

The Impact Of Iklanku Shopee Feature Through Brand Awareness On Purchase Decisions Using SEMPLS, Naive Bayes and K-NN Method

Ni Putu Intan Waindika Dharma^{a1}, I Ketut

Gede Darma Putra^{a2}, Ni Made Ika Marini Mandenni^{a3}, Yuzliza Binti Mohd Yasin^{b4}

^aDepartment of Information Technology, Faculty of Engineering, Udayana University, Indonesia
Jl.Raya Kampus Unud Bukit Jimbaran, Bali, Indonesia-803611

^bDepartment of Business, Faculty of Economics, Finance and Business, Utara Malaysia
University, Malaysia

Sintok, 06010 Bukit Kayu Hitam, Kedah, Malaysia

e-mail: ¹iwaindika@gmail.com, ²ikgdarmaputra@unud.ac.id, ³made_ikamarini@unud.ac.id,
⁴yuslizayasin06@gmail.com

Abstrak

Tujuan dari penelitian ini adalah untuk menguji tiga variabel yaitu pengaruh Fitur Iklanku Shopee terhadap Keputusan Pembelian dan Brand Awareness. Penelitian ini dilatarbelakangi karena adanya fenomena yang terjadi pada kondisi perekonomian yang sedang melemah akibat covid-19, namun pembelian produk fashion justru meningkat. Data diperoleh dari penyebaran kuesioner adalah sebanyak 350 data. SmartPLS digunakan untuk menemukan outer model dan inner model. Hasil data tersebut diolah kembali untuk menentukan cluster dan klasifikasi dengan menggunakan K-Means, Naive Bayes dan K-NN pada aplikasi RapidMiner. Hasil dari pengujian menggunakan SmartPLS bahwa hipotesis yang dibangun berpengaruh positif. Hasil dari pengolahan data mining adalah algoritma Naive Bayes lebih akurat dalam klasifikasi menggunakan split data dan juga cross validation.

Kata Kunci: Brand Awareness, Fitur Iklanku Shopee, K-NN, Keputusan Pembelian, Naive Bayes.

Abstract

The purpose of this study was to examine the effect of my Shopee Advertising Feature on Purchase Decisions and Brand Awareness. This research is motivated by the phenomenon of inequality that occurs in economic conditions that are weakening due to COVID-19, but the purchase of fashion products has actually increased. Data obtained from the distribution of questionnaires is as much as 350 data. SmartPLS is used to find the outer model and inner model. The results of the data are reprocessed to determine clusters and classification using K-Means, Naive Bayes and K-NN in the RapidMiner. The results obtained from SmartPLS are that every hypothesis that is built is positive. The results obtained from data mining processing are that the Naive Bayes algorithm is more accurate in classification using split data and cross validation.

Keywords: Brand Awareness, Iklanku Shopee Feature, K-NN, Naive Bayes, Purchase Decision.

1. Introduction

Information technology is increasingly sophisticated has provided various conveniences for humans, this causes changes in the mindset and habits of modern society in carrying out activities of daily life [1]. The total population in Bali Province is 4.36 million people and internet users in Bali in 2020 ranks 16th with a total of 3,411,084 internet users, it can be concluded that almost 80% of Bali's population uses the Internet in daily life. Internet use also affects people's shopping habits, especially those who have internet access done online via a smartphone [2]. In the second quarter of 2020 where the Covid-19 pandemic has hit the whole countries including

Indonesia which causes many people to lose their jobs due to social restrictions that have a hard impact on the community's economy. Some e-commerce also experienced an increase in sales due to this pandemic, one of which was Shopee. Shopee experienced an increase in sales in the second quarter of 2020 and became the most popular e-commerce.

The purchase of fashion products has actually increased, the causative factor is important to study. This study also analyzes the data that has been obtained from the results of the questionnaire by using the data mining method on the RapidMiner application that this study does not only display statistics obtained from the results of the questionnaire data. Data mining is very important to do because it can find out the added value of data mining a collection of data in the form of knowledge that has not been known manually.

Advertising is an important part of marketing activities seen by consumers [3]. Advertisements appear and are seen by the public, marketing is often interpreted as advertising. Television, radio, newspapers or internet media advertising becomes a menu that appears every day and in large numbers. Almost every second and minute advertisements color these media, and are captured by the audience. The Shopee My Advertisement feature is also one of the online marketing strategies carried out by Shopee in marketing products, besides that advertisements also provide several attractive offers and price discounts that cause users to become interested in making purchases on the Shopee application. Research that brand awareness can affect advertising on a product, indicates that brand awareness can be a variable that mediates the effect of advertising on product purchasing decisions [4].

The purpose of this study is to test the effect Iklanku Shopee feature on purchasing decisions and brand awareness of fashion products on the Shopee application. Knowing the results of clustering on this research data that it can find out the best cluster and the results of the classification comparison of the Naive Bayes algorithm and the K-NN algorithm. Limitation of the problem in this study is the measurement of product purchase decisions in the Shopee application is limited to the purchase of fashion products, Shopee users in question are Shopee users in the Denpasar area, given the limitations of researchers and Denpasar area where the majority of the population has often used digital technology in their daily life. Data clustering using the K-Means algorithm method and classification using Naive Bayes and K-NN algorithms. K-Nearest Neighbor (K-NN) is a classification algorithm based on the similarity of its nearest neighbors, and is usually used for classification problems. The advantage of the K-NN method is that it is easy to use, effective and efficient, it can provide good results. The classification of similarity results is obtained based on the closest distance between the data sample and the object [5]. The advantages of the KNN method are quite simple, popular, effective, and efficient. The KNN method is applied frequently and gives good results. Similar objects are classified in the same category [6].

2. Research Method

2.1. Purchase Decision

Purchasing decision process consists of five stages, namely the need recognition stage, information search, alternative evaluation, purchase decision, and post-purchase behavior. decision purchase has four indicators that is : Stability on product, that is confidence consumer for use a product, Habit, that is repetition buy a product continuously, Recommend that is give influence on others to buy a product, Purchase repeat that is do purchase repeat ater once buy a product [7].

2.2. Shopee Iklanku Features

Advertising is part of promotion, which is one of the most common media for companies to communicate persuasively to target buyers and the wider community [8]. The measurement of this variable uses research by [9] which is based on four indicators. The indicators used in this Iklan ku feature are: Finding information about a product or company from various media is easy, *Design* interesting media used, The information conveyed in various media is clear, Messages contained in various media can be trusted.

2.3. Brand Awareness

Brand Awareness is the knowledge that consumers have about a brand so that it makes it easier for companies to understand to make consumers to remember the brand. The pyramid of brand awareness as well as indicators in measuring brand awareness seen from the

lowest level to the highest level are: *Unaware of brand* (not aware of the brand), Brand Recognition (Brand recognition), Brand Recall (Improvement of the brand) and Top of Mind [2].

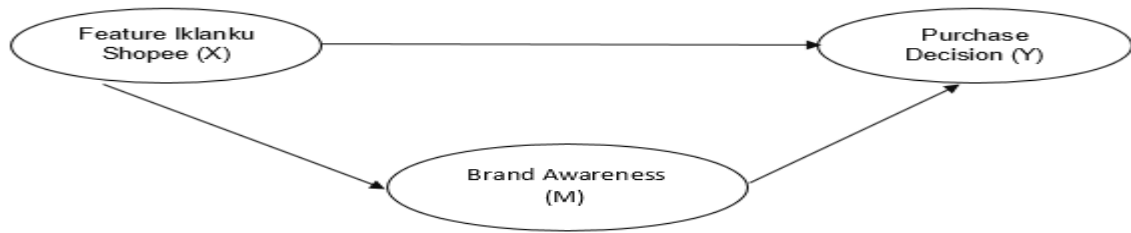


Figure 1. Research Conceptual Framework.

Figure 1 shows the conceptual framework. The mediating variable or the intervening variable is variables that theoretically affect (weaken and strengthen) the relationship between independent variable with dependent [10].

2.4. Location and Object of Research

This research is aimed at Shopee application users in Denpasar. Research variable the independent variable in This research is the feature of *Iklanku* Shopee (X). The dependent variable of this research is the purchase decision (Y). The mediating variable in this study is brand awareness (M). This study determines the sample by using the sample determination table from Isaac and Michael. Determining the number of samples based on the Isaac and Michael formula provides convenience determining the number of samples based on the error rate of 1%, 5% and 10% [11]. The error rate of this research or sampling error in determining the sample in this study is to use an error rate of 5%. The number of internet users in Bali 2020 is 3,411,084. The data collection technique used a questionnaire with Issac and Michael's calculations, namely $N = 3,411,084$ at an error rate of 5%, the sample studied was 349 rounded up to 350 data for better test results.

2.5. Data Collection Techniques

Data collection techniques in this study used a questionnaire technique. Based on this opinion, the questionnaires were distributed directly with related parties. The questionnaire contains a Likert scale that is used as a measuring tool, five scale format options are provided as follows: Strongly Disagree (STS), Disagree (TS), Indecisive (R), Agree (S), Strongly Agree (SS).

2.6. Data Analysis Technique

Data analysis technique is the resulting correlation coefficient is then compared with the applicable validation standards If $r > 0.30$, then the instrument item is declared valid and If $r < 0.30$, then the instrument item is declared invalid. The reliability test is a tool for measuring a questionnaire which is an indicator of the variable. Alpha Cronbach was used to test Pearson's product moment correlation and reliability measurements. The value must be > 0.6 .

2.7. Data Analysis with SEMPLS (Outer Model and Inner Model)

This SEMPLS model evaluates the Outer Model (evaluation of the measurement model) and the Inner Model (evaluation of the structural model). *Rule of thumb Test Validity* Outer Model *Convergent* $> 0,70$, *AVE* $> 0,5$, *Validity Discriminant* *Cross Loading* $> 0,70$, *Reliability* *Crobach Alpha* $> 0,70$ and *composite reliability* $> 0,70$.

Rule of thumb Inner Model is *R Square* $0,67:0,33$ and $0,19$. Changes in the value of *R-squares* are used to explain the effect of the exogenous latent variable on the endogenous latent variable, whether it has a substantive effect. The value of *R square* is used to find the value of *Q square*. The value of *Q square* must be > 0 [12].

Outer model is used to find the value of the validity and reliability of the data. Inner model to predict causality between latent variables by using bootstrapping, T-statistic test parameters obtained to determine there is a causal relationship. Evaluation of the structural model is evaluated through the R2 test, if the result of R2 is greater than 0.2, it can be said that the latent predictor has a major influence on the structural level. The R-square on the PLS model can be evaluated with the Q-square which can measure how well the observed values produced by the model and parameter estimates are [12].

2.8. Clustering and Classification Data

Data mining is used to solve a problem by using an algorithm [13]. Clustering and classification data process uses the K-Means, Naive Bayes and K-NN algorithm data mining methods in the RapidMiner application. Select the data used, namely 350 questionnaire data with the number of attributes 12 and the number of parameters 5. Perform data clustering using the K-Means algorithm in the RapidMiner application. The data is clustered into 5 groups. Determine the value of Davies Bouldin. The purpose of Davies Bouldin's measurement is to maximize the distance between clusters from one another. The cluster with the best number of clusters is the cluster that has the minimum Davies Bouldin value [14].

Classification of data from data cluster 2 to data cluster 5 by using Naive Bayes algorithm to determine the accuracy by using cross validation with a number of k-fold 10 and also to know each prediction from the classification results in each data cluster. Classification of data from data cluster 2 up to cluster 5 data by using K-NN algorithm to determine the accuracy by using cross validation with the number of k-fold 10 and the number of "k" neighbors 10. Classification of data from data cluster 2 to data cluster 5 using the Naive Bayes algorithm to find out the accuracy by using split data with the amount of training data is 0,8 and the testing data is 0.2. Classification data from data cluster 2 up to cluster 5 data by using K-NN algorithm to determine the accuracy by using split data with the amount of training data is 0,8 and the testing data is 0.2 and the number the neighbor's "k" is 10. Note result obtained then compare for get results accuracy which most height and results predictions obtained.

3. Result and Discussion

3.1. Characteristics of Respondents

The characteristics of the respondents obtained from the SPSS test results are characteristics of respondents based on Age is 21 – 30 years old 52,9%. Characteristics of respondents by Gender is woman 72,3%. Characteristics of respondents based on Last Education is Senior high school 64%. Results characteristics of respondents based on Occupation is Student 224%.

3.2. Validity and Reliability Test

Coefficient correlation from feature *Iklanku* Shopee (variabel X) is valid, Coefficient correlation Brand Awareness (variabel M) is valid and Coefficient correlation Purchase Decision (variabel Y) is valid. The value of an instrument is said to be reliable if the Cronbach Alpha value is > 0.6. Reliability test is carried out to determine the consistency of the items in the statement that are truly trustworthy. Reliable Test Results on Questionnaire Instruments of feature *Iklanku* 0,928, brand awareness 0,907 and decision purchase is 0,919 Reliable.

3.3. Outer Model

Validity test that is test loading factor (correlation among score item/score component with score construct) indicators which measure the construct. Rules of thumb which used for validity convergent is outer loading >0,7. Result validity convergent about variabel brand awareness (M), feature *Iklanku* and decision purchase is > 0,70. Discriminant validity relates to the principle that different construct quantifiers should not highly correlated and also must have value >0,7.

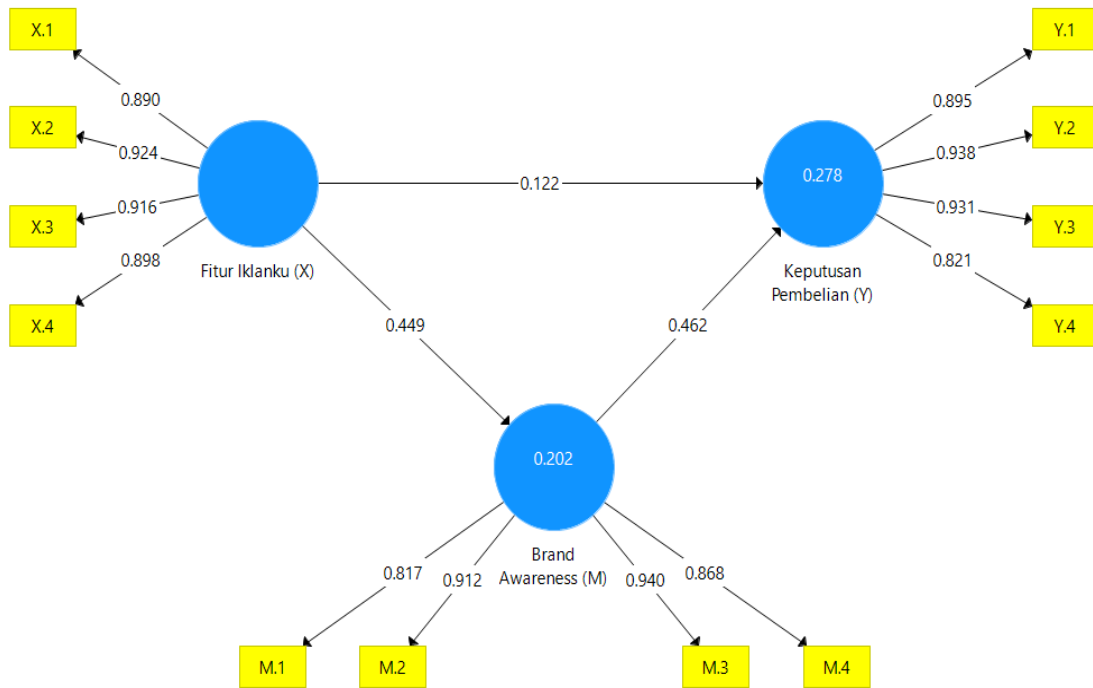


Figure 2. Outer Model

Table 1. Convergent Validity Test Results

	Brand Awareness (M)	Fitur Iklanku (X)	Keputusan Pembelian (Y)
M.1	0,817		
M.2	0,912		
M.3	0,940		
M.4	0,868		
X.1		0,890	
X.2		0,924	
X.3		0,916	
X.4		0,898	
Y.1			0,895
Y.2			0,938
Y.3			0,931
Y.4			0,821

Discriminant Validity Test Results is variabel brand awareness 0,886, feature *Iklanku* 0,907 and decision purchase 0,898. Cronbach Alpha Questionnaire Reliability Test Results is brand awareness 0,908, feature *Iklanku* 0,928 and decision purchase 0,919.

Table 2. Discriminant Validity Test Results and Cronbach Alpha

	Brand Awareness (M)	Fitur Iklanku (X)	Keputusan Pembelian (Y)	Cronbach's Alpha
Brand Awareness (M)	0,886			0,908
Fitur Iklanku (X)	0,449	0,907		0,928
Keputusan Pembelian (Y)	0,516	0,329	0,898	0,919

Composite Reliability Questionnaire Reliability Test Results is brand awareness (M) 0,936, Iklan ku feature (X) 0,949 and purchase decision (Y) 0,943.

Table 3. Composite Reliability

	Composite Reliability
Brand Awareness (M)	0,936
Fitur Iklanku (X)	0,949
Keputusan Pembelian (Y)	0,943

3.4. Inner Model

The inner model is carried out when the outer model has been tested well. To find out the influence in the relationship of the latent variables, the data were tested using the inner model (structural model).

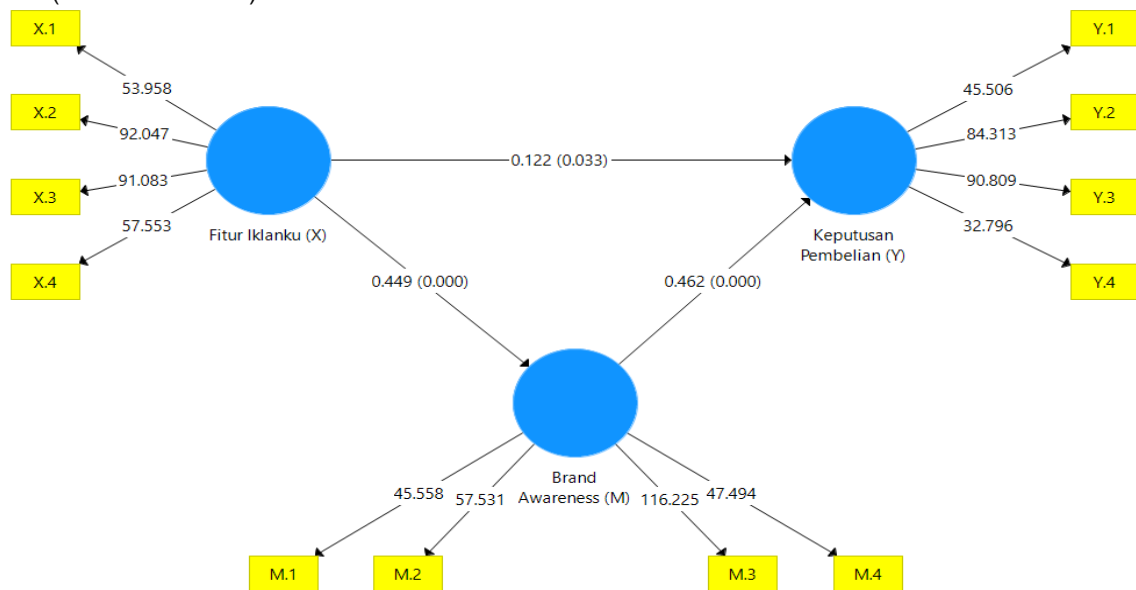


Figure 3. Inner Model

Table 4. R-Square Test Results

	R Square	R Square Adjusted
Brand Awareness (M)	0,202	0,199
Keputusan Pembelian (Y)	0,278	0,274

Inner Model test results R-Square 0.202 and 0.278. The Q-square result of 0.424 and the next step is to find the value of the hypothesis.

3.5. Hypothesis Test

The first hypothesis testing is to test the effect of each variable and must have a value > 1.96. The results of hypothesis testing can be seen in the column in table 5.

Table 5. Results of Direct Hypothesis

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (O/STDEV)	P Values
Brand Awareness (M) -> Keputusan Pembelian (Y)	0,462	0,465	0,059	7,796	0,000
Fitur Iklanku (X) -> Brand Awareness (M)	0,449	0,455	0,053	8,530	0,000
Fitur Iklanku (X) -> Keputusan Pembelian (Y)	0,122	0,121	0,057	2,143	0,033

Results of direct hypothesis testing direct and indirect are as follows value of t-statistics > t-value (7,796 > 1.96) it can be concluded that brand awareness positive and significant effect on purchasing decisions, and the hypothesis is accepted. The effect of feature Iklanku Shopee on brand awareness has a result of 8,530. The effect of feature Iklanku Shopee on purchasing decisions the results of hypothesis testing are 2,143. All direct hypothesis testing can be accepted.

Table 6. Results of Indirect Hypothesis

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (O/STDEV)	P Values
Fitur Iklanku (X) -> Brand Awareness (M) -> Keputusan Pembelian (Y)	0,207	0,211	0,035	5,844	0,000

The result of testing the indirect hypothesis is that the role of the brand awareness variable as a mediation shows positive and significant results because it has a value of 5.844.

3.6. Result Clustering Process

Implementation K-Means on RapidMiner with processed data is 350 questionnaire data. Then tested with Davies Bouldin to find out the best cluster.

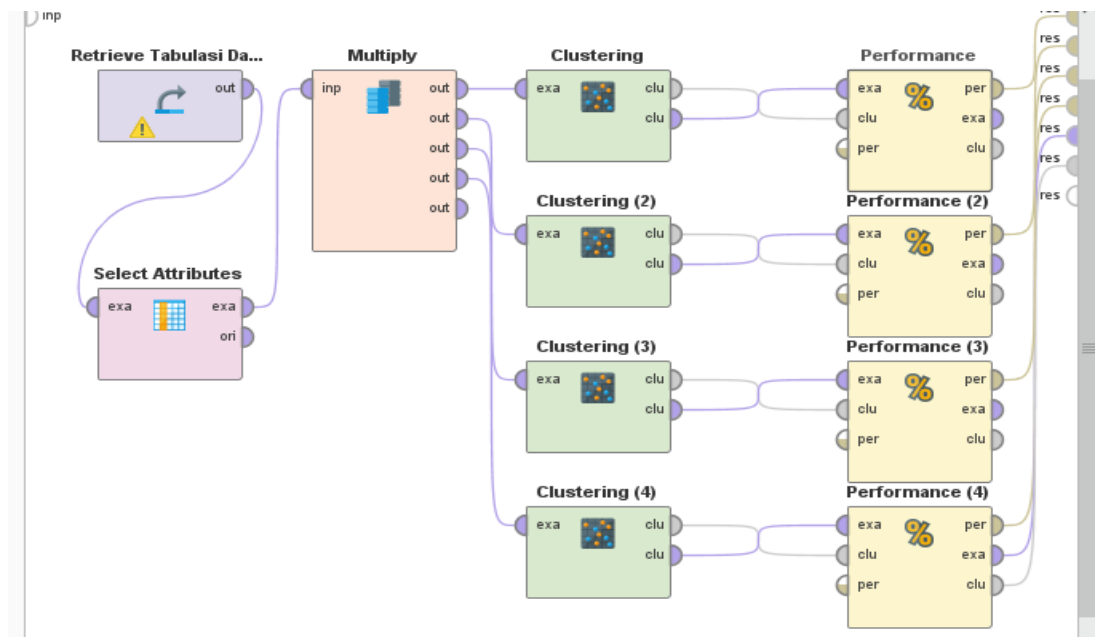


Figure 4. Clustering Process with K-Means Algorithm and Davies Bouldin Testing.

The results of the four tests with 5 clusters show that: The results of Cluster 2 evaluation using the Davies Bouldin method are 0.114, Result Evaluation cluster 3 use method Davies

Boludin Method is 0,109, Result Evaluation cluster 4 use method Davies Boludin Method is 0,109, Result Evaluation cluster 5 use method Davies Boludin Method is 0,108. Davies Bouldin is the method used to determine the best number of clusters.

3.7. Implementation of the algorithm with Naive Bayes and K-NN using Cross Validation

The advantage of using the algorithm method is that it only requires a small amount of training data and is good for long-term prediction. Naive Bayes method is very good for predicting. Naive Bayes process is to divide the problem into classes from the characteristics of similarities and differences by using statistics that can predict the probability of a class. The process of classifying data that has been clustered is as follows using RapidMiner.

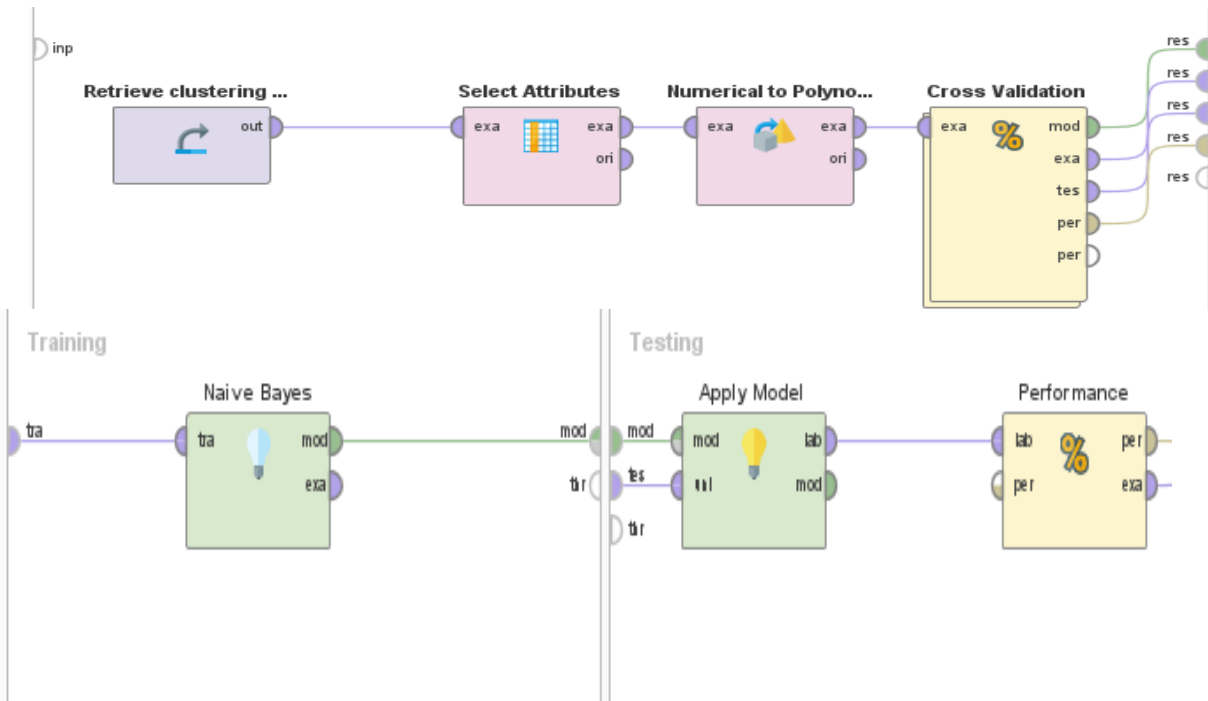
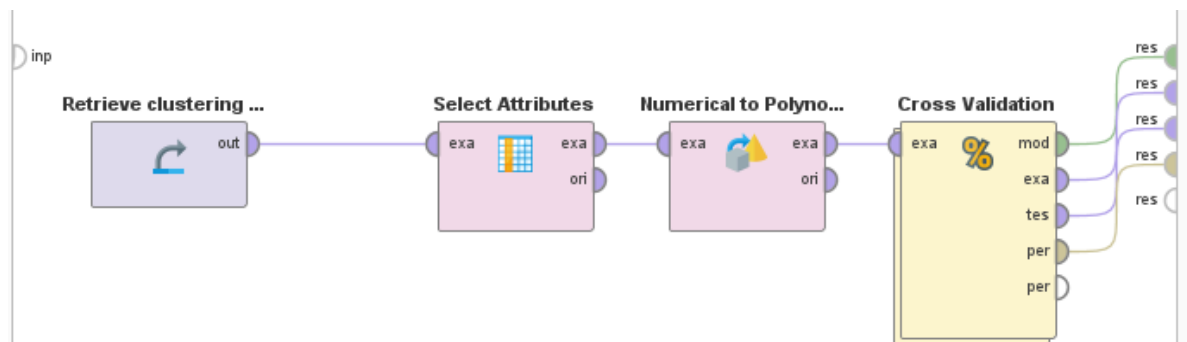


Figure 5. Data Classification Process with Naive Bayes Algorithm (Cross Validation)

The KNN algorithm has the advantage that it is simple, fast and easy to learn and requires a lot of training data. Data processing using the Naive Bayes algorithm and K-NN are almost the same. The K-NN method requires more training data.



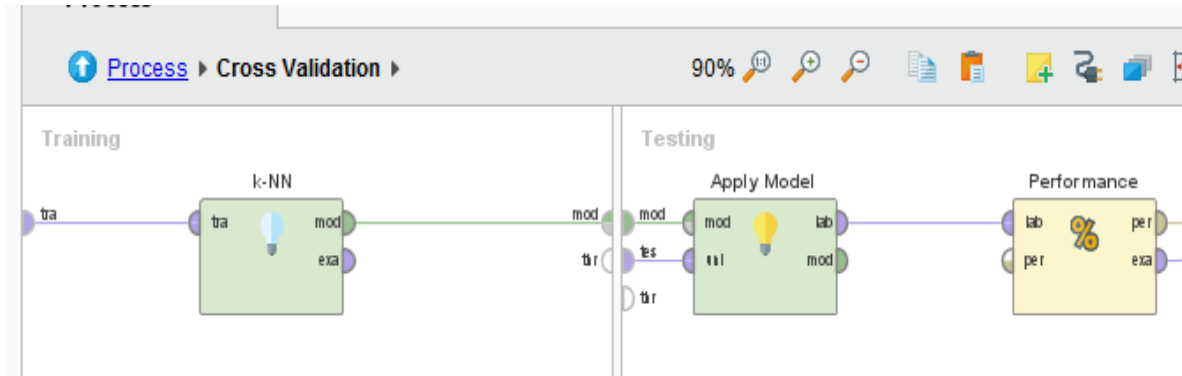


Figure 6. Data Classification Process with and K-NN Algorithm (Cross Validation).

Accuracy Results of Data Classification with Naive Bayes using cross validation. Accuracy classification of naive bayes using cross validation is data clustering 2 is 98,00%, data clustering 3 is 97,43%, data clustering 4 is 96,57%, data clustering 5 is 95,71%. Accuracy results of data classification with K- NN using cross validation is data clustering 2 is 96,57%, data clustering 3 is 92,29%, data clustering 4 is 90,57%, data clustering 5 is 89,43%.

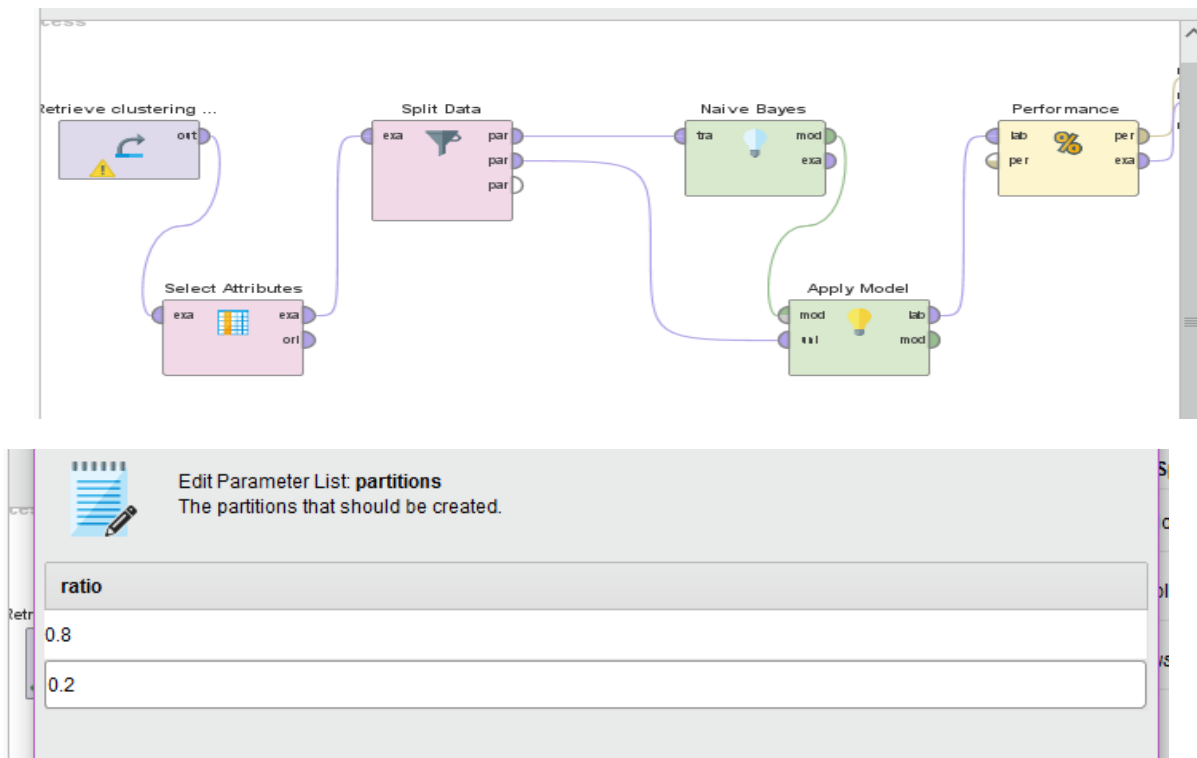


Figure 7. Data Classification Proses with Naive Bayes Algorithm and K-NN Using Split Data

Figure 7 shows the test of 350 questionnaire data that has been clustered with the K-Means algorithm and then classified using the Naive Bayes and K-NN algorithm so that the accuracy of each cluster can be known. The results of the classification algorithm with Naive Bayes and K-NN using the data split operator are as follows:

Table 7. Accuracy Results of Data Classification with Naive Bayes and K-NN Using Split Data.

Data	Accuracy Naive Bayes (split data)	Accuracy K-NN (split data)
Data Clustering 2	95.71%	92.86%
Data Clustering 3	94.29%	94.29%

Data Clustering 4	87.14%	85.71%
Data Clustering 5	91.55%	92.96%

3.8 Information Gain

Information gain is a method used to select features or attributes so that know the limit of importance of an attribute. The highest value is the value that has a gain of 1. The results obtained from testing using information gain found that in testing cluster 2 data the attribute that has the smallest value is attribute Y4 which has a value of 0.132 and the attribute that has the most influence on attribute selection in cluster 2 data is M4 with a value of 0.369. Testing cluster 3 data is the attribute that has the most influence on the selection of cluster 3 data, which is X3 with a value of 0.458, while the attribute with the least effect is Y4 with a value of 0.174. Testing this cluster 4 data is the attribute that has the most influence is the M2 attribute with a value of 0.516 while the attribute that has the least influence is Y4 which has a value of 0.172. Testing this cluster 5 data is the most influential attribute is the M3 attribute with a value of 0,228.

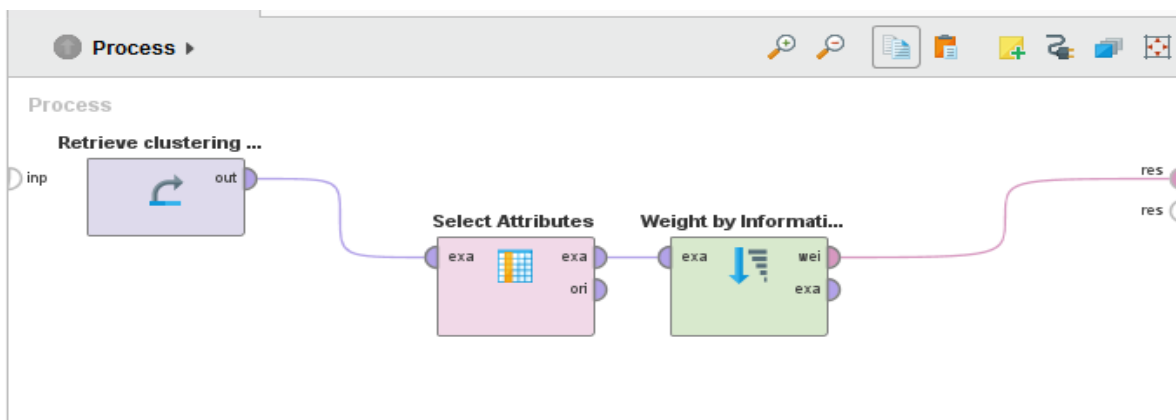


Figure 8. Attribute Testing Process for Each Cluster with Information Gain

3.9 Prediction Results

Prediction result from data classifiers using the Naive Bayes algorithm and K-NN algorithm with data split operators and cross validation can be seen in tables 2 and 3.

Table 8. Prediction Results from Data Classification with Naive Bayes and K-NN using cross validation.

Number of Predictions That Dont Match Initial Cluster Data			
Cross Validation Naive Bayes			
Data Cluster 2	Data Cluster 3	Data Cluster 4	Data Cluster 5
7 Data	9 Data	12 Data	15 Data
Cross Validation K-NN			
Data Cluster 2	Data Cluster 3	Data Cluster 4	Data Cluster 5
12 Data	27 Data	33 Data	37 Data

The table 8 shows the results that algorithm Naive Bayes has a difference prediction result with *cluster* the beginning less than the K-NN algorithm. Cross Validation here uses K-Folds 10, with 12 data attributes and the total data is 350 questionnaire data.

Table 9. Prediction Results from Data Classification with Naive Bayes and K-NN using split data.

Number of Predictions That Dont Match Initial Cluster Data			
Split Data Naive Bayes			
Data Cluster 2	Data Cluster 3	Data Cluster 4	Data Cluster 5
3 Data	4 Data	9 Data	6 Data

Split Data K-NN			
Data Cluster 2	Data Cluster 3	Data Cluster 4	Data Cluster 5
5 Data	4 Data	10 Data	5 Data

The data split method here uses 0.8 training data and 0.2 testing data, respectively. The number of "k" used in this K-NN algorithm method is 10. The above results show that the K-NN algorithm method has more predictions that are not in accordance with the initial cluster.

3.10 General Discussion of Results and Methods

SPSS was used to find characteristic data from questionnaires that had been distributed based on age, gender, last education and occupation. Testing the instrument is testing the questionnaire statement which will be distributed. Testing using SPSS in research is also to find out the description of respondents' answers and the average answers given from respondents.

The test using SEM PLS in this research is to find the value of the Outer Model and Inner Model. The value of the Outer Model is used to test convergent validity (measurements of a construct must be highly correlated), discriminant validity (measurements from different constructs must not be highly correlated). SEM PLS in this study is also used to find the value of the inner model. The results obtained between the two tests using SPSS and SEM PLS show that the validity and reliability tests are valid and reliable, so they are feasible to use.

Analysis using clustering and classification algorithms aims to determine the classification of each parameter used to answer the statement, there are five parameters used to fill out the questionnaire. The results of the five-cluster test showed that the more parameters used, the smaller the Davies Bouldin value obtained so that the results were better.

The comparison between the Naive Bayes classification algorithm and K-NN in the study shows that the Naive Bayes classification algorithm has higher accuracy results than using cross validation and split data methods. Predictions from using the Naive Bayes algorithm have prediction results that are not much different from the initial data, so that from the known prediction values that are not much different from the initial data, it can be stated that the Naive Bayes algorithm is more suitable for use in research and the Naive Bayes method can also produce results that high accuracy with little training data.

4. Conclusions

The conclusion of the study is based on the results of the research analysis and the results of the discussion of the research influence *Iklanku* Shopee feature through brand awareness of fashion product purchasing decisions using the Naive Bayes and K-NN classification algorithms with K-Means. The conclusions are as follows: Shopee *Iklanku* feature is influential positive and significant to fashion product purchase decisions on the Shopee application, so that the hypothesis is accepted. Shopee *Iklanku* feature is influential positive and significant to the brand awareness of fashion products on the Shopee application and the hypothesis is accepted. *Brand awareness* positive and significant effect on purchasing decisions for fashion products on the Shopee application and the hypothesis is accepted. *Brand awareness* positive and significant effect in mediating the effect of Shopee *Iklanku* feature on the decision to purchase fashion products on the Shopee application, so that the hypothesis is accepted.

Result of clustering 350 data In this questionnaire, there are 5 clusters, namely cluster 0, cluster 1, cluster 2, cluster 3, and cluster 4. The cluster validation test using Davies Bouldin got the results that the cluster 5 test got the best results. The result of classification in this study is the Naive Bayes algorithm has a more accurate accuracy value than the K-NN algorithm, this validation uses cross validation with k-folds 10 and split training data data is 0.8 and test data is 0.2. K on the K-NN used in each validation process is 10.

Prediction results and accuracy of the Naive Bayes algorithm is higher than the K-NN algorithm. This shows that the Naive Bayes algorithm works optimally, to determine decisions from prior knowledge.

References

- [1] E. Erdawati, "Pengaruh Gaya Hidup, Iklan Dan Harga Terhadap Keputusan Pembelian Online Melalui Shopee," *J. Apresiasi Ekon.*, vol. 8, no. 3, pp. 365–371, 2020, doi: 10.31846/jae.v8i3.331.
- [2] J. R. Priana and S. Suwandi, "Pengaruh Iklan 'Introduction: JO & JEK' di Youtube Terhadap Brand Awareness (Studi Kasus Perusahaan GO-JEK)," *J. Entrep. Manag. Ind.*, vol. 2, no. 4, 2019, doi: 10.36782/jemi.v2i4.1961.
- [3] C. Saraswati, Suharno, and G. N. Achmad, "the Effect of Advertising Attractiveness and Celebrity Endorser on Brand Image and Buying Intention of Shampoo Pantene," *Procur. J. Ilm. Manaj.*, vol. 8, no. 2, pp. 216–226, 2020.
- [4] P. Purwanti, S. Sarwani, and D. Sunarsi, "Pengaruh Inovasi Produk Dan Brand Awareness Terhadap Keputusan Pembelian Konsumen Pada Pt. Unilever Indonesia," *Inovasi*, vol. 7, no. 1, p. 24, 2020, doi: 10.32493/inovasi.v7i1.p24-31.5442.
- [5] Putra, Piarsa, and Lesmana, "Identification of Baby's Feet Using Principal Component Analysis (PCA) Method Character Extraction with K-Nearest Neighbor (KNN) Classification in MATLAB Application," *Identif. Baby's Feet USSing Princ. Compon. Anal. Method Character Extr. With K-Nearest Neighbor Classif. MATLAB Appl.*, vol. 9, pp. 1–12, 2021.
- [6] G. Angga, N. Piarsa, and M. Suwija, "Comparison of Support Vector Machine and K-Nearest Neighbor for Baby Foot Identification based on Image Geometric Characteristics," vol. 9, no. 1, pp. 84–95, 2021.
- [7] D. J. Osak and Y. Pasharibu, "Pengaruh Brand Ambassador Dan Tagline Terhadap Keputusan Pembelian Online Dengan Mediasi Brand Awareness," *Ekon. Dan Bisnis*, vol. 4, pp. 357–380, 2019.
- [8] T. Kuspriyono, "Pengaruh Iklan Terhadap Keputusan Pembelian Apartemen Meikarta," *Cakrawala*, vol. XVIII, no. 1, pp. 59–66, 2018.
- [9] S. O. Moloku, J. R. . Tampi, and O. Walangitan, "Pengaruh Iklan Terhadap Keputusan Pembelian Produk Smartphone Oppo Di IT Center Manado," *Bisnis, J. Adm.*, vol. 8, no. 2, pp. 98–106, 2019.
- [10] T. Nur, "Pengaruh Profitabilitas Terhadap Nilai Perusahaan Dengan Struktur Modal Sebagai Variabel Intervening," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2018.
- [11] L. Mananeke, H. Tawas, F. Ekonomi, and J. Manajemen, "Analisis Faktor Determinan Keputusan Pembelian Digerai Starbucks Manado Town Square," *J. EMBA J. Ris. Ekon. Manajemen, Bisnis dan Akunt.*, vol. 7, no. 4, 2019, doi: 10.35794/emba.v7i4.26022.
- [12] D. L. Trenggonowati and K. Kulsum, "Analisis Faktor Optimalisasi Golden Age Anak Usia Dini Studi Kasus Di Kota Cilegon," *J. Ind. Serv.*, vol. 4, no. 1, pp. 48–56, 2018, doi: 10.36055/jiss.v4i1.4088.
- [13] Ardiyansyah, P. A. Rahayuningsih, and R. Maulana, "Analisis Perbandingan Algoritma Klasifikasi Data Mining Untuk Dataset Blogger Dengan Rapid Miner," *J. Khatulistiwa Inform.*, vol. VI, no. 1, pp. 20–28, 2018.
- [14] A. Badruttamam, S. Sudarno, and D. A. I. Maruddani, "Penerapan Analisis Kluster K-modes dengan Validasi Davies Bouldin Index Dalam Menentukan Karakteristik Kanal Youtube di Indonesia (Studi Kasus: 250 Kanal YouTube Indonesia Teratas Menurut Socialblade)," *J. Gaussian*, vol. 9, no. 3, pp. 263–272, 2020, doi: 10.14710/j.gauss.v9i3.28907.