

# A Study On Factors And Using Machine Learning For Learning Analytics In Computer Programming Of Junior High School Students

Nontachai Samngamjan<sup>1</sup>, Wudhijaya Philuek<sup>2</sup>, Pachara Malangpoo<sup>3</sup>, and Sirirat Janyarat<sup>4</sup>

<sup>1,2,3</sup>Computer and Educational Technology Program, Faculty of Education, Nakhon Sawan Rajabhat University, Mueang Nakhon Sawan District, Nakhon Sawan Province, 60000 Thailand

<sup>4</sup>Srinakharinwirot University, Bangkok, 10110 Thailand

---

## Abstract

This research aims 1) to study factors that affect the learning of Computer Programming of Junior High-school students, and 2) to study the Machine Learning Techniques in analytics of factors that affect the learning of Computer Programming of Junior High-school students. Data were used came from the survey of 411 Junior High-school students. The preliminary analysis using the document research and collect from 2 round of experts confirmed and leading the selected factors to Exploratory Factor Analysis (EFA) and then to analytics by 5 Machine Learning Techniques which were K- Nearest Neighbors Technique, Logistic Regression Technique, Decision Tree Technique, Support Vector Machine Technique, and Naïve Bayes Technique.

Results found that,

1) The Exploratory Factor Analysis (EFA) by Correlation Coefficient found that there were 11 factors of Instructor, 4 factors of learning environment, 5 factors of learning media, and 11 factors of learner which affect to learning Computer Programming.

2) The results of using Machine Learning Techniques found that there were Logistic Regression Technique, Decision Tree Technique, and Naïve Bayes Technique showed the highest accuray in Gender class, and there were Decision Tree Technique, and Naïve Bayes Technique showed the highest accuray in Level of study class.

**Keywords:** Machine Learning Techniques, Learning Computer Programming, Learning Achievement

---

## I. INTRODUCTION

In Thailand at present, the core curriculum has changed the Career and Technology core subject area and Science core subject area by moving the Information and Communication Technology learning subject from the Career and Technology core subject area to the Science core subject area and changed its name to "Technology learning area" and defines two standards, namely Standard 4.1 Design and Technology and Standard 4.2 Computing Science (Thai Ministry of Education, 2017). The

standard 4.2 Computing Science (CS) is a content that is closed and teaches the thinking process. As a result, teachers who have graduated in Computing Science or have not graduated directly can teach children. The teaching method is to use 5 Computational Thinking skills, namely Problem Solving/Planning, Pattern Recognition, Data Processing, teaching children to think of Algorithms, and finally to Coding/ Programming (Theeaksorn, 2018). Moreover, in the era that is full of information whether it is electronic transactions, searching for online trading information, using social networks, etc., these behaviors create a huge amount of new information and make the information that leads to the creation of a trading system and cause the decision support in those organizations slows down in processing of data as well. Even if the database has been modified to increase the processing and accessibility of the data, but as the business system continues to operate, the data will increase until the much larger volume, which further reduces the efficiency of data access and processing, thereby giving rise to a concept known as Big Data (Pongsuk, 2017). In addition, Sakdulyatham (2010) presented the application of data mining techniques to predict student performance by using data from the Faculty of Business Administration student database from 2006 - 2008 to analyze useful information. The decision was made in the study counseling of the advisors by analyzing the study behaviors. Classification of students in each field of study provides information about the skills necessary to study in each field of study based on various factors. Data Mining is about finding all relationships and patterns that actually exist in a database but are hidden within large amounts of data. Related to economics, finance, and education, etc. Data Mining has a number of different techniques for solving problems, each of which has specific properties. Giudici (2003) defined it as Data Mining is the process of selecting and exploring data as well as modeling data to find patterns and correlations from large amounts of data to obtain useful results from the definition of Data Mining is the process of extracting or searching for information based on large amounts of data with pattern recognition technology, as well as Statistics and Mathematical techniques, and modeling of data to obtain knowledge or hidden information. in a large database.

The importance above shows that in learning management, teachers or instructors need to train learners to have clearer practical competency through the process of project development or innovation development, which can be seen from the indications of learning in each Grade level, students must be able to develop projects from Grade 7, and in learning management, Coding/ Programming is something that learners have to learn and develop projects from Coding/ Programming that teachers need to find a way to manage teaching or develop a way for students to learn effectively. This research objectives were to study factors that affect the learning of Computer Programming of Junior High-school students, and 2) to study the Machine Learning Techniques in analytics of factors that affect the learning of Computer Programming of Junior High-school students.

Researchers have studied the concepts, theories and related research. to be used as a guideline for conducting research, consisting of,

#### 1. Machine Learning

“Learning is a sign of change in the system. This makes it possible to do the same job more efficiently the next time” (Simon, 2002).

“Learning is creating or improving an example of a topic on how to use experience” (Michalski, 2003).

Learning is an essential process for the survival of living beings. which this learning process begins to develop from birth when learning from the information we put together with the determination and create a model that can be used to make decisions (Bastanlar and Ozuysal, 2014).

Samuel (1959) was one of the pioneers of Machine Learning attesting to the fact that machines learn to perform better than those who initially programmed.

Mitchell (1997) gave the meaning of Machine Learning It was a system that could learn from examples on its own without the input of a programmer. This advance came with the idea that computers could learn only from data in order to produce. Accurate results Machine Learning is the science of studying and building algorithms that learn data and predict it.

## 2. Machine Learning Patterns or Techniques

2.1 Decision Tree is a tool that helps to define the scope of a problem and help create possible alternatives to the solution in which the decision tree is input to create a forecast. Forecasting in the form of decision trees has a supervised learning approach, which is to create a classification forecasting model from a predefined sample of data. also known as and apply the predicted model to the test data (Test Set) to test the accuracy of the forecast model (Sara, 2007).

2.2 Naive Bayes is a probabilistic grouping technique based on Bayes' Theorem and the assumption that the occurrence of grouping events is independent of which learning is classified by process. Naive Bayes is widely used in machine learning research.

2.3 K-Nearest Neighbor (KNN) is a classification based on k closest properties, which is a method of grouping into new observations close to the learned dataset.

2.4 Support Vector Machine (SVM) is a supervised algorithm used in machine learning classification and regression. In which SVM's algorithm finds the boundary between classes, maximizing the perpendicular distance between the boundary and the point of this boundary class is known as a hyperplane.

2.5 Logistic Regression (LR) is a classification algorithm used when output labels are binary (0 and 1), an instructor-led machine learning model used for classification, which is built on regression. linear by Linear Regression generated using available input data with actual result variables (Hendel, 2006).

## II. METHODOLOGY

This research was exploratory research and testing of Machine Learning techniques in analyzing factors in computer programming learning. The research process is divided into 5 steps as follows:

### 1. Procedures for conducting research

1.1 A study collecting factors affecting computer programming learning

1.2 Interpretation of factors The grouping of questions for each factor was sent to experts for review and to be revised. and send it to experts to check the quality of the questionnaire

1.3 Analysis of each question item used for prognostication in computer programming learning ability

1.4 Checking the structural validity of the questionnaire

### 2. Research scope

The data collection used for analysis results from the results of lower secondary school students under the Nakhon Sawan Secondary Educational Service Area Office answered a survey on learning computer programming problems.

### 3. Research tools

#### 3.1 Analytical data

3.1.1 Survey on factors affecting computer programming learning

3.1.2 Data Analysis Program

#### 3.2 Working equipment

3.2.1 Computer

3.2.2 Statistical Packaged Programs

- 3.2.3 Data Analysis Program
- 3.2.4 Microsoft Excel 2019
- 3.2.5 Techniques used in data analysis
- 3.2.6 Anaconda Python

4. Methods of collecting information

- Step 1 Submit a request for research data collection
- Step 2. Archive the survey with students who are willing to answer the survey.

5. Data analysis by CRISP-DM process, which has the following research steps:

5.1 Business Understanding

5.2 Data Understanding by collect Initial data as Figure 1, and describe data. The data representation used in the modeling uses the data from The results of the test in the mathematics course and computer and self-assessment results from 4 factor groups, which are data from Instructor factor learning environment factors in learning management media and student factors From the questionnaire of Andadet (2011); Mad-ada (2016); Konkhao (2015); Na Takuathung et al. (1997); Abdunabi and Ku (2019); O`zden and Tezer (2018); Lahtinenand Jä`rvinen (2005); MD Derss and Mohamad Ali (2012) total number 354 itemsby grouping the factors to be sent to 3 experts to assess the quality of the questionnaire by finding the Index of Item-Objective Congruence (IOC) and collecting data to find the  $\bar{X}$  and S.D. value individually and find the sum of each factor, and take it to Exploratory Factor Analysis by setting variables in the analysis to cut off unrelated factors to only those that are the most consistent. The results of the analysis of questions and the results of the Eigenvalue at is greater than 1, resulting in the result having 5 elements of 31 items.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
2	ime	thilainkaylunla	gradet1	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13						
3	1	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
4	1	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
5	1	3	3.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
6	1	3	2.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
7	2	3	1.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
8	2	3	2.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
9	2	3	1.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
10	2	3	3.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
11	2	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
12	2	3	3.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
13	2	3	3.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
14	1	3	2.5	3.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
15	1	3	1.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
16	2	3	3.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
17	2	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
18	1	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
19	1	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
20	2	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
21	1	3	2	3.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
22	2	3	1	2.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
23	2	3	3.5	3.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
24	2	3	2.5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	

Figure 1 shows the raw data.

5.3 Data Preparation by Data Selection, Data Cleaning, Data Construct, Data Integrate, and Data Transformation.

5.4 Modeling

It is the selection of the appropriate model. Improve variables for best results. At this stage, the researcher selects the following data mining techniques and algorithms.

- Support Vector Machine (SVM) Technique
- K - Nearest Neighbors (KNN) Technique
- Logistic Regression (LR) Techniques
- Decision Tree (DT) Technique
- Naïve Bayes (NB) Technique

5.5 Evaluation by considering the Support, Accuracy, Precision, Depth (Recall), F1-Score values, and correlation and consistency.

5.6 Deployment

### III. RESULTS

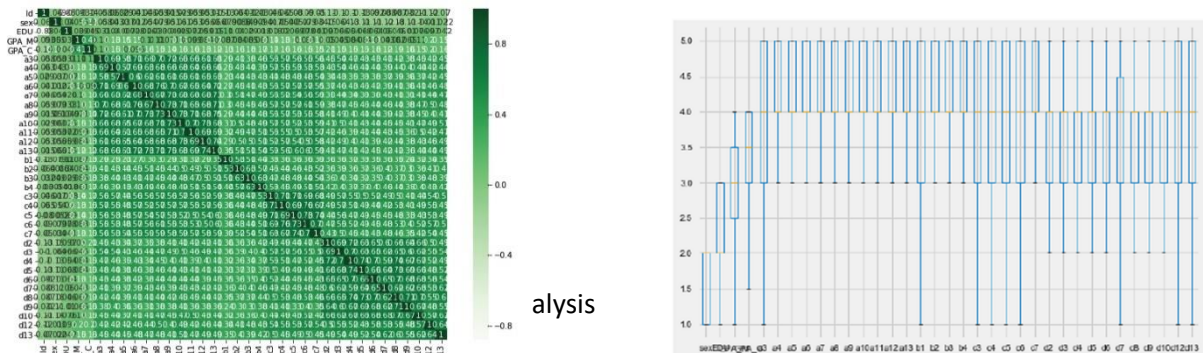
1. The results of a study on factors affecting the learning of computer programming among junior high school students using machine learning techniques. The results are shown in Table 1.

**Table 1** shows questions in each factor group.

Instructor Factors	Learning Environment factors
<ol style="list-style-type: none"> <li>1. Teachers have a teaching method that is easy to understand, clear, and creates an inviting atmosphere for learning.</li> <li>2. Teachers have the ability to convey knowledge in a clear sequence.</li> <li>3. Teachers come to teach on time.</li> <li>4. Teachers have the ability to make students develop their ideas.</li> <li>5. Teachers have good personality and are reliable.</li> <li>6. Teachers have teaching methods to stimulate students to develop thinking skills. and the pursuit of additional knowledge</li> <li>7. Teachers pay attention and take care of learners thoroughly.</li> <li>8. Teachers always interact with learners.</li> <li>9. Teachers are interested in students' problems.</li> <li>10. Teachers help learners when there is a problem with equal work.</li> <li>11. The teacher gives feedback on the progress of the learners.</li> </ol>	<ol style="list-style-type: none"> <li>1. The number of computers is sufficient and ready to use.</li> <li>2. The computer classroom is well lit, e.g. not too bright and too dark. no glare affecting the eyes</li> <li>3. The school provides appropriate computer rooms for students.</li> <li>4. The environment is conducive to team learning.</li> </ol>
Instructional Media Factors	Learner Factors
<ol style="list-style-type: none"> <li>1. The language used in the media is easy to understand.</li> <li>2. There is a user manual that is easy to understand.</li> <li>3. Learning materials help students gain knowledge and have more programming skills.</li> <li>4. Learning materials help learners to be interested in learning.</li> <li>5. The learning materials are comprehensive and consistent with the learning content.</li> </ol>	<ol style="list-style-type: none"> <li>1. I can write code blocks right.</li> <li>2. I can understand the structure of programming.</li> <li>3. I can fix complex programming errors.</li> <li>4. I can organize and design programs in a modular manner.</li> <li>5. I can manage my programming time to meet the deadline.</li> <li>6. I can write programs only those that are familiar to me.</li> <li>7. I can organize and design modular programs.</li> </ol>

	<p>8. I can write an algorithm.</p> <p>9. I can find bugs in my own program.</p> <p>10. I use my computer to get information from the internet.</p> <p>11. I can mentor or talk to a teacher or friend.</p>
--	---

From Table 1, the exploratory analysis results were obtained by analyzing the Exploratory Factor Analysis (EFA) data to determine the exploratory components. Then the question item with a component value greater than 0.1 was selected with the remaining factors as follows. Instructor factors were 11 items, learning environment factors were 4 items, media factors for learning management were 5 items, and student factors were 11 items, a total of 31 items.



2. Comparison results of testing 5 techniques of gender and class as shown in Table 2 and Table 2.

**Table 2** Results of the test technique of (Gender)

Machine Learning Techniques	Accuracy: Mean	S.D.
K nearest Neighbors (KNN)	0.6975	0.3870
Support Vector Machine (SVM)	0.9925	0.0787
Logistic Regression (LR)	1.0000	0.0000
Decision Tree (DT)	1.0000	0.0000
Naïve Bayes (NB)	1.0000	0.0000

From Table 2, the results of testing 5 techniques for gender showed that the K- Nearest Neighbors (KNN) technique had an accuracy of 0.6975 with a standard deviation of 0.3870. The Support Vector Machine (SVM) technique had an accuracy of 0.9925 with a standard deviation of 0.0787. Logistic Regression (LR), Decision Tree (DT) and Naïve Bayes (NB) techniques have an accuracy of 1.0000. A standard deviation of 0.0000. The most common were Logistic Regression (LR), Decision Tree (DT) and Naïve Bayes (NB) techniques with an accuracy of 1.0000 and a standard deviation of 0.0000.

**Table 2** Results of the test technique of (Class)

Machine Learning Techniques	Accuracy: Mean	S.D.
K nearest Neighbors (KNN)	0.6650	0.3940
Logistic Regression (LR)	0.8725	0.2692
Support Vector Machine(SVM)	0.9150	0.2404

Decision Tree (DT)	1.0000	0.0000
Naïve Bayes (NB)	1.0000	0.0000

From Table 3 the results of testing 5 techniques for class showed that the K Nearest Neighbors (KNN) technique had an accuracy of 0.6650 with a standard deviation of 0.3940 and the Logistic Regression (LR) technique with an accuracy of 0.8725. The standard deviation was 0.2692, the Support Vector Machine (SVM) technique had an accuracy of 0.9150, the standard deviation was 0.2404, the Decision Tree (DT) and Naïve Bayes (NB) technique had an accuracy of 1.0000. The standard deviation was 0.0000. The techniques with the highest accuracy were the Decision Tree (DT) and Naïve Bayes (NB) techniques with an accuracy of 1.0000. The standard deviation was 0.0000.

#### IV. DISCUSSIONS

Discussion of the results of this research consisted of two topics: 1) the results of the study of factors affecting computer programming learning, and 2) the results of the study of Machine Learning techniques in the analysis of learning factors in learning computer programming for junior high school students as follows:

1. The results of the study of factors affecting the learning of computer programming found that the results of the study of factors affecting computer programming learning consisted of 31 factors, 11 of which were questions on the teacher factor. 4 learning environments, 5 learning media factors, and 11 learner factors. In terms of the factors that the instructors are involved in causing students to achieve learning achievements, the instructors must have an understanding of the content, creating a learning management atmosphere and having the ability to make the learners develop their skills. thinking and having a thorough attention to learners which has been said by scholars that the teaching factor is related to the development of knowledge for learners, which teachers must have an understanding of the learners and should have knowledge of the content taught by Problems during studying in that class (Kruethong, 2012). In terms of learning environment factors, many scholars have stated that learning environment factors are related to academic achievement. whether it is teaching techniques, classroom atmosphere, media and teaching technology. Therefore, the setting of the learning environment has an effect on enhancing physical, mental and intellectual development (Hosuwan, 2005; Taewakha, 2008). In terms of learning management media factors, students will be asked to assess whether the material used by the teacher has an easy-to-understand language level or not by using learning management media, it can help students to learn more and more interesting. Many scholars have stated that the use of traditional media such as books and notes or the use of presentations may not be an effective teaching aid for students to learn programming better (Bennedsen and Caspersen, 2005). Therefore, in learning to programming, learning materials are one of the learning tools that can promote students' success, such as the use of Visualization Media. The use of interactive media in simulations or learning through experience and using examples (Konenecki, 2014). In terms of student factors, questions for each factor will ask about the assessment of the students' abilities. Asking about learning programming and how competent students are. Many scholars have studied factors that affect the prognosis of success in learning programming. The studying or learning is one of the factors that enable students to be successful, such as the way students learn, the time spent studying, and motivation (Kinnunen and Malmi, 2006), etc. The most common problems encountered in learning to program include a lack of problem-solving skills and a lack of the logical

reasoning underlying abstract thinking (Gomes and Mendes, 2007; Robins, 2019; Savage and Piwek, 2019).

2. The results of a study on machine learning techniques in analyzing factors in learning computer programming for junior high school students.

By comparing the efficiency and accuracy of the five forecasting modeling techniques, the Logistic Regression technique has an accuracy of 83%, the Support Vector Machine technique has an accuracy of 83%, and the Decision Tree technique has an accuracy. at 83 percent and the Naïve Bayes technique had an accuracy of 83%. The choice of technique should be based on the data and the selection of different techniques to use in forecasting. by the research of Kamjornkittikhun (2018); Chutipascharoen and Saenrat (2018) found the results of the two techniques, Naïve Bayes and Neural Network, which the results of the Neural Network classification had higher directness values but took time to process. This may be a method that is not suitable for large amounts of data, making the analysis time-consuming to process. However, the comparison of Naïve Bayes, the process is simple can be processed quickly and the results are good as well, and the techniques used in the experiment should study the relevant theories. and other principles to increase efficiency to be more efficient. Hence, Machine Learning that is based on Bayes Theorem's principles. It is an algorithm for classifying information by learning the problems that arise in order to create new classification conditions, classify the data by using probabilities and calculating the assumption-based probability distribution for the calculated data to modify the distribution. This has the effect of increasing or decreasing the probability of the data. The resulting new data and the model set for the data are adapted according to the new data in conjunction with the original data based on the Naïve Bayes using probabilistic calculations that are used to make predictions. As a classification solution that can predict outcomes, it analyzes the relationships between variables in order to create probabilistic conditions for each relationship. Naive Bay is an effective classification method, there is a simple working algorithm that is suitable for the case of large sample sets and their attributes are independent of each other (Thurachon, Sumethawattanapong and Songsuk, 2014).

## V. REFERENCES

- Abdunabi, H. and Ku, Y. (2019). **Towards Enhancing Programming Self-Efficacy Perceptions Among Undergraduate Information Systems Students**. University of Northern Colorado, Greeley, USA.
- Andadet, N. (2011). **The development of a virtual classroom for academic use. Computers in constructivist business for Undergraduate students Faculty of Business Administration, Rajamangala University of Technology Isan**. Proceedings of The 23<sup>rd</sup> National Graduate Research Conference at Rajamangala University of Technology Isan 23-24 December 2011, 935-940.
- Bastanlar, Y. and Ozuysal, M. (2014)**Introduction to Machine Learning**, pages 105-128. Humana Press, Totowa, New Jersey, USA
- Bennedsen, J and Caspersen, M. E. (2005). **An Investigation of Potential Success Factors for an Introductory Model-Driven Programming Course**. Proceedings of ICER 2005 The First International Computing Education Research Workshop, October 1-2, 2005, Seattle, WA, USA.



- Chutipascharoen, A. and Saenraj, C. (2018). **Comparison of Algorithm Performance and Selection of Appropriate Characteristics for Predicting the Chances of Success in International Money Transfers of Individuals**. Research Journal, KhonKaen University.
- TawanDevaksorn. (2018). **Upgrading the quality of Thai education**. Post Today News (15 Jun 2018)
- Giudici, P. (2003). **Applied Data Mining: Statistical Methods for Business and Industry**. Wiley, Hoboken.
- Gomes, A., & Mendes, A. J. (2007). **Learning to program-difficulties and solutions**. Paper presented at the International Conference on Engineering Education–ICEE. Retrieved from [https://www.researchgate.net/profile/Anabela\\_Gomes2/publication/228328491\\_Learning\\_to\\_program\\_-\\_difficulties\\_and\\_solutions/links/02e7e52389017b9984000000.pdf](https://www.researchgate.net/profile/Anabela_Gomes2/publication/228328491_Learning_to_program_-_difficulties_and_solutions/links/02e7e52389017b9984000000.pdf)
- Hendel, D. (2006) **Efficacy of participating in a first-year seminar on student satisfaction and retention**. Journal of College Student Retention: Research, Theory Practice, 8(4):413-423. [https://commons.wikimedia.org/wiki/File:Linear\\_regression.svg](https://commons.wikimedia.org/wiki/File:Linear_regression.svg) [retrieved 2020-04-18]. <https://cvxopt.org/examples/book/logreg.html> [retrieved 2020-04-19] <https://www.pexels.com> [retrieved 2020-04-17].
- Hosuwan, P. (2005). **Organizing learning conditions, the door to success in learning development**. Department of Educational Technology, Faculty of Education, Burapha University: Chonburi
- Kamchonkittikhun, W. and Khemaphataphan, C. (2018). Anomaly analysis of electricity usage data by Naif Bay. The 8th National Academic Conference on Information Technology, pages 404-408.
- Kittiporn, M. (2019). **Educational guidance system and machine learning techniques**. Journal of Information Science. Pages 92-109.
- Konecki, K. T. (2014). **“Socjologia emocji według Thomasa Scheffa” [Sociology of Emotions by Thomas Scheff]**. Pp. 11-38 in *Emocje w życiu codziennym. Analizakulturowych, społecznych i organizacyjnych warunków ujawniania i kierowania emocjami*, edited by K. T. Konecki and B. Pawłowska. Lodz: Wydawnictwo Uniwersytetu Łódzkiego.
- Konkhaio, K. (2015). **A study of the learning environment of students at Sattasamut School. Under the Office of Secondary Education Service Area 10**. Master's Degree Thesis, Burapha University.
- Kotsiantis, S.B. (2007) **Supervised machine learning: A review of classification techniques**. Informatica (Ljubljana), 31:249-268
- Kruethong, T. (2012). **Development of Knowledge and Integration of Teaching Methods of Science Teachers. junior high school with a contextual approach to teaching energy**. Kasetsart University, Bangkok.
- Lahtinen, A. M. and Järvinen. (2005). **A study of the difficulties of novice programmers**. Tampere University of Technology
- Mad-ada, U. (2016). **Developing Computer Programming Skills Using Skills Training Kit. computer programming with a collaborative learning approach for 2<sup>nd</sup> year students in Software Engineering**. Walailak University library journal Thaksin University.
- MD Derss and Mohamad Ali. (2012). **Difficulties in learning programming: Views of students**. Computing and Creative Industry University.
- Michalski, R. (2003) **Inferential Theory of Learning and Natural Induction**. Retrieved (21/08/2021) from <https://www.mli.gmu.edu/michalski/montreal07-08-03.pdf>
- Mitchell, T. (1997) **Machine Learning**. New York: McGraw Hill, New York, USA.
- Na Takuathung, O. et al. (1997). **Guidelines for the development of computer teaching in secondary schools**. Chulalongkorn University

- Özden and Tezer. (2018). **The Effect of Coding Teaching on Students' Self-Efficacy Perceptions of Technology and Design Courses**. Education Faculty, Near East University.
- Pongsuk, T. (2017). **A Study of Data Mining Techniques in Prediction of Academic Results and Learning Behavior of Students of Computer Studies Program at Northern University**. Nakhon Sawan Rajabhat University : Nakhon Sawan
- Robins, A. (2019). **Novice programmers and introductory programming**. The Cambridge Handbook of Computing Education Research, Cambridge Handbooks in Psychology, 327-376. Retrieved from <https://www.otago.ac.nz/computer-science/otago706761.pdf>
- Sakdulyatham, R. (2010). **Using Data Mining Techniques to Build a Knowledge Base for prediction Student achievement** Ratchaphruek College Journal: Ratchaphruek College
- Savage, S. and Piwek, P. (2019). **Full report on challenges with learning to program and problem solve: an analysis of first year undergraduate Open University distance learning students' online discussions**. Retrieved from [http://oro.open.ac.uk/68073/1/IOC\\_TM112\\_Python\\_Help\\_Forum\\_Research.pdf](http://oro.open.ac.uk/68073/1/IOC_TM112_Python_Help_Forum_Research.pdf)
- Sara, C. (2007). **Classification of graduation status groups, studied by decision tree model**. Bangkok: Department of Computer Science and Information, Graduate School King Mongkut's University of Technology North Bangkok.
- Simon, H. (2002) **What We Know About Learning Retrieved** (21/08/2021) from <http://158.132.155.107/posh97/private/learning/learning-Simon.htm>
- Thai Ministry of Education. (2017). **Basic Education Core Curriculum 2008 (Revised 2017)**. Bangkok: Teachers Council Ladprao Printing House.
- Thewakha, P. (2008). **Comparison of Learning Achievement on Life and Environment. of students in Mathayomsuksa 3 who manage to learn by using conventional teaching methods and setting up the learning environment and atmosphere together with the search-for-knowledge teaching method**. Master of Education, Technology Education, Graduate School, Silpakorn University.
- Thurachon, W., Sumethawattanapong, W., and Songsuk, N. (2014). **The classification system of para rubber by using Naïve Bay classifier**, pp. 20-25, National Academic Conference. Computer and Information Technology No. 10, King Mongkut's University of Technology North Bangkok, Bangkok.
- Zhang, X., Chen, L., Guo, Z.-H., and Liang, H. (2019). **Identification of human membrane protein types by incorporating network embedding methods**. IEEE Access 7, 140794–140805. doi: 10.1109/ACCESS.2019.2944177