Pediatric HSCT survival rates detection through the application of different ANN model optimizers, BSA, VSA, and GOA

Nima Javanmehr  
Babol University of Medical Sciences

Hossein Moayedi  
Duy Tan University

Fahimeh Farokhnia (fahimehfarokhnia2044@gmail.com)  
Kermanshah University of Medical Sciences

Research Article

Keywords: Machine Learning, Survival Rate Estimation, Neural Networks, Artificial Intelligence, HSCT survival, Pediatric BMT, Evolutionary Algorithms, Optimization

Posted Date: February 16th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2574558/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Machine learning (ML) possesses unique characteristics that render it useful in a variety of applications. Thanks to creative approaches to observing complex clinical data through the lens of mathematical coding, researchers have uncovered a crossroad between computer and medical sciences that offers an exciting landscape to improve the current clinical diagnostic and therapeutic approaches. Bringing together findings from multiple sources, such as private health information, laboratory, and physical examination, neural networks have yielded novel modeling systems in which different features in the medical dataset dynamically contribute to the maturation of the system’s predicting and classifying functions. This potency is commonly attributed to the training function in a neural network, which enables the ANN to autonomously recognize the link between the input and outputs of a particular database. Besides the ANN’s groundbreaking promises, a bulk of applications have surfaced its existing limitations, including local minima entrapment and extended processing time. In this context, evolutionary algorithms (EAs) are developed to address the ANN’s shortcomings. In the present research, we recruit ANN-based BSA, VSA, and GOA algorithms to optimize the neural network’s prediction competence. The proposed models are utilized in a database from UCI databank to predict the outcome of bone marrow transplant in children with hematologic conditions. Root-mean-square deviation (RMSD), ROC, and AUC measures are harnessed to analyze and compare the outcomes of different models. The ANN-BSA model is recognized to bring about the most viable results concerning the relationship between input and output layer data (that is, clinical features and survival rates). This research provides solid proof of the significant assistance of ML systems to healthcare practitioners to estimate an individual-based prognosis.

1 Introduction

Hematopoietic stem cell transplant (HSCT) in children has been on a burgeoning rise for the past three decades, employed as the standard medical care for a variety of neoplastic and nonmalignant bone marrow disorders. HSCT is the state-of-the-art treatment for multiple neoplastic diseases, such as leukemia, lymphoma, and solid tumors (CNS neoplasms, neuroblastoma) (1). It is also becoming the prevailing approach in nonmalignant conditions, such as severe aplastic anemia and the spectrum of hereditary conditions, including metabolic and storage disorders, immunity disorders, histiocytic diseases, bone marrow failure syndrome, and hemoglobinopathies (2). This broader spectrum of indications for therapeutic bone marrow transplant (BMT) is partly ascribed to a constellation of factors, such as reevaluation of the candidate qualification assessment criteria, well conditioning, use of novel HSC sources, along with improved post-transplant care (3). Intriguingly, in pediatric leukemia patients (< 18y), who underwent matched sibling donor grafts, researchers observed a 5-year survival rate of over 80%, hence highlighting that HSCT is a solid bet in the field of bone marrow disorders research. On the other hand, transplants using other marrow sources commonly confront staggering obstacles, which imperatively impinge on the HSCT outcome. This highlights the limited knowledge regarding peri-transplant risk factors and necessity to develop solutions to optimize the outcomes and minimize the untoward consequences (4–6).

Recently, healthcare professionals have been vigorously avoiding status quo bias and opened the doors for out-of-the-box computation intelligence methodology to improve the efficiency of medical diagnosis and therapeutics. ML utilizes a constellation of datasets, particular computer hardware, software, and analytics, to accelerate data exchange, data assessment, and pattern and behavior recognition in the input and output data (7).
Artificial intelligence (AI) is, in essence, a semantic emulation of the human cognitive system. ML is a category of AI techniques based on autonomous training algorithms that improves by experience compilation and alters in response to different input entries without manual programming. ML algorithms rely primarily on a specific subset of the database, and training data, to bring about prediction and classification models. Neural networks are among the main algorithms used in ML methodology, comprising deep learning systems denoted as artificial neural networks (ANN) (8).

Revolutionary hardware design and developments and big striking data are considered the most powerful levers to create and evolve ML systems. Supervised learning requires researchers to present the datasets, model the representations, and label data. The classification arm in ML is exemplary of supervised learning, putting input data into certain output sets. Mounting evidence highlights the prediction arm of ML, as it recognizes the rules in a given database and generalizes the noted rules to the rest of the samples (9). Notably, other commonly used programs use a set of clear commands to draw the relationship between input and output data, contrasting with the induced rules in ML (10, 11).

Computational intelligence is moving into the mainstream on various paths, comprising digital advertising solutions, digital image processing, pattern recognition, and spam filters (12). In the medical field, ML is applied to various problems, including etiological diagnosis of disorders, as Diao and colleagues employed ML to detect secondary hypertension in a retrospective study using electronic medical records (13). Clinical imaging diagnostics as employed in tumor diagnosis (14), outbreak prediction (15), clinical trials (16), digital health records (17), personalized medicine (18), and drug development (19) are among the significant contributions of ML to the medical field, which we will further discuss in the following sections.

ANNs are considered the most potent modeling systems emanated from a deep delve into neurobiology. A number of nodes in different layers reminiscent of neurons in the nervous system of living organisms concomitantly work as a processing unit to solve complex problems. If the training data fulfill specific characteristics, the ANN shows promise in learning the underlying rules of the database. This justifies its early deployment in medical research and clinic as a reasonable prediction and classification modality, as it does not require precise diagnostic features of the medical condition to draw a lucid picture of it (20).

ANN systems have enabled the research paradigm to gain a novel understanding. However, each technique comes with pros and cons, leaving grounds for further debates. The main drawbacks of this method are falling to local minima, overfitting, and loss function minimization (21). To sidestep the matters, there is a growing consensus on optimization algorithm employment. Among others, population-based metaheuristics, which are conceptualized by profound insight into biologic systems, have taken center stage in modeling techniques to improve the targets (22). In this vein, the backtracking search optimization algorithm (BSA) (23), the vortex search algorithm (VSA) (24), and the grasshopper optimization algorithm (GOA) (25) are utilized in the current research to improve the efficacy of the neural network. These EAs are stochastic search algorithms with significant promise for resolving pragmatic nonlinear complex numerical optimization problems. As a recently proposed EA, BSA is built on iterative processing aiming at objective function minimization. It forms trial populations and orchestrates the level of the search-direction matrix and search-space limits, thus, gaining experience from previous generations and outstanding applicability in neural networks (26). Consistently, GOA is a swarm intelligence approach emanated from inspecting grasshoppers' collective forage and mass movement behavior in the ecosystem (25). VSA is also a newly introduced metaheuristic algorithm originating from the
vortex flow and the shearing process in fluid dynamics (27). The rationale for selecting these algorithms lies in their plain coding and improved global optimum analysis, contributing to the timely and efficient solving of complex problems. Given their high adaption to various numerical optimization systems, a variety of clinical and engineering applications have been suggested in the literature (8).

The intersection of clinics and computational sciences ushered in a new era of endeavors to portray the previously unamenable medical diagnostics and therapeutic approaches via ML modeling. Current research attempts to portray the BMT prognosis in children with hematologic disorders by exploiting hybrid ANN-BSA, ANN-VSA, and ANN-GOA algorithms. Bone marrow transplant: children Data Set, a freely available dataset in the UCI storehouse, is harnessed to design the hybrid multilayer perceptron (MLP) ANN-BSA, -VSA, -GOA systems (28).

The document’s outline follows: we further discuss the ANN and hybridization methods. In section 2, we provide a thorough literature review regarding the medical application of machine learning techniques. The database details are described in detail in section 3. Section 4 demonstrates the methodology in the current manuscript, explaining the ANN, the hybrid model, and the BSA, VSA, and GOA algorithms. Section 5 presents the experimental setup and the result analysis, discussing the proposed method’s effectiveness. Section 6 summarizes the study.

1.1. The artificial neural network

ANN emulates the neuronal system in the human brain, hypothesizing it as a complex nonlinear, parallel input processing system. Various factors are alike between the ANN and the biological central nervous system (CNS). These networks harness a substantially adaptable framework with altering parameters to provide specific duties or adjust to new databases (29). Inter-nodal connections are a remarked example, where the node or processing elements (PEs) are denoted as a first-level processing unit, similar to neurons in the CNS. Single nodes equal to neurons function to modify inputs to limited extents primarily. However, the interchangeable interaction of a broad range of neurons (i.e., nodal network) and not merely limited data exchange in a single neuron level contributes to accrued data processing function in CNS and ANN (30). The latter is defined as a training process. Each dataset is categorized into a sizable amount of training data (frequently 70–80% of the dataset), used to improve the network’s performance, and a 20–30% substitute of testing data. The training process harnesses the training input and its associated outputs. PEs represent a specific machine intelligence called a "memory with extended distribution (31)." To shed light, PEs encounter different inputs with various associated weights. If the cumulative weights attributed to a particular set of signals reach a tolerance point in the PE, the PE transfers a single output. During the training process, if a specific PE shows outcomes in close proximity to the target outcome, the PE’s associated connections and weights get increased, promoting the effect of such fine-tuned PE (32). In fact, the modeling system appraises the outputs in each layer by comparing the model’s predicted outcome and the target outcomes in the training dataset, thus, identifying the error-containing neurons and lowering their attributed weights to fine-tune the system (33). Learning is regarded as the most impactful engagement of ANNs in the ML age, as they show specific characteristics that help in various applications, such as pattern recognition, noble approximators, system identification, target estimation, and reversed modeling. Collectively, the ANNs field is perceived as a solid bet in the computational intelligence era as they recognize patterns and predict algorithms in the given information without knowing concrete links between the inputs and outputs, providing different tasks in the scope of decision-making (11).
1.2. Hybridization

Despite the great promise of ANN to provide valid prediction systems in the MATLAB environment, time-consuming learning and entrapment in local minima are among the significant drawbacks of this approach. In the context of ANNs, the level of inefficient performance is denoted the loss function. Hence, reaching a minimal loss function enhances the model's performance (34). Optimizers by fine-tuning the algorithm's features, such as learning rate and weights, play a progressively prominent role in the case of training the network for loss reduction. To further clarify the concept, one might imagine a blindfolded hiker attempting to descend from a mountaintop. The hiker will not be able to ride to any particular path; however, they will know whether they are descending downhill (progress) or ascending uphill (loss). Cumulatively, through multiple downward moves, they reach the base.

Similarly, at the beginning of the network function, the most efficient attributes of the model, such as optimum weights and learning rates, are not clear-cut. In the ANN, increasing the number of hidden layers, nodes, and activation function complexity lays a foundation for more reliable outcomes, despite requiring more time-taking processing (35). Thus, there is still a hot debate on the proper adjustment of ANN variables, including the weights, nodes, bias, and shape-changing architecture. Commonly, trial error depending on the loss function reduction leads to finding proper models (progress) (36). However, it is a debatable and time-consuming process.

In this realm, a plethora of optimizing techniques have been introduced to the machine learning literature in the past decade, contributing to fine-tuning the net weights and learning rates. These optimizer algorithms have sufficiently shown superiority over the formerly used Levenberg-Marquardt (LM) function in determining optimal values of model parameters (37). These methods are channeled into two significant subtypes: deterministic global optimization, which drives its impact through robust analytics, utilized commonly in complex global mathematical problems such as black-box modeling and related optimization. On the other page, metaheuristics, defined as a set of search-guidance methods, are exploited to analyze the function's domain to fine-tune the solutions promptly. Metaheuristic algorithms encompass many strategies, including simple local exploration and complex learning systems. These optimization systems have shown substantial promise in developing optimized results, by modulation of nodes, weights, bias, network structure, and activation and learning algorithm, particularly within defective datasets or in the face of restricted computation potencies (38).

The engagement of metaheuristics machinery to the ANN has exhibited great feasibility in terms of enhanced prediction models by concomitant exploration and data trading. Overall, ANN-based optimization algorithms constitute hundreds of techniques originated in essence by inspecting random populations. Mainly, metaheuristic optimizers are commonly categorized into biology- and physical-based algorithms (39). The biological-based class encompasses particle swarm optimization (PSO), bee colony algorithm (BCA), ant colony optimization (ACO), backtracking search algorithm (BSA), etc. The latter class comprises simulated annealing (SA), gravitational search algorithm (GSA), and chaotic optimization algorithm (COA) (39). Notably, the pragmatic application of these methods remains a daunting task because there is a dearth of firm data concerning selecting specific optimization systems to run with different optimization problems. Recent Investigations have shed a bright light on this, suggesting a robust association with the increased information regarding the target function's characteristics and more precise optimization (38).

2 Literature Review
Medical application of machine learning strategies has been on the rise. Gemini, an AI simulation application from GNS Healthcare, models the course of multiple myeloma with emphasis on medication's effect, paving the way for more personalized drug design by gathering in-vivo patient responses (40). AI-based medical diagnosis support programs are harnessed to analyze big data to recognize certain clinical conditions, opt for appropriate future treatment strategies, control problems, and promote patient care. In this context, the clinical decision support system (CDSS) has become a viable option to aid medical professionals in enhancing the accuracy of their diagnosis and ordering more effective therapies (41). There has been an upsurge in using ML for improved medical diagnosis. A branch of ML, including the convolutional neural networks, are powerful tools to recognize changes in radiographic images, with prominent accuracies similar to expert radiologists. For instance, IDx-DR is an AI diagnostic system that automatically detects diabetic retinopathy and macular edema by assessing retinal images, which was approved by the US food and drug administration (FDA) in 2018 (42). Consistently, clinical endeavors to forecast the outcomes of particular medical hypotheses yield unpromising results by virtue of inadequate pathophysiological understanding of the condition. To overcome these shortcomings, bioinformatics models are harnessed to design research quest models and optimize them to achieve the proper configuration and estimate the best-fit outcomes. Besides, ML tools leverage the accuracy of forecasting a particular patient’s prognosis (43). For instance, in the context of cancer prognostics, Jiang and colleagues designed an ML-based model to estimate the prognosis of patients with stage III colorectal cancer, presenting vital data to help plan individualized optimal therapeutic strategies (44).

3 Established Database

The present research harnessed an online database from https://archive.ics.uci.edu/ml/datasets/Bone+marrow+transplant%3A+children#. This database aimed primarily at appraising the determinants involved in the post-HSCT prognosis in children with hematologic pathologies (45). To clarify, previous experimental research conducted by Kalwak et al. recorded a statistically significant positive correlation between the concentration of CD34\(^+\) cells and increased post-transplant survival rates (46). This underscores the immediate necessity to investigate systematically the determinants affecting the transplant prognosis. The dataset we used evaluated 37 attributes and two outcome measures in 187 pediatric patients. The attributes comprise gender (male = 1, female = 0) and gender matching (compatible female to male = 1, other = 0), age of the (< 35 = 0, > 35 = 1) and recipient (< 10 = 0, > 10 = 1) and stem cell source (peripheral blood = 1, marrow = 0), development of acute graft versus host disease (aGVHD) stage > 2 (yes = 1, no = 0), donor and recipient ABO (O = 0, A = 1, B = -1, AB = 2) and ABO compatibility (matched = 1, mismatched = 0) and recipient Rh (+ = 1, - = 0). As well as, CMV status in donor and recipient (presence = 1, absence = 0), the second transplant post relapse (no = 0, yes = 1), risk group (high = 1, low = 0), disease group (malignant = 1, nonmalignant = 0), HLA match (matched = 0, mismatched = 1) and the number of compatible HLA antigen and alleles, disease recurrence (no = 0, yes = 1). It also includes CD34\(^+\) and CD3\(^+\) cell dose/kg recipient body weight, CD3\(^+\)/CD34\(^+\) ratio, BMI, time for absolute neutrophil count (ANC) and platelet level recovery (< 30 = 1, < 60 = 2, < 90 = 3, < 120 = 4, < 150 = 5, < 180 = 6, > 180 = 7). The output targets of the database were established as survival time (numerical) and survival status. Health records related to the patients hospitalized in the hemato-oncology department at Wroclaw Medical University, Poland, for ten years from 2000, were organized in the selected dataset. The records comprise 112 males and 75 females, with the prevalence of malignant and nonmalignant disorders, 82% and 18%, respectively. The median age at the time of BMT was 9.7 years (ranging from 0.7 to 19.2). Among these, 84 patients disclosed +1 year survival, and 35 individuals showed a +5 year survival.
4 Methodology

4.1 Methodology

The three tasks stated earlier in this section must be completed using maps, modeling methodologies, model validation, and optimization algorithm analysis (Fig. 1). Below are comprehensive descriptions of each step.

4.2 Artificial Neural Network

The ANN model has shown promise with images, text, and data tables. Providing the learning weights and nonlinear functions, ANN can predict outputs from a set of input data. The nonlinearity of the activation functions contributes to ANN's advantage in learning complex and non-established links between input and output data, denoted a universal approximation. This justifies the broad spectrum of ANN application in the scientific community (47, 48).

Figure 2 illustrates the architecture of the ANN. Activation functions calculate all input weights associated with a specific PE output, and each bias weight is a sum of all neuron inputs. Back-propagation (BP) is harnessed to adjust the model's weights to acquire gradients. The gradient could entirely disappear or explode during BP (49).

4.3 Hybrid model development

The optimization process comprises multiple tasks. Given the substantial impact of the optimized ANN algorithm on the accuracy of the results, presetting the ANN model could be adjusted to improve the hybrid model. Deep learning algorithms are preprocessed in different stages, including employing the hyperbolic tangent activation system (50). The Tansing function provides the cumulative output of a particular network's layer based on the overall weighted values in that layer. The fine-tuning of the input size, number of epochs, and objective function plays a critical role before optimal outcomes' appraisal when a stop criterion is met. The present paper considers the mean square error (MSE) as the objective function, providing a valuable measure to evaluate the solutions (51). The lower MSE values indicate the more proximity of the outcomes to the line of the best fit (regression line), in tune with more accurate predictions. The MSE equation is shown below:

\[ \text{MSE} = \frac{1}{n} \sum (Y_i - y_i)^2 \]

Where \( n \) denotes the data points' quantity, \( Y_i \) and \( y_i \) are the observed and predicted values, respectively. Herein, we provide an exemplary comparison of the convergence curves of multiple optimization behavior of the models (MSEs) in different population sizes for the number of iterations to ensure the consistent complexity of the operations (Fig. 2).

We exploit population-based metaheuristics to optimize the proposed model. These algorithms optimize nonlinear, nondifferentiable, and multimodal numerical queries. The underlying concept of the BSA is the return of a tribe to the previously found enriched hunting zones on a random basis (26). BSA, a population-based EA, concentrates on an iterative process to reduce the MSE. It runs in five stages, including initialization, primarily selection, mutation, crossover, and secondary selection. In contrast to other EAs, the mutation is utilized in the BSA as one direction per target and a non-uniform crossover method, which levers its potency. In the same vein,
it exploits a memory to form a trial population built arbitrarily from former generations and to choreograph the search-direction matrix and search-space boundaries (52).

GOA is a swarm-based EA that originated from the natural foraging patterns of grasshoppers. Grasshopper's random behavior dynamics depend on a set of factors, including its present location, wind advection, gravity effect, global best position, and the correlative position of neighboring grasshoppers and their social interactions. These factors contribute to its great promise in avoiding entrapment in local minima. The standard GOA algorithm's source is provided in: http://www.alimirjalili.com/GOA.html. The different movement behaviors among adult and nymph grasshoppers are mirrored in the GOA's diversification and intensification concepts (25).

The vortex search algorithm (VSA) is one of the recently proposed metaheuristic algorithms based on the vortical flow of the stirred fluids. The search approach in the VSA is designed as a vortex phenomenon by employing a dynamic step-size adjustment plan, decreasing the search space's radius, to maintain a tuned exploring and exploitive model. The innermost circle indicates the algorithm's optimum solution. VSA uses the inverse incomplete gamma function to modify the radius in each epoch. In general, the framework of the VSA encompasses forming primary outcomes, constructing the candidate outcomes, changing towards the present outcomes, and the radius decreasing process. For further insights into the VSA, it is recommended to review Dogan and Olmez's paper (24).

5 Results And Discussion

The present research is conducted in MATLAB 2022a environment for network architecture assessments and simulations, which includes a CPU of i5 @4.1 GHz, a memory of 4 GB, and 64-bit OS. For the experimentation quest, the selected HSCT database, which contains 37 attributes and 182 data points, was divided into training and testing classes, which are assessed over many iterations. Training-to-testing data ratio was set to 80:20, 146 and 36 inputs were assigned for training and testing the network's performance, respectively. An array of network characterizations, such as the different number of layers, hidden layer neurons, and the number of training/test inputs and the associated functions, have been modulated and compared to develop the most accurate prediction system. The network's performance was meticulously appraised through the root mean square error (RMSE), highlighting the difference between actual and predicted outputs, sensitivity, specificity, prevalence, accuracy, and error rate.

Taking the RMSE indices into account for the accuracy of the classic ANN configuration, Table 1 unveils that across a range of changes in the number of hidden layer neurons, a feedforward BPN approach with six neurons in the hidden layer, which are the artificially designed inputs by the network presents the minimal level of RMSE, 0.456 and 0.450 for the training and test inputs, respectively (Table 1). The lower RMSEs reflect the higher prediction accuracies. This primary optimization process paves the way for a novel optimization system that is used to obtain optimal solutions.
Table 1
sensitivity analysis for the change in the number of neurons predicting the HSCT outcomes in children with hematologic disorders.

<table>
<thead>
<tr>
<th>number of neurons</th>
<th>Network results</th>
<th>Scoring</th>
<th>Total score</th>
<th>RANK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSEtotal</td>
<td>RMSEtrain</td>
<td>RMSEtest</td>
<td>RMSEtotal</td>
</tr>
<tr>
<td>1</td>
<td>1.299</td>
<td>1.477</td>
<td>1.355</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>1.630</td>
<td>1.668</td>
<td>1.641</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0.702</td>
<td>0.630</td>
<td>0.681</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>1.029</td>
<td>1.097</td>
<td>1.050</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>0.943</td>
<td>0.746</td>
<td>0.889</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>0.447</td>
<td>0.456</td>
<td>0.450</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>0.650</td>
<td>0.745</td>
<td>0.680</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>2.108</td>
<td>2.426</td>
<td>2.208</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0.476</td>
<td>0.455</td>
<td>0.470</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>1.587</td>
<td>1.575</td>
<td>1.584</td>
<td>3</td>
</tr>
</tbody>
</table>

The MSE's correlation with the number of epochs obtained from the BSAMLp, VSAMLp, and GOAMLp models for the prediction of the HSCT outcomes is depicted in Fig. 2. It implies that the optimum target is met with a population size of 500, 500, and 450 in BSAMLp, VSAMLp, and GOAMLp, respectively.

The receiver operating characteristic curve or ROC curve illustrates the classifier system's performance at different classification thresholds. ROC plots true positive rate (TPR) against false positive rate (FPR). ROC curve holds promise to evaluate the optimal hybrid model. Also, AUC, which stands for Area Under the ROC Curve, demonstrates the two-dimensional area below the ROC curve from (0,0) to (1, 1), providing firm data about the aggregate performance measure regardless of the discrimination thresholds. AUC, indeed, has gained prominence in providing analysis of the ROC curve's accuracy results. Because firstly, the AUC is scale-invariant, focusing on predictions' ranks as opposed to their absolute values, and secondly, due to its invariance toward different thresholds in assessing the predictions' accuracy. An increase in AUC correlates with the network's enhanced potency to classify the positive and negative classes.

Figures 3, 4, and 5 illustrate the AUC for different populations in BSA, VSA, and GOA models, respectively. The forecasted AUC accuracy levels for the BSAMLp, VSAMLp, and GOAMLp training and testing inputs for different population sizes 50, 100, 150, 200, 250, 300, 350, 400, 450, and 500 are provided in the mentioned figures. To further clarify, for the BSAMLp testing databases, the AUC for population size of 50, 100, 150, 200, 250, 300, 350, 400, 450 and 500 swarm size were obtained 0.880, 0.832, 0.787, 0.865, 0.743, 0.843, 0.824, 0.904, 0.945, and 0.934 respectively. The AUC of testing datasets for the VSAMLp and GOAMLp with 50, 100, 150, 200, 250, 300, 350, 400, 450, and 500 swarm sizes are as follows, 0.858, 0.889, 0.813, 0.886, 0.868, 0.873, 0.864, 0.872, 0.881, and 0.859 and 0.701, 0.682, 0.825, 0.902, 0.799, 0.885, 0.878, 0.843, 0.778, and 0.808, respectively.
The best-fit BSAML, VSAML, and GOAML hybrid models for estimating the survival rates following HSCT have a swarm size of 500, 100, and 300, according to Tables 2, 3, and 4’s training and testing results. Similarities between phase one and phase two results indicate that the processes are highly similar.

Table 2
Network results based on AUC statistical indices for various proposed BSAML swarm size

<table>
<thead>
<tr>
<th>Population size</th>
<th>Network AUC results</th>
<th>Scoring</th>
<th>Total score</th>
<th>RANK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Testing</td>
<td>Training</td>
<td>Testing</td>
</tr>
<tr>
<td>50</td>
<td>0.974</td>
<td>0.880</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>100</td>
<td>0.981</td>
<td>0.832</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>150</td>
<td>0.974</td>
<td>0.788</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>200</td>
<td>0.977</td>
<td>0.866</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>250</td>
<td>0.981</td>
<td>0.743</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>300</td>
<td>0.992</td>
<td>0.844</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>350</td>
<td>0.985</td>
<td>0.824</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>400</td>
<td>0.988</td>
<td>0.905</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>450</td>
<td>0.981</td>
<td>0.945</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>500</td>
<td>0.993</td>
<td>0.935</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 3
Network results based on AUC statistical indices for various proposed VSAML swarm size

<table>
<thead>
<tr>
<th>Population size</th>
<th>Network AUC results</th>
<th>Scoring</th>
<th>Total score</th>
<th>RANK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Testing</td>
<td>Training</td>
<td>Testing</td>
</tr>
<tr>
<td>50</td>
<td>0.999</td>
<td>0.858</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>100</td>
<td>1.000</td>
<td>0.889</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>150</td>
<td>1.000</td>
<td>0.814</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>200</td>
<td>1.000</td>
<td>0.887</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>250</td>
<td>1.000</td>
<td>0.868</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>300</td>
<td>1.000</td>
<td>0.874</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>350</td>
<td>1.000</td>
<td>0.865</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>400</td>
<td>1.000</td>
<td>0.872</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>450</td>
<td>1.000</td>
<td>0.882</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>500</td>
<td>1.000</td>
<td>0.859</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 4
Network results based on AUC statistical indices for various proposed GOAMLP swarm size

<table>
<thead>
<tr>
<th>Population size</th>
<th>Network AUC results</th>
<th>Scoring</th>
<th>Total score</th>
<th>RANK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Testing</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Training</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.593</td>
<td>0.702</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>100</td>
<td>0.930</td>
<td>0.682</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>150</td>
<td>0.992</td>
<td>0.826</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>200</td>
<td>0.993</td>
<td>0.902</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>250</td>
<td>0.995</td>
<td>0.799</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>300</td>
<td>1.000</td>
<td>0.885</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>350</td>
<td>0.999</td>
<td>0.879</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>400</td>
<td>0.999</td>
<td>0.844</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>450</td>
<td>1.000</td>
<td>0.779</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>500</td>
<td>1.000</td>
<td>0.809</td>
<td>8</td>
<td>5</td>
</tr>
</tbody>
</table>

Figures 6, 7, and 8 depict the error analysis and frequency of mean absolute error (MAE) for training and testing the best fit of the proposed BSAMLP, VSAMLP, and GOAMLP models, indicating the error frequency and values.

Overall, Table 4 demonstrates the optimized characteristics of each model's best performance implemented in our methodology.

Table 4
Comparison of the best structures of the proposed algorithms.

<table>
<thead>
<tr>
<th>Optimization algorithm</th>
<th>Swarm size</th>
<th>AUC Results</th>
<th>Total score</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Training</td>
<td>Testing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>score</td>
<td>score</td>
<td></td>
</tr>
<tr>
<td>BSA</td>
<td>500</td>
<td>0.993</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>VSA</td>
<td>100</td>
<td>1.000</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>GOA</td>
<td>300</td>
<td>1.000</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

6. Conclusion

Novel ANN-based systems were investigated to forecast survival rates. A significant downside of the traditional statistical methods detecting nonlinear relations is that the link between input and output should be manually expressed, while ANNs have the advantage of self-learning. As ANN algorithms are pronouncedly flexible, many approaches have been developed to enhance the proposed model's performance, including metaheuristics. We employed BSA, VSA, and GOA to improve the ANN's accuracy and compare these methods' efficacy in the survival rate estimation research field. The field data from bone marrow transplants and the children dataset
available on UCI were employed to develop models in the present paper. The presented results unveiled the higher accuracy of these models at predicting pediatric HSCT survival rates compared to conventionally used medical methods among healthcare professionals.

We utilized standard statistical performance assessment techniques to reach the best model. Learning and simulation present ROC and AUC illustrations, respectively. The ANN-BSA, ANN-VSA, and ANN-GOA models' optimal neurons are calculated based on the cumulative number of neurons in each hidden layer. Training of the model contributes to the high accuracy of estimating the survival rate of HSCT in children with hematologic disorders. In the present research, we showed that the ANN model with six neurons in the hidden layers outperforms the conventional ANN system, laying a foundation for the best-fit BSA, VSA, and GOA model with a population size of 500, 500, and 450, respectively. Therefore, our results encourage the adoption of these algorithms to model survival studies from medical data, and the current research might be of significant value for future research in the field. This study revealed the applicability of these algorithms as efficient optimizers for the mentioned objective. Therefore, the developed model can be used correctly in practice. The authors, however, believe that conducting comparative studies in the future might help detect the most suitable models.

**Declarations**

**Conflicts of interest**

The authors declare that they have no competing interests.

**Ethics approval**

Not applicable.

**Consent to participate**

The authors agreed to publish this manuscript in the present journal.

**Consent to publication**

Not applicable.

**Funding**

None.

**Availability of data and material**

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Code availability**

Not applicable.

**Authors' contributions**
NJ and HM performed the search, and interpreted the results. All authors drafted the manuscript, reviewed and edited the draft. HM provided methodology and software requirements. NJ and FF prepared the figures and tables. FF conceptualized and supervised this study.

Acknowledgments
None.

References


Figures

Figure 1

It depicts the ANN model's structure
Figure 2

The best-fit model for the (a) BSAML, (b) VSAML, and (c) GOAML
Figure 3

AUC results for training and testing dataset for different proposed BSAMLP structures

Figure 4

AUC results for training and testing dataset for different proposed VSAMLP structures
Figure 5

AUC results for training and testing datasets for different proposed GOAML structures
Figure 6

The error and frequency of MAE for the best fit BSAMLP proposed model
Figure 7

The error and frequency of MAE for the best fit VSAMLp proposed model
Figure 8

The error and frequency of MAE for the best fit GOAML model.