International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-1, October, 2019

Investigating Significant Factors Influencing Response Time of Customer Complaint Based on Analytics Method

HafizahFarhah Saipan Saipol, NurTasnimShamsuddin, ZaihismaChe Cob, NurLailaAb Ghani, Sulfeeza Mohd Drus

Abstract: The purpose of this study is to investigate significant factors that influenced duration of solving financial institutions' customer complaint. Using raw customer complaint dataset from Consumer Financial Protection Bureau (CFPB) website, it was found that many of sub-categories are not well organized. Thus, it is important to proceed with data cleaning and data preparation steps before any analysis been performed. In this study, Artificial Neural Network (ANN) had been chosen since it can deal with non-linear relationship by using sigmoid function. Further to this, it was found that Product, Company response and Issues are the significant factors that are more likely to be solved more than one day. The use of this analysis can be particularly beneficial for related financial party that might need to assist their customer in future.

Keywords: Customer Complaint, Neural Network, Data Cleaning

I.INTRODUCTION

Dissatisfying experiences or unsatisfactory problem resolution while dealing with the financial services or products may cause customer to complaint or remain silence. The value of the customer complaints can turn a dissatisfied customer into a satisfied and loyal customer if only the complaints are handled properly. The main priority of customer-centric enterprises focused on handling complaints before it turned into large numbers to save manpower (Xu et al., 2017). Analyzing the complaint data help the institutions to gain valuable insights from the complaint data performance and improve the service to reduce the complaint in the future. Data mining in complaint management could be conducted to discover the unseen patterns of complaints from a company's database. Besides, it could uncover the root of the problems as indicated by Bae et al. (2005) and LarivièRe and Poel (2005).

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
HafizahFarhah Saipan Saipol*, Institute of Informatics and Computing in Energy, Universiti TenagaNasional, Jalan IKRAM- UNITEN, 43000 Kajang, Selangor
NurTasnimShamsuddin, UNITEN Research and Development, Jalan IKRAM-UNITEN, 43000 Kajang, Selangor
ZaihismaChe Cob, College of Computing and Informatics, Universiti TenagaNasional, Jalan IKRAM-UNITEN,43000 Kajang, Selangor
NurLailaAb Ghani, College of Computing and Informatics, Universiti TenagaNasional, Jalan IKRAM-UNITEN, 43000 Kajang, Selangor
Sulfeeza Mohd Drus, College of Computing and Informatics, Universiti TenagaNasional, Jalan IKRAM-UNITEN, 43000 Kajang, Selangor

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Developing model for better management of customer complaints has attracted the researchers (Atlikhan et al., 2013). Chen et al. (2012) considered four analytical stages of model: primary diagnosis, advanced diagnosis, review and action based on data mining of customer-complaint databases to diagnose and correct service failures. The model which the core concept focused on the customer, integrated the views of customers, managers and consultants to improve the strategy and to eliminate the causes of service failure.

Understanding the application domain is one of the crucial stages in order to understand what the customer wants to accomplish from a business perspective before proceed to the next stage: data understanding and data preparation or data mining. Several data mining techniques such as neural network, self-organizing map, survival analysis, and regression analysis have been conducted in complaint management (Bae et al., 2005; Bijmolt et al., 2014; Hadden et al, 2008; Huang, 2012; LarivièRe and Poel, 2005). Artificial Neural Network (ANN) represents the most commonly used approaches in the literature (Ngai et al., 2009). There are few studies on customer complaint using ANN as one of their predictive analysis to solve related current issues (Hadden et al., 2008; Huang, 2012). They share the common goal to identify the root problem and predict the customer expectations that could help the organization to create a deeper understanding of customers and to maintain good customer relations. In addition, ANN can be used as preliminary predictive analysis which be used to predict significant factors at general level. Recently, passenger's decision making and perceptions had successfully predicted by using ANN (Islam et al., 2016). This statement been supported by Ansari et al. (2016) saying that customer loyalty elements can be adequately assessed by ANN.

The aim of this study is to obtain the significant factors that influenced duration of solving customer complaint using ANN. This paper is organized into four sections. Section 2 contains the methodology based on CRISP-DM process model. In Section 3, results are analyzed and the conclusion of this study is presented in Section 4.

II.RELATED WORKS

Meeting customer expectation and maintaining good customer relations requires investigating significant factors that influenced the duration of solving the customer complaints.

This will improve the customer complaint management system. In response to achieve the target, Hadden et al. (2008) have investigated the importance of the provisions data which represents whether the companies took a timely response to resolve the complaint. Besides, they also emphasized the importance of the complaint information such as the type of complaint, duration of complaints to be resolved and the number of complaints made in a certain period of time. Another important issue proposed by Huang (2012) who have investigated the importance of service recovery and its effectiveness which increases customer satisfaction. Customers' feedback regarding the companies' respond towards the complaint is one of the parameters of interest to show the customer satisfaction towards the service recovery. This is supported by Harrison and Shirom (1999) emphasized the significance of information feedback to shape the organization operations.

Table. 1 Summary of related words based on previous study

No.	Author (Year)	Significant attributes
1.	Henneberg et al. (2008)	There were seven attributes of expected complaint resolution management: i) Take quick action ii) Understand problem iii) Empathy iv) Active listening v) Manners vi) Openness
2.	Fierro et al. (2015)	This study found that better satisfaction level among customers in complaint handling lead towards customer's engagement ($p< 0.001$).
3.	Supriaddin et al. (2015)	Based on the structural analysis in this study, it was found that below attributes had significant effect on complaint handling: i) Customer satisfaction ii) Customer trust iii) Customer loyalty Based on the confirmatory factor analysis (CFA), it was found that the speed of handling customer complaint is the most significant indicator for management to create procedure of handling customer complaint.

Previous studies showed that the speed of handling of customer complaint and satisfactionlevel among customers during handling complaints were associated with customer engagement and reflecting the management of customer service on complaint handling.

III.MATERIALS AND METHOD

In order to conduct this study, a secondary dataset was retrieved from an open source website. Datasets from Consumer Financial Protection Bureau (CFPB) were downloaded from CFPB website and were cleaned to build ANN model and predict the significant factors that influences duration of solving customer complaint.

The frequently used methodology was known as Cross Industry Standard Process for Data Mining (CRISP-DM) (Piatetsky, 2014). There are six phases of CRISP-DM process model which are business understanding, data understanding, data preparation, modeling, evaluation and deployment. This study will not adapt the last phase of the process model. Details on each phase are explained in the following sub-sections.

Business understanding

Consumer Financial Protection Bureau (CFPB) is an organization that was created in 2010 to write and force rules for financial institutions, examine both bank and non-bank financial institutions monitor and report on markets, as well as collect and track consumer complaints. CFPB received thousands of complaints about financial products such as mortgage, bank account, credit card, loan and virtual currency. Once CFPB received the complaints, they sent to the related companies to solve the problem and respond to the customer. This action could help the customer to be treated fairly by the financial institutions. The raw customer complaint dataset is publicly available to be accessed and downloaded through the CFPB website for public use.

Data understanding

The complaint data set consists of 18 attributes: date received, date sent to company, product, sub-product issue, sub-issue, consumer complaint narrative, company public response, company, state, ZIP code, tags, consumer consent provided, submitted via, company response to consumer, timely response, and consumer disputed, complaint ID.

A total of 1,048,576 complaints regarding financial products and services from 1st December 2011 to 5th September 2018 are submitted to CFPB through 6 channels: email, fax, phone, postal mail, referral and web. The data set are the complaints from 4,873 companies. These companies came from 63 different states from United States of America which includes 50 states, 9 Commonwealth states, 3 military states and 1 US minor outlying island. The complaint comprises of 18 products, 52 sub-products and 166 issues and 219 sub-issues. The complaint data set requires data cleaning before analyzing the data to investigate significant factors that influenced duration of solving customer complaint.

Data preparation

Data cleaning

Inconsistencies in the observations result from different CFPB worker as well as captured from a variety of complaint channels. It is essential to ensure the quality data because erroneous data limit the performance of statistical methods induce misleading analytics and data mining results (Berti-Equille, 2007; Wahlin and Grimvall, 2008). For example, duplicate observations are found in sub-product attribute which have the same meaning but written in two different ways:

i. "(CD) Certificate of deposit" is recorded in 1st March 2012 until 21st April 2017 and "CD (Certificate of Deposit)" recorded between 24th April 2017 until 28th August 2018;

ii. "Traveler's/Cashier's checks" is recorded in 23rd July 2014 until 12th April 2017 and "Traveler's check or cashier's check" recorded between 27th April 2017 until 8th August 2018;

These inconsistencies not only occurred for sub-product but also for product and issue column. The same categories are entered differently after 23rd April 2017 may due to new system updated after the date. Products, sub-products, issues and duration are merged into smaller groups with the same classification due to several reasons: the sparse classes have very few total observations; and the feature has several classes that are quite similar. After data transforming process, product which consisted of 18 classes, are merged into 10 classes. Meanwhile 77 sub-products are merged into 45, and 166 issues merged into 13 classes.

Sub-issue attribute is removed due to two reasons: the class of sub-issue is large containing 219 classes; and very few total values of observations for each class which will not give significant impact to data analysis. Besides, the new issue classification has taken into consideration the sub-issue in order to categorize the new issue. Another attribute that is being removed is consumer complaint narrative due to its suitability only for text mining analysis.

Since the data provides the date CFPB received the complaints and the date CFPB sent the complaint to the respected company, thus, the duration of CFPB took to send the complaint data to the company can be calculated. This is crucial to predict the time taken for the complaints to be solved within same day or more than one day. The duration group is created as a new attribute.

This dataset comprises of two types of data: numerical and categorical. Numerical data involves quantitative values such as date received, data sent to company, complaint ID and duration. Meanwhile categorical data describes qualitative without numerical values such as product, sub-product, issue, company public response, company, state, tags, consumer consent provided, submitted via, company response to consumer, timely response, and consumer disputed. These category data are further classified into nominal category data (Zhang et al., 2018) for further analysis.

Modelling and Evaluation

In this study, ANN been used to seek significant factors from upper view to estimate time taken to solve the complaint within same day or more than one day. Outcome from ANN model can be evaluated by looking into the sigmoid values and accuracy rate. Artificial Neural Network (ANN) model were created using RapidMiner software. It consists of various nodes to help finding significantfactors to predict duration of solving customer complaint (within same day or more than one day).

Artificial Neural Network (ANN) was known as connection among human brains to proceed with certain function or job by using software simulation (Leong et al., 2015). This technique applied in various sectors such as detecting fraudulent credit card transactions, modelling financial time series, recognizing handwritten numbers and letters and to estimate real estate values. To be precise, the idea of ANN models was coming from human brain studies where connection of human brain works by learning from experience and it had been adapted into computer's application to follow specific instructions over and over. To summarize, ANN develop an ability to generalize and learn from data in a way human's ability to learn experience.

IV.RESULT AND DISCUSSION

Based on ANN model, it was found that there are five factors that are significant to predict the duration of solving customer complaint. It was shown that Consumer decision to dispute and Type of complaint submission are more significant to predict output "0" (duration of solving customer complaint within same day). In contrast, Issues, Company response and Product are more significant to predict output "1" (duration of solving customer complaint more than one day). Sigmoid values are mathematical function commonly used to handle binary classification. In this study, the target variable was duration of solving customer complaint ("0": within same day and "1": more than one day). Sigmoid function as per below:

$$S(x) = \frac{1}{1 + e^{-x}}$$

In addition, sigmoid probability values for Issues, Company response and Product were 0.619, 0.826 and 0.99 respectively. This indicate that there are 61.9% that customer complaint due to Issues factor were going to be solve more than one day. Moreover, it showed that almost 99% and 82.6% of customer complaint more than one day were due to Product and Company response factor accordingly.

Further to this, evaluation on ANN model were done and based on this study, accuracy rate was 76.87% with root mean squared error is 0.4. This finding showed that the model is adequate with the dataset given. In addition, upper management might focus more on the three significant issued which are Issue, Product and Company response in order to improve their customer services. This study only showed a general view and factors that might influence the duration of customer complaint.

V.CONCLUSION

Datasets from Consumer Financial Protection Bureau (CFPB) are analyzed to investigate significant factors that influenced duration of solving financial institutions' customer complaint. It was found that ANN model was one of the effective tools for modeling duration of customer complaint. It consists of various nodes and to find the best node, sigmoid function was used as activation function.

By using ANN, it was found that Product, Issue and Company response were significant factors to predict duration of solving customer complaint. To date, this fulfil the study objectives and it is recommended that upper management of financial companies will pay more attention on their Product and Issues services. The use of ANN for analyzing the duration of solving customer complaint can be particularly useful for relevant financial party who strive to overcome current issues regarding their customers.The important of solving customer complaint is more apparent in service sector to ensure customer loyalty and support. This model enables upper management for a better understands regarding their customer complaint.

Limitation of this study was incomplete dataset due to missing data. In addition, the sample size within group's attribute were not equal thus resulting in less accuracy of the model. However, this study was using raw dataset from CFPG website hencereflected relatedfactorsof resolving customer complaints.

ACKNOWLEDGMENT

The authors would like to express appreciation for the funding support by TNB R&D Seeding Fund and for securing and managing the fund by UNITEN and UNITEN Research and Development Sdn. Bhd [Project Number =U-TC-RD-18-17].

REFERENCES

1. Ansari, A. and A. Riasi. 2016. Modelling and evaluating customer loyalty using neural networks: Evidence from startup insurance companies. Future Bus. J., 2: 15-30.
2. Atlikhan, A., A. Albadvi and A.H. Ghapanchi. 2013. A conceptual model for proactive-interactive customer complaint management systems. Intern. J. of Bus. Inf. Syst., 13: 490-503.
3. Bae, S.M., S. H. Ha and S.C. Park. 2005. A web-based system for analyzing the voices of call center customers in the service industry. Expert Syst. with App., 28: 29–41.
4. Berti-Equille, L. 2007. Measuring and modelling data quality for quality-awareness in data mining. Quality Measures in Data Mining Studies in Computational Intelligence, 43: 101-126.
5. Bijmolt, T.H.A., E.K.R.E. Huizingh and A. Krawczyk. 2014. Effects of complaint behaviour and service recovery satisfaction on consumer intentions to repurchase on the internet, Internet Res., 24 (5): 608-628.
6. Chen, C.K., A.J. Shie and C.H. Yu. 2012.A customer-oriented organisational diagnostic model based on data mining of customer-complaint databases. Expert Syst. with App. 39: 786-792.
7. Fierro, J.C., I.M. Polo and F. J. Sese. 2015. Can complaint-handling efforts promote customerengagement?.Service Business. 10(4), 847-866.
8. Hadden, J. A. Tiwari, R. Roy and D. Ruta. 2008. Churn Prediction using Complaints Data. World Academy of Sci., Eng. and Tech., 19: 809-814.
9. Harrison, M. I., and A. Shirom. 1999. Organizational diagnosis and assessment:Bridging theory and practice. Thousand Oaks, Sage Publications.
10. Henneberg, S.C. B. Ashnai, T. Gruber, P. Naude, A. Reppel. Using Means-end approaches to understand business complaint management. Manchester IMP Research Group. http://www. impgroup. org/uploads/papers/6791. pdf [Accessed 20/06/2019].
11. Huang, H.C. 2012. Using artificial neural networks to predict restaurant industry service recovery. Int. Journal of Advancements in Computing Tech., 4(10):315-321.
12. Islam M.R., M. Hadiuzzaman, R. Banik, M.M. Hasnat, S.R. Musabbir and S. Hossain. 2016. Bus service quality prediction and attribute ranking: a neural network approach. Public Transp., 8: 295-313. doi: 10.1007/s12469-016-0124-0.
13. LarivièRe, B. and D.V.D. Poel. 2005. Investigating the post-complaint period by means of survival analysis. Expert Syst. with App., 29: 667–677.
14. Leong, L.Y., T.S. Hew and K.B. Ooi. 2015.A SEM-artificial-neural-network analysis of the relationship between SERVPERF, customer satisfaction and loyalty among low-cost and full-service airline. Expert Syst. with App., 42: 6620-6634.
15. Ngai, E.W.T., L. Xiu and D.C.K. Chau. 2009. Application of data mining techniques in customer relationship management: A literature review and classification. Expert Syst. with App., 36: 2592–2602.
16. Piatetsky, G. 2014. CRISP-DM still the top methodology for analytics, data mining, or data science projects. Retrieved 14.03.2019. http://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-projects.html.
17. Supriaddin N., A. Palilati, H. Bua, Patwayati, H. Jusuf. 2015. The effect of complaint handling toward customers satisfaction, trust and loyalty to Bank Rakyat Indonesia (Bri) Southest Sulawesi. The International Journal of Engineering and Science, 4:1-10.
18. Wahlin, K. and A. Grimvall. 2008. Uncertainty in water quality data and its implications for trend detection: Lessons from Swedish environmental data. Environ. Sci. and Policy, 2: 115-124.
19. Xu, X., H.Y. Li, S.C. Wang and Y.L. Shen. 2017. Prediction model of power customer complaints based on PCA-BP neural network. 2017 3rd International Conference on Computer Science and Mechanical Automation (CSMA 2017), 87-96.
20. Zhang, C., L. Cao and A. Romagnoli. 2018. On the feature engineering of building energy data mining. Sustain. Cities and Society, 39: 508-518.