

# Analog the Performance between Three Classifiers on Bank Marketing Data

Gan Fui Yee, Suliadi Firdaus Sufahani, Mustafa Mamat, Mohamad Afendee Mohamed, Puspa Liza Ghazali

**ABSTRACT---** *There are several different classification methods can be used to do the classification which can classified the data into specified groups or classes. This paper presents a comparison of performance between three classifiers which include Naïve Bayes, Decision Tree and Neural Network on Bank Marketing dataset. This study focus on which classifier will have the better performance based on some performance measure in two different datasets. The result shows that machine learning classifier was not able compare to Naïve Bayes and Decision Tree classifier. Based on the results, the huge dataset obtained the more information which can be predict accurately and identify the performance of the classifier correctly.*

**Index Terms** — Artificial Neural Network Decision Tree, marketing data classification, Naïve Bayes.

## I. INTRODUCTION

Research in data mining has been warmly welcomed lately for its capacity to solve real world problem in classification, clustering and prediction. Intelligent decision making is considered as complex problem which now has successfully been addressed by classification technique. Classification is a function that tries to distinguish data according to categories or classes. By applying classification technique, we can accurately decide the class labels for which the attribute values are known, yet the class values remain to be found [1].

Decision Tree is a linear method which can be understood and interpreted easily [2]. Decision Tree offers a fast and versatile approach to classification of instances in large datasets having many attributes [1], [3]-[5]. The Decision Tree algorithm is well known for its robustness and learning efficiency with a learning time complexity [2]. Decision Tree Algorithms are extensively used in machine learning field to classify biomedical signals [6]. There are two general problems when constructing the Decision Trees which are the substantial growth of the tree during training phase and the pruning process to remove superfluous nodes and branches to improve classification accuracy. In general, a Decision Tree classifier is based on greedy approach

**Revised Manuscript Received on July 10, 2019.**

**Gan Fui Yee**, Department of Mathematics and Statistic, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Campus, 84600 Pagoh, Johor, Malaysia.

**Suliadi Firdaus Sufahani**, Department of Mathematics and Statistic, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Campus, 84600 Pagoh, Johor, Malaysia.

**Mustafa Mamat**, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Kampung Besut, 22200 Kuala Terengganu, Terengganu, Malaysia.

**Mohamad Afendee Mohamed**, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Kampung Besut, 22200 Kuala Terengganu, Terengganu, Malaysia.

**Puspa Liza Ghazali**, Faculty of Economics and Management Sciences, Universiti Sultan Zainal Abidin, Kampung Gong Badak, 21300 Kuala Terengganu, Terengganu, Malaysia.

which offers an efficient and effective technique for grouping data instances [1]. A Decision Tree classifier is similar to a Tree-like structure consisting of internal nodes, leaf nodes, and branches [7].

Bayesian classification is a famous supervised classification model. In Bayesian classification, Naïve Bayes classifier is a classifier which is using probability to perform the classification task [8]. Besides, Naïve Bayes classifier is a general classifier which always executed surprisingly in classification task. The strong assumption of Naïve Bayes classifier considers the attributes as conditionally independent given the class label [1], [9]-[11]. Neural network is an interconnected group of artificial neurons simulating the thinking process of human brain [12]-[14]. Neural network classifier is use to analyse the stimulus vector from the statistical model in order to decide whether the network data belong to one group or otherwise [13], [15]-[18]. Neural networks are known as an effective method to classify the class labels. Neural networks were used to detect irregular activities in the datasets [19], [36], [37].

## II. LITERATURE REVIEW

Decision Trees and Naïve Bayes are efficient, easy and useful to solve the classification issues in data mining [20]. The Decision Tree has some benefits and limitation during the construction. The advantage of Decision Tree is simplicity where the Decision Tree can be constructed easily and interpret rapidly from the resulting Tree while the drawback of Decision Tree is the disallow of the multiple output and noisy data [7]. The presence of noisy data comes with drawbacks on the generated Decision Tree such as over fitting and accuracy reduction. The Decision Tree can remove the noise by using the Decision Tree induction [20].

Naïve Bayes classifier is frequently employed for solving various classification problems in the field of data mining because the Naïve Bayes classifier provides simplicity and impressive classification accuracy [1]. The benefits of Naïve Bayes classifier are that it is easy to use and it only need a small size of training data to obtain the parameter estimates which is required for probability generation [1], [8], [21]. Moreover, Naïve Bayes classifier can overcome the absent attribute value via eliminating the corresponding probabilities for those attributes after the probability of each class have been calculated [1], [22],[23].

Neural networks have the limitation which it has the poor interpretability. However, it has some advantages such as high tolerance when dealing with noisy data and ability to classify patterns without prior training. Besides, neural network can function in the environment where there only exist little knowledge of the relationships between attributes and classes [13]. Moreover, Neural network is much more suitable for continuous-valued inputs and outputs, than other algorithms [3].

### III. METHODOLOGY

#### A. Decision Tree

Decision Tree composes of root, nodes, branches, and leaves [3], [6], [22], [24]. It classifies the feature vectors by traversing from the root and moves until a leaf node is identified. A highest entropy reduction algorithm is employed for the purpose of discovering the best feature among all features [1]. Fig. 1 shows the process of Decision Tree classifier.

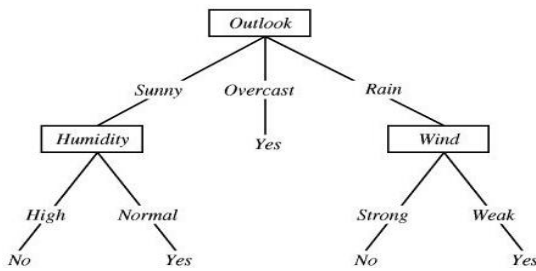


Fig. 1: Process of Decision Tree classifier

#### B. Naïve Bayes

Naïve Bayes classifier is process to estimate the probability of new data in predefined category by according to Bayes theorem and produced a better output in classification task [5], [21]-[23]. Naïve Bayes classifier provides prior knowledge, information and practical learning algorithms [8]. The output of the Naïve Bayes classifier can be enhances by removing inappropriate features from the modelling process. It distinguished the effective computational methods in selecting appropriate features for Naïve Bayes classification [5], [31]-[35].

Fig. 2 shows the formula of Naïve Bayes Classifier.

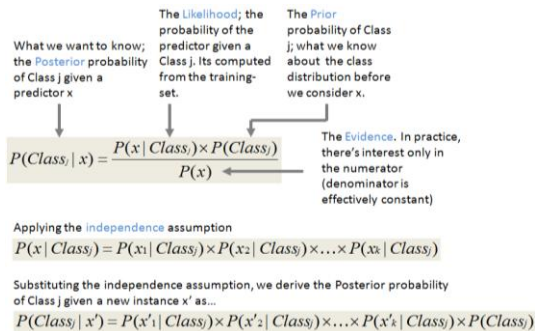


Fig. 2: Formula of Naïve Bayes classifier

#### C. Neural Network

Neural network is made of simple, connected processors as neurons and each neurons produce a sequence of real-valued activations [14]. Input to first layer neurons is absorbed from the environment, while the other neurons

through weighted connections from previously active neurons. Some neurons may effects the circumferences by triggering actions [26]. Learning is about tuning the weights that can help the network produce appropriate behaviour [27]. During the learning phase, the weight connecting each node is adjusted such that it will satisfy the actual label assigned to the input attributes. Neural network normally requires a long training times, but it is suitable for lots of applications [3], [28].

Fig. 3 shows the process of Neural Network classifier.

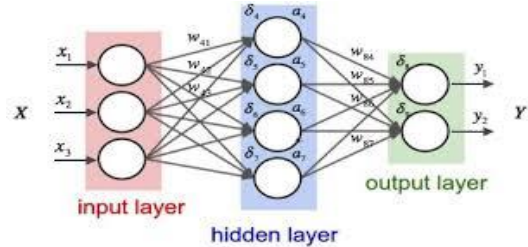


Fig. 3: Process of Neural Network classifier

#### D. Performance Measures

There are several performance measurement that can be used to identify how well the classifier performance [29], [30], [38]. In this study, there are seven performance measurements were applied to evaluate the performance of the classifiers which are accuracy, precision, specificity, sensitivity, recall, error rate and F-measurement. The formula for all performance measurements are as below [29]:

$$\text{Accuracy} = (tp+tn)/(tp+fp+fn+tn) \tag{1}$$

$$\text{Precision} = tp/(tp+fp) \tag{2}$$

$$\text{Specificity} = tn/(tn+fp) \tag{3}$$

$$\text{Sensitivity} = tp/(tn+fp) \tag{4}$$

$$\text{Recall} = tp/(tp+fn) \tag{5}$$

$$\text{Error rate} = (fp+fn)/all \tag{6}$$

$$F\text{-measure} = \frac{(\beta^2 + 1) * precision * sensitivity}{\beta^2 * precision + sensitivity} \tag{7}$$

where

$tp$ : True Positive: A mushroom predicted is edible when it is actually edible.

$tn$ : True Negative: A mushroom predicted is poisonous when it is actually poisonous.

$fp$ : True Positive: A mushroom predicted is edible when it is actually poisonous.

$fn$ : True Positive: A mushroom predicted is poisonous when it is actually edible.

### IV. RESULTS AND DISCUSSION

The bank marketing prediction is compared between three classifiers which includes Naïve Bayes classifier, Neural Network classifier and Decision Tree classifier. Two datasets are used to do the comparison of performance between three classifiers with same number of attributes but



different number of observations. The first dataset contains 41188 observations while the second dataset contains 3000 observations. Both datasets involve one class label which to identify whether the client subscribed a term deposit and 20 attributes which stated as following:

- I. Age
- II. Type of job
- III. Marital status
- IV. Education level
- V. Credit in default
- VI. Housing loan
- VII. Personal loan
- VIII. Contact communication type
- IX. Last contact month of year
- X. Last contact day of the week
- XI. Last contact duration
- XII. Number of contacts made during this campaign and for this client
- XIII. Number of days since the last contact with the client in a previous campaign
- XIV. Number of contacts made before this campaign and for this client
- XV. Outcome of the previous marketing campaign
- XVI. Employment variation rate
- XVII. Consumer price index
- XVIII. Consumer confidence index
- XIX. Euribor 3 month rate
- XX. Number of employees

Table 1 illustrates the performance measurement for each of the classifier by using Rapidminer Automodel software.

**Table 1: Performance measurement comparison between three classifiers (41188 observations)**

Model	Naïve Bayes	Neural Network	Decision Tree
Accuracy, %	86.3	90.3	91.0
Precision, %	42.6	55.5	66.2
Specificity, %	89.6	92.7	97.3
Sensitivity, %	60.9	72.0	41.8
Recall, %	60.9	72.0	41.8
Error Rate, %	13.7	9.7	9.0
F-measure, %	50.1	62.7	51.3

From Table 1, Neural Network classifier has the highest sensitivity, recall and F-measure while Decision Tree classifier has the highest accuracy, precision, specificity and the lowest error rate. Therefore, Decision Tree classifier has the better performance compare to Naïve Bayes and Neural Network classifiers.

Table 2 shows the performance measurement for each of the classifier by using Rapidminer Automodel software.

**Table 2: Performance measurement between three classifiers (3000 observations)**

Model	Naïve Bayes	Neural Network	Decision Tree
Accuracy, %	81.0	88.2	89.0
Precision, %	31.4	49.3	56.2
Specificity, %	84.7	93.4	97.4
Sensitivity, %	52.9	48.6	25.7

Recall, %	52.9	48.6	25.7
Error Rate, %	19.0	11.8	11.0
F-measure, %	39.4	48.9	35.3

From Table 2, the result shows that Decision Tree classifier obtains the highest accuracy, precision, specificity and lowest error rate which can shows that Decision Tree classifier performed better compare to Naïve Bayes classifier and Neural Network classifier. Besides, Neural Network classifier obtains the highest F-measure and Naïve Bayes classifier obtains the highest sensitivity and recall.

Compare both tables with different datasets, the Decision Tree classifier always have the highest accuracy, precision, specificity and the lowest error rate. While, in the dataset with 300 observations, Naïve Bayes classifier obtains the highest sensitivity and recall but Naïve Bayes classifier did not obtains the highest sensitivity and recall in the dataset with 41188 observations.

## V. CONCLUSION

The results show that the Decision Tree is a better classifier compare to Naïve Bayes classifier and Neural Network classifier in the any dataset because the performance measurement of Decision Tree obtained four better performance from the performance measurement which are the highest accuracy, precision, specificity and the lowest error rate. Besides, the Naïve Bayes classifier is simple and highly performed in small dataset compare to Neural Network where Naïve Bayes classifier obtained the highest sensitivity and recall in the dataset with 3000 observations but Neural Network classifier did not obtained. From the research that been done, the comparison of performance between the classifiers can involves more than three classifiers, I suggest the researchers to use other classifiers to compare the performance to understand more about the performance of classifiers. Moreover, the researchers may compare the performance between the classifiers with different datasets with different number of attributes.

## ACKNOWLEDGMENT

First of all, I would like to take this previous opportunity to express my gratitude and appreciation to my lecturer, Dr. Suliadi Firdaus Sufahani for his responsibility, knowledge and commitment in helping me in completing this study. In addition, I am thankful for his guidance and knowledge sharing during my progress in doing this study. Moreover, I like to thank our friends who supported me when I faced difficulty in the study. Last but not least, I would like to thank my family for their help and support. Thank you to Universiti Tun Hussein Onn Malaysia for the GPPS.

## REFERENCES

1. D. Farid, L. Zhang, C. Mofizur, M. A. Hossain, and R. Strachan, "Hybrid decision tree and Naïve Bayes classifiers for multi-class classification tasks," *Expert Systems with Applications*, 41(4), 2014, pp. 1937-1946.



2. K. Chen, K. Wang, K. Wang, and M. Angelia, "Applying particle swarm optimization-based decision tree classifier for cancer classification on gene expression data," *Applied Soft Computing*, 24, 2014, pp. 773-780.
3. H. Jiawei, K. Micheline, and P. Jian, *Data Mining Concepts and Techniques*. Massachusetts: Morgan Kaufmann Publishers, 2012.
4. K. Trapeznikov, and V. Saligrama, "Supervised sequential classification under budget constraints," 16th International Conference on Artificial Intelligence and Statistics, 2013, pp. 581-589.
5. G. Keerthika and D. S. Priya, "Feature subset evaluation and classification using Naive Bayes Classifier," *Journal of Network Communications and Emerging Technologies*, 1(1), 2015, pp. 22-27.
6. E. Gokgoz, A. Subasi, and B. S. Processing, "Comparison of decision tree algorithms for EMG signal classification using DWT," *Biomedical Signal Processing and Control*, 18, 2015, pp. 138-144.
7. D. T. Bui, T. C. Ho, I. Revhaug, and B. Pradhan, "Landslide susceptibility mapping along the National Road 32 of Vietnam using GIS-based J48 Decision Tree classifier and its ensembles," in *Cartography from Pole to Pole*, M. Buchroithner, N. Prechtel and D. Burghardt, Eds. Berlin: Springer, 2014, pp. 303-317.
8. L. Dey, S. Chakraborty, A. Biswas, B. Bose, and S. Tiwari, Sentiment analysis of review datasets using Naive Bayes and K-NN classifier. 2016, Available: <https://arxiv.org/ftp/arxiv/papers/1610/1610.09982.pdf>.
9. B. Baesens, T. Van Gestel, S. Viaene, M. Stepanova, J. Suykens, and J. Vanthienen, "Benchmarking state-of-the-art classification algorithms for credit scoring," *Journal of the Operational Research Society*, 54, 2003, pp. 627-635.
10. B. Baesens, R. Castelo, and J. Vanthienen, "Learning Bayesian network classifiers for credit scoring using Markov Chain Monte Carlo search," *IEEE Object Recognition Supported by User Interaction for Service Robots*, 2002, pp. 49-52.
11. M. J. Islam, Q. M. J. Wu, M. Ahmadi, and M. A. Sid-Ahmed, "Investigating the performance of Naive- Bayes classifiers and K- Nearest Neighbor classifiers," *Journal of Convergence Information Technology*, 5(2), 2010, pp. 133-137.
12. S. G. Wu, F. S. Bao, E. Y. Xu, Y. Wang, Y. Chang, and Q. Xiang, "A Leaf Recognition Algorithm for plant classification using Probabilistic Neural Network," *IEEE International Symposium on Signal Processing and Information Technology*, 2007, pp. 11-16.
13. G. Vinodhini, and R. M. Chandrasekaran, "A comparative performance evaluation of neural network based approach for sentiment classification of online reviews," *Journal of King Saud University-Computer and Information Sciences*, 28(1), 2016, pp. 2-12.
14. D. Zeng, K. Liu, S. Lai, G. Zhou, and J. Zhao, "Relation classification via Convolutional Deep Neural Network," 25th International Conference on Computational Linguistics, 2014, pp. 2335-2344.
15. Y. Shao, W. Chen, and N. Deng, "Nonparallel hyperplane support vector machine for binary classification problems," *Information Sciences*, 263, 2014, pp. 22-35.
16. C. Feng, W. Sun, J. Zhang, and L. Cao, "A Support Vector Machine based Naive Bayes Algorithm for spam filtering," *IEEE 35th International Performance Computing and Communications Conference*, 2016, pp. 1-8.
17. L. Chih-Wei, H. Chih-Chung, and C. Chih-Jen, "Evaluating unsupervised and supervised image classification methods for mapping cotton root rot," *Precision Agriculture*, 16(2), 2014, pp. 201-215.
18. C. Hsieh, S. Si, and I. S. Dhillon, "A Divide-and-Conquer Solver for Kernel Support Vector Machines," 31st International Conference on Machine Learning, 2014, pp. 566-574.
19. Z. Zhang, J. Li, C. N. Manikopoulos, J. Jorgenson, and J. Ucles, "HIDE: A hierarchical network intrusion detection system using statistical preprocessing and neural network classification," *IEEE Workshop on Information Assurance and Security*, 2001, pp. 85-90.
20. H. Parvin, M. Mirnabibaboli, and H. Alinejad-Rokny, "Engineering Applications of Artificial Intelligence Proposing a classifier ensemble framework based on classifier selection and decision tree," *Engineering Applications of Artificial Intelligence*, 37, 2015, pp. 34-42.
21. P. Tsangaratos, and I. Ilia, "Comparison of a logistic regression and Naive Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size," *Catena*, 145, 2016, pp. 164-179.
22. R. Pandya, "C5.0 algorithm to improved decision tree with feature selection and reduced error pruning," *International Journal of Computer Applications*, 117(16), 2015, pp. 18-21.
23. H. Geng, T. Lu, X. Lin, Y. Liu, and F. Yan, "Prediction of protein-protein interaction sites based on Naive Bayes Classifier," *Biochemical Research International*, 2015, pp. 1-7.
24. B. Krawczyk, M. Wozniak, and G. Schaefer, "Cost-sensitive decision tree ensembles for effective imbalanced classification," *Applied Soft Computing*, 14, 2014, pp. 554-562.
25. R. K. Sharma, V. Sugumaran, and M. Amarnath, "A comparative study of naive Bayes classifier and Bayes net classifier for fault diagnosis of roller bearing using sound signal Hemantha Kumar," *International Journal of Decision Support Systems*, 1(1), 2015, pp. 115-129.
26. D. P. B. T. B. Strik, A. M. Domnanovich, L. Zani, R. Braun, and P. Holubar, "Prediction of trace compounds in biogas from anaerobic digestion using the MATLAB Neural Network Toolbox," *Environmental Modelling and Software*, 20(6), 2005, pp. 803-810.
27. J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, 61, 2015, pp. 85-117.
28. B. Zheng, S. W. Myint, P. S. Thenkabail, and R. M. Aggarwal, "A support vector machine to identify irrigated crop types using time-series Landsat NDVI data," *International Journal of Applied Earth Observation and Geoinformation*, 34, 2015, pp. 103-112.
29. M. Sokolova, and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing and Management*, 45, 2009, pp. 427-437.
30. D. Jayeola, Z. Ismail, S. F. Sufahani, and D. P. Manliura, "Optimal method for investing on assets using Black Litterman Model," *Far East Journal of Mathematical Sciences*, 101(5), 2017, pp. 1123.
31. S. Sufahani, and Z. Ismail, "The statistical analysis of the prevalence of pneumonia for children age 12 in west Malaysian hospital," *Applied Mathematical Sciences*, 8(114), 2014, pp. 5673-5680.
32. S. Sufahani, N. Che-Him, A. Khamis, M. S. Rusiman, N. Arbin, C. K. Yee, I. N. Ramli, N. A. Suhaimi, S. S. Jing, and Z. A. Azmi, "Descriptive statistics with Box-Jenkins and marketing research for jewellery company in Malaysia," *Far East Journal of Mathematical Sciences*, 101(10), 2017, pp. 2151-2161.
33. S. F. Sufahani, and A. Ahmad, "A comparison between normal and non-normal data in bootstrap," *Applied Mathematical Sciences* 6(89-92), 2012, pp. 4547-4560.

34. M. S. Rusiman, O. C. Hau, A. W. Abdullah, S. F. Sufahani, and N. A. Azmi, "An analysis of time series for the prediction of Barramundi (Ikan Siakap) price in Malaysia," *Far East Journal of Mathematical Sciences*, 102(9), 2017, pp. 2081-2093.
35. Z. Ismai, N. Abu, and S. Sufahani, "New product forecasting with limited or no data," *AIP Conference Proceedings*, 1782(1), 2016, pp. 1-8.
36. S. Sufahani, M. G. Kamardan, M. S. Rusiman, M. Mohamad, K. Khalid, M. Ali, K. Khalid, M. K. M. Nawawi, and A. Ahmad, "A mathematical study on "additive technique" versus "branch and bound technique" for solving binary programming problem," *Journal of Physics: Conference Series*, 995(1), 2018, pp. 1-8.
37. M. Ali, S. Sufahani M. G. Kamardan, and Z. Ismail, "A new diet scheduling model for Malaysian school children using zero-one optimization approach," *Global Journal of Pure and Applied Mathematics*, 12(1), 2016, pp. 413-419.
38. A. Zinober, and S. Sufahani, "A non-standard optimal control problem arising in an economics application," *Pesquisa Operacional* 33(1), 2013, pp. 63-71.