Avrupa Bilim ve Teknoloji Dergisi
Sayı 21, S. 610-620, Ocak 2021

Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
No. 21, pp. 610-620, January 2021

Research Article

A Comparison of Various Supervised Machine Learning Techniques for Prostate Cancer Prediction

Ebru Erdem[1], Ferhat Bozkurt[2*]

[1*] Atatürk University, Faculty of Engineering, Department of Computer Engineering, Erzurum, Turkey (ORCID: 0000-0002-4042-7549), ebruerdem@atauni.edu.tr
[2] Atatürk University, Faculty of Engineering, Department of Computer Engineering, Erzurum, Turkey (ORCID: 0000-0003-0088-5825), fbozkurt@atauni.edu.tr

(First received 30 September 2020 and in final form 31 January 2021)

(**DOI:** 10.31590/ejosat.802810)

ATIF/REFERENCE: Erdem, E. & Bozkurt, F. (2021). A Comparison of Various Supervised Machine Learning Techniques for Prostate Cancer Prediction. *European Journal of Science and Technology*, (21), 610-620.

Abstract

Prostate cancer is a kind of cancer that is seen worldwide and causes death of many people. Early diagnosis of cancer helps patients during the treatment phase. For this reason, cancer prediction is very crucial, according to the symptoms seen in the patient. One of the biggest problems in medicine is diagnosing diseases. The absence of certain definitive rules for the evaluation of symptoms of prostate cancer and the low rate of prediction of the diagnostic methods currently in effect made this study essential. It is thought that machine learning methods can be effective for the solution of the problems where there are no specific and definite rules and the factors affecting the event can be predicted. With this awareness, various solutions are developed by computer-aided systems. In this paper, we compare and discuss the performance of different supervised machine learning algorithms (i.e., k-nearest neighbor, support vector machines, random forest, logistic regression, linear regression, Naive Bayes, linear discrimination analysis, linear classification, multi-layer perceptron and deep neural network) for prostate cancer prediction. In this study, an open-access online prostate cancer data which consists of observations of 100 patients is used. The main intention is to evaluate the correctness in classifying data with respect to effectiveness and efficiency of each algorithm in terms of precision, recall, AUC, F1-Score, accuracy. The accuracy of the methods may vary according to the training and test data. In order to obtain more stable results, each algorithm was run more than ten times and their five best performances were recorded. The results show that multi-layer perceptron (MLP) can result in high prediction accuracy that is better compared to other approaches. Experimental results show that MLP gives the highest accuracy (97%) with the lowest error rate (0.03). The MLP classifier outperformed the other algorithms used in this study and is one of the best studies ever reported in the literature in terms of accuracy, AUC and F1 score performance criteria. As a result, we can say that if the computer is trained with machine learning methods based on patient information, it can be clinically useful with high accuracy in predicting cancer. In this way, an unnecessary biopsy of the patient can be prevented.

Keywords: Prostate cancer, supervised machine learning, artificial neural network, deep learning, classification, performance, effectiveness, efficiency

Prostat Kanseri Tahmini için Çeşitli Denetimli Makine Öğrenimi Tekniklerinin Karşılaştırılması

Öz

Prostat kanseri dünya genelinde yaygın olarak görülen ve ölüme yol açan kanser türlerinden biridir. Kanserin erken teşhisi hastaların tedavi aşamasında yardımcı olmaktadır. Bu sebeple, hastada görülen belirtilere göre kanser tahmini büyük önem taşımaktadır. Sağlık alanında en büyük sorunlardan biri hastalığı teşhis etmektir. Prostat kanseri semptomlarının değerlendirilmesi için belirli kesin

[*] Corresponding Author: fbozkurt@atauni.edu.tr

kuralların olmaması ve şu anda yürürlükte olan tanı yöntemlerinin düşük öngörü oranı bu çalışmayı gerekli kılmıştır. Belirli ve kesin kuralların bulunmadığı ve olayı etkileyen faktörlerin öngörülebildiği sorunların çözümünde makine öğrenimi yöntemlerinin etkili olabileceği düşünülmektedir. Bu farkındalığın bilinci ile bilgisayar destekli sistemler tarafından çeşitli çözümler geliştirilmektedir. Bu çalışmada, prostat kanserinin tahmini için çeşitli denetimli makine öğrenme algoritmalarının (destek vektör makineleri, rastgele orman, k-en yakın komşu, lojistik regresyon, doğrusal regresyon, Naive Bayes, doğrusal ayrımcılık analizi, doğrusal sınıflandırma, çok katmanlı algılayıcılar ve derin yapay sinir ağları gibi) performansını karşılaştırır ve tartışırız. Bu çalışmada 100 hastanın gözlemlerinden oluşan açık erişimli çevrimiçi prostat kanseri verisi kullanılmıştır. Temel amaç her algoritmanın verilerin sınıflandırılmasındaki doğruluğunu, etkinlik ve verimlilik açısından hassasiyet, recall, AUC, F1-Score ve doğruluğa göre değerlendirmektir. Yöntemlerin doğruluğu, eğitim ve test verilerine göre değişebilir. Daha istikrarlı sonuçlar elde etmek için, her bir algoritmayı 10'dan fazla çalıştırdık ve en iyi 5 performansını kaydettik. Sonuçlar çok katmanlı algılayıcının (MLP), diğer yaklaşımlara göre göre daha iyi olan yüksek tahmin doğruluğu ile sonuçlanabildiğini göstermektedir. Deneysel sonuçlar, MLP'nin en yüksek doğruluğu (%97) ve en düşük hata oranını (0.03) verdiğini göstermektedir. MLP sınıflandırıcısı, bu çalışmada kullanılan diğer algoritmalardan daha iyi performans gösterdi ve doğruluk, AUC ve F1 puan performans kriterleri açısından literatürde bildirilen en iyi çalışmalardan biridir. Sonuç olarak, bilgisayarın hasta bilgilerine dayanarak makine öğrenmesi yöntemleri ile eğitilmesi durumunda, kanseri tahmin etmede yüksek bir doğrulukla klinik olarak yararlı olabileceğini söyleyebiliriz. Böylece hastaya gereksiz bir biyopsi önlenebilir.

Anahtar Kelimeler: Prostat kanseri, denetimli makine öğrenmesi, yapay sinir ağları, derin öğrenme, sınıflandırma, performans, etkinlik, verimlilik.

1. Introduction

Cancer cells are abnormal cells that grow faster than normal and refuse to die (Chang et al., 2014). Cancer could be seen in many organs such as lung, skin, stomach, large intestine, breast, and prostate. Cancer formation is divided into two as benign and malignant. In benign growth, while only tissue is growing, in malignant growth if the organ is not exposed to an early diagnosis-treatment process, it does nonfunctional the organ that is contaminated with. And if this situation sustains, it can result in the death of the person. Cancer gives different symptoms in each organ. When these symptoms are considered, with various treatment processes it is possible to get rid of the organ infected, to prevent its spread and to rescue the person's life (Grönberg, 2003). For this reason, early diagnosis of cancer disease has made it one of the important problems in medicine (Cuzick et al., 2014). In prostate cancer, some patients are more prone to the potential to turn into a fatal phenotype. In some patients, there is a healing situation with the treatment of the main tumor. Accurate follow-up of the patient is the most important factor during the disease progression. The patients at high risk should be identified in order to develop more effective treatment paradigms. The classification should be done for diagnosis (Chang et al., 2014).

Although there are various diagnostic methods of prostate cancer, the success of these methods is not acceptable. In addition, there is no specific formula for interpreting the data obtained from the hospital. In the diagnosis of prostate cancer, cancer is suspected by examining the PSA value which is a blood marker and the FPSA / PSA ratio (Stephan et al., 2002). In the findings of TRUS (Transrectal Ultrasonography), PRI (Finger Rectal Examination), suspicion increases with the presence of a nodule and the individual can be diagnosed by performing a biopsy several times (Taşkıran, 2008). In addition, prostate cancer risk can be calculated with methods such as Online Calculator which is developed as a result of examining many patients' information (Ankerst et al., 2014). However, the real diagnosis can only be made as a result of biopsy.

In this study, prostate cancer was studied which is one of the most cancer-related deaths of men and whose symptoms have similar characteristics with benign growth. One of the biggest problems in medicine is to diagnose diseases. The absence of certain definitive rules for the evaluation of symptoms of prostate cancer and the low rate of prediction of the diagnostic methods currently in effect made this study essential. It is thought that machine learning methods can be effective for the solution of the problems where there are no specific and definite rules and the factors affecting the event can be predicted. The aim of this study is to use and compare various supervised machine learning algorithms like Support Vector Machines (SVM), Random Forest (RF), k-Nearest Neighbor (kNN), Logistic Regression (LR), Linear Regression (LR), Naive Bayes (NB), Linear Discrimination Analysis (LDA), Linear Classification (LF), Multi-Layer Perceptron (MLP) and Deep Neural Network (DNN) to predict prostate cancer. We think that the success of a system in which the computer can decide based on patient information may be higher in the prediction of cancer. In this system, the program will learn the relationship between parameters and be able to diagnose prostate cancer. Thus, an unnecessary biopsy can be prevented to patient with the success of this study.

The rest of the paper organized as follows. In Section 2, we introduced the related works in literature. In Section 3, material and method are introduced. In section 3, experimental results and findings are discussed. Finally, Section 4 concludes the paper and presents some future work opportunities.

2. Related Works

Classification is one of the most crucial and essential task in machine learning and data mining (Zaki & Meira, 2019). Many studies have been performed to implement data mining and machine learning to different medical data sets to classify prostate cancer (Goldenberg et al., 2019; Ray, 2020). Srivenkatesh (2020) made a prediction of prostate cancer on the Kaggle data set using machine learning techniques. The accuracy results that obtained from kNN, SVM, LR, NB and RF algorithms were between 70% and 90%. The highest accuracy (90%) was taken from the LR and RF algorithms (Srivenkatesh, 2020). Laabidi and Aissaoui (2020) compared popular machine learning algorithms to predict results on the same Kaggle dataset for prostate cancer as Srivenkatesh was done. The highest result is obtained with the Recurrent Neural Network (RNN) method with an accuracy of 81.3% when compared to other methods (Laabidi & Aissaoui, 2020). Zhang et al. (2006) achieved 92% accuracy by using the SVM classifier. Depending on the correlation between variables, it was observed that performance

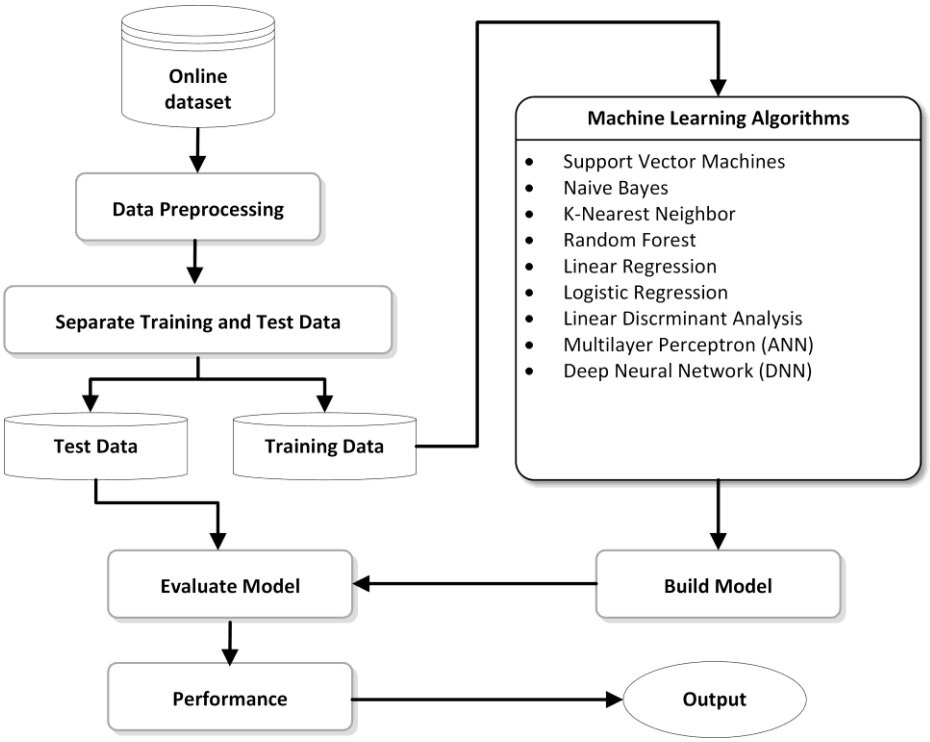

Figure 1. The workflow of the performance analysis of various supervised machine learning algorithms

could be improved by developing variable selection methods (Zhang et al., 2006). Shakeel and Manogaran (2020) used the DBCR dataset which includes biomedical details such as patient age, cancer volume, Gleason score, prostate weight, and antigen. Preprocessing was performed by applying standardization procedures on the data. When the accuracy rates of the study were analyzed, it was seen that radial trained neural networks-RTELNN (99.3%), SVM(95.90%), Neural Networks (97.12%), and MLP (98.15%) are obtained.

Recently, the success of deep learning has been proved in the prediction of prostate cancer. Diagnosis by experts was made from H&E stained samples with the help of light microscopy (Arvidsson et al., 2018). In the study conducted by Arvidsson et al., (2018) several different approaches based on the CNN were presented to three different datasets and the results were compared. Data augmentation operations were carried out using digital spot separation and color diversification methods. In the results obtained, autoencoder was seen as the method of providing the best generalization. Benign and malignant tissue was detected with a 95% accuracy rate, and 81% accuracy was detected in the Gleason rating (Arvidsson et al., 2018). Yuan et al. (2019) performed the classification of prostate cancer on MRI images using a deep learning-based multi-parametric transfer learning model. Their model obtained high accuracy with %86.92 on classification of prostate cancer. According to their comparison results, their methods show that they perform better in prostate cancer classification with higher accuracy than both existing deep learning models and hand-crafted feature-based methods (Yuan et al., 2019). Abraham and Nair (2018) used a deep network of stacked sparse automatic encoders (SSAE). They extracted high-level features from handcrafted tissue features and classified the MRI images using a softmax classifier (SMC). They achieved 47.3% accuracy by using this model (Abraham & Nair, 2018). The classification of prostate cancer was performed using multi-parametric MR images by Li et al. (2018). 10-fold cross-validation was used for the model. As a result, SVM classification using mpMRI-derived image properties has provided accurate and automatic differentiation of low and high grade prostate cancer in central gland (Li et al., 2018). The classification of Epstein gleason score can significantly reduce healing in prostate cancer patients. In prostate cancer, biopsy accuracy for each histological classification can be found using the kappa coefficient and by evaluating specificity, sensitivity, negative, and positive predictive value (De Nunzio et al., 2018). Song et al. (2018) developed a computer-aided diagnosis of prostate cancer by using a deep convolutional neural network for MRI images. The dataset was obtained from the PROSTATEx database which contains 195 patients' data. Normally, a radiologist manually labels the areas related to PCas and NCs while viewing the prostate. Song et al. (2018) designed the DCNN model which is inspired by VGG-Net to distinguish between PCas and NCs based on the mp-MRI data combination. The superior performance was observed between PCa and NC and 94.4% AUC was obtained in the test data via this model. Their model obtained superior performance between PCa and NC for test data with 94.4% AUC (Song et al., 2018). The Gleason score for prostate cancer patients is one of the most important prognostic factors that potentially determine treatment. A deep learning algorithm is developed by Nagpal et al. (2019) to perform quantitation and Gleason scoring in prostate cancer diagnosis. When it is compared to a reference standard provided by pathologists of genitourinary, the 0.61 average accuracy was obtained thanks to 29 general pathologists in the validation set. DLS achieved a significantly higher diagnostic with accuracy of 0.70 (Nagpal et al., 2019). Reda et al. (2018) proposed a fully automated convolutional neural network-based CAD model for early detection of prostate cancer. The input of the proposed system is 3D-DWI datasets that are collected from 45 patients (25 malignant, 20 benign) at seven distinct b-values. According to the accuracy rates obtained from distinct b-values, it shows the high performance with 97.8% accuracy through CNN based classifier (Reda et al., 2018).

Table 1. Statistical analysis of the dataset

	radius	texture	perimeter	area	smoothness	compactness	symmetry	fractal_dimension
count	100	100	100	100	100	100	100	100
mean	16.85	18.23	96.78	702.88	0.102730	0.126700	0.193170	0.064690
std	4.87	5.19	23.67	319.71	0.014642	0.061144	0.030785	0.008151
min	9	11	52	202	0.07	0.038	0.135	0.053
max	25	27	172	1878	0.143	0.345	0.304	0.097

3. Material and Method

The design flow diagram of this study, which was carried out for the prediction of prostate cancer through various supervised machine learning methods, is given in Figure 1. Each step in this diagram is detailed in the following subsections.

3.1. Dataset

In order to analyze the performance of the methods evaluated in this study, a prostate cancer dataset was used as shown in first step of the design flow. Prostate cancer dataset is an open-access data source that can be obtained through the Kaggle platform (Sajid, 2018). The dataset consists of observations from 100 patients. The dataset consists of eight independent variables (radius, texture, area, perimeter, compactness, smoothness, fractal dimension, symmetry) and one dependent variable (diagnosis result). The variables used as predictors are as follows: 1-Radius, 2-Perimeter, 3-Texture, 4-Smoothness, 5- Area, 6- Compactness, 7-Symmetry, 8-Fractal dimension and 9-Diagnosis result. The output response takes two values: “B” for benign tumors and “M” for malignant tumors. A statistical analysis containing detailed information about the data is given in Table 1.

3.2. Preprocessing

As shown in Table 1, some high-valued features and lower valued ones are present together in the data set. In this context, the data needs some transformations because the lower-valued features will have no effect compared to the high-valued ones. The data (as shown in Figure 2) which is highly variable in terms of size will increase the variance and will not have an equal effect in distance calculations, since it will outweigh the low-value features in weight calculations. In order to normalize the data, it is necessary to convert the values of the numeric columns in the data set to a common scale without disturbing the differences in the range of values in the preprocessing stage (Gültepe, 2019). Therefore, in normal cases, a scaling process is performed to reduce the distances between different dataset values. Similarly, non-numeric properties should not be processed in order to perform data transformations. In order to do data mining, non-numerical data must be converted to numerical data. The diagnosis_result column, which is categorical and represents the dependent variable in the dataset, has been converted to binary. These values need to be scaled to the same standard in order to guarantee that all values equally contribute to the success of the algorithms used in this study. To ensure standardization, the data in the dataset was scaled by feature standardization. For this purpose, the StandardScaler method of sklearn.preprocessing library was used in Python. In this process, the raw data is converted into a standardized value score with 1 standard deviation and 0 arithmetic mean.

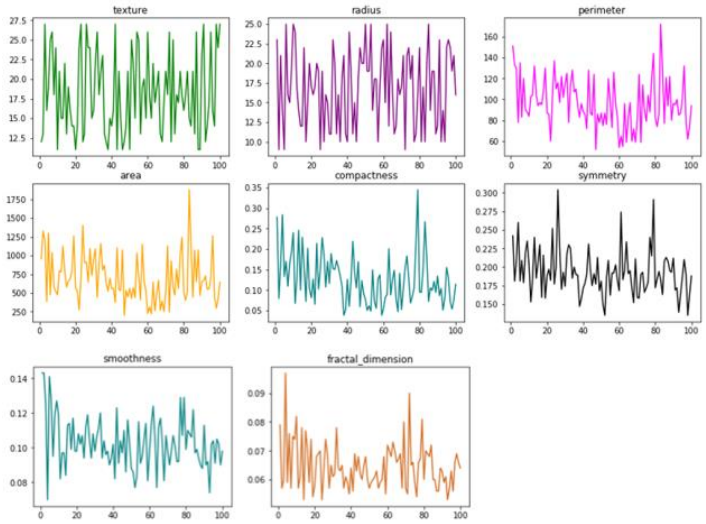

Figure 2. The clinical features and general distribution of data

3.3. Machine Learning Techniques

Within the scope of this study, nine popular supervised machine learning techniques were evaluated to compare their performances on prostate cancer dataset. Thus, we implemented Naive Bayes (NB), Logistic Regression, K-Nearest Neighbor (K-NN), Support Vector Machines (SVM), Linear Regression, Random Forest (RF), Linear Discrimination Analysis (LDA), Multi-Layer Perceptron (MLP), and Deep Neural Network (DNN) techniques by comparing their performance with accuracy, ROC-AUC and F-measure as metric. These techniques are evaluated to analyze the effectiveness of different machine learning approaches on the same dataset. These algorithms are preferred because they are easy to implement and can produce good results in terms of performance.

3.3.1. Support Vector Machines

It is a popular classifier with a quadratic programming technique that can bring a high level of solution to model complexity. In SVM, data is first transformed into the feature space using a suitable kernel function. In that space, the classification is made through the hyper plane. Statistical learning theory constitutes the infrastructure of SVMs. SVM also can be evaluated in the feed-forward network category (Chang et al., 2000; Erickson et al., 2017).

3.3.2. Naive Bayes

Naive Bayesian algorithm is a basic statistical classifier. For given dataset, this classifier calculates a probability set by counting combinations of values and frequency (Aydın, 2018). This algorithm supposes that all variables are independent. And it assumes no dependency between attributes. This hypothesis of conditional independence is seldomly right in real-world

applications, so it qualifies as Naive, however the algorithm prone to learn quickly in various controlled classification problems (Saritas & Yasar, 2019). It can be said that NB is particularly better than other approaches in different cases for especially medical datasets in the literature (Al-Aidaroos et al., 2012; Laabidi & Aissaoui, 2020).

3.3.3. K-Nearest Neighbor

The grouping method proposed by Cover and Hart, in which the group containing the sample data point and the closest neighbor to this data point are determined according to the value of k, is called the K-NN algorithm (Cover & Hart, 1967). K-NN is a supervised learning algorithm that solves the grouping problem. It is one of the most frequently used classification algorithms in the literature. Classification of data in K-NN algorithm is performed using data whose class is known. Since it is an example-based algorithm, the learning process is carried out based on the data found in the training set (Yağanoğlu et al., 2014). In spite of its simplicity, K-NN yields competitive results and even outperforms other complex learning algorithms in some cases. This method is a simpler but more effective machine learning method than others, especially for a smaller number of classes (Karakoyun & Hacıbeyoğlu, 2014). The most important advantage of this method is that successful studies can be made in the classification process with multiple categorized data points. K-NN could be preferred in classification as well as in the solution of regression problems. It is used when independent variables are quantitative and the classification process is applied depending on the distances between observations. Although it has a quite simple structure, it has a high computational cost (Taşcı & Onan, 2016). In this study, the neighborhood value (k) that is the best appropriate to the relevant dataset is determined first. The k value is chosen to be between 3 and 10 for most data sets (Çebi et al., 2019). For this purpose, training data is utilized and the best neighborhood value for the related dataset is determined as 5. Moreover, distance calculation in the algorithm is performed with Euclidean distance.

3.3.4. Random Forest

Random Forest is a supervised machine learning algorithm that could be preferred in classification as well as in the solution of regression problems. RF classifier is a bagging based ensemble learning method (Alqaraleh, 2020). This algorithm was proposed by Breiman (2001) as a combination of "Bagging" and "Random Subspace" methods (Breiman, 2001). RF classifier is frequently preferred in the literature for problems as high-dimensional and complex data, mixed categorical and numerical variables, nonlinear relationships, non-Gaussian statistical distributions, etc. RF branches all nodes using the best of randomly selected variables in each node instead of branching each node using the best branch of all variables. The algorithm creates a decision forest by using decision trees. It creates a random tree community / forest, then more than one decision tree is trained to make the most accurate classification during the operation (Pervan & Keleş, 2019). Most of the time, it can give superior results even without the use of hyperparameters. Since it gives very fast and good results even in complex and noisy data sets, it is frequently preferred to obtain useful information in medicine. The RF classifier performance has shown that this classifier can be used for some problems in medicine and could be an appropriate technique to obtain useful information (Alickovic & Subasi, 2016).

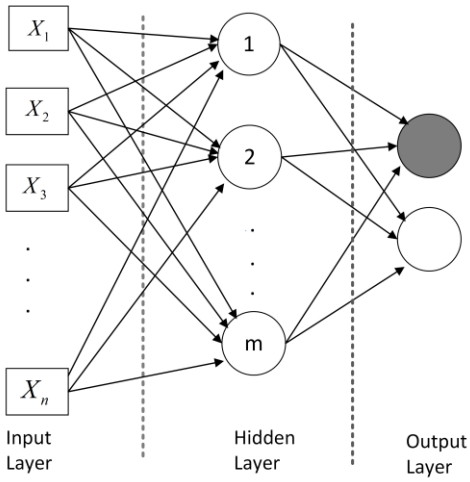

Figure 3. A typical MLP model (Bateni et al., 2007)

3.3.5. Logistic/Linear Regression Analysis

In order to find a cause-effect relationship between variables, Regression analysis is used as an analysis method. Regression analysis is the determination of the functional form of the relationship between the independent and dependent variables in order to perform prediction (Çokluk, 2010). The intended use of Logistic Regression Analysis is the same as other model construction techniques used in statistics. It is to build a biologically reasonable model that can state the relationship between independent and dependent variables in the best fit by using the least variable (Yavuz & Çilengiroğlu, 2020). Logistic regression is commonly used in medical applications to facilitate the interpretation of model parameters like variations in log size (Kurt et al., 2008). It is appropriate for variable selection methods commonly found in commercial applications and to interpret results as probabilities. Linear regression is an approximation in order to model the correlation between a numeric dependent factor and one or more independent factors (Karadağ, 2020). Linear regression analysis is to build a model that estimates the desired variable, based on the variable(s) that can be detected earlier or easier than the desired variable (İyi & Erol, 2008). The main feature that distinguishes logistic regression analysis from linear regression analysis is that the dependent variable has two or more classes

3.3.6. Linear Discriminant Analysis

This algorithm developed by R. A. Fischer in 1936 and is preferred especially for dimension reduction (Mika et al., 1999). In this way; it reduces underfitting and overfitting problems by making the calculation easier. In order to find the combination of properties that are best separated between classes, linear separation analysis maximizes their ratio to class distribution between classes, rather than maximizing total clutter. The simple idea is that: the same classes should be tightly wrapped together, distinct classes ought to be as far apart as possible in the lower dimension representations. It can be used for data preprocessing in cases where feature extraction is difficult before classification. LDA examines the distribution of classes to classify the data ideally and finds the difference between their mean values. It creates feature subspaces through them. LDA reduces the number of dimensions to prevent learning and memorization. Therefore, it can provide a high advantage in feature extraction as well as the application as a classification algorithm (Huang et al., 2009; Çebi et al., 2019).

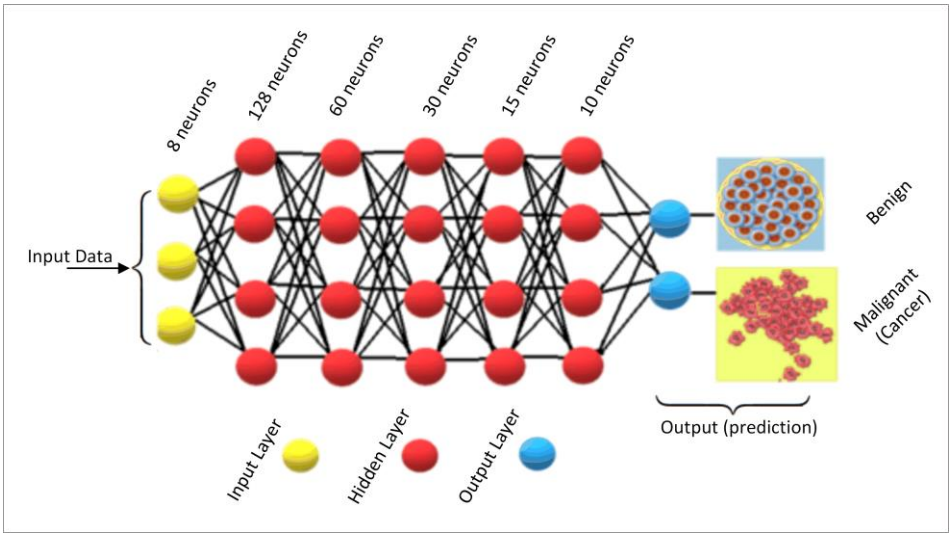

Figure 4. DNN architecture of the implemented model in this study

3.3.7. Multi Layer Perceptron

In recent years, interest in neural networks application for problems that cannot be solved with classical techniques has increased and it has been used successfully in many medical applications. MLP is widely used in classification and regression applications in many fields such as pattern, speech recognition, and classification problems (Ramchoun et al., 2016). A multi layer perceptron (MLP) is commonly used as supervised learning method or classifier. And, it is a feed-forward artificial neural network. It works better for the data which are not linearly separable (Özhan, 2020). In particular, it has a great influence on the convergence of these networks, in terms of the choice of the structure for the type of activation function used for each neuron. MLP generally has superior performance in classification, prediction, recognition, and interpretation (Mohammadi et al., 2017). This model which inputs and possible outputs are shown together during the training phase, is the most frequently used model in neural network. In the MLP method, there is an input layer, an output layer, and one or more hidden layers between the input and output as shown in Figure 3. The processing unit in the layers is interconnected. In MLP, the information that is resolved by the input layer is taken into the system, the information that is processed by the output layer is exported. A backpropagation algorithm is used for training. In this algorithm, the errors are tried to be reduced from the backward output to the input. Thus, it is aimed to reduce the error between the predicted result from the network and actual result. As shown in Figure 3, a typical MLP model is built with n inputs, a hidden layer with m nodes, and two output nodes (Bateni et al., 2007; Bozkurt et al., 2015; Timuş & Kıyak, 2015). MLP was utilized by using the MLPClassifier method of Python's sklearn.neural_network.MLPClassifier library. This model contains 100 hidden layers as default. For this model, we used the default parameters of MLPClassifier as 100 hidden-layer neural networks with ReLU activation function. In addition, Enaˇchescu (2004) proposed a regression model based on MLP network. This model carried out MLP with no activation function in output layer for backpropagation to train. Hence, a loss function that bases on square error is used and output consists of a set of continuous values. In comparison with other regression models, it was claimed that the MLP achieved a prediction error of about 10 times lower than other models (Enaˇchescu, 2004). Thus, we also used MLP-Regression model to obtain high accuracy and low prediction error in this study.

3.3.8. Deep Neural Network

Deep Learning algorithms can be considered as the more complex form of artificial neural networks (ANN). A deep feed-forward neural network is an artificial neural network containing many hidden layers between input and output layers. Deep learning is an artificial neural network architecture with multi-layer perceptrons (Liu et al., 2017). ANN algorithms have been developed based on human learning processes. In ANN systems, structures defined as neurons are modeled to be interconnected just as the neurons in the biological nervous system interact with each other. Deep learning is generally based on learning from the representation of data. In this way; it was thought that it would have the capacity to learn, memorize and reveal the relationship between data. Neural networks can easily learn the features and the relationships between them that other algorithms cannot easily discover. Artificial neural networks learn a complex model using layers of neurons that mathematically transform data (Pervan & Keleş, 2019).

In machine learning, deep learning algorithms have aspects that differ from existing algorithms. The algorithms need a huge amount of data and hardware with very high computing power that can handle this data. Deep learning algorithms have become very popular due to the great advances in graphics card-based computing power. It is also used by large technology companies (Google, Facebook, Microsoft, Nvidia, etc.) with high data and computing power and integrated into their products. At the same time, these technology companies are supporting rapid progress in this area by opening their own deep learning software libraries to developer communities (Liu et al., 2017; Klang, 2018).

In this paper, a deep neural network (DNN) was created with the open-source library Keras (Chollet, 2015) and classification was performed on the dataset. In this proposed method, seven-layer neural network is constructed with the Keras library. As shown in Figure 4, DNN consists of a total of 7 layers which are one input layer (yellow neurons on this layer), five hidden layers (red neurons on these layers), and one output layer (blue neurons on this layer). As shown in Figure 4, there

Table 2. Performance comparison of the observed machine learning algorithms on prostate cancer dataset

Name of the Algorithm	Precision (%)	Recall (%)	AUC (%)	Accuracy (%)	F1-Score (%)
K-NN	88	83	83.30	83	86
SVM	94	89	90.3	90	91
Logistic Regression	73	92	84.7	83	81
NB	83	94	86.2	87	88
RF	95	91	88.5	90	93
Linear Regression	89	84	87.6	83	86
LDA	100	81	90.5	87	89
MLP	**95**	**100**	**95.8**	**97**	**97**
MLP-Regressor	90	90	90.3	90	90
DNN	89	94	88.9	90	92

Table 3. Error measurements of the observed machine learning algorithms on prostate cancer dataset

Model	K-NN	SVM	Log.Reg.	NB	RF	Lin.Reg.	LDA	MLP	MLP.Reg.	DNN
MAE	25.57	17.03	34.03	8.60	17.03	0.20	34.0	**0.03**	17.03	0.10
MSE	0.17	0.10	0.17	0.13	0.10	0.27	0.13	**0.03**	0.10	0.10
RMSE	0.44	0.32	0.41	0.37	0.32	0.52	0.37	**0.18**	0.32	0.32

are 8 neurons (nodes) in the input layer, 128 neurons in the first hidden layer, 60 neurons in the second layer, 30 neurons in the third layer, 15 neurons in the fourth layer, and 10 neurons in the fifth hidden layer. ReLU is used as the activation process for hidden layers. Sigmoid function is used for classification in the last layer. Binary crossentropy is preferred as the loss function because of the binary classification.

In artificial neural networks, the accurate updating of weights is crucial to the learning process. After calculating the error, the weights are updated according to the obtained error. The back-propagation algorithm is used to update the weights. RMS-Prop, Adam, and Adadelta are Gradient-based stochastic gradient descent methods. They frequently used to calculate the effect of each weight in the neural network on the calculated error by backpropagation algorithm (Yazan & Talu, 2017). In this study, we preferred Adam algorithm for optimization. Adam is a more efficient, adaptive optimization algorithm that can be used instead of the classical stochastic gradient descent method. It updates the learning rate dynamically for each parameter.

The reliability of the performance level of the classifier was provided by the 10-fold cross validation method. In this study, we use *scikit-learn StratifiedKFold* to perform 10-fold cross-validation. This resampling technique provides a robust estimate of the machine learning model's performance on raw data. In order to evaluate our model and print the results, we use the *cross_val_score()* function of *scikit-learn* by using the cross-validation schema.

4. Experimental Results and Discussion

In this section, we compare and discuss the performance of nine supervised machine learning algorithms as classifiers that are K-Nearest Neighbor (K-NN), Support Vector Machines (SVM), Random Forest (RF), Naive Bayes (NB), Linear Regression, Logistic Regression, Linear Discrimination Analysis (LDA), Multi-Layer Perceptron (MLP) and Deep Neural Network (DNN). The training and test sets were randomly selected as 70% training and 30% test data based on the original data to investigate classifier's accuracy and performance. Different performance results can be obtained in different methods according to the distribution of training and test data. In other words, the accuracy of the methods may vary according to the training and test data. At this stage in order to obtain more stable results, we ran each algorithm more than ten times and recorded their five best performances. As a result of the classification, "B" benign cell and "M" malignant cell (cancer) were detected.

The design flow diagram of this study, which was carried out for the prediction of prostate cancer through various supervised machine learning methods, is given in Figure 1. Each step in this diagram is detailed in the following subsections

Performance evaluation of various machine learning algorithms are discussed by considering misclassified instances, correctly classified instances, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE). We performed the calculation of Precision, Recall, AUC, Accuracy, and F-Measure of various machine learning algorithms through confusion matrix for dataset with two (binary) classes. For some situations, additional measures are required to evaluate classifiers. Since only the accuracy in some situations could be misleading. AUC is considered a better measure than accuracy. Also, the F-measure is used as a combination of recall and precision in a single metric. Especially in uneven class distribution, it sometimes could be more useful than accuracy.

We observed the comparison of performances of the nine classifiers as shown in Table 2. We can see clearly see that MLP produces the best performance for this dataset according to scores of AUC (95.8%), Accuracy (97%) and F1-Score (97). K-NN, Logistic Regression and Linear Regression (Accuracy=83%) outcome lower results than others. We applied the regression model based on MLP on this dataset. MLP-Regressor model gives the 90.3% AUC and 90% Accuracy for prostate cancer dataset. In comparison with other regression models, MLP- Regressor model gives high accuracy and low prediction error than other models.

Table 4. Confusion matrix of the observed machine learning algorithms on prostate cancer dataset

Algorithm	Actual Classes	Predicted	
		Cancer= yes	Cancer= no
K-NN	Cancer= yes	83 (TP)	17 (FN)
	Cancer= no	17 (FP)	83 (TN)
SVM	Cancer= yes	92 (TP)	8 (FN)
	Cancer= no	11 (FP)	89 (TN)
Log.Reg.	Cancer= yes	92 (TP)	8 (FN)
	Cancer= no	22 (FP)	78 (TN)
NB	Cancer= yes	79 (TP)	21 (FN)
	Cancer= no	6 (FP)	94 (TN)
RF	Cancer= yes	86 (TP)	14 (FN)
	Cancer= no	9 (FP)	91 (TN)
Lin.Reg.	Cancer= yes	86 (TP)	14 (FN)
	Cancer= no	13 (FP)	87 (TN)
LDA	Cancer= yes	100 (TP)	0 (FN)
	Cancer= no	19 (FP)	81 (TN)
MLP	Cancer= yes	92 (TP)	8 (FN)
	Cancer= no	0 (FP)	100 (TN)
MLP.Reg.	Cancer= yes	92 (TP)	8 (FN)
	Cancer= no	12 (FP)	88 (TN)
DNN	Cancer= yes	83 (TP)	17 (FN)
	Cancer= no	6 (FP)	94 (TN)

Classifier performance was evaluated according to values of MAE, MSE and RMSE. The MAE, MSE and RMSE values were calculated in the test samples and the average of the five error results obtained was taken. The lowest average value was obtained with MLP (MAE=0.03, MSE=0.03 and RMSE=0.18) as shown in Table 3.

Confusion matrix is one of the simplest and most heuristic metrics used to obtain the accuracy of the model. Table 4 shows the test complexity matrix of the observed machine learning algorithms. TN, TP, FN and FP values are taken into consideration while calculating the performance criteria (Precision, Recall, AUC, Accuracy and F-Measure). For instance to evaluate Accuracy value from the formula ACC=((TP+TN)/(TP+TN+FP+FN))*100, ACC= ((92+100)/(92+100+0+100))*100=96, %96 is obtained for MLP method.

The another performance measure which we observed is that the Receiver Operating Characteristics (ROC) curve. This curve gives us the evaluation metric of the observed machine learning algorithms. A ROC curve describes the visual relationship between true and false positives. ROC curve is composed by plotting the number of false positive on the horizontal axis (1- specificity) and true positive values on vertical axis (sensitivity). An optimum point on the ROC curve should be (0, 1). This means that none of negative instances are misclassified as positive and all positive instances are classified as positive. A ROC curve covers all information in the confusion matrix. The closer the curve follows to the upper-left edge of the ROC space, the greater the accuracy of the test. In order to compare classifiers, the area under the ROC curve, in other words AUC (Area Under of Curve) value is calculated to reduce the performance of the ROC to a single value. Since AUC is part of the field in the unit square, its value is always between 0 and 1. Figure 5(a) shows the ROC cure of the DNN method. Figure 5(b) shows the ROC cure of the other methods. As shown in Figure 5(b), MLP classifier has the greater accuracy and AUC value (0.96) than others.

Table 5. The results of studies and MLP method on same dataset

Article	Year	Method	Accuracy(%)
Laabidi & Aissaoui	2020	RNN	81.30
Srivenkatesh	2020	Random Forest Lojistic Regression	90
Proposed	**2020**	**MLP**	**97**

In order to demonstrate the performance of the proposed model, both accuracy and loss graphs are shown in Figure 6. The x-axis is the number of model training periods which refers to the number of training cycles across the complete dataset, and the Y-axes are the loss and accuracy respectively. When the accuracy graph is closely examined, the initial test (validation) accuracy is higher than the training accuracy for some periods. Both test and training accuracy curves follow an upward trend as the number of periods increase. The loss curve shows the decrease in error rate as shown in Figure 6. It shows that the training process and the learning of the network are in a good learning rate. As shown in Figure 6, while the loss value decreases at each epoch (i.e. 100 epoch), accuracy increases and learning occurs through the given training set.

To the best of our knowledge, the performance comparison among various supervised machine learning techniques is limited on same prostate cancer dataset. Table 5 shows the studies which are use the same dataset. The data set is already public available online since 2018. Laabidi and Aissaoui (2020) studied on popular machine learning algorithms to predict prostate cancer. The higher results were obtained with the Recurrent Neural Network (RNN) method with an accuracy of 81.3% when compared to other methods. Srivenkatesh (2020) also studied on predicting prostate cancer with different machine-learning algorithms. Srivenkatesh obtained the better

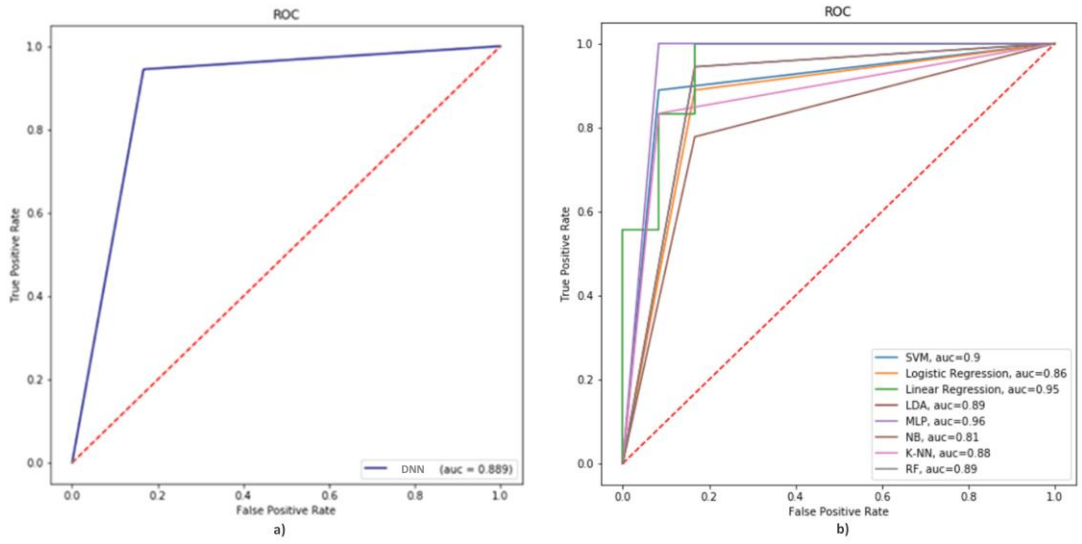

Figure 5. ROC curves of the observed machine learning algorithms: (a) ROC curve of a DNN model, (b) ROC curve of others

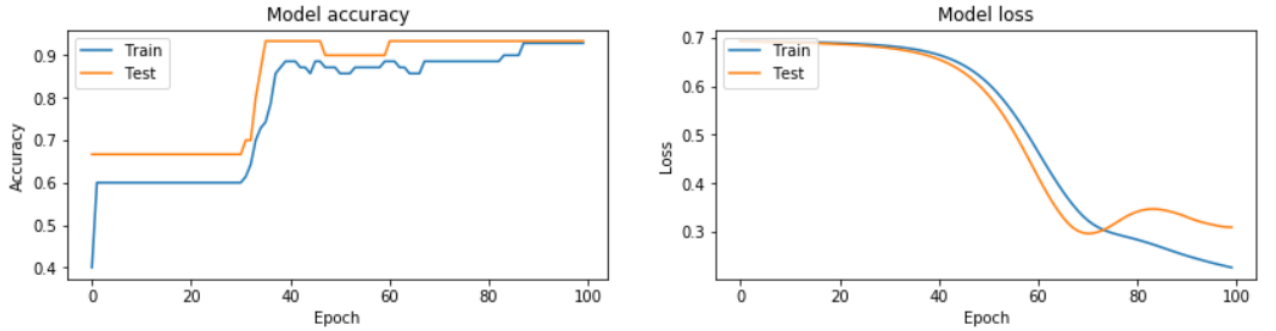

Figure 6. An example of loss and accuracy graph of the DNN model

accuracy (90%) with Random Forest and Logistic Regression when compared to other algorithms. In this study, we were able to obtain the highest accuracy rate of 97% with the MLP classifier when compared to other methods. Experimental results showed that the MLP classifier give better results according to the other algorithms used in this study. In this context, MLP classifier outperformed the best study ever reported in the literature in terms of accuracy, AUC and F1-Score performance criteria.

Based on the experimental results, MLPClassifier method of Python's sklearn.neural_network.MLPClassifier library give the best results (%97 accuracy) in predicting because of having 100 hidden-layer neural networks. Our modelled DNN method has fewer hidden layers. We think that if the DNN is trained with more hidden layers (also more neurons) and a very large amount of data, it will give higher result than %90 accuracy.

5. Conclusion and Future Work

In this study, prostate cancer which is one of the most cancer-related deaths of men and whose symptoms have similar characteristics with benign growth was studied. One of the biggest problems in medicine is to diagnose diseases. The lack of specific guidelines for evaluating prostate cancer symptoms and the low predictive rate of currently available diagnostic methods make this study valuable. We think that machine learning methods can be effective in predicting of prostate cancer problems where there are no specific and precise rules and the factors affecting the event can be predicted. Within the scope of this study, various supervised machine learning techniques were evaluated to predict prostate cancer. For this purpose, nine machine learning algorithms were implemented. The effectiveness of these algorithms are tested with a prostate cancer dataset of 100 patients from the public Kaggle platform. This dataset which is public available online since 2018, also it has been used by some studies in the literature. Experimental results showed that the MLP classifier outperformed the other algorithms used in this study and the best study ever reported in the literature in terms of accuracy, AUC and F1-Score performance criteria.

In this context, MLP classifier was found to be a good alternative for prostate cancer classification according to our experimental results. Thus, we can say that if the computer is trained with machine learning methods based on patient information, it can be clinically useful with high accuracy in predicting cancer. In this way, an unnecessary biopsy of the patient can be prevented. In future studies, it is planned to consider other datasets used in the related studies in the literature to predict prostate cancer. And, it is planned to apply and compare deep learning techniques for different data sets and large amounts of data.

References

Abraham, B., & Nair, M. S. (2018). Computer-aided classification of prostate cancer grade groups from MRI images using texture features and stacked sparse autoencoder. Computerized Medical Imaging and Graphics, 69, 60-68.

Al-Aidaroos, K. M., Bakar, A. A., & Othman, Z. (2012). Medical data classification with Naive Bayes approach. Information Technology Journal, 11(9), 1166.

Alickovic, E., & Subasi, A. (2016). Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier. Journal of medical systems, 40(4), 108.

Alqaraleh, S. (2020). Turkish Sentiment Analysis System via Ensemble Learning. European Journal of Science and Technology, (Special Issue), 122-129

Ankerst, D. P., Hoefler, J., Bock, S., Goodman, P. J., Vickers, A., Hernandez, J., ... & Thompson, I. M. (2014). Prostate Cancer Prevention Trial risk calculator 2.0 for the prediction of low-vs high-grade prostate cancer. Urology, 83(6), 1362-1368.

Arvidsson, I., Overgaard, N. C., Marginean, F. E., Krzyzanowska, A., Bjartell, A., Åström, K., & Heyden, A. (2018). Generalization of prostate cancer classification for multiple sites using deep learning. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 191-194). IEEE.

Aydın, C. (2018). Makine Öğrenmesi Algoritmaları Kullanılarak İtfaiye İstasyonu İhtiyacının Sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi, (14), 169-175.

Bateni, S. M., Borghei, S. M., & Jeng, D. S. (2007). Neural network and neuro-fuzzy assessments for scour depth around bridge piers. Engineering Applications of Artificial Intelligence, 20(3), 401-414.

Bozkurt, F., Altay, Ş.Y., Yaganoğlu, M., (2015). Yapay Sinir Ağları İle Ankara İlinde Hava Kalitesi Sağlık İndeksi Tahmini, 2.Ulusal Yönetim Bilişim Sistemleri Kongresi, Erzurum.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Chang, A. J., Autio, K. A., Roach III, M., & Scher, H. I. (2014). High-risk prostate cancer—classification and therapy. Nature reviews Clinical oncology, 11(6), 308.

Chollet, F. (2015). Keras: Deep learning library for theano and tensorflow. URL: https://keras. io/k.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27.

Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. Cambridge university press.

Cuzick, J., Thorat, M. A., Andriole, G., Brawley, O. W., Brown, P. H., Culig, Z., ... & Ilic, D. (2014). Prevention and early detection of prostate cancer. The lancet oncology, 15(11), e484-e492.

Çebi, C. B., Bulut, F. S., Fırat, H., Karataş, G & Şahingoz, Ö. K. (2019) Saldırı Tespit Sistemlerinde Makine Öğrenmesi Modellerinin Karşılaştırılması. Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 12(3), 1513-1525.

Çokluk, Ö. (2010). Lojistik regresyon analizi: Kavram ve uygulama. Kuram ve Uygulamada Eğitim Bilimleri, 10(3), 1357-1407.

De Nunzio, C., Pastore, A. L., Lombardo, R., Simone, G., Leonardo, C., Mastroianni, R., ... & Fuschi, A. (2018). The new Epstein gleason score classification significantly reduces upgrading in prostate cancer patients. European Journal of Surgical Oncology, 44(6), 835-839.

Ena˘chescu, D. (2004). Multilayer perceptron model for prostate cancer prediction. International Journal of Computer Mathematics, 81(4), 407-415.

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. Radiographics, 37(2), 505-515.

Goldenberg, S. L., Nir, G., & Salcudean, S. E. (2019). A new era: artificial intelligence and machine learning in prostate cancer. Nature Reviews Urology, 16(7), 391-403.

Grönberg, H. (2003). Prostate cancer epidemiology. The Lancet, 361(9360), 859-864.

Gültepe, Y. (2019). Makine Öğrenmesi Algoritmaları ile Hava Kirliliği Tahmini Üzerine Karşılaştırmalı Bir Değerlendirme. Avrupa Bilim ve Teknoloji Dergisi, (16), 8-15.

Huang, D., Quan, Y., He, M., & Zhou, B. (2009). Comparison of linear discriminant analysis methods for the classification of cancer based on gene expression data. Journal of experimental & clinical cancer research, 28(1), 149.

İyi, P., & Erol, H. (2008). Çoklu Lineer Regresyonda En İyi Model Seçimi. ÇÜ Fen Bilimleri Enstitüsü, 17(5), 48-56.

Karadağ, K. (2020). Makine Öğrenme Yöntemleri ile Semen Kalitesi Tahmini. Avrupa Bilim ve Teknoloji Dergisi, (18), 306-311.

Karakoyun, M., & Hacıbeyoğlu, M. (2014). Biyomedikal Veri Kümeleri İle Makine Öğrenmesi Siniflandirma Algoritmalarinin İstatistiksel Olarak Karşilaştirilmasi. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 16(48), 30-42.

Klang, E. (2018). Deep learning and medical imaging. Journal of thoracic disease, 10(3), 1325.

Kurt, I., Ture, M., & Kurum, A. T. (2008). Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. Expert systems with applications, 34(1), 366-374.

Laabidi, A., & Aissaoui, M. (2020). Performance analysis of Machine learning classifiers for predicting diabetes and prostate cancer. In 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET) (pp. 1-6). IEEE.

Li, J., Weng, Z., Xu, H., Zhang, Z., Miao, H., Chen, W., ... & Ye, Q. (2018). Support Vector Machines (SVM) classification of prostate cancer Gleason score in central gland using multiparametric magnetic resonance images: A cross-validated study. European journal of radiology, 98, 61-67.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. Neurocomputing, 234, 11-26.

Mika, S., Ratsch, G., Weston, J., Scholkopf, B., & Mullers, K. R. (1999). Fisher discriminant analysis with kernels. In Neural networks for signal processing IX: Proceedings of the 1999 IEEE signal processing society workshop, 41-48. IEEE.

Mohammadi, M. R., Sadrossadat, S. A., Mortazavi, M. G., & Nouri, B. (2017). A brief review over neural network modeling techniques. In 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI) (pp. 54-57). IEEE.

Nagpal, K., Foote, D., Liu, Y., Chen, P. H. C., Wulczyn, E., Tan, F., ... & Corrado, G. S. (2019). Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. NPJ digital medicine, 2(1), 1-10.

Özhan, E. (2020). Yapay Sinir Ağları ve Üstel Düzleştirme Yöntemi ile Türkiye'deki CO2 Emisyonunun Zaman Serisi ile Tahmini. Avrupa Bilim ve Teknoloji Dergisi, (19), 282-289.

Pervan, N., & Keleş, Y. (2019). Derin öğrenme yaklaşımları kullanarak Türkçe metinlerden anlamsal çıkarım yapma. Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Ankara.

Ramchoun, H., Idrissi, M. A. J., Ghanou, Y., & Ettaouil, M. (2016). Multilayer Perceptron: Architecture Optimization and Training. IJIMAI, 4(1), 26-30.

Ray, S. A. (2020). Survey on Application of Machine Learning Algorithms in Cancer Prediction and Prognosis. In Data Management, Analytics and Innovation (pp. 349-361). Springer, Singapore.

Reda, I., Ghazal, M., Shalaby, A., Elmogy, M., AbouEl-Fetouh, A., Ayinde, B. O., ... & El-Baz, A. (2018). A novel adcs-based cnn classification system for precise diagnosis of prostate cancer. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 3923-3928). IEEE.

Sajid S. (2018). Prostate cancer dataset, [Online]. Available: https://www.kaggle.com/sajidsaifi/prostate-cancer

Saritas, M. M., & Yasar, A. (2019). Performance analysis of ANN and Naive Bayes classification algorithm for data classification. International Journal of Intelligent Systems and Applications in Engineering, 7(2), 88-91.

Shakeel, P. M., & Manogaran, G. (2020). Prostate cancer classification from prostate biomedical data using ant rough set algorithm with radial trained extreme learning neural network. Health and Technology, 10(1), 157-165.

Song, Y., Zhang, Y. D., Yan, X., Liu, H., Zhou, M., Hu, B., & Yang, G. (2018). Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. Journal of Magnetic Resonance Imaging, 48(6), 1570-1577.

Srivenkatesh, M. (2020) Prediction of Prostate Cancer using Machine Learning Algorithms International Journal of Recent Technology and Engineering. 8(5).

Stephan, C., Jung, K., Cammann, H., Vogel, B., Brux, B., Kristiansen, G., ... & Sinha, P. (2002). An artificial neural network considerably improves the diagnostic power of percent free prostate-specific antigen in prostate cancer diagnosis: Results of a 5-year investigation. International journal of cancer, 99(3), 466-473.

Taşcı, E., & Onan, A. (2016). K-en yakın komşu algoritması parametrelerinin sınıflandırma performansı üzerine etkisinin incelenmesi. Akademik Bilişim.

Taşkıran, U. Ş. (2008). Prostat kanser riski hesaplamalarında yapay sinir ağları kullanımı (Doctoral dissertation, Selçuk Üniversitesi Fen Bilimleri Enstitüsü).

Timuş, O., & Kıyak, E. (2015). Optimizing MLP Classifier and ECG Features for Sleep Apnea Detection. Journal of Naval Sciences and Engineering, 11(1), 1-18.

Yağanoğlu, M., Bozkurt, F., & Günay, F. B. (2014). EEG tabanli beyin-bilgisayar arayüzü sistemlerinde öznitelik çikarma yöntemleri. Mühendislik Bilimleri ve Tasarım Dergisi, 2(3), 313-318.

Yavuz, A., & Çilengiroğlu, Ö. V. (2020) Lojistik Regresyon ve CART Yöntemlerinin Tahmin Edici Performanslarının Yaşam Memnuniyeti Verileri için Karşılaştırılması. Avrupa Bilim ve Teknoloji Dergisi, (18), 719-727.

Yazan, E., & Talu, M. F. (2017). Comparison of the stochastic gradient descent based optimization techniques. In 2017 International Artificial Intelligence and Data Processing Symposium (IDAP) (pp. 1-5). IEEE.

Yuan, Y., Qin, W., Buyyounouski, M., Ibragimov, B., Hancock, S., Han, B., & Xing, L. (2019). Prostate cancer classification with multiparametric MRI transfer learning model. Medical physics, 46(2), 756-765.

Zaki, M. J., & Meira Jr, W. (2019). Data Mining and Machine Learning: Fundamental Concepts and Algorithms. Cambridge University Press.

Zhang, C., Li, H. R., Fan, J. B., Wang-Rodriguez, J., Downs, T., Fu, X. D., & Zhang, M. Q. (2006). Profiling alternatively spliced mRNA isoforms for prostate cancer classification. Bmc Bioinformatics, 7(1), 202.