การวิเคราะห์ข้อมูลลูกค้าผู้ซื้อกล้วยไม้กระถางช่วงไวรัสโคโรนาระบาดในประเทศไทย

Customer Analytics of Orchid Pot Business during the First Corona Virus Outbreak Period in Thailand

กมลวรรณ กระถินทอง[1] และ โษฑศ์รัตต ธรรมบุษดี[1*]

Kamolwan Krathinthong[1] and Sotarat Thammaboosadee[1*]

[1]กลุ่มสาขาวิชาเทคโนโลยีการจัดการระบบสารสนเทศ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล จังหวัดนครปฐม 73170

[1] Information Technology Management, Faculty of Engineering, Mahidol University, Nakhonpathom, 73170, Thailand

[*]Corresponding Author, E-mail: Sotarat.tha@mahidol.ac.th

Received: 7 November 2021 | Revised: 25 February 2022 | Accepted: 28 February 2022

บทคัดย่อ

โคโรนาไวรัส 2019 ที่รู้จักกันในชื่อโควิด-19 เริ่มแพร่ระบาดในปี 2019 ผู้ติดเชื้อรายแรกถูกพบขึ้นที่มณฑลอู่ฮั่น ประเทศจีนและได้แพร่ระบาดไปทั่วโลกรวมถึงประเทศไทย งานวิจัยฉบับนี้ใช้วิธีการเหมืองข้อมูล (Data mining) โดยได้นำกฎความสัมพันธ์ (Association rule) มาใช้เพื่อทำความเข้าใจพฤติกรรมผู้ซื้อสินค้ากล้วยไม้กระถางในช่วงไวรัสโคโรนาระบาดช่วงแรกในประเทศไทย โดยได้นำข้อมูลของลูกค้าที่ซื้อกล้วยไม้กระถางมาวิเคราะห์ รวมทั้งยังได้นำหลักการ FP-Growth มาปรับใช้เพื่อทำความเข้าใจข้อมูลลูกค้าที่มักจะซื้อสินค้าร่วมกัน วิธีกฎความสัมพันธ์ให้ผลลัพธ์ที่ดีที่สุดมา 7 กฎของลูกค้าที่ซื้อสินค้ากล้วยไม้กระถางแต่ละชนิดพร้อมกัน โดยแต่ละกฎแสดงค่าความเชื่อมั่น (confidence) ค่าลิฟท์ (lift) และ ค่าคอนวิกชัน Conviction ในช่วง 0.833 - 0.857, 2.629 - 5.602 และ 4.098 - 5.929 ตามลำดับ จากนั้นได้นำการสร้างแบบจำลองเชิงทำนาย (predictive modeling) มาประยุกต์ใช้เพื่อหาความสัมพันธ์และทำนายตัวแปรอิสระที่มีผลต่อตัวแปรต้นโดยใช้อัลกอริทึม Generalized Linear Model, Deep Learning, Random Forest และ Gradient Boosted Three และ ทดลองเปรียบเทียบกับการใช้ Automatic feature selection และ feature extractions โดยทั้งหมดให้ค่าความคลาดเคลื่อน (Relative error) ต่ำสุดอยู่ที่ 15.2% ผลที่คาดว่าจะได้รับจากงานวิจัยนี้เพื่อให้ผู้ประกอบการกล้วยไม้สามารถนำ Association Rule และ Predictive Modeling ไปปรับใช้ได้หากเกิดเหตุการณ์ร้ายแรงอย่างเช่น โควิด-19 ขึ้นอีก

ABSTRACT

CORONA VIRUS 2019, also known as COVID-19, began to spread in December 2019. The first affected covid-19 person was found in Wuhan district, China, then COVID-19 impacted worldwide, including Thailand. This research uses the Data Mining technique by applying the Association Rule to understand customer behavior in people who buy orchid pots during the first coronavirus pandemic in Thailand. This study applies to the database of customer purchase transactions who buy orchid pots. This research adopts the FP-Growth model to understand the groups of products customers typically buy. Finally, the Association

rule generates seven rules of orchid pot types that customers purchase in the same basket. Each rule shows Confidence, Lift, and Conviction range from 0.833 – 0.857, 2.629 – 5.602, 4.098 – 5.929, respectively. This study also deployed Predictive modeling by utilizing the Generalized linear model, Deep learning, Random Forest, and Gradient Boosted Tree. As a result of the Predictive model, Gradient Boosted Tree without Auto feature selection and feature extractions methods produce the lowest relative error at 15.2%. The association rule finds an orchid pot that customers purchased one of the items in the group. The expected result of this study is that orchid entrepreneurs can adopt this outcome from Association Rule and Predictive Modeling Analytics when a problematic situation similar to the COVID-19 pandemic happens again.

คำสำคัญ: กฎความสัมพันธ์ แบบจำลองเชิงทำนาย กล้วยไม้กระถาง โควิด-19 พฤติกรรมผู้บริโภค
Keywords: Association Rule, Predictive Modeling, Orchid pots, COVID-19, Customer behavior

INTRODUCTION

Orchids from Thailand are very well-known and economic crops of the world. It can create revenue and value-added more than 2,500 million baht per year from orchids global market approximate 7,720 million baht. Thailand is the number one to exporter orchids in flower to other countries worldwide (De et al., 2019). When, The Coronavirus outbreak which also called “COVID-19” around the world includes Thailand. Industrial crops orchid is directly affected by the COVID-19 pandemic (Bhatti et al., 2020). So, the business could not run continuously. Orchid entrepreneurs try to change from the traditional way of orchid distribution by selling only intermediaries directly purchase from the farm and then distribute orchids to well-known plant and flower markets around Thailand or export to abroad to sell new channels online platform directly. This research takes advantage from data mining to find the useful customer data transaction for exporling customer relationship and orchid business can adopt this model when the severe situation like COVID-19 happen again. During COVID-19, customer behavior has rapidly changed from buying at a brick-and-mortar store to online shopping online shopping in Thailand registered a 40% growth in usage after 2020 lockdowns. Moreover, most people must stay at home due to government policies like lockdown. So, online shopping sharply increases (Wongtana-sarasin et al., 2021).

Pillai and Jolhe (2020) mentioned that Market Basket Analysis Case Study of a Supermarket: This study aims to give supermarket entrepreneur to promote goods and increase sales strategically. These results also give precious insights for cross-selling, up-selling, and new product combinations tasks. By applying association rules, the 576 customer purchased transactions took place in a well-known supermarket in Mumbai for February 2019. They considered the huge number of products, demographics, season, and past experiences. This study shows that the three parameters, Support, Confidence, and Lift, play an essential role in creating many strategies in a supermarket. They are helpful in the discovery of insight into the relationship between various products.

Karim and Sirwan (2019) introduced a Data Mining Approach for Prediction Modelling Using

Association Rule. This research aims to give the supermarket authorities find and analyze what similar products should be kept on one single shelf that matches the common purchasing tendency. This research performs prediction modeling with customer purchased transaction data from supermarket applications. The outcome of this study helps supermarkets increase sales and goodwill of the supermarket through saving the customer's time and effort during purchasing.

Ghassani et al. (2021) performed one technique of the association method is Market Basket Analysis, in their research "Market Basket Analysis Using The FP-Growth Algorithm to Determine Cross-Selling. They used various kinds of basic needs such as basic foodstuffs transaction data in November 2019 by applying with FP-Growth algorithm for calculating the sets that often appear from data. The study researcher determines a minimum support value of 3% and minimum confidence of 50%. The association process using these values produces three strong rules, namely if ades 350 ml, then fried / lontong with a support value of 0.030 and confidence 0.556 and if fried st, then fried / lontong with a support value of 0.048 and confidence 0.639, and if nasi uduk / bacang, then fried / rice cake with a support value of 0.031 and confidence 0.824. The results of the association rules can be applied using one of the marketing techniques, namely cross-selling, to increase the company's profit.

Ebrahimi et al. (2019) proposed the lastest analysis of risk criteria of mastitis relied on milk using Data Mining techniques like prediction model in mastitis occurrence in their research. They use Data from 364,249 milk example which were collected by an automatic platform which milk volume, lactose concentration, electrical conductivity, protein concentration, peak flow and milking time to measure the model via Deep Learning, Naïve Bayes, Generalized Linear Model, Decision Tree, Gradient-Boosted Tree and Random Forest. Finally, Gradient-Boosted Tree gave the best accurate performance at 84.9% in predicting sub-clinical bovine mastitis. All of those dataset illustrate how the algorithyms can be utilize to forecast sub-clinical mastitis in multiple bovine herds regardless of the size and sample techniques. This research is advantage for other farms to identificate infected cows ad minimizing economic lose.

Wu and Lin (2018) mentioned in the divergent of e-commerce that they utilized the Association rule model to explore the relationships in e-commerce logistics. The Support show how frequently the ecommerce logistics items occur at the same time. The confidence explain the probability of e-commerce logistics item A and B happen given the occurrence of e-commerce logistics item A. The lift describes ecommerce logistics item A and B happening more or less frequently than expectation. As the useful of two data mining techniques. The researcher will use Association Rule and Predictive Modeling with this study to explore pattern insight of customer who buy orchids during covid-19 outbreaks via an online platform.

Therefore, this research starts by collecting customer data classified with gender, location, amount, quantity, species, and color of customer purchasing. Before customers decide to buy orchids, they are usually selected by Species, Color, and Roots. Thus, in this paper researcher analyzes data by using the "Data Mining technique" (Rygielski et al., 2002). The processes are applying Association

rule by using the FP-Growth model to find customer behavior insight according to the customer purchased transactions. Association Rule was measured by Support, Confidence, Lift, and Conviction. (Wicaksono et al., 2020) and developing Predictive Models to test the relationship between the target or label features and other dependent variables by utilizing different prediction systems, Generalized linear model (GLM), Deep learning (DL), Random forest (RF), Gradient Boosted Tree (GBT) and Automatic Feature Selection algorithms (Wang et al., 2019a). They will be evaluated by Relative error and Weights by Correlation. Researchers hope this model will be more developed to create marketing strategies for orchid entrepreneurs to plan their species, color, and roots to sell to a customer to increase revenue.

METHODOLOGY

In this paper utilizes data mining techniques with database customer purchase transactions to understand customer behavior insight. The methodology of this research is illustrated in Fig. 1

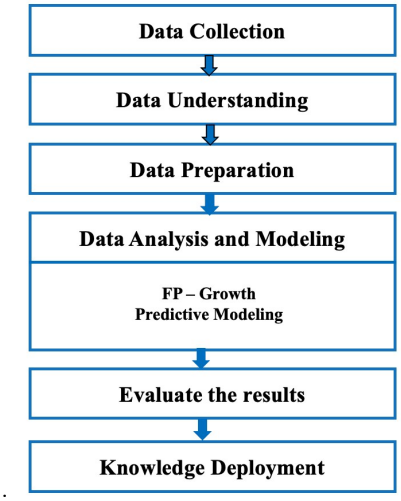

Figure 1. Overall Methodology

Data Collection: Researcher collect data from an orchid entrepreneur in Nakhonpathom provinces which data was analyzed into two periods by using orchid purchasing customer data 555 transactions between 1 April 2020 to 6 July 2020 divided data into two periods among the peak of COVID-19 outbreaks on 1 April 2020 to 14 June 2020 and after the government declined district policy phase 4 on 15 June 2020 to 6 July 2020 (Wongtanasarasin et al., 2021).

Data Understanding: To understand the customer data who buy orchid pot compries order date, platform, quantity, price, postal code, roots-type, color, and sex by studying from raw data recorded in the daily spreadsheet program and then researcher properly organizes the data to be ready for analysis, as illustrated in Table 1.

Data Preparation: In this process, Raw Data was input to correctly organize for ready analyze, as shown in Figure. 2.

Table 1 Data Dictionary

No.	Attribute	Definition	No.	Attribute	Definition
1.	Order Date	The date when the customer orders orchid	9.	Dendrobium species	Dendrobium species
2.	Platform	Channel to sell product	10.	Helix Dendrobium	Helix Dendrobium species are set as Semi-Aerial roots.
3.	Type	Type of sell product	11.	Semi-Aerial roots	Type of orchid roots
4.	Quantity	Orchid order in Quantity	12.	Mokara	Mokara orchid species are set as Aerial roots.
5.	Price	price per pot based on orchid species	13.	Asco	Asco orchid specie is an orchid that was mixed between Ascocentrum and Vanda specie.
6.	Amount	Total Amount	14.	Aerial roots	Show the total quantity of customer who buy Aerial roots orchids
7.	Provinces	Provinces	15.	Sex	Gender
8.	Postal Code	Postal code point customer orders orchid location			

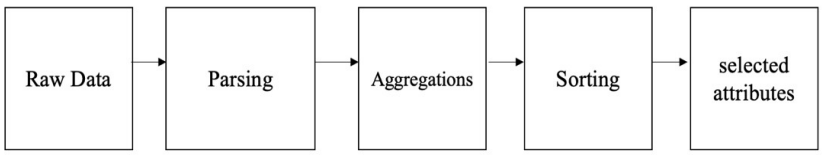

Figure 2. Data Preparation Flow chart

This process utilized customer purchased data transactions from one of the orchid entrepreneurs, which data was collected in a spreadsheet daily. The researcher cleans raw data using a machine learning program to get the outcome cleaning data to be ready to analyze as follow:

1. All Mokara, Dendrobium white mix green, and Helix pink were parsed converse from nominal attributes to numerical type.

2. Aggregation is used with attributes Amount, Quantity, Orchid species, Orchids color, and Orchid roots to find the total sum and group attributes Order date, Type, Platform and Sex into the same group.

3. Sorting Order date, Type, Platform attributes by order date.

4. The data set were transformed from nominal to numerical type and mapped into binary (0,1) for Association Rule model.

Data Analysis and Modeling: This step was utilized two techniques of Data Mining to analyze data customers as follow: Association Rule used FP- Growth algorithm to find the frequent customer's purchase patterns and adapted market basket analysis to understand the purchase behavior insight of the orchid pots buyer who tend to buy orchid together. This research used the parameter to evaluate the optimization association rule using the criteria support (Khurana and Sharma, 2013) as a frequency constraint. FP – Growth use Support, Confidence, Lift and Conviction to evaluate the model. Support is a criteria that indicate how frequently the items apper in our data set. Confidence (Kumar and Tiwari, n.d.) indentifies the frequently of if-then pattern were found as true statement. Lift (Hussein et al., 2015) is used to compare with Confidence criteria to find the Confidence maximization. Finally Conviction (Valdivia, 2014) is also the criteria that ensure the result of Confidence and Lift.

Predictive Modeling is used to predict and forecast likely future outcomes with the orchid pots buyer transactions. The example algorithms for this model are generalized linear model (GLM) (Vevea and Hedges, 1995) is the method that finds the relationship the dependent (attribute amount) and independent variables (attribute , Date, Orchid Species, Orchids root, Orchids colors and platform) deep learning (DL) (Arora et al., 2020) is a model to create a large neural network with multiple layers and is trained by the backpropagation method, random forest (RF) (Shaikhina et al., 2019) consists of many decision trees. The 'forest' generated by the random forest algorithm is trained through bagging or bootstrap aggregating. Bagging is an ensemble meta-algorithm that improves the accuracy of machine learning algorithms. The gradient boosted tree (GBT) (Ebrahimi et al., 2019) is a model that develops from the decision tree to improve efficiency. It is processed by randomly creating many decision trees.

Automatic Feature Engineering (Severyn and Moschitti, 2013) aims to help automatically create many examples features out of dataset from which the best can be selected and used for training. It supports multi-objective feature engineering and provide the optimum solution from the Pareto front concerning complexity vs. model error. This research setup the complexity to low level it focuses on model interpretation rather than the complex model without explainability. Moreover, the ten-fold cross-validation (Berrar, 2018) is used to train the dataset. Measuring by Relative error was used as a measure of precisionis. Relative error is the ratio of the absolute error of a measurement being taken. In other words, this type of error is relative to the size of the item being measured. Relative error is expressed as a percentage and has no units. (Wang et al., 2019b) After training by four algrorithm, researchers apply automatic feature selection and feature generation (Venkatesh and Anuradha, 2019) in this study. Automatic feature selection and generation is the process where automatically selected the features in the dataset that contribute most to our prediction variable or output in which data are interested. Moreover, this study will find important factors by weight by correlation (Žák and Souček, 2019) with each attribute compared with quantity attributes.

Knowledge Deployment: This step refers to the result of this study, which will be discussed in the next section.

RESULT AND DISCUSSION

In this section, the researcher reports that the Association Rule and Predictive modeling results according to the methodology described in a previous section. The Association Rule with the FP–Growth is shown in Table 2.

Table 2 Rules from Association Model

Premise	Conclusion	Support	Confidence	Lift	Conviction
Dendrobium boey-leng , Dendrobium purple mix red	Dendrobium red mix brown	0.033	0.857	5.602	5.929
Helix dendrobium purple, Dendrobium red mix brown	Dendrobium boey-leng	0.027	0.833	5.446	5.082
Helix dendrobium pink, Dendrobium old rose	Dendrobium boey-leng	0.027	0.833	5.446	5.082
Dendrobium old rose, Dendrobium boey-leng	Dendrobium red mix brown	0.027	0.833	5.446	5.082
Dendrobium old rose, Dendrobium red mix brown, Dendrobium white dot purple	Dendrobium boey-leng	0.027	0.833	5.446	5.082
Helix dendrobium pink, Dendrobium old rose	Helix dendrobium purple	0.027	0.833	2.629	4.098
Helix dendrobium pink, Dendrobium white dot purple	Helix dendrobium purple	0.027	0.833	2.629	4.098

As a result, FP – Growth generates seven rules that customers purchase each orchid type at the same time. Criteria for selecting these 7 rules include Support, Confidence, and Lift. The Support represents how frequently the customer purchase transactions occure together. The Confidence describes the conditional probability of both orchid types A and B orrcuring given the occurrence of orchid types A. The Lftt explains orchid type A and B occurring more or less frequently than expected. Conviction measures the implication power of the rule from statistical independence. The researcher summarizes that customers buy dendrobium boey-leng and dendrobium purple mix red, they will buy dendrobium red mix brown with Support, Confidence, Lift, and Conviction are maximization from all rules. And customers buy helix dendrobium purple and dendrobium red mix brown, they will buy dendrobium boey-leng, and customers buy helix dendrobium pink and dendrobium old rose, they will dendrobium boey-leng. Support, Confidence, Lift, and Conviction measurement are

shown in the table as ordered. Orchids entrepreneurs can adopt these rule to create their market plan to match products to sell together and campaign sale promotion.

The Predictive modeling were employed to process the relationship between amount of purchasing as a label feature and other dependent variables (Date, Orchid Species, Orchids root, Orchids colors, and platform) to find how much independent variables affect to amount label. This study tests the predictive model with GBT, RF, DL, and GLM algorithms. Dividing into two steps are before Automatic Feature Selection and after Automatic Feature Selection. The model is measured by Relative error and the GBT without Automatic Feature Selection model algorithm produces the lowest relative error on every algorithm at 15.2%. The other algorithms result is shown in Table 3, Figure 3, and Table 4.

Moreover, this work measures the performance of predictive modeling by relative error. The measurement by weights by correlation divided into two sections are before and after feature selection. The performance shows that there are three attributes that similar between before and after. The Attributes semi-arial roots, dendrobium mix color, and dendrobium white spot give the same weight at 0.669, 0.561, and 0.321, respectively

According to the two performance of the predictive model shown in the Figure above, the best algorithm for training the customer data sets is gradient boosted trees. It is accuracy at 84.8% and takes 30 seconds in processing the data.

Table 3 The Predictive modeling performance by relative error.

Model	Relative Error	
	Non-Automatic Feature Engineering	**Automatic Feature Engineering**
Gradient Boosted Trees	15.20%	16.50%
Random Forest	30.50%	27.10%
Deep Learning	27.80%	28.20%
Generalized Linear Model	32.10%	28.50%

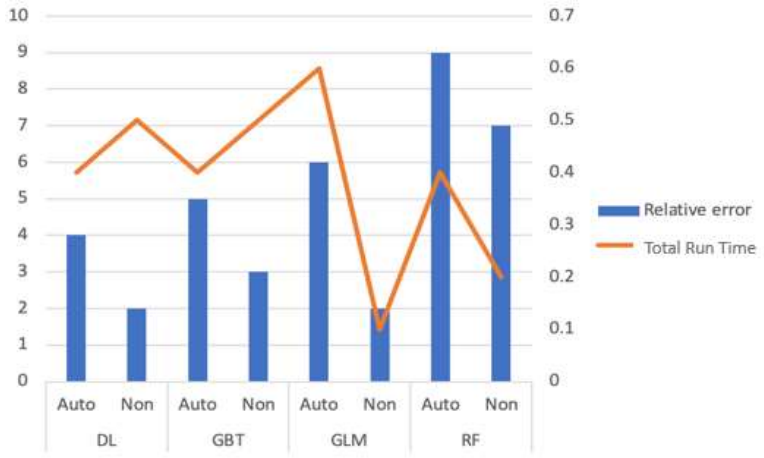

Figure 3. Predictive performance by Relative Error and Total Run Time

Table 4 Predictive modeling performance by weights by correlation.

Weights by Correlation					
Non-Automatic Feature Engineering			**Automatic Feature Engineering**		
No	**Attribute**	**Weight**	**No**	**Attribute**	**Weight**
1	Semi-arial roots	0.669	1	Semi-arial roots	0.669
2	Dendrobium Mix color	0.561	2	Dendrobium Mix color	0.561
3	Dendrobium white dot purple	0.321	3	Dendrobium white dot purple	0.321
4	Quarter*	0.317	4	isFacebook	0.283
5	isFacebook	0.283	5	Arial roots	0.263
6	Month*	0.275	6	Dendrobium red mix brown	0.227
7	Duration*	0.269	7	Helix dendrobium purple	0.226
8	Arial roots	0.263	8	Dendrobium boey-leng	0.203
9	Dendrobium red mix brown	0.227	9	Dendrobium white	0.196
10	Helix dendrobium purple	0.226	10	Dendrobium arirang blue	0.186

*Noted: Quarter, Month and Duration were extracted after using feature engineering.

Researcher utilizes weights by correlation to explore the main factor that affect to orchid amount perchasing. The result above show that Semi-arial roots is the esstential variable that affect to amount when customer buy product the most. Dendrobium Mix color and Dendrobium white dot purple are high affect when customer buy orchid pot to amount both Non and Automatic Feature Engineering.

After the Association Rule, Pillai and Jolhe (2020) mentioned that Market Basket Analysis Case Study of a Supermarket: The results of this study provide valuable insights for cross-selling, up-selling, and new product integration tasks. By applying association rules, the 576 customer purchased transactions took place in a popular supermarket in Mumbai for February 2019. They considered the vast number of products, demographics, season, and past experiences. This study shows that the three measures, Support, Confidence, and Lift, play an essential role in deciding various promotional strategies in a supermarket. They are helpful in the discovery of insight into the relationship between various products. The previous research can compare with this research. The 555 customer purchased transactions which were collected in April – July 2020 was used to analyze in Association Rule. The results give seven rules for cross-selling of orchid types. The Support, Confidence, Lift, and Conviction were measure for seven rules. The advantage of these seven rules is helpful for exploring customer behavior insight and their relationship when purchase each orchid types together.

After the result of Predictive Modeling. This study can compare with Mansour et al., (2019) proposed the lastest analysis of risk parameters of mastitis based on milking using Data Mining techniques like prediction model in mastitis occurrence in their research “Comprehensive

analysis of machine learning models for predictions of sub-clinical mastitis: Gradient -Boosted Trees Outperform other models" They use Data from 364, 249 milk example which were collected by an automatic platform which milk volume, lactose concentration, electrical conductivity , protein concentration, peak flow and milking time to measure the model by Random Forest, Deep Learning, Generalized Linear Model, Decision Tree, and Gradient-Boosted Tree. Finally, Gradient-Boosted Tree gave the best accurate performance at 84.9% in predicting sub-clinical bovine mastitis. All of those data illustrate how the algorithyms can be applied to forecast sub-clinical mastitis in multiple bovine herds regardless of the size and sample tecniques. This research is advantage for other farms to identificate infected cows ad minimizing economic lose. The Predictive Modeling of their study can compare with this research because of using Deep Learning, Gradient -Boosted Trees, Generalized Linear Model, and Random forest to analyze. The study of Mansour et al., (2019). gave the best accurate performance at 84.9% for Gradient-Boosted Tree and this research gave the best algorithm for training the customer data sets is Gradient-Boosted Tree too. It is accuracy at 84.8% and takes 30 seconds in processing the data.

CONCLUSIONS

This paper aims to examine customer behavior patterns insight into who buy orchid pots during the first coronavirus pandemic in Thailand. Because of Coronavirus pandemic around the world and also in Thailand. So, orchids pot entrepreneurs who distribute their products via offline platforms can not run the business normally. They changed their sell orchid pots from their traditional way to an online platform and started recording customer data transactions in a spreadsheet. Thus, the researcher is interested in using the customer purchase data transaction from orchid business to analyze customer behavior because orchids have many kinds of species, colors, and roots. Hence, the researcher decides to use the advantage from the Data Mining technique to apply with the data. The Association Rule was deployed with market basket analysis model to find customers buy dendrobium boey-leng and dendrobium purple mix while evaluated by the maximization of Support, Confidence, Lift, and Conviction. Moreover, the predictive modeling is used to forecast future outcomes with the orchid pots buyer transactions by applying that with Gradient Boosted Trees, Random Forest, Deep Learning, and Generalized Linear Model algorithms to evaluate. Attribute "amount" was used as a label feature to compare with other dependent variables (Date, Orchid Species, Orchids root, Orchids colors, and platform. Dividing into two steps are after Automatic Feature Selection. The model is measured by relative error, and the Gradient Boosted Trees without Automatic feature selection model algorithm produces the lowest relative error on every algorithm at 15.2%. This research aims to help orchid entrepreneurs use this model to adapt when a difficult situation like COVID-19 happens again. The amount of data transaction limited this study, and species of orchid is less. Researchers hope this model will be more developed and utilized more data mining techniques to apply with this study to figure more relationship in customer data transaction and will be helped orchid entrepreneurs to manage their orchid inventory. Finally orchid entrepreneurs

create marketing strategies to sell more orchids and increase revenue.

REFERENCES

Arora, P., Kumar, H. and Panigrahi, B.K. (2020). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. 139: 110017.

Berrar, D. (2018). Cross-validation. In Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics 1(3): 542–545.

Bhatti, A., Akram, H., Basit, H.M., Khan, A.U., Mahwish, S., Naqvi, R. and Bilal, M. (2020). E-commerce trends during COVID-19 Pandemic. International Journal of Future Generation Communication and Networking 13(2): 1449–1452.

De, L.C., Kumar, R., Khan, A. and Medhi, R.P. (2019). Social Science Orchid Farming-A Remunerative Approach for Farmers Livelihood. International Journal of Scientific Research 3(9): 77-80.

Ebrahimi, M., Mohammadi-Dehcheshmeh, M., Ebrahimie, E. and Petrovski, K.R. (2019). Comprehensive analysis of machine learning models for prediction of sub-clinical mastitis: Deep Learning and Gradient-Boosted Trees outperform other models. Computers in Biology and Medicine 114.

Ghassani, F.Z., Jamaludin, A. and Irawan, A.Y. (2021). Market basket analysis using the fp-growth algorithm to determine cross-selling: Jurnal Informatika Polinema 7(4): 49-54.

Hussein, N., Alashqur, A., and Sowan, B. (2015). Using the interestingness measure lift to generate association rules. Journal of Advanced Computer Science and Technology 4(1): 156.

Khurana, K., and Sharma, M.S. (2013). A Comparative Analysis of Association Rules Mining Algorithms. International Journal of Scientific and Research Publications 3(5).

Kumar Jain, J., and Tiwari, N. (n.d.). A Survey: On Association Rule Mining. Manoj Ramaiya / International Journal of Engineering Research and Applications (IJERA) Www.Ijera.Com, 3: 2065–2069.

Pillai, A.R. and Jolhe, D.A. (2020) Market Basket Analysis: Case Study of a Supermarket. Advances in Mechanical Engineering pp 727–734.

Rygielski, C., Wang, J.C. and Yen, D.C. (2002). Data mining techniques for customer relationship management. In Technology in Society. 24(4): 483-502.

Sarkhel H. Ther Karim and Rzgar Sirwan (2019) A Data Mining Approach for Prediction Modeling Using Association rule. international journal on informatics visualization 3(3): 266-268.

Severyn, A., and Moschitti, A. (2013). Automatic Feature Engineering for Answer Selection and Extraction. In: Proceedings of the 2013 conference on Empirical methods in natural language processing. 458–467.

Shaikhina, T., Lowe, D., Daga, S., Briggs, D., Higgins, R. and Khovanova, N. (2019). Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. 52: 456–462.

Valdivia Miranda, C. (2014). Sistemas informáticos y redes locales : Sistemas de

telecomunicaciones e informáticos, Electricidad y electrónica. Paraninfo.

Venkatesh, B. and Anuradha, J. (2019). A review of Feature Selection and its methods. Cybernetics and Information Technology 19(1): 3–26.

Vevea, J. L. and Hedges, L.V. (1995). A general linear model for estimating effect size in the presence of publication bias. Psychome trika 60(3): 419-435.

Wang, P., Tian, C., Liu, R. and Wang, J. (2019a). Mathematical model for multivariate nonlinear prediction of SMD of X-type swirl pressure nozzles. Process Safety and Environmental Protection 125: 228–237.

Wang, P., Tian, C., Liu, R., and Wang, J. (2019b). Mathematical model for multivariate nonlinear prediction of SMD of X-type swirl pressure nozzles. Process Safety and Environmental Protection 125: 228–237.

Wicaksono, D., Ihsan Jambak, M., and Saputra, D.M. (2020). The comparison of apriori algorithm with preprocessing and FP-growth algorithm for finding frequent data pattern in association rule. In: Proceedings of the Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019). 315-319.

Wongtanasarasin, W., Srisawang, T., Yothiya, W., and Phinyo, P. (2021). Impact of national lockdown towards emergency department visits and admission rates during the COVID-19 pandemic in Thailand: A hospital-based study. Emergency Medicine Australasia 33(2): 316–323.

Wu, P.J. and Lin, K.C. (2019). Unstructured big data analytics for retrieving e-commerce logistics knowledge: Telematics and Informatics 35(1): 237-244.

Žák, P., and Souček, M. (2019). Correlation of tumor necrosis factor alpha, interleukin 6 and interleukin 10 with blood pressure, risk of preeclampsia and low birth weight in gestational diabetes. Physiological Research 68(3): 395–408.

❑❑❑❑❑