Pay Later Risk Management: A Review of FMECA and Potential Customer Prediction Frameworks Through the Application of Machine Learning

Arif Furqon Nugraha Adz Zikri[1], Wiwin Suwarningsih[2]
[1,2]Master of Management Science, Faculty of Economics and Business, Padjadjaran University, Bandung, Indonesia

Article Info

Article history:

Received May 20, 2023
Revised Aug 19, 2023
Accepted Oct 20, 2023

Keywords:

FMECA
Critical Matrix
Cause and Effect Diagram
Machine Learning
Pay Later

ABSTRACT

The development of technology continues to develop and gradually change the way people buy such as on online shopping sites. The increase in internet use, especially in the use of E Commerce, has given birth to great potential in the market, especially in Indonesia. These changes prompted the birth of various payment methods. One of them is Pay Later. 27% of the 3560 samples decided to use Pay Later with all the conveniences offered. However, the development of Pay Later is not synchronized with good risk management. The use of Pay Later, which is not targeted at the right consumers, causes PT. XYZ suffered losses due to 22.37% of users defaulting on Pay Later installments. The purpose of this study is to reduce Pay Later default users by answering what factors cause consumers to default. To support this study, the authors used FMECA, Cause Effect Diagrams and conducted tests using Machine Learning to improve company efficiency. Through critical matrix analysis, the author gets 3 priority failure modes, Users default, users disappear, and users experience payment delays. In solving the problems in this study, the authors provide recommendations in the form of a new framework in the form of analyzing the best Pay Later offers by analyzing consumer behavior patterns in an E Commerce by utilizing Machine Learning. However, future research will need to be conducted correlation analysis and static testing in testing attribute correlation before testing algorithms when building machine learning models. The authors also suggested comparing using other methods to improve risk management in this study

Corresponding Author:

Arif Furqon Nugraha Adz Zikri,
Master of Management Science,
Faculty of Economics and Business,
Padjadjaran University,
Bandung, Indonesia.
Email: arif19003@mail.unpad.ac.id

1. INTRODUCTION

The Drastic changes in the market followed by the development of information technology have created various opportunities and challenges on the growth of the internet, paid online learning sites and information technology by generating new ways of learning in understanding consumers [1]. The evolution and development of the internet, especially regarding the way of transactions such as on online shopping sites has not only changed the way people buy but also changed the strategies used for marketing and payment methods to attract and retain customers by adjusting communication in a certain way [2].

The increase in internet use, especially in transactions on E-Commerce has created great potential in the market, especially Indonesia. In February 2022, the use of the internet to access E-Commerce accounted for 33% of all activities carried out on the Internet in Indonesia [3]. The increase in the use of E-Commerce has encouraged the birth of payment methods that can provide convenience, flexibility, efficiency and simplicity in making transactions [4], [5]. One of them is the birth of "Shopping Now Pay Later" or better known as Pay Later. A feature that applies the concept of a credit card where users can directly buy and use services directly with payments that can be made in installments [6]. The growth of transactions using pay later is quite large. In research conducted by Katadata and Kredivo, it shows that 27% of 3,560 people shopping in e-commerce decide to make payments in installments or pay later during 2020. It is estimated that the pay later usage rate will grow by 33.3% during 2022-2028 in Asia Pacific, with total money transactions on pay later predicted to increase from 82.81 billion dollars in 2021 to 749.22 billion dollars in 2028 [7]. One of the reasons the pay later feature is growing rapidly in Indonesia is because of the various conveniences offered that really help its use to meet daily needs, such as to manage finances, fulfill needs and simplify the payment transaction process from traveling to shopping [2].

The growth of internet platforms, especially in transacting on E-Commerce has created challenges for businesses and professionals to run businesses [8]. However, by presenting unsecured loans like conventional banks in general and relatively easy requirements, it creates high potential that can affect the sustainability of the company [5]. The decision to provide credit has a high risk of the inability of prospective debtors to pay their credit obligations in accordance with a predetermined period. So, to maintain and minimize these risks, Pay Later provider companies must be able to carry out very careful assessments and considerations by conducting various analyses to describe how the credibility of prospective debtors is in the future. In addition, Pay Later provider companies also need to supervise prospective debtors, both by direct supervision and indirect supervision [10]. This credit supervision aims to minimize and prevent credit irregularities or credit risks [2]. The high usage of Pay Later is often not targeted at the right consumers. In 2022, there are at least 22.37% of customers from PT. XYZ who failed to make pay later payments that caused losses to the company. The future of E-Commerce depends entirely on technology and professionals capable of creating and developing customer behavior patterns analysis to acquire potential customers [11].

With the high value of consumers who default on pay later, researchers are encouraged to develop a frame of reference that can be used to analyze potential customers using Machine Learning in the use of Pay Later and look for factors that cause failures in payments for errors in analyzing Pay Later user customers. So, this paper will specifically answer 3 questions. First, what kind of factors cause default through the FMECA analysis. Second, how to leverage Machine Learning to predict potential customers to reduce problem users and reduce bad debts. Third, what kind of framework can be used to reduce pay later problems by utilizing Machine Learning.

2. RESEARCH METHOD

The trend of online shopping continues to grow rapidly, especially with the presence of the Pay Later payment method. Predicting consumer behavior is an important aspect to optimize Pay Later user goals as a step to prevent consumer default while still paying attention to customer comfort to increase customer revenue and loyalty [12]. By collaborating Machine learning and FMECA, the framework of the study is as follows:

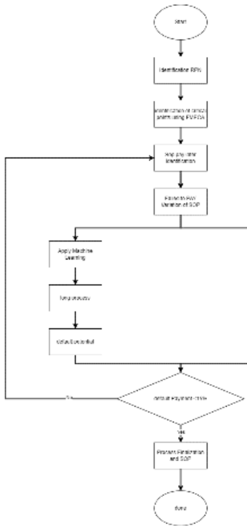

Figure 1. Research Methods

2.1. FMECA

There are various methods that can be used to conduct risk testing to determine the cause of a failure in a process or system. 2 methods that are often used are FMEA (Failure Mode Effect analysis) and FMECA (Failure Mode Effect and Critically Analysis) [13]. The FMEA (Failure Mode Effect Analysis) method is a method that serves to identify by testing various possible possibilities and the occurrence of failures caused by failures in a system or process that are not efficient and effective so as to cause failure loopholes. FMECA (Failure Mode Effect and Critically Analysis) is a method derived from the development of the FMEA method [14], [15]. The FMECA method is used to analyze potential failures by adding critical point analysis such as using the Risk Matrix [16]. In this study, the authors used the FMECA method where FMECA has the advantage of combining the FMEA method and critical analysis using a criticality matrix to make it easier to design SOPs in the analysis of potential risks that may occur in a process.

According to John (2015), FMECA is a methodology used to identify and analyze:

1. Analysis of all possible potential failures of various processes and systems
2. Analyzing the effects of such failures on systems and processes and compiling SOPs, how to avoid such failures.

FMECA consists of 2 analysts, namely using the FMEA (Failure Mode Effect Analysis) method and the critical analysis method. The FMEA method functions in conducting an analysis and evaluation of the risks that have a probity of problems The method to be used is to use the Risk Priority Number (RPN). In compiling the RPN, the variables needed to evaluate the cause of the failure and identify the possible effects that will result from the failure. The variables are to determine how much severity (severity), how much the failure rate (occurrence), and the detection rate (detection). Conventionally, the steps in compiling the FMEA are as follows: [17]

A. Identify the failure mode of each process that may or has occurred and define the potential impact it will have on the company's performance.
B. Determines the severity rating (S) of the failure mode, which defines how serious the impact of the failure mode is.
C. Determines the occurrence (O) of a frequency that has defined how often the failure occurs.
D. Determine a detection rating (D) that describes how it performs in detecting the possibility of such failures.
E. Determine the Risk Priority Number (RPN). The RPN formula is as follows:

$$RPN = S \times O \times D$$

F. Processes, components, or outputs that have a high RPN value will be selected as a critical risk level so that they need immediate treatment.
G. Defining the Risk Matrix and designing SOPs

Table 1. Severity Index

Rating	*Effect*	*Severity Effect*
7	Hazardous Without Warning	The failure rate is very high when the failure mode affects the performance of the company and lenders
6	Hazardous With Warning	Failure rates are very high when failure modes have the potential to affect the performance of companies and lenders
5	High	High failure rate when causing losses to lenders
4	Moderate	Intermediate failure rate resulting in minor losses
3	Low	Failure occurs with late payment
2	Very Low	Failure occurs when there is an extension of the loan period
1	None	No Influence

Table 2. Occurrence Index

Rating	*Probability of Occurrence*	*Failure Probability*
7	Very High: Failure is almost inevitable	>1 in 2
6		1 in 3
5	High: Repeated Failures	1 In 8
4		1 In 20
3		1 In 80
2	Moderate: Occasional failures	1 In 400
1	Low: Relatively Little Failure	1 In 8000

Table 3. Detection Index

Rating	*Detection*	*Probability of Detection*
7	Very Remote	Very little ability to detect failure rates on *Pay Later*
6	Low	Low ability to detect failure rates on *Pay Later*
5	Moderate	Medium ability to detect failure rates on *Pay Later*
4	Moderate High	Very moderate ability to detect failure rates on *Pay Later*
3	High	High ability to detect failure rates on *Pay Later*
2	Very High	Very high ability to detect failure rates on *Pay Later*
1	Almost Certain	Almost certainly the ability to detect failure rates on *Pay Later*

2.2. *Risk Matrix*

Risk Matrix is a follow-up step after compiling the RPN to conduct a critical analysis in accordance with predetermined criteria [18]. The result of the Risk Matrix is that items included in the rating of risk are defined based on the highest to lowest level of risk. Overall, the FMECA results and risk matrix are presented in the FMECA worksheet [18]. The table says risk categories based on the effect of using Pay Later.

Table 4. Severity of Consequences

Severity of Consequences	
Category	**Definition**
Catastrophic (I)	Causing a *Company Shutdown*
Critical (II)	May cause the company to experience a decline in performance
Marginal (III)	Can experience lander drops
Negligible (IV)	Can operate with little risk

Table 5. Severity Frequency

Severity of Frequency		
Frequency of events	**Definition**	
	Qualitative	**Quantitative**
Frequent	Frequent	≥ 1000
Probable	Highly Possible	≥ 500
Occasional	Common	≥ 100
Remote	Rare	≥ 50
Improbable	Unlikely to happen	< 10

Table 6. Risk Matrix

Frequency	**Frequent**	**Probable**	**Occasional**	**Remote**	**Improbable**
Catastrophic (I)	1 High	2 High	4 High	8 Passable	12 Passable
Critical (II)	3 High	5 High	6 Passable	10 Acceptable	15 Acceptable
Marginal (III)	7 Passable	9 Passable	11 Acceptable	14 Acceptable	17 Acceptable
Negligible (IV)	13 Acceptable	16 Acceptable	18 Acceptable	19 Acceptable	20 Acceptable

2.3. *Application of Machine Learning*

The application of machine learning can be useful for making predictions about choices, interests, and factors that affect consumers. This engineering technique turns large amounts of data into significant predictions [19], [20]. This allows for companies to go towards thinking in formulating strategies through data analysis [21]. With prediction automation, analytics teams will be able to act operationally and market. The application of machine learning in this study was used to analyze consumers who have the potential to be given access to pay later services [22], [23]. The algorithm that will be used in this study is as follows:

2.3.1 Decision Tree

Decision tree is a classification method by representing a tree structure with each node presenting an attribute, while its branch represents the value of an attribute, and the leaf represents a class of attributes [24]. In general, the concept is to turn data into a decision tree by applying rules. The benefit of a decision tree is its ability to change complex decision-making processes into simpler ones so that decision making will better interpret the solution of problems. Decision trees are also useful for exploiting data, finding hidden relationships between a number of prospective input variables and a target variable [14].

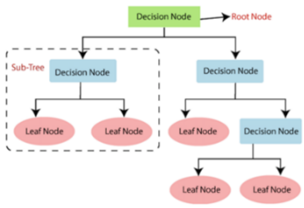

Figure 2. Decision Tree Visualization
Source: Decision Tree – Toward data Science (2020)

In calculating the decision tree, it is necessary to formulate entropy and Information gain which is defined as follows:

Calculate the entropy and information gain of each attribute by using the formula:

$$Entropy(S) \ = -P_{+}log_{2}P_{+} - P_{-}log_{2}P_{-}$$

With:

S : sample space (data) used for training

P_{+} : many samples that are likely to be positive (Healthy) in the sample data for certain criteria

P_{-} : many samples that are likely to be negative (Sick) in the sample data for certain criteria

$$Gain(S,A) \ = \ Entropy(s) - \sum_{v \notin Value(A)} \frac{|S_v|}{|S|} Entropy(S_i)$$

Information:

A : Attributes

V : Represents a possible value for attribute A

Value (A) : the set of possible value values for attribute A

$|S_v|$: number of samples for value V

$|S||$: the sum of the entire sample data

$Entropy(S_i)$: *entropy* for sample samples that have a value of V

2.3.2 Random Forest

Algorithms are technically a method Computationally based classification, consisting of a mutually independent classification tree. A Random Forest is a rule-based model, which uses the attribute with the highest covariance as its root, handling the most important data first. At this point, if the data used will be grouped into 2, that is, the training and validation data with a composition of 2/3 of the entire sample is training data while 1/3 is validation data[19], [22]. Both samples will be used to estimate the results of Random Forest performance.

The final result of the classification using Random Forest will be taken and flattened using the arithmetic mean. All results from Ntree in each class will be counted with probabilities assignments. The inputted data will choose a healthier and more fluid decision tree [22], [26].

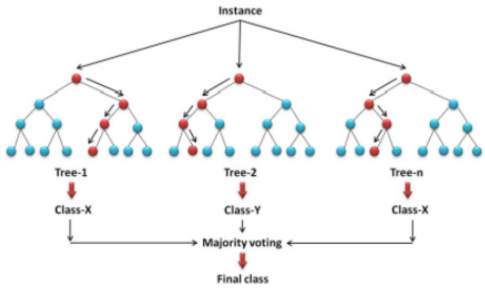

Figure 3. Random Forest Model Visualization
Source: Toward Data Science

2.3.3 XG Boosting

The XG Boost method is a popular designation for the extreme gradient boosting method which is one of the methods that belongs to the machine learning algorithm type, namely supervised learning. The extreme gradient boosting method is a development of the Gradient Boosted Decision Tree (GBDT), where gradient boosting is a machine learning technique that produces prediction models from the results of combining weak prediction models or referred to as ensembles to perform regression or classification tasks [27].

Gradient boosting is built using the boosting method, where to minimize errors or residues in the model, a new model is created and will continue to be added until no mistakes can be made [16].The difference between XG Boost and gradient boosting is the optimization offered on XG Boost, in the XG Boost model can implement parallelization that makes the model development process efficient.

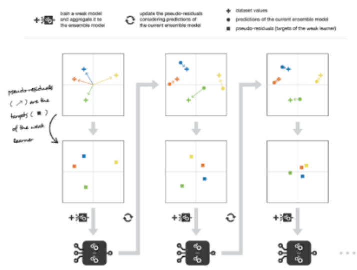

Figure 4. XG Boosting Model Visualization
Source: Ensemble methods: bagging, boosting and stacking – Towards Data Science

2.3.4 KNN

KNN is a machine learning algorithm without strong supervision. The KNN algorithm has lazy training properties, meaning it does not use training data points to create models. All training data is used at the testing stage to make the process of training faster. The concept behind KNN is that if customers always share the same opinion among a variety of different questions, those neighbors are likely to have the same taste and purchasing behavior for the same product. In practice, this machine learning algorithm model first maps each attribute to a higher dimension and then gathers the customers closest to each other [19]. Then, if there is a new customer data place into this research model, it will provide predicted results based on the most votes from its closest point [28].

In theory, if using manual calculations, calculating the distance between 2 points in the KNN algorithm can use the Euclidean Distance method which can be used on 1 dimensional space (one variable or independent variable), 2-dimensional space (2 free variables), or multi-dimensional space (more than 2 free variables) [29], [30]. In general, the Euclidean distance formula at 1 dimensional space is as follows: [26].

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^{n} (x_{1i} - x_{2i})^2}$$

While variables with more than one use the following equation.

$$dis = \sqrt{\sum_{i=0}^{n} (x_{1i} - x_{2i})^2 + (y_{1i} - y_{2i})^2 + \ldots\ldots}$$

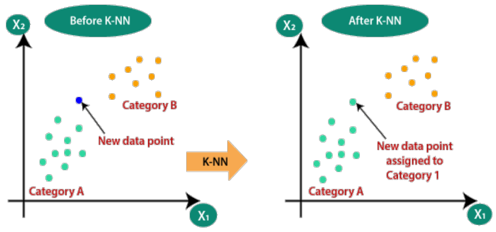

Figure 5. KNN Model Visualization
Source: Toward Data Science

2.3.5 Support Vector Machine

Support Vector Machine is another promising machine learning model. Unlike other traditional methods such as LSE, SVM only considers the error distance from its Support Vectors which is a point close to the decision limit of the input data [31]. Support Vector Machine adaptively adjusts the constraints to fit the parameters of each category, then, selects spinach that has maximum fault tolerance to be the result and ignores outliers beyond its limits.

Therefore, even if there is an outlier in the training data, the SVM will automatically ignore the outlier, data will thus be affected by it, See figure 6.

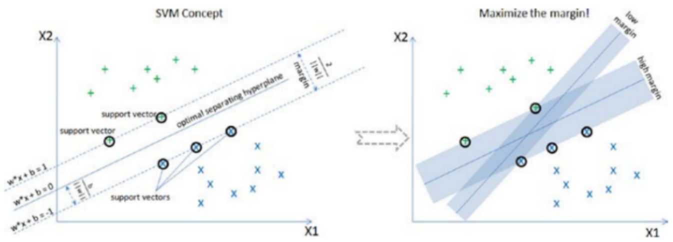

Figure 6. Visualization Model Support Vector Machine
Source: Toward Data Science

To maximize margins, we need to minimize:
$(1/2)||w||2$ $subject\ to\ yi(WTXi + b) - 1 \geq 0$ for all i.
The final SVM equation can be written mathematically as follows:

$$L = \sum_i di - \frac{1}{2} a_i a_j y_i y_j (\underline{XI}\ \underline{Xj})$$

2.4. Model of Determining Potential Customers

This platform will explore the possibility of developing algorithms that can be used to predict potential customers through purchasing behavior while using an E Commerce platform. Understanding buyer intentions is critical to reducing problem customers. Activities that can be carried out without face-to-face so that every action in an E Commerce can be a reference regarding financial capabilities and buyer behavior [32]. To describe how to determine the model of determining potential customers, the authors define as follows:

1) Machine Learning Thought Pipeline

The fast-paced movement of data demands that the analytics team be able to analyze consumers quickly and precisely for the convenience of consumers. In analyzing potential consumers on an E Commerce platform, one way that can be done is by analyzing customer behavior patterns. Customer behavior patterns are factors that involve purchasing or using a product and how it can be achieved, selected, and purchased. Machine learning needs to be done to analyze customers quickly so that the analytics team only needs to do additional analysis to potential customers or consumers [22], [33]. In this study, behavioral patterns are described as transaction activities so far as a way to see how trustworthy consumers are in being given access using Pay Later. Here is the thinking flow of using machine learning.

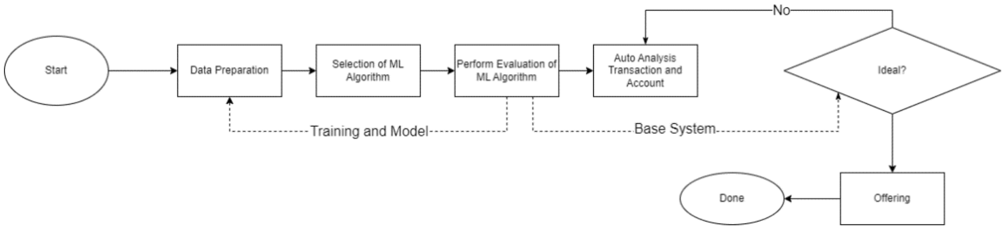

Figure 7.Machine Learning pipelines

The use of machine learning needs to build a model as a base system to predict whether the customer is a potential customer to be given Pay Later access. In building models, Machine Learning needs to learn first using Training data. Data Training is data used by machine learning to learn how to define potential customers. Then Machine Learning needs to define the previous data in various algorithms, build models, and test those algorithms with the accuracy level output of the prediction models on each algorithm. The algorithm that has the highest level of accuracy is the algorithm that will be used to predict potential customers [34].

2) Dataset

The dataset used in this study to test the proposed framework model is E Commerce transaction data at PT. XYZ consists of 10 Features including categorical and numeric, with a total of 11,289 sessions belonging to different users. These features are shown in table 7.

Table 7. Dataset Feature

Features name	Description
Verification Account	Verify the account with the Boolean data type caption "Already" or "Not yet”
Gander	Gender Description
Age	Age Description
Education	Description of Last Education
Transaction on E - Commerce	Description of transaction range in the last 3/6 months and 1 year
Salary	Income Description
Delivery Location	Boolean data type, delivery location is consistent or not
Length of use of E Commerce platforms	Duration of use of platform E - Commerce
BI Check	Boolean data type, User is a bad or good credit score

3. RESULTS AND DISCUSSION

3.1. Result

The following are the results of FMECA Analysis and Machine Learning Implementation as follows:

Failure Analysis and RPN calculation

In conducting a failure analysis, there are 2 ways that can be used. Firstly, Risk Priority, Second using Critical Number (CN) in determining the RPN value, what must be done is by multiplying the Severity, Occurrence, and detection values. The value of each of these periods is the result of identification carried out by certain departments such as the Management Process. The formula for calculating the Risk Priority Number (RPN) is as follows [35].

$$RPN = S \times O \times D$$

In the following table is the value of the RPN calculation.

Table 8. RPN

Failure Mode	S	O	D	RPN
Pay Later Payment Failure	9	5	9	405
Customer disappears	10	7	10	700
Late Payment of Pay Later	7	8	8	448
Extension of Pay Later payment period	4	10	3	120

Criticality Matrix

The criticality matrix functions to determine failure priorities based on the severity and extent of frequent occurrence of a failure mode. In this study, the data from the severity of consequence and severity of frequency analysis of each failure mode will be processed to be classified into a rating of risk as described in table 9 by combining the degree of consequence when a failure occurs and the potential for failure to occur. The criticality matrix referred to in this study is a failure mode categorized as High in the rating of risk table based on the results of risk matrix analysis in the FMEA method [14]. The high category in the Risk Matrix has an average level of frequency of occurrence and a higher level of effects caused by impacts compared to others [36]. The following is the result of the matrix risk analysis on each failure mode presented in the table below.

Table 9. Rating of Risk

Failure Mode	Rating of Risk	RPN
Pay Later Payment Failure	High	405
Customer disappears	High	700
Late Payment of Pay Later	High	448
Extended period of Pay Later payments	Acceptable	120

Machine Learning Implementation

Implementing Machine Learning, there are various ways that can be used, such as using Python by creating coding scripts or using analytics platforms such as RapidMiner, Knime, JASP, Orange, and other analytics platforms demit [37], [38]. In this study, the authors used the RapidMiner platform to make it easier to display visualizations in algorithm testing. RapidMiner itself can visualize simple and relatively easy to use by dragging and dropping a node that defines certain activities. The visualization of testing algorithms for determining Machine Learning models in predicting Potential Customers is as follows:

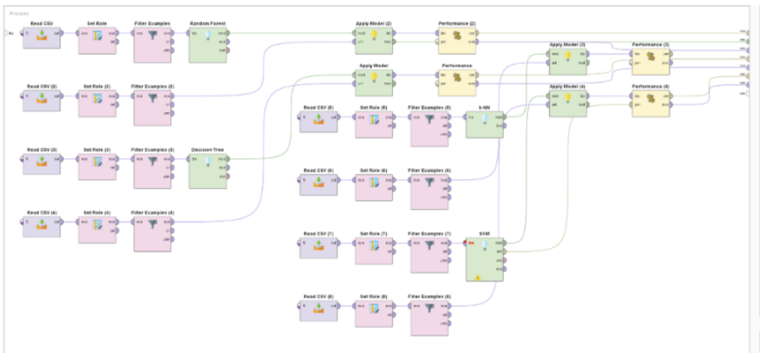

Figure 8. Machine Learning Implementation Visualization

3.2. Discussion

From the results of the RPN calculation, it is known that 4 failure modes obtained from PT. XYZ. From the results of the analysis, it is then analyzed using a criticality matrix to focus on failure modes that have a high rating of risk. After understanding the risks that must be overcome, the author analyzes the causes of the emergence of this failure mode in order to design SOPs to overcome these problems using the cause-and-effect diagram method.

Cause and Effect Diagram

Based on the results of the criticality matrix as shown in table 9, it is shown that priorities 1, 2, and 3 have something in common, namely the problem of customers experiencing bad debts. In the analysis of the cause-and-effect diagram or fish bone, the author analyzes various factors, namely: humans (Companies), Methods (opening pay later access), Machines (tools used), Materials (Lenders). The purpose of compiling Fishbone is to create a work plan to improve Pay Later services through 5W+1H [15].

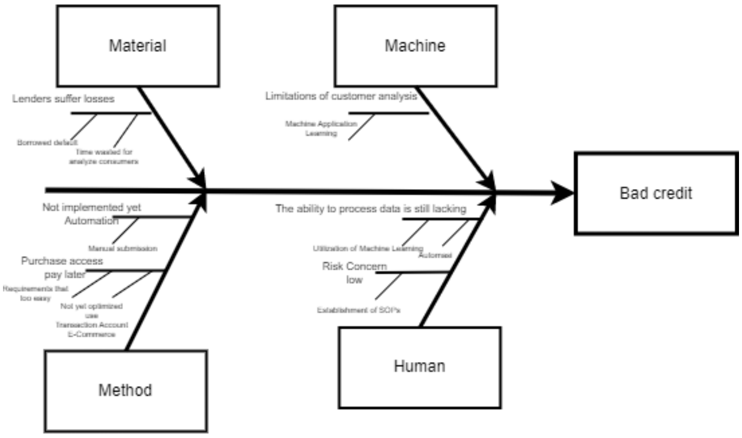

Figure 9. Cause and Effect Diagram

5W + 1H Analysis

Table 10. Limited Offer

5W+1H	Answer
What to do?	Providing *a limited Pay Later* offer by observing the pattern of transaction activities on E Commerce
Why apply a Limited Offer	To filter potential users to reduce consumer default on *Pay Later* installments
Where to apply restrictions on *Pay Later* offers	If the E Commerce account is effective, Machine Learning will always calculate potential customers with a range of 1 month
Who provides the limited offer	Analytics Team and Business Team
When to provide a limited offer	Auto Detect, at the time prior to the offer, once every 1 Month Machine Learning will test Consumer eligibility
How to provide Machine Learning to predict potential customers.	1. Provides Machine Learning Models for Potential Customer analysis 2. Providing offers to potential customers 3. Perform advanced analysis after a customer submits *Pay Later*

Table 11. Application of Machine Learning 5W+1H

5W+1H	Answer
What to do?	Using Machine Learning for analysis of potential pay *later* users
Why you should implement Machine Learning	To reduce the occurrence of consumer defaults by paying attention to consumer transaction patterns
Where to apply Machine Learning	Before giving a *Pay Later* offer
Who provides Machine Learning applications	Analytics Team
When to provide machine learning applications	Auto Detect, once every 1-month Machine Learning will test consumer feasibility
How to provide Machine Learning to predict potential customers.	1. Provide past Pay Later user data to model Training data 2. Testing algorithms with the highest level of accuracy 3. Build the model

Potential Customer Analysis using Machine Learning

After testing 5 machine learning algorithms, namely Decision Tree, Random Forest, KNN, Support Vector Machine, and XG Boosting, it can be seen in table 12 that the random forest algorithm has the highest level of accuracy compared to other algorithms. Looking at how the algorithm works, it can be seen that the inherent reason why Random Forest has the highest level of accuracy is that Random Forest works by creating more than 1 model or rule where other algorithms only create 1 model or rule. Random Forest in practice can create the desired number of models such as 100, 1000, and others [22], [39]–[41]. Of course, it is adjusted to the amount of data provided to build the model. Here is the level of accuracy of the algorithm tested in building the model.

Table 12. Comparison of Algorithm Accuracy Levels

Algorithm	Accuracy
Decision Tree	82,8 %
Random Forest	93,4 %
XG Boosting	89,12%
KNN	90%
Support Vector Machine	87%

Recommended Pay Later Process Flow

After successfully identifying the factors that influence the occurrence of default on Pay Later which is then compiled in an SOP and proposes the use of machine learning in predicting customers who have potential in the transaction rate to be given Pay Later access. To support the previous analysis, the author has compiled a framework that can be used to implement previously carried out analysis by paying attention to the efficiency in conducting the analysis as described in figure 10. Analysis is carried out with 2 parts, analysis with machine learning and analysis of confirmation of personal information by the analytics team. The key to activity is in the analysis that is performed automatically by machine learning on each E Commerce account every 1 time per Month on a recurring basis. This provides an opportunity for each user to be able to use the Pay Later feature when they have met the requirements in accordance with the SOP set by the company. The activity is carried out by conducting a follow-up analysis for registration containing personal information for verification. The framework can be seen in figure 10 below.

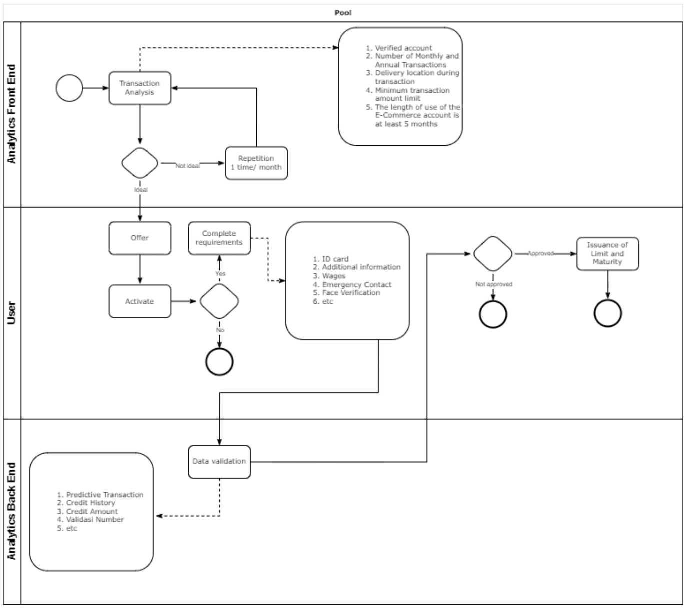

Figure 10. Machine Learning Implementation Framework after FMECA SOP preparation

4. CONCLUSION

Based on the discussion as described in the previous chapter, the author draws the conclusion that using the FMECA analysis there are 3 failure modes that are the priority in this study, namely Pay Later payment failure, customer disappears, and Late Payment of Pay Later installments. The main factor that causes this failure is that the SOP in determining users who are entitled to use Pay Later is still not efficient and effective. To dig deeper, the author conducted a cause and effect analysis of fish bone to find out the factors that caused the failure and made a 5W+1H analysis as a guideline in the preparation of SOPs. The basic SOP produced is to make a limited offer to users who are interested in using the Pay Later feature. In supporting the previous SOP, it is necessary to analyze consumer behavior patterns by analyzing trends carried out on the E Commerce platform before making an offer to 'find out whether the user is worthy. In analyzing these consumer behavior patterns, the author proposes the use of Machine Learning to speed up the process, efficiency, and repeated analysis every month to provide opportunities for each user. Finally, in order to clarify the concept proposed by the author, the author proposes a framework that can be used as a reference. By applying the FMECA method, it can analyze the risk of failure that may occur so that it can prepare changes in the SOPs and frameworks used. In order to develop the implementation of FMECA in the business line, the author uses machine learning by analyzing algorithms that are well used by paying attention to the highest level of accuracy. Evidently, the use of machine learning can help in increasing efficiency and effectiveness as previously designed SOPs using FMECA. This also shows the level of probability of a 15% decrease in consumers defaulting on PT. XYZ However, in future studies it is necessary to conduct correlation analysis and static testing in testing attribute correlations before conducting algorithm trials when building machine learning models. The authors also suggested comparing using other methods to improve risk management in this study.

REFERENCES

[1] A. Munsch, "Millennial and generation Z digital marketing communication and advertising effectiveness: A qualitative exploration," *Journal of Global Scholars of Marketing Science*, vol. 31, no. 1, pp. 10–29, Jan. 2021, doi: 10.1080/21639159.2020.1808812.

[2] N. Taufik, F. H. E. Prabowo, A. D. Santosa, and A. E. Mandasari, "Smes' E-Commerce Adoption Towards Consumer Experience," *Jurnal Bisnis dan Manajemen*, vol. 21, no. 2, pp. 144–155, 2020, doi: 10.24198/jbm.v21i2.373.

[3] "Digital-2022-Indonesia-February-2022-v01_compressed".

[4] N. Wang, "Research on the influence of the cross-border e-commerce development of small and medium-sized enterprises in Dongguan in the post-epidemic era," in *Proceedings - 2nd International Conference on E-Commerce and Internet Technology, ECIT 2021*, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 176–180. doi: 10.1109/ECIT52743.2021.00047.

[5] X. Qu, "Research on the Development of Cross-Border E-commerce in the Context of Digital Trade," 2020.

[6] S. AFTIKA, H. HANIF, and Y. DEVI, "PENGARUH PENGGUNAAN SISTEM PEMBAYARAN SHOPEEPAYLATER 'BAYAR NANTI' TERHADAP PERILAKU KONSUMTIF MAHASISWA UIN RADEN INTAN LAMPUNG DALAM PERSPEKTIF BISNIS SYARIAH," *REVENUE: Jurnal Manajemen Bisnis Islam*, vol. 3, no. 1, pp. 87–106, Jan. 2022, doi: 10.24042/revenue.v3i1.11228.

[7] N. Modi and J. Singh, "Understanding Online Consumer Behavior at E-commerce Portals Using Eye-Gaze Tracking," *Int J Hum Comput Interact*, 2022, doi: 10.1080/10447318.2022.2047318.

[8] N. Johnsirani Venkatesan, C. S. Nam, and D. R. Shin, "Deep Learning Frameworks on Apache Spark: A Review," *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, vol. 36, no. 2. Taylor and Francis Ltd., pp. 164–177, Mar. 04, 2019. doi: 10.1080/02564602.2018.1440975.

[9] L. T. T. Tran, "Managing the effectiveness of e-commerce platforms in a pandemic," *Journal of Retailing and Consumer Services*, vol. 58, Jan. 2021, doi: 10.1016/j.jretconser.2020.102287.

[10] C. Cheung, C. C. Fang, and Y. C. Wang, "Consumer behavior and mobile payment: An empirical study of the restaurant industry," *Journal of China Tourism Research*, vol. 18, no. 4, pp. 735–755, 2022, doi: 10.1080/19388160.2021.1972072.

[11] J. Iaksch, E. Fernandes, and M. Borsato, "Digitalization and Big data in smart farming–a review," *Journal of Management Analytics*, vol. 8, no. 2. Taylor and Francis Ltd., pp. 333–349, 2021. doi: 10.1080/23270012.2021.1897957.

[12] J. Quan, X. Wang, and Y. Quan, "Effects of Consumers' Strategic Behavior and Psychological Satisfaction on the Retailer's Pricing and Inventory Decisions," *IEEE Access*, vol. 7, pp. 178779–178787, 2019, doi: 10.1109/ACCESS.2019.2958685.

[13] D. M. Sumarta, W. Suweca, R. Setiawan, S. Tinggi, and T. Bandung, "PENERAPAN METODE FAILURE MODE, EFFECT AND CRITICALITY ANALYSIS (FMECA) PADA DRIVE STATION ALAT ANGKUT KONVEYOR REL," 2017.

[14] G. J. Jiang, S. Cai, N. Zhang, H. X. Chen, and H. H. Sun, "Reliability analysis of the starting and landing system of UAV by FMECA and FTA," *Journal of Industrial and Production Engineering*, vol. 36, no. 8, pp. 503–511, Nov. 2019, doi: 10.1080/21681015.2019.1695682.

[15] W. Hamro Afiva, F. Tatas, D. Atmaji, and D. J. Alhilman, "Penerapan Metode Reliability Centered Maintenance (RCM) pada Perencanaan Interval Preventive Maintenance dan Estimasi Biaya Pemeliharaan Menggunakan Analisis FMECA (Studi Kasus : PT. XYZ)," vol. XIII, no. 3, pp. 298–310, 2019.

[16] E. Tri Hartanto and U. Ciptomulyono, "APLIKASI FAILURE MODE EFFECT AND CRITICALITY ANALYSIS (FMECA) DALAM PENENTUAN INTERVAL WAKTU PENGGANTIAN KOMPONEN KRITIS RADAR JRC JMA 5310 PADA KRI SATUAN KAPAL PATROLI KOARMATIM Dosen Prodi Teknik Industri FTI-ITS 2) Dosen Prodi S-2 ASRO STTAL 3)."

[17] R. Ravid, E. Bashkansky, and S. Dror, "Integrating AHP in an FMECA framework for ranking Down syndrome tests," *IIE Trans Healthc Syst Eng*, vol. 1, no. 2, pp. 91–100, 2011, doi: 10.1080/19488300.2011.609521.

[18] G. Di Bona, A. Silvestri, A. Forcina, and A. Petrillo, “Total efficient risk priority number (TERPN): a new method for risk assessment,” *J Risk Res*, vol. 21, no. 11, pp. 1384–1408, Nov. 2018, doi: 10.1080/13669877.2017.1307260.

[19] M. Swamynathan, *Mastering Machine Learning with Python in Six Steps*. Apress, 2017. doi: 10.1007/978-1-4842-2866-1.

[20] R. Bai *et al.*, “Analytics and machine learning in vehicle routing research,” *Int J Prod Res*, vol. 61, no. 1, pp. 4–30, 2023, doi: 10.1080/00207543.2021.2013566.

[21] S. Hosseini, “A decision support system based on machined learned Bayesian network for predicting successful direct sales marketing,” *Journal of Management Analytics*, vol. 8, no. 2, pp. 295–315, 2021, doi: 10.1080/23270012.2021.1897956.

[22] K. Chi-Hsien and S. Nagasawa, “Applying machine learning to market analysis: Knowing your luxury consumer,” *Journal of Management Analytics*, vol. 6, no. 4, pp. 404–419, Oct. 2019, doi: 10.1080/23270012.2019.1692254.

[23] S. K. Sardar, N. Kumar, and S. C. Lee, “A Systematic Literature Review on Machine Learning Algorithms for Human Status Detection,” *IEEE Access*, vol. 10. Institute of Electrical and Electronics Engineers Inc., pp. 74366–74382, 2022. doi: 10.1109/ACCESS.2022.3190967.

[24] M. Swamynathan, *Mastering Machine Learning with Python in Six Steps*. Apress, 2017. doi: 10.1007/978-1-4842-2866-1.

[25] M. Karami, D. Crick, and J. M. Crick, “Non-predictive decision-making, market-oriented behaviours, and smaller-sized firms’ performance,” *Journal of Strategic Marketing*, 2022, doi: 10.1080/0965254X.2022.2052938.

[26] M. Swamynathan, *Mastering Machine Learning with Python in Six Steps*. Apress, 2017. doi: 10.1007/978-1-4842-2866-1.

[27] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[28] H. H. Zhao, X. C. Luo, R. Ma, and X. Lu, “An Extended Regularized K-Means Clustering Approach for High-Dimensional Customer Segmentation with Correlated Variables,” *IEEE Access*, vol. 9, pp. 48405–48412, 2021, doi: 10.1109/ACCESS.2021.3067499.

[29] K. Pradhan and P. Chawla, “Medical Internet of things using machine learning algorithms for lung cancer detection,” *Journal of Management Analytics*, vol. 7, no. 4. Taylor and Francis Ltd., pp. 591–623, Oct. 01, 2020. doi: 10.1080/23270012.2020.1811789.

[30] P. A. Morales, S. True, and R. K. Tudor, “Insights, challenges and recommendations for research on sustainability in marketing,” *Journal of Global Scholars of Marketing Science*, vol. 30, no. 4, pp. 394–406, Oct. 2020, doi: 10.1080/21639159.2020.1803757.

[31] A. Rojas-Dominguez, L. C. Padierna, J. M. Carpio Valadez, H. J. Puga-Soberanes, and H. J. Fraire, “Optimal Hyper-Parameter Tuning of SVM Classifiers with Application to Medical Diagnosis,” *IEEE Access*, vol. 6, pp. 7164–7176, Dec. 2017, doi: 10.1109/ACCESS.2017.2779794.

[32] S. F. Yap and S. S. Gaur, “Consumer Dissonance in the Context of Online Consumer Behavior: A Review and Research Agenda,” *Journal of Internet Commerce*, vol. 13, no. 2, pp. 116–137, 2014, doi: 10.1080/15332861.2014.934647.

[33] IEEE Control Systems Society, Dongbei da xue, Chinese Association of Automation. Technical Committee on Control and Decision of Cyber Physical Systems, IEEE Singapore Section. Industrial Electronics Chapter, Chongqing da xue, and Institute of Electrical and Electronics Engineers., *Proceedings of the 30th Chinese Control and Decision Conference (2018 CCDC) : 09-11 June 2018, Shenyang, China*.

[34] D. H. Kim, S. Lee, J. Jeon, and B. C. Song, “Real-time purchase behavior recognition system based on deep learning-based object detection and tracking for an unmanned product cabinet,” *Expert Syst Appl*, vol. 143, Apr. 2020, doi: 10.1016/j.eswa.2019.113063.

[35] R. Ravid, E. Bashkansky, and S. Dror, “Integrating AHP in an FMECA framework for ranking Down syndrome tests,” *IIE Trans Healthc Syst Eng*, vol. 1, no. 2, pp. 91–100, 2011, doi: 10.1080/19488300.2011.609521.

[36] S. Hasan, E. Sahiri, and E. Tri Hartanto, “ANALYSIS OF DETERMINING CRITICAL COMPONENTS USING FMECA METHOD IN SEAWATER PUMP OF DIESEL GENERATOR CATERPILLAR 3412.”

[37] M. Hofmann, R. K. Hofmann, and • Klinkenberg, “RapidMineR RapidMineR Data Mining Use Cases and Business Analytics Applications Edited by.” [Online]. Available: www.RapidMiner.com.

[38] S. Demirkan, I. Demirkan, and A. McKee, “Blockchain technology in the future of business cyber security and accounting,” *Journal of Management Analytics*, vol. 7, no. 2. Taylor and Francis Ltd., pp. 189–208, Apr. 02, 2020. doi: 10.1080/23270012.2020.1731721.

[39] G. Meena, D. Sharma, and M. Mahrishi, “Traffic Prediction for Intelligent Transportation System using Machine Learning,” in *Proceedings of 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things, ICETCE 2020*, Institute of Electrical and Electronics Engineers Inc., Feb. 2020, pp. 145–148. doi: 10.1109/ICETCE48199.2020.9091758.

[40] R. Rai, M. K. Tiwari, D. Ivanov, and A. Dolgui, “Machine learning in manufacturing and industry 4.0 applications,” *International Journal of Production Research*, vol. 59, no. 16. Taylor and Francis Ltd., pp. 4773–4778, 2021. doi: 10.1080/00207543.2021.1956675.

[41] K. Pradhan and P. Chawla, “Medical Internet of things using machine learning algorithms for lung cancer detection,” *Journal of Management Analytics*, vol. 7, no. 4. Taylor and Francis Ltd., pp. 591–623, Oct. 01, 2020. doi: 10.1080/23270012.2020.1811789.