Hybrid Method for Churn Prediction Model in The Case of Telecommunication Companies

Dwi Pebrianti
Computer Science
Budi Luhur University
Jakarta Selatan, Indonesia
dwipebrianti@iium.edu.my

Desti Destiansari Istinabiyah
Computer Science
Budi Luhur University
Jakarta Selatan, Indonesia
desti.istinabiyah@biz.linknet.co.id

Luhur Bayuaji
Computer Science
Budi Luhur University
Jakarta Selatan, Indonesia
luhurbayuaji@ump.edu.my

Rusdah
Computer Science
Budi Luhur University
Jakarta Selatan, Indonesia
rusdah@budiluhur.ac.id

***Abstract*— Internet Service Provider Companies are competing in offering attractive products to customers. The companies even mutually do acquisition the customers from other telecommunication companies. Subscribers who decided to unsubscribe the service from a service company is known as Churn. Churn rate continues to increase every year and the competition for internet service providers is getting tougher. The reason for the high Churn rate is that the retention incentive offer was rejected by Churn's potential customers. This is because the company was late in offering Churn retention incentives and customers had switched to competitors. One way to deal with this problem is to predict the possibility of customer's Churn. The purpose of this study is to create a Churn prediction model based on customer datasets and to develop a classification system that indicates customer will Churn or not. This study aims to develop and examined several hybrid models which are logistic regression with LASSO penalty, logistic regression with elastic net penalty, logistic regression with adaboost, and XGBoost with tuning parameters. The default methods for the classification process are logistic regression, Support Vector Machine (SVM), random forest and XGBoost. The results obtained from the 8 model prediction tests, where Churn predictions model using the hybrid method XGBoost with parameter tuning showed the best performance with Area Under Curve (AUC) 0.968, accuracy 94.3%, recall 98% and precision 96%.**

***Keywords*— Churn, Telecommunication, Hybrid Method.**

I. INTRODUCTION

Companies are competing to innovate in offering attractive products to customers. The company mutually acquires customers between telecommunication companies from each other. The customer stops subscribing to the telecommunications company, then the customer switches to a competitor, the customer behavior is referred to as Churn. Internet service providers have churn rates of up to 30-35% per year [9]. This can have a negative impact on the telecommunications company because the company's revenue will decline. In addition, the costs of acquiring new customers are greater than the costs of retaining customers. Therefore, efforts need to be made to retain customers. One way that telecommunications companies keep their customers from churn is to provide customer retention incentives. Retention incentives consist of various kinds, such as providing priority services, providing speed upgrades, providing bill adjustments , providing package discounts, and others. However, the policy of providing retention incentives has not been implemented effectively. This policy is often encountered with various obstacles so that customers become Churn. The obstacle that is often encountered is that the retention incentive offer is rejected by Churn's potential customers. The company was late in offering Churn retention incentives because customers had switched to competitors. Another reason is that the retention incentive offer doesn't match the needs of Churn's potential customers. One way to deal with this problem is to predict which customers will Churn. One way to predict Churn is to create a Churn prediction model based on customer datasets. What is done is the classification of the total customers into customers who are indicated to Churn and not Churn. This method can also be used to identify variables that influence customers to Churn. Thus, the company can adopt a retention incentive policy that suits the needs of Churn's potential customers and is not late in making an offer. Research on Churn subscriber predictions in telecommunication companies has been done before. Churn customer predictions on the Telcom Churn dataset using extreme learning machines get an accuracy of 76.96 % [17]. Then, Saputra states that the Churn customer prediction model can use the logistic regression method with the LASSO penalty method. The study showed results with an AUC value of 0.83 and an accuracy of 84.18% [15]. However, the weakness of this research is that it does not show more specific variables such as competitor factors to produce more accurate Churn predictions and produce more precise policies. Jain examined Churn's prediction of one of the telecommunications companies in the United States using logistic regression and logit boost methods. The results of the two techniques are not much different, logistic regression has an accuracy of 85.24 % where the logit boost also has an accuracy of 85.18% [8]. The weakness of this research is that it is not discussed related to efforts to increase accuracy so that there are significant differences between the 2 models tested. Verbeke stated that 6-8 variables are generally sufficient to predict Churn with high accuracy [18]. However, there are still very few studies that focus on the selection of variables to predict telecommunication Churn subscribers. Based on the research that has been done previously, further research is needed in the selection of variables and can find out the variables that are significant to the model. The methods used are prediction methods combined with a penalty or boosting. The output of this research is a predictive model with the best accuracy that can be used for churn reduction policy analysis.

The results of the analysis will be used to determine whether existing policies can be continued or need to be evaluated. Thus, an effective policy can be produced in reducing Churn customers.

II. LITERATURE REVIEW

A. *Churn Customer*

Companies need to be aware of the increasing competition between competitors. Companies must carry out new marketing strategies to meet customer needs, improve satisfaction and retain customers to win the competition [9]. The following are examples of reasons customers to unsubscribe:

1. The customer changes his residential address to a place that does not allow the company to provide its services.
2. Corporate customers have changed their business so they don't need service.
3. Customers feel that the quality of service is less than optimal. The service is often interrupted and takes a long time to handle.
4. Customers switch services to competitors. Competitors offer lower prices and more attractive offers.
5. Customers feel that the service is not by the salesforce agreement information. Packages and prices that have been installed are not by the agreement with salesforce.
6. Customers feel that customer expectations are not being met in terms of service speed.
7. Customers experience bills that are not by the agreement or often experience tariff increases.

B. *XGBoost Classifier*

Here is how the XGBoost classifier works, as shown in Fig. 1. XGBoost is an advanced implementation of the gradient boosting algorithm. Xboost uses the ensemble principle, which combines several weak learner sets (trees) into a strong model to produce strong predictions. XGBoost has many advantages including, it can perform parallel processing which can speed up computation, has high flexibility of objective setting, built-in cross-validation, and overcomes split during negative loss. With these advantages, XGBoost is very suitable for processing classification data. XGBoost will create a tree as a way to classify train data so that a specific target can be obtained. XGBoost has several parameters that can be set so that it can adjust to the dataset obtained. Parameter tuning in XGBoost can be done using gridSearchCV and tuning manually.

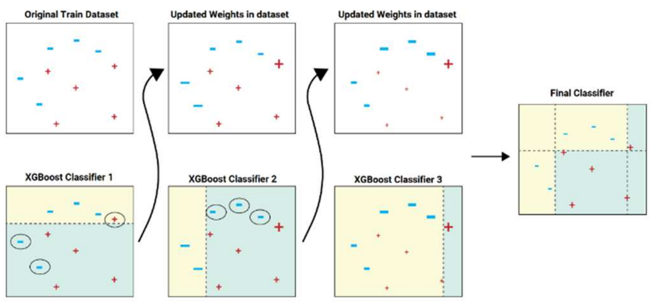

Fig. 1. How XGBoost Classifier Works

Parameter tuning is an activity to determine the right parameters to get the XGBoost classifier algorithm that has high accuracy. Parameter tuning is done by changing the parameters given to the XGBoost classifier algorithm. From the given parameters, the prediction accuracy is evaluated. The highest accuracy is when certain given parameters are taken to be the ideal parameters for the XGBoost classifier algorithm.

Parameters in the XGBoost classifier algorithm include:

1. *learning_rate*
   learning rate is the learning rate performed by the XGBoost classifier algorithm. The level of learning affects the XGBoost classifier algorithm in making a classifier in the form of a tree.
2. *n_estimators*
   n_estimators is the number of trees created by the classifier.
3. *max_depth*
   max_depth is the maximum depth of the tree.
4. *min_child_weight*
   min_child_weight is the minimum amount of weight (hessian) required on child nodes.
5. *gamma*
   gamma is the minimum loss required to partition the nodes in the tree
6. *subsample*
   subsample is the number of samples selected to construct the tree. subsampling will be done every boosting iteration [11].

C. *Churn Management Analysis*

Offering retention incentives is a strategy to keep customers from becoming Churn. Customers who will Churn are offered Churn retention incentives. Customers will churn marked by the customer's initiative to contact customer service or by delaying payment. Churn retention incentive policy analysis of telecommunications companies is usually carried out using the marketing mix method (Saputra, 2021).

With the Churn prediction model and marketing mix analysis, management strategies to reduce Churn become more optimal. The strategy begins with looking at the results of the prediction model. The results of the prediction model are customers who are predicted to churn or not and show significant variables to support the analysis of the causes of customers to churn. Then proceed with the analysis of the marketing mix. For example, given an extreme discount promo if the customer is indicated to be churn because of the price and aggressive competitors. It is hoped that the prediction model will be useful as an initial step in knowing whether the customer pattern will Churn or not. If the company is more responsive to customers who are indicated to be Churn, then the company needs to immediately provide promo incentives according to customer expectations so that customers will agree to continue to subscribe. In other words, the prediction model is expected to slow down the Churn rate in telecommunication companies.

III. Research Methodology

A. Phase 1 (Business Understanding)

The first step in this research is to identify the needs of the business or organization based on the problems faced. The obstacle that is often encountered by telecommunications companies is the Churn rate increasing by 30-50% per year while the offer of retention incentives is rejected by Churn prospective customers. From 2015 to 2021, the telecommunications companies lost an average of 15,000 subscribers every month. Every year telecommunications companies experience an increasing Churn rate. One of the factors causing this is that the company is late in offering Churn retention incentives because customers have subscribed to competitors. Based on these problems, the research objective to be achieved is to create a Churn customer prediction model with high accuracy.

B. Phase 2 (Data Understanding)

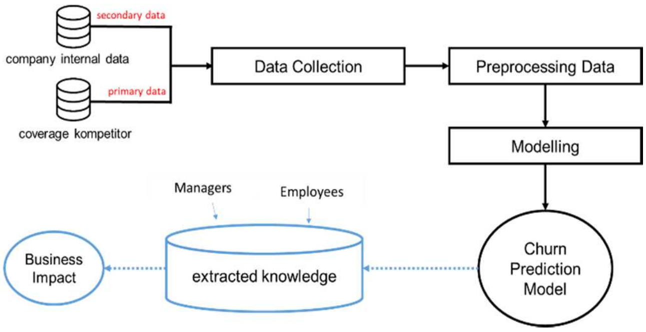

Fig. 2. Knowledge extraction process from data sources

In this study required research dataset under study. The dataset is used as a modeling material that is applied and as an evaluation material. The dataset that will be used is secondary data which is the company's internal data (customer data in 2020) and primary data obtained from several official websites of internet service providers in Indonesia (active areas and competitor brochures). More specifically, the data extraction picture is shown in Figure 2.

C. Phase 3 (Data Preparation)

The following are several stages of dataset preprocessing that will be carried out in this study,

1. Data Cleaning

The original dataset in research, often encounters some obstacles such as some missing or empty data. This is if the lost data is entered into Subsequent modeling will produce less accurate models. Lost data will be imputed data. Data imputation is a way of dealing with missing data with possible values based on available information.

2. Data Transformation

Data transformation is used to change data in a suitable form in the data mining process. Several techniques for data transformation are normalization, attribute selection, and discretization. Normalization is done to scale data values within a certain range of values, for example -1 to 1 or 0 to 1. The second technique is attribute selection. Attribute selection is the process of selecting the attributes given for the data mining process. The last is the discretization technique. This technique is done to replace the raw value of the numeric attribute with an interval value.

3. Data Partition

The research dataset will be partitioned. Data partitioning is the division of the dataset into two groups of datasets, namely the training dataset and the testing dataset. The training dataset provides the raw material for building predictive models . Testing data is used as a predictive model performance measurement . The training data used is 70% and the testing data used is 30%.

D. Phase 4 (Modeling)

At this stage the training data will be built a predictive data model . The results of the model can predict which customers will Churn from the training data. The exploration of predictive models is listed in Figure 3.

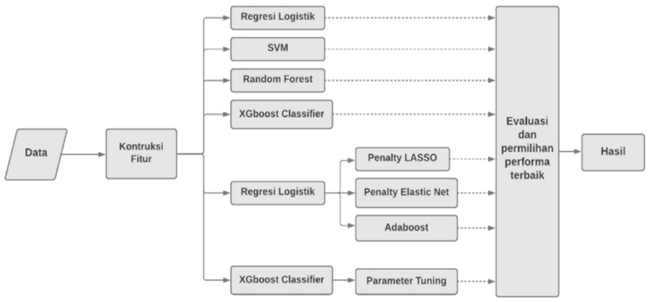

Fig. 3. Method exploration flow

E. Fase 5 (Evaluation)

The knowledge model generated from the data mining process must also be evaluated and interpreted based on certain measures. In evaluating the achievement of the predictive data model, it can be investigated using a confusion matrix to produce accuracy, sensitivity, specificity and AUC.

1. Confusion Matrix

In Churn customer prediction research, a confusion matrix is used to determine the Churn customer probability prediction or not. This is done by using a cut value of 0.5 or $(Y) = 0.5$. That is if the probability value of (Y) is <0.5 then the customer is categorized as not going to Churn and if the probability value of (Y) is 0.5 then the customer is predicted to Churn. The confusion matrix produces accuracy, precision, and recall. Churn is a condition where the customer stops using telecommunications services. Non-Churn are customers who continue to actively subscribe to telecommunication services.

2. AUC

$$AUC = \frac{1}{2}\sum_{i=1}^{n}(x_{i+1} - x_i)(y_{i+1} - y) \qquad (1)$$

AUC is a square-shaped area whose value is always between 0 and 1. For testing using equation (1) from the AUC area formula. x is 1-specificity while y is sensitivity . The best level of model performance can be seen from the AUC value with the criteria in Table 1.

TABLE I. CRITERIA FOR AUC VALUE [5]

AUC Score	Interpretation
0.9-1.00	*Excellent Classification*
0.8-0.9	*Good Classification*
0.7-0.8	*Fair Classification*
0.6-0.7	*Poor Classification*
0.5-0.6	*Failure*

IV. RESULT AND DISCUSSION

A. Selection of the best prediction model

From existing research, prediction accuracy on telecommunication company data is still in the range of 85-87%. Although these results can be said to be good, for the rapid development of the telecommunications world, especially during the pandemic, other methods are needed to improve the accuracy of churn prediction.

In this study, the research problem that is focused on is the Churn rate of telecommunications companies continues to increase every year and the competition for internet service providers is getting tighter. The dataset used is different from previous research. However, the selection of variables refers to previous research with an update plan, namely the addition of a competitor area factor variable.

Tested 1000 customer data with 8 prediction models. The data is done by split technique. 700 customer data as training data and 300 customer data as testing data. The model is then evaluated for classification. The way to determine the feasibility of the model can be seen through the accuracy of the model in classifying the data. The higher the classification accuracy of the obtained model, the better the model that has been obtained. The best model was selected based on the AUC, accuracy, precision, and recall parameters. The model obtained is evaluated using data testing. The performance results of the 8 methods can be seen in Table II.

TABLE II. SELECTION OF THE BEST PREDICTION MODEL

No	Model	AUC	Accuracy	Precision	Recall	Output
1	Regresi Logistik	0.931	89.00%	90%	97%	function
2	SVM	0.931	82.00%	82%	**100%**	non function
3	Random Forest	0.964	91.30%	92%	**98%**	non function
4	XGBoost Classifier	0.963	92.30%	94%	97%	non function
5	Logistic Regression Penalty LASSO	0.935	89.67%	92%	96%	function
6	Logistic Regression Penalty Elastic Net	0.929	89.30%	91%	96%	function
7	Logistic Regression Adaboost	0.918	86.00%	87%	97%	non function
8	XGBoost Classifier parameter tuning	**0.968**	**94.30%**	**96%**	**98%**	non function

* the bold note indicates the highest number

Table II shows a comparison of the performance of the eight test methods. Logistic regression gets the model performed very well. The logistic regression algorithm has the advantage that the independent variables in logistic regression can be a mixture of continuous, district, and dichotomous variables. Logistic regression does not require the limitations of the independent variables and does not require that the independent variables be in the form of intervals. Logistic regression uses a sigmoid function that can be used to transform the predicted value to the probability value, the probability value is in the range 0-1. 0 indicates active and 1 indicates Churn. Then the optimization method is carried out using penalty LASSO, Elastic Net, and Adaboost. The LASSO penalty logistic regression uses the maximum likelihood estimation function which functions to estimate the parameters in the model and the LASSO term is added to reduce variables that do not affect the model. The difference in function in the LASSO penalty logistic regression and logistic regression, makes the performance of the LASSO penalty logistic regression model better than the logistic regression, with an increase in accuracy of 0.67%. Furthermore, the logistic regression method was optimized using a penalty elastic net. Penalty elastic net logistic regression uses the maximum likelihood estimation function and the term elastic net is added to reduce variables that do not affect the model. The term elastic net is a combination of penalty from the ridge which can handle high correlation problems and LASSO which has variable selection properties. But the performance of the model is not better than the LASSO penalty logistic regression. The difference is 0.37%. Furthermore, the logistic regression method was optimized using Adaboost. The sigmoid function to get the coefficients is then classified into the second using Adaboost. The disadvantage of Adaboost is that trees are made of only one branch and two leaves (stumps). Stumps aim to reduce errors. Stumps that are built do not have the same weight on the final prediction. Thus, stumps that have large errors have little effect on decision-making. Thus, the performance results of the logistic regression classification with Adaboost are lower than the logistic regression. The difference in performance is 3%. It can be concluded that the LASSO penalty logistic regression is the model with the best performance from a comparison of 3 logistic regression optimization models for the Churn prediction case of telecommunications companies.

Furthermore, the SVM method produces the lowest model performance compared to the other 8 test methods. SVM classification uses a subset of training points and has a high training time so it is not suitable for large data sets. Then, the random forest method produces the best third-order model performance out of 8 test methods. The random forest algorithm works by building several trees and combining them to get a more stable and accurate prediction. A 'forest' built by a random forest is a collection of trees that are usually trained with the bagging method or a combination of learning models to improve overall results.

Next is the model performance of the XGBoost classifier method. It has advantages over the random forest. XGBoost can perform parallel processing to speed up computing, has high flexibility in setting objectives, built-in cross-validation, has regularization features, and overcomes splits during negative loss. Thus, XGBoost gets better model performance than the random forest and other test methods. Further, the XGBoost method was optimized using tuning parameters. XGBoost parameter tuning combines several weak learner sets (trees) into a strong model to produce strong predictions. The tuning parameters used are learning_rate carried out by the XGBoost algorithm of 0.01, n_estimators or the number of trees created by the classifier as many as 1000, max_depth or the maximum tree depth of 15.

Table 2 shows the highest AUC value of the 8 methods is the XGBoost model with a tuning parameter of 0.968. In addition, the XGBoost model with tuning parameters has high accuracy, specificity, and sensitivity values. The highest accuracy is 94.3%, the highest specificity is 96% and the sensitivity is 98%. So, the best model using testing data is XGBoost with tuning parameters. The performance results are obtained from the exploration of the tuning parameters which can be seen in Table 3.

TABLE III. EXPLORATION OF TUNING PARAMETERS

Metode	Parameter	Auc	Akurasi	Precision	Recall
XGBoost	default	0.963	92.30%	94%	97%
XGBoost	learning_rate =0.01	0.932	89.3%	91%	97%
XGBoost	learning_rate =0.01, n_estimators=1000	0.961	93%	94%	98%
XGBoost	learning_rate =0.01, n_estimators=1000, max_depth=15	0.968	94.30%	96%	98%

XGBoost with default parameters gets the highest AUC from logistic regression, SVM, and random forest methods. Researchers use a hybrid method to increase the performance of the model, namely XGBoost combined with tuning parameters. Obtained the accuracy rose to 94.3%. The performance of the XGBoost model with parameters learning_rate=0.01, n_estimators=1000, max_depth=15 shows the best performance compared to the other 7 test methods.

But the weakness of the XGBoost model cannot show significant variables to the model. Then the best model is selected with the output in the form of a function. Table 2 shows that in the output function category, the highest AUC value is the Logistics Regression model with Penalty LASSO. In addition, the model also has high accuracy, specificity, and sensitivity values. Logistics Regression Model with LASSO penalty reduces 2 variables to be insignificant to the model. With the shrinkage of variables that are not significant to the model, the LASSO penalty logistic regression shows the AUC value and accuracy are higher than the Logistics Regression model. The variables that have a significant effect on the model are *X1* (account number), *X3* (rate), *X4* (age), *X5* (customer type), *X6* (score), *X7* (subscription age), *X8* (product), *X9* (area), *X10* (Iconnet competitor) and *X11* (Biznet competitor). Next, the variable exploration will be carried out on the XGBoost tuning parameter model.

TABLE IV. EXPLORATION OF THE SIGNIFICANT VARIABLES OF THE XGBOOST TUNING PARAMETER MODEL

Exclude Variabel	AUC
Account number	0.937
Age	0.949
For_Churn	0.960
Product	0.961
Rate	0.965
Biznet	0.966
Score	0.967
Iconnet	0.967
Area	0.967
Ctype	0.968

Table 4 shows the accuracy of each test excluding one of the variables. Table 4 shows 9 out of 10 variables that affect the prediction model. If one of the data is removed then the model will reduce its AUC. But when removing the Ctype variable (customer type) the AUC remains the same. Then Ctype does not affect the XGBoost model tuning parameters. 9 variables that are significant to the model can provide a predictive model output in the form of a prospective customer's decision to churn or not. In Table 4, it can be seen that the performance of account numbers and the length of delay in payment (age) are the variables that most influence the model. Variables of the length of deferred payment (age) and age of subscription (for_churn) can assist in process analysis. Product and rate variables (average billing per month) can support product analysis. Competitor area variables can support competitor analysis.

B. Churn Management Analysis

The modeling results can predict which customers will Churn. Another result of modeling is that significant variables can be found in the model. The variables that are significant to the model represent the habits and groups of customers who frequently churn. The variables in Table 4 are one of the bases for making decisions in the retention-intensive bidding policy to reduce Churn. Prospective Churn customers must take precautions so as not to Churn. The way to prevent customers from Churn is by predicting customers who will Churn and providing intensive retention early (not waiting for customer complaints). This step shows the company is more responsive to cases experienced by customers. Intensive retention policy analysis uses the marketing mix analysis method.

V. CONCLUSION & FUTURE WORKS

Based on significant variables in the model, the characteristics of the company's data such as the length of delay in payment, products, prices, scores, and external data such as the competitor's area are proven to affect the prediction model. XGBoost parameter settings combine several weak learner sets (trees) into a robust model resulting in strong predictions. Parameter settings used are learning_rate or the level of learning carried out by the XGBoost algorithm of 0.01, n_estimators, or the number of trees created by the classifier as much as 1000, max_depth, or maximum tree depth of 15. Of the 8 prediction model tests, the Churn prediction model for telecommunications companies with hybrid method XGBoost parameter tuning shows the best performance with AUC 0.968, accuracy 94.3%, recall 98%, and precision 96%. This figure shows the very good performance of the prediction model.

The results of the modeling are in the form of a classification of customer churn predictions or not and significant variables. Significant variables can assist Churn management analysis. In process analysis, the strategy carried out on Churn's predicted data, the company took the initiative to ask for obstacles and provide chat-based customer service tools to facilitate communication with customers. This is to overcome customers experiencing service problems but are not willing to call customer service. In product analysis, all First Media products have higher prices than their competitors. The proposed strategy on customer data that is predicted to Churn, the company provides retention incentives by providing package promo prices according to competitors' areas. But the price offered needs to be adjusted to a business scheme that does not harm the company. In competitor analysis, the proposed strategy is the existence of tools that can recap external data related to network areas covered by various providers. The existence of such external data is the first step to preventing customers from churn due to competitor factors.

For the future work, based on the results obtained, the solution to the increasing Churn rate problem can be solved by predicting which customers will Churn using the XGBoost parameter tuning model. Research can be continued on competitor's area tools to support business development.

REFERENCE

[1] Baragoin, C. Mining Your Own Business in Telecoms Using DB2 Intelligent Miner for Data. IBM Redbooks. 2001.

[2] Chapman, Pete et al. Crisp-Dm 1.0 : Step-by-Step Data Mining Guide. www.spss.com/worldwide. www.spss.com/worldwide. 2000.

[3] Deng, Y., Li, D., Yang, L., Tang, J., & Zhao, J. Analysis and prediction of bank user Churn based on ensemble learning algorithm. Proceedings of 2021 IEEE International Conference on Power Electronics, Computer Applications, ICPECA 2021, 288–291.https://doi.org/10.1109/ICPECA51329 .2021.9362520. 2021.

[4] Euler, T. Publishing operational models of data mining case studies. In Proceedings of the Workshop on Data Mining Case Studies at the 5th IEEE International Conference on Data Mining (ICDM) (pp. 99-106). 2010.

[5] Gorunescu, F. Data Mining: Concepts, models and techniques (Vol. 12). Springer Science & Business Media. 2011.

[6] Hadden, J., Tiwari, A., Roy, R., & Ruta, D. Churn prediction using complaints data. In Proceedings of world academy of scienc. 2006.

[7] Hu, X., Yang, Y., Chen, L., & Zhu, S. Research on a Customer Churn Combination Prediction Model Based on Decision Tree and Neural Network. 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics, ICCCBDA 2020, 129–132. https://doi.org/10.1109/ICCCBDA49 378.2020.9095611. 2020.

[8] Jain, H., Khunteta, A., & Srivastava, S. Churn Prediction in Telecommunication using Logistic Regression and Logit boost. Procedia Computer Science. 2020.

[9] J. Zhong, and W. Li. "Predicting customer churn in the telecommunication industry by analyzing phone call transcripts with convolutional neural networks." In Proc. 2019 3rd International Conference on Innovation i n Artificial Intelligence, pp. 55-59. 2019..

[10] Lu, N., Lin, H., Lu, J., & Zhang, G. A customer Churn prediction model in telecom industry using boosting. IEEE Transactions on Industrial Informatics, 10(2), 1659–1665. https://doi.org/10.1109/TII.2012.2224355. 2014.

[11] Muslim, F. Penerapan Brute Force dan Decrease and Conquer pada Parameter Tuning XGBoostClassifier. https://informatika.stei.itb.ac.id/~rina ldi.munir/Stmik/2018-2019/Makalah/Makalah-Stima-2019-043.pdf

[12] Montgomery, D. C., & Peck, E. A. Introduction to Linear Regression Analysis, Wiley. New York, USA. 1992.

[13] Mozer, M. C., Wolniewicz, R., Grimes, D. B., Johnson, E., & Kaushansky, H. Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. IEEE Transactions on neural networks, 11(3), 690-696. 2000.

[14] Navlani, A. AdaBoost Classifier in Python. https://www.datacamp.com/ community/tutorials/ada boost-classifier-python. 2018.

[15] Saputra, F. D. Prediksi Churn dan strategi retensi pada kasus perusahaan telekomunikasi. Tesis. 2021.

[16] Senthan, P., Rathnayaka, R., Kuhaneswaran, B., & Kumara, B. Prediksi Keputusan Pelanggan Menggunakan Extreme Learning Machine Pada Data Telco Customer Churn. 3(11), 8. 2021.

[17] Tisantri, D. H., Wihandika, R. C., & Adinugroho, S. Prediksi Keputusan Pelanggan Menggunakan Extreme Learning Machine Pada Data Telco Customer Churn. 3(11), 8. 2019.

[18] Verbeke, W., Dejaeger, K., Martens, D., Hur, J., & Baesens, B. New insights into Churn prediction in the telecommunication sector: A profit driven data mining approach. European Journal of Operational Research, 218(1). 2012.

[19] Xu, T., Ma, Y., & Kim, K. Telecom Churn prediction system based on ensemble learning using feature grouping. Applied Sciences (Switzerland), 11(11). 2021.

[20] Yabas, U., Cankaya, H. C., & Ince, T. Customer Churn prediction for telecom services. Proceedings - International Computer Software and Applications Conference, 358–359. 2012.