An Ensemble-based Machine Learning Model for Investigating Children Interaction with Robots in Childhood Education

Doris Fuster-Guillén[1*], Oscar Gustavo Guadalupe Zevallos[2], Juan Sánchez Tarrillo[3], Silvia Josefina Aguinaga Vasquez[4], Miguel A. Saavedra-López[5] and Ronald M. Hernández[6]

[1*]Professor, Universidad Nacional Mayor de San Marcos, Lima, Perú.
dfusterg@unmsm.edu.pe, Orcid: https://orcid.org/0000-0002-7889-2243

[2]Professor, Universidad César Vallejo, Lima, Perú.
oguadalupe@unamad.edu.pe , Orcid: https://orcid.org/0000-0003-2946-1583

[3]Professor, Universidad Católica Santo Toribio de Mogrovejo, Chiclayo, Perú.
jsanchez832@gmail.com, Orcid: https://orcid.org/0000-0002-6763-760X

[4]Professor, Universidad Católica Santo Toribio de Mogrovejo, Chiclayo, Perú.
silviaguinaga25@gmail.com, Orcid: https://orcid.org/0000-0001-7810-4625

[5]Professor, Universidad Continental, Cusco, Perú.
saavedralopezmiguel@gmail.com, Orcid: https://orcid.org/0000-0003-4913-933X

[6]Professor, Universidad Privada Norbert Wiener, Lima, Perú.
ronald.hernandez@outlook.com.pe, Orcid: https://orcid.org/0000-0003-1263-2454

Received: December 10, 2022; Accepted: January 15, 2023; Published: March 30, 2023

Abstract

Investing in children's well-being and supporting high-quality pre-school education is a significant component of its promotion (ECE). All children have the right to participate. ECE teachers' thoughts about children's participation were examined to see if they were linked to children's perceptions of their participation. On the other hand, current studies focus on a single categorization method with lower overall accuracy. The findings of this study provided the basis for the development of an ensemble machine learning (ML) approach for measuring the participation of children with learning disabilities in educational situations that were specifically developed for them. Visual and auditory data are collected and analyzed to determine whether or not the youngster is engaged during the robot-child interaction in this manner. It is proposed that an ensemble ML technique (Enhanced Deep Neural Network (EDNN), Modified Extreme Gradient Boost Classifier, and Logistic Regression) be used to judge whether or not a youngster is actively engaged in the learning process. Children's participation in ECE courses depends on both the quantitative and qualitative characteristics of the classroom, according to this research.

Keywords: Artificial Intelligence, Childhood Education, Multimodal Data, Ensemble ML.

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 14, number: 1 (March), pp. 60-68. DOI: *10.58346/JOWUA.2023.I1.005*

*Corresponding author: Professor, Universidad Nacional Mayor de San Marcos, Lima, Perú.

1 Introduction

Educational technology includes computer-based learning. Moreover, educational technology combines resources other than computers to maximise each resource's unique traits and benefits (Ninaus, M., et al., 2019), even in childhood education. Besides computers, interactive whiteboards and programmable toys are commonly employed in childhood education. Childhood education may include game consoles and robots (Morgan, P. L., et al., 2012). Artificial Intelligence techniques have been used in computer-based learning to improve learning outcomes. Traditional CAI systems rely on shallow representations of the teaching domain, student data, and pedagogical practices. They struggle to successfully adapt to the learning process due to inadequate adaptability and learner evaluation. Intelligent Educational Systems (IESs) use AI techniques and mechanisms (Ingersoll, B., et al., 2006). The purpose is to emulate learners and knowledge about the teaching subject (Cuayáhuitl, H, 2019).

Like Intelligent Educational Systems (computer-based), Intelligent Robots are a standard paradigm of Artificial Intelligence in education (McConnell, S. R. (2000)). Educational robots have advantages over computer-based learning methods. Educational robots are self-contained, mobile, and diverse. They may feel emotions and react to human relationships. Robots provide unique engagement opportunities that foster bonding with young children (Sallin, A. (2021)). Some studies demonstrate that young children treat robots more like friends than machinery or toys. The new study intends to add to past research by measuring children's engagement with the social robot NAO (Vartiainen, H., et al., 2020). The experiments involved ten youngsters, each of whom was given a scenario by a child psychologist. Section 2 describes the recommended technique. Section 3 discusses the findings. Section 4 concludes and plans future work.

2 Proposed Methodology

This study presented an ensemble ML-based approach to measuring the involvement of children with learning disabilities in properly prepared instructional scenarios. The suggested methodology's actual procedure is (1) the use of multimodal information comprised of visual and aural modalities, and (2) the use of an ensemble ML technique (EDNN, Modified XG Boost, and Logistic Regression) that offers a conclusion regarding the child's involvement state.

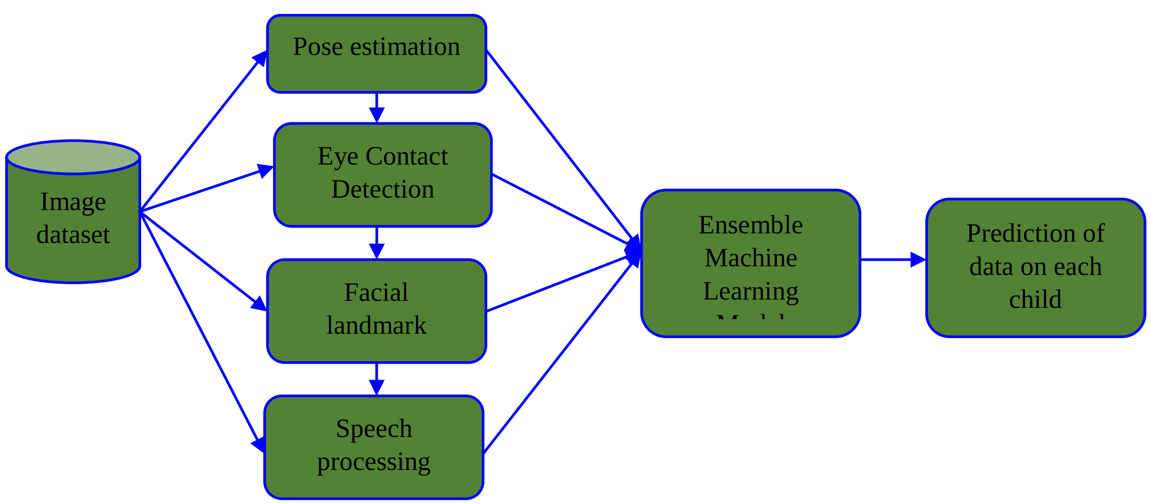

Figure 1: The Process of the Children Engagement with Robot using Ensemble ML

Dataset Formulation

In ten scenarios, the studies involved ten youngsters, two girls and eight boys, ages 9 to 10 (Table 1). This is done in a classroom with children, the NAO robot, and a child psychologist at the back of robot. The robot must also face the child for the speech recognition module to perform correctly, as the microphones are aimed at the youngster. The video recordings were used to create an 819-sample dataset with 11 characteristics per sample, in which 99 samples represented engaged children, while 720 samples represented non-engaged. After manual annotation, competent child psychologists determined the ground truth data.

Table 1: Educational Scenarios' Attributes

Scenario	Types of Activities Included
S1	Meet/greet, text decoding, phonology (de)composition, memory, and robot-child relaxation game
S2	Meet/greet, phonetic discrimination, text reading, decoding, and comprehension
S3	Meet/greet, story listening and telling, and sentence structuring
S4	Text comprehension and visual representation
S5	Phonemic addition, sentence playback from memory, and robot-child relaxation game
S6	Meet/greet, sentence playback from memory, and reading enhancement
S7	Meet/greet, phonetic awareness, and robot-child relaxation game
S8	Meet/greet, acoustic vocal discrimination, and acoustic syllable discrimination
S9	Memory enhancement and text decoding
S10	Text reading and robot-child relaxation game

Multimodal Sensing

The social robot's sensing capabilities determine the sort of sensory data processed to determine the child's involvement state during the contact. This study uses the well-known NAO robot, but other social robots might also be employed. The robot has two identical RGB video cameras on the forehead and a microphone, allowing it to offer visual and auditory sensing.

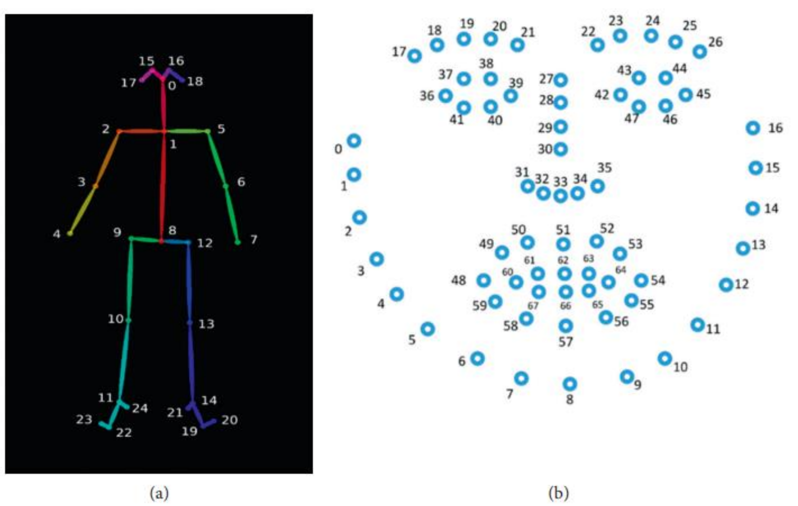

Figure 2: Visual Sensing Retrieved Data. (a) Body Pose Identification. (b) Facial Landmark Identification

1. **Visual Sensing:** These skills of NAO robots allow it to capture video frames of the child's body and face. The library extracts the body stance from each video frame, comprising of 25 key points (2 on torso, six on hands, 12 on legs, and five on head), as shown in Fig.2 (a). Using the Open Face library, 68 facial landmarks are retrieved from the child's face (refer to Figure 2(b)). Facial Action Coding System (FACS) uses evaluated facial landmarks to describe a child's emotional state. Finally, the Open Gaze library detects the child-robot eye contact.
2. **Audio Sensing:** At the time of communication with the child, the robot must always face the child to capture and evaluate the child's speech and provide other information about the child's involvement level.

Most of the visual elements listed above are estimated by monitoring and evaluating retrieved vital points. For features 3 and 7, the most frequent condition (0 or 1) in the time window is chosen. If the child is not speaking, verify if the voice volume is 350 RMS or higher, and the mean voice level. A feature vector FV R11 is allocated to each 60-second video frame. The collected characteristics from instructional situations are utilized for training the ML technique to predict the child's involvement.

Ensemble-based ML Approaches

An ensemble ML classifier is used to solve a standard two-class classification issue. Ensemble algorithms have established records on complicated datasets (Jonnalagadda, A, 2022). Three ML-based classifiers are utilised here to classify the data. The EDNN, MXGB, and Logistic regression models are merged to produce a prediction technique (ensemble approach). Finally, display the voting-based classification result.

EDNN

Deep Learning has been shown to be useful in creating correct predictions from complex data sources. Because these parameters are chosen based on the users' experiences and/or trial and error, an overfitting problem is introduced in the process, which is a big difficulty. A computational technique can be used to automate and minimize errors in this procedure. It used fuzzy inference to reduce the problem. Fuzzy-inference systems can employ responsive neurons to categorize desirable aspects.

Integration with Fuzzy Inference System

Complex non-linear challenges can be simulated using fuzzy inference techniques. Rule-based systems offer the advantage of being influenced by subjective data. An analyst can use this to improve categorization results or change the system's behaviour. This feedback bias can also be employed to accelerate autonomous system learning while retaining stability. The DL features will be further processed by FIS, allowing the system to simulate human thinking and providing a way for biassing the system with analyst feedback. To train the system, both feature vector inputs and outputs must be valid. The benefit of this approach is that these fuzzy rules describe the fuzzy system's behaviour. The learning algorithm block analyses the fuzzy inference block's input pattern. Adjusted system weights are returned to the system, automatically changing prediction behavior and allowing the system to adapt.

Assume that the rules include three fuzzy *if-then* rules of Takagi and Sugeno's kind.

Rule 1: If x is A_1, y is B_1, and z is C_1 then $f_1=p_1x+q_1y+t_1z+ r_1$,

Rule 2: If x is A_2, y is B_2 and z is C_2 then $f_2=p_2x+q_2y+t_2z+ r_2$,

Rule 3: If x is A_3, y is B_3 and z is C_3 then $f_3=p_3x+q_3y+t_3z+ r_3$,

Modified Extreme Gradient Boost (MXGB) Classifier

Extreme Gradient Boosting (XG Boost) is a ML algorithm for regression and classification issues based on the Gradient Boosting Decision Tree (GBDT). Firstly, define the second-order approximation of additive tree boosting. The amount of data is m, and the amount of features is n. The raw forecast is zi, and the possibility recognition is (y I)̂= σ (z i), in which σ(.) represents the sigmoid function. Remember that the (y i)̂ in their analysis is denoted as z here. The accurate label is characterized by yi, and the parameters for the two-loss functions are α and γ. The gradients/Hessians are documented in a merged format independent of yi, to assist vectorization in other related applications.

According to the additive learning objective used in practice is:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, z_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \qquad (1)$$

In which t denotes the t-th iteration of the training process. Employing second-order Taylor expansion on equation 1, will obtain:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n}[l\left(y_i, z_i^{(t-1)} + g_i f_t(x_i)\right) + \frac{1}{2}h_i(f_t(x_i))^2 + \Omega(f_t) \qquad (2)$$

$$\propto \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i(f_t(x_i))^2] + \Omega(f_t) \qquad (3)$$

Because XG Boost does not automatically differentiate, manual differentiation is required. The resulting expressions can also be used in other ML projects. Both loss functions employ sigmoid activation, and the derivatives use the following fundamental characteristic of the sigmoid:

$$\frac{\partial \hat{y}}{\partial z} = \frac{\partial \sigma(z)}{\partial z} \qquad (4)$$

$$= \sigma(z)(1-\sigma(z)) \qquad (5)$$

$$= \hat{y}\,(1-\hat{y}\,) \qquad (6)$$

Distraction Population Factor (DPF)

Apart from the regularised objective, these strategies help prevent overfitting. After each phase of tree boosting, the Distraction population factor is raised by a factor 〖RAw〗 j^*.

For a fixed structure q(x), calculate the best weight RAw_j^* of leaf j by

$$RAw_j^* = -\frac{\sum_{i\in I_j} g_i}{\sum_{i\in I_j} h_i + \lambda} \qquad (7)$$

In which ∈ is an approximation factor. Here each data point is weighted by h_i which denotes the weight, can rewrite Eq (8) as

$$\sum_{i=1}^{n}\frac{1}{2}h_i\left(f_t(x_i) - g_i/h_i\right)^2 + \Omega(f_t) \quad (8)$$

Describe $I_j = \{i|q(x_i) = j\}$ as the instance set of leaf j. And it can rewrite as Eq (10) by expanding Ω,

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i f_t^{\,2}(x_i)] + \Omega(f_t) \qquad (9)$$

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i f_t^{\,2}(x_i)] + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^{\,2} \qquad (10)$$

Logistic Regression (LR)

The logistic regression design is chosen because it is widely used in data mining and ML domains. For example, this study evaluated risk factors for provided data and predicted the possibility of happenings.

When dealing with two-category concerns (i.e., just two output types representing each category), logistic regression is commonly used to calculate the chance of each classification event occurring. Below is a logistic regression framework:

$$prob(Y = 1) = \frac{e^z}{1+e^z} \qquad (11)$$

In which Y denotes to binary dependent variable (Y is equal to 1 if an event happens; Y=0 otherwise), e stands for the foundation of natural logarithms, and Z means with constant β0, co-efficient βj and predictors Xj, for p predictors (j=1,2,3,.....,p).

3 Results and Discussion

Experiments were set up to evaluate the proposed methodology's performance. The studies used the scikit-learn ML Library for Python 2.7. The tests were done on a system with an Intel i7-6700HQ CPU, 8 GB DDR4 RAM, and a GTX 960M GPU.

Table 2: Performance Comparison Table for Suggested and Conventional Approaches

Metrics	RF	SVM	EMLM
Precision	79.24	81.57	87.67
Recall	72.15	75.68	89.27
F-measure	86.47	89.62	92.59
Accuracy	68.54	81.57	94.21

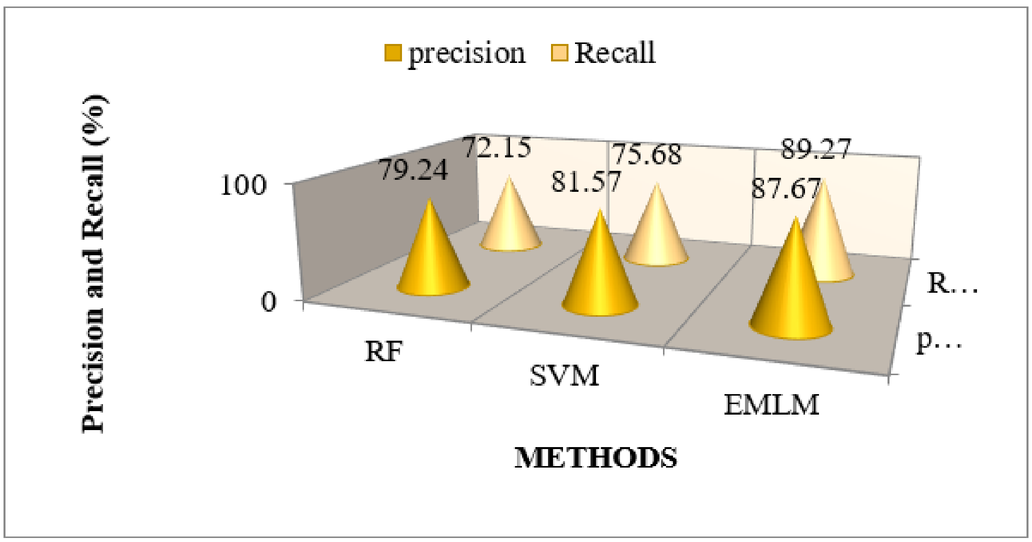

Figure 2: Precision and Recall Results between the Suggested and Conventional Approaches

This method outperforms the present classifier in terms of precision and recall, as shown in Figure 2. The findings show that the proposed algorithm works quite well. So the presented method will outperform other classifiers developed on earlier versions.

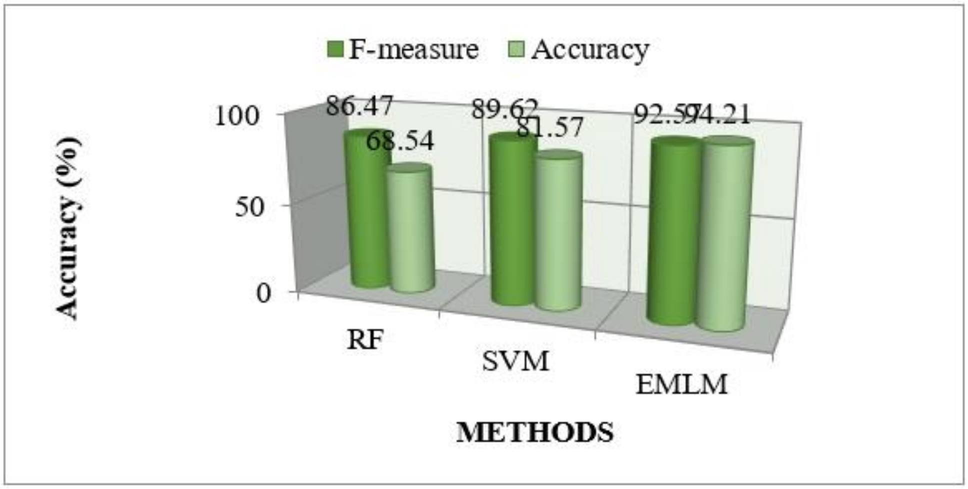

Figure 3: Accuracy and F-measure Results between the Suggested and Conventional Approaches

Figure 3 compares the experimental, and EMLM learning projected results using RF and SVM approaches. The results show that ensemble learning may considerably increase prediction performance in aamulti-label formulation prediction. so, the suggested algorithm is effective than the existing one.

4 Conclusion

This paper tackled detecting a child's engagement with a social robot in two-way intelligent communication. Using an ensemble ML technique, it was solved as a two-class categorization issue. To explain the child's behavior, the suggested methodology uses multimodal data (visual and auditory). The basic concept that a child with learning challenges is diagnosed by treating body and head positions, facial expressions, eye contact, and speech was approved. Thus, the suggested approach outperforms the conventional strategy in terms of accuracy.

References

[1] Boutaba, R., Salahuddin, M.A., Limam, N., Ayoubi, S., Shahriar, N., Estrada-Solano, F., & Caicedo, O.M. (2018). A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. *Journal of Internet Services and Applications*, *9*(1), 1-99.

[2] Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., & Cho, H. (2015). Xg boost: extreme gradient boosting. *R package version*, *1*(4), 1-4.

[3] Cuayáhuitl, H., Lee, D., Ryu, S., Cho, Y., Choi, S., Indurthi, S., & Kim, J. (2019). Ensemble-based deep reinforcement learning for chatbots. *Neurocomputing*, *366*, 118-130.

[4] Field, A. (2009). Logistic regression. *Discovering statistics using SPSS*, *264*, 315.

[5] Ingersoll, B., & Dvortcsak, A. (2006). Including parent training in the early childhood special education curriculum for children with autism spectrum disorders. *Topics in Early Childhood Special Education*, *26*(3), 179-187.

[6] Jonnalagadda, A., Rajvir, M., Singh, S., Chandramouliswaran, S., George, J., & Kamalov, F. (2022). An Ensemble-Based Machine Learning Model for Emotion and Mental Health Detection. *Journal of Information & Knowledge Management*.

[7] Kim, H. (2020). 5G core network security issues and attack classification from network protocol perspective. *Journal of Internet Services and Information Security*, *10*(2), 1-15.

[8] Larochelle, H., Bengio, Y., Louradour, J., & Lamblin, P. (2009). Exploring strategies for training deep neural networks. *Journal of machine learning research*, *10*(1), 1-40.
[9] McConnell, S.R. (2000). Assessment in early intervention and early childhood special education: Building on the past to project into our future. *Topics in Early Childhood Special Education*, *20*(1), 43-48.
[10] Morgan, P.L., Farkas, G., Hillemeier, M.M., & Maczuga, S. (2012). Are minority children disproportionately represented in early intervention and early childhood special education? *Educational Researcher*, *41*(9), 339-351.
[11] Ninaus, M., Greipl, S., Kiili, K., Lindstedt, A., Huber, S., Klein, E., & Moeller, K. (2019). Increased emotional engagement in game-based learning–A machine learning approach on facial emotion detection data. *Computers & Education*, *142*.
[12] Sallin, A. (2021). Estimating returns to special education: combining machine learning and text analysis to address confounding.
[13] Vartiainen, H., Tedre, M., & Valtonen, T. (2020). Learning machine learning with very young children: Who is teaching whom? *International journal of child-computer interaction*, *25*.

Authors Biography

Doris Fuster-Guillén

Researcher level III, recognized by CONCYTEC, PhD in education sciences, master in teaching and university research. He has published scientific articles in high impact indexed journals, author of books and chapters of research texts in national and international publishers. Specialist in education, social sciences and research skills oriented to quantitative, qualitative and mixed approaches.

Oscar Gustavo Guadalupe Zevallos

Bachelor in SOCIOLOGY, registered in the College of Sociologists of Peru with the No. 3689. He has a Master's degree in "PUBLIC MANAGEMENT" and has extensive knowledge of the national reality and the region of Madre de Dios. In the last few years, his professional actions as a university teacher, prevention of human trafficking and the environmental disaster that Madre de Dios is experiencing stand out, addressing them with different approaches that have positioned him as an international reference on the aforementioned topics.

Juan Sánchez Tarrillo

Bachelor's degree in education, specializing in Historical and Social Sciences and Philosophy. Doctor in education, Master in education sciences in the mention of research and teaching. Currently teaching in the faculty of humanities and academic coordinator of the graduate school of the Catholic University Santo Toribio de Mogrovejo (USAT)

Silvia Josefina Aguinaga Vasquez
Bachelor's degree in education, specializing in Historical and Social Sciences and Philosophy. Doctor in education, Master in educational sciences in the mention of research and teaching. Currently teaching in the faculty of humanities at the Catholic University Santo Toribio de Mogrovejo (USAT),

Miguel A. Saavedra-López
Doctor in Psychology from Universidad César Vallejo, Perú, 2023. Dean of the Regional Board of Directors of Tumbes - College of Psychologists of Peru (2020 - 2022). Qualified as a RENACYT researcher by the National Council of Science, Technology and Technological Innovation (CONCYTEC). Currently teaching at the National University of Tumbes and Continental University (Peru), director of the International Journal of Social Sciences. Interest in research related to mental health, artificial intelligence and scientific production.

Ronald M. Hernandez
Bachelor's Degree in Psychology. Master in Education. Qualified researcher in RENACYT (Concytec). University professor. International assistant editor of the Educational Journal of Works Oriented to the XXI Century (Panama), Varela Journal (Cuba). Peer reviewer of scientific articles in national and international journals on topics of Psychology, Education and Medicine. Couator of the book Publishing in scientific journals. Has published scientific articles in international indexed journals in Scopus, Web of Science and SciELO