Application of Principal Component Analysis and Maximum Likelihood Estimation Method to Identify the Determinant Factors Intention to Use of Paylater in E-Commerce

Faizatul Fajariah¹,², Hisarma Saragih³, Donny Dharmawan³, Loso Judijanto⁴, Musran Munizu⁵

¹Sekolah Tinggi Ilmu Ekonomi AMM
²Universitas Simalungun
³Universitas Krisnadwipayana
⁴IPOSS Jakarta Indonesia
⁵Universitas Hasanuddin

faizatulfajariah@gmail.com

Abstract
This research aims to analyze which factors are most determining in influencing interest in using pay-later. This type of research is multivariate analytical research using a quantitative approach. All pay-later service customers who engage in e-commerce make up the study's population. Purposive sampling and non-probability sampling were employed in the sample-taking process based on this study. There are one hundred samples. Researchers have made observations to gather data. The author intends to employ principal component analysis and maximum likelihood estimation techniques for data analysis. Following the completion of the study, the following findings were made. The principal component analysis method's RMSE value is 0.02. The maximum likelihood estimates method's root mean square error is 0.015. Therefore, it can be said that the principal component analysis approach is appropriate for usage because it can explain factor formation due to its strong factor loading value and low RMSE value. The influence of advantages is the most important element determining interest in pay-later utilizing the principal component analysis method since it has the highest loading factor value in decision-making.

Keywords: Pay-Later, Principal Component Analysis, Maximum Likelihood Estimation, Decision Making.

1. Introduction

The business world is experiencing fiercer rivalry due to the development of increasingly sophisticated technology. Every business provides benefits to customers in an attempt to become more competitive. The more an effective, user-friendly, and visually appealing sales system is created, the more money the business will make. This is demonstrated by the numerous online marketplaces, or “e-commerce,” that facilitate sales, purchases, and product acquisitions by allowing buyers and sellers to conduct business virtually. One of the e-commerce platforms that is currently expanding quickly in Indonesia is Shopee [1]. Over the past three years, Shopee, an online retailer, has seen a rise in visitors. With 600 million visits in 2021, it will have the highest volume of e-commerce traffic in Indonesia. With 530 million visitors, Tokopedia is in second place, while Bukalapak comes in third with 100 million visits. The digital market application, which offers a variety of sales products like food, electronics, different kinds of clothing needs, and so on, is the reason for the high number of visitors [2]. Furthermore, it offers a number of features that are designed to make transactions simpler for customers. A digital financial system, the most recent technology advancement in the non-cash financial sector, is another service provided by e-commerce. It is effective and gaining traction [3]. Fintech’s contemporary financial services are a great help to people's lives get easier [4]. The financial actions of persons are significantly impacted by this disorder. Consequently, Paylater a digital credit card payment feature was introduced by e-commerce. The most recent Paylater option is becoming more and more popular among users [5]. One of the many e-commerce transaction techniques called Paylater tries to make it simpler for users to purchase items through the application and can be paid for in the end within a predetermined grace period [6]. The peer-to-peer lending company regulates this system collectively. Some people find it helpful to use Paylater to make purchases ahead of time and pay for them at the end of the month [7]. PayLater's advantages have led to a sharp rise in the number of PayLater users in Indonesia. The paylater feature is most frequently used by respondents in e-commerce [8]. PayLater is not the most often used payment option, despite its increasing user base. After bank transfers, small market payments, and e-wallets, Paylater is the fourth most popular service [9]. The choice to utilize Paylater as a transaction mechanism remains with the user. You must fulfill the requirements in order to use...
it, which include providing proof of identity, paying interest, and incurring processing costs and fines if payments are made after the deadline [10]. This is a consideration for the public in making Paylater a safe transaction tool. The occurrence of technical problems in the Paylater system, which cause losses for consumers. The inaccessible Paylater feature means consumers cannot pay their bills until the due date [11]. This causes consumer Paylater to be suspended [12]. Apart from that, the requirements for activating Paylater by entering an identity card are still below the public's trust limit because they are prone to hacking attacks. This raises doubts about using Paylater, and consumers are still considering it before deciding to use it [13].

There are two ways to eliminate variables in the analytical procedure utilized in this study. The first approach is principle component analysis (PCA), which is the most straightforward way to extract components. It extracts the fewest possible factors while retaining the majority of the data included in all of the variables [14]. The second method is maximum likelihood estimation because MLE is an unbiased, efficient, and consistent factor extraction that can produce estimates that are most likely to obtain a good correlation matrix [15]. As is known, the maximum likelihood estimation method is the most appropriate method to use because it has the lowest error rate. Therefore, it is necessary to carry out analysis using similar methods in order to compare which method is most appropriate to use in this research [16]. Finding the variables that can account for the connection or link between the different independent indicators under study is done through the use of factor analysis. When examining factors or causes, it is common practice to combine multiple related variables into a single factor and condense multiple factors that would affect a single component variable into multiple smaller but significant factors [17]. Multivariate analysis refers to all statistical methods that jointly (simultaneously) evaluate more than two variables in an individual or object. One method of achieving this is via analyzing the presence of causative factors. It is common to refer to this technique as multivariable analysis [18]. In order to establish one or more groups of variables whose number is fewer than the initial variable, factor analysis, in theory, searches for correlation (inter-relationship) between a number of initially independent or independent variables [19]. Factor analysis can create a number of new variables while keeping the factors that have been found in the original variables [20].

Electronic commerce, also known as e-commerce, is the exchange of goods and services between buyers and sellers. It can take place in person or online [21]. It involves a thorough search for services, focusing information based on requirements and specifications, and comparing prices, types, and locations to find the desired product [22]. Automation, or converting manual business processes into automatic ones (often referred to as enterprise resource planning), is one of the five core ideas of e-commerce [23]. Integration, also known as streamlining, is an integrated process that produces effective and efficient outcomes (the just-in-time idea) [24]. Publication facilitates the marketing of commercialized goods and services (the idea of electronic cataloging) [25]. By reducing human error, interaction involves the sharing of knowledge or data amongst businesspeople (the electronic data interchange concept). An agreement between two commercial parties to conduct business with the involvement of a third party as the paying function is known as a transaction (electronic payment concept) [26]. Principal component analysis, sometimes known as PCA, is one of the most widely used techniques for extracting data [27]. By preserving as much information from the original variables as feasible, it aims to minimize variables; the accuracy increases with the number of variables [28].

One statistical test that is a part of the factor analysis group is called PCA. PCA is a mathematical approach designed to employ a minimal number of parameters to describe the variation present in a data collection (i.e., the answers used to characterize a sample) [29]. Shared variance, sometimes referred to as communality, rises when one variable is related to another. While general analysis approaches solely address general variances, component analysis processes are related to all variances [30]. Component analysis will be carried out if the primary objective of factor analysis is data reduction and the specified degree of variation and error variance are minor. The desired probability distribution is the one that makes the observed data "most likely," according to the principle of maximum likelihood estimation (MLE), which was first developed by R.A. Fisher in the 1920s. This means that the value of the parameter vector that maximizes the likelihood function should be found [31]. The MLE estimate is the resultant parameter vector that is found by exploring the multidimensional parameter space. This is the population that is most likely to generate the observed data, according to the MLE principle. In conclusion, maximum likelihood estimation is a technique for determining the probability distribution that maximizes the likelihood of the observed data.

2. Research Methods

This kind of study employs a quantitative technique in multivariate analysis. According to the justification, this study is comparative in nature that is, it compares and contrasts so it can be classified in two different ways to compare or search for differences between the components. Researchers used a population all e-commerce customers who use pay-later services to conduct the experiment; however, the precise population size is uncertain. Purposive sampling and non-probability sampling were employed in the sample-taking process based on this study. There are one hundred samples. Researchers have examined previous literature studies and conducted observations to gather data. In order to conduct observations, data were gathered based on firsthand observations of the research.
object. Regarding the application's pay-later option, observations were made. This observation was made by directly witnessing how the pay-later feature works, starting with observing the application's pay-later menu, how to activate pay-later, how to calculate pay-later interest, and how to pay bills using pay-later options. Studying literature requires an understanding of theory and the application of factor analysis to connect the theories discovered. To support this research, literature reviews can gather data from books, e-books, journals, and relevant past studies. The author intends to employ principal component analysis and maximum likelihood estimation techniques for data analysis. When determining variables into several causes that have less value than the preceding variable, this method is employed as a guide. Use the a priori method, eigenvalue, scree plot, percentage of variance, factor rotation, factor naming, and model accuracy testing to determine the number of components. Then, in order to generate research that is beneficial for connected parties, the researchers changed a number of factors from the Technology Acceptance Model (TAM), the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2), and multiple linear regression techniques. Thus, 11 factors including the impact of promotion, perceived utility, perceived ease of use, perceived risk, perceived trust, facilitating conditions, perceived technology security, social influence, buying impulse, price value, and hedonic motivation are used in this study to examine how people choose between pay-later payment options.

3. Results and Discussion

Communities are initially the accumulation of numerous variances from a variable that the factors created can explain. The promotion influence variable is able to explain factors of 0.9, perception of benefits of 0.92, perception of risk of 0.64, perception of trust of 0.75, perception of technology security of 0.4, impulse purchase of 0.6, and price value of 0.7. As a result, it is evident that six variables have values greater than 0.5 and one variable, the technological security perception variable, has a value less than 0.5. Six variables have been chosen in this instance and have a respectable association with the elements that have been established. Utilizing the total variance explained test is the next step. In order to create a factor, the total variance explained gives a score for every variable that has been examined. The variation can be explained by two types of analysis: full squared subtraction of loadings and starting eigenvalues. The resulting factors are displayed in the variance of the first eigenvalue. The eigenvalue must be larger than 1000 in order to identify the ensuing multiplier; if it is less than 1000, no multiplier will occur. This speaks about three possible causes derived from the seven factors examined. When a factor finding mechanism is included, the absolute eigenvalue score is greater than 1. With a component eigenvalue of 2.7, factor 1 is formed, accounting for 38% of the variation. In the meantime, 1.2, a factor of 2, is the component with two eigenvalue components and can account for 17% of the fluctuation. Alternatively, 1.02, the component with three eigenvalue components, is a factor of three or can account for a variation of 14.5%. The extraction sum of squared loadings score indicates that 69% of the variation in the variable can be explained by the combined value of the three factors that were produced. Since the eigenvalue component score of components 4, 5, 6, and 7 is less than 1, they do not contribute to the total value and are not considered factors. The term "loading factor" refers to the degree of correlation that exists between a single variable and factors 1, 2, and 3. From the results that have been found, the relationship between the variables promotion influence (0.94), perception of benefits (0.95), and perception of risk (0.8) is very strong on factors 1. For the correlation of the variable's perception of technology security (0.5), impulse purchase (0.6), and price value (0.7), it is very strong towards factor 2, and for the correlation of the variable perception of trust (0.8), it is very strong towards factor 3. For this reason, it is necessary to have a series of Suaya rotation clarifies each of the variables that will be classified into factors 1, 2, and 3. A component transformation matrix is a component transformation that shows whether the factors formed are no longer related or orthogonal. The linkage values of 0.99, 0.900, and 0.900 on the major diagonal are more than 0.5. This indicates that there is a strong association between each of the three elements (components) and the desired rotation factor since the three factors need to be validated. Based on the rotation of factors that have been processed cumulatively, the following factors 1, 2, and 3 classifications will later affect pay-laters interest in digital markets: Factor 1 (customer experience and marketing strategy factors) is composed of three variables that are referred to as such: risk perception (0.8), promotion influence (0.950), and perception benefits (0.96). The perceived benefits and risks are related to the customer experience after using a product, where using pay-later not only has a positive impact but can have a negative impact that the user must accept. Promotion is closely related to marketing strategy, whereby promoting the advantages of pay-later it can influence interest in use.

Three variables make up component 2, sometimes known as the product quality factor: price value (0.8), perceived technical safety (0.55), and impulsive purchasing (0.4). Because of the relationship between a product's perceived level of technological safety, price, and impulsive buy, it is known as the "product quality factor." If a thing is well-made, it will make someone feel protected. The price value and product quality are tightly correlated, and a good product quality might pique a user's impulsive interest in making a purchase. The single variable that makes up Factor 3 (the "trust factor") is perceived trust (0.8). The variable, perceived trust, is the source of the name of the trust factor. Factors of 0.9, perceived benefits of 0.98, perceived danger of 0.5, perceived trust of 0.2, perceived
technological security of 0.2, impulsive purchases of 0.11, and a price value of 0.2 may all be explained by the promotional influence variable. Hence, it is evident that four variables have values less than 0.5 and three variables have values greater than 0.5. This indicates a correlation between the development of the components and the three variables that are involved.

The loading factor displays the strength of the connection that exists between a variable and each of the first three factors. According to the findings, there is a very good correlation between the variables that influence promotion (0.984), perception of advantages (0.99), and perception of risk (0.99) with factor 1. The correlation between the variables' price value (0.3), impulsive buying (0.3), and perception of technological security (0.3) is in factor 2, but it is weak because it is less than 0.5. Similarly, the correlation between the variables' perception of trust (0.25) and impulsive buying (0.3) is in factor 3, with a weak loading value less than 0.5. As a result, a rotation process is required to make clear the distinctions between the variables that will be a part of factors 1, 2, and 3. The factors that influence pay-later interest in e-commerce have been named 1, 2, and 3, based on the results of the factor rotation that was conducted. Factor 1 (customer experience and marketing strategy) is made up of three variables: the influence of promotions (0.98), perceived benefits (0.9), and perceived risk (0.55). These variables are referred to as the customer experience and marketing strategy factors. The perceived benefits and risks are related to the customer experience after using a product, where using pay-later not only has a positive impact but can have a negative impact that the user must accept. Promotion is closely related to marketing strategy, whereby promoting the advantages of pay-later it can influence interest in its use.

The only variable that makes up component 2 (trust) is perceived trust (0.43), also known as the trust factor. The variable, perceived trust, is the source of the name of the trust factor. The pricing value (0.5), perceived technological safety (0.34), and impulsive purchase (0.3) are the three variables that make up factor 3, or product quality. Because of the relationship between a product's perceived level of technological safety, price, and impulsive buy, it is known as the "product quality factor." If a thing is well-made, it will make someone feel protected. The price value and product quality are tightly correlated, and a good product quality might pique a user's impulsive interest in making a purchase. The factors 2 and 3 are arranged differently when factor rotation findings from principal component analysis and maximum likelihood estimation approaches are compared; factor formation remains unchanged. The variables perceived advantages, perceived risk, and promotional effect are found in factor 2 of the principal component analysis approach, while the perceived trust variable is found in factor 3. In the meantime, the pricing value, perception of technical security, and impulsive purchase variables are found in factor 3 of the maximum likelihood estimation approach, while the perception of trust variable is found in factor 2. In actuality, there are some similarities between the outcomes of investigating causes through maximum likelihood estimate and principal component analysis. Changes in communality values, extraction sums of squared loadings, reproduced correlation values, and the factors formed can all happen based on specific scores. The calculations show that the principal component analysis method yielded RMSE findings of 0.02. The calculations show that the maximum likelihood estimation approach yielded RMSE results of 0.5. It is possible to determine that the maximum likelihood estimation RMSE value of 0.5 is larger than the principal component analysis RMSE value of 0.02 by calculating the RMSE values of the two methods: principal component analysis and maximum likelihood estimation. The more appropriate the method is to employ, in theory, the smaller the RMSE value. This indicates that the main component analysis method is the one that should be used more appropriately.

4. Conclusion

After completing the research by the author, we can underline the following matters: Calculating the root mean squared error (RMSE), or the RMSE value of the main component analysis approach, yields results on the comparative level of accuracy of the analysis model of elements that influence interest in using pay-later using two methodologies. The approach is more suited to employ the smaller the RMSE value. As a result, it can be said that the principal component analysis approach is appropriate for usage because it can explain factor formation and has a low RMSE value and a strong factor loading value (around 1). The findings of the main component analysis method's examination of three elements that are assumed to affect interest in pay-later are as follows: Factor 1 is made up of three elements that are referred to as the customer experience and marketing strategy factors: perceived benefits, promotional influence, and risk perception. Factor 2 (also known as the product quality factor) is composed of three variables: pricing value, the sense of technology safety, and impulsive purchase. The trust factor, or sense of trust, is the only variable that makes up factor 3. Using the main component analysis approach, the influence of advantages is the most significant element influencing interest in pay-later since it has the highest loading factor value in decision-making. The following recommendations for more research can be made in light of the preceding conclusions: In carrying out factor extraction, this method is very often used by researchers, so researchers suggest trying to use other factor extraction methods such as principal axis factoring and unweighted least squares so that they can see comparisons between various analysis methods. It is recommended for future researchers to examine other features of e-commerce applications, because the greater the increase in e-commerce service features, the greater the influence on interest in using the application. This study does not cover all
applications of e-commerce websites because it only looks at mobile applications. For this reason, it is advised that future studies examine every e-commerce website in order to provide thorough data. Subsequently, this study demonstrates that hedonic motivation, perceived convenience, social impact, and supporting facilities are non-formal variables and so require a review.

References


