Application Of Correlation Data Mining the Covid-19 Pandemic With Student Achievement Level Using A Priori Method

Fiqri Agung Pangestu[*]

STMIK Kaputama Binjai, Indonesia

*Corresponding Author

Email : fiqriofficial21@gmail.com

Abstract

Education is carried out as early as possible so that a person gets the broadest insight into knowledge. The longer you have higher education will further improve the quality of human resources. Education can be obtained through lessons that are in accordance with the level of education and the learning material itself. The current condition of the COVID-19 outbreak has resulted in several agencies implementing Work From Home (WFH) or working from home, including educational institutions or agencies. The government provides a policy for every school to conduct online learning by using gadgets (laptops or cellphones) to access websites or learning links such as Whatsapp, Google Form, Google Classroom. A very significant difference in the student's learning process from usual, will affect the sharpening of students' ways of thinking about their learning. Students are required to be able to understand the material without direct interaction from the teacher, and the teacher is not able to provide the material as a whole to students. This will affect student achievement.

Keywords: Data Mining, A Priori, Correlation

INTRODUCTION

Conditions during the COVID-19 outbreak resulted in several agencies implementing work from home, including educational institutions. The government provides a policy for every school to conduct online learning by using a cellphone or laptop to access learning websites such as WhatsApp, Google Form, Google Classroom. significant differences in student learning from the previous greatly affect the way students learn.

Currently students are required to be able to understand the learning material without any direct interaction from the teacher, and also the teacher is not able to provide the material as a whole to students with existing limitations. This is what will affect student achievement.

learning achievement is the result of learning achieved after going through the process of learning activities. learning achievement can be shown through the value given by a teacher from the number of fields of study that have been studied by students. Student achievement in online learning remains the same, which is seen from the daily test scores, mid-semester tests, and end-semester tests.

RESEARCH METHODS

The a priori algorithm is a type of association rule in data mining. In addition to a priori, which are included in this group are the Generalized Rule Induction method and the Hash Based Algorithm.

Rules that state associations between attributes are often called affinity analysis or market basket analysis.

Association analysis or association rule mining is a data mining technique to find associative rules between a combination of items. An example of an associative rule of purchasing analysis in a supermarket is knowing how likely it is that a customer gives bread along with milk. With this

knowledge, supermarket owners can arrange the placement of their goods or design marketing campaigns by using discount coupons for certain combinations of goods. Association rules are usually expressed in the form:

{bread, butter} -> {milk} (support = 40%, confidence = 50%)

The rule means that "50%" of the transactions in the database containing bread and butter items also contain dairy items. Meanwhile, 40% of all transactions in the database contain those three items."

It can also be interpreted: "A consumer who buys bread and butter has a 50% chance of also buying milk. This rule is quite significant because it represents 40% of the transaction records so far". Association analysis is defined as a process to find all association rules that meet the minimum requirements for support (minimum support) and minimum requirements for confidence (minimum confidence).

Data Mining

According to Turban (2005) Data Mining is a term used to describe the discovery of knowledge in the database. Data mining is a process that uses statistical, mathematical, artificial intelligence, and machine learning techniques to extract and identify useful information and related knowledge from large databases.

Data mining is not an entirely new field. One of the difficulties in defining data mining is the fact that data mining inherits many aspects and techniques from previously established scientific fields.

Correlation

Correlation is one of the statistical techniques used to find the relationship between two or more variables that are quantitative. The relationship between the two variables occurs because of a causal relationship or it can also be due to chance.

Meanwhile, according to Teguh Wahyono (2012, p, 115) correlation can be said as a reciprocal relationship or cause and effect between two events.

high frequency pattern analysis

This stage is looking for a combination of items that meet the minimum requirements of the value of support in databases. Support value an item is obtained by the formula following:

$$Support(A) = \frac{Transaction\ amount\ contains\ A}{Number\ of\ trancaction}$$

Meanwhile, the supoprt value of the 2 items is obtained from the following formula 2:

$$Support\ (A, B) = P(A \cap B)$$

$$Support(A, B) = \frac{\sum Transaction\ contains\ A\ and\ B}{\sum Transaction}$$

Establishment of association rules

After all the high frequency patterns are found, then we look for association rules tat meet the minimum requirments for confidence by calcualting the confidence of the associative rule A → B. The confidence value f the rule A → B is obtained from the following formua :

$$Confidence = P(B|A) = \frac{\sum Trasaction\ contains\ A\ and\ B}{Transaction\ contains\ A}$$

RESULTS AND DISCUSSION

The following is a description of the results of the analysis of the application of data mining using the a priori method for the correlation of the learning process, report cards and rankings at SMP N 2 Binjai, Langkat Regency.

Process flowchart a priori

The following flowchart is implemened with the rules that apply in the design. And can be seen the design with a structured design concept as follows:

Association Analysis

The basic association is divide into 2 stages, namely:

a. High-frequency pattern analysis in this stage looks for a combination of items that meet the minium requirments of the supoprt value in the database.

b. The formation of association rules after all frequencies are found, then search for association ruls that meet the confidence requirements by calculating the associative.

after that, the representation can be done as follows:

Table.1 Learning Process Data Representation in Student Database

No.	Learning Process	Code
1	Daring	Dr
2	Face to Face	Tm

Table.2 Representastion of Report Values in Student Database

No.	Category	predicate	Code
1	80 – 100	Very Good	Si
2	71 – 80	Good	Bi
3	51 – 70	Enough	Ck
4	0 - 50	Not Good	Kr

Table.3 Rank Representation in Student Database

No.	Category	Predicate	Code
1	1 – 3	Amazing	Lb
2	4 – 10	Very Good	Sb
3	11 – 20	Good	Bk

4	>=21	Not Good	Ck

Table.4 Student Data

Alternative Name	Learning process	Report Card Value	Ranking
A1	FACE TO FACE	Very Good	Good
A2	DARING	Very Good	Very Good
B1	FACE TO FACE	Very Good	Amazing
B2	DARING	Very Good	Amazing
C1	FACE TO FACE	Very Good	Amazing
C2	DARING	Very Good	Very Good
D1	FACE TO FACE	Very Good	Very Good
D2	DARING	Very Good	Very Good
E1	FACE TO FACE	Very Good	Very Good
E2	DARING	Very Good	Very Good
F1	FACE TO FACE	Good	Good
F2	DARING	Good	Good
G1	FACE TO FACE	Very Good	Very Good
G2	DARING	Very Good	Very Good
H1	FACE TO FACE	Very Good	Amazing
H2	DARING	Very Good	Amazing
I1	FACE TO FACE	Very Good	Very Good
I2	DARING	Very Good	Very Good
J1	FACE TO FACE	Very Good	Very Good
J2	DARING	Very Good	Amazing

The data in table.1 in the student database can be represented as table.2.

Table.5 Student Data Tabular Format

No	Tm	Dr	Sa	Ba	Cu	Ku	Lb	Sb	Bk	Cu
1	1	0	1	0	0	0	0	0	1	0
2	0	1	1	0	0	0	0	1	0	0
3	1	0	1	0	0	0	1	0	0	0
4	0	1	1	0	0	0	1	0	0	0
5	1	0	1	0	0	0	1	0	0	0
6	0	1	1	0	0	0	0	1	0	0
7	1	0	1	0	0	0	0	1	0	0
8	0	1	1	0	0	0	0	1	0	0
9	1	0	1	0	0	0	0	1	0	0
10	0	1	1	0	0	0	0	1	0	0
11	1	0	0	1	0	0	0	0	1	0
12	0	1	0	1	0	0	0	0	1	0
13	1	0	1	0	0	0	0	1	0	0
14	0	1	1	0	0	0	0	1	0	0
15	1	0	1	0	0	0	1	0	0	0

16	0	1	1	0	0	0	1	0	0	0
17	1	0	1	0	0	0	0	1	0	0
18	0	1	1	0	0	0	0	1	0	0
19	1	0	1	0	0	0	0	1	0	0
20	0	1	1	0	0	0	1	0	0	0
∑	10	10	18	2	0	0	6	11	3	0

Rule for 3 itemset consists of 7 rules which are impelemted in the following table.

Table.6 3-Itemset association rules

If antecedent then consequent	Support	Confidence
If Tm and Sa and Lb	3/20*100% = 0,15%	3/10*100% = 30%
If Tm and Sa and Sb	5/20*100% = 0,25%	5/10*100% = 50%
If Tm and Sa and Bk	1/20*100% = 0,05%	1/10*100% = 10%
If Tm and Ba and Bk	1/20*100% = 0,05%	1/10*100% = 10%
If Dr and Sa and Lb	3/20*100% = 0,15%	3/10*100% = 30%
If Dr and Sa and Sb	6/20*100% = 0,3%	6/10*100% = 60%
If Dr and Ba and Bk	1/20*100% = 0,05%	1/10*100% = 10%

And after the support and confidence values are obtained for each cindidate, the multiplication between support and confidence is carried out.

Tabel.7 3-Itemset final association rules

If antecedent then consequent	Support	Confidence	S*C
If Tm and Sa and Lb	0,15%	30%	4,5%
If Tm and Sa and Sb	0,25%	50%	12,5%
If Tm and Sa and Bk	0,05%	10%	0,5%
If Tm and Ba and Bk	0,05%	10%	0,5%
If Dr and Sa and Lb	0,15%	30%	4,5%
If Dr and Sa and Sb	0,3%	60%	18%
If Dr and Ba and Bk	0,05%	10%	0,5%

After obtaining the Support and Confidence values for each candidate, mulplication between Support and Confidence is carried out, then the result of the multiplication is 18% and becomes the bes rule

Tabel.8 *Best Rule*

| *If antecedent then consequent*	*Support*	*Confidence*	*S*C*
If Dr and Sa and Sb	0,3%	60%	18%

Then what is obtained as a rule is:
If the Online Learning Process ad Report Values are Very Good then hhe student's ranking is very good. With support = 0,3% and confidence = 60% with a value of C*S = 18%

Implementation
       At this stage, the impelentation of the software is a continuation of the design stage so this implementation must be based on the design that has been carried out previously and testing is carried out t see the result of each running process that can produce the expeted output.
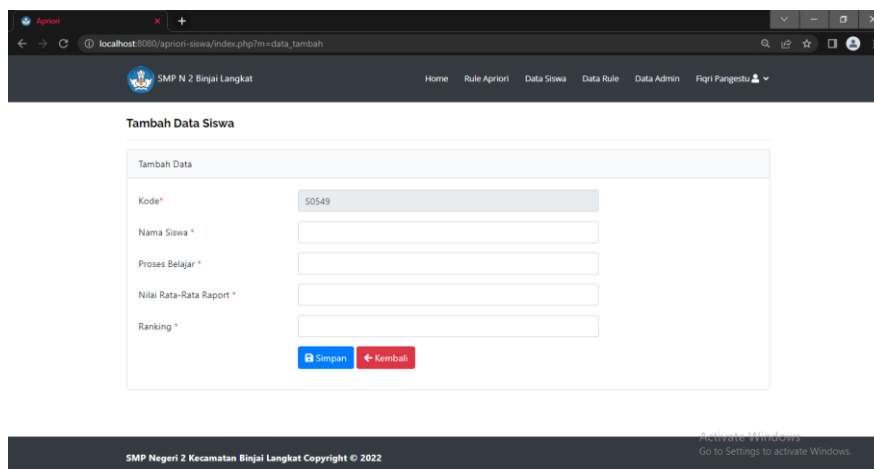
Input rules page
       This page seres to input be rules that have been created from weka, here is a picture of the rule input page.
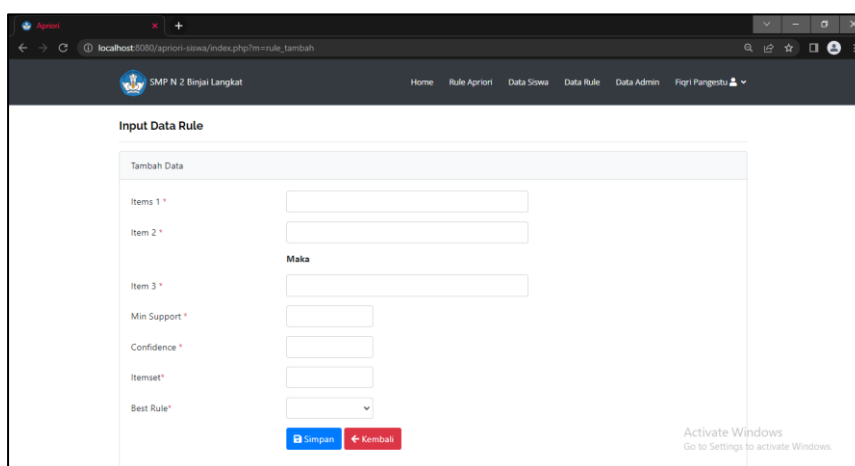
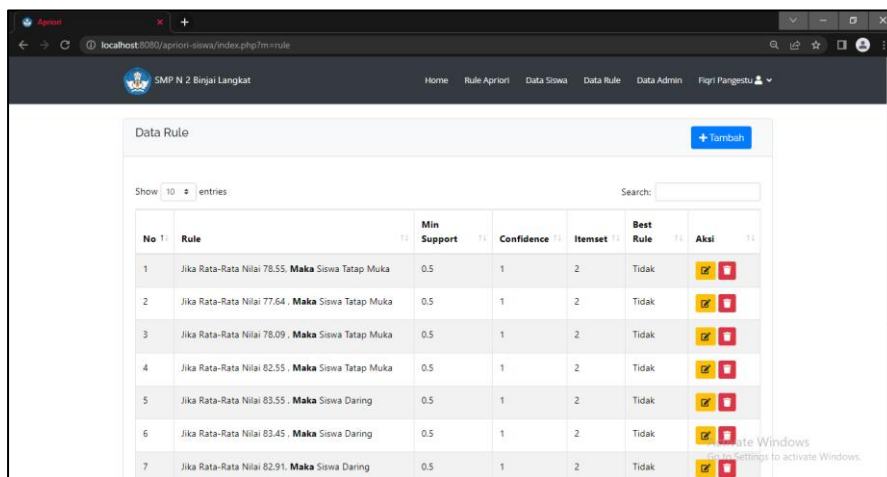Picture 1. Input Rules Page

Student data input page
       This page can be accessed by the admin to input student data into the database.

Picture 2. Studen data input page

Page showing rule

This page functions to view the rules that have been inputted by the admin and the admin can edit and delete as shown below:

Picture 3. Page showing rule

Student data display page

This page displays student data information that has been iinput by the admin and on this page the data can be edited by the admin. The image of the student data display page can be seen as below:
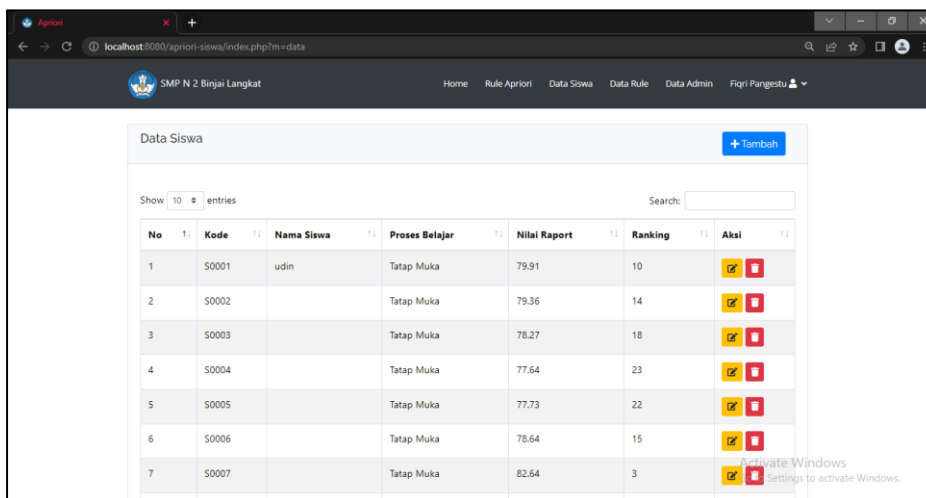
Picture 4. Student data display page

CONCLUSION

Based on the results of research and discussion that has been carried out, some comclusions are obtained as follows, from the data testing that has been done, the appliction of data mining association rules using the a priori algorihtm can be used to find a combination of existing data on student data at the state 2 Binjai junior high school, langkat district. Fro the result of testing

342

student data in the 2019/2020 school yer, 548 data were fund, 36 rules were formed and the highest Best Rule was obtained with a maximum support value of 0,5% and a confidence value of 1%.
.

REFERENCES

Putria, H., Maulana, H.L., Uswayun, S.D. 2020. "Analysis of the Onle Learnig Process during the Covid-19 Pandemic or Elementary School Teacher. Jurnal of Education.

Shafi'i. A., Marfiyanto. T., Rodiah. K.S. 2018. "Study of Student Achiement in Various Aspects and Influencing Factors. Jurnal of Educational Communication.

Novriyeni. Sihombing. A. 2016. "Data Mining The Correlation of Parents Work Against the Cumulative Achievment Index of STMIK kaputama Binjai's Graduates.

Kusrini, Lutfi. T.E. 2009. "Data Mining Algorithm. STMIK AMIKUM, Yogyakata.