

# Finding the Relationship of Subjects for Academic Guidance: A Case Study of Logistics Engineering and Management Program, Chiang Rai Rajabhat University

Panee Sujinda <sup>\*</sup>, Nilubon Phandee <sup>†</sup>, Kittakorn Sriwana, Thanawut Thanavanich,  
Atikhom Siri, Kamol Boonlom, Seksan Winyangkul  
Center for Microwave and Robotic Technology,

School of Computer and Information Technology, Chiang Rai Rajabhat University, Thailand

**Abstract:** *The problems of students studying Logistics Engineering and Management Program, Chiang Rai Rajabhat University, are determined to improve the efficiency of studying the core courses which are considered having a significant relationship. The relationship of subjects is applied as guidance for the student achievement to earn higher academic record. Finding the relationship of subjects, a case study of Logistics Engineering and Management Program, Chiang Rai Rajabhat University, is analyzed by applying two major approaches of data mining consisting of data classification and association rule to examine 12 core subjects of the curriculum. The academic records of seniors are analyzed to find the relationship of subjects used as guidance and a direction to improve the quality of education of the next enrolled students. The findings show average accuracy of data 76.60 percent.*

**Keywords:** *Academic guidance, data mining, data classification*

## 1. Introduction

Data mining is a technique used to manage a number of dataset to categorize and find out the relationship of data [1]. Nowadays, data mining is applied in various works including business which is used for the management decision, science, medical, economic, social as well as educational activities. For education, making a decision of a student to choose a suitable program is significant and necessary. Consequently, an application of information technology to help students making a right decision [2] is considered a vital approach used as guideline to achieve their goals and knowledge exchange.

The behavior or information analysis of academic records of university students is therefore beneficial for educational direction for the best results. The findings are used for an improvement of learning as required or planned for their future or further study.

This research aims on applying data mining for an analysis of academic guidance to improve the quality of education [3]. This approach is data examination to select potential data for analyzing with classification method. The method sorts out the information and the Decision Tree approach is then employed to model and setup the conditions as required by users or the tree.

## 2. Related Theories

The research applies the technique of data mining to analyze and classify data with Rapid Miner program for an assessment of academic guidance efficiency.

### 2.1 Academic Guidance Background

An academic guidance is a process used with students to comprehend and realize in order to settle and to support in their educational improvement. The guidance is hence an important factor to support, encourage and assist students in processing information and way of thinking as well as the ability to apply their knowledge to resolve and improve their lives and education.

## 2.2 Data Mining

Data mining is used as a tool to analyze for finding out the relationship and type of datasets [4]. This technique consists of 3 major parts which are Association Rule, Data Processing and Data Clustering.

## 2.3 Data Classification

Data Classification, an approach used for sorting out data into parts, combines similar data or related information together and distinctive information are separated correspondingly. In this research, algorithms are utilized for data mining which include the following.

Decision Tree, an approach used in data mining for decision making, is similar to an upside down tree. The first node of the tree is called Root Node. Each node reveals Attribute and tree branch represents findings, and Leaf Node represents class [7].

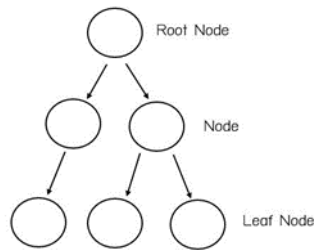


Fig. 1: Decision Tree Structure

## 2.4 Association Rule

Association Rule is an essential technique of Data Mining which is applicable in various fields. The process works to figure out data relationship of large information to be analyzed or predicted. In some cases, information is from customer purchasing or “Market Basket Analysis” [8] which is examined from data collected. The findings become the resolution resulted from the Association Rule. The rule is represented in form of cause to result with the data of the Market Basket Analysis. The data including nominal and Ordinal are analyzed in this process.

## 3. Methodology

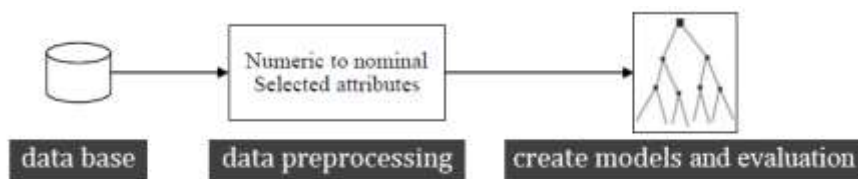


Fig. 2: Program Process

### 3.1 Decision Tree Process

In Fig. 2, it describes the structure of program used in Rapid Miner divided into 4 parts consisting of database, data selection, data model and data processing. 1) Database combines information used for analyzing and kept in Microsoft Excel 2) Data selection is utilized for classifying information for analysis 3) Data model creates the tree structure in Rapid Miner 4) Data processing is an evaluation or measurement of the model efficiency.

#### 3.1.1 Data Preparation

The database used in this research is collected from student records of Logistics Engineering and Management Program of Chiang Rai Rajabhat University graduated in the academic year of 2012. The data analyzed and used as guidance for students composes of 12 core subject results of Logistics Engineering and

Management Program which every student in the curriculum takes and 94 students are researched with following subjects shown in Table 1.

TABLE I: 12 Core Subjects of Logistics Engineering and Management Program

ID	Subjects	Year	Semester
LEM2401	Industrial Plant Design and Layout	2	2
LEM2402	Engineering Economics	2	2
LEM2403	Industrial Manufacturing Process	2	2
LEM2405	Logistics and Supply Chain Management	2	2
LEM2404	Cost Analysis and Budgeting	2	3
LEM3403	Work Study	3	1
LEM3404	Operation Research	3	1
LEM3405	Planning and Production Controlling	3	1
LEM3417	Warehouse and Inventory Management	3	1
LEM3407	Computer Simulation in Industrial Work	3	1
LEM3406	Quality Control	3	2
LEM4402	Technology and Information Systems for Logistics	4	1

### 3.1.2 Data Processing

Data collected is filled out as in Table 2 with 111 Attributes and 94 records arranging from program outline and setting Technology and Information Systems for Logistics (LEM4402) – the subject which applied knowledge of other subjects in this subject, as Leaf Node (Destination subject). Grading 2.5, 3, 3.5 and 4 as “Good”, and 2, 1.5, 1 and 0 as “Bad” is shown in Table 2.

TABLE II: demonstrates sample of subjects used for analysis with Technology and Information Systems for Logistics (LEM4402) set as Leaf Node (Destination subject).

LEM2401	LEM2402	LEM2403	...	LEM4402
2.5	3.5	3.5	...	Good
2	1	4	...	Bad
3.5	4	4	...	Good
2	2.5	3	...	Bad
3	4	3	...	Good
4	4	3.5	...	Bad

### 3.1.3 Creating Model and Evaluation

The processes are to create a model of Decision Tree in Rapid Miner program with data processed in Table 2 and conditions of Decision Tree. Maximal Depth value is set to 5 and Minimal Leaf Size is set to 2. Findings of experiment are shown in form of Decision Tree consequently.

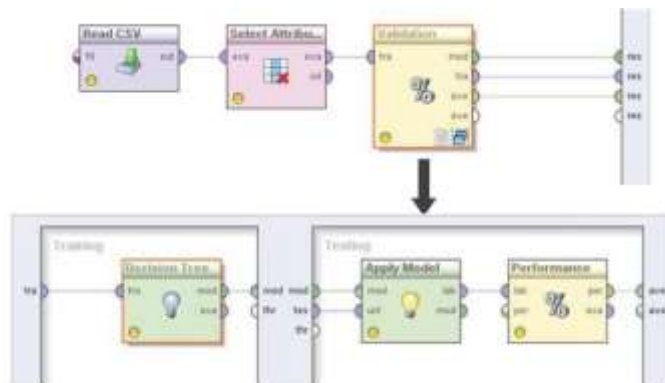


Fig. 3: Decision Tree Model Creating

### 3.2 Association Rule Structure

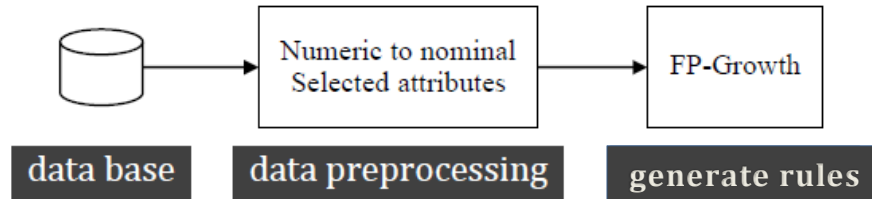


Fig. 4: Association Rule Process

As shown in Fig. 4, Association Rule is divided into 3 parts consisting of database, data processing and Association Rule.

#### 3.2.1 Database

Database includes student records of Logistics Engineering Program of Chiang Rai Rajabhat University for 117 records of graduated students in 2012.

#### 3.2.2 Data Processing

The grades of Logistics Engineering Program are verified in Table 3; 2.5, 3, 3.5 and 4 are “Good”, and 2, 1.5, 1 and 0 are “Bad”.

TABLE III: Sample of subjects used in the research with “Good” and “Bad” evaluation

LEM2401	LEM2402	LEM2403	...	LEM4402
Good	Good	Good	...	Good
Bad	Bad	Good	...	Bad
Good	Good	Good	...	Good
Bad	Good	Good	...	Bad
Good	Good	Good	...	Good
Good	Good	Good	...	Bad

#### 3.2.3 Association Rule Process

Data processed is input to Rapid Miner program as in Fig. 5.

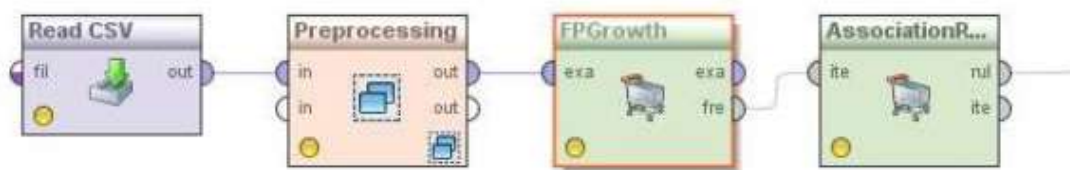


Fig. 5: Association Rule Process

## 4. Examination Results

According to the research to find out the relationship of subjects for academic guidance, a case study of Logistics Engineering and Management Program, Chiang Rai Rajabhat University, the results reveal the Decision Tree for 4 conditions as appeared in Figure 6, 7, 8 and 9.

### 1) Gain Ratio

The grades of core subjects input to Rapid Miner program reveal an accuracy of Decision Tree 72.34 percent.

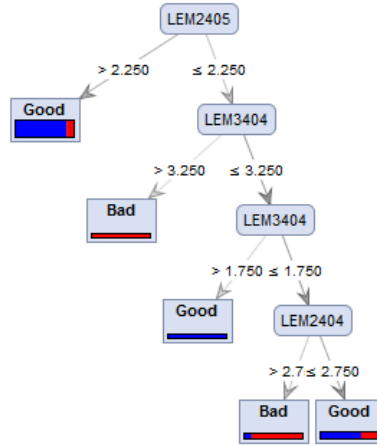


Fig. 6: Gain Ratio Decision Tree

2) **Information Gain**

The grades of core subjects input to Rapid Miner program reveal an accuracy of Decision Tree 68.09 percent.

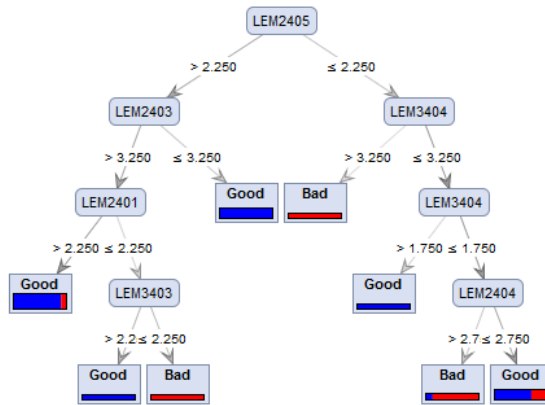


Fig. 7: Information Gain Decision Tree

3) **Gini Index**

The grades of core subjects input to Rapid Miner program reveal an accuracy of Decision Tree 69.15 percent.

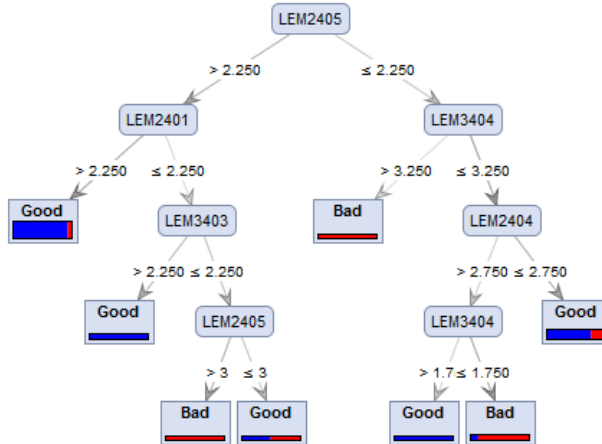


Fig. 8: Gini Index Decision Tree

**4) Accuracy**

The grades of core subjects input to Rapid Miner program reveal an accuracy of Decision Tree 76.60 percent.

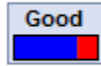


Fig. 9: Accuracy Decision Tree

The experiment is conducted in accordance with the 4 conditions of Decision Tree. The chart demonstrating the accuracy reveals the percentage that the Accuracy condition yields the most precise.

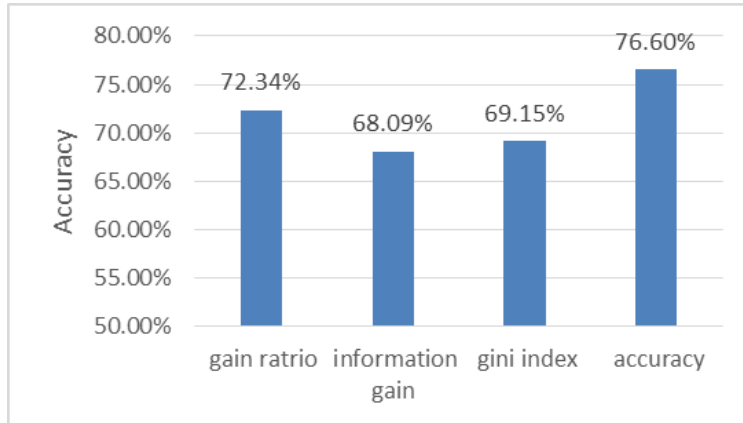


Fig. 10: Accuracy percentages of 4 conditions of Decision Tree

According to examination of Association Rule, it shows that the grade of Technology and Information Systems for Logistics (LEM4402) is good resulted from good grades of the Premises subjects as in Table 4.

TABLE IV: Association Rule of subjects

Rules	Premises	Conclusion	Support	Confidence
1	LEM2401 = Good, LEM2405 = Good, LEM3417 = Good, LEM3405 = Good	Good	0.202	0.730
2	LEM2405 = Good, LEM3403 = Good	Good	0.319	0.731
3	LEM2401 = Good, LEM2405 = Good, LEM3403 = Good	Good	0.265	0.735
4	LEM2403 = Good, LEM2405 = Good, LEM3417 = Good	Good	0.265	0.735

**5. Discussion**

The study of finding the relationship of subjects for academic guidance, a case study of Logistics Engineering and Management Program of Chiang Rai Rajabhat University, the Accuracy model is selected due to it is comprehensible and clearly identifies subject relation as well as reveals the most accuracy. There are 4 models of Decision Tree as following; 1) Gain Ratio resulting 72.34 percent of accuracy 2) Information Gain resulting 68.09 percent of accuracy 3) Gain Index resulting 69.15 percent of accuracy and 4) Accuracy resulting 76.60 percent of accuracy. The findings show, as in Table 4, that the good grades of courses including the Industrial Plant Design and Layout (LEM2401), Logistics and Supply Chain Management (LEM2405),

Warehouse and Inventory Management (LEM3417) and Planning and Production Controlling (LEM3405) result in good academic record (“Good”) of Technology and Information Systems for Logistics (LEM4402) subject.

## 6. Conclusion

According to the research, it is concluded that the Accuracy model reveals the most accurate. The Accuracy model classifies academic grades effectively more than the Gain Ratio, Information Gain and Gain Index. The model is beneficial for analyzing the efficiency, accuracy and finding out the relation of Binary Decision Tree which is widely used these days. This study hence proposes the approach to apply in the process of grade classification appropriately as the average efficiency is 76.60 percent.

## 7. References

- [1] M. Kantardzic, *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons, 2011.  
<http://dx.doi.org/10.1002/9781118029145>
- [2] D. T. Larose, *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons, 2014.  
<http://dx.doi.org/10.1002/9781118874059>
- [3] J. R. Quinlan, *C4. 5: programs for machine learning*. Elsevier, 2014.
- [4] C. Borgelt, “An implementation of the fp-growth algorithm,” in *Proceedings of the 1st international workshop on open source data mining: frequent pattern mining implementations*. ACM, 2005, pp. 1–5.  
<http://dx.doi.org/10.1145/1133905.1133907>
- [5] E. Harris, “Information gain versus gain ratio: A study of split method biases.” in *ISAIM*, 2002.
- [6] L. E. Raileanu and K. Stoffel, “Theoretical comparison between the gini index and information gain criteria,” *Annals of Mathematics and Artificial Intelligence*, vol. 41, no. 1, pp. 77–93, 2004.  
<http://dx.doi.org/10.1023/B:AMAI.0000018580.96245.c6>