

行列分解：評価情報とハイブリッド化

利用者が評価したアイテムの情報

[Koren+ 09, Paterek 07]

利用者が、どのアイテムを評価したかという情報は重要

利用者 i に依存した
スカラーの係数

利用者 i が評価した
アイテムの集合

アイテム l の効果
 k 次元ベクトル

$$s_{ij} = (\mu + a_i + b_j + [\mathbf{u}_i + c_i \sum_{l \in \mathcal{R}_i} w_l]^\top \mathbf{v}_j)$$

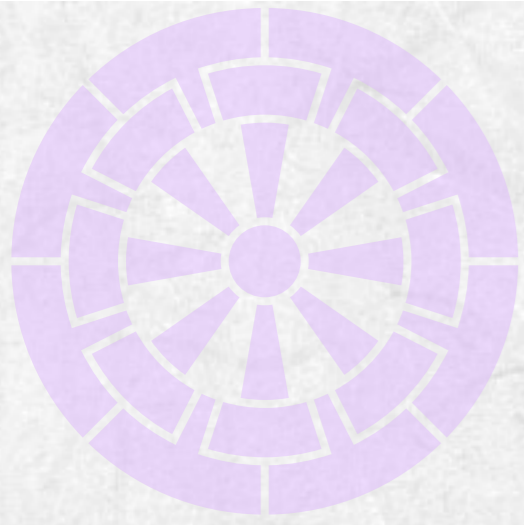
* [Koren+ 09] では c_i を $\sqrt{|\mathcal{R}_i|}$ にするヒューリスティックを導入

ハイブリッド化

アイテムや利用者の特徴も考慮可能

$$s_{ij} = (\mu + a_i + b_j + \mathbf{u}_i^\top \mathbf{v}_j + \mathbf{d}_j^\top \mathbf{x}_i + \mathbf{e}_i^\top \mathbf{y}_j)$$

- * \mathbf{x}_i : 利用者 i のデモグラフィックやそのコンテキストの特徴ベクトル
- * \mathbf{y}_j : アイテム j 自身やそのコンテキストの特徴ベクトル
- * $\mathbf{d}_j, \mathbf{e}_i$: 学習する重みベクトル



推薦システムのこれから



推薦システムは役に立つのか？

利用者にとって推薦システムは役立つ？



もちろん!

ただし、次の条件が必要！

推薦システム側
利用者が欲しいと
予測したもの

=

利用者側
利用者が
欲しいもの

推薦システムと利用者要求の一致

推薦システム側
利用者が欲しいと
予測したもの

=

利用者側
利用者が
欲しいもの

利用者が推薦システムに合わせる

- ✳ 利用者主導の推薦：推薦の方向付けを利用者が行う
- ✳ 推薦結果の選択：システムが開示した情報から推薦を取捨選択
- ✳ 推薦状況の限定：利用者が情報要求が自明な状況で利用

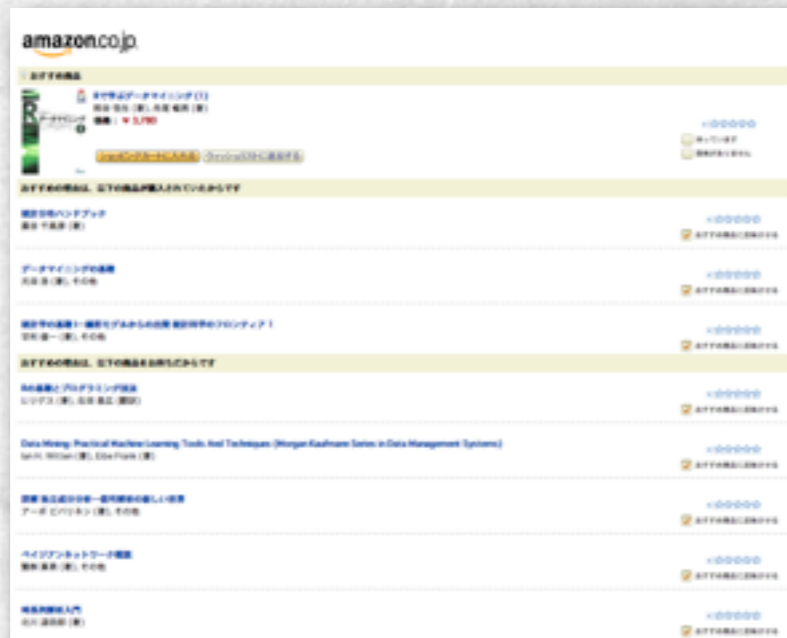
推薦システムが利用者に合わせる

- ✳ 入出力の多様化：評価値以外の情報をもっと活用
- ✳ 推薦内容の多様化：正解率以外の視点も考慮，より積極的な提案

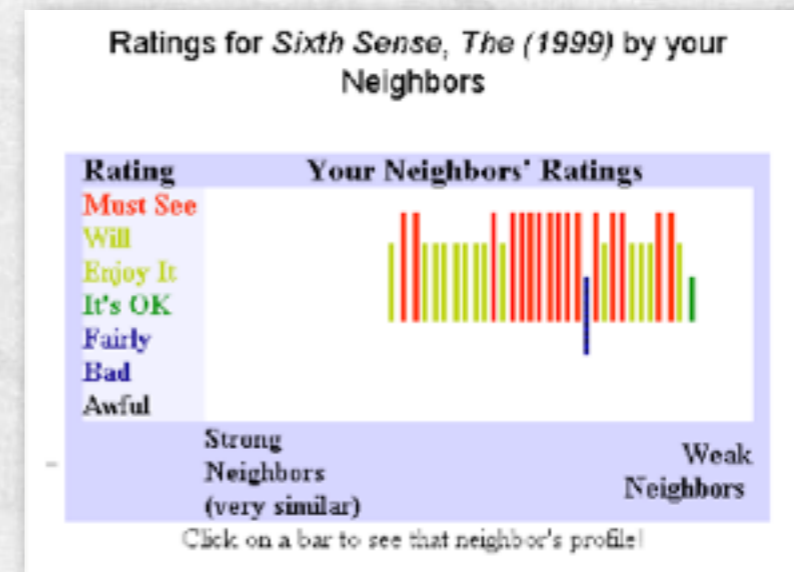
推薦理由の提示：ホワイトボックス

推薦そのものだけでなく推薦の理由も提示して，利用者が推薦を信用し受け入れやすくする [Sinha+ 02]

ホワイトボックス：推薦のプロセス自体を見せる



2007/7/26 にスクリーンショットを取得



例1：アイテム間型近傍法で，過去に購入したアイテムと類似していることを明示する

[Linden+ 03]

例2：利用者間型メモリベース法で，嗜好が似ている人の評価を示す

[Herlocker+ 00]

入出力の多様化

内容ベースフィルタリングの弱点 アイテムのデータベース構築は高コスト

- * DBが整備されているのは書籍・CDなどの限られた品目
- * アイテムの情報を提示すると意志決定にとって重要 [Swearingen 01]
- * 高コストのDB整備には、それに見合う規模の店舗できない
- * DBにない商品はないも同然 → DBの寡占が流通の寡占に



推薦のためのデータベースを低コストで構築する技術

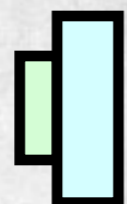
- * **入力情報の整備**：アイテム情報の Linked Data や Webサービス による共有，嗜好のゆらぎの補正，評価基準の多様化
- * **入力情報の高度化**：自然言語・画像・音声・音楽情報処理で抽出

入力情報の多様化

入力デバイスの多様化

屋外
晴れてる

ワンピース
を着ている

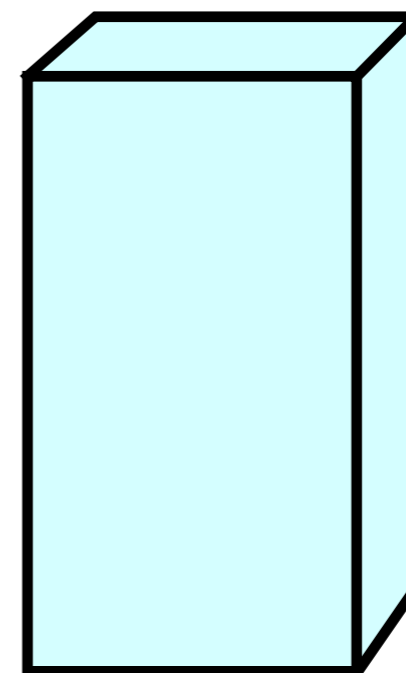


携帯・PC
のカメラ

こんな色が好き？

画像処理で環境や対象の識別に使われている特徴量を推薦の生成に利用

自然言語処理と画像処理との協調



10 × 40 × 90

高さ・奥行き・幅との対応付け

90cm幅の棚がほしいな



データベース情報の収集

集中的に大規模データベースを作る
製品情報の掲載や修正にタイムラグ
情報の寡占による流通の支配



分散管理と凝集サービス

メーカー・サービス事業者

- * 自社の製品・サービスの情報を linked data や Webサービスとして提供

データ加工・提供

- * 各種のデータベースを収集して、Webサービスで効率的に提供するサービス
- * 目的に合わせて語彙の変更や同義語の凝集などを行う

出力手法の高度化

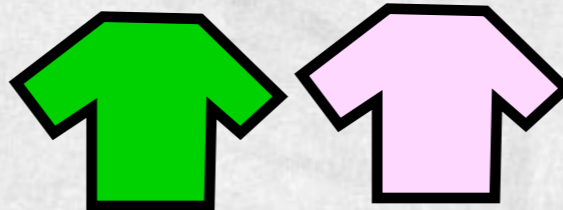
推薦の提示：★の数やアイテムの順位付けリスト
魅力的な提案ではない



魅力的な推薦の提示は意志決定に影響 [Swearingen+ 01, Cooke+ 02, Senecal+ 04, Sinha+ 02]

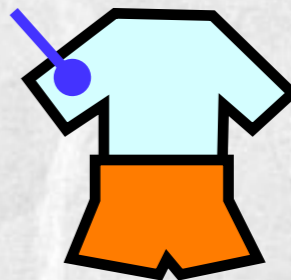


商品と合成して表示
(バーチャル試着)



携帯や電子掲示板
を利用した提示

現実の商品



コーディネート例など
ARを積極的に

- ✳ 利用者からのフィードバックをより活用したARや電子掲示板
- ✳ 小売店だけでなく、メーカーも直接顧客にアクセス
➡ 流通の流れが変わるかも？

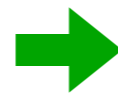
PRPの呪縛

PRP (Probability Ranking Principle)

[Robertson 77, Belew 00]

情報検索分野での古典的な原理

クエリへの適合確率で成立
した回答リスト



あらゆる回答の中で利用者
にとって最も効果的



これは推薦システムでも真理???

利用者の情報要求は不明確, もしくは存在しない
情報検索とは違った, 意志決定支援としての側面

- * **cherry picking** : 容易に分かるアイテムのみを推薦して予測精度を「かせぐ」のは安易 [Herlocker+ 04]
- * **収穫逓減の法則 (law of diminishing marginal returns)** : 類似したものを続けて与えられると効用は低下 [Ziegler+ 05]

推薦リスト全体を考えた推薦

アイテムごとに好き嫌いを判断し
それらを集めてリストにする



推薦はリスト全体をまとめて考慮すべき [McNee+ 06]

リスト中に含めるアイテムの選択によって

正解率

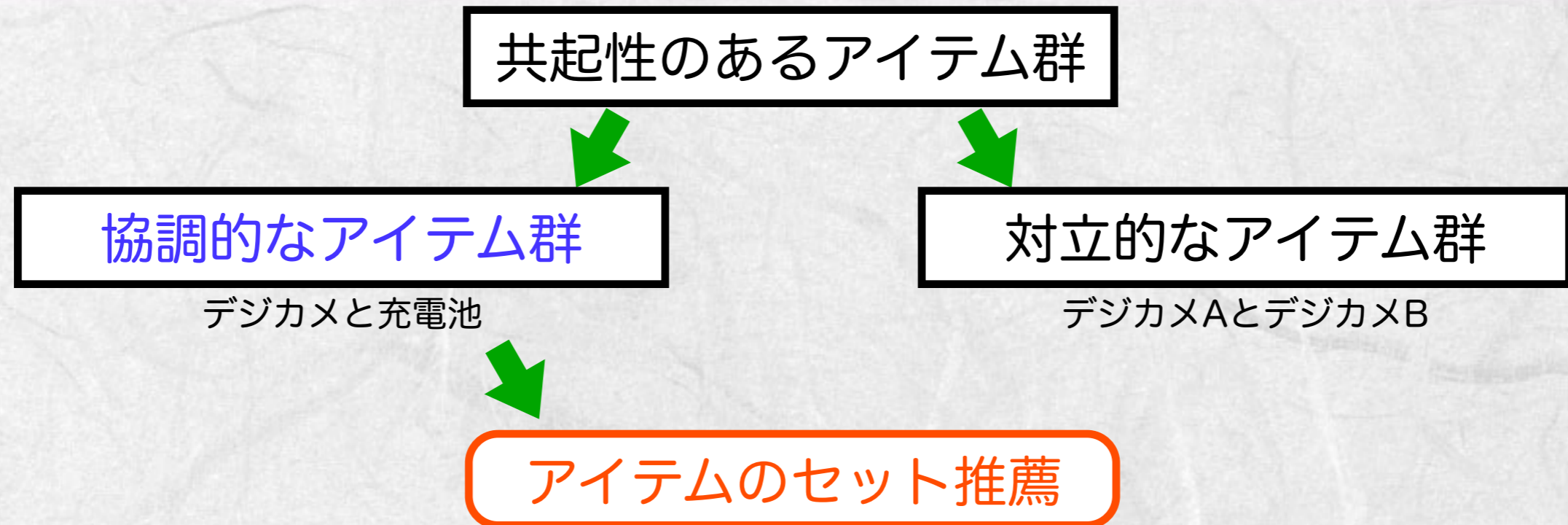


セレンディピティ

バランス

例：いろいろな利用者に対して一連の推薦をする場合を想定。予測嗜好スコアがしきい値の以上アイテムから、できるだけ異なるアイテムが推薦されるようにリストを作る

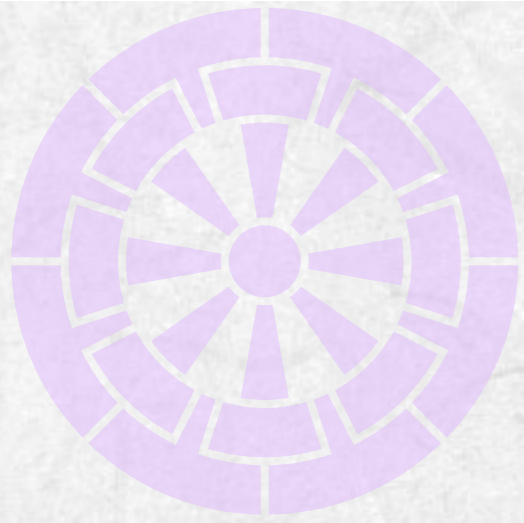
セット推薦



- * 協調的なアイテム群であるかの予測
- * セット販売のみの割引き価格の設定

共起性以外の観点からのアイテムのセット化

- * 嗜好パターンが類似している利用者の好むアイテム群と、アイテム自体の特徴を考慮して(シャツとジャケット, 柄+柄はダメ), 衣料品のコーディネート
- * 旅行先でのレストラン・レジャーのセット



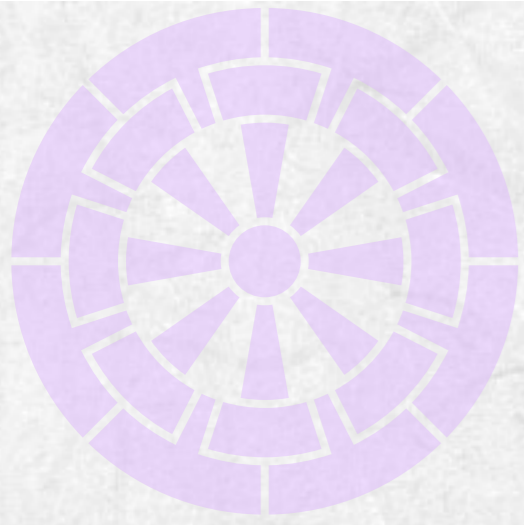
単純な予測精度の競争は終わった

Recommender System

推薦システム

より健全な運用と
より幅広い活用を可能に





Attendant System

おもてなしシステム

利用者の意図をくみ取って
より積極的で魅力的な提案を可能に



正誤表

- *P.72 : porfolio効果 → portfolio効果
- *P.77 と P.78 の損失関数 : 「 λ 」 → 「 $\lambda/2$ 」
- *P.123 : 「 I の列ベクトル v_j 」 → 「 V の列ベクトル v_j 」