Experimental and Applied Medical Science, 2, 4: 242-256, 2021. DOI 10.46871/eams.2021.29

Use of the Support Vector Regression in Medical Data Analysis

Özlem Akay[1*], Mehmet Tunçeli[2]

[1]*Gaziantep Islam Science and Technology University, Faculty of Medicine, Department of Basic Medical Sciences, Gaziantep, Turkey.*
[2]*Osmaniye State Hospital, Physical Therapy and Rehabilitation, Osmaniye, Turkey.*

Abstract

This study was considered as an application for the use of support vector machines (SVM), which is frequently used in many fields, in medical research and a prediction model was created with support vector machine regression (SVR) using medical study data. For determining the appropriate model and estimating the best parameter, estimation models were obtained with linear regression (LR), SVR, tuned-SVR that provide flexibility in terms of maximum error and penalty cost, and SVR using kernel functions. Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) values were calculated to evaluate the performance of the models. The obtained results showed that the SVR model using the radial kernel function provided the best prediction among the predictive models.

Key words: *Support vector regression, Medical data, Occipitocervical angle, Cervical vertical translation distance.*

Introduction

In recent years, Support Vector Machines (SVM), a data mining and machine learning method, are widely used in multidisciplinary applications and offer effective solutions to a variety of real-life problems. Support Vector Machines is proposed for regression and classification-based problems by Vapnik (1982) (1). SVM operates as structural risk minimization in statistical learning theory rather than the experimental risk minimization method in which the classification function is derived by minimizing the mean square of an error in the dataset. The theory of the method is based on Structural Risk Minimization (2). SVM has been widely used in data mining, pattern and face identification, image and text classification, quality control methods, economics, biology, genetics, and other bioinformatics applications due to its strong theoretical structure, high generalization ability, and high performance in applications (3). Moreover, one of the most important areas where data mining is used is the field of medicine (health-care). Due to the rapid growth of medical data, it has become indispensable to use data mining techniques to assist decision support and forecasting systems in the healthcare field (4). Information discoveries to be obtained from medical data are very important because the data in

***Corresponding Author:** Özlem Akay, Tel: +90 505 736 9153, E–mail:ozlem.akay@gibtu.edu.tr, ORCID: orcid.org/0000-0002-9539-7252.

the field of medicine is vital (5). Data Mining methods can be considered as a system that supports decisions in effective diagnosis and treatment or in institutions. In this sense, data mining can provide important contributions to health institutions in areas such as the diagnosis applied to patients in health institutions, the development of treatment methods, the shortening of treatment periods, the determination of regional disease groups, and the making of management decisions (6). It is thought that data mining used in the medical field will provide a new horizon for medical research by filling the place of medical research, which requires many clinical studies in the future and has disadvantages in terms of both economic and human health (7). Data Mining plays an important role in uncovering new trends in healthcare, which helps all parties involved in the medical field.

There are many studies in the literature on different medical data sets using the support vector regression method. Some of the studies on the subject are summarized as follows.

Georga et al. (2013) used the support vector regression (SVR) method to evaluate the effect of subcutaneous glucose on hypoglycemia. The method was evaluated on a dataset of 15 patients with type 1 diabetes under free-living conditions. The model estimated the non-nocturnal (i.e., diurnal) and nocturnal events during sleep separately over 30-minute and 60-minute horizons, using information about the final glucose profile, meals, insulin intake, and physical activities for the hypoglycemic threshold of 70 mg/dL. Nocturnal hypoglycemic events were predicted with 94% precision for both horizons and time delays of 5.43 min and 4.57 min, respectively. Regarding diurnal events, excluding physical activities, the sensitivity is 92% and 96% for a 30-minute and 60-minute horizon, respectively, with both time lags less than 5 minutes. SVR results give an accurate prediction of hypoglycemia and perform well in most diurnal and nocturnal cases (8). Guo et al.(2021) proposed a Multi Kernel Support Vector Regression (MKSVRMAS) model that Maximizes Mean Similarity to estimate the dry weight of hemodialysis patients by collecting clinically effective characteristic data. Demographic data, anthropometric measurements, and Bioimpedance spectroscopy (BIS) of 476 hemodialysis patients were collected. The obtained data were used for the model. Experimental results show that dry weight is positively correlated with Body Mass Index and Heart Rate. Age, sex, systolic blood pressure, diastolic blood pressure, hemodialysis time are negatively associated with dry weight. The Model's Mean Squared Error (RMSE) is 1.3817. The proposed model could serve as a viable alternative for the estimation of the dry weight of hemodialysis patients, thus providing a new avenue for clinical application (9). Riaz et al. (2009) estimated the motion of lung tumors monitored at 30 Hz with multidimensional linear adaptive filters and support vector regression. They extended previous work from other groups looking at adaptive filters using the general framework of a multiple-input single-output (MISO) adaptive system that uses multiple correlated signals to predict the movement of a tumor. At 400 ms latency, root mean square errors (RMSEs) for the 14 treatment sessions studied using linear regression, single-output adaptive filter,

MISO, and support vector regression were 2.58, 1.60, 1.58, 1.71, and 1.26mm, respectively. At 1 s, the RMSEs are 4.40, 2.61, 3.34, 2.66, and 1.93 mm, respectively. It has been found that support vector regression can most accurately predict future tumor position and provide an RMSE of less than 2 mm at a 1s latency (10). Hamdi et al. (2018) estimated blood glucose levels with support vector regression (SVR) and differential evolution (DE) algorithms using continuous glucose monitoring (CGM) data. The proposed method was validated using real CGM data from 12 patients. The root mean square error (RMSE) obtained for the prediction horizon (PH) is 9.44, 10.78, 11.82, and 12.95 mg/dL, which is equal to 15, 30, 45, and 60 minutes, respectively. The results of the study show that the proposed method is promising (11). Chang et al. (2013) proposed an emotion recognition method that takes into account physiological signals. They designed a custom emotion induction experiment and collected five physiological signals of the subjects, including the electrocardiogram, galvanic skin responses (GSR), blood volume pulse, and pulse. They trained the trend curves of three emotions (sadness, fear, and pleasure) using support vector regression (SVR). Experimental results show that the proposed method achieves a high recognition rate of up to 89.2% (12).

Bergeron et al. (2005) aimed to obtain robust but non-linear relationships between inputs and outputs by choosing support vector regression (SVR) as a learning method because of its strong generalizability ability resulting from penalizing model complexity. In the study, laser scans of the trunk surface and reconstructions of spinal data from X-rays from scoliosis patients are functionally represented as surfaces and curves. There, the leading functional principal component coefficients established comprehensive features and provided sufficient dimensionality reduction for the estimation of the spine from the trunk. In this research application, the first robust estimate was obtained with the coefficients of determination for the leading outputs of 0.70 and 0.82, respectively, in the test set (13). Salat and Salat (2015) proposed a new and universal support vector regression (SVR)-based method for the analysis of drug interactions, which significantly accelerates isobolographic analysis. A theoretical model of the dose-effect relationship was constructed to simulate the various dose ratios of the two drugs by using the SVR. The model created can be used for any level of drug action and generate classical isobolograms to determine the nature of drug interactions (additivity, subadditivity, or synergy), which is particularly important in the case of novel compounds with high biological activity properties where the mechanism of action is unknown. The usefulness of this SVR method for modeling dose-effect relationships has been validated in a mouse model of toxic peripheral neuropathy induced by a single intraperitoneal dose of oxaliplatin. This method can also be applied during preliminary research on the mechanism of action of new compounds (14). Goli et al. (2016) proposed a new SVR model. This model is trained with SVR and the traditional Cox model with different cores. Evaluations were made using available medical datasets as well as a Breast Cancer (BC) dataset of 573 patients who visited the Oncology Clinic in Hamadan province, Iran. The results show

that the survival time for the BC dataset can be more accurately estimated with linear SVR compared to non-linear SVR. Based on the three feature selection methods, metastasis status, progesterone receptor status, and human epidermal growth factor receptor 2 status are the best features associated with survival. Also, when all features are included in the model, the SVR performs similarly or better than Cox (15). Han et al. (2021) proposed a fully automatic BAA method combining transfer learning and regression learning for Bone Age Assessment (BAA), which is an important issue in clinical practice in assessing children's biological maturity. First, the adaptive threshold segmentation method was used to segment the hand bone from the raw X-ray images, then the transfer learning technique was used to extract the high-level features of the hand bone images. Finally, support vector regression was applied to perform BAA. Experimental results show that the proposed method is more accurate than existing approaches and achieves better performance with lower root mean square error (RMSE) (16). Bono and Zorzi (2008) proposed a Support Vector machines-based regression (SVR)-method from functional Magnetic Resonance Imaging (fMRI) data to resolve cognitive states. In the context of the Pittsburgh Brain Activity Interpretation Competition (PBAIC, 2007), three participants were scanned during three 20-minute runs into a Virtual Reality Environment (VRE) in which they played a game that engaged them in various search tasks. A set of objective feature ratings were automatically extracted from the VRE during the screening session, while a set of subjective features were subsequently derived from each experience. The applicability of the SVR approach was explored in the case of a highly complex regression problem where the subjective experiences of participants immersed in a VRE had to be estimated from fMRI data. The results highlight the ability of nonlinear SVR to generalize, making this approach particularly interesting for the real-world application of Brain Computer Interface (BCI) (17). Zhang et al. (2014) developed a multivariate lesion-symptom mapping (MLSM) using a machine learning-based multivariate regression algorithm: Support vector regression (SVR). In the proposed SVR-LSM, the symptom associated with the whole lesion map, as opposed to each isolated voxel, is modeled using a nonlinear function, so intervoxel correlations are considered intrinsically, providing a potentially more sensitive way to examine lesion-symptom relationships. Both approaches were used in a synthetic dataset analysis to discover the relative values of VLSM and SVR-LSM. When applied to lesion data and language measures from patients with brain damages, SVR-LSM reproduced the baseline pattern of previous findings identified by the VLSM and showed higher sensitivity than VLSM for identifying lesion-behavior relationships (18). Seo et al. (2010) developed a support vector regression model for congenital muscular torticollis to investigate the prognosis of physical therapy in infants. Fifty-nine infants with congenital muscular torticollis received physical therapy until the degree of neck tilt was less than 5^0, and the mass diameter was re-evaluated after treatment. Support vector regression model was applied to the obtained data to predict prognoses. For the proposed model, 10, 20,

and 50 fold cross-tabulation analyzes were performed based on support vector regression and the traditional multiple regression method based on least squares. The developed support vector regression model is an effective prognostic tool for infants with congenital muscular torticollis and those undergoing physical therapy (19). Georga et al. (2013) discussed subcutaneous (s.c.) glucose estimation as a multivariate regression problem using support vector regression (SVR). Variables are the s.c. glucose profile, the plasma insulin concentration, the appearance of meal-derived glucose in the systemic circulation, and the energy expenditure during physical activities. Six cases corresponding to different combinations of the variables considered were used to investigate the effect of the input on the daily glucose prediction. The proposed method was evaluated using a dataset consisting of 27 patients in free-living conditions. To both optimize and test the SVR model, ten-fold cross-validation was applied to each dataset separately. When all input variables are taken into account, the mean prediction errors are 5.21, 6.03, 7.14, and 7.62 mg/dl for the 15-, 30-, 60-, and 120-min prediction horizons, respectively. The results clearly show that the availability of multivariate data and their effective combination can significantly improve the accuracy of both short-term and long-term predictions (20). Zhang et al. (2014) developed a new multivariate lesion symptom mapping (MLSM) method using support vector regression (SVR). Using lesion data, focal brain-behavior relationships were extensively evaluated by massive voxel-based lesion symptom mapping (VLSM). In SVR-LSM, a symptom associated with the entire lesion map rather than each isolated voxel is modeled using a nonlinear function, so intervoxel correlations are considered intrinsically, providing a potentially more sensitive way to examine the symptom-lesion association. Evaluations using synthetic and real data showed that the SVR-LSM outperformed the VLSM in detecting brain-behavior associations (in terms of sensitivity and specificity). While the method was designed for lesion analysis, it would be easy to expand to neuroimaging data (21).

The aim of this study was to examine the support vector regression method, which is a data mining method, and the relationship between the occipitocervical angle and the cervical vertical translation distance by using this method.

Support Vector Regression

Support Vector Machines (1), a machine learning approach with strong theoretical foundations, can be applied independently of the distribution; It has been proposed for classification and regression problems as supervised or semi-supervised (22). Although most studies with SVM are related to classification, SVM also gives very good results in regression applications (23).

A comprehensive study was conducted by Smola and Schölkopf (2004) on the use of SVM in regression problems (24). The SVM used in regression applications is called Support Vector Regression (SVR) (25). Support Vector Regression is a robust method that minimizes the effects of outlier observations on the regression equation (26). In SVM regression, training examples are $[(x_i - y_i)]_{i=1}^{N} = 1$ and the

response variable is defined as $y_i \in \mathbb{R}$ instead of $y_i \in \{+1, -1\}$ (27). The SVM regression method tries to find a function that approximates the training data by trying to minimize the estimation error (28). The aim here is to determine a line (function) that can take the maximum point in the range of a margin with the smallest error. Figure 1 shows the graph of the SVR.

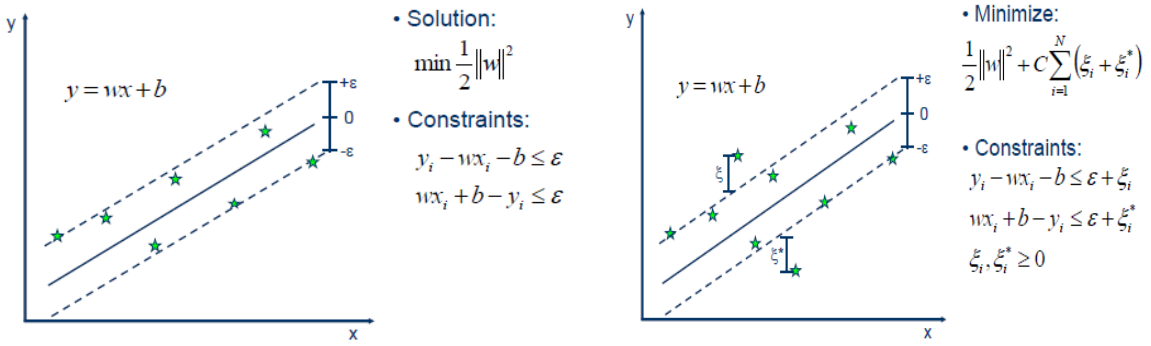

Figure 1: Linear support vector regression.
Source: (29).

SVM regression can generalize well for unsupervised test data by using the input-output relationship supervised during the training process. For the approximation problem of a continuous value function, $D=\{(x_i, y_i)|i\in\{1,2,\ldots,N\}\}$ N $x_i \in R$ n are n input vector, $y_i \in R$ are output (target) value. Regression analysis aims to accurately predict the desired outputs ($y_i \in R$) by specifying a mathematical function. Regression problems can be classified as linear and nonlinear (30).

According to the theory, the nonlinear relationship between input and output data should be represented by a linear function. This function is given as follows:

$$f(x) = w^T \Phi(x) + b$$

Where, f(x) is the estimation of the function, $\Phi(x)$ denotes a fixed feature-space transformation w ($w \in R^n$) and b ($b \in R$) are the weight and bias parameters, respectively (20). Regularization constant C>0 controls model complexity by determining how much of error will be allowed for training data that is out of the bound ε (19). ε is the threshold greater than or equal to 0. The standard form of SVM regression is defined as C>0 and ε>0 as follows:

$$Min_{w,b,\xi,\xi^*} = \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i + \xi_i^*$$

which is ssubject to

$$\begin{cases} w^T \Phi(x_i) + b - y_i \le \epsilon + \xi_i \\ y_i - w^T \Phi(x_i) - b \le \epsilon + \xi^* \\ \xi_i, \xi_i^* \ge 0; i = 1, 2, \ldots, l \end{cases}$$

where y_i is the target function value given x_i, $\xi_i$and ξ_i^* are slack variables to deal with the infeasible constraints (i.e., out of the ε precision) (31). ξ_i shows the training errors under ε, and ξ_i^* shows the training errors over ε (28). A non-linear support vector regression graph is given in Figure 2. Kernel functions are also used for nonlinear cases in support vector regression (32). Kernel functions transform data into a higher-dimensional feature space to enable linear separation.

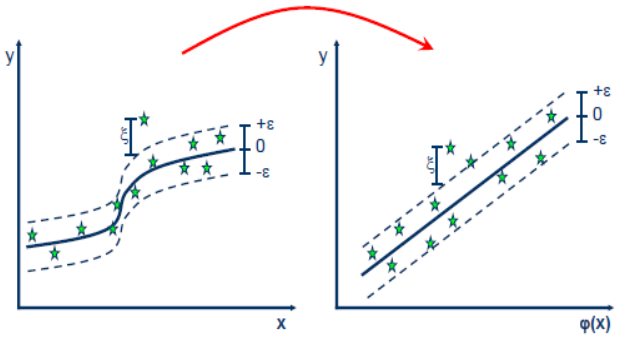

Figure 2: Nonlinear support vector regression.
Source: (29).

The estimation performance of the SVM regression model is related to the kernel function and parameter values. The essence of the support vector regression algorithm is the introduction of the kernel function. The kernel function has two aspects: the creation of the kernel function and the selection of the kernel function model. Model selection determines the kernel function that better fits the data characteristics of the original sample data before training. The kernel function consists of two steps: first, the type of kernel function is determined, and then the relevant parameters of the kernel function are selected. C and ε, the type and parameter of the kernel function must be carefully determined. Choosing the kernel function type and determining the most appropriate parameter value are very important (33). Selection of the appropriate model is key to improving the performance of support vector regression. Figure 3 shows the network structure of the SVM.

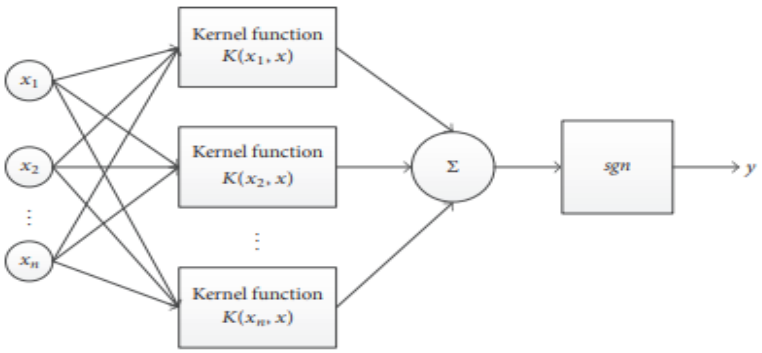

Figure 3: Diagram of the input and output of the SVR.
Source: (34).

The $K(x_i, x_j)$ seen in Figure 3 represents the kernel function. The kernel function *K* calculates the inner product of the input vectors. The SVM output value of the relevant example is equal to the sum of the combinations of these inner products and weights (35). Sigmoid, polynomial, and radial basis kernel functions are generally used in SVM (36). Equations of these functions are given in Table 1.

Table 1: Typical kernel functiony used in SVM.

Type of kernel	Kernel function
Sigmoid	$K(x_i, x_j) = tanh(\gamma(x_i.x_j) + c)$
Polynomial	$K(x_i, x_j) = (\gamma(x_i.x_j) + c)^d$
Radial basis function	$K(x_i, x_j) = exp(-\gamma\|x_i - x_j\|^2)$

Data and Results

In this study, the relationship between the occipitocervical angle and the cervical vertical translation distance was investigated using support vector regression. Occipitocervical Angle is the dependent variable, and Cervical Vertical Translation Distance is the independent variable. The data set was obtained by examining the measuring the vertical translation distances on lateral cervical direct radiographs (DGrfs) of a total of 175 adult individuals, 100 female and 75 male, aged between 18 and 65, whose cervical sagittal DGrfs were taken with various preliminary diagnoses in the Radiology Unit of Osmaniye State Hospital. Measurements on the obtained graphics were performed and evaluated with the PACS VIEWER (Picture Archiving and Communications System) software. Occipitocervical angle measurement, which evaluates the upper cervical region, and Cervical vertical translation distance measurements were made on DGrfs of 175 adult individuals using the Risser Ferguson method. For the study, necessary permissions were obtained from the Çukurova University Non-Interventional Clinical Research Ethics Committee, the Chief Physician of Osmaniye State Hospital, where the DGrfs were obtained and the measurements were carried out, and the Osmaniye Provincial Health Directorate. In this study, the relationship between the occipitocervical angle and the cervical vertical translation distance was investigated using support vector regression. Prediction models were created with linear regression (LR), SVR, tuned-SVR for different error and penalty cost values, and SVR using kernel functions. The prediction performances of the models were calculated using Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The formulas for these performance criteria are given in Table 2. Analyzes were made in R program using "e1071", "caret", "Metrics", and "ggplot2" packages.

Table 2: Performance measures and formulas.

MSE (Mean Square Error)	$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$
RMSE (Root Mean Square Error)	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$
MAE (Mean Absolute Error)	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert Y_i - \hat{Y}_i\rvert$
MAPE (Mean Absolute Percentage Error)	$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\lvert Y_i - \hat{Y}_i\rvert}{Y_i}$

The Occipitocervical Angle value is indicated by Y, the estimated Occipitocervical Angle value $\hat{Y}$, and the number of samples in the test subset is given as n. The variables of Occipitocervical Angle and Cervical Vertical Translation Distance used are explained as follows.

Occipitocervical Angle: It is the angle between the McGregor line and the line drawn parallel to the inferior plateau of the second cervical vertebra (C2). The McGregor line is the line joining the posterosuperior of the hard palate and the most caudal of the middle of the occipital curve (37). The visual of the occipitocervical angle is given in Figure 4.

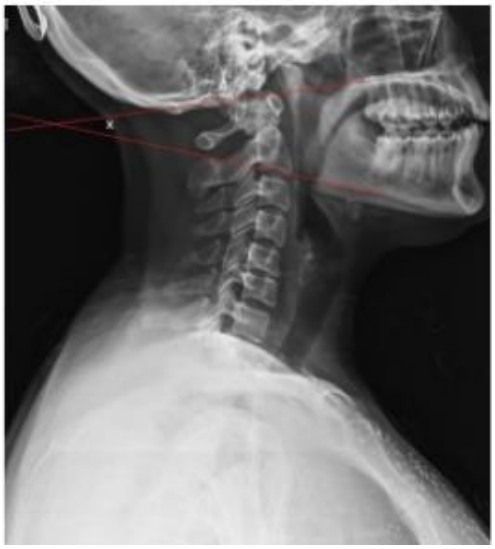

Figure 4: Occipitocervical angle.
Source: (38).

Measurement of Cervical Vertical Translation Distance: The position of the head in front and behind the head resulting from the translation of the head to the anterior or posterior can be evaluated by measuring the vertical translation distances on lateral cervical direct radiographs (DGrf) (39). Cervical vertical translation (CVT) can be calculated in several ways. The most commonly used method is the measurement of the distance between the vertical line passing through the center of the second cervical vertebra (C2) and the vertical line passing through the

posteriosuperior of the seventh cervical vertebra (C7) According to this method, if the distance is less than 15 mm or greater than 40 mm, there is cervical malalignment, and if the distance is greater than 40 mm, it means anterior translation of the head (40). The visual for the measurement of the cervical vertical translation distances is given in Figure 5. Descriptive statistics of the dataset are given in Table 3.

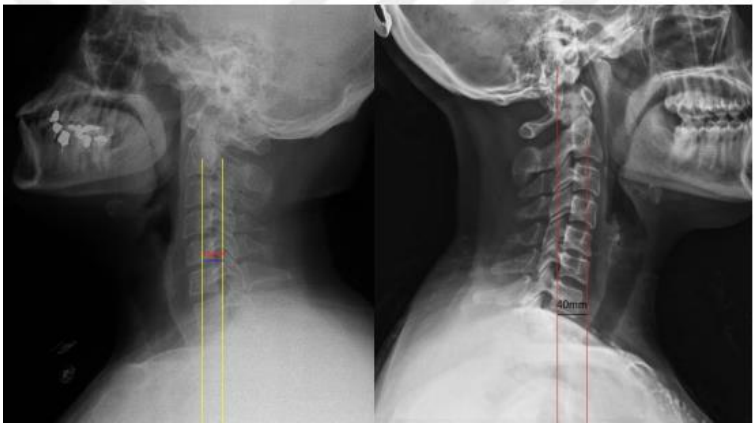

Figure 5: Measurement of cervical vertical translation distances
Source: (38).

Table 3: Descriptive statistics.

	n	Mean	Median	Mode	Std. Deviation	Minimum	Maximum
Occipitocervical Angle	175	22.28	23.10	26.20	7.33	5.20	40.50
Cervical Vertical Translation distance	175	22.70	22.20	20.40	7.28	4.20	44.30

Correlation analysis was conducted to determine whether there is a relationship between Occipitocervical Angle and Cervical Vertical Translation distance. According to the results of the analysis, it was determined that there was a moderately positive correlation between the Occipitocervical Angle and the Cervical Vertical Translation distance (r=0.552, p=0.000). The obtained relationship is modeled by using LR, SVR and tuned-SVR, and SVR with kernel functions. For SVR, tuned-SVR, and SVR with kernel functions, the data set is divided into 70% training 30% test set. The prediction models were created with SVR using SVR, tunned-SVR, and the kernel function of SVR, for different error and penalty cost values. A grid search algorithm was used to search for C, ε, and γ (the hyperparameters) which gave the best results. In the grid search algorithm, the parameters with the lowest error rate were selected. The training set was used to model the data and the test set was used to predict the data. In order to evaluate and compare the performances of the obtained prediction models, the performance criteria MSE, RMSE, MAE, and MAPE were calculated. The values of the performance criteria of the created prediction models are given in Table 4.

Table 4: Performance measures of prediction models for the test dataset.

	LR	**SVR**	**Tuned SVR** (Epsilon: 0.4 Cost: 4)	**SVR Radial** (Epsilon: 0.1 Cost: 4 Gamma = 1)	**SVR Sigmoid** (Epsilon: 0.1 Cost: 4 Gamma = 1)
MSE	33.93258	33.24458	37.33058	**30.53782**	48.24955
RMSE	5.825168	5.765811	6.109875	**5.526104**	6.946189
MAE	4.678624	**4.66205**	4.979137	**4.667553**	5.663636
MAPE	0.264095	**0.267374**	0.281040	**0.271438**	0.352433

When the performance criteria of all prediction models given in Table 4 are examined, the model with the lowest MSE (30.53782), and RMSE (5.526104) values is the radial-based SVR, while the model with the lowest MAE (4.66205), and MAPE (0.267374) values is the SVR model. However, considering that the MAE (4.667553), and MAPE (0.271438) values of the radial-based SVR model are very close to the MAE (4.66205), and MAPE (0.267374) values of the SVR model, it is seen that the radial-based SVR model is the most successful estimation model in estimating the Occipitocervical Angle. Since MSE, RMSE, MAE, and MAPE are error measures, the smaller these values, the better the model's performance (34).

When the models obtained were ranked according to their prediction performance, it was determined that the best model was the radial-based SVR model, while the worst model was the sigmoid-based SVR model. The values obtained for the five methods using all the data without dividing the data set are given in Table 5.

Table 5: Performance measures of prediction models for the data set.

	LR	**SVR**	**Tuned SVR** (Epsilon:0.4 Cost:4)	**SVR Radial** (Epsilon:0.1 Cost:4 Gamma=1)	**SVR Sigmoid** (Epsilon:0.1 Cost:4 Gamma=1)
MSE	34.5805	34.59394	33.67318	**30.40686**	54.22387
RMSE	5.880519	5.881661	5.80286	**5.514241**	7.363686
MAE	4.853072	4.660561	4.676705	**4.20588**	6.007143
MAPE	0.269215	0.252847	0.256133	**0.24724**	0.373623

As seen in Table 5, it is seen that the most suitable model obtained using all data is the radial-based SVR method with the lowest MSE RMSE, MAE, and MAPE values. The estimation model obtained against the data points is given in Figure 6.

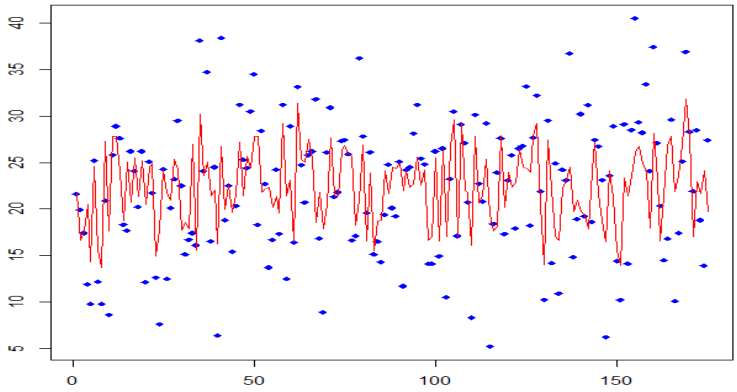

Figure 6. Data points and predicted values.

Figure 6 shows the results of prediction for the Occipitocervical Angle indices achieved by the aforementioned SVR with radial basis kernel. There are 175 data points in the dataset, which is plotted Cervical Vertical Translation distance on the x-axis. The y-axis represents the Occipitocervical Angle value. The dotted blue and red lines are the target and predicted values respectively.

Conclusions

SVM, one of the most used machine learning algorithms, is generally used in classification and regression processes. It is based on structural risk minimization and statistical learning theory and has good generalization ability. SVM can work with both numerical and categorical variables, and it can produce successful results even in small samples and in the case of outliers (41). The power and popularity of SVM are largely due to its ability to achieve high accuracies with balanced performance that can be generalized even when the size of the feature space greatly exceeds the number of sample observations available for training (42). Support Vector Machines generate a model that predicts target values of data samples in a given test set. Support Vector Machines are one of the best approaches in data modeling (43). It produces accurate and powerful results by using various kernel functions, especially in cases where the data cannot be separated linearly (44). Choosing these functions correctly is very important for the method to give good results (45). Difficulty in deciding which function to choose can be considered as a disadvantage of SVM.

In this study, support vector regression was considered for medical data analysis, and the relationship between occipitocervical angle and cervical vertical translation distance was examined. Knowing the contour of the cervical part of the columna vertebralis in the sagittal plane and the borders of the cervical lordosis is important and necessary to evaluate pathological conditions. The cervical part of the columna vertebralis has a lordotic structure to balance the thoracic kyphosis. This lordotic structure, which is a warning for the development of Luschka joints, is important for the smoothness and integrity of the cervical region. The normal form of the cervical alignment is important for the loads on facet joints and discs, segmental kinematics, normal range of motion, and pain in the cervical region. Changes in the occipitocervical angle and translations cause the head forward posture place an excessive load on the muscles and joints.

For this reason, they can cause postural disorders that cause neck pain. Predicting and detecting angular changes in cervical posture will be useful in eliminating existing complaints, preventing further pathological conditions, and reviewing the efficacy of treatments. Correlation analysis was performed to determine whether there is a relationship between the occipitocervical angle and the cervical vertical translation distance. According to the obtained correlation analysis results, it was concluded that there is a moderate positive relationship between the occipitocervical angle and the cervical vertical translation distance. To create the model, prediction models were created using linear regression, SVR, tuned-SVR for different error and penalty cost values, and SVR using kernel functions. The appropriate model was determined by using the performance evaluation criteria MSE, RMSE, MAE, and MAPE. From the results obtained, it has been seen that the SVR model made using the radial basis kernel function is the appropriate estimation model.

In the studies conducted, no study was found that evaluated the regression relationship between the cervical vertical translation distance and the occipitocervical angle, which evaluates the upper cervical lordosis. Therefore, in the light of the available data, it is thought that the findings obtained in the study will be useful in understanding the columna vertebralis morphometry in the field of anatomy, radiology, cervical surgery and will contribute to the literature.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgement

No institution has given financial support to the study. All researchers contributed equally to the study.

References

1. Vapnik V. Estimation of Dependences Based on Empirical Data. Springer Verlag. 1982.
2. Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural processing letters. 1999;9(3):293-300.
3. Arat MM. Destek Vektör Makineleri Üzerine Bir Çalışma. Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Yüksek lisans tezi. Ankara, 2014.
4. Alan M. Tıbbi veriler üzerine birliktelik kuralları madenciliği. Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi. 2019;20(1):410-419.
5. Çataloluk H, Kesler M. A diagnostic software tool for skin diseases with basic and weighted K-NN. In 2012 international symposium on innovations in intelligent systems and applications. 2012, July; pp. 1-4. IEEE.
6. Kocamaz K. Hastane Yönetim Bilgi Sistemlerinde Veri Madenciliği, Selçuk Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi. Konya, 2007.
7. Kaya E, Bulun M, Arslan A. Tıpta Veri Ambarları Oluşturma ve Veri Madenciliği Uygulamaları. Akademik Bilişim. Adana, 2003.
8. Georga EI, Protopappas VC, Ardigo D, et al. Multivariate prediction of subcutaneous glucose concentration in type 1 diabetes patients based on support vector regression. IEEE journal of biomedical and health informatics. 2012; 17(1): 71-81.
9. Guo X, Zhou W, Shi B, et al. An efficient multiple kernel support vector regression model for assessing dry weight of hemodialysis patients. Current Bioinformatics. 2021;16(2): 284-293.
10. Riaz N, Shanker P, Wiersma R. et al. Predicting respiratory tumor motion with multi-dimensional adaptive filters and support vector regression. Physics in Medicine & Biology. 2009;54(19):5735.
11. Hamdi T, Ali J B, Di Costanzo V. et al. Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm. Biocybernetics and Biomedical Engineering. 2018;38(2):362-372.
12. Chang CY, Chang CW, Zheng JY, et al. Physiological emotion analysis using support vector regression. Neurocomputing. 2013; 122, 79-87.

13. Bergeron C, Cheriet F, Ronsky J, et al. Prediction of anterior scoliotic spinal curve from trunk surface using support vector regression. Engineering applications of artificial intelligence. 2005; 18(8): 973-983.
14. Sałat R, Sałat K. Modeling analgesic drug interactions using support vector regression: a new approach to isobolographic analysis. Journal of pharmacological and toxicological methods. 2015;71: 95-102.
15. Goli S, Mahjub H, Faradmal J, et al. Survival prediction and feature selection in patients with breast cancer using support vector regression. Computational and mathematical methods in medicine, 2016;1-13.
16. Han J, Jia Y, Zha C. et al. Automatic bone age assessment combined with transfer learning and support vector regression. In 2018 9th International Conference on Information Technology in Medicine and Education (ITME). 2018, October; pp. 61-66. IEEE.
17. Di Bono MG, Zorzi M. Decoding Cognitive States from fMRI Data Using Support Vector Regression. PsychNology Journal. 2008;6(2).
18. Zhang Y, Kimberg DY, Coslett HB. et al. Multivariate lesion-symptom mapping using support vector regression. Human brain mapping. 2014;35(12):5861-5876.
19. Seo ST, Lee IH, Son CS. et al. Support vector regression-based model to analyze prognosis of infants with congenital muscular torticollis. Healthcare informatics research. 2010;16(4):224-230.
20. Georga EI, Protopappas VC, Ardigo D. et al. A glucose model based on support vector regression for the prediction of hypoglycemic events under free-living conditions. Diabetes technology & therapeutics. 2013;15(8):634-643.
21. Zhang Y, Kimberg DY, Coslett HB. et al. Support vector regression based multivariate lesion-symptom mapping. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2014, August; pp. 5599-5602. IEEE.
22. Çomak, E. Destek vektör makinelerinin etkin eğitimi için yeni yaklaşımlar. Yayınlanmamış Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik Elektronik Mühendisliği Anabilim Dalı, Doktora Tezi. Konya, 2008.
23. Bilişik MT. Destek Vektör Makinesi, Çoklu Regresyon Ve Doğrusal Olmayan Programlama İle Perakendecilik Sektöründe Gelir Yönetimi İçin Dinamik Fiyatlandırma. XI. Üretim Ara t Jrmalar J Sempozyumu, 23-24 Haziran 2011;785-799.
24. Smola AJ. Schölkopf BA. Tutorial on Support Vector Regression, Statistics and Computing. 2004;14:199- 222.
25. Güner N, Çomak E. Mühendislik öğrencilerinin matematik I derslerindeki başarısının destek vektör makineleri kullanılarak tahmin edilmesi. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi. 2011;17(2): 87-96.
26. Meral G, Saraçlı S. Destek Vektör Makineleri ve Türkiye'deki Enerji Santrallerinde Doğal Gaz Tüketimi Üzerine Bir Uygulama. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi. 2020;24(2):411-418.
27. Daş M, Balpetek N., Akpınar EK. et al Türkiye'de bulunan farklı illerin rüzgâr enerjisi potansiyelinin incelenmesi ve sonuçların destek vektör makinesi regresyon ile tahminsel modelinin oluşturulması. Journal of the Faculty of Engineering & Architecture of Gazi University. 2019;34(4):2203-2214.
28. Çoban F, Demir L. Yapay Sinir Ağları ve Destek Vektör Regresyonu ile Talep Tahmini: Gıda İşletmesinde Bir Uygulama. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi. 2021;23(67):327-338.
29. https://www.saedsayad.com/support_vector_machine_reg.htm (Last access date: 10.09.2021)
30. Karal Ö. Compression of ECG data by support vector regression method. J. Fac. Eng. Arch. Gazi Univ. 2018;1:743-756.
31. Park J, Kim KY, Kwon O. Comparison of machine learning algorithms to predict psychological wellness indices for ubiquitous healthcare system design. In Proceedings of the 2014 International Conference on Innovative Design and Manufacturing (ICIDM). 2014, August. pp. 263-269. IEEE.
32. Yakut Y, Yakut E, Yavuz S. Yapay Sinir Ağlari Ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeksi Tahmini. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. 2014;19(1):139-157.
33. Baydaroğlu Ö, Koçak K. SVR-based prediction of evaporation combined with chaotic approach. Journal of Hydrology. 2014;508, 356–363.
34. Wang H, Xu D. Parameter selection method for support vector regression based on adaptive fusion of the mixed kernel function. Journal of Control Science and Engineering. 2017.
35. Arslan A, Şen B. Detection of non-coding RNA's with optimized support vector machines. In 2015 23nd Signal Processing and Communications Applications Conference (SIU). 2015, May; pp. 1668-1671. IEEE.
36. Açıkkar M, Sivrikaya O. Yıkanmış Türk Linyit Kömürlerinin Üst Isıl Değerinin Destek Vektör Regresyonu ile Tahmini. Avrupa Bilim ve Teknoloji Dergisi. 2020;18:16-24.
37. McGregor M. The significance of certain measurements of the skull in the diagnosis of

basilar impression. Br J Radiol. 1948;21(244):171-81.

38. Tunçeli, M. Cervical Lordoz Derecesini belirlemede kullanılan morfometrik ölçüm yöntemlerinin karşılaştırmalı olarak incelenmesi. Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü Anatomi Anabilim Dalı, Yüksek lisans tezi. Adana, 2021.

39. Aksoy E. Servikal Dizilim Bozukluğu İle Kronik Boyun Ağrısı ve Tetik Nokta Arasındaki İlişki., İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Fiziksel Tıp ve Rehabilitasyon Anabilim Dalı, Uzmanlık Tezi. İstanbul, 2015.

40. Scheer JK, Tang JA, Smith JS, Acosta Jr FL, Protopsaltis TS, Blondel B, et al. Cervical spine alignment, sagittal deformity, and clinical implications: a review. J Neurosurg: Spine. 2013;19(2):141-59.

41. Kuzu BS, Yakut SG. Destek Vektör Makineleri Yardımıyla İmalat Sanayisinde Mali Başarısızlık Tahminlerinin Teknoloji Yoğunluğuna göre İncelenmesi. Osmaniye Korkut Ata Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. 2020;4(2): 36-54.

42. Pisner DA. Schnyer DM. Support vector machine. In Machine Learning. Academic Press.2020; 101-121.

43. Jakkula V. Tutorial on support vector machine (svm). School of EECS, Washington State University. 2006;37.

44. Pochet NLMM, Suykens JAK. Support vector machines versus logistic regression: improving prospective performance in clinical decision-making. Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2006;27(6):607-608.

45. Meyer D, Wien FT. Support vector machines. The Interface to libsvm in package e1071. 2015;28.