Jurnal Mantik, 7 (1) (2023), ISSN 2685-4236 (Online)

Published by:Institute of Computer Science (IOCS)

Jurnal Mantik

Journal homepage: www.iocscience.org/ejournal/index.php/mantik

Implementation of hybrid methods in data mining for Predicting customer churn in the telecommunications sector

Restu Herdian[1], Abba Suganda Girsang[2]
[1,2]Computer Science Department/Bina Nusantara University, Indonesia

A R T I C L E I N F O

Article history:

Received Mar 25, 2023
Revised Apr 20, 2023
Accepted Apr 28, 2023

Keywords:

Artificial Neural Network
Decision Tree
Hybrid Algorithm

ABSTRACT

In recent years, the telecommunication industry is growth and become very competitive where has reached the point maintaining customer is very essential than acquiring new customer. And the two key factor for maintaining customer, the first is defining the segment of customer want to churn and the second is accuracy of predictive model. In this article we propose the hybrid model based on decision tree and artificial neural network (ANN) with the two stages of process to answer the problem of maintaining customer, the first is a segmentation phase with decision rules and the second is a prediction phase with artificial neural network (ANN). Our finding in benchmarked against the previous algorithms (decision tree and ANN) with the AUC metrics show the proposed model or hybrid achieves better accuracy and with the comprehensive information of what a drive customer churn.

Corresponding Author:

Restu Herdian,
Computer Science Department,
Bina Nusantara University,
Kebun Jeruk 27, West Jakarta – DKI Jakarta,11480, Indonesia.
Email: restu.herdian001@binus.ac.id

1. INTRODUCTION

In the last few decades, mobile telecommunication growth become the main communication across the world, the market becomes very competitive has reached the point where every potential customer has to be won over by every competitor. Another reason the standardization and regulation of mobile telecommunication allowed customers to have two different network (Rehman & Ali, 2014), and therefore it has become essential to maximize the profits on a regular basis, for that various strategies have been proposed, acquiring new customers, up selling the existing customer & increase the retention period of existing customers. Among all the strategies, retention of existing customers is the least expensive as compared to others (Lalwani et al., 2022) because as the intensified competition in the market can affect to acquiring new customer because become more difficult and can cost up to six times more than what it cost to retaining a current customer (Amin et al., 2019), And the real problem for the situation is when companies losing a customer or customer churn (De Caigny et al., 2018).

Managing customer churn becomes more essential for the development and maintaining the relationship between customer and companies (Routh et al., 2020). And in order to cure customer churn it has become essential to companies and managers to use the data of historical customer to create model prediction, to search and identify customers who show high potential to make churn and using retention strategy campaign to reduce customer churn (De Caigny et al., 2018).

Several factors become the crucial factor in the reason behind the problem of customer churn and also become solution to find out the probability to make customer abandon company or customer churn, and for that reason, in the competitive marketplace has led to motive building and development customer churn prediction model (Kumar & Kumar, 2019). Customer churn has been handled on two different perspectives from previous research. The first, researchers focus on boosting predictive model by improving models with more complex models, and the second, researcher want to understand what drives customer churn such as customer satisfaction. Based on the main concerns of two type research perspective, the customer churn model should have good result of predictive performance and provide lead to actionable insight (De Caigny et al., 2018).

The development customer churn prediction have been successful, applied with many technique with different algorithms like SVM, logistic regression, decision tree, naïve bayes and neural network in every sector like banking, airlines and many more (Kumar & Kumar, 2019). The decision tree algorithm has characteristic form their ability to make split the data into more homogeneous subset (De Caigny et al., 2018) and the decision tree algorithm become very useful in churn prediction settings to understand what variable in data can effect customer leaving the company (Höppner et al., 2020), because the algorithm of decision tree has process into two phases is building and pruning, from the building phase the data set to grouping by subset with some identical value and the second phase is pruning removing the noisy data (Nisha Saini et al., 2017) and artificial neural networks model is implemented from the concept of ability to learn from mistakes.how brain works to stimulates neural networks. This process occurs from brain processing information which will then be transmitted through neurons (Abdulsalam et al., 2022).

For the references in the propose of hybrid concept we find related works with similar topic like In the previous research, with the research with the same concept of collaboration from decision tree and artificial neural network but with a different approach and in different area, it is called CT-ANN is implemented from the classifier tree and artificial neural network (ANN), which is in the implementation CT-ANN firstly algorithm make split the feature with CT Algorithm to choose important feature and feature reduction and the second step ANN model build using the important variables from CT algorithm along with prediction results made by CT Method and used as input information in the input layer of neural networks (Chakraborty et al., 2018), another previous research with the same concept of collaboration but in different approach for using hybrid concept with decision tree and logistic regression by propose to make better predictive model with two step approach in the process, the first step using decision tree to identify homogenous segments with decision rule and the second step is to make a prediction with logistic regression from each segments (De Caigny et al., 2018), And the last references for similar concept is from the implementation of collaboration between clustering analysis and decision tree, which is the role of cluster analysis is to find market segment and decision tree as a prediction model to identify churn or no churn customer (Pejić Bach et al., 2021).

2. RESEARCH METHOD

The implementation methodology for hybrid algorithms divided into 6 phases, from data preparation, data preprocessing, feature selection, undersampling, training & testing, and final process the implementation of hybrid algorithms with the evaluation to measure result of model with AUC metric. The methodology of implementation presented in figure 1.

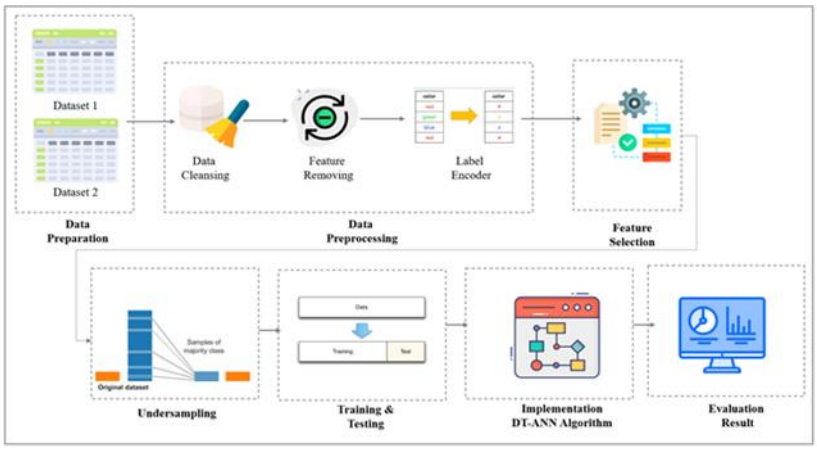

Figure 1. The methodology of implementation

2.1. Data Preparations

In the implementation, the experiment uses two datasets from the telecommunication sector for customer churn presented in Table 1.

Table 1. Dataset customer churn

	Dataset Name	Source	Total Records	Total Attribute
DS1	Train & Test	https://www.kaggle.com/c/customer-churn-prediction-2020/data	5,000	20
DS3	WA_Fn-UseC_-Telco-Customer-Churn	https://www.kaggle.com/blastchar/telco-customer-churn	7,043	21

For the first dataset (DS1) has 20 attribute is [state ,account length ,phone number ,international plan ,voice main plan ,num of voice mail messages ,total day minutes ,total day calls ,total day charge ,total eve minutes ,total eve calls ,total eve charge ,total night minutes ,total night calls ,total night calls ,total night charge ,total international minutes ,total intl calls ,number customer service calls ,churn].

A And the second dataset (DS2) has 21 attribute is [cutomerid ,gender ,senior citizen ,partner ,dependents ,tenure ,phone service ,multiplelines ,internet service ,online security ,online backup ,device protection ,tech support ,streaming tv ,streaming movies ,contact ,paperless billing ,payment method ,monthly charges ,total charges ,churn].

2.2. Data Preprocessing

The implementation data preprocessing begins with maintaining missing values with imputation procedure if missing values more than 5% of the values of an attribute and if missing values less than 5% the procedure will remove data has missing values, and for the next step to make dataset become more simple will begin feature removing to

reduce attribute has one categorical variable and difference categorical for all variable like customerid presented in figure 2.

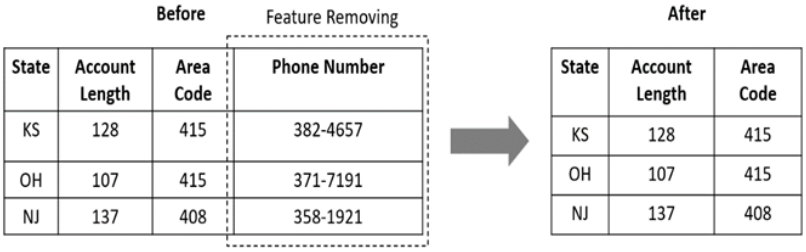

Figure 2. Feature Removing

After maintaining missing values, The implementation of all categorical variables has next phases using encoding which is turned into binary variables.the implementation presented in figure 3.

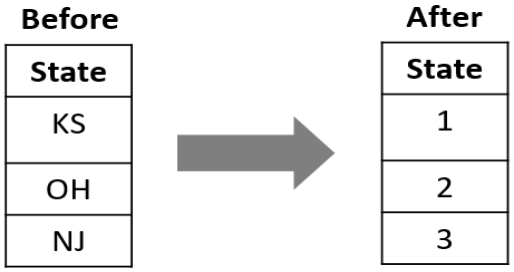

Figure 3 Label Encoding process

The minority class in imbalanced dataset always more important. For example like in the case of customer churn prediction, the minority class is customers who are about to leaves organization (Shumaly et al., 2020). And therefore in the implementation of customer churn always has a problem with imbalanced classes between variable non-churn and churn because non-churn classes always become majority than churn classes, and to solve problem and to make variable balance the researcher implemented with under sampling presented in figure 4.

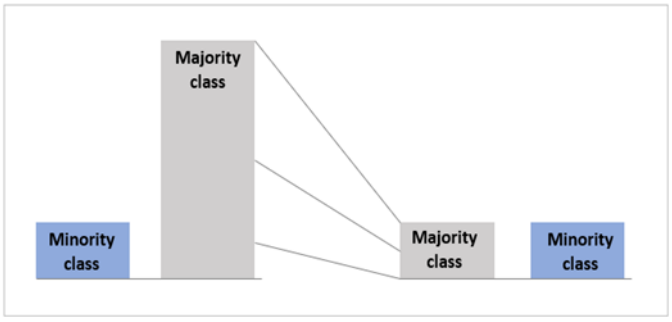

Figure 4 Undersampling Process

2.3. Feature Selection

The role of selecting certain attribute is become essential because it has a direct impact on the effective classification results (Sivasankar & Vijaya, 2019). And feature selection has implement to reduce number of variables become more simple and more effective to use for predict (De Caigny et al., 2018).The implementation of fischer score described by equation(1).

$$Fisher\ score = \frac{\bar{X}_c - \bar{X}_{nc}}{\sqrt{S_c^2 + S_{nc}^2}} \quad (1)$$

2.4. Algorithm Implementation

The approach for innovation is the combination with two algorithms based on two phases from decision tree with decision rule to split the data into more homogenous subsets and artificial neural network (ANN) approach with predictive model fit with a dataset. The model presented in figure 5.

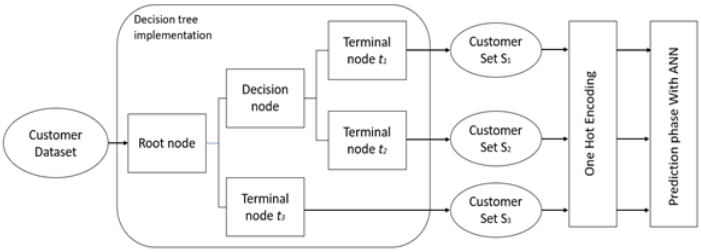

Figure 5. The structure of model hybrid algorithm

The entire customer set S divided with different criteria from attribute of data by the structure from tree with consist set of leaves or terminal node T (De Caigny et al., 2018)(Coussement et al., 2020).the implementation of decision tree for every subset described by equations(2):

$$\boldsymbol{S} = \cup_{t \in T} \boldsymbol{S_t} : \forall\, \boldsymbol{t} \neq \mathbf{t^I} : \boldsymbol{S_t} \cap \boldsymbol{S_{t^I}} = \boldsymbol{\theta} \quad (2)$$

In the first phase of implementation is segmententation phase, the detailed mechanism of implementation decision tree with the process to a segmented customer data by recursively split into smaller and purer subsets, the process starts in the root node without parent node and the iterative process determines to optimal splitting segment that divide the data over two child nodes. The process will ends when no further splits are desirable or possible and the set nodes without child nodes called terminal nodes or leaves. The implementation process presented in figure 6.

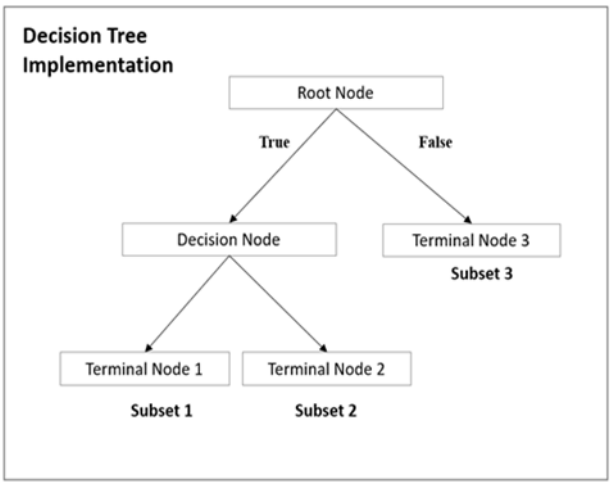

Figure 6. The implementations structure of decision tree

After implementation of decision tree with decision rules has create new attribute to describing heterogeneity of data and the second step of process is encoding new attribute with one-hot encoding methods to create dimension based on categorical of new attribute presented in figure. 7.

Index	Terminal Node
1	Subset 1
2	Subset 2

index	Subset 1	Subset 2
1	1	0
2	0	1

Figure 7. One-hot encoding process

Based on the result as a new attributes with the implementation of decision tree, the second phases is the implementation of prediction model with artificial neural network. The model consist of the interconnection of neurons by respective weight for each connection and the output of neural network from experience of learning during the training process between connection of neurons. And in the implementations neural network occur of multi layers which is the signal input spread through layer by layer (Ismail et al., 2015). The process of neural network system presented in figure 8.

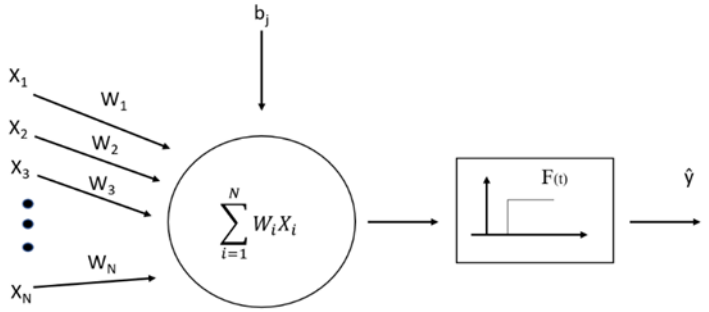

Figure 8. The process of neural network systems

In the training process present in figure. 4, the process of input data repeatedly fed into neural network. The performance predictive based on neural network are comparing between desired output and calculated errors . In the process error output will repeatedly fed back to the neural network model and weights adjusted until have the minimum error is achieved to produce better the desired output (Ismail et al., 2015). The process presented in equation(2) and variable description of figures 3 presented in Table 2.

$$\varepsilon_i = y_i - \hat{y}_i \quad (2)$$

Table 2. Variable & description

Variable	Description
$\hat{y}_i$	is neural network output
y_i	is actual output
W_i	is connections weight between hidden and input layers
X_i	is input nodes
b_j	is thresholds in hidden layers

In the implementation, when data from an input is presented at the input layer, the nodes of network (neurons) make calculations in the successive layer until an output

value is computed at each of the outputs nodes (Ismail et al., 2015)]. And the result in output i[th] hidden node described by equation(3).

$$h_i = \sum_{i=1}^{N} W_i X_i \quad (3)$$

And result for the output of i[th] neural network described by equation(4).

$$\hat{y}_i = F\left(\sum_{i=1}^{N} W_i X_i + b_j\right) \quad (4)$$

2.5. Evaluation Criteria

The measurement of implementation using metric with the area under the receiver operating characteristics curve (AUC) in the result of testing model. The AUC can be derived from the confusion matrix. Table 3 presented by confusion matrix for binary classification

Table 3. Confusion matrix for binary classification

		Actual		
		1	0	
Predicted	1	True Positive (TP)	False Positive (FP)	Predictive Positive (PP)
	0	False Negative (FN)	True Negative (TN)	Predictive Negative (PN)
		Actual Positives (AP)	Actual Negative (AN)	

The AUC can provides estimating probability with randomly chosen instance of class 1 by correctly rated and higher rank than instance of class 0 (De Caigny et al., 2018). The formula of metric AUC described by Eq. (6). UC can provides estimating probability with randomly chosen instance of class 1 by correctly rated and higher rank than instance of class 0

$$AUC = \int_0^1 \frac{TP}{AP} d\frac{FP}{AN} \quad (6)$$

3. RESULTS AND DISCUSSIONS

3.1. Result

The methodology of algorithm has two phases in the implementation from phases implementation of decision tree to split data and create new attribute with description of every subset of data through process of encoding by criteria in every category subset and the second phase is implementation of artificial neural network to predict the final result of model based on new data from implementation of decision tree. The implementation phase presented in figure 9.

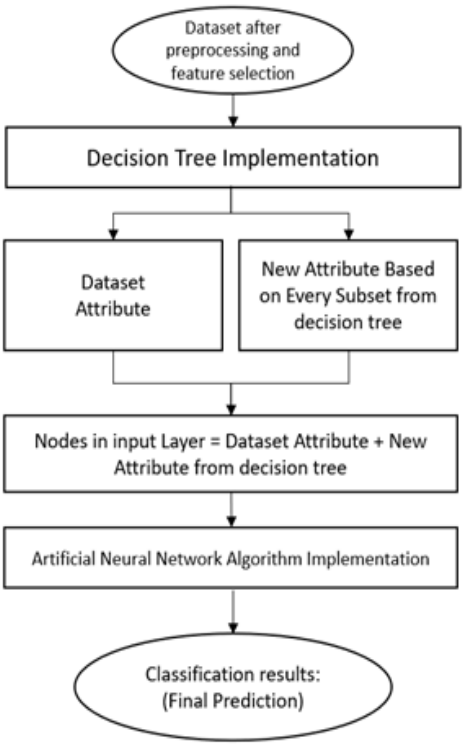

Figure 9. The implementation phase of algorithms

a. Feature Selection

In the process of implementation feature selection will produce score for every attribute shown in Table 4. and in the procedure implementation has minimum score for all of attribute and after the score produce , the last two attribute from bottom with low score will remove from dataset.

Table 4. Measurement attribute with fisher score dataset 1

No	Attribute	Score
1	Total Day Minutes	3474.569228
2	Num of Voice mail Messages	1127.545049
3	Total day Charge	590.631370
4	Total Eve Minutes	507.608392
5	International Plan	303.963357
6	Number Customer Service calls	245.459832
7	Total Night Minutes	132.878281
8	Voice mail Plan	45.058340
9	Total Eve Charge	43.139769
10	Account Length	35.321294
11	Total Intl Calls	14.959338
12	Total International Minutes	14.877082
13	Total Night Charge	5.978417

14	Total Day Calls	5.113360
15	Total Intl Charge	4.014171
16	Area Code	2.118859
17	State	2.033226
18	Total Night Calls	0.972533
19	Total Eve Calls	0.774513

The last two attribute from dataset, attribute gender and phone service will remove from dataset to reduce dataset from the inefficient of non-correlated attribute for customer churn prediction model and final result dataset has 17 attribute for the implementation of hybrid algorithm.

b. Decision Tree

In the process of implementation decision tree begin with implementation decision rule based on algorithm decision tree with setting max depth 3 , for split data into subset by criteria of every type customers from attribute dataset. The implementation of decision tree will split data based on the method of entropy from decision tree and with the setting of 3 max depth will produce subset based on the criteria of the most correlated attribute for customer churn prediction model. The decision tree implementation in dataset 1 presented by figure 10.

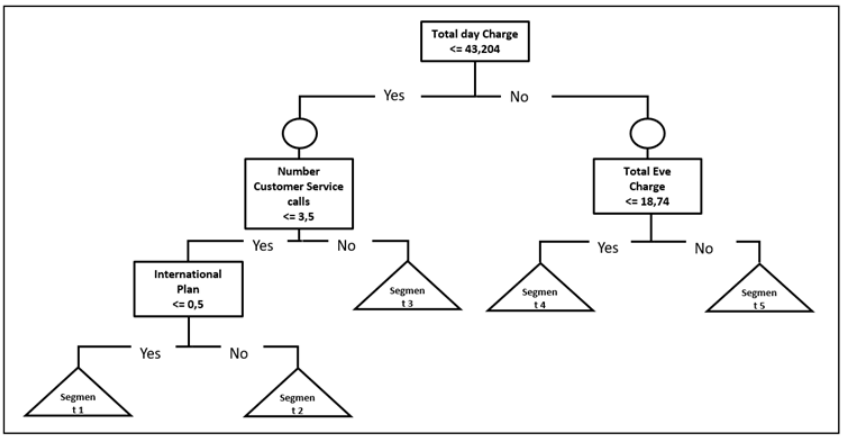

Figure 10. Decision Tree implementation

c. Artifcial Neural Network

And in the final process begin with implementation of artificial neural network (ANN) for prediction phase and in the process implementation, neural network settings adapt with the best settings for implement neuron and hidden layer. For the implementation of artificial neural network use 2 hidden layer with the first hidden layer has 10 neuron and second hidden layer with 8 neuron present in figure 11.

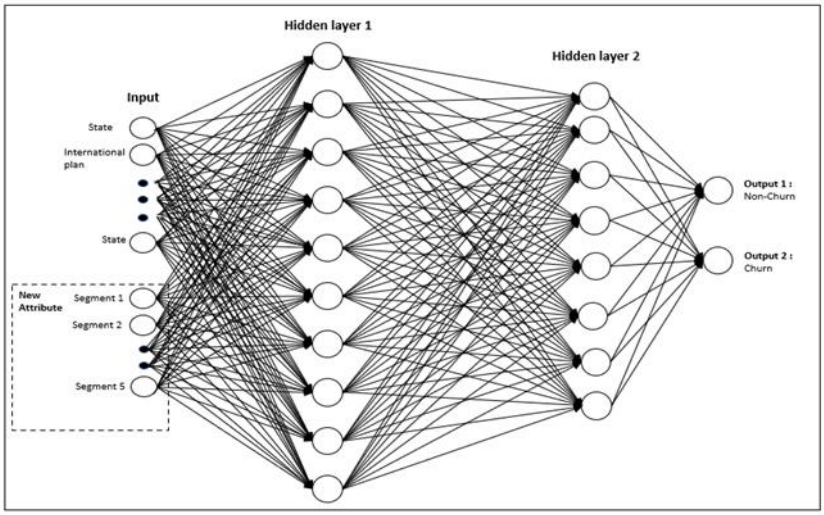

Figure 11. ANN Implementation

d. Evaluation Result

In the implementations and comparing between new algorithm and previous algorithm, the output of performance score based on metric area under the receiver operating characteristics curve (AUC) present in Table 5 with new algorithms has improved better than just with standalone algorithms

Table 5. Measurement metric using area under the receiver operating characteristics curve (AUC)

Algorithms	Datasets	
	DS1	DS2
DT-ANN	0.81	0.76
DT	0.76	0.73
ANN	0.71	0.72

3.2 Discussion

The hybrid concept has better result than standalone algorithms because the hybrid concept provide idea to maximize the potential of data and algorithms, which is the first phase can maximize the potential of data with providing new variables or information and for the second phase, the implementation of prediction has advantage to maximize potential performance in adding new variables from previous phase and the ability of artificial neural network to adapt in data capacity.

Based on the result of hybrid model, we derive insight the ability of decision tree for providing valuable information for boosting prediction model like in another research with the implementation of collaboration algorithms between decision tree and logistic regression (De Caigny et al., 2018) or the ability of decision tree can provide insightable information to help analyst understand the meaning from data, like similar insight from another research with (Höppner et al., 2020) for provide information of target customer for a customer retention campaign and for artificial neural network is become better algorithm to implementation prediction model with more valuable information or variables like similar explanation from research of CT-ANN by (Chakraborty et al., 2018).

But the result of propose concept has advantage from previous research in ability to maximize the potential of concept by the creativity to change setting algorithms to adapt with data capacity, because the two algorithms from hybrid concept from decision tree and artificial neural network has same flexible ability to change the scenario of

settings to adapt with the capacity of data, with decision tree in pruning ability to reduce noisy data and ability determine the number of segmentation with decision rules, for artificial neural network can determine number neurons and layer to maximize learning process for data and algorithm become better as a prediction model because adding new valuable information or variables to boosting model.

4. CONCLUSION

The result of implementation has conclusion the combination of algorithm between decision tree and artificial neural network become more better than standalone algorithm based on result of metric AUC. for the combination algorithm with DT-ANN has result metric is 0.81 by DS1 and 0.76 by DS2 and for standalone algorithm like decision tree is 0.76 by DS1 and 0.73 by DS2 , and with neural network is 0.71 by DS1 and 0.72 by DS2.

The better result from hybrid algorithms with DT-ANN explain the innovation in adding segmented phase before implementation of prediction model can improve the prediction result, which is the role of segmented phase is created new information or variable to describe the heterogeneity of data , which will boosting performance of prediction model.

And another conclusions from the model is the number criteria of settings in hybrid algorithm, which is make algorithm has flexibility in ability to adapt for such case as fitting model with data capability, and has capacity to detailed information from the heterogeneity of data. For that the propose of hybrid algorithmns provide better solution than standalone algorithmns with accuracy, insightful information, and capacity of algorithmns to validate many hypothesis with determine setting algorithmns.

The development of implementation in this research has better performance than the previous algorithms or standalone algorithm, but the purpose model has limitations, first such as using smaller dataset, and for that reason the next research we need to challenge algorithms to the test with using large dataset and different dataset.

And the second limitation is that purpose of this research does not go into detail to optimize the potential of each algorithm, such as adding more valuable information by exploring parameters such as like maximum depth setting in decision tree or to finding the best prediction model by maximize the number of neuron or layer in an artificial neural network, which is in futher research we need to discuss more deeper about the potential of paremeter with each algorithm in the hybrid concept.

And possibly there are a lot of number opportunities in futher development that might be pursued from innovation to maximize the concept model, such as adding another phase by using different algorithm with the goal of create new variables or information, exploring each phase of the hybrid algorithm concept and modifying with different algorithms in segmentation phase such as applying clustering algorithms and compare with this hybrid algorithm for find the best algorithms that can fit to create new information in boosting the prediction model performance, and applying different algorithms in prediction phase or using advanced concept of artificial neural network such as deep learning or CNN.

ACKNOWLEDGEMENTS

The acknowledgements for this research “Implementation of Hybrid Methods in Data Mining for Predicting Customer Churn in the Telecommunications Sector ” is for the lecturers and academic staff who provided support from the research started into finalization. And also i like to say thank to our family who have encouraged and motivated me to always positive and never give up. and hopefully the results from this research can provide benefits for the academic world and the development of technology in the future.

REFERENCES

Abdulsalam, S. O., Arowolo, M. O., Saheed, Y. K., & Afolayan, J. O. (2022). Customer Churn Prediction in Telecommunication Industry Using Classification and Regression Trees and Artificial Neural Network Algorithms. *Indonesian Journal of Electrical Engineering and Informatics*, *10*(2), 431–440. https://doi.org/10.52549/ijeei.v10i2.2985

Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2019). Customer churn prediction in telecommunication industry using data certainty. *Journal of Business Research*, *94*(October 2017), 290–301. https://doi.org/10.1016/j.jbusres.2018.03.003

Amin, A., Shah, B., Khattak, A. M., Lopes Moreira, F. J., Ali, G., Rocha, A., & Anwar, S. (2019). Cross-company customer churn prediction in telecommunication: A comparison of data transformation methods. *International Journal of Information Management*, *46*(June), 304–319. https://doi.org/10.1016/j.ijinfomgt.2018.08.015

Cenggoro, T. W., Wirastari, R. A., Rudianto, E., Mohadi, M. I., Ratj, D., & Pardamean, B. (2021). Deep Learning as a Vector Embedding Model for Customer Churn. *Procedia Computer Science*, *179*(2019), 624–631. https://doi.org/10.1016/j.procs.2021.01.048

Chakraborty, T., Chattopadhyay, S., & Chakraborty, A. K. (2018). A novel hybridization of classification trees and artificial neural networks for selection of students in a business school. *Opsearch*, *55*(2), 434–446. https://doi.org/10.1007/s12597-017-0329-2

Coussement, K., Phan, M., De Caigny, A., Benoit, D. F., & Raes, A. (2020). Predicting student dropout in subscription-based online learning environments: The beneficial impact of the logit leaf model. *Decision Support Systems*, *135*(May), 113325. https://doi.org/10.1016/j.dss.2020.113325

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, *269*(2), 760–772. https://doi.org/10.1016/j.ejor.2018.02.009

Halibas, A. S., Cherian Matthew, A., Pillai, I. G., Harold Reazol, J., Delvo, E. G., & Bonachita Reazol, L. (2019). Determining the intervening effects of exploratory data analysis and feature engineering in telecoms customer churn modelling. *2019 4th MEC International Conference on Big Data and Smart City, ICBDSC 2019*, 1–7. https://doi.org/10.1109/ICBDSC.2019.8645578

Höppner, S., Stripling, E., Baesens, B., Broucke, S. vanden, & Verdonck, T. (2020). Profit driven decision trees for churn prediction. *European Journal of Operational Research*, *284*(3), 920–933. https://doi.org/10.1016/j.ejor.2018.11.072

Ismail, M. R., Awang, M. K., Rahman, M. N. A., & Makhtar, M. (2015). A multi-layer perceptron approach for customer churn prediction. *International Journal of Multimedia and Ubiquitous Engineering*, *10*(7), 213–222. https://doi.org/10.14257/ijmue.2015.10.7.22

Kumar, S., & Kumar, M. (2019). Predicting customer churn using artificial neural network. In *Communications in Computer and Information Science* (Vol. 1000). Springer International Publishing. https://doi.org/10.1007/978-3-030-20257-6_25

Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. (2022). Customer churn prediction system: a machine learning approach. *Computing*, *104*(2), 271–294. https://doi.org/10.1007/s00607-021-00908-y

Latha, K. J., Baburao, M., & Kavitha, C. (2019). A Comparative study on Logit leaf model (LLM) and Support leaf model (SLM) for predicting the customer churn. *International Journal of Computer Sciences and Engineering*, *7*(5), 1628–1632. https://doi.org/10.26438/ijcse/v7i5.16281632

Li, P., Wang, X., Qin, Z., & Metzler, D. (2019). Combining decision trees and neural networks for learning-to-rank in personal search. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2032–2040. https://doi.org/10.1145/3292500.3330676

Liu, R., Ali, S., Bilal, S. F., Sakhawat, Z., Imran, A., Almuhaimeed, A., Alzahrani, A., & Sun, G. (2022). An Intelligent Hybrid Scheme for Customer Churn Prediction Integrating Clustering and Classification Algorithms. *Applied Sciences (Switzerland)*, *12*(18), 1–17. https://doi.org/10.3390/app12189355

Nguyen, T. D., Kasmarik, K. E., & Abbass, H. A. (2020). Towards Interpretable Deep Neural Networks: An Exact Transformation to Multi-Class Multivariate Decision Trees. *ArXiv*.

Nisha Saini, Monika, & Dr. Kanwal Garg. (2017). Churn Prediction in Telecommunication Industry using Decision Tree. *International Journal of Engineering Research And*, *V6*(04). https://doi.org/10.17577/ijertv6is040379

Pejić Bach, M., Pivar, J., & Jaković, B. (2021). Churn Management in Telecommunications: Hybrid Approach Using Cluster Analysis and Decision Trees. *Journal of Risk and Financial Management*, *14*(11), 544. https://doi.org/10.3390/jrfm14110544

Rehman, A., & Ali, A. R. (2014). Customer Churn Prediction , Segmentation and Fraud Detection in Telecommunication Industry. *ASE BigData/SocialInformatics/PASSAT/BioMedCom 2014 Conference, Harvard University*, *December*, 1–9. https://www.researchgate.net/publication/304822719

Routh, P., Roy, A., & Meyer, J. (2020). Estimating customer churn under competing risks. *Journal of the Operational Research Society*, *0*(0), 1–18. https://doi.org/10.1080/01605682.2020.1776166

Seymen, O. F., Dogan, O., & Hiziroglu, A. (2021). Customer Churn Prediction Using Deep Learning. *Advances in Intelligent Systems and Computing*, *1383 AISC*(April), 520–529. https://doi.org/10.1007/978-3-030-73689-7_50

Shumaly, S., Neysaryan, P., & Guo, Y. (2020). Handling Class Imbalance in Customer Churn Prediction in Telecom Sector Using Sampling Techniques, Bagging and Boosting Trees. *2020 10h International Conference on Computer and Knowledge Engineering, ICCKE 2020*, 82–87. https://doi.org/10.1109/ICCKE50421.2020.9303698

Sivasankar, E., & Vijaya, J. (2019). A study of feature selection techniques for predicting customer retention in telecommunication sector. *International Journal of Business Information Systems*, *31*(1), 1–26. https://doi.org/10.1504/IJBIS.2019.099524

Ulas Vural, M. Ergun Okay, E. M. Y. (2020). Churn Prediction for Telecommunication Industry Using Artificial Neural Networks. *International Journal of Computer and Information Engineering*, *14*(11).