Sigma Journal of Engineering and Natural Sciences
Web page info: https://sigma.yildiz.edu.tr
DOI: 10.14744/sigma.2025.1970

Research Article

Multiclass anemia classification based on multivariate adaptive regression spline method: Developing a decision support system for doctors

Nagihan YAĞMUR[1,*], Idiris DAĞ[2]

[1]*Department of Computer Engineering, Kutahya Dumlupinar University, Kutahya, 43100, Türkiye*
[2]*Department of Computer Engineering, Eskisehir Osmangazi University, Eskisehir, 26040, Türkiye*

ARTICLE INFO

Article history
Received: 05 March 2024
Revised: 13 May 2024
Accepted: 15 October 2024

Keywords:
Anemia Classification; Decision-Support Systems; Machine Learning; Multivariate Adaptive Regression Spline Method

ABSTRACT

Artificial intelligence and machine learning have the potential to forecast emerging diseases by analyzing patient medical data, offering decision-support systems for healthcare practitioners. These approaches are pivotal in minimizing diagnostic inaccuracies for physicians and hold significant value for patients and healthcare organizations in scrutinizing medical records. Nevertheless, the effectiveness of algorithms fluctuates based on various attributes of datasets, fluctuations in the size of datasets, diverse parameters and their characteristics, and varying quantities of patient records. In this research, two distinct models have been suggested for categorizing five types of anemia diseases using the multivariate adaptive regression spline method: a 1st-degree multiclass multivariate adaptive regression spline model and a 2nd-degree multiclass multivariate adaptive regression spline model. These models were employed to classify entries into non-anemia, folate deficiency anemia, Hgb-anemia, B12 deficiency anemia, and iron deficiency anemia. When the results were analyzed, 98.30% accuracy, 92.91% precision, 92.76% recall, and 92.74% F1-score were obtained with the 2nd order multiclass multivariate adaptive regression spline model. The outcomes obtained aim to offer valuable insights to medical students and physicians engaged in addressing the anemia classification challenge. The application of these models holds significant promise in enhancing diagnostic procedures, minimizing error rates, and devising more efficient treatment strategies. At the same time, in addition to the studies using the same data set in the literature, it has been shown that the adaptive spline method also shows successful results.

Cite this article as: Yağmur N, Dağ I. Multiclass anemia classification based on multivariate adaptive regression spline method: Developing a decision support system for doctors. Sigma J Eng Nat Sci 2026;44(1):93−104.

***Corresponding author.**
*E-mail address: nagihan.yagmur@dpu.edu.tr

This paper was recommended for publication in revised form by Editor-in-Chief Ahmet Selim Dalkilic

Published by Yıldız Technical University Press, İstanbul, Turkey

INTRODUCTION

Thanks to advances in science and technology, diagnoses in the medical world have greatly benefited from the measurement of data such as patients' blood values, medical images, and medical signals [1]. Making these measurements in detail facilitates the detection of patterns and hidden connections in datasets thanks to machine learning methods, a sub-branch of artificial intelligence that can make connections between complex data. Artificial intelligence and machine learning methods lead to the rapid and accurate resolution and diagnosis of diseases [1–3].

Artificial intelligence and machine learning methods are often used in various studies such as heart failure disease prediction, Diabetes classification, and heart failure disease prediction, which are frequently researched in the literature [4-9]. These studies aim to train processes to speed up the decision-making process of specialist doctors when a new case comes in. That is, the aim is to ensure that when specialists are faced with new cases, the system can similarly make decisions for them.

Modeling data and disease is a major component of medicine today. One of the two processes that enhance this is the determination of model parameters by machine learning methods. The terms Data, Disease, and Modeling are used in the determination of model parameters and machine learning methods help this process to be performed faster and more accurately. In a study using deep learning architectures Alexnet and Resnet50 for automatic detection of lung cancer using computed tomography images, the performances of the model parameters were compared and the highest accuracy value was obtained with the AlexNet architecture [10]. In the study where the success of the k-means algorithm was improved with machine learning methods using various cancer data, the death game optimization algorithm was combined with the k-means method. As a result, the success of the proposed clustering method was increased [11]. In a study on brain tumor classification using two-dimensional MRI images, promising results were obtained for complex classification problems by optimizing the vision transformer (ViT) network [12]. A hybrid k-means method is presented to assign weights to input parameters in a study conducted to correctly diagnose diseases in four different medical datasets. It is emphasized that the results obtained will reduce staff workload and test costs. [13]. In a study to reduce the effects of COVID-19 on the education system, an IoT and deep learning-based system was proposed [14]. The proposed system aims to protect students and staff against infectious diseases and improve students' performance in classes by monitoring environmental conditions through an IoT-based sensor network to ensure the use of masks in closed areas during the current pandemic. In the study conducted on the Face Mask Detection Dataset containing a total of 7553 masked and unmasked images, the best results were obtained as 99.5% for F1 Score and 99% for MCC with the model trained on the InceptionV3 network. In a study using Machine Learning classification methods such as K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Decision Tree (DT) for diabetes prediction by taking different attributes related to diabetes disease, random forest achieved high success in predicting diabetes compared to other machine learning methods [15]. In a study of brain tumor detection using YOLOv5, YOLOv7, and YOLOv8, it was concluded that brain tumors were better detected with YOLOv8 [16]. The images were first preprocessed by applying color transformation, histogram equalization, and filter operations. As a result, a high success rate of 87% was achieved. In a study using Computed Tomography (CT) images for the detection of brain hemorrhages (ResNet), Visual Geometry Group (VGG), and EfficientNet architectures, it was emphasized that it is important to choose the most appropriate model for the dataset [17]. As a result, it was emphasized that the success of the EfficientNet-B2 model was higher than the other models in the literature.

Thanks to scientific and technological developments, artificial intelligence system applications used to support many areas are widely used. Thus, the health sector is now in need of experts in the evaluation of data due to technological advances. Doctors' busy working hours, difficulties in accessing various specialists, and the vital importance of early diagnosis of urgent pathologies contribute to the growing need for apps that support healthcare professionals.

The diagnosis of anemia is crucial for physicians, and anemia signifies a condition where oxygen is not adequately transported; in this case, red blood cells fail to fulfill their function completely. It occurs as a result of low hemoglobin, indicating values below 13 g/dl in men and 12 g/dl in women. Iron deficiency anemia is the most common form of anemia. It is caused by insufficient intake or reduced absorption of the iron the body needs. Most iron is found in red blood cells, and women can develop iron deficiency and anemia due to blood loss. The main causes of anemia are bleeding, destruction of red blood cells (hemolysis), insufficient production of red blood cells, and low hemoglobin production. Women are at higher risk of anemia due to menstrual bleeding [18]. Children may develop anemia due to stomach bleeding caused by the frequent use of aspirin and antipyretics [19]. Hemolytic anemia, B12 deficiency anemia, iron deficiency anemia, thalassemia, Fanconi anemia, etc. The main causes of anemia, which has varieties such as can be examined under 3 headings: 1. blood loss 2. insufficiency of red blood cells 3. high rate of red blood cell destruction [19].

The World Health Organization (WHO) has characterized anemia as a state where the red blood cells or their capability to carry oxygen are inadequate [16]. This condition has been highlighted to predominantly impact women and young children, leading to a considerable decline in quality of life and potential association with severe illnesses.

There are also studies on the classification of anemia disease, which aim to provide decision support to physicians. When the studies were examined, it was seen that they were divided into two groups as binary class (anemia / no anemia) studies and multi-class (iron deficiency anemia, B12 vitamin deficiency anemia, etc.). In an analysis of binary class studies, a research effort utilized MARS and the Harris hawks algorithm (HHO) to process data from 1732 patients categorized as anemic or non-anemic, sourced from the Kaggle database [20]. In the study, six different algorithms including multiple linear form HHO, multiple quadratic form HHO, multiple exponential form HHO, 1st order MARS model, 2nd order MARS model, and the strongest MARS model were proposed. At the end of the study, the performance evaluations of the six proposed methods were found to be higher than other studies in the literature. In a study to detect the anemia status of a patient, a model based on classification with support vector machine (SVM) and RF was developed [21]. In a study conducted using blood data from Kanti children's hospital, a model was designed to predict anemia in children under 5 years of age [22]. The dataset consisting of 700 records was first normalized and data balanced. As a result, the RF method demonstrated the most successful performance with an accuracy rate of 98.4%. Finally, it aims to increase the performance of the algorithms by using methods such as feature selection and Ensemble Learning methods, Voting, Stacking, Bagging, and Boosting. Machine learning methods such as CNN, k-NN, Decision Tree (DT), NB and SVM have been used to propose using palm images for anemia detection as an effective, cost- and time-saving solution as an alternative to invasive methods [23]. In a study focused on predicting anemia in children, various machine-learning methods were used to find factors associated with anemia [24]. While MLP showed the highest success with 81.67%, the highest success (82.5%) was achieved when the DT feature selection method was applied. In a study where 1732 blood data taken from the Kaggle database were classified using the Crow Search Optimization Algorithm (CSO), CSO was analyzed in 2 different scenarios. Multilinear form CSO and quadratic linear form CSO [25].

Again, in a study in which various machine learning algorithms such as DT, SVM, RF, LR and k-NN were tested to predict anemia in children, a powerful prediction model was created by improving classifiers with learning techniques such as bagging, boosting, and stacking [26]. In another study, the process of diagnosing anemia and anemia-related diseases was performed with machine-learning methods [27]. In this context, a machine learning architecture that can diagnose 12 different types of anemia was developed. The data were obtained from Düzce University Research and Application Hospital with ethics committee permission. Then, the weights of the data were calculated using techniques such as Information Gain, Information Gain Ratio, Correlation, and Principal Component Analysis. After the calculations, machine learning techniques such as Support Vector Machines, Decision Trees, Naive Bayes, and Artificial Neural Networks were applied to the original dataset and the datasets created using the new attributes. The results obtained by using cross-validation method in all models were evaluated with performance indicators such as accuracy, classification error, sensitivity, precision, and f1-score. In a study where anemia was predicted using various methods such as XGBoost, LightGBM, Lasso, Rigde and Elastic-net, the SMOTE technique, which attempts to balance the distribution of classes by randomly increasing and replicating minority class instances, and Pearson correlation to identify strongly correlated points that are thresholds for feature selection [28]. As a result, the Light GBM algorithm outperformed the others in terms of accuracy. Some recent studies have proposed a robust and efficient classification algorithm using SVM, k-NN, LR, DT, RF algorithm to overcome the weaknesses of manual evaluation and thus classify Sickle Cell Anemia (SCA) case into 3 different classes [29]. Also, in a different study to solve the 5-class anemia classification problem like this study, it was proposed to use fuzzy logic-based parameter optimization and achieve class-based and overall accuracy improvement within the dataset [30]. Six new models are introduced with the integration of TreeBagger, Chicken Swarm Optimization Algorithm (CSA), CSO and JAYA method.

In this study, 15.300 patient hematology data obtained from the study titled "Nutritional Anemia Disease Classification Based on Genetic Algorithm and Deep Learning Algorithms" by Kilicarslan et al. at Tokat Gaziosmanpaşa University Faculty of Medicine [31,32]. The dataset encompasses records of individuals with Hgb-anemia, iron deficiency anemia, B12 vitamin deficiency anemia, folate deficiency anemia, and non-anemia. The primary objective of this study is to develop a system useful in diagnosing anemia in general healthcare services, particularly in situations where access to specialist physicians is becoming challenging due to the rising number of patients and hospital density. It is anticipated that this study can make substantial contributions not only to non-specialized healthcare professionals but also to the existing literature.

It is possible to note that, the MARS method is a mathematical modeling approach preferred in general for analyzing complex data sets and often used in various studies such as prediction, analysis, and classification [20,33,34]. Different mathematical modeling approaches have been studied in detail by mathematicians in the literature [37–39]. The behavior of MARS can also be seen in many studies in the literature. When the studies examined, it is understood from the values of different performance criteria that this machine learning based method can be used to develop effective prediction models, especially for engineering datasets [20,40].

In this study, it is determined that the MARS method, which mathematically models complex data and is frequently used in prediction and classification problems, has not been used in the anemia classification problem as far as it is known and examined under two scenarios. 1st order multiclass MARS model and 2nd order multiclass MARS model. As a result, it aims to reveal complex models by examining the pruning parameters and degree values that significantly affect the performance of the MARS method.

MATERIALS AND METHODS

Preprocessing

The data set used in the study is Kılıçarslan et al. It includes the blood count results of 15300 patients taken from Tokat Gaziosmanpaşa University Faculty of Medicine [32]. Data of pregnant women, cancer patients, and children and irrelevant and missing parameters were published by Kılıçarslan et al. It was removed from the data set by. The dataset used in the study includes 24 parameters and 5 different classes. Due to numerical value differences in the data set, blood values were normalized by scaling them between 0.1 and 0.9.

MARS Method

MARS method has a nonlinear and nonparametric structure and it is a method introduced to the literature by Friedman (1991) [41]. With its ability to make reliable predictions, it finds itself used in many fields thanks to its success. Also, it produces strong predictions compared to other models and gives robust results. Therefore, despite the fact that the models include especially large number of terms, it is frequently preferred. MARS method is composed of 2 stages. And the first of these stages is the stage of "step-forward algorithm". In this stage, the aim is to maximize the basic functions under the forward step algorithm which forms complex model. And as it is a result of this, the models that include incorrect terms often come out. In this step, an iterative process is carried out with the aim of obtaining an optimal model by reducing complexity. Thanks to the iterative process, it is aimed to achieve a balance between accuracy and model complexity until the number of basic functions is maximum.

Thanks to the backward step algorithm, the complexity of the models produced in the forward step phase is reduced and the functions that contribute the least to the increase in total squares are determined and removed from the model. Improvement is achieved in this way and helps to expand and simplify the model. This process is called "pruning" and is managed by a hyperparameter called "nprune". As a result, the optimal prediction model is created with the pruning parameter [42]. In the MARS method, which provides good modeling in high dimensional and complex data, the degree of pruning affects this performance, as does the selection of interactions or functions to be extracted. There is another important hyperparameter in the MARS, where it creates piecewise basis functions. This is "degree". This parameter contributes an ability that has already learned by the model, to understand the different relationships. Therefore it can learn complex relationships. Reducing the degree, it may also capture simpler relationships. Choosing the right degree is very essential for the model that best fits with the data and learns, and for not facing with simplification problem. In short, degree and pruning parameters are very important for this method's performance.

The general MARS model is expressed as [43]

$$Y = \beta_0 + \sum_{k=1}^{K} a_k\, \beta_k(X_t) + \varepsilon_t \tag{1}$$

In this context, X represents the independent variable, while K denotes the number of basis functions, and k stands for the number of nodes. The coefficients of the kth basis function are represented by a_k, B_0 signifies the constant term, and $B_k(X_t)$ denotes the k. independent variable for the k. basis function. The basis function is expressed as $k=1,2,..K$ has the following form [43].

$$B_m(x) = \prod_{t=1}^{L_m} \left[S_{l,m} \left(x_{v(1,m)} - k_{l,m} \right) \right] \tag{2}$$

According to Equation 2, L_m is the degree of interaction, $S_{l,m} \in [\pm 1]$, $k_{l,m}$ is the node value, $x_{v(1,m)}$ is the argument value.

Adaptation of the Mars Method to Multiple Anemia Classification Problem

Table 1. Types of models utilizing the MARS method

Models	Model characteristics
MARS_model-1	*Degree = 1 nprune = 20*
MARS_model-2	*Degree = 1 nprune = 20*

In order to evaluate the performance of the MARS method, analyzes were carried out on 2 separate models using different combination of hyperparameters. These analyzes were carried out on very different set of hyperparameters. The aim of this analyzes is to find out the prediction performance of the model and / or find out the combination of hyperparameters when the model is at peak level. This really creates a comprehensive understanding of the MARS method applied to the data – how it hinges on making elaborate predictions and still maintains reasonable ability to generalize. Experimentally, MARS_Model-1 was configured with degree 1 and nprune 20 and MARS_Model-2 was configured with degree 2 and NPRUNE 20.

For MARS_Model-1, a 10-fold cross-validation method was employed, and the accuracy values along with the average accuracy obtained at each fold are presented in Table 2.

Table 2. Accuray values obtained as a result of each fold by MARS_model-1.

Resample	Accuracy
Fold-1	%97.98
Fold-2	%97.97
Fold-3	%98.50
Fold-4	%97.91
Fold-5	%98.43
Fold-6	%97.84
Fold-7	%97.65
Fold-8	%97.71
Fold-9	%98.24
Fold-10	%97.97
Avg Acc =	%98.02

The high accuracy values signify that the model is robust even in the presence of an imbalanced dataset.

Prior to pruning, 7 basis functions with low success rates were produced and subsequently eliminated from MARS_Model-1. Subsequently, the 20 basis functions generated through the MARS method for MARS_Model-1, along with the corresponding weights, are detailed in Table 3.

As indicated in Table 3, j represents the number of basis functions, and the BFj values denote the basis functions defined for the model. Based on the details provided in Table 3, the mathematical equation for MARS_Model-1 can be expressed as follows:

$$\begin{aligned} Model-1_{class1} = &-125.71 - 363.02xBF1 - 344.7xBF2 \\ &- 238023.15xBF3 - 408104.61xBF4 - 173907xBF5 \\ &+ 71438.5xBF6 - 71404.2xBF7 + 112.9xBF8 \\ &+ 4258.14xBF9 - 37.93xBF10 + 345.57xBF11 \\ &- 46.38xBF12 + 175.68xBF13 - 498.41xBF14 \\ &- 317.83xBF15 + 435.40xBF16 - 562.54xBF17 \\ &+ 2909.55xBF18 + 308.51xBF19 \end{aligned} \tag{3}$$

$$\begin{aligned} Model-1_{class2} = &-4.3 + 3.37xBF1 + 10.01xBF2 \\ &+ 270.13xBF3 - 863.32xBF4 + 576.51xBF5 \\ &- 916.55xBF6 - 12198.40xBF7 - 7.69xBF8 \\ &+ 14577.59xBF9 + 30.70xBF10 - 4.91xBF11 \\ &- 40.22xBF12 - 16.44xBF13 - 14.40xBF14 \\ &- 112.42xBF15 - 110.40xBF16 - 276.26xBF17 \\ &- 3255.93xBF18 + 14.83xBF19 \end{aligned} \tag{4}$$

Table 3. The prediction results of MARS_model-1

Basic Functions		Weights				
		1	2	3	4	5
	Constant	-125.71	-4.3	3.27	-15.13	-6.5
BF1	*GENDER*	-363.02	3.37	4.2	1.79	2.1
BF2	*max (0,0.51 – RBC)*	-344.7	10.01	-18.46	8.70	4.3
BF3	*max (0, HBG – 0.45)*	-238023.15	270.13	637.28	332.25	376.8
BF4	*max (0, HBG – 0.46)*	408104.61	-863.32	-1991.35	-1081.53	-907.4
BF5	*max (0, HBG – 0.47)*	-173907.0	576.51	1375.29	775.40	544.1
BF6	*max (0, HBG – 0.51)*	71438.5	-916.55	-644.43	-814.26	-416.6
BF7	*max (0, HBG – 0.52)*	-71404.2	-12198.40	-20551.91	-6425.84	-27267.2
BF8	*max (0,0.57 – HBG)*	112.9	-7.69	11.09	-9.82	-2.1
BF9	*max (0, HBG – 0.57)*	4258.14	14577.59	23525.97	7988.08	31474.9
BF10	*max (0,0.07 – MCH)*	-37.93	30.70	-37.38	-57.23	-33.7
BF11	*max (0, MCH – 0.07)*	345.57	-4.91	4.41	-29.11	6.1
BF12	*max (0,0.11 – SD)*	-46.38	-40.22	11.97	-15.25	-22.8
BF13	*max (0, SD – 0.11)*	175.68	-16.44	42.29	-8.98	-9.7
BF14	*max (0, SD – 0.29)*	-498.41	14.40	-27.40	10.86	10
BF15	*max (0, TSD – 0.16)*	-317.83	112.42	-147.25	174.98	77.9
BF16	*max (0, TSD – 0.23)*	435.40	-110.40	130.38	-174.64	-77.1
BF17	*max (0,0.08 – FOLATE)*	-562.54	-276.26	-5.99	238.40	-174.7
BF18	*max (0,0.006 – B12)*	2909.55	-3255.93	-61.46	251.27	784.4
BF19	*max (0, B12 – 0.006)*	308.51	14.83	-11.04	16.43	-4127036.3

$$\begin{aligned} Model-1_{class3} = {} & 3.27 + 4.2xBF1 - 18.46xBF2 \\ & + 637.28xBF3 - 1991.35xBF4 + 1375.29xBF5 \\ & - 644.43xBF6 - 20551.91xBF7 - 11.09xBF8 \\ & + 23525.97xBF9 - 37.38xBF10 + 4.41xBF11 \\ & + 11.97xBF12 + 42.29xBF13 - 27.40xBF14 \\ & - 147.25xBF15 + 130.38xBF16 - 5.99xBF17 \\ & - 61.46xBF18 - 11.04xBF19 \end{aligned} \tag{5}$$

$$\begin{aligned} Model-1_{class4} = {} & -15.13 + 1.79xBF1 + 8.70xBF2 \\ & + 332.25xBF3 - 1081.53xBF4 + 775.40xBF5 \\ & - 814.26xBF6 - 6425.84xBF7 - 9.82xBF8 \\ & + 7988.08xBF9 - 57.23xBF10 - 29.11xBF11 \\ & - 15.25xBF12 - 8.98xBF13 + 10.86xBF14 \\ & + 174.98xBF15 - 174.64xBF16 + 238.40xBF17 \\ & + 251.27xBF18 + 16.43xBF19 \end{aligned} \tag{6}$$

$$\begin{aligned} Model-1_{class5} = {} & -6.5 + 2.1xBF1 + 4.3xBF2 + 376.8xBF3 \\ & - 907.4xBF4 + 544.1xBF5 - 416.6xBF6 - 27267.2xBF7 \\ & - 2.1xBF8 + 31474.9xBF9 - 33.7xBF10 + 6.1xBF11 \\ & - 22.8xBF12 - 9.7xBF13 + 10xBF14 + 77.9xBF15 \\ & - 77.1xBF16 - 174.7xBF17 + 784.4xBF18 \\ & - 4127036.3xBF19 \end{aligned} \tag{7}$$

According to the information provided in Table 3 and Equations 3-7, the MARS method pruned and eliminated 16 out of the 24 parameters from the model as they made minimal contributions to the model's performance. Consequently, the model was constructed with parameters such as GENDER, RBC, HBG, MCH, SD, TSD, FOLATE, and B12, which have a high contribution to the model performance. Upon reviewing Equations 3-7, it is evident that with a degree of 1, the formation of simple linear functions occurs, without any interaction among the basis functions. As stated in the equations, the dataset was subjected to class-based modeling with the MARS method, and the prediction result with a high success rate was taken as a reference. Classification was performed in this way.

Table 4. Accuray values obtained as a result of each fold by MARS_model-2.

Resample	Accuracy
Fold-1	%98.17
Fold-2	%97.97
Fold-3	%98.36
Fold-4	%97.84
Fold-5	%98.63
Fold-6	%98.24
Fold-7	%98.24
Fold-8	%98.10
Fold-9	%98.69
Fold-10	%98.23
Avg Acc =	%98.25

For MARS_Model-2, 10-fold cross-validation was applied and the accuracy and average accuracy values obtained at each folding step are presented in Table 4. High accuracy values indicate that the model is not affected by the unbalanced data set.

Before the pruning process for MARS_Model-2, a collection of 32 basis functions was initially created, following which 12 functions with lower success rates were omitted from the model. Subsequently, Table 5 presents the 20 basis functions derived using the MARS method for MARS_Model-2, inclusive of their corresponding weights.

$$\begin{aligned} Model-2_{class1} = {} & -980.82 + 114.02xBF1 + 101764.59xBF2 \\ & - 134629.02xBF3 + 34290.32xBF4 + 1682.25xBF5 \\ & - 103544.43xBF6 + 91825.56xBF7 + 20620.6xBF8 \\ & - 52302.795xBF9 + 41493.423xBF10 + 8.29xBF11 \\ & - 31.94xBF12 + 37.45xBF13 - 70.35xBF14 \\ & - 141.35xBF15 + 718.07xBF16 - 645.69xBF17 \\ & + 859.32xBF18 + 5601.58xBF19 \end{aligned} \tag{8}$$

$$\begin{aligned} Model-2_{class2} = {} & -0.64 + 1.18xBF1 - 227.76xBF2 \\ & - 173.04xBF3 + 319.64xBF4 + 242251.42xBF5 \\ & - 12420.24xBF6 - 1776.51xBF7 - 32.58xBF8 \\ & + 182.65xBF9 - 54.8xBF10 - 28.69xBF11 \\ & + 145.03xBF12 + 1440497.07xBF13 \\ & - 1440545.22xBF14 + 194.93xBF15 - 1330.09xBF16 \\ & - 1439373.27xBF17 - 1358.75xBF18 - 20917.73xBF19 \end{aligned} \tag{9}$$

$$\begin{aligned} Model-2_{class3} = {} & +1.13 - 3.15xBF1 - 594.10xBF2 \\ & + 362.84xBF3 + 10.65xBF4 - 12.37xBF5 \\ & - 3779.46xBF6 - 2504.97xBF7 - 134.71xBF8 \\ & + 397.15xBF9 + 4.44xBF10 + 80.66xBF11 \\ & - 542.99xBF12 - 176.54xBF13 + 210.18xBF14 \\ & - 363.42xBF15 + 2035.7xBF16 - 1487.89xBF17 \\ & - 233.22xBF18 + 96.33xBF19 \end{aligned} \tag{10}$$

$$\begin{aligned} Model-2_{class4} = {} & +0.53 + 0.77xBF1 - 762.61xBF2 \\ & + 847.54xBF3 - 399.29xBF4 + 15652.09xBF5 \\ & - 17315.68xBF6 - 869.46xBF7 - 137.27xBF8 \\ & + 383.62xBF9 + 103.92xBF10 + 176.29xBF11 \\ & - 3008.09xBF12 + 94423.9xBF13 - 94344.426xBF14 \\ & - 175.84xBF15 - 5330.8xBF16 - 88906.14xBF17 \\ & + 3485.53xBF18 - 256.41xBF19 \end{aligned} \tag{11}$$

$$\begin{aligned} Model-2_{class5} = {} & -1.71 + 0.41xBF1 - 271.61xBF2 \\ & + 139.7xBF3 + 4.3xBF4 + 261731.73xBF5 \\ & - 24207.21xBF6 - 690.03xBF7 - 27.48xBF8 \\ & + 145.52xBF9 + 36.53xBF10 - 10.48xBF11 \\ & - 19.02xBF12 + 15566.1xBF13 - 1556574.2xBF14 \\ & + 183.53xBF15 - 1116.8xBF16 - 15555652.35xBF17 \\ & - 1838.8xBF18 + 14763.27xBF19 \end{aligned} \tag{12}$$

Table 5. The prediction results of MARS_model-2

Basic Function		Weights				
		1	2	3	4	5
	Constant	-980.82	-0.64	1.13	0.53	-1.71
*BF*1	*GENDER*	114.02	1.18	-3.15	0.77	0.41
*BF*2	*max* (0, *HBG* – 0.45)	101764.59	-227.76	-594.10	-762.61	-271.61
*BF*3	*max* (0, *HBG* – 0.47)	-134629.02	-173.04	362.8	847.54	139.7
*BF*4	*max* (0, *HBG* – 0.49)	34290.32	319.64	10.65	-399.29	4.3
*BF*5	*max* (0,0.57 – *HBG*)	1682.25	242251.42	-12.37	15652.09	261731.73
*BF*6	*GENDER** *max* (0, *HBG* – 0.52)	-103544.43	-12420.24	-3779.46	-17315.68	-24207.21
*BF*7	*GENDER** *max* (0, *HBG* – 0.51)	91825.56	-1776.51	-2504.97	-869.46	-690.03
*BF*8	*GENDER** *max* (0, *HBG* – 0.4)	20620.6	-32.58	-134.71	-137.27	-27.48
*BF*9	*GENDER** *max* (0, *HBG* – 0.43)	-52302.795	182.65	397.15	383.62	145.52
*BF*10	*GENDER** *max* (0, *HBG* – 0.48)	41493.423	-54.8	4.44	103.92	36.53
*BF*11	*GENDER** *max* (0,0.24 – *TSD*)	8.29	-28.69	80.66	176.29	-10.48
*BF*12	*max* (0, *HBG* – 0.44)* *max* (0,0.24 – *TSD*)	-31.94	145.03	-542.99	-3008.09	-19.02
*BF*13	*max* (0,0.57 – *HBG*)* *max* (0, *TSD* – 0.35)	37.45	1440497.07	-176.54	94423.9	1556592.1
*BF*14	*max* (0,0.57 – *HBG*)* *max* (0,0.35 – *TSD*)	-70.35	-1440545.22	210.18	-94344.426	-1556574.25
*BF*15	*max* (0,0.57 – *HBG*)* *max* (0, *TSD* – 0.25)	-141.35	194.93	-363.42	-175.84	183.53
*BF*16	*max* (0,0.57 – *HBG*)* *max* (0, *TSD* – 0.21)	718.07	-1330.09	2035.7	-5330.8	-1116.8
*BF*17	*max* (0,0.57 – *HBG*)* *max* (0, *TSD* – 0.18)	-645.69	-1439373.27	-1487.89	-88906.14	-15555652.35
*BF*18	*max* (0,0.57 – *HBG*)* *max* (0,0.83 – *FOLAT*)	859.32	-1358.75	-233.22	3485.53	-1838.8
*BF*19	*max* (0,0.57 – *HBG*)* *max* (0,0.006 – *B*12)	5601.58	-20917.73	96.33	-256.41	14763.27

Upon examination of Table 5 and Equations 8-12, it is evident that the highest interaction level of the basis functions is 2, established by setting the degree to 2. This signifies that the model's basis functions consist of both linear and quadratic functions, as depicted in Table 5.

With the MARS method, it was observed that *BF1*, *BF2*, *BF3*, *BF4*, and *BF5* represent linear functions, while the remaining basis functions are quadratic functions resulting from the product of two basis functions. Additionally, 19 out of 24 parameters were pruned and eliminated from the model due to their minimal contribution to the model's performance, resulting in the model being created with GENDER, HBG, TSD, FOLATE, and B12 parameters, which exhibited high contribution to the model performance.

Upon reviewing Equations 8-12, it is evident that the dataset is modeled based on a class basis using the MARS method, with the prediction result showing a high success rate serving as the benchmark for classification. Classification was carried out in this manner. Consequently, when evaluating new patient data using Equations 3-12, classification can be accomplished by inputting the required information into the *BFj* basis functions.

Evaluation

In this study, a 10-fold cross-validation method is used to evaluate the performance of the proposed models for classifying the five-class anemli problem. The strategy of dividing the dataset into 10 different subsets and testing one subset at a time involves using the remaining subsets for training purposes. In this way, problems due to class imbalance were minimized. Performances were measured with ROC analysis metrics. ROC analysis is a performance evaluation metric. The results obtained showed that the proposed models were effective in correctly classifying the anemia class.

In equations 13-16, macro metric formulas representing unweighted average calculations for each class are presented.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

$$Recall_m = \frac{TP_c}{TP_c + FN_c} \tag{14}$$

$$Precision_m = \sum_c \frac{TP_c}{TP_c + FP_c}, c \in classes \tag{15}$$

$$F1 - Score_m = \frac{2xPrecision_m xRecall_m}{Precision_m + Recall_m} \tag{16}$$

Within the equations, when considering the macro metric m and the *classes = {non-disease records, HGB-anemia, iron deficiency anemia, B12 deficiency anemia, folic acid deficiency anemia}*, the term TP_c denotes the count of samples accurately categorized as c.

The objective of ROC analysis is to assess the efficacy of outcomes derived from different methods and to compare them using metrics like Accuracy, Specificity, Sensitivity, F1-Score, Precision, and AUC [44]. The fundamental ROC parameters utilized in the analysis, namely *FN (False Negative)*, *TN (True Negative)*, *FP (False Positive)*, and *TP (True Positive)* serve to indicate the accuracy of the classification results through true and false predictions.

To evaluate model performance, it is important to select appropriate metrics such as Accuracy, AUC and F1-Score. Accuracy measures the proportion of correct predictions, while AUC evaluates the balance between True Positive Rate (TPR) and False Positive Rate (FPR) at different thresholds, targeting high TPR and low FPR (Yagmur et al., 2024). In unbalanced datasets, relying on Accuracy alone is insufficient. Sensitivity and recall-based F1-Score are commonly used for imbalanced datasets. A high F1-Score indicates the minimal difference between two values. In particular, a high recall implies low sensitivity and vice versa. Moreover, a high recall value implies that a significant proportion of samples in the minority class are correctly predicted, while high precision indicates that the predicted minority samples most likely belong to the minority class. Increasing precision and recall values lead to a higher F1-Score, indicative of the difference between minimal precision and recall values.

Although AUC is widely used to address imbalanced data issues, the F-measure is considered more appropriate for imbalanced data sets. This is because the minority class usually has a higher weight than the majority class and the F-measure is capable of classifying samples in the minority class with better accuracy and less misclassification. The F-measure focuses only on the accuracy of the classifier on the minority class. The AUC measure evaluates the overall accuracy of the classifier on both the majority and minority classes. Given the imbalanced nature of the dataset used in this study, the next section compares the performance of AUC and F1-Score. However, other studies in the literature often prioritize accuracy metrics, so additional ROC metrics are also presented.

RESULTS AND DISCUSSION

This study encompassed the execution of a classification analysis on the blood metrics and results of patients diagnosed with anemia. Patients were sorted using the MARS method, with the effectiveness of the method evaluated across two separate models.

For MARS_Model-1, 1st order models with a pruning coefficient of 20 and for MARS_Model-2, 2nd order models with a pruning coefficient of 20 were used.

In this part, we display ROC performance analyses and confusion matrices and obtained for two separate models. Figure 1, 2 exhibits the confusion matrices for MARS_Model-1 and MARS_Model-2, alongside the computed ROC performance values derived from these matrices.

These analyses visually present the classification performance of each model and evaluate the sensitivity, specificity and performance of the algorithms through the ROC curves given in Figure 3, 4. The results obtained reveal that Model-1 and Model-2 have successful classification performance. These graphs allow us to evaluate in more detail the ability of each algorithm to detect the disease.

The evaluation of confusion matrices in Figure 1 and Figure 2 reveals that, as a result, 282 patients were misclassified by Model-1, and 260 patients were misclassified by Model-2 out of the total cases.

Table 6 encompasses ROC performance analyses obtained from testing two different models. This table provides a detailed overview of the sensitivity, specificity, and overall performance of each model through ROC curves. ROC performance analyses assess the ability of Model-1 and Model-2 to detect the disease, offering a comprehensive view of the classification performance of each algorithm.

The results of evaluation on MARS method performances are given in Table 6. The accuracy measures are quite successful in both tables and the proposed methods exhibit high precision through the datasets.Creation of

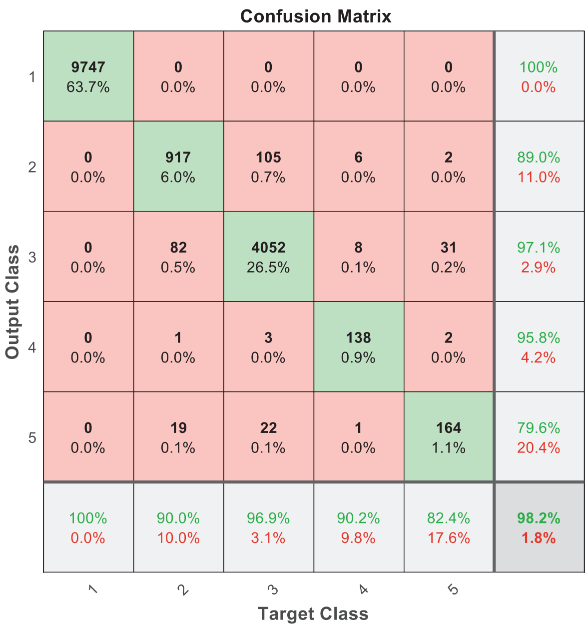

Figure 1. Confussion matrix of model-1.

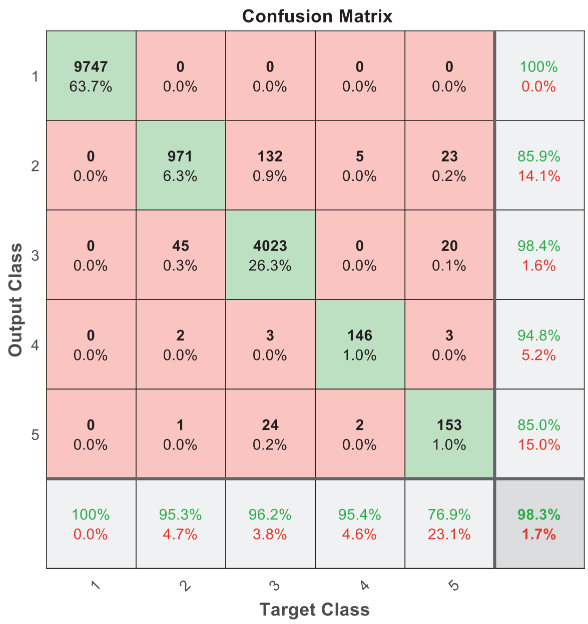

Figure 2. Confussion matrix of model-2

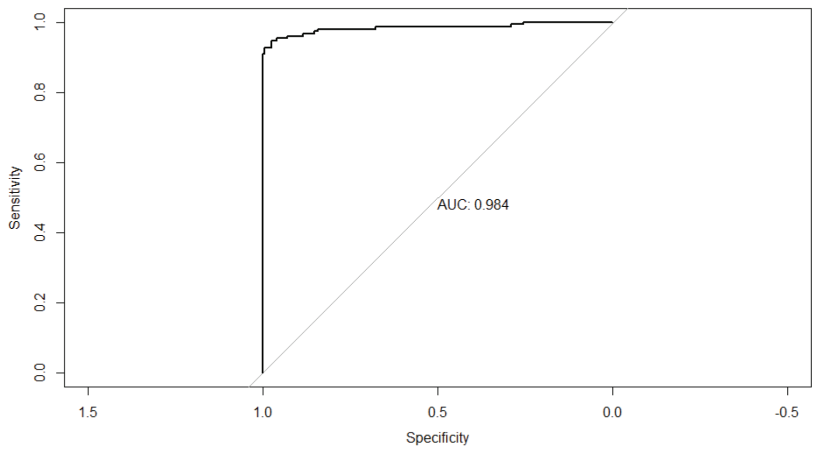

Figure 3. ROC of model-1.

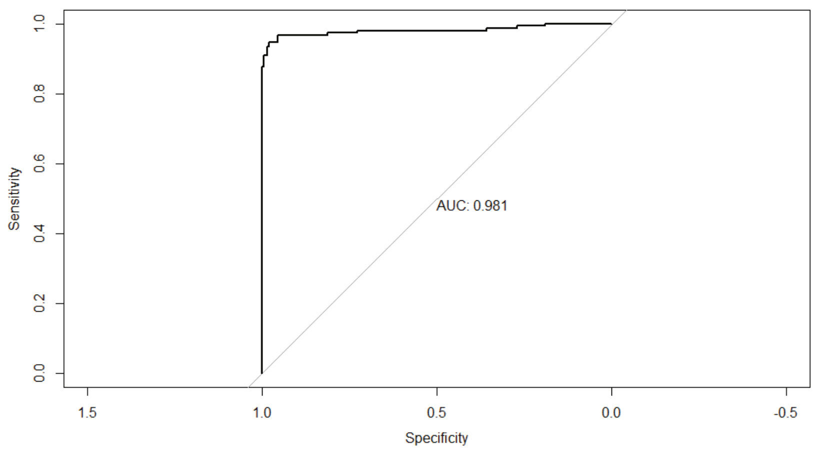

Figure 4. ROC of model-2.

Table 6. ROC analyzes

Model name	Accuracy	Recall	Precision	F1-Score	AUC
MARS_Model-1	98.1569	0.9190	0.9231	0.9208	0.984
MARS_Model-2	98.3007	0.9276	0.9291	0.9274	0.981

effective classification on such an imbalanced dataset is quite striking across ten different subsets. In our study on discrimination of anemia disease through the same levels of different parameters, we applied the two tests which are capable of successfully modeling and interpreting interaction among anemia parameters in different levels. The high efficacy of MARS method was revealed in the classification of anemia disease. The F1-Score values of 92.08% and 92.74% signify a higher percentage of accurately classified instances within the minority class and a decreased misclassification rate. This metric provides an overall assessment of model performance by balancing precision and recall. The research demonstrates positive outcomes in the classification of diverse types of anemia diseases using MARS methods, specifically highlighting their effectiveness in addressing issues arising from dataset imbalance.

CONCLUSION

The dataset used in the study consists of complete blood count results obtained from studies conducted at Gaziosmanpaşa University Faculty of Medicine. This data set includes information on 15.300 patients between 2013-2018. In the study conducted by Kılıçarslan and colleagues, data from pregnant women, cancer patients and children were excluded and the dataset was subjected to a noise removal process to eliminate insignificant and missing parameter records. The parameter values of the dataset, which was made publicly available in the Kaggle database, were normalized by scaling them between 0.1 and 0.9. This normalization aimed to compensate for significant numerical differences between parameter values.

Tests were conducted on two different models (Model-1 and Model-2). The tests performed on two different models with parameters set as multivariate adaptive regression spline method 1st order and pruning coefficient 20 for Model-1 and 2nd order and pruning coefficient 20 for Model-2 gave successful results. For Model-1, 98.1569% accuracy, 92.31% precision, 91.90% recall, 92.08% F1-score were obtained. Model-2 showed 98.30% accuracy, 92.91% precision, 92.76% recall and 92.74% F1-score. As a result, Model-1 misclassified a total of 282 patients, while Model-2 misclassified a total of 260 patients.

The findings from this study are expected to provide valuable information for medical students and practitioners in predicting five different types of anemia. We are confident that the classification performance of our proposed models will contribute positively to the existing literature. In future studies, more parameters and patient datasets can be used and success can be compared using different degrees and pruning coefficients.

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Anemia Disease Dataset at https://www.kaggle.com/code/serhathoca/anemia-disease-dataset/notebook.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence was not used in the preparation of the article.

REFERENCES

[1] Yılmaz H, Buldu A, Yılmaz K, Kuncan F. Early stage diabetes risk prediction with artificial intteligence methods. J Sci Rep B 2023:7:59–71.

[2] Dursun MA, Serttaş S. A Multi-Metric Model for analyzing and comparing extractive text summarization approaches and algorithms on scientific papers. Dicle Univ J Eng 2024;15:31–48. [CrossRef]

[3] Dursun MA, Yavuz G. L-PFO: Enhancing polar fox optimization with linear population size reduction for complex Benchmark problems. Artif Intell Stud 2025;8:100–115.

[4] Badik ŞT, Mutlu A. Machine learning classification models for the patients who have heart failure. Sigma J Eng Nat Sci 2024;42:235–44. [CrossRef]

[5] Bolukbasi IB. Dengesiz bir diyabet veri setinde makine öğrenmesi yöntemlerini kullanarak diyabet hastalığının teşhisi [Doctoral dissertation]. Bursa: Uludag University; 2023.
[6] Cihan M, Uzbaş B, Ceylan M. Fusion and CNN based classification of liver focal lesions using magnetic resonance imaging phases. Sigma J Eng Nat Sci 2023;41:135−45. [CrossRef]
[7] Dörterler S, Dumlu H, Özdemir D, Temurtas H. Hybridization of k-means and meta-heuristics algorithms for heart disease diagnosis. In book: New Trends in Engineering and Applied Natural Sciences. 1st ed. Chapter: 5 Publisher: İzmir: Duvar Publishing; 2022. p. 55.
[8] Yağmur N. A hybrid approach to obesity level determination with decision tree and pelican optimization algorithm. J Sci Rep A 2024;57:97−109. [CrossRef]
[9] Yağmur N, Temurtaş H, Dağ İ. Classification of anemia disease using artificial neural networks methods. J Sci Rep 2023;8:20−34.
[10] Narin D, Onur TÖ. The effect of hyper parameters on the classification of lung cancer images using deep learning methods. Erzincan Univ J Sci Technol 2022;15:258−68. [CrossRef]
[11] Dörterler S. Kanser hastalığı teşhisinde ölüm oyunu optimizasyon algoritmasinin etkisi. In: Genç N, Mühendislik Alanında Uluslararası Araştırmalar VIII. Konya: Eğitim Yayınevi; 2023. p. 15.
[12] Şahin E, Özdemir D, Temurtaş H. Multi-objective optimization of ViT architecture for efficient brain tumor classification. Biomed Signal Process Control 2024;91:105938. [CrossRef]
[13] Dörterler S, Dumlu H, Özdemir D, Temurtaş H. Hybridization of meta-heuristic algorithms with K-means for clustering analysis: Case of medical datasets. Gazi J Eng Sci 2024;10:1−11. [CrossRef]
[14] Aydemir F, Arslan S. A system design with deep learning and IoT to ensure education continuity for Post-COVID. IEEE Trans Consum Electron 2023;69:217−255. [CrossRef]
[15] Nahzat S, Yağanoğlu M. Diabetes prediction using machine learning classification algorithms. Eur J Sci Technol 2021;24:53−9.
[16] Özcan B, Bakır H. Yapay zeka destekli beyin görüntüleri üzerinde tümör tespiti. Int Conf Pioneer Innov Stud 2023;1:297−306. [CrossRef]
[17] Gençtürk TH, Gülağız FK, İsmail K. A Comparative analysis of brain hemorrhage diagnosis on CT scans using deep learning methods. J Intell Syst Theory Appl 2023;6:75−84. [CrossRef]
[18] WHO. Anaemia. World Health Organization. Available at: https://www.who.int/healthtopics/anaemia#tab=tab_1 Accessed on Oct 4, 2023.
[19] Salman ME, Arı S. Bulanık mantık tekniği kullanılarak anemi hastalığı teşhisi. Sakarya: Sakarya Üniversitesi Fen Bilimleri Enstitüsü; 2016. p. 61.
[20] Yagmur N, Dag I, Temurtas H. Classification of anemia using Harris Hawks optimization method and multivariate adaptive regression spline. Neural Comput Appl 2024;36:1−20. [CrossRef]
[21] Dixit A, Jha R, Mishra R, Vhatkar S. Prediction of anemia disease using machine learning algorithms. Intelligent Computing and Networking: Proceedings of IC-ICN 2022, Springer; 2023, p. 229−238. [CrossRef]
[22] Dhakal P, Khanal S, Bista R. Prediction of anemia using machine learning algorithms. AIRCC's Int J Comput Sci InformTechnol 2023;15:15−30. [CrossRef]
[23] Appiahene P, Asare JW, Donkoh ET, Dimauro G, Maglietta R. Detection of iron deficiency anemia by medical images: a comparative study of machine learning algorithms. BioData Min 2023;16:1−20. [CrossRef]
[24] Saihood Q, Sonuç E. The efficiency of classification techniques in predicting anemia among children: A comparative study. International Conference on Emerging Technology Trends in Internet of Things and Computing, New York: Springer; 2021, p. 167−81. [CrossRef]
[25] Yağmur N. Anemi hastalığı sınıflandırmasında karga arama optimizasyon algoritması. Mühendislik Alanında Akademik Araştırma ve Derlemeler 2023;291−307.
[26] Farzaliyev E, Saihood Q, Sonuç E. Çocuklarda anemi hastalığının teşhisinde topluluk öğrenme yöntemlerinin kullanılması. Int Conf Recent Acad Stud 2023;1:129−135. [CrossRef]
[27] Karagül Yıldız T. Performance comparison of artificial intelligence techniques in the diagnosis of hematological diseases (Doctorial thesis). Sakarya: Sakarya Üniversitesi Fen Bilimleri Enstitüsü; 2021.
[28] Aishwarya K. Threshold based feature selection for anemia prediction. 2023 International Conference on Computer Communication and Informatics (ICCCI), IEEE; 2023, p. 1−5. [CrossRef]
[29] Yeruva S, Gowtham BP, Chandana YH, Varalakshmi MS, Jain S. Prediction of Anemia Disease Using Classification Methods. Machine Learning Technologies and Applications: Proceedings of ICACECS 2020. New York: Springer; 2021, p. 1−11. [CrossRef]
[30] Yagmur N, Dag I, Temurtas H. A new computer-aided diagnostic method for classifying anaemia disease: Hybrid use of Tree Bagger and metaheuristics. Expert Syst 2023;41:e13528. [CrossRef]
[31] Kilicarslan S, Celik M, Sahin Ş. Hybrid models based on genetic algorithm and deep learning algorithms for nutritional Anemia disease classification. Biomed Signal Process Control 2021;63:102231. [CrossRef]
[32] Kilicarslan S. Anemia disease dataset. Available at: https://www.kaggle.com/code/serhathoca/anemia-disease-dataset. Accessed on Jul 3, 2023.

[33] Celik S, Yilmaz O. Prediction of body weight of Turkish tazi dogs using data mining Techniques: Classification and Regression Tree (CART) and multivariate adaptive regression splines (MARS). Pak J Zool 2018;50:575–583. [CrossRef]
[34] Ture M, Kurt I, Kurum AT, Ozdamar K. Comparing classification techniques for predicting essential hypertension. Expert Syst Appl 2005;29:583–588. [CrossRef]
[35] Rajawat RS, Singh KK, Mishra VN. Approximation by modified Bernstein polynomials based on real parameters. Mathematical Foundations of Computing 2024;7:297-309. [CrossRef]
[36] Mishra VN, Rajawat RS, Sharma V. On generalized quantum Bernstein polynomials. Advances in Pure and Applied Algebra: Proceedings of the CONIAPS XXVII International Conference 2021, Walter de Gruyter GmbH & Co KG; 2023, p. 149. [CrossRef]
[37] Raiz M, Rajawat RS, Mishra VN. α-Schurer Durrmeyer operators and their approximation properties. Ann Univ Craiova Math Comput Sci Ser 2023;50:189–204. [CrossRef]
[38] Pandey S, Rajawat RS, Mishra VN. Approximation properties of modified Jain-Gamma operators preserving linear function. Palestine J Math 2023;12:169–182.
[39] Raiz M, Rajawat RS, Mishra LN, Mishra VN. Approximation on bivariate of Durrmeyer operators based on beta function. J Anal 2024;32:311–333. [CrossRef]
[40] Agwu OE, Elraies KA, Alkouh A, Alatefi S. Mathematical modelling of drilling mud plastic viscosity at downhole conditions using multivariate adaptive regression splines. Geoenergy Sci Eng 2024;233:212584. [CrossRef]
[41] Friedman JH. Multivariate adaptive regression splines. Ann Stat 1991;19:1–67. [CrossRef]
[42] Kuter S. Atmospheric correction and image classification on MODIS images by nonparametric regression splines (Doctorial Thesis). Ankara: Middle East Technical University; 2014.
[43] Özfalcı Y. Multivariate adaptive regression splines (Master thesis). Ankara: Gazi Üniversitesi; 2008.
[44] Dursun MA, Kavas PÖ. Beyond diagnosis: Cross-dataset evaluation of risk factors for thyroid cancer recurrence. Artif Intell Stud 2025;8:38–69. [CrossRef]