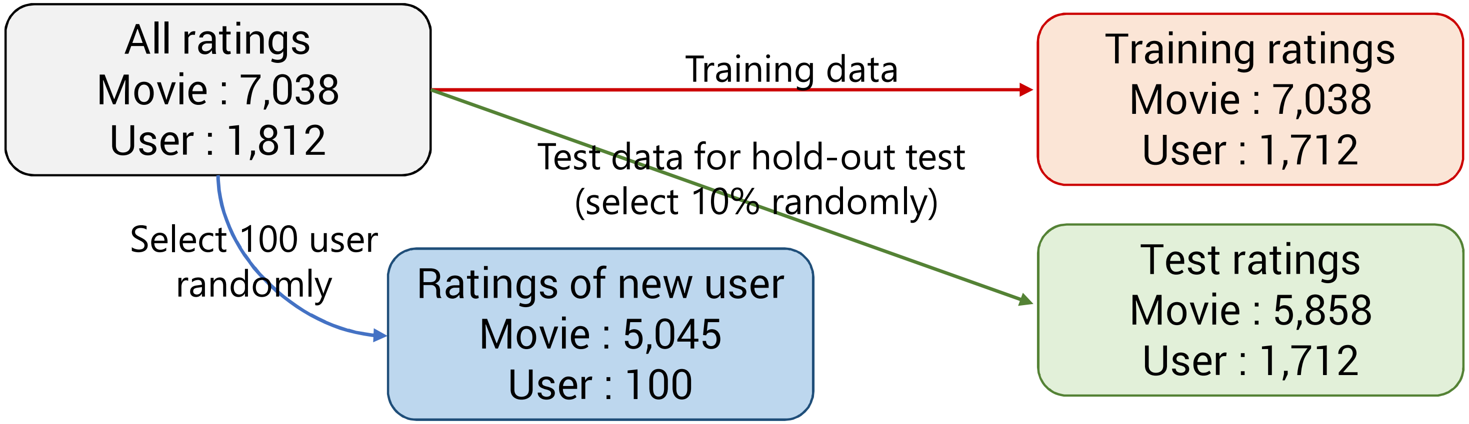
Data for empirical study

Stage 1

- ➢ Deleted tag which occurs less than 2.
- ➢ Deleted tag about name of actor or director.

Number of tag	Vocabulary
71,060	4,042

Stage 2



All ratings
Movie : 7,038
User : 1,812

Training data

Training ratings
Movie : 7,038
User : 1,712

Test data for hold-out test
(select 10% randomly)

Test ratings
Movie : 5,858
User : 1,712

Select 100 user
randomly

Ratings of new user
Movie : 5,045
User : 100

Data for empirical study

Stage 2

- I removed users who evaluated the same movie more than 2 times from all ratings.

  So, rating r_{ij}(rating of movie j by user i) is one.
- Total number of each groups of ratings are as follows.

	Number of rating	Number of movie	Number of users
All	630,149	7,038	1,812
Training ratings	567,134	7,038	1,712
Test ratings	63,015	5,858	1,712
Ratings of new user	38,908	5,045	100

Setting for empirical study

Stage 1

- Setting genre's names as seed words. Ex) Word "Action" → always go to "topic 1"
- Deciding total number of topics as 35 because of interpretability ($K = 35$)
- Sampling 2000 times and using parameter at last sampling
- Setting hyper parameters as $\alpha = \frac{1}{K}, \beta = 0.1$

Stage 2

- Sampling 10000 times and using mean of parameters from 5000 to 10000.
- Setting hyper parameters as

$$\bar{\boldsymbol{\Delta}} = \mathbf{0},\ \boldsymbol{A}_{\boldsymbol{\Delta}} = 0.01 * \boldsymbol{I}_{\boldsymbol{p}},\ \nu_{0e} = 3.0,\ s_{0e} = Var(\boldsymbol{r}),\ \nu_{0l} = K + 3, \nu_{0g} = p + 3,\ \boldsymbol{\Lambda}_{\mathbf{0}} = \nu_{0l} * \boldsymbol{I}_{\boldsymbol{K}},$$

$$\boldsymbol{\Gamma}_{\mathbf{0}} = \nu_{0g} * \boldsymbol{I}_{\boldsymbol{p}}$$

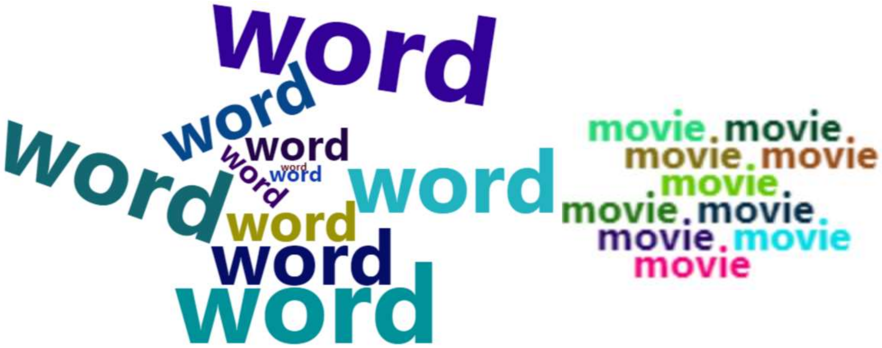
Result of empirical study

Stage 1

- ➢ Interpretation of topics

↓Top 40 words having high ϕ_k

↓Top 10 movies having high θ_k



Topic k
“meaning of topic”

Result of empirical study

Stage 1

- <u>Interpretation of topics</u>

Topic 3
"Animation(Disney)"

Result of empirical study

Stage 1

- ➢ Interpretation of topics

Notting Hill (1999)
Rumor Has It... (2005)
Sweet Home Alabama (2002)
While You Were Sleeping (1995)
Sleepless in Seattle (1993)
You've Got Mail (1998)
Spanglish (2004)
Holiday, The (2006)
Elizabethtown (2005)
Family Stone, The (2005)

Topic 14
"Romance"

Result of empirical study

Stage 1

- Interpretation of topics

Going Places (Les Valseuses) (1974)
History of the World: Part I (1981)
Love and Death (1975)
After Hours (1985)
Day for Night (Nuit Amテゥricaine, La) (1973)
Clean Slate (Coup de Torchon) (1981)
Mr. Hulot's Holiday (Les Vacances de M. Hulot) (1953)
Irma Vep (1996)
Mystery Train (1989)
Woman Is a Woman, A (Une femme est une femme) (1961)

Topic 20
"With message (Satirically)"

Result of empirical study

Stage 1

- Interpretation of topics

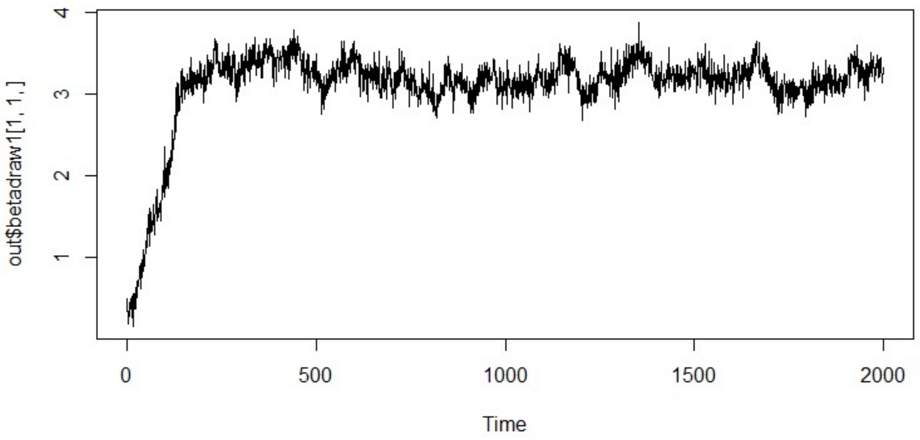
Hoosiers (1986)
Girlfight (2000)
Rudy (1993) Rocky III (1982)
World's Fastest Indian, The (2005)
Invincible (2006)
Bad News Bears, The (1976)
Rookie, The (2002)
Glory Road (2006)
Hidalgo (2004)

Topic 31
"Sports"

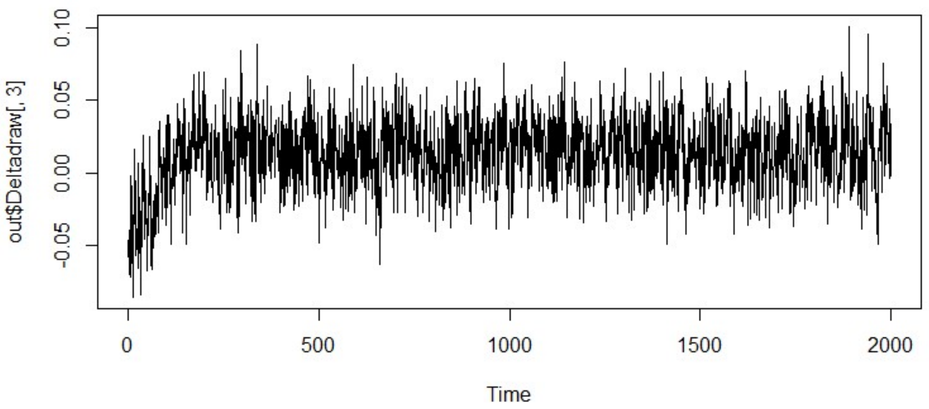
Result of empirical study

Stage 2

- Check of sampling convergence
  - Examples of convergence of parameters are as follows.



β



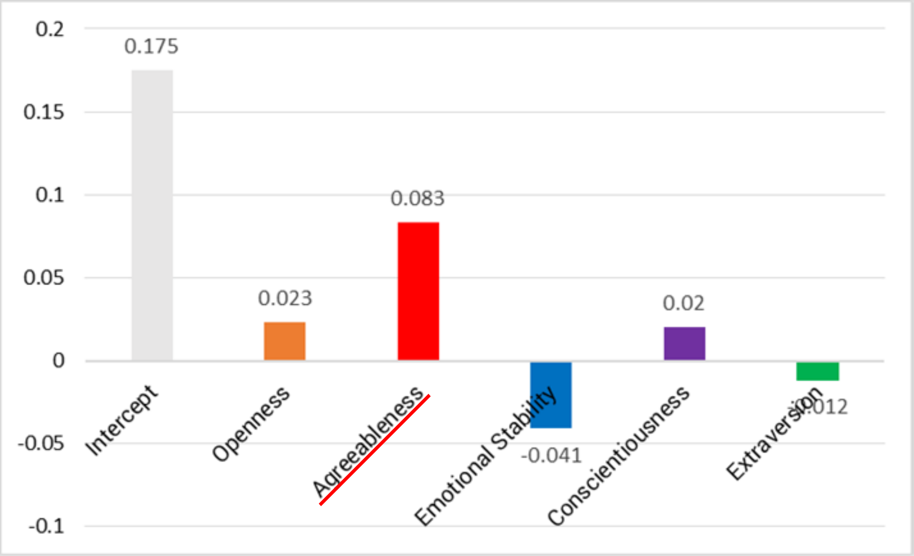
Δ

Result of empirical study

Stage 2

- ➢ <u>Check the rough preference for each personalities</u>

Topic 3 “Animation (Disney)”



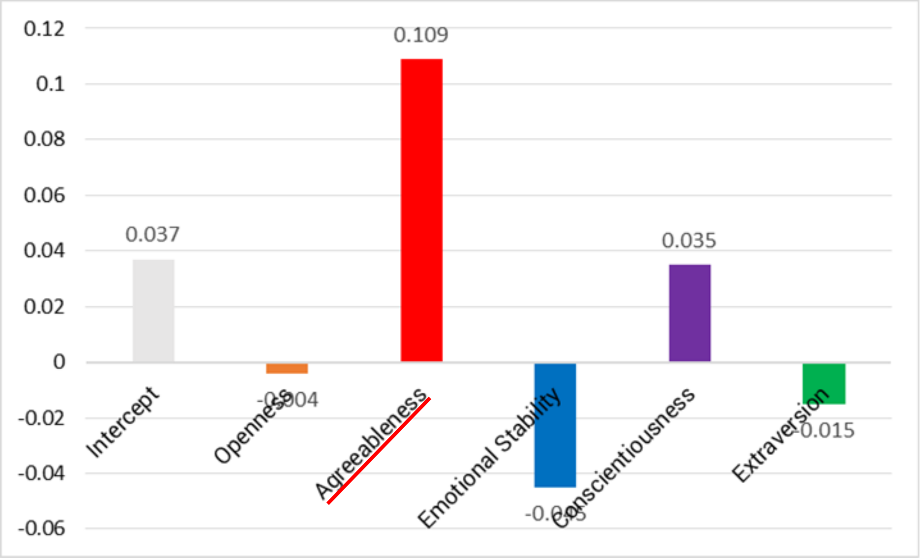
※red underline means significant at 95% level

Result of empirical study

Stage 2

- Check the rough preference for each personalities

Topic 14 “Romance”



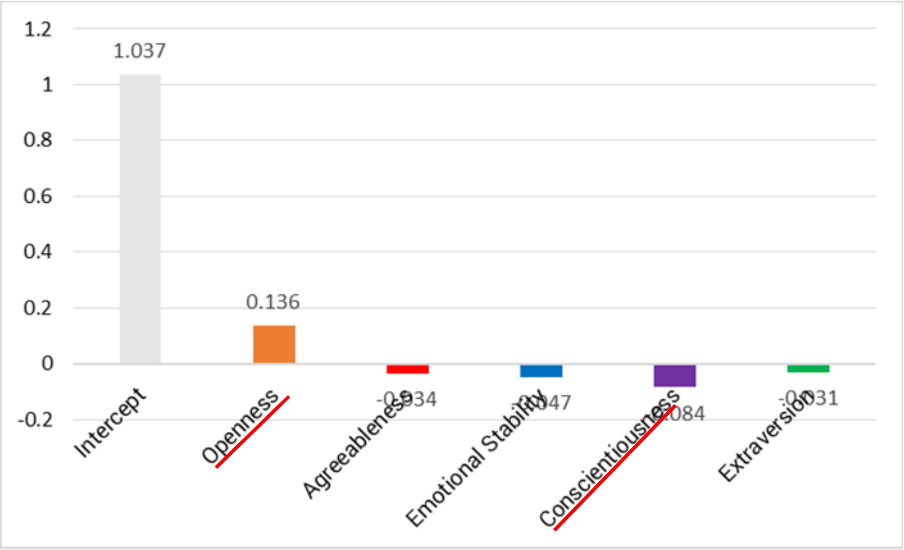
※red underline means significant at 95% level

Result of empirical study

Stage 2

- <u>Check the rough preference for each personalities</u>

Topic 20 “With message (Satirically)”



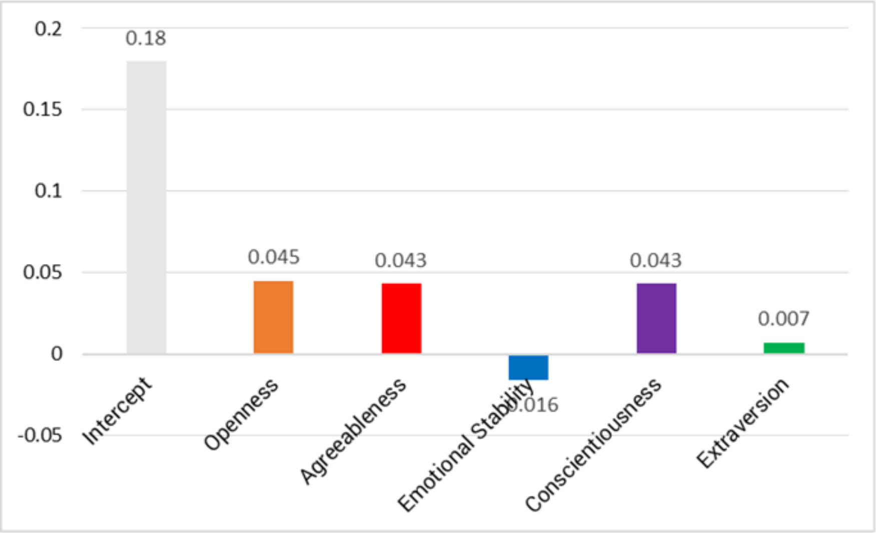
※red underline means significant at 95% level

Result of empirical study

Stage 2

- ➢ Check the rough preference for each personalities

Topic 31 “Sports”



※red underline means significant at 95% level

Result of empirical study

Stage 2

- Hold-out test for existing user
  - Model 1: model of this research
  - Model 2 : $r_{ij} = \boldsymbol{G_j}'\boldsymbol{\beta_i} + \boldsymbol{X_i}'\boldsymbol{\beta_j} + \boldsymbol{\epsilon_{ij}} \quad \boldsymbol{\epsilon_{ij}} \sim N_n(0, \sigma^2)$

    $\boldsymbol{\beta_i} = \boldsymbol{\Delta X_i} + \boldsymbol{\lambda_i} \quad \boldsymbol{\eta_i} \sim N_K(0, \boldsymbol{\Lambda})$

    $\boldsymbol{\beta_j} = \boldsymbol{\gamma_j} \quad \boldsymbol{\gamma_j} \sim N_p(0, \boldsymbol{\Gamma})$

$\boldsymbol{G_j}$: Genres of product $\boldsymbol{j}$ $\left(\boldsymbol{G_j} \in \mathbb{R}^{18}\right)$

These are dummy variable

    - Genres are following 18 genres.
    - *“Action“, "Adventure“, "Animation“, "Children's“, "Comedy”,"Crime”, "Documentary“, "Drama“, "Fantasy“, "Film-Noir“, "Horror“, "Musical“, "Mystery“, "Romance“, "Sci-Fi“, "Thriller“, "War“, "Western”*

  - Model 3 : $r_{ij} = \boldsymbol{\theta_j}'\boldsymbol{\beta_i} + \boldsymbol{X_i}'\boldsymbol{\beta_j} + \boldsymbol{\epsilon_{ij}} \quad \boldsymbol{\epsilon_{ij}} \sim N_n(0, \sigma^2)$

    $\boldsymbol{\beta_i} = \boldsymbol{\Delta X_i} + \boldsymbol{\lambda_i} \quad \boldsymbol{\eta_i} \sim N_K(0, \boldsymbol{\Lambda})$

    ⇨ Not consider product heterogeneity

  - Model 4 : $r_{ij} = \boldsymbol{G_j}'\boldsymbol{\beta_i} + \boldsymbol{X_i}'\boldsymbol{\beta_j} + \boldsymbol{\epsilon_{ij}} \quad \boldsymbol{\epsilon_{ij}} \sim N_n(0, \sigma^2)$

    $\boldsymbol{\beta_i} = \boldsymbol{\Delta X_i} + \boldsymbol{\lambda_i} \quad \boldsymbol{\eta_i} \sim N_K(0, \boldsymbol{\Lambda})$

Result of empirical study

Stage 2

- Hold-out test for existing user

	RMSE
Model 1	0.7854
Model 2	0.8167
Model 3	0.8544
Model 4	0.9110

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}$$

$where\ y_i$: actual value

$\widehat{y}_i$: predict value

Model 1 shows the highest accuracy.
So, topic proportion is better than “genre”.

Result of empirical study

Stage 2

- Hold-out test for new user
  - Model 1: model of this research
  - Model 2 : $r_{ij} = \boldsymbol{\theta}_{\boldsymbol{j}}{}'\boldsymbol{\beta}_{\boldsymbol{i}} + \boldsymbol{X}_{\boldsymbol{i}}{}'\boldsymbol{\beta}_{\boldsymbol{j}} + \boldsymbol{\epsilon}_{\boldsymbol{ij}} \quad \boldsymbol{\epsilon}_{\boldsymbol{ij}} \sim N_n\left(0, \sigma^2\right)$

$$\boldsymbol{\beta}_{\boldsymbol{i}} = \overline{\boldsymbol{\beta}} + \boldsymbol{\lambda}_{\boldsymbol{i}} \quad \boldsymbol{\eta}_{\boldsymbol{i}} \sim N_K(0, \boldsymbol{\Lambda})$$

$$\boldsymbol{\beta}_{\boldsymbol{j}} = \boldsymbol{\gamma}_{\boldsymbol{j}} \quad \boldsymbol{\gamma}_{\boldsymbol{j}} \sim N_p(0, \boldsymbol{\Gamma})$$

→ Not consider interaction effect

	RMSE
Model 1	0.8546
Model 2	1.037

Considering interaction effect improved recommendation accuracy for new user

Conclusion and challenge, future work

- Through extracting feature of products which cannot be get from “genre” by topic model and modeling rating by hierarchical regression model, I could improve recommendation accuracy.
- Existence of interaction effect between extracted feature and user’s personality was revealed, and it could improve recommendation accuracy for new user.

【Challenge】

- There are topics which are hard to interpret.
- Primitive topic model → more appropriate model.
- Automating decision of number of topics K using information criterion.

【Future work】

- Simultaneous estimation of topics and parameters in regression model.
  Ex) Ansari and Zhang (2018)

Reference

- Asim Ansari, Skander Essegaier, Rajeev Kohli (2000), “Internet Recommendation Systems” *Journal of marketing research Vol.37 pp. 363-375*
- Asim Ansari, Yang Li, Jonathan Z, Zhang. (2018) “Probabilistic Topic Model for Hybrid Recommender Systems: A Stochastic Variational Bayesian Approach” *Marketing Science*
- Joachim Büschken, Greg M. Allenby. (2016) “Sentence-Based Text Analysis for Customer Reviews“ *Marketing Science*
- Oliver Toubia, Garud lyengar, Renee Bunnell, and Alain Lemaire. (2019) “Extracting Features of Entertainment Products: A Guided Latent Dirichlet Allocation Approach Informed by the Psychology of Media Consumption” *Journal of Marketing Research*
- Peter J. Rentfrow, Lewis R. Goldberg, Ran Zilca (2011) “Listening, Watching, and Reading: The Structure and Correlates of Entertainment Preferences” *Journal of Personality Vol 79, pp.223-258*
- Tien T. Nguyen, F. Maxwell Harper, Loren Terveen, Joseph A. Konstan (2018) “User Personality and User Satisfaction with Recommender Systems” *Information systems Frontiers Volume 20 pp 1173–1189*
- Weike Pan (2015) “ A survey of transfer learning for collaborative recommendation with auxiliary data“ *Neurocomputing vol.177,pp447-453*
- 神嶌敏弘 (2007), “推薦システムのアルゴリズム” *人工知能学会誌 Vol.22 No.6*