THE EFFECTIVE PREDICTION MODEL BASED ON EMPLOYMENT POSITIONS FOR COMPUTER CURRICULUMS USING CLASSIFICATION TECHNIQUES

Jittima Silprachawong

Dept. of Information Technology, Faculty of Science and Technology, Suratthani Rajabhat University, 272 Surat-Nasarn Rd., Muang, Surat Thani, 84100, Thailand
jittima.sil@sru.ac.th

Kannika Kaewchuea

Dept. of Computer Education, Faculty of Education, Suratthani Rajabhat University, 272 Surat-Nasarn Rd., Muang, Surat Thani, 84100, Thailand
kannika.kae@sru.ac.th

Asok Srisawat

Dept. of Computer Education, Faculty of Education, Suratthani Rajabhat University, 272 Surat-Nasarn Rd., Muang, Surat Thani, 84100, Thailand
asok.sri@sru.ac.th

Abstract

Data mining techniques were used to build a prediction model prototype based on employment positions for computer majors in the Faculty of Science and Technology, Suratthani Rajabhat University. This study investigated the impact factors of employment positions of computer careers, compared model performances that consisted of six classification techniques and developed a web application using the effective prediction model that could be implemented in four majors using data collected from the graduates of sessions 2016-2020. The results of this research demonstrated that factors of seven courses and cumulative grade point average predicted employment positions. The models with best performances were rule-based and naive Bayes methods with a classification accuracy of 89.66%, precision of 90.48%, and recall of 90.00%. Therefore, a web application of the prediction model was developed using PHP programming. In practice, this application can suitably predict employment positions for computer majors.

***Keywords*: Prediction model; IT employment position; model performance; classification techniques.**

1. Introduction

Currently, many university curriculums have been designed for graduated students to develop academic qualities in areas of labor market demand. Therefore, universities have developed continuous instructional processes related to national demands. In addition, information technology has changed rapidly over the past periods, changing the qualities of graduates required by the labor market at both organizational and national levels. However, there have not been enough computer professionals to meet the demand of the companies that require them, especially those in the computer field. Furthermore, an important problem is that the graduates work in other employment positions that are not related to their bachelor's degrees in computer fields.

Overall, in a survey on employment situations of graduates conducted in 2018, the percentage of graduates who were already working was 69.93%, working and studying was 1.49%, not working and not studying was 25.08%, and studying 3.50% [Nunman, T. (2017)]. Most of the graduates were employed (71.42%). Nevertheless, the research mentioned a problem that the graduates were working in different fields than the fields of their bachelor's degrees; this remains unstudied. In Thailand, research has not been conducted on the number of all students in higher education who graduated and work in different fields than the fields of their bachelors' degrees. Foreign research indicates that graduates who work in different fields comprise one third of all graduates, especially in the fields of social sciences or humanities, which have no obvious career patterns. However, these estimations are higher than the results of the studies. Hence, it is necessary to provide information for decision

making and creating educational plans to achieve the goal of obtaining employment positions in the same field as the bachelor's degree.

Moreover, the qualifications of graduates should be appropriate for the computer field and meet the needs of the labor market or entrepreneurs. Thus, all courses within the computer science curriculums, such as their cumulative grades or grades of each subject in those curriculums, and students' backgrounds may be impact factors on their future computer careers. Until now, there has been a lack of graduates with skills in digital and related fields to serve the labor market needs in Thailand [The Nation. (2018)].

Therefore, this research considers the importance of student qualifications in accordance with the careers and needs for computer personnel and the labor market expectations. In addition, the results of this research could be implemented with students in computer majors before they graduate with their bachelors' degrees from the universities. This research aims to study the impact factors on qualifications of employment positions for computer majors in the Faculty of Science and Technology, Suratthani Rajabhat University in Thailand, to develop a prototype of a prediction model based on employment positions for computer majors, and to compare model performances using different classification techniques. This research acquired datasets for four curriculums, including computer science, information technology, computer technology (which consisted of three branches: computer technology management, telecommunication, and multimedia), and multimedia technology and animation, by collecting data for five years for the graduates of sessions 2016-2020.

2. Related Work

A variety of techniques has been studied and reported in literature reviews for prediction models in the healthcare [VinothinL and Baghavathi (2020)], [Sridevi and Arun Kumar (2021)], and academics [Vivek Raj and Manivannan (2021)], [Athani *et al.* (2017)]. In general, the criteria of higher education have been measured using a graduate production process and post-graduate success. Therefore, it is very important in career counseling for future graduates to receive opportunities for employment positions directly within their fields. However, employment positions seem insufficient because graduates work in careers in different fields than the fields of their degrees. In addition, a good employment match would reduce their risks of changing new positions. Some studies have attempted to investigate the use of students' academic performance and social behavior with several attributes using data mining techniques [Athani *et al.* (2017)], [Na, W. (2020)], analysis of educational big data using machine learning for guiding the students in high school [Ababneh *et al.* (2021)] and optimize of agent-user matching process using a machine learning algorithm [Avdagić-Golub *et al.* (2020)]. The latest research was conducted with 4,634 students from 16 colleges and collected data from their use of campus smart cards for almost three years (from 2010/09/01 to 2014/06/30). It used four types of behavior: mastery of professional skills, behavioral regularity, reading interest, and family economic status to predict the career paths of the students using the approach cluster centers based on the XGBoost (ACCBOX) model [Nie *et al.* (2020)]. The results demonstrated that the prediction performance of the ACCBOX model that incorporated prototypical cluster center generation and a novel regularization item had the highest accuracy when compared with other methods.

Another research model that focused on the career prediction of computer science students using advanced machine learning techniques was designed in a different way with various factors including abilities of students in sports and academics and their hobbies, interests, competitions, skills, and knowledge. These factors and the total number of parameters considered amounted to 36 inputs, and 15 final job roles were predicted. The research results of three algorithms were compared: support vector machine, random forest decision tree, XGBoost. Support vector machine had the highest accuracy (90.3%) followed by XGBoost (88.33%) [Sripath *et al.* (2018)].

Data mining techniques consist of three components: classification or clustering, association rules, and sequence analysis. A linear classification was used to analyze a career prediction model with data collected from students' Facebook accounts. The data were used to evaluate their personalities. They also answered 20 questions to measure their aptitude and completed forms with their background information. The results revealed that these factors should be considered, even though they would predict different careers. Additionally, the model performance of the created intuitive career system was the highest with an accuracy of 77.41%, 75.4%, and 60.09% for aptitude, personality, and background information, respectively [Rangnekar *et al*. (2018)].

Meanwhile, a study on predicting the employability of graduates used three classification algorithms (decision tree, support vector machine, and artificial neural network) and identified seven factors that affected graduate employability: age, faculty, field of study, co-curriculum, material status, industrial internship, and English skill [Othman *et al.* (2018)]. The collected data from Malaysia's tracer study conducted in 2011-2017 contained 43,863 data instances. The results of the study showed that decision tree with J48 performed better than other techniques with an accuracy of 66.0651% that increased to 66.1824% after parameter tuning.

Moreover, a study on career prediction using data mining applied five classification techniques: ID3, CART, random forest, support vector machine, and neural networks and compared the model performances using measures of the confusion matrix: accuracy, precision, recall, and F-measure [Arafath *et al.* (2018)]. The research

collected data through an online survey of computer science students from 13 different universities in Bangladesh. The survey focused on the professional skills, interpersonal skills, and academic records of the students. Their evaluation technique was ten-fold cross- validation. Their results indicated that CART and neural networks (multi-layer perceptron) achieved the highest prediction accuracy of 95.24%.

In another classification technique, the C5.0 algorithm was used to develop a web application of career counseling to help students in high school to select a course in their career path [Gorad *et al.* (2017)]. This research provided datasets from questionnaires based on personality traits with the Myers-Briggs theory, interest, and capacity. Subsequently, the accuracy of the prediction model was improved using C5.0 with adaptive boosting. The results showed that the three factors impacted career selection. Then, the system of career counseling recommended particular courses along with the list of colleges providing those courses.

Additionally, a study of predicting information technology (IT) employability using data mining techniques used five algorithms: naïve Bayes, J48, SimpleCart, logistic regression, and chiad [Piad and Ballera (2014)]. The data were collected from the placement office based on the five-year profiles of the 515 students. Three academic variables had influence on predicting employability: IT_ Core, IT Professional, and Gender. The results showed that the logic regression algorithm had the highest accuracy (78.4%). This result is consistent with the study on predicting student success using data generated in traditional educational environments with Logistic Regression Classifier offering good values in terms of accuracy, true negative rate, and true positive rate [Bucos and Drăgulescu (2018)].

In addition, the association rule is one of the most common data mining techniques that have been applied in the research of guidelines for academic support of student career paths using data mining algorithms [Sodanil *et al.* (2019)]. The collected data consisted of the students' grades from 25 main courses of 215 student records in the field of IT at the Suan Sunandha Rajabhat University in 2011–2019, and five jobs including graphics and multimedia, IT support and service, network and security, programming, and others were considered. These were applied to the 14 association rules using WEKA software. The results supported the students in finding a place of internship optimal for their abilities.

Hence, this research emphasizes the study of the important impact factors in predicting the employment positions of computer careers recognized by labor market demands and matches the qualifications of graduating students in computer majors with those careers. In many previous studies, detailed datasets of each course and cumulative grade point average (CGPA) of all computer major students were not analyzed. In addition, model performances of six different classification techniques will be compared: decision tree, rule-based, k-nearest neighbor, support vector machine, artificial neural network, and naïve Bayes. Furthermore, the optimized prediction model based on employment positions for computer majors in the Faculty of Science and Technology, Suratthani Rajabhat University, will be developed into a web application to support graduating students of computer majors in selecting future career paths directly connected with their bachelors' degrees.

3. Proposed Methodology

This research was divided into two stages: building the prototype of the prediction model and developing a web application based on the effective prediction model. The framework of the prototype of a prediction model based on employment positions (PMEP) for computer majors in the Faculty of Science and Technology, Suratthani Rajabhat University, consisted of three independent variables: the academic results of each course in four computer curriculums, CGPA of graduated students, and qualifications necessary for employment positions in computer careers. The dependent variable was the prediction model based on employment positions for computer majors in the Faculty of Science and Technology, Suratthani Rajabhat University, as shown in Fig. 1.

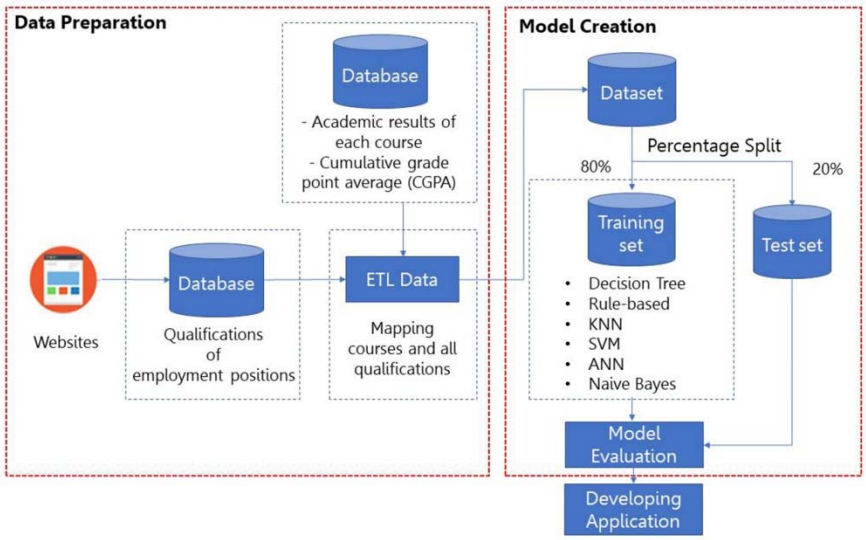

Fig 1. A framework of the proposed PMEP research

3.1. *Building the Prototype of PMEP*

We collected datasets from two parts: websites about computer careers and academic results from the Office of Academic Support and Registration (ASR) and the Office of Academic Resources and Information Technology (ARIT). Subsequently, we utilized the phases of the CRISP-DM methodology, named the data mining process, to build the prototype of a prediction model. The datasets were separated into two groups: a training set and a test set with split validation on 80% for training and 20% for test sets. Six classification techniques were used to compare the model performances. Finally, the effective prediction model was implemented as a web application. Building the prototype of a prediction model based on employment positions, we conducted six phases of the data mining process: business and data understanding, data preparation, model creation, model performance evaluation, and deployment.

To encourage business and data understanding, we focused on a prediction model based on employment positions for graduates in computer majors in the Faculty of Science and Technology, Suratthani Rajabhat University. In this research, we considered the academic results of graduates over a period of five years for the graduates of sessions 2016-2020. Based on our prediction model, career counseling can benefit graduating students in computer majors, and the risk of labor shortages, including mismatched position qualifications, can be reduced. For this purpose, we specified the top-ten ranked employment positions of current computer careers through well-known job. Data preparation is an important phase for determining research results in the data mining process, as the result quality depends on data quality. Data collection for this research included three sets of data. Two of the three sets were obtained from the offices of ASR and ARIT of Suratthani Rajabhat University. These datasets included the graduates' academic results of each course in the computer curriculum and cumulative grade point average of each graduate in five years (2012-2016; Table 1).

Item	Number of Students' Records					
	2012	2013	2014	2015	2016	Total
Academic Results	8,224	6,746	5,452	3,463	485	24,730
CGPA	169	147	166	82	35	549

Table 1. Total number of data items

Data integration is a technical combination of different resources. Data sets from these resources were rearranged by years 2012-2016 in Excel format. The remaining dataset included employment positions for computer majors that were extracted from job websites, a total of 290 positions, and mapped with target careers of each curriculum in the Thailand Qualifications Framework for higher education (TQF:HEd): programmer, mobile programmer, software engineer, IT support staff, web programmer, IT sale consultant, database administrator, technical expert, network administrator, user interface designer, etc. These datasets were prepared

with five features, including position data, employment fields, CGPA, and computer and language qualifications. Subsequently, data transformations were conducted on all existing datasets in a suitable pattern of csv format for building prediction models using RapidMiner [Kotu and Deshpande (2014)].

Six classification techniques were generated to obtain the prediction model: decision tree, rule-based, k-nearest neighbor, support vector machine, artificial neural network, and naïve Bayes. A data file from the data preparation phase was imported into RapidMiner. The percentage split technique was used to separate the data into two groups: training and test groups. This was performed using split validation, which is a technique to predict a suitable model in which its operator allows training on one dataset and testing on another explicit testing dataset. Six classification algorithms were used for predictive modeling to produce the effective model. These patterns were executed and run with 80% training and 20% test sets, as shown in Fig. 2.

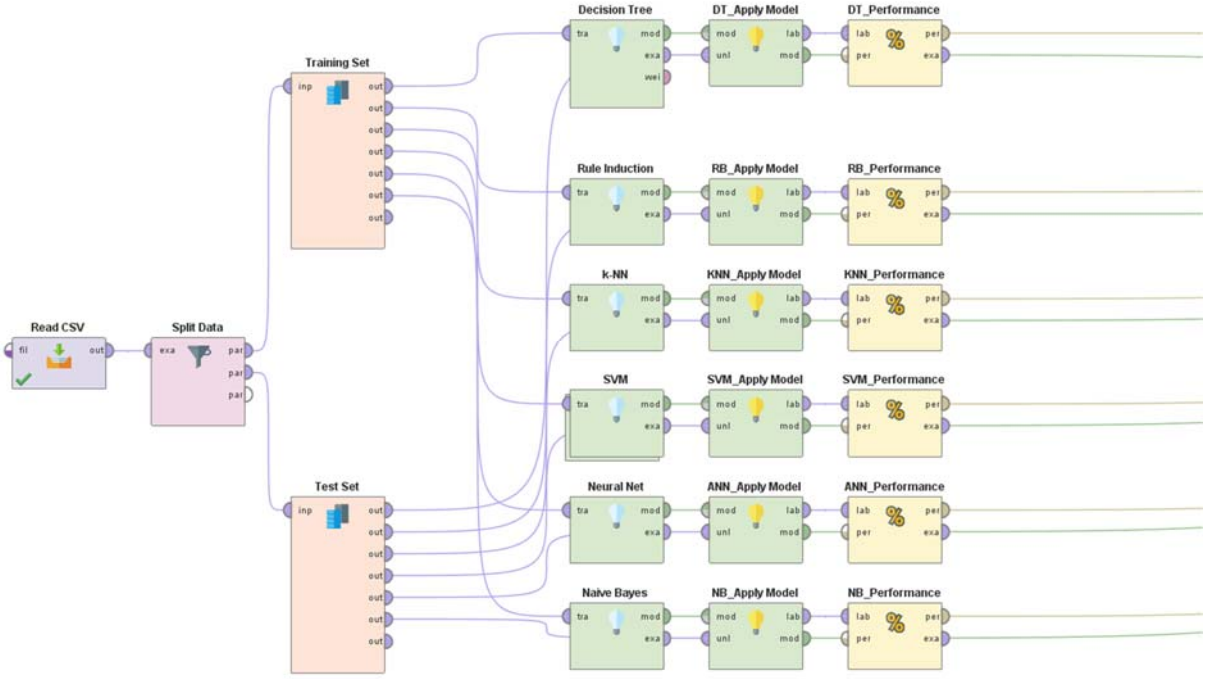

Fig. 2. Analyzing the prediction models using RapidMiner

In addition, model performances were evaluated using four parameters: accuracy, precision, recall, and F-Measure [Han et al. (2012)], [Witten et al. (2017)]. Then, a comparison of the model performances was conducted using the confusion matrix for each classifier. Accuracy is a measurement of the number of correct predictions made with the sum of both true positive and true negative, which were divided by the total number of predictions made as follows Eq. (1):

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

where TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative, respectively.

Precision is a measurement of the number of true positives divided by the number of true positives plus false positives, as follows Eq (2):

$$Precision = TP/(TP + FP) \quad (2)$$

Recall is a measurement of the number of true positives divided by the number of true positives plus the number of false negatives, as follows Eq (3):

$$Recall = TP/(TP + FN) \quad (3)$$

Further, the F-measure is a measurement of the balance between precision and recall, calculated as follows Eq (4):

$$F\text{-}measure = \frac{2*Precision*Recall}{(Precision+Recall)} \quad (4)$$

Therefore, for all values of accuracy, precision, recall, and F-measure obtained for the prediction models of employment positions, we selected the effective classifier using data mining techniques with the highest accuracy percentage of that classifier. Subsequently, the effective classifier was used to build a prediction model of employment positions for computer majors.

3.2. *Developing a Web Application of the Effective Prediction Model*

A web application of the optimized prediction model was developed using PHP programming language, which is known to be the fastest programming language. In addition, PHP frameworks are open source, platform independent, simple, and easy, connected with databases easily and securely, and provide great online support. Therefore, PMEP was implemented and run on a window allowing the entry of the input data, such as CGPA of graduates' courses for each curriculum. The output value indicated the employment position suitable for each student.

4. Experimental Results

The academic results of each course in the computer curriculum, cumulative grade point average, and target careers of each curriculum in TQF:HEd were used in the present study. The experimental results for both the developed PMEP and web applications are provided. The model performances of six classification techniques were compared. Then, the optimized model was implemented using a web application. As a result, there were two parts: PMEP model performance and a web application of PMEP.

4.1. *PMEP Model Performance*

The model performances of the six classification techniques indicated that the models with best performances were equally rule-based and naïve Bayes approaches with a classification accuracy of 89.66%, precision of 90.48%, and recall of 90.00% (Table 2). Although the accuracy of both the k-nearest neighbor (K=3) and artificial neural network methods were the same as those of the previous approaches, their precision and recall were lower. These results indicated the ability of a classification model to return only relevant instances (precision) and to identify all relevant instances (recall).

Model	Accuracy	Precision	Recall	F-Measure
Decision Tree	82.76%	60.99%	63.81%	0.62368
Rule-Based	89.66%	90.48%	90.00%	0.90239
KNN (K=3)	89.66%	76.19%	77.14%	0.76662
KNN (K=5)	82.76%	61.43%	61.43%	0.61430
KNN (K=8)	79.31%	44.76%	54.29%	0.49067
SVM	82.76%	56.75%	61.43%	0.58997
ANN	89.66%	64.29%	70.00%	0.67024
Naïve Bayes	89.66%	90.48%	90.00%	0.90239

Abbreviations: KNN: k-nearest neighbor; ANN: artificial neural network; SVM: support vector machine

Table 2. Model performances

Therefore, the optimized prediction model was used a rule-based approach, with academic result factors of eight courses: object-oriented programming, computer programming principles, software engineering, programming and algorithms, operating systems, data communication and networks, office automation, and web programming. In addition, the target careers of computer graduates were matched with seven careers: programmer (object-oriented programming), programmer, IT support, software engineer, IT sales consultant, user experience/user interface designer, and database administrator (Table 3). Moreover, students' grades were represented with numbers: A as 4, B+ as 3.5, B as 3, C+ as 2.5, C as 2, D+ as 1.5, D as 1, and E as 0 when they entered their academic results from each course into the web application.

Rules results from RapidMiner	Meaning of Rules
If OOP > 0.500 then programmer (OOP)	If students receive a higher grade than D in their object-oriented programming, they tend to become programmers (OOP).
If programming > 0.500 and SE ≤ 0.500 then programmer	If students receive a higher grade than D in their computer programming principles and receive a D or lower in software engineering, they tend to become programmers.
If algorithm ≤ 0.500 and OS > 0.500 then IT support	If students receive a D or lower in their programming and algorithms and receive a higher grade than D in operating systems, they tend to become IT support staff.
If network > 0.500 then software engineer	If students receive a higher grade than D in their data communication and networks, they tend to become software engineers.
If algorithm ≤ 0.500 and CGPA ≤ 1.350 and office > 0.500 then IT sales consultant	If students receive a D or lower in their programming and algorithms and their CGPA is lower than or equal to 1.350 and they receive higher grade than D in office automation, they tend to become IT sales consultants.
If algorithm ≤ 0.500 and CGPA ≤ 1.350 then UX/UI designer	If students receive a D or lower in their programming and algorithms and their CGPA is lower than or equal to 1.350, they tend to become UX/UI designers.
If CGPA > 1.350 then IT support	If students' CGPA is higher than 1.350, they tend to become IT support staff.
If web programming > 0.500 then programmer	If students receive a higher grade than D in their web programming, they tend to become programmers.

Rules results from RapidMiner	Meaning of Rules
Else database administrator	If students do not match with the previous conditions, they tend to become database administrators.

Table 3. Results of the rule-based model

4.2. *A Web Application of PMEP*

Using the rule-based approach and PHP programming, a web application of PMEP was developed and implemented to recommend employment options to graduates with a bachelor's degree from the four curriculums: computer science, information technology, computer technology (which consisted of computer technology management, telecommunication, and multimedia), and multimedia technology and animation (Fig. 3).

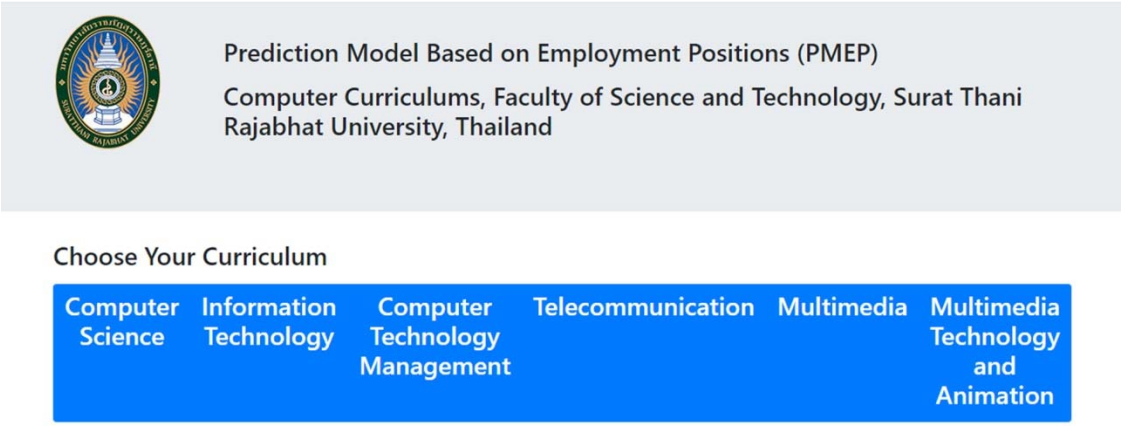

Fig. 3. A web application of PMEP

A web application of PMEP consists of six computer fields for students to choose their curriculum. Afterwards, they can enter their grades using numbers and their CGPA, as depicted in Figure 4. As a result, the system evaluates the conditions of rules when all input values are submitted by pressing the button, "Recommend your positions." For example, the output values in Fig. 4 indicate that suitable employment positions of the student are programmer and IT support.

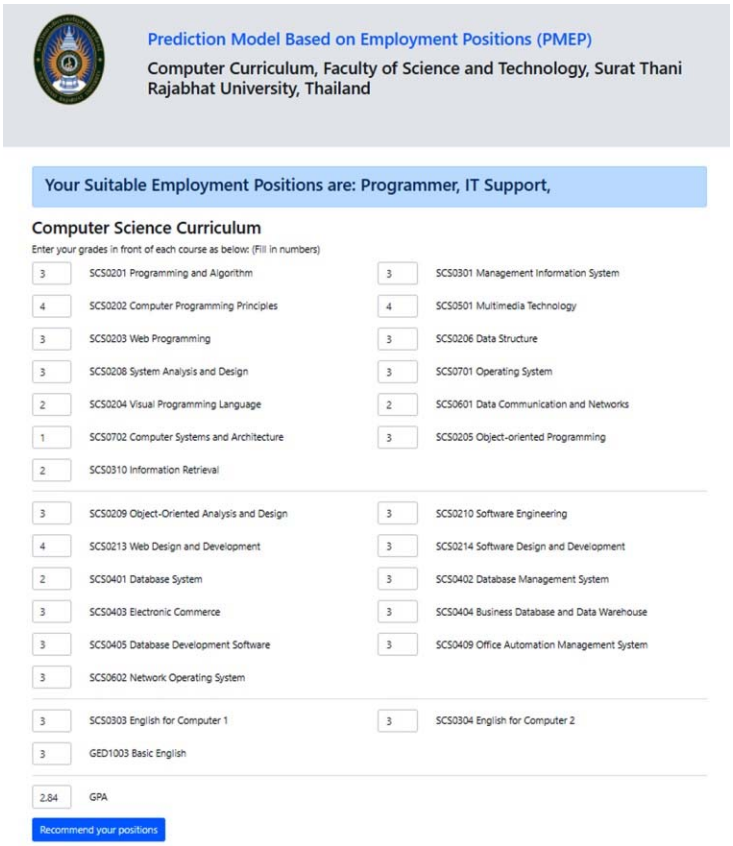

Fig. 4. Example of the entry input value and results of suitable employment positions for a computer science major

5. Discussion and Conclusion

Six classification techniques were used to investigate the impact factors for employment positions of students with computer-related majors. The academic results of each course and CGPA influenced the choice of a suitable employment position for graduates with computer-related bachelor's degrees. The similarity of this study was revealed in the prediction of employment positions from academic records or students' grades [Othman *et al.* (2018)],[Gorad *et al.* (2017)],[Piad *et al.* (2014)]. Eventually, an optimized PMEP using a rule-based approach was developed for a web application. The model evaluation demonstrated that the rule-based method performs better than other classification techniques and has the highest accuracy. However, this study was compared with the different classification techniques used in previous studies, which resulted in better model performance. Therefore, our PMEP application can recommend employment positions directly to graduates with bachelor's degrees in computer fields. Moreover, the application can provide an opportunity for choosing their employment positions in advance to reduce the risk of resigning from work later. In the future, we would like to use other variables such as professional and language skills to build a more accurate prediction model based on employment positions in the computer curriculum.

Acknowledgments

The Research and Development Institute, Surat Thani Rajabhat University, supported the research funding for this study. We would like to give special thanks to the Office of Academic Support and Registration (ASR) and the Office of Academic Resources and Information Technology (ARIT) of Surat Thani Rajabhat University for providing the data. We would also like to thank the Science and Technology faculty for contributing to and facilitating this study. We would like to acknowledge Dr. Aongart Aun-a-nan and Dr. Pornthip Wimonsong for their suggestions and assistance.

References

[1] Ababneh, M.; Aljarrah, A.; Karagozlu, D.; Ozdamli, F. (2021): Guiding the students in high school by using machine learning. TEM Journal, 10(1), pp. 384-391.
[2] Arafath, M.Y.; Saifuzzaman, M.; Ahmed, S.; Hossain, S.A. (2018): Predicting career using data mining. In 2018 International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, pp. 889-894.
[3] Athani, S. S.; Kodli, S. A.; Banavasi, M. N.; Hiremath, P.G.S. (2017): Student academic performance and social behavior predictor using data mining techniques. In 2017 International Conference on Computing, Communication and Automation (ICCCA), Grater Noida, India, pp. 170-174.

[4] Avdagić-Golub, E.; Begović, M.; Kosovac, A. (2020): Optimization of agent-user matching process using a machine learning algorithms. TEM Journal, 9(1), pp. 158-163.
[5] Bucos, M.; Drăgulescu, B. (2018): Predicting student success using data generated in traditional educational environments. TEM Journal, 7(3), pp. 617-625.
[6] Gorad, N.; Zalte, I.; Nandi, A.; Nayak, D. (2017): Career counseling using data mining. International Journal of Engineering Science and Computing, 7(4), pp. 10271-10274.
[7] Han J.; Kamber M.; Pie J. (2012). *Data mining: concepts and techniques (The Morgan Kaufmann Series in Data Management Systems)* 3rd edn. Morgan Kaufmann publishers, Waltham, MA.
[8] Kotu, V.; Deshpande, B. (2014). *Predictive analytics and data mining: concepts and practice with RapidMiner*, 1st edn. Morgan Kaufmann publishers, Waltham, MA.
[9] Na, Wei. (2020). A data mining method for students' behavior understanding. International Journal of Emerging Technologies in Learning, 15(6), pp. 18-31.
[10] Nie, M.; Xiong, Z.; Zhong, R.; Deng, W.; Yang, G. (2020): Career choice prediction based on campus big data-mining the potential behavior of college students. Applied Sciences (Switzerland), 10(2841), pp. 1-14.
[11] Nunman, T. (2017). *Working which are not related with Bachelor's Degree: Reflection on Inflation Society*. Retrieved from: https://www.posttoday.com/politic/report/418945/, [accessed: 1 March 2021].
[12] Othman, Z.; Shan, S. W.; Yusoff, I.; Kee, C.P. (2018): Classification techniques for predicting graduate employability. International Journal on Advanced Science, Engineering and Information Technology, 8(4), pp. 1712-1720.
[13] Piad, K. C.; Ballera, M. A. (2014): Predicting IT employability using data mining techniques. The Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications, January 2014, pp. 26-30.
[14] Rangnekar, R.H.; Suratwala, K.P.; Krishna, S.; Dhage. S. (2018): Career prediction model using data mining and linear classification. In 2018 Proceedings of The 4th International Conference on Computing, Communication Control and Automation (ICCUBEA), Pune, India, pp. 1-6.
[15] Sridevi, M.; Arun Kumar, B.R. (2021): A framework for performance evaluation of machine learning techniques to predict the decision to choose palliative care in advanced stages of alzheimer's disease. Indian Journal of Computer Science and Engineering, 12(1), pp. 35-46.
[16] Sripath R. K.; Roopkanth, K.; Uday T. V.; Bhavana, V.; Priyanka, J. (2018): Student career prediction using advanced machine learning techniques. International Journal of Engineering and Technology (UAE), 7(2), pp. 26-29.
[17] Sodanil, M.; Chotirat, S.; Poomhiran. L.; Viriyapant, K. (2019): Guideline for academic support of student career path using mining algorithm. The Third International Conference on Natural Language Processing and Information Retrieval, ACM International Conference Proceeding Series, Tokushima, Japan, pp. 133-137.
[18] The Nation. (2018). *Experts worried about lack of Thai digital skills*. Retrieved from: https://www.nationthailand.com/national/30343718, [accessed: 22 April 2021].
[19] VinothinL, A.; Baghavathi, P.S. (2020): Disease prediction model on imbalanced data using machine learning techniques. Indian Journal of Computer Science and Engineering, 11(6), pp. 708-718.
[20] Vivek Raj, S. N.; Manivannan, S.K. (2021): Modelling a machine learning classifier for predicting student's entrepreneurial intentions. Indian Journal of Computer Science and Engineering, 12(3), pp. 598-604.
[21] Witten, I. H.; Frank, E.; Hall, M.A.; Pal, C.J. (2017). *Data mining: practical machine learning tools and techniques (Morgan Kaufmann Series in Data Management Systems)*, 4th edn. Morgan Kaufmann publishers, Cambridge, MA.

Authors Profile

Jittima Silprachawong received her Doctor of Technology in Science, DTECH degree in Computer Science from University of Technology, Sydney in Faculty of Sciences, Australia in 2006. Since 1999 until now, she has been a lecturer and researcher in Department of Information Technology, Faculty of Science and Technology in Suratthani Rajabhat University, Thailand. She is still working her research about data mining, machine learning and management information systems. Her main research areas concern with building and implementing a prototype of prediction model using data mining techniques.

Kannika Kaewchuea was born in Suratthani, Thailand in 1977. She graduated in bachelor degree of Computer Science from Suratthani Rajabhat Institute, Thailand in 2000. In 2004, she received master degree in information technology from King Mongkut's Institute of Technology North Bangkok, Thailand. Currently, she is a lecturer and researcher in Department of Computer Education, Faculty of Education in Suratthani Rajabhat University, Thailand. Additionally, she interests in working her research about database management and utilizing data mining techniques.

Asok Srisawat received master degree in information technology from King Mongkut's Institute of Technology North Bangkok, Thailand in 2003. He has been teaching experiences for 24 years. Currently, he is a lecturer and researcher in Department of Computer Education, Faculty of Education in Suratthani Rajabhat University, Thailand. His main areas of interest are about database management and software development.