TRANSACTIONAL PROCESSING SYSTEMS

Applicability Evaluation of Web Mining in Healthcare E-Commerce towards Business Success and a derived Cournot Model

P. Damodharan[1] · C. S. Ravichandran[2]

Abstract

Internet has become integral part of day-to-day business to almost everybody, this result in diversified interests of customers. So as to catering to this intrinsic need, any E-Commerce firm to survive must be of cutting edge and competitive edge. The providers should not only to stay abreast with technologies where the life cycle of a technology is at its bare minimum and further dwindling They should also entertain the customers through inventively fine tuning the delicate parameters of website. This involves evaluating the usage pattern and trails of the customer left as a log, deriving pattern from click stream etc. However, the cutting edge technology applied by the big healthcare E-Commerce industries like private Cloud utilization (John et al., *Optimization and Computing*, 2012), web content mining enables them to attract and retain innumerable number of customers even during peak hours. According to the research carried out in this paper, there are two distinct types of online business based on web content promoted towards buy, they are classified as exhaustive promote and partial promote. Typically exhaustive promote website perform even complex web mining operations for identifying and enticing the potential customers to buy various healthcare products based on various factors such as buying habits, interests etc. However for the partial promote in the observed cases, they are not even aware of the existence of such techniques. Based on the analysis performed on various renounced online websites, if 60% and above of the web content leads the customer to perform the 'buy, then it is exhaustive promote the rest is considered as partial promote. Moreover a huge gap is observed between Partial and exhaustive promote when it comes to the deployment of the web mining techniques. Consequently to understand the varying role of web mining in the online business successes, this paper models the web mining as a Game in Cournot Model. The results show that the model suits the economics behind the online businesses in both the cases and thus helps to identify or enhance the underlying web mining techniques towards business success.

Keywords Web mining · E-Commerce · Healthcare · Cournot model · Game theory · Web content mining

Introduction

Internet has been the prime attraction to all businesses this makes the Internet crowded with various services, most of them is hosted in World Wide Web (WWW). Moreover for

✉ P. Damodharan
  damumtp_cse@yahoo.com

  C. S. Ravichandran
  eniyanravi@gmail.com

1 Akshaya College of Engineering and Technology, Coimbatore, India

2 Sri Ramakrishna Engineering College, Coimbatore, India

the sake of simplicity, processing towards the client side is greatly reduced to mere click stream which is nothing but the user navigation pattern [1]. However to ensure the competitive edge the website owners has to analyse various aspects of the client without bothering them. Though it appears simple, the kind of effort and subtle methods needed to identify the potential customer are much more complicated and always been the subject of research due to the ever evolving nature of WWW. For instance, website visitors may visit the website for various reasons without being the customer, it requires not only the log analysis but also study their psychology, browsing habits etc. [2]. This gets even more complicated due to the growing nature of the Internet the website visitors nowadays use the dynamic IP address from the available pool. In such case, there is no way for the online businesses to clearly identify the genuine potential customers it therefore

suffers with considerable number of false detections [3]. In spite of various such difficulties the online businesses still thrives to identify and categorize the website visitors using self employed web mining algorithms. Inspite of all the difficulties, no business can ever be successful in WWW without putting an effective web mining algorithm in place. Therefore it requires through understanding about the working mechanism of web mining before investing time and cost on it.

The fundamental procedure remains the same even for the most successful web mining algorithms, they are discovery, classification, personalization and recommendation [2, 4]. Web discovery involves identifying target region whether local or global and then applying methods to search and collect relevant data involves feature selection [5]. Subsequently, Classification involves arranging data in a organized way for further analysis. Personalization is the result of classification involves listing the potential customers such as list of identified users from mere server log, log in, harvesting from WWW and making essential changes in websites [6]. However recommendation is the final outcome and it directs the website owner with necessary changes and upgrades to further the development of the industry [2]. Each and every steps in fundamental procedure might need a unique web mining technique, i.e. for recommendation apart from web content mining the web usage mining is also essential. For identification web usage mining may offer some useful insight [7]. Opinion mining may helps in better personalization [8]. However for a successful web mining various mining algorithm should be integrated. No matter how effective it has been implemented; it still suffers limitations such as false positives and false negatives [3].

As a result, web mining has been often upgraded and kept as the trade secret because the E-Commerce firm that holds the better algorithm holds the competitive edge. This exhibits the game theory nature, because a loss of a customer to one healthcare E-Commerce firm is a gain to another. However it has never failed to grab the attention of research community, it has been intensely researched and documented since the arrival of World Wide Web. However web mining evolves in such a way that, even after the exhaustive study and research, proposing a brand new web mining strategy and considering it as a superior strategy may be misleading. Because the competing online businesses still may have better strategy or have deployed the similar strategy in a better way, which is a proper game theory problem.

Game theory is the study of strategy behind conflict and cooperation between two are more competent participants. The common feature of game theory is the interdependence among the participants. That means the outcome for each participant depends on the strategies used by the others. Therefore game theory as a science of logical decision making has been applied to study the wide array of relational aspects involved in cases not limited to humans, animals, and

computers. Game theory therefore is widely used in economics and also in various other emerging concepts computer oriented field such as E-Commerce, E-Learning etc. It has also been used in web mining as well [9].

Each and every step in web mining should evolve every now and then to meet the needs of the customers. Consequently to have a proper mechanism in place thorough study and research is necessary before modelling and developing a complete package for web mining. The Following sections therefore perform the comprehensive study and research necessary for a successful web mining algorithm.

Background and literature review

Web mining can be categorized based on the kind of data it uses to mine as i) Web usage mining, ii) Web content mining and iii) Web structure mining.

Web usage mining

Web Usage Mining applies data mining techniques to search through usage logs from Web server in order to identify, understand the behaviour and needs of the clients [10]. It is a time consuming and space consuming and cost incurring techniques which involve pre-fetching and catching from even long-term backlog [11, 12]. Web Usage Mining uses data available in log such as Client machine's IP address, Access date and time, Request method (GET or POST), URL of the page accessed, Transfer protocol (HTTP 1.0, HTTP 1.1, Success of return code, Number of bytes transmitted as well as binaries [1, 5]. Web server access log, proxy server log, browser log, error log, refer log, user profile, registration data, cookies, user queries and bookmark data etc. [13].

Web usage mining can be further differentiated depending upon the facility with which the usage log was collected. They are i) Web Server usage mining- The user logs collected by the Web server are used for mining [3]. ii) Application Server usage mining- Commercial application servers are enriched with necessary features to build and support e-commerce applications with little effort. It can therefore log and store voluminous transactions performed over it, therefore usage mining performed can yield better result but can be time consuming [11, 12]. Application Level usage mining- Applications can collect user data through enticing them with interesting offers; new event etc. mining performed at application level is easy and effective due to the availability of the necessary data.

Once the potential customer has been identified it becomes extremely difficult to understand the behaviour and interests of the potential customer using Web usage mining. This requires mining on the Click stream generated by the web visitors [1]. Click stream as the number of successive clicks

performed helps to measure the number or type of hyperlink or web document elements accessed [31].

Web structure mining

Web structure mining is an effort to define and map the hyperlink structure, XML document structure etc. using ontology, graph theory etc. [6]. This mapping helps to interpret user clicks to understand user interest, behaviour, etc. i.e. elements necessary towards recommendation. For instance, Google as a leading search engine uses a web structure mining based ranking recommendation algorithm called as PageRank to populate the search result [14].

Web content mining

Web content mining is the process of discovering, extracting and organizing the essential data from the content or description of web documents available over World Wide Web. Web content mining requires automation for information retrieval [6]. Therefore any intelligent web agents such as Web Crawler, Meta data crawler is a web content mining technique [15]. This is the only technique that can even mine the web database content [16]. It can scale across various data such as unstructured data, semi-structured data and structured data. The result obtained helps to rank, organize or edit the web content [17].

Deployment types

Web mining can be further classified based on the way it is deployed to gather knowledge as i) Simple, ii) Composite, iii) hybrid and iv) adaptable. *Simple* web mining is a lightweight dynamic algorithm implemented to do a mining for a special occasion or for special purpose [18], this does not suits long term business needs but can serve as a cost, time and space effective technique to accomplish a specific task. On the contrary *composite* mining involves a framework and various features to perform long term mining [1]. It is not dynamic or scalable might also incur cost, consume time and occupy space. If it is not monitored it might result in loss of a potential client. *Hybrid* is an effort to integrate few algorithms to perform effective web mining [19]. If it is not done properly it might backfire and reduce the performance. *Adaptable* technique involves dynamic mechanism that allows switching between the various mining algorithms to adapt the circumstances [9, 20].

Other useful mining techniques

Web scraping, web harvesting or web data extraction nowadays is widely used to extract data from other websites using Hypertext Transfer Protocol [6]. Web scraping involves fetching and extracting from web. Fetching is the process of downloading and catching of one or more web pages. Fetching requires automated intelligent agents therefore web crawling is used in web scraping. Once the fetching is over the extraction process begins which is a mining process it involves parsing, searching, and listing of the data in required format.

Web scraping involves web content mining necessary for applications such as web indexing, data mining, online price change monitoring and price comparison, product review scraping (to watch the competition), gathering real estate listings, weather data monitoring, website change detection, research, tracking online presence and reputation etc. Leading companies such as Amazon AWS and Google provide web scraping tools to scrap from public data available over WWW.

Social media mining- encompasses process such as measure, model, and mining patterns and trends in huge raw data collected from social media networks [4]. Usually mining pattern includes social media usage, online behaviors, sharing of content, connections between individuals, online buying behavior, etc. [21]. Experts use a range of basic concepts such as data mining, machine learning social network analysis, network science, sociology, ethnography, optimization and mathematics, social network analysis, network science, sociology, ethnography, optimization and mathematics and statistics to develop mining algorithms suitable for investigating massive social media dataset.

Search Engine Optimization (SEO) - is a process of optimizing website suitable for mining performed by search engines [22]. Optimizing a website involve editing its content, HTML tags, associated coding, presenting site map to list the hyper links associated with the website etc. SEO is done in context of making it relevant for keyword search and perform better in indexing. More frequent a target website link appears in the top page of the search results, more hits and click stream it can generate towards the target website [23]. These visitors can then be converted into customers. SEO may target different kinds of search, including content search, image search, local search, video search, academic search, news search etc. [24]. Consequently Search engine optimization (SEO) is considered as an essential Internet marketing strategy for business success. Furthermore, SEO is viewed as the proper follow up of web mining process.

Related works

According to the literature study performed it is found that the term web mining is used more inclusively. This creates the possibility towards ignoring local aspects of web mining. Therefore to properly distinguish the difference based on the local and global aspect of web mining. We term the mining done in the local web server facility as website mining and the

comparative mining performed with more contender servers is termed as web mining. Web usage mining is the type of website mining which help to evaluate service delay to relevant customers failing to do so may result in loss of customer. However, performing it online during peak hours may impede some customer and result in loss of customer [10].

Moreover the essential component for web mining to be effectively applied by the small scale and middle scale E-Commerce firm is business success. The content placed in the website to promote business' varies from one E-Commerce firm to another, they are classified as i) Exhaustive promote and ii) Partial promote. If every click performed by the customer is expected to culminate in the 'buy' then it is *exhaustive promote*. The exhaustive promote websites like ebay, amazon are tailor made for hosting healthcare buyers. Whereas the *partial promote* websites such as education institute are not designed to be so. For instance, the website hosted by education institutions is destined to handle wide array of users such as faculty, students as well as prospective students. There are other factors that helps to distinguish them as well, for instance, education institution as a partial promote looks to attract prospective students as the customer. On the contrary, online shopping as an 'exhaustive promote' looks to attract the buyers as a customer.

In earlier days exhaustive promote websites used techniques such as questionnaire, quiz etc. to identify and categorize customer [25]. Due to the fast advancing nature of Internet, nowadays customers are not interested in wasting their time in performing such online surveys. If such surveys are posted before 'buy' they may move to other websites. Therefore in many cases opinion mining is conducted after buy. However customers offer their feedbacks as the review. This review is used to short list the products for the potential buyers [8]. Honestly presenting the reviews on website is practised by every online shopping websites to help their customers to make appropriate product choice. This practise helps them in retaining their customers. However if the producer is the seller, then it is collected secretly but the disgruntled buyers still may post their feedback as review in the third party review websites. The impact of such review is growing, paying less attention or no attention to those reviews will be a big blunder [25].

The case is entirely different for partial promote websites, where customers may have lot of questions in their mind and would like to seek answers. Therefore most of the partial promote websites present list of frequently asked questions (FAQ). They also allow the potential clients to ask their questions. Most successful partial promote websites even have the online chatting facility which is backed by well trained call centre professional to answer the question and convince them to join their institutes.

Moreover for the exhaustive promote websites having knowledge about the customer is essential, since there is no way to properly identify and understand the customer through log analysis. To achieve this, personal data harvest is performed with or without the knowledge of the customers. Personal data is usually harvested from mail accounts, gaming websites, social networks, purchasing from black market even hacking is performed to harvest personal data [26]. This creates a huge privacy concerns because the trend of harvesting customer data is on the rise. Simply having the knowledge about customer is merely not enough, because at the end of the business process, customer satisfaction is crucial to have a competitive edge.

Problem formulation

Web mining usage can be viewed as pure economical problem because intention behind applying any web mining in an E-Commerce industry is purely for the business success [27]. Many industries fail to meet the expectation even after investing considerable cost towards web mining due to the lack of technical foresight. Follow up is the term normally used by the higher level management personal associated with the E-Commerce firms to enhance the technical aspect of business promotion. This allows the technical employees to completely ignore one and adapt another web mining. This is a common problem faced by many E-commerce businesses; therefore evaluating the feasibility of business success using ignored web mining techniques is identified for the key problem.

However to interpret the problem as a Game the nature of both partial and exhaustive promote is defined based on their commonalities as follows

- The game between exhaustive promote or between partial promote involves, attracting and retaining the customer. However to achieve this, apart from the good business drivers such as uncompromising quality, competitive pricing understanding the market is also essential. Let us assume all the business are cost effective and offers good quality products. Now the game comes to applying effective web mining increase the demand over Internet.
- When a small scale partial promote or exhaustive promote firms are forced to compete with large scale firms over WWW to stay in the business, it definitely lack infrastructure in spite of its limited capabilities it can make use of effective web mining techniques to stay in the market.. This marks their capability to define their own independent output and derive demand only towards it.
- E-Commerce industries whether they are partial promote or exhaustive promote they have a target output. The output in terms of online shopping is selling stipulated number of products, In terms of educational institutions admitting the students to the available seats etc. Therefore they

employ strategies such as web mining, online advertisement etc., necessary to draw more Internet traffic towards their website.

- In both the cases a game played is symmetric in nature *i.e.* all contenders have equal opportunity to use strategy from available infinite set of strategies, and the payoff depends only on the strategies being played, not on the nature of the contender or simply player.

- Moreover both the cases players are rational that means they employ such strategies independently to maximize their profits based on their own expectation. They assume it is not conjecture i.e. their output decision will not have an effect on the decisions of its rivals nevertheless their strategy affects others outcome.

There is one Game theory that suits those implications that is Cournot competition [28, 29].

Modelling implications as Cournot competition

Assumptions

Cournot competition is an economic model where companies only compete on the amount of their own producible output, which they decide independently of each other and simultaneously. According to the concept of Cournot competition the following assumptions were made

- There is more than one exhaustive promote and all exhaustive promote accumulate a homogeneous product (sellable healthcare goods). Similarly there is more than one partial promote and all partial promote produce a homogeneous product (cost incurred).
- Whether partial promote or exhaustive promote they don't cooperate with each other;
- Both have market power i.e. each one's output decision affects the market price (acquired from customer);
- In both the cases contenting peers compete in quantities, and choose the strategies simultaneously;
- The contenting peers are economically rational and act strategically; usually seeking to maximize profit even if given their competitors' decisions (web mining knowledge).
- Price is expected to come down as the result of competence for total output, therefore Price is considered as a decreasing function of total output.
- All firms knows', the total number of firms in the market, and tries to excel the output decision of the others.
- Each firm has a cost function $c(s_i)$. The cost function involves incurred cost towards basic business elements such as goods in terms of online shopping, procuring tickets

from service providers in terms of online booking, paying to faculty and performing other expenditure to host the students in terms of educational institutions etc. mostly they are invariable. Cost also includes spending towards online marketing which involves web mining as a decisive element. Normally the cost function is considered to be known by the contenting peers. The cost functions may be the same or different among contenting peers.

- The number of participant is assumed to be fixed in both cases, but it encourages newcomers to enter the business hence it increases output and reduce the profits because the former monopolist will adjust their output to adopt the new situation [28].
- The market price is set at a level such that demand equals the total quantity produced by all contenting peers. Therefore each firm is allowed to take the quantity or target set by its competitors as it is readily available and evaluates its possible residual demand, and then behaves as a monopoly. Thus the *market* determines the *price* at which it is sold.

Postulation

Cournot Competition deals with a situation in which n companies independently decide their production levels of a given product. The price of the product is a linear function that depends negatively on the total output. *Price is defined as the cost and effort incurred towards web mining,* We model it by means of the following strategic game.

Assume a set $\{1, \ldots, n\}$ of players as either partial promote or exhaustive promote, where $n > 1$.

A strategic game for n players $(S_1, \ldots, S_n, p_1, \ldots, p_n)$, consists of a non-empty infinite set Si of strategies i.e. $S_1 \times \ldots \times S_n \rightarrow \mathbb{R}^+$ for each player i.

The payoff function for each player in both exhaustive and partial promote scenario is defined by

$$p_i(s) := s_i \left(a - b \sum_{j=1}^{n} s_j \right) - cs_i$$

For all $p_i(s)$: $a > c$, $b > 0$ and $c > 0$ where a - *set price*, b – *quantity* and c - *cost.*

The payoff function involves the market price of the product which is represented by the expression $a - b \sum_{j=1}^{n} s_j$, where $b > 0$.

While 'the strategy set involves the steps taken to attract and retain customers i.e. increasing the customer visits, it involves spending mostly towards SEO and web content mining, online advertisements, identifying and listing the customers, harvesting data from social networks etc. The cost for follow up involves opinion mining, Most cases c is constant and genuine, whereas a and b varies**.**

Table 1 Link Juice measure for Partial Promote and Exhaustive Promote websites

Leading Online Shopping Website	Top-notch Healthcare Products Website
Internal links: 3	Internal links: 95
No follow internal links: 0	No follow internal links: 0
External links: 292	External links: 33
No follow external links: 0	No follow external links: 0
Link Juice: 1.0%	Link Juice: 74.2%

Substantiating the setting

According to the Cournot Competition theorem the profit depends negatively on the total output, $\sum_{j=1}^{n} s_j$. Additionally, $\mathbf{cs_i}$ is the production cost corresponding to the production level $\mathbf{s_i}$. we assume that the production cost function in partial promote are the same for all competing peers in its category and the production cost function in exhaustive promote are the same for all competing peers in its category.

When it comes to $\mathbf{s_i}$ versus $\mathbf{s_j}$ the objective is to keep the si maximum and sj minimal. Since they are the parameters that indicates the demand vs market, These parameters are introduced in such a way to consider the effects of web mining in increasing the payoff, because better payoff means better profit. Quantity can usually be increased based on the increased demand obtained using web mining. Therefore the online business with better web mining can have better demand thus results in increased quantity without having an impact on the price. Thus the player holds the fair share in the market.

However if the exhaustive promote or partial promote fails to retain the quantity then $\mathbf{a \le c}$, then the payoffs would be always negative or zero this is often the result of lack using adequate web mining in increasing the demand. Since the objective of any online business is success.

Therefore $\mathbf{p_i(s) = (a-c)s_i - b\ s_i\sum_{j=1}^{n} s_j}$ must be positive to ensure the business success. This makes the assumption a > c as suitable only for the organization with web mining capabilities.

Table 2 Observed SEO violation in websites of Top-notch Healthcare Products Website

SEO violations of partial promote website

Doc type: Not valid We couldn't find valid doc type declaration. Check if it's set correctly.

Page weight: 57.38 kb Page weight should be as small as possible. Try to optimize HTML code to get better score. Page weight should be as small as possible. Try to optimize HTML code to get better score.

WWW redirection (301): no For search bots website addresses with www and without it are considered as different pages. Adding redirection help you avoid double content penalty.

Keywords (0 phrases, 0 characters) Keywords metatag is empty

Description (0 characters) There are no description set for this page

Wrong heading level. Max allowed level is 1

Experimental results and discussion

Differentiating the partial promote from exhaustive promote help in categorizing the websitethus helps in further our understanding on the nature of the environment and the research necessity.

Distinguishing partial promote from exhaustive promote

Online based industries such as Online Shopping, Online Booking often brag about their online performance in converting the hits to buy ratio as 80%, 90% etc. Therefore, even for the newcomer, in exhaustive promote industry there is a well pronounced ratio for the business success i.e. 60% [30]. Whereas there are industries they are not even aware of this ratio. Rather not prepared to provide the data for research because of the fear of losing the sensitive data. However there is a way to have a perspective on that through using web content mining. Web content mining is emerging as a big business so the competitive tools such as CAT (Content Analysis Tool), site-analyzer, seowebpageanalyzer, etc. are readily available. Using 'seowebpageanalyzer' we performed the analysis on many firms apart from fully online driven industries such as Educational institutions, Tourism Industries etc. The result is as follows (Tables 1 and 2).

Link Juice is the ratio between internal and external links. It shouldn't be lower than 50% otherwise the website might be considered as portal to other websites. However this helps to distinguish between fully online driven business and the partially online driven business. Arriving at a conclusion based on Link Juice violation can be disastrous because apart from Link Juice the Top-notch Healthcare Products Website impedes the web mining through violating almost all other aspects necessary for search engine optimization (SEO).

These violations were spotted in almost all observed Top-notch Healthcare Products Website in south India. These violations imply that their websites are not properly configured for web content mining. Therefore expecting to harvest healthcare product over online can be futile. This also indicates the lack of utilization of readily available cost effective cutting edge technology. Therefore the observations serve as a measure for the huge gap between availability and technology usage.

Fig. 1 Normal distribution graph for exhaustive promote website using Link Juice ratio
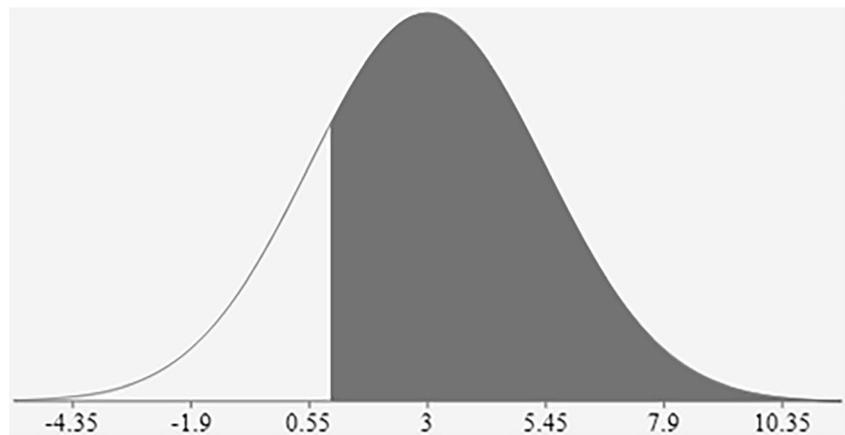
However Tourism industries such as Hotels, Software industries, etc. have configured websites better to suit SEO needs. This helps them to attract the customers' online even though their sole resource is not website. However the analysis performed in top healthcare products website, in spite of their pioneering nature they haven't configured the website to suit their business needs. The discrepancies observed are as follows (Figs. 1, 2 and 3).

The above graph is obtained through plotting the Link Juice values obtained for Online shopping in normal distribution for further analysis. The Normal Distribution uses mean and standard deviation to distribute the data. It always has mean = median = mode, therefore it is always symmetric about the centre. It thus distributes data in such a way that 50% of values less than the mean and 50% greater than the mean. Therefore it helps to measure how spread out the values area in the bigger picture. In this case, Normal Distribution indicates that all the observed online driven businesses violate the Link Juice standard in a huge way. They have presented most of their links externally thus affecting the Link Juice ratio After scrutinizing the results it became evident that this violation is not the result of lack of knowledge towards adopting to the web mining

standards. Instead it is done intentionally to exploit the web mining to attract more customers in context of business success. This observation therefore serves as a measure for superior knowledge and expertise the 'exhaustive promote' holds.

The better performance of Top-notch Healthcare Products Website in Link Juice is due to the presence of internal links which is necessary for various functionalities of the institute.

However there where totally nine violations associated with this web content mining tool. Among those the Online shopping websites apart from violating Link Juice there is no other violation. This show they have configured their websites suitable for web content mining. This indicates that the Online driven business have mastered the web content mining in such a way to exploit it to maximize their revenue. Contrarily the websites of partial promoteare not even aware of the existence of web content mining. As a result, their websites vastly violates the basic web mining requirement for successful online business. The obtained result for more than 100 healthcare websites is presented as follow.

Based on the analysis performed on various renounced online websites, if 60% and above of the web content leads the customer to perform the 'buy then it is

Fig. 2 Normal distribution graph for Partial promote websites using Link Juice ratio
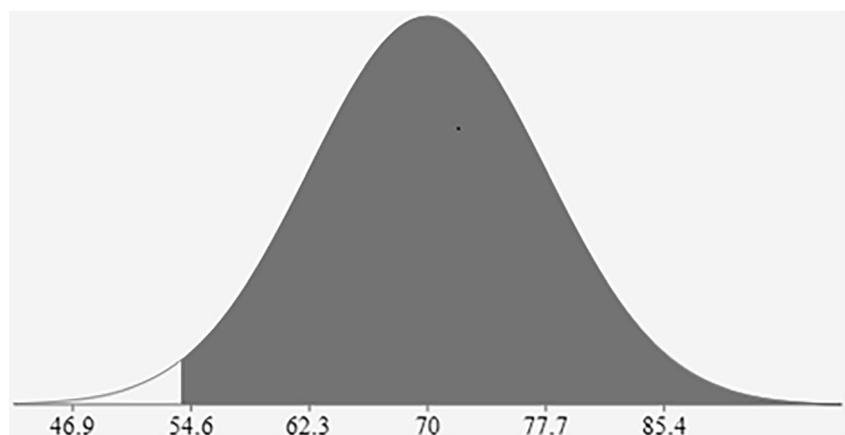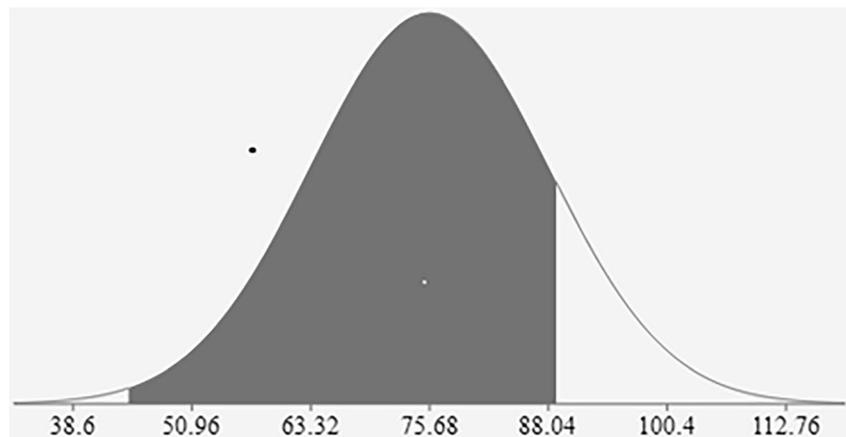
Fig. 3 Normal distribution graph for partial promote websites using violation except Link Juice ratio

exhaustive promote. The rest is classified as partial promote. In other words if any access to the contents some how can result in the buy then it is exhaustive promote. If not, it is not an exhaustive promote. This scenario marks the huge gap between the Partial and exhaustive promote websites in using web mining techniques. Moreover if only less than the 20% of the content is focused towards online business productivity then it is poor partial promote. In case of partial promote the dependence of website to obtain the client is minimal because to admit prospective student universities advertise in paper, social networks, issue track, newspaper ads, etc. However partial promote are no competition to the exhaustive promote, still the difference between them is too huge. It should not have existed after all the advancement happening in web mining. Existence of poor partial promote implicate the requirement for more research towards enhancing the applicability.

Comparison analyses

The intention behind this section is to analyze the effectiveness behind various web mining techniques. However to analyze the effectiveness of web mining from technical perspective alone is simply not enough. Because, for the online businesses, if the web mining result is not interpreted as business success, then using them is simply waste of money. Therefore

a survey is conducted with various partial promote and exhaustive promote firms. The partial promote under study is healthcare website. The details of which is not presented not only because they neglect to disclose their identities but also due to the nature of the product listing them is deemed to be unnecessary. The sample space involves 14 online driven businesses exhaustive promote and 25healthcare product website as partial promote. Other have either neglected the survey are not have provided sufficient data. This number is converted into percentage and the result is tabulated as follows.

The inference from the Table 3 is as follows. Web mining as promotional strategy performed only on the occasion of new product added in website. The healthcare website company under study are not top-notch and they don't receive any aid. Those company are private owned and behave as a perfect business organization. Most popular strategy adopted towards online promotion is placing ads on various blogs and websites. Moreover among various healthcare products offered their worst performing product is chosen for analysis for the sake of simplicity.

Likelihood chance in case of partial promote is discerned as the measure of intention to rely more on the online marketing such as using web mining to enhance the WWW visibility. *Effectiveness* is the percentage of money spent towards web mining comparing to other promotional strategies. *Business Success ratio* is simply the ratio between target set and target achieved on outcome.

Table 3 Partial promote web mining implications

Web Mining Type	No. of partial promote	Effectiveness	Likelihood Chance	Business Success ratio
Web Usage Mining	8	7%	55%	78%
Web content mining	2	4%	68%	62%
Hybrid mining	7	9%	45%	88%
Adaptive or Perfect mining				
No web mining	8		81%	53%

Table 4 Exhaustive promote web mining implication

Web Mining Type	No. of exhaustive promote	Effectiveness	Likelihood chance	Business Success ratio
Web Usage Mining				
Web content mining	5	72%	4%	88%
Hybrid mining	2	80%	10%	86%
Adaptive or Perfect mining	7	87%	1%	92%
No web mining				

According to our pilot study performed among various leading healthcare company. Few company which employed opinion mining to enhance their posting performed better in product added which is hybrid in nature. *Hybrid mining* in case of partial promote is not the presence of proper custom build framework. Rather a kind of collective mining effort made toward correlating the obtained web related data into observable pattern for business analysis. For instance after making massive online advertisement effort, observing the web hits and if it meets their expectation placing an attractive event in websites etc. (Table 4).

Perfect web mining is mostly evident with exhaustive promote firms, it happens only when the level of expertise is high. The exhaustive promote firms under study genuinely uses necessary technique for web and website mining. For web mining the minimal requirement for exhaustive firm is performing SEO enhancing page rating and indexing. *Likelihood chance* in terms of exhaustive promote is their intention to change or adopt to other online web mining techniques. Though the business success is not 100% as expected in exhaustive promote, they don't attribute that loss in percentage to their web mining techniques they have adopted.

Hybrid mining is the case when the firm looks to employ static traffic exploration, user categorization, feedback follow up techniques, mostly those are custom designed stays the way it is build until the next business decision is made to upgrade it. However adaptive mining is a dynamic strategy which involves purchasing the emerging techniques immediately as they arrive, training the employs to manipulate it to the best use.

According to the conducted comparison analysis each business has a target output to achieve, their price fixing is based upon the cost they pay towards their distributors or cost incur for procuring a service and not based on the contenders price. Target quantity is achieved through the independent application of web mining techniques. This therefore suits the Cournot Competition and thus eliminates the possibility for other economic games [28]. Therefore the performed pilot study validates our intention to involve the Cournot economic game theory in developing adaptive framework for web mining.

Conclusions and future work

Wide range of web mining algorithms are studied and categorized. The application of web mining for business success has become a new reality. However many scintillating web mining algorithms fails to meet the baseline because of the lack of understanding of the ground realities. However from over study and analysis we observed that the content placed in the website to promote business varies from one E-Commerce firm to another, based on that any business which employs Internet marketing can be classified as Exhaustive promote or Partial promote. If every click performed by the customer is expected to culminate in the 'buy' then it is exhaustive promote other websites are classified as partial promote. Subsequently Web Mining usage has been viewed as pure economical problem because intention behind applying any web mining in healthcare E-Commerce industry is purely for the business success. However to address this economical aspect behind the web mining modeling it using a economic game theory for various aspect behind the partial promote and exhaustive promote became necessary. According the observations and analysis made the core elements to choose a model is identified as; websites in both the cases only compete on the amount of their own producible output, which they decide independently of each other and simultaneously. The economic game theory which suits this requirement is Cournot competition. Therefore Cournot game theory is used to model the business oriented nature of the web mining. From our experimental analysis and comprehensive study it became evident that the exhaustive promote has put the web mining to better use in drawing the customers and constantly upgrade the website to suit the emerging trends. Contrarily there is a class of partial promote websites which are not even aware of the existence of web mining they are termed as poor partial promote.

Intention behind modeling partial promote and exhaustive promote using Cournot Game theory is to have a dynamic framework where all aspects of the web mining as a game is considered especially in handling business uncertainties. Therefore our future work involves developing an integrated adaptable techno-economical web mining which performs mining operations from the perspective of business success.

Compliance with ethical standards

Conflict of Interest   This paper has not communicated anywhere till this moment, now only it is communicated to your esteemed journal for the publication with the knowledge of all co-authors.

Ethical approval   This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Sujatha, V., and Punithavalli, Improved User Navigation Pattern Prediction Technique from Web Log Data, International Conference on Communication Technology and System Design 2011. Procedia Engineering:92–99, 2012.
2. Alphy, A., and Prabakaran, S., A Dynamic Recommender System for Improved Web Usage Mining and CRM Using Swarm Intelligence. Sci. World J., 2015.
3. Lopes, P., and Roy, B., Dynamic Recommendation System Using Web Usage Mining for E-commerce Users, International Conference on Advanced Computing Technologies and Applications (ICACTA), 60–69, 2015.
4. Ting, I.-H., Chang, P. S., and Wang, S.-L., Understanding Microblog Users for Social Recommendation Based on Social Networks Analysis. J. Univ. Comput. Sci. 18(4): 554–576, 2012.
5. Nazar, N. B., and Senthilkumar, R., An online approach for feature selection for classification in big data. Turk. J. Electr. Eng. Comput. Sci.:163–171, 2017.
6. Gawali, R. R., and Mundhe, S. D., Web Mining techniques, process and applications in Ecommerce, Proceedings of National Conference on Emerging Trends: Innovations and Challenges in IT, 19–20, 2013.
7. Jojo, J., Sugana, N., User profile creation based on navigation pattern for modeling user behaviour with personalised search, IEEE Xplore International Conference on Current Trends in Engineering and Technology (ICCTET), 2013.
8. Somprasertsri, G., and Lalitrojwong, P., Mining Feature-Opinion in Online Customer Reviews for Opinion Summarization. J. Univ. Comput. Sci. 16(6):938–955, 2010.
9. Wang, Y., Integration of Data Mining with Game Theory. In: Wang K., Kovacs G.L., Wozny M., Fang M. (eds) Knowledge Enterprise: Intelligent Strategies in Product Design, Manufacturing, and Management. IFIP International Federation for Information Processing, vol. 207, 2006.
10. Naveena Devi, B., Rama Devi, Y., Padmaja Rani, B., Rajeshwar, R., Design and Implementation of Web Usage Mining Intelligent System in the Field of e-commerce, International Conference on Communication Technology and System Design, pp. 20–27, 2011.
11. Trabold, D., Horváth, T., Mining Data Streams with Dynamic Confidence Intervals. In Big Data Analytics and Knowledge Discovery DaWaK 2016, Lecture Notes in Computer Science, vol 9829. 2016a.
12. Trabold, D., and Horváth, T., Mining Data Streams with Dynamic Confidence Intervals, International Conference on Big Data Analytics and Knowledge Discovery DaWaK, 99–113, 2016b.
13. Niu, N., Stroulia, E., and El-Ramly, M., Understanding Web Usage for Dynamic Web-Site Adaptation: A Case Study, Proceedings of the Fourth International Workshop on Web Site Evolution (WSE'02), 2002.
14. Lina, Y., Shib, X., and Yimin Weib, C., On computing PageRank via lumping the Google matrix. J. Comput. Appl. Math. 224(2): 702–708, 2009.
15. Salah, T., and Tiun, S., Focused crawling of online business Web pages using latent semantic indexing approach. ARPN Journal of Engineering and Applied Sciences 11(15):9229–9234, 2016.
16. Harry, T., Achsan, Y., and Wibowo, W. C., A Fast Distributed Focused-web Crawling. Procedia Engineering 69:492–499, 2014.
17. Choi, J., Lee, G., and Moon, J., Web Context Classification Based on Information Quality Factors. J. Univ. Comput. Sci. 16(16): 2232–2251, 2010.
18. Sabucedo, L. Á., and Rifón, L. A., Locating and Crawling e-Government Services A Light-weight Semantic Approach. J. Univ. Comput. Sci. 16(8):1117–1137, 2010.
19. Stathis, P., Kavallieratou, E., and Papamarkos, N., An Evaluation Technique for Binarization Algorithms. J. Univ. Comput. Sci. 14(18):3011–3030, 2008.
20. Abraham, A., Business Intelligence from Web Usage Mining. J. Inf. Knowl. Manag. 02(04), 2003.
21. UlviŞimşek, M., Özdemir, S., Analysis of the relation between Turkish twitter messages and stock market index, IEEE Xplore 6th International Conference on Application of Information and Communication Technologies (AICT), 17–19, 2012.
22. Cui, M., and Hu, S., Search Engine Optimization Research for Website Promotion. IEEE International Conference on Information Technology, Computer Engineering and Management Sciences (ICM), 2011.
23. Attardi, G., Di Marco, S., and Salvi, D., Categorisation by Content. J. Univ. Comput. Sci. 4(9):719–736, 1998.
24. Safran, C., and Kappe, F., Success Factors in a Weblog Community. J. Univ. Comput. Sci. 14(4):546–556, 2008.
25. Panniello, U., Ardito, L., and Petruzzelli, A. M., Does the Users' Tendency to Seek Information Affect Recommender Systems' Performance? J. Univ. Comput. Sci. 23(2):187–207, 2017.
26. Velásquez, J. D., Web mining and privacy concerns: Some important legal issues to be consider before applying any data and information extraction technique in web-based environments. Elsevier Expert Systems with Applications 40(13):5228–5239, 2013.
27. Bravo, C., Figueroa, N., and Weber, R., Modeling pricing strategies using game theory and support vector machines. In Petra Pertner, editor, Advances in Data Mining. Applications and Theoretical Aspects. Lect. Notes Comput. Sci 6171:323–337, 2010.
28. Abolhassani, M., Bateni, M. H., Hajiaghayi, M. T., Mahini, H., and Sawant, A., Network Cournot Competition, arXiv:1405.1794 [cs.GT], 2014.
29. Bimpikis, K., Ehsani, S., Ilkiliç, R., Modeling Pricing Strategies Using Game Theory and Support Vector Machines, Proceedings of the fifteenth ACM conference on Economics and computation EC'14, 733–733, 2014.
30. Barry, C., Markey, R., Almquist, E., and Brahm, C., Putting social media to work. Bain Brief, 2011.
31. John, J. M., Venifa Mini, G., Arun, E., User Profile Tracking by Web Usage Mining in Cloud Computing, International Conference on Modelling, Optimization and Computing (ICMOC 2012), 2012.