

Conceptual models for integration of Data Mining and SEO to Improve E-Commerce Performance

Tri Andjarwati^{1*}, Agus Hermanto², Vieqi Rakhma Wulan³

^{1,2}Universitas 17 Agustus 1945 Surabaya

³University of PGRI Adi Buana Surabaya

*Corresponding author: triandjarwati@untag-sby.ac.id

Abstract

Since the initiation of online trading in Indonesia, the public trust getting higher in using e-business applications have a positive effect, in the form of more open markets and raised competition to gain consumer trust. This requires a concern of company management settling their product sales strategies to continue going up, mainly online trading. At some point, the technique used for online marketing is SEO (search engine optimizer), but in its use, SEO still has some drawbacks. Therefore, in this article it is proposed to carry out data mining integration with SEO, so that the results can be used to improve e-business performance support that can be used as an alternative decision-making by management.

Keywords: SEO, data mining, e-commerce, management, marketing

Introduction

The development of information technology infrastructure that is getting better brings of the impact of changes in all fields, especially business and trade, including the behaviour of people as users of technology. Some consumer behaviors that drive the growth of e-commerce, including: consumers are starting to like online shopping, both for purchasing goods or services, especially domestic consumers who also begin to use mobile wallets, which continue to grow from year to year as seen in the Fig. 1 (JakPat 2015). The use of this mobile wallet was triggered by the growth of online transportation which helped popularize deposit payment accounts by offering cheaper rates (CK Finanzpro GmbH 2018). The behaviour's consumers like the same day shipping service (Maketeers.com 2017). Based on data released by comScore, the average marketplace growth is close to 100 percent, while e-commerce growth figures reach 289 % (IdnTimes 2017). With a fairly high growth rate, the Indonesian e-commerce market is very competitive, so management requires alternative methods of decision making to increase sales, especially transactions through e-commerce. From the last seven years, 2007 to 2015, e-money's multiple annual growth rate (CAGR) increased dramatically by 77%. It can be seen that in recent years the use of non-cash payment instruments has increasingly been chosen by the public to make payments for goods and services transactions (Bank Indonesia 2016).

So far in the world of e-commerce, website owners who want to do branding, many use SEO (Search Engine Optimizer), namely: marketing methods that focus on increasing visibility in the results of organic search engines (non-paid) (Fishkin dan Staff 2015). The use of SEO in e-commerce applications has several advantages and disadvantages (IDProgrammer 2017). The advantages of being owned by SEO, including: increasing the increase in visitor traffic on a website, the cost of using SEO is quite affordable, SEO supporting software is available in sufficient variants, monitoring visitor traffic is easy to do and the website becomes easily known. While the lack of SEO, including: SEO optimization process spends a lot of time has a dependence on search engines and high competition because many are using similar techniques.

By considering the advantages and disadvantages of SEO, in this article, we suggest to use data mining as an additional tool besides using SEO, as a helping method to increase sales. The results to be achieved by combining the two tools are the achievement of management performance for online sales effectively, efficiently and practically.

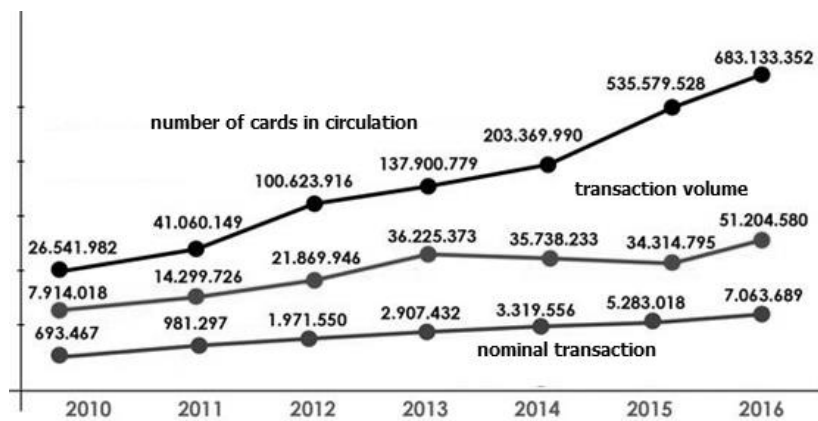


Figure 1. The Use of E-Money in Indonesia (JakPat 2015)

Literature Review

In this section we collect information that is relevant to the topic or problem that is the object of research so that we can utilize all the information and thoughts that are relevant to the research we have done.

Marketing

Marketing is the process of identification and all activities to meet human and social needs. Marketing in a nutshell, can be defined as follows, to bring together needs that can provide benefits (Kotler dan Keller 2012). Based on these definitions, it can be said that the purpose of marketing is to achieve profits. Furthermore, the marketing associated with other entities (or in other terms the relation of the five basic markets and their connection flows), can be seen in Fig. 2 (Kotler dan Keller 2012).

At this time, fundamentally there has been no change from these relations, even though the world has entered a more massive digital industry era, but has an impact on changes in marketing behavior, opportunities and challenges, which include: information technology networks, globalization, deregulation, privatization, competition intensive, industrial convergence, retail transformation, disinter-mediation, consumer buying power, consumer information, consumer participation, and consumer rejection.

Facing various changes in the disruptive era, more innovative ways are needed, especially in designing business marketing (Calabretta dan Kleinsmann 2017), therefore, as a primary function in business innovation, the development of marketing design has also provided support for companies to increase the speed and complexity of technology, including cultural changes in the economic and social fields.

Marketing management has also evolved from a focus on methods of exchanging goods into core business processes that create and retain customers and shareholder value through SD logic (Vargo dan Lusch 2011). Marketers also face customer involvement that is far reaching in innovation in the era of digital transformation, which leads to several opportunities for synergy between marketing and design practices. One such opportunity is the increasing number of markets and customer data (data mining) available to marketers, which together with a widening gap between the ability of marketers to understand and use this data. Even for now, marketing is strongly influenced by digital integration and non-digital functionality, because it has an impact on most aspects of the marketing mix (Lamberton dan Stephen 2016).

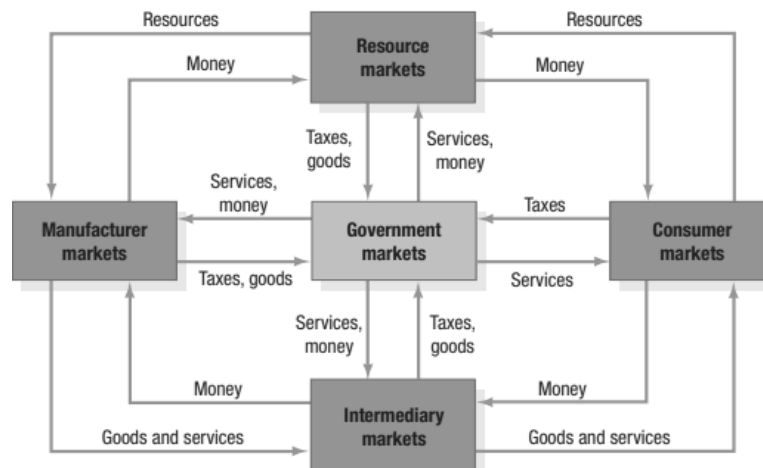


Fig. 2 The Relationship structure of five basic markets and connection flows (Kotler & Keller, 2012)

Based on the consideration of digital aspects and non-digital functionality, the use of data mining is an alternative in determining marketing planning. Data mining is the process of finding patterns in large data sets that involve interacting methods between machine learning, statistics, and database systems. Data mining has the ultimate goal of extracting information (with intelligent methods) from a collection of data and transforming information into structures that can be understood for further use (Hastie, Tibshirani dan Friedman 2009).

Data Mining

The actual use of data mining is for semi-automatic or automatic analysis of a large amount of data in extracting previously unknown patterns of interest such as groups of data records (cluster analysis), unusual records (anomaly detection), dependencies (mining association rules), and mining sequential patterns (Kriegel, Kröger dan Zimek 2012). This usually involves the use of database techniques such as spatial indices. These patterns can then be seen as a kind of summary of input data, and can be used in further analysis or, for example, in machine learning and predictive analysis. For example, the data mining step might identify several groups in the data, which can then be used to obtain more accurate prediction results by a decision support system (Witten, Frank dan Hall 2011). Data collection, data preparation, and results interpretation and reporting are part of the data development step, and also part of the entire KDD process as an additional step (Fayyad, Piatetsky-Shapiro and Smyth 2008). The process of implementing core mining can be explained as follows (WideSkills.Com 2018):

1. Determine the database object

Determination of database objects is useful for understanding how to extract according to the format of data from the DBMS used.

2. Selection of processing techniques

There are various kinds of techniques, which can be selected according to needs. Some techniques that are often used include: autonomous knowledge miner, data driven miner, miner driver query and miner interactive data.

Data mining is effectively used primarily by businesses that focus on customer satisfaction services; this is because such a business model will always try to create new techniques and strategies that increase customer satisfaction. Over time, the retail industry has made the application of several data mining techniques that are used to identify customer behavior, purchasing patterns and trends and improve supply chain management. If these things are implemented correctly, data mining techniques can provide support to retailers to increase customer retention and satisfaction, thereby increasing profitability (WideSkills.Com 2018). As an illustration of the data mining process, can be seen in Fig. 3 below:

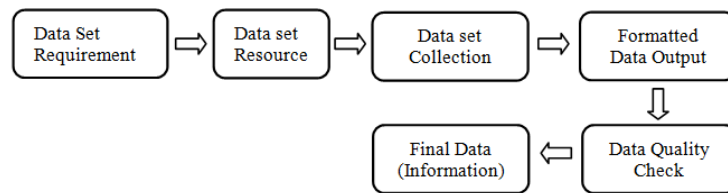


Fig. 3 The Stages of Data Mining

The ability of adaptive and flexible data mining will be very useful because processing methods can be obtained from business analytics, artificial intelligence, machine learning and statistics. The use of data mining can be applied in retail businesses (WideSkills.Com 2018), including:

1. Setting product prices

Retail business owners can satisfy customers in various ways. You can by introducing lower prices or offering coupons, coming up with new stores in prime locations, making special price offers for quality goods and so on. To attract new customers and make them regular violators, retail business owners can set product prices correctly. Product pricing must be done carefully, so that it can provide benefits to both retailers and customers.

Determining optimal prices for all products in the store is a challenging task. There are several factors that influence this, including consumer demand, price demand interaction, etc. Another aspect to consider when modifying the price of a product is that an increase in price can lead to lower sales and the adoption of customers for alternative goods. In this case, the application of data mining can help the process of identifying the occurrence of demand by consumers for products and also to gain an understanding of changes in prices of certain products affecting sales of other products.

2. Order Trend

By analyzing the facts of past purchases, retailers can gain an understanding of customer buying behavior that is very helpful in a variety of ways. For example, if there are purchases of two or three shared items made by a group of people, retailers can store and place these items in the closest location. This will ensure that customers who want to buy these items will not be left behind, so it will have an impact on increasing sales. Another case is that if retailers gain an understanding that more customers shop on weekends, actions can be taken to store enough inventory (inventory management) and announce attractive discounts at this time, this is an action to attract more customers.

3. Customer Segmentation

If retailers can create customer segments using different factors, it will be useful to create effective marketing strategies. Customers who respond to certain product launches can be put in the same group; customer grouping can also be done based on buying the same type of product or those who shop on weekends; customers who respond to new promotions or discounts can be grouped together. This information can be useful for identifying general group reactions to certain advertisements or promotions. Therefore, retailers can maintain or increase customer loyalty when they introduce things that might attract a specific group.

4. Effective advertising

Data mining techniques can be used to identify the effectiveness of certain promotions that have been carried out in many media or based on geographic location. For example, data mining can be applied to examine groups of customers who have responded positively to a promotion, the level of effectiveness of promotions in the form of budgets and the benefits obtained, the most successful media channels used for various campaigns in the past time period and beyond. By analyzing these kinds of information, retailers can design and implement promotions and advertise more effectively and profitably.

Search Engine Optimizer

The use of SEO (Search Engine Optimization) in e-commerce began around 1997, although actually it has been running since the creation of search engines in the 1990s, which was

pioneered by Erick Ward in 1994 known as Link Moses (Baker 2003). The working principle or Ward's ranking algorithm is based on two aspects, namely: studying human behavior in search (relevance) and connection. The working principle was then used as the basis of the search engine algorithm conducted by Google since 1998 (Shelley 2016).

The development of SEO from year to year until now, SEO has four pillars, namely: Links, Content, Mobile and Local. The first pillar, that is Links, has a big role in SEO, because it has the task of finding and building relevant links that have the principle of mutual benefit. The second pillar, namely Content, has a role to attract visitors to a website. The third pillar, Mobile, has the role of testing website components to be compatible with mobile devices and is now developing into a mobile search strategy. The fourth pillar, Local, serves to fulfill business needs by attracting visitors to visit certain websites, not just ordinary maps, but also includes: application, social and so on (Shelley 2016).

In 2015, Andrey Lipattsev, Head of Google Search Engine Algorithm Development introduced three web ranking factors, namely: Links, Content and new algorithms in search engines, named RankBrain (Sullivan 2017), Meanwhile Microsoft has started developing a web learning algorithm with machine learning under the name RankNet in 2005 on their search engine called Bing (Sullivan 2017). But between the two (Google and Bing) which cannot be compared which is better, although in popularity Google is more widely used.

Another important aspect for SEO is the data structure, because the data structure can provide a great opportunity for SEO to communicate important information with search engines, increase the visibility of content, and reach the target audience. Using a good data structure will provide crawler crawlers from search engines to extract and understand specific information related to the content, in this case, product type, aggregate ranking, available offers, and product reviews, thus increasing the possibility of the ability of crawlers to understand web content with increased accuracy (Digital Marketing Depot 2006).

Data Structure can make SEO have a competitive advantage in the SERP (Search Engine Result Page) which includes increased SERP space and a higher click through rate. In today's SEO landscape, search engines must fit billions of questions into trillions of web pages every day. SEO, on the other hand, is always looking for clear ways to communicate with search engines to improve the performance of their content. Structured data appears at this common priority junction. Enriched results may be the most visible advantage of structured data. These are search results that go beyond the standard blue link. Enriched results make it possible to display prominent ratings, product details, company contacts, reviews, job openings, and more. Structured data can allow improved results, including: snippets enriched in search results, such as images and product information, knowledge graph boxes, rich cards for mobile devices, carousel amps results on mobile devices (Brickley, et al. 2011).

Conceptual Model

This paper proposes innovation in marketing management by combining the results of processing from data mining with SEO techniques, with the integration model as shown in Fig. 4 below:

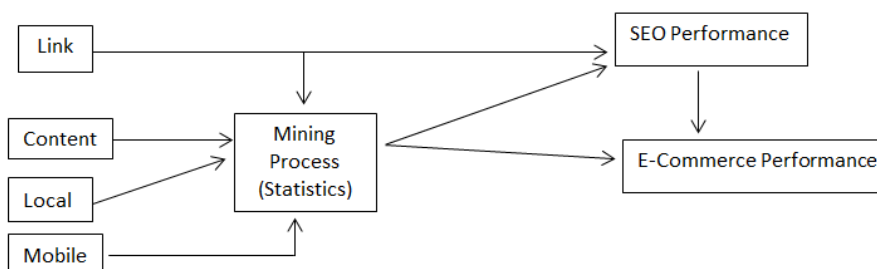


Fig.4. SEO and Data Mining Process Integration Model for Marketing Performance Improvement

In Figure 4, it can be explained that the source of data mining is sales data and web operational data in the past, which are then processed by the following process:

- a. Determination of data collection requirements, adjusted for the values for parameters that will be forwarded to SEO, namely Link, Content, Cellular and Local.
- b. Format settings are done by passing the stages starting from preprocessing by performing data cleaning actions, namely cleaning historical data from null and irrelevant data, followed by identifying users in the data history, this is because the same user can use a different host. At this stage Session identification is also carried out, so a single session is created for each user.
- c. Furthermore, a process of merging of all entities is carried out and omits the attribute parts from Link, Content, Cellular and Local which are not needed, and only take the main attributes. Then, a statistical analysis technique is carried out in which this technique presents data on user access patterns provided in the form of graphical displays that become parameters of analysis with associations rules. The last step is to determine market segmentation in applications based on similarity patterns, which will be obtained in the transaction pattern.
- d. The next step, will continue with pattern analysis and visualize the results of the mining process in the form of graphs of data presentation, which can be made as a decision to improve e-commerce and SEO performance, so that in the last phase requires a relationship between several processes to get the data integrated properly, in the form of application relations with visitors, obtained from traffic on the website, the most dense visit time, and the most frequent results of visits on the pages in the application, which will be used as input on SEO machines and e-commerce applications.

Conclusion

This paper has presented a proposal to improve e-commerce performance by utilizing the use of data mining integrated with SEO as a powerful tool. Data mining has been known as an intelligent method in the field of computational data processing, it has been applied to many things, while SEO is a popular technique for ranking websites for search results. The integration method that we propose in this paper is important in improving digital marketing performance. In this paper we have also presented data sources that can be processed using statistical data mining methods and how to process those originating from various sources or digital content, so that they can be used as evaluation material to determine the most influential parameters in improving e-commerce performance.

References

- Baker, Loren. *A Brief History of Search Engine Optimization*. 2003.
<https://www.searchenginejournal.com/seo-101/seo-history/> (accessed Oktober 10, 2018).
- Bank Indonesia. "Jumlah Uang Elektronik." *Bank Sentral Republik Indonesia*. February 25, 2016.
<https://www.bi.go.id/id/statistik/sistem-pembayaran/uang-elektronik/Contents/Jumlah%20Uang%20Elektronik.aspx> (accessed March 16, 2019).
- Brickley, Dan, Stéphane Corlosquet, Jason Douglas, Vicki Tardif Holland, Martin Hepp, dan Charlie Jiang. *Structured Data In SEO*. Schema.Org, 2011.
- Calabretta, Giulia, dan Maaïke Kleinsmann. "Technology-driven evolution of design practices:envisioning the role of design in the digital era." *Journal of Marketing Management*, 2017: 292-304.

- CK Finanzpro GmbH. *Mobile Wallet Space in Indonesia difficult for newcomer*. 22 Januari 2018. <http://fintechnews.sg/15927/indonesia/mobile-wallet-space-in-indonesia-difficult-for-newcomer/> (diakses September 23, 2018).
- Digital Marketing Depot. *Search Marketing Now Free Webcast, e-book and white papers*. 2006. <https://digitalmarketingdepot.com/> (diakses Oktober 12, 2018).
- Fayyad, Usama, Gregory Piatetsky-Shapiro, and Padhraic Smyth. *From Data Mining to Knowledge Discovery in Databases*. AI Magazine, 2008.
- Fishkin, Rand, dan Moz Staff. *Beginners Guide To SEO*. Moz, Inc. 18 Desember 2015. <https://moz.com/beginners-guide-to-seo> (diakses September 10, 2018).
- Hastie, Trevor, Robert Tibshirani, dan Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Stanford.Edu, 2009.
- IdnTimes. *Pertumbuhan E-Commerce Indonesia Meningkatkan Tajam, Siapa di Posisi Teratas?* IDN Media. 24 Agustus 2017. <https://www.idntimes.com/business/economy/yoshi/pertumbuhan-e-commerce-indonesia-1/full> (diakses September 10, 2018).
- IDProgrammer. *Kelebihan dan Kekurangan SEO*. 2017. <http://idprogrammer.com/kelebihan-dan-kekurangan-seo/> (accessed Oktober 10, 2018).
- JakPat. *E-money Usage Survey Report Indonesia*. 31 Januari 2015. <https://blog.jakpat.net/e-money-usage-survey-report-indonesia/> (diakses Oktober 10, 2018).
- Kotler, Philip, dan Kevin Lane Keller. *Marketing Management 14th Edition*. New Jersey: Prentice Hall, 2012.
- Kriegel, Hans-Peter, Peer Kröger, dan Arthur Zimek. "Subspace clustering." Dalam *Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, oleh Hans-Peter Kriegel, Peer Kröger dan Arthur Zimek, 351–364. Wiley, 2012.
- Lamberton, Cait, dan Andrew T. Stephen. "A Thematic Exploration of Digital, Social Media, and Mobile Marketing: Research Evolution from 2000 to 2015 and an Agenda for Future Inquiry." *Journal of Marketing*, 2016: 146 - 172.
- Maketeers.com. *Lima Tren e-Commerce Tahun 2018*. 2017. <http://maketeers.com/lima-tren-e-commerce-tahun-2018/> (diakses September 12, 2018).
- Shelley, Ryan. *25+ Years of SEO: A Timeline of Search Engine Optimization*. 31 Mei 2016. <https://www.smamarketing.net/blog/25-years-of-seo-a-timeline-of-search-engine-optimization-infographic> (diakses Oktober 10, 2018).
- Sullivan, Danny. *Faq All About The New Google Rankbrain Algorithm*. June 2017. <https://searchengineland.com/faq-all-about-the-new-google-rankbrain-algorithm-234440> (diakses Oktober 12, 2018).
- Vargo, Stephen L., dan Robert F. Lusch. "It's all B2B and beyond: Toward a systems perspective of the market." *Industrial Marketing Management*, 2011: 181 - 187.
- WideSkills.Com. *Data Mining Processes | Data Mining Tutorial*. 2018. <http://wideskills.com/data-mining-tutorial/data-mining-processes> (accessed Oktober 12, 2018).
- Witten, Ian H., Eibe Frank, dan Mark A Hall. *Data Mining: Practical Machine Learning Tools and Techniques (3 ed.)*. Elsevier, 2011.