International Journal of Advanced Trends in Computer Science and Engineering

Bank Marketing Data Mining using CRISP-DM Approach

Tuga Mauritsius[1], Annisa Safira Braza[2], Fransisca[3]

[1]Information Systems Management Department, BINUS Graduate Program-Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia, tuga.mauritsius@binus.ac.id

[2]Information Systems Management Department, BINUS Graduate Program-Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia, annisa.braza@binus.ac.id

[3]Information Systems Management Department, BINUS Graduate Program-Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia, fransisca001@binus.ac.id

ABSTRACT

We live in a world where vast amounts of data are collected daily. At the bank, a large amount of data is recorded about their customers. This data is commonly used to create, maintain relationships and direct connections with customers to target them individually in the sale of certain products or banking offerings. Analyzing such data is an important need. This paper describes data mining approaches aim to build a predictive model that labels data into a predefined class. To define the processes and tasks that you must do to develop a successful Data Mining project using CRISP-DM Framework. This study will use Multilayer Perceptron and logistic regression as data mining technique. In this paper, the resulting model allows overfitting, so it can be avoided using cross validation. the model that provides the greatest average benefit is the Multilayer perceptron model with a 70% percentage split.

Key words : CRISP-DM, Data Mining, Logistic Regression, Multilayer Perceptron.

1. INTRODUCTION

A bank has a lot of customer data. These data can be used to analyze and classify customers into several segments to define appropriate targets in product offerings and other bank offerings. After customers are grouped into several segments, then they will get offers via telephone, email, mail, or others based on their segment type so that the campaign is conducted according to personal trends, this is called direct marketing. Ou et al., 2003 said that Not surprisingly, most companies including banks have switched from a mass marketing strategy to a direct marketing strategy because it has proven more effective and also efficient in offering other products and services personally. Direct marketing is very effective and widely used strategy of contacting customers or potential customers rather than having an indirect channel especially for banking sector (Elsalamony and Elsayad, 2013;

Elsalamony, 2014). Now marketing staff do not need to waste time and energy sending e-mails to nonpotential customers or the company can reduce the expense of calling nonpotential customers.

To identify prospective customers who will be offered for new products in terms of direct marketing usually use a method called data mining. From stored customer history data, a predictive model can be made that a customer is going to respond to the promotion or an offer (Usama et al., 1996). Wisaeng, 2013 said that data mining is the process of extracting previously unknown information from a large datasets and analyzing hidden patterns of data according to different perspectives for categorization into useful information, which is collected and assembled in common areas, such as data warehouses, for efficient analysis, data mining algorithms, facilitating business decision making and other information requirements to ultimately cut costs and increase revenue. Most commonly used techniques in data mining are: artificial neural networks, genetic algorithms, rule induction, nearest neighbor method and memory based reasoning, logistic regression, discriminate analysis and decision trees. These techniques and machine learning algorithms can be used for marketing and campaigning.

In this paper, data mining approaches aim to build a predictive model that labels data into a predefined class ("yes" or "no"). Improve the efficiency of the marketing campaigns and helping the decision makers to choose the potential customers using portuguese marketing campaign dataset related with bank deposit subscription. Moro et al., (2011) processed the same dataset using the Support Vector Machine (SVM) model within CRISP-DM Framework. This study will use neural network and logistic regression as data mining technique.

2. LITERATURE REVIEW

2.1 Data Mining

Data mining is a process of extracting or extracting data that has not been known before, but can be understood and useful

from large databases and is used to make a very important business decision (Connoly and Begg, 2015). Hurwitz et al. (2013) explained that data mining includes the activity of investigating and analyzing large amounts of data to look for patterns in the data. In general, the purpose of data mining is to group and predict data. In grouping, the goal is to sort the data into small groups. Whereas in prediction the aim is to predict the value of a continuous variable.

2.2. Classification

Data mining has several techniques based on the tasks that can be done, one of them is Classification. This technique includes the process of finding a set of models or functions that explain and differentiate classes of data or concepts, for the purpose of predicting classes from objects whose class labels are unknown in the data (Chen et al., 2015). Classification is a form of data analysis that extracts a model where the model explains important data classes. These models are referred to as classifiers, predicting categorical class labels (discrete, unordered). For example, classification models can be built to categorize bank loan applications whether they are safe or risky.

In the application of data mining, classification is very useful to be applied in targeted marketing, fraud detection, medical diagnosis, and many more .(Han et al., 2012). While Moin and Ahmed (2012) explained that classification is the most commonly used data mining technique by classifying each item from a data set into predefined groups of sets.

2.3. Logistic Regression

In Logistic Regression model, predictors of the dependent variable are considered independent variable. The relationship that occurs between the independent variable and the dependent variable is nonlinear. Independent variable can be measured in various scales, namely on a nominal, ordinal, interval or ratio scale. Unlike the dependent variable which can only be measured in binary format. (Tsangaratos & Ilia, 2016)

$$\text{logit}(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i + e$$

Equation 1: Logistic Regression Model (Budimir et al., 2015)

y = dependent variable

x_i = i-th explanatory variable

β_0 = constant

β_i = i-th regression coefficient

e = error

The probability (p) of the occurrence of y is :

$$p = \frac{exp^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i)}}{1 + exp^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i)}}$$

Equation 2: Probability the occurrence of y (Budimir et al., 2015)

2.4. Multilayer Perceptron (MLP)

MultilayerPerceptron is a set of interconnected input / output units where each connection has a weight associated with it. In the learning phase, the network adjusts the weights to be able to predict the label class of input tuples accurately (Han et al.,2012).

2.4.1 Models of a Neuron

A Neuron is an fundamental information processing units for neural network operations (Haykin, 2009). Inside a Neuron consists of several main parts: :

● A set of synapses, each of which has its own weight or strength. Specifically, the xj signal at the synaptic input j connected to neuron k is multiplied by the synaptic weight of time.

● Adders to add up the input signal with the synaptic weight of each neuron and are also biased.

● Activation function to limit the output amplitude of neurons. The activation function is also called the squashing function, in this case limiting the range of amplitudes allowed from the output signal to several values.
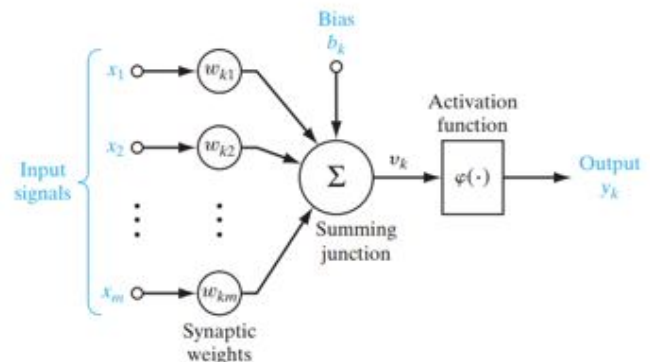

Figure 1: Nonlinear model of a neuron, labeled k (Haykin, 2009)

2.4.2 Activation Function

The activation functions of the MLP algorithm is to change the data into a certain range (MathWorks, 2002). Some commonly used functions are:

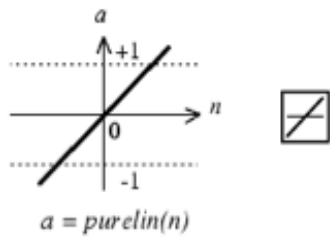● Linear: The output activity is proportional to the number of input weights.

Figure 2: Linear Activation Function (MathWorks, 2002)

● Hard-limit: Output set one of two settings depending on whether the number of inputs. The result output is 0 or 1.
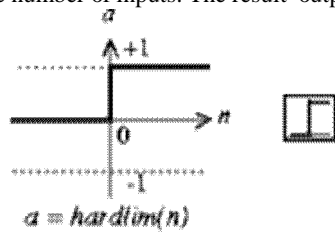

Figure 3: Hard-limit Activation Function (MathWorks, 2002)

● Sigmoid: Output continuously changes but is not linear. This unit contains a greater similarity to the digital network than the linear and threshold units, but all three must be considered.
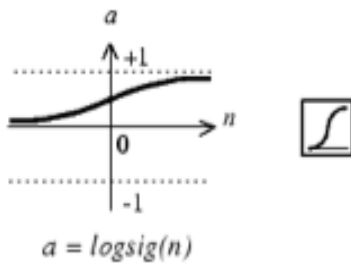

Figure 4: Sigmoid Activation Function (MathWorks, 2002)
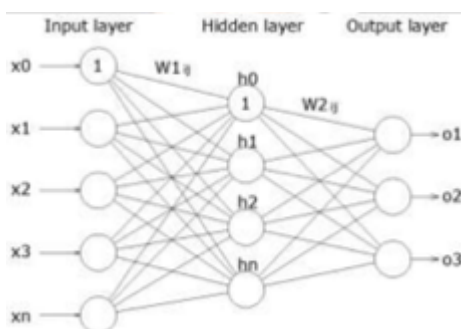
2.4.3 Multilayer Perceptron Architecture

Figure 5: MLP Architecture (Kukreja et al., 2016)

Components in MLP Architecture consist of:

● nput layer: Accept the value of an input.
● idden layer (s): Set the layer between the input layer or output layer. The number of hidden layers varies, it can be one or more.

● utput layer: Usually consists of one neuron, but does not rule out more than one possibility.

● eight: Weight given between connections, represented by an arrow line (W).

● ias: The value of external factors that affect processing (b)

Mathematically, a neuron k can be described as:

$$u_k = \sum_{j=1}^{m} w_{kj} x_j$$

Equation 3: Neuron K Equation

So that the equation will produce:

$$y_k = \varphi(u_k + b_k)$$

Equation 4: Neuron K in Activation Function

Where φ is an activation function.

2.5 Cross Validation & Percentage Split

Cross-validation is a process evaluating algorithms through cross-validation, using values Folds entered (Sartika & Sensuse, 2017). To estimate variability over a linear regression model, then can be done repeatedly sampling different from the data training, then fitting the linear regression model for each new sample, and then check which one produces the model better. Such an approach enables us to obtain information that will not be available when the model uses sample data the original and only one-time training (James et al.,2013).

While Percentage Split Percentage split is a deep process evaluate how well the algorithm is able to predict certain percentage of data. Data set will be divided into 2, data training and test data (Sartika & Sensuse, 2017).

2.6 Confusion Matrix

Confusion Matrix is a table data displayed in a few rows that are predicted to be correct and incorrect by the classification model. This table is needed to measure the performance of a classification model (Swastina L., 2013). In (Tayefi, et al., 2017) research, it is also explained that Confusion matrix is a method used to evaluate the performance of decision tree classifications.

This helps to find out whether classification algorithms are generally mislabelling with one another. By using the Confusion Matrix, tuples from different classes can be

identified (Han, Kamber, & Pei, 2012). In making a confusion matrix table there are four things that must be known:

1. True Positive (TP): the correct number of positive tuples from the decision tree classification results.
2. True Negatives (TN): the number of negative tuples that have an incorrect value from the decision tree classification results.
3. False Positive (FP): the number of positive tuples that have an incorrect value.
4. False Positive (FP): the number of negative tuples that have an incorrect value.

2.7 ROC Curve & AUC

ROC (Receiver Operating Characteristic) is a two-dimensional graph with false positives as horizontal lines and true positive to measure the difference in the performance of the method used The ROC curve is a technique for visualizing and testing the performance of classifications based on their performance. A better classification model is one that has a larger ROC curve. The AUC (Area Under the ROC Curve) accuracy performance can be classified into five groups namely (Gorunescu, 2011):

Table 1 : AUC Accuracy Classification (Gorunescu,2011)

Value	Accuracy Indicator
0.90 - 1.00	Excellent classification
0.80 - 0.90	Good classification
0.70 - 0.80	Fair classification
0.60 - 0.70	Poor classification
0.50 - 0.60	Failure

2. 8 Related Work

One of the previous studies conducted by (Moro et al., 2015) regarding the application of data mining in bank marketing. The study processed the same dataset using the Support Vector Machine (SVM), Naive Bayes, and Decision Tree model within CRISP-DM Framework. The results of this study indicate that the Support Vector Machine (SVM) model is the best model because it achieved high predictive performances.

The other study that also use the same dataset used in this paper is conducted by (Parlar & Acaravci, 2017). The study used Information Gain and Chi-square methods to select important features on dataset and compared the dataset with Naive Bayes algorithm. The results for this study is between Information Gain and Chi-square they are both resulting very close performances, although they are different for first five highest ranked features.

In previous study conducted by (Ravi et al., 2015) used dataset from Indian Banking Loan to predict the payment behavior by using data mining. The algorithms used in this study are Logistic Regression, Decision Tree, and Neural Network. The results indicate that Logistic Regression model achieved the highest performances compared to the other model (Neural Network and Decision Tree).

Other study conducted by (Dalvi et al., 2016) that used dataset customer churn in Telecom industry. The dataset in this study are evaluated using Logistic Regression and Decision Tree. The result of this study is shown that the best algorithm used in predicting customer churn is not the one with best statistical precision but the one that provide best insights to further prevent churn behavior.

3. RESEARCH METHODOLOGY

In this paper, we use the CRISP-DM Framework. CRISP-DM has been the most favored methodology in data mining domain . Therefore, we have chosen it as our reference model. By using CRISP-DM, we can find interesting patterns from within the data that we want to run. Where the data will be processed through the phases of the existing phase of the business understanding phase, understanding data, data preparation, modeling, evaluation and finally deployment. With this phase, it is expected that the results of this study will get the most appropriate modeling in the data mining process, so that the information generated is more efficient. The six phases of CRISP-DM are shown in Fig.6 and described briefly as follow.
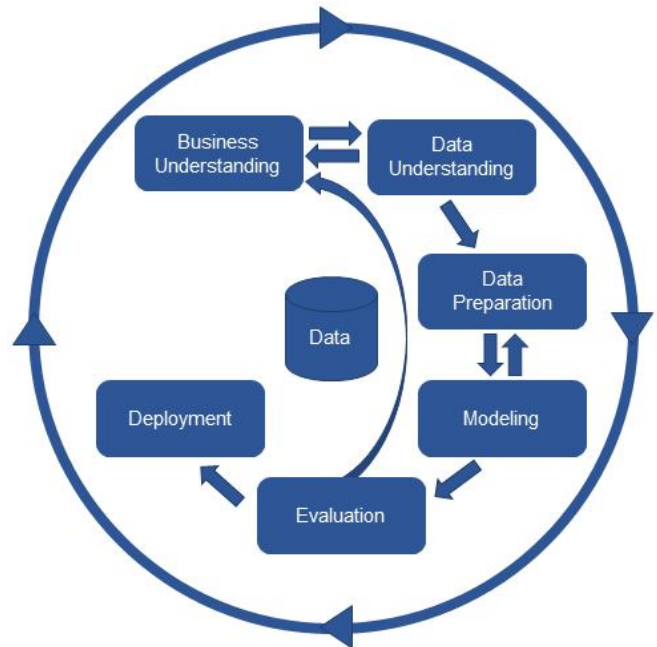

Figure 6: CRISP-DM Approach

1) Business Understanding

In this paper, data mining approaches aim to build a predictive model that labels data into a predefined class ("yes" or "no"). Improve the efficiency of the marketing campaigns and helping the decision makers to choose the potential customers using portuguese marketing campaign dataset related with bank deposit subscription.

2) Data Understanding

Bank Marketing data is obtained from the site kaggle.com. The type of data in this paper is supervised. Because, there are variables that are intended as the purpose of grouping data into existing data. Data samples that are entitled to receive 45211 instances. The algorithm used is logistical and multilayer-perceptron. And there are 16 features and 1 class (Y) . Data can be seen from the following table:

Table 2: Metadata

Name	Description and Values
Age	Age at the contact (Numeric)
Job	Type of Job: "Categorical: "Administrative, "unknown," "unemployed," "management," "housemaid," "entrepreneur"
Marital	Marital status: "Married", "Divorce","Single"
Education	Categorical: "Unknown", "secondary," "primary", "tertiary"
Default	Has credit in default? (binary: "Yes" , "No")
Balance	Average yearly balance, in euro (numeric)
Housing	Has housing loan? (binary: "Yes" , "No")
Contact	Contact communication type: "Categorical: "Unknown", "telephone", "cellular"
Loan	Has personal Loan? (binary: "Yes" , "No")
Day	Last contact day of the month (numeric)
Month	Last contact month of year (categorical: "Jan", "Feb", ..."Nov", "Dec")
Duration	Lasto contact duration, in seconds (numeric)
Campaign	Number of contact performed during this campaign and for this client (numeric, includes contact)
pdays	Number of days that passed by after the client was last contacted from previous campaign (numeric, -1 means client was not previously contacted)
previous	Number of contact performed before this campaign and for this client (numeric)
poutcome	Outcome of the previous marketing campaign (categorical: "Unknown", "other", "failure", "success")
y	Has the client subscribed a term deposit? ("Yes", "No")

3) Data Preparation

In the data preparation phase, the data is prepared for the training process. The data itself consists of 45211 instances and consists of nominal and numerical. First, using Weka, it makes the 's provided filter. First we randomize the instances to shuffle the order of instances passed through it. Next, we normalize all the numeric values in the given dataset so the resulting values are in [0,1].

4) Modeling

The algorithm used in this training data is Logistics and Multilayer Perceptron. By using the Logistics algorithm, the results of the training data can produce the following values:

- **Can calculate opportunities.**
  Based on the objectives of the training data in this study, the Bank Marketing wants to calculate the opportunities of customers who want to subscribe to deposits, if it is known how big the opportunities are,

the customers will be contacted by the Bank Marketing and will be asked questions about the variable characteristics of prospective customers who want to subscribe to deposits.

- **Able to see data characteristics**
  In this training data the Bank Marketing wants to see the type of customer. First, customer who wants to subscribe to deposits and customers who do not want to subscribe to deposits.

- **Can find out the influencing factors of both groups**
  By using Logistics, Bank Marketing can find out what are the factors of the two customer's characteristics.

Whereas the purpose of using the Multilayer Perceptron algorithm is to search according to the level of accuracy. MultiLayer Perceptron is able to detect or analyze problems that are very complex In the multilayer perceptron, the number of hidden layers is used according to the formula:

$$Hidden\ layer = \frac{number\ of\ atribute + number\ of\ class}{2}$$

Equation 5: Hidden Layer Fornula

So there are 8 hidden layers and there are also 26 units in it. To determine, carried out trial and error. Here is the simplified architecture:
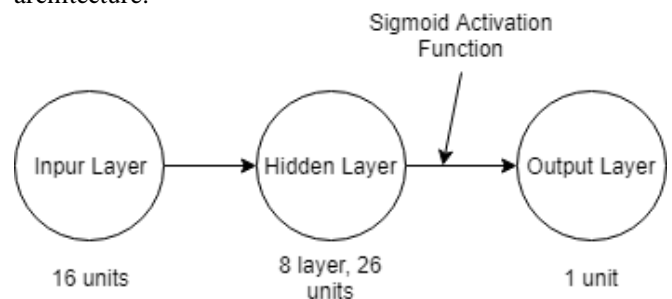

Figure 7: Multilayer Perceptron Architecture

5) Evaluation

At the evaluation phase, based on the algorithm (Logistic and MultilayerPerceptron) and the selected test options (Cross Validation and Percentage Split), a comparison is made in order to choose which is most appropriate. This is assessed from the results of correctly classified, confusion matrix, AUC, and expected profit.

6) Deployment

After conducting training data from the phases above, in this deployment phase the results of the training data can be

predicted according to the research objectives. The purpose of this study is to predict customers who want to subscribe to bank deposits, and customers who do not want to subscribe to bank deposits. So the prediction results become more objective.

4. EXPERIMENT AND RESULT

To run this training data there are laptop specifications used as follows:

Table 3: Laptop Spesification

Laptop Spesification	
Screen	14 Inch Res 1366 x 768 piksel, 45% NTSC with 178° wide-viewing angle, slim bezel, Asus Splendid.
Processor	Intel Core i5-8250U quad-core 1,6GHz TurboBoost 3,4GHz.
Graphic Type	Nvidia GeForce MX130 VRAM 2GB GDDR5.
Memory	RAM 4GB DDR4 2133MHz / 1TB HDD 5400RPM.
Operation System	Windows 10 Home
Tools/ Software	Weka Explorer Version 3.8.3

The data set that has been prepared, conducted several experiments using logistic and multilayer perceptron algorithms to be evaluated.

Table 4: Result

| Algorithm	Test option	Correctly classified	AUC	Time to build model (seconds)	Time to run
Logistic	Cross validation (10)	90.18%	0.907	5.72	52 seconds
MultilayerPerceptron	Cross validation (10)	89.40%	0.886	508.7	1h 8m 25s
Logistic	Percentage Split 66%	90.66%	0.911	0.03	8seconds
	Percentage Split 70%	90.58%	0.91	0.03	9seconds
	Percentage Split 75%	90.64%	0.909	0.03	10seconds
	Percentage Split 80%	90.51%	0.908	0.02	10seconds
MultilayerPerceptron	Percentage Split 66%	89.40%	0.884	467.29	12m 33s
	Percentage Split 70%	89.80%	0.879	0.16	12m 43s
	Percentage Split 75%	90.07%	0.893	0.13	12m 45s
	Percentage Split 80%	90.39%	0.889	0.12	12m 35s
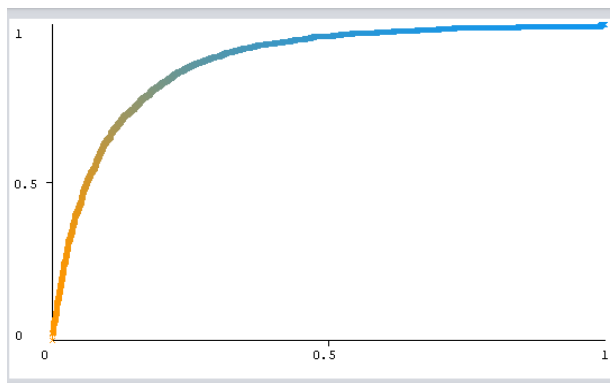
Figure 8: AUC for Highest Expected Profit (Multilayer Perceptron, 70% Percentage split
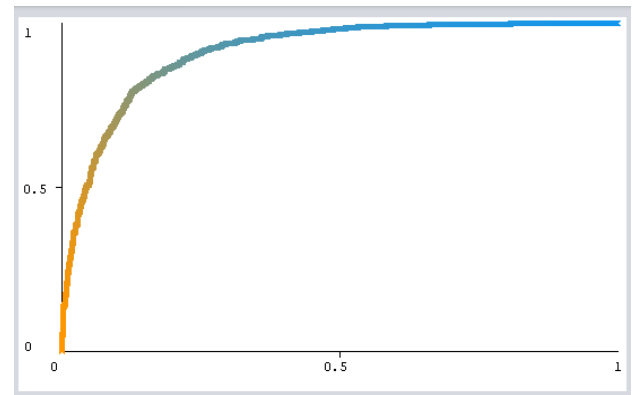

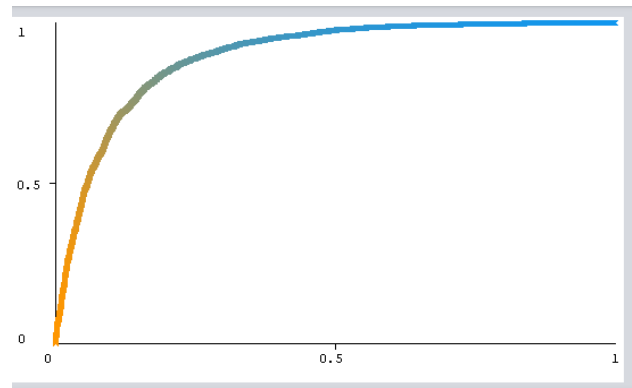**Figure 9:** AUC for Highest Logistic Regression (66% percentage split)

Figure 10: AUC for Highest Multilayer Perceptron (75% percentage split)

Based on the results obtained, Logistics with a percentage split of 66% has correctly classified numbers and the highest AUC, which is 90.66% (Correctly classified) and 0.911 (AUC). However, the resulting model allows overfitting, which can be avoided using cross validation. Because of this, the time to build the model and the time to test cross validation is longer than the percentage split.

Because this data aims to provide promotion efficiency, the test results are also evaluated with expected profit. The formula is as follows:

$$Expected\ profit = p(p).[p(Y|p).b(Y,p) + p(N|p).b(N|p)] + p(n)[p(N|n).b(N|n) + p(Y|n).b(Y,n)]$$

Equation 4: Expected Profit

Expected profit illustrates if using this model, the profit can be obtained per customer. Where p is actual positive, n is actual negative, Y is predicted positive, N is negative predicted. p is also a probability and b is a benefit. This

2327

expected profit calculation should be calculated based on actual costs and benefits. This calculation is only based on the assumption of getting a profit of $ 100 if the respondent who is called purchases the product and costs $ 1 to call. Here are the results of the calculation :

Table 5: Expected Profit

It can be seen from the results that the model that provides the greatest average profit is the Multilayer perceptron model with a percentage split of 70%, which is $ 5.56.

Here's the comparison between the previous work:

Table 6: Comparison with previous study

Author	Algorithm	ROC	Result
(Moro et al., 2015)	-Support Vector Machine (SVM) -Naive Bayes -Decision Tree	-0.938 -0.870 -0.868	Support Vector Machine (SVM)
(Ravi et al., 2015)	-Logistic Regression -Decision Tree -Neural Network		Logistic Regression
This Study	-Logistic Regression -Multilayer Perceptron	-0.911 -0.893	Multilayer Perceptron

5. CONCLUSION

The results of this study illustrate the data set that was prepared by the experiment using the logistical and multilayer perceptron algorithm to be evaluated. The methodology used in this study is CRISP-DM. In this paper, data mining approaches aim to build a predictive model that labels data into a predefined class ("yes" or "no"). Data samples that are titled to receive 45211 instances. This is assessed from the results of correctly classified, confusion matrix, AUC, and expected profit.

From the calculation results it is proven that, in the multilayer perceptron there are 8 hidden layers and there are also 26

units in it. And Based on the results obtained, Logistics with a 66% split percentage has correctly classified numbers and the highest AUC, namely 90.66% (Correctly classified) and 0.911 (AUC). However, the resulting model allows overfitting, so it can be avoided using cross validation. Therefore, the time to build the model and the time to test cross validation is longer than the percentage divided. Because this data aims to provide promotion efficiency, so the test results can also be evaluated with expected income.

The expected benefits can be explained if using this model, so the benefits can be obtained per customer. Calculation of expected profits must be calculated based on actual costs and benefits. This calculation is only based on the assumption of getting a profit of $ 100, where respondents are called to buy the product and the costs incurred to call are $ 1. The conclusion is that the model that provides the greatest average benefit is the Multilayer perceptron model with a 70% percentage split, which is $ 5.56.

REFERENCES

1. Elsalamony, H.A.. **Bank direct marketing analysis of data mining techniques**. International Journal of Computer Applications, 85(7), pp.12-22, 2014. https://doi.org/10.5120/14852-3218
2. Elsayyad, A.M.A. **Bank Direct Marketing Based on Neural Network**, 2013.
3. Moro, S., Laureano, R. and Cortez, P. **Using data mining for bank direct marketing: An application of the crisp-dm methodology**. In Proceedings of European Simulation and Modelling Conference-ESM'2011 (pp. 117-121). EUROSIS-ETI, 2011.
4. Ou, C., Liu, C., Huang, J. and Zhong, N. **On data mining for direct marketing**. In International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing (pp. 491-498). Springer, Berlin, Heidelberg, May, 2003. https://doi.org/10.1007/3-540-39205-X_83
5. Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. **From data mining to knowledge discovery in *databases***. AI magazine, 17(3), pp.37-37, 1996.
6. Wisaeng, K. **A comparison of different classification techniques for bank direct marketing**. International Journal of Soft Computing and Engineering (IJSCE), *3*(4), pp.116-119, 2013.
7. Swastina, L. **Penerapan Algoritma C4. 5 Untuk Penentuan Jurusan Mahasiswa,** 2013.
8. Tayefi, M., Tajfard, M., Saffar, S., Hanachi, P., Amirabadizadeh, A. R., Esmaeily, H., … Ghayour-Mobarhan, M. **hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm**. Computer Methods and Programs in Biomedicine, *141*, 105–109, 2017. https://doi.org/10.1016/j.cmpb.2017.02.001

9. Han, J., Kamber, M., & Pei, J. **Data mining concepts and techniques third edition**. The Morgan Kaufmann Series in Data Management Systems, 83–124, 2011.

10. Tsangaratos, P., & Ilia, I. **Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size**. Catena, 145, 164-179, 2016.

11. Haykin, S. **Neural Network and Learning Machines Third Edition. Pearson**, 2009.

12. MathWorks. **Neural Network Toolbox User's Guide. MathWorks**, 2002.

13. Kukreja, H., C S, S., & S, K. **An Introduction To Artificial Neural Network**. International Journal Of Advance Research And Innovative Ideas In Education, 27-30, 2016.

14. Hurwitz, J. S., Nugent, A., Halper, F., & Kaufman, M. **Big data for dummies**. John Wiley & Sons, 2013.

15. Moin, K. I., & Ahmed, Q. B. International Journal of Engineering Research and Applications (IJERA). **Page Use of Data Mining in Banking**, 2(2), 738–742, 2012.

16. Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A. V, & Rong, X. **Data mining for the internet of things: literature review and challenges**. International Journal of Distributed Sensor Networks, 11(8), 431047, 2015. https://doi.org/10.1155/2015/431047

17. Budimir, M. E. A., Atkinson, P. M., & Lewis, H. G. **A systematic review of landslide probability mapping using logistic regression. Landslides**, 12(3), 419-436, 2015.

18. Conolly, T., & Begg, C. E. **Database systems: a practical approach to design, implementation, and management**. Pearson, 2015.

19. Sartika, D., & Sensuse, D. I. **Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian. Jurnal Teknik Informatika Dan Sistem Informasi**, 3(2), 151–161, 2017.

20. James, G., Witten, D., Hastie, T., & Tibshirani, R. **An introduction to statistical learning** (Vol. 112). Springer, 2013. https://doi.org/10.1007/978-1-4614-7138-7

21. Gorunescu, F. **Data Mining: Concepts, models and techniques** (Vol. 12). Springer Science & Business Media, 2013.

22. Parlar, T., & ACARAVCI, S. K. **Using Data Mining Techniques for Detecting the Important Features of the Bank Direct Marketing Data**. International Journal of Economics and Financial Issues, 7(2), 692-696, 2017.

23. Ravi, V., Koparkar, S., Raju, N. P., & Sridher, S. **Improving retail banking loans recovery via data mining techniques: a case study from Indian market**. International Journal of Electronic Customer Relationship Management, 9(2-3), 189-201, 2015. https://doi.org/10.1504/IJECRM.2015.071716

24. Dalvi, P. K., Khandge, S. K., Deomore, A., Bankar, A., & Kanade, V. A. **Analysis of customer churn prediction in telecom industry using decision trees and logistic regression**. In 2016 Symposium on Colossal Data Analysis and Networking (CDAN) (pp. 1-4). IEEE, March, 2016. https://doi.org/10.1109/CDAN.2016.7570883

25. Gunawan, F. E., Soewito, B., Surantha, N., & Mauritsius, T. **One More Reason to Reject Manuscript about Machine Learning for Structural Health Monitoring**. 2018 Indonesian Association for Pattern Recognition International Conference (INAPR), 2018.

26. Oumayma Oueslati , Ahmed Ibrahim S. Khalil, Habib Ounelli. **Sentiment Analysis for Helpful Reviews Prediction**. International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), vol. 7, 2018. https://doi.org/10.30534/ijatcse/2018/02732018

27. Apoorva Deshpande, Ramnaresh Sharma. **Multilevel Ensemble Classifier using Normalized Feature based Intrusion Detection System**. International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), vol.7, October, 2018. https://doi.org/10.30534/ijatcse/2018/02752018