Prediction Analysis for Business To Business (B2B) Sales of Telecommunication Services using Machine Learning Techniques

Oryza Wisesa[1*], Andi Adriansyah[2], Osamah Ibrahim Khalaf[3]
1- Department of Electrical Engineerng, Universitas Mercu Buana, Jakarta, Indonesia.
Email:oryza49@gmail.com (Corresponding author)
2- Department of Electrical Engineerng, Universitas Mercu Buana, Jakarta, Indonesia.
Email: andi@mercubuana.ac.id
3- College of Information Engineering, Al-Nahrain University, Baghdad, Iraq.
Email: usama.ibrahem@coie.nahrain.edu.iq

Received: July 2020 Revised: September 2020 Accepted: October 2020

ABSTRACT:
Sales prediction analysis requires intelligent data mining techniques with accurate prediction models and high reliability. In most cases, business highly relies on information as well as demand forecast of the sales trends. This research uses B2B sales data for analysis. The B2B data could provide information on how telecommunication company should manage its sales team, products, and budgeting flows. The accurate estimates enable Telecommunication company to survive the market war and increase with market growth. Comprehensible predictive models were studied and analyzed using a technique of machine learning to improve the prediction of the future sale. It is hard to cope with big data and sale prediction accuracy if the system of traditional forecast is used. In this study, machine learning technique was also used to analyze the reliability of B2B sales. In addition, at the end of this research, other measures and techniques used to predict sales were introduced. The predictive model with best performance evaluation is recommended to forecast the trending B2B sales. The study results are put into an order of reliability and accuracy of the best method to predict and forecast including estimation, evaluation, and transformation. The best performance model found was Gradient Boost Algorithm. The result form graph the data close together from beginning till end of data target MSE and MAPE result are the best result than other method, MSE =24.743.000.000,00 and MAPE =0,18. This model performed maximum accuracy in predicting and forecasting of the future B2B sales.

KEYWORDS: Machine Learning Techniques, Sales Forecasting, Prediction, Telecommunication, Reliability, B2B (Business to Business).

1. INTRODUCTION

Sales forecast analysis requires intelligent data mining techniques with accurate prediction models and high reliability. Sales estimates provide data on how a company should manage its sales team, products and also budgeting flows. Accurate estimates enable organizations to increase in accordance with growth of the market in the possible maximum level of income. To achieve effective process of changing big data into information, which is useful for predicting cost and sales, data mining is commonly used. Therefore, this research will focus on comparison of prediction analysis and B2B sales reliability sales using machine learning techniques. The formulation of the problem in this study is how to determine the type of prediction analysis that is precise and accurate by using comparison between four machine learning techniques in B2B sales with 2016-2017 & 2018 data.

2. RELATED WORK

Some of forecasting strategies and prediction methods were analyzed in [1]. Data tuning and algorithm clustering on sales data were compared in a study by [2]. Data classification plays an important role in making decision as studied in [3]. It is effective to use clustering technique to find the best pattern of distribution, while clustering algorithms use the similarity measure based on the distance metric. The information retrieved from a large data can be successfully converted into a rational format if it uses suitable technique of data mining. In addition, either using supervised or unsupervised learning can be conducted [4].

Paper type: Research paper
DOI: https://doi.org/10.29252/mjee.14.4.145
How to cite this paper: O. Wisesa, A. Adriansyah and O. Ibrahim Khalaf, **"Prediction Analysis for Business To Business (B2B) Sales of Telecommunication Services using Machine Learning Techniques"**, *Majlesi Journal of Electrical Engineering*, Vol.14, No. 4, pp. 145-153, 2020.

Effective business decision can be achieved if it uses suitable technique of sales prediction. In this case, both algorithms and concepts are conducted in [5]. This study employed some technique of data mining to conduct sales prediction for B2B products. This includes sale prediction of any products to get accustomed to the task already studied before.

From this problem, several researches have arisen to help prediction and forecasting sales. The research tried to develop system with analysis using predictive machine learning techniques.

Brian D'Arcy, Claire Gallagher, Michael G. Madden [6], in 2015, conducted a research on a lot of businesses relying on the techniques of forecasting to examine whether the business organization can achieve sales opportunities or not. Alternative methods were developed to calculate sale opportunity presenting 3 (three) differences: first, qualitative assessment accompanied with quantitative data to define the opportunity attributes; second, replacing the weight factor with the Augmented Naïve Bayes Tree (TAN) classifier which is able to figure out dependencies between variables and produce probabilistic results where thresholds are implemented; the last, TAN is studied using data retrieved from the past, while the current QSP has a static weight. This approach has 90.6% accuracy in examining if the sales will be successfully won or not. In this case, a significant increase in the accuracy of the current approach is around 75.6%.

In 2018, Kaftannikov Igor Leopoldovich, Parasich Andrey Viktorovich, Parasich Viktor Aleksandrovich [7], conducted research on the solution for estimating House Prices with the Regression Technique that aims to predict housing prices based on attribute values such as the area of the house, building year, etc. In this study, the algorithm of classic machine learning and original method was used and explained. Ching-Seh (Mike) Wu, Pratik Patil, Saravana Gunaseelan [8] conducted research on sales on Black Friday (discounted days) predictions to develop algorithm which is accurate and efficient for analyzing customers' spending and expenditure in the past to see same features the customers may present in the future. This study implemented various techniques of machine learning such as neural networks and regression in constructing predictive model and compared the prediction performance and accuracy. To achieve the best performed prediction, various algorithms and platforms were applied in this technique; in this case, there were 7 (seven) machine learning algorithms used. In the same year, Sunitha Cheriyan, Shaniba Ibrahim, Saju Mohanan, Susan Treesa [9] conducted a study on the model of sales data and estimates using several techniques and sizes for predicting sales. In performance evaluation, the best model of prediction is recommended for estimating sales trends. The study results focused on consistency and precision of the efficient techniques for forecasting and prediction. It was found in a study that the best model to be used was the Gradient Boost Algorithm, indicating maximum precision in forecasting and predicting the future sales.

In 2019, Irfan Ullah et al. [10] proposed a model of churn prediction using classification and clustering techniques to detect churn customers and show influencing factors of churn customers in the sector of telecommunications. Feature selection is conducted by gaining information and ranking attribute correlation filter. The first model ranks data of churn customer by implementing a classification algorithm in which of the algorithm, i.e. Random Forest shows good performance with 88.63% examples are correctly classified. CRM has an essential responsibility to avoid churners by implementing an applicable retention policy. Once it has been classified, the model divides stirring customer data by classifying those who stir in groups employing cosine similarity in order to retention offers based on group. In addition, churn factor which plays important roles to see the cause of churn was identified. CRM may boost productivity, suggest related promotion derived from the pattern of behaviors, and improve marketing campaign; all of those can be done by identifying the essential churn factors taken from the data of customers. The evaluation of churn prediction can implement some metrics including precision, accuracy, f-size, recall, and ROC (Receiving Operating Characteristics) areas. The results indicate that RF algorithm was appropriate to obtain churn classification from its model, while k-means clustering was best for customer profiles. In addition, the influencing factors affect the customer churn by using some rules which are produced from the attributes of classifier algorithm.

3. RESEARCH METODOLOGY

This study aims to assess and analyze machine learning techniques for sales prediction, to know which machine learning techniques are more reliable for prediction forecasting B2B Sales.

A. *Data Collection and Preparation*

Data of B2B sales collected from the sales data of 2016, 2017, and 2018 were used in this study to predict the future sales. In this case, the databases included were Category, City, Type of items and its opportunity-ID, Quarter, Product Name, Sub Service Product, Service Product (MIDI or Non-MIDI, and Sales Revenue. The inessential and irrelevant data and redundant entries were removed from the initial large data, and it produced dataset of Quarter, Year, Service Product and Quantity [11].

B. *Exploratory Analysis*

Exploratory analysis consisting of some steps for data mining was done to comprehensively realize the nature of data studied [12]. Fig. 1 shows six key steps in the process of data mining.

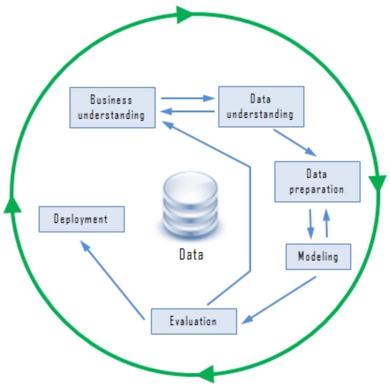

Fig. 1. Data mining process.

It can be seen from the figure above that data mining processes are understanding data, preparing, modelling, evaluation, and deploying.

The collected data describing the sales can be seen in the following table.

Table 1. Yearly sales by product.

Sales Quarter / Product	Y2016 MIDI	Y2016 NON MIDI	Y2017 MIDI	Y2017 NON MIDI	Y2018 MIDI	Y2018 NON MIDI
	Year of Data					
Q1	976 M	228 M	648,047 M	122,800 M	1,937,887 M	850,250 M
Q2	1,031 M	4,292 M	572,748 M	138,784 M	2,112,247 M	764,069 M
Q3	13,875 M	248 M	1,344,548 M	154,728 M	2,099,508 M	1,554,222 M
Q4	71,547 M	4,191 M	1,642,016 M	227,790 M	1,629,096 M	2,342,905 M

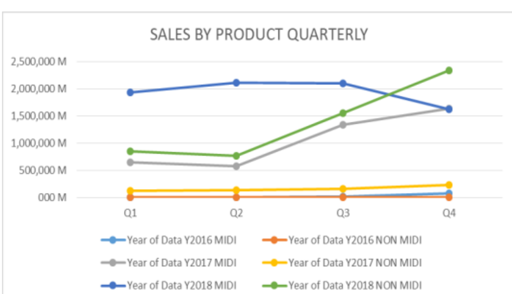

Fig. 2. Yearly sales by product.

From the figure above, we can see the total revenue of the sales from 2016 until 2018 reported quarterly as indicated Q1, Q2, Q3, and Q4, respectively. It also shows that MIDI and NON-MIDI products do not always show similar increase.

C. *Outlier detection*

Outlier detection performs well for both data preprocessing and model optimization. Furthermore, it can employ the model, act as the initiation of the next optimization, and help display independent information of the models. It focuses more on the quality of each data attributes. In addition, it discards the data attribute which has less value. Data means a set of data after being transformed for modelling, while correlation means a pattern of connection among attributes of the sales revenue showing positive correlation. This pattern or matrix can be seen in the following table.

Table 2. Correlation matrix.

	Quarter	*Product*	*Quantity*	*Revenue*
Quarter	1	0.183809187	0.000935759	0.008941
Product	0.183809187	1	0.006986816	0.020283
Quantity	0.000935759	0.006986816	1	0.050925
Revenue	0.008941057	0.020282845	0.050925088	1

D. *Forecasting and Trends*

Both Figs. 3 and 4 indicate the trends obtained by classifying MIDI and NON-MIDI products. In this case, the forecasting represents future sales in the Quarter 1 to 4 from 2016 to 2021, while the trends indicate the total revenue until the current quarter. Blue dots and lines represent the actual sales made while the orange dots and lines represent sale estimation, each quarter indicating a little bit increase of sales.

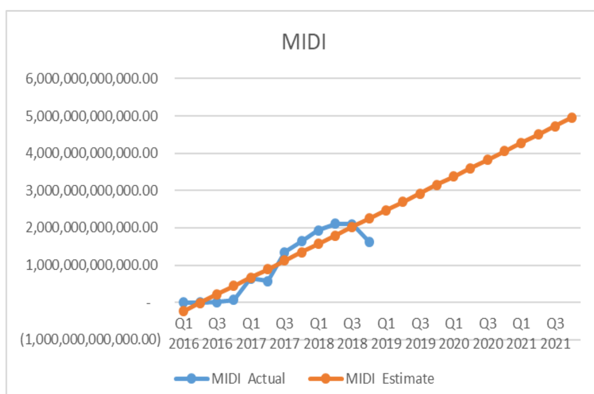

Fig. 3. Forecast for five years MIDI product.

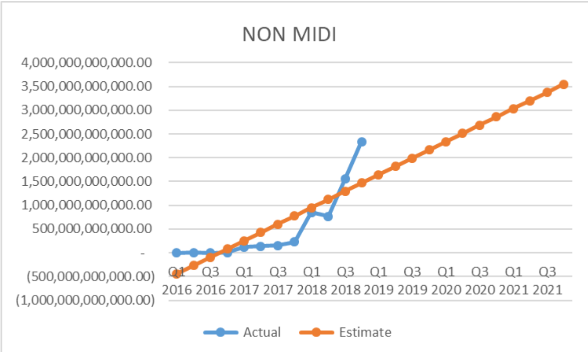

Fig. 4. Forecast for five years NON-MIDI product.

It is also produced by classifying the total products sold and revenue for each quarter into cluster quantity and revenue. The trend lines in Figs. 5 and 6 show more MIDI product but less quantity.

MAD, MSE, RMSE, and MAPE should be calculated and evaluated to improve the forecast results in terms of smoothening exponent, moving average, etc. In addition, it defines how accurate the trend makes prediction for future sales.

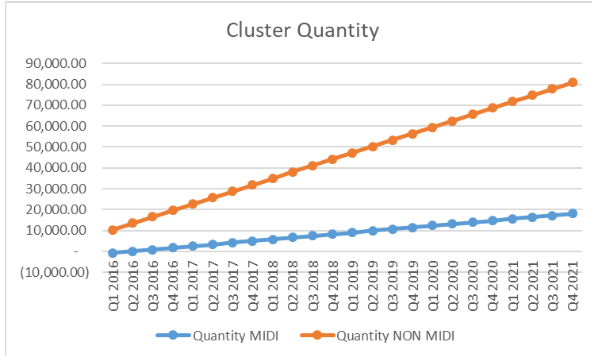

Fig. 5. Trend analysis using clusters quantity.

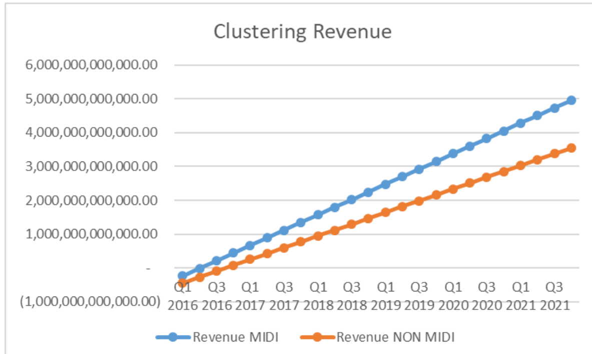

Fig. 6. Trend analysis using clusters revenue.

Table 3. Forecast table.

Period	Actual	Forecast	Error	Absolute Value of Error	Square of Error	Absolute Values of Errors Divided by Actual Values.
t	**At**	**Ft**	**At -Ft**	**| At -Ft|**	**(At -Ft)^2**	**| (At - Ft)/At|**
Q1 2016	1,203 M	-676,249 M	677,452 M	677,452 M	458,941,129,576,390B	562.978203
Q2 2016	5,324 M	-276,960 M	282,284 M	282,284 M	79,684,242,583,633B	53.02570361
Q3 2016	14,124 M	122,328 M	-108,204 M	108,204 M	11,708,119,865,150B	7.661197668
Q4 2016	75,737 M	521,616 M	-445,879 M	445,879 M	198,807,652,897,953B	5.887167771
Q1 2017	770,847 M	920,904 M	-150,057 M	150,057 M	22,516,972,605,367B	0.194664406
Q2 2017	711,532 M	1,320,192 M	-608,660 M	608,660 M	370,467,072,062,137B	0.855421756
Q3 2017	1,499,276 M	1,719,480 M	-220,204 M	220,204 M	48,489,887,965,007B	0.146873673
Q4 2017	1,869,806 M	2,118,769 M	-248,962 M	248,962 M	61,982,243,485,751B	0.133148738
Q1 2018	2,788,137 M	2,518,057 M	270,080 M	270,080 M	72,943,140,839,484B	0.096867522

Q2 2018	2,876,316 M	2,917,345 M	-41,029 M	41,029 M	1,683,346,612,449B	0.014264289
Q3 2018	3,653,731 M	3,316,633 M	337,098 M	337,098 M	113,634,981,190,092B	0.092261278
Q4 2018	3,972,002 M	3,715,921 M	256,081 M	256,081 M	65,577,312,028,869B	0.064471439
Total	18,238,035 M	18,238,035 M	0.003051758	3,645,989 M	1,506,436,101,712,280B	631.1502452

MAD	303,832,391,369.09
MSE	125,536,341,809,357,000,000,000.00
RMSE	354,311,080,562.49
MAPE	52.60

E. ***Prediction***

To see the probable event occurring in the future, prediction may play important role in it. In this case, algorithms of machine learning improve the intelligence of prediction system without demanding intervention manually. Ethem Alpaydin [13] states that Machine Learning (ML) is able to improve criterion of performance employing data retrieved from the past experiences.

In fact, any discipline may apply a technique called machine learning utilizing statistic calculations to get a lot of classification and clustering problem solved. These statistics include three types namely supervised, unsupervised, and semi-supervised statistics. Four algorithms of machine learning are discussed in this paper; they are Decision Tree (DT), Generalized Linear Model (GLM), Random Forest, and Gradient Boost Tree (GBT).

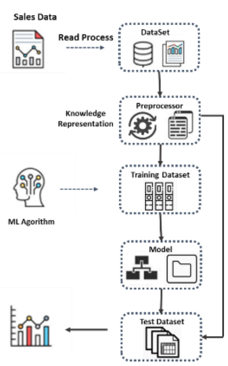

Fig. 7. System Architecture.

This study employed 4 (four) algorithms of machine learning and tested the model to see the accuracy performance in which the model with the best accuracy were selected.

1) ***Generalized Linear Model***

This model indicates a big class of linear regression model [14] which is conventional and focusing on the uninterrupted response where the variable shows uninterrupted categorical predictors [15]. Random component in this case probability distribution of response variable (Yi) and linear predictor are used in this model, as can be seen in the following formula:

$$\eta i = \alpha + \beta 1Xi1 + \beta 2Xi2 + \cdots + \beta kXik \quad (1)$$

A smooth and invertible linearizing link functions g(.), which transforms the

$$\mu i = E(Yi) \quad (2)$$
$$g\ (\mu i) = \eta(i) = \alpha + \beta 1Xi1 + \beta 2Xi3 + \cdots + \beta kXik \quad (3)$$

It provides estimated coefficient of regression and asymptotic standard errors of coefficient. Meanwhile, the parameter of dispersion in GLS tends to be fixed to a numeric value 1[16].

2) ***Decission Tree***

This model is a powerful analysis for multiple variables, strong tool for data mining, and classifier of recursive partition in the instant space. The implementation of decision tree can be seen in a lot of domains. This approach indicates the factors to obtain goals and the steps and means of implementation [14]. The objective is indicated in (O) and (Ci) as the steps to follow and (Mij) is the means of action corresponded to these ways, as noted by qi, (i= 1 …. n), meeting the relation.

$$\sum_{i=1}^{n} qi = 1, cu\ qi \geq 0 \quad (4)$$

To get (Mij), the coefficient (a ij) with a number of series 1 for each point.

$$a_{11+}a_{12+}a_{13+\cdots+}a_{1m} = 1,$$
$$a_{21+}a_{22+}a_{23+\cdots+}a_{2m} = 1,$$
$$\ldots + \cdots + \cdots + \cdots = 1$$

$$a_{n1+}a_{n2+}a_{n3+\cdots+}a_{nm} = 1$$
$$\sum aij = 1 \quad (5)$$

Fig. 8 represents the model of decision tree used in this study.

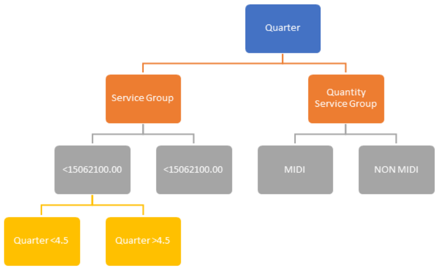

Fig. 8. Decision tree model.

3) *Gradient Boosted Trees*

Gradient boosting serves as a technique of machine learning to solve classification and regression problems. This model can put together learning method using large decision trees in order to make final model for prediction [15]. The construction of this model is based on the principles of collecting weak learners to have a strong learner by boosting the process. This model shows strong method of training to add weak learners into the model; in this case, they serve as decision tree [16].

F(x) can be defined as a full model after t-1 round, while h(x) serves as a new tree which is put to the model

$$F0 = 0 \quad (6)$$
$$ft(x) = Ft - 1(x) + h(x) \quad (7)$$

When there is error found in the model which have been built in the previous step, new function will be proposed to make correction. Thus, the new function(x) proposed should make prediction of the residual FT-1(x). Fig. 9 illustrate this approach proposed in this study.

Fig. 9. Gradient boosted tree model.

4) *Random Forest*

This approach is built from a lot of single decision trees which serve as an ensemble and spit out a classification prediction. When the class gets the highest votes, it is then chosen as the model prediction.

Each uncorrelated tree or model serving as a committee will surpass individual constituent model. In this case, low relationship among the model becomes the key itself. The models which are not correlated may lead into ensemble predictions showing more accurate than individual predictions. It is because the trees are able to defend each other from each individual error (unless constantly error in the same direction). Though there could be wrong and right trees, a group of trees can go to the right direction.

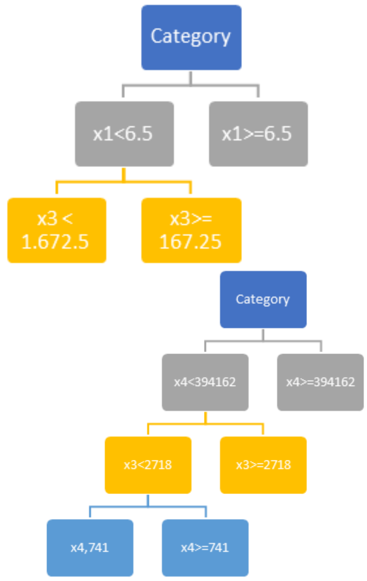

Fig. 10. Random forest model.

5) *Forecast Estimation, Evaluation & Transformation*

MAD, MSE, RMSE, and MAPE should be calculated and evaluated to improve the forecast results in terms of smoothening exponent, moving average, etc. In addition, it defines how accurate the trend makes prediction for future sales.

a) *Mean Absolute Percentage Error (MAPE)*

MAPE determines accuracy noted in the error percentage which is considered easier to understand compared to other statistic calculations.

MAPE can be written in an equation below:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \quad (8)$$

Where, At and Ft means the actual and forecast values, respectively. This calculation sums up the absolute value of the point forecasted in time and divides it by the number of fitted point n, then multiplies it by 100% to produce percentage error.

b) *Mean Absolute Deviation (MAD)*

MAD determines accuracy in similar unit as the data, which assists the conceptualization of error. In this case, outliers have fewer effect on MAD than on MSD. MAD can be written in the equation as follows:

$$MAD = \frac{\sum_{t=1}^{n}|A_t - F_t|}{n} \quad (9)$$

c) Mean Squared Error (MSE)

MSE is used to measure the accuracy of values of fitted time series. In this case, outliers indicate bigger effect on MSE than on MAD. MSE can be written in the equation as follows:

$$MSE = \frac{\sum_{t=1}^{n}(A_t - F_t)^2}{n} \quad (10)$$

d) Root Mean Squared Error (RMSE)

RMSE is used to calculate the gaps between sample or population values predicted by a model and the values that are observed. RMSE expresses the square root of the second sample moment of the gaps between values being predicted and observed or the quadratic mean of these gaps. This deviation can be called residual when the data samples used for estimation are calculated. Meanwhile, when it is computed out-of-sample, this deviation can be called prediction errors. RMSE aggregates the error size in several predictions into a single predictive power measure. RMSE can be used to compare the errors of forecasting from various models for specific dataset.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(A_t - F_t)^2}{n}} \quad (11)$$

e) Performance Metrics (Error Measures)

The goal of this research is to build the type of prediction analysis which is precise and accurate by using comparison between four machine learning techniques in B2B sales with 2016, 2017 & 2018 data. There are 4 methods of machine learning to be used to predict GLM, DT, GBT and Random Forest. We compare the MAPE and MSE is a measure of prediction accuracy of a forecasting method in statistics, for example in trend estimation, also used as a loss function for regression problems in machine learning. Botchkarev [17], Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Topology. There is another group of metrics that do not have dimension and referred to as dimensionless (Dimensionless Quantity, n.d.) or scale-free, scaled, or scale-independent. Commonly, dimensionless metrics involve mathematical division of quantities of the same dimensional units (e.g. ratios, relative, percentage indicators), e.g. MAPE. Some authors, without attempting to build a complete taxonomy, suggest grouping metrics by certain aspects, e.g. characteristic of error measured. Morley, Brito and Welling [17] grouped metrics by the nature of measured statistic: accuracy (e.g. MSE, RMSE, MdAE, etc.) and bias (e.g. ME, MPE, etc.). So in this research we measure performance with MSE for measured statistic accuracy and MAPE, because it is a dimensionless metrics that involve mathematical division of quantities. The smallest value of the error result for MSE and MAPE is the best result.

4. RESULT AND ANALYSIS

Prediction performance mainly deals with accuracy. Confusion matrix, showing the total prediction indicated by RMSE, MSE, and absolute errors, and accuracy of each class are measured. The results of this calculation show average error as can be seen in Table 4. In addition, this calculation assists to recognize whether the prediction is wrong or not.

Table 4. Prediction table.

	Manual	GLM
MAD	289,298 M	21,127,000,000,000,000 M
MSE	122,883,547,626,822,000 M	103,080 M
RMSE	350,547 M	321,060.74
MAPE	111.64	0.28
	Manual	**Decision Tree**
MAD	289,298 M	141,570,000,000,000,000 M
MSE	122,883,547,626,822,000 M	243,980 M
RMSE	350,547 M	493,943.32
MAPE	111.64	0.72
	Manual	**Random Forest**
MAD	289,298 M	260,730,000,000,000,000 M
MSE	122,883,547,626,822,000 M	337,700 M
RMSE	350,547 M	581,119.61
MAPE	111.64	0.93
	Manual	**GBT**
MAD	289,298 M	8,806,400,000,000 M
MSE	122,883,547,626,822,000 M	24.743.000.000,00
RMSE	350,547 M	157.299,08
MAPE	111.64	0,18

5. CONCLUSION AND SUGGESTION

Conclussion

Before we get through the analysis using Machine learning, we analyze using manual method and get result

MSE= 122,883,547,626,822,000,000,000.00 MAPE= 111.64. Each method uses different approach for prediction analysis, for the first method, Generalized Linear Model (GLM), serves to test the causal correlation [18] between Factor Cause Variable (X) and as a result variable. In this research, the X variable is Quarter, Service, Service Group and Quantity and the Y variable is Revenue Data. The result form graph of data is slightly close together between prediction and target but become away after data number 24, but the MSE and MAPE result are better than the result of manual analysis MSE = 103,080,000,000.00 and MAPE = 0.28. The second method is decision tree which is a structure looking like a flowchart with every internal node that signifies a "test" in an attribute. The Root Node for this case is divided by quarter and the decision node for this case is divided by service group, code, and quarter. The result form graphs, the data is only close together in several point MSE and MAPE results are better than the result of manual analysis but worse than GLM, MSE = 243,980,000,00.00 and MAPE = 0.72. The third method is random forest. In this thesis, we make bagging, there are 4 models used instead of just 1 model (each model has a child dataset from randomly selected brooders, usually 0,8:0,2). Bagging is usually done because there are assumptions about the distribution of datasets. In this case, we will use 3 leaf inputs. The result form graph, the data only close together in several point worse than decision tree methods between prediction and target MSE and MAPE results are worse than the result of manual analysis, GLM and Decision Tree, MSE = 337,700,000,00.00 and MAPE = 0,.3. In this thesis, we implement Gradient boosted Tree in implementation on decision tree algorithm. The result form graphs, the data close together from beginning till end of data target MSE and MAPE results are the best result than other methods, MSE = 24,743,000,000.00 and MAPE = 0.18

Suggestion

The last design gave Mean Squared Error = 0.18 for GBT Prediction. It was suggested to develop GBT algorithm to improve the MSE performance. The improvement can be using different set of rules (different kind of GT example without the use of GBT script but using random forest script) or combine boosted tree using ADAboost script.

REFERENCES

[1] Sastry, S. H., Babu, P., & Prasada, M. S. **"Analysis & Prediction of Sales Data in SAP-ERP System using Clustering Algorithms".** arXiv preprint arXiv:1312.2678, 2013.

[2] Shrivastava, V., & Arya, N. **"A study of various clustering algorithms on retail sales data".** *Int. J. Comput. Commun. Netw*, 1(2), 2012.

[3] Rajagopal, D. **"Customer data clustering using data mining technique"**. arXiv preprint arXiv:1112.2663, 2011.

[4] Mann, A. K., & Kaur, N. **"Review paper on clustering techniques".** *Global Journal of Computer Science and Technology*, 2013.

[5] Shah, N., Solanki, M., Tambe, A., & Dhangar, D. **"Sales Prediction Using Effective Mining Techniques",** 2015.

[6] D'Arcy, B., Gallagher, C., and Madden, M. G., **"A Bayesian Classification Approach to Improving Performance for a Real-World Sales Forecasting Application,"** *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, 10.1109/ICMLA.2015.150, 2015.

[7] Aleksandrovich, P.V., Leopoldovich, K.I., and Viktorovich, P.A., **"Predicting Sales Prices of the Houses Using Regression Methods of Machine Learning",** *3rd Russian-Pacific Conference on Computer Technology and Applications (RPC),* 10.1109/RPC.2018.8482191, 2018.

[8] Ching-She, W., Patil, P., and Gunaseelan, S., **"Comparison of Different Machine Learning Algorithms for Multiple Regression on Black Friday Sales Data,"** *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS),* 10.1109/ICSESS.2018.8663760, 2018.

[9] Cheriyan, S., Ibrahim, S., Mohanan, J., and Treesa, S., **"Intelligent Sales Prediction Using Machine Learning Techniques",** *2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE),* 10.1109/iCCECOME.2018.8659115, 2018.

[10] Ullah, I,. Raza, B., Malik, A.K, Imran, M., Islam, S.U., Kim, S.W., **"A Churn Prediction Model using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector,"** *IEEE Access*, 2019.

[11] Rey, T. D., Wells, C., & Kauhl, J. **"Using data mining in forecasting problems".** *In SAS Global Forum 2013: Data Mining and Text Analytics*, 2013.

[12] Huang, W., Zhang, Q., Xu, W., Fu, H., Wang, M., & Liang, X. **"A Novel Trigger Model for Sales Prediction with Data Mining Techniques".** *Data Science Journal*, 14, 2015.

[13] Alpaydin, E. **"Introduction to Machine Learning (Adaptive Computation and Machine Learning)",** *The MIT Press*, 2004.

[14] Lazăr, C., & Lazăr, M. **"Using the Method of Decision Trees in the Forecasting Activity. Petroleum-Gas University of Ploiesti Bulletin",** *Technical Series*, Vol. 67(1), 2015.

[15] Flesch, B., Vatrapu, R., Mukkamala, R. R., & Hussain, A. **"Social set visualizer: A set theoretical approach to big social data analytics of real-world events",** *In Big Data (Big Data), 2015 IEEE Internationsal Conference on*, pp. 2418-2427, 2015.

[16] Asooja, K., Bordea, G., Vulcu, G., & Buitelaar, P. **"Forecasting Emerging Trends from Scientific Literature"**. *In LREC*, 2016.

[17] Botchkarev, A. **"Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology",** arXiv preprint arXiv: 1809.03006, 2018.

[18] PREDA, C. and SAPORTA, G. **"PLS approach for clusterwise linear regression on functional data".** *In Classification, Clustering, and Data Mining Applications* (D. Banks, L. House, F. R. McMorris, P. Arabie and W, 2004.