The Recommendation System for Learning Activities According to Learning Styles Analyzed by Data Mining
ระบบเสนอแนะกิจกรรมการเรียนรู้ตามรูปแบบการเรียนรู้ซึ่งวิเคราะห์ด้วยการทำเหมืองข้อมูล

Pensri Amornsinlaphachai
เพ็ญศรี อมรศิลปชัย

Article History
Receive: August 15, 2020
Revised: December 9, 2021
Accepted: December 11, 2021

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์ตัวแปรที่เกี่ยวข้องกับรูปแบบการเรียนรู้ของผู้เรียน เพื่อคัดเลือกโมเดลจำแนกรูปแบบการเรียนรู้ เพื่อพัฒนาระบบเสนอแนะกิจกรรมการเรียนรู้ตามรูปแบบการเรียนรู้ซึ่งวิเคราะห์ด้วยการทำเหมืองข้อมูล และ 4) เพื่อเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของผู้เรียน ระหว่างการเรียนซึ่งจัดกิจกรรมการเรียนรู้โดยใช้ระบบเสนอแนะกิจกรรมการเรียนรู้ตามรูปแบบการเรียนรู้ กับการเรียนซึ่งจัดกิจกรรมการเรียนรู้โดยผู้สอน สำหรับตัวแปรที่เกี่ยวข้องกับรูปแบบการเรียนรู้ จะวิเคราะห์จากความคิดเห็นของผู้เชี่ยวชาญ จำนวน 12 คน ซึ่งเก็บรวมรวมโดยใช้แบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อตัวแปรที่เกี่ยวข้องกับรูปแบบการเรียนรู้ ในการสร้างโมเดลจำแนกรูปแบบการเรียนรู้ จะใช้ข้อมูลของนักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏนครราชสีมา จำนวน 1,328 คน ซึ่งเก็บรวมรวมโดยใช้แบบสอบถามวัดรูปแบบการเรียนรู้ ตามหลักการของ ของ Honey และ Mumford ส่วนการประเมินระบบเสนอแนะกิจกรรมการเรียนรู้ จะประเมินจากความพึงพอใจของผู้เชี่ยวชาญที่มีต่อระบบ จำนวน 5 คน และในการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนจะทดลองกับนักศึกษา 2 กลุ่ม โดย กลุ่มที่ 1 จำนวน 28 คน ซึ่งเรียนโดยใช้ระบบจัดกิจกรรมการเรียนรู้ และกลุ่มที่ 2 จำนวน 27 คน ซึ่งเรียนโดยผู้สอนเป็นผู้จัดกิจกรรมการเรียนรู้ ผลการวิจัยพบว่า 1) มี 7 ตัวแปรที่เกี่ยวข้องกับรูปแบบการเรียนรู้ ได้แก่ เพศ คณะ ชั้นปี เกรดเฉลี่ยสะสม วิชาที่ทำคะแนนได้สูงสุด วุฒิการศึกษาเดิม และแผนการเรียนเดิม 2) โมเดลที่มีประสิทธิภาพมากที่สุด คือ โมเดลที่สร้างจากอัลกอริทึม J48graft ซึ่งมีค่าความถูกต้องร้อยละ 82.23 โดยผลลัพธ์นี้ได้มาจากการทดสอบประสิทธิภาพของโมเดลจำแนกรูปแบบการเรียนรู้ซึ่งสร้างจากอัลกอริทึมด้านการทำเหมืองข้อมูลจำนวน 14 อัลกอริทึม สำหรับข้อมูลเข้าที่นำไปทำเหมืองข้อมูล ได้แก่ ข้อมูลของ 7 ตัวแปรที่เกี่ยวข้องกับรูปแบบการเรียนรู้ และ ผลการตอบแบบสอบถามวัดรูปแบบการเรียนรู้ 3) ระบบเสนอแนะกิจกรรมการเรียนรู้ ประกอบด้วย 4 โมดูลหลัก ได้แก่ โมดูลการจัดการข้อมูลนักศึกษา โมดูลกิจกรรมการเรียนรู้ โมดูลการทำนายและเสนอแนะ และ โมดูลการวัดผลสัมฤทธิ์ทางการเรียน โดยโมเดลที่สร้างจาก J48graft ถูกนำไปใช้เป็นพื้นฐานในการสร้างโมดูลการทำนายและเสนอแนะ สำหรับการประเมินความพึงพอใจของผู้เชี่ยวชาญที่มีต่อระบบโดยภาพรวม คะแนนเฉลี่ยอยู่ที่ 4.6 ซึ่งความพึงพอใจอยู่ในระดับมากที่สุด 4) นักศึกษาซึ่งเรียนโดยใช้ระบบจัดกิจกรรมการเรียนรู้ มีผลสัมฤทธิ์ทางการเรียนสูงกว่านักศึกษาที่เรียนโดยผู้สอนเป็นผู้จัดกิจกรรมการเรียนรู้ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

คำสำคัญ : รูปแบบการเรียนรู้ ; ตัวแปรที่เกี่ยวข้องกับรูปแบบการเรียนรู้ ; กิจกรรมการเรียนรู้ ; การทำเหมืองข้อมูล ; เทคนิคการจำแนกประเภท

ผู้ช่วยศาสตราจารย์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา, Assistant Professor, Faculty of Science and Technology, Nakhon Ratchasima Rajabhat University

ABSTRACT

This research has four purposes including 1) to analyze the variables related to learners' learning style 2) to select a learning style classification model 3) to develop a recommendation system for learning activities according to learning style analyzed by data mining and 4) to compare learners' study achievement between the study whose learning activities are organized by the recommendation system for learning activities and the study whose learning activities are organized by an instructor. For variables related to learning style, 12 experts' opinions are analyzed by using a questionnaire for experts' opinions on variables related to learning style. Learning style questionnaires were developed based on the principle of Honey and Mumford, which are collected from 1,328 undergraduate students studying at Nakhon Ratchasima Rajabhat University. These samplings are applied to create 14 learning classification models. In order to evaluate the recommendation system for learning activities, the satisfactions of the system are collected from 5 experts. For comparing study achievement, the experiment is conducted with 2 groups of students. The first group consists of 28 students studying by the system-based organizing the learning activities whereas the second group consists of 28 students studying by an instructor-based organizing's learning activities. The research results are as follows. Firstly, there are 7 variables related to the learning style : gender, faculty, year, GPA, highest score subject, previous qualification and previous study plans. Secondly, the most efficient model is the one created by J48graft algorithm, which has an accuracy of 82.23%. This results are obtained by testing the performance of learning style classification models created by 14 data mining algorithms. The inputs for data mining are the 7 variables related to learning style and the results of the learning style questionnaires collected from 1,328 students. Thirdly, the recommendation system for learning activities consists of 4 main modules: student data management module, learning activity module, prediction and recommendation module and study achievement module. The model created by J48graft is used as a basis for building the prediction and recommendation module. For assessing experts' satisfaction with the system, overall average score is 4.60, which satisfaction is at the most level. Fourthly, the students from the system-based organizing's learning activities reach higher study achievement than the students from the instructor-based organizing's learning activities at the statistical significance level of .05.

Keywords : Learning Style ; Variables Related to Learning Style ; Learning Activity ; Data Mining ; Classification Technique

Introduction

Reducing inequality by improving the quality of education is one of the important strategies of Thailand which is consistent with the national development strategy in Thai National Economic and Social Development Plan (The twelfth edition 2017–2021) that wants to change the paradigm of education management emphasizing on knowledge transfer by teachers to learner-centered education that considers learners the most importance (Chaijaroen et al., 2018 ; Kwangmuang, 2018). That is educational institutions must provide content and activities in accordance with learners' interests and skills by taking into account individual differences. For teaching, we find that each learner has different learning style. If teachers can know the learning style of each learner, it will enable teachers to organize learning activities that are suitable for each learner and this will affect the performance of learners (Fleming and Bonwell, 2019).

The learning style of the individual depends on the physical characteristics, thought, responding and interacting with the environment. Learning style can be categorized into several forms. For example, Fleming and Bonwell (2019) identified four learning styles according to individual perception as follows:

1) visual 2) aural 3) read / write and 4) kinesthetic. Honey and Mumford (1982) divided the learning style into four categories based on human traits: activists, reflectors, theorists and pragmatists. Each learning style has its own suitable learning activity.

Thus, if teachers can know learning styles of learners, it will enable teachers to organize learning activities that are suitable for each learner. This will affect the performance of the learner (Fleming and Bonwell, 2019). However, teachers are often not interested in the learning style of each learner; thus teachers always organize learning activities for all learners in the same format. In the case that teachers want to organize learning activities according to the learners' learning styles, teachers may use a questionnaire with a large number of questions according to the educational theories mentioned above which takes time to collect and analyze data. Some teachers may use their own judgment to analyze or observe learners which is likely to mistake. To eliminate such problems, in computer science, there is a prediction method called data mining to predict events by relying on the past data. For example, Prasada Rao et al (2016) compared the predictions of students' learning behavior with J48, Naïve Bayes and Random forest algorithm by using several variables such as gender, grade and attendance. It was found that, Random forest algorithm shows better accuracy while the data set size goes on increasing. Hence data mining can be considered as a technique for predicting the outcome of various events with high reliability and precision.

Therefore we are interested in applying data mining technique to build a model to classify learners' learning style since knowing learners' learning style makes it possible to organize learning activities that are suitable for each learner which will affect the performance of learners. In this research, a recommendation system for learning activities according to learning styles analyzed by data mining is developed by using various theoretical concepts that can be put into action. It is expected that this research will meet government policy to develop the quality of the Thai people to achieve the country's intellectual capital. Moreover it will result in the development of the teaching and learning process for educational institutions in Thailand as well as lead to a sustainable learning society in the future.

Objective

Four main objectives of this research are as follows.

1. To analyze the variables related to learners' learning style
2. To select a learning style classification model
3. To develop a recommendation system for learning activities according to learning style analyzed by data mining
4. To compare learners' study achievement between the study whose learning activities are organized by the recommendation system and the study whose learning activities are organized by an instructor.

Literature Review

1. Data mining

Data mining is analyzing, finding and scrutinizing data and data's relational pattern from large amounts of data to obtain new and useful information that can be applied. There are several techniques for data mining. Which technique is chosen depends on the nature of data and the relationships between data including the purpose of problem solving (Mirza et al., 2016). Three most commonly used data mining techniques are 1) association rule discovery 2) classification) and 3) clustering. Several algorithms of the classification technique utilized in this research are as follows.

Naïve Bayes (NB) is used to create a model with a simple classification method using the probability theory. Its hypothesis is that every attribute used to determine the probability must be independent. Whereas Bayesian Belief Network called BayesNet (BN) which can describe condition independent is an algorithm reducing the disadvantage of Naïve Bayes (Berrar, 2019).

Decision tree algorithm is used to create a rule-based model. This algorithm learns from past events that had happened and uses conditions to decide what results will be achieved when events occur (Radhwan et al., 2017). The decision tree algorithm can be subdivided into several algorithms such as ID3, C4.5 or J48, J48graft, REP Tree or CART, CHAID, random forest (RF) and random trees (RT).

Artificial neural network (ANN) uses functions obtained from a group of data to calculate the results. In the ANN, there is a parallel grouping of sub-processors called nodes or neurons and there are links between nodes similar to the human brain. SVM is a mathematical model that is in the same group as artificial neural network, but SVM has advantages in that it is less likely to cause overfitting problems (Li et al., 2017).

ZeroR is the simplest classification method which relies on the target and ignores all predictors. ZeroR classifier simply predicts the majority class (Sangeorzan, 2019). Whereas JRIP or RIPPER is a rule-based classification using a set of IF-THEN rules to predict results (Eyasu et al., 2020). For K-nearest neighbor (KNN), a data segmentation method is used by measuring the distance between data and the number of neighboring data. If any class of data has the greatest number, that class is the result

2. Educational Theory

2.1 Learning style and learning activity

Learning style is physical aspects, thoughts and feelings that people use to perceive, respond, and interact with the learning environment in a relatively stable manner (Jaleel and Mary Thomas, 2019). Learning style can be categorized into several forms. For example, Kolb (1984) classified learning styles into four types: 1) diverging learning style is careful analysis of multiple approaches 2) assimilating learning style is the use of reasoning to judge things 3) converging learning style is the use of the mind to make judgments and 4) accommodating learning style is taking action to see real results. Fleming and Bonwell (2019) identified four learning styles according to individual perception as follows: 1) visual is learned through visualization 2) aural is learning through listening 3) read / write is learned through reading and writing and 4) kinesthetic is learned through testing. Honey and Mumford (1982) divided learning styles into four types based on the person's characteristics : 1) activists love new experiences and new ideas and enjoy working as a team 2) reflectors love to collect information and think through it carefully before concluding something 3) theorists think of the problem step by step and like asking and thinking carefully and 4) pragmatists are less tolerant of long conversations but like ideas that can be applied in real work. Each learning style has its own appropriate learning activity. For example, Honey and Mumford (1982) divided learning styles into activists, reflectors, theorists and pragmatists where appropriate activities are brainstorming, paired discussion, application of theory and case study respectively.

2.2 Bloom's Taxonomy

Bloom (1956) defines three aspects of human learning behavior as follows:

1. Cognitive domain is brain behavior which is intelligence behavior and the ability to think of different stories. The cognitive domain is divided into six levels that are knowledge, comprehension, application, analysis, synthesis and evaluation.

2. Affective domain is a psychological behavior consisting of 5 levels of behavior that are receive, respond, value, organize and characterize.

3. The psychomotor domain is a physical behavior or neuromuscular consisting of 5 levels of behavior that are imitation, manipulation, precision, articulation and naturalization.

For this research, the cognitive domain is utilized in the part of study achievement test.

3. Technology

3.1 Cloud technology

Cloud technology is a computing technology which shares computer resources, both hardware and software, to provide various services via a web browser or various applications connected to cloud computing. Cloud computing has been used for a long time through Internet-based services such as Hotmail and Yahoo. Nowadays, cloud computing is used in several aspects such as facebook and youtube. These online applications make data can be accessed anytime and anywhere via Internet connected to cloud computing. Due to the capabilities of these applications that can deliver a wide range of services, cloud computing is applied in many fields, including education field (Na Nongkhai and Keawkiriya, 2016). Currently, one of the most popular educational application running through cloud computing is google application which is an online application that offers a wide range of services such as electronic mail, online meeting, online storage, online questionnaire, online test and virtual classroom.

3.2 Web-based technology

At present, web-based technology is often applied in education. It is called web-based learning or web-based instruction (WBI) which is a form of education blending the internet technology with instructional design processes to enhance learning efficiency and solve problems in time and place (Chongsomchai, 2018). The Internet is the largest computer network which has many service forms. In the field of education, the Internet can be used in many ways, for instance, webboard, chat, conference and other social media such as line and facebook. WBI will make students can communicate and discuss with other learners and experts. Therefore learning will be like normal classroom learning where the classroom is represented by the web and the book's content is replaced by the web's content whereas discussion will be done through chat, webboard and other social media. This will make the learner who is not assertive can express more opinions and ask questions.

From relevant research studies, knowing learners' learning style makes it possible to organize learning activities that are appropriate for each learner which will affect the performance of learners. However if instructors have to find learning styles of all learners by themselves, it will take quite a lot of time to collect and analyze data. Using a learning style classification model created by data mining helps instructors quickly understand each learner's learning style. Since there are several data mining algorithms, these algorithms have to be compared to find the most effective algorithm to be used to create the model. For learning style, there are many learning style principles; nevertheless from related research studies, it has not been found that the learning style based on Honey and Mumford principle has been applied to data mining. Therefore this research applies the aforementioned learning style to data mining. In addition, web and cloud technologies are used to store student data, learning activities and study achievement tests; thus this allows instructors to access the data from anywhere and also allows immediate processing the tests. Therefore teaching is more efficient. For measuring study achievement, the cognitive domain according to Bloom's Taxonomy is evaluated.

Conceptual Framework

Several educational theories and computer science principles are used in this research. The conceptual framework of the research is shown in Figure 1.

From the conceptual framework, many educational theories and computer science principles used in this research are learning style, learning activities according to learning style, Bloom's Taxonomy, data mining, model evaluation, Cloud technology and Web-based technology whereas the variables involved in this research are as follows.

1. Independent variables

1.1 Variables related to Learners' learning style

1.2 Performance of each model used to classify learning style

1.3 The study whose learning activities are organized by the recommendation system for learning activities and the study whose learning activities are organized by an instructor.

2. Dependent variables

2.1 The best performance model used to classify learning style

2.2 Academic achievement of learners

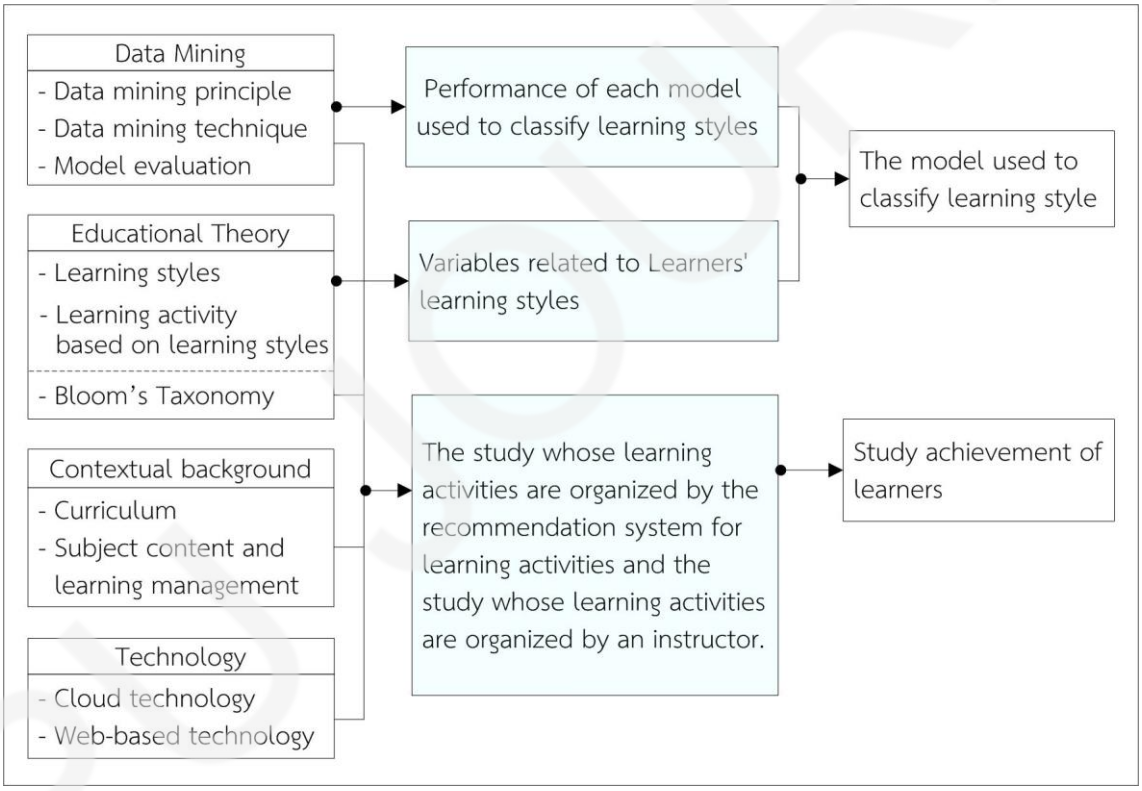

Figure 1 Conceptual Framework

Research Methodology

Population, Samples and Target groups

Population and Samples

To collect data from the learning style questionnaire, 1,328 students used in this phase are calculated from 10,516 students at Nakhon Ratchasima Rajabhat University by using the Krejcie & Morgan formula at a confidence level of 95%.

Target groups

The target groups used in this research was divided into 2 main groups: expert group and student group. The details of each group are as follows.

1. Expert group

1.1 To analyze the variables related to learners' learning style, 12 experts graduated with a minimum of a master's degree with knowledge of learning style and had a teaching experience for 5 years or more.

1.2 To examine the translation of the learning style questionnaire, 3 English experts graduated with a minimum of a master's degree in English and had a teaching experience for 5 years or more from Nakhon Ratchasima Rajabhat University.

1.3 To assess the recommendation system for learning activities, 5 experts graduated with a minimum of a master's degree in a related field and had a teaching experience for 5 years or more.

1.4 To assess the consistency index of the questionnaire for assessing the recommendation system for learning activities, 5 experts graduated with a minimum of a master's degree in a related field and had a teaching experience for 5 years or more.

1.5 To assess the consistency index of the academic achievement test, 5 experts graduated with a minimum of a master's degree in a related field and had a teaching experience for 5 years or more.

2. Student group

2.1 To identify the quality of the academic achievement test that is difficulty, discrimination and reliability of the test, 29 undergraduate students of computer education department at Nakhon Ratchasima Rajabhat University who previously enrolled in the computer network data communication and ethics for information technology course in Semester 1/2019, are used in this stage.

2.2 To compare study achievement, 55 undergraduate students of computer education department at Nakhon Ratchasima Rajabhat University who enrolled in the computer network system course in Semester 1/2020 are separated into 28 students studying with learning activities suggested with the system and 27 students studying with learning activities organized by an instructor.

Research Instruments

The research instruments consist of several tools as follows.

1. A questionnaire for experts' opinions on variables related to learning style composes of 12 basic variables: gender, faculty, year, GPA, highest score subject, previous qualification, previous study plan, location of previous school, computer and internet experience, homeland of learners, parents' education and family income. These 12 variables used as questions of the questionnaire are derived from relevant researches.

2. A learning style questionnaire consists 2 parts as follows. Part 1 is questions obtained from analyzing 12 experts' opinions about variables related to learning style. Part 2 is 40 questions based on Honey and Mumford learning style questionnaire (Honey and Mumford, 1982). The questions of the questionnaire are reviewed by 3 English experts.

3. Fourteen learning style classification models are created by 14 data mining algorithms. The performance of these models are compared to select the best performance model to be used as a part of the recommendation system for learning activities. The result of comparing these models is shown in the section of Results and Discussion.

4. The recommendation system for learning activities is designed and developed based on educational theories and computer science principles as shown in Figure 1. This system is evaluated by 5 experts by using an evaluation form consisting of 3 components: 1) system design, 2) system components, and 3) system performance. The IOC (Index of Item Objective Congruence) of the questions in the form is evaluated by 5 experts and it's found that all questions have a consistency with the assessment objective of 0.8, which is greater than the criteria of 0.5. The result of assessing the system is shown in the section of Results and Discussion.

5. A study achievement test whose content is a part of the computer network system course according to curriculum (revised) 2019 and the computer network data communication and ethics for information technology course according to curriculum (revised) 2017 of Science and Technology Faculty, Nakhon Ratchasima Rajabhat University. The test composes of 70 multiple choice questions by scoring 1 point per 1 question. The IOC of the questions in the test is evaluated by 5 experts and it's found that all questions have a consistency with the learning objective of 0.8, which is greater than the criteria of 0.5. However after assessing the quality of the test, only 65 questions can be used. The reliability of the test is calculated with formula of Kuder-Richardson's KR20 and the result shows that these 65 questions have the reliability value of 0.95.

Data collection

1. The questionnaires about variables related to learning style are collected from 12 experts. Characteristics of questions show the 12 basic variables' suitability in 5 levels: 5, 4, 3, 2 and 1 which mean highest to lowest according to the Likert technique. This data is obtained in Semester 2/2019.

2. The learners' learning style questionnaires are collected from 1,328 undergraduate students at Nakhon Ratchasima Rajabhat University in Semester 2/2019. Students' learning styles are determined by using the criteria of Honey and Mumford (Honey and Mumford, 1982). .

3. The performance of 14 learning style classification models indicated by correctness, precision, recall, f-measure, processing time and mean absolute error of the models are calculated during creating the models with data mining software in Semester 2/2019.

4. The evaluation results of the recommendation system for learning activities are obtained from 5 experts' opinions in Semester 2/2019. Characteristics of questions show experts' satisfaction with the system in 5 levels: 5, 4, 3, 2 and 1 according to the Likert technique.

5. The quality of the study achievement test is calculated from the test scores of 29 students who previously enrolled in the computer network data communication and ethics for information technology course in Semester 1/2019. For a study experiment, learners' study achievement is derived from the experiment with 2 group of students enrolling in computer network system course in semester 1/2020. In group 1 with 28 students, learning activities are organized by the recommendation system for learning activities whereas in group 2 with 27 students, learning activities are organized by an instructor.

Data Analysis

1. The 12 experts' opinions on variables related to learning style are analyzed by using mean ($\bar{X}$) and standard deviation (S.D.). The variables to be used to classify learning style must have the relevance to learning style in the high and highest levels.

2. Students' learning style data based on the criteria of Honey and Mumford are analyzed by data mining algorithms to create models to predict learners' learning style.

3. The Performance of learning style classification models indicated by correctness, precision, recall, f-measure, processing time and mean absolute error of the models is analyzed by descriptive analysis and interpretation

4. The 5 experts' opinions on the recommendation system for learning activities are analyzed by using mean ($\overline{X}$) and standard deviation (S.D.).

5. The Quality of the study achievement test is calculated with the formulas for difficulty and discrimination of the test and the formula of Kuder-Richardson's KR20 for reliability of the test. The questions of the test to be used must have difficulty value of 0.21-0.80 and discrimination value of 0.20 or above. Learners' study achievement is obtained from the aforementioned study experiment with 2 group of students. The achievement scores for each group are compared by using t-test (independent).

Results and Discussion

1. Analyzing the variables related to learners' learning style

For analyzing the variables related to learners' learning style, it is started by studying the basic variables related to learning style from relevant researches. The result of this study is found that 12 basic variables : gender, faculty, year, GPA, highest score subject, previous qualification, previous study plan, location of previous school, computer and internet experience, homeland of learners, parents' education and family income are related to learning style as shown in Table 1.

Table 1 The results of analyzing variables related to the learning style

Variables related to learning style	$\overline{X}$	S.D.	Level of relevant
Gender	4.42	0.76	high
Faculty	4.58	0.64	highest
Year	4.08	0.95	high
GPA	4.17	0.90	high
highest score subject	4.50	0.65	highest
previous qualification	3.75	0.92	high
previous study plan	4.25	0.83	high
location of previous school	2.92	0.95	moderate
computer and internet experience	2.83	1.28	moderate
homeland of learners	2.67	1.19	moderate
parents' education	2.92	1.32	moderate
family income	3.08	1.26	moderate

From Table 1, the 12 basic variables are used as questions in a questionnaire to ask 12 experts' opinions. The result of analyzing 12 experts' opinions shows that 7 variables: gender, faculty, year, GPA, highest score subject, previous qualification and previous study plans are related to learners' learning style in the high and highest levels. This result is consistent with the researches of Pantho (2016), Chuangchai (2016) and Iaosanurak et al. (2017). Pantho (2016) indicates that gender, study field, year, GPA and previous qualification are related to learning style while Chuangchai (2016) identifies that gender and year are related to learning style and Iaosanurak et al (2017) specify that gender, study field and study achievement are related to learning style.

2. Selecting a learning style classification model

For this research, 14 learning classification models are created by 14 data mining algorithms whose inputs are questions derived from the 7 variables related to learning style and the learning style questionnaire based on the principle of Honey and Mumford collected from 1,328 undergraduate students studying at Nakhon Ratchasima Rajabhat University. The performance of the 14 models based on 14 algorithms ordered by correctness is shown in Table 2.

Table 2 The performance of each model generated from 14 algorithms ordered by correctness

Algorithm	Correctness	Precision	Recall	F-Measure	Mean Absolute Error	Time (Sec.)
J48graft	82.23	0.83	0.82	0.82	0.12	0.17
J48	82.00	0.83	0.82	0.82	0.13	0.10
JRip	81.25	0.82	0.81	0.81	0.15	0.26
RF	81.17	0.81	0.81	0.81	0.13	0.56
CART	81.02	0.82	0.81	0.81	0.14	0.03
CHAID	79.52	0.80	0.80	0.80	0.14	0.10
RT	79.52	0.80	0.80	0.80	0.12	0.01
ANN	79.37	0.79	0.79	0.79	0.12	4.30
KNN	78.16	0.78	0.78	0.78	0.13	0.00
ID3	77.56	0.81	0.81	0.81	0.11	0.04
BN (p=2)	75.98	0.76	0.76	0.76	0.19	0.02
SVM	60.47	0.62	0.61	0.54	0.30	0.72
NB	57.76	0.54	0.58	0.53	0.28	0.01
BN (p=1)	57.53	0.53	0.58	0.53	0.27	0.00
ZeroR	52.18	0.27	0.52	0.36	0.32	0.00

From Table 2, the model generated from ZeroR algorithm has the least correctness, recall, and f-measure values since ZeroR algorithm is the simplest classification technique which predicts the result from the majority class and ignores all predictors. For the most efficient model, the model generated from J48graft algorithm has the highest correctness, recall, and f-measure values. For Mean Absolute Error, it's not the least but it is at a good level as compared to other algorithms. The processing time of each algorithm is not much different since the time is measured in seconds. The model created from J48graft algorithm is selected to be used as a part of the recommendation system for learning activities. Not many researchers have studied the effectiveness of J48graft algorithm; however Kladchuen and Sanrach (2018) compared the performance of J48graft, Naïve Bayes and Rule Induction and indicated that J48graft has the best performance.

3. Developing a recommendation system for learning activities

The recommendation system for learning activities is designed and developed by applying educational theories together with computer principles. By using the aforementioned theories and principles, the system is synthesized into 4 modules: 1) student data management module 2) learning activity module 3) prediction and recommendation module and 4) study achievement module as shown in the designing framework in Figure 2.

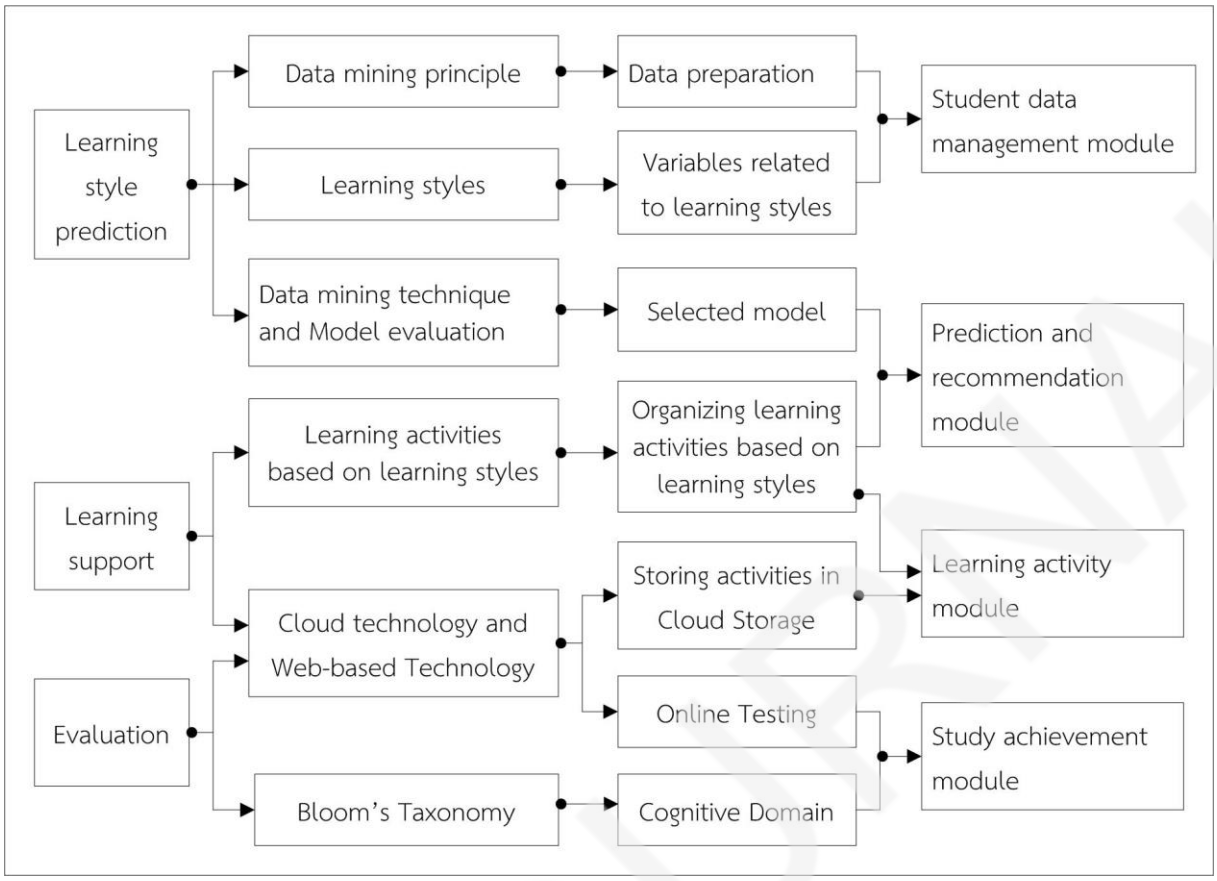

Figure 2 Designing Framework

The details of 4 modules of the system are as follows.

3.1 The student data management module consists of 3 parts: 1) adding student data 2) changing student data and 3) converting the data automatically. This module is used to store student data which are the 7 variables related to learning style: gender, faculty, year, GPA, highest score subject, previous qualification and previous study plans. These data will be automatically converted to the format suitable for data mining.

3.2 The learning activity module consists of 2 parts: 1) adding learning activities and 2) changing learning activities. Learning activities conforming to learning style based on the principle of Honey and Mumford are added to the system according to the guidance provided by the system. Learning activity documents are stored in a google drive which is a cloud storage, by saving the documents' links in the system.

3.3 The prediction and recommendation module consists of 3 parts: 1) predicting learning styles 2) suggesting learning activities and 3) learning activities of individual student. For predicting learning styles, students' data stored in the system are used as inputs for the model built with J48graft algorithm to predict each student's learning style. For suggesting learning activities, students' learning styles will be connected to the learning activities according to the principle of Honey and Mumford.

3.4 The study achievement module consists of 3 parts: 1) adding study achievement test 2) changing study achievement test and 3) measuring study achievement. Study achievement tests created with a google form are added to the system by saving the forms' links in the system.

Examples of the system's screens showing the results of predicting learning styles and suggesting learning activities are shown in Figure 3 and Figure 4 respectively. Data in the screens are fictitious.

For assessing experts' satisfaction with the recommendation system for learning activities in 3 facets: 1) system design 2) system components and 3) system performance, overall average score is 4.60, which satisfaction is at the most level.

Figure 3 The system's screen showing the result of predicting learning styles

Figure 4 The system's screen showing the result of suggesting learning activities

4. Comparing learners' study achievement

For the result of comparing learners' study achievement, it is found that the students studying by the system organizing learning activities have higher study achievement than the students studying by an instructor organizing learning activities at the statistical significance level of .05 as shown in Table 3. This may be because an instructor often organizes the same learning activity for all students without considering each student's learning style.

Table 3 The results of comparing study achievement

Group of students	N	$\bar{X}$	S.D.	t	Sig
Students using the system organizes learning activities	28	45.96	4.70	3.98	.00
Students whose instructor organizes learning activities	27	41.04	4.47		

* Statistically significant level at .05

This result is consistent with the research of Sintia et al (2019) indicating that learners who follow the learning style have significantly higher study achievement than those who do not follow learning style. While Hadriana et al (2019) stated that studying based on the learning style will lead to better study achievement.

Conclusion

There are 7 variables related to the learning style: gender, faculty, year, GPA, highest score subject, previous qualification and previous study plans. For testing the performance learning classification models created by data mining whose inputs are the 7 variables and the learning style measurement questions based on the principle of Honey and Mumford, It's found that the most effective model is the one built with J48graft algorithm. Thus this model is chosen as a part of the recommendation system for learning activities consisting of 4 main modules: 1) student data management module 2) learning activity module 3) prediction and recommendation module and 4) study achievement test module. In the experiment to compare learners' study achievement, the students studying by the system organizing learning activities have higher study achievement than the students studying by an instructor organizing learning activities with statistical significance at the .05 level.

Contribution

This research provides guidelines for the application of data mining techniques in education. In particular, data mining techniques are utilized to enhance students 'learning. Moreover, this research provides a learning style classification model and a recommendation system for learning activities that other educational institutions can apply for their purposes.

Limitations

The limitation is that this research only studied physical variables such as GPA, and highest score subject but the research did not study other psychological variables such as motivation, personality and learning goals.

Suggestion

1. Psychological variables such as motivation, personality and learning goals, should be studied to predict learners' learning style.

2. Other principles of learning style such as Fleming's learning style (Fleming and Bonwell, 2019) or Grasha's learning style (Grasha, 1996), should be studied.

References

Berrar, D. (2019). Bayes' Theorem and Naive Bayes Classifier. *Encyclopedia of Bioinformatics and Computational Biology.* 1(1),403–412.

Bloom, B. A. (1956). *Taxonomy of Educational Objective Handbook I: Cognitive Domain.* New York : David Mc Kay Company Inc.

Chaijaroen, S., Khanjug, I., Samat, C., Kwangmuang, P. and Wongtatum, P. (2018). Design and Development of Learning Innovation Enhancing the Learners' Knowledge Construction based on Thai Wisdom and Thai Living on the topic of Force and Pressure for Fifth Grade Students. *Technical Education Journal King Mongkut's University of Technology North Bangkok.* 9(3),186–197.

Chongsomchai, S. (2018). The Development of Web-Based Instruction in 31247 Introduction to C Programming Language for Students of Matthayomsuksa 4 Using TGT Teaching Techniques. *Journal of Educational Technology and Communications Faculty of Education Mahasarakham University*. 1(3),79–91.

Chuangchai, T. (2016). The students' learning style analysis on students' senior project by using data mining. *Journal of Graduate Studies Valaya Alongkorn Rajabhat University.* 10(2),53–62.

Eyasu, K., Jimma, W. and Tadesse, T. (2020). Developing a Prototype Knowledge-Based System for Diagnosis and Treatment of Diabetes Using Data Mining Techniques. *Ethiopian Journal of Health Science.* 30(1),115–124.

Fleming, N. and Bonwell, C. (2019). *How Do I Learn Best? A student's guide to improved learning.* 2nd Edition. Christchurch, NZ : VARK Learn Ltd.

Grasha, A. F. (1996). *Teaching with Style: A Practical Guide to Enhancing Learning by Understanding Teaching and Learning Styles.* Pittsburgh, PA : Alliance Publishers.

Hadriana, H., Primahardani, I. and Mahdum, M. (2019). Learning Style and Learning Achievement of Students of FKIP Universitas Riau In Learning English. *Journal of Educational Sciences.* 3(3),340–352.

Honey, P. and Mumford, A. (1982). *The Manual of Learning Styles.* Maidenhead, UK : Peter Honey.

Iaosanurak, C., Chunlasewok, C., Buabangplu, P., and Suksangprasit, J. (2017). The relationships between gender majors' subject academic achievement and personality with learning style of student teachers. *Journal of Graduate Studies Valaya Alongkorn Rajabhat University.* 11(1),33–43.

Jaleel, S. and Mary Thomas, A. (2019). *A Glimpse on Learning Styles. In Learning Styles Theories and Implications for Teaching Learning (p. 3).* SAN JOSE, CA: Horizon Research Publishing.

Kladchuen, R. and Sanrach, C. (2018). An Efficiency Comparison of Algorithms and Feature Selection Methods to Predict the Learning Achievement of Vocational Students. *Research Journal Rajamangala University of Technology Thanyaburi.* 17(1),1–10.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development.* New Jersey : Prentice-Hall.

Kwangmuang, P. (2018). The result of learner's critical thinking development used with constructivist learning innovation to enhance knowledge construction and critical thinking for undergraduate student. *Panyapiwat Journal.* 10(1),175–184.

Li, Z., Zhang, Q. and Zhao, X. (2017). Performance analysis of K-nearest neighbor, support vector machine, and artificial neural network classifiers for driver drowsiness detection with different road geometries. *International Journal of Distributed Sensor Networks.* 13(9),1–12.

Mirza, S., Mittal, S. and Zaman, M. (2016). A Review of Data Mining Literature. *International Journal of Computer Science and Information Security.* 14(11),437-442.

Na Nongkhai, L. and Keawkiriya, T. (2016). A Development of the Learning Method of Problem-Based Learning with Cooperative Learning on Cloud Computing by Google Apps. *Romphruek Journal.* 34(3),11–34.

Pantho, O. (2016). *A Synthesis of an Adaptive e-Learning Model with e-Mentoring System Using VARK Learning Style by Data Mining Methodology.* Bangkok : King Mongkut's University of Technology North Bangkok.

Prasada Rao, K., Chandra Sekhara, M.V.P. and Ramesh, B. (2016). Predicting Learning Behavior of Students using Classification Techniques. *International Journal of Computer Applications.* 139(7), 15–19.

Radhwan H. A. Alsagheer, Abbas F. H. Alharan and Ali S. A. Al-Haboobi. (2017). Popular Decision Tree Algorithms of Data Mining Techniques: A Review. *International Journal of Computer Science and Mobile Computing.* 6(6),133–142.

Sangeorzan, L. (2019). Effectiveness analysis of ZeroR and J48 classifiers using WEKA toolkit. Bulletin of the Transilvania University of Bra¸sov, Series III: Mathematics, *Informatics, Physics.* 12(61),481–486.

Sintia, I., Rusnayati, H. and Samsudin, A. (2019). VARK learning style and cooperative learning implementation on impulse and momentum. *Journal of Physics: Conference Series.* 1280(5),1–7.