INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

AN OVERVIEW STUDY ON WEB MINING IN ECOMMERCE

Computer Science

|	
H.Mohana	Ph.D Research Scholar, Department of Computer Science, Dr Ambedkar Government Arts College (Autonomous), Vyasarpadi, Chennai – 39.
DR M.Suriakala	Assistant Professor and Research Advisor, Department of Computer Science, Dr Ambedkar Government Arts College (Autonomous), Vyasarpadi, Chennai – 39.

ABSTRACT

Electronic Commerce (E-Commerce) is an icon of communication, data management and security management that allows business application with different organization. It automatically exchanges information to the sales of goods and services. Like an operating system, it acts as an interface between seller and the buyer. In an internet world E- commerce industry plays a vital role in the WWW. E-commerce website becomes an unbeatable one and it shows the rapid development in it. Enterprises mainly focus on rebuild the relation with their old customer and simultaneously focussing on new customer. The effectual technology used in E-Commerce website is data mining and web mining. Data mining is the process of extracting immersing patterns from large database. Whereas, web mining is the usage of data mining techniques to extract interesting information from the web data. This paper mainly focuses on sentiment analysis about the most famous E-Commerce site in India Is Flipkart in the year 2015.Semantic based approach is used to find the users opinionated phrases. In this paper, we try to elucidate the basics of e-commerce, the challenges of e-commerce in India and finally we made opinion analysis about the famous e-commerce company Flipkart.

I. INTRODUCTION

Web mining is a data mining technique extracting interesting implicit information and potentially useful patterns from www. Taxonomy of web mining defines three different characteristics.
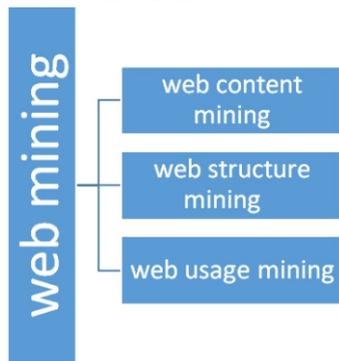
Fig 1: Taxonomy of Web Mining

1. Web content mining
It analyse the content of web resources. Mining the web content like
- Mining text and multimedia documents like text, image, audio, video.
- Web page content mining
- Search result mining.

Two types of Approaches used in web content mining one is database approach another one is agent based approach. Database approach helps to retrieve the semi structured data from the web. Whereas agent based approach helps to search and collect the relevant information from the web.

2. Web structure mining
Mining the website structure using hyperlink. It is the process of inferring knowledge from www and link between references in the web. The main goal is to summarize the structural behaviour about web pages and web sites. It shows the relationship between user and the web. The main goal of web structure mining is defining and analysing the Hyperlink and the Document structure.

3. Web usage mining
It is also called as web log mining. Web usage mining focuses on knowing the users behaviours to access the internet. The main goal of web usage mining is data pre-processing, pattern discovery and pattern analysis.

E-Commerce is a medium where buying and sellers are more important into it. The E-Commerce deals with entire online process of buying, selling, paying, services of different product to the customer. E-Commerce is an interface where we can able to buy product directly from the sellers. So that customers could able to get more discount comparing to retail shop. In retail business we go and buy the product for that we waste our precious time going to one shop to other shop and asking bargain to retails. But here that type of problem is not there. Through E-Commerce we can save time, energy and the most important thing is money. It provides a cost efficient and effective way of doing business in the web. The term E-Commerce is not just buying and selling products over internet rather than that it has the capability to compete with other giants in the market. E-commerce is interchange of goods/services in a large scale. seller and buyer move towards an internet in an on-line order system. Financial risk has been increased in recent years due to changes in payment mode and the forward rate where items are purchased. It is a search gate for customers as well as doing business in efficient manner.

E-commerce is a very good place for business. Nowadays selling the products through online plays an important role in the world. The E-Commerce strategy mainly focuses on E-Marketing and E-Commerce recommendation system. E-Marketing is adapting the marketing strategies and developing the improved one. The term E-Commerce recommendation system is used for product recommendation to the customer. The product searched by the customer through E-Commerce is automatically saved in to the database. From the database the vendor focuses on the customer who searched for a particular product and giving suggestions to the customer based on their search. E-commerce recommendation system is fully based on a) Customer profile b) History and c) Other customer's previous transactions.

Nowadays E-Commerce shows a tremendous growth in different areas. A successful E-Commerce website is one who understands and satisfies the requirements of both E-Commerce seller and buyers.

Types of E – Commerce
(i) B2B (Business to business transactions)
Right now companies exploit to do business with each other like manufacturers selling to distributors, wholesalers selling to retailers. Pricing is negotiable because it is fully based on quantity of orders. Thus, the result shows that 90% of E-Commerce transactions on the internet are B2B. e.g. wholesalers often sell their products to retailers.

(ii) B2C (Business to Consumer transactions)
It involves direct selling to consumer over the internet. Shopping cart software utilize the business selling people to the general public. e.g. retailers sell their product to consumer (i.e.) supermarket

(iii) C2C (Consumer to Consumer transactions)
The third party gradually increases a transaction bridge between one or more consumers involved. E.g online auctioning.

(iv) C2B (Consumer to Business transactions)
It is a reversible model of B2C. in this model a consumer offers goods to company then the company will pay for them.

(v) B2E (Business to Employee transactions)
In an intranet environment the company offer a product to their employees. in order to automate the employee related work process B2E is used.

(vi) Others
a.   G2G (Government to Government)
b.   G2E (Government to Employee)
c.   G2B (Government to Business)
d.   B2G (Business to Government)

ROLE OF WEB MINING IN E-COMMERCE

CUSTOMER ANALYSIS:
Maintaining the customer contacts, customer behaviour, customer reaction, customer satisfaction and profitability etc.

PRODUCT ANALYSIS:
Identify "Bottle Necks" (i.e.) production, delaying orders, examining production dynamics and comparing the product results given by the production department.

MARKETING ANALYSIS:
Analysing sales receipts, profitability, sales targets, profits margins, Times of orders, stock exchange and market identification.

PRICE ANALYSIS:
An amount of money paid to a particular product.

II. LITERATURE REVIEW

DevangJhaveri et.al [1] proposed a twitter sentiment analysis on e-commerce website in India. His work is to extract the data from the most famous social network is twitter. Flipkart and snapdeal is targeted and found out the sentiment behind each tweets various lexicon based approaches are applied and their accuracy is investigated.

Anil kumark.m et.al [2] analysed web user's opinion from phrases and emoticons. they classify customer opinion in to positive /negative. Nowadays user's personal opinions are expressed as word phrases, short words, emoticons etc. sentiment product lexicon is used to capture the opinionated phrases for each and every reviews. Whereas, emoticon lexicon is used to capture emoticon reviews.

Avanijadeja et.al [3] targets product based approach. The main objective is to provide a feature based summary for a large set of customer's feedback. In first step, they fetched products features from most famous e-commerce sites. In next step, they collected users review about the product in internet by extracting nouns, adverbs, adjective for all individual reviews, finally the collected data is analysed using sentiwordnet and found the semantic score for related sentences.

Neha verma, et.al [4] designed a semantic and neural based E-Commerce page ranking algorithm is implemented (SNEC). SNEC is a website ranking tool. This tool helps the customer to find the relevant websites on the top as well as business impact to compare their strengths and weakness. Using SNEC algorithm he found a new tool called website priority determination tool. The tool allows to compare almost six E-Commerce websites using drop down box and the search box is used to find the pattern of "what the customer actually searched for?". The website priority determination tool allows us to enter as many URL's of websites, then click comparison button, finally the tool will assign website priority to each of the customer based on the content priority module, time spent priority module and neural priority module.

Ravi kumar et.al [5] proposed a semantic based approach to extract five different product features based on the reviews using typed dependency relations.

K.M.anilkumar et.al[6]. analysed people's opinions, evaluations, sentiments, attitudes and emotions from kannada web documents. Semantic and machine learning approach is used. Finally, he compared both of these approaches and found the results. In Semantic method, Negator window, Baseline algorithm, sentence based method perform well. In Machine learning methods, Naïve Bayes method performs the best.

Kazuo Misue et.al [7]. developed an emotion weather maps. The map portrays spatial diversions of complex emotions from large set of people. He proposed a set of eight maps, each and every one represents one of eight emoticons (joy, trust, fear, surprise, sadness, disgust, anger, and anticipation). The data are extracted from social media. Social media is the place where large group of peoples are connected.

III. Service provided by E-Commerce website is:
1.   Sales
2.   Advertisement
3.   Introducing different kinds of products
4.   Giving best offers
5.   Comparison between other E-Commerce site and recommending for a right choice
6.   User reviews
7.   User friendly

Nowadays the growth of E-Commerce business is very high. So, keeping all these things in mind the source integrates web mining with E-Commerce website.

IV. ISSUES RELATED TO E-COMMERCE
1.   Taxation
2.   Customs
3.   Regulation
4.   Fraud
5.   Security
6.   Privacy
7.   Profitability

V. List of top 10 E-commerce sites in India [8]
We are living in an E-commerce era and many of you might have your own product .so, who are all the biggest E-Commerce players in the world? Let's have a look at the top 10 E-Commerce website in the world. This data is based on votes given by the viewers, which gives you the appropriate data of traffic to a website.

Sample example

S.No	E-COMMERCE SITES	AVERAGE SCORE	RANK
1	Amazon	4.75	1
2	Flipkart	4.78	2
3	Snapdeal	4.90	3
4	Paytm	5.01	4
5	Ebay	5.02	5
6	Jabong	5.11	6
7	Myntra	5.16	7
8	Shopclues	5.55	8
9	Pepperfry	5.81	9
10	Homeshop18	5.87	10

TABLE 1: Top 10 E-commerce websites

If three people have given rank to Amazon and Flipkart like 1,7,3 and 8,2,9. So, we will get the average for Amazon is (1+7+3)/3 = 3.6 and for Flipkart is (8+2+9)/3=6.3. In this case Amazon comes first and flipkart comes next so on.

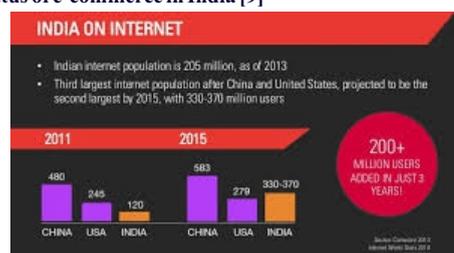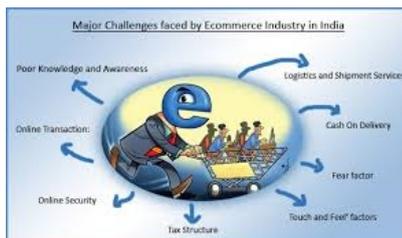### VI. Status of e-commerce in India [9]

Figure 2: growth of Internet users [10]

Today E-commerce is a backbone in Indian society and it has become an integral part of our daily life. The first e-commerce site in India was rediff.com. It was one of the most trafficked portals for both Indian and non-resident Indians. It provided a wealth of Indian-related business news a search engine, e-commerce and web solution services. The past 2 years have seen a rise in the number of companies enabling e-commerce technologies and the internet in India. Major Indian portal sites have also shifted towards e-commerce instead of depending on advertising revenues. There are websites providing any number of goods and services. Then there are those, which provide a specific product along with its allied services Multi-product e-commerce-These Indian E-commerce portals provide goods and services in a variety of categories. It creates a new opportunity for entrepreneurial start-ups. Ease of Internet access, Safe and secure payment modes coupled with aggressive marketing by E-Commerce Giants has revolutionized this segment. Rapid development in mobile technology has given way to Mobile Commerce with many E-Commerce companies shifting to App only model.

VII. CHALLENGES OF E-COMMERCE IN INDIA

Figure 3: challenges [10]

- E-Infrastructural Issues
- Touch and Feel
- Branding & Marketing
- Logistics & Supply Chain
- Declining Margins
- Tax related issues

VIII. FUTURE of E-Commerce in India [10]

Figure 4: Future and scope [11]

- Social Media: Majority of online buying decisions are made on Social Media.
- App only Approach
- Drone Delivery
- Artificial Intelligence
- Google's Buy Now Button

IX. Benefits of E-Commerce

The advantage of E-Commerce over traditional business is faster speed and lower expenses in completing customer transactions and orders. Because of these advantages the following industries in different fields such as retailing, banking, transportation, medical services, communication and education are establishing their business in the web [12].

- 100% business uptime
- 24/7 services
- Fast buying and selling procedure
- Cost efficiency
- Global access
- Quick response time.

X. Limitations of E-Commerce

- Customer fear personal information being used wrongly or theft. I.e. privacy issues.
- No guarantee about product quality
- Customer expectations unmet
- Rules and regulations vary
- Security and privacy i.e. vulnerability to fraud and other crimes
- Limitations of support services etc.

XI. OPINION ANALYSIS ABOUT FLIPKART

Flipkart is an e-commerce company founded in 2007 by Sachin BansalandBinny Bansal. The company is registered in Singapore, but has its headquarters in Bangalore, Karnataka. Flipkart has launched its own product range under the name Digi Flip with products including tablets, USBs, and laptop bags. Flipkart has wide range of connections. They worked for Amazone.com and left to create their new company incorporated in October 2007 as Flipkart Online services pvt ltd. The first product they sold was the book leaving Microsoft to change the world. Today it leads a 2nd rank position in India.

Sentiment is defined as customer's attitude, thoughts, and judgement prompted by feelings. A semantic approach is used to analyse the results. A small set of data's (i.e.) Oct, Nov and Dec 2015 has been collected through internet via mouthshut.com. The collected data is then analysed using Microsoft excels and produced the result. Data's shows that there is a service oriented issues in Flipkart but the result shows that most of the customers were either partially or fully satisfied. Some customers feel bad because of duplicate product issues, damage piece, poor customer service. A chart has been created to view the results. First chart shows the overall customer reviews in positive, negative and neutral way. Second chart shows the product analysis method for the collected data.
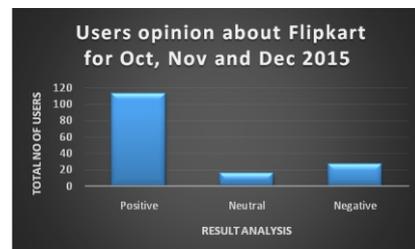
Figure 5: Overall customer's feedback about Flipkart

XII. CONCLUSION

The concept of e-commerce is all about using the Internet to do business better and faster. We have discussed that how web mining is integrated with E-Commerce. Each and every user leave a digital footprint in E-commerce site so based on the footprint their needs can be evaluated and validated using web mining concepts. In this paper I focussed on user's opinion about Flipkart in the year 2015 and I found Flipkart shows the positive impact results. Here I have done a sentiment analysis for Flipkart only with small set of data. In future I thought to do a sentiment analysis for all top 10 E-commerce sites in India with large data sets. Both semantic and machine learning approach will be implemented using some special tools like weka, R tool and compare the results obtained. Finally, the E-commerce must find out to Whom? What? How? and When? To be analysed.India is showing tremendous growth in this area. Future of E-Commerce is very innovative in India. If you do that, you will see results!

REFERENCES

1. Devangjhaveri, anushchaudhari, Lakshmi kurup, "Twitter sentiment analysis on e-commerce website in India", In proceedings of international journal of computer applications (0975 8887), IJCA, Vol 127 – No 18, Oct 2015.
2. Anil Kumar K.M, Suresha. "Analyzing Web user' Opinion from Phrases and Emoticons", IJCA Special Issue on "Computational Science - New Dimensions & Perspectives", NCCSE, 20112.
3. AvaniJadeja, IndrJeet Rajput. "Feature based sentiment analysis on customer feedback", In proceedings of journal of information, knowledge and research in computer engineering.
4. Neha Verma, Dheeraj Malhotra, Monica Malhotra, Jatindersingh, "E-Commerce website ranking using semantic web mining and neural computing", In proceedings of international conference on advanced computing technologies and applications, ICACTA, Elsevier, pp. 42-51, 2015.
5. Ravi kumar V and K. Raghuveer, "web user's opinion analysis for product features extraction and opinion summarization", In proceedings of International Journal of Web & Semantic Technology (IJWesT) Vol.3, No.4, October 2012.
6. K.M.Anilkumar, N.Rajasimha, Manovikasreddy, A.Raja Narayana and KewalNadgir, "Analysis of users sentiments from kannada web document", in proceedings of Elsevier ,Science direct, procedia computer science, 54,pp.247-256, 2015.
7. Kazuo Misue, Kiyohisa Taguchi, "Development of Emotion – Weather Maps", Elsevier, Science direct,procedia computer science 60,pp.292-301,2015.
8. www.Indiafreestuff.in
9. www.uniassignment.com
10. www.linkedin.com
11. www.problab.com
12. Istrate Mihai, "Web Mining in E-Commerce", university of Pitesti faculty of mathematics and informatics No.1, stadiomului, Romania.