Decision Tree Modeling for an Intelligent Recommendation System Supporting SRM for Universities in Thailand

Kanokwan Kongsakun, Tuchtawan chanakul, and Chun Che Fung

Abstract—Over the years, many technological innovations have supported the advances and management of academic operations and processes. One of the initiatives designed to help students and staff is the Student Recommendation System, which aims to assist the academic development and achievement by the students. This form of service can be considered as an integral part of the Student Relationship Management (SRM) system. The purpose of this paper is to report the development of a decision tree model in an Intelligent Recommendation System with an aim to forecast the likely future Grade Point Average to be achieved by students. The study is based on a sample of 3,550 student records and the results demonstrated that the accuracy of the recommendation model is reasonable. This system will help the counselors in recommending the appropriate courses for students thereby increasing their chances of success.

Keywords—Intelligent Recommendation System, Student Relationship Management, Decision Tree, Data Mining.

I. INTRODUCTION

Nowadays, many higher education institutes are facing the problems of low rate of graduations in comparison to the number of enrollment. For example, Vandamme [1] reports that 60% of freshmen in Belgium either failed or dropped out from the courses. A report by the National Center for Public Policy and Higher Education (NCPPHE) [2] also shows that the average retention rate of U.S. tertiary freshmen who enrolled in 2002 is 73.6%. Similarly, the overall drop-out rate of university students in the south of Thailand over the five year period of 1999 to 2003 was 12%, and the discontinuation rates of students who enrolled after 2001 were even higher [3].

To assist a student’s achievement in higher education institutes, a service normally provided by most universities is Student Counseling. It can be said that a key aspect of student services is to provide course guidance as this will assist the students in their course selection and future university experience. On the other hand, many students have chosen particular courses of study just because of perceived job opportunities, peer pressure and parental advice. Issues may arise if a student is not interested in the course, or if the course or career is not suitably matched with the student’s capability[4]. In Thailand’s tertiary education sector, teaching staff may have insufficient time to counsel the students due to high workload and there are inadequate tools to support them. Hence, it is desirable that some forms of intelligent recommendation tools could be developed to assist staff and students in the enrolment process. This forms the motivation of this research.

Such system could be used to provide course advice and counseling for freshmen in order to achieve a better match between the student’s ability and success in course completion. In the case of Thai universities, this service is normally provided by counselors or advisors who have many years of experience within the organisation. However, with increasing number of students and expanded number of choices, the workload on the advisors is becoming too much to handle. It becomes apparent that some forms of intelligent system will be useful in assisting the advisors.

In this paper, a proposed intelligent recommendation system is reported. This paper is structured as follows. Section 2 describes the background, which is composed of Student Relationship Management (SRM) in universities and issues faced by Thai university students. Section 3 describes Decision Tree techniques which are used in the reported Intelligent Recommendation System, and Section 4 focuses on the proposed framework, which presents the main idea and the research methodology. Section 5 describes the experiments and the results. This paper then concludes with discussions on the work to be undertaken and future development.

II. BACKGROUND

A. Student Relationship Management in Universities

The problem of low student retention in higher education could be attributed to low student satisfaction, student transfers and drop-outs [1, 5]. This issue leads to a reduction in the number of enrolments and revenue, and increasing cost of replacement. On the other hand, it was found that the quality and convenience of support services are other factors that influence students to change educational institutes [6].
Consequently, the concept of SRM has been implemented in various universities so as to assist the improvement of the quality of learning processes and student activities.

Definitions of SRM have been adopted from the established practices of Customer Relationship Management (CRM) which focuses on customers and are aimed to establish effective competition and new strategies in order to improve the performance of a firm [7]. In the case of SRM, the context is within the education sector. Although there have been many research focused on CRM, few research studies have concentrated on SRM. In addition, the technological supports are inadequate to sustain SRM in universities. For instance, a SRM system’s architecture has been proposed so as to support the SRM concepts and techniques that assist the university’s Business Intelligent System [8]. This project provided a tool to aid the tertiary students in their decision-making process. The SRM strategy also provided the institution with SRM practices, including the planned activities to be developed for the students, as well other relevant participants. However, the study verified that the technological support to the SRM concepts and practices were insufficient at the time of writing [8].

In the context of educational institutes, Piedade and Santos (2008) explained that SRM involves the identification of performance indicators and behavioral patterns that characterize the students and the different situations under which the students are supervised. In addition, the concept of SRM is “understood as a process based on the student acquired knowledge, whose main purpose is to keep a close and effective students institution relationship through the closely monitoring of their academic activities along their academic path” [9]. Hence, it can be said that SRM can be utilised as an important means to support and enhance a student’s satisfaction. Since understanding the needs of the students is essential for their satisfaction, it is necessary to prepare strategies in both teaching and related services to support Student Relationship Management. This paper therefore proposes an innovative information system to assist students in universities in order to support the SRM concept.

B. Issues Faced By Thai University Students

Prior studies have studied issues faced by Thai students during their time in the universities. For example, Sarawut [10] studied the causes of dropouts and program incompletion among undergraduate students from the Faculty of Engineering, King Mongkut’s University of Technology North Bangkok, that the general reason for underachievement is a teaching and learning issue. Furthermore, the study shows that three unaccomplished group have different reasons. The main reason of the first group is a student’s attitude towards the field of study. This group has perception that their field of study is too hard. The main reasons of the second and third group are related to teaching and learning. Hence, this indicates the need to match the course requirements and academic capabilities of the students.

Another study at Dhurajikk Pundit University, Thailand looked at the relationship between learning behaviour and low academic achievement (below 2.0 GPA) of the first year students in the regular four-year undergraduate degree programs. The results indicated that students who had low academic achievement had a moderate score in every aspect of learning behaviour. On average, the students scored highest in class attendance, followed by the attempt to spend more time on study after obtaining low examination grades. Some of the problems and difficulties that mostly affected students’ low academic achievement were the students’ lack of understanding of the subject and lack of motivation and enthusiasm to learn [11].

Moreover, some other studies had focused on issues relating to students’ backgrounds prior to their enrolment, which may have effects on the progress of the students’ studies. For example, a research group from the Department of Education[12], Thailand studied the backgrounds of 289,007 Grade twelve students which may have affected their academic achievements. The study showed that the factors which could have effects on the academic achievement of the students may be attributed to personal information such as gender and interests, parental factors such as their jobs and qualifications, and information on the schools such as their sizes, types and ranking.

Therefore, in the recruitment and enrolment of students in higher education, it is necessary to meet the student’s needs and to match their capability with the course of their choice. The students’ backgrounds may also have a part to play in the matching process. Understanding the student’s needs will implicitly enhance the student’s learning experience and increase their chances of success, and thereby reduce the wastage of resources due to dropouts, and change of programs. These factors are therefore taken into consideration in the proposed recommendation system in this study.

III. DECISION TREE BASED INTELLIGENT RECOMMENDATION SYSTEM TO SUPPORT SRM

In term of education systems, Ackerman and Schibrowsky [13] have applied the concept of business relationships and proposed the business relationship marketing framework. The framework provided a different view on retention strategies and an economic justification on the need for implementing retention programs. The prominent result is the improvement of graduation rates by 65% by simply retaining one additional student out of every ten. The researcher added that this framework is appropriate both on the issues of places on quality of services. Although some problems could not be solved directly, it is recognized that Information and Communication Technologies (ICT) can be used and contributes towards maintaining a stronger relationship with students in the educational systems [8].

In this study, a new intelligent Recommendation System is proposed to support universities students in Thailand. This paper shows a hybrid system which is based on data mining
Another part of the framework focuses on the likelihood of GPA for students in each year. After the students selected the course to study and completed the enrollment process, the likelihood of GPA for Year 1 results can be used to monitor the performance of this group of students. The input data of this process is the same as the one shown in Table 1, with the addition of the GPA scores from the previous year. These are used as the extended features in the input to the CHAID model. The result of the Recommendation is the GPA score of the year. In the same way, the system may be used to perform a likelihood of GPA for Year 2 based on results from the first year. Similar approach can be adopted for the likelihood of Year 3 and 4 results. Some example results of this part are shown in this paper.

### IV. THE PROPOSED FRAMEWORK

Several solutions have been proposed to support SRM in the universities; however, not many systems in Thailand have focused on recommendation systems using historic records from graduated students. A recommendation system could apply statistical, artificial intelligence and data mining techniques by making appropriate recommendation for the students. Figure 1 illustrates the proposed recommendation system architecture. This proposal aims to analyse a student’s background such as the high school where the student studied previously, school results and the student’s performance in terms of GPA’s from the university’s database. The result can then be used to match the profiles of the new students. In this way, the recommendation system is designed to provide suggestions on the most appropriate courses and subjects for the students, based on historical records from the university’s database.

#### A. Data Analysis

It can be seen in Fig. 1 that Association rules, Decision Tree, Support Vector Machines and Neural Network are used to train the input data; however, this paper only reports the component on Decision Tree which uses the CHAID algorithm to classify the data and to establish the approximate function. For assessing the results statistically, F test is used in the case of continuous target field, and the chi-squared test is used for categorical target field.

#### B. Intelligent Recommendation Model

The Integrated Recommendation Model comprises three parts: Course Recommendation for freshmen, Likelihood of GPA, and Subject Recommendation for students (year 1 to 4) respectively.

Part A focuses on the course recommendation for freshmen and it is made up of two modules, which are the Overall GPA Recommendation, and the Course Ranking Recommendation respectively. In the Overall GPA recommendation module, the output is a numeric value for an expected overall GPA for the student. The output of the Course Ranking Recommendation module is a list of five ranked recommendations based on the results from the previous module.

The results of both parts can be used as suggestions to the freshmen during the consultation process. Some example results from Part A – Course Recommendation, are shown in this paper, and the input data of these two modules in the model are shown in Table 1.
supervisor to make recommendation to students in their enrollment of subjects or units in each year.

To address the issue of imbalanced number of students in each course, the Recommendation Model shown in Fig. 1 can be duplicated for different departments. The models’ computation is entirely data-driven and not based on subjective opinion, hence, the prediction models are unbiased and they will be used as an integral part of an Electronic Intelligent Recommendation System.

**C. Electronic Intelligent Recommendation System (e-IRS)**

It is planned that the new intelligent Recommendation Models will form an integral part of an online system for private universities in Thailand. The developed system will be evaluated by the university management and feedback from experienced counselors will be sought. The proposed system will also be available for use by new students who will access the online-application in their course selection during the enrolment process. As for the recommendation of students after enrolment and subsequent years’ results, this could be used by the counselors, staff, student’s supervisor and university management to provide supports for students who are likely to need help with their studies. This information will enable the university to better focus on the utilisation of their resources. In particular, this could be used to improve the retention rate by providing additional supports to the group of students who may be at risk.

**V. EXPERIMENT DESIGN**

The data preparation and selection process involves a dataset of 3,550 student records from five academic years. All the student data have included records from the first year to graduation. Due to privacy issues, the data in this study do not indicate any personal information, and no student is identified in the research. The student data has been randomised, and all private information has been removed. Example data from the dataset is shown below.

<table>
<thead>
<tr>
<th>No.</th>
<th>Variables</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>UniID</td>
<td>Randomized Student ID which is not included in the clustering process. They are only used as an identification of different students</td>
</tr>
<tr>
<td>2.</td>
<td>GPA</td>
<td>Overall GPA results from previous study prior to admission to university</td>
</tr>
<tr>
<td>3.</td>
<td>Type of school</td>
<td>The school types are separated as follows A: High School, B: Technical College, C: Commercial College, D: Open School, E: Sports, Thai Dancing, Religion or Handcraft Training Schools, F: Other Universities (change universities or courses) G: Vocation Training Schools</td>
</tr>
<tr>
<td>4.</td>
<td>Number of Awards</td>
<td>Awards that students have received from previous study (normalized between 0.0 to 4.0, 0.0 – received no award, 4.0 – received max no. of awards in the dataset)</td>
</tr>
<tr>
<td>5.</td>
<td>Talent and Interest (in Group number)</td>
<td>Talent and the interest (1 = sports, 2 = music and entertainment, 3 = presentation, 4 = academic, 5 = others, 6 = involved with 2 to 3 items of talents and interests, 7 = involved with more than 3 talents and interests)</td>
</tr>
<tr>
<td>6.</td>
<td>Channels</td>
<td>The channels to know the university such as television, family</td>
</tr>
<tr>
<td>7.</td>
<td>Admission Round</td>
<td>Admission round of each university which can be round 1 to 5</td>
</tr>
<tr>
<td>8.</td>
<td>Guardian Occupation</td>
<td>The occupation of Guardian such as teacher, governor</td>
</tr>
<tr>
<td>9.</td>
<td>Gender</td>
<td>Gender: Female or Male</td>
</tr>
<tr>
<td>10.</td>
<td>Uni GPA</td>
<td>Overall GPA in university which the range is from 0 to 4</td>
</tr>
</tbody>
</table>

**EXPERIMENTAL RESULTS**

A. Results of Course Recommendation (Likelihood of overall GPA and Course Ranking Recommendation)
The result of a CHAID analysis can be displayed in the local mode of tree diagram, which shows the detailed results within each node, number of training records, percentage of prediction, the predicted value and number of each node.

Fig. 3 Comparison of MAE of testing data of sub-models for overall GPA and course ranking Recommendation

Testing was carried out in the final step of the experiment which used 30% of the available data, and they were not used in the training phase. In Fig. 4, it can be seen that the lowest value of MAE is based on data from the Department of Law. On the other hand, the highest value is based on data from the Department of Public Administration. The average value of the MAE of all models is 0.068. It is observed that majority of the other courses are quite similar. Due to the relative low error of 6.8%, it could be said that the overall results indicated reasonable recommendation results.

B. Results of Likelihood of GPA of students each year

Fig. 4 Comparison of MAE of testing data on the Likelihood of GPA in each Year

Fig. 4 shows a comparison of MAE of the results of the sub-models from each department in each year. It can be seen that the range of values of MAE is the lowest based on data from the Department of education and the highest value is based on data from the Department of Communication Art. The average of MAE of all models based on all years is 0.393 which is higher than the result of MAE of testing data of sub-models for overall GPA and course ranking Recommendation. However, this error indicated reasonable recommendation results.

Based on the given data and MAE from the experiments, the results have shown that the Decision Tree based models can be utilised to predict the GPA results of students with a good degree of accuracy.

VII. CONCLUSIONS

This article describes a recommendation system in support of SRM and to address issues related to the problem of course advice or counseling for university students in Thailand. The recent work is focusing on the development and implementation of each process in the framework. The experiments have been based on Decision Tree models and the accuracy of the recommendation model is reasonable with a low value of MAE of 6.8%. In comparison to the Likelihood of GPA in each Year, the MAE is 39.3%. It can be seen that the accuracy of overall GPA and course ranking Recommendation is higher. It can be because of more number of students in training data.

It is expected that the recommendation system will provide a useful service for the university management; course counselors, academic staff and students, and the proposed system could also be used to support Student Relationship Management strategies among Thai private universities.

REFERENCES