ISSN 2278-3091
Volume 8, No.6, November – December 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse06862019.pdf
https://doi.org/10.30534/ijatcse/2019/06862019

A Response Assessment on the Implementation of Senior High School TVL Track through Data Mining Technique

Alicia Z. Maghuyop
College of Teacher Education, Surigao State College of Technology, Surigao City, Philippines
alicemaghuyop52@gmail.com

ABSTRACT

This paper re-evaluates the accurateness of the students' responses in the actual evaluation conducted in assessing the implementation of senior high school Technology Vocational Livelihood (TVL) track as basis to curricular intervention program done in some Senior High Schools in the Division of Surigao del Norte, Philippines. In this study, the use of data mining technique particularly the prediction was observed to test the accurateness of the responses as perceived by the 769 senior high school students. The use of the C4.5 algorithm depicted a 71% accuracy when applied in the evaluation dataset when performed using the 10-folds cross-validation scheme in WEKA software. The average accuracy percentage denotes that the findings, technologal philosophy formulated, and the proposed curricular instruction program from actual study is worthy of implementation. It is hoped that the study will contribute to the two major kinds of literature; data mining and in higher education mining.

Key words: Data mining, educational data mining, multiple intelligences, naïve bayes, reading comprehension,

1. INTRODUCTION

The major purpose of education is to help individual learners develop into a responsible and contributing members of the changing society. Since every learner is considered unique and important, his/her ultimate worth as a person has to be recognized. Education has a dual function as it influences the learners' lifestyle and also provides the foundation for a satisfying career [1].

To prepare learners for jobs and responsibilities not yet identified, the educational system faces the challenges of providing them with lifelong occupational and career developmental opportunities. They must be equipped with skills to thrive in an environment of complex experiences and rapid changes that necessitate school and curricular interventions.

The present unpredictable fate of the new educational program in the Philippines from global to ASEAN integration goes parallel with the challenges posited among the basic educational system. School heads, teachers and students are facing the changes of the curriculum. Among these are the qualification requirements of teachers, number of years for basic education, and the teaching-learning processes.

The Philippines has embarked on challenging educational reforms of improving the curriculum by adding two more years of learning. The Enhanced Basic Education Act institutionalized the implementation of the K to 12 program that aims to develop holistic graduates, equipped with 21st century skills. These graduates are projected to be prepared for higher education, middle level skills development, ready for employment and entrepreneurship. This is mostly being carried out by one salient feature, the senior high school program.

The main thrust of the Senior High School (SHS) is to produce productive and responsible citizens equipped with essential competencies, skills and values which will make them both a lifelong learners and employment-ready. Given this additional years of SHS and the new feature of the basic education program, it is necessary to report on the levels of assessment opportunities in order to ensure quality implementation, hence, the study of [1] was conducted.

However, conventional researchers still uses the old predefined queries and charts in conducting research. Some researcher-made questionnaires are deployed and the responses on these evaluations are tallied and are kept in databases. Interpretations are made through statistics and basic mathematical methods. The use of traditional qualitative and quantitative research methods does not give a well-established knowledge extraction from databases. Since the methods are outdated, the use of the latter limits the researchers in achieving their quality objectives. With the advent of technology, a new paradigm called data mining (DM) has emerged [2].

Therefore, this study used data mining technique particularly the C4.5 algorithm to evaluate and predict how accurate the responses of the student and teacher-respondents were in the study of [1] in assessing the implementation of TVL track in SHS in the Division of Surigao del Norte, Philippines . By determining the accuracy of the responses, the researcher would know if the baseline data and the findings are accurate and worthy of implementation. The result of this study will serve as a basis to a curricular intervention program and will be added to the literature of educational data mining and the application of the C4.5 algorithm in the education context.

2. LITERATURE REVIEW

Data mining, also called Knowledge Discovery from Databases (KDD) extracts implicit information or knowledge from databases using data mining and machine learning algorithms. Some of the well-known functions in data mining used diverse approaches of DM analysis, such as Bayes classifiers, clustering methods, neural networks, optimization algorithms, support vector machines, regressions, prediction and many more [3], [4]. These approaches encourage data

analysis in its full potential with the help of the right algorithms. Among the many data mining approaches, the prediction is considered as the prime method that is commonly used in educational data mining, business, health, and even in almost all sectors of the society [5].

Premised in the application of data mining in all areas in research, this trend is undoubtedly inevitable. DM is a medium of discovering novel and potentially valuable information from large amounts of data [6]. The data mining when applied to educational context is called Educational Data Mining (EDM). The field of EDM is new and developing in the area of the education sector [7]. This emerging field concerns with developing methods that discover knowledge from data from educational environments. Data mining algorithms such as C4.5, K-Nearest Neighbor, Naïve Bayes, neural networks, K-means clustering, and many others, are instrumental in extracting data from the datasets [8] used for classification, clustering, and prediction, among others.

Moreso, the use of data mining techniques has attracted more and more attention in this big data era [9]. Prediction, as one of the data mining techniques, can be made utilizing autoregressive integrated moving average (ARIMA) algorithm that used historical data in predicting cases extent in education, society, climate, health, and others. The ARIMA model was also used in forecasting dengue hemorrhagic fever cases in Southern Thailand [10]. Further, the use of the ARIMA algorithm in the prediction of travel time to the urban roadway was also discovered [11]. The Philippines' electric consumption [12] and inflation prediction [13] were also realized by the use of the ARIMA algorithm. Furthermore, the same algorithm has been used in forecasting incidence of hemorrhagic fever with renal syndrome in China [14].

Recent studies utilized and developed different models for prediction using Artificial Neural Networks [15], [16], Naïve Decision Tress [17]–[19], Bayes Classification [20]–[23], C4.5 algorithm [4], Random Forest [24]–[26], Support Vector Machine (SVM) [27]–[30], Apriori [31], [32], Linear Regression [33], [34], Logistic Regression Algorithm [35], [36], and Nearest Neighbors[37], [38].

A study of [4] predicted the accuracy of the instructional performance employed by the faculty in SUC's in the Caraga Region, Philippines. Two data mining techniques were utilized namely the prediction and clustering revealing an 87% prediction accuracy using the K-means integrated to C4.5 prediction model. Also, a study of [39] used K-Means segmentation technique and C4.5 algorithm to build a prediction model for customer loyalty in multimedia service provider. The integration of K-Means and C4.5 algorithm have yielded an increase of 79.33% accuracy prediction from the identified 69.23% accuracy with C4.5 algorithm alone.

Extent of crime analysis, the data mining techniques has been implemented to analyze crimes such as theft, homicide, and various drug offenses along with suspicious activities, noise complaints, and burglar alarm by using qualitative and quantitative approaches [40]. The rates of each crime and the cities with high crime rates have been identified using K-means clustering data mining approach applied in crime datasets from the New South Wales region of Australia. Further, an attempt to predict the index and non-index crimes in Surigao del Norte, Philippines was done using the K-Means clustering algorithm and ARIMA algorithm. Predicting crime occurrence for the years 2018-2022 was achieved based on the historical crime data from years 2013-2017 [41]. Another study examined the specific trends and patterns of terrorist attacks in India. The K-means clustering algorithm was used to determine the year wherein the terrorist groups were most active and also which terrorist group has affected the most [42].

3. METHODOLOGY

3.1 Datasets

The respondents of this study were the students of Grade 12 in the Division of Surigao del Norte. Specifically the schools that are involved are the following: Alegria National High School, Bacuag National Agro Industrial School, Claver National High School, Gigaquit National High School, Placer National High School, Mainit National High School, Malimono National High School, Surigao Norte National High School, Tagana-an National High School, San Francisco National High School, Toledo National High School and Tubod National High School consisting of 769 instances and 30 variables that are grouped according to facilities and equipment, industry preparedness, workplace environment, instruction, and imposition, excerpt from [1].

Table 1: Variables used in the study

Category	Content	Variable	Possible Value
Facilities and Equipment	Classrooms are provided with alternative technology facilities like TV,video player,etc to help abreast with modernity.	F1	1,2,3,4
	Math, Science and English laboratory are equipped with necessary teaching-learning aids,services, tools and gadgets	F2	1,2,3,4
	School library is equipped with updated collection of books, textbooks, periodical newspapers, magazines, etc.	F3	1,2,3,4
	Health-promoting facilities are provided like canteen, comfort rooms, drinking fountains, safe playground, garbage bins, etc.	F4	1,2,3,4
	Student services facilities are found like medical clinic, counseling room, gymnasium, visitors waiting area, etc.	F5	1,2,3,4
	Classroom physical structures like tables, desks, chairs, cabinets, bulletin board displays are conspicuously provided.	F6	1,2,3,4
Industry Preparedness The school…	Provides the students functional knowledge and skills to earn a living.	IP1	1,2,3,4
	Prepares them of their physical, psychological well-being, and welfare.	IP2	1,2,3,4
	Helps the students in the completion of necessary certification requirements	IP3	1,2,3,4
	Awares the students of the impacts and implications of their basic learning to	IP4	1,2,3,4

	survival in the world of work		
	Strengthens existing linkages to industries and forging new alliances maximizing social participation	IP5	1,2,3,4
	Reinforces skills training of students to ensure relevant industry-demand based courses	IP6	1,2,3,4
Workplace Environment	The physical structure of the specific work area like classroom or shoproom in conducive to doing work.	W1	1,2,3,4
	Properly ventilated rooms where cleanliness is maintained.	W2	1,2,3,4
	Availability of facilities, equipment, modern technology devices, tools, gadgets prepared and ready for use.	W3	1,2,3,4
	Locked cabinets are in place for storage and safe-keeping of industry or home tools, devices, utensils, etc.	W4	1,2,3,4
	Work area is monitored, supervised by a designated official to check inefficiency.	W5	1,2,3,4
	Harmonious relationship among students, teachers, and administrators are observed.	W6	1,2,3,4
Instruction The school…	Complies with program offerings of subject and courses study aligned to the New Enhanced basic Education Curriculum	IN1	1,2,3,4
	Provides teachers the needed instructional technologies, materials and aids to secure quality learning students	IN2	1,2,3,4
	Equips students with knowledge, skills to meet the demands of higher ladder and labor market	IN3	1,2,3,4
	Enhances competencies of teachers by monitoring the carry-over of skills learned from in-service trainings through follow-up techniques	IN4	1,2,3,4
	Applies variety of teaching methods, approaches, techniques and strategies aligned with students' diversified needs and interests.	IN5	1,2,3,4
	Monitors performance and progress students and teachers through periodic conduct of assessment and evaluation.	IN6	1,2,3,4
Imposition	Observes strictly the implementation of the K-12 new curriculum program as mandated by Basic Education Act of 2013.	IM1	1,2,3,4
	Adopts the outcomes-based education (OBE) course syllabi following the content, performance and competency standards.	IM2	1,2,3,4
	Aligns lesson objectives with assessment procedures and standards.	IM3	1,2,3,4
	Conducts or sends teachers to attend seminars, training and workshops for retooling and updating standards of quality teaching and learning.	IM4	1,2,3,4
	Obliges teachers strong commitment of this missionary role to develop greater competence in teaching one's subject area.	IM5	1,2,3,4
	Encourages active participation of teachers in research studies and professional development.	IM6	1,2,3,4

3.2 C4.5 Algorithm

Excerpt form the study of [3], the C4.5 algorithm is referred to as a descendant to ID3 model which was created by J. Ross Quinlan and is also grounded on Hunt's algorithm. The C4.5 algorithm is considered to be one of the most famous decision tree algorithm in machine learning. This claim was asserted to be true up to date by [44]. C4.5 uses both categorical and continuous attributes in building a decision tree (DT). Splitting the attribute values into two according to the identified threshold where all values greater than the threshold is considered as one child hence, the other shows how robust C4.5 in handling continuous attributes. This algorithm is also known to handle missing attribute values. In building the DT, C4.5 uses Gain Ratio as attribute selection measure by removing the biases of information gain if many outcome values of an attribute are identified.

To perform, compute the gain ratio of each attribute first. Attributes whose gain ratio is at maximum will be identified as the root node of the tree. The algorithm uses a pessimistic pruning approach in removing unnecessary branches in the decision tree to increase the classification accuracy.

3.3 Prediction Evaluation Tools

To evaluate, the accuracy, precision, recall, and F-measure, mean absolute error (MAE), and root mean squared error (RMSE), are used which conformed to the formula from the study of [34] as to wit:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$F - measure = 2 \times \frac{precision \times recall}{precision + recall} \quad (4)$$

$$MAE = \sum_{t=T+1}^{T+h} | \hat{y}_t - \hat{y}_t | / h \quad (5)$$

$$RMSE = \sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - \hat{y}_t)^2 / h} \quad (6)$$

The true positive, true negative, false positive and false negative, are represented by TP, TN, FP, and FN, respectively. The lower the statistical error value is, the better the forecasting ability of the model.

4. RESULT AND DISCUSSION

To predict the accurateness of the responses of the students, the dataset was loaded to the Waikato Environment for Knowledge Analysis (WEKA) software using the 10-folds cross-validation scheme. The simulation result revealed a 71% prediction accuracy denoting that the responses are in acceptable level of accuracy and the model predicted a 71% correctly classified instances. The prediction model obtained a prediction accuracy of 71% making the responses of the students in the study of [1] reliable and worthy of implementation as shown in Tables 2-3. The precision of the dataset has an overall score of 70%. This denotes that 70% of the dataset are correctly classified by the algorithm when used in WEKA.

Further, the recall metric denotes that 71% of the instances in the dataset were retrieved by the system correctly and revealed a good performance. The f-measure showed an 70.6% balance performance for both precision and recall of the model. Lastly, the RMSE and MAE showed a zero-based value of 0.515 and 0.3327, respectively revealing how concentrated the prediction is. The low values for RMSE and MAE depicts a good forecast.

Table 2: Prediction model accuracy evaluation

Model	Accuracy %	Precision	Recall	F-Measure
C4.5	71.0145%	0.704	0.710	0.706

Table 3: Forecast error evaluation

RMSE	MAE
0.515	0.3327

5. CONCLUSION

Predicting the accurateness of the students' responses in the study conducted by [1] is essential. It is important to cross-validate and re-evaluate the responses of the students to know if the knowledge generated based on the findings from the actual research conducted, as well as the proposed intervention plan, is a must to implement.

Since it is not enough that research is conducted using predefined methods, the use of the data mining technique particularly the C4.5 algorithm for prediction, was undertaken. The simulation result showed the effectiveness of data mining technique particularly the use of C4.5 algorithm in knowledge extraction and validation. The use of the algorithm in predicting the accurateness of the responses helped the researcher in validating the acceptability level of the knowledge acquired when the predefined methods were used prior to its implementation. It is recommended that the output of this study be used in data proliferation and influence administrator's decision.

REFERENCES

[1] R. E. Tumbaga, "Assessment on the Implementation of Senior High School TVL Track: Basis to Curricular Intervention Program," 2019.

[2] K. Rajalakshmi, S. S. Dhenakaran, and N. Roobini, "Comparative Analysis of K-Means Algorithm in Disease Prediction," *Int. J. Sci. Eng. Technol. Res.*, vol. 4, no. 7, pp. 2697–2699, 2015.

[3] A. J. P. Delima, "An Experimental Comparison of Hybrid Modified Genetic Algorithm-based Prediction Models," *Int. J. Recent Technol. Eng.*, vol. 8, no. 1, pp. 1756–1760, 2019.

[4] U. O. Cagas, A. J. P. Delima, and T. L. Toledo, "PreFIC : Predictability of Faculty Instructional Performance through Hybrid Prediction Model," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 7, pp. 22–25, 2019.

[5] A. J. P. Delima, A. M. Sison, and R. P. Medina, "Variable Reduction-based Prediction through Modified Genetic Algorithm," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 5, pp. 356–363, 2019. https://doi.org/10.14569/IJACSA.2019.0100544

[6] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2012.

[7] P. Kaur, M. Singh, and G. S. Josan, "Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector," *Procedia Comput. Sci.*, vol. 57, pp. 500–508, 2015. https://doi.org/10.1016/j.procs.2015.07.372

[8] S. K. Yadav and S. Pal, "Data Mining : A Prediction for Performance Improvement of Engineering Students using Classification," *World Comput. Sci. Inf. Technol. J. WCSIT*, vol. 2, no. 2, pp. 51–56, 2012.

[9] C. Jinyin, L. Xiang, Z. Haibing, and B. Xintong, "A novel cluster center fast determination clustering algorithm," *Appl. Soft Comput. J.*, vol. 57, pp. 539–555, 2017. https://doi.org/10.1016/j.asoc.2017.04.031

[10] S. Promprou, M. Jaroensutasinee, and K. Jaroensutasinee, "Forecasting dengue haemorrhagic fever cases in Southern Thailand using ARIMA Models," *Dengue Bull.*, vol. 30, pp. 99–106, 2006.

[11] D. Billings and Y. Jiann-Shiou, "Application of the ARIMA Models to Urban Roadway Travel Time Prediction-A Case Study. Systems, Man and Cybernetics, 2006. SMC'06," *IEEE Int. Conf.*, pp. 2529–2534, 2006. https://doi.org/10.1109/ICSMC.2006.385244

[12] A. J. P. Delima, "Application of Time Series Analysis in Projecting Philippines ' Electric Consumption," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 5, pp. 694–699, 2019. https://doi.org/10.18178/ijmlc.2019.9.5.860

[13] A. J. P. Delima and M. T. Q. Lumintac, "Application of Time Series Analysis for Philippines ' Inflation Prediction," *Int. J. Recent Technol. Eng.*, vol. 8, no. 1, pp. 1761–1765, 2019.

[14] Q. Liu, X. Liu, B. Jiang, and W. Yang, "Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model," 2011.

[15] I. Kayri and M. T. Gencoglu, "Prediction of Power Production from a Single Axis Photovoltaic System by Artificial Neural Networks," pp. 212–215, 2017. https://doi.org/10.1109/EMES.2017.7980417

[16] J. Jiao, C. Chen, X. Tian, C. U. Hospital, and J. Universiy, "propagation artificial neural network model (BP-ANN) model in cirrhosis patients," 2016.

[17] X. Zhang, X. Wang, W. Chen, J. Tao, W. Huang, and T. Wang, "A Taxi Gap Prediction Method via Double Ensemble Gradient Boosting Decision Tree," *2017 IEEE 3rd Int. Conf. Big Data Secur. Cloud*

(BigDataSecurity), IEEE Int. Conf. High Perform. Smart Comput. IEEE Int. Conf. Intell. Data Secur., pp. 255–260, 2017.

[18] I. Bjij, I. Hdoufane, and A. Jarid, "Molecular modeling : Application of Support Vector Machines and Decision Trees for anti-HIV activity prediction of organic compounds," pp. 1–4, 2016.

[19] C. Wang and M. Lin, "Prediction of Accrual Expenses in Balance Sheet Using Decision Trees and Linear Regression," pp. 73–77, 2016.

[20] N. K. Walia, P. Kalra, and D. Mehrotra, "Prediction of carbon stock available in forest using naive bayes approach," *Proc. - 2016 2nd Int. Conf. Comput. Intell. Commun. Technol. CICT 2016*, pp. 275–279, 2016.

[21] T. Mori, "Superposed naive bayes for accurate and interpretable prediction," *Proc. - 2015 IEEE 14th Int. Conf. Mach. Learn. Appl. ICMLA 2015*, pp. 1228–1233, 2016. https://doi.org/10.1109/ICMLA.2015.147

[22] S. Aneja and S. Lal, "Effective asthma disease prediction using naive Bayes - Neural network fusion technique," *Proc. 2014 3rd Int. Conf. Parallel, Distrib. Grid Comput. PDGC 2014*, pp. 137–140, 2015.

[23] S. L. Menezes and G. Varkey, "Prediction of Missing Items Using Naive Bayes Classifier and Graph Based Prediction," *2013 Third Int. Conf. Adv. Comput. Commun.*, pp. 39–45, 2013.

[24] W. Xu, J. Zhang, Q. Zhang, and X. Wei, "Risk prediction of type II diabetes based on random forest model," *3rd Int. Conf. Adv. Electr. Electron. Information, Commun. Bio-Informatics*, pp. 382–386, 2017.

[25] P. Khongchai and P. Songmuang, "Random Forest for Salary Prediction System to Improve Students' Motivation," *Proc. - 12th Int. Conf. Signal Image Technol. Internet-Based Syst. SITIS 2016*, pp. 637–642, 2017. https://doi.org/10.1109/SITIS.2016.106

[26] N. Mago, S. Srivastava, R. D., L. E. S. Lewis, and S. M, "Prediction of Apnea of Prematurity in Neonates using Support Vector Machines and Random Forests," pp. 693–697, 2016.

[27] A. K. Patel, "Development of Online Machine Vision System using Support Vector Regression (SVR) Algorithm for Grade Prediction of Iron Ores," pp. 149–152, 2017.

[28] W. Zhang, H. Zhang, J. Liu, K. Li, D. Yang, and H. Tian, "Weather Prediction With Multiclass Support Vector Machines in the Fault Detection of Photovoltaic System," vol. 4, no. 3, pp. 520–525, 2017.

[29] H. Zhang, X. Zheng, and J. Tian, "Virtual Network Mapping based on the Prediction of Support Vector Machine," 2016. https://doi.org/10.1109/ITME.2016.0197

[30] S. Kaur and D. S. Kalra, "Disease Prediction using Hybrid K-means and Support Vector Machine," 2016. https://doi.org/10.1109/IICIP.2016.7975367

[31] W. Min, Z. Hui, Y. Longxiang, and Z. Hongbo, "Prediction Algorithm In Mobile Social Environment," 2015.

[32] L. Jena and N. K. Kamila, "A Model for Prediction of Human Depression Using Apriori Algorithm," *2014 Int. Conf. Inf. Technol.*, pp. 240–244, 2014. https://doi.org/10.1109/ICIT.2014.65

[33] J. Sahlin, R. Eriksson, M. T. Ali, and M. Ghandhari, "Transmission Line Loss Prediction Based on Linear Regression and Exchange Flow Modelling," 2017. https://doi.org/10.1109/PTC.2017.7980810

[34] M. Rahim and B. Thirion, "Multi-output predictions from neuroimaging : assessing reduced-rank linear models."

[35] A. I. J. Tostes, T. H. Silva, R. Assuncao, F. L. P. Duarte-Figueiredo, and A. A. F. Loureiro, "STRIP: A Short-Term Traffic Jam Prediction Based on Logistic Regression," *2016 IEEE 84th Veh. Technol. Conf.*, pp. 1–5, 2016.

[36] S. Usman and A. Subject, "Risk Prediction Of Having Increased Arterial Stiffness Among Diabetic Patients Using Logistic Regression," pp. 699–701, 2016.

[37] H. Al-Shehri *et al.*, "Student performance prediction using Support Vector Machine and K-Nearest Neighbor," *Can. Conf. Electr. Comput. Eng.*, pp. 17–20, 2017.

[38] Z. Wang and I. Koprinska, "Solar Power Prediction with Data Source Weighted Nearest Neighbors," pp. 1411–1418, 2017. https://doi.org/10.1109/IJCNN.2017.7966018

[39] S. Moedjiono, Y. R. Isak, and A. Kusdaryono, "Customer Loyalty Prediction In Multimedia Service Provider Company With K-Means Segmentation And C4 . 5 Algorithm," in *2016 International Conference on Informatics and Computing (ICIC)*, 2016, pp. 1–6.

[40] J. Agarwal, "Crime Analysis using K-Means Clustering," *Int. J. Comput. Appl.* , vol. 83, no. 4, pp. 975–8887, 2013.

[41] A. J. P. Delima, "Applying Data Mining Techniques in Predicting Index and non-Index Crimes," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 4, pp. 533–538, 2019.

[42] P. Gupta, A. S. Sabitha, and T. Choudhury, "Terrorist Attacks Analysis Using Clustering Algorithm," © *Springer Nat. Singapore Pte Ltd.*, pp. 317–328, 2018.

[43] S. L. Salzberg, "Book Review: C4.5: by J. Ross Quinlan. Inc., 1993.," *Mach. Learn.*, vol. 16, pp. 235–240, 1994.

[44] R. Benkercha and S. Moulahoum, "Fault detection and diagnosis based on C4.5 decision tree algorithm for grid connected PV system," *Sol. Energy*, vol. 173, pp. 610–634, 2018. https://doi.org/10.1016/j.solener.2018.07.089

[45] J. L. M. Segovia, "Multiple Intelligences and Reading Comprehension of Senior High School Students in San Jose District of Dinagat Islands," 2019.

[46] A. J. P. Delima and V. C. Francisco, "Rank-based variable minimization using clustering algorithm," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 5, pp. 575–580, 2019. https://doi.org/10.18178/ijmlc.2019.9.5.842