

Fashion Recommender on Amazon

#Team 5:

Yichuan Zhang – zhang.yichu@husky.neu.edu

Houze Liu – liu.hou@husky.neu.edu

Goals of the Project

- Recommendation System for the customers based on Amazon/Ebay selling and ratings
- According to database built from Amazon/Ebay Api, we rank most popular elements that prevail in fashion trend.

Use Case

- User Input:
 - A Keyword in kind of certain product catalog
- Output For User:
 - Most popular item in the catalog indicated by keyword
 - A table consists of Most Popular Brand, Most Popular Color and Most Popular Price Range of this catalog etc.

Methodology Conclusion

- Parse Data(XML)
 - Manufacturer field
 - ProductGroup field
 - Title field and etc.
- Filter out irrelevant items
- Calculate popularity/trend in brand, color, price and etc.
- Visualize output data to user

```
<Item>
  <ASIN>B012E5ERGG</ASIN>
  <ParentASIN>B00R1HD3NQ</ParentASIN>
  <DetailPageURL>https://www.amazon.com/Lacoste-Womens-Sleeve-Stretch-Merlot/dp/B012E5ERGG%3Fpsc%3D1%26
  <ItemLinks>
    <ItemLink>
      <Description>Technical Details</Description>
      <URL>https://www.amazon.com/Lacoste-Womens-Sleeve-Stretch-Merlot/dp/tech-data/B012E5ERGG%3FSu
    </ItemLink>
    <ItemLink>
      <Description>Add To Baby Registry</Description>
      <URL>https://www.amazon.com/gp/registry/baby/add-item.html%3Fasin.0%3DB012E5ERGG%26Subscripti
    </ItemLink>
    <ItemLink>
      <Description>Add To Wedding Registry</Description>
      <URL>https://www.amazon.com/gp/registry/wedding/add-item.html%3Fasin.0%3DB012E5ERGG%26Subscri
    </ItemLink>
    ...
  </ItemLinks>
  <ItemAttributes>
    <Manufacturer>Lacoste Womens Apparel</Manufacturer>
    <ProductGroup>Apparel</ProductGroup>
    <Title>Lacoste Women's Half Sleeve Stretch Pique Slim Fit Polo Shirt, Merlot Purple, 32</Title>
  </ItemAttributes>
```

Data Sources & Data Magnitude

- Analysis based on attributes of product like keyword, brand, price that can define a search request
- Finding product data by using ItemSearch API call
- Data size: Based on number of API calls. Each API will return 10 items
- Reference:
<https://docs.aws.amazon.com/AWSECommerceService/latest/DG/CommonItemSearchParameters.html>

Milestones Reports

- March 5, 2018 - March 12, 2018: compose ideals
- March 13, 2018 - March 30, 2018: prepare data and self-learning Amazon API
- March 20, 2018 - April 5, 2018: Spark self-learning
- March 20, 2018 - April 10, 2018: Implementation and testing
- April 1, 2018 - April 15, 2018: Zeppelin self-learning and visualization

Programming in Scala and Code

- Scala usage
 - Ingesting
 - Parsing
 - Calculating
 - Testing
- Repository address
 - <https://github.com/YichuanZhang/CSYE7200-SPRING2018-TEAM5.git>
- Documents
 - Google slides

Acceptance Criteria

- Our system results should match Amazon's recommendation system on 1 item out of every 10 items
- The most popular elements should frequently appear in Amazon Trend web page. ex: In 'clothes' catalog, 'black' clothes in the page should beats other color clothes in percentage, if 'black' tops popularity in our system.
- We should be able to collect enough data(10,000 items).

Q & A

Thank you