

プログラミング基礎

12月16日（第9回）

担当：満行（PSI）

本日の話題

- Amazonって何をやっているの？
- 情報推薦システムを作ろう！

目次

- 情報推薦システムとは
- 情報推薦システムの分類
- 協調フィルタリングとは
- 協調フィルタリングの分類
- アルゴリズム
- DEMO

情報推薦システムとは

推薦システム (Recommender Systems) の原始的な定義

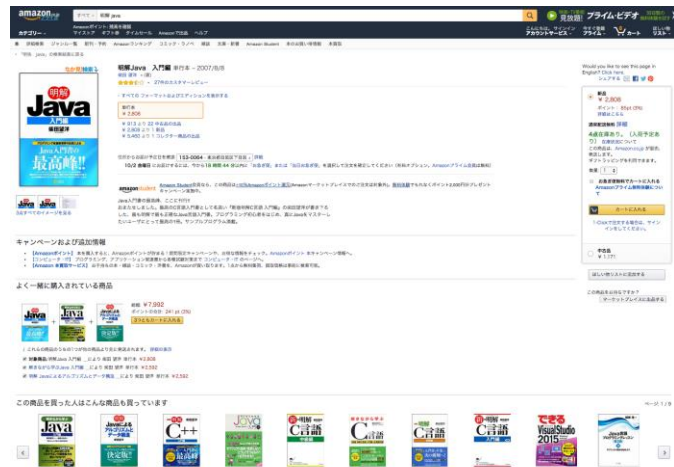
“推薦システムは、自分の経験 だけでは整理しきれない情報や経験したことの無い未知の情報に関して、他者からの評価を参考にするという行為を補助や促進を行うシステム”

Paul Resnick and Hal R Varian. Recommender systems. Communications of the ACM, Vol. 40, No. 3, pp. 56–58, 1997.

要は、
買い物や意思決定を行う際、日常的には行っている他人からアドバイスをもらうことや趣味や嗜好に合わせて情報収集を行うが、聞く相手がいない、自分の趣味がわからない、人の言っていることが信用できないといった人のために、様々なデータに基づいて嗜好をシステムティックに推定してレコメンドしてくれるシステム。

情報推薦システムとは

現在、ショッピングサイト(ex. amazon)や検索システム(ex. Google)などの多くのサービスやシステムに導入されている技術である。



Amazon



他のキーワード: Javaオブジェクト指向練習

Javaのオブジェクト指向入門 - KAB-studio
www.kab-studio.biz/Programming/OOPinJava/

「Javaのオブジェクト指向入門」は、Javaを使用したオブジェクト指向プログラミングの解説を掲載しているコンテンツです。クラスがよくわからない、オブジェクト指向って難しい、Javaとオブジェクト指向の関係がわからない、Javaに挫折しちゃった、という方。

「オブジェクト指向」はなぜ難しい? - 1.1 とりあえずクラスを作ってみよう - 1.2 クラスと参照

Javaのオブジェクト指向を 90 分で理解する - 偏見プログラ...

pg-kura.hatenablog.com/entry/20111222/1324557014

2011/12/22 - 1. 分からない。いくら説明を読んでも分からない。o 1.1. 未だに分からないJavaのオブジェクト指向 今日び Java のオブジェクト指向の説明なんて星の数ほどあるような気がしますが、それでもなお「これで分かった!」という説明に辿りつけ...

いまさら聞けないJavaによるオブジェクト指向の常識 (1/5) - IT
www.atmarkit.co.jp ~ Java Agile

2008/05/08 - 今回は、オブジェクト指向プログラミングの常識について、Java言語を通じて学んでいきます。この連載「プログラマーの常識をJavaで身につける」では、Java言語やその文法は一通り理解しているが、「プログラマー」としては初心者、という方を...

Java - オブジェクト指向はクソか? - Qiita

qiita.com ~ 投稿 | Java

2014/05/13 - 私は長らくJavaを用いて開発を行ってきましたが、最近、関数型言語であるClojureにより魅力を感じています。... として、オブジェクト指向がとても有名になってしまっただけで、OOPを批判することは教会で罵ること(“swearing in church”)の...

Google 検索



食べログ フレンド 東京フレンド 両王・小田急池袋フレンド 駒場東大前・池ノ上フレンド 駒場東大前駅(フレンド) ルヴェ・ゾン・ヴェール 駒場

約 398,000 件 (0.48 秒)

他のキーワード: Javaオブジェクト指向練習

Javaのオブジェクト指向入門 - KAB-studio
www.kab-studio.biz/Programming/OOPinJava/

「Javaのオブジェクト指向入門」は、Javaを使用したオブジェクト指向プログラミングの解説を掲載しているコンテンツです。クラスがよくわからない、オブジェクト指向って難しい、Javaとオブジェクト指向の関係がわからない、Javaに挫折しちゃった、という方。

「オブジェクト指向」はなぜ難しい? - 1.1 とりあえずクラスを作ってみよう - 1.2 クラスと参照

Javaのオブジェクト指向を 90 分で理解する - 偏見プログラ...

pg-kura.hatenablog.com/entry/20111222/1324557014

2011/12/22 - 1. 分からない。いくら説明を読んでも分からない。o 1.1. 未だに分からないJavaのオブジェクト指向 今日び Java のオブジェクト指向の説明なんて星の数ほどあるような気がしますが、それでもなお「これで分かった!」という説明に辿りつけ...

いまさら聞けないJavaによるオブジェクト指向の常識 (1/5) - IT
www.atmarkit.co.jp ~ Java Agile

2008/05/08 - 今回は、オブジェクト指向プログラミングの常識について、Java言語を通じて学んでいきます。この連載「プログラマーの常識をJavaで身につける」では、Java言語やその文法は一通り理解しているが、「プログラマー」としては初心者、という方を...

Java - オブジェクト指向はクソか? - Qiita

qiita.com ~ 投稿 | Java

2014/05/13 - 私は長らくJavaを用いて開発を行ってきましたが、最近、関数型言語であるClojureにより魅力を感じています。... として、オブジェクト指向がとても有名になってしまっただけで、OOPを批判することは教会で罵ること(“swearing in church”)の...

Google 検索

食べログの「星」



食べログ フレンド 東京フレンド 両王・小田急池袋フレンド 駒場東大前・池ノ上フレンド 駒場東大前駅(フレンド) ルヴェ・ゾン・ヴェール 駒場

約 398,000 件 (0.48 秒)

他のキーワード: Javaオブジェクト指向練習

Javaのオブジェクト指向入門 - KAB-studio
www.kab-studio.biz/Programming/OOPinJava/

「Javaのオブジェクト指向入門」は、Javaを使用したオブジェクト指向プログラミングの解説を掲載しているコンテンツです。クラスがよくわからない、オブジェクト指向って難しい、Javaとオブジェクト指向の関係がわからない、Javaに挫折しちゃった、という方。

「オブジェクト指向」はなぜ難しい? - 1.1 とりあえずクラスを作ってみよう - 1.2 クラスと参照

Javaのオブジェクト指向を 90 分で理解する - 偏見プログラ...

pg-kura.hatenablog.com/entry/20111222/1324557014

2011/12/22 - 1. 分からない。いくら説明を読んでも分からない。o 1.1. 未だに分からないJavaのオブジェクト指向 今日び Java のオブジェクト指向の説明なんて星の数ほどあるような気がしますが、それでもなお「これで分かった!」という説明に辿りつけ...

いまさら聞けないJavaによるオブジェクト指向の常識 (1/5) - IT
www.atmarkit.co.jp ~ Java Agile

2008/05/08 - 今回は、オブジェクト指向プログラミングの常識について、Java言語を通じて学んでいきます。この連載「プログラマーの常識をJavaで身につける」では、Java言語やその文法は一通り理解しているが、「プログラマー」としては初心者、という方を...

Java - オブジェクト指向はクソか? - Qiita

qiita.com ~ 投稿 | Java

2014/05/13 - 私は長らくJavaを用いて開発を行ってきましたが、最近、関数型言語であるClojureにより魅力を感じています。... として、オブジェクト指向がとても有名になってしまっただけで、OOPを批判することは教会で罵ること(“swearing in church”)の...

Google 検索

食べログの「星」

情報推薦システムの分類

- 知識ベースフィルタリング

データに手を加えるなどの操作はせず、ユーザーが指定したカテゴリや価格などによってフィルタリングし、推薦する。

Ex. 『価格.com』の絞り込み、『食べログ』などのラーメンランキング

- コンテンツベースフィルタリング

商品の内容からユーザーの嗜好とのマッチングを推定し、推薦する。

Ex. 『Google』の検索

- 協調フィルタリング

履歴情報から嗜好を推定し、類似した商品を推薦する。(多く推薦システムで用いられている)

Ex. 『amazon』の「この商品を買った人はこんな商品も買っています」、『Youtube』のあなたへのおすすめ

協調フィルタリングとは

<協調フィルタリングにおける仮定>

「嗜好が似ているユーザーは行動も似ている」

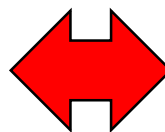


「行動が似ているユーザーは嗜好が似ている」

ユーザー自身が自分の趣味・嗜好を最も意識しない推薦手法

メリット

- 全くの未知の商品に対しても推薦してくれる
- よく使えば使うほど精度が上がる



デメリット

- 見当はずれな推薦がされることもある
- 一度買った商品は二度と推薦されない
- 最初は精度があまり高くない
(コールドスタート問題)

協調フィルタリングの分類

- アイテムベース協調フィルタリング

商品同士の類似度を取り、類似度の高い商品を推薦する協調フィルタリング
→「この商品を買った人はこんな商品も買っています」推薦

- ユーザーベース協調フィルタリング

ユーザー同士の類似度を取り、類似度の高いユーザーの履歴のうち、最もコンバージョンの可能性が高いと考えられる商品を推薦する協調フィルタリング。
→「あなたへのおすすめ」推薦

評価Matrix


- 協調フィルタリングはユーザーごとのアイテムに対する評価（利用の有無）を数値化したベクトルを並べたMatrixを用いる。

Ex. 各ユーザーのラーメン店に対するレビューの5段階の評価値

	駒鉄	山手ラーメン	千里眼	砦	侍
ユーザーA	3	2	1	4	0
ユーザーB	4	2	1	0	5
ユーザーC	2	2	2	4	0
ユーザーD	0	3	4	3	0
ユーザーE	0	0	0	0	4

評価Matrixにおける協調フィルタリングの分類

- アイテムベース協調フィルタリングはアイテム方向のベクトルに注目する。

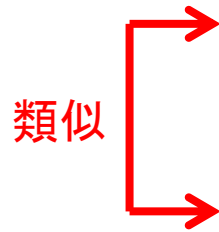


A red bracket with the word "類似" (Similarity) in red text above it, connecting the "山手ラーメン" and "千里眼" columns of the table below.

	駒鉄	山手ラーメン	千里眼	砦	侍
ユーザーA	3	2	1	4	0
ユーザーB	4	2	1	0	5
ユーザーC	2	2	2	4	0
ユーザーD	0	3	4	3	0
ユーザーE	0	0	0	0	4

評価Matrixにおける協調フィルタリングの分類

- ユーザーベース協調フィルタリングはユーザー方向のベクトルに注目する。



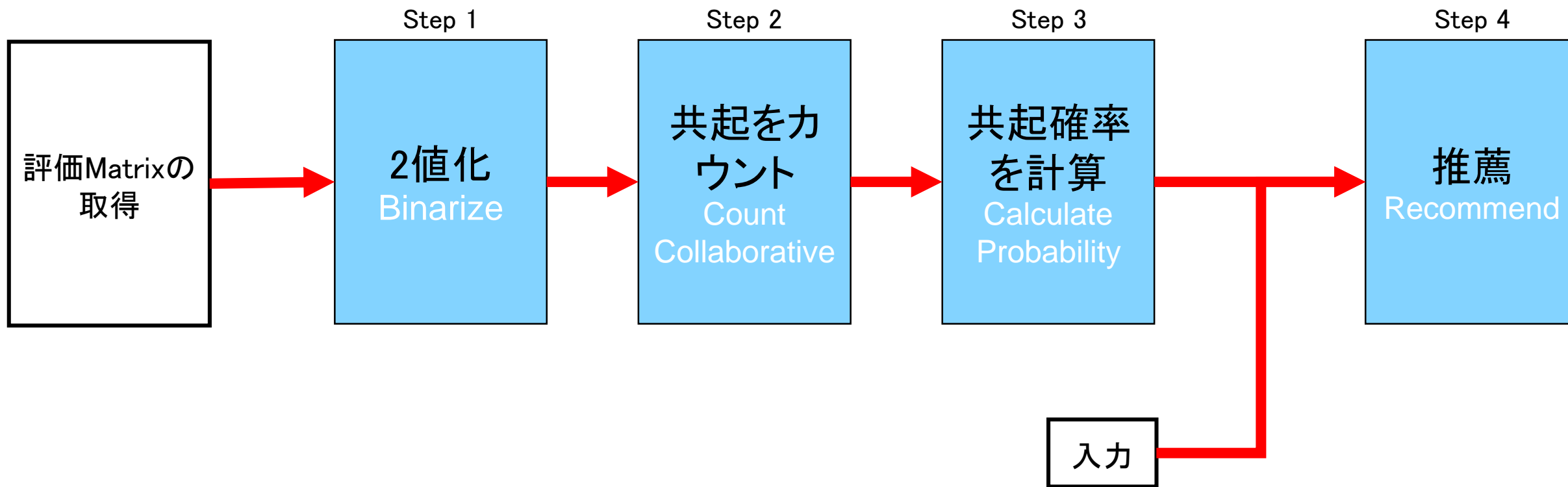
	駒鉄	山手ラーメン	千里眼	砦	侍
ユーザーA	3	2	1	4	0
ユーザーB	4	2	1	0	5
ユーザーC	2	2	2	4	0
ユーザーD	0	3	4	3	0
ユーザーE	0	0	0	0	4

今回のアルゴリズム

- 先ほどのラーメン屋の例のMatrixを用いて説明する。
- 協調フィルタリングのアルゴリズムはいくつかあるが、今回はごく単純なアイテムベース協調フィルタリングのアルゴリズムを取り上げる。

	駒鉄	山手ラーメン	千里眼	砦	侍
ユーザーA	3	2	1	4	0
ユーザーB	4	2	1	0	5
ユーザーC	2	2	2	4	0
ユーザーD	0	3	4	3	0
ユーザーE	0	0	0	0	4

今回のアルゴリズム

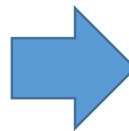


Step 1 : 2値化

- 関係性を簡略化する(特徴抽出)
今回は評価値が2以上を1とし、それ未満を0とする。

```
function binarize(matrix)
```

	駒鉄	山手ラーメン	千里眼	砦	侍
ユーザーA	3	2	1	4	0
ユーザーB	4	2	1	0	5
ユーザーC	2	2	2	4	0
ユーザーD	0	3	4	3	0
ユーザーE	0	0	0	0	4



	駒鉄	山手ラーメン	千里眼	砦	侍
ユーザーA	1	1	0	1	0
ユーザーB	1	1	0	0	1
ユーザーC	1	1	1	1	0
ユーザーD	0	1	1	1	0
ユーザーE	0	0	0	0	1

Step 2 : 共起をカウント

- 共起の数を数える

function countCollaborative(matrix)

	駒鉄	山手ラーメン	千里眼	砦	侍
ユーザーA	1	1	0	1	0
ユーザーB	1	1	0	0	1
ユーザーC	1	1	1	1	0
ユーザーD	0	1	1	1	0
ユーザーE	0	0	0	0	1



	駒鉄	山手ラーメン	千里眼	砦	侍
駒鉄	—	3	1	2	1
山手ラーメン	3	—	2	3	1
千里眼	1	2	—	2	0
砦	2	3	2	—	0
侍	1	1	0	0	—
計	7	9	5	7	2

Step 3 : 共起確率を計算

- 共起の確率を計算

```
function calcProbability(matrix)
```

	駒鉄	山手ラーメン	千里眼	砦	侍
駒鉄	—	3	1	2	1
山手ラーメン	3	—	2	3	1
千里眼	1	2	—	2	0
砦	2	3	2	—	0
侍	1	1	0	0	—
計	7	9	5	7	2



	駒鉄	山手ラーメン	千里眼	砦	侍
駒鉄	—	1/3	1/5	2/7	1/2
山手ラーメン	3/7	—	2/5	3/7	1/2
千里眼	1/7	2/9	—	2/7	0
砦	2/7	1/3	2/5	—	0
侍	1/7	1/9	0	0	—
計	1	1	1	1	1

Step 4 : 推薦

- 縦方向に見たときの確率の順に推薦される。

	駒鉄	山手ラーメン	千里眼	砦	侍
駒鉄	—	1/3	1/5	2/7	1/2
山手ラーメン	3/7	—	2/5	3/7	1/2
千里眼	1/7	2/9	—	2/7	0
砦	2/7	1/3	2/5	—	0
侍	1/7	1/9	0	0	—
計	1	1	1	1	1

(「駒鉄」を好きと答えた人に対して)

こちらもおススメ！！
山手ラーメン 砦

(解釈)

駒鉄を好きと答えた人が好きと答える確率の高い順に推薦する。

```
function recommendByRank(probabilities, returnCount)
```

DEMO

- Sample 2のデータはMovieLens (<http://grouplens.org/datasets/movielens/>)が配布している映画に関するレビューの最新の縮小版データ(MovieLens Latest Datasets – Small)を用いる。

＞サンプルプログラムを動かす際の注意点

- ブラウザは**必ずSafariを使用**してください。
- また、Safariの「開発」メニューから「ローカルファイルの制限を無効にする」を選択した状態で、HTMLファイルを開いてください。

配布プログラムの内容

<基本的なサンプルファイル>

- sample1.html

<上級者向け ライブラリファイル>

- 協調フィルタリング実装: CollaborativeFiltering.js
- CSV→javascriptオブジェクト変換: CSVForCollaborativeFiltering.js
- CSV読み込み関連ライブラリ: CSVData.js

<特に気にしなくていいファイル>

- (モーダルウィンドウ表示関連: modal.js)
- (テーブル表示用jQueryプラグイン (<http://michaeleisenbraun.com/columns/>): jquery.columns.min.js)
- その他 CSSなど

DEMO

- sample1.html : ラーメンレビュー 単純リスト

<上級者向け>

- sample2.html : Movie Lens 単純リスト
- sample3.html : Movie Lens ライブラリ使用テーブル