

Discussion on Mining Big Data on Consumer Views for Market-Driven Products

Truong Dinh Trang¹, Dinh Tran Ngoc Huy², Pham Hung Nhan³

¹Master, Posts and Telecommunications Institute of Technology

²MBA (corresponding), Banking University HCMC Ho Chi Minh city Vietnam - International University of Japan, Niigata Japan

³Master (corresponding), Political School of Ca Mau Province, Vietnam

ARTICLE INFO

Article History:

Accepted: 10 May 2023

Published: 26 May 2023

Publication Issue

Volume 10, Issue 3

May-June-2023

Page Number

420-427

ABSTRACT

The purpose of this study is to present Discussion on Mining big data on consumer views for market-driven products. Methodology: Methods of systematization, generalization and synthesis; to carry out the research overview, the theoretical basis of the thesis; Analytical, statistical, and comparative methods: to analyze and evaluate the results of the application model to identify the customer's point of view. The results hence include: A reasoning system for using large databases of customer views to improve product quality and adjust business advertising strategies. This has not been fully and clearly shown in many studies in Vietnam so far.

Key words: big data, consumer views, market driven products

I. INTRODUCTION

In today's modern world, competition between business organizations is becoming more and more fierce, the costs required to attract customers are increasing. Moreover, the management and understanding of the relationship between the business organization and its customers and related activities is becoming more and more of an issue of concern and challenge than ever before for strategic entrepreneurs (Gupta et al, 2006). It is clear that product supply chains have become much more complex, prompting businesses to link together and apply advanced and modern technologies to maintain

a competitive position in the market. Moreover, the association and cooperation to maintain promotion position is the traditional method, always effective. However, this method requires many subjective factors between business organizations in the same or even different fields. Subjective factors are sometimes difficult to reconcile to achieve effective cooperation and association. It can even cause undesirable effects on cooperative and interconnected organizations.

Hence authors select this topic:

“Discussion on Mining big data on consumer views for market-driven products”

II. LITERATURE REVIEW

Domestic studies use large databases for analyzing customer views in order to improve, innovate and create products and services, providing customers with important analytical methods to apply. Customer points, especially the application of intelligent algorithms, are not much. In the food sector, online opinion mining and emotional analysis models are implemented. Here the authors used a variety of machine learning models such as decision tree, logistics regression, naïve bayer to analyze comments, comment time, comment content, customer reviews of real products. used products through food.vn, a website specializing in cuisine in Vietnam. Machine learning models are trained and then evaluated with the model's output requirement to recognize negative and positive reviews in order to improve the content and presentation of foods so that attract more customers, exploit potential customer sources, and on that basis, managers understand the advantages and disadvantages of products and services to improve business strategies better. The accuracy of customer evaluation recognized by the model is over 90%, showing the very good efficiency of the proposed machine learning model in the research. In the banking sector, studies have pointed out the importance of building and exploiting a large database system in commercial banks. It is clear that big banking data is a huge source of data, updated every minute, which is playing a big role in the banking sector with specific applications such as: analysis, classification of customer satisfaction and behavior. customers, analysis, detection and warning, prevention of risky behavior, tampering, optimization of data processing during operation, analysis and decision support. Accessing, researching and exploiting big banking data will bring many benefits to banks in business such as: cost savings, increased development time and optimization. However, in this study, the author has not come up with an effective tool for big data analysis to meet the specific

requirements of saving costs, optimizing products as well as adjusting strategies. advertising business (Nguyen Van Cuong, 2020).

Many researchers also use Kano's model to quantify the importance of CR. Kano's model serves as a tool for understanding CRs and their impact on CS. In Kano's model, different CRs are classified into must-have attribute, one-way attribute, attractive attribute, neutral attribute, etc. Chen and Chuang presented a powerful design approach to achieve higher CS levels in terms of aesthetic qualities (Shieh et al, 2008). In such a powerful design approach, Gray relational analysis with Taguchi method has been proposed to optimize subjective quality with multi-criteria characteristics. Kano's model is then used to balance the weights of multiple criteria to facilitate designers to understand the relationship between performance criteria and CS. To decide the weight of multiple criteria, regression method with dummy variables is often used to identify important attributes. However, they are thought to have the potential to lead to incorrect classification of multiple criteria under certain conditions (Kwong & Bai, 2003). Therefore, a censored regression method is proposed to improve the performance of pseudo-regression with dummy variables to get more accurate attribute classification.

III.METHODOLOGY

Research object: Large database systems about customer views, intelligent algorithms, products and services that customers use and have feedback on.

Research scope: includes the scope of time, space and content.

+ About time scope: the thesis focuses on research on customer perspective database systems in the period 2016-2023.

+ About the spatial scope: the thesis is domestic and international research because it will use data from both domestic and international customers.

+ About the scope of content: The thesis is limited to the use and research of database systems on customer

views for popular products (do not use large databases of different fields). other areas). Develop a plan to analyze big data to improve and enhance the quality of products and services to adjust business and advertising strategies according to market trends.

Theoretical basis (model/research framework)

The research framework is shown in Figure 1. First of all, data from large databases, collected on social networking sites, e-commerce platforms, sales pages of business organizations, advertising agencies. reports are downloaded and collected. This data is raw and unprocessed data. The next step in the research framework is to implement the data preprocessing methods. Here a variety of methods are examined, applied to clean the data, reformat the data in a uniform manner and then label the data points that present a negative, positive or neutral opinion. established for the types of products, services and objects that customers are using. The data is then processed to clarify the most important characteristics. More specifically, it is to highlight the non-interdependent characteristics of the customer set. To do this, a number of intelligent algorithms are applied to evaluate the performance, classification efficiency, and identification of each feature.

Thus, the interdependent characteristics will be reduced to reduce the complexity of the final algorithm as well as increase the efficiency of identification and classification of the proposed algorithm. The data is then divided into sub-sets consisting of a training set, a test set, and a test set. A series of intelligent algorithm models are trained using the training data and then evaluated for performance and accuracy in classification and data recognition using the test dataset. This is to ensure that the performance of the smart algorithm model is good on different data sets, avoiding problems of learning and remembering too long on a data set of intelligent algorithm models. Then, the intelligent algorithm models will be compared with each other to find the optimal intelligent algorithm model for the input customer perspective data set. The intelligent

algorithm models will also be re-tested on the test dataset to ensure that efficiency and accuracy are well maintained for different data sets. A series of performance evaluation parameters are then also used to represent the performance of intelligent algorithm models (accuracy, sensitivity, specificity, AUC, propensity, ...). Finally, the analysis and evaluation of the results of the model provided that products, business strategies, advertising, market trends, and consumer trends will be performed with comparative methods, statistics to clarify which types of products need to be improved quality, improve the form as well as the existing business strategy is still appropriate and how it needs to be adjusted to match the wishes of customers. Conclusions about the optimal algorithm model as well as adjustments related to products, business strategies, advertising, market trends, and consumption trends will be given along with recommendations for collection. data, processing data as well as extracting characteristics of data expressing customer views, opinions, requests, feedback.

IV.MAIN FINDINGS

4.1 Analyze consumer requirements for product quality improvement

There are also studies investigating how to determine the target value of the specification for QFD. The decision model for robot selection is introduced by fuzzy linear regression and QFD. Fuzzy linear regression is for deciding the target value of the specifications when uncertain CRs are presented and an incorrect relationship between ECs is found in the QFD. Similarly, a fuzzy linear regression, QFD and zero-objective programming were applied to decide which enterprise resource planning (ERP) systems meet the CRs of companies. In this approach, QFD makes a decision that considers the relationship between companies' CRs and specification characteristics as well as the interaction of characteristics between ERP systems. However, in

these approaches, the only goal is to maximize CS. Then, a framework was proposed to determine the target value of the specification for QFD by fuzzy linear regression and fuzzy multi-objective programming. Fuzzy linear regression is to find functional relationships between CR and specification, and between specifications. Fuzzy multi-objective programming is built to determine the value of the specification by maximizing CS within a limited budget. At the same time, other goals such as technical difficulty and scalability of the specification were considered in this fuzzy multi-objective program (Nahm et al, 2004; Akao, 2003; Karsak et al, 2009).

To balance CS and development costs, a fuzzy multi-objective approach has been proposed for uncertain and ambiguous CRs (Sener et al, 2011). In this method, Kano's model is combined into a QFD with consideration of the inherent uncertainty of CR as well as the nonlinear relationship between CR and specification. Some researchers have also found that consumers tend to give higher importance to baseline CRs. Accordingly, a model incorporating Kano's model into the QFD was proposed to adjust for the importance of CR. Others have developed a similar model incorporating Kano's model into QFD. But it has been used in multiple concurrent product design scenario to understand CR and balance the importance of CR. The integration of Kano's model with QFD was also found to monitor the dynamic changes of CR. In their study, Kano's model was to determine the rate at which a certain category of Kano changes over time and how to improve the current model to accommodate possible future CRs. With the integration of Kano's Model with QFD, the importance of CR has been dynamically adjusted. However, sometimes CRs are very dynamic and they will change over time. To capture the rapid changes of CR, Gray theory is combined with QFD. With that model in mind, the importance of CRs was monitored for dynamic and future CR response. From the probabilistic point of view, the Markov series model is

also reported to analyze the rapid change of CR (Mu et al, 2008; Tontini, 2007; Sireli et al, 2007).

4.2 Analysis of customers' online views

An unsupervised model has been proposed to identify aspects and assess emotional polarities from online reviews (Lin et al, 2010). First, a probabilistic phylogenetic model was proposed, in which words in a single sentence are generated from one aspect. Later, an extended model was introduced, in which both aspect and emotion analyze emotions in different aspects. However, it was found that it failed to identify emotions specific to an aspect and did not separate sentimental words from factual information. Then, a general aspect and sentiment model was proposed to extract aspects and from emotions from online reviews (Zo & Oh, 2011).

V. DISCUSSION AND CONCLUSION

The basic objective of the study is to study the large database systems related to the views, feedbacks and opinions of customers about products, services, objects provided to the market and applications. Using customer perspective analysis method to analyze big data for the purpose of improving and improving the quality of products and services in line with market orientation as well as consumer trends.

The results hence include:

- A reasoning system for using large databases of customer views to improve product quality and adjust business advertising strategies. This has not been fully and clearly shown in many studies in Vietnam so far.

VI. ACKNOWLEDGEMENT

Thank you editors, friends to assist this publishing

VII. REFERENCES

- [1]. Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Sriram, S. (2006). Modeling Customer Lifetime Value. *Journal of Service Research*, 9(2), 139–155.
- [2]. Moradlou, H., & Backhouse, C. (2014). Re-shoring UK manufacturing activities, supply chain management & postponement issues. 18th Annual Cambridge International Manufacturing Symposium, 344–354
- [3]. Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: A systematic review and agere search. *Electronic Markets*, 26(2), 173–194.
- [4]. Hyman, Paul. 2012. “Researchers Struggle to Measure Big Data’s Impact.” *ACM Communications*, November 13.
- [5]. Basant, A., Namita, M., Pooja, B., Sonal Garg 2, 2015. Sentiment Analysis Using Common-Sense and Context Information. Hindawi Publishing Corporation Computational Intelligence and Neuroscience.
- [6]. Ainur, Y., Yisong, Y., Claire, C., 2010. Multi-level structured models for document-level sentiment classification. In: *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. MIT, Massachusetts, Association for Computational Linguistics, USA, pp. 1046–1056.
- [7]. Noura, F., Elie, C., Rawad, A.A., Hazem, H., 2010. Sentence-level and document-level sentiment mining for arabic texts. In: *Proceeding IEEE International Conference on Data Mining Workshops*.
- [8]. Nikos, E., Angeliki, L., Georgios, P., Konstantinos, C., 2011. ELS: a word-level method for entity-level sentiment analysis. In: *WIMS ’11 Proceedings of the International Conference on Web Intelligence, Mining and Semantics*.
- [9]. Haochen, Z., Fei, S., 2015. Aspect-level sentiment analysis based on a generalized probabilistic topic and syntax model. In: *Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference*. Association for the Advancement of Artificial Intelligence.
- [10]. Nguyen Dang Lap Bang, Nguyen Van Ho, Ho Trung Thanh, Opinion mining model and online customer sentiment analysis in the food industry, *Journal of Science*, Ho Chi Minh City Open University, 16(1), 64-78.
- [11]. Dao My Hang, Dang Thu Hoai, Big data applications - challenges for Vietnamese commercial banks, *Journal of Banking Science & Training*, No. 224+225- January & 2. 2021.
- [12]. Nguyen Van Cuong, 2020. Applying data mining techniques to study customer behavior according to circadian rhythms in mobile marketing in a real estate project in Ho Chi Minh City. *Journal of Science Can Tho University*. 56(4D): 198-207.
- [13]. Nguyễn Thành Thủy, Trần Thị Châu Giang, Một mô hình học máy trong phân tích ý kiến khách hàng dựa trên văn bản tiếng Việt: Bài toán dịch vụ khách sạn, *Kỷ yếu hội thảo khoa học quốc gia 2019 “CNTT và ứng dụng trong các lĩnh vực”*, 2019.
- [14]. Karsak, E. Ertugrul, 2004. “Fuzzy Multiple Objective Decision Making Approach to Prioritize Design Requirements in Quality Function Deployment.” *International Journal of Production Research* 42 (18): 3957–3974.
- [15]. Chen, Yizeng, Richard Y. K. Fung, and Jiafu Tang. 2006. “Rating Technical Attributes in Fuzzy QFD by Integrating Fuzzy Weighted Average Method and Fuzzy Expected Value Operator.” *European Journal of Operational Research* 174 (3): 1553–1566.
- [16]. Wang, Ying-Ming. 2012. “Assessing the Relative Importance Weights of Customer Requirements Using Multiple Preference Formats and Nonlinear Programming.” *International Journal of Production Research* 50 (16): 4414–4425.

- [17].Chen, Chun-Hsien, Li Pheng Khoo, and Wei Yan. 2002. "A Strategy for Acquiring Customer Requirement Patterns Using Laddering Technique and ART2 Neural Network." *Advanced Engineering Informatics* 16 (3): 229–240.
- [18].Chen, Chun-Chih, and Ming-Chuen Chuang. 2008. "Integrating the Kano Model into a Robust Design Approach to Enhance Customer Satisfaction with Product Design." *International Journal of Production Economics* 114: 667–681.
- [19].Shieh, Meng-Dar, Wei Yan, and Chun-Hsien Chen. 2008. "Soliciting Customer Requirements for Product Redesign Based on Picture Sorts and ART2 Neural Network." *Expert Systems with Applications*. 34 (1): 194–204.
- [20].Kwong, C. K., and H. Bai. 2003. "Determining the Importance Weights for the Customer Requirements in QFD Using a Fuzzy AHP with an Extent Analysis Approach." *IIE Transactions* 35 (7): 619–626.
- [21].Lin, Chenghua, Yulan He, and Richard Everson. 2010. "A Comparative Study of Bayesian Models for Unsupervised Sentiment Detection." *CONLL'10, Uppsala, Sweden*. 144–152.
- [22].Li, Yanlai, Jiafu Tang, Xinggang Luo, and Jie Xu. 2009. "An Integrated Method of Rough Set, Kano's Model and AHP for Rating Customer Requirements' Final Importance." *Expert Systems with Applications* 36 (3): 7045–7053
- [23].Nahm, Yoon-Eui, Haruo Ishikawa, and Masato Inoue. 2013. "New Rating Methods to Prioritize Customer Requirements in QFD with Incomplete Customer Preferences." *The International Journal of Advanced Manufacturing Technology* 65: 1587–1604.
- [24].Akao, Yoji. 2004. *QFD: Quality Function Deployment – Integrating Customer Requirements into Product Design*. New York: Productivity Press.
- [25].Chan, Lai-Kow, and Ming-Lu Wu. 2002. "Quality Function Deployment: A Literature Review." *European Journal of Operational Research* 143 (3): 463–497.
- [26].Nepal, Bimal, Om P. Yadav, and Alper Murat. 2010. "A Fuzzy-AHP Approach to Prioritization of CS Attributes in Target Planning for Automotive Product Development." *Expert Systems with Applications* 37 (10): 6775–6786.
- [27].Lai, Xin, Min Xie, Kay-Chuan Tan, and Bo Yang. 2008. "Ranking of Customer Requirements in a Competitive Environment." *Computers & Industrial Engineering* 54 (2): 202–214.
- [28].Kwon, Changhyun, Terry L. Friesz, ReetabrataMookherjee, Tao Yao, and Baichun Feng. 2009. "Non-cooperative Competition among Revenue Maximizing Service Providers with Demand Learning." *European Journal of Operational Research* 197 (33): 981–996.
- [29].Karsak, E. Ertugrul, 2008. "Robot Selection Using an Integrated Approach Based on Quality Function Deployment and Fuzzy Regression." *International Journal of Production Research* 46 (3): 723–738.
- [30].Karsak, E. Ertugrul, and C. Okan Özogul, 2009. "An Integrated Decision Making Approach for ERP System Selection." *Expert Systems with Applications* 36 (1): 660–667.
- [31].Sener, Zeynep, and E. ErtugrulKarsak. 2011. "A Combined Fuzzy Linear Regression and Fuzzy Multiple Objective Programming Approach for Setting Target Levels in Quality Function Deployment." *Expert Systems with Applications* 38 (4): 3015–3022.
- [32].Mu, Li-Feng, Jia-Fu Tang, Yi-Zeng Chen, and C. K. Kwong. 2008. "A Fuzzy Multi-objective Model of QFD Product Planning Integrating Kano Model." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 16 (6): 793–813.
- [33].Tontini, Gérson. 2007. "Integrating the Kano Model and QFD for Designing New Products." *Total Quality Management & Business Excellence* 18 (6): 599–612.

- [34].Sireli, Y., P. Kauffmann, and E. Ozan. 2007. "Integration of Kano's Model into QFD for Multiple Product Design." *IEEE Transactions on Engineering Management* 54 (2): 380–390.
- [35].Raharjo, Hendry, Aarnout C. Brombacher, T. N. Goh, and Bo Bergman. 2010. "On Integrating Kano's Model Dynamics into QFD for Multiple Product Design." *Quality Reliability Engineering International* 26 (4): 351–363.
- [36].Wu, H. H., A. Y. H. Liao, and P. C. Wang. 2005. "Using Grey Theory in Quality Function Deployment to Analyse Dynamic Customer Requirements." *The International Journal of Advanced Manufacturing Technology* 25: 1241–1247.
- [37].Wu, Hsin-Hung, and Jiunn-I Shieh. 2006. "Using a Markov Chain Model in Quality Function Deployment to Analyse Customer Requirements." *The International Journal of Advanced Manufacturing Technology* 30: 141–146.
- [38].Mukherjee, Arjun, and Bing Liu. 2012. "Aspect Extraction through Semi-supervised Modeling." *ACL'12, Jeju Island, Korea*. 339–348.
- [39].Lin, Chenghua, and Yulan He. 2009. "Joint Sentiment/Topic Model for Sentiment Analysis." *CIKM'09, Hong Kong, China*. 375–384.
- [40].Lin, Chenghua, Yulan He, and Richard Everson. 2010. "A Comparative Study of Bayesian Models for Unsupervised Sentiment Detection." *CONLL'10, Uppsala, Sweden*. 144–152.
- [41].Jo, Yohan, and Alice H. Oh. 2011. "Aspect and Sentiment Unification Model for Online Review Analysis." *WSDM'11, Hong Kong, China*. 815–824.
- [42].ThiNgu,D.,Huong,D.T.,Huy,D.T.N.,Thanh,P. T., &Dongul, E. S. (2021). Language teaching application to English students at master's grade levels on history and macroeconomic-banking management courses in universities and colleges. *Journal of Language and Linguistic Studies*, 17(3), 1457-1468.
- [43].ThiHoa,N.,Hang,N.T et al. (2021). Human resource for schools of politics and for international relation during globalization and EVFTA. *Elementary education online*, 20(4).
- [44].TTH Ha et al. (2019). Modern corporate governance standards and role of auditing-cases in some Western european countries after financial crisis, corporate scandals and manipulation, *International Journal of Entrepreneurship*
- [45].TDT Vu et al. (2021). Human Education And Educational Issues For Society And Economy-Case In Emerging Markets Including Vietnam, *Elementary education Online* 20 (2)
- [46].TH Le et al. (2021). Solutions To Enhance Students' Awareness In E-Learning Training-And Technical Application For Education Quality Enhancement At Colleges And Universities, *Design Engineering*, 7648-7659
- [47].VTT Dung et al. (2021). Enhancing the capabilities of students after graduation-a case study at university of economics and business administration-thainguyen university, vietnam
- [48].VQ Nam et al. (2021). Suggested Risk Policies from Comparison of 2 Groups of Vietnam Banks-Previous SOE Banks and Private Banks During Post-Low Inflation Period 2015-2020, *RevistaGeintec-GestaoInovacao E Tecnologias* 11 (2), 531-546
- [49].VQ Nam et al. (2021). Increasing Agricultural Productivity, Quality and Quantity of Coffee and Tea Crops Planting and Marketing Mix Solutions-Methods of Eliminating Coffee Berry Borer and Insects in Nprth, *Alinteri Journal of Agriculture Sciences* 36 (1)
- [50].VQ Nam et al. (2021). Solutions to Promote Startup for the Youth in Minoritty and Moutainous Region of Thai Nguyen Province-Vietnam, *Journal of Contemporary Issues in Business and Government* 27 (3), 2113-2118
- [51].V Van Chung et al. (2021). Eastern philosophical theories and marxism-lenin philosophies in

- vietnam society and education, Review of International Geographical Education Online 11 (8), 1586-1591
- [52]. VQ Nam et al. (2021). Internet of Things (IoTs) Effects and Building Effective Management Information System (MIS) in Vietnam Enterprises and Human-Computer Interaction Issues in Industry 4.0, Webology 18
- [53]. Thuy, D. V. T., Huy, D. T. N., Anh, V. T. K., Thach, N. N., & Hanh, H. T. (2021). Quality of education of ethnic minority communities in vietnam-problems and recommendations. Ilkogretim Online, 20(4).
- [54]. Tinh, D. T., Thuy, N. T., & Ngoc Huy, D. T. (2021). Doing Business Research and Teaching Methodology for Undergraduate, Postgraduate and Doctoral Students-Case in Various Markets Including Vietnam. Ilkogretim Online, 20(1).
- [55]. Tram, P. N et al. (2021). Educational, Political and Socio-Economic Development of Vietnam Based on Ho Chi Minh's Ideology. Ilkogretim Online, 20(1).
- [56]. Trung, N. D., Thu, B. T et al. (2022). Using English To Teach Students With Social Sciences Major-Via A Case Of Some Vietnam Newspapers With The Uk, Italian And French Approaches And Regulations On Publishing Fake News And Internet Crime. Journal of Language and Linguistic Studies, 17(3).
- [57]. Vu, T. D. T et al. (2021). Human Education And Educational Issues For Society And Economy-Case In Emerging Markets Including Vietnam. Ilkogretim Online, 20(2).

Cite this Article

Truong Dinh Trang, Dinh Tran Ngoc Huy, Pham Hung Nhan, "Discussion on Mining Big Data on Consumer Views For Market-Driven Products", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 10 Issue 3, pp. 412-419, May-June 2023. Available at doi : <https://doi.org/10.32628/IJSRST52310397>
Journal URL : <https://ijsrst.com/IJSRST52310397>