ISSN (Print): 0974-6846
Indian Journal of Science and Technology, *Vol 11(37), DOI: 10.17485/ijst/2018/v11i37/130842, October 2018* ISSN (Online): 0974-5645

Predicting Employability Skills among Information Technology Graduates of Philippine State University in their On-the-Job Training using J48 Algorithm

Rommel L. Verecio*

Information Technology & Computer Education Leyte Normal University Philippines; rlverecio@lnu.edu.ph

Abstract

Objectives: This study predicts the employability skills acquired by the BS Information Technology graduates of Leyte Normal University, Tacloban City. Work skills or employability skills were identified based on the Commission on Higher Education (CHED) Memorandum Order 53, series of 2006, which categorizes as personal, interpersonal, and technical understanding skills. **Methods:** Using data mining to predict employability skills of the graduates particularly on the J48 algorithm a C4.5 decision tree model. The 10-Folds Cross-validation and Receiving Operating Characteristics Curve (ROC) was deployed to create a model and test the result based on the attributes. The collected datasets of this study are from the graduates from S.Y. 2015-2016 to 2017-2018. There are 138 datasets with six variables (Systems Development, Research, Business Operations, Technical Support, Interpersonal Skills, and Remarks). **Findings:** Decision tree model and decision rule for classification were created. There were 85.50% correctly classified with an AUC weighted mean of 85.10%. Hence, classified students who possess Expert (82.8%), Advanced (95%), and Intermediate (57.10%) OJT performance. The model has high acceptability to predict that Business Operations is the most important attribute in the On-the-Job Training. Wherein, students are trained to be productive, responsible and cooperative and have the initiative to act. **Application/Improvements:** The result of this study can be a basis for policy measures for effective OJT program. Further, for more improvement of this paper and model, additional parameters should be considered to have more factors involved in predicting employability skills.

Keywords: Decision Tree, Employability Skills, Machine Learning, On-the-Job Training, J48 Algorithm, BSIT, LNU

1. Introduction

Human capital investment in an organization is necessary for identifying competent workers that show remarkable skills in business operations. The on-the-job training entails costs that include the time committed by the worker and co-workers to learning skills that increase productivity and equipment needed to teach this skills[1]. The increased productivity of the worker measures the returns of investment during subsequent periods of use. Moreover, skilled workers have higher chances of employability because they have specific expertise and effectively accomplish a task than those of non-skilled workers. Academe has a significant part in developing highly skilled workers. Designing a competent curriculum for the On-the-Job training is essential that will target the needed skills when deploying for OJT in an IT industry should be identified. On-the-Job training exposes the student to the real world business operations that they will function as an employee in the establishment. Hence,

*Author for correspondence

students are trained to become responsible, efficient and be productive in their workplace. Moreover, on-the-job training program exposed the students to the different IT situation. Wherein, students had a chance to apply their skills, knowledge, and attitude in the workplace[2].

Commission on Higher Education formulates Memorandum Order No.23, s.2009 that set Guidelines for Student Internship Program in the Philippines (SIPP) for all Programs with Practicum Subject. As stated in the CMO, it indicates guidelines of that student who qualifies for the On-the-Job training as well and the responsibilities of the involved parties like the school, student, and the partner agency. The Leyte Normal University, Tacloban City a public university and categorized as one of the many state universities in the Philippines offered BS Information Technology (BSIT) program as one of its many undergraduate curricular offerings. The curriculum is anchored on the Rules and Standards for the Undergraduate Information Technology Education (CHED Memo Order No.53, s.2006). The fourth-year student of Bachelor of Science in Information Technology has a nine (9) unit, Internship course with a sum of four hundred eighty-six (486) hours to render for the OJT program. Hence, the students select their company from our partner agency. The seventy percent (70%) of the student grade comes from the evaluation of the company regarding student performance that they manifest regarding Business Operations, Interpersonal Skills, Technical Support, System Development and Research. The remaining thirty percent (30) of their OJT grades comes from the OJT coordinator based on the submitted requirements.

The application of data mining techniques supports educational institutions in many activities. In the education domain, researchers used data mining techniques for extracting rules, information, and predicting specific behaviors in several areas from substantial data sets. Employability issues have also been taken into consideration in other countries. Research by the Higher Education Academy with the Council for Industry and Higher Education inthe United Kingdom concluded that there are six competencies that employers observean individual who can transform the organizations and add values in their careers[3]. The six competencies are cognitive skills or brainpower, generic competencies, personal capabilities, technical ability, business or organization awareness and practical elements. Furthermore, it covers a set of achievements comprises skills, understandings and personal attributes that make graduates more likely to gain employment and successful in their chosen occupations which benefits the graduates, the community and also the economy.

There is a lot of research that was done in the academic performance prediction that may lead to employability; the employability prediction is still in a nascent state. Even the term "Employability" still has no precise definition. As described in many ways like the ability to secure a job, getting a job within a specified period after graduating, it may be the ability to skill map oneself according to the job need, or the willingness of the student to extend the graduate learning at work[4]. In this paper, we have taken employability as securing a job while on campus, i.e., while students are in the fifth semester and get placement offers from companies. In a study conducted to increase understanding of factors that influence the employability of university graduates both qualitative and quantitative methods were adopted. It was concluded that employers find soft skill more significant in employment as compared to academic excellence[5]. This exploratory study further explores the attributes like emotional intelligence, life experience work-life balance has been linked to employability. However, the findings are limited to theoretical aspects.

A recent study conducted includes almost the same methodology as above in which the data was sourced from the Tracer Study database[6]. Application of data mining algorithms indicates the superiority of the Decision Tree classification model over Bayes Network Classification Models. Another study conducted have taken attributes from the curriculum vitae, application, and interview of the candidate and applied data mining techniques to predict the performance of a new applicant[7]. The model helps the management in deciding the hiring of the employee.

Hence, the ability to predict OJT employability skills is essential for the teachers as well as to students, and partner companies in strengthening the teaching and training of the OJT students. The data mining will be usedto attain this objective. Data mining techniques are used to operate on massive amounts of data to discover hidden patterns and relationships helpful in decision making. The primary elements of data mining are applying different techniques and calculations to extract and identifytrends from massive amounts of data[8]. Further, this study delved to develop a model using a decision tree for predicting the OJT employability skills.

2. Methodology

The study utilized Knowledge Discovery in Database (KDD) process. The process starts with data collection and data pre-processing followed by classification model construction and ends with model evaluation and interpretations.

2.1 Data Collection

There are 138 datasets with six variables (Systems Development, Research, Business Operations, Technical Support, Interpersonal Skills, and Remarks). Data collection was obtained from the previous grades of the 4th year BSIT students enrolled in Internship/OJT Practicum from S.Y. 2015-2016 to 2017-2018. Table 1 shows the attribute description and their possible values.

Systems Development-the student or on-the-job trainee may be part of the project team from systems planning, analysis, design, implementation (includes programming, testing, debugging), deployment (contains final documentation, installation, user training) and support. The system to be developed is usually based on a business need articulated by the host establishment which can be addressed by automation through customized software and related hardware.

Research - is based on particular knowledge gap or technology gap identified in a certain knowledge domain or emerging technology. The objective is to propose a novel or unique solution to the identified problem and develop a working prototype application of the concept. Companies or institutions that are, in one way or another affected by the identified problem are primary candidate recipients of the planned research and development projects while the students have their on-the-job training.

Table 1. Students Related Variable

Variables	Description	Possible Values
S	Systems Development	1 – 5
R	Research	1 – 5
B	Business Operations	1 – 5
T	Technical Support	1 – 5
I	Interpersonal Skills	1 – 5
R1	Remarks	Intermediate, Advanced, Expert

Business Operations - referred to the following business processes, namely: sales and marketing management, financial management, human resource management, logistics management, customer relationship management and the like. The student is exposed to the business operations of the host establishment, and they document their experience. The whole practicum period of the on-the-job trainee exposed them through onsite immersion of the host establishment.

Technical Support - done in the form of database management, network management or data center management through design, administration or maintenance functions. Other forms of technical support for a firm's system infrastructure can be explored as deemed appropriate.

Interpersonal Skills-includes the following domains (teamwork and collaborative skills, oral and written communication skills and conflict resolution skills), that are demonstrated by the student during their on-the-job training.

Remarks-is the overall assessment of the on-the-job trainee performance as indicated on the different domains such as (Systems Development, Business Operations, Interpersonal Skills, Technical Support and Research) which categorized as an Intermediate, Advanced and Expert trainee?

2.2 Data Mining Process

Figure 1 illustrates the steps in the KDD process. First step data collection, the data used in this study is the students' grade in On-the-Job course for the last three (3) semesters of School Year 2015-2016 to 2017-2018. Student performance in Business Operation, Interpersonal Skills, Research, Systems Development and Technical Support is used in this study. Second step data preparation, student grades are stored in MS Excel and converted into an Attribute-Relation File Format (ARFF) which is a text file that describes a list of instances sharing a set of attributes and the accepted file format for *WEKA* application.

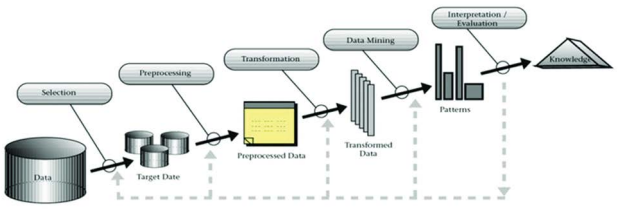

Figure 1. Using the J48 Algorithm.

Also, in this stage data cleaning is done to eliminate unwanted symbols (e.i., comma, colon, spaces). The declaration of syntax like @Relation, @Attribute, and the @ Datais for Weka application requirement. Next, information is uploaded to the Weka Application and conducted the pre-processing of raw data to a more understandable file format. The third step, data modeling, WEKA used to predict the Employability Skills of the BSIT students. This stage consists of five phases of training, pattern, testing, result evaluation and knowledge representation. In this stage, it divides the cleaned data into two stages; the *training and testing stage*. In the training stage, the J48 algorithm is used to build a model. The *J48 algorithm is C4.5 decision* tree approach that is useful in the classification of the problem which creates a binary tree model[9]. Moreover, the testing stage, the k-fold cross validation using *10-fold cross-validation*. The idea is based on a technique by dividing your data into *k* number of equally-sized folds. Also in the testing stage *Receiving Operating Characteristics Curve (ROC) Area Under ROC Curve* is a graph that is used to visualize the selected classifier based on the Expert, Advance and Intermediate remarks[10].

3. Result and Discussion

3.1 The Model

Figure 2 illustrates the graphical presentation of the pruned decision tree on OJT Employability Skills. Wherein, Business Operation as the highest instances and become the first split between the (Interpersonal Skills <= 4) and (Interpersonal Skills > 4) in predicting student OJT employability skills. Moreover, in Figure 3, shows the OJT employability skills decision rule that the Business Operation c has the highest factor that the students will be considered as Expert, Advanced and Intermediate in the OJT training.

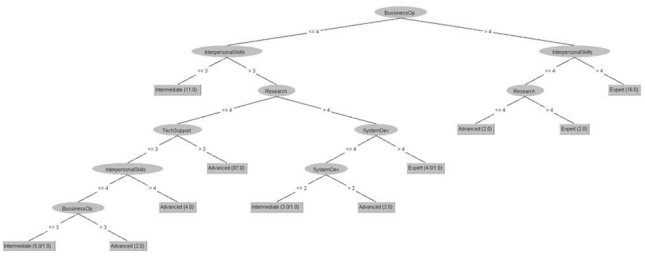

Figure 2. Decision Tree Model on On-the-Job Trainee Employability Skills.

The decision tree has correctly classified 138 instances as shown in Table 2 the confusion matrix that correctly classified instances and misclassification of OJT Employability Skills. Moreover, is interpreted as:

- The decision tree has correctly classified Sixteen (16) instances as EXPERT, and incorrectly Classified Four (4) as ADVANCED and One (1) as INTERMEDIATE leading to Misclassification.
- The decision tree has correctly classified Ninety-two (92) instances as ADVANCED, and incorrectly Classified Five (5) as EXPERT and Six (6) as INTERMEDIATE leading to Misclassification.
- The decision tree has correctly classified Ten (10) instances as INTERMEDIATE, and incorrectly Classified Four (4) as ADVANCED leading to Misclassification.

Table 3 shows the cross-validation summary wherein there are 85.50% correctly classified instances and 14.49% incorrectly classified instances this is supported by Table 4 that shows the precise accuracy by class wherein the *Precision Weight Average* of the OJT Employability Skills of the students is 85.10%. Additionally, the study utilized Receiving Operating Characteristics Curve (ROC) and the Area under ROC Curve (AUC) for model accuracy a shown in Figure 4 about ROC curve and AUC curve. Also, results revealed that the attribute EXPERT 97% accuracy, ADVANCED has 90% accuracy, and INTERMEDIATE

R1: IF (BusinessOp <=4) AND (InterpersonalSkills <=3) THEN
Skills = "INTERMEDIATE"
R2: IF (BusinessOp <=4) AND (InterpersonalSkills >3) AND (Research <=4) AND (TechSupport <=3) And (InterpersonalSkills <=4) AND (BusinessOp <=3)THEN
Skills = "INTERMEDIATE"
R3: IF (BusinessOp <=4) AND (InterpersonalSkills >3) AND (Research <=4) AND (TechSupport <=3) AND (InterpersonalSkills >4) THEN
Skills = "ADVANCED"
R4: IF (BusinessOp <=4) AND (InterpersonalSkills >3) AND (Research <=4) AND (TechSupport >3) THEN
Skills = "ADVANCED"
R5: IF (BusinessOp <=4) AND (InterpersonalSkills >3) AND (Research >4) AND (SystemDev <=4) AND (SystemDev <=2)
Skills = "INTERMEDIATE"
R6: IF (BusinessOp <=4) AND (InterpersonalSkills >3) AND (Research >4) AND (SystemDev <=4) AND (SystemDev >2)
Skills = "ADVANCED"
R7: IF (BusinessOp <=4) AND (InterpersonalSkills >3) AND (Research >4) AND (SystemDev >4)
Skills = "EXPERT"
R8: IF (BusinessOp >4) AND (InterpersonalSkills <=4) AND (Research <=4) THEN
Skills = "ADVANCED"
R9: IF (BusinessOp >4) AND (InterpersonalSkills >4) THEN
Skills = "EXPERT"

Figure 3. OJT Employability Skills Decision Rule.

Table 2. Confusion Matrix

a	b	C		<-- classified as
16	5	0	|	a = Expert
4	92	4	|	b = Advanced
1	6	10	|	c = Intermediate

has 86.10%. Finally, the model has high acceptability and accuracy in predicting the Student OJT Employability Skills.

3.2 Employability Skills of the On-the-Job Trainee

Employability skill of the on-the-job trainee depends on background and training during the internship program. As the result of data mining, Business Operation top as the highest indicator for the trainee to become employable. Under Business Operation Discipline students are being rated according to the Quality of Work, Initiative, Judgement, Adaptability, Team Work, Responsibility, Attendance, Public Relations, Management of Workload and Planning. In the study conducted, the result showed that communication skills (written and oral) and critical thinker is essential skills for IT graduates to be hired in the future[11]. A similar study conducted, the trainee should be able to develop soft skills like teamwork willingness to adopt firm values and culture, and lastly, the trainee should possess an accurate understanding of one's own identity and personal capabilities[12]. Additionally, information technology skill, human relation, and critical thinking skills are necessary that the trainee should develop[13].

Second Interpersonal Skills under this discipline trainee are integrated theories learned in school and interact with another co-worker. Those with good interpersonal skills functions easily in the team group, and other people. They can communicate effectively with other colleague, customers, and clients. Interpersonal skills are therefore vital in all areas of business, education, and society.

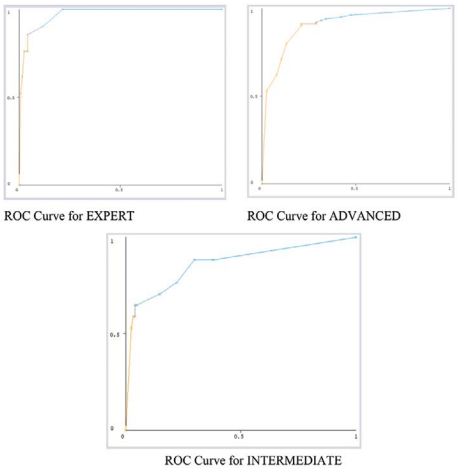

Figure 4. Area under the ROC curve.

Table 3. Cross-Validation Summary

Correctly Classified Instances	118		85.5072	%
Incorrectly Classified Instances	20		14.4928	%
Kappa statistic	0.6578			
Mean absolute error	0.1013			
Root mean squared error	0.288			
Relative absolute error	34.3766	%		
Root relative squared error	75.4581	%		
Total Number of Instances	**138**			

4. Conclusion

It is a reality that almost every year there is an increase in some graduates produced by higher education institutions. Graduates are facing more competition to ensure their employment in the job market. This study predicted the employability skills based on the performance of their on-the-job training which will help them to be more competitive in the job market. With the application of various classifiers to find the employability skills of the students and develop employability model based on the suitable classifier. Highly acceptable results shown in the decision tree, confusion matrix, Receiving Operating Characteristics Curve (ROC), and Area under ROC Curve. The rules derived from the model emphasize that Business Operations is the most important attribute in

Table 4. Detailed Accuracy by Class

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.762	0.043	0.762	0.762	0.762	0.719	0.970	0.828	Expert
	0.920	0.289	0.893	0.920	0.906	0.647	0.903	0.950	Advanced
	0.588	0.033	0.714	0.588	0.645	0.604	0.861	0.571	Intermediate
Weighted Avg.	**0.855**	**0.220**	**0.851**	**0.855**	**0.852**	**0.653**	**0.908**	**0.884**	

the On-the-Job Training. Wherein, students are trained to be productive, responsible and cooperative and have the initiative to act.

5. Recommendation

The result of this study can be a basis for policy measures for effective OJT program thereby focusing skills and competencies necessary for employability. It is encouraged that administration must design software that will integrate and employ the developed model to help in decision making and strategic management of the university.

6. References

1. Barron JM, Berger MC, Black DA. On-the-job training. 1997.
2. Verecio RL. On-the-Job Training of the BS Information Technology Program of Leyte Normal University, Tacloban City: An Assessment. International Journal of Education and Research. 2014; 2(3):1-10.
3. Rees C, Forbes P, Kubler B. Student Employability Profiles: A Guide for Higher Education Practitioners, 2nd Edition. The Higher Education Academy, York: United Kingdom. 2006; p. 1-83.
4. Harvey L. Defining and Measuring Employability. Quality in Higher Education. 2001; 7(2):97-109. https://doi.org/10.1080/13538320120059990
5. Finch D, Hamilton L, Baldwin R, Zehner M. An exploratory study of factors affecting undergraduate employability. Education Training. 2013; 55(7):681-704. https://doi.org/10.1108/ET-07-2012-0077
6. Sapaat M, Mustapha A, Ahmad J, Chamili K, Muhamad R. A data mining approach to construct a graduates employability model in Malaysia. International Journal of New Computer Architectures and their Applications. 2011; 1(4):1086-98.
7. Chien C, Chen L. Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. Expert Systems with Applications. 2008; 34(1):280-90. https://doi.org/10.1016/j.eswa.2006.09.003
8. Hand DJ. Principles of data mining. Drug Safety. 2007; 30(7):621-2. https://doi.org/10.2165/00002018-200730070-00010 PMid:17604416
9. Patil TR, Sherekar SS. Performance analysis of Naive Bayes and J48 classification algorithm for data classification. International Journal of Computer Science and Applications. 2013; 6(2):256-61.
10. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters. 2006; 27(8):861-74. https://doi.org/10.1016/j.patrec.2005.10.010
11. Bringula RP, Balcoba AC, Basa RS. Employable Skills of Information Technology Graduates in the Philippines: Do Industry Practitioners and Educators have the Same View? In Proceedings of the 21st Western Canadian Conference on Computing Education. 2016.
12. Castro EAG. Industry Participation in Developing Competencies for Employment Success: Learnings from a 3-Year OJT Program of A Philippine Higher Education Institution. Widyatama International Seminar. 2014; p. 1-11.
13. Buenviaje MG, del Mundo GV, A-Onuevo F, Martinez M. Employability of Business and Computer Management Graduates of one Higher Education Institution in the Philippines. Asia Pacific Journal of Multidisciplinary Research. 2015; 3(5):1-9.