

A Framework for Personalized Apparel Recommendation

Wei Li

Nov 27, 2012

Outline

Introduction

Our Approach

Evaluation

Conclusion and Future Work

Introduction

Introduction

Our Approach

Evaluation

Conclusion and Future Work

Motivation

- Picking an item is **time-consuming** for individuals
- **Personalized** recommendation tends to be very useful
- Optimize the **user experience**
- Shortening the time will improve the **throughput**

Related Work

- Collaborative Filtering
 - **NO** personalization
- Integrated diffusion-based algorithm
 - **NOT** online algorithm
 - Can **NOT** give real-time response
- INTRIGUE
 - Use naive polynomial formulas for scoring of items and thus has **less intelligence** than machine learning based approaches

Vision

To build an **apparel recommender system framework** which is ...

1. **Personalized**

- user rating
- purchase history, favorites, scanning time

2. **Accurate**

- image features
- text tags

3. **Efficient**

- online context: E-commerce, e.g. Taobao

Our Approach

Introduction

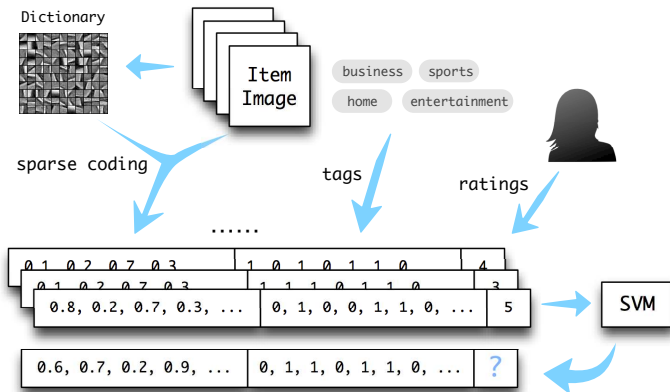
Our Approach

Evaluation

Conclusion and Future Work

Our Approach

- System framework Overview



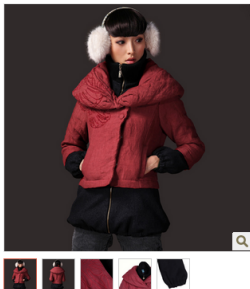
Our Approach

- Data collections
- Model **apparel items**
- Model **user preference**
- Recommend by rating

Data collections

- **Crawl** apparel items from Taobao: including both **images** and **textual labels**
- An online **website** for collecting user ratings

Website

[切换用户](#)


Ab, 您好!

感谢您能花时间为我们的项目做测试哦 O(∩_∩)O~

为使网站能使用的方便有效, 这里有一些小tips:

1. 请根据您这件衣服的喜好打分, 即您有多大意愿拥有这件衣服
2. 打分时请忽略价格和颜色因素, 以衣服款式为主
3. 鼠标移至左下方的小图上可以切换至衣服的不同角度
4. 鼠标移至左边的大图上可以看到放大的细节图
5. 以点亮几颗星星表示您给这件衣服的分, 5分表示非常喜欢, 1分表示很差



您已经评价过 0 件衣服了

品牌: 裂帛

领子: 翻领

衣门襟: 拉链

颜色分类: 红色

货号: 15180014

厚薄: 加厚

图案: 纯色

尺码: S (160/80A) M (165/84A) L (170/...

板型: 修身型

袖长: 长袖

风格: 原创设计

价格: 201-500元

衣长: 中长款 (65cm<衣长<=80cm)

袖型: 蝙蝠袖

适合人群: 淑女

年份: 2011

Website

[切换用户](#)


品牌:RTW

风格:欧美

季节:春季秋季

款式细节:绣花

货号:RTWT11751

适合人群:淑女

价格:71-120元

工艺:拼接

款式:长袖

颜色分类:草绿色藏青色粉色

材质:棉+氨纶

衣长:中长款(衣长50-70CM)

领子:翻领

尺码:S/160M/165L/170

袖型:常规

年份:2012

Model an item according to text

- **Tag vector** from text: binary representation
- **HAC Clustering**: detect the different text labels with the same meaning

Model an item according to image

- Do **NOT** care about colors
- Preprocessed to **gray-scale map** by morphological manipulation using **OpenCV** to eliminate **disturbance** form the background and people face
- **K-SVD** : Learn dictionary of apparels according to classification
- Sparse representation of images to get the **feature vector**

Model an item according to image

- Image preprocessing with OpenCV



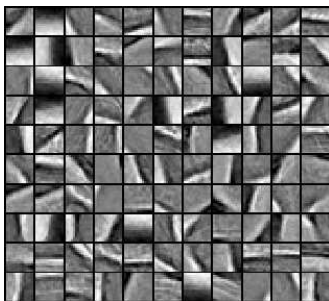
Model an item according to image

- Image preprocessing with OpenCV



Model an item according to image

- Dictionary learned from K-SVD



Model an item

- Combine the vector of **textual labels** and vector of **image features** to represent an **apparel item**

Model user preference

Get **user preference** and model it

- Judge garments with **scores(1-5)** based on **personal preference**
- Model **user preference** by their actions: purchase history, favorites collections, scanning time, etc.

Model user preference

Map apparel item to user preference

- According to **specific** user
- Mix **apparel vectors** and **textual tag vectors** along with the corresponding **user ratings** as training set

Recommend by rating

- Support vector machine (Libsvm)
- Ranking by **predictive rating** for the specific user

Recommendation

[切换用户](#)


1. 请根据您这件衣服的喜好打分，即您有多大意愿拥有这件衣服
2. 打分时请忽略价格和颜色因素，以衣服款式为主
3. 鼠标移至左下方的小图上可以切换至衣服的不同角度
4. 鼠标移至左边的大图上可以看到放大的细节图
5. 以点亮几颗星星表示您给这件衣服的分，5分表示非常喜欢，1分表示很差



您已经评价过 42 件衣服了

您可能会喜欢：



品牌:RTW

风格:欧美

季节:春季秋季

款式细节:绣花

货号:RTWT11751

适合人群:淑女

价格:71-120元

工艺:拼接

款式:长袖

颜色分类:草绿色藏青色粉色

材质:棉+氨纶

衣长:中长款(衣长50-70CM)

领子:翻领

尺码:S/160M/165L/170

袖型:常规

年份:2012

Problems and solutions

- P: People's **emphasis** on image and text varied. How to **weight** the two vectors?
- S: SVM will automatically learn from the training set to determine the feature influence.

Problems and solutions

- P: Each apparel have **different quantity** of images which will lead to a **varied** length of apparel vectors.
- S: Treat the images affiliated to the same apparel as **different apparels** with the **same** textual tags and user rating.

Evaluation

Introduction

Our Approach

Evaluation

Conclusion and Future Work

Experiment Settings

- 1772 garments : several images and text labels
- 24 volunteers
- 60 ratings per volunteer 1553 ratings in total

Evaluation

Table 1: 5-Fold Cross-Validation Accuracies

ID	Accuracy	ID	Accuracy	ID	Accuracy
1	89.1%	9	97.1%	17	93.3%
2	92.5%	10	78.3%	18	94.6%
3	80.9%	11	82.1%	19	85.7%
4	88.5%	12	85.0%	20	89.4%
5	91.3%	13	94.1%	21	88.4%
6	90.0%	14	89.4%	22	89.7%
7	86.9%	15	93.8%	23	90.5%
8	89.5%	16	96.7%	24	95.2%

Conclusion and Future Work

Introduction

Our Approach

Evaluation

Conclusion and Future Work

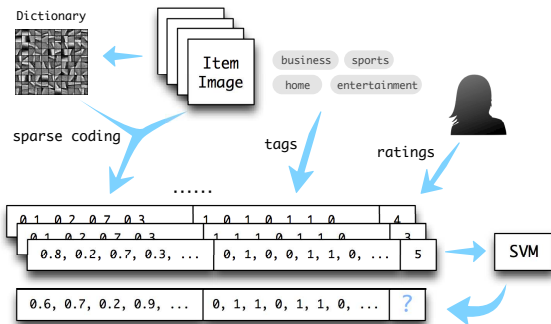
Conclusions

- **No** similar work has done before
- **Personalization** meet the user's need
- Integrate **feature representation** and **machine learning**
- Leverage both the **images** and **textual** labels to enhance accuracy
- Combined with **e-commerce**, this framework can have a very good market prospect

Future Work

- Provide more **recommending filters** under our framework
- Experiment on **various aspects** of our framework and provide extensive test result for reference

Any Questions :)



Thank you!