e-ISSN : 2549-9904 ISSN : 2549-9610

Joiv

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

Customer Relationship Management and Machine Learning technology for Identifying the Customer

Hidra Amnur[#]

[#] Information Technology Department, Politeknik Negeri Padang, Padang, Indonesia E-mail:hidraamnur@gmail.com

Abstract— Customer Relationship Management needed for the company to know their customer more closed, and make two-way communication between company and customer. in CRM solutions are multi-criteria decision-making analysis tools that do not require prior assumptions to explore the weights and performances among project risk, project management and organization performance, based on research framework of stimulus-organism response model. in this study, Machine learning with Support Vector Machine algorithm is currently for classification task due to its ability to model nonlinearities CRM Solutions. With Machine Learning and CRM, Bank X optimize their profit, with manage their more benefit customer or find a new customer or get their lost potential customer back.

Keywords- Customer Relationship Management, Machine Learning, Support Vector Machine

I. INTRODUCTION

Machine learning is the subfield of computer science that gives computers the ability to learn without being explicitly programmed (Arthur Samuel, 1959) [1]. Machine learning is closely related to (and often overlaps with) computational statistics, which also focuses in prediction-making through the use of computers. It has strong ties to mathematical optimization, which delivers methods, theory and application domains to the field. There are several algorithms Machine Learning, Support Vector Machine (SVM) is one of them. SVM developed by Boser, Guyon, Vapnik, and first presented in1992 at the Annual Workshop on Computational Learning Theory. [2]

Customer Relationship Management analysis done by measurement of the Net Present Value of the profit generated by a number of consumers at specific times. This way the Lifetime Value Customer.

$$Lifetime \ Value = \frac{\sum_{i=0}^{n} Net \ Present \ Value}{\sum_{i=0}^{n} Customers}$$

The measurements are also carried out by Regency Frequency Monetary (RFM). [3] The Pareto principle (also known as the 80/20 rule, the law of the vital few, or the principle of factor sparsity). states that, for many events, roughly 80% of the effects come from 20% of the causes. CRM has been conceptualised from five different viewpoints:

(1) Process, (2) Strategy, (3) Philosophy, (4) Capability, and/or (5) Technological tools [4]

Different definition of CRM put emphasis on different perspectives. CRM's technological perspective was stressed in [5], its knowledge management perspective was emphasized in [6] and its business re-engineering and continuous improvement perspective is presented in [7]. CRM at three levels, Strategic, Analytical and Collaborative. Strategic CRM: It is focused on development of a customercentric business culture. Product, production and selling are the three major business orientations identified by Kotler [8]. Analytical CRM: Analytical CRM builds on the foundation of customer information. Collaborative CRM: Staff members from different departments can share information collected when interacting with customers [9].

Within the field of data analytics of CRM, machine learning is a method used to devise complex models and algorithms SVM that lend themselves to prediction; in commercial use, this is known as predictive analytics.

II. MATERIAL AND METHOD

Lifetime Value become a barometer for sorting valuable customer and differentiate with other customer groups. customers with higher value became the focus of attention. in this case use the sample data on bank x for 5 years, special savings products.

No	Description	1 years	2 years	3 years	4 years	5 years
1	Customers	883	976	1016	1511	3638
2	Retention Rate	0	111%	115%	171%	412%
3	Referal Rate	0	5%	10%	15%	20%
4	Refferal Customers	0	49	102	227	728
5	The Rate For Savings (Rp 000)	6,252,000	5,618,000	9,425,000	6,300,000	11,474,000
6	Total Revenue (Rp 000)	5,520,516,000	5,757,326,400	10,533,380,000	10,947,195,000	50,090,894,400
7	Variabel Cost	40%	45%	50%	55%	60%
8	Variabel Cost (Rp 000)	2,208,206,000	2,467,426,000	4,787,900,000	5,235,625,000	25,045,447,000
9	Total Cost	9,272,206,000	2,467,426,244	4,787,900,508	5,235,616,133	25,045,450,638
10	Referral Insentif (Rp 000)	0	244	508	1,133	3,638
11	Discount Rate	1.00	1.16	1.20	1.35	1.56
12	NPV Profit (Rp 000)	(3,751,690,000)	2,836,120,825	4,787,899,577	4,230,799,161	16,054,771,642
13	accumulation of NPV profit	(3,751,690,000)	(915,569,175)	3,872,330,401	8,103,129,561	24,157,901,204
14	Lifetime Value Retention (Rp 000)	(4,249,000)	2,664,000	3,927,000,000	2,100,000	2,942,000
15	Lifetime Value Referral (Rp 000)	(4,248,799)	2,767,487	4,284,090,530	2,535,782	3,677,563
16	Different Retention Vs Referral		103,487	357,090,530	334,782	735,563

TABLE I LIFETIME VALUE RETENTION AND LIFETIME VALUE REFERRAL

TABLE II FAIRLY SPREAD REGENCY

Group	Customers	Active Constomers	Respons	Cost	Savings	Profit
1	728	727	21.21%	26,182,472	2,433,635,400	2,407,452,928
2	728	691	20.15%	24,873,966	2,312,011,000	2,287,137,034
3	728	686	20.00%	24,688,800	2,294,800,000	2,270,111,200
4	728	685	19.97%	24,651,767	2,291,357,800	2,266,706,033
5	728	640	18.67%	23,046,995	2,142,195,800	2,119,148,805
Total	3640	3429	100%	123,444,000	11,474,000,000	11,350,556,000

From the Table 1, Banks do referral to the customer, because it provides a long term good value than without no referral. At the moment, the banks do referrals for its customers. so, there are some banks perform different treatment of customers, intensive is given to consumers who benefit, such as giving a gift or bonus. This is to maintain profitable customer so that they are not disappointed. Than do RFM (Regency, Frequency and Monetary), Regency is a measure of consumer value by looking at the behaviour of the last savings. divide the consumers with the same size, 5 groups (20% of total consumer). Calculate cost, savings, profit from each group (Table II).

III. RESULTS AND DISCUSSION

The next question, which one benefit consumers, The data retrieved from the table II and then entered the stage of processing machine learning with SVM algorithm.

1. Margins: Intuition

This section will give the intuitions about margins and about the "confidence" of predictions. Consider Customers regression, where the probability $p(y = 1|x; \theta)$ is modeled by $h\theta(x) = g(\theta T x)$. We would then predict "1" on an input x if and only if $h\theta(x) \ge 0.5$, or equivalently, if and only if θ T x \geq 0. Consider a positive training example (y = 1). The larger θ T x is, the larger also is $h\theta(x) = p(y = x)$ 1|x;w, b), and thus also the higher our degree of "confidence" that the label is 1. Thus, informally we can think of our prediction as being a very confident one that y = 1 if θ T x \gg 0. Similarly, we think of logistic regression as making a very confident prediction of y = 0, if $\theta T x \ll 0$. Given a training set, again informally it seems that we'd have found a good fit to the training data if we can find θ so that θ T x (i) \gg 0 whenever y (i) = 1, and θ T x (i) \ll 0 whenever y(i) = 0, since this would reflect a very confident (and correct) set of classifications for all the training examples. This seems to be a nice goal to aim for, and we'll soon formalize this idea using the notion of functional margins.

For a different type of intuition, consider the following figure, in which x's represent positive training examples, o's denote negative training examples, a decision boundary (this is the line given by the equation θ T x = 0, and is also called the separating hyperplane) is also shown, and three points have also been labeled A, B and C. x for benefit customers, o for not benefit customers



Fig. 1 Intuition of customer

2. Notation

To make SVMs easier, first need to introduce a new notation for talking about classification. We will be considering a linear classifier for a binary classification problem with labels y and features x. From now, we'll use $y \in \{-1, 1\}$ (instead of $\{0, 1\}$) to denote the class labels. Also, rather than parameterizing our linear classifier with the vector θ , we will use parameters w, b, and write our classifier as

$$h_{w,b}(x) = g(w^T x + b)$$

Here, g(z) = 1 if $z \ge 0$, and g(z) = -1 otherwise. This "w, b" notation allows us to explicitly treat the intercept term b separately from the other parameters. Thus, b takes the role of what was previously $\theta 0$, and w takes the role of $[\theta 1 \dots \theta n]$ T. Note also that, from our definition of g above, our classifier will directly predict either 1 or -1 (cf. the perceptron algorithm), without first going through the intermediate step of estimating the probability of y being 1.

3. Functional and geometric margins

Formalize the notions of the functional and geometric margins. Given a training example (x (i), y(i)), we define the functional margin of (w, b) with respect to the training example

$$\gamma^i = \gamma^i (w^T x + b)$$

Note that if y (i) = 1, then for the functional margin to be large (i.e., for our prediction to be confident and correct), we need w T x + b to be a large positive number. Conversely, if y (i) = -1, then for the functional margin to be large, we need w T x + b to be a large negative number. Moreover, if y (i) ($w^T x + b$) > 0, then our prediction on this example is correct. (Check this yourself.) Hence, a large functional margin represents a confident and a correct prediction. For a linear classifier with the choice of g given above (taking values in $\{-1, 1\}$), there's one property of the functional margin that makes it not a very good measure of confidence.

Given a training set $S = \{(x (i), y(i)); i = 1, ..., m\}$, we also define the function margin of (w, b) with respect to S to be the smallest of the functional margins of the individual training examples. Denoted by γ , this can therefore be written: $\gamma^{2} = \min i=1,...,m \gamma^{2}$ (i) . geometric margins, Consider the picture below



Fig. 2 Geometric Margin of customer

The decision boundary corresponding to (w, b) is shown, along with the vector w. Note that w is orthogonal (at 90°) to the separating hyperplane. (You should convince yourself that this must be the case.) Consider the point at A, which represents the input x (i) of some training example with label y (i) = 1. Its distance to the decision boundary, γ (i), is given by the line segment AB. Solving for γ (i) yields

$$\gamma^{(i)} = y^{(i)} \left(\left(\frac{w}{||w||} \right)^T x^{(i)} + \frac{b}{||w||} \right).$$

Note that if ||w|| = 1, then the functional margin equals the geometric margin—this thus gives us a way of relating these two different notions of margin. Also, the geometric margin is invariant to rescaling of the parameters;

if replace w with 2w and b with 2b, then the geometric margin does not change. This will in fact come in handy later. Specifically, because of this invariance to the scaling of the parameters, when trying to fit w and b

to training data, we can impose an arbitrary scaling constraint on w without changing anything important; for instance, we can demand that ||w|| = 1, or |w1| = 5, or |w1+b| + |w| = 2, and any of these can be satisfied simply by rescaling w and b.

4. The optimal margin classifier

Given a training set, it seems from our previous discussion that a natural desideratum is to try to find a decision boundary that maximizes the (geometric) margin, since this would reflect a very confident set of predictions on the training set and a good "fit" to the training data. Specifically, this will result in a classifier that separates the positive and the negative training examples with a "gap" (geometric margin).

$$\max_{\gamma,w,b} \quad \gamma \\ \text{s.t.} \quad y^{(i)}(w^T x^{(i)} + b) \ge \gamma, \quad i = 1, \dots, m \\ ||w|| = 1.$$

For now, we will assume that we are given a training set that is linearly separable; i.e., that it is possible to separate the positive and negative examples using some separating hyperplane. How we we find the one that achieves the maximum geometric margin? We can pose the following optimization problem:



Fig. 3 Support Vector Machine Result

From the positive and the negative on SVM algorithm (fig.3) consumers can be divided into three segmentation CRM :

Prioritas 1:consumers high-value, (the positive, target=yes) Prioritas 2:consumers medium-value, (the margin)

Prioritas 3:consumers low-value, (the negative, target=no)

IV. CONCLUSIONS

The customers must be identified from the current customers with whom collaboration and relationships of lifetime value and RFM, CRM Method can be set up. It is important to consider that building relationships with inappropriate customers is one of the main reasons for failure in projects of systems related to the customer. The identification of customers can be accomplished by Big Data techniques of advanced analytics, since they help to discover trends, patterns and other insights, applied to historical information from past interactions with customers and current information of customers.

ACKNOWLEDGMENT

The authors would like to acknowledge to Andrew Y. Ng, Stanford University, Stanford, CA, United States. His Machine Learning Course help me to find a solution for wrote this paper.

REFERENCES

- [1] Samuel, Arthur L. (1959). "Some studies in machine learning using the game of checkers". IBM Journal of research and development.
- [2] [5] T. Hastie et al., and P. W. Daly, The Elements of Statistical Learning : Datamining, Interference, and Prediction, Springer Verlag, 2001.
- [3] Chan, Syafruddin, Relationship Management, PT Gramedia Pustaka Utama, Jakarta, 2003
- [4] Zablah AR, Bellenger DN, Johnston WJ (2004) An evaluation of divergent perspectives on customer relationship management: towards a com-mon understanding of an emerging phenomenon. Ind Mark Manage.
- [5] D. Peppers, M. Rogers, A new marketing paradigm, Planning Review, 23(2) (1995) 1418.
- [6] A. Massey, M. Montoya-Weiss, K. Holcom, Re-engineering the customer relationship:leveraging knowledge assets at IBM, Decision Support Systems, 32 (2001) 155–170.
- [7] J. Anton, Customer Relationship Management: Making Hard Decisions with Soft Numbers, Prentice Hall, Englewood Cliffs, New Jersey, 1996.
- [8] P. Kotler, Marketing Management: The Millennium Edition. Englewood Cliffs, NJ:Prentice-Hall International, 2000.
- J. Edwards, Get It Together with Collaborative CRM, insideCRM, 2007. Tippit. http://www.insidecrm.com/features/collaborativecrm-112907
- [10] Lambert DM (2010) Customer relationship management as a business pro-cess. J Bus Ind Mark
- [11] Hu Y-H, Huang T, Kao Y-H (2013) Knowledge discovery of weighted RFM sequential patterns from customer sequence databases. J Syst Softw