http://ejournal.seaninstitute.or.id/index.php/Ekonomi
Jurnal Ekonomi, Volume 11, No 02 September 2022
ISSN: 2301-6280 (print) ISSN: 2721-9879 (online)

CUSTOMER'S RESPONSES TOWARDS IN-VEHICLE COUPON RECOMMENDATION AN IMPLEMENTATION OF BUSINESS ANALYTICS CONCEPT

Genesis Sembiring Depari[1]*, Efin Shu[2], Cut Alya Fachriza[3], Joceline Chow[4], Jerica Wijaya[5], Ryanto Winata[6]

Fakultas Ekonomi dan Bisnis, Universitas Pelita Harapan Medan

ARTICLEINFO

Keywords:
Customer's response,
In-vehicle coupon recommendation,
Business analytics

E-mail:
genesis.sembiring@uph.edu

ABSTRACT

Marketers are constantly searching for innovative tactics to increase sales performance. One of the most popular methods is through offering coupons to potential customers. However, selecting the most potential customers is not an easy task. Customer selection and segmentation become urgently important for business. To tackle these problems, an application of business analytics method is introduced. Besides, 3 machine learning algorithms such as random forest, naive bayes, and decision tree were utilized in predicting the likelihood of coupon to be accepted by users. Eventually, Random forest was found as the most accurate algorithm with the highest prediction accuracy.

1. INTRODUCTION

In the world of business, marketers are constantly searching for innovative tactics. The purpose of the innovation is to improve their strategies with the aim for customers to be more engaged and appealed to their products. One of the tactics that they used to increase customers' engagement is by offering coupons. Coupon recommendation systems have been popularly known amongst marketers in business to boost customers' habits so they could purchase more products.

Companies' services become not only more appealing when customers are given a challenge with a corresponding reward, yet it also can result in the customers having repeated purchases. Those purchases then will enhance the brand's impact on its customers. However, determining which coupon can fulfill the customers' demand would be rather difficult as each customer has their own preferences, behavior, situation, and context. Thus, offering them the wrong coupons might bring loss to ourselves as a business is losing customers. To overcome such issues, marketers could collect the data that would help in predicting their future customers and offer a useful coupon that is worth it to the customer.

To prevent losing the customer, machine learning can be utilized by the marketer. By collecting the data and having results that are extracted from machine learning, marketers could make a wise decision that could bring a good impact on their company. However, the data collected might be a large amount of data that is taking too much time to be analyzed manually by human power. Not only did it take much time, but it also needed lots of workforces which affected inflation costs. With the presence of machine learning, marketers will be given more time to decide on their wise decision and prevent a miscalculation (Schapire, 2008).

In the last few years, utilizing machine learning methods to analyze customer behavior has become more appealing. Many of the models that are used are clustering models, regression models, and utilized algorithms such as random forest, deep learning, support vector machine, and others. In 2017, Wang et al proposed a Bayesian framework for the sets of learning rules to resolve a classification problem and application to predict customer intention in in-vehicle recommendation systems (Wang et al., 2017). In

2020, Dou X utilized ensemble learning that aims to predict customers' behavior in online purchasing (Dou, 2020).

In this paper, the in-vehicle coupon recommendation data set was taken as an example and examined through business analytics for the purpose to analyze customers' responses toward in-vehicle coupon recommendations. We are utilizing one of the machine learning software platforms Rapidminer, in which the business analytics concept would be performed. A comparative analysis by using several algorithms of machine learning would also be used to conclude the final analysis. The output results would encompass the prediction of whether customers will accept the coupon recommended by the marketers in various scenarios. Having predictions about customers' responses to the recommended coupons would allow marketers to furtherly understand their customers better.

2. METHOD

The in-vehicle coupon recommendation dataset was retrieved from the UCI Open source dataset. It encompasses 24 attributes with 12.684 data recorded through an Amazon Mechanical Turk poll. The survey described different scenarios that include destination, education, passengers, and more, before suggesting the voucher to the customer.

Table 1 Attributes and Explanation

No	Attributes	Description	Role (Dependent / Independent)
1.	Destination	No Urgent Place, Home, Work	Independent
2.	Passenger	Alone, Friend(s), Kid(s), Partner (who are the passengers in the car)	Independent
3.	Weather	Sunny, Rainy, Snowy	Independent
4.	Temperature	55, 80, 30	Independent
5.	Time	2PM, 10AM, 6PM, 7AM, 10PM	Independent
6.	Coupon	Restaurant(<$20), Coffee House, Carryout & Take away, Bar, Restaurant($20-$50)	Independent
7.	Expiration	1d, 2h (the coupon expires in 1 day or in 2 hours)	Independent
8.	Gender	Female, Male	Independent
9.	Age	21, 46, 26, 31, 41, 50plus, 36, below21	Independent
10.	Marital Status	Unmarried partner, Single, Married partner, Divorced, Widowed	Independent
11.	Has_Children	1, 0	Independent
12.	Education	Some college - no degree, Bachelors degree, Associates degree, High School Graduate, Graduate degree (Masters or Doctorate), Some High School	Independent
13.	Occupation	Unemployed, Architecture & Engineering, Student, Education & Training & Library,	Independent

		Healthcare Support, Healthcare Practitioners & Technical, Sales & Related, Management, Arts Design Entertainment Sports & Media, Computer & Mathematical, Life Physical Social Science, Personal Care & Service, Community & Social Services, Office & Administrative Support, Construction & Extraction, Legal, Retired, Installation Maintenance & Repair, Transportation & Material Moving, Business & Financial, Protective Service, Food Preparation & Serving Related, Production Occupations, Building & Grounds Cleaning & Maintenance, Farming Fishing & Forestry	
14.	Income	$37500 - $49999, $62500 - $74999, $12500 - $24999, $75000 - $87499, $50000 - $62499, $25000 - $37499, $100000 or More, $87500 - $99999, Less than $12500	Independent
15.	Bar	never, less1, 1~3, gt8, nan4~8 (implied in how many times the respondent goes to a bar every month)	Independent
16.	Coffee House	never, less1, 4~8, 1~3, gt8, nan (implied in how many times respondent go to a coffee house every month)	Independent
17.	CarryAway	n4~8, 1~3, gt8, less1, never (implied in how many respondents get take-away food every month)	Independent
18.	RestaurantLessThan20	4~8, 1~3, less1, gt8, never (implied in how many times do respondents go-to restaurant with an average expense per person of less than $20 every month?)	Independent
19.	Restaurant20To50	1~3, less1, never, gt8, 4~8, nan (implied in how many times do respondents go to a restaurant with an average expense per person of $20 - $50 every month?)	Independent
20.	toCoupon_GEQ15min	0, 1 (implied the driving distance to the restaurant/bar for using the coupon is greater than 15 minutes)	Independent
21.	toCoupon_GEQ25min	0, 1 (implied the driving distance to the restaurant/bar for using the coupon is greater than 25 minutes)	Independent
22.	direction_same	0, 1 (implied whether the restaurant/bar is in the same direction as the respondent's current destination)	Independent
23.	direction_opp	1, 0 (implied whether the restaurant/bar is in the same direction as the respondent's current destination)	Independent
24.	Y	1, 0 (implied whether the coupon is accepted or declined)	Dependent

Source: UCI Open source dataset

In this study, the RapidMiner 9.10.001 was utilized for analyzing the data. There were 3 machine learning algorithms which are random forest, naive bayes, and decision tree that were being examined and compared in terms of time processing and accuracy percentages. In data preparation, we had replaced several missing values of the attributes, converted the numerical attribute to the polynomial type, and selected the 'Y' attribute as the targeted variable. Optimize Parameter (Grid) operator was applied for each algorithm. The optimized parameters of random forest were the number of trees and criterion. The optimized parameter of naive bayes was Laplace correction. While the optimized parameter of the decision tree was criterion and confidence. Cross Validation was also applied, in which the data was split into training and testing. And in reducing the dimensionality of the data, Weight by Information Gain Ratio was applied.

Business Analytics

Business analytics is the summarization of all mechanisms that help to retrieve, process, convert, and analyze data into applicable insight that leads to preferable and rapid decision-making in the business. Not only decision making, but business analytics could also discover problem-solving by using highly developed mathematical, statistical, machine learning, and network science methods, along with a diversification of data and expert comprehension. Business analytics combines tools and techniques from four major domains such as information management, descriptive analytics, predictive analytics, and prescriptive analytics. Information management involves classifying, extracting, and reconstructing data and information into data warehouses, for business analysts to understand what has happened in the organization by using descriptive analytics tools. These days, since we are living in the big data era, analysts should know how to read, summarize, and perform the data so they can make better decisions. Different from descriptive, predictive analytics tools using the organization's past performance to forecast and approximate the future behavior that leads to high-efficiency results. Now, prescriptive analytics is using data to identify the best actions and optimal decisions to solve a problem in the organization. (Arben Asllani, 2014)

Descriptive Analytics

Descriptive analytics is data analysis method that can be used to explain what is happening in business. Descriptive analysis can be done through central tendency (mean, median, mode) and variability analysis (variance, standard deviation). The results of descriptive statistics can be used for comparing one data to the others in helping the users detect the sample characteristic that might influence their conclusion about the data (Thompson, 2009).

Predictive Analytics

Predictive analytics is an area of statistics that is argued with extracting information from data and utilizing it for the prediction of trends and behavior patterns. Frequently the unknown event of interest is in the future, yet predictive analytics can be utilized for various types of unknown either in the past, present, or future. For example, analyzing suspects of a crime or credit card fraud (Steven Finlay, 2014).

Predictive analytics takes confirmed relationships between explanatory and criterion variables from previous occurrences to predict future events (Hair, 2007). The ability of predictive analytics in learning based on previous occurrences is what defines its predictive technology. The purpose of predictive analytics is to generate relevant information, actionable information, better results, and wise decisions and predict future occurrences through analyzing the variety, veracity, speed, and value of large amounts of data (Rajni & Malaya, 2015).

Prescriptive Analytics

Prescriptive analytics is data analysis that is used to explain and determine what is the next action plan and why should they do it. It is known to be the most sophisticated type of business analytics which can outcome the best intelligence and value to business cases. The purpose of this analytics is to bring the best option from the prediction analytics result. To conduct prescriptive analytics, require the predictive analytics to be included and utilizes artificial intelligence, the algorithm is optimized and experts system in

http://ejournal.seaninstitute.or.id/index.php/Ekonomi
Jurnal Ekonomi, Volume 11, No 02 September 2022
ISSN: 2301-6280 (print) ISSN: 2721-9879 (online)

a probabilistic context in aiming to provide adaptive, constrained, automated, time-dependent, and optimal decisions. (Lepenioti et al., 2020).

The Algorithm

Random Forest

Random forest algorithm is another one of the best classifiers algorithms because it was able to classify a huge amount of data and provide accuracy. This algorithm assembles a forest of decision trees in a certain number. These trees are created from bootstrap samples of a training dataset. Each node produced by the decision tree represents a test on a specific attribute, each branch represents the end result of the test, and each leaf node represents a class label for an outcome for regression. (Wu et al., 2017).

Naive Bayes

Naive bayes is an algorithm in Bayesian statistics, which is a simple probabilistic classifier with strong assumptions that all of the attributes are conditionally independent, given by the value of the class variable (Bagus et al., n.d.) In the other words, naive bayes is also defined as a classifier with a probability method, which is able to predict the opportunities in the future based on the experiences that have occurred in the past. This algorithm is widely being applied in real-world practice because of its easiness and simplicity, with no complex frequentative parameter estimation. Depending on the probability model, naive bayes classifiers can be trained very efficiently in a supervised learning setting; it often outperforms more sophisticated classification methods. The other advantage of this algorithm is when it estimates the parameters that are necessary for classification, it only requires a small amount of training data. Since the independent attribute assumptions, there is no matrix multiplication or to determine the entire covariance matrix, but only the variance of the variables for each class. Therefore, when it comes to predicting large quantities and relatively high levels of accuracy, it is more efficient to choose the naive bayes method. (P.Bhargavi et al, 2009)

Decision Tree

Decision trees are the most used technique in data mining, as it is instinctive and efficient for producing the classifiers and regression from data. Decision trees, which are also known as classification rules, use a logic method so it is easy to set up and easy to interpret from the business user's point of view. There are a mass number of decision tree installation algorithms that are being described fundamentally in the machine-learning and applied-statistics literature. Those algorithms direct the execution of learning methods that assemble decision trees from a collection of input-output samples. A decision tree technique is to look for a solution. It contains a collection of nodes, where the attributes are being tested in each node and each represents a splitting rule. This split data was done by nodes based on the homogeneousness of data and set into subsets. (Kantardzic, n.d.).

ISSN: 2301-6280 (print) ISSN: 2721-9879 (online)

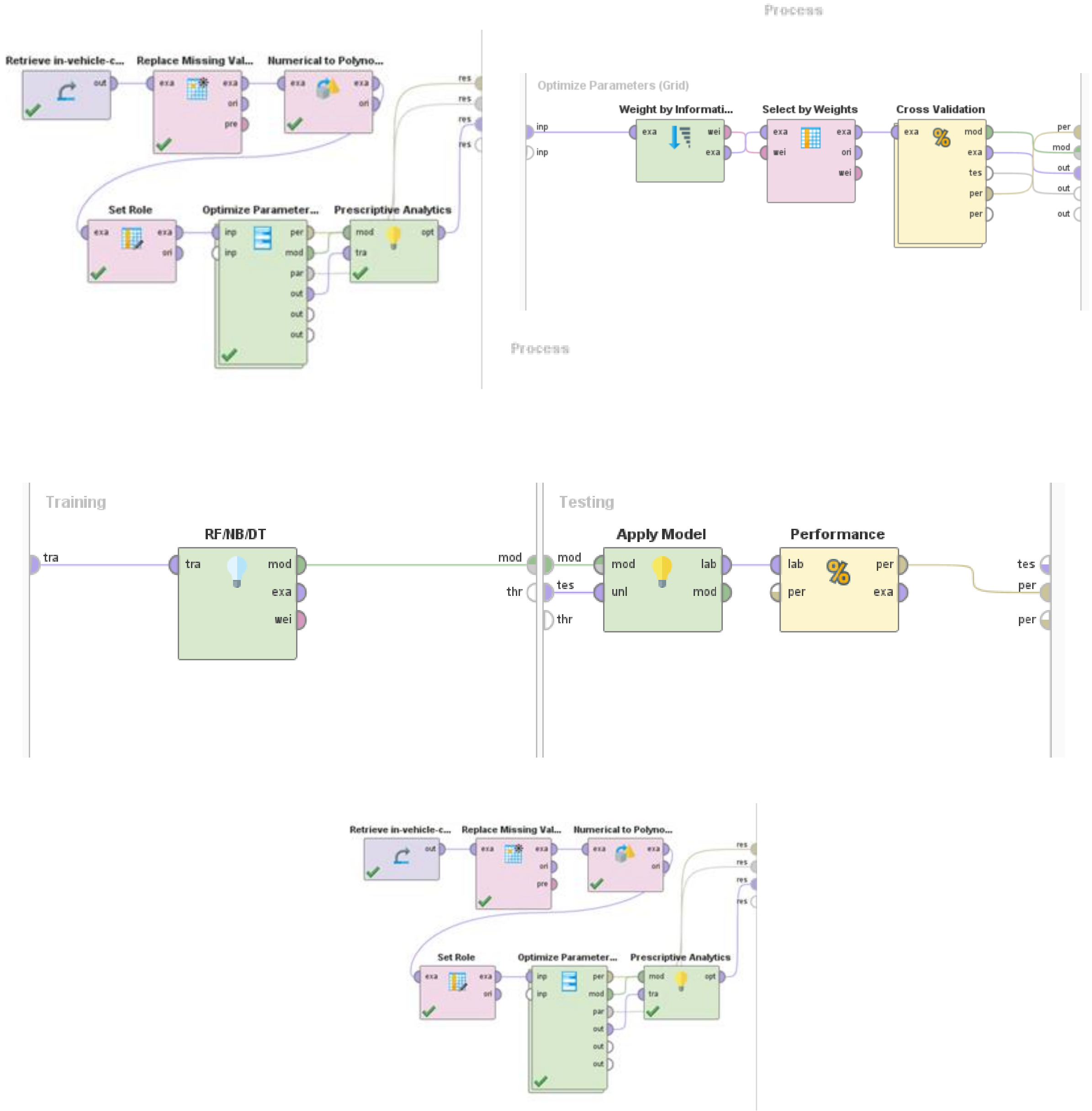

Figure 1. Model Sequence using RapidMiner

3. RESULT AND DISCUSSION

3.1 Descriptive Analytics

Descriptive analytics provided information regarding the customer's profile that has been offered the in-vehicle coupon. In Table 2, the data contain mostly married females who like to travel alone on a sunny day around 6 PM. Most of them have attended college, yet didn't graduate and left without getting a degree. Their educational background might be impacting them to be unemployed. However, for those who have an occupation, it states that most of them earn an income of around $25000 - $37499. The data also mentioned that the destination is mostly the No Urgent Place such as Coffee House which provides a coupon that expires in one day. Additionally, most of the customers that visit the Coffee House are at least once a month and have their order takeaway at least 1 to 3 times every month. If they are to visit the Restaurant then they mostly will spend less than $20 for 1 to 3 times and $20-$50 at least once a month.

The number of toCoupon_GEQ15min has at least 0 while the highest is 1. It results in having the average being larger than the standard deviation (0.561>0.496). This means that the data is most likely to be homogenous or less varied. The outcome of having less variety of data means it is more consistent and the next event that would likely occur is predictable.

In the number of those who have the toCoupon_GEQ25min attribute, the least number is 0 and the highest is 1. The average provided from it is 0.119 which is smaller than the standard deviation which is 0.324 means that the data is low on consistency and hard to predict. Comparing the standard deviation of toCoupon_GEQ15min and toCoupon_GEQ25min, we can conclude that the toCoupon_GEQ15min is variated and hard to predict due to low consistency due to having a larger standard deviation.

3.2 Predictive Analytics

Predictive analytics provided the prediction of customers' responses towards the in-vehicle coupon. In interpreting the predictive analytics result, we center our attention on overall accuracy percentage, class precision, and execution time. Along with the pred. 1 - true 1 result which implied the number of customers predicted and proven to accept the coupon, and the pred. 0 - true 1 result which implied the number of customers predicted to decline yet proven to accept the coupon.

Table 4. Result Utilizing Random Forest Algorithm

Accuracy: 77.65% +/- 1.86% (micro average: 77.65%)			
	true 1	true 0	class precision
pred. 1	6126	1751	77.77%
pred. 0	1084	3723	77.45%
class recall	84.97%	68.01%	

Above shown is the result of predictive analytics with two parameters: number of trees and criterion being optimized. By using the optimal parameters, the model achieved a 77.65% overall accuracy. In addition, 6126 customers were predicted to accept and proven to accept the coupon with 77.77% class precision, and 1084 customers were predicted to decline yet proven to accept the coupon with 77.45% class precision.

Table 5. Result Utilizing Naive Bayes Algorithm

Accuracy: 65.88% +/- 2.07% (micro average: 65.88%)			
	true 1	true 0	class precision
pred. 1	5567	2685	67.46%
pred. 0	1643	2789	62.93%
class recall	77.21%	50.95%	

In using the Naive Bayes algorithm, laplace correction parameter was optimized. The model achieved 65.88% overall accuracy. Besides that, 5567 customers were predicted to accept and proven to accept the coupon with 67.46% class precision, and 1643 customers were predicted to decline yet proven to accept the coupon with 77.21% class precision.

Table 5. Result Utilizing Decision Tree Algorithm

Accuracy: 71.86% +/- 1.32% (micro average: 71.86%)			
	true 1	true 0	class precision
pred. 1	6151	2510	71.02%
pred. 0	1059	2964	73.68%
class recall	85.31%	54.15%	

http://ejournal.seaninstitute.or.id/index.php/Ekonomi
Jurnal Ekonomi, Volume 11, No 02 September 2022
ISSN: 2301-6280 (print) ISSN: 2721-9879 (online)

There were also two parameters that were optimized by using the decision tree algorithm. By optimizing criterion and maximal depth parameters, 71.86% overall accuracy was achieved. Additionally, 6151 customers were predicted to accept and proven to accept the coupon with 71.02%, and 1059 customers were predicted to decline yet proven to accept the coupon with 73.68%% class precision.

Table. 6 Three Algorithm Comparison (Predictive Analytics)

Algorithms	Accuracy	Execution Time		Value	Class Precision
RF	77.65%	00 : 35 : 04	pred.1 - true 1	6126	77.77%
			pred. 0 - true 1	1084	77.45%
NB	65.88%	00 : 00 : 01	pred.1 - true 1	5567	67.46%
			pred. 0 - true 1	1643	62.93%
DT	70.80%	00 : 01 : 13	pred.1 - true 1	6151	71.02%
			pred. 0 - true 1	1059	73.68%

In comparing the three algorithms, random forest achieved the highest overall accuracy and class precision yet require time to process. Meanwhile, the Naive Bayes had the lowest time in processing yet achieved the lowest overall accuracy percentage and class precision. Therefore, we proposed for marketers to utilize the Random Forest algorithm in its model to predict the better result in predicting consumer behavior towards in-vehicle coupon recommendations.

3.3 Prescriptive Analytics

Table. 7 Three Algorithm Comparison (Prescriptive Analytics)

	RF	NB	DT
prediction(Y)	1	1	1
confidence(1)	0.523	0.525	0.750
confidence(0)	0.477	0.475	0.250
income	$37500 - $49999	$37500 - $49999	$37500 - $49999
toCoupon_GEQ25min	0	0	0
direction_opp	1	1	1
education	High School Graduate	High School Graduate	High School Graduate
occupation	Transportation & Material Moving	Transportation & Material Moving	Transportation & Material Moving
gender	Female	Female	Female
destination	Work	Work	Work
RestaurantLessThan20	never	never	never
direction_same	1	1	1
toCoupon_GEQ15min	0	0	0
weather	Sunny	Sunny	Sunny
temperature	72	72	72
CarryAway	4~8	4~8	4~8
CoffeeHouse	gt8	gt8	gt8
passenger	Alone	Alone	Alone
coupon	Coffee House	Coffee House	Coffee House
has_children	0	0	0
Restaurant20to50	never	never	never

http://ejournal.seaninstitute.or.id/index.php/Ekonomi
Jurnal Ekonomi, Volume 11, No 02 September 2022
ISSN: 2301-6280 (print) ISSN: 2721-9879 (online)

Bar	1~3	1~3	1~3
expiration	2h	2h	2h
time	6PM	6PM	6PM
age	39	39	39
maritalStatus	Married partner	Married partner	Married partner

Above shown is the table comparison of the prescriptive analytics result of the three algorithms, in which it provided the information of customers that accepted the in-vehicle coupon. From the three algorithms, all of the attributes were in the same category. Yet, the decision tree achieved the highest confidence level of all of the three algorithms examined.

The associated attributes encompass income of $37,500 to $49,999; high school graduate in education background; driving distance to a restaurant/bar is greater than 25 minutes; having the bar/restaurant in the same direction as the destination; working in transportation and material moving field; is a female; a destination to work; never went to a restaurant with expense less than $20 every month; the restaurant/bar is in a different direction with the current destination; driving distance to restaurant/bar is greater than 15 minutes; the weather is sunny; the temperature is in 72 degrees Fahrenheit; has taken take-away food 4 to 8 times every month; went to a coffee house more than 8 times; had no kid(s) as a passenger; accepted a coffee house coupon; has no children; never went to a restaurant with expense less than $20-$50 every month; went to a bar 1 to 3 times every month; accepted the 2-hour coupon; coupon was in given at 6 PM; is 39 years old; married. Therefore, marketers could take the outcome of the prescriptive analytics and utilized the associated attributes to develop a better marketing strategy.

4. CONLUSION

In-vehicle coupon recommendation is a system that provides information regarding the coupon available by learning the customer's behavior to help customers provide the suitable coupon. When sellers offer the customers coupons, the customers would give their response to it. Customer response is frequently analyzed by marketers to understand customer behavior towards their products and develop better marketing strategies. In developing better strategies and decisions, marketers could leverage the business analytics concept through the application of machine learning algorithm.

This paper leverages the concept of business analytics by using RapidMiner 9.10.001 in predicting customers' responses to in-vehicle coupon recommendations. There were 3 machine learning algorithms which are random forest, naive bayes, and decision tree that were being examined and compared in terms of the time processing and accuracy percentages. Each of the algorithms also has its parameter optimized to increase accuracy, as well as information gain ratio to reduce dimensionality.

From descriptive analytics, it is a size up that 5 out of 8 integer attributes are clustered to the respective mean (average > standard deviation). Predictive analytics results showed that random forest achieved the highest accuracy with 77.65% overall accuracy percentage, yet required the most time to process. Meanwhile, the prescriptive analytics showed attributes that are associated with customers who accepted the in-vehicle coupon. From the result, it is shown that the three algorithms have the same associated attributes in the same category. However, the decision tree algorithm acquired the highest confidence level of 0.750.

All of the results that have been discussed, bring us to a conclusion about the best algorithm that is suitable to be applied in analyzing in-vehicle coupon recommendation responses of customers. The most suitable algorithm for predictive analytics is random forest as the prediction percentage is the highest of the rest, which is the class precision of prediction 1 (yes) and true 1 (yes) with the result of 77.77% with an overall accuracy of 77.65%. While the most suitable for prescriptive analytics is the decision tree due it has the highest confidence level. Therefore marketers can use business analytics concepts with random forest and decision tree algorithm ability to generate relevant information, actionable information as well as better and wiser decisions in business.

REFERENCES

[1]. Arben Asllani. (2014). *Business Analytics with Management Science Models and Methods*. Pearson.

[2]. Bagus, I., Peling, A., Arnawan, N., Putu, I., Arthawan, A., & Janardana, I. (n.d.). Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm. In International Journal of Engineering and Emerging Technology (Vol. 2, Issue 1).

[3]. Dou, X. (2020). Online Purchase Behavior Prediction and Analysis Using Ensemble Learning. *2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, 532–536. https://doi.org/10.1109/ICCCBDA49378.2020.9095554

[4]. Gentile, C., Spiller, N., & Noci, G. (2007). How to Sustain the Customer Experience:: An Overview of Experience Components that Co-create Value With the Customer. *European Management Journal*, *25*(5), 395–410. https://doi.org/10.1016/J.EMJ.2007.08.005

[5]. Hair, J. F. (2007). Knowledge creation in marketing: The role of predictive analytics. *European Business Review*, *19*(4), 303–315. https://doi.org/10.1108/09555340710760134

[6]. Kantardzic, Mehmed. (n.d.). *Data mining : concepts, models, methods, and algorithms*.

[7]. Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. *International Journal of Information Management*, *50*, 57–70. https://doi.org/10.1016/J.IJINFOMGT.2019.04.003

[8]. Leskovec, J., Rajaraman, A., & Ullman, J. D. (n.d.). *Mining of Massive Datasets*.

[9]. Madyatmadja, E. D., Jordan, S. I., & Andry, J. F. (2021). Big data analysis using rapidminer studio to predict suicide rate in several countries. *ICIC Express Letters, Part B: Applications*, *12*(8), 757–764. https://doi.org/10.24507/icicelb.12.08.757

[10]. Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2*(1). https://doi.org/10.1186/s40537-014-0007-7

[11]. Nilsson, N. J. (1998). *INTRODUCTION TO MACHINE LEARNING AN EARLY DRAFT OF A PROPOSED TEXTBOOK*.

[12]. Padmavathi Mahila Viswa Vidyalayam, S. (2009). Applying Naive Bayes Data Mining Technique for Classification of Agricultural Land Soils. In IJCSNS International Journal of Computer Science and Network Security (Vol. 9, Issue 8).

[13]. Rajni, J., & Malaya, D. B. (2015). Predictive Analytics in a Higher Education Context. *IT Professional*, *17*(4), 24–33. https://doi.org/10.1109/MITP.2015.68

[14]. Steven Finlay. (2014). *Predictive Analytics, Data Mining and Big Data: Myths, Misconceptions and Methods (Business in the Digital Economy)*.

[15]. UCI Machine Learning Repository: in-vehicle coupon recommendation Data Set. (n.d.). UCI Machine Learning Repository. Retrieved April 7, 2022, from https://archive.ics.uci.edu/ml/datasets/in-vehicle+coupon+recommendation

[16]. Thompson, C. B. (2009). Descriptive Data Analysis. *Air Medical Journal*, *28*(2), 56–59. https://doi.org/10.1016/j.amj.2008.12.001

[17]. Wang, T., Rudin, C., Liu, Y., Klampfl, E., & Macneille, P. (2017). A Bayesian Framework for Learning Rule Sets for Interpretable Classification. In *Journal of Machine Learning Research* (Vol. 18). http://jmlr.org/papers/v18/16-003.html.

[18]. Wu, D., Jennings, C., Terpenny, J., Gao, R. X., & Kumara, S. (2017). A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests. *Journal of Manufacturing Science and Engineering*, *139*(7). https://doi.org/10.1115/1.4036350