Design and Development of Machine Learning and Evolutionary Computation Methods for Risk Factors Identification in Early Childhood Disability

Dr. I Wayan Suryasa and I Wayan Astu Werdistira

Abstract--- *Timely identification of social and emotional disorders is crucial for the immediate welfare and future well-being of young children. The present study encounters challenges in identifying and establishing risk factors associated with early childhood impairment. Consequently, the overall system performance is substantially reduced. In order to tackle the aforementioned challenges, this study proposes the utilisation of the Cuckoo Search Optimisation with Adaptive Network-based Fuzzy Inference System (CSO+ANFIS) technique. The aim is to effectively generate and identify risk factors associated with early childhood disability. This study employs the CSO method to select the most important attributes and determines the optimal objective function based on the highest fitness values. The ANFIS technique focuses on identifying important risk variables by analysing the hidden layer and fuzzy inference values. The experimental results have shown that the suggested CSO+ANFIS technique surpasses the current paradigm in terms of accuracy and sensitivity metrics.*

Keywords--- *Early Childhood Disability, Risk Factors, CSO+ANFIS.*

I. Introduction

EARLY childhood refers to the time span encompassing prenatal development up until the age of eight. This stage is crucial for advancement and growth, as experiences in early childhood can have long-lasting effects on an individual's life outcomes. During early childhood, children have a crucial opportunity to establish the foundation for lifelong learning and engagement, while also reducing the risk of developmental delays and disorders (Justice, Laura M., et al., 2019).

Early recognition and treatment for special needs children are strongly advised. The importance of the years between birth and five have been acknowledged by psychologists, doctors, scientists, and educators. These early years are essential if there is any possibility of impairment (Bertoncelli, Carlo M., et al., 2019). Inadequate legislation and policies, negative attitudes, limited resources, and inconvenient locations are all obstacles for impaired children and their families. Without timely and efficient early intervention, care, and security for children with intellectual disabilities or impairments, their problems might deteriorate, resulting in long-term consequences, increased poverty, and considerable marginalization.

Some children are born with crippling health conditions, while others become disabled due to disease, accident, or malnutrition. Autism, Down syndrome, traumatic spinal cord damage, muscular dystrophy, spina bifida, cerebral palsy. Some children have a single disability, while others have many disabilities. A kid with cerebral palsy may have mobility, communication, and cognitive issues (Donat, U., et al., 2020). Disabilities are experienced differently by each child due to the complex combination of health, environmental, and personal factors (Michelle, G., et al., 2019).

This study's major goal is to identify risk factors for early childhood impairment. Despite countless studies and approaches, performance is not guaranteed. Existing methods suffer from computational complexity and accuracy issues. The CSO+ANFIS algorithm is suggested in this study to accurately identify risk factors in early childhood impairment and improve overall system performance. This study's essential contribution is identifying risk variables for LD kids. The suggested scheme improves results by applying efficient strategies.

The balance of the work is presented: Section 2 describes the recommended methodology for identifying risk factors in early childhood impairment. Section 3 contains a description of the experimental data and performance analysis. Finally, Section 4 briefs the outcomes.

II. Materials and Methods

A new algorithm called CSO+ANFIS is presented in this study to develop and identify early childhood impairment risk variables. Since these models of child development have been universally acknowledged, early detection programmes for children from low-income families are now based on a philosophy of change that stresses educational enrichment for children as well as family education and support services.

Children and family traits are the most strongly connected with clinical detection of the language issue in preschoolers when using the presented approach. A sample of 483 3- to 5-year-old children was analyzed using machine learning to uncover factors that best categorized those receiving therapy for a language issue (54% affected).

In this part, the CSO method is used to choose critical features from the inputs. Cuckoo search is a new natural metaheuristic technique for finding optimal solutions. Cuckoos

Dr. I Wayan Suryasa, SS., M. Hum., Assistant Professor, Institut Teknologi dan Bisnis STIKOM Bali, Denpasar, Indonesia. E-mail: suryasa@stikom-bali.ac.id
I Wayan Astu Werdistira, Institut Teknologi dan Bisnis STIKOM Bali, Denpasar, Indonesia. E-mail: wayan_astu@stikom-bali.ac.id

DOI: 10.9756/BIJMMI/V14I1/BIJ24005

are attractive birds due to their wonderful melody and aggressive breeding style. Ani and Guira cuckoos, lay their eggs in communal nests and may remove other eggs to boost their hatching chances. Unknown eggs are thrown away by the host bird, or the nest is abandoned and a new one established elsewhere (Sadegh, S., et al., 2017). Instead of isotropic random walks, this CSO technique uses Levy flights.

The three theoretical principles for characterizing typical cuckoo search are:

- Every cuckoo lays a single egg and places it in a nest selected randomly.
- Only the most superior nests, containing eggs of the utmost quality, will be inherited by subsequent generations.
- The quantity of accessible host nests remains constant, and the host detects the cuckoo's egg with a high probability of 0.1. The host bird has the option to either remove the egg from the nest or leave the nest and build a new one.

The final criterion is simulated by a fraction pa of the n host nests restored with fresh nests (with new random solutions). CSO is easy to use and provides ample search space. It employs the levy flight to global search to prospect the search space more effectively.

The efficacy of CSO can be enhanced by utilising multiple host nests containing a substantial number of eggs (Douglas, R., et al. 2013). Typically, cuckoos select one of three nest types in which to deposit their eggs. The typical cuckoo selects a group of host nests that have eggs that are identical in appearance to its own. Other cuckoos select a group of host nests that contain eggs that are distinct from their own. Some cuckoo species lay camouflaged eggs that have a dark coloration, in contrast to the light-colored eggs of the birds they parasitize.

Initial Population

Each egg presents a distinct set of risk factors that are selected and utilised to accurately classify the inputs. The features are selected from the top-m ranked aspects using statistical measures applied to the provided input.

Finding New Solutions and Levy Flight

The ECS-based feature selection approach utilises levy flight to explore new solutions from Equation (2). Exploring the most optimal solution obtained thus far through a brief walk should result in the discovery of additional solutions, thereby accelerating the local search process. By utilising Levy flight, a new solution $x_i^{(t+1)}$ is generated for cuckoo i, and is described as follows,

$$x_i^{(t+1)} = x_i^{(t)} + C \oplus Levy\ (s, \lambda) \qquad (1)$$

t is the step size.

$$Levy\ (S, \lambda) \sim s^{-\lambda}, \quad 1 < \lambda \leq 3 \qquad (2)$$

Crossover and Mutation

- When a regular cuckoo bird reproduces, it uses a technique called crossover to lay two eggs and then selects the best one.
- For the European cuckoo, it follows a process of laying two eggs using a crossover mechanism that incorporates a balanced mutation operator. The cuckoo then selects the most optimal egg.
- In an alternative approach, eggs are generated through the utilisation of a random solution, which is deliberately obscure or mysterious in nature.

Fitness Function

It is quite crucial in the selection process. The notable subset characteristics from the training dataset are efficiently selected by applying the optimal fitness function values. As a result, the fitness function incorporates both relevance and redundancy to lead CSO in its search for the optimal feature subset.

$$fitness\ (fi) = \alpha \times D - (1 - \alpha) \times R \qquad (3)$$

Where $D = \frac{1}{|S|} \sum_i I(xi, C)$ (4)

$$R = \frac{1}{|S|^2} \sum_i \sum_j I(xi, xj) \qquad (5)$$

In which X denotes the cluster of picked features and C is the class label. Each characteristic and class label are random variables. D computes labels by comparing the picked feature subset to the class labels. R assesses the mutual information, indicating the feature subset's redundancy. Fitness is a function that maximizes relevance D while minimizing redundancy R in a feature subset. The relevance is thought to be more critical than the redundancy, it is set to be larger (1-) than in the fitness function. Using CSO efficiently selects the key and crucial subset features from the inputs.

Algorithm 1: CSO

Create an initial population of n host nests with m eggs (features).

while (t<MaxGeneration) or (stop criterion)

for every nest

Acquire a cuckoo type randomly (say i)

Identify the kind of the cuckoo

If cuckoo_type=common_cuckoo

Build two eggs via crossover with two best eggs in the nest (solution) and pick the optimal one among them

Else if cuckoo_type = European cuckoo

Form two eggs via crossover with balanced mutation operator with any two eggs in the nest

Pick the optimal one among them with greatest pressure and temperature

Else

Generate egg with random solution (cryptic egg)

End if

Compute its fitness fi using (3) for solubility data

Pick an egg with the worst solution in the nest (say j)

If (fi>fj)

Restore j by new solution i

End if

Rate the egg as per the solution

Determine the optimal solution among m eggs in the nest

Desert a fraction of the eggs in the nest containing the worst solutions and form new ones using levy flight (1) and (2)

Preserve the optimal solutions (risk factors)

End for

Rate the eggs in all nests with fitness value and identify the present best

End while

Return risk factors for correct determination in early childhood disability

The method selects the solubility data features based on the goal function's best fitness values. With the fresh optimal solutions, superfluous characteristics are greatly eliminated. Thus, the CSO technique is utilized to improve the inputs' valuable properties.

Risk Factors Identification Using ANFIS Algorithm for Early Childhood Disability

ANFIS integrates a neural network and a fuzzy inference system. The fuzzy logic model considers the system's imprecision and ambiguity, while the neural network provides it adaptability (José de Jesús, R., et al. 2019). Using this hybrid method, an initial fuzzy model with its input variables is created using rules taken from the system's input-output data. The neural network is then utilized to fine-tune the basic fuzzy model's administrations to develop the final ANFIS model. It ensures database versatility, rapid convergence, and excellent accuracy. The ANFIS algorithm's innovation improves early childhood impairment risk factor determination.

III. Experimental Result

In this work, we used the CSO+ANFIS approach to assess the identification of risk variables in early childhood impairment. The CSO+ANFIS technique is utilized for the input dataset to pick significant characteristics and develop an optimal prediction model utilizing the ANFIS model with the CSO procedure. The accuracy and sensitivity measures are tested using the conventional Support Vector Machine (SVM) and the novel CSO+ANFIS technique.

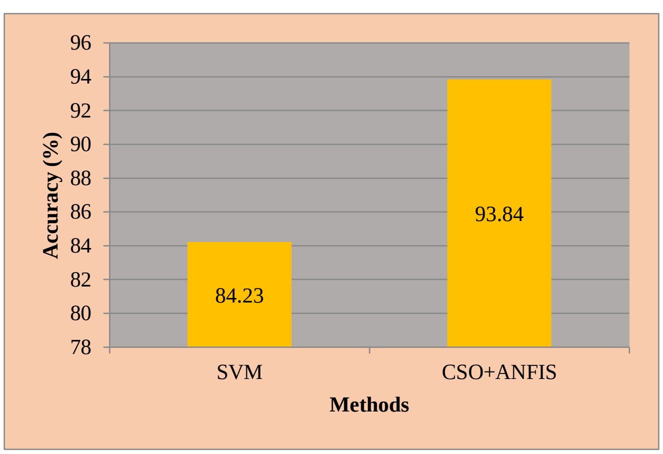

Fig. 1: Accuracy

According to the preceding Fig. 1, the comparison metric is assessed using known methodologies. The techniques are chosen for the x-axis, and the accuracy value is plotted on the y-axis. The recommended CSO+ANFIS approach is more accurate, whereas the existing SVM algorithm is less accurate. As a result, the suggested CSO+ANFIS strategy greatly improves identification accuracy and risk factors in early childhood impairment.

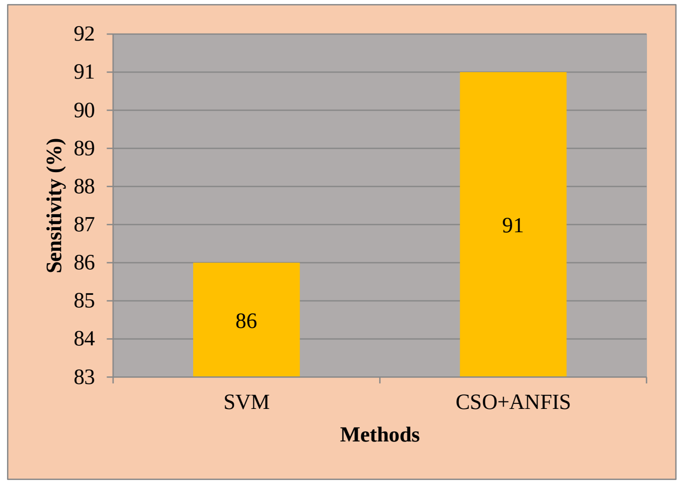

Fig. 2: Sensitivity

According to Fig. 2, the comparison metric is evaluated in terms of sensitivity using known methods. The methods are represented on the x-axis, while the sensitivity value is plotted on the y-axis. The proposed CSO+ANFIS approach is more sensitive, whereas the existing SVM algorithm is less sensitive. Consequently, the suggested CSO+ANFIS technique dramatically enhances the recognition performance in discovering risk variables in early childhood impairment.

IV. Conclusion

The CSO+ANFIS method is used to model early childhood impairment problems accurately. This work suggests using CSO+ANFIS to detect and develop early childhood impairment risk factors effectively. This study's findings help design early childhood disability therapies. They are predicting risk variables utilising the suggested CSO+ANFIS approach. The told CSO+ANFIS algorithm assesses and predicts risk variables in children with suspected learning difficulties. The outcomes indicate that the suggested risk factors procedure outperforms existing techniques regarding accuracy and sensitivity metrics.

References

[1] L.M. Justice, W.Y. Ahn and J.A. Logan, "Identifying children with clinical language disorder: an application of machine-learning classification", *Journal of Learning Disabilities,* Vol. 52, No. 5, Pp. 351-365, 2019.

[2] C.M. Bertoncelli, P. Altamura, E.R. Vieira, D. Bertoncelli, S. Thummler and F. Solla, "Identifying factors associated with severe intellectual disabilities in teenagers with cerebral palsy using a predictive learning model", *Journal of Child Neurology*, Vol. 34, No. 4, Pp. 221-229, 2019.

[3] D. Uwayezu, A. Gatarayiha and M. Nzayirambaho, "Prevalence of dental caries and associated risk factors in children living with disabilities in Rwanda: a cross-sectional study", *Pan African Medical Journal*, Vol. 36, No. 1, Pp. 1-10, 2020.

[4] M. Gonzalez and T.B. Brown, “Early Childhood Educators' Perceptions of Dyslexia and Ability to Identify Students At-Risk”, *Journal of education and learning*, Vol. 8, No. 3, Pp. 1-12, 2019.

[5] S. Salesi and G. Cosma, “A novel extended binary cuckoo search algorithm for feature selection”, *In 2nd international conference on knowledge engineering and applications (ICKEA)*, Pp. 6-12, 2017.

[6] D. Rodrigues, L.A. Pereira, T.N.S. Almeida, J.P. Papa, A.N. Souza, C.C. Ramos and X.S. Yang, “BCS: A binary cuckoo search algorithm for feature selection”, *In IEEE International symposium on circuits and systems (ISCAS)*, Pp. 465-468, 2013.

[7] J.D.J. Rubio, D.R. Cruz, I. Elias, G. Ochoa, R. Balcazar and A. Aguilar, “ANFIS system for classification of brain signals”, *Journal of Intelligent & Fuzzy Systems*, Vol. 37, No. 3, Pp. 4033-4041, 2019.