Süleyman Demirel Üniversitesi
Sağlık Bilimleri Dergisi
Cilt 15, Sayı 1, 91 - 104, 2024

Suleyman Demirel University
Journal of Health Sciences
Volume 15, Issue 1, 91 - 104, 2024

Nursing Strategies for Diabetic Patient Management: Predicting Parameter Values Post-Exenatide Treatment with Machine Learning Algorithm

Diyabetik Hasta Yönetiminde Hemşirelik Stratejileri: Eksenatid Tedavisi Sonrası Parametrelerin Değerlerinin Makine Öğrenme Algoritmasıyla Tahmin Edilmesi

Sıddıka ERSOY [1*], Remzi GÜRFİDAN [2]

[1] Division of Medical Nursing, Department of Nursing, Süleyman Demirel University Faculty of Health Sciences, Isparta, Turkey
[2] Division of Computer Programming, Department of Computer Technologies, Isparta University of Applied Sciences University, Yalvaç Technical Sciences Vocational School, Isparta, Turkey

Ö Z E T

Diabetes Mellitus'un (DM) küresel ölçekteki artışı obezite oranlarındaki artışa paraleldir; Türkiye'de yetişkinler arasında diyabet prevalansı %13,7, obezite ise %32'dir. Diyabet ve obezitenin iç içe geçmiş doğası ve ilave kronik hastalık riskinin artması nedeniyle diyabet hastalarının yönetimi kapsamlı bir yaklaşımı gerektirmektedir. Diyabet hemşireleri, düzenli değerlendirmeler, kan şekeri takibi, ilaç yönetimi ve hasta eğitimini kapsayan diyabet bakımında çok önemli bir rol oynamaktadır. İnkretin-mimetik glukagon benzeri peptid-1 reseptör agonistleri (GLP-1A), diyabet ve kilo kontrolünde üstünlük göstererek bunları ikinci basamak tedaviler olarak konumlandırmıştır. Diyabet hemşirelerinin diyabet hastalarına diyet rehberliği, fiziksel aktivite teşviki ve kilo verme yardımı yoluyla hayati destek sağlamasıyla, kilo yönetimi diyabet bakımında temel olmaya devam etmektedir. GLP-1A tedavisine hasta yanıtlarını tahmin etmek, tedavi sonuçlarını optimize etmek, kararları kolaylaştırmak ve potansiyel komplikasyonları önlemek için çok önemlidir.

Yapay zeka (AI) ve makine öğrenimi (ML), sağlık hizmeti sunumunu geliştirmek için umut verici yollar sunmaktadır. Çalışma eksenatid kullanan diyabet hastalarında makine öğrenmesi algoritmalarını kullanarak açlık kan şekeri düzeylerini, HbA1C değerlerini ve kilo kaybı sonuçlarını tahmin etmeyi amaçlamaktadır. Batı Akdeniz'deki gerçek hasta verilerinin analiz edildiği bu çalışma, kilo kaybı, açlık kan şekeri düzeyleri ve HbA1C değerlerini tahmin etmede SVR algoritması sırasıyla %99.9, %99.9 ve %97.3'lük başarı oranlarına ulaşmıştır.

Bulgularımız hemşirelikte, özellikle de diyabetik hasta yönetimine yönelik prognostik modellemede yapay zeka odaklı yaklaşımların potansiyelinin altını çizmektedir. Hemşireler, makine öğreniminden yararlanarak tedavi yanıtlarını tahmin edebilir, karar alma sürecini kolaylaştırabilir ve hasta bakım kalitesini yükseltebilir. Yapay zeka uygulamaları geliştikçe, bu teknolojileri hemşirelik rollerine entegre etmek, hasta merkezli bakımı ilerletmeyi ve sağlık sonuçlarını optimize etmeyi vaat etmektedir.

Anahtar Kelimeler: Diyabet yönetimi, inkretin tedavi, diyabet hemşireliği, makine öğrenmesi algoritmaları

Alınış / Received: 09.03.2024 Kabul / Accepted: 15.04.2024 Online Yayınlanma / Published Online: 22.04.2024

DOI: 10.22312/sdusbed.1449989
ARAŞTIRMA MAKALESİ

*Corresponding author: siddikaersoy@gmail.com

ABSTRACT

The global escalation of DM parallels the rise in obesity rates, with Turkey experiencing a prevalence of 13.7% for diabetes and 32% for obesity among adults. Managing diabetic patients necessitates a comprehensive approach due to the intertwined nature of diabetes and obesity, along with the heightened risk of additional chronic illnesses. Diabet nurses play a pivotal role in diabetic care, encompassing regular assessments, blood glucose monitoring, medication management, patient education. Incretin-mimetic glucagon-like peptide-1 receptor-agonists (GLP-1A) have demonstrated superiority in diabetes, weight control, positioning them as second-line treatments. Weight management remains fundamental in diabetes care, with Diabet nurses providing vital support through dietary guidance, physical activity promotion, and weight loss assistance for diabetic patients. Predicting patient responses to GLP-1A therapy is crucial for optimizing treatment outcomes, streamlining decisions, averting potential complications.

Artificial intelligence (AI) and machine learning (ML) offer promising avenues for enhancing healthcare delivery. Our study aimed to forecast fasting blood sugar levels, HbA1C values, and weight loss outcomes in diabetic patients using exenatide, utilizing the random forest algorithm. Analyzing real patient data from the Western-Mediterranean, this study achieved substantial success rates of %99.9, %99.9 and %97.3 in predicting weight loss, fasting blood sugar levels, and HbA1C values, respectively.

Our findings underscore the potential of AI-driven approaches in nursing, particularly in prognostic modeling for diabetic patient management. By leveraging ML, nurses can anticipate treatment responses, streamline decision-making, and elevate patient care quality. As AI applications evolve, integrating these technologies into nursing roles promises to advance patient-centered care and optimize health outcomes.

Keywords: Diabetes management, incretin therapy, diabet nursing, machine learning algorithm

1. Introduction

Studies show that with the increase in weight in the society, the incidence of Diabetes mellitus (DM), which is one of the serious chronic diseases, increases .[1-5] According to Turkey Diabetes, Hypertension, Obesity and Endocrine Disease Survey (TURDEP-I / II), in the Turkish adult population, the prevalence of diabetes has reached 13.7% and the obesity prevalence has reached 32%. At the same time, when the data are compared, it is seen that diabetes has increased by 90% and obesity by 44% [6]. Since the incidence of complications should be considered in diabetic patients, they should be evaluated with a more holistic perspective than other chronic diseases. It is known that obesity, which accompanies Type II DM, triggers other physiological disorders and becomes an increased risk factor in the occurrence of additional chronic diseases. Increasing stress factors will make it difficult to catch the target blood sugar level. Therefore, the fact that the treatment used has a gain such as weight loss will make disease management easier.

Internal medicine nurses play a crucial role in overseeing the care of diabetic patients, ensuring adherence to treatment regimens, and implementing interventions to promote health and well-being. Nurses are at the forefront of diabetic patient care, conducting regular assessments, monitoring blood glucose levels, managing medications, and educating patients on self-management strategies. Continuous monitoring facilitates early detection of complications, optimization of glycemic control, and adjustment of treatment plans as needed [7-9].

Proper drug use is important for disease management, prevention of complications and good glycemic control in DM patients. For the effectiveness of drug use, the right drug should be chosen, treatment should be started on time and patient compliance should be available [10,11]. In order to reach the target HbA1C level in treatment, incretins took place as an additional option. Incretins are hormones that stimulate insulin secretion in response to food intake [12,13]. Incretin-mimetic glucagon-like peptide-1 receptor agonists (GLP-1A: Glucagon-like peptide-1 receptor agonists) mimic incretin hormones. Thus, they increase the effect of this hormone. In a study on GLP-1A, it was shown that GLP-1A were superior to other diabetes treatments in terms of diabetes and weight control. Thus, GLP-1 agonists took place in the clinic as a secondline treatment option [10,14-17].

Weight management is a critical aspect of diabetes care, as excess weight exacerbates insulin resistance and increases the risk of cardiovascular complications. Internal medicine nurses can provide dietary guidance, promote physical activity, and offer motivational support to facilitate weight loss among diabetic patients. Weight reduction not only improves glycemic control but also enhances overall health and reduces the need for pharmacological interventions [18-20].

Benefit and loss relationship is one of the situations that should be taken into consideration when choosing the treatment method used. Anticipating how patients will respond to the GLP-1A treatment process will be time-saving to make the right decisions in treatment management. With quick decision making in treatment preferences, complications that may develop in patients can be prevented. It will also allow planning lifestyle arrangements that will improve the quality of life by predicting the weight loss of patients.

Artificial intelligence practices and innovations in technology are increasing day by day. Artificial intelligence is one of the most important components of the Industry 4.0 concept, which is formed by these developments. Artificial intelligence can generally be defined as the processes of imitating human behavior and thoughts and transferring executive skills to machines. In addition, artificial intelligence has many different uses in everyday life [21-24]. Artificial intelligence, which is used in many fields today, is also important in the field of health. Artificial intelligence use is one of the preferred ways for diagnosis and prognosis monitoring in serious chronic diseases such as diabetes, cancer and heart diseases in the field of health [25].

To prevent loss of time in the diagnosis phase of patients applying to the hospital and to follow up the treatment especially in the settlements where the transportation of health personnel is difficult increase the use of artificial intelligence in the field of health. Also economic losses duae to excessive use of staff makes artificial intelligence more valuable [26,27]. Studies emphasize the benefits of using artificial intelligence applications in practice [25-27]. The concept of Machine Learning, which is one of the methods used in the realization of Artificial Intelligence technology; is defined as the ability of the computer to learn without being programmed explicitly [28,29]. By using machine learning algorithms, regression, time series estimates, classification operations, prediction operations can be done. In the application of machine learning method, algorithms such as Naive Bayes classifier, K-Means, linear regression, logistic regression, support vector machines, random forest, artificial neural network, decision trees and random forest, nearest neighbor are used.

This study was planned experimentally to predict the patient's fasting blood sugar value, HgA1C and weight loss during Exenatide use using machine learning algorithms, according to the measurement values taken from patients who were given treatment counseling by nurses.

2. Material and Method

The data forming the basis of this study were selected from real data obtained from patients living in the Western Mediterranean, diagnosed with Type II DM, and using Exenatide, for use in a published study with the approval of the researchers [30]. The data were taken from patients' analysis results and survey data on a voluntary basis. The data set consisting of the data collected from the patients was rearranged to create software on the computer. Standard statistical analysis methods were not used during and after data editing. In the data set used, 13 independent variables were defined as input data. As a result of the algorithms applied according to the input values, the system is expected to estimate the value of the intended parameter. Twenty per cent of the 200 data rows in the data set were allocated as test data. In addition, 20 rows were removed from the data set and reserved for validation and were not included in the training. The final validation tests of the trained model were carried out by predicting these values that were not encountered in training and the success graphs obtained are given.

In this study, experimental studies were carried out to estimate the fasting blood glucose value, HgA1C and weight loss data of patients diagnosed with DM and using Exenatide by using machine learning algorithm according to the measurement values taken from individuals. In the proposed model, 11 independent variables were defined as input data in the dataset created using real data from patients. The dependent, independent variables and data types created in the software are written in detail in Table 1.

Table 1. The Independent Variables of Patients Used as Input Data in Machine Learning Algorithm

Parameter	**Definition**
Gender	Male = 0 Female = 1
Age	0 - 105 years
Duration of Diabetes	0 - 40 years
Height	20 - 220 cm
Weight	2 - 200 kg
Smoking	0- not using 1- Using
Alcohol	0- not using 1- Using
Duration of Exenatide Usage	1- 100 weeks
Fasting Blood Sugar	10 – 500 mg/dl
BMI	10 – 100 kg/m^2
HgA1C	0 – 50%
Fasting Blood Sugar 2	10 – 500 mg/dl
HgA1C2	0 – 50 %
Weight loss	1 – 100 kg

Firstly carried out, the suitability of the data in the data set was reviewed and appropriate independent variables were determined. Afterwards, whether there is missing data in the dataset was examined and the missing data determined were completed with appropriate algorithms. The types of data in the data set was made suitable for the correct training of the data and the correct implementation of the algorithms. After the preliminary part of the training was completed, the data were divided into two as training data and test data. 80% of all data is divided into educational data, and 20% is divided into test data. In order to obtain more successful results, the data were subjected to repeated training.

2.1. K-Nearest Neighbours Regressor Algorithm

K-Nearest Neighbors (KNN) is an effective machine learning method that is preferred as a classification or regression solver. The algorithm uses the classes or values of the nearest neighbouring points to classify or predict a new data point. The basic principle of KNN proceeds by recognizing that data points with similar characteristics tend to have the same class or a similar value. Considering x and y as axis values, after calculating the distance, the input x is considered as the class value with the highest probability. This is calculated by Equation 2.

$$P(y = j|X = x) = \frac{1}{K}\sum_{i \epsilon A} I\left(y^{(i)} = j\right) \quad (1)$$

2.2. Gradient Boosting Regressor Algorithm

Gradient Boosting is a machine learning algorithm used as a solver in classification and regression processes. This algorithm aims to create a strong learner by combining weak learners together. Gradient Boosting aims to combine weak predictors (usually decision tree type models) to create a strong prediction model. The basic principle of how this algorithm works is to correct the erroneous learning of the previous weak estimator by adding new estimators. This process affects the calculation of the weights, while the new values are determined by the loss function. Equation 4 is used for the overall model calculation.

$$\gamma_m = arg_\gamma min \sum_i^n L(y_i, F_{m-1}(x_i) + \gamma) \quad (2)$$

Here i = 1 -n belongs to r_{ij}, where j represents the leaf. y is the observed value, γ is the predicted value

2.3. Random Forest Regressor Algorithm

Random Forest is a machine learning algorithm that is widely used especially in classification and regression problems. Random Forest can create a more powerful and generalizable model by combining multiple decision trees. When decision trees are configured for regression models, the average of the decision trees is the prediction value. Random Forest uses randomization to minimize the risk of overfitting. Random feature selections and random generation of data subsets make the model more diverse and generalizable. Mean square error value for Random Forest is calculated as in Equation 5.

$$RF_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(f_i - y_i)^2 \quad (3)$$

Where N is the number of data points, f_i is the value returned by the model and y_i is the actual value for data point i.

2.4. AdaBoost Regressor Algorithm

AdaBoost (Adaptive Boosting) is an ensemble learning algorithm for building strong models. AdaBoost aims to build a stronger model by combining weak models together. The AdaBoost algorithm is an algorithm that works on weights and each weak classifier is assigned a weight. Once a classifier is trained on the weighted training set, the weights of the misclassified examples are increased, and the next classifier is trained on this updated weighted data set. This process continues until a desired number of iterations or specific learning objective is reached. Equation 6 is used for the overall model calculation. The error rate is calculated by ε_t, that is, It shows how well the t'th classifier is able to correct the errors made on the weighted training data set.

$$\varepsilon_t = \frac{\sum_{i=1}^{N} w_{i\,.1(h_t(x_i) \neq y_i)}}{\sum_{i=1}^{N1} w_i} \quad (4)$$

It's here, the 1 function is a function that indicates whether the expression in parentheses is true (1) or false (0).

2.5. Support Vector Regressor Algorithm

Support Vector Regressor (SVR) is an example of a type of machine learning algorithm known as support vector machines (SVM). SVR is an algorithm used for regression problems, i.e. when an output variable must be predicted by one or more input variables. The main goal of SVR is to optimally classify the points in a regression problem around a hyperplane. However, unlike the classification problem, in regression problems the output value is a continuous variable. In determining this hyperplane, SVR tries to find a hyperplane that minimises the error of the model and at the same time contains as many points as possible within a certain tolerance (epsilon). SVR is particularly effective when there is noise in the dataset because, thanks to a certain tolerance (epsilon), it can reduce the impact of this noise. Furthermore, through the kernel functions of SVR, you can also model non-linear relationships. This allows SVR to be used even when there is no linear relationship in your dataset.

Support Vector Regression (SVR) can be expressed as a mathematical model as follows:

Data set (x_{one}, and_{one}), (x_2, and_2), …, (x_n, and_n), including, x_i while representing input properties,andi refers to the corresponding outputs (predicted values). The main purpose of SVR is to predict outputs from input features using a hyperplane. This hyperplane is expressed by an equation of the following general form:

$$f(x)=\langle In ,x\rangle+b \quad (5)$$

Here:

- $f(x)$refers to the predicted output.
- In*In*And*b*are the parameters of the SVR model.
- *X* refers to the input properties.
- ⟨ *In* ,*x*⟩ is the inner product operation.

SVR aims to find the best hyperplane within a certain margin of error (epsilon). This can be expressed as an optimization problem with the following constraints:

$\min_{w, b, \xi, \xi^* \text{one} /2} \| w \|^2 + C \sum i=1 \rightarrow n \, (X_i + X_i^*)$

Under this minimization problem, the following constraints are satisfied:

$$\begin{aligned} &and_i - \langle In, x_i \rangle - b \leq \epsilon + X_i \\ &\langle In, x_i \rangle + b - and_i \leq \epsilon + X_i^* \\ &X_i, X_i^* \geq 0 \end{aligned} \quad (6)$$

Here:
- C is an editing parameter and controls the flattening of the model.
- X_i And X_i^* refers to points with low margins of error.
- ϵ stands for fault tolerance.

This constrained optimization problem is the mathematical expression of support vector regression. This model is an extended version of SVM for regression problems, trying to optimally fit the input data around the hyperplane while maintaining maximum margins within a certain error tolerance.

2.6. Evaluation of the models

Error metrics used to evaluate the success of machine learning algorithms are used to measure how well the model performs. These metrics help to assess how well a model's predictions match the true values and the generalization ability of the model.

Mean absolute error (MAE) is a metric that shows how close the predicted values are to the true values. This metric is calculated by Equation 7 [31-33].

$$MAE = \frac{1}{n} \sum_{r=1}^{n} \left| P_d^{r,m} - P_d^{r,c} \right| \quad (7)$$

Root means square error (RMSE) was chosen to compare the prediction errors of different trained models. The closer the RMSE value is to 0, the better the predictive ability of the model in terms of its absolute deviation. The RMSE value is calculated by Equation 8 [31-34].

$$RMSE = \sqrt{\frac{1}{n} \sum_{r=1}^{n} (P_d^{r,m} - P_d^{r,c})^2} \quad (8)$$

The coefficient of determination (R^2) is used to estimate model efficiency and is calculated by Equation 9 [31].

$$R^2 = 1 - \frac{\sum_{r=1}^{n} (P_d^{r,m} - P_d^{r,c})^2}{\sum_{r=1}^{n} (P_d^{r,m} - \bar{P}_d^{r,m})^2} \quad (9)$$

MSE either assesses the quality of an estimator. The MSE metric is calculated by Equation 10.

$$MSE = \frac{1}{n} \sum_{r=1}^{n} (P_i - P_i')^2 \quad (10)$$

3. Results

In this study, the weight loss column, axis column and a1c column in the data set used in this study will be estimated separately. During the training process of each prediction, the other two dependent variables will be included in the data set. In this section, a separate subheading is opened for the estimation process of each independent variable and the obtained results are presented with graphs and numerical results. In addition, the effect factor importance ranks of the independent variables affecting the dependent variable are also shown graphically.

3.1. Machine Learning Processes And Findings For Weight Loss

In Figure 1, the sum of the effect factor importance ranks of the independent variables affecting the dependent variable is shown in a bar graph as a percentage ratio. According to this graph, it is concluded that the age criterion is ineffective in this process. In addition, it is seen that height, weight and A1c values are of great importance in predicting weight loss.

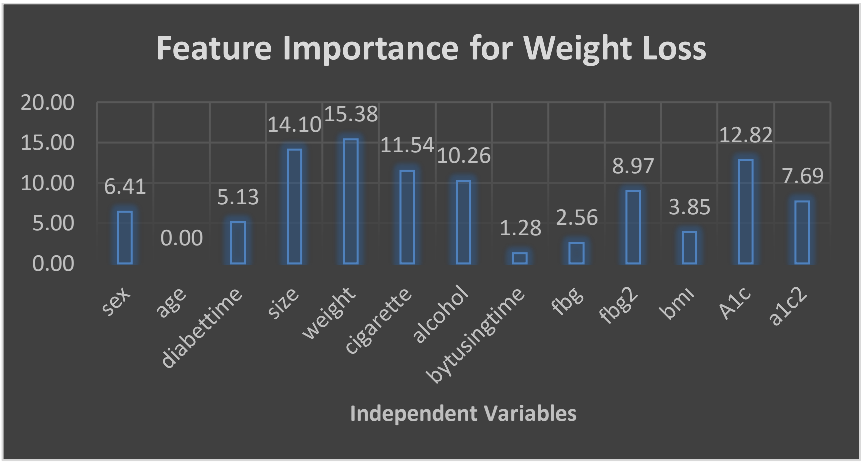


Figure 1. Influence factor plot of the independent variables on the outcome for Weight Loss

Table 2 shows the results of the training of KNN, AdaBoost, RandomForest, GradientBoosting, GradientBoosting, and Support Vector Regressor algorithms on weight loss prediction. In the tables, the matching of actual and prediction points, the distribution of prediction errors made by the model, and the validation graph are shown in detail. When these graphs are analysed, it can be seen that GBoosting and SVR are the algorithms that are successful and have high prediction success as a result of training.

Table 2. Performance graphs of machine learning models for weight loss

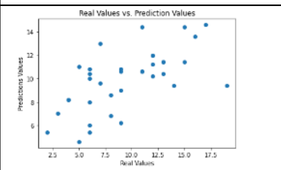
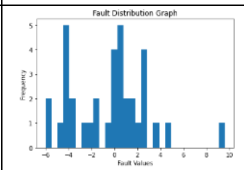
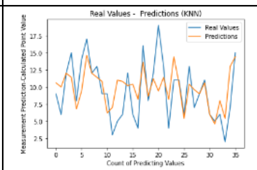
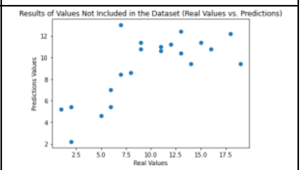
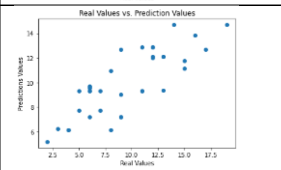
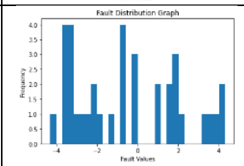
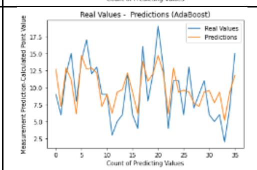
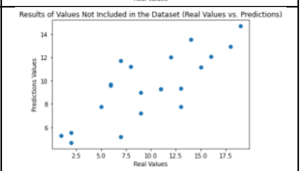
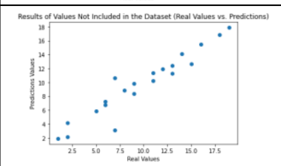
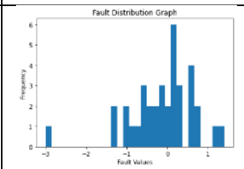
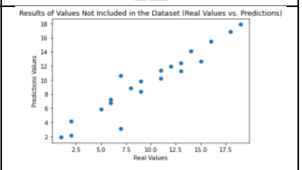
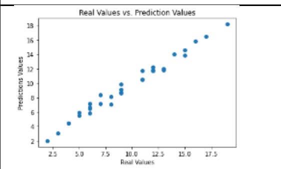
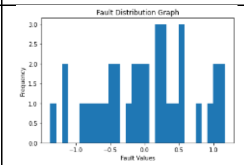
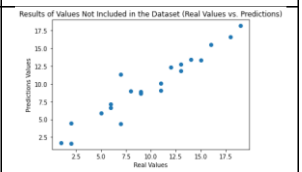
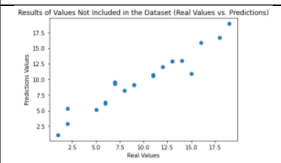
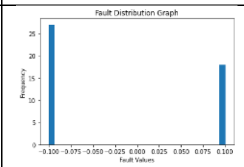
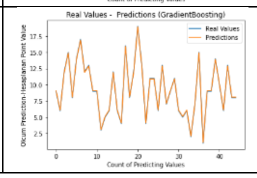
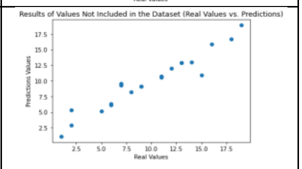
Algorithms	Real vs. Prediction Points	Fault Distribution	Real vs. Prediction Values	Validation Graph
KNN				
AdaBoost				
Random Forest				
G.Boosting				
SVR				

Table 3 shows the MSE, MAE, RMSE, R square, standard deviation and correlation coefficient metric values obtained from the training of KNN, AdaBoost, RandomForest, GradientBoosting, GradientBoosting, and Support Vector Regressor algorithms. While the MAE, MSE, and RMSE values have the lowest values close to 0 in the SVR algorithm, the fact that the R-squared value is closest to the highest value of 1 makes the SVR algorithm the algorithm that gives the best results. In addition to this success comment, the similarity and closeness values in predicting the data not seen in the training in the validation process are close to 0, which is the lowest standard deviation, and the correlation coefficient value of 1, which is the highest value, proves how close the model predicts in real life.

Table 3. Training result metric values of machine learning models for Weight Loss

Algorithms & Metrics	MSE	MAE	RMSE	R-Squared	Standard Deviation	Correlation Coefficient
KNN	10.715	2.511	3.273	0.386	3.244	0.630
AdaBoost	7.428	2.379	2.725	0.574	2.705	0.773
RandomForest	0.641	0.569	0.801	0.963	0.791	0.985
GradientBoosting	0.443	0.547	0.665	0.974	0.665	0.990
SVR	0.010	0.100	0.100	0.999	0.097	0.999

Table 4 shows the hyperparameter values used in the fine-tuning processes that enable the Support Vector Regressor algorithm to give such successful results in training. While preparing the SVR model, the optimum results of the Kernel, Gamma, Epsilon, C, Degree and Maximum iteration hyperparameters were obtained by testing with experiments.

Table 4. Hyperparameter settings of the SVR model for Weight Loss

Hypermeter	Value	Hypermeter	Value
Kernel	'rbf'	Gamma	0,1
Degree	3	Epsilon	0,1
C	1e3	Max. Iter	-1

In this study, metrics measuring error values and success rate were used to evaluate the experimental results. The values of error metrics and success metrics in the predictions of the proposed model are shown in Table 2, Table 3 and Table 4. The SVR algorithm was the most successful in predicting the weight loss values of the proposed model with a success rate of 99%.

3.2. Machine Learning Processes And Findings for Fasting Blood Glucose

In Figure 2, the sum of the importance ranks of the influence factor of the independent variables affecting the dependent variable is shown in a bar graph as a percentage ratio. According to this graph, it is concluded that the BMI criterion is ineffective in this process. In addition, it is seen that gender, a1c2 and fbg2 values are of great importance in fasting blood glucose prediction.

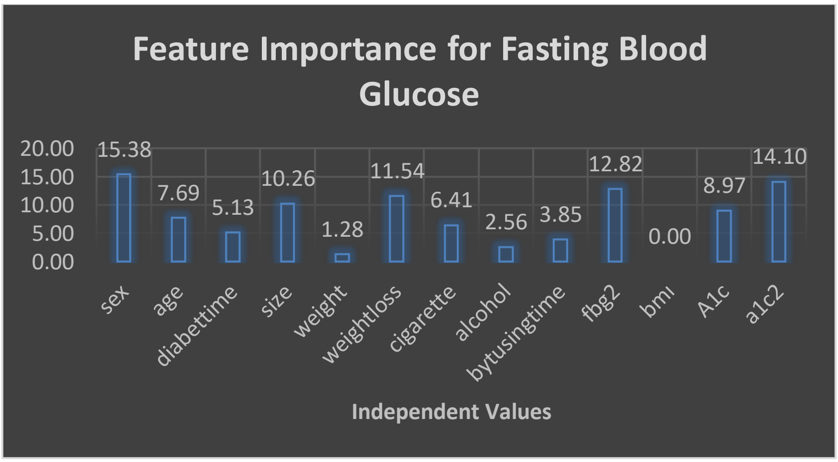


Figure 2. Influence factor plot of the independent variables on the outcome for Fasting Blood Glucose

Table 5 shows the training of KNN, AdaBoost, RandomForest, GradientBoosting, GradientBoosting, and Support Vector Regressor algorithms on fasting glucose prediction and the results are shown. In the tables, the matching of actual and prediction points, the distribution of prediction errors made by the model, and the validation graph are shown in detail. When these graphs are analysed, it can be seen that GBoosting and SVR are the algorithms that are successful and have high prediction success as a result of training.

Table 5. Performance graphs of machine learning models for fasting blood glucose

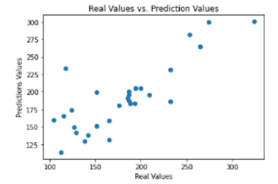
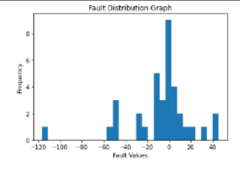
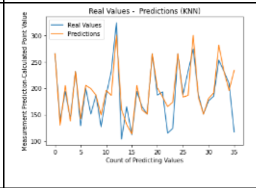
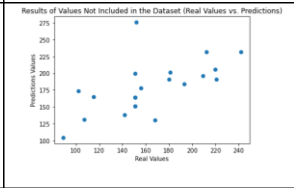
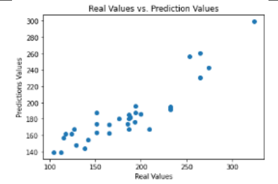
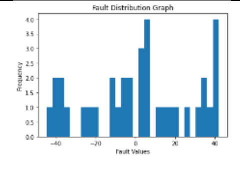
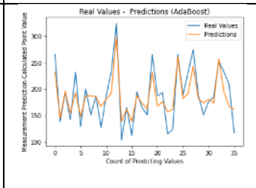
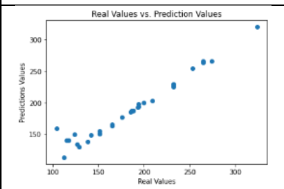
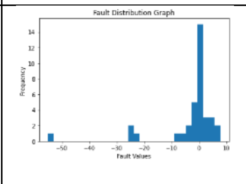
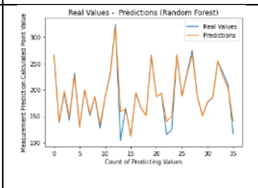
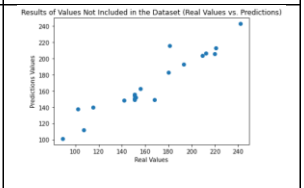
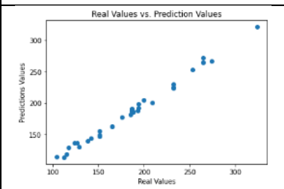
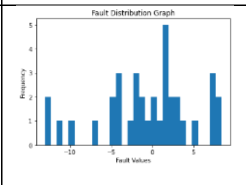
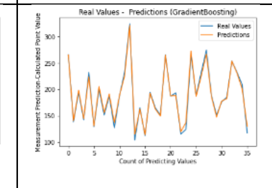
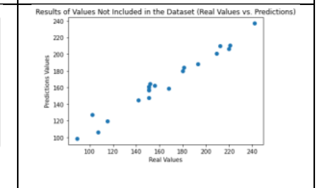
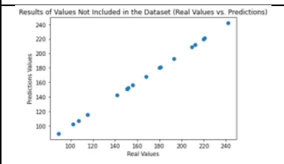
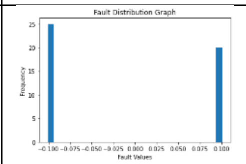
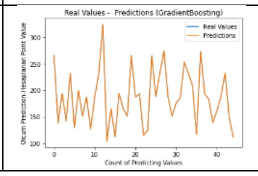
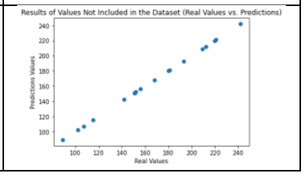
Algorithms	Real vs. Prediction Points	Fault Distribution	Real vs. Prediction Values	Validation Graph
KNN				
AdaBoost				
Random Forest				
G.Boosting				
SVR				

Table 6 shows the MSE, MAE, RMSE, R square, standard deviation and correlation coefficient metric values obtained from the training of KNN, AdaBoost, RandomForest, GradientBoosting, GradientBoosting, and Support Vector Regressor algorithms. While the MAE, MSE, and RMSE values have the lowest values close to 0 in the SVR algorithm, the fact that the R-squared value is closest to the highest value of 1 makes the SVR algorithm the algorithm that gives the best results. In addition to this success comment, the similarity and closeness values in predicting the data not seen in the training in the validation process are close to 0, which is the lowest standard deviation, and the correlation coefficient value of 1, which is the highest value, proves how close the model predicts in real life.

Table 6. Training result metric values of machine learning models for Weight Loss

Algorithms & Metrics	MSE	MAE	RMSE	R-Squared	Standard Deviation	Correlation Coefficient
KNN	916.985	18.627	30.281	0.672	29.313	0.834
AdaBoost	677.231	21.419	26.023	0.758	25.953	0.903
RandomForest	144.007	5.558	12.000	0.948	11.568	0.979
GradientBoosting	30.723	4.292	5.542	0.989	5.518	0.995
SVR	0.009	0.009	0.009	0.999	0.099	0.999

3.3. Machine Learning Processes And Findings for A1C

In Figure 3, the sum of the influence factor importance ranks of the independent variables affecting the dependent variable is shown in a bar graph as a percentage ratio. According to this graph, it is concluded that diabettime criterion is ineffective in this process. In addition, it is seen that height, alcohol use status and a1c2 values are of great importance in A1c prediction.

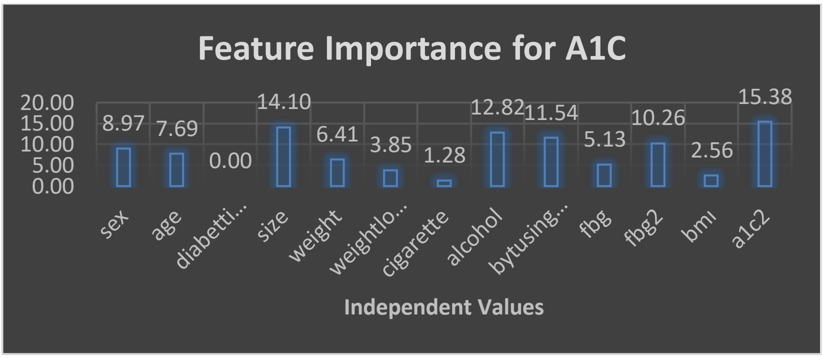


Figure 3. Influence factor plot of the independent variables on the outcome for A1C

In Table 7, the training of KNN, AdaBoost, RandomForest, GradientBoosting, GradientBoosting, and Support Vector Regressor algorithms on A1c prediction and the results are shown. In the tables, the matching of actual and prediction points, the distribution of prediction errors made by the model, and the validation graph are shown in detail. When these graphs are analysed, it can be seen that GBoosting and SVR are the algorithms that are successful and have high prediction success as a result of training.

Table 7. Performance graphs of machine learning models for A1C

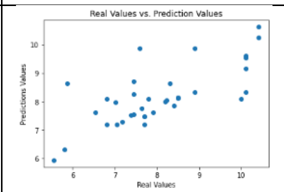
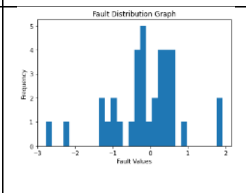
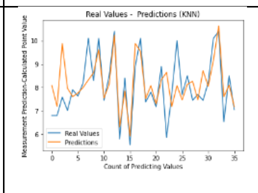
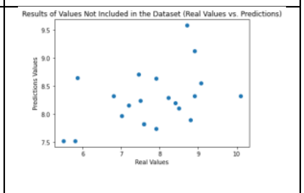
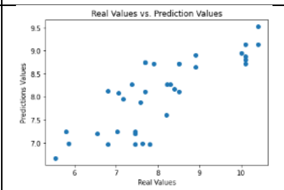
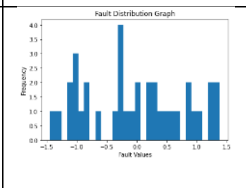
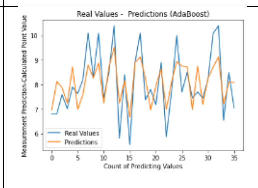
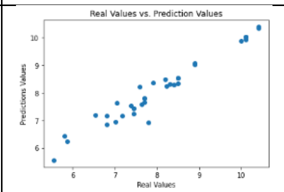
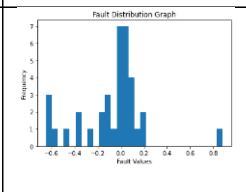
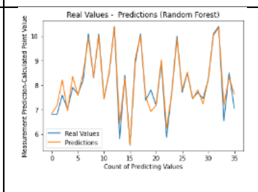
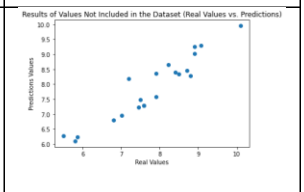
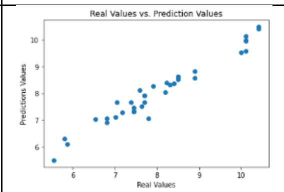
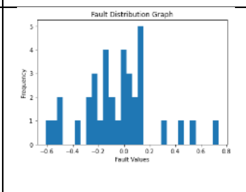
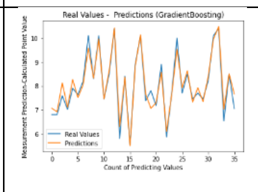
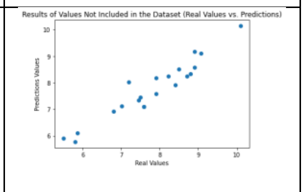
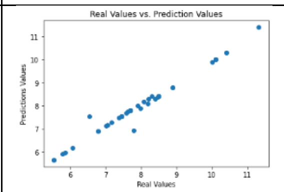
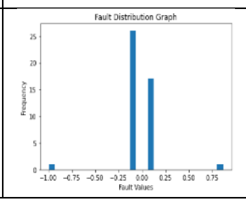
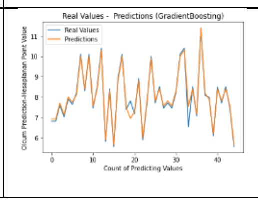
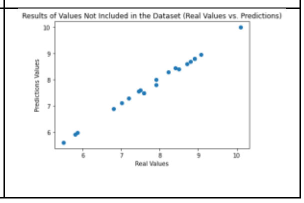
Algorithms	Real vs. Prediction Points	Fault Distribution	Real vs. Prediction Values	Validation Graph
KNN				
AdaBoost				
Random Forest				
G.Boosting				
SVR				

Table 8 shows the MSE, MAE, RMSE, R square, standard deviation and correlation coefficient metric values obtained from the training of KNN, AdaBoost, RandomForest, GradientBoosting, GradientBoosting, and Support Vector Regressor algorithms. While the MAE, MSE, and RMSE values have the lowest values close to 0 in the SVR algorithm, the fact that the R-squared value is closest to the highest value of 1 makes the SVR algorithm the algorithm that gives the best results. In addition to this success comment, the similarity and closeness values in predicting the data not seen in the training in the validation process are close to 0, which is the lowest standard deviation, and the correlation coefficient value of 1, which is the highest value, proves how close the model predicts in real life.

Table 8. Training result metric values of machine learning models for A1C

Algorithms & Metrics	MSE	MAE	RMSE	R-Squared	Standard Deviation	Correlation Coefficient
KNN	0.849	0.663	0.921	0.495	0.913	0.713
AdaBoost	0.670	0.690	0.819	0.601	0.818	0.801
RandomForest	0.086	0.191	0.294	0.948	0.285	0.976
GradientBoosting	0.081	0.213	0.286	0.951	0.282	0.977
SVR	0.047	0.136	0.218	0.973	0.217	0.987

Machine learning is a method frequently used in the application of artificial intelligence technology. Machine learning method, which is used in many scientific studies in different fields [10,28,29,35] is the basis of many studies in the field of health [36-38]. In machine learning method, classification and prediction [36-40] operations can be performed on images such as X-ray films, ultrasound records, MR images, and the same operations can be performed on numerical data [41-44].

4. Discussion and Conclusion

In most of the studies, only one step predictions such as diagnosis of the disease or prognosis prediction of the disease are included with artificial intelligence. However, in our study, both the diagnosis of the patients can be predicted and the expected benefit rates of the drug used simultaneously can be predicted. In our study, machine learning models were preferred due to their higher predictive power.

The fact that the data sets used in the study were obtained from real patients who applied to the clinics will enable the results to be used for the patient profile estimation of the region. Based on the data obtained, a very high success value was obtained in patients with DM Type II in terms of ease of follow-up. In the diagnosis of Type II DM, it was shown that a possible diagnosis can be made without the physician based on the physiological findings of the patient [25]. In a study about diagnosing metabolic syndrome without blood tests, the rate of making the correct diagnosis was determined as 85.12% [45]. The analysis of adipocytokines and anthropometric levels obtained from obese women (diabetic and non-diabetic) with experimental data set and the probability of having DM in women with obesity was tried to be determined. In a study on diabetic retinopathy, a network was created to define the stages of retinopathy, and it was found that the automatic detection system leads to clearer results than individual evaluations. When the results are interpreted, gains in terms of supporting early treatment came to the fore [46,47]. With the widespread use of artificial intelligence practices, primary health care services or specialist nurse follow-up will be preferred in the future in order to facilitate access to health in rural areas without hospitals and physician access [25].

In this study, experimental studies were carried out on the prediction of fasting blood glucose value, HgA1C average and weight loss data of patients using exenatide using machine learning algorithms according to the measurement values taken from individuals. The proposed model was found to be 99%, 98% and 98% successful in predicting Weight Loss, Fasting Blood Glucose Value and HgA1C Average values, respectively. In future studies, it is aimed to increase the amount of data in the data set, to observe the results by applying different algorithms and to increase the final success rate above 98%. This study is also important in terms of showing that machine learning algorithm applications can be applied for different nursing roles.

Declaration of Ethical Code

In this study, we undertake that all the rules required to be followed within the scope of the "Higher Education Institutions Scientific Research and Publication Ethics Directive" are complied with, and that none of the actions stated under the heading "Actions Against Scientific Research and Publication Ethics" are not carried out.

References

[1] TEMD: Türkiye Endokrinoloji ve Metabolizma Dernegi, ", Diabetes Mellitus ve Komplikasyonlarının Tanı, Tedavi ve Izlem Kılavuzu-2019," *Miki Matbaacılık San. ve Tic. Ltd*, vol. 12, no. 1, 2019.

[2] T. H. S. K. T.C. Sağlık Bakanlığı, "TC. Sağlık Bakanlığı: Türkiye Diyabet Programı 2015-20," Ankara, 2014.

[3] IDF: Internatıonal Dıabetes Federation., " Diabetes Atlas 2013," 2013.

[4] A. H. Mokdad *et al.*, "Prevalence of Obesity, Diabetes, and Obesity-Related Health Risk Factors, 2001," *JAMA*, vol. 289, no. 1, pp. 76–79, Jan. 2003, doi: 10.1001/JAMA.289.1.76.

[5] H. E. Bays, R. H. Chapman, and S. Grandy, "The relationship of body mass index to diabetes mellitus, hypertension and dyslipidaemia: comparison of data from two national surveys," *Int J Clin Pract*, vol. 61, no. 5, pp. 737–747, May 2007, doi: 10.1111/J.1742-1241.2007.01336.X.

[6] I. Satman *et al.*, "Twelve-year trends in the prevalence and risk factors of diabetes and prediabetes in Turkish adults," *Eur J Epidemiol*, vol. 28, no. 2, pp. 169–180, Feb. 2013, doi: 10.1007/S10654-013-9771-5/TABLES/2.

[7] A. B. Evert *et al.*, "Nutrition therapy for adults with diabetes or prediabetes: A consensus report," *Diabetes Care*, vol. 42, no. 5, pp. 731–754, 2019, doi: 10.2337/DCI19-0014/-/DC1.

[8] M. A. Powers *et al.*, "Diabetes Self-management Education and Support in Type 2 Diabetes: A Joint Position Statement of the American Diabetes Association, the American Association of Diabetes Educators, and the Academy of Nutrition and Dietetics," *Diabetes Educator*, vol. 43, no. 1, pp. 40–53, Feb. 2017, doi: 10.1177/0145721716689694/ASSET/IMAGES/LARGE/ 10.1177_0145721716689694-FIG2.JPEG.

[9] E. L. Johnson *et al.*, "Standards of Medical Care in Diabetes—2019 Abridged for Primary Care Providers," *Clin Diabetes*, vol. 37, no. 1, p. 11, Jan. 2019, doi: 10.2337/CD18-0105.

[10] A. D. Association, "6. Glycemic Targets: Standards of Medical Care in Diabetes—2018," *Diabetes Care*, vol. 41, no. Supplement_1, pp. S55–S64, Jan. 2018, doi: 10.2337/DC18-S006.

[11] M. Lambert, "ADA Releases Revisions to Recommendations for Standards of Medical Care in Diabetes," *Am Fam Physician*, vol. 85, no. 5, pp. 514–515, Mar. 2012, Accessed: Mar. 01, 2024. [Online]. Available: https://www.aafp.org/pubs/afp/issues/2012/0301/p514.html

[12] H. Elrıck, L. Stımmler, C. J. Hlad, and Y. Araı, "Plasma Insulin Response to Oral and Intravenous Glucose Administration," *J Clin Endocrinol Metab*, vol. 24, no. 10, pp. 1076–1082, Oct. 1964, doi: 10.1210/JCEM-24-10-1076.

[13] R. Çolak, "Tip 2 diabetes mellitus tedavisinde inkretinler," *J Exp Clin Med*, vol. 29, no. 1s, pp. 30–38, Jun. 2012, doi: 10.5835/JECM.OMU.29.S1.007.

[14] R. S. Poyatos, P. R. Servan, and C. V. Martínez, "[Effects of exenatide lar in type 2 diabetes mellitus and obesity]," *Nutr Hosp*, vol. 31, no. 1, pp. 292–298, 2014, doi: 10.3305/NH.2015.31.1.8283.

[15] J. L. Faillie, O. H. Yu, H. Yin, D. Hillaire-Buys, A. Barkun, and L. Azoulay, "Association of Bile Duct and Gallbladder Diseases With the Use of Incretin-Based Drugs in Patients With Type 2 Diabetes Mellitus," *JAMA Intern Med*, vol. 176, no. 10, pp. 1474–1481, Oct. 2016, doi: 10.1001/ JAMAINTERNMED.2016.1531.

[16] A. J. Garber *et al.*, "Consensus Statement by the American Association of Clinical Endocrinologists and American College of Endocrinology on the Comprehensive Type 2 Diabetes Management Algorithm – 2020 Executive Summary," *Endocrine Practice*, vol. 26, no. 1, pp. 107–139, Jan. 2020, doi: 10.4158/CS-2019-0472.

[17] S. E. Inzucchi *et al.*, “Management of Hyperglycemia in Type 2 Diabetes, 2015: A Patient-Centered Approach: Update to a Position Statement of the American Diabetes Association and the European Association for the Study of Diabetes,” *Diabetes Care*, vol. 38, no. 1, pp. 140–149, Jan. 2015, doi: 10.2337/DC14-2441.

[18] M. E. Lean *et al.*, “Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial,” *The Lancet*, vol. 391, no. 10120, pp. 541–551, Feb. 2018, doi: 10.1016/S0140-6736(17)33102-1.

[19] T. L. A. R. Group, “Cardiovascular Effects of Intensive Lifestyle Intervention in Type 2 Diabetes,” *https://doi.org/10.1056/NEJMoa1212914*, vol. 369, no. 2, pp. 145–154, Jul. 2013, doi: 10.1056/NEJMOA1212914.

[20] M. Espeland, “Look AHEAD (Action for Health in Diabetes): design and methods for a clinical trial of weight loss for the prevention of cardiovascular disease in type 2 diabetes,” *Control Clin Trials*, vol. 24, no. 5, pp. 610–628, Oct. 2003, doi: 10.1016/S0197-2456(03)00064-3.

[21] C. M. Bishop, *Prml*. 2006. Accessed: Mar. 07, 2024. [Online]. Available:https://link.springer.com /book/ 9780387310732

[22] Z. Obermeyer and E. J. Emanuel, “Predicting the Future — Big Data, Machine Learning, and Clinical Medicine,” *N Engl J Med*, vol. 375, no. 13, p. 1216, Sep. 2016, doi: 10.1056/NEJMP1606181.

[23] R. C. Deo, “Machine Learning in Medicine,” *Circulation*, vol. 132, no. 20, pp. 1920–1930, Nov. 2015, doi: 10.1161/CIRCULATIONAHA.115.001593.

[24] M. Long, Y. Cao, Z. Cao, J. Wang, and M. I. Jordan, “Transferable Representation Learning with Deep Adaptation Networks,” *IEEE Trans Pattern Anal Mach Intell*, vol. 41, no. 12, pp. 3071–3085, Dec. 2019, doi: 10.1109/TPAMI.2018.2868685.

[25] S. Spanig, A. Emberger-Klein, J. P. Sowa, A. Canbay, K. Menrad, and D. Heider, “The virtual doctor: An interactive clinical-decision-support system based on deep learning for non-invasive prediction of diabetes,” *Artif Intell Med*, vol. 100, p. 101706, Sep. 2019, doi: 10.1016/J.ARTMED.2019.101706.

[26] J. Baumbach and H. H. H. W. Schmidt, “The End of Medicine as We Know It: Introduction to the New Journal, Systems Medicine,” *https://home.liebertpub.com/sysm*, vol. 1, no. 1, pp. 1–2, Feb. 2018, doi: 10.1089/SYSM.2017.28999.JBA.

[27] A. Talaei-Khoei and J. M. Wilson, “Identifying people at risk of developing type 2 diabetes: A comparison of predictive analytics techniques and predictor variables,” *Int J Med Inform*, vol. 119, pp. 22–38, Nov. 2018, doi: 10.1016/J.IJMEDINF.2018.08.008.

[28] E. Rosten and T. Drummond, “Machine Learning for High-Speed Corner Detection,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3951 LNCS, pp. 430–443, 2006, doi: 10.1007/11744023_34.

[29] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” *Nature 2017 549:7671*, vol. 549, no. 7671, pp. 195–202, Sep. 2017, doi: 10.1038/nature23474.

[30] J. Peters, “Machine Learning for Robotics: Learning Methods for Robot Motor Skills,” 2008, Accessed: Mar. 07, 2024. [Online]. Available: https://www.ias.informatik.tu-darmstadt.de/uploads/Research/Thesis/thesis_1.pdf

[31] A. T. Hammid, M. H. Bin Sulaiman, and A. N. Abdalla, “Prediction of small hydropower plant power production in Himreen Lake dam (HLD) using artificial neural network,” *Alexandria Engineering Journal*, vol. 57, no. 1, pp. 211–221, Mar. 2018, doi: 10.1016/J.AEJ.2016.12.011.

[32] G. Mishra, D. Sehgal, and J. K. Valadi, “Quantitative Structure Activity Relationship study of the Anti-Hepatitis Peptides employing Random Forests and Extra-trees regressors,” *Bioinformation*, vol. 13, no. 3, p. 60, Mar. 2017, doi: 10.6026/97320630013060.

[33] M. K. AlOmar, M. M. Hameed, and M. A. AlSaadi, “Multi hours ahead prediction of surface ozone gas concentration: Robust artificial intelligence approach,” *Atmos Pollut Res*, vol. 11, no. 9, pp. 1572–1587, Sep. 2020, doi: 10.1016/J.APR.2020.06.024.

[34] C. J. Willmott and K. Matsuura, “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,” *Clim Res*, vol. 30, no. 1, pp. 79–82, Dec. 2005, doi: 10.3354/CR030079.

[35] S. Ersoy, İ. H. Ersoy, and Ş. Özkahraman-Koç, “Examination of Quality of life in Patients Using Exenatide with Different Dimensions,” *Journal Of Pakistan Medical Association*, vol. 72, no. 5, pp. 817–821, 2022.

[36] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, “Breast cancer histopathological image classification using Convolutional Neural Networks,” *Proceedings of the International Joint Conference on Neural Networks*, vol. 2016-October, pp. 2560–2567, Oct. 2016, doi: 10.1109/IJCNN.2016.7727519.

[37] J. Shan, S. K. Alam, B. Garra, Y. Zhang, and T. Ahmed, “Computer-Aided Diagnosis for Breast Ultrasound Using Computerized BI-RADS Features and Machine Learning Methods,” *Ultrasound Med Biol*, vol. 42, no. 4, pp. 980–988, Apr. 2016, doi: 10.1016/J.ULTRASMEDBIO.2015.11.016.

[38] X. Yang and B. Fei, “3D prostate segmentation of ultrasound images combining longitudinal image registration and machine learning,” *https://doi.org/10.1117/12.912188*, vol. 8316, pp. 803–811, Feb. 2012, doi: 10.1117/12.912188.

[39] X. G. Zhang, J. J. Xu, and G. Y. Ge, “Defects recognition on X-ray images for weld inspection using SVM,” *Proceedings of 2004 International Conference on Machine Learning and Cybernetics*, vol. 6, pp. 3721–3725, 2004, doi: 10.1109/ICMLC.2004.1380463.

[40] E. I. Zacharaki *et al.*, “Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme,” *Magn Reson Med*, vol. 62, no. 6, p. 1609–1618, Dec. 2009, doi: 10.1002/MRM.22147.

[41] A. A. Süzen, “LSTM Derin Sinir Ağları ile Üniversite Giriş Sınavındaki Matematik Soru Sayılarının Konulara Göre Tahmini”, *Engineering Sciences (NWSAENS)*, 2019, vol. 14, no. 3, p. 112-118.doi: 10.12739/NWSA.2019.14.3.1A0436.

[42] A. A. Süzen and K. Kayaalp, “Classification of Induction Motors by Fault Type with bidirectional Long-Short Term Memory Method”, 4. Uluslararası Mühendislik ve Doğa Bilimlerinde Yenilikçi Yaklaşımlar Sempozyumu, Samsun, Türkiye, 22 Kasım 2019. doi: 10.36287/setsci.4.6.074.

[43] Qeethara Kadhim Al-Shayea, “Artificial Neural Networks in Medical Diagnosis,” *IJCSI International Journal of Computer Science Issues*, vol. 8, no. 2, p. 150–154, 2011.

[44] C. Zhao, G. Z. Li, C. Wang, and J. Niu, “Advances in patient classification for traditional Chinese medicine: a machine learning perspective,” *Evid Based Complement Alternat Med*, vol. 2015, pp. 1–18, 2015.

[45] M. Barrios, M. Jimeno, P. Villalba, and E. Navarro, “Novel Data Mining Methodology for Healthcare Applied to a New Model to Diagnose Metabolic Syndrome without a Blood Test,” *Diagnostics 2019, Vol. 9, Page 192*, vol. 9, no. 4, p. 192, Nov. 2019, doi: 10.3390/ DIAGNOSTICS9040192.

[46] H. Lotfi *et al.*, “Machine learning as new promising technique for selection of significant features in obese women with type 2 diabetes,” *Horm Mol Biol Clin Investig*, vol. 41, no. 1, Mar. 2020, doi: 10.1515/HMBCI-2019-0019/MACHINEREADABLECITATION/RIS.

[47] H. Riaz, J. Park, H. Choi, H. Kim, and J. Kim, “Deep and Densely Connected Networks for Classification of Diabetic Retinopathy,” *Diagnostics 2020, Vol. 10, Page 24*, vol. 10, no. 1, p. 24, Jan. 2020, doi: 10.3390/DIAGNOSTICS10010024.