INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Research Paper on Implementation of Mining Facets Automation for the Searched Queries

Shraddha Londhe[1], Dr. Hemant Deshmukh[2], Prof. Pallavi Dhande[3]

[1]*M.E 2ND Year Computer Science and Engineering, [2]Head of the Department, [3]Assistent Professor, Dr. Rajendra Gode Institute of Technology & Research*

Abstract: Data mining is the process of sorting through large data sets to identify patterns and establish relationships to solve problems through data analysis. Data mining tools allow enterprises to predict future trends. The process of finding query facets which are in the form of multiple groups of words or phrases will be address as a problem to explain and summarize the content covered by a query. Facet is a set of items which describe and summarize one important aspect of a query, here a facet item is typically a word or phrase. A query may have multiple facets that summarize the information about the query from different perspectives. It is assumed that the important aspects of a query are usually presented and repeated in the query's top retrieved documents in the style of lists and query facets can be mined out by aggregating these significant lists. Query facets provide interesting and useful knowledge about a query and thus can be used to improve search experiences in many ways. To assist information for finding faceted queries a technique is explore that represent interesting facets of a query using groups of semantically related terms extracted from search results. Web search queries are often multi-faceted, which makes a simple ranked list of results inadequate. So, a method is used, refer to as QDMiner to automatically mine query facets by extracting and grouping frequent lists from free text, HTML tags, and repeat regions within top search results. Search results based on used method will simply improve the efficiency of users ability to find information easily.

I. INTRODUCTION

A query facet is a set of items which describe and summarize one important aspect of a query. Here a facet item is typically a word or a phrase. A query may have multiple facets that summarize the information about the query from different perspectives.

Query facets provide interesting and useful knowledge about a query and thus can be used to improve search experiences in many ways. users can understand some important aspects of a query without browsing tens of pages. For example, a user could learn different brands and categories of watches. We can also implement a faceted search [1], [2], [3], [4]based on the mined query facets. User can clarify their specific intent by selecting facet items. Then search results could be restricted to the documents that are relevant to the items. Important pieces of information about a query are usually presented in list styles and repeated many times among top retrieved documents. Thus we propose aggregating frequent lists within the top search results to mine query facets and implement a system called QDminer. More specifically QDMiner extracts lists from free text, HTML tags and repeat regions contained in the top search results groups them into clusters based on the items they contain then rank the clusters and items based on how the lists and items appear in the top results. Query facets are mined by the following four steps:

1) *List and Context Extraction* Lists and their context are extracted from each document.
2) *List Weighting* All extracted lists are weighted and thus some unimportant or noisy lists that occasionally occurs in a page can be assigned by low weights
3) *List Clustering* Similar lists are grouped together to compose a facet. We do not use individual weighted lists as query
4) *Facet and Item Ranking* Facets and their items are evaluated and ranked. Based on our motivation that a good facet should frequently appear in the top results, a facet is more important.
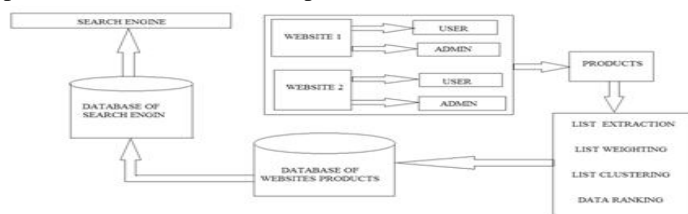

Figure 1 : Working Of Proposed System

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

In project, there are two ecommerce website that are ecommerce 1 and ecommerce 2. Both website have its own admin and various users. Admin have authority to upload product that uploaded

product are going through 4 steps that are list extraction, list weighting, list clustering and data ranking. These 4 steps are very important steps throughout the project because these 4 steps helps to convert unstructured data to structured data then this structured data submit to the database of websites and with the help of admin structured data in database of websites are finally submitted to database of search engine. And if user search any product then data comes in structured format in proper way from database of search engine

In this paper we also implement the impression and click technique.

Impression is a measurement of responses from a Web server to a page request from the user browser, which is filtered from robotic activity and error codes, and is recorded at a point as close as possible to opportunity to see the page by the user. An impression is when an ad is fetched from its source, and is countable. In this topic by using impression we show most visited items by user in our search engine. Click-through rate (CTR) is the ratio of users who click on a specific link to the number of total users who view a page, email.It is commonly used to measure the success of an particular website as well as the effectiveness of email campaigns.

II. IMPLEMENTATION

In project, there are two ecommerce website that are ecommerce 1 and ecommerce 2. Both website have its own admin and various users. Admin have authority to upload product that uploaded product are going through 4 steps that are list extraction, list weighting, list clustering and data ranking. These 4 steps are very important steps throughout the project because these 4 steps helps to convert unstructured data to structured data then this structured data submit to the database of websites and with the help of admin structured data in

database of websites are finally submitted to database of search engine. And if user search any product then data comes in structured format from database of search engine This shows various screenshots how the implementation of the system is done. There are six modules in the project.

A. Ecommerce 1(Website 1)
1) Ecommerce 1 Home Page
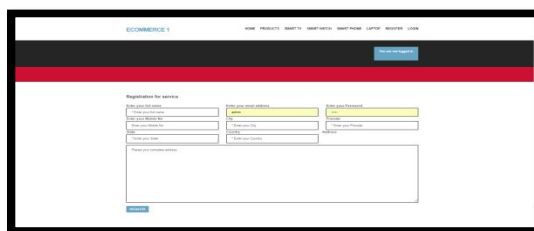
Screenshot 1: Ecommerce 1 Home Page:

Screenshot 6.1 is of home page of ecommerce 1 that is website 1.

2) Ecommerce 1 Registration Form

Screenshot 2: Ecommerce 1

3) Registration Form: In Registration Form Of Ecommerce 1, new user have to enter full name, email address, password, mobile number, city, pin code, state, country, address. In this form all fields are necessary.
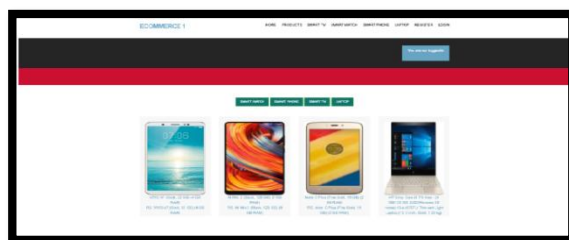
4) Ecommerce 1 Login Page
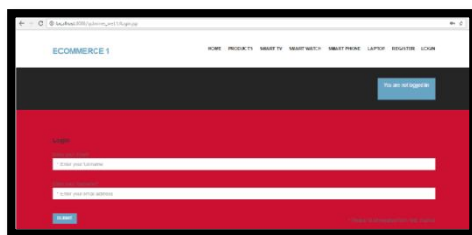
Screenshot 3:Ecommerce 1 Login Page

This is the login page of the ecommerce 1 website from where user can login with the help of email address and password
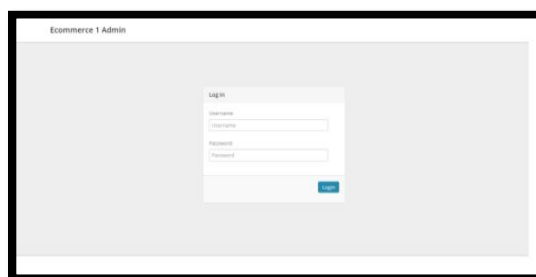
5) Ecommerce 1 Products

Screenshot 4 : Ecommerce 1 Products

In product page there is list of all product as well as there is also link of categories so that user can directly go to that page of particular category if user wants.
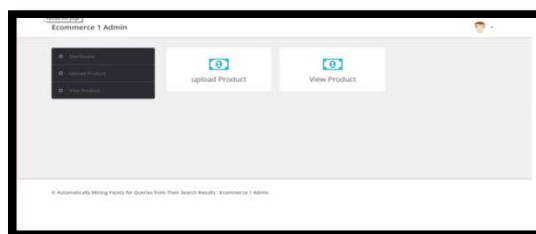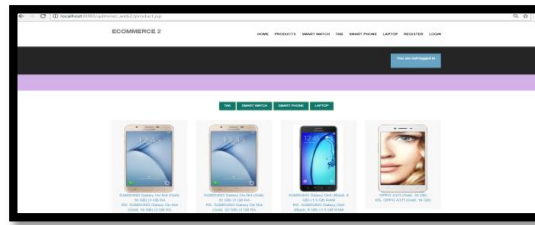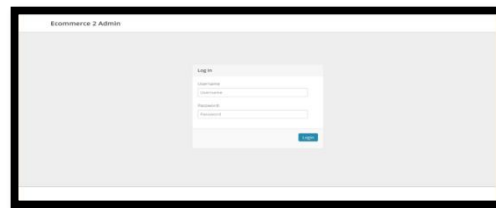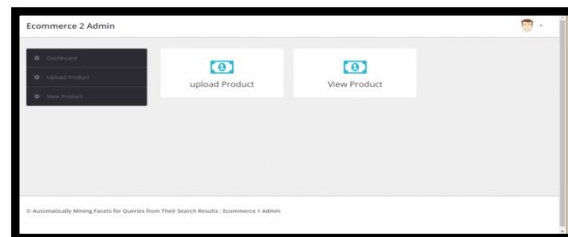
B. Admin of Ecommerce 1
1) Login Page

Screenshot 6: Login Page Of Admin 1

This is login page of admin of ecommerce 1 in which username and password of admin is require.
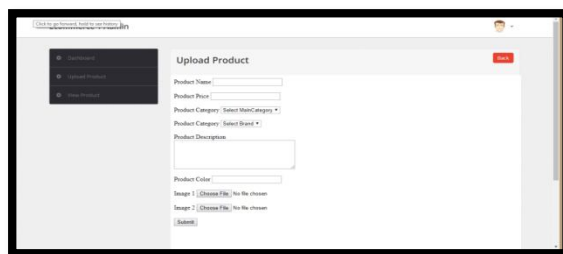
2) Home Page

Screenshot 7: Home Page Of Admin 1

This is home page of admin of ecommerce Admin have authority to handle all the work of ecommerce 1 such as upload product, view product.

![IJRASET logo]

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

3) *Upload Product*

Screenshot 8: Upload Product Admin1

In this upload form we have to fill product name, product price, product category, product brand, product description, product colour, product image and after submitting the product will show in ecommerce 1.

4) *Product Details*
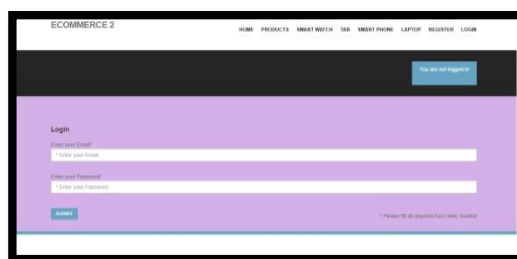
Screenshot 9: Product Details Admin 1

This is product details of admin of ecommerce 1. In this product details admin can view uploaded product.

C. *Ecommerce 2 (Website 2)*
1) *Login Page*
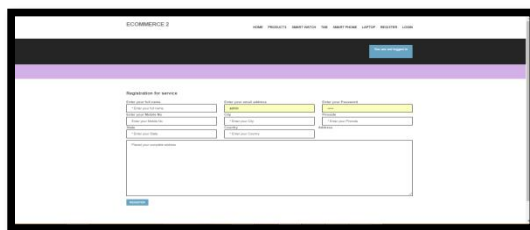
Screenshot 10: Ecommerce 2 Login

This is the login page of the ecommerce 2 website from where user can login with the help of email address and password

2) *Registration Form*

Screenshot 11: Registration Form

In Registration Form of Ecommerce 2, new user have to enter full name, emailaddress, password, mobile number, city,pin code, state, country, address. In this form all fields are necessary.

3) Product

Screenshot 12 : Product Of Ecommerce

In product page there is list of all product as well as there is also link of categories so that user can directly go to that page of particular category if user wants.

D.   Admin2
1)   Login Page

Screenshot 14: Login Page Of Admin Of Ecommerce 2

This is login page of admin of ecommerce 2 in which username and password of admin is require. Admin have authority to handle all the work of ecommerce 2

2)   Home Page

Screenshot 15: Home Page Of Admin 2

This is home page of admin of ecommerce2.Admin have authority to handle all the work of ecommerce website 2 such as upload product, view product.

3)   Upload Product
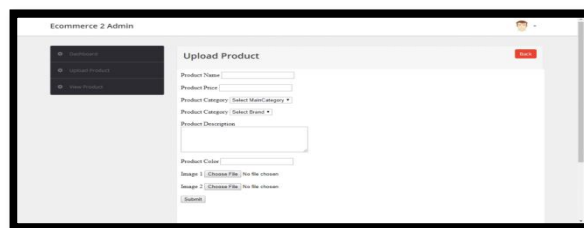
Screenshot 16: Upload Page Of Admin Of Ecommerce 2

This is upload page of admin of ecommerce 2. In this upload form we have to fill product name, product price, product category, product brand, product description , product colour, product image and after submitting the product will show in ecommerce 2.

4)   Product Details
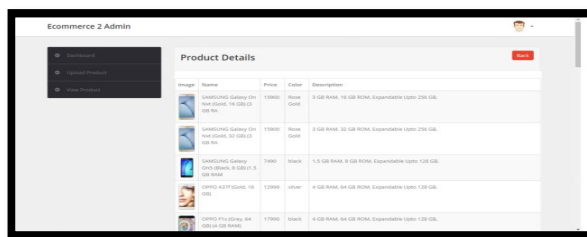
Screenshot 17: Product Details Of Admin Of Ecommerce 2

This is product details of admin of ecommerce 2. In this product details admin can view uploaded product.

5)   Payment Details
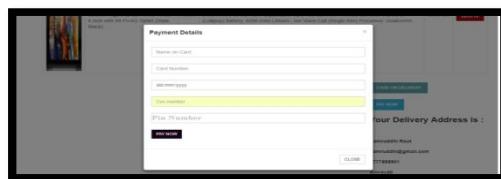
Screenshot 18: Payment Details

If user wants to do online payment then user have to enter name of card, card number, date, cvv number, and pin number. And after payment order place successfully.

E.   Search Engine
1)   Search Engine Home
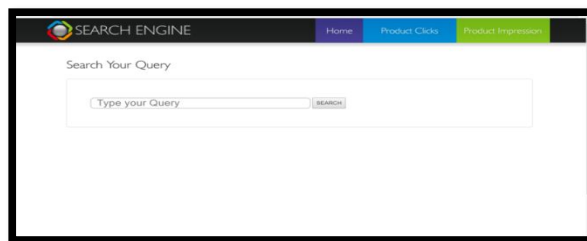
Screenshot 19: Search Engine Home

This is main part of project, search engine which gives the proper result to the user which store in the database of search engine. There are two ecommerce websitethat are ecommerce 1 and ecommerce 2. Both website have its own admin and various users. Admin upload product that uploaded product are going through 4 steps that are list extraction, list weighting, list clustering and data ranking.These 4 steps helps to convert unstructured data to structured data then this structured data submit to the database of websites and with the help of admin structured data in database of websites are finally submitted to database of search engine. And if user search any product then data comes in structured format in proper way from database of search engine.
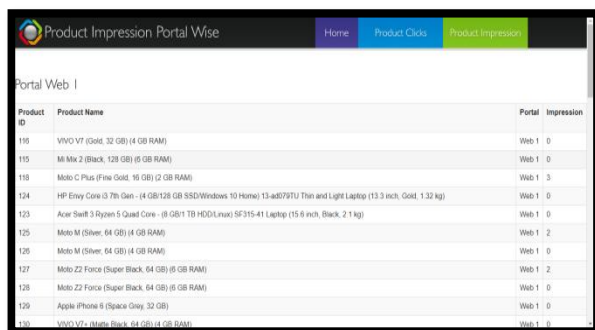
2)   Product Clicks Portal Wise
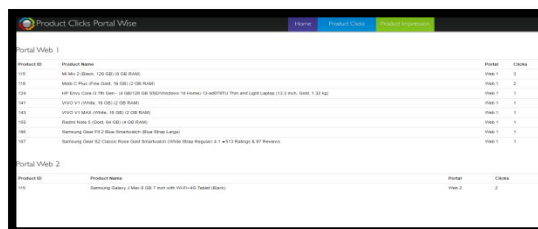
Screenshot 20:Product Clicks Portal Wise

The click-through rate is the number of times a click is made on the product divided by the total impressions. Product clicks are given separated for both the ecommerce website. Each item name is given with its product click portal wise

3) *Product Impression Portal Wise*

Screenshot 21: Product Impression Portal Wise

An impression is when an product is fetched from its source, and is countable. Whether the product is clicked is not takeninto account. Each time an product is fetched, it is counted as one impression

F. *Admin Section of Search Engine*
1) *Login page*
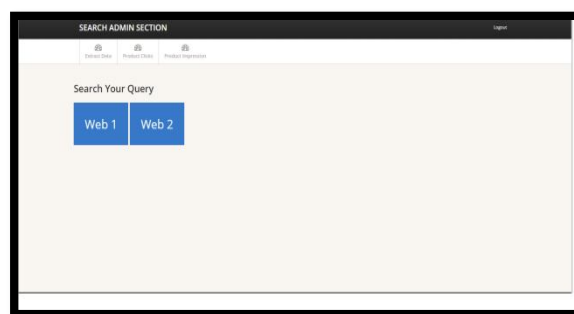
Screenshot 22: Login page

This is login page of admin of search engine in which username and password of admin is require. Admin have authority to submit the structured data to the database of the search engine.

2) *Home Page Of Admin Of Search Engine*

Screenshot 23: Home page

In home page of admin of search engine, there are tree links one is of extract data which is direct to home page which contain web1 and web2 two buttons, second is product clicks and third is product impression. There is also log out button to log out from the session.

3) Product Clicks Portal Wise In Admin Of Search Engine

Screenshot 24: Product Clicks Portal Wise In Admin of Search Engine

In product clicks portal wise that the click-through rate is the number of times a click is made on the product divided by the total impressions.

4) Product Impression Portal Wise In

a) Admin of Search Engine

Screenshot 25: Product Impression

b) Portal Wise In Admin of Search Engine: In product impression portal wise, an impression is when an product is fetched from its source, and is countable.

5) List Extraction

Screenshot 26: List Extraction

In list extraction, lists and their context are extracted from each document. "men's watches, women's watches, luxury watches, ..." is an example list extracted. There are two extracted list one is of ecommerce website1 and second is of ecommerce website2.

6) List Weighting

Screenshot 27: List Weighting

In list weighting, all extracted lists are weighted and thus some unimportant or noisy lists, Such as the price list that occasionally occurs in a page can be assigned by low weights. Some of the extracted lists are not informative or even useless. There are two weighted list ecommerce1 and ecommerce2.
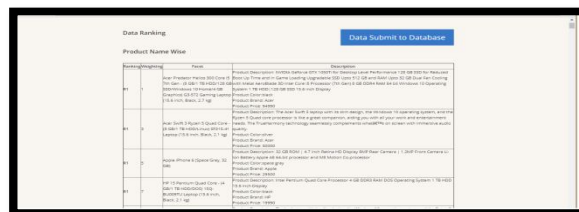
7) *List Clustering*
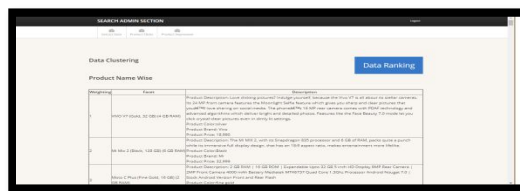
Screenshot 28**:** List Clustering

In list clustering, similar lists are grouped together to compose a facet. For example, different lists about watch gender types are grouped because they share the same items "men's" and "women's". There are two clustered list one is of ecommerce website1 and second is of ecommerce website2.

8) *Data Ranking*

Screenshot 29: Data Ranking

In data ranking, facets and their items are evaluated and ranked . For example, the facet on brands is ranked higher than the facet on colours based on how frequent the facets occur and how relevant the supporting documents are.

III.   ACKNOWLEDGEMENT

It is matter of great pleasure by getting the opportunity of highlighting a fraction of knowledge I acquired during my technical education through this dissertation. The making of the project needed co-operation and guidance of a number of people. I therefore consider it my prime duty to thank my guid, co-guid and principal.

IV.   CONCLUSION

We study the problem of finding query facets. We propose a systematic solution which we

refer to as QDMiner to automatically mine query facets by aggregating frequent lists from free text, HTML tags and repeat regions within top search results. the first approach of finding query facets, QDMiner can be improved in many aspects. We further analyze the problem of duplicated lists, and find that facets can be improved by modeling fine-grained similarities between lists within a facet by comparing their similarities.

REFERENCES

[1] O.Ben-Yitzhak, N.Golbandi, N.Har'El, R. Lempel, A.Neumann, S.Ofek-Koifman , D.Sheinwald, E.Shekita, B.Sznajder and S.Yogev, "Beyond basic faceted search," in Proceedings of WSDM '08, 2008.

[2] M. Diao, S. Mukherjea, N.Rajput, and K.Srivastava, "Faceted search and Browsing of audio content on spoken web," in Proceedings of CIKM '10, 2010.

[3] D.Dash, J.Rao, N.Megiddo, A. Ailamaki, and G.Lohman "Dynamic faceted search for discovery-driven analysis," in CIKM '08, 2008.

[4] W. Kong and J. Allan, "Extending faceted search to the general web," in Proceedings of CIKM '14, ser. CIKM '14. New York,NY, USA: ACM, 2014, pp. 839–848.

[5] T. Cheng, X. Yan, and K. C.-C. Chang, "Supporting entity search: a large-scale prototype search engine," in Proceedings of SIGMOD '07, 2007, pp. 1144–1146.

[6] K. Balog, E. Meij, and M. de Rijke, "Entity search: building bridges between two worlds," in Proceedings of SEMSEARCH'10, 2010, pp. 9:1–9:5.

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)