Contents

Z.J. Pudlowski  Editorial  359

G.C. Kabouridis  The niche institute strategy - the way out of economic crisis for Greek higher  360
educational institutions: the case of the Technological Educational Institute
of Western Greece

C-H. Huang  A competency-based calculus teaching module for engineering students  365

A.H. Al-Arfaj & H.I. Abu-Mulaweh  A newly established college of engineering led by a non-engineer  371

Z. Ke  Engineering education knowledge management based on Topic Maps  376

H. Dong & X. Song  The theory and practice of linear regression  382

Z. Nedic, S. Gadzhanov & A. Nafalski  An industry-based project for teaching electrical power systems  388

P. Anastasiadis & G. Metaxas  Formulating the principles of an eco-city  394

L. Fürst & V. Matnić  Introductory programming course: motivating students with prior knowledge  400

X. Liu, Y. Sun & J. Wu  Reforming the teaching of mechanical design for industrial design students  406

J. Qiu  Evaluating students’ skills in a physical education course using grey system theory  412

A.P. Wibawa, A. Nafalski & Z. Nedic  Potential of an educational application of a language statistical translation system  418

A. Jurcius & A.V. Valiulis  Searching for residual stress measurement methods for structural steel components  424

W. Cao, Z. Liu & L. Zheng  Reforming the teaching of a single chip microprocessor course based on CDIO engineering education  428

K.W.H. Tsui & S.J. Thatcher  Aviation safety in the undergraduate curriculum: aviation students’ perceptions  434

D. Schott  Signal reconstruction - a project for students of electrical engineering  438

X. Lv  Peer assessment of perception and attitudes in public speaking English classes  445

S. Gadzhanov, A. Nafalski & Z. Nedic  Remote laboratories on motion control systems  450

S.J. Thatcher & K.W.H. Tsui  A comparison of the learning outcomes of international students from Asia and Australian students in an undergraduate aviation programme  456

J. Kang & G. Chen  Computer-aided instruction of mechanical engineering test technology and signal processing based on the use of MATLAB  460

W. Li  A study on matching the criteria of core courses with job requirements: a case study based on the urban rail specialisation in vocational colleges  465

X. Liu  An evaluation index system of undergraduate education quality based on an analytic hierarchy process model  470

357
S. Jia & C. Yang
Teaching software testing based on CDIO

Training engineering undergraduates in building environment and energy

D. W. Tai, J. L. Chen, R-C. Zhang & V. Tai
The review of current engineering and technology programme accreditation in Taiwan

H. Zhang, Y. Lu, J. Yang, G. Xu & Z. Sun
Applying Bu mode to the teaching of computer basics and the use of competitions in teaching

F. Jia
A multi-agent collaborative learning scheme for young university teachers based on reinforcement learning

B. Cai
Engineering education of electrical information courses: a comparison between Germany and China

H. Feng & Y. Li
An effective teaching method for a Linux course

L. Zhang, W. Zhang, N. Zhu & Y. Jia
Reforming cryptography education by teaching practical sessions

S. Kocijancic & A. Boonsongsrikul
A survey of student-centred approaches to engineering education - a case study concerning Slovenia and Thailand

J. Kasih, M. Ayub & S. Susanto
Predicting students' final passing results using the Apriori Algorithm

S. Li, C. Yang & Y. Zhang
A survey on the present situation of students' educational technology ability at a normal college

H. Huang, S. Chen & C. Huang
Computer teaching based on experiments using a virtualisation platform

P. Yao, C. Zhang, R. Yao, Y. Xiao & X. Wang
Analysing and improving teaching based on a random sample investigation

C. Xu
Classroom flipping as the basis of a teaching model for the course Mobile Application Development

H. Liu
Constructing a Web-based autonomous learning model for teaching English

J-R. Wang, Y-C. Chang, C-W. Li & H-C. Li
A study of competency indicators of Web-based micro-entrepreneurs in the consumer leisure and entertainment electronics retail industry, using AHP for analysing the weights

L. Wang, X. Jin, Y. Shi & L. Cao
An evaluation of the quality of university education based on multidimensional time series analysis

X. Zhang, Z. Xiong, Y. Ding & G. Wang
Reforms of an operating system course at a local university

J. Wang, C. Ma, S. Huang, L. Yu & R. Xia
Applying Whitehead's process philosophy to engineering-oriented practical education

D. Chen & Y. Wang
Reconstructing the training system of garment engineering education with the characteristics of art engineering

Q. Lei & Z. Wu
Research on engineering applications of Fresnel reflection functions

Y. Wang, H. Lui & D. Chen
Establishing a training mode for art engineering using art classification codes

X. Wu & L. Shang
Adding value to college business majors through a model education station built for practice sessions

Index of Authors

358
Predicting students' final passing results using the Apriori Algorithm

Julianti Kasih†, Mewati Ayub† & Sani Susanto‡

Maranatha Christian University, Bandung, Indonesia†
Parahyangan Catholic University, Bandung, Indonesia‡

ABSTRACT: The studies discussed form part of a programme, other aspects of which have been previously considered [1][2]. The ultimate objective is to facilitate a lecturer in helping students to predict their final passing results based on their performance in several subjects in the first four semesters of their study period. In previous research, this aim was achieved through two techniques: discriminant analysis [1] and the Classification and Regression Trees (CART) algorithm [2]. Those two techniques resulted in a diagramme-based relationship. In this research, a rule-based relationship of the form IF - THEN is introduced and subsequently applied using software based on the Apriori Algorithm.

INTRODUCTION

This research project continues a theme that was considered in previous studies [1][2], the objective of which was to facilitate a lecturer in helping students to predict their final passing results based on their performance in several subjects in the first four semesters during their study period. The arguments for why this kind of prediction is considered important were discussed in [2] and are rewritten in the Appendix. The passing results in the Indonesian education system are classified into three grades: Extraordinary (Cum Laude), Very Satisfactory and Satisfactory [3].

The research was undertaken in the same institution, the Faculty of Information Technology, a university in Bandung, West Java, Indonesia. For reasons of confidentiality, the full name of the institution has not been included. In the two previous works, it was demonstrated that discriminant analysis [1] and the Classification and Regression Trees (CART) algorithm [2] helped academic advisors in this faculty to predict the final passing results of a student based on his/her grade in some subjects during the first four semesters during their undergraduate programme. This sort of facility enables academic advisors to assist students in setting up their study plans each semester in order for them to perform to their full potential [1][2]. Moreover, this work aims at helping the academic advisors with a more practical way of predicting the final passing results of a student.

In this research, a data mining task called an association was employed. Association is performed through a technique called the Apriori Algorithm. This algorithm produces some rule-based relationships in the form IF- THEN statements. This kind of statement serves in a more ready to read feature compared to the territorial map or decision tree employed in the previous work in [1] and [2], respectively.

OVERVIEW OF BACKGROUND THEORY

David Hand et al define data mining as the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner [4].

The observational data or the data to be summarised are often called the training data. Data mining has six tasks: description, estimation, prediction, classification, clustering and association [4]. Association is based on affinity analysis, the study of attributes or characteristics that go together. One amongst several methods for affinity analysis is market basket analysis, which tries to discover associations among these attributes with the aim to discover association rules for quantifying the relationship between two or more attributes.

The association rule takes the form If antecedent, then consequent, which for reasons of simplicity often desires a single consequent [4]. The performance measures of this rule are the support, confidence, rule support, lift and deployability outcomes. With the assistance of the SPSS Clementine 10.1 software package, these measures are first defined by the
term instances. Instances define the number of records in the data set that match the antecedents. For example, given the association If purchase bread, then purchase cheese, the number of records in the training data that include the antecedent purchase bread are referred to as instances.

Support or antecedent support is the proportion of training data for which the antecedents are true. For example, if 50% of the training data includes the purchase of bread, then, the rule If purchase bread, then, purchase cheese will have an antecedent support of 50%. Support as defined here is the same as the instances but is represented as a percentage.

Rule support is the proportion of training data for which the entire rule, antecedents and consequent(s), are true. For example, if 20% of the training data contains both the purchase of bread and cheese, then, rule support for the rule If purchase bread, then, purchase cheese is 20%.

Confidence is the ratio of rule support to antecedent support. This indicates the proportion of training data with the specified antecedent(s) for which the consequent(s) is/are also true. For example, if 50% of the training data contains bread (indicating antecedent support) but only 20% contains both bread and cheese (indicating rule support), then, confidence for the rule If purchase bread, then, purchase cheese would be rule support/antecedent support or, in this case, 40%. Lift is the ratio of confidence for the rule to the prior probability of having the consequent. For example, if 10% of the entire population purchases bread, then, a rule that predicts whether people will purchase bread with 20% confidence will have a lift of 20/10 = 2. If another rule tells that people will purchase bread with 11% confidence, then, the rule has a lift of close to one (1), meaning that having the antecedent(s) does not make a lot of difference in the probability of having the consequent. In general, rules with lift far from 1 will be more interesting, than, rules with lift close to one (1).

Deployability is a measure of what percentage of the training data satisfies the conditions of the antecedent but does not satisfy the consequent. In product purchase terms, it basically means what percentage of the total customer base owns (or has purchased) the antecedent(s) but has not yet purchased the consequent.

EXPERIMENT: THE RESULT AND INTERPRETATION

The research was undertaken in the same institution, the Faculty of Information Technology, a university in Bandung, West Java, Indonesia. The rules were generated by SPSS Clementine 10.1 software.

As in the previous research programme, the academic transcripts from 146 alumni served as input or observational or training data, which were available from the authors [1][2]. From these data, the students' final passing results were determined by the final marks from the following eight (8) subjects: IF102 (Introduction to Computer Application), IF103 (Introduction to Information Technology), IF104 (Algorithms and Programming), IF105 (Basic Programming), IF106 (Informatics Mathematics) and IF202 (Linear Algebra and Matrices), IF 203 (Computer Network) and IF 205 (File System and Access). The final marks of these eight (8) subjects take the role of antecedents. The final marks of a subject were classified into five (5) groups, as follows: A (High Distinction), B (Distinction), C (Credit), D (Pass) and E (Fail) with some intermediates, such as B+ and C+.

The students' final passing results, as mentioned previously, were classified into three groups: 1 - Extraordinary (Cum Laude); 2 - Very Satisfactory; and 3 - Satisfactory. The students' final passing results take the role as the consequent.
The SPSS Clementine 10.1 model is described in Figure 1. In this figure, the icon on:

- the left side describes the input or observational or training data;
- the right side describes the algorithm employed, in this case, the Apriori Algorithm; and
- the top-middle side describes the output of the Apriori Algorithm;
- the bottom-middle side describes the type of the data, the final marks of the 8 (eight) subjects are the antecedents, while the passing result is the consequent.

The training data were saved in the form of a SPSS worksheet file, a section of which is represented in Figure 2.

The generated association rules are displayed in Figure 3. In generating association rules, SPSS Clementine 10.1 software gives the user the options to determine the minimum antecedent support, the minimum rule confidence and the maximum number of antecedents, which in this research were set at 20%, 80% and 5, respectively. Six association rules were generated. The authors will be able to generate more rules if they reduce the minimum antecedent support and the minimum rule confidence and vice versa.
One of the rules generated is IF IF103 = B and IF106 = A and IF102 = A, THEN Passing Result = Very Satisfactory. The antecedents of this rule are IF103 = B and IF106 = A and IF102 = A, and its consequent is Passing Result = Very Satisfactory. This rule has the following performances:

- Instances equal to 30, which means that out of 146 records in the training data sets, the number of records in the data set that match the antecedents is 30;
- Support or antecedent support is 20.548%, this value is due to the fact that the number of records in the 146 training data for which the antecedents are true is 30, or 20.548%;
- Rule support is 18.493%, this value is due to the fact that the number of records in the training data for which the antecedents and consequent(s) are true is 27 out of 146 records or 18.493%;
- Confidence is 90%, which means that the ratio of rule support, that is 18.493%, to antecedent support, that is 20.548% is 90%;
- Lift is 1.327, this value comes from the ratio of confidence for the rule (90%) to the prior probability of having the consequent Passing Result = Very Satisfactory (99 records out of 146, or 67.8%);
- Deployability is 2.055%, this value comes from a measure of what percentage of the training data satisfies the conditions of the antecedent IF103 = B and IF106 = A and IF102 = A, but does not satisfy the consequent Passing Result = Very Satisfactory. In this case is three (3) records out of 146 or 2.055%.

Care should be taken when applying the generated rules. Those rules do not express the causal relationship between the antecedent(s) and the consequent. As discussed in the background theory, association is based on affinity analysis, the study of attributes that go together. In the association rule context, the attributes are the antecedent(s) and consequent.

CONCLUSIONS AND SUGGESTION FOR FURTHER RESEARCH

This research demonstrates the ability of one of the data mining techniques that enable an academic advisor to help students predict their final passing results. Compared to the previous research [1,2], this research demonstrates that the prediction can be performed very practically, with neither graphs nor charts being required. The prediction is carried out by sentences of the form, IF antecedent, THEN consequent. Further research may take the form of the investigation of the prediction of the final passing results, which might be based on a multivariate statistics technique called logistic regression. The authors intend to perform this research in the not too distant future.

ACKNOWLEDGMENT

Sincerest thanks are conveyed by the authors to Mr. Radiant Victor Imbar, former Dean of the Faculty of Information Technology at Maranatha Christian University, for his support and providing the data from the academic transcripts of alumni.

REFERENCES

3. Government Regulation, the Republic of Indonesia, Nr 60 Year 1999 about Higher Education.

APPENDIX

In relation to the Introduction, the arguments for making the prediction of students’ final passing results as the main issue of this research are as follows:

- First, one of the important aims of higher education in the Republic of Indonesia is to prepare the academic participants (students) to become members of society with the academic and/or professional abilities to enable them to apply/develop/enrich the foundations of knowledge in the sciences, technology and the arts.
- Second, to achieve this aim, undergraduate students are assigned to several academic advisors (an informal translation of the Indonesian term *Dosen Wali*) throughout their years of higher educational studies.
- Third, academic advisors, who are lecturers, have as their main task the fostering of students’ academic and non-academic activities. With regard to students’ academic activities, one of the duties of the academic advisor is to help students in setting up their study plans for each semester.
- Fourth, setting up a study plan includes providing guidance for students regarding how many subjects, and which subjects, to undertake.
- Fifth, through this guidance, students are expected to obtain the best passing results at the end of their undergraduate study [1-3].