International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

An Optimized System for Distributed Information Retrieval and Analysis

Raghavendra.Sheddi[1*], Meenakumari.V.Umarani[2]
[1]ComputerScience & Engineering department, R. T. E. Society's Rural Engineering college Hulkoti, Gadag, Karnataka , India-582205
[2]Computer Science & Engineering department, R. T. E. Society's Rural Engineering college Hulkoti, Gadag, Karnataka , India-582205
*Corresponding AuthorEmail: raghuis016@gmail.com

Abstract: The recent technology development fascinates the people towards information and its services. Internet, has transformed how people interact with information. Much of the routine information access by the general public is now based on full-text information retrieval. Full-text information retrieval consists of discovering database contents, ranking databases by their expected ability to satisfy the query, searching a small number of databases, and merging results returned by different databases. This paper provides useful and simplified way for the users for the adaption to retrieve the distributed data. The collected data at the source is clustered using k-means clustering technique and cluster weights are assigned. Next, the quality of reviews is assessed and classified based on cluster weights. The user is provided with the summarized opinion about purchase of the product.

Keywords:- Information Retrieval, Web mining, K-means algorithm

1. Introduction

Internet has been emerging from an information domain to a market domain with thousands, potentially millions, of electronic stores, sales and other profitable service area. This creates crucial opportunities, but is not without problems. The information overload is an obstacle to the practical use of potentially useful information on the Web.

Theory ofweb mining comprises of methods for summarizing, classification and clustering of the web contents. The system provides valuable and motivating patterns about usersnecessities and contribution behavior. Itaims the knowledge innovation, in which the main objects are the customary collections of text documents and, more recently, also the groups of multimedia documents such as images, movies, sound clips, which are embedded in or associated to the

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

online content. It is primarily grounded on research in information retrieval and text mining, such as information mining, text classification and combination, and information visualization. Some of the major web content mining methods are as follows:

a) Unstructured data mining techniques
b) Structured data mining techniques
c) Semi structured data mining techniques
d) Multimedia data mining techniques

2. Literature Survey

[1] In this paper, authors put forward that, "the number of customer's reviews that a product gets is increasing with rapid speed. The superiority of Customer reviews displayed on the websites differ significantly. In the present concept, author makes an attempt to assess a review based on its superiority, to help the customer make right selection of the product. The quality of customer reviews is evaluated as most-significant, more-significant, significant and insignificant. An innovative and active web mining technique based on review grouping is proposed for evaluating a consumer review of a specific manufactured goods".

[2] In this paper, authors put forward that, "sellers marketing products on the Net request their consumers to review the goods and related services. As e-commerce is became superfluous popular, the amount of customer reviews that a manufactured good obtains grows rapidly. For a standard product, the amount of reviews can be in hundreds. This makes it difficult for consumers to read them in edict to make a conclusion whether to buy the manufactured good. In this concept, author aims to summarize all the consumer reviews of a manufactured good. This summarization job is dissimilar from old-style text summarization since author is only fascinated in the precise features of the manufactured good that customers have thoughts on and also whether the ideas are progressive or destructive. We do not summarize the reviews by picking or rephrasing a subset of the original verdicts from the reviews to capture their main facts as in the typical text summarization".

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

[3] In this paper, authors suggested that, "Web mining essentially concentrates on education about web user with their interaction with web sites and application of web to extract knowledge from World Wide Web. The intention of web mining is to find user's access object robotically and punctually from the enormous web record data such as regular access routes, regular access sets and grouping of data. This article provides examination and investigation of existing web mining method and tools.

[4] In this paper, authors suggested, "Perceptions into web mining methods, procedures and its applications in the current cut-throat industry environment as well in investigation and mining contents for learning determinations. It further explains how using web content mining shows vital role by getting rich set of contents and uses those contents in the conclusion building in the business atmosphere, learning and investigation".

[5] In this paper, authors suggests that "Web mining uses various data mining procedures to determine valuable acquaintance from Web hyperlinks, sheet content and usage record. The key uses of web content mining are to collect, classify, consolidate and deliver the best potential information available on the Web to the consumer demanding the information. The mining tools are imperative to scanning the many HTML documents, pictures, and script. Then, the result is used by the search engines. In this paper, authors introduce the concepts related to web mining; and then present an impression of diverse Web Content mining tools".

3. Methodology

The overall system is divided into two sub-systems. The first sub-system provides the overall functionality of the system. The retrieving of data from the destination systems across the Internet and the processing operation carried out to obtain the final result of the proposed project.

The second sub-system defines the data processing functionality. This sub-system is a part of first sub-system. Here the complete web mining process is carried out to obtain the overall summary of the data in the form of results.

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

The system performs the operation in 3 phases. Each of the sub-systems are discussed in detail in the further sections in this chapter.

3.1 Product Review System

The system is proposed to enable the Product Review System to have the functionality mentioned in Figure 3.1. This process is Phase One of the overall system design. The sub system represented has the following components:

1) Product Review System
2) User Interface Module
3) Data Processing Module
4) Database
5) Review Collection Process
6) Destination Systems

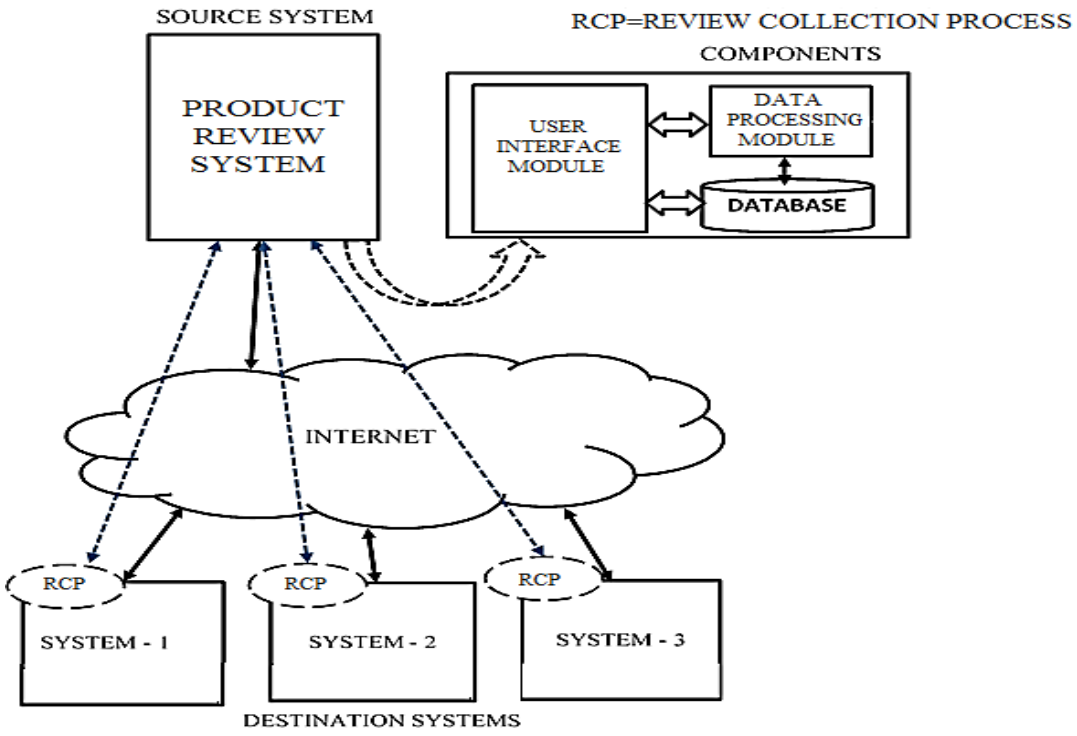

Figure 3.1 Product Review System

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

1) **Product Review System:** This is a source system from which all the processes related to the operation are initiated. The Product Review System is connected to the Internet from which distributed data is retrieved from one or multiple destination systems. This operation is performed using Review Collection Process. This process is initiated from Product Review System, connects to the destination system to retrieve the data.
2) **User Interface Module:** This is the part of the Product Review System. User Interface module accepts input for which data has to be searched over the Internet. Once Review Collection Process retrieves data, the User Interface module takes the data and stores in the Database. Through this module data processing is also initiated.
3) **Data Processing Module:** This module is taken as the next sub-system of the project. Through the User Interface module, the data processing module is initiated. The retrieved data i.e. the Raw Reviews stored in database is further processed. The complete data mining process is defined in this Data Processing Module. The detailed operation carried out is discussed in the next section shown in Figure 4.2.
4) **Database:** Here the data retrieved by Review Collection Process, data processed by the Data Processing Module is stored, also including the final mining process result. Through the User Interface Module, the final results are represented to the customer.
5) **Review Collection Process:** The main objective of this process is to retrieve Raw Reviews in the form of Pros and Cons from the destination system represented in the web pages. It is the process initiated at Product Review System for retrieving the Raw Reviews from one or more destination systems.
6) **Destination Systems:** These are the systems where distributed data resides i.e. Raw Reviews. These are spread across the Internet from which data is retrieved and stored in the database of the Product Review System.

3.2 Data Processing and Analysis System

The second sub-system of the project defines the overall functionality for data processing and analysis. This is the Phase two of the complete system. Since the retrieved data is Raw Reviews of a product, this data is further processed and analyzed to obtain the overall summary of a product based on customer's reviews on that particular product. Figure 3.2 defines the

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

complete procedure of processing and analysis of raw reviews. The Pros and Cons are processed separate and quality of reviews in both the categories is determined [1].

The functionality of Data Processing and Analysis System is categorized into following operations:

1) Construction of a Review's Feature Matrices

2) Grouping of customer reviews.

3) Group Weight computation.

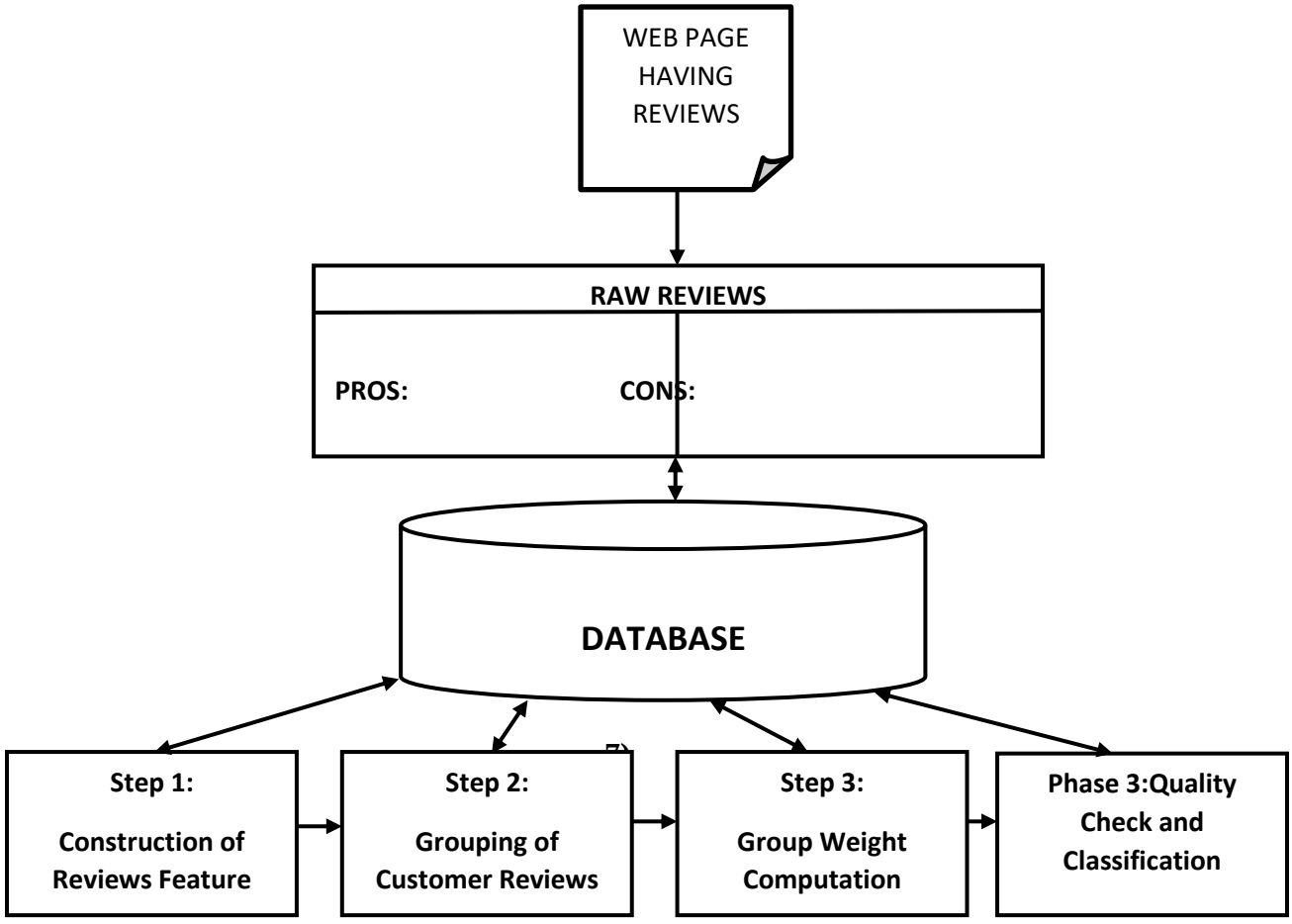

Figure 3.2 Data Processing and Analysis System

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

1) Construction of a Review's Feature Matrices:

The inputs for this component are the set of raw reviews and the feature set extracted in the previous phase. Consider that there are a total of m customer reviews for a particular product and n features are extracted from each of the reviews. We construct a review matrix M of order of m x n using the procedure 1 [1].

Procedure 1: Algorithm for Construction of a Review's Feature Matrices.

```
for each review Ri in the raw reviews database
{
    for each feature fj in the review
    {
        if feature fj is present in Ri then
            Mij=1
        else
            Mij=0
    }
}
```

2) Grouping of Customer Reviews.

Now, we propose to group the reviews into four groups by applying a k-means clustering technique with k = 4 and absolute difference of two data values as the distance measure, for the data set of reviews present in review matrix. The input to this component is the review matrix constructed in the Procedure 1. The algorithm for grouping reviews based on clustering technique is given in Procedure 2[1].

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

Procedure 2: Proposed algorithm for Grouping of customer reviews by clustering

1- Construct the review matrix M for the review set using Algorithm 1.

2- Apply k-means clustering technique with k = 4 for the review set and obtain four clusters of reviews.

3- For each cluster, compute cluster weight Wg, g=1 to 4, as shown below:

a) Compute feature wise sum of the reviews in gth cluster, given by

$$Ygj = \sum_{i=1}^{p} Xij, \text{ for } j=1 \text{ to } n \quad \text{---------- (1)}$$

Where,

n = number of features of the product

p = number of reviews of product in the g^{th} cluster

Y_{gj} = sum of j^{th} feature of all the reviews belonging to g^{th} cluster.

X = sub matrix of review matrix M for g^{th} cluster.

b) Compute the feature wise weight vector WV_g for g^{th} cluster given by

WVg = (WVg1, WVg2, ….WVgn) ---------- (2)

where,WVgj = Ygj/p, for j=1 to n, are the jth feature weights for gth cluster of reviews

3) Group Weight Computation.

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

Mark the clusters as G1, G2, G3 and G4 groups with declining order of their cluster weights Wg, for g = 1, 2, 3, 4. The corresponding feature weight vectors WVg, g = 1, 2, 3, 4, are the representative vectors of the clusters G1, G2, G3 and G4, respectively [1].

Compute cluster weight Wg for gth cluster given by

$$Wg = \sum_{j=1}^{n} WVgj \quad \text{---------- (3)}$$

where, WVgj is weight of the jth feature of the gth cluster

3.3 Quality Check for the Reviews and Classification

The third phase of the system is to find out the group to which a specified review fits based on its quality. A review from the raw review database, and the feature weight vector of each cluster are the inputs for evaluation of the review quality. The algorithm for review quality assessment is given in the Procedure 3 [1].

Performing the Quality Check for the Reviews and classifying them for each review Classification is performed to decide whether the review is Most Important, More Important, Important or Non Important Review.

Procedure 3: Algorithm for Quality Check for the Reviews and Classification.

1- Gather the Review online from the webpage, identify and quote the features that appear in the given review and store as a New Review Vector (NRV).

2- Compute dot product $NRVS_g$ of NRV and WV_g, g = 1, 2, 3, 4, given by

$$NRVSg = \sum_{j=1}^{n} (WVgj)\,(NRVj) \quad \text{---------- (4)}$$

3- Determine value of $NRVS_{max,}$ as

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

$NRVS_{max}$ = max (NRVS1, NRVS2, NRVS3, NRVS4) ---------- (5)

4- Perform the Quality check and classify the raw review using the following condition:

If $NRVS_{max}$ = NRVS1, then the Review is Most Important (MIR).

If $NRVS_{max}$ = NRVS 2, then the Review is More Important (MoIR).

If $NRVS_{max}$ = NRVS 3, then the Review is Important (IR).

If $NRVS_{max}$ = NRVS 4, then the Review is NonImportant (NIR).

The final output of the system provides the quality check result for all the reviews in the form of pros and cons. The final result table shows the total count of the reviews performed quality check and categorized based on above conditions specified in step-4.

4. Results and discussion

In this section, the effectiveness of the system and its efficiency in checking the quality of Reviews from pros and cons are described. The experiments were conducted by taking reviews from the web pages and assessing them as Most Important Reviews, More Important Reviews, Important Reviews and Non Important Reviews.

4.1 Result of Distributed Information Retrieval and Analysis

The result is obtained from online extraction of reviews. The reviews extraction is a real time process which is the first phase of the system. During Review extraction, reviews in the form of pros and cons are taken from the user commented reviews in the epinion.com website. This site posts the reviews on a specific webpage, that is identified and operation is performed.

Next, is the features extraction process is performed. Here we analyze and select the features based on their frequency of occurrence in the reviews. We select the features with high frequency of occurrence. The Feature Review Matrix is constructed based the features obtained.

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

For the purpose of performing the quality check on the reviews and classification, some reviews are selected on the same product and NRV is created which the matrix constructed for selected reviews.

The final classification result is obtained after selection of reviews under which particular category it falls. The result is shown in the following Tables 6.1 and 6.2. The result obtained is tested for two different devices and the below table shows the sample results based on quality check performed.

The system performance is also analyzed based on the execution time complexity. It is the time required for the execution of the different processes in the system.

Product: Canon EOS 20D Camera			
Total number Reviews for analysis: Pros: 100, Cons: 100			
Classification	**Pros**	**Classification**	**Cons**
Most Important Reviews (MIR)	0	Most Important Reviews (MIR)	24
More Important Reviews (MoIR)	11	More Important Reviews (MoIR)	13
Important Reviews (IR)	24	Important Reviews (IR)	3
NonImportant Reviews (NIR)	25	NonImportant Reviews (NIR)	20

Table 4.1: Result of an Optimized Distributed Information Retrieval and Analysis System for the product Canon EOS 20D Camera

Product: Apple Iphone 5s			
Total number Reviews for analysis: Pros: 100, Cons: 100			
Classification	**Pros**	**Classification**	**Cons**
Most Important Reviews	35	Most Important Reviews	0

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

(MIR)		(MIR)	
More Important Reviews (MoIR)	1	More Important Reviews (MoIR)	28
Important Reviews (IR)	0	Important Reviews (IR)	1
NonImportant Reviews (NIR)	4	NonImportant Reviews (NIR)	11

Table 4.2: Result of an Optimized Distributed Information Retrieval and Analysis System for the product Apple Iphone 5s

4.2 Snapshots

Snapshots represent the system implementation and execution process of the system. These provide the overall system representation and implementation. Each and every phase of the proposed system is shown step by step in the following snapshots.

1) Account Authentication

Figure 4.1: Account Authentication for the Customer

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

2) Option Selector to perform the operations in the system

Figure 4.2: Option Selector for the customer to perform further operation

3) Reviews Extraction Process

Figure 4.3: Reviews Extraction Process

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

4) Extracted Reviews

Figure 4.4: Extracted Reviews from the website

5) Features Extraction

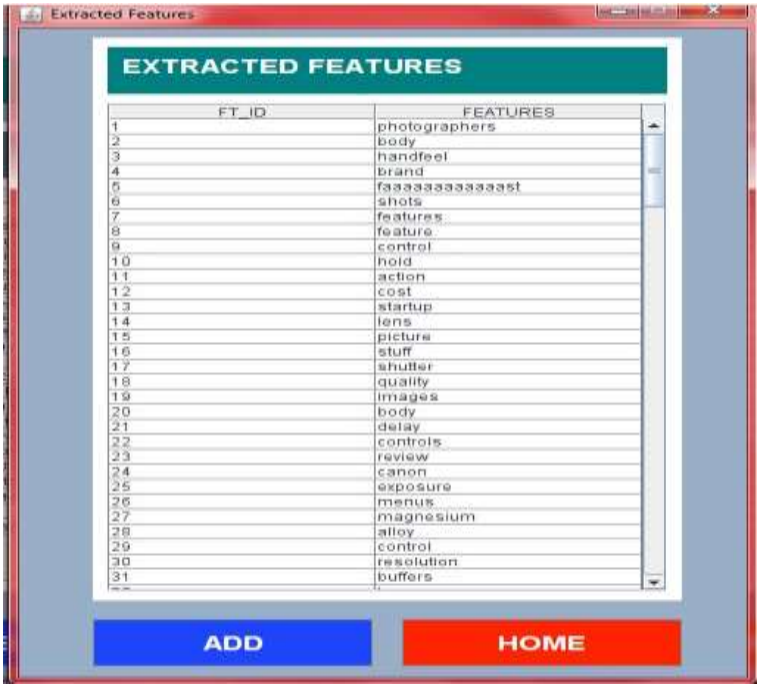

Figure 4.5: Features Extraction Process

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

6) Review Clustering

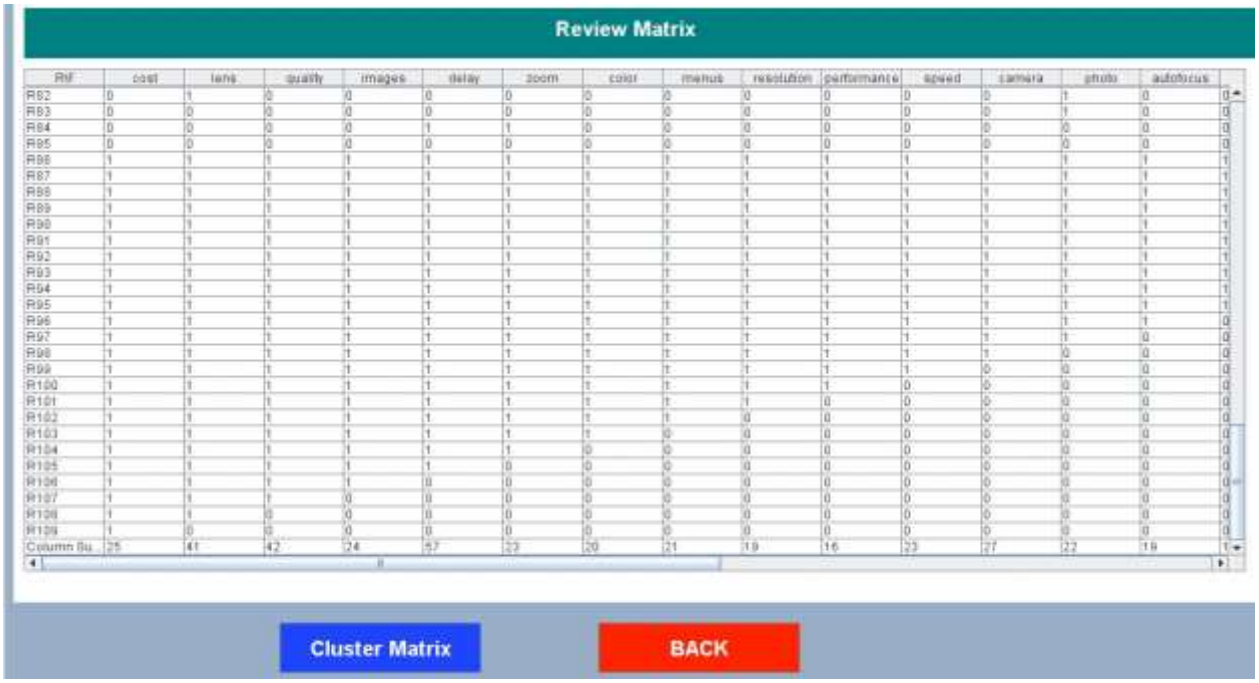

Figure 4.6: Review Matrix

7) Clustered Reviews

Figure 4.7: Clustered Reviews into 4 groups

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

8) Calculation of Weight Vector Table

Figure 4.8: Weight Vector Table

9) New Review Vector Matrix Construction

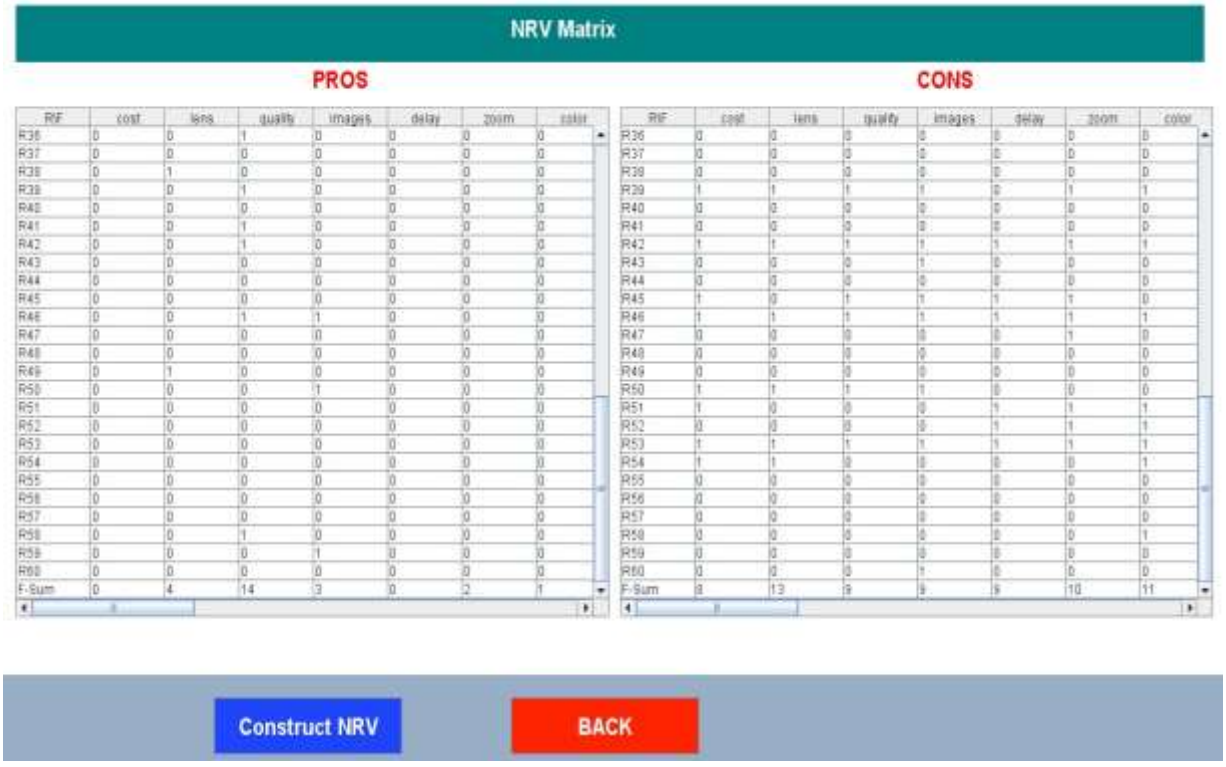

Figure 4.9: Quality check NRV Matrix

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

10) NRV Calculation

Figure 4.10: NRV Values Calculation

11) Final Quality Check Result

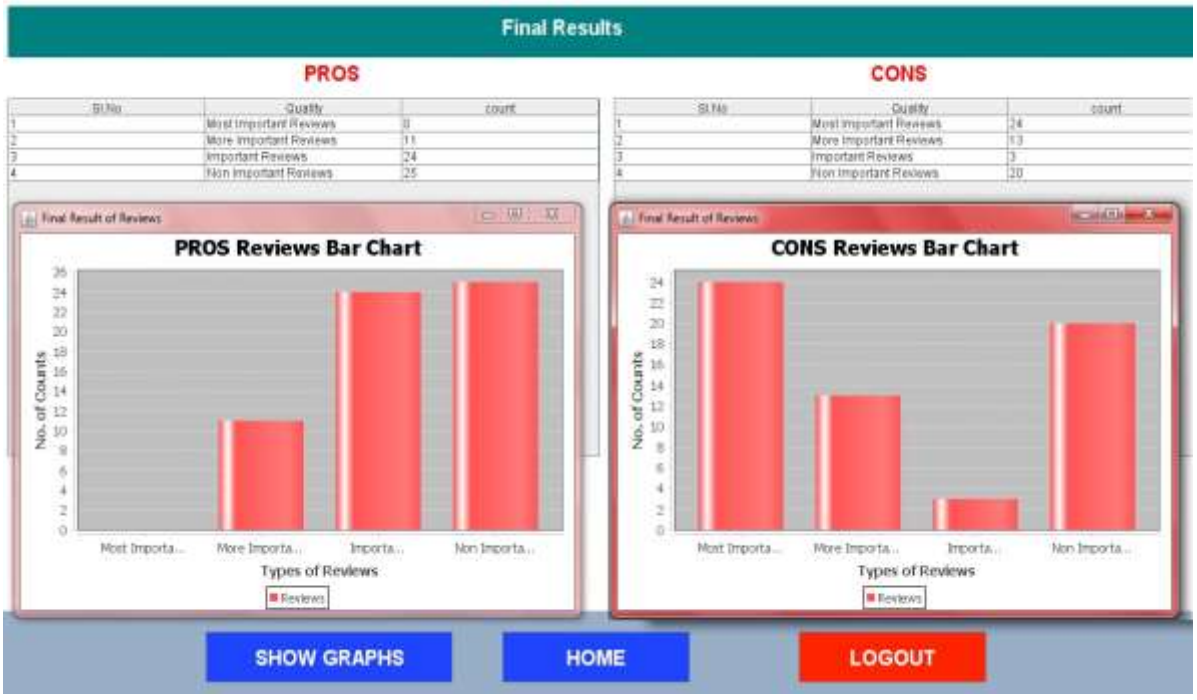

Figure 4.11: Final Quality Check Result and Graphs

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

5. Conclussion

The system designed and implemented for Distributed Information Retrieval and Analysis performs effective processing of the distributed data, with new methodology of performing quality check and classification of reviews. This method when compared with previous techniques provides better accuracy in classification of reviews. The ease to analyze and work with the flow of the system provides enhanced and optimized results for the users to determine over the products based on Reviews on the product.

The system currently implemented has a limitation with respect to parts of speech process applied during features identification and extraction. Apart from this limitation, the overall system provides a better performance with respect to analysis and quality check.

The present system implemented can be enhanced using Mobile Agents for distributed information Retrieval. Mobile Agents are Aglets implemented using Java Programming language that highly reduces the utilization of network resources such a bandwidth consumption. An Aglet is a code that can migrate to the destination system carrying its code and state of the execution with it. Then aglet resumes its execution and performs necessary actions being on the destination machine. Mobile Agents in Distributed Information and Data Retrieval lead to revolution in the field of networks and utilization of bandwidth and other resources over the network.

6. Referrences

[1] P.S Hiremath, Siddu P. Algur, S. Shivashankar, Quality Assessment of Customer Reviews Extracted From Web Pages: A Review Clustering Approach,Dept. of P.G. Studies and Research in Computer Science, Gulbarga, Karnataka, India, Vol. 02, No. 03, 600-606, (IJCSE) International Journal on Computer Science and Engineering, 2010.

[2] Minqing Hu and Bing Liu, Mining Opinion Features in Customer Reviews, Department of Computer Science University of Illinois at Chicago, American Association for Artificial Intelligence, (www.aaai.org), 2004.

[3] PradnyeshBhisikar, Prof. AmitSahu, Overview on Web Mining and Different Technique for Web Personalisation, Vol. 3, Issue 2, pp.543-545, International Journal of Engineering Research and Applications (IJERA), March -April 2013.

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 9 - September 2017 - Pages 29-47

[4] GovindMurariUpadhyay, KanikaDhingra, Web Content Mining: Its Techniques and Uses, IITM India, Volume 3, Issue 11, International Journal of Advanced Research in Computer Science and Software Engineering, www.ijarcsse.com, ISSN: 2277 128X, November 2013.

[5] AbdelhakimHerrouz, ChabaneKhentout, MahieddineDjoudi, Overview of Web Content Mining Tools, Volume 2, Issue 6, ISSN: 2319 – 1813 ISBN: 2319 – 1805, 2013.

[6] Altaf Husain, ShreedharYadawad, RanganathYadawad, Mobile Agents for Distributed Information Retrieval, International Journal of Modern Trends in Engineering and Research, IJMTER,www.ijmter.com, SJIF): 1.711, Volume 02, Issue 03,, e-ISSN: 2349-9745, p-ISSN: 2393-8161, March – 2015.