Efficiency Assessment of Undergraduate Students Based on Academic Record Using Deep Learning Methodology

Arthit Buranasing

***Abstract*—Computer science is the study of computers and computational systems which computer scientists deal mostly with application software and system software. Although knowing how to program is essential to the study of computer science, but it is only one element of the field. For example, software development uses various skills and techniques which are included in various subjects of a general computer science course. This paper focuses on senior students in computer science course who would like to assess the efficiency of their computer science skill in order to improve themselves. Moreover, the model also helps in the recruitment of new staff so that the companies would be able to assess the efficiency of newly graduated students or inexperienced candidates. This is because the lack of skill and inefficiency could cause problems to the hiring companies since they would have to invest time and money into training the new staff. This model can solve this problem by evaluating the performance and define the skills that must be improved directly. The result of the model is satisfactory, the average accuracy from experiment testing of confusion matrix is 89.33%.**

***Index Terms*—Artificial intelligence, deep learning, student performance prediction, recommender system.**

I. Introduction

Computer science is the study of computers and computational systems. Unlike electrical and computer engineers. Computer scientists deal mostly with application software and system software. The principal areas of study within computer science include artificial intelligence, computer systems and networks, security, database systems, human computer interaction, computer vision and computer graphics, numerical analysis, programming languages, software engineering, bioinformatics and theory of computing [1].

Although knowing how to program is essential to the study of computer science, but it is only one element of the field. For example, software development uses various skills and techniques which are included in various subjects of a general computer science course such as problem solving skill which is included in a computer algorithm subject, software design is included in a system analysis and design subject, machine learning is a part of artificial intelligence, etc. These skills require programming to accomplish the computer science task [2].

However, there is a sequence of process to perform a software development, various skills in research and other tasks of computer science. For example, for software development, the process of software development (known as the software development life cycle) [3] consists of 6 steps including planning, analysis, design, development and implementation, testing and deployment and maintenance which each step requires a particular skill in each subject.

In Addition, to understand what each student lacks is important for the student him/herself to improve a specific skill and important for the companies to assess the efficiency of the candidates. This is because the lack of skill and inefficiency could cause problems to the hiring companies since they would have to invest time and money into training the new staff. Therefore, it is important to assess the skills requried for computer science tasks and identify the lack of skill in each area in order to make improvements.

However, one methodology to evaluate the students' performance is Deep Learning methodology [4]. Deep Learning is an artificial intelligence function that imitates how a human brain works in processing data and creating patterns for decision making. Deep Learning is a subset of machine learning in artificial intelligence (AI) that has networks capable of learning, unsupervised, from data that is unstructured or unlabeled. In addition, learning can be supervised, semi-supervised or unsupervised. Also known as deep neural learning or deep neural network.

This paper focuses on senior students in a computer science course who would like to assess their efficiency by themselves or on companies that would like to assess the efficiency of the newly graduated students or inexperience candidates applying for jobs. This model benefits both the students and companies since they will able to improve only the specific skills that students or candidates lack such as programming, network administrator, data mining etc. This paper is organized into the following sections: Section I: Introduction, Section II: Overview of efficiency assessment system, Section III: Data collection and transformation, Section IV: Feature selection and data pre-processing for experiments, Section V: Deep Learning architecture analysis and design, Section VI: Performance of the model, Section VII: Student's performance graphic display and the last section is the conclusion and remarks.

II. Overview of Efficiency Assessment System

The overview of the efficiency assessment of undergraduate students which is shown in Fig. 1 consists of 4 parts, i.e. a) Part 1, Transcript of record data, in this part, the transcript will be transform into computer data b) Part 2, Feature selection for each computer science task c) Part 3,

Manuscript received December 21, 2019; revised March 7, 2020. This work was supported in part by Srinakharinwirot University, Thailand.

Arthit Buranasing is with Srinakharinwirot University, Thailand (e-mail: arthit.bur@hotmail.com).

doi: 10.18178/ijiet.2020.10.7.1416

Evaluation model and efficiency prediction d) Part 4, Student's performance chart.

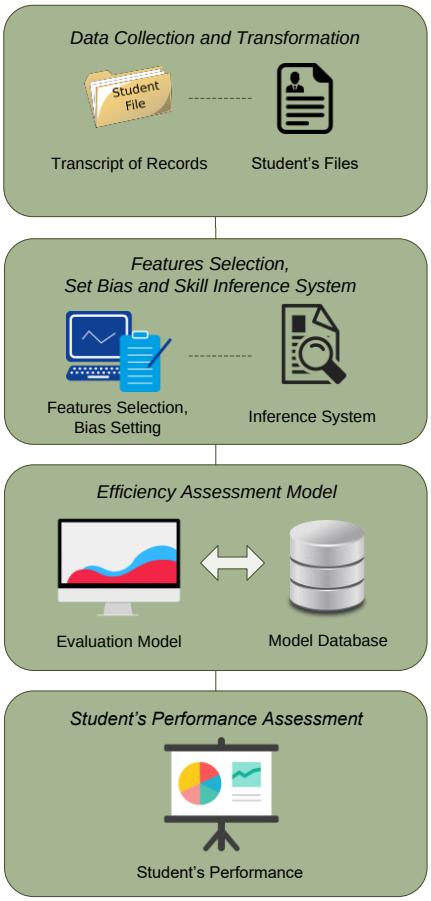

Fig. 1. Overview of efficiency assessment system.

III. Data Collection and Transformation

The experiment in this paper used 214 students' transcript records. The data is based on a group of undergraduate students in the same year (only senior students), course or department (computer science) and university. The data includes: name, sex, date of birth, subject and academic year, etc. However, the data in the transcript was only selected from computer science subject which is shown the example in Table I.

TABLE I: Example of Student Data for Experiment

Subject	Grade
Mathematics III	C
Statistic for Engineering and Science	B
System Analysis and Design	A
Computer Programming	C+
Database System	B
Computer Algorithm and Formal Method	A
Discrete Mathematics for Computer Science	B
Computer Graphics	C
Image Processing	B
Data Mining	C+
Machine Learning	B+

The features in Table I were transformed into numerical data by standard of academic record [5]. The transformed's data is shown the example in Table II. Note that, this data will be used in feature selection and the next section.

TABLE II: Example of Student Data Transforms for Experiment

Subject	Grade
Mathematics III	60
Statistic for Engineering and Science	70
System Analysis and Design	80
Computer Programming	75
Database System	70
Computer Algorithm and Formal Method	80
Discrete Mathematics for Computer Science	70
Computer Graphics	60
Image Processing	70
Data Mining	65
Machine Learning	75

IV. Feature Selection and Data Pre-processing for Experiments

In this paper, the efficiency of the students was evaluated from the skill in computer science tasks in 3 classes of performance i.e. mobile application development, computer vision and machine learning. Each group includes individual features of the students' skills which are shown the example in Table III-V.

TABLE III: List of Features for Mobile Application Development

ID	Subject (Bias)
S1	**Computer programming** is the process of designing and building an executable computer program for accomplishing a specific computing task such as JAVA or C#. (1.0)
S2	**System Analysis and design** deal with the planning of the development of information systems through understanding and specifying in detail what a system should do and how the components of the system should be implemented and work together (1.0)
S3	**Data structure** is a data organization, management and storage format that enables efficient access and modification. (0.8)
S4	**Compuer algorithm** is a set of instructions, typically to solve a class of problems or perform computation. Algorithms are unambiguous specifications for performing calculation, data processing, automated reasoning and other tasks. (1.0)
S5	**Data Communication and Network** is a digital telecommunication network which allows nodes to share resources (1.0)
S6	**Software testing** is a process to evaluate the functionality of a software application with an intent to find whether the developed software met the specified requirements or not (0.8)
S7	**Project management** is the practice of initiating, planning, executing, controlling and closing the work of a team to achieve specific goals and meet specific success criteria at the specified time (0.5)

TABLE IV: List of Features for Computer Vision

ID	Subject (Bias)
S1	**Computer programming** is the process of designing and building an executable computer program for accomplishing a specific computing task such as MATLAB or C++. (1.0)
S3	**Data structure** is a data organization, management, and storage format that enables efficient access and modification. (0.8)
S4	**Compuer algorithm** is a set of instructions, typically to solve a class of problems or perform a computation. Algorithms are unambiguous specifications for performing calculation, data processing, automated reasoning, and other tasks. (1.0)
S8	**Digital image processing** is the use of computer algorithms to perform image processing on digital images Image technique for enhancement a image (1.0)
S9	**Machine Learning (ML)** is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead (1.0)
S10	**Computer graphics** are pictures and films created using computers (1.0)

TABLE V: LIST OF FEATURES FOR MACHINE LEARNING

ID	Subject (Bias)
S1	**Computer programming** is the process of designing and building an executable computer program for accomplishing a specific computing task such as MATLAB or Python. (1.0)
S3	**Data Structures** used for design structures inside computer code (0.8)
S4	**Compuer algorithm** is a set of instructions, typically to solve a class of problems or perform a computation. Algorithms are unambiguous specifications for performing calculation, data processing, automated reasoning, and other tasks. (1.0)
S9	**Machine Learning (ML)** is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead (1.0)
S11	**Data mining** is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems (1.0)

The example of student dataset in each class is shown in Table VI-VIII. (214 records in each class). However, since each feature may not be equally important so bias will be set by computer science expertise. All data in the dataset will be calculated with bias by using Fuzzy Inference System (FIS), (Fig. 2) with Mamdani techniques [6] as follows: the system consists of 2 inputs and 1 output. The first input is a standard score (Fig. 3) and second input is a bias with S-shaped membership function (Fig. 4) and the output is adjusted range of score (Fig. 5). As a consequence, the output of Fuzzy Inference System for skill evaluation is shown in Fig. 6. Note that, all the rules of fuzzy is (if X and Bias) then X-Adjust, where X is grade of student i.e., A, B+, B, C+, C, D+, D, and F.

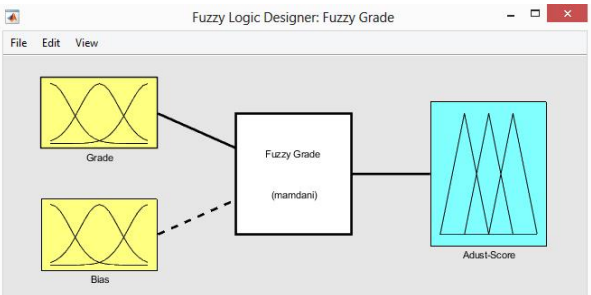

Fig. 2. Fuzzy inference system structure for student's performance assessment.

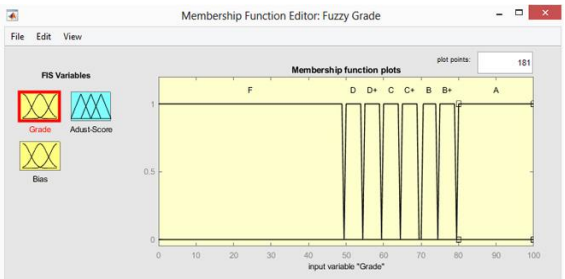

Fig. 3. Input 1 (score) of fuzzy inference system.

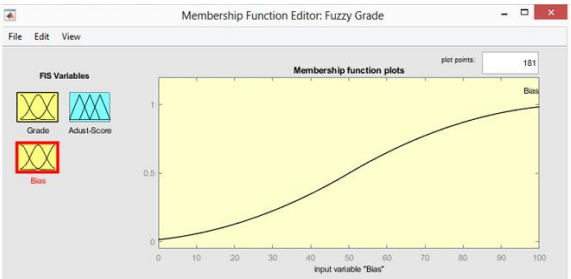

Fig. 4. Input 2 (bias) of fuzzy inference system.

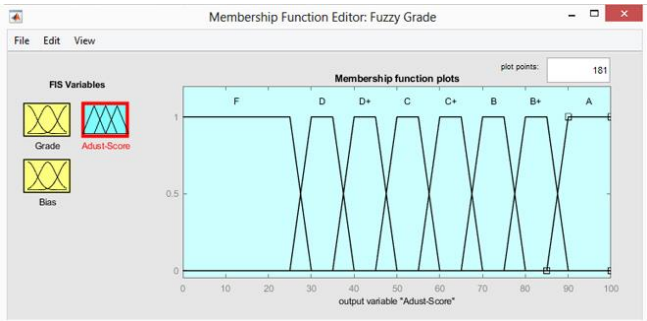

Fig. 5. Output of fuzzy inference system.

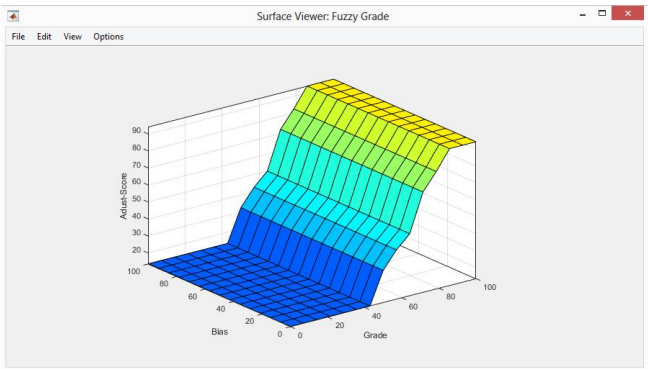

Fig. 6. Fuzzy infrence system for skill evaulation.

TABLE VI: EXAMPLE OF CLASS 1 FEATURES FOR EXPERIMENT

STD ID \ Subject	S1	S2	S3	S4	S5	S6	S7
1	89	84	90	85	82	80	87
2	67	74	75	75	72	63	67
3	60	66	71	61	69	63	75
4	74	73	60	78	69	76	76
5	71	71	60	80	67	76	64

TABLE VII: EXAMPLE OF CLASS 2 FEATURES FOR EXPERIMENT

STD ID \ Subject	S1	S2	S3	S4	S5	S6
1	89	78	83	92	80	78
2	60	76	80	61	69	60
3	70	74	76	75	73	80
4	65	69	61	79	80	68
5	60	78	66	78	77	74

TABLE VIII: EXAMPLE OF CLASS 3 FEATURES FOR EXPERIMENT

STD ID \ Subject	S1	S2	S3	S4	S5
1	77	73	61	78	73
2	66	76	74	70	80
3	75	62	78	70	80
4	70	67	70	62	74
5	75	68	75	77	78

V. DEEP LEARNING ARCHITECTURE ANALYSIS AND DESIGN

The model architectures of Deep Learning in each class are designed as follows: Class 1, the model architecture consists of 7 input nodes, 2 hidden layers and 5 output nodes. Class 2, the model architecture consists of 6 input nodes, 2 hidden layers and 5 output nodes. Class 3, the model architecture consists of 5 input nodes, 2 hidden layers and 5 output nodes. All model architecture are shown in Fig. 7, 8 and 9

The input of each class in the model Classes 1 to 3 is a feature selection data in each class which is descripted in

Section IV. The output class consists of 5 output nodes with A-F labels which is descripted in the Table IX.

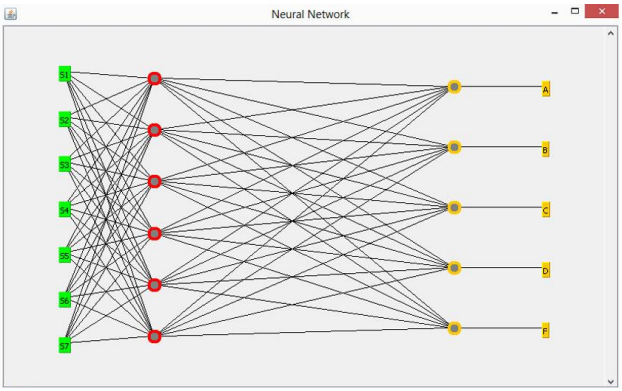

Fig. 7. Neural network architecture model for class 1.

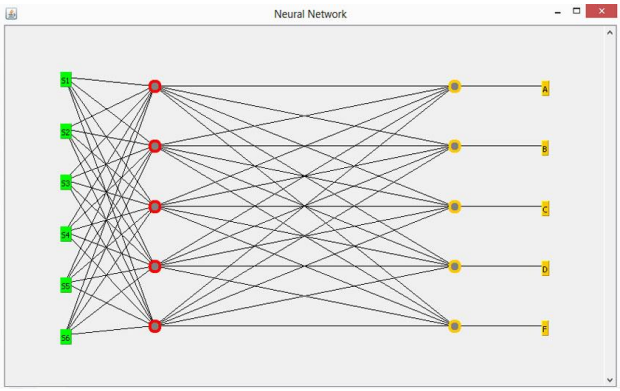

Fig. 8. Neural network architecture model for class 2.

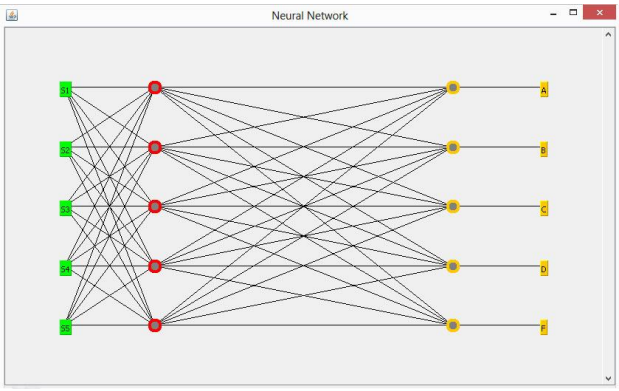

Fig. 9. Neural network architecture model for class 3.

TABLE IX: OUTPUT SYSMBOL DESCRIPTION

Description / Output	Description of Symbol
A	Good
B	Above Average
C	Average
D	Below Average
F	Bad

VI. PERFORMANCE OF THE MODEL

The performance of the model was evaluated in Table X-XII to 3. The experiment of the efficiency assessment model was divided into 2 parts: one is the training part in which the result is absent in this paper due to the objective of the training part was only created as a mathematical based equation and the other is the testing part. In the training class, the model used the transcript of records data 150 records (about 70%), and the testing class used 64 records (about 30%)

TABLE X: THE PERFORMANCE OF EFFICIENCY ASSESSMENT MODEL IN CONFUSION MATRIX (CLASS 1)

Actual Class / Predicted Class	A	B	C	D	F
A	7	1	0	0	0
B	0	18	0	0	0
C	0	0	13	3	0
D	0	0	0	12	3
F	0	0	0	1	6

TABLE XI: THE PERFORMANCE OF EFFICIENCY ASSESSMENT MODEL IN CONFUSION MATRIX (CLASS 2)

Actual Class / Predicted Class	A	B	C	D	F
A	16	2	0	0	0
B	1	7	0	0	0
C	0	0	16	2	0
D	0	0	1	12	0
F	0	0	0	2	5

TABLE XII: THE PERFORMANCE OF EFFICIENCY ASSESSMENT MODEL IN CONFUSION MATRIX (CLASS 3)

Actual Class / Predicted Class	A	B	C	D	F
A	11	0	1	0	0
B	1	17	0	0	0
C	0	0	11	1	0
D	0	0	1	15	0
F	0	0	0	0	6

This paper proposed the efficiency assemssment model by using the artificial neural network. The performance of the model gives an average F-measure of Classes 1,2 and 3 of around 0.87 or 87%, 0.88 or 88% and 0.93 or 90% respectively.

VII. STUDENT'S PERFORMANCE GRAPHIC DISPLAY

To display the student's performance graphically, it can be shown in a pie chart in Fig. 10 which the data is derived from Table XIII.

TABLE XIII: EXAMPLE OF STUDENT'S PERFORMANCE ASSESSMENT IN CLASS 1

Subject / STD ID	S1	S2	S3	S4	S5
1	77	73	61	78	73

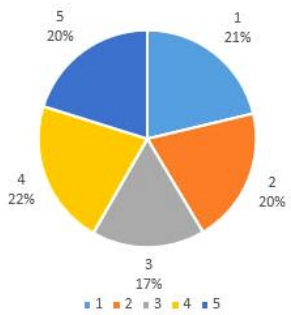

Fig. 10. Student's performance assessment.

In Fig. 10, it is an example of the students' performance assessment that can be described as follows: the students have the ability of S1 = 21%, S2 = 20% S3 = 17%, S4 = 22%,

S5 =20% (derived from: Fuzzy Inference System) and the overall result is above average (derived from Deep Learning model). In addition, the pie chart describes which skill of the students lacked.

VIII. Conclusion and Remark

This paper focuses on efficiency assessment of senior undergraduate students in the same course or department (computer science) and university. This model benefits both the students and companies since they will be albe to improve only the specific skills that the students or candidates lack. This is because the lack of skill and inefficiency could cause problems to the hiring companies since they will have to invest time and money into training the new staff. This model solves the problem by evaluating the performance and define the skills that staff need to be improved directly. The result of the model is satisfactory, the accuracy from experiment testing of confusion matrix in each class are 87%, 88% and 93% respectively.

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

Author Contributions

This paper had been written by a single author who had done all the research, writing and review stages. The author approved the final version of this paper.

References

[1] J. Voas, R. Kuhn *et al*., "A clustering technique for digital communications channel equalization using radial basis function networks," *IEEE Computer Society*, February 2018.

[2] F. Y.Assiri, "Recommendations to improve programming skills of students of computer science," in *Proc. 2016 SAI Computing Conference (SAI)*, July. 2016.

[3] IEEE Standard Board, "IEEE standard for developing software life cycle processes," *IEEE Computer Society*, Jan. 1992.

[4] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, Dec. 2016, pp. 417-466.

[5] Grading scale. (2019). [Online]. Available: https://studyinthailand.org/study-abroad-thailand-university/gpa-grading-examination.html

[6] N. Wang, C. Hu, and W. Shi, "A mamdani fuzzy modeling method via evolution-objective cluster analysis," in *Proc. the 31st Chinese Control Conference*, Jul. 2012.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).

Arthit Buranasing was born in 1984. He received B.Sc., M.Sc. and Ph.D. degree in computer science from King Mongkut's University of Technology North Bangkok, Thailand. He is a lecturer in Department of Educational Technology, Faculty of Education, Srinakharinwirot University, Thailand. His research interests include artificial intelligence (AI), computer algorithms, earth science, geoscience, weather forecasting, geographic information system (GIS), remote sensing (RS), intelligent transportation system (ITS), educational technology.