Dropout Situation of Business Computer Students, University of Phayao

Pratya Nuankaew
University of Phayao, Phayao, Thailand
`pratya.nu@up.ac.th`

Abstract—This paper aims to study the problem of the dropout situation of students in the business computer program at the University of Phayao. It consists of three sub-goals: (1) The first objective is to identify factors related to the dropout of students in the business computer program at the University of Phayao. (2) The second objective is to study the relationship between factors related to drop-out of students in the business computer program at the University of Phayao. (3) The last objective is to test the relationship model of factors related to the dropout of students in the business computer program at the University of Phayao.

The sample group was 389 students from Business Computer Department at the School of Information and Communication Technology, University of Phayao in the academic year 2012-2019. The tools used in the research are statistical data analysis and machine learning. It consists of percentages, decision tree algorithms, cross-validation methods, and the confusion matrix performance.

The results showed that the dropout rate of learners in business computer program tended to increase even though the number of new students decreased. In addition, it was found that factors affecting the dropout consisted of six courses: 221110 Fundamental Information Technology, 221120 Introduction to Programming, 001103 Thai Language Skills, 128221 Principles of Marketing, 005171 Life and Health, and 122130 Principles of Management.

Keywords—Student dropout prediction, educational data mining, learning model, student model

1   Introduction

Education is a tool for developing people, raising the level of knowledge, thinking processes, and methods of survival in society [1-2]. The educational system and educational management process have various patterns such as education in the system, formal education, and informal education. [3-4]. All patterns of education are consistent with lifelong learning styles.

However, the educational process still has many obstacles. One of the major problems is the dropout of students. The problems of student's dropout appear in all education levels, all ages, and in all countries [5-9, 13-15].

At the same time, the problem of reduced attendance for most students in Thailand is appearing. There are problems with learners dropping out at all levels, for example, pre-primary school, primary school, secondary school, and at the university level tend to decrease in students as shown in Figure 1. Moreover, the attendance rate for higher education was at the low level of 58.5% in 2017 as shown in Figure 1. It represents the trend of future problems.
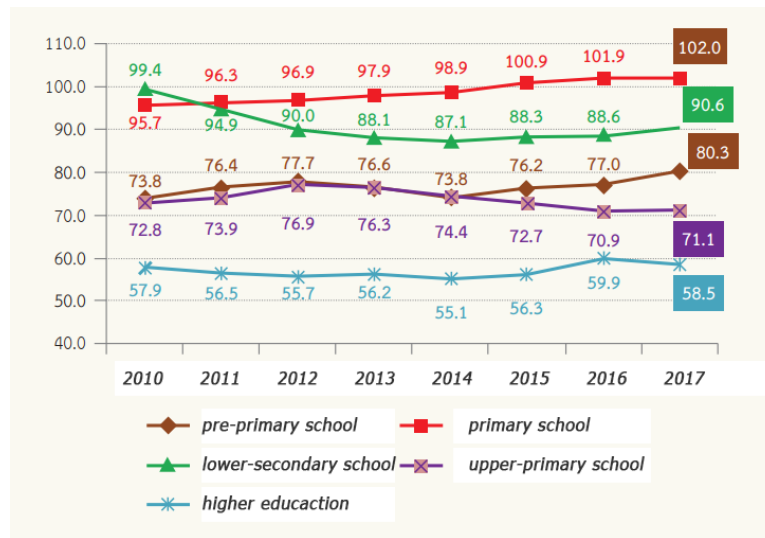

Fig. 1. An attendance rate of students during the academic year 2010-2017 of Thailand

The issue of dropout is seriously affecting students, society, families, schools, and universities. Due to the low birth rate situation, the competition in choosing students to enter schools and universities reduces the admission criteria of students. It affects the quality of learners and the attitude of attending school without competition. Moreover, UNESCO Bangkok reports that most students in Central and Southern Asia do not study at school or fail to achieve some academic results [7]. It is more reflective on part of the student's disregard for the institution and not focusing on their future. Regarding the specific reasons and problems, the event has inspired the researcher to conduct a research study on this topic.

1.1 Research purpose

More than twenty years ago, the number of students interested in business computer program was quite high [10, 12], but the number of students who were not able to achieve academic success was also increasing. One of the factors involved is the dropout of higher education students. The leading causes of student dropout is the performance, goal achievement, and an appropriate curriculum design. Thus, the research question is what courses and structures affect the dropout of students. The assumption of the course factors and structure can solve the problem of student drop-

out. Therefore, this research aims to study the causes and patterns of dropout of students in higher education.

The sample is 389 students of business computer program at the University of Phayao, consisting of students who have successfully graduated, close to graduating, and students who are currently studying. The collected data is divided into five sets.

The first set is 108 students who studied during the academic year 2012-2015, the second set is 102 students who studied during the academic year 2013-2016, the third set is 59 students who studied during the academic year 2014-2017, the fourth set is 42 students who studied during the academic year 2015-2018, and the last set is 71 students who are currently studying in the academic year 2016-2019.

The research tools are statistical data analysis and machine learning. These tools consist of percentages, decision tree algorithms, cross-validation methods, and the confusion matrix performance.

The purpose of the research is to know the situation of student dropout in business computer program at the University of Phayao. The work consists of three sub-goals: identifying the factors, studying the relationship of factors, and testing the relationship model.

1.2   Research approach

The research methodology has four main steps. The first step is to define the population and the sample groups that the researcher plans to collect data from the student's academic results. The second step is making data preparation and data management that are intended for processing data which is divided into five parts: data cleaning, data integration, data transformation, data reduction, and data analyzing [11].

After analyzing the data, the third step is to develop a prediction model and the learning process of the machine to create a model [2, 4, 11] to predict on the dropout. The final step is the performance indicator of the prediction model to test the model and forecast [11] on the data: steps are all shown in Figure 2.
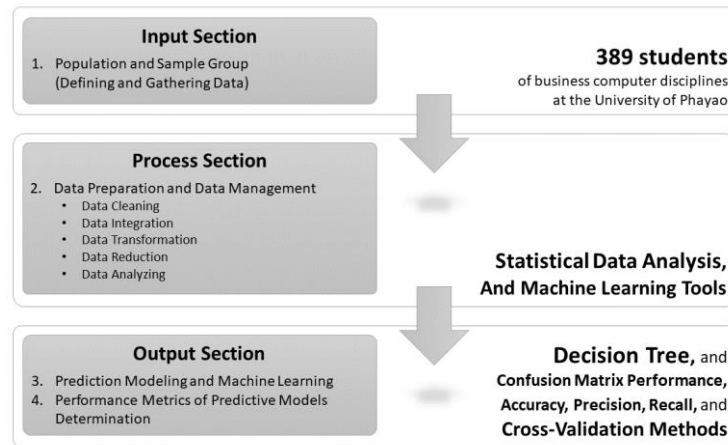
Fig. 2. Research Methodology

2 Literature Review

2.1 Challenge in business computer program

The business computer program has been around for more than 30 years' as part of the Thailand university system. It has been organized as part of the educational process since the beginning of the computer and technology era. Currently, the business computer program is still being taught in Thailand. The role of knowledge in the field is likely to change as the digital economy and online marketing field creates new demand for better technology. The obvious challenge is the higher intelligent technology [13]. It has developed learning and can be developed into various forms of relationships between computer technology and business disciplines. Such challenges have a significant impact on the field of study and for those interested in continuing their studies in the business computer program. Therefore, those who are involved in the field need to adjust themselves urgently to be able to compete in the digital world.

2.2 The Importance of predicting dropout on universities

It can be defined that the problem of student dropout is a significant problem in every country [5-9, 13-15]. The issue of dropout is not limited to just one level of education. The student dropout statistics report in Thailand shows that it is still high, although it is a basic level of education [10]. Meanwhile, the dropout statistics of university students in Thailand still tend to increase at a higher level as well [12].

Moreover, the direction of the study of student's dropout problems has received a lot of international attention [13-15]. A prime example is the study of the situation of the student's dropout in Tunisia, where the researchers studied 105 Tunisian schools during the 2012-2013 periods to develop four critical models for analyzing the issues

of consideration dropout levels and education environment [14]. Another example is the study of the problem of dropout in Italy. The researchers have reported important characteristics and problems of studying the impact of dropout at the university. The significant problems that come along are the disciplines that have not been consistent with the capabilities of the learners [15]. Besides, some researchers have developed more advanced research tools by using computational science and machine learning to solve dropout problems [11, 13]. Based on the above examples and priorities, it reflects the importance of this research.

3 Research Methodology

The research methodology follows the research approach which consists of four steps: defining the population and sample group, data preparation and data management, development of predictive models and machine learning steps, and performance indicator of the prediction model: all are shown with details in Figure 2.

3.1 Population and sampling group

The sample is 389 students majoring in the business computer program at the University of Phayao. The data collected is divided into five sets, as shown in Table. 1 below.

Table 1. Population and Sampling Group

Academic Year	Number of Students	Number of Graduates		Number of Dropouts	Dropout Classification				
		Scheduled	*Delayed*		*1*	*2*	*3*	*4*	*5*
2012-2015	108	38	20	50 (46.30%)	3	5	35	6	1
2013-2016	102	29	23	50 (49.02%)	3	5	35	2	7
2014-2017	59	26	7	26 (44.07%)	0	3	21	2	0
2015-2018	42	33	* Studying	16 (29.09%)	0	3	6	2	5
2016-2019	71	* Studying		21 (29.58%)	0	7	6	1	7
Total:	**389**	**126**	**50**	**163 (41.27%)**	**6**	**23**	**103**	**13**	**20**

1 = dropout without condition (not paying for registration fees), 2 = dropout with condition (not admitted), 3 = dropout with condition (academic results), 4 = dropout with condition (retire), 5 = dropout with condition (change programs)

Table 1 provides the collected details of students from the business computer program at the University of Phayao. The details include data from the academic year 2012-2019. The data in Table 1 shows that the number of students entering the study tends to decrease. However, the overall dropout rate constant is approximately half of all students, which is a fixed rate (163 students or 41.27%). Moreover, when considering the cause of dropout, it was found that most of the results were from the academic results (103 students or 26.06%).

3.2 Data preparation and data management

The purpose of data preparation and data management is to do a screening process for data analysis, modeling, and searching for courses that have an influence on the student dropout. It consists of five steps that include data cleaning, data integration, data transformation, data reduction, and data analyzing [11]. The following details are as followed:

Data cleaning is the process of reviewing, correcting, deleting incomplete data from the data set. It is an essential step in the database to eliminate imperfections data, inaccuracies data, and distorted data.

Data integration is a combination of techniques used to integrate data from various sources, in order to be meaningful and being able to use the data with reliability.

Data transformation is the process of converting data from one structure to another format or structure. It is an important step for analyzing data, such as the number of credits of unequal courses that need to be converted to the same standard or normalization calculated with the academic results.

Data reduction is the process of converting digital data, numbers, and letters which are compiled into a format that can be interpreted and calculated. For example, if the grade of the academic result is A, then it is required to be converted into values for the calculation in the next step by assigning A to equal 4.

Data analysis is a process of inspecting, cleaning, transforming, and creating data models to find useful information, conclusions, and supporting decisions.

3.3 Prediction modeling and machine learning

The construction of a prediction model is a set of techniques for developing and using models to make predictions based on patterns obtained from analyzing the data [2, 4, 11]. The technique used in this step is the decision tree. The decision tree is a structure that includes the root node, branch, and leaf node. The top node in the tree is the root node, which is the most critical part. Each internal node represents the test. Each branch feature represents the results of the test, and each leaf node holds a class label. The example of the simple decision tree model is shown in Figure 3.
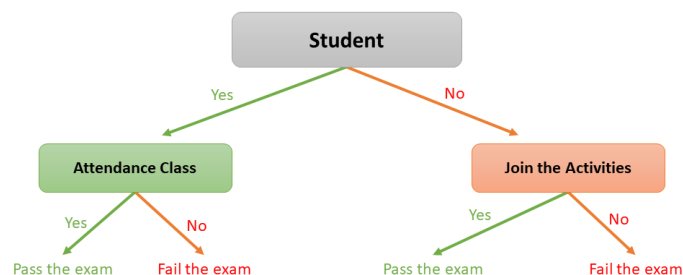

Fig. 3. The Simple Decision Tree Model

Figure 3 shows a simple decision tree model of exam results. It can predict the results of the exam from the factors (attributes) in the model, which consists of three factors: students, attendance class, and join the activities.

3.4 Performance metrics of predictive models' determination

The purpose of this step is to examine and consider the model obtained from the previous step. The tools used in this section consist of five types: confusion matrix performance, accuracy measurement, precision measurement, recall measurement, and cross-validation methods with the following details [11].

The confusion matrix is one of the simplest and easiest metrics used to find the accuracy and correctness of the model. It is used for classification problems where output can be two or more types of classes. The confusion matrix calculation method is shown in Table 2.

Accuracy measurement is a method for identifying problems from the correct number of predictions made by the model over all kinds of predictions made, where the calculation method is shown in Table 2.

Precision measurement is a measure that provides the predictions of the proportion of accuracy obtained from predictions in the prototype model, where the calculation method is shown in Table 2.

Recall measurement is a measurement unit that shows the proportion of truth that has been predicted by the algorithm, whether it is true or not. The calculation method is shown in Table 2.

Table 2. Confusion Matrix

Predicted / Actual	True Condition		Precision
	Positive	*Negative*	
Predicted Positive	True Positive (TP)	False Positive (FP)	$\text{Precision}_{(Positive)} = \frac{TP}{TP+FP}$
Predicted Negative	False Negative (FN)	True Negative (TN)	$\text{Precision}_{(Negative)} = \frac{TP}{TP+FP}$
Recall	$\text{Recall}_{(Positive)} = \frac{TP}{TP+FN}$	$\text{Recall}_{(Negative)} = \frac{TN}{FP+TN}$	

Cross-validation methods are methods for testing model performance by dividing data into two parts: one is using data for training, and one is for testing models.

4 Research Results and Discussion

Research results and discussions are divided into four parts: summary of data collection, summary results from model development, summary results from the model performance, and research discussion.

4.1 Summary of data collection

The data collected is 389 students of business computer program at the University of Phayao. The collected data consists of students who have successfully graduated, getting closer to graduation, and students who are currently studying. This is shown in Table 3 below.

Table 3. The dropout characteristics of each data set

Academic Year	Students: Dropouts*	Class level of students							
		1st year	2nd year	3rd year	4th year	5th year	6th year	7th year	8th year
2012-2019	108: 50	31 (28.70%)	13 (12.04%)	2 (1.85%)	2 (1.85%)	0	1 (0.95%)	1 (0.95%)	0
2013-2020	102: 50	33 (32.35%)	12 (11.76%)	5 (4.90%)	0	0	0	0	0
2014-2021	59: 26	18 (30.51%)	5 (8.47%)	2 (3.39%)	1 (1.69%)	0	0	0	0
2015-2022	49: 16	9 (18.37%)	5 (10.20%)	2 (4.08%)	0	0	0	0	0
2016-2023	71: 21	5 (7.04%)	15 (21.13%)	1 (1.41%)	0	0	0	0	0
Total:	**389: 163**	**96 (24.68%)**	**50 (12.85%)**	**12 (3.08%)**	**3 (0.77%)**	**0**	**1 (0.26%)**	**1 (0.26%)**	**0**

* The total number of students per number of students who have dropped out because of their academic results

Table 3 details the data collection. It shows the number of students who dropped out, which is approximately half of all students. Moreover, when considering the year of dropout, it was found that most of the results were from the 1st year (96 students or 24.68%). Details of the 1st year students' dropout is shown in Table 4

Table 4. The 1st year student's dropout

Academic Year	Number of Dropouts	Dropout Classification				
		Class 1	Class 2	Class 3	Class 4	Class 5
2012-2015	31	3 (9.68%)	2 (6.45%)	24 (77.42%)	0	2 (6.45%)
2013-2016	33	1 (3.03%)	4 (12.12%)	25 (75.76%)	2 (6.06%)	1 (3.03%)
2014-2017	18	0	2 (11.11%)	14 (77.78%)	2 (11.11%)	0
2015-2018*	9	0	1 (11.11%)	5 (55.56%)	0	3 (33.33%)
2016-2019*	5	0	4 (80.00%)	0	1 (20.00%)	0

Academic Year	Number of Dropouts	Dropout Classification				
		Class 1	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Total:	**96**	**4** (4.17%)	**13** (13.54%)	**68** (70.83%)	**5** (5.21%)	**6** (6.25%)

1 = dropout without condition (not paying for registration fees), 2 = dropout with condition (not admitted), 3 = dropout with condition (academic results), 4 = dropout with condition (retire), 5 = dropout with condition (change programs)

Table 4 shows the details of student's dropout from the 1st year business computer program at University of Phayao from the academic year 2012-2019. It shows that the majority were from class 3, based on the condition of academic results (68 students or 70.83%). Some examples of the academic results of the students are shown in Table 5. The results of all students appear on the website at https://bit.ly/2XIqiZw.

Table 5. Examples of Student's Academic Results

Stu-dent	Semester 1							Semester 2						Results	
	C 1	*C 2*	*C 3*	*C 4*	*C 5*	*C 6*	*C 7*	*C 8*	*C 9*	*C 10*	*C 11*	*C 12*	*C 13*	*Grad uated*	*Drop out*
sid001	F	B+	C+	A	D+	F	F	F	F	C	F	F	F	/	
sid002	F	A	B	B+	D+	D	D+	D+	D	C+	B	D+	C	/	
sid003	F	C	C+	C+	D+	W	B+	C	D	D+	D+	D+	A	/	
sid004	F	A	C+	B+	D	W	F	D	D	D+	D+	D	C		/
sid005	D	A	B	A	C+	C	B+	C	D+	B	B	C+	B	/	
sid006	F	B+	B	A	D	F	F	C	W	C	D	F	W		/
sid007	F	B+	C+	A	D+	W	D	D	F	C	C+	D	D		/
sid008	D+	A	A	A	C+	C	C	B	C+	B	B	B	C+	/	
sid009	D+	B	D	B	D	W	D	F	F	D+	C	F	D		/

C1 = 001111, C2 = 004152, C3 = 005171, C4 = 005173, C5 = 126100, C6 = 221100, C7 = 221110, C8 = 001103, C9 = 001112, C10 = 00313x,
C11= 122130, C12 = 128221, C13 = 221120, C1-C7 is the course that is studied in Semester 1, C8-C13 is the course that is studied in Semester 2

After preparing the data, the next part is the presentation of data analysis results, modeling results and model performance results.

4.2 Summary of model development

After preparing the data, this section is a development model that is used as a decision tool to predict the academic results. The tool is the decision tree process for de-

veloping models and it has been created as a rule model. The results are shown in Table 6.

Table 6. Rule Model

If	Then
If 221110 = A	Then Graduated (10 / 0)
If 221110 = B	Then Graduated (27 / 0)
If 221110 = B+	Then Graduated (20 / 0)
If 221110 = C and 221120 = A	Then Graduated (3 / 0)
If 221110 = C and 221120 = B	Then Graduated (4 / 0)
If 221110 = C and 221120 = B+	Then Graduated (1 / 0)
If 221110 = C and 221120 = C	Then Graduated (9 / 0)
If 221110 = C and 221120 = C+	Then Graduated (6 / 0)
If 221110 = C and 221120 = D	Then Graduated (3 / 1)
If 221110 = C and 221120 = D+	Then Graduated (4 / 0)
If 221110 = C and 221120 = F	Then Graduated (3 / 0)
If 221110 = C and 221120 = W	Then Graduated (0 / 1)
If 221110 = C+ and 001103 = B	Then Graduated (1 / 0)
If 221110 = C+ and 001103 = C	Then Graduated (11 / 0)
If 221110 = C+ and 001103 = C+	Then Graduated (7 / 0)
If 221110 = C+ and 001103 = D	Then Graduated (4 / 0)
If 221110 = C+ and 001103 = D+	Then Graduated (9 / 0)
If 221110 = C+ and 001103 = F	Then Dropout (0 / 2)
If 221110 = D and 128221 = B	Then Graduated (2 / 1)
If 221110 = D and 128221 = C	Then Graduated (9 / 0)
If 221110 = D and 128221 = C+	then Graduated (4 / 0)
If 221110 = D and 128221 = D	Then Dropout (3 / 4)
If 221110 = D and 128221 = D+	Then Graduated (9 / 4)
If 221110 = D and 128221 = F	Then Dropout (0 / 9)
If 221110 = D and 128221 = W	Then Dropout (0 / 1)
If 221110 = D and 005171 = B	Then Dropout (0 / 1)
If 221110 = D and 005171 = B+	Then Dropout (0 / 1)
If 221110 = D and 005171 = C+	Then Graduated (1 / 0)
If 221110 = D+ and 122130 = B	Then Graduated (6 / 0)
If 221110 = D+ and 122130 = B+	Then Graduated (1 / 0)
If 221110 = D+ and 122130 = C	Then Graduated (11 / 2)
If 221110 = D+ and 122130 = C+	Then Graduated (10 / 0)
If 221110 = D+ and 122130 = D	Then Graduated (1 / 0)
If 221110 = D+ and 122130 = D	Then Graduated (1 / 0)
If 221110 = D+ and 122130 = D+	Then Graduated (3 / 2)
If 221110 = D+ and 122130 = F	Then Dropout (0 / 2)
If 221110 = D+	Then Graduated (1 / 0)
If 221110 = F and 128221 = C	Then Dropout (0 / 4)
If 221110 = F and 128221 = C	Then Graduated (1 / 0)
If 221110 = F and 128221 = C+	Then Graduated (1 / 0)
If 221110 = F and 128221 = D	Then Dropout (3 / 16)
If 221110 = F and 128221 = D+	Then Dropout (0 / 5)

If	Then
If 221110 = F and 128221 = F	Then Dropout (0 / 11)
If 221110 = F and 128221 = W	Then Dropout (0 / 2)
Correct (93.80%): 242 out of 258 training examples.	

According to Table 6, it shows the rule model of the student achievement prediction model. It has the ability to predict the academic results of students at a very high level (Correct = 93.80% or 242 out of 258 training examples). The results of the model can be used to guide and plan student's learning in the business computer program at the University of Phayao. The model also shows the factors that are important to the student's study plan which consists of 221110, 221120, 001103, 128221, 005171, and 122130.

The next step is to test the model selection, which shows the performance of the model with details in the summary of model performance step.

4.3   Summary of model performance

This step is intended to test and select the best model by using the four types of performance evaluation tools, including confusion matrix performance, accuracy measurement, precision measurement, recall measurement, and cross-validation methods. The model performance results are shown in Table 7 and Table 8.

Table 7. Decision Tree's Performance

Depth of Model / Accuracy	Cross Validation			
	5-Fold	*10-Fold*	*15-Fold*	*Leave-one-out*
Depth 3*	86.45%	86.05%	87.21%*	87.21%*
Depth 5	82.95%	82.57%	83.73%	86.43%
Depth 7	82.95%	82.57%	83.73%	86.43%
Depth 9	82.95%	82.57%	83.73%	86.43%

Table 7 shows that the model is efficient and suitable for the decision tree model at depth 3 being used with the 15-fold and leave-one-out cross-validation methods, with the accuracy of 87.21%. After knowing the effective model, Table 8 shows the results of the significant model testing.

Table 8. Significant Model Testing

Predicted / Actual	True Condition		Precision
	True Graduated	*True Dropout*	
Pred. Graduated	177	21	89.39%
Pred. Dropout	12	48	80.00%
Recall	93.65%	69.57%	

4.4 Research discussion

From the results of data collection and the results of the implementation of the research methodology, there are two important issues in discussions: the number of student dropout and the model obtained from machine learning.

The number of student dropout: Table 1 clearly shows that the number of students who drop out has nearly half of the total number of students (163 students or 41.277%). Moreover, it was found that the cause of dropout is primarily due to student's academic results, with 103 students (26.06%). At the same time, it was found that the 1st year had the highest number of student dropout with 96 students (24.68%) as shown in Table 3. The reason that students leave most in the 1st year is from the academic results, and there are some parts that have dropped out in the 2nd year with the same reason as shown in Table 3 (50 students or 12.85%).

Therefore, it can be concluded that the impact of students who drop out in the business computer program mostly comes from the academic results which occurred in the overall picture shown in Table 1. For this reason, executives and related persons of business computer program at the University of Phayao should be aware of such problems and to configure with the facts for making plans to resolve the dropout issue.

However, the findings are consistent with the concept of the researcher, which has been expanded and developed the model of student achievement predictions in the business computer program. The success of the prediction model is discussed in the next section.

The model obtained from machine learning: According to the prediction model that has been designed and developed, the results of performance testing and prototype model selection are shown in Table 7 and Table 8, while the prediction model is shown in Table 6.

From Table 6, it can be seen that the courses that are significant to the students' predictions of academic results consist of six courses: 221110, 221120, 001103, 128221, 005171, and 122130. Where the most important course is 221110, in which the course acts as a root node. From the development of models in Table 6, it shows that the model can accurately predict the student's achievement at a high level (Correct = 93.80% or 242 out of 258 training examples). Therefore, it can be concluded that this developed model is suitable for predicting student's achievement.

Moreover, the model has been tested for the performance of the model as shown in Table 8. Table 8 shows performance values from the student achievement prediction model including the accuracy of the model which is equal to 87.21%; precision class of graduated is equal to 89.39%, precision class of dropout is equal to 80.00%, recall class of graduated is equal to 93.65%, and precision class of dropout is equal to 69.57%. From the model performance testing results, it reflects that the model is reasonable to be recognized and deserves to be used to expand the results in the future.

5   Conclusion

This paper aims to study the problem of students' dropout situation in the business computer program at the University of Phayao. There are three sub-goals: identifying the factors, studying the relationship of factors, and testing the relationship model. The data collected is 389 students from the Department of Business Computer at University of Phayao from the academic year 2012-2019 which is divided into five sets. The first set is 108 students who studied during the academic year 2012-2015, the second set is 102 students who studied during the academic year 2013-2016, the third set is 59 students who studied during the academic year 2014-2017, the fourth set is 42 students who studied during the academic year 2015-2018, and the last set is 71 students who are studying during the academic year 2016-2019. The research tools are statistical data analysis and machine learning. It consists of percentages, decision tree algorithms, cross-validation methods, and the confusion matrix performance.

From the collection of data, it clearly shows that the number of students who drop out was almost half of all students (163 students or 41.277%). In addition, it was found that academic results were the cause of dropout (103 students or 26.06%). It was also learned that the level with the most dropout is the 1st academic year.

Results from the study and development of the prototype model found that the factors affecting student achievement consisted of six courses: 221110, 221120, 001103, 128221, 005171, and 122130. The results of the prototype model development are shown in Table 6. Moreover, the prototype model has been tested for efficiency. It obtained a high-performance level of prediction with an accuracy that is equal to 87.21% as shown in Table 7 and Table 8. Thus, it has been concluded that the model is suitable for future expansion.

For future research, the researcher should expand the sample group to be more diverse, which is the goal of the next step. In addition, the research needs to gather data on the network of researchers in various universities in order to have more diverse samples. Furthermore, the researcher will continue to develop a mobile application to expand into an international research.

6   Acknowledgement

7 References

[1] Mahoney, M. H. (2017). An academic librarian's search for meaning: A survey of lifelong learning in library literature. *The Journal of Academic Librarianship*, *43*(6), 540–547. https://doi.org/10.1016/j.acalib.2017.05.004

[2] Nuankaew, P., Nuankaew, W., Phanniphong, K., & Bussaman, S. (2017). Mobile Applications for the Prediction of Learning Outcomes for Learning Strategies and Learning Achievement in Lifelong Learning. *International Conference on Interactive Collaborative Learning*, 400–412. Springer. https://doi.org/10.1109/tale.2017.8252332

[3] Martins, J., Branco, F., Gonçalves, R., Au-Yong-Oliveira, M., Oliveira, T., Naranjo-Zolotov, M., & Cruz-Jesus, F. (2019). Assessing the success behind the use of education management information systems in higher education. *Telematics and Informatics*, *38*, 182–193. https://doi.org/10.1016/j.tele.2018.10.001

[4] Nuankaew, W., Nuankaew, P., & Phanniphong, K. (2017). Lifelong Learning Application: Mobile Application to Promote Lifelong Learning and Introduce Educational Institutions. *Advances in Science, Technology and Engineering Systems Journal*, *2*(3), 1556–1564. https://doi.org/10.25046/aj0203194

[5] Hegde, V., & Prageeth, P. P. (2018). Higher education student dropout prediction and analysis through educational data mining. *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, 694–699. https://doi.org/10.1109/ICISC.2018.8398887

[6] Heredia, D., Amaya, Y., & Barrientos, E. (2015). Student Dropout Predictive Model Using Data Mining Techniques. *IEEE Latin America Transactions*, *13*(9), 3127–3134. https://doi.org/10.1109/TLA.2015.7350068

[7] Most children not learning are in school: Report points to 'crisis' in Central and Southern Asia. (n.d.). Retrieved March 26, 2019, from UNESCO Bangkok website: https://bangkok.unesco.org/index.php/content/most-children-not-learning-are-school-report-points-%E2%80%98crisis%E2%80%99-central-and-southern-asia

[8] Mustafa, M. N., Chowdhury, L., & Kamal, M. S. (2012). Students dropout prediction for intelligent system from tertiary level in developing country. *2012 International Conference on Informatics, Electronics Vision (ICIEV)*, 113–118. https://doi.org/10.1109/ICIEV.2012.6317441

[9] Tan, M., & Shao, P. (2015). Prediction of Student Dropout in E-Learning Program Through the Use of Machine Learning Method. *International Journal of Emerging Technologies in Learning (IJET)*, *10*(1), 11–17. https://doi.org/10.3991/ijet.v10i1.4189

[10] Thailand Drop-out rate for primary education, 1970-2018 - knoema.com. (n.d.). Retrieved March 26, 2019, from Knoema website: https://knoema.com//atlas/Thailand/topics/Education/Primary-Education/Drop-out-rate-for-primary-education

[11] Nuankaew, W. S., Nuankaew, P., Teeraputon, D., Phanniphong, K., & Bussaman, S. (2019). Perception and Attitude Toward Self-Regulated Learning of Thailand's Students in Educational Data Mining Perspective. *International Journal of Emerging Technologies in Learning (IJET)*, *14*(09), 16. https://doi.org/10.3991/ijet.v14i09.10048

[12] National statistics office Thailand Home. (n.d.). "Education Statistics". Retrieved June 2, 2019, from http://statbbi.nso.go.th/staticreport/page/sector/en/03.aspx

[13] Chung, J. Y., & Lee, S. (2019). Dropout early warning systems for high school students using machine learning. *Children and Youth Services Review*, *96*, 346–353. https://doi.org/10.1016/j.childyouth.2018.11.030

[14] Ben Yahia, F., Essid, H., & Rebai, S. (2018). Do dropout and environmental factors matter? A directional distance function assessment of tunisian education efficiency. *International*

Journal of Educational Development, *60*, 120–127. https://doi.org/10.1016/j.ijedudev.2017.11.004

[15] Gitto, L., Minervini, L. F., & Monaco, L. (2016). University dropouts in Italy: Are supply side characteristics part of the problem? *Economic Analysis and Policy*, *49*, 108–116. https://doi.org/10.1016/j.eap.2015.12.004

8 Author

Pratya Nuankaew is currently a lecturer at the School of Information and Communication Technology, University of Phayao, Phayao, Thailand. He received a B.Ed. Degree in Educational Technology in 2001, M.Sc. degree in Information Technology in 2008 from Naresuan University, and a Ph.D. degree in Computer Engineering in 2018 from Mae Fah Luang University. His research interests are in Engineering Education Research (EER), Educational Data Mining (EDM), Self-Regulated Learning Strategies (SRLS), Educational Communications and Technology (ECT), Knowledge Management in Technology Education (KMTE), learning analytics, mobile learning, ubiquitous learning and context aware.