



2022 Campus Challenge

Extract Trends from social media data

Team Name: DTYDHTCode

Institute Name: Indian Institute of Technology (IIT), Bhilai

Team members details

Team Name	DTYDHTCode		
Institute Name	Indian Institute of Technology (IIT), Bhilai		
Team Members >			
	1 (Leader)	2	3
Name	Anupam Kumar	Ritik Dhanore	
Batch	2023	2023	

Deliverables/Expectations for Level 2 (Idea + Code Submission)

Deliverable 1:

Identification of trends from social media

1. Identify trends on social media based on category. Can restrict to Fashion as a category for the project. Ex: Polka dots dresses are trending on twitter.
2. Ranking/scoring logic for trends extracted.
3. Outcome format:
 - a. Option1: List of trending keyword(s) along with list of sample images and respective links from which the trend is derived with most trending first:
Example: Trends:[{Polka dot dresses, <list of links/images>,trending score}, {Bellbottom Jeans, <list of links/images>,trending score}..]
 - b. Option 2: structured data according to flipkart category, sub category, vertical and product attributes
Example: {category: Fashion, Sub-category: Women Western, vertical: Women dresses, trending attribute type: Pattern, trending attribute value: Polka Print, list of sample images and links from which the trend is derived}.
Outcome with Option 2 format will be given bonus points.

Deliverable 2:

Mapping trends with Flipkart products:

1. Create mapping of extracted trending keyword(s) with Flipkart category, sub category, vertical and product attribute(s), search page links.
Example:{category: Fashion, Sub-category: Women Western, vertical: Women dresses, trending attribute type: Pattern, trending attribute value: Polka Print}
Note: Use category, Subcategory combination from the Flipkart Website
2. From a trending keyword, creating a corresponding searchable term on Flipkart which will lead to matching products.
Example: Tropical Tops keywords will not give right results directly on Flipkart but we can construct search query for it using some intelligence.
3. Points will be given based on similarity between sample images for trends and product results on Flipkart.

Use-cases

- To identify the trends from social media, we have used the data specifically from Twitter.
- The trends have been identified by searching for different hashtags(#) in twitter by using relevant libraries.
- The user is able to explore the top list of products from all categories and check its relative score with respect to the other matched products.
- The name of the products which are trending are shown to the user in correct order.

Solution statement/ Proposed approach

The brief approach to solve the given problem statement is as follows:

- Scrape data from social media:
 - We have used Twitter to scrape the social media data for different items of different subcategories.
 - Initially, we created categories for - fashion, electronics, furnitures, sports, beauty, books and kids products.
 - For each of these category, we have set the most used and frequent hashtags for each of them, like #design for some clothing item representing fashion product.
 - Using the library tweepy, we have scrapped each of these hashtags that we created and explored the media items(image, gif, video, etc) fetched from the results.
 - These data are collected and stored in json formats because the data received is not uniform.

Solution statement/ Proposed approach

The brief approach to solve the given problem statement is as follows:

- Use the data to explore the trending products:
 - Now that all the data is stored, we search for the similarity between the results and number of times a similar product comes up in our result.
 - Most importantly, the trends are observed by going through the trending section of twitter where the latest trends are already mentioned.
 - We go through that section by running our piece of code and check for the hashtags from our keywords of hashtags from all the trending items/products.
 - This is done for a fixed frequency to accumulate the trending items from time to time.
 - After this process is complete, we update the json list of trending products according to the list obtained and also update the result accordingly on the website.

Limitations

- The results obtained are upto a specific time only, since the trends keep on changing so, to get the best results, we need to run the python script for scrapping the latest trends just before using tat data.
- Currently, the data shown on the hosted website is with respect to the trends from one day back.
- We have used only one social media platform i.e., twitter for obtaining the latest trends, the data from other social media would improve the results.
- We tried to scrape out the product name, and the other information available with it like an image or a video or some other media, but for some of the products none of these media sources were available as a result of which we are unable to display the image for the item or similar media source for the trending product on our website.

Future Scope

- We have not yet automated a method to use the latest trends obtained from social media and map the results obtained to Flipkart products. This part is still in progress and we would like to complete this part as soon as possible.
- After the problem statement is completely solved, the flipkart users can be provided with a option to explore the latest trends and be able to buy similar products on Flipkart by getting the mapped product.
- The results obtained can even be used by Flipkart to decide the products which can be discounted in festive seasons since these products are expected to be bought the most.