ISSN: 2633-4828 Vol. 6 No.1, January, 2024

International Journal of Applied Engineering and Technology

Machine Learning Approaches in Predicting Households' Compliance to Maternal and Child Health Conditions of a Philippine Welfare Program

John Vianne Murcia[1] (Corresponding author)
[1]*Business Analytics Program, College of Business Administration Education, University of Mindanao, Davao City, Philippines,* jv_murcia@umindanao.edu.ph

Marlon Suelto[2]
[1]*Environmental Science Program, College of Arts and Sciences Education, University of Mindanao, Davao City, Philippines,* msuelto@umindanao.edu.ph

Joane May Delima[3], John Raven Manulat[4], Conrado Panerio, Jr.[5]
UM Digos College, Digos City, Philippines, [3]jmdelima@umindanao.edu.ph, [4]johnravenmanulat@umindanao.edu.ph, [5]conrado.panerio@umindanao.edu.ph

Date of Submission: 15th November 2023 Revised: 27th December 2023 Accepted: 18th January 2024

How to Cite: **Murcia, J. V., Suelto, M., Delima, J. M., Manulat, J. R. and Panerio, C. J. (2024). Machine Learning Approaches in Predicting Households' Compliance to Maternal and Child Welfare Program. International Journal of Applied Engineering and Technology Research 5(4), pp.31-37.**

Abstract **- The study successfully utilized various supervised and unsupervised data mining algorithms to categorize families as either compliant or non-compliant with regards to the maternal and child health and nutrition provisions of the Philippine government's conditional cash transfer program. The dataset was extracted from the advanced system known as e4PsMap, which stands for electronic 4Ps Mapping. A grand total of 1,357 instances underwent training and were submitted to a five-fold cross-validation process. The machine learning operations were executed using the WEKA software. Results revealed that Decision Tree (J48) in all confidence factors and Random Forest had the maximum number of successfully categorized examples which instantaneously equals the highest classification accuracy (100%), the highest F-measure (1.00), and the highest kappa statistics (1.00). Implications are discussed.**

Index Terms **- machine learning, decision tree, Random Forest, Random Tree, Logistic, Multilayer Perceptron, *k*-nearest neighbor, Weka.**

INTRODUCTION

It has long been evident that the health of individuals is closely tied to poverty, and those experiencing poverty are particularly vulnerable to health risks, particularly children and pregnant women[1]. Poverty can be assessed by considering socioeconomic position and can be confirmed by examining educational achievement, occupation, and income[2]. The poverty rate in the Philippines surged to 18.1% in 2022, up from 16.7% in 2016, resulting in 19.99 million Filipinos living below the poverty line[3].

In order to combat poverty, the Philippine government has established social assistance programs through the Department of Social Welfare and Development (DSWD), namely the Pantawid Pamilyang Pilipino Program (4Ps). Since 2007, the Department of Social Welfare and Development (DSWD) has enforced the 4Ps, as a component of the national flagship social assistance program designed to offer conditional cash payments to impoverished Filipino families[4]. The agency's latest quarterly report indicates that there were 4.41 million real recipients, surpassing the objective by 0.3%[5]. Efficiently collecting, comprehensively integrating, and monitoring the vast amount of data and information from the beneficiaries is necessary to successfully evaluate the program's performance for future adjustments and assessments. Hence, it is imperative to introduce a novel and methodical intervention to ensure the effective execution of the program and avoid any potential disasters in the future.

The absence of well-defined criteria as qualifying indicators, while utilizing a decision support system for the social welfare department in the Philippines, to forecast eligible families for the Pantawid Pamilyang Pilipino Program (4Ps), continues to provide a barrier in terms of allocating and distributing funds appropriately. The social welfare institution has recognized that the growing number of complaints and concerns regarding the selection criteria for qualified 4Ps beneficiaries could have been foreseen by implementing a recommender system. This system would assist the institution in identifying program beneficiaries by analyzing relevant information from databases.

Nevertheless, the absence of a decision support system is prevalent in the Philippines, particularly in the Social Welfare and Development Department. This poses a significant obstacle to the implementation of RA 11310 and RA 5416, often known as the Social Welfare Act of 1968.

The processing and analysis step encounter substantial difficulties when dealing with bulk data and huge data sets[6]. The study of substantial and extensive datasets necessitates expertise in various fields such as engineering science, information technology, and security [7]. Consequently, comprehensive research was carried out to assess the effectiveness of implementing the 4Ps program. However, little focus was placed on the incorporation of many domains and the utilization of data analytics and prediction. The objective of this study is to utilize diverse data mining techniques to forecast maternal and child outcomes in order to evaluate the effectiveness of the program. Various techniques were employed to cluster, categorize, and display the datasets. Notwithstanding these constraints, the social welfare organization believed that implementing a decision support system would consistently include ongoing efforts to evaluate the efficacy of their budget allocation.

Until recently, the idea of combining data from several databases and evaluating it was not possible. However, advancements in data extraction and analysis techniques, along with increased computing power, also known as big data, have made this possible. These accomplishments have been commended by some, who see them as marking the beginning of a new period in the field of social work, similar to what has occurred in public health and health promotion[8]. Implementing data mining techniques in the social welfare sector presents significant difficulties[9]. Due to strict legal requirements, a substantial volume of data has been gathered. Nevertheless, the sector faces difficulties in managing data due to its insufficiency and lack of completeness. There is a lack of a system to record and utilize this data for data-driven analysis and decision-making, which is crucial for ensuring the accuracy of applicant qualifications, improving social service objectives, payment accuracy, and compliance[10].

Several studies have been undertaken that resemble the study. An examination of the current body of research on information systems [11][12] indicates a scarcity of documented evidence pertaining to the implementation of the Decision Support System for Social Welfare. The study investigates the application of data mining techniques to identify trends in social welfare data, namely health indicators, in Jamaica. It employs an integrated model for knowledge discovery and data mining in social welfare programs. The Computer Aided Registration Evaluation System (CARES) employed the decision tree method and achieved a test accuracy of 94%.

Furthermore, a study[13] developed a system to aid the Social Services office of Jogjakarta City in assessing the eligibility of low-income families for the Family Hope Program. This system incorporates a calculation feature that utilizes the fuzzy logic simple additive weighting method, facilitating the agency's determination of the suitability of providing social welfare assistance to low-income families.

In addition, [14] created a decision support system for selecting recipients of direct cash from the village fund. This system utilizes the multi-attribute utility theory (MAUT) to identify suitable community candidates who should receive financial help. The findings demonstrated a test accuracy of 92.75%.

The study employed the Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) to develop a model for identifying priority users of direct financial assistance. These priority users are determined based on pre-established criteria and must be handled first. The study determined that the system can help rate the recipients of assistance who should be recommended to receive direct monetary support.

In [16], utilizing the Simple Additive Weighting (SAW) method, the researchers suggested appropriate candidates for the Poor Students Assistance Funds of a private college. The study found that implementing the SAW technique resulted in an 81% increase in processing speed and accuracy. [17] developed a decision support system to enhance the administration and allocation of assistance. The study employed the Simple Additive Weighting (SAW) approach in conjunction with a Fuzzy Multi-Attribute Decision-Making model (FMADM). The study's findings indicate that it can determine the best choice by evaluating and assigning a weight value to each criterion, and provide assistance based on the available finances.

Building on the existing literature and advancements in data mining techniques, this study aims to innovate in the field of social welfare by developing a sophisticated decision support system. The study aims to utilize big data analytics to forecast maternal and child outcomes inside the Pantawid Pamilyang Pilipino Program (4Ps), hence assessing the effectiveness of the program. The approach entails the grouping, categorization, and visualization of extensive datasets, necessitating proficiency in multiple fields such as engineering science, information technology, and security. The utilization of data mining in social welfare, as demonstrated in this novel approach, is both innovative and timely, particularly in light of the pressing necessity for more streamlined and precise processes within the Philippines' Social Welfare and Development Department. The study aims to fill the existing gaps in data-driven decision-making and analysis, with the goal of guaranteeing the precision of applicant credentials, enhancing payment accuracy, and ensuring adherence to the applicable legal requirements.

Through this approach, it aims to transform the management and distribution of social welfare aid, making a substantial impact on the existing research in this field and establishing a fresh benchmark for the application of technology in social welfare initiatives.

METHODS

Dataset

The City Government of Digos's Social Welfare and Development Office released information on the child and maternal health and nutrition results of members of the Pantawid Pamilyang Pilipino Program (4Ps). The extraction process resulted in the identification of nine (9) relevant attributes from the CSV file. These attributes consist of eight (8) explanatory attributes and one class attribute. A new attribute, matchild_health_score, was created. This attribute is a composite index that assigns weights to the attributes that represent the conditions provided by the 4Ps. The equation used to calculate the matchild_health_score is as follows:

$$SCORE = \sum (IMMU * 0.5 + MATERNAL * 0.5)$$

which can be further extended to:

$$SCORE = \sum \left(\left[DEW * 35 + \sum_{i=1,\dots,13}^{n} VACC * 65 \right] * 0.5 + [PRE * 50 + POST * 50] * 0.5 \right)$$

where

IMMU - mandatory immunization of children
MATERNAL - mandatory maternal health monitoring
DEW - administered deworming as a component of immunization
VACC - administered 13 vaccines for children as a component of immunization
PRE - prenatal checkups of mothers as a component of maternal health monitoring
POST - postnatal checkups of mothers as a component of maternal health monitoring

Table 1.
Dataset Description

Attribute	Type	Description
hh_head	Nominal	The head of the household, either the father (1) or mother (2).
mon_income	Nominal	Income brackets of households in PhP, which can either be below PhP 5000 (1), PhP 5001 to 7499 (2), PhP 7500 to 9999 (3), PhP 10000 to 12000 (4) or PhP 12000 or more (5)
num_fammem	Numeric	Number of family members in the household expressed as real numbers.
num_child	Numeric	Number of children ages 0 to 18 years old in the household expressed as real numbers.
pregnant	Nominal	Presence (1) or absence (0) of pregnant family members in the family.
year_resid	Nominal	Year where the family started their residence expressed as actual ordinal years.
type_resid	Nominal	Type of residence of the household, defined either as rented (1), owned (2), or living with parents/children (3).
house_desc	Nominal	Type of house based on the material used in the construction, defined either as concrete (1), semi-concrete (2), wood (3) or makeshift (4).
matchild_health_score	Numeric	Maternal and child health and nutrition score based on several conditionality attributes of the 4Ps, expressed as scaled values 0 to 100, computed based on Equations 1 and 2.
compliance	Nominal	Compliance category: household is compliant (1) or not (0).

Feature Selection

Prior to the classification phase, feature selection was conducted using three renowned feature selection algorithms in Weka[18] to assure the inclusion of important features. The techniques used include correlation-based feature selection (`CorrelationAttributeEval`), information gain-based or entropy-based feature selection (`InfoGainAttributeEval`), and learner-based feature selection (`WrapperSubsetEval`), in accordance with the given proposal. The correlation-based feature selection identified the top five attributes that are most strongly correlated with the class attribute.

These attributes are matchild_health_score (r=0.8488), num_child (r=0.3508), num_fammem (r=0.056), hh_head (r=0.0543), and year_resid (r=0.0495). On the other hand, the four attributes that have the weakest correlation with the class attribute are type_resid (r=0.008), pregnant (r=0.009), house_desc (r=0.0098), and mon_income (r=0.014). It has been proposed that attributes with r-values greater than or equal to 0.20 should be kept in order to get the best classification performance. However, only two attributes (matchild_health_score and num_child) met this condition.

In terms of information gain (entropy)-based feature selection, the top three attributes with the highest rankings are matchild_health_score (0.6411), num_child (0.21059), and num_fammem (0.06369). On the other hand, the attributes with the lowest values are type_resid (r=0.00329), pregnant (0.00285), house_desc (0.00341), hh_head (0.00226), and mon_income (0.00209), which have very low entropy values. A recommendation was made to set an arbitrary threshold for entropy/information gain of 0.05. However, only three qualities met this requirement.

The learner-based feature selection, using the WrapperSubsetEval attribute evaluator with the BestFit search method, determined that a five-fold classification is appropriate for the task at hand. WrapperSubsetEval can automatically pick attributes, similar to the variable suppression process in stepwise regression[19]. The findings indicated that out of the nine attributes considered for classifying the class attribute, only matchild_health_score is deemed suitable with a merit of 83.7%. Modifying the search technique parameter to GreedyStepwise additionally confirmed the presence of the same attribute that is likely to be included in the classification process.

Data Classification and Cross-validation

A total of six (6) classifiers were chosen to carry out the classification of the training dataset and predict the class label in the test set. We utilized widely-used classifiers, including *k*-NN (lBk), NaiveBayes, Decision Trees (J48), and Logistic, along with classifier techniques such as Random Tree, Random Forest, and Multilayer Perceptron.

Each classifier generally possesses parameters that require tuning. During this phase, adjustments were made to the parameter values for the number of *k* in the k-NN algorithm and the confidence factor of the decision tree (J48). In the *k*-NN algorithm, increasing the value of k reduces the likelihood of error[20]. On the other hand, in the J48 algorithm, the confidence factor is adjusted to evaluate the success of post-pruning[21]. The *k*-NN algorithm was used for classifications with values of *k* set to 3, 5, 7, 9, and 11. Similar to the reference [22], we made the decision to omit the default *k*-NN 1.

The value of *k* was initialized as 3, which indicates that the three nearest classes will be examined, and the label that appears most frequently will be applied. This principle also applies to larger values of *k*. However, the confidence factor in the J48 classifier was adjusted in increments of 0.25 from its default value of 0.25. Three categorization techniques were conducted using different thresholds (C.0.25, C.0.5, C.0.75).

To ensure that both classes are evenly distributed in each fold during cross-validation, the unsupervised filter Randomize was employed to randomly rearrange the original dataset and allocate the instances among the five folds[23].

RESULTS AND DISCUSSION

There are 12 cross-validations conducted using the four classifiers: three using J48, four using IBk (*k*-NN), one for NaiveBayes, one for Logistic, one for Random Tree, one for Random Forest, and one for Multilayer Perceptron. Table 2 demonstrates that the majority of classifiers exhibited outstanding performance on the training dataset, with a few obtaining flawless or very flawless classification accuracy. The J48 classifier, in all its variations (C.0.25, C.0.50, C.0.75), demonstrated outstanding performance by obtaining a classification accuracy rate of 100%. The J48 algorithm successfully classified all cases in the training dataset, regardless of the confidence factor. The J48 classifier's exceptional accuracy, demonstrated by a Kappa statistic of 1.00, indicates an impeccable degree of agreement that exceeds chance. This underscores the classifier's efficiency in this specific dataset.

By comparison, the Random Tree classifier exhibited a little lower, yet still remarkable, classification accuracy of 96.9049%, accompanied by a Kappa value of 0.884. This implies a significant level of agreement that is not due to chance, but it also signals that there is some potential for improvement when compared to the J48 classifier. The Random Forest classifier achieved a classification accuracy of 100% and a Kappa statistic of 1.00, which indicates its strong performance and reliability in accurately categorizing occurrences in the training dataset.

The lBk classifier exhibited different levels of accuracy depending on the chosen k-values. The most accurate results were obtained with *k*-NN 5, achieving a precision of 93.073%. Meanwhile, *k*-NN 3 and *k*-NN 9 both performed well, with a precision of 92.336% each. The *k*-NN 7 version exhibited a little reduced accuracy rate of 92.115%. These findings, coupled with Kappa values ranging from 0.658 to 0.709, indicate that the lBk classifier is successful but its performance is affected by the selection of the k-value. Furthermore, it does not achieve the same level of perfection as the J48 and Random Forest classifiers.

Table 2.
Classification Accuracy of the Classifiers and their Variants on Training Dataset

Classifier	Variants	Correctly Classified Instances	κ
J48	C.0.25	1357 (100%)	1.00
J48	C.0.50	1357 (100%)	1.00
J48	C.0.75	1357 (100%)	1.00
Random Tree	-	1315 (96.9049%)	0.884
Random Forest	-	1357 (100%)	1.00
lBk	3	1253 (92.336%)	0.672
lBk	5	1263 (93.073%)	0.709
lBk	7	1250 (92.115%)	0.658
lBk	9	1253 (92.336%)	0.667
Naïve Bayes	-	1356 (99.9263%)	0.997
Logistic	-	1318 (97.126%)	0.898
Multilayer Perceptron	-	1352 (99.6315%)	0.986

The NaiveBayes classifier demonstrated an exceptional classification accuracy of 99.9263%, along with a Kappa statistic of 0.997, suggesting a high level of agreement beyond what would be expected by chance. The Logistic classifier achieved a classification accuracy of 97.126% and a Kappa statistic of 0.898, indicating a significant level of agreement. The Multilayer Perceptron demonstrated exceptional performance in classifying occurrences, with a classification accuracy of 99.6315% and a Kappa statistic of 0.986.

In addition to evaluating the accuracy of the classifiers, various factors were taken into account to determine the effectiveness of cross-validation on the training set for each classifier and their parameter-tuned sub-analyses. These factors include kappa values, average F-measure, true positive and false positive rates, precision value, and recall value. The evaluation included all classifiers, including variations of the runs with parameter tuning on k nearest neighbor (IBk) and decision trees (J48). The overview of the categorization performance data is presented in Table 3.

As previously mentioned, J48 and Random Forest had the strongest performance across several evaluation metrics, including correctly categorized cases, kappa statistic, true positive and false positive rates, precision, recall, and F-measure. Although *k*-NN 5 has the lowest classification accuracy, it has the highest accuracy among the lBk variations. However, when the number of k was further increased, there was a noticeable decline in the performance of the classifier. This finding aligns with the results of the feature extraction analysis conducted using WrapperSubsetEval, which recommended using five cross-validation folds for optimal classifier performance.

Table 3.
Classification performance of the utilised classifiers and their variants on training dataset (*N*=1357)

Classifier	Variants	TP Rate	FP Rate	Precision	Recall	F-measure
J48	C.0.25	1.00	1.00	1.00	1.00	1.00
	C.0.50	1.00	1.00	1.00	1.00	1.00
	C.0.75	1.00	1.00	1.00	1.00	1.00
RandomTree	-	0.969	0.108	0.969	0.969	0.969
RandomForest	-	1.00	1.00	1.00	1.00	1.00
lBk	3	0.923	0.347	0.922	0.923	0.916
	5	0.931	0.312	0.930	0.931	0.925
	7	0.921	0.365	0.921	0.921	0.913
	9	0.923	0.361	0.924	0.923	0.916
Naïve Bayes	-	0.999	0.000	0.999	0.999	0.999
Logistic	-	0.971	0.046	0.973	0.971	0.972
Multilayer Perceptron	-	0.996	0.015	0.996	0.996	0.996

The Decision Tree (J48) and Random Forest classification algorithms were identified as the most effective methods for assessing compliance of households with maternal and child health and nutrition conditions under the 4Ps program. We utilized these classifiers to generate visual representations of the decision trees, enabling us to identify the specific threshold at which the compliance score (matchild_health_score) is considered compliant (1) or non-compliant (0). Figure 2 shows that the J48 algorithm determined that households classed as 4Ps need to achieve a score of 42.5 or higher in order to be considered compliant with the program. On the other hand, Figure 3 demonstrates that the Random Forest algorithm set a threshold of 46.5 as the minimum score for households to be labeled as compliant.

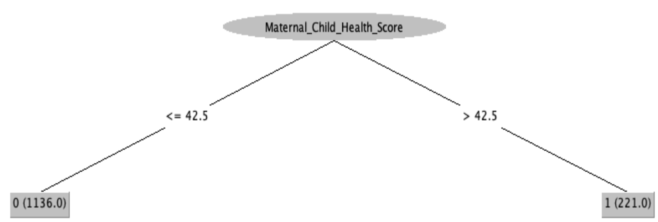

Figure 2. Decision tree (J48) classifier showing household compliance to 4Ps in all confidence factors

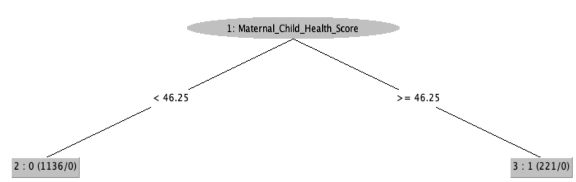

Figure 3. Random Forest classifier showing household compliance to 4Ps

CONCLUSION

The study concludes that J48 (in all confidence factors) and Random Forest algorithms have the best performance in terms of classifying compliant and non-compliant household beneficiaries on the maternal and child health and nutrition provisions of the *Pantawid Pamilyang Pilipino Program*. All confidence factors of J48 and Random Forest revealed the same classification accuracy of 100%, F-measure of 1.00, and kappa of 1.00. Furthermore, visualizing the trees of these classifiers also revealed possible threshold scores, which demarcates compliance levels of both compliant and non-compliant households.

RECOMMENDATIONS

Based on the high classification accuracy demonstrated by the J48 and trees.RandomForest across all confidence factors, it is advisable to employ these classifiers as principal instruments for evaluating and forecasting the adherence of household beneficiaries to the maternal and child health and nutrition provisions of the *Pantawid Pamilyang Pilipino* Program. These tools not only offered a high level of accuracy, but more importantly, they also provided valuable insights into compliance scores, rendering them highly suitable for practical application. The adoption of these algorithms as commonly accepted evaluative instruments will enhance the effective identification of households that conform to regulations and those that do not, thereby enabling more efficient allocation of resources and focused interventions.

The decision trees generated by the J48 and trees.RandomForest algorithms revealed distinct threshold values for compliance, specifically 42.5 and 46.5, respectively. As a strategic recommendation, it is advisable for program administrators to contemplate the adoption of a threshold based on a range, wherein the compliance score for households is classified by setting it at the average of two values, specifically 44.5. This approach, which takes into account the perspectives of both algorithms, aims to incorporate a moderate and carefully determined cut-off point. Moreover, it is essential to conduct regular monitoring and adjustment of this threshold. This can be achieved by leveraging the knowledge gained from these classifiers to effectively respond to any shifts in household behaviors or modifications in program prerequisites that may occur in the future.

REFERENCES

[1] Frufonga, R.F., "The Pantawid Pamilyang Pilipino program (4Ps) in Iloilo, Philippines: An evaluation," Asia Pacific Journal of Multidisciplinary Research, vol.3, 59-65, 2015.

[2] A. M. Hermeto and A. Junqueira Caetano, "Socioeconomic status, family structure and child outcomes in Brazil: health in the childhood," International Journal of Social Economics, vol. 36, no. 10, pp. 979–995, Aug. 2009

[3] Philippine Statistics Authority, "Proportion of Poor Filipinos was Recorded at 18.1 Percent in 2021," psa.gov.ph, Aug. 15, 2022

[4] R. Velarde and L. Fernandez, "Philippine Social Protection Note Welfare and Distributional Impacts of the Pantawid Pamilyang Pilipino Program." Available: https://documents1.worldbank.org/curated/en/249641468329342625/pdf/634180REVISED0Box361516B00PUBLIC00PSPN03.pdf

[5] "Data Updates | Pantawid Pamilya." https://pantawid.dswd.gov.ph/data-updates/ (accessed Feb. 27, 2023)

[6] A. Sandryhaila and J. M. F. Moura, "Big Data Analysis with Signal Processing on Graphs: Representation and processing of massive data sets with irregular structure," IEEE Signal Processing Magazine, vol. 31, no. 5, pp. 80–90, Sep. 2014

[7] M. S. Mahmud, J. Z. Huang, S. Salloum, T. Z. Emara, and K. Sadatdiynov, "A survey of data partitioning and sampling methods to support big data analysis," Big Data Mining and Analytics, vol. 3, no. 2, pp. 85–101, Jun. 2020

[8] B. J. Ruth and J. W. Marshall, "A History of Social Work in Public Health," American Journal of Public Health, vol. 107, no. S3, pp. S236–S242, Dec. 2017

[9] Longbing Cao, "Social Security and Social Welfare Data Mining: An Overview," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 42, no. 6, pp. 837–853, Nov. 2012

[10] R. Anderson and G. Mansingh, "Data Mining Approach to Decision Support in Social Welfare," International Journal of Business Intelligence Research, vol. 5, no. 2, pp. 39–61, Apr. 2014

[11] S. B. Eom, S. M. Lee, E. B. Kim, and C. Somarajan, "A survey of decision support system applications (1988–1994)," Journal of the Operational Research Society, vol. 49, no. 2, pp. 109–120, 1998

[12] S. Eom, "DSS, BI, and Data Analytics Research: Current State and Emerging Trends (2015–2019)," Decision Support Systems X: Cognitive Decision Support Systems and Technologies, pp. 167–179, 2020

[13] A. N. Kasiwi, A. Nurmandi, D. Mutiarin, and M. F. Azka, "Artificial Data Management in Reaching Conditional Cash Transfer of Program Keluarga Harapan (PKH) Utilizing Simple Addictive Weighting," IOP Conference Series: Earth and Environmental Science, vol. 717, no. 1, p. 012013, Mar. 2021

[14] R. Ramadiani, A. R. Rahmana, I. Islamiyah, M. D. Balfas, T. Rahman, and A. Yunianta, "Decision Support System of Direct Cash-Village Fund Recipients Using Multi Attribute Utility Theory," 2021 5th International Conference on Informatics and Computational Sciences (ICICoS), Nov. 2021

[15] T. Tarigan and N. Sidabukker, "Decision support system in determining recipients of direct government cash assistance for the village community of Patane IV Porsea using the Promethee method," Jurnal Manajemen, Informatika, Rekayasa Perangkat Lunak Dan Teknologi Komunikasi, vol 1, no.2, December, 2022.

[16] D. Daha & P. M. Hasugian, "Application of the SAW method in a decision support system recommendation for poor student assistance funds (BS<) at Budi Mulia Tumba Jae Private Junior High School," Distance: Journal of Data Science, Technology, and Computer Science, vol 1, no. 1, December 2021.

[17] L. Nulhakim, D. Andriansyah, & F. D. Putri, "Model fuzzy MADM in the decision support system of Mustahik service program beneficiaries at Laznas PPPA Daarul Quran Center Tangerang," in IAIC International Conference Series, vol. 3, no.1, 26-37, November 2020.

[18] H. Liu & H. Motoda, Feature selection for knowledge discovery and data mining. Vol. 454. Springer Science & Business Media, 2012.

[19] I. Rodríguez-Rodríguez, J.V. Rodríguez, D. J. Pardo-Quiles, P. Heras-González, & I. Chatzigiannakis. "Modeling and forecasting gender-based violence through machine learning techniques," Applied Sciences vol. 10, no. 22, p.8244, 2020.

[20] L.M. Zouhal & T. Denoeux, "An evidence-theoretic k-NN rule with parameter optimization," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 28, no. 2, pp.263-271, 1998.

[21] T. A. Alam, M. Mushtaq, K. Shaukat, I. Hameed, M. U. Sarwar, & S. Luo. "A novel method for performance measurement of public educational institutions using machine learning models," Applied Sciences 11, no. 19, p.9296, 2021.

[22] J. Alejandrino, J. J. Bolacoy, & J. V. Murcia. "Supervised and unsupervised data mining approaches in loan default prediction," International Journal of Electrical & Computer Engineering, vol. 13, no. 2, pp. 1837-1847, 2023.

[23] M. C. Gaya, & J. I. Giráldez, "Techniques for distributed theory synthesis in multiagent systems," in International Symposium on Distributed Computing and Artificial Intelligence 2008 (DCAI 2008), pp. 395-402. Springer Berlin Heidelberg, 2009.

ACKNOWLEDGMENT

We thank the Department of Science and Technology for funding this research project under the Good Governance through Data Science and Decision Support System (GODDESS) PCIERD Project No. 10625.

AUTHOR INFORMATION

John Vianne Murcia, college professor of Business Analytics, University of Mindanao, Davao City.

Marlon Suelto, senior lecturer of Environmental Science, University of Mindanao, Davao City.

Joane May Delima, associate professor of Information Technology, UM Digos College, Digos City.

John Raven Manulat, associate professor of Information Technology, UM Digos College, Digos City.

Conrado Jr. Panerio, research coordinator and associate professor of Mathematics Education, UM Digos College, Digos City.