

SENTIMENT ANALYSIS AND PRODUCT ASPECT RANKING

Sonali Gavande (sonali.gavande30@gmail.com), Amit Dhatrak (amit.dhatrak95@gmail.com), Sagar Padekar (sagar.padekar@gmail.com), Sayli Choudhar (saylichoudhar0806@gmail.com), Prof. D.M. Ujlambkar

All India Shri Shivaji Memorial Society's College Of Engineering Pune

ABSTRACT

Nowadays, there is a trend of online shopping as more and more products are being sold by the manufacturers on the Internet and customers are expressing their views and opinions about a product on the Internet. Traditionally, individuals collect feedback from their friends or relatives before purchasing an item, but nowadays user generated reviews give the information which help customers in buying their favourable products. So opinions of huge number of individuals around the globe can be analyzed.

For that, various opinion mining techniques and sentiment analysis techniques are currently being used and products are being ranked using various ranking algorithms. In this report, we also present the survey of different opinions and reviews mining techniques and product ranking algorithms and their analysis in terms of their effectiveness and complexities. Based on the analysis, a system is chosen consisting of effective product ranking algorithm based on customer reviews. The result of the system is a list of ranked products for the customers.

I. INTRODUCTION

When the world wide web is created, it has undergone various changes, such as even common people started to join the various social networking sites. At the same time, various new technologies came into existence to enrich computer technologies and networking related tasks, costs required for operations related to different web sites and the costs of hardware in computers decreased to a greater extent, all of which together contributed to the popularity of the Internet.

There are some systems available which are ready to use and they have all the facilities necessary for easy installation, publication of the given content and edition of the same, which enable daily users to publish the content online though they may not be computer programmers or computer experts. Because of the use of such facilities, more social forums, personal or social forums and specialized websites have been evolved and the user generated contents have increased to a greater extent and that, too, significantly. As the advanced versions of the web have introduced various new concepts, the active participation of different users became more valuable, different customers and different communities, specially through the contribution of the knowledge of each user to add to the information which is used globally. The main consequence of this new order is a public space which is given for all discussions between community and it became almost a new trend for new internet applications. In order to meet with the new demands of this internet era, most used web sites like social

networking sites, e-commerce, newspapers, magazines available on the internet had to improve the structure of their existing systems to comply with new design standards. It provides a common space which allows interaction of various internet users through the exchange of opinions, views and sharing of experiences. And so as the sharing of opinions started, users have produced a variety of information and heaps of textual data, a valuable content that can prove as extremely useful, and also in many of the cases as a unique and a primary source of information. The recapture and retrieval of this meaningful data just depends on specialized techniques and this is because of this data which is unstructured and free.

Also in Today's world, online shopping has more importance since thousands of reviews about various products are posted on different shopping websites. But it is impossible for consumers to read all those reviews while buying a particular product. In this report, a product ranking system using opinion mining techniques to favorably rank products for users is presented. Our system would provide users to specify product distinctive properties through a query, and send back the ranking results of all matched products. Here, we are using product ranking system proposed by: Tian and Liu considers 3 issues while calculating product scores:

- 1)Product release month
- 2)Product reviews
- 3)Product popularity

Nowadays, the quantity of information that is available actually makes the real system more concerned about how to handle heaps of textual data, which ensures that the web users will have access to the best resources by taking the least effort. In these recent times, special attention is being given to the quantity of the content which is produced by different users. One of the most affected sectors due to this is e-commerce which is affected by the quantity of textual data produced by consumers, which increased to a significant extent during the Web 2.0 phase. Customer opinions represent a very useful and unique information which should be treated in a good manner or care-fully taken care of by the people who work in the field of research. Thus, this work give attention to the necessity of some techniques that will provide researchers some mechanism to fully use this available review data. From the perspective of consumer, considering opinions of other people before purchasing any type of product was a natural common behavior before the existence of Internet.

In the present era of the web development and digital world, the basic point which makes a difference is that a customer can have access to millions of reviews and opinions, which influences the decision making ability of users to a greater extent. Basically, consumers want to find the best products which are available at the lowest price. In other words, they will choose products which are affordable and appropriate to their needs.

It is important to foreground the advantage of analyzing other people's reviews and opinions, which generally come from their nature which is neutral towards any product, and which are not usually associated to a particular organization or company. These opinions represent the voice of ordinary customers, and to a significant extent it differs from those various advertisements (advertisements generally focus on the positive aspects and neglect the negative ones in order to favor the product).

If we view this from the view point of e-commerce, receiving customers' feedback can develop the strategies so that profits of the sector will increase to a greater extent. For example, if we consider the case of online shopping, smart advertisements can be placed by measuring the level of satisfaction of customers for a given product. It is really common to try to explore products with millions of opinions, and hence it could be a difficult task for a consumer to analyze all of them.

An important difference is that it makes the present mechanisms which are available for ranking not so efficient to represent the information portrayed by opinions. This difference is mainly due to nature of textual data given by consumers. This information is either opinions, reviews or facts. The actual search systems are centered around actualities i.e facts (e.g mechanisms available for ranking which are used by search engine). An opinion however is a judgement or belief about a particular issue or a subject. Therefore, one review for the particular object which is under discussion is usually different from multiple reviews for the same object which is under discussion. In the same manner, a summarization mechanism represents the actuality of opinions and it gives us more efficient ways to deliberate i.e to make an inference out of them.

Basically, this work represents the ways for extracting, locating, classifying and summarizing reviews on the Internet. The framework which is put forward combines several techniques to extract information which is valuable from all the user generated reviews which are in natural language text, so that it will provide enrichment of the knowledge of users by taking benefit of the content which is available in a more intelligent and organized way. As a result of the implemented techniques, data can be structured, this will also provide a necessary bridge for many softwares to be able to fully interact with others.

II. LITERATURE SURVEY

1. Web Product Ranking Using Opinion Mining Authors : Yin-Fu Huang and Heng Lin Publication : IEEE 2013

In this paper, the author has present a product ranking system using opinion mining techniques. Users can specify product features to get back the ranking results of all matched products. Opinion mining is very suitable for the applications many users to discuss a single topic when a large number of sentences and reviews are on the criticism of a single topic or when exactly identifying the correct rate of a bias direction is required to improve accuracy. However it had some drawbacks like it could not analyse the aspects that were presented in a paragraph form or a more complex form.

2. Text Rank: A Novel Concept for Extraction Based Text Summarization

Authors : Dipti.D.Pawar, M.S.Bewoor, S.H.Patil
Publication : September 2012

This paper proposes an indexing structure in which index is built on the basis of context of the document rather than on the terms basis. Existing text summarization techniques shows that the indexing is done on the basis of the words in the document and consists of an array of the posting lists. Document features like term frequency, text length are used to assign indexing weight to words. Hence indexing weights of the document words are used to calculate the sentence similarity value between document words which remains independent on context. While doing so they have also used novel concept of Lexical association (semantic association) between document words to calculate the similarity between sentences using computed indexing Weights. The proposed concept of sentence similarity measure has been used with the graph-based ranking method to create document graph and get summary of document.

3. Ranking System for Opinion Mining of Features from Re-view Documents

Authors : Tanvir Ahmad, Mohammad Najmud Doja
Publication : July 2012

In this paper the author proposes a feature wise opinion mining system which first extracts features from user generated contents, then determines the intensity of the opinions by giving emphasis to the modifier of the words,

which expresses opinions. It finds the numeric score of all the features using SentiWordNet and then calculates the overall orientation of the feature to determine how intense the opinion is for both the positive and negative features. The positive and negative features are identified by extracting the associated modifiers and opinions.

4. Multi-Document Summarization of Evaluative Text
Authors : Giuseppe Carenini, Raymond Ng, and Adam Pauls
Publication : May 2011

This paper presents and compares two approaches to the task of summarizing evaluative arguments. The first is a sentence extraction based approach while the second is a language generation-based approach. They evaluate these approaches in a user study and find that they quantitatively perform equally well. Qualitatively, however, they found that it performs well for different but complementary reasons. The general problem considered in this paper is how to effectively summarize a large corpora of evaluative text about a single entity (e.g., a product). They conclude that an effective method for summarizing evaluative arguments must effectively synthesize the two approaches.

5. Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics
Authors : Anindya Ghose, Panagiotis G. Ipeirotis, Member, IEEE,

Publication : IEEE 2010

This paper is the first study that integrates econometric, text mining, and predictive modeling techniques toward a more complete analysis of the information captured by user-generated online reviews in order to estimate their helpfulness and economic impact. All the aspects analysed from the reviews are saved in the dataset and analysed later and the algorithm creates a pattern.

6. Kanayama, Hiroshi and Tetsuya Nasukawa (2013), Fully automatic lexicon expansion for domain-oriented sentiment analysis, In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics Publication : IEEE 2013

Kanayama-et-al proposed an approach to build a domain oriented sentiment lexicon to identify the words that express a particular sentiment in a given domain. By construction, a domain specific lexicon considers sentiment orientation of words in a particular domain. Therefore, this method cannot be readily applied to classify sentiment in a different dataset.

7. Ding, Xiaowen, Bing Liu and Philip S. Yu (2014), A holistic lexiconbased approach to opinion mining, In Proceedings of the 2014 International Conference on Web Search and Data Mining, Association for Computing Machinery, pp. 231-240.

Publication : IEEE 2014

Ding-et-al focused on customer reviews of products. In particular, the author reviewed the problem of determining

the semantic orientations (positive, negative or neutral) of opinions expressed on product features in reviews. So, the author proposed holistic approach that can accurately infer the semantic orientation of an opinion word based on the review context. It provided a new function which is used to combine multiple opinion words in the same sentence.

8. Pang, Bo and Lillian Lee (2015), Opinion Mining and Sentiment Analysis, Foundations and Trends in Information Retrieval, Vol. 2, No. 1/2, pp.1-135.

Publication : IEEE 2015

In Pang-et-al focused on the methods that seek to address the new challenges raised by sentiment aware applications, as compared to those that are already present in more traditional fact based analysis. This paper includes a material on summarization of evaluative text and on broader issues regarding privacy, manipulation, and economic impact that the development of opinion oriented information access services gives rise to. To facilitate future work, a discussion of benchmark datasets is also provided.

9. Ramage, Daniel, David Hall, Ramesh Nallapati and Christopher D. Manning (2010), Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora, In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Vol.1, pp.248-256.

Publication : IEEE 2010

This paper is the first study that integrates econometric, text mining, and predictive modeling techniques toward a more complete analysis of the information captured by user generated online reviews in order to estimate their helpfulness and economic impact.

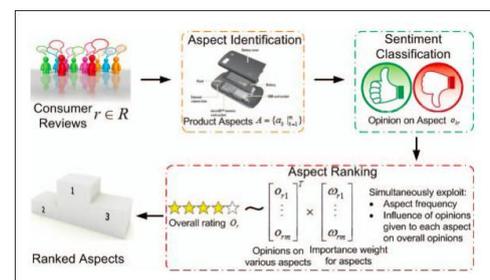


Fig: Architecture Diagram

III. DESIGN AND IMPLEMENTATION

A web application is a program that runs on a computer with a web server, while its users interact with it via a web browser or similar user agent.

Web Application Architecture

A web application uses request response model for its execution. The client requests the information by typing the Uniform Resource Locator (URL). This request is sent to

web server which is present on a system in a network. Web server locates the resource requested, if available either sends it to the client for processing (client-side page) or processes it and sends the response back to the client. Request travels over the internet using protocols like HTTP, known as HTTP request. Response generated by web server is also in the form of a HTTP response.

The components of Web application architecture :

Web server :

Web server can be applied to either of the two components :

1. The physical machine on which the application resides it also known as web server.

2. The software which is used to manage the application on the physical machine is also known as web server.

From the developer perspective, a web server is the software which is used to manage the access to web application. It does not have the capability to host applications. Apache HTTP server is an example of web server.

Web Container :

The data being transferred from one end to another needs to be read, understood, and presented to the user in a usable format. Web browser is a software application that understands and translates HTML content. It allows the user to interact with a remote system by converting his/her requirement into text based content and using internal mechanisms to transfer this content to a processing component.

A web browser typically has a Graphical user interface, which helps the user to understand the content better and faster.

The component that provides services to the user is known as resource" and it is identified using a Uniform Resource Locator"(URL) or Uniform Resource Identifier"(URI). A web browser often referred to as user agent(UA).

Web browser :

Web containers are sometimes erroneously clubbed as web server. The responsibility of the web container is to provide the execution environment for the web application, execute the applications life-cycle and manage the ecosystem in which the application runs.

1. A web container is responsible for checking that URL requestor has valid access rights, mapping a particular URL to a particular servlet in a given application, and importantly managing the life-cycle of servlets. It is the main part of a web server that actually communicates with the Java servlets.

2. A web container gives us a federated view of the Java EE platform APIs and the same services as a JSP container as well.

3. A web container actually implements the agreement of web component of the Java EE architecture, and specifies a runtime environment for web components that includes concurrency, security, deployment, transaction, life-cycle management, and other services.

8.2 Tools and Technologies Used

The Java Programming Language Platforms

There are mainly 3 platforms which the Java programming language uses :

- 1) Java Platform, Micro Edition (Java ME)
- 2) Java Platform, Standard Edition (Java SE)
- 3) Java Platform, Enterprise Edition (Java EE)

We have used java EE platform for implementing the web application.

Java EE Application

An application of JAVA EE or which is generally called as Enterprise Edition Application in a JAVA platform is nothing but a unit which can be deployed and gives us functionality of Java EE Platform. The applications of JAVA EE are actually engineered and programmed to be distributed over multiple computing web-tiers. This can also be a single module or a group of modules for Java EE which can be packaged into an EAR file along with Deployment descriptor of an application of JAVA EE.

Enterprise applications contains the combinations of the following web components:

1. Modules required for connector or resource adapters(which are packaged into RAR(.rar) files).
2. Protocol for Initiation of session(SIP) modules (which are packaged into SAR(.sar) files).
3. Modules which are packaged in JAR(.jar) files and are generally known as EJB modules.
4. Modules which are included in client-side of an application.
5. Additional JAR(.jar) files which consist of classes which are dependent and other components the web application requires.
6. Web modules which are packaged in WAR (.war) files.

Java servlet

Servlets serve the following needs of an internet application:

1. Servlet handles state information which the stateless HTTP protocol cannot give.

2. It Processes or stores data that is required to be submitted from a standard HTML form(basically serves as a business unit in web application).

3. Servlet gives us the content which is dynamic such as the results of a database query.

Basically we will discuss about what is a servlet and its functions. The servlet are the server-side components that extend the capabilities of the server and provide a powerful mechanism for developing server-side programs.

The general features of a servlet are:

1. Efficient
2. Persistent
3. Portable
4. Robust
5. Extensible
6. Secure
7. Scalable

Java Servlets are web components that allow application logic to be embedded in the HTTP request-response process and provide a way by which a web application can be built. Servlets solve many of the common problems faced when using CGI.

Code for initialization of servlet is executed only once when the web server loads it for the rest time. After the server is loaded, handling new requests is only a matter of calling the service method. This approach is much better than loading a totally new program which is executable with every request.

Servlets can maintain state between requests. When the servlet is loaded, it stays resident in memory while serving the incoming requests. When the servlet is initialized, it queries the database table which holds a list of categories and stores these categories in a vector. As a request for a list of categories comes from the client, instead of querying the database again, advantage of the persistent characteristics of servlets can be taken.

Servlets are developed using Java, so they are portable. This enables servlets can be transferred to a new operating system and hardware combination without changing the source. Code that was compiled on a windows NT platform can be moved to a Solaris box without making any changes.

Because servlets are developed with admittance to the whole JDK, they are very robust and powerful solutions. Java gives a well defined exception hierarchy for error handling. It has a garbage collector to prevent problems with memory leaks.

Being developed in an object-oriented language like Java is, servlets can be expanded and made to polymorph to new objects as per requirement

. A good example is online catalog. Suppose the same catalog search tool is to be developed at the top of every dynamic page in the web site, the same code need not be added to each of the servlets.

At the the server-side of the web Architecture, servlets run by inher-iting the security provided by the web server. Servlets can also take the advantage of Java security manager.

Servlets are widely accepted because of the advantages gained from using Java. Vendors are providing servlets support in 2 main forms. The list form is servers that have built-in support for servlets, and the second is by using third-party add on.

Java Language

Java language aims to let the developer of an application write only once, and can run anywhere, means any java code code that runs on one plat-form can run on any other platform. The same code does not need to be recompiled to run on any other platform. Java is actually is a programming language which is mainly object-oriented, and class-based, and which is de-signed with speci cations to have less implementation dependencies like no platform dependency. Applications in java generally compiled to bytecode (.class le) which can run on any Java virtual machine (JVM) in spite of any type of computer architecture.

Java Platform And It's Features

There are following main features of Java :

1. Object-oriented
2. Simple
3. Robust
4. Distributed
5. Secure
6. Architecture neutral
7. Portable
8. Interpreted
9. Multi-threaded

Java is object-oriented language.

Java language is simple in the sense that it removes some of the annoy-ing features of other object-oriented languages such as c++ that have been troubling programmers for a long time. For example, there are no header-less to be included which increased the code unnecessarily and were

required because the library consisted of functions such as API. In Java, API consists of classes which are instantiated on demand.

Any language is called as robust if it gives mechanism for the early detection of potential problems. Java provides a way in which the developer

does not have to worry about chronic problems such as bad pointers, memory allocation errors and memory leakage. This is made possible because JVM does not rely on operating System for management.

Applications are said to be distributed if they run in different runtimes and try to interact with each other to achieve functionality. Java provides an API which allows seamless communications between separate modules without exposing the network complexities to the developer e.g., socket programming.

The JVM needs a bytecode file as its executable. Every executable has its header that gives details about the file and also helps identify the file as a valid executable file for that machine. If the file gets corrupted, then a bytecode verifier checks it out with the information in header such as parity check, etc. and discards the file if it is corrupted. And through this, we can say that Java language provides security.

Java is compiled as well as interpreted. First, the source code of any Java program is compiled and gives the bytecode and then the obtained bytecode is interpreted at every stage of the execution to its native code.

Java Server Pages

JSP - JAVA Server Pages : Java Server Pages is a Java Technology that generates dynamic web pages based on HTML. JSP life cycle divided in 7 phases:

Translation : JSP translated to servlet source code. Here JSP evaluated for syntax and tag less.

compilation : Servlet source code is compiled into java bytecode i.e a servlet class . Translation and compilation can occur anytime between deployment and first request by the user.

Loading : The application loads the servlet class into the memory of container using the applications class loader.

Instantiation : The application instantiates an instance of the servlet class for each JSP in the application.

Initialization : The `jspinit()` method is called on the servlet instance to initialize it.

Service : The `jspservice()` method is invoked for each request received for the JSP. This method cannot be overridden.

Destruction : The servlet instance is taken out of service `jspDestroy()` method called.

Figure 8.1: Java Server Pages

HTML

HTML is the basic building block for creating a web application. HTML also ensures that the application reaches multitudes rather than being dependent on technology being used. HTML stands for Hyper Text Mark Up language. It is not a programming language, it is markup language. A markup language is a set of markup tags. HTML uses markup tags to describe web pages. Due to cross platforms data transfer issues, text based data transfer is preferred over object based data transfer. The language made up of textual data and control sequences is HTML. HTML has different tags, attributes, and many other components. Using HTML, text can be customized, content can be given a specific style, other resources can be linked, HTML tables and lists can be created, data forms can be created, data inputs can be given, etc. HTML can be embedded in known languages such as java easily and javascripts can be used. There are wide applications of HTML besides these applications also.

Apache Tomcat

Apache Tomcat is a type of web container which in general called as non-commercial. It is basically a servlet container derived by the Apache Software Foundation(ASF) and it is also a web server which is open source. It actually programs the JavaServer Pages (JSP) specifications and the Java Servlet given in an application, and provides an environment which is basically pure JAVA for any Java code to run properly. With one of the most simplest configurations, it also runs in a single operating system process (known as uniprocessor). Every process runs a Java virtual machine (JVM) which is the core of JAVA Architecture. Apache Tomcat also includes tools for management and configuration. Each and every HTTP request that comes from a browser to the server of Apache Tomcat is always processed in each of the separate threads.

In short, Apache tomcat server is-

1. Apache Tomcat basically empowers numerous mission-critical, large-scale web applications across a varied range of institutions organizations and industries.

2. It is an basically an implementation of open source software of the Java Servlet and JavaServer Pages technologies. The Java Servlets defined as business units in an application and JavaServer Pages specifications are derived under the Community Process of JAVA.

Wordnet and POS tagging

Wordnet

WordNet is basically a database which is lexical for the English language. It collects words in English sentences and groups them into collection of synonyms which is basically called as synsets, and gives us general and short concepts and definitions, and records the different relations which are semantic between these synsets.

The aim is to serve two purposes :

1. To produce a collection of thesaurus and dictionary words that is more usable and comparable to the input stream of given words.

2. To support the analysis of text data which is basically automatic text data and applications in artificial intelligence domain. All the lexical databases and the software are basically published under the recognition of a BSD style license and they can be downloaded and used as an open source. We can get the lexical database by online browsing.

Part of speech tagging

The POS tagging generally referred to as Part-of-speech tagging (i.e. POST), is also called as word category disambiguation or grammatical tagging, and it is basically the process of giving the correct part of speeches which can be verb, adverb, adjective, noun, pronoun, etc.// To each word in a data that is textual and which is based on both of its context and definition as follows:

This DT
 article NN
 is VBZ
 about RB
 the DT
 sport NN
 . SENT

Algorithm Details

A Sentiment classifier is learned from the Pros reviews (positive reviews) and cons reviews (negative reviews). The classification is done using Naive Bayes model classifier.

The Pros and Cons reviews have explicitly categorized positive and negative opinions on the aspects. These reviews are valuable training samples for learning a sentiment classifier. Pros and Cons reviews are used to train a sentiment classifier, which is in turn used to determine consumer opinions on the aspects in free text reviews.

First sentiment terms in Pros and Cons reviews are collected, then the classifier is trained using these sentiment terms and this trained classifier is used to classify the aspect in free text review.

Algorithm Implementation

```
public class BayesClassifier T, K extends Classifier T, K f
```

f

```
private float featuresProbabilityProduct(Collection T features, K category)
```

```
float product = 1.0f;
```

```
for (T feature : features)
```

```
product *= this.featureWeighedAverage(feature, category);
return product;
```

g

```
float
private float categoryProbability(Collection T features, K category)
```

```
return ((float) this.categoryCount(category)
```

```
 / (float) this.getCategoriesTotal())
```

```
* featuresProbabilityProduct(features, category);
```

g

f

Collection T

```
private SortedSet Classification T, K
categoryProbabilities(
features)
```

```
SortedSet SortedSet Classification T, K
probabilities =
new
```

Classification T, K (

```
new TreeSet Classification
```

f

```
@Override
```

```
Comparator Classification T, K ()
```

```
public int compare(Classification T, K o1, Classification T, K o2) f
```

```
int toReturn = Float.compare(
```

```
o1.getProbability(), o2.getProbability());
```

```
if ((toReturn == 0)
&& !o1.getCategory().equals(o2.getCategory())) toReturn =
-1;
```

```
return toReturn;
```

```

g
g);
for (K category : this.getCategories())

probabilities.add(new Classification T, K (
features, category, this.categoryProbability(features,
category))); return probabilities;

g
@Override

public Classification T, K classify(Collection T features) f
if (features.size() == 0) f

SortedSet Classification T, K probabilities =
this.categoryProbabilities(features);

(probabilities.size())

return probabilities.last();

g
return null;

g
Collection T

public Collection Classification T, K
classifyDetailed(
features)
return this.categoryProbabilities(features);

g g

```

IV. CONCLUSION AND FUTURE SCOPE

A product aspect ranking System is used to identify the important aspects of products from numerous consumer reviews. System contains three main components, i.e. product aspect identification, aspect sentiment classification, and aspect ranking. First, system used the Pros and Cons reviews to improve aspect identification and sentiment classification on free-text reviews. Then an aspect ranking algorithm is used to calculate the weight of various aspects of a product from numerous reviews. The product aspects are finally ranked according to their weight.

Proposed method shows the performance improvement over the two existing systems frequency based system and TFIDF based system in terms of DCG by 9.7. We have planned to identify the important aspects of a product from online consumer reviews. Our supposition is that the important aspects of a product should be the aspects that are

frequently commented by consumers and consumers opinions on the important aspects greatly pressure their over-all opinions on the product. Based on this assumption, we have developed a aspect ranking algorithm which will identify the important aspects by concurrently considering the aspect frequency and the pressure of consumers opinions given to each aspect on their overall opinions.

Till now, we have achieved the ranking of the products which are purely based on the reviews given by different users about different products. In the future, ranking can be achieved by considering the various technological features of the products in addition with the reviews generated by users. For example, for product such as laptop, ranking will be done in accordance with the features such as capacity of Hard Disk, RAM, cost etc. User will enter the feature such as cost, that means user will expect good quality in less cost, and by considering the different models in laptop in an ascending order of cost and that too with the good score of product reviews, those models of laptops will be given higher ranks.

V. REFERENCES

- [1] Y. Rui, T.S. Huang, and S.-F. Chang, Image retrieval: Current techniques, promising directions, and open issues, *Journal of visual communication and image representation*, vol. 10, 1999, pp. 3962.
- [2] S.K. Chang and A. Hsu, Image information systems: where do we go from here?, *IEEE transactions on Knowledge and Data Engineering*, vol. 4, 1992, pp. 431442.
- [3] G. Mailivasan Parthiban Karthikram, Tag based image retrieval (TBIR) using automatic image annotation, *International Journal of Re-search in Engineering and Technology*, vol. 03, 2014.
- [4] T. Kato, Database architecture for content-based image retrieval, *SPIE/IS& T 1992 Symposium on Electronic Imaging: Science and Technology*, International Society for Optics and Photonics, 1992, pp. 112123.
- [5] H. Zhang H. Dagan F. Long and D. Fen, Fundamentals of content based image retrieval, in D. Feng, W. Siu, H. Zhang (Eds.): *Multimedia Information Retrieval and Management. Technological Fundamentals and Applications*, Springer-Verlag, .
- [6] Ricardo da Silva Torres ,Alexandre Xavier Falcao,"Content-Based Image Retrieval: Applications",Published by RITA Numero 2,2006
- [7] M. H. Saad, H. I. Saleh, H. Konbor, M. Ashour, "Image Retrieval based on Integration between Color and Geometric Moment Features", Published by the Arab Journal of Nuclear Sciences and Applications,45(2)447-454(2012)
- [8] Douglas Baquero, Juan Molina, "An image retrieval system for tomato disease assessment ", Published by IEEE 978-1-4799-7666-9/14/31.00,2014

[9] Sudeep Marwaha, Subhash Chand, Arijit Saha, Disease Diagnosis in Crops using Content Based Image Retrieval,IEEE 978-1-4673-5119-5/12, 2012.

[10] Sushant Shrikant Hiwale, Dhanraj Dhotre, Content-Based Image Re-trieval: Concept and Current Practices, IEEE978-1-4799-7678-2/15, 2015.

[11] Balasubramanian Subbiah1, Seldev Christopher, Image Classification through integrated K-Means Algorithm, IJCSI Vol 9, Issue 2 No.2, March 2012.

[12] Poorani, Prathiba, Ravindran, Integrated Feature Extraction for Image Retrieval, IJCSMC Vol. 2, Issue. 2, February2013.

[13] M. H. Saad, H. I. Saleh, H. Konbor, M. Ashour, Image Retrieval based on Integration between Color and Geometric Moment Features, Arab Journal of Nuclear Sciences and Applications,45(2)447-454(2012), 2012.

[14] S. Mangijao Singh, K. Hemachandran , Image Retrieval based on the Combination of Color Histogram and Color Moment, International Journal of Computer Applications Volume 58 No.3, November 2012.

[15] Hong Liu, Xiaohong Yu, Application Research of k-means Clustering Algorithm in Image Retrieval System,ISCSCT ISBN 978-952-5726-07-7, Dec 2009.