Black Sea Journal of Engineering and Science
doi: 10.34248/bsengineering.1574470

Open Access Journal
e-ISSN: 2619 – 8991

Research Article
Volume 8 - Issue 1: 243-248 / January 2025

PREDICTING AND REDUCING PATIENT WAITING TIMES IN DENTAL CLINICS USING MACHINE LEARNING: A CASE STUDY FROM TÜRKİYE

Abdulkadir KESKİN[1]*

[1]*Istanbul Medeniyet University, Faculty of Engineering and Natural Sciences, Department of Statistics, 34674, İstanbul, Türkiye*

Abstract: Long waiting times in polyclinics are a critical factor affecting patient satisfaction and the efficient use of healthcare personnel and resources. This study applied machine learning (ML) algorithms to predict and reduce patient waiting times in a dental clinic in Türkiye. The daily data collected from the clinic included variables such as patient satisfaction, appointment patients, Walk-in patients, number of doctors and nurses, and dental technicians on duty. Six ML algorithms were tested: Decision Trees (DT), Linear Regression (LR), Support Vector Machines (SVM), Gaussian Process Regression (GPR), Kernel Regression (KR), and Neural Networks (NN). Among these, the GPR model achieved the best performance, accurately predicting patient waiting times with an R^2 value of 0.936 and RMSE of 0.075. This study highlights the potential of ML methods to enhance operational efficiency in healthcare management.

Keywords: Healthcare management, Waiting time prediction, Dental clinic, Machine learning

***Corresponding author:** *Istanbul Medeniyet University, Faculty of Engineering and Natural Sciences, Department of Statistics, 34674, İstanbul, Türkiye*
E mail: abdulkadir.keskin@medeniyet.edu.tr, (A. Keskin)
Abdulkadir KESKİN https://orcid.org/0000-0002-4795-1028

Received: October 27, 2024
Accepted: December 03, 2024
Published: January 15, 2025

Cite as: Keskin A. 2025. Predicting and reducing patient waiting times in dental clinics using machine learning: a case study from Türkiye. BSJ Eng Sci, 8(1): 243-248.

1. Introduction

Efficiency and service quality in healthcare have become increasingly important in recent years due to growing demand and a competitive environment. Patient satisfaction in a competitive market is regarded as one of the most critical indicators of a healthcare institution's success. Particularly, patient waiting times are a significant factor that directly affects patient satisfaction (Soremekun et al., 2011; Bahammam, 2023). Long waiting times hinder patients' access to services, disrupt operational processes within hospitals, and increase the workload of healthcare staff. This situation leads to a decline in the quality of healthcare services and erodes patients' trust in the healthcare institution. Delivering healthcare services promptly and managing patient flow efficiently helps patients feel better both physically and psychologically (Boudreaux and O'Hea, 2004; Mohsin et al., 2007).

A significant portion of the treatment time in hospitals impacts waiting times, directly influencing both patient satisfaction and the quality of care. Various studies on patient satisfaction have shown that waiting time has a direct effect on patients' decisions to return to the hospital (Anderson et al., 2007; Pitrou et al., 2009). Patients with long waiting times are expected to deliver lower values of overall service quality and thus, in turn, reduce the probability of returning to that hospital. Analysis of the relevant data obtained to reduce waiting times, use available resources effectively, and improve the process are important for health management. Specifically, data-driven methods like machine learning ML are highly effective tools for extracting meaningful insights from large datasets and identifying inefficiencies in service delivery (Stiglic et al., 2020; Keskin et al., 2024). ML algorithms offer various opportunities to improve service processes and enhance operational efficiency by making predictions based on historical data. Large and complex datasets that are difficult to analyze using traditional methods can be processed quickly and effectively using ML algorithms, significantly contributing to healthcare management. The applications of ML methods range from disease diagnosis and treatment to improving operational processes (Kononenko, 2001; Liao et al., 2016).

Several studies in literature have developed strategies to estimate patient waiting times. Cayirli and Veral (2003) explored the optimization of appointment systems and examined the effects of factors such as delays, service times, and patient and doctor preferences on waiting times. Similarly, Qu and Shi (2011) proposed models to manage both pre-scheduled appointments and real-time demands within the same system. These models provide valuable insights into optimizing patient demand while predicting waiting times. Another study developed a Kalman filter-based model to predict outpatient waiting times, finding that this model outperformed traditional models (Montecinos et al., 2018). ML techniques used in dental clinics have been found to perform better than traditional methods. For instance, while Autoregressive integrated moving average (ARIMA) models predict waiting times in a specific clinic based on historical data, ML algorithms produce more dynamic and accurate results by considering current patient data and external factors (Channouf et al., 2007). In this context, studies on predicting waiting times contribute to more efficient use of resources while increasing patient satisfaction (Reid et al., 2013). Atalan and Keskin (2023) estimated patient waiting times in a dental clinic in Türkiye using a discrete-event simulation (DES) model. Komşuoğlu (2022) highlighted in a study on patient satisfaction in dental clinics that treatment duration is a significant variable influencing patient satisfaction.

The use of artificial intelligence, especially machine learning ML algorithms or artificial neural networks, is increasingly used in various fields, including healthcare management. ML-based models, such as artificial neural networks, generally provide much more accurate predictions than traditional statistical methods. For this reason, using such methods in healthcare management can improve operational efficiency, reduce costs, reduce staff workload, and increase patient satisfaction. In this study, which aims to improve patient waiting times, one of these problems, ML methods were used. Waiting time is considered one of the most critical indicators of healthcare quality and reducing it is important to increase patient satisfaction. The aim of the study was to estimate maximum waiting times and identify areas for improvement to increase operational efficiency and reduce waiting times. Six different ML algorithms were used to achieve this goal.

This study consists of four main sections. The first section emphasizes the importance of the topic and provides a comprehensive review of the relevant literature. The second section details the dataset and methodology employed in the study. In the third section, machine learning algorithms are compared to identify the ones with the best predictive performance, and the results are analyzed. Finally, the last section discusses the study's recommendations, limitations, and suggestions for future research.

2. Materials and Methods

In this study, clinical data collected over a 6-month period from an intermediate-sized dental clinic in Türkiye were utilized. Typically, for machine learning algorithms, 80% of the dataset is randomly selected for training, while the remaining 20% is used for testing. To ensure uniformity and enhance analytical accuracy, the dataset was normalized to a range between 0 and 1 prior to analysis. Table 1 presents the basic statistics related to daily appointment patients, walk-in patients, the number of doctors, nurses, and dental technicians, as well as the average waiting time (in minutes) in the dental clinic. In addition, the general satisfaction score of patients, which is critical for healthcare services, was also included in the model. The satisfaction score for the service received by the patients was obtained using a scale from one to ten. The dataset was obtained from daily records between March 1, 2024, and August 31, 2024. Descriptive statistics for the variables used are summarized in Table 1.

Before running machine learning algorithms, a feature selection analysis was conducted to assess the importance of the variables used in the models. Feature selection helps improve the accuracy of the ML model by eliminating unnecessary variables (Miao and Niu, 2016). Figure 1 shows the feature importance results of the variables based on the F-test algorithms. In the ML models predicting the average waiting time, it was found that the patient satisfaction score is the most important parameter. The second most significant parameter turned out to be the number of scheduled patients. In contrast, the number of technicians was identified as the least important parameter compared to the others. The patient satisfaction score was collected using a Likert scale ranging from 1 to 10, obtained after the patients received treatment. Waiting time was recorded in minutes.

Waiting times at the dental clinic were predicted using six different algorithms. The models used include Decision Trees (DT), Linear Regression (LR), Support Vector Machines (SVM), Gaussian Process Regression (GPR), Kernel Regression (KR), and Neural Networks (NN). In this study, various machine learning algorithms that perform well with different data structures were employed to analyze the data from the dental clinic. Decision Trees create a hierarchical model by dividing the dataset into branches, showing the effects of each variable under certain conditions (Song and Lu, 2015). Linear Regression assumes a linear relationship between the dependent and independent variables. Although it models this relationship using the least squares method and is simple and interpretable, it may be insufficient for non-linear relationships (Bertsimas and King, 2016). Support Vector Machines excel in solving non-linear problems by finding the best hyperplane that separates the data points (Gualtieri and Chettri, 2000). Gaussian Process Regression assigns distributions to each data point using a probabilistic approach, considering uncertainties. This method is particularly preferred in cases requiring high accuracy (Marrel and Iooss, 2024). Kernel Methods are effective in data analysis by projecting non-linear relationships into a higher-dimensional space (Arenas-Garcia et al., 2013). Thanks to their multilayered structures, neural networks learn from inputs, model complex relationships, and exhibit strong performance, especially with large data sets (Naskath et al., 2023). The characteristics of the six machine learning methods used in this research are detailed in Table 2. All analyses were performed with MATLAB 2024, a powerful software tool widely used in fields such as machine learning, simulation and artificial intelligence (The MathWorks, 2024).

Table 1. Descriptive statistics of the variables used to estimate patient waiting time

	Variable	Mean	Median	Maximum	Minimum	Std. Dev.	Obs.
Average waiting time	Output	12.93	13.90	29.73	4.13	8.11	158
Appointment Patients	Input	54.87	56.00	86.00	15.00	18.16	158
Walk-in patient	Input	13.46	12.00	28.00	4.00	5.67	158
Number of Dentists	Input	4.03	4.00	5.00	3.00	0.85	158
Number of Nurses	Input	2.96	3.00	4.00	2.00	0.80	158
Num Technicians	Input	1.49	1.00	2.00	1.00	0.50	158
Patient Satisfaction	Input	5.74	5.71	7.72	4.04	0.88	158

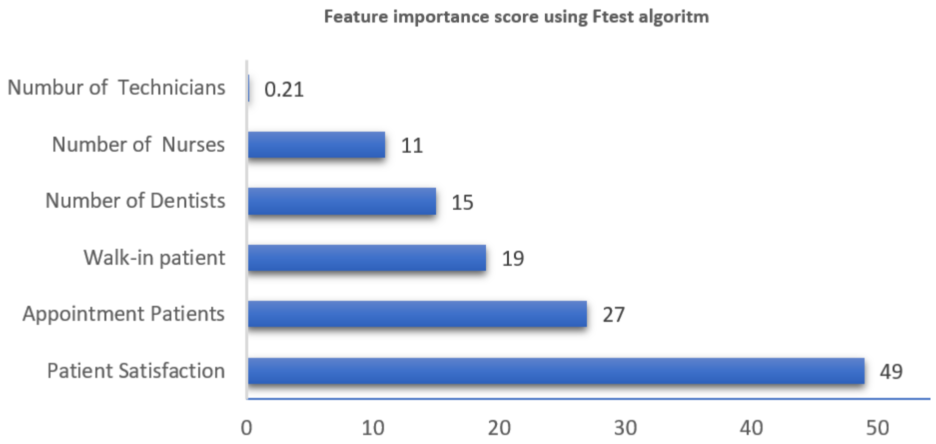

Figure 1. The feature selection information of ML model

Table 2. Parameters of machine learning algorithms used to predict waiting time

Algorithm	Parameters of the algorithm
Decision Trees (DT)	- MaxNumSplits: 100 - MinLeafSize: 1 - SplitCriterion: 'gdi' (Gini's diversity index)-Prune: 'off'
Linear Regression (LR)	- FitIntercept: true (bias term is included) - Solver: 'normal equations' (closed-form solution) - Lambda: 0 (no regularization, simple regression)
Support Vector Machines (SVM)	- KernelFunction: 'linear' - BoxConstraint: 1 (penalty parameter) - KernelScale: 'auto', Standardize: true
Gaussian Process Regression (GPR)	- KernelFunction: 'squaredexponential' - BasisFunction: 'constant' - Sigma: 1e-3 (noise variance)
Kernel Regression (KR)	- KernelFunction: 'Gaussian' - Bandwidth: 1 (kernel width)
Neural Network (NN)	- HiddenLayerSizes: 10 (single layer, 10 neurons) - ActivationFunction: 'relu' - Solver: 'adam' - MaxEpochs: 100, LearningRate: 0.01

In this study, the performance of the machine learning models was compared using commonly used evaluation metrics. To assess the accuracy and error rates of the models, the predictive performances of the algorithms were evaluated based on RMSE (Root Mean Square Error), MSE (Mean Square Error), R^2 (coefficient of determination), and MAE (Mean Absolute Error). The details are shown in the following formulas.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \tilde{y}_i|^2 \quad (1)$$

$$MSE = \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 \quad (2)$$

$$R^2 = \sum_{i=1}^{n} \left[\frac{y_i - \tilde{y}_i}{y_i - \bar{y}_i}\right]^2 \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2}{n}} \quad (4)$$

3. Results

The results of the machine learning models used to predict the average waiting time of patients in the dental clinic are presented in Table 3. The Decision Tree model obtained 0.087 RMSE, 0.916 R^2, 0.008 MSE and 0.065 MAE values in the training phase. In the testing phase, it showed a strong prediction performance with 0.086 RMSE and 0.917 R^2 values.

On the other hand, the Linear Regression model performed worse than the Decision Tree model. While the model reached 0.106 RMSE and 0.875 R^2 in the training phase, its performance decreased even further in the testing phase, with 0.848 R^2 and 0.116 RMSE, showing a lower accuracy than the training phase. Although the LR model performs well in linear relationships, its predictive power weakens in complex models. The Support Vector Machines model has an RMSE of 0.110 and an R^2 of 0.867 in the training phase. In the testing phase, it performed close to the LR model with 0.850 R^2 and 0.115 RMSE. Similar to the LR model, the SVM model performed worse than the DT model.

The Gaussian process regression model provided the best results during the training phase, with an RMSE of 0.075 and an R^2 of 0.938. In the testing phase, it exhibited performance nearly identical to the training phase, with an R^2 of 0.936 and an RMSE of 0.075. GPR represented the data very well, emerging as the best-performing ML model.

The kernel regression model showed similar performance to the decision tree model in the testing phase, with an RMSE of 0.087 and an R^2 of 0.916. In the testing phase, it performed almost identically to the DT model, with an R^2 of 0.915.

The neural network model had an RMSE of 0.112 and an R^2 of 0.861 during the training phase, indicating lower performance compared to other models. However, in the testing phase, it provided better results with an R^2 of 0.917 and an RMSE of 0.086.

Overall, the GPR model, alongside the DT model, has shown the best performance for predicting the average waiting time of patients in the clinic. In comparison to other models, the KR model also demonstrated strong performance. The training and testing performance of the GPR model is illustrated in Figure 2.

Table 3. Value of the measurement performance of ML models

Ml model algorithm	Train				Test			
Performance criteria	MAE	MSE	RMSE	R^2	MAE	MSE	RMSE	R^2
Decision Tree	0.065	0.008	0.087	0.916	0.063	0.007	0.086	0.916
Linear Regression	0.086	0.011	0.106	0.875	0.097	0.014	0.116	0.848
Support Vector Machines	0.085	0.012	0.110	0.867	0.094	0.013	0.115	0.850
Gaussian Process Regression	0.056	0.006	0.075	0.938	0.056	0.006	0.075	0.936
Kernel Regresion	0.067	0.008	0.087	0.916	0.060	0.008	0.087	0.915
Neural Network	0.081	0.013	0.112	0.861	0.065	0.007	0.086	0.917

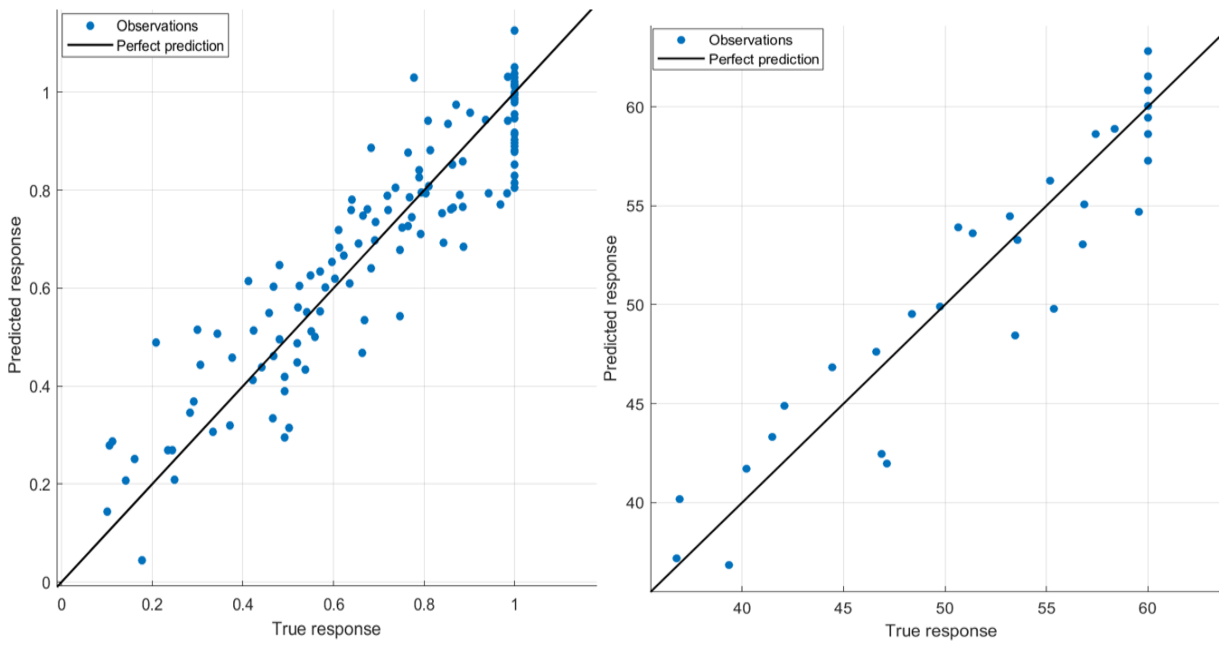

Figure 2. Performance of the train and test prediction phase of the GPR algorithm.

Health expenditures are increasingly becoming a significant economic burden on governments worldwide for various reasons, such as the rising average lifespan. A substantial portion of the expenditures is allocated to fundamental resources in the healthcare sector, including medical personnel, facilities, equipment, and medications (Atalan and Şahin, 2024). Particularly in dental health services, the use of alternative resource allocation techniques can provide crucial data for decision-makers and policymakers by offering deeper insights into the economic impact of health expenditures.

Inefficient management of resources leads to increased costs in healthcare systems and delays in service delivery. Such inefficiencies financially and operationally challenge healthcare systems (Hung et al., 2019). Therefore, the use of increasingly popular machine learning ML methods in healthcare management is critical to achieving more efficient results. ML outperforms classical statistical methods by analyzing large data sets, detecting patterns, and optimizing resource allocation. Integrating ML techniques into healthcare management minimizes inefficient use of resources, saves costs, and improves service quality with smoother operations. This study proposes an ML algorithm that aims to increase resource management efficiency in healthcare. With ML based solutions, healthcare managers can improve decision-making processes, predict future resource needs more accurately, and support the development of sustainable healthcare systems.

4. Discussion and Conclusion

In this study, the waiting time was estimated by variables obtained using ML algorithms to minimize patient waiting times in dental clinics. Six different machine learning models were used in this context. Decision Trees DT, Linear Regression LR, Support Vector Machines SVM, Gaussian Process Regression GPR, Kernel Regression KR and Artificial Neural Networks NN were applied to the daily data obtained from dental clinics and the performance of each ML model was evaluated by means of certain criteria. Before the analysis, feature selection was used to evaluate the importance levels of the variables included in the models. According to the results obtained, it was found that the GPR model performed better than the other five ML models, while the SVM model showed the lowest performance. The obtained results show that the ML algorithm will be useful in estimating patient waiting times and improving the process. In addition, the integration of data-driven decision support systems into hospital management processes will contribute to more efficient resource use, improved service quality and optimized staff workload.

This study has two primary limitations. First, incorporating detailed information about staff shift schedules could potentially enhance the accuracy and comprehensiveness of the predictions. Second, applying the methodology to larger dental hospitals, where a broader range of parameters and average treatment durations for various procedures can be included, may yield more robust and generalizable results.

Future research can expand on the findings of this study in several ways. First, larger-scale data analyses can be conducted in bigger dental hospitals, incorporating diverse treatment types and patient densities to improve prediction accuracy. Second, multivariate models that include additional variables such as patient age, appointment time, and staff experience could provide more comprehensive insights. Third, real-time prediction systems powered by artificial intelligence could be developed to enhance operational efficiency and patient satisfaction. Fourth, international comparisons could be made by analyzing datasets from clinics in different countries to test the generalizability of the proposed methods. Lastly, optimization studies can focus on designing algorithms to minimize waiting times by optimizing patient and staff workflows.

Author Contributions

The percentages of the authors' contributions are presented below. The author reviewed and approved the final version of the manuscript.

	A.K
C	100
D	100
S	100
DCP	100
DAI	100
L	100
W	100
CR	100
SR	100
PM	100
FA	100

C=Concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management, FA= funding acquisition

Conflict of Interest

The author declared that there is no conflict of interest.

Ethical Consideration:

Ethics committee approval was not required for this study because of there was no study on animals or humans.

References

Anderson RT, Camacho FT, Balkrishnan R. 2007. Willing to wait? The influence of patient wait time on satisfaction with primary care. BMC Health Serv Res, 7(1): 1-5. https://doi.org/10.1186/1472-6963-7-31/TABLES/2

Arenas-Garcia J, Petersen KB, Camps-Valls G, Hansen LK. 2013. Kernel multivariate analysis framework for supervised subspace learning: A tutorial on linear and kernel multivariate methods. IEEE Signal Proces Mag, 30(4): 16-29.

https://doi.org/10.1109/MSP.2013.2250591

Atalan A, Keskin A. 2023. Estimation of the utilization rates of the resources of a dental clinic by simulation. Sigma J Eng Nat Sci, 41(2): 423-432. https://doi.org/10.14744/sigma.2023.00045

Atalan A, Şahin H. 2024. Forecasting of the dental workforce with machine learning models. Müh Bil Araş Derg, 6(1): 125-132. https://doi.org/10.46387/bjesr.1455345

Bahammam FA. 2023. Satisfaction of clinical waiting time in ear, nose & throat departments of the Ministry of Health in Jeddah, Saudi Arabia. Health Serv Insights, 2023: 16. https://doi.org/10.1177/11786329231183315

Bertsimas D, King A. 2016. OR forum—An algorithmic approach to linear regression. Oper Res, 64(1): 2-16. https://doi.org/10.1287/opre.2015.1436

Boudreaux ED, O'Hea EL. 2004. Patient satisfaction in the Emergency Department: a review of the literature and implications for practice. J Emer Medic, 26(1): 13-26. https://doi.org/10.1016/J.JEMERMED.2003.04.003

Cayirli T, Veral E. 2003. Outpatient scheduling in health care: A review of literature. Prod Oper Manag, 12(4): 519-549. https://doi.org/10.1111/J.1937-5956.2003.TB00218.X

Channouf N, L'Ecuyer P, Ingolfsson A, Avramidis AN. 2007. The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta. Health Care Manag Sci, 10(1): 25-45. https://doi.org/10.1007/S10729-006-9006-3/METRICS

Gualtieri JA, Chettri S. 2000. Support Vector Machines for classification of hyperspectral data. International Geoscience and Remote Sensing Symposium (IGARSS), July 24928, Honolulu, US, 2: 813-815. https://doi.org/10.1109/IGARSS.2000.861712

Hung M, Xu J, Lauren E, Voss MW, Rosales MN, Su W, Licari FW. 2019. Development of a recommender system for dental care using machine learning. SN Appl Sci, 1(7): 1-12. https://doi.org/10.1007/S42452-019-0795-7/FIGURES/1

Keskin A, Ersin İ, Atalan A. 2024. Price estimation of selected grains products based on machine learning for agricultural economic development in Türkiye. J Anim Plant Sci, 34(5): 1290-1301. https://doi.org/10.36899/japs.2024.5.0811

Komşuoğlu AF. 2022. Özel diş kliniklerinde hasta memnuniyeti ve diş sağlık hizmet kalitesi. İstat Uyg Bil Derg, 6: 1-11. https://doi.org/10.52693/JSAS.1200905

Kononenko, I. 2001. Machine learning for medical diagnosis: history, state of the art and perspective. Artif Intel Medic, 23(1): 89-109. https://doi.org/10.1016/S0933-3657(01)00077-X

Liao Z, Li D, Wang X, Li L, Zou Q. 2016. Cancer diagnosis through isomir expression with machine learning method. Curr Bioinformatics, 11(3): 57-63. https://doi.org/10.2174/1574893611666160609081155

Marrel A, Iooss B. 2024. Probabilistic surrogate modeling by Gaussian process: A review on recent insights in estimation and validation. Reliab Eng Syst Safety, 247: 110094. https://doi.org/10.1016/j.ress.2024.110094

Miao J, Niu L. 2016. A survey on feature selection. Proc Comp Sci, 91: 919-926. https://doi.org/10.1016/J.PROCS.2016.07.111

Mohsin M, Forero R, Ieraci S, Bauman AE, Young L, Santiano N. 2007. A population follow-up study of patients who left an emergency department without being seen by a medical officer. Emer Medic J, 24(3): 175-179. https://doi.org/10.1136/EMJ.2006.038679

Montecinos J, Ouhimmou M, Chauhan S. 2018. Waiting-time estimation in walk-in clinics. Int Transact Oper Res, 25(1): 51-74. https://doi.org/10.1111/itor.12353

Naskath J, Sivakamasundari G, Begum AAS. 2023. A study on different deep learning algorithms used in deep neural nets: MLP SOM and DBN. Wireless Personal Commun, 128(4): 2913-2936. https://doi.org/10.1007/S11277-022-10079-4/FIGURES/7

Pitrou I, Lecourt AC, Bailly L, Brousse B, Dauchet L, Ladner J. 2009. Waiting time and assessment of patient satisfaction in a large reference emergency department: A prospective cohort study, France. Eur J Emer Medic, 16(4): 177-182. https://doi.org/10.1097/MEJ.0B013E32831016A6

Qu X, Shi J. 2011. Modeling the effect of patient choice on the performance of open access scheduling. Int J Prod Econ, 129(2): 314-327. https://doi.org/10.1016/j.ijpe.2010.11.006

Reid RO, Ashwood JS, Friedberg MW, Weber ES, Setodji CM, Mehrotra A. 2013. Retail clinic visits and receipt of primary care. J General Internal Medic, 28(4): 504-512. https://doi.org/10.1007/S11606-012-2243-X/TABLES/5

Song Y-Y, Lu Y. 2015. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry, 27(2): 130-135. https://doi.org/10.11919/j.issn.1002-0829.215044

Soremekun OA, Takayesu JK, Bohan SJ. 2011. Framework for analyzing wait times and other factors that impact patient satisfaction in the emergency department. J Emer Medic, 41(6): 686-692. https://doi.org/10.1016/J.JEMERMED.2011.01.018

Stiglic G, Kocbek P, Fijacko N, Zitnik M, Verbert K, Cilar L. 2020. Interpretability of machine learning-based prediction models in healthcare. WIREs Data Mining Knowl Discov, 10(5): e1379. https://doi.org/10.1002/widm.1379

The MathWorks, Inc. 2024. MATLAB and Statistics Toolbox Release [2024a], Regression Learner. Massachusetts: The MathWorks, Inc.