Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/ 10.33395/sinkron.v7i3.11599
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Decision Model for Unplanned ICU Transfer in a Hospital with Association Rule Learning

Nanda Lestari[1)] , Sawaluddin[2)], and Parapat Gultom[3)]
Universitas Sumatera Utara, Medan, Indonesia
nda.lestari86@gmail.com

Abstract. The initial decision after treatment in the hospital emergency room is very important because apart from being an indicator of the quality of care for emergency room practitioners, it is also needed to achieve health goals, namely improving the quality of critical care and preventing death. The initial decision was also for an unplanned ICU transfer. Unplanned ICU transfer is the transfer of patients who originally came from the ER (Emergency Room), then to the Inpatient Room (having been treated for 24-48 hours), then to the ICU. Many studies have been carried out to predict the initial decision of unplanned ICU transfer using univariate analysis, logistic regression analysis, and association rules. The association rule algorithm generates rules between patient diagnosis features that form a decision model for unplanned ICU transfers, so it is essential to get an association rule algorithm that is more efficient in generating rules. In this study, we compare two association rule algorithms to get a more efficient algorithm; then, the rules are used to form a decision model for unplanned ICU transfers. The study results obtained that the Apriori algorithm requires a completion time of 3 ms and the FP-Growth algorithm requires a completion time of 31 ms. Hence, the FP-Growth algorithm is 28 ms more efficient than the Apriori algorithm, while the resulting rule generation is the same number of 67 rules. Only 11 rules meet the minsupp and minconf threshold and include the set of Class Association Rules (CAR), which are used to form a decision model for unplanned ICU transfers with binary integer programming.

INTRODUCTION

In a hospital, an emergency unit is a part of the service that provides initial treatment for patients who suffer from illness or injury that can threaten their survival. The emergency department functions to receive, stabilize and manage patients who need emergency treatment immediately. Prior to treating patients in the ER, doctors on duty in the ER (Emergency Room) perform triage, namely the process of sorting patients based on the severity of the injury or illness suffered by giving a color code, namely red for the first priority (patients with serious life-threatening injuries), yellow for the second priority (patients not present). immediate life threat), green for third priority (patients with minimal injury), and black for zero priority (patients died) (Bapoje et al., 2011).

Unplanned ICU transfer is the transfer of a patient initially from the ER then to the Inpatient Room (having been treated for 24-48 hours) and then to the ICU. These unplanned ICU transfers usually lead to higher patient mortality rates compared to patients admitted directly to the ICU from the ER (planned ICU transfer). As a result, the initial decision to transfer to the ICU from the ED within 24-48 hours has been considered very important as an indicator of the quality of care for ED practitioners and is urgently needed to achieve the health goals of improving the quality of critical care and preventing death (Han et al., 2004).

*Corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/ 10.33395/sinkron.v7i3.11599
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Research that has been carried out in predicting the initial decision of ICU transfer from the ED using univariate analysis and logistic regression analysis aimed to identify the main risk factors (features) associated with high-risk patients with a critical illness (Stiell et al., 2013). The results of logistic regression (LR) are widely used in medical diagnostic research which proves useful for predicting patients but fails for different patient conditions, where in most cases, the risk factors (diagnostic features) that cause unplanned ICU transfer do not apply to all patient conditions so that an association rule-based deep learning system optimization approach is used that can find differences in patterns or rules between risk factors (diagnostic features) that are appropriate for high-risk patients (Chou et al., 2020). The association rule used is the Apriori Algorithm and then to form a decision model using the Mix-Integer Programming approach (Zhu, 2019). The Apriori algorithm is less efficient in generating association rules because it has to scan the data repeatedly (Chou et al., 2020). So, in this study, the author aims to find a more efficient algorithm in generating association rules between patient diagnosis features, which then the association rule is used to form a decision model.

METHODS

The research method used in this study is a literature study by comparing two association rule learning algorithms, namely the Apriori Algorithm and the FP-Tree Algorithm which aims to find an efficient algorithm in generating Association Rules between diagnostic features of the FP-Growth Algorithm. Then Generate a class association rule (CAR) set by pruning the Association rule between patient diagnosis features generated based on the Associative Classification (AC). The binary integer programming mathematical model was used as a decision model designed to identify patients who were transferred to the ICU unexpectedly or to an unplanned ICU transfer.

RESULT AND DISCUSSION

Association Rule is a form of implication of *if - then* then denoted $X \rightarrow Y$, where X and Y are disjoint itemset, namely $X \cap Y = \emptyset$. The strength of the association rule can be measured in terms of support and confidence. Support indicates how often a rule is generated to be applied to a given dataset. While confidence shows how often item Y appears in transactions containing X. support and confidence are defined as follows (Han et al., 2004):

$$Support,\ s(X) = \frac{\sigma(X)}{N} \tag{1}$$

$$Support,\ s(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N} \tag{2}$$

$$Confidence,\ c(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)} \tag{3}$$

Comparison of Apriori Algorithm and FP-Growth Algorithm

Illustration of the data used in this study, for example there is data on patients who enter the ER in a hospital, the data contains the diagnostic features possessed by the patient and the patient class, class 1 for patients who experience unplanned ICU transfers and class 0 for patients who are not unplanned ICU transfer. For example, the minimum support (θ_s)= 25% dan minimum confidence (θ_c)= 50%. The data is processed using the Apriori Algorithm (Agrawal et al., 1993) and the FP-Growth Algorithm (Borgelt, 2005).

TABLE 1. Patient data in the form of diagnostic features and class labels

TID	Item	Class
1	$I2, I3, I6$	1
2	$I3, I4, I5$	0
3	$I4, I6$	0
4	$I4, I5, I6$	0

*Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/ 10.33395/sinkron.v7i3.11599
e-ISSN : 2541-2019
p-ISSN : 2541-044X

5	$I2, I3, I4, I5, I7$	1
6	$I4, I6$	0
7	$I2, I3$	1
8	$I2, I3, I4$	1

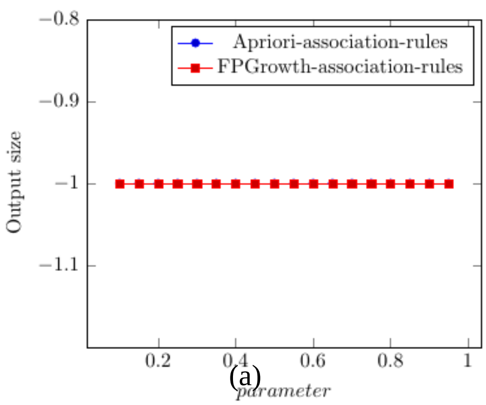
TABLE 2. Illustration of patient diagnosis feature data

No.	Symbol	Diagnostic features
1	$I2$	Coronary heart disease
2	$I3$	Cerebral vascular rupture disease
3	$I4$	Cancer
4	$I5$	Kidney failure
5	$I6$	Respiratory disorders
6	$I7$	Liver dysfunction

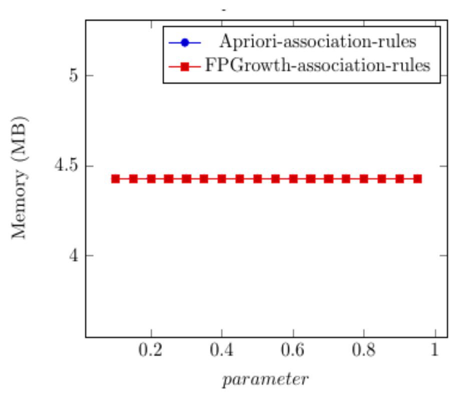
TABLE 3. Patient class illustration

No.	Symbol	Class Description
1	1	Patients undergoing unplanned ICU transfer
2	0	Patients who did not have an unplanned ICU Transfer

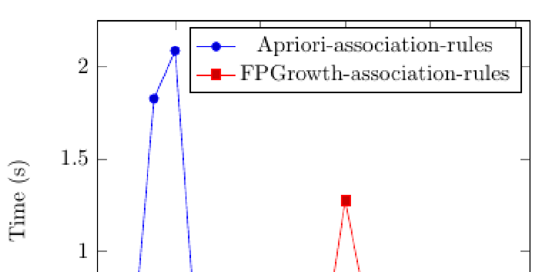
Data processing with the Apriori Algorithm and FP-Growth Algorithm is carried out using SPMF v2.54 (Java) software with a core i7 CPU operating system, 8 Gb RAM and 1 Tb SSD. The calculation results are as follows:



(a)



(b)



*Corresponding autho

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/ 10.33395/sinkron.v7i3.11599
e-ISSN : 2541-2019
p-ISSN : 2541-044X

(c)

FIGURE 1. Comparison Apriori algorithm and FP-Growth algorithm for output size (a), memory usage (b), and time (c)

TABLE 4. Comparison results of Apriori algorithm and FP-Growth algorithm

Comparison	Algoritma Apriori	Algoritma FP-Growth
Total time	31 ms	3 ms
Maximum memory usage	8.942672729492188 mb	5.401176452636719 mb
Number of association rules generated	67	67

In processing the data above, it can be seen that the Apriori Algorithm takes 3 ms with a total association rule generated 67 rules and the FP-Growth Algorithm takes 31 ms with a total association rule generated 67 rules. So the FP-Growth Algorithm is 28 ms more efficient than the Apriori Algorithm. The resulting rule generation is 67 rules.

Class Association Rule set (CARs)

Association rule with a special case called the Associative Classification (AC), namely in the rule's right-hand side (consequent) only class attributes are considered, for example in the X→Y rule, Y must be a class attribute (Witten & Frank, 2002). Generating a class association rule (CAR) set by pruning the Association rule between patient diagnostic features generated based on the Associative Classification (AC) so that on the rule's right-hand side (consequent) there is only class 1 (unplanned ICU transfer) or class 0 (unplanned non-ICU transfer). The rules obtained after pruning are as follows:

TABLE 5. Class Association Rule (CAR)

No.	Rule	Support	Confidence
1	*I*2 ==> *I*1	50%	100%
2	*I*3 ==> *I*1	50%	80%
3	*I*2 *I*3 ==> *I*1	50%	100%
4	*I*2 *I*4 ==> *I*1	25%	100%
5	*I*3 *I*4 ==> *I*1	25%	66%
6	*I*2 *I*3 *I*4 ==> *I*1	25%	100%
7	*I*4 ==> *I*0	50%	66%
8	*I*5 ==> *I*0	25%	66%
9	*I*6 ==> *I*0	37,5%	75%
10	*I*4 *I*5 ==> *I*0	25%	66%
11	*I*4 *I*6 ==> *I*0	37,5%	100%

*Corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/ 10.33395/sinkron.v7i3.11599
e-ISSN : 2541-2019
p-ISSN : 2541-044X

From the table above, it can be seen that there are 11 rules which are a set of Class Association Rules (CAR)are $\{\{R_1: I2 \rightarrow 1\}, \{R_2: I3 \rightarrow 1\}, \{R_3: I2, I3 \rightarrow 1\}, \{R_4: I2, I4 \rightarrow 1\}, \{R_5: I3, I4 \rightarrow 1\}, \{R_6: I2, I3, I4 \rightarrow 1\}, \{R_7: I4 \rightarrow 0\}, \{R_8: I5 \rightarrow 0\}, \{R_9: I6 \rightarrow 0\}, \{R_{10}: I4, I5 \rightarrow 0\}, \{R_{11}: I4, I6 \rightarrow 0\}\}$.

$R_1: I2 \rightarrow 1$, *support* 50% *dan confidence* 100% means that patients with features of coronary heart disease diagnosis (*I2*) will be included in the unplanned ICU transfer class (*1*), support 50% means that in the database 50% of patients with features of coronary heart disease diagnosis are included in the class of unplanned ICU transfers and 100% confidence means that 100% of patients with the features of a cancer diagnosis are concurrently admitted to an unplanned ICU transfer class.

Mathematical Model

The decision model was designed as a binary integer programming to identify patients who were transferred to the ICU unexpectedly or to an unplanned ICU transfer. Patients, diagnostic features and rules are used as variables in the model.

TABLE 6. Matrix **a** (Patient data in the form of diagnostic features and class labels)

TID	Class	$I2$	$I3$	$I4$	$I5$	$I6$
T_1	1	1	1	0	0	1
T_2	1	1	1	1	1	0
T_3	1	1	1	0	0	0
T_4	1	1	1	1	0	0
N_1	0	0	1	1	1	0
N_2	0	0	0	1	0	1
N_3	0	0	0	1	1	1
N_4	0	0	0	1	0	1

TABLE 7. Matrix **b** (representation of diagnostic features with rule)

fitur	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}
$I2$	1	0	0	1	0	1	0	0	0	0	0
$I3$	0	1	0	1	1	1	0	0	0	0	0
$I4$	0	0	1	0	1	1	1	0	0	1	1
$I5$	0	0	0	0	0	0	0	1	0	1	0
$I6$	0	0	0	0	0	0	0	0	1	0	1
size	1	1	1	2	2	3	1	1	1	2	2
$\#T$ covered	4	4	4	4	2	2	0	0	0	0	0
$\#N$ covered	0	0	0	0	0	0	2	2	3	2	3

TABLE 8. Matrix **c** (patient representation with rule)

TID	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}
T_1	1	1	1	0	0	0	0	0	0	0	0
T_2	1	1	1	1	1	1	0	0	0	0	0
T_3	1	1	1	0	0	0	0	0	0	0	0

*Corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/ 10.33395/sinkron.v7i3.11599
e-ISSN : 2541-2019
p-ISSN : 2541-044X

T_4	1	1	1	1	1	1	0	0	0	0	0
N_1	0	0	0	0	0	0	1	1	0	1	0
N_2	0	0	0	0	0	0	1	0	1	0	1
N_3	0	0	0	0	0	0	1	1	1	1	1
N_4	0	0	0	0	0	0	1	0	1	0	1

Matrix **a** is an $n \times m$ binary matrix. Where $a_{ij} = 1$ if patient *i* has diagnostic feature *j*. Patients were grouped into 2 classes, namely positive (unplanned ICU transfer) and negative (unplanned non-ICU transfer) so that $\mathbf{a}^+ \cup \mathbf{a}^- = \mathbf{a}$ dan $\mathbf{a}^+ \cap \mathbf{a}^- = \emptyset$.

Matrix **b** is an $m \times p$ binary matrix. Where $b_{jk} = 1$ if diagnosis feature *j* is included in rule *k*.

Matrix **c** is an $n \times p$ binary matrix which is the result of $\mathbf{c} = \mathbf{a} \otimes \mathbf{b}$ where $c_{ik} = 1$ if patient *i* who has diagnostic features is covered by rule *k*. Similar to matrix a, patients were grouped into 2 classes, namely positive (unplanned ICU transfer) and negative (unplanned non-ICU transfer) so that $\mathbf{c}^+ \cup \mathbf{c}^- = \mathbf{c}$ dan $\mathbf{c}^+ \cap \mathbf{c} = \emptyset$.

The decision variables used in the model are as follows:

$$x_i = \begin{cases} 1, & \text{if patient } i \text{ can be covered in model} \\ 0, & \text{if patient } i \text{ cannot be covered in model} \end{cases}$$

$$y_j = \begin{cases} 1, & \text{if diagnostic feature } j \text{ is used in the model} \\ 0, & \text{if diagnostic feature } j \text{ is not used in the model} \end{cases}$$

$$z_k = \begin{cases} 1, & \text{if rule k is used in the model} \\ 0, & \text{if rule k is not used in the model} \end{cases}$$

The decision model is formulated as Binary Integer Programming as follows :

$$\min \sum_{j=1}^{m} y_j + \sum_{k=1}^{p} z_k + \sum_{i \in |I|^-} x_i - \sum_{i \in |I|^+} x_i \quad (4)$$

Subject to:

$$\sum_{k \in K} c_{ik}^+ z_k \geq x_i \qquad \forall\, i \in I^+ \quad (5)$$

$$\sum_{k \in K} c_{ik}^- z_k \leq x_i \qquad \forall\, i \in I^- \quad (6)$$

$$\sum_{k \in K} b_{jk} z_k \leq y_j \qquad \forall\, j \in J \quad (7)$$

$$x_i, y_j, z_k \in \{0,1\} \quad (8)$$

The objective function in equation (4) is to minimize the number of diagnostic features and the number of rules included in the decision model while ensuring that the selected rule minimizes negative coverage (unplanned non-ICU transfers) and maximizes positive coverage (unplanned ICU transfers) of patient.

The constraint function in equation (5) ensures that the unplanned transfer of patient *i* is covered by at least one rule. Equation (6) shows an indication if the non-transfer of patient *i* is covered by the selected rule. Equation (7) shows an indication if the diagnostic feature *j* is used by the selected rule. While equation (8) shows that x_i, y_j, z_k adalah are binary decision variables.

CONCLUSION

Comparing two association rule algorithms that are part of deep learning optimization, namely the Apriori Algorithm and the Fp-Growth Algorithm to find efficiency in the formation of rules between patient diagnostic features, where these rules will be used to form a decision model for unplanned ICU

*Corresponding author

transfers with binary integer programming. The Apriori Algorithm takes 3 ms and the FP-Growth Algorithm takes 31 ms so that the FP-Growth Algorithm is 28 ms more efficient than the Apriori Algorithm, while the resulting rule generation is the same number of 67 rules. Only 11 rules meet the minsupp and minconf thresholds and include the set of Class Association Rules (CAR) which are used to form a decision model for unplanned ICU transfers with binary integer programming.

ACKNOWLEDGMENTS

The author would like to thank the Universitas Sumatera Utara for supporting this research.

REFERENCES

Agrawal, R., Faloutsos, C. & Swami, A. (1993). Efficient similarity search in sequence databases. *International Conference on Foundations of Data Organization and Algorithms*, 69–84.

Bapoje, S. R., Gaudiani, J. L., Narayanan, V. & Albert, R. K. (2011). Unplanned transfers to a medical intensive care unit: causes and relationship to preventable errors in care. *Journal of Hospital Medicine*, *6*(2), 68–72.

Borgelt, C. (2005). An Implementation of the FP-growth Algorithm. *Proceedings of the 1st International Workshop on Open Source Data Mining: Frequent Pattern Mining Implementations*, 1–5.

Chou, C.-A., Cao, Q., Weng, S.-J. & Tsai, C.-H. (2020). Mixed-integer optimization approach to learning association rules for unplanned ICU transfer. *Artificial Intelligence in Medicine*, *103*, 101806.

Han, J., Pei, J., Yin, Y. & Mao, R. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data Mining and Knowledge Discovery*, *8*(1), 53–87.

Stiell, I. G., Clement, C. M., Brison, R. J., Rowe, B. H., Borgundvaag, B., Aaron, S. D., Lang, E., Calder, L. A., Perry, J. J. & Forster, A. J. (2013). A risk scoring system to identify emergency department patients with heart failure at high risk for serious adverse events. *Academic Emergency Medicine*, *20*(1), 17–26.

Witten, I. H. & Frank, E. (2002). Data mining: practical machine learning tools and techniques with Java implementations. *Acm Sigmod Record*, *31*(1), 76–77.

Zhu, S. (2019). Research on data mining of education technical ability training for physical education students based on Apriori algorithm. *Cluster Computing*, *22*(6), 14811–14818.

*Corresponding author