DIGITAL
HEALTH

# Multilayer perceptron-based self-care early prediction of children with disabilities

Rahman Ali[1], Jamil Hussain[2] and Seung Won Lee[3] (iD)

## Abstract

Early identification of children with self-care impairments is one of the key challenges professional therapists face due to the complex and time-consuming detection process using relevant self-care activities. Due to the complex nature of the problem, machine-learning methods have been widely applied in this area. In this study, a feed-forward artificial neural network (ANN)-based self-care prediction methodology, called multilayer perceptron (MLP)-progressive, has been proposed. The proposed methodology integrates unsupervised instance-based resampling and randomizing preprocessing techniques to MLP for improved early detection of self-care disabilities in children. Preprocessing of the dataset affects the MLP performance; hence, randomization and resampling of the dataset improves the performance of the MLP model. To confirm the usefulness of MLP-progressive, three experiments were conducted, including validating MLP-progressive methodology over multi-class and binary-class datasets, impact analysis of the proposed preprocessing filters on the model performance, and comparing the MLP-progressive results with state-of-the-art studies. The evaluation metrics accuracy, precision, recall, F-measure, TP rate, FP rate, and ROC were used to measure performance of the proposed disability detection model. The proposed MLP-progressive model outperforms existing methods and attains a classification accuracy of 97.14% and 98.57% on multi-class and binary-class datasets, respectively. Additionally, when evaluated on the multi-class dataset, significant improvements in accuracies ranging from 90.00% to 97.14% were observed when compared to state-of-the-art methods.

## Introduction

Children with disabilities face difficulties in carrying out their daily routine activities. Self-care activities are the most important of these and should be observed by the parents. It is challenging for parents and occupational therapists to detect such behaviors due to the lack of appropriate knowledge and availability of professional experts. This complicates self-care detection and requires adequate automated handling.[1] Automated decision-making technologies and software can make this detection process more accessible and help therapists suggest personalized treatment for each child.[2] WHO has developed a framework known as ICF-CY, which stands for "international classification of functioning, disability, and health for children and youth." Researchers in the subject domain consult the ICF-CY framework to predict self-care issues and declare disabilities.[3]

Several machine learning algorithms, such as artificial neural network (ANN),[4] instance-based learning (IBK, i.e. k-nearest neighbor (KNN)),[5] probabilistic naïve Bayes (NB) model,[6] an optimized distributed gradient boosting library named extreme gradient boosting (XGBoost), deep neural networks (DNNs),[7] and Care2Vec: a hybrid of

[1]Quaid-e-Azam College of Commerce, University of Peshawar, Khyber Pakhtunkhwa, Pakistan
[2]Department of Data Science, Sejong University, Seoul, Korea
[3]Sungkyunkwan University School of Medicine, Suwon, Korea

**Corresponding authors:**
Seung Won Lee, Sungkyunkwan University School of Medicine, Suwon 16419, Korea.
Email: LSW2920@gmail.com

Jamil Hussain, Department of Data Science, Sejong University, Seoul 30019, Korea.
Email: jamil@sejong.ac.kr

autoencoders and DNNs have been applied to address the same problem. These models help therapists detect disabilities in children by observing their self-care activities.[8]

Similarly, ANN models, such as fuzzy artificial neural networks (FNN)[9] and their variants have been successfully used in healthcare and well-being applications. A feed-forward ANN model, named multilayer perceptron (MLP), is intensively used for different healthcare classifications and prediction services. Several models include the use of a MLP neural network model for the categorization of teenage hypertension,[10] prediction of subjective health symptoms in individuals residing near mobile phone base stations,[11] and the prediction of health risk.[12]

In machine learning, data preprocessing and filtering methods improve the predictive accuracy of models; hence, they are used properly. The data resampling methods estimate the model classification accuracy and precision without the addition of new data samples. It is observed that resampling the training data decreases the standard deviation of the training and prediction up to a maximum of 50%.[13] When data resampling methods are combined with neural networks, it increases the prediction accuracy by a significant value.[14] Computer simulations demonstrate this increase in estimate.[15]

In the healthcare setting, resampling has been shown to be an effective technique for enhancing the performance of a predictive model of end-stage renal disease using unbalanced data. Likewise, the resampling technique Synthetic Minority Over-Sampling Technique-Edited Nearest Neighbor (SMOTE-ENN) is used as a pre-processing step for logistic regression analysis.[16]

In machine learning, randomization randomly shuffles the order of instances in a dataset, resulting in the development of distinct models from a single initial training set. Randomization enables machine learning algorithms to explore more of the search space and consider diversity for precise modeling.

To the best of our knowledge, the integration of unsupervised filtering methods, such as data resampling and randomization with the feed-forward ANN (i.e. MLP) has been used for the first time to improve the predictive accuracy of the disability prediction models proposed in the literature so far.

In this study, an efficient disability prediction model is proposed, which combines unsupervised data resampling and unsupervised randomization pre-processing methods with MLP to produce an MLP-progressive model for the prediction of self-care disability in children. The prediction results of the MLP-progressive model are expected to assist the therapist in recommending appropriate treatment/therapy to the parents of the children. To realize the effectiveness of the proposed MLP-progressive model, the publicly available dataset SCADI (self-care activities based on ICF-CY) is adopted for simulation.

The objective set for this study is to propose a new methodology for early detection of self-care disabilities in children using a feed-forward ANN based on the MLP algorithm. The study aims to improve the early identification of children with self-care impairments, which is a key challenge for professional therapists due to the complex and time-consuming detection process using relevant self-care activities.

This study hypothesizes that the proposed methodology, MLP-progressive, will improve the performance of MLP models in predicting self-care disabilities in children. Specifically, the study proposes that the integration of unsupervised instance-based resampling and randomizing preprocessing techniques to MLP will lead to improved early detection of self-care disabilities in children. The study further hypothesizes that the preprocessing filters proposed in this study will have a significant impact on the performance of MLP models. Finally, the study hypothesizes that the MLP-progressive methodology will outperform state-of-the-art studies in predicting self-care disabilities in children, as measured by various evaluation metrics, including accuracy, precision, recall, F-measure, TP rate, FP rate, and ROC. The study suggests that the results of the MLP-progressive model will help therapists in preparing personalized therapy plans for children with disabilities.

Key contributions of the proposed MLP-progressive model are as follows:

(i) Integration of unsupervised data pre-processing filters, that is, resampling and randomization with MLP to produce an MLP-progressive model. This has been done for the first time.
(ii) The performance of the standard MLP, called MLP-basic, is improved by using unsupervised data pre-processing filters comprising resampling and randomization.
(iii) We conducted a comparative analysis of the proposed MLP-progressive model with previously designed disability prediction models.

## Literature review

Accurate prediction of disabilities in children based on their self-care activities is a critical task in machine learning. In this literature review section, we discuss previous studies that have used machine learning models, particularly ANNs, and data pre-processing methods such as randomization and resampling to predict disabilities based on self-care activities. By examining the strengths and limitations of previous approaches, we aim to identify gap for the research in this focused study.

Machine learning techniques have been extensively utilized in the healthcare industry to enhance the well-being of patients and improve medical diagnoses. These techniques involve the application of advanced algorithms that enable

the extraction of valuable insights from medical data, thereby providing remarkable illness prediction abilities. The use of machine learning techniques in healthcare has led to several significant breakthroughs, such as the development of predictive models for disease diagnosis and prognosis. These models enable healthcare professionals to identify the early onset of medical conditions, allowing for early intervention and more successful treatment outcomes. Furthermore, machine learning techniques have also been utilized in personalized medicine, which involves tailoring medical treatments to individual patients based on their unique characteristics and medical history.[17] Other uses include the creation of electronic health records, the generation of insight, the enhancement of patient risk score systems, the prediction of illnesses, and the simplification of hospital operations.[18] Machine learning methods have been and could likely be used for predicting disability in children. In this case, the standard ICF-CY dataset has been publicly released for research.[3] SCADI is the only publicly accessible dataset that meets with the ICF-CY requirements for predicting self-care. The original SCADI dataset was provided by Zarchi et al.[4] They used an ANN for the classification of children with disability (CwD) and further applied the decision tree algorithm C4.5 for the extraction of the rules related to self-care prediction problems, obtaining an accuracy of up to 83.10%. A hybrid of principal component analysis and KNN was used by Islam et al.[5] for the reduction of the feature space and prediction of disability in multi-class self-care problems. They achieved an accuracy of 84.29% using 5-fold cross-validation. The information gain regression curve feature selection method was used by the authors[6] along with a NB classifier using a 10-fold cross-validation to achieve the highest accuracy results of 78.32%. The Synthetic Minority Over-Sampling Technique (SMOTE), a dataset class balancing method, was used by Le et al.[19] in combination with extreme gradient boosting (XGBoost) to achieve an improved accuracy of 85.40%. Souza et al.[9] transformed the SCADI dataset into a binary classification dataset with only two sets of instances, that is, positive and negative disability. They used FNN for the purpose of prediction, which produced an accuracy of 85.11% on the binary SCADI dataset. DNNs and extreme learning machines (ELM) have recently been used by Akyol (2020) on a multi-class SCADI dataset[9] using a hold-out method (i.e. 60% and 40% data split) which achieved an accuracy of 97.45% and 88.88% on DNN and ELM, respectively. Hybrids of autoencoder and DNNs[8] and GA + XGBoost[20] have recently been used. GA-XGBoost was tested using the 10-fold CV, which achieved an average accuracy of 84.29% for multi-class datasets and 91.43% for binary-class datasets.

He et al.[21] used multimodal MRI data and clinical data to predict neurodevelopmental deficits in very preterm infants. Results showed improved prediction of cognitive, language, and motor deficits at 2 years corrected age with an accuracy of 88.40%, 87.20%, and 86.70%, respectively. Zdrodowska and Dardzińska-Głębocka[22] extracted classification and action rules to support therapists in recognizing self-care problems in children with disabilities. A model was obtained based on 17 features with the greatest impact on classifying a child into a particular group of self-care problems. Sharmila[23] proposed the use of machine learning algorithms for predicting depression level in children. Results showed that the ResNet Algorithm outperformed other detection approaches. The study suggests using data sets from sources such as AVEC to build real-time depression prediction systems. Sharma[24] presented a machine learning-based expert system for diagnosing and classifying self-care issues in children with physical and mental disorders. Results showed significant improvement in performance with PM-PSO feature selector, with accuracy ranging from 64.28% to 81%. Islam et al.[25] discussed the use of machine learning to detect autism spectrum disorder (ASD) in toddlers. The study aimed to create an online tool for early diagnosis of ASD using supervised learning algorithms. KNN and Random Forest showed the highest accuracy and speed for diagnosis. Seung-Hyun[15] investigated the possibility of using big data from a health insurance database to detect childhood developmental delays leading to disabilities before clinical registration. Multiple classification algorithms were used, and it was found that disabilities could be detected with significant accuracy even at the age of 4 years, which is earlier than the mean diagnostic age of 4.99 years.

The resampling approaches randomize the training and test sets to mimic the quality of training and assessing models on random samples of a dataset from the domain as opposed to a single data sample. In literature, resampling with neural networks has been observed to increase the system performance.[14] Using resampling in conjunction with SMOTE-ENN as a pretreatment strategy, followed by a logistic regression analysis, has improved the prognosis of renal illness in healthcare.[16] Chakravarthy et al.[26] present a study on the effectiveness of resampling techniques on imbalanced datasets using multiple performance measures. The authors compared the performance of Synthetic Minority Over-Sampling and Random Over-Sampling techniques over multiple learning algorithms and resampling ratios for eight different performance measures against two datasets from diverse domains such as medicine and engineering.

AME has been described by Lin et al.[27] as the influence of the data points on the behavior of the model due to randomly chosen data subsets. When added to the training data, they demonstrated that a data point has a significant AME, which influences the behavior. Controlling for confounding factors in machine learning studies requires the

randomization of trials. Papernot[28] discussed the impact of randomization on the development of more robust machine learning models.

A critical review of the state-of-the-art methods in the subject area reveals that none of the previously cited studies have exploited unsupervised instance-based data resampling and randomization together with MLP to increase the prediction accuracy of disability in children. Therefore, the present study proposes a novel model, named MLP-progressive, by combining unsupervised instance-based resampling and randomization methods with MLP to improve the predictive performance of the self-care prediction model.

## Preliminaries

This study proposes a novel feed-forward ANN-based self-care prediction methodology called MLP-progressive for early detection of self-care disabilities in children. The proposed methodology integrates unsupervised instance-based resampling and randomizing preprocessing techniques to improve the performance of MLP model. Hence, the subsequent sub-sections provide basic concepts of the techniques resampling, randomization, and MLP to provide base for the proposed study.

### Resampling data

Resampling techniques have gained significant attention in recent years for improving the performance of machine learning algorithms. In wildlife habitat prediction models, researchers have demonstrated that resampling methods can yield a more accurate estimate of model classification accuracy. Studies have shown that resampling can lead to a significant increase in the precision of the models.[15]

According to a study,[13] repeated resampling can improve the accuracy of the machine learning algorithm by up to 50%. The study found that resampling training data leads to improved performance of the algorithm and reduces the standard deviation by a maximum of up to 50%.

Other studies have also reported that resampling techniques can enhance the performance of machine learning algorithms. For example, a study by Roshan et al.[29] showed that resampling techniques such as oversampling and undersampling can improve the accuracy of decision tree algorithms. Similarly, a study by Thejas et al.[30] found that resampling methods such as SMOTE can improve the performance of support vector machine classifiers.

### Randomization of data

Randomization is a method for allowing an algorithm to explore more of the search space, and depending on how randomization is implemented, it may be seen both as a means of diversity and intensification. There are several methods for intensification and diversification. In reality, each algorithm and its versions adopt distinct strategies to attain a balance between exploration and exploitation.[31] Randomization is basically shuffling the instances of a dataset for bringing diversification in its records. Randomness in data may exist at various levels in machine learning, often boosting the performance or reducing the challenges of conventional approaches.[32] Randomization enables the practitioner of machine learning to generalize findings, making them practical and helpful. The sequence in which cases are presented to a model influences the internal choices. This is particularly sensitive to neural networks and other algorithms. Before each training cycle, it is recommended to randomly mix the training data.[33]

### Multilayer perceptron

The narrator elaborated that the MLP is a kind of feed-forward ANN that emulates the parallel processing and interconnectedness of the human brain. The MLP comprises three layers[34]: the input layer, the hidden layer, and the output layer. The input layer accepts input data in the form of an input signal that is a combination of feature values. Model of the MLP three layers architecture is shown in Figure 1. The output layer undertakes the processing of prediction and classification, grounded on the information conveyed by the input layer. The classified output is then juxtaposed with the observed one, and the error is computed. To update the network weights from the output layer towards the input layer, the error signal serves as a basis for weight adjustment. There are arbitrary amounts of hidden layers interspersed between the input and output layers, responsible for computation and commonly known as the computational engine of the MLP. Each neuron in the hidden layer is linked with the neurons of the next layer, and the connections between them are termed as weights, whose values are optimized through the iterative learning phase. The learning phase is perpetually iterated until the error value reaches a value lower than the threshold level. The backpropagation learning algorithm is a prevalent approach utilized to train the neurons in the MLP.[35]

### Performance metrics

In machine learning, numerous evaluation metrics are utilized to gauge the performance of prediction models. Some of the widely employed metrics are accuracy, precision, recall, F-measure, true positive (TP) rate, false positive (FP) rate, and receiver operating characteristic (ROC) curve. These metrics are employed to evaluate and compare the performance of a proposed MLP-progressive model with its basic and optimized versions, as well as with other prediction models. Each metric has its own advantages and limitations and is utilized in different contexts. However, if an algorithm performs consistently well
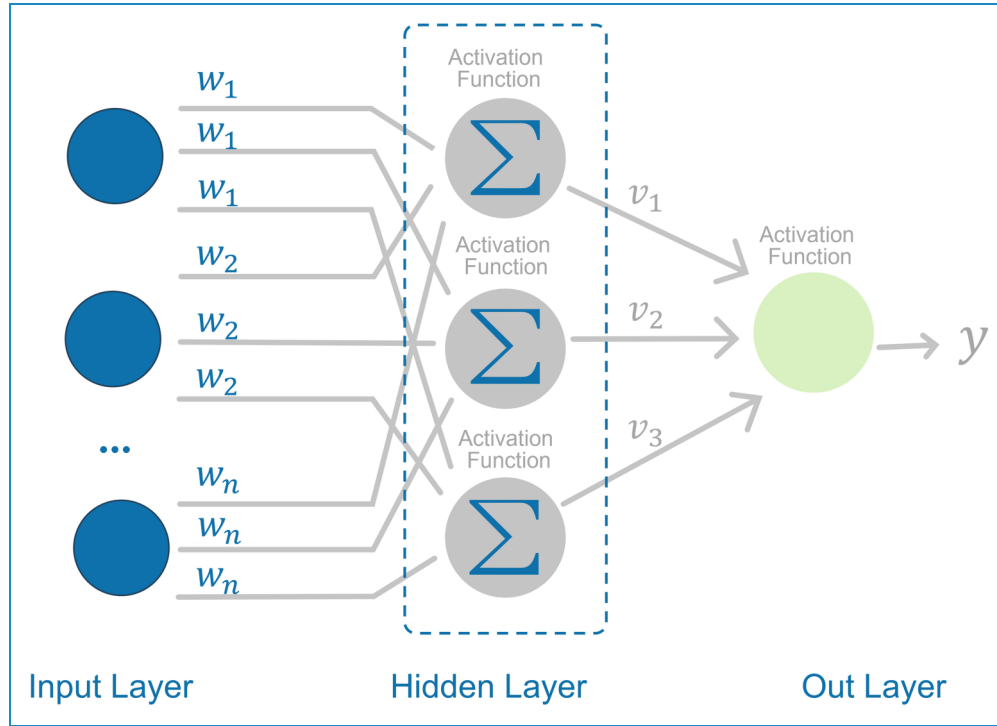
**Figure 1.** Architecture of multilayer perceptron.[34]

across multiple metrics, it is considered to be a reliable and accurate predictor and hence we use them here for evaluating the proposed model.

Mathematical forms of these metrics are shown in Table 1.

## Goodness-of-fit tests

Goodness-of-fit tests are statistical tests used to determine whether a proposed model is a good fit for a given dataset.[36] There are several tests that can be used to evaluate the fitness of a proposed MLP-progressive model, including $p$-value,[37] Fowlkes-Mallows Index (FMI),[38] Normalized Mutual Information (NMI),[39] and Adjusted Rand Index (ARI).[40] To confirm that the proposed MLP-progressive model is correct and the results produced are not random, these goodness-of-fit tests can be applied to evaluate the model's performance. For example, a low $p$-value, high FMI, high NMI, and high ARI would indicate that the model is a good fit for the dataset and that the results produced are not random. Mathematical forms of these goodness-of-fit tests metrics are shown in Table 2.

## Methodology–multilayer perceptron-based prediction model

This section presents the proposed research methodology, which exploits a machine learning prediction model in combination with unsupervised instance-based pre-processing

**Table 1.** Evaluation criteria for both multi-class and binary-class SCADI datasets.

| Evaluation metric | Mathematical form |
|---|---|
| Accuracy | $ACC = \frac{TP+TN}{TP+TN+FP+FN}$ |
| Precision | $Precision = \frac{TP}{TP+FP}$ |
| Recall | $Recall = \frac{TP}{TP+FN}$ |
| F-measure | $F1Score = \frac{2*precision*recall}{precision+recall}$ |
| TP rate | $TPR = \frac{TP}{P} = \frac{TP}{TP+FN} = 1 - FNR$ |
| FP rate | $FPR = \frac{FP}{N} = \frac{FP}{FP+TN} = 1 - TNR$ |
| ROC area | We use it only for the binary classification dataset |

TP: true positive; TN: true negative; FN: false negative; FNR: false negative rate; TNR: true negative rate.

methods. The supervised machine learning predictive model adopted here is MLP, a supplement of a feed-forward neural network. The unsupervised preprocessing methods used are instance-based resampling and randomization filters. The hybrid methodology of these unsupervised filters and a supervised MLP classifier formed the proposed MLP-progressive model depicted in Figure 2 and explained in the subsequent sub-sections.

**Table 2.** Goodness-of-fit tests criteria for both multi-class and binary-class SCADI datasets.

| Goodness-of-fit tests metric | Mathematical form |
|---|---|
| Fowlkes-Mallows Index (FMI) | $FMI = \frac{TP}{\sqrt{(TP + FP) * (TP + FN)}}$ |
| Normalized Mutual Information (NMI) | $NMI = \frac{(2 * I(Y, Y\_pred))}{(H(Y) + H(Y\_pred))}$ |
| Adjusted Rand Index (ARI) | $ARI = \frac{(RI - Expected\_RI)}{(max(RI) - Expected\_RI)}$ |

TP: true positive; TN: true negative; FN: false negative; Y: true labels; Y_pred: predicted labels; RI: Rand Index.

## Self-care training data

To realize the proposed MLP-progressive model, the past data on the self-care activities of children are required. For this purpose, the publicly accessible standard dataset on self-care activity difficulties of children with impairments, SCADI[4] was used. This dataset was collected by occupational therapists in Yazd, Iran[3] from 2016 to 2017 from educational and medical facilities. The demographic information of the children, such as age and sex, was considered during the dataset creation time to add a personalized aspect.

### Self-care dataset (SCADI).
The SCADI dataset contains predictors as conditional or independent attributes and types of disabilities as the class variable or decision attribute. Hence, it is used by supervised machine learning algorithms to automatically predict the disability type possessed by the child. The details are as follows.

#### Predictors.
The SCADI dataset contains the data of 70 children. The age range of the children is 6–18 years. The dataset is comprised of 41% women and 59% men. It has 29 predictors, which are self-care activities, the details of which are presented in Table 2. For each feature code, the level of disability is applied and indicated with values 0–9. A value of 0 for a particular impairment means that the child has no impairment for this particular activity. The gender attribute is represented with 0 for men and 1 for women. The details are shown in Table 3.

#### Disability target classes.
The SCADI disability dataset is a multi-class problem containing seven classes, which the therapist categorizes. These classes are shown in Table 4. The class variable value of 7 indicates the lack of self-care issues. Similarly, values 1 through 6 indicate the existence of a disability. This initial dataset served as the basis for a multiclass dataset.

Furthermore, to show the effectiveness of the proposed methodology, we converted the original 7-class problem to a binary-class problem by considering class 1–6 as the presence of disability and class 7 as the absence of disability. In the case of the binary-class dataset, the output was set to CwD and children without disabilities (CwoD). Details of the binary-class dataset are shown in Table 5.

## Instance-based data preprocessing

It is universally accepted that preprocessed data brings noticeable effects in the performance of machine learning models.[41] We adopted unsupervised instance-based data preprocessing methods in this connection, implemented in Weka.[42] Two filters, instance-based resample and instance-based randomize, were picked from the unsupervised group of Weka filters and supplied to both the multi-class and binary-class datasets shown in Tables 2 and 3. The details of these preprocessing methods are given in the subsequent sub-sections.

### Resampling.
Weka's unsupervised instance-based resample filter for resampling is used for multi- and binary classes. This filter produces a random subsample of the self-care activities dataset using sampling with replacement. The original dataset is entirely loaded into the memory for further processing. In our case, the resample filter ensured the maintenance of class distribution in each subsample. During resampling, this filter was used with default parameters, which are presented in Table 6.

### Randomization.
The randomization filter shuffles the order of the instances passed.[32] For the randomization of the multi- and binary-classes self-care activities dataset, the Weka unsupervised instance-based randomize filter is used, which shuffles the order of the 70 instances in both datasets. We used this filter with the default parameter. The random number seed set in this study was 42.

## Creation of MLP model

In Figure 1, the process of developing self-care prediction models for both multi-class and binary-class problems using MLP in the Weka machine learning tool is illustrated. Preprocessed datasets, which have undergone resampling and randomization, are fed into the MLP algorithm. A 10-fold cross-validation method is utilized to validate the models and store them in the disability prediction knowledge base as learned models.

## MLP-based disability prediction

The MLP trained model, saved as disability prediction knowledge base, is used to test the test dataset, that is, Self-care Test Dataset, in this case. The trained model generates results for the test dataset and returns the disability results for the evaluation metrics enlisted in Table 1.
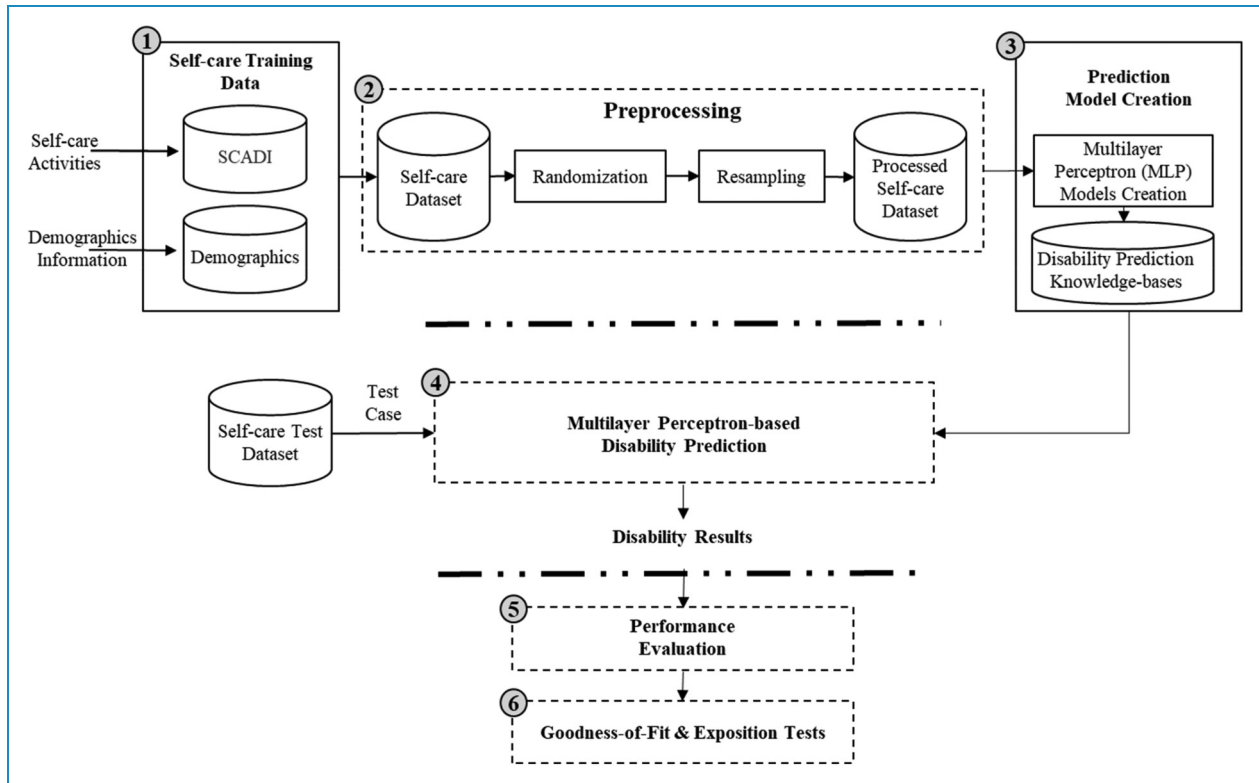
**Figure 2.** The proposed multilayer perceptron (MLP-progressive) model for self-care prediction.

## Performance evaluation

In the performance evaluation phase, the results generated for the metrics enlisted in Table 1 are analyzed for both multi-class and binary-class datasets. Similarly, the obtained results are also compared with results of the state-of-the-art methods. These results and evaluations are presented in sections "Experiments and Results" and "Evaluation, Goodness-of-fit, and Exposition Tests."

## Goodness-of-fit and exposition tests

*Goodness-of-fit tests.* Goodness-of-fit tests, such as *p*-value, FMI, NMI, and ARI, which were introduced in the preliminaries, are commonly utilized to validate the accuracy of a proposed MLP-progressive model and to reject the null hypothesis that the produced results are random. Notably, FMI, NMI, and ARI, with the exception of *p*-value, exhibit a range of values from 0 to 1, where a score of 1 denotes a perfect concordance between the actual and observed sets of labels, while a score of 0 implies no concordance. Furthermore, it is generally accepted that a *p*-value of less than 0.05 signifies statistical significance, indicating that the observed outcomes are unlikely to have occurred by chance.

*Exposition tests.* The proposed MLP-progressive model claims that it is not only correct from technical perspective, but also provides its practical relevance to the early care

professionals' therapeutic interventions. For the purpose, clinical studies would be necessary to demonstrate a causal link between therapy interventions and the results produced. However, in the present study, it is beyond the scope of this paper. However, some statistical analysis may be used to find this connection between different factors and outcomes based on observational data, even in the absence of clinical trial data. We used hypothesis testing to find associations between the factors, such as the patient's age, gender, and the type of therapy they received. The results are summarized in section "Evaluation, Goodness-of-fit, and Exposition Tests.".

## Experiments and results

To validate the proposed MLP-progressive methodology, we performed a series of experiments over self-care datasets (SCADI), transformed to multi-class and binary-class variants. The details are in the subsequent sections.

### Datasets

We used two datasets, which are variants of the SADI dataset, as shown in Tables 3 and 4.

### Experimental setup

To simulate the proposed MLP-progressive model, we used Weka 3.9, an open-source data mining library, with its

**Table 3.** Description of the predictors in the ICF-CY (SCADI) dataset[3].

| Self-care | No. | Feature code | Activity description | Code for disability level |
|---|---|---|---|---|
| Washing oneself | i. | d 5100 | Washing body parts | 0, 1, 2, 3, 4, 8, 9 |
| | ii. | d 5101 | Washing whole body | |
| | iii. | d 5102 | Drying oneself | |
| Caring for body parts | iv. | d 5200 | Caring for skin | 0, 1, 2, 3, 4, 8, 9 |
| | v. | d 5201 | Caring for teeth | |
| | vi. | d 5202 | Caring for hair | |
| | vii. | d 5203 | Caring for fingernails | |
| | viii. | d 5204 | Caring for toenails | |
| | ix. | d 5205 | Caring for nose | |
| Toileting | x. | d 53000 | Indicating need for urination | 0, 1, 2, 3, 4, 8, 9 |
| | xi. | d 53001 | Urinating appropriately | |
| | xii. | d 53010 | Indicating need for defecation | |
| | xiii. | d 53011 | Defecating appropriately | |
| | xiv. | d 5302 | Menstrual care | |
| Dressing | xv. | d 5400 | Putting clothes | 0, 1, 2, 3, 4, 8, 9 |
| | xvi. | d 5401 | Taking off clothes | |
| | xvii. | d 5402 | Putting on footwear | |
| | xviii. | d 5403 | Taking off footwear | |
| | xix. | d 5404 | Choosing appropriate clothing | |
| Eating | xx. | d 5500 | Indicating need for eating | 0, 1, 2, 3, 4, 8, 9 |
| | xxi. | d 5501 | Eating appropriately | |
| Drinking | xxii. | d 5600 | Indicating need for drinking | 0, 1, 2, 3, 4, 8, 9 |
| | xxiii. | d 5602 | Feeding from bottle | |
| Looking after one's health | xxiv. | d 5700 | Ensuring one's physical comfort | 0, 1, 2, 3, 4, 8, 9 |
| | xxv. | d 5701 | Managing diet and fitness | |
| | xxvi. | d 57020 | Managing medications and following health advice | |
| | xxvii. | d 57021 | Seeking advice or assistance from caregivers or professionals | |

**Table 3.** Continued.

| Self-care | No. | Feature code | Activity description | Code for disability level |
|---|---|---|---|---|
| | xxviii. | d 57022 | Avoiding risks of abuse of drugs or alcohol | |
| Looking after one's safety | xxix. | d 571 | Looking after one's safety | 0, 1, 2, 3, 4, 8, 9 |

Codes for the intensity level of disability are interpreted as follows. 0: no impairment, 1: mild impairment, 2: moderate impairment, 3: severe impairment, 4: complete impairment, 8: no specified, and 9: not applicable.

**Table 4.** Description of SCADI multi-class dataset.

| S. No. | Label | Count |
|---|---|---|
| 1 | Class1: Unable to care for body parts | 2 |
| 2 | Class2: Unable to do toileting activities | 7 |
| 3 | Class3: Unable to dress themselves | 1 |
| 4 | Class4: Unable to wash and care for body parts and dressing | 12 |
| 5 | Class5: Unable to wash, care for body parts, toileting, and dressing | 3 |
| 6 | Class6: Unable to eat, drink, wash, care for body parts, toileting, and dressing | 29 |
| 7 | Class7: No issues | 16 |

**Table 5.** Description of the SCADI binary-class dataset.

| S. No. | Label | Count |
|---|---|---|
| 1 | CwD: Children with Disability | 54 |
| 2 | CwoD: Children without Disability | 16 |

**Table 6.** Parameters of the unsupervised instance-based resampling filter.

| S. No. | Parameter | Value |
|---|---|---|
| 1 | invertSelection | False |
| 2 | noReplacement | False |
| 3 | randomSeed | 1 |
| 4 | sampleSizePercent | 100 |

default parameters, as the simulation tool on a standalone Intel(R) Core(TM) i5-8250U CPU @ 1.60 GHz 1.80 GHz system with an 8.00 GB memory and a 64-bit operating system.

## Performance metrics

We used performance metrics shown in Table 1.

## Experiments

To validate the proposed methodology, we performed the following four experiments: (a) simulation of the MLP-progressive model using a multi-class dataset, (b) simulation of the MLP-progressive model using a binary-class dataset, (c) impact of resampling and randomization on prediction using multi- and binary-class datasets and goodness-of-fit and exposition tests. The details of these experiments are provided in "Experiments and Results".

*MLP-progressive on a multi-class dataset.* In the first experiment, the multi-class dataset, presented in Table 4, was supplied to the proposed MLP-progressive model in the following sequence.

After the design and execution of experiment 1, as per the sequence specified in Table 7, the proposed methodology yielded the results shown in Tables 8 and 9. As per the summary results, the proposed model achieves an accuracy of 97.2%.

The dataset exhibits an imbalance in the distribution of instances across its constituent classes, with classes 1, 3, and 5 having fewer instances than the other classes. This characteristic results in a reduced diversity of training data and decreased separability between the classes, potentially leading to underfitting of the MLP model. As a result, the model may be inadequately equipped to capture the underlying patterns and relationships between the features and class labels for these classes, thus resulting in relatively lower classification accuracy for these classes in comparison to the other classes.

*MLP-progressive on binary dataset.* In the second experiment, the binary-class dataset, presented in Table 5, was

**Table 7.** The flow of experimental 1 on multi-class dataset.

| Sequence No./Steps No. | Operation | Outcome |
|---|---|---|
| 1 | Loading dataset | SCADI-Multi-Class |
| 2 | Applying randomization | weka.filters.unsupervised.instance.Randomize-(default parameter) |
| 3 | Applying resampling | weka.filters.unsupervised.instance.Resample-(default parameter) |
| 4 | Building and testing prediction model | weka.classifiers.functions.MultilayerPerceptron (default parameter) |

**Table 8.** Class-wise performance of MLP-progressive model on multi-class dataset.

| === Summary === | Correctly classified instances (68/70) 97.14% | | | | |
|---|---|---|---|---|---|
| | Detailed accuracy by class | | | | |
| Class | Precision | Recall | F1-score | TP rate | FP rate |
| Class 1 | 1 | 1 | 1 | 1 | 0 |
| Class 2 | 1 | 1 | 1 | 1 | 0 |
| Class 3 | 0 | 0 | 0 | 0 | 0 |
| Class 4 | 1 | 1 | 1 | 1 | 0 |
| Class 5 | 1 | 1 | 1 | 1 | 0 |
| Class 6 | 1 | 0.96 | 0.98 | 0.96 | 0 |
| Class 7 | 0.90 | 1 | 0.95 | 1 | 0.03 |
| Weighted Avg. | 0.96 | 0.971 | 0.96 | 0.97 | 0.01 |

supplied to the proposed MLP-progressive model in the following sequence.

In the second experiment, the sequence or steps of operations were performed as per the specification shown in Table 10. The results obtained are shown in Tables 11 and 12. The results/values for the other measures are also shown in Table 12. According to the final results, the accuracy of the proposed model is 98.57%.

*Impact of randomization and resampling on performance.* To validate the effect of unsupervised filters on the prediction of the MLP-based model using the SCDAI dataset, we designed four experiments, named MLP-standard, MLP-optimized I, MLP-optimized II, and MLP-progressive. In the first experiment, MLP-standard, no preprocessing method was used. In the second experiment, MLP-optimized I, only the resample

filter is applied before applying MLP. The third experiment, MLP-optimized II, is designed by backing up the resample filter with a randomize filter. The fourth and last experiment, MLP-progressive, is designed by using the filters and proposed method in the sequence specified as followed: randomization→resampling→MLP model. Figure 3 further compares all the four variants side-by-side for multi-class and binary-class problems. The results obtained for both the multi-class and binary-class problems are shown in Tables 12 and 13, respectively.

*Performance impact on multi-class dataset.* Executing the MLP standards, the system achieved an accuracy of 80%. After executing MLP-optimized I, the accuracy obtained is improved from 80% to 91.40%. In the third experiment, MLP-optimized II, the resample and randomize filters back up the MLP method. However, no change in results obtained is observed. In the fourth experiment, MLP-progressive, randomization was followed by resampling, which further boosts performance of the MLP from 91.10% to 97.14%. The detailed results are shown in Table 13.

*Performance impact on binary-class dataset.* For the binary-class problem, the proposed methodology was performed using the same set of four experiments resulting in four different predictive models. Table 14 shows the results of the basic, optimized, and progressive models. Table 14 shows the accuracy of all four variants of the proposed MLP model, with the values 92.81, 97.14, 97.14, and 98.57 for MLP-standard, MLP-optimized I, MLP-optimized II, and MLP-progressive, respectively.

From the results it is evident that instance-based resample filter for resampling the dataset and then using MLP technique is an effective approach. The rationales behind this are that MLP is well-suited for resampled datasets because it can handle complex nonlinear relationships between the input features and the output labels. The oversampling or undersampling performed by the instance-based resample filter increases complexity of the dataset, and hence the MLP can best handle this increased complexity by adapting its weights and biases during training. Therefore, the use of the instance-based resample filter for

**Table 9.** Confusion matrix of MLP-progressive on multi-class dataset.

| Class | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 | Class 6 | Class 7 |
|-------|---------|---------|---------|---------|---------|---------|---------|
| Class 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Class 2 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| Class 3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Class 4 | 0 | 0 | 0 | 11 | 0 | 0 | 0 |
| Class 5 | 0 | 0 | 0 | 0 | 4 | 0 | 0 |
| Class 6 | 0 | 0 | 0 | 0 | 0 | 29 | 1 |
| Class 7 | 0 | 0 | 0 | 0 | 0 | 0 | 19 |

resampling and MLP technique can be justified by their complementary strengths in addressing imbalanced datasets and handling complex relationships between features and labels, respectively. This combination can lead to improved predictive performance and generalizability of the trained model.

## Evaluation, goodness-of-fit, and exposition tests

This section compares results of the proposed MLP-progressive model with state-of-the-art methods and checks its goodness-of-fits and exposition.

### Comparison and evaluation

The comparison results show that in the case of a multi-class dataset, the proposed MLP progressive outperforms the state-of-the-art model. However, in the case of binary classification, the proposed model and GA-XGBoost perform similarly. The side-by-side results of the proposed MLP-progressive model are shown in Figure 4. These results are shown in Tables 15 and 16, respectively.

### Results of goodness-of-fit tests

Similarly, the results for goodness-of-fit tests are shown in Table 17, which are obtained by following the procedure as discussed below. First, we train the MLP-based self-care model on a training set and then use the trained model to make predictions on a test set. Then we calculate the mean squared error (MSE) between the observed values in the test set and the predicted values from the model. Lower MSE value indicates a good fit between the model and the data.

The findings reveal noteworthy results concerning the evaluation metrics, including FMI, NMI, and ARI. Specifically, the computed values of these measures approach 1, which indicates a high degree of concordance

between the ascertained and true labeling sets. Moreover, the derived $p$-values exhibit statistical significance, with a significance level of less than 0.05, underscoring that the observed outcomes are improbable to have transpired by chance. Hence, the findings underscore the reliability and validity of the study outcomes, with implications for further research and practice.

### Results of exposition test

We used $t$-test and a chi-square test to test the exposition hypotheses. In the subject dataset, we have factors, such as the patient's age, gender, and the type of therapy they received. The findings imply that there is no significant difference in male and female patients' therapy outcomes. A $p$-value larger than 0.05 was obtained from the $t$-test comparing the therapy outcome for male and female patients, demonstrating that the null hypothesis—that there is not a significant distinction between the therapeutic outcomes of male and female patients—cannot be rejected.

We also conducted a $t$-test to compare the means of two age groups. The difference in averages between the two age groups (age = 12 and age > 12) based on the dataset record sets is not statistically significant at the 0.05 level of significance, according to the $t$-test findings, since the $p$-value of 0.18 is higher than the threshold of 0.05. Considering this, we are unable to rule out the null hypothesis that there is not a significant difference in the prevalence of self-care issues between the two age groups. It is crucial to keep in mind that although the $t$-statistic of 1.34 indicates a moderate difference between the two groups, this difference is too small to be regarded as statistically significant based on the provided $p$-value.

## Discussion

In this study, we explored the performance of the proposed MLP-progressive model on a binary-class dataset and

**Table 10.** The flow of experimental-2 on multi-class dataset.

| Sequence No. | Operation | Outcome |
|---|---|---|
| 1 | Loading dataset | SCADI-Binary-Class |
| 2 | Applying randomization | weka.filters.unsupervised.instance.Randomize-(default parameter) |
| 3 | Applying resampling | weka.filters.unsupervised.instance.Resample-(default parameter) |
| 4 | Building prediction model | weka.classifiers.functions.MultilayerPerceptron(default parameter) |

**Table 11.** Class-wise performance of MLP-progressive model on binary-class dataset.

| === Summary === | Correctly classified instances (69/70) 98.54% | | | | | |
|---|---|---|---|---|---|---|
| | Detailed accuracy by class | | | | | |
| Class | Precision | Recall | F1-score | TP rate | FP rate | ROC area |
| CwD | 1 | 0.98 | 0.99 | 0.98 | 0 | 1 |
| CwoD | 0.95 | 1 | 0.97 | 1 | 0.02 | 1 |
| Weighted Avg. | 0. 98 | 0. 98 | 0. 98 | 0.98 | 0.005 | 1 |

**Table 12.** Confusion matrix of MLP-progressive on binary-class dataset.

| Class | CwD | CwoD |
|---|---|---|
| CwD | 50 | 1 |
| CwoD | 0 | 19 |

compared it with other types of MLP models and analyzed the impact of randomization and resampling on the model's performance. Additionally, the goodness-of-fit and exposition tests were conducted to further evaluate the model. First, we tested the performance of the MLP-progressive model on the binary-class dataset. The results showed that the MLP-progressive model achieved an accuracy of 98.57%. This indicates that the model was able to classify the instances correctly in 98.54% of the cases. The precision, recall, and F1-score values were also high, demonstrating the model's ability to perform well in both classes (CwD and CwoD). The confusion matrix showed that the model made only one misclassification out of the 70 instances, which is a strong performance.

Next, the impact of randomization and resampling on the performance of MLP models was explored. Four tests were performed: MLP-standard, MLP-optimized I, MLP-optimized II, and MLP-progressive. The purpose was to investigate how different preprocessing strategies influence the performance of the MLP models. The results demonstrated that the MLP-progressive model yielded the best performance among all variations, in terms of accuracy, precision, recall, and F1-score for both the multi-class and binary-class datasets. This indicates that the sequential application of randomization and resampling, followed by the MLP model, resulted in the highest overall performance.

Furthermore, it was observed that resampling alone (MLP-optimized I) had a positive impact on the model's performance. However, the addition of randomization (MLP-optimized II) did not lead to any further improvement in performance. This suggests that the order in which the preprocessing techniques are applied plays a crucial role in achieving optimal results. These findings highlight the importance of carefully designing the preprocessing pipeline for MLP models. Specifically, the combination of randomization and resampling, in the specified sequence, proved to be effective in enhancing the predictive performance of the MLP model. It is worth noting that resampling can address imbalanced datasets by introducing oversampling or undersampling, while MLP is well-suited for handling complex relationships between features and labels. Overall, the study emphasizes the significance of preprocessing methods in improving the accuracy and generalizability of MLP models. By selecting and ordering the
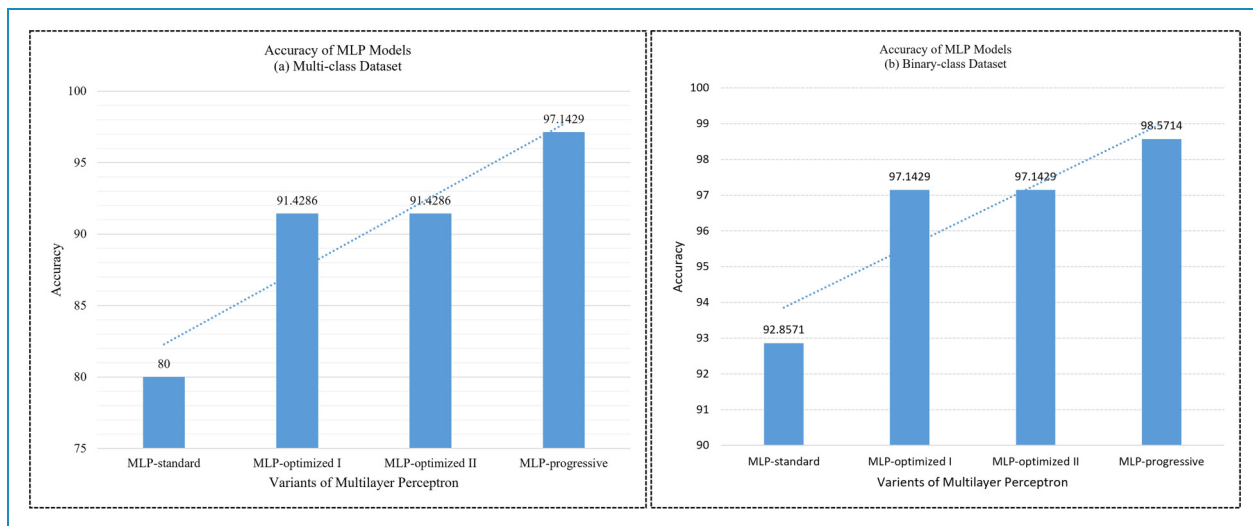
**Figure 3.** Comparative analysis of different variants of the proposed MLP-based models.

**Table 13.** Performance impacts of the proposed preprocessing filters on multi-class dataset.

| Method | Preprocessing | | No. of features (%) | Performance metric (%)—accuracy measures | | | | Time complexity (s) |
|---|---|---|---|---|---|---|---|---|
| | Filter 1 | Filter 2 | | Accuracy | Precision | Recall | F-measure | Time taken to build model |
| MLP-standard | - | - | 205 (100%) | 80.00 | 0.77 | 0.80 | 0.78 | 11.53 s |
| MLP-optimized I | Resample | - | | 91.42 | 0.89 | 0.91 | 0.90 | 10.96 s |
| MLP-optimized II | Resample | Randomize | | 91.42 | 0.89 | 0.91 | 0.90 | 11.41 s |
| MLP-progressive | Randomize | Resample | | 97.14 | 0.96 | 0.97 | 0.96 | 11.29 s |

**Table 14.** Performance impact of the proposed preprocessing filters on the binary-class dataset.

| Method | Preprocessing | | No. of features (%) | Performance metric (%)—accuracy measures | | | | Time complexity (s) |
|---|---|---|---|---|---|---|---|---|
| | Filter 1 | Filter 2 | | Accuracy | Precision | Recall | F-measure | Time taken to build model |
| MLP-standard | - | - | 205 (100%) | 92.85 | 0.92 | 0.92 | 0.92 | 10.46 s |
| MLP-optimized I | Resample | - | | 97.14 | 0.97 | 0.97 | 0.97 | 10.37 s |
| MLP-optimized II | Resample | Randomize | | 97.14 | 0.97 | 0.97 | 0.97 | 10.57 s |
| MLP-progressive | Randomize | Resample | | 98.57 | 0.98 | 0.98 | 0.98 | 10.44 s |
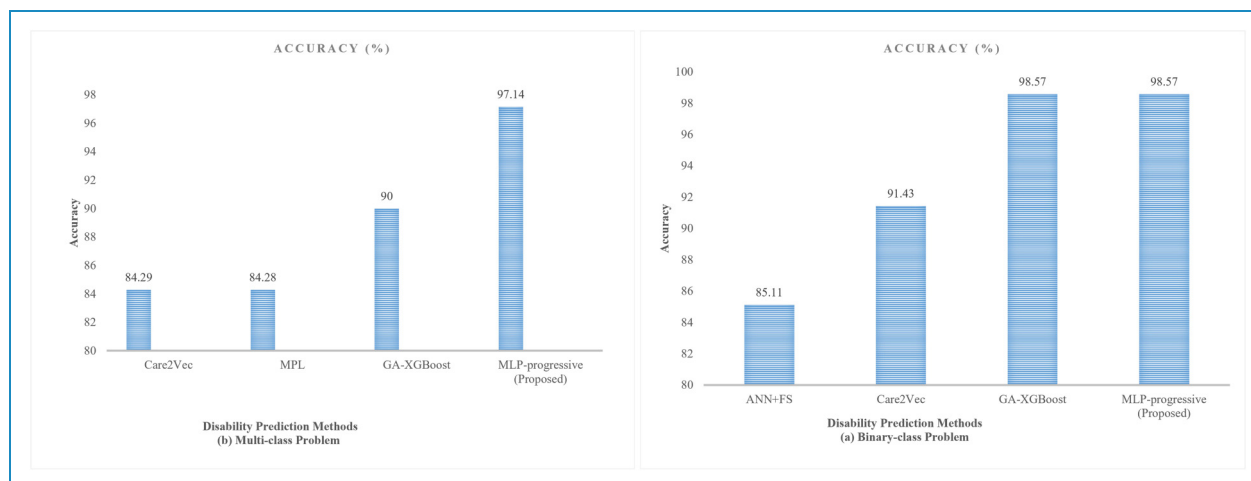
**Figure 4.** Comparison of the proposed model with past studies.

**Table 15.** Comparison of the MLP-progressive model with state-of-the-art methods on multi-class problem.

| Method | Year | Feature selection | No. of features (%) | Validation method | Classification type | Accuracy (%) |
|---|---|---|---|---|---|---|
| Care2Vec: a hybrid of autoencoders and deep neural networks[8] | 2020 | - | 205 (100%) | 10-fold CV | Multi-class | 84.29 |
| GA-XGBoost[20] | 2020 | GA | 88 (42.93%) | 10-fold CV | Multi-class | 90.00 |
| Multilayer perceptron (MLP)[43] | 2021 | - | 205 (100%) | 10-fold CV | Multi-class | 80.00 |
| | | Chi-square test | 39 (19.02%) | | | 82.85 |
| | | Correlation-based feature subset selection | 19 (9.27%) | | | 84.28 |
| MLP-progressive model (Randomize + Resample + MLP) | 2022 | - | 205 (100%) | 10-fold CV | Multi-class | 97.14 |

appropriate preprocessing techniques, researchers and practitioners can enhance the performance of MLP models on binary-class and multi-class datasets.

Furthermore, the comparison of the proposed MLP-progressive model with state-of-the-art methods showed promising results. In the multi-class problem, the MLP-progressive model achieved an accuracy of 97.14%, outperforming the other compared models. In the binary-class problem, the MLP-progressive model achieved an accuracy of 98.57%, like the performance of the GA-XGBoost model. These results demonstrate the competitiveness of the proposed MLP-progressive model in both multi-class and binary-class classification tasks.

The goodness-of-fit tests confirmed the reliability and validity of the MLP-progressive model. The low MSE values indicate a good fit between the model and the data, while the high values of the FMI, NMI, and ARI indicate a high degree of concordance between the observed and true labeling sets. These results provide further evidence of the model's effectiveness.

Finally, the exposition tests explored the relationship between different factors in the dataset, such as age, gender, and therapy outcomes. The results showed that there was no significant difference in therapy outcomes between male and female patients, as well as between different age groups. These findings suggest that gender and

**Table 16.** Comparison of the MLP-progressive model with state-of-the-art methods on binary-class problem.

| Method | Year | Feature selection | No. of Features (%) | Validation method | Classification type | Accuracy (%) |
|--------|------|-------------------|---------------------|-------------------|---------------------|--------------|
| Artificial neural networks and fuzzy systems[9] | 2019 | - | 205 (100%) | $k$-fold * | Binary-class | 85.11 |
| Care2Vec: a hybrid of autoencoders and deep neural networks[8] | 2020 | - | 205 (100%) | 10-fold CV | Binary-class | 91.43 |
| GA-XGBoost | 2020 | GA | 111 (54.15%) | 10-fold CV | Binary-class | 98.57 |
| MLP-progressive model (Randomize + Resample + MLP) | 2022 | - | 205 (100%) | 10-fold CV | Binary-class | 98.57 |

**Table 17.** Results of goodness-of-fit tests.

| S. No. | Test | Multi-class dataset | Binary-class dataset |
|--------|------|---------------------|----------------------|
| 1 | Mean squared error | 0.28 | 0.03 |
| 2 | $p$-value | 2.88 e-16 | 0.000464 |
| 3 | Fowlkes-Mallows Index (FMI) | 0.935 | 0.95 |
| 4 | Normalized Mutual Information (NMI) | 0.93 | 0.69 |
| 5 | Adjusted Rand Index (ARI) | 0.98 | 0.96 |

age may not be influential factors in predicting therapy outcomes in the given dataset.

Overall, the results and analyses we have shown in this study support the effectiveness of the proposed MLP-progressive model for binary-class classification tasks. The sequential application of randomization and resampling, followed by the MLP model, proved to be beneficial for improving the model's performance. The comparison with state-of-the-art methods demonstrated the competitiveness of the proposed model. The goodness-of-fit and exposition tests further validated the model's reliability and provided insights into the relationship between different factors in the dataset.

## Conclusions, limitations, and future works

This research paper proposed a MLP-based self-care prediction methodology, called MLP-progressive, for early detection of self-care disabilities in children. The methodology has integrated unsupervised instance-based resampling and randomizing preprocessing techniques to MLP that has resulted in improved performance. The study conducted three experiments including validation of the proposed methodology on multi-class and binary-class

datasets, performing impact analysis of the proposed preprocessing filters on model performance, and comparing results with state-of-the-art studies. The evaluation metrics used to measure the performance of the models include accuracy, precision, recall, F-measure, TP rate, FP rate, and ROC. The outcomes reveled that MLP-progressive performs better than other prediction models in terms of classification accuracies. The proposed model achieved accuracy in the range of 91.10–97.14% on multi-class dataset and 92.85, 97.14, 97.14, and 98.57 for MLP-standard, MLP-optimized I, MLP-optimized II, and MLP-progressive, respectively, over binary-class dataset. The evaluation results show significant improvements in accuracy (90–97.14%) over multi-class problem and same results in the binary-class problem as compared to GA-XGBoost. The improved performance, in case of multi-class problem, indicates that MLP-progressive can help professional therapists in preparing personalized therapy plans for children with disabilities, improving the accuracy of early identification of self-care impairments.

The study is subject to certain limitations, namely, the utilization of a restricted sample size and the absence of external validation. The limited sample size could have negative consequences on the generalizability of the findings to a

larger population of children with disabilities. Additionally, the lack of external validation could lead to reduced reliability and generalizability of the study's outcomes.

Future research directions for the study include conducting external validation of MLP-progressive on different datasets and settings to increase the reliability and generalizability of the findings. Other machine learning models could be explored and compared to MLP-progressive to determine the most effective model for early identification of self-care impairments. The proposed methodology could also be extended to investigate the potential application of detecting other types of disabilities in children, offering new ways to apply machine learning methods for disability detection. A longitudinal study could also be conducted to evaluate MLP-progressive's effectiveness and reliability in detecting self-care impairments in children over time.

**Contributorship:** Conceptualization: R.A. Data curation: R.A. Investigation: R.A, J.H. Methodology: R.A, J.H. Writing original draft: R.A. Review & editing: A.N., J.H. Reviewing and financial contribution: S.W.L.

**Declaration of conflicting interests:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethical approval:** This study was approved by the Ethics Committee of the University of Peshawar.

**Guarantor:** RA.

**ORCID iD:** Seung Won Lee ⓘD https://orcid.org/0000-0001-5632-5208

## References

1. Wu T-K, Huang S-C and Meng Y-R. *Identifying and diagnosing students with learning disabilities using ANN and SVM*. in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. 2006. IEEE.
2. Haley SM, et al. Pediatric evaluation of disability inventory. *Phys Occup Ther Pediatr* 1992; 11: 13.
3. Organization WH. *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY*. Geneva, Switzerland: World Health Organization, 2007.
4. Zarchi MS, Bushehri SF and Dehghanizadeh M. SCADI: a standard dataset for self-care problems classification of children with physical and motor disability. *Int J Med Inf* 2018; 114: 81–87.
5. Islam B, Ashafuddula NM and Mahmud F. *A machine learning approach to detect self-care problems of children with physical and motor disability*. in 2018 *21st International Conference of Computer and Information Technology (ICCIT)*. 2018. IEEE.
6. Liu L, et al. *Feature selection based on feature curve of subclass problem*. in 2019 *International Joint Conference on Neural Networks (IJCNN)*. 2019. IEEE.
7. Akyol K. Comparing of deep neural networks and extreme learning machines based on growing and pruning approach. *Expert Syst Appl* 2020; 140: 112875.
8. Putatunda S. Care2Vec: a hybrid autoencoder-based approach for the classification of self-care problems in physically disabled children. *Neural Comput Appl* 2020; 32: 17669–17680.
9. Souza PV, et al. *Using hybrid systems in the construction of expert systems in the identification of cognitive and motor problems in children and young people*. in 2019 *IEEE international conference on fuzzy systems (FUZZ-IEEE)*. 2019. IEEE.
10. Chai SS, et al. A multilayer perceptron neural network model to classify hypertension in adolescents using anthropometric measurements: a cross-sectional study in Sarawak, Malaysia. *Comput Math Methods Med* 2021; 2021: 2794888.
11. Parsaei H, Faraz M and Mortazavi S. A multilayer perceptron neural network-based model for predicting subjective health symptoms in people living in the vicinity of mobile phone base stations. *Ecopsychology* 2017; 9: 99–105.
12. Bikku T. Multi-layered deep learning perceptron approach for health risk prediction. *J Big Data* 2020; 7: 1–14.
13. Molinaro AM, Simon R and Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. *Bioinformatics* 2005; 21: 3301–3307.
14. Rodrigues BP, et al. Resampling in neural networks with application to spatial analysis. *Geo-spatial Inf Sci* 2022; 25: 1–12.
15. Verbyla DL and Litvaitis JA. Resampling methods for evaluating classification accuracy of wildlife habitat models. *Environ Manag* 1989; 13: 783–787.
16. Shi X, et al. A resampling method to improve the prognostic model of end-stage kidney disease: a better strategy for imbalanced data. *Front Med* 2022; 9: 730748.
17. Shailaja K, Seetharamulu B and Jabbar M. *Machine learning in healthcare: A review*. in 2018 *Second international conference on electronics, communication and aerospace technology (ICECA)*. 2018. IEEE.
18. Panesar A. *Machine learning and AI for healthcare*. Coventry, UK: Springer, 2019.
19. Le T and Baik SW. A robust framework for self-care problem identification for children with disability. *Symmetry* 2019; 11: 89.
20. Syafrudin M, et al. A self-care prediction model for children with disability based on genetic algorithm and extreme gradient boosting. *Mathematics* 2020; 8: 1590.
21. He L, et al. Deep multimodal learning from MRI and clinical data for early prediction of neurodevelopmental deficits in very preterm infants. *Front Neurosci* 2021; 15: 753033.
22. Zdrodowska M and Dardzińska-Głębocka A. Classification and action rules in identification and self-care assessment problems. *Technol Health Care* 2022; 30: 257–269.
23. Sharmila M. Depression level calculation for predicting child psychometric retardation using DepressNet approach through

GPU accelerated Google cloud platform. *Turkish J Comput Math Educ (TURCOMAT)* 2021; 12: 1879–1887.

24. Sharma M. Categorization of self care problem for children with disabilities using partial swarm optimization approach. *Int J Inf Technol* 2020; 14: 1–9.

25. Islam S, et al. *Autism spectrum disorder detection in toddlers for early diagnosis using machine learning.* in 2020 *IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE).* 2020. IEEE.

26. Chakravarthy AD, et al. *Predictive models with resampling: a comparative study of machine learning algorithms and their performances on handling imbalanced datasets.* in 2019 *18th IEEE International Conference On Machine Learning And Applications (ICMLA).* 2019. IEEE.

27. Lin J, et al. *Measuring the Effect of Training Data on Deep Learning Predictions via Randomized Experiments.* in *International Conference on Machine Learning.* 2022. PMLR.

28. Papernot N. *The Role of Randomization in Trustworthy Machine Learning.* in *Proceedings of the 9th ACM Workshop on Moving Target Defense.* 2022.

29. Roshan SE and Asadi S. Improvement of bagging performance for classification of imbalanced datasets using evolutionary multi-objective optimization. *Eng Appl Artif Intell* 2020; 87: 103319.

30. Thejas G, et al. An extension of synthetic minority oversampling technique based on Kalman filter for imbalanced datasets. *Mach Learn Appl* 2022; 8: 100267.

31. Barnes C, et al. Patchmatch: a randomized correspondence algorithm for structural image editing. *ACM Trans Graph* 2009; 28: 24.

32. Popkov YS, Dubnov YA and Popkov AY. *Randomized machine learning: statement, solution, applications.* in 2016 *IEEE 8th International Conference on Intelligent Systems (IS).* 2016. IEEE.

33. Gallicchio C, et al. *Randomized Machine Learning Approaches: Recent Developments and Challenges.* in *ESANN.* 2017.

34. Bento C. *Multilayer Perceptron explained with a real-life example and python code: sentiment analysis.* Towards Data Science, 2021.

35. Pinkus A. Approximation theory of the MLP model in neural networks. *Acta Numer* 1999; 8: 143–195.

36. Gaudart J, Giusiano B and Huiart L. Comparison of the performance of multi-layer perceptron and linear regression for epidemiological data. *Comput Stat Data Anal* 2004; 44: 547–570.

37. Shen X, et al. A goodness-of-fit test based on neural network sieve estimators. *Stat Probab Lett* 2021; 174: 109100.

38. Fowlkes EB and Mallows CL. A method for comparing two hierarchical clusterings. *J Am Stat Assoc* 1983; 78: 553–569.

39. Fattah SA, Lin C-C and Kung S-Y. *A mutual information based approach for evaluating the quality of clustering.* in 2011 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* 2011. IEEE.

40. Chacón JE and Rastrojo AI. Minimum adjusted Rand index for two clusterings of a given size. *Adv Data Anal Classif* 2023; 17: 125–133.

41. Zelaya CVG. *Towards explaining the effects of data preprocessing on machine learning.* in 2019 *IEEE 35th international conference on data engineering (ICDE).* 2019. IEEE.

42. Hall M, et al. The WEKA data mining software: an update. *ACM SIGKDD Explor Newsl* 2009; 11: 10–18.

43. Dardzińska-Głębocka A and Zdrodowska M. Analysis children with disabilities self-care problems based on selected data mining techniques. *Procedia Comput Sci* 2021; 192: 2854–2862.