DeepStyle

Hyunsoo Na

2021. 6. 9

Index

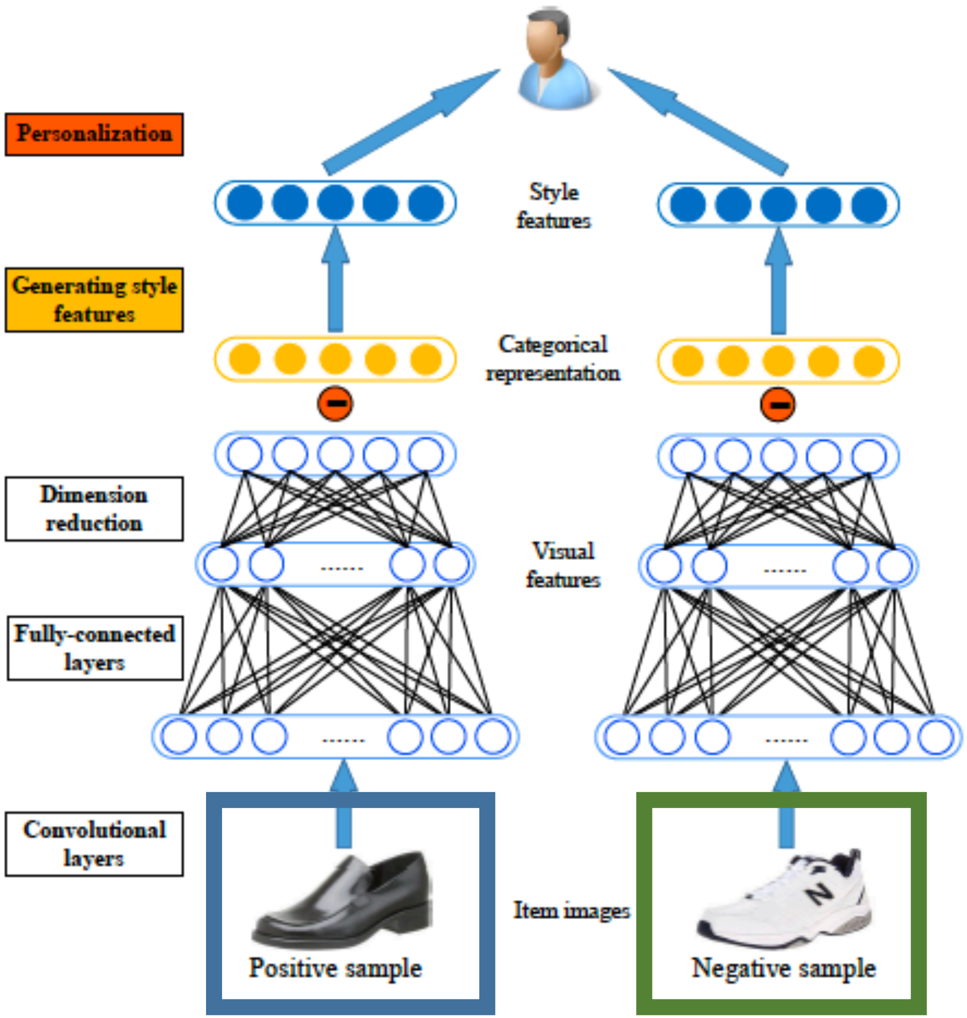
- Overview
- Introduction
- DeepStyle
- Experiment
- Conclusion

Overview

- **DeepStyle**
  - Q Liu, et al. DeepStyle: Learning User Preferences for Visual Recommendation, SIGIR, 2017

- **Overview**
  - Visual Recommendation
  - Style Features

Overview



Positive Sample - Negative Sample = Style

Introduction

- **Visual Information**
  - Visual Dimension => User Preferences
  - Better Personalized Recommendation

- **Previous Studies**
  - Personalized Matching of items on visual features
  - Visual BPR (VBPR)
  - Failed to capture different styles of items

Introduction

- **Clothing Cluster**
  - CNN visual features
  - One Row is one Cluster
  - Cluster based on clothing category
  - Different styles not distinguished

- **Category vs Style**
  - Clothing Category is very powerful
    - Similarity based
    - Leather shoes vs Sneakers
    - Leather shoes vs Suit Pants
  - Different styles
    - Casual, formal, male, female…
  - Similar style items are bought together
  - Sparse Hierarchical Embedding
    - Varying Style Features

DeepStyle

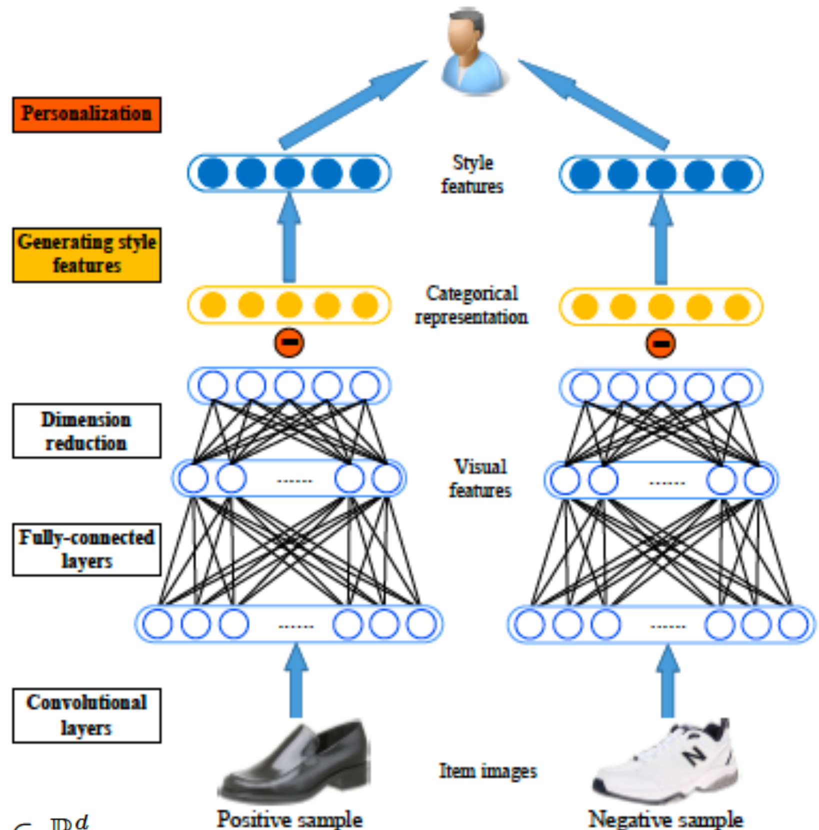
- **Visual Feature Space**
  - Items with similar styles may not be similar in visual space
  - How can we capture style over dominant category features?

- **DeepStyle**
  - Item = Style + Category
  - Subtract category representation from visual features
  - BPR => Personalized Recommendation

- **Notation**
  - Set of items selected by user u $\mathcal{I}^u$
  - Item's specific visual category l_i

DeepStyle

- **Architecture**
  - 5 Convolutional layers + 3 FC Layers
  - Pre-trained on ImageNet
  - 4096 visual feature vector
- **Obtaining Style Features**
  - Visual Feature $\mathbf{s}_i = \mathbf{E}\mathbf{v}_i - \mathbf{l}_i$
  - Style Feature $\mathbf{s}_i \in \mathbb{R}^d$
  - Weight Matrix $\mathbf{E} \in \mathbb{R}^{d \times 4096}$
- **Bayesian Personalized Ranking**
  - Incorporate style features in BPR (implicit)
  - Prediction $\hat{y}_{u,i} = (\mathbf{p}_u)^T (\mathbf{s}_i + \mathbf{q}_i)$
  - Latent representation of user/item $\mathbf{p}_u \in \mathbb{R}^d$ $\mathbf{q}_i \in \mathbb{R}^d$



DeepStyle

- **Objection**
  - Model needs to satisfy $\hat{y}_{u,i} > \hat{y}_{u,i'}$
  - Positive/Negative Sample $p\left(u, i > i'\right) = g\left(\hat{y}_{u,i} - \hat{y}_{u,i'}\right)$
  - g : Usually Sigmoid

  - Negative log likelihood : Minimize $J = \sum_{u,i} \ln\left(1 + e^{-\left(\hat{y}_{u,i} - \hat{y}_{u,i'}\right)}\right) + \lambda \|\theta\|^2$
  - Normalization λ , Parameters θ

Experiments

- **Datasets**
  - Clothing subset and Home/Kitchen subsets of Amazon Dataset
  - Clothing 74 categories
  - Home/Kitchen 86 categories
  - 80% training, 20% testing
  - Remove users with less than 5 records and more than 100 records

  - Warm Start : overall ranking
  - Cold Start : Items with less than 5 records during training

  - AUC Metric

$$AUC = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{1}{|set\left(i \in \mathcal{I}^u, i' \notin \mathcal{I}^u\right)|} \sum_{i \in \mathcal{I}^u, i' \notin \mathcal{I}^u} \delta\left(p_{u,i} > p_{u,i'}\right)$$

  - Dirac delta function : 1 when condition met, otherwise 0 δ

Experiments

- **Benchmark**
  - BPR
  - Visual BPR (VBPR)

  - Sparse Hierarchical Embedding (Sherlock)
    - Extends VBPR to categorical information
    - Categorical Tree
    - Embedding matrices for transferring visual features to style features vary according category
    - One embedding matrix for each category leads to large amount of parameter

Experiments

- **Category + Style**
  - Visual Feature is important (especially for cold-start)
  - BPR vs VBPR

  - Categorical Info is important in style
  - VBPR vs DeepStyle

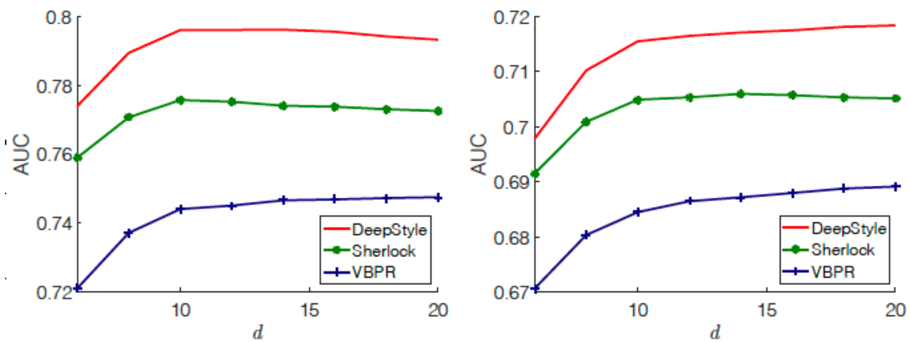
  - Learning style feature is important
  - Sherlock vs DeepStyle

Table 1: Performance comparison on predicting users preferences on items measured by AUC. The dimensionality is $d = 10$ on both datasets.

dataset	setting	BPR	VBPR	Sherlock	DeepStyle
Clothing	warm-start	0.6243	0.7441	0.7758	**0.7961**
	cold-start	0.5037	0.6915	0.7167	**0.7317**
Home	warm-start	0.5848	0.6845	0.7049	**0.7155**
	cold-start	0.5053	0.6140	0.6322	**0.6396**

Experiments

- **Dimensionality**
  - Stable after d=10 : dimension flexibility
  - Sherlock overfit after 10
    - Assumption : One embedding matrix for each category estimate too many parameters



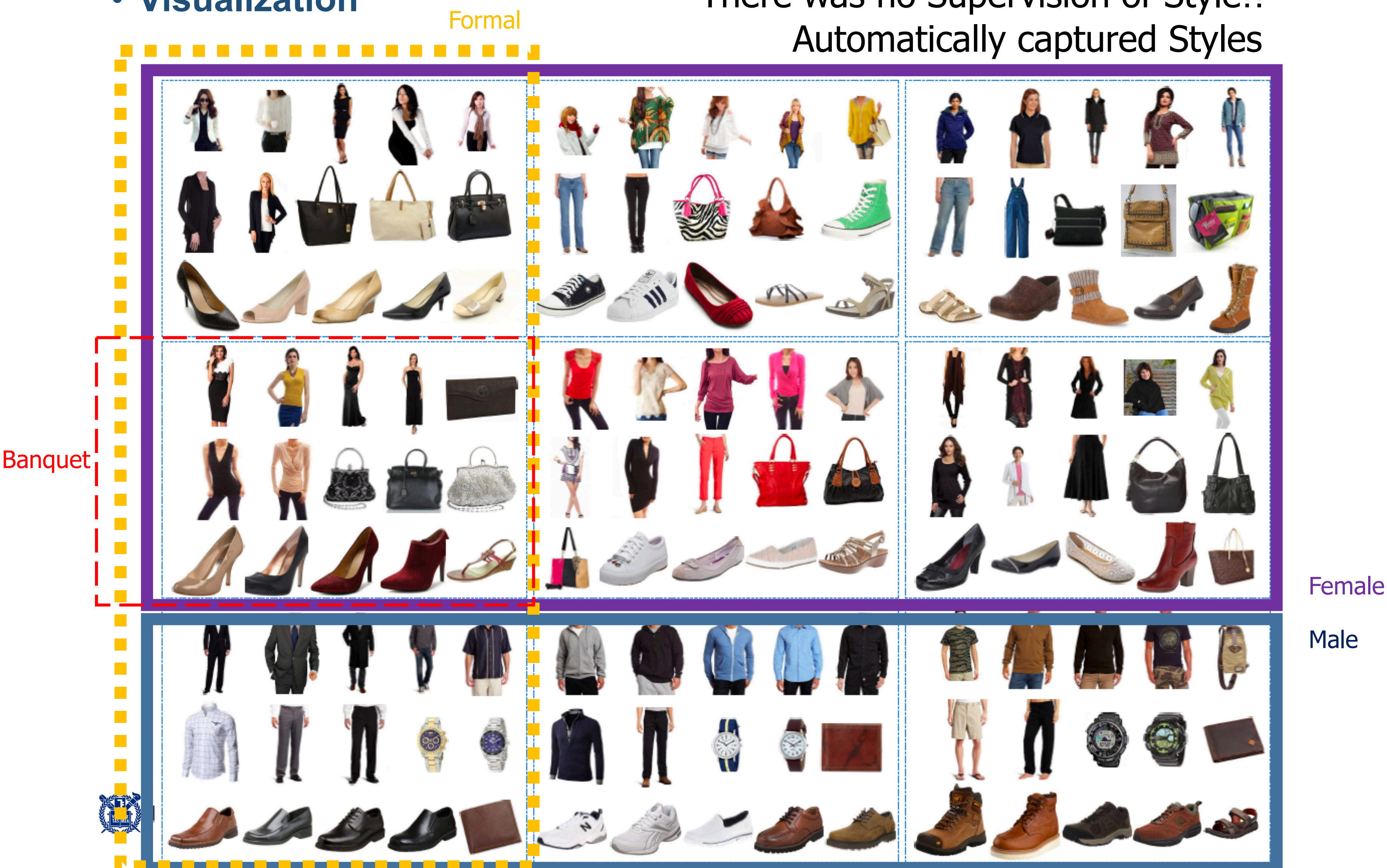
(a) Clothing. (b) Home.

Figure 3: Performance of DeepStyle, Sherlock and VBPR with varying dimensionality under the warm-start setting measured by AUC.

Experiments

- **Visualization**

There was no Supervision of Style!!
Automatically captured Styles



Conclusion

- **DeepStyle**
  - Visual Recommendation

  - Learning styles of items and preferences of users
  - Subtracts categorical information from visual features

  - Added style features to VBPR

End