BJIT Ltd

Weekly Report

Fashion Recommender System

Prepared by: Muttakin Islam, Md Shohag Mia
Reviewed by: Javed Hasan

Table of Contents

Project Goal

The goal of this project is to develop a **Fashion Recommender System** that recommends fashion items that are similar to given or selected one. This system will function calculating image similarity ranking.
A customer may upload an image of a fashion item or the image might be selected from an image database, then the recommender system will calculate the similarity ranking and predict some similar fashion items based on similarity ranking.
The image dataset will have metadata for the images. This metadata includes the information of gender, age category(men/women/girls/boys), item category, season, year etc of a certain fashion item. We will include these metadata of a fashion item image to filter age, gender, season etc based fashion that are complex(Not possible) to get from image alone. It will help us to avoid age, gender etc related major issues(mistakes) to make the system more accurate.
We will apply Machine learning based cosine similarity models to find similarity(on metadata) among fashion items first. It will help us to filter out items based on age, gender etc.
Then we will use a Deep learning CNN model to get similar fashion items with similar patterns. Here we will use a triplet of images to get image similarity and image ranking. There will be the given image, the positive(most similar) image and the negative(most dissimilar) image, from these three images our model will find the most close(similar) images for recommendation.
We may also use NLP to get the current Fashion trends from Social media like twitter and apply them for more accurate, specific, trendy and personalized recommendations. We may deploy humans(Stylists) to do it.

Goal of This Week

1. Business understanding
2. Data understanding
3. Exploring data to get business goal

Accomplishments

- We have studied some books:
  - Data Science
  - CRISP DM-1.0

- Separated the tasks in the first two phases of **CRISP DM** and specify the child-tasks and deliverables of those child-tasks.
  - Task - Business and Data understanding

- We have explored some datasets. Those are:
  Kaggle Fashion Dataset.
  - Dataset description report

No. of features	No. of usable feature	No. of data
10	9	44424

Features	Description	Example	Data type	Missing count	Filling factor (%)
id	Unique identity for every product image and it's metadata	15970	number	0	100
gender	Gender of the person for whom the product is	Men / Women / Boys / Girls / Unisex	string	0	100
masterCategory	Master category of the product	Apparel / Accessories / Footwear	string	0	100
subCategory	Subcategory of the product	Topwear / Bottomwear / Shoes / Belts / Sarees / Watches	string	0	100
articleType		Shirts / Jeans / Track pants / Watch / Casual shoes / Hand	string	0	100

		bag			
baseColour	Main color of the product	Grey / Green / Purple / Navy blue	string	15	99.96623447
season	Season for which the product is suitable to wear	Fall / Summer / Winter	string	21	99.95272825
year	Year of manufacture	2011 / 2012	number	1	99.99774896
usage	Usage of the product	Casual / Formal / Sports	string	317	99.28642175
productDisplayName	Name of the product	Peter England Men Party Blue Jeans	text	7	99.98424275

- We have studied some research papers:
    - https://arxiv.org/ftp/arxiv/papers/1709/1709.08761.pdf

Plans for Next Week

1. Business understanding
2. Data understanding
3. Exploring more datasets
4. Repeat the process to meet business goal the most

Project Status

Green

Note -

- Green means project is on schedule and there are no major issues
- Yellow means the project is somewhat delayed and/or there may be some major risks at the current point
- Red means the project is seriously at risk of being delayed and/or there are some major risks affecting the project

Project Timeline(Business & Data understanding)

Tasks	Sub-tasks	Week-1	Week-2	Week-3	Week-4	Week-5
Business Understanding	Determining business objectives					
	Assessing situation					
	Determining project goals					
	Producing project plan					
Data Understanding	Collecting initial data					
	Describing data					
	Exploring data					
	Verifying data					

Query

None

Key Issues, Risks or Concerns

Item	**Action/Resolution**	**Responsible**	**Completion Date**
GPU requirement	Needs GPU to explore potential models	Javed Hasan	