New Method in SEM Analysis Using the Apriori Algorithm to Accelerate the Goodness of Fit Model

Dien Novita[1], Ermatita[2*], Samsuryadi[3], Dian Palupi Rini[4]
Doctoral Program in Engineering Science, Universitas Sriwijaya, Palembang, Indonesia[1]
Faculty of Computer Science, Universitas Sriwijaya, Palembang, Indonesia[2, 3, 4]
Faculty of Computer Science and Engineering, Universitas Multi Data Palembang, Palembang, Indonesia[1]

Abstract*—This research aims to develop a new method in Structural Equation Modelling (SEM) analysis using the Apriori algorithm to accelerate the achievement of Goodness of Fit models, focusing on traditional retail purchasing decision models in Indonesia, especially in Palembang. SEM will be used to model causal relationships between variables that influence purchasing decisions in traditional retail. However, the Goodness of Fit model testing process takes a long time due to the complexity of the model. Therefore, this research uses the Apriori algorithm to filter variables that have a significant relationship in traditional retail purchasing decision models to reduce model complexity and speed up Goodness of Fit calculations. There are two stages in the research. First, the Apriori algorithm identifies frequent item sets that frequently appear among variables influencing traditional retail consumer purchasing decisions, such as product, price, and location. This pattern becomes the basis for SEM modeling, focusing on selected variables and, in the second stage, measuring the Goodness of Fit of the SEM model, namely GFI, RMSEA, AGFI, NFI, and CFI, to evaluate the suitability of the model which explains the factors that support traditional retail purchasing decisions in Palembang. The practical implications of this research are significant, as it provides a more efficient and effective method for modeling and understanding consumer behavior in the context of traditional retail. Based on other studies, if this research uses a conventional SEM approach, it does not meet the cut-off value of Goodness of Fit. Meanwhile, the results of the proposed method, namely combining Apriori into SEM, called APR-SEM, obtained a significant Goodness of Fit evaluation. The model coefficient of determination value from APR-SEM is R^2 0.71, higher than the conventional model, R^2 0.52. This method effectively simplifies the SEM model by identifying the most relevant relationships, thereby providing a clearer understanding of the critical factors influencing purchasing decisions in traditional retail in Palembang City.

Keywords*—APR-SEM; method; goodness of fit; traditional retail

I. Introduction

The existence of retail in Indonesia is relatively fast. According to 2021, Global Retail Development Index (GRDI) data, Indonesia managed to rank fourth, up one place from 2019 [1]. Meanwhile, Indonesia's retail sales growth was reported to increase by 3% in September 2021 [2]. The Indonesia Retail Summit 2023, which carries the theme ASEAN Retail: Epicentrum of Growth, shows the enthusiasm and commitment to Indonesia's role in strengthening the ASEAN and even the global retail industry. Currently, the Indonesian retail industry is the largest in ASEAN; Indonesia is in the number one position for retail in ASEAN [3].

In Indonesia, there are two forms of retail business run by the community, namely traditional retail and modern retail [4]. In contrast to traditional retail, modern retail is a shop with a self-service system that sells various types of goods at retail in minimarkets, malls (supermarkets and hypermarkets), department stores, or wholesale shops in the form of wholesalers [5].

With modern management, service, quantity, quality, and prices, modern retail has a higher competitiveness than traditional retail. Many traditional retailers then lack visitors, and their turnover decreases until they finally close due to their inability to compete with the many modern retailers nearby [6][7]. However, quite a few traditional retailers can survive amidst the invasion of many modern minimarkets popping up around these traditional retailers. These retailers continue to survive, sell, and serve buyers and are still busy with buyers. Interestingly, traditional retail in Indonesia also dominates the retail sales business in Indonesia [8]. Retail sales activities in Indonesia are still dominated by traditional retail. This can be seen from the number of outlets and the sales value which is much higher than modern retail outlets. According to data compiled by Euromonitor [8], there are 3.57 million traditional retail outlets in Indonesia.

With the current conditions, where the Indonesian retail industry is the largest in ASEAN, and traditional retail dominates the retail sales business in Indonesia, the large number of studies that discuss the existence of retail, including traditional retail types, are the basis for this research. Traditional retail businesses aim to increase competitiveness to survive with unique products, services, and facilities according to traditional retail characteristics and, of course, to make a profit. Increasing competitiveness in traditional retail businesses must be distinct from the owner's ability to manage the business. Traditional retail ownership, which is more personal, requires managers to survive by relying on brilliant ideas to advance their business. The use of ideas or concepts known as tacit knowledge is part of knowledge management [9]. Some research models are based on converting tacit knowledge into individual knowledge [10]. The purchasing decision model in traditional retail has 3 variables, namely product, price, and place which influence purchasing decisions [11].

In various disciplines such as marketing, management, and information technology, analysis of relationships between

variables is critical to understanding the dynamics that influence research results. Structural Equation Modelling (SEM) is a powerful method for analyzing complex relationships between latent variables and their indicators. Meanwhile, the Apriori algorithm, which originates from the field of data mining, is well known for analyzing association patterns between items in large data sets. Although SEM is very effective for modelling structural relationships between variables, this method is not designed to handle the analysis of high-frequency association patterns in large data sets. SEM can be obtained through the partial least squares structural equation modelling technique [12]. On the other hand, the Apriori algorithm can identify significant association rules in the data but cannot explicitly model the hypothesized causal relationships as SEM does. Apriori can generate association rules on data sets, for example, transactional basket data, to produce variations in sets of items (baskets) that adequately represent consumer purchasing patterns [13].

Therefore, combining these two methods offers a more comprehensive approach to analyzing complex data. This research aims to produce a model related to purchasing decisions in traditional retail using SEM analysis. SEM analysis tests the model's suitability by examining various Goodness of Fit criteria. However, the model Goodness of Fit testing process often takes a long time due to the complexity of the model. Therefore, combining the Structural Equation Modelling(SEM) method with the Apriori algorithm will make it easy to analyze complex relationships between variables and identify significant association patterns in the data. This approach can provide more profound and comprehensive insight into the relationships between variables so that Goodness of Fit criteria are quickly achieved.

This paper consists of several parts. In Section II, this paper reviewing several concepts related to research involving Structural Equation Modeling techniques and the Apriori Algorithm; besides looking at the retail conditions in Indonesia. Section III, this paper provides a general overview of research methodology, including respondent demographics and research model. Section IV of this paper explains the proposed research method regarding the stages. Section V, this paper provides results and discussion, and finally, Section VI concludes this paper.

II. Literature Review

The existence of traditional retail has a significant influence on economic development in Indonesia with the number of traditional retailers surpassing modern retail. Based on type, traditional grocery stores are the most numerous retailers in Indonesia. The number was recorded at 3.57 million units. A total of 38,323 retailers are in the form of department stores. Then, there are 1,411 supermarket-type retailers. Then, forecourt retail and hypermarkets were 358 units and 285 units, respectively, as in Table I.

The dominant type of retail in Indonesia is the traditional grocery store, which is both wholesale and retail. The types of products sold in traditional grocery stores are shown in Table II below.

TABLE I. Types of Retail and Number in Indonesia in 2022 [14]

Retail Type	Total (unit)
Traditional Grocery	3,935,238
Department Store	41,453
Supermarket	1,544
Food/Beverage Specialist	5,455
Hypermarket	298

TABLE II. Types of Products in Traditional Grocery Stores

Types of Products	Source
Necessities, snacks, drinks, toiletries and washing supplies, household supplies, medicines, kitchen spices, instant food, LPG gas, and stationery supplies.	[15]
Household equipment and necessities.	[16]
A variety of daily household needs	[17]

State-of-the-art research [18][19][20][21][22][23] which formulates business strategies for retail businesses, especially small-scale retail, in the form of developing business strategy models using data mining, artificial neural networks, and structural equation modelling, no one has combined the tools between data mining and structural equation modelling. However, research in study [24] and [25] has used a combination of SEM and data mining methods with model validation between 54%-89%. The research model [26] is the Apriori + hybrid structural equation model (SEM) using a data-based Apriori algorithm that explores large-scale hidden relationships between variables. Meanwhile, the SEM model captures user behaviour and decision-making procedures, providing interpretable results. Additionally, association rules in Apriori facilitate the specification of complex SEM models, thereby substantially reducing modelling calibration. The Goodness of Fit results show that the SEM + Apriori hybrid model performs better than the conventional model, namely with an R^2 value of 0.82, which is greater than the conventional model, 0.69.

From the previous research above, this is the first time anyone has specifically proposed a combination of Apriori and SEM analysis in formulating research models. Although research [26] has used a hybrid SEM + Apriori model, it is limited to the automotive industry. So, this research proposes SEM analysis using the Apriori algorithm to model purchasing decisions in traditional retail. The following are several theories related to the proposed method in this research, namely Association Rule, Apriori Algorithm, and Structural Equation Modelling.

A. Association Rule

Association rule mining is a data mining technique that finds similar rules in an event [27]. An example of an association rule that is often encountered is the process of purchasing merchandise at a shopping centre. From historical data it is known that the purchase of bread mainly follows the purchase of milk, so shop owners can increase their profits by arranging the location of milk and bread close together and providing discounts that attract buyers. Commonly used association methods are FP-Growth, Coefficient of Correlation, Chi-Square, Apriori, and others.

B. Apriori Algorithm

The Apriori algorithm, a classic in data mining, is a powerful tool for uncovering association rules. Its main purpose is to delve deep into the complex network of associationrelationships between variables [26], thereby showcasing the depth of the topic.

- Support is the probability that items A and B appear simultaneously, representing the importance of the corresponding association rule in the dataset.

$$support(A \rightarrow B) = P(A \cup B)$$

- Confidence is the proportion of the number of times that A and B appear simultaneously over the number of times that A appears. This represents the credibility of the association rule.

$$confidence(A \rightarrow B) = \frac{support(A \rightarrow B)}{support(A)}$$

Key metrics in the Apriori algorithm are based on:

- Support: Percentage of transactions containing itemsets or rules. The higher the support, the more frequently the item or rule occurs.
- Confidence: The probability that the consequent occurs when the antecedent exists. A value close to 1 indicates the strength of the rule.
- Lift: It's the key that unlocks the door to understanding association. The ratio between actual support and expected support if the antecedent and consequent were independent. Values greater than 1 indicate a positive association, enriching our knowledge about the data.
- Conviction: Measuring the strength of a rule by considering the frequency of antecedents that are not followed by consequents. The infinity (inf) value indicates the rule is firm.

The steps in the Apriori Algorithm are:

1) Determine minimum support and confidence
2) Find items that appear frequently (frequent itemsets)
3) Remove items that don't appear frequently
4) Create association rules
5) Evaluate association rules

C. Structural Equation Modelling

Structural Equation Modelling (SEM) is a technique for observing the interdependence between various variables. It is a confirmatory method to check the suitability of data and conceptual models. This is especially true when drawing conclusions where variables cannot be measured. SEM is also called a combination of factor analysis and multiple regression. If completed on a software platform, it can be completed using IBM SPSS AMOS, LISREL, and R [28]. The stages in the SEM method are [18]:

1) Development of a theoretical model
2) Development of path diagrams
3) Convert the path diagram into an equation
4) Select input matrix and model estimation
5) Possible identification problems
6) Evaluate goodness of fit criteria

The minimum number of samples in SEM recommended is 100-150 data [29]. Five or fewer constructs exist, each with more than three items (observed variables) and high item communality (0.6 or higher). Table III shows the goodness of fit in step 6 for the model.

TABLE III. MODEL-OF-FIT INDICES

Model-of-fit indices	Full name/key concerns	Cut off value
RMSEA	Root Mean Square Error of Approximation	Value between 0.08 and 0.10 (mediocre fit), <0.08 (good fit)
GFI	Goodness of fit statistics Exhibits bias towards samples	Value >0.90 or >0.95 (use 0.95 if factor loading and number of sample are low)
AGFI	Adjusted goodness of fit statistics Needs to be accompanied by other indices	Value of >0.80
NFI	Normed fit index Sensitive to sample size	Value of >0.90
CFI	Comparative fit index Revised version of NFI Less affected by sample size	Value of ≥0.90

III. RESEARCH METHODOLOGY

Questionnaire data collection used Google Forms [29] to reach all areas of the city of Palembang as research objects. Table IV shows the demographics of a total of 213 respondents.

TABLE IV. RESPONDENT DEMOGRAPHICS

Sample Characteristic		Percent (%)
Gender	Male	57.75
	Female	42.25
Age	Lower than 20 years	39.91
	20-29	49.30
	30-39	5.16
	40-49	4.69
	50-59	0.94
Occupation	Not yet working	0.94
	BUMN	0.47
	Teacher/Lecturer	1.41
	Housewife	48.83
	Student	43.19
	PNS/TNI/Polri	1.41
	Private	3.76
Area	Alang-alang Lebar	8.45
	Bukit Kecil	0.94
	Gandus	1.41
	Ilir Barat I	8.92

	lir Barat II	8.45
	Ilir Timur I	7.98
	Ilir Timur II	14.55
	Ilir Timur III	6.57
	Jakabaring	2.82
	Kalidoni	10.80
	Kemuning	0.47
	Kertapati	0.94
	Plaju	1.41
	Sako	8.45
	Seberang Ulu I	1.41
	Seberang Ulu II	0.94
	Sematang Borang	7.51
	Sukarami	7.98

The object of this research is traditional retail consumers in Palembang City. Fig. 1 shows the methodology adopted for this study. First, the research focuses on reviewing journal literature related to traditional retail existence models, producing a research conceptual model, and compiling a questionnaire based on this conceptual model.

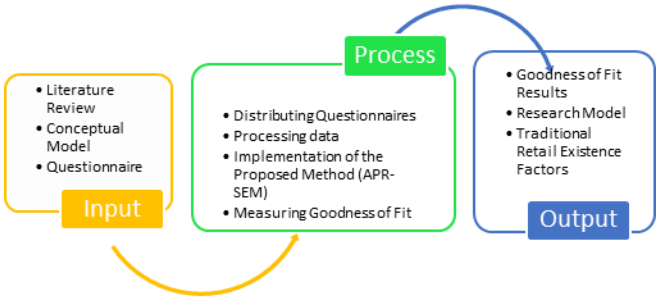

Fig. 1. Research methodology.

The conceptual model of purchasing decisions in traditional retail consists of four variables: exogenous variables: product, price, and place, and endogenous variables: purchasing decisions, as in Fig. 2.

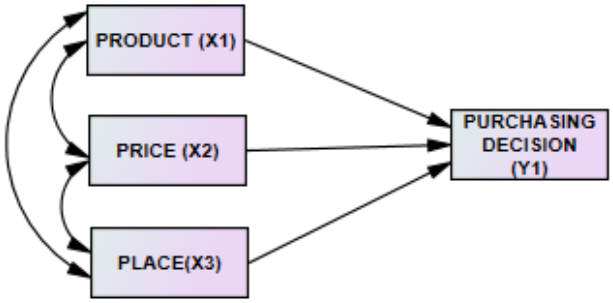

Fig. 2. Purchasing decision models in traditional retail.

Table V shows the indicators for each variable in this research.

TABLE V. INDICATORS OF EACH VARIABLE

Variable	Indicator	Code
Product (X1)	product availability	X11
	product variety	X12
	quality product	X13
	product packaging	X14
Price (X2)	price according to product quality	X21
	cheap price	X22
	discounts	X23
	bargaining price	X24
Place (X3)	comfortable and safe	X31
	clean	X32
	location	X33
	parking area	X34
Purchasing Decision (Y1)	self-interest	Y11
	make ends meet	Y12
	location near home	Y13
	time is not long	Y14
	social level	Y15

The questionnaires used Google Form media, and the data in this research used AMOS for SEM analysis and Jupyter Lab Python for the Apriori algorithm. The next step is to test the model's goodness of fit. One of the main challenges in SEM is a model's Goodness of Fit testing process, which evaluates how well a proposed model fits empirical data. These models, due to their complexity involving many variables and complex causal relationships, often take a long time to test, especially when they are highly complex.

IV. PROPOSED METHOD

This research proposes an APR-SEM method by integrating the Apriori algorithm in SEM analysis to overcome the challenges of this research model, which involves many variables, complex causal relationships, and testing processes. The Apriori algorithm is used in association analysis to find significant relationship patterns between variables. In the context of SEM, this algorithm can filter and identify variables with a significant relationship so that only relevant variables are included in the final model. So, Fig. 3 is a proposed method combining the Apriori algorithm into the SEM analysis technique with the following process.

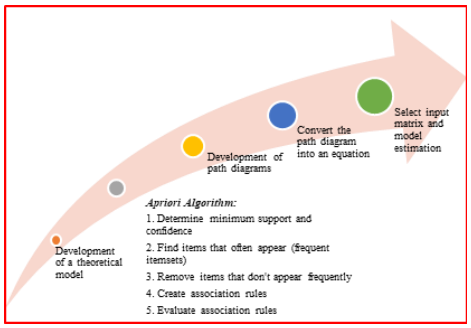

Fig. 3. Proposed method APR-SEM.

APR-SEM is an approach that combines the power of the Apriori algorithm to filter significant variables before use in an SEM model. This aims to:

1) Reduce the complexity of the SEM model by filtering out irrelevant variables.

2) Speed up achieving Goodness of Fit in the SEM model by looking for the strongest associations between variables.

Overall, this new method shows how the Apriori algorithm can be used upfront to filter out significant variables before inserting them into the SEM model, thereby speeding up and simplifying the SEM evaluation process.

V. Result and Discussion

A. Result

The stages in the SEM method are the development of a theoretical model. This stage builds a research model using an in-depth literature study related to research related to purchasing decisions in traditional retail. This stage produces consumer behaviour with the determining factors for purchasing decisions in traditional retail: price, product, and place variables. The next stage is the development of path diagrams. Before entering this stage, the Apriori algorithm will look for the strongest associations between the indicators in the research model. In this case, using Jupyter Lab Python, get the most robust association rule from the research variables. The results of the best rule association for each research variable are as in Table VI.

TABLE VI. Association Best Rule

Variable	Association Best Rule	Key Matric			
		Support	Confidence	Lift	Conviction
X1	X12→X11	98.4%	1.000	1.016	Inf.
	X11→X12	98.4%	1.000	1.016	Inf.
X2	X21→X22	59.8%	0.987	1.062	5.457
	X22→X21	59.8%	0.987	1.062	5.457
X3	X32→X31	98.4%	0.992	1.008	1.984
	X31→X32	98.4%	0.992	1.008	1.984
Y1	Y11→Y12	98.4%	1.000	1.008	Inf.
	Y12→Y11	98.4%	1.000	1.008	Inf.
	Y11→Y14	81.1%	0.824	1.016	1.073
	Y14→Y11	81.1%	0.824	1.016	1.073

The Apriori results will be compared between the conventional SEM results without Apriori and the proposed APR-SEM method for variable X1, as in Fig. 4.

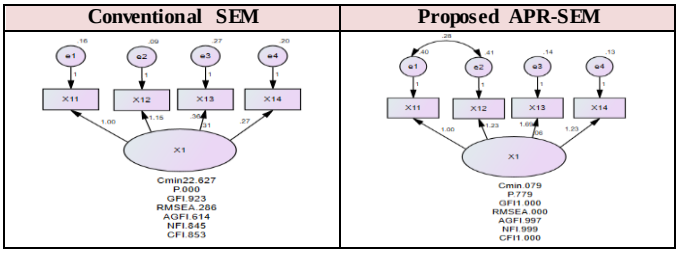

Fig. 4. The comparison of X1 results.

The goodness of fit results for variable X1, as in Table VII.

TABLE VII. Goodness of Fit Variable X1

Index	Cut off value	Conventional SEM	Proposed APR-SEM
GFI	> 0,90	0.923	1.000
RMSEA	< 0,08	0.286	0.000
AGFI	> 0,80	0.614	0.997
NFI	> 0,90	0.845	0.999
CFI	> 0,90	0.853	1.000

From the goodness of fit results of the product variable (X1), the goodness of fit results using the proposed Apriori-SEM method obtained results that all met the cut-off value.

The Apriori results will be compared between the conventional SEM results without Apriori and the proposed APR-SEM method for variable X2, as in Fig. 5.

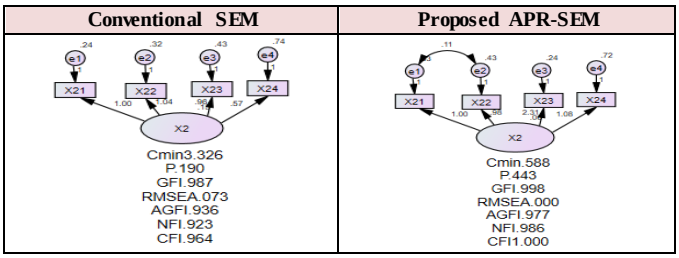

Fig. 5. The comparison of X2 results.

From the Apriori results, the best association rule is between variables X21 and X22, so there is a relationship between them in the path diagram. The goodness of fit results for variable X2, as in Table VIII.

TABLE VIII. Goodness of Fit Variable X2

Index	Cut off value	Conventional SEM	Proposed APR-SEM
GFI	> 0,90	0.987	0.998
RMSEA	< 0,08	0.073	0.000
AGFI	> 0,80	0.936	0.977
NFI	> 0,90	0.923	0.986
CFI	> 0,90	0.964	1.000

The goodness of fit results of the product variable (X2) using the proposed APR-SEM method obtained results that all met the cut-off value, the same as conventional SEM. But in this case, the P value in Fig. 5 is higher in the proposed APR-SEM model, namely 0.443, which shows that the model is better.

The Apriori results will be compared between the conventional SEM results without Apriori and the proposed APR-SEM method for variable X3, as in Fig. 6.

From the Apriori results, the best association rule is between variables X31 and X32, so there is a relationship between them in the path diagram. The goodness of fit results for variable X3, as in Table IX.

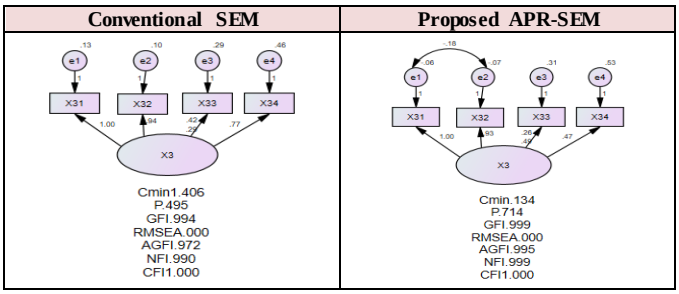

Fig. 6. The comparison of X3 results.

TABLE IX. GOODNESS OF FIT VARIABLE X3

Index	Cut off value	Conventional SEM	Proposed APR-SEM
GFI	> 0,90	0.994	0.999
RMSEA	< 0,08	0.000	0.000
AGFI	> 0,80	0.972	0.995
NFI	> 0,90	0.990	0.999
CFI	> 0,90	1.000	1.000

The goodness of fit results of the product variable (X3) using the proposed APR-SEM method obtained results that all met the cut-off value, the same as conventional SEM. But in this case, the P value in Fig. 6 is higher in the proposed APR-SEM model, namely 0.714, which shows that the model is better.

The Apriori results will be compared between the conventional SEM results without Apriori and the proposed APR-SEM method for variable Y1, as in Fig. 7. From the Apriori results, the best association rule is between variables Y11 and Y12 and also Y11 and Y14, so there is a relationship between them in the path diagram. The goodness of fit results for variable Y1, as in Table X.

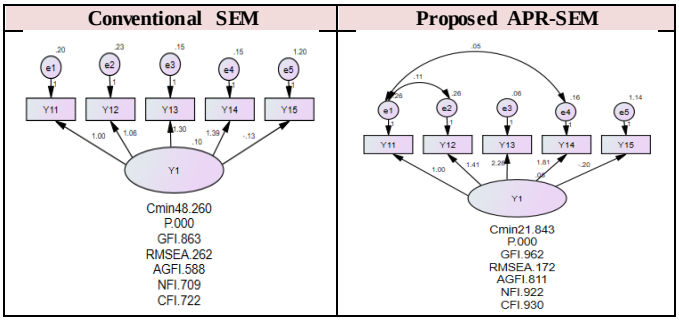

Fig. 7. The comparison of Y1 results.

TABLE X. GOODNESS OF FIT VARIABLE Y1

Index	Cut off value	Conventional SEM	Proposed APR-SEM
GFI	> 0,90	0.863	0.962
RMSEA	< 0,08	0.262	0.172
AGFI	> 0,80	0.588	0.811
NFI	> 0,90	0.709	0.922
CFI	> 0,90	0.722	0.930

The goodness of fit results of the purchasing decision variable (Y1) using the proposed Apriori-SEM method obtained results that all met the cut-off value except RMSEA, but this can be considered because the other indices have met. This is different from conventional SEM, which does not meet all requirements. This shows that the model is better.

Before describing the path diagram of the entire research model, it is necessary to convert it into equation form. The research model has latent variables (X1, X2, X3, and Y1) which are connected to several measurement variables such as X11, X12, X13, and X14 for X1 variable. X21, X22, X23, and X24 for X2 variable. X31, X32, X33, and X34 for X3 variable. Y11, Y12, Y13, Y14, and Y15 for Y1 variable. Convert the path diagram to an equation, namely:

1) Measurement model: X1 is determined by the indicators X11, X12, X13, and X14

$$X11 = \lambda_{11}X1 + e_1 \quad (1)$$

$$X12 = \lambda_{12}X1 + e_2 \quad (2)$$

$$X13 = \lambda_{13}X1 + e_3 \quad (3)$$

$$X14 = \lambda_{14}X1 + e_4 \quad (4)$$

X2 is determined by the indicators X21, X22, X23, and X24

$$X21 = \lambda_{21}X1 + e_5 \quad (5)$$

$$X22 = \lambda_{22}X1 + e_6 \quad (6)$$

$$X23 = \lambda_{23}X1 + e_7 \quad (7)$$

$$X24 = \lambda_{24}X1 + e_8 \quad (8)$$

X3 is determined by the indicators X31, X32, X33, and X34

$$X31 = \lambda_{31}X1 + e_9 \quad (9)$$

$$X32 = \lambda_{32}X1 + e_{10} \quad (10)$$

$$X33 = \lambda_{33}X1 + e_{11} \quad (11)$$

$$X34 = \lambda_{34}X1 + e_{12} \quad (12)$$

Y1 is determined by the indicators Y11, Y12, Y13, Y14, and Y15

$$Y11 = \lambda_{y11}Y1 + e_{13} \quad (13)$$

$$Y12 = \lambda_{y12}Y1 + e_{14} \quad (14)$$

$$Y13 = \lambda_{y13}Y1 + e_{15} \quad (15)$$

$$Y14 = \lambda_{y14}Y1 + e_{16} \quad (16)$$

$$Y15 = \lambda_{y15}Y1 + e_{17} \quad (17)$$

2) Structural model: Structural models explain the relationships between latent variables in the system. Y1 is influenced by X1, X2, and X3.

$$Y1 = \beta_{11}X1 + \beta_{12}X2 + \beta_{13}X3 + \xi_1 \quad (18)$$

3) Covariance between latent variables: The covariance between latent variables X1, X2, and X3 is written as the equation:

$$Cov(X1, X2) = \phi_{12} \quad (19)$$

$$Cov(X1, X3) = \phi_{13} \quad (20)$$

$$Cov(X1, X2) = \phi_{12} \quad (21)$$

So, the path diagram of the traditional retail purchasing decision research model is as in Fig. 8.

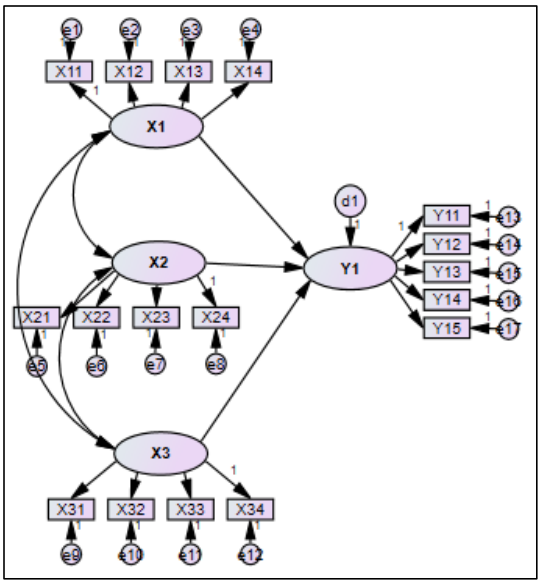

Fig. 8. Path diagram.

The next step is to choose the input matrix by selecting the Covariance Matrix or Correlation Matrix as the input data. After selecting the input matrix, determine the estimation method used. Maximum Likelihood (ML) is the default estimation method when using AMOS. After selecting the input matrix and estimation method, the next step is to run the analysis to obtain parameter estimates and model fit. The estimated calculation results from the research model using conventional SEM are shown in Fig. 9.

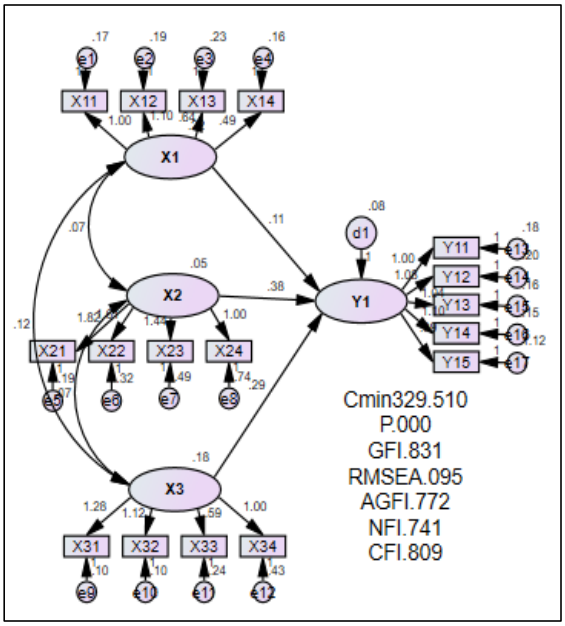

Fig. 9. The SEM conventional results.

From the Apriori results and modification of the path diagram model using AMOS, the calculated estimate results are in Fig. 10.

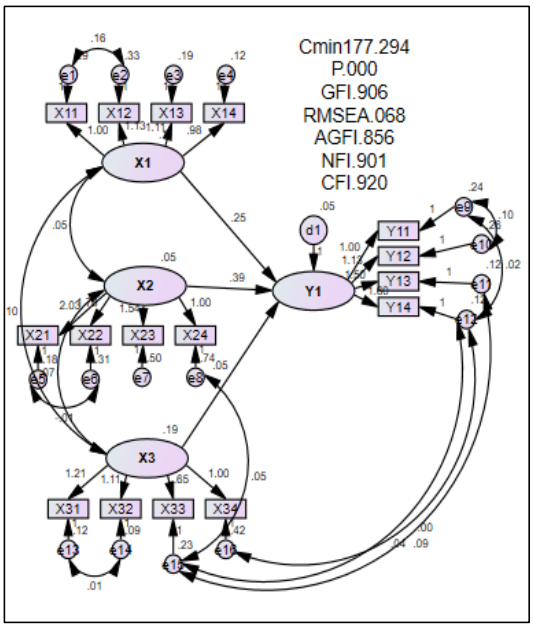

Fig. 10. The APR-SEM results.

From the Apriori results and the modification of the path diagram model using AMOS, the calculated estimate results meet the goodness of fit rules in conventional SEM and APR-SEM, as in Table XI.

TABLE XI. GOODNESS OF FIT FULL MODEL

Index	Cut off value	Conventional SEM	Proposed APR-SEM
GFI	> 0,90	0.831	0.906
RMSEA	< 0,08	0.095	0.068
AGFI	> 0,80	0.772	0.856
NFI	> 0,90	0.741	0.901
CFI	> 0,90	0.809	0.920

These results indicate that none of the measures of suitability when using conventional SEM are suitable. By adding the Apriori algorithm to find patterns or association rules between indicator variables and modifying the path diagram model using AMOS, the path diagram results for each variable X1, X2, X3, and Y1 obtained results that all met the cut-off value.

B. Discussion

From the results of research on the Goodness of Fit model using the APR-SEM method, it is stated that the purchasing decision model by traditional retail customers depends on three factors: product, price, and place. The model coefficient of determination value from APR-SEM is R^2 0.71, higher than the conventional model, R^2 0.52. Variable Y15, namely social level, has a value of 0 for the level of contribution to the model, so it is removed from the latent variable of the AMOS model. This is by the research results [30] and [31], which state that purchasing decisions at traditional retail are not based on a social level but

on consumer age, usually older, who is more likely to shop at traditional retailers.

This research also shows hidden relationships obtainedfrom association rules using the Apriori algorithm, namely the relationship between product availability and diversity, guaranteed product quality at low prices, and a comfortable, safe, clean place. Apart from that, there are four newcorrelations from the modification of the model from AMOS using Modification Indices. Namely, there is a relationship between the main tendencies of buyers in traditional retail because the retail location is close to home, and they can haggle over prices and not wait long for service. Besides that, there is a correlation between parking space availability for buyers who can be served quickly in traditional retail, such as shopping straight from their vehicle.

So, a clear agenda for future Structural Equation Modeling analysis to achieve the Goodness of Fit model (APR-SEM method) is to follow the following steps:

1) Development of a theoretical model
2) Evaluation of best association rules with the Apriori Algorithm
3) Development of path diagram
4) Convert the path diagram into an equation
5) Estimates goodness of fit model

VI. Conclusion

Combining the Apriori algorithm with the SEM method produces a more comprehensive analytical approach to understanding the relationships between variables and association patterns in the data. This can be seen from the purchasing decision model in traditional retail, which does not meet goodness of fit and becomes fit after adding the best association rules to the research model path diagram. The combination of SEM and Apriori allows the discovery of new insights that might have yet to be revealed if these two methods were used separately. For example, associations discovered by Apriori can lead to more in-depth modeling of structural relationships in SEM. This purchasing decision model in traditional retail will provide an analytical framework that helps in understanding the factors that influence the sustainability and competitiveness of traditional retail amidst market competition with modern retail and technological developments. This model will guide traditional retail owners in formulating appropriate strategies to increase their competitiveness, such as adaptingto consumer preferences, integrating technology, or improving service quality. However, combining these two methods, namely SEM and Apriori analysis, avoids including all associative patterns in SEM. Choose the most important relationships with strong theoretical or logical support. We cannot directly obtain goodness of fit in the model, but we must still modify the indices (MI) if necessary. However, be careful not to make modifications that violate the theory. Checkresidual correlations and eliminate non-significant relationships.

References

[1] AT Kerney, 'Global Retail Development Index', 2021. [Online]. Available: https://www.kearney.com/industry/consumer-retail/global-retail-development-index

[2] CEIC, 'Pertumbuhan Penjualan Ritel Indonesia', 2021. https://www.ceicdata.com/id/indicator/indonesia/retail-sales-growth

[3] PANRB, 'Indonesia Ritel Summit 2023 Dongkrak Pergerakan dan Tingkat Konsumsi Wisatawan', 2023. https://www.menpan.go.id/site/berita-terkini/berita-daerah/indonesia-ritel-summit-2023-dongkrak-pergerakan-dan-tingkat-konsumsi-wisatawan

[4] H. Chaniago, I. Mulyawan, T. Suhaeni, and R. Jumiyani, 'Faktor Kunci Keberhasilan Ritel Modern Di Indonesia', Jurnal Akuntansi, Ekonomi dan Manajemen Bisnis, vol. 7, no. 2, pp. 201–208, 2019, doi: 10.30871/jaemb.v7i2.1726.

[5] A. Tohri, M. Mastur, H. Habibuddin, H. Syamsiar, and L. Parhanuddin, 'Dampak Sosial dan Ekonomi Ritel Modern (Alfamart dan Indomaret) Terhadap UMKM di Lombok Timur', RESIPROKAL: Jurnal Riset Sosiologi Progresif Aktual, vol. 5, no. 1, pp. 45–56, 2023, doi: 10.29303/resiprokal.v5i1.280.

[6] P. S. Arimawa and F. Leasiwal, 'Dampak Keberadaan Pasar Modern Terhadap Eksistensi Pasar Tradisional Di Kota Tobelo Kabupaten Halmahera Utara', Jurnal Pundi, vol. 2, no. 3, pp. 287–292, 2018, doi: 10.31575/jp.v2i3.100.

[7] Z. Muhzinat and S. Achiria, 'Dampak Keberadaan Minimarket terhadap Toko Kelontong di Pasar Klampis Kabupaten Bangkalan Madura', IQTISHADIA Jurnal Ekonomi & Perbankan Syariah, vol. 6, no. 2, pp. 203–211, 2019, doi: 10.19105/iqtishadia.v6i2.2448.

[8] U. S. D. of Agriculture, 'Ritel Tradisional Dominasi Usaha Penjualan Eceran di Indonesia', 2022. [Online]. Available: https://databoks.katadata.co.id/datapublish/2022/07/12/ritel-tradisional-dominasi-usaha-penjualan-eceran-di-indonesia

[9] D. Novita and E. Ermatita, 'Implementation Knowledge Management for Knowing the Factors That Have Influenced Income for Traditional Retail', in Proceeding of International Conference of Health, Science and Technology, 2021, pp. 260–263. [Online]. Available: https://ojs.udb.ac.id/index.php/icohetech/article/view/1137/977

[10] C. Ronceros, J. Medina, P. León, A. Mendieta, J. Fernández, and Y. Martinez, 'Knowledge Management Model for the Generation of Innovative Capacities in Organizations that Provide Services', International Journal of Advanced Computer Science and Applications, vol. 14, no. 5, pp. 423–430, 2023, doi: 10.14569/IJACSA.2023.0140545.

[11] Muflihatul Fauza, 'Analisis Faktor Yang Mempengaruhi Eksistensi Ritel Tradisional Dalam Menghadapi Ritel Modern Di Kecamatan Medan Amplas', At-Tawassuth, vol. 2, no. 1, pp. 146–169, 2017.

[12] J. J. Rodríguez-Delgado, P. López-Casaperalta, M. G. Berrios-Espezúa, A. M. Acosta-Quelopana, and J. Sulla-Torres, 'Digital Learning Tools for Security Inductions in Mining Interns: A PLS-SEM Analysis', International Journal of Advanced Computer Science and Applications, vol. 13, no. 5, pp. 530–536, 2022, doi: 10.14569/IJACSA.2022.0130562.

[13] M. I. Akazue et al., 'Handling Transactional Data Features via Associative Rule Mining for Mobile Online Shopping Platforms', International Journal of Advanced Computer Science and Applications, vol. 15, no. 3, pp. 530–538, 2024, doi: 10.14569/IJACSA.2024.0150354.

[14] M. A. Rizaty, 'Jumlah Toko Retail Indonesia Mencapai 3,61 Juta pada 2021', dataindonesia.id, 2023. https://dataindonesia.id/industri-perdagangan/detail/jumlah-toko-retail-indonesia-mencapai-361-juta-pada-2021 (accessed Jul. 02, 2023).

[15] D. Januaji, 'Mau Buka Toko Kelontong? 10 Barang Ini Wajib Ada', ottopay.id, 2023.

[16] T. H. Putra, 'Toko Kelontong Tradisional Dalam Era Teknologi Bisnis Digital', vol. 2, no. 3, 2023.

[17] S. Lestari, B. A. Yani, and I. A. DPW, 'Analisis Perbedaan Persepsi Konsumen Minimarket Modern dan Toko Kelontong di Desa Kartonatan, Kartasura, Sukoharjo', Edunomika, vol. 5, no. 2, pp. 1032–1037, 2021, [Online]. Available: https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf

[18] S. T. Ha, M. C. Lo, M. K. Suaidi, A. A. Mohamad, and Z. Bin Razak, 'Knowledge management process, entrepreneurial orientation and performance in smes: Evidence from an emerging economy', Sustainability (Switzerland), vol. 13, no. 17, 2021, doi: 10.3390/su13179791.

[19] P. Pitakkornrat and P. Pongsiri, 'A Structural Equation Model of Knowledge Management Strategy to Develop Best Practice for Industrial Business in Thailand', Academy of Strategic Management Journal, vol. 19, no. 2, pp. 1–13, 2020.

[20] F. Ibrahim, B. S. Putra, F. H. Azhra, and N. Fadhlurrohman, 'Analysis of Marketing Strategy at Setia Stores Using AHP, Clustering, and AR-Mba Method', International Journal of Industrial Optimization, vol. 2, no. 2, p. 125, 2021, doi: 10.12928/ijio.v2i2.4369.

[21] B. Melović, M. Dabić, M. Vukčević, D. Ćirović, and T. Backović, 'Strategic Business Decision Making : The Use and Relevance of Marketing Metrics and Knowledge Management', Journal of Knowledge Management, vol. 25, no. 11, pp. 175–202, 2021, doi: 10.1108/JKM-10-2020-0764.

[22] A. R. Wibowo and A. Jananto, 'Implementasi Data Mining Metode Asosiasi Algoritma FP-Growth Pada Perusahaan Ritel', Inspiration: Jurnal Teknologi Informasi dan Komunikasi, vol. 10, no. 2, p. 200, 2020, doi: 10.35585/inspir.v10i2.2585.

[23] R. Takdirillah, 'Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan', Edumatic : Jurnal Pendidikan Informatika, vol. 4, no. 1, pp. 37–46, 2020, doi: 10.29408/edumatic.v4i1.2081.

[24] C. L. Yang, C. Y. Huang, and Y. H. Hsiao, 'Using Social Media Mining and PLS-SEM to Examine The Causal Relationship Between Public Environmental Concerns and Adaptation Strategies', International Journal of Environmental Research and Public Health, vol. 18, no. 10, 2021, doi: 10.3390/ijerph18105270.

[25] J. T. Thorson et al., 'Identifying Direct and Indirect Associations Among Traits by Merging Phylogenetic Comparative Methods and Structural Equation Models', Methods in Ecology and Evolution, vol. 14, no. 5, pp. 1259–1275, 2023, doi: 10.1111/2041-210X.14076.

[26] X. Lai, S. Zhang, N. Mao, J. Liu, and Q. Chen, 'Kansei engineering for new energy vehicle exterior design: An internet big data mining approach', Computers and Industrial Engineering, vol. 165, no. June 2021, p. 107913, 2022, doi: 10.1016/j.cie.2021.107913.

[27] M. A. Muslim et al., Data Mining Algoritma C4.5, Cetakan Pe. Semarang, 2019.

[28] J. J. Thakkar, Structural equation modelling: Application for research and practice (with AMOS and R), vol. 285. 2020. doi: 10.1007/978-981-15-3793-6_1.

[29] T. L. Wiemken, S. P. Furmanek, W. A. Mattingly, J. Haas, J. A. Ramirez, and R. M. Carrico, 'Googling your hand hygiene data: Using Google Forms, Google Sheets, and R to collect and automate analysis of hand hygiene compliance monitoring', American Journal of Infection Control, vol. 46, no. 6, pp. 617–619, 2018, doi: 10.1016/j.ajic.2018.01.010.

[30] I. Rakasyifa and G. W. Mukti, 'Faktor-Faktor Yang Mempengaruhi Keputusan Pembelian Sayur dan Buah Di Ritel Online (Suatu Kasus Pada Konsumen Ritel Online Di Jakarta)', Jurnal Pemikiran Masyarakat Ilmiah Berwawasan Agribisnis, vol. 6, no. 1, pp. 275–289, 2020.

[31] S. I. Isnawati and A. Purwanto, 'Generation Z Buying Behaviour Analysis of Retail Business Opportunities', Ilmiah Bisnis, Manajemen dan Akuntansi, vol. 2, no. 2, pp. 11–21, 2022, doi: http://dx.doi.org/10.35.