



後処理による独立性強化型推薦システム

神嶌 敏弘, 赤穂 昭太郎

産業技術総合研究所

2016年度人工知能学会全国大会 (第30回) @ 北九州市, 2016.6.6

<http://www.kamishima.net/>

概要

推薦の独立・公正性

- ✿ 法や規定を遵守した推薦内容の決定
- ✿ 推薦対象となる情報の提供者の公平な扱い
- ✿ 利用者が望まない情報の影響の排除



独立性強化型推薦システム

みにくいアヒルの子の定理などから、
あらゆる情報に対して独立な推薦は不可能



利用者などが指定した、特定のセンシティブ情報に対して、
推薦結果ができるだけ独立性を保つように配慮する推薦システム

目次

推薦独立性

- ✿ センシティブ変数, 推薦独立性, 推薦の独立性強化の例

推薦独立性の応用例

- ✿ 法や規定の遵守, 利用者が望まない情報の影響の排除, 情報提供者の公平性

独立性強化型推薦システム

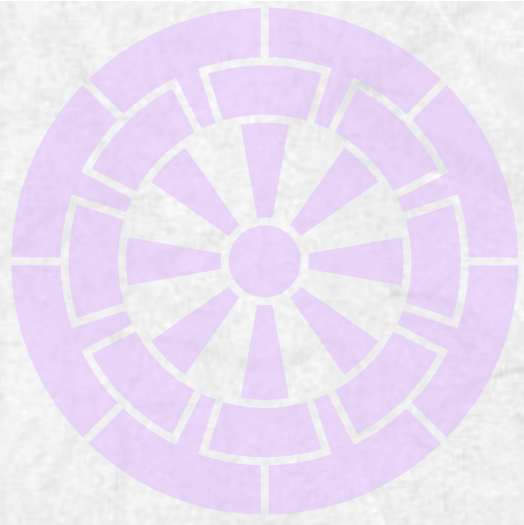
- ✿ 独立性強化型推薦システム, 公正配慮型予測手法の分類, 中処理型 (既存手法), 後処理型 (線形分布変換法)

実験

- ✿ 6種類のデータについて, 中処理型と後処理型の予測精度と独立性を比較

関連研究

まとめ



推薦獨立性



センシティブ特徴

標準的な推薦モデルの確率変数

X : 利用者, Y : アイテム, R : 評価値



推薦独立性のために新たな変数を導入

S : センシティブ特徴

- * 利用者が自身の目的に応じて指定する
- * この指定したセンシティブ特徴と推薦結果は独立になるように
- * 特徴の値は利用者やアイテム, およびこれらの特徴に依存する

例 : センシティブ特徴 = 利用者の性別・映画の公開年

本研究ではセンシティブ特徴が二値の場合のみを対象とする

推薦独立性

推薦独立性

- ✳ 指定したセンシティブ特徴に対し，その特徴の情報が推薦結果に全く影響しないとき，その推薦結果は独立である
- ✳ センシティブ特徴の状態は明示的に推薦結果の推論から除外する

例： センシティブ特徴 = 映画の公開年



映画が古いか新しいかということが
その映画を推薦するかどうかという判断に影響しない



もし公開年以外の条件が全く同じ二つの映画 A と B があった場合，
映画 B が推薦されたときは映画 A も必ず推薦され，
またこの逆も成り立つ

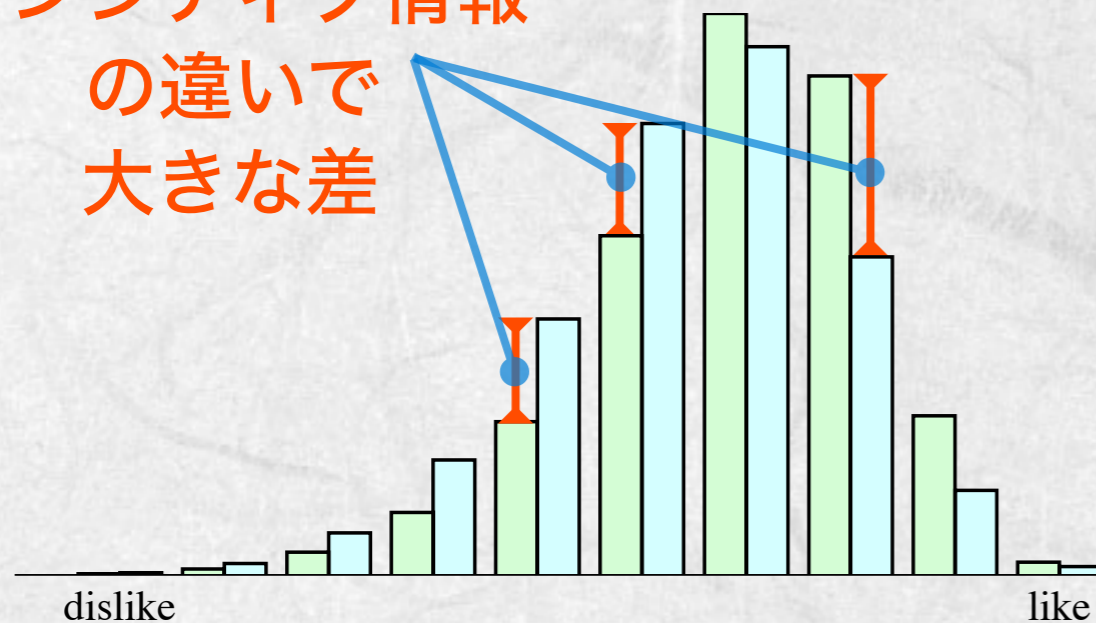
推薦の独立性強化の例

古い映画 と 新しい映画 では古い映画の高く評価されやすい
映画の新旧をセンシティブ特徴として選んだ独立性強化型推薦を実行

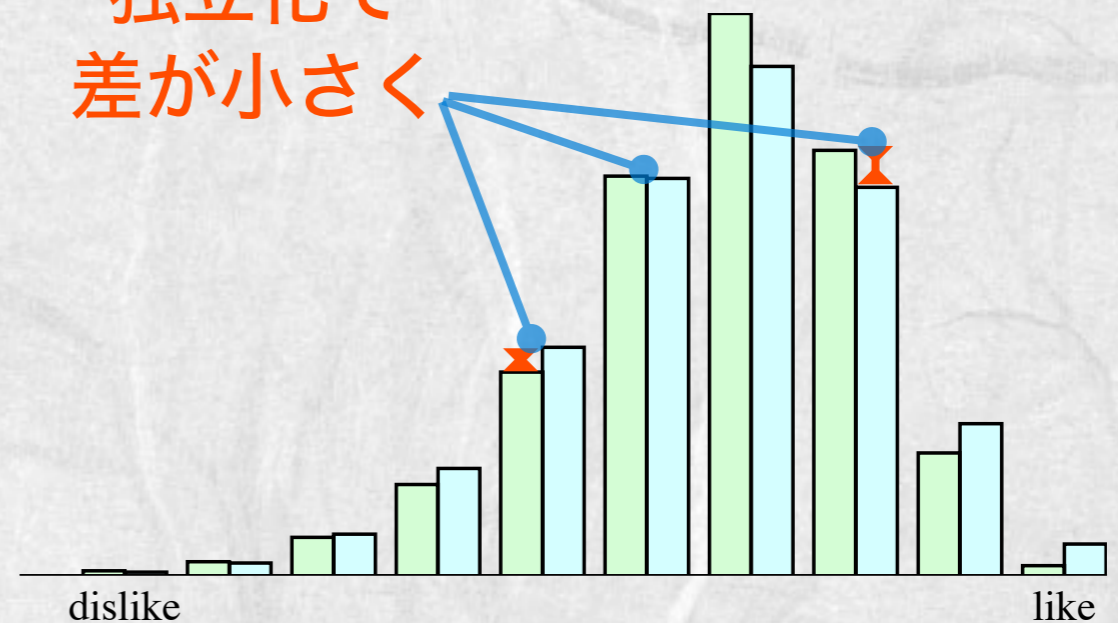
標準的な手法による予測

独立性強化型による予測

センシティブ情報
の違いで
大きな差



独立化で
差が小さく

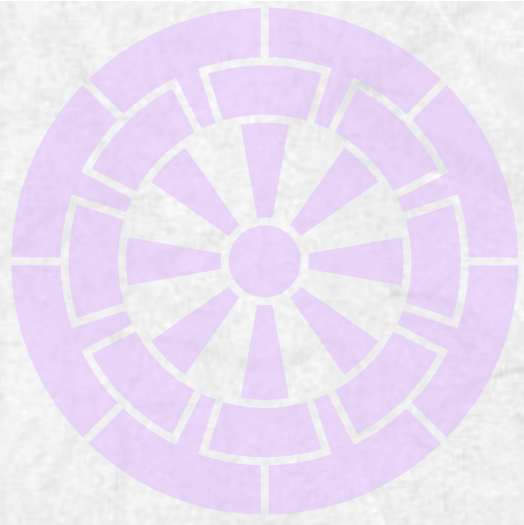


※ 予測評価値のヒストグラムの各ビンを **古い映画** と **新しい映画** で並べた

独立化により、違うセンシティブ情報でも評価値の分布の差が小さい



センシティブ情報に推薦結果が影響されず独立性が強化できている



推薦独立性の応用例



法や規定の遵守

[Sweeney 13]

推薦サービスを法や規定に従って運用する必要性

キーワードマッチ広告の配信

人名で検索したとき、ヨーロッパ系由来の名前より、アフリカ系由来の名前に対して、より頻繁に逮捕歴を示唆するテキストと共に、逮捕歴検索サイトの広告が表示された



社会的に差別的な扱いは回避されるべき

センシティブ特徴 = 対象者のセンシティブな個人属性情報



法的・社会的にセンシティブな情報を排除して推薦することが可能

利用者が望まない情報の影響の排除

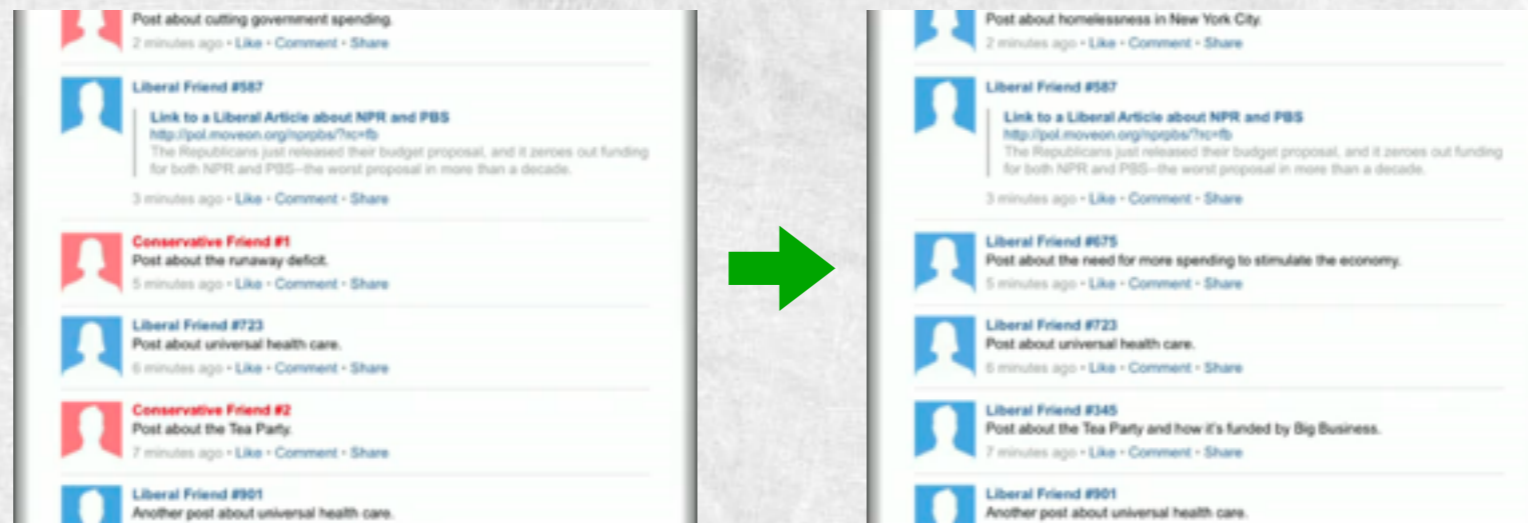
[TED Talk by Eli Pariser]

推薦結果に影響することを利用者が望まない情報を排除する

フィルターバブル問題：個人化技術によって利用者が多様な情報源や視点に接する機会が減らされているとの Eli Pariser による主張

Facebookの友人推薦リスト

進歩派の友人を増やしていると、利用者に断りなく保守派が消された



センシティブ特徴 = 友人候補の政治的態度



友人リストに加えるかどうかの判断で政治的態度の情報を除外

情報提供者の公平性

推薦システム運営者が情報提供者を公平に扱う必要性

検索結果の表示順位

2012年11月のBloombergの記事 (<http://bloom.bg/PPNEaS>) によれば、自身のサービスを競合企業と同種のサービスより上位に表示しているとの点についてFTCの調査を受けた

情報提供者の公平な扱い

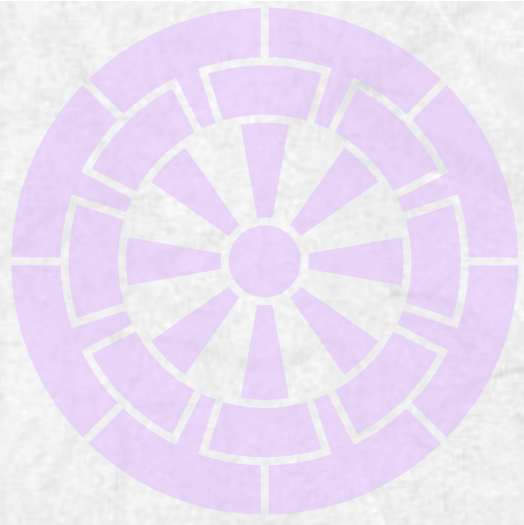
小売店をまとめたオンラインモールや、顧客が提供する情報を掲載する不動産・求職サイトでは、コンテンツ提供者の公平性に配慮が必要



センシティブ特徴 = 推薦候補アイテムの情報提供者



候補の提供者の情報を推薦から無視することで提供者を公平に扱う



独立性強化型推薦システム



独立性強化型推薦システム

独立性強化型推薦システム Independence-Enhanced Recommender System

指定したセンシティブ特徴からの独立性
推薦独立性を保つための制約項を導入

+

高い予測精度
経験誤差の最小化によって高い予測精度を達成



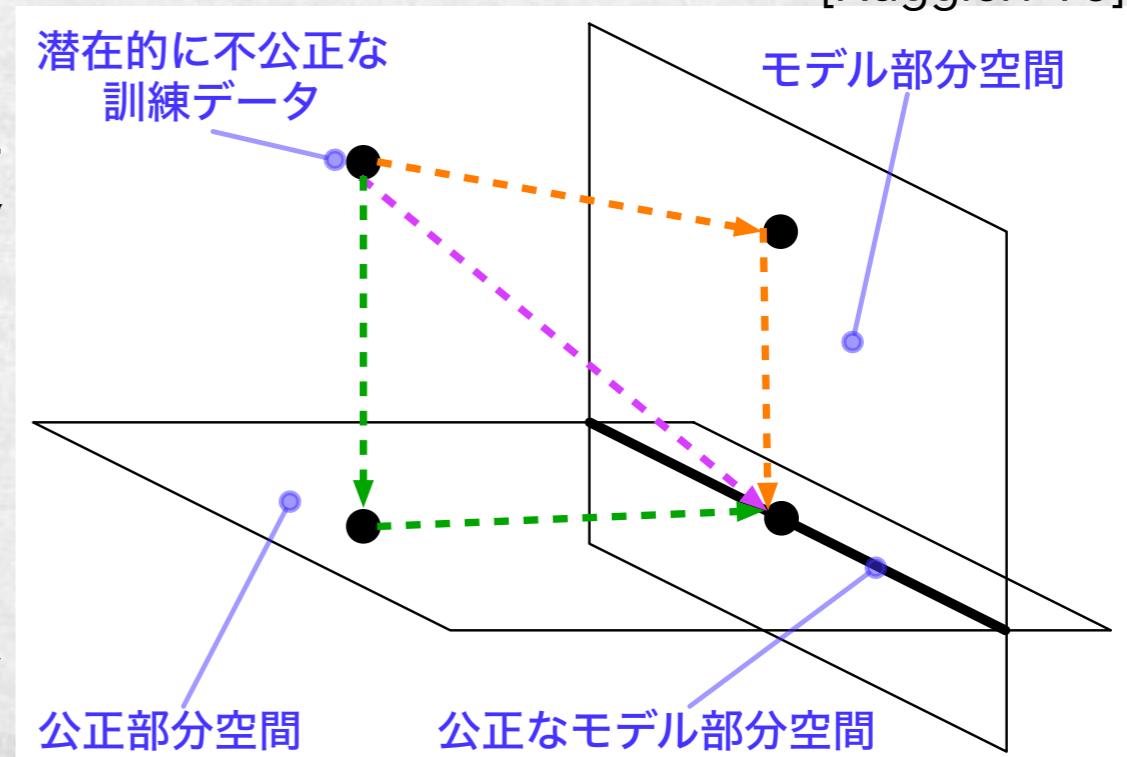
確率的行列分解モデルの独立性を強化したアルゴリズムを考案

公正配慮型予測手法の分類

[Ruggieri 10]

公正配慮型予測

- ★ 予測する目的変数とセンシティブ変数の独立性の制約下で正確な予測
- ★ 潜在的に不公正な訓練データから、公正なモデル空間 (RLSを満たす) で最良のモデルを学習するのが目標



公正配慮型予測手法の分類

- ★ **前処理型 (pre-process)** : 訓練データを公正性を満たすように変換してから、普通の予測手法を適用する (図中の緑色)
- ★ **中処理型 (in-process)** : 潜在的に不公正な訓練データから直接的に公正な予測モデルを獲得する (図中の紫色)
- ★ **後処理型 (post-process)** : 普通の予測手法でモデルを獲得したのち、公正性の制約を満たすように修正する (図中のオレンジ色)

中処理型 (既存手法)

[Kamishima 13]

標準的な推薦モデルを拡張した独立性強化型推薦モデルの目的関数

独立性パラメータ
独立性と予測精度のバランスを調整

正則化
パラメータ

$$\sum_s \sum_{r \in D^{(s)}} \text{loss}(r, \hat{r}^{(s)}(x, y)) - \eta \text{ind}(R, S) + \lambda \|\Theta\|_2^2$$

センシティブ変数ごとの損失関数の和

独立性項

L₂正則化項

この目的関数を最小化するようにパラメータを学習

独立性項 $\text{ind}(R, S)$: 推薦独立性を評価する項

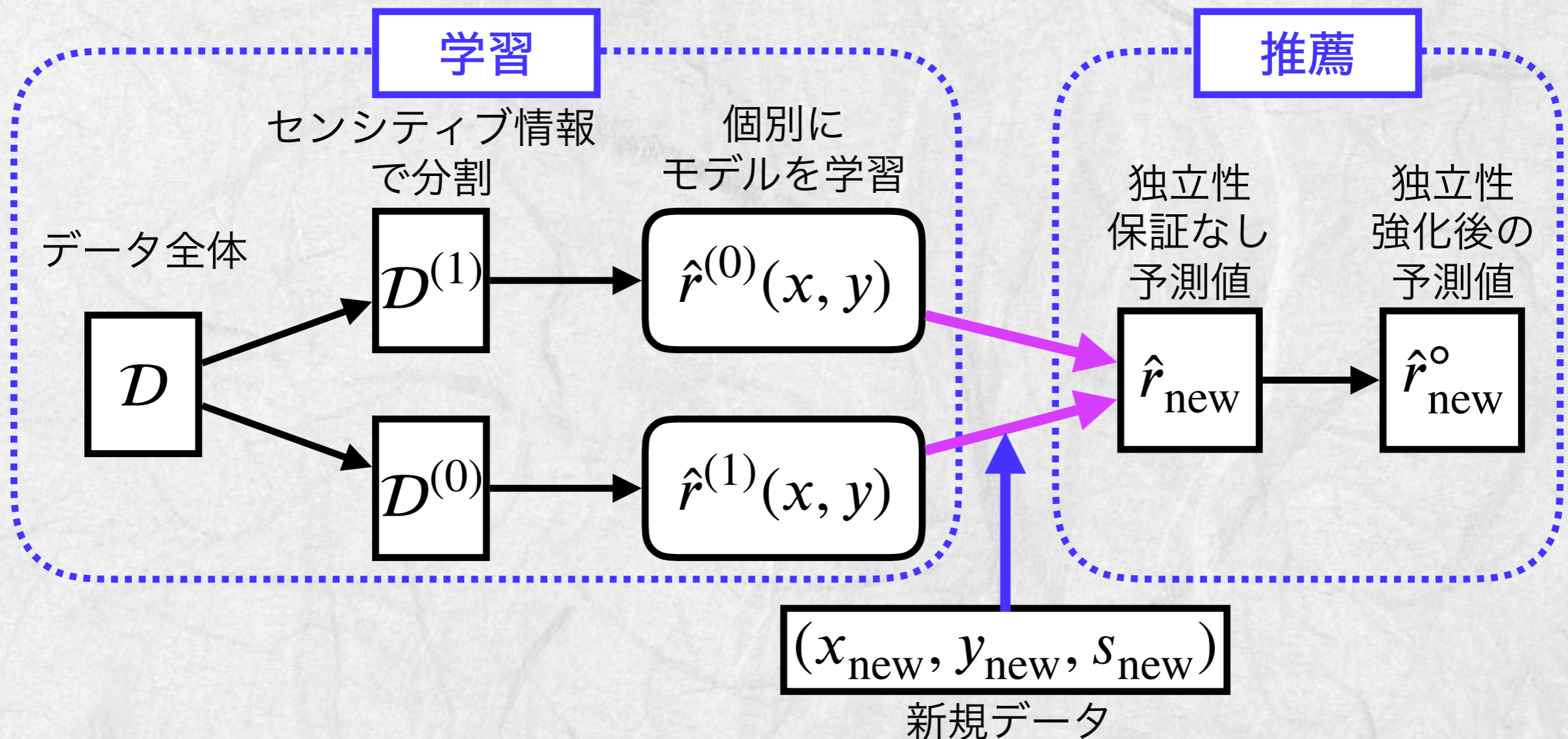
センシティブ特徴が 0 と 1 の場合の予測評価値の平均を一致させる

$$-\left(\text{mean}^{(0)}(\hat{r}) - \text{mean}^{(1)}(\hat{r})\right)^2$$

後処理型 (線形分布変換法)

簡単なベースラインでどれくらいの精度・独立性が達成できるか検証

- * センシティブ情報ごとに通常の推薦モデルを学習
- * 各推薦モデルの平均と標準偏差を線形変換で一致させる



後処理型（線形分布変換法）

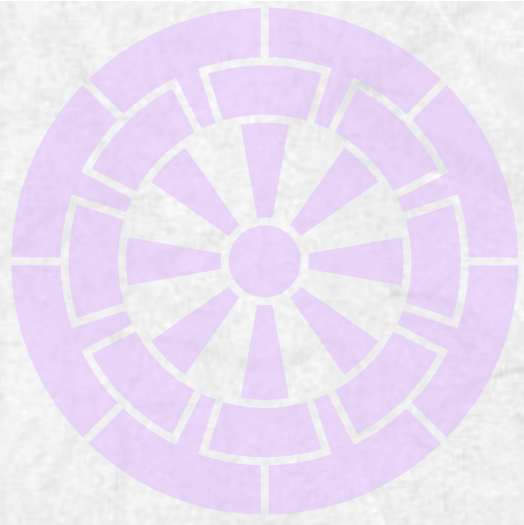
センシティブ特徴値ごとの予測評価値の平均と標準偏差を
全体の平均と標準偏差に一致させるように線形変換

$$\hat{r}_{\text{LDT}}^{\circ}(x, y, s) = \frac{\text{stdev}(\hat{r})}{\text{stdev}^{(s)}(\hat{r})} \left(\hat{r}(x, y, s) - \text{mean}^{(s)}(\hat{r}) \right) + \text{mean}(\hat{r})$$

全予測評価値の分布の標準偏差
変換前の予測評価値
全予測評価値の分布の平均値

独立性強化後の予測評価値
センシティブ情報が s の場合の予測評価値の分布の標準偏差
センシティブ情報が s の場合の予測評価値の分布の平均値

※ 訓練データの平均値や標準偏差は、データ中の真の評価値ではなく、訓練データに対する予測評価値に対して求める



実験結果



実験条件

データ集合：合計6種類

- ★ Movielens 1M：センシティブ情報={利用者性別, 公開年}
- ★ Flixster：センシティブ情報={映画の人気度}
- ★ Sushi：センシティブ情報={海産物, 利用者は10代, 性別}

その他の条件

- ★ 中処理型独立性パラメータ $\eta=100$ (独立性を非常に強化した状態)
- ★ 5分割交差確認

評価尺度

- ★ 予測精度：MAE (平均絶対誤差) 小→良
- ★ 独立性：KS (2標本Kolmogorov-Smirnov検定統計量) 小→良
 - ※ 経験累積分布関数の差の面積

実験結果 Movielens / Flixster

データ	指標	標準	中処理型	後処理型
Movielens Year	MAE	0.687	0.697	0.696
	KS	0.1662	0.0271	0.0174
Movielens Gender	MAE	0.691	0.694	0.692
	KS	0.0319	0.0050	0.0047
Flixster	MAE	0.645	0.653	0.650
	KS	0.1416	0.0165	0.0280

- ✳ 中処理型と後処理型を比較したとき、予測精度が良い方を青字で、独立性・公正性が高い方を赤字で示した

実験結果 Sushi

データ	指標	標準	中処理型(既存)	後処理型(新)
Sushi Age	MAE	0.918	0.926	0.915
	KS	0.2445	0.0256	0.0462
Sushi Gender	MAE	0.922	0.925	0.925
	KS	0.0739	0.0321	0.0235
Sushi Seafood	MAE	0.908	0.925	0.912
	KS	0.3306	0.2002	0.2991

- ✿ 中処理型と後処理型を比較したとき、予測精度が良い方を青字で、独立性・公正性が高い方を赤字で示した

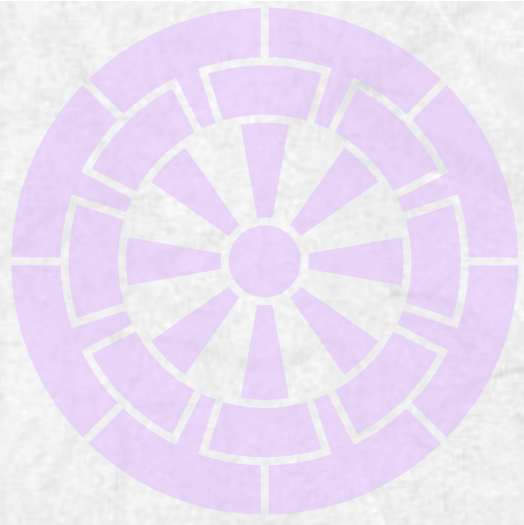
実験結果 まとめ

実験結果の概要

- * 後処理型の線形分布変換法は
 - * 予測精度では，中処理型の平均一致法より **やや有利**
 - * 独立性では，両者は **ほぼ同等**

その他の観点からのコメント

- * 中処理型は1次のモーメントしか考慮していないが，それでも2次モーメントまで考慮した後処理型と同等
- * 今回の後処理型は非常に簡潔な方法だが，それでもなかなかの精度
 - * 他の分類問題などにおいても簡潔な後処理法は強力



関連研究



推薦の多様性

[Ziegler+ 05]

話題の多様化 (Topic Diversification)

利用者と嗜好が一致するアイテムを選びつつ、推薦リストから類似したアイテムは排除することで、推薦アイテムの多様性を確保

類似度関数：アイテムの内容に基づいてアイテム間の類似度を測る
アイテムの階層的な分類があるとき、その構造木中でのホップ数の逆数などで測る

多様性

アイテムが推薦リスト中や推薦時系列上にのアイテムが互いに類似していない

推薦結果間の関係

独立性

センシティブ特徴の情報を排除するだけで、推薦アイテム間の類似性は関係ない

推薦とセンシティブ特徴の関係

プライバシー保護データマイニング

評価値変数 R とセンシティブ特徴 S の独立性

評価値変数 R とセンシティブ特徴 V の相互情報量が 0



プライバシー保護データマイニングの観点からの解釈

たとえ評価値 R の予測値を知られることがあっても、
センシティブ特徴 V の情報が漏洩しない

※ t -近接性と概念的には強い関連がある

まとめ

本発表の寄与

- ✳ 独立性強化型推薦で簡潔な後処理型の線形分布変換法を考案
- ✳ 中処理型の平均一致法と比較し，その性能を検証した
- ✳ 精度では後処理型はやや有利，独立性ではほぼ同等

今後の予定

- ✳ 1次モーメントしか考慮していない中処理法は，2次モーメント以上を考慮できるように拡張する
- ✳ 評価値予測ではなく，適合アイテムの推薦での独立性強化推薦手法の開発

謝辞

- ✳ Grouplens research lab とDr. Mohsen Jamali による実験データの提供に感謝する
- ✳ 本研究はJSPS科研費 16700157, 21500154, 24500194, および 15K00327 の助成を受けたものである