OPTIMISING MARKET SEGMENTATION FOR THE TELECOMMUNICATIONS INDUSTRY: A CONTEXTUAL MARKETING BASED APPROACH: CASE STUDY AT PT TELKOMSEL

AgusWahyu Widodo [1], GadangRamantoko[2]

Faculty of Economics and Business, Telkom University,
INDONESIA.

[1]mr.wewe79@gmail.com, [2]gadangramantoko@gmail.com

ABSTRACT

Telkomsel runs thousands of campaigns every month by targeting High Value Customer (HVC), but has proven to have limited level of personalization. Based on the data obtained, from 87 trillion Telkomsel revenue, as much 52% was contributed by HVC, amounting to 14% of total customers. Nevertheless, the take up rate of product campaigns conducted on the HVC segment is still low (around 4.5%). On the other hand the "One-product-for-all" solution is no longer accepted, because customers become more demanding and looking for personalized products and services that meet to their needs. In response, an enhanced HVC segmentation scheme in accordance with Contextual Marketing is formed, that is by adding the number of behavioral attributes. This study aims to (1) determine the effect of adding customer behavior attributes in the segmentation of take up rate, (2) knowing HVC segmentation along with the profile formed, and (3) knowing the contextual campaign strategy that is suitable for Telkomsel HVC subscriber by matching existing products with new segments. Theory approach for Customer Segmentation, Customer Profile, Customer Behavior and Contextual Marketing is applied. Data analysis is done by using SPSS Modeler software with K-Means clustering algorithm and Logistic Regression algorithm. The segmentation was carried out using 10 attributes consist of Lange of Stay, Flag Data User and 8 additional attributes based on customer behavior. The result shows that when 8 behavior attributes added it can intensify campaign effectiveness as showed in increase of 3.86% take-up rate. As well as obtained 5 new segments consist of customer with characteristics such as: Data User Heavy on Internet with High Recharge (29.0%), Loyal customer, heavy on Voice and SMS with Medium Recharge (22.3%), Normal usage with High Recharge (18.5%), Customer Heavy on Voice and SMS with medium Recharge (20.7%), and Value Customer, Heavy on Internet and Voice with High Recharge (9.6%). Campaign strategy is done by matchmaking existing product with the characteristics and behavior of each segment.

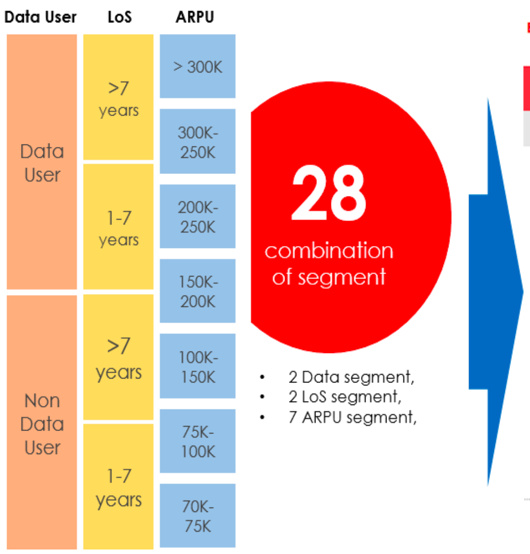
Keywords:K-Means Clustering, Logistic Regression, Contextual Marketing, Customer Profile, Customer Segmentation

INTRODUCTION

Established in 1995, Telkomsel succeed to become leading cellular operator in Indonesia within costumer base more than 173 million, represent almost 50% of telecomunications market share in Indonesia (Telkom, 2016). In terms of operational performance Telkomsel is the leader in nusantara. Data shows in 2016 Telkomsel was capable to increase its costumer base up to 13.96% compared to previous year, and be able to escalate operational profit as much as 42.63%. Within the same year, level of revenue per user Telkomsel (ARPU) increase to 44.9 thousands rupiahs or increase on average 5.2% compared to previous year. Escalation of ARPU were dominated by loyal costumer of Telkomsel with subscription period more than 1 year, which is referred to High Value Customer (HVC).

Leena and Luna International, Chikusei, Japan.
(株) リナアンドルナインターナショナル, 筑西市,日本
ISSN: 2186-845X ISSN: 2186-8441 Print
www.ajmse. leena-luna.co.jp

Telkomsel HVC is the customer with LoS (Length of Stay) more than 1 year and ARPU (Average Revenue per User) more than 70 thousands rupiahs. What makes it remarkable is though the HVC is only from customer base total in company, they contribute in more than a half of company revenue (52%) (Telkomsel, 2017). Therefore, enhancement of HVC customer segmentation become important for company as it will give significant impact in potential revenue.



Breakdown per HVC Segment

		Overall				
ARPU SEGMENT		TARGET (in Mn subs)	TAKER ('000 Subs)	TRX ('000 Subs)	TAKER RATE	REVENUE (in Billion IDR)
GRAND TOTAL		23.2	1,048.0	2,613.6	4.5%	105.4
HVC	70K-75K	1.5	97.8	292.6	6.3%	5.9
	75K-100K	6.0	358.9	904.1	5.9%	31.1
	100K-150K	7.1	334.5	881.6	4.7%	33.8
	150K-200K	3.5	130.7	302.5	3.7%	14.8
	200K-250K	1.9	48.2	89.5	2.5%	5.3
	250K-300K	1.1	25.5	42.8	2.3%	3.9
	>300K	2.0	52.5	100.4	2.7%	10.6

Figure 2. Exisiting Customer Segmentation of Telkomsel HVC & Take up Rate of Campaign

In current situation, approach in HVC segmentation depend on three basic attributes such as Flag Data User, Length of Stay (LoS) and Average Revenue per User (ARPU), later on those attributes were generate to twenty eight (28) different segment combinations. Recently, Telkomsel implemented thousands of campaign every month for targeting HVC, however the result shows poor level of personalizaton that historically impact to low of average level take-up (that is 4.5%). In further impact, it can be said that within the increase of ARPU, the level of take-up on average is decreasing (Figure 1). This fact become opportunity at the same time challenges for Telkomsel in selecting the right campaign strategy for each customer, in expected to gain more revenue. Segmentation accuracy is one of the method did by company, as it can give picture of costumer profiling. However, current segmentation used by company have not met expectation between costumer behavior and pattern of HVC. Customer behavior attribute was added in this reserach to determine customer segmentation.

LITERARURE REVIEW

According to Warner (2016), contextual market can be said as identification process "the moment of truth" for each customer and personalization of individual interaction with right message and channel in certain moment. This matter needs comprehensive undertstanding of behavior, habits and choices to determine each customer needs. Context of relevance and Voice of Customer are determined as the last and mature market approach. That approach method can increase time of responds as well as Return of Investment (ROI) and customer engagement, through personally and targeted promotion message. In terms of contextual market, market segmentation and customer profiling are interconnected and completed each other that further become the base of strategic decision, alocated resource and determined prioroty for the next campaign method.

In terms of campaign management, data mining is implementation of algorithm and ETL (Extraction-Transformasi-Load) tools that used for customer behavior data analysis, customer group data identification, customer grouping within the same characteristics, and give the best solution based on analysed pattern. Data Mining is knowledge extraction from the data, through technology that combine fundamental of data science (Provost & Fawcett, 2013). According to Berry and Linoff (2004) Clustering Method is one of data mining method which used to segmented heterogeneous population become subgroup or more homogeneous group.

METHODOLOGY

Population and Sample

In this research, the population is customer of PT Telkomsel whose belong to category of HVC scattered to 28 existing segment based on ARPU grouping variable, Length of Stay, and Flag Data User with 23 million customer total of customer. Within random sampling method (Slovin formulas), it was obtained 11,187 sample. For this research it is used 280,245 data sample, which consist of data for each segment (28 segments) both *taker* and *non taker*, which means represent every population and every segments. Therefore, the amount of sample in this reserach has met the requirements criteria.

Operationalization Variable and Data Collection

Variabe used in this research consist of 3 forming variable for the previous Telkomsel HVC Customer segmentation that are Length of Stay (LoS), ARPU, danFlagData User added by 8 new variables which represent customer behavior. Those 11 variables are determined as independent variable, while dependent variable in this research is coming from product selected by customer in campaign activity that is product retrieval status or activity status (Taker and Non Taker) which affect to value of take up rate in campaign. Data were analysed by using quantitative reserach method and cross section data as the research conduct at the same time in form of customer transactional data as added HVC customer segmentation variable. Secondary source of data used by this research which come from HVC transactional data from billing data of Online Charging Systmen (OCS) that are collected and streamed to Telkomsel Big Data system.

Instrument and Measure

This study used Data Mining for making the Model, it is kind of application software SPSS Modeler in 18.0 version. Data mining were classified into classification, regression, clustering association, anomaly detection, time series, tect mining, feature selection (Kotu&Desphande, 2015). Moreover this study used clustering method with K-Means algorithm and classification with Logistic regression algorithm. K-Means was used in segmentation process of HVC Telkomsel and Logistic Regression was used as predictive method to predict amount of customer who decided to take the product and take up rate value of HVC Telkomsel. To predict appropriate model it was used logistic regression due to variable target in this study is in the form of binary (0 or 1) with variable input numerical and categorical.

Forward Stepwise was selected as variable selection method. With this method, each iteration would entered one by one input variable that is field in this study up to the last variable input, and then it was accuration measurement from each iteration until found the optimum accuration model. As can be seen each steps were visualize in Figure. 2:

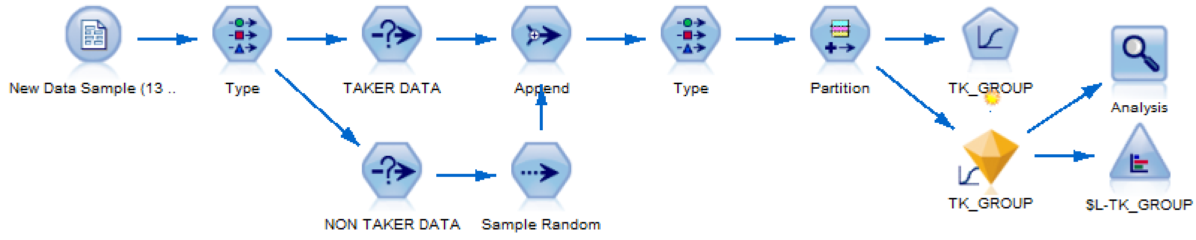


Figure 3. Stream SPSS Modeler for Prediction Model

According to (Joanne Peng, Lee, & Ingersoll, 2002)formula of Logistic Regression for predicting certain event with using more than one variable, as follows:

$$\pi = \frac{e^{a+\beta_1 X_1+\beta_2 X_2+\beta_n X_n}}{1+e^{a+\beta_1 X_1+\beta_2 X_2+\beta_n X_n}}$$

Whereas:

π : opportunity of event
α : constants
$\beta 1 \ldots \beta n$: regression coefficient for variable $X1$ ……. Xn
$X1 \ldots Xn$: input variable$X1$ ……. Xn
e : natural logarithm (2.71828)

Within the same method K-Means and Logistics Regression, authors found and measured validity of made model. In clustering with K-Means algorithm, there are silhouette value index that used to determined model validity or formed segmentation. That Index was used to measure how effective grouped observation and predict the range of one cluster to another (Wendler&Grottup, 2016). The higher silhouette value, the higher change of the model is effective to be used. Algorithm in K-Means clustering shown on Figure 3 (a) (Tan, Steinbach, & V. Kumar, 2005)

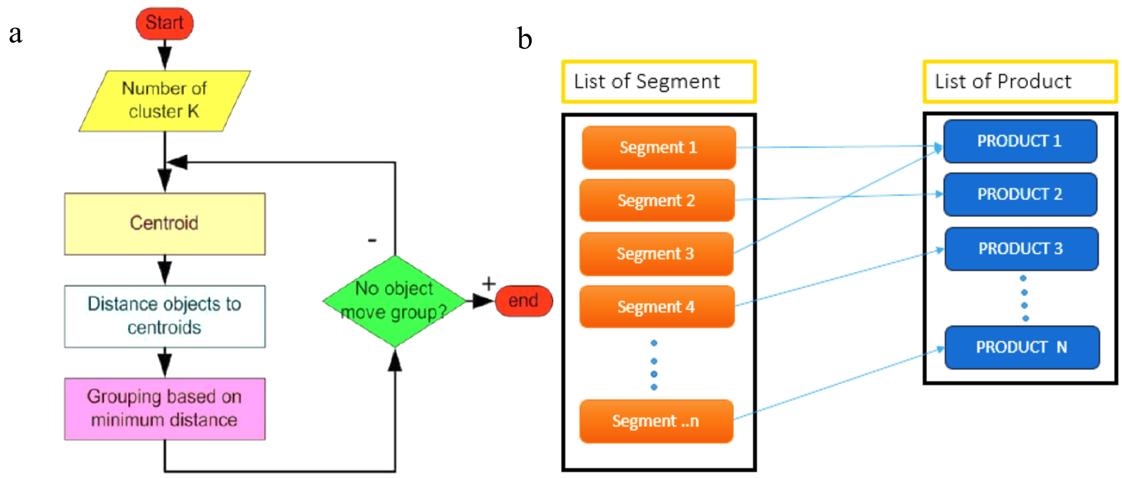


Figure 4. (a) Flowchart K-Means Clustering, (b) Metode Matchmaking Segmendan Product

Different with K-Means, Confussion matrix (Table 1) were used to measure validity of logistic regression model. Measurement was determined by level of accuracy and level of sensitivity. The higher level of accuracy and sensitivity of certain model will give effectiveness of model used.

Table 1. Confussion Matrix

		KondisiHasilPrediksi	
	Total Populasi	Taker	Non Taker
KondisiAktual	Taker	**TP** (*True Positive*)	**FN** (*False Negative*)
	Non Taker	**FP** (*False Positive*)	**TN** (*True Negative*)

Source: (Grottrup & Wendler, 2016)

True Positive Rate (TPR) atau*Sensitivity* = $\frac{TP}{TP+FN}$

Accurancy (ACC) = $\frac{TP+TN}{Total\ Populasi}$

Matchmaking (Figure 3[b]) is further step of segmentation process after the clustering done. Matchmaking approach is actually a simple method that matched variable that have been determined a certain group that later on included to segmentation from the result of clustering.

FINDINGS

Existing Segment Validation

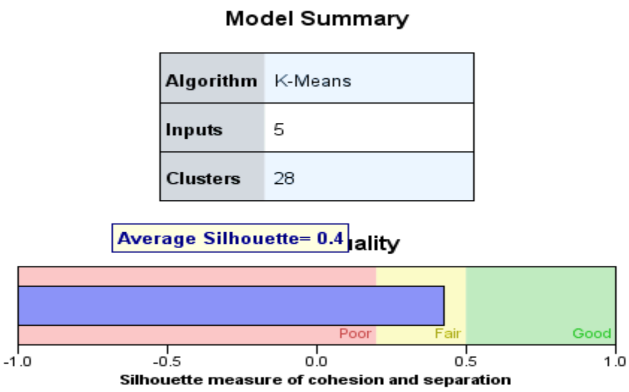


Figure 5. Existing Cluster Model Summary

To determined existing customer segmentation ofHVCTelkomsel were done by utilize 3 variable, that are Flag Data User, Length of Stay, and Average Revenue per User (ARPU). As can be seen in Figure.3, K-Means method in SPSS Modeler 18.0 versionwith 28 cluster or segment, it give result silhouette value as much as 0.4. However, this value stil indicate overlapping data or characteristcs of customer who has little differencies amongst segments.

The 5 number of cluster is considered has a good cluster quality. Cluster quality is shown by silhouette value 0.5. That silhouette index were used as tools to evaluate in clastering used K-Means. This Index also used to measure of how good grouping observation and estimate the average range from one cluster to another. Regarding to analysis result, it was found when adding 8 new variables based on customer behavior, it can increase segmentation effectivity which is shown by silhouette value index (validity model) increase by 0.1. Furthermore, analysis result show the cluster 1 has bigger amount of HVC customer as much as 29.0% from the total, followed by cluster 2 as much as 22.3%, and cluster 4 as much as 20.7%. Meanwhile, cluster 5 has smallest amount of HVC customer as much as 9.6%. In conclusion, largely of HVC being researched are come from 1, 2, and 4 cluster.

Modeling & Clustering New Segment

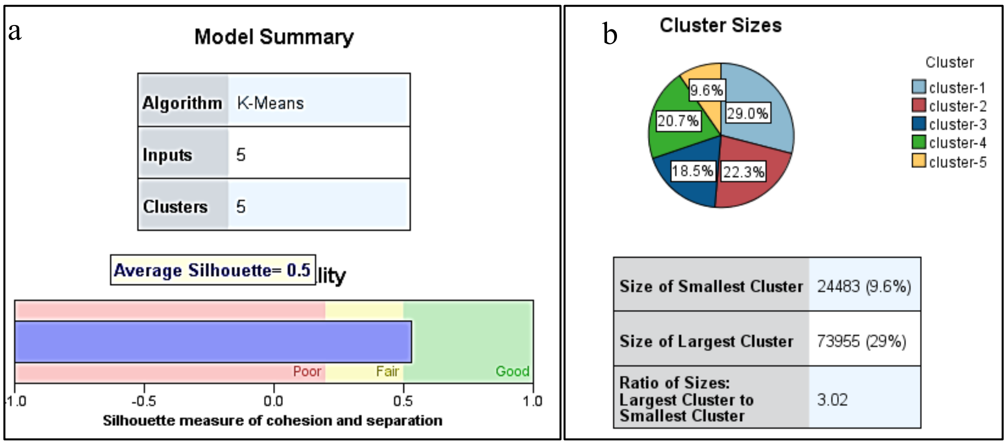


Figure 6. (a) New Clustering Model Summary, (b) New Segment Cluster Size

New Segment Cluster Profiling

Table 2. Sum-Up Profiling Cluster Result

Cluster	Cluster-1	Cluster-2	Cluster-3	Cluster-4	Cluster-5
Category	Data User Heavy on Internet with High Recharge	Loyal customer, heavy on Voice and SMS with Medium Recharge	Normal usage with High Recharge	Customer Heavy on Voice and SMS with medium Recharge	Value Customer, Heavy on internet and voice
Cluster Size	29,0%	22,3%	18,5%	20,7%	9,6%
Datauser Flag	Data User	Non Data User	Data User	Non Data User	Data User
LOS	1 - 7 Years	> 7 Years	> 7 Years	1 - 7 Years	> 7 Years
Ratio Voice Onnet	0.80	0.88	0.85	0.87	0.80
Ratio SMS Onnet	0.70	0.70	0.71	0.72	0.68
Ratio Data 4G	0.32	0.00	0.28	0.00	0.36
Total Revenue Voice	Rp57,560	Rp146,605	Rp56,549	Rp128,719	Rp145,159
Total Revenue SMS	Rp9,598	Rp21,982	Rp11,093	Rp23,146	Rp12,374
Total Revenue Data	Rp101,760	Rp93	Rp77,040	Rp219	Rp90,689
Average Amount per Recharge	Rp50,411	Rp35,877	Rp37,451	Rp28,150	Rp84,095
Days Recharge	5 Days	7 Days	5 Days	8 Days	4 Days

Segmentation of Telkomsel HVC customer devided into 5 segments with different characteristics and profiling based on group of customer behavior and group of customer buying method.

Predictive Analitic

In predictive analysis of existing segments (28 segments) and new segments (5 segments), afterconducted data cleansing it is selected sample of customer data as much 245,852, consist of Non Taker customer 234,789 (95.5%) and customer of random sampling using spss modeler 11,063 (4.5%) as Taker customer. 70% portionof training data set and 30% testing data set were used in this research.

After accurate model forming process using SPSS Modeler versi 18.0, it was given 3 iteration results. In predictive analysis **exsisting segments** (28 segments) accuration level of logistic regression model until the last third iteration reached 8.6 % and sensitivity level 30.4%. As forprediction result show from total training and testing data set **245,852** customer, it was predicted **26,995** or **10.98 %**of customer will be "***taker"***or decided to take product offered. Different with existing segment, in new segment it was given 9 iteration result. It is discovered that accuracy level of logistic regression model untul the last ninth iteration reached **85.3 % and sensitivity level 45.8%. Therefore,** from total *training* and*testing data set* as much **245,852** of customer were predicted **50,178** of customer or **20.41%** will be "***taker"***or decided to take rpoduct offered.

Predictive Model Operation

Evaluation result regarding existing segment and new segment, model in new segment showed better result from the existing. Examination were done by comparing two model within consideration of accuracy level, sensitivity model and value of taker prediction (take up rate) (Table 3). By using new segmentation model, then conducting model operation to predict prospectious customer from total customer of TelkomselHVC. As there are 23 millions of TelkomselHVC, it is predicted 4.7 millions of customer or 20.42% who will considerate to take product offered if Telkomsel perform campaign with contextual approach, that is Consumer Behavior. On teh other words, the take up rate value will be 20.41%

Table 3. Comparison&Escalation of Take up Rate Predictionand Actual

Take Up Rate	Eksisting Segments (28 Segments)	New Segments (5 Segments)	Escalation
Prediction	10.98%	20.41%	9.43%
Actual	4.5%	8.36%	3.86%

As can be seen in Table. 3, by added 8 new variables related to Consumer Behavior of Telkomsel HVC customer, it is found that prediction value of Take Up Rate increased by 9.43% or in actual value equal to 3.86% which increase from 4.5% to 8.36%.

Strategy

Accurate market approach will impact to result and effectiveness certain market method to the company revenue and target. After conducted segmentation of Telkomsel HVC customer using clustering method, the next step is formulate suitable marketing strategy for each group or cluster with matchmaking existing product and formed segmentation. Marketing strategy for each cluster or customer segment can be seen in Table 4.

Table 4. Product Offered Strategy Based on Segments Behavior with Matchmaking

Segment					Cluster	Existing Product Offering		
Datauser Flag	LoS	Behavior				Product/ Service Offer	Validity	Price
		Usage	Biggest Revenue	Recharge				
Data User	1 - 7 Years	Heavy on Internet	Internet	High Recharge (after 5 days) & nominal Rp 50.411	1	Big on Data, Medium on Voice Onnet, Small on SMS Onnet	3 Days	30000
							7 Days	65000
							15 Days	145000
							30 Days	300000
	> 7 Years	Normal Usage	Voice & Internet	High Recharge (after 5 days) & nominal Rp 37.451	3	Medium on Data, Voice Onnet& SMS Onnet	3 Days	25000
							7 Days	65000
							15 Days	130000
							30 Days	250000

Leena and Luna International, Chikusei, Japan.
(株) リナアンドルナインターナショナル, 筑西市,日本
ISSN: 2186-845X ISSN: 2186-8441 Print
www.ajmse. leena-luna.co.jp

		Value Customer, Heavy on Internet and Voice	Voice & Internet	High Recharge (after 4 days) & nominal Rp 84.095	5	Big on Data & Voice Onnet, Small on SMS Onnet	3 Days	65000
							7 Days	130000
							15 Days	260000
							30 Days	500000
Non Data User	1 - 7 Years	Heavy on Voice and SMS	Voice	Medium Recharge (after 7 days) & nominal Rp 35.877	2	Big on Voice Onnet& SMS Onnet, Small on Data	3 Days	14000
							7 Days	35000
							15 Days	65000
							30 Days	125000
	> 7 Years	Loyal customer, heavy on Voice and SMS	Voice	Medium Recharge (after 8 days) & nominal Rp 28.150	4	Big on Voice Onnet& SMS Onnet, Small on Data	3 Days	10000
							7 Days	27000
							15 Days	52000
							30 Days	100000

CONCLUSION

Based on the analysis data result and clustering with K-Means method, it can be concluded that by added 8attribute customer behavior in determined and formed customer segmentation, it is shown that clustering give information result when 8 behavior attributes added it can intensify campaign effectiveness as showed in increase of 3.86% take-up rate. It was obtained 5 customer segmentation are formed with most customer proportion in segment 1 as much as 29.0%, followed by segment 2 (22.3%) and segment 4 (20.7%). Within characteristics or each segment profiling are Segment 1 Data User Heavy on Internet with High Recharge, Segment 2 Loyal customer; heavy on Voice and SMS with Medium Recharge, Segment 3 Normal usage with High Recharge, Segment 4 Customer Heavy on Voice and SMS with medium Recharge and Segment 5 Value Customer, Heavy on Internet and Voice with High Recharge. Furthermore, suitable contextual campaign strategy for HVC based on its segment will be devided into 20 product offering (campaign) based on types of product, validity of the product and priced offered.

ACKNOWLEDGEMENTS

We are enormously grateful to all the member and team of IT Big Data and Digital Experience & Business Insight Telkomsel for their advice and providing source materials on this research. We also would like to thank to the lecturers and employees of Master Management, Economics and Business Faculty, Telkom University for their support and constructive criticism and advice of this research.

REFERENCES

[1] Berry, M. J., & Linoff, G. S. (2004). *Data mining techniques for marketing, sales, and customer relationship management.* Canada: Wiley Publishing, Inc. Indianapolis, Indiana.

[2] Peng, J., Lee, K., & Ingersoll, G. (2002). *An introduction to logistic regression analysis & reporting: An introduction to logistic regression analysis & reporting*. USA: Authors.

[3] Provost, F., & Fawcett, T. (2013). *Data science for business.* United States of America: O'Reilly Media, Inc.

[4] Tan, P. N. (2005). *Introduction to data mining.* Boston, MA, USA: Addison-Wesley Longman Publishing Co.Inc.

[5] Telkom. (2016). *Building Indonesian digital economy: 2016 annual report.* Retrieved from https://cdn.indonesia-investments.com/bedrijfsprofiel/201/Telekomunikasi-Indonesia-Telkom-Annual-Report-2016-Company-Profile-Indonesia-Investments.pdf.

[6] Telkomsel, M. D. (2017). *Monthly management report periode Juli 2017.* Indonesia: Telkomsel.

[7] Warner, R. (2016). *Let's put enterprise marketing technology into context.* Retrieved from http://blogs.forrester.com/rusty_warner/16-04-20-lets_put_enterprise_marketing_technology_into_context.

Leena and Luna International, Chikusei, Japan.
(株) リナアンドルナインターナショナル, 筑西市,日本
ISSN: 2186-845X ISSN: 2186-8441 Print
www.ajmse. leena-luna.co.jp