INTRODUCTION

Background

Competition in a world of trade is increasing, especially in the coffee business. The increase in people's interest in modern coffee drinks causes many business ventures such as cafes, coffee shops, and more. One way to attract customers is not only with excellent product quality but also in terms of customer service to attract loyalty. Therefore, we need customer relationship management to increase customer loyalty and reduce customer churn for sustainable business. Customer relationship management (CRM) is developed to explore consumer information with the help of information technology (Chorianopoulos 2016). Customer Relationship Management is a strategic approach that aims to generate profit and long-term customer relationships through information technology. The company needed a customer classification to determine the level of customer loyalty. Customer data can be used to classify a loyal customer.

Coffee is a commodity that is developing lately. Based on statistics from the ministry of agriculture (Triyanti 2018), Indonesia's coffee consumption from 2000 to 2016 tends to increase. In 2000, Indonesia's coffee consumption only reached 1.68 million bags (packs), wherein one package there was 60 kg of coffee. However, in 2016, Indonesia's coffee consumption has reached 4.6 million bags or surged more than 174 percent. Due to the increased consumption of coffee processed into modern coffee drinks, coffee shops are increasingly rampant in Bogor. According to the chairman of Indonesia's specialty coffee association (SCAI), Syafrudin said that coffee shops' contribution to the uptake of domestic coffee production reached 30%. This growth is due to the increasing number of people who like to enjoy brewed coffee in coffee shops rather than drinking instant coffee.

In this research, coffee that will be examined is a modern coffee drink that is blended from roasted coffee into various types of drinks such as milk coffee, cappuccino, americano, pandan coffee, and other modern coffees. The modern coffee drink is indeed trending in Indonesia, especially in Bogor, because people's lifestyles nowadays tend to prefer drinking coffee blended by baristas while enjoying the coffee shop's facilities and atmosphere. There are currently around 150 coffee shops in the Bogor area. Increasing competition in the world of coffee business makes Kemenady Coffee Shop tend to lose customers over time.

The problem then arises when customers find another coffee shop and decided not to return to this coffee shop because of the lack of service to the purchaser, the quality of coffee that is less competitive, and less innovation in the marketing of the products. For this reason, customer churn prediction is needed to discover which customer will cut the relationship with a coffee shop in the future and determine customer retention, which is a strategy to increase sales by serving and satisfying customers who are expected to return customers. By identifying customer segments, the company can maintain the level of customer loyalty and improve customer satisfaction. Customer segmentation can be used to identify a suitable marketing strategy and applied to the target customers.

Today Customer Relationship Management (CRM) is explicitly driven by emerging technologies. Companies developing CRM systems and companies using

CRM are both equally dependent on technology. Intrusive technologies are emerging that have started penetrating the minds of the customers. The Usage of technologies like Artificial Intelligence (AI), Augmented Reality (AR), Virtual Reality (VR), Predictive Modeling, Deep Learning, Machine Learning, etc. have become essential for businesses to survive, mostly in the domains like Supply Chain Management (SCM), Production and CRM (Deb *et al.* 2018). The role has further extended to accessing the customer's needs and wants and offering products that customers may like to purchase with Artificial Intelligence (AI) applications. AI uses customer's purchase history, customer's time spend on various products, customer's social media presence and behavior, accessing the similarities of customers with another buyer, and recommending new products. In this research, we focused on predicting customer patterns to prevent churn in the future based on past transaction data from customers, thus helping companies recommend the best strategy to prevent the customer from churning.

Business analytics is defined as the process of understanding the data-driven activities of a business to draw inferences to make calculated decisions with higher certainty (Bag 2017). Business analytics encompasses a gamut of analysis around business data to draw information that could be used by the managers at various levels in an organization. Business analytics enables fact-based decision-making while extending accountability in decision making. Customer analytics refers to the processes that bestow businesses to engross the customer outlook necessary to deliver accurately anticipated services (Fry *et al.* 2019). It is the practice of visualizing customer transactions to find underlying patterns, behaviours, or anomalies. In this research, business analytics showed a new insight into a modern coffee shop. Transaction data obtained from coffee shops are processed to produce a rule in which the rule can provide important information which is very useful in preventing churn in the coffee shop. This process will significantly help coffee shops in the future to anticipate losing customers and decreasing profits.

The customer churn prediction can be defined as assigning a probability to leave the company next to each customer in the company dataset to make it viable to the firm to indicate which customer has the highest propensity to churn (Fry *et al*. 2019). For the company to set a threshold to indicate which probability to churn, they want to tackle. The customer's clusters who crossed this threshold will be targeted by a specially designed marketing plan to increase their retention. Customer churn refers to when customers stop the relationship with the company. A business usually tells customers as churning when the specified amount of time has passed since the last interaction with the company. The full cost of customer churn includes loss of revenue and marketing costs involved with replacing with a new customer. Reducing customer churn is one of the main objectives of each business. Customer churn prediction using machine learning models follow a set of steps. The data is collected, next, the selected data was pre-processed and transformed into a suitable form for building a machine learning model. After modelling the testing was performed and then finally the model was deployed (Kim, Shin & Park, 2005). The machine learning investigated the data and detects the underlying data patterns for the customer churn analysis (Kim, Shin & Park, 2005). Using machine learning the prediction of customer churn was more accurate than the traditional approach.

Predictive Analytics is the process of finding interesting patterns and meaningful data (Abbott 2014). Predictive Analytics is a data-driven algorithm and obtains the key characteristics of the data model itself. Algorithms on predictive analytics automate the process of finding patterns in the data. The domain that will focus on predictive analytics that will build CRM is a business strategy that focuses on customers designed to optimize revenue, profitability, and customer loyalty. Analytics can be defined as a process that involves the use of statistical techniques (measures of central tendency, graphs, and so on), information system software (data mining, sorting routines), and operations research methodologies (linear programming) to explore, visualize, discover and communicate patterns or trends in data (Duarte *et al.* 2018). Analytics plays an important role in this research to find new insights into customer relationship management. In this study, the data mining method is used to process customer data to find new patterns in customer churn predictions, which are expected to provide insight into coffee shop owners in developing a marketing strategy that fits each customer category. So far, research on customer churn has been overwhelming, especially in services such as banking and insurance. However, it is still difficult to find customer churn predictions in the coffee shop sector. Therefore, this study is dedicated to coffee businesses confused about knowing customers' patterns and behavior.

PT. Industry Kemenady Mandiri is a company engaged in the sale of coffee in the form of a coffee shop as well as a supplier of roasted arabica and robusta beans. PT. Industry Kemenady Mandiri has its own farmers' groups cultivated by Kemenady to produce a coffee bean that is processed into roasted coffee in beans and powder form. The coffee powder is then reprocessed into modern coffee drinks sold in Kemenady coffee shop in various kinds by the ongoing trends such as milk coffee or the rakjat's coffee, espresso, latte cappuccino and more. While the products of arabica and robusta roasted coffee sold in packaged form in Kemenady coffee shop. Additionally, roasted coffee products also marketed to the hotel, restaurant, and coffee shop in Indonesia. Kemenady Coffee Shop is a modern coffee shop that targets teenagers, young people, and the elderly. Based on the Kemenady coffee shop's operational manager's statement, the majority of visitors are aged 14-25 years, and there are more female visitors than men. The time most visited by customers is the range from 15:00 WIB to 21:00 WIB. Usually, at 15:00 filled by school children who stop by to enjoy coffee after school. While from 18:00 - 21:00, the majority is filled with customers who want to enjoy coffee after work. Friday to Sunday is the busiest day at Kemenady coffee shop.

The customer churn topic is discussed many times in the intelecommunication sector. Many kinds of research were done to know the optimization algorithms to use to expect when the customers are going to churn before they do, knowing that about 30 percent of the customers are going to churn per year. To acquire a new customer, the company is going to spend ten times more than to retain the already existing customers, so by now to retain the existing customers is more profitable than acquiring new customers. With the recent techniques in data mining, companies succeed in having insights about their customer behavior either by using the usual Sociodemographic variables or using more advanced variables like their customer call center data.

Research in the CRM field has been done a lot. Approaches that are often used can be qualitative or quantitative. Most of the research to determine customer satisfaction and customer loyalty uses qualitative methods such as SWOT analysis and customer satisfaction analysis based on surveys and questionnaires. However, the latest research is increasingly sophisticated due to better technological developments. Research in the CRM field can be carried out in a more contemporary way using the latest methods and approaches such as AI, Machine Learning, Data Mining, and Big Data.

A comparative study on customer churn prediction was performed by Vafeiadis, et.al. (2016) on telecom data set. The performance comparison of multi-layer perceptron, Decision Tree, SVM, Naïve Bayes and Logistic regression were compared. All the models were built and evaluated using cross-validation. Monte Carlo simulations were used and SVM has outperformed other models with an accuracy of 97% and F-measure of 84%. Kiran Dahiya et al. (2015) proposed a new framework for the churn prediction model, implemented it using WEKA data mining software. Each customer was classified as a potential churner or non-churner. The framework discussed was based on the Knowledge Discovery Data process. Three different datasets, small, medium, and abundant with different attribute tests, were considered. The efficiency and performance of the decision tree and logistic regression techniques have been compared. The accuracy achieved with the decision tree was much higher than logistic regression. From this paper proved that decision tree is a lot more accurate to predict the churn.

Qureshi et al.(2013) aims to present commonly used data mining techniques for churn prediction. The dataset used was obtained from the Customer DNA website and contains traffic data of 1,006,000 customers and their usage behavior for three months. The class imbalance problem was solved by re-sampling. Regression analysis, Artificial Neural Networks, K-Means Clustering, Decision Trees including CHAID, Exhaustive CHAID, CART, and QUEST were taken into consideration to identify churn. The results were compared based on the values of precision, recall, and F-measure. Decision trees, especially Exhaustive CHAID, were found to be the most accurate algorithm in identifying potential churners.

Khan et al. (2015) presented a unified analytic framework for detecting the early warnings of churn and assigning a “Churn Score” to each customer that indicates the likelihood of a particular customer to churn within a predefined amount of time. The approach uses a brute force approach to feature engineering that generates a large number of overlapping features from customer transaction logs, then uses two related techniques to identify the features and metrics that are most predictive of customer churn. These features are then fed into a series of supervised learning algorithms that can accurately predict subscriber churn. For a dataset of roughly 1,000,000 subscribers from a South Asian mobile operator observed for six months, an approximate of 90 percent accuracy was achieved.

The research started by analyzing business processes in sales of coffee beverages. Then we distributed questionnaires to customers to determine the customer satisfaction rate with the service, quality, price, and choices that will be used to analyze the level of customer loyalty. Then proceed with building the customer relationship management model using RFM Analysis, Customer Lifetime Value based on Data Mining and clustering. RFM analysis is a method often used to determine the segmentation of customers based on past behavior. RFM analysis helps the company significantly, not only in identifying and targeting valuable customers who have a very high chance but also in avoiding the cost of expensive communication and marketing to a customer who has a high probability of churning. Following on from RFM, another key concept of CRM is the customer life cycle. Data mining provided customer insight, which is vital for these objectives and establishing an effective CRM strategy (Han *et al.* 2012). The results of RFM analysis and lifetime value will be used to segment the customer and predict customer churn using the predictive analytics method, which is a decision tree.

The challenge is to build a churn prediction model in a coffee shop based on past transactions based on the explanation above. The expected goal is to design strategies to prevent the customer from churning in the future. This research obtained the data from multiple sources. We applied data mining technique to bring out insight into this research and motivation for improving the modern coffee shop in terms of customer services.

Problem Statement

The problem in this coffee shop is when customers find another coffee shop and decided not to return to this coffee shop because of the lack of service to the purchaser, the quality of coffee that is less competitive, and less innovation in the marketing of the products. Based on a kemenady coffee shop owner statement, coffee sales decreased over time due to high competition in the coffee business. For this reason, customer churn prediction is needed to determine customer retention, which is a strategy to increase sales by serving and satisfying customers who are expected to return customers. This is achieved by analyzing the customer segmentation and satisfaction with customer analytics, which is the RFM method and CLV based on data mining. We propose the rules to predict the customers who are about to churn in the future based on the surveys, interviews and past transaction data from customers. We will use predictive analytics, a decision tree to analyze the data to understand the prediction rules. From this prediction, hopefully, we can suggest a fit marketing strategy to retain the customers.

Objectives

This study has the following objectives:

1. To model a new business process using BPMN.
2. To predict the customer churn using data mining tools.
3. To recommend a fit strategy to prevent churn and maintain loyalty.
4. To evaluate the result from customer churn prediction

Benefits

The benefit of this research is to provide a solution to maintain customer loyalty and to predict the potential churners at Kemenady Coffee Shop based on customer satisfaction and behaviour from a past transaction. It helps to decrease the churn rate and maintain loyalty by applying the strategy recommended based on the prediction results.

Boundaries

The boundaries of this study are as follows:

1. Agroindustry coffee in question is a modern coffee drink.
2. The coffee shop in question is a modern coffee shop targeting teenagers, young adults, and adults.
3. The customer data used for this research are from registered customers in the database.
4. Customer satisfaction referred to the result from surveys.

LITERATURE REVIEW

Business Process and Requirement

A business process is an existing process or mechanism in running a business (Weske 2019). The business process refers to how the process occurs or the sequence of activities in implementing something. The existence of innovation causes this business process always to be redesigned to make it faster, more efficient, more effective, and cheaper. The result of this innovation is the requirements of the stakeholders involved in the business.

To redesign the business process, it uses the system development life cycle (SDLC). It shows if SDLC consists of planning, analysis, design, implementation, and maintenance in a cycle. SDLC is a methodology for systems development that highlights systems analysis and systems design (Valnacich and Goerge 2017). The life cycle is a circular process in which the output leads to develop a new version of an existing system. The cycle is iterative until an acceptable system is found.

The crucial parts of SDLC are the analysis and design phases. The analysis studies about system requirements. This phase studies current procedures in doing organization tasks. The analysis has two subphases: requirements determination and requirement structuring. The requirement determination phase determines what the users want from the new system, and the requirement structuring organizes the information in the previous step into a meaningful representation in UML (Unified Model Language) diagram and system entity diagram conducted by Wasson 2016 (Figure 1). Figure 1 shows the 11 attributes in the system entity.

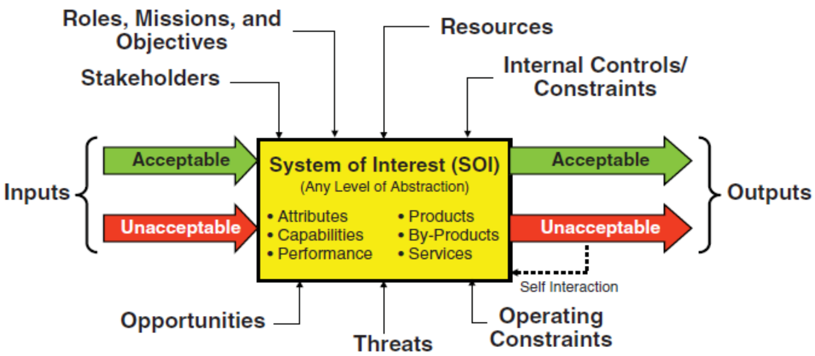

Figure 1 System entity diagram (Wasson 2016)

This research also applied the BPMN (Business Process Model & Notation) diagram as the UML diagram. The BPMN diagram shows the detail activity each stakeholder does in the system. The output of the analysis phase is a description of the alternative solution recommendation. The design phase converts the result into logical and then physical system specifications. This step also formulates the solution until it runs properly and works as intended.

Business Process Modeling Notation (BPMN)

BPMN is a graphical notation to represent the flow of a business process. BPMN provides many graphic notations used for process modeling. Graphical notations owned by BPMN include start events, tasks, intermediate messages, end events, gateways, and other notations, each of which has its function. In BPMN, a business process requires a sequence in business activities and supporting information. Therefore, BPMN is a workflow that is easier to understand for analyzing and modeling a business or business process (Aagesen and Krogstie 2015).

Design

Design is a series of procedures to translate a system's analysis into a programming language for describing how the system components are implemented. While understanding the system development is creating new activities and replace or repair the existing system in whole or in part (Maxim *et al*. 2015). It can be concluded that design and development is a depiction, planning, and sketching or arrangement of several separate elements into a unified whole and function. Thus, the notion of design improves the analysis results in the form of a software package and then creates a system or to improve the existing system.

Data Mining for Customer Relationship Management

Customers are the most important asset of a company and organization. Because of that a company should plan and employ a clear strategy for customer handling. Customer Relationship Management (CRM) is the strategy for building, managing, and strengthening loyal and long-lasting customer relationship. CRM should be a customer-centric approach based on customer insight. Its scope should be the customers' personalized handling as distinct entities by identifying and understanding their differentiated needs, preferences, and behavior. CRM aims at two main objectives (Chorianopoulos 2016):

1. Customer retention through customer satisfaction
2. Customer development

Data mining can provide customer insight vital for these objectives and for establishing an effective CRM strategy. It can lead to personalized interactions with customers, increasing satisfaction and profitable customer relationships through data analysis. It can offer individualized and optimized customer management throughout all phases of the customer life cycle, from acquisition and establishment of a strong relationship to attrition prevention and win back of lost customers. Marketers strive to get a greater market share and a more significant share of their customers. Marketers are responsible for getting, developing, and keeping customers. Data mining can help them in all these tasks, as shown in Figure 2 (Chorianopoulos 2016).

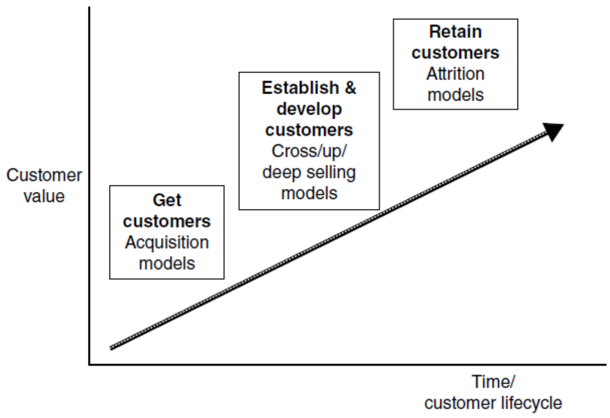

Figure 2 Data mining and customer life cycle (Chorianopoulos 2016)

Customer Satisfaction

According to (Nettleton 2014), a key aspect that affects the business must be mentioned: customer satisfaction. This can be a very influential factor in a customer's loyalty and can be a source of new customers via favorable references. To obtain information and data about customers' degree of satisfaction, the easiest way is to ask them—directly, if they are accessible, or via feedback questionnaires sent by post or from the business's website. IT systems play a crucial role in attending to customer problems and queries (as well as the interpersonal skills of employees) by giving informational support for a rapid solution. It is important to avoid delays when attending to customers, be it due to inefficiency in the administrative circuits or due to a lack of resources (Putler and Krider 2012).

Again, it is evident that a business can accumulate statistics about products, services, sales channels, and regions in terms of the number of complaints, returned goods, and so on. To capture feedback data about the quality of service or about products, the business can design a questionnaire and send it to customers, with some incentive for them to return it duly completed (Design *et al.* 2008). In this way, the business can identify which of its regional offices are excellent and need attention. Of course, the company also needs to have previously defined what is meant by an acceptable level of service or product quality. For example, the best level for returned goods is, of course, zero percent. Feedback about clients' defects is a valuable source of information for the production and research and development departments. It can be used to make corrections to the product or service.

Customer Development

Customer development is the process of growing the value of retaining a customer. Companies generally attempt to cross-sell and up-sell products into the customer base while still having regard for the customer's satisfaction. Cross-selling is selling additional products and services to an existing customer. Up-selling is selling higher-priced or higher-margin products and services to an existing customer (Kotler *et al.* 2015). Customers generally do not respond positively to persistent and repeated efforts to sell additional products and services that are not related to their requirements. Indeed, there is an argument that companies should down-sell where appropriate. This means identifying and providing lower-cost solutions to the customers' problems, even if making a smaller margin. Customers may regard up-selling as opportunistic and exploitative, thereby reducing the level of trust they have in the supplier and putting the relationship at risk (Mohammadi *et al.* 2014). However, multi-product ownership creates a structural bond that decreases the risk of relationship dissolution.

Business Analytics

Business analytics is defined as understanding the data-driven activities of a business to draw inferences to make calculated decisions with higher certainty (Bag 2017). Business analytics encompasses a gamut of analysis around business data to draw information that could be used by the managers at various levels in an organization. Business analytics enables fact-based decision-making while extending accountability in decision making. Business analytics is *defined as* exploring, experimenting, stimulating, and summarizing data to extract information (Shmueli *et al.* 2018). Business analytics encompasses the entire key informational and decisional attributes of any business, and it is vitally important that business analytics features in the overall strategic vision of all businesses (Bag 2017). The significant goals of business analytics include:

- Providing real-time, actionable information aimed at superior business decision making.
- Providing tools at all levels of an organization to help decision making around customer goals and profits while comparing performance.
- Providing analysis that helps the business forecast the future with greater objectivity and accuracy.
- Providing the insight and understanding to support informed decisions and confident actions and providing the feedback that is needed to create a learning organization.

Analytics can be defined as a process that involves the use of statistical techniques (measures of central tendency, graphs, and so on), information system software (data mining, sorting routines), and operations research methodologies (linear programming) to explore, visualize, discover and communicate patterns or trends in data. Simply, analytics convert data into useful information. Analytics is an older term commonly applied to all disciplines, not just business. A typical example of the use of analytics is the weather measurements collected and converted into statistics, which in turn predict weather patterns (Duarte *et al.* 2018). There are many types of analytics, and there is a need to organize these types to understand their uses. According to the Institute of Operations Research and Management Sciences (INFORMS) organization suggests for grouping analytics types (Table 1). These types of analytics can be viewed independently. For example, some firms may only use descriptive analytics to provide information on the decisions they face. Others may use a combination of analytic types to glean insightful information needed to plan and make decisions (Sedkaoui 2018).

Whereas the process of analytics can involve any one of the three types of analytics, the major components of business analytics include all three used in combination to generate new, unique, and valuable information that can aid business organization decision-making. In addition, the three types of analytics are applied sequentially (descriptive, then predictive, then prescriptive). Therefore, business analytics (BA) can be defined as a process beginning with business-related data collection and a sequential application of descriptive, predictive, and prescriptive major analytic components, the outcome of which supports and demonstrates business decision-making and organizational performance (Sedkaoui 2018).

Table 1 Types of business analytics (Fry *et al.* 2019)

Type of Analytics	Definition
Descriptive	The application of simple statistical techniques that describes what is contained in a data set or database.
Predictive	An application of advanced statistical, information software, or operations research methods to identify predictive variables and build predictive models to identify trends and relationships not readily observed in a descriptive analytics.
Prescriptive	An application of decision science, management science, and operations research methodologies (applied mathematical techniques) to make best use of allocable resources.

Customer Analytics

Customer analytics refers to the processes that bestow businesses with the engrossing customer outlook necessary to deliver accurately anticipated services. It is the practice of visualizing customer transactions to find underlying patterns, behaviors, or anomalies. For all business-to-customer activities, customer analytics embodies methods and practices aimed at customer processes to make gainful investments. Customer analytics is more focused than customer intelligence, which is generic in nature. Understanding a customer implies deciphering and decoding an individual within the immediate surroundings he or she prospers. Customer analytics's origin goes back to individual psychology, aspiration, and motivation, surrounded by the sociological transformation in the form of a household, marriage, family, social interaction, and peer dynamics (Bag 2017).

RFM Models

RFM Model is a method often used to determine the segmentation of customers based on past behavior (Zalaghi and Abbasnejad 2014). RFM model attributes can be used to find valuable customer segments to find trends in spending and to predict future customer behavior. *Recency* is the distance between the last time a transaction with the current time. If the range, the smaller the value of R increases. Frequency is how often customers make purchases in a certain period, such as three times within one year. If the transaction amount higher, then the value of F is also more significant. Monetary means the amount of money spent on the transaction customers when a certain period of money, the greater the value of M will be higher. The larger the value of R and F, the more likely customers will do the transaction back to the company. Also, the higher the amount of M, the tendency of customers in respond to the company's products and services (Pankaj 2019).

RFM techniques significantly help organizations identify and target valuable customers who have a very high chance and avoid the cost of expensive communication and marketing to customers who have a lower probability of purchase. Instead, it helps identify customers with a high chance and to target marketing strategy and discussion at the appropriate company. Limitations are the RFM technique can be applied only to the available historical customer data and not on the data leads (Santoso and Erdaka 2015).

Customer Lifetime Value

Customers are the basis of a firm's existence. A firm creates and provides value (product or service) to customers and provides value (revenue/profits) to firms (Glady *et al.* 2018). One way to determine customer value is to look at metrics such as average customer revenue or average customer profit (Konečnik *et al.* 2014). Unless used for a specific segment, these metrics put all customers at equal level, which is incorrect and indeed not of interest to marketers who celebrate customer heterogeneity. Some customers do provide more profits than others and some customers lead to losses. Another interesting way to analyse the value provided by a customer is to look at their purchase behavior through RFM analysis (recency, frequency, monetary value) (Bai 2018).

RFM analysis is a common approach for understanding and monitoring consuming behaviors. It is quite popular, especially in the retail industry. As its name implies, it involves the calculation of three core KPIs, recency, frequency, and monetary value which summarize the corresponding dimensions of the customer relationship with the enterprise. The recency measurement indicates the time since the last purchase transaction. Frequency denotes the number and the rate of the purchase transactions. Monetary value measures the purchase amount. These indicators are typically calculated at a customer (cardholder) level through simple data processing of the recorded transactional data (Chorianopoulos 2016).

- **Recency**: time (in units of time, typically in days or in months) since the most recent purchase transaction or shopping visit.
- **Frequency**: total number of purchase transactions or shopping visits in the period examined. An alternative, and more "normalized" approach that also takes into account the tenure of the customer, calculate frequency as the average number of transactions per unit of time, for instance, the monthly average number of transactions.
- **Monetary value**: the total or the average per time unit (e.g., monthly average value) amount of purchases within the examined period. According to an alternative yet not so popular definition, the monetary value indicator is defined as the average transaction value (average amount per purchase transaction). Since the total value tends to be correlated with the frequency of the transactions, the reasoning of this alternative definition is to capture a different and supplementing aspect of the purchase behavior.

Customers can be rated on each parameter and classified from most valuable (highest recency, frequency, and monetary value) to least valuable (lowest recency, frequency, and monetary value) (Olson 2017). This is a simple model that can help segment customers and predict their future behaviour. These methods were good enough when we did not have access to great data and advanced statistical software to analyse data. Now we can collect and analyse data at the lowest level, which allows more sophisticated analysis. CLV is one such metric that directly accounts for future value.

CLV can be defined as the total financial contribution from the current period into the future, revenues minus costs of a customer over his/her future lifetime with the company and therefore reflects the future profitability. It is a forward-looking customer metric that takes into account the current value and the future value provided by customers. It provides a dollar value for customer relationship. It helps distinguish customers according to their value over the life of their business with the firm. Future marketing strategies can then be planned accordingly for both current and future customers (Pochiraju *et al*. 2019).

Customer Churn

The Customer churn prediction can be defined as assigning a probability to leave the company next to each customer in the company dataset, to make it viable to the firm to indicate which customer has the highest propensity to churn. For the company to set a threshold to indicate which probability to churn they want to tackle. The customers clusters whom crossed this threshold will be targeted by a specially designed marketing plan to increase their retention. Different approaches introduced to increase customer retention (Figure 3), while targeting all of the churner's clusters will be Monterey exhausting, that's why the companies must select wisely which of these customers to target, instead of targeting every one also known as uplift modeling (Coussement *et al.* 2016).

There is a direct relationship between customer lifetime value and the ability to grow business. As such, the higher the customer churn rate are, the lower the business chances for growth. Even if the company has some of the best marketing campaigns in industry, the company's bottom line suffers if the company is losing customers at a high rate, as the cost of acquiring new customers is high. To reduce customer churn in a business is to build customer loyalty through relevant experiences and personalized service. A predictive approach is used to avoid future customer churn (Nabareseh 2017). Some companies are surveying customers who has already churned to determine their reason of leaving. The best way to avoid customer churn is to truly know its customers by meeting their expectations and making them satisfied with the business product and services (Keramati *et al.* 2016). To do that, the company need to have insights into customers through the use of Big Data and a customer data platform (Leonard 2014). Companies can anticipate customers' needs and issues. This includes identifying customers who are at risk of churning and working to improve their satisfaction.

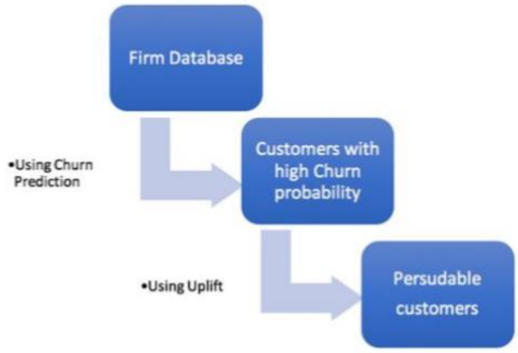

Figure 3 Customer churn model (Coussement *et al.* 2016)

Business Analytics Process

Business Analytics Process (BAP) is defined as a consistent process through which business objectives can be met and insights executed and then tested with the best in class data and advanced analytics driving strategies and executions. BAP begins with business and ends with deployment in six steps (Chorianopoulos 2016). However, just having these steps is not sufficient to ensure the success of an analytics endeavour. BAP also need to build in six feedback loops and formally establish an analytics sandbox. The following list describes the six steps of Business Analytics Process:

1. Business objectives. In these steps, essential business questions are raised and anticipated successes are defined. Business and key analytics leaders should jointly define the objective.
2. Data Audit. Once the objectives are defined, an experienced business analytics team should quickly focus on the data needed for the potential models and conduct a thorough data audit to determine data availability within the existing data infrastructure. If there are issues with the data quality or quantity, the team goes back to the first step to determine how to modify the objectives or further collect or cleanse more data.

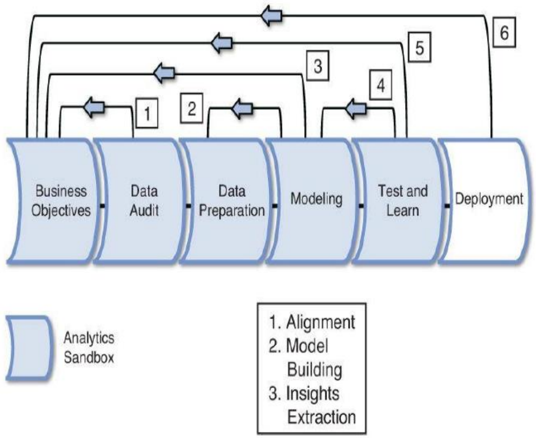

Figure 4 Business analytics process (Chorianopoulos 2016)

3. Data preparation. Once data is determined to be of adequate quality and quantity, it needs to be prepared appropriately into formats fed into models to be built.
4. Analytics modelling. The single step of modelling in CRISP-DM is expanded into three sub steps in BAP, as follows:
   - Business Alignment: Before any modelling can be done, the business objectives have to be translated into a series of analytics modelling objectives and briefs. To do this properly, someone who possessed intimate knowledge of moth modelling and business strategy must lead this step.
   - Model building—After aligning with business, the modeling goals can now be accomplished in actual modeling. Different models can be used to solve the same analytics problem, and they tend to produce results of different levels of lifts and insights. The models to use might depend on the data characteristics, business conditions, and model performance. A series of models is usually built and tested for comparison, and the best-performing model is then used for deployment. During this stage, some data may need to be further modified or augmented. This is Feedback Loop 2.
   - Model insight extraction—The importance of this step cannot be overemphasized. Analytics professionals love their craft (that is, building models), so they typically spend 10 percent of their time in data preparation, 70–80 percent in model building, and another 10–20 percent in extracting insights. Ideally, the proportion of time should be 1/3 spent in data preparation, 1/3 in model building, and 1/3 in extracting insights. More questions are often raised during this stage, and existing ideas and assumptions are modified as new insights are found. With these new insights, there should be a Feedback Loop 3 to refine, modify, or augment the original business objectives.
5. Test-and-learn stage—Once the models are validated and insights are shared and confirmed by the business, the model insights must be translated into actual business actions with the key drivers and metrics defined. First, the models are tested using a small subset of the data used for actual deployment. If the model insights were not supported by the test results, the models are revisited to determine whether the cause was errors in input data or models, or changes in market conditions or deficient marketing collaterals. This is Feedback Loop 4. To carry out the test-and-learn phase, the following steps are needed:
   - Operational alignment—During the field tests, all the functional and business units should be involved and trained for the actual deployment. Any potential kinks in the execution should be ironed out during the test and- learn phase.
   - Design of experiments (DOE)—In addition to testing whether the models work in the field as predicted, this is the step in which all the variations in the campaign promotions (for example, creative, copy, offer, fonts, and channels) and any other potential factors might be included into a multicell design. The aim is to determine the incremental effects of each of the levers or factors and determine the optimal combination of factors to the Dependent Variables (DV) such as conversion rates, site traffic, or response rates.
   - Scenario planning—The lever settings and the validated rates can then be fed back to the business strategy team to build a what-if simulator. A simulator is an invaluable tool for determining what the current investment level and its expected ROI should be, the alignment of all the stakeholders during the

deployment stage, and the field-tested reliable inputs to the future planning process. This is Feedback Loop 5.

- Deployment—This stage is where the rubber meets the road; all the insights and resources are deployed for real. During the deployment step, each stakeholder should measure and benchmark the results against what was discovered during the test-and-learn stage.

Predictive Analytics

Predictive Analytics is the process of finding interesting patterns and meaningful data (Abbott 2014). Predictive Analytics is a data-driven algorithm and obtains the key characteristics of the data model itself. Algorithms on predictive analytics automate the process of finding patterns in the data. A domain that will focus on predictive analytics that will build CRM is a business strategy that focuses on customers designed to optimize revenue, profitability, and customer loyalty (Chorianopoulos 2016). Many statistical and data mining techniques are introduced to investigate customer churn prediction.

In contrast to the market surveys, data mining techniques analyze the information obtained from both historical & current data to predict the patterns on historical data and future customer attitudes (Nabareseh 2017). The most common techniques used for prediction are Decision Tree (DT), Logistic Regression (LR), Support Vector Machine (SVM) and Neural Network (NN). Furthermore, the decision tree is used to resolve classification problems to divide the instances into two or more than two classes (Ibrahem 2012).

Similarly, logistics regression gives the probability by providing input/output fields, a set of equations, and factors causing customer churn. Predictive Analytics can divide into classifications and prediction, which can be subdivided into a decision tree, logistic regression, neural network and support vector machine (Avon 2016). To protect the profit and brand name, it is deemed necessary to retain the customers. Hence a collection of these customer data and necessary prediction would help identification of satisfied customer and would help the utilization of resources for these targeted customers (Abbot 2014)

The basic layer for predicting future customer churn is data from the past. The company look at data from customers that already have churned and their characteristics/behavior before the churn happened. There are the following steps to make a precise prediction (Umayaparvathi and Iyakutti 2016):

1. Use Case/Business Case: Step one is about understanding the business or use case with the desired outcome. Only by understanding the final objective we can build a model that is actually of use. In this case, the objective is to predict the customer churn by identifying customer dissatisfaction that will lead to potential churn candidates.
2. Data collection and cleaning: Understanding the context is possible to identify the right data source, cleansing the data sets, and preparing for feature selection or engineering. The predicting model is only as good as the data source.
3. Feature selection and engineering: In this third step, the company decides which feature they want to include in the model and prepare the machine learning algorithm's cleansed data to predict customer churn.

4. Modeling: with the prepared data, the company is ready to feed with the model. But to make good predictions, they first need to find the right model (selection) and secondly need to evaluate that the algorithm works.
5. Insights and actions: the last step is to evaluate and interpret the outcome.

The model proposed in this research is decision tree-based models, the importance of using decision tree models is that they are fast to compute and highly interpretable. Which have a great advantage for companies not only to estimate the propensity of each customer to churn but also to understand which features are essential to retain the customers in the future, also it has a great advantage over many machine learning algorithms, which the decision trees can deal with heterogeneous data, defined as data that contain numerical and categorical features. Decision Tree is one of the most potent and well-known classifications and prediction methods in applying data mining. The Decision Tree converts data into a decision tree and decision rules. This method's advantage is that it is useful in analyzing a large number of attributes of existing data and is easily understood by end-users. The decision tree algorithm is a data mining technique that recursively creates a data partition using the depth-first greedy approach or the breadth-first approach until all data entered into a particular class. When a Decision Tree model classifies an instance, the Decision Tree sorts it through the tree to the suitable leaf node. Each leaf node shows a classification (Tsai and Chen 2010). Nie *et al.* 2011 suggested that the Decision Tree not only produces results which are easy to understand but that it also can build models using numerical and categorical datasets. In the present research, Decision Tree techniques were applied to build a prediction model for customer churn from a coffee shop for two reasons. One reason relates to our goal of finding the features of churners and our need to understand if-then rules. Due to Decision Tree provides easy understanding rules. The other reason is the type of data. The data include numerical and categorical types, and the Decision Tree was suitable for these types of data.

Coffee in Kemenady Coffee Shop

Kemenday Coffee Shop and Co-working space is one of the most famous coffee shops in Bogor area. Kemenady Coffee Shop is a modern coffee shop that targets teenagers, young people and the elderly. They serve many different types of coffee for example milk coffee, cappuccino, americano and other menus. The very best seller is kopi rakjat which is a coffee serve with milk and brown sugar. Kopi rakjat also one of their specialty in this coffee shop. They also sell a different type of roasted coffee from around Indonesia for example, arabica, robusta, gayo, and papua's coffee. PT. Industry Kemenady Mandiri is a company engaged in the sale of coffee in the form of a coffee shop and a supplier of roasted arabica and robusta beans. PT. Industry Kemenady Mandiri has its own farmers' groups cultivated by Kemenady to produce a coffee bean that is processed into roasted coffee and powder form. Then the coffee powder is reprocessed into modern coffee drinks sold in Kemenady coffee shop in various kinds by the ongoing trends such as milk coffee or the rakjat's coffee, espresso, latte, cappuccino and more. While the products of arabica and robusta roasted coffee sold in packaged form in Kemenady coffee shop. Additionally, roasted coffee products are also marketed to the hotel, restaurant, and coffee shop in Indonesia. Based on the statement from the Kemenady coffee shop's operational manager, the majority of visitors are aged 14-25 years and there are more female visitors than men. The time most visited by customers is the range from 15:00 WIB to 21:00 WIB. Usually, at 15:00, filled by school children who stop by to enjoy coffee after school. While from 18:00 - 21:00, the majority is filled with customers who want to enjoy coffee after work. Friday to Sunday is the busiest day at Kemenady coffee shop.

METHODOLOGY

Research Framework

This research was divided into four stages. The first was to model the business process. The second was to build a CRM model based on satisfaction, RFM, CLV, and clustering. The third was to predict the customer churn using a classification decision tree, and the last stage was to evaluate the results to construct the marketing strategies. Figure 5 illustrates the research framework.

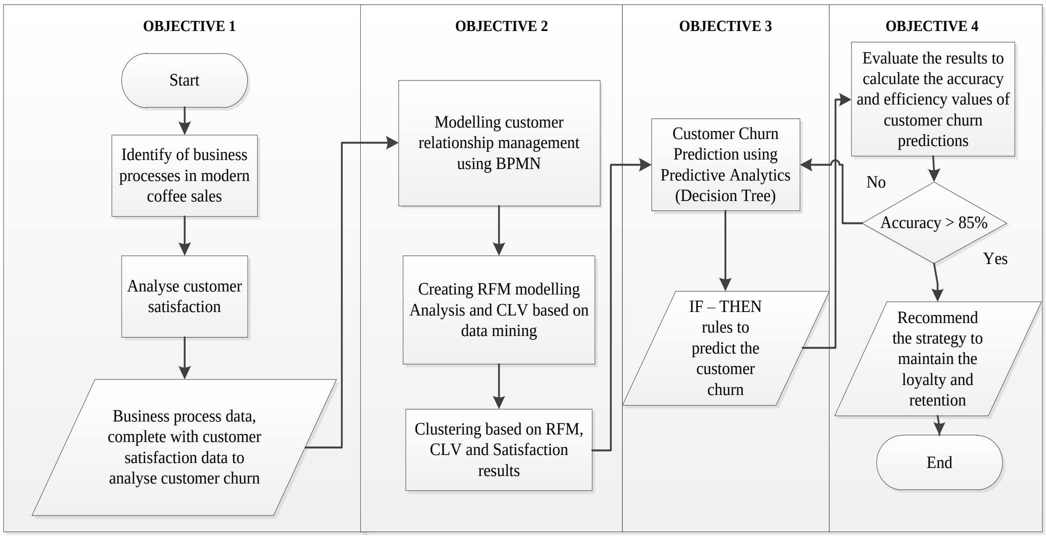

Figure 5 The research framework

Identification of Business Processes

Identify the business process is a process to analyze people who are involved in the process of selling and buying roasted coffee and coffee drinks present. Data collection techniques in the identification business process involve using survey techniques and interviews with customers in Kemenady Coffee Shop located in Bogor, West Java. Questionnaires for questions to be asked during the interview shown in the appendix 1.

Modeling Customer Relationship Management

Modeling customer relationship management is the process of interaction depiction of the company in all aspects of customers ranging from sales, acquisition, enhancement, up retention. CRM is a dynamic process that governs the relationship between the customer and the company so that customers can choose to continue the mutually beneficial commercial relations and anticipate that this relationship does not become unprofitable. In modeling CRM, the writer will be performed using the Business Process Modeling Notation (BPMN). It will explain a business process model for modern coffee drink, business process payments for products, business process delivery of products ordered, marketing of products and business processes.

RFM Modelling Analysis and Customer Lifetime Value

RFM modeling analysis was conducted to determine customer loyalty of parameters Recency, Frequency, and Monetary. The data used are each customer transaction data from July – December 2019. CLV is the value of the customer lifetime value calculated to determine the customer life cycle that results can determine the potential customers and customer loyalty in Kemenady Coffee Shop. Here is a mathematical formula for modeling CRM with RFM and CLV:

$$CLV_k \text{ value} = (NR_k \times WR_k) + (NF_k \times WF_k) + (NM_k \times WM_k) \qquad (1)$$

CLV_k = Value to the Customer Life Cycle k
NR_k = Normalization Recency
WR_k = Weight Recency
NF_k = Normalized Frequency of sales transactions
WF_k = Weight Frequency
NM_k = Normalization sales results
WM_k = Weight Monetary
k = customer

Description of RFM categories:
NRk = for (high, medium, low) NRK value is (1,2,3)
NFk = for (high, medium, low) NFk value is (3,2,1)
NMk = for (high, medium, low) NMK values are (3,2,1)

Clustering K-Means

K-means is a prototype-based, partitional clustering technique that attempts to find a user-specified number of clusters (K) represented by their centroids (Tarokh and Nekooei 2015). The K-Means clustering method attempts to group existing data, in this case, was the customer's transaction data, into groups with partitioning systems, where data in one group has the same characteristics as each other and has different characteristics than the data in the other group. We were using WEKA 3.9 to run the clustering phase.

K-Means belongs to the class of clustering algorithms called *hard partitioning* algorithms because every data point falls into one partition (cluster) and one only. A practitioner will typically only have to specify the inputs to the model and how many clusters the algorithm should find; the algorithm takes care of the rest. The following steps describe only what the algorithm does (Abbott 2014):

Initialization: Select the number of clusters to find. More sophisticated implementations of the algorithm will attempt a range in the number of clusters specified by the user and retain the single number of clusters that is considered best.

`

Either way, the user must have an idea of the number of clusters that exist in the data or that are desired to be discovered in the data.
Step 1: Assign one cluster center per cluster. Most software implementations identify a random data point in the training data for each cluster and assign this point as a cluster center.
Step 2: Compute the distance between each cluster center and every data point in the training data.
Step 3: Assign a label to each data point, indicating its nearest cluster center.
Step 4: Compute the mean value for every input in each cluster-based.
Step 5: Repeat Steps 2 through 4 until cluster membership does not change.

Customer Churn Prediction using Decision Tree Algorithm

The decision tree structure is made up of the root, internal, and leaf nodes. Entropy is the information needed to predict an event and give a probability distribution. Entropy is the number of bits needed to extract a class into a random sample of data. Entropy will be used to determine the first root for building a tree. The lowest entropy will be chosen as the root or leaf of the tree. In general, the C4.5 algorithm for constructing a decision tree is as follows.

a. Select the attribute as the root
b. Create a branch for each value
c. Divide cases in branches
d. Repeat the process for each branch until all cases in the branch have the same class.

Decision Tree uses information gain as its attribute selection measure. This measure is based on pioneering work by Claude Shannon on information theory, which studied the value or "information content" of messages. Let node N represent or hold the tuples of partition D. The attribute with the highest information gain is chosen as the splitting attribute for node N. Such an approach minimizes the expected number of tests needed to classify a given tuple and guarantees that a simple (but not necessarily the simplest) tree is found. The expected information needed to classify a tuple in D is given by Equation 2 (Han *et al*. 2012):

$$Info(D) = -\sum_{i=1}^{m} pi \log_2(pi) \qquad (2)$$

pi is the nonzero probability that an arbitrary tuple in D belongs to class Ci and estimated by $|Ci,Dj|/|D|$. A log function to the base 2 is used, because the information is encoded in bits. *Info*(*D*) is just the average amount of information needed to identify the class label of a tuple in *D*. Note that, at this point, the information we have is based solely on the proportions of tuples of each class. *Info*(*D*) is also known as the entropy of *D*.

Now, suppose we were to partition the tuples in *D* on some attribute *A* having *v* distinct values, $\{a_1, a_2, \ldots, a_v\}$, as observed from the training data. If *A* is discrete-valued, these values correspond directly to the *v* outcomes of a test on *A*. Attribute *A* can be used to split *D* into *v* partitions or subsets, $\{D_1, D_2, \ldots, D_v\}$, where D_j contains those tuples in *D* that have outcome a_j of *A*. These partitions would correspond to the branches grown from node *N*. Ideally, we would like this partitioning to produce an exact classification of the tuples. That is, we would like for each partition to be pure. However, it is quite likely that the partitions will be impure (e.g., where a partition may contain a collection of tuples from different classes rather than from a single class). How much more information would we still need (after the partitioning) to arrive at an exact classification? This amount is measured by Equation 3 (Han *et al. 2012*):

$$Info_A(D) = \sum_{j=1}^{v} \frac{|Dj|}{|D|} x Info(D_j) \tag{3}$$

The term $\frac{|Dj|}{|D|}$ acts as the weight of the *j*th partition. $Info_A$ (*D*) is the expected information required to classify a tuple from *D* based on A's partitioning. The smaller the expected information (still) required, the greater the purity of the partitions. Information gain is defined as the difference between the original information requirement (i.e., based on just the proportion of classes) and the new requirement (i.e., obtained after partitioning on *A*). That is in Equation 4 below.

$$Gain(A) = Info(D) - Info_A(D) \tag{4}$$

In other words, *Gain* (*A*) tells us how much would be gained by branching on *A*. It is the expected reduction in the information requirement caused by knowing the value of *A*. The attribute *A* with the highest information gain, *Gain* (*A*), is chosen as the splitting attribute at node *N*. This is equivalent to saying that we want to partition on the attribute *A* that would do the "best classification," so that the amount of information still required to finish classifying the tuples is minimal (i.e., minimum $Info_A$ (*D*).

Calculation of entropy values can be seen in the Equation 5 (Hssina *et al.*):

$$Entrophy(S) = \sum_{k=0}^{n} -pi * \log_2 p_i \tag{5}$$

Information:

S : case set

A : Attributes

n : number of attribute attributes

p_i : the proportion of S_i to S

Evaluation of Accuracy Level of Customer Churn Prediction Value

In this model, to calculate an indicator of success using the confusion matrix (Table 2). A confusion matrix is a method that is usually used to perform the calculation accuracy on the concept of data mining or Decision Support Systems (Gorunescu 2011). In measuring performance using a confusion matrix, there are four (4) terms representing the results of the classification process. The fourth term is True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). True Negative (TN) is the number of negative data is detected correctly, while False Positive (FP) is the harmful data but recognized as positive data.

Table 2 Confusion matrix

		True Values	
		True	False
Prediction	True	TP Correct result	FP Unexpected result
	False	FN Missing result	TN Correct absence of result

Precision is the data taken based on inadequate information. In binary classification, precision can be made equal to the positive predictive value. Equation 6 is a rule of precision:

$$\text{Precision} = \left(\frac{TP}{(TP+FP)}\right) X100\% \qquad (6)$$

A recall is a successful deletion of data taken from the data that is relevant to the query. In binary classification, recall known as sensitivity. The emergence of relevant data received approved by the question can be seen by the recall. High sensitivity explains that the model produces a few false negatives and this explain the high accuracy in the model (Ullah *et al.* 2019). Equation 7 is the rule of the recall:

$$\text{Recall} = \left(\frac{TP}{(TP+FN)}\right) X100\% \qquad (7)$$

Accuracy is the percentage of the total data identified and assessed. The accuracy tells that overall how often the model is making a correct prediction. Equation 8 is a rule accuracy.

$$\text{Accuracy} = \left(\frac{(TP+TN)}{(TP+TN+FP+FN)}\right) X100\% \qquad (8)$$

Although accuracy is intuitive and commonly used to compare prediction methods, it is not considered to be an optimum figure of merit for churn modelling because it is unreliable in a situation of class imbalanced, which is by far the most common in this application area (Nie et al, 2011).

Recommend the Strategy to Prevent Churn and Maintain Loyalty

One of the objectives of this research is to recommend the strategy to maintain customer loyalty and retention. To do that, the researcher will use expert opinions from the coffee community in Bogor to suggest how the researchers find a strategy to keep customer loyalty by not meeting the churn rules.

This strategy is produced by considering the factors that cause customer churn which is obtained by analyzing the results of a survey of registered customers and past transaction data. This strategy is expected to prevent future customers from following predicted pattern rules. The strategy designed by the 2 coffee experts with the 5P approach (Product, Price, Place, People, Promotion) and customer satisfaction survey results.

Indicator of Success

Each of the objectives has the key performance indicator (KPI). For objective 1 which is to model a new business process has the KPI to build a successful BPMN that fits with the coffee shop business process. For objective 2 which is to predict the customer churn using data mining tools has the KPI to build models based on RFM, CLV, and clustering analysis. For objective 3 which is to recommend a fit strategy to prevent churn and maintain loyalty has the KPI to produce marketing strategy recommendations that are in accordance with churn prediction results. For the last objective which is to evaluate the result from customer churn prediction has the KPI to build a model with an accuracy more than 85%.

RESULT AND DISCUSSION

The New Business Process in Customer Churn Prediction

The new business process was the proposed customer churn prediction model. It was illustrated through the system entity and the BPMN diagram. Figure 6 shows 11 attributes in the system: acceptable input, unacceptable input, stakeholder, roles, missions, and objectives, resources, acceptable output, unacceptable output, threats, controls, opportunities, and processes. The stakeholders are the customer, management staff and management analyst. The management staff consists of cashiers, financial staff and marketing staff. The difference with the management analyst is that the core parts of the coffee shop are the owners, marketing managers and data analysts who play a role in the final decision-making process. The customer churn prediction model was the acceptable inputs from the customer's transaction data to predict the rules that will tell the information about the customer's pattern. The acceptable output was the churn prediction model with 85% using the acceptable input, which was the customer transaction data and customer's survey. The indicator of success in this research was the churn prediction model with 85% and more accuracy. The stakeholders in this system were customer, staff management and management analyst.

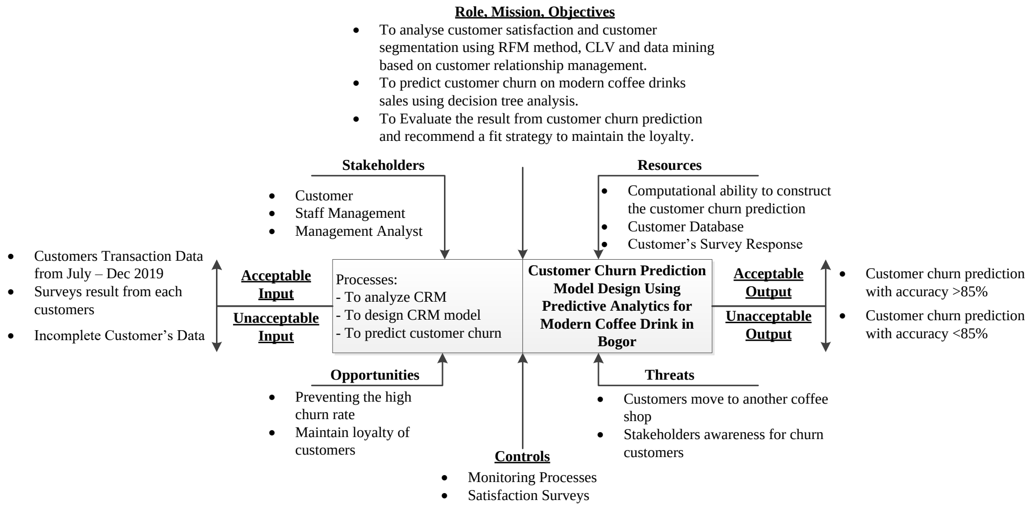

Figure 6 The system entity of the customer churn prediction model

BPMN Diagram

The new business process was shown through the BPMN (Business Process Modeling & Notation) diagram (Figure 7). The BPMN diagram was used to find out in detail the customer churn prediction model. The system had 3 stakeholders: customer, management staff and Internal management staff. It showed that the management analyst was responsible for deciding the strategy for preventing the churn in the future based on the churn prediction. It was calculated to find the customers' patterns and rules and recommend the perfect strategy to prevent the churn from happening in the future.

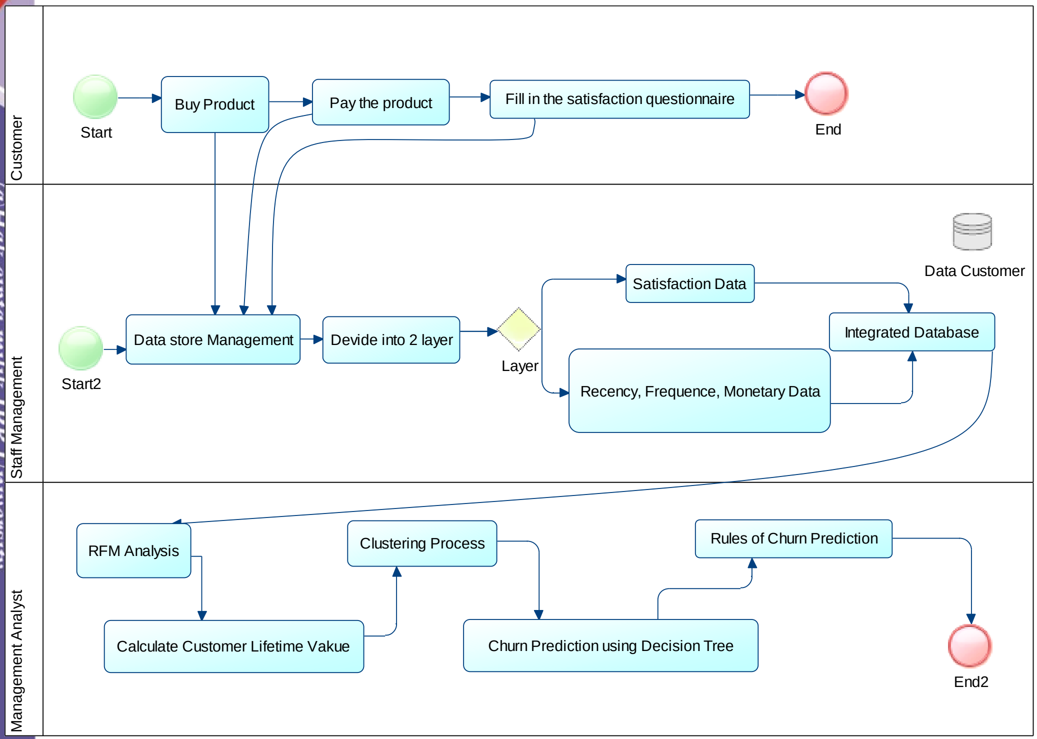

Figure 7 BPMN diagram

Customer Relationship Management Construction

The interview and survey results were customer satisfaction attributes and customer transaction's raw data from the past 6 months from July – December 2019. We conducted satisfaction surveys to 427 registered customers who made the purchases for the last six months. There were 405 satisfied and 22 not satisfied customers. The surveys used 5 attributes: product quality, price, services, emotional factor, and facilities.

Table 3 Question Lists for Customer Satisfaction Survey

No	Attributes	Questions
1.	Product Quality	Apakah rasa produk yang ditawarkan oleh Kemenady sudah sesuai harapan?
		Keragaman produk yang ditawarkan di Kemenady sudah sesuai harapan?
2.	Price	Harga produk yang ditawarkan di Kemenady sudah sesuai harapan?
3.	Sevices	Penampilan pelayanan di Kemenady sudah sesuai harapan?
		Kebersihan fasilitas sudah sesuai harapan?
		Menu yang ditawarkan sudah sesuai harapan?
		Alunan musik yang diputar di Kemenady sudah sesuai harapan?
		Kecepatan mengantarkan pesanan di Kemenady sudah sesuai harapan?
		Kesesuaian pesanan dengan produk yang disajikan sudah sesuai harapan?
		Kesigapan waitress di Kemenady sudah sesuai harapan?

		Keramahan pelayan Kemenady sudah sesuai harapan?
		Kemampuan pramusaji berkomunikasi dengan konsumen sudah sesuai harapan?
4.	Emotional Factors	Bentuk produk (penyajian makanan dan minuman oada wadahnya) yang ditawarkan sudah sesuai harapan?
		Warna dan rasa produk di Kemenady sudah sesuai harapan?
5.	Facilities	Kemudahan menemukan lokasi Kemenady sudah sesuai harapan?
		Kemudahan metode pembayaran sudah sesuai harapan?
		Kemudahan mendapatkan tempat duduk sudah sesuai harapan?

RFM analysis was also conducted to determine the Recency, Frequency, and Monetary status of each customer. We used the past transaction data based on their recent visit, frequent visit, and the money they spent in this coffee shop.

Recency is obtained by calculating the number of last days the customer made a coffee shop transaction. Frequency is obtained by calculating the number of times a customer makes a coffee shop transaction in 6 months. Monetary is obtained by calculating the total amount of money spent on making transactions at the coffee shop during the last 6 months. Table 3 describes the information and ranges used in the RFM analysis.

Table 4 RFM range

RFM	Range	Point	Note
Recency	1-40	3	High
	41-80	2	Moderate
	81-121	1	Low
Frequency	3-6	1	Low
	7-10	2	Moderate
	11-14	3	High
Monetary	Rp450.000 – Rp549.000	1	Low
	Rp550.000 – Rp649.000	2	Moderate
	Rp650.000 – Rp750.000	3	High

Customer Lifetime Value (CLV) was also conducted to discover a lifetime's rank based on their RFM status. From 427 customers, we received 15 ranks based on their CLV; and rank 1 is the most loyal customer in this coffee shop, and rank 15 is a customer with the highest potential to churn. The calculation is done using the CLV formulation, where the RFM points are multiplied by the expert's weight, which then adds up the entire RFM value (Table 4). The CLV results are then ranked from the largest to the smallest CLV values, where the largest CLV score gets the first rank, and the smallest value gets the last ranking. Table 5 shows the total customer in each rank that determined their potential to churn in the future.

Table 5 CLV calculation

No	No. Pelanggan	Nama	R Point	F Point	M Point	R Status	F Status	M Status	R Weight	F Weight	M Weight	CLV
1.	KMDBGR001	Yance	1	2	1	Low	Moderate	Low	0,3	0,7	0,35	1,35
2.	KMDBGR002	Daru	3	3	3	High	High	High	0,9	1,05	1,05	3
3.	KMDBGR003	Padma Hana	2	3	3	Moderate	High	High	0,6	1,05	1,05	2,7
4.	KMDBGR004	Kiandra	1	1	3	Low	Low	High	0,3	0,35	1,05	1,7
5.	KMDBGR005	Jaeman	1	3	3	Low	High	High	0,3	1,05	1,05	2,4
..	..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..	..
423.	KMDBGR423	Widia	1	3	2	Low	High	Moderate	0,3	1,05	0,7	2,05
424.	KMDBGR424	Tohir	2	3	1	Moderate	High	Low	0,6	1,05	0,35	2
425.	KMDBGR425	Hamdi	2	2	1	Moderate	Moderate	Low	0,6	0,7	0,35	1,65
426.	KMDBGR426	Riska	3	3	3	High	High	High	0,9	1,05	1,05	3
427.	KMDBGR427	Jingga	2	1	2	Moderate	Low	Moderate	0,6	0,35	0,7	1,65

Table 6 Total customer in each rank

CLV Rank	Total Customer
1	23
2	26
3	32
4	26
5	36
6	43
7	42
8	43
9	31
10	51
11	27
12	10
13	20
14	7
15	10

Clustering Process Based on CRM

K-means clustering is conducted to classify customers into 3 clusters. The categories are the potential to churn, loyal, and churn customers. From these results, we use it to see the pattern for each customer and make predictions based on the cluster obtained. Table 6 and Figure 8 describe the results of clustering running on the Weka software.

Table 7 Clustering result

Cluster	Status	Total Customers
0	Potential to churn	137 (32%)
1	Loyal	165 (39%)
2	Churn	125 (29%)

```
Final cluster centroids:
                                        Cluster#
Attribute           Full Data          0          1          2
                      (427.0)    (137.0)    (165.0)    (125.0)
=========================================================
No. Pelanggan       KMDBGR001  KMDBGR001  KMDBGR002  KMDBGR006
R Status                  Low        Low       High   Moderate
F Status                 High   Moderate       High        Low
M Status                 High       High       High        Low
CLV                    2.0979     1.9485     2.5488     1.6664
Category            Satisfied  Satisfied  Satisfied  Satisfied




Time taken to build model (full training data) : 0.02 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       137 ( 32%)
1       165 ( 39%)
2       125 ( 29%)
```

Figure 8 Clustering results

From the results of clustering, it was found that there were 32% of customers who entered the potential to churn cluster where coffee shops were advised to start taking steps to keep customers from leading to churn. In addition, there are 39% of loyal customers where the coffee shop should maintain customer loyalty by implementing a suitable strategy. The last cluster is customers who churn; there are 29% of customers.

Churn Prediction Using Classification Decision Tree

The goal of this step was the rule generation for preventive customer churn. First, this step needed the data from the customer past transaction and the RFM, CLV, and clustering results. It had seven categories: customer number, name, R status, F status, M status, CLV, and category. It had one target named class: potential to churn, churn, and loyal. These data were obtained from the coffee shop owner.

The decision tree constructed by recursively splitting the instance space into smaller sub-groups until a specified criterion has been met. The decrease in impurity of the parent node against the child nodes defines the goodness of the split. The tree is only allowed to grow until the decrease in impurity falls below a user-defined threshold. At this time the node becomes a terminal, or leaf node. We worked on the decision tree classification with the Jupyter Notebook python programming language. We divide the data into 70% training data and 30% testing data. It used 5 max_depth of the tree. The results of the decision tree were as described in Table 8.

Table 8 Rules description

IF CLV <= 2.17
1. IF 1.68 > CLV < 1.18, R = Low, F = Low, THEN **Churn**
2. IF 1.68 > CLV >= 1.18, R = Low, F = Low, THEN **Churn**
3. IF CLV <= 1.68, R = Low, F = Moderate, THEN **Potential to Churn**
4. IF CLV <= 1.68, R = Moderate, F = Low, THEN **Churn**
5. IF CLV <=1.68, R = Moderate, F = Moderate, Category = Satisfied, THEN **Churn**
6. IF CLV <=1.68, R = Moderate, F = Moderate, Category = Not Satisfied, THEN **Churn**
7. IF CLV <=1.68, R = High, THEN **Churn**
8. IF CLV > 1.68, F= Moderate, THEN **Churn**
9. IF 1.83 > CLV > 1.68, F = Low, Category = Satisfied, THEN **Potential to Churn**
10. IF 1.83 > CLV > 1.68, F = Low, Category = Not Satisfied, THEN **Potential to Churn**
11. IF 1.83 < CLV > 1.68, F = Low, R = Moderate, THEN **Churn**
12. IF 1.83 < CLV > 1.68, F = Low, R = High, THEN **Churn**
13. IF 1.85 > CLV > 1.68, F = High, M = Low, Category = Satisfied, THEN **Churn**
14. IF 1.85 < CLV > 1.68, F = High, M = Low, Category = Satisfied, THEN **Churn**
15. IF CLV > 1.68, F= High, M = Low, Category = Not Satisfied, THEN **Churn**
16. IF CLV > 1.68, F = High, M = Moderate, THEN Potential to **Churn**
IF CLV > 2.17
17. IF CLV > 2.17, F = Moderate, R = Moderate, Category = Satisfied, THEN **Potential to Churn**
18. IF CLV > 2.17, F = Moderate, R = Moderate, Category = Not Satisfied, THEN **Potential to Churn**
19. IF 2.48 > CLV > 2.17, R = High, F = Low, THEN **Loyal**
20. IF 2.48 > CLV > 2.17, R = High, F = Moderate, Category = Satisfied, THEN **Loyal**
21. IF 2.48 > CLV > 2.17, R = High, F = Moderate, Category = Not Satisfied, THEN **Loyal**
22. IF 2.17 > CLV > 2.48, R = High, Category = Satisfied, THEN **Loyal**
23. IF 2.17 > CLV > 2.48, R = High, Category = Not Satisfied, THEN **Loyal**
24. IF CLV > 2.17, F = High, Category = Satisfied, M = Moderate, R = Moderate, THEN **Loyal**
25. IF CLV > 2.17, F = High, Category = Satisfied, R = High, M = Low, THEN **Loyal**
26. IF CLV > 2.17, F = High, Category = Satisfied, R = High, M = Moderate, THEN **Loyal**
27. IF 2.55 > CLV > 2.17, M = High, THEN **Loyal**
28. IF 2.85 > CLV > 2.55, M = High, THEN **Loyal**
29. IF 2.85 < CLV > 2.55, M = High, THEN **Loyal**
30. IF CLV > 2.17, F = High, Category = Not Satisfied, R = Low, THEN **Potential to Churn**
31. IF CLV > 2.17, F = High, Category = Not Satisfied, R = Moderate, THEN **Loyal**

These rules can continue to change as data is available in real time. If the rule results show the level of loyalty is less or equal to 60%, coffee shops are advised to take action by implementing marketing strategies in accordance with their respective clusters. In addition, coffee shops should also take action if the churn percentage is more than or equal to 29% and the potential to churn percentage exceeds or is equal to 32%.

Evaluation using Confusion Matrix

Every model needs to evaluate the accuracy of the model. This accuracy will determine whether this model is feasible or not suitable for use in real situations. In this study, we use a confusion matrix consisting of true positive, false positive, true negative and false negative that describes the state of the data being evaluated.

- True positives: data points labeled positive which are actually positive.
- False positives: data points labeled positive which are actually negative.
- True negatives: data points labeled negative which are actually negative.

`

- False negatives: data points labeled as negative are actually positive.

The accuracy value is obtained by performing calculations according to the formula described in the methodology chapter. Detailed explanations with manual calculations are described in appendix 7.

Table 9 Confusion Matrix

Class	TP	TN	FP	FN	Precision	Recall	Accuracy
Potential to Churn	28	46	4	8	0,88	0,78	
Loyal	20	59	4	3	0,83	0,87	0,86
Churn	26	55	4	1	0,87	0,96	

In this research, the churn prediction model produces an accuracy of 86% where this model is classified as good because this model has a high sensitivity / recall value. High sensitivity explains that this model produces a few false negatives (Ullah *et al.* 2019).

Strategy Recommendations

The marketing strategy is designed for the three customer categories according to the predictive results of the rules. This marketing strategy is designed by discussing with coffee experts, namely the Kemenady Coffee Shop owner and one of the coffee experts, a Bogor coffee community member. Based on interviews with the two experts, several customer category strategic recommendations will be explained in Table 10.

Based on the results of the customer satisfaction survey (appendix 2), services is an attribute that has the highest dissatisfaction value compared to other attributes. Experts hope that there will be an increase in service not only in terms of facilities or quality but also by providing other benefits that can improve service. That is why this strategy was developed by considering the increase in services according to each customer cluster.

Table 10 Strategy Recommendations

Category	Suggested Strategy	Person in Charged
Potential to Churn	1. Follow up the customer by providing a reminder via the WhatsApp to the number that has been registered	Management Staff
	2. Notify periodically any promotional that are currently available	Management Staff
	3. Offer member programs to make them loyal	Cashier
Churn	1. Give special discounts targeted at churn customers but within a certain period, for example "a special promo only for you, 20% discount for all drinks, only today"	Management Staff
	2. Notify periodically any promotional that are currently available	Management Staff
	3. Offer member programs to make them loyal	Cashier
Loyal	1. Keep up with what they need, provide information about promos regularly to keep them loyal	Management Staff
	2. Implementing a member tier system such as silver, gold and platinum, each grade has a different benefit	Management Staff

Advantage and Disadvantage

This research showed the advantages of preventing the customer from churning in the future. This research also helped the coffee shop to be solved the problem of decreasing the profit. Meanwhile, the churn prediction can also show the customers' buying patterns, which will help the coffee shop design and apply the marketing strategies that fit with it.

Besides that, this research still has disadvantages. It required skills and application to run the model and it also required to update customers data every time because the model will run the real time data. In addition, data that can be used is customer data with certain attributes to predict customer churn. Therefore, the prediction of the churn model can only be run when the coffee shop has sufficient data for processing. The model cannot be applied if the coffee shop does not have an adequate database and of course it will not be able to produce predictions of rules that are really needed from a business to make further decisions. The model will continue to be used if the company or coffee shop applies a real time data system that keeps updating rules and strategies as the data progresses.

CONCLUSION AND RECOMMENDATION

Conclusion

The BPMN diagram was used to find out in detail the customer churn prediction model. The model has three stakeholders: customers, staff management, and customer churn prediction model. The methods used to meet this goal are RFM, CLV and Clustering where RFM generates a buying trend on each customer. CLV delivers lifetime results for each customer as well as customer loyalty rankings which has 15 ranks. Clustering generates customer groups based on RFM and CLV results to determine each customer's segment. We obtained 31 rules from the prediction we ran using a classification decision tree. The evaluation using the confusion matrix showed 86% accuracy in the prediction model. A marketing strategy recommendation was given to prevent the churn from happening in the future. We used 2 experts' opinions to design this recommendation.

Data mining approach and business analytics allows this model to be used to increase insights in the CRM field. In addition, with the ability of artificial intelligence to keep data continuous along with the increase in customers. Prediction rules that will continue to change following the movement of data every day.

Recommendation

Future work is recommended to apply the marketing strategies to see the actual result from the model we obtained and compare the result. It is also better to analyze and implement the model in more than 1 coffee shop to compare the results between each coffee shop against the accuracy of the churn prediction model.

The model needs the implementation of big data where the company always run the model using the real time data. The need for model improvements to increase the accuracy value of the prediction model. The next work may try to use other method such naïve Bayes and other rules methods to achieve higher accuracy.

In addition, it is better if you build a marketing strategy that really refers to each rule from the churn prediction, which of course requires a coffee expert in terms of service.

REFERENCES

Aagesen G, Krogstie J. 2015 Apr 1. BPMN 2.0 for Modeling Business Processes. *Handb Bus Process Manag 1 Introd Methods, Inf Syst*.

Abbott D. 2014. *Applied Predictive Analytics*. Indianapolis (IN): John Wiley & Sons, Inc.

Avon V. 2016. Machine learning techniques for customer churn prediction in banking environments. *Universit'a degli Studi di Padova.*

Bag D. 2017. *Business Analytics*. New York (NY): Taylor & Francis Group.

Bai Y. 2018. The Application of Customer Relationship Management and Data Mining in Chinese Insurance Companies. *Management & Engineering 31*: 1838-5745 31.

Chorianopoulos A. 2016. *Effective CRM Using Predictive Analytics*. West Suxxes (UK): John Wiley & Sons, Inc.

Coussement K, Lessmann S, Verstraeten G. 2016. A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. *Decis Support Syst*. 95.

Deb SM, Deb V. 2018. Artificial Intelligence: Creating Automated Insights for Customer Relationship Management. *8th International Conference on Cloud Computing, Data Science & Engineering.*

Design S, Wilburn M, Allen D, Hayes BE, Kessler S. 2008. *Measuring Customer Satisfaction and Loyalty*. Milwaukee (WI): American Society for Quality.

Duarte A, Laguna M, Mart R, Rey U, Carlos J. 2018. *Metaheuristics for Business Analytics*. Antigua and Barbuda (AG): Springer International Publishing.

Fry MJ, Cochran JJ, Ohlmann JW, Anderson DR. 2019. *Essentials of Business Analytics, Third Edition*.

Glady N, Baesens B, Croux C, Glady N. 2018. Modeling Customer Loyalty Using Customer Lifetime Value. *Katholieke Universiteit Leuven.*

Han J, Kamber M, Pei J. 2012. *Data Mining: Concept and Techniques*. Massachusetts (MA): Elsevier Inc..

Hssina B, Merbouha A, Ezzikouri H, Erritali M. A comparative study of decision tree ID3 and C4 . 5. (2):13–19.

Ibrahem HAK. 2012. a Comparative Study of Tree-Based Models for Churn Prediction: a Case Study in the Telecommunication Sector. *Dissertações de Mestrado em Estatística e Gestão da Informação*.

Keramati A, Ghaneei H, Mirmohammadi SM. 2016. Developing a prediction model for customer churn from electronic banking services using data mining. *Financ Innov*. 2(1).

Kiran D, Kanika T. 2015. Customer Churn Prediction in Telecommunication Industries using Data Mining Techniques- A Review. *International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 4*, ISSN: 2277 128X.

Konečnik RM, Ruzzier M, Hisrich R. 2014. Value, satisfaction and customer loyalty. *Mark Entrep SMEs*. November: 21–36.

Kotler P, Johnson SC, Distinguished S. 2015. *Customer Relationship Management*. Third Edit. New York (NY): Taylor & Francis Group, editor.

Leonard P. 2014. Customer data analytics: Privacy settings for "Big Data" business.

Int Data Priv Law. 4(1):53–68.

Maxim, Bruce R., Pressman RS. 2015. *Software Engineering: A Practitioner's Approach*. 8th Edition. New York (NY): McGraw-Hill Education.

Mohammadi VD, Albadvi A, Teymorpur B. 2014. Predicting Customer Churn Using CLV in Insurance Industry. *Shiraz J Syst Manag*. 2(1):39–49.

Nabareseh IS. 2017. *Predictive Analytics : A Data Mining Technique in Customer Churn Management for Decision Making*. Czechia (CZ): Tomas Bata University in Zlín.

Nie G, Rowe W, Zhang L, Tian Y, Shi Y. 2011. Credit card churn forecasting by logistic regression and decision tree. *Expert Syst Appl*. 38(12):15273–15285.

Nettleton D. 2014. *Commercial Data Mining Predictive Analytics Projects*. Massacusetts (MA): Morgan Kaufmann.

Olson D. 2017. *Recency Frequency and Monetary Model*. Singapore (SG): Springer Nature Singapore.

Pankaj MM. 2019. Demand Chain Management : Enhancing Customer Lifetime Value Through Integration of Marketing and Supply Chain Management 7. June. *The IUP Journal of Business Strategy*, Vol. XII, No. 3.

Pochiraju, Bhimasankaram; Seshadri S. 2019. *Essentials of Business Analytics: An Introduction to the Methodology and its Applications*. 1st ed. Switzerland (AG): Springer International Publishing.

Putler DS, Krider RE. 2012. *Customer and Business Analytics*. Florida (FL): Taylor & Francis Group.

Santoso AS, Erdaka A. 2015. Customer Loyalty in Collaborative Consumption Model: Empirical Study of CRM for Product-Service System-Based e-Commerce in Indonesia. *Procedia Comput Sci*. 72:543–551.

S. A. Qureshi, A. S. Rehman, A. M. Qamar, A. Kamal, and A. Rehman. 2013. Telecommunication subscribers' churn prediction model using machine learning. *Eighth International Conference on Digital Information Management (ICDIM 2013), Islamabad*, pp. 131-136.

Sedkaoui S. 2018. *Big Data Analytics for Entrepreneurial Success*. Business Science Reference.

Shmueli G, Bruce PC, Yahav I, Patel NR, Kenneth C. Lichtendahl J. 2018. *Data Mining for Business Analytics*. New Jersey (NJ): John Wiley & Sons, Inc.

Tarokh MJ, Nekooei A. 2015. Customer Clustering Based on Customer Lifetime Value : A Case Study of an iranian bank. *Int J Infirm Commun Technol Res*. 7(2): 71–90.

Triyanti DR. 2018. *Outlook Kopi 2018*.

Tsai C-F, Chen M-Y. 2010. Variable selection by association rules for customer churn prediction of multimedia on demand. *Expert Syst Appl*. 37(3): 2006–2015.

Ullah I, Raza B, Malik AK, Imran M, Islam SU, Kim SW. 2019. A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector. *IEEE Access*. 7:60134–60149.

Umayaparvathi V, Iyakutti K. 2016. Attribute selection and Customer Churn Prediction in telecom industry. *International Conference on Data Mining and Advanced Computing (SAPIENCE), Ernakulam*, 2016, pp. 84-90

Valacich JS, Goerge JF. 2017. *Modern System Analysis and Design*. New York (US): Pearson.
Wasson CS. 2016. *System Engineering Analysis, Design, And Development: Concepts, Principles, and Practices*. New Jersey (US): John Wiley & Sons
Weske M. 2019. *Business Process Modelling Foundation*. Heidelberg (DE): Springer.
Zalaghi Z, Abbasnejad VY. 2014. Measuring customer loyalty using an extended RFM and clustering technique. *Manag Sci Lett*. 4(5): 905–912

BIOGRAPHY

The author was born on September 24th 1995 in Pekanbaru, Riau. The author is the first child of Mr. Rinaldi and Mrs. Mulhayati. In 2017, the author graduated from the bachelor program of Industrial Engineering - Trisakti University with a thesis entitled "System Business Intelligence Model Design to Analyze Marketing for Bread Products in Randy Bakery Factory Using Data Mining and Cube Method." While pursuing the master's degree, the author also actives in a different kind of activity such as voluntary works and freelances. The author also participated in the Business and Innovation Engineering Conference and published a paper with a title of "A Comparative Study of Supervised Learning Models for Churn Prediction: A Case Study in a Coffee Shop."